

Quantum Mechanics

Complete Course Notes

Lectures by Prof. Ben Heidenreich

Notes compiled by Siyi Zhang

Jump to course part

Motivation and Quantum States	15
Finite-Dimensional Hilbert Spaces	33
The Postulates of Quantum Mechanics	46
Time Evolution and Symmetries	63
Quantum Information and Entanglement	100
Continuous Observables	134
Wave Mechanics and Bound States	163
Scattering and Resonances	188
Approximation Methods	202

Content

Introduction	13
References and source materials	14
I Motivation and Quantum States	15
1 Why quantum mechanics changes the description of a state	15
2 Photons, matter waves, and interference	15
2.1 Localized detection and a wave-like distribution	15
2.2 Why adding classical path probabilities fails	16
2.3 Probability amplitudes and the Born rule	18
2.4 Matter waves	19
2.5 Deriving the phase difference	19
3 Complementarity, uncertainty, and information	20
3.1 A which-path measurement on an electron	20
3.2 Can recoil of the source reveal the path?	21
3.3 What information is contained in a quantum state?	22
4 The Stern–Gerlach experiment	22
4.1 Historical motivation and magnetic moments	22
4.2 An inhomogeneous magnetic field	24
4.3 The observed two outcomes and electron spin	25
5 Sequential measurements and coherent recombination	26
5.1 Preparing and measuring a spin state	26
5.2 Why precession cannot simply be tuned away in a resolved measurement	27
5.3 A modified apparatus that splits and recombines	28
5.4 The amplitudes behind the four experiments	30
6 Complex two-state vectors	30
6.1 Changing between the z and x descriptions	30
6.2 Why complex amplitudes are needed	31
6.3 Why an unpolarized beam is not a pure state	32
II Finite-Dimensional Hilbert Spaces	33
7 Kets, bras, and linear maps	33
7.1 Complex vector spaces	33
7.2 The dual space	33
7.3 Operators and their algebra	33
7.4 Outer products and rank	34

8	Inner products and adjoints	35
8.1	The Hilbert-space inner product	35
8.2	Adjoints of operators	36
9	Orthonormal bases and projectors	37
9.1	Resolution of the identity	37
9.2	Projectors onto subspaces	37
10	Eigenvectors and the spectral theorem	38
10.1	Self-adjoint eigenvalue properties	38
10.2	The corresponding facts for normal operators	38
10.3	Proof of the finite-dimensional spectral theorem	39
10.4	Positive operators	40
11	Degeneracy, simultaneous eigenvectors, and spectral projectors	41
11.1	Degenerate eigenspaces	41
11.2	Complete orthogonal projector families	41
11.3	Commuting normal operators	41
12	Unitary operators and changes of representation	43
12.1	Preservation of inner products	43
12.2	Unitary conjugation of operators	43
13	Components, matrix multiplication, and invariants	44
13.1	Columns, rows, and operator matrices	44
13.2	Active transformations and passive basis changes	45
13.3	Trace and determinant	45
III	The Postulates of Quantum Mechanics	46
14	Spin measurements in operator language	46
14.1	State kets and the Stern–Gerlach projectors	46
14.2	An observable and its expectation value	47
14.3	The three spin observables	47
14.4	Why distinct measurement states must be orthogonal	48
14.5	Pauli matrices and their algebra	48
14.6	Relative phases in a Stern–Gerlach recombiner	50
15	Postulate 1: pure states	51
15.1	Why use a complex Hilbert space?	51
15.2	Orthogonal alternatives in a decaying system	52
16	From interference to linear measurement maps	52
16.1	Outcomes need not specify complete final states	52
16.2	Real rescaling and additivity	54
16.3	Why these conditions give linearity or antilinearity	54

16.4	Why a coherent physical branch cannot be antilinear	55
17	Repeatable measurements and the Born rule	56
17.1	Probability conservation and repeatability	56
17.2	Deriving the projector form step by step	57
17.3	Probabilities and the state after an outcome	58
18	Postulate 2: observables and ideal measurements	59
18.1	Degeneracy and what an apparatus actually resolves	60
18.2	Measurements without repeatability	61
19	Postulate 3: closed-system transformations	61
19.1	Finite and infinite dimensions	62
19.2	Why an arbitrary component norm will not do	62
IV	Time Evolution and Symmetries	63
20	Unitary time evolution and the Hamiltonian	63
20.1	Closed processes and the propagator	63
20.2	Defining and identifying the generator	63
20.3	Time-independent Hamiltonians and stationary states	64
20.4	When the Hamiltonian depends on time	65
21	Spin precession and driven two-level dynamics	66
21.1	Precession calculated from the evolving state	66
21.2	Exact Rabi oscillations for a rotating field	67
22	The Schrödinger and Heisenberg pictures	67
22.1	A simultaneous change of states and observables	67
22.2	Constructing the Heisenberg picture	68
22.3	Derivation of the Heisenberg equation	68
22.4	Solving spin precession as an operator equation	68
23	Canonical quantization and Ehrenfest's theorem	69
23.1	Classical evolution as a Poisson bracket	69
23.2	Angular momentum brackets worked out	70
23.3	Ordering and the limits of the correspondence	70
23.4	Ehrenfest's theorem and mean trajectories	71
24	Ammonia inversion: dynamics determined by symmetry	71
24.1	The two-state approximation	71
24.2	Choosing phases for the reflection operator	72
24.3	The most general reflection-invariant Hamiltonian	72
24.4	Time evolution from a localized configuration	73
25	Quantum symmetries and selection rules	74

25.1	Symmetries preserve the equation of motion	74
25.2	Three equivalent ways to describe a transformation	74
25.3	A general Abelian selection rule	75
26	Groups, representations, and triangular hopping	75
26.1	Why symmetry operations form a group	75
26.2	Three sites and the dihedral group	77
26.3	Eliminating phases in the site representation	77
26.4	The invariant Hamiltonian and its eigenvectors	78
26.5	Why the doublet is forced to be degenerate	79
26.6	A cyclic N -site extension	79
27	Irreducible representations and non-Abelian selection rules	80
27.1	Invariant subspaces and complete reducibility	80
27.2	Schur's lemma with proof	80
27.3	Equivalent irreps and multiplicity spaces	80
28	Continuous symmetries and conserved generators	81
28.1	One-parameter groups	81
28.2	Phase conventions and additive constants	82
28.3	Spin generates spatial rotations	82
28.4	Energy and time translations	84
28.5	Infinitesimal transformations in the three viewpoints	84
29	Lie groups and Lie algebras	85
29.1	Composing continuous transformations	85
29.2	Closure and conservation of commutators	86
29.3	The invariant trace inner product	86
29.4	Local algebra and global group	87
30	Classification of spin representations	88
30.1	The Casimir and ladder algebra	88
30.2	Norms, termination, and quantization	89
30.3	Irreducibility and completeness of the classification	90
30.4	Ordinary vectors are spin one	91
30.5	How far must a state rotate to become orthogonal?	91
31	Half-integer spin, topology, and projective representations	93
31.1	The sign under a full rotation	93
31.2	The group $SU(2)$ and its double cover	93
31.3	A geometric picture of the two groups	94
31.4	Projective multiplication and relative phases	96
32	Time-dependent symmetries and time reversal	97
32.1	Explicitly time-dependent unitary symmetries	97
32.2	Quantum Noether theorem with explicit time dependence	97

32.3	Antiunitary maps	98
32.4	Why time reversal is antiunitary	99
V	Quantum Information and Entanglement	100
33	Uncertainty, superposition, and quantum information	100
33.1	Fluctuations and variances	100
33.2	Cauchy–Schwarz and the uncertainty relation	100
33.3	An alternative proof by positivity	102
33.4	Superposition does not mean classical ignorance	102
34	Statistical ensembles and density operators	103
34.1	Three Stern–Gerlach preparations	103
34.2	The density operator of an ensemble	105
34.3	Coherence and nonunique ensemble decompositions	106
34.4	Pure states, mixed states, and Bloch vectors	106
34.5	The measurement rules in density-operator form	107
34.6	A measurement whose result is forgotten	108
34.7	Unitary evolution and the von Neumann equation	108
34.8	Comparison with classical Liouville evolution	109
35	Composite systems and tensor products	110
35.1	The Hilbert space of two subsystems	110
35.2	Operators on a composite system	110
35.3	Two spins and their total spin	111
35.4	The singlet state and anticorrelations	112
35.5	Product states and entangled states	113
36	Reduced states, the partial trace, and no signalling	114
36.1	Why a subsystem needs a density operator	114
36.2	Definition and basis formula	114
36.3	The local-expectation identity	115
36.4	Worked partial traces	116
36.5	Local operations and conditional reduced states	116
36.6	Why remote measurements cannot send a signal	117
36.7	The EPR arrangement	118
37	Bell inequalities and quantum correlations	119
37.1	The experiment and its assumptions	119
37.2	Joint probabilities for the singlet	120
37.3	A three-setting Bell inequality	121
37.4	Violation by quantum mechanics	122
37.5	What the Bell result does and does not establish	123
38	Measurement as an interaction with a detector	123

38.1	A system qubit and a readout qubit	123
38.2	The controlled-NOT realization	124
38.3	Copying a basis label is not cloning a state	124
38.4	Why entanglement alone is not a selected outcome	125
38.5	A detector with many records	125
38.6	The scope of the unitary detector model	126
39	Schmidt decomposition and purification	127
39.1	The Schmidt decomposition	127
39.2	Unequal dimensions and maximal entanglement	128
39.3	Purifying an arbitrary density operator	128
39.4	Uniqueness up to a unitary on the purifier	129
40	Entropy and entanglement	130
40.1	Classical entropy and operator functions	130
40.2	Basic entropy properties	130
40.3	Entanglement entropy of a pure pair	131
40.4	Entropy of product states and correlated states	132
40.5	Entropy, environments, and thermodynamic descriptions	132
VI	Continuous Observables	134
41	Canonical commutation relations and infinite-dimensional spaces	134
41.1	The trace obstruction	134
41.2	Algebra first, representation second	135
42	The harmonic oscillator from ladder operators	135
42.1	Creation, annihilation, and number operators	135
42.2	How the ladder moves an eigenvalue	136
42.3	Why the eigenvalues are nonnegative integers	136
42.4	Normalization, phases, and the Fock basis	137
42.5	Worked exercise: equal degeneracy of every rung	138
42.6	Position and momentum matrix elements	138
43	Unbounded operators, domains, and generalized states	139
43.1	Normalizable does not mean every operator can act	139
43.2	Worked exercise: cancellation can enlarge the position domain	140
43.3	Generalized eigenvectors and a test-function space	141
44	Diagonalizing position: Hermite functions	142
44.1	The position eigenvalue recurrence	142
44.2	Rodrigues' formula and the generating function	142
44.3	Orthogonality, with its normalization proved	143
44.4	Completeness of the Hermite functions	144
44.5	Worked exercise: generalized orthogonality and independence	144

45	Position measurements and projection-valued measures	145
45.1	The wavefunction as a representation of the state	145
45.2	Finite-resolution position measurements	146
45.3	Worked exercise: the PVM identities	146
45.4	Why the density of a PVM is not a point projector	147
45.5	Example: the oscillator ground-state position distribution	147
46	Momentum, translations, and phase conventions	147
46.1	Commutators with functions of momentum	147
46.2	Fixing the relative phases of position kets	148
46.3	What changes under a position-dependent basis phase	149
46.4	The momentum domain on the line	149
47	Momentum eigenstates and the Fourier transform	150
47.1	Plane waves and delta normalization	150
47.2	The two wavefunction representations	151
47.3	Exact position and exact momentum under Fourier transformation	151
48	Gaussian wave packets and the uncertainty bound	152
48.1	A normalized packet with specified means and width	152
48.2	The momentum Gaussian, derived by completing the square	152
48.3	The uncertainty inequality and its equality conditions	153
48.4	Worked exercise: every minimum-product state is Gaussian	153
49	The oscillator ground state and zero-point motion	154
49.1	Fixing the remaining phase of the oscillator functions	154
49.2	Worked exercise: excited states are not Gaussian packets	155
49.3	A lower bound on energy without ladder operators	155
50	The Schrödinger representation in several dimensions	156
50.1	The Hilbert space is not fixed by the choice of potential	156
50.2	The scope of the Stone–von Neumann theorem	156
50.3	Joint position and momentum bases in dimension d	156
51	Schrödinger’s wave equation	157
51.1	From the abstract state equation to a PDE	157
51.2	Stationary states and the energy eigenvalue equation	158
52	Probability conservation and probability current	158
52.1	Deriving the continuity equation	158
52.2	A stationary density can carry a current	159
52.3	Current and the expectation of momentum	159
53	Wavefunction phase and the classical Hamilton–Jacobi limit	159
53.1	Amplitude and phase variables	159
53.2	Separating the Schrödinger equation into real and imaginary parts	160
53.3	The controlled classical limit	161

53.4	Time-independent Hamiltonians and the characteristic function . . .	162
VII	Wave Mechanics and Bound States	163
54	Stationary states and confinement	163
54.1	Energy eigenfunctions and their time dependence	163
54.2	The free particle and continuum normalization	163
54.3	A periodic box	164
54.4	A hard-walled box and the momentum domain	164
54.5	Density of states in arbitrary dimension	165
55	Multidimensional harmonic oscillators	165
55.1	Normal coordinates and the spectrum	165
55.2	Rotation multiplets and the hidden symmetry	166
55.3	High-energy state counting and accessible volume	167
56	One-dimensional bound states	168
56.1	Discrete levels, continua, and thresholds	168
56.2	An attractive delta-function well	168
56.3	Parity reduction for the finite square well	169
56.4	Graphical quantization, level count, and energy bounds	170
56.5	Normalization, tail probability, and signs	171
56.6	A shallow well and the delta-function limit	172
56.7	A state approaching the continuum	172
56.8	Nondegeneracy, nodes, and a threshold example	173
57	Central potentials and the radial Schrödinger equation	173
57.1	Rotational symmetry and orbital angular momentum	173
57.2	Removing the radial measure	174
57.3	The infinite spherical well	175
58	The Coulomb spectrum by factorization	176
58.1	Two-body reduction and natural scales	176
58.2	Factorization and the lowest state at fixed angular momentum	177
58.3	Shape invariance and quantization of every bound energy	177
58.4	Descending in angular momentum and normalization factors	178
58.5	Explicit radial polynomials and their recurrence	179
58.6	Laguerre form, normalization, and low states	180
59	Propagators and free-particle wavepackets	181
59.1	The kernel of time evolution	181
59.2	The free kernel by a Gaussian integral	181
59.3	Gaussian spreading, current, and momentum uncertainty	182
59.4	A moving Gaussian and a general quadratic phase	183
59.5	Propagator on a circle	183

60	Oscillator propagation, coherent states, and squeezing	184
60.1	Mehler's formula and the oscillator kernel	184
60.2	A coherent Gaussian moving without distortion	185
60.3	A squeezed packet and its breathing width	186
VIII	Scattering and Resonances	188
61	Stationary scattering amplitudes and probability flux	188
61.1	Incoming and outgoing solutions	188
61.2	The conserved current and flux normalization	188
61.3	Recovering arbitrary incidence from one scattering solution	189
62	Scattering of localized wavepackets	190
62.1	Building a packet from stationary scattering states	190
62.2	Exact free dispersion and the separated outgoing packets	190
62.3	Phase derivatives and time delays	191
63	A delta scatterer and analytic continuation	192
63.1	Exact reflection and transmission	192
63.2	The high-energy classical limit	193
63.3	A pole is not automatically a bound state	193
64	Poles, bound states, and a double-delta potential	194
64.1	Why upper-half-plane poles give bound states	194
64.2	Exact amplitudes for two identical delta functions	194
64.3	Imaginary poles and the number of bound states	195
64.4	The infinite families of complex poles	196
65	Resonances, time delay, and unstable trapped states	196
65.1	A narrow pole produces a Lorentzian peak	196
65.2	Outgoing resonance functions and self-adjointness	198
65.3	Why the formal spatial growth has a physical interpretation	198
66	A normalizable scattering packet near a resonance	199
66.1	Separating the pole from the smooth background	199
66.2	Deriving the Gaussian-pole integral	199
66.3	The exact pole packet	200
IX	Approximation Methods	202
67	The variational method	202
67.1	The ground-state energy as a minimum	202
67.2	A Gaussian trial state for the quartic oscillator	202
67.3	Why an approximate state can give a better energy	203
67.4	Excited-state bounds and parity	204

68	Nondegenerate stationary perturbation theory	204
68.1	Intermediate normalization and two exact equations	205
68.2	The pseudoinverse and the reduced resolvent	205
68.3	Expansion and recursion	206
68.4	A longitudinal perturbation of a two-state system	207
68.5	A transverse perturbation and exact comparison	207
68.6	The anharmonic oscillator	208
68.7	Analyticity is a separate question	209
69	Asymptotic expansions	210
69.1	The meaning of an asymptotic series	210
69.2	Stirling's expansion and optimal truncation	210
70	Degenerate perturbation theory	212
70.1	Choosing the zeroth-order state	212
70.2	Second-order splitting of residual degeneracy	213
70.3	The first-order state within the original eigenspace	213
70.4	A practical sequence of calculations	214
71	Two coupled harmonic oscillators	214
71.1	Degeneracy and the oscillator symmetry	214
71.2	First-order splitting of the first excited level	215
71.3	Second-order energy shifts	215
71.4	Exact normal modes and perturbative comparison	217
72	The semiclassical WKB expansion	218
72.1	A local momentum and the phase equation	218
72.2	Leading phase, amplitude, and the next correction	218
72.3	Validity and turning points	219
72.4	Counterpropagating waves and the phase around a well	220
73	Turning points, Airy functions, and bound-state quantization	221
73.1	Reduction to the Airy equation	221
73.2	Deriving the connection at a right turning point	222
73.3	A left turning point and signed action conventions	223
73.4	Action differences and classical travel times	224
73.5	Bound states between two simple turning points	224
73.6	The harmonic oscillator	225
73.7	The pure quartic oscillator	226
74	Tunneling through a smooth potential barrier	227
74.1	Scattering boundary conditions and phase conventions	227
74.2	Matching from the transmitted wave back to the incident wave	228
74.3	The transmission phase and wave-packet delay	229
75	Euclidean trajectories and the tunneling action	230

75.1 The inverted-potential problem	230
75.2 Instanton and bounce actions	231
76 Metastable states and resonance lifetimes	231
76.1 A symmetric well between two barriers	231
76.2 Matching the outgoing solution through the right barrier	232
76.3 Even and odd resonance conditions	233
76.4 Expanding around a real level	234
76.5 Lifetime and the classical escape picture	234
Source-to-topic index	236

Introduction

These notes were compiled by **Siyi Zhang** from the quantum mechanics lectures of **Prof. Ben Heidenreich**. The presentation follows the numbered lecture sequence, incorporating the accompanying handwritten lectures and supplementary handouts.

Quantum mechanics is developed first as a framework for states, observables, probabilities, and physical transformations. Wave mechanics then emerges as the position-space realization of this framework, providing tools for bound states, scattering, and approximation methods.

Course content

1. **Motivation.** Experimental foundations, Stern–Gerlach measurements, interference, and the need for probability amplitudes.
2. **Finite-dimensional Hilbert spaces.** Bras and kets, operators, inner products, spectral decompositions, and changes of basis.
3. **The postulates.** States and observables, measurement probabilities, state updates, and closed quantum processes.
4. **Time evolution and symmetries.** The Hamiltonian, evolution operators, pictures of dynamics, generators, and conservation laws.
5. **Quantum information and entanglement.** Uncertainty, mixed states, composite systems, reduced states, and quantum correlations.
6. **Continuous observables.** Infinite-dimensional state spaces, the harmonic oscillator, position and momentum, and wavefunctions.
7. **Wave mechanics and bound states.** Stationary states, probability flow, matching conditions, and exactly solvable potentials.
8. **Scattering and resonances.** Reflection and transmission, scattering amplitudes, bound-state poles, and unstable states.
9. **Approximation methods.** Variational estimates, semiclassical methods, and perturbation theory, including degeneracy.

Conventions

The inner product is antilinear in the bra and linear in the ket. States differing by a nonzero complex scalar represent the same pure state; normalization is stated where probabilities are evaluated. The Fourier phases and normalization factors are specified when position and momentum representations are introduced.

References and source materials

Primary lecture sequence

Ben Heidenreich, *Quantum Mechanics*, supplied lecture notes:

1. *Motivation*.
2. *Finite-dimensional Hilbert Spaces*.
3. *The Postulates of Quantum Mechanics*.
4. *Time Evolution and Symmetries*.
5. *Quantum Information and Entanglement*, with the accompanying handwritten lectures.
6. *Continuous Observables*, with the accompanying handwritten lectures.
7. *Wave Mechanics*, together with the bound-state handout.
8. *Scattering, Bound States, and Resonances*, together with *Note on Resonances*.
9. *Approximation Methods*, together with *Degenerate Perturbation Theory*.

Reading cited in the lectures

The time-evolution and symmetries lecture cites the following reading:

- John S. Townsend, *A Modern Approach to Quantum Mechanics*, University Science Books. The lecture cites Section 4.5 for ammonia and the preparation of a localized molecular configuration.
- J. J. Sakurai and Jim Napolitano, *Modern Quantum Mechanics*, Cambridge University Press. The lecture cites Section 4.4 for time reversal and antilinear operators.

Using the notes

Definitions establish the mathematical language; propositions and proofs record the logical steps behind the main results. Worked examples develop the calculations, while exercises retain further applications from the lectures. The original diagrams are placed alongside the relevant arguments. An index at the end maps the fourteen supplied PDFs to the corresponding parts of these notes.

Part I

Motivation and Quantum States

Chapter 1

1 Why quantum mechanics changes the description of a state

Quantum mechanics is often introduced through Schrödinger's equation,

$$i\hbar \frac{\partial \psi(\mathbf{r}, t)}{\partial t} = -\frac{\hbar^2}{2m} \nabla^2 \psi(\mathbf{r}, t) + V(\mathbf{r}, t) \psi(\mathbf{r}, t). \quad (1.1)$$

Schrödinger proposed this equation in 1926. It gives the dynamics of a particular kind of quantum system: a nonrelativistic particle with a position-space wavefunction. It is an important application of quantum mechanics, but it is not the entire framework.

The more fundamental change is in *kinematics*: what constitutes a state, which questions can have definite answers, and how the state determines the results of measurements. In classical particle mechanics, a point in phase space, specified by position and momentum, supplies the initial data. In quantum mechanics, the state will instead be represented by a vector in a complex Hilbert space. The same general rules apply to a two-state spin system, to a particle moving in space, and to quantum fields, although their Hilbert spaces are very different.

We will therefore develop the state and measurement rules first. A prescription for constructing a quantum theory from classical variables is a further step, usually called *quantization*. Such prescriptions help identify the appropriate Hilbert space and observables; they do not replace the general quantum postulates.

2 Photons, matter waves, and interference

2.1 Localized detection and a wave-like distribution

Planck's work on thermal radiation in 1900 and Einstein's explanation of the photoelectric effect in 1905 introduced the relation

$$E = hf = \hbar\omega, \quad h \simeq 6.6 \times 10^{-34} \text{ J s}, \quad \hbar = \frac{h}{2\pi}. \quad (2.1)$$

An individual photon transfers energy in a localized detection event. Repeating the experiment produces many such events at apparently random positions. Nevertheless, the distribution of many detections agrees with the intensity predicted by wave optics. A localized event and a wave-like statistical distribution are both parts of the phenomenon that the theory must describe.

Consider monochromatic light of wavelength λ incident on two parallel slits. Let each slit have width w , let their centres be separated by d , and let the screen be a distance L beyond the slits. We denote transverse position on the screen by x .

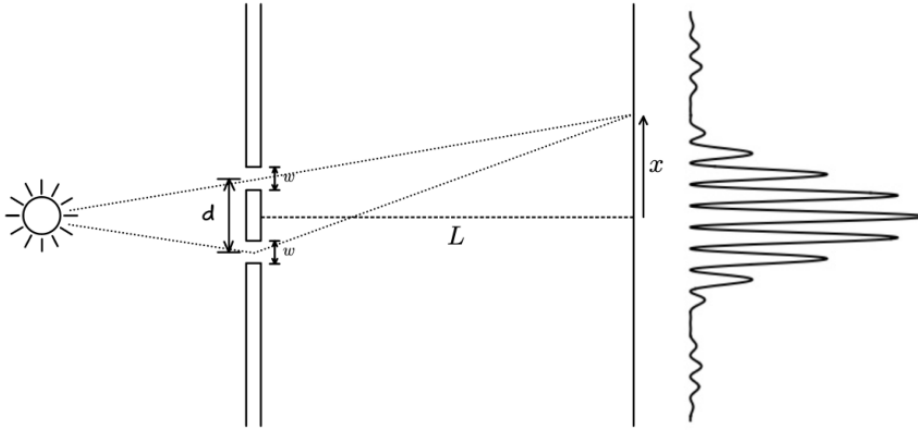


Figure 1: The original double-slit geometry: slit width w , separation d , and a distant screen.

In the paraxial, Fraunhofer regime, the intensity takes the form

$$I_{\text{two}}(x) \propto \text{sinc}^2\left(\frac{\pi wx}{\lambda L}\right) \cos^2\left(\frac{\pi dx}{\lambda L}\right), \quad \text{sinc } u = \frac{\sin u}{u}. \quad (2.2)$$

The continuous definition at the origin is $\text{sinc } 0 = 1$. The broad sinc^2 factor is the single-slit diffraction envelope. The more rapidly varying cosine factor produces the interference fringes. Fraunhofer propagation requires the quadratic variation of the propagation phase across the aperture to be negligible; a sufficient far-field condition is $L \gg a^2/\lambda$, where a is the transverse size of the entire aperture. The paraxial approximation also requires the relevant observation angles to be small.

Covering one slit removes the narrow fringes, leaving the single-slit envelope. Reducing the intensity does not continuously turn the two-slit pattern into the one-slit pattern. Even when photons are detected one at a time, the two-slit distribution builds up after many events. Weak-light experiments already explored this issue in the early twentieth century; genuine single-photon preparations make its meaning especially direct.

2.2 Why adding classical path probabilities fails

If a photon is detected at a slit, the detector records one localized outcome. It is tempting to infer that every photon has a definite slit label even when that label is not measured, and that opening both slits should therefore give

$$P(x) \stackrel{?}{=} P_1(x) + P_2(x). \quad (2.3)$$

This prediction has no interference term and disagrees with the observed fringes. Indeed, there are positions where opening the second slit *reduces* the count rate to zero. Adding nonnegative probabilities for two independent alternatives cannot do this.

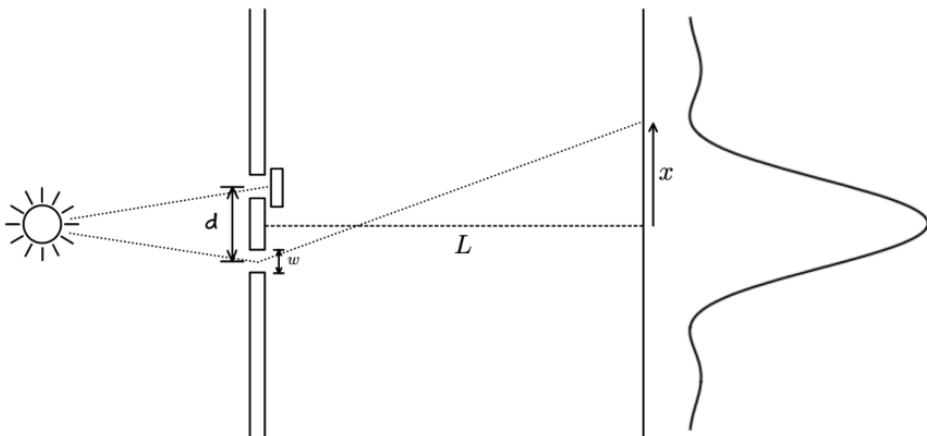


Figure 2: Blocking one slit leaves the broad diffraction envelope.

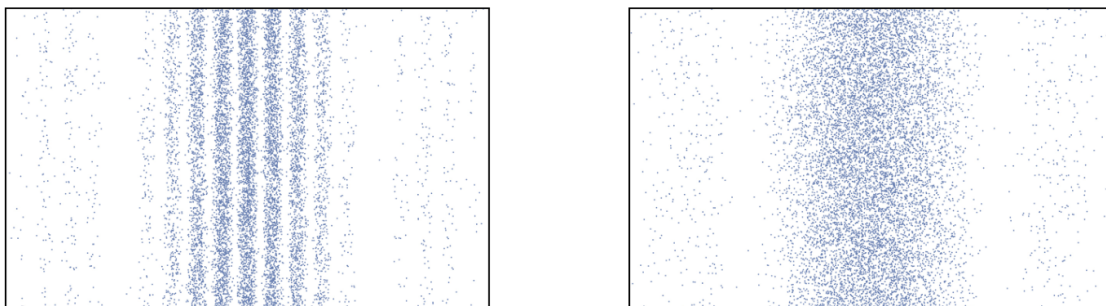


Figure 3: The original simulated photon counts: both slits open on the left, and one slit open on the right. Each localized detection contributes to a distribution.

For equal illumination of two nonoverlapping slits, a convenient idealized normalization is

$$P_1(x) = P_2(x) = \frac{w}{2\lambda L} \operatorname{sinc}^2\left(\frac{\pi wx}{\lambda L}\right), \quad (2.4)$$

$$P(x) = \frac{2w}{\lambda L} \operatorname{sinc}^2\left(\frac{\pi wx}{\lambda L}\right) \cos^2\left(\frac{\pi dx}{\lambda L}\right). \quad (2.5)$$

Here P_1 and P_2 use the original two-path incident ensemble as the reference: each integrates to $1/2$ in the ideal far-field model. If one conditions on transmission through a single open slit, the normalized conditional distribution is $2P_1$ or $2P_2$. This distinction between a probability relative to the incident beam and a conditional probability will also matter in the Stern–Gerlach experiments.

One can try to find which slit a photon traverses by scattering it from an electron. If the scattering record distinguishes the alternatives, the interference fringes disappear. A path detector need not absorb the photon to have this effect. The issue is whether a distinguishable record of the alternative has been produced.

The correct conclusion is operational: a path measurement and an interference measurement are different experimental arrangements. A definite result of the first does not justify assigning the same unmeasured result in the second.

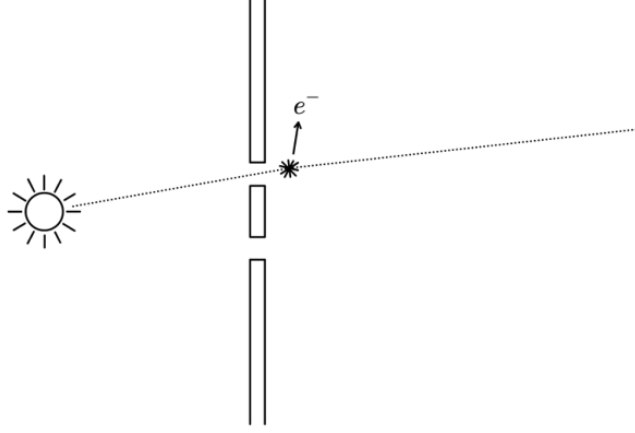


Figure 4: The source's which-path arrangement for photons, using scattering near the slits.

2.3 Probability amplitudes and the Born rule

The alternative quantities that add are complex **probability amplitudes**. For the double slit, choose

$$\psi_1(x) = \sqrt{\frac{w}{2\lambda L}} e^{i\pi dx/(\lambda L)} \operatorname{sinc}\left(\frac{\pi wx}{\lambda L}\right), \quad (2.6)$$

$$\psi_2(x) = \sqrt{\frac{w}{2\lambda L}} e^{-i\pi dx/(\lambda L)} \operatorname{sinc}\left(\frac{\pi wx}{\lambda L}\right). \quad (2.7)$$

For coherent alternatives leading to the same final outcome,

$$\psi(x) = \psi_1(x) + \psi_2(x) = \sqrt{\frac{2w}{\lambda L}} \cos\left(\frac{\pi dx}{\lambda L}\right) \operatorname{sinc}\left(\frac{\pi wx}{\lambda L}\right). \quad (2.8)$$

The **Born rule** gives the probability density as the squared modulus:

$$P(x) = |\psi(x)|^2 = |\psi_1|^2 + |\psi_2|^2 + 2 \operatorname{Re}(\psi_1^* \psi_2). \quad (2.9)$$

The last term can be positive or negative. It accounts for both enhancement and destructive interference. The phases of the amplitudes contain information that is absent from the probabilities $|\psi_1|^2$ and $|\psi_2|^2$ separately.

Multiplying every amplitude by the same constant phase $e^{i\phi}$ leaves all probabilities unchanged. Multiplying a position-space wavefunction by a position-dependent phase $e^{i\phi(x)}$ also leaves its position density unchanged, but generally changes other predictions, such as momentum probabilities. A global phase and a relative phase are therefore very different kinds of information.

The word *coherent* specifies that the alternatives retain a definite relative phase and have no remaining record that distinguishes them in the degrees of freedom relevant to interference. If the alternatives become correlated with orthogonal detector states $|D_1\rangle$ and $|D_2\rangle$, the total amplitude is instead

$$\psi_1(x)|D_1\rangle + \psi_2(x)|D_2\rangle.$$

Its squared norm is

$$P(x) = |\psi_1|^2 + |\psi_2|^2 + 2 \operatorname{Re}[\psi_1^* \psi_2 \langle D_1 | D_2 \rangle]. \quad (2.10)$$

Thus an orthogonal path record eliminates the cross term even if nobody reads the detector. This expression anticipates the Hilbert-space treatment below.

2.4 Matter waves

The phenomenon is not restricted to light. De Broglie proposed in 1924 that a particle with momentum magnitude p has wavelength

$$\lambda = \frac{h}{p}. \quad (2.11)$$

Electron diffraction, including the Davisson–Germer experiment, supplied direct evidence for this relation. Electrons can also produce a double-slit interference distribution through individual localized detections. Calling light a wave and matter a particle does not remove the need for a common quantum description.

2.5 Deriving the phase difference

Exercise 2.1 (Path lengths and a displaced source). Derive the two-slit interference phase, first for an on-axis source and then for a source at transverse position x_0 , a distance L_0 before the slit plane.

Solution. Put the slit centres at $s_{\pm} = \pm d/2$. For a screen point x , the outgoing path length is

$$\ell_{\text{out}}(s) = \sqrt{L^2 + (x - s)^2} \simeq L + \frac{(x - s)^2}{2L}.$$

Therefore

$$\ell_{\text{out}}(d/2) - \ell_{\text{out}}(-d/2) = -\frac{dx}{L}.$$

Multiplication by $k = 2\pi/\lambda$ gives the phase difference $-2\pi dx/(\lambda L)$. Two equal amplitudes with phase difference $\Delta\phi$ have a total squared magnitude proportional to

$$|e^{i\Delta\phi/2} + e^{-i\Delta\phi/2}|^2 = 4 \cos^2 \frac{\Delta\phi}{2},$$

which yields the cosine factor in (2.2).

For the displaced source, include the incoming distance:

$$\ell(s) = \sqrt{L_0^2 + (s - x_0)^2} + \sqrt{L^2 + (x - s)^2}.$$

Expanding both square roots gives

$$\ell(d/2) - \ell(-d/2) \simeq -d \left(\frac{x_0}{L_0} + \frac{x}{L} \right).$$

The interference factor is consequently

$$\cos^2 \left[\frac{\pi d}{\lambda} \left(\frac{x}{L} + \frac{x_0}{L_0} \right) \right]. \quad (2.12)$$

For a narrow slit, integration over its local transverse coordinate $u \in [-w/2, w/2]$ supplies the diffraction envelope. For example, for on-axis illumination,

$$\int_{-w/2}^{w/2} e^{-ikxu/L} du = w \operatorname{sinc} \left(\frac{\pi wx}{\lambda L} \right).$$

This separates the within-slit diffraction factor from the between-slit interference factor. $\square$

3 Complementarity, uncertainty, and information

3.1 A which-path measurement on an electron

Suppose an electron has longitudinal momentum p_e and wavelength $\lambda_e = h/p_e$. We illuminate the slit region with photons of wavelength λ_γ to determine which slit the electron passes through.

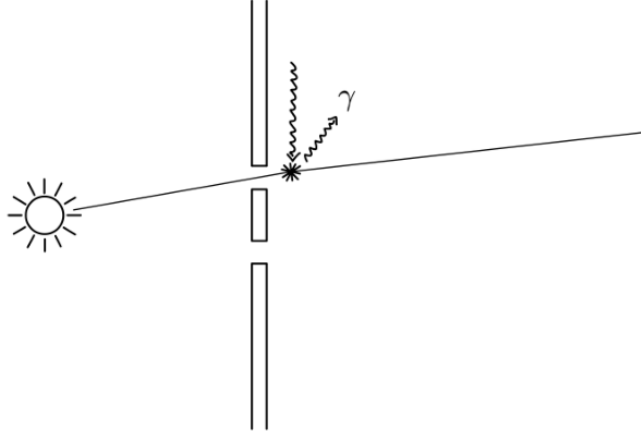


Figure 5: The original electron which-path thought experiment: photons scatter from the electron near the slit plane.

A photon carries momentum of order h/λ_γ . Scattering produces a transverse electron momentum uncertainty of the same order:

$$\Delta p_e \sim \frac{h}{\lambda_\gamma}. \quad (3.1)$$

To resolve slits separated by d , the wavelength must be no larger than the relevant separation, up to optical factors:

$$\lambda_\gamma \lesssim d \implies \Delta p_e \gtrsim \frac{h}{d}.$$

The subsequent flight time is approximately L/v_z . For a small transverse kick, the uncertainty in the arrival position is

$$\Delta x \sim \frac{\Delta p_e}{m} \frac{L}{v_z} = \frac{\Delta p_e L}{p_e} = \frac{\Delta p_e L \lambda_e}{h} \quad (3.2)$$

$$\gtrsim \frac{L \lambda_e}{d}. \quad (3.3)$$

But $L \lambda_e/d$ is precisely the spacing between adjacent interference maxima. The localization necessary for this which-path measurement brings in a disturbance large enough to spoil the fringes.

This is an order-of-magnitude argument, not a derivation of the exact standard-deviation inequality $\sigma_x \sigma_p \geq \hbar/2$. Its purpose is to show how the two requirements conflict in a concrete apparatus. More generally, loss of interference is described by the distinguishability of path records, as in (2.10); it need not always be interpreted as a classical random kick.

Bohr's principle of complementarity emphasizes that different experimental arrangements make different aspects of a quantum system accessible. Heisenberg's uncertainty principle quantifies restrictions on simultaneous statistical sharpness. Both emerged in 1927. The Bohr–Einstein thought experiments repeatedly tested whether one could obtain a path record while preserving a complete interference pattern.

3.2 Can recoil of the source reveal the path?

Exercise 3.1 (Source recoil and the interference phase). In the displaced-source geometry, try to infer the photon's slit from the transverse recoil of the emitting source. Explain why making this recoil sufficiently sharp conflicts with a stable interference pattern.

Solution. The angular difference between the two outgoing photon directions at the source is approximately d/L_0 . Since the photon momentum magnitude is h/λ , the separation between the two possible transverse recoil momenta is of order

$$\delta p_{\text{recoil}} \sim \frac{\hbar d}{\lambda L_0}.$$

To distinguish these recoils reliably, the initial source momentum spread must be much smaller:

$$\sigma_{p_0} \ll \frac{\hbar d}{\lambda L_0}.$$

The position–momentum uncertainty relation then implies

$$\sigma_{x_0} \geq \frac{\hbar}{2\sigma_{p_0}} \gg \frac{\lambda L_0}{4\pi d}.$$

The relative phase in (2.12) varies with source position at rate $2\pi d/(\lambda L_0)$. Its spread is therefore

$$\sigma_{\Delta\phi} \sim \frac{2\pi d}{\lambda L_0} \sigma_{x_0}.$$

In the limit of a sharply distinguishable recoil record, this spread becomes large. Averaging over the uncertain source position suppresses the oscillatory interference term. Equivalently, the two paths become correlated with distinguishable source recoil states, so that their overlap in (2.10) becomes small. The source itself can serve as the path detector; it need not be a separately placed instrument at a slit. $\square$

3.3 What information is contained in a quantum state?

A classical particle state is often described by $(\mathbf{r}, \mathbf{v})$, or equivalently $(\mathbf{r}, \mathbf{p})$. A quantum state cannot assign arbitrarily sharp values to all of these quantities simultaneously. It supplies probabilities for measurements, including measurements that are mutually incompatible.

For a *pure state*, the amplitudes, including their relative phases, determine the predictions for every allowed measurement. A measurement of position on many identically prepared systems determines $|\psi(x)|^2$, but does not by itself determine the phase of $\psi(x)$. Different measurements are needed to recover the additional information.

The quantum state of an unknown system cannot in general be read out completely from one specimen. A measurement gives a particular result and generally changes the state. Repeating the same measurement on that same specimen does not recreate the original preparation. **State tomography** instead uses many identically prepared systems and several measurement settings. In the general case, the object determined by the statistics is a density operator. Different statistical ensembles of pure states can give the same density operator, so tomography does not identify a unique ensemble decomposition.

Quantum mechanics is intrinsically probabilistic in its usual operational formulation: even an exactly specified pure state need not assign a certain result to a chosen measurement. Interpretations address what, if anything, underlies this mathematical description. Interpretations that retain the same preparation, transformation, and probability rules agree on these experimental predictions. Proposals that modify those rules are different physical theories and must be judged by their predictions. The calculations developed here use the operational rules and do not require an additional interpretive choice.

4 The Stern–Gerlach experiment

4.1 Historical motivation and magnetic moments

Stern and Gerlach performed their experiment in 1922. Its historical setting was the old quantum theory, including the Bohr atomic model and its extensions. In the Bohr model, circular electron orbits obey the rule

$$L = n\hbar,$$

while differences of orbital energies account for discrete spectral frequencies. For hydrogen these take the Rydberg form

$$f = R_H c \left(\frac{1}{n_f^2} - \frac{1}{n_i^2} \right), \quad n_i > n_f.$$

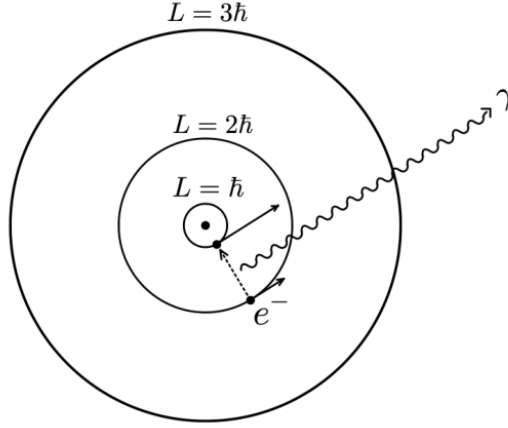


Figure 6: The original sketch of the Bohr model, whose quantization ideas motivated the historical experiment.

A puzzling idea was *spatial quantization*: how could a projection of angular momentum have discrete possible values when the axis used to define that projection is arbitrary? The Stern–Gerlach experiment turns this question into a measurement.

A circulating charge produces a magnetic dipole moment. If a charge q completes an orbit of period T and oriented area $\mathbf{A}$, its magnetic moment is

$$\boldsymbol{\mu} = \frac{q}{T} \mathbf{A}.$$

For a Kepler orbit, conservation of angular momentum gives the areal velocity

$$\frac{d\mathbf{A}}{dt} = \frac{\mathbf{L}}{2m}.$$

Integrating over a period yields $\mathbf{A}/T = \mathbf{L}/(2m)$, and hence

$$\boldsymbol{\mu} = \frac{q}{2m} \mathbf{L} \equiv \gamma \mathbf{L}. \quad (4.1)$$

For an electron, $q = -e$, so its orbital magnetic moment is antiparallel to its orbital angular momentum. The natural magnitude is the Bohr magneton,

$$\mu_B = \frac{e\hbar}{2m_e}. \quad (4.2)$$

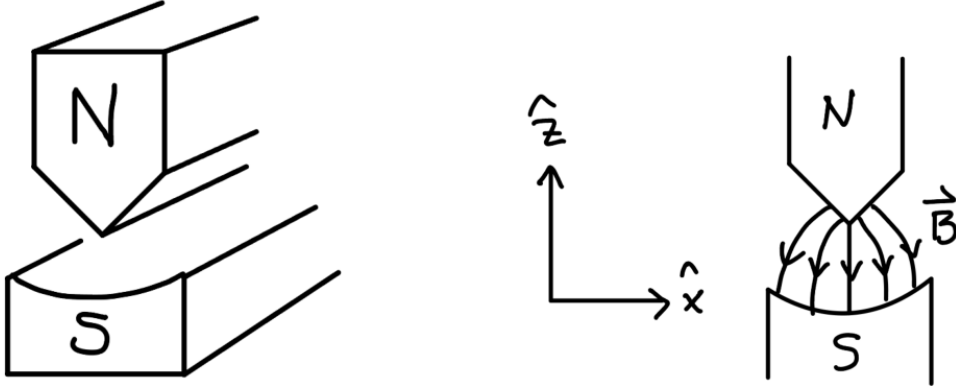


Figure 7: The source's knife-edge magnet geometry and the resulting inhomogeneous field.

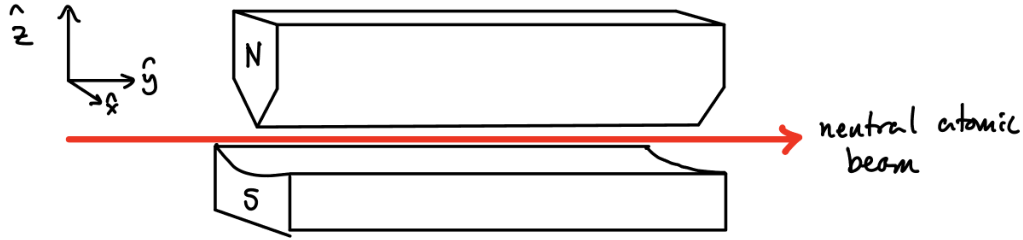


Figure 8: A neutral atomic beam traversing the magnetic field. The beam direction is $+\hat{y}$ in the source convention.

4.2 An inhomogeneous magnetic field

The apparatus sends a beam of *neutral* atoms through an inhomogeneous magnetic field. Neutrality avoids a net-charge Lorentz force, while the atoms' magnetic moments still interact with the field. A sharp pole opposite a broader pole creates the required gradient.

The interaction energy is

$$U_{\text{int}} = -\boldsymbol{\mu} \cdot \mathbf{B}. \quad (4.3)$$

There are two effects. First, the torque is

$$\mathbf{N} = \boldsymbol{\mu} \times \mathbf{B}.$$

Writing $\boldsymbol{\mu} = \gamma \mathbf{L}$ gives

$$\frac{d\mathbf{L}}{dt} = \gamma \mathbf{L} \times \mathbf{B} = (-\gamma \mathbf{B}) \times \mathbf{L}.$$

Thus the moment undergoes **Larmor precession** with angular-velocity vector

$$\boldsymbol{\omega}_L = -\gamma \mathbf{B}. \quad (4.4)$$

For electronic magnetic moments and laboratory fields of order 0.1 T, the angular frequency is of order 10^{10} s^{-1} . Over many cycles, transverse components average to

zero, leaving

$$\bar{\boldsymbol{\mu}} = (\boldsymbol{\mu} \cdot \hat{\mathbf{B}}) \hat{\mathbf{B}}.$$

Second, the spatial gradient produces a force:

$$F_i = -\frac{\partial U_{\text{int}}}{\partial x_i} = \boldsymbol{\mu} \cdot \frac{\partial \mathbf{B}}{\partial x_i}. \quad (4.5)$$

Near the middle of the apparatus, symmetry and a predominantly vertical field give

$$\bar{F}_z \simeq \mu_z \frac{\partial B_z}{\partial z}. \quad (4.6)$$

In the source geometry, $\partial B_z / \partial z < 0$ and the electronic gyromagnetic ratio is negative. A positive angular-momentum projection therefore produces an upward deflection.

If the moments were classical vectors of fixed magnitude and random isotropic orientation, their projections would range continuously from $-\mu$ to $+\mu$. More precisely, isotropy gives a uniform distribution in $\cos \theta$, and therefore a uniform projection distribution

$$\frac{dP}{dL_z} = \frac{1}{2L}, \quad -L \leq L_z \leq L.$$

The corresponding deflections would form a continuous band, broadened further by the beam profile and instrumental effects.

4.3 The observed two outcomes and electron spin

The silver beam instead splits into two components.

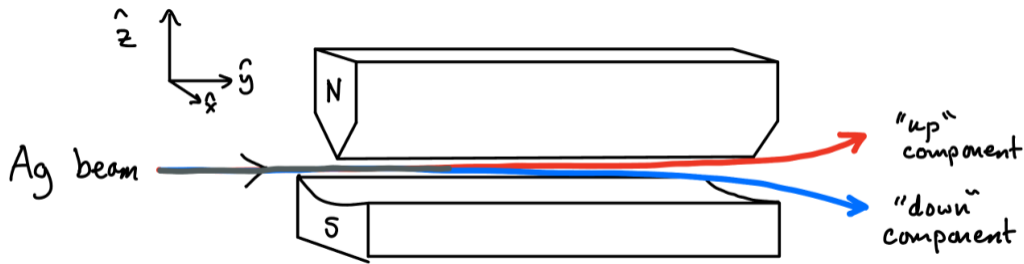


Figure 9: The observed two-beam structure in the original Stern–Gerlach sketch.

The relevant angular momentum in silver is the intrinsic spin of a single unpaired electron; the net orbital contribution in its ground state vanishes. We therefore write it as $\mathbf{S}$. Its magnetic moment is

$$\boldsymbol{\mu} = \gamma_e \mathbf{S}, \quad \gamma_e = -g_e \frac{e}{2m_e}, \quad g_e \simeq 2.0023.$$

Historically, Stern and Gerlach measured the magnetic moment rather than the angular momentum directly. The factor $g_e \simeq 2$ makes the moment associated with a spin projection $\hbar/2$ approximately equal in magnitude to the old orbital estimate for

angular momentum $\hbar$. This compensation is why the angular-momentum factor of two was not apparent from their magnetic splitting alone.

The magnetic deflection together with the gyromagnetic ratio gives

$$S_z = +\frac{\hbar}{2} \quad \text{or} \quad S_z = -\frac{\hbar}{2}. \quad (4.7)$$

These are projections, not the magnitude of the spin vector. For a spin-1/2 system, $S^2 = 3\hbar^2/4$, so the magnitude associated with S^2 is $\sqrt{3}\hbar/2$.

Rotating the apparatus changes which component is measured but still produces two possible results, $\pm\hbar/2$. Discrete projections therefore do not select a preferred direction in space. Understanding how different measurement axes are related is one of the main reasons to introduce a Hilbert space.

5 Sequential measurements and coherent recombination

5.1 Preparing and measuring a spin state

Write SG_z for a Stern–Gerlach device oriented to measure S_z . We label the states in its two output beams by $|z, +\rangle$ and $|z, -\rangle$. The notation is at first just a label for a reproducible preparation.

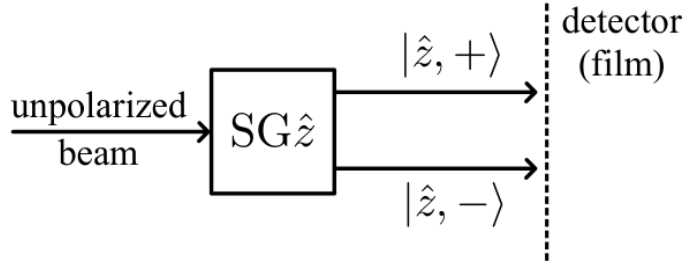


Figure 10: The two spin preparations produced by a z -oriented Stern–Gerlach apparatus.

Send only the upper output into a second SG_z device. All transmitted atoms leave the second device in its upper output:

$$P(z, + | z, +) = 1, \quad P(z, - | z, +) = 0. \quad (5.1)$$

The first device prepares a definite S_z state, and an immediate repeat of the ideal measurement reproduces the outcome. An unpolarized input gives equal intensities in the two outputs. It can be represented as an equal statistical mixture of the two z states, although this representation is not unique.

Now turn the second apparatus so that it measures S_x . A prepared $|z, +\rangle$ beam splits equally:

$$P(x, + | z, +) = P(x, - | z, +) = \frac{1}{2}. \quad (5.2)$$

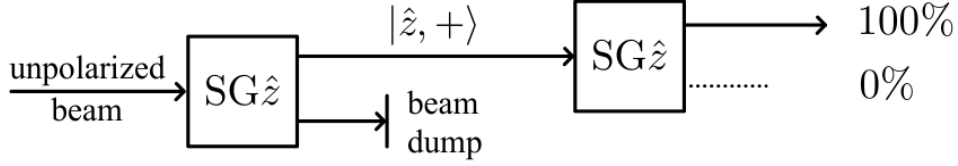


Figure 11: Filtering the $z, +$ output and repeating the same measurement gives $z, +$ with certainty.

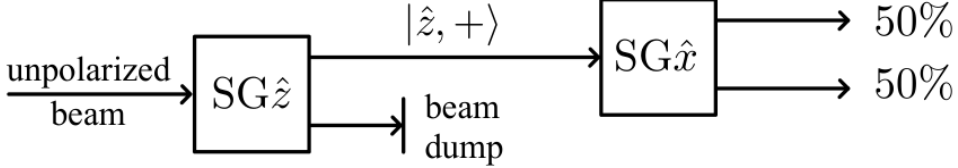


Figure 12: A definite S_z preparation gives equal probabilities for the two S_x outcomes.

The same holds for a prepared $|z, -\rangle$ beam.

One might imagine that the original beam contains four subpopulations with predetermined pairs of values (S_z, S_x) . This simple model predicts that selecting $z, +$ and then $x, +$ leaves only atoms that still have $z, +$. Test that prediction with a third device measuring S_z :

$$|z, +\rangle \longrightarrow \text{select } x, + \longrightarrow S_z. \quad (5.3)$$

The final outcomes are again equally likely.

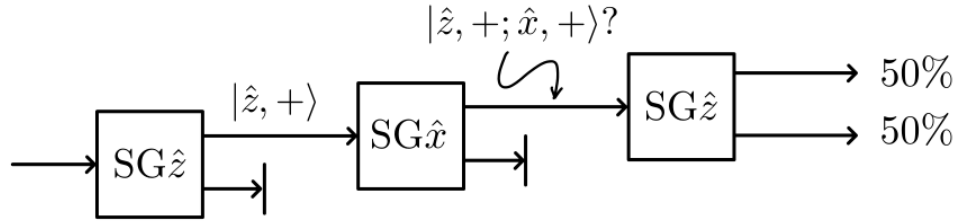


Figure 13: After the intermediate $x, +$ selection, a final S_z measurement gives both outcomes.

The naive picture of filtering fixed, undisturbed values of both components fails. The intermediate measurement prepares a new state and removes the earlier certainty about S_z .

5.2 Why precession cannot simply be tuned away in a resolved measurement

The rapid Larmor precession in a Stern–Gerlach magnet suggests a physical mechanism for changing transverse spin information. Atoms entering at slightly different heights

experience different fields and accumulate different precession angles. Could one make the beam so well controlled that every atom executes exactly the same number of precessions, while still resolving the two spin outcomes?

Exercise 5.1 (Resolving the impulse and the precession angle). Let $G = \partial B_z / \partial z$, let T be the transit time, and let μ denote the magnitude of the magnetic moment projection. The two impulses are approximately $\pm |\mu GT|$. Relate the resolution of these impulses to the uncertainty in the accumulated precession angle.

Solution. To resolve the outgoing momentum branches sharply, require

$$\sigma_{p_z}^{\text{in}} \ll |\mu GT|.$$

The uncertainty relation implies

$$\sigma_z^{\text{in}} \geq \frac{\hbar}{2\sigma_{p_z}^{\text{in}}} \gg \frac{\hbar}{2|\mu GT|}.$$

The accumulated Larmor angle is approximately $\phi(z) = -\gamma B_z(z)T$. A position spread produces an angle spread

$$\sigma_\phi \simeq |\gamma GT| \sigma_z^{\text{in}} \gg \frac{\hbar|\gamma|}{2|\mu|}.$$

For a spin-1/2 moment, $|\mu| = |\gamma|\hbar/2$, so the last quantity is 1. Sharp separation of the momentum branches is therefore incompatible with a sharply controlled transverse precession phase in this arrangement. This gives a concrete consistency check on the failure of simultaneous sharp spin-component assignments. $\square$

5.3 A modified apparatus that splits and recombines

Instead of detecting or blocking a separated beam, we can try to bring the two paths back together coherently. The modified Stern–Gerlach apparatus uses three magnetic regions. The first separates the paths; the second has the opposite gradient and approximately twice the transit time; the last has the original gradient and reunites them.

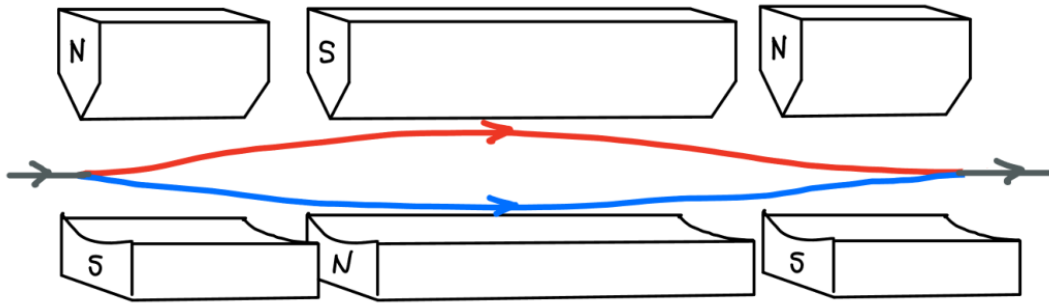


Figure 14: The source's three-region Stern–Gerlach recombining. Opposite intermediate forces reverse the separation and allow the two paths to meet again.

For an idealized branch acceleration a , take successive accelerations $a, -a, a$ for times $T, 2T, T$. Starting from zero transverse velocity, the velocities at the boundaries are

$$0, \quad aT, \quad -aT, \quad 0.$$

The transverse displacements in the three regions are

$$\frac{1}{2}aT^2, \quad (aT)(2T) + \frac{1}{2}(-a)(2T)^2 = 0, \quad -aT^2 + \frac{1}{2}aT^2 = -\frac{1}{2}aT^2.$$

Both net displacement and net momentum change vanish. The opposite spin branch follows the reflected trajectory. Thus the apparatus can restore a common outgoing spatial mode while allowing access to the separated paths in the middle.

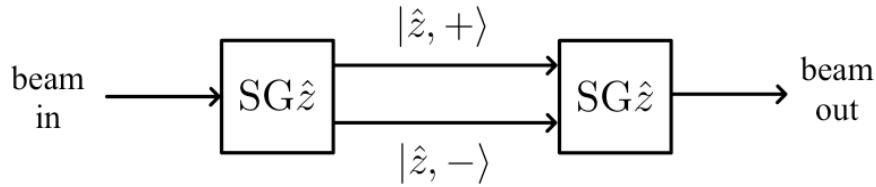


Figure 15: An abstract split-and-recombine device. Either internal path can be blocked, or both can remain open.

Spatial recombination is not by itself sufficient: relative phases must also be controlled, and no remaining environment or detector record may distinguish the paths. The ideal device below is tuned so that its two branches recombine with the required relative phase.

Prepare $|z, +\rangle$, send it through an x -oriented modified device, and measure S_z at the end. If either internal path is blocked, half the incident atoms are lost. Among the transmitted atoms, the two final z outcomes are equally probable. Relative to the beam entering the modified device, each final outcome therefore has probability $1/4$.

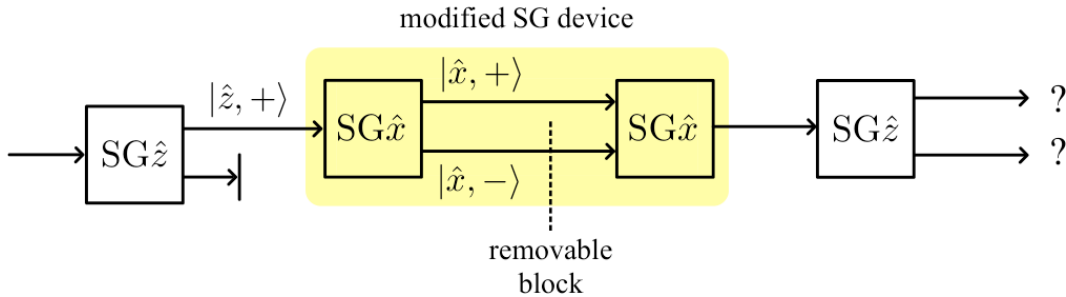


Figure 16: The fourth thought experiment: a $z, +$ preparation, an x split-and-recombine stage with a removable block, and a final z measurement.

With both paths open and the phases properly tuned, all atoms instead emerge as $z, +$:

$$P(z, +) = 1, \quad P(z, -) = 0. \quad (5.4)$$

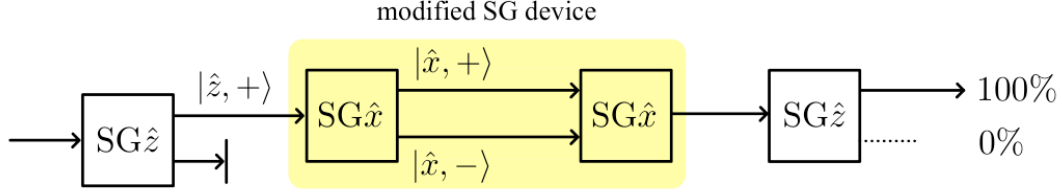


Figure 17: With both coherent paths open, the ideal recombiner restores the initial $z, +$ state.

This is a spin version of double-slit interference. It shows why merely separating the beams does not have the same effect as recording an outcome and then discarding that record.

5.4 The amplitudes behind the four experiments

Choose phases so that the amplitude for either path to reach final $z, +$ is $1/2$. The amplitudes for final $z, -$ have opposite signs:

Internal path	final $z, +$	final $z, -$
$x, +$	$1/2$	$1/2$
$x, -$	$1/2$	$-1/2$

With a single path open, the squared magnitude of each entry is $1/4$. With both paths open, the amplitudes add to 1 and 0. The two probability distributions are different because the coherent experiment adds amplitudes before squaring.

This motivates the column vectors

$$|z, +\rangle \leftrightarrow \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |z, -\rangle \leftrightarrow \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad |x, \pm\rangle \leftrightarrow \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ \pm 1 \end{pmatrix}. \quad (5.5)$$

Each pure state in this display has unit norm. The unnormalized contribution from one branch of the modified apparatus has an additional factor $1/\sqrt{2}$ because that branch contains only half the incident probability.

6 Complex two-state vectors

6.1 Changing between the z and x descriptions

A general normalized two-state vector is

$$|\psi\rangle = \alpha|z, +\rangle + \beta|z, -\rangle, \quad |\alpha|^2 + |\beta|^2 = 1. \quad (6.1)$$

Its z measurement probabilities are $P(z, +) = |\alpha|^2$ and $P(z, -) = |\beta|^2$. Using (5.5),

$$|z, +\rangle = \frac{|x, +\rangle + |x, -\rangle}{\sqrt{2}}, \quad (6.2)$$

$$|z, -\rangle = \frac{|x, +\rangle - |x, -\rangle}{\sqrt{2}}, \quad (6.3)$$

and therefore

$$|\psi\rangle = \frac{\alpha + \beta}{\sqrt{2}}|x, +\rangle + \frac{\alpha - \beta}{\sqrt{2}}|x, -\rangle. \quad (6.4)$$

It follows that

$$P(x, \pm) = \frac{|\alpha \pm \beta|^2}{2}. \quad (6.5)$$

The sum is one because the cross terms cancel:

$$\frac{|\alpha + \beta|^2 + |\alpha - \beta|^2}{2} = |\alpha|^2 + |\beta|^2.$$

The change of coordinates is

$$\begin{pmatrix} \psi_{x,+} \\ \psi_{x,-} \end{pmatrix} = H \begin{pmatrix} \alpha \\ \beta \end{pmatrix}, \quad H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}. \quad (6.6)$$

Since $H^2 = \mathbf{1}$, the same matrix recovers the z components from the x components. Neither choice of axis contains more state information.

6.2 Why complex amplitudes are needed

Prepare the $y, +$ output of an apparatus oriented along $\hat{y}$. Measurements along either z or x give equal probabilities.

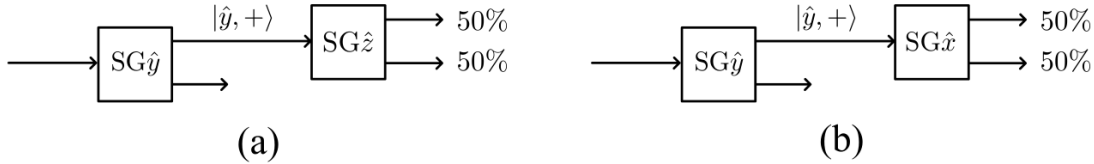


Figure 18: The two rotated preparations that determine the relative phase of the $y, +$ state.

Equal z probabilities imply that, after choosing an overall phase,

$$|y, +\rangle = \frac{1}{\sqrt{2}}|z, +\rangle + \frac{e^{i\phi}}{\sqrt{2}}|z, -\rangle.$$

Equal x probabilities then require

$$P(x, \pm | y, +) = \frac{1}{4}|1 \pm e^{i\phi}|^2 = \frac{1}{2}(1 \pm \cos \phi) = \frac{1}{2}.$$

Thus $\cos \phi = 0$, so the relative phase is $\phi = \pm\pi/2$. A conventional choice is

$$|y, \pm\rangle = \frac{1}{\sqrt{2}}(|z, +\rangle \pm i|z, -\rangle). \quad (6.7)$$

The y states cannot be represented by real two-component amplitudes in the same convention that made the x and z states real. Complex numbers encode the required relative phase.

Simultaneously complex-conjugating all state components and all operator matrices gives an alternative representation with identical predictions. Complex-conjugating only a state while holding the measurement operators fixed is generally a physical change: in this convention it exchanges $|y, +\rangle$ and $|y, -\rangle$.

For a general state, the y amplitudes are

$$\psi_{y,+} = \frac{\alpha - i\beta}{\sqrt{2}}, \quad \psi_{y,-} = \frac{\alpha + i\beta}{\sqrt{2}}. \quad (6.8)$$

Equivalently,

$$\begin{pmatrix} \psi_{y,+} \\ \psi_{y,-} \end{pmatrix} = U_y \begin{pmatrix} \alpha \\ \beta \end{pmatrix}, \quad U_y = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix}. \quad (6.9)$$

Exercise 6.1 (Normalization and unitary changes of components). Verify that both H in (6.6) and U_y in (6.9) preserve the normalization of every state vector.

Solution. Direct multiplication gives

$$H^\dagger H = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}^2 = \mathbf{1},$$

and

$$U_y^\dagger U_y = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix} \begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix} = \mathbf{1}.$$

Therefore $(U\psi)^\dagger(U\psi) = \psi^\dagger\psi$ for either change of components. Such matrices are called unitary. They preserve the complete inner product, not merely the norms of the displayed basis states. $\square$

6.3 Why an unpolarized beam is not a pure state

Exercise 6.2 (No single ket gives an unpolarized spin-1/2 state). Show that there is no normalized column $(\alpha, \beta)^\top$ that gives probabilities $1/2, 1/2$ for measurements along each of x , y , and z .

Solution. Equal z probabilities require $|\alpha|^2 = |\beta|^2 = 1/2$. From (6.5), equal x probabilities require

$$\text{Re}(\alpha^*\beta) = 0.$$

The y probabilities are

$$P(y, \pm) = \frac{1}{2} |\alpha \mp i\beta|^2 = \frac{1}{2} \pm \text{Im}(\alpha^*\beta),$$

so equality also requires $\text{Im}(\alpha^*\beta) = 0$. Together these conditions give $\alpha^*\beta = 0$, in contradiction with $|\alpha\beta| = 1/2$.

The unpolarized beam is instead a mixed state, represented by

$$\rho = \frac{1}{2} |z, +\rangle \langle z, +| + \frac{1}{2} |z, -\rangle \langle z, -| = \frac{1}{2} \mathbf{1}.$$

It can equally be decomposed as a half-and-half mixture in the x basis or in the y basis. This is an explicit example of the nonuniqueness of a mixed-state ensemble decomposition. $\square$

Part II

Finite-Dimensional Hilbert Spaces

Chapter 2

7 Kets, bras, and linear maps

7.1 Complex vector spaces

A complex vector space $\mathcal{H}$ has vector addition and multiplication by complex scalars. In Dirac notation, a vector is a **ket**, written $|\psi\rangle$. If $|\psi\rangle, |\chi\rangle \in \mathcal{H}$ and $a, b \in \mathbb{C}$, then

$$a|\psi\rangle + b|\chi\rangle \in \mathcal{H}.$$

A basis consists of linearly independent vectors that span the space. In this chapter $\dim \mathcal{H} = N < \infty$. A general ket is therefore determined by N complex components once a basis has been chosen. Infinite-dimensional spaces require an additional completeness condition and attention to convergence and operator domains; these issues will be introduced when needed.

The abstract ket and its coordinate column are different objects. The ket does not depend on a choice of basis, whereas its components do. This distinction is the same one encountered when a geometric vector is represented by different Cartesian coordinates.

7.2 The dual space

A **bra** is a complex-linear map from $\mathcal{H}$ to $\mathbb{C}$. The bras form the dual vector space $\mathcal{H}^*$. If $\langle\omega|$ is a bra, its value on a ket is written

$$\langle\omega|(|\psi\rangle) = \langle\omega | \psi\rangle.$$

Linearity means

$$\langle\omega|(a|\psi\rangle + b|\chi\rangle) = a\langle\omega | \psi\rangle + b\langle\omega | \chi\rangle. \quad (7.1)$$

At this stage a bra is simply a linear functional. The association of a particular bra with a particular ket will come from an inner product.

Bras can themselves be added and multiplied by scalars, pointwise:

$$(a\langle\omega| + b\langle\eta|)|\psi\rangle = a\langle\omega | \psi\rangle + b\langle\eta | \psi\rangle.$$

Complex scalars commute with bras and kets; for example $a|\psi\rangle = |\psi\rangle a$. This does not mean that arbitrary products of bras and kets can be reordered. Their order determines whether the result is a scalar, a vector, a functional, or an operator.

7.3 Operators and their algebra

A linear operator $X : \mathcal{H} \rightarrow \mathcal{H}$ obeys

$$X(a|\psi\rangle + b|\chi\rangle) = aX|\psi\rangle + bX|\chi\rangle.$$

It acts on a ket from the left. It also induces an action on bras from the right: $\langle\omega|X$ is the functional satisfying

$$(\langle\omega|X)|\psi\rangle = \langle\omega|(X|\psi\rangle).$$

The **identity** $\mathbf{1}$ satisfies $\mathbf{1}|\psi\rangle = |\psi\rangle$ for every ket, and the zero operator sends every ket to zero. Two operators are equal precisely when they act identically on every vector; checking a basis is sufficient.

Addition and composition are defined by

$$(X + Y)|\psi\rangle = X|\psi\rangle + Y|\psi\rangle, \quad (XY)|\psi\rangle = X(Y|\psi\rangle).$$

Composition is associative but is generally not commutative. The **commutator**

$$[X, Y] = XY - YX \tag{7.2}$$

measures its failure to commute. In a matrix product, the rightmost operator acts first. The same order convention will be essential for sequential measurements and time evolution.

7.4 Outer products and rank

A ket followed by a bra defines an **outer product**:

$$(|\psi\rangle\langle\omega|)|\chi\rangle = |\psi\rangle\langle\omega|\chi\rangle. \tag{7.3}$$

Every output lies in the one-dimensional space spanned by $|\psi\rangle$. Thus a nonzero outer product has rank one. By contrast, the reversed product $\langle\omega|\chi\rangle$ is a complex number.

For any operator X , the **range** is

$$\text{ran } X = \{X|\psi\rangle : |\psi\rangle \in \mathcal{H}\},$$

and its rank is $\dim \text{ran } X$. Equivalently, the rank is the minimum number of rank-one terms needed to express X as a sum of outer products. To see this, suppose

$$X = \sum_{a=1}^r |u_a\rangle\langle\omega_a|.$$

Then $\text{ran } X \subseteq \text{span}\{|u_1\rangle, \dots, |u_r\rangle\}$, so $\text{rank } X \leq r$. Conversely, choose a basis $|u_a\rangle$ of the range. The coefficients of $X|\psi\rangle$ in this basis are linear functions of $|\psi\rangle$, hence bras $\langle\omega_a|$, and give such a decomposition with $r = \text{rank } X$. In particular, every operator on an N -dimensional space has rank at most N .

8 Inner products and adjoints

8.1 The Hilbert-space inner product

Definition 8.1. A complex inner product assigns a number $\langle \psi | \chi \rangle$ to each pair of vectors and satisfies

$$\langle \psi | \chi \rangle = \langle \chi | \psi \rangle^*, \quad \text{conjugate symmetry,} \quad (8.1)$$

$$\langle \psi | a\chi + b\eta \rangle = a\langle \psi | \chi \rangle + b\langle \psi | \eta \rangle, \quad \text{linearity in the second slot,} \quad (8.2)$$

$$\langle \psi | \psi \rangle \geq 0, \quad \langle \psi | \psi \rangle = 0 \iff |\psi\rangle = 0, \quad \text{positive definiteness.} \quad (8.3)$$

A finite-dimensional complex vector space equipped with an inner product is a Hilbert space. Its norm is

$$\| |\psi\rangle \| = \sqrt{\langle \psi | \psi \rangle}.$$

Conjugate symmetry and second-slot linearity imply first-slot *antilinearity*:

$$\langle a\psi + b\chi | \eta \rangle = a^* \langle \psi | \eta \rangle + b^* \langle \chi | \eta \rangle. \quad (8.4)$$

Thus the inner product associates a bra $\langle \psi |$ to each ket $|\psi\rangle$, but the association is antilinear:

$$(a|\psi\rangle + b|\chi\rangle)^\dagger = a^* \langle \psi | + b^* \langle \chi |. \quad (8.5)$$

The dagger notation records this association.

Proposition 8.2 (The ket–bra correspondence). *Every bra on a finite-dimensional Hilbert space is the inner-product bra associated with a unique ket.*

Proof. First establish uniqueness. If $\langle \psi | = \langle \chi |$, then $\langle \psi - \chi | = 0$, so evaluating the zero functional on $|\psi - \chi\rangle$ gives

$$\| |\psi\rangle - |\chi\rangle \|^2 = 0.$$

Positive definiteness implies $|\psi\rangle = |\chi\rangle$.

The association is thus injective. The space of bras has the same finite complex dimension N as the space of kets: given any basis, a linear functional is determined by its N values on that basis. An injective antilinear map between equal finite-dimensional complex spaces is surjective, since its image is a complex subspace of dimension N . Once an orthonormal basis $|i\rangle$ is chosen, the correspondence is explicit: a bra with values $b_i = \langle \omega | i \rangle$ is associated with

$$|\omega\rangle = \sum_i b_i^* |i\rangle.$$

□

This finite-dimensional statement is the version of the Riesz representation theorem needed here. It does not mean that the dual-space definition of a bra was antilinear: each bra is linear in its ket argument. The antilinearity concerns the map that constructs a bra from a ket.

8.2 Adjoints of operators

For a linear operator X , the bra $\langle\psi|X$ has a unique corresponding ket. This defines the **adjoint** $X^\dagger$ by

$$X^\dagger|\psi\rangle = (\langle\psi|X)^\dagger. \quad (8.6)$$

Equivalently, for all kets,

$$\langle\phi|X\psi\rangle = \langle X^\dagger\phi|\psi\rangle. \quad (8.7)$$

The adjoint is a linear operator: the two antilinear ket–bra conversions involved in its definition cancel one another.

Proposition 8.3 (Adjoint identities). *For operators X, Y and scalars a, b ,*

$$(aX + bY)^\dagger = a^*X^\dagger + b^*Y^\dagger, \quad (8.8)$$

$$(XY)^\dagger = Y^\dagger X^\dagger, \quad (8.9)$$

$$(X^\dagger)^\dagger = X, \quad (8.10)$$

$$(|\psi\rangle\langle\omega|)^\dagger = |\omega\rangle\langle\psi|. \quad (8.11)$$

Furthermore,

$$(\langle\phi|X|\psi\rangle)^* = \langle\psi|X^\dagger|\phi\rangle, \quad (X|\psi\rangle)^\dagger = \langle\psi|X^\dagger. \quad (8.12)$$

Proof. For the sum, use antilinearity in the first slot:

$$\begin{aligned} \langle\phi|(aX + bY)\psi\rangle &= a\langle X^\dagger\phi|\psi\rangle + b\langle Y^\dagger\phi|\psi\rangle \\ &= \langle(a^*X^\dagger + b^*Y^\dagger)\phi|\psi\rangle. \end{aligned}$$

Uniqueness of the adjoint proves (8.8). For the product, apply the defining relation twice:

$$\langle\phi|XY\psi\rangle = \langle X^\dagger\phi|Y\psi\rangle = \langle Y^\dagger X^\dagger\phi|\psi\rangle.$$

This reverses the order, giving (8.9). Conjugate symmetry applied to (8.7) yields the matrix-element identity. Applying it a second time gives the double-adjoint identity.

Finally, the matrix element of an outer product is

$$\langle\phi||\psi\rangle\langle\omega||\chi\rangle = \langle\phi|\psi\rangle\langle\omega|\chi\rangle.$$

Its complex conjugate is $\langle\chi|\omega\rangle\langle\psi|\phi\rangle$, which is the matrix element of $|\omega\rangle\langle\psi|$ in the reversed slots. Hence (8.11). The bra identity in (8.12) follows by evaluating both sides on an arbitrary ket. $\square$

Definition 8.4. An operator is

- **self-adjoint** or **Hermitian** if $X^\dagger = X$;
- **unitary** if $U^\dagger U = UU^\dagger = \mathbf{1}$;
- **normal** if $X^\dagger X = XX^\dagger$.

Every self-adjoint operator is normal, and every unitary operator is normal. A normal operator need be neither self-adjoint nor unitary.

9 Orthonormal bases and projectors

A vector is normalized if its norm is one. Two vectors are orthogonal if their inner product is zero. A set $\{|i\rangle\}$ is orthonormal if

$$\langle i | j \rangle = \delta_{ij}. \quad (9.1)$$

Such a set is linearly independent: taking the inner product of $\sum_i c_i |i\rangle = 0$ with $\langle j|$ gives $c_j = 0$. It therefore contains at most N vectors, and an orthonormal set of N vectors is a basis.

An orthonormal basis can be built from any basis by Gram–Schmidt. Normalize the first vector. From the next vector subtract its projections onto all previously constructed orthonormal vectors, then normalize the nonzero remainder. At step k , linear independence ensures that the remainder is not zero. In formulas,

$$|v_k\rangle = |u_k\rangle - \sum_{j < k} |e_j\rangle \langle e_j | u_k \rangle, \quad |e_k\rangle = \frac{|v_k\rangle}{\|v_k\|}.$$

An orthonormal set can similarly be extended to a complete orthonormal basis.

9.1 Resolution of the identity

In an orthonormal basis, a vector has the expansion

$$|\psi\rangle = \sum_{i=1}^N |i\rangle \langle i | \psi \rangle. \quad (9.2)$$

Indeed, if $|\psi\rangle = \sum_i c_i |i\rangle$, taking $\langle j|$ gives $c_j = \langle j | \psi \rangle$. Since (9.2) holds for every ket,

$$\boxed{\mathbf{1} = \sum_{i=1}^N |i\rangle \langle i|.} \quad (9.3)$$

This is the **completeness relation**. It is an operator equality, and inserting it between factors is the basic way to turn abstract operator expressions into component sums. It also gives

$$\langle \phi | \psi \rangle = \sum_i \langle \phi | i \rangle \langle i | \psi \rangle, \quad \|\psi\|^2 = \sum_i |\langle i | \psi \rangle|^2.$$

9.2 Projectors onto subspaces

A **projector** is an idempotent operator, $\Pi^2 = \Pi$. An **orthogonal projector** additionally satisfies $\Pi^\dagger = \Pi$. Unless stated otherwise, the projectors used for measurements below will be orthogonal projectors.

Let $|1\rangle, \dots, |r\rangle$ be an orthonormal basis of a subspace V . Then

$$\Pi_V = \sum_{i=1}^r |i\rangle \langle i| \quad (9.4)$$

is its orthogonal projector. Its adjoint equals itself, while

$$\Pi_V^2 = \sum_{i,j} |i\rangle\langle i|j\rangle\langle j| = \sum_i |i\rangle\langle i| = \Pi_V.$$

Every ket has the decomposition

$$|\psi\rangle = \Pi_V|\psi\rangle + (\mathbf{1} - \Pi_V)|\psi\rangle,$$

where the first term lies in V and the second is orthogonal to V . Thus Π_V acts as the identity on V and as zero on $V^\perp$. These properties characterize it without reference to a particular orthonormal basis of V .

For a normalized ket $|\psi\rangle$, the rank-one projector is $\Pi_\psi = |\psi\rangle\langle\psi|$. If the ket is not normalized, the corresponding projector is

$$\Pi_\psi = \frac{|\psi\rangle\langle\psi|}{\langle\psi|\psi\rangle}.$$

Multiplying the ket by a nonzero complex scalar leaves this operator unchanged.

10 Eigenvectors and the spectral theorem

10.1 Self-adjoint eigenvalue properties

An eigenvector of X is a *nonzero* ket satisfying

$$X|\psi\rangle = \lambda|\psi\rangle. \quad (10.1)$$

The number λ is its eigenvalue. The zero vector is excluded because it satisfies this equation for every λ .

If X is self-adjoint, its eigenvalues are real. To prove this, evaluate the expectation in an eigenvector:

$$\lambda\langle\psi|\psi\rangle = \langle\psi|X|\psi\rangle = (\langle\psi|X|\psi\rangle)^* = \lambda^*\langle\psi|\psi\rangle.$$

Since $\langle\psi|\psi\rangle > 0$, $\lambda = \lambda^*$. If $|\psi\rangle$ and $|\chi\rangle$ have distinct eigenvalues λ and μ , then

$$\lambda\langle\psi|\chi\rangle = \langle X\psi|\chi\rangle = \langle\psi|X\chi\rangle = \mu\langle\psi|\chi\rangle.$$

Therefore $\langle\psi|\chi\rangle = 0$. Distinct eigenvalues of a self-adjoint operator have orthogonal eigenspaces.

10.2 The corresponding facts for normal operators

Proposition 10.1. *An operator X is normal if and only if*

$$\|X|\psi\rangle\| = \|X^\dagger|\psi\rangle\| \quad \text{for every } |\psi\rangle. \quad (10.2)$$

For a normal operator, $X|\psi\rangle = \lambda|\psi\rangle$ implies $X^\dagger|\psi\rangle = \lambda^*|\psi\rangle$. Eigenvectors with distinct eigenvalues are orthogonal.

Proof. Subtract the two squared norms:

$$\|X|\psi\rangle\|^2 - \|X^\dagger|\psi\rangle\|^2 = \langle\psi|(X^\dagger X - X X^\dagger)|\psi\rangle.$$

Normality makes this zero. Conversely, if it vanishes for every ket, the Hermitian operator $A = X^\dagger X - X X^\dagger$ has zero quadratic form. In an orthonormal basis, using $|i\rangle$ gives $A_{ii} = 0$. Using $|i\rangle + |j\rangle$ gives $2 \operatorname{Re} A_{ij} = 0$, and using $|i\rangle + i|j\rangle$ gives $\operatorname{Im} A_{ij} = 0$. Hence every matrix element of A is zero and X is normal.

Now $Y = X - \lambda \mathbf{1}$ is normal whenever X is normal. If $Y|\psi\rangle = 0$, (10.2) implies $Y^\dagger|\psi\rangle = 0$, which is the claimed adjoint-eigenvector relation. For $X|\chi\rangle = \mu|\chi\rangle$, it follows that

$$\mu\langle\psi|\chi\rangle = \langle\psi|X\chi\rangle = \langle X^\dagger\psi|\chi\rangle = \lambda\langle\psi|\chi\rangle.$$

If $\lambda \neq \mu$, the overlap must vanish. $\square$

Normal operators can have complex eigenvalues. Orthogonality of distinct eigenspaces does not require those eigenvalues to be real.

10.3 Proof of the finite-dimensional spectral theorem

Theorem 10.2 (Spectral theorem). *A complex finite-dimensional operator is normal if and only if it has an orthonormal basis of eigenvectors. In such a basis,*

$$X = \sum_{i=1}^N \lambda_i |i\rangle\langle i|. \quad (10.3)$$

For a self-adjoint operator, all λ_i are real.

Proof. We prove existence of the eigenbasis by induction on the dimension. Over $\mathbb{C}$, the characteristic polynomial has a root, so a nonzero eigenvector exists. Normalize one such vector $|1\rangle$, with eigenvalue λ_1 .

For any $|v\rangle$ orthogonal to $|1\rangle$, the normal-operator properties give

$$\langle 1 | Xv \rangle = \langle X^\dagger 1 | v \rangle = \lambda_1 \langle 1 | v \rangle = 0.$$

Similarly, $\langle 1 | X^\dagger v \rangle = 0$. Thus $\operatorname{span}\{|1\rangle\}^\perp$ is invariant under both X and $X^\dagger$. Their restrictions to this $(N-1)$ -dimensional space are still adjoints of one another and still commute. The restricted operator is therefore normal. By induction, its orthogonal complement has an orthonormal eigenbasis. Together with $|1\rangle$, this gives an orthonormal eigenbasis of all of $\mathcal{H}$.

Insert the completeness relation on the right of X :

$$X = X\mathbf{1} = \sum_i X|i\rangle\langle i| = \sum_i \lambda_i |i\rangle\langle i|,$$

proving (10.3). Conversely, an operator with this expansion has

$$X^\dagger = \sum_i \lambda_i^* |i\rangle\langle i|, \quad X^\dagger X = \sum_i |\lambda_i|^2 |i\rangle\langle i| = X X^\dagger,$$

so it is normal. If X is self-adjoint, the preceding eigenvalue argument makes every λ_i real. $\square$

Exercise 10.3 (What is and is not contained in an eigenbasis). Show that an orthonormal eigenbasis contains every eigenvalue of the operator, but need not contain every eigenvector.

Solution. Expand any eigenvector $|\psi\rangle = \sum_i c_i |i\rangle$ of eigenvalue λ . Then

$$0 = (X - \lambda \mathbf{1})|\psi\rangle = \sum_i (\lambda_i - \lambda) c_i |i\rangle.$$

Each coefficient vanishes. Since at least one $c_i \neq 0$, λ equals at least one λ_i . Moreover, $|\psi\rangle$ is a linear combination only of basis vectors with that eigenvalue.

There may be infinitely many different eigenvector directions in a degenerate eigenspace. For example, every nonzero vector is an eigenvector of $\mathbf{1}$. A finite orthonormal basis cannot list all of them. $\square$

A normal operator is self-adjoint exactly when its spectrum is real. The forward direction was already proved; the reverse direction follows by taking the adjoint of (10.3).

10.4 Positive operators

A self-adjoint operator A is **positive semidefinite**, written $A \geq 0$, if

$$\langle \psi | A | \psi \rangle \geq 0 \quad \text{for every } |\psi\rangle.$$

It is positive definite if the inequality is strict for every nonzero ket. In an eigenbasis,

$$\langle \psi | A | \psi \rangle = \sum_i a_i |c_i|^2,$$

so positivity is equivalent to $a_i \geq 0$ for every eigenvalue; positive definiteness is equivalent to $a_i > 0$ for all of them. In particular,

$$X^\dagger X \geq 0, \quad \langle \psi | X^\dagger X | \psi \rangle = \|X|\psi\rangle\|^2.$$

A positive operator has a positive square root

$$A^{1/2} = \sum_i \sqrt{a_i} |i\rangle \langle i|, \quad (A^{1/2})^2 = A.$$

Consequently, for $A \geq 0$,

$$\langle \psi | A | \psi \rangle = 0 \iff A^{1/2}|\psi\rangle = 0 \iff A|\psi\rangle = 0.$$

This implication will be useful when a zero measurement probability is converted into an operator condition.

11 Degeneracy, simultaneous eigenvectors, and spectral projectors

11.1 Degenerate eigenspaces

If an eigenvalue occurs more than once in an eigenbasis, its eigenspace has dimension greater than one and the eigenvalue is **degenerate**. Any orthonormal basis within that eigenspace is allowed. Replacing it by a unitary linear combination changes the eigenbasis without changing the operator.

For distinct eigenvalues λ_a , let Π_a be the orthogonal projector onto the entire corresponding eigenspace. Then

$$X = \sum_a \lambda_a \Pi_a, \quad \Pi_a \Pi_b = \delta_{ab} \Pi_a, \quad \sum_a \Pi_a = \mathbf{1}. \quad (11.1)$$

The rank of Π_a is the degeneracy of λ_a . Although basis vectors inside a degenerate eigenspace are not unique, the eigenspace and its projector are unique. In finite dimension the latter can also be written as a polynomial in X :

$$\Pi_a = \prod_{b \neq a} \frac{X - \lambda_b \mathbf{1}}{\lambda_a - \lambda_b}. \quad (11.2)$$

This polynomial equals one on the λ_a eigenspace and zero on every other eigenspace.

11.2 Complete orthogonal projector families

A family of orthogonal projectors is *complete* if its sum is the identity. For such a family, completeness already forces pairwise orthogonality.

Proposition 11.1. *If each Π_a is self-adjoint and idempotent and $\sum_a \Pi_a = \mathbf{1}$, then $\Pi_a \Pi_b = 0$ for $a \neq b$.*

Proof. Take $|v\rangle \in \text{ran } \Pi_b$, so $\Pi_b |v\rangle = |v\rangle$. Completeness gives

$$\| |v\rangle \|^2 = \sum_a \langle v | \Pi_a | v \rangle = \sum_a \| \Pi_a |v\rangle \|^2.$$

The $a = b$ term already equals $\| |v\rangle \|^2$. Every other term is nonnegative, so every other term must vanish. Hence $\Pi_a |v\rangle = 0$ for $a \neq b$. Since every vector of the form $\Pi_b |\psi\rangle$ belongs to this range, $\Pi_a \Pi_b = 0$. $\square$

11.3 Commuting normal operators

Theorem 11.2. *A finite family of pairwise commuting normal operators has a common orthonormal eigenbasis. In particular, commuting self-adjoint operators can be simultaneously diagonalized.*

Proof. First consider two normal operators X and Y with $XY = YX$. By (11.2), their spectral projectors are polynomials in X and Y , respectively. Polynomials in commuting operators commute, so

$$[\Pi_a^X, \Pi_b^Y] = 0.$$

Their products

$$\Pi_{ab} = \Pi_a^X \Pi_b^Y \quad (11.3)$$

are self-adjoint because the factors commute, and idempotent because both factors are projectors. Their range is the intersection of the two eigenspaces. They are pairwise orthogonal and complete:

$$\sum_{a,b} \Pi_{ab} = \left(\sum_a \Pi_a^X \right) \left(\sum_b \Pi_b^Y \right) = \mathbf{1}.$$

Choose an orthonormal basis in every nonzero joint range. The union is a common orthonormal eigenbasis. Repeating the construction with the remaining commuting operators proves the finite-family statement. $\square$

The projector argument also proves the useful equivalence

$$[X, Y] = 0 \iff [\Pi_a^X, \Pi_b^Y] = 0 \text{ for all } a, b \quad (11.4)$$

for normal X, Y . The reverse implication follows directly by expanding X and Y in their spectral projectors.

One can also understand the theorem as successive resolution of degeneracy. If $X|\psi\rangle = \lambda|\psi\rangle$, then

$$X(Y|\psi\rangle) = Y(X|\psi\rangle) = \lambda Y|\psi\rangle,$$

so Y preserves the λ eigenspace of X . Diagonalizing Y inside that space provides additional labels without changing the already assigned value of X .

Example 11.3 (Joint labels remove individual degeneracies). In an orthonormal basis of a three-dimensional space, let

$$X = \text{diag}(1, 1, -1), \quad Y = \text{diag}(1, -1, -1).$$

Both operators have degeneracies: X has a two-dimensional $+1$ eigenspace and Y has a two-dimensional -1 eigenspace. The joint eigenvalue pairs are nevertheless distinct:

$$(1, 1), \quad (1, -1), \quad (-1, -1).$$

Thus the joint values identify each basis direction uniquely. Only an arbitrary phase for each basis ket and a permutation of the ordered list remain. If a joint eigenspace has dimension greater than one, unitary freedom within it remains as well.

12 Unitary operators and changes of representation

12.1 Preservation of inner products

An operator U preserves all inner products if

$$\langle U\phi | U\psi \rangle = \langle \phi | \psi \rangle \quad \text{for all } |\phi\rangle, |\psi\rangle.$$

The left side is $\langle \phi | U^\dagger U | \psi \rangle$, so this condition is equivalent to $U^\dagger U = \mathbf{1}$. In finite dimension, this already implies $UU^\dagger = \mathbf{1}$: U has trivial kernel and hence is invertible, and its left inverse $U^\dagger$ must equal its inverse. Thus inner-product-preserving operators are exactly the unitary operators.

A unitary operator maps an orthonormal basis to another orthonormal basis. Conversely, given two ordered orthonormal bases $\{|i\rangle\}$ and $\{|i'\rangle\}$, define

$$U = \sum_i |i'\rangle \langle i|. \quad (12.1)$$

Then $U|i\rangle = |i'\rangle$, and

$$U^\dagger U = \sum_{i,j} |i\rangle \langle i' | j'\rangle \langle j| = \sum_i |i\rangle \langle i| = \mathbf{1}.$$

The same calculation in reverse gives $UU^\dagger = \mathbf{1}$.

12.2 Unitary conjugation of operators

Under a unitary transformation, define

$$|\psi\rangle_U = U|\psi\rangle, \quad X_U = UXU^\dagger.$$

Then

$$X_U |\psi\rangle_U = UX|\psi\rangle,$$

so the operator action is transformed consistently with the states. Inner products and matrix elements are unchanged if all objects are transformed together:

$$\langle \phi_U | \psi_U \rangle = \langle \phi | \psi \rangle, \quad \langle \phi_U | X_U | \psi_U \rangle = \langle \phi | X | \psi \rangle.$$

The transformation preserves the operator algebra:

$$(aX + bY)_U = aX_U + bY_U, \quad (12.2)$$

$$(XY)_U = X_U Y_U, \quad (12.3)$$

$$(X^\dagger)_U = (X_U)^\dagger, \quad (12.4)$$

$$[X, Y]_U = [X_U, Y_U]. \quad (12.5)$$

For example, $X_U Y_U = UXU^\dagger UYU^\dagger = UXYU^\dagger$. Consequently, self-adjointness, unitarity, and normality are preserved.

If $X|\psi\rangle = \lambda|\psi\rangle$, then

$$X_U(U|\psi\rangle) = \lambda U|\psi\rangle.$$

Thus eigenvalues and their multiplicities are also preserved. Conversely, two normal operators with the same eigenvalues and multiplicities are unitarily equivalent: choose corresponding orthonormal eigenbases and use (12.1) to map one to the other.

An operator is invariant precisely when $UXU^\dagger = X$, or equivalently when $[U, X] = 0$. A unitary transformation is always a valid change of representation at the kinematic level. To be a physical symmetry of a particular dynamics, it must also respect the specified dynamical laws. The distinction will become concrete when a Hamiltonian is introduced.

13 Components, matrix multiplication, and invariants

13.1 Columns, rows, and operator matrices

Fix an orthonormal basis $|i\rangle$. Define

$$\psi_i = \langle i | \psi \rangle, \quad X_{ij} = \langle i | X | j \rangle.$$

The ket is represented by the column (ψ_i) , while its associated bra is represented by the conjugate row (ψ_i^*) . Completeness on both sides of X gives

$$X = \sum_{i,j} X_{ij} |i\rangle \langle j|. \quad (13.1)$$

An arbitrary bra can be written $\langle \omega | = \sum_i b_i \langle i |$, where $b_i = \langle \omega | i \rangle$. If it is the adjoint of a ket with components ω_i , then $b_i = \omega_i^*$.

The components of $X|\psi\rangle$ are

$$(X\psi)_i = \langle i | X | \psi \rangle = \sum_j \langle i | X | j \rangle \langle j | \psi \rangle = \sum_j X_{ij} \psi_j.$$

Similarly, inserting the identity between two operators gives ordinary matrix multiplication:

$$(XY)_{ij} = \sum_k X_{ik} Y_{kj}. \quad (13.2)$$

The inner product becomes $\langle \phi | \psi \rangle = \sum_i \phi_i^* \psi_i$. The adjoint is the conjugate transpose:

$$(X^\dagger)_{ij} = X_{ji}^*. \quad (13.3)$$

Thus a self-adjoint operator has a Hermitian matrix in every orthonormal basis, and a unitary operator has a unitary matrix in every orthonormal basis.

13.2 Active transformations and passive basis changes

An *active* transformation changes a vector while the basis is fixed. If $|\psi\rangle \mapsto U|\psi\rangle$, its column changes as $\psi \mapsto U\psi$, using the matrix of U in that basis. An actively transformed operator has matrix $X \mapsto UXU^\dagger$.

A *passive* transformation changes the basis while the abstract objects remain fixed. Let

$$|i'\rangle = V|i\rangle$$

for a unitary V . The components in the primed basis are

$$\psi'_i = \langle i' | \psi \rangle = \langle i | V^\dagger | \psi \rangle, \quad \psi' = V^\dagger \psi, \quad (13.4)$$

$$X'_{ij} = \langle i' | X | j' \rangle = \langle i | V^\dagger X V | j \rangle, \quad X' = V^\dagger X V. \quad (13.5)$$

If the basis change is instead parametrized as $|i'\rangle = U^\dagger|i\rangle$, these become $\psi' = U\psi$ and $X' = UXU^\dagger$, the same matrix formulas as the active transformation. The interpretation is different because different objects were held fixed.

Finally, if both the states and the basis vectors are acted on by the same unitary, the components stay unchanged because $\langle Ui | U\psi \rangle = \langle i | \psi \rangle$. This explains why coordinate formulas alone do not specify whether a transformation is active or passive.

13.3 Trace and determinant

The trace and determinant of an operator are defined using any matrix representation:

$$\text{Tr } X = \sum_i X_{ii}, \quad \det X = \det(X_{ij}).$$

They do not depend on the chosen basis. For the trace, first prove cyclicity:

$$\text{Tr}(XY) = \sum_{i,j} X_{ij}Y_{ji} = \sum_{j,i} Y_{ji}X_{ij} = \text{Tr}(YX). \quad (13.6)$$

It follows that $\text{Tr}(V^\dagger X V) = \text{Tr}(X V V^\dagger) = \text{Tr } X$. For the determinant, multiplicativity gives

$$\det(V^\dagger X V) = \det(V^\dagger) \det(X) \det(V) = \det X,$$

since $\det(V^\dagger) \det(V) = \det \mathbf{1} = 1$. More generally these invariants are unchanged under any invertible similarity transformation.

Exercise 13.1 (Trace of a rank-one operator). Show that $\text{Tr}(|\psi\rangle\langle\omega|) = \langle\omega | \psi\rangle$.

Solution. Insert a complete orthonormal basis:

$$\text{Tr}(|\psi\rangle\langle\omega|) = \sum_i \langle i | \psi \rangle \langle \omega | i \rangle = \langle \omega | \left(\sum_i |i\rangle\langle i| \right) | \psi \rangle = \langle \omega | \psi \rangle.$$

In particular, a rank-one orthogonal projector has trace one, and a projector onto an r -dimensional subspace has trace r . For a normal operator, the spectral representation gives $\text{Tr } X = \sum_i \lambda_i$ and $\det X = \prod_i \lambda_i$, with eigenvalues counted with multiplicity. $\square$

Part III

The Postulates of Quantum Mechanics

14 Spin measurements in operator language

14.1 State kets and the Stern–Gerlach projectors

A two-dimensional Hilbert space describes a **qubit**. For the spin example, use the orthonormal basis $|z, +\rangle, |z, -\rangle$ and write

$$|\psi\rangle = \alpha|z, +\rangle + \beta|z, -\rangle, \quad \langle\psi | \psi\rangle = |\alpha|^2 + |\beta|^2. \quad (14.1)$$

A z -oriented Stern–Gerlach device separates the input into its two components.

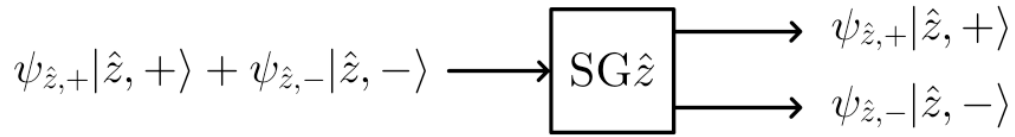


Figure 19: The abstract input ket and its two outgoing spin components.

Define the rank-one orthogonal projectors

$$\Pi_{z,+} = |z, +\rangle\langle z, +|, \quad \Pi_{z,-} = |z, -\rangle\langle z, -|. \quad (14.2)$$

Then the unnormalized branch kets are

$$|\psi_+\rangle = \Pi_{z,+}|\psi\rangle = \alpha|z, +\rangle, \quad |\psi_-\rangle = \Pi_{z,-}|\psi\rangle = \beta|z, -\rangle.$$

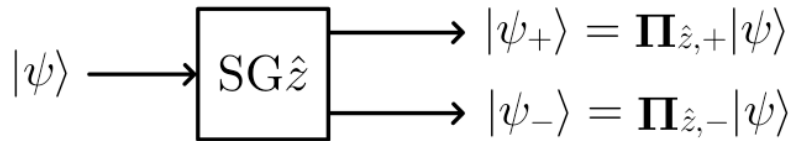


Figure 20: The same Stern–Gerlach device expressed using a complete orthogonal projector family.

For a normalized input, the squared norm of each branch is the corresponding probability. If a preceding filter has removed part of an incident beam, an unnormalized input can instead retain its probability of reaching the apparatus. The sum of its branch norms then equals that incident probability, rather than one. Conditional probabilities are obtained by dividing by $\langle\psi | \psi\rangle$.

14.2 An observable and its expectation value

Associate the numerical outcomes $\pm\hbar/2$ with the spectral projectors:

$$S_z = \frac{\hbar}{2}\Pi_{z,+} - \frac{\hbar}{2}\Pi_{z,-}. \quad (14.3)$$

This is a self-adjoint operator. Its eigenvalues are the possible measurement values, and its eigenspaces describe states with definite outcomes.

For repeated trials on identically prepared atoms, the average measured value is

$$\langle S_z \rangle_\psi = \frac{\hbar}{2} P(z, +) - \frac{\hbar}{2} P(z, -) \quad (14.4)$$

$$= \frac{\hbar}{2} \frac{|\alpha|^2 - |\beta|^2}{|\alpha|^2 + |\beta|^2} = \frac{\langle \psi | S_z | \psi \rangle}{\langle \psi | \psi \rangle}. \quad (14.5)$$

The last expression is independent of any nonzero complex rescaling of the ket.

For bookkeeping one may normalize $\langle \psi | \psi \rangle = N$, where N is the number of independently prepared repetitions. Then $\|\Pi_{z,\pm}|\psi\rangle\|^2$ is the *expected* number of outcomes in the corresponding branch. This is not an N -particle state vector; it is a convenient rescaling of the one-system ket. Actual counts fluctuate. For independent repetitions with outcome probability p , the count has mean Np and variance $Np(1-p)$, so relative fluctuations become small as the number of trials increases.

14.3 The three spin observables

The x eigenstates give projectors

$$\Pi_{x,\pm} = |x, \pm\rangle\langle x, \pm|, \quad S_x = \frac{\hbar}{2}(\Pi_{x,+} - \Pi_{x,-}).$$

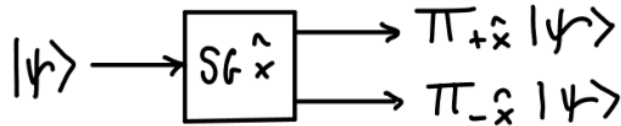


Figure 21: The original x -oriented apparatus represented by its output projectors.

To simplify the following algebra, write $|+\rangle = |z, +\rangle$ and $|-\rangle = |z, -\rangle$. Expanding the x projectors gives

$$\begin{aligned} \Pi_{x,+} &= \frac{1}{2}(|+\rangle\langle +| + |+\rangle\langle -| + |-\rangle\langle +| + |-\rangle\langle -|), \\ \Pi_{x,-} &= \frac{1}{2}(|+\rangle\langle +| - |+\rangle\langle -| - |-\rangle\langle +| + |-\rangle\langle -|). \end{aligned}$$

Taking the difference,

$$S_x = \frac{\hbar}{2}(|+\rangle\langle -| + |-\rangle\langle +|). \quad (14.6)$$

For $|y, \pm\rangle = (|+\rangle \pm i|-\rangle)/\sqrt{2}$, the bras contain the conjugate phases. Thus

$$\begin{aligned}\Pi_{y,+} &= \frac{1}{2}(|+\rangle\langle+| - i|+\rangle\langle-| + i|-\rangle\langle+| + |-\rangle\langle-|), \\ \Pi_{y,-} &= \frac{1}{2}(|+\rangle\langle+| + i|+\rangle\langle-| - i|-\rangle\langle+| + |-\rangle\langle-|),\end{aligned}$$

and

$$S_y = \frac{\hbar}{2i}(|+\rangle\langle-| - |-\rangle\langle+|). \quad (14.7)$$

Using $\langle+|+\rangle = \langle-|-\rangle = 1$ and $\langle+|-\rangle = 0$, multiply the outer products to obtain

$$[S_z, S_x] = \frac{\hbar^2}{2}(|+\rangle\langle-| - |-\rangle\langle+|) = i\hbar S_y, \quad (14.8)$$

$$[S_x, S_y] = i\hbar S_z, \quad [S_y, S_z] = i\hbar S_x. \quad (14.9)$$

These relations are the spin angular-momentum algebra in three dimensions. Since the observables do not commute, they have no common eigenbasis. Their projective measurements are incompatible, and changing the order of a sequence can change its statistics.

14.4 Why distinct measurement states must be orthogonal

Suppose a device separates two normalized states $|a\rangle, |b\rangle$ with amplitudes α, β into distinct, perfectly distinguishable outcomes. Probability conservation would require

$$\|\alpha|a\rangle + \beta|b\rangle\|^2 = |\alpha|^2 + |\beta|^2$$

for every pair of amplitudes. Expanding the left side gives

$$|\alpha|^2 + |\beta|^2 + 2\operatorname{Re}(\alpha^*\beta\langle a|b\rangle).$$

For this to equal the sum of the outgoing probabilities for all phases and magnitudes, $\langle a|b\rangle$ must vanish. This yields the **orthogonality principle**: alternatives that can be distinguished with certainty correspond to orthogonal states, or more generally to orthogonal subspaces.

The principle refers to perfect distinguishability, not to the mere existence of different preparation procedures. Nonorthogonal pure states are physically different and can give different statistics, but a single measurement cannot distinguish them without error in every trial.

14.5 Pauli matrices and their algebra

In the ordered z basis,

$$S_j = \frac{\hbar}{2}\sigma_j, \quad \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (14.10)$$

Direct multiplication gives

$$\sigma_x^2 = \sigma_y^2 = \sigma_z^2 = \mathbf{1}, \quad (14.11)$$

$$\sigma_x \sigma_y = -\sigma_y \sigma_x = i\sigma_z, \quad (14.12)$$

$$\sigma_y \sigma_z = -\sigma_z \sigma_y = i\sigma_x, \quad (14.13)$$

$$\sigma_z \sigma_x = -\sigma_x \sigma_z = i\sigma_y. \quad (14.14)$$

These can be written compactly as

$$\sigma_i \sigma_j = \delta_{ij} \mathbf{1} + i \sum_k \epsilon_{ijk} \sigma_k, \quad (14.15)$$

where ϵ_{ijk} is the completely antisymmetric symbol with $\epsilon_{xyz} = +1$. Subtracting the formula with i and j reversed reproduces (14.9); adding gives

$$\{\sigma_i, \sigma_j\} = 2\delta_{ij} \mathbf{1}.$$

In particular,

$$S_x^2 + S_y^2 + S_z^2 = \frac{3\hbar^2}{4} \mathbf{1}.$$

Every Hermitian 2×2 matrix is a real linear combination of $\mathbf{1}$ and the three Pauli matrices. Indeed, writing its general form as

$$A = \begin{pmatrix} a & c - id \\ c + id & b \end{pmatrix}, \quad a, b, c, d \in \mathbb{R},$$

one finds

$$A = \frac{a+b}{2} \mathbf{1} + c\sigma_x + d\sigma_y + \frac{a-b}{2} \sigma_z. \quad (14.16)$$

The coefficients are real, but the matrix need not have only real entries.

Exercise 14.1 (The spin matrices in the x basis). Find the matrix representations of S_x, S_y, S_z in the ordered basis $|x, +\rangle, |x, -\rangle$.

Solution. The matrix whose columns are the new basis kets expressed in the old basis is H from (6.6). Therefore the new matrices are $H^\dagger S_j H$. Since $H = H^\dagger$, direct multiplication gives

$$H\sigma_x H = \sigma_z, \quad H\sigma_y H = -\sigma_y, \quad H\sigma_z H = \sigma_x.$$

Consequently,

$$[S_x]_x = \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad (14.17)$$

$$[S_y]_x = \frac{\hbar}{2} \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}, \quad (14.18)$$

$$[S_z]_x = \frac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \quad (14.19)$$

The sign change in the matrix representing S_y is required by the chosen phases of the x basis. The abstract operators and their commutation relations have not changed. $\square$

14.6 Relative phases in a Stern–Gerlach recombiner

Exercise 14.2 (A recombiner with unequal branch phases). Let the two internal x branches acquire phases $e^{i\theta_+}$ and $e^{i\theta_-}$. Find the output operator and the probabilities of a final S_z measurement.

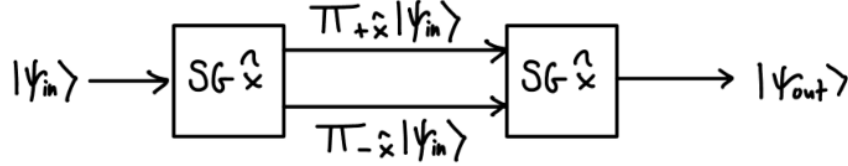


Figure 22: The source's modified Stern–Gerlach device for studying the relative phase between its two internal branches.

Solution. The coherent output is the sum of the two branch kets:

$$|\psi_{\text{out}}\rangle = e^{i\theta_+}\Pi_{x,+}|\psi_{\text{in}}\rangle + e^{i\theta_-}\Pi_{x,-}|\psi_{\text{in}}\rangle = U|\psi_{\text{in}}\rangle,$$

where

$$U = e^{i\theta_+}\Pi_{x,+} + e^{i\theta_-}\Pi_{x,-}. \quad (14.20)$$

The projectors are orthogonal and complete, so

$$U^\dagger U = \Pi_{x,+} + \Pi_{x,-} = \mathbf{1}, \quad UU^\dagger = \mathbf{1}.$$

Thus U is unitary and hence normal. It is self-adjoint only when both eigenphases are real numbers, $e^{i\theta_\pm} = \pm 1$.

Put $\bar{\theta} = (\theta_+ + \theta_-)/2$ and $\delta = (\theta_+ - \theta_-)/2$. Since $\Pi_{x,\pm} = (\mathbf{1} \pm \sigma_x)/2$,

$$U = e^{i\bar{\theta}} (\cos \delta \mathbf{1} + i \sin \delta \sigma_x). \quad (14.21)$$

For input $|z, +\rangle$,

$$|\psi_{\text{out}}\rangle = e^{i\bar{\theta}} (\cos \delta |z, +\rangle + i \sin \delta |z, -\rangle),$$

and therefore

$$P(z, +) = \cos^2 \frac{\theta_+ - \theta_-}{2}, \quad P(z, -) = \sin^2 \frac{\theta_+ - \theta_-}{2}. \quad (14.22)$$

The common phase cancels. Equal branch phases restore $z, +$ with certainty; a relative phase of π converts it into $z, -$.

For a general normalized input $\alpha|z, +\rangle + \beta|z, -\rangle$, the probabilities are

$$P(z, +) = |\alpha \cos \delta + i\beta \sin \delta|^2, \quad (14.23)$$

$$P(z, -) = |\beta \cos \delta + i\alpha \sin \delta|^2. \quad (14.24)$$

Their sum is one by unitarity. The general operator expression is $P(z, \pm) = \langle \psi | U^\dagger \Pi_{z,\pm} U | \psi \rangle$. $\square$

15 Postulate 1: pure states

Postulate 1. A quantum system has an associated complex Hilbert space $\mathcal{H}$. A pure state is represented by a nonzero ket $|\psi\rangle \in \mathcal{H}$, with kets related by multiplication by any nonzero complex scalar representing the same physical pure state.

The equivalence class is a **ray**. One usually chooses a unit-norm representative, leaving the freedom $|\psi\rangle \mapsto e^{i\phi}|\psi\rangle$.

The vector-space structure is motivated by interference: coherent routes to the same outcome add amplitudes and, more generally, state vectors. Scaling an unnormalized vector changes its statistical weight; multiplying it by a phase changes its phase relative to another coherently combined component. However, the postulate is a physical assumption about quantum systems, not a consequence of abstract vector-space algebra alone.

Global phase does not eliminate relative phase. The rays represented by $|\psi\rangle$ and $e^{i\phi}|\psi\rangle$ are the same, but in general

$$a|\psi\rangle + b|\chi\rangle \quad \text{and} \quad a|\psi\rangle + e^{i\phi}b|\chi\rangle$$

are different rays. For this reason one cannot add rays without first specifying representative kets and relative phases. A coherent superposition is also different from randomly preparing one state or another. The latter is a statistical mixture and does not contain the same controlled phase information.

A filter that transmits a fraction p can be described by an unnormalized branch vector whose norm is reduced by $\sqrt{p}$. Conditioning on transmission renormalizes that vector. This bookkeeping should not be confused with replacing an unknown statistical mixture by a pure-state superposition.

15.1 Why use a complex Hilbert space?

The spin experiments showed why a two-dimensional description using only real amplitudes is inadequate. There is nevertheless a mathematical way to express a complex N -dimensional vector space as a real $2N$ -dimensional one: separate every component into real and imaginary parts. Multiplication by i is then represented by a real linear map J satisfying

$$J^2 = -\mathbf{1}.$$

Explicitly,

$$z = a + ib \longleftrightarrow \begin{pmatrix} a \\ b \end{pmatrix}, \quad A + iB \longleftrightarrow \begin{pmatrix} A & -B \\ B & A \end{pmatrix}, \quad J = \begin{pmatrix} 0 & -\mathbf{1} \\ \mathbf{1} & 0 \end{pmatrix}.$$

Here a, b are real component columns and A, B are real matrices. A complex-linear operator becomes a real operator that commutes with J . The complex inner product can be reconstructed from the real inner product and this additional structure.

Thus a doubled real description *with the appropriate complex structure and restrictions* can encode the same single-system mathematics. Simply replacing $\mathbb{C}^N$ by $\mathbb{R}^N$,

or allowing arbitrary real operations on $\mathbb{R}^{2N}$ while discarding J , is a different proposal. Composite systems introduce further structure as well. The complex Hilbert-space formulation packages the empirically successful rules directly and is the formulation used throughout these notes.

Postulate 1 alone does not tell us which Hilbert space describes a given physical system. Spin-1/2 uses $\mathbb{C}^2$; a particle moving in space requires an infinite-dimensional space. Identifying the state space and the operators that correspond to measured quantities is part of constructing the physical model.

15.2 Orthogonal alternatives in a decaying system

Consider an isolated ^{210}Po atom inside a sealed box. It can decay to ^{206}Pb plus an alpha particle. The half-life in the example is approximately 138 days, so the mean lifetime in the exponential approximation is

$$\tau = \frac{138 \text{ days}}{\ln 2} \simeq 199 \text{ days}.$$

The alternatives “undecayed” and “decayed” can be distinguished, and therefore correspond to orthogonal subspaces of the Hilbert space of the box. At a given time the state can be decomposed as

$$|\Psi(t)\rangle = e^{-t/(2\tau)} e^{i\phi_u(t)} |u(t)\rangle + \sqrt{1 - e^{-t/\tau}} e^{i\phi_d(t)} |d(t)\rangle, \quad (15.1)$$

where $|u(t)\rangle$ and $|d(t)\rangle$ are normalized vectors in the respective orthogonal subspaces. Its norm is one, and its two sector probabilities are $e^{-t/\tau}$ and $1 - e^{-t/\tau}$.

The decayed sector includes the momenta, positions, and other degrees of freedom of the daughter products. It is not literally a single fixed second level of a closed two-dimensional system. The exponential law is an approximation appropriate to a range of decay times; the notation in (15.1) displays the two alternatives without suppressing their internal degrees of freedom.

Isolation also means more than concealment from the experimenter. A system inside a box may still interact with uncontrolled surroundings. If it is genuinely isolated, checking whether it has decayed is a measurement on that quantum state. We now formulate the corresponding rules.

16 From interference to linear measurement maps

16.1 Outcomes need not specify complete final states

Consider an apparatus with outcomes labelled $i = 1, \dots, n$.

A measurement outcome does not necessarily fix every degree of freedom of the system. Measuring an atom’s spin, for example, need not determine its other internal degrees of freedom. The output of a given branch must therefore be described by a ket, rather than by a single scalar amplitude alone.

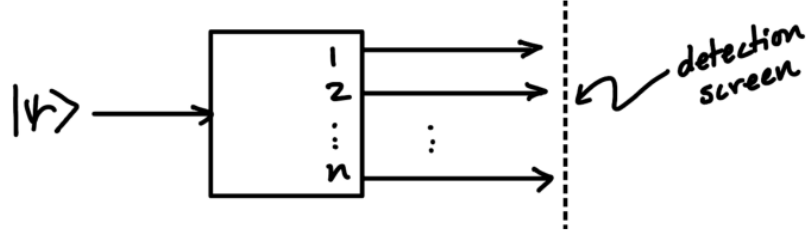


Figure 23: A generalized Stern–Gerlach-type apparatus with several possible outcomes.

The generalized interference principle states that if the same outcome can be reached coherently by several alternatives, the corresponding output kets add:

$$|\psi\rangle = \sum_a |\psi_a\rangle, \quad P_{\text{tot}} = \langle\psi | \psi\rangle. \quad (16.1)$$

The individual route weights are $P_a = \langle\psi_a | \psi_a\rangle$. Expanding the total norm produces cross terms $\langle\psi_a | \psi_b\rangle$; they vanish when the alternatives carry orthogonal records.

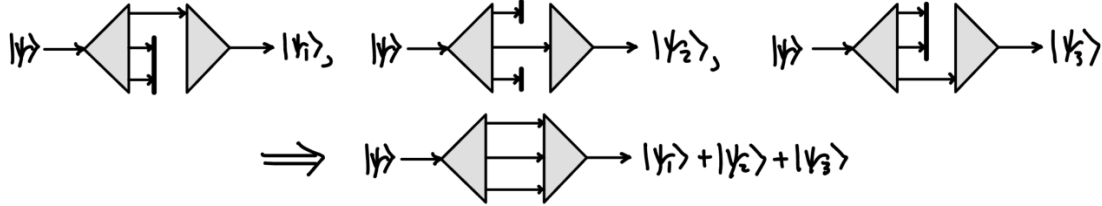


Figure 24: The original diagram for adding state kets associated with coherent routes to the same outcome.

The distinction between an *outcome* and a fully specified *final state* matters. A detector may supply only one coarse label while leaving a nontrivial output state within the corresponding sector.

For repeated identical preparations, one may choose $\langle\psi | \psi\rangle = N$. Let $|\psi_i\rangle$ be the outgoing ket with $\langle\psi_i | \psi_i\rangle = N_i$, where N_i is the expected count in branch i .

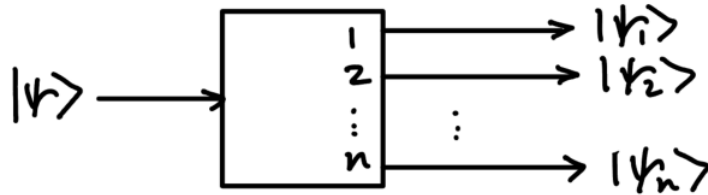


Figure 25: Output branch kets retain both their internal state and their statistical weight.

We seek maps f_i such that

$$|\psi_i\rangle = f_i(|\psi\rangle).$$

The following argument shows how the interference rules restrict these maps. It assumes that the degrees of freedom needed to describe coherent branch outputs are included.

16.2 Real rescaling and additivity

If the same experiment is repeated λ^2 times as often, with $\lambda > 0$, all expected counts scale by λ^2 . This is represented by scaling every relevant ket by λ , so

$$f_i(\lambda|\psi\rangle) = \lambda f_i(|\psi\rangle), \quad \lambda > 0. \quad (16.2)$$

No new relative phase can depend on this bookkeeping choice. If it did, recombining branches would change observable probabilities merely because the number of repetitions had changed.

Next, feed an interferometer into the device. With only path A open, let the input be $|\psi_A\rangle$ and the branch- i output be $f_i(|\psi_A\rangle)$; define the corresponding quantities for path B similarly. With both paths open coherently, the input is their sum and the output is the sum of the two branch outputs.

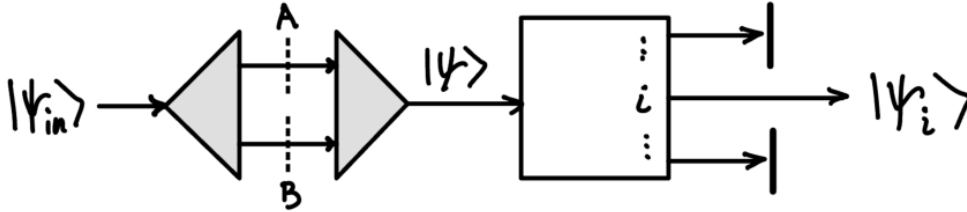


Figure 26: Directing coherent incoming alternatives into the measurement device establishes additivity of each branch map.

Thus

$$f_i(|\phi\rangle + |\chi\rangle) = f_i(|\phi\rangle) + f_i(|\chi\rangle). \quad (16.3)$$

This implies $f_i(0) = 0$ and $f_i(-|\psi\rangle) = -f_i(|\psi\rangle)$. Together with positive homogeneity, it gives homogeneity for all real scalars.

The input's overall phase is unobservable. Therefore its output ray and norm must remain unchanged when $|\psi\rangle \mapsto e^{i\phi}|\psi\rangle$:

$$f_i(e^{i\phi}|\psi\rangle) = e^{i\phi'} f_i(|\psi\rangle), \quad (16.4)$$

where the relation between ϕ' and ϕ remains to be determined.

16.3 Why these conditions give linearity or antilinearity

Exercise 16.1. Show that additivity, real homogeneity, and phase covariance imply that a nonzero map f is either complex-linear or complex-antilinear.

Solution. For any vector with $f(|\psi\rangle) \neq 0$, phase covariance at $\phi = \pi/2$ gives

$$f(i|\psi\rangle) = c f(|\psi\rangle), \quad |c| = 1.$$

On the one hand, additivity gives

$$f((1+i)|\psi\rangle) = (1+c)f(|\psi\rangle).$$

On the other hand, $1+i = \sqrt{2}e^{i\pi/4}$, so real homogeneity and phase covariance imply that this output has norm $\sqrt{2}\|f(|\psi\rangle)\|$. Hence

$$|1+c|^2 = 2.$$

Since $|c| = 1$, this says $2 + 2\operatorname{Re} c = 2$, and therefore $c = +i$ or $c = -i$.

The sign cannot vary from one nonzero output to another. Suppose $f(i|\psi\rangle) = ia$ with $a = f(|\psi\rangle) \neq 0$, but $f(i|\chi\rangle) = -ib$ with $b = f(|\chi\rangle) \neq 0$. Additivity gives

$$f(|\psi\rangle + |\chi\rangle) = a + b, \quad f(i(|\psi\rangle + |\chi\rangle)) = i(a - b).$$

If $a + b \neq 0$, the second expression must equal either $+i(a + b)$ or $-i(a + b)$, forcing $b = 0$ or $a = 0$, a contradiction. If $a + b = 0$, phase covariance requires the second expression to vanish, but it equals $2ia \neq 0$. Thus one sign applies to all vectors with nonzero output. A zero output remains zero after multiplication by a phase.

For $\lambda = s + it$, real homogeneity now gives

$$f(\lambda|\psi\rangle) = sf(|\psi\rangle) + tf(i|\psi\rangle) = \begin{cases} \lambda f(|\psi\rangle), & \text{with the plus sign,} \\ \lambda^* f(|\psi\rangle), & \text{with the minus sign.} \end{cases}$$

Together with additivity, these are complex linearity and complex antilinearity, respectively. Equivalently, the phase transforms as $e^{i\phi} \mapsto e^{\pm i\phi}$. $\square$

16.4 Why a coherent physical branch cannot be antilinear

Now compare a path through the apparatus with a coherent bypass path.

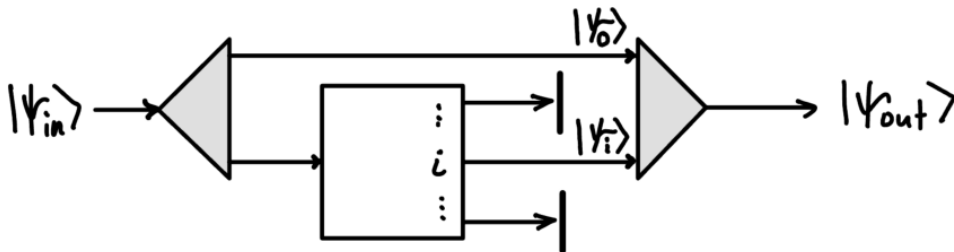


Figure 27: Interfering a branch through the device with a bypass branch tests the transformation of an overall input phase.

If the device branch were antilinear, changing the entire input ket by the physically irrelevant phase $e^{i\phi}$ would change that branch by $e^{-i\phi}$. The bypass branch, which undergoes an ordinary linear propagation, would change by $e^{i\phi}$. Their *relative* phase would therefore change by 2ϕ , affecting the recombined probabilities.

This would make the arbitrary overall phase of the incoming state observable, contrary to the pure-state postulate. The coherent branch maps are consequently complex-linear:

$$\boxed{|\psi_i\rangle = O_i|\psi\rangle}, \quad (16.5)$$

where each operator O_i is independent of the input state.

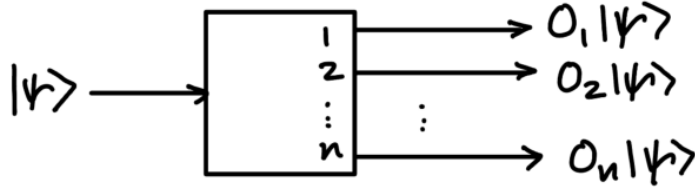


Figure 28: Each outcome branch is represented by a linear operator acting on the input ket.

This does not rule out using antilinear maps to represent certain abstract symmetries, such as time reversal. It rules out treating an arbitrary antilinear map as a coherent apparatus branch acting on unknown inputs in the same way as a physical linear operation.

17 Repeatable measurements and the Born rule

17.1 Probability conservation and repeatability

For the branch maps in (16.5), the branch weight is

$$N_i = \|O_i|\psi\rangle\|^2 = \langle\psi|O_i^\dagger O_i|\psi\rangle. \quad (17.1)$$

If every trial produces exactly one of the listed outcomes, then $\sum_i N_i = \langle\psi|\psi\rangle$ for every input. Equality of these quadratic forms is equivalent to

$$\sum_i O_i^\dagger O_i = 1. \quad (17.2)$$

This is the probability-conservation condition.

A measurement is **repeatable** if an immediate second use of the same device yields the same outcome with certainty. For a first outcome j , the branch with second outcome i has ket $O_i O_j |\psi\rangle$. If $i \neq j$, its probability must vanish for every input, so

$$O_i O_j = 0 \quad (i \neq j). \quad (17.3)$$

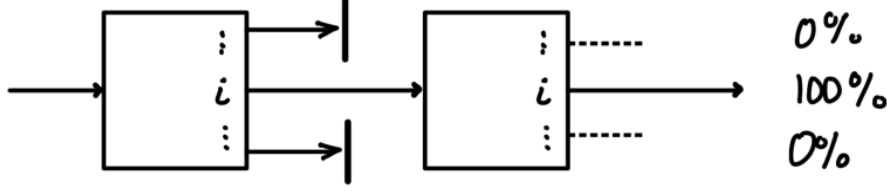


Figure 29: Repeatability: after outcome i , the same apparatus gives outcome i again with certainty.

Conversely, these operator conditions forbid every different second outcome; probability conservation then puts all the probability in the repeated outcome.

The next result derives the full structure of these repeatable measurements in finite dimension, including their possible residual action inside a degenerate outcome sector.

17.2 Deriving the projector form step by step

Theorem 17.1 (Finite-dimensional repeatable measurement maps). *Let $O_1, \dots, O_n$ act on the same finite-dimensional Hilbert space and satisfy (17.2) and (17.3). Then*

$$O_i = U\Pi_i, \quad (17.4)$$

where U is unitary, the Π_i form a complete orthogonal projector family, and $[U, \Pi_i] = 0$ for every i .

Proof. Step 1: the output ranges are mutually orthogonal. Let $|u\rangle = O_i|\alpha\rangle$ and $|v\rangle = O_j|\beta\rangle$ with $i \neq j$. Insert (17.2):

$$\langle u | v \rangle = \sum_k \langle O_k u | O_k v \rangle.$$

Repeatability says $O_k|u\rangle = 0$ unless $k = i$, and $O_k|v\rangle = 0$ unless $k = j$. Since $i \neq j$, every term in the sum is zero. Thus $\text{ran } O_i \perp \text{ran } O_j$, or equivalently

$$O_i^\dagger O_j = 0 \quad (i \neq j). \quad (17.5)$$

Step 2: the sum of the branch maps is unitary. Define

$$U = \sum_i O_i.$$

Using (17.5),

$$U^\dagger U = \sum_{i,j} O_i^\dagger O_j = \sum_i O_i^\dagger O_i = \mathbf{1}.$$

In finite dimension this makes U invertible and $U^{-1} = U^\dagger$, so also $UU^\dagger = \mathbf{1}$.

Step 3: construct self-adjoint, complete operators. Set

$$\Pi_i = U^\dagger O_i. \quad (17.6)$$

Only one term survives in $U^\dagger O_i = \sum_j O_j^\dagger O_i$, giving

$$\Pi_i = O_i^\dagger O_i. \quad (17.7)$$

Thus Π_i is positive and self-adjoint. Their sum is

$$\sum_i \Pi_i = U^\dagger \sum_i O_i = U^\dagger U = \mathbf{1}.$$

Step 4: show that U commutes with each branch map. From (17.3),

$$U O_i = \sum_j O_j O_i = O_i^2, \quad O_i U = \sum_j O_i O_j = O_i^2.$$

Hence $[U, O_i] = 0$. Since U is invertible, $U^\dagger = U^{-1}$ also commutes with O_i .

Step 5: prove orthogonality and idempotence of the Π_i . For $i \neq j$,

$$\Pi_i \Pi_j = U^\dagger O_i U^\dagger O_j = (U^\dagger)^2 O_i O_j = 0.$$

Using completeness,

$$\Pi_i = \Pi_i \sum_j \Pi_j = \Pi_i^2.$$

Together with self-adjointness this proves that the Π_i are orthogonal projectors. Finally,

$$U \Pi_i = O_i, \quad \Pi_i U = U^\dagger O_i U = O_i,$$

so U commutes with every Π_i and (17.4) follows. $\square$

The converse is immediate but useful to check. If $O_i = U \Pi_i$ with U unitary and commuting with a complete orthogonal projector family, then

$$\sum_i O_i^\dagger O_i = \sum_i \Pi_i = \mathbf{1}, \quad O_i O_j = U^2 \Pi_i \Pi_j = 0 \quad (i \neq j).$$

Thus the derived form is both necessary and sufficient under the stated finite-dimensional assumptions.

17.3 Probabilities and the state after an outcome

For a nonzero input ket, (17.7) gives

$$p_i = \frac{\langle \psi | \Pi_i | \psi \rangle}{\langle \psi | \psi \rangle} = \frac{\| \Pi_i | \psi \rangle \|^2}{\| | \psi \rangle \|^2}. \quad (17.8)$$

This is the Born rule for a projective measurement. The probabilities are nonnegative and sum to one. The unnormalized branch is $U \Pi_i | \psi \rangle$, while the normalized conditional state, for a nonzero-probability outcome, is

$$| \psi_i^{\text{cond}} \rangle = \frac{U \Pi_i | \psi \rangle}{\sqrt{\langle \psi | \Pi_i | \psi \rangle}}. \quad (17.9)$$

An outcome with zero probability has no conditional state assigned by this formula.

Repeated applications yielding the same outcome have unnormalized state

$$O_i^m|\psi\rangle = (U\Pi_i)^m|\psi\rangle = U^m\Pi_i|\psi\rangle.$$

The outcome remains certain after the first measurement, but the state may continue evolving inside that outcome subspace.

If Π_i has rank one, say $\Pi_i = |\chi\rangle\langle\chi|$, commutation implies that $U|\chi\rangle$ lies in the same one-dimensional subspace. Thus

$$U|\chi\rangle = e^{i\theta_i}|\chi\rangle,$$

and the residual action is only an unobservable phase on that branch. For a degenerate outcome, U may rotate vectors within its subspace. Repeatability of the outcome therefore need not mean that every detail of the conditional state is unchanged.

An **ideal projective measurement** takes $U = \mathbf{1}$, yielding the projection or Lüders update. A block-dependent phase would give the same conditional rays. Nontrivial unitary rotations within a degenerate block describe additional disturbance beyond the ideal update.

18 Postulate 2: observables and ideal measurements

Assign distinct real numerical values a_i to a complete projector family and define

$$A = \sum_i a_i \Pi_i. \quad (18.1)$$

The operator is self-adjoint, and the Π_i are precisely its eigenspace projectors. Conversely, the spectral theorem assigns such a family to every self-adjoint operator. This leads to the standard ideal-measurement postulate.

Postulate 2. An observable is represented by a self-adjoint operator A . Its possible ideal-measurement outcomes are its eigenvalues. If Π_i projects onto the eigenspace of eigenvalue a_i , the probability of that outcome is

$$p_i = \frac{\langle\psi|\Pi_i|\psi\rangle}{\langle\psi|\psi\rangle}.$$

Conditioned on an outcome of nonzero probability, an ideal measurement gives the normalized state

$$|\psi_i\rangle = \frac{\Pi_i|\psi\rangle}{\sqrt{\langle\psi|\Pi_i|\psi\rangle}}.$$

For a nondegenerate eigenvalue with normalized eigenket $|a_i\rangle$, this reduces to

$$p_i = \frac{|\langle a_i | \psi \rangle|^2}{\langle \psi | \psi \rangle}.$$

For a degenerate eigenspace with orthonormal basis $|a_i, \alpha\rangle$,

$$p_i = \frac{\sum_{\alpha} |\langle a_i, \alpha | \psi \rangle|^2}{\langle \psi | \psi \rangle}.$$

The basis chosen within the eigenspace does not affect this probability.

The choice to assign real numbers to outcomes is conventional for ordinary observables. A normal operator can also package complex-valued outcome labels:

$$N = \sum_i z_i \Pi_i = A + iB, \quad A = \frac{N + N^\dagger}{2}, \quad B = \frac{N - N^\dagger}{2i}.$$

The operators A and B are self-adjoint. Normality of N is equivalent to $[A, B] = 0$, since

$$NN^\dagger - N^\dagger N = -2i[A, B].$$

Thus measuring the complex label z_i amounts to a compatible joint measurement of its real and imaginary parts. It introduces no new probability rule: the same common spectral projectors determine the outcomes.

18.1 Degeneracy and what an apparatus actually resolves

An ideal measurement of A preserves relative amplitudes within a degenerate eigenspace. For example, if

$$|\psi\rangle = c_1|a, 1\rangle + c_2|a, 2\rangle + d|b\rangle, \quad a \neq b,$$

then outcome a gives a conditional state proportional to

$$c_1|a, 1\rangle + c_2|a, 2\rangle.$$

It does not select one of the two a basis vectors. An apparatus that also measures a second observable resolving these basis vectors performs a finer measurement. Even if the experimenter later forgets that extra result, its disturbance need not be the same as the ideal coarse measurement of A alone.

The expectation of an observable is

$$\langle A \rangle_\psi = \sum_i a_i p_i = \frac{\langle \psi | A | \psi \rangle}{\langle \psi | \psi \rangle}. \quad (18.2)$$

Its higher moments follow from the same spectral expansion:

$$\langle A^k \rangle_\psi = \sum_i a_i^k p_i, \quad (\Delta A)^2 = \langle A^2 \rangle_\psi - \langle A \rangle_\psi^2.$$

A state in one eigenspace has zero variance for that observable. This does not imply zero variance for every observable.

18.2 Measurements without repeatability

The derivation separates two assumptions: probability conservation and repeatability. If repeatability is dropped, linear branch operators M_i still obey

$$\sum_i M_i^\dagger M_i = \mathbf{1},$$

but $E_i = M_i^\dagger M_i$ need not be projectors. They are positive operators satisfying

$$E_i \geq 0, \quad \sum_i E_i = \mathbf{1}, \quad p_i = \frac{\langle \psi | E_i | \psi \rangle}{\langle \psi | \psi \rangle}. \quad (18.3)$$

Such a collection is a **positive-operator-valued measure**, or POVM, for a finite outcome set. The conditional branch state is proportional to $M_i|\psi\rangle$. Its probability operator E_i alone does not completely specify the state disturbance.

If a coarse outcome leaves several unobserved environmental alternatives, more than one branch operator can be associated with that outcome. The density-operator formalism handles the resulting mixed conditional states. The projective postulate describes ideal measurements of observables; the more general measurement language describes additional apparatuses without changing the probability rule based on squared norms.

19 Postulate 3: closed-system transformations

An experiment can manipulate a quantum system without producing a measurement outcome. Examples include controlled magnetic fields acting on a spin or a coherent apparatus with all paths recombined. The same interference argument makes the transformation linear:

$$|\psi_{\text{out}}\rangle = U|\psi_{\text{in}}\rangle.$$

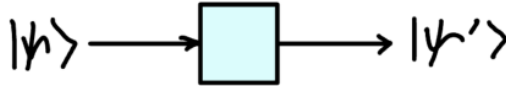


Figure 30: The original symbol for a process that manipulates a system without an outcome selection.

If the process is deterministic and preserves the norm for every input, then

$$\langle \psi | U^\dagger U | \psi \rangle = \langle \psi | \psi \rangle \quad \text{for all } |\psi\rangle,$$

so $U^\dagger U = \mathbf{1}$. In the finite-dimensional setting this implies unitarity.

Postulate 3. The evolution or reversible transformation of a closed quantum system is represented by a unitary operator. It maps a state ket to $U|\psi\rangle$ and preserves all inner products and total probability.

This postulate does not say that every process that goes unobserved is unitary on the chosen subsystem. A system may interact with an environment and become entangled with it, even when no human observer looks at either one. The joint closed system may evolve unitarily while the reduced subsystem requires a density operator and a more general transformation.

One can often enlarge the Hilbert space to include the relevant apparatus or environment and recover a closed-system description. Alternatively, one can retain only the subsystem and explicitly account for its interaction with the discarded degrees of freedom. A prescribed classical external field is an approximation in which the source of that field is sufficiently stable that its quantum correlations with the system can be neglected at the desired accuracy.

19.1 Finite and infinite dimensions

In infinite dimension, $U^\dagger U = \mathbf{1}$ alone makes U an isometry, but need not make it surjective. For example, on an orthonormal basis $|0\rangle, |1\rangle, \dots$, define the unilateral shift

$$V|n\rangle = |n+1\rangle.$$

It satisfies

$$V^\dagger V = \mathbf{1}, \quad VV^\dagger = \mathbf{1} - |0\rangle\langle 0|.$$

It preserves norms but has no preimage for $|0\rangle$. Reversible closed-system evolution requires the full unitary condition, including surjectivity. This is automatic for the finite-dimensional systems developed in these chapters, but must be kept separate from mere norm preservation in the infinite-dimensional theory.

19.2 Why an arbitrary component norm will not do

The normalization rule must be compatible with the coherent changes of basis observed in the spin experiments. This already rules out replacing the Hilbert norm by an arbitrary familiar component norm. For example, consider

$$\|\psi\|_p = (|\psi_1|^p + |\psi_2|^p)^{1/p}, \quad 1 \leq p < \infty.$$

The state $(1, 0)^\top$ has norm one for every p , but the physically required matrix H of (6.6) maps it to $(1, 1)^\top/\sqrt{2}$, whose norm is

$$\left\| H \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\|_p = (2 \cdot 2^{-p/2})^{1/p} = 2^{1/p-1/2}.$$

This equals one only for $p = 2$. For the maximum-component norm, the output norm is $1/\sqrt{2}$, so that choice also fails. This test does not classify every conceivable norm, but it shows why merely changing the normalization convention while retaining the observed coherent transformations is not consistent. The Hilbert-space norm and the amplitude-addition rule work together.

The three postulates specify the state space, measurement probabilities and updates, and closed-system transformations. They apply well beyond spin systems. To obtain a dynamical theory, the next step is to characterize the unitary family associated with time evolution and identify its generator, the Hamiltonian.

Part IV

Time Evolution and Symmetries

Chapter 4

20 Unitary time evolution and the Hamiltonian

20.1 Closed processes and the propagator

A closed quantum process changes a state by a unitary transformation. The system may be driven by prescribed external fields; “closed” means that its evolution is described without discarding entanglement with an unobserved environment. Let $|\psi; t_0\rangle$ be its state at an initial time. Define the time-evolution operator, or propagator, by

$$|\psi; t\rangle = U(t, t_0)|\psi; t_0\rangle. \quad (20.1)$$

The operator must work for every initial state, not only for one chosen ket. Its basic properties are

$$U(t, t) = \mathbf{1}, \quad (20.2)$$

$$U(t_3, t_2)U(t_2, t_1) = U(t_3, t_1), \quad (20.3)$$

$$U(t_2, t_1)^\dagger = U(t_2, t_1)^{-1} = U(t_1, t_2). \quad (20.4)$$

The second relation says that evolving in stages gives the same result as evolving directly. In particular, evolving forward and then backward returns the original state. Unitarity preserves every transition amplitude between states that undergo the same evolution:

$$\langle\phi; t | \psi; t\rangle = \langle\phi; t_0|U^\dagger U|\psi; t_0\rangle = \langle\phi; t_0 | \psi; t_0\rangle.$$

For an explicitly driven system, the propagator generally depends on both times separately, rather than only on their difference.

20.2 Defining and identifying the generator

Assume the evolution is differentiable on the states under consideration. Define

$$H(t) = i\hbar \frac{\partial U(t, t_0)}{\partial t} U(t, t_0)^\dagger. \quad (20.5)$$

This operator has units of energy. Its role as the generator of time translations follows from the definition; its physical form is found by specifying the interactions and comparing with experiment or the appropriate classical limit.

Proposition 20.1 (Properties of the Hamiltonian). *The generator in (20.5) is self-adjoint and is independent of the reference time t_0 .*

Proof. Differentiate $UU^\dagger = \mathbf{1}$:

$$\dot{U}U^\dagger + U\dot{U}^\dagger = 0.$$

Taking the adjoint of the definition therefore gives

$$H^\dagger = -i\hbar U \dot{U}^\dagger = i\hbar \dot{U} U^\dagger = H.$$

For reference-time independence, choose another initial time s and use $U(t, t_0) = U(t, s)U(s, t_0)$. The second factor is independent of t , so

$$\begin{aligned} \partial_t U(t, t_0) U(t, t_0)^\dagger &= \partial_t U(t, s) U(s, t_0) U(s, t_0)^\dagger U(t, s)^\dagger \\ &= \partial_t U(t, s) U(t, s)^\dagger. \end{aligned}$$

The result depends only on the time at which the generator acts. In infinite-dimensional problems, these statements are interpreted on appropriate operator domains; unitarity and differentiability provide the self-adjoint generator under the usual evolution assumptions. $\square$

Multiplying (20.5) on the right by U gives

$$i\hbar \partial_t U(t, t_0) = H(t)U(t, t_0), \quad U(t_0, t_0) = \mathbf{1}. \quad (20.6)$$

Apply this equation to the initial ket, which is fixed when t varies:

$$\boxed{i\hbar \frac{\partial}{\partial t} |\psi; t\rangle = H(t) |\psi; t\rangle.} \quad (20.7)$$

This is the abstract Schrödinger equation. It applies to spins, finite-state models, and spatial wave functions alike. The differential wave equation in position space is a particular representation of it.

Conversely, a self-adjoint Hamiltonian preserves normalization:

$$\frac{d}{dt} \langle \psi | \psi \rangle = \frac{i}{\hbar} \langle \psi | H | \psi \rangle - \frac{i}{\hbar} \langle \psi | H | \psi \rangle = 0.$$

The same calculation preserves inner products of any two solutions. Thus the generator equation and unitary dynamics are mutually consistent.

20.3 Time-independent Hamiltonians and stationary states

Suppose H is time independent and has an orthonormal energy eigenbasis:

$$H|n\rangle = E_n|n\rangle, \quad H = \sum_n E_n |n\rangle \langle n|.$$

Write $|\psi; t\rangle = \sum_n c_n(t) |n\rangle$. Because the basis is fixed,

$$\sum_n i\hbar \dot{c}_n |n\rangle = \sum_n E_n c_n |n\rangle.$$

Taking the inner product with $\langle m|$ isolates one scalar equation,

$$i\hbar \dot{c}_m = E_m c_m, \quad c_m(t) = e^{-iE_m(t-t_0)/\hbar} c_m(t_0).$$

Therefore

$$|\psi; t\rangle = \sum_n e^{-iE_n(t-t_0)/\hbar} |n\rangle \langle n | \psi; t_0\rangle, \quad (20.8)$$

$$U(t, t_0) = \sum_n e^{-iE_n(t-t_0)/\hbar} |n\rangle \langle n| = \exp \left[-\frac{iH(t-t_0)}{\hbar} \right]. \quad (20.9)$$

For a function f of a diagonalizable self-adjoint operator, $f(H) = \sum_n f(E_n) |n\rangle \langle n|$; the exponential can equivalently be defined by its convergent power series in finite dimensions. Continuous spectra require the corresponding spectral integrals.

An energy eigenket evolves only by an overall phase. Its ray is fixed, and the expectation value of every observable with no explicit time dependence remains constant. Such a state is called **stationary**. A superposition of different energies acquires relative phases, giving

$$\langle X \rangle(t) = \sum_{m,n} c_m(t_0)^* c_n(t_0) e^{i(E_m - E_n)(t-t_0)/\hbar} X_{mn}.$$

The observable frequencies are energy differences divided by $\hbar$. Adding a constant $C\mathbf{1}$ to H multiplies every state by the same phase $e^{-iC(t-t_0)/\hbar}$ and leaves these frequencies unchanged.

20.4 When the Hamiltonian depends on time

The formal exponential of an integral is valid if $[H(t), H(s)] = 0$ at all pairs of times:

$$U(t, t_0) = \exp \left[-\frac{i}{\hbar} \int_{t_0}^t H(s) ds \right].$$

Without this commutativity condition, operator order matters. Integrate (20.6) once and iterate:

$$\begin{aligned} U(t, t_0) &= \mathbf{1} - \frac{i}{\hbar} \int_{t_0}^t H(t_1) U(t_1, t_0) dt_1 \\ &= \mathbf{1} - \frac{i}{\hbar} \int_{t_0}^t H(t_1) dt_1 \\ &\quad + \left(-\frac{i}{\hbar} \right)^2 \int_{t_0}^t dt_1 \int_{t_0}^{t_1} dt_2 H(t_1) H(t_2) + \cdots \end{aligned}$$

Later times appear to the left. This time-ordered series is denoted $\mathcal{T} \exp[-(i/\hbar) \int H dt]$. It explains why instantaneously diagonalizing $H(t)$ does not, by itself, solve an arbitrary time-dependent problem: its eigenvectors may also change.

21 Spin precession and driven two-level dynamics

21.1 Precession calculated from the evolving state

Use the S_z eigenbasis $|\uparrow\rangle, |\downarrow\rangle$, and let

$$H = \frac{E_0}{2} (|\uparrow\rangle\langle\uparrow| - |\downarrow\rangle\langle\downarrow|) = \frac{E_0}{\hbar} S_z.$$

For $|\psi; 0\rangle = |+\rangle = (|\uparrow\rangle + |\downarrow\rangle)/\sqrt{2}$,

$$|\psi; t\rangle = \frac{1}{\sqrt{2}} \left(e^{-iE_0 t/(2\hbar)} |\uparrow\rangle + e^{iE_0 t/(2\hbar)} |\downarrow\rangle \right).$$

The operators needed to evaluate its spin are

$$\begin{aligned} S_x &= \frac{\hbar}{2} (|\uparrow\rangle\langle\downarrow| + |\downarrow\rangle\langle\uparrow|), \\ S_y &= \frac{\hbar}{2i} (|\uparrow\rangle\langle\downarrow| - |\downarrow\rangle\langle\uparrow|), \\ S_z &= \frac{\hbar}{2} (|\uparrow\rangle\langle\uparrow| - |\downarrow\rangle\langle\downarrow|). \end{aligned}$$

Set $\omega_0 = E_0/\hbar$. Then

$$\begin{aligned} \langle S_x \rangle &= \frac{\hbar}{4} (e^{i\omega_0 t} + e^{-i\omega_0 t}) = \frac{\hbar}{2} \cos \omega_0 t, \\ \langle S_y \rangle &= \frac{\hbar}{4i} (e^{i\omega_0 t} - e^{-i\omega_0 t}) = \frac{\hbar}{2} \sin \omega_0 t, \\ \langle S_z \rangle &= \frac{\hbar}{2} \left(\frac{1}{2} - \frac{1}{2} \right) = 0. \end{aligned}$$

The spin expectation value rotates about the positive z axis with signed angular frequency ω_0 .

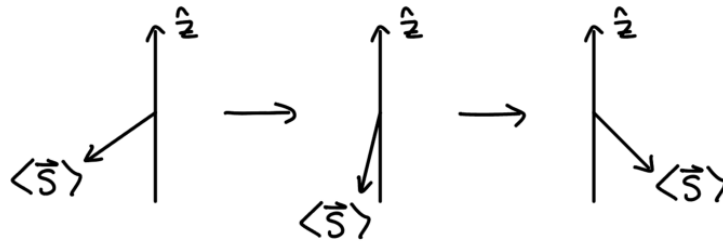


Figure 31: Precession of the spin expectation value under a Hamiltonian proportional to S_z .

For a magnetic moment $\boldsymbol{\mu} = \gamma \mathbf{S}$ in a static field $\mathbf{B} = B_z \hat{z}$, the interaction Hamiltonian is

$$H = -\boldsymbol{\mu} \cdot \mathbf{B} = -\gamma B_z S_z.$$

Consequently $E_0 = -\hbar\gamma B_z$ and $\boldsymbol{\omega} = -\gamma \mathbf{B}$, the classical Larmor frequency. This identifies the energy generator with the magnetic interaction energy and uses the same $\hbar$ that fixes the spin quantum.

21.2 Exact Rabi oscillations for a rotating field

A useful driven two-level model is

$$H(t) = \frac{\hbar}{2} [\omega_0 \sigma_z + \Omega (\cos \omega t \sigma_x + \sin \omega t \sigma_y)]. \quad (21.1)$$

The transverse field rotates at frequency ω . Define $R(t) = e^{-i\omega t \sigma_z/2}$ and write $|\psi; t\rangle = R(t)|\chi; t\rangle$. Substitution into the Schrödinger equation gives

$$i\hbar|\dot{\chi}\rangle = (R^\dagger H R - i\hbar R^\dagger \dot{R}) |\chi\rangle.$$

Since $R\sigma_x R^\dagger = \cos \omega t \sigma_x + \sin \omega t \sigma_y$ and $R^\dagger \dot{R} = -i\omega \sigma_z/2$, the rotating-frame Hamiltonian is constant:

$$H_{\text{rot}} = \frac{\hbar}{2} (\Delta \sigma_z + \Omega \sigma_x), \quad \Delta = \omega_0 - \omega.$$

Let $\Omega_R = \sqrt{\Delta^2 + \Omega^2}$. The Pauli algebra gives $(\Delta \sigma_z + \Omega \sigma_x)^2 = \Omega_R^2 \mathbf{1}$, so

$$e^{-iH_{\text{rot}}t/\hbar} = \cos \frac{\Omega_R t}{2} \mathbf{1} - i \sin \frac{\Omega_R t}{2} \frac{\Delta \sigma_z + \Omega \sigma_x}{\Omega_R}.$$

Starting from $|\uparrow\rangle$, the down-state amplitude in the rotating frame is

$$-i \frac{\Omega}{\Omega_R} \sin \frac{\Omega_R t}{2}.$$

The final factor $R(t)$ only changes its phase, so

$$P_{\uparrow \rightarrow \downarrow}(t) = \frac{\Omega^2}{\Delta^2 + \Omega^2} \sin^2 \left(\frac{\sqrt{\Delta^2 + \Omega^2} t}{2} \right). \quad (21.2)$$

On resonance, $\Delta = 0$, complete population transfer occurs at $t = \pi/|\Omega|$. Detuning both changes the oscillation frequency and reduces the maximum transfer probability. Equation (21.2) is exact for the rotating field in (21.1); a linearly oscillating field leads to this form after an appropriate rotating-wave approximation.

22 The Schrödinger and Heisenberg pictures

22.1 A simultaneous change of states and observables

For a unitary $W(t)$, transform

$$|\psi\rangle \mapsto W(t)|\psi\rangle, \quad X \mapsto W(t)XW(t)^\dagger.$$

Every matrix element is unchanged:

$$\langle \phi | W^\dagger (W X W^\dagger) W | \psi \rangle = \langle \phi | X | \psi \rangle.$$

The time dependence can therefore be redistributed between states and operators without changing predictions. If the original state obeys $i\hbar|\dot{\psi}\rangle = H|\psi\rangle$, its transformed state obeys

$$i\hbar \frac{d}{dt} (W|\psi\rangle) = (W H W^\dagger + i\hbar \dot{W} W^\dagger) W|\psi\rangle.$$

The extra derivative term is necessary whenever W depends on time. It is the general origin of the rotating-frame correction above.

22.2 Constructing the Heisenberg picture

Write $U(t) = U(t, 0)$. The Schrödinger-picture state and operator are $|\psi_S(t)\rangle = U(t)|\psi; 0\rangle$ and $X_S(t)$. Choose $W = U^\dagger$:

$$|\psi_H\rangle = U(t)^\dagger |\psi_S(t)\rangle = |\psi; 0\rangle, \quad (22.1)$$

$$X_H(t) = U(t)^\dagger X_S(t) U(t). \quad (22.2)$$

Then

$$\langle X \rangle(t) = \langle \psi_S(t) | X_S(t) | \psi_S(t) \rangle = \langle \psi_H | X_H(t) | \psi_H \rangle.$$

The Heisenberg state is fixed; the operators carry the dynamics. This is an equivalent description of the same experiment, rather than a modification of the evolution law.

22.3 Derivation of the Heisenberg equation

Use $\dot{U} = -iH_S U/\hbar$ and $\dot{U}^\dagger = iU^\dagger H_S/\hbar$ to differentiate (22.2):

$$\begin{aligned} \dot{X}_H &= \dot{U}^\dagger X_S U + U^\dagger \dot{X}_S U + U^\dagger X_S \dot{U} \\ &= \frac{i}{\hbar} U^\dagger (H_S X_S - X_S H_S) U + U^\dagger \dot{X}_S U \\ &= \frac{i}{\hbar} [H_H, X_H] + U^\dagger \dot{X}_S U, \quad H_H = U^\dagger H_S U. \end{aligned}$$

Thus

$$\boxed{\frac{dX_H}{dt} = \frac{i}{\hbar} [H_H(t), X_H(t)] + \left(\frac{\partial X_S}{\partial t} \right)_H}. \quad (22.3)$$

The last term means $U^\dagger (\partial_t X_S) U$. It describes explicit time dependence in the definition of the observable and must not be confused with the total derivative of X_H .

If H_S is time independent, $U = e^{-iH_S t/\hbar}$ commutes with H_S , and therefore $H_H = H_S$. For a driven Hamiltonian this equality does not generally hold. Taking $X_S = H_S$ gives the energy-balance law

$$\frac{d}{dt} \langle H \rangle = \left\langle \frac{\partial H_S}{\partial t} \right\rangle.$$

Only explicit time dependence can change the mean energy of the closed system.

22.4 Solving spin precession as an operator equation

For $H = \omega_0 S_z$, the spin commutators give

$$[H, S_x] = i\hbar\omega_0 S_y, \quad [H, S_y] = -i\hbar\omega_0 S_x, \quad [H, S_z] = 0.$$

Hence $S_{z,H}(t) = S_z$. To solve for $S_{x,H}$, expand in the fixed Schrödinger operators:

$$S_{x,H}(t) = a(t)S_x + b(t)S_y.$$

Substitution into (22.3) gives

$$\dot{a}S_x + \dot{b}S_y = \omega_0 bS_x - \omega_0 aS_y.$$

Linear independence implies $\dot{a} = \omega_0 b$, $\dot{b} = -\omega_0 a$. The initial condition $S_{x,H}(0) = S_x$ gives $a(0) = 1$, $b(0) = 0$, so

$$a(t) = \cos \omega_0 t, \quad b(t) = -\sin \omega_0 t.$$

Repeating the calculation with $S_{y,H}(0) = S_y$ gives

$$\begin{aligned} S_{x,H}(t) &= \cos \omega_0 t S_x - \sin \omega_0 t S_y, \\ S_{y,H}(t) &= \sin \omega_0 t S_x + \cos \omega_0 t S_y, \\ S_{z,H}(t) &= S_z. \end{aligned} \tag{22.4}$$

These operator identities are valid for every initial state. Taking expectation values yields

$$\begin{pmatrix} \langle S_x \rangle_t \\ \langle S_y \rangle_t \\ \langle S_z \rangle_t \end{pmatrix} = \begin{pmatrix} \cos \omega_0 t & -\sin \omega_0 t & 0 \\ \sin \omega_0 t & \cos \omega_0 t & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \langle S_x \rangle_0 \\ \langle S_y \rangle_0 \\ \langle S_z \rangle_0 \end{pmatrix}.$$

The previous calculation with $|+x\rangle$ is one particular initial condition for this general rotation law.

Example 22.1 (Time-dependent eigenkets of a Heisenberg operator). Since $S_{x,H} = U^\dagger S_x U$, an instantaneous positive eigenket is

$$|+x; t\rangle_H = U(t)^\dagger |+x\rangle = \frac{1}{\sqrt{2}} (e^{i\omega_0 t/2} |\uparrow\rangle + e^{-i\omega_0 t/2} |\downarrow\rangle).$$

Indeed, $S_{x,H}|+x; t\rangle_H = U^\dagger S_x |+x\rangle = (\hbar/2)|+x; t\rangle_H$. This eigenket is time dependent, even up to an overall phase for generic t . There is no contradiction with (22.1): it is the instantaneous eigenvector of a changing measurement operator, not the fixed state describing one system's unmeasured evolution. At different times it specifies different preparations that would make that measurement certain.

23 Canonical quantization and Ehrenfest's theorem

23.1 Classical evolution as a Poisson bracket

For canonical coordinates q_i, p_i and a classical observable $X(q, p, t)$, Hamilton's equations give

$$\begin{aligned} \frac{dX}{dt} &= \sum_i \left(\frac{\partial X}{\partial q_i} \dot{q}_i + \frac{\partial X}{\partial p_i} \dot{p}_i \right) + \frac{\partial X}{\partial t} \\ &= \sum_i \left(\frac{\partial X}{\partial q_i} \frac{\partial H}{\partial p_i} - \frac{\partial X}{\partial p_i} \frac{\partial H}{\partial q_i} \right) + \frac{\partial X}{\partial t} = \{X, H\}_{\text{PB}} + \partial_t X. \end{aligned}$$

The Poisson bracket is

$$\{X, Y\}_{\text{PB}} = \sum_i (\partial_{q_i} X \partial_{p_i} Y - \partial_{p_i} X \partial_{q_i} Y).$$

It is bilinear and antisymmetric, satisfies the product rule, and obeys

$$\{\{X, Y\}_{\text{PB}}, Z\}_{\text{PB}} + \{\{Y, Z\}_{\text{PB}}, X\}_{\text{PB}} + \{\{Z, X\}_{\text{PB}}, Y\}_{\text{PB}} = 0.$$

Commutators satisfy the same algebraic identities. This motivates the correspondence

$$\{X, Y\}_{\text{PB}} \longleftrightarrow \frac{1}{i\hbar}[\widehat{X}, \widehat{Y}].$$

In particular, canonical quantization imposes

$$[\widehat{q}_i, \widehat{p}_j] = i\hbar\delta_{ij}\mathbf{1}, \quad [\widehat{q}_i, \widehat{q}_j] = 0, \quad [\widehat{p}_i, \widehat{p}_j] = 0. \quad (23.1)$$

The classical Hamiltonian suggests a quantum Hamiltonian built from these operators, after operator-ordering and domain choices are made.

23.2 Angular momentum brackets worked out

For a particle, $L_i = \epsilon_{iab}q_ap_b$. The derivatives are

$$\partial_{q_a}L_i = \epsilon_{iab}p_b, \quad \partial_{p_a}L_i = \epsilon_{iba}q_b.$$

A direct example is $L_x = yp_z - zp_y$, $L_y = zp_x - xp_z$. Only the coordinate z and momentum p_z contribute to their bracket:

$$\begin{aligned} \{L_x, L_y\}_{\text{PB}} &= (-p_y)(-x) - y(p_x) \\ &= xp_y - yp_x = L_z. \end{aligned}$$

The cyclic cases follow by relabeling coordinates, giving

$$\{L_i, L_j\}_{\text{PB}} = \epsilon_{ijk}L_k.$$

Canonical quantization produces $[\widehat{L}_i, \widehat{L}_j] = i\hbar\epsilon_{ijk}\widehat{L}_k$, the same algebra satisfied by spin. The algebra is shared, although orbital and intrinsic angular momentum need not act on the same kind of Hilbert space.

23.3 Ordering and the limits of the correspondence

Classically $pq = qp$, but quantum mechanically $\widehat{p}\widehat{q} = \widehat{q}\widehat{p} - i\hbar\mathbf{1}$. For a real classical observable pq , a natural self-adjoint choice is

$$pq \longmapsto \frac{1}{2}(\widehat{p}\widehat{q} + \widehat{q}\widehat{p}).$$

Different choices for more complicated observables can differ by terms that vanish in a classical limit. There is no universal quantization map that preserves the Poisson bracket exactly for every classical function. The canonical relations and Hamiltonian specify a quantum model; the bracket correspondence is exact for suitable subalgebras and provides a leading semiclassical guide more generally.

It is often useful to reverse the viewpoint: start from a quantum model, identify its canonical variables, and determine the classical behavior to which it reduces. Distinct quantum theories may share the same leading classical limit.

23.4 Ehrenfest's theorem and mean trajectories

Taking the expectation value of (22.3) in the fixed Heisenberg state gives

$$\boxed{\frac{d}{dt}\langle X \rangle = \frac{i}{\hbar}\langle [H, X] \rangle + \langle \partial_t X \rangle.} \quad (23.2)$$

The result is picture independent if states and operators are evaluated consistently. For $H = p^2/(2m) + V(q)$,

$$[p^2, q] = p[p, q] + [p, q]p = -2i\hbar p, \quad [V(q), p] = i\hbar V'(q),$$

so

$$\frac{d\langle q \rangle}{dt} = \frac{\langle p \rangle}{m}, \quad \frac{d\langle p \rangle}{dt} = -\langle V'(q) \rangle. \quad (23.3)$$

The force is averaged over the quantum state. In general, $\langle V'(q) \rangle \neq V'(\langle q \rangle)$, so the mean position need not follow an exact classical trajectory. For a narrow packet, expanding around $\bar{q} = \langle q \rangle$ gives

$$\langle V'(q) \rangle = V'(\bar{q}) + \frac{1}{2}V'''(\bar{q})\langle (q - \bar{q})^2 \rangle + \cdots.$$

The first correction depends on the packet width. For a quadratic potential the force is linear, the correction vanishes identically, and the first moments satisfy the classical equations exactly. Larmor precession likewise closes exactly because the spin equations are linear in the spin components.

24 Ammonia inversion: dynamics determined by symmetry

24.1 The two-state approximation

In an ammonia molecule, the nitrogen atom prefers a position on either side of the plane through the three hydrogen atoms. Holding the hydrogen framework fixed gives two equivalent configurations:

$$|1\rangle = \text{nitrogen above the plane}, \quad |2\rangle = \text{nitrogen below the plane}.$$

These states span an effective two-dimensional Hilbert space. The actual molecule also has rotational, vibrational, and other degrees of freedom; the model isolates its inversion coordinate.

Turning the entire molecule over can also interchange the apparent positions of nitrogen, but this is a different motion from inversion through a fixed hydrogen framework. To distinguish them, consider a molecule carrying angular momentum normal to the hydrogen plane. An isolated reversal of its molecular axis would alter the angular-momentum direction, while inversion of nitrogen through the plane does not. Reflection in the plane leaves the normal component of angular momentum invariant because angular momentum is an axial vector. Thus this rotation does not remove the inversion symmetry.

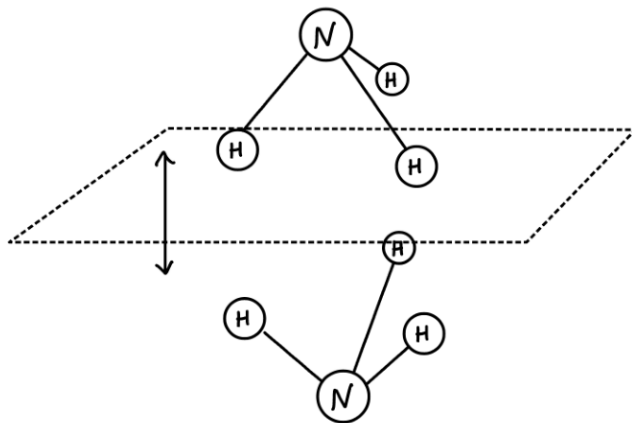


Figure 32: The two inversion configurations of ammonia, related by reflection through the hydrogen plane.

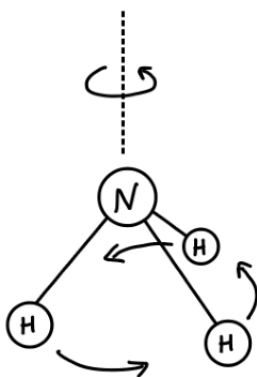


Figure 33: Overall molecular rotation is distinct from the nitrogen inversion coordinate.

24.2 Choosing phases for the reflection operator

A reflection P interchanges the localized configurations, initially with possible phases:

$$P|1\rangle = e^{i\alpha}|2\rangle, \quad P|2\rangle = e^{i\beta}|1\rangle.$$

Rephasing $|2\rangle$ relative to $|1\rangle$ makes the two phases equal. The remaining common phase in P has no effect on its action on rays, and may be removed. We can therefore choose

$$P = |1\rangle\langle 2| + |2\rangle\langle 1|, \quad P^\dagger = P, \quad P^2 = \mathbf{1}.$$

This phase convention makes the geometric statement “reflecting twice returns the original configuration” explicit at the operator level.

24.3 The most general reflection-invariant Hamiltonian

Write an arbitrary self-adjoint two-state Hamiltonian as

$$H = \begin{pmatrix} a & b \\ b^* & d \end{pmatrix}, \quad a, d \in \mathbb{R}.$$

Reflection invariance means $PHP^\dagger = H$, or $[P, H] = 0$. Since

$$PHP = \begin{pmatrix} d & b^* \\ b & a \end{pmatrix},$$

the condition forces $a = d = E_0$ and $b = b^* = A$. Hence

$$H = E_0 \mathbf{1} + AP = \begin{pmatrix} E_0 & A \\ A & E_0 \end{pmatrix}. \quad (24.1)$$

The common diagonal energy is E_0 . The off-diagonal amplitude A describes quantum tunneling between the two configurations. Symmetry fixes the form of the Hamiltonian, but not the magnitude of the tunneling amplitude. In a high-barrier regime its magnitude can be exponentially small in the barrier parameters. Its comparison with an arbitrarily chosen energy zero E_0 has no invariant meaning.

The characteristic equation is

$$\det(H - \lambda \mathbf{1}) = (E_0 - \lambda)^2 - A^2 = 0,$$

so $E_\pm = E_0 \pm A$. Solving the two rows of $(H - E_\pm \mathbf{1})(a, b)^\top = 0$ gives $b = \pm a$, hence

$$|\pm\rangle = \frac{|1\rangle \pm |2\rangle}{\sqrt{2}}, \quad P|\pm\rangle = \pm|\pm\rangle, \quad H|\pm\rangle = (E_0 \pm A)|\pm\rangle. \quad (24.2)$$

Diagonalizing the symmetry first would have produced the same basis without solving the characteristic equation.

A simultaneous convention change $|2\rangle \mapsto -|2\rangle$ and $P \mapsto -P$ preserves the swap form of P , exchanges the labels “even” and “odd”, and sends $A \mapsto -A$. The splitting $2|A|$ is unchanged. In the conventional spatial-parity realization, the lowest inversion wave function is even; choose $A < 0$ so that $|+\rangle$ is the ground state. Unlike the two classical minima, the two localized kets are not energy eigenstates. Tunneling lifts their apparent degeneracy.

24.4 Time evolution from a localized configuration

Prepare $|\psi; 0\rangle = |1\rangle = (|+\rangle + |-\rangle)/\sqrt{2}$. Such localization can be achieved using electric fields that distinguish the two configurations. The subsequent field-free state is

$$|\psi; t\rangle = \frac{1}{\sqrt{2}} \left(e^{-i(E_0+A)t/\hbar} |+\rangle + e^{-i(E_0-A)t/\hbar} |-\rangle \right).$$

Returning to the localized basis,

$$|\psi; t\rangle = e^{-iE_0 t/\hbar} \left[\cos \frac{At}{\hbar} |1\rangle - i \sin \frac{At}{\hbar} |2\rangle \right]. \quad (24.3)$$

Therefore

$$\begin{aligned} P_1(t) &= \cos^2 \frac{At}{\hbar} = \frac{1}{2} \left[1 + \cos \frac{2|A|t}{\hbar} \right], \\ P_2(t) &= \sin^2 \frac{At}{\hbar} = \frac{1}{2} \left[1 - \cos \frac{2|A|t}{\hbar} \right]. \end{aligned} \quad (24.4)$$

The inversion oscillates with angular frequency $2|A|/\hbar$, the energy splitting divided by $\hbar$. Electromagnetic radiation at this frequency can coherently stimulate transitions between the inversion levels. This is the essential two-level physics behind an ammonia maser, a microwave-frequency counterpart of a laser.

The example illustrates why classically equivalent configurations need not be degenerate quantum energy eigenstates. When a tunneling matrix element is allowed, the energy eigenstates are often superpositions with a nonzero splitting. This is not a universal prohibition on degeneracy: symmetry, disconnected sectors, or accidental parameter values can preserve it. In particular, a one-dimensional unitary symmetry representation does not by itself force an energy degeneracy.

25 Quantum symmetries and selection rules

25.1 Symmetries preserve the equation of motion

Let V be time-independent and unitary. It is a symmetry when $V|\psi; t\rangle$ solves the same Schrödinger equation whenever $|\psi; t\rangle$ does. Since

$$i\hbar\partial_t(V|\psi\rangle) = VH|\psi\rangle,$$

this condition is equivalent to $VH|\psi\rangle = HV|\psi\rangle$ for every state:

$$[V, H(t)] = 0 \quad \text{for every } t. \quad (25.1)$$

The symmetry need not arise from a spatial transformation; it may act on internal degrees of freedom.

If $V|\phi\rangle = v|\phi\rangle$, then $V(H|\phi\rangle) = H(V|\phi\rangle) = vH|\phi\rangle$. Hence each symmetry eigenspace is invariant under the Hamiltonian and its evolution. More directly, $VU(t, t_0)$ and $U(t, t_0)V$ obey the same differential equation and initial condition, so they are equal. The probabilities of belonging to the symmetry sectors are therefore conserved. If every sector is one-dimensional, the evolution changes only the phase of each symmetry eigenket.

25.2 Three equivalent ways to describe a transformation

There are two active descriptions and one passive description:

$$\begin{array}{ll} \text{active state:} & |\psi\rangle \mapsto V|\psi\rangle, & X \mapsto X, \\ \text{active operator:} & |\psi\rangle \mapsto |\psi\rangle, & X \mapsto V^\dagger X V, \\ \text{passive:} & |\psi\rangle \mapsto V|\psi\rangle, & X \mapsto V X V^\dagger. \end{array}$$

The first two give the same transformed expectation value, $\langle\psi|V^\dagger X V|\psi\rangle$. The passive description leaves every matrix element invariant and amounts to changing the representation of all states and operators together. These viewpoints are independent of the choice of Schrödinger or Heisenberg picture.

In the ammonia model, define a dimensionless inversion-position observable

$$Z = |1\rangle\langle 1| - |2\rangle\langle 2|.$$

Then $P^\dagger Z P = -Z$: reflection reverses the sign of the measured inversion coordinate. In contrast, $P^\dagger H P = H$.

25.3 A general Abelian selection rule

Let $V|a\rangle = v_a|a\rangle$ and $V|b\rangle = v_b|b\rangle$, with $|v_a| = |v_b| = 1$. Suppose an operator transforms as $V^\dagger X V = \chi X$. Inserting the symmetry gives

$$\langle a|V^\dagger X V|b\rangle = v_a^* v_b \langle a|X|b\rangle = \chi \langle a|X|b\rangle.$$

Thus

$$(v_a^* v_b - \chi) \langle a|X|b\rangle = 0. \quad (25.2)$$

A matrix element can be nonzero only if the symmetry phases match. For parity eigenvalues $v_a, v_b = \pm 1$, an even operator ($\chi = +1$) connects equal parities, while an odd operator ($\chi = -1$) connects opposite parities.

For ammonia this immediately gives

$$\langle +|H|-\rangle = 0, \quad \langle +|Z|+\rangle = \langle -|Z|-\rangle = 0.$$

Indeed, direct calculation shows $Z|+\rangle = |-\rangle$ and $Z|-\rangle = |+\rangle$. The inversion dipole is odd, so it can drive the transition between the even and odd levels. The selection rule expresses this without requiring an explicit wave function.

26 Groups, representations, and triangular hopping

26.1 Why symmetry operations form a group

If V_1 and V_2 commute with H , then

$$[V_1 V_2, H] = V_1 [V_2, H] + [V_1, H] V_2 = 0.$$

Their product is unitary and is another symmetry. Likewise, $[V^{-1}, H] = -V^{-1}[V, H]V^{-1} = 0$. Including the identity, symmetries are therefore closed under multiplication and inverses. Operator multiplication is associative. These are precisely the group axioms.

A group is **Abelian** if every pair of elements commutes. A group generated by a single element is cyclic. If $V^n = \mathbf{1}$ and no smaller positive power is the identity, its elements are

$$\mathbf{1}, V, V^2, \dots, V^{n-1},$$

and multiplication adds exponents modulo n . This is the abstract cyclic group $\mathbb{Z}_n$. If no nonzero power of T equals the identity, its integer powers form a group isomorphic to $\mathbb{Z}$ under addition. An isomorphism is a bijection that preserves the group law.

A reflection generates $\mathbb{Z}_2$; a fixed translation on an infinite line generates $\mathbb{Z}$; and a rotation through $2\pi/n$ generates $\mathbb{Z}_n$. Combining transformations can give a non-Abelian group. For example, rotating and then translating a lattice does not generally give the same result as translating and then rotating it.

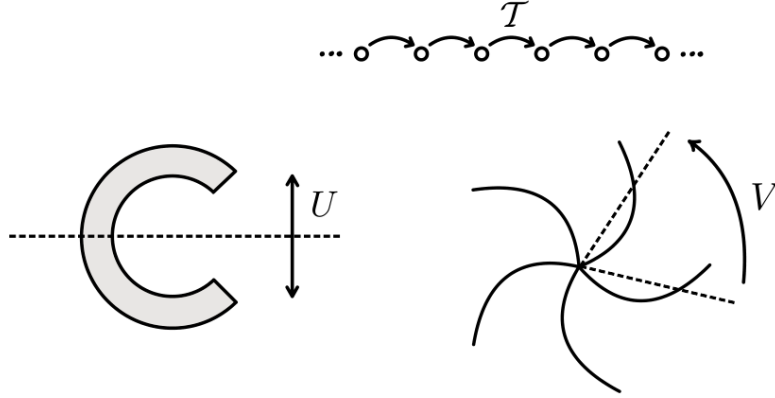


Figure 34: Geometries with discrete reflection, translation, and rotation symmetries.

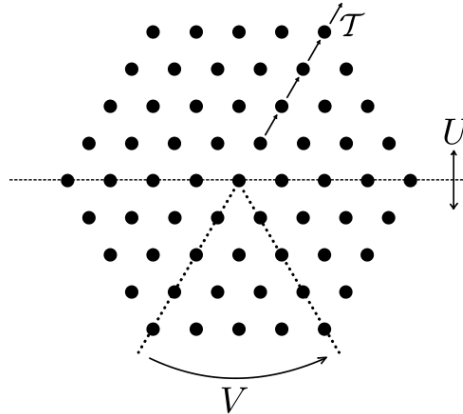


Figure 35: A triangular lattice supporting translations, rotations, and reflections.

The geometry may determine the abstract group before the Hilbert space is known. Crystal point groups describe the corresponding rotations and reflections; space groups include the translational structure. To implement an abstract group G on a quantum Hilbert space, specify unitary operators $D(g)$ satisfying

$$D(g)D(h) = D(gh), \quad D(e) = \mathbf{1}.$$

This is a **unitary representation**. The abstract multiplication law and its realization as operators are separate pieces of information.

26.2 Three sites and the dihedral group

Consider localized states $|1\rangle, |2\rangle, |3\rangle$ at the vertices of an equilateral triangle. A cyclic rotation and the three reflections act as

$$\begin{aligned} C &= |2\rangle\langle 1| + |3\rangle\langle 2| + |1\rangle\langle 3|, \\ R_1 &= |1\rangle\langle 1| + |3\rangle\langle 2| + |2\rangle\langle 3|, \\ R_2 &= |2\rangle\langle 2| + |3\rangle\langle 1| + |1\rangle\langle 3|, \\ R_3 &= |3\rangle\langle 3| + |2\rangle\langle 1| + |1\rangle\langle 2|. \end{aligned}$$

The six group elements are $\mathbf{1}, C, C^{-1}, R_1, R_2, R_3$. Direct action on the three basis states verifies

$$C^3 = \mathbf{1}, \quad R_1^2 = \mathbf{1}, \quad R_1 C R_1 = C^{-1}, \quad R_2 = R_1 C, \quad R_3 = C R_1.$$

Thus the symmetry group is D_3 , with presentation

$$D_3 = \langle c, r \mid c^3 = r^2 = (cr)^2 = e \rangle.$$

Its order is six; conventions that index a dihedral group by its order instead call this group D_6 . Here D_n denotes the symmetry group of an n -gon, of order $2n$. Two reflections also generate D_3 : $C = R_1 R_2$, $C^{-1} = R_2 R_1$, and $R_3 = R_2 R_1 R_2 = R_1 R_2 R_1$.

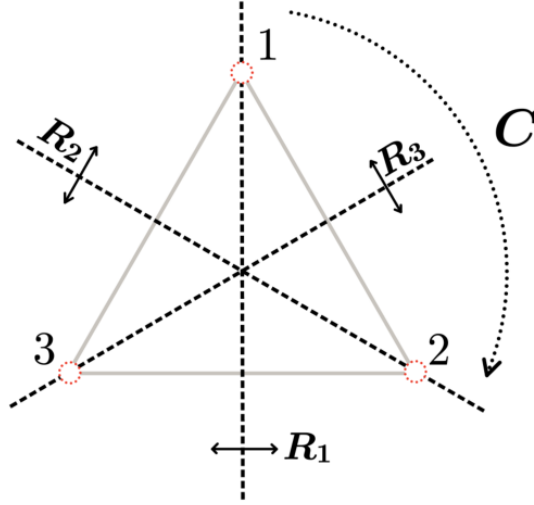


Figure 36: The rotation and reflection operations on the three-site system.

26.3 Eliminating phases in the site representation

Suppose initially $C|j\rangle = e^{i\alpha_j}|j+1\rangle$, with labels understood modulo three. The relation $C^3 = \mathbf{1}$ requires $e^{i(\alpha_1+\alpha_2+\alpha_3)} = 1$. Choose a phase for $|1\rangle$ and define the other two basis states by $|2\rangle = C|1\rangle$ and $|3\rangle = C^2|1\rangle$. Then C has exactly the phase-free form above.

A reflection fixing site one acts there by $R_1|1\rangle = \rho|1\rangle$, with $\rho = \pm 1$ from $R_1^2 = \mathbf{1}$. Using $R_1C = C^{-1}R_1$ gives

$$R_1|2\rangle = \rho|3\rangle, \quad R_1|3\rangle = \rho|2\rangle.$$

Multiplying all reflection operators by the common sign ρ leaves the dihedral group relations intact and gives the displayed permutation representation. Geometrically this chooses the trivial reflection phase of each localized site state. Additional internal degrees of freedom can carry their own symmetry representation; they must be included if present rather than absorbed into this three-state model.

26.4 The invariant Hamiltonian and its eigenvectors

Requiring invariance under all permutations generated by C, R_1 forces equal diagonal entries and equal real off-diagonal entries:

$$H = \begin{pmatrix} E_0 & A/2 & A/2 \\ A/2 & E_0 & A/2 \\ A/2 & A/2 & E_0 \end{pmatrix} = E_0\mathbf{1} + \frac{A}{2}(C + C^\dagger). \quad (26.1)$$

The factor $1/2$ is a parameter convention. To diagonalize, first use the cyclic symmetry. Set $\zeta = e^{2\pi i/3}$ and define

$$|n\rangle_C = \frac{1}{\sqrt{3}}(|1\rangle + \zeta^{-n}|2\rangle + \zeta^{-2n}|3\rangle), \quad n = 0, 1, 2.$$

For $n \neq m$, $1 + \zeta^{n-m} + \zeta^{2(n-m)} = 0$, so these kets are orthonormal. Acting with C shifts the site labels:

$$C|n\rangle_C = \frac{1}{\sqrt{3}}(|2\rangle + \zeta^{-n}|3\rangle + \zeta^{-2n}|1\rangle) = \zeta^n|n\rangle_C.$$

Equation (26.1) therefore gives

$$E_n = E_0 + \frac{A}{2}(\zeta^n + \zeta^{-n}) = E_0 + A \cos \frac{2\pi n}{3}.$$

There is one singlet and one doublet:

$$E_{0,C} = E_0 + A, \quad E_{1,C} = E_{2,C} = E_0 - \frac{A}{2}. \quad (26.2)$$

For the usual negative hopping amplitude, the fully symmetric singlet is the ground state. Symmetry-enforced degeneracy need not occur in the ground state; this model places it in the excited doublet. There is no universal rule that quantum ground states must be nondegenerate: familiar nodeless-ground-state results require suitable scalar Hamiltonians and boundary assumptions.

26.5 Why the doublet is forced to be degenerate

Reflection exchanges the two nontrivial rotation eigenstates:

$$R_1|n\rangle_C = |-n\rangle_C.$$

Since $[H, R_1] = 0$,

$$H(R_1|1\rangle_C) = R_1H|1\rangle_C = E_{1,C}|2\rangle_C.$$

But the left side is $E_{2,C}|2\rangle_C$, proving $E_{1,C} = E_{2,C}$. Rotation alone would allow different energies for its three eigenvalues; the noncommuting reflection links two of them.

The Hilbert space decomposes as an orthogonal direct sum

$$\mathcal{H} = \mathcal{H}_A \oplus \mathcal{H}_B, \quad \mathcal{H}_A = \text{span}\{|0\rangle_C\}, \quad \mathcal{H}_B = \text{span}\{|1\rangle_C, |-1\rangle_C\}.$$

A direct sum means that every ket has a unique expression as a sum of one ket from each subspace. On $\mathcal{H}_B$, the symmetry matrices are

$$C_B = \begin{pmatrix} \zeta & 0 \\ 0 & \zeta^{-1} \end{pmatrix}, \quad (R_1)_B = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

The only one-dimensional C_B -invariant subspaces are its eigenlines, and reflection exchanges them. Thus no line is invariant under the whole group: the doublet is irreducible. The projectors

$$\Pi_A = |0\rangle_C\langle 0|_C, \quad \Pi_B = |1\rangle_C\langle 1|_C + |-1\rangle_C\langle -1|_C$$

commute with every symmetry. The Hamiltonian is $H = (E_0 + A)\Pi_A + (E_0 - A/2)\Pi_B$.

26.6 A cyclic N -site extension

For $N \geq 3$ sites on a ring, let $C|j\rangle = |j+1\rangle$ modulo N and consider nearest-neighbor hopping

$$H = E_0\mathbf{1} + J(C + C^\dagger), \quad J \in \mathbb{R}.$$

The discrete Fourier kets are

$$|k_n\rangle = \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} e^{-ik_n j} |j\rangle, \quad k_n = \frac{2\pi n}{N}.$$

They obey $C|k_n\rangle = e^{ik_n} |k_n\rangle$ and have energies $E_n = E_0 + 2J \cos k_n$. A localized initial ket evolves with amplitude

$$\langle j | \psi(t) \rangle = \frac{1}{N} \sum_{n=0}^{N-1} e^{-iE_n t/\hbar} e^{-ik_n(j-j_0)} \quad \text{if } |\psi(0)\rangle = |j_0\rangle. \quad (26.3)$$

At $t = 0$, the discrete Fourier identity gives δ_{j,j_0} ; unitarity preserves the total probability. Reflection sends $k_n \rightarrow -k_n$ and pairs equal energies. The unpaired sectors are $k = 0$ and, for even N , $k = \pi$. This generalizes the singlet-doublet structure of the triangle.

27 Irreducible representations and non-Abelian selection rules

27.1 Invariant subspaces and complete reducibility

A subspace $V \subset \mathcal{H}$ is invariant when $D(g)V \subseteq V$ for every g . If the representation is unitary, its orthogonal complement is also invariant: for $w \perp V$ and $v \in V$,

$$\langle v | D(g)w \rangle = \langle D(g)^{-1}v | w \rangle = 0.$$

Thus an invariant subspace gives an orthogonal direct-sum decomposition. Iterating in finite dimensions decomposes the representation into irreducible pieces. Infinite-dimensional representations may require direct integrals rather than only discrete sums; the finite-dimensional models here need only direct sums.

For a complex Abelian unitary representation, the symmetry matrices commute and can be simultaneously diagonalized. Each common eigenline is invariant, so all finite-dimensional irreducible representations are one-dimensional. Non-Abelian groups can have higher-dimensional irreps, although a particular representation may still consist only of singlets.

27.2 Schur's lemma with proof

An intertwiner $T : V \rightarrow W$ satisfies $TD_V(g) = D_W(g)T$ for every g . Its kernel and image are invariant. If V and W are irreducible and $T \neq 0$, irreducibility therefore forces $\ker T = \{0\}$ and $\text{ran } T = W$. Hence T is an isomorphism; there are no nonzero intertwiners between inequivalent irreps.

For an endomorphism $T : V \rightarrow V$ on a finite-dimensional complex irrep, choose an eigenvalue λ . The operator $T - \lambda \mathbf{1}$ is also an intertwiner and has a nonzero kernel. The previous argument forces $T - \lambda \mathbf{1} = 0$. Thus

$$[T, D(g)] = 0 \text{ for every } g \implies T = \lambda \mathbf{1} \text{ on an irreducible space.} \quad (27.1)$$

This proves Schur's lemma in the form needed for quantum symmetries.

27.3 Equivalent irreps and multiplicity spaces

To state the dynamical consequence accurately, group equivalent copies together:

$$\mathcal{H} = \bigoplus_{\alpha} (V_{\alpha} \otimes M_{\alpha}), \quad D(g) = \bigoplus_{\alpha} (D_{\alpha}(g) \otimes \mathbf{1}_{M_{\alpha}}).$$

Here V_{α} carries an inequivalent irrep of dimension d_{α} , and M_{α} records how many equivalent copies occur. A Hamiltonian commuting with the symmetry must have the form

$$H(t) = \bigoplus_{\alpha} (\mathbf{1}_{V_{\alpha}} \otimes h_{\alpha}(t)). \quad (27.2)$$

To derive this, write each block between copies of two irreps. Commuting with all $D(g)$ says that this block is an intertwiner. It vanishes for inequivalent irreps and is a

scalar on the irrep indices for equivalent ones. Those scalars form the matrix h_α on the multiplicity space.

If an irrep occurs once, the Hamiltonian is a scalar on it, and all its d_α states are degenerate. If it occurs several times, the Hamiltonian may mix the equivalent copies. Diagonalizing h_α gives energy levels each with at least d_α -fold degeneracy. An arbitrary chosen copy need not itself be preserved by the dynamics. The symmetry type is preserved; multiplicity labels can evolve.

For a symmetry-invariant observable X , the same block rule applies. Thus it cannot connect inequivalent symmetry types, and within a single multiplicity-free irrep it is proportional to the identity. This is the non-Abelian counterpart of the Abelian selection rule. For operators that transform nontrivially, allowed transitions are determined by whether the final irrep occurs in the product of the operator's symmetry type with the initial irrep.

28 Continuous symmetries and conserved generators

28.1 One-parameter groups

A one-parameter unitary group is a continuous family $V(\theta)$ satisfying

$$V(0) = \mathbf{1}, \quad V(\theta_1)V(\theta_2) = V(\theta_1 + \theta_2). \quad (28.1)$$

Consequently $V(\theta)^{-1} = V(-\theta) = V(\theta)^\dagger$, and all members of this particular family commute. The parameter can be an angle, a displacement, or a time interval. Periodic families describe circle subgroups; a general one-parameter group need not be periodic. We first use differentiable finite-dimensional matrices. For infinite-dimensional systems, strong continuity gives a self-adjoint generator, and derivative identities hold on its appropriate domain.

Define

$$Q = i\hbar \frac{dV(\theta)}{d\theta} V(\theta)^\dagger. \quad (28.2)$$

There are three properties to establish before calling Q a conserved observable.

First, differentiate $V(\theta + \epsilon) = V(\epsilon)V(\theta)$ with respect to ϵ at zero:

$$V'(\theta) = V'(0)V(\theta).$$

Multiplying by $V(\theta)^\dagger$ gives

$$Q = i\hbar V'(0), \quad (28.3)$$

which is independent of θ . Differentiating the group law with the two factors reversed also gives $V' = VV'(0)$; therefore Q commutes with the whole one-parameter group it generates.

Second, differentiating $V^\dagger V = \mathbf{1}$ gives $V'^\dagger V + V^\dagger V' = 0$. Equivalently, $VV'^\dagger = -V'V^\dagger$, and hence

$$Q^\dagger = -i\hbar VV'^\dagger = i\hbar V'V^\dagger = Q.$$

Thus the generator is Hermitian.

Third, suppose every $V(\theta)$ is a time-independent symmetry: $[V(\theta), H(t)] = 0$ for all θ, t . Differentiating at zero yields

$$[V'(0), H(t)] = 0, \quad [Q, H(t)] = 0.$$

Because Q has no explicit time dependence, the Heisenberg equation then gives

$$\frac{dQ_H}{dt} = \frac{i}{\hbar}[H_H, Q_H] = 0. \quad (28.4)$$

This proves the time-independent version of the **quantum Noether theorem**: a continuous unitary symmetry has a self-adjoint conserved generator. In particular, every state's expectation of Q is constant. The more general case with explicit time dependence is derived in Section 32.

Conversely, if Q is time independent and commutes with $H(t)$, then every power of Q commutes with $H(t)$, so its exponential does as well. Solving $i\hbar V' = QV$ with $V(0) = \mathbf{1}$ gives

$$V(\theta) = \exp\left(-\frac{i\theta Q}{\hbar}\right). \quad (28.5)$$

A conserved generator and its one-parameter symmetry therefore contain the same information.

28.2 Phase conventions and additive constants

Replacing $V(\theta)$ by $\tilde{V}(\theta) = e^{ia\theta}V(\theta)$ preserves its action on rays and its one-parameter group law. It changes the generator to

$$\tilde{Q} = Q - \hbar a \mathbf{1}.$$

An additive scalar in a generator is thus an overall-phase convention. In a finite-dimensional space of dimension d , one can set $\text{Tr } Q = 0$ by subtracting $(\text{Tr } Q/d)\mathbf{1}$. There is generally no finite trace with which to make this prescription in an infinite-dimensional Hilbert space.

There is also a global qualification. If the parameter describes a circle with a specified period, an arbitrary rephasing on the real parameter line need not remain single-valued on that circle. A shift that is harmless for the infinitesimal generator can change whether we obtain an ordinary or a projective representation of the global group. Spin rotations provide the central example below.

28.3 Spin generates spatial rotations

For the qubit Hamiltonian proportional to S_z , consider

$$V_0(\theta) = |\uparrow\rangle\langle\uparrow| + e^{i\theta}|\downarrow\rangle\langle\downarrow|.$$

It commutes with H and obeys the group law. Direct differentiation would give

$$Q_0 = -\hbar|\downarrow\rangle\langle\downarrow| = S_z - \frac{\hbar}{2}\mathbf{1}.$$

Remove the scalar part by taking

$$V_z(\theta) = e^{-i\theta/2} V_0(\theta) = e^{-i\theta/2} |\uparrow\rangle\langle\uparrow| + e^{i\theta/2} |\downarrow\rangle\langle\downarrow| = e^{-i\theta S_z/\hbar}. \quad (28.6)$$

The generator is now exactly S_z .

To identify the spatial action, recall

$$S_x = \frac{\hbar}{2} (|\uparrow\rangle\langle\downarrow| + |\downarrow\rangle\langle\uparrow|), \quad S_y = \frac{\hbar}{2i} (|\uparrow\rangle\langle\downarrow| - |\downarrow\rangle\langle\uparrow|).$$

Conjugation gives

$$\begin{aligned} V_z^\dagger S_x V_z &= \frac{\hbar}{2} (e^{i\theta} |\uparrow\rangle\langle\downarrow| + e^{-i\theta} |\downarrow\rangle\langle\uparrow|) = \cos\theta S_x - \sin\theta S_y, \\ V_z^\dagger S_y V_z &= \frac{\hbar}{2i} (e^{i\theta} |\uparrow\rangle\langle\downarrow| - e^{-i\theta} |\downarrow\rangle\langle\uparrow|) = \sin\theta S_x + \cos\theta S_y, \\ V_z^\dagger S_z V_z &= S_z. \end{aligned}$$

In the transformed state $V_z|\psi\rangle$, therefore,

$$\begin{pmatrix} \langle S_x \rangle' \\ \langle S_y \rangle' \\ \langle S_z \rangle' \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \langle S_x \rangle \\ \langle S_y \rangle \\ \langle S_z \rangle \end{pmatrix}. \quad (28.7)$$

This is the right-handed active rotation about the positive z axis.

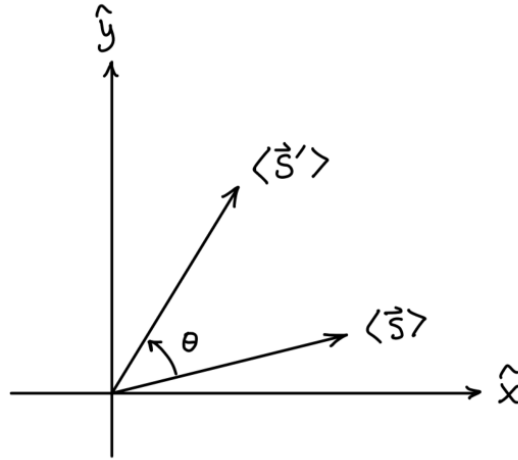


Figure 37: Rotation of the spin expectation value around the z axis.

The same construction in the S_x eigenbasis gives

$$\begin{aligned} V_x(\theta) &= e^{-i\theta/2} |+\rangle\langle+| + e^{i\theta/2} |-\rangle\langle-| \\ &= \cos\frac{\theta}{2} \mathbf{1} - i \sin\frac{\theta}{2} \sigma_x = \begin{pmatrix} \cos(\theta/2) & -i \sin(\theta/2) \\ -i \sin(\theta/2) & \cos(\theta/2) \end{pmatrix}. \end{aligned} \quad (28.8)$$

Here we used $|\pm x\rangle = (|\uparrow\rangle \pm |\downarrow\rangle)/\sqrt{2}$. The derivative at zero is $-i\sigma_x/2$, so $Q_x = S_x$. Multiplying the matrices, or using the Pauli product rule, gives

$$\begin{aligned} V_x^\dagger S_x V_x &= S_x, \\ V_x^\dagger S_y V_x &= \cos \theta S_y - \sin \theta S_z, \\ V_x^\dagger S_z V_x &= \sin \theta S_y + \cos \theta S_z. \end{aligned}$$

These are exactly the components of a right-handed rotation about x .

Likewise,

$$\begin{aligned} V_y(\theta) &= \cos \frac{\theta}{2} \mathbf{1} - \sin \frac{\theta}{2} (|\uparrow\rangle\langle\downarrow| - |\downarrow\rangle\langle\uparrow|) \\ &= \cos \frac{\theta}{2} \mathbf{1} - i \sin \frac{\theta}{2} \sigma_y = \begin{pmatrix} \cos(\theta/2) & -\sin(\theta/2) \\ \sin(\theta/2) & \cos(\theta/2) \end{pmatrix}, \end{aligned} \quad (28.9)$$

so $Q_y = S_y$, and

$$\begin{aligned} V_y^\dagger S_y V_y &= S_y, \\ V_y^\dagger S_z V_y &= \cos \theta S_z - \sin \theta S_x, \\ V_y^\dagger S_x V_y &= \sin \theta S_z + \cos \theta S_x. \end{aligned}$$

All three rotation operators act on the qubit state space. Whether a particular rotation is a *symmetry of its dynamics* is a separate question: the corresponding generator must commute with the actual Hamiltonian. A nonzero fixed field along z preserves rotations about z , but generally breaks rotations about x and y . With no preferred direction and a scalar Hamiltonian on the doublet, all rotations are symmetries.

28.4 Energy and time translations

For a time-independent Hamiltonian, H itself is a conserved charge. Its generated symmetry is

$$V(\Delta t) = e^{-i\Delta t H/\hbar} = U(t + \Delta t, t).$$

Acting on an evolving state advances the solution by a fixed amount: $V(\Delta t)|\psi, t\rangle = |\psi, t + \Delta t\rangle$. Thus energy conservation is the Noether statement associated with time-translation invariance. Explicitly time-dependent driving usually removes this symmetry and permits energy exchange with the external apparatus.

28.5 Infinitesimal transformations in the three viewpoints

Expand $V(\theta) = \mathbf{1} - i\theta Q/\hbar + O(\theta^2)$. In the active state viewpoint,

$$|\psi\rangle \mapsto |\psi\rangle - \frac{i\theta}{\hbar} Q|\psi\rangle, \quad X \mapsto X.$$

To describe the same change with the state fixed, transform the observable by $V^\dagger X V$:

$$|\psi\rangle \mapsto |\psi\rangle, \quad X \mapsto X + \frac{i\theta}{\hbar} [Q, X] = X - \frac{i\theta}{\hbar} [X, Q].$$

Finally, a passive change that leaves matrix elements unchanged uses $|\psi\rangle \mapsto V|\psi\rangle$ and $X \mapsto V X V^\dagger$:

$$|\psi\rangle \mapsto |\psi\rangle - \frac{i\theta}{\hbar} Q |\psi\rangle, \quad X \mapsto X + \frac{i\theta}{\hbar} [X, Q].$$

The opposite signs in the two operator transformations follow from which side carries V and which carries $V^\dagger$.

29 Lie groups and Lie algebras

29.1 Composing continuous transformations

If $Q_1, \dots, Q_n$ are conserved Hermitian charges, any real linear combination is also a conserved Hermitian charge. We can therefore form

$$V(\theta^1, \dots, \theta^n) = \exp\left(-\frac{i}{\hbar} \sum_a \theta^a Q_a\right).$$

If the charges commute, multiplication simply adds their parameters. When they do not commute, the order of transformations matters.

To see the first correction, write $e^X e^Y = e^Z$ and compare series through quadratic order. The left side is

$$\mathbf{1} + X + Y + \frac{X^2}{2} + XY + \frac{Y^2}{2} + O(3).$$

Putting $Z = X + Y + Z_2 + O(3)$ on the right gives

$$\mathbf{1} + X + Y + Z_2 + \frac{1}{2}(X^2 + XY + YX + Y^2) + O(3).$$

Equality requires $Z_2 = \frac{1}{2}(XY - YX)$. The full Baker–Campbell–Hausdorff expansion begins

$$e^X e^Y = \exp\left(X + Y + \frac{1}{2}[X, Y] + \frac{1}{12}[X, [X, Y]] - \frac{1}{12}[Y, [X, Y]] + \dots\right). \quad (29.1)$$

With $X = -i\theta_1 Q_1/\hbar$ and $Y = -i\theta_2 Q_2/\hbar$, this becomes

$$e^{-i\theta_1 Q_1/\hbar} e^{-i\theta_2 Q_2/\hbar} = \exp\left\{-\frac{i}{\hbar} \left(\theta_1 Q_1 + \theta_2 Q_2 + \frac{\theta_1 \theta_2}{2i\hbar} [Q_1, Q_2] + O(\theta^3)\right)\right\}. \quad (29.2)$$

Reversing the order changes the sign of the displayed commutator term. A useful way to isolate noncommutativity is the small group commutator:

$$e^X e^Y e^{-X} e^{-Y} = \mathbf{1} + [X, Y] + O(3) = \mathbf{1} - \frac{\theta_1 \theta_2}{\hbar^2} [Q_1, Q_2] + O(3).$$

Thus the Lie bracket records the leading discrepancy between doing one infinitesimal transformation before or after another.

29.2 Closure and conservation of commutators

The properly normalized commutator

$$C = \frac{1}{i\hbar}[Q_1, Q_2]$$

is Hermitian: taking its adjoint changes the sign of both $1/i$ and the commutator. If $[H, Q_1] = [H, Q_2] = 0$, the derivation identity gives

$$[H, [Q_1, Q_2]] = [[H, Q_1], Q_2] + [Q_1, [H, Q_2]] = 0.$$

Hence C is another conserved charge. This identity is equivalent to the Jacobi identity, so the result does not depend on any particular representation.

A finite list of independent generators describes a closed finite-dimensional Lie algebra when

$$[Q_i, Q_j] = i\hbar \sum_k f_{ij}{}^k Q_k, \quad f_{ij}{}^k \in \mathbb{R}. \quad (29.3)$$

Reality follows because the normalized commutator and the basis charges are Hermitian. Antisymmetry gives $f_{ij}{}^k = -f_{ji}{}^k$. Substitute this expansion into $[[Q_i, Q_j], Q_k] + [[Q_j, Q_k], Q_i] + [[Q_k, Q_i], Q_j] = 0$. The coefficient of the independent Q_m is $-\hbar^2 \sum_l (f_{ij}{}^l f_{lk}{}^m + \text{cdots})$, so

$$\sum_l (f_{ij}{}^l f_{lk}{}^m + f_{jk}{}^l f_{li}{}^m + f_{ki}{}^l f_{lj}{}^m) = 0. \quad (29.4)$$

The antisymmetry and this quadratic constraint are the algebraic conditions on the structure constants.

Sometimes closure requires the identity as a central generator, even though its exponential acts only by an overall phase. The canonical example is the Heisenberg algebra

$$[x, p] = i\hbar \mathbf{1}, \quad [x, \mathbf{1}] = [p, \mathbf{1}] = 0.$$

It has no finite-dimensional matrix realization with nonzero $\hbar$: taking traces would give $0 = i\hbar \dim \mathcal{H}$. More generally, commutators of finite matrices are traceless, and a basis of traceless Hermitian charges closes without an independent identity component. The infinite-dimensional canonical operators are unbounded, so the finite-matrix trace argument does not apply to them.

29.3 The invariant trace inner product

In a finite-dimensional unitary representation, let the Q_i be real linearly independent and define

$$h_{ij} = \text{Tr}(Q_i Q_j).$$

Cyclicity shows symmetry, $h_{ji} = h_{ij}$, and Hermiticity shows reality:

$$h_{ij}^* = \text{Tr}[(Q_i Q_j)^\dagger] = \text{Tr}(Q_j Q_i) = h_{ij}.$$

For any real nonzero vector c , the Hermitian matrix $Q(c) = \sum_i c_i Q_i$ is nonzero. Diagonalizing it gives real eigenvalues q_a , at least one nonzero, and therefore

$$\sum_{ij} c_i h_{ij} c_j = \text{Tr}[Q(c)^2] = \sum_a q_a^2 > 0.$$

Thus h is a positive-definite real inner product on the represented Lie algebra.

The identity

$$\text{Tr}([A, B]C + [A, C]B) = \text{Tr}(ABC - BAC + ACB - CAB) = 0$$

follows by moving the first factor of BAC and CAB around the trace. Taking $A = Q_i$, $B = Q_j$, and $C = Q_k$, then using the structure constants, gives

$$\sum_l (f_{ij}^l h_{lk} + f_{ik}^l h_{lj}) = 0. \quad (29.5)$$

This means the infinitesimal adjoint action preserves the inner product.

Apply ordinary Gram–Schmidt to choose an orthonormal real basis of charges, so $h_{ij} = \delta_{ij}$. With all indices lowered, (29.5) becomes $f_{ijk} = -f_{ikj}$. Together with $f_{ijk} = -f_{jik}$, this makes f_{ijk} completely antisymmetric. A finite-dimensional real Lie algebra with a positive-definite invariant inner product is of compact reductive type: it splits into its Abelian center and compact semisimple factors. Here it is the image of the algebra in the unitary representation that has this property; an abstract algebra may have a kernel in that representation.

29.4 Local algebra and global group

A Lie group is a smooth manifold with smooth multiplication and inversion. Near its identity, exponentials of a linearly independent basis of tangent generators give local coordinates. Translating this patch by any group element gives a coordinate patch there. The BCH formula expresses the multiplication law near the identity in terms of the Lie bracket.

These statements are local. A general group element need not be a single exponential, and distinct global Lie groups can have the same Lie algebra. Nevertheless, the connected identity component is generated by products of elements in sufficiently small exponential neighborhoods. The Lie algebra determines the simply connected group integrating it, while additional discrete identifications distinguish other connected groups with that algebra.

Abstractly, a real Lie algebra $\mathfrak{g}$ is a real vector space with a bilinear antisymmetric bracket satisfying Jacobi. In the Hermitian convention used here, a representation assigns $x \mapsto Q_x$ linearly, with

$$Q_{[x,y]_{\mathfrak{g}}} = \frac{1}{i\hbar} [Q_x, Q_y]. \quad (29.6)$$

The factor on the right converts the operator commutator of Hermitian charges back to a Hermitian charge. One could instead use anti-Hermitian generators and remove this

factor from the representation law. A complex Lie algebra uses a complex vector space instead; the physical infinitesimal unitary transformations are naturally organized first as a real Lie algebra.

For rotations, an abstract basis ℓ_x, ℓ_y, ℓ_z obeys

$$[\ell_i, \ell_j]_{\mathfrak{g}} = \epsilon_{ijk} \ell_k.$$

Its Hermitian representatives satisfy

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

This is $\mathfrak{so}(3)$, also isomorphic to $\mathfrak{su}(2)$. The qubit operators $S_i = \hbar\sigma_i/2$ give one representation. The algebra itself does not specify the dimension of the state space, so we next determine all irreducible unitary rotation multiplets.

30 Classification of spin representations

30.1 The Casimir and ladder algebra

We seek Hermitian operators L_x, L_y, L_z satisfying

$$[L_i, L_j] = i\hbar \epsilon_{ijk} L_k \tag{30.1}$$

on an initially unspecified Hilbert space. An atom in a rotationally invariant environment must have a state space on which rotations act, but it need not transform like a qubit. The possible irreducible multiplets follow from the algebra and positivity alone.

Define the total angular momentum

$$L^2 = L_x^2 + L_y^2 + L_z^2.$$

Using $[AB, C] = A[B, C] + [A, C]B$,

$$\begin{aligned} [L^2, L_i] &= \sum_j (L_j [L_j, L_i] + [L_j, L_i] L_j) \\ &= i\hbar \sum_{jk} \epsilon_{jik} (L_j L_k + L_k L_j) = 0. \end{aligned}$$

The last sum vanishes because the Levi-Civita symbol is antisymmetric under $j \leftrightarrow k$ while the operator factor is symmetric. Thus L^2 is a central operator, or **Casimir**. It can be diagonalized simultaneously with L_z . Write a normalized simultaneous eigenvector as

$$L^2|\lambda, m\rangle = \hbar^2\lambda|\lambda, m\rangle, \quad L_z|\lambda, m\rangle = \hbar m|\lambda, m\rangle. \tag{30.2}$$

The dimensionless λ is nonnegative because $\langle L^2 \rangle = \sum_i \|L_i|\psi\rangle\|^2 \geq 0$.

Introduce

$$L_+ = L_x + iL_y, \quad L_- = L_x - iL_y = L_+^\dagger.$$

From the original commutators,

$$\begin{aligned}[L_z, L_+] &= i\hbar L_y + i(-i\hbar L_x) = \hbar L_+, \\ [L_z, L_-] &= i\hbar L_y - i(-i\hbar L_x) = -\hbar L_-, \\ [L_+, L_-] &= -i[L_x, L_y] + i[L_y, L_x] = 2\hbar L_z.\end{aligned}$$

Conversely, these three relations together with $L_x = (L_+ + L_-)/2$ and $L_y = (L_+ - L_-)/(2i)$ reproduce (30.1). The two descriptions are equivalent.

Since $[L^2, L_\pm] = 0$ and $[L_z, L_\pm] = \pm\hbar L_\pm$,

$$\begin{aligned}L^2 L_\pm |\lambda, m\rangle &= \hbar^2 \lambda L_\pm |\lambda, m\rangle, \\ L_z L_\pm |\lambda, m\rangle &= L_\pm (L_z \pm \hbar) |\lambda, m\rangle = \hbar(m \pm 1) L_\pm |\lambda, m\rangle.\end{aligned}$$

Whenever it is nonzero, $L_\pm |\lambda, m\rangle$ is therefore another state with the same λ and with m shifted by one. This explains the names raising and lowering operators.

30.2 Norms, termination, and quantization

The normalization is determined by multiplying the ladder operators:

$$\begin{aligned}L_\mp L_\pm &= (L_x \mp iL_y)(L_x \pm iL_y) \\ &= L_x^2 + L_y^2 \pm i[L_x, L_y] \\ &= L^2 - L_z^2 \mp \hbar L_z = L^2 - L_z(L_z \pm \hbar).\end{aligned}$$

Consequently,

$$\|L_\pm |\lambda, m\rangle\|^2 = \langle \lambda, m | L_\mp L_\pm | \lambda, m \rangle = \hbar^2 [\lambda - m(m \pm 1)]. \quad (30.3)$$

A norm cannot be negative. Repeatedly raising m would eventually make $m(m+1) > \lambda$ if the ladder continued forever. Therefore it must terminate at a highest weight $m = \ell$ satisfying

$$L_+ |\lambda, \ell\rangle = 0, \quad \lambda = \ell(\ell + 1).$$

At the highest state the lowering norm is $\hbar^2[\ell(\ell + 1) - \ell(\ell - 1)] = 2\hbar^2\ell$, so $\ell \geq 0$. Similarly there is a lowest weight m_- with

$$L_- |\lambda, m_- \rangle = 0, \quad m_-(m_- - 1) = \ell(\ell + 1).$$

Factoring gives

$$(m_- + \ell)(m_- - \ell - 1) = 0.$$

Since $m_- \leq \ell$, only $m_- = -\ell$ is possible. The number of unit steps between the endpoints is 2ℓ , so

$$2\ell \in \mathbb{N} \cup \{0\}, \quad \ell = 0, \frac{1}{2}, 1, \frac{3}{2}, \dots \quad (30.4)$$

Thus positivity and the ladder structure force integer or half-integer angular momentum; no separate quantization prescription was inserted.

Choosing successive ket phases to make the nonzero ladder coefficients positive, the normalized result is

$$\begin{aligned}
L^2|\ell, m\rangle &= \hbar^2\ell(\ell+1)|\ell, m\rangle, \\
L_z|\ell, m\rangle &= \hbar m|\ell, m\rangle, \\
L_{\pm}|\ell, m\rangle &= \hbar\sqrt{\ell(\ell+1)-m(m\pm 1)}|\ell, m\pm 1\rangle \\
&= \hbar\sqrt{(\ell\mp m)(\ell\pm m+1)}|\ell, m\pm 1\rangle.
\end{aligned} \tag{30.5}$$

The allowed magnetic quantum numbers are $m = -\ell, -\ell+1, \dots, \ell$. Distinct m have distinct eigenvalues of the Hermitian L_z , so the normalized kets are orthogonal. Their span has dimension $2\ell+1$ and is preserved by $L_z, L_{\pm}$, hence by all three L_i .

Conversely, for any allowed ℓ , define operators by (30.5) on an orthonormal basis of this size. Their adjoints have the required relation, and

$$\begin{aligned}
(L_+L_- - L_-L_+)|\ell, m\rangle &= \hbar^2[\ell(\ell+1) - m(m-1) \\
&\quad - \ell(\ell+1) + m(m+1)]|\ell, m\rangle \\
&= 2\hbar^2m|\ell, m\rangle = 2\hbar L_z|\ell, m\rangle.
\end{aligned}$$

The other ladder commutators follow directly from the shift in m . Every nonnegative integer or half-integer ℓ therefore produces an actual unitary representation.

30.3 Irreducibility and completeness of the classification

Let O commute with L_z, L_+, L_- . In the single ladder just built, each L_z eigenspace is one dimensional, so

$$O|\ell, m\rangle = o_m|\ell, m\rangle.$$

For $m < \ell$, apply $[O, L_+] = 0$ to the ket. Because the raising coefficient is nonzero,

$$o_{m+1} = o_m.$$

It follows that every o_m is the same and $O = c\mathbf{1}$.

If the representation were reducible, a nonzero proper invariant subspace would have an invariant orthogonal complement because the L_i are Hermitian. The orthogonal projector onto that subspace would then commute with all L_i but would not be scalar, contradicting the result just proved. The ladder representation is therefore irreducible. Conversely, in any irreducible unitary rotation representation, the invariant ladder generated from a simultaneous eigenvector must fill the whole space. This yields precisely the spin- ℓ representations. For a general unitary representation of the compact rotation or spin group, one decomposes into these finite-dimensional irreducible multiplets, possibly with repeated copies.

The first few cases are immediate. For $\ell = 0$, all generators vanish and the representation is a singlet. For $\ell = \frac{1}{2}$, the ordered basis $|\frac{1}{2}, \frac{1}{2}\rangle, |\frac{1}{2}, -\frac{1}{2}\rangle$ gives $L_i = \hbar\sigma_i/2$.

For $\ell = 1$, using the ordered basis $m = 1, 0, -1$ gives

$$L_z = \hbar \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix},$$

$$L_x = \frac{\hbar}{\sqrt{2}} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad L_y = \frac{\hbar}{\sqrt{2}} \begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix}.$$

30.4 Ordinary vectors are spin one

An ordinary three-component vector transforms by a real rotation matrix. Complexify this vector space and use its usual Hermitian inner product. A rotation about z has

$$R_z(\theta) = \begin{pmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix} = e^{-i\theta L_z/\hbar},$$

so $L_z = i\hbar R'_z(0)$. More generally,

$$(L_i)_{jk} = -i\hbar \epsilon_{ijk}. \quad (30.6)$$

These matrices are Hermitian because the real antisymmetric Levi-Civita matrices are multiplied by $-i$. Explicitly,

$$L_x = \hbar \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \quad L_y = \hbar \begin{pmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ -i & 0 & 0 \end{pmatrix}, \quad L_z = \hbar \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Direct multiplication gives $[L_x, L_y] = i\hbar L_z$ and its cyclic partners. Also $L_x^2 + L_y^2 + L_z^2 = 2\hbar^2 \mathbf{1}$, identifying $\ell(\ell+1) = 2$, hence $\ell = 1$.

The spherical basis

$$|1, 1\rangle = -\frac{\hat{x} + i\hat{y}}{\sqrt{2}}, \quad |1, 0\rangle = \hat{z}, \quad |1, -1\rangle = \frac{\hat{x} - i\hat{y}}{\sqrt{2}}$$

has L_z eigenvalues $\hbar, 0, -\hbar$ and the positive ladder coefficients in (30.5). For example,

$$L_-|1, 1\rangle = \sqrt{2}\hbar|1, 0\rangle, \quad L_-|1, 0\rangle = \sqrt{2}\hbar|1, -1\rangle.$$

Thus the familiar vector representation is unitarily equivalent to the abstract spin-one ladder representation.

30.5 How far must a state rotate to become orthogonal?

Fix a unit vector $\hat{n}$ and a normalized spin- ℓ state. Choose the eigenbasis of $L_n = \hat{n} \cdot \mathbf{L}$ and write

$$|\psi\rangle = \sum_{m=-\ell}^{\ell} c_m |\ell, m\rangle_n, \quad p_m = |c_m|^2, \quad \sum_m p_m = 1.$$

The overlap with its rotated image is

$$A(\theta) = \langle \psi | e^{-i\theta L_n/\hbar} | \psi \rangle = \sum_{m=-\ell}^{\ell} p_m e^{-im\theta}. \quad (30.7)$$

Only the probabilities in the rotation-axis basis enter; the phases of the c_m cancel.

For $\ell > 0$ and $|\theta| < \pi/(2\ell)$, every term obeys $|m\theta| < \pi/2$, hence $\cos(m\theta) > 0$. Therefore

$$\Re A(\theta) = \sum_m p_m \cos(m\theta) > 0,$$

and the overlap cannot vanish. At $|\theta| = \theta_{\min} = \pi/(2\ell)$, all cosines remain strictly positive except at $m = \pm\ell$, where they are zero. Vanishing real part then forces $p_m = 0$ for every interior weight. The imaginary part vanishes only if the two endpoint probabilities are equal. Consequently,

$$\theta_{\min} = \frac{\pi}{2\ell}, \quad |\psi\rangle = \frac{1}{\sqrt{2}} (e^{i\alpha} |\ell, \ell\rangle_n + e^{i\beta} |\ell, -\ell\rangle_n) \quad (30.8)$$

are exactly the states attaining this universal minimum. For these states $A(\theta) = \cos(\ell\theta)$. A particular other state may require a larger angle or may never become orthogonal under this one-axis rotation. For $\ell = 0$, rotations leave every state unchanged, so no orthogonal image exists.

There is a second question: can one generate an entire orthonormal basis by rotating one state about this same axis? Put $d = 2\ell + 1$ and suppose

$$|\psi_j\rangle = e^{-i\theta_j L_n/\hbar} |\psi\rangle, \quad \theta_0 = 0 < \theta_1 < \dots < \theta_{d-1} < 2\pi.$$

In the m basis, the matrix whose columns are these kets is $M_{mj} = c_m e^{-im\theta_j}$. If its d columns are orthonormal, then the square matrix M is unitary, so its rows also have unit norm. The norm of row m is $d|c_m|^2$, implying

$$|c_m|^2 = \frac{1}{d} \quad \text{for every } m. \quad (30.9)$$

With these uniform probabilities, orthogonality to the initial ket requires, for each $j > 0$,

$$0 = \langle \psi_0 | \psi_j \rangle = \frac{e^{i\ell\theta_j}}{d} \sum_{r=0}^{d-1} e^{-ir\theta_j} = \frac{e^{i\ell\theta_j}}{d} \frac{1 - e^{-id\theta_j}}{1 - e^{-i\theta_j}}.$$

The denominator is nonzero because $0 < \theta_j < 2\pi$. Thus each θ_j is a nonzero multiple of $2\pi/d$. There are exactly $d - 1$ such distinct angles in the required interval, so their ordering fixes

$$\theta_j = \frac{2\pi j}{d}, \quad j = 0, \dots, d-1. \quad (30.10)$$

These conditions are also sufficient: for $j \neq k$, the inner product is a phase times $d^{-1} \sum_{r=0}^{d-1} e^{-2\pi i r(j-k)/d} = 0$. Hence a complete family of examples is

$$|\psi\rangle = \frac{1}{\sqrt{2\ell+1}} \sum_{m=-\ell}^{\ell} e^{i\alpha_m} |\ell, m\rangle_n,$$

with arbitrary coefficient phases, rotated through the equally spaced angles in (30.10). The discrete Fourier orthogonality identity is the reason this construction works.

31 Half-integer spin, topology, and projective representations

31.1 The sign under a full rotation

In a spin- ℓ multiplet,

$$e^{-2\pi i L_z / \hbar} |\ell, m\rangle = e^{-2\pi i m} |\ell, m\rangle.$$

All the m values differ by integers, so the phase is the same across the entire multiplet:

$$U(2\pi) = (-1)^{2\ell} \mathbf{1}. \quad (31.1)$$

Integer spins return to the same ket after 2π ; half-integer spins return to its negative and require 4π to return to the same ket. Conjugating by a rotation that takes z to another axis proves the same statement for a full turn about any axis.

The sign is invisible on a single ray, but it prevents the half-integer-spin matrices from forming an ordinary representation of $\text{SO}(3)$: the identity rotation must be represented by $\mathbf{1}$, not by $-\mathbf{1}$. Rephasing the matrices along one chosen axis can remove that axis's sign, but no globally consistent choice preserves the multiplication law for all spatial rotations at once. The obstruction is encoded in the global topology of the rotation group.

31.2 The group $\text{SU}(2)$ and its double cover

For spin one half,

$$U(\hat{n}, \theta) = \exp\left(-\frac{i\theta}{2} \hat{n} \cdot \boldsymbol{\sigma}\right) = \cos \frac{\theta}{2} \mathbf{1} - i \sin \frac{\theta}{2} \hat{n} \cdot \boldsymbol{\sigma}. \quad (31.2)$$

The final equality follows from $(\hat{n} \cdot \boldsymbol{\sigma})^2 = \mathbf{1}$, which separates the exponential into its even and odd powers. These matrices are unitary with determinant one. Conversely every 2×2 special unitary matrix has the form

$$U = \begin{pmatrix} \alpha & \beta \\ -\beta^* & \alpha^* \end{pmatrix}, \quad |\alpha|^2 + |\beta|^2 = 1. \quad (31.3)$$

Indeed, the orthogonal unit vector to a first row (α, β) is $(-\beta^*, \alpha^*)$ up to phase, and determinant one fixes that phase. The four real components of α, β therefore satisfy

$$(\Re \alpha)^2 + (\Im \alpha)^2 + (\Re \beta)^2 + (\Im \beta)^2 = 1.$$

As a manifold, $\text{SU}(2)$ is the unit three-sphere S^3 in $\mathbb{R}^4$. All spin- ℓ multiplets, including half-integer ones, are ordinary representations of this group.

The map from $\text{SU}(2)$ to spatial rotations can be made explicit. Associate a real vector $\mathbf{x}$ with the traceless Hermitian matrix $X = \mathbf{x} \cdot \boldsymbol{\sigma}$. Conjugation by U preserves Hermiticity and trace, so

$$UXU^\dagger = (R(U)\mathbf{x}) \cdot \boldsymbol{\sigma}$$

for a real linear map $R(U)$. It preserves the Euclidean inner product because $\frac{1}{2} \text{Tr}(XY) = \mathbf{x} \cdot \mathbf{y}$ is unchanged by conjugation. Thus $R(U)$ is orthogonal; continuity from $U = \mathbf{1}$ gives $\det R(U) = 1$. The product rule for conjugations gives $R(UV) = R(U)R(V)$.

To identify the angle, set $N = \hat{n} \cdot \boldsymbol{\sigma}$ and use

$$NX = (\hat{n} \cdot \mathbf{x})\mathbf{1} + i(\hat{n} \times \mathbf{x}) \cdot \boldsymbol{\sigma}, \quad NXN = 2(\hat{n} \cdot \mathbf{x})N - X.$$

Expanding $(c\mathbf{1} - isN)X(c\mathbf{1} + isN)$ with $c = \cos(\theta/2)$, $s = \sin(\theta/2)$ gives Rodrigues' formula

$$R(U)\mathbf{x} = \cos \theta \mathbf{x} + \sin \theta \hat{n} \times \mathbf{x} + (1 - \cos \theta)\hat{n}(\hat{n} \cdot \mathbf{x}). \quad (31.4)$$

Every spatial rotation has such an axis-angle form, so the map is onto. Its kernel consists of matrices commuting with every Pauli matrix, hence scalar matrices. Within $\text{SU}(2)$ these are precisely $\pm \mathbf{1}$. Therefore

$$\text{SO}(3) \cong \text{SU}(2)/\{\mathbf{1}, -\mathbf{1}\}. \quad (31.5)$$

Each rotation has two lifts, U and $-U$, to the spin group.

31.3 A geometric picture of the two groups

Let $\boldsymbol{\theta} = \theta \hat{n}$ range over the closed ball of radius 2π in $\mathbb{R}^3$ and use (31.2). At its center $U = \mathbf{1}$; at every point on its outer sphere $U = -\mathbf{1}$. All outer boundary points therefore represent one point of $\text{SU}(2)$. In the interior $0 < \theta < 2\pi$, the parametrization is unique. Consequently the ball with its boundary sphere collapsed is

$$B^3/S^2 \cong S^3 \cong \text{SU}(2).$$

The center and the collapsed boundary are *distinct* points at this stage. The lower-dimensional analogy is a disk with its boundary circle collapsed to one point, giving a two-sphere $B^2/S^1 \cong S^2$.

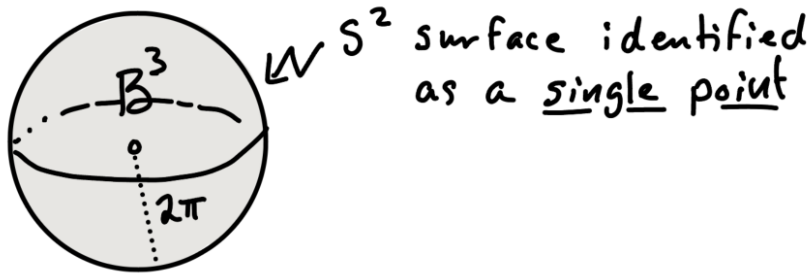


Figure 38: A three-ball with its entire boundary identified to one point describes the three-sphere.

To obtain $\text{SO}(3)$, identify U with $-U$. In axis-angle coordinates, this pairs

$$(\theta, \hat{n}) \sim (2\pi - \theta, -\hat{n}).$$

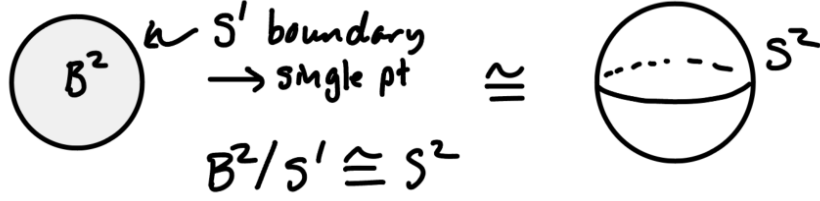


Figure 39: The analogous construction in one lower dimension: collapsing a disk boundary gives a two-sphere.

It also identifies the center with the collapsed outer boundary. On the four-dimensional coordinates of the unit S^3 , it is the antipodal identification $x \sim -x$. Thus

$$\mathrm{SO}(3) \cong S^3/\mathbb{Z}_2 = \mathbb{RP}^3. \quad (31.6)$$

An equivalent smaller model uses the ball of radius π , identifying antipodal points only on its boundary: a rotation by π about $\hat{n}$ equals one by π about $-\hat{n}$.

Every closed loop in S^3 contracts to a point; $\mathrm{SU}(2)$ is simply connected. In $\mathrm{SO}(3)$, the path that rotates from angle zero to 2π around a fixed axis is closed, because its endpoint is the identity rotation. Its lift to $\mathrm{SU}(2)$ runs from $\mathbf{1}$ to $-\mathbf{1}$, so it is not closed. If the original loop contracted, its continuously lifted endpoint could not change from $-\mathbf{1}$ to $\mathbf{1}$. The loop is therefore noncontractible. Traversing it twice lifts to a closed loop in S^3 , which does contract. This is the origin of

$$\pi_1(\mathrm{SO}(3)) = \mathbb{Z}_2.$$

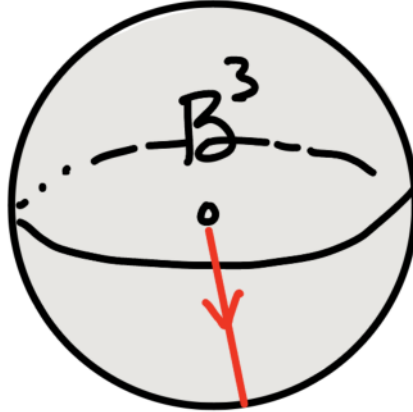


Figure 40: The path from the center to the outer boundary becomes a closed, noncontractible loop after the additional identification defining $\mathrm{SO}(3)$.

The simply connected double cover is called $\mathrm{Spin}(3) = \mathrm{SU}(2)$. More generally, $\mathrm{Spin}(n)$ is the simply connected double cover of $\mathrm{SO}(n)$ for $n \geq 3$. The circle $U(1) \cong S^1$ and $\mathrm{SU}(2) \cong S^3$ are the positive-dimensional spheres that admit a Lie group structure. The shared rotation Lie algebra captures the local generators; the distinction between $\mathrm{SO}(3)$ and its cover is global and determines which spin representations descend to ordinary representations of the rotation group.

31.4 Projective multiplication and relative phases

Physical pure states are rays, so symmetry representatives may obey

$$U_g U_h = e^{i\phi(g,h)} U_{gh}. \quad (31.7)$$

This is a projective unitary representation. The phases are not completely arbitrary: associativity of three operators requires

$$\phi(g, h) + \phi(gh, k) = \phi(h, k) + \phi(g, hk) \pmod{2\pi}.$$

Rephasing $U_g \mapsto e^{ia(g)} U_g$ changes the multiplier by

$$\phi(g, h) \mapsto \phi(g, h) + a(g) + a(h) - a(gh).$$

A projective representation is equivalent to an ordinary one only when such a rephasing removes the multiplier globally.

Half-integer spin gives a projective representation of $\text{SO}(3)$, or equivalently an ordinary representation of $\text{SU}(2)$. More generally, projective representations may require a central extension, and a universal cover alone does not remove every possible projective multiplier. For a covering group $\tilde{G}$ with discrete kernel K , the quotient is $G = \tilde{G}/K$. A representation of $\tilde{G}$ descends to a projective representation of G when every element of K acts as a scalar on the *whole* Hilbert space. For rotations, the kernel is $\{\pm 1\}$ and this condition is immediate in each irreducible spin multiplet.

The word “whole” matters. In the ammonia example, $P|+\rangle = |+\rangle$ and $P|-\rangle = -|-\rangle$. On each one-dimensional sector separately, P acts by a scalar phase. But on $a|+\rangle + b|-\rangle$ it changes a relative phase, so P is not projectively trivial on the full two-dimensional space. One cannot independently discard the phases in separate sectors when coherent superpositions between them are allowed.

For a fixed projective multiplier, finite-dimensional unitary representations can still be decomposed into invariant orthogonal subspaces: the invariant-complement argument is unchanged. The care needed is that all blocks share a consistent multiplier and that relative phases between the blocks retain their meaning. It is often simplest to work first with an ordinary representation of a suitable cover or central extension, decompose it there, and only then identify which phases are common to the full physical state.

For example, on a direct sum of integer- and half-integer-spin sectors, a 2π rotation is $+1$ on the former and -1 on the latter. It is not one scalar on that direct sum. In a field theory with fermions this is the operator $(-1)^F$, which distinguishes even and odd fermion number, and the spin group remains the relevant group on the full Hilbert space.

Ordinary classical spatial vectors and tensors transform in integer-spin representations of $\text{SO}(3)$. Half-integer spin describes a different transformation law, familiar from quantum states. Spinor fields can also be used in classical field formulations; in fermionic formulations they involve anticommuting Grassmann variables and do not describe ordinary classical particle coordinates. The representation-theoretic statement is the precise one: half-integer spin is an ordinary representation of the spin group, and a projective representation of the spatial rotation group.

32 Time-dependent symmetries and time reversal

32.1 Explicitly time-dependent unitary symmetries

A symmetry transformation need not be represented by a constant Schrödinger-picture matrix. Let $V(t)$ be unitary and ask whether

$$|\psi'(t)\rangle = V(t)|\psi(t)\rangle$$

is a solution of the same Schrödinger equation whenever $|\psi(t)\rangle$ is. Differentiating carefully gives

$$\begin{aligned} i\hbar \frac{d}{dt}|\psi'\rangle &= i\hbar \dot{V}|\psi\rangle + V i\hbar \frac{d}{dt}|\psi\rangle \\ &= (i\hbar \dot{V} + VH)|\psi\rangle. \end{aligned}$$

For this to equal $H|\psi'\rangle = HV|\psi\rangle$ for every initial state, we require

$$i\hbar \dot{V} + [V, H] = 0, \quad \text{equivalently} \quad \dot{V} + \frac{i}{\hbar}[H, V] = 0. \quad (32.1)$$

The condition $[V, H] = 0$ is recovered when $\dot{V} = 0$.

Its meaning is particularly simple in the Heisenberg picture. Put $V_H(t) = U(t, t_0)^\dagger V(t) U(t, t_0)$. The derivative is

$$\frac{dV_H}{dt} = U^\dagger \left(\dot{V} + \frac{i}{\hbar}[H, V] \right) U = 0.$$

Thus a unitary symmetry is a unitary operator constant in the Heisenberg picture. For any prescribed unitary V_0 at t_0 , the solution is

$$V(t) = U(t, t_0) V_0 U(t, t_0)^\dagger. \quad (32.2)$$

This shows that formally many time-dependent symmetries always exist. Most do not simplify the problem, because finding them by this formula already requires solving the evolution. The useful symmetries are usually those identifiable directly from the physical setup, before solving its dynamics.

32.2 Quantum Noether theorem with explicit time dependence

Suppose $V_\theta(t)$ is a differentiable one-parameter group for each t , with generator $Q(t)$ and

$$V_\theta(t) = e^{-i\theta Q(t)/\hbar}.$$

Insert $V_\theta = \mathbf{1} - i\theta Q/\hbar + O(\theta^2)$ into (32.1). At first order in θ this gives

$$\frac{\partial Q}{\partial t} + \frac{i}{\hbar}[H, Q] = 0 \quad \Longleftrightarrow \quad \frac{dQ_H}{dt} = 0. \quad (32.3)$$

Conversely, if the Hermitian $Q(t)$ obeys this equation, then $Q(t) = UQ(t_0)U^\dagger$. Exponentiating yields $V_\theta(t) = Ue^{-i\theta Q(t_0)/\hbar}U^\dagger$, which satisfies the full symmetry equation. Thus conserved Hermitian charges, including those with explicit time dependence, correspond to continuous unitary symmetries.

Example 32.1 (A conserved boost charge). For a free particle in one dimension, $H = p^2/(2m)$, consider

$$G(t) = mx - tp.$$

Although G explicitly depends on time, it is conserved. Indeed, $\partial G/\partial t = -p$, and

$$[H, G] = \frac{1}{2}[p^2, x] = \frac{1}{2}(p[p, x] + [p, x]p) = -i\hbar p.$$

Hence $(i/\hbar)[H, G] = p$, exactly canceling the explicit derivative. The associated family $e^{-ivG(t)/\hbar}$ implements a Galilean boost, with the sign of the velocity shift fixed by this chosen convention. In fact, its active operator action gives $V^\dagger p V = p - mv$ and $V^\dagger x V = x - vt$; the mean velocity therefore changes by $-v$. This example illustrates how a spacetime symmetry can have a time-dependent generator while its Heisenberg charge remains constant.

32.3 Antiunitary maps

An antilinear operator Θ obeys

$$\Theta(a|\psi\rangle + b|\chi\rangle) = a^*\Theta|\psi\rangle + b^*\Theta|\chi\rangle.$$

It is **antiunitary** if it is invertible and

$$\langle \Theta\psi | \Theta\chi \rangle = \langle \psi | \chi \rangle^* = \langle \chi | \psi \rangle. \quad (32.4)$$

This preserves norms and transition probabilities because the modulus squared of an inner product is unaffected by complex conjugation. A symmetry acts on the entire state space; we are not attempting to superpose a linear and an antilinear rule as if together they formed one linear quantum operation.

Choose an orthonormal basis and let K conjugate all coefficients in that basis. Then K is antiunitary and $K^2 = \mathbf{1}$. For any antiunitary Θ , the composition $W = \Theta K$ is linear and unitary, so

$$\Theta = WK.$$

This representation is basis dependent, but it makes one crucial rule transparent:

$$\Theta i = -i\Theta.$$

One must conjugate scalar coefficients when moving them through Θ ; treating it like an ordinary unitary matrix gives incorrect signs in time-evolution equations.

32.4 Why time reversal is antiunitary

Let $|\psi(t)\rangle$ solve the Schrödinger equation and define the candidate time-reversed solution

$$|\psi_T(t)\rangle = \Theta|\psi(-t)\rangle,$$

where Θ is time independent and antiunitary. Write $s = -t$. Since $\partial_s|\psi\rangle = -iH(s)|\psi\rangle/\hbar$, antilinearity gives

$$\begin{aligned}\frac{d}{dt}|\psi_T(t)\rangle &= -\Theta \frac{d}{ds}|\psi(s)\rangle \Big|_{s=-t} \\ &= -\frac{i}{\hbar}\Theta H(-t)|\psi(-t)\rangle.\end{aligned}$$

Multiplying by $i\hbar$ therefore yields

$$i\hbar \frac{d}{dt}|\psi_T(t)\rangle = \Theta H(-t)\Theta^{-1}|\psi_T(t)\rangle.$$

The reversed state obeys the original driven problem precisely when

$$H(t) = \Theta H(-t)\Theta^{-1}. \quad (32.5)$$

For a time-independent Hamiltonian this reduces to $\Theta H\Theta^{-1} = H$, or $\Theta H = H\Theta$. For a driven system the reversal must also transform the time dependence of the driving protocol as indicated by the $-t$ argument.

If instead we used $\Theta|\psi(t)\rangle$ without reversing time, an identical calculation would produce the Hamiltonian $-\Theta H(t)\Theta^{-1}$. Requiring this to be the original H would force $\Theta H = -H\Theta$ in the stationary case. An energy E would then have a partner $-E$. Most physical spectra are bounded below and unbounded above, so such a symmetry is impossible for them. The sign change of time and the conjugation of i work together to avoid this problem in genuine time reversal.

Example 32.2 (Spin-one-half time reversal). In the S_z basis choose

$$\Theta = -i\sigma_y K = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} K.$$

It sends $|\uparrow\rangle$ to $|\downarrow\rangle$ and $|\downarrow\rangle$ to $-|\uparrow\rangle$. Because $K\sigma_x K = \sigma_x$, $K\sigma_y K = -\sigma_y$, and $K\sigma_z K = \sigma_z$, conjugation by the remaining matrix gives

$$\Theta S_i \Theta^{-1} = -S_i \quad (i = x, y, z).$$

Angular momentum reverses, as expected when velocities reverse. Also $-i\sigma_y$ is real and squares to $-\mathbf{1}$, so $\Theta^2 = -\mathbf{1}$ on this doublet. A scalar change in the phase of Θ does not change this square: the second application complex-conjugates that phase.

For a fixed magnetic field, $H = -\gamma \mathbf{B} \cdot \mathbf{S}$ obeys $\Theta H(\mathbf{B})\Theta^{-1} = H(-\mathbf{B})$. Thus a nonzero fixed external field generally breaks time-reversal symmetry; reversing the whole physical experiment also reverses the magnetic field.

These antiunitary symmetries preserve the probability structure just as unitary symmetries do, but their antilinearity requires separate care. They extend, rather than replace, the unitary symmetry and conserved charge framework developed above.

Part V

Quantum Information and Entanglement

Chapter 5

33 Uncertainty, superposition, and quantum information

A classical state specifies a definite point in phase space. A probability distribution on phase space can describe incomplete knowledge of that point. A quantum state plays a different role: even a completely specified state generally gives a probability distribution for the outcomes of a measurement. Moreover, different observables need not admit simultaneous sharp values. We begin by making this statement quantitative, and then distinguish this intrinsic uncertainty from uncertainty about how a system was prepared.

33.1 Fluctuations and variances

For a normalized state $|\psi\rangle$ and a Hermitian observable A , define

$$\langle A \rangle = \langle \psi | A | \psi \rangle, \quad \Delta A = A - \langle A \rangle \mathbf{1}. \quad (33.1)$$

The expectation of the fluctuation operator vanishes. Its mean square is

$$\sigma_A^2 = \langle (\Delta A)^2 \rangle \quad (33.2)$$

$$= \langle A^2 - 2\langle A \rangle A + \langle A \rangle^2 \mathbf{1} \rangle = \langle A^2 \rangle - \langle A \rangle^2. \quad (33.3)$$

Because ΔA is Hermitian,

$$\sigma_A^2 = \|\Delta A |\psi\rangle\|^2 \geq 0. \quad (33.4)$$

Thus $\sigma_A = 0$ precisely when $A|\psi\rangle = \langle A \rangle |\psi\rangle$: a sharp observable corresponds to an eigenstate. For an unnormalized ket, every expectation is divided by $\langle \psi | \psi \rangle$. Below, states are normalized unless stated otherwise. For unbounded operators, the states under discussion are assumed to lie in the domains of the products that occur and to have finite variances.

33.2 Cauchy–Schwarz and the uncertainty relation

Lemma 33.1 (Cauchy–Schwarz inequality). *For any vectors $|u\rangle, |v\rangle$ in an inner-product space,*

$$|\langle u | v \rangle|^2 \leq \langle u | u \rangle \langle v | v \rangle. \quad (33.5)$$

For nonzero vectors, equality holds if and only if they are proportional.

Proof. If $|v\rangle = 0$, the statement is immediate. Otherwise positivity gives

$$\begin{aligned} 0 &\leq \| |u\rangle + \lambda |v\rangle \|^2 \\ &= \langle u | u \rangle + \lambda \langle u | v \rangle + \lambda^* \langle v | u \rangle + |\lambda|^2 \langle v | v \rangle. \end{aligned}$$

Choose $\lambda = -\langle v | u \rangle / \langle v | v \rangle$. Substitution yields

$$0 \leq \langle u | u \rangle - \frac{|\langle v | u \rangle|^2}{\langle v | v \rangle},$$

which proves the inequality. Equality means that the chosen vector $|u\rangle + \lambda |v\rangle$ is zero, giving the stated proportionality condition. $\square$

Apply this result to $|u\rangle = \Delta A |\psi\rangle$ and $|v\rangle = \Delta B |\psi\rangle$. Then

$$\sigma_A^2 \sigma_B^2 \geq |\langle \Delta A \Delta B \rangle|^2. \quad (33.6)$$

The product separates into a commutator and an anticommutator:

$$\Delta A \Delta B = \frac{1}{2} [A, B] + \frac{1}{2} \{ \Delta A, \Delta B \}. \quad (33.7)$$

The first term is anti-Hermitian, so its expectation is purely imaginary; the second is Hermitian, so its expectation is real. Therefore the absolute square has no cross term. Define the real, symmetrized covariance

$$C_{AB} = \frac{1}{2} \langle \{ \Delta A, \Delta B \} \rangle. \quad (33.8)$$

We obtain the stronger Schrödinger–Robertson inequality

$$\boxed{\sigma_A^2 \sigma_B^2 \geq \frac{1}{4} |\langle [A, B] \rangle|^2 + C_{AB}^2.} \quad (33.9)$$

Discarding the nonnegative covariance term gives

$$\boxed{\sigma_A \sigma_B \geq \frac{1}{2} |\langle [A, B] \rangle|.} \quad (33.10)$$

The stronger bound is saturated when the two fluctuation vectors are proportional. Saturating the weaker bound additionally requires $C_{AB} = 0$, apart from the corresponding zero-variance special cases.

Example 33.2 (Position and momentum). Since $[x, p] = i\hbar 1$,

$$\sigma_x \sigma_p \geq \frac{\hbar}{2}.$$

This bound applies to the spread of outcome distributions in a state. It does not require modeling an apparatus that disturbs an earlier measurement.

Example 33.3 (Spin uncertainty). For spin $1/2$, $S_i = \hbar\sigma_i/2$ and $[S_x, S_y] = i\hbar S_z$. In $|+z\rangle$,

$$\langle S_x \rangle = \langle S_y \rangle = 0, \quad \sigma_{S_x}^2 = \sigma_{S_y}^2 = \frac{\hbar^2}{4}.$$

Consequently

$$\sigma_{S_x} \sigma_{S_y} = \frac{\hbar^2}{4} = \frac{\hbar}{2} |\langle S_z \rangle|,$$

so the Robertson bound is saturated. In another state the expectation of S_z can vanish even though S_x and S_y do not commute. A vanishing lower bound in that state does not assert that both observables are sharp.

33.3 An alternative proof by positivity

The uncertainty relation can also be derived without first invoking Cauchy–Schwarz. This gives a useful worked version of the alternate-proof exercise. For every real t ,

$$0 \leq \|(\Delta A + it\Delta B)|\psi\rangle\|^2 \quad (33.11)$$

$$= \sigma_A^2 + t^2 \sigma_B^2 + it\langle[A, B]\rangle. \quad (33.12)$$

The coefficient $c = i\langle[A, B]\rangle$ is real. If $\sigma_B^2 > 0$, the quadratic has its minimum at $t = -c/(2\sigma_B^2)$. Positivity at this minimum implies

$$\sigma_A^2 - \frac{c^2}{4\sigma_B^2} \geq 0,$$

which is exactly (33.10). If $\sigma_B = 0$, positivity for both positive and negative t forces $c = 0$, and the bound again holds. Allowing a general complex coefficient in place of it recovers Cauchy–Schwarz and the covariance term in (33.9).

33.4 Superposition does not mean classical ignorance

Consider a molecule such as ammonia, with two localized configurations $|1\rangle$ and $|2\rangle$ corresponding approximately to the nitrogen atom on opposite sides of the plane of the hydrogen atoms. In a two-state description, a symmetric energy eigenstate can take the form

$$|+\rangle = \frac{|1\rangle + |2\rangle}{\sqrt{2}}. \quad (33.13)$$

A measurement that distinguishes the two configurations gives either result with probability $1/2$. This does not make $|+\rangle$ equivalent to preparing $|1\rangle$ on half the trials and $|2\rangle$ on the other half. The relative phase in a superposition affects measurements in another basis. The distinction will be explicit in the Stern–Gerlach examples below.

A general normalized qubit, after removing an overall phase, can be written

$$|\psi\rangle = \cos \frac{\theta}{2} |0\rangle + e^{i\phi} \sin \frac{\theta}{2} |1\rangle, \quad 0 \leq \theta \leq \pi, \quad 0 \leq \phi < 2\pi. \quad (33.14)$$

Its pure states form the Bloch sphere. The continuous pair (θ, ϕ) describes a continuum of possible preparations, but a single projective measurement on one qubit has at most

two distinguishable outcomes. One cannot determine an arbitrary unknown direction on the Bloch sphere from a single copy. Repeated measurements on many identically prepared copies, using different measurement axes, are needed for state tomography.

This distinction persists in higher dimensions. A pure state has complex amplitudes, but one measurement does not reveal those amplitudes. Perfect single-shot discrimination is possible for mutually orthogonal states; nonorthogonal states cannot be distinguished with certainty by a measurement. The number of possible mathematical state vectors therefore must not be confused with the number of reliably distinguishable messages carried by a single system.

34 Statistical ensembles and density operators

34.1 Three Stern–Gerlach preparations

Write $|0\rangle = | +z\rangle$ and $|1\rangle = | -z\rangle$. The spin eigenstates along x and y are

$$|\pm x\rangle = \frac{|0\rangle \pm |1\rangle}{\sqrt{2}}, \quad (34.1)$$

$$|\pm y\rangle = \frac{|0\rangle \pm i|1\rangle}{\sqrt{2}}. \quad (34.2)$$

Consider two preparations:

1. Preparation I: each atom is prepared in $|0\rangle$ with probability $1/2$ and in $|1\rangle$ with probability $1/2$.
2. Preparation II: every atom is prepared in $|+x\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$.

A Stern–Gerlach apparatus oriented along z gives equal upward and downward populations for both preparations.

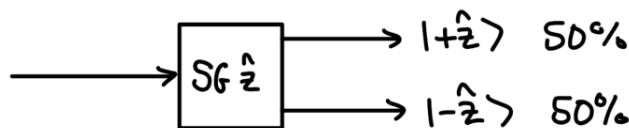


Figure 41: A Stern–Gerlach measurement along z cannot distinguish an equal mixture of z spins from a coherent $+x$ preparation. Original lecture diagram.

To distinguish them, measure S_x . For Preparation I, all four conditional probabilities are $1/2$:

$$P(\pm x \mid +z) = P(\pm x \mid -z) = \frac{1}{2}. \quad (34.3)$$

The law of total probability gives

$$P(+x) = P(+x \mid +z)P(+z) + P(+x \mid -z)P(-z) \quad (34.4)$$

$$= \frac{1}{2} \frac{1}{2} + \frac{1}{2} \frac{1}{2} = \frac{1}{2}, \quad (34.5)$$

and similarly $P(-x) = 1/2$. In Preparation II the input is already an eigenstate of S_x , so $P(+x) = 1$ and $P(-x) = 0$.

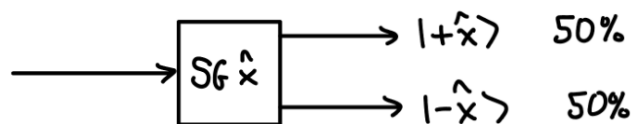


Figure 42: The equal incoherent mixture gives equal outcomes when measured along x . Original lecture diagram.

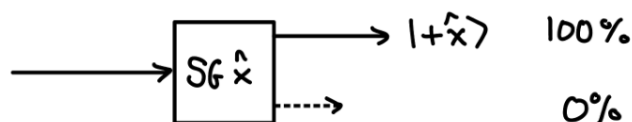


Figure 43: The coherent $+x$ state gives a definite $+x$ outcome. Original lecture diagram.

Example 34.1 (A relative phase is experimentally observable). For the state

$$|\psi_\phi\rangle = \frac{|0\rangle + e^{i\phi}|1\rangle}{\sqrt{2}},$$

a z measurement always gives probabilities $1/2, 1/2$. However,

$$\langle +x | \psi_\phi \rangle = \frac{1 + e^{i\phi}}{2}, \quad \langle -x | \psi_\phi \rangle = \frac{1 - e^{i\phi}}{2},$$

so

$$P(+x) = \cos^2 \frac{\phi}{2}, \quad P(-x) = \sin^2 \frac{\phi}{2}. \quad (34.6)$$

The phase is not an additional probability in the z basis: it controls interference between the two amplitudes. A y measurement additionally gives $P(\pm y) = (1 \pm \sin \phi)/2$, distinguishing phases with the same cosine.

Example 34.2 (A partially polarized beam). Prepare $|+z\rangle$ with probability f and $|-z\rangle$ with probability $1 - f$. Let the measurement axis be

$$\mathbf{n} = (\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta).$$

The $+\mathbf{n}$ eigenket is $\cos(\theta/2)|0\rangle + e^{i\phi} \sin(\theta/2)|1\rangle$, so

$$P(+\mathbf{n}) = f \cos^2 \frac{\theta}{2} + (1 - f) \sin^2 \frac{\theta}{2} \quad (34.7)$$

$$= \frac{1}{2} [1 + (2f - 1) \cos \theta], \quad (34.8)$$

$$P(-\mathbf{n}) = \frac{1}{2} [1 - (2f - 1) \cos \theta]. \quad (34.9)$$

For a specified axis, the outcomes are equally likely precisely when $(2f - 1) \cos \theta = 0$. They are equally likely for *every* axis precisely when $f = 1/2$. This equal mixture is an unpolarized beam.

34.2 The density operator of an ensemble

Suppose the normalized ket $|\psi_i\rangle$ is prepared with classical probability p_i , where $p_i \geq 0$ and $\sum_i p_i = 1$. Define

$$\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|. \quad (34.10)$$

The probability of a rank-one outcome $|a\rangle$ is

$$P(a) = \sum_i p_i |\langle a | \psi_i \rangle|^2 \quad (34.11)$$

$$= \sum_i p_i \langle a | \psi_i \rangle \langle \psi_i | a \rangle = \langle a | \rho | a \rangle. \quad (34.12)$$

Thus the ensemble enters measurement probabilities through the single operator ρ .

Theorem 34.3 (Characterization of density operators). *In a finite-dimensional Hilbert space, an operator is a density operator if and only if it is Hermitian, positive semidefinite, and has trace one:*

$$\rho^\dagger = \rho, \quad \langle \chi | \rho | \chi \rangle \geq 0 \quad \text{for every } |\chi\rangle, \quad \text{Tr } \rho = 1. \quad (34.13)$$

Proof. An ensemble operator is Hermitian because each $|\psi_i\rangle\langle\psi_i|$ is Hermitian and p_i is real. Positivity follows from

$$\langle \chi | \rho | \chi \rangle = \sum_i p_i |\langle \chi | \psi_i \rangle|^2 \geq 0.$$

Using an orthonormal basis $\{|a\rangle\}$,

$$\text{Tr}(|\psi\rangle\langle\psi|) = \sum_a \langle a | \psi \rangle \langle \psi | a \rangle = \langle \psi | \sum_a |a\rangle\langle a| \psi \rangle = 1.$$

Consequently $\text{Tr } \rho = \sum_i p_i = 1$.

Conversely, a Hermitian ρ has an orthonormal eigenbasis:

$$\rho = \sum_a r_a |a\rangle\langle a|.$$

Positivity applied to each $|a\rangle$ gives $r_a \geq 0$, and unit trace gives $\sum_a r_a = 1$. This is an ensemble decomposition, proving sufficiency. $\square$

The theorem supplies both a test for proposed density matrices and a way to prepare any allowed density matrix. For infinite-dimensional systems, density operators are additionally required to be trace class; the finite-dimensional arguments below avoid those functional-analytic issues.

34.3 Coherence and nonunique ensemble decompositions

In the ordered basis $(|0\rangle, |1\rangle)$, the two beam preparations are

$$\rho_{\text{I}} = \frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|1\rangle\langle 1| = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad (34.14)$$

$$\rho_{\text{II}} = |+x\rangle\langle +x| = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}. \quad (34.15)$$

The diagonal elements give the z -basis probabilities. The off-diagonal entries retain the relative-phase information responsible for the different x -basis statistics. Coherence is always discussed relative to a chosen basis: a pure-state matrix can be diagonal in its own eigenbasis.

A third preparation, an equal classical mixture of $|+x\rangle$ and $|-x\rangle$, gives

$$\rho_{\text{III}} = \frac{1}{2}|+x\rangle\langle +x| + \frac{1}{2}|-x\rangle\langle -x| \quad (34.16)$$

$$= \frac{1}{4} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} + \frac{1}{4} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} = \frac{1}{2} \mathbf{1}. \quad (34.17)$$

Thus Preparations I and III have different preparation stories but the same density operator. No measurement on the spin alone distinguishes them. If a separate classical register records which preparation was used, measurements involving that register describe a larger system and can reveal information absent from the spin's density matrix.

34.4 Pure states, mixed states, and Bloch vectors

Definition 34.4. A state is **pure** when $\rho = |\psi\rangle\langle\psi|$ for a normalized ket. A density operator that is not pure is **mixed**.

The ket of a pure state is only determined up to an overall phase, since $e^{i\alpha}|\psi\rangle\langle\psi|e^{-i\alpha} = |\psi\rangle\langle\psi|$. The following conditions are equivalent:

$$\rho \text{ is pure} \iff \text{rank } \rho = 1 \iff \rho^2 = \rho \iff \text{Tr } \rho^2 = 1. \quad (34.18)$$

To see this, let r_a be the eigenvalues. They obey $0 \leq r_a \leq 1$ and $\sum_a r_a = 1$. Hence $\sum_a r_a^2 \leq 1$, with equality precisely when each r_a is zero or one. Exactly one eigenvalue must then equal one.

In dimension d , the **maximally mixed state** is

$$\rho_* = \frac{1}{d} \mathbf{1}. \quad (34.19)$$

It is invariant under every unitary basis change, gives equal probabilities for every orthonormal-basis measurement, and has the smallest purity:

$$1 = \left(\sum_a r_a\right)^2 \leq d \sum_a r_a^2 \implies \text{Tr } \rho^2 \geq \frac{1}{d}.$$

Equality holds exactly when all eigenvalues equal $1/d$.

For a qubit, every Hermitian trace-one matrix has the expansion

$$\rho = \frac{1}{2}(\mathbf{1} + \mathbf{r} \cdot \boldsymbol{\sigma}) = \frac{1}{2} \begin{pmatrix} 1 + r_z & r_x - ir_y \\ r_x + ir_y & 1 - r_z \end{pmatrix}, \quad (34.20)$$

where $r_i = \text{Tr}(\rho\sigma_i)$ is real. Because $(\mathbf{r} \cdot \boldsymbol{\sigma})^2 = |\mathbf{r}|^2 \mathbf{1}$, the eigenvalues are $(1 \pm |\mathbf{r}|)/2$. Positivity is therefore equivalent to $|\mathbf{r}| \leq 1$. The Bloch sphere is the surface of the Bloch ball: pure states have $|\mathbf{r}| = 1$, mixed states have $|\mathbf{r}| < 1$, and the unpolarized state is its center. Furthermore,

$$\text{Tr} \rho^2 = \frac{1 + |\mathbf{r}|^2}{2}, \quad P(\pm \mathbf{n}) = \frac{1 \pm \mathbf{r} \cdot \mathbf{n}}{2}. \quad (34.21)$$

Measuring $\sigma_x, \sigma_y, \sigma_z$ on separate subensembles determines r_x, r_y, r_z and hence reconstructs ρ .

34.5 The measurement rules in density-operator form

Let the observable be $A = \sum_{\lambda} \lambda \Pi_{\lambda}$, where the orthogonal projectors satisfy

$$\Pi_{\lambda}^{\dagger} = \Pi_{\lambda}, \quad \Pi_{\lambda} \Pi_{\mu} = \delta_{\lambda\mu} \Pi_{\lambda}, \quad \sum_{\lambda} \Pi_{\lambda} = \mathbf{1}.$$

A projector may have rank greater than one when the outcome is degenerate.

Theorem 34.5 (Born rule and selective state update). *The probability of outcome λ is*

$$p_{\lambda} = \text{Tr}(\rho \Pi_{\lambda}). \quad (34.22)$$

For an ideal projective measurement, conditional on an outcome of nonzero probability, the resulting state is

$$\rho_{\lambda} = \frac{\Pi_{\lambda} \rho \Pi_{\lambda}}{\text{Tr}(\rho \Pi_{\lambda})}. \quad (34.23)$$

Proof. For ensemble member i , write $q_{\lambda|i} = \langle \psi_i | \Pi_{\lambda} | \psi_i \rangle$. Then

$$p_{\lambda} = \sum_i p_i q_{\lambda|i} = \text{Tr} \left(\sum_i p_i |\psi_i\rangle \langle \psi_i| \Pi_{\lambda} \right) = \text{Tr}(\rho \Pi_{\lambda}).$$

When $q_{\lambda|i} > 0$, that ensemble member becomes $\Pi_{\lambda} |\psi_i\rangle / \sqrt{q_{\lambda|i}}$. Bayes' rule gives its conditional ensemble weight $p_{i|\lambda} = p_i q_{\lambda|i} / p_{\lambda}$. Therefore

$$\begin{aligned} \rho_{\lambda} &= \sum_{i: q_{\lambda|i} > 0} \frac{p_i q_{\lambda|i}}{p_{\lambda}} \frac{\Pi_{\lambda} |\psi_i\rangle \langle \psi_i| \Pi_{\lambda}}{q_{\lambda|i}} \\ &= \frac{\Pi_{\lambda} \rho \Pi_{\lambda}}{p_{\lambda}}. \end{aligned}$$

Terms with $q_{\lambda|i} = 0$ contribute zero. The denominator normalizes the result because $\text{Tr}(\Pi_{\lambda} \rho \Pi_{\lambda}) = \text{Tr}(\rho \Pi_{\lambda})$. $\square$

For any observable, the expectation is consequently

$$\langle A \rangle = \sum_{\lambda} \lambda p_{\lambda} = \text{Tr} \left(\rho \sum_{\lambda} \lambda \Pi_{\lambda} \right) = \boxed{\text{Tr}(\rho A)}. \quad (34.24)$$

All moments follow by replacing A with A^n .

34.6 A measurement whose result is forgotten

If the measurement is performed but its outcome is not retained, average over the conditional states:

$$\boxed{\rho' = \sum_{\lambda} p_{\lambda} \rho_{\lambda} = \sum_{\lambda} \Pi_{\lambda} \rho \Pi_{\lambda}}. \quad (34.25)$$

This is distinct from doing nothing. Before measurement,

$$\rho = \sum_{\lambda, \mu} \Pi_{\lambda} \rho \Pi_{\mu}.$$

The nonselective measurement removes the blocks with $\lambda \neq \mu$, while leaving coherence inside a degenerate eigenspace intact. It leaves ρ unchanged if and only if ρ is block diagonal in these eigenspaces, equivalently $[\rho, \Pi_{\lambda}] = 0$ for every λ .

For example, forgetting a z -measurement outcome on $|+x\rangle$ gives

$$|+x\rangle\langle+x| \mapsto \frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|1\rangle\langle 1|. \quad (34.26)$$

A subsequent x measurement now gives both outcomes equally often, whereas it would have given $+x$ with certainty without the intervening measurement. A measurement can therefore convert a pure reduced description to a mixed one when its classical record is discarded.

The density operator is sufficient even for a sequence of operations. For a first projector Π_{λ} , a unitary U , and a later projector Q_{μ} , the joint probability is

$$P(\lambda, \mu) = \text{Tr}(Q_{\mu} U \Pi_{\lambda} \rho \Pi_{\lambda} U^{\dagger} Q_{\mu}). \quad (34.27)$$

The normalization in the intermediate conditional state cancels its outcome probability. Thus different ensembles producing the same ρ remain indistinguishable under such experiments on the system.

34.7 Unitary evolution and the von Neumann equation

If every ensemble member evolves by $|\psi'_i\rangle = U|\psi_i\rangle$, with fixed preparation probabilities, then

$$\rho' = \sum_i p_i U |\psi_i\rangle\langle\psi_i| U^{\dagger} = \boxed{U \rho U^{\dagger}}. \quad (34.28)$$

This proves the unitary-evolution rule. It preserves Hermiticity, positivity, trace, eigenvalues, rank, and purity.

To obtain a differential equation, use $i\hbar\partial_t|\psi_i\rangle = H|\psi_i\rangle$ and its adjoint:

$$i\hbar\dot{\rho} = \sum_i p_i (H|\psi_i\rangle\langle\psi_i| - |\psi_i\rangle\langle\psi_i|H) \quad (34.29)$$

$$= [H, \rho]. \quad (34.30)$$

Equivalently, $\dot{\rho} + (1/i\hbar)[\rho, H] = 0$. These equations hold for a time-dependent Hamiltonian as well.

Let $\rho_S(t) = U(t)\rho_S(0)U(t)^\dagger$. The Heisenberg-picture density operator is

$$\rho_H(t) = U(t)^\dagger \rho_S(t) U(t) = \rho_S(0). \quad (34.31)$$

The cancellation can be checked directly. Since $\dot{U} = -iHU/\hbar$ and $\dot{U}^\dagger = iU^\dagger H/\hbar$,

$$\dot{\rho}_H = \frac{i}{\hbar}U^\dagger H\rho_S U - \frac{i}{\hbar}U^\dagger (H\rho_S - \rho_S H)U - \frac{i}{\hbar}U^\dagger \rho_S H U = 0.$$

Expectations are picture independent:

$$\text{Tr}(\rho_S A_S) = \text{Tr}(\rho_H A_H), \quad A_H = U^\dagger A_S U. \quad (34.32)$$

The density operator describes the state; it is not an extra fixed observable whose one-shot measurement reveals an unknown state. Its matrix elements are reconstructed statistically from a sufficient set of observables. A specified density matrix is Hermitian and can of course also be used mathematically as an observable, but that is a different use of the same operator.

34.8 Comparison with classical Liouville evolution

For a classical phase-space probability density $f(q, p, t)$, conservation of probability under Hamiltonian flow gives

$$\frac{\partial f}{\partial t} + \{f, H\}_{\text{PB}} = 0, \quad \{f, H\}_{\text{PB}} = \sum_a \left(\frac{\partial f}{\partial q_a} \frac{\partial H}{\partial p_a} - \frac{\partial f}{\partial p_a} \frac{\partial H}{\partial q_a} \right). \quad (34.33)$$

The formal correspondence $\{ , \}_{\text{PB}} \leftrightarrow [,]/(i\hbar)$ relates this equation to (34.30). In fixed phase-space coordinates the density evolves, just as ρ_S evolves. Along the Hamiltonian trajectories it is constant, analogous to moving the evolution into the operators while keeping ρ_H fixed. This is a structural comparison, not a prescription that maps every classical probability density to a quantum density operator.

There is also an important difference in interpretation. A classical mixture can represent ignorance of a definite phase-space point. A quantum density operator can describe either an ensemble of preparations or a subsystem of an entangled pure state, even when the full state is completely known. To understand the latter possibility we must first describe composite systems.

35 Composite systems and tensor products

35.1 The Hilbert space of two subsystems

If distinguishable subsystems A and B have Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$, the composite system has Hilbert space

$$\boxed{\mathcal{H}_{AB} = \mathcal{H}_A \otimes \mathcal{H}_B.} \quad (35.1)$$

This is the composition postulate. A product ket $|\psi\rangle_A \otimes |\chi\rangle_B$ specifies the two subsystem states independently. A general composite ket is a linear combination of product kets and need not itself be a product.

The tensor product is bilinear:

$$(a|\psi\rangle + b|\phi\rangle) \otimes |\chi\rangle = a|\psi\rangle \otimes |\chi\rangle + b|\phi\rangle \otimes |\chi\rangle, \quad (35.2)$$

$$|\psi\rangle \otimes (a|\chi\rangle + b|\eta\rangle) = a|\psi\rangle \otimes |\chi\rangle + b|\psi\rangle \otimes |\eta\rangle. \quad (35.3)$$

A scalar can be moved between factors: $(a|\psi\rangle) \otimes |\chi\rangle = |\psi\rangle \otimes (a|\chi\rangle)$. The inner product on product vectors is

$$(\langle\psi|_A \otimes \langle\chi|_B)(|\phi\rangle_A \otimes |\eta\rangle_B) = \langle\psi|_A \langle\chi|_B \phi\rangle_A \eta\rangle_B, \quad (35.4)$$

and extends by sesquilinearity to sums. In particular, products of normalized kets are normalized.

If $\{|a\rangle_A\}_{a=1}^{d_A}$ and $\{|b\rangle_B\}_{b=1}^{d_B}$ are orthonormal bases, their products form an orthonormal basis:

$$|a, b\rangle = |a\rangle_A \otimes |b\rangle_B, \quad \langle a, b | a', b'\rangle = \delta_{aa'} \delta_{bb'}. \quad (35.5)$$

It follows that

$$\dim \mathcal{H}_{AB} = d_A d_B, \quad |\Psi\rangle = \sum_{a,b} c_{ab} |a\rangle_A \otimes |b\rangle_B, \quad \sum_{a,b} |c_{ab}|^2 = 1. \quad (35.6)$$

The tensor-product symbol is often suppressed, but subsystem labels remain important. The adjoint in a fixed factor order is

$$(|\psi\rangle_A \otimes |\chi\rangle_B)^\dagger = \langle\psi|_A \otimes \langle\chi|_B.$$

Changing $A \otimes B$ to $B \otimes A$ requires the specified swap map $|\psi\rangle_A \otimes |\chi\rangle_B \mapsto |\chi\rangle_B \otimes |\psi\rangle_A$; it is not an equality of vectors in a fixed ordered tensor product.

For spatially separated distinguishable systems, the tensor-product description expresses their independent local degrees of freedom. Identical particles and quantum fields require additional care in identifying subsystems; the spin and detector models used here have explicitly distinguished factors.

35.2 Operators on a composite system

The tensor product of operators is defined on product kets by

$$(X_A \otimes Y_B)(|\psi\rangle_A \otimes |\chi\rangle_B) = (X_A |\psi\rangle_A) \otimes (Y_B |\chi\rangle_B), \quad (35.7)$$

and then extended linearly. It obeys

$$(X \otimes Y)(X' \otimes Y') = (XX') \otimes (YY'), \quad (35.8)$$

$$(X \otimes Y)^\dagger = X^\dagger \otimes Y^\dagger. \quad (35.9)$$

Every operator on a finite-dimensional composite space can be expanded as a sum of tensor products, since

$$L = \sum_{a,b,a',b'} L_{ab,a'b'} (|a\rangle\langle a'|) \otimes (|b\rangle\langle b'|). \quad (35.10)$$

It is generally a sum of products rather than one product.

An observable acting only on A is represented by $X_A \otimes \mathbf{1}_B$; an observable acting only on B is $\mathbf{1}_A \otimes Y_B$. They commute:

$$[X_A \otimes \mathbf{1}_B, \mathbf{1}_A \otimes Y_B] = 0. \quad (35.11)$$

If $X_A|\lambda\rangle = \lambda|\lambda\rangle$ and $Y_B|\xi\rangle = \xi|\xi\rangle$, then $|\lambda\rangle_A \otimes |\xi\rangle_B$ is a simultaneous eigenket. For example,

$$\sum_{\lambda,\xi} \lambda|\lambda,\xi\rangle\langle\lambda,\xi| = \left(\sum_{\lambda} \lambda|\lambda\rangle\langle\lambda| \right) \otimes \left(\sum_{\xi} |\xi\rangle\langle\xi| \right) \quad (35.12)$$

$$= X_A \otimes \mathbf{1}_B. \quad (35.13)$$

Additional degeneracy indices can be included in the same way. Commuting local observables permit joint measurements, but their outcomes can be correlated through the joint state. Commutativity does not require the probabilities to factorize.

35.3 Two spins and their total spin

For two spin-1/2 systems use the ordered basis

$$|00\rangle, \quad |01\rangle, \quad |10\rangle, \quad |11\rangle. \quad (35.14)$$

The first spin's z component is

$$S_{Az} = \frac{\hbar}{2} (|00\rangle\langle 00| + |01\rangle\langle 01| - |10\rangle\langle 10| - |11\rangle\langle 11|) \quad (35.15)$$

$$= \frac{\hbar}{2} (|0\rangle\langle 0| - |1\rangle\langle 1|) \otimes (|0\rangle\langle 0| + |1\rangle\langle 1|) = S_z \otimes \mathbf{1}. \quad (35.16)$$

Similarly,

$$S_{Bz} = \frac{\hbar}{2} (|00\rangle\langle 00| - |01\rangle\langle 01| + |10\rangle\langle 10| - |11\rangle\langle 11|) = \mathbf{1} \otimes S_z. \quad (35.17)$$

Their matrix representations are

$$S_{Az} = \frac{\hbar}{2} \text{diag}(1, 1, -1, -1), \quad S_{Bz} = \frac{\hbar}{2} \text{diag}(1, -1, 1, -1). \quad (35.18)$$

The total angular momentum is the vector operator

$$\mathbf{S} = \mathbf{S}_A + \mathbf{S}_B, \quad S_z = \hbar \text{diag}(1, 0, 0, -1). \quad (35.19)$$

Its z eigenvalues are $\hbar, 0, -\hbar$, with a two-dimensional zero-eigenspace. Merely knowing $S_z = 0$ does not determine the joint state: $|01\rangle$, $|10\rangle$, and all their normalized superpositions are allowed.

35.4 The singlet state and anticorrelations

The normalized antisymmetric state is

$$|\Psi^-\rangle = \frac{|01\rangle - |10\rangle}{\sqrt{2}}. \quad (35.20)$$

It has total spin zero. Indeed, $S_z|\Psi^-\rangle = 0$, while

$$\begin{aligned} S_x|01\rangle &= \frac{\hbar}{2}(|11\rangle + |00\rangle), & S_x|10\rangle &= \frac{\hbar}{2}(|00\rangle + |11\rangle), \\ S_y|01\rangle &= \frac{i\hbar}{2}(|11\rangle - |00\rangle), & S_y|10\rangle &= \frac{i\hbar}{2}(|11\rangle - |00\rangle). \end{aligned}$$

Subtracting the members of each pair gives $S_x|\Psi^-\rangle = S_y|\Psi^-\rangle = 0$, and hence $\mathbf{S}^2|\Psi^-\rangle = 0$.

The same anticorrelation is visible in another spin basis. Expanding the x states,

$$|+x, -x\rangle = \frac{1}{2}(|00\rangle - |01\rangle + |10\rangle - |11\rangle), \quad (35.21)$$

$$|-x, +x\rangle = \frac{1}{2}(|00\rangle + |01\rangle - |10\rangle - |11\rangle). \quad (35.22)$$

Their difference is $-|01\rangle + |10\rangle$, so

$$|\Psi^-\rangle = -\frac{|+x, -x\rangle - |-x, +x\rangle}{\sqrt{2}}. \quad (35.23)$$

The overall minus sign has no observable effect. For the y basis,

$$|+y, -y\rangle = \frac{1}{2}(|00\rangle - i|01\rangle + i|10\rangle + |11\rangle), \quad (35.24)$$

$$|-y, +y\rangle = \frac{1}{2}(|00\rangle + i|01\rangle - i|10\rangle + |11\rangle), \quad (35.25)$$

so

$$|\Psi^-\rangle = \frac{i}{\sqrt{2}}(|+y, -y\rangle - |-y, +y\rangle). \quad (35.26)$$

The common phase again does not matter.

More generally, for a two-by-two matrix U ,

$$(U \otimes U)(|01\rangle - |10\rangle) = (\det U)(|01\rangle - |10\rangle). \quad (35.27)$$

This follows by expanding both products: the $|00\rangle$ and $|11\rangle$ coefficients cancel, while the $|01\rangle$ coefficient is the determinant and the $|10\rangle$ coefficient its negative. A spin rotation has $U \in \text{SU}(2)$, so the singlet is invariant under a common rotation. Consequently measurements along any common axis give opposite results with certainty, and either observer's two possible results occur with probability $1/2$.

35.5 Product states and entangled states

Definition 35.1. A pure bipartite state is a **product state** when $|\Psi\rangle = |\psi\rangle_A \otimes |\chi\rangle_B$. It is **entangled** when no such factorization exists.

For product states, the coefficient matrix in (35.6) has the form $c_{ab} = u_a v_b$. Thus a nonzero pure state is a product precisely when its coefficient matrix has rank one. For the singlet,

$$(c_{ab}) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},$$

which has rank two. Equivalently, attempting $(a|0\rangle + b|1\rangle) \otimes (c|0\rangle + d|1\rangle)$ would require $ac = bd = 0$ but ad and bc both nonzero, an impossibility.

As another example, suppose orthonormal pairs $|\lambda_1\rangle, |\lambda_2\rangle$ and $|\xi_1\rangle, |\xi_2\rangle$ are associated with local observables. The state

$$|\Psi\rangle = \frac{|\lambda_1\rangle_A |\xi_1\rangle_B + |\lambda_2\rangle_A |\xi_2\rangle_B}{\sqrt{2}} \quad (35.28)$$

is entangled. A rank-one measurement on A in the λ basis gives either outcome with probability $1/2$. Conditional on outcome λ_i , the joint state becomes $|\lambda_i\rangle_A |\xi_i\rangle_B$, and a later ξ measurement on B has a definite corresponding result.

In fact, a rank-one selective measurement on A always leaves a pure joint state in product form. The unnormalized branch is

$$(|a\rangle\langle a| \otimes \mathbf{1})|\Psi\rangle = |a\rangle_A \otimes [(\langle a|_A \otimes \mathbf{1}_B)|\Psi\rangle]. \quad (35.29)$$

A degenerate projector need not remove all entanglement. Also, when the outcome is forgotten, the mixture of product branches can retain classical correlations. These are distinct statements from the disappearance of entanglement in each rank-one conditional branch.

For mixed states, a product state has the form $\rho_A \otimes \rho_B$. The broader class of **separable states** consists of convex mixtures

$$\rho_{AB} = \sum_i p_i \rho_A^{(i)} \otimes \rho_B^{(i)}. \quad (35.30)$$

Such states can have classical correlations without entanglement. A mixed state is entangled if no decomposition of this form exists. The distinction between a product state and a separable correlated state will also matter for entropy.

36 Reduced states, the partial trace, and no signalling

36.1 Why a subsystem needs a density operator

The singlet is a completely specified pure state of two spins, yet either spin by itself gives equal probabilities along every axis. There is no pure qubit state with this property: every pure qubit has one axis along which its spin is certain. Its local state must instead be $\mathbf{1}/2$. This mixedness reflects entanglement with the other spin, not an unspecified choice of a pure state for the pair.

We seek an operation taking a joint density matrix ρ_{AB} to a density matrix ρ_A that reproduces every measurement on A . If Π_λ is a local measurement projector, its full-space version is $\Pi_\lambda \otimes \mathbf{1}_B$, so the required relation is

$$\mathrm{Tr}_{AB}[\rho_{AB}(\Pi_\lambda \otimes \mathbf{1}_B)] = \mathrm{Tr}_A(\rho_A \Pi_\lambda). \quad (36.1)$$

The partial trace constructs precisely this ρ_A .

36.2 Definition and basis formula

First note the product-trace identity:

$$\mathrm{Tr}_{AB}(X_A \otimes Y_B) = \sum_{a,b} \langle a|X|a\rangle \langle b|Y|b\rangle \quad (36.2)$$

$$= (\mathrm{Tr}_A X_A)(\mathrm{Tr}_B Y_B). \quad (36.3)$$

For a product operator, define

$$\mathrm{Tr}_B(X_A \otimes Y_B) = X_A \mathrm{Tr}_B Y_B, \quad (36.4)$$

and extend linearly to arbitrary operators. A basis-independent meaning is supplied by the equivalent formula

$$\boxed{\mathrm{Tr}_B L = \sum_b (\mathbf{1}_A \otimes \langle b|_B) L (\mathbf{1}_A \otimes |b\rangle_B).} \quad (36.5)$$

Here $\mathbf{1}_A \otimes \langle b|_B$ is a partial contraction. For example,

$$(\mathbf{1}_A \otimes \langle b|_B)(|\psi\rangle_A \otimes |\chi\rangle_B) = |\psi\rangle_A \langle b | \chi\rangle_B. \quad (36.6)$$

Applying the right-hand side of (36.5) to a product operator gives $X_A \sum_b \langle b|Y|b\rangle = X_A \mathrm{Tr} Y$, so the two definitions agree on a spanning set of operators.

To verify independence of the chosen basis, let $|b\rangle = \sum_c U_{cb}|c\rangle$ for a unitary matrix U . Then the transformed contraction sum has coefficients

$$\sum_b U_{cb}^* U_{db} = \delta_{cd},$$

and reduces to the original contraction sum. No basis choice is part of the physical reduced state.

In components, the formula is particularly simple:

$$(\text{Tr}_B L)_{aa'} = \sum_b L_{ab, a'b}. \quad (36.7)$$

The indices belonging to B are identified and summed. The A indices remain as the row and column indices of an operator on $\mathcal{H}_A$.

36.3 The local-expectation identity

For $L = \sum_i X_i \otimes Y_i$ and an operator M on A ,

$$\text{Tr}_{AB} [L(M \otimes \mathbf{1}_B)] = \sum_i \text{Tr}_{AB} (X_i M \otimes Y_i) \quad (36.8)$$

$$= \sum_i \text{Tr}_A (X_i M) \text{Tr}_B Y_i \quad (36.9)$$

$$= \text{Tr}_A \left[\left(\sum_i X_i \text{Tr}_B Y_i \right) M \right] \quad (36.10)$$

$$= \boxed{\text{Tr}_A [(\text{Tr}_B L) M]}. \quad (36.11)$$

Thus define the **reduced density operator**

$$\boxed{\rho_A = \text{Tr}_B \rho_{AB}}. \quad (36.12)$$

Equation (36.11) proves both the local Born rule and the local expectation rule

$$P(\lambda) = \text{Tr}_A (\rho_A \Pi_\lambda), \quad \langle M_A \rangle = \text{Tr}_A (\rho_A M_A). \quad (36.13)$$

The reduced state contains all information accessible by measurements on A alone, but not generally the correlations between A and B .

Proposition 36.1. *The partial trace of a density operator is a density operator.*

Proof. Hermiticity follows by taking the adjoint of each term in (36.5). For any $|\psi\rangle_A$,

$$\langle \psi | \rho_A | \psi \rangle = \sum_b \langle \psi, b | \rho_{AB} | \psi, b \rangle \geq 0,$$

so ρ_A is positive semidefinite. Finally,

$$\text{Tr}_A \rho_A = \sum_{a,b} \langle a, b | \rho_{AB} | a, b \rangle = \text{Tr}_{AB} \rho_{AB} = 1.$$

All three density-operator conditions are satisfied. $\square$

For a product state,

$$\text{Tr}_B (\rho_A \otimes \rho_B) = \rho_A \text{Tr} \rho_B = \rho_A. \quad (36.14)$$

The same operation on a correlated state averages over the inaccessible subsystem without selecting a measurement outcome on it.

36.4 Worked partial traces

For a two-qubit matrix in the ordering $(00, 01, 10, 11)$, write

$$\rho_{AB} = \begin{pmatrix} r_{00,00} & r_{00,01} & r_{00,10} & r_{00,11} \\ r_{01,00} & r_{01,01} & r_{01,10} & r_{01,11} \\ r_{10,00} & r_{10,01} & r_{10,10} & r_{10,11} \\ r_{11,00} & r_{11,01} & r_{11,10} & r_{11,11} \end{pmatrix}.$$

Tracing over B gives

$$\rho_A = \begin{pmatrix} r_{00,00} + r_{01,01} & r_{00,10} + r_{01,11} \\ r_{10,00} + r_{11,01} & r_{10,10} + r_{11,11} \end{pmatrix}. \quad (36.15)$$

Tracing over A instead gives

$$\rho_B = \begin{pmatrix} r_{00,00} + r_{10,10} & r_{00,01} + r_{10,11} \\ r_{01,00} + r_{11,10} & r_{01,01} + r_{11,11} \end{pmatrix}. \quad (36.16)$$

The two formulas sum different index pairs; they should not be interchanged.

For the singlet, expand the projector:

$$\rho_{AB} = |\Psi^-\rangle\langle\Psi^-| = \frac{1}{2}(|01\rangle\langle 01| + |10\rangle\langle 10| \quad (36.17)$$

$$- |01\rangle\langle 10| - |10\rangle\langle 01|). \quad (36.18)$$

The first diagonal term is $|0\rangle\langle 0| \otimes |1\rangle\langle 1|$ and its partial trace is $|0\rangle\langle 0|$. The second yields $|1\rangle\langle 1|$. The cross terms vanish because $\text{Tr}(|1\rangle\langle 0|) = \langle 0 | 1 \rangle = 0$. Therefore

$$\rho_A = \rho_B = \frac{1}{2}(|0\rangle\langle 0| + |1\rangle\langle 1|) = \frac{\mathbf{1}}{2}. \quad (36.19)$$

The contraction method gives the same answer with less expansion:

$$(\mathbf{1} \otimes \langle 0|)|\Psi^-\rangle = -\frac{|1\rangle}{\sqrt{2}}, \quad (36.20)$$

$$(\mathbf{1} \otimes \langle 1|)|\Psi^-\rangle = \frac{|0\rangle}{\sqrt{2}}. \quad (36.21)$$

Taking the two outer products and adding yields $\mathbf{1}/2$. Notice that these contraction vectors are not individually normalized. Their squared norms already supply the correct probabilities.

36.5 Local operations and conditional reduced states

Two partial-trace identities will be used repeatedly. For operators M, N acting on A ,

$$\text{Tr}_B[(M \otimes \mathbf{1})L(N \otimes \mathbf{1})] = M(\text{Tr}_B L)N. \quad (36.22)$$

This follows immediately by commuting M, N past the partial contractions in (36.5). For an operator Y acting on the traced subsystem,

$$\mathrm{Tr}_B[(\mathbf{1} \otimes Y)L] = \mathrm{Tr}_B[L(\mathbf{1} \otimes Y)]. \quad (36.23)$$

To prove the second identity, expand $L = \sum_i X_i \otimes Y_i$ and use $\mathrm{Tr}(YY_i) = \mathrm{Tr}(Y_i Y)$ in each term. This is a limited cyclicity rule: arbitrary A operators cannot be cycled through a partial trace.

If A undergoes an ideal projective measurement with outcome λ , the joint state changes to

$$\rho_{AB|\lambda} = \frac{(\Pi_\lambda \otimes \mathbf{1})\rho_{AB}(\Pi_\lambda \otimes \mathbf{1})}{p_\lambda}.$$

Using (36.22),

$$\rho_{A|\lambda} = \frac{\Pi_\lambda \rho_A \Pi_\lambda}{\mathrm{Tr}_A(\rho_A \Pi_\lambda)}. \quad (36.24)$$

Thus the subsystem follows precisely the same measurement rule as a system described initially by an ensemble. If A undergoes a unitary U_A , its reduced state similarly obeys

$$\rho'_A = U_A \rho_A U_A^\dagger. \quad (36.25)$$

A selected result on B can also change the conditional state assigned to A . If B has projectors Q_b , then

$$\rho_{A|b} = \frac{\mathrm{Tr}_B[(\mathbf{1} \otimes Q_b)\rho_{AB}(\mathbf{1} \otimes Q_b)]}{p_b}, \quad p_b = \mathrm{Tr}[\rho_{AB}(\mathbf{1} \otimes Q_b)]. \quad (36.26)$$

For the singlet, conditioning on B being spin up along z gives $\rho_{A|+z} = |1\rangle\langle 1|$, while conditioning on spin down gives $\rho_{A|-z} = |0\rangle\langle 0|$. These conditional states differ from the unconditioned $\rho_A = \mathbf{1}/2$.

36.6 Why remote measurements cannot send a signal

If the result on B is not communicated, the relevant state is the weighted average of the conditional states. Its numerator is the partial trace of the nonselective joint state:

$$\rho'_A = \sum_b \mathrm{Tr}_B[(\mathbf{1} \otimes Q_b)\rho_{AB}(\mathbf{1} \otimes Q_b)] \quad (36.27)$$

$$= \sum_b \mathrm{Tr}_B[\rho_{AB}(\mathbf{1} \otimes Q_b^2)] \quad (36.28)$$

$$= \mathrm{Tr}_B \left[\rho_{AB} \left(\mathbf{1} \otimes \sum_b Q_b \right) \right] = \rho_A. \quad (36.29)$$

The second line uses partial cyclicity, the third uses $Q_b^2 = Q_b$, and the final step uses completeness. Every local probability on A is unchanged, even though individual conditional branches can be different.

A unitary operation on B also has no effect on ρ_A :

$$\rho'_A = \text{Tr}_B[(\mathbf{1} \otimes U_B)\rho_{AB}(\mathbf{1} \otimes U_B^\dagger)] \quad (36.30)$$

$$= \text{Tr}_B[\rho_{AB}(\mathbf{1} \otimes U_B^\dagger U_B)] = \rho_A. \quad (36.31)$$

This proves the local-unitary exercise directly. More generally, if the outcomes of an operation on B have operators K_b with $\sum_b K_b^\dagger K_b = \mathbf{1}_B$, the same calculation replaces Q_b^2 by $K_b^\dagger K_b$ and again gives $\rho'_A = \rho_A$. It is the trace-preserving, unconditioned operation that cannot signal. Postselection changes a conditional ensemble, but identifying that ensemble requires a classical record from B .

36.7 The EPR arrangement

Imagine preparing an entangled pair, separating its members into distant laboratories, and making local measurements while the two measurement events are spacelike separated. Each laboratory may choose its setting and record its result before a light-speed message from the other laboratory could arrive. Only later are the two records brought together and compared.

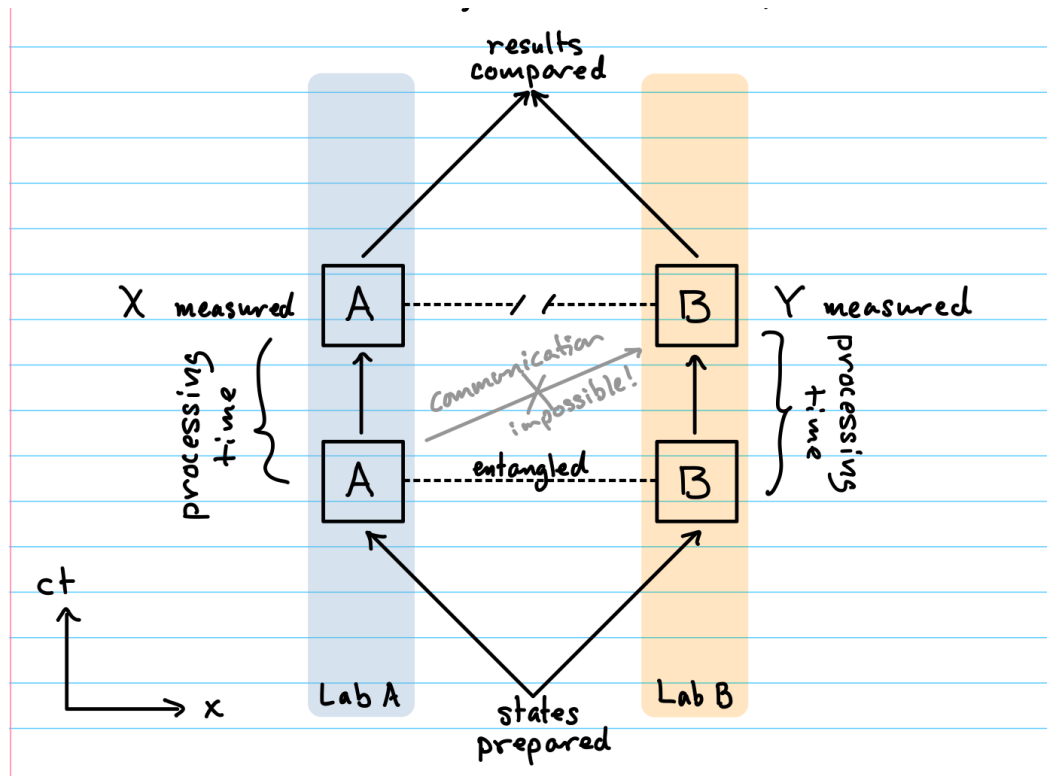


Figure 44: The EPR arrangement: an entangled pair is distributed to two laboratories, local measurements are completed without communication between the measurement events, and the records are compared later. Original handwritten lecture diagram.

Perfect correlations make this arrangement striking. Knowledge of one result can determine what the corresponding measurement on the other member will yield.

Nevertheless, neither observer controls the random local result, and the reduced state is unchanged by an unannounced remote measurement. The arrangement therefore does not provide faster-than-light communication.

Interaction with an environment can transfer information about the entangled degrees of freedom into additional systems. Ignoring those systems can suppress the coherence responsible for the original entanglement. After a rank-one measurement and conditioning on its result, the two measured systems may be in a product state. If the result is forgotten, classical correlations can remain. For example, forgetting a z measurement on one member of a singlet leaves the mixture of $|01\rangle$ and $|10\rangle$: it retains z anticorrelation, while subsequent x measurements on the two spins are independent.

No signalling resolves the communication question, but it does not settle whether all the correlations can be explained by preexisting local variables. That is the question addressed by Bell inequalities.

37 Bell inequalities and quantum correlations

37.1 The experiment and its assumptions

A source produces pairs in the singlet state. The particles travel to two separated laboratories, where spin is measured along independently selected axes $\mathbf{n}_A$ and $\mathbf{n}_B$. Label each outcome by $a, b \in \{+1, -1\}$, corresponding to the spin values $a\hbar/2$ and $b\hbar/2$. The setting choices and measurements are arranged so that neither measurement event can receive a light-speed signal carrying the other setting or outcome. If the separation is L , the relevant setting and measurement interval is shorter than L/c in the simplified laboratory arrangement.

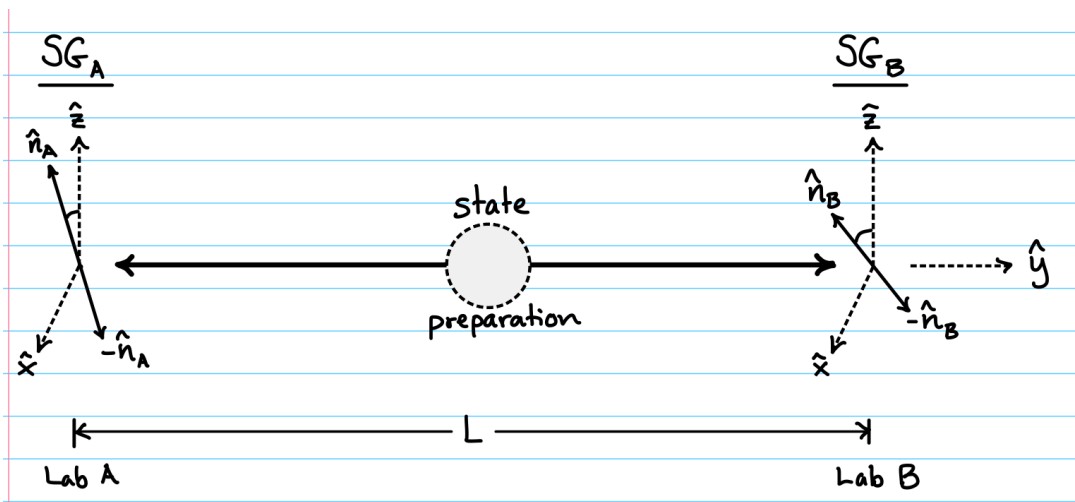


Figure 45: A singlet source sends the two particles to separated Stern–Gerlach analyzers with independently chosen axes. The lecture sketch takes the particles along the y direction and the analyzer axes in the xz plane.

A local hidden-variable description introduces a variable λ characterizing information shared by the pair at preparation. Bell locality is expressed by the conditional factorization

$$P(a, b \mid \mathbf{n}_A, \mathbf{n}_B, \lambda) = P_A(a \mid \mathbf{n}_A, \lambda) P_B(b \mid \mathbf{n}_B, \lambda). \quad (37.1)$$

The distribution of λ is also assumed independent of the freely chosen measurement settings. The observed probabilities are averages over λ . The local probabilities can be deterministic or stochastic; local random variables can be included in an enlarged λ to represent the latter as predetermined response functions. The essential assumption is that one laboratory's response does not depend on the other laboratory's setting once λ is specified.

For a finite list of possible axes, such a model assigns joint probabilities to the outcomes that would occur for each setting. The following three-setting argument makes the resulting constraint particularly transparent. It concerns alternative first measurements on different trials, not a sequence of incompatible measurements on the same particle.

37.2 Joint probabilities for the singlet

Let the angle between the two analyzer axes be θ . Rotational invariance lets us take $\mathbf{n}_A = \hat{z}$ and $\mathbf{n}_B = (\sin \theta, 0, \cos \theta)$. For the second spin, choose

$$|+\mathbf{n}_B\rangle = \cos \frac{\theta}{2} |0\rangle + \sin \frac{\theta}{2} |1\rangle, \quad (37.2)$$

$$|-\mathbf{n}_B\rangle = -\sin \frac{\theta}{2} |0\rangle + \cos \frac{\theta}{2} |1\rangle. \quad (37.3)$$

If the first result is $+$, which occurs with probability $1/2$, the second spin is $|1\rangle$. Its probabilities are

$$P(+\mathbf{n}_B \mid +z) = \sin^2 \frac{\theta}{2}, \quad P(-\mathbf{n}_B \mid +z) = \cos^2 \frac{\theta}{2}.$$

If the first result is $-$, the second spin is $|0\rangle$, interchanging these two conditional probabilities. Multiplication by the first-result probability gives

$$P_{++}(\theta) = P_{--}(\theta) = \frac{1}{2} \sin^2 \frac{\theta}{2} = \frac{1 - \cos \theta}{4}, \quad (37.4)$$

$$P_{+-}(\theta) = P_{-+}(\theta) = \frac{1}{2} \cos^2 \frac{\theta}{2} = \frac{1 + \cos \theta}{4}. \quad (37.5)$$

At $\theta = 0$ the outcomes are perfectly opposite; at $\theta = \pi/2$ all four pairs have probability $1/4$; at $\theta = \pi$ they have the same numerical sign because the two axis conventions are opposite.

For arbitrary axes, the singlet projector can be written

$$\rho_{\Psi^-} = \frac{1}{4} \left(\mathbf{1} \otimes \mathbf{1} - \sum_{i=x,y,z} \sigma_i \otimes \sigma_i \right). \quad (37.6)$$

For example, $\sigma_z \otimes \sigma_z$ supplies the diagonal distinction between parallel and antiparallel spins, while $\sigma_x \otimes \sigma_x + \sigma_y \otimes \sigma_y$ supplies twice the off-diagonal exchanges between $|01\rangle$ and $|10\rangle$. This reproduces (36.18). The local projectors are $\Pi_a(\mathbf{n}) = (\mathbf{1} + a\mathbf{n} \cdot \boldsymbol{\sigma})/2$. Using $\text{Tr } \sigma_i = 0$ and $\text{Tr}(\sigma_i \sigma_j) = 2\delta_{ij}$ gives

$$\boxed{P(a, b \mid \mathbf{n}_A, \mathbf{n}_B) = \frac{1}{4}(1 - ab\mathbf{n}_A \cdot \mathbf{n}_B).} \quad (37.7)$$

The marginal probabilities are $1/2$, independent of the remote axis, whereas the correlation is

$$E(\mathbf{n}_A, \mathbf{n}_B) = \sum_{a,b} abP(a, b) = -\mathbf{n}_A \cdot \mathbf{n}_B. \quad (37.8)$$

This calculation displays no signalling and a nontrivial joint correlation in the same formula.

37.3 A three-setting Bell inequality

Choose three possible axes $\mathbf{n}_1, \mathbf{n}_2, \mathbf{n}_3$. In a predetermined local description, let $\sigma_i = \pm 1$ be A 's result if axis i is selected and $\tau_i = \pm 1$ be B 's result for that axis. A joint probability distribution is assigned to

$$P(\sigma_1, \sigma_2, \sigma_3; \tau_1, \tau_2, \tau_3), \quad (37.9)$$

with nonnegative entries summing to one. For example, one possible assignment is $(\sigma_1, \sigma_2, \sigma_3) = (+, -, +)$ and $(\tau_1, \tau_2, \tau_3) = (-, +, -)$.

To reproduce the singlet's perfect same-axis anticorrelation, every assignment of nonzero weight must obey

$$\tau_i = -\sigma_i \quad (i = 1, 2, 3). \quad (37.10)$$

Otherwise a same-axis parallel outcome would have a positive probability; nonnegative probabilities from other assignments cannot cancel it. There are thus only eight potentially nonzero assignments.

Abbreviate a sign triple by, for example, $+-+$. The observable probability of two positive results along axes 1 and 2 is

$$P_{++}(1, 2) = P(+ - +; - + -) + P(+ - -; - + +). \quad (37.11)$$

Indeed $\sigma_1 = +$ and $\tau_2 = +$ imply $\sigma_2 = -$, leaving σ_3 as the two alternatives. Similarly,

$$P_{++}(3, 2) = P(+ - +; - + -) + P(- - +; + + -), \quad (37.12)$$

$$P_{++}(1, 3) = P(+ + -; - - +) + P(+ - -; - + +). \quad (37.13)$$

The sum of these latter probabilities contains both terms in (37.11), plus two additional nonnegative terms. Therefore every such local model obeys

$$\boxed{P_{++}(1, 2) \leq P_{++}(1, 3) + P_{++}(3, 2).} \quad (37.14)$$

Equivalently, the event $\{\sigma_1 = +, \sigma_2 = -\}$ is contained in the union of $\{\sigma_1 = +, \sigma_3 = -\}$ and $\{\sigma_3 = +, \sigma_2 = -\}$: σ_3 must be one of its two possible signs. The inequality is an ordinary probability constraint on a common joint distribution for the alternative settings.

37.4 Violation by quantum mechanics

Let θ_{ij} be the angle between axes i and j . Substitution of (37.5) into the Bell inequality would require

$$\sin^2 \frac{\theta_{12}}{2} \leq \sin^2 \frac{\theta_{13}}{2} + \sin^2 \frac{\theta_{23}}{2}. \quad (37.15)$$

Choose three coplanar axes with $\mathbf{n}_3$ halfway between the other two: $\theta_{13} = \theta_{23} = \theta$ and $\theta_{12} = 2\theta$.

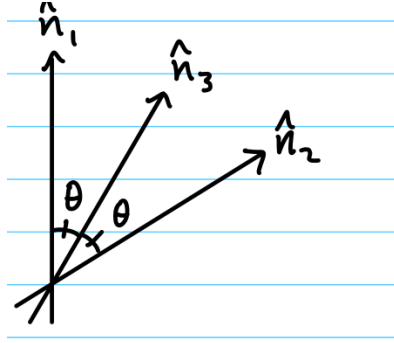


Figure 46: Three analyzer axes with equal neighboring angles θ . Original handwritten lecture diagram.

The bound becomes

$$\sin^2 \theta \leq 2 \sin^2 \frac{\theta}{2}. \quad (37.16)$$

But

$$\sin^2 \theta - 2 \sin^2 \frac{\theta}{2} = 4 \sin^2 \frac{\theta}{2} \cos^2 \frac{\theta}{2} - 2 \sin^2 \frac{\theta}{2} \quad (37.17)$$

$$= 2 \sin^2 \frac{\theta}{2} \cos \theta > 0 \quad (0 < \theta < \pi/2). \quad (37.18)$$

Thus quantum predictions violate the local bound throughout this range. For a concrete example, take $\theta = \pi/3$:

$$P_{++}(1, 2) = \frac{1}{2} \sin^2 \frac{\pi}{3} = \frac{3}{8}, \quad P_{++}(1, 3) + P_{++}(3, 2) = 2 \left(\frac{1}{2} \sin^2 \frac{\pi}{6} \right) = \frac{1}{4}. \quad (37.19)$$

The predicted left-hand side exceeds the bound by $1/8$.

The contradiction is not a failure of ordinary probability theory for actual recorded outcomes. Each fixed quantum experiment has a normalized, nonnegative outcome distribution. What fails is the attempt to embed all these incompatible measurement contexts into the particular common local assignment used above.

37.5 What the Bell result does and does not establish

Bell experiments test inequalities appropriate to their actual apparatus, efficiencies, and spacetime arrangement. A sequence of experiments, beginning with pioneering tests in the 1970s and followed by increasingly controlled tests, has observed the quantum violations. Experiments in 2015 jointly addressed major detection and locality loopholes. The 2022 Nobel Prize in Physics recognized Alain Aspect, John Clauser, and Anton Zeilinger for work on entangled photons, Bell inequalities, and quantum information.

The mathematical conclusion is that the observed quantum correlations cannot be reproduced while retaining all the assumptions entering a Bell-local, setting-independent model. It is not a theorem that determinism by itself is impossible. Nonlocal hidden-variable theories can reproduce quantum correlations; accepting dependence between hidden variables and measurement settings changes another assumption. Such alternatives must be assessed as physical theories, with their assumptions stated explicitly.

Bell locality is stronger than operational no signalling. Quantum theory violates Bell inequalities while retaining the no-signalling identities proved in Section 36.6. Relativistic quantum field theory also imposes local commutativity for appropriate observables at spacelike separation. These statements explain why a violation of a Bell inequality does not by itself construct a controllable signal outside the light cone. A nonlocal hidden-variable account, such as a Bohmian account, likewise should not be identified automatically with a faster-than-light communication protocol.

38 Measurement as an interaction with a detector

38.1 A system qubit and a readout qubit

The projective measurement postulate appears different from unitary evolution: a selected outcome is random and its state update is noninvertible. To understand how a detector participates, include it in the Hilbert space. Let the system basis be $|0\rangle_S, |1\rangle_S$ and a two-state detector have pointer basis $|0\rangle_D, |1\rangle_D$.

A tempting detector rule would be

$$U|i\rangle_S|j\rangle_D = |i\rangle_S|i\rangle_D \quad \text{for both } j = 0, 1. \quad (38.1)$$

This cannot be unitary. For fixed i , the two orthogonal input states $|i, 0\rangle$ and $|i, 1\rangle$ would map to the same output. Their inner product would change from zero to one. Equivalently, the proposed map collapses a four-dimensional input space into the two-dimensional span of $|00\rangle$ and $|11\rangle$, so it is not invertible.

A functioning detector must be initialized in a known ready state. If the initial detector is $|0\rangle_D$, we only need

$$U|0, 0\rangle = |0, 0\rangle, \quad U|1, 0\rangle = |1, 1\rangle. \quad (38.2)$$

These two requirements preserve orthogonality and can be extended to a unitary on the full four-dimensional space.

38.2 The controlled-NOT realization

One such extension is the controlled-NOT operator,

$$U_{\text{CNOT}}|i, j\rangle = |i, j \oplus i\rangle, \quad i, j \in \{0, 1\}, \quad (38.3)$$

where $\oplus$ is addition modulo two. In the basis $(00, 01, 10, 11)$,

$$U_{\text{CNOT}} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}. \quad (38.4)$$

Its columns form an orthonormal basis, so $U^\dagger U = \mathbf{1}$. It is also real symmetric and squares to the identity: applying the controlled bit flip twice restores every basis state. This explicitly solves the detector-unitarity exercise.

Preparing the detector in $|0\rangle$ is itself a physical process. If the pointer has $|0\rangle$ as a nondegenerate ground state, coupling it to a sufficiently cold bath can make its initial state approximately $|0\rangle\langle 0|$. At nonzero temperature the preparation generally has a thermal error. Resetting arbitrary detector states to the same ready state is not unitary on the detector alone; the bath carries away information and entropy. This does not conflict with unitarity of the later closed system–detector interaction.

For a superposition, linearity gives

$$(a|0\rangle + b|1\rangle)|0\rangle_D \mapsto a|0, 0\rangle + b|1, 1\rangle, \quad |a|^2 + |b|^2 = 1. \quad (38.5)$$

The system and pointer become entangled. Measuring the pointer in its record basis yields outcome 0 with probability $|a|^2$ and outcome 1 with probability $|b|^2$, and the corresponding system state is $|0\rangle$ or $|1\rangle$.

38.3 Copying a basis label is not cloning a state

The controlled-NOT operation copies the label of the two orthogonal basis states into a blank pointer. It does not make two copies of an arbitrary unknown superposition. The desired cloned output would be

$$\begin{aligned} (a|0\rangle + b|1\rangle) \otimes (a|0\rangle + b|1\rangle) &= a^2|00\rangle + ab|01\rangle \\ &\quad + ab|10\rangle + b^2|11\rangle, \end{aligned} \quad (38.6)$$

which differs from (38.5).

Theorem 38.1 (No cloning of arbitrary pure states). *A single unitary cannot copy every unknown pure state into a fixed blank register while retaining the original state.*

Proof. Suppose a unitary satisfies

$$U|\psi\rangle|b\rangle = |\psi\rangle|\psi\rangle, \quad U|\phi\rangle|b\rangle = |\phi\rangle|\phi\rangle$$

for normalized states and a fixed normalized blank $|b\rangle$. Preservation of inner products implies

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

Thus the overlap must be zero or one in these phase conventions. In particular, two distinct nonorthogonal rays, with $0 < |\langle\psi|\phi\rangle| < 1$, cannot both be cloned by that device. Orthogonal basis states avoid the obstruction, as the controlled-NOT example shows. $\square$

38.4 Why entanglement alone is not a selected outcome

The joint state in (38.5) is still pure. Its density operator contains off-diagonal terms:

$$\begin{aligned}\rho_{SD} = & |a|^2|00\rangle\langle 00| + |b|^2|11\rangle\langle 11| \\ & + ab^*|00\rangle\langle 11| + a^*b|11\rangle\langle 00|.\end{aligned}\quad (38.7)$$

By contrast, forgetting the outcome of an ideal measurement gives the joint record mixture

$$\rho_{SD}^{\text{record}} = |a|^2|00\rangle\langle 00| + |b|^2|11\rangle\langle 11|. \quad (38.8)$$

These are different global states. For instance, the operator $|00\rangle\langle 11| + |11\rangle\langle 00|$ has expectation $2\text{Re}(a^*b)$ in the coherent state and zero in the mixture. A unitary acting on the closed pair cannot change its initial rank-one density operator into the rank-two mixture when both amplitudes are nonzero.

Tracing over the pointer already gives

$$\rho_S = |a|^2|0\rangle\langle 0| + |b|^2|1\rangle\langle 1|, \quad (38.9)$$

because the orthogonal pointer records remove the cross terms from the partial trace. Thus the system alone follows the nonselective measurement rule, even while the combined state retains coherence. A larger detector explains why recovering that coherence from accessible variables becomes difficult.

38.5 A detector with many records

Let the detector contain N readout qubits, initially in $|0\rangle^{\otimes N}$. Define a unitary on the entire computational basis by

$$U|i, j_1, \dots, j_N\rangle = |i, j_1 \oplus i, \dots, j_N \oplus i\rangle. \quad (38.10)$$

This is again a permutation of an orthonormal basis, with $U^2 = \mathbf{1}$. The resulting state is

$$|\Omega\rangle = a|0\rangle_S|0\rangle^{\otimes N} + b|1\rangle_S|1\rangle^{\otimes N}. \quad (38.11)$$

Many detector degrees of freedom now carry the same basis record. No individual degree of freedom carries an independent copy of the original superposition.

Suppose only the last readout qubit is inaccessible. Let R denote the system plus the first $N - 1$ readouts, and write

$$|R_0\rangle = |0\rangle_S|0\rangle^{\otimes(N-1)}, \quad |R_1\rangle = |1\rangle_S|1\rangle^{\otimes(N-1)}.$$

Then $|\Omega\rangle = a|R_0\rangle|0\rangle_N + b|R_1\rangle|1\rangle_N$. Expanding the projector and tracing the last qubit gives

$$\begin{aligned}\rho_R &= |a|^2|R_0\rangle\langle R_0| \text{Tr}(|0\rangle\langle 0|) + |b|^2|R_1\rangle\langle R_1| \text{Tr}(|1\rangle\langle 1|) \\ &\quad + ab^*|R_0\rangle\langle R_1| \text{Tr}(|0\rangle\langle 1|) + a^*b|R_1\rangle\langle R_0| \text{Tr}(|1\rangle\langle 0|)\end{aligned}\quad (38.12)$$

$$= |a|^2|R_0\rangle\langle R_0| + |b|^2|R_1\rangle\langle R_1|. \quad (38.13)$$

The cross terms vanish because $\langle 1 | 0 \rangle = 0$. Tracing out further readouts preserves the same form of classical record correlations among the remaining degrees of freedom.

In a less idealized model, let the environment records be $|E_0\rangle, |E_1\rangle$, which need not be exactly orthogonal:

$$|\Omega\rangle = a|R_0\rangle|E_0\rangle + b|R_1\rangle|E_1\rangle. \quad (38.14)$$

The reduced off-diagonal term is then

$$ab^*\langle E_1 | E_0 \rangle |R_0\rangle\langle R_1|. \quad (38.15)$$

If the environment consists of many independently correlated factors, $|E_i\rangle = \bigotimes_{k=1}^N |e_i^{(k)}\rangle$, its overlap is a product:

$$\langle E_1 | E_0 \rangle = \prod_{k=1}^N \langle e_1^{(k)} | e_0^{(k)} \rangle. \quad (38.16)$$

Even moderately distinguishable records can make this product very small. Suppression of locally accessible interference through such correlations is **decoherence**.

38.6 The scope of the unitary detector model

The detector model explains how an initially coherent system can acquire a mixed reduced state, how a pointer becomes correlated with an outcome basis, and why many accessible records agree. The coefficients of the reduced mixture are the Born weights $|a|^2, |b|^2$. Conditional on reading one pointer outcome, the state assignment agrees with the corresponding projective update.

The calculation uses the Born rule and the rule for conditional measurements when interpreting an observed pointer result; it does not derive those rules merely from a unitary matrix. Nor does taking a partial trace turn the total pure state into a single selected branch. The difference between an unconditioned reduced state and an individual observed outcome remains part of the interpretation of measurement.

An interpretation retaining a universally unitary pure state may regard the branches as relative states of correlated observers and records. Other interpretations treat the occurrence or description of individual outcomes differently. The reduced-state calculations above do not require choosing between those interpretations. If a larger closed system is assumed pure, its subsystems can nevertheless be mixed solely because they are entangled with the rest. Purification makes that mathematical possibility precise.

39 Schmidt decomposition and purification

39.1 The Schmidt decomposition

A general pure state of $A \otimes B$ can involve $d_A d_B$ product-basis coefficients. A suitable choice of local bases greatly simplifies it.

Theorem 39.1 (Schmidt decomposition). *For every normalized bipartite pure state there are positive numbers $p_1, \dots, p_r$ summing to one and orthonormal sets $\{|u_i\rangle_A\}$, $\{|v_i\rangle_B\}$ such that*

$$|\Psi\rangle = \sum_{i=1}^r \sqrt{p_i} |u_i\rangle_A \otimes |v_i\rangle_B, \quad r \leq \min(d_A, d_B). \quad (39.1)$$

The number r is the Schmidt rank. The numbers p_i are the nonzero eigenvalues, including multiplicities, of both reduced density operators.

Proof. Diagonalize the positive operator $\rho_A = \text{Tr}_B |\Psi\rangle\langle\Psi|$:

$$\rho_A = \sum_{i=1}^r p_i |u_i\rangle\langle u_i|, \quad p_i > 0. \quad (39.2)$$

Complete the $|u_i\rangle$ to a full orthonormal basis of $\mathcal{H}_A$ if necessary. For each positive eigenvalue define

$$|v_i\rangle_B = \frac{1}{\sqrt{p_i}} (\langle u_i |_A \otimes \mathbf{1}_B) |\Psi\rangle. \quad (39.3)$$

Their inner products are

$$\langle v_i | v_j \rangle = \frac{\langle \Psi | |u_i\rangle\langle u_j| \otimes \mathbf{1}_B | \Psi \rangle}{\sqrt{p_i p_j}} \quad (39.4)$$

$$= \frac{\langle u_j | \rho_A | u_i \rangle}{\sqrt{p_i p_j}} = \delta_{ij}. \quad (39.5)$$

If a basis vector $|u\rangle$ belongs to the kernel of ρ_A , then

$$\|(\langle u | \otimes \mathbf{1}) |\Psi\rangle\|^2 = \langle u | \rho_A | u \rangle = 0,$$

so its contraction vanishes. Inserting the complete A basis into $|\Psi\rangle$ therefore gives

$$|\Psi\rangle = \sum_{i=1}^r |u_i\rangle \otimes [(\langle u_i | \otimes \mathbf{1}) |\Psi\rangle] = \sum_{i=1}^r \sqrt{p_i} |u_i\rangle \otimes |v_i\rangle.$$

Because there are r orthonormal vectors in each subsystem, $r \leq \min(d_A, d_B)$. Taking either partial trace yields

$$\rho_A = \sum_i p_i |u_i\rangle\langle u_i|, \quad \rho_B = \sum_i p_i |v_i\rangle\langle v_i|. \quad (39.6)$$

Thus the two reduced operators have exactly the same nonzero spectrum. $\square$

This is also a singular-value decomposition of the coefficient matrix c_{ab} : the Schmidt coefficients $\sqrt{p_i}$ are its nonzero singular values. Degenerate p_i allow corresponding changes of Schmidt bases; the multiset of p_i and the Schmidt rank do not depend on those choices.

Corollary 39.2. *A bipartite pure state is a product if and only if either reduced state is pure. Equivalently, it is a product if and only if its Schmidt rank equals one.*

Proof. For $r = 1$, the Schmidt decomposition is a single product with coefficient one. Conversely, a product has rank-one reduced operators. If either reduced state is pure, its only nonzero eigenvalue is one; the shared Schmidt spectrum then has a single term. $\square$

A globally pure state can therefore have mixed marginals only through entanglement. The same statement cannot be used for a globally mixed state: classical correlations can also give mixed marginals.

39.2 Unequal dimensions and maximal entanglement

If $d_A \leq d_B$, the reduced state on B has rank at most d_A . When $d_A < d_B$, it cannot equal $\mathbf{1}_B/d_B$, which has rank d_B . The smaller subsystem can nevertheless be maximally mixed. Choose orthonormal states $|1\rangle_B, \dots, |d_A\rangle_B$ and let

$$|\Phi_{d_A}\rangle = \frac{1}{\sqrt{d_A}} \sum_{i=1}^{d_A} |i\rangle_A |i\rangle_B. \quad (39.7)$$

Then

$$\rho_A = \frac{\mathbf{1}_A}{d_A}, \quad \rho_B = \frac{1}{d_A} \sum_{i=1}^{d_A} |i\rangle_B \langle i|_B. \quad (39.8)$$

The latter is maximally mixed on its d_A -dimensional support but has zero eigenvalues outside that support. Both subsystems are maximally mixed on their entire spaces only when their dimensions are equal.

For example, a qubit and a three-level system may be in

$$|\Psi\rangle = \frac{|0\rangle_A |0\rangle_B + |1\rangle_A |1\rangle_B}{\sqrt{2}}. \quad (39.9)$$

Their reduced spectra are $(1/2, 1/2)$ and $(1/2, 1/2, 0)$. The zero in the larger space does not imply an asymmetry of entanglement; the nonzero spectra, and hence their entanglement entropies, agree.

39.3 Purifying an arbitrary density operator

Definition 39.3. A **purification** of ρ_A is a pure state $|\Psi\rangle_{AB}$ on a larger space such that $\text{Tr}_B |\Psi\rangle \langle \Psi| = \rho_A$.

Let the spectral decomposition of a rank- r density operator be

$$\rho_A = \sum_{i=1}^r p_i |i\rangle_A \langle i|_A, \quad p_i > 0, \quad \sum_i p_i = 1. \quad (39.10)$$

Choose any auxiliary system B with at least r orthonormal states and define

$$|\Psi\rangle_{AB} = \sum_{i=1}^r \sqrt{p_i} |i\rangle_A |i\rangle_B. \quad (39.11)$$

The state is normalized because its product kets are orthonormal. Its partial trace is

$$\text{Tr}_B |\Psi\rangle \langle \Psi| = \sum_{i,j} \sqrt{p_i p_j} |i\rangle_A \langle j|_A \text{Tr}_B (|i\rangle_B \langle j|_B) \quad (39.12)$$

$$= \sum_{i,j} \sqrt{p_i p_j} |i\rangle_A \langle j|_A \delta_{ij} = \rho_A. \quad (39.13)$$

Thus every density operator admits a purification. Conversely, the Schmidt-rank bound shows that a purification requires

$$\dim \mathcal{H}_B \geq \text{rank } \rho_A. \quad (39.14)$$

This condition is both necessary and sufficient.

Purification is a mathematical realization of the fact that a mixed state can arise from ignoring an entangled partner. It does not assert that a specific laboratory ensemble was prepared by this particular auxiliary system. Local observations determine ρ_A , not which of its possible purifications or ensemble realizations occurred.

39.4 Uniqueness up to a unitary on the purifier

Proposition 39.4. *If $|\Psi\rangle$ and $|\Phi\rangle$ purify the same ρ_A using the same finite-dimensional auxiliary space $\mathcal{H}_B$, there exists a unitary V_B such that*

$$|\Phi\rangle = (\mathbf{1}_A \otimes V_B) |\Psi\rangle. \quad (39.15)$$

Proof. Fix one orthonormal eigenbasis $\{|u_i\rangle\}$ for the support of ρ_A , with positive eigenvalues p_i . Apply the contraction construction (39.3) to each purification using this same basis. It gives two orthonormal sets in $\mathcal{H}_B$:

$$|\Psi\rangle = \sum_i \sqrt{p_i} |u_i\rangle |v_i\rangle, \quad |\Phi\rangle = \sum_i \sqrt{p_i} |u_i\rangle |w_i\rangle.$$

Complete each set to a full orthonormal basis of $\mathcal{H}_B$. The linear map sending the first completed basis to the second is unitary and satisfies $V_B |v_i\rangle = |w_i\rangle$. Applying it to the Schmidt expansion proves the result. Degenerate eigenvalues cause no difficulty because the same fixed eigenbasis was used for both contractions. $\square$

The converse is immediate from local-unitary no signalling: a unitary on B leaves ρ_A unchanged. For purifying spaces of different dimensions, the corresponding statement uses an isometry between the supported subspaces, or embeds both purifiers into a common larger space before extending that isometry to a unitary.

40 Entropy and entanglement

40.1 Classical entropy and operator functions

For N equally likely alternatives, the dimensionless entropy is $S = \log N$. For a general discrete distribution, the corresponding expression is

$$S(p) = - \sum_i p_i \log p_i. \quad (40.1)$$

Natural logarithms are used here. Dividing by $\log 2$ expresses entropy in bits, and multiplying by Boltzmann's constant k_B gives physical entropy units. The convention $0 \log 0 = 0$ follows from the limit as $p \rightarrow 0^+$.

For a Hermitian operator with spectral decomposition $A = \sum_i a_i |i\rangle\langle i|$, a function is defined spectrally:

$$f(A) = \sum_i f(a_i) |i\rangle\langle i|. \quad (40.2)$$

In particular, the logarithm of a strictly positive density operator is $\log \rho = \sum_i (\log p_i) |i\rangle\langle i|$. If zero eigenvalues are present, $\rho \log \rho$ is defined by continuous extension on its kernel. This leads to the **von Neumann entropy**

$$S(\rho) = - \text{Tr}(\rho \log \rho) = - \sum_i p_i \log p_i. \quad (40.3)$$

The probabilities in this spectral formula are the eigenvalues of ρ , not the weights of an arbitrarily chosen nonorthogonal ensemble decomposition. Different ensembles describing the same density operator have the same von Neumann entropy, even when the classical entropy of their preparation labels differs.

For instance, the pure state $|0\rangle\langle 0|$ may be prepared with a redundant classical random label attached to identical copies of $|0\rangle$. The label can have positive classical entropy while the system remains pure and has zero von Neumann entropy. Entropy of a preparation record and entropy of the quantum system are distinct quantities.

40.2 Basic entropy properties

Because $0 \leq p_i \leq 1$, each term $-p_i \log p_i$ is nonnegative. Thus

$$S(\rho) \geq 0, \quad S(\rho) = 0 \iff \rho \text{ is pure.} \quad (40.4)$$

The equivalence follows because each nonzero term vanishes only when its eigenvalue is one, and the eigenvalues sum to one.

In dimension d , the largest possible entropy is $\log d$. One way to prove this without assuming full rank is to let r be the number of positive eigenvalues and use concavity of the logarithm:

$$S(\rho) = \sum_{i=1}^r p_i \log \frac{1}{p_i} \quad (40.5)$$

$$\leq \log \left(\sum_{i=1}^r p_i \frac{1}{p_i} \right) = \log r \leq \log d. \quad (40.6)$$

Equality with $\log d$ requires $r = d$ and all $1/p_i$ equal, hence $p_i = 1/d$. Therefore

$$0 \leq S(\rho) \leq \log d, \quad S(\mathbf{1}/d) = \log d. \quad (40.7)$$

Unitary conjugation preserves the eigenvalues, so

$$S(U\rho U^\dagger) = S(\rho). \quad (40.8)$$

The entropy of a closed system undergoing unitary evolution is constant, though subsystem entropies can change as interactions create or remove entanglement.

For a qubit with Bloch vector length $r = |\mathbf{r}|$, the eigenvalues are $(1 \pm r)/2$, giving

$$S(\rho) = -\frac{1+r}{2} \log \frac{1+r}{2} - \frac{1-r}{2} \log \frac{1-r}{2}. \quad (40.9)$$

It is zero on the surface of the Bloch sphere and $\log 2$ at the maximally mixed center.

40.3 Entanglement entropy of a pure pair

For a bipartite pure state, define its **entanglement entropy** by

$$S_{\text{ent}}(|\Psi\rangle) = S(\rho_A), \quad \rho_A = \text{Tr}_B |\Psi\rangle\langle\Psi|. \quad (40.10)$$

The Schmidt theorem proves that this does not depend on which side is used. Indeed, the nonzero spectra of ρ_A and ρ_B are the same, and any extra zero eigenvalues contribute zero. Therefore

$$\boxed{S(\rho_A) = S(\rho_B) = -\sum_i p_i \log p_i, \quad S(\rho_{AB}) = 0 \quad \text{for a pure joint state.}} \quad (40.11)$$

This solves the equal-entropy exercise, including unequal subsystem dimensions. Entanglement entropy vanishes precisely for a pure product state, and is at most $\log \min(d_A, d_B)$.

Example 40.1 (A continuously variable amount of entanglement). For

$$|\Psi_p\rangle = \sqrt{p}|00\rangle + \sqrt{1-p}|11\rangle, \quad 0 \leq p \leq 1, \quad (40.12)$$

the Schmidt probabilities are p and $1-p$. Hence

$$S_{\text{ent}} = -p \log p - (1-p) \log(1-p). \quad (40.13)$$

At $p = 0$ or $p = 1$ the state is a product. At $p = 1/2$ it is maximally entangled and each marginal is $\mathbf{1}/2$, with entropy $\log 2$. The singlet has the same Schmidt spectrum and the same entanglement entropy, although its spin correlations have a different basis form.

40.4 Entropy of product states and correlated states

Proposition 40.2 (Additivity for a product state). *For normalized density operators on two subsystems,*

$$\boxed{S(\rho_A \otimes \rho_B) = S(\rho_A) + S(\rho_B)}. \quad (40.14)$$

Proof. Let p_i and q_j be the eigenvalues of the two density operators. Their product eigenvectors diagonalize $\rho_A \otimes \rho_B$, with eigenvalues $p_i q_j$. Restricting the sum to positive factors and using $\log(p_i q_j) = \log p_i + \log q_j$,

$$\begin{aligned} S(\rho_A \otimes \rho_B) &= - \sum_{i,j} p_i q_j (\log p_i + \log q_j) \\ &= - \left(\sum_j q_j \right) \sum_i p_i \log p_i - \left(\sum_i p_i \right) \sum_j q_j \log q_j \\ &= S(\rho_A) + S(\rho_B). \end{aligned}$$

Zero eigenvalues contribute zero and do not alter the result. $\square$

The product hypothesis is essential. A separable mixture can contain classical correlations, so the absence of entanglement is not enough to establish additivity. Consider

$$\rho_{cc} = \frac{1}{2} |00\rangle\langle 00| + \frac{1}{2} |11\rangle\langle 11|. \quad (40.15)$$

This is explicitly a mixture of product states. Its joint spectrum is $(1/2, 1/2, 0, 0)$, while both marginal spectra are $(1/2, 1/2)$. Thus

$$S(\rho_{cc}) = \log 2, \quad S(\rho_A) + S(\rho_B) = 2 \log 2. \quad (40.16)$$

The same marginals also occur for the pure state $(|00\rangle + |11\rangle)/\sqrt{2}$, whose joint entropy is zero. Consequently a mixed marginal by itself does not establish entanglement when the joint state is not known to be pure.

The quantity

$$I(A : B) = S(\rho_A) + S(\rho_B) - S(\rho_{AB}) \quad (40.17)$$

is called the quantum mutual information. In these examples it is zero for a product, $\log 2$ for the classical correlated mixture, and $2 \log 2$ for the pure maximally entangled pair. It measures correlations in the joint state rather than the entropy of either part alone.

40.5 Entropy, environments, and thermodynamic descriptions

If a complete closed system is assumed to be in a pure state, the entropy of any subsystem's exact density operator is its entanglement entropy with the complement. The full entropy remains zero under unitary evolution, while an interaction can raise the subsystem entropy by distributing information into correlations. The many-record detector model is an explicit example: the system begins pure, but its reduced entropy after recording is

$$-|a|^2 \log |a|^2 - |b|^2 \log |b|^2. \quad (40.18)$$

The joint state of system, detector, and all retained environmental degrees of freedom can still be pure.

This observation provides a useful microscopic perspective on entropy in an open system. Its application to thermodynamics requires specifying the system–environment division and the state used in the thermodynamic description. Thermal ensembles or coarse-grained descriptions may also encode incomplete preparation information or discarded macroscopic distinctions. Therefore thermodynamic entropy can coincide with entanglement entropy under appropriate pure-global-state assumptions, but the identification is not automatic for every statistical description. The density-operator formalism keeps these possibilities compatible: local predictions depend on the reduced state regardless of whether it was obtained from an ensemble, a purification, or a larger entangled system.

Part VI

Continuous Observables

41 Canonical commutation relations and infinite-dimensional spaces

The finite-dimensional formalism describes spin and other systems with a finite set of independent states. A particle moving on a line requires a larger state space. The obstruction can be seen before solving any particular Hamiltonian. Classical position and momentum obey the canonical Poisson bracket $\{x, p\}_{\text{PB}} = 1$. The canonical quantum relation is

$$[X, P] = i\hbar \mathbf{1}. \quad (41.1)$$

Capital letters will denote operators, while lower-case x, p denote their spectral values. This distinction is useful when an equation contains both an operator and an ordinary coordinate.

41.1 The trace obstruction

Suppose X and P were matrices on a Hilbert space of finite dimension $d > 0$. Cyclicity of the trace would give

$$\text{Tr}(XP - PX) = \text{Tr}(XP) - \text{Tr}(PX) = 0, \quad (41.2)$$

whereas tracing the right side of (41.1) would give

$$\text{Tr}(i\hbar \mathbf{1}) = i\hbar d \neq 0. \quad (41.3)$$

Thus no finite-dimensional matrices satisfy the canonical relation. The conclusion concerns the exact relation; truncating a particle problem to a finite basis necessarily modifies the commutator at the truncation.

In infinite dimensions, this trace calculation is not valid. The identity has infinite trace, and the traces of the unbounded products XP and PX need not exist. One cannot subtract two divergent traces and then use cyclicity as if they were finite numbers.

Definition 41.1 (Trace-class operator). An operator A is trace class if

$$\text{Tr} |A| < \infty, \quad |A| = \sqrt{A^\dagger A}. \quad (41.4)$$

Equivalently, its singular values s_n satisfy $\sum_n s_n < \infty$. For a trace-class A and bounded B , both products are trace class and $\text{Tr}(AB) = \text{Tr}(BA)$. These hypotheses do not apply to the attempted trace of the canonical commutator.

Another important difference is that position and momentum are unbounded operators. They are defined on appropriate dense domains, rather than on every vector of the Hilbert space. Equation (41.1) holds on a common invariant domain where the products make sense. The Schwartz space of smooth, rapidly decreasing wavefunctions will provide such a domain.

41.2 Algebra first, representation second

For the oscillator, the Hamiltonian and commutator are known before we have chosen coordinates for states:

$$H = \frac{P^2}{2m} + \frac{1}{2}m\omega^2 X^2, \quad m > 0, \quad \omega > 0, \quad [X, P] = i\hbar \mathbf{1}. \quad (41.5)$$

The task is to find a representation of this algebra and determine the energy spectrum. This resembles the algebraic construction of spin representations, but the oscillator ladder will have a lowest rung and no highest rung. The state space will be infinite dimensional.

The construction below separates two statements. The ladder argument classifies the possible normalizable number eigenstates. The explicit Hermite-function basis constructed afterward establishes their existence and completeness in the usual position representation.

42 The harmonic oscillator from ladder operators

42.1 Creation, annihilation, and number operators

Define

$$a = \sqrt{\frac{m\omega}{2\hbar}} \left(X + \frac{iP}{m\omega} \right), \quad a^\dagger = \sqrt{\frac{m\omega}{2\hbar}} \left(X - \frac{iP}{m\omega} \right). \quad (42.1)$$

These operators are dimensionless and are adjoints of one another on their appropriate domains. Neither is self-adjoint. Their commutator is

$$\begin{aligned} [a, a^\dagger] &= \frac{m\omega}{2\hbar} \left[\left[X, -\frac{iP}{m\omega} \right] + \left[\frac{iP}{m\omega}, X \right] \right] \\ &= \frac{1}{2\hbar} [-i[X, P] + i[P, X]] = \mathbf{1}. \end{aligned} \quad (42.2)$$

In particular, they are not normal operators, since $aa^\dagger \neq a^\dagger a$.

The number operator is

$$N = a^\dagger a. \quad (42.3)$$

Expanding the product without interchanging X and P gives

$$\begin{aligned} N &= \frac{m\omega}{2\hbar} \left(X^2 + \frac{P^2}{m^2\omega^2} + \frac{i}{m\omega} [X, P] \right) \\ &= \frac{1}{\hbar\omega} \left(\frac{P^2}{2m} + \frac{1}{2}m\omega^2 X^2 \right) - \frac{1}{2}\mathbf{1}. \end{aligned} \quad (42.4)$$

Therefore

$$\boxed{H = \hbar\omega \left(N + \frac{1}{2}\mathbf{1} \right)}. \quad (42.5)$$

The constant 1/2 comes from noncommutativity. Replacing the operators by commuting classical variables would lose this term.

For every vector on which the expression is defined,

$$\langle \psi | N | \psi \rangle = \|a|\psi\rangle\|^2 \geq 0.$$

Thus N is nonnegative. In the oscillator representation it is a self-adjoint operator with a domain specified below.

42.2 How the ladder moves an eigenvalue

The product rule for commutators gives

$$[N, a] = [a^\dagger a, a] = a^\dagger[a, a] + [a^\dagger, a]a = -a, \quad (42.6)$$

$$[N, a^\dagger] = a^\dagger[a, a^\dagger] + [a^\dagger, a^\dagger]a = a^\dagger. \quad (42.7)$$

Let $|\nu\rangle$ be a normalized eigenvector of N with eigenvalue ν . Then

$$N(a|\nu\rangle) = (aN - a)|\nu\rangle = (\nu - 1)a|\nu\rangle, \quad (42.8)$$

$$N(a^\dagger|\nu\rangle) = (a^\dagger N + a^\dagger)|\nu\rangle = (\nu + 1)a^\dagger|\nu\rangle. \quad (42.9)$$

Whenever these vectors are nonzero, they are eigenvectors one unit below or above the original eigenvalue. Their squared norms determine whether a rung can disappear:

$$\|a|\nu\rangle\|^2 = \langle \nu | a^\dagger a | \nu \rangle = \nu, \quad (42.10)$$

$$\|a^\dagger|\nu\rangle\|^2 = \langle \nu | a a^\dagger | \nu \rangle = \langle \nu | N + \mathbf{1} | \nu \rangle = \nu + 1. \quad (42.11)$$

Positivity requires $\nu \geq 0$. The raising operation never vanishes, and the lowering operation vanishes precisely at $\nu = 0$.

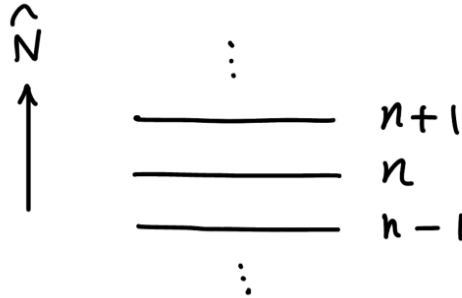


Figure 47: The ladder of number eigenstates. Each application of a or $a^\dagger$ changes the number by one.

42.3 Why the eigenvalues are nonnegative integers

Starting from a number eigenstate, repeatedly lower it. If ν were not an integer, eventually one would obtain a nonzero eigenvector with an eigenvalue $0 < \nu' < 1$. Its

lowered vector would still have positive norm ν' , but its eigenvalue would be $\nu' - 1 < 0$, contradicting positivity of N . Hence

$$\nu = n \in \{0, 1, 2, \dots\}. \quad (42.12)$$

The lowest state satisfies

$$a|0\rangle = 0. \quad (42.13)$$

Raising it repeatedly produces a state at every nonnegative integer. There is no top rung because the norm in (42.11) is strictly positive for all these states.

By (42.5), the energy eigenvalues are

$$E_n = \hbar\omega \left(n + \frac{1}{2} \right), \quad n = 0, 1, 2, \dots \quad (42.14)$$

Neighboring levels are separated by $\hbar\omega$. The ground-state energy $E_0 = \hbar\omega/2$ is the **zero-point energy**; the classical equilibrium $x = p = 0$ would instead have energy zero.

42.4 Normalization, phases, and the Fock basis

Choose an orthonormal vector at each rung and fix relative phases so that the lowering coefficient is real and positive:

$$a|n\rangle = c_n|n-1\rangle, \quad c_n > 0 \quad (n \geq 1).$$

Equation (42.10) gives $c_n^2 = n$. Taking the adjoint matrix element, or multiplying by $a^\dagger$, fixes the raising coefficient. Thus

$$a|n\rangle = \sqrt{n}|n-1\rangle, \quad a^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle. \quad (42.15)$$

The ground-state convention $a|0\rangle = 0$ applies separately. Repeated raising gives

$$|n\rangle = \frac{(a^\dagger)^n}{\sqrt{n!}}|0\rangle. \quad (42.16)$$

To verify the normalization, commute one a through $(a^\dagger)^n$:

$$[a, (a^\dagger)^n] = n(a^\dagger)^{n-1}.$$

Applying this repeatedly to the vacuum gives $a^n(a^\dagger)^n|0\rangle = n!|0\rangle$. Different n values are orthogonal because they have different eigenvalues of the self-adjoint operator N .

On the oscillator Hilbert space, the basis resolves the identity:

$$\mathbf{1} = \sum_{n=0}^{\infty} |n\rangle\langle n|, \quad |\psi\rangle = \sum_{n=0}^{\infty} \psi_n |n\rangle, \quad \psi_n = \langle n | \psi \rangle. \quad (42.17)$$

The series converges in Hilbert-space norm and

$$\|\psi\|^2 = \sum_{n=0}^{\infty} |\psi_n|^2. \quad (42.18)$$

Consequently,

$$\mathcal{H}_{\text{osc}} \cong \ell^2(\mathbb{N}_0), \quad (42.19)$$

the space of square-summable complex sequences. The algebraic finite linear combinations of number states form a useful dense invariant domain.

42.5 Worked exercise: equal degeneracy of every rung

Let $\mathcal{K}_n = \ker(N - n\mathbf{1})$ be the eigenspace at level n . For $|v\rangle \in \mathcal{K}_n$, (42.10) implies

$$\left\| \frac{a|v\rangle}{\sqrt{n}} \right\| = \|v\| \quad (n \geq 1).$$

Moreover, on $\mathcal{K}_n$, $a^\dagger a/n = \mathbf{1}$, while on $\mathcal{K}_{n-1}$, $aa^\dagger/n = \mathbf{1}$. Thus

$$\frac{a}{\sqrt{n}} : \mathcal{K}_n \longrightarrow \mathcal{K}_{n-1} \quad (42.20)$$

is a unitary isomorphism with inverse $a^\dagger/\sqrt{n}$. Every eigenspace has the same degeneracy as the vacuum space $\mathcal{K}_0 = \ker a$.

Choose an orthonormal basis $|0, \alpha\rangle$ of $\mathcal{K}_0$ and set

$$|n, \alpha\rangle = \frac{(a^\dagger)^n}{\sqrt{n!}} |0, \alpha\rangle. \quad (42.21)$$

The previous norm calculation gives $\langle m, \beta | n, \alpha \rangle = \delta_{mn} \delta_{\alpha\beta}$. The map

$$|n\rangle \otimes |\alpha\rangle \longmapsto |n, \alpha\rangle$$

extends to a unitary identification

$$\boxed{\mathcal{H} \cong \mathcal{H}_{\text{osc}} \otimes \mathcal{H}_{\text{int}}}. \quad (42.22)$$

Here $\mathcal{H}_{\text{int}} \cong \mathcal{K}_0$, and $X, P, a, a^\dagger, N$ act as the oscillator operator tensored with the identity on $\mathcal{H}_{\text{int}}$. For an electron in a spin-independent oscillator potential, the second factor may be the two-dimensional spin space. More general independent internal degrees of freedom give a larger factor. An irreducible representation of the position–momentum algebra has only one copy of the oscillator ladder.

42.6 Position and momentum matrix elements

Inverting (42.1),

$$X = x_{\text{osc}}(a + a^\dagger), \quad P = ip_{\text{osc}}(a^\dagger - a), \quad (42.23)$$

where

$$x_{\text{osc}} = \sqrt{\frac{\hbar}{2m\omega}}, \quad p_{\text{osc}} = \sqrt{\frac{m\hbar\omega}{2}}, \quad x_{\text{osc}}p_{\text{osc}} = \frac{\hbar}{2}. \quad (42.24)$$

Hence

$$\langle m|X|n\rangle = x_{\text{osc}} \left(\sqrt{n} \delta_{m,n-1} + \sqrt{n+1} \delta_{m,n+1} \right), \quad (42.25)$$

$$\langle m|P|n\rangle = ip_{\text{osc}} \left(\sqrt{n+1} \delta_{m,n+1} - \sqrt{n} \delta_{m,n-1} \right). \quad (42.26)$$

The diagonal entries vanish, so $\langle X \rangle_n = \langle P \rangle_n = 0$. In X^2 , the terms a^2 and $(a^\dagger)^2$ connect different levels, while $aa^\dagger + a^\dagger a = 2N + \mathbf{1}$. Therefore

$$\langle X^2 \rangle_n = \frac{\hbar}{2m\omega} (2n+1), \quad (42.27)$$

$$\langle P^2 \rangle_n = \frac{m\hbar\omega}{2} (2n+1), \quad (42.28)$$

$$\Delta X_n \Delta P_n = \hbar \left(n + \frac{1}{2} \right). \quad (42.29)$$

Only the ground state attains $\Delta X \Delta P = \hbar/2$.

43 Unbounded operators, domains, and generalized states

43.1 Normalizable does not mean every operator can act

A state belongs to ℓ^2 when $\sum_n |\psi_n|^2 < \infty$. The ladder operators multiply coefficients by factors growing as $\sqrt{n}$, so their outputs may fail to be square summable. Explicitly,

$$a|\psi\rangle = \sum_{n=0}^{\infty} \sqrt{n+1} \psi_{n+1} |n\rangle, \quad (43.1)$$

$$a^\dagger|\psi\rangle = \sum_{n=1}^{\infty} \sqrt{n} \psi_{n-1} |n\rangle. \quad (43.2)$$

Their domains and norms are

$$\mathcal{D}(a) = \left\{ \psi \in \ell^2 : \sum_{n=0}^{\infty} n |\psi_n|^2 < \infty \right\}, \quad (43.3)$$

$$\|a\psi\|^2 = \sum_{n=0}^{\infty} n |\psi_n|^2, \quad (43.4)$$

$$\mathcal{D}(a^\dagger) = \left\{ \psi \in \ell^2 : \sum_{n=0}^{\infty} (n+1) |\psi_n|^2 < \infty \right\}, \quad (43.5)$$

$$\|a^\dagger\psi\|^2 = \sum_{n=0}^{\infty} (n+1) |\psi_n|^2. \quad (43.6)$$

Since $\psi \in \ell^2$ already, the two convergence conditions are equivalent:

$$\mathcal{D}(a) = \mathcal{D}(a^\dagger), \quad \|a^\dagger\psi\|^2 - \|a\psi\|^2 = \|\psi\|^2. \quad (43.7)$$

These statements solve the first two parts of the domain exercise.

For example,

$$\psi_n = \frac{C}{n+1}, \quad C = \frac{\sqrt{6}}{\pi}, \quad (43.8)$$

is normalized because $\sum_{n=0}^{\infty} (n+1)^{-2} = \pi^2/6$. However,

$$\sum_{n=0}^{\infty} n|\psi_n|^2 = C^2 \sum_{n=0}^{\infty} \frac{n}{(n+1)^2} = \infty,$$

so neither $a\psi$ nor $a^\dagger\psi$ is a Hilbert-space vector. Unboundedness is also immediate from $\|a|n\rangle\| = \sqrt{n}$: no constant bounds the output norm by the input norm uniformly over normalized states.

The number operator has the smaller domain

$$\mathcal{D}(N) = \left\{ \psi \in \ell^2 : \sum_{n=0}^{\infty} n^2 |\psi_n|^2 < \infty \right\}. \quad (43.9)$$

Finite mean number, $\sum n|\psi_n|^2 < \infty$, is the quadratic-form condition; it need not imply that the vector $N\psi$ exists in the Hilbert space. The same distinction applies to finite mean oscillator energy and the domain of H .

43.2 Worked exercise: cancellation can enlarge the position domain

If $\psi \in \mathcal{D}(a) = \mathcal{D}(a^\dagger)$, the triangle inequality gives

$$\|X\psi\| \leq x_{\text{osc}} (\|a\psi\| + \|a^\dagger\psi\|) < \infty.$$

Thus $a\psi$ normalizable implies $X\psi$ normalizable. The converse fails, because the two terms in X can cancel in its self-adjoint extension.

Use the sequence suggested in the exercise,

$$\psi_n = C \frac{i^n}{n+1}. \quad (43.10)$$

Its absolute values are the same as in the preceding example, so $\psi \in \ell^2$ but $\psi \notin \mathcal{D}(a)$. The coefficient of $X\psi$ at level $n \geq 1$, computed from the tridiagonal position matrix, is

$$\begin{aligned} (X\psi)_n &= x_{\text{osc}} \left(\sqrt{n} \psi_{n-1} + \sqrt{n+1} \psi_{n+1} \right) \\ &= C x_{\text{osc}} i^{n-1} \left[\frac{1}{\sqrt{n}} - \frac{\sqrt{n+1}}{n+2} \right]. \end{aligned} \quad (43.11)$$

For large n ,

$$\frac{\sqrt{n+1}}{n+2} = \frac{1}{\sqrt{n}} \left(1 - \frac{3}{2n} + O(n^{-2}) \right),$$

so $(X\psi)_n = O(n^{-3/2})$. The $n = 0$ coefficient is finite as well. The resulting sequence is square summable.

This formal calculation does represent an element of the operator domain. To see it without subtracting two undefined vectors, truncate $\psi^{(M)} = \sum_{n=0}^M \psi_n |n\rangle$. Then $\psi^{(M)} \rightarrow \psi$ in norm. The interior coefficients of $X\psi^{(M)}$ converge to (43.11); the two extra edge coefficients have size $O(M^{-1/2})$ and disappear in norm. Thus $X\psi^{(M)}$ converges to a Hilbert-space vector. Closedness of the self-adjoint position operator implies $\psi \in \mathcal{D}(X)$.

For the same sequence, the two terms in P add instead of cancelling:

$$(P\psi)_n = Cp_{\text{osc}} i^n \left[\frac{1}{\sqrt{n}} + \frac{\sqrt{n+1}}{n+2} \right] = O(n^{-1/2}).$$

Their squared magnitudes are not summable, so $\psi \notin \mathcal{D}(P)$. The identity $X = x_{\text{osc}}(a + a^\dagger)$ can be used as a sum of separate operator actions on their common domain; extending X beyond that domain requires the limiting construction just described.

43.3 Generalized eigenvectors and a test-function space

A formal expansion $\sum_n c_n |n\rangle$ with $\sum_n |c_n|^2 = \infty$ is not a normalized physical state. Some such expansions nevertheless define useful generalized vectors. Choose, for example, the space Φ of sequences that decrease faster than every inverse power of n . If the coefficients c_n grow at most polynomially, then

$$\langle c | \chi \rangle = \sum_n c_n^* \chi_n \quad (43.12)$$

converges for $\chi \in \Phi$ and defines a continuous functional on that test space. This is the sequence version of a distribution.

The inclusions are often summarized as

$$\Phi \subset \mathcal{H} \subset \Phi', \quad (43.13)$$

with the usual conjugation convention when generalized bras are used. In position space, one may take Φ to be the Schwartz space and Φ' its distributional dual. Position and momentum eigenkets will belong to this generalized setting. A divergent norm is therefore a signal to use distributions, not permission to treat an infinite norm as a finite probability normalization. Not every arbitrary nonsummable sequence defines a continuous functional on a chosen test space.

44 Diagonalizing position: Hermite functions

44.1 The position eigenvalue recurrence

Seek a generalized eigenket satisfying

$$X|x\rangle = x|x\rangle, \quad x \in \mathbb{R}, \quad |x\rangle = \sum_{n=0}^{\infty} f_n(x)|n\rangle. \quad (44.1)$$

Projecting onto $\langle n|$ and using (42.26) gives

$$x f_n(x) = \sqrt{\frac{\hbar}{2m\omega}} \left[\sqrt{n} f_{n-1}(x) + \sqrt{n+1} f_{n+1}(x) \right]. \quad (44.2)$$

For $n = 0$, the first term is absent. Once $f_0(x)$ is given, the recurrence determines every subsequent coefficient.

Introduce the oscillator length and dimensionless coordinate

$$\ell = \sqrt{\frac{\hbar}{m\omega}}, \quad u = \frac{x}{\ell}, \quad (44.3)$$

and write

$$f_n(x) = \frac{\mathcal{N}(x)}{\sqrt{2^n n!}} H_n(u), \quad H_0(u) = 1. \quad (44.4)$$

Substitution into (44.2) and cancellation of the common factors reduces it to

$$\boxed{H_{n+1}(u) = 2uH_n(u) - 2nH_{n-1}(u).} \quad (44.5)$$

The first polynomials are

$$H_0 = 1, \quad H_1 = 2u, \quad (44.6)$$

$$H_2 = 4u^2 - 2, \quad H_3 = 8u^3 - 12u, \quad (44.7)$$

$$H_4 = 16u^4 - 48u^2 + 12. \quad (44.8)$$

They have degree n , leading coefficient 2^n , and parity $H_n(-u) = (-1)^n H_n(u)$.

44.2 Rodrigues' formula and the generating function

The recurrence is solved by the physicists' Hermite polynomials,

$$H_n(u) = (-1)^n e^{u^2} \frac{d^n}{du^n} e^{-u^2}. \quad (44.9)$$

A useful way to derive their identities is to Taylor expand $e^{-(u-t)^2}$ about u :

$$\begin{aligned} e^{2ut-t^2} &= e^{u^2} e^{-(u-t)^2} \\ &= e^{u^2} \sum_{n=0}^{\infty} \frac{(-t)^n}{n!} \frac{d^n}{du^n} e^{-u^2} \\ &= \sum_{n=0}^{\infty} H_n(u) \frac{t^n}{n!}. \end{aligned} \quad (44.10)$$

Differentiating with respect to u gives

$$H'_n(u) = 2nH_{n-1}(u). \quad (44.11)$$

Differentiating with respect to t instead gives $H_{n+1} = 2uH_n - H'_n$, which is precisely (44.5). Combining the two identities yields

$$H''_n - 2uH'_n + 2nH_n = 0. \quad (44.12)$$

This differential equation will turn into the stationary oscillator Schrödinger equation after multiplying by the Gaussian factor.

44.3 Orthogonality, with its normalization proved

Consider

$$I_{mn} = \int_{-\infty}^{\infty} e^{-u^2} H_m(u) H_n(u) du.$$

Using (44.9) for H_n and integrating by parts n times,

$$\begin{aligned} I_{mn} &= (-1)^n \int H_m(u) \frac{d^n}{du^n} e^{-u^2} du \\ &= \int H_m^{(n)}(u) e^{-u^2} du. \end{aligned} \quad (44.13)$$

All boundary terms vanish, since a polynomial times a Gaussian tends to zero at either end. If $m < n$, the derivative is zero. If $m = n$, the leading coefficient gives $H_n^{(n)} = 2^n n!$, and the Gaussian integral is $\sqrt{\pi}$. Symmetry in m, n handles the case $m > n$. Thus

$$\boxed{\int_{-\infty}^{\infty} e^{-u^2} H_m(u) H_n(u) du = \sqrt{\pi} 2^n n! \delta_{mn}.} \quad (44.14)$$

To obtain a continuous resolution of the identity, require

$$\int_{-\infty}^{\infty} |x\rangle \langle x| dx = \mathbf{1}.$$

Its number-basis matrix elements are

$$\int f_m(x) f_n(x)^* dx = \delta_{mn}.$$

Comparing with (44.14) after $x = \ell u$ fixes

$$|\mathcal{N}(x)|^2 = \frac{1}{\sqrt{\pi}\ell} e^{-x^2/\ell^2}. \quad (44.15)$$

There remains a position-dependent phase. The generalized ket can be written as

$$|x\rangle = \frac{e^{i\phi(x)}}{\pi^{1/4}\sqrt{\ell}} e^{-x^2/(2\ell^2)} \sum_{n=0}^{\infty} \frac{H_n(x/\ell)}{\sqrt{2^n n!}} |n\rangle. \quad (44.16)$$

Its conjugate coefficient is the wavefunction of the number state:

$$\varphi_n(x) = \langle x | n \rangle = \frac{e^{-i\phi(x)}}{\pi^{1/4}\sqrt{2^n n! \ell}} H_n(x/\ell) e^{-x^2/(2\ell^2)}. \quad (44.17)$$

The phase convention will be fixed by translations below.

44.4 Completeness of the Hermite functions

Orthogonality alone does not imply that a family spans its Hilbert space. For completeness, define the dimensionless Hermite functions

$$h_n(u) = \frac{H_n(u)e^{-u^2/2}}{\pi^{1/4}\sqrt{2^n n!}}. \quad (44.18)$$

Suppose $g \in L^2(\mathbb{R})$ is orthogonal to every h_n . Because the polynomials H_n span all polynomials, this implies

$$\int g(u)e^{-u^2/2}u^n du = 0 \quad (n = 0, 1, 2, \dots).$$

Consider

$$F(z) = \int_{-\infty}^{\infty} g(u)e^{-u^2/2}e^{zu} du. \quad (44.19)$$

For z in any bounded subset of the complex plane, Cauchy–Schwarz and the Gaussian decay control this integral and all its derivatives. Hence F is entire, and the moment conditions say $F^{(n)}(0) = 0$ for every n . Its Taylor series is therefore identically zero. Setting $z = -ik$ shows that the Fourier transform of the integrable function $g(u)e^{-u^2/2}$ vanishes. Uniqueness of the Fourier transform implies $g = 0$ almost everywhere. Thus $\{h_n\}$ is a complete orthonormal basis of $L^2(\mathbb{R})$.

This proves that the map

$$(\psi_0, \psi_1, \dots) \mapsto \psi(x) = \sum_{n=0}^{\infty} \psi_n \varphi_n(x) \quad (44.20)$$

is unitary from ℓ^2 to $L^2(\mathbb{R})$. It also proves that the oscillator number states exhaust the usual one-dimensional Hilbert space, rather than merely giving a subset of the energy spectrum.

The completeness kernel is the distributional identity

$$\boxed{\sum_{n=0}^{\infty} \frac{e^{-(u^2+v^2)/2}}{\sqrt{\pi} 2^n n!} H_n(u) H_n(v) = \delta(u-v).} \quad (44.21)$$

It means that integrating the kernel against a suitable test function reproduces that function. It is not an assertion that a pointwise convergent ordinary function equals a delta function.

44.5 Worked exercise: generalized orthogonality and independence

Insert the number-basis identity between two position kets and use (44.21), with $u = x/\ell$, $v = y/\ell$:

$$\begin{aligned} \langle x | y \rangle &= \sum_n \langle x | n \rangle \langle n | y \rangle \\ &= e^{i[\phi(y)-\phi(x)]} \frac{1}{\ell} \delta\left(\frac{x-y}{\ell}\right) = \delta(x-y). \end{aligned} \quad (44.22)$$

The phase factor equals one on the support of the delta function, and $\ell^{-1}\delta((x-y)/\ell) = \delta(x-y)$.

If a square-integrable coefficient function satisfies $\int f(x)|x\rangle dx = 0$, project onto every $\langle n|$:

$$\int \langle n | x \rangle f(x) dx = 0 \quad \text{for all } n.$$

Completeness of the Hermite functions implies $f = 0$ almost everywhere. This is the relevant continuous-basis independence statement. Changes at isolated points do not change an L^2 vector or any of these integrals. The same result follows directly by taking the overlap with $\langle y|$ and using (44.22).

There is no contradiction between the countable number basis and the uncountable family of position labels. The number states are actual Hilbert-space vectors. The position kets are generalized vectors outside that space, with delta normalization rather than finite norm.

45 Position measurements and projection-valued measures

45.1 The wavefunction as a representation of the state

The position-space wavefunction is

$$\psi(x) = \langle x | \psi \rangle, \quad |\psi\rangle = \int_{\mathbb{R}} \psi(x)|x\rangle dx. \quad (45.1)$$

It specifies the components of an abstract state in the position representation. The vector $|\psi\rangle$ and the function $\psi(x)$ carry the same information after a basis convention has been chosen, but they are not the same kind of object. Changing the basis changes components without changing the physical state.

The continuous resolution of identity gives

$$\langle \psi | \chi \rangle = \int_{\mathbb{R}} \psi(x)^* \chi(x) dx, \quad \|\psi\|^2 = \int_{\mathbb{R}} |\psi(x)|^2 dx. \quad (45.2)$$

Thus the position representation is $L^2(\mathbb{R})$, with functions identified when they agree almost everywhere. Its scalar product agrees with the ℓ^2 scalar product of the number-state coefficients.

Projecting $X|\psi\rangle$ onto $\langle x|$ gives

$$\langle x | X\psi \rangle = x\psi(x). \quad (45.3)$$

In particular, for a suitable function f and a state with the required finite moment,

$$\langle f(X) \rangle_{\psi} = \int_{\mathbb{R}} f(x) |\psi(x)|^2 dx. \quad (45.4)$$

The self-adjoint multiplication operator X has domain

$$\mathcal{D}(X) = \{\psi \in L^2(\mathbb{R}) : x\psi(x) \in L^2(\mathbb{R})\}. \quad (45.5)$$

An exact position eigenfunction would have support at one point, which has zero measure. There is therefore no nonzero L^2 eigenvector of X ; its spectrum is continuous.

45.2 Finite-resolution position measurements

A detector resolves a region R , such as an interval $[a, b]$, rather than an exact point. The corresponding orthogonal projection is

$$\Pi_R = \int_R |x\rangle\langle x| dx. \quad (45.6)$$

In the wavefunction representation,

$$(\Pi_R \psi)(x) = \mathbf{1}_R(x) \psi(x), \quad (45.7)$$

where $\mathbf{1}_R$ is the indicator of the region. The probability is

$$\Pr(X \in R) = \langle \psi | \Pi_R | \psi \rangle = \int_R |\psi(x)|^2 dx. \quad (45.8)$$

Thus $|\psi(x)|^2$ is a probability *density*, not the probability at an exact point. In one dimension it has units of inverse length, and $\psi(x)$ has units of length to the power $-1/2$.

For a small interval $[x, x + \epsilon]$ at a continuity point,

$$\Pr(x \leq X \leq x + \epsilon) = |\psi(x)|^2 \epsilon + o(\epsilon).$$

Conditioned on a successful ideal yes/no measurement of R with nonzero probability, the normalized state is

$$|\psi_R\rangle = \frac{\Pi_R |\psi\rangle}{\sqrt{\Pr(X \in R)}}.$$

It is still a Hilbert-space state for a finite interval; it need not lie in the domain of momentum or energy after an ideal sharp-edged projection.

45.3 Worked exercise: the PVM identities

Multiplication by an indicator immediately proves

$$\Pi_R^\dagger = \Pi_R, \quad \Pi_R^2 = \Pi_R, \quad (45.9)$$

$$\Pi_{R_1} \Pi_{R_2} = \Pi_{R_1 \cap R_2}, \quad \Pi_{\mathbb{R}} = \mathbf{1}, \quad (45.10)$$

$$\Pi_{R_1 \cup R_2} = \Pi_{R_1} + \Pi_{R_2} \quad (R_1 \cap R_2 = \emptyset). \quad (45.11)$$

For example, $\mathbf{1}_{R_1} \mathbf{1}_{R_2} = \mathbf{1}_{R_1 \cap R_2}$ gives the product identity. More generally, if the regions R_j are pairwise disjoint,

$$\Pi_{\cup_j R_j} \psi = \sum_j \Pi_{R_j} \psi, \quad (45.12)$$

where the sum converges in norm for every state. Its norm convergence follows from the orthogonality of the disjoint pieces and $\sum_j \|\Pi_{R_j} \psi\|^2 \leq \|\psi\|^2$. These properties define an orthogonal **projection-valued measure**, extending a complete set of discrete spectral projectors to a continuous spectrum.

45.4 Why the density of a PVM is not a point projector

The formal symbol

$$\pi(x) = |x\rangle\langle x| \quad (45.13)$$

is an operator-valued distribution. It is not an orthogonal projector onto an ordinary one-dimensional subspace of L^2 . In particular, the formal factor $\langle x | x \rangle = \delta(0)$ is undefined; it must not be used to calculate $\pi(x)^2$ as a product of ordinary operators.

Integration against a bounded measurable function w gives an actual bounded operator,

$$\pi[w] = \int_{\mathbb{R}} w(x) |x\rangle\langle x| dx = w(X), \quad (w(X)\psi)(x) = w(x)\psi(x). \quad (45.14)$$

Its norm is at most the essential supremum of $|w|$; choosing $w = \mathbf{1}_R$ gives Π_R . For unbounded w , the multiplication operator instead has the natural domain $\{\psi : w\psi \in L^2\}$. This is how the same spectral measure defines both bounded projections and unbounded observables.

45.5 Example: the oscillator ground-state position distribution

From (44.17) and $H_0 = 1$,

$$\psi_0(x) = \frac{1}{\pi^{1/4}\sqrt{\ell}} e^{-x^2/(2\ell^2)} e^{-i\phi(x)}, \quad |\psi_0(x)|^2 = \frac{1}{\sqrt{\pi}\ell} e^{-x^2/\ell^2}. \quad (45.15)$$

The density is Gaussian, with mean zero and variance $\ell^2/2$:

$$\langle X \rangle_0 = 0, \quad (\Delta X_0)^2 = \frac{\hbar}{2m\omega}. \quad (45.16)$$

The variance is the square of the standard deviation ΔX_0 . For example,

$$\Pr(-L \leq X \leq L) = \operatorname{erf}(L/\ell), \quad \operatorname{erf}(z) = \frac{2}{\sqrt{\pi}} \int_0^z e^{-t^2} dt. \quad (45.17)$$

46 Momentum, translations, and phase conventions

46.1 Commutators with functions of momentum

The identity $[A, BC] = [A, B]C + B[A, C]$ gives, by induction,

$$[X, P^n] = i\hbar n P^{n-1}. \quad (46.1)$$

For a power series that converges on a suitable common domain,

$$[X, f(P)] = i\hbar f'(P). \quad (46.2)$$

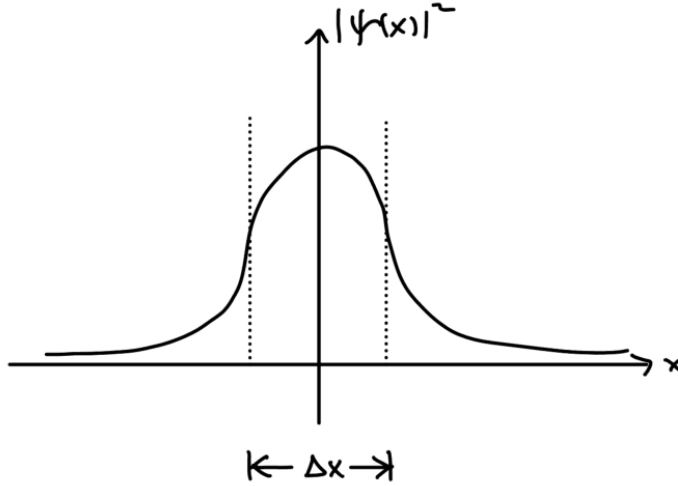


Figure 48: The Gaussian position density of the oscillator ground state. The width is set by $\Delta X = \sqrt{\hbar/(2m\omega)}$.

This identity provides an algebraic route from the canonical commutator to spatial translations.

Define the unitary translation group by

$$T(b) = e^{-ibP/\hbar}, \quad b \in \mathbb{R}. \quad (46.3)$$

Self-adjointness of P gives

$$T(0) = \mathbf{1}, \quad T(b)T(c) = T(b+c), \quad T(b)^\dagger = T(-b) = T(b)^{-1}. \quad (46.4)$$

Equation (46.2) gives

$$[X, T(b)] = bT(b), \quad XT(b) = T(b)(X + b\mathbf{1}). \quad (46.5)$$

Thus

$$X T(b)|x\rangle = (x + b)T(b)|x\rangle.$$

The translated ket belongs to the generalized position eigenspace at $x + b$. For one spinless particle on a line this eigenspace has multiplicity one, so

$$T(b)|x\rangle = \lambda(x, b)|x + b\rangle. \quad (46.6)$$

Preservation of the delta normalization by $T(b)$ implies $|\lambda(x, b)| = 1$.

46.2 Fixing the relative phases of position kets

Choose a phase for the generalized position ket at the origin, $|x = 0\rangle$, and define all others by

$$|x\rangle = T(x)|x = 0\rangle. \quad (46.7)$$

Here $|x = 0\rangle$ is a position eigenket, not the oscillator vacuum $|0\rangle$. The group law then gives

$$T(b)|x\rangle = T(b)T(x)|x = 0\rangle = |x + b\rangle. \quad (46.8)$$

No position-dependent phase factor remains in this translation law.

Taking the adjoint and acting on a state gives

$$\langle x|T(b) = \langle x - b|, \quad \langle x|T(b)\psi\rangle = \psi(x - b). \quad (46.9)$$

A positive translation of the state shifts its density to the right: $|\psi(x - b)|^2$ is centered a distance b to the right of $|\psi(x)|^2$. The opposite sign in the argument is required for this physical shift.

Differentiate (46.9) at $b = 0$. Since $T'(0) = -iP/\hbar$,

$$-\frac{i}{\hbar}\langle x|P\psi\rangle = \left.\frac{\partial}{\partial b}\psi(x - b)\right|_{b=0} = -\frac{d\psi}{dx}. \quad (46.10)$$

Therefore the momentum operator in this convention is

$$(P\psi)(x) = -i\hbar\frac{d\psi}{dx}. \quad (46.11)$$

Momentum is the generator of spatial translations, just as the Hamiltonian is the generator of time evolution.

46.3 What changes under a position-dependent basis phase

If the basis is relabeled by $|x\rangle_{\text{new}} = e^{i\alpha(x)}|x\rangle$, then the same abstract state has components $\psi_{\text{new}} = e^{-i\alpha(x)}\psi$. The momentum representation changes along with the components:

$$\begin{aligned} (P_{\text{new}}\psi_{\text{new}})(x) &= e^{-i\alpha(x)}\left(-i\hbar\frac{d}{dx}\right)[e^{i\alpha(x)}\psi_{\text{new}}(x)] \\ &= \left(-i\hbar\frac{d}{dx} + \hbar\alpha'(x)\right)\psi_{\text{new}}(x). \end{aligned} \quad (46.12)$$

All physical predictions remain unchanged by a consistent basis change. Once the translation convention is fixed, changing the phase profile of a wavefunction while keeping the operators fixed generally changes the physical state. A constant overall phase still has no observable effect.

46.4 The momentum domain on the line

The self-adjoint momentum operator on $L^2(\mathbb{R})$ has domain

$$\mathcal{D}(P) = \{\psi \in L^2(\mathbb{R}) : \psi' \text{ exists weakly and } \psi' \in L^2(\mathbb{R})\}. \quad (46.13)$$

For sufficiently smooth functions with suitable decay, integration by parts shows its symmetry:

$$\begin{aligned}\int \phi^*(-i\hbar\psi') dx &= [-i\hbar\phi^*\psi]_{-\infty}^{\infty} + \int (-i\hbar\phi')^*\psi dx \\ &= \langle P\phi, \psi \rangle.\end{aligned}\tag{46.14}$$

The full self-adjoint statement follows also from the unitary Fourier transform below: P becomes multiplication by the real variable p on its maximal multiplication domain. Boundary conditions would matter if the configuration space were a finite interval or a circle; they cannot be ignored merely because the local differential expression looks the same.

On the common Schwartz domain,

$$(XP - PX)\psi = x(-i\hbar\psi') + i\hbar(x\psi)' = i\hbar\psi,$$

which verifies the canonical commutator directly.

47 Momentum eigenstates and the Fourier transform

47.1 Plane waves and delta normalization

For a generalized momentum eigenket, $P|p\rangle = p|p\rangle$, define $\psi_p(x) = \langle x | p \rangle$. Equation (46.11) gives

$$-i\hbar \frac{d\psi_p}{dx} = p\psi_p, \quad \psi_p(x) = \mathcal{M}(p)e^{ipx/\hbar}.\tag{47.1}$$

To impose the continuous resolution of identity in momentum space, take position matrix elements:

$$\int_{\mathbb{R}} \langle x | p \rangle \langle p | y \rangle dp = \delta(x - y).\tag{47.2}$$

The Fourier identity in these units is

$$\int_{\mathbb{R}} e^{ip(x-y)/\hbar} dp = 2\pi\hbar\delta(x - y).\tag{47.3}$$

It fixes the magnitude of $\mathcal{M}$. Choosing its phase real and positive gives

$$\boxed{\langle x | p \rangle = \frac{1}{\sqrt{2\pi\hbar}}e^{ipx/\hbar}, \quad \langle p | x \rangle = \frac{1}{\sqrt{2\pi\hbar}}e^{-ipx/\hbar}.}\tag{47.4}$$

The momentum normalization follows by integrating over position:

$$\langle p | q \rangle = \int \langle p | x \rangle \langle x | q \rangle dx = \delta(p - q), \quad \int |p\rangle \langle p| dp = \mathbf{1}.\tag{47.5}$$

Plane waves do not decay at spatial infinity, so they are generalized states rather than normalizable states on the whole line.

47.2 The two wavefunction representations

Define the momentum-space wavefunction by

$$\tilde{\psi}(p) = \langle p | \psi \rangle. \quad (47.6)$$

Insert the position or momentum identity to obtain the unitary transform pair

$$\tilde{\psi}(p) = \frac{1}{\sqrt{2\pi\hbar}} \int_{\mathbb{R}} e^{-ipx/\hbar} \psi(x) dx, \quad (47.7)$$

$$\psi(x) = \frac{1}{\sqrt{2\pi\hbar}} \int_{\mathbb{R}} e^{ipx/\hbar} \tilde{\psi}(p) dp. \quad (47.8)$$

They satisfy Parseval's identity,

$$\int \psi(x)^* \chi(x) dx = \int \tilde{\psi}(p)^* \tilde{\chi}(p) dp. \quad (47.9)$$

For example, substitute both inverse transforms on the left. The x integral gives $2\pi\hbar\delta(p-q)$ and removes one momentum integral, leaving the right side. Normalization in either representation is thus normalization of the same state.

In momentum representation,

$$(P\tilde{\psi})(p) = p\tilde{\psi}(p), \quad (X\tilde{\psi})(p) = i\hbar \frac{d\tilde{\psi}}{dp}. \quad (47.10)$$

The second identity follows by differentiating (47.7) with respect to p . Position and momentum exchange the roles of multiplication and differentiation, with the signs required by $[X, P] = i\hbar\mathbf{1}$.

Momentum measurement probabilities use the same PVM construction:

$$\Pr(P \in I) = \int_I |\tilde{\psi}(p)|^2 dp, \quad \langle f(P) \rangle = \int f(p) |\tilde{\psi}(p)|^2 dp.$$

The dimensions of $\tilde{\psi}$ are momentum to the power $-1/2$.

47.3 Exact position and exact momentum under Fourier transformation

For the generalized momentum eigenstate $|q\rangle$,

$$\psi_q(x) = \frac{e^{iqx/\hbar}}{\sqrt{2\pi\hbar}}, \quad \tilde{\psi}_q(p) = \delta(p - q). \quad (47.11)$$

For the generalized position eigenstate $|y\rangle$,

$$\psi_y(x) = \delta(x - y), \quad \tilde{\psi}_y(p) = \frac{e^{-ipy/\hbar}}{\sqrt{2\pi\hbar}}. \quad (47.12)$$

Thus a plane wave and a delta function are Fourier descriptions of the same type of nonnormalizable idealization. Neither is a physical normalized state on the line. A finite-width wave packet provides the normalizable approximation used in probability calculations.

A spatial translation multiplies the momentum wavefunction by a phase:

$$\widetilde{T(b)\psi}(p) = e^{-ibp/\hbar} \widetilde{\psi}(p).$$

Conversely, multiplying $\psi(x)$ by $e^{iqx/\hbar}$ shifts its momentum distribution by q . These two facts connect the phase of a wavefunction with its mean position and momentum.

48 Gaussian wave packets and the uncertainty bound

48.1 A normalized packet with specified means and width

Let $\sigma > 0$, $x_0, p_0 \in \mathbb{R}$, and choose

$$\psi(x) = \frac{1}{(2\pi\sigma^2)^{1/4}} \exp\left[-\frac{(x-x_0)^2}{4\sigma^2} + \frac{ip_0x}{\hbar}\right]. \quad (48.1)$$

The notation in the source e^{ikx} is recovered by setting $p_0 = \hbar k$. The density is

$$|\psi(x)|^2 = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-x_0)^2/(2\sigma^2)}. \quad (48.2)$$

With $y = (x - x_0)/\sigma$, the Gaussian integral gives normalization, oddness gives $\int y e^{-y^2/2} dy = 0$, and integration by parts gives $\int y^2 e^{-y^2/2} dy = \sqrt{2\pi}$. Consequently,

$$\langle X \rangle = x_0, \quad \langle X^2 \rangle = x_0^2 + \sigma^2, \quad \Delta X = \sigma. \quad (48.3)$$

48.2 The momentum Gaussian, derived by completing the square

Insert (48.1) into (47.7) and let $y = x - x_0$, $q = p - p_0$:

$$\widetilde{\psi}(p) = \frac{e^{-iqx_0/\hbar}}{\sqrt{2\pi\hbar}(2\pi\sigma^2)^{1/4}} \int_{\mathbb{R}} \exp\left[-\frac{y^2}{4\sigma^2} - \frac{iqy}{\hbar}\right] dy. \quad (48.4)$$

The exponent can be completed as

$$-\frac{y^2}{4\sigma^2} - \frac{iqy}{\hbar} = -\frac{1}{4\sigma^2} \left(y + \frac{2i\sigma^2 q}{\hbar}\right)^2 - \frac{\sigma^2 q^2}{\hbar^2}.$$

Equivalently, apply the convergent complex Gaussian integral $\int e^{-ay^2+by} dy = \sqrt{\pi/a} e^{b^2/(4a)}$ with $a > 0$. The result is

$$\boxed{\widetilde{\psi}(p) = \left(\frac{2\sigma^2}{\pi\hbar^2}\right)^{1/4} \exp\left[-\frac{\sigma^2(p-p_0)^2}{\hbar^2} - \frac{i(p-p_0)x_0}{\hbar}\right]}. \quad (48.5)$$

Its squared magnitude is a normalized Gaussian centered at p_0 with variance $\hbar^2/(4\sigma^2)$:

$$\langle P \rangle = p_0, \quad \Delta P = \frac{\hbar}{2\sigma}, \quad \boxed{\Delta X \Delta P = \frac{\hbar}{2}}. \quad (48.6)$$

Increasing the spatial width decreases the momentum width. The phases in the two representations encode the corresponding shifts of the means.

48.3 The uncertainty inequality and its equality conditions

Let $\delta X = X - \langle X \rangle$ and $\delta P = P - \langle P \rangle$, for a state with finite variances and the common-domain assumptions needed for the following products. Cauchy–Schwarz gives

$$(\Delta X)^2(\Delta P)^2 \geq |\langle \psi | \delta X \delta P | \psi \rangle|^2. \quad (48.7)$$

The complex expectation can be separated into its real and imaginary parts:

$$\langle \psi | \delta X \delta P | \psi \rangle = \frac{1}{2} \langle \{ \delta X, \delta P \} \rangle + \frac{1}{2} \langle [X, P] \rangle = C_{XP} + \frac{i\hbar}{2}, \quad (48.8)$$

where C_{XP} is the symmetrized covariance. Thus

$$(\Delta X)^2(\Delta P)^2 \geq C_{XP}^2 + \frac{\hbar^2}{4} \geq \frac{\hbar^2}{4}. \quad (48.9)$$

The product bound is the Heisenberg inequality,

$$\boxed{\Delta X \Delta P \geq \frac{\hbar}{2}}. \quad (48.10)$$

For equality in the simple product bound, equality must hold in Cauchy–Schwarz and the covariance must vanish. The two Hilbert-space vectors must therefore obey

$$(P - p_0)|\psi\rangle = ic(X - x_0)|\psi\rangle, \quad c > 0. \quad (48.11)$$

The sign follows from $\text{Im}\langle \delta X \delta P \rangle = \hbar/2 > 0$. Taking the expectation after multiplication by δX gives $c(\Delta X)^2 = \hbar/2$.

48.4 Worked exercise: every minimum-product state is Gaussian

In position representation, (48.11) is a first-order ODE:

$$-i\hbar\psi'(x) - p_0\psi(x) = ic(x - x_0)\psi(x).$$

Divide by $-i\hbar$ to obtain

$$\frac{\psi'}{\psi} = \frac{ip_0}{\hbar} - \frac{c}{\hbar}(x - x_0).$$

Integration gives

$$\psi(x) = A \exp \left[-\frac{c}{2\hbar}(x - x_0)^2 + \frac{ip_0 x}{\hbar} \right]. \quad (48.12)$$

Normalizability requires $c > 0$. Using $c = \hbar/(2\sigma^2)$ and fixing $|A|$ recovers exactly (48.1), up to a constant phase. Conversely, differentiating that Gaussian verifies (48.11), so it saturates the bound.

A Gaussian with an additional quadratic phase generally has nonzero C_{XP} . Such a state can saturate the stronger first inequality in (48.9) while having $\Delta X \Delta P > \hbar/2$. Thus the minimum *product* condition selects a Gaussian with only a linear spatial phase in the fixed translation convention.

49 The oscillator ground state and zero-point motion

49.1 Fixing the remaining phase of the oscillator functions

The Hermite construction initially allowed a phase $e^{-i\phi(x)}$ in all oscillator wavefunctions. In the translation convention, $P = -i\hbar\partial_x$. Apply momentum to the vacuum using ladder operators:

$$\langle x|P|0\rangle = ip_{\text{osc}}\langle x|1\rangle = ip_{\text{osc}}\sqrt{2}\frac{x}{\ell}\psi_0(x) = im\omega x\psi_0(x). \quad (49.1)$$

On the other hand, differentiating $\psi_0 \propto e^{-x^2/(2\ell^2)}e^{-i\phi(x)}$ gives

$$-i\hbar\psi_0' = [im\omega x - \hbar\phi'(x)]\psi_0(x). \quad (49.2)$$

Agreement requires $\phi'(x) = 0$. Choose the remaining constant phase to be zero. The fully normalized oscillator wavefunctions are therefore

$$\varphi_n(x) = \frac{1}{\pi^{1/4}\sqrt{2^n n! \ell}} H_n(x/\ell) e^{-x^2/(2\ell^2)}. \quad (49.3)$$

Using (44.12), differentiate $e^{-u^2/2}H_n(u)$ twice to find

$$\left(-\frac{1}{2}\frac{d^2}{du^2} + \frac{1}{2}u^2 \right) [e^{-u^2/2}H_n(u)] = \left(n + \frac{1}{2} \right) e^{-u^2/2}H_n(u).$$

After $x = \ell u$, this is the stationary Schrödinger equation with the energy (42.14).

The ground state can also be obtained immediately from $a|0\rangle = 0$:

$$\left(x + \ell^2 \frac{d}{dx} \right) \psi_0 = 0, \quad \psi_0(x) = \frac{1}{\pi^{1/4}\sqrt{\ell}} e^{-x^2/(2\ell^2)}.$$

Its uncertainties are

$$\Delta X_0 = \sqrt{\frac{\hbar}{2m\omega}}, \quad \Delta P_0 = \sqrt{\frac{m\hbar\omega}{2}}. \quad (49.4)$$

49.2 Worked exercise: excited states are not Gaussian packets

For $n \geq 1$, (49.3) contains the nonconstant factor $H_n(x/\ell)$. It cannot be absorbed into the exponential of a real negative quadratic plus a linear imaginary phase. One can prove this without using the locations of the Hermite zeros. Every normalized Gaussian packet of the form (48.1) saturates $\Delta X \Delta P = \hbar/2$, whereas (42.29) gives

$$\Delta X_n \Delta P_n = \hbar \left(n + \frac{1}{2} \right) > \frac{\hbar}{2} \quad (n \geq 1).$$

Thus no excited number state is such a Gaussian packet. More generally, Hermite functions at excited levels have nodes, while a nonzero Gaussian with any finite quadratic phase has no zeros on the real line.

49.3 A lower bound on energy without ladder operators

For any normalized state with finite oscillator energy, write the second moments as variance plus squared mean:

$$\begin{aligned} \langle H \rangle &= \frac{(\Delta P)^2 + p_0^2}{2m} + \frac{1}{2} m \omega^2 [(\Delta X)^2 + x_0^2] \\ &\geq \frac{(\Delta P)^2}{2m} + \frac{1}{2} m \omega^2 (\Delta X)^2 \\ &\geq \frac{\hbar^2}{8m(\Delta X)^2} + \frac{1}{2} m \omega^2 (\Delta X)^2. \end{aligned} \quad (49.5)$$

The first inequality discards nonnegative energies associated with the means. The second uses $\Delta P \geq \hbar/(2\Delta X)$.

Set $s = \Delta X > 0$. The derivative of the final expression is

$$\frac{d}{ds} \left(\frac{\hbar^2}{8ms^2} + \frac{1}{2} m \omega^2 s^2 \right) = -\frac{\hbar^2}{4ms^3} + m \omega^2 s.$$

The unique minimum occurs at

$$s^2 = \frac{\hbar}{2m\omega}. \quad (49.6)$$

At this width, each of the two terms is $\hbar\omega/4$, giving

$$\boxed{\langle H \rangle \geq \frac{1}{2} \hbar \omega.} \quad (49.7)$$

Every inequality can be saturated simultaneously: choose $x_0 = p_0 = 0$, a minimum-product Gaussian, and the width just found. This recovers the ground-state energy and wavefunction without finding any excited states.

The physical competition is explicit. Making the packet narrower reduces its potential energy near the bottom of the well, but increases its momentum uncertainty and therefore its kinetic energy. At the optimum,

$$\left\langle \frac{P^2}{2m} \right\rangle_0 = \left\langle \frac{1}{2} m \omega^2 X^2 \right\rangle_0 = \frac{1}{4} \hbar \omega. \quad (49.8)$$

There is no normalizable state with both zero position uncertainty and zero momentum uncertainty.

50 The Schrödinger representation in several dimensions

50.1 The Hilbert space is not fixed by the choice of potential

The oscillator provided a convenient algebraic construction of the position and momentum representation. Once this representation is identified, changing the potential does not change the Hilbert space of a spinless particle moving on the full line. The operators act as

$$\mathcal{H} = L^2(\mathbb{R}), \quad X\psi = x\psi, \quad P\psi = -i\hbar\psi', \quad (50.1)$$

with their appropriate domains. A Hamiltonian $P^2/(2m) + V(X)$ changes the dynamics and usually changes the energy basis, but position and momentum remain the same kinematic observables. Internal spin adds a tensor factor; changing the configuration space or its boundary conditions is a separate change to the physical problem.

50.2 The scope of the Stone–von Neumann theorem

The exponentiated canonical relations avoid products of unbounded operators. Along with $T(b) = e^{-ibP/\hbar}$, introduce $M(q) = e^{iqX/\hbar}$. In position representation,

$$(T(b)\psi)(x) = \psi(x - b), \quad (M(q)\psi)(x) = e^{iqx/\hbar}\psi(x).$$

Their composition gives the Weyl relation

$$T(b)M(q) = e^{-ibq/\hbar}M(q)T(b). \quad (50.2)$$

For a finite number of canonical degrees of freedom, a strongly continuous irreducible unitary representation of these relations with the same nonzero $\hbar$ is unique up to a unitary transformation. This is the **Stone–von Neumann theorem**.

The hypotheses matter. Irreducibility rules out extra multiplicity factors such as $\mathcal{H}_{\text{int}}$. Strong continuity ensures that the unitary groups have self-adjoint generators. The bare formal commutator on an unspecified domain is not by itself the full theorem, and its finite-degree-of-freedom uniqueness statement does not extend unchanged to quantum fields with infinitely many degrees of freedom.

50.3 Joint position and momentum bases in dimension d

For a spinless particle in $\mathbb{R}^d$, the canonical relations are

$$[X_i, X_j] = 0, \quad [P_i, P_j] = 0, \quad [X_i, P_j] = i\hbar\delta_{ij}\mathbf{1}. \quad (50.3)$$

The strongly commuting position components have a joint spectral basis, written $|\mathbf{x}\rangle = |x_1, \dots, x_d\rangle$:

$$X_i|\mathbf{x}\rangle = x_i|\mathbf{x}\rangle, \quad (50.4)$$

$$\langle \mathbf{x} | \mathbf{y} \rangle = \delta^{(d)}(\mathbf{x} - \mathbf{y}), \quad (50.5)$$

$$\int_{\mathbb{R}^d} |\mathbf{x}\rangle \langle \mathbf{x}| d^d x = \mathbf{1}. \quad (50.6)$$

The state space is $L^2(\mathbb{R}^d)$, and

$$\psi(\mathbf{x}) = \langle \mathbf{x} | \psi \rangle, \quad \mathbf{X}\psi = \mathbf{x}\psi, \quad \mathbf{P}\psi = -i\hbar\nabla\psi. \quad (50.7)$$

For example,

$$\langle \mathbf{X} \rangle = \int_{\mathbb{R}^d} \mathbf{x} |\psi(\mathbf{x})|^2 d^d x.$$

The generalized overlap and Fourier pair are

$$\langle \mathbf{x} | \mathbf{p} \rangle = \frac{e^{i\mathbf{p}\cdot\mathbf{x}/\hbar}}{(2\pi\hbar)^{d/2}}, \quad (50.8)$$

$$\tilde{\psi}(\mathbf{p}) = \frac{1}{(2\pi\hbar)^{d/2}} \int_{\mathbb{R}^d} e^{-i\mathbf{p}\cdot\mathbf{x}/\hbar} \psi(\mathbf{x}) d^d x, \quad (50.9)$$

$$\psi(\mathbf{x}) = \frac{1}{(2\pi\hbar)^{d/2}} \int_{\mathbb{R}^d} e^{i\mathbf{p}\cdot\mathbf{x}/\hbar} \tilde{\psi}(\mathbf{p}) d^d p. \quad (50.10)$$

Each coordinate satisfies its own uncertainty relation $\Delta X_i \Delta P_i \geq \hbar/2$.

51 Schrödinger's wave equation

51.1 From the abstract state equation to a PDE

For a nonrelativistic spinless particle with real scalar potential,

$$H = \frac{\mathbf{P}^2}{2m} + V(\mathbf{X}, t). \quad (51.1)$$

The abstract Schrödinger equation is

$$i\hbar \frac{\partial}{\partial t} |\psi(t)\rangle = H |\psi(t)\rangle. \quad (51.2)$$

In the Schrödinger picture the position basis is time independent. Taking the overlap with $\langle \mathbf{x} |$ therefore gives

$$i\hbar \partial_t \psi(\mathbf{x}, t) = \langle \mathbf{x} | H \psi(t) \rangle.$$

The kinetic term uses $\mathbf{P}^2 = -\hbar^2 \nabla^2$ and the potential acts by multiplication. Hence

$$\boxed{i\hbar \partial_t \psi(\mathbf{x}, t) = -\frac{\hbar^2}{2m} \nabla^2 \psi(\mathbf{x}, t) + V(\mathbf{x}, t) \psi(\mathbf{x}, t).} \quad (51.3)$$

This wave equation is the coordinate representation of the same abstract unitary evolution law. A self-adjoint Hamiltonian and its domain, including any boundary conditions, specify the evolution problem.

51.2 Stationary states and the energy eigenvalue equation

If the potential is time independent and $H|E\rangle = E|E\rangle$, then

$$|\psi(t)\rangle = e^{-iEt/\hbar}|E\rangle, \quad \psi(\mathbf{x}, t) = e^{-iEt/\hbar}\varphi_E(\mathbf{x}). \quad (51.4)$$

Substituting this form into (51.3) gives

$$\left[-\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{x}) \right] \varphi_E(\mathbf{x}) = E\varphi_E(\mathbf{x}). \quad (51.5)$$

The time-independent Schrödinger equation is an eigenvalue equation, not a second independent postulate. A superposition of energies evolves by giving each energy component its own phase. Its probability density can change because of interference, even though the density of a single stationary state is time independent.

52 Probability conservation and probability current

52.1 Deriving the continuity equation

Define the position probability density

$$\rho(\mathbf{x}, t) = |\psi(\mathbf{x}, t)|^2. \quad (52.1)$$

Differentiate and substitute (51.3) and its complex conjugate:

$$\begin{aligned} \partial_t \rho &= \psi^* \partial_t \psi + \psi \partial_t \psi^* \\ &= \frac{i\hbar}{2m} \psi^* \nabla^2 \psi - \frac{iV}{\hbar} |\psi|^2 - \frac{i\hbar}{2m} \psi \nabla^2 \psi^* + \frac{iV}{\hbar} |\psi|^2 \\ &= \frac{i\hbar}{2m} (\psi^* \nabla^2 \psi - \psi \nabla^2 \psi^*). \end{aligned} \quad (52.2)$$

The potential terms cancel because V is real. The vector identity

$$\nabla \cdot (\psi^* \nabla \psi - \psi \nabla \psi^*) = \psi^* \nabla^2 \psi - \psi \nabla^2 \psi^*$$

follows because the two cross terms cancel. Define

$$\mathbf{j} = \frac{\hbar}{2mi} (\psi^* \nabla \psi - \psi \nabla \psi^*) = \frac{\hbar}{m} \text{Im}(\psi^* \nabla \psi). \quad (52.3)$$

Then

$$\partial_t \rho + \nabla \cdot \mathbf{j} = 0. \quad (52.4)$$

This is a local statement of probability conservation.

For a fixed region R , integration and the divergence theorem give

$$\frac{d}{dt} \int_R \rho d^d x = - \int_{\partial R} \mathbf{j} \cdot \mathbf{n} dS. \quad (52.5)$$

The probability inside the region changes by the net current through its boundary. On the whole space, sufficient decay makes the boundary flux vanish and preserves total norm. In a bounded region, a self-adjoint Hamiltonian imposes boundary conditions consistent with the corresponding probability balance.

52.2 A stationary density can carry a current

The equation $\partial_t \rho = 0$ requires $\nabla \cdot \mathbf{j} = 0$, not $\mathbf{j} = 0$. For example, the formal traveling plane wave

$$\psi(\mathbf{x}, t) = A e^{i(\mathbf{p} \cdot \mathbf{x} - Et)/\hbar}$$

has constant density $|A|^2$ but

$$\mathbf{j} = \frac{\mathbf{p}}{m} |A|^2. \quad (52.6)$$

It carries a steady probability flux. A real stationary bound-state wavefunction, on the other hand, has zero current because $\text{Im}(\varphi_E^* \nabla \varphi_E) = 0$. Plane waves on the line are distributional states; a periodic box or a broad normalized packet gives the corresponding normalizable setting.

52.3 Current and the expectation of momentum

Since $-i\hbar \nabla$ is momentum,

$$\mathbf{j}(\mathbf{x}, t) = \frac{1}{m} \text{Re}[\psi(\mathbf{x}, t)^* (-i\hbar \nabla) \psi(\mathbf{x}, t)]. \quad (52.7)$$

Integration gives

$$\int_{\mathbb{R}^d} \mathbf{j}(\mathbf{x}, t) d^d x = \frac{1}{m} \text{Re} \langle \mathbf{P} \rangle = \frac{\langle \mathbf{P} \rangle}{m}, \quad (52.8)$$

where reality follows from self-adjointness and the necessary domain conditions. Equivalently, symmetrizing the integral of $\psi^* (-i\hbar \nabla) \psi$ reproduces the two terms in (52.3).

The same relation follows from the continuity equation:

$$\frac{d}{dt} \langle X_i \rangle = - \int x_i \nabla \cdot \mathbf{j} d^d x = \int j_i d^d x = \frac{\langle P_i \rangle}{m}, \quad (52.9)$$

assuming the surface term vanishes. The quantity $m\mathbf{j}$ behaves as a symmetrized local momentum density in this representation. It is not a positive joint probability distribution assigning a simultaneously sharp position and momentum to the particle.

53 Wavefunction phase and the classical Hamilton–Jacobi limit

53.1 Amplitude and phase variables

In a region where $\rho > 0$, write

$$\psi(\mathbf{x}, t) = R(\mathbf{x}, t) e^{iS(\mathbf{x}, t)/\hbar}, \quad R = \sqrt{\rho} > 0, \quad S \in \mathbb{R}. \quad (53.1)$$

The phase can be chosen locally; nodes and nontrivial phase winding may prevent one global smooth S . Differentiation gives

$$\nabla \psi = e^{iS/\hbar} \left(\nabla R + \frac{i}{\hbar} R \nabla S \right),$$

and therefore

$$\psi^* \nabla \psi = R \nabla R + \frac{i}{\hbar} \rho \nabla S.$$

The first term is real. Equation (52.3) becomes

$$\boxed{\mathbf{j} = \rho \frac{\nabla S}{m}.} \quad (53.2)$$

A spatially varying phase produces current; the density alone cannot specify it. Where $\rho > 0$, the ratio $\mathbf{j}/\rho = \nabla S/m$ is a well-defined probability-flow field. Its existence does not change the quantum restrictions on simultaneous sharp position and momentum measurements.

53.2 Separating the Schrödinger equation into real and imaginary parts

The required derivatives are

$$\partial_t \psi = e^{iS/\hbar} \left(R_t + \frac{i}{\hbar} R S_t \right), \quad (53.3)$$

$$\nabla^2 \psi = e^{iS/\hbar} \left[\nabla^2 R + \frac{2i}{\hbar} \nabla R \cdot \nabla S + \frac{i}{\hbar} R \nabla^2 S - \frac{1}{\hbar^2} R |\nabla S|^2 \right]. \quad (53.4)$$

Insert these into (51.3), cancel $e^{iS/\hbar}$, and divide by R :

$$\begin{aligned} i\hbar \frac{R_t}{R} - S_t &= -\frac{\hbar^2}{2m} \frac{\nabla^2 R}{R} - \frac{i\hbar}{m} \frac{\nabla R}{R} \cdot \nabla S - \frac{i\hbar}{2m} \nabla^2 S \\ &\quad + \frac{|\nabla S|^2}{2m} + V. \end{aligned} \quad (53.5)$$

The imaginary part is

$$R_t = -\frac{1}{m} \nabla R \cdot \nabla S - \frac{R}{2m} \nabla^2 S.$$

Multiply by $2R$ and use $2RR_t = \rho_t$ and $2R\nabla R = \nabla \rho$ to obtain

$$\partial_t \rho + \nabla \cdot \left(\rho \frac{\nabla S}{m} \right) = 0, \quad (53.6)$$

which is the same continuity equation as before.

The real part is

$$\boxed{S_t + \frac{|\nabla S|^2}{2m} + V + Q = 0, \quad Q = -\frac{\hbar^2}{2m} \frac{\nabla^2 \sqrt{\rho}}{\sqrt{\rho}}.} \quad (53.7)$$

These two real equations are exactly equivalent to the Schrödinger equation wherever the amplitude–phase parametrization is regular. The term Q contains the quantum correction to the Hamilton–Jacobi equation.

53.3 The controlled classical limit

If the amplitude and phase vary on scales for which Q is negligible, the leading equation is

$$\boxed{S_t + \frac{|\nabla S|^2}{2m} + V(\mathbf{x}, t) = 0.} \quad (53.8)$$

This is the classical Hamilton–Jacobi equation. The limit is a semiclassical approximation, not a universal permission to set every factor of $\hbar$ to zero: derivatives can become large near nodes, turning points, or interference structures, making Q important even when $\hbar$ is small compared with a characteristic action.

In Hamilton–Jacobi theory,

$$\mathbf{p} = \nabla S, \quad \mathbf{v} = \frac{\mathbf{p}}{m}, \quad (53.9)$$

so (53.2) becomes the classical transport form $\dot{\mathbf{j}} = \rho \mathbf{v}$. Along a classical trajectory,

$$\frac{dS}{dt} = S_t + \dot{\mathbf{x}} \cdot \nabla S = -H + \mathbf{p} \cdot \dot{\mathbf{x}} = L. \quad (53.10)$$

Therefore the principal function can be interpreted, with suitable initial or endpoint data and a chosen trajectory branch, as the accumulated classical action

$$S = \int L dt + \text{constant}. \quad (53.11)$$

This explains the appearance of the phase $e^{iS/\hbar}$ in semiclassical wavefunctions and in the stationary-phase interpretation of quantum propagation.

53.4 Time-independent Hamiltonians and the characteristic function

If H has no explicit time dependence, seek

$$S(\mathbf{x}, t) = W(\mathbf{x}) - Et. \quad (53.12)$$

The Hamilton–Jacobi equation reduces to

$$\frac{|\nabla W|^2}{2m} + V(\mathbf{x}) = E, \quad (53.13)$$

where W is Hamilton’s characteristic function. In one dimension,

$$W'(x) = p(x) = \pm \sqrt{2m[E - V(x)]}, \quad W(x) = \int^x p(x') dx'.$$

Thus the leading oscillatory factor is $e^{iW(x)/\hbar}$, the phase that appears in the WKB approximation. The amplitude is fixed by the next transport equation, rather than by this phase equation alone.

For a more general Hamiltonian operator with a suitable classical principal symbol $H_{\text{cl}}(\mathbf{x}, \mathbf{p}, t)$, substituting $Re^{iS/\hbar}$ and retaining the leading semiclassical terms gives

$$S_t + H_{\text{cl}}(\mathbf{x}, \nabla S, t) = 0. \quad (53.14)$$

At leading order, the transport velocity is $\partial H_{\text{cl}}/\partial \mathbf{p}$ evaluated at $\mathbf{p} = \nabla S$. The exact current formula (52.3) was derived for the quadratic kinetic energy $P^2/(2m)$ and need not hold unchanged for an arbitrary Hamiltonian.

Part VII

Wave Mechanics and Bound States

Chapter 7

54 Stationary states and confinement

54.1 Energy eigenfunctions and their time dependence

For a time-independent scalar Hamiltonian

$$H = \frac{\mathbf{p}^2}{2m} + V(\mathbf{x}),$$

the time-dependent Schrödinger equation is solved by expanding in energy eigenfunctions. A stationary state has the form

$$\Psi_E(\mathbf{x}, t) = e^{-iEt/\hbar} \psi_E(\mathbf{x}), \quad \left[-\frac{\hbar^2}{2m} \nabla^2 + V(\mathbf{x}) \right] \psi_E = E \psi_E.$$

Its probability density is independent of time. The word stationary refers to this density; it does not by itself imply zero probability current.

If the potential and boundary conditions are real, complex conjugation maps an energy eigenfunction to another eigenfunction with the same real energy. Consequently, real and imaginary parts are also solutions, and one can choose a real basis within an eigenspace invariant under complex conjugation. For a real eigenfunction,

$$\mathbf{j} = \frac{\hbar}{m} \text{Im}(\psi^* \nabla \psi) = 0.$$

For a normalizable state in the usual full-space domain, this also gives $\langle \mathbf{p} \rangle = m \int \mathbf{j} d^d x = 0$. A complex basis that simultaneously diagonalizes other conserved observables can be more useful. For example, plane waves diagonalize momentum but are not real.

54.2 The free particle and continuum normalization

For $V = 0$, choose simultaneous eigenstates of H and $\mathbf{p}$:

$$\psi_{\mathbf{p}}(\mathbf{x}) = \frac{1}{(2\pi\hbar)^{d/2}} e^{i\mathbf{p}\cdot\mathbf{x}/\hbar}, \quad E_{\mathbf{p}} = \frac{\mathbf{p}^2}{2m}. \quad (54.1)$$

The normalization is distributional:

$$\int \psi_{\mathbf{p}}^*(\mathbf{x}) \psi_{\mathbf{p}'}(\mathbf{x}) d^d x = \delta^{(d)}(\mathbf{p} - \mathbf{p}').$$

No individual plane wave is square-integrable over all space. Physical normalizable states are wavepackets formed by superposing momenta. For $\mathbf{p} \neq 0$, real combinations of the opposite momenta give sine and cosine energy eigenfunctions. Such combinations no longer diagonalize momentum. Complex conjugation reverses momentum while preserving the free particle energy.

54.3 A periodic box

Let $0 \leq x_i < L_i$ and impose periodic boundary conditions in every direction. Periodicity of a plane wave requires

$$e^{ip_i L_i/\hbar} = 1, \quad p_i = \frac{2\pi\hbar n_i}{L_i}, \quad n_i \in \mathbb{Z}.$$

The normalized eigenfunctions and energies are

$$\psi_{\mathbf{n}}(\mathbf{x}) = \frac{1}{\sqrt{\mathcal{V}}} \exp\left(2\pi i \sum_{i=1}^d \frac{n_i x_i}{L_i}\right), \quad \mathcal{V} = \prod_i L_i, \quad (54.2)$$

$$E_{\mathbf{n}} = \frac{(2\pi\hbar)^2}{2m} \sum_i \frac{n_i^2}{L_i^2}. \quad (54.3)$$

Finite volume makes the eigenfunctions normalizable and the spectrum discrete. A one-dimensional periodic box can represent a particle moving on a ring of circumference $L = 2\pi R$.

For a state with s nonzero components of $\mathbf{n}$, independent sign reversals give 2^s states with equal energy. Equal or specially related side lengths can create additional degeneracy. The ground state $\mathbf{n} = 0$ is unique. Sending every L_i to infinity while keeping $p_i = 2\pi\hbar n_i/L_i$ fixed recovers the continuum.

Example 54.1 (Symmetry behind the sign degeneracy). On a ring, translations $T(a)$ and reflection P commute with the kinetic Hamiltonian, but they satisfy $PT(a)P^{-1} = T(-a)$. They therefore form a noncommutative symmetry group. Translation eigenstates with momenta $\pm p$ are exchanged by reflection and have the same energy. In several dimensions the corresponding translations and coordinate reflections account for the sign multiplets; a cubic box also has coordinate permutations. Degeneracy can be analyzed through these symmetries without assuming that every coincident level is explained by rotations alone.

54.4 A hard-walled box and the momentum domain

For an infinite square well on $0 < x < L$, the boundary conditions are $\psi(0) = \psi(L) = 0$. Inside the well, $\psi'' + k^2\psi = 0$, so $\psi = A \sin kx + B \cos kx$. The first endpoint gives $B = 0$, and the second gives $kL = n\pi$, $n = 1, 2, \dots$. Normalization uses $\int_0^L \sin^2(n\pi x/L) dx = L/2$, yielding

$$\psi_n(x) = \sqrt{\frac{2}{L}} \sin \frac{n\pi x}{L}, \quad E_n = \frac{\pi^2 \hbar^2 n^2}{2mL^2}. \quad (54.4)$$

For a rectangular d -dimensional box, take tensor products:

$$\psi_{\mathbf{n}}(\mathbf{x}) = \prod_i \sqrt{\frac{2}{L_i}} \sin \frac{n_i \pi x_i}{L_i}, \quad E_{\mathbf{n}} = \frac{\pi^2 \hbar^2}{2m} \sum_i \frac{n_i^2}{L_i^2}, \quad n_i \geq 1.$$

The ground energy is positive because confinement prohibits a constant wavefunction.

There is an important distinction between the kinetic Hamiltonian and the formal first derivative $p = -i\hbar\partial_x$. On an interval,

$$\langle f, pg \rangle - \langle pf, g \rangle = -i\hbar[f^*g]_0^L.$$

Imposing Dirichlet conditions makes this boundary term zero but does not make this first-derivative operator self-adjoint: its adjoint has a larger domain. Moreover, differentiating a Dirichlet function generally gives a function that does not satisfy Dirichlet conditions. The self-adjoint Hamiltonian with domain $H^2 \cap H_0^1$ is well defined independently of trying to square a self-adjoint momentum on that same domain. This is why box energy eigenstates need not be momentum eigenstates.

54.5 Density of states in arbitrary dimension

In a large periodic box, each momentum state occupies a cell of volume $(2\pi\hbar)^d/\mathcal{V}$ in momentum space. The states with energy at most E occupy a d -ball of radius $p_E = \sqrt{2mE}$. Since a unit d -ball has volume $\pi^{d/2}/\Gamma(d/2 + 1)$,

$$N(E) \simeq \frac{\mathcal{V}}{(2\pi\hbar)^d} \frac{\pi^{d/2}}{\Gamma(d/2 + 1)} (2mE)^{d/2}. \quad (54.5)$$

Differentiation gives

$$g(E) = \frac{dN}{dE} \simeq \frac{\mathcal{V}}{(2\pi\hbar)^d} \frac{\pi^{d/2}(2m)^{d/2}}{\Gamma(d/2)} E^{d/2-1}.$$

In three dimensions,

$$\frac{g(E)}{\mathcal{V}} = \frac{m^{3/2}}{\sqrt{2}\pi^2\hbar^3} \sqrt{E}.$$

This approximation describes a smoothed high-energy spectrum; it does not replace the exact discrete levels near the ground state.

For a hard-walled box, momentum labels have spacing $\pi\hbar/L_i$ but only positive integers occur. The smaller cell volume contributes a factor 2^d , while restriction to the positive orthant contributes 2^{-d} . These factors cancel, giving exactly the same leading density (54.5). For unequal side lengths the allowed region in integer-label space is an ellipsoid. Rescaling its axes again gives the factor $\mathcal{V} = \prod_i L_i$. Boundary conditions affect subleading surface corrections, while the leading term depends only on volume.

55 Multidimensional harmonic oscillators

55.1 Normal coordinates and the spectrum

Consider

$$H = \frac{\mathbf{p}^2}{2m} + \frac{1}{2} K_{ij} x_i x_j,$$

where K is real, symmetric, and positive definite. An orthogonal change of coordinates diagonalizes K while preserving the kinetic energy and canonical commutators. In principal axes,

$$H = \sum_{i=1}^d \left(\frac{p_i^2}{2m} + \frac{1}{2} m \omega_i^2 x_i^2 \right), \quad \omega_i = \sqrt{K_i/m}.$$

The eigenstates are products of one-dimensional oscillator states:

$$|n_1, \dots, n_d\rangle = \bigotimes_i |n_i\rangle, \quad E_{\mathbf{n}} = \sum_i \hbar \omega_i (n_i + \tfrac{1}{2}), \quad n_i = 0, 1, 2, \dots$$

The one-dimensional Hermite wavefunctions and ladder construction were derived in the continuous-observables chapter.

For an isotropic oscillator $\omega_i = \omega$, let $N = \sum_i n_i$. Then

$$E_N = \hbar \omega (N + d/2).$$

The number of nonnegative integer solutions of $\sum_i n_i = N$ is

$$d_N = \binom{N + d - 1}{d - 1}. \quad (55.1)$$

To count them, arrange N identical excitation symbols and $d - 1$ separators; the numbers between successive separators are the n_i . For $d = 3$, $d_N = (N + 1)(N + 2)/2$, giving 1, 3, 6, 10, 15, 21, ...

55.2 Rotation multiplets and the hidden symmetry

A creation operator $a_i^\dagger$ transforms as a vector under spatial rotations. A level-one state has coefficients v_i in $v_i a_i^\dagger |0\rangle$. A level-two state is $S_{ij} a_i^\dagger a_j^\dagger |0\rangle$, and commuting creation operators mean that only the symmetric part $S_{ij} = S_{ji}$ contributes.

Under a real rotation O , $S \mapsto OSO^T$. Its trace is invariant, so the two projections

$$S_{\text{trace}} = \frac{\text{Tr } S}{d} \mathbf{1}, \quad S_{\text{traceless}} = S - \frac{\text{Tr } S}{d} \mathbf{1}$$

span invariant subspaces. In three dimensions the six-dimensional level splits into a scalar of dimension one and a traceless symmetric tensor of dimension five, corresponding to $\ell = 0$ and $\ell = 2$. The antisymmetric $\ell = 1$ part of $1 \otimes 1 = 2 \oplus 1 \oplus 0$ is absent because the creation operators commute.

More generally, a level- N state is a symmetric rank- N tensor. Removing successive traces in three dimensions gives

$$\ell = N, N - 2, N - 4, \dots,$$

ending at zero or one. Thus the first levels decompose as

N	Dimension	Orbital angular momenta
0	1	0
1	3	1
2	5 + 1	2, 0
3	7 + 3	3, 1
4	9 + 5 + 1	4, 2, 0
5	11 + 7 + 3	5, 3, 1

The complete oscillator is invariant under $O(d)$, including inversion. At level N , inversion contributes $(-1)^N$, consistent with the parity $(-1)^\ell$ of all the angular multiplets listed above.

Rotations alone do not require different ℓ multiplets to have the same energy. The additional degeneracy is explained by a larger symmetry. Define $Q_{ij} = a_i^\dagger a_j$. Since $H = \hbar\omega(\sum_i a_i^\dagger a_i + d/2)$,

$$[H, Q_{ij}] = 0, \quad [Q_{ij}, Q_{kl}] = \delta_{jk}Q_{il} - \delta_{il}Q_{kj}.$$

These operators generate $\mathfrak{u}(d)$; the traceless combinations generate $\mathfrak{su}(d)$. They transfer excitations between coordinates without changing N . The level- N states form the symmetric N th-power representation of this larger group. In three dimensions this is the hidden $SU(3)$ symmetry that relates the different rotation multiplets.

55.3 High-energy state counting and accessible volume

For $d = 3$,

$$\mathcal{N}(N) = \sum_{q=0}^N \frac{(q+1)(q+2)}{2} = \frac{(N+1)(N+2)(N+3)}{6}.$$

At high energy $E \simeq N\hbar\omega$, hence

$$\mathcal{N}(E) \simeq \frac{E^3}{6(\hbar\omega)^3}, \quad g(E) \simeq \frac{E^2}{2(\hbar\omega)^3}. \quad (55.2)$$

In d dimensions, the corresponding expressions are $\mathcal{N}(E) \simeq E^d/[d!(\hbar\omega)^d]$ and $g(E) \simeq E^{d-1}/[(d-1)!(\hbar\omega)^d]$.

The density grows faster than in a fixed box because the classically accessible radius grows as $r_E = \sqrt{2E/(m\omega^2)}$.

A rough comparison multiplies the free-particle density by the volume $4\pi r_E^3/3$, giving

$$g_{\text{rough}}(E) = \frac{8}{3\pi} \frac{E^2}{(\hbar\omega)^3}.$$

It captures the E^2 scaling but overcounts, because it assigns kinetic energy E everywhere. A precise semiclassical calculation integrates the local kinetic energy:

$$g(E) = \int_{V(\mathbf{x}) < E} \frac{m^{3/2}}{\sqrt{2\pi^2\hbar^3}} \sqrt{E - V(\mathbf{x})} d^3x.$$

Substitute $r = r_E s$. Using $\int_0^1 s^2 \sqrt{1-s^2} ds = \pi/16$ recovers the coefficient $1/2$ in (55.2).

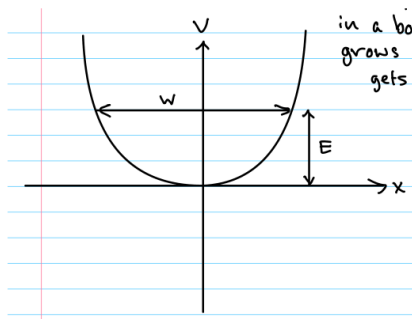


Figure 49: The spatial region accessible at energy E grows with oscillator energy.

56 One-dimensional bound states

56.1 Discrete levels, continua, and thresholds

For a potential tending to zero at infinity, a generic negative-energy state has exponentially decaying exterior tails and can be normalizable. Positive-energy exterior solutions oscillate and form a continuum of generalized eigenstates. A finite attractive well therefore has both a discrete bound spectrum and a scattering continuum. A confining potential that grows without bound has a discrete spectrum, while the free particle has only the continuum.

The continuum threshold is set by the asymptotic potential, not by the height of a barrier surrounding a local minimum. If the energy lies above an open asymptotic channel, a classically trapped state can tunnel out. It then represents an unstable resonance rather than a normalizable stationary bound state. Specially tuned potentials can support threshold states or bound states embedded in a continuum, so this description is a generic one rather than an unrestricted theorem.



Figure 50: A state trapped classically can escape quantum mechanically through a barrier into an open channel.

56.2 An attractive delta-function well

Let

$$V(x) = -g\delta(x), \quad g > 0.$$

The coupling g has units energy times length. For $E < 0$, define $\kappa = \sqrt{-2mE}/\hbar$. Away from the origin, $\psi'' = \kappa^2\psi$. Normalizability removes the growing exponential on

each side:

$$\psi(x) = \begin{cases} Ae^{-\kappa x}, & x > 0, \\ Be^{\kappa x}, & x < 0. \end{cases}$$

The wavefunction must be continuous. A jump would create a δ' term in ψ'' that the potential cannot cancel. Thus $A = B$.

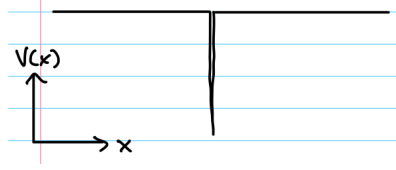


Figure 51: An attractive delta-function potential.

Integrate the Schrödinger equation across $[-\epsilon, \epsilon]$:

$$-\frac{\hbar^2}{2m}[\psi'(\epsilon) - \psi'(-\epsilon)] - g\psi(0) = E \int_{-\epsilon}^{\epsilon} \psi(x) dx.$$

As $\epsilon \rightarrow 0$, the integral vanishes and

$$\psi'(0^+) - \psi'(0^-) = -\frac{2mg}{\hbar^2}\psi(0). \quad (56.1)$$

Substitution gives $-2\kappa A = -2mgA/\hbar^2$, hence

$$\kappa = \frac{mg}{\hbar^2}, \quad E = -\frac{mg^2}{2\hbar^2}.$$

Normalization gives $1 = 2|A|^2 \int_0^\infty e^{-2\kappa x} dx = |A|^2/\kappa$, so

$$\psi(x) = \sqrt{\kappa} e^{-\kappa|x|}. \quad (56.2)$$

There is exactly one bound state. An odd state would have $\psi(0) = 0$ and would reduce to a globally smooth, decaying free solution, which must vanish identically. Positive-energy scattering states remain present.

56.3 Parity reduction for the finite square well

Consider

$$V(x) = \begin{cases} -V_0, & |x| < a, \\ 0, & |x| > a, \end{cases} \quad V_0 > 0.$$

Since $V(-x) = V(x)$, energy eigenfunctions can be chosen even or odd. At $x = 0$, an even solution satisfies $\psi'(0) = 0$; an odd solution satisfies $\psi(0) = 0$. Thus it is sufficient to solve on $x \geq 0$. For $-V_0 < E < 0$, define

$$k = \frac{\sqrt{2m(V_0 + E)}}{\hbar}, \quad \kappa = \frac{\sqrt{-2mE}}{\hbar}.$$

The normalizable solutions are

$$\psi_+(x) = \begin{cases} A \cos kx, & 0 < x < a, \\ Be^{-\kappa x}, & x > a, \end{cases}$$

$$\psi_-(x) = \begin{cases} A \sin kx, & 0 < x < a, \\ Be^{-\kappa x}, & x > a. \end{cases}$$

At a finite step, both ψ and ψ' are continuous. Even matching gives

$$A \cos ka = Be^{-\kappa a}, \quad -kA \sin ka = -\kappa Be^{-\kappa a},$$

and division yields $k \tan ka = \kappa$. Odd matching similarly yields $-k \cot ka = \kappa$. These are the quantization conditions.

56.4 Graphical quantization, level count, and energy bounds

Introduce

$$z = ka, \quad q = \kappa a, \quad z_0 = \frac{a\sqrt{2mV_0}}{\hbar}, \quad z^2 + q^2 = z_0^2.$$

The two equations become

$$z \tan z = \sqrt{z_0^2 - z^2} \quad (\text{even}), \quad -z \cot z = \sqrt{z_0^2 - z^2} \quad (\text{odd}). \quad (56.3)$$

The positive branches alternate. On each admissible half-period the left side increases from zero to infinity, while the semicircle on the right decreases. Each intersection gives one bound state.

There is always an even ground state for every $z_0 > 0$. Numbering states $n = 1, 2, \dots$ from the ground state,

$$\frac{(n-1)\pi}{2} < z_n < \frac{n\pi}{2}, \quad z_n < z_0.$$

A new bound state appears when z_0 exceeds $(n-1)\pi/2$. Thus

$$N_b = \left\lceil \frac{2z_0}{\pi} \right\rceil, \quad z_0 > 0. \quad (56.4)$$

Away from integer values of $2z_0/\pi$, this equals $1 + \lfloor 2z_0/\pi \rfloor$. Exactly at a threshold the zero-energy solution has a nondecaying exterior tail and is not an additional square-integrable bound state.

Since $E_n = \hbar^2 z_n^2 / (2ma^2) - V_0$,

$$\frac{\hbar^2 (n-1)^2 \pi^2}{8ma^2} - V_0 < E_n < \frac{\hbar^2 n^2 \pi^2}{8ma^2} - V_0, \quad E_n < 0.$$

For a deep well and fixed low n , $z_n \rightarrow n\pi/2$. The upper bound becomes the hard-wall spectrum of a box of length $2a$, shifted by $-V_0$.

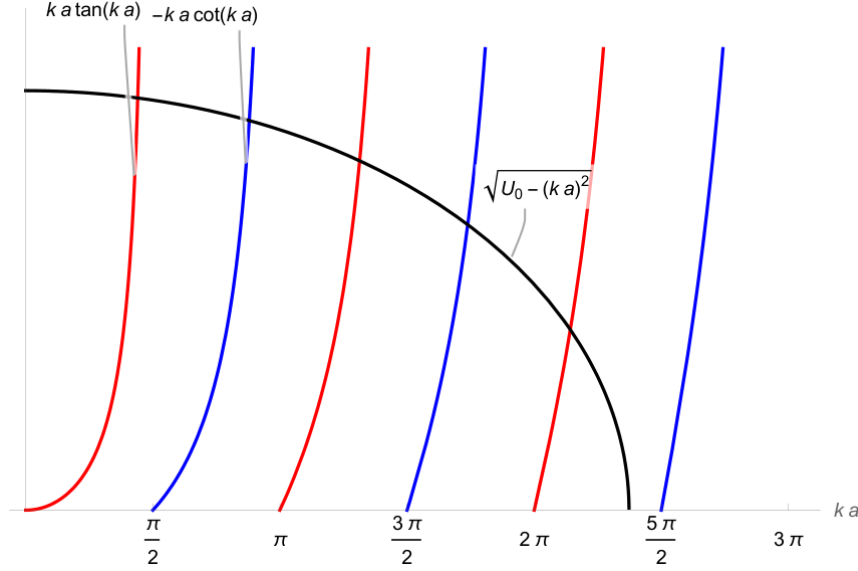


Figure 52: Even and odd square-well levels are the intersections of the transcendental matching curves.

56.5 Normalization, tail probability, and signs

Choose the interior coefficient A real and positive. For an even state,

$$\begin{aligned} 1 &= 2A^2 \int_0^a \cos^2 kx \, dx + 2B^2 \int_a^\infty e^{-2\kappa x} \, dx \\ &= A^2 \left[a + \frac{\sin 2ka}{2k} + \frac{\cos^2 ka}{\kappa} \right]. \end{aligned}$$

Using $k \tan ka = \kappa$ simplifies the bracket to $a + 1/\kappa$. For odd states the analogous expression is $a - \sin(2ka)/(2k) + \sin^2(ka)/\kappa$, which simplifies to the same result. Therefore

$$A = \sqrt{\frac{\kappa}{1 + \kappa a}}, \quad B = \begin{cases} Ae^{\kappa a} \cos ka, & \text{even,} \\ Ae^{\kappa a} \sin ka, & \text{odd.} \end{cases} \quad (56.5)$$

For the n th state, the sign of B is $(-1)^{\lfloor (n-1)/2 \rfloor}$ under this phase convention. Its magnitude is

$$|B| = e^{\kappa a} \sqrt{\frac{\kappa k^2}{(k^2 + \kappa^2)(1 + \kappa a)}}.$$

The exterior probability is

$$P_{\text{out}} = \frac{B^2 e^{-2\kappa a}}{\kappa} = \frac{k^2}{(k^2 + \kappa^2)(1 + \kappa a)}, \quad P_{\text{in}} = 1 - P_{\text{out}}. \quad (56.6)$$

These probabilities lie between zero and one. Weak binding increases the fraction outside; it does not make a probability diverge.

56.6 A shallow well and the delta-function limit

If $z_0 \ll 1$, only the even ground state exists. Since $z \leq z_0$, $z \tan z = z^2 + O(z^4)$, so

$$q \simeq z^2, \quad z^4 + z^2 - z_0^2 \simeq 0.$$

The positive root gives

$$z^2 \simeq \frac{-1 + \sqrt{1 + 4z_0^2}}{2} = z_0^2 - z_0^4 + O(z_0^6).$$

The leading z_0^2 term cancels the well depth in the energy. The next term is essential:

$$E \simeq -\frac{\hbar^2 z_0^4}{2ma^2} = -\frac{2ma^2 V_0^2}{\hbar^2} = -\frac{mg^2}{2\hbar^2}, \quad g = 2aV_0.$$

This is exactly the attractive delta-well energy with the same integrated strength. A narrow well with fixed area approaches that point interaction.

56.7 A state approaching the continuum

Let $z_0 = z_c + \epsilon$, where $z_c = N\pi/2$ with $N \geq 1$ and $0 < \epsilon \ll 1$. Write the newest root as $z = z_c + \delta$. On the appropriate even or odd branch, the matching equation becomes

$$q \simeq z\delta, \quad q^2 = (z_0 - z)(z_0 + z).$$

To leading nontrivial order,

$$2z_c(\epsilon - \delta) \simeq z_c^2 \delta^2, \quad \delta = \epsilon - \frac{z_c}{2} \epsilon^2 + O(\epsilon^3).$$

Thus $q \simeq z_c \epsilon$ and

$$E = -\frac{\hbar^2 q^2}{2ma^2} \simeq -\epsilon^2 V_0.$$

The decay length $\kappa^{-1} = a/q$ diverges as the threshold is approached. Using (56.6), $P_{\text{in}} \sim \kappa a \rightarrow 0$.

Parity gives $\langle x \rangle = 0$. The exterior contribution to the second moment is evaluated exactly by putting $x = a + y$:

$$\begin{aligned} \langle x^2 \rangle_{\text{out}} &= 2B^2 e^{-2\kappa a} \int_0^\infty (a + y)^2 e^{-2\kappa y} dy \\ &= P_{\text{out}} \left(a^2 + \frac{a}{\kappa} + \frac{1}{2\kappa^2} \right). \end{aligned}$$

For completeness, its interior counterpart can also be evaluated without an approximation. Write $s = +1$ for even parity and $s = -1$ for odd parity. Since $2 \cos^2 kx = 1 + \cos 2kx$ and $2 \sin^2 kx = 1 - \cos 2kx$,

$$\begin{aligned} \langle x^2 \rangle_{\text{in}} &= A^2 \int_0^a x^2 [1 + s \cos(2kx)] dx \\ &= A^2 \left\{ \frac{a^3}{3} + s \left[\frac{a^2 \sin(2ka)}{2k} + \frac{a \cos(2ka)}{2k^2} - \frac{\sin(2ka)}{4k^3} \right] \right\}. \end{aligned}$$

The antiderivative follows by integrating $x^2 \cos(2kx)$ twice by parts. Together with the exact exterior expression this gives the variance of any bound level:

$$(\Delta x)^2 = \langle x^2 \rangle_{\text{in}} + P_{\text{out}} \left(a^2 + \frac{a}{\kappa} + \frac{1}{2\kappa^2} \right).$$

The interior contribution is bounded by $a^2 P_{\text{in}}$. Therefore

$$(\Delta x)^2 \sim \frac{1}{2\kappa^2} \longrightarrow \infty.$$

Conditioned on finding the particle outside the well, the distance $y = |x| - a$ has density $2\kappa e^{-2\kappa y}$ and mean $1/(2\kappa)$. This makes the growing tail directly measurable.

56.8 Nondegeneracy, nodes, and a threshold example

For two square-integrable solutions of a regular one-dimensional real potential with the same energy, the Wronskian $W = \psi_1 \psi_2' - \psi_1' \psi_2$ has derivative zero by the Schrödinger equation. The decaying boundary conditions imply $W = 0$, so the solutions are proportional. Bound levels are therefore nondegenerate under the usual separated boundary conditions. The Sturm oscillation theorem orders them by nodes: the ground state has none, and the n th state has $n - 1$ interior nodes. In a symmetric well this makes parity alternate.

The finite-range square well has no normalizable threshold state, but slower-decaying potentials can. To construct one, choose a nodeless normalizable function and solve the zero-energy equation for V :

$$V(x) = \frac{\hbar^2}{2m} \frac{\psi''(x)}{\psi(x)}.$$

For

$$\psi(x) = \frac{1}{\sqrt{\pi a}} \frac{1}{\sqrt{1 + (x/a)^2}},$$

normalization follows from $\int_{-\infty}^{\infty} dx/(1 + x^2/a^2) = \pi a$. Direct differentiation gives

$$V(x) = \frac{\hbar^2}{2m} \frac{2x^2 - a^2}{(a^2 + x^2)^2}, \quad E = 0.$$

This potential tends to zero but supports a square-integrable state right at the threshold. More generally, writing $\psi = e^{u(x)}$ gives $\psi''/\psi = u'' + (u')^2$, a useful inverse construction. A Gaussian choice instead produces a confining quadratic potential.

57 Central potentials and the radial Schrödinger equation

57.1 Rotational symmetry and orbital angular momentum

A central potential depends only on $r = |\mathbf{x}|$: $H = \mathbf{p}^2/(2m) + V(r)$. Define $L_i = \epsilon_{ijk} x_j p_k$. There is no ordering ambiguity because the antisymmetric tensor selects different

Cartesian indices, for which x_j and p_k commute. The canonical commutators give

$$[L_i, x_j] = i\hbar\epsilon_{ijk}x_k, \quad [L_i, p_j] = i\hbar\epsilon_{ijk}p_k, \quad [L_i, L_j] = i\hbar\epsilon_{ijk}L_k.$$

The first two identities imply $[L_i, \mathbf{x}^2] = [L_i, \mathbf{p}^2] = 0$; hence $[H, L_i] = 0$ for a central potential. We can diagonalize H, L^2, L_z simultaneously.

In spherical coordinates,

$$\nabla^2 = \frac{1}{r^2}\partial_r(r^2\partial_r) + \frac{1}{r^2\sin\theta}\partial_\theta(\sin\theta\partial_\theta) + \frac{1}{r^2\sin^2\theta}\partial_\phi^2,$$

where

$$L^2 = -\hbar^2 \left[\frac{1}{\sin\theta}\partial_\theta(\sin\theta\partial_\theta) + \frac{1}{\sin^2\theta}\partial_\phi^2 \right], \quad L_z = -i\hbar\partial_\phi.$$

The angular eigenfunctions are spherical harmonics:

$$L^2 Y_{\ell m} = \hbar^2 \ell(\ell+1) Y_{\ell m}, \quad L_z Y_{\ell m} = \hbar m Y_{\ell m}.$$

A scalar orbital wavefunction carries integer $\ell = 0, 1, \dots$, with $m = -\ell, \dots, \ell$; half-integer spin requires additional internal degrees of freedom.

With the Condon–Shortley phase included in P_ℓ^m ,

$$Y_{\ell m}(\theta, \phi) = \sqrt{\frac{2\ell+1}{4\pi} \frac{(\ell-m)!}{(\ell+m)!}} P_\ell^m(\cos\theta) e^{im\phi}.$$

They obey

$$\int_{S^2} Y_{\ell m}^* Y_{\ell' m'} d\Omega = \delta_{\ell\ell'} \delta_{mm'}, \quad \sum_{\ell, m} Y_{\ell m}(\Omega) Y_{\ell m}^*(\Omega') = \delta_{S^2}(\Omega, \Omega').$$

Here $d\Omega = \sin\theta d\theta d\phi$, and the angular delta function is defined with this measure.

57.2 Removing the radial measure

Set $\psi(r, \Omega) = R(r) Y_{\ell m}(\Omega)$. The radial equation is

$$-\frac{\hbar^2}{2mr^2} \frac{d}{dr}(r^2 R') + \left[V(r) + \frac{\hbar^2 \ell(\ell+1)}{2mr^2} \right] R = ER.$$

Define $\chi(r) = rR(r)$. Since $(r^2 R')' = r\chi''$, the equation simplifies to

$$-\frac{\hbar^2}{2m} \chi''(r) + U_\ell(r) \chi(r) = E \chi(r), \quad U_\ell(r) = V(r) + \frac{\hbar^2 \ell(\ell+1)}{2mr^2}. \quad (57.1)$$

The extra positive term is the centrifugal barrier. The problem is one-dimensional on the half-line, but with an angular-momentum-dependent potential.

Normalization becomes

$$\int |\psi|^2 d^3x = \int_0^\infty r^2 |R|^2 dr \int |Y_{\ell m}|^2 d\Omega = \int_0^\infty |\chi|^2 dr = 1.$$

Thus $|\chi(r)|^2 dr$ is the probability in a spherical shell of thickness dr . For the standard regular central-potential domain, $\chi(0) = 0$. A nonzero constant $\chi(0)$ would make $R \sim 1/r$ and introduce an inadmissible singularity without a suitable point interaction. Singular potentials and contact interactions require their own domain analysis rather than automatic use of this condition.

A general wavefunction has the decomposition

$$\psi(\mathbf{x}) = \sum_{\ell, m} \frac{\chi_{\ell m}(r)}{r} Y_{\ell m}(\Omega), \quad \sum_{\ell, m} \int_0^\infty |\chi_{\ell m}|^2 dr = 1.$$

Central symmetry decouples these angular sectors. This differs from Cartesian separability: a sum of one-dimensional finite wells creates semi-infinite troughs along coordinate directions and is not the same potential as a bounded spherical or square-shaped well.

57.3 The infinite spherical well

Let $V = 0$ for $r < a$ and impose a hard wall at $r = a$. Inside,

$$-\chi'' + \frac{\ell(\ell+1)}{r^2} \chi = k^2 \chi, \quad E = \frac{\hbar^2 k^2}{2m}.$$

Its solutions are $\chi = A r j_\ell(kr) + B r y_\ell(kr)$. Rayleigh's formulas are

$$j_\ell(z) = (-z)^\ell \left(\frac{1}{z} \frac{d}{dz} \right)^\ell \frac{\sin z}{z}, \quad y_\ell(z) = -(-z)^\ell \left(\frac{1}{z} \frac{d}{dz} \right)^\ell \frac{\cos z}{z}.$$

For example,

$$j_0 = \frac{\sin z}{z}, \quad j_1 = \frac{\sin z}{z^2} - \frac{\cos z}{z}, \quad y_0 = -\frac{\cos z}{z}, \quad y_1 = -\frac{\cos z}{z^2} - \frac{\sin z}{z}.$$

Near zero, $j_\ell(z) \sim z^\ell / (2\ell+1)!!$ and $y_\ell(z) \sim -(2\ell-1)!! / z^{\ell+1}$. Regularity eliminates B . The outer wall requires

$$j_\ell(ka) = 0.$$

If $z_{\ell n}$ is its n th positive zero, then

$$E_{n\ell} = \frac{\hbar^2 z_{\ell n}^2}{2ma^2}, \quad R_{n\ell}(r) = \frac{\sqrt{2}}{a^{3/2} |j_{\ell+1}(z_{\ell n})|} j_\ell(z_{\ell n} r/a). \quad (57.2)$$

The normalization follows from $\int_0^a r^2 j_\ell(z_{\ell n} r/a)^2 dr = a^3 j_{\ell+1}(z_{\ell n})^2 / 2$.

For $\ell = 0$, $z_{0n} = n\pi$ and $\chi_{n0} = \sqrt{2/a} \sin(n\pi r/a)$. The ground energy is $E_0 = \hbar^2 \pi^2 / (2ma^2)$. The first angular sectors in increasing energy are 0, 1, 2, 0, 3, 1, 4, 2, 5, 0, ... Each radial level retains its $2\ell + 1$ magnetic degeneracy, even though different radial-angular pairs generally have different energies.

Example 57.1 (Sphere versus cube). For a cube of side L , $E_{\text{cube}} = 3\pi^2 \hbar^2 / (2mL^2)$. If $L = 2a$, the cube has 3/4 of the sphere's ground energy; the cube also has the larger volume. If the volumes are equal, $L = (4\pi/3)^{1/3} a$, and

$$\frac{E_{\text{cube}}}{E_{\text{sphere}}} = \frac{3}{(4\pi/3)^{2/3}} > 1.$$

At equal volume the sphere has the lower ground energy.

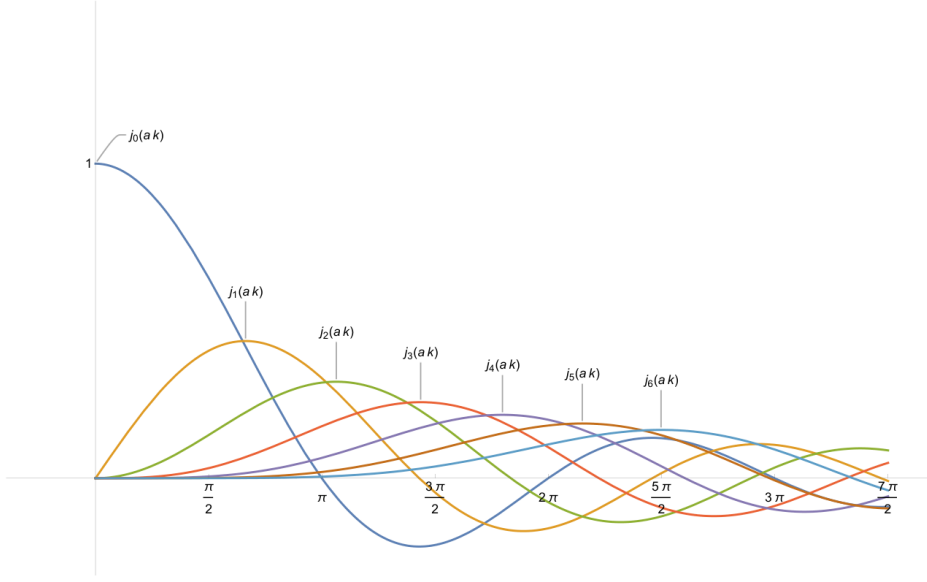


Figure 53: The zeros of the spherical Bessel functions determine the radial levels.

58 The Coulomb spectrum by factorization

58.1 Two-body reduction and natural scales

For $V(r) = -K/r$, $K > 0$, the hydrogenic value is $K = Ze^2/(4\pi\epsilon_0)$ in SI units. If the nucleus is not treated as infinitely heavy, separate the center of mass from the relative coordinate:

$$\mathbf{R} = \frac{m_e \mathbf{r}_e + m_N \mathbf{r}_N}{m_e + m_N}, \quad \mathbf{r} = \mathbf{r}_e - \mathbf{r}_N.$$

The kinetic energy becomes $\mathbf{P}^2/[2(m_e + m_N)] + \mathbf{p}^2/(2\mu)$, where $\mu = m_e m_N/(m_e + m_N)$. The internal problem therefore uses the reduced mass. In the following, m denotes this mass.

Define

$$a_0 = \frac{\hbar^2}{mK}, \quad E_R = \frac{mK^2}{2\hbar^2} = \frac{\hbar^2}{2ma_0^2}, \quad \varepsilon = \frac{2mE}{\hbar^2}.$$

For ordinary hydrogen, a_0 is approximately 5.29×10^{-11} m and E_R approximately 13.6 eV, with small reduced-mass corrections. The radial equation is

$$h_\ell \chi = \varepsilon \chi, \quad h_\ell = -\frac{d^2}{dr^2} + \frac{\ell(\ell+1)}{r^2} - \frac{2}{a_0 r}. \quad (58.1)$$

We seek square-integrable solutions with $\varepsilon < 0$.

58.2 Factorization and the lowest state at fixed angular momentum

Let $A = d/dr + W(r)$ with real W . Integration by parts on radial functions whose boundary term vanishes gives $A^\dagger = -d/dr + W$. Acting on a test function,

$$\begin{aligned} A^\dagger A f &= (-\partial_r + W)(f' + Wf) \\ &= -f'' - W'f - Wf' + Wf' + W^2f \\ &= [-\partial_r^2 + W^2 - W'] f. \end{aligned}$$

Choose

$$W_\ell(r) = -\frac{\ell+1}{r} + \frac{1}{(\ell+1)a_0}.$$

Then

$$A_\ell^\dagger A_\ell = h_\ell + c_\ell, \quad c_\ell = \frac{1}{(\ell+1)^2 a_0^2}. \quad (58.2)$$

Since $\langle \chi, A_\ell^\dagger A_\ell \chi \rangle = \|A_\ell \chi\|^2 \geq 0$, $\varepsilon \geq -c_\ell$. Equality is possible exactly when $A_\ell \chi = 0$:

$$\frac{\chi'}{\chi} = \frac{\ell+1}{r} - \frac{1}{(\ell+1)a_0}.$$

Integration gives

$$\chi_{\ell+1,\ell}(r) = N r^{\ell+1} e^{-r/[(\ell+1)a_0]}.$$

Using $\int_0^\infty r^q e^{-br} dr = q!/b^{q+1}$,

$$\chi_{\ell+1,\ell}(r) = \sqrt{\frac{2}{(\ell+1)a_0(2\ell+2)!}} \left[\frac{2r}{(\ell+1)a_0} \right]^{\ell+1} e^{-r/[(\ell+1)a_0]}, \quad (58.3)$$

with energy $E = -E_R/(\ell+1)^2$. The overall ground state has

$$\chi_{10} = \frac{2r}{a_0^{3/2}} e^{-r/a_0}, \quad \psi_{100}(\mathbf{x}) = \frac{1}{\sqrt{\pi a_0^3}} e^{-r/a_0}, \quad E_1 = -E_R.$$

The first-order equation has only one solution up to scale, establishing uniqueness of the fixed- ℓ lowest radial state.

58.3 Shape invariance and quantization of every bound energy

Reversing the factors changes the sign of W' :

$$A_\ell A_\ell^\dagger = -\partial_r^2 + \frac{(\ell+1)(\ell+2)}{r^2} - \frac{2}{a_0 r} + c_\ell = h_{\ell+1} + c_\ell.$$

Therefore

$$A_\ell A_\ell^\dagger - c_\ell = A_{\ell+1}^\dagger A_{\ell+1} - c_{\ell+1}. \quad (58.4)$$

This is called shape invariance: the partner operator has the same functional form with a shifted angular-momentum parameter.

If $h_\ell \chi = \varepsilon \chi$, then

$$h_{\ell+1}(A_\ell \chi) = \varepsilon A_\ell \chi.$$

The norm is

$$\|A_\ell \chi\|^2 = (\varepsilon + c_\ell) \|\chi\|^2.$$

An excited state in the ℓ sector has $\varepsilon > -c_\ell$, so the new state is nonzero, normalizable, and has angular momentum $\ell + 1$ at the same energy.

Repeat this operation. A fixed negative ε cannot satisfy $\varepsilon \geq -c_j$ for arbitrarily large j , because $c_j \rightarrow 0$. The chain must therefore terminate at a state annihilated by one of the A_j . If its final angular momentum is $j = n - 1$, its energy is

$$\varepsilon_n = -\frac{1}{n^2 a_0^2}, \quad E_n = -\frac{E_R}{n^2}, \quad n = 1, 2, 3, \dots \quad (58.5)$$

The starting angular momentum satisfies $0 \leq \ell \leq n - 1$. This argument obtains the entire allowed set of bound energies from positivity and the terminating ladder.

58.4 Descending in angular momentum and normalization factors

For $\ell > 0$, the adjoint operator $A_{\ell-1}^\dagger$ maps an ℓ -sector state to one with $\ell - 1$ and the same energy. Its squared norm factor is

$$\|A_{\ell-1}^\dagger \chi\|^2 = (\varepsilon + c_{\ell-1}) \|\chi\|^2 = \frac{n^2 - \ell^2}{n^2 \ell^2 a_0^2} \|\chi\|^2 > 0.$$

Thus every $\ell = 0, \dots, n - 1$ occurs. With a compatible phase convention,

$$A_{\ell-1}^\dagger |n, \ell\rangle = \frac{\sqrt{n^2 - \ell^2}}{n \ell a_0} |n, \ell - 1\rangle, \quad (58.6)$$

$$A_\ell |n, \ell\rangle = \frac{\sqrt{n^2 - (\ell + 1)^2}}{n(\ell + 1) a_0} |n, \ell + 1\rangle. \quad (58.7)$$

The second formula follows either from its norm or by taking the adjoint matrix element of the first with ℓ replaced by $\ell + 1$.

The procedure stops at $\ell = 0$. The formal expression A_{-1} contains $1/(\ell + 1)$ and is undefined, and a scalar orbital wavefunction does not have an $\ell = -1$ angular sector.

There is one radial state for each pair (n, ℓ) . To see this directly, map any such state repeatedly to the unique highest- ℓ state $\ell = n - 1$. The maps have nonzero norm factors on all intermediate sectors, so they are injective there. Two independent states could not both map into a one-dimensional final space. Including m , the orbital degeneracy at principal number n is

$$\sum_{\ell=0}^{n-1} (2\ell + 1) = n^2.$$

This count excludes electron spin and any relativistic corrections.

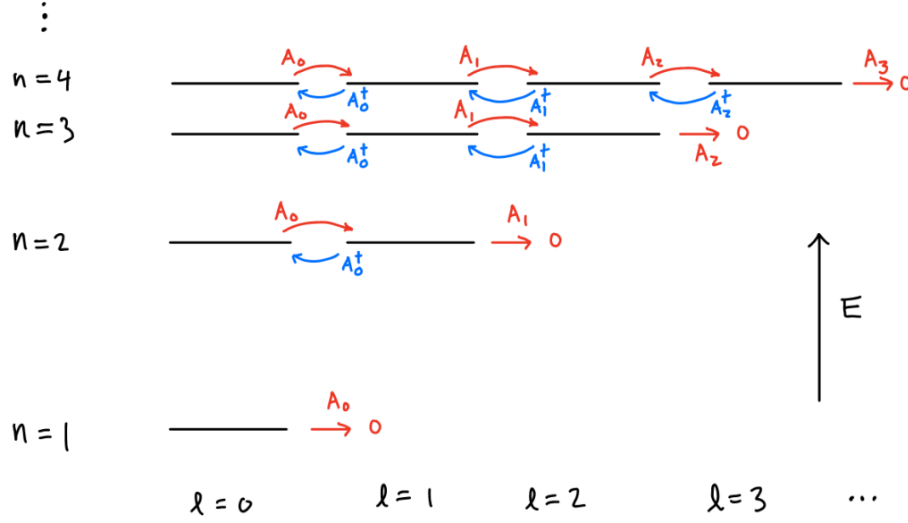


Figure 54: Coulomb radial states of equal energy are connected by the factorization operators.

58.5 Explicit radial polynomials and their recurrence

Set $u = 2r/(na_0)$ and write

$$\chi_{n\ell}(r) = \sqrt{\frac{2}{na_0}} f_{n\ell}(u), \quad f_{n,n-1}(u) = \frac{u^n e^{-u/2}}{\sqrt{(2n)!}}.$$

Substitution into (58.6) gives

$$f_{n,\ell-1}(u) = \frac{1}{\sqrt{n^2 - \ell^2}} \left[n - 2\ell \left(\frac{d}{du} + \frac{\ell}{u} \right) \right] f_{n\ell}(u). \quad (58.8)$$

Every application reduces the lowest power of u by one but leaves the highest power u^n unchanged. Hence

$$f_{n\ell}(u) = u^{\ell+1} P_{n-\ell-1}^{2\ell+1}(u) e^{-u/2}.$$

Use $k = n - \ell - 1$ and $\alpha = 2\ell + 1$. Then

$$P_{k+1}^{\alpha-2}(u) = \frac{(k + \alpha)uP_k^\alpha(u) - (\alpha - 1)[u(P_k^\alpha)'(u) + \alpha P_k^\alpha(u)]}{\sqrt{(k+1)(k+\alpha)}}, \quad P_0^\alpha = \frac{1}{\sqrt{(\alpha+1)!}}. \quad (58.9)$$

The first three polynomials are

$$\begin{aligned} P_0^\alpha(u) &= \frac{1}{\sqrt{(\alpha+1)!}}, \\ P_1^\alpha(u) &= \frac{u - (\alpha+1)}{\sqrt{(\alpha+3)(\alpha+1)!}}, \\ P_2^\alpha(u) &= \frac{u^2 - 2(\alpha+2)u + (\alpha+1)(\alpha+2)}{\sqrt{2(\alpha+5)(\alpha+2)!}}. \end{aligned}$$

A closed finite-sum solution is

$$P_k^\alpha(u) = (-1)^k \sqrt{\frac{k!(\alpha+k)!}{\alpha+2k+1}} \sum_{p=0}^k \frac{(-u)^p}{(k-p)!p!(p+\alpha)!}. \quad (58.10)$$

To verify the recurrence, insert this sum into its right-hand side and collect the coefficient of u^p . The multiplication-by- u term uses the coefficient at $p-1$, while $u\partial_u + \alpha$ multiplies the coefficient at p by $p+\alpha$. Putting these over the common denominator $(k+1-p)!p!(p+\alpha-2)!$ yields exactly the coefficient in $P_{k+1}^{\alpha-2}$; the square-root prefactor changes by $\sqrt{(k+1)(k+\alpha)}$. At $k=0$ the formula reduces to the prescribed constant, completing the induction.

58.6 Laguerre form, normalization, and low states

Define generalized Laguerre polynomials by Rodrigues' formula

$$L_k^{(\alpha)}(u) = \frac{u^{-\alpha} e^u}{k!} \frac{d^k}{du^k} (u^{k+\alpha} e^{-u}).$$

Leibniz differentiation gives

$$L_k^{(\alpha)}(u) = \sum_{p=0}^k \frac{(-1)^p (k+\alpha)!}{(k-p)!p!(p+\alpha)!} u^p,$$

so

$$P_k^\alpha = (-1)^k \sqrt{\frac{k!}{(\alpha+k)!(\alpha+2k+1)}} L_k^{(\alpha)}.$$

Dropping an overall state-dependent sign, the conventional normalized radial function is

$$R_{n\ell}(r) = \left(\frac{2}{na_0}\right)^{3/2} \sqrt{\frac{(n-\ell-1)!}{2n(n+\ell)!}} e^{-u/2} u^\ell L_{n-\ell-1}^{(2\ell+1)}(u), \quad u = \frac{2r}{na_0}. \quad (58.11)$$

The full states are $\psi_{n\ell m} = R_{n\ell} Y_{\ell m}$. The polynomial equation follows by substitution into (58.1):

$$uL'' + (2\ell+2-u)L' + (n-\ell-1)L = 0.$$

For example,

$$\begin{aligned} R_{10} &= 2a_0^{-3/2} e^{-r/a_0}, \\ R_{20} &= \frac{1}{2\sqrt{2}a_0^{3/2}} (2-r/a_0) e^{-r/(2a_0)}, \\ R_{21} &= \frac{1}{2\sqrt{6}a_0^{3/2}} (r/a_0) e^{-r/(2a_0)}. \end{aligned}$$

The $2s$ state has one radial node, whereas the $2p$ state has none.

Orthonormality is inherited from the self-adjoint radial problem and the normalized ladders. It also gives a useful identity for the P polynomials. With $\rho = 2r/a_0$ and fixed ℓ ,

$$\int_0^\infty \rho^{2\ell+2} P_{n-\ell-1}^{2\ell+1}(\rho/n) P_{n'-\ell-1}^{2\ell+1}(\rho/n') e^{-\rho/(2n)-\rho/(2n')} d\rho = n^{2\ell+3} \delta_{nn'}.$$

Here $n, n' \geq \ell + 1$. Although the argument and exponential scale depend on n , the identity follows immediately by inserting the normalized radial functions into $\int_0^\infty \chi_{n\ell} \chi_{n'\ell} dr = \delta_{nn'}$.

59 Propagators and free-particle wavepackets

59.1 The kernel of time evolution

Insert position completeness into an initial state:

$$|\Psi(t_0)\rangle = \int d^d x' \Psi(\mathbf{x}', t_0) |\mathbf{x}'\rangle.$$

Apply $U(t, t_0)$ and project onto $\langle \mathbf{x} |$:

$$\Psi(\mathbf{x}, t) = \int d^d x' K(\mathbf{x}, t; \mathbf{x}', t_0) \Psi(\mathbf{x}', t_0), \quad K = \langle \mathbf{x} | U(t, t_0) | \mathbf{x}' \rangle. \quad (59.1)$$

The propagator is a position-space matrix element of the evolution operator. It satisfies the Schrödinger equation in its first arguments and the initial condition $K(\mathbf{x}, t_0; \mathbf{x}', t_0) = \delta^{(d)}(\mathbf{x} - \mathbf{x}')$. Composition of time evolution gives

$$K(\mathbf{x}, t; \mathbf{y}, t_0) = \int d^d z K(\mathbf{x}, t; \mathbf{z}, t_1) K(\mathbf{z}, t_1; \mathbf{y}, t_0).$$

For a time-independent Hamiltonian, write $\tau = t - t_0$ and insert a complete energy basis:

$$K(\mathbf{x}, \mathbf{y}; \tau) = \sum_n e^{-iE_n \tau / \hbar} \psi_n(\mathbf{x}) \psi_n^*(\mathbf{y}), \quad (59.2)$$

with integrals over continuum labels when present. A retarded Green function for $i\hbar\partial_t - H$ is $G_R = (i\hbar)^{-1} \Theta(t - t_0) K$.

59.2 The free kernel by a Gaussian integral

For the free particle, momentum completeness gives

$$K_0(\mathbf{x}, \mathbf{y}; \tau) = \int \frac{d^d p}{(2\pi\hbar)^d} \exp \left[-\frac{ip^2 \tau}{2m\hbar} + \frac{i\mathbf{p} \cdot (\mathbf{x} - \mathbf{y})}{\hbar} \right].$$

Complete the square:

$$-\frac{i\tau}{2m\hbar} \left[\mathbf{p} - \frac{m(\mathbf{x} - \mathbf{y})}{\tau} \right]^2 + \frac{im|\mathbf{x} - \mathbf{y}|^2}{2\hbar\tau}.$$

Using the Gaussian integral with a small damping factor and then taking its limit yields

$$K_0(\mathbf{x}, \mathbf{y}; \tau) = \left(\frac{m}{2\pi i \hbar \tau} \right)^{d/2} \exp \left[\frac{im|\mathbf{x} - \mathbf{y}|^2}{2\hbar \tau} \right]. \quad (59.3)$$

For positive τ , choose the branch obtained continuously from the regularized integral. Negative times follow by $K_0(\mathbf{x}, \mathbf{y}; -\tau) = K_0(\mathbf{y}, \mathbf{x}; \tau)^*$. The limit $\tau \rightarrow 0$ is a delta distribution, not an ordinary pointwise limit of the oscillating expression.

59.3 Gaussian spreading, current, and momentum uncertainty

Take a normalized initial packet with variance σ_0^2 in each Cartesian coordinate:

$$\Psi(\mathbf{x}, 0) = \frac{1}{(2\pi\sigma_0^2)^{d/4}} \exp \left[-\frac{\mathbf{x}^2}{4\sigma_0^2} \right].$$

Convolution with (59.3) gives another Gaussian. Define $b(t) = \sigma_0^2 + i\hbar t/(2m)$. Before integration, the exponent is

$$\frac{im}{2\hbar t} |\mathbf{x} - \mathbf{y}|^2 - \frac{\mathbf{y}^2}{4\sigma_0^2} = \frac{imb(t)}{2\hbar\sigma_0^2 t} \left(\mathbf{y} - \frac{\sigma_0^2}{b(t)} \mathbf{x} \right)^2 - \frac{\mathbf{x}^2}{4b(t)}.$$

The square here is the algebraic dot product, without complex conjugation; the shifted center is generally complex. The original real integration contour can be shifted because the Gaussian is entire and damped. Its integral is

$$\int_{\mathbb{R}^d} \exp \left[\frac{imb(t)}{2\hbar\sigma_0^2 t} \left(\mathbf{y} - \frac{\sigma_0^2}{b(t)} \mathbf{x} \right)^2 \right] d^d y = \left(\frac{2\pi i \hbar \sigma_0^2 t}{mb(t)} \right)^{d/2}.$$

Multiplying this by the kernel prefactor and initial normalization gives

$$\Psi(\mathbf{x}, t) = \frac{1}{(2\pi)^{d/4}} \left(\frac{\sigma_0}{b(t)} \right)^{d/2} \exp \left[-\frac{\mathbf{x}^2}{4b(t)} \right]. \quad (59.4)$$

The square root is continued from $t = 0$. Taking the absolute square,

$$|\Psi|^2 = \frac{1}{[2\pi\sigma(t)^2]^{d/2}} \exp \left[-\frac{\mathbf{x}^2}{2\sigma(t)^2} \right], \quad \sigma(t)^2 = \sigma_0^2 + \left(\frac{\hbar t}{2m\sigma_0} \right)^2.$$

The dispersion time is $t_d = 2m\sigma_0^2/\hbar$. At long times, $\sigma(t) \sim \hbar|t|/(2m\sigma_0)$. A narrower initial packet spreads faster because its momentum uncertainty is larger: $\Delta p_i = \hbar/(2\sigma_0)$. Momentum is conserved, so the long-time width is precisely $|t|\Delta p_i/m$.

To obtain the current, differentiate the logarithm of (59.4):

$$\nabla \Psi = -\frac{\mathbf{x}}{2b(t)} \Psi, \quad \mathbf{j} = |\Psi|^2 \frac{\hbar^2 t}{4m^2 \sigma_0^4 + \hbar^2 t^2} \mathbf{x}.$$

For $t > 0$, probability flows outward from the center. The associated velocity field is $(\dot{\sigma}/\sigma)\mathbf{x}$, which is exactly the velocity required to expand a Gaussian of width $\sigma(t)$.

59.4 A moving Gaussian and a general quadratic phase

For initial mean position $\mathbf{x}_0$ and momentum $\mathbf{p}_0$,

$$\Psi(\mathbf{x}, 0) = \frac{1}{(2\pi\sigma_0^2)^{d/4}} \exp \left[-\frac{|\mathbf{x} - \mathbf{x}_0|^2}{4\sigma_0^2} + \frac{i\mathbf{p}_0 \cdot (\mathbf{x} - \mathbf{x}_0)}{\hbar} \right].$$

Let $\mathbf{v}_0 = \mathbf{p}_0/m$. The evolved solution is

$$\Psi(\mathbf{x}, t) = e^{i[\mathbf{p}_0 \cdot (\mathbf{x} - \mathbf{x}_0) - p_0^2 t / (2m)] / \hbar} \Psi_0(\mathbf{x} - \mathbf{x}_0 - \mathbf{v}_0 t, t),$$

where Ψ_0 is the zero-mean packet in (59.4). The center follows $\langle \mathbf{x} \rangle = \mathbf{x}_0 + \mathbf{v}_0 t$, while the width evolves as before.

A complex Gaussian in one dimension can initially be written

$$\Psi(x, 0) = \mathcal{N} \exp[-A_0(x - x_0)^2 + ip_0(x - x_0)/\hbar], \quad \text{Re } A_0 > 0.$$

Separate the translating center and its plane-wave phase, and write the remaining packet as $C(t) \exp[-A(t)(x - v_0 t - x_0)^2]$. Matching the quadratic and constant terms in the free Schrödinger equation gives

$$\dot{A} = -\frac{2i\hbar}{m} A^2, \quad \frac{\dot{C}}{C} = -\frac{i\hbar}{m} A.$$

The first equation integrates to $A(t)^{-1} = A_0^{-1} + 2i\hbar t/m$; inserting it into the second gives

$$A(t) = \frac{A_0}{1 + 2i\hbar A_0 t/m}, \quad \frac{C(t)}{C(0)} = (1 + 2i\hbar A_0 t/m)^{-1/2}.$$

The imaginary part of A_0 is an initial position-momentum correlation. Equivalently, the Heisenberg identity $x(t) = x(0) + p(0)t/m$ gives

$$(\Delta x(t))^2 = (\Delta x_0)^2 + \frac{2t}{m} \text{Cov}(x, p)_0 + \frac{t^2}{m^2} (\Delta p_0)^2.$$

A negatively correlated packet may initially contract before spreading.

59.5 Propagator on a circle

On a circle of circumference L , the momentum labels are $p_n = 2\pi\hbar n/L$. Therefore

$$K_L(x, y; t) = \frac{1}{L} \sum_{n \in \mathbb{Z}} \exp \left[\frac{2\pi i n(x - y)}{L} - \frac{i\hbar t}{2m} \left(\frac{2\pi n}{L} \right)^2 \right].$$

Alternatively, sum the free-space kernel over winding numbers:

$$K_L(x, y; t) = \sum_{w \in \mathbb{Z}} K_0(x, y + wL; t).$$

To verify the equality directly, put $X = x - y$ and expand the periodic image sum in Fourier modes. Its n th coefficient is

$$\begin{aligned} c_n(t) &= \frac{1}{L} \int_0^L \sum_{w \in \mathbb{Z}} K_0(X - wL, 0; t) e^{-2\pi i n X/L} dX \\ &= \frac{1}{L} \int_{-\infty}^{\infty} K_0(X, 0; t) e^{-2\pi i n X/L} dX \\ &= \frac{1}{L} \exp \left[-\frac{i\hbar t}{2m} \left(\frac{2\pi n}{L} \right)^2 \right]. \end{aligned}$$

In the second step, translating each interval by wL does not change the Fourier phase, and the translated intervals cover the real line. In the last step, the Fourier transform of the free kernel is the free evolution phase. These are exactly the coefficients in the spectral formula. The image representation makes periodicity explicit; the spectral one makes the discrete energies explicit. Both reduce to the periodic delta function at $t = 0$.

60 Oscillator propagation, coherent states, and squeezing

60.1 Mehler's formula and the oscillator kernel

The normalized Hermite eigenfunctions are

$$\psi_n(x) = \left(\frac{m\omega}{\pi\hbar} \right)^{1/4} \frac{H_n(u) e^{-u^2/2}}{\sqrt{2^n n!}}, \quad u = \sqrt{m\omega/\hbar} x.$$

Their spectral sum is

$$K(x, y; t) = \sqrt{\frac{m\omega}{\pi\hbar}} e^{-(u^2+v^2)/2 - i\omega t/2} \sum_{n=0}^{\infty} \frac{\rho^n H_n(u) H_n(v)}{2^n n!}, \quad \rho = e^{-i\omega t}.$$

Mehler's formula, initially valid for $|\rho| < 1$, is

$$\sum_{n=0}^{\infty} \frac{\rho^n H_n(u) H_n(v)}{2^n n!} = \frac{1}{\sqrt{1-\rho^2}} \exp \left[\frac{2\rho uv - \rho^2(u^2 + v^2)}{1-\rho^2} \right].$$

Apply it with a small damping regulator, then take its boundary value. Since $1 - \rho^2 = 2ie^{-i\omega t} \sin \omega t$ and $1 + \rho^2 = 2e^{-i\omega t} \cos \omega t$, one obtains, for $0 < \omega t < \pi$,

$$K(x, y; t) = \sqrt{\frac{m\omega}{2\pi i \hbar \sin \omega t}} \exp \left[\frac{im\omega}{2\hbar} \frac{(x^2 + y^2) \cos \omega t - 2xy}{\sin \omega t} \right]. \quad (60.1)$$

The short-time limit reproduces the free kernel.

The square root must be followed continuously through successive intervals; using the principal square root independently at every time loses phases. For positive noncaustic times, put $n = \lfloor \omega t / \pi \rfloor$. An unambiguous equivalent prefactor is

$$\sqrt{\frac{m\omega}{2\pi\hbar|\sin\omega t|}} \exp[-i\pi/4 - in\pi/2].$$

At $t = n\pi/\omega$, the kernel becomes the distribution

$$K(x, y; n\pi/\omega) = e^{-in\pi/2} \delta(x - (-1)^n y).$$

In particular, $U(\pi/\omega) = -iP$ and $U(2\pi/\omega) = -\mathbf{1}$. Hence the propagator is antiperiodic after one classical period, while physical rays and all expectation values return to their initial values. The minus sign is due to the zero-point energy.

60.2 A coherent Gaussian moving without distortion

Take an initial Gaussian centered at x_0 , at rest, with the ground-state width $\sigma_g = \sqrt{\hbar/(2m\omega)}$:

$$\Psi(x, 0) = \frac{1}{(2\pi\sigma_g^2)^{1/4}} \exp[-(x - x_0)^2/(4\sigma_g^2)].$$

Insert this into the kernel integral and put $\theta = \omega t$. Using $m\omega/(2\hbar) = 1/(4\sigma_g^2)$, the exponent of the integrand is

$$S(y) = \frac{i[(x^2 + y^2) \cos \theta - 2xy]}{4\sigma_g^2 \sin \theta} - \frac{(y - x_0)^2}{4\sigma_g^2}.$$

Complete the square explicitly:

$$\begin{aligned} S(y) = & -\frac{e^{i\theta}}{4i\sigma_g^2 \sin \theta} [y - (x + ix_0 \sin \theta)e^{-i\theta}]^2 \\ & - \frac{(x - x_0 e^{-i\theta})^2}{4\sigma_g^2} - \frac{ix_0^2 e^{-i\theta} \sin \theta}{4\sigma_g^2}. \end{aligned}$$

The integral of the first factor contributes $[4\pi i \sigma_g^2 \sin \theta e^{-i\theta}]^{1/2}$. Together with the kernel prefactor this leaves the phase $e^{-i\theta/2}$, on the branch continued from $t = 0$. Thus an intermediate form of the result is

$$\Psi(x, t) = \frac{e^{-i\theta/2}}{(2\pi\sigma_g^2)^{1/4}} \exp \left[-\frac{(x - x_0 e^{-i\theta})^2}{4\sigma_g^2} - \frac{ix_0^2 e^{-i\theta} \sin \theta}{4\sigma_g^2} \right].$$

Separating the real and imaginary parts of its exponent gives

$$\begin{aligned} \Psi(x, t) = & \frac{e^{-i\omega t/2}}{(2\pi\sigma_g^2)^{1/4}} \exp \left[-\frac{(x - x_0 \cos \omega t)^2}{4\sigma_g^2} \right] \\ & \times \exp \left[-\frac{ixx_0 \sin \omega t}{2\sigma_g^2} + \frac{ix_0^2 \sin \omega t \cos \omega t}{4\sigma_g^2} \right]. \end{aligned} \quad (60.2)$$

The absolute square is a Gaussian of fixed width σ_g , with

$$\langle x(t) \rangle = x_0 \cos \omega t, \quad \langle p(t) \rangle = -m\omega x_0 \sin \omega t.$$

The second expression follows by differentiating the spatial phase. These are exactly the classical oscillator trajectory and momentum. The current is $j = (\langle p \rangle / m) |\Psi|^2$, translating the complete profile.

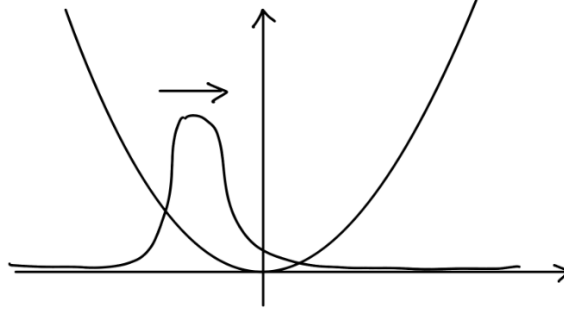


Figure 55: A coherent Gaussian follows the classical oscillator motion while retaining its width.

For arbitrary initial (x_0, p_0) , replace the center and momentum by

$$x_c(t) = x_0 \cos \omega t + \frac{p_0}{m\omega} \sin \omega t, \quad p_c(t) = p_0 \cos \omega t - m\omega x_0 \sin \omega t.$$

The state remains a minimum-uncertainty Gaussian. In ladder notation it is an eigenstate of a with $\alpha = \sqrt{m\omega/(2\hbar)}x_0 + ip_0/\sqrt{2m\hbar\omega}$, and time evolution sends $\alpha \mapsto \alpha e^{-i\omega t}$, apart from the common zero-point phase.

60.3 A squeezed packet and its breathing width

If the initial real Gaussian has width $\sigma_0 \neq \sigma_g$, the center still follows the same classical orbit, but the width changes. The Heisenberg equations give

$$x(t) = x(0) \cos \omega t + \frac{p(0)}{m\omega} \sin \omega t,$$

$$p(t) = p(0) \cos \omega t - m\omega x(0) \sin \omega t.$$

Initially $\text{Cov}(x, p) = 0$ and $(\Delta p_0)^2 = \hbar^2/(4\sigma_0^2)$. Thus

$$(\Delta x(t))^2 = \sigma_0^2 \cos^2 \omega t + \frac{\hbar^2}{4m^2\omega^2\sigma_0^2} \sin^2 \omega t, \quad (60.3)$$

$$(\Delta p(t))^2 = \frac{\hbar^2}{4\sigma_0^2} \cos^2 \omega t + m^2\omega^2\sigma_0^2 \sin^2 \omega t, \quad (60.4)$$

$$\text{Cov}(x, p)_t = \left(\frac{\hbar^2}{4m\omega\sigma_0^2} - m\omega\sigma_0^2 \right) \sin \omega t \cos \omega t. \quad (60.5)$$

The width oscillates at frequency 2ω . At a quarter period, position and momentum squeezing exchange roles. Although $\Delta x \Delta p$ generally exceeds $\hbar/2$, the full covariance relation remains saturated:

$$(\Delta x)^2(\Delta p)^2 - \text{Cov}(x, p)^2 = \frac{\hbar^2}{4}.$$

For a zero-mean state, define $b(t) = \cos \omega t + i[\sigma_g^2/\sigma_0^2] \sin \omega t$. The full Gaussian solution is

$$\Psi(x, t) = \frac{b(t)^{-1/2}}{(2\pi\sigma_0^2)^{1/4}} \exp \left[-\frac{x^2}{4\sigma_0^2} \frac{\cos \omega t + i(\sigma_0^2/\sigma_g^2) \sin \omega t}{\cos \omega t + i(\sigma_g^2/\sigma_0^2) \sin \omega t} \right].$$

The continuous square-root branch supplies the correct global phase. Setting $\sigma_0 = \sigma_g$ reduces this to the stationary ground-state Gaussian with phase $e^{-i\omega t/2}$.

Part VIII

Scattering and Resonances

Chapter 8

61 Stationary scattering amplitudes and probability flux

61.1 Incoming and outgoing solutions

Consider a real one-dimensional potential that vanishes outside $0 < x < L$. At a positive energy $E = \hbar^2 k^2 / (2m)$, $k > 0$, an energy eigenfunction has the exterior form

$$\psi(x) = \begin{cases} Ae^{ikx} + Be^{-ikx}, & x < 0, \\ Ce^{ikx} + De^{-ikx}, & x > L. \end{cases} \quad (61.1)$$

The solution inside the interaction region is obtained from the Schrödinger equation and matched at the boundaries. As a second-order ODE it has two independent solutions: specifying the two left coefficients (A, B) determines (C, D) , subject to the usual matching conditions.

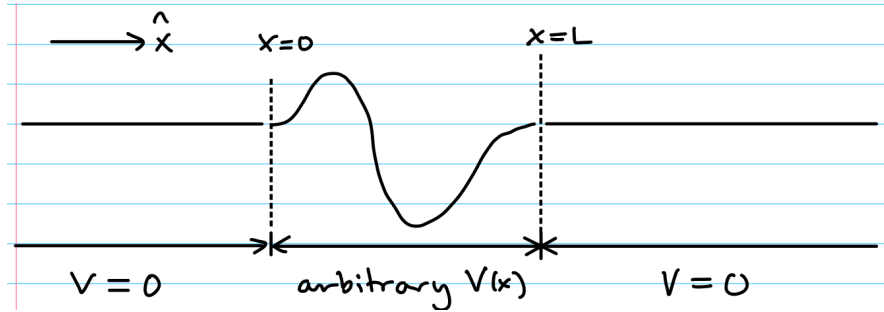


Figure 56: A localized potential with free-particle regions on both sides.

With the time factor $e^{-iEt/\hbar}$ understood, e^{ikx} travels to the right and e^{-ikx} to the left. Thus A, D are incoming amplitudes and B, C outgoing amplitudes.

61.2 The conserved current and flux normalization

A stationary density obeys $\partial_x j = 0$, with

$$j = \frac{\hbar}{m} \text{Im}(\psi^* \psi').$$

In the left exterior region,

$$\psi^* \psi' = ik|A|^2 - ik|B|^2 + ikB^* A e^{2ikx} - ikA^* B e^{-2ikx}.$$

The two cross terms have a real sum, so they do not contribute to its imaginary part. Therefore

$$j_{\text{left}} = \frac{\hbar k}{m} (|A|^2 - |B|^2), \quad j_{\text{right}} = \frac{\hbar k}{m} (|C|^2 - |D|^2).$$

Conservation gives

$$|A|^2 + |D|^2 = |B|^2 + |C|^2. \quad (61.2)$$

The total incoming flux equals the outgoing flux. The simple squared amplitudes apply because both asymptotic regions have the same velocity. If asymptotic potentials differ, a transmitted probability includes the ratio of transmitted to incoming group velocities.

For incidence only from the left, set $A = 1$, $D = 0$ and define

$$\psi_k^L(x) = \begin{cases} e^{ikx} + r(k)e^{-ikx}, & x < 0, \\ t(k)e^{ikx}, & x > L. \end{cases} \quad (61.3)$$

Then

$$R(k) = |r(k)|^2, \quad T(k) = |t(k)|^2, \quad R(k) + T(k) = 1. \quad (61.4)$$

The quantities r, t are complex amplitudes; their absolute squares are reflection and transmission probabilities. Their phases carry additional information about the motion of wavepackets.

61.3 Recovering arbitrary incidence from one scattering solution

For real V and real k , ψ_k^{L*} is another solution. Form

$$\psi_k^R = \frac{\psi_k^{L*} - r^* \psi_k^L}{t^*}.$$

On the left, its right-going component cancels, leaving

$$\psi_k^R(x) = \frac{1 - |r|^2}{t^*} e^{-ikx} = t e^{-ikx}, \quad x < 0.$$

On the right,

$$\psi_k^R(x) = e^{-ikx} - \frac{r^* t}{t^*} e^{ikx}.$$

Thus the right-incidence transmission amplitude equals t , while the right reflection amplitude is $r_R = -r^* t / t^*$. A general incoming combination $A\psi_k^L + D\psi_k^R$ gives

$$B = rA + tD, \quad C = tA - \frac{r^* t}{t^*} D. \quad (61.5)$$

Equivalently,

$$\begin{pmatrix} B \\ C \end{pmatrix} = S(k) \begin{pmatrix} A \\ D \end{pmatrix}, \quad S(k) = \begin{pmatrix} r & t \\ t & -r^* t / t^* \end{pmatrix}.$$

Its columns have unit norm and zero mutual inner product, so $S^\dagger S = \mathbf{1}$. This is unitarity expressed in the basis of incoming and outgoing channels. For a finite nonsingular potential and $k > 0$, t is nonzero; limiting cases with an impenetrable barrier can be taken separately. Spatial reflection symmetry, when the origin is chosen at the center of the potential, further equates the two reflection amplitudes.

62 Scattering of localized wavepackets

62.1 Building a packet from stationary scattering states

To interpret the stationary amplitudes, form a packet initially far to the left:

$$\Psi(x, t) = \mathcal{N} \int_0^\infty dk e^{-\ell^2(k-k_0)^2 - ikx_0 - i\hbar k^2 t/(2m)} \psi_k^L(x), \quad (62.1)$$

where $k_0 > 0$, $x_0 < 0$, $k_0\ell \gg 1$, and $|x_0| \gg \ell$. In the narrow-band, far-left limit, normalization gives $\mathcal{N} = (\ell^2/(2\pi^3))^{1/4}$. Indeed, the squared modulus of the initial Gaussian integrates to $|\mathcal{N}|^2(\pi/\ell^2)\sqrt{2\pi}\ell = 1$. Since the Gaussian is exponentially small near $k = 0$, integrals may be extended to the whole real axis with exponentially small error.

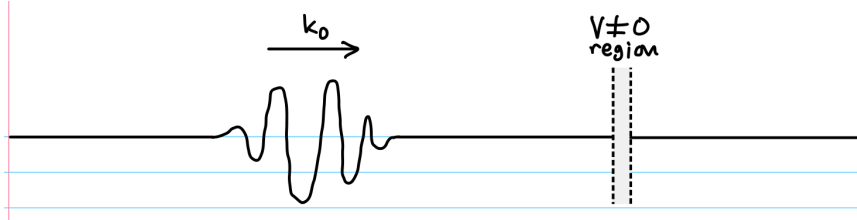


Figure 57: A narrow momentum packet approaches the localized scattering region from the left.

At $t = 0$, the incoming contribution for $x < 0$ is proportional to

$$\int_{-\infty}^{\infty} e^{-\ell^2(k-k_0)^2 + ik(x-x_0)} dk = \frac{\sqrt{\pi}}{\ell} e^{-(x-x_0)^2/(4\ell^2) + ik_0(x-x_0)}.$$

It is the desired Gaussian centered at x_0 with position variance ℓ^2 . If $r(k)$ is nearly constant over its momentum width, the reflected piece initially has the form

$$r(k_0) \frac{\sqrt{\pi}}{\ell} e^{-(x+x_0)^2/(4\ell^2) - ik_0(x+x_0)}.$$

Its formal center is $-x_0 > 0$, outside the region $x < 0$ where this exterior formula applies. Consequently it is exponentially small in the left region at the initial time. The transmitted exterior contribution is also initially exponentially small for $x > L$. The superposition therefore describes a single incoming packet rather than three simultaneously present distant packets.

62.2 Exact free dispersion and the separated outgoing packets

The Gaussian integral underlying the time-dependent pieces is

$$\begin{aligned} I(Y, t) &= \int_{-\infty}^{\infty} dk e^{-\ell^2(k-k_0)^2 + ikY - i\hbar k^2 t/(2m)} \\ &= \sqrt{\frac{\pi}{A(t)}} e^{ik_0 Y - i\hbar k_0^2 t/(2m)} \exp \left[-\frac{(Y - v_0 t)^2}{4A(t)} \right]. \end{aligned} \quad (62.2)$$

where

$$A(t) = \ell^2 + \frac{i\hbar t}{2m}, \quad v_0 = \frac{\hbar k_0}{m}.$$

To derive it, put $k = k_0 + q$, collect the quadratic term $-Aq^2$ and linear term $iq(Y - v_0 t)$, complete the square, and use $\int e^{-Aq^2} dq = \sqrt{\pi/A}$ with $\text{Re } A > 0$.

The width becomes $\sqrt{\ell^2 + [\hbar t/(2m\ell)]^2}$. If the collision occurs near $t_c = |x_0|/v_0$, dispersion during the encounter is small when

$$t_c \ll \frac{2m\ell^2}{\hbar}, \quad \text{or approximately} \quad \frac{|x_0|}{k_0} \ll \ell^2 \ll x_0^2.$$

The last inequality also guarantees initial separation from the scatterer. In the leading narrow-band approximation,

$$\begin{aligned} \Psi_{\text{in}} &\simeq \mathcal{N} I(x - x_0, t), \\ \Psi_{\text{ref}} &\simeq \mathcal{N} r(k_0) I(-x - x_0, t), \\ \Psi_{\text{trans}} &\simeq \mathcal{N} t(k_0) I(x - x_0, t). \end{aligned}$$

After the interaction, the reflected center is $-x_0 - v_0 t$ and the transmitted center is $x_0 + v_0 t$. They move away from the interaction region in opposite directions.

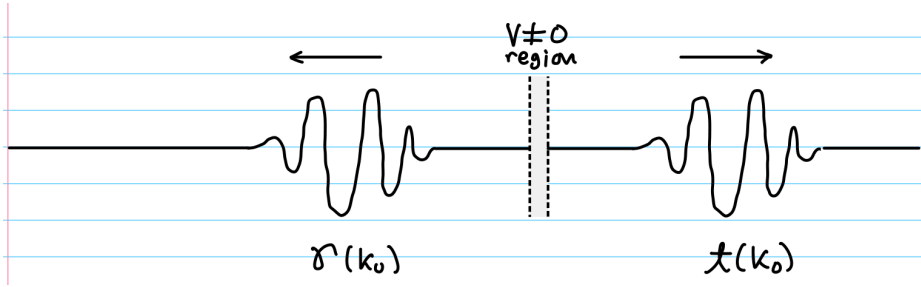


Figure 58: After scattering, the outgoing packets have reflection and transmission amplitudes $r(k_0)$ and $t(k_0)$.

Integrating the probability in each separated packet gives $R(k_0)$ and $T(k_0)$ when the initial packet has unit norm. More generally, a packet with nonnegligible momentum width samples the weighted averages of $|r(k)|^2$ and $|t(k)|^2$ over its momentum distribution.

62.3 Phase derivatives and time delays

To retain the leading variation in transmission amplitude, expand its logarithm near a nonzero $t(k_0)$:

$$\log t(k) = \log t(k_0) + (\alpha + i\beta)(k - k_0) + O((k - k_0)^2),$$

where

$$\alpha = \text{Re} \frac{t'(k_0)}{t(k_0)}, \quad \beta = \text{Im} \frac{t'(k_0)}{t(k_0)} = \left. \frac{d \arg t}{dk} \right|_{k_0}.$$

The real linear term shifts the Gaussian momentum center to $\tilde{k}_0 = k_0 + \alpha/(2\ell^2)$. The imaginary term replaces x_0 by $\tilde{x}_0 = x_0 - \beta$. Thus

$$x_{\text{trans}}(t) \simeq x_0 - \beta + \frac{\hbar}{m} \left(k_0 + \frac{\alpha}{2\ell^2} \right) t.$$

For sufficiently narrow momentum width and times short compared with strong dispersive reshaping, the velocity-filtering term is negligible. The packet is then shifted backward by β , which corresponds to the transmission delay

$$\tau_t(k) = \frac{m}{\hbar k} \text{Im} \frac{d}{dk} \log t(k) = \hbar \frac{d \arg t}{dE}. \quad (62.3)$$

The reflection calculation is identical after reversing the outgoing coordinate:

$$\tau_r(k) = \frac{m}{\hbar k} \text{Im} \frac{d}{dk} \log r(k).$$

These delays are measured relative to free reference propagation with the same exterior phase convention. Near zeros or very rapid variation of an amplitude, a single translated packet may not be a good description; filtering and distortion must then be kept.

63 A delta scatterer and analytic continuation

63.1 Exact reflection and transmission

Let $V(x) = g\delta(x)$, where g may have either sign, and define $\kappa_0 = mg/\hbar^2$. For left incidence,

$$\psi(x) = \begin{cases} e^{ikx} + re^{-ikx}, & x < 0, \\ te^{ikx}, & x > 0. \end{cases}$$

Continuity and the derivative jump give

$$1 + r = t, \quad \psi'(0^+) - \psi'(0^-) = 2\kappa_0\psi(0),$$

so

$$ik(t - 1 + r) = 2\kappa_0 t.$$

Using $t = 1 + r$, solve for

$$r(k) = \frac{\kappa_0}{ik - \kappa_0} = \frac{-i\kappa_0}{k + i\kappa_0}, \quad t(k) = \frac{ik}{ik - \kappa_0} = \frac{k}{k + i\kappa_0}. \quad (63.1)$$

Hence

$$R = \frac{\kappa_0^2}{k^2 + \kappa_0^2}, \quad T = \frac{k^2}{k^2 + \kappa_0^2}.$$

At high energy $T \rightarrow 1$. At low energy $T \rightarrow 0$ for every nonzero g , including an attractive delta. Strong reflection is therefore not limited to a classical barrier higher than the incident energy.

Since

$$\frac{d}{dk} \log t = \frac{1}{k} - \frac{1}{k + i\kappa_0},$$

the phase derivative is $\kappa_0/(k^2 + \kappa_0^2)$. The same imaginary part occurs for r . Thus

$$\tau_r = \tau_t = \frac{m\kappa_0}{\hbar k(k^2 + \kappa_0^2)}. \quad (63.2)$$

A repulsive delta gives a positive delay, an attractive one a negative shift relative to the chosen free reference.

63.2 The high-energy classical limit

At large $p_0 = \hbar k$,

$$\tau \sim \frac{m^2 g}{p_0^3}.$$

For a smooth localized potential with $|V| \ll p_0^2/(2m)$, the classical travel-time difference is

$$\tau_{\text{cl}} = \int \left[\frac{m}{\sqrt{p_0^2 - 2mV(x)}} - \frac{m}{p_0} \right] dx.$$

Expand $(1 - z)^{-1/2} = 1 + z/2 + O(z^2)$ to obtain

$$\tau_{\text{cl}} \simeq \frac{m^2}{p_0^3} \int V(x) dx.$$

The delta result agrees when g is identified with the integrated potential strength. This comparison concerns the high-energy limit of a narrow potential, rather than a literal classical trajectory through a singular delta distribution.

63.3 A pole is not automatically a bound state

The delta amplitudes have a pole at

$$k_* = -i\kappa_0, \quad E_* = -\frac{\hbar^2 \kappa_0^2}{2m}.$$

For $g > 0$, the pole lies on the negative imaginary axis. The Hamiltonian is nonnegative, so it has no negative-energy bound state. This pole is a virtual or antibound state, not a normalizable eigenstate. For $g < 0$, the pole lies on the positive imaginary axis, with $\kappa = -\kappa_0 > 0$. The associated solution is

$$\psi(x) = \sqrt{\kappa} e^{-\kappa|x|}, \quad E = -\frac{\hbar^2 \kappa^2}{2m},$$

precisely the attractive delta bound state. The sign of the imaginary momentum determines whether the exterior exponentials decay or grow.

64 Poles, bound states, and a double-delta potential

64.1 Why upper-half-plane poles give bound states

For compactly supported sufficiently regular real potentials, the solutions and scattering amplitudes can be continued in complex k . The differential equation depends analytically on k^2 , even though normalization and conjugation identities require separate attention. A generic complex- k scattering solution grows at at least one spatial infinity and is not a physical Hilbert-space eigenvector.

Suppose a common pole occurs at $k_* = i\kappa$ with $\kappa > 0$. Divide the left-incidence solution by $t(k)$ and take the pole limit:

$$\frac{\psi_k^L(x)}{t(k)} = \begin{cases} t(k)^{-1}e^{ikx} + [r(k)/t(k)]e^{-ikx}, & x < 0, \\ e^{ikx}, & x > L. \end{cases}$$

At the pole, the incoming coefficient vanishes. The limiting solution is purely outgoing in the analytic sense:

$$\psi_*(x) \propto \begin{cases} c_L e^{\kappa x}, & x < 0, \\ c_R e^{-\kappa x}, & x > L. \end{cases}$$

Both tails decay, so it is normalizable. For a nonsingular finite interaction, a bound state cannot vanish on an entire exterior interval: zero value and derivative at an endpoint would force the ODE solution to vanish everywhere. This is why the physical bound-state pole appears in both reflection and transmission, though their residues differ.

If a pole in the upper half-plane had $k = a + ib$ with $a \neq 0$, its outgoing tails would still decay but its energy $\hbar^2(a^2 - b^2 + 2iab)/(2m)$ would be nonreal. Self-adjointness rules this out. Hence physical upper-half-plane poles lie on the positive imaginary axis. Negative-imaginary-axis poles instead have growing tails and do not represent bound states.

64.2 Exact amplitudes for two identical delta functions

Consider

$$V(x) = g[\delta(x) + \delta(x - L)], \quad \kappa_0 = \frac{mg}{\hbar^2}.$$

For a single delta write $t_1 = k/(k + i\kappa_0)$ and $r_1 = -i\kappa_0/(k + i\kappa_0)$. A wave transmitted through both barriers may make any number of round trips between them. A round trip contributes $r_1^2 e^{2ikL}$. Summing the geometric series gives

$$t = \frac{t_1^2}{1 - r_1^2 e^{2ikL}}.$$

For reflection, include direct reflection from the first delta and every path that transmits through it twice and reflects from the second:

$$r = r_1 + \frac{t_1^2 r_1 e^{2ikL}}{1 - r_1^2 e^{2ikL}}.$$

Thus

$$t(k) = \frac{k^2}{(k + i\kappa_0)^2 + \kappa_0^2 e^{2ikL}}, \quad (64.1)$$

$$r(k) = -i\kappa_0 \frac{k + i\kappa_0 + (k - i\kappa_0)e^{2ikL}}{(k + i\kappa_0)^2 + \kappa_0^2 e^{2ikL}}. \quad (64.2)$$

The lack of an overall e^{ikL} in t reflects the convention that the transmitted wave is te^{ikx} in the original global coordinate. The pole equation factors into two parity branches:

$$k = -i\kappa_0(1 \pm e^{ikL}). \quad (64.3)$$

A zero of a denominator must still be checked against its numerator; in particular, the universal algebraic root $k = 0$ is not automatically a pole of the scattering amplitudes.

64.3 Imaginary poles and the number of bound states

Put $k = i\kappa$, $y = \kappa L$, and $\lambda = \kappa_0 L$. Then

$$y = -\lambda(1 \pm e^{-y}), \quad \lambda = -\frac{y}{1 \pm e^{-y}}. \quad (64.4)$$

These curves determine the imaginary-axis poles graphically.

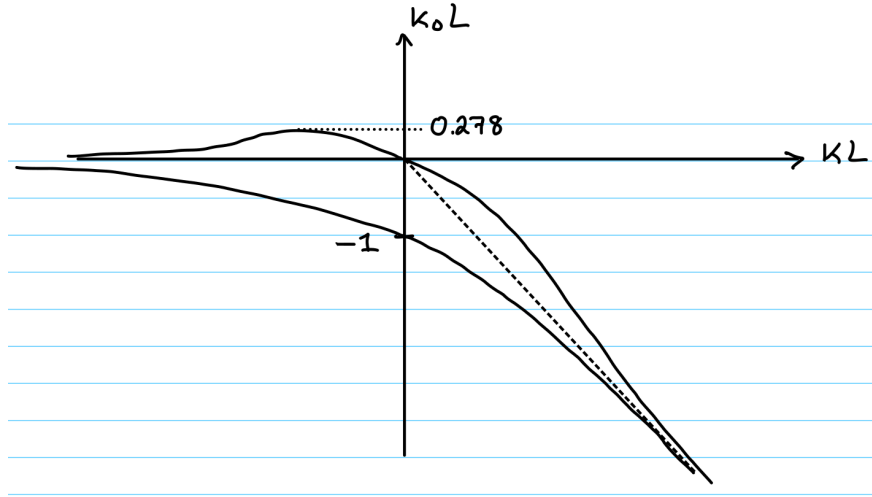


Figure 59: Imaginary-axis poles of the double-delta amplitudes as the coupling changes.

For attraction, let $a = -\lambda > 0$. The even branch is $y = a(1 + e^{-y})$ and always has one positive solution. The odd branch is $y = a(1 - e^{-y})$. Near zero its right-hand side has slope a ; it has a nonzero positive solution only when $a > 1$. Consequently there is one bound state for $-1 \leq \lambda < 0$ and two for $\lambda < -1$. At $\lambda = -1$, the second state is at threshold and is not square-integrable. For $-1 < \lambda < 0$, the corresponding odd pole lies on the negative imaginary axis.

For repulsion $\lambda > 0$, all imaginary poles have $y < 0$ and are virtual states. Let $u = -y > 0$ on the plus branch. Then $\lambda = u/(1 + e^u)$. Its maximum occurs when

$$(u - 1)e^u = 1.$$

Writing W for the Lambert function, defined by $W(z)e^{W(z)} = z$, this gives

$$\lambda_c = W(e^{-1}) \simeq 0.27846.$$

For $0 < \lambda < \lambda_c$ there are two negative imaginary poles. At λ_c they merge, and for $\lambda > \lambda_c$ they leave the imaginary axis as a resonance pair. The free case $\lambda = 0$ must be treated separately because numerator cancellations remove the interaction poles.

64.4 The infinite families of complex poles

Define $\chi = \lambda - ikL$. Equation (64.3) becomes

$$\chi e^\chi = \pm \lambda e^\lambda.$$

Thus, formally,

$$k = \frac{i}{L} [W_n(\pm \lambda e^\lambda) - \lambda], \quad n \in \mathbb{Z},$$

with cancelled roots excluded. The infinitely many branches display the infinite resonance families produced by repeated propagation between the two scatterers.

To obtain the graphical relation in real variables, let $\eta = \text{Re}(kL)$ and $\xi = \lambda + \text{Im}(kL)$, so $\chi = \xi - i\eta$. The imaginary part of χe^χ must vanish:

$$\xi \sin \eta + \eta \cos \eta = 0, \quad \xi = -\eta \cot \eta.$$

Its real part is then $-\eta e^{-\eta \cot \eta} / \sin \eta$. Therefore

$$\lambda e^\lambda = \pm \frac{\eta e^{-\eta \cot \eta}}{\sin \eta}, \quad \text{Im}(kL) = -\eta \cot \eta - \lambda. \quad (64.5)$$

The case $\eta = 0$ is handled by the imaginary-axis analysis above.

All poles with nonzero real part lie in the lower half-plane for this self-adjoint scattering problem. Reality of the potential implies $t(-k^*) = t(k)^*$ and the corresponding relation for r , so resonances occur in pairs $k = k_R - i\gamma$ and $k = -k_R - i\gamma$, $\gamma > 0$.

65 Resonances, time delay, and unstable trapped states

65.1 A narrow pole produces a Lorentzian peak

Focus on a pole $k_* = k_R - i\gamma$ with $k_R > 0$ and small $\gamma > 0$. Near it,

$$t(k) = \frac{c_{-1}}{k - k_R + i\gamma} + \text{a regular background}.$$

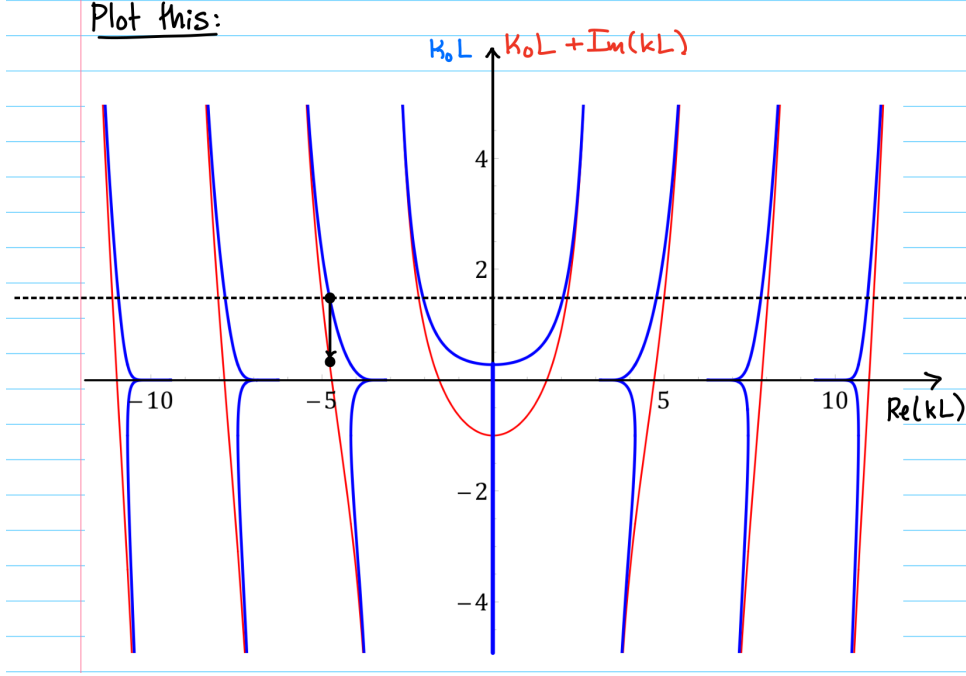


Figure 60: The graphical construction of double-delta resonance poles. The horizontal coordinate determines the real part of kL and the curve difference determines its imaginary part.

If the pole term dominates and the background changes slowly,

$$T(k) \simeq \frac{|c_{-1}|^2}{(k - k_R)^2 + \gamma^2}.$$

The complete amplitude still obeys $T \leq 1$; the residue and background are constrained by unitarity. The phase derivative is

$$\text{Im} \frac{d}{dk} \log t \simeq \frac{\gamma}{(k - k_R)^2 + \gamma^2}.$$

Thus

$$\tau_t(k) \simeq \frac{m}{\hbar k} \frac{\gamma}{(k - k_R)^2 + \gamma^2}, \quad \tau_{\max} \simeq \frac{m}{\hbar k_R \gamma}. \quad (65.1)$$

A narrow resonance has a long delay: the particle can remain near the scatterer before escaping.

The complex pole energy is

$$E_* = \frac{\hbar^2}{2m} (k_R^2 - \gamma^2 - 2ik_R \gamma) = E_{\text{res}} - \frac{i}{2} \Gamma_E, \quad \Gamma_E = \frac{2\hbar^2 k_R \gamma}{m}. \quad (65.2)$$

For $\gamma \ll k_R$, the energy-space Lorentzian has full width Γ_E . The amplitude time factor is $e^{-iE_{\text{res}}t/\hbar} e^{-\Gamma_E t/(2\hbar)}$. The survival probability therefore has decay time

$$\tau_{\text{decay}} = \frac{\hbar}{\Gamma_E} = \frac{m}{2\hbar k_R \gamma}. \quad (65.3)$$

For a pure isolated pole, the peak phase delay is twice this probability lifetime. Both become large as the pole approaches the real axis.

65.2 Outgoing resonance functions and self-adjointness

At the pole, the incoming wave disappears after suitable rescaling, and the formal stationary solution has outgoing tails

$$\psi_*(x) = \begin{cases} C_L e^{-ik_* x} = C_L e^{-ik_R x - \gamma x}, & x < 0, \\ C_R e^{ik_* x} = C_R e^{ik_R x + \gamma x}, & x > L. \end{cases}$$

They grow exponentially as $|x| \rightarrow \infty$. This function is not a normalizable eigenvector of the self-adjoint Hamiltonian. Its complex energy consequently does not contradict the reality of the Hilbert-space spectrum. It is a resonance solution with outgoing boundary conditions, not an ordinary stationary state prepared with unit probability.

The role of the boundary is seen directly. On a finite interval $[a, b]$, multiply the stationary equation by ψ^* and integrate:

$$E \int_a^b |\psi|^2 dx = \int_a^b \left[\frac{\hbar^2}{2m} |\psi'|^2 + V |\psi|^2 \right] dx - \frac{\hbar^2}{2m} [\psi^* \psi']_a^b.$$

Taking imaginary parts gives

$$\text{Im } E = -\frac{\hbar}{2} \frac{j(b) - j(a)}{\int_a^b |\psi|^2 dx}. \quad (65.4)$$

For a normalizable eigenfunction in the self-adjoint domain, the boundary form vanishes as the endpoints go to infinity, so $\text{Im } E = 0$. For the outgoing resonance, outward flux and a growing exterior norm remain; the limiting argument for a normalizable state does not apply.

There can be no nontrivial purely outgoing solution at a strictly positive real k for a real potential with equal exterior channels. The current would be negative on the left and positive on the right, whereas stationary current conservation requires them to be equal. Both amplitudes would have to vanish. This rules out real positive scattering poles; the threshold $k = 0$ needs separate treatment.

65.3 Why the formal spatial growth has a physical interpretation

Multiply the resonance function by its complex-energy time factor. On the right its magnitude is proportional to

$$\exp[\gamma x - \hbar k_R \gamma t / m] = \exp[\gamma(x - v_R t)], \quad v_R = \hbar k_R / m.$$

The left outgoing branch has the analogous dependence on $-x - v_R t$. At a fixed location, probability decays as $\exp(-t/\tau_{\text{decay}})$. Farther outgoing particles were emitted earlier. Because the trapped population was larger at earlier times, the formal outgoing profile grows with distance.

The decay rate per surviving trapped particle is constant in the exponential regime, but the number of surviving particles decreases. The bare outgoing solution extends

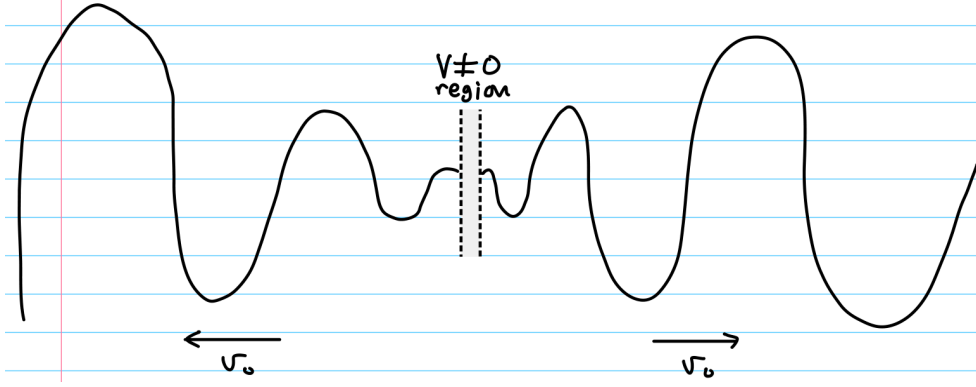


Figure 61: The outgoing resonance profile grows with distance because distant parts correspond to earlier emission. A physical packet supplies an outer envelope.

that history indefinitely into the past and is therefore not normalizable. A real scattering experiment has a finite incoming packet and a finite excitation time. Its outgoing resonance contribution must be cut off by a moving envelope. The next calculation constructs that envelope explicitly.

66 A normalizable scattering packet near a resonance

66.1 Separating the pole from the smooth background

Expand the transmission amplitude near a lower-half-plane pole:

$$t(k) = \frac{c_{-1}}{k - k_*} + c_0 + O(k - k_*).$$

The smooth terms produce the ordinary transmitted packet and its phase-dependent shift. The pole contribution to the transmitted packet is

$$\Psi_{\text{pole}}(x, t) = \mathcal{N} c_{-1} \int_{-\infty}^{\infty} \frac{\exp[-\ell^2(k - k_0)^2 + ik(x - x_0) - i\hbar k^2 t / (2m)]}{k - k_*} dk, \quad x > L. \quad (66.1)$$

As before, extending the lower endpoint from zero to minus infinity is justified by $k_0 \ell \gg 1$. The pole expansion is useful when the packet samples the neighborhood where this term controls the rapid variation.

66.2 Deriving the Gaussian-pole integral

For $\text{Re } A > 0$ and $\text{Im } z < 0$, define

$$I(A, B, z) = \int_{-\infty}^{\infty} \frac{e^{-Au^2 + Bu}}{u - z} du.$$

Because $\text{Im}(u - z) > 0$ on the real axis,

$$\frac{1}{u - z} = -i \int_0^\infty e^{i(u-z)s} ds.$$

Insert this representation and perform the Gaussian u integral:

$$\begin{aligned} I &= -i \sqrt{\frac{\pi}{A}} \int_0^\infty \exp \left[-izs + \frac{(B + is)^2}{4A} \right] ds \\ &= -i \sqrt{\frac{\pi}{A}} e^{B^2/(4A)} \int_0^\infty \exp \left[-\frac{s^2}{4A} + i \left(\frac{B}{2A} - z \right) s \right] ds. \end{aligned}$$

Complete the square in s . The remaining half-line Gaussian is expressed through the complementary error function:

$$\text{erf } w = \frac{2}{\sqrt{\pi}} \int_0^w e^{-v^2} dv, \quad \text{erfc } w = 1 - \text{erf } w.$$

For complex arguments this definition is understood by analytic continuation. The result is

$$I(A, B, z) = -i\pi e^{-Az^2 + Bz} \text{erfc} \left[\frac{i(2Az - B)}{2\sqrt{A}} \right]. \quad (66.2)$$

The square root is the branch with positive real part, continued from positive real A . The lower-half-plane condition on z fixes the contour prescription and the sign in this formula.

66.3 The exact pole packet

Set $X = x - x_0$, $A = \ell^2 + i\hbar t/(2m)$, and $B = 2\ell^2 k_0 + iX$. There is also an overall factor $e^{-\ell^2 k_0^2}$ from expanding the initial Gaussian. Applying (66.2) gives

$$\begin{aligned} \Psi_{\text{pole}}(x, t) &= -i\pi \mathcal{N} c_{-1} e^{-\ell^2(k_* - k_0)^2 + ik_* X - i\hbar k_*^2 t/(2m)} \text{erfc } Z(x, t), \\ Z(x, t) &= \frac{2i\ell^2(k_* - k_0) + X - \hbar t k_*/m}{2\sqrt{\ell^2 + i\hbar t/(2m)}}. \end{aligned} \quad (66.3)$$

The first exponential is exactly the formal outgoing resonance solution, multiplied by the incident packet's spectral weight at the complex pole. The error-function factor provides the missing spatial envelope.

Choose the incident central momentum to match the resonance: $k_0 = k_R$, $k_* = k_R - i\gamma$, and $v_R = \hbar k_R/m$. Then

$$\begin{aligned} -\ell^2(k_* - k_0)^2 &= \ell^2 \gamma^2, \\ Z(x, t) &= \gamma \sqrt{A(t)} + \frac{x - x_0 - v_R t}{2\sqrt{A(t)}}. \end{aligned} \quad (66.4)$$

This form keeps all dispersive effects. In the regime $\hbar|t|/(2m) \ll \ell^2$, replace $\sqrt{A}$ by ℓ to obtain

$$Z(x, t) \simeq \ell \gamma + \frac{x - x_0 - v_R t}{2\ell}.$$

For real $z \gg 1$, $\operatorname{erfc} z$ is exponentially small; for $z \ll -1$ it is approximately two. Define the moving front

$$x_1(t) = x_0 + v_R t - 2\ell^2 \gamma. \quad (66.5)$$

Away from a transition region of width of order ℓ ,

$$\Psi_{\text{pole}}(x, t) \simeq \begin{cases} -2\pi i \mathcal{N}_{C-1} e^{\ell^2 \gamma^2 + i k_* (x - x_0) - i \hbar k_*^2 t / (2m)}, & L < x \lesssim x_1(t), \\ 0, & x \gtrsim x_1(t). \end{cases}$$

Behind the front, the pole wave has the exponentially decaying time behavior of the trapped state. Ahead of it, the error-function envelope suppresses the unphysical growing tail. When $x_1(t) < L$, this outgoing resonance region has not yet emerged on the transmitted side.

The exact integral (66.1) is a superposition with Gaussian momentum weighting and no real-axis singularity because $\operatorname{Im} k_* < 0$. It is therefore compatible with a normalizable wavepacket. The full scattering state also contains the incident, reflected, and smooth transmitted pieces, whose interference describes formation and decay of the trapped component. The resonant eigenfunction is thus a useful intermediate description of a physical time-dependent process, even though that eigenfunction by itself is not a normalizable stationary state.

Part IX

Approximation Methods

Chapter 9

67 The variational method

Exactly solvable Hamiltonians are useful both as physical models and as starting points for approximations. Three approaches will be developed here. The variational method bounds energies by optimizing trial states. Perturbation theory expands about a solvable Hamiltonian. The WKB method uses the slow spatial variation of the local classical momentum. Their assumptions differ, and their answers should be interpreted accordingly: a variational energy is an upper bound, whereas a truncated perturbative or semiclassical expression need not bound the exact answer.

67.1 The ground-state energy as a minimum

Suppose that H is self-adjoint and bounded below, and that its lowest energy is E_0 . For any nonzero trial state in the domain of H , define the Rayleigh quotient

$$\mathcal{E}[\psi] = \frac{\langle \psi | H | \psi \rangle}{\langle \psi | \psi \rangle}. \quad (67.1)$$

In a discrete orthonormal energy basis, $|\psi\rangle = \sum_n c_n |n\rangle$, this is

$$\mathcal{E}[\psi] = \frac{\sum_n |c_n|^2 E_n}{\sum_n |c_n|^2} = E_0 + \frac{\sum_n |c_n|^2 (E_n - E_0)}{\sum_n |c_n|^2} \geq E_0. \quad (67.2)$$

The same argument uses a spectral integral when continuous spectrum is present: the spectral measure is nonnegative and supported at energies at least E_0 .

A parameterized trial family $\psi(x; \alpha_1, \dots, \alpha_r)$ converts an infinite-dimensional minimization into an ordinary one. Every member gives an upper bound, so the smallest value within the family is its best energy bound. An overall nonzero multiplier of ψ has no effect on (67.1). Normalizing the trial states is convenient, but unnecessary if the denominator is retained.

67.2 A Gaussian trial state for the quartic oscillator

Consider

$$H = \frac{p^2}{2m} + \lambda x^4, \quad \lambda > 0. \quad (67.3)$$

There is no harmonic term. Choose the normalized even Gaussian

$$\psi_\ell(x) = \frac{1}{(2\pi\ell^2)^{1/4}} \exp\left(-\frac{x^2}{4\ell^2}\right), \quad \ell > 0. \quad (67.4)$$

Its probability distribution has variance ℓ^2 . Gaussian integration gives

$$\langle x^2 \rangle = \ell^2, \quad \langle x^4 \rangle = 3\ell^4.$$

For the kinetic energy, differentiation yields

$$\psi'_\ell = -\frac{x}{2\ell^2}\psi_\ell, \quad \psi''_\ell = \left(\frac{x^2}{4\ell^4} - \frac{1}{2\ell^2}\right)\psi_\ell.$$

Consequently,

$$\left\langle \frac{p^2}{2m} \right\rangle = -\frac{\hbar^2}{2m} \left(\frac{\langle x^2 \rangle}{4\ell^4} - \frac{1}{2\ell^2} \right) = \frac{\hbar^2}{8m\ell^2}, \quad (67.5)$$

$$\mathcal{E}(\ell) = \frac{\hbar^2}{8m\ell^2} + 3\lambda\ell^4. \quad (67.6)$$

A narrow wavefunction lowers the potential energy but raises the kinetic energy. The optimum balances these two effects:

$$\frac{d\mathcal{E}}{d\ell} = -\frac{\hbar^2}{4m\ell^3} + 12\lambda\ell^3 = 0, \quad \ell_{\text{opt}} = \left(\frac{\hbar^2}{48m\lambda} \right)^{1/6}.$$

Substitution gives

$$E_0 \leq \mathcal{E}_{\min} = \left(\frac{3}{4} \right)^{4/3} \frac{\hbar^{4/3}\lambda^{1/3}}{m^{2/3}} \simeq 0.6814 \frac{\hbar^{4/3}\lambda^{1/3}}{m^{2/3}}. \quad (67.7)$$

The numerical ground-state coefficient is approximately 0.6680, so this one-parameter trial state already gives a fairly accurate energy. The combination $\hbar^{4/3}\lambda^{1/3}m^{-2/3}$ is fixed by dimensions; the nontrivial output is the dimensionless coefficient.

67.3 Why an approximate state can give a better energy

Let $|0\rangle$ be a normalized ground state and let $|\phi\rangle$ be a normalized orthogonal state. Write a nearby normalized trial state as

$$|\psi\rangle = \frac{|0\rangle + \epsilon|\phi\rangle}{\sqrt{1 + |\epsilon|^2}}, \quad \langle 0 | \phi \rangle = 0.$$

Because $H|0\rangle = E_0|0\rangle$ and H is self-adjoint, $\langle 0|H|\phi\rangle = \langle \phi|H|0\rangle = 0$. Therefore

$$\mathcal{E}[\psi] - E_0 = \frac{|\epsilon|^2}{1 + |\epsilon|^2} (\langle \phi|H|\phi\rangle - E_0). \quad (67.8)$$

The state error begins at order ϵ , but its energy error begins at order $|\epsilon|^2$. This stationarity of the Rayleigh quotient explains why energy estimates are often more accurate than the trial wavefunctions that produce them. It does not imply comparable accuracy for every observable.

67.4 Excited-state bounds and parity

If a trial state is orthogonal to the first r exact energy eigenstates, its expansion starts at E_r , and the same weighted-average argument gives $\mathcal{E}[\psi] \geq E_r$. Orthogonality is essential: minimizing over unrestricted states returns the ground state.

Symmetry can enforce the needed orthogonality without knowing the ground wavefunction. For a one-dimensional even confining potential, the nondegenerate ground state is even, whereas the lowest odd state is the first excited state. For the quartic oscillator choose

$$\psi_\ell^{\text{odd}}(x) = \frac{1}{(2\pi\ell^2)^{1/4}} \frac{x}{\ell} \exp\left(-\frac{x^2}{4\ell^2}\right). \quad (67.9)$$

Its normalization follows from $\langle x^2 \rangle_{\psi_\ell} = \ell^2$. For the potential term,

$$\langle x^4 \rangle_{\text{odd}} = \frac{\langle x^6 \rangle_{\psi_\ell}}{\ell^2} = 15\ell^4.$$

Differentiating (67.9), or integrating $\hbar^2|\psi'|^2/(2m)$, gives

$$\left\langle \frac{p^2}{2m} \right\rangle_{\text{odd}} = \frac{\hbar^2}{2m\ell^2} \left\langle \left(1 - \frac{x^2}{2\ell^2}\right)^2 \right\rangle_{\psi_\ell} = \frac{3\hbar^2}{8m\ell^2}.$$

Thus

$$\mathcal{E}_{\text{odd}}(\ell) = \frac{3\hbar^2}{8m\ell^2} + 15\lambda\ell^4, \quad (67.10)$$

$$\ell_{\text{opt}}^{\text{odd}} = \left(\frac{\hbar^2}{80m\lambda} \right)^{1/6}, \quad (67.11)$$

$$E_1 \leq \mathcal{E}_{\text{odd},\text{min}} = \frac{9 \cdot 5^{1/3}}{2^{8/3}} \frac{\hbar^{4/3} \lambda^{1/3}}{m^{2/3}} \simeq 2.424 \frac{\hbar^{4/3} \lambda^{1/3}}{m^{2/3}}. \quad (67.12)$$

The numerical coefficient is approximately 2.394, about 1.3% below this bound. The odd trial function has the required node at the origin.

68 Nondegenerate stationary perturbation theory

Consider a Hamiltonian

$$H = H_0 + \epsilon H_1, \quad H_0|n^{(0)}\rangle = E_n^{(0)}|n^{(0)}\rangle. \quad (68.1)$$

The parameter ϵ keeps track of orders; after the calculation it can be set to one if its smallness has been included in H_1 . We first assume the unperturbed level of interest is isolated and nondegenerate. Small denominators signal the need to reconsider this assumption.

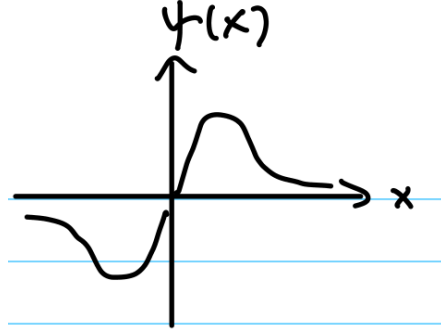


Figure 62: An odd Gaussian trial wavefunction for the first excited state of an even potential.

68.1 Intermediate normalization and two exact equations

Let $|n\rangle$ be the exact eigenstate continuously associated with $|n^{(0)}\rangle$. As long as their overlap is nonzero, its scale and phase can be chosen so that

$$\langle n^{(0)} | n \rangle = 1. \quad (68.2)$$

This is *intermediate normalization*; it does not set $\langle n | n \rangle = 1$. Write

$$|n\rangle = |n^{(0)}\rangle + |n_\Delta\rangle, \quad \langle n^{(0)} | n_\Delta \rangle = 0, \quad E_n = E_n^{(0)} + \Delta_n.$$

The exact eigenvalue equation becomes

$$(E_n^{(0)} - H_0)|n_\Delta\rangle = (\epsilon H_1 - \Delta_n)|n\rangle. \quad (68.3)$$

Multiplication by $\langle n^{(0)} |$ gives the exact energy relation

$$\Delta_n = \epsilon \langle n^{(0)} | H_1 | n \rangle. \quad (68.4)$$

To extract the orthogonal component of the state, the zero eigenvalue of $E_n^{(0)} - H_0$ must be treated carefully.

68.2 The pseudoinverse and the reduced resolvent

For a normal operator with spectral decomposition $N = \sum_j \lambda_j |j\rangle \langle j|$, define its spectral pseudoinverse by

$$N^+ = \sum_{\lambda_j \neq 0} \lambda_j^{-1} |j\rangle \langle j|.$$

The symbol $+$ here denotes the pseudoinverse, not the Hermitian adjoint. If Π_0 projects onto the null space, then

$$N^+ N = N N^+ = \mathbf{1} - \Pi_0.$$

Thus applying N^+ to $N|x\rangle = |y\rangle$ determines the component of $|x\rangle$ outside the null space. A component in the null space remains undetermined. For the self-adjoint

operators used here, the solvability condition is that the source have zero projection onto the null space.

Define the reduced resolvent

$$R_n = (E_n^{(0)} - H_0)^+ = \sum_{m \neq n} \frac{|m^{(0)}\rangle \langle m^{(0)}|}{E_n^{(0)} - E_m^{(0)}}. \quad (68.5)$$

It annihilates $|n^{(0)}\rangle$, and $R_n(E_n^{(0)} - H_0) = \mathbf{1} - |n^{(0)}\rangle \langle n^{(0)}|$. Applying it to (68.3) yields

$$|n_\Delta\rangle = R_n(\epsilon H_1 - \Delta_n)|n\rangle. \quad (68.6)$$

Equations (68.4) and (68.6) are an exact coupled formulation of the eigenvalue problem near the chosen state.

68.3 Expansion and recursion

Expand

$$E_n = E_n^{(0)} + \epsilon E_n^{(1)} + \epsilon^2 E_n^{(2)} + \dots, \quad (68.7)$$

$$|n\rangle = |n^{(0)}\rangle + \epsilon |n^{(1)}\rangle + \epsilon^2 |n^{(2)}\rangle + \dots, \quad \langle n^{(0)} | n^{(k)} \rangle = 0 \quad (k \geq 1). \quad (68.8)$$

Substituting in the exact equations and comparing powers of ϵ gives

$$E_n^{(k)} = \langle n^{(0)} | H_1 | n^{(k-1)} \rangle \quad (k \geq 1), \quad (68.9)$$

and

$$|n^{(k)}\rangle = R_n \left[H_1 |n^{(k-1)}\rangle - \sum_{j=1}^{k-1} E_n^{(j)} |n^{(k-j)}\rangle \right]. \quad (68.10)$$

The term with $j = k$ would be proportional to $|n^{(0)}\rangle$ and is annihilated by R_n . In particular,

$$E_n^{(1)} = \langle n^{(0)} | H_1 | n^{(0)} \rangle, \quad (68.11)$$

$$|n^{(1)}\rangle = R_n H_1 |n^{(0)}\rangle, \quad (68.12)$$

$$E_n^{(2)} = \langle n^{(0)} | H_1 R_n H_1 | n^{(0)} \rangle, \quad (68.13)$$

$$|n^{(2)}\rangle = R_n (H_1 - E_n^{(1)}) R_n H_1 |n^{(0)}\rangle. \quad (68.14)$$

In the unperturbed energy basis these formulas read

$$|n^{(1)}\rangle = \sum_{m \neq n} \frac{\langle m^{(0)} | H_1 | n^{(0)} \rangle}{E_n^{(0)} - E_m^{(0)}} |m^{(0)}\rangle, \quad (68.15)$$

$$E_n^{(2)} = \sum_{m \neq n} \frac{|\langle m^{(0)} | H_1 | n^{(0)} \rangle|^2}{E_n^{(0)} - E_m^{(0)}}. \quad (68.16)$$

For the ground state all denominators in the second-order energy are negative, so $E_0^{(2)} \leq 0$. For an excited state, levels above and below it contribute with opposite signs.

A necessary practical smallness condition is that the mixing coefficients

$$\left| \frac{\epsilon \langle m^{(0)} | H_1 | n^{(0)} \rangle}{E_n^{(0)} - E_m^{(0)}} \right|$$

be small for states coupled appreciably to $|n^{(0)}\rangle$. The overall size of a matrix element alone is insufficient; the energy separation matters.

Intermediate-normalized states can be converted to unit norm at the end. To second order,

$$\langle n | n \rangle = 1 + \epsilon^2 \langle n^{(1)} | n^{(1)} \rangle + O(\epsilon^3),$$

so the normalized state acquires the additional component $-\frac{1}{2}\epsilon^2 \langle n^{(1)} | n^{(1)} \rangle |n^{(0)}\rangle$. This component is a normalization effect, not additional mixing.

68.4 A longitudinal perturbation of a two-state system

Let

$$H_0 = \frac{E_0}{2} \sigma_z, \quad H_1 = \frac{E_1}{2} \sigma_z, \quad E_0 > 0, \quad |E_1| \ll E_0,$$

and absorb the bookkeeping parameter into E_1 . The unperturbed states $|\uparrow\rangle, |\downarrow\rangle$ have energies $E_0/2, -E_0/2$. The first-order shifts are

$$E_{\uparrow}^{(1)} = \frac{E_1}{2}, \quad E_{\downarrow}^{(1)} = -\frac{E_1}{2}.$$

For example,

$$R_{\uparrow} = \frac{1}{E_0} |\downarrow\rangle \langle \downarrow|, \quad R_{\uparrow} H_1 |\uparrow\rangle = 0.$$

There is no state mixing and all higher energy shifts vanish. Indeed, $H = (E_0 + E_1) \sigma_z / 2$ is diagonal in the same basis and has the exact energies $\pm(E_0 + E_1)/2$.

68.5 A transverse perturbation and exact comparison

Now take

$$H = \frac{E_0}{2} \sigma_z + \frac{E_1}{2} \sigma_x.$$

The diagonal matrix elements of H_1 vanish, while $\langle \downarrow | H_1 | \uparrow \rangle = E_1/2$. Hence

$$E_{\uparrow}^{(1)} = E_{\downarrow}^{(1)} = 0, \tag{68.17}$$

$$E_{\uparrow}^{(2)} = \frac{E_1^2}{4E_0}, \quad E_{\downarrow}^{(2)} = -\frac{E_1^2}{4E_0}, \tag{68.18}$$

$$|\uparrow^{(1)}\rangle = \frac{E_1}{2E_0} |\downarrow\rangle, \quad |\downarrow^{(1)}\rangle = -\frac{E_1}{2E_0} |\uparrow\rangle. \tag{68.19}$$

Diagonalizing the 2×2 matrix directly gives

$$E_{\pm} = \pm \frac{1}{2} \sqrt{E_0^2 + E_1^2} = \pm \left(\frac{E_0}{2} + \frac{E_1^2}{4E_0} + O(E_1^4/E_0^3) \right).$$

This agrees with perturbation theory.

Exercise 68.1 (Exact eigenvectors and their expansion). Compare the exact eigenvectors with the perturbative state correction.

Define θ by

$$\cos \theta = \frac{E_0}{\sqrt{E_0^2 + E_1^2}}, \quad \sin \theta = \frac{E_1}{\sqrt{E_0^2 + E_1^2}},$$

taking the branch with $\theta \rightarrow 0$ as $E_1 \rightarrow 0$. Normalized eigenvectors are

$$|+\rangle = \cos \frac{\theta}{2} |\uparrow\rangle + \sin \frac{\theta}{2} |\downarrow\rangle, \quad |-\rangle = -\sin \frac{\theta}{2} |\uparrow\rangle + \cos \frac{\theta}{2} |\downarrow\rangle.$$

With $\eta = E_1/E_0$, their expansions are

$$\begin{aligned} |+\rangle &= \left(1 - \frac{\eta^2}{8}\right) |\uparrow\rangle + \frac{\eta}{2} |\downarrow\rangle + O(\eta^3), \\ |-\rangle &= -\frac{\eta}{2} |\uparrow\rangle + \left(1 - \frac{\eta^2}{8}\right) |\downarrow\rangle + O(\eta^3). \end{aligned}$$

Dividing the first by $\cos(\theta/2)$ gives its intermediate-normalized form $|\uparrow\rangle + \tan(\theta/2)|\downarrow\rangle$. The first-order coefficient is $\eta/2$, exactly as obtained from (68.15).

68.6 The anharmonic oscillator

Take

$$H = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2 + \lambda x^4 = H_0 + H_1. \quad (68.20)$$

Writing

$$x = \sqrt{\frac{\hbar}{2m\omega}}(a + a^\dagger), \quad H_0 = \hbar\omega \left(a^\dagger a + \frac{1}{2}\right)$$

gives

$$E_n^{(1)} = \lambda \left(\frac{\hbar}{2m\omega}\right)^2 \langle n|(a + a^\dagger)^4|n\rangle.$$

For the ground state,

$$(a + a^\dagger)^2|0\rangle = |0\rangle + \sqrt{2}|2\rangle.$$

Since $(a + a^\dagger)^2$ is Hermitian,

$$\langle 0|(a + a^\dagger)^4|0\rangle = \|(a + a^\dagger)^2|0\rangle\|^2 = 1 + 2 = 3.$$

Equivalently, commuting each annihilation operator to the right in the fourth power leaves three vacuum contractions. The needed commutations follow from $aa^\dagger = a^\dagger a + \mathbf{1}$: for example $\langle 0|a^2(a^\dagger)^2|0\rangle = 2$ and $\langle 0|aa^\dagger aa^\dagger|0\rangle = 1$, while the other surviving-order possibilities either vanish against the vacuum or reduce to these terms. Therefore

$$E_0 = \frac{\hbar\omega}{2} + 3\lambda \left(\frac{\hbar}{2m\omega}\right)^2 + \dots \quad (68.21)$$

Exercise 68.2 (The shift of an arbitrary oscillator level). Compute $\langle n|x^4|n\rangle$ and determine when the first-order energy shift becomes comparable to the unperturbed energy.

Apply the square of $a + a^\dagger$ first:

$$(a + a^\dagger)^2|n\rangle = \sqrt{n(n-1)}|n-2\rangle + (2n+1)|n\rangle + \sqrt{(n+1)(n+2)}|n+2\rangle. \quad (68.22)$$

The squared norm is

$$n(n-1) + (2n+1)^2 + (n+1)(n+2) = 6n^2 + 6n + 3.$$

Consequently,

$$E_n^{(1)} = \lambda \left(\frac{\hbar}{2m\omega} \right)^2 (6n^2 + 6n + 3). \quad (68.23)$$

At large n , comparison with $E_n^{(0)} \simeq n\hbar\omega$ gives

$$\frac{E_n^{(1)}}{E_n^{(0)}} \simeq \frac{3\lambda\hbar n}{2m^2\omega^3}.$$

The correction becomes order one when $n \sim 2m^2\omega^3/(3\lambda\hbar)$. Thus a perturbation that is small for low states need not be small throughout the spectrum. The same conclusion follows from the oscillator size $\langle x^2 \rangle_n = (n + \frac{1}{2})\hbar/(m\omega)$: sufficiently energetic states explore positions where λx^4 competes with the quadratic term.

For $\lambda > 0$, evaluating the full Hamiltonian in the unperturbed ground state gives precisely the first two terms displayed in (68.21). The variational principle therefore makes this particular truncation a ground-state upper bound. More generally, whenever the full Hamiltonian is bounded below and the unperturbed ground state is an admissible trial vector,

$$E_{\text{gs}}(H) \leq \langle 0^{(0)}|H|0^{(0)}\rangle = E_0^{(0)} + \epsilon E_0^{(1)}.$$

Thus the ground-state energy truncated after first order is a variational upper bound, irrespective of convergence of the perturbation series. A generic higher-order truncation does not have this property.

68.7 Analyticity is a separate question

If an eigenvalue is analytic in the perturbation parameter, the perturbation series is its Taylor series. The formal recursion alone does not prove analyticity or convergence. The quartic oscillator illustrates why this matters. For $\lambda < 0$, the potential tends to $-\infty$ as $|x| \rightarrow \infty$, and there is no ground-state energy bounded below. A harmonic-looking central well can support metastable resonances, but these are described by outgoing boundary conditions and complex resonance energies rather than normalizable ground-state eigenvectors.

Even for positive λ , the perturbation is not uniformly small in position: the quartic term eventually dominates the quadratic term at large $|x|$. Useful perturbative information can therefore exist without an ordinary convergent Taylor expansion around $\lambda = 0$.

69 Asymptotic expansions

69.1 The meaning of an asymptotic series

A function has the asymptotic expansion

$$f(x) \sim \sum_{n=0}^{\infty} a_n x^n \quad (x \rightarrow 0^+)$$

if, for each fixed integer $N \geq 0$,

$$\lim_{x \rightarrow 0^+} \frac{f(x) - \sum_{n=0}^N a_n x^n}{x^N} = 0. \quad (69.1)$$

The limit is taken with N fixed. It makes no claim that the infinite sum converges at a nonzero value of x . Taylor expansions of sufficiently regular functions provide familiar examples, but the definition also includes divergent series. The direction of approach, such as $x \rightarrow 0^+$ rather than $x \rightarrow 0^-$, is part of the statement.

At large x one often uses inverse powers: $f(x) \sim a_0 + a_1/x + a_2/x^2 + \dots$. More generally, let $\{g_n(x)\}$ be an asymptotic scale satisfying $g_{n+1}(x)/g_n(x) \rightarrow 0$ in the specified limit. Then

$$f(x) \sim \sum_{n=0}^{\infty} a_n g_n(x)$$

means that the remainder after g_N is $o(g_N)$. Powers are only one possible scale; logarithms, exponentials, and products of these also arise.

69.2 Stirling's expansion and optimal truncation

For large positive x , the gamma function has the expansion

$$\Gamma(x) \sim \sqrt{2\pi} x^{x-\frac{1}{2}} e^{-x} \left(1 + \frac{1}{12x} + \frac{1}{288x^2} - \frac{139}{51840x^3} + \dots \right). \quad (69.2)$$

It is useful to divide out the rapidly varying leading factor:

$$f(x) = \frac{\Gamma(x)e^x}{\sqrt{2\pi} x^{x-\frac{1}{2}}}.$$

The first approximations are 1, $1 + 1/(12x)$, and $1 + 1/(12x) + 1/(288x^2)$. Each additional fixed-order approximation becomes more accurate in its appropriate large- x limit.

For comparison, the rational function

$$g(x) = \frac{x^2}{1+x^2} = 1 - \frac{1}{x^2} + \frac{1}{x^4} - \frac{1}{x^6} + \dots \quad (69.3)$$

has a geometric expansion convergent for $|x| > 1$. At $x = 2$, adding terms in this geometric series improves the result indefinitely. At the same fixed argument, sufficiently

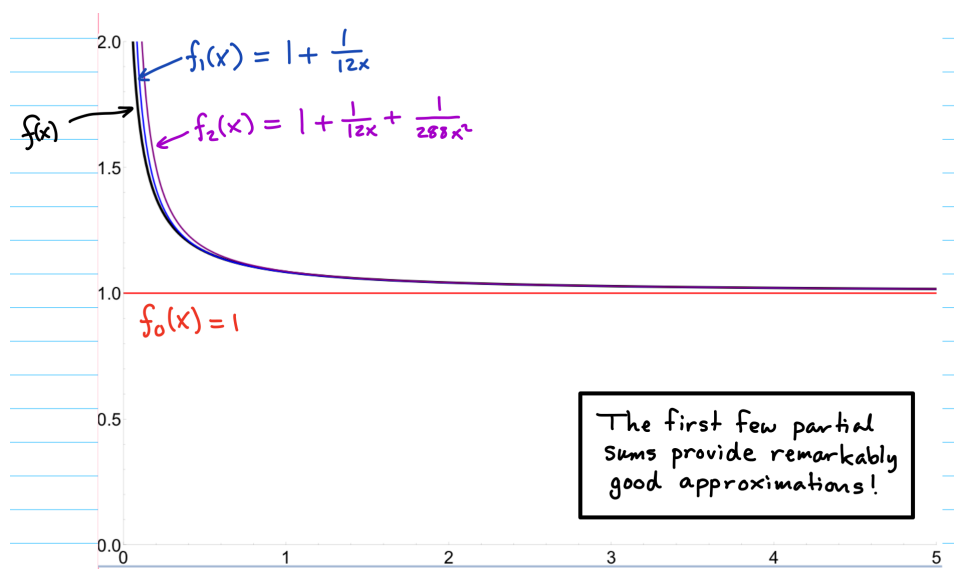


Figure 63: The scaled gamma function and successive low-order Stirling approximations. The asymptotic regime is large positive x .

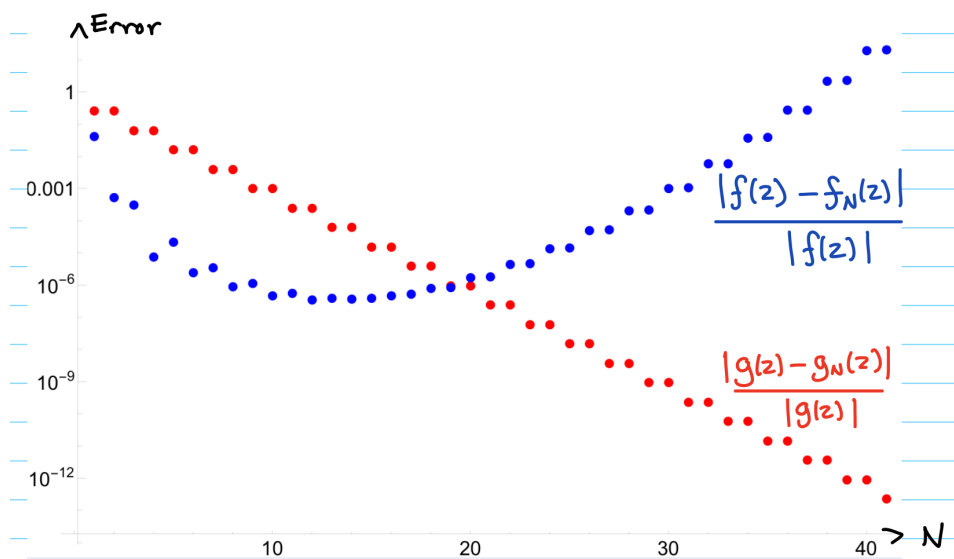


Figure 64: Errors of successive truncations at a fixed argument: the convergent geometric series continues to improve, whereas the asymptotic Stirling series has an optimal truncation order.

high-order Stirling terms eventually worsen the approximation: its coefficients grow so rapidly that the terms cease to decrease.

The distinction is between two limits. In an asymptotic statement, x approaches the expansion limit at fixed N . In a convergence statement, $N \rightarrow \infty$ at fixed x . For a divergent asymptotic series, one usually stops near the smallest term. As the parameter approaches the asymptotic limit, the useful truncation order generally increases. These properties explain how perturbation theory can be quantitatively valuable even when summing arbitrarily many terms would not converge.

70 Degenerate perturbation theory

70.1 Choosing the zeroth-order state

Suppose $E^{(0)}$ is a degenerate eigenvalue of H_0 , and let P_0 project onto its entire eigenspace. Write $Q_0 = \mathbf{1} - P_0$. A perturbation generally selects particular linear combinations within this eigenspace. Therefore the zeroth-order vector cannot be chosen arbitrarily and then kept fixed.

Let $|\psi^{(0)}\rangle \in \text{ran } P_0$ be the leading state, with

$$|\psi\rangle = |\psi^{(0)}\rangle + \epsilon|\psi^{(1)}\rangle + \epsilon^2|\psi^{(2)}\rangle + \dots, \quad E = E^{(0)} + \epsilon E^{(1)} + \epsilon^2 E^{(2)} + \dots.$$

Choose $\langle\psi^{(0)}|\psi\rangle = 1$. Set $\Delta = E - E^{(0)}$ and define

$$R_0 = (E^{(0)} - H_0)^+ = \sum_{E_m^{(0)} \neq E^{(0)}} \frac{|m^{(0)}\rangle\langle m^{(0)}|}{E^{(0)} - E_m^{(0)}}. \quad (70.1)$$

The sum excludes the *whole* degenerate eigenspace, not just one chosen basis vector. The full Schrödinger equation is equivalent to

$$P_0(\epsilon H_1 - \Delta)|\psi\rangle = 0, \quad (70.2)$$

$$Q_0|\psi\rangle = R_0(\epsilon H_1 - \Delta)|\psi\rangle. \quad (70.3)$$

The first is the condition in the null space of $E^{(0)} - H_0$; the second solves the complementary equation. Together they recover the full equation.

At first order, (70.2) gives

$$D_1|\psi^{(0)}\rangle = E^{(1)}|\psi^{(0)}\rangle, \quad D_1 = P_0 H_1 P_0. \quad (70.4)$$

Thus one diagonalizes H_1 *restricted to the degenerate eigenspace*. If $\{|\alpha^{(0)}\rangle\}_{\alpha=1}^d$ is any orthonormal basis of that space, the finite matrix is

$$(D_1)_{\alpha\beta} = \langle\alpha^{(0)}|H_1|\beta^{(0)}\rangle.$$

Its eigenvalues are the first-order shifts, and its eigenvectors specify the zeroth-order combinations that evolve into perturbed energy states.

70.2 Second-order splitting of residual degeneracy

The complementary equation first gives

$$Q_0|\psi^{(1)}\rangle = R_0H_1|\psi^{(0)}\rangle. \quad (70.5)$$

It does not determine $P_0|\psi^{(1)}\rangle$. Intermediate normalization only requires orthogonality to the *single* vector $|\psi^{(0)}\rangle$, so mixing into other vectors of $\text{ran } P_0$ is still allowed.

At second order, projection onto the original eigenspace yields

$$P_0(H_1 - E^{(1)})|\psi^{(1)}\rangle = E^{(2)}|\psi^{(0)}\rangle.$$

Substitute (70.5) and define $A = P_0H_1R_0H_1P_0$. The result is

$$(D_1 - E^{(1)}P_0)P_0|\psi^{(1)}\rangle + A|\psi^{(0)}\rangle = E^{(2)}|\psi^{(0)}\rangle. \quad (70.6)$$

Let P_1 project onto the eigenspace of D_1 with eigenvalue $E^{(1)}$ *inside* $\text{ran } P_0$. Then

$$P_0P_1 = P_1P_0 = P_1, \quad P_1(D_1 - E^{(1)}P_0) = 0.$$

Applying P_1 to (70.6) gives

$$D_2|\psi^{(0)}\rangle = E^{(2)}|\psi^{(0)}\rangle, \quad D_2 = P_1H_1R_0H_1P_1. \quad (70.7)$$

If the first-order correction leaves a degeneracy, D_2 must be diagonalized within the remaining subspace. In an orthonormal basis $\{|a\rangle\} \subset \text{ran } P_1$, its matrix is

$$(D_2)_{ab} = \sum_{E_m^{(0)} \neq E^{(0)}} \frac{\langle a|H_1|m^{(0)}\rangle \langle m^{(0)}|H_1|b\rangle}{E^{(0)} - E_m^{(0)}}. \quad (70.8)$$

The intermediate states lie outside the original degenerate eigenspace.

If P_1 has rank one, its only normalized vector, up to phase, is $|\psi^{(0)}\rangle$. There is then no further diagonalization, and

$$E^{(2)} = \langle \psi^{(0)}|H_1R_0H_1|\psi^{(0)}\rangle. \quad (70.9)$$

This resembles the nondegenerate formula, but the state and the excluded subspace have first been chosen by degenerate perturbation theory.

70.3 The first-order state within the original eigenspace

Define the restricted pseudoinverse

$$S_1 = (E^{(1)}P_0 - D_1)_{\text{ran } P_0}^+,$$

extended by zero on $\text{ran } Q_0$. If $|b\rangle$ are eigenvectors of D_1 with first-order eigenvalues $E_b^{(1)}$, it is

$$S_1 = \sum_{E_b^{(1)} \neq E^{(1)}} \frac{|b\rangle\langle b|}{E^{(1)} - E_b^{(1)}}.$$

Rearranging (70.6) and applying S_1 gives

$$(P_0 - P_1)|\psi^{(1)}\rangle = S_1 H_1 R_0 H_1 |\psi^{(0)}\rangle. \quad (70.10)$$

Consequently, the full first-order correction is

$$|\psi^{(1)}\rangle = R_0 H_1 |\psi^{(0)}\rangle + S_1 H_1 R_0 H_1 |\psi^{(0)}\rangle + P_1 |\psi^{(1)}\rangle. \quad (70.11)$$

If P_1 is one-dimensional, intermediate normalization sets the last term to zero. If it is still degenerate, its component orthogonal to $|\psi^{(0)}\rangle$ is determined by higher-order equations. In particular, if the degeneracy is first lifted at second order, determining that part of the first-order state generally requires the third-order equations. It cannot in general be discarded merely because it lies in the original degenerate space.

For a nondegenerate first-order branch a , the internal contribution can also be written

$$(P_0 - P_1)|\psi_a^{(1)}\rangle = \sum_{b \neq a} \frac{\langle b | H_1 R_0 H_1 | a \rangle}{E_a^{(1)} - E_b^{(1)}} |b\rangle.$$

Here the sum is over different first-order branches in the same unperturbed eigenspace. This makes the sign and the first-order energy denominator explicit.

70.4 A practical sequence of calculations

1. Find the degenerate eigenspace of H_0 , its projector P_0 , and the reduced resolvent R_0 excluding that entire eigenspace.
2. Diagonalize $D_1 = P_0 H_1 P_0$ to find all first-order shifts and the corresponding subspaces P_1 .
3. Within each remaining degenerate first-order subspace, diagonalize $D_2 = P_1 H_1 R_0 H_1 P_1$. This chooses the leading vectors at second order.
4. Compute mixing outside the original eigenspace using $Q_0 |\psi^{(1)}\rangle = R_0 H_1 |\psi^{(0)}\rangle$.
5. Compute mixing between distinct first-order branches using (70.10); if degeneracy remains, use higher-order solvability equations for the unresolved component.

A conserved symmetry can make these steps easier. If both H_0 and H_1 commute with a symmetry operator, different symmetry sectors do not mix. Degenerate levels in different sectors can then be treated separately. The reason is a vanishing matrix element enforced by symmetry, rather than an assumption that arbitrary vectors in a degenerate eigenspace remain eigenstates.

71 Two coupled harmonic oscillators

71.1 Degeneracy and the oscillator symmetry

Consider

$$H_0 = \frac{p_x^2 + p_y^2}{2m} + \frac{1}{2}m\omega^2(x^2 + y^2), \quad H_1 = \Delta xy, \quad (71.1)$$

with $|\Delta| \ll m\omega^2$. The bookkeeping parameter is included in Δ . Product states $|n_x, n_y\rangle$ have unperturbed energies

$$E_{n_x, n_y}^{(0)} = \hbar\omega(n_x + n_y + 1).$$

At fixed $N = n_x + n_y$, there are $N + 1$ states. Rotations of the x, y plane give an evident $\text{SO}(2)$ symmetry. The equal-frequency oscillator also has a larger symmetry: if $a = (a_x, a_y)^\top$, then

$$H_0 = \hbar\omega(a^\dagger a + 1)$$

is invariant under $a \mapsto Ua$ for $U \in U(2)$. The nontrivial $\text{SU}(2)$ part is generated by

$$J_z = \frac{1}{2}(a_x^\dagger a_x - a_y^\dagger a_y), \quad J_+ = a_x^\dagger a_y, \quad J_- = a_y^\dagger a_x.$$

They obey $[J_z, J_\pm] = \pm J_\pm$, $[J_+, J_-] = 2J_z$, commute with N , and act within a fixed energy eigenspace. At total occupation N , the magnetic quantum number runs from $-N/2$ to $N/2$, so the representation has $j = N/2$ and dimension $2j + 1 = N + 1$.

The perturbation is

$$H_1 = g(a_x + a_x^\dagger)(a_y + a_y^\dagger), \quad g = \frac{\hbar\Delta}{2m\omega}. \quad (71.2)$$

It changes each occupation number by one, with either sign when allowed. Its diagonal matrix elements in the product basis vanish.

71.2 First-order splitting of the first excited level

The ground state $|0, 0\rangle$ is nondegenerate. The first excited energy $2\hbar\omega$ has basis $|1, 0\rangle, |0, 1\rangle$. Restricted to that space,

$$D_1 = \begin{pmatrix} \langle 1, 0 | H_1 | 1, 0 \rangle & \langle 1, 0 | H_1 | 0, 1 \rangle \\ \langle 0, 1 | H_1 | 1, 0 \rangle & \langle 0, 1 | H_1 | 0, 1 \rangle \end{pmatrix} = g \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Thus

$$|\pm^{(0)}\rangle = \frac{|1, 0\rangle \pm |0, 1\rangle}{\sqrt{2}}, \quad E_\pm^{(1)} = \pm g = \pm \frac{\hbar\Delta}{2m\omega}. \quad (71.3)$$

The perturbation selects symmetric and antisymmetric combinations. Applying the nondegenerate formula directly to $|1, 0\rangle$ would instead encounter a zero denominator associated with $|0, 1\rangle$.

71.3 Second-order energy shifts

Because the first-order branches are nondegenerate for $\Delta \neq 0$,

$$E_\pm^{(2)} = \sum_{n_x + n_y \neq 1} \frac{|\langle n_x, n_y | H_1 | \pm^{(0)} \rangle|^2}{2\hbar\omega - \hbar\omega(n_x + n_y + 1)}.$$

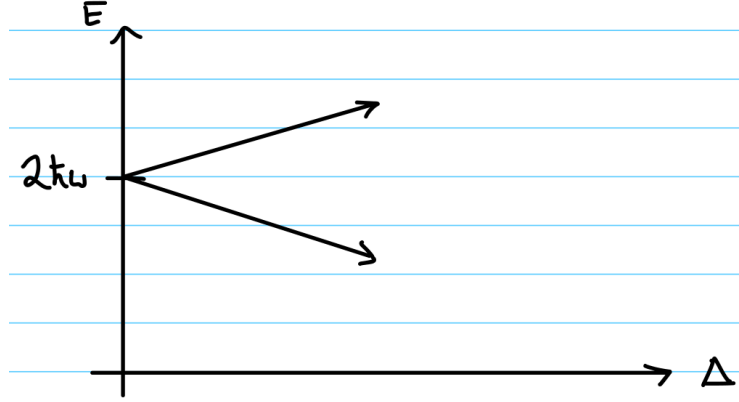


Figure 65: The first excited oscillator doublet splits at first order into symmetric and antisymmetric combinations.

To find the nonzero terms, apply the perturbation before taking overlaps:

$$H_1|1, 0\rangle = g(|0, 1\rangle + \sqrt{2}|2, 1\rangle), \quad (71.4)$$

$$H_1|0, 1\rangle = g(|1, 0\rangle + \sqrt{2}|1, 2\rangle). \quad (71.5)$$

The components $|0, 1\rangle$ and $|1, 0\rangle$ lie inside the original degenerate space and are excluded by R_0 . Outside it,

$$Q_0 H_1 |\pm^{(0)}\rangle = g(|2, 1\rangle \pm |1, 2\rangle).$$

Both intermediate states have energy $4\hbar\omega$. Therefore

$$E_{\pm}^{(2)} = \frac{g^2 + g^2}{-2\hbar\omega} = -\frac{g^2}{\hbar\omega} = -\frac{\hbar\Delta^2}{4m^2\omega^3}. \quad (71.6)$$

The same negative shift is added to both first-order branches:

$$E_{\pm} = 2\hbar\omega \pm \frac{\hbar\Delta}{2m\omega} - \frac{\hbar\Delta^2}{4m^2\omega^3} + \dots \quad (71.7)$$

The second-order effective matrix on the original doublet is proportional to the identity. Hence there is no first-order internal mixing between $|+^{(0)}\rangle$ and $|-^{(0)}\rangle$, while the external correction is

$$Q_0 |\pm^{(1)}\rangle = -\frac{g}{2\hbar\omega}(|2, 1\rangle \pm |1, 2\rangle).$$

For completeness, the ground state couples only to $|1, 1\rangle$: $H_1|0, 0\rangle = g|1, 1\rangle$. Its first-order shift vanishes and its second-order shift is

$$E_{\text{gs}}^{(2)} = \frac{g^2}{\hbar\omega - 3\hbar\omega} = -\frac{\hbar\Delta^2}{8m^2\omega^3}. \quad (71.8)$$

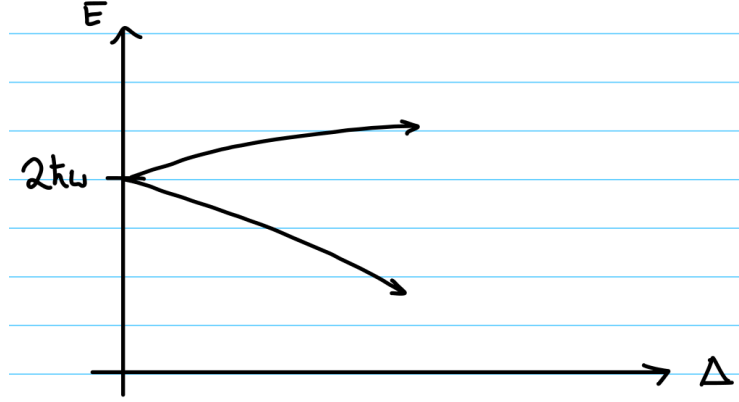


Figure 66: The two first-order branches receive the same second-order downward displacement.

71.4 Exact normal modes and perturbative comparison

Introduce an orthogonal change of coordinates and momenta,

$$u = \frac{x+y}{\sqrt{2}}, \quad v = \frac{x-y}{\sqrt{2}}, \quad p_u = \frac{p_x+p_y}{\sqrt{2}}, \quad p_v = \frac{p_x-p_y}{\sqrt{2}}.$$

Since $x^2 + y^2 = u^2 + v^2$ and $xy = (u^2 - v^2)/2$, the full Hamiltonian becomes

$$H = \frac{p_u^2}{2m} + \frac{1}{2}m\omega_u^2 u^2 + \frac{p_v^2}{2m} + \frac{1}{2}m\omega_v^2 v^2, \quad \omega_u = \sqrt{\omega^2 + \frac{\Delta}{m}}, \quad \omega_v = \sqrt{\omega^2 - \frac{\Delta}{m}}. \quad (71.9)$$

Both modes are stable when $|\Delta| < m\omega^2$. The exact energy spectrum is

$$E_{n_u, n_v} = \hbar\omega_u \left(n_u + \frac{1}{2} \right) + \hbar\omega_v \left(n_v + \frac{1}{2} \right).$$

At $\Delta = 0$, the states $|1_u, 0_v\rangle$ and $|0_u, 1_v\rangle$ reduce to the symmetric and antisymmetric combinations in (71.3). Expanding

$$\omega_{u,v} = \omega \left[1 \pm \frac{\Delta}{2m\omega^2} - \frac{\Delta^2}{8m^2\omega^4} + O\left(\frac{\Delta^3}{m^3\omega^6}\right) \right]$$

gives

$$\begin{aligned} E_{1,0} &= \frac{3}{2}\hbar\omega_u + \frac{1}{2}\hbar\omega_v = 2\hbar\omega + \frac{\hbar\Delta}{2m\omega} - \frac{\hbar\Delta^2}{4m^2\omega^3} + \cdots, \\ E_{0,1} &= \frac{1}{2}\hbar\omega_u + \frac{3}{2}\hbar\omega_v = 2\hbar\omega - \frac{\hbar\Delta}{2m\omega} - \frac{\hbar\Delta^2}{4m^2\omega^3} + \cdots. \end{aligned}$$

The exact ground-state energy similarly gives

$$\frac{\hbar}{2}(\omega_u + \omega_v) = \hbar\omega - \frac{\hbar\Delta^2}{8m^2\omega^3} + \cdots,$$

confirming (71.8).

72 The semiclassical WKB expansion

72.1 A local momentum and the phase equation

For a particle in one dimension, the time-independent Schrödinger equation can be written

$$\psi''(x) + \frac{p(x)^2}{\hbar^2} \psi(x) = 0, \quad p(x)^2 = 2m[E - V(x)]. \quad (72.1)$$

For constant positive p , its solutions are plane waves $\exp(\pm ipx/\hbar)$. If p varies slowly, one expects waves whose local phase gradient is approximately $\pm p/\hbar$.

Write

$$\psi(x) = \exp\left(\frac{i}{\hbar} W(x)\right),$$

where W can be complex so that its imaginary part describes amplitude variation. Differentiating twice gives

$$\frac{\psi''}{\psi} = \frac{i}{\hbar} W'' - \frac{(W')^2}{\hbar^2}.$$

The Schrödinger equation is therefore equivalent to the Riccati equation

$$(W')^2 - i\hbar W'' = p^2. \quad (72.2)$$

The WKB method, named for Wentzel, Kramers, and Brillouin, solves this equation as a semiclassical expansion.

72.2 Leading phase, amplitude, and the next correction

Expand $W = W_0 + \hbar W_1 + \hbar^2 W_2 + \dots$. Equating successive powers of $\hbar$ in (72.2) gives

$$(W'_0)^2 = p^2, \quad (72.3)$$

$$2W'_0 W'_1 = iW''_0, \quad (72.4)$$

$$2W'_0 W'_2 + (W'_1)^2 = iW''_1. \quad (72.5)$$

Let $s = \pm 1$ denote the direction of the leading wave. Then

$$W'_0 = sp, \quad W_0 = s \int^x p(y) dy,$$

up to an irrelevant constant. The next equation gives

$$W'_1 = \frac{i}{2} \frac{p'}{p}, \quad W_1 = \frac{i}{2} \log p.$$

Thus the first two terms yield

$$\psi_s(x) \simeq \frac{C_s}{\sqrt{p(x)}} \exp\left(\frac{si}{\hbar} \int^x p(y) dy\right). \quad (72.6)$$

The factor $p^{-1/2}$ is essential. Its squared magnitude scales as the time a classical particle spends in a given interval, since $v = p/m$. A single traveling component has current

$$j = \frac{\hbar}{m} \text{Im}(\psi_s^* \psi'_s) = \frac{s|C_s|^2}{m}$$

to this order, independent of position.

Exercise 72.1 (The second phase correction). Find W_2 and identify the first correction beyond the $p^{-1/2}$ amplitude.

Using

$$W_1'' = \frac{i}{2} \left(\frac{p''}{p} - \frac{p'^2}{p^2} \right), \quad (W_1')^2 = -\frac{p'^2}{4p^2}$$

in the last equation of (72.5) gives

$$2spW_2' = -\frac{p''}{2p} + \frac{3p'^2}{4p^2}.$$

Hence

$$W_2' = s \left(-\frac{p''}{4p^2} + \frac{3p'^2}{8p^3} \right), \quad (72.7)$$

$$W_2 = -s \left[\frac{p'}{4p^2} + \frac{1}{8} \int^x \frac{p'(y)^2}{p(y)^3} dy \right]. \quad (72.8)$$

The integration uses $(p'/p^2)' = p''/p^2 - 2p'^2/p^3$. The exponent through this order is

$$\frac{iW}{\hbar} = \frac{\text{si}}{\hbar} \int^x p dy - \frac{1}{2} \log p - \text{si} \hbar \left[\frac{p'}{4p^2} + \frac{1}{8} \int^x \frac{p'^2}{p^3} dy \right] + \dots \quad (72.9)$$

72.3 Validity and turning points

The leading hierarchy requires the momentum to vary little over a reduced local wavelength $\hbar/|p|$:

$$\left| \frac{\hbar p'}{p^2} \right| \ll 1. \quad (72.10)$$

Since $p' = -mV'/p$, this becomes

$$\hbar^2(V')^2 \ll 8m|E - V|^3. \quad (72.11)$$

Higher derivatives must also be controlled. Direct substitution of (72.6) into (72.1) leaves a relative residual

$$\frac{\psi_s'' + (p^2/\hbar^2)\psi_s}{\psi_s} = \frac{3p'^2}{4p^2} - \frac{p''}{2p}.$$

For a slowly varying smooth potential, its magnitude is small relative to $p^2/\hbar^2$. This gives a useful direct check on an approximate solution.

In a classically forbidden region, define a positive real quantity

$$\kappa(x) = \sqrt{2m[V(x) - E]}.$$

The continuation $p = i\kappa$ turns oscillation into exponential behavior:

$$\psi(x) \simeq \frac{C_+}{\sqrt{\kappa(x)}} e^{\frac{1}{\hbar} \int^x \kappa dy} + \frac{C_-}{\sqrt{\kappa(x)}} e^{-\frac{1}{\hbar} \int^x \kappa dy}. \quad (72.12)$$

At a turning point x_0 where $V(x_0) = E$, both p and κ vanish. The WKB amplitude diverges and (72.11) fails. A separate local solution is needed to connect the two regions.

72.4 Counterpropagating waves and the phase around a well

Within a classically allowed well, a bound-state wavefunction contains both right-moving and left-moving components. Reflection at each turning point converts one into the other.

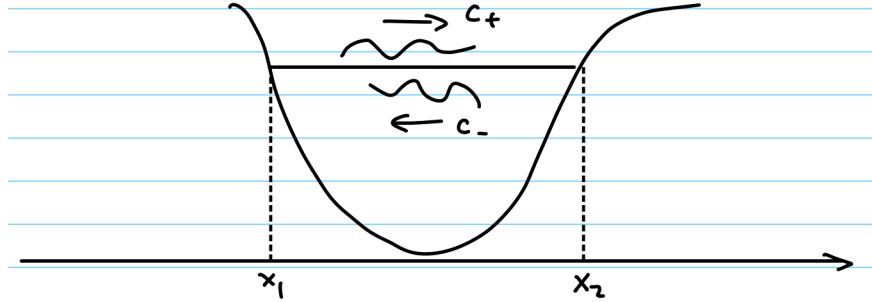


Figure 67: Counterpropagating semiclassical waves in a potential well. Reflection at the turning points closes the classical orbit.

Ignoring the reflection phases for a moment, consistency after a complete round trip would suggest

$$\exp\left(\frac{2i}{\hbar} \int_{x_1}^{x_2} p dx\right) = 1, \quad \int_{x_1}^{x_2} p dx = n\pi\hbar.$$

This is the motivation for action quantization, often written $\oint p dx = nh$ with $h = 2\pi\hbar$. A smooth turning point contributes an additional phase, however. The missing phases are responsible for the half-integer shift in the correct two-turning-point rule. They must be derived by matching to a nonsingular local solution.

73 Turning points, Airy functions, and bound-state quantization

73.1 Reduction to the Airy equation

Consider a simple right turning point x_0 with $V(x_0) = E$ and $F = V'(x_0) > 0$. Sufficiently close to it,

$$V(x) = E + F(x - x_0) + O((x - x_0)^2).$$

The linearized Schrödinger equation is

$$-\frac{\hbar^2}{2m}\psi'' + F(x - x_0)\psi = 0.$$

Introduce

$$z = \alpha(x - x_0), \quad \alpha = \left(\frac{2mF}{\hbar^2}\right)^{1/3}. \quad (73.1)$$

Then

$$\frac{d^2\psi}{dz^2} = z\psi, \quad \psi = A \text{Ai}(z) + B \text{Bi}(z). \quad (73.2)$$

The Airy functions are finite at $z = 0$. Their large- $|z|$ forms overlap with WKB in a region sufficiently far from the turning point for WKB to apply but sufficiently close that linearizing the potential remains accurate.

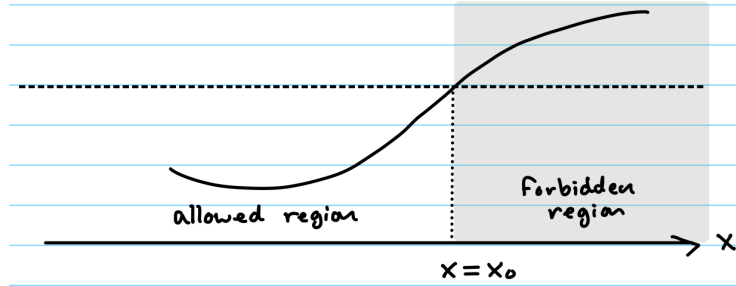


Figure 68: A simple turning point with $V'(x_0) > 0$: the classically allowed region is on the left and the forbidden region is on the right.

For $s, z > 0$ and large,

$$\text{Ai}(-s) \sim \frac{1}{\sqrt{\pi} s^{1/4}} \sin\left(\frac{2}{3}s^{3/2} + \frac{\pi}{4}\right), \quad (73.3)$$

$$\text{Bi}(-s) \sim \frac{1}{\sqrt{\pi} s^{1/4}} \cos\left(\frac{2}{3}s^{3/2} + \frac{\pi}{4}\right), \quad (73.4)$$

$$\text{Ai}(z) \sim \frac{1}{2\sqrt{\pi} z^{1/4}} \exp\left(-\frac{2}{3}z^{3/2}\right), \quad (73.5)$$

$$\text{Bi}(z) \sim \frac{1}{\sqrt{\pi} z^{1/4}} \exp\left(\frac{2}{3}z^{3/2}\right). \quad (73.6)$$

Both the phase $\pi/4$ and the factor of two in the decaying solution are needed in the connection formulas.

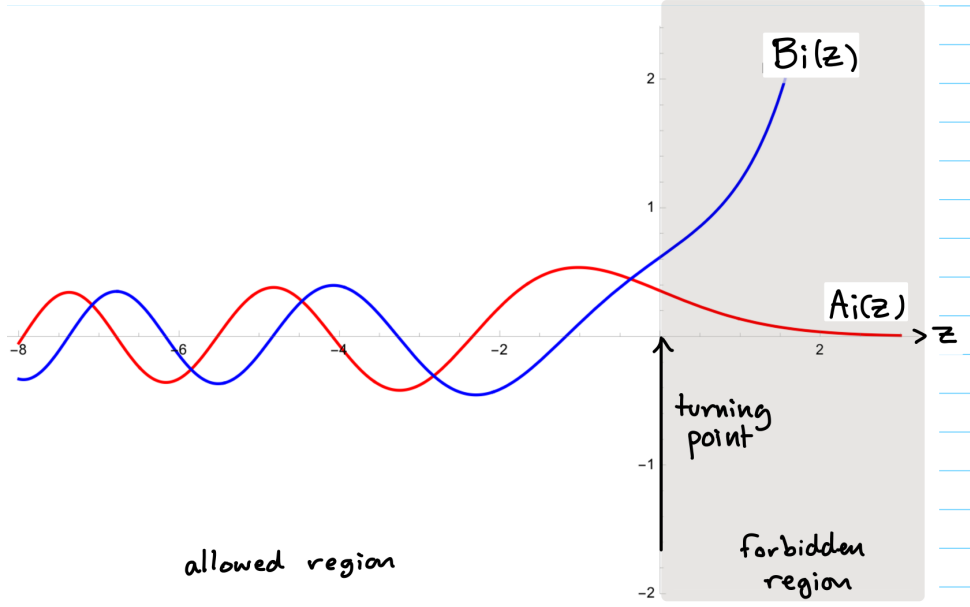


Figure 69: The Airy functions oscillate for negative argument. For positive argument, Ai decays and Bi grows.

73.2 Deriving the connection at a right turning point

Define positive action distances from the turning point:

$$\theta(x) = \frac{1}{\hbar} \int_x^{x_0} p(y) dy \quad (x < x_0), \quad Z(x) = \frac{1}{\hbar} \int_{x_0}^x \kappa(y) dy \quad (x > x_0).$$

For the linearized potential, these integrals are $\theta = \frac{2}{3}(-z)^{3/2}$ and $Z = \frac{2}{3}z^{3/2}$. Also

$$p = \hbar\alpha\sqrt{-z} \quad (z < 0), \quad \kappa = \hbar\alpha\sqrt{z} \quad (z > 0).$$

Thus $|z|^{-1/4}$ differs from $p^{-1/2}$ or $\kappa^{-1/2}$ by the same constant $(\hbar\alpha)^{1/2}$ on each side. Matching the Airy asymptotics term by term gives

$$\frac{C}{\sqrt{\kappa}} e^{-Z} \longleftrightarrow \frac{2C}{\sqrt{p}} \sin\left(\theta + \frac{\pi}{4}\right), \quad (73.7)$$

$$\frac{D}{\sqrt{\kappa}} e^Z \longleftrightarrow \frac{D}{\sqrt{p}} \cos\left(\theta + \frac{\pi}{4}\right). \quad (73.8)$$

Here the forbidden expression is on the right of the turning point and the oscillatory expression is on the left. The arrows specify the asymptotic forms of the same local solution; they are not propagation in time.

Exercise 73.1 (Matching all coefficients). Starting from $A \text{Ai}(z) + B \text{Bi}(z)$, recover the two connection formulas, including the relative factor two.

On the forbidden side, the coefficient of $e^{-Z}/\sqrt{\kappa}$ is $A(\hbar\alpha)^{1/2}/(2\sqrt{\pi})$, whereas the coefficient of $e^Z/\sqrt{\kappa}$ is $B(\hbar\alpha)^{1/2}/\sqrt{\pi}$. On the allowed side, the sine and cosine coefficients are respectively $A(\hbar\alpha)^{1/2}/\sqrt{\pi}$ and $B(\hbar\alpha)^{1/2}/\sqrt{\pi}$. Their ratios give (73.7) and (73.8) directly.

73.3 A left turning point and signed action conventions

For $V'(x_0) < 0$, the forbidden region is on the left and the allowed region is on the right. Reverse the local coordinate, or repeat the Airy reduction with $z = (2m|V'|/\hbar^2)^{1/3}(x_0 - x)$. With

$$Z_L(x) = \frac{1}{\hbar} \int_x^{x_0} \kappa \, dy, \quad \theta_R(x) = \frac{1}{\hbar} \int_{x_0}^x p \, dy,$$

both positive in their respective regions, the formulas become

$$\frac{C}{\sqrt{\kappa}} e^{-Z_L} \longleftrightarrow \frac{2C}{\sqrt{p}} \sin\left(\theta_R + \frac{\pi}{4}\right), \quad (73.9)$$

$$\frac{D}{\sqrt{\kappa}} e^{Z_L} \longleftrightarrow \frac{D}{\sqrt{p}} \cos\left(\theta_R + \frac{\pi}{4}\right). \quad (73.10)$$

The decaying exponential again has twice its coefficient in the allowed sine wave.

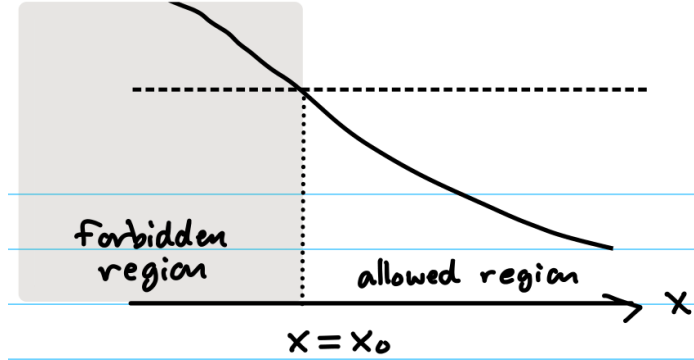


Figure 70: A left turning point, with $V'(x_0) < 0$. Its matching formulas follow by reflecting the right-turning-point problem.

For calculations with several turning points it is helpful to use actions whose derivatives are always positive:

$$I_i(x) = \frac{1}{\hbar} \int_{x_i}^x p(y) \, dy, \quad K_i(x) = \frac{1}{\hbar} \int_{x_i}^x \kappa(y) \, dy. \quad (73.11)$$

The first is defined in the relevant allowed interval and the second in the relevant forbidden interval. They can be negative to the left of their reference point, although $I'_i = p/\hbar > 0$ and $K'_i = \kappa/\hbar > 0$. Absorb a common factor $\sqrt{\hbar}$ into coefficients so that amplitudes are written $1/\sqrt{I'_i}$ and $1/\sqrt{K'_i}$.

At a turning point with positive slope, the connection formulas are then

$$\frac{\cos(I_i + \pi/4)}{\sqrt{I'_i}} \longleftrightarrow \frac{e^{-K_i}}{2\sqrt{K'_i}}, \quad (73.12)$$

$$\frac{\sin(I_i + \pi/4)}{\sqrt{I'_i}} \longleftrightarrow \frac{e^{K_i}}{\sqrt{K'_i}}. \quad (73.13)$$

Indeed, the positive allowed distance is $\theta = -I_i$, and $\sin(-I_i + \pi/4) = \cos(I_i + \pi/4)$. At a negative-slope turning point, the corresponding formulas are

$$\frac{\cos(I_i + \pi/4)}{\sqrt{I'_i}} \longleftrightarrow \frac{e^{-K_i}}{\sqrt{K'_i}}, \quad (73.14)$$

$$\frac{\sin(I_i + \pi/4)}{\sqrt{I'_i}} \longleftrightarrow \frac{e^{K_i}}{2\sqrt{K'_i}}. \quad (73.15)$$

Here $K_i < 0$ on the forbidden side, so e^{K_i} is the decaying solution. These signed forms will make scattering and resonance calculations compact.

73.4 Action differences and classical travel times

For two turning points bounding an allowed interval, define

$$I_{ij}(E) = \frac{1}{\hbar} \int_{x_i(E)}^{x_j(E)} p(x, E) dx, \quad x_i < x_j. \quad (73.16)$$

Within that interval, $I_i - I_j = I_{ij}$. Since $\partial_E p = m/p = 1/v$, differentiation gives

$$\hbar \frac{\partial I_{ij}}{\partial E} = \int_{x_i}^{x_j} \frac{m}{p(x, E)} dx = \int_{x_i}^{x_j} \frac{dx}{v(x, E)}. \quad (73.17)$$

The right-hand side is the classical time to travel from one turning point to the other. The endpoint terms in differentiating the original action vanish because $p(x_i) = p(x_j) = 0$. The derivative integrand behaves as $|x - x_i|^{-1/2}$ near a simple turning point and is integrable.

For a forbidden interval, define

$$K_{ij}(E) = \frac{1}{\hbar} \int_{x_i(E)}^{x_j(E)} \kappa(x, E) dx. \quad (73.18)$$

Then $K_i - K_j = K_{ij}$ and

$$\frac{\partial K_{ij}}{\partial E} = -\frac{1}{\hbar} \int_{x_i}^{x_j} \frac{m}{\kappa(x, E)} dx. \quad (73.19)$$

This is a real derivative of the forbidden action, not a real classical travel time through the forbidden region.

73.5 Bound states between two simple turning points

Suppose the potential has an allowed region $x_1 < x < x_2$, with forbidden regions extending outward on both sides. Normalizability selects a decaying exponential in each exterior region.

Write

$$\theta_1(x) = \frac{1}{\hbar} \int_{x_1}^x p dy, \quad \theta_2(x) = \frac{1}{\hbar} \int_x^{x_2} p dy, \quad \theta_1 + \theta_2 = I_{12}.$$

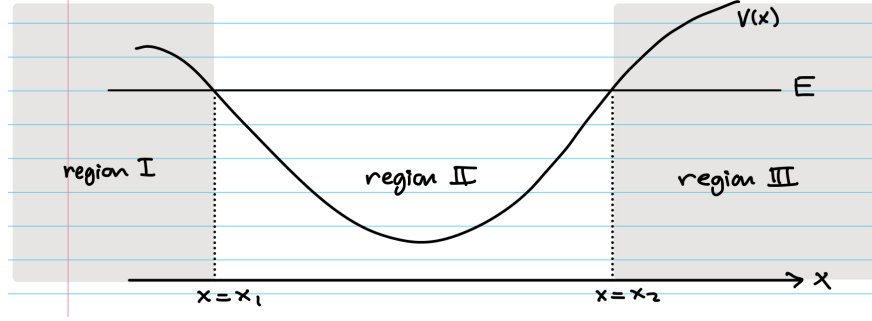


Figure 71: A bound-state well with two simple turning points. The exterior solutions decay in regions I and III, while region II is oscillatory.

Matching a decaying exterior solution at the left endpoint gives

$$\psi_{\text{in}}(x) = \frac{2C_L}{\sqrt{p}} \sin\left(\theta_1 + \frac{\pi}{4}\right).$$

Matching the right decay gives, in the same interval,

$$\psi_{\text{in}}(x) = \frac{2C_R}{\sqrt{p}} \sin\left(\theta_2 + \frac{\pi}{4}\right).$$

Set $A = \theta_1 + \pi/4$. The second sine is $\sin(I_{12} + \pi/2 - A)$. Expanding,

$$\sin(I_{12} + \pi/2 - A) = \sin(I_{12} + \pi/2) \cos A - \cos(I_{12} + \pi/2) \sin A.$$

For this to be proportional to $\sin A$ throughout the interval, its $\cos A$ coefficient must vanish. Therefore

$$\boxed{\int_{x_1}^{x_2} \sqrt{2m[E_n - V(x)]} dx = \pi\hbar \left(n + \frac{1}{2}\right), \quad n = 0, 1, 2, \dots} \quad (73.20)$$

The relative coefficients satisfy $C_R = (-1)^n C_L$. Equivalently, $\oint p dx = 2\pi\hbar(n + \frac{1}{2})$. The half-integer shift is the combined effect of the two smooth turning points. Hard walls and singular endpoints require their own boundary phases.

73.6 The harmonic oscillator

For $V(x) = m\omega^2 x^2/2$, the turning points are $\pm x_0$ with $x_0^2 = 2E/(m\omega^2)$, and

$$p(x) = m\omega \sqrt{x_0^2 - x^2}.$$

The action is the area of a semicircle multiplied by $m\omega$:

$$\int_{-x_0}^{x_0} p dx = m\omega \frac{\pi x_0^2}{2} = \frac{\pi E}{\omega}.$$

The phase rule without turning-point corrections would give $E_n = n\hbar\omega$. The correctly matched rule gives

$$E_n = \hbar\omega \left(n + \frac{1}{2} \right), \quad (73.21)$$

which is the exact harmonic-oscillator spectrum. The WKB wavefunctions still need local turning-point treatment; exact energies do not make the leading WKB expression uniformly exact.

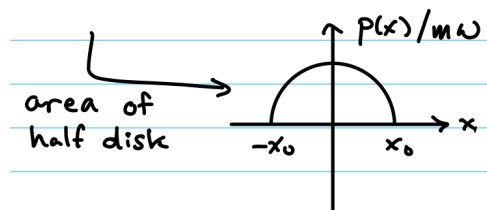


Figure 72: The harmonic-oscillator action integral is proportional to the area of a semicircle in the (x, p) plane.

73.7 The pure quartic oscillator

Return to $V(x) = \lambda x^4$. Let $x_0 = (E/\lambda)^{1/4}$. Then

$$p(x) = \sqrt{2m\lambda} \sqrt{x_0^4 - x^4}.$$

With $u = x/x_0$,

$$\int_{-x_0}^{x_0} p \, dx = \sqrt{2m\lambda} x_0^3 \mathcal{K}, \quad (73.22)$$

$$\mathcal{K} = \int_{-1}^1 \sqrt{1 - u^4} \, du = \frac{1}{2} B\left(\frac{1}{4}, \frac{3}{2}\right) \simeq 1.748. \quad (73.23)$$

Solving the quantization condition for E gives

$$E_n^{\text{WKB}} = \left[\frac{\pi(n + \frac{1}{2})}{\sqrt{2} \mathcal{K}} \right]^{4/3} \frac{\hbar^{4/3} \lambda^{1/3}}{m^{2/3}}. \quad (73.24)$$

For the first two states, the dimensionless coefficients are approximately 0.546 and 2.36, compared with numerical values 0.668 and 2.394. The lowest state is poorly semiclassical, whereas the first excited state is already much closer. The approximation improves for highly excited states, whose action is large compared with $\hbar$.

The comparison with the variational method is instructive:

Energy coefficient	Ground state	First excited state
Gaussian variational bound	0.6814	2.424
Leading WKB	0.546	2.36
Numerical value	0.668	2.394

All entries multiply $\hbar^{4/3} \lambda^{1/3} / m^{2/3}$. The variational estimates are upper bounds; the WKB values are asymptotic approximations without that ordering guarantee.

74 Tunneling through a smooth potential barrier

74.1 Scattering boundary conditions and phase conventions

Let a localized barrier have two simple turning points $x_1 < x_2$ at energy $E > 0$. Regions I and III are allowed, while $x_1 < x < x_2$ is forbidden. Assume $V(x) \rightarrow 0$ sufficiently rapidly as $x \rightarrow \pm\infty$, so the same asymptotic wave number $k = \sqrt{2mE}/\hbar$ applies on both sides. Unit-amplitude incidence from the left means

$$\psi(x) \sim \begin{cases} e^{ikx} + r(E)e^{-ikx}, & x \rightarrow -\infty, \\ t(E)e^{ikx}, & x \rightarrow +\infty. \end{cases} \quad (74.1)$$

Since the incident and transmitted velocities are equal, the transmission probability is $T = |t|^2$ and the reflection probability is $R = |r|^2$.

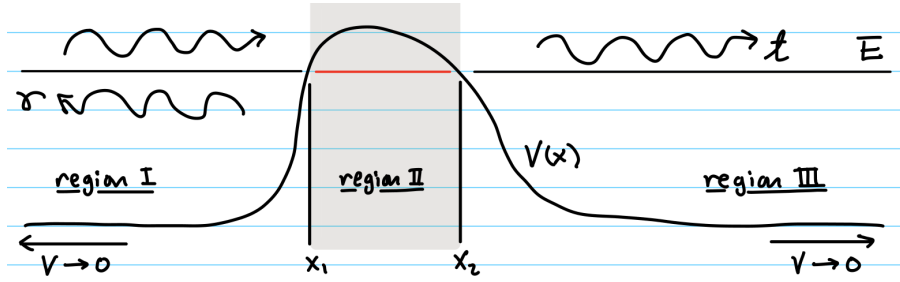


Figure 73: A smooth tunneling barrier with allowed exterior regions I and III and a forbidden interval II between x_1 and x_2 .

To retain the asymptotic scattering phases, define

$$I_{-\infty}(x) = kx + \frac{1}{\hbar} \int_{-\infty}^x [p(y) - \hbar k] dy \quad \text{in region I,} \quad (74.2)$$

$$I_{+\infty}(x) = kx + \frac{1}{\hbar} \int_{+\infty}^x [p(y) - \hbar k] dy \quad \text{in region III.} \quad (74.3)$$

Their derivatives are $p/\hbar$, and each approaches kx at its indicated infinity. The corresponding WKB states are

$$\psi_I = \frac{\sqrt{k}}{\sqrt{I'}} (e^{iI_{-\infty}} + re^{-iI_{-\infty}}), \quad (74.4)$$

$$\psi_{III} = \frac{\sqrt{k}}{\sqrt{I'}} te^{iI_{+\infty}}. \quad (74.5)$$

The factor $\sqrt{k}$ ensures the plane-wave amplitudes in (74.1).

Let I_1 and I_2 be the signed turning-point actions from (73.11). The constant differences are

$$A(E) = I_{-\infty} - I_1 = kx_1 + \frac{1}{\hbar} \int_{-\infty}^{x_1} (p - \hbar k) dx, \quad (74.6)$$

$$B(E) = I_2 - I_{+\infty} = -kx_2 + \frac{1}{\hbar} \int_{x_2}^{\infty} (p - \hbar k) dx. \quad (74.7)$$

The forbidden action is

$$K = K_{12}(E) = \frac{1}{\hbar} \int_{x_1}^{x_2} \kappa(x, E) dx. \quad (74.8)$$

74.2 Matching from the transmitted wave back to the incident wave

Start on the right, where there is only an outgoing wave. With $C = t\sqrt{k} e^{-iB-i\pi/4}$, it can be written

$$\psi_{\text{III}} = \frac{C}{\sqrt{I'}} e^{i(I_2+\pi/4)} = \frac{C}{\sqrt{I'}} [\cos(I_2 + \pi/4) + i \sin(I_2 + \pi/4)].$$

The slope at x_2 is negative. Using (73.15), the forbidden solution is

$$\psi_{\text{II}} = \frac{C}{\sqrt{K'}} \left(e^{-K_2} + \frac{i}{2} e^{K_2} \right). \quad (74.9)$$

Because $K_2 = K_1 - K$, this is also

$$\psi_{\text{II}} = \frac{C}{\sqrt{K'}} \left(e^K e^{-K_1} + \frac{i}{2} e^{-K} e^{K_1} \right).$$

At x_1 the slope is positive. Matching again yields

$$\psi_{\text{I}} = \frac{C}{\sqrt{I'}} \left[2e^K \cos(I_1 + \pi/4) + \frac{i}{2} e^{-K} \sin(I_1 + \pi/4) \right]. \quad (74.10)$$

To identify incoming and reflected amplitudes, put $\theta = I_1 + \pi/4$:

$$2e^K \cos \theta + \frac{i}{2} e^{-K} \sin \theta = \left(e^K + \frac{1}{4} e^{-K} \right) e^{i\theta} + \left(e^K - \frac{1}{4} e^{-K} \right) e^{-i\theta}.$$

The incident coefficient must agree with $\sqrt{k} e^{iA} e^{iI_1}$ in (74.5). Comparing both traveling components gives

$$t \simeq \frac{e^{i(A+B)}}{e^K + \frac{1}{4} e^{-K}}, \quad (74.11)$$

$$r \simeq -i e^{2iA} \frac{e^K - \frac{1}{4} e^{-K}}{e^K + \frac{1}{4} e^{-K}}. \quad (74.12)$$

These are the amplitudes obtained from the leading Airy connection formulas while retaining both exponential components. Their reflected and transmitted fluxes sum to one because

$$\left(e^K + \frac{1}{4} e^{-K} \right)^2 - \left(e^K - \frac{1}{4} e^{-K} \right)^2 = 1.$$

The connection approximation is not uniform when the two turning points approach each other near the top of the barrier.

For a thick barrier, $K \gg 1$, the terms suppressed by e^{-2K} can be discarded in these amplitudes:

$$t \simeq e^{i(A+B)} e^{-K}, \quad r \simeq -i e^{2iA}. \quad (74.13)$$

In particular,

$$T \simeq e^{-2K} = \exp \left[-\frac{2}{\hbar} \int_{x_1}^{x_2} \sqrt{2m[V(x) - E]} dx \right]. \quad (74.14)$$

The reflected wave has nearly unit magnitude. The exponentially small transmitted amplitude is determined by the forbidden action.

74.3 The transmission phase and wave-packet delay

For a narrow incident packet, the energy derivative of the transmission phase controls the displacement of the transmitted peak. To see this, write the transmitted packet far to the right as

$$\psi_{\text{tr}}(x, t) = \int a(k) |t(k)| \exp\{i[kx - \omega(k)t + \phi_t(k)]\} dk, \quad \omega(k) = \frac{\hbar k^2}{2m}.$$

Stationary phase near its central wave number gives $x - v_0 t + \partial_k \phi_t = 0$, where $v_0 = \hbar k/m$. Relative to free propagation, the phase delay is therefore

$$\tau_{\text{ph}} = \frac{1}{v_0} \frac{\partial \phi_t}{\partial k} = \hbar \frac{\partial \phi_t}{\partial E} = \frac{m}{\hbar k} \frac{\partial \phi_t}{\partial k}. \quad (74.15)$$

At leading thick-barrier order, $\phi_t = A + B$; the forbidden action changes the magnitude of t , not its real phase.

Exercise 74.1 (Differentiating the exterior phase). Differentiate $A(E)$ and $B(E)$, including the energy dependence of the turning points, and express the answer in terms of classical velocities.

Since $\hbar \partial_E k = 1/v_0$ and $\partial_E p = 1/v$, differentiating A gives

$$\begin{aligned} \hbar A' &= \frac{x_1}{v_0} + \hbar k x_1' + \int_{-\infty}^{x_1} \left(\frac{1}{v} - \frac{1}{v_0} \right) dx + [p(x_1) - \hbar k] x_1' \\ &= \frac{x_1}{v_0} + \int_{-\infty}^{x_1} \left(\frac{1}{v} - \frac{1}{v_0} \right) dx. \end{aligned}$$

The endpoint terms cancel because $p(x_1) = 0$. Similarly,

$$\hbar B' = -\frac{x_2}{v_0} + \int_{x_2}^{\infty} \left(\frac{1}{v} - \frac{1}{v_0} \right) dx.$$

Thus

$$\tau_{\text{ph}} = \int_{-\infty}^{x_1} \left(\frac{1}{v} - \frac{1}{v_0} \right) dx + \int_{x_2}^{\infty} \left(\frac{1}{v} - \frac{1}{v_0} \right) dx - \frac{x_2 - x_1}{v_0}. \quad (74.16)$$

The first two terms are the classical excess travel times in the allowed exterior regions. The last subtracts the free travel time across the width of the forbidden interval.

This result says that the leading WKB forbidden-region action contributes attenuation without a real propagation phase. It does not define a literal zero time spent inside the barrier. A peak delay is a property of an outgoing packet relative to a reference packet; energy filtering and reshaping can change that peak. Other questions about tunneling duration require an explicit operational definition and need not give the same quantity. Furthermore, the Schrödinger equation used here is nonrelativistic. Its phase-delay result cannot by itself establish a relativistic signal or front velocity.

75 Euclidean trajectories and the tunneling action

75.1 The inverted-potential problem

The forbidden action has a useful classical interpretation after analytic continuation to imaginary time. At fixed energy E , consider the Euclidean action

$$S_E[x] = \int \left[\frac{m}{2} \left(\frac{dx}{d\tau} \right)^2 + V(x) - E \right] d\tau. \quad (75.1)$$

Here τ is Euclidean time. The Euler–Lagrange equation is

$$m \frac{d^2 x}{d\tau^2} = V'(x). \quad (75.2)$$

It is the Newton equation for the inverted potential $U_E(x) = E - V(x)$. A region forbidden at the original real energy becomes accessible to a Euclidean trajectory.

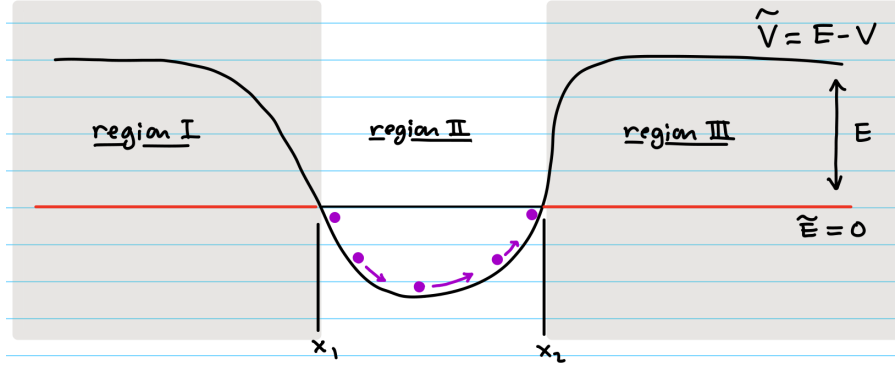


Figure 74: The inverted-potential description of tunneling. The Euclidean trajectory connects the turning points through the region forbidden to the original real-time motion.

Because the Euclidean Lagrangian has no explicit τ dependence,

$$\mathcal{E}_E = \frac{m}{2} \left(\frac{dx}{d\tau} \right)^2 - V(x) + E \quad (75.3)$$

is conserved. A trajectory starting at a turning point with zero Euclidean velocity has $\mathcal{E}_E = 0$, and therefore

$$\frac{m}{2} \left(\frac{dx}{d\tau} \right)^2 = V(x) - E. \quad (75.4)$$

For motion from x_1 to x_2 , choose the positive square root,

$$\frac{dx}{d\tau} = \sqrt{\frac{2[V(x) - E]}{m}} = \frac{\kappa(x)}{m}.$$

The turning points of the original and inverted problems coincide.

75.2 Instanton and bounce actions

Evaluate (75.1) on this trajectory. The two terms in its integrand are equal by (75.4), so

$$S_{\text{inst}} = \int m \left(\frac{dx}{d\tau} \right)^2 d\tau \quad (75.5)$$

$$= \int_{x_1}^{x_2} m \frac{dx}{d\tau} dx = \int_{x_1}^{x_2} \sqrt{2m[V(x) - E]} dx = \hbar K_{12}. \quad (75.6)$$

Consequently,

$$|t| \simeq e^{-S_{\text{inst}}/\hbar}, \quad T \simeq e^{-2S_{\text{inst}}/\hbar}. \quad (75.7)$$

A Euclidean trajectory that goes across the forbidden region and returns is a *bounce*. Its action is $S_{\text{bounce}} = 2S_{\text{inst}}$, so the same probability suppression can be written $e^{-S_{\text{bounce}}/\hbar}$.

This is the semiclassical origin of the connection between tunneling and Euclidean path integrals: stationary Euclidean trajectories determine exponential suppressions. The same construction extends to field configurations, where bounce solutions describe processes such as decay of a metastable vacuum. The Euclidean trajectory determines a tunneling amplitude; its imaginary-time parametrization is not an ordinary real-time history of a particle moving through the barrier.

76 Metastable states and resonance lifetimes

76.1 A symmetric well between two barriers

Consider an even potential $V(-x) = V(x)$ with a central well separated from the asymptotic exterior by two barriers. At the energy of interest there are four turning points

$$x_1 < x_2 < 0 < x_3 < x_4, \quad x_1 = -x_4, \quad x_2 = -x_3.$$

The central interval (x_2, x_3) and both exterior regions are allowed; the intervals (x_1, x_2) and (x_3, x_4) are forbidden.

If the barriers were impenetrable, the central well would have ordinary bound states. Finite transmission lets probability escape, producing long-lived resonances

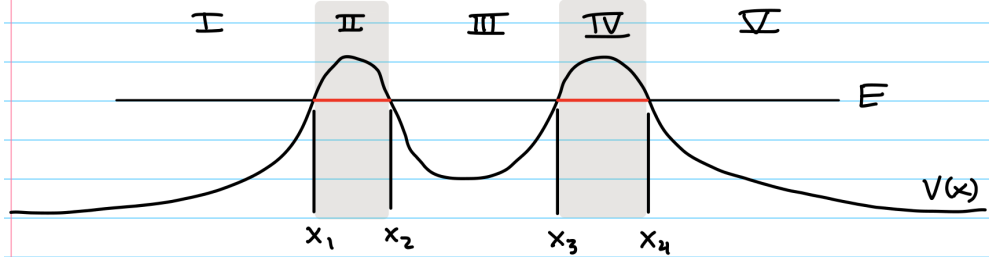


Figure 75: A symmetric central well bounded by two tunneling barriers. The five regions alternate between allowed and forbidden motion.

when the barriers are sufficiently opaque. Their real energies approximately obey the bound-state quantization rule for the inner turning points, while their small imaginary parts determine the escape rate.

For an even potential, solutions may be chosen to have definite parity. It is sufficient to solve on $x \geq 0$ with

$$\psi'(0) = 0 \quad \text{for even parity}, \quad \psi(0) = 0 \quad \text{for odd parity.} \quad (76.1)$$

The resonance boundary condition is purely outgoing:

$$\psi(x) \propto e^{ikx} \quad (x \rightarrow +\infty), \quad \psi(x) \propto e^{-ikx} \quad (x \rightarrow -\infty).$$

A decaying resonance has $k = k_0 - ik_1$ with $k_0, k_1 > 0$ and an energy below the real axis. Its outgoing spatial solution grows at infinity and is not a normalizable stationary eigenvector of the self-adjoint Hamiltonian. It represents the pole governing an approximately exponential decay of a physical state localized in the well.

76.2 Matching the outgoing solution through the right barrier

Use the signed actions I_i, K_i of (73.11). An overall normalization is irrelevant, so choose the outgoing solution in region V as

$$\psi_V = \frac{1}{\sqrt{I'}} e^{i(I_4 + \pi/4)}. \quad (76.2)$$

At x_4 the slope is negative. Matching its cosine and sine pieces gives

$$\psi_{IV} = \frac{1}{\sqrt{K'}} \left(e^{-K_4} + \frac{i}{2} e^{K_4} \right). \quad (76.3)$$

Let

$$K_{34}(E) = \frac{1}{\hbar} \int_{x_3}^{x_4} \kappa(x, E) dx. \quad (76.4)$$

Since $K_4 = K_3 - K_{34}$, the same function near the inner edge of the barrier is

$$\psi_{IV} = \frac{1}{\sqrt{K'}} \left[e^{K_{34}} e^{-K_3} + \frac{i}{2} e^{-K_{34}} e^{K_3} \right].$$

At x_3 the slope is positive. The positive-slope connection formulas give the central-well solution

$$\psi_{\text{III}} = \frac{1}{\sqrt{I'}} \left[2e^{K_{34}} \cos(I_3 + \pi/4) + \frac{i}{2} e^{-K_{34}} \sin(I_3 + \pi/4) \right]. \quad (76.5)$$

The second term is exponentially small compared with the first, but it must be retained: it is the only source of the imaginary energy at this order. Discarding it would recover perfectly reflecting barriers and an infinite lifetime.

76.3 Even and odd resonance conditions

Define the action measured from the centre,

$$I_0(x) = \frac{1}{\hbar} \int_0^x p(y, E) dy, \quad I_{03}(E) = \frac{1}{\hbar} \int_0^{x_3} p(y, E) dy.$$

Then $I_3 = I_0 - I_{03}$. Since p is even, I_0 is odd and $I_0''(0) = p'(0)/\hbar = 0$. In particular, differentiating the WKB amplitude $1/\sqrt{I'}$ makes no contribution at the centre.

Put

$$\alpha = I_{03} - \frac{\pi}{4}, \quad K = K_{34}.$$

At $x = 0$, the angle $I_3 + \pi/4$ is $-\alpha$. Apart from nonzero common factors, the value and derivative of (76.5) are

$$\psi(0) \propto 2e^K \cos \alpha - \frac{i}{2} e^{-K} \sin \alpha, \quad (76.6)$$

$$\psi'(0) \propto 2e^K \sin \alpha + \frac{i}{2} e^{-K} \cos \alpha. \quad (76.7)$$

Therefore the even condition is

$$\tan \left(I_{03} - \frac{\pi}{4} \right) = -\frac{i}{4} e^{-2K_{34}}, \quad (76.8)$$

and the odd condition is

$$\cot \left(I_{03} - \frac{\pi}{4} \right) = \frac{i}{4} e^{-2K_{34}}. \quad (76.9)$$

In the opaque-barrier limit, the right-hand sides vanish. For even states, $I_{03} - \pi/4 = \ell\pi$, so

$$I_{23} = 2I_{03} = (2\ell + \frac{1}{2})\pi.$$

For odd states, $I_{03} - \pi/4 = \ell\pi + \pi/2$, so

$$I_{23} = (2\ell + \frac{3}{2})\pi.$$

Combining both parities gives

$$I_{23}(E_n) = \left(n + \frac{1}{2} \right) \pi, \quad n = 0, 1, 2, \dots \quad (76.10)$$

Even n corresponds to even parity, and odd n to odd parity. These are the usual real WKB levels of the central well.

76.4 Expanding around a real level

For small leakage, write the angle as its real root plus a small complex displacement. Near an even root $\tan \alpha \simeq \alpha - \ell\pi$. Near an odd root $\cot \alpha \simeq -[\alpha - (\ell + \frac{1}{2})\pi]$. Both parity equations therefore reduce to

$$I_{03}(E) - \left(\frac{n}{2} + \frac{1}{4}\right)\pi \simeq -\frac{i}{4}e^{-2K_{34}(E)}. \quad (76.11)$$

Equivalently,

$$I_{23}(E) - \left(n + \frac{1}{2}\right)\pi \simeq -\frac{i}{2}e^{-2K_{34}(E)}. \quad (76.12)$$

Parameterize the complex energy by

$$E = E_R - \frac{i\Gamma}{2}, \quad \Gamma > 0. \quad (76.13)$$

Here Γ is the energy width. Expanding the left-hand side of (76.11),

$$I_{03}(E) = I_{03}(E_R) - \frac{i\Gamma}{2}I'_{03}(E_R) + O(\Gamma^2).$$

The forbidden action similarly satisfies

$$K_{34}(E) = K_{34}(E_R) - \frac{i\Gamma}{2}K'_{34}(E_R) + O(\Gamma^2),$$

and thus

$$e^{-2K_{34}(E)} = e^{-2K_{34}(E_R)} [1 + i\Gamma K'_{34}(E_R) + O(\Gamma^2)].$$

The correction on the right is of higher exponential order because Γ itself is proportional to the barrier transmission. To leading order it may be omitted.

The real part of the resonance equation gives $I_{23}(E_R) = (n + \frac{1}{2})\pi$, so $E_R \simeq E_n$. The imaginary part gives

$$\frac{\Gamma}{2}I'_{03}(E_n) = \frac{1}{4}e^{-2K_{34}(E_n)}.$$

Using $I'_{23} = 2I'_{03}$, the width is

$$\boxed{\Gamma_n \simeq \frac{e^{-2K_{34}(E_n)}}{I'_{23}(E_n)}}. \quad (76.14)$$

Since $I'_{23} > 0$, the width is positive and the imaginary part of the energy is negative, as required for decay.

76.5 Lifetime and the classical escape picture

The time-dependent resonance factor is

$$e^{-iEt/\hbar} = e^{-iE_R t/\hbar} e^{-\Gamma t/(2\hbar)}.$$

The amplitude decreases with exponent $\Gamma/(2\hbar)$; the probability decreases twice as rapidly:

$$P(t) \propto e^{-\Gamma t/\hbar} = e^{-t/\tau}, \quad \tau = \frac{\hbar}{\Gamma}. \quad (76.15)$$

Combining this with (76.14),

$$\boxed{\tau_n \simeq \hbar I'_{23}(E_n) e^{2K_{34}(E_n)}}. \quad (76.16)$$

By the action-time identity,

$$t_c = \hbar I'_{23}(E_n) = \int_{x_2}^{x_3} \frac{dx}{v(x, E_n)}$$

is the one-way classical travel time across the central well. The full oscillation period is $2t_c$. Each barrier has transmission probability $T \simeq e^{-2K_{34}}$, so

$$\tau \simeq \frac{t_c}{T} = \frac{\text{full classical period}}{2T}. \quad (76.17)$$

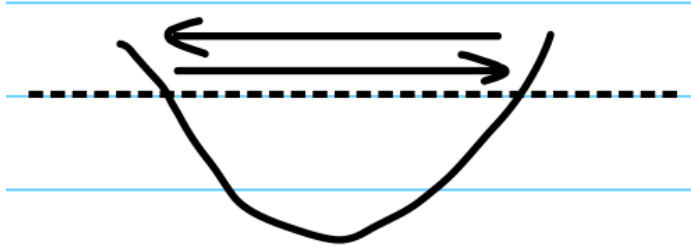


Figure 76: The escape interpretation: a particle traverses the central well between successive encounters with its two weakly transmitting barriers.

This result can be recovered by counting barrier encounters. After N encounters, the survival probability is approximately

$$P_N = (1 - T)^N.$$

For $T \ll 1$,

$$\log P_N = N \log(1 - T) \simeq -NT.$$

There is one encounter after each time interval t_c , since the particle alternates between the two walls. Thus $N \simeq t/t_c$ and

$$P(t) \simeq \exp\left(-\frac{Tt}{t_c}\right),$$

which gives (76.17). Counting only one escape attempt per full period would miss the second barrier and produce an incorrect factor two.

For a large well, this picture can be made especially direct by following a wave packet between the barriers: a small piece escapes at every encounter. The stationary resonance calculation does not require a sharply localized packet at every stage; its controlled ingredients are the semiclassical connection formulas, weak leakage, and the separation of the lifetime from the much shorter oscillation time.

Source-to-topic index

The fourteen source PDFs are incorporated as follows. When printed and handwritten lectures cover the same argument, it is developed once; additional examples and derivations are included in the relevant section.

Source	Location in these notes
<i>1: Motivation</i> (19 pages)	Motivation and experimental foundations
<i>2: Hilbert Spaces</i> (10 pages)	Finite-dimensional Hilbert spaces
<i>3: Postulates</i> (13 pages)	Quantum states, measurements, and evolution
<i>4: Time Evolution and Symmetries</i> (31 pages)	Time evolution and symmetries
<i>5: Quantum Information</i> (10 pages)	Quantum information and entanglement
<i>5B: Quantum Information</i> , handwritten (37 pages)	Quantum information and entanglement
<i>6: Continuous Observables</i> (12 pages)	Continuous observables
<i>6B: Continuous Observables</i> , handwritten (22 pages)	Continuous observables
<i>7: Wave Mechanics</i> (29 pages)	Wave mechanics and bound states
<i>7C: Bound States</i> (12 pages)	Wave mechanics and bound states
<i>8: Scattering</i> (19 pages)	Scattering and resonances
<i>8B: Note on Resonances</i> (2 pages)	Scattering and resonances
<i>9: Approximation Methods</i> (52 pages)	Approximation methods
<i>9B: Degenerate Perturbation Theory</i> (4 pages)	Approximation methods