

Part II Advanced Quantum Mechanics

William Royce

November 4, 2024

Part II Physics, The University of Cambridge

Preface

“Quantum mechanics is very impressive. But an inner voice tells me that it is not yet the real thing. The theory produces a good deal but hardly brings us closer to the secret of the Old One.”

Without wishing to overstate the case, the discovery of quantum mechanics is the single greatest achievement in the history of human civilisation.

Quantum mechanics is an outrageous departure from our classical, comforting, common sense view of the world. It is more baffling and disturbing than anything dreamt up by science fiction writers. And yet it is undoubtably the correct description of the universe we inhabit, providing insights into many aspects of the world around us.

Among the many successes of quantum mechanics are answers to old and very basic questions. Like why is matter stable? And why does the sun shine and why is it yellow? And why are all solids either conductors or insulators? But quantum mechanics also opens up vistas that we didn't previously know existed, from novel states of matter where the constituent particles become so entangled that they can be coaxed to perform seemingly impossible tasks, to the subatomic world of fluctuating ethereal fields, to a revised understanding of the meaning of information and what one can achieve with it.

Although quantum mechanics gives the right answers to all questions, it does so at a price. The answers are always statistical in nature. There are few certainties in the quantum world. Of course, there are few certainties in the classical world too, but we can always strive to eliminate them. In classical physics, knowledge is power: the more you know, the better off and less uncertain you are. This is because classical probabilities are always about our ignorance and any classical system that appears random does so because it is somehow difficult, but never impossible, for us to know what's going on inside.

This is not the way of the quantum world. The randomness is real, the unpredictability inherent. There is no way to game the system or gain some illicit knowledge of underlying quantum properties that would reduce the uncertainty. If, for example, a particle has probability $1/2$ to be in one place and probability $1/2$ to be elsewhere, this may well be because the particle really is in both places at the same time. Any attempt to eliminate the quantum certainty will simply shift it, like a bubble in wallpaper, elsewhere. As we will see, quantum particles are fragile objects and the very act of looking changes them, disturbing many of the other delicate properties that they possess.

The purpose of these lectures is to begin to get to grips with the quantum world and understand some of the strange behaviour that it exhibits. As should be clear already, this is something of a step into the unknown and our classical intuition will not be a great guide in the quantum realm. Fortunately we have a better guide because, for reasons that we do not fully understand, the quantum world is described astonishingly well by the language of mathematics. Our strategy in these lectures, and later ones, will be to embrace this mathematical description and leverage it to build a new intuition for how the universe really works.

In the following chapters we'll additionally cover what Hilbert space is and what it's operators do in a more general framework, building insight into the mathematical structure of quantum mechanics. Much of this is just linear algebra, but the Hilbert spaces we'll care about in QM are often infinite-dimensional.

So, mathematically, much of Quantum Mechanics boils down to a mix of Functional Analysis and Representation Theory. It's even true that it provides a particularly interesting example of these subjects. *But this is not the reason we study it.* We study Quantum Mechanics in an effort to understand and appreciate our World, not some abstract mathematical one. You're all intimately familiar with vector spaces, and you're (hopefully!) also very good at solving Sturm–Liouville type eigenfunction / eigenvalue problems. But the real skill is in understanding how this formalism relates to the world we see around us.

It's not obvious. Newton's laws are (at least generically) non-linear differential equations and we can't usually superpose solutions. General Relativity teaches us that space-time is not flat. So it's not at all clear that our particle should in fact be described by a point in a vector space, any more than it was obvious to Aristotle that bodies actually stay in uniform motion unless acted on by a force, or clear to the Ancients that the arrival of solar eclipses, changes of the weather, or any other natural phenomenon are actually governed by calculable Laws, rather than the whims of various gods. For this reason, instead of emphasising how weird and different Quantum Mechanics is, I'd prefer to make you appreciate how it actually underpins the physics you're already familiar with. No matter how good you are at solving eigenvalue problems, if you don't see how these relate to your everyday physical intuition, knowledge you've built up since first opening your eyes and learning to crawl, then you haven't really understood the subject.

Contents

1 Wave Mechanics and the Schrödinger Equation	1
1.1 Historical Foundations of Quantum Physics	1
1.1.1 Black-Body Radiation	1
1.1.2 Photoelectric Effect	2
1.1.3 Compton Scattering	4
1.1.4 Atomic Spectra	4
1.2 Wave Mechanics	5
1.2.1 Maxwell's Wave Equation	6
1.2.2 Schrödinger's Equation	7
1.3 Postulates of Quantum Theory	8
1.3.1 Postulate 1	8
1.3.2 Postulate 2	9
1.3.3 Postulate 3	9
1.3.4 Postulate 4	10
2 Quantum Mechanics in One Dimension	11
2.1 Wave Mechanics of Unbound Particles	12
2.1.1 Particle Flux and conservation of probability	12
2.1.2 Free Particle	12
2.1.3 Potential Step	14
2.1.4 Potential Barrier	15
2.1.5 The Rectangular Potential Well	16
2.2 Wave Mechanics of Bound Particles	17
2.2.1 The Rectangular Potential Well (Continued)	17
2.2.2 The δ -Function Potential Well	18
2.2.3 The δ -Function Model of a Crystal	18
2.3 Wentzel, Kramers and Brillouin (WKB) Method	20
2.3.1 Semi-Classical Approximation to Leading Order	20
2.3.2 Next to Leading Order Correction	22
2.3.3 Connection Formulae, Boundary Conditions and Quantisation Rules	23
2.3.4 Example: Simple Harmonic Oscillator	24
2.3.5 Example: Quantum Tunneling	25
3 Operator Methods and Formalism in Quantum Mechanics	27
3.1 States	28
3.1.1 The Inner Product	28
3.1.2 Dual Space	30
3.1.3 Dirac Notation and Continuum States	30
3.2 Operators	32
3.2.1 Matrix Representations of Operators	34
3.2.2 Eigenfunctions and Eigenvalues	36
3.2.3 Hermitian Operators	36
3.2.4 Measurement	40
3.2.5 Time-Evolution Operator	43
3.2.6 Commutation Relations	43
3.3 The Generalised Uncertainty Principle	46
3.4 Time-Evolution of Expectation Values	48

3.5	The Heisenberg Picture	49
3.6	Quantum Harmonic Oscillator	49
4	Quantum Mechanics in More Than One Dimension	53
4.1	Rigid Diatomic Molecule	53
4.2	Angular momentum	54
4.2.1	Commutation Relations	54
4.2.2	Eigenvalues of Angular Momentum	55
4.2.3	Representation of the Angular Momentum States	58
5	Spin	61
5.1	Composite Systems	62
5.1.1	Tensor Product of Hilbert Spaces	62
5.1.2	Operators in Composite Systems	64
5.2	Spinors, Spin Operators, Pauli Matrices	65
5.3	Relating the Spinor to the Spin Direction	66
5.4	Spin Precession in a Magnetic Field	67
5.4.1	Energy Eigenstates	69
5.4.2	Wavefunction Evolution	70
5.5	Addition of Angular Momenta	71
5.5.1	Addition of Two Spin 1/2 Degrees of Freedom	74
5.5.2	Addition of Angular Momentum and Spin	75
5.5.3	Addition of two angular momenta $J = 1$	75
6	Motion in a Magnetic Field	77
6.1	Classical Mechanics of a Particle in a Field	77
6.2	Quantum Mechanics of a Particle in a Field	79
6.3	Gauge Invariance and the Aharonov-Bohm Effect	80
6.4	Free Electron in a Magnetic Field	83
6.5	The Magnetic Moment of the Muon	86
6.5.1	The Measurement	87
6.6	Magnetic Moments	89
6.7	The Stern-Gerlach Experiment	90
7	Approximation Methods for Stationary States	95
7.1	Time-independent Perturbation Theory	95
7.1.1	The Perturbation Series	95
7.1.2	First Order Perturbation Theory	96
7.1.3	Second Order Perturbation Theory	98
7.2	Degenerate Perturbation Theory	107
7.2.1	Example: Perturbed 2D Infinite Square Well	110
7.2.2	Example: Linear Stark Effect	112
7.2.3	Nearly Free Electron Model	113
7.3	Variational Method	114
7.3.1	Example: Ground State of Hydrogen Atom	115
7.3.2	Helium Atom Addressed by the Variational Approach	116
7.3.3	The Rayleigh-Ritz Method	117
7.3.4	Example: Hydrogen Atom with Finite Proton Mass	119

8 Atomic Structure	121
8.1 The Non-Relativistic Hydrogen Atom	121
8.1.1 Hydrogen Atom Revisited	121
8.1.2 Including Electron Spin	123
8.1.3 The Zeeman Effect	125
8.2 The “Real” Hydrogen Atom	127
8.2.1 Relativistic Correction to the Kinetic Energy	131
8.2.2 Spin-Orbit Coupling	133
8.2.3 Darwin Term	136
8.2.4 Hydrogen Atom Fine Structure	137
8.2.5 Lamb Shift	139
8.2.6 Hyperfine Structure	140
8.3 Multi-Electron Atoms	145
8.3.1 Central Field Approximation	146
8.3.2 Spin-Orbit Coupling	151
8.3.3 jj Coupling Scheme	155
8.3.4 Zeeman Effect	157
9 Symmetry in Quantum Mechanics	159
9.1 Observables as Generators of Transformations	160
9.2 Consequences of Symmetries: Multiplets	162
9.3 Rotational Symmetry in Quantum Mechanics	162
9.3.1 Scalar Operators	164
9.3.2 Vector Operators	165
9.4 The Wigner-Eckart Theorem for Scalar Operators	168
9.4.1 Consequences of the Wigner-Eckart Theorem (for Scalars)	170
9.5 The Wigner-Eckart Theorem for Vector Operators	170
9.5.1 Selection Rules for Vector Operator Matrix Elements	171
9.5.2 The Landé Projection Formula	172
9.6 Magnetic Dipole Moments	174
9.6.1 g -Factors	175
9.6.2 Combining magnetic moments	176
10 Identical Particles	181
10.1 Quantum Statistics	181
10.2 Space and spin wavefunctions	183
10.3 Physical Consequences of Particle Statistics	184
10.4 Many-Body Systems	187
10.4.1 Non-interacting Fermi Gas	187
10.4.2 Interacting Fermi Gas	188
10.4.3 Non-interacting Bose gas	191
10.4.4 Interacting Bose Gas	191
11 From Molecules to Solids	193
11.1 The H_2^+ Ion	194
11.2 The H_2 molecule	197
11.3 From molecules to solids	200

12 Time-Dependent Perturbation Theory	205
12.1 Time-Dependent Potentials: General Formalism	205
12.1.1 Dynamics of a Driven Two-Level System	206
12.1.2 Paramagnetic Resonance	207
12.2 Time-dependent perturbation theory	209
12.3 “Sudden” Perturbation	210
12.3.1 Harmonic Perturbations: Fermi’s Golden Rule	210
12.3.2 Harmonic Perturbations: Second-Order Transitions	212
13 Scattering Theory	213
13.1 Basics	214
13.2 The Born Approximation	216
14 Radiative Transitions	219
14.1 Coupling of Matter to the Electromagnetic Field	219
14.1.1 Quantum Fields	219
14.1.2 Spontaneous Emission	220
14.1.3 Absorption and Stimulated Emission	221
14.1.4 Einstein’s A and B Coefficients	223
14.2 Selection Rules	224
14.3 Lasers	226
14.3.1 Operating Principles of a Laser	227
14.3.2 Gain Mechanism	228
15 Field Theory: From Phonons to Photons	231
15.1 Quantisation of the Classical Atomic Chain	231
15.1.1 Classical Chain	231
15.1.2 Quantum Chain	234
15.2 Quantum electrodynamics	238
15.2.1 Classical Theory of the Electromagnetic Field	239
15.2.2 Quantum Field Theory of the Electromagnetic Field	241
15.2.3 Fock States	243
15.2.4 Coherent States	243
15.2.5 Non-Classical Light	245
A Appendix	A.1
A.1 Additional Proofs	A.1
A.1.1 Probability Current	A.1
A.1.2 Gaussian Integral	A.2
A.2 Symmetry Transformations	A.4
A.2.1 Time Translations	A.7
A.2.2 Spatial Translations	A.7
A.2.3 Spatial Rotations	A.10
A.3 The Wigner-Eckart Theorem	A.12
A.3.1 Scalar and Vector Operators	A.12
A.3.2 Angular Momentum Relations	A.13
A.3.3 The Wigner-Eckart Theorem for Scalar Operators	A.13
A.3.4 The Wigner-Eckart Theorem for Vector Operators	A.14
A.4 Virial Theorem	A.16
A.5 Density of States	A.17

A.5.1	2D Density of States	A.17
A.5.2	3D Density of States	A.19
A.6	Landau Levels from a Gauge Independent Approach	A.20
A.7	The Coulomb Gauge	A.21
A.7.1	The Symmetric Gauge	A.22
B	Appendix: Equations	B.1
B.1	Wave Mechanics and the Schrödinger Equation 1	B.1
B.2	Quantum Mechanics in One Dimension 2	B.1
B.3	Operator Methods and Formalism in Quantum Mechanics 3	B.1
B.4	Quantum Mechanics in More Than One Dimension 4	B.3
B.5	Spin 5	B.3
B.6	Motion in a Magnetic Field 6	B.3
B.7	Approximation Methods for Stationary States 7	B.3
B.8	Symmetry in Quantum Mechanics 9	B.3
B.9	Identical Particles 10	B.3
B.10	Atomic Structure 8	B.3
B.11	From Molecules to Solids 11	B.3
B.12	Time-Dependent Perturbation Theory 12	B.3
B.13	Scattering Theory 13	B.3
B.14	Radiative Transitions 14	B.3
B.15	Field Theory: From Phonons to Photons 15	B.3

List of Tables

8.1	Hydrogen subshell designations for principal quantum numbers	122
8.2	Uncoupled basis states $ n\ell m_\ell\rangle \otimes sm_s\rangle$ for the $n = 2$ energy level.	123
8.3	Coupled basis states $ njm_j\ell s\rangle$ for the $n = 2$ energy level.	124
8.4	Spectroscopic term notaion.	124
8.5	The total Hydrogen fine structure corrections, calculated as $(\Delta E)_{\text{FS}} = (\Delta E)_{\text{R}} + (\Delta E)_{\text{SO}} + (\Delta E)_{\text{D}}$	139
8.6	The number of electrons accommodated due to degeneracy in shell n	145
8.7	Ordering of Hydrogen orbitals for each shell n , where ℓ denotes the subshell.	149
8.8	Electronic subshell configuration, according to the aufbau principle.	150
8.9	Possibilities of m_ℓ values for electrons to form an antisymmetric total angular momentum state.	154
8.10	Possibilities of m_ℓ values for electrons to form an symmetric total angular momentum state.	154
8.11	Experimental energy ordering of Carbon states using spectroscopy.	154
9.1	Examples of the magnetic moments and g -factors of various particles and nuclei, expressed in terms of both μ_B and μ_N	176
11.1	Binding energy and equilibrium nuclear separation of the hydrogen molecule in various approximations.	198
14.1	Typical values of the coherence length for a number of light sources.	227

List of Figures

1.1	As the temperature of a black body decreases, the emitted thermal radiation decreases in intensity and its maximum moves to longer wavelengths. Shown for comparison is the classical Rayleigh–Jeans law and its ultraviolet catastrophe.	2
1.2	The COBE (Cosmic Background Explorer) satellite made careful measurements of the shape of the spectrum of the emission from the cosmic microwave background. As one can see, the behaviour at low-frequencies (long-wavelengths) conforms well with the predicted Rayleigh-Jeans law translating to a temperature of 2.728K. However, at high frequencies, there is a departure from the predicted ν^2 dependence.	3
1.3	Measurements of the photoelectric effect taken by Robert Millikan showing the variation of the stopping voltage, eV, with variation of the frequency of incident light. Figure reproduced from Robert A. Millikan's Nobel Lecture, <i>The Electron and the Light-Quanta from the Experimental Point of View</i> , <i>The Nobel Foundation</i> , 1923.	3
2.1	Graphical solutions for a square potential well. The top three intersections correspond to a wide well, and the bottom two to a well having half of the width.	18
4.1	Quantisation of angular momentum, a schematic showing the angular momentum scheme for $\ell = 2$ with $\mathbf{L}^2 = \sqrt{2(2+1)}\hbar = \sqrt{6}\hbar$ and the five possible values for the L_z projection. The ladder operators increase/decrease the z -component of angular momentum. The series terminates at the upper and lower ends when L_z becomes greater than the magnitude L by generating the zero vector. Notice that the z -component of angular momentum can be zero, and the average value of the angular momentum can be zero.	57
5.1	The Stern-Gerlach experiment of 1922. Silver atoms travelling through an inhomogeneous magnetic field, and being deflected up or down depending on their spin. The screen revealed discrete points of accumulation, rather than a continuous distribution, owing to their quantised spin. Historically, this experiment was decisive in convincing physicists of the reality of angular-momentum quantisation in all atomic-scale systems.	61
5.2	For a particle of spin s , the spin eigenstates $ s, m_s\rangle$ defined with the z -axis (the \mathbf{B} -field direction) as quantisation axis are also <i>energy</i> eigenstates with $E_{m_s} = -m_s\hbar\omega_s$	70
5.3	$\langle \hat{\mathbf{S}} \rangle$ precesses about the magnetic field direction at the Larmor frequency, ω_s , matching the classical expectation.	71
5.4	Vector addition of orbital and spin angular momentum.	72
5.5	Combining the orbital (red) and spin (blue) ladder operators to create the total (green) ladder operator. The $\ell = 1$ and $s = 1/2$ combinations are shown.	72

6.1 (Left) Schematic showing the geometry of an experiment to observe the Aharonov-Bohm effect. Electrons from a coherent source can follow two paths which encircle a region where the magnetic field is non-zero. (Right) Interference fringes for electron beams passing near a toroidal magnet from the experiment by Tonomura and collaborators in 1986. The electron beam passing through the center of the torus acquires an additional phase, resulting in fringes that are shifted with respect to those outside the torus, demonstrating the Aharonov-Bohm effect.	80
6.2 (Left) A voltage V drives a current I in the positive x -direction. Normal Ohmic resistance is V/I . A magnetic field in the positive z -direction shifts positive charge carriers in the negative y -direction. This generates a Hall potential and a Hall resistance (VH/I) in the y direction. (Right) The Hall resistance varies stepwise with changes in magnetic field B . Step height is given by the physical constant h/e^2 (value approximately $25\text{k}\Omega$) divided by an integer i . The figure shows steps for $i = 2, 3, 4, 5, 6, 8$ and 10 . The effect has given rise to a new international standard for resistance. Since 1990 this has been represented by the unit 1 klitzing, defined as the Hall resistance at the fourth step ($h/4e^2$). The lower peaked curve represents the Ohmic resistance, which disappears at each step	84
6.3 The effect of “switching on” the magnetic field is to transform the (top) $\mathbf{B} = \mathbf{0}$ zero-field, uniform density of states into a set of equally spaced and highly degenerate Landau levels (bottom), while preserving the total number of available states.	85
6.4 For $g = 2$ ($a_\mu = 0$), the spin precession and cyclotron frequencies are equal ($\omega_s = \omega_c$, $\omega_a = 0$). The muon momentum and spin vectors precess at the same rate, remaining tangential to the circular orbit at all times (left-hand diagram). For $g_\mu \approx 2.002$ ($a_\mu \approx 0.001$), we have $\omega_s > \omega_c$, so the spin vector precesses (slightly) faster than the momentum vector. The spin vector therefore rotates relative to the orbital tangent, with angular frequency $\omega_s - \omega_c = \omega_a$ (right-hand diagram).	87
6.5 In the muon rest frame, electrons from muon decay, $\mu^- \rightarrow e^- + \bar{\nu}_e + \nu_\mu$, are emitted non-isotropically with respect to the muon spin direction. The electron distribution (the ellipse) rotates with the muon spin vector.	88
6.6 After a Lorentz transformation to the lab frame, the distribution of electron four-momenta depends on the angle of the spin vector relative to the direction of the Lorentz boost. The electron energy distribution in the lab frame depends on the angle between the muon spin vector and the muon three-momentum.	88
6.7 The plot contains $\sim 3 \times 10^9$ muon decays, about 40% of the total E821 data set. The periodic structure is superimposed on a steady exponential fall due to the muon decay: $N_e(t) = N_0 e^{-t/\tau}$ (the muon has a mean lifetime of $\tau = 2.2\ \mu\text{s}$ in its rest frame). G. W. Bennett et al., Phys. Rev. D 73 (2006) 072003	89
6.8 Stern Gerlach apparatus, each value of m_s gives rise to a different trajectory, whatever the form of $\mathbf{B}_1(\mathbf{r})$, there are $2s + 1$ possible trajectories through the magnetic field, one for each possible value of m_s	92

6.9 Stern Gerlach apparatus for an arbitrary initial spin state $ \psi(0)\rangle = \sum_{m_s} c_{m_s} s, m_s\rangle$. A factor of $ c_{m_s} ^2$ is associated with each individual m_s trajectory. The incoming beam separates into $2s + 1$ outgoing beams, each with different m_s , and with intensity in each proportional to $ c_{m_s} ^2$.	93
7.1 Infinite square well with central bump.	101
7.2 The plot shows the true and perturbation theory energies as a function of the height, ε , of the bump, for a bump of width $b/a = 0.1$ in units of the unperturbed ground state energy $E_0 \equiv \hbar^2\pi^2/(8ma^2)$. For $n = 2$, the energy differences are too small to be visible.	101
7.3 Infinite square well in an electric field.	103
7.4 For the ground state, only antisymmetric states (with even m) contribute to the sum, and $m = 2$ overwhelmingly dominates. The perturbed wavefunctions preserve the boundary conditions, but no longer have definite symmetry.	104
7.5 Two hydrogen atoms, both in the ground state and separated by a distance r .	106
7.6 To leading order, the interaction potential V corresponds to the interaction of two electric dipoles $\mathbf{er}_1, \mathbf{er}_2$ separated by a distance r .	106
7.7 The eigenvalues $E_{n,\alpha}^{(1)}$ will be different for each value of α , splitting the original energy level $E_n^{(0)}$ into distinct levels in the presence of $\hat{H}^{(1)}$.	109
7.8 A perturbation $\hat{H}^{(1)} = \varepsilon \hat{x} \hat{y}$ to the 2D infinite square well splits the degenerate ($g = 2$) first-excited level into two separate levels.	112
8.1 Figure showing the hierarchy of energy shifts of the spectra of hydrogen-like atoms as a result of relativistic corrections. The first column shows the energy spectrum predicted by the (non-relativistic) Bohr theory. The second column shows the predicted energy shifts from relativistic corrections arising from the Dirac theory. The third column includes corrections due quantum electrodynamics and the fourth column includes terms for coupling to the nuclear spin degrees of freedom. The $H - \alpha$ line, particularly important in the astronomy, corresponds to the transition between the levels with $n = 2$ and $n = 3$. Each hyperfine level has degeneracy $g = 2F + 1$. For an <i>isolated</i> atom, this degeneracy (w.r.t. m_F) cannot be broken due to rotational symmetry. For an isolated atom, the energy levels cannot be split further. The degeneracy can only be lifted by putting the atom into an external E or B field, for example, thereby breaking the rotational symmetry.	138
8.2 The hyperfine interaction splits each S-states into two separate levels.	142
8.3 A recent (2016) all-sky survey of neutral atomic hydrogen. Neutral H atoms make up $\sim 44\%$ of interstellar gas (ISG) in the Milky Way. The two small pink areas on the lower left hand side are M31 and M33, and the large orange region represents the Large and Small Magellanic Clouds. The key explains how the colour relates to the relative speed, green is moving away and blue towards. HI4PI Collaboration, A&A 594 (2016) A116.	143
8.4 M33 rotation curve (points) compared with the best-fitting model (continuous line). Visible matter alone cannot explain the large rotational velocities seen at large radius; a spherical halo of <i>dark matter</i> is also needed. E. Corbelli & P. Salucci, MNRAS 311 (2000) 441	144
8.5 Ionisation energy trends plotted against the atomic number, in units eV.	145
8.6 Electron filling order illustrated by the Aufbau principle.	150

8.7 Level scheme of the carbon atom $(1s)^2(2s)^2(2p)^2$. Drawing is not to scale. On the left the energy is shown without any two-particle interaction. The electron-electron interaction leads to a three-fold energy splitting with L and S remaining good quantum numbers. Spin-orbit coupling leads to a further splitting of the states with J remaining a good quantum number. Finally on the right, the levels show Zeeman splittings in an external magnetic field. In this case, the full set of 15 levels become non-degenerate.	155
8.8 The transition from LS to jj coupling is seen in the series C-Si-Ge-Sn-Pb	156
8.9 In the weak field case, the vector model (left) implies that the coupling of the orbital angular momentum L to the spin angular momentum S is stronger than their coupling to the external field. In this case where spin-orbit coupling is dominant, they can be visualised as combining to form a total angular momentum J which then precesses about the magnetic field direction. In the strong field case (right), S and L couple more strongly to the external magnetic field than to each other, and can be visualised as independently precessing about the external field direction.	158
8.10 The well known doublet which is responsible for the bright yellow light from a sodium lamp may be used to demonstrate several of the influences which cause splitting of the emission lines of atomic spectra. The transition which gives rise to the doublet is from the $3p$ to the $3s$ level. The fact that the $3s$ state is lower than the $3p$ state is a good example of the dependence of atomic energy levels on orbital angular momentum. The $3s$ electron penetrates the $1s$ shell more and is less effectively shielded than the $3p$ electron, so the $3s$ level is lower. The fact that there is a doublet shows the smaller dependence of the atomic energy levels on the total angular momentum. The $3p$ level is split into states with total angular momentum $J = 3/2$ and $J = 1/2$ by the spin-orbit interaction. In the presence of an external magnetic field, these levels are further split by the magnetic dipole energy, showing dependence of the energies on the z -component of the total angular momentum.	158
10.1 Energy level diagram for ortho- and parahelium showing the first order shift in energies due to the Coulomb repulsion of electrons. Here we assume that one of the electrons stays close to the ground state of the unperturbed Hamiltonian.	186
10.2 (Left) Schematic showing the phase space volume associated with each plane wave state in a Fermi gas. (Right) Schematic showing the state occupancy of a filled Fermi sea.	188
11.1 Molecular potential energy curves for H_2^+ ion.	196
11.2 Example notation for the labelling of homonuclear diatomic molecular orbitals.	197
11.3 Molecular potential energy and configuration mixing in H_2	199
11.4 One dimensional band structure for E_k for $S = 0.3$, $t = 1$ and $\epsilon = 0$. Here we have shown schematically a band-filling with $k_Fa = \pi/2$	202

11.5 Left: dispersion relation of graphene E_k obtained with the LCAO approximation. Notice that near the centre of the band, the dispersion becomes point-like. Around these points, the dispersion takes form of a linear (Dirac) cone. Right: band structure of a sample of multilayer epitaxial graphene. Linear bands emerge from the K -points in the band structure. Three “cones” can be seen from three layers of graphene in the MEG sample.	203
11.6 Lattice vectors in periodic graphene lattice.	204
13.1 Notation and labelling for classical scattering phenomena. 213	
14.1 Schematic showing spontaneous emission from an initial state at energy $E_i = E_2$ to a final state at energy $E_f = E_1$	221
14.2 Schematic showing absorption from an initial state at energy $E_i = E_1$ to a final state at energy $E_f = E_2$	222
14.3 Schematic showing the stimulated emission from an initial state at energy $E_i = E_2$ to a final state at energy $E_f = E_1$	222
15.1 Toy model of a one-dimensional solid: a chain of point-like particles each of mass m coupled elastically by springs with spring constant k_s 231	
15.2 Description of the lattice displacements ϕ_n in terms of <i>smooth functions</i> , $\phi(x)$ of a continuous variable x	232
15.3 Illustration of the quantum system’s discrete wave-like excitations, each indexed by a wave number $k = 2\pi m/L$. These excitations are depicted as independent collective oscillator modes with k -dependent frequencies, distinct from the coupled atomic constituents also behaving as oscillators.” .	236
15.4 Harmonic oscillator energy state ϵ_0 , as an accumulation of n elementary entities, or quasi-particles, each having energy $\hbar\omega$. These quasi-particles must be bosons since the same state $\hbar\omega$ can be occupied by more than one particle.	236
15.5 The figure shows a typical measured phonon dispersion of an ordered crystalline solid obtained by neutron scattering. The x -axis indexes wavenumbers along a lattice direction (specified in units of π/a). Three generic aspects are visible: (1) near $k = 0$, the dispersion is, as expected, linear. (2) The several branches are associated with different “polarisations” of the lattice fluctuations. (3) For wavelengths comparable to the lattice spacing, $k \sim \pi/a$, non-universal features specific to the particular material become visible.	238
15.6 EM waveguide with rectangular cross-section. The structure of the eigenmodes of the EM field is determined by boundary conditions at the walls of the cavity.	240
A.1 Quantised locations of allowed \mathbf{k} vectors in k -space, showing the annulus containing a number of available states dN between k and $k + dk$, where $k = \mathbf{k} $ A.18	

CHAPTER 1

Wave Mechanics and the Schrödinger Equation

The formulation of a consistent theory of statistical mechanics, electrodynamics and special relativity during the latter half of the 19th century and the early part of the 20th century had been a triumph of “unification”. However, the undoubted success of these theories gave an impression that physics was a mature, complete, and predictive science. Nowhere was confidence expressed more clearly than in the famous quote made at the time by Lord Kelvin: *There is nothing new to be discovered in physics now. All that remains is more and more precise measurement.* However, there were a number of seemingly unrelated and unsettling problems that challenged the prevailing theories.

1.1 Historical Foundations of Quantum Physics

1.1.1 Black-Body Radiation

In 1860, Gustav Kirchhoff introduced the concept of a “black body”, an object that absorbs all electromagnetic radiation that falls upon it – none passes through and none is reflected. Since no light is reflected or transmitted, the object appears black when it is cold. However, above absolute zero, a black body emits thermal radiation with a spectrum that depends on temperature. To determine the spectrum of radiated energy, it is helpful to think of a black body as a thermal cavity at a temperature, T . The energy radiated by the cavity can be estimated by considering the resonant modes. In three-dimensions, the number of modes, per unit frequency per unit volume is given by

$$N(\nu) \, d\nu = \frac{8\pi\nu^2}{c^3} \, d\nu, \quad (1.1)$$

where, as usual, c is the speed of light.

The amount of radiation emitted in a given frequency range should be proportional to the number of modes in that range. Within the framework of classical statistical mechanics, each of these modes has an equal chance of being excited, and the average energy in each mode is $k_B T$ (equipartition), where k_B is the Boltzmann constant. The corresponding energy density is therefore given by the **Rayleigh-Jeans law**,

$$\rho(\nu, T) = \frac{8\pi\nu^2}{c^3} k_B T. \quad (1.2)$$

This result predicts that $\rho(\nu, T)$ increases without bound at high frequencies, ν — the ultraviolet (UV) catastrophe. However, such behaviour stood in contradiction with experiment which revealed that the high-frequency dependence is quite benign (see, e.g., Fig. 1.2). To resolve difficulties presented by the UV catastrophe, Planck hypothesised

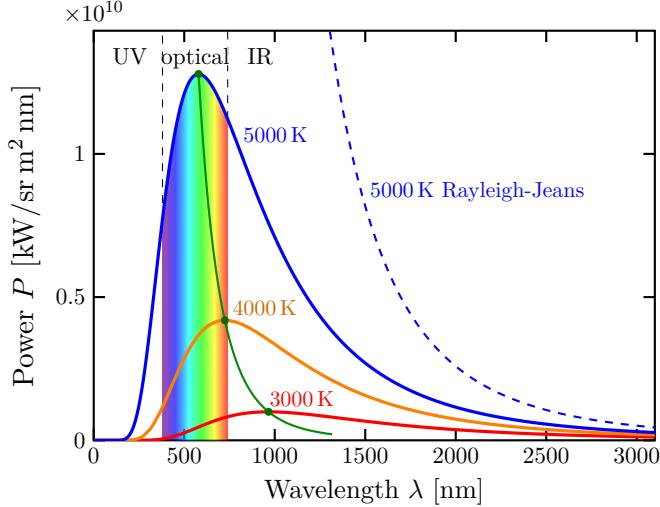


Fig. 1.1: As the temperature of a black body decreases, the emitted thermal radiation decreases in intensity and its maximum moves to longer wavelengths. Shown for comparison is the classical Rayleigh–Jeans law and its ultraviolet catastrophe.

that, for each mode ν , energy is quantised in units of $h\nu$, where h denotes the Planck constant. In this case, the energy of each mode is given by

$$\langle \varepsilon(\nu) \rangle = \frac{\sum_{n=0}^{\infty} n h \nu e^{-n h \nu / k_B T}}{\sum_{n=0}^{\infty} \nu e^{-n h \nu / k_B T}} = \frac{h \nu}{e^{h \nu / k_B T} - 1}, \quad (1.3)$$

leading to the anticipated suppression of high frequency modes. From this result one obtains the celebrated Planck radiation formula,

$$\rho(\nu, T) = \frac{8\pi\nu^2}{c^3} \langle \varepsilon(\nu) \rangle = \frac{8\pi\nu^2}{c^3} \frac{h\nu}{e^{h\nu/k_B T} - 1}. \quad (1.4)$$

This result conforms with experiment (Fig. 1.2), and converges on the Rayleigh–Jeans law at low frequencies, $h\nu/k_B T \rightarrow 0$. Planck’s result suggests that electromagnetic energy is quantised: light of wavelength $\lambda = c/\nu$ is made up of quanta each of which has energy $h\nu$. The equipartition law fails for oscillation modes with high frequencies, $h\nu \gg k_B T$.

A quantum theory for the specific heat of matter, which takes into account the quantisation of lattice vibrational modes, was subsequently given by Debye and Einstein.

1.1.2 Photoelectric Effect

We turn now to the second ground-breaking experiment in the development of quantum theory. When a metallic surface is exposed to electromagnetic radiation, above a certain threshold frequency, the light is absorbed and electrons are emitted. In 1902, Philipp Eduard Anton von Lenard observed that the energy of individual emitted electrons increases with the frequency of the light. This was at odds with Maxwell’s wave theory of light, which predicted that the electron energy would be proportional to the intensity of the radiation.

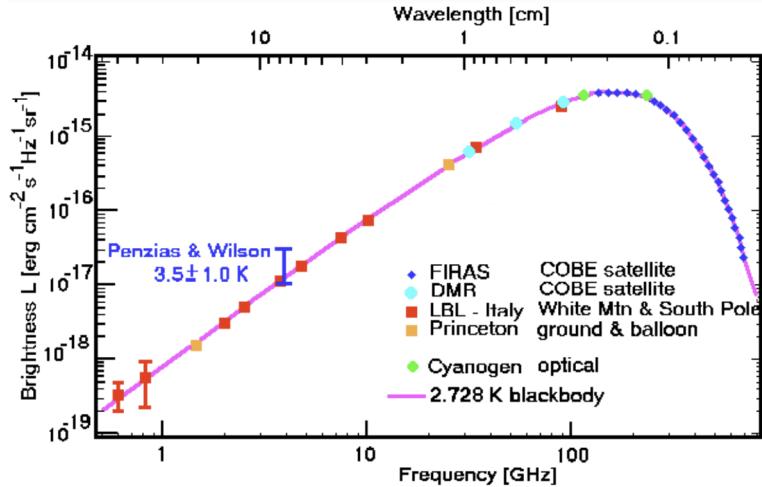


Fig. 1.2: The COBE (Cosmic Background Explorer) satellite made careful measurements of the shape of the spectrum of the emission from the cosmic microwave background. As one can see, the behaviour at low-frequencies (long-wavelengths) conforms well with the predicted Rayleigh-Jeans law translating to a temperature of 2.728K. However, at high frequencies, there is a departure from the predicted ν^2 dependence.

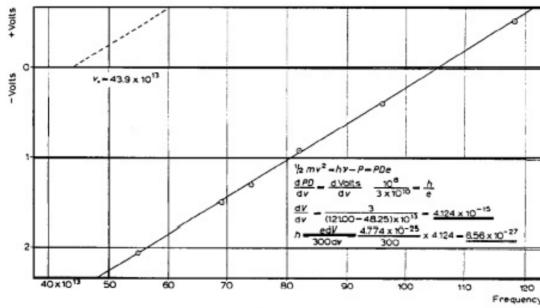


Fig. 1.3: Measurements of the photoelectric effect taken by Robert Millikan showing the variation of the stopping voltage, eV, with variation of the frequency of incident light. Figure reproduced from Robert A. Millikan's Nobel Lecture, *The Electron and the Light-Quanta from the Experimental Point of View, The Nobel Foundation*, 1923.

In 1905, Einstein resolved this paradox by describing light as composed of discrete quanta (photons), rather than continuous waves. Based upon Planck's theory of black-body radiation, Einstein theorised that the energy in each quantum of light was proportional to the frequency. A photon above a threshold energy, the "work function" W of the metal, has the required energy to eject a single electron, creating the observed effect. In particular, Einstein's theory was able to predict that the maximum kinetic energy of electrons emitted by the radiation should vary as

$$\text{K.E.}_{\text{max}} = h\nu - W. \quad (1.5)$$

Later, in 1916, Millikan was able to measure the maximum kinetic energy of the emitted electrons using an evacuated glass chamber. The kinetic energy of the photoelectrons was found by measuring the potential energy of the electric field, eV, needed to stop them. As well as confirming the linear dependence of the kinetic energy on frequency, by making use of his estimate for the electron charge, e , established from his oil drop experiment

in 1913, he was able to determine Planck's constant to a precision of around 0.5%. This discovery led to the quantum revolution in physics and earned Einstein the Nobel Prize in 1921.

1.1.3 Compton Scattering

In 1923, Compton investigated the scattering of high energy X-rays and γ -ray from electrons in a carbon target. By measuring the spectrum of radiation at different angles relative to the incident beam, he found two scattering peaks. The first peak occurred at a wavelength which matched that of the incident beam, while the second varied with angle. Within the framework of a purely classical theory of the scattering of electromagnetic radiation from a charged particle – Thomson scattering – the wavelength of a low-intensity beam should remain unchanged.

Compton's observation demonstrated that light cannot be explained purely as a classical wave phenomenon. Light must behave as if it consists of particles in order to explain the low-intensity Compton scattering. If one assumes that the radiation is comprised of photons that have a well defined momentum as well as energy, $p = \frac{h\nu}{c} = \frac{h}{\lambda}$, the shift in wavelength can be understood: The interaction between electrons and high energy photons (ca. keV) results in the electron being given part of the energy (making it recoil), and a photon with the remaining energy being emitted in a different direction from the original, so that the overall momentum of the system is conserved. By taking into account both conservation of energy and momentum of the system, the Compton scattering formula describing the shift in the wavelength as function of scattering angle θ can be derived,¹

$$\Delta\lambda = \lambda' - \lambda = \frac{h}{m_e c} (1 - \cos \theta). \quad (1.6)$$

The constant of proportionality $h/m_e c = 0.002426$ nm, the Compton wavelength, characterises the scale of scattering. Moreover, as $h \rightarrow 0$, one finds that $\Delta\lambda \rightarrow 0$ leading to the classical prediction.

1.1.4 Atomic Spectra

The discovery by Rutherford that the atom was comprised of a small positively charged nucleus surrounded by a diffuse cloud of electrons led naturally to the consideration of a planetary model of the atom. However, a classical theory of electrodynamics would predict that an accelerating charge would radiate energy leading to the eventual collapse of the electron into the nucleus. Moreover, as the electron spirals inwards, the emission would

¹If we assume that the total energy and momentum are conserved in the scattering of a photon (γ) from an initially stationary target electron (e), we have $E_\gamma + E_e = E_{\gamma'} + E_{e'}$ and $\mathbf{p}_\gamma = \mathbf{p}_{\gamma'} + \mathbf{p}_{e'}$. Here $E_\gamma = h\nu$ and $E_e = m_e c^2$ denote the energy of the photon and electron before the collision, while $E_{\gamma'} = h\nu'$ and $E_{e'} = \sqrt{(p_{e'} c)^2 + (mc^2)^2}$ denote the energies after. From the equation for energy conservation, one obtains $(p_{e'} c)^2 = (h(\nu - \nu') + m_e c^2)^2 - (m_e c^2)^2$. From the equation for momentum conservation, one obtains $p_{e'}^2 = p_\gamma^2 + p_{\gamma'}^2 - 2|p_\gamma||p_{\gamma'}|\cos\theta$. Then, noting that $E_\gamma = p_\gamma c$, a rearrangement of these equations obtains the Compton scattering formula.

gradually increase in frequency leading to a broad continuous spectra. Yet, detailed studies of electrical discharges in low-pressure gases revealed that atoms emit light at discrete frequencies. The clue to resolving these puzzling observations lay in the discrete nature of atomic spectra. For the hydrogen atom, light emitted when the atom is thermally excited has a particular pattern: Balmer had discovered in 1885 that the emitted wavelengths follow the empirical law, $\lambda = \lambda_0(1/4 - 1/n^2)$ where $n = 3, 4, 5, \dots$ and $\lambda_0 = 3645.6 \text{ \AA}$. Neils Bohr realised that these discrete values of the wavelength reflected the emission of individual photons having energy equal to the energy difference between two allowed orbits of the electron circling the nucleus (the proton), $E_n - E_m = h\nu$, leading to the conclusion that the allowed energy levels must be quantised and varying as $E_n = -\frac{hcR_H}{n^2}$, where $R_H = 109678 \text{ cm}^{-1}$ denotes the Rydberg constant.

How could the quantum $h\nu$ restricting allowed radiation energies also restrict the allowed electron orbits? In 1913 Bohr proposed that the angular momentum of an electron in one of these orbits is quantised in units of Planck's constant,

$$L = m_e v r = n\hbar, \quad \hbar = \frac{h}{2\pi}. \quad (1.7)$$

But why should only certain angular momenta be allowed for the circling electron? A heuristic explanation was provided by de Broglie: just as the constituents of light waves (photons) are seen through Compton scattering to act like particles (of definite energy and momentum), so particles such as electrons may exhibit wave-like properties. For photons, we have seen that the relationship between wavelength and momentum is $p = h/\lambda$. de Broglie hypothesised that the inverse was true: for particles with a momentum p , the wavelength is

$$\lambda = \frac{h}{p}, \quad \text{i.e. } p = \hbar k \quad (1.8)$$

where k denotes the wavevector of the particle. Applied to the electron in the atom, this result suggested that the allowed circular orbits are standing waves, from which Bohr's angular momentum quantisation follows. The de Broglie hypothesis found quantitative support in an experiment by Davisson and Germer, and independently by G. P. Thomson in 1927. Their studies of electron diffraction from a crystalline array of Nickel atoms confirmed that the diffraction angles depend on the incident energy (and therefore momentum).

1.2 Wave Mechanics

de Broglie's doctoral thesis, defended at the end of 1924, created a lot of excitement in European physics circles. Shortly after it was published in the Autumn of 1925, Pieter Debye, a theorist in Zurich, suggested to Erwin Schrödinger that he give a seminar on de Broglie's work. Schrödinger gave a polished presentation, but at the end, Debye remarked that he considered the whole theory rather childish: Why should a wave confine itself to a circle in space? It wasn't as if the circle was a waving circular string; real waves in space diffracted and diffused; in fact they obeyed three-dimensional wave equations, and that was what was needed. This was a direct challenge to Schrödinger, who spent some weeks in the Swiss mountains working on the problem, and constructing his equation.

1.2.1 Maxwell's Wave Equation

For a monochromatic wave in vacuum, with no currents or charges present, Maxwell's wave equation,

$$\boxed{\nabla^2 \mathbf{E} - \frac{1}{c^2} \ddot{\mathbf{E}} = 0,} \quad (1.9)$$

admits the plane wave solution, $\mathbf{E} = \mathbf{E}_0 e^{i(\mathbf{k} \cdot \mathbf{r} - \omega t)}$, with the linear dispersion relation $\omega = c|\mathbf{k}|$ and c the velocity of light. Here, (and throughout the text) we adopt the convention, $\ddot{\mathbf{E}} = \partial_t^2 \mathbf{E}$. We know from the photoelectric effect and Compton scattering that the photon energy and momentum are related to the frequency and wavelength of light through the relations $E = h\nu = \hbar\omega$, $p = \frac{h}{\lambda} = \hbar k$. The wave equation tells us that $\omega = c|\mathbf{k}|$ and hence $E = c|\mathbf{p}|$. If we think of $e^{i(\mathbf{k} \cdot \mathbf{r} - \omega t)}$ as describing a particle (photon) it would be more natural to write the plane wave in terms of the energy and momentum of the particle as $E_0 e^{i(\mathbf{p} \cdot \mathbf{r} - Et)/\hbar}$. Then, one may see that the wave equation applied to the plane wave describing particle propagation yields the familiar energy momentum relationship, $E^2 = (c\mathbf{p})^2$ for a massless relativistic particle.

This discussion suggests how one might extend the wave equation from the photon (with zero rest mass) to a particle with rest mass m_0 . We require a wave equation that, when it operates on a plane wave, yields the relativistic energy-momentum invariant, $E^2 = (c\mathbf{p})^2 + (m_0 c^2)^2$. Writing the plane wave function $\phi(\mathbf{r}, t) = A e^{i(\mathbf{p} \cdot \mathbf{r} - Et)/\hbar}$, where A is a constant, we can recover the energy-momentum invariant by adding a constant mass term to the wave operator,

$$\left(\nabla^2 - \frac{\partial_t^2}{c^2} - \frac{m_0 c^2}{\hbar^2} \right) e^{i(\mathbf{p} \cdot \mathbf{r} - Et)/\hbar} = -\frac{((c\mathbf{p})^2 - E^2 + m_0^2 c^4)}{(\hbar c)^2} e^{i(\mathbf{p} \cdot \mathbf{r} - Et)/\hbar} = 0. \quad (1.10)$$

This wave equation is called the **Klein-Gordon equation** and correctly describes the propagation of relativistic particles of mass m_0 . However, its form is inappropriate for non-relativistic particles, like the electron in hydrogen.

Continuing along the same lines, let us assume that a non-relativistic electron in free space is also described by a plane wave of the form $\Psi(\mathbf{r}, t) = A e^{i(\mathbf{p} \cdot \mathbf{r} - Et)/\hbar}$. We need to construct an operator which, when applied to this wave function, just gives us the ordinary non-relativistic energy-momentum relation, $E = \frac{\mathbf{p}^2}{2m}$. The factor of \mathbf{p}^2 can be recovered from two derivatives with respect to \mathbf{r} , but the only way we can get E is by having a single differentiation with respect to time, i.e.

$$i\hbar \partial_t \Psi(\mathbf{r}, t) = -\frac{\hbar^2}{2m} \nabla^2 \Psi(\mathbf{r}, t). \quad (1.11)$$

This is Schrödinger's equation for a free non-relativistic particle. One remarkable feature of this equation is the factor of i which shows that the wavefunction is complex.

How, then, does the presence of a spatially varying scalar potential affect the propagation of a de Broglie wave? This question was considered by Sommerfeld in an attempt to generalise the rather restrictive conditions in Bohr's model of the atom. Since the electron orbit was established by an inverse square force law, just like the planets around the Sun,

Sommerfeld couldn't understand why Bohr's atom had only circular orbits as opposed to Keplerlike elliptical orbits. (Recall that all of the observed spectral lines of hydrogen were accounted for by energy differences between circular orbits.)

de Broglie's analysis of the allowed circular orbits can be formulated by assuming that, at some instant, the spatial variation of the wavefunction on going around the orbit includes a phase term of the form $e^{ipq/\hbar}$, where here the parameter q measures the spatial distance around the orbit. Now, for an acceptable wavefunction, the total phase change on going around the orbit must be $2\pi n$, where n is integer. For the usual Bohr circular orbit, where $p = |\mathbf{p}|$ is constant, this leads to quantisation of the angular momentum $L = pr = n\hbar$.

Sommerfeld considered a general Keplerian elliptical orbit. Assuming that the de Broglie relation $p = h/\lambda$ still holds, the wavelength must vary as the particle moves around the orbit, being shortest where the particle travels fastest, at its closest approach to the nucleus. Nevertheless, the phase change on moving a short distance Δq should still be $p\Delta q/\hbar$. Requiring the wavefunction to link up smoothly on going once around the orbit gives the **Bohr-Sommerfeld quantisation condition**

$$\oint p \, dq = nh, \quad (1.12)$$

where \oint denotes the line integral around a closed orbit. Thus only certain *elliptical* orbits are allowed. The mathematics is non-trivial, but it turns out that every allowed elliptical orbit has the same energy as one of the allowed circular orbits. That is why Bohr's theory gave the correct energy levels. This analysis suggests that, in a varying potential, the wavelength changes in concert with the momentum.

1.2.2 Schrödinger's Equation

Following Sommerfeld's considerations, let us then consider a particle moving in one spatial dimension subject to a "roller coaster-like" potential. How do we expect the wavefunction to behave? As discussed above, we would expect the wavelength to be shortest where the potential is lowest, in the minima, because that's where the particle is going the fastest. Our task then is to construct a wave equation which leads naturally to the relation following from (classical) energy conservation, $E = \frac{p^2}{2m} + V(x)$. In contrast to the free particle case discussed above, the relevant wavefunction here will no longer be a simple plane wave, since the wavelength (determined through the momentum via the de Broglie relation) varies with the potential. However, at a given position x , the momentum is determined by the "local wavelength". The appropriate wave equation is the one-dimensional Schrödinger equation,

$$i\hbar\partial_t\Psi(x, t) = -\frac{\hbar^2\partial_x^2}{2m}\Psi(x, t) + V(x)\Psi(x, t), \quad (1.13)$$

with the generalisation to three-dimensions leading to the Laplacian operator ∇^2 in place of ∂_x^2 (cf. Maxwell's equation).

So far, the validity of this equation rests on plausibility arguments and hand-waving. Why should anyone believe that it really describes an electron wave? Schrödinger's test of

his equation was the hydrogen atom. He looked for Bohr's "stationary states": states in which the electron was localised somewhere near the proton, and having a definite energy. The time dependence would be the same as for a plane wave of definite energy, $e^{-Et/\hbar}$; the spatial dependence would be a time-independent function decreasing rapidly at large distances from the proton. From the solution of the stationary wave equation for the Coulomb potential, he was able to deduce the allowed values of energy and momentum. These values were exactly the same as those obtained by Bohr (except that the lowest allowed state in the "new" theory had zero angular momentum): impressive evidence that the new theory was correct.

1.3 Postulates of Quantum Theory

Since there remains no "first principles" derivation of the quantum mechanical equations of motion, the theory is underpinned by a set of "postulates" whose validity rest on experimental verification. Needless to say, quantum mechanics remains perhaps the most successful theory in physics.

To build a quantum theory, we can start from a Hilbert space \mathcal{H} and define observables as operators, or we can start from operator algebra and define \mathcal{H} as a representation.

1.3.1 Postulate 1

The state of a quantum mechanical system is completely specified by a complex function $\Psi(\mathbf{r}, t)$ that depends upon the coordinates of the particle(s) and on time. This function, called the wavefunction or state function, has the important property that $|\Psi(\mathbf{r}, t)|^2 d\mathbf{r}$ represents the probability that the particle lies in the volume element $d\mathbf{r} \equiv d^d\mathbf{r}$ located at position \mathbf{r} at time t . The function may also depend on other coordinates too.

The wavefunction must satisfy certain mathematical conditions because of this probabilistic interpretation. For the case of a single particle in a bound state, the net probability of finding it at some point in space must be unity leading to the normalisation condition, $\int_{-\infty}^{\infty} |\Psi(\mathbf{r}, t)|^2 d\mathbf{r} = 1$. It is customary to also normalise many-particle wavefunctions to unity. The wavefunction must also be single-valued, continuous, and finite and (usually) square-integrable.

The ket $|\Psi\rangle$ belongs to a *Hilbert space* $|\Psi\rangle \in \mathcal{H}$, which is a complex vector space equipped with an inner product, which is to say there exists an *inner product* $\langle \phi | \psi \rangle$ associating a complex number to each pair of elements $|\phi\rangle, |\psi\rangle \in \mathcal{H}$ that satisfies the following properties:

- The inner product is conjugate symmetric; that is, the inner product of a pair of elements is equal to the complex conjugate of the inner product of the swapped elements:

$$\langle \psi | \phi \rangle = \langle \phi | \psi \rangle^*. \quad (1.14)$$

Importantly, this implies that $\langle \psi | \psi \rangle \in \mathbb{R}$.

- The inner product is linear in its second² argument. For all complex numbers α and β ,

$$\langle \phi | \alpha \psi_1 + \beta \psi_2 \rangle = \alpha \langle \phi | \psi_1 \rangle + \beta \langle \phi | \psi_2 \rangle. \quad (1.15)$$

- The inner product of an element with itself is positive definite

$$\begin{aligned} \langle \psi | \psi \rangle &> 0 & \text{if } x \neq 0, \\ \langle \psi | \psi \rangle &= 0 & \text{if } x = 0. \end{aligned} \quad (1.16)$$

1.3.2 Postulate 2

To every observable in classical mechanics there corresponds a linear, complete, Hermitian operator in quantum mechanics.

If we require that the expectation value of an operator \hat{A} is real, then it follows that \hat{A} must be a Hermitian operator. If the result of a measurement of an operator \hat{A} is the number a , then a must be one of the eigenvalues, $\hat{A}\Psi = a\Psi$, where Ψ is the corresponding eigenfunction. This postulate captures a central point of quantum mechanics – the values of dynamical variables can be quantised (although it is still possible to have a continuum of eigenvalues in the case of unbound states). For every Hermitian operator, we can always choose an orthogonal set of eigenstates.

1.3.3 Postulate 3

If a system is in a state described by a normalised wavefunction Ψ , then the average value of the observable corresponding to \hat{A} is given by

$$\langle \hat{A} \rangle_{\Psi} = \langle \Psi | \hat{A} | \Psi \rangle = \int_{-\infty}^{\infty} \Psi^* \hat{A} \Psi \, dr. \quad (1.17)$$

If the system is in an eigenstate of \hat{A} with eigenvalue a , then any measurement of the quantity A will yield a . Although measurements must always yield an eigenvalue, the state does not have to be an eigenstate of \hat{A} initially. An arbitrary state can be expanded in the complete set of eigenvectors of \hat{A} ($\hat{A}\Psi_i = a_i\Psi_i$) as $\Psi = \sum_i^n c_i\Psi_i$, where n may go to infinity. In this case, the probability of obtaining the result a_i from the measurement of \hat{A} is given by $P(a_i) = |\langle \Psi_i | \Psi \rangle|^2 = |c_i|^2$. The expectation value of \hat{A} for the state Ψ is the sum over all possible values of the measurement and given by

$$\langle \hat{A} \rangle_{\Psi} = \sum_i a_i |\langle \Psi_i | \Psi \rangle|^2 = \sum_i |c_i|^2. \quad (1.18)$$

Finally, a measurement of Ψ which leads to the eigenvalue a_i , causes the wavefunction to “collapses” into the corresponding eigenstate Ψ_i . (In the case that a_i is degenerate, then Ψ becomes the projection of Ψ onto the degenerate subspace). Thus, measurement

²In some conventions, inner products are linear in their first arguments instead.

affects the state of the system, and any complete set $\{|\phi_n\rangle\}$ is in 1-1 correspondence with a possible outcome of a measurement. In general, we have the *Born rule* to determine the probability of observing a collapse onto an eigenstate $|\phi\rangle$ from a measurement of the state $|\psi\rangle$,

$$\text{Prob}(|\psi\rangle \rightarrow |\phi\rangle) = \frac{|\langle\phi|\psi\rangle|^2}{\langle\phi|\phi\rangle\langle\psi|\psi\rangle} \in \mathbb{R}, \quad (1.19)$$

where the denominator is such that this holds also for un-normalised states.

1.3.4 Postulate 4

The dynamical evolution of a quantum wavefunction or state function of a system evolves in accordance with the time-dependent Schrödinger equation

$$i\hbar\frac{\partial\Psi}{\partial t} = \hat{H}\Psi(\mathbf{r}, t), \quad (1.20)$$

where \hat{H} is the Hamiltonian of the system. If Ψ is an eigenstate of \hat{H} , it follows that $\Psi(\mathbf{r}, t) = \Psi(\mathbf{r}, 0)e^{-iEt/\hbar}$.

CHAPTER 2

Quantum Mechanics in One Dimension

Following the rules of quantum mechanics, we have seen that the state of a quantum particle, subject to a scalar potential $V(\mathbf{r}, t)$, is described by the time-dependent Schrödinger equation,

$$i\hbar\partial_t\Psi(\mathbf{r},t) = -\frac{\hbar^2\nabla^2}{2m}\Psi(\mathbf{r},t) + V(\mathbf{r},t)\Psi(\mathbf{r},t). \quad (2.1)$$

As with all second order linear differential equations, if the potential $V(\mathbf{r}, t)$ is time-independent, the time-dependence of the wavefunction can be separated from the spatial dependence. Setting $\Psi(\mathbf{r}, t) = T(t)\psi(\mathbf{r})$, and separating the variables, the Schrödinger equation takes the form,

$$\frac{\left(-\frac{\hbar^2\nabla^2}{2m}\psi(\mathbf{r}) + V(\mathbf{r})\psi(\mathbf{r})\right)}{\psi(\mathbf{r})} = \frac{i\hbar\partial_tT(t)}{T(t)} = \text{const.} = E. \quad (2.2)$$

Since we have a function of only \mathbf{r} set equal to a function of only t , they both must equal a constant. In the equation above, we call the constant E (with some knowledge of the outcome). We now have an equation in t set equal to a constant, $i\hbar\partial_tT(t) = ET(t)$, which has a simple general solution, $T(t) = Ce^{-iEt/\hbar}$, where C is some constant. The corresponding equation in \mathbf{r} is then given by the stationary, or **time-independent Schrödinger equation**,

$$-\frac{\hbar^2\nabla^2}{2m}\psi(x) + V(x)\psi(x) = E\psi(x). \quad (2.3)$$

The full time-dependent solution is given by $\Psi(\mathbf{r}, t) = e^{-iEt/\hbar}\psi(\mathbf{r})$ with definite energy, E . Their probability density $|\Psi(\mathbf{r}, t)|^2 = |\psi(\mathbf{r})|^2$ is constant in time – hence they are called stationary states! The operator

$$\hat{H} = -\frac{\hbar^2\nabla^2}{2m} + V(\mathbf{r}) \quad (2.4)$$

defines the **Hamiltonian** and the stationary wave equation can be written as the eigenfunction equation, $\hat{H}\psi(\mathbf{r}) = E\psi(\mathbf{r})$, i.e. $\psi(\mathbf{r})$ is an eigenstate of \hat{H} with eigenvalue E .

To explore its properties, we will first review some simple and, hopefully, familiar applications of the equation to one-dimensional systems. In addressing the one-dimensional geometry, we will divide our consideration between potentials, $V(x)$, which leave the particle free (i.e. unbound), and those that bind the particle to some region of space.

2.1 Wave Mechanics of Unbound Particles

2.1.1 Particle Flux and conservation of probability

In analogy to the Poynting vector for the electromagnetic field, we may want to know the probability current. For example, for a free particle system, the probability density is uniform over all space, but there is a net flow along the direction of momentum. We can derive an equation showing conservation of probability by differentiating the probability density, $P(x, t) = |\Psi(x, t)|^2$, and using the Schrödinger equation, $\partial_t P(x, t) + \partial_x j(x, t) = 0$. This translates to the usual conservation equation if $j(x, t)$ is identified as the probability current,

$$j(x, t) = -\frac{i\hbar}{2m} [\Psi^* \partial_x \Psi - \Psi \partial_x \Psi^*]. \quad (2.5)$$

If we integrate over some interval in x , $\int_a^b \partial_t P(x, t) dx = -\int_a^b \partial_x j(x, t) dx$ it follows that $\partial_t \int_a^b P(x, t) dx = j(x = a, t) - j(x = b, t)$, i.e. the rate of change of probability is equal to the net flux entering the interval.

To extending this analysis to three space dimensions, we use the general form of the continuity equation, $\partial_t P(\mathbf{r}, t) + \nabla \cdot \mathbf{j}(\mathbf{r}, t) = 0$, from which follows the particle flux,

$$\boxed{\mathbf{j}(\mathbf{r}, t) = -\frac{i\hbar}{2m} [\Psi^*(\mathbf{r}, t) \nabla \Psi(\mathbf{r}, t) - \Psi(\mathbf{r}, t) \nabla \Psi^*(\mathbf{r}, t)]}. \quad (2.6)$$

For completeness, the proof of Eq. (2.6) is provided in the Appendix A.1.1.

2.1.2 Free Particle

In the absence of an external potential, the time-dependent Schrödinger equation (2.1) describes the propagation of travelling waves. In one dimension, the corresponding complex wavefunction has the form

$$\Psi(x, t) = A e^{i(kx - \omega t)}, \quad (2.7)$$

where A is the amplitude, and $E(k) = \hbar\omega(k) = \frac{\hbar^2 k^2}{2m}$ represents the free particle energy dispersion for a non-relativistic particle of mass, m , and wavevector $k = 2\pi/\lambda$ with λ the wavelength. Each wavefunction describes a plane wave in which the particle has definite energy $E(k)$ and, in accordance with the de Broglie relation, momentum $p = \hbar/k = h/\lambda$. The energy spectrum of a freely-moving particle is therefore continuous, extending from zero to infinity and, apart from the spatially constant state $k = 0$, has a two-fold degeneracy corresponding to right and left moving particles.

For an infinite system, it makes no sense to fix the amplitude A by the normalisation of the total probability. Instead, it is useful to fix the flux associated with the wavefunction. Making use of (2.5) for the particle current, the plane wave is associated with a constant (time-independent) flux,

$$j(x, t) = -\frac{i\hbar}{2m} [\Psi^* \partial_x \Psi - \Psi \partial_x \Psi^*] = |A|^2 \frac{\hbar k}{m} = |A|^2 \frac{p}{m}. \quad (2.8)$$

For a given value of the flux j , the amplitude is given, up to an arbitrary constant phase, by $A = \sqrt{mj/\hbar k}$.

To prepare a **wave packet** which is localised to a region of space, we must superpose components of different wave number. In an open system, this may be achieved using a Fourier expansion. For any function,¹ $\psi(x)$, we have the Fourier decomposition,

$$\psi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \psi(k) e^{ikx} dk, \quad (2.9)$$

where the coefficients are defined by the inverse transform,

$$\psi(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \psi(x) e^{-ikx} dx. \quad (2.10)$$

The normalisation of $\psi(k)$ follows automatically from the normalisation of $\psi(x)$, $\int_{-\infty}^{\infty} \psi^*(k)\psi(k) dk = \int_{-\infty}^{\infty} \psi^*(x)\psi(x) dx$, and both can represent probability amplitudes. Applied to a wavefunction, $\psi(x)$ can be understood as a wave packet made up of contributions involving definite momentum states, e^{ikx} , with amplitude set by the Fourier coefficient $\psi(k)$. The probability for a particle to be found in a region of width dx around some value of x is given by $|\psi(x)|^2 dx$. Similarly, the probability for a particle to have wave number k in a region of width dk around some value of k is given by $|\psi(k)|^2 dk$. (Remember that $p = \hbar k$ so the momentum distribution is very closely related. Here, for economy of notation, we work with k .)

The Fourier transform of a normalised Gaussian wave packet, $\psi(k) = \left(\frac{2\alpha}{\pi}\right)^{1/4} e^{-\alpha(k-k_0)^2}$ is also a Gaussian,

$$\psi(x) = \left(\frac{1}{2\pi\alpha}\right)^{1/4} e^{ik_0x} e^{-\frac{x^2}{4\alpha}} \quad (2.11)$$

From these representations, we can see that it is possible to represent a *single* particle, localised in real space as a superposition of plane wave states localised in Fourier space. But note that, while we have achieved our goal of finding localised wave packets, this has been at the expense of having some non-zero width in x and in k .

For the Gaussian wave packet, we can straightforwardly obtain the width (as measured by the root mean square – RMS) of the probability distribution, $\Delta x = (\langle (x - \langle x \rangle)^2 \rangle)^{1/2} = (\langle x^2 \rangle - \langle x \rangle^2)^{1/2} = \sqrt{\alpha}$, and $\Delta k = \frac{1}{\sqrt{4\alpha}}$. We can again see that, as we vary the width in k -space, the width in x -space varies to keep the following product constant, $\Delta x \Delta k = \frac{1}{2}$. If we translate from the wavevector into momentum $p = \hbar k$, then $\Delta p = \hbar \Delta k$ and

$$\Delta x \Delta p = \frac{\hbar}{2}. \quad (2.12)$$

¹More precisely, we can make such an expansion providing we meet some rather weak conditions of smoothness and differentiability of $\psi(x)$ – conditions met naturally by problems which derive from physical systems!

If we consider the width of the distribution as a measure of the "uncertainty", we will prove in section 3.3 that the Gaussian wave packet provides the minimum uncertainty. This result shows that we cannot know the position of a particle and its momentum at the same time. If we try to localise a particle to a very small region of space, its momentum becomes uncertain. If we try to make a particle with a definite momentum, its probability distribution spreads out over space.

With this introduction, we now turn to consider the interaction of a particle with a non-uniform potential background. For non-confining potentials, such systems fall into the class of **scattering problems**: For a beam of particles incident on a non-uniform potential, what fraction of the particles are transmitted and what fraction are reflected? In the one-dimensional system, the classical counterpart of this problem is trivial: For particle energies which exceed the maximum potential, all particles are eventually transmitted, while for energies which are lower, all particles are reflected. In quantum mechanics, the situation is richer: For a generic potential of finite extent and height, some particles are always reflected and some are always transmitted. Later in the course we will consider the general problem of scattering from a localised potential in arbitrary dimension. But for now, we will focus on the one-dimensional system, where many of the key concepts can be formulated.

2.1.3 Potential Step

As we have seen, for a time-independent potential, the wavefunction can be factorised as $\Psi(x, t) = e^{-iEt/\hbar}\psi(x)$, where $\psi(x)$ is obtained from the stationary form of the Schrödinger equation,

$$\left[-\frac{\hbar^2 \partial_x^2}{2m} + V(x) \right] \psi(x) = E\psi(x), \quad (2.13)$$

and E denotes the energy of the particle. As $|\Psi(x, t)|^2$ represents a probability density, it must be everywhere finite. As a result, we can deduce that the wavefunction, $\psi(x)$, is also finite. Moreover, since E and $V(x)$ are presumed finite, so must be $\partial_x^2\psi(x)$. The latter condition implies that

- both $\psi(x)$ and $\partial_x\psi(x)$ must be continuous functions of x , even if V has a discontinuity.

Consider then the influence of a potential step on the propagation of a beam of particles. Specifically, let us assume that a beam of particles with kinetic energy, E , moving from left to right are incident upon a potential step of height V_0 at position $x = 0$. If the beam has unit amplitude, the reflected and transmitted (complex) amplitudes are set by r and t . The corresponding wavefunction is given by

$$\begin{aligned} \psi_-(x) &= e^{ik_-x} + re^{-ik_-x} & x < 0, \\ \psi_+(x) &= te^{ik_+x} & x > 0, \end{aligned} \quad (2.14)$$

where $k_- = \frac{\sqrt{2mE}}{\hbar}$ and $k_+ = \frac{\sqrt{2m(E-V_0)}}{\hbar}$. Applying the continuity conditions on $\psi(x)$ and $\partial_x\psi(x)$ at the step ($x = 0$), one obtains the relations $1 + r = t$ and $ik_-(1 - r) = ik_+t$

leading to the reflection and transmission amplitudes,

$$r = \frac{k_- - k_+}{k_- + k_+}, \quad t = \frac{2k_-}{k_- + k_+} \quad (2.15)$$

The reflectivity, R , and transmittivity, T , are defined by the ratios,

$$R = \frac{\text{reflected flux}}{\text{incident flux}}, \quad T = \frac{\text{transmitted flux}}{\text{incident flux}}. \quad (2.16)$$

With the incident, reflected, and transmitted fluxes given by $|A|^{2\frac{\hbar k_-}{m}}$, $|Ar|^{2\frac{\hbar k_-}{m}}$, and $|At|^{2\frac{\hbar k_+}{m}}$ respectively, one obtains

$$R = \left| \frac{k_- - k_+}{k_- + k_+} \right|^2 = |r|^2, \quad T = \left| \frac{2k_-}{k_- + k_+} \right|^2 \frac{k_+}{k_-} = |t|^2 \frac{k_+}{k_-} = \frac{4k_- k_+}{(k_- + k_+)^2}. \quad (2.17)$$

From these results one can confirm that the total flux is, as expected, conserved in the scattering process, i.e. $R + T = 1$.

2.1.4 Potential Barrier

Having dealt with the potential step, we now turn to consider the problem of a beam of particles incident upon a square potential barrier of height V_0 (presumed positive for now) and width a . As mentioned above, this geometry is particularly important as it includes the simplest example of a scattering phenomenon in which a beam of particles is “deflected” by a local potential. Moreover, this one-dimensional geometry also provides a platform to explore a phenomenon peculiar to quantum mechanics – **quantum tunneling**. For these reasons, we will treat this problem fully and with some care.

Since the barrier is localised to a region of size a , the incident and transmitted wavefunctions have the same functional form, $e^{ik_1 x}$, where $k_1 = \frac{\sqrt{2mE}}{\hbar}$, and differ only in their complex amplitude, i.e. after the encounter with the barrier, the transmitted wavefunction undergoes only a change of amplitude (some particles are reflected from the barrier, even when the energy of the incident beam, E , is in excess of V_0) and a phase shift. To determine the relative change in amplitude and phase, we can parameterise the wavefunction as

$$\begin{aligned} \psi_1(x) &= e^{ik_1 x} + r e^{-ik_1 x} & x \leq 0 \\ \psi_2(x) &= A e^{ik_2 x} + B e^{-ik_2 x} & 0 \leq x \leq a \\ \psi_3(x) &= t e^{ik_1 x} & a \leq x \end{aligned} \quad (2.18)$$

where $k_2 = \frac{\sqrt{2m(E-V_0)}}{\hbar}$. Here, as with the step, r denotes the reflected amplitude and t the transmitted.

Applying the continuity conditions on the wavefunction, ψ , and its derivative, $\partial_x \psi$, at the barrier interfaces at $x = 0$ and $x = a$, one obtains

$$\begin{cases} 1 + r = A + B \\ A e^{ik_2 a} + B e^{-ik_2 a} = t e^{ik_1 a} \end{cases}, \quad \begin{cases} k_1(1 - r) = k_2(A - B) \\ k_2(A e^{ik_2 a} - B e^{-ik_2 a}) = k_1 t e^{ik_1 a} \end{cases}. \quad (2.19)$$

Together, these four equations specify the four unknowns, r , t , A and B . Solving, one obtains

$$t = \frac{2k_1 k_2 e^{-ik_1 a}}{2k_1 k_2 \cos(k_2 a) - i(k_1^2 + k_2^2)^2 \sin(k_2 a)}, \quad (2.20)$$

translating to a transmissivity of

$$T = |t|^2 = \frac{1}{1 + \frac{1}{4} \left(\frac{k_1}{k_2} - \frac{k_2}{k_1} \right)^2 \sin^2(k_2 a)}, \quad (2.21)$$

and the reflectivity, $R = 1 - T$. As a consistency check, we can see that, when $V_0 = 0$, $k_2 = k_1$ and $t = 1$, as expected. Moreover, T is restricted to the interval from 0 to 1 as required. So, for barrier heights in the range $E > V_0 > 0$, the transmittivity T shows an oscillatory behaviour with k_2 reaching unity when $k_2 a = n\pi$ with n integer. At these values, there is a conspiracy of interference effects which eliminate altogether the reflected component of the wave leading to perfect transmission. Such a situation arises when the width of the barrier is perfectly matched to an integer or half-integer number of wavelengths inside the barrier.

When the energy of the incident particles falls below the energy of the barrier, $0 < E < V_0$, a classical beam would be completely reflected. However, in the quantum system, particles are able to tunnel through the barrier region and escape leading to a non-zero transmission coefficient. In this regime, $k_2 = ik_2$ becomes pure imaginary leading to an evanescent decay of the wavefunction under the barrier and a suppression, but not extinction, of transmission probability,

$$T = |t|^2 = \frac{1}{1 + \frac{1}{4} \left(\frac{k_1}{\kappa_2} - \frac{\kappa_2}{k_1} \right)^2 \sinh^2(\kappa_2 a)}. \quad (2.22)$$

For $\kappa_2 a \gg 1$ (the weak tunneling limit), the transmittivity takes the form

$$T \simeq \frac{16k_1^2 \kappa_2^2}{(k_1^2 + \kappa_2^2)^2} e^{-2\kappa_2 a}. \quad (2.23)$$

Finally, on a cautionary note, while the phenomenon of quantum mechanical tunneling is well-established, it is difficult to access in a convincing experimental manner. Although a classical particle with energy $E < V_0$ is unable to penetrate the barrier region, in a physical setting, one is usually concerned with a thermal distribution of particles. In such cases, thermal activation may lead to transmission *over* a barrier. Such processes often overwhelm any contribution from true quantum mechanical tunneling.

2.1.5 The Rectangular Potential Well

Finally, if we consider scattering from a potential well (i.e. with $V_0 < 0$), while $E > 0$, we can apply the results of the previous section to find a continuum of unbound states with the corresponding **resonance** behaviour. However, in addition to these unbound states, for $E < 0$ we have the opportunity to find bound states of the potential. It is to this general problem that we now turn.

2.2 Wave Mechanics of Bound Particles

In the case of unbound particles, we have seen that the spectrum of states is continuous. However, for bound particles, the wavefunctions satisfying the Schrödinger equation have only particular quantised energies. In the one-dimensional system, we will find that all binding potentials are capable of hosting a bound state, a feature particular to the low dimensional system.

2.2.1 The Rectangular Potential Well (Continued)

As a starting point, let us consider a rectangular potential well similar to that discussed above. To make use of symmetry considerations, it is helpful to reposition the potential setting

$$V(x) = \begin{cases} 0 & x \leq -a \\ -V_0 & -a \leq x \leq a \\ 0 & a \leq x \end{cases}, \quad (2.24)$$

where the potential depth V_0 is assumed positive. In this case, we will look for bound state solutions with energies lying in the range $-V_0 < E < 0$. Since the Hamiltonian is invariant under **parity transformation**, $[\hat{H}, \hat{P}] = 0$ (where $\hat{P}\psi(x) = \psi(-x)$), the eigenstates of the Hamiltonian \hat{H} must also be eigenstates of parity, i.e. we expect the eigenfunctions to separate into those symmetric and those antisymmetric under parity.

For $E < 0$ (bound states), the wavefunction outside the well region must have the form

$$\psi(x < -a) = Ce^{\kappa x}, \quad \psi(x > a) = De^{-\kappa x}, \quad (2.25)$$

with $\kappa = \frac{\sqrt{-2mE}}{\hbar}$ while in the central well region, the general solution is of the form

$$\psi(-a < x < a) = A \cos(kx) + B \sin(kx), \quad (2.26)$$

where $k = \frac{\sqrt{2m(E+V_0)}}{\hbar}$. Once again we have four equations in four unknowns. The calculation shows that either A or B must be zero for a solution. This means that the states separate into solutions with even or odd parity.

For the even states, the solution of the equations leads to the quantisation condition, $\kappa = \tan kak$, while for the odd states, we find $\kappa = -\cot kak$. These are transcendental equations, and must be solved numerically. Figure 2.1 compares $\kappa a = \left(\frac{2mV_0a^2}{\hbar^2} - (ka)^2\right)^{1/2}$ with $ka \tan ka$ for the even states and to $-ka \cot ka$ for the odd states. Where the curves intersect, we have an allowed energy. From the structure of these equations, it is evident that an even state solution can always be found for arbitrarily small values of the binding potential V_0 while, for odd states, bound states appear only at a critical value of the coupling strength. The wider and deeper the well, the more solutions are generated.

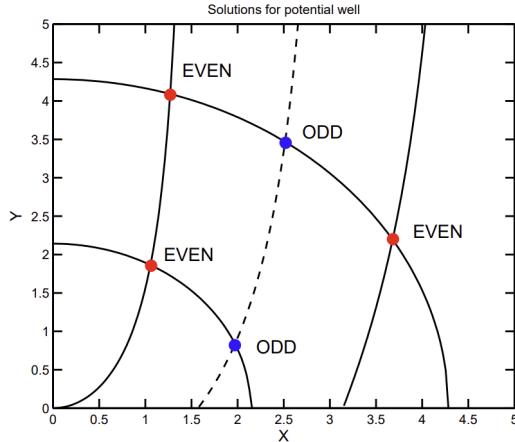


Fig. 2.1: Graphical solutions for a square potential well. The top three intersections correspond to a wide well, and the bottom two to a well having half of the width.

2.2.2 The δ -Function Potential Well

Let us now consider perhaps the simplest binding potential, the δ -function, $V(x) = -aV_0\delta(x)$. Here the parameter ‘ a ’ denotes some microscopic length scale introduced to make the product $a\delta(x)$ dimensionless.² For a state to be bound, its energy must be negative. Moreover, the form of the potential demands that the wavefunction is symmetric under parity, $x \rightarrow -x$. (A wavefunction which was antisymmetric must have $\psi(0) = 0$ and so could not be influenced by the δ -function potential.) We therefore look for a solution of the form

$$\psi(x) = A \begin{cases} e^{\kappa x} & x < 0 \\ e^{-\kappa x} & x > 0 \end{cases}, \quad (2.27)$$

where $\kappa = \frac{\sqrt{-2mE}}{\hbar}$. With this choice, the wavefunction remains everywhere continuous including at the potential, $x = 0$. Integrating the stationary form of the Schrödinger equation across an infinitesimal interval that spans the region of the δ -function potential, we find that

$$\partial_x \psi|_{+\epsilon} - \partial_x \psi|_{-\epsilon} = -\frac{2maV_0}{\hbar^2} \psi(0). \quad (2.28)$$

From this result, we obtain that $\kappa = maV_0/\hbar^2$, leading to the bound state energy

$$E = -\frac{ma^2V_0^2}{2\hbar^2}. \quad (2.29)$$

Indeed, the solution is unique. An attractive δ -function potential hosts only one bound state.

2.2.3 The δ -Function Model of a Crystal

Finally, as our last example of a one-dimensional quantum system, let us consider a particle moving in a periodic potential. The **Kronig-Penney model** provides a caricature of

²Note that the dimensions of $\delta(x)$ are [Length]⁻¹.

a (one-dimensional) crystalline lattice potential. The potential created by the ions is approximated as an infinite array of potential wells defined by a set of repulsive δ -function potentials,

$$V(x) = aV_0 \sum_{n=-\infty}^{\infty} \delta(x - na). \quad (2.30)$$

Since the potential is repulsive, it is evident that all states have energy $E > 0$. This potential has a new symmetry; a translation by the lattice spacing a leaves the potential unchanged, $V(x + a) = V(x)$. The probability density must therefore exhibit the same translational symmetry,

$$|\psi(x + a)|^2 = |\psi(x)|^2 \quad (2.31)$$

which means that, under translation, the wavefunction differs by at most a phase, $\psi(x + a) = e^{i\phi}\psi(x)$. In the region from $(n-1)a < x < na$, the general solution of the Schrödinger equation is plane wave like and can be written in the form,

$$\psi_n(x) = A_n \sin [k(x - na)] + B_n \cos [k(x - na)], \quad (2.32)$$

where $k = \frac{\sqrt{2mE}}{\hbar}$ and, following the constraint on translational invariance, $A_{n+1} = e^{i\phi}A_n$ and $B_{n+1} = e^{i\phi}B_n$. By applying the boundary conditions, one can derive a constraint on k similar to the quantised energies for bound states considered above.

Consider the boundary conditions at position $x = na$. Continuity of the wavefunction, $\psi_n|_{x=na} = \psi_{n+1}|_{x=na}$, translates to the condition, $B_n = A_{n+1} \sin(-ka) + B_{n+1} \cos(-ka)$ or

$$B_{n+1} = \frac{B_n + A_{n+1} \sin(ka)}{\cos(ka)}. \quad (2.33)$$

Similarly, the discontinuity in the first derivative, $\partial_x \psi_{n+1}|_{x=na} - \partial_x \psi_n|_{x=na} = \frac{2maV_0}{\hbar^2} \psi_n(na)$, leads to the condition $k[A_{n+1} \cos(ka) + B_{n+1} \sin(ka) - A_n] = \frac{2maV_0}{\hbar^2} B_n$. Substituting the expression for B_{n+1} and rearranging, one obtains

$$A_{n+1} = \frac{2maV_0}{\hbar^2 k} B_n \cos(ka) - B_n \sin(ka) + A_n \cos(ka). \quad (2.34)$$

Similarly, replacing the expression for A_{n+1} in that for B_{n+1} , one obtains the parallel equation,

$$B_{n+1} = \frac{2maV_0}{\hbar^2 k} B_n \sin(ka) + B_n \cos(ka) + A_n \sin(ka). \quad (2.35)$$

With these two equations, and the relations $A_{n+1} = e^{i\phi}A_n$ and $B_{n+1} = e^{i\phi}B_n$, we obtain the quantisation condition,³

$$\cos \phi = \cos(ka) + \frac{maV_0}{\hbar^2 k} \sin(ka).$$

(2.37)

As $\hbar k = \sqrt{2mE}$, this result relates the allowed values of energy to the real parameter, ϕ . Since $\cos \phi$ can only take values between -1 and 1 , there are a sequence of allowed bands of energy with energy gaps separating these bands.

³Eliminating A_n and B_n from the equations, a sequence of cancellations obtains

$$e^{2i\phi} + e^{i\phi} \left(\frac{2maV_0}{\hbar^2 k} \sin(ka) + 2 \cos(ka) \right) + 1 = 0. \quad (2.36)$$

Then multiplying by $e^{-i\phi}$, we obtain the expression for $\cos \phi$.

Such behaviour is characteristic of the spectrum of periodic lattices: In the periodic system, the wavefunctions – known as **Bloch states** – are indexed by a “quasi”- momentum index k , and a band index n where each Bloch band is separated by an energy gap within which there are no allowed states. In a **metal**, electrons (fermions) populate the energy states starting with the lowest energy up to some energy scale known as the **Fermi energy**. For a partially-filled band, low-lying excitations associated with the continuum of states allow electrons to be accelerated by a weak electric field. In a **band insulator**, all states are filled up to an energy gap. In this case, a small electric field is unable to excite electrons across the energy gap – hence the system remains insulating.

2.3 Wentzel, Kramers and Brillouin (WKB) Method

The WKB (or Wentzel, Kramers and Brillouin) approximation describes a “quasi-classical” method for solving the one-dimensional time-independent Schrödinger equation. Note that the consideration of one-dimensional systems is less restrictive than it may sound as many symmetrical higher-dimensional problems are rendered effectively one-dimensional (e.g. the radial equation for the hydrogen atom). The WKB method is named after physicists Wentzel, Kramers and Brillouin, who all developed the approach independently in 1926. Earlier, in 1923, the mathematician Harold Jeffreys had developed a general method of approximating the general class of linear, second-order differential equations, which of course includes the Schrödinger equation. But since the Schrödinger equation was developed two years later, and Wentzel, Kramers, and Brillouin were apparently unaware of this earlier work, the contribution of Jeffreys is often neglected.

The WKB method is important both as a practical means of approximating solutions to the Schrödinger equation, and also as a conceptual framework for understanding the classical limit of quantum mechanics. The WKB approximation is valid whenever the wavelength, λ , is small in comparison to other relevant length scales in the problem. This condition is not restricted to quantum mechanics, but rather can be applied to any wave-like system (such as fluids, electromagnetic waves, etc.), where it leads to approximation schemes which are mathematically very similar to the WKB method in quantum mechanics. For example, in optics the approach is called the **eikonal method**, and in general the method is referred to as **short wavelength asymptotics**. Whatever the name, the method is an old one, which predates quantum mechanics – indeed, it was apparently first used by Liouville and Green in the first half of the nineteenth century. In quantum mechanics, λ is interpreted as the de Broglie wavelength, and L is normally the length scale of the potential. Thus, the WKB method is valid if the wavefunction oscillates many times before the potential energy changes significantly

2.3.1 Semi-Classical Approximation to Leading Order

Consider then the propagation of a quantum particle in a slowly-varying one-dimensional potential, $V(x)$. Here, by “slowly-varying” we mean that, in any small region the wavefunction is well-approximated by a plane wave, and that the wavelength only changes over

distances that are long compared with the local value of the wavelength. We're also assuming for the moment that the particle has positive kinetic energy in the region. Under these conditions, we can anticipate that the solution to the time-independent Schrödinger equation

$$-\frac{\hbar^2}{2m}\partial_x^2\psi(x) + V(x)\psi(x) = E\psi(x), \quad (2.38)$$

will take the form $A(x)e^{\pm ip(x)x}$ where $p(x)$ is the “local” value of the momentum set by the *classical* value, $p^2/2m + V(x) = E$, and the amplitude, $A(x)$, is slowly-varying compared with the phase factor. Clearly this is a *semi-classical* limit: \hbar has to be sufficiently small that there are many oscillations in the typical distance over which the potential varies.⁴

To develop this idea more rigorously, and to emphasise the rapid *phase* variation in the semi-classical limit, we can parameterise the wavefunction as

$$\phi(x) = e^{i\sigma(x)/\hbar}. \quad (2.39)$$

Here the complex function $\sigma(x)$ encompasses both the amplitude and phase. Then, with $-\hbar^2\partial_x^2\psi(x) = -i\hbar e^{i\sigma(x)/\hbar}\partial_x^2\sigma(x) + e^{i\sigma(x)/\hbar}(\partial_x\sigma)^2$ and direct substitution into Eq. (2.38), the Schrödinger equation may be rewritten in terms of the phase function as

$$-i\hbar\partial_x^2\sigma(x) + (\partial_x\sigma)^2 = p^2(x). \quad (2.40)$$

Now, since we're assuming the system is semi-classical, it makes sense to expand $\sigma(x)$ as a power series in \hbar , setting

$$\sigma = \sigma_0 + (\hbar/i)\sigma_1 + (\hbar/i)^2\sigma_2 + \dots \quad (2.41)$$

At the leading (zeroth) order of the expansion, we can drop the first term in (2.40), leading to $(\partial_x\sigma_0)^2 = p^2(x)$. Defining $p(x) = +\sqrt{2m(E - V(x))}$, this equation permits the two solutions $\partial_x\sigma_0 = \pm p(x)$, from which we conclude

$$\sigma_0(x) = \pm \int^x p(x') \, dx'. \quad (2.42)$$

From the form of the Schrödinger equation (2.40), it is evident that this approximate solution is only valid if we can ignore the first term. More precisely, we must have

$$\left| \frac{\hbar\partial_x^2\sigma(x)}{(\partial_x\sigma(x))^2} \right| \equiv |\partial_x(\hbar/\partial_x\sigma)| \ll 1. \quad (2.43)$$

But, in the leading approximation, $\partial_x\sigma(x) \simeq p(x)$ and $p(x) = 2\pi\hbar/\lambda(x)$, so the condition translates to the relation

$$\frac{1}{2\pi}|\partial_x\lambda(x)| \ll 1. \quad (2.44)$$

⁴To avoid any point of confusion, it is of course true that \hbar is a fundamental constant – not easily adjusted! So what do we mean when we say that the semi-classical limit translates to $\hbar \rightarrow 0$? The validity of the semi-classical approximation relies upon $\lambda/L \ll 1$. Following the de Broglie relation, we may write this inequality as $h/pL \ll 1$, where p denotes the particle momentum. Now, in this correspondence, both p and L can be considered as “classical” scales. So, formally, we can think of accessing the semi-classical limit by adjusting \hbar so that it is small enough to fulfil this inequality. Alternatively, at fixed \hbar , we can access the semi-classical regime by reaching to higher and higher energy scales (larger and larger p) so that the inequality becomes valid.

This means that the change in wavelength over a distance of one wavelength must be small. This condition can not always be met: In particular, if the particle is confined by an attractive potential, at the edge of the classically allowed region, where $E = V(x)$, $p(x)$ must be zero and the corresponding wavelength infinite. The approximation is only valid well away from these **classical turning points**, a matter to which we will return shortly.

2.3.2 Next to Leading Order Correction

Let us now turn to the next term in the expansion in \hbar . Retaining terms from Eq. (2.40) which are of order \hbar , we have

$$-i\hbar\partial_x^2\sigma_0(x) + 2\partial_x\sigma_0(\hbar/i)\partial_x\sigma_1 = 0. \quad (2.45)$$

Rearranging this equation, and integrating, we find

$$\partial_x\sigma_1 = -\frac{\partial_x^2\sigma_0}{2\partial_x\sigma_0} = -\frac{\partial_x p}{2p}, \quad \sigma_1(x) = -\frac{1}{2}\ln p(x) + \text{const}, \quad (2.46)$$

where we have evaluated the derivatives using the previous result in Eq. (2.42). So, to this order of approximation, by substituting this result into Eq. (2.39) the wavefunction takes the form,

$$\psi(x) = \frac{C_1}{\sqrt{p(x)}} e^{(i/\hbar) \int^x p \, dx'} + \frac{C_2}{\sqrt{p(x)}} e^{(-i/\hbar) \int^x p \, dx'}, \quad (2.47)$$

where C_1 and C_2 denote constants of integration.

To interpret the factors of $\sqrt{p(x)}$, consider the first term: a wave moving to the right. Since $p(x)$ is real (remember we are currently considering the classically allowed region where $E > V(x)$), the exponential has modulus unity, and the local probability density is proportional to $1/p(x)$, i.e. to $1/v(x)$, where $v(x)$ denotes the velocity of the particle. This dependence has a simple physical interpretation: the probability of finding the particle in any given small interval is proportional to the time it spends there. Hence it is inversely proportional to its speed.

We turn now to consider the wavefunction in the classically forbidden region where

$$\frac{p^2}{2m} = E - V(x) < 0. \quad (2.48)$$

Here $p(x)$ is of course pure imaginary, but the same formal phase solution of the Schrödinger equation applies provided, again, that the particle is remote from the classical turning points where $E = V(x)$. In this region, the wavefunction takes the general form,

$$\psi(x) = \frac{C'_1}{\sqrt{|p(x)|}} e^{-(i/\hbar) \int^x |p| \, dx'} + \frac{C'_2}{\sqrt{|p(x)|}} e^{(i/\hbar) \int^x |p| \, dx'}. \quad (2.49)$$

This completes our study of the wavefunction in the regions in which the semiclassical approach can be formally justified. However, to make use of this approximation, we have to understand how to deal with the regions close to the classical turning points. Remember that in our treatment of the Schrödinger equation the energy quantisation derived from the implementation of boundary conditions.

2.3.3 Connection Formulae, Boundary Conditions and Quantisation Rules

Let us assume that we are dealing with a one-dimensional confining potential where the classically allowed region is unique and spans the interval $b \leq x \leq a$. Clearly, in the classically forbidden region to the right, $x > a$, only the first term in Eq. (2.49) remains convergent and can contribute, while for $x < b$ it is only the second term that contributes. Moreover, in the classically allowed region, $b \leq x \leq a$, the wavefunction has the oscillating form (2.47).

But how do we connect the three regions together? To answer this question, it is necessary to make the assumption that the potential varies sufficiently smoothly that it is a good approximation to take it to be linear in the vicinity of the classical turning points. That is to say, we assume that a linear potential is a sufficiently good approximation out to the point where the short wavelength (or decay length for tunneling regions) description is adequate. Therefore, near the classical turning at $x = a$, we take the potential to be

$$E - V(x) \simeq F_0(x - a), \quad (2.50)$$

where F_0 denotes the (constant) force. For a strictly linear potential, the wavefunction can be determined analytically, and takes the form of an Airy function. In particular, it is known that the Airy function to the right of the classical turning point has the asymptotic form

$$\lim_{x \gg a} \psi(x) = \frac{C}{2\sqrt{|p(x)|}} e^{-(i/\hbar) \int^x |p| dx'}, \quad (2.51)$$

translating to a decay into the classically forbidden region while, to the left, it has the asymptotic oscillatory solution,

$$\lim_{b \ll x < a} \psi(x) = \frac{C}{2\sqrt{|p(x)|}} \cos \left[\frac{1}{\hbar} \int_x^a p dx' - \frac{\pi}{4} \right] \equiv \frac{C}{2\sqrt{|p(x)|}} \cos \left[\frac{\pi}{4} - \frac{1}{\hbar} \int_x^a p dx' \right]. \quad (2.52)$$

At the second classical turning point at $x = b$, the same argument gives

$$\lim_{b < x \ll a} \psi(x) = \frac{C'}{2\sqrt{|p(x)|}} \cos \left[\frac{1}{\hbar} \int_b^x p dx' - \frac{\pi}{4} \right]. \quad (2.53)$$

For these two expressions to be consistent, we must have $C' = \pm C$ and

$$\left(\frac{1}{\hbar} \int_b^x p(x') dx' - \frac{\pi}{4} \right) - \left(\frac{\pi}{4} - \frac{1}{\hbar} \int_x^a p(x') dx' \right) = n\pi, \quad (2.54)$$

where, for n even, $C' = C$ and for n odd, $C' = -C$. Therefore, we have the condition $\frac{1}{\hbar} \int_b^a p(x) dx = (n + 1/2)\pi$, or when cast in terms of a complete periodic cycle of the classical motion,

$\oint p(x) dx = 2\pi\hbar(n + 1/2).$

(2.55)

This is just the **Bohr-Sommerfeld quantisation condition**, and n can be interpreted as the number of nodes of the wavefunction.

Note that the integrated action, $\oint p dx$, represents the area of the classical path in phase space. This shows that each state is associated with an element of phase space $2\pi\hbar$.

From this, we can deduce the approximate energy splitting between levels in the semi-classical limit: The change in the integral with energy ΔE corresponding to one level must be $2\pi\hbar$ - one more state and one more node, i.e. $\Delta E \partial_E \oint p dx = 2\pi\hbar$. Now $\partial_p E = v$, so $\oint \partial_E p dx = \oint dx/v = T$, the period of the orbit. Therefore, $\Delta E = 2\pi\hbar/T = \hbar\omega$: In the semi-classical limit, if a particle emits one photon and drops to the next level, the frequency of the photon emitted is just the orbital frequency of the particle.

For a particle strictly confined to one dimension, the connection formulae can be understood within a simple picture: The wavefunction “spills over” into the classically forbidden region, and its twisting there collects an $\pi/4$ of phase change. So, in the lowest state, the total phase change in the classically allowed region need only be $\pi/2$. For the radial equation, assuming that the potential is well behaved at the origin, the wavefunction goes to zero there. A bound state will still spill over beyond the classical turning point at r_0 , say, but clearly there must be a total phase change of $3\pi/4$ in the allowed region for the lowest state, since there can be no spill over to negative r . In this case, the general quantisation formula will be

$$\frac{1}{\hbar} \int_r^{r_0} p(r) dr = (n + 3/4)\pi, \quad n = 0, 1, 2, \dots, \quad (2.56)$$

with the series terminating if and when the turning point reaches infinity. In fact, some potentials, including the Coulomb potential and the centrifugal barrier for $\ell \neq 0$, are in fact singular at $r = 0$. These cases require special treatment.

2.3.4 Example: Simple Harmonic Oscillator

For the quantum harmonic oscillator, $H = \frac{p^2}{2m} + \frac{1}{2}m\omega^2x^2 = E$, the classical momentum is given by

$$p(x) = \sqrt{2m\left(E - \frac{m\omega^2x^2}{2}\right)}. \quad (2.57)$$

The classical turning points are set by $E = m\omega^2x_0^2/2$, i.e. $x_0 = \pm 2E/m\omega^2$. Over a periodic cycle, the classical action is given by

$$\oint p(x) dx = 2 \int_{-x_0}^{x_0} dx \sqrt{2m\left(E - \frac{m\omega^2x^2}{2}\right)} = 2\pi \frac{E}{\omega}. \quad (2.58)$$

According to the WKB method, the latter must be equated to $2\pi\hbar(n + 1/2)$, with the last term reflecting the two turning points. As a result, we find that the energy levels are as expected specified by $E_n = (n + 1/2)\hbar\omega$.

In the WKB approximation, the corresponding wavefunctions are given by

$$\psi(x) = \frac{C}{\sqrt{p(x)}} \cos \left(\frac{1}{\hbar} \int_{-x_0}^{x_0} p(x') dx' - \frac{\pi}{4} \right) \quad (2.59)$$

$$= \frac{C}{\sqrt{p(x)}} \cos \left(\frac{2\pi}{4} (n + 1/2) \frac{1}{\hbar} \int_0^{x_0} p(x') dx' - \frac{\pi}{4} \right) \quad (2.60)$$

$$= \frac{C}{\sqrt{p(x)}} \cos \left(\frac{n\pi}{2} + \frac{E}{\hbar\omega} \left[\arcsin \left(\frac{x}{x_0} \right) + \frac{x}{x_0} \sqrt{1 - \frac{x^2}{x_0^2}} \right] \right), \quad (2.61)$$

for $0 < x < x_0$ and

$$\psi(x) = \frac{C}{2\sqrt{p(x)}} \exp\left(-\frac{E}{\hbar\omega} \left[\frac{x}{x_0} \sqrt{\frac{x^2}{x_0^2} - 1} - \text{arcosh}\left(\frac{x}{x_0}\right) \right]\right) \quad (2.62)$$

for $x > a$. Note that the failure of the WKB approximation is reflected in the appearance of discontinuities in the wavefunction at the classical turning points. Nevertheless, the wavefunction at high energies provides a strikingly good approximation to the exact wavefunction.

2.3.5 Example: Quantum Tunneling

Consider the problem of quantum tunneling. Suppose that a beam of particles is incident upon a localised potential barrier, $V(x)$. Further, let us assume that, over a single continuous region of space, from b to a , the potential rises above the incident energy of the incoming particles so that, classically, all particles would be reflected. In the quantum system, the some particles incident from the left may tunnel through the barrier and continue propagating to the right. We are interested in finding the transmission probability.

From the WKB solution, to the left of the barrier (region 1), we expect a wavefunction of the form

$$\psi_1(x) = \frac{1}{\sqrt{p}} \exp\left[\frac{i}{\hbar} \int_b^x p \, dx'\right] + r(E) \frac{1}{\sqrt{p}} \exp\left[-\frac{i}{\hbar} \int_b^x p \, dx'\right], \quad (2.63)$$

with $p(x) = \sqrt{2m(E - V(x))}$, while, to the right of the barrier (region 3), the wavefunction is given by

$$\psi_3(x) = t(E) \frac{1}{\sqrt{p}} \exp\left[\frac{i}{\hbar} \int_a^x p \, dx'\right]. \quad (2.64)$$

In the barrier region, the wavefunction is given by

$$\psi_2(x) = \frac{C_1}{\sqrt{|p(x)|}} \exp\left[-\frac{i}{\hbar} \int_a^x |p| \, dx'\right] + \frac{C_2}{\sqrt{|p(x)|}} \exp\left[\frac{i}{\hbar} \int_a^x |p| \, dx'\right]. \quad (2.65)$$

Applying the matching conditions on the wavefunction at the classical turning points, one obtains the transmissivity,

$$T(E) \simeq \exp\left[-\frac{2}{\hbar} \int_a^b |p| \, dx\right]. \quad (2.66)$$

CHAPTER 3

Operator Methods and Formalism in Quantum Mechanics

In the previous chapters we made good headway in building some intuition for simple quantum systems by solving the Schrödinger equation and trying to interpret the results. But, at times, it felt like we were making up the rules as we went along. Our goal in this section is bring some order to the table and describe the mathematical structure that underlies quantum mechanics.

While the wave mechanical formulation has proved successful in describing the quantum mechanics of bound and unbound particles, some properties can not be represented through a wave-like description. For example, the electron spin degree of freedom does not translate to the action of a gradient operator. It is therefore useful to reformulate quantum mechanics in a framework that involves only operators.

Before discussing properties of operators, it is helpful to introduce a further simplification of notation. One advantage of the operator algebra is that it does not rely upon a particular basis. For example, when one writes $\hat{H} = \frac{\hat{p}^2}{2m}$, where the hat denotes an operator, we can equally represent the momentum operator in the spatial coordinate basis, when it is described by the differential operator, $\hat{p} = -i\hbar\partial_x$, or in the momentum basis, when it is just a number $\hat{p} = p$. Similarly, it would be useful to work with a basis for the wavefunction which is coordinate independent. Such a representation was developed by Dirac early in the formulation of quantum mechanics.

In the parlors of mathematics, square integrable functions (such as wavefunctions) are said form a vector space, much like the familiar three-dimensional vector spaces. In the **Dirac notation**, a state vector or wavefunction, ψ , is represented as a “ket”, $|\psi\rangle$. Just as we can express any three-dimensional vector in terms of the basis vectors, $\mathbf{r} = x\hat{\mathbf{e}}_1 + y\hat{\mathbf{e}}_2 + z\hat{\mathbf{e}}_3$, so we can expand any wavefunction as a superposition of basis state vectors,

$$|\psi\rangle = \lambda_1 |\psi_1\rangle + \lambda_2 |\psi_2\rangle + \dots \quad (3.1)$$

Alongside the ket, we can define the “bra”, $\langle\psi|$. Together, the bra and ket define the **scalar product**

$$\langle\phi|\psi\rangle \equiv \int_{-\infty}^{\infty} dx \phi^*(x)\psi(x), \quad (3.2)$$

from which follows the identity, $\langle\phi|\psi\rangle^* = \langle\psi|\phi\rangle$. In this formulation, the real space representation of the wavefunction is recovered from the inner product $\phi(x) = \langle x|\phi\rangle$ while the momentum space wavefunction is obtained from $\phi(p) = \langle p|\phi\rangle$. As with a three-dimensional vector space where $\mathbf{a} \cdot \mathbf{b} \leq |a||b|$, the magnitude of the scalar product is limited by the magnitude of the vectors,

$$\langle\psi|\phi\rangle \leq \sqrt{\langle\psi|\psi\rangle \langle\phi|\phi\rangle}, \quad (3.3)$$

a relation known as the **Schwarz inequality**.

There are four facets of quantum mechanics. These are

- States;
- Observables;
- Time Evolution;
- Measurement.

We've already said all there is to say about time evolution, at least for now: the evolution of the wavefunction is governed by the time-dependent Schrödinger equation. Here we will elaborate on the other three.

3.1 States

Recall the basic tenet: a quantum state is described by a normalisable wavefunction $\psi(\mathbf{x}, t)$, where normalisable means that it obeys

$$\int d^3x |\psi|^2 < \infty. \quad (3.4)$$

Until now, it largely appears that quantum mechanics is, like many other areas of physics, all about solving differential equations. That's not entirely wrong, but it misses the key fact: at heart, quantum mechanics is really a theory of linear algebra. This will become apparent as we unveil the bigger picture.

We've already seen a part of this linear algebra aspect in the principle of superposition. If $\psi(\mathbf{x}, t)$ and $\phi(\mathbf{x}, t)$ are both viable states of a system, then so too is any linear combination $\alpha\psi(\mathbf{x}, t) + \beta\phi(\mathbf{x}, t)$ where $\alpha, \beta \in \mathbb{C}$. Mathematically, this is the statement that the states form a *vector space* over the complex numbers.

3.1.1 The Inner Product

There is one important structure on our vector space: an *inner product*. This is a map that takes two vectors and spits out a number.

We're very used to inner products for finite dimensional vector spaces. Given two vectors in \mathbb{R}^N - call them \vec{u} and \vec{v} - the inner product is simply $\vec{u} \cdot \vec{v}$. We introduce new notation and write the inner product on \mathbb{R}^N as

$$\langle \vec{u} | \vec{v} \rangle = \vec{u} \cdot \vec{v}. \quad (3.5)$$

A slight amendment to the inner product is needed if we're dealing with a complex vector space like \mathbb{C}^N . In this case, the inner product is

$$\langle \vec{u} | \vec{v} \rangle = \vec{u}^* \cdot \vec{v}, \quad (3.6)$$

where we take the complex conjugation of the first vector¹ before we take the dot product. This has the advantage that the inner product of any vector with itself is necessarily non-negative

$$\langle \vec{u} | \vec{u} \rangle = \vec{u}^* \cdot \vec{u} \geq 0. \quad (3.7)$$

This means that we get to use the inner product to define the length of a vector $|\vec{u}| = \sqrt{\langle \vec{u} | \vec{u} \rangle}$. Inner products with this property are said to be *positive definite*.

This now brings us to the infinite dimensional vector space of functions that we work with in quantum mechanics. The inner product between two complex functions $\psi(\mathbf{x})$ and $\phi(\mathbf{x})$, each functions over \mathbb{R}^3 , is defined by

$$\langle \psi | \phi \rangle = \int d^3x \psi^*(\mathbf{x}) \phi(\mathbf{x}). \quad (3.8)$$

Note, again, the star in the first argument which ensures that the inner product of a function with itself is positive definite. In fact, our normalisability requirement (3.4) can be restated in terms of the inner product

$$\langle \psi | \psi \rangle < \infty. \quad (3.9)$$

From the definition, we see that the inner product is linear in the second argument

$$\langle \psi | \alpha \phi_1 + \beta \phi_2 \rangle = \alpha \langle \psi | \phi_1 \rangle + \beta \langle \psi | \phi_2 \rangle, \quad (3.10)$$

for any $\alpha, \beta \in \mathbb{C}$. However, it is *anti-linear* in the first argument, which simply means that the complex conjugation on ψ gives us

$$\langle \alpha \psi_1 + \beta \psi_2 | \phi \rangle = \alpha^* \langle \psi_1 | \phi \rangle + \beta^* \langle \psi_2 | \phi \rangle. \quad (3.11)$$

Alternatively, this anti-linear statement follows by noticing that complex conjugation exchanges the two entries in the inner product

$$\langle \psi | \phi \rangle = \langle \phi | \psi \rangle^*. \quad (3.12)$$

We'll make use of some of these simple mathematical identities as we proceed.

In more fancy mathematical language, a vector space, whether finite or infinite dimensional, with a positive-definite inner product is called a **Hilbert space** \mathcal{H} .²

Hilbert space is the stage for quantum mechanics. The first thing you should do when describing any quantum system is specify its Hilbert space. If a particle is moving on some space M then the relevant Hilbert space is called $\mathcal{H} = L^2(M)$, which is the space of square-normalisable functions on M . Here the L is for “Lebesgue”, while the word “square” and the superscript 2 both reflect the fact that you should integrate $|\psi|^2$ to determine whether it's normalisable.

¹In the maths literature, the inner product is often taken to be linear in the left entry and antilinear in the right. We'll follow the QM literature, which always uses the opposite convention.

²In fact this statement is only almost true: there is an extra requirement in the case of an infinite dimensional vector space that the space is “complete” which captures the heuristic idea that there are no points “missing” from \mathcal{H} , roughly speaking, this means that as your vector evolves it can't suddenly exit the space. Importantly, it also allows us to do analysis: being able to differentiate vectors in \mathcal{H} requires that we can take limits, and to do this we need to know whether the limits exist.

As our theories get more advanced and abstract, we need different Hilbert spaces. Sometimes these are simpler: for example, there is a lot of interesting physics lurking in the finite dimensional Hilbert space $\mathcal{H} = \mathbb{C}^2$ where a state is just a two-dimensional complex vector. But sometimes the Hilbert spaces are vastly more complicated, like the space in quantum field theory where M itself is an infinite dimensional space of functions and $L^2(M)$ is something ghastly and poorly understood.

3.1.2 Dual Space

As with any vector space, the *dual* \mathcal{H}^* of a Hilbert space \mathcal{H} is the space of linear maps $\mathcal{H} \rightarrow \mathbb{C}$. That is, an element $\varphi \in \mathcal{H}^*$ defines a map $\varphi : \psi \mapsto \varphi(\psi) \in \mathbb{C}$ for every $\psi \in \mathcal{H}$, such that

$$\varphi : a\psi_1 + b\psi_2 \mapsto a\varphi(\psi_1) + b\varphi(\psi_2), \quad (3.13)$$

for all $\psi_1, \psi_2 \in \mathcal{H}$ and $a, b \in \mathbb{C}$.

One way to construct such a map is to use the inner product: given some state $\phi \in \mathcal{H}$ we can define an element $(\phi, \cdot) \in \mathcal{H}^*$ which acts by

$$(\phi, \cdot) : \psi \mapsto (\phi, \psi), \quad (3.14)$$

i.e. we take the inner product of $\psi \in \mathcal{H}$ with our chosen element ϕ . The linearity properties of the inner product transfer to ensure that (ϕ, \cdot) is indeed a *linear* map - note that since the inner product is antilinear in its first entry, it's important that our chosen element ϕ sits in this first entry. In finite dimensions, every element of \mathcal{H}^* arises this way. That is, *any* linear map $\varphi : \mathcal{H} \rightarrow \mathbb{C}$ can be written as (ϕ, \cdot) for some fixed choice of $\phi \in \mathcal{H}$. This means in particular that the inner product (\cdot, \cdot) provides a vector space isomorphism

$$\mathcal{H}^* \xrightarrow{(\cdot, \cdot)} \mathcal{H}. \quad (3.15)$$

Comfortingly, the same result also holds in infinite dimensions. The isomorphism $\mathcal{H}^* \cong \mathcal{H}$ is what is special about Hilbert spaces among various other infinite dimensional vector spaces, and makes them especially easy to handle.

3.1.3 Dirac Notation and Continuum States

We'll use a notation for Hilbert spaces that was introduced by Dirac and is standard throughout the theoretical physics literature. Dirac denotes an element of \mathcal{H} as $|\psi\rangle$, where the symbol “ $| \rangle$ ” is known as a *ket*. An element of the dual space is written $\langle \phi |$ and the symbol “ $\langle |$ ” is called a *bra*. The relation between the ket $|\phi\rangle \in \mathcal{H}$ and the bra $\langle \phi | \in \mathcal{H}^*$ is what we would previously have written as ϕ vs (ϕ, \cdot) . The inner product between two states $|\psi\rangle, |\phi\rangle \in \mathcal{H}$ is then written $\langle \phi | \psi \rangle$ forming a *bra-ket* or bracket. Note that this implicitly uses the isomorphism $\mathcal{H}^* \cong \mathcal{H}$ provided by the inner product, building it into the notation. Recall also that (at least in all physics courses!) $\langle \phi | \psi \rangle$ is *antilinear* in $|\phi\rangle$.

Given an orthonormal basis $\{|e_a\rangle\}$ of \mathcal{H} , at least for the cases $\mathcal{H} \cong \mathbb{C}^n$ or $\mathcal{H} \cong \ell^2$,³ in Dirac notation we can expand a general ket $|\psi\rangle$ as

$$|\psi\rangle = \sum_a \psi_a |e_a\rangle, \quad (3.17)$$

in terms of this basis. Then $\langle \chi | \psi \rangle = \sum_{a,b} \chi_b^* \psi_a \langle e_b | e_a \rangle = \sum_a \chi_a^* \psi_a$ as usual.

One place where Dirac notation is particularly convenient is that it allows us to label states by their eigenvalues. We let $|q\rangle$ denote an eigenstate of some operator Q with eigenvalue q , so that $Q|q\rangle = q|q\rangle$. For example, a state that is an eigenstate of the position operator \mathbf{X} with position $\mathbf{x} \in \mathbb{R}^3$ (representing a particle that is definitely located at \mathbf{x}) is written $|\mathbf{x}\rangle$, while the state $|\mathbf{p}\rangle$ is an eigenstate of the momentum operator \mathbf{P} with eigenvalue \mathbf{p} . There's a potential confusion here: $|\mathbf{x}\rangle$ does *not* refer to the function \mathbf{x} , but rather a state in \mathcal{H} whose only support is at \mathbf{x} . The wavefunction corresponding to $|\mathbf{x}\rangle$ was really a δ -function.

In infinite-dimensional Hilbert spaces, one may not be lucky enough to have valid eigenvectors of operators. As an example, an eigenvector of the position operator on $L^2([0, 1])$ is $|x\rangle$ which is not normalisable, and so cannot belong to the Hilbert space.

However, we can look for the next best thing, which is an approximate eigenvector. In the context of our example, we may look for a subspace of vectors supported on a tiny interval $[x, x+\varepsilon] \subset [0, 1]$ such that $X|x\rangle \approx x|x\rangle$. For bounded observables, we may collect such subspaces into a projection-valued measure $\{E_x\}$.

In this case, we introduce a *continuum basis* with elements $|a\rangle$ labelled by a continuous variable a , normalised so that

$$\langle a' | a \rangle = \delta(a' - a), \quad (3.18)$$

using the Dirac δ -function. Then we write

$$|\psi\rangle = \int \psi(a) |a\rangle da \quad (3.19)$$

to expand a general $|\psi\rangle$ in terms of the $|a\rangle$'s. The point of the normalisation (3.18) is that

$$\langle \chi | \psi \rangle = \int \chi(b)^* \psi(a) \langle b | a \rangle db da = \int \chi(b)^* \psi(a) \delta(b - a) db da = \int \chi(a)^* \psi(a) da, \quad (3.20)$$

which is just the inner product (and also norm) we gave $L^2(\mathbb{R}, da)$ before. Indeed, a key example of a continuum basis is the *position* basis $\{|x\rangle\}$, where $x \in \mathbb{R}$. Expanding a general state $|\psi\rangle$ as an integral

$$|\psi\rangle = \int_{\mathbb{R}} \psi(x') |x'\rangle dx', \quad (3.21)$$

we see that the complex coefficients are

$$\langle x | \psi \rangle = \int_{\mathbb{R}} \psi(x') \langle x | x' \rangle dx' = \psi(x). \quad (3.22)$$

³The space known as ℓ^2 is the space of infinite sequences of complex numbers $\mathbf{u} = (u_1, u_2, u_3, \dots)$ such that

$$\sum_{i=1}^{\infty} |u_i|^2 < \infty, \quad (3.16)$$

heuristically, you can think of it as “ \mathbb{C} with $n = \infty$ ”

In other words, the position space wavefunctions we're familiar with are nothing but the coefficients of a state $|\psi\rangle \in \mathcal{H}$ in a particular (position) continuum basis.

As always, we could equally choose to expand this same vector in any number of different bases. For example, our state $|\psi\rangle = \int_{\mathbb{R}} \psi(x) |x\rangle dx$ from above can equally be expanded in the momentum basis as $|\psi\rangle = \int_{\mathbb{R}} \tilde{\psi}(p) |p\rangle dp$, where the new coefficients $\tilde{\psi}(p) = \langle p|\psi\rangle$ are the *momentum space wavefunction*. Later, we'll show that $\langle x|p\rangle = e^{ixp/\hbar}/\sqrt{2\pi\hbar}$, so these two sets of coefficients are related by

$$\langle x|\psi\rangle = \int_{\mathbb{R}} \tilde{\psi}(p) \langle x|p\rangle dp = \frac{1}{\sqrt{2\pi\hbar}} \int_{\mathbb{R}} e^{ipx/\hbar} \tilde{\psi}(p) dp \quad (3.23)$$

$$\langle p|\psi\rangle = \int_{\mathbb{R}} \tilde{\psi}(p) \langle p|x\rangle dx = \frac{1}{\sqrt{2\pi\hbar}} \int_{\mathbb{R}} e^{-ipx/\hbar} \psi(x) dx \quad (3.24)$$

(In the first line here, we expanded $|\psi\rangle$ in the momentum basis, then took the inner product with $|x\rangle$, while in the second we expanded the same $|\psi\rangle$ in the position basis and took the inner product with $|p\rangle$.) Equation (3.24) is just the statement that the position and momentum space wavefunctions are each other's Fourier transforms.

The real point I wish to make is that the fundamental object is the abstract vector $|\psi\rangle \in \mathcal{H}$. All the physical information about a quantum system is encoded in its state vector $|\psi\rangle$; the wavefunctions $\psi(x)$ or $\tilde{\psi}(p)$ are merely the expansion coefficients in some basis. Like any other choice of basis, this expansion may be useful for some purposes and unhelpful for others.

Finally, a technical point. Although using “continuum bases” such as $\{|x\rangle\}$ or $\{|p\rangle\}$ is convenient, because it emphasises the similarities between the infinite and finite-dimensional cases, it blurs the analysis. In particular, if $\langle x'|x\rangle = \delta(x' - x)$ then the norm

$$||x\rangle|^2 = \delta(x - x) = \delta(0), \quad ??? \quad (3.25)$$

so, whatever these objects $|x\rangle$ are, they certainly do not lie in our Hilbert space. It is possible to make good mathematical sense of these by appropriately enlarging our spaces to include spaces of distributions, but for the most part (and certainly in this course) physicists are content to say that continuum states such as $|x\rangle$ are allowable as basis elements, but call them *non-normalisable states*: actual physical particles are never represented by a non-normalisable state.

3.2 Operators

The quantum state is described by a wavefunction. But how does this wavefunction encode the information contained in the state? In quantum mechanics, all observables are represented by **operators** on the Hilbert space. In the present case, this is a mathematical object \hat{A} that acts on one state vector, $|\psi\rangle$, giving you back another, $|\phi\rangle$, i.e. $\hat{A}|\psi\rangle = |\phi\rangle$. If

$$\boxed{\hat{A}|\psi\rangle = a|\psi\rangle}, \quad (3.26)$$

then $|\phi\rangle$ is said to be an **eigenstate** (or **eigenfunction**) of \hat{A} with eigenvalue a . For example, the plane wave state $\psi_x(x) = \langle x|\psi_p\rangle = Ae^{ipx/\hbar}$ is an eigenstate of the **momentum**

operator, $\hat{p} = -i\hbar\partial_x$, with eigenvalue p . For a free particle, the plane wave is also an eigenstate of the Hamiltonian, $\hat{H} = \frac{\hat{p}^2}{2m}$ with eigenvalue $\frac{p^2}{2m}$.

To be more precise, observables in quantum mechanics are related to a particular kind of *linear operator*. Here “linear” means that operators obey

$$\hat{A}[\alpha|\psi_1\rangle + |\psi_2\rangle] = \alpha\hat{A}|\psi_1\rangle + \hat{A}|\psi_2\rangle, \quad (3.27)$$

for any $\alpha \in \mathbb{C}$.

Returning to our analogy with finite dimensional vector spaces: if you’re dealing with N -dimensional vectors, then the corresponding linear operators are just $N \times N$ matrices. Now, however, we’re dealing with infinite dimensional vector spaces – or, more precisely, Hilbert spaces - where the vectors themselves are functions. What replaces the matrices? The answer, as we will now see, is differential operators.

In quantum mechanics, for any observable A , there is an operator \hat{A} which acts on the wavefunction so that, if a system is in a state described by $|\psi\rangle$, the expectation value of A is

$$\langle A \rangle = \langle \psi | \hat{A} | \psi \rangle = \int_{-\infty}^{\infty} dx \phi^*(x) \hat{A} \psi(x). \quad (3.28)$$

Every operator corresponding to an observable is both linear and Hermitian: That is, for any two wavefunctions $|\phi\rangle$ and $|\psi\rangle$, and any two complex numbers α and β , **linearity** implies that

$$\hat{A}(\alpha|\psi\rangle + \beta|\phi\rangle) = \alpha(\hat{A}|\psi\rangle) + \beta(\hat{A}|\phi\rangle). \quad (3.29)$$

Moreover, for any linear operator \hat{A} , the **Hermitian conjugate** operator (also known as the adjoint) is defined by the relation

$$\langle \phi | \hat{A} \psi \rangle = \int dx \phi^*(\hat{A}\psi) = \int dx \psi (\hat{A}^\dagger \phi)^* = \langle \hat{A}^\dagger \phi | \psi \rangle. \quad (3.30)$$

From the definition, $\langle \hat{A}^\dagger \phi | \psi \rangle = \langle \phi | \hat{A} \psi \rangle$, we can prove some useful relations: Taking the complex conjugate, $\langle \hat{A}^\dagger \phi | \psi \rangle^* = \langle \psi | \hat{A}^\dagger \phi \rangle = \langle \hat{A} \psi | \phi \rangle$, and then finding the Hermitian conjugate of \hat{A}^\dagger , we have

$$\langle \psi | \hat{A}^\dagger \phi \rangle = \langle (\hat{A}^\dagger)^\dagger \psi | \phi \rangle = \langle \hat{A} \psi | \phi \rangle, \quad \text{i.e. } (\hat{A}^\dagger)^\dagger = \hat{A}. \quad (3.31)$$

Therefore, if we take the Hermitian conjugate twice, we get back to the same operator. It’s easy to show that $(\lambda\hat{A})^\dagger = \lambda^*\hat{A}^\dagger$ and $(\hat{A} + \hat{B})^\dagger = \hat{A}^\dagger + \hat{B}^\dagger$ just from the properties of the dot product. We can also show that $(\hat{A}\hat{B})^\dagger = \hat{B}^\dagger\hat{A}^\dagger$ from the identity, $\langle \phi | \hat{A}\hat{B}\psi \rangle = \langle \hat{A}^\dagger\phi | \hat{B}\psi \rangle = \langle \hat{B}^\dagger\hat{A}^\dagger\phi | \psi \rangle$. Note that operators are **associative** but not (in general) **commutative**,

$$\hat{A}\hat{B}|\psi\rangle = \hat{A}(\hat{B}|\psi\rangle) = (\hat{A}\hat{B})|\psi\rangle \neq \hat{B}\hat{A}|\psi\rangle. \quad (3.32)$$

It is helpful to define the **commutator** of two operators by

$$[\hat{A}, \hat{B}] \equiv \hat{A}\hat{B} - \hat{B}\hat{A}. \quad (3.33)$$

This commutator obeys the following properties:

$$\text{antisymmetry} \quad [\hat{A}, \hat{B}] = -[\hat{B}, \hat{A}] \quad (3.34)$$

$$\text{linearity} \quad [\alpha_1 \hat{A}_1 + \alpha_2 \hat{A}_2, \hat{B}] = \alpha_1 [\hat{A}_1, \hat{B}] + \alpha_2 [\hat{A}_2, \hat{B}] \quad \forall \alpha_1, \alpha_2 \in \mathbb{C} \quad (3.35)$$

$$\text{Leibniz identity} \quad [\hat{A}, \hat{B}\hat{C}] = [\hat{A}, \hat{B}]\hat{C} + \hat{B}[\hat{A}, \hat{C}] \quad (3.36)$$

$$\text{Jacobi identity} \quad [\hat{A}, [\hat{B}, \hat{C}]] + [\hat{B}, [\hat{C}, \hat{A}]] + [\hat{C}, [\hat{A}, \hat{B}]] = 0. \quad (3.37)$$

Each of these may easily be verified from the definition (3.33).

A physical variable must have real expectation values (and eigenvalues). This implies that the operators representing physical variables have some special properties. By computing the complex conjugate of the expectation value of a physical variable, we can easily show that physical operators are their own Hermitian conjugate,

$$\langle \psi | \hat{H} | \psi \rangle^* = \left[\int_{-\infty}^{\infty} \psi^*(x) \hat{H} \psi(x) dx \right]^* = \int_{-\infty}^{\infty} \psi(x) (\hat{H} \psi(x))^* dx = \langle \hat{H} \psi | \psi \rangle. \quad (3.38)$$

i.e. $\langle \hat{H} \psi | \psi \rangle = \langle \psi | \hat{H} \psi \rangle = \langle \hat{H}^\dagger \psi | \psi \rangle$, and $\hat{H}^\dagger = \hat{H}$. Operators that are their own Hermitian conjugate are called **Hermitian** (or self-adjoint).

3.2.1 Matrix Representations of Operators

Eigenfunctions of Hermitian operators $\hat{H} |i\rangle = E_i |i\rangle$ form an orthonormal (i.e. $\langle i | j \rangle = \delta_{ij}$) complete basis. For any complete basis $\{|i\rangle\}$, we can expand a state $|\psi\rangle$ as

$$|\psi\rangle = \sum_i \psi_i |i\rangle \quad (3.39)$$

This is a “representation” of $|\psi\rangle$ in this particular basis: e.g. if $\{|i\rangle\}$ are energy eigenstates, then we say that ψ_i is the “energy-representation” of $|\psi\rangle$.

To obtain an explicit form for ψ_i , we note that we can for any orthonormal complete set of states $\{|i\rangle\}$, we can always write

$$|\psi\rangle = \sum_i |i\rangle \langle i | \psi \rangle. \quad (3.40)$$

Thus, we see that the complex numbers $\langle i | \psi \rangle$ are the components of the state vector in the basis states $\{|i\rangle\}$, that is $\psi_i \equiv \langle i | \psi \rangle$.

A ‘ket’ state vector followed by a ‘bra’ state vector is an example of an operator. The **projection operator** onto a state $|j\rangle$ is given by $|j\rangle \langle j|$. First the bra vector dots into the state, giving the coefficient of $|j\rangle$ in the state, then it is multiplied by the unit vector $|j\rangle$, turning it back into a vector, with the right length to be a projection. An operator maps one vector into another vector, so this is an operator. If we sum over a complete orthonormal set of states, like the eigenstates of a Hermitian operator, we obtain the (useful) **resolution of the identity**

$$\hat{\mathbb{I}} = \sum_i |i\rangle \langle i|. \quad (3.41)$$

Using the resolution of the identity, it is possible to express any operator as

$$\hat{A} = \hat{\mathbb{I}}\hat{A}\hat{\mathbb{I}} = \sum_i \sum_j |j\rangle \langle j| \hat{A} |i\rangle \langle i|. \quad (3.42)$$

The complex numbers $A_{ji} = \langle j|A|i\rangle$ are the “matrix representation” of \hat{A} in this basis,

$$\hat{A} = \sum_{ij} A_{ji} |j\rangle \langle i|. \quad (3.43)$$

Consider the action of \hat{A} on a state $|\psi\rangle$

$$|\phi\rangle \equiv \hat{A} |\psi\rangle \quad (3.44)$$

$$= \sum_i \sum_j |j\rangle \langle j| \hat{A} |i\rangle \langle i| \psi = \sum_{i,j} |j\rangle A_{ji} \psi_i \quad (3.45)$$

The representation of $|\psi\rangle$ in this basis is

$$\phi_i \equiv \langle i|\phi\rangle = \langle i| \left(\sum_{k,j} |j\rangle A_{jk} \psi_k \right) = \sum_{k,j} \delta_{ij} A_{jk} \psi_k \quad (3.46)$$

$$= \sum_k A_{ik} \phi_k \quad (3.47)$$

Thus the action of the operator \hat{A} is via matrix multiplication of A_{ik} on the components of the vector ψ_k . Any complete orthonormal set of vectors can be used as the basis. Each gives rise to a different (vector) representation of the states and (matrix) representation of the operators.

In the above discussion, we have assumed that the basis $\{|i\rangle\}$ is discrete (i.e. we can sum over the projectors $|i\rangle\langle i|$). How do we interpret these sums when the basis is described by a continuous variable? For example, we could consider the eigenstates $\{|x\rangle\}$ of the position operator \hat{x} , or the eigenstates $\{|p\rangle\}$ of the momentum operator \hat{p} .

For a continuous set of states, the resolution of the identity becomes

$$\hat{\mathbb{I}} = \int dx |x\rangle \langle x|. \quad (3.48)$$

and the orthonormality condition becomes

$$\langle x'|x\rangle = \delta(x - x'), \quad (3.49)$$

where $\delta(x)$ is the **Dirac delta function**. Thus, for a state $|\psi\rangle$ we can write

$$|\psi\rangle = \int dx |x\rangle \langle x| \psi = \int dx \psi(x) |x\rangle \quad (3.50)$$

If $\{|x\rangle\}$ are the position eigenstates, then the complex function $\psi(x)$ the position representation of $|\psi\rangle$. This function $\psi(x)$ is nothing but the wavefunction of the Schrödinger equation. Eq. (3.50) shows how the Schrödinger wavefunction should be interpreted within the more abstract Dirac notation.

The basis states can be formed from any complete set of orthogonal states. In particular, they can be formed from the basis states of the position or the momentum operator, i.e. $\int_{-\infty}^{\infty} dx |x\rangle\langle x| = \int_{-\infty}^{\infty} dp |p\rangle\langle p| = \hat{I}$. If we apply these definitions, we can then recover the familiar Fourier representation,

$$\psi(x) = \langle x|\psi\rangle = \int_{-\infty}^{\infty} dp \underbrace{\langle x|p\rangle}_{e^{ipx/\hbar}/\sqrt{2\pi\hbar}} \langle p|\psi\rangle = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} dp e^{ipx/\hbar} \psi(p), \quad (3.51)$$

where $\langle x|p\rangle$ is the position representation of the momentum eigenstate $|p\rangle$.

3.2.2 Eigenfunctions and Eigenvalues

Given an $N \times N$ matrix A , it is often useful to determine its eigenvalues. Recall that these are a collection of (up to) N numbers λ which solve the equation

$$A\vec{u} = \lambda\vec{u} \quad \text{for some vector } \vec{u}. \quad (3.52)$$

The corresponding vector \vec{u} is then called the eigenvector.

We can play the same game with operators. Given an operator \hat{A} , its *eigenvalues* λ solve the equation

$$\hat{A}\psi(\mathbf{x}) = \lambda\psi(\mathbf{x}) \quad \text{for some function } \psi(\mathbf{x}). \quad (3.53)$$

The corresponding function $\psi(\mathbf{x})$ is called the *eigenfunction* or *eigenstate*. The collection of all eigenvalues is called the *spectrum* of the operator \hat{A} .

This has an important interpretation in physics. If you measure the value of a physical observable \hat{A} then the answer that you get will be one of the eigenvalues of \hat{A} . This is the key idea that relates operators to observables.

The outcome of any measurement of \hat{A} lies in the spectrum of \hat{A} .

We've already met this idea in subsection 1.3.3. The time independent Schrödinger equation is simply the eigenvalue equation for the Hamiltonian

$$\hat{H}\psi = E\psi. \quad (3.54)$$

The spectrum of the Hamiltonian determines the possible energy levels of the quantum system. In most cases this spectrum is discrete, like the harmonic oscillator. In other simple, but mildly pathological cases, the spectrum is continuous, like a particle on a line.

3.2.3 Hermitian Operators

Not any old linear operator qualifies as a physical observable in quantum mechanics. We should restrict attention to those operators that are *Hermitian*.

To explain this, we first introduce the concept of an adjoint operator. Given an operator \hat{A} , its adjoint \hat{A}^\dagger is defined by the requirement that the following inner-product relation holds

$$\langle \psi | \hat{A} \phi \rangle = \langle \hat{A}^\dagger \psi | \phi \rangle, \quad (3.55)$$

for all states ψ and ϕ . Written in terms of integrals of wavefunctions, this means

$$\int d^3x \psi^* \hat{A} \phi = \int d^3x (\hat{A}^\dagger \psi)^* \phi. \quad (3.56)$$

An operator is said to be *Hermitian* or *self-adjoint* if

$$\hat{A} = \hat{A}^\dagger. \quad (3.57)$$

All physical observables correspond to Hermitian operators. Let's check that this is the case for the operators that we just met. First the position operator. We have

$$\langle \psi | \hat{\mathbf{x}} \phi \rangle = \int d^3x \psi^* \mathbf{x} \phi = \int d^3x (\mathbf{x} \psi)^* \phi = \langle \hat{\mathbf{x}} \psi | \phi \rangle. \quad (3.58)$$

So we learn that $\hat{\mathbf{x}} = \hat{\mathbf{x}}^\dagger$ as promised. Here the calculation was trivial: it is just a matter of running through definitions. The calculation for the momentum operator is marginally less trivial. We have

$$\begin{aligned} \langle \psi | \hat{\mathbf{p}} \phi \rangle &= \int d^3x \psi^* (-i\hbar \nabla \phi) \\ &= \int d^3x i\hbar (\nabla \psi^*) \phi \\ &= \int d^3x (-i\hbar \nabla \psi)^* \phi = \langle \hat{\mathbf{p}} \psi | \phi \rangle \end{aligned} \quad (3.59)$$

In the second line we integrated by parts, picking up one minus sign, and in the third line we took the factor of i inside the brackets with the complex conjugation giving us a second minus sign. The two minus signs cancel out to give $\hat{\mathbf{p}} = \hat{\mathbf{p}}^\dagger$. This is one way to see why the momentum operator necessarily comes with a factor of i .

Note that, when we integrated by parts, we threw away the total derivative. This is justified on the grounds that the states ψ and ϕ are normalisable and so asymptote to zero at infinity.

The Hermiticity of the Hamiltonian follows immediately because it's a real function of \hat{x} and \hat{p} . Alternatively, you check explicitly that the Hamiltonian is Hermitian by doing the kind of integration-by-parts calculation that we saw above.

3.2.3.1 Properties of Hermitian Matrices

Hermitian operators have a number of properties that make them the right candidates for physical observables. To explain these, we're first going to regress to the simpler world of $N \times N$ complex matrices acting on N -dimensional complex vectors.

A complex matrix A is said to be Hermitian if it obeys $A^\dagger := (A^*)^T = A$. This means that the components of the matrix obey $A_{ij} = A_{ji}^*$. To see that this coincides with our previous definition, recall that the inner product for two complex vectors is (3.6)

$$\langle \vec{u} | \vec{v} \rangle = \vec{u}^* \cdot \vec{v} \quad (3.60)$$

For a Hermitian matrix A we have

$$\langle \vec{u} | A \vec{v} \rangle = \vec{u}^* \cdot \vec{v} = u_i^* A_{ij} v_j = u_i^* A_{ji}^* v_j = (A \vec{u})^* \cdot \vec{v} = \langle A \vec{u} | \vec{v} \rangle, \quad (3.61)$$

in agreement with our earlier, more general definition.

The eigenvalues and eigenvectors of a Hermitian matrix A are determined by the equation

$$A \vec{u}_n = \lambda_n \vec{u}_n \quad (3.62)$$

where $n = 1, \dots, N$ labels the eigenvalues and eigenvectors. These have a number of special properties. The first two are:

- The eigenvalues λ_n are real.
- If two eigenvalues are distinct $\lambda_N \neq \lambda_M$ then their corresponding eigenvectors are orthogonal: $\vec{u}_n^* \cdot \vec{u}_m = 0$.

Both are simple to prove. We take the inner product of A between two different eigenvectors.

$$\begin{aligned} \langle \vec{u}_n | A \vec{u}_m \rangle &= \langle A \vec{u}_n | \vec{u}_m \rangle \implies \lambda_n^* \vec{u}_n^* \cdot \vec{u}_m = \lambda_m \vec{u}_n^* \cdot \vec{u}_m \\ &\implies (\lambda_n^* - \lambda_m) \vec{u}_n^* \cdot \vec{u}_m = 0. \end{aligned} \quad (3.63)$$

Set $n = m$ to find $\lambda_n = \lambda_n^*$ so that the eigenvalue is real. Alternatively, pick distinct eigenvalues $\lambda_N \neq \lambda_m$ to show that the eigenvectors are orthogonal, $\vec{u}_n^* \cdot \vec{u}_m = 0$ if $m \neq n$.

If the matrix has two or more eigenvalues that coincide, then it's still possible to pick the eigenvectors to be orthogonal. We won't prove this statement, but it follows straightforwardly by just picking an orthogonal basis for the space spanned by the eigenvectors.

There is one final property of Hermitian matrices that we won't prove: they always have N eigenvectors. This is important. Because the eigenvectors can be taken to be orthogonal, they necessarily span the space \mathbb{C}^N , meaning that any vector can be expanded in terms of eigenvectors

$$\vec{u} = \sum_{n=1}^N a_n \vec{u}_n, \quad (3.64)$$

for some complex coefficients a_n .

With these results in mind, let's now return to Hermitian operators that act on infinite dimensional Hilbert spaces.

3.2.3.2 Properties of Hermitian Operators

The eigenvalue for a Hermitian operator is

$$\hat{A} |\phi_n\rangle = \lambda_n |\phi_n\rangle, \quad (3.65)$$

where λ_n is the eigenvalue and we've shifted notation slightly from our earlier discussion and will refer to eigenfunctions as $|\phi_n\rangle$. Now the label runs over $n = 1, \dots, \infty$ reflecting the fact that we've got an infinite dimensional Hilbert space.

For a Hermitian operator \hat{A} , the eigenvalues and eigenfunctions have the following two properties:

- The eigenvalues λ_n are real.
- If two eigenvalues are distinct $\lambda_n \neq \lambda_m$ then their corresponding eigenfunctions are orthogonal: $\langle \phi_n | \phi_m \rangle = 0$.

Recall our important boxed statement from earlier: the outcome of any measurement of \hat{A} is given by one of the eigenvalues of \hat{A} . The first of the properties above ensures that, happily, the result of any measurement is guaranteed to be a real number. That's a good thing. If you go into a lab you'll see dials and needles and rulers and other complicated things. All of them measure real numbers, never complex.

The proof of the two properties above is identical to the proof for finite matrices. If \hat{A} is Hermitian, we have

$$\begin{aligned} \langle \phi_n | \hat{A} \phi_m \rangle &= \langle \hat{A} \phi_n | \phi_m \rangle \implies \lambda_n^* \langle \phi_n | \phi_m \rangle = \lambda_m \langle \phi_n | \phi_m \rangle \\ &\implies (\lambda_n^* - \lambda_m) \langle \phi_n | \phi_m \rangle = 0. \end{aligned} \quad (3.66)$$

Again, set $n = m$ to find $\lambda_n = \lambda_n^*$ so that the eigenvalue is real. Alternatively, pick distinct eigenvalues $\lambda_n \neq \lambda_m$ to show that the eigenfunctions are orthogonal, which now means

$$\langle \phi_n | \phi_m \rangle = \int d^3x \phi_n^* \phi_m = 0 \quad \text{if } m \neq n. \quad (3.67)$$

An operator with distinct eigenvalues is said to have a *non-degenerate* spectrum. If, on the other hand, two or more eigenvalues coincide then we say it has a *degenerate* spectrum. In this case, as with matrices, it's still possible to pick orthogonal eigenfunctions. Again, we won't prove this. However, it means that in all cases, if we normalise the eigenstates then we can always take them to obey

$$\langle \phi_n | \phi_m \rangle = \int d^3X \phi_n^* \phi_m = \delta_{mn}. \quad (3.68)$$

Eigenfunctions that obey this are said to be *orthonormal*.

Finally, one last important statement that we make without proof: the eigenfunctions of any Hermitian operator are *complete*. This means that we can expand any wavefunction $|\psi\rangle$ in terms of eigenstates of a given operator \hat{A}

$$|\psi\rangle = \sum_{n=1}^{\infty} a_n |\phi_n\rangle, \quad (3.69)$$

for some complex coefficients a_n . If you know both $|\psi\rangle$ and the normalised eigenfunctions $|\phi_n\rangle$, then it's simple to get an expression for these coefficients: we use the orthonormality relation (3.68), together with linearity of the inner product, to get

$$\langle \phi_n | \psi \rangle = \sum_m a_m \langle \phi_n | \phi_m \rangle = \sum_m a_n \delta_{mn} = a_n. \quad (3.70)$$

Or, in terms of integrals

$$a_n = \int d^3x \phi_n^*(\mathbf{x}) \psi(\mathbf{x}). \quad (3.71)$$

Relatedly, the norm of a wavefunction $|\psi(\mathbf{x})\rangle$ is

$$|\psi(\mathbf{x})|^2 := \int d^3x \psi^*(\mathbf{x}) \psi(\mathbf{x}) = \sum_m \sum_n \int d^3x a_n^* \phi_n^* a_m \phi_m = \sum_N |a_n|^2, \quad (3.72)$$

where we have again used the orthonormality condition (3.68). We see that a wavefunction is normalised only in $\sum_n |a_n|^2 = 1$.

3.2.4 Measurement

Now we turn to the final facet of quantum mechanics: measurement. Suppose that we have a quantum system in a given state $\psi(\mathbf{x})$. We want to measure an observable that is associated to an operator \hat{A} whose eigenvalue equation is

$$\hat{A}\phi_n = \lambda_n \phi_n. \quad (3.73)$$

We stated before that the possible result of any measurement is taken from the spectrum of eigenvalues $\{\lambda_n\}$. But how does this result depend on the state ψ ?

Here is the answer. We first decomposes the state ψ in terms of orthonormal eigenstates of \hat{A} ,

$$\psi(\mathbf{x}) = \sum_n a_n \phi_n(\mathbf{x}). \quad (3.74)$$

We will take both ϕ_n and the wavefunction ψ to be normalised, which means that $\sum_n |a_n|^2 = 1$. Assuming that the spectrum is non-degenerate, the probability that the result of the measurement gives the answer λ_n is given by the *Born rule* (1.19),

$$\text{Prob}(\lambda_n) = |a_n|^2. \quad (3.75)$$

A normalised wavefunction ensures that the sum of probabilities is $\sum_n \text{Prob}(\lambda_n) = 1$.

We make a measurement and get the result λ_n for some particular n . Now there's no doubt about the value of the observable: we've just measured it and know that it's λ_n . To reflect this, there is a “collapse of the wavefunction” which jumps to

$$\psi(\mathbf{x}) \rightarrow \phi_n(\mathbf{x}). \quad (3.76)$$

This ensures that if we make a second measurement of \hat{A} immediately after the first then we get the same answer: λ_n .

The discontinuous jump of the wavefunction (3.11) is very different in character from the smooth evolution under the guidance of the time dependent Schrödinger equation. In

particular, the collapse of the wavefunction is not even a linear operation: if you add two wavefunctions together, you first have to normalise the sum to compute the probability and that's a simple, but non-linear calculation.

Given that measurement of quantum states is performed by stuff that's made out of atoms and therefore should itself be governed by the Schrödinger equation, it seems like there should be a more unified description of how states evolve.

There's a simple generalisation of the story above when the spectrum is degenerate. In that case, the probability that we get a given result λ is

$$\text{Prob}(\lambda) = \sum_{n|\lambda_n=\lambda} |a_n|^2 \quad (3.77)$$

After the measurement, the wavefunction collapses to $\psi \rightarrow C \sum_{n|\lambda_n=\lambda} a_n \phi_n$ where C is the appropriate normalisation factor.

First One Thing, Then The Other Let's look at one obvious implication of the measurement procedure described above. Suppose that we measure a (non-degenerate) observable \hat{A} and get a definite result, say λ_6 . Immediately after the measurement the system sits in the state

$$\psi(\mathbf{x}) = \phi_6(\mathbf{x}). \quad (3.78)$$

We already noted that if we again measure \hat{A} then there's no doubt about the outcome: we get the result λ_6 again for sure. The general statement is that the system has a definite value for some observable only if it sits in an eigenstate of that observable.

Now we measure a different observable, \hat{B} . In general, this will have a different set of eigenstates,

$$\hat{B}\chi_n(\mathbf{x}) = \mu_n \chi_n(\mathbf{x}), \quad (3.79)$$

and a different set of outcomes $\{\mu_n\}$. There's no reason that the eigenstates of \hat{B} should have anything to do with the eigenstates of \hat{A} : they will typically be entirely different functions. We now have to play the game all over again: we take our state, which we know is ϕ_6 , and expand

$$\phi_6(\mathbf{x}) = \sum_m b_m \chi_m(\mathbf{x}). \quad (3.80)$$

There is no certainty about what we get when we measure \hat{B} : the result μ_m appears with probability $\text{Prob}(\mu_m) = |b_m|^2$ and the wavefunction will immediately collapse to the corresponding eigenstate $\chi_m(\mathbf{x})$. Suppose that we do the experiment and find the result μ_{17} . Now the wavefunction collapses again, to

$$\phi_6(\mathbf{x}) \rightarrow \chi_{17}(\mathbf{x}). \quad (3.81)$$

Now comes the rub. We are capricious and decide to go back and measure the original observable \hat{A} . It wasn't so long ago that we measured it and found the result λ_6 . But now, having measured \hat{B} , there's no guarantee that we'll get λ_6 again! Instead, we need to go through the same process and expand our wavefunction

$$\chi_{17}(\mathbf{x}) = \sum_n c_n \phi_n(\mathbf{x}). \quad (3.82)$$

The probability that we get the result λ_6 again is $|c_6|^2$. But now it's just one among many options. There's no reason to have $|c_6| = 1$.

There are different words that we can drape around this. First, it's clear that the measurement of a quantum system is never innocent. You can't just take a small peek and walk away pretending that you've not done anything. Instead, any measurement of a quantum system necessarily disturbs the state.

Moreover, if the state originally had a specific value for one observable \hat{A} and you measure a different observable \hat{B} then you destroy the original property of the state. It's tempting to say that there's no way to *know* the values of both \hat{A} and \hat{B} at the same time. But that's not what the mathematics is telling us. Instead, the quantum particle can't *have* specific values of \hat{A} and \hat{B} at the same time. Indeed, most of the time it doesn't have a specific value for either since its wavefunction is a superposition of eigenstates. But, providing the eigenstates of the two operators don't coincide, if the quantum state takes a definite value for one observable, then it cannot for the other. We'll quantify this idea in Subsection 3.3 where we introduce the Heisenberg uncertainty relations

3.2.4.1 Expectation Values

Take a normalised state $\psi(\mathbf{x})$ and measure the observable \hat{A} , with eigenstates and eigenvalues given by

$$\hat{A}\phi_n = \lambda_n\phi_n. \quad (3.83)$$

As we've seen, if we expand the original state in orthonormal eigenfunctions

$$\psi(\mathbf{x}) = \sum_n a_n \phi_n(\mathbf{x}), \quad (3.84)$$

then the probability that we measure λ_n is $|a_n|^2$ (assuming a non-degenerate spectrum). This means that if we have many systems, all in the same state ψ , and perform the same measurement on each then the results will differ but the average will be

$$\langle \hat{A} \rangle_{\psi} = \sum_n |a_n|^2 \lambda_n. \quad (3.85)$$

Note the little subscript ψ on the angular brackets, reminding us that the average value depends on the state of the system. We call this average the *expectation value*. It has a nice expression in terms of the inner product.

For a normalised wavefunction $\psi(\mathbf{x})$, the expectation value can be written as

$$\langle \hat{A} \rangle_{\psi} = \langle \psi | \hat{A} | \psi \rangle = \int d^3x \psi(\mathbf{x})^* \hat{A} \psi(\mathbf{x}). \quad (3.86)$$

If $\psi(\mathbf{x})$ is un-normalised, this should be replaced by

$$\langle \hat{A} \rangle_{\psi} = \frac{\langle \psi | \hat{A} | \psi \rangle}{\langle \psi | \psi \rangle} \quad (3.87)$$

which follows from (3.86) by first normalising ψ . To show this, it's simplest to start with a normalised wavefunction and work backwards:

$$\begin{aligned} \int d^3x \psi(\mathbf{x})^* \hat{A} \psi(\mathbf{x}) &= \sum_n \sum_m \int d^3x a_n^* a_m \phi_n^*(\mathbf{x}) \hat{A} \phi_m(\mathbf{x}) \\ &= \sum_n \sum_m \int d^3x a_n^* a_m \phi_n^*(\mathbf{x}) \lambda_m \phi_m(\mathbf{x}) \\ &= \sum_n \sum_m a_n^* a_m \lambda_m \delta_{nm} = \sum_n |a_n|^2 \lambda_n, \end{aligned} \quad (3.88)$$

where, in the last step, we used the orthonormality of the eigenstates (3.68).

3.2.5 Time-Evolution Operator

The ability to develop an eigenfunction expansion provides the means to explore the time evolution of a general wave packet, $|\psi\rangle$ under the action of a Hamiltonian. Formally, we can evolve a wavefunction forward in time by applying the time-evolution operator. For a Hamiltonian which is time independent, we have $|\psi(t)\rangle = \hat{U}|\psi(0)\rangle$, where

$$\boxed{\hat{U} = e^{-i\hat{H}t/\hbar}}, \quad (3.89)$$

denotes the time-evolution operator.⁴ By inserting the resolution of identity, $\mathbb{I} = \sum_i |i\rangle\langle i|$, where the states $|i\rangle$ are eigenstates of the Hamiltonian with eigenvalue E_i , we find that

$$|\psi(t)\rangle = e^{-i\hat{H}t/\hbar} \sum_i \langle i| \langle i|\psi(0)\rangle = \sum_i \langle i| \langle i|\psi(0)\rangle e^{-iE_i t/\hbar}. \quad (3.90)$$

The time-evolution operator is an example of a **unitary operator**. The latter are defined as transformations which preserve the scalar product, $\langle \phi|\psi\rangle = \langle \hat{U}\phi|\hat{U}\psi\rangle = \langle \phi|\hat{U}^\dagger\hat{U}\psi\rangle \stackrel{!}{=} \langle \psi|\phi\rangle$, i.e.

$$\boxed{\hat{U}^\dagger \hat{U} = \hat{\mathbb{I}}} \quad (3.91)$$

3.2.6 Commutation Relations

There is an algebraic way of formalising whether two observables can be simultaneously measured or whether, as is usually the case, a measurement of one messes up a previous measurement of the other. The underlying structure is known as *commutation relations*.

Given two operators \hat{A} and \hat{B} , their *commutator* is defined to be

$$\boxed{[\hat{A}, \hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A}}. \quad (3.92)$$

The key idea is familiar from matrices. If you multiply two matrices A and B then the order matters. Typically AB is not the same as BA . The commutator $[A, B]$ captures the

⁴This equation follows from integrating the time-dependent Schrödinger equation, $\hat{H}|\psi\rangle = i\hbar\partial_t|\psi\rangle$.

difference between these. Clearly $[A, B] = -[B, A]$ and the same also holds for operators: $[\hat{A}, \hat{B}] = -[\hat{B}, \hat{A}]$.

Some care should be taken with formulae like (3.92). The operators have a job to do: they act on functions. It's important to remember this when manipulating these kind of equations. A good rule of thumb is to think of operator equations like (3.92) as meaning

$$[\hat{A}, \hat{B}]\psi = \hat{A}\hat{B}\psi - \hat{B}\hat{A}\psi \quad \forall\psi(\mathbf{x}). \quad (3.93)$$

When we do calculations, at least initially, it's useful to use this phrasing to avoid any slip ups.

The commutation relations of $\hat{\mathbf{x}} = (\hat{x}_1, \hat{x}_2, \hat{x}_3)$ and $\hat{\mathbf{p}} = (\hat{p}_1, \hat{p}_2, \hat{p}_3)$ are

$$[\hat{x}_i, \hat{x}_j] = [\hat{p}_i, \hat{p}_j] = 0 \quad \text{and} \quad [\hat{x}_i, \hat{p}_j] = i\hbar\delta_{ij}. \quad (3.94)$$

These are known as the *canonical commutation relations* on account of the important role that they play in quantum mechanics. Note that the right-hand side of the final relation is just a constant, but this too should be viewed as an operator: it's the trivial operator that just multiplies any function by a constant.

The last commutation relation is perhaps surprising. In 1D, it reads simply $[\hat{x}, \hat{p}] = i\hbar$. Note that in higher dimensions we can write this as $[\hat{\mathbf{x}}, \hat{\mathbf{p}}] = i\hbar\mathbb{I}$. This kind of relation is not possible for finite matrices A and B : the analogous result would be $[A, B] = i\hbar\mathbb{I}$ with \mathbb{I} the unit matrix. But if we take the trace of the left-hand side we find

$$\text{Tr}[A, B] = \text{Tr}(AB - BA) = 0 \quad (3.95)$$

because $\text{Tr}(AB) = \text{Tr}(BA)$. Yet clearly $\text{Tr}(\mathbb{I}) \neq 0$. However, infinite matrices, a.k.a. operators, can do things that finite matrices cannot, the canonical commutation relations (3.94) among them.

The canonical commutation relations follow straightforwardly from the definitions

$$\hat{x}_i = x_i \quad \text{and} \quad \hat{p}_i = -i\hbar\frac{\partial}{\partial x_i}. \quad (3.96)$$

Recalling that we should view all operators as acting on functions, as in (3.93), the commutation of position operators is

$$[\hat{x}_i, \hat{x}_j]\psi(\mathbf{x}) = (x_i x_j - x_j x_i)\psi(\mathbf{x}) = 0 \quad \forall\psi(\mathbf{x}), \quad (3.97)$$

while the commutation of momentum operators is

$$[\hat{p}_i, \hat{p}_j]\psi(\mathbf{x}) = \left(\frac{\partial}{\partial x_i} \frac{\partial}{\partial x_j} - \frac{\partial}{\partial x_j} \frac{\partial}{\partial x_i} \right) \psi(\mathbf{x}) = 0 \quad \forall\psi(\mathbf{x}), \quad (3.98)$$

which follows because the order in which we take partial derivatives doesn't matter. That leaves us just with the more interesting commutation relation. It is

$$[\hat{x}_i, \hat{p}_j]\psi(\mathbf{x}) = -i\hbar \left(x_i \frac{\partial}{\partial x_j} - \frac{\partial}{\partial x_j} x_i \right) \psi(\mathbf{x}). \quad (3.99)$$

We can see the issue: the first derivative hits only the function ψ . Meanwhile the second derivative acts on $x_i\psi$, giving us an extra term from the chain rule. Moreover, the $\partial\psi/\partial x_j$ terms cancel, leaving us only with this extra term

$$[\hat{x}_i, \hat{p}_j]\psi(\mathbf{x}) = i\hbar \frac{\partial x_i}{\partial x_j} \psi(\mathbf{x}) = i\hbar \delta_{ij} \psi(\mathbf{x}) \quad \forall \psi(\mathbf{x}). \quad (3.100)$$

Note that the proof ultimately boils down to the operator equation

$$x_i \frac{\partial}{\partial x_j} - \frac{\partial}{\partial x_j} x_i = -\delta_{ij}. \quad (3.101)$$

When you first meet these kinds of equations, it's not immediately obvious why they're true. As we saw above, the right way to proceed is to check that it holds on when evaluated on any arbitrary function $\psi(\mathbf{x})$.

One reason why the canonical commutation relations hold such a prominent position is because very similar equations can be found in classical mechanics. In that context, the relation involves neither commutators nor \hbar , but something called a *Poisson bracket*. This is a concept that arises in the more sophisticated, Hamiltonian approach to classical mechanics. While the physics of quantum mechanics is a huge departure from our classical world, the underlying mathematical structure turns out to be surprisingly close.

Occasionally on our quantum travels, we will run into two operators that commute, meaning

$$[\hat{A}, \hat{B}] = 0. \quad (3.102)$$

This is interesting because $[\hat{A}, \hat{B}] = 0$ if and only if \hat{A} and \hat{B} share the same eigenfunctions. First suppose that both \hat{A} and \hat{B} share the same eigenfunctions, so that

$$\hat{A}\phi_n(\mathbf{x}) = \lambda_N \phi_n(\mathbf{x}) \quad \text{and} \quad \hat{B}\phi_n(\mathbf{x}) = \mu_n \phi_n(\mathbf{x}). \quad (3.103)$$

Then any function $\psi(\mathbf{x})$ can be expanded in terms of these eigenstates $\psi = \sum_n a_n \phi_n$, giving

$$[\hat{A}, \hat{B}]\psi = \sum_n a_n [\hat{A}, \hat{B}]\phi_n = \sum_n a_n (\lambda_N \mu_n - \mu_N \lambda_n) \phi_n = 0 \quad \forall \psi(\mathbf{x}). \quad (3.104)$$

Conversely, suppose that $[\hat{A}, \hat{B}] = 0$. We will restrict ourselves to the case where \hat{A} has a non-degenerate spectrum. (The claim holds more generally but we have to work a little harder.) If ϕ is an eigenstate of \hat{A} , then we have

$$\begin{aligned} \hat{A}\phi = \lambda\phi &\implies \hat{B}\hat{A}\phi = \lambda\hat{B}\phi \\ &\implies \hat{A}\hat{B}\phi = \lambda\hat{M}\phi, \end{aligned} \quad (3.105)$$

where the second line follows because $[\hat{A}, \hat{B}] = 0$. But this tells us that $\hat{B}\phi$ is also an eigenstate of \hat{A} with eigenvalue λ . By assumption, there is only one eigenstate with eigenvalue λ , which means that we must have

$$\hat{M}\phi = \mu\phi, \quad (3.106)$$

for some μ . But this is the statement that ϕ is also an eigenstate of \hat{B} . The two operators need not have the same eigenvalues, but they do have the same eigenstates.

The physical interpretation of the claim follows from our discussion in Subsection 3.2.4. It is possible for a quantum system to have simultaneous values for commuting observables. If you measure \hat{A} and then measure \hat{M} , the state will not be perturbed as long as $[\hat{A}, \hat{B}] = 0$. (Actually this last statement does require that the spectrum of \hat{A} is non-degenerate.) If you then measure \hat{A} again, you'll get the same answer as the first time.

3.2.6.1 Example: Evolution of Harmonic Oscillator

Consider the harmonic oscillator Hamiltonian $\hat{H} = \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega^2x^2$. Later in this chapter, we will see that the eigenstates, $|n\rangle$, have equally-spaced eigenvalues, $E_n = \hbar\omega(n + 1/2)$, for $n = 0, 1, 2, \dots$. Let us then consider the time-evolution of a general wavepacket, $|\psi(0)\rangle$, under the action of the Hamiltonian. From the equation above, we find that $|\psi(t)\rangle = \sum_n |n\rangle \langle n|\psi(0)\rangle e^{-iE_nt/\hbar}$. Since the eigenvalues are equally spaced, let us consider what happens when $t = t_r \equiv 2\pi r/\omega$, with r integer. In this case, since $e^{2\pi i n r} = 1$, we have

$$|\psi(t_r)\rangle = \sum_n |n\rangle \langle n|\psi(0)\rangle e^{-i\omega t_r/2} = (-1)^r |\psi(0)\rangle \quad (3.107)$$

From this result, we can see that, up to an overall phase, the wave packet is perfectly reconstructed at these times. This recurrence or “echo” is not generic, but is a manifestation of the equal separation of eigenvalues in the harmonic oscillator.

3.3 The Generalised Uncertainty Principle

For non-commuting Hermitian operators, $[\hat{A}, \hat{B}] \neq 0$, it is straightforward to establish a bound on the uncertainty in their expectation values. Given a state $|\psi\rangle$, the mean square uncertainty is defined as

$$(\Delta_\Psi A)^2 = \langle\psi|(\hat{A} - \langle\hat{A}\rangle)^2|\psi\rangle = \langle\psi|\hat{U}^2|\psi\rangle, \quad (3.108)$$

$$(\Delta_\Psi B)^2 = \langle\psi|(\hat{B} - \langle\hat{B}\rangle)^2|\psi\rangle = \langle\psi|\hat{V}^2|\psi\rangle, \quad (3.109)$$

where we have defined the operators $\hat{U} = \hat{A} - \langle\hat{A}\rangle$ and $\hat{V} = \hat{B} - \langle\hat{B}\rangle$. Since $\langle\hat{A}\rangle$ and $\langle\hat{B}\rangle$ are just constants, $[\hat{U}, \hat{V}] = [\hat{A}, \hat{B}]$. We can also notice that this quantity is well defined such that is always positive, since $\langle\psi|(\hat{A} - \langle\hat{A}\rangle)^2|\psi\rangle = |\hat{A} - \langle\hat{A}\rangle|_\Psi^2$. We can also see from this representation that $\Delta_\Psi A$ must only vanish if and only if $\hat{A}|\Psi\rangle = \langle\hat{A}\rangle|\Psi\rangle$, i.e. $|\Psi\rangle$ is an eigenstate of \hat{A} .

Now let us take the scalar product of $\hat{U}|\psi\rangle + i\lambda\hat{V}|\psi\rangle$ with itself to develop some information about the uncertainties. As a modulus, the scalar product must be greater

than or equal to zero, i.e. expanding, we have $\langle \psi | \hat{U}^2 | \psi \rangle + \lambda^2 \langle \psi | \hat{V}^2 | \psi \rangle + i\lambda \langle \hat{U}\psi | \hat{V}\psi \rangle - i\lambda \langle \hat{V}\psi | \hat{U}\psi \rangle \geq 0$. Reorganising this equation in terms of the uncertainties, we thus find

$$(\Delta_\Psi A)^2 + \lambda^2 (\Delta_\Psi B)^2 + i\lambda \langle [\hat{U}, \hat{V}] \rangle \geq 0. \quad (3.110)$$

If we minimise this expression with respect to λ , we can determine when the inequality becomes strongest. In doing so, we find

$$2\lambda(\Delta_\Psi B)^2 + i \langle [\hat{U}, \hat{V}] \rangle = 0, \quad \lambda = -\frac{i}{2} \frac{\langle [\hat{U}, \hat{V}] \rangle}{(\Delta_\Psi B)^2}. \quad (3.111)$$

Substituting this value of λ back into the inequality, we then find,

$$(\Delta_\Psi A)^2 (\Delta_\Psi B)^2 \geq -\frac{1}{4} \langle [\hat{U}, \hat{V}] \rangle^2. \quad (3.112)$$

We therefore find that, for non-commuting operators, the uncertainties obey the following inequality,

$$\boxed{\Delta_\Psi A \Delta_\Psi B \geq \left| \frac{1}{2} \langle [\hat{A}, \hat{B}] \rangle \right|.} \quad (3.113)$$

An alternate, possibly easier to recall, proof follows by first defining the state

$$|\Psi_A\rangle = \hat{A} |\Psi\rangle - \langle \hat{A} \rangle_\Psi |\Psi\rangle. \quad (3.114)$$

Thus we can see from Eq. (3.108) that

$$||\Psi_A\rangle| = \Delta_\Psi A. \quad (3.115)$$

Now, considering the inner product between $|\Psi_A\rangle$ and a similarly defined $|\Psi_B\rangle$ gives

$$\langle \Psi_A | \Psi_B \rangle = \langle \Psi | \hat{A} \hat{B} | \Psi \rangle - \langle \hat{A} \rangle \langle \hat{B} \rangle. \quad (3.116)$$

It is now useful to consider the imaginary part of this inner product using $2i \operatorname{Im}\{z\} = z - z^*$,

$$\begin{aligned} 2i \operatorname{Im}\{\langle \Psi_A | \Psi_B \rangle\} &= \langle \Psi | \hat{A} \hat{B} | \Psi \rangle - \langle \Psi | \hat{A} \hat{B} | \Psi \rangle^* \\ &= \langle \Psi | [\hat{A}, \hat{B}] | \Psi \rangle. \end{aligned} \quad (3.117)$$

Finally we can again prove the generalised uncertainty principle

$$\begin{aligned} \Delta_\Psi A \Delta_\Psi B &= |\Psi_A| |\Psi_B| && \text{Eq. (3.115)} \\ &\geq |\langle \Psi_A | \Psi_B \rangle| && \text{Cauchy-Schwarz} \\ &\geq \frac{1}{2} \left| \langle [\hat{A}, \hat{B}] \rangle \right| && |z| \geq \operatorname{Im}\{z\} \end{aligned} \quad (3.118)$$

If the commutator is a constant, as in the case of the conjugate operators $[\hat{p}, \hat{x}] = -i\hbar$, the expectation values can be dropped, and we obtain the relation, $(\Delta_\Psi A)(\Delta_\Psi B) \geq \frac{i}{2} [\hat{A}, \hat{B}]$. For momentum and position, this result recovers **Heisenberg's uncertainty principle**,

$$\boxed{\Delta p \Delta x \geq \left| \frac{1}{2} \langle [\hat{p}, \hat{x}] \rangle \right| = \frac{\hbar}{2}.} \quad (3.119)$$

3.4 Time-Evolution of Expectation Values

Finally, to close this section on operators, let us consider how their expectation values evolve. To do so, let us consider a general operator \hat{A} which may itself involve time. The time derivative of a general expectation value has three terms.

$$\frac{d}{dt} \langle \psi | \hat{A} | \psi \rangle = \partial_t (\langle \psi |) \hat{A} | \psi \rangle + \langle \psi | \partial_t \hat{A} | \psi \rangle + \langle \psi | \hat{A} (\partial_t | \psi \rangle). \quad (3.120)$$

If we then make use of the time-dependent Schrödinger equation, $i\hbar \partial_t | \psi \rangle = \hat{H} | \psi \rangle$, and the Hermiticity of the Hamiltonian, we obtain

$$\frac{d}{dt} \langle \psi | \hat{A} | \psi \rangle = \underbrace{\frac{i}{\hbar} (\langle \psi | \hat{H} \hat{A} | \psi \rangle - \langle \psi | \hat{A} \hat{H} | \psi \rangle)}_{\frac{i}{\hbar} \langle [\hat{H}, \hat{A}] \rangle} + \langle \psi | \partial_t \hat{A} | \psi \rangle. \quad (3.121)$$

$$(3.122)$$

This is an important and general result for the time derivative of expectation values which becomes simple if the operator itself does not explicitly depend on time,

$$\frac{d}{dt} \langle \hat{A} \rangle = \frac{i}{\hbar} \langle [\hat{H}, \hat{A}] \rangle. \quad (3.123)$$

From this result, which is known as **Ehrenfest's theorem**, we see that expectation values of operators that commute with the Hamiltonian are constants of the motion

3.4.0.1 Example: Evolution of \hat{x} and \hat{p} Operators

From the non-relativistic Schrödinger operator for a single particle moving in a potential, $\hat{H} = \frac{\hat{p}^2}{2m} + V(\hat{x})$, we can derive the quantum mechanical counterparts of Hamilton's classical equations of motion.

Suppose we wanted to know the instantaneous change in the expectation of the momentum operator \hat{p} . Using Ehrenfest's theorem, we have

$$\frac{d}{dt} \langle \hat{p} \rangle = \frac{i}{\hbar} \langle [\hat{H}, \hat{p}] \rangle + \left\langle \frac{\partial}{\partial t} \hat{p} \right\rangle = \frac{i}{\hbar} \langle [V(\hat{x}), \hat{p}] \rangle, \quad (3.124)$$

since the operator \hat{p} commutes with itself and has no time dependence. By expanding the right hand side, using the identity $[F(\hat{x}), \hat{p}] = i\hbar \partial_{\hat{x}} F(\hat{x})$, we obtain

$$\frac{d}{dt} \langle \hat{p} \rangle = - \langle \partial_{\hat{x}} V(\hat{x}) \rangle = - \langle \partial_{\hat{x}} \hat{H} \rangle. \quad (3.125)$$

Similarly, we can obtain the instantaneous change in the position expectation value,

$$\frac{d}{dt} \langle \hat{x} \rangle = \frac{i}{\hbar} \langle [\hat{H}, \hat{x}] \rangle + \left\langle \frac{\partial}{\partial t} \hat{x} \right\rangle = \frac{i}{2m\hbar} \langle [\hat{p}^2, \hat{x}] \rangle, \quad (3.126)$$

and thus from this we can evaluate the appropriate commutator, noting that $[\hat{p}, \hat{x}] = -i\hbar$, to give

$$\frac{d}{dt} \langle \hat{x} \rangle = \frac{1}{m} \langle \hat{p} \rangle = \langle \partial_{\hat{p}} \hat{H} \rangle. \quad (3.127)$$

Although, at first glance, it might appear that the Ehrenfest theorem is saying that the quantum mechanical expectation values obey Newton's classical equations of motion, this is not actually the case. This result does not say that the pair $\{\langle \hat{x} \rangle, \langle \hat{p} \rangle\}$ satisfies Newton's second law, because the right-hand side of the formulae (3.125) and (3.127) are of the form $\langle F(\hat{x}, t) \rangle$ rather than $F(\langle \hat{x} \rangle, t)$. Nevertheless, for states that are highly localised in space, the expected position and momentum will approximately follow classical trajectories, which may be understood as an instance of the correspondence principle.

3.5 The Heisenberg Picture

Until now, the time dependence of an evolving quantum system has been placed within the wavefunction while the operators have remained constant - this is the **Schrödinger picture or representation**. However, it is sometimes useful to transfer the time-dependence to the operators. To see how, let us consider the expectation value of some operator \hat{B} ,

$$\langle \psi(t) | \hat{B} | \psi(t) \rangle = \left\langle e^{-i\hat{H}t/\hbar} \psi(0) \left| \hat{B} \right| e^{-i\hat{H}t/\hbar} \psi(0) \right\rangle = \langle \psi(0) | e^{i\hat{H}t/\hbar} \hat{B} e^{-i\hat{H}t/\hbar} | \psi(0) \rangle. \quad (3.128)$$

According to rules of associativity, we can multiply operators together before using them. If we define the operator $\hat{B}(t) = e^{i\hat{H}t/\hbar} \hat{B} e^{-i\hat{H}t/\hbar}$, the time-dependence of the expectation values has been transferred from the wavefunction. This is called the **Heisenberg picture or representation** and in it, the operators evolve with time (even if, as in this case, they have no explicit time dependence) while the wavefunctions remain constant. In this representation, the time derivative of the operator is given by

$$\begin{aligned} \frac{d}{dt} \hat{B} &= \frac{i\hat{H}}{\hbar} e^{i\hat{H}t/\hbar} \hat{B} e^{-i\hat{H}t/\hbar} - e^{i\hat{H}t/\hbar} \hat{B} \frac{i\hat{H}}{\hbar} e^{-i\hat{H}t/\hbar} \\ &= \frac{i}{\hbar} e^{i\hat{H}t/\hbar} [\hat{H}, \hat{B}] e^{-i\hat{H}t/\hbar} \\ &= \frac{i}{\hbar} [\hat{H}, \hat{B}(t)]. \end{aligned} \quad (3.129)$$

$i\hbar \frac{d\hat{B}}{dt} = [\hat{B}(t), \hat{H}].$

(3.130)

3.6 Quantum Harmonic Oscillator

As we will see time and again in this course, the harmonic oscillator assumes a privileged position in quantum mechanics and quantum field theory finding numerous and sometimes unexpected applications. It is useful to us now in that it provides a platform for us to implement some of the technology that has been developed in this chapter. In the one-dimensional case, the quantum harmonic oscillator Hamiltonian takes the form,

$$\hat{H} = \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega^2 \hat{x}^2, \quad (3.131)$$

where $\hat{p} = -i\hbar\partial_x$. To find the eigenstates of the Hamiltonian, we could look for solutions of the linear second order differential equation corresponding to the time-independent

Schrödinger equation, $\hat{H}\psi = E\psi$, where $\hat{H} = -\frac{\hbar^2}{2m}\partial_x^2 + \frac{1}{2}m\omega^2\hat{x}^2$. The integrability of the Schrödinger operator in this case allows the stationary states to be expressed in terms of a set of orthogonal functions known as Hermite polynomials. However, the complexity of the exact eigenstates obscure a number of special and useful features of the harmonic oscillator system. To identify these features, we will instead follow a method based on an operator formalism.

This is an important example for it is also a case where we can start from operator algebra, and thence define \mathcal{H} as a representation. The form of the Hamiltonian as the sum of the squares of momenta and position suggests that it can be recast as the “square of an operator”. To this end, let us introduce the dimensionless operator

$$a = \sqrt{\frac{m\omega}{2\hbar}}\left(\hat{x} + i\frac{\hat{p}}{m\omega}\right), \quad a^\dagger = \sqrt{\frac{m\omega}{2\hbar}}\left(\hat{x} - i\frac{\hat{p}}{m\omega}\right), \quad (3.132)$$

where, for notational convenience, we have not drawn hats on the operators a and its Hermitian conjugate a^\dagger . Making use of the identity,

$$a^\dagger a = \frac{m\omega}{2\hbar}\hat{x}^2 + \frac{\hat{p}^2}{2\hbar m\omega} + \frac{i}{2\hbar}[\hat{x}, \hat{p}] = \frac{\hat{H}}{\hbar\omega} - \frac{1}{2} \quad (3.133)$$

and the parallel relation, $aa^\dagger = \frac{\hat{H}}{\hbar\omega} + \frac{1}{2}$, we see that the operators fulfill the commutation relations

$$[a, a^\dagger] \equiv aa^\dagger - a^\dagger a = 1. \quad (3.134)$$

Then, setting the hermitian number operator as $\hat{N} = a^\dagger a$, the Hamiltonian can be cast in the form

$$\hat{H} = \hbar\omega(\hat{N} + 1/2). \quad (3.135)$$

We also consequently find the useful commutation relations,

$$\begin{aligned} [\hat{N}, \hat{a}^\dagger] &= [\hat{a}^\dagger \hat{a}, \hat{a}^\dagger] \\ &= \hat{a}^\dagger [\hat{a}, \hat{a}^\dagger] + [\hat{a}^\dagger, \hat{a}^\dagger] \hat{a}^\dagger \\ &= \hat{a}^\dagger, \end{aligned} \quad (3.136)$$

and similarly we find

$$[\hat{N}, \hat{a}] = -\hat{a}. \quad (3.137)$$

Since the operator $\hat{N} = a^\dagger a$ must lead to a positive definite result, we see that the eigenstates of the harmonic oscillator must have energies of $\hbar\omega/2$ or higher. Moreover, the ground state $|0\rangle$ can be identified by finding the state for which $a|0\rangle = 0$. Expressed in the coordinate basis, this translates to the equation,⁵

$$\left(x + \frac{\hbar}{m\omega}\partial_x\right)\psi_0(x) = 0, \quad \psi_0(x) = \langle x|0\rangle = \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} e^{-m\omega x^2/2\hbar}. \quad (3.139)$$

⁵Formally, in coordinate basis, we have $\langle x'|a|x\rangle = \delta(x'-x)(a + \frac{\hbar}{m\omega}\partial_x)$ and $\langle x|0\rangle = \psi_0(x)$. Then making use of the resolution of identity $\int dx |x\rangle\langle x| = \hat{\mathbb{I}}$, we have

$$\langle x|a|0\rangle = 0 = \int dx \langle x|a|x'\rangle \langle x'|0\rangle = \left(x + \frac{\hbar}{m\omega}\partial_x\right)\psi_0(x). \quad (3.138)$$

Since $\hat{N}|0\rangle = a^\dagger a|0\rangle = 0$, this state is an eigenstate with energy $\hbar\omega/2$. The higher lying states can be found by acting upon this state with the operator a^\dagger . The proof runs as follows: If $\hat{N}|n\rangle = n|n\rangle$, we have

$$\hat{N}(a^\dagger|n\rangle) = a^\dagger \underbrace{aa^\dagger}_{a^\dagger a+1} |n\rangle = (a^\dagger \underbrace{a^\dagger a}_{\hat{N}} + a^\dagger)|n\rangle = (n+1)a^\dagger|n\rangle \quad (3.140)$$

or, equivalently, $[\hat{N}, a^\dagger] = a^\dagger$. In other words, if $|n\rangle$ is an eigenstate of \hat{N} with eigenvalue n , then $a^\dagger|n\rangle$ is an eigenstate with eigenvalue $n+1$.

From this result, we can deduce that the eigenstates for a “tower” $|0\rangle, |1\rangle = C_1 a^\dagger |0\rangle, |2\rangle = C_2 (a^\dagger)^2 |0\rangle$, etc., where C_n denotes the normalisation. If $\langle n|n\rangle = 1$ we have

$$\langle n|aa^\dagger|n\rangle = \langle n|(\hat{N}+1)|n\rangle = (n+1). \quad (3.141)$$

Therefore, with $|n+1\rangle = \frac{1}{\sqrt{n+1}}a^\dagger|n\rangle$ the state $|n+1\rangle$ is also normalised, $\langle n+1|n+1\rangle = 1$. By induction, we can deduce the general normalisation,

$$|n\rangle = \frac{1}{\sqrt{n!}}(a^\dagger)^n|0\rangle, \quad (3.142)$$

with $\langle n|n'\rangle = \delta_{nn'}$, $\hat{H}|n\rangle \hbar\omega(n+1/2)|n\rangle$ and

$$a^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle, \quad a|n\rangle = \sqrt{n}|n-1\rangle. \quad (3.143)$$

The operators a and a^\dagger represent **ladder operators** and have the effect of lowering or raising the energy of the state.

In fact, the operator representation achieves something quite remarkable and, as we will see, unexpectedly profound. The quantum harmonic oscillator describes the motion of a single particle in a one-dimensional potential well. Its eigenvalues turn out to be equally spaced - a ladder of eigenvalues, separated by a constant energy $\hbar\omega$. If we are energetic, we can of course translate our results into a coordinate representation $\psi_n(x) = \langle x|n\rangle$.⁶ However, the operator representation affords a second interpretation, one that lends itself to further generalisation in quantum field theory. We can instead interpret the quantum harmonic oscillator as a simple system involving many fictitious particles, each of energy $\hbar\omega$. In this representation, known as the **Fock space**, the vacuum state $|0\rangle$ is one involving no particles, $|1\rangle$ involves a single particle, $|2\rangle$ has two and so on. These fictitious particles are created and annihilated by the action of the raising and lowering operators, a^\dagger and a with canonical commutation relations, $[a, a^\dagger] = 1$. Later in the course, we will find that these commutation relations are the hallmark of **bosonic** quantum particles and this representation, known as the **second quantisation** underpins the quantum field theory of the electromagnetic field.

This completes our abridged survey of operator methods in quantum mechanics.

⁶Expressed in real space, the harmonic oscillator wavefunctions are in fact described by the Hermite polynomials,

$$\psi_n(x) = \langle x|n\rangle = \sqrt{\frac{1}{2^n n!}} H_n\left(\sqrt{\frac{m\omega}{\hbar}}x\right) \exp\left[-\frac{m\omega x^2}{2\hbar}\right], \quad (3.144)$$

where $H_n(x) = (-1)^n e^{x^2} \frac{d^n}{dx^n} e^{-x^2}$.

CHAPTER 4

Quantum Mechanics in More Than One Dimension

Previously, we have explored the manifestations of quantum mechanics in one spatial dimension and discussed the properties of bound and unbound states. The concepts developed there apply equally to higher dimension. However, for a general two or three-dimensional potential, without any symmetry, the solutions of the Schrödinger equation are often inaccessible. In such situations, we may develop approximation methods to address the properties of the states (e.g. WKB method, and see chapter 7). However, in systems where there is a high degree of symmetry, the quantum mechanics of the system can often be reduced to a tractable low-dimensional theory.

4.1 Rigid Diatomic Molecule

As a pilot example let us consider the quantum mechanics of a rigid diatomic molecule with nuclear masses m_1 and m_2 , and (equilibrium) bond length, R_e . Since the molecule is rigid, its coordinates can be specified by its centre of mass, $\mathbf{R} = \frac{m_1\mathbf{r}_1 + m_2\mathbf{r}_2}{m_1 + m_2}$, and internal orientation, $\mathbf{r} = \mathbf{r}_2 - \mathbf{r}_1$ (with $|\mathbf{r}| = R_e$). Defining the total mass $M = m_1 + m_2$, and moment of inertia, $I = \mu R_e^2$, where $\mu = m_1 m_2 / (m_1 + m_2)$ denotes the reduced mass, the corresponding Hamiltonian can be then separated into the kinetic energy associated with the centre of mass motion and the rotational kinetic energy,

$$\hat{H} = \frac{\hat{\mathbf{P}}^2}{2M} + \frac{\hat{\mathbf{L}}^2}{2I}, \quad (4.1)$$

where $\hat{\mathbf{P}} = -i\hbar\nabla_{\mathbf{R}}$ and $\hat{\mathbf{L}} = \hat{\mathbf{r}} \times \hat{\mathbf{p}}$ denotes the angular momentum associated with the internal degrees of freedom. Since the internal and centre of mass degrees of freedom separate, the wavefunction can be factorized as $\psi(\mathbf{r}, \mathbf{R}) = e^{i\mathbf{K}\cdot\mathbf{R}}Y(\mathbf{r})$, where the first factor accounts for the free particle motion of the body, and the second factor relates to the internal angular degrees of freedom.

As a result of the coordinate separation, we have reduced the problem of the rigid diatomic molecule to the study of the quantum mechanics of a particle moving on a sphere – the **rigid rotor**,

$$\hat{H}_{\text{rot}} = \frac{\hat{\mathbf{L}}^2}{2I}. \quad (4.2)$$

The eigenstates of this component of the Hamiltonian are simply the states of the angular momentum operator. Indeed, in *any* quantum mechanical system involving a radial potential, the angular momentum will be conserved, i.e. $[\hat{H}, \hat{\mathbf{L}}] = 0$ meaning that the angular component of the wavefunction can be indexed by the states of the angular momentum operator. We therefore now digress to discuss the quantum mechanics of angular momentum

4.2 Angular momentum

In this section, we'll consider finite-dimensional Hilbert spaces \mathcal{H}_j whose elements transform among themselves under rotations. In mathematics, the \mathcal{H}_j are known as *irreducible representations* of the rotation group, whilst in physics they're called *multiplets* of definite total angular momentum. This section will also substantiate the link between the rotation generators $\hat{\mathbf{J}}$ and the physics of angular momentum, by examining a simple Hamiltonian for rotating diatomic molecule. We'll also flesh out the details of exactly how the orientation of a system is encoded in the amplitudes for it to be found in different eigenstates of appropriate angular momentum operators.

4.2.1 Commutation Relations

Following the usual canonical quantisation procedure, the angular momentum operator is defined by $\hat{\mathbf{L}} = \hat{\mathbf{r}} \times \hat{\mathbf{p}}$ where, as usual $\hat{\mathbf{r}}$ and $\hat{\mathbf{p}}$ obey the commutation relations, $[\hat{p}_i, \hat{r}_j] = -i\hbar\delta_{ij}$. Using this relation, one may then show that the angular momentum operators obey the **spin commutation relations**,

$$[\hat{L}_i, \hat{L}_j] = i\hbar\epsilon_{ijk}\hat{L}_k. \quad (4.3)$$

where, as usual, ϵ_{ijk} denotes the totally antisymmetric tensor - the Levi Civita symbol.¹

Unlike components of $\hat{\mathbf{L}}$ and $\hat{\mathbf{r}}$, and $\hat{\mathbf{L}}$ and $\hat{\mathbf{p}}$ are non-commuting

$$[\hat{L}_i, \hat{r}_j] = i\hbar\epsilon_{ijk}\hat{r}_k, \quad [\hat{L}_i, \hat{p}_j] = -\hbar\epsilon_{ijk}\hat{p}_k. \quad (4.4)$$

Consider a vector operator $\hat{\mathbf{V}} = (\hat{V}_x, \hat{V}_y, \hat{V}_z) = (\hat{V}_1, \hat{V}_2, \hat{V}_3)$ satisfying the commutation relations

$$[\hat{L}_i, \hat{V}_j] = i\hbar\epsilon_{ijk}\hat{V}_k. \quad (4.5)$$

Then, for the squared operator $\hat{\mathbf{V}}^2 = \hat{V}_x^2 + \hat{V}_y^2 + \hat{V}_z^2$, we have

$$[\hat{L}_i, \hat{\mathbf{V}}^2] = [\hat{L}_i, \hat{V}_j]\hat{V}_j + \hat{V}_j[\hat{L}_i, \hat{V}_j] = i\hbar\epsilon_{ijk}(\hat{V}_k\hat{V}_j + \hat{V}_j\hat{V}_k) = 0, \quad (4.6)$$

since ϵ_{ijk} is antisymmetric in j, k , while $\hat{V}_k\hat{V}_j + \hat{V}_j\hat{V}_k$ is symmetric. For such vector operators

$$[\hat{L}_i, \hat{\mathbf{V}}^2] = 0, \quad [\hat{\mathbf{L}}, \hat{\mathbf{V}}^2] = 0, \quad [\hat{L}_i^2, \hat{\mathbf{V}}^2] = 0, \quad [\hat{\mathbf{L}}^2, \hat{\mathbf{V}}^2] = 0. \quad (4.7)$$

The operators $\hat{\mathbf{V}} = \hat{\mathbf{L}}, \hat{\mathbf{r}}, \hat{\mathbf{p}}$ all satisfy Eq. (4.5), giving immediately

$$[\hat{L}_i, \hat{\mathbf{L}}^2] = [\hat{L}_i, \hat{\mathbf{r}}^2] = [\hat{L}_i, \hat{\mathbf{p}}^2] = [\hat{\mathbf{L}}^2, \hat{\mathbf{r}}^2] = [\hat{\mathbf{L}}^2, \hat{\mathbf{p}}^2] = 0. \quad (4.8)$$

¹Recall that $\epsilon_{ijk} = 1$ if $\{i, j, k\}$ is an even permutation of $\{1, 2, 3\}$, -1 if it is an odd permutation, and 0 if any index is repeated.

4.2.2 Eigenvalues of Angular Momentum

In the following, we will construct a basis set of angular momentum states. Since the angular momentum is a *vector* quantity, it may be characterised by its magnitude and direction. For the former, let us define the operator $\hat{\mathbf{L}}^2 = \hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2$. With the latter, since the separate components of the angular momentum are all mutually non-commuting, we cannot construct a common set of eigenstates for any two of them. However, since $[\hat{\mathbf{L}}, \hat{\mathbf{L}}^2] = 0$ we can find a complete set of simultaneous eigenstates of (any) one component of $\hat{\mathbf{L}}$ and $\hat{\mathbf{L}}^2$. Without loss of generality, we orient our coordinate system so that the chosen component of $\hat{\mathbf{L}}$ is \hat{L}_z . Therefore, in the following, we will look for an eigenbasis of $\hat{\mathbf{L}}^2$ and *one* direction, say \hat{L}_z ,

$$\hat{\mathbf{L}}^2 |a, b\rangle = a |a, b\rangle, \quad \hat{L}_z |a, b\rangle = b |a, b\rangle. \quad (4.9)$$

We also choose all the states $\{|a, b\rangle\}$ to be properly normalised. Since they are eigenstates of Hermitian operators, they must then be orthonormal:

$$\langle a', b' | a, b \rangle = \delta_{aa'} \delta_{bb'}. \quad (4.10)$$

Finally, since we're looking for the simplest possible way to realize our commutation relations, we'll assume that the states $\{|a, b\rangle\}$ are non-degenerate; that is, we want a Hilbert space where $|a, b\rangle$ is the unique eigenstate of $\hat{\mathbf{L}}^2$ and \hat{L}_z with the given eigenvalues

To find $|a, b\rangle$, we could simply proceed by looking for a suitable coordinate basis to represent $\hat{\mathbf{L}}^2$ and \hat{L}_z in terms of differential operators. However, although we will undertake such a programme in due course, before getting to this formalism, we can make substantial progress without resorting to an explicit coordinate representation.

4.2.2.1 Raising and Lowering Operators for Angular Momentum

The set of eigenvalues a and b can be obtained by making use of a trick based on a “ladder operator” formalism which parallels that used in the study of the quantum harmonic oscillator in section 3.6. Specifically, let us define the raising and lowering operators,

$$\hat{L}_{\pm} = \hat{L}_x \pm i\hat{L}_y, \quad (4.11)$$

which obey $\hat{L}_{\pm}^\dagger = \hat{J}_{\mp}$ and so are each other's adjoint. With this definition, one may then show that

$$[\hat{L}_z, \hat{L}_{\pm}] = \pm \hbar \hat{L}_{\pm}. \quad (4.12)$$

Since each component of the angular momentum commutes with $\hat{\mathbf{L}}^2$, we can deduce that the action of \hat{L}_{\pm} on $|a, b\rangle$ cannot affect the value of a relating to the magnitude of the angular momentum. However, they do affect the projection:

$$\hat{L}_z \hat{L}_{\pm} |a, b\rangle = \hat{L}_{\pm} \hat{L}_z |a, b\rangle + [\hat{L}_z, \hat{L}_{\pm}] |a, b\rangle = (b \pm \hbar) |a, b\rangle. \quad (4.13)$$

Therefore, if $|a, b\rangle$ is an eigenstate of \hat{L}_z with eigenvalue b , $\hat{L}_{\pm} |a, b\rangle$ is either zero, or an eigenstate of \hat{L}_z with eigenvalue $b \pm \hbar$, i.e. $\hat{L}_{\pm} |a, b\rangle = C_{\pm}(a, b) |a, b \pm \hbar\rangle$ where $C_{\pm}(a, b)$ is a normalisation constant. We see that the role of \hat{L}_{\pm} in angular momentum is similar

to the role of \hat{a} and \hat{a}^\dagger for energy of the harmonic oscillator. Again, \hat{J}_+ and \hat{J}_- are often called *raising* and *lowering* operators. Their role is to rotate our system, aligning more or less of its total angular momentum along the z -axis without changing the total angular momentum available.

To fix the normalisation, we may note that the norm,

$$\left\| \hat{L}_\pm |a, b\rangle \right\|^2 = \langle a, b | \hat{L}_\pm^\dagger \hat{L}_\pm |a, b\rangle = \langle a, b | \hat{L}_\mp \hat{L}_\pm |a, b\rangle, \quad (4.14)$$

where we have used the identity $\hat{L}_\pm^\dagger = \hat{L}_\mp$. Then, making use of the relation $\hat{L}_\mp \hat{L}_\pm = \hat{L}_x^2 + \hat{L}_y^2 \pm i[\hat{L}_x, \hat{L}_y] = \hat{\mathbf{L}}^2 - \hat{L}_z^2 \mp \hbar \hat{L}_z$, and the presumed normalisation, $\langle a, b | a, b\rangle = 1$, one finds

$$\left\| \hat{L}_\pm |a, b\rangle \right\|^2 = \langle a, b | (\hat{\mathbf{L}}^2 - \hat{L}_z^2 \mp \hbar \hat{L}_z) |a, b\rangle = a - b^2 \mp \hbar b. \quad (4.15)$$

As a represents the eigenvalue of a sum of squares of Hermitian operators, it is necessarily non-negative. Moreover, b is real. Therefore, for a given a , b must be bounded: there must be a b_{max} and a (negative or zero) b_{min} . In particular,

$$\left\| \hat{L}_+ |a, b_{max}\rangle \right\|^2 = a - b_{max}^2 \mp \hbar b_{max} \quad (4.16)$$

$$\left\| \hat{L}_- |a, b_{min}\rangle \right\|^2 = a - b_{min}^2 \mp \hbar b_{min} \quad (4.17)$$

For a given a , b_{max} and b_{min} are determined uniquely – there cannot be two states with the same a but different b annihilated by \hat{L}_+ . It also follows immediately that $a = b_{max}(b_{max} + \hbar)$ and $b_{min} = -b_{max}$. Furthermore, we know that if we keep operating on $|a, b_{min}\rangle$ with \hat{L}_+ , we generate a sequence of states with \hat{L}_z eigenvalues $b_{min} + \hbar, b_{min} + 2\hbar, b_{min} + 3\hbar, \dots$. This series must terminate, and the only possible way for that to happen is for b_{max} to be equal to $b_{min} + n\hbar$ with n integer, from which it follows that b_{max} is either an integer or half an odd integer times \hbar .

At this point, we switch to the standard notation. We have established that the eigenvalues of \hat{L}_z form a finite ladder, with spacing \hbar . We write them as $m\hbar$, and ℓ is used to denote the maximum value of m , so the eigenvalue of $\hat{\mathbf{L}}^2$, $a = \ell(\ell + 1)\hbar^2$. Both ℓ and m will be integers or half odd integers, but the spacing of the ladder of m values is always unity. Although we have been writing $|a, b\rangle$ with $a = \ell(\ell + 1)\hbar^2$, $b = m\hbar$ we shall henceforth follow convention and write $|\ell, m\rangle$.

In summary, the operators $\hat{\mathbf{L}}^2$ and \hat{L}_z have a common set of orthonormal eigenstates $|\ell, m\rangle$ with

$$\boxed{\hat{\mathbf{L}}^2 |\ell, m\rangle = \ell(\ell + 1)\hbar^2 |\ell, m\rangle, \quad \hat{L}_z |\ell, m\rangle = m\hbar |\ell, m\rangle,} \quad (4.18)$$

where ℓ, m are integers or half-integers. The allowed quantum numbers m form a ladder with step spacing unity, the maximum value of m is ℓ , and the minimum value is $-\ell$. With these results, we may then return to the normalisation of the raising and lowering operators. In particular, making use of Eq. (4.15), we have

$$\boxed{\begin{aligned} \hat{L}_+ |\ell, m\rangle &= \hbar \sqrt{\ell(\ell + 1) - m(m + 1)} |\ell, m + 1\rangle \\ \hat{L}_- |\ell, m\rangle &= \hbar \sqrt{\ell(\ell + 1) - m(m - 1)} |\ell, m - 1\rangle \end{aligned}} \quad (4.19)$$

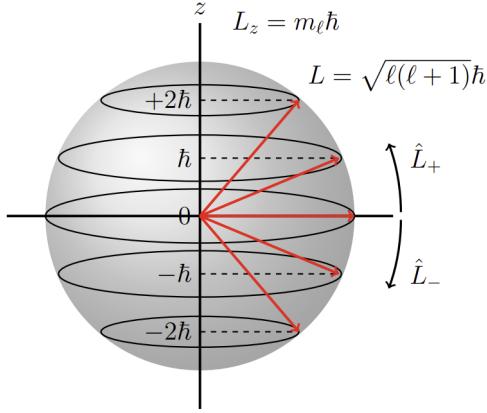


Fig. 4.1: Quantisation of angular momentum, a schematic showing the angular momentum scheme for $\ell = 2$ with $\mathbf{L}^2 = \sqrt{2(2+1)}\hbar = \sqrt{6}\hbar$ and the five possible values for the L_z projection. The ladder operators increase/decrease the z -component of angular momentum. The series terminates at the upper and lower ends when L_z becomes greater than the magnitude L by generating the zero vector. Notice that the z -component of angular momentum can be zero, and the average value of the angular momentum can be zero.

The use of m to denote the component of angular momentum in one direction came about because a Bohr-type electron in orbit is a current loop, with a magnetic moment parallel to its angular momentum. So the m measured the component of magnetic moment in a chosen direction, usually along an external magnetic field. For this reason, m is often termed the **magnetic quantum number**.

Note that, since \hat{L}_x and \hat{L}_y can be written in terms of the raising and lowering operators \hat{L}_{\pm} , rotations around an arbitrary axis transform states in a given \mathcal{H}_ℓ into other states in the same \mathcal{H}_ℓ . Since $\hat{L}_x = \frac{1}{2}(\hat{L}_+ + \hat{L}_-)$, we see that when $\ell > 0$, the state $|\ell, m\rangle$ is never an eigenstate of \hat{L}_x . Hence, for $\ell > 0$, we can never be certain of the outcome of measurements of both \hat{L}_x and \hat{L}_z . This argument applies equally to \hat{L}_y , so it's impossible to be certain of the outcome of measurements of more than one component of $\hat{\mathbf{L}}$. If $\ell = 0$, every component of $\hat{\mathbf{L}}$ yields zero, but a null vector in \mathbb{R}^3 has no direction. Consequently, there are no states in which the vector $\hat{\mathbf{L}}$ has a well-defined direction. This situation contrasts with the momentum vector $\hat{\mathbf{P}}$ which can have a well-defined direction since all its components commute with one another.

In the mathematical literature, states with $m = \ell$ are known as *highest weight* states. They play a key role, because once we know the highest weight state the rest of the multiplet can be constructed by applying lowering operators such as \hat{L}_- . Physically, a highest weight state is one in which the body's angular momentum is most nearly aligned along the z -axis. In this state, the ratio of the squared angular momentum that lies in the xy -plane to that parallel to the z -axis is

$$\frac{\langle \ell, \ell | \hat{L}_x^2 + \hat{L}_y^2 | \ell, \ell \rangle}{\langle \ell, \ell | \hat{L}_z^2 | \ell, \ell \rangle} = \frac{\langle \ell, \ell | \hat{\mathbf{L}}^2 - \hat{L}^2 | \ell, \ell \rangle}{\hbar^2 \ell^2} = \frac{1}{\ell}. \quad (4.20)$$

As $\hat{L}_x = \frac{1}{2}(\hat{L}_+ + \hat{L}_-)$ we have $\langle \ell, \ell | \hat{J}_x | \ell, \ell \rangle = 0$ and similarly $\langle \ell, \ell | \hat{L}_y | \ell, \ell \rangle = 0$. Thus we have no information about the direction in the xy -plane any angular momentum points.

Macroscopic bodies, for which $\ell \gg 1$, have only a tiny fraction of their total angular

momentum in the xy -plane when they're in state $|\ell, \ell\rangle$, so the uncertain direction of this component of angular momentum does not lead to significant uncertainty in the total direction of the angular momentum vector. By contrast, when $\ell = 1/2$, even in state $|1/2, 1/2\rangle$ there's twice as much angular momentum associated with the xy -plane as with the z -axis and the uncertain direction of this planar component makes it impossible to say anything more specific than that the angular momentum vector lies somewhere in the northern, rather than southern, hemisphere. Even when $\ell = 1$ and $m = 1$, the state $|1, 1\rangle$ has as much angular momentum along the xy -plane as it has parallel to the z -axis, so the direction of the angular momentum vector is still very uncertain. This is no surprise: since $\dim \mathcal{H}_\ell = 2\ell + 1$, when $\ell = \frac{1}{2}$ there are only two independent states the orientation of our body can take. With such a limited Hilbert space it's no wonder we only have a fuzzy notion of where the body's angular momentum lies.

4.2.3 Representation of the Angular Momentum States

Having established expressions for the eigenvalues of the angular momentum operators, it is now necessary to establish (angular) coordinate representations for the corresponding eigenstates, $Y_{\ell m}(\theta, \phi) = \langle \theta, \phi | \ell, m \rangle$. Here the angles θ and ϕ denote the spherical coordinates parameterising the unit sphere. Previously, we obtained the eigenvalues of the angular momentum operator by making use of the raising and lowering operators in a manner that paralleled the study of the quantum harmonic oscillator. Similarly, to obtain explicit expressions for the eigenstates, we must make use of the coordinate representation of these operators. With $\mathbf{r} = r\hat{\mathbf{e}}_r$, the gradient operator can be written in spherical polar coordinates as

$$\nabla = \hat{\mathbf{e}}_r \partial_r + \hat{\mathbf{e}}_\theta \frac{1}{r} \partial_\theta + \hat{\mathbf{e}}_\phi \frac{1}{r \sin \theta} \partial_\phi. \quad (4.21)$$

From this result, we thus obtain

$$\boxed{\hat{L}_z = -i\partial_\phi, \quad \hat{L}_\pm \hbar e^{\pm i\phi} (\pm \partial_\theta + i \cot \theta \partial_\phi),} \quad (4.22)$$

and, at least formally,

$$\boxed{\hat{\mathbf{L}}^2 = -\hbar^2 \left[\frac{1}{\sin \theta} \partial_\theta (\sin \theta \partial_\theta) + \frac{1}{\sin^2 \theta} \partial_\phi^2 \right].} \quad (4.23)$$

Beginning with the eigenstates of \hat{L}_z , the eigenvalue equation (4.18), and making use of the expression above, we have

$$-i\hbar\partial_\phi Y_{\ell m}(\theta, \phi) = m\hbar Y_{\ell m}(\theta, \phi). \quad (4.24)$$

Since the left hand side depends only on ϕ , the solution is separable and takes the form $Y_{\ell m}(\theta, \phi) = F(\theta)e^{im\phi}$. Note that, provided m is integer, the continuity of the wavefunction, $Y_{\ell m}(\theta, \phi + 2\pi) = Y_{\ell m}(\theta, \phi)$, is ensured. Thus, the condition that the co-ordinate representation of the wavefunction be singlevalued requires m to be integer, and hence ℓ to be integer. (As discussed in Chapter 5, for “spin” angular momentum half-integer ℓ is possible.)

To determine the second component of the eigenstates, $F(\theta)$, we could immediately turn to the eigenvalue equation involving the differential operator for $\hat{\mathbf{L}}^2$,

$$\left[\frac{1}{\sin \theta} \partial_\theta (\sin \theta \partial_\theta) - \frac{m^2}{\sin^2 \theta} \right] F(\theta) = \ell(\ell+1)F(\theta) \quad (4.25)$$

However, to construct the states, it is easier to draw upon the properties of the angular momentum raising and lowering operators (much in the same way that the Hermite polynomials are generated by the action of ladder operators in the harmonic oscillator problem).

Consider then the state of maximal m , $|\ell, \ell\rangle$, for which $\hat{L}_+ |\ell, \ell\rangle = 0$. Making use of the coordinate representation of the raising operator above together with the separability of the wavefunction, this relation implies that

$$0 = \langle \theta, \phi | \hat{L}_+ | \ell, \ell \rangle = \hbar e^{i\phi} (\partial_\theta + i \cot \theta \partial_\phi) Y_{\ell m}(\theta, \phi) = \hbar e^{i(\ell+1)\phi} (\partial_\theta - \ell \cot \theta) F(\theta). \quad (4.26)$$

From this result it follows that $\partial_\theta F(\theta) = \ell \cot \theta F(\theta)$ with the solution $F(\theta) = C \sin^\ell \theta$, and C a constant of normalization. States with values of m lower than ℓ can then be obtained simply by repeated application of the angular momentum lowering operator \hat{L}_- to the state $|\ell, \ell\rangle$. This amounts to the relation

$$Y_{\ell m}(\theta, \phi) = C(\hat{L}_-)^{\ell-m} [\sin^\ell \theta e^{i\ell\phi}] \quad (4.27)$$

$$= C(-\partial_\theta + \ell \cot \theta)^{\ell-m} [\sin^\ell \theta e^{i\ell\phi}]. \quad (4.28)$$

The eigenfunctions produced by this procedure are well known and referred to as the **spherical harmonics**. In particular, one finds that the normalised eigenstates take the form,

$$Y_{\ell m}(\theta, \phi) = (-1)^{m+|m|} \left[\frac{2\ell+1}{4\pi} \frac{(\ell-|m|)!}{(\ell+|m|)!} \right]^{1/2} P_\ell^{|m|}(\cos \theta) e^{im\phi} \quad (4.29)$$

where

$$P_\ell^m(\xi) = \frac{(1-\xi^2)^{m/2}}{2^\ell \ell!} \frac{d^{m+\ell}}{d\xi^{m+\ell}} (\xi^2 - 1)^\ell, \quad (4.30)$$

represent the **associated Legendre polynomials**. In particular, for the first few angular momentum states, we have

$$\begin{aligned} Y_{00} &= \frac{1}{\sqrt{4\pi}}, \\ Y_{10} &= \sqrt{\frac{3}{4\pi}} \cos \theta, & Y_{11} &= -\sqrt{\frac{3}{8\pi}} e^{i\phi} \sin \theta, \\ Y_{20} &= \sqrt{\frac{5}{16\pi}} (3 \cos^2 \theta - 1), & Y_{21} &= -\sqrt{\frac{15}{8\pi}} \sin \theta \cos \theta, & Y_{22} &= \sqrt{\frac{15}{32\pi}} e^{2i\phi} \sin^2 \theta. \end{aligned} \quad (4.31)$$

As a complete basis set, the spherical harmonics can be used as a resolution of the identity

$$\sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} |\ell, m\rangle \langle \ell, m| = \hat{\mathbb{I}}. \quad (4.32)$$

In terms of the angular position eigenstates $|\theta, \phi\rangle$, we have

$$\int_0^\pi \sin \theta d\theta \int_0^{2\pi} d\phi |\theta, \phi\rangle \langle \theta, \phi| = \hat{\mathbb{I}}, \quad (4.33)$$

where the integration measure $\sin \theta d\theta d\phi$ weights all directions equally. The normalisation condition is

$$\langle \theta', \phi' | \theta, \phi \rangle = \frac{1}{\sin \theta} \delta(\theta - \theta') \delta(\phi - \phi'), \quad (4.34)$$

as may be readily verified by requiring $\hat{\mathbb{I}}^2 = \hat{\mathbb{I}}$ from (4.33).

Thus, expressing the angular momentum eigenstates in the coordinate basis, we have

$$\sum_{\ell=0}^{\infty} Y_{\ell,m}^*(\theta', \phi') Y_{\ell,m}(\theta, \phi) = \frac{1}{\sin \theta} \delta(\theta - \theta') \delta(\phi - \phi'), \quad (4.35)$$

Similarly, we have the orthogonality condition,

$$\int_0^\pi \sin \theta d\theta \int_0^{2\pi} d\phi Y_{\ell,m}^*(\theta', \phi') Y_{\ell,m}(\theta, \phi) = \delta_{\ell\ell'} \delta_{mm'}. \quad (4.36)$$

An equal mix of all spherical harmonics $Y_{\ell,m}$ for a given ℓ is isotropic (a uniform sphere)

$$\sum_{m=-\ell}^{\ell} |Y_{\ell,m}(\theta, \phi)|^2 = \text{const.} \quad (4.37)$$

Under a *parity operation* $\mathbf{r} \rightarrow -\mathbf{r}$, the spherical harmonics transform as

$$Y_{\ell,m}(\pi - \theta, \pi + \phi) = (-1)^\ell Y_{\ell,m}(\theta, \phi), \quad (4.38)$$

i.e. $Y_{\ell,m}$ are eigenstates of parity with eigenvalue

$$P = (-1)^\ell. \quad (4.39)$$

After this lengthy digression, we may now return to the problem of the quantum mechanical rotor Hamiltonian and the rigid diatomic molecule. From the analysis above, we have found that the eigenstates of the Hamiltonian (4.1) are given by $\psi(\mathbf{R}, \mathbf{r}) = \frac{1}{\sqrt{2\pi}} e^{i\mathbf{K}\cdot\mathbf{R}} Y_{\ell,m}(\theta, \phi)$ with eigenvalues

$$E_{\mathbf{K},\ell} = \frac{\hbar^2 \mathbf{K}^2}{2m} + \frac{\hbar^2}{2I} \ell(\ell + 1), \quad (4.40)$$

where each \mathbf{K}, ℓ value has a $2\ell + 1$ -fold degeneracy.

Indeed, this simplification applies to any quantum mechanical system involving a *spherically symmetric* potential, $\hat{V}(r)$. Since the operator $\hat{\mathbf{L}}$ depends only on the angular variables θ and ϕ , for any function of r alone, $f(r)$,

$$[f(r), \hat{L}_x] = [f(r), \hat{L}_y] = [f(r), \hat{L}_z] = 0, \quad (4.41)$$

and thus

$$[\hat{\mathbf{L}}, f(r)] = [\hat{\mathbf{L}}^2, f(r)] = 0. \quad (4.42)$$

The angular momentum operator commutes also with $\hat{\mathbf{p}}^2$ (See Eq. (4.8))

$$[\hat{\mathbf{L}}, \hat{\mathbf{p}}] = [\hat{\mathbf{L}}^2, \hat{\mathbf{p}}^2] = 0. \quad (4.43)$$

Therefore, for a radial potential, $\hat{\mathbf{L}}$ commutes with the Hamiltonian, $\hat{H} = \hat{\mathbf{p}}^2/2m + V(r)$

$$[\hat{H}, \hat{\mathbf{L}}] = [\hat{H}, \hat{\mathbf{L}}^2] = 0.$$

(4.44)

Thus, the angular momentum will be conserved. The angular component of the wavefunction can be identified with the eigenstates of the angular momentum operator, $\hat{\mathbf{L}}$.

CHAPTER 5

Spin

Until we have focussed on the quantum mechanics of particles which are “featureless”, carrying no internal degrees of freedom. However, a relativistic formulation of quantum mechanics shows that particles can exhibit an intrinsic angular momentum component known as spin. However, the discovery of the spin degree of freedom marginally predates the development of relativistic quantum mechanics by Dirac and was achieved in a ground-breaking experiment by Stern and Gerlach (1922). In their experiment, they passed a well-collimated beam of silver atoms through a region of inhomogeneous field before allowing the particles to impact on a photographic plate (see Fig. 5.1). The magnetic field was directed perpendicular to the beam, and has a strong gradient, $\partial_z B_z \neq 0$ so that a beam comprised of atoms with a magnetic moment would be bent towards the z or $-z$ axis. As the magnetic moment will be proportional to the total angular momentum, such an experiment can be thought of as a measurement of its projection along z .

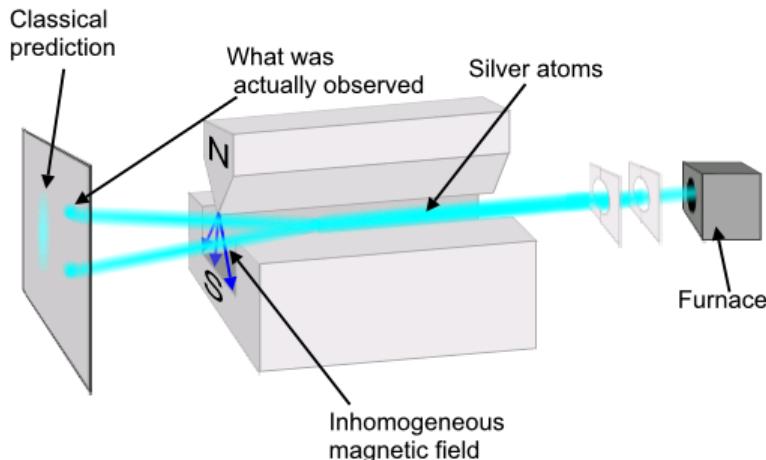


Fig. 5.1: The Stern-Gerlach experiment of 1922. Silver atoms travelling through an inhomogeneous magnetic field, and being deflected up or down depending on their spin. The screen revealed discrete points of accumulation, rather than a continuous distribution, owing to their quantised spin. Historically, this experiment was decisive in convincing physicists of the reality of angular-momentum quantisation in all atomic-scale systems.

At the time of the experiment, there was an expectation that the magnetic moment of the atom was generated in its entirety by the orbital angular momentum. Within that expectation, the result was qualitatively consistent with the model of Bohr – that orbits of the electron are quantised. Hence the congratulations on the postcard. However, as we have seen, the full quantum theory (of Schrödinger, developed after 1922) leads one to expect that the angular momentum is $\ell = 0, 1, 2, \dots$ so there should be a minimum of 3 possible values of l_z . As such, one would expect that there would be an odd number of possible z -components of angular momentum: $2\ell + 1$ with ℓ integer. Curiously, Stern and Gerlach’s experiment showed that the beam of silver atoms split into two! This discovery, which caused great discussion and surprise in subsequent years, presented a puzzle.

However, in our derivation of allowed angular momentum eigenvalues we found that,

although for any system the allowed values of m form a ladder with spacing \hbar , we could not rule out half-integral values of m . The lowest such case, $\ell = 1/2$, would in fact have just two allowed m values: $m = \pm 1/2$. However, such an ℓ value could not translate to an orbital angular momentum because the z -component of the orbital wavefunction, ψ has a factor $e^{im\phi} = e^{\pm i\phi/2}$, and therefore acquires a factor -1 on rotating through 2π ! This would imply that ϕ is not single-valued, which doesn't make sense for a Schrödinger-type wavefunction.

Yet the experimental result was irrefutable. Therefore, this must be a new kind of non-orbital angular momentum - spin. Conceptually, just as the Earth has orbital angular momentum in its yearly circle around the sun, and also spin angular momentum from its daily turning, the electron has an analogous spin. But this analogy has obvious limitations: the Earth's spin is, after all, made up of material orbiting around the axis through the poles. The electron spin cannot be imagined as arising from a rotating body, since orbital angular momenta always come in integral multiples of \hbar . Fortunately, this lack of a simple quasi-mechanical picture underlying electron spin doesn't prevent us from using the general angular momentum machinery developed earlier, which followed just from analysing the effect of spatial rotation on a quantum mechanical system.

5.1 Composite Systems

We often have to deal with systems with more than a single degree of freedom. This could be just a single particle moving in three dimensions rather than one, or a composite system such as a Hydrogen atom consisting of both an electron and a proton, or a diamond which consists of a very large number of carbon atoms. In the case of a macroscopic object, our system could be made up of a huge number of constituent parts, each able to behave differently, at least in principle.

It should be intuitively clear that the more complicated our system is, meaning the more independent degrees of freedom it possesses, the larger a Hilbert space we'll need if we wish to encode all its possible states. We'll now understand how quantum mechanics handles systems with more than one degree of freedom by taking the tensor product of the Hilbert spaces of its constituent parts.

5.1.1 Tensor Product of Hilbert Spaces

Hilbert spaces are vector spaces over \mathbb{C} , so we can try to define their tensor product in the same way as we would for any pair of vector spaces. If \mathcal{H}_1 and \mathcal{H}_2 are two, finite dimensional Hilbert spaces, with $\{|e_a\rangle\}_{a=1}^m$ a basis of \mathcal{H}_1 and $\{|f_\alpha\rangle\}_{\alpha=1}^n$ a basis of \mathcal{H}_2 , then the *tensor product* $\mathcal{H}_1 \otimes \mathcal{H}_2$ is again a vector space over \mathbb{C} spanned by all pairs of elements $|e_a\rangle \otimes |f_\alpha\rangle$ chosen from the two bases. It follows that

$$\dim(\mathcal{H}_1 \otimes \mathcal{H}_2) = \dim(\mathcal{H}_1)\dim(\mathcal{H}_2), \quad (5.1)$$

provided these are finite.

It's important to stress that, as with any vector space, a general element of $\mathcal{H}_1 \otimes \mathcal{H}_2$ is a *linear combination* of these basis elements. In particular, although general elements of \mathcal{H}_1 and \mathcal{H}_2 can be written

$$|\psi_1\rangle = \sum_{a=1}^m c_a |e_a\rangle \quad \text{and} \quad |\psi_2\rangle = \sum_{\alpha=1}^n d_\alpha |f_\alpha\rangle, \quad (5.2)$$

it is not true that a general element of $\mathcal{H}_1 \otimes \mathcal{H}_2$ necessarily takes the form $|\psi_1\rangle \otimes |\psi_2\rangle$. Rather, a general element of $\mathcal{H}_1 \otimes \mathcal{H}_2$ may be written as

$$|\Psi\rangle = \sum_{a,\alpha} r_{a\alpha} |e_a\rangle \otimes |f_\alpha\rangle. \quad (5.3)$$

In particular, elements of the form $|\psi_1\rangle \otimes |\psi_2\rangle \in \mathcal{H}_1 \otimes \mathcal{H}_2$ are specified by only $\dim(\mathcal{H}_1) + \dim(\mathcal{H}_2)$ complex coefficients, vastly fewer than is required to specify a generic element (5.3). Elements of the form $|\psi\rangle \otimes |\phi\rangle$ are sometimes called *simple*, while in physics generic elements of the form (5.3) are said to be *entangled*.

In order to make $\mathcal{H}_1 \otimes \mathcal{H}_2$ a Hilbert space, rather than just a vector space, we must give it an inner product. We do this in the obvious way: if $\langle \cdot | \rangle_1$ and $\langle \cdot | \rangle_2$ denote inner products on \mathcal{H}_1 and \mathcal{H}_2 , we first define the inner product between each pair of basis elements of $\mathcal{H}_1 \otimes \mathcal{H}_2$ by

$$(\langle e_a | \otimes \langle f_\alpha |)(|e_b\rangle \otimes |f_\beta\rangle) = \langle e_a | e_b \rangle_1 \langle f_\alpha | f_\beta \rangle_2, \quad (5.4)$$

and then extend it to arbitrary states $|\Psi\rangle$ by linearity. Note in particular that if $\{|e_a\rangle\}$ and $\{|f_\alpha\rangle\}$ are orthonormal bases of their respective Hilbert spaces, then $\{|e_a\rangle \otimes |f_\alpha\rangle\}$ will also be an orthonormal basis of $\mathcal{H}_1 \otimes \mathcal{H}_2$.

The most common occurrence of this is simply a single particle moving in more than one dimension. For example, suppose a quantum particle moves in \mathbb{R}^2 , described by Cartesian coordinates (x, y) . If $\{|x\rangle\}_{x \in \mathbb{R}}$ is a complete set of position eigenstates for the x -direction, and $\{|y\rangle\}_{y \in \mathbb{R}}$ likewise a complete set for the y -direction, then the state of the particle can be expanded as

$$|\psi\rangle = \int_{\mathbb{R} \times \mathbb{R}} \psi(x, y) |x\rangle \otimes |y\rangle \overbrace{\mathrm{d}x \mathrm{d}y}^{\mathrm{d}^2x}, \quad (5.5)$$

where $\{|x\rangle \otimes |y\rangle\}_{x,y \in \mathbb{R}}$ form a continuum basis of the tensor product space. Note that in general, our wavefunction is not a simple product $\psi(x, y) \neq \psi_1(x)\psi_2(y)$ of wavefunctions in the two directions separately, though it may be possible to write it as the sum of such products, as you met repeatedly when studying separation of variables. In this continuum case, the inner product between any two states is

$$\begin{aligned} \langle \chi | \psi \rangle &= \int_{\mathbb{R}^2 \times \mathbb{R}^2} \chi(x', y')^* \psi(x, y) \langle x' | x \rangle \langle y' | y \rangle \mathrm{d}x' \mathrm{d}y' \mathrm{d}x \mathrm{d}y \\ &= \int_{\mathbb{R}^2 \times \mathbb{R}^2} \chi(x', y')^* \psi(x, y) \delta(x - x') \delta(y - y') \mathrm{d}x' \mathrm{d}y' \mathrm{d}x \mathrm{d}y \\ &= \int_{\mathbb{R}^2} \chi(x, y)^* \psi(x, y) \mathrm{d}x \mathrm{d}y. \end{aligned} \quad (5.6)$$

This is exactly the structure of $L^2(\mathbb{R}^2, \mathrm{d}^2x)$, and we identify $L^2(\mathbb{R}^2, \mathrm{d}^2x) \cong L^2(\mathbb{R}, \mathrm{d}x) \otimes L^2(\mathbb{R}, \mathrm{d}y)$.

More commonly, physics considers quantum systems that live in \mathbb{R}^3 , described by a Hilbert space $\mathcal{H} \cong L^2(\mathbb{R}^3, d^3x)$ obtained from the tensor product of three copies of the Hilbert space for a one-dimensional system. We can expand a general state in this Hilbert space as

$$\psi = \int_{\mathbb{R}^3} \psi(x, y, z) |x\rangle \otimes |y\rangle \otimes |z\rangle dx dy dz = \int_{\mathbb{R}^3} \psi(\mathbf{x}) |\mathbf{x}\rangle d^3x, \quad (5.7)$$

where the last expression is just a convenient shorthand. Going further, to describe a composite system such as a Hydrogen atom that consists of two particles (an electron and proton), we need to take the tensor product of the individual Hilbert spaces of each particle. Thus (neglecting the spin of the electron and proton) the atom is described by a state

$$|\Psi\rangle = \int_{\mathbb{R}^6} \Psi(\mathbf{x}_e, \mathbf{x}_p) |\mathbf{x}_e\rangle \otimes |\mathbf{x}_p\rangle d^3x_e d^3x_p, \quad (5.8)$$

where $\Psi(\mathbf{x}_e, \mathbf{x}_p) = (\langle \mathbf{x}_e | \otimes \langle \mathbf{x}_p |) |\Psi\rangle$ is the amplitude for the atom's electron to be found at \mathbf{x}_e while its proton is at \mathbf{x}_p .

Once we know the Hilbert spaces appropriate to describe the fundamental constituents of our system, we can build up the Hilbert space for the combined system by taking tensor products. We should then ask “*What is the correct Hilbert space to use to describe the fundamental particles in our system?*”. Ultimately, this question can only be determined by carrying out an experiment. For example, experiments performed by Stern & Gerlach (see Fig. 5.1) showed that a single electron is in fact described by a Hilbert space $\mathcal{H}_e = L^2(\mathbb{R}^3, d^3x) \otimes \mathbb{C}^2$, formed from the tensor product of the electron's position space wavefunction with the two-dimensional Hilbert space \mathbb{C}^2 describing the electron's “internal” degrees of freedom. Thus, letting $\{|\uparrow\rangle, |\downarrow\rangle\}$ be an orthonormal basis of \mathbb{C}^2 , a generic state of the electron (whether or not it's part of a Hydrogen atom) is

$$|\Psi\rangle = |\phi\rangle \otimes |\uparrow\rangle + |\chi\rangle \otimes |\downarrow\rangle, \quad (5.9)$$

given in the position representation by

$$\langle \mathbf{x}_e | \Psi \rangle = \phi(\mathbf{x}_e) |\uparrow\rangle + \chi(\mathbf{x}_e) |\downarrow\rangle \quad (5.10)$$

involving a *pair* of wavefunctions $\phi, \chi \in L^2(\mathbb{R}^3, d^3x)$.

5.1.2 Operators in Composite Systems

Let's now understand how operators act on our composite system. Given linear operators $\hat{A} : \mathcal{H}_1 \rightarrow \mathcal{H}_1$ and $\hat{B} : \mathcal{H}_2 \rightarrow \mathcal{H}_2$, we define the linear operator $\hat{A} \otimes \hat{B} : \mathcal{H}_1 \otimes \mathcal{H}_2 \rightarrow \mathcal{H}_1 \otimes \mathcal{H}_2$ by

$$(\hat{A} \otimes \hat{B}) : |e_\alpha\rangle \otimes |f_\alpha\rangle \mapsto (\hat{A} |e_\alpha\rangle) \otimes (\hat{B} |f_\alpha\rangle), \quad (5.11)$$

and we extend this definition to arbitrary states in $\mathcal{H}_1 \otimes \mathcal{H}_2$ by linearity. In particular, the operator \hat{A} acting purely on \mathcal{H}_1 corresponds to the operator $\hat{A} \otimes \mathbb{I}_{\mathcal{H}_2}$ when acting on the tensor product, and similarly $\mathbb{I}_{\mathcal{H}_1} \otimes \hat{B}$ acts non-trivially just on the second factor in the tensor product. Note that since they act on separate Hilbert spaces,

$$[\hat{A} \otimes \mathbb{I}_{\mathcal{H}_2}, \mathbb{I}_{\mathcal{H}_1} \otimes \hat{B}] = 0, \quad (5.12)$$

for any operators \hat{A}, \hat{B} , even if these operators happen not to commute when acting on the same Hilbert space.

5.2 Spinors, Spin Operators, Pauli Matrices

The Hilbert space of angular momentum states for spin 1/2 is two-dimensional. Various notations are used: $|\ell, m\rangle$ becomes $|s, m\rangle$ or, more graphically,

$$|1/2, 1/2\rangle = |\uparrow\rangle, \quad |1/2, -1/2\rangle = |\downarrow\rangle. \quad (5.13)$$

A general state of spin can be written as the linear combination,

$$\alpha |\uparrow\rangle + \beta |\downarrow\rangle = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}, \quad (5.14)$$

with the normalisation condition, $|\alpha|^2 + |\beta|^2 = 1$, and this two-dimensional ket is called a **spinor**. Operators acting on spinors are necessarily of the form of 2×2 matrices. We shall adopt the usual practice of denoting the angular momentum components \hat{L}_i by \hat{S}_i for spins.

From our definition of the spinor, it is evident that the z -component of the spin can be represented as the matrix,

$$\hat{S}_z = \frac{\hbar}{2} \hat{\sigma}_z, \quad \hat{\sigma}_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (5.15)$$

From the general formulae (4.19) for raising and lowering operators $\hat{S}_{\pm} = \hat{S}_x \pm i\hat{S}_y$, with $s = 1/2$, we have $\hat{S}_+ |1/2, -1/2\rangle = \hbar |1/2, 1/2\rangle$, $\hat{S}_- |1/2, 1/2\rangle = \hbar |1/2, -1/2\rangle$, or, in matrix form,

$$\hat{S}_x + i\hat{S}_y = \hat{S}_+ = \hbar \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad \hat{S}_x - i\hat{S}_y = \hat{S}_- = \hbar \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}. \quad (5.16)$$

It therefore follows that an appropriate matrix representation for spin 1/2 is given by the **Pauli spin matrices**, $\hat{\mathbf{S}} = \frac{\hbar}{2} \hat{\boldsymbol{\sigma}} = (\hat{\sigma}_x, \hat{\sigma}_y, \hat{\sigma}_z)$ where

$$\hat{\sigma}_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \hat{\sigma}_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \hat{\sigma}_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (5.17)$$

The total spin $\hat{\mathbf{S}}^2 = \frac{\hbar^2}{4} \hat{\boldsymbol{\sigma}}^2 = \frac{3}{4} \hbar^2$, i.e. $s(s+1)\hbar^2$ for $s = 1/2$.

All three of the Pauli matrices can be compacted into a single expression,

$$\hat{\sigma}_j = \begin{pmatrix} \delta_{j3} & \delta_{j1} - i\delta_{j2} \\ \delta_{j1} + i\delta_{j2} & -\delta_{j3} \end{pmatrix} \quad (5.18)$$

This expression is useful for when any of the matrices (but no particular one) is to be used in algebraic manipulations, e.g. calculating the expectation values of the components of the magnetic moment. The matrices are also involutory, that is a square matrix which is its own inverse,

$$\hat{\sigma}_x^2 = \hat{\sigma}_y^2 = \hat{\sigma}_z^2 = -i\hat{\sigma}_x\hat{\sigma}_y\hat{\sigma}_z = \hat{\mathbb{I}}. \quad (5.19)$$

Another useful property, for inferring the wavefunctions for particles whose spin is aligned with the other cartesian axis, the corresponding eigenvectors

$$|\uparrow\rangle_x = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad |\downarrow\rangle_x = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \quad (5.20)$$

$$|\uparrow\rangle_y = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ i \end{pmatrix} \quad |\downarrow\rangle_y = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -i \end{pmatrix}. \quad (5.21)$$

5.3 Relating the Spinor to the Spin Direction

For a general state $\alpha|\uparrow\rangle + \beta|\downarrow\rangle$, how do α, β relate to which way the spin is pointing? To find out, let us assume that it is pointing up along the unit vector $\hat{\mathbf{n}} = (\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta)$, i.e. in the direction (θ, ϕ) . In other words, the spin is an eigenstate of the operator $\hat{\mathbf{n}} \cdot \hat{\boldsymbol{\sigma}}$ having eigenvalue unity:

$$\begin{pmatrix} n_z & n_x - in_y \\ n_x + in_y & -n_z \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}. \quad (5.22)$$

From this expression, we find that $\alpha/\beta = (n_x - in_y)/(1 - n_z) = e^{-i\phi} \cot \theta/2$. Then, making use of the normalisation, $|\alpha|^2 + |\beta|^2 = 1$, we obtain (up to an arbitrary phase)

$$\begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} e^{-i\phi/2} \cos \theta/2 \\ e^{i\phi/2} \sin \theta/2 \end{pmatrix}. \quad (5.23)$$

Since $e^{-i\phi} \cot \theta/2$ can be used to specify any complex number with $0 \leq \theta \leq \pi, 0 \leq \phi \leq 2\pi$, so for any possible spinor, there is an associated direction along which the spin points up.

5.3.0.1 The Spin Rotation Operator

In general, the rotation operator for rotation through an angle θ about an axis in the direction of the unit vector $\hat{\mathbf{n}}$ is given by $e^{-i\theta\hat{\mathbf{n}} \cdot \hat{\mathbf{J}}/\hbar}$ where $\hat{\mathbf{J}}$ denotes the angular momentum operator. For spin, $\hat{\mathbf{J}} = \hat{\mathbf{S}} = \frac{1}{2}\hbar\boldsymbol{\sigma}$, and the rotation operator takes the form¹ $e^{-i\theta\hat{\mathbf{n}} \cdot \hat{\mathbf{J}}/\hbar} = e^{-i(\theta/2)(\hat{\mathbf{n}} \cdot \hat{\boldsymbol{\sigma}})}$. Expanding the exponential, and making use of the Pauli matrix identities, $(\hat{\mathbf{n}} \cdot \hat{\boldsymbol{\sigma}})^2 = \hat{\mathbb{I}}$, one can show that

$$e^{-i(\theta/2)(\hat{\mathbf{n}} \cdot \hat{\boldsymbol{\sigma}})} = \hat{\mathbb{I}} \cos \theta/2 - i\hat{\mathbf{n}} \cdot \boldsymbol{\sigma} \sin \theta/2. \quad (5.24)$$

The rotation operator is a 2×2 matrix operating on the ket space. The 2×2 rotation matrices are unitary and form a group known as $SU(2)$; the 2 refers to the dimensionality, the U to their being unitary, and the S signifying determinant +1. Note that for rotation about the z -axis, $\hat{\mathbf{n}} = (0, 0, 1)$, it is more natural to replace θ with ϕ , and the rotation operator takes the form,

$$e^{-i(\theta/2)(\hat{\mathbf{n}} \cdot \boldsymbol{\sigma})} = \begin{pmatrix} e^{-i\phi/2} & 0 \\ 0 & e^{i\phi/2} \end{pmatrix}. \quad (5.25)$$

¹Warning: do not confuse θ - the rotation angle - with the spherical polar angle used to parameterise $\hat{\mathbf{n}}$.

In particular, the wavefunction is multiplied by -1 for a rotation of 2π . Since this is true for any initial wave function, it is clearly also true for rotation through 2π about any axis.

5.4 Spin Precession in a Magnetic Field

Consider a magnetised classical object spinning about its centre of mass, with angular momentum \mathbf{L} and parallel magnetic moment $\boldsymbol{\mu}_L$, $\boldsymbol{\mu}_L = \gamma_L \mathbf{L}$. The constant γ_L is called the *gyromagnetic ratio*, and defined as $\gamma_L = q/2m$. Now suppose that we impose a magnetic field \mathbf{B} along, say, the z -direction. This will exert a torque $\mathbf{T} = \boldsymbol{\mu}_L \times \mathbf{B} = \gamma_L \mathbf{L} \times \mathbf{B} = \frac{d\mathbf{L}}{dt}$. This equation is easily solved and shows that the angular momentum vector \mathbf{L} precesses about the magnetic field direction with angular velocity of precession $\boldsymbol{\omega}_0 = -\gamma_L \mathbf{B}$.²

In the following, we will show that precisely the same result appears in the study of the quantum mechanics of an electron spin in a magnetic field. The electron has an intrinsic magnetic dipole moment proportional to its spin, $\hat{\boldsymbol{\mu}}_S = \gamma_S \hat{\mathbf{S}}$, where $\gamma = g \frac{-e}{2m_e}$ and the gyromagnetic ratio, g , is very close to 2.³

The magnetic moment of a particle or system can also be specified via its *scalar* magnetic moment μ . μ is usually referred to simply at “the magnetic moment” and is the quantity usually listed in data tables. For spin 1/2 states, μ is usually defined via the spin-up state $|\uparrow\rangle$ as

$$\mu = \langle \uparrow | \hat{\mu}_z | \uparrow \rangle = \gamma_S \langle \uparrow | \hat{S}_z | \uparrow \rangle = \gamma_S \frac{\hbar}{2}. \quad (5.26)$$

For the electron, $\gamma_S = -g\mu_B/\hbar$, the scalar magnetic moment μ_e is thus

$$\mu_e = -\frac{g}{2}\mu_B. \quad (5.27)$$

In the approximation $g_e = 2$ the electron’s magnetic moment is simply

$$\mu_e = -\mu_B. \quad (5.28)$$

The interaction of *any* magnetic moment $\hat{\boldsymbol{\mu}}$ with an external magnetic field \mathbf{B} , is given by the Hamiltonian

$$\hat{H} = -\hat{\boldsymbol{\mu}} \cdot \mathbf{B}. \quad (5.29)$$

The Hamiltonian for the interaction of the electron’s dipole moment with the magnetic field is given by

$$\hat{H} = -\hat{\boldsymbol{\mu}}_S \cdot \mathbf{B} = -\gamma_S \hat{\mathbf{S}} \cdot \mathbf{B}. \quad (5.30)$$

For an electron moving in a magnetic field \mathbf{B} , the overall (orbital plus spin) interaction can then be written in the general form

$$\hat{H}_B = -\hat{\boldsymbol{\mu}}_L \cdot \mathbf{B} - \hat{\boldsymbol{\mu}}_S \cdot \mathbf{B} = -(-\gamma_L \hat{\mathbf{L}} + \gamma_S \hat{\mathbf{S}}) \cdot \mathbf{B}, \quad (5.31)$$

²Proof: From the equation of motion, with $L_+ = L_x + iL_y$, $\frac{dL_+}{dt} = -i\gamma_L BL_+$, $L_+ = L_+^0 e^{-i\gamma_L B t}$. Of course, $\frac{dL_z}{dt} = 0$, since $\frac{d\mathbf{L}}{dt} = \gamma_L \mathbf{L} \times \mathbf{B}$ is perpendicular to \mathbf{B} , which is in the z -direction.

³This g -factor terminology is used more widely. In general, the magnetic moment of an atom is written $\hat{\boldsymbol{\mu}}_S = -g\mu_B(\hat{\mathbf{J}}/\hbar)$, where $\mu_B = \frac{e\hbar}{2m_e}$ is known as the **Bohr magneton**, $\hat{\mathbf{J}} = \hat{\mathbf{L}} + \hat{\mathbf{S}}$ is the total (orbital and spin) angular momentum and g depends on the internal structure of the particular atom.

thus,

$$\hat{H}_B = \frac{e}{2m_e} (\hat{\mathbf{L}} + g\hat{\mathbf{S}}) \cdot \mathbf{B} = \frac{\mu_B}{\hbar} (\hat{\mathbf{L}} + g\hat{\mathbf{S}}) \cdot \mathbf{B}. \quad (5.32)$$

Hence the time development is specified by the equation $|\psi(t)\rangle = \hat{U}|\psi(0)\rangle$, with the time-evolution operator (or propagator), $\hat{U} = e^{-i\hat{H}t/\hbar} = e^{i\gamma_S \hat{\sigma} \cdot \mathbf{B} t/2}$. However, this is nothing but the rotation operator (as shown earlier) through an angle $-\gamma_S B t$ about the direction of \mathbf{B} !

For an arbitrary initial spin orientation

$$\begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} e^{-i\phi/2} \cos \theta/2 \\ e^{i\phi/2} \sin \theta/2 \end{pmatrix}, \quad (5.33)$$

the propagator for a magnetic field in the z -direction is given by

$$\hat{U}(t) = e^{i\gamma \hat{\sigma} \cdot \mathbf{B} t/2} = \begin{pmatrix} e^{-i\omega_s t/2} & 0 \\ 0 & e^{i\omega_s t/2} \end{pmatrix}, \quad (5.34)$$

so the time-dependent spinor is set by

$$\begin{pmatrix} \alpha(t) \\ \beta(t) \end{pmatrix} = \begin{pmatrix} e^{-i(\phi+\omega_s t)/2} \cos \theta/2 \\ e^{i(\phi+\omega_s t)/2} \sin \theta/2 \end{pmatrix}. \quad (5.35)$$

The angle θ between the spin and the field stays constant while the azimuthal angle around the field increases as $\phi = \phi_0 + \omega_s t$, exactly as in the classical case. The frequency $\omega_s = g \frac{|e|B}{2m_e}$ is the Larmor frequency for spin precession. For a magnetic field of 1T, $\omega_s \simeq 10^{11} \text{ rad s}^{-1}$.

Consider a particle with spin moving in a magnetic field \mathbf{B} , the time evolution of the expectation values of $\hat{\mathbf{S}}$ is given by Ehrenfest's theorem as

$$\frac{d}{dt} \langle \hat{\mathbf{S}} \rangle = \frac{i}{\hbar} \langle [\hat{H}, \hat{\mathbf{S}}] \rangle. \quad (5.36)$$

The only non-commuting contribution to \hat{H} is the spin interaction,

$$\hat{H} = -\hat{\mu}_S \cdot \mathbf{B} = -\gamma_S \hat{\mathbf{S}} \cdot \mathbf{B}. \quad (5.37)$$

The x -component of the commutator $[\hat{H}, \hat{\mathbf{S}}]$ above involves

$$\begin{aligned} [\mathbf{B} \cdot \hat{\mathbf{S}}, \hat{S}_x] &= B_y [\hat{S}_y, \hat{S}_x] + B_z [\hat{S}_z, \hat{S}_x] \\ &= i\hbar (-B_y \hat{S}_z + B_z \hat{S}_y) \\ &= i\hbar (\hat{\mathbf{S}} \times \mathbf{B})_x. \end{aligned} \quad (5.38)$$

And similarly for y and z , so $\langle \hat{\mathbf{S}} \rangle = (\langle \hat{S}_x \rangle, \langle \hat{S}_y \rangle, \langle \hat{S}_z \rangle)$ satisfies

$$\frac{d}{dt} \langle \hat{\mathbf{S}} \rangle = \gamma_S \langle \hat{\mathbf{S}} \rangle \times \mathbf{B}.$$

(5.39)

The solutions of this correspond to *precession* of the three-vector $\langle \hat{\mathbf{S}} \rangle$ of spin expectation values $\langle \hat{\mathbf{S}} \rangle = (\langle \hat{S}_x \rangle, \langle \hat{S}_y \rangle, \langle \hat{S}_z \rangle)$ around \mathbf{B} , at the *Larmor frequency* ω_s

$$\omega_s = \gamma_S B. \quad (5.40)$$

For example, for $\mathbf{B} = (0, 0, B)$, Eq. (5.39) the solutions have the form

$$\langle \hat{S}_x \rangle = A \cos(\omega_s t + \phi), \quad \langle \hat{S}_y \rangle = -A \sin(\omega_s t + \phi), \quad \langle \hat{S}_z \rangle = \text{const.} \quad (5.41)$$

Eq. (5.39) can also be written as $d/dt \langle \hat{\mathbf{S}} \rangle = \langle \hat{\mu}_S \rangle \times \mathbf{B}$. This is reminiscent of the equation $\dot{\mathbf{L}} = \mathbf{G} = \boldsymbol{\mu} \times \mathbf{B}$ which gives rise to precession in classical dynamics.

For a *spin-half* particle, the spin precession (Larmor) frequency is

$$\omega_s = \gamma_S B = \frac{g}{2} \frac{qB}{m} = \frac{2\mu B}{\hbar}. \quad (5.42)$$

This can also be expressed as

$$\omega_s = \frac{g}{2} \omega_c \quad \text{where} \quad \omega_c \equiv \frac{qB}{m} \quad (5.43)$$

is the *cyclotron frequency*. ω_c is the frequency with which particles moving transverse to a magnetic field \mathbf{B} undergo circular orbits, since $mv^2/r = qvB$ and $v = r\omega_c$.

The difference between the spin and cyclotron (orbit) frequencies is

$$\omega_a \equiv \omega_s - \omega_c = \frac{g - 2}{2} \frac{qB}{m}. \quad (5.44)$$

Precision determinations of the muon magnetic moment are based on measuring the frequency difference ω_a .

5.4.1 Energy Eigenstates

Orienting the z -axis along the magnetic field \mathbf{B} , the Hamiltonian is

$$\mathbf{B} = (0, 0, B), \quad \hat{H} = -\gamma_S B \hat{S}_z. \quad (5.45)$$

For a particle of spin s , the spin eigenstates $|s, m_s\rangle$ defined with the z -axis (the \mathbf{B} -field direction) as quantisation axis are also *energy* eigenstates

$$\hat{S}_z |s, m_s\rangle = m_s \hbar |s, m_s\rangle, \quad (5.46)$$

and so

$$\hat{H} |s, m_s\rangle = E_{m_s} |s, m_s\rangle, \quad E_{m_s} = -\gamma_S B m_s \hbar = -m_s \hbar \omega_s, \quad (5.47)$$

where ω_s is the Larmor frequency again $\omega_s = \gamma_S B$.

Thus, in a magnetic field, for spin s , we obtain a set of $2s + 1$ equally spaced energy levels, with energy spacing

$$\Delta E = \hbar \omega_s = \hbar \gamma_S B. \quad (5.48)$$

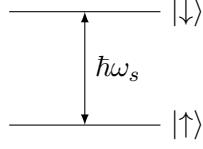


Fig. 5.2: For a particle of spin s , the spin eigenstates $|s, m_s\rangle$ defined with the z -axis (the \mathbf{B} -field direction) as quantisation axis are also *energy* eigenstates with $E_{m_s} = -m_s \hbar\omega_s$.

5.4.2 Wavefunction Evolution

For a particle which is initially in a spin eigenstate, $|\psi(0)\rangle = |s, m_s\rangle$, the spin state $|\psi(t)\rangle$ evolves as

$$|\psi(t)\rangle = e^{-iE_{m_s}t/\hbar} |s, m_s\rangle = e^{im_s\omega_st} |s, m_s\rangle. \quad (5.49)$$

The particle therefore remains in the state $|s, m_s\rangle$ at all times.

The expectation value of the spin z -component for the state $|\psi(t)\rangle$ is

$$\langle \psi(t) | \hat{S}_z | \psi(t) \rangle = \langle s, m_s | \hat{S}_z | s, m_s \rangle = m_s \hbar. \quad (5.50)$$

The expectation values of the x and y components *vanish*, as is easily seen by introducing the ladder operators \hat{S}_+ and \hat{S}_-

$$\langle \psi(t) | \hat{S}_{x,y} | \psi(t) \rangle \propto \langle s, m_s | (\hat{S}_+ \pm \hat{S}_-) | s, m_s \rangle = 0. \quad (5.51)$$

Thus, for an initial spin state $|\psi(0)\rangle = |s, m_s\rangle$, the vector of spin expectation values remains constant, and directed along z

$$\langle \psi(t) | \hat{\mathbf{S}} | \psi(t) \rangle = \langle s, m_s | \hat{\mathbf{S}} | s, m_s \rangle = (0, 0, m_s \hbar). \quad (5.52)$$

This is a special case, with no “visible” precession.

Now, generalising to an *arbitrary initial spin state*

$$|\psi(0)\rangle = \sum_{m_s} c_{m_s} |s, m_s\rangle, \quad (5.53)$$

in this case, the state $|\psi\rangle$ evolves with time as

$$|\psi(t)\rangle = \sum_{m_s} c_{m_s} e^{im_s\omega_st} |s, m_s\rangle. \quad (5.54)$$

The spin expectation values are now given by

$$\langle \psi(t) | \hat{\mathbf{S}} | \psi(t) \rangle = \sum_{m'_s, m_s} c_{m'_s}^* c_{m_s} e^{i(m_s - m'_s)\omega_st} \langle s, m'_s | \hat{\mathbf{S}} | s, m_s \rangle, \quad (5.55)$$

and the x and y components $\langle \hat{S}_x \rangle$, $\langle \hat{S}_y \rangle$ involve matrix elements of the form

$$\langle s, m'_s | \hat{S}_{x,y} | s, m_s \rangle \propto \langle s, m'_s | (\hat{S}_+ \pm \hat{S}_-) | s, m_s \rangle. \quad (5.56)$$

These matrix elements are non-zero for only $(m_s - m'_s) \pm 1$, giving a time dependence

$$e^{i(m_s - m'_s)} = e^{\pm i\omega_st}, \quad (5.57)$$

where again $\omega_s = \gamma_S B$. Hence $\langle \hat{S}_x \rangle$ and $\langle \hat{S}_y \rangle$ oscillate at the Larmor frequency $\omega_s = \gamma_S B$, as already found using Ehrenfest's theorem (Eq. (5.41))

$$\langle \hat{S}_{x,y} \rangle = A_{x,y} \cos(\omega_s t) + B_{x,y} \sin(\omega_s t). \quad (5.58)$$

For z , the matrix elements are non-zero only for $m'_s = m_s$:

$$\langle s, m'_s | \hat{S}_z | s, m_s \rangle = \delta_{m'_s, m_s} m_s \hbar. \quad (5.59)$$

Thus $\langle \hat{S}_z \rangle$ is constant (time-independent) as also shown in Eq. (5.41)

$$\langle \hat{S}_z \rangle = \langle \psi(t) | \hat{S}_z | \psi(t) \rangle = \sum_{m_s} |c_{m_s}|^2 m_s \hbar = \text{const.} \quad (5.60)$$

Thus, we again find that the vector $\langle \hat{\mathbf{S}} \rangle = \langle \psi(t) | \hat{\mathbf{S}} | \psi(t) \rangle$ of spin expectation values precesses around the magnetic field direction at the Larmor frequency $\omega_s = \gamma_S B$, matching the classical expectation.

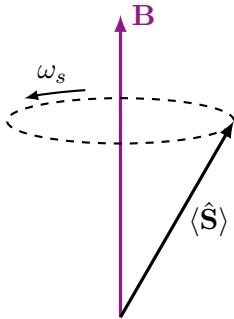


Fig. 5.3: $\langle \hat{\mathbf{S}} \rangle$ precesses about the magnetic field direction at the Larmor frequency, ω_s , matching the classical expectation.

5.5 Addition of Angular Momenta

In subsequent chapters, it will be necessary to add angular momentum, be it the addition of orbital and spin angular momenta, $\hat{\mathbf{J}} = \hat{\mathbf{L}} + \hat{\mathbf{S}}$, as with the study of spin-orbit coupling in atoms, or the addition of general angular momenta, $\hat{\mathbf{J}} = \hat{\mathbf{J}}_1 + \hat{\mathbf{J}}_2$ as occurs in the consideration of multi-electron atoms. In the following section, we will explore three problems: The addition of two spin 1/2 degrees of freedom; the addition of a general orbital angular momentum and spin; and the addition of spin $J = 1$ angular momenta. However, before addressing these examples in turn, let us first make some general remarks.

Without specifying any particular application, let us consider the total angular momentum $\hat{\mathbf{J}} = \hat{\mathbf{J}}_1 + \hat{\mathbf{J}}_2$ where $\hat{\mathbf{J}}_1$ and $\hat{\mathbf{J}}_2$ correspond to distinct degrees of freedom, $[\hat{\mathbf{J}}_1, \hat{\mathbf{J}}_2] = 0$, and the individual operators obey angular momentum commutation relations. As a result, the total angular momentum also obeys angular momentum commutation relations,

$$[\hat{J}_i, \hat{J}_j] = i\hbar\epsilon_{ijk}\hat{J}_k. \quad (5.61)$$

Regardless of the constituents of $\hat{\mathbf{J}}$, we would, on the basis of these commutators, come to the following conclusions:

- $\hat{\mathbf{J}}^2$ and \hat{J}_z are compatible observables;
- The eigenvalues of \hat{J}_z are $m_j \hbar$;
- The eigenvalues of $\hat{\mathbf{J}}^2$ are $j(j+1)\hbar^2$;
- m_j takes one of the $2j+1$ values $-j, -j+1, \dots, j-1, j$;
- j can take integer or half-integer values.

But how does this indexing relate to the indexing associated with the contributory orbital and spin parts? ℓ, m_ℓ are generated through ladder operators, s, m_s are generated through ladder operators, and j, m_j are generated through ladder operators.

How do ℓ, m_ℓ, s, m_s relate to j, m_j ? The quantisation conditions for \hat{J}_z are constrained by the quantisation conditions for \hat{L}_z and \hat{S}_z (see Fig. 5.4).

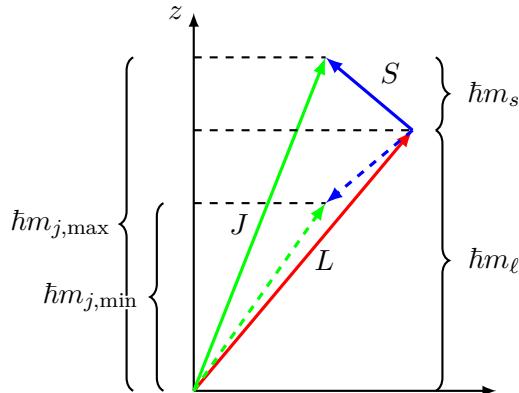


Fig. 5.4: Vector addition of orbital and spin angular momentum.

The procedure for combining angular momentum quantum numbers can be found by algebraic means, but the best insight comes from a simple diagram as illustrated by Fig. 5.5.

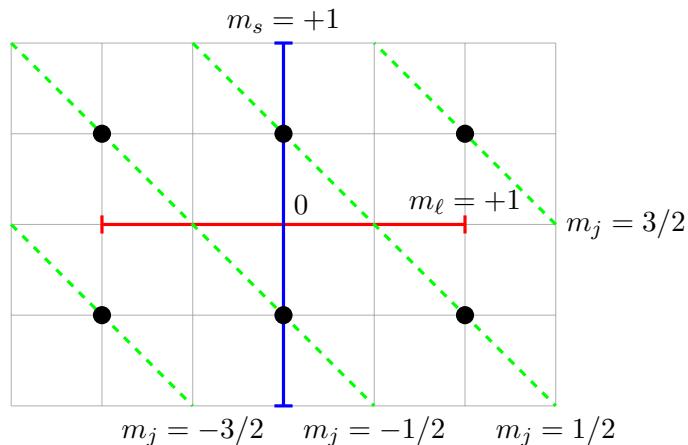


Fig. 5.5: Combining the orbital (red) and spin (blue) ladder operators to create the total (green) ladder operator. The $\ell = 1$ and $s = 1/2$ combinations are shown.

Since $\hat{J}_z = \hat{L}_z + \hat{S}_z$, then $m_j = m_\ell + m_s$ covers all possibilities. The maximum possible value for m_j is $m_{\ell,\max} + m_{s,\max}$, and therefore the maximum possible value for j is $\ell + s$. The maximum possible value of J^2 occurs when \mathbf{L} and \mathbf{S} are “maximally parallel”: Fig. 5.4 (green arrow).

For $j = j_{\max} = \ell + s$, m_j can take the following values

$$m_j = -(\ell + s), -(\ell + s) + 1, \dots, (\ell + s) - 1, (\ell + s). \quad (5.62)$$

The minimum possible value of J^2 occurs when \mathbf{L} and \mathbf{S} are “maximally antiparallel”: Fig. 5.4 (dashed green arrow), and so $j_{\min} = |\ell - s|$.

For $j = j_{\min} = |\ell - s|$, m_j can take the following values:

$$m_j = -|\ell - s|, -|\ell - s| + 1, \dots, |\ell - s| - 1, |\ell - s|. \quad (5.63)$$

The allowed values for the total angular momentum operator j are

$$j = (\ell + s), (\ell + s) - 1, \dots, |\ell - s| + 1, |\ell - s|. \quad (5.64)$$

There are $2s + 1$ or $2\ell + 1$ such values of j , whichever is fewer. For example:

$$\ell = 3, s = 1 \begin{cases} j = 4 & m_j = -4, -3, -2, -1, 0, 1, 2, 3, 4 \\ j = 3 & m_j = -3, -2, -1, 0, 1, 2, 3 \\ j = 2 & m_j = -2, -1, 0, 1, 2 \end{cases}. \quad (5.65)$$

Taking $s \leq \ell$, the total number of states is

$$\sum_{\ell=1}^{\ell+s} (2\ell + 1) = (2\ell + 1)(2s + 1). \quad (5.66)$$

For each angular momentum component, the states $|j_1, m_1\rangle$ and $|j_2, m_2\rangle$ where $m_i = -j_i, \dots, j_i$, provide a basis of states of the total angular momentum operator, $\hat{\mathbf{J}}_i^2$ and the projection \hat{J}_{iz} . Together, they form a complete basis which can be used to span the states of the coupled spins.⁴

$$|j_1, m_1, j_2, m_2\rangle \equiv |j_1, m_1\rangle \otimes |j_2, m_2\rangle. \quad (5.67)$$

These product states are also eigenstates of \hat{J}_z with eigenvalue $\hbar(m_1 + m_2)$, but not of $\hat{\mathbf{J}}^2$. However, for practical application, we require a basis in which the total angular momentum operator $\hat{\mathbf{J}}^2$ is also diagonal. That is, we must find eigenstates $|j, m_j, j_1, j_2\rangle$ of the four mutually commuting operators $\hat{\mathbf{J}}^2$, \hat{J}_z , $\hat{\mathbf{J}}_1^2$, and $\hat{\mathbf{J}}_2^2$.

In general, the relation between the two basis can be expressed as

$$\begin{aligned} |j, m_j, j_1, j_2\rangle &= \sum_{m_1, m_2} C_{j_1, m_1, j_2, m_2} |j_1, m_1\rangle \otimes |j_2, m_2\rangle \\ &= \sum_{m_1, m_2} |j_1, m_1, j_2, m_2\rangle \langle j_1, m_1, j_2, m_2 | j, m_j, j_1, j_2 \rangle, \end{aligned} \quad (5.68)$$

⁴Here \otimes denotes the “direct product” and shows that the two constituent spin states access their own independent Hilbert space.

where the matrix elements C_{j_1, m_1, j_2, m_2} are known as **Clebsch-Gordan coefficients**. In general, the determination of these coefficients from first principles is a somewhat soul destroying exercise and one that we do not intend to pursue in great detail.⁵ In any case, for practical purposes, such coefficients have been tabulated in the literature and can be readily obtained. However, in some simple cases, these matrix elements can be determined straightforwardly. Moreover, the algorithmic programme by which they are deduced offer some new conceptual insights.

Operationally, the mechanism for finding the basis states of the total angular momentum operator follow the strategy:

1. As a unique entry, the basis state with maximal J_{\max} and $m_j = J_{\max}$ is easy to deduce from the original basis states since it involves the product of states of **highest weight**,

$$|J_{\max}, m_j = J_{\max}, j_1, j_2\rangle = |j_1, m_1 = j_1\rangle \otimes |j_2, m_2 = j_2\rangle, \quad (5.70)$$

where $J_{\max} = j_1 + j_2$.

2. From this state, we can use of the total spin lowering operator \hat{J}_- to find all states with $J = J_{\max}$ and $m_j = -J_{\max}, \dots, J_{\max}$.
3. From the state with $J = J_{\max}$ and $m_j = J_{\max} - 1$, one can then obtain the state with $J = J_{\max} - 1$ and $m_j = J_{\max} - 1$ by orthogonality. Now one can return to the second step of the programme and repeat until $J = |j_1 - j_2|$ when all $(2j_1 + 1)(2j_2 + 1)$ basis states have been obtained.

5.5.1 Addition of Two Spin 1/2 Degrees of Freedom

For two spin 1/2 degrees of freedom, we could simply construct and diagonalise the complete 4×4 matrix elements of the total spin. However, to gain some intuition for the general case, let us consider the programme above. Firstly, the maximal total spin state is given by

$$|S = 1, m_S = 1, s_1 = 1/2, s_2 = 1/2\rangle = |s_1 = 1/2, m_{s_1} = 1/2\rangle \otimes |s_2 = 1/2, m_{s_2} = 1/2\rangle. \quad (5.71)$$

Now, since $s_1 = 1/2$ and $s = 1/2$ is implicit, we can rewrite this equation in a more colloquial form as

$$|S = 1, m_S = 1\rangle = |\uparrow_1\rangle \otimes |\uparrow_2\rangle. \quad (5.72)$$

We now follow step 2 of the programme and subject the maximal spin state to the total spin lowering operator, $\hat{S}_- = \hat{S}_1^- + \hat{S}_2^-$. In doing so, making use of Eq. (4.19), we find

$$\hat{S}_- |S = 1, m_S = 1\rangle = \sqrt{2}\hbar |S = 1, m_S = 0\rangle = \hbar(|\downarrow_1\rangle \otimes |\uparrow_2\rangle + |\uparrow_1\rangle \otimes |\downarrow_2\rangle), \quad (5.73)$$

⁵In fact, one may show that the general matrix element is given by

$$\begin{aligned} \langle j_1, m_1, j_2, m_2 | j, m_j, j_1, j_2 \rangle &= \delta_{m_j, m_1+m_2} \sqrt{\frac{(j_1 + j_2 - j)!(j + j_1 - j_2)!(j + j_2 - j_1)!(2j + 1)}{(j + j_1 + j_2 + 1)!}} \\ &\times \sum_k \frac{(-1)^k \sqrt{(j_1 + m_1)!(j_1 - m_1)!(j_2 + m_2)!(j_2 - m_2)!(j + m)!(j - m)!}}{k!(j_1 + j_2 - j - k)!(j_1 - m_1 - k)!(j_2 + m_2 - k)!(j - j_2 + m_1 + k)!(j - j_1 - m_2 + k)!}. \end{aligned} \quad (5.69)$$

i.e. $|S = 1, m_S = 0\rangle = \frac{1}{\sqrt{2}}(|\downarrow_1\rangle \otimes |\uparrow_2\rangle + |\uparrow_1\rangle \otimes |\downarrow_2\rangle)$. Similarly,

$$\hat{S}_- |S = 1, m_S = 0\rangle = \sqrt{2}\hbar |S = 1, m_S = -1\rangle = \sqrt{2}\hbar |\uparrow_1\rangle \otimes |\uparrow_2\rangle, \quad (5.74)$$

i.e. $|S = 1, m_S = -1\rangle = |\uparrow_1\rangle \otimes |\uparrow_2\rangle$. This completes the construction of the manifold of spin $S = 1$ states - the **spin triplet** states. Following the programme, we must now consider the lower spin state.

In this case, the next multiplet is the unique total spin singlet state $|S = 0, m_S = 0\rangle$. The latter must be orthogonal to the spin triplet state $|S = 1, m_S = 0\rangle$. As a result, we can deduce that

$$|S = 0, m_S = 0\rangle = \frac{1}{\sqrt{2}}(|\downarrow_1\rangle \otimes |\uparrow_2\rangle + |\uparrow_1\rangle \otimes |\downarrow_2\rangle). \quad (5.75)$$

5.5.2 Addition of Angular Momentum and Spin

We now turn to the problem of the addition of angular momentum and spin, $\hat{\mathbf{J}} = \hat{\mathbf{L}} + \hat{\mathbf{S}}$. In the original basis, for a given angular momentum ℓ , one can identify $2 \times (2\ell + 1)$ product states $|\ell, m_\ell\rangle \otimes |\uparrow\rangle$ and $|\ell, m_\ell\rangle \otimes |\downarrow\rangle$, with $m_\ell = -\ell, \dots, \ell$, involving eigenstates of $\hat{\mathbf{L}}^2$, \hat{L}_z , $\hat{\mathbf{S}}^2$ and \hat{S}_z , but not $\hat{\mathbf{J}}^2$. From these basis states, we are looking for eigenstates of $\hat{\mathbf{J}}^2$, \hat{J}_z , $\hat{\mathbf{L}}^2$ and $\hat{\mathbf{S}}^2$. To undertake this programme, it is helpful to recall the action of the angular momentum raising and lower operators,

$$\hat{L}_\pm |\ell, m_\ell\rangle = ((\ell \pm m_\ell + 1)(\ell \mp m_\ell))^{1/2} \hbar |\ell, m_\ell \pm 1\rangle \quad (5.76)$$

as well as the identity

$$\hat{\mathbf{J}}^2 = \hat{\mathbf{L}}^2 + \hat{\mathbf{S}}^2 + \overbrace{2\hat{L}_z\hat{S}_z + \hat{L}_+\hat{S}_- + \hat{S}_+\hat{L}_-}^{2\hat{\mathbf{L}}\cdot\hat{\mathbf{S}}}. \quad (5.77)$$

For the eigenstates of $\hat{\mathbf{J}}^2$, \hat{J}_z , $\hat{\mathbf{L}}^2$ and $\hat{\mathbf{S}}^2$ we will adopt the notation $|j, m_j, \ell\rangle$ leaving the spin $S = 1/2$ implicit. The maximal spin state is given by⁶

5.5.3 Addition of two angular momenta $J = 1$

As mentioned above, for the general case the programme is algebraically technical and unrewarding. However, for completeness, we consider here the explicit example of the addition of two spin 1 degrees of freedom. Once again, the maximal spin state is given by

$$|J = 2, m_J = 2, j_1 = 1, j_2 = 1\rangle = |j_1 = 1, m_1 = 1\rangle \otimes |j_2 = 1, m_2 = 1\rangle, \quad (5.81)$$

⁶The proof runs as follows:

$$\hat{J}_z |\ell, \ell\rangle \otimes |\uparrow\rangle = (\hat{L}_z + \hat{S}_z) |\ell, \ell\rangle \otimes |\uparrow\rangle = (\ell + 1/2)\hbar |\ell, \ell\rangle \otimes |\uparrow\rangle, \quad (5.78)$$

and

$$\hat{\mathbf{J}}^2 |\ell, \ell\rangle \otimes |\uparrow\rangle = \hbar^2(\ell(\ell + 1) + \frac{1}{2}(\frac{1}{2} + 1) + 1 + 2\ell\frac{1}{2}) |\ell, \ell\rangle \otimes |\uparrow\rangle \quad (5.79)$$

$$= \hbar^2(\ell + \frac{1}{2})(\ell + \frac{3}{2}) |\ell, \ell\rangle \otimes |\uparrow\rangle. \quad (5.80)$$

or, more concisely, $|2, 2\rangle = |1\rangle \otimes |1\rangle$, where we leave j_1 and j_2 implicit. Once again, making use of Eq. (4.19) and an economy of notation, we find

$$\begin{cases} |2, 2\rangle = |1\rangle \otimes |1\rangle \\ |2, 1\rangle = \frac{1}{\sqrt{2}}(|0\rangle \otimes |1\rangle + |1\rangle \otimes |0\rangle) \\ |2, 0\rangle = \frac{1}{\sqrt{6}}(|-1\rangle \otimes |1\rangle + 2|0\rangle \otimes |0\rangle + |1\rangle \otimes |-1\rangle) \\ |2, -1\rangle = \frac{1}{\sqrt{2}}(|0\rangle \otimes |-1\rangle + |-1\rangle \otimes |0\rangle) \\ |2, -2\rangle = |-1\rangle \otimes |-1\rangle \end{cases}. \quad (5.82)$$

Then, from the expression for $|2, 1\rangle$, we can construct the next maximal spin state $|1, 1\rangle = \frac{1}{\sqrt{2}}(|0\rangle \otimes |1\rangle - |1\rangle \otimes |0\rangle)$, from the orthogonality condition. Once again, acting on this state with the total spin lowering operator, we obtain the remaining members of the multiplet,

$$\begin{cases} |1, 1\rangle = \frac{1}{\sqrt{2}}(|0\rangle \otimes |1\rangle - |1\rangle \otimes |0\rangle) \\ |1, 0\rangle = \frac{1}{\sqrt{2}}(|-1\rangle \otimes |1\rangle - |1\rangle \otimes |-1\rangle) \\ |1, -1\rangle = \frac{1}{\sqrt{2}}(|-1\rangle \otimes |0\rangle - |0\rangle \otimes |-1\rangle) \end{cases} \quad (5.83)$$

Finally, finding the state orthogonal to $|1, 0\rangle$ and $|1, 0\rangle$, we obtain the final state,

$$|0, 0\rangle = \frac{1}{\sqrt{3}}(|-1\rangle \otimes |1\rangle - |0\rangle \otimes |0\rangle + |1\rangle \otimes |-1\rangle). \quad (5.84)$$

CHAPTER 6

Motion in a Magnetic Field

Hitherto, we have focussed on applications of quantum mechanics to free particles or particles confined by scalar potentials. In the following, we will address the influence of a magnetic field on a charged particle. Classically, the force on a charged particle in an electric and magnetic field is specified by the **Lorentz force law**:

$$\mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}), \quad (6.1)$$

where q denotes the charge and \mathbf{v} the velocity. (Here we will adopt a convention in which q denotes the charge (which may be positive or negative) and $e \equiv |e|$ denotes the modulus of the electron charge, i.e. for an electron, the charge $q_e = -e = -1.602176487 \times 10^{-19} \text{ C}$.) The velocity-dependent force associated with the magnetic field is quite different from the conservative forces associated with scalar potentials, and the programme for transferring from classical to quantum mechanics – replacing momenta with the appropriate operators – has to be carried out with more care. As preparation, it is helpful to revise how the Lorentz force arises in the Lagrangian formulation of classical mechanics.

6.1 Classical Mechanics of a Particle in a Field

For a system with m degrees of freedom specified by coordinates q_1, \dots, q_m , the classical action is determined from the Lagrangian $L(q_i, \dot{q}_i)$ by

$$S[q_i] = \int dt L(q_i, \dot{q}_i) \quad (6.2)$$

The action is said to be a **functional** of the coordinates $q_i(t)$. According to **Hamilton's extremal principle** (also known as the **principle of least action**), the dynamics of a classical system is described by the equations that minimise the action. These equations of motion can be expressed through the classical Lagrangian in the form of the Euler-Lagrange equations,

$$\frac{d}{dt} (\partial_{\dot{q}_i} L(q_i, \dot{q}_i)) - \partial_{q_i} L(q_i, \dot{q}_i) = 0. \quad (6.3)$$

According to Hamilton's extremal principle, for any smooth set of curves $\omega_i(t)$, the variation of the action around the classical solution $q_i(t)$ is zero, i.e. $\lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} (S[q_i + \epsilon w_i] - S[q_i]) = 0$. Applied to the action, the variation implies that, for any i , $\int dt (w_i \partial_i L(q_i, \dot{q}_i) + \dot{w}_i \partial_{\dot{q}_i} L(q_i, \dot{q}_i)) = 0$. Then, integrating the second term by parts, and dropping the boundary term, one obtains

$$\int dt \left(\partial_i L(q_i, \dot{q}_i) + \frac{d}{dt} \partial_{\dot{q}_i} L(q_i, \dot{q}_i) \right) = 0. \quad (6.4)$$

Since this equality must follow for any function $w_i(t)$, the term in parentheses in the integrand must vanish leading to the Euler-Lagrange equation (6.3).

The **canonical momentum** is specified by the equation $p_i = \partial_{\dot{q}_i} L$, and the classical Hamiltonian is defined by the Legendre transform,

$$H(q_i, p_i) = \sum_i p_i \dot{q}_i - L(q_i, \dot{q}_i). \quad (6.5)$$

It is straightforward to check that the equations of motion can be written in the form of Hamilton's equations of motion,

$$\dot{q}_i = \partial_{p_i} H, \quad \dot{p}_i = -\partial_{q_i} H. \quad (6.6)$$

From these equations it follows that, if the Hamiltonian is independent of a particular coordinate q_i , the corresponding momentum p_i remains constant. For **conservative forces**,¹ the classical Lagrangian and Hamiltonian can be written as $L = T - V$, $H = T + V$, with T the kinetic energy and V the potential energy.

Any dynamical variable f in the system is some function of the phase space coordinates, the q_i s and p_i s, and (assuming it does not depend explicitly on time) its time-development is given by:

$$\frac{d}{dt} f(q_i, p_i) = \partial_{q_i} f \dot{q}_i + \partial_{p_i} f \dot{p}_i = \partial_{q_i} f \partial_{p_i} H - \partial_{p_i} f \partial_{q_i} H \equiv \{f, H\}. \quad (6.7)$$

The curly brackets are known as Poisson brackets, and are defined for any dynamical variables as $\{A, B\} = \partial_{q_i} A \partial_{p_i} B - \partial_{p_i} A \partial_{q_i} B$. From Hamilton's equations, we have shown that for any variable, $\dot{f} = \{f, H\}$. It is easy to check that, for the coordinates and canonical momenta, $\{q_i, q_j\} = 0 = \{p_i, p_j\}$, $\{q_i, p_j\} = \delta_{ij}$. This was the classical mathematical structure that led Dirac to link up classical and quantum mechanics: He realised that the Poisson brackets were the classical version of the commutators, so a classical canonical momentum must correspond to the quantum differential operator in the corresponding coordinate.

With these foundations revised, we now return to the problem at hand; the influence of an electromagnetic field on the dynamics of the charged particle.

As the Lorentz force is velocity dependent, it can not be expressed simply as the gradient of some potential. Nevertheless, the classical path traversed by a charged particle is still specified by the principle of least action. The electric and magnetic fields can be written in terms of a scalar and a vector potential as $\mathbf{B} = \nabla \times \mathbf{A}$, $\mathbf{E} = -\nabla\phi - \dot{\mathbf{A}}$. The corresponding Lagrangian takes the form:²

$$L = \frac{1}{2} m \mathbf{v}^2 - q\phi + q\mathbf{v} \cdot \mathbf{A}. \quad (6.8)$$

In this case, the general coordinates $q_i \equiv x_i = (x_1, x_2, x_3)$ are just the Cartesian coordinates specifying the position of the particle, and the \dot{q}_i are the three components

¹i.e. forces that conserve mechanical energy

²In a relativistic formulation, the interaction term here looks less arbitrary: the relativistic version would have the relativistically invariant $q \int A^\mu dx_\mu$ added to the action integral, where the four-potential $A_\mu = (\phi, \mathbf{A})$ and $dx_\mu = (ct, dx_1, dx_2, dx_3)$. This is the simplest possible invariant interaction between the electromagnetic field and the particle's four-velocity. Then, in the non-relativistic limit, $q \int A^\mu dx_\mu$ just becomes $q \int \mathbf{v} \cdot \mathbf{A} - \phi dt$.

$\dot{x}_i = (\dot{x}_1, \dot{x}_2, \dot{x}_3)$ of the particle velocities. The important point is that the *canonical* momentum

$$p_i = \partial_{\dot{x}_i} L = mv_i + qA_i, \quad (6.9)$$

is no longer simply given by the mass \times velocity – there is an extra term!

Making use of the definition (6.5), the corresponding Hamiltonian is given by

$$H(q_i, p_i) = \sum_i (mv_i + qA_i)v_i - \frac{1}{2}m\mathbf{v}^2 + q\phi - q\mathbf{v} \cdot \mathbf{A} = \frac{1}{2}m\mathbf{v}^2 + q\phi. \quad (6.10)$$

Reassuringly, the Hamiltonian just has the familiar form of the sum of the kinetic and potential energy. However, to get Hamilton's equations of motion, the Hamiltonian has to be expressed solely in terms of the coordinates and canonical momenta, i.e.

$$H = \frac{1}{2m}(\mathbf{p} - q\mathbf{A}(\mathbf{r}, t))^2 + q\phi(\mathbf{r}, t). \quad (6.11)$$

Let us now consider Hamilton's equations of motion, $\dot{x}_i = \partial_{p_i} H$ and $\dot{p}_i = -\partial_{x_i} H$. The first equation recovers the expression for the canonical momentum while second equation yields the Lorentz force law. To understand how, we must first keep in mind that dp/dt is not the acceleration: The A -dependent term also varies in time, and in a quite complicated way, since it is the field at a point moving with the particle. More precisely,

$$\dot{p}_i = m\ddot{x}_i + q\dot{A}_i = m\ddot{x}_i + q(\partial_t A_i + v_j \partial_{x_j} A_i), \quad (6.12)$$

where we have assumed a summation over repeated indices. The right-hand side of the second of Hamilton's equation, $\dot{p}_i = -\frac{\partial H}{\partial x_i}$, is given by

$$-\partial_{x_i} H = \frac{1}{m}(\mathbf{p} - q\mathbf{A}(\mathbf{r}, t))q\partial_{x_i} \mathbf{A} - q\partial_{x_i}\phi(\mathbf{r}, t) = qv_j \partial_{x_i} A_j - q\partial_{x_i}\phi. \quad (6.13)$$

Together, we obtain the equation of motion, $m\ddot{x}_i = -q(\partial_t A_i + v_j \partial_{x_j} A_i) + qv_j \partial_{x_i} A_j - q\partial_{x_i}\phi$. Using the identity $\mathbf{v} \times (\nabla \times \mathbf{A}) = \nabla(\mathbf{v} \cdot \mathbf{A}) - (\mathbf{v} \cdot \nabla)\mathbf{A}$, and the expressions for the electric and magnetic fields in terms of the potentials, one recovers the Lorentz equation

$$m\ddot{x}_i = \mathbf{F} = q(\mathbf{E} + \mathbf{v} \times \mathbf{B}). \quad (6.14)$$

With these preliminary discussions of the classical system in place, we are now in a position to turn to the quantum mechanic.

6.2 Quantum Mechanics of a Particle in a Field

To transfer to the quantum mechanical regime, we must implement the canonical quantisation procedure setting $\hat{\mathbf{p}} = -i\hbar\nabla$, so that $[\hat{x}_i, \hat{p}_j] = i\hbar\delta_{ij}$. However, in this case, $\hat{p}_i \neq m\hat{v}_i$. This leads to the novel situation that the velocities in different directions do not commute.³ To explore influence of the magnetic field on the particle dynamics, it is helpful to assess the relative weight of the A -dependent contributions to the quantum Hamiltonian,

$$\hat{H} = \frac{1}{2m}(\mathbf{p} - q\mathbf{A}(\mathbf{r}, t))^2 + q\phi(\mathbf{r}, t). \quad (6.15)$$

³With $m\hat{v}_i = -i\hbar\partial_{x_i} - qA_i$, it is easy to verify that $[\hat{v}_x, \hat{v}_y] = \frac{i\hbar q}{m^2}B$.

6.3 Gauge Invariance and the Aharonov-Bohm Effect

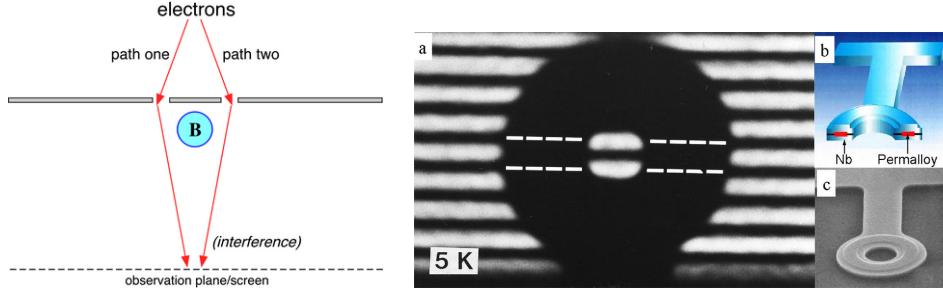


Fig. 6.1: (Left) Schematic showing the geometry of an experiment to observe the Aharonov-Bohm effect. Electrons from a coherent source can follow two paths which encircle a region where the magnetic field is non-zero. (Right) Interference fringes for electron beams passing near a toroidal magnet from the experiment by Tonomura and collaborators in 1986. The electron beam passing through the center of the torus acquires an additional phase, resulting in fringes that are shifted with respect to those outside the torus, demonstrating the Aharonov-Bohm effect.

Our derivation above shows that the quantum mechanical Hamiltonian of a charged particle is defined in terms of the vector potential, \mathbf{A} . Since the latter is defined only up to some gauge choice, this suggests that the wavefunction is not a gauge invariant object. Indeed, it is only the observables associated with the wavefunction which must be gauge invariant. To explore this gauge freedom, let us consider the influence of the **gauge transformation**,

$$\mathbf{A} \mapsto \mathbf{A}' = \mathbf{A} + \nabla\Lambda, \quad \phi \mapsto \phi' = \phi - \partial_t\Lambda, \quad (6.16)$$

where $\Lambda(\mathbf{r}, t)$ denotes a scalar function.

Is Schrödinger's equation invariant under gauge transformations too? In terms of the gauge-transformed potentials, \mathbf{A}' and ϕ' , Eq. (6.15) reads

$$i\hbar \frac{\partial \psi}{\partial t} = \frac{1}{2m} [-i\hbar \nabla - q\mathbf{A}' - q(\nabla\Lambda)]^2 \psi + q\phi'\psi + q\frac{\partial\chi}{\partial t}\psi. \quad (6.17)$$

This does *not* look like the original Schrödinger equation...

The only hope to restore gauge invariance is to transform the wavefunction ψ as well as the potentials \mathbf{A} and ϕ :

$$\mathbf{A} \mapsto \mathbf{A}' = \mathbf{A} + \nabla\Lambda, \quad \phi \mapsto \phi' = \phi - \partial_t\Lambda, \quad \psi \mapsto \psi'. \quad (6.18)$$

To leave the probability density unchanged, $|\psi'|^2 = |\psi|^2$, the only realistic choice is to multiply by a phase factor:

$$\psi'(\mathbf{r}, t) = \exp \left[i \frac{q}{\hbar} \Lambda(\mathbf{r}, t) \right] \psi(\mathbf{r}, t) \quad (6.19)$$

to cancel the additional terms that appeared in the Schrödinger equation.

Setting $\psi = \psi' e^{-iq\Lambda(\mathbf{r}, t)/\hbar}$ in Eq. (6.17), the gauge-transformed Schrödinger equation reads

$$i\hbar \frac{\partial}{\partial t} \left(e^{-iq\Lambda(\mathbf{r}, t)/\hbar} \psi' \right) = \left[\frac{1}{2m} (-i\hbar \nabla - q\mathbf{A}' + q(\nabla\Lambda))^2 + q\phi' + q\partial_t\Lambda \right] \left(e^{-iq\Lambda(\mathbf{r}, t)/\hbar} \psi' \right). \quad (6.20)$$

Let's look first at the action of $-i\hbar\nabla$ on $\left(e^{-iq\Lambda(\mathbf{r},t)/\hbar}\psi'\right)$:

$$\begin{aligned} -i\hbar(e^{-iq\Lambda(\mathbf{r},t)/\hbar}\psi') &= -qe^{iq\Lambda(\mathbf{r},t)/\hbar}(\nabla\Lambda)\psi' - i\hbar e^{-iq\Lambda(\mathbf{r},t)/\hbar}(\nabla\psi') \\ &= e^{-iq\Lambda(\mathbf{r},t)/\hbar}(-q(\nabla\Lambda) - i\hbar\nabla)\psi'. \end{aligned} \quad (6.21)$$

Therefore, the following term becomes

$$(-i\hbar\nabla - q\mathbf{A}' + q(\nabla\Lambda))\left(e^{-iq\Lambda(\mathbf{r},t)/\hbar}\psi'\right) = e^{-iq\Lambda(\mathbf{r},t)/\hbar}[-i\hbar\nabla - q\mathbf{A}']\psi', \quad (6.22)$$

which, as it holds for any ψ' , gives us the operator relation $(\dots)e^{-iq\Lambda(\mathbf{r},t)/\hbar} = e^{-iq\Lambda(\mathbf{r},t)/\hbar}[\dots]$.

Using the previous operator relation twice,

$$e^{-iq\Lambda(\mathbf{r},t)/\hbar}\left(q\frac{\partial\Lambda}{\partial t} + i\hbar\frac{\partial}{\partial t}\right) = e^{-iq\Lambda(\mathbf{r},t)/\hbar}\left[\frac{1}{2m}(-i\hbar\nabla - q\mathbf{A}')^2 + q\phi' + q\frac{\partial\Lambda}{\partial t}\right]\psi' \quad (6.23)$$

and Eq. (6.20) simplifies to

$$i\hbar\frac{\partial\psi'}{\partial t} = \frac{1}{2m}[-i\hbar\nabla - q\mathbf{A}']^2\psi' + q\phi'\psi'. \quad (6.24)$$

This is now of the same form as the original Schrödinger equation, Eq. (6.15), but with the gauge-transformed potentials \mathbf{A}' and ϕ' , and the transformed wavefunction ψ' .

Gauge invariance is of importance not just in electromagnetism, but more generally to the description of particle interactions in the Standard Model of particle physics. Gauge invariance is promoted to a fundamental underlying symmetry principle which determines the form of the EM, weak and strong interactions.

The gauge transformation introduces an additional space and time-dependent phase factor into the wavefunction. However, since the observable translates to the probability density, $|\phi|^2$, this phase dependence seems invisible.

One physical manifestation of the gauge invariance of the wavefunction is found in the **Aharonov-Bohm effect**. Consider a particle with charge q travelling along a path, P , in which the magnetic field, $\mathbf{B} = 0$ is identically zero. However, a vanishing of the magnetic field does not imply that the vector potential, \mathbf{A} is zero. Indeed, as we have seen, any $\Lambda(\mathbf{r})$ such that $\mathbf{A} = \nabla\Lambda$ will translate to this condition. In traversing the path, the wavefunction of the particle will acquire the phase factor $e^{i\varphi}$, with $\varphi = \frac{q}{\hbar} \int_P \mathbf{A} \cdot d\mathbf{r}$,⁴ where the line integral runs along the path.

If we consider now two separate paths P and P' which share the same initial and final points, the relative phase of the wavefunction will be set by

$$\Delta\varphi = \frac{q}{\hbar} \int_P \mathbf{A} \cdot d\mathbf{r} - \frac{q}{\hbar} \int_{P'} \mathbf{A} \cdot d\mathbf{r} = \frac{q}{\hbar} \oint \mathbf{A} \cdot d\mathbf{r} = \underbrace{\frac{q}{\hbar} \int_A \mathbf{B} \cdot d^2\mathbf{r}}_{=\Phi}, \quad (6.25)$$

where the line integral \oint runs over the loop involving paths P and P' , and \int_A runs over the area enclosed by the loop. The last relation follows from the application of Stokes'

⁴We can find this by integrating the expression $\mathbf{A} = \nabla\Lambda$ along an arbitrary path P .

theorem. This result shows that the relative phase $\Delta\varphi$ is fixed by the factor q/\hbar multiplied by the magnetic flux $\Phi = \int_A \mathbf{B} \cdot d^2\mathbf{r}$ enclosed by the loop.⁵ In the absence of a magnetic field, the flux vanishes, and there is no additional phase and the interference pattern can be written as

$$I_{D,0} \propto |\psi_{P,0} + \psi_{P',0}|^2. \quad (6.26)$$

However, if we allow the paths to enclose a region of non-vanishing magnetic field, even if the field is identically zero on the paths P and P' , the wavefunction will acquire a non-vanishing relative phase. This flux-dependent phase difference translates to an observable shift of interference fringes when on an observation plane. Coming back to our two particle beams along paths P and P' ,⁶

$$\psi_P = \psi_{P,0} e^{i \int_P \mathbf{A} \cdot d\mathbf{r}}, \quad \psi_{P'} = \psi_{P',0} e^{i \int_{P'} \mathbf{A} \cdot d\mathbf{r}}. \quad (6.27)$$

Thus the interference pattern on the screen is then

$$\begin{aligned} I_D \propto |\psi_P + \psi_{P'}|^2 &= \left| \psi_{P,0} e^{i \int_P \mathbf{A} \cdot d\mathbf{r}} + \psi_{P',0} e^{i \int_{P'} \mathbf{A} \cdot d\mathbf{r}} \right|^2 \\ &= \left| \psi_{P,0} e^{i \oint \mathbf{A} \cdot d\mathbf{r}} + \psi_{P',0} \right|^2 \\ &= \left| \psi_{P,0} e^{i \frac{q}{\hbar} \Phi} + \psi_{P',0} \right|. \end{aligned} \quad (6.28)$$

The two beams thus arrive with a phase factors φ proportional to the magnetic flux enclosed by the two paths

$$\varphi = \frac{q}{\hbar} \Phi = \frac{e}{\hbar} \Phi = 2\pi \frac{\Phi}{\Phi_0}, \quad (6.29)$$

where $\Phi_0 \equiv h/e$, the *flux quantum*.

Since the original proposal, the Aharonov-Bohm effect has been studied in several experimental contexts. Of these, the most rigorous study was undertaken by Tonomura in 1986. Tonomura fabricated a doughnut-shaped (toroidal) ferromagnet six micrometers in diameter, and covered it with a niobium superconductor to completely confine the magnetic field within the doughnut, in accordance with the Meissner effect.⁷ With the magnet maintained at 5K, they measured the phase difference from the interference fringes between one electron beam passing through the hole in the doughnut and the other passing on the outside of the doughnut. Interference fringes are displaced with just half a fringe of spacing inside and outside of the doughnut, indicating the existence of the Aharonov-Bohm effect. Although electrons pass through regions free of any electromagnetic field, an observable effect was produced due to the existence of vector potentials.

The observation of the half-fringe spacing reflects the constraints imposed by the superconducting toroidal shield. When a superconductor completely surrounds a magnetic flux, the flux is quantised to an integral multiple of quantised flux $\hbar/2e$, the factor of two reflecting that fact that the superconductor involves a condensate of electron pairs. When an odd number of vortices are enclosed inside the superconductor, the relative phase shift

⁵Note that the phase difference depends on the magnetic flux, a function of the magnetic field, and is therefore a gauge invariant quantity.

⁶Of course choosing both paths to begin and end in the same position.

⁷Perfect diamagnetism, a hallmark of superconductivity, leads to the complete expulsion of magnetic fields – a phenomenon known as the Meissner effect.

becomes $\pi \pmod{2\pi}$ – half-spacing! For an even number of vortices, the phase shift is zero.⁸

6.4 Free Electron in a Magnetic Field

Finally, to complete our survey of the influence of a uniform magnetic field on the dynamics of charged particles, let us consider the problem of a free quantum particle. In this case, the classical electron orbits can be macroscopic and there is no reason to neglect the diamagnetic contribution to the Hamiltonian. Previously, we have worked with a gauge in which $\mathbf{A} = (-y, x, 0)B/2$, giving a constant field B in the z -direction $\mathbf{B} = (0, 0, B)$. However, to address the Schrödinger equation for a particle in a uniform perpendicular magnetic field, it is convenient to adopt the **Landau gauge**, $\mathbf{A}(\mathbf{r}) = (-By, 0, 0)$.

In this case, the stationary form of the Schrödinger equation is given by

$$\hat{H}\psi(\mathbf{r}) = \frac{1}{2m}[(\hat{p}_x^2 + qBy)^2 + \hat{p}_y^2 + \hat{p}_z^2]\psi(\mathbf{r}) = E\psi(\mathbf{r}). \quad (6.30)$$

Since \hat{H} commutes with both \hat{p}_x and \hat{p}_z , both operators have a common set of eigenstates reflecting the fact that p_x and p_z are conserved by the dynamics. The wavefunction must therefore take the form, $\psi(\mathbf{r}) = e^{i(p_xx+p_yy)/\hbar}\chi(y)$, with $\chi(y)$ defined by the equation,

$$\left[\frac{\hat{p}_y^2}{2m} + \frac{1}{2}m\omega_c^2(y - y_0)^2 \right] \chi(y) = \left(E - \frac{p_z^2}{2m} \right) \chi(y). \quad (6.31)$$

Here $y_0 = -p_x/qB$ and $\omega_c = |q|B/m$ coincides with the **cyclotron frequency** of the classical charged particle. We now see that the conserved canonical momentum p_x in the x -direction is in fact the coordinate of the centre of a simple harmonic oscillator potential in the y -direction with frequency ω_c . As a result, we can immediately infer that the eigenvalues of the Hamiltonian are comprised of a free particle component associated with motion parallel to the field, and a set of harmonic oscillator states,

$$E_{n,p_z} = (n + 1/2)\hbar\omega_c + \frac{p_z^2}{2m}. \quad (6.32)$$

The quantum numbers, n , specify states known as **Landau levels**.

Let us confine our attention to states corresponding to the lowest oscillator (Landau level) state, (and, for simplicity, $p_z = 0$), $E_0 = \hbar\omega_c/2$. What is the degeneracy of this Landau level? Consider a rectangular geometry of area $A = L_x \times L_y$ and in this case, the allowed values of p_x are quantised, as we can show by imposing periodic boundary conditions. The centre of the oscillator wavefunction, $y_0 = -p_x/qB$, must lie between 0 and L_y . With periodic boundary conditions $e^{ip_x L_x \hbar} = 1$, so that $p_x = n2\pi\hbar/L_x$. This means that y_0 takes a series of evenly-spaced discrete values, separated by $\Delta y_0 = \hbar/qBL_x$. So, for electron degrees of freedom, $q = -e$, the total number of states $N_{\text{states}} = L_y/|\Delta y_0|$, i.e.

$$N_{\text{states}} = \frac{L_x L_y}{\hbar/eB} = A \frac{B}{\Phi_0} = \frac{\Phi}{\Phi_0}, \quad (6.33)$$

⁸The superconducting flux quantum was actually predicted prior to Aharonov and Bohm, by Fritz London in 1948 using a phenomenological theory.

where $\Phi_0 = h/e$ denotes the “flux quantum”. So the total number of states in the lowest energy level coincides with the total number of flux quanta making up the field B penetrating the area A .

The Landau level degeneracy, N_{states} , depends on field; the larger the field, the more electrons can be fit into each Landau level. In the physical system, each Landau level is spin split by the Zeeman coupling, with (6.33) applying to one spin only. Finally, although we treated x and y in an asymmetric manner, this was merely for convenience of calculation; no physical quantity should differentiate between the two due to the symmetry of the original problem.

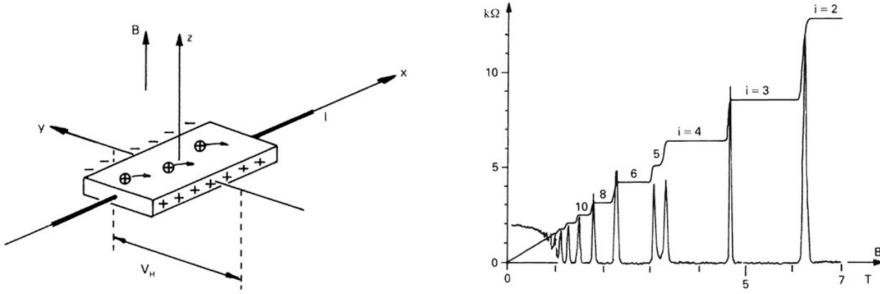


Fig. 6.2: (Left) A voltage V drives a current I in the positive x -direction. Normal Ohmic resistance is V/I . A magnetic field in the positive z -direction shifts positive charge carriers in the negative y -direction. This generates a Hall potential and a Hall resistance (VH/I) in the y direction. (Right) The Hall resistance varies stepwise with changes in magnetic field B . Step height is given by the physical constant h/e^2 (value approximately $25\text{k}\Omega$) divided by an integer i . The figure shows steps for $i = 2, 3, 4, 5, 6, 8$ and 10 . The effect has given rise to a new international standard for resistance. Since 1990 this has been represented by the unit 1 k Ω , defined as the Hall resistance at the fourth step ($h/4e^2$). The lower peaked curve represents the Ohmic resistance, which disappears at each step

It is interesting to compare the degeneracy of the Landau levels with the density of states when no magnetic field is present. For the same 2D system but with $\mathbf{B} = \mathbf{0}$, the density of states per unit area is independent of the energy (See Appendix A.5)

$$g(E) = \frac{m}{2\pi\hbar^2}. \quad (6.34)$$

Therefore, the total number of states in an energy interval equal to the Landau level spacing $\Delta E = \hbar\omega_c$ is

$$N = Ag(E)\Delta E = \frac{Am}{2\pi\hbar^2} \frac{\hbar eB}{m} = \frac{AB}{h/e} = \frac{\Phi}{\Phi_0} = N_L, \quad (6.35)$$

equal to the number of states in each Landau level. The effect of “switching on” the magnetic field is to transform the zero-field, uniform density of states into a set of equally spaced and highly degenerate Landau levels, while preserving the total number of available states.

For a system containing a total of N electrons, all available states are occupied up to a maximum energy E_F , the *Fermi energy*, such that

$$N = Ag(E_F)E_F = \frac{Am}{2\pi\hbar^2} E_F. \quad (6.36)$$

The occupation fraction of the highest Landau level is

$$f_L = \frac{N}{N_L} - \text{int}\left(\frac{N}{N_L}\right). \quad (6.37)$$

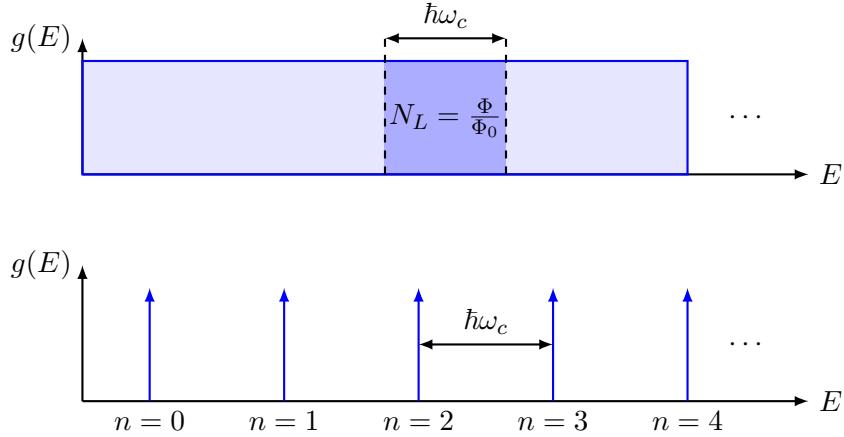


Fig. 6.3: The effect of “switching on” the magnetic field is to transform the (top) $\mathbf{B} = \mathbf{0}$ zero-field, uniform density of states into a set of equally spaced and highly degenerate Landau levels (bottom), while preserving the total number of available states.

Combining Eqs. (6.36) and (6.35), we obtain

$$\frac{N}{N_L} = \frac{Am}{2\pi\hbar^2} E_F \frac{\Phi_0}{\Phi} = \frac{Am}{2\pi\hbar^2} \frac{h/e}{BA} E_F = \frac{E_F}{\hbar\omega_c}. \quad (6.38)$$

So,

$$f_L = \frac{E_F}{\hbar\omega_c} - \text{int}\left(\frac{E_F}{\hbar\omega_c}\right) \quad (6.39)$$

Thus, as B increases, f_L is periodic with a period given by

$$\Delta\left(\frac{E_F}{\hbar\omega_c}\right) = \frac{E_F m}{\hbar e} \Delta\left(\frac{1}{B}\right) = 1 \implies \Delta\left(\frac{1}{B}\right) = \frac{\hbar e}{m E_F}. \quad (6.40)$$

Because of this, many properties of metals show periodic behaviour with $1/B$:

$$\begin{aligned} \text{electrical resistivity} &\leftrightarrow \text{the Shubnikov-de Haas effect} \\ \text{magnetic susceptibility} &\leftrightarrow \text{the de Haas-van Alphen effect.} \end{aligned}$$

However, by far the most striking and unexpected manifestation of Landau levels in 2D electron systems is the Quantum Hall effect

6.4.0.1 Integer Quantum Hall Effect

Until now, we have considered the impact of just a magnetic field. Consider now the Hall effect geometry in which we apply a crossed electric, \mathbf{E} and magnetic field, \mathbf{B} . Taking into account both contributions, the total current flow is given by

$$\mathbf{j} = \sigma_0 \left(\mathbf{E} - \frac{\mathbf{j} \times \mathbf{B}}{ne} \right), \quad (6.41)$$

where σ_0 denotes the conductivity, and n is the electron density. With the electric field oriented along y , and the magnetic field along z , the latter equation may be rewritten as

$$\begin{pmatrix} 1 & \frac{\sigma_0 B}{ne} \\ -\frac{\sigma_0 B}{ne} & 1 \end{pmatrix} \begin{pmatrix} j_x \\ j_y \end{pmatrix} = \sigma_0 \begin{pmatrix} 0 \\ E_y \end{pmatrix}. \quad (6.42)$$

Inverting these equations, one finds that

$$j_x = \underbrace{\frac{-\sigma_0^2 B / ne}{1 + (\sigma_0 B / ne)^2}}_{\sigma_{xy}} E_y, \quad j_y = \underbrace{\frac{\sigma_0}{1 + (\sigma_0 B / ne)^2}}_{\sigma_{xy}} E_x. \quad (6.43)$$

These provide the classical expressions for the longitudinal and Hall conductivities, σ_{yy} and σ_{xy} in the crossed field. Note that, for these classical expressions, σ_{xy} is proportional to B .

How does quantum mechanics revise this picture? For the classical model – **Drude theory**, the random elastic scattering of electrons impurities leads to a constant drift velocity in the presence of a constant electric field, $\sigma_0 = \frac{ne^2\tau}{m_e}$, where τ denotes the mean time between collisions. Now let us suppose the magnetic field is chosen so that the electrons exactly fill all of the states in the first ν Landau levels, i.e. the total number of electrons $nL_x L_y$ is

$$nL_x L_y = \nu N_{\text{states}} \implies n = \nu \frac{eB}{h}. \quad (6.44)$$

The scattering of electrons must lead to a transfer between quantum states. However, if all states of the same energy are filled,⁹ elastic (energy conserving) scattering becomes impossible. Moreover, since the next accessible Landau level energy is a distance $\hbar\omega_c$ away, at low enough temperatures, inelastic scattering becomes frozen out. As a result, the scattering rate $1/\tau$ vanishes at special values of the field, i.e. $\sigma_{yy} \rightarrow 0$ and

$$\sigma_{xy} \rightarrow \frac{ne}{B} = \nu \frac{e^2}{h}. \quad (6.45)$$

At critical values of the field, the Hall conductivity is quantised in units of e^2/h . Inverting the conductivity tensor, one obtains the resistivity tensor,

$$\begin{pmatrix} \rho_{xx} & \rho_{xy} \\ -\rho_{xy} & \rho_{yy} \end{pmatrix} = \begin{pmatrix} \sigma_{xx} & \sigma_{xy} \\ -\sigma_{xy} & \sigma_{yy} \end{pmatrix}^{-1}, \quad (6.46)$$

where

$$\rho_{xx} = \frac{\sigma_{xx}}{\sigma_{xx}^2 + \sigma_{xy}^2}, \quad \rho_{xy} = \frac{\sigma_{xy}}{\sigma_{xx}^2 + \sigma_{xy}^2}, \quad (6.47)$$

So, when $\sigma_{xx} = 0$ and $\sigma_{xy} = \nu e^2/h$, $\rho_{xx} = 0$ and $\rho_{xy} = h/\nu e^2$. The quantum Hall state describes dissipationless current flow in which the Hall conductance σ_{xy} is quantised in units of e^2/h . Experimental measurements of these values provides the best determination of fundamental ratio e^2/h , better than 1 part in 10^7 .

6.5 The Magnetic Moment of the Muon

The *muon* is a heavy ($m_\mu/m_e \approx 207$), unstable ($\tau_\mu \approx 2.2 \mu\text{s}$) version of the electron. There has been a long standing discrepancy between the measured value of the muon's magnetic moment and the corresponding QED prediction.

⁹Note that electrons are subject to Pauli's exclusion principle restricting the occupancy of each state to unity

Following from (5.44), we can define the anomalous magnetic moment of the muon

$$a_\mu \equiv \frac{g_\mu - 2}{2} \quad (6.48)$$

The discrepancy between QED and experiment may or may not need “new physics” to be resolved, new experiments and improved theoretical calculations are underway.

6.5.1 The Measurement

The experimental technique used to measure the muon magnetic moment is a beautiful example of spin precession in action.

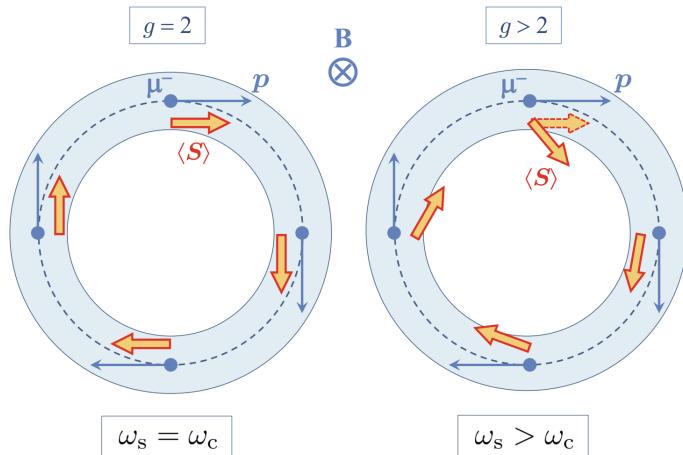


Fig. 6.4: For $g = 2$ ($a_\mu = 0$), the spin precession and cyclotron frequencies are equal ($\omega_s = \omega_c$, $\omega_a = 0$). The muon momentum and spin vectors precess at the same rate, remaining tangential to the circular orbit at all times (left-hand diagram). For $g_\mu \approx 2.002$ ($a_\mu \approx 0.001$), we have $\omega_s > \omega_c$, so the spin vector precesses (slightly) faster than the momentum vector. The spin vector therefore rotates relative to the orbital tangent, with angular frequency $\omega_s - \omega_c = \omega_a$ (right-hand diagram).

In the E281 experiment, relativistic muons ($v/c = 0.99941$) travel in a circular orbit around a ring containing a highly uniform (~ 1 ppm) magnetic field ($B = 1.4513$ T). The muon momentum precesses at the cyclotron frequency, $\omega_c = eB/m$. The ring is initially ($t = 0$) filled with polarised muons, with the muon spin vector oriented along the muon momentum direction.

For $g = 2$ ($a_\mu = 0$), the spin precession and cyclotron frequencies are equal ($\omega_s = \omega_c$, $\omega_a = 0$). The muon momentum and spin vectors precess at the same rate, remaining aligned tangential to the circular orbit at all times $t > 0$ (left-hand diagram Fig. 6.4).

For $g_\mu > 2$, we have $\omega_s > \omega_c$, so the muon spin vector precesses (slightly) faster than the momentum vector. The spin vector therefore rotates relative to the orbital tangent, with angular frequency $\omega_s - \omega_c = \omega_a$ (right-hand diagram Fig. 6.4).

In the muon rest frame, electrons from muon decay, $\mu^- \rightarrow e^- + \bar{\nu}_e + \nu_\mu$, are emitted non-isotropically with respect to the muon spin direction, as shown schematically by the

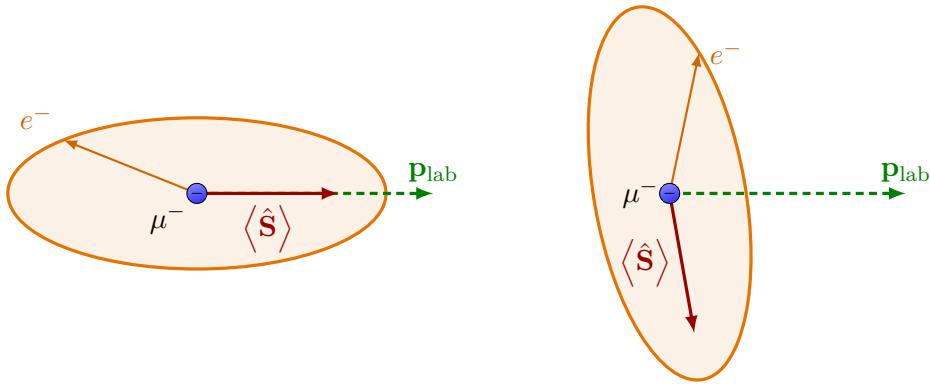


Fig. 6.5: In the muon rest frame, electrons from muon decay, $\mu^- \rightarrow e^- + \bar{\nu}_e + \nu_\mu$, are emitted non-isotropically with respect to the muon spin direction. The electron distribution (the ellipse) rotates with the muon spin vector.

elliptical probability distribution in Fig. 6.5. (The shape is not in fact elliptical, nor even symmetric due to *parity violation*, but this doesn't affect the argument.) The electron distribution (the ellipse) rotates with the muon spin vector.

In a particle's rest frame, $\mathbf{p} = \mathbf{0}$, the only quantity which can possibly introduce a preferred spatial direction is the *spin*. As the spin vector $\langle \hat{\mathbf{S}} \rangle$ rotates about the muon direction, so too does the distribution of decay electrons.

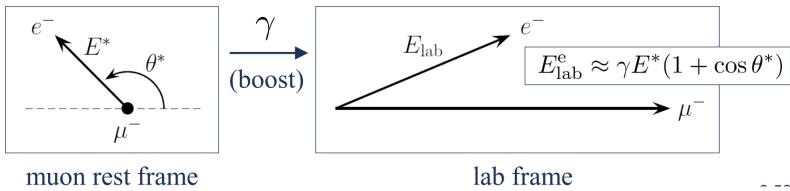


Fig. 6.6: After a Lorentz transformation to the lab frame, the distribution of electron four-momenta depends on the angle of the spin vector relative to the direction of the Lorentz boost. The electron energy distribution in the lab frame depends on the angle between the muon spin vector and the muon three-momentum.

After a Lorentz transformation to the lab frame (Fig. 6.6), the distribution of electron four-momenta depends on the angle of the spin vector relative to the direction of the Lorentz boost. The boost direction is always tangential to the orbital circle, or, equivalently, always lies along the muon three-momentum direction. The electron energy distribution in the lab frame therefore depends on the angle between the muon spin vector and the muon three-momentum. In the lab frame, electrons above a given threshold energy are detected and counted. The number of detected electrons above a certain lab frame energy varies as the spin $\langle \hat{\mathbf{S}} \rangle$ rotates relative to the momentum \mathbf{p} . Thus the number of electrons detected varies with time and therefore shows a structure which depends on the difference $\omega_a = \omega_s - \omega_c$ (Eq. (5.44)) between the spin $\langle \hat{\mathbf{S}} \rangle$ and momentum \mathbf{p} precession frequencies, with a period $2\pi/\omega_a$ (superimposed on a steady exponential decay of the number of electrons due to the muon lifetime of about $2.2\ \mu\text{s}$).

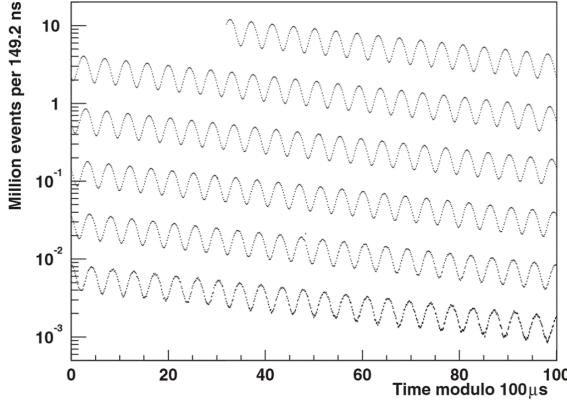


Fig. 6.7: The plot contains $\sim 3 \times 10^9$ muon decays, about 40% of the total E821 data set. The periodic structure is superimposed on a steady exponential fall due to the muon decay: $N_e(t) = N_0 e^{-t/\tau}$ (the muon has a mean lifetime of $\tau = 2.2 \mu\text{s}$ in its rest frame). G. W. Bennett et al., Phys. Rev. D **73** (2006) 072003

6.6 Magnetic Moments

The *electron* magnetic moment has been measured with a relative precision better than 1 part in 10^{12}

$$(g_e/2)_{\text{expt.}} = 1.00115965218073(28). \quad (6.49)$$

The small ($\sim 0.1\%$) deviation from $g_e = 2$ is resolved by Quantum Electrodynamics, the quantum theory of the electromagnetic field.

For the proton ($q = +e, m = m_p$) and neutron ($q = 0$), by analogy with the electron and muon, we might expect the magnetic moments to be

$$\hat{\boldsymbol{\mu}}_p = \frac{e}{m_p} \hat{\mathbf{S}}_p, \quad \hat{\boldsymbol{\mu}}_n = 0, \quad (6.50)$$

using $\gamma_S = q/m$. These naïve expectations fail badly, so instead introducing the *nuclear magneton* μ_N ,

$$\mu_N \equiv \frac{e\hbar}{2m_p} \approx 5.05 \times 10^{-27} \text{ J T}^{-1} \approx 3.15 \times 10^{-8} \text{ eV T}^{-1}, \quad (6.51)$$

and defining proton and neutron *g*-factors g_p and g_n such that

$$\hat{\boldsymbol{\mu}}_p = g_p \frac{\mu_N}{\hbar} \hat{\mathbf{S}}_p, \quad \mu_p = \frac{g_p}{2} \mu_N \quad (6.52)$$

$$\hat{\boldsymbol{\mu}}_n = g_n \frac{\mu_N}{\hbar} \hat{\mathbf{S}}_n, \quad \mu_n = \frac{g_n}{2} \mu_N, \quad (6.53)$$

we naïvely expect $g_p = 2, \mu_p = \mu_N$ and $g_n = 0, \mu_n = 0$, but instead find

$$\mu_p = +2.79284734462(82)\mu_N, \quad g_p \approx +5.586 \quad (6.54)$$

$$\mu_n = -1.91304273(45) \mu_N, \quad \mu_n \approx -3.826. \quad (6.55)$$

Atomic nuclei come in all possible spins, $S = 0, 1/2, 1, 3/2, 2, 5/2, \dots$. To generalise the definition of μ (Eq. (5.26)) to arbitrary spin S , in place of the spin-up state $|\uparrow\rangle$, we

consider the spin state $|S, m_S\rangle$ with the maximal possible value ($m_S = S$) of m_S , and define

$$\mu = \langle S, m_S = S | (\hat{\mu}_S)_z | S, m_S = S \rangle. \quad (6.56)$$

With $\hat{\mu} = \gamma_S \hat{\mathbf{S}}$, the (scalar) magnetic moment μ is therefore

$$\mu = \gamma_S S \hbar, \quad (6.57)$$

since $\hat{S}_z |S, S\rangle = S\hbar |S, S\rangle$. The g -factors defined in Section 5.4 naturally generalise to nuclei as

$$\mu = g_S \mu_N S. \quad (6.58)$$

The spin precession frequency for nuclei is then

$$\omega_s = \gamma_S B = \frac{\mu B}{S\hbar} = \frac{g_S \mu_N B}{\hbar}. \quad (6.59)$$

In *atomic* physics, for the proton and for nuclei, it is more usual to define g -factors in terms of the Bohr magneton μ_B , rather than μ_N

$$\hat{\mu}_p = -g_p \frac{\mu_B}{\hbar} \hat{\mathbf{S}}_p, \quad g_p \approx -\frac{m_e}{m_p} \times (2.793) \approx -0.0031 \quad (6.60)$$

and to use a different sign convention! The spin of the nucleus is usually denoted as \mathbf{I} (whereas in nuclear physics and particle physics the standard notation is \mathbf{J})

$$\hat{H}_B = -\hat{\mu}_I \cdot \mathbf{B} = -\gamma_I \hat{\mathbf{I}} \cdot \mathbf{B}, \quad \hat{\mu}_I = -g_I \frac{\mu_B}{\hbar} \hat{\mathbf{I}}. \quad (6.61)$$

For a single-electron atom for example, in atomic-style notation, the overall magnetic dipole contribution to the Hamiltonian is written as

$$\hat{H}_B = \frac{\mu_B}{\hbar} (\hat{\mathbf{L}} + g_e \hat{\mathbf{S}} + g_I \hat{\mathbf{I}}) \cdot \mathbf{B} \quad (6.62)$$

6.7 The Stern-Gerlach Experiment

Classically, a magnetic dipole moment $\boldsymbol{\mu}$ in an external magnetic field \mathbf{B} is subject to a couple

$$\mathbf{G} = \boldsymbol{\mu} \times \mathbf{B} \quad (6.63)$$

(leading to precession of $\boldsymbol{\mu}$ around \mathbf{B} , for example). In addition, if the field \mathbf{B} is *non-uniform*, then $\boldsymbol{\mu}$ is also subject to a force

$$\mathbf{F} = \nabla(\boldsymbol{\mu} \cdot \mathbf{B}). \quad (6.64)$$

For a sample of randomly oriented dipoles, the scalar product $\boldsymbol{\mu} \cdot \mathbf{B}$ takes on a continuous range of values between $-\mu B$ and $+\mu B$, leading, classically, to a continuum of possible trajectories through the magnetic field.

In the *quantum* case, however, particles possessing an internal magnetic dipole moment $\boldsymbol{\mu}$ follow only a *finite number* of distinct spatial trajectories. This is a truly fundamental quantum effect, first demonstrated experimentally by Stern and Gerlach (1922).

Consider a beam of *neutral* particles (e.g. neutrons) sent into a slowly varying magnetic field of the form

$$\mathbf{B}(\mathbf{r}) = \mathbf{B}_0 + \mathbf{B}_1(\mathbf{r}), \quad |\mathbf{B}_1(\mathbf{r})| \ll |\mathbf{B}_0|, \quad (6.65)$$

where \mathbf{B}_0 is constant, uniform and “large”, and \mathbf{B}_1 is “small”. The field \mathbf{B}_1 must satisfy Maxwell’s equations,

$$\nabla \cdot \mathbf{B}_1 = 0, \quad \nabla \times \mathbf{B}_1 = 0, \quad (6.66)$$

but its precise form is otherwise unimportant.

Without loss of generality, orient the z -axis along \mathbf{B}_0 , $\mathbf{B}_0 = (0, 0, B_0)$. Let the particles have a magnetic dipole moment $\hat{\mu}_S = \gamma_S \hat{\mathbf{S}}$. The motion of the particles through the magnetic field is then described by the Hamiltonian

$$\hat{H} = \frac{\hat{\mathbf{p}}^2}{2m} - \hat{\mu}_S \cdot \mathbf{B} = \frac{\hat{\mathbf{p}}^2}{2m} - \gamma_S (\hat{S}_z B_0 + \hat{\mathbf{S}} \cdot \mathbf{B}_1(\mathbf{r})). \quad (6.67)$$

Expanded out, the Hamiltonian is

$$\hat{H} = \frac{\hat{\mathbf{p}}^2}{2m} - \gamma_S (\hat{S}_z B_0 + \hat{S}_x B_{1,x}(\mathbf{r}) + \hat{S}_y B_{1,y}(\mathbf{r}) + \hat{S}_z B_{1,z}(\mathbf{r})). \quad (6.68)$$

Ehrenfest’s theorem gives the equations of motion of the particle as

$$\frac{d}{dt} \langle \hat{\mathbf{r}} \rangle = \frac{i}{\hbar} \langle [\hat{H}, \hat{\mathbf{r}}] \rangle, \quad \frac{d}{dt} \langle \hat{\mathbf{p}} \rangle = \frac{i}{\hbar} \langle [\hat{H}, \hat{\mathbf{p}}] \rangle. \quad (6.69)$$

The spin operator $\hat{\mathbf{S}}$ commutes with the observables $\hat{\mathbf{r}}$ and $\hat{\mathbf{p}}$

$$[\hat{\mathbf{S}}, \hat{\mathbf{r}}] = 0, \quad [\hat{\mathbf{S}}, \hat{\mathbf{p}}] = 0. \quad (6.70)$$

Hence, using the identity $[\hat{\mathbf{r}}, \hat{\mathbf{p}}^2] = 2i\hbar\hat{\mathbf{p}}$, we have

$$[\hat{H}, \hat{\mathbf{r}}] = \frac{1}{2m} [\hat{\mathbf{p}}^2, \hat{\mathbf{r}}] = -\frac{i\hbar}{m} \hat{\mathbf{p}}. \quad (6.71)$$

This gives (as expected on classical grounds)

$$\frac{d}{dt} \langle \hat{\mathbf{r}} \rangle = \frac{1}{m} \langle \hat{\mathbf{p}} \rangle. \quad (6.72)$$

Similarly, we have

$$[\hat{H}, \hat{\mathbf{p}}] = -\gamma_S (\hat{S}_x [B_{1,x}(\mathbf{r}), \hat{\mathbf{p}}] + \hat{S}_y [B_{1,y}(\mathbf{r}), \hat{\mathbf{p}}] + \hat{S}_z [B_{1,z}(\mathbf{r}), \hat{\mathbf{p}}]). \quad (6.73)$$

Applying the operator relation $[f(\mathbf{r}), \hat{\mathbf{p}}] = i\hbar \nabla f$ to each of the field components (B_x, B_y, B_z) gives

$$[\hat{H}, \hat{\mathbf{p}}] = -i\hbar \gamma_S (\hat{S}_x (\nabla B_{1,x}(\mathbf{r})) + \hat{S}_y (\nabla B_{1,y}(\mathbf{r})) + \hat{S}_z (\nabla B_{1,z}(\mathbf{r}))). \quad (6.74)$$

The second equation of motion is therefore

$$\begin{aligned} \frac{d}{dt} \langle \hat{\mathbf{p}} \rangle &= \gamma_S \langle \hat{S}_x (\nabla B_{1,x}(\mathbf{r})) + \hat{S}_y (\nabla B_{1,y}(\mathbf{r})) + \hat{S}_z (\nabla B_{1,z}(\mathbf{r})) \rangle \\ &= \gamma_S (\langle \hat{S}_x \rangle \langle \nabla B_{1,x}(\mathbf{r}) \rangle + \langle \hat{S}_y \rangle \langle \nabla B_{1,y}(\mathbf{r}) \rangle + \langle \hat{S}_z \rangle \langle \nabla B_{1,z}(\mathbf{r}) \rangle), \end{aligned} \quad (6.75)$$

where, in the last step, each term has been factorised into its spin and spatial components.

Neglecting the contribution from the “small” field component \mathbf{B}_1 in Eq. (6.67), the spin components of the wavefunction evolves according to the Hamiltonian

$$\hat{H} = -\gamma_S \hat{S}_z B_0, \quad (6.76)$$

i.e. to good approximation, the spin precesses about the “large” magnetic field \mathbf{B}_0 .

Suppose the particle enters the Stern-Gerlach apparatus in the initial spin eigenstate $|s, m_s\rangle$, as defined relative to the z -axis (the direction of the “large” uniform field \mathbf{B}_0)

$$|\psi(0)\rangle = |s, m_s\rangle, \quad \hat{S}_z |s, m_s\rangle = m_s \hbar |s, m_s\rangle. \quad (6.77)$$

Then, from Eq. (5.52), the expectation value $\langle \hat{\mathbf{S}} \rangle$ remains constant, and directed along z

$$\langle \psi(t) | \hat{\mathbf{S}} | \psi(t) \rangle = (0, 0, m_s \hbar). \quad (6.78)$$

In this case, only the z term of equation Eq. (6.75) is non-zero, giving the second equation of motion as

$$\boxed{\frac{d}{dt} \langle \hat{\mathbf{p}} \rangle = \gamma_S \hbar m_s \langle \nabla B_{1,z}(\mathbf{r}) \rangle.} \quad (6.79)$$

Combining this with Eq. (6.72) then gives a second-order differential equation for the trajectory followed by the particle

$$\boxed{m \frac{d^2}{dt^2} \langle \hat{\mathbf{r}} \rangle = \gamma_S \hbar m_s \langle \nabla B_{1,z}(\mathbf{r}) \rangle.} \quad (6.80)$$

Each value of m_s gives rise to a different solution (trajectory)

$$\langle \hat{\mathbf{r}} \rangle_{m_s}(t), \quad m_s = -s, \dots, s-1, s. \quad (6.81)$$

Thus, whatever the form of $\mathbf{B}_1(\mathbf{r})$, there are $2s + 1$ possible trajectories through the magnetic field, one for each possible value of m_s .

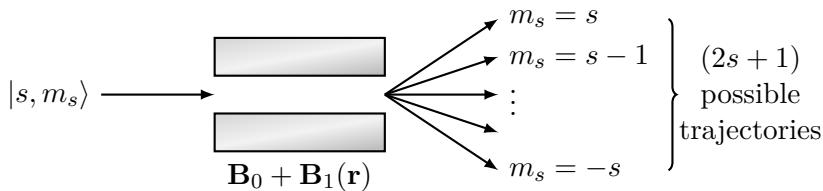


Fig. 6.8: Stern Gerlach apparatus, each value of m_s gives rise to a different trajectory, whatever the form of $\mathbf{B}_1(\mathbf{r})$, there are $2s + 1$ possible trajectories through the magnetic field, one for each possible value of m_s .

Note that, because we arbitrarily chose to orient the “large” magnetic field \mathbf{B}_0 along the z axis, it is the quantum number m_s defined with respect to z that has been picked out. Taking \mathbf{B}_0 along x would determine m_x , etc.

The trajectories obtained are the same as would be obtained classically if the magnetic dipole moment could take on only the *quantised* values

$$\boldsymbol{\mu} = (0, 0, \gamma_S \hbar m_s) \quad (6.82)$$

Now generalise to an arbitrary initial spin state

$$|\psi(0)\rangle = \sum_{m_s} c_{m_s} |s, m_s\rangle. \quad (6.83)$$

We have precession at the Larmor frequency $\omega_s = \gamma_S B_0$, but in practice, since B_0 is “large”, the precession is extremely rapid on timescales associated with the motion through the magnetic field. Hence, when averaged over even short timescales, the x and y matrix elements *effectively* vanish

$$\langle s, m'_s | \hat{S}_x | s, m_s \rangle = \langle s, m'_s | \hat{S}_y | s, m_s \rangle = 0. \quad (6.84)$$

Only the z -component is non-zero, as in equation Eq. (5.60)

$$\langle \psi(t) | \hat{S}_z | \psi(t) \rangle = \sum_{m_s} |c_{m_s}|^2 m_s \hbar. \quad (6.85)$$

Therefore, for a general initial spin state, the differential equation defining the average position of the particle is

$$m \frac{d^2}{dt^2} \langle \hat{\mathbf{r}} \rangle = \gamma_S \hbar \sum_{m_s} |c_{m_s}|^2 m_s \langle \nabla B_{1,z}(\mathbf{r}) \rangle. \quad (6.86)$$

Thus a factor of $|c_{m_s}|^2$ is associated with each individual m_s trajectory. The incoming beam separates into $2s + 1$ outgoing beams, each with different m_s , and with intensity in each proportional to $|c_{m_s}|^2$.

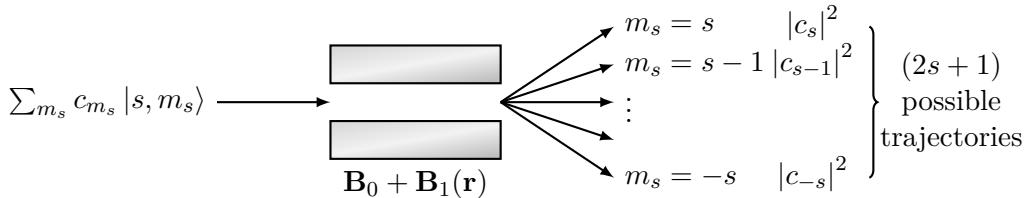


Fig. 6.9: Stern Gerlach apparatus for an arbitrary initial spin state $|\psi(0)\rangle = \sum_{m_s} c_{m_s} |s, m_s\rangle$. A factor of $|c_{m_s}|^2$ is associated with each individual m_s trajectory. The incoming beam separates into $2s + 1$ outgoing beams, each with different m_s , and with intensity in each proportional to $|c_{m_s}|^2$.

Note however that the analysis here has only determined how the *average position* of the particles in the original beam evolves with time. We have not actually shown that the original beam can *split* into two or more distinct outgoing beams, with well-defined spin states and intensities.¹⁰ To demonstrate this fully requires a more involved analysis of the time evolution of the wavefunction, not just the average position.

The Stern-Gerlach effect is of interest as a truly fundamental quantum phenomenon, but also finds many applications in the analysis or selection of spin states:

- Observation of the trajectory followed by a particle tells us the value of the quantum number m_s ;
- Particular spin states can be selected by blocking one or more of the emerging beams.

¹⁰Though Eq. (6.86) strongly suggests that this is what happens.

CHAPTER 7

Approximation Methods for Stationary States

7.1 Time-independent Perturbation Theory

While we have succeeded in deriving formal analytical solutions for stationary states of the Schrödinger operator in a variety of settings, in the majority of practical applications, *exact* solutions are inaccessible.¹ For example, if an atom is placed in an external electric field, the energy levels shift, and the wavefunctions become distorted — the Stark effect. The new energy levels and wavefunctions could in principle be obtained by writing down a complete Hamiltonian, including the external field. Indeed, such a programme may be achieved for the hydrogen atom. But even there, if the external field is small compared with the electric field inside the atom (which is billions of volts per metre) it is easier to compute the changes in the energy levels and wavefunctions within a scheme of successive corrections to the zero-field values. This method, termed perturbation theory, is the single most important method for solving problems in quantum mechanics, and is widely used in atomic physics, condensed matter and particle physics.

It should be acknowledged that there are – typically very interesting – problems which cannot be solved using perturbation theory, even when the perturbation is very weak; although such problems are the exception rather than the rule. One such case is the one-dimensional problem of free particles perturbed by a localised potential of strength λ . As we found earlier in Chapter (2), switching on an arbitrarily weak attractive potential causes the $k = 0$ free particle wavefunction to drop below the continuum of plane wave energies and become a localised bound state with binding energy of order λ^2 . However, on changing the sign of λ to give a repulsive potential, there is no bound state; the lowest energy plane wave state stays at energy zero. Therefore the energy shift on switching on the perturbation cannot be represented as a power series in λ , the strength of the perturbation. This particular difficulty does not typically occur in three dimensions, where arbitrarily weak potentials do not in general lead to bound states.

7.1.1 The Perturbation Series

Let us then consider an unperturbed “zeroth order” Hamiltonian, $\hat{H}^{(0)}$, having known (orthonormal) eigenstates $|n^{(0)}\rangle$ and eigenvalues $E_n^{(0)}$,

$$\hat{H}^{(0)} |n^{(0)}\rangle = E_n^{(0)} |n^{(0)}\rangle. \quad (7.1)$$

In the following we will address the question of how the eigenstates and eigenenergies are modified by the imposition of a small perturbation, $\hat{H}^{(1)}$ (such as that imposed by an

¹Indeed, even if such a solution is formally accessible, its complexity may render it of no practical benefit.

external electric or magnetic field on a charged particle, or the deformation of some other external potential). In short, we are interested in the solution of the Schrödinger equation,

$$\hat{H} |n\rangle = (\hat{H}^{(0)} + \hat{H}^{(1)}) |n\rangle = E_n |n\rangle \quad (7.2)$$

If the perturbation is small, $\langle n^{(0)} | \hat{H}^{(1)} | n^{(0)} \rangle \ll E_n^{(0)}$, it seems natural to suppose that, on turning on $\hat{H}^{(1)}$, the eigenfunctions and eigenvalues will change adiabatically from their unperturbed to their perturbed values, a situation described formally as “adiabatic continuity”,

$$|n^{(0)}\rangle \mapsto |n\rangle, \quad E_n^{(0)} \mapsto E_n. \quad (7.3)$$

However, note that this is not always the case. For example, as mentioned above, an infinitesimal perturbation has the capacity to develop a bound state not present in the unperturbed system. For now, let us proceed with the perturbative expansion and return later to discuss its potential range of validity.

The basic assumption that underpins the perturbation theory is that, for $\hat{H}^{(1)}$ small, the leading corrections are of the same order of magnitude as $\hat{H}^{(1)}$ itself. The perturbed eigenenergies and eigenvectors can then be obtained to a greater accuracy by a successive series of corrections, each of order $\langle \hat{H}^{(1)} \rangle / \langle \hat{H}^{(0)} \rangle$ compared with the previous. To identify terms of the same order in $\langle \hat{H}^{(1)} \rangle / \langle \hat{H}^{(0)} \rangle$, it is convenient to extract from $\hat{H}^{(1)}$ a dimensionless parameter λ , characterising the relative magnitude of the perturbation against $\hat{H}^{(0)}$, and then expand $|n\rangle$ and E_n as a power series in λ , i.e.

$$|n\rangle = |n^{(0)}\rangle + \lambda |n^{(1)}\rangle + \lambda^2 |n^{(2)}\rangle + \dots = \sum_{m=0}^{\infty} \lambda^m |n^{(m)}\rangle, \quad (7.4)$$

$$E_n = E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \dots = \sum_{m=0}^{\infty} \lambda^m E_n^{(m)}. \quad (7.5)$$

One may think of the parameter λ as an artificial book-keeping device to organise the perturbative expansion, and which is eventually set to unity at the end of the calculation.

Applied to the stationary form of the Schrödinger equation (7.2), an expansion of this sort leads to the relation

$$\begin{aligned} & (\hat{H}^{(0)} + \lambda \hat{H}^{(1)}) (|n^{(0)}\rangle + \lambda |n^{(1)}\rangle + \lambda^2 |n^{(2)}\rangle + \dots) \\ &= (E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \dots) (|n^{(0)}\rangle + \lambda |n^{(1)}\rangle + \lambda^2 |n^{(2)}\rangle + \dots) \end{aligned} \quad (7.6)$$

From this equation, we must relate terms of equal order in λ . At the lowest order, $\mathcal{O}(\lambda^0)$, we simply recover the unperturbed equation (7.1). In practical applications, one is usually interested in determining the first non-zero perturbative correction. In the following, we will explore the form of the first and second order perturbative corrections.

7.1.2 First Order Perturbation Theory

Isolating terms from (7.6) which are first order in λ ,

$$\hat{H}^{(0)} |n^{(1)}\rangle + \hat{H}^{(1)} |n^{(0)}\rangle = E_n^{(0)} |n^{(1)}\rangle + E_n^{(1)} |n^{(0)}\rangle. \quad (7.7)$$

and taking the inner product with the unperturbed states $\langle n^{(0)} |$, one obtains

$$\langle n^{(0)} | \hat{H}^{(0)} | n^{(1)} \rangle + \langle n^{(0)} | \hat{H}^{(1)} | n^{(0)} \rangle = \langle n^{(0)} | E_n^{(0)} | n^{(1)} \rangle + \langle n^{(0)} | E_n^{(1)} | n^{(0)} \rangle. \quad (7.8)$$

Noting that $\langle n^{(0)} | \hat{H}^{(0)} = \langle n^{(0)} | E_n^{(0)}$, and exploiting the normalisation $\langle n^{(0)} | n^{(0)} \rangle = 1$, one finds that the first order shift in energy is given simply by the expectation value of the perturbation taken with respect to the unperturbed eigenfunctions,

$$E_n^{(1)} = \langle n^{(0)} | \hat{H}^{(1)} | n^{(0)} \rangle. \quad (7.9)$$

Turning to the wavefunction, if we instead take the inner product of (7.7) with $\langle m^{(0)} |$ (with $m \neq n$), we obtain

$$\begin{aligned} \langle m^{(0)} | \hat{H}^{(0)} | n^{(1)} \rangle + \langle m^{(0)} | \hat{H}^{(1)} | n^{(0)} \rangle &= \langle m^{(0)} | E_n^{(0)} | n^{(1)} \rangle + \langle m^{(0)} | E_n^{(1)} | n^{(0)} \rangle \\ E_m^{(0)} \langle m^{(0)} | n^{(1)} \rangle + \langle m^{(0)} | \hat{H}^{(1)} | n^{(0)} \rangle &= E_n^{(0)} \langle m^{(0)} | n^{(1)} \rangle. \end{aligned} \quad (7.10)$$

Once again, with $\langle m^{(0)} | \hat{H}^{(0)} = \langle m^{(0)} | E_m^{(0)}$ and the orthogonality condition on the wavefunctions, $\langle m^{(0)} | n^{(0)} \rangle = 0$, one obtains an expression for the first order shift of the wavefunction expressed in the unperturbed basis,

$$\langle m^{(0)} | n^{(1)} \rangle = \frac{\langle m^{(0)} | \hat{H}^{(1)} | n^{(0)} \rangle}{E_n^{(0)} - E_m^{(0)}}, \quad E_n^{(0)} \neq E_m^{(0)}. \quad (7.11)$$

Note however, if $|m^{(0)}\rangle$ has the *same* energy as $|n^{(0)}\rangle$ (i.e. if the states are degenerate, then Eq. (7.10) reduces just to the condition

$$\langle m^{(0)} | \hat{H}^{(1)} | n^{(0)} \rangle = 0, \quad m \neq n, E_n^{(0)} = E_m^{(0)}, \quad (7.12)$$

and the overlap $\langle m^{(0)} | n^{(1)} \rangle$ remains undetermined and so we will clearly have a problem if (see Section 7.2).

In summary, setting $\lambda = 1$, to first order in perturbation theory, we have the eigenvalues and eigenfunctions,

$$E_n \simeq E_n^{(0)} + \langle n^{(0)} | \hat{H}^{(1)} | n^{(0)} \rangle, \quad (7.13)$$

$$|n\rangle \simeq |n^{(0)}\rangle + \sum_{m \neq n} |m^{(0)}\rangle \frac{\langle m^{(0)} | \hat{H}^{(1)} | n^{(0)} \rangle}{E_n^{(0)} - E_m^{(0)}}. \quad (7.14)$$

Before turning to the second order of perturbation theory, let us first consider a simple application of the method.

7.1.2.1 Ground State Energy of the Helium Atom

For the Helium atom, two electrons are bound to a nucleus of two protons and two neutrons. If one neglects altogether the Coulomb interaction between the electrons, in

the ground state, both electrons would occupy the ground state hydrogenic wavefunction (scaled appropriately to accommodate the doubling of the nuclear charge) and have opposite spin. Treating the Coulomb interaction between electrons as a perturbation, one may then use the basis above to estimate the shift in the ground state energy with

$$\hat{H}^{(1)} = \frac{1}{4\pi\epsilon_0} \frac{e^2}{|\mathbf{r}_1 - \mathbf{r}_2|}. \quad (7.15)$$

The hydrogenic wave functions are specified by three quantum numbers, n , ℓ , and m . In the ground state, the corresponding wavefunction takes the spatially isotropic form,

$$\langle \mathbf{r} | n = 1, \ell = 0, m = 0 \rangle = \psi_{100}(\mathbf{r}) = \left(\frac{1}{\pi a^3} \right)^{1/2} e^{-r/a}, \quad (7.16)$$

where $a = \frac{4\pi\epsilon_0}{Ze^2} \frac{\hbar^2}{m_e} = \frac{a_0}{Z}$ denotes the atomic Bohr radius for a nuclear charge Z . For the Helium atom ($Z = 2$), the symmetrised ground state of the unperturbed Hamiltonian is given by the spin singlet ($S = 0$) electron wavefunction,

$$|g.s.(0)\rangle = \frac{1}{\sqrt{2}} (|100, \uparrow\rangle \otimes |100, \downarrow\rangle - |100, \downarrow\rangle \otimes |100, \uparrow\rangle). \quad (7.17)$$

Here we have used the **direct product** \otimes to discriminate between the two electrons. Then, applying the perturbation theory formula above (7.9), to first order in the Coulomb interaction, the energy shift is given by

$$E_n^{(1)} = \langle g.s.(0) | \hat{H}^{(1)} | g.s.(0) \rangle = \frac{e^2}{4\pi\epsilon_0} \frac{1}{(\pi a^3)^2} \int d\mathbf{r}_1 d\mathbf{r}_2 \frac{e^{-2(r_1+r_2)/a}}{|\mathbf{r}_1 - \mathbf{r}_2|} = \frac{e^2}{4\pi\epsilon_0} \frac{C_0}{2a}, \quad (7.18)$$

where we have defined the dimensionless constant $C_0 = \frac{1}{(4\pi)^2} \int d\mathbf{z}_1 d\mathbf{z}_2 \frac{e^{-(z_1+z_2)}}{|\mathbf{z}_1 - \mathbf{z}_2|}$. Then, making use of the identity,

$$\frac{1}{(4\pi)^2} \int d\Omega_1 d\Omega_2 \frac{1}{|\mathbf{z}_1 - \mathbf{z}_2|} = \frac{1}{\max(z_1, z_2)}, \quad (7.19)$$

where the integrations run over the angular coordinates of the vectors \mathbf{z}_1 and \mathbf{z}_2 , and $z_{1,2} = |\mathbf{z}_{1,2}|$, one finds that $C_0 = 2 \int_0^\infty dz_1 z_1^2 e^{-z_1} \int_{z_1}^\infty dz_2 z_2 e^{-z_2} = 5/4$. As a result, noting that the Rydberg energy, $Ry = \frac{e^2}{4\pi\epsilon_0} \frac{1}{2a_0}$, we obtain the first order energy shift $\Delta E = \frac{5}{4} Z Ry \simeq 34 \text{eV}$ for $Z = 2$. This leads to a total ground state energy of $(2Z^2 - \frac{5}{4}Z) Ry = -5.5 Ry \simeq -74.8 \text{eV}$ compared to the experimental value of -5.807Ry .

7.1.3 Second Order Perturbation Theory

With the first order of perturbation theory in place, we now turn to consider the influence of the second order terms in the perturbative expansion (7.6). Isolating terms of order λ^2 , we have

$$\hat{H}^{(0)} |n^{(2)}\rangle + \hat{H}^{(1)} |n^{(1)}\rangle = E_n^{(0)} |n^{(2)}\rangle + E_n^{(1)} |n^{(1)}\rangle + E_n^{(2)} |n^{(0)}\rangle. \quad (7.20)$$

As before, taking the inner product with $\langle n^{(0)} |$, one obtains

$$\begin{aligned} & \langle n^{(0)} | \hat{H}^{(0)} | n^{(2)} \rangle + \langle n^{(0)} | \hat{H}^{(1)} | n^{(1)} \rangle \\ &= \langle n^{(0)} | E_n^{(0)} | n^{(2)} \rangle + \langle n^{(0)} | E_n^{(1)} | n^{(1)} \rangle + \langle n^{(0)} | E_n^{(2)} | n^{(0)} \rangle. \end{aligned} \quad (7.21)$$

Noting that the first two terms on the left and right hand sides cancel, we are left with the result

$$E_n^{(2)} = \langle n^{(0)} | \hat{H}^{(1)} | n^{(1)} \rangle - E_n^{(1)} \langle n^{(0)} | n^{(1)} \rangle. \quad (7.22)$$

Previously, we have made use of the normalisation of the basis states, $|n^{(0)}\rangle$. We have said nothing so far about the normalisation of the exact eigenstates, $|n\rangle$. Of course, eventually, we would like to ensure normalisation of these states too. However, to facilitate the perturbative expansion, it is operationally more convenient to impose a normalisation on $|n\rangle$ through the condition $\langle n^{(0)} | n \rangle = 1$. Substituting the λ expansion for $|n\rangle$, we thus have

$$\langle n^{(0)} | n \rangle = 1 = \langle n^{(0)} | n^{(0)} \rangle + \lambda \langle n^{(0)} | n^{(1)} \rangle + \lambda^2 \langle n^{(0)} | n^{(2)} \rangle + \dots. \quad (7.23)$$

From this relation, it follows that $\langle n^{(0)} | n^{(1)} \rangle = \langle n^{(0)} | n^{(2)} \rangle = \dots = 0$. We can therefore drop the term $E_n^{(1)} \langle n^{(0)} | n^{(1)} \rangle$ from consideration. As a result, we obtain

$$E_n^{(2)} = \langle n^{(0)} | \hat{H}^{(1)} | n^{(1)} \rangle = \langle n^{(0)} | \hat{H}^{(1)} \sum_{m \neq n} | m^{(0)} \rangle \frac{\langle m^{(0)} | \hat{H}^{(1)} | n^{(0)} \rangle}{E_n^{(0)} - E_m^{(0)}} \quad (7.24)$$

i.e.

$$E_n^{(2)} = \sum_{m \neq n} \frac{|\langle m^{(0)} | \hat{H}^{(1)} | n^{(0)} \rangle|^2}{E_n^{(0)} - E_m^{(0)}}. \quad (7.25)$$

From this result, we can conclude that,

- for the ground state, the second order shift in energy is always negative;
- if the matrix elements of $\hat{H}^{(1)}$ are of comparable magnitude, neighbouring levels make a larger contribution than distant levels;
- Levels that lie in close proximity tend to be **repelled**;
- If a fraction of the states belong to a continuum (see Subsection 3.1.3 for an introduction in continuum states), the sum in Eq. (7.25) should be replaced by an integral.

Once again, to illustrate the utility of the perturbative expansion, let us consider a concrete physical example.

7.1.3.1 The Quadratic Stark Effect

Consider the influence of an external electric field on the ground state of the hydrogen atom. As the composite electron and proton are drawn in different directions by the field, the relative displacement of the electron cloud and nucleus results in the formation of a dipole which serves to lower the overall energy. In this case, the perturbation due to the external field takes the form

$$\hat{H}^{(1)} = -q\mathcal{E}z = e\mathcal{E}r \cos \theta, \quad (7.26)$$

where $q = -e$ denotes the electron charge, and the electric field, $\mathcal{E} = \mathcal{E}\hat{e}_z$ is oriented along the z -axis. With the non-perturbed energy spectrum given by $E_{n\ell m}^{(0)} \equiv E_n^{(0)} = -\text{Ry}/n^2$, the ground state energy is given by $E^{(0)} \equiv E_{100}^{(0)} = -\text{Ry}$. At first order in the electric field strength, \mathcal{E} , the shift in the ground state energy is given by $E^{(1)} = \langle 100|e\mathcal{E}z|100\rangle$ where the ground state wavefunction was defined above (7.16). Since the potential perturbation is antisymmetric in z , it is easy to see that the energy shift vanishes at this order.

We are therefore led to consider the contribution second order in the field strength. Making use of Eq. (7.25), and neglecting the contribution to the energy shift from the continuum of unbound positive energy states, we have

$$E^{(2)} = \sum_{n \neq 1, \ell, m} \frac{|\langle n\ell m | e\mathcal{E}z | 100 \rangle|^2}{E_1^{(0)} - E_n^{(0)}}, \quad (7.27)$$

where $|n\ell m\rangle$ denote the set of bound state hydrogenic wavefunctions. Although the expression for $E^{(2)}$ can be computed exactly, the programme is somewhat tedious. However, we can place a strong bound on the energy shift through the following argument: Since, for $n > 2$, $|E_1^{(0)} - E_n^{(0)}| > |E_1^{(0)} - E_2^{(0)}|$, we have

$$|E^{(2)}| < \frac{1}{E_2^{(0)} - E_1^{(0)}} \sum_{n \neq 1, \ell, m} \langle 100 | e\mathcal{E}z | n\ell m \rangle \langle n\ell m | e\mathcal{E}z | 100 \rangle. \quad (7.28)$$

Since $\sum_{n, \ell, m} |n\ell m\rangle \langle n\ell m| = \hat{\mathbb{I}}$, we have $\sum_{n \neq 1, \ell, m} |n\ell m\rangle \langle n\ell m| = \hat{\mathbb{I}} - |100\rangle \langle 100|$. Finally, since $\langle 100 | z | 100 \rangle = 0$, we can conclude that $|E^{(2)}| < \frac{1}{E_2^{(0)} - E_1^{(0)}} \langle 100 | (e\mathcal{E}z)^2 | 100 \rangle$. With $\langle 100 | z^2 | 100 \rangle = a_0^2$, $E_1^{(0)} = -\frac{e^2}{4\pi\epsilon_0} \frac{1}{2a_0} = -\text{Ry}$, and $E_2^{(0)} = E_1^{(0)}/4$, we have

$$|E^{(2)}| < \frac{1}{\frac{3}{4}e^2/8\pi\epsilon_0 a_0} (e\mathcal{E})^2 a_0^2 = \frac{8}{3} 4\pi\epsilon_0 \mathcal{E}^2 a_0^3. \quad (7.29)$$

Furthermore, since all terms in the perturbation series for $E^{(2)}$ are negative, the first term in the series sets a lower bound, $|E^{(2)}| > \frac{|\langle 210 | e\mathcal{E}z | 100 \rangle|^2}{E_2^{(0)} - E_1^{(0)}}$. From this result, one can show that $0.55 \times \frac{8}{3} 4\pi\epsilon_0 \mathcal{E}^2 a_0^3 < |E^{(2)}| < \frac{8}{3} 4\pi\epsilon_0 \mathcal{E}^2 a_0^3$.

7.1.3.2 Infinite Square Well with Central Bump

$$\hat{H}^{(1)} = \begin{cases} \varepsilon & |x| < b \\ 0 & \text{otherwise} \end{cases} \quad (7.30)$$

Schrödinger's equation can be solved exactly for this example, but only with a lot of work (numerical or graphical), and thus we turn to perturbation theory to provide an approximation to the solution due its closeness to one we have already so closely studied.

For an infinite square well covering $-a < x < a$, the perturbed energies (without the bump) are

$$E_n^{(0)} = \frac{\hbar^2}{2m} \frac{n^2\pi^2}{4a^2} = n^2 E_0 \quad (7.31)$$

$$|n^{(0)}\rangle = \psi_n(x) = \sqrt{\frac{1}{a}} \sin\left(\frac{n\pi}{2a}(x + a)\right), \quad (7.32)$$

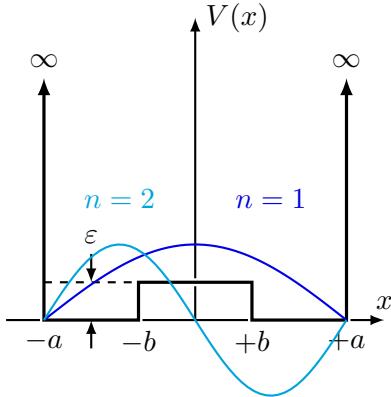


Fig. 7.1: Infinite square well with central bump.

where E_0 is the unperturbed ground state energy $E_0 \equiv E_1^{(0)} = \frac{\hbar^2\pi^2}{8ma^2}$ ($n = 1$).

The first-order correction to the unperturbed energy $E_n^{(0)}$ is given by

$$E_n^{(1)} = \left\langle n^{(0)} \left| \hat{H}^{(1)} \right| n^{(0)} \right\rangle = \varepsilon \int_{-b}^{+b} |\psi_n(x)|^2 dx. \quad (7.33)$$

Substituting for the unperturbed wavefunction $\psi_n(x)$ this is

$$\begin{aligned} E_n^{(1)} &= \frac{\varepsilon}{a} \int_{-b}^{+b} \sin^2\left(\frac{n\pi}{2a}(x+a)\right) dx \\ &= \varepsilon \left[\frac{b}{a} - \frac{(-1)^n}{n\pi} \sin\left(\frac{n\pi b}{a}\right) \right]. \end{aligned} \quad (7.34)$$

In particular,

$$\text{ground state } (n = 1) \quad E_1^{(1)} = \frac{\varepsilon}{a} \left[b + \frac{a}{\pi} \sin\left(\frac{\pi b}{a}\right) \right], \quad (7.35)$$

$$\text{first excited state } (n = 2) \quad E_2^{(1)} = \frac{\varepsilon}{a} \left[b - \frac{a}{2\pi} \sin\left(\frac{2\pi b}{a}\right) \right]. \quad (7.36)$$

The first-order correction is tangential to the true solution at $\varepsilon = 0$.

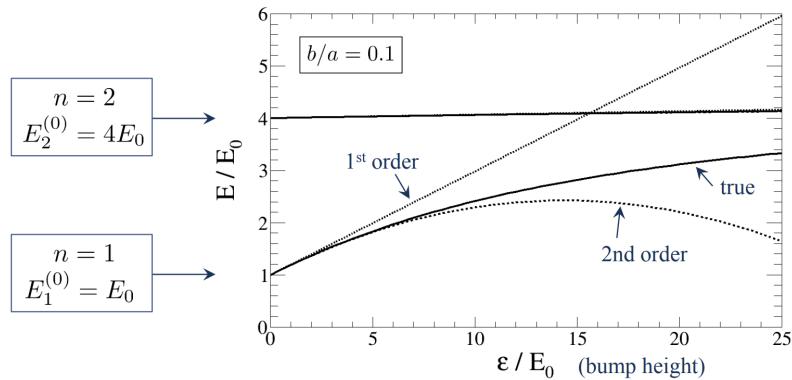


Fig. 7.2: The plot shows the true and perturbation theory energies as a function of the height, ε , of the bump, for a bump of width $b/a = 0.1$ in units of the unperturbed ground state energy $E_0 \equiv \hbar^2\pi^2/(8ma^2)$. For $n = 2$, the energy differences are too small to be visible.

For a narrow bump, $b \ll a$, equation (7.34) gives approximately

$$E_n^{(1)} \approx \frac{\varepsilon b}{a} [1 - (-1)^n]. \quad (7.37)$$

In particular

$$\text{ground state } (n = 1) \quad E_1^{(1)} \approx \frac{2\varepsilon b}{a}, \quad (7.38)$$

$$\text{first excited state } (n = 2) \quad E_2^{(1)} \approx 0. \quad (7.39)$$

The small energy correction for $n = 2$ is qualitatively as expected from the form of the unperturbed wavefunctions. At $x = 0$, where the bump is, the $n = 1$ wavefunction is *maximal*, but the $n = 2$ wavefunction *vanishes* (see Fig. 7.1). Hence we expect the matrix element $\langle n^{(0)} | \hat{H}^{(1)} | n^{(0)} \rangle$ to be much smaller for $n = 2$ than for $n = 1$.

The *second-order* energy corrections are given by equation (7.25), as an infinite sum over terms involving matrix elements of the form

$$\begin{aligned} \langle m^{(0)} | \hat{H}^{(1)} | n^{(0)} \rangle &= \varepsilon \int_{-b}^{+b} \psi_m^*(x) \psi_n(x) dx \\ &= \frac{\varepsilon}{b} \int_{-b}^{+b} \sin\left(\frac{m\pi}{2a}(x+a)\right) \sin\left(\frac{n\pi}{2a}(x+a)\right) dx. \end{aligned} \quad (7.40)$$

If m and n are both even, or both odd, evaluation the integral gives

$$\langle m^{(0)} | \hat{H}^{(1)} | n^{(0)} \rangle = \frac{2\varepsilon}{\pi} \left[\frac{(-1)^{(m-n)/2}}{(m-n)} \sin\left(\frac{(m-n)\pi b}{2a}\right) - \frac{(-1)^{(m+n)/2}}{(m+n)} \sin\left(\frac{(m+n)\pi b}{2a}\right) \right]. \quad (7.41)$$

If m is odd and n is even, or vice versa, the integral *vanishes*

$$\langle m^{(0)} | \hat{H}^{(1)} | n^{(0)} \rangle = 0 \quad (m+n \text{ odd}). \quad (7.42)$$

Finally, the *first-order wavefunctions* are given by Eq. (7.14), involving an infinite sum over the same matrix elements. Because of the matrix element property of Eq. (7.42), the symmetry of the wavefunctions is preserved, states which are initially even (odd) remain even (odd) at first-order (consistent with the fact that the central-bump potential $V(x)$ is symmetric).

Each term in the infinite sum satisfies the boundary conditions at $|x| = a$ and $|x| = b$ (because each $|m^{(0)}\rangle$ does)

$$\psi(\pm a) = 0, \quad \psi, \psi' \text{ continuous at } x = \pm b, \quad (7.43)$$

so do the first-order perturbed wavefunction.

Overall conclusion: perturbation theory gives sensible results for both the eigenstate energies and the corresponding wavefunctions.

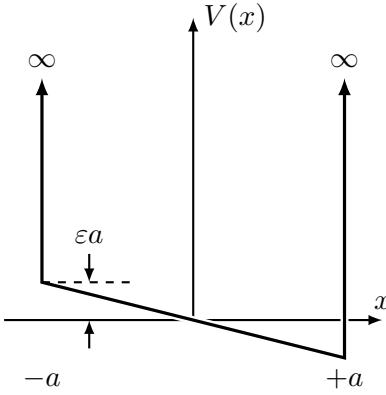


Fig. 7.3: Infinite square well in an electric field.

7.1.3.3 Infinite Square Well in an Electric Field

Applying an electric field \mathcal{E} corresponds to adding a perturbation $\hat{H}^{(1)} = -q\mathcal{E}x$, so let us consider $\hat{H}^{(1)} = \varepsilon x$. In this case, no true analytical solution of Schrödinger's equation is available for comparison

The first-order energy corrections all *vanish* (by symmetry)

$$E_n^{(1)} = \varepsilon \int_{-a}^{+a} x |\psi_n(x)|^2 dx = \frac{\varepsilon}{a} \int_{-a}^{+a} x \sin^2\left(\frac{n\pi}{2a}(x+a)\right) dx = -. \quad (7.44)$$

(This is a 1D example of a *parity selection rule*, see ??).

However, there is a non-zero change at first-order to the *wavefunctions*,

$$|n\rangle = |n^{(0)}\rangle + \frac{\varepsilon}{E_0} \sum_{m \neq n} |m^{(0)}\rangle \frac{\langle m^{(0)} | \hat{x} | n^{(0)} \rangle}{n^2 - m^2}, \quad (7.45)$$

since $E_i^{(0)} = i^2 E_0$.

This requires evaluation of the matrix elements of the operator \hat{x} between unperturbed states with $m \neq n$:

$$\begin{aligned} \langle m^{(0)} | \hat{x} | n^{(0)} \rangle &= \int_{-a}^{+a} x \psi_m^*(x) \psi_n(x) dx \\ &= \frac{1}{a} \int_{-a}^{+a} x \sin\left(\frac{m\pi}{2a}(x+a)\right) \sin\left(\frac{n\pi}{2a}(x+a)\right) dx. \end{aligned} \quad (7.46)$$

If m and n are both even, or both odd, then the integral above vanishes by symmetry, otherwise we find

$$\langle m^{(0)} | \hat{x} | n^{(0)} \rangle = -\frac{16a}{\pi^2} \frac{mn}{(m^2 - n^2)^2} \quad (m+n \text{ odd}). \quad (7.47)$$

Hence, the change in the wavefunctions, to first-order, is

$$|n\rangle = |n^{(0)}\rangle + \frac{\varepsilon a}{E_0 \pi^2 \sqrt{a}} \sum_{m+n \text{ odd}} \frac{m}{(m^2 - n^2)^3} \sin\left(\frac{m\pi}{2a}(x+a)\right). \quad (7.48)$$

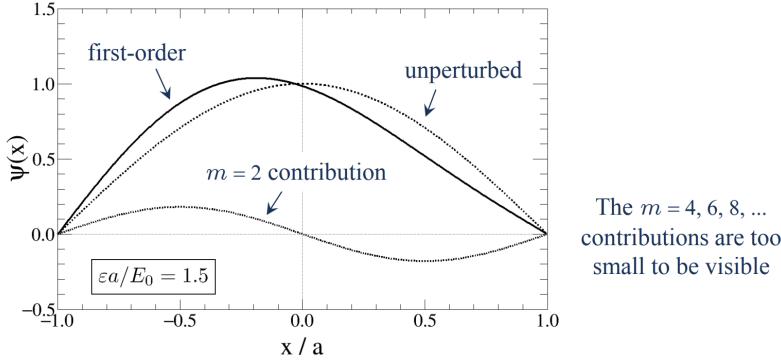


Fig. 7.4: For the ground state ($n = 1$), for example, only antisymmetric states (with even m) contribute to the sum, and $m = 2$ overwhelmingly dominates. The perturbed wavefunctions preserve the boundary conditions, but no longer have definite symmetry.

For the ground state ($n = 1$), for example, only antisymmetric states (with even m) contribute to the sum, and $m = 2$ overwhelmingly dominates, see Fig. 7.4. The perturbed wavefunctions preserve the boundary conditions, but no longer have definite symmetry as would be expected because neither does the overall potential $V(x)$.

The leading contribution to the energy corrections is quadratic (second-order)

$$\begin{aligned} E_n^{(2)} &= \frac{\varepsilon^2}{E_0} \sum_{m \neq n} \frac{|\langle m^{(0)} | \hat{x} | n^{(0)} \rangle|}{n^2 - m^2} \\ \implies \frac{E_n^{(2)}}{E_n^{(0)}} &= -\left(\frac{\varepsilon a}{E_0}\right)^2 \left(\frac{16}{\pi^2}\right)^2 \sum_{m+n \text{ odd}} \frac{m^2}{(m^2 - n^2)^3}. \end{aligned} \quad (7.49)$$

For ground state ($n = 1$), again only even values of m contribute and the sum is overwhelmingly dominated by the $m = 2$ contribution

$$\begin{aligned} \frac{E_1^{(2)}}{E_1^{(0)}} &= -\left(\frac{\varepsilon a}{E_0}\right)^2 \left(\frac{16}{\pi^2}\right)^2 \sum_{m=2,4,6,\dots} \frac{m^2}{(m^2 - 1)^3} \\ &= -\left(\frac{\varepsilon a}{E_0}\right)^2 \left(\frac{16}{\pi^2}\right)^2 \left[\frac{4}{27} + \frac{16}{15^3} + \frac{36}{35^3} + \dots \right] \end{aligned} \quad (7.50)$$

7.1.3.4 Harmonic Oscillator with Linear Perturbation

$$\hat{H} = \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega^2\hat{x}^2 + \varepsilon\hat{x}. \quad (7.51)$$

The zeroth-order harmonic oscillator eigenstates, $|n\rangle \equiv |n^{(0)}\rangle$, are given in Eq. (3.139)

$$\psi_0(x) = \langle x | 0 \rangle = \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} e^{-m\omega x^2/2\hbar}, \quad (7.52)$$

the first-order energy corrections are:

$$E_n^{(1)} = \langle n | \hat{H}^{(1)} | n \rangle = \varepsilon \langle n | \hat{x} | n \rangle = \varepsilon \int_{-\infty}^{+\infty} x |\psi_n(x)|^2 dx. \quad (7.53)$$

This is easiest evaluated using the ladder operators,²

$$\begin{aligned}\hat{x} &= \sqrt{\frac{\hbar}{2m\omega}}(\hat{a} + \hat{a}^\dagger) \\ \implies \langle n|\hat{x}|n\rangle &\propto \langle n|\hat{a}|n\rangle + \langle n|\hat{a}^\dagger|n\rangle \\ &= \sqrt{n}\langle n|n-1\rangle + \sqrt{n+1}\langle n|n+1\rangle = 0\end{aligned}$$

Hence the first-order energy shift *vanishes* for all n :

$$E_n^{(1)} = \langle n|\hat{H}^{(1)}|n\rangle = 0 \quad (n = 0, 1, 2, \dots) \quad (7.54)$$

The second-order energy correction is

$$E_n^{(2)} = \sum_{m \neq n} \frac{|\langle m|\hat{H}|n\rangle|}{E_n^{(0)} - E_m^{(0)}} = \varepsilon^2 \sum_{m \neq n} \frac{|\langle m|\hat{x}|n\rangle|^2}{E_n^{(0)} - E_m^{(0)}} \quad (7.55)$$

and receives contributions only from neighbouring levels, $m = n \pm 1$

$$\langle n-1|\hat{a}|n\rangle = \sqrt{n} \implies \langle n-1|\hat{x}|n\rangle = \sqrt{\frac{\hbar}{2m\omega}}\sqrt{n} \quad (7.56)$$

$$\langle n+1|\hat{a}^\dagger|n\rangle = \sqrt{n+1} \implies \langle n+1|\hat{x}|n\rangle = \sqrt{\frac{\hbar}{2m\omega}}\sqrt{n+1}. \quad (7.57)$$

All other terms vanish, $\langle m|\hat{x}|n\rangle = 0$ for $m \neq n \pm 1$.

Hence, to second-order, the perturbed energies are (for $n = 0, 1, 2, \dots$)

$$E_n = E_n^{(0)} + \frac{\hbar\varepsilon^2}{2m\omega} \left(\frac{n}{E_n^{(0)} - E_{n-1}^{(0)}} + \frac{n+1}{E_n^{(0)} - E_{n+1}^{(0)}} \right) \quad (7.58)$$

Substituting the unperturbed energies $E_n^{(0)} = (n + 1/2)\hbar\omega$ then gives the second-order energies as

$$E_n = \left(n + \frac{1}{2}\right)\hbar\omega - \frac{\varepsilon^2}{2m\omega^2}. \quad (7.59)$$

This is, in fact, the *true* solution, as can be seen by writing the overall potential as

$$V(x) = \frac{1}{2}m\omega^2\hat{x}^2 + \varepsilon\hat{x} = \frac{1}{2}m\omega^2\left(x + \frac{\varepsilon}{m\omega^2}\right)^2 - \frac{\varepsilon^2}{2m\omega^2}. \quad (7.60)$$

This is a harmonic oscillator, still with frequency ω , but with the central position shifted,

$$x \rightarrow x' = x + \varepsilon/(m\omega^2), \quad (7.61)$$

and with energy levels lowered by a common amount

$$E_n \rightarrow E'_n = E_n - \frac{\varepsilon^2}{2m\omega^2}, \quad E_n = \left(n + \frac{1}{2}\right)\hbar\omega. \quad (7.62)$$

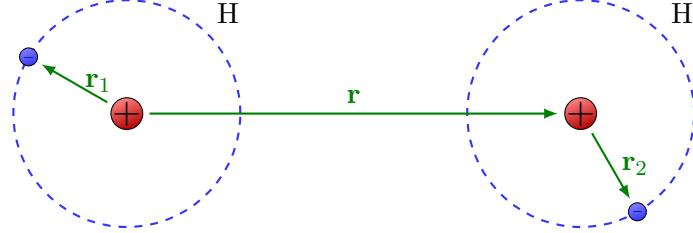


Fig. 7.5: Two hydrogen atoms, both in the ground state and separated by a distance r .

7.1.3.5 Van der Waals Interaction

Consider the force between two hydrogen atoms, both in the ground state, and separated by a distance r as illustrated in Fig. 7.5.

$$\hat{H} = \hat{H}_0 + V \begin{cases} \hat{H}_0 = -\frac{\hbar^2}{2m_e} (\nabla_1^2 + \nabla_2^2) - \frac{e^2}{4\pi\epsilon_0 r_1} - \frac{e^2}{4\pi\epsilon_0 r_2} \\ V = \frac{e^2}{4\pi\epsilon_0} \left(\frac{1}{r} + \frac{1}{|\mathbf{r} + \mathbf{r}_2 - \mathbf{r}_1|} - \frac{1}{|\mathbf{r} + \mathbf{r}_2|} - \frac{1}{|\mathbf{r} - \mathbf{r}_1|} \right) \end{cases} \quad (7.63)$$

If the separation r is large enough, the potential V can be treated as a perturbation with respect to \hat{H}_0 . The unperturbed ground state (of \hat{H}_0) is the tensor product (See Subsection 5.1.1) of two hydrogen atom ground state wavefunctions $|n\ell m\rangle = |100\rangle$:

$$|0^{(0)}\rangle = |100(\mathbf{r}_1)\rangle \otimes |100(\mathbf{r}_2)\rangle, \quad \hat{H}_0 |0^{(0)}\rangle = -2R_\infty |0^{(0)}\rangle. \quad (7.64)$$

For $r \gg a_0$, we can use the multipole expansion to expand the perturbation, $\hat{H}^{(1)}$, in powers of r_i/r

$$V = \frac{e^2}{4\pi\epsilon_0 r^3} (x_1 x_2 + y_1 y_2 - 2z_1 z_2) + \mathcal{O}\left(\frac{1}{r^4}\right). \quad (7.65)$$

To leading-order, V corresponds to the interaction of two electric dipoles $e\mathbf{r}_1, e\mathbf{r}_2$ separated by a distance r .

$$\mathbf{p}_1 = e\mathbf{r}_1 = e(x_1, y_1, z_1) \quad \mathbf{p}_2 = e\mathbf{r}_2 = e(x_2, y_2, z_2) \quad \text{---} \quad \mathbf{r}$$

Fig. 7.6: To leading order, the interaction potential V corresponds to the interaction of two electric dipoles $e\mathbf{r}_1, e\mathbf{r}_2$ separated by a distance r .

The first-order matrix elements of V all vanish (by symmetry)

$$\langle 100 | x_{1,2} | 100 \rangle = \langle 100 | y_{1,2} | 100 \rangle = \langle 100 | z_{1,2} | 100 \rangle = 0. \quad (7.66)$$

Hence the first-order correction to the ground state energy vanishes also

$$E_0^{(1)} = \langle 0^{(0)} | V | 0^{(0)} \rangle = 0. \quad (7.67)$$

² $\hat{a}|n\rangle = \sqrt{n}|n\rangle$, $\hat{a}^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle$.

The first non-vanishing contribution comes at second-order

$$E^{(2)}(r) = \left(\frac{e^2}{4\pi\epsilon_0} \frac{1}{r^3} \right)^2 \underbrace{\sum_{k \neq 0} \frac{\left| \langle k^{(0)} | (x_1 x_2 + y_1 y_2 - 2z_1 z_2) | 0^{(0)} \rangle \right|^2}{E_0^{(0)} - E_k^{(0)}}}_{\text{independent of } r}. \quad (7.68)$$

Hence the interaction varies as $1/r^6$, and, since for all k , the interaction energy is negative (i.e. attractive). This $1/r^6$ long-range, attractive *van der Waals'* potential is a general feature of the interaction between two atoms in their ground state.

7.2 Degenerate Perturbation Theory

The perturbative analysis above is reliable providing that the successive terms in the expansion form a convergent series. A necessary condition is that the matrix elements of the perturbing Hamiltonian must be smaller than the corresponding energy level differences of the original Hamiltonian. If it has different states with the same energy (i.e. degeneracies), and the perturbation has nonzero matrix elements between these degenerate levels, then clearly the theory breaks down. However, the problem is easily fixed.

If a pair of zeroth-order eigenstates $|m^{(0)}\rangle$ and $|n^{(0)}\rangle$ are degenerate in energy, $E_m^{(0)} = E_n^{(0)}$, then as in Eq. (7.12) Eq. (7.10) reduces to

$$\langle m^{(0)} | \hat{H}^{(1)} | n^{(0)} \rangle = 0, \quad m \neq n, \quad E_n^{(0)} = E_m^{(0)}. \quad (7.69)$$

Therefore, to obtain a solution to the perturbation theory equations, the unperturbed (zeroth-order) states must satisfy the condition

$$\{m \neq n, \quad E_m^{(0)} = E_n^{(0)}\} \implies \langle m^{(0)} | \hat{H}^{(1)} | n^{(0)} \rangle = 0. \quad (7.70)$$

However the zeroth-order states $|m^{(0)}\rangle$ and $|n^{(0)}\rangle$, and the perturbation $\hat{H}^{(1)}$, are given quantities over which we have no control, and *there is no guarantee initially that any unperturbed states which are degenerate will satisfy condition (7.70)*. But all is not lost.

Consider an unperturbed energy level which has degeneracy $g > 1$,

$$\hat{H}^{(0)} |n_j^{(0)}\rangle = E_n^{(0)} |n_j^{(0)}\rangle, \quad j = 1, 2, \dots, g. \quad (7.71)$$

But $\{|n_j^{(0)}\rangle\}$ is just one possible choice for the unperturbed basis in the degenerate subspace, linear combinations of $|n_j^{(0)}\rangle$ are still eigenvectors of $\hat{H}^{(0)}$

$$\hat{H}^{(0)} (c_1 |n_j^{(0)}\rangle + c_2 |n_k^{(0)}\rangle) = E_n^{(0)} (c_1 |n_j^{(0)}\rangle + c_2 |n_k^{(0)}\rangle). \quad (7.72)$$

For perturbation theory to work, we want to choose the basis of degenerate states to be linear combinations $|n_{\alpha,\beta,\dots}^{(0)}\rangle$ that satisfy the condition

$$\langle n_{\alpha}^{(0)} | \hat{H}^{(1)} | n_{\beta}^{(0)} \rangle = 0, \quad \alpha \neq \beta, \quad (7.73)$$

or equivalently

$$\hat{H}_{\alpha\beta}^{(1)} \equiv \langle n_{\alpha}^{(0)} | \hat{H}^{(1)} | n_{\beta}^{(0)} \rangle = E'_{n,\alpha} \delta_{\alpha\beta} \quad (7.74)$$

$$\hat{H}' \equiv \text{diag}(E'_{n,1}, E'_{n,2}, \dots, E'_{n,g}), \quad (7.75)$$

where $E'_{n,\alpha}$ are a set of constants.

The $g \times g$ matrix

$$\hat{H}_{jk}^{(1)} \equiv \langle n_j^{(0)} | \hat{H}^{(1)} | n_k^{(0)} \rangle \quad (7.76)$$

has eigenvalues and eigenvectors given by the solutions of the equation

$$\begin{pmatrix} \hat{H}_{11}^{(1)} & \hat{H}_{12}^{(1)} & \cdots & \hat{H}_{1g}^{(1)} \\ \hat{H}_{21}^{(1)} & \hat{H}_{22}^{(1)} & \cdots & \hat{H}_{2g}^{(1)} \\ \vdots & \vdots & \ddots & \vdots \\ \hat{H}_{g1}^{(1)} & \hat{H}_{g2}^{(1)} & \cdots & \hat{H}_{gg}^{(1)} \end{pmatrix} \begin{pmatrix} c_{1\alpha} \\ c_{2\alpha} \\ \vdots \\ c_{g\alpha} \end{pmatrix} = E'_{n,\alpha} \begin{pmatrix} c_{1\alpha} \\ c_{2\alpha} \\ \vdots \\ c_{g\alpha} \end{pmatrix}, \quad (7.77)$$

where the index $\alpha = 1, 2, \dots, g$ labels the g different solutions.

The state

$$|n_{\alpha}^{(0)}\rangle = \sum_{j=1}^g c_{j\alpha} |n_j^{(0)}\rangle, \quad \alpha = 1, 2, \dots, g \quad (7.78)$$

are the (mutually orthogonal) eigenvectors associated to the eigenvalues $N'_{n,\alpha}$ and satisfy the condition (7.73). They form a *good basis* for perturbation theory.

There is a rule that helps us identify a good basis without the need to explicitly diagonalise the matrix in Eq. (7.77): Let \hat{H} be a hermitian operator that commutes with the full hamiltonian \hat{H} . If the vectors of the unperturbed basis are also eigenstates of \hat{A} with different eigenvalues, then the basis is “good”.

Consider two eigenstates $|n_p^{(0)}\rangle$ and $|n_q^{(0)}\rangle$, with \hat{A} eigenvalues $a_p \neq a_q$, then

$$\left. \begin{aligned} 0 &= \langle n_p^{(0)} | [\hat{A}, \hat{H}^{(1)}] | n_q^{(0)} \rangle \\ &= \left. \begin{aligned} &= \langle n_p^{(0)} | \hat{A} \hat{H}^{(1)} | n_q^{(0)} \rangle - \langle n_p^{(0)} | \hat{H}^{(1)} \hat{A} | n_q^{(0)} \rangle \end{aligned} \right\} \quad \Rightarrow \quad \langle n_p^{(0)} | \hat{H}^{(1)} | n_q^{(0)} \rangle = 0, \quad p \neq q. \\ &= (a_p - a_q) \langle n_p^{(0)} | \hat{H}^{(1)} | n_q^{(0)} \rangle \end{aligned} \right\} \quad (7.79)$$

Provided we work in the zeroth-order basis $|n_{\alpha}^{(0)}\rangle$, rather than $|n_j^{(0)}\rangle$, all the results obtained earlier can safely be applied to degenerate states. In particular, the first order correction to each of the g degenerate states are given by Eq. (7.9)

$$E_{n,\alpha}^{(1)} = \langle n_{\alpha}^{(0)} | \hat{H}^{(1)} | n_{\alpha}^{(0)} \rangle, \quad \alpha = 1, 2, \dots, g. \quad (7.80)$$

In general the eigenvalues $E_{n,\alpha}^{(1)}$ will be different for each value of α . In that case, the original energy level $E_n^{(0)}$ splits into (up to) g distinct levels in the presence of $\hat{H}^{(1)}$ (at first order).

In summary, for an energy level n with degeneracy $g > 1$, the procedure is:

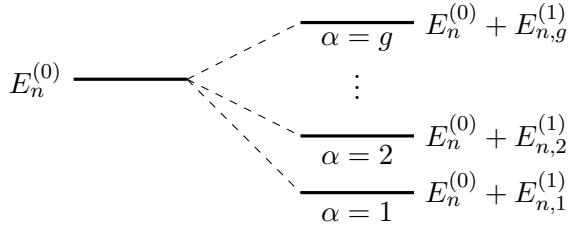


Fig. 7.7: The eigenvalues $E_{n,\alpha}^{(1)}$ will be different for each value of α , splitting the original energy level $E_n^{(0)}$ into distinct levels in the presence of $\hat{H}^{(1)}$.

- Find the matrix representation $\langle n_j^{(0)} | \hat{H}^{(1)} | n_k^{(0)} \rangle$ of the perturbation in the original basis states $\{|n_j^{(0)}\rangle\}$ for that energy level and diagonalise it.
- Find the eigenvalues $E_{n,\alpha}^{(1)}$ and eigenvectors $|n_\alpha^{(0)}\rangle$ of the matrix $\hat{H}^{(1)}$.
- Alternatively, use the rule in Eq. (7.79). to identify a good basis directly.

Operating in the good basis $|n_\alpha^{(0)}\rangle$:

- Results obtained for non-degenerate states apply also to degenerate states
 - The eigenvalues $E_{n,\alpha}^{(1)}$ give the first order energy correction directly, and the original energy level splits into (up to) g separate levels.
 - The first-order correction to the eigenstate misses the piece along the degenerate subspace, i.e. $\langle n_\alpha^{(0)} | n_\beta^{(1)} \rangle$ for $\alpha \neq \beta$.

To understand how this works, let us consider a particular example.

Recall that, for the simple harmonic oscillator, $\hat{H} = \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega^2x^2$, the ground state wavefunction is given by $\langle x|0\rangle = (\frac{m\omega}{\pi\hbar})^{1/4}e^{-\xi^2/2}$, where $\xi = x\sqrt{m\omega/\hbar}$ and the first excited state by $\langle x|1\rangle = \left(\frac{4m\omega}{\pi\hbar}\right)^{1/4}\xi e^{-\xi^2/2}$. The wavefunctions for the two-dimensional harmonic oscillator,

$$\hat{H}^{(0)} \equiv \frac{1}{2m}(\hat{p}_x^2 + \hat{p}_y^2) + \frac{1}{2}m\omega^2(x^2 + y^2), \quad (7.81)$$

are given simply by the product of two one-dimensional oscillators. So, setting $\eta = y\sqrt{m\omega/\hbar}$, the ground state is given by $\langle x, y|0, 0\rangle = (\frac{m\omega}{\pi\hbar})^{1/2}e^{-(\xi+\eta)^2/2}$, and the two *degenerate* first excited states, an energy $\hbar\omega$ above the ground state, are given by

$$\begin{cases} \langle x, y|1, 0\rangle \\ \langle x, y|0, 1\rangle \end{cases} = \left(\frac{m\omega}{\pi\hbar}\right)^{1/2}e^{-(\xi+\eta)^2/2} \begin{cases} \xi \\ \eta \end{cases}. \quad (7.82)$$

Suppose now we add to the Hamiltonian a perturbation

$$\hat{H}^{(1)} = \alpha m\omega^2 xy = \alpha\hbar\omega\xi\eta, \quad (7.83)$$

controlled by a small parameter α . Notice that, by symmetry, the following matrix elements all vanish, $\langle \hat{H}^{(1)} | 0, 0 \rangle = \langle \hat{H}^{(1)} | 1, 0 \rangle = \langle \hat{H}^{(1)} | 0, 1 \rangle = 0$. Therefore, according to a naïve perturbation theory, there is no first-order correction to the energies of these states. However, on proceeding to consider the second-order correction to the energy, the theory breaks down. The off-diagonal matrix element $\langle 1, 0 | \hat{H}^{(1)} | 0, 1 \rangle$ is non-zero, but the two states $|0, 1\rangle$ and $|1, 0\rangle$ have the same energy! This gives an infinite term in the series for $E_{n=1}^{(2)}$.

Yet we know that a small perturbation of this type will not wreck a two-dimensional simple harmonic oscillator – so what is wrong with our approach? To understand the origin of the problem and its fix, it is helpful to plot the original harmonic oscillator potential $\frac{1}{2}m\omega^2(x^2 + y^2)$ together with the perturbing potential $\alpha m\omega^2 xy$. The first of course has circular symmetry, the second has two symmetry axes oriented in the directions $x = \pm y$, climbing most steeply from the origin along $x = y$, falling most rapidly in the directions $x = -y$. If we combine the two potentials into a single quadratic form,

$$\frac{1}{2}m\omega^2(x^2 + y^2) + \alpha m\omega^2 xy = \frac{1}{2}m\omega^2 \left[(1 + \alpha) \left(\frac{x+y}{2} \right)^2 + (1 - \alpha) \left(\frac{x-y}{2} \right)^2 \right]. \quad (7.84)$$

the original circles of constant potential become ellipses, with their axes aligned along $x = \pm y$.

As soon as the perturbation is introduced, the eigenstates lie in the direction of the new elliptic axes. This is a large change from the original x and y bases, which is not proportional to the small parameter α . But the original unperturbed problem had circular symmetry, and there was no particular reason to choose the x and y axes as we did. If we had instead chosen as our original axes the lines $x = \pm y$, the basis states would not have undergone large changes on switching on the perturbation. The resolution of the problem is now clear: *Before switching on the perturbation, one must choose a set of basis states in a degenerate subspace in which the perturbation is diagonal.*

In fact, for the simple harmonic oscillator example above, the problem can of course be solved exactly by rearranging the coordinates to lie along the symmetry axes, $(x \pm y)/\sqrt{2}$. It is then clear that, despite the results of naïve first order perturbation theory, there is indeed a *first order shift* in the energy levels, $\hbar\omega \rightarrow \hbar\omega\sqrt{1 \pm \alpha} \approx \hbar\omega(1 \pm \alpha/2)$.

7.2.1 Example: Perturbed 2D Infinite Square Well

Consider a particle of mass m confined within a 2D infinite square well of dimension $2a \times 2a$

$$\hat{H}_0 = \frac{\hat{\mathbf{p}}^2}{2m} + \hat{V}(\hat{x}, \hat{y}), \quad \hat{V}(\hat{x}, \hat{y}) = \begin{cases} 0 & |x| < a, |y| < a \\ \infty & \text{otherwise} \end{cases}. \quad (7.85)$$

The unperturbed Hamiltonian \hat{H}_0 has eigenstates which are tensor products of eigenstates $|n_x\rangle$ and $|n_y\rangle$ of separate 1D infinite wells for x and y

$$|n_x, n_y\rangle = |n_x\rangle \otimes |n_y\rangle, \quad n_x, n_y = 1, 2, 3, \dots, \quad (7.86)$$

with wavefunctions

$$\psi(x, y) = \frac{1}{a} \sin\left(\frac{n_x \pi}{2a}(x + a)\right) \sin\left(\frac{n_y \pi}{2a}(y + a)\right). \quad (7.87)$$

The unperturbed 2D energies are the sum of the two 1D energies

$$E_{n_x, n_y}^{(0)} = (n_x^2 + n_y^2) E_0, \quad E_0 = \frac{\hbar^2 \pi^2}{8ma^2}. \quad (7.88)$$

All states with $n_x \neq n_y$ are degenerate, with degeneracy $g = 2$ e.g. the first excited state consists of the two degenerate states

$$|n_x, n_y\rangle = |1, 2\rangle, |2, 1\rangle, \quad E_{1,2}^{(0)} = E_{2,1}^{(0)} = 5E_0. \quad (7.89)$$

Now, apply a perturbation

$$\hat{H} = \varepsilon \hat{x} \hat{y}. \quad (7.90)$$

The perturbation $\hat{H}^{(1)}$ has matrix elements which factorise as

$$\langle m_x, m_y | \hat{H}^{(1)} | n_x, n_y \rangle = \varepsilon \langle m_x | \hat{x} | n_x \rangle \langle m_y | \hat{y} | n_y \rangle. \quad (7.91)$$

The factors on the right side evaluate from considering the relevant factor in Eq. (7.87)

$$\langle m_x | \hat{x} | n_x \rangle = \frac{1}{a} \int_{-a}^{+a} x \sin\left(\frac{m_x \pi}{2a}(x + a)\right) \sin\left(\frac{n_x \pi}{2a}(x + a)\right) dx, \quad (7.92)$$

in particular,

$$\langle 1 | \hat{x} | 2 \rangle = \frac{1}{a} \int_{-a}^{+a} x \sin\left(\frac{\pi}{2a}(x + a)\right) \sin\left(\frac{\pi}{a}(x + a)\right) dx = -\frac{32}{9} \frac{a}{\pi^2}. \quad (7.93)$$

For the first excited state, we thus have

$$\langle 1, 2 | \hat{x} \hat{y} | 2, 1 \rangle = \langle 1 | \hat{x} | 2 \rangle \langle 2 | \hat{y} | 1 \rangle = a^2 \left(\frac{32}{9\pi^2}\right)^2. \quad (7.94)$$

Thus the matrix representation of $\hat{H}^{(1)}$ is not (initially) diagonal

$$\langle 1, 2 | \hat{H}^{(1)} | 2, 1 \rangle = \langle 2, 1 | \hat{H}^{(1)} | 1, 2 \rangle = A, \quad A = \varepsilon^2 \left(\frac{32}{9\pi^2}\right)^2 \quad (7.95)$$

to apply perturbation theory, we must find suitable linear combinations of the initial states $|1, 2\rangle$ and $|2, 1\rangle$.

These are found by diagonalising the 2×2 matrix

$$\begin{pmatrix} \langle 1, 2 | \hat{H}^{(1)} | 1, 2 \rangle & \langle 1, 2 | \hat{H}^{(1)} | 2, 1 \rangle \\ \langle 2, 1 | \hat{H}^{(1)} | 1, 2 \rangle & \langle 2, 1 | \hat{H}^{(1)} | 2, 1 \rangle \end{pmatrix} = \begin{pmatrix} 0 & A \\ A & 0 \end{pmatrix}. \quad (7.96)$$

This matrix has eigenvalues $\pm A$ and eigenvectors $(1, \pm 1)$. Hence the (normalised) zeroth-order eigenstates which must be used in perturbation theory are

$$|\psi_{\pm}\rangle \equiv |1, 2\rangle_{\pm} = \frac{1}{\sqrt{2}}(|1, 2\rangle \pm |2, 1\rangle), \quad (7.97)$$

it is straightforward to check that these states do indeed diagonalise the perturbation

$$\begin{pmatrix} \langle \psi_+ | \hat{H}^{(1)} | \psi_+ \rangle & \langle \psi_+ | \hat{H}^{(1)} | \psi_- \rangle \\ \langle \psi_- | \hat{H}^{(1)} | \psi_+ \rangle & \langle \psi_- | \hat{H}^{(1)} | \psi_- \rangle \end{pmatrix} = \begin{pmatrix} A & 0 \\ 0 & -A \end{pmatrix}. \quad (7.98)$$

The first-order energy corrections are given, with no further work, by the diagonal elements above

$$E_n^{(1)} = \langle \psi_\pm | \hat{H}^{(1)} | \psi_\pm \rangle = \pm A = \pm \varepsilon a^2 \left(\frac{32}{9\pi^2} \right)^2. \quad (7.99)$$

In summary, application of the perturbation

$$\hat{H}^{(1)} = \varepsilon \hat{x} \hat{y} \quad (7.100)$$

to the 2D infinite square well splits the degenerate ($g = 2$) first-excited level into two separate levels where $A = \varepsilon a^2 (32/9\pi^2)^2$.

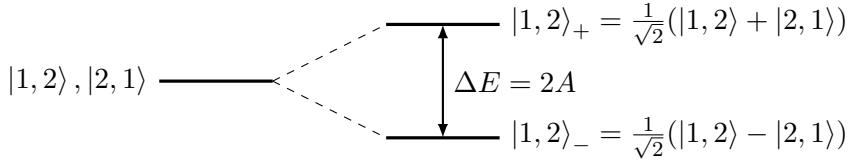


Fig. 7.8: A perturbation $\hat{H}^{(1)} = \varepsilon \hat{x} \hat{y}$ to the 2D infinite square well splits the degenerate ($g = 2$) first-excited level into two separate levels.

The zeroth-order states $|1, 2\rangle_+$ and $|1, 2\rangle_-$ could be used to compute the energy corrections and wavefunctions to any order of perturbation theory.

7.2.2 Example: Linear Stark Effect

As with the two-dimensional harmonic oscillator discussed above, the hydrogen atom has a non-degenerate ground state, but degeneracy in its lowest excited states. Specifically, there are four $n = 2$ states, all having energy $-\frac{1}{4}\text{Ry}$. In spherical coordinates, these wavefunctions are given by

$$\begin{cases} \psi_{200}(\mathbf{r}) \\ \psi_{210}(\mathbf{r}) \\ \psi_{21,\pm 1}(\mathbf{r}) \end{cases} = \left(\frac{1}{32\pi a_0^3} \right)^{1/2} e^{-r/2a_0} \begin{cases} \left(2 - \frac{r}{a_0} \right) \\ \frac{r}{a_0} \cos \theta \\ \frac{r}{a_0} e^{\pm i\phi} \sin \theta \end{cases}. \quad (7.101)$$

When perturbing this system with an electric field oriented in the z -direction, $\hat{H}^{(1)} = e\mathcal{E}r \cos \theta$, a naïve application of perturbation theory predicts no first-order shift in any of these energy levels. However, to first order in E there is a non-zero matrix element between two degenerate levels $\Delta = \langle 200 | \hat{H}^{(1)} | 210 \rangle$. All the other matrix elements between these basis states in the four-dimensional degenerate subspace are zero. So the only diagonalisation necessary is within the two-dimensional degenerate subspace spanned by $|200\rangle$ and $|210\rangle$, i.e.

$$\hat{H}^{(1)} = \begin{pmatrix} 0 & \Delta \\ \Delta & 0 \end{pmatrix}, \quad (7.102)$$

$$\text{with } \Delta = e\mathcal{E} \left(\frac{1}{32\pi a_0^3} \right) \int \left(2 - \frac{r}{a_0} \right) \frac{(r \cos \theta)^2}{a_0} e^{-r/a_0} r^2 dr \sin \theta d\theta d\phi = -3e\mathcal{E}a_0.$$

Diagonalising $\hat{H}^{(1)}$ within this sub-space, the new basis states are given by the symmetric and antisymmetric combinations, $(|200\rangle + |210\rangle)/\sqrt{2}$ with energy shifts $\pm\Delta$, *linear* in the perturbing electric field. The states $|2\ell, \pm 1\rangle$ are not changed by the presence of the field to this level of approximation, so the complete energy map of the $n = 2$ states in the electric field has two states at the original energy of $-\text{Ry}/4$, one state moved up from that energy by Δ , and one down by Δ . Notice that the new eigenstates $(|200\rangle \pm |210\rangle)/\sqrt{2}$ are *not* eigenstates of the parity operator - a sketch of their wavefunctions reveals that, in fact, they have non-vanishing electric dipole moment \mathbf{p} . Indeed this is the reason for the energy shift, $\pm\Delta = \mp 2e\mathcal{E}a_0 = \mp \mathbf{p} \cdot \mathbf{\mathcal{E}}$.

7.2.3 Nearly Free Electron Model

As a second and important example of the degenerate perturbation theory, let us consider the problem of a particle moving in one dimension and subject to a weak periodic potential, $V(x) = 2V \cos(2\pi x/a)$ - the **nearly free electron model**. This problem provides a caricature of a simple crystalline solid in which (free) conduction electrons propagate in the presence of a periodic background lattice potential. Here we suppose that the strength of the potential V is small as compared to the typical energy scale of the particle so that it may be treated as a small perturbation. In the following, we will suppose that the total one-dimensional system is of length $L = Na$, with periodic boundary conditions.

For the unperturbed free particle system, the eigenstates are simply plane waves $\psi_k(x) = \langle x|k\rangle = \frac{1}{\sqrt{L}}e^{ikx}$ indexed by the wavenumber $k = 2\pi n/L$, n integer, and the unperturbed spectrum is given by $E_k^{(0)} = \hbar^2 k^2 / 2m$. The matrix elements of the perturbation between states of different wavevector are given by

$$\begin{aligned} \langle k|V|k'\rangle &= \frac{1}{L} \int_0^L dx e^{i(k-k')x} 2V \cos(2\pi x/a) \\ &= \frac{V}{L} \int_0^L \left(e^{i(k-k'+2\pi x/a)x} + e^{i(k-k'-2\pi x/a)x} \right) = V \delta_{k'-k, 2\pi/a}. \end{aligned} \quad (7.103)$$

Note that all diagonal matrix elements of the perturbation are identically zero. In general, for wavevectors k and k' separated by $G = 2\pi/a$, the unperturbed states are non-degenerate. For these states one can compute the relative energy shift within the framework of second order perturbation theory. However, for states $k = -k' = G/2 \equiv \pi/2$, the unperturbed free particle spectrum is degenerate. Here, and in the neighbourhood of these k values, we must implement a degenerate perturbation theory.

For the sinusoidal potential considered here, only states separated by $G = 2\pi/a$ are coupled by the perturbation. We may therefore consider matrix elements of the full Hamiltonian between pairs of coupled states, $|k = G/2 + q\rangle$ and $|k = -G/2 + q\rangle$

$$H = \begin{pmatrix} E_{G/2+q}^{(0)} & V \\ V & E_{-G/2+q}^{(0)} \end{pmatrix}. \quad (7.104)$$

As a result, to leading order in V , we obtain the eigenvalues,

$$E_q = \frac{\hbar^2}{2m} \left(q^2 + (\pi/a)^2 \right) \pm \left(V^2 + \frac{\pi^2 \hbar^4 q^2}{4m^2 a^2} \right)^{1/2}. \quad (7.105)$$

In particular, this result shows that, for $k = \pm G/2$, the degeneracy of the free particle system is lifted by the potential. In the vicinity, $|q| \ll G$, the spectrum of eigenvalues is separated by a gap of size $2V$. The appearance of the gap mirrors the behaviour found in our study of the Kronig-Penney model of a crystal studied in section 2.2.3.

The appearance of the gap has important consequences in theory of solids. Electrons are fermions and have to obey Pauli's exclusion principle. In a metal, at low temperatures, electrons occupy the free particle-like states up to some (Fermi) energy which lies away from gap. Here, the accessibility of very low-energy excitations due to the continuum of nearby states allows current to flow when a small electric field is applied. However, when the Fermi energy lies in the gap created by the lattice potential, an electric field is unable to create excitations and induce current flow. Such systems are described as **(band) insulators**.

7.3 Variational Method

So far, in devising approximation methods for quantum mechanics, we have focused on the development of a perturbative expansion scheme in which the states of the non-perturbed system provided a suitable platform. Here, by “suitable” we refer to situations in which the states of the unperturbed system mirror those of the full system – adiabatic continuity. For example, the states of the harmonic oscillator potential with a small perturbation will mirror those of the unperturbed Hamiltonian: The ground state will be nodeless, the first excited state will be antisymmetric having one node, and so on. However, often we work with systems where the true eigenstates of the problem may not be adiabatically connected to some simple unperturbed reference state. This situation is particularly significant in strongly interacting quantum systems where many-particle correlations can effect phase transitions to new states of matter – e.g. the development of superfluid condensates, or the fractional quantum Hall fluid. To address such systems it is often extremely effective to “guess” and then optimise a trial wavefunction. The method of optimisation relies upon a simple theoretical framework known as the variational approach. For reasons that will become clear, the variational method is particularly wellsuited to addressing the ground state.

The variational method involves the optimisation of some trial wavefunction on the basis of one or more adjustable parameters. The optimisation is achieved by minimising the expectation value of the energy on the trial function, and thereby finding the best approximation to the true ground state wave function. This seemingly crude approach can, in fact, give a surprisingly good approximation to the ground state energy but, it is usually not so good for the wavefunction, as will become clear. However, as mentioned above, the real strength of the variational method arises in the study of many-body quantum systems, where states are more strongly constrained by fundamental symmetries such as “exclusion statistics”.

To develop the method, we'll begin with the problem of a single quantum particle confined to a potential, $\hat{H} = \frac{\hat{p}^2}{2m} + V(\hat{r})$. If the particle is restricted to one dimension, and we're looking for the ground state in any fairly localised potential well, it makes sense to start with a trial wavefunction which belongs to the family of normalised Gaussians, $\langle x | \psi(\alpha) \rangle = (\alpha/\pi)^{1/4} e^{-\alpha x^2/2}$. Such a trial state fulfils the criterion of being nodeless, and is exponentially localised to the region of the binding potential. It also has the feature that it includes the exact ground states of the harmonic binding potential.

The variation approach involves simply minimising the expectation value of the energy, $E = \langle \psi(\alpha) | \hat{H} | \psi(\alpha) \rangle$, with respect to variations of the **variational parameter**, α . (Of course, as with any minimisation, one must check that the variation does not lead to a maximum of the energy!) Not surprisingly, this programme leads to the exact ground state for the simple harmonic oscillator potential, while it serves only as an approximation for other potentials. What is perhaps surprising is that the result is only off by only ca. 30% or so for the attractive δ -function potential, even though the wavefunction looks substantially different. The Gaussian family cannot be used if there is an infinite wall anywhere: since the wavefunction must vanish where the potential is infinite, one must find a family of wavefunctions that vanishes at such a boundary.

To gain some further insight into the approach, suppose the Hamiltonian \hat{H} has a set of eigenstates, $\hat{H} |n\rangle = E_n |n\rangle$. Since the Hamiltonian is Hermitian, these states span the space of possible wave functions, including our variational family of Gaussians, so we can write, $|\psi(\alpha)\rangle = \sum_n a_n(\alpha) |n\rangle$. From this expansion, we have

$$\frac{\langle \psi(\alpha) | \hat{H} | \psi(\alpha) \rangle}{\langle \psi(\alpha) | \psi(\alpha) \rangle} = \sum_n |a_n|^2 E_n \geq E_0. \quad (7.106)$$

for any $|\psi(\alpha)\rangle$. (We don't need the denominator if we've chosen a family of normalised wavefunctions, as we did with the Gaussians above.) Evidently, minimising the left hand side of this equation as function of α provides an **upper bound** on the ground state energy.

We can see immediately that this will probably be better for finding the ground state energy than for the wavefunction: Suppose the optimum state in our family is given by, say, $|\alpha_{\min}\rangle = N(|0\rangle + 0.2|1\rangle)$ with the normalisation $N \simeq 0.98$, i.e. a 20% admixture of the first excited state. Then the wavefunction is off by ca. 20%, but the energy estimate will be too high by only $0.04(E_1 - E_0)$, usually a much smaller error.

7.3.1 Example: Ground State of Hydrogen Atom

To get some idea of how well the variational approach works, consider its application to the ground state of the hydrogen atom. Taking into account the spherical symmetry of the ground state, we may focus on the one-dimensional radial component of the wavefunction. Defining the trial radial wavefunction $u(\rho)$ (presumed real), where $\rho = r/a_0$, the variational energy is given by

$$E(u) = -Ry \frac{\int_0^\infty d\rho u(\rho) \left(\frac{d^2}{d\rho^2} + \frac{2}{\rho} \right) u(\rho)}{\int_0^\infty d\rho u^2(\rho)}. \quad (7.107)$$

For the three families of trial functions,

$$u_1(\rho) = \rho e^{-\alpha\rho}, \quad u_2(\rho) = \frac{\rho}{\alpha^2 + \rho^2}, \quad u_3(\rho) = \rho^2 e^{-\alpha\rho}, \quad (7.108)$$

and finds $\alpha_{\min} = 1, \pi/4$, and $3/2$ respectively. The first family, u_1 , includes the exact result, and the minimisation procedure correctly finds it. For the three families, the predicted energy of the optimal state is off by 0, 25%, and 21% respectively.

The corresponding error in the wavefunction is defined by how far the square of the overlap with the true ground state wavefunction falls short of unity. For the three families, $\varepsilon = 1 - |\langle \psi_0 | \psi_{\text{var}} \rangle|^2 = 0, 0.21$, and 0.05 . Notice here that our handwaving argument that the energies would be found much more accurately than the wavefunctions seems to come unstuck. The third family has far better wavefunction overlap than the second, but only a slightly better energy estimate. Why? A key point is that the potential is singular at the origin; there is a big contribution to the potential energy from a rather small region, and the third family of trial states is the least accurate of the three there. The second family of functions are very inaccurate at large distances: the expectation value $\langle r \rangle = 1.5a_0, \infty, 1.66a_0$ for the three families. But at large distances, both kinetic and potential energies are small, so the result can still look reasonable. These examples reinforce the point that the variational method should be implemented with some caution.

In some cases, one can exploit symmetry to address the properties of higher-lying states. For example, if the one-dimensional attractive potential is symmetric about the origin, and has more than one bound state, the ground state will be even, the first excited state odd. Therefore, we can estimate the energy of the first excited state by minimising a family of odd functions, such as $\psi(x, \alpha) = \left(\frac{\sqrt{\pi}}{2a^{3/2}}\right)^{1/2} xe^{-\alpha x^2/2}$.

7.3.2 Helium Atom Addressed by the Variational Approach

For the hydrogen atom, we know that the ground state energy is -1Ry , or -13.6eV . The He^+ ion (with just a single electron) has a nuclear charge of $Z = 2$, so the ground state energy of the electron, being proportional to Z^2 , will now be equal to -4Ry . Therefore, for the He atom, if we neglect their mutual interaction, the electrons will occupy the ground state wavefunction having opposite spin, leading to a total ground state energy of -8Ry or -109eV . In practice, as we have seen earlier, the repulsion between the electrons raises the ground state energy to -79eV (see Section 7.1.2.1).

To get a better estimate for the ground state energy, one can retain the form of the ionic wavefunction, $\left(\frac{Z^3}{\pi a_0^3}\right)^{1/2} e^{-Zr/\pi a_0}$, but now with Z a variational parameter, rather than setting it equal to the nuclear charge $Z = 2$. In other words, let us accommodate the effects of electron-electron repulsion, which must “push” the wavefunctions to larger radii, by keeping exactly the same wavefunction profile but lessening the effective nuclear charge as reflected in the spread of the wavefunction from $Z = 2$ to $Z < 2$. The precise value will be set by varying it to find the minimum total energy, including the term from electron-electron repulsion.

To find the potential energy from the nuclear-electron interactions, we of course must use the actual nuclear charge $Z = 2$, but impose a variable Z for the wavefunction, so the

nuclear potential energy for the two electrons is given by,

$$\text{p.e.} = -2 \times \frac{2e^2}{4\pi\epsilon_0} \int_0^\infty 4\pi r^2 dr \frac{\mathcal{Z}^3}{\pi a_0^3} \frac{e^{-2\mathcal{Z}r/a_0}}{r} = -\mathcal{Z} \frac{e^2}{\pi\epsilon_0 a_0} = -8\mathcal{Z}\text{Ry}. \quad (7.109)$$

This could have been inferred from the formula for the one electron ion of charge Z , where the potential energy for the one electron is $-2Z^2\text{Ry}$, one factor of Z being from the nuclear charge, the other from the consequent shrinking of the orbit. The kinetic energy is even easier to determine: it depends entirely on the form of the wavefunction, and not on the actual nuclear charge. So for our trial wavefunction it has to be $\mathcal{Z}^2\text{Ry}$ per electron. Finally, making use of our calculation in Section 7.1.2.1, we can immediately write down the positive contribution to the energy expectation value from the electron-electron interaction,

$$\frac{e^2}{4\pi\epsilon_0} \frac{\mathcal{Z}^2}{(\pi a_0^3)^2} \int d\mathbf{r}_1 d\mathbf{r}_2 \frac{e^{-2\mathcal{Z}(r_1+r_2)/a_0}}{|\mathbf{r}_1 - \mathbf{r}_2|} = \frac{5}{4} \frac{e^2}{4\pi\epsilon_0} \frac{\mathcal{Z}}{2a_0} = \frac{5}{4} \mathcal{Z}\text{Ry}. \quad (7.110)$$

Collecting all of the terms, the total variational state energy is given by:

$$E = -2 \left(4\mathcal{Z} - \mathcal{Z}^2 - \frac{5}{8} \mathcal{Z} \right) \text{Ry}. \quad (7.111)$$

Minimisation of this energy with respect to \mathcal{Z} obtains the minimum at $\mathcal{Z} = 2 - \frac{5}{16}$, leading to an energy of -77.5eV . This result departs from the true value by about 1eV . So, indeed, the presence of the other electron leads effectively to a shielding of the nuclear charge by an amount of ca. $(5/16)e$.

7.3.3 The Rayleigh-Ritz Method

The *Rayleigh-Ritz* method is a particular example of the variational method which employs a trial wavefunction of the form

$$\psi(\mathbf{r}) = \sum_j \alpha_j \psi_j(\mathbf{r}),$$

(7.112)

where the $\psi_j(\mathbf{r})$ are a set of linearly independent wavefunctions, the α_j are variational parameters, and *the states ψ_j do not necessarily have to be complete or orthonormal*.

For example, in molecular physics, states can be approximated as a linear combination of single-electron atomic states (*Linear Combination of Atomic Orbitals* (LCAO)).

The Rayleigh-Ritz method can then be used to determine the optimal linear combination.

The quantity to be minimised is the generalised expectation value

$$\langle E \rangle = \frac{\langle \psi | \hat{H} | \psi \rangle}{\langle \psi | \psi \rangle} = \frac{\sum_{j,k} \alpha_j \alpha_k \langle \psi_j | \hat{H} | \psi_k \rangle}{\sum_{j,k} \alpha_j \alpha_k \langle \psi_j | \psi_k \rangle} = \frac{\sum_{j,k} \alpha_j \alpha_k H_{jk}}{\sum_{j,k} \alpha_j \alpha_k S_{jk}}, \quad (7.113)$$

where the constants H_{jk} and S_{jk} are defined as

$$H_{jk} \equiv \langle \psi_j | \hat{H} | \psi_k \rangle = H_{kj}^*, \quad S_{jk} \equiv \langle \psi_j | \psi_k \rangle = S_{kj}^*. \quad (7.114)$$

If $\{|\psi\rangle\}$ is an orthonormal set, then we have $\langle \psi_j | \psi_k \rangle = \delta_{jk} \implies S_{jk} = \delta_{jk}$.

Minimising $\langle E \rangle$ with respect to the variational parameter α_i gives³

$$0 = \frac{\partial \langle E \rangle}{\partial \alpha_i} = \frac{\sum_k \alpha_k H_{ik} + \sum_j \alpha_j H_{ji}}{\sum_{j,k} \alpha_j \alpha_k S_{jk}} - \frac{\sum_{j,k} \alpha_j \alpha_k H_{jk}}{\left(\sum_{j,k} \alpha_j \alpha_k S_{jk}\right)^2} \left(\sum_k \alpha_k S_{ik} + \sum_j \alpha_j S_{ij} \right). \quad (7.116)$$

This tidies up (for each value of the index i) as

$$\left(\sum_k \alpha_k (H_{ik} + H_{ki}) \right) \left(\sum_{j,k} \alpha_j \alpha_k S_{jk} \right) = \left(\sum_k \alpha_k (S_{ik} + S_{ki}) \right) \left(\sum_{j,k} \alpha_j \alpha_k H_{jk} \right) \quad (7.117)$$

We shall only need to consider the case that the H_{jk} and S_{jk} are both real matrices, the equation above then simplifies (for each i) as

$$\left(\sum_k \alpha_k H_{ik} \right) \left(\sum_{j,k} \alpha_j \alpha_k S_{jk} \right) = \left(\sum_k \alpha_k S_{ik} \right) \left(\sum_{j,k} \alpha_j \alpha_k H_{jk} \right). \quad (7.118)$$

If the minimum value of $\langle E \rangle$ is E_{\min} , then from Eq. (7.113) we also have

$$\left(\sum_{j,k} \alpha_j \alpha_k S_{jk} \right) E_{\min} = \left(\sum_{j,k} \alpha_j \alpha_k H_{jk} \right) \quad (7.119)$$

Comparing Eqs. (7.118) and (7.119), the minimum corresponds to

$$\sum_k \alpha_k (H_{ik} - E_{\min} S_{ik}) = 0 \quad (7.120)$$

for each i . This represents a set of simultaneous equations for the α_i , which can be written in matrix form as

$$(H - E_{\min} S) \alpha = 0. \quad (7.121)$$

Non-trivial solutions are obtained by solving the secular equation

$$\det(H - E_{\min} S) = 0. \quad (7.122)$$

The lowest of the resulting set of eigenvalues E_{\min} provides the best upper bound on the ground state energy

This completes our discussion of the principles of the variational approach. However, later in the course, we will find the variational methods appearing in several important applications.

³Making use of the identities

$$\frac{\partial \alpha_j}{\partial \alpha_i} = \delta_{ji}, \quad \frac{\partial \alpha_k}{\partial \alpha_i} = \delta_{ki}. \quad (7.115)$$

7.3.4 Example: Hydrogen Atom with Finite Proton Mass

A finite proton mass m_p can be taken into account exactly by replacing the electron mass m_e by the reduced mass μ

$$\hat{H} = -\frac{\hbar^2}{2\mu}\nabla^2 - \frac{e^2}{4\pi\epsilon_0 r}, \quad \mu = \frac{m_p m_e}{m_p + m_e} \approx (0.999456)m_e. \quad (7.123)$$

The hydrogen atom ground state energy thus becomes

$$E_0 = -\left(\frac{e^2}{4\pi\epsilon_0\hbar c}\right)^2 \frac{1}{2}\mu c^2 = -\frac{\mu}{m_e} R_\infty \approx -(0.999456)R_\infty. \quad (7.124)$$

To test / demonstrate the Rayleigh-Ritz method, we consider the same problem using a trial wavefunction of the form

$$\psi_{\text{trial}} = \alpha_1 |100\rangle + \alpha_2 |200\rangle, \quad (7.125)$$

which is a linear combination of the *infinite-mass* states $|n\ell m_\ell\rangle$, i.e. the wavefunctions $|100\rangle$ and $|200\rangle$ are expressed in terms of m_e , *not* the reduced mass μ .

Thus we have,

$$\left. \begin{aligned} |100\rangle &= \left(\frac{1}{\pi a_0^3}\right)^{1/2} e^{-r/a_0} \\ |200\rangle &= \left(\frac{1}{32\pi a_0^3}\right)^{1/2} e^{-r/2a_0} \left(2 - \frac{r}{a_0}\right) \end{aligned} \right\} a_0 = \frac{4\pi\epsilon_0}{e^2} \frac{\hbar^2}{m_e}. \quad (7.126)$$

In this case, the basis states happen to be orthonormal

$$S_{jk} \equiv \langle \psi_j | \psi_k \rangle, \quad S_{jk} = \delta_{jk}. \quad (7.127)$$

The Hamiltonian matrix elements are given by

$$H_{jk} \equiv \langle \psi_j | \hat{H} | \psi_k \rangle = \int \psi_j^* \left(-\frac{\hbar^2}{2\mu} \nabla^2 - \frac{e^2}{4\pi\epsilon_0 r} \right) \psi_k d^3\mathbf{r}, \quad (7.128)$$

e.g.

$$H_{21} = A \int_0^\infty e^{-r/2a_0} \left(2 - \frac{r}{a_0}\right) \left[\left(-\frac{\hbar^2}{2\mu} \nabla^2 - \frac{e^2}{4\pi\epsilon_0 r}\right) e^{-r/a_0} \right] 4\pi r^2 dr, \quad (7.129)$$

where,

$$A = \left(\frac{1}{32\pi a_0^3}\right)^{1/2} \left(\frac{1}{\pi a_0^3}\right)^{1/2} = \frac{1}{\sqrt{32}\pi a_0^3}. \quad (7.130)$$

Using the identities

$$\nabla^2 e^{-\beta r} = \left(\frac{2}{r} \frac{\partial}{\partial r} + \frac{\partial^2}{\partial r^2}\right) e^{-\beta r} = -\beta \left(\frac{2}{r} - \beta\right) e^{-\beta r}, \quad \beta \equiv \frac{1}{a_0} \quad (7.131)$$

and

$$\int_0^\infty r^n e^{-\beta r} dr = \frac{n!}{\beta^{n+1}} \quad (7.132)$$

gives

$$H_{11} = 4H_{22} = \left(\frac{m_e}{m_p} - 1 \right) R_\infty, \quad H_{12} = H_{21} = \frac{16}{27\sqrt{2}} \frac{m_e}{m_p} R_\infty. \quad (7.133)$$

The secular equation is thus

$$\begin{vmatrix} H_{11} - ES_{11} & H_{12} - ES_{12} \\ H_{21} - ES_{21} & H_{22} - ES_{22} \end{vmatrix} = \begin{vmatrix} H_{11} - E & H_{12} \\ H_{21} & H_{22} - E \end{vmatrix} = 0 \quad (7.134)$$

$$\begin{aligned} \implies & E^2 - (H_{11} + H_{22})E + H_{11}H_{22} - H_{12}H_{21} = 0 \\ \implies & E = -\left(1 - \frac{m_e}{m_p}\right) \left[\frac{5}{8} \pm \frac{3}{8} \sqrt{1 + 2\left(\frac{64}{81} \frac{m_e}{m_p - m_e}\right)^2} \right] R_\infty. \end{aligned} \quad (7.135)$$

The best (lowest) upper bound is obtained by choosing the “+” sign

$$E_{\min} \approx -(0.999455)R_\infty, \quad (7.136)$$

which is within 3 parts in 10^7 of the true answer.

The ground-state wavefunction corresponding to this minimum is found by solving

$$\begin{pmatrix} H_{11} - E_{\min} & H_{12} \\ H_{21} & H_{22} - E_{\min} \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} = 0. \quad (7.137)$$

This gives

$$\frac{\alpha_1}{\alpha_2} = -\frac{H_{12}}{H_{11} - E_{\min}} \approx -3284.7. \quad (7.138)$$

Requiring also $\alpha_1^2 + \alpha_2^2 = 1$ then fixes

$$\alpha_1 \approx 0.999999954, \quad \alpha_2 \approx -0.000304, \quad (7.139)$$

i.e. with a finite-mass proton, the ground-state wavefunction contains about a 3 parts in 10^4 admixture of the infinite-mass $2s$ wavefunction.

CHAPTER 8

Atomic Structure

8.1 The Non-Relativistic Hydrogen Atom

The quantum mechanics of atomic hydrogen, and hydrogen-like atoms is characterised by a large degeneracy with eigenvalues separating into multiplets of n^2 -fold degeneracy, where n denotes the principal quantum number. However, although the idealised Schrödinger Hamiltonian,

$$\hat{H}_0 = \frac{\hat{\mathbf{p}}^2}{2m} + V(r), \quad V(r) = -\frac{1}{4\pi\epsilon_0} \frac{Ze^2}{r}, \quad (8.1)$$

provides a useful platform from which develop our intuition, there are several important effects which mean that the formulation is a little too naïve. These “corrections”, which derive from several sources, are important as they lead to physical ramifications which extend beyond the realm of atomic physics. Here we outline some of the effects which need to be taken into account even for atomic hydrogen, before moving on to discuss the quantum physics of multielectron atoms. In broad terms, the effects to be considered can be grouped into those caused by the internal properties of the nucleus, and those which derive from relativistic corrections.

To orient our discussion, it will be helpful to summarise some of key aspects of the solutions of the non-relativistic Schrödinger equation, $\hat{H}_0\psi = E\psi$ on which we will draw:

8.1.1 Hydrogen Atom Revisited

As with any centrally symmetric potential, the solutions of the Schrödinger equation take the form $\psi_{n\ell m_\ell}(\mathbf{r}) = R_{n\ell}(r)Y_{\ell,m_\ell}(\theta, \phi)$, where the spherical harmonic functions $Y_{\ell,m_\ell}(\theta, \phi)$ depend only on spherical polar coordinates, and $R_{n\ell}(r)$ represents the radial component of the wavefunction.

$$\psi_{n\ell m_\ell}(\mathbf{r}) = R_{n\ell}(r)Y_{\ell,m_\ell}(\theta, \phi) = |n\ell m_\ell\rangle \begin{cases} n = 1, 2, 3, \dots \\ \ell = 0, 1, 2, \dots, n-1 \\ m_\ell = -\ell, -\ell+1, \dots, \ell \end{cases} \quad (8.2)$$

Solving the radial wave equation introduces a radial quantum number, $n_r \geq 0$. In the case of a Coulomb potential, the energy depends on the principal quantum number $n = n_r + \ell + 1 \geq 1$, and not on n_r and ℓ separately.

For atomic hydrogen ($Z = 1$), the energy levels of the Hamiltonian (8.1) are given by

$$E_n = -\frac{\text{Ry}}{n^2}, \quad \text{Ry} = \left(\frac{e^2}{4\pi\epsilon_0} \right)^2 \frac{m}{2\hbar} = \frac{e^2}{4\pi\epsilon_0} \frac{1}{2a_0} = \frac{1}{2} mc^2 \alpha^2, \quad (8.3)$$

where $a_0 = \frac{4\pi\epsilon_0}{e^2} \frac{\hbar^2}{m}$ is the Bohr radius, $\alpha = \frac{e^2}{4\pi\epsilon_0} \frac{1}{\hbar c}$ denotes the **fine structure constant**, and m represents the reduced mass of the electron and proton. Applied to single electron

ions with higher atomic weight, such as He^+ , Li^{2+} , etc., the Bohr radius is reduced by a factor $1/Z$, where Z denotes the nuclear charge, and the energy is given by

$$E_n = -\frac{Z^2}{n^2} \text{Ry} = -\frac{1}{2n^2} mc^2(Z\alpha)^2. \quad (8.4)$$

Since $n \geq 1$ and $n_r \geq 0$, the allowed combinations of quantum numbers are shown in table 8.1, where we have introduced the conventional notation whereby values of $\ell = 0, 1, 2, 3, 4, \dots$ are represented by letters s, p, d, f, g, \dots respectively.

n	ℓ	Subshell(s)
1	0	1s
2	0, 1	2s, 2p
3	0, 1, 2	3s, 3p, 3d
4	0, 1, 2, 3	4s, 4p, 4d, 4f
n	$0, \dots, (n-1)$	ns, \dots

Table 8.1: Hydrogen subshell designations for principal quantum numbers

Since E_n depends only on n , this implies, for example, an exact degeneracy of the $2s$ and $2p$, and of the $3s$, $3p$ and $3d$ levels.

These results emerge from a treatment of the non-relativistic hydrogen atom which is inherently non-relativistic. In fact, the Hamiltonian (8.1) represents only the leading term in an expansion in $v^2/c^2 \simeq (Z\alpha)^2$ of the full relativistic Hamiltonian (see below). Higher order terms provide relativistic corrections, which impact significantly in atomic and condensed matter physics, and lead to a lifting of the degeneracy. In the following we will discuss and obtain the hierarchy of leading relativistic corrections.¹ This discussion will provide a platform to describe multi-electron atoms.

To see that $v^2/c^2 \simeq (Z\alpha)^2$, we may invoke the **virial theorem** (For a discussion of the form and proof of this theorem, refer to Appendix A.4.). The latter shows that the average kinetic energy is related to the potential energy as $\langle T \rangle = -\frac{1}{2} \langle V \rangle$. Therefore, the average energy is given by $\langle E \rangle = \langle T \rangle + \langle V \rangle = -\langle T \rangle \equiv -\frac{1}{2}mv^2$. We therefore have that $\frac{1}{2}mv^2 = \text{Ry} \equiv \frac{1}{2}mc^2(Z\alpha)^2$ from which follows the relation $v^2/c^2 \simeq (Z\alpha)^2$.

So far, our treatment of the hydrogen atom has been non-relativistic and has neglected the effects of the electron and proton spin, for example. We now aim to move, step by step, towards a more realistic treatment. We first add “switch on” the electron spin “by hand”, remaining non-relativistic, then add the leading-order relativistic effects predicted by the Dirac equation. The relativistic correction to kinetic energy 8.2.1, spin-orbit coupling 8.2.2 and the Darwin term 8.2.3, which lead to *fine structure*, and finally we “switch on” the nuclear (proton) spin by adding the interactions between the proton and electron magnetic moments, leading to *hyperfine structure* 8.2.6.

¹It may seem odd to discuss relativistic corrections without introducing the Dirac equation and the relativistic formulation of quantum mechanics. However, such a discussion would present a lengthy and complex digression. We will therefore follow the normal practice of discussing relativistic corrections as perturbations to the familiar non-relativistic theory.

8.1.2 Including Electron Spin

If we simply “switch on” the electron spin, the Hamiltonian of the Hydrogen atom remains as

$$\hat{H} = \frac{\hat{p}^2}{2m} - V(r), \quad V(r) = -e\phi(r) = -\frac{Ze^2}{4\pi\epsilon_0 r}. \quad (8.5)$$

Since \hat{H} does not contain the electron spin operator $\hat{\mathbf{S}}$, direct products of spacial and spin wavefunctions are energy eigenstates

$$|n\ell m_\ell sm_s\rangle = |n\ell m_\ell\rangle \otimes |sm_s\rangle = \begin{cases} |n\ell m_\ell\rangle \otimes |\uparrow\rangle \\ |n\ell m_\ell\rangle \otimes |\downarrow\rangle \end{cases}, \quad (8.6)$$

since we have $s = 1/2$. The energies depend only on the principal quantum number $n = 1, 2, \dots$,

$$\hat{H} |n\ell m_\ell sm_s\rangle = E_n |n\ell m_\ell sm_s\rangle, \quad E_N = -\frac{Z^2}{2n^2} \text{Ry}. \quad (8.7)$$

The energy levels are *degenerate* with respect to the quantum numbers ℓ , m_ℓ and m_s , with degeneracy $g = 2n^2$.

An alternative set of basis states is provided by the eigenstates of the total angular momentum of the electron, $\hat{\mathbf{J}}$, $|njm_j\ell s\rangle$

$$\left. \begin{array}{l} \hat{\mathbf{J}}^2 |njm_j\ell s\rangle = j(j+1)\hbar^2 |njm_j\ell s\rangle \\ \hat{\mathbf{J}}_z |njm_j\ell s\rangle = m_j\hbar |njm_j\ell s\rangle \end{array} \right\} j = \ell \otimes s = \ell \pm \frac{1}{2} \quad (8.8)$$

Since the energies depend only on n , these are also eigenstates of \hat{H}

$$\hat{H} |njm_j\ell s\rangle = E_n |njm_j\ell s\rangle, \quad E_N = -\frac{Z^2}{2n^2} \text{Ry}. \quad (8.9)$$

These states are equivalently degenerate in j , m_j and ℓ , with degeneracy

$$g = \sum_{\ell=0}^{n-1} \sum_{j=\ell-1/2}^{\ell+1/2} (2j+1) = \sum_{\ell=0}^{n-1} 2(2\ell+1) = n^2 \quad (8.10)$$

These alternative bases, known as *uncoupled* and *coupled* basis respectively, are linear combinations of each other, we can use whichever is more convenient for the problem at hand.

ℓ	m_ℓ	m_s	g
0	0	$\pm 1/2$	$1 \times 2 = 2$
1	$0, \pm 1$	$\pm 1/2$	$3 \times 2 = 6$

Table 8.2: Uncoupled basis states $|n\ell m_\ell\rangle \otimes |sm_s\rangle$ for the $n = 2$ energy level.

In spectroscopy, a *term* symbol is an abbreviated description of the total spin and orbital angular momentum quantum numbers of the electrons in a multi-electron atom. So while the word symbol suggests otherwise, it represents an actual value of a physical quantity. The spectroscopic *term* notation is

$$n^{2S+1}L_j, \quad (8.11)$$

ℓ	j	m_j	g
0	1/2	$\pm 1/2$	$1 \times 2 = 2$
1	1/2	$\pm 1/2$	$1 \times 2 = 2$
1	3/2	$\pm 1/2, \pm 3/2$	$1 \times 4 = 4$

Table 8.3: Coupled basis states $|njm_j\ell s\rangle$ for the $n = 2$ energy level.

n	ℓ	j	Subshell	Terms
1	0	1/2	1s	$1^2S_{1/2}$
2	0	1/2	2s	$2^2S_{1/2}$
	1	1/2, 3/2	2p	$2^2P_{1/2}, 2^2P_{3/2}$
3	0	1/2	3s	$3^2S_{1/2}$
	1	1/2, 3/2	3p	$3^2P_{1/2}, 3^2P_{3/2}$
	2	3/2, 5/2	3d	$3^2D_{3/2}, 3^2D_{5/2}$

Table 8.4: Spectroscopic term notaion.

where for hydrogen, $L \equiv \ell$, $S \equiv s$, and $J \equiv j$.

The coupled states expand in terms of the uncoupled states as

$$|njm_j\ell s\rangle = \sum_{\substack{m_\ell, m_s \\ m_\ell + m_s = m_j}} |n\ell m_\ell s m_s\rangle \underbrace{\langle \ell m_\ell; s m_s | j m_j \rangle}_{\text{Clebsch-Gordan coefficients}}, \quad (8.12)$$

where, for $s = 1/2$, the sum contains at most two non-zero terms

$$m_s = \pm 1/2, \quad m_\ell \pm \frac{1}{2} = m_j. \quad (8.13)$$

For example, in the case of states with $\ell = 1$, $j = 1 \otimes 1/2 = 1/2, 3/2$, and $m_j = \pm 1/2, \pm 3/2$. For $m_J = +3/2$ the only possibility is $j = 3/2$

$$\underbrace{|n, \frac{3}{2}, +\frac{3}{2}, 1, \frac{1}{2}\rangle}_{|njm_j\ell s\rangle} = \underbrace{|n, 1, +1\rangle \otimes |\uparrow\rangle}_{|n\ell m_\ell\rangle \otimes |sm_s\rangle}. \quad (8.14)$$

For $m_j = +1/2$, both $j = 1/2$ and $j = 3/2$ are possible (for any n)

$$j = 3/2 : |n, \frac{3}{2}, +\frac{1}{2}, 1, \frac{1}{2}\rangle = \sqrt{\frac{1}{3}} |n, 1, +1\rangle \otimes |\downarrow\rangle + \sqrt{\frac{2}{3}} |n, 1, 0\rangle \otimes |\uparrow\rangle \quad (8.15)$$

$$j = 1/2 : |n, \frac{1}{2}, +\frac{1}{2}, 1, \frac{1}{2}\rangle = \sqrt{\frac{2}{3}} |n, 1, +1\rangle \otimes |\downarrow\rangle - \sqrt{\frac{1}{3}} |n, 1, 0\rangle \otimes |\uparrow\rangle. \quad (8.16)$$

Explicitly, the wavefunctions above for the $m_j = +1/2$ states are

$$j = 3/2 : |n, \frac{3}{2}, +\frac{1}{2}, 1, \frac{1}{2}\rangle = R_{n1}(r) \left[\sqrt{\frac{1}{3}} Y_{11}(\theta, \phi) |\downarrow\rangle + \sqrt{\frac{2}{3}} Y_{10}(\theta, \phi) |\uparrow\rangle \right] \quad (8.17)$$

$$j = 1/2 : |n, \frac{1}{2}, +\frac{1}{2}, 1, \frac{1}{2}\rangle = R_{n1}(r) \left[\sqrt{\frac{2}{3}} Y_{11}(\theta, \phi) |\downarrow\rangle - \sqrt{\frac{1}{3}} Y_{10}(\theta, \phi) |\uparrow\rangle \right]. \quad (8.18)$$

We can gain further insight by considering the commutation properties of the various angular momentum operators. The operators $\hat{\mathbf{L}}$ and $\hat{\mathbf{S}}$ are independent, and commute with each other

$$[\hat{\mathbf{L}}, \hat{\mathbf{S}}] = 0, \quad [\hat{\mathbf{L}}^2, \hat{\mathbf{S}}^2] = 0. \quad (8.19)$$

The operator product $\hat{\mathbf{L}} \cdot \hat{\mathbf{S}}$ therefore commutes with $\hat{\mathbf{L}}^2$ and $\hat{\mathbf{S}}^2$

$$[\hat{\mathbf{L}} \cdot \hat{\mathbf{S}}, \hat{\mathbf{L}}^2] = \sum_i [\hat{L}_i \hat{S}_i, \hat{L}^2] = 0 \quad (8.20)$$

$$[\hat{\mathbf{L}} \cdot \hat{\mathbf{S}}, \hat{\mathbf{S}}^2] = \sum_i [\hat{L}_i \hat{S}_i, \hat{S}^2] = 0. \quad (8.21)$$

The operator $\hat{\mathbf{J}}^2 = (\hat{\mathbf{L}} + \hat{\mathbf{S}})^2 = \hat{\mathbf{L}}^2 + \hat{\mathbf{S}}^2 + 2\hat{\mathbf{L}} \cdot \hat{\mathbf{S}}$ commutes with $\hat{\mathbf{L}}^2$, $\hat{\mathbf{S}}^2$ and $\hat{\mathbf{L}} \cdot \hat{\mathbf{S}}$

$$[\hat{\mathbf{J}}^2, \hat{\mathbf{L}}^2] = [\hat{\mathbf{J}}^2, \hat{\mathbf{S}}^2] = [\hat{\mathbf{J}}^2, \hat{\mathbf{L}} \cdot \hat{\mathbf{S}}] = 0 \quad (8.22)$$

For spherically symmetric potentials, the Hamiltonian $\hat{H} = \hat{\mathbf{p}}^2/2m - V(r)$ commutes with $\hat{\mathbf{L}}$ (See Eq. (4.44) in Subsection 4.2.3)

$$[\hat{\mathbf{p}}^2, \hat{\mathbf{L}}] = 0, \quad [\hat{V}(r), \hat{\mathbf{L}}] = 0 \quad \implies \quad [\hat{H}, \hat{\mathbf{L}}] = [\hat{H}, \hat{\mathbf{L}}^2] = 0. \quad (8.23)$$

As \hat{H} does not depend on spin, it also commutes with $\hat{\mathbf{S}}$ and, hence, $\hat{\mathbf{J}}$.

Therefore, the Hamiltonian for the Hydrogen atom commutes with all the angular momentum operators: $\hat{\mathbf{L}}$, $\hat{\mathbf{L}}^2$, $\hat{\mathbf{S}}$, $\hat{\mathbf{S}}^2$, $\hat{\mathbf{J}}$, and $\hat{\mathbf{J}}^2$. Thus ℓ , m_ℓ , s , m_s , j , and m_j are all *good quantum numbers*. We thus have two mutually commuting sets of operators, each of which must possess its own set of simultaneous eigenstates

$\{\hat{H}, \hat{\mathbf{L}}^2, \hat{L}_z, \hat{\mathbf{S}}^2, \hat{S}_z\}$	$\{\hat{H}, \hat{\mathbf{J}}^2, \hat{J}_z, \hat{\mathbf{L}}^2, \hat{\mathbf{S}}^2\}$
$ n\ell m_\ell s m_s\rangle$	$ nj m_j \ell s\rangle$
uncoupled	coupled

(8.24)

8.1.3 The Zeeman Effect

We start by exploring how the spectrum of the hydrogen atom is influenced by a uniform magnetic field, \mathbf{B} . There are two main effects:

1. The magnetic field couples to the magnetic moment associated with the *spin* of the electron.² This gives rise to an additional term in the Hamiltonian

$$\hat{H}_{\text{spin}} = -\boldsymbol{\mu}_e \cdot \mathbf{B} = g \frac{e}{2m} \hat{\mathbf{S}} \cdot \mathbf{B} \quad (8.25)$$

where we have used the spin magnetic dipole moment of the electron $\boldsymbol{\mu}_e = -g\mu_B(\mathbf{S}/\hbar)$ where $\mu_B \equiv e\hbar/(2m)$ is the Bohr magneton, and g the electron g -factor which is close to $g = 2$.³

²There is also a coupling to the nuclear magnetic moment. Since $\mu_N/\mu_e \sim 10^{-3}$ this is a much weaker effect. However, it can be important in the hyperfine structure.

³The Dirac equation for the electron leads to the prediction that $g = 2$ exactly. Experimental measurements show $g = 2[1.0011596521884(\pm 43)]$. The small correction is known as the anomalous magnetic dipole moment of the electron. It arises from the electron's interaction with virtual photons in quantum electrodynamics, and can be accurately computed within that theory.

2. The field couples to the orbital motion of the (charged) electron. This is accounted for by replacing the kinetic energy with

$$\frac{1}{2m}\hat{\mathbf{p}}^2 \rightarrow \frac{1}{2m}(\hat{\mathbf{p}} + e\mathbf{A}(\hat{\mathbf{r}}))^2. \quad (8.26)$$

giving an additional change in the Hamiltonian

$$\hat{H}_{\text{orbital}} = \frac{1}{2m} [e(\hat{\mathbf{p}} \cdot \mathbf{A} + \mathbf{A} \cdot \hat{\mathbf{p}}) + e^2|\mathbf{A}|^2] \quad (8.27)$$

Let us explore further the effects of these orbital terms. The term linear in \mathbf{A} is known as the **paramagnetic term**, and the quadratic term the **diamagnetic** contribution. For a constant magnetic field, the vector potential can be as $\mathbf{A} = -\hat{\mathbf{r}} \times \mathbf{B}/2$. In this case, the paramagnetic component takes the form

$$-\frac{ie\hbar}{m}\mathbf{A} \cdot \nabla = -\frac{ie\hbar}{2m}(\hat{\mathbf{r}} \times \nabla) \cdot \mathbf{B} = \frac{e}{2m}\hat{\mathbf{L}} \cdot \mathbf{B}, \quad (8.28)$$

where $\hat{\mathbf{L}}$ denotes the angular momentum operator. Similarly, the diamagnetic term leads to

$$\frac{e^2}{2m}\mathbf{A}^2 = \frac{e^2}{8m}(\hat{\mathbf{r}}^2\mathbf{B}^2 - (\hat{\mathbf{r}} \cdot \mathbf{A})^2) = \frac{e^2B^2}{2m}(\hat{x}^2 + \hat{y}^2), \quad (8.29)$$

where, here, we have chosen the magnetic field to lie along the z -axis.

Before addressing the role of these separate contributions in atomic hydrogen, let us first estimate their relative magnitude. With $\langle \hat{x}^2 + \hat{y}^2 \rangle \simeq a_0^2$, where a_0 denotes the Bohr radius, and $\langle \hat{L}_z \rangle \simeq \hbar$, the ratio of the diamagnetic to paramagnetic terms is

$$\frac{(e^2/8m_e)\langle \hat{x}^2 + \hat{y}^2 \rangle B^2}{(e/2m_e)\langle \hat{L}_z \rangle B} = \frac{e}{4}\frac{a_0^2 B}{\hbar} \simeq 10^{-6} B/\text{T}. \quad (8.30)$$

Therefore, provided the electron is bound in an orbital of size $\sim a_0$, for fields that can be achieved in the laboratory ($B \simeq 1\text{T}$) the diamagnetic term is negligible compared to the paramagnetic term.

Thus, for typical laboratory magnetic fields, the diamagnetic term can be neglected and the net change to the Hamiltonian is

$$\hat{H}_{\text{Zeeman}} \equiv \hat{H}_{\text{spin}}\hat{H}_{\text{orbital}} \simeq \frac{e}{2m}(g\hat{\mathbf{S}} + \hat{\mathbf{L}}) \cdot \mathbf{B} \quad (8.31)$$

The overall Hamiltonian for an electron moving in a Coulomb potential in the presence of a constant magnetic field $\mathbf{B} = Bz$ becomes

$$\hat{H} \simeq \hat{H}_0 + \frac{eB}{2m}(g\hat{S}_z + \hat{L}_z), \quad (8.32)$$

where $\hat{H}_0 = \frac{\hat{\mathbf{p}}^2}{2m} - \frac{e^2}{4\pi\epsilon_0 r}$. Since $[\hat{H}_0, \hat{S}_z] = [\hat{H}_0, \hat{L}_z] = 0$, the eigenstates of the unperturbed hamiltonian, $\psi_{n\ell m_\ell m_s}$, remain eigenstates of \hat{H} and the corresponding energy levels are

$$E_{n\ell m_\ell m_s} \simeq -\frac{\text{Ry}}{n^2} + \hbar\omega_L(gm_s + m_\ell) \quad (8.33)$$

where $\omega_L = \frac{eB}{2m}$ denotes the **Larmor frequency**. From this result, we expect that a constant magnetic field will lead to a splitting of the $2 \times (2\ell + 1)$ -fold degeneracy of the energy levels. Before the knowledge of the existence of electron spin, one would have expected the level to split into $(2\ell + 1)$ levels, due to the orbital Zeeman effect. The fact that experiment showed an additional two-fold splitting, was one of the key insights that led to the identification of electron spin.

8.2 The “Real” Hydrogen Atom

The relativistic corrections (sometimes known as the **fine-structure corrections**) to the spectrum of hydrogen-like atoms derive from three different sources:

- relativistic corrections to the kinetic energy;
- coupling between spin and orbital degrees of freedom;
- a contribution known as the Darwin term.

In the following, we will discuss each of these corrections in turn.

8.2.0.1 The Dirac Equation

The Schrödinger equation is not Lorentz-invariant; it contains a *first-order* time derivative but contains *second-order* spatial derivatives, and therefore it cannot describe *relativistic* particles.

Starting from the relativistic energy equation $E^2 = \hat{\mathbf{p}}^2 c^2 + m^2 c^4$, we could think of defining an appropriate hamiltonian for a relativistic particle

$$\hat{H} = \sqrt{\hat{\mathbf{p}}^2 c^2 + m^2 c^4} = mc^2 \sqrt{1 + \frac{\hat{\mathbf{p}}^2}{m^2 c^2}} = mc^2 + \frac{1}{2} \frac{\hat{\mathbf{p}}^2}{m} - \frac{1}{8} \frac{\hat{\mathbf{p}}^4}{m^3 c^2} + \dots, \quad (8.34)$$

which does not look like something we would want to attempt to solve! Luckily for us, Dirac found a way to write down a relativistic wave equation that is *linear* in the time and space derivatives, the *Dirac equation*. A complete description of the Dirac equation is beyond the scope of this course, here we only give a brief introduction.

What if we could take the square root in Eq. (8.34) and write the Schrödinger equation as

$$i\hbar \frac{\partial}{\partial t} \psi = \hat{H}_{\text{Dirac}} \psi, \quad (8.35)$$

if there exists some

$$\boxed{\hat{H}_{\text{Dirac}} = c\boldsymbol{\alpha} \cdot \hat{\mathbf{p}} + \beta mc^2}, \quad (8.36)$$

where $\boldsymbol{\alpha} = (\alpha_x, \alpha_y, \alpha_z)$ and β are constants.

For this to work, the constants α and β cannot be ordinary complex numbers; they have to be 4×4 complex matrices

$$\alpha_i = \begin{pmatrix} 0 & \sigma_i \\ \sigma_i & 0 \end{pmatrix}, \quad \beta = \begin{pmatrix} \mathbb{I} & 0 \\ 0 & -\mathbb{I} \end{pmatrix}, \quad (8.37)$$

where σ_i are the Pauli matrices and \mathbb{I} is the 2×2 identity matrix.

The wavefunction $\Psi(\mathbf{r}, t)$ must then be a four-component *spinor*

$$\Psi = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{pmatrix}. \quad (8.38)$$

The four degrees of freedom can be interpreted as two *particle*, plus two *antiparticle*, spin states.

The Dirac equation describes *relativistic spin-half particles*. The interpretation of relativistic wave equations such as the Dirac equation is beset by conceptual problems (negative energies and probabilities), these can only be properly resolved by embedding such equations within a quantum field theory (QFT). We then naturally obtain antiparticles as well as particles, and can accommodate the creation and annihilation of particles and antiparticles. To describe atomic electrons, we need only the *low energy limit* of the Dirac equation for a spin-half particle moving in an external EM field, in this limit, these conceptual problems are not an issue.

An external EM field described by potentials $\mathbf{A}(\mathbf{r}, t)$ and $\phi(\mathbf{r}, t)$ can be introduced, as before, via the minimal substitution prescription:

$$\left. \begin{array}{l} \hat{\mathbf{p}} \rightarrow \hat{\mathbf{p}} - q\mathbf{A} \\ E \rightarrow E + q\phi \end{array} \right\} \implies \hat{H} = c\alpha \cdot (\hat{\mathbf{p}} - q\mathbf{A}) + \beta mc^2 + q\phi. \quad (8.39)$$

Keeping only terms up to order v^2/c^2 , the Dirac equations then becomes (After a lengthy calculation, in any gauge)

$$(E' - q\phi)\Psi = \left[\frac{1}{2m}(\hat{\mathbf{p}} - q\mathbf{A})^2 - \frac{q}{m}\hat{\mathbf{S}} \cdot (\nabla \times \mathbf{A}) - \frac{\hat{\mathbf{p}}^2}{8m^3c^2} + \frac{q\hbar^2}{8m^2c^2}(\nabla^2\phi) + \frac{q}{2m^2c^2}\hat{\mathbf{S}} \cdot (\nabla\phi \times \hat{\mathbf{p}}) \right] \Psi, \quad (8.40)$$

where $E' = E - mc^2$ is the relativistic kinetic energy and

$$\Psi \rightarrow \begin{pmatrix} \psi_1 \\ \psi_2 \\ \sim 0 \\ \sim 0 \end{pmatrix}. \quad (8.41)$$

8.2.0.2 The Pauli Equation

The first few terms of Eq. (8.40) reproduce the *Pauli equation*, an equation introduced by Pauli to explain the results of the Stern-Gerlach experiment

$$i\hbar \frac{\partial}{\partial t} \chi = \hat{H}_{\text{Pauli}} \chi, \quad \hat{H}_{\text{Pauli}} = \frac{1}{2m} (\boldsymbol{\sigma} \cdot (\hat{\mathbf{p}} - q\mathbf{A}))^2 + q\phi, \quad (8.42)$$

where the solutions χ are two-component spinors, describing the two degrees of freedom of a spin-half particle.

Using the properties of the Pauli matrices, in particular for vectors \mathbf{a} and \mathbf{b}

$$(\boldsymbol{\sigma} \cdot \mathbf{a})(\boldsymbol{\sigma} \cdot \mathbf{b}) = (\mathbf{a} \cdot \mathbf{b})\mathbb{I}_{2 \times 2} + i\boldsymbol{\sigma} \cdot (\mathbf{a} \times \mathbf{b}), \quad (8.43)$$

we can show that for a free spin-half particle ($A_i = \phi = 0$), the Pauli equation reduces to

$$i\hbar \frac{\partial}{\partial t} \chi = \frac{\hat{\mathbf{p}}^2}{2m} \mathbb{I}_{2 \times 2} \chi. \quad (8.44)$$

More interestingly, expanding Eq. (8.42) we recover

$$\hat{H}_{\text{Pauli}} = \frac{1}{2m} (\hat{\mathbf{p}} - q\mathbf{A})^2 - \frac{q\hbar}{2m} \boldsymbol{\sigma} \cdot (\nabla \times \mathbf{A}) + q\phi, \quad (8.45)$$

where $\mathbf{B} = \nabla \times \mathbf{A}$ is the magnetic field. The first term is the kinetic energy of a particle in an external EM field. An analysis in the symmetric gauge $\mathbf{A} = \frac{1}{2}\mathbf{B} \times \mathbf{r}$ (see Appendix A.7.1) yields

$$\begin{aligned} \frac{1}{2m} (\hat{\mathbf{p}} - q\mathbf{A})^2 &= \frac{1}{2m} \left[-q(\hat{\mathbf{p}} \cdot \mathbf{A} + \mathbf{A} \cdot \hat{\mathbf{p}}) + q^2 |\mathbf{A}|^2 \right] \\ &= \frac{\hat{\mathbf{p}}^2}{2m} - \frac{q}{2m} (\hat{\mathbf{L}} \cdot \mathbf{B}) + \frac{q^2}{8m} [B^2 r^2 - (\mathbf{B} \cdot \hat{\mathbf{r}})], \end{aligned} \quad (8.46)$$

giving rise to the interaction term between the *orbital* magnetic moment and the magnetic field $-\hat{\mu}_L \cdot \mathbf{B} = -(q/2m)\hat{\mathbf{L}} \cdot \mathbf{B}$.

The second term gives rise to the interaction term between the spin and the magnetic field, $\mathbf{B} = \nabla \times \mathbf{A}$

$$-\frac{q\hbar}{2m} \hat{\boldsymbol{\sigma}} \cdot (\nabla \times \mathbf{A}) = -\frac{q}{m} \hat{\mathbf{S}} \cdot \mathbf{B}, \quad (8.47)$$

where the value of the g -factor is predicted to be $g = 2$ for a spin-half particle!

By recovering Pauli’s equation, Dirac now had a *theoretical argument* that implied that spin was somehow the consequence of the marriage of quantum mechanics and relativity.

8.2.0.3 The Dirac equation for the Hydrogen Atom

We can now compute the leading relativistic effects predicted by the Dirac equation for the hydrogen atom, up to terms of order v^2/c^2 . For an atomic electron moving in the electric field of an infinitely heavy nucleus of charge $+Ze$, we have

$$\mathbf{A} = \mathbf{0}, \quad \mathbf{B} = \mathbf{0}, \quad q = -e, \quad \phi = \phi(r) = -\frac{Ze^2}{4\pi\epsilon_0 r}. \quad (8.48)$$

Going back to Eq. (8.40), we find

$$E' \Psi = \left[\underbrace{\frac{\hat{\mathbf{p}}^2}{2m} - e\phi(r)}_{\hat{H}_0} - \underbrace{\frac{\hat{\mathbf{p}}^2}{8m_e^3 c^2}}_{\hat{H}_R} - \underbrace{\frac{e\hbar^2}{8m_e^2 c^2} \nabla^2 \phi}_{\hat{H}_D} - \underbrace{\frac{e}{2m_e^2 c^2} \hat{\mathbf{S}} \cdot (\nabla \phi \times \hat{\mathbf{p}})}_{\hat{H}_{SO}} \right] \Psi. \quad (8.49)$$

We will study the effect of these extra terms using perturbation theory with

$$\hat{H}^{(1)} = \hat{H}_R + \hat{H}_D + \hat{H}_{SO}. \quad (8.50)$$

- \hat{H}_R (Subsection 8.2.1): Going back to our relativistic energy equation,

$$E = \sqrt{p^2 c^2 + m^2 c^4} = mc^2 + \frac{1}{2} \frac{p^2}{m} - \frac{1}{8} \frac{p^4}{m^3 c^2} + \dots \quad (8.51)$$

The *relativistic correction* to the electron energy is thus

$$\boxed{\hat{H}_R = -\frac{\hat{\mathbf{p}}^4}{8m^3 c^2}}. \quad (8.52)$$

- \hat{H}_{SO} (Subsection 8.2.2): For a central potential $\phi = \phi(r)$, we have

$$\begin{aligned} \nabla \phi &= \frac{1}{r} \frac{d\phi}{dr} \hat{\mathbf{r}} \implies \nabla \phi \times \hat{\mathbf{p}} = \frac{1}{r} \frac{d\phi}{dr} \hat{\mathbf{r}} \times \hat{\mathbf{p}} = \frac{1}{r} \frac{d\phi}{dr} \hat{\mathbf{L}} \\ \hat{H}_{SO} &= -\frac{e\hbar^2}{2m^2 c^2} \hat{\mathbf{S}} \cdot (\nabla \phi \times \hat{\mathbf{p}}) = -\frac{e\hbar^2}{2m^2 c^2} \frac{1}{r} \frac{d\phi}{dr} \hat{\mathbf{S}} \cdot \hat{\mathbf{L}}. \end{aligned} \quad (8.53)$$

In particular, for a hydrogen-like atom, we have the *Spin-orbit correction*

$$\boxed{\hat{H}_{SO} = \frac{1}{2m^2 c^2} \frac{Ze^2}{4\pi\epsilon_0} \frac{1}{r^3} \hat{\mathbf{S}} \cdot \hat{\mathbf{L}}}. \quad (8.54)$$

The physical origin of the spin-orbit term can be understood using a semi-classical approach. As the orbiting electron moves through the electric field $\mathbf{E} = -\nabla\phi$ of the nucleus, it experiences (after a Lorentz Transformation) a *magnetic field*. The interaction $-\hat{\mu} \cdot \mathbf{B}$ of the electron magnetic moment $\hat{\mu}$ with this apparent magnetic field \mathbf{B} produces the spin-orbit interaction.

- \hat{H}_D (Subsection 8.2.3): Has no classical interpretation, it is purely quantum in origin, the *Darwin Correction*

$$\boxed{\hat{H}_D = -\frac{e\hbar^2}{8m_e^2 c^2} \nabla^2 \phi}. \quad (8.55)$$

What is the size of these perturbations? E.g. for the relativistic correction,

$$\frac{\langle \hat{H}_R \rangle}{\langle \hat{H}_0 \rangle} \sim \frac{p^4 / 8m^3 c^2}{p^2 / 2m} \sim \frac{p^2}{m^2 c^2} \sim \frac{v^2}{c^2} \sim (Z\alpha)^2 \ll 1. \quad (8.56)$$

Expect resulting energy corrections to be small, and *first-order perturbation theory* to be reliable. Are there degeneracies? The unperturbed energy levels are degenerate for all n , with $g = 2n^2$, thus we must use degenerate perturbation theory.

The energy corrections will break some of the initial degeneracy, causing energy levels to split into two or more closely spaced levels, giving rise to the so-called *fine structure* of the Hydrogen atom.

8.2.1 Relativistic Correction to the Kinetic Energy

Previously, we have taken the kinetic energy to have the familiar non-relativistic form, $\frac{\hat{\mathbf{p}}^2}{2m}$. However, from the expression for the relativistic energy-momentum invariant, $p_\mu p^\mu = (mc)^2$, we can already anticipate that the leading correction to the non-relativistic Hamiltonian appears at order \mathbf{p}^4 ,

$$E = \sqrt{\hat{\mathbf{p}}^2 c^2 + m^2 c^4} = mc^2 + \frac{\hat{\mathbf{p}}^2}{2m} - \frac{1}{8} \frac{(\hat{\mathbf{p}}^2)^2}{m^3 c^2} + \dots \quad (8.57)$$

As a result, we can infer the following perturbation to the kinetic energy of the electron,

$$\hat{H}_1 = -\frac{1}{8} \frac{(\hat{\mathbf{p}}^2)^2}{m^3 c^2}.$$

(8.58)

When compared with the non-relativistic kinetic energy, $\hat{\mathbf{p}}^2/2m$, one can see that the perturbation is smaller by a factor of $p^2/m^2 c^2 = v^2/c^2 \simeq (Z\alpha)^2$, i.e. \hat{H}_1 is only a small perturbation for small atomic number, $Z \ll 1/\alpha \simeq 137$. We can therefore make use of a perturbative analysis to estimate the scale of the correction.

In principle, the large-scale degeneracy of the hydrogen atom would demand an analysis based on the degenerate perturbation theory. However, fortunately, since the off-diagonal matrix elements vanish,⁴

$$\langle n\ell m_\ell | \hat{H}_1 | n\ell' m'_\ell \rangle = 0, \quad \text{for } \ell \neq \ell' \text{ or } m_\ell \neq m'_\ell, \quad (8.59)$$

degenerate states are uncoupled and such an approach is unnecessary.

The operator $\hat{\mathbf{p}}^4 = (-i\hbar\nabla)^4$ in \hat{H}_1 is not easy to work with directly, the trick is to express \hat{H}_1 in terms of $\hat{H}_0 = \hat{\mathbf{p}}^2/2m + \hat{V}$. Then making use of the identity, $\hat{H}_1 = -\frac{1}{2mc^2}[\hat{H}_0 - V(r)]^2$,

$$\begin{aligned} \hat{H}_1 &= -\frac{1}{2mc^2} \left(\frac{\hat{\mathbf{p}}}{2m} \right)^2 = -\frac{1}{2mc^2} (\hat{H}_0 - \hat{V})^2 \\ &= -\frac{1}{2mc^2} (\hat{H}_0^2 - \hat{H}_0 \hat{V} - \hat{V} \hat{H}_0 + \hat{V}^2). \end{aligned} \quad (8.60)$$

Let us choose the uncoupled basis for the unperturbed states, $|n\ell m_\ell\rangle \otimes |sm_s\rangle$. Let us now check if \hat{H}_1 is diagonal in the degenerate subspace (states with given n). Since \hat{H}_1 is independent of spin, we have

$$\langle n\ell' m'_\ell sm_s | \hat{H}_1 | n\ell m_\ell sm_s \rangle = \langle n\ell' m'_\ell | \hat{H}_1 | n\ell m_\ell \rangle \delta_{m_s, m'_s}. \quad (8.61)$$

The matrix elements of the \hat{H}_0^2 term are easy to calculate as

$$\langle n\ell' m'_\ell | \hat{H}_0^2 | n\ell m_\ell \rangle = E_n^2 \delta_{\ell\ell'} \delta_{m_\ell, m'_\ell}. \quad (8.62)$$

⁴Proof: Since $[\hat{H}_1, \hat{\mathbf{L}}^2] = 0$, $\hbar[\ell'(\ell' + 1) + \ell(\ell + 1)] \langle n\ell m_\ell | \hat{H}_1 | n\ell' m'_\ell \rangle = 0$. Similarly, since $[\hat{H}_1, \hat{L}_z] = 0$, $\hbar(m'_\ell - m_\ell) \langle n\ell m_\ell | \hat{H}_1 | n\ell' m'_\ell \rangle = 0$.

Given that $|n\ell m_\ell\rangle = R_{n\ell}(r)Y_{\ell m}(\theta, \phi)$, matrix elements of any radial functions $f(r)$ evaluate as

$$\begin{aligned}\langle n'\ell'm'_\ell | f(r) | n\ell m_\ell \rangle &= \int_0^\infty R_{n'\ell'}(r)f(r)R_{n\ell}(r)r^2 dr \int Y_{\ell'm'_\ell}^* Y_{\ell m_\ell} d\Omega \\ &= \left[\int_0^\infty r^2 f(r) R_{n\ell}(r)^2 dr \right] \delta_{\ell\ell'} \delta_{m_\ell, m'_\ell} \\ &= \langle f(r) \rangle_{n\ell} \delta_{\ell\ell'} \delta_{m_\ell, m'_\ell}.\end{aligned}\quad (8.63)$$

Thus, in this basis the matrix elements of $f(r)$ are diagonal in n , ℓ and m_ℓ (and s , m_s), and the eigenvalues depend only on n and ℓ . This is a direct consequence of the *rotational symmetry* of the full Hamiltonian.

Applying this to $f(r) = \hat{V}(r)$, $\hat{V}(r)^2$, we can show that, for the uncoupled basis, \hat{H}_1 is diagonal in the degenerate subspace (of states with given n)

$$\langle n\ell'm'_\ell s m'_s | \hat{H}_1 | n\ell m_\ell s m_s \rangle = \langle \hat{H}_1 \rangle_{n\ell} \delta_{\ell\ell'} \delta_{m_\ell m'_\ell} \delta_{m_s m'_s}, \quad (8.64)$$

hence this is a *good basis* for perturbation theory.

The first-order energy corrections are given by the diagonal elements

$$\begin{aligned}(\Delta E_n)_R &= -\frac{1}{2mc^2} \left\langle \hat{H}_0^2 - \hat{H}_0 \hat{V} - \hat{V} \hat{H}_0 + \hat{V}^2 \right\rangle_{n\ell} \\ &= -\frac{1}{2mc^2} \left(E_n^2 - 2E_n \langle \hat{V} \rangle_{n\ell} + \langle \hat{V}^2 \rangle_{n\ell} \right) \\ &= -\frac{1}{2mc^2} \left(E_n^2 - 2E_n \left(\frac{Ze^2}{4\pi\epsilon_0} \right) \left\langle \frac{1}{r} \right\rangle_{n\ell} + \left(\frac{Ze^2}{4\pi\epsilon_0} \right)^2 \left\langle \frac{1}{r^2} \right\rangle_{n\ell} \right).\end{aligned}\quad (8.65)$$

The expectation values of $1/r$ and $1/r^2$ can be calculated to be

$$\left\langle \frac{1}{r} \right\rangle_{n\ell} = \frac{Z}{a_0 n^2}, \quad \left\langle \frac{1}{r^2} \right\rangle_{n\ell} = \frac{Z^2}{a_0^3 n^4 (\ell + 1/2)} \quad \text{and} \quad E_n = -\frac{Z^2}{2n^2} \frac{e^2}{4\pi\epsilon_0 a_0}. \quad (8.66)$$

Putting everything together,

$$(\Delta E_n)_R = -\frac{1}{2mc^2} \left(\frac{e^2}{4\pi\epsilon_0 a_0} \right)^2 \frac{Z^4}{n^4} \left(\frac{1}{4} + \frac{1}{2} + \frac{n}{(\ell + 1/2)} \right). \quad (8.67)$$

Which, in terms of the fine structure constant, α , becomes

$$(\Delta E_n)_R = -\left(\frac{Z}{n} \right)^4 \left(\frac{3}{4} - \frac{n}{\ell + 1/2} \right) \alpha^2 \text{Ry.}$$

(8.68)

Thus, for each energy level n , the perturbation \hat{H}_R lifts the degeneracy between states of different ℓ , e.g. the $n = 3$ level splits into three separate levels ($\ell = 0, 1, 2$,).

Note that given (ΔE_n) depends only on n and ℓ , we could also have used the coupled basis $|njm_j\ell s\rangle$ to calculate the matrix elements of \hat{H}_R

$$\langle n\ell m_\ell s m_s | \hat{H}_R | n\ell m_\ell s m_s \rangle = \langle \hat{H}_R \rangle_{n\ell} \quad (\text{independent of } m_\ell, m_s) \quad (8.69)$$

$$= \langle njm_j\ell s | \underbrace{\hat{H}_R}_{\substack{\text{linear combinations of} \\ |n\ell m_\ell s m_s\rangle \text{with different} \\ m_\ell \text{ and } m_s}} | njm_j\ell s \rangle \quad (8.70)$$

The energy splitting is small, the scale being set by

$$\alpha^2 \text{Ry} \approx (13.6 \text{ eV})/(137)^2 \approx 7.2 \times 10^{-4} \text{ eV}, \quad (8.71)$$

i.e. a factor α^2 smaller than the Bohr energy scale. However, as we will see, this conclusion is a little hasty. We need to gather all terms of the same order of perturbation theory before we can reach any definite conclusions. We now turn to the second important class of corrections.

8.2.2 Spin-Orbit Coupling

As well as revealing the existence of an internal spin degree of freedom, Dirac’s relativistic formulation of quantum mechanics shows that there is a further relativistic correction to the Schrödinger operator which involves a coupling between the spin and orbital degrees of freedom. For a general potential $V(r)$, this spin-orbit coupling takes the form,

$$\hat{H}_{\text{SO}} = \frac{1}{2m^2c^2} \frac{1}{r} (\partial_r V(r)) \hat{\mathbf{S}} \cdot \hat{\mathbf{L}}. \quad (8.72)$$

For a hydrogen-like atom, $V(r) = -\frac{1}{4\pi\epsilon_0} \frac{Ze^2}{r}$, and

$$\hat{H}_{\text{SO}} = \frac{1}{2m^2c^2} \frac{Ze^2}{4\pi\epsilon_0} \frac{1}{r^3} \hat{\mathbf{S}} \cdot \hat{\mathbf{L}}. \quad (8.73)$$

It will also be convenient for later calculations to define the shorthand

$$\hat{H}_{\text{SO}} = \xi(r) \hat{\mathbf{L}} \cdot \hat{\mathbf{S}}, \quad \xi(r) \equiv \frac{1}{2m^2c^2} \frac{Ze^2}{4\pi\epsilon_0} \frac{1}{r^3} \propto \frac{1}{r^3}. \quad (8.74)$$

Physically, the origin of the spin-orbit interaction can be understood from the following considerations. As the electron is moving through the electric field of the nucleus then, in its rest frame, it will experience this as a magnetic field. There will be an additional energy term in the Hamiltonian associated with the orientation of the spin magnetic moment with respect to this field. We can make an estimate of the spin-orbit interaction energy as follows: If we have a central field determined by an electrostatic potential $\phi(r)$, the corresponding electric field is given by $\mathbf{E} = -\nabla\phi(r) = -\hat{\mathbf{e}}_r(\partial_r\phi) = +\hat{\mathbf{e}}_r(\partial_r V)/e$ where in the last step we introduce the potential energy of the electron $V = -e\phi$. For an electron moving at velocity v , this translates to an effective magnetic field $\mathbf{B} = -\frac{1}{c^2}\mathbf{v} \times \mathbf{E}$. The magnetic moment of the electron associated with its spin is equal to $\boldsymbol{\mu}_s = g_s \frac{-e}{2m} \hat{\mathbf{S}} \equiv -\frac{e}{m} \hat{\mathbf{S}}$, and thus the interaction energy is given by, and thus the interaction energy is given by

$$-\boldsymbol{\mu} \cdot \mathbf{B} = -\frac{e}{mc^2} \mathbf{S} \cdot (\mathbf{v} \times \mathbf{E}) = -\frac{1}{(mc)^2} \mathbf{S} \cdot (\mathbf{p} \times \mathbf{e}_r(\partial_r V)) = +\frac{1}{(mc)^2} \frac{1}{r} (\partial_r V) \mathbf{L} \cdot \mathbf{S}, \quad (8.75)$$

where we have used the relation $\mathbf{p} \times \mathbf{e}_r = \mathbf{p} \times \frac{\mathbf{r}}{r} = -\frac{\mathbf{L}}{r}$. In fact this isn’t quite correct; there is a relativistic effect connected with the precession of axes under rotation, called **Thomas precession** which multiplies the formula by a further factor of 1/2.

Once again, we can estimate the effect of spin-orbit coupling by treating \hat{H}_{SO} as a perturbation. In the absence of spin-orbit interaction, one may express the eigenstates

of hydrogen-like atoms in the basis states of the mutually commuting operators, \hat{H}_0 , $\hat{\mathbf{L}}^2$, \hat{L}_z , $\hat{\mathbf{S}}^2$, and \hat{S}_z . However, in the presence of spin-orbit coupling, the total Hamiltonian no longer commutes with \hat{L}_z or \hat{S}_z . It is therefore helpful to make use of the degeneracy of the unperturbed Hamiltonian to switch to a new basis in which the angular momentum components of the perturbed system are diagonal. This can be achieved by turning to the basis of eigenstates of the operators, \hat{H}_0 , $\hat{\mathbf{J}}^2$, \hat{J}_z , $\hat{\mathbf{L}}^2$, and $\hat{\mathbf{L}}^2$, where $\hat{\mathbf{J}} = \hat{\mathbf{L}} + \hat{\mathbf{S}}$ denotes the total angular momentum. (For a discussion of the form of these basis states, we refer back to Chapters 4 and 5.)

The trick is to notice that the operator product $\hat{\mathbf{L}} \cdot \hat{\mathbf{S}}$ can be written entirely in terms of squares of angular momentum operators. Making use of the relation, $\hat{\mathbf{J}}^2 = \hat{\mathbf{L}}^2 + \hat{\mathbf{S}}^2 + 2\hat{\mathbf{L}} \cdot \hat{\mathbf{S}}$, in this basis, it follows that,

$$\hat{\mathbf{L}} \cdot \hat{\mathbf{S}} = \frac{1}{2}(\hat{\mathbf{J}}^2 - \hat{\mathbf{L}}^2 - \hat{\mathbf{S}}^2). \quad (8.76)$$

The coupled basis states $|njm_j\ell s\rangle$ are simultaneous eigenstates of the operators $\hat{\mathbf{J}}^2$, $\hat{\mathbf{L}}^2$, and $\hat{\mathbf{S}}$, and therefore also of $\hat{\mathbf{L}} \cdot \hat{\mathbf{S}}$

$$\hat{\mathbf{L}} \cdot \hat{\mathbf{S}} |njm_j\ell s\rangle = \frac{1}{2}\hbar^2(j(j+1) - \ell(\ell+1) - s(s+1)) |njm_j\ell s\rangle. \quad (8.77)$$

Therefore, in the coupled basis, the matrix elements of \hat{H}_{SO} are given by

$$\langle nj'm'_j\ell' s | \xi(r) \hat{\mathbf{L}} \cdot \hat{\mathbf{S}} | njm_j\ell s \rangle = \frac{1}{2}\hbar^2[j(j+1) - \ell(\ell+1) - s(s+1)] \langle nj'm'_j\ell' s | \xi(r) | njm_j\ell s \rangle. \quad (8.78)$$

As for the radial dependence of the perturbation, once again, the off-diagonal matrix elements vanish since $\xi(r)$ does not depend on angle or spin variables, circumventing the need to invoke degenerate perturbation theory, similarly to Eq. (8.63) we have

$$\langle nj'm'_j\ell' s | \xi(r) | njm_j\ell s \rangle = \langle \xi(r) \rangle_{n\ell} \delta_{jj'} \delta_{m_j m'_j} \delta_{\ell\ell'}, \quad (8.79)$$

where the expectation values of $\xi(r)$ are given by

$$\langle \xi(r) \rangle_{n\ell} = \int_0^\infty r^2 \xi(r) R_{n\ell}(r)^2 dr. \quad (8.80)$$

The coupled basis is in this case the *good basis* for perturbation theory.

The first-order energy corrections due to \hat{H}_{SO} are given by the diagonal matrix elements

$$(\Delta E_n)_{SO} = \frac{1}{2}\hbar^2[j(j+1) - \ell(\ell+1) - s(s+1)] \langle \xi(r) \rangle_{n\ell}. \quad (8.81)$$

For an S -state ($\ell = 0$, $s = 1/2$, $j = 1/2$), note that, for $\ell = 0$, there is no orbital angular momentum with which to couple! The factor on the right hand side vanishes,

$$(\Delta E_n)_{SO} = 0 \quad (\ell = 0). \quad (8.82)$$

For any cases with $\ell > 0$, the factor we need to evaluate is ($\xi(r) \sim 1/r^3$)

$$\left\langle \frac{1}{r^3} \right\rangle_{n\ell} = \frac{Z^3}{a_0^3 n^3} \frac{1}{\ell(\ell+1/2)(\ell+1)}. \quad (8.83)$$

And hence,

$$(\Delta E_n)_{\text{SO}} = \frac{1}{2m^2c^2} \frac{Ze^2}{4\pi\epsilon_0} \frac{Z^3}{a_0^3 n^3} \left[\frac{j(j+1) - \ell(\ell+1) - s(s+1)}{\ell(\ell+1/2)(\ell+1)} \right] \quad (8.84)$$

Since $s = 1/2$, there are only two possibilities for j

$$j(j+1) - \ell(\ell+1) - s(s+1) = \begin{cases} \ell & (j = \ell + 1/2) \\ -(\ell+1) & (j = \ell - 1/2) \end{cases}. \quad (8.85)$$

The spin-orbit correction for $\ell > 0$ can thus be written more compactly as

$$(\Delta E_n)_{\text{SO}} = \pm \frac{1}{2} \left(\frac{Z}{n} \right)^4 \frac{n}{j+1/2} \frac{1}{\ell+1/2} \alpha^2 \text{Ry},$$

(8.86)

for $j = \ell \pm 1/2$. The spin-orbit energy correction depends on n , j , and ℓ , thus for each energy level n , \hat{H}_{SO} lifts the degeneracy with respect to both j and ℓ . E.g. the $n = 3$ level would split in five separate levels: $(\ell, j) = (0, 1/2), (1, 1/2), (1, 3/2), (2, 3/2), (2, 5/2)$.

Differently from the relativistic correction, the spin-orbit correction could not have been calculated using the uncoupled basis $|n\ell m_\ell s m_s\rangle$. This is because $\hat{H}_{\text{SO}} = \xi(r)\hat{\mathbf{L}} \cdot \hat{\mathbf{S}}$ does *not* commute with all operators in the uncoupled set $\{\hat{\mathbf{L}}^2, \hat{L}_z, \hat{\mathbf{S}}^2, \hat{S}_z\}$, in particular

$$[\hat{\mathbf{L}} \cdot \hat{\mathbf{S}}, \hat{L}_z] \neq 0 \quad \text{and} \quad [\hat{\mathbf{L}} \cdot \hat{\mathbf{S}}, \hat{S}_z] \neq 0. \quad (8.87)$$

Hence, the uncoupled basis is *not a good basis* for perturbation theory with \hat{H}_{SO} .

8.2.2.1 The Zeeman Effect Revisited

Our theory of the Zeeman effect in Section 8.1.3 was based on the effect of a uniform magnetic field as a perturbation to the *non-relativistic* Hamiltonian \hat{H}_0 . How are these conclusions affected by the additional physics from spin-orbit coupling?

We showed that the effect of a magnetic field $\mathbf{B} = Bz$ on a one-electron atom is to produce an additional term

$$\hat{H}_{\text{Zeeman}} = \frac{eB}{2m} (\hat{L}_z + g\hat{S}_z). \quad (8.88)$$

This denotes the Zeeman energy associated with the coupling of the spin and orbital angular momentum degrees of freedom to the magnetic field. Depending on the scale of the magnetic field, the Zeeman term \hat{H}_{Zeeman} or the spin-orbit term \hat{H}_{SO} may dominate the spectrum of the atom.

We have just seen that, to leading order, the relativistic corrections lead to a fine-structure energy shift of

$$\Delta E_{n,j}^{\text{rel.}} = \frac{1}{2} mc^2 \left(\frac{Z\alpha}{n} \right)^4 \left(\frac{3}{4} - \frac{n}{j+1/2} \right), \quad (8.89)$$

for states $|n, j = \ell \pm 1/2, m_j, \ell\rangle$. For weak magnetic fields, we can treat the Zeeman energy as a perturbation to this spectrum. Here, although states with common j values are degenerate, the two spatial wavefunctions have different parities, since they have orbital angular momenta ($\ell = j - 1/2$ or $\ell = j + 1/2$) that differ by 1. (For example, for a given shell, with $n \geq 1$, a $j = 1/2$ state can be formed either from the $\ell = 0$ orbital or from the $\ell = 1$ orbital.) As a consequence of the differing parities of the spatial wavefunctions, the off-diagonal matrix element of \hat{H}_{Zeeman} coupling these states vanishes. We may therefore avoid using degenerate perturbation theory. Making use of the relation,

$$\langle n, j = \ell \pm 1/2, m_j, \ell | \hat{S}_z | n, j = \ell \pm 1/2, m_j, \ell \rangle = \pm \frac{\hbar m_j}{2\ell + 1}, \quad (8.90)$$

we obtain the following expression for the first order energy shift,

$$\Delta E_{j=\ell \pm 1/2, m_j, \ell}^{\text{Zeeman}} = \mu_B B m_j \left(1 \pm \frac{1}{2\ell + 1} \right), \quad (8.91)$$

where μ_B denotes the Bohr magneton. Therefore, we see that all degenerate levels are split due to the magnetic field. In contrast to the “normal” Zeeman effect, the magnitude of the splitting depends on ℓ .

If the **field is strong**, the Zeeman energy becomes large in comparison with the spin-orbit contribution. In this case, we must work with the basis states $|n, \ell, m_\ell, m_s\rangle = |n, \ell, m_\ell\rangle \otimes |m_s\rangle$ in which both \hat{H}_0 and \hat{H}_{Zeeman} are diagonal. Within first order of perturbation theory, one then finds that

$$\Delta E_{n, \ell, m_\ell, m_s} = \mu_B (m_\ell + m_s) + \frac{1}{2} \left(\frac{Z\alpha}{n} \right)^4 \left(\frac{3}{4} - \frac{n}{\ell + 1/2} - \frac{nm_\ell m_s}{\ell(\ell + 1/2)(\ell + 1)} \right), \quad (8.92)$$

the first term arising from the Zeeman energy and the remaining terms from $\hat{H}^{\text{rel.}}$. At intermediate values of the field, we have to apply degenerate perturbation theory to the states involving the linear combination of $|n, j = \ell \pm 1/2, m_j, \ell\rangle$.

8.2.3 Darwin Term

The final contribution to the Hamiltonian from relativistic effects is known as the Darwin term and arises from the “**Zitterbewegung**” of the electron – trembling motion – which smears the effective potential felt by the electron. Such effects lead to a perturbation of the form,

$$\hat{H}_{\text{D}} = \frac{\hbar^2}{8m^2c^2} \nabla^2 \phi, \quad \phi = \frac{Ze}{4\pi\epsilon_0 r}. \quad (8.93)$$

In this case, $\nabla^2 \phi$ is a function of r only and the matrix elements of \hat{H}_{D} are diagonal in both bases. For $1/r$ potential, $\nabla^2(1/r)$ vanishes everywhere except at the origin, where it diverges

$$\nabla^2(1/r) = -4\pi\delta^{(3)}(\mathbf{r}). \quad (8.94)$$

Therefore, the expectation value of \hat{H}_{D} becomes

$$\langle \psi | \hat{H}_{\text{D}} | \psi \rangle \propto \langle \psi | \delta^{(3)} | \psi \rangle = \int \psi(\mathbf{r})^* \psi(\mathbf{r}) \delta^{(3)}(\mathbf{r}) d^3\mathbf{r} = |\psi(0)|^2, \quad (8.95)$$

where $\psi(0)$ is the value of the wavefunction at the origin.

Since the perturbation acts only at the origin, it affects only states with $\ell = 0$, $|n00\rangle$. As a result, one finds that

$$|\psi_{n00}(0)|^2 = \frac{Z^3}{\pi a_0^3 n^3}. \quad (8.96)$$

Hence the energy shift due to the Darwin term is non-zero only for $\ell = 0$

$$(\Delta E_n)_D = \begin{cases} \frac{Z^4}{n^3} \alpha^2 \text{Ry} & \text{for } \ell = 0 \\ 0 & \text{for } \ell > 0 \end{cases}.$$

(8.97)

This result is valid for *both* the uncoupled and coupled bases.

8.2.4 Hydrogen Atom Fine Structure

The energy corrections are

$$(\Delta E_n)_R = -\left(\frac{Z}{n}\right)^4 \left(\frac{3}{4} - \frac{n}{\ell + 1/2}\right) \alpha^2 \text{Ry} \quad (8.98)$$

$$(\Delta E_n)_{SO} = \begin{cases} 0 & \text{for } \ell = 0 \\ \pm \frac{1}{2} \left(\frac{Z}{n}\right)^4 \frac{n}{j + 1/2} \frac{1}{\ell + 1/2} \alpha^2 \text{Ry} & \text{for } \ell > 0 \end{cases} \quad (8.99)$$

$$(\Delta E_n)_D = \begin{cases} \frac{Z^4}{n^3} \alpha^2 \text{Ry} & \text{for } \ell = 0 \\ 0 & \text{for } \ell > 0 \end{cases} \quad (8.100)$$

For all three perturbations, the scale of the energy corrections is set by

$$\alpha^2 \text{Ry} \approx 7.2 \times 10^{-4} \text{ eV} \approx 175 \text{ GHz.} \quad (8.101)$$

Intriguingly, the $\ell \rightarrow 0$ limit of the $(\Delta E_n)_{SO}$ formula for $\ell > 0$ is formally identical to that which would be obtained from $(\Delta E_n)_D$ at $\ell = 0$. At first sight, the total energy correction, $(\Delta E_n)_{FS}$, seems to depend on all three quantum numbers, n , j , and ℓ , but in fact it *does not depend on ℓ !*

As a result, combining all three contributions, the total energy shift is given simply by

$$(\Delta E_n)_{FS} = \left(\frac{Z}{n}\right)^4 \left(\frac{3}{4} - \frac{n}{j + 1/2}\right) \alpha^2 \text{Ry.}$$

(8.102)

a result that is independent of ℓ and m_j . For all n , the net effect of the correction is to *lower the eigenstate energies*. The n^{th} energy level splits into n separate energy levels, one for each j . E.g. the $n = 3$ level splits into three separate levels ($j = 1/2, 3/2, 5/2$). States with the same j but different ℓ remain degenerate in energy, e.g. the $n = 3$ levels with $j = 1/2$ but $\ell = 0, 1$ remain degenerate.

To discuss the predicted energy shifts for particular states, it is helpful to introduce some nomenclature from atomic physics. For a state with principal quantum number n , total spin s , orbital angular momentum ℓ , and total angular momentum j , one may use

spectroscopic notation (recall from Subsection 8.1.2) $n^{2s+1}L_j$ to define the state. For a hydrogen-like atom, with just a single electron, $2s + 1 = 2$. In this case, the factor $2s + 1$ is often just dropped for brevity.

If we apply our perturbative expression for the relativistic corrections (8.102), how do we expect the levels to shift for hydrogen-like atoms? As we have seen, for the non-relativistic Hamiltonian, each state of given n exhibits a $2n^2$ -fold degeneracy. For a given multiplet specified by n , the relativistic corrections depend only on j and n . For $n = 1$, we have $\ell = 0$ and $j = 1/2$: Both $1S_{1/2}$ states, with $m_j = 1/2$ and $-1/2$, experience a negative energy shift by an amount $\Delta E_{1,1/2,m_j,0} = -\frac{1}{4}Z^4\alpha^2\text{Ry}$. For $n = 2$, ℓ can take the values of 0 or 1. With $j = 1/2$, both the former $2S_{1/2}$ state, and the latter $2P_{1/2}$ states share the same negative shift in energy, $\Delta E_{2,1/2,m_j,0} = \Delta E_{2,1/2,m_j,1} = -\frac{5}{64}Z^4\alpha^2\text{Ry}$, while the $2P_{3/2}$ experiences a shift of $\Delta E_{2,3/2,m_j,1} = -\frac{1}{64}Z^4\alpha^2\text{Ry}$. Finally, for $n = 3$, ℓ can take values of 0, 1 or 2. Here, the pairs of states $3S_{1/2}$ and $3P_{1/2}$, and $3P_{3/2}$ and $3D_{3/2}$ each remain degenerate while the state $3D_{5/2}$ is unique. These predicted shifts are summarised in Figure 8.1.

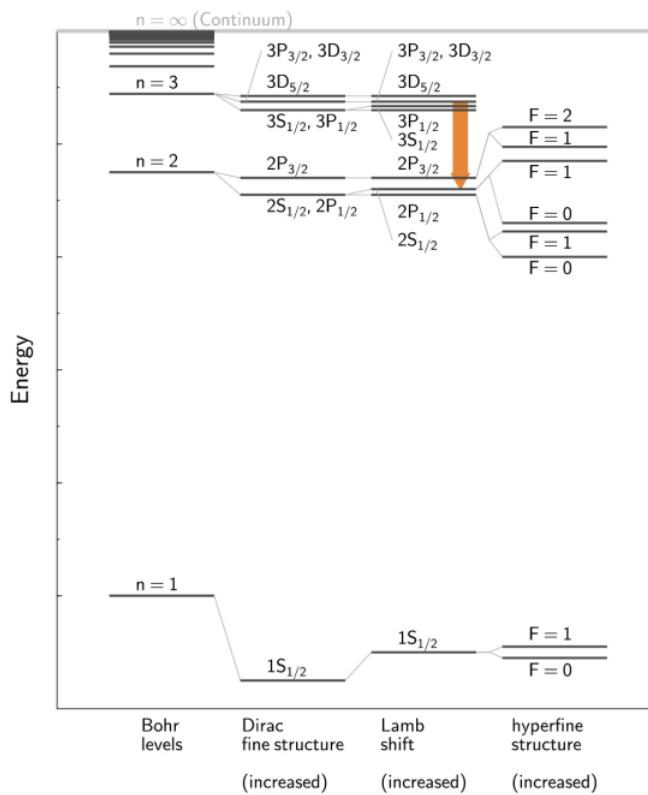


Fig. 8.1: Figure showing the hierarchy of energy shifts of the spectra of hydrogen-like atoms as a result of relativistic corrections. The first column shows the energy spectrum predicted by the (non-relativistic) Bohr theory. The second column shows the predicted energy shifts from relativistic corrections arising from the Dirac theory. The third column includes corrections due quantum electrodynamics and the fourth column includes terms for coupling to the nuclear spin degrees of freedom. The $H - \alpha$ line, particularly important in the astronomy, corresponds to the transition between the levels with $n = 2$ and $n = 3$. Each hyperfine level has degeneracy $g = 2F + 1$. For an *isolated* atom, this degeneracy (w.r.t. m_F) cannot be broken due to rotational symmetry. For an isolated atom, the energy levels cannot be split further. The degeneracy can only be lifted by putting the atom into an external \mathbf{E} or \mathbf{B} field, for example, thereby breaking the rotational symmetry.

This completes our discussion of the relativistic corrections which develop from the treatment of the Dirac theory for the hydrogen atom. However, this does not complete our description of the “real” hydrogen atom. Indeed, there are further corrections which derive from quantum electrodynamics and nuclear effects which we now turn to address.

n	ℓ	j	Term	units of $(\alpha^2 R_\infty / 4n^3)$			GHz	
				$(\Delta E)_R$	$(\Delta E)_{SO}$	$(\Delta E)_D$	$(\Delta E)_{FS}$	
1	0	1/2	1S _{1/2}	-5	0	+4	-1	-43.7
	0	1/2	2S _{1/2}	-13/2	0	+4	-5/2	-13.7
2	1	1/2	2P _{1/2}	-7/6	-4/3	0	-5/2	-13.7
	1	3/2	2P _{3/2}	-7/6	+2/3	0	-1/2	-2.7
0	1/2	3S _{1/2}	-7	0	+4	-3	-4.9	
	1	1/2	3P _{1/2}	-5/3	-4/3	0	-3	-4.9
3	1	3/2	3P _{3/2}	-5/3	+2/3	0	-1	-1.6
	2	3/2	3D _{3/2}	-1	+4/15	0	-1	-1.6
2	5/2	3D _{5/2}	-1	+1/3	0	-1/3	-0.3	

Table 8.5: The total Hydrogen fine structure corrections, calculated as $(\Delta E)_{FS} = (\Delta E)_R + (\Delta E)_{SO} + (\Delta E)_D$.

8.2.5 Lamb Shift

According to the perturbation theory above, the relativistic corrections which follow from the Dirac theory for hydrogen leave states with the same j but different ℓ degenerate, e.g. the 2S_{1/2} and 2P_{1/2} states. However, in 1947, a careful experimental study by Willis Lamb and Robert Rutherford discovered that this was *not* in fact the case: 2P_{1/2} state is slightly lower in energy than the 2S_{1/2} state resulting in a small shift of the corresponding spectral line – the Lamb shift. It might seem that such a tiny effect would be deemed insignificant, but in this case, the observed shift (which was explained by Hans Bethe in the same year) provided considerable insight into quantum electrodynamics.

In quantum electrodynamics, a quantised radiation field has a zero-point energy equivalent to the mean-square electric field so that even in a vacuum there are fluctuations. These fluctuations cause an electron to execute an oscillatory motion and its charge is therefore smeared. If the electron is bound by a non-uniform electric field (as in hydrogen), it experiences a different potential from that appropriate to its mean position. Hence the atomic levels are shifted. In hydrogen-like atoms, the smearing occurs over a length scale,

$$\langle(\delta \mathbf{r})^2\rangle \simeq \frac{2\alpha}{\pi} \left(\frac{\hbar}{mc}\right)^2 \ln\left(\frac{1}{\alpha Z}\right), \quad (8.103)$$

some five orders of magnitude smaller than the Bohr radius. This causes the electron spin g -factor to be slightly different from 2,

$$g_s = 2 \left(1 + \frac{\alpha}{2\pi} - 0.328 \frac{\alpha^2}{\pi^2} + \dots\right). \quad (8.104)$$

There is also a slight weakening of the force on the electron when it is very close to the nucleus, causing the 2S_{1/2} electron (which has penetrated all the way to the nucleus) to be

slightly higher in energy than the $2P_{1/2}$ electron. Taking into account these corrections, one obtains a positive energy shift

$$\Delta E_{\text{Lamb}} \simeq \left(\frac{Z}{n}\right)^4 n\alpha^2 \text{Ry} \times \left(\frac{8}{3\pi}\alpha \ln\left(\frac{1}{\alpha Z}\right)\right) \delta_{\ell,0}, \quad (8.105)$$

for states with $\ell = 0$.

8.2.6 Hyperfine Structure

So far, we have considered the nucleus as simply a massive point charge responsible for the large electrostatic interaction with the charged electrons which surround it. However, the nucleus has a spin angular momentum which is associated with a further set of **hyperfine corrections** to the atomic spectra of atoms. The term hyperfine collectively denotes any effect due to the atomic nucleus, except those solely due to its electric charge or finite mass, so includes effects due to the finite nuclear size, spin, magnetic dipole moment, electric quadrupole moment, etc. The most important hyperfine effects are due to the magnetic dipole moment associated with the nuclear spin, I . As with electrons, the protons and neutrons that make up a nucleus are fermions, each with intrinsic spin $1/2$. This means that a nucleus will have some total nuclear spin which is labelled by the quantum number, I . The latter leads to a nuclear magnetic moment,

$$\hat{\mu}_N = g_N \frac{Ze}{2M_N} \hat{\mathbf{I}}, \quad (8.106)$$

where M_N denotes the mass of the nucleus, and g_N denotes the nuclear g -factor. Since the nucleus has internal structure, the nuclear g -factor is not simply 2 as it (nearly) is for the electron, the difference arises in QED for roughly the same reason as the Lamb shift, the electron is not simply a pointlike particle, but is continually emitting and absorbing virtual photons. For Hydrogen, this nuclear magnetic moment is

$$(\hat{\mu}_S)_p = g_p \frac{\mu_N}{\hbar} \hat{\mathbf{I}}, \quad (8.107)$$

where for the proton, the sole nuclear constituent of atomic hydrogen, $g_P \approx 5.586$. Even though the neutron is charge neutral, its g -factor is about -3.83 . (The constituent quarks have g -factors of 2 (plus corrections) like the electron but the problem is complicated by the strong interactions which make it hard to define a quark's mass.) We can compute (to some accuracy) the g -factor of nuclei from that of protons and neutrons as we can compute the proton's g -factor from its quark constituents. Since the nuclear mass is several orders of magnitude higher than that of the electron, the nuclear magnetic moment provides only a small perturbation.

The proton magnetic moment $(\hat{\mu}_S)_p$ (8.107) sets up a magnetic field \mathbf{B}_p . The atomic electron magnetic dipole moment $(\hat{\mu}_S)_e$ located within the proton's magnetic field \mathbf{B}_p gives rise to a hyperfine interaction

$$\hat{H}_{\text{hf}} = -(\hat{\mu}_S)_e \cdot \mathbf{B}_p = \frac{\mu_B}{\hbar} (\hat{\mathbf{L}} + g_e \hat{\mathbf{S}}) \cdot \mathbf{B}_p. \quad (8.108)$$

Classically, a magnetic moment \mathbf{M} generates a magnetic field given by

$$\begin{aligned}\mathbf{A} &= \frac{\mu_0}{4\pi} \frac{\mathbf{M} \times \mathbf{r}}{r^3} \\ \mathbf{B} = \nabla \times \mathbf{A} &= \frac{\mu_0}{4\pi} \left[\underbrace{\frac{3\mathbf{r}(\mathbf{r} \cdot \mathbf{M}) - r^2 M^2}{r^5}}_{\propto 1/r^3} + \frac{8\pi}{3} \mathbf{M} \delta^{(3)}(\mathbf{r}) \right].\end{aligned}\quad (8.109)$$

Substituting Eq. (8.109) into Eq. (8.108), with the vector \mathbf{M} replaced by the proton magnetic dipole operator $(\hat{\mu}_S)_p$, then gives

$$\boxed{\hat{H}_{\text{hf}} = g_p \frac{\mu_N \mu_B}{\hbar^2} \frac{\mu_0}{4\pi} (\hat{\mathbf{L}} + g_e \hat{\mathbf{S}}) \cdot \left[\frac{3\mathbf{r}(\mathbf{r} \cdot \hat{\mathbf{I}} - r^2 \hat{\mathbf{I}}^2)}{r^5} + \frac{8\pi}{3} \hat{\mathbf{I}} \delta^{(3)}(\mathbf{r}) \right]} \quad (8.110)$$

We can again use first-order perturbation theory to estimate the energy corrections.

8.2.6.1 The S-States

To explore the effect of this field, let us consider just the s-electrons, i.e. $\ell = 0$, for simplicity. For an S-state, with $\ell = 0$, the first-order energy correction due to this hyperfine interaction is

$$(\Delta E)_{\text{hf}} = \langle \psi_{n00} | \hat{H}_{\text{hf}} | \psi_{n00} \rangle \int \psi_{n00}^*(\mathbf{r}) \hat{H}_{\text{hf}} \psi_{n00}(\mathbf{r}) d^3\mathbf{r}. \quad (8.111)$$

In this case, there is no contribution from the electron’s orbital angular momentum, $\hat{\mathbf{L}}$, since

$$\langle n00 | \hat{\mathbf{L}} = 0. \quad (8.112)$$

Also, the $3\mathbf{r}(\mathbf{r} \cdot \hat{\mathbf{I}})$ and $-r^2 \hat{\mathbf{I}}^2$ cancel due to symmetry,

$$\langle \hat{\mathbf{x}}^2 \rangle = \langle \hat{\mathbf{y}}^2 \rangle = \langle \hat{\mathbf{z}}^2 \rangle = \frac{1}{3} \langle \hat{\mathbf{r}}^2 \rangle. \quad (8.113)$$

Hence, for S-states, only the spin $\hat{\mathbf{S}}$ and the δ -function terms contribute

$$(\Delta E)_{\text{hf}} = g_e g_p \frac{\mu_N \mu_B}{\hbar^2} \frac{\mu_0}{4\pi} \langle \hat{\mathbf{S}} \cdot \hat{\mathbf{I}} \rangle \frac{8\pi}{3} |\psi_{n00}(0)|^2. \quad (8.114)$$

This is an interaction between the electron and proton magnetic dipoles

$$(\hat{\mu}_S)_e \cdot (\hat{\mu}_S)_p = \left(g_e \frac{\mu_B}{\hbar} \hat{\mathbf{S}} \right) \cdot \left(g_p \frac{\mu_N}{\hbar} \hat{\mathbf{I}} \right). \quad (8.115)$$

Using $|\psi_{n00}|^2 = Z^3 / (\pi a_0^3 n^3)$, we find

$$(\Delta E)_{\text{hf}} = \frac{4g_e g_p}{3\hbar^2} \frac{Z^3}{n^3} \frac{m_e}{m_p} \langle \hat{\mathbf{S}} \cdot \hat{\mathbf{I}} \rangle \alpha^2 \text{Ry}. \quad (8.116)$$

The total angular momentum of the hydrogen atom is

$$\hat{\mathbf{F}} = \hat{\mathbf{L}} + \hat{\mathbf{S}} + \hat{\mathbf{I}}. \quad (8.117)$$

Squaring this relation gives

$$\hat{\mathbf{F}}^2 = \hat{\mathbf{L}}^2 + \hat{\mathbf{S}}^2 + \hat{\mathbf{I}}^2 + 2\hat{\mathbf{L}} \cdot \hat{\mathbf{S}} + 2\hat{\mathbf{L}} \cdot \hat{\mathbf{I}} + 2\hat{\mathbf{S}} \cdot \hat{\mathbf{I}}. \quad (8.118)$$

In n S-state, with $\ell = 0$, $\langle \hat{\mathbf{L}}^2 \rangle = \langle \hat{\mathbf{L}} \cdot \hat{\mathbf{S}} \rangle = \langle \hat{\mathbf{L}} \cdot \hat{\mathbf{I}} \rangle$, so

$$\langle \hat{\mathbf{F}}^2 \rangle = \langle \hat{\mathbf{S}}^2 \rangle + \langle \hat{\mathbf{I}}^2 \rangle + 2 \langle \hat{\mathbf{S}} \cdot \hat{\mathbf{I}} \rangle, \quad (8.119)$$

or equivalently, for $\ell = 0$ we can use $\hat{\mathbf{F}} = \hat{\mathbf{S}} + \hat{\mathbf{I}}$.

With this, the expectation value $\langle \hat{\mathbf{S}} \cdot \hat{\mathbf{I}} \rangle$ can be evaluated as

$$\begin{aligned} \langle \hat{\mathbf{S}} \cdot \hat{\mathbf{I}} \rangle &= \frac{1}{2} [\langle \hat{\mathbf{F}}^2 \rangle - \langle \hat{\mathbf{S}}^2 \rangle - \langle \hat{\mathbf{I}}^2 \rangle] \\ &= \frac{1}{2} [F(F+1) - s(s+1) - I(I+1)] \hbar^2. \end{aligned} \quad (8.120)$$

For hydrogen atom S-states, with $s = 1/2$, and $I = 1/2$, this gives $F = s \otimes I = \{0, 1\}$ and so

$$\langle \hat{\mathbf{S}} \cdot \hat{\mathbf{I}} \rangle = \begin{cases} -\frac{3}{4}\hbar^2 & \text{for } F = 0 \\ +\frac{1}{4}\hbar^2 & \text{for } F = 1 \end{cases}. \quad (8.121)$$

Hence, for hydrogen S-states, first-order perturbation theory predicts

$$\Delta E = \begin{cases} -\frac{3}{4}(\Delta E)_{\text{hf}} & \text{for } F = 0 \\ +\frac{1}{4}(\Delta E)_{\text{hf}} & \text{for } F = 1 \end{cases} \quad (8.122)$$

with

$$(\Delta E)_{\text{hf}} = \frac{4g_e g_p}{3} \frac{Z^3}{n^3} \frac{m_e}{m_p} \alpha^2 \text{Ry}, \quad (8.123)$$

hyperfine splitting.

Therefore, the $1s$ state of Hydrogen is split into two, corresponding to the two possible values $F = 0$ and 1 . The transition between these two levels has frequency 1420 MHz, or wavelength 21 cm, so lies in the radio waveband. It is an important transition for radio astronomy. The hyperfine interaction splits each S-state into two separate levels as shown in Fig. 8.2. Note that the degeneracy of the zeroth-order states is now $g = 4n^2$.

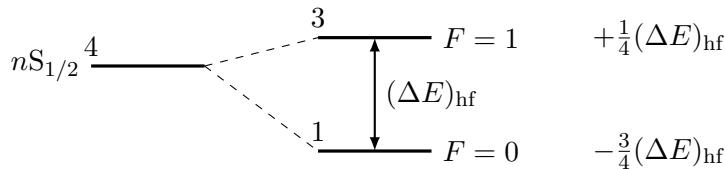


Fig. 8.2: The hyperfine interaction splits each S-states into two separate levels.

Relative to fine structure, hyperfine splitting is suppressed by an extra factor of $m_e/m_p \sim 1/2000$. For the hydrogen atom ground state ($n = 1$), the hyperfine splitting is

$$(\Delta E)_{\text{hf}} \approx 5.87 \times 10^{-6} \text{ eV}, \quad \Delta\nu = 1420.4 \text{ MHz}, \quad \lambda = 21.1 \text{ cm}. \quad (8.124)$$

For *excited S-states* ($n > 1$), the hyperfine splitting is even smaller,

$$(\Delta E_n)_{\text{hf}} \propto 1/n^3. \quad (8.125)$$

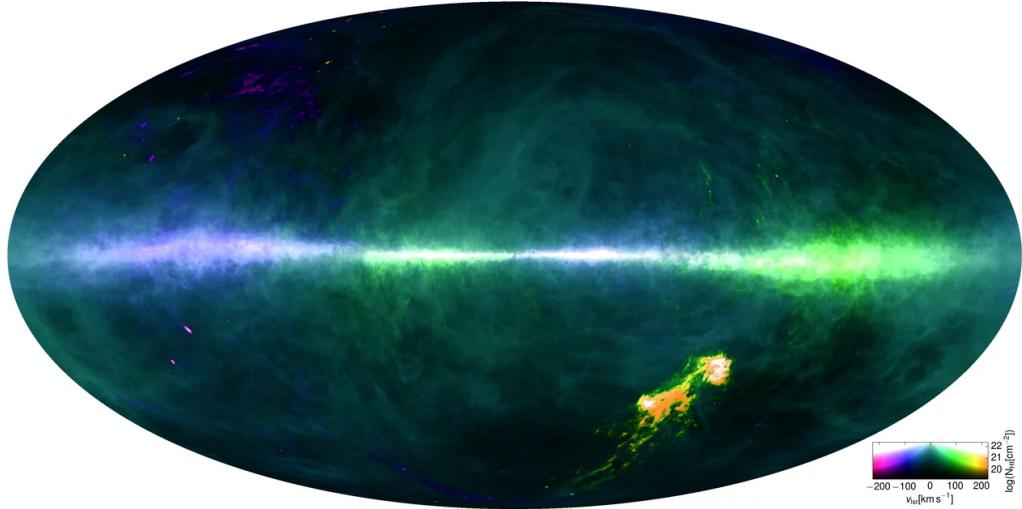


Fig. 8.3: A recent (2016) all-sky survey of neutral atomic hydrogen. Neutral H atoms make up $\sim 44\%$ of interstellar gas (ISG) in the Milky Way. The two small pink areas on the lower left hand side are M31 and M33, and the large orange region represents the Large and Small Magellanic Clouds. The key explains how the colour relates to the relative speed, green is moving away and blue towards. HI4PI Collaboration, A&A **594** (2016) A116.

The 21 cm Line The hydrogen 21 cm line is of great importance in astronomy for detecting the density, motion and temperature of hydrogen gas clouds. The (very small) hyperfine energy splitting, $\Delta E \approx 5.87 \times 10^{-6}$ eV, is equivalent to a temperature of only

$$(\Delta E)_{\text{hf}} = k_B T, \quad T \approx 0.07 \text{ K}. \quad (8.126)$$

Even for temperatures as low as that of the cosmic microwave background (CMB), $T = 2.73$ K, the upper ($F = 1$) level is substantially occupied

$$N(\text{upper}) : N(\text{lower}) \approx 3 : 1, \quad (8.127)$$

from $g = 2F + 1$.

Spontaneous decay of the upper level gives the emission of a 21 cm photon. The emission rate is extremely low, ~ 1 photon per H atom per 120 Myr. But fortunately, there is a lot of neutral hydrogen out there...

8.2.6.2 Non S-States (P,D,F...)

Calculations of the hyperfine corrections for non-S states, with $\ell > 0$, are more involved than for $\ell = 0$; we only briefly summarise the results. The non-S states are also split into

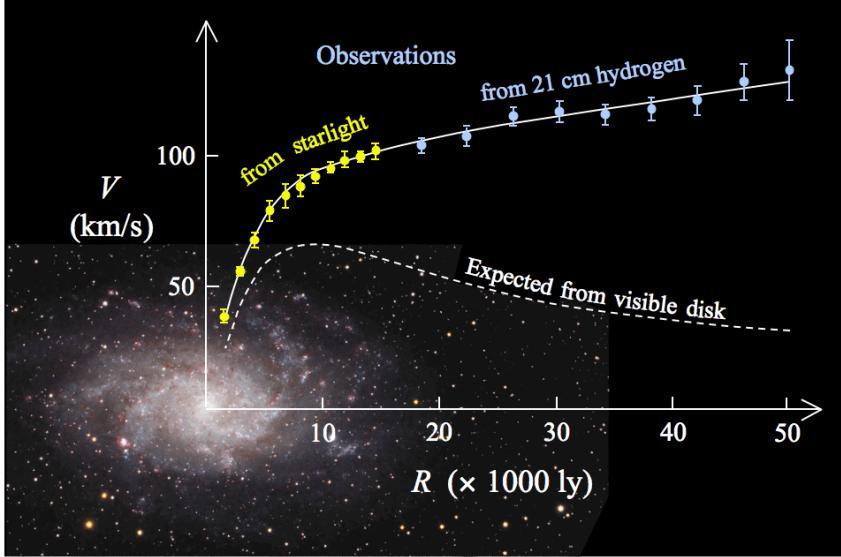


Fig. 8.4: M33 rotation curve (points) compared with the best-fitting model (continuous line). Visible matter alone cannot explain the large rotational velocities seen at large radius; a spherical halo of *dark matter* is also needed. E. Corbelli & P. Salucci, MNRAS **311** (2000) 441

two separate levels by hyperfine interactions, e.g.

$$\text{for } \ell = 1 : \begin{cases} j = 1/2, \quad I = 1/2 \implies F = j \otimes I = \{0, 1\} \\ j = 3/2, \quad I = 1/2 \implies F = j \otimes I = \{1, 2\} \end{cases}. \quad (8.128)$$

The predicted first-order energy shifts turn out to be

$$(\Delta E)_{\text{hf}} = \frac{g_e g_p}{4} \frac{Z^3}{n^3} \frac{m_e}{m_p} \alpha^2 \text{Ry} \left[\frac{F(F+1) - j(j+1) - 3/4}{j(j+1)(\ell + 1/2)} \right]. \quad (8.129)$$

The $\ell > 0$ hyperfine splitting are similar, but smaller, than for $\ell = 0$, e.g.

$$\text{for } n = 2, \quad \ell = 1 : \begin{cases} (\Delta E)_{\text{hf}} \approx 2.5 \times 10^{-7} \text{ eV} & (2P_{1/2}) \\ (\Delta E)_{\text{hf}} \approx 9.8 \times 10^{-8} \text{ eV} & (2P_{3/2}) \end{cases}, \quad (8.130)$$

while for $\ell = 0$ we have

$$\ell = 0 : (\Delta E)_{\text{hf}} \approx 7.5 \times 10^{-7} \text{ eV} \quad (2S_{1/2}) \quad (8.131)$$

A further contribution to the hyperfine structure arises if the **nuclear shape** is not spherical thus distorting the Coulomb potential; this occurs for deuterium and for many other nuclei.

Finally, before leaving this section, we should note that the nucleus is not point-like but has a small size. The effect of **finite nuclear size** can be estimated perturbatively. In doing so, one finds that the s ($\ell = 0$) levels are those most affected, because these have the largest probability of finding the electron close to the nucleus; but the effect is still very small in hydrogen. It can be significant, however, in atoms of high nuclear charge Z , or for **muonic** atoms. A muon is a particle somewhat like an electron, but about 200 times heavier. If a muon is captured by an atom, the corresponding Bohr radius is 200 times smaller, thus enhancing the nuclear size effect.

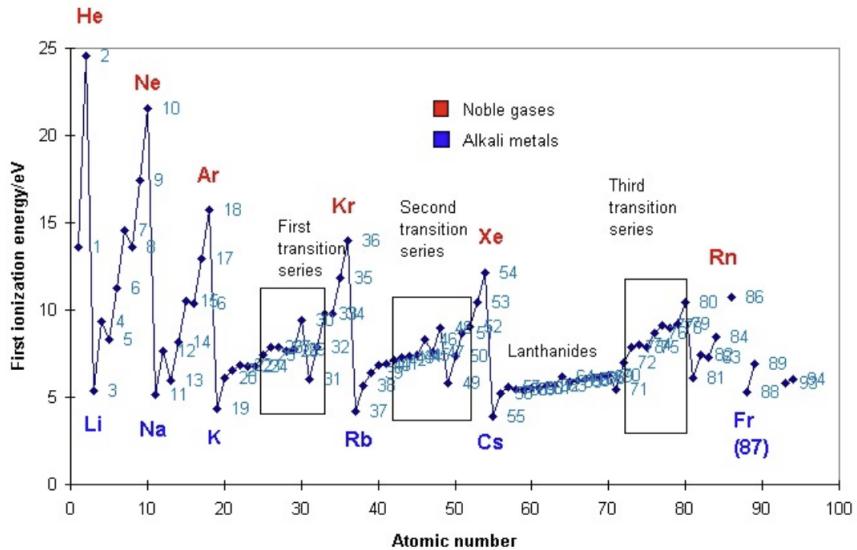


Fig. 8.5: Ionisation energy trends plotted against the atomic number, in units eV.

This completes our discussion of the “one-electron” theory. We now turn to consider the properties of multi-electron atoms.

8.3 Multi-Electron Atoms

To address the electronic structure of a multi-electron atom, we might begin with the hydrogenic energy levels for an atom of nuclear charge Z , and start filling the lowest levels with electrons, accounting for the exclusion principle. The degeneracy for quantum numbers (n, ℓ) is $2 \times (2\ell + 1)$, where $(2\ell + 1)$ is the number of available m_ℓ values, and the factor of 2 accounts for the spin degeneracy. Hence, the number of electrons accommodated in shell, n , would be $2 \times n^2$,

n	ℓ	Degeneracy in shell	Cumulative Total
1	0	2	2
2	0, 1	$(1 + 3) \times 2 = 8$	10
3	0, 1, 2	$(1 + 3 + 5) \times 2 = 18$	28
4	0, 1, 2, 3	$(1 + 3 + 5 + 7) \times 2 = 32$	60

Table 8.6: The number of electrons accommodated due to degeneracy in shell n .

We would therefore expect that atoms containing 2, 10, 28 or 60 electrons would be especially stable, and that in atoms containing one more electron than this, the outermost electron would be less tightly bound. In fact, if we look at data (Fig. 8.5) recording the first ionisation energy of atoms, i.e. the minimum energy needed to remove one electron, we find that the noble gases, having $Z = 2, 10, 18, 36, \dots$ are especially tightly bound, and the elements containing one more electron, the alkali metals, are significantly less tightly bound.

The reason for the failure of this simple-minded approach is that we have neglected

the repulsion between electrons. In fact, the first ionisation energies of atoms show a relatively weak dependence on Z ; this tells us that the outermost electrons are almost completely shielded from the nuclear charge.⁵ Indeed, when we treated the Helium atom as an example of the variational method in section 7.3.2 we found that the effect of electron-electron repulsion was sizeable, and really too large to be treated accurately by perturbation theory.

8.3.1 Central Field Approximation

Leaving aside for now the influence of spin or relativistic effects, the Hamiltonian for a multi-electron atom can be written as

$$\hat{H} = \sum_i \left[-\frac{\hbar^2}{2m} \nabla_i^2 - \frac{1}{4\pi\epsilon_0} \frac{Ze^2}{r_i} \right] + \sum_{i < j} \frac{1}{4\pi\epsilon_0} \frac{e^2}{r_{ij}}, \quad (8.132)$$

where $r_{ij} \equiv |\mathbf{r}_i - \mathbf{r}_j|$. The first term represents the “single-particle” contribution to the Hamiltonian arising from interaction of each electron with the nucleus, while the last term represents the mutual Coulomb interaction between the constituent electrons. It is this latter term that makes the generic problem “many-body” in character and therefore very complicated. Yet, as we have already seen in the perturbative analysis of the excited states of atomic Helium, this term can have important physical consequences both on the overall energy of the problem and on the associated spin structure of the states.

The **central field approximation** is based upon the observation that the electron interaction term contains a large central (spherically symmetric) component arising from the “core electrons”. From the following relation,

$$\sum_{m_\ell=-\ell}^{\ell} |Y_{\ell m_\ell}(\theta, \phi)|^2 = \text{const.} \quad (8.133)$$

it is apparent that a closed shell has an electron density distribution which is isotropic (independent of θ and ϕ). We can therefore develop a perturbative scheme by setting $\hat{H} = \hat{H}_0 + \hat{H}_1$, where

$$\hat{H}_0 = \sum_i \left[-\frac{\hbar^2}{2m} \nabla_i^2 - \frac{1}{4\pi\epsilon_0} \frac{Ze^2}{r_i} + U_i(r_i) \right], \quad \hat{H}_1 = \sum_{i < j} \frac{1}{4\pi\epsilon_0} \frac{e^2}{r_{ij}} - \sum_i U_i(r_i). \quad (8.134)$$

Here the one-electron potential, $U_i(r)$, which is assumed central (see below), incorporates the “average” effect of the other electrons. Before discussing how to choose the potentials $U_i(r)$, let us note that \hat{H}_0 is separable into a sum of terms for each electron, so that the total wavefunction can be factorised into components for each electron. The basic idea is first to solve the Schrödinger equation using \hat{H}_0 , and then to treat \hat{H}_1 as a small perturbation.

⁵In fact, the shielding is not completely perfect. For a given energy shell, the effective nuclear charge varies for an atomic number Z as $Z_{\text{eff}} \sim (1+\alpha)^Z$ where $\alpha > 0$ characterises the ineffectiveness of screening. This implies that the ionisation energy $I_Z = -E_Z \sim Z_{\text{eff}}^2 \sim (1+2\alpha Z)$. The near-linear dependence of I_Z on Z is reflected in Fig. 8.5

On general grounds, since the Hamiltonian \hat{H}_0 continues to commute with the angular momentum operator, $[\hat{H}_0, \hat{\mathbf{L}}] = 0$, we can see that the eigenfunctions of \hat{H}_0 will be characterised by quantum numbers (n, ℓ, m_ℓ, m_s) . However, since the effective potential is no longer Coulomb-like, the ℓ values for a given n need not be degenerate. Of course, the difficult part of this procedure is to estimate $U_i(r)$; the potential energy experienced by each electron depends on the wavefunction of all the other electrons, which is only known after the Schrödinger equation has been solved. This suggests that an iterative approach to solving the problem will be required.

To understand how the potentials $U_i(r)$ can be estimated – the **self-consistent field method** – it is instructive to consider a variational approach due originally to Hartree. If electrons are considered independent, the wavefunction can be factorised into the product state,

$$\Psi(\{\mathbf{r}_i\}) = \psi_{i_1}(\mathbf{r}_1)\psi_{i_2}(\mathbf{r}_2) \cdots \psi_{i_N}(\mathbf{r}_N), \quad (8.135)$$

where the quantum numbers, $i_k \equiv (nlm_\ell m_s)_k$, indicate the individual state occupancies. Note that this product state is not a properly antisymmetrised Slater determinant – the exclusion principle is taken into account only in so far as the energy of the ground state is taken to be the lowest that is consistent with the assignment of different quantum numbers, $nlm_\ell m_s$ to each electron. Nevertheless, using this wavefunction as a trial state, the variational energy is then given by

$$\begin{aligned} E = \langle \Psi | \hat{H} | \Psi \rangle &= \sum_i \int d^3r \psi_i^* \left(-\frac{\hbar^2}{2m} \nabla^2 - \frac{1}{4\pi\epsilon_0} \frac{Ze^2}{r} \right) \psi_i \\ &\quad + \frac{1}{4\pi\epsilon_0} \sum_{i < j} \int d^3r \int d^3r' \psi_i^*(\mathbf{r}) \psi_j^*(\mathbf{r}') \frac{e^2}{|\mathbf{r} - \mathbf{r}'|} \psi_j(\mathbf{r}') \psi_i(\mathbf{r}). \end{aligned} \quad (8.136)$$

Now, according to the variational principle, we must minimise the energy functional by varying $E[\psi_i]$ with respect to the complex wavefunction, ψ_i , subject to the normalisation condition, $\langle \psi | \psi \rangle = 1$. The latter can be imposed using a set of Lagrange multipliers, ε_i , i.e.

$$\frac{\delta}{\delta \psi_i^*} \left[E - \varepsilon_i \left(\int d^3r |\psi_i(\mathbf{r})|^2 - 1 \right) \right] = 0. \quad (8.137)$$

Following the variation,⁶ one obtains the **Hartree equations**,

$$\left(-\frac{\hbar^2}{2m} \nabla^2 - \frac{1}{4\pi\epsilon_0} \frac{Ze^2}{r} \right) \psi_i + \frac{1}{4\pi\epsilon_0} \sum_{j \neq i} \int d^3r' |\psi_j(\mathbf{r}')|^2 \frac{e^2}{|\mathbf{r} - \mathbf{r}'|} \psi_i(\mathbf{r}) = \varepsilon_i \psi_i(\mathbf{r}). \quad (8.138)$$

Then according to the variational principle, amongst all possible trial functions ψ_i , the set that minimises the energy are determined by the effective potential,

$$U_i(\mathbf{r}) = \frac{1}{4\pi\epsilon_0} \sum_{j \neq i} \int d^3r' |\psi_j(\mathbf{r}')|^2 \frac{e^2}{|\mathbf{r} - \mathbf{r}'|}. \quad (8.139)$$

Eq. (8.138) has a simple interpretation: The first two terms relate to the nuclear potential experienced by the individual electrons, while the third term represents the electrostatic

⁶Note that, in applying the variation, the wavefunction ψ_i^* can be considered independent of ψ_i

potential due to the other electrons. However, to simplify the procedure, it is useful to engineer the radial symmetry of the potential by replacing $U_i(\mathbf{r})$ by its spherical average,

$$U_i(\mathbf{r}) \mapsto U_i(r) = \int \frac{d\Omega}{4\pi} U_i(\mathbf{r}). \quad (8.140)$$

Finally, to relate the Lagrange multipliers, ε_i (which have the appearance of one-electron energies), to the total energy, we can multiply Eq. (8.138) by $\psi_i^*(\mathbf{r})$ and integrate,

$$\varepsilon_i = \int d^3r \psi_i^* \left(-\frac{\hbar^2}{2m} \nabla^2 - \frac{1}{4\pi\epsilon_0} \frac{Ze^2}{r} \right) \psi_i + \frac{1}{4\pi\epsilon_0} \sum_{j \neq i} \int d^3r' d^3r |\psi_j(\mathbf{r}')|^2 \frac{e^2}{|\mathbf{r} - \mathbf{r}'|} |\psi_i(\mathbf{r})|^2. \quad (8.141)$$

If we compare this expression with the variational state energy, we find that

$$E = \sum_i \varepsilon_i - \sum_{i < j} \frac{1}{4\pi\epsilon_0} \sum_{j \neq i} \int d^3r' d^3r |\psi_j(\mathbf{r}')|^2 \frac{e^2}{|\mathbf{r} - \mathbf{r}'|} |\psi_i(\mathbf{r})|^2. \quad (8.142)$$

To summarise, if we wish to implement the central field approximation to determine the states of a multi-electron atom, we must follow the algorithm:

1. Firstly, one makes an initial “guess” for a (common) central potential, $U(r)$. As $r \rightarrow 0$, screening becomes increasingly ineffective and we expect $U(r) \rightarrow 0$. As $r \rightarrow \infty$, we anticipate that $U(r) \rightarrow \frac{1}{4\pi\epsilon_0} \frac{(Z-1)e^2}{r}$, corresponding to perfect screening. So, as a starting point, we make take some smooth function $U(r)$ interpolating between these limits. For this trial potential, we can solve (numerically) for the eigenstates of the single-particle Hamiltonian. We can then use these states as a platform to build the product wavefunction and in turn determine the self-consistent potentials, $U_i(r)$.
 2. With these potentials, $U_i(r)$, we can determine a new set of eigenstates for the set of Schrödinger equations,
- $$\left[-\frac{\hbar^2}{2m} \nabla^2 - \frac{1}{4\pi\epsilon_0} \frac{Ze^2}{r} + U_i(r) \right] \psi_i = \varepsilon_i \psi_i. \quad (8.143)$$
3. An estimate for the ground state energy of an atom can be found by filling up the energy levels, starting from the lowest, and taking account of the exclusion principle.
 4. Using these wavefunctions, one can then make an improved estimate of the potentials $U_i(r_i)$ and return to step 2 iterating until convergence.

An improvement to this procedure, known the **Hartree-Fock** method, takes account of exchange interactions. In order to do this, it is necessary to ensure that the wavefunction, including spin, is antisymmetric under interchange of any pair of electrons. This is achieved by introducing the **Slater determinant**. Writing the individual electron wavefunction for the i th electron as $\psi_k(r_i)$, where $i = 1, 2, \dots, N$ and k is shorthand for the

set of quantum numbers ($n\ell m_\ell m_s$), the overall wavefunction is given by

$$\Psi = \frac{1}{\sqrt{N!}} \begin{vmatrix} \psi_1(\mathbf{r}_1) & \psi_1(\mathbf{r}_2) & \psi_1(\mathbf{r}_3) & \cdots \\ \psi_2(\mathbf{r}_1) & \psi_2(\mathbf{r}_2) & \psi_2(\mathbf{r}_3) & \cdots \\ \psi_3(\mathbf{r}_1) & \psi_3(\mathbf{r}_2) & \psi_3(\mathbf{r}_3) & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{vmatrix}. \quad (8.144)$$

Note that each of the $N!$ terms in Ψ is a product of wavefunctions for each individual electron. The $1/\sqrt{N!}$ factor ensures the wavefunction is normalised. A determinant changes sign if any two columns are exchanged, corresponding to $\mathbf{r}_i \leftrightarrow \mathbf{r}_j$ (say); this ensures that the wavefunction is antisymmetric under exchange of electrons i and j . Likewise, a determinant is zero if any two rows are identical; hence all the ψ_k s must be different and the Pauli exclusion principle is satisfied.⁷ In this approximation, a variational analysis leads to the Hartree-Fock equations,

$$\begin{aligned} \varepsilon_i \psi_i(\mathbf{r}) = & \left[-\frac{\hbar^2}{2m} \nabla_i^2 - \frac{1}{4\pi\epsilon_0} \frac{Ze^2}{r_i} \right] \\ & + \frac{1}{4\pi\epsilon_0} \sum_{j \neq i} \int d^3 r_j \frac{e^2}{|\mathbf{r} - \mathbf{r}'|} \psi_j^*(\mathbf{r}') [\psi_j(\mathbf{r}') \psi_i(\mathbf{r}) - \psi_j(\mathbf{r}) \psi_i(\mathbf{r}') \delta_{m_{s_i}, m_{s_j}}]. \end{aligned} \quad (8.145)$$

The first term in the last set of brackets translates to the ordinary Hartree contribution above and describes the influence of the charge density of the other electrons, while the second term describes the non-local **exchange** contribution, a manifestation of particle statistics.

The outcome of such calculations is that the eigenfunctions are, as for hydrogen, characterised by quantum numbers n, ℓ, m_ℓ , with $\ell < n$, but that the states with different ℓ for a given n are not degenerate, with the lower values of ℓ lying lower. This is because, for the higher ℓ values, the electrons tend to lie further from the nucleus on average, and are therefore more effectively screened. The states corresponding to a particular value of n are generally referred to as a **shell**, and those belonging to a particular pair of values of n, ℓ are usually referred to as a **subshell**. The energy levels are ordered as in Table 8.7 below (with the lowest lying on the left):

Subshell name	1s	2s	2p	3s	3p	4s	3d	4p	5s	4d	...
$n =$	1	2	2	3	3	4	3	4	5	4	...
$\ell =$	0	0	1	0	1	0	2	1	0	2	...
Degeneracy	2	2	6	2	6	2	10	6	2	10	...
Cumulative	2	4	10	12	18	20	30	36	38	48	...

Table 8.7: Ordering of Hydrogen orbitals for each shell n , where ℓ denotes the subshell.

Note that the values of Z corresponding to the noble gases, 2, 10, 18, 36, at which the ionisation energy is unusually high, now emerge naturally from this filling order, corresponding to the numbers of electrons just before a new shell (n) is entered. There is a handy mnemonic to remember this filling order. By writing the subshells down as

⁷Note that for $N = 2$, the determinant reduces to the familiar antisymmetric wavefunction, $1/\sqrt{2}[\psi_1(\mathbf{r}_1)\psi_2(\mathbf{r}_2) - \psi_2(\mathbf{r}_1)\psi_1(\mathbf{r}_2)]$.

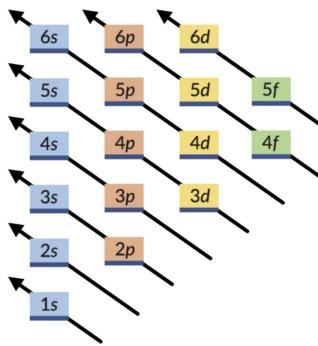


Fig. 8.6: Electron filling order illustrated by the Aufbau principle.

shown in figure 8.6, the order of states can be read off along diagonals from lower right to upper left, starting at the bottom.

We can use this sequence of energy levels to predict the ground state electron configuration of atoms. We simply fill up the levels starting from the lowest, accounting for the exclusion principle, until the electrons are all accommodated (the **aufbau principle**). A few examples are contained in table 8.8. Since it is generally the outermost electrons

Z	Element	Configuration	$2S+1L_J$	Ionis. Pot. (eV)
1	H	(1s)	$^2S_{1/2}$	13.6
2	He	(1s) ²	1S_0	24.6
3	Li	He (2s)	$^2S_{1/2}$	5.4
4	Be	He (2s) ²	1S_0	9.3
5	B	He (2s) ² (2p)	$^2P_{1/2}$	8.3
6	C	He (2s) ² (2p) ²	3P_0	11.3
7	N	He (2s)(2p) ³	$^4S_{3/2}$	14.5
8	O	He (2s) ² (2p) ⁴	3P_2	13.6
9	F	He (2s) ² (2p) ⁵	$^2P_{3/2}$	17.4
10	Ne	He (2s) ² (2p) ⁶	1S_0	21.6
11	Na	Ne (3s)	$^2S_{1/2}$	5.1

Table 8.8: Electronic subshell configuration, according to the aufbau principle.

which are of most interest, contributing to chemical activity or optical spectra, one often omits the inner closed shells, and just writes O as (2p)⁴, for example. However, the configuration is not always correctly predicted, especially in the heavier elements, where levels may be close together. It may be favourable to promote one or even two electrons one level above that expected in this simple picture, in order to achieve a filled shell. For example, Cu ($Z = 29$) would be expected to have configuration $\cdots(4s)^2(3d)^9$, and actually has configuration $\cdots(4s)^1(3d)^{10}$. There are several similar examples in the transition elements where the d subshells are being filled, and many among the lanthanides (rare earths) and actinides where f subshells are being filled.

Since the assignment of an electron configuration requires only the enumeration of the values of n and ℓ for all electrons, but not those of m_ℓ and m_s , each configuration will be accompanied by a **degeneracy** g . If $\nu_{n\ell}$ denotes the number of electrons occupying a

given level $E_{n,\ell}$, and $\delta_\ell = 2 \times (2\ell + 1)$ is the degeneracy of that level, there are

$$d_{n\ell} = \frac{\delta_\ell!}{\nu_{n\ell}!(\delta_\ell - \nu_{n\ell})!} \quad (8.146)$$

ways of distributing the $\nu_{n\ell}$ electrons among the δ_ℓ individual states. The total degeneracy, g , is then obtained from the product.

This scheme provides a basis to understand the periodic table of elements. We would expect that elements which have similar electron configurations in their outermost shells (such as Li, Na, K, Rb, Cs, Fr which all have $(ns)^1$ or F, Cl, Br, I, which all have $(np)^5$) would have similar chemical properties, such as valency, since it is the unpaired outer electrons which especially participate in chemical bonding. Therefore, if one arranges the atoms in order of increasing atomic number Z (which equals the number of electrons in the atom), periodic behaviour is seen whenever a new subshell of a given ℓ is filled.

8.3.2 Spin-Orbit Coupling

The procedure outlined above allows us to predict the occupation of subshells in an atomic ground state. This is not in general sufficient to specify the ground state fully. If there are several electrons in a partially filled subshell, then their spins and orbital angular momenta can combine in several different ways, to give different values of total angular momentum, with different energies. In order to deal with this problem, it is necessary to consider the spin-orbit interaction as well as the residual Coulomb interaction between the outer electrons.

Schematically we can write the Hamiltonian for this system as follows:

$$\hat{H} \approx \underbrace{\hat{H}_0 + \sum_{i < j} \frac{1}{4\pi\epsilon_0} \frac{e^2}{r_{ij}} - \sum_i U_i(r_i)}_{\hat{H}_1} + \underbrace{\sum_i \xi_i(r_i) \hat{\mathbf{L}}_i \cdot \hat{\mathbf{S}}}_{\hat{H}_{SO}}, \quad (8.147)$$

where \hat{H}_0 includes the kinetic energy and central field terms, \hat{H}_1 is the residual Coulomb interaction, and (with $\xi_i(r) = \frac{1}{2m^2c^2} \frac{1}{r} (\partial_r V(r))$) \hat{H}_{SO} is the spin-orbit interaction. We can then consider two possible scenarios:

- $\hat{H}_{SO} \ll \hat{H}_1$: This tends to apply in the case of light atoms. In this situation, one considers first the eigenstates of $\hat{H}_0 + \hat{H}_1$, and then treats \hat{H}_{SO} as a perturbation. This leads to a scheme called **LS** (or **Russell-Saunders coupling**).
- $\hat{H}_{SO} \gg \hat{H}_1$: This can apply in very heavy atoms, or in heavily ionised light atoms, in which the electrons are moving at higher velocities and relativistic effects such as the spin-orbit interaction are more important. In this case, a scheme called *jj* coupling applies.

It is important to emphasise that both of these scenarios represent approximations; real atoms do not always conform to the comparatively simple picture which emerges from these schemes.

8.3.2.1 LS Coupling Scheme

In this approximation, we start by considering the eigenstates of $\hat{H}_0 + \hat{H}_1$. We note that this Hamiltonian must commute with the total angular momentum $\hat{\mathbf{J}}^2$ (because of invariance under rotations in space), and also clearly commutes with the total spin $\hat{\mathbf{S}}^2$. It also commutes with the total orbital angular momentum $\hat{\mathbf{L}}^2$, since \hat{H}_1 only involves internal interactions, and must therefore be invariant under global rotation of all the electrons. Therefore the energy levels can be characterised by the corresponding total angular momentum quantum numbers L, S, J . Their ordering in energy is given by Hund's rules:

1. Combine the spins of the electrons to obtain possible values of total spin S . The largest permitted value of S lies lowest in energy.
2. For this value of S , find the possible values of total angular momentum L . The largest value of L lies lowest in energy.
3. Couple the values of L and S to obtain the values of J (hence the name of the scheme). If the subshell is less than half full, the smallest value of J lies lowest; otherwise, the largest value of J lies lowest.

In deciding on the permitted values of L and S , in addition to applying the usual rules for adding angular momenta, one also has to ensure that the exclusion principle is respected, as we will see later when considering some examples.

These rules are empirical; there are exceptions, especially to the L and J rules (2 and 3). Nevertheless, Hund's rules are a useful guide, and we should try to understand their physical origin.

1. Maximising S makes the spin wavefunction as symmetric as possible. This tends to make the spatial wavefunction antisymmetric, and hence reduces the Coulomb repulsion, as we saw when discussing the exchange interactions in Helium and the origin of itinerant ferromagnetism.
2. Maximising L also tends to keep the electrons apart. This is less obvious, though a simple classical picture of electrons rotating round the nucleus in the same or different senses makes it at least plausible.
3. The separation of energies for states of different J arises from treating the spin-orbit term \hat{H}_{SO} as a perturbation (fine structure). It can be shown (using the Wigner-Eckart theorem) that

$$\begin{aligned} \langle J, m_J, L, S | \sum_i \xi_i(r_i) \hat{\mathbf{L}}_i \cdot \hat{\mathbf{S}}_i | J, m_J, L, S \rangle \\ = \zeta(L, S) \langle \hat{\mathbf{L}}J, m_J, L, S \cdot \hat{\mathbf{S}} \rangle \\ = \frac{\zeta(L, S)}{2} [J(J+1) - L(L+1) - S(S+1)], \quad (8.148) \end{aligned}$$

where the matrix element $\zeta(L, S)$ depends on the total L and S values. Since one may show that the sign of $\zeta(L, S)$ changes according to whether the subshell is more or less than half-filled, the third Hund's rule is established.

To understand the application of LS coupling, it is best to work through some examples. Starting with the simplest multi-electron atom, helium, the ground state has an electron configuration $(1s)^2$, and must therefore have $L = S = J = 0$. In fact, for any completely filled subshell, we have $L = S = 0$ and hence $J = 0$, since the total m_L and m_S must equal zero if all substates are occupied. Consider now an excited state of helium, e.g. $(1s)^1(2p)^1$, in which one electron has been excited to the $2p$ level. We can now have $S = 1$ or $S = 0$, with the $S = 1$ state lying lower in energy according to Hund's rules. Combining the orbital angular momenta of the electrons yields $L = 1$ and thus, with $S = 0$, $J = 1$, while with $S = 1$, $J = 0, 1, 2$ with $J = 0$ lying lowest in energy.

Once again, as with the hydrogen-like states, we may index the states of multi-electron atoms by spectroscopic term notation, $^{2S+1}L_J$. The superscript $2S + 1$ gives the multiplicity of J values into which the level is split by the spin-orbit interaction; the L value is represented by a capital letter, S , P , D , etc., and J is represented by its numerical value. Thus, for the $(1s)^1(2p)^1$ state of helium, there are four possible states, with terms:

$$^3P_0 \ ^3P_1 \ ^3P_2 \quad ^1P_1, \quad (8.149)$$

where the three 3P states are separated by the spin-orbit interaction, and the singlet 1P state lies much higher in energy owing to the Coulomb interaction. The separations between the 3P_2 and 3P_1 and the 3P_1 and 3P_0 should be in the ratio $2 : 1$. This is an example of the **Landé interval rule**, which states that the separation between a pair of adjacent levels in a fine structure multiplet is proportional to the larger of the two J values involved. This is easily shown using Eq. (8.148) – the separation in energy between states J and $J - 1$ is

$$\propto J(J + 1) - (J - 1)J = 2J. \quad (8.150)$$

Actually in the case of helium the situation is a bit more complicated, because it turns out that the spin-orbit interaction between *different* electrons makes a non-negligible additional contribution to the fine structure. Other excited states of helium, of the form $(1s)^1(n\ell)^1$, can be handled similarly, and again separate into singlet and triplet states.

We next consider the case of carbon, which has ground state electron configuration $(1s)^2(2s)^2(2p)^2$. This introduces a further complication; we now have two identical electrons in the same unfilled subshell, and we need to ensure that their wavefunction is antisymmetric with respect to electron exchange. The total spin can either be the singlet $S = 0$ state, which has an antisymmetric wavefunction $\frac{1}{\sqrt{2}}[|\uparrow_1\rangle \otimes |\downarrow_2\rangle - |\downarrow_1\rangle \otimes |\uparrow_2\rangle]$, or one of the triplet $S = 1$ states, which are symmetric, $\frac{1}{\sqrt{2}}[|\uparrow_1\rangle \otimes |\downarrow_2\rangle + |\downarrow_1\rangle \otimes |\uparrow_2\rangle]$, $|\uparrow_1\rangle \otimes |\uparrow_2\rangle$, or $|\downarrow_1\rangle \otimes |\downarrow_2\rangle$. We must therefore choose values of L with the appropriate symmetry to partner each value of S . To form an antisymmetric state, the two electrons must have different values of m_ℓ , so the possibilities are as shown in table 8.9. Inspecting the values of m_L we can deduce that $L = 1$.⁸ By contrast, to form a symmetric total angular momentum state, the two electrons may have any values of m_ℓ , leading to the possibilities shown in table 8.10. Inspecting the values of m_L we infer that $L = 2$ or 0 .

⁸This result would also be apparent if we recall that angular momentum states are eigenstates of

$m_\ell^{(1)}$	$m_\ell^{(2)}$	m_L
1	0	1
1	-1	0
0	-1	-1

Table 8.9: Possibilities of m_ℓ values for electrons to form an antisymmetric total angular momentum state.

$m_\ell^{(1)}$	$m_\ell^{(2)}$	m_L
1	1	2
1	0	1
1	-1	0
0	0	0
0	-1	-1
-1	-1	-2

Table 8.10: Possibilities of m_ℓ values for electrons to form an symmetric total angular momentum state.

We must therefore take $S = 1$ with $L = 1$ and $S = 0$ with $L = 2$ or 0. Finally, to account for the fine structure, we note that the states with $S = 1$ and $L = 1$ can be combined into a single $J = 0$ state, three $J = 1$ states, and five $J = 2$ states leading to the terms 3P_0 , 3P_1 , and 3P_2 respectively. Similarly the $S = 0$, $L = 2$ state can be combined to give five $J = 2$ states, 1D_2 , while the $S = 0$, $L = 0$ state gives the single $J = 0$ state, 1S_0 . Altogether we recover the $1 + 3 + 5 + 5 + 1 = 15$ possible states (cf. Eq. (8.146)) with the ordering in energy given by Hund's rules, shown in table 8.11. The experimental energy values are given using the conventional spectroscopic units of inverse wavelength. Note that the Landé interval rule is approximately obeyed by the fine structure triplet, and that the separation between L and S values caused by the electron-electron repulsion is *much* greater than the spin-orbit effect. In an excited state of carbon, e.g. $(2p)^1(3p)^1$,

	E / cm^{-1}
1S_0	20649
1D_2	10195
3P_2	43
3P_1	16
3P_0	0

Table 8.11: Experimental energy ordering of Carbon states using spectroscopy.

the electrons are no longer equivalent, because they have different radial wavefunctions. So now one can combine any of $S = 0, 1$ with any of $L = 0, 1, 2$, yielding the following terms (in order of increasing energy, according to Hund's rules):

$${}^3D_{1,2,3} \quad {}^3P_{0,1,2} \quad {}^3S_1 \quad {}^1D_2 \quad {}^1P_1 \quad {}^1S_0. \quad (8.151)$$

For **nitrogen**, the electron configuration is given by $(1s)^2(2s)^2(2p)^3$. The maximal value of spin is $S = 3/2$ while L can take values 3, 2, 1 and 0. Since the spin wavefunction

the parity operator with eigenvalue $(-1)^L$. Since there are just two electrons, this result shows that both the $L = 0$ and $L = 2$ wavefunction must be symmetric under exchange.

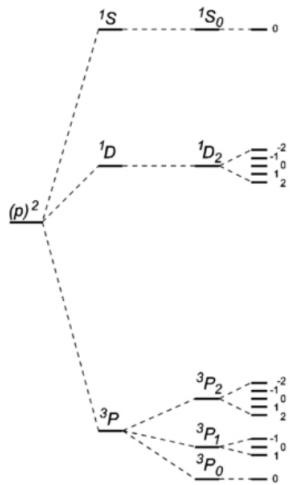


Fig. 8.7: Level scheme of the carbon atom $(1s)^2(2s)^2(2p)^2$. Drawing is not to scale. On the left the energy is shown without any two-particle interaction. The electron-electron interaction leads to a three-fold energy splitting with L and S remaining good quantum numbers. Spin-orbit coupling leads to a further splitting of the states with J remaining a good quantum number. Finally on the right, the levels show Zeeman splittings in an external magnetic field. In this case, the full set of 15 levels become non-degenerate.

(being maximal) is symmetric, the spatial wavefunction must be completely antisymmetric. This demands that all three states with $m_\ell = 1, 0, -1$ must be involved. We must therefore have $L = 0$, leading to $J = 3/2$ and the term, ${}^4S_{3/2}$.

As a final example, let us consider the ground state of **oxygen**, which has electron configuration $(2p)^4$. Although there are four electrons in the $(2p)$ subshell, the maximum value of $S = 1$. This is because there are only three available values of $m_\ell = \pm 1, 0$, and therefore one of these must contain two electrons with opposite spins. Therefore, the maximum value of $S = 1$, achieved by having electrons with $m_s = +\frac{1}{2}$ in both the other m_ℓ states. By pursuing this argument, it is quite easy to see that the allowed values of L , S and J are the same as for carbon $(2p)^2$. This is in fact a general result – the allowed quantum numbers for a subshell with n electrons are the same as for that of a subshell with n “holes”. Therefore, the energy levels for the oxygen ground state configuration are the same as for carbon, except that the fine structure multiplet is inverted, in accordance with Hund’s third rule.

8.3.3 jj Coupling Scheme

When relativistic effects take precedence over electron interaction effects, we must start by considering the eigenstates of $\hat{H}_0 + \hat{H}_{SO} = \hat{H}_0 + \sum_i \zeta_i(r_i) \hat{\mathbf{L}}_i \cdot \hat{\mathbf{S}}_i$. These must be eigenstates of $\hat{\mathbf{J}}^2$ as before, because of the overall rotational invariance, and also of $\hat{\mathbf{J}}_i^2$ for each electron. Therefore, in this case, the coupling procedure is to find the allowed j values of individual electrons, whose energies will be separated by the spin-orbit interaction. Then these individual j values are combined to find the allowed values of total J . The effect of the residual Coulomb interaction will be to split the J values for a given set of j s. Sadly, in this case, there are no simple rules to parallel those of Hund.

As an example, consider a configuration $(np)^2$ in the jj coupling scheme, to be compared with the example of carbon which we studied in the LS scheme. Combining $s = 1/2$ with $\ell = 1$, each electron can have $j = 1/2$ or $3/2$. If the electrons have the same j value, they are equivalent, so we have to take care of the symmetry of the wavefunction. We therefore have the following possibilities:

- $j_1 = j_2 = 3/2 \implies J = 3, 2, 1, 0$, of which $J = 2, 0$ are antisymmetric.
- $j_1 = j_2 = 1/2 \implies J = 1, 0$, of which $J = 0$ is antisymmetric.
- $j_1 = 1/2, j_2 = 3/2 \implies J = 2, 1$.

In jj coupling, the term is written $(j_1, j_2)_J$, so we have the following terms in our example:

$$(1/2, 1/2)_0 \quad (3/2, 1/2)_1 (3/2, 1/2)_2 \quad (3/2, 3/2)_2 (3/2, 3/2)_0 \quad (8.152)$$

in order of increasing energy. Note that both LS and jj coupling give the same values of J (in this case, two states with $J = 0$, two with $J = 2$ and one with $J = 1$) and in the same order. However, the pattern of levels is different; in LS coupling we found a triplet and two singlets, while in this ideal jj scenario, we have two doublets and a singlet. The sets of states in the two coupling schemes must be expressible as linear combinations of one another, and the physical states for a real atom are likely to differ from either approximation.

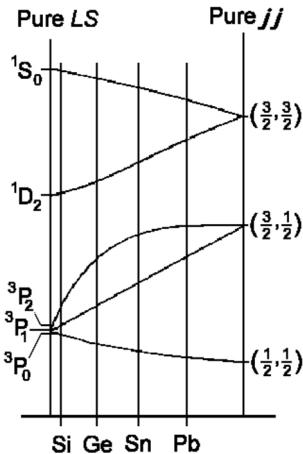


Fig. 8.8: The transition from LS to jj coupling is seen in the series C-Si-Ge-Sn-Pb

In fact, this idealised form of jj coupling is not seen in the heaviest such atom in the periodic table, lead $(6p)^2$. However, it is seen in some highly ionised states, for example in Cr^{18+} , which has the same electron configuration $(2p)^2$ as carbon, but where, because of the larger unscreened charge on the nucleus, the electrons are moving more relativistically, enhancing the spin-orbit effect. However, a classic example of the transition from LS to jj coupling is seen in the series C-Si-Ge-Sn-Pb in the excited states $(2p)(3s), (3p)(4s), \dots, (6p)(7s)$ (see Fig. 8.8). Here, the electrons are not in the same subshell, so their wavefunctions overlap less, and the Coulomb repulsion is reduced compared to the spin-orbit interaction. Analysing this situation in the LS coupling approximation, one expects a triplet and a singlet:

$$^3P_{0,1,2} \quad ^1P_1, \quad (8.153)$$

while in the jj scheme one expects two doublets:

$$(1/2, 1/2)_{0,1} \quad (1/2, 3/2)_{2,1}. \quad (8.154)$$

Experimentally, C and Si conform to the LS expectation and Pb to the jj scheme, while Ge and Sn are intermediate.

8.3.4 Zeeman Effect

For a multi-electron atom in a weak magnetic field, the appropriate unperturbed states are given by $|J, M_J, L, S\rangle$, where J, L, S refer to the total angular momenta. To determine the Zeeman energy shift, we need to determine the matrix element of \hat{S}_z . To do so, we can make use of the following argument. Since $2\hat{\mathbf{L}} \cdot \hat{\mathbf{S}} = \hat{\mathbf{J}}^2 - \hat{\mathbf{L}}^2 - \hat{\mathbf{S}}^2$, this operator is diagonal in the basis of states, $|J, M_J, L, S\rangle$. Therefore, the matrix element of the operator (recall that $[\hat{S}_i, \hat{S}_j] = i\hbar\epsilon_{ijk}\hat{S}_k$ and $[\hat{L}_i, \hat{S}_k] = 0$),

$$-i\hbar\hat{\mathbf{S}} \times \hat{\mathbf{L}} \equiv \hat{\mathbf{S}}(\hat{\mathbf{L}} \cdot \hat{\mathbf{S}}) - (\hat{\mathbf{L}} \cdot \hat{\mathbf{S}})\hat{\mathbf{S}} \quad (8.155)$$

must vanish. Moreover, from the identity $[\hat{\mathbf{L}} \cdot \hat{\mathbf{S}}, \hat{\mathbf{J}}] = 0$, it follows that the matrix element of the vector product,

$$-i\hbar(\hat{\mathbf{S}} \times \hat{\mathbf{L}}) \times \hat{\mathbf{J}} = \hat{\mathbf{S}} \times \hat{\mathbf{J}}(\hat{\mathbf{L}} \cdot \hat{\mathbf{S}}) - (\hat{\mathbf{L}} \cdot \hat{\mathbf{S}})\hat{\mathbf{S}} \times \hat{\mathbf{J}}, \quad (8.156)$$

must also vanish. If we expand the left hand side, we thus find that the matrix element of

$$(\hat{\mathbf{S}} \times \hat{\mathbf{L}}) \times \hat{\mathbf{J}} = \hat{\mathbf{L}}(\hat{\mathbf{S}} \cdot \hat{\mathbf{J}}) - \hat{\mathbf{S}}(\hat{\mathbf{L}} \cdot \hat{\mathbf{J}}) \stackrel{\hat{\mathbf{L}}=\hat{\mathbf{J}}-\hat{\mathbf{S}}}{=} \hat{\mathbf{J}}(\hat{\mathbf{S}} \cdot \hat{\mathbf{J}}) - \hat{\mathbf{S}}\hat{\mathbf{J}}^2, \quad (8.157)$$

also vanishes. Therefore, it follows that $\langle \hat{\mathbf{S}}\hat{\mathbf{J}}^2 \rangle = \langle \hat{\mathbf{J}}(\hat{\mathbf{S}} \cdot \hat{\mathbf{J}}) \rangle$, where the expectation value is taken over the basis states. Then, with $\hat{\mathbf{S}} \cdot \hat{\mathbf{J}} = \frac{1}{2}(\hat{\mathbf{J}}^2 + \hat{\mathbf{S}}^2 - \hat{\mathbf{L}}^2)$, we have that

$$\langle \hat{S}_z \rangle = \langle \hat{J}_z \rangle \frac{J(J+1) + S(S+1) - L(L+1)}{2J(J+1)}. \quad (8.158)$$

As a result, we can deduce that, at first order in perturbation theory, the energy shift arising from the Zeeman term is given by

$$\Delta E_{J,M_J,L,S} = \mu_B g_J M_J B, \quad (8.159)$$

where

$$g_J = \frac{J(J+1) + S(S+1) - L(L+1)}{2J(J+1)}, \quad (8.160)$$

denotes the effective **Landé g-factor**, which lies between 1 and 2. Note that, in the special case of hydrogen, where $S = 1/2$ and $J = L \pm 1/2$, we recover our previous result. The predicted Zeeman splitting for sodium is shown in figure 8.10.

In the **strong field limit**, where the influence of Zeeman term dominates, the appropriate basis states are set by $|L, M_L, A, M_S\rangle$, in which the operators $\hat{\mathbf{L}}^2$, \hat{L}_z , $\hat{\mathbf{S}}^2$, \hat{S}_z , and \hat{H}_{Zeeman} are diagonal. In this case, the energy splitting takes the form

$$\Delta E_{J,M_J,L,S} = \mu_B B(M_L + 2M_S) + \frac{1}{2}mc^2 \left\{ \frac{Z\alpha}{n} \right\}^4 \frac{nM_L M_S}{\ell(\ell+1/2)(\ell+1)}, \quad (8.161)$$

where the second term arises from the spin-orbit interaction.

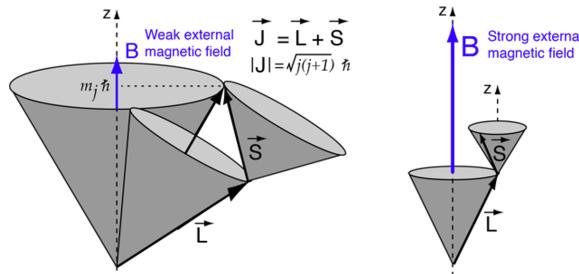


Fig. 8.9: In the weak field case, the vector model (left) implies that the coupling of the orbital angular momentum L to the spin angular momentum S is stronger than their coupling to the external field. In this case where spin-orbit coupling is dominant, they can be visualised as combining to form a total angular momentum J which then precesses about the magnetic field direction. In the strong field case (right), S and L couple more strongly to the external magnetic field than to each other, and can be visualised as independently precessing about the external field direction.

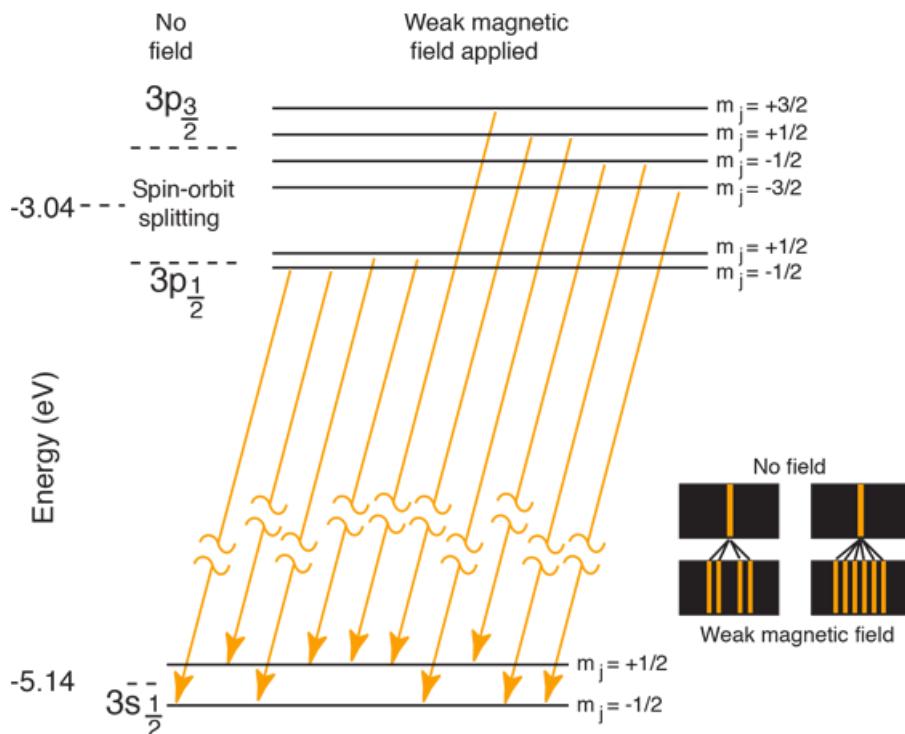


Fig. 8.10: The well known doublet which is responsible for the bright yellow light from a sodium lamp may be used to demonstrate several of the influences which cause splitting of the emission lines of atomic spectra. The transition which gives rise to the doublet is from the $3p$ to the $3s$ level. The fact that the $3s$ state is lower than the $3p$ state is a good example of the dependence of atomic energy levels on orbital angular momentum. The $3s$ electron penetrates the $1s$ shell more and is less effectively shielded than the $3p$ electron, so the $3s$ level is lower. The fact that there is a doublet shows the smaller dependence of the atomic energy levels on the total angular momentum. The $3p$ level is split into states with total angular momentum $J = 3/2$ and $J = 1/2$ by the spin-orbit interaction. In the presence of an external magnetic field, these levels are further split by the magnetic dipole energy, showing dependence of the energies on the z -component of the total angular momentum.

CHAPTER 9

Symmetry in Quantum Mechanics

Symmetry considerations are very important in quantum theory. The structure of eigenstates and the spectrum of energy levels of a quantum system reflect the symmetry of its Hamiltonian. As we will see later, the transition probabilities between different states under a perturbation, such as that imposed by an external electromagnetic field, depend in a crucial way on the transformation properties of the perturbation and lead to “selection rules”. Symmetries can be classified into two types, discrete and continuous, according to the transformations that generate them. For example, a mirror symmetry is an example of a discrete symmetry while a rotation in three-dimensional space is continuous. There is a close connection in physics between invariance under certain transformations and conservation laws

time translation	$t \rightarrow t + \tau$	energy conservation
space translation	$\mathbf{r} \rightarrow \mathbf{r} + \mathbf{a}$	momentum conservation
spatial rotation	$r_i = \sum_j R_{ij} r_j$	angular momentum conservation
spatial inversion	$\mathbf{r} \rightarrow -\mathbf{r}$	parity conservation

In addition, in quantum mechanics the structure and degeneracy of the eigenstates and eigenvalues will reflect the symmetries of the Hamiltonian.

Formally, the symmetries of a quantum system can be represented by a group of unitary transformations (or operators), \hat{U} , that act in the Hilbert space.¹ Under the action of such a unitary transformation, operators corresponding to observables \hat{A} of the quantum model will then transform as,

$$\hat{A} \rightarrow \hat{U}^\dagger \hat{A} \hat{U}. \quad (9.1)$$

For unitary transformations, we have seen that $\hat{U}^\dagger \hat{U} = \hat{\mathbb{I}}$, i.e. $\hat{U}^\dagger = \hat{U}^{-1}$. Under what circumstances does such a group of transformations represent a **symmetry group**? Consider a Schrödinger particle in three dimensions:² The basic observables are the position and momentum vectors, $\hat{\mathbf{r}}$ and $\hat{\mathbf{p}}$. We can always define a transformation of the coordinate system, or the observables, such that a vector $\hat{\mathbf{A}} = \hat{\mathbf{r}}$ or $\hat{\mathbf{p}}$ is mapped to $\mathbf{R}[\hat{\mathbf{A}}]$.³ If \mathbf{R} is an element of the group of transformations, then this transformation will be represented by

¹In quantum mechanics, the possible states of a system can be represented by unit vectors (called “state vectors”) residing in “state space” known as the Hilbert space. The precise nature of the Hilbert space is dependent on the system; for example, the state space for position and momentum states is the space of square-integrable functions (See Section 3.1).

²In the following, we will focus our considerations on the realm of “low-energy” physics where the relevant space-time transformations belong to the Galilei group.

³e.g., for a clockwise spatial rotation by an angle θ around \hat{e}_z , we have,

$$\mathbf{R}[\mathbf{r}] = R_{ij} \hat{x}_j, \quad \begin{pmatrix} \cos \theta & \sin \theta & 0 \\ -\sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad (9.2)$$

Similarly, for a spatial translation by a vector \mathbf{a} , $\mathbf{R}[\mathbf{r}] = \mathbf{r} + \mathbf{a}$.

a unitary operator $\hat{U}(\mathbf{R})$, such that

$$\hat{U}^\dagger \hat{\mathbf{A}} \hat{U} = \mathbf{R}[\hat{\mathbf{A}}]. \quad (9.3)$$

Such unitary transformations are said to be **symmetries of a general operator** $\hat{O}(\hat{\mathbf{p}}, \hat{\mathbf{r}})$ if

$$\hat{U}^\dagger \hat{O} \hat{U} = \hat{O}, \quad \text{i.e. } [\hat{O}, \hat{U}] = 0. \quad (9.4)$$

If $\hat{O}(\hat{\mathbf{p}}, \hat{\mathbf{r}}) \equiv \hat{H}$, the quantum Hamiltonian, such unitary transformations are said to be symmetries of the quantum system.

9.1 Observables as Generators of Transformations

The vector operators $\hat{\mathbf{p}}$ and $\hat{\mathbf{r}}$ for a Schrödinger particle are themselves generators of space-time transformations. From the standard commutation relations one can show that, for a constant vector \mathbf{a} , the unitary operator

$$\hat{U}(\mathbf{a}) = \exp \left[-\frac{i}{\hbar} \mathbf{a} \cdot \hat{\mathbf{p}} \right], \quad (9.5)$$

acting in the Hilbert space of a Schrödinger particle performs a **spatial translation**, $\hat{U}^\dagger(\mathbf{a}) f(\mathbf{r}) \hat{U}(\mathbf{a}) = f(\mathbf{r} + \mathbf{a})$, where $f((\mathbf{r}))$ denotes a general algebraic function of \mathbf{r} .

The proof runs as follows: With $\hat{\mathbf{p}} = -i\hbar \nabla$,

$$\hat{U}^\dagger(\mathbf{a}) = e^{\mathbf{a} \cdot \nabla} = \sum_{n=0}^{\infty} \frac{1}{n!} a_{i_1} \cdots a_{i_n} \nabla_{i_1} \cdots \nabla_{i_n}, \quad (9.6)$$

where summation on the repeated indices, i_m is assumed. Then, making use of the **Baker-Hausdorff identity**

$$e^{\hat{A}} \hat{B} e^{-\hat{A}} = \hat{B} + [\hat{A}, \hat{B}] + \frac{1}{2!} [\hat{A}, [\hat{A}, \hat{B}]] + \cdots, \quad (9.7)$$

it follows that

$$\hat{U}^\dagger(\mathbf{a}) f(\mathbf{r}) \hat{U}(\mathbf{a}) = f(\mathbf{a}) + a_{i_1} (\nabla_{i_1} f(\mathbf{r})) + \frac{1}{2!} a_{i_1} a_{i_2} (\nabla_{i_1} \nabla_{i_2} f(\mathbf{r})) + \cdots = f(\mathbf{r} + \mathbf{a}), \quad (9.8)$$

where the last identity follows from the Taylor expansion.

Therefore, a quantum system has spatial translation as an invariance group if and only if the following condition holds,

$$\hat{U}(\mathbf{a}) \hat{H} = \hat{H} \hat{U}(\mathbf{a}), \quad \text{i.e. } \hat{\mathbf{p}} \hat{H} = \hat{H} \hat{\mathbf{p}} \quad (9.9)$$

This demands that the Hamiltonian is independent of position, $\hat{H} = \hat{H}(\hat{\mathbf{p}})$, as one might have expected! Similarly, the group of unitary transformations, $\hat{U}(\mathbf{B}) = \exp \left[-\frac{i}{\hbar} \mathbf{B} \cdot \hat{\mathbf{r}} \right]$, performs **translations in momentum space**. Moreover, **spatial rotations** are generated by the transformation $\hat{U} = \exp \left[-\frac{i}{\hbar} \theta \mathbf{e}_n \cdot \hat{\mathbf{L}} \right]$, where $\hat{\mathbf{L}} = \hat{\mathbf{r}} \times \hat{\mathbf{p}}$ denotes the angular momentum operator. For more details see Appendix A.2.

As we have seen, **time translations** are generated by the time evolution operator, $\hat{U}(t) = \exp\left[-\frac{i}{\hbar}\hat{H}t\right]$. Therefore, every observable which commutes with the Hamiltonian is a constant of the motion (invariant under time translations),

$$\hat{H}\hat{A} = \hat{A}\hat{H} \implies e^{i\hat{H}t/\hbar}\hat{A}e^{-i\hat{H}t/\hbar} = \hat{A}, \quad \forall t. \quad (9.10)$$

We now turn to consider some examples of discrete symmetries. Amongst these, perhaps the most important in low-energy physics are parity and time-reversal. The **parity** operation, denoted \hat{P} , involves a reversal of sign on all coordinates.

$$\hat{P}\psi(\mathbf{r}) = \psi(-\mathbf{r}). \quad (9.11)$$

This is clearly a discrete transformation. Application of parity twice returns the initial state implying that $\hat{P}^2 = 1$. Therefore, the eigenvalues of the parity operation (if such exist) are ± 1 . A wavefunction will have a defined parity if and only if it is an even or odd function. For example, for $\psi(x) = \cos(x)$, $\hat{P}\psi = \cos(-x) = \cos(x) = \psi$; thus ψ is even and $P = 1$. Similarly $\psi = \sin(x)$ is odd with $P = -1$. Later, in the next chapter, we will encounter the spherical harmonic functions which have the following important symmetry under parity, $\hat{P}Y_{\ell m} = (-1)^{\ell}Y_{\ell m}$. Parity will be conserved if the Hamiltonian is invariant under the parity operation, i.e. if the Hamiltonian is invariant under a reversal of sign of all the coordinates.⁴

In classical mechanics, the **time-reversal** operation involves simply “running the movie backwards”. The time-reversed state of the phase space coordinates $(x(t), p(t))$ is defined by $(x_T(t), p_T(t))$ where $x_T(t) = x(t)$ and $p_T(t) = -p(t)$. Hence, if the system evolved from $(x(0), p(0))$ to $(x(t), p(t))$ in time t and at t we reverse the velocity, $p(t) \rightarrow -p(t)$ with $x(t) \rightarrow x(t)$, at time $2t$ the system would have returned to $x(2t) = x(0)$ while $p(2t) = -p(0)$. If this happens, we say that the system is time-reversal invariant. Of course, this is just the statement that Newton’s laws are the same if $t \rightarrow -t$. A notable case where this is not true is that of a charged particle in a magnetic field.

As with classical mechanics, time-reversal in quantum mechanics involves the operation $t \rightarrow -t$. However, referring to the time-dependent Schrödinger equation, $i\hbar\partial_t\psi(x, t) = \hat{H}\psi(x, t)$, we can see that the operation $t \rightarrow -t$ is equivalent to complex conjugation of the wavefunction, $\psi \rightarrow \psi^*$ if $\hat{H}^* = \hat{H}$. Let us then consider the time-evolution of $\psi(x, t)$,

$$\psi(x, 0) \xrightarrow{\text{c.c.}} e^{+\frac{i}{\hbar}\hat{H}^*t}\psi^*(x, 0) \xrightarrow{\text{evolve}} e^{-\frac{i}{\hbar}\hat{H}t}e^{+\frac{i}{\hbar}\hat{H}^*t}\psi^*(x, 0). \quad (9.12)$$

If we require that $\psi(x, 2t) = \psi^*(x, 0)$, we must have $\hat{H}^* = \hat{H}(x)$. Therefore, \hat{H} is invariant under time-reversal if and only if \hat{H} is real.

Although the group of space-transformations covers the symmetries that pertain to “low-energy” quantum physics, such as atomic physics, quantum optics, and quantum chemistry, in nuclear physics and elementary particle physics new observables come into play (e.g. the isospin quantum numbers and the other quark charges in the standard model). They generate symmetry groups which lack a classical counterpart, and they do not have any obvious relation with space-time transformations. These symmetries are often called **internal symmetries** in order to underline this fact.

⁴In high energy physics, parity is a symmetry of the strong and electromagnetic forces, but does not hold for the weak force. Therefore, parity is conserved in strong and electromagnetic interactions, but is violated in weak interactions.

9.2 Consequences of Symmetries: Multiplets

Having established how to identify whether an operator belongs to a group of symmetry transformations, we now consider the consequences. Consider a single unitary transformation \hat{U} in the Hilbert space, and an observable \hat{A} which commutes with \hat{U} , $[\hat{U}, \hat{A}] = 0$. If \hat{A} has an eigenvector $|a\rangle$, it follows that $\hat{U}|a\rangle$ will be an eigenvector with the same eigenvalue, i.e.

$$\hat{U}\hat{A}|a\rangle = \hat{A}\hat{U}|a\rangle = a\hat{U}|a\rangle. \quad (9.13)$$

This means that either:

1. $|a\rangle$ is an eigenvector of both \hat{A} and \hat{U} , or
2. the eigenvalue a is degenerate: the linear space spanned by the vectors $\hat{U}^n|a\rangle$ (n integer) are eigenvectors with the same eigenvalue.

This mathematical argument leads to the conclusion that, given a group G of unitary operators $\hat{U}(x)$, $x \in G$, for any observable which is invariant under these transformations, i.e.

$$[\hat{U}(x), \hat{A}] = 0, \quad \forall x \in G, \quad (9.14)$$

its discrete eigenvalues and eigenvectors will show a characteristic multiplet structure: there will be a degeneracy due to the symmetry such that the eigenvectors belonging to each eigenvalue form an invariant subspace under the group of transformations.

For example, if the Hamiltonian commutes with the angular momentum operators, \hat{L}_i , $i = x, y, z$, i.e. it is invariant under three-dimensional rotations, an energy level with a given orbital quantum number ℓ is at least $(2\ell + 1)$ -fold degenerate. Such a degeneracy can be seen as the result of non-trivial actions of the operator \hat{L}_x and \hat{L}_y on an energy (and \hat{L}_z) eigenstate $|E, \ell, m\rangle$ (where m is the magnetic quantum number associated with \hat{L}_z).

9.3 Rotational Symmetry in Quantum Mechanics

Under a rotation R , the states of a quantum system transform as

$$|\psi\rangle \rightarrow |\psi'\rangle = \hat{U}(R)|\psi\rangle. \quad (9.15)$$

If all measurement outcomes remain invariant under spatial rotations, then the operator $\hat{U}(R)$ is a unitary operator:

$$\hat{U}(R)^\dagger \hat{U}(R) = \hat{\mathbb{I}}. \quad (9.16)$$

For an infinitesimal rotation described by a vector ω , it must be possible to write the unitary operator $\hat{U}(R)$ in the form

$$\hat{U}(\omega) = \hat{\mathbb{I}} + \frac{i}{\hbar}(\omega \cdot \hat{\mathbf{J}}) + \mathcal{O}(|\omega|^2), \quad (9.17)$$

where the generators $\hat{\mathbf{J}} = (\hat{J}_1, \hat{J}_2, \hat{J}_3)$ are a vector of Hermitian operators. The factor of $\frac{1}{\hbar}$ is introduced in order that the generators have dimension of angular momentum. Since rotations do not in general commute with each other, the generator components \hat{J}_i will not be mutually commuting (unlike the generators \hat{P}_i of spatial translations).

Since two successive rotations, R_1 followed by R_2 , are equivalent to a single rotation described by the matrix product R_2R_1 , we must have

$$\hat{U}(R_2)\hat{U}(R_1) = \hat{U}(R_2R_1). \quad (9.18)$$

To obtain the commutation relations satisfied by the components \hat{J}_i , we consider a sequence of three infinitesimal rotations:

$$R_1 = \mathbb{I} + \omega' + \dots; \quad R_2 = \mathbb{I} + \omega + \dots; \quad R_3 = (R_1)^{-1} = \mathbb{I} - \omega' + \dots. \quad (9.19)$$

where we have used $(R_1)^{-1} = (R_1)^T$ and $(\omega')^T = -\omega'$. This corresponds to an overall rotation given by

$$R_3R_2R_1 = (\mathbb{I} - \omega' + \dots)(\mathbb{I} + \omega + \dots)(\mathbb{I} + \omega' + \dots) = \mathbb{I} + \omega + \omega\omega' - \omega'\omega + \dots, \quad (9.20)$$

where we have kept the quadratic terms of order $\omega\omega'$ and $\omega'\omega$ as these will be needed in what follows. We thus have

$$R_3R_2R_1 = \mathbb{I} + \omega + \Omega + \dots, \quad (9.21)$$

where

$$\Omega \equiv \omega\omega' - \omega'\omega. \quad (9.22)$$

Explicit multiplication gives the matrix product $\omega\omega'$ as

$$\omega\omega' = \begin{pmatrix} 0 & \omega_3 & -\omega_2 \\ -\omega_3 & 0 & \omega_1 \\ \omega_2 & -\omega_1 & 0 \end{pmatrix} \begin{pmatrix} 0 & \omega'_3 & -\omega'_2 \\ -\omega'_3 & 0 & \omega'_1 \\ \omega'_2 & -\omega'_1 & 0 \end{pmatrix} \quad (9.23)$$

$$= \begin{pmatrix} 0 & \omega_2\omega'_1 - \omega_1\omega'_2 & \omega_3\omega'_1 - \omega_1\omega'_3 \\ \omega_1\omega'_2 - \omega_2\omega'_1 & 0 & \omega_3\omega'_2 - \omega_2\omega'_3 \\ \omega_1\omega'_3 - \omega_3\omega'_1 & \omega_2\omega'_3 - \omega_3\omega'_2 & 0 \end{pmatrix} \quad (9.24)$$

Thus the matrix Ω corresponds to an infinitesimal rotation,

$$\Omega = \begin{pmatrix} 0 & \Omega_3 & -\Omega_2 \\ -\Omega_3 & 0 & \Omega_1 \\ \Omega_2 & -\Omega_1 & 0 \end{pmatrix}, \quad (9.25)$$

with components

$$\Omega_1 = \omega_3\omega'_2 - \omega_2\omega'_3, \quad \Omega_2 = \omega_1\omega'_3 - \omega_3\omega'_1, \quad \Omega_3 = \omega_2\omega'_1 - \omega_1\omega'_2. \quad (9.26)$$

The equation $\hat{U}(R_3)\hat{U}(R_2)\hat{U}(R_1) = \hat{U}(R_3R_2R_1)$ then gives

$$\left(\hat{\mathbb{I}} - \frac{i}{\hbar}(\omega' \cdot \hat{\mathbf{J}})\right)\left(\hat{\mathbb{I}} + \frac{i}{\hbar}(\omega \cdot \hat{\mathbf{J}})\right)\left(\hat{\mathbb{I}} + \frac{i}{\hbar}(\omega' \cdot \hat{\mathbf{J}})\right) = \hat{\mathbb{I}} + \frac{i}{\hbar}(\omega \cdot \hat{\mathbf{J}}) + \frac{i}{\hbar}(\Omega \cdot \hat{\mathbf{J}}) + \dots, \quad (9.27)$$

where $\Omega = (\Omega_1, \Omega_2, \Omega_3)$. Multiplying out the left-hand side above then gives

$$(\omega' \cdot \hat{\mathbf{J}})(\omega \cdot \hat{\mathbf{J}}) - (\omega \cdot \hat{\mathbf{J}})(\omega' \cdot \hat{\mathbf{J}}) = i\hbar(\Omega \cdot \hat{\mathbf{J}}). \quad (9.28)$$

Writing the scalar products in Eq. (9.28) explicitly in terms of the vector components, and equating the coefficients of $\omega_1\omega'_2$ on both sides, gives

$$\hat{J}_1\hat{J}_2 - \hat{J}_2\hat{J}_1 = i\hbar\hat{J}_3. \quad (9.29)$$

Similarly, equating the coefficients of $\omega_2\omega'_3$ gives

$$\hat{J}_2\hat{J}_3 - \hat{J}_3\hat{J}_2 = i\hbar\hat{J}_1, \quad (9.30)$$

and equating the coefficients of $\omega_3\omega'_1$ gives

$$\hat{J}_3\hat{J}_1 - \hat{J}_1\hat{J}_3 = i\hbar\hat{J}_2. \quad (9.31)$$

Overall, we obtain the commutation relations

$$[\hat{J}_i, \hat{J}_j] = i\hbar\epsilon_{ijk}\hat{J}_k.$$

(9.32)

The vector operator $\hat{\mathbf{J}}$ therefore satisfies the standard commutation relations for angular momentum. Since $\hat{\mathbf{J}}$ is a property of the system as a whole, rather than of a particular particle, we identify $\hat{\mathbf{J}}$ as the *total* angular momentum operator for the system.

Consider for example, the rotation of a system through an angle ϕ_0 about the z -axis,

$$\psi(\phi) \rightarrow \psi'(\phi) = \hat{U}(\phi_0)\psi(\phi) = \psi(\phi + \phi_0). \quad (9.33)$$

For the case of an infinitesimal rotation, this gives

$$\psi(\phi + \phi_0) = \psi(\phi) + \phi_0 \frac{\partial\psi(\phi)}{\partial\phi} + \dots = \left(\hat{\mathbb{I}} + \phi_0 \frac{\partial}{\partial\phi} + \dots \right) \psi(\phi). \quad (9.34)$$

The vector $\boldsymbol{\omega}$ corresponding to this rotation is

$$\boldsymbol{\omega} = (\omega_1, \omega_2, \omega_3) = (0, 0, \phi_0). \quad (9.35)$$

Hence ,the symmetry operator is

$$\hat{U}(\phi_0) = \hat{\mathbb{I}} + \frac{i}{\hbar} \boldsymbol{\omega} \cdot \hat{\mathbf{J}} + \dots = \hat{\mathbb{I}} + \frac{i}{\hbar} \phi_0 \hat{J}_z + \dots, \quad (9.36)$$

and thus we identify

$$\hat{J}_z = -i\hbar \frac{\partial}{\partial\phi}. \quad (9.37)$$

9.3.1 Scalar Operators

A *scalar operator* is an operator whose matrix elements are invariant under rotations,

$$\langle \psi' | \hat{K} | \phi' \rangle = \langle \psi | \hat{K} | \phi \rangle, \quad (9.38)$$

for all possible states of the system, $|\psi\rangle$ and $|\phi\rangle$. Since $|\psi'\rangle = \hat{U}(R)|\psi\rangle$ and $|\phi'\rangle = \hat{U}(R)|\phi\rangle$, the left-hand side is

$$\langle \psi | \hat{U}^\dagger(R) \hat{K} \hat{U}(R) | \psi \rangle = \langle \psi' | \hat{K} | \phi' \rangle = \langle \psi | \hat{U}(R)^\dagger \hat{K} \hat{U}(R) | \phi \rangle. \quad (9.39)$$

For Eq. (9.38) to hold for all possible states $|\psi\rangle$ and $|\phi\rangle$ then requires that \hat{K} satisfy the relation

$$\hat{U}(R)^\dagger \hat{K} \hat{U}(R) = \hat{K} \quad (9.40)$$

for all rotations R . Since $\hat{U}(R)$ is unitary, this is equivalent to the condition that \hat{K} commutes with $\hat{U}(R)$ for all rotations R :

$$[\hat{U}(R), \hat{K}] = 0. \quad (9.41)$$

Applying this result to the case of infinitesimal rotations, with $\hat{U}(R)$ of the form given in Eq. (9.17), then shows that \hat{K} must commute with all components of the total angular momentum operator, $\hat{\mathbf{J}}$:

$$[\hat{J}_i, \hat{K}] = 0; \quad [\hat{\mathbf{J}}, \hat{K}] = 0. \quad (9.42)$$

From this, it follows immediately that \hat{K} also commutes with $\hat{\mathbf{J}}^2$ and with the ladder operators $\hat{J}_\pm = \hat{J}_1 \pm i\hat{J}_2$:

$$[\hat{\mathbf{J}}^2, \hat{K}] = 0, \quad [\hat{J}_\pm, \hat{K}] = 0. \quad (9.43)$$

Eq. (9.42) can also serve as the *definition* of a scalar operator \hat{K} with respect to an angular momentum operator $\hat{\mathbf{J}}$.

For an *isolated system*, the Hamiltonian \hat{H} must be invariant under spatial rotations, i.e. is must be a scalar operator

$$\hat{U}^\dagger(R) \hat{H} \hat{U}(R) = \hat{H} \quad \Rightarrow \quad [\hat{H}, \hat{\mathbf{J}}] = 0. \quad (9.44)$$

Hence, from Ehrenfest's theorem,

$$\frac{d}{dt} \langle \psi | \hat{H} | \psi \rangle = 0. \quad (9.45)$$

Conservation of total angular momentum for an isolated system is a consequence of invariance under spatial rotations.

9.3.2 Vector Operators

An operator

$$\hat{\mathbf{V}} = (\hat{V}_x, \hat{V}_y, \hat{V}_z) = (\hat{V}_1, \hat{V}_2, \hat{V}_3) \quad (9.46)$$

whose expectation values transform under rotations like an ordinary vector,

$$\langle \psi | \hat{V}_i | \psi \rangle \rightarrow \langle \psi' | \hat{V}_i | \psi' \rangle = \sum_j R_{ij} \langle \psi | \hat{V}_j | \psi \rangle, \quad (9.47)$$

is said to be a *vector operator*. The left-hand side of Eq. (9.47) is

$$\langle \psi' | \hat{V}_i | \psi' \rangle = \langle \psi | \hat{U}(R)^\dagger \hat{V}_i \hat{U}(R) | \psi \rangle \quad (9.48)$$

For Eq. (9.47) to hold for all possible states $|\psi\rangle_i$, the operator $\hat{\mathbf{V}}$ must satisfy the relation

$$\hat{U}(R)^\dagger \hat{V}_i \hat{U}(R) = \sum_j R_{ij} \hat{V}_j \quad (9.49)$$

for all possible rotations R .

For the case that R is an infinitesimal rotation, $R = \mathbb{I} + \omega$, Eq. (9.49) gives

$$\left(\mathbb{I} - \frac{i}{\hbar}(\omega \cdot \hat{\mathbf{J}}) + \dots\right)\hat{V}_i \left(\mathbb{I} + \frac{i}{\hbar}(\omega \cdot \hat{\mathbf{J}}) + \dots\right) = \hat{V}_i + \sum_j \omega_{ij} \hat{V}_j + \dots . \quad (9.50)$$

Expanding the left-hand side up to first-order in ω then gives

$$\frac{i}{\hbar} \hat{V}_i (\omega \cdot \hat{\mathbf{J}}) - \frac{i}{\hbar} (\omega \cdot \hat{\mathbf{J}}) \hat{V}_i = \sum_j \omega_{ij} \hat{V}_j \quad (i = 1, 2, 3). \quad (9.51)$$

Written out explicitly, the right-hand side of the above equation is

$$\sum_j \omega_{ij} \hat{V}_j = \begin{pmatrix} 0 & \omega_3 & -\omega_2 \\ -\omega_3 & 0 & \omega_1 \\ \omega_2 & -\omega_1 & 0 \end{pmatrix} \begin{pmatrix} \hat{V}_1 \\ \hat{V}_2 \\ \hat{V}_3 \end{pmatrix} = \begin{pmatrix} \omega_3 \hat{V}_2 - \omega_2 \hat{V}_3 \\ -\omega_3 \hat{V}_1 + \omega_1 \hat{V}_3 \\ \omega_2 \hat{V}_1 - \omega_1 \hat{V}_2 \end{pmatrix}. \quad (9.52)$$

Hence, for $i = 1$, Eq. (9.51) gives

$$\hat{V}_1 (\omega \cdot \hat{\mathbf{J}}) - (\omega \cdot \hat{\mathbf{J}}) \hat{V}_1 = -i\hbar (\omega_3 \hat{V}_2 - \omega_2 \hat{V}_3). \quad (9.53)$$

Expanding $\omega \cdot \hat{\mathbf{J}} = \omega_1 \hat{J}_1 + \omega_2 \hat{J}_2 + \omega_3 \hat{J}_3$ on the left-hand side, and rearranging, we obtain

$$\omega_1 (\hat{V}_1 \hat{J}_1 - \hat{J}_1 \hat{V}_1) + \omega_2 (\hat{V}_1 \hat{J}_2 - \hat{J}_2 \hat{V}_1) + \omega_3 (\hat{V}_1 \hat{J}_3 - \hat{J}_3 \hat{V}_1) = -i\hbar (\omega_3 \hat{V}_2 - \omega_2 \hat{V}_3). \quad (9.54)$$

This relation must hold for all possible choices of ω . Equating the coefficients of ω_1 , ω_2 and ω_3 on both sides gives

$$\hat{V}_1 \hat{J}_1 - \hat{J}_1 \hat{V}_1 = 0, \quad \hat{V}_1 \hat{J}_2 - \hat{J}_2 \hat{V}_1 = i\hbar \hat{V}_3, \quad \hat{V}_1 \hat{J}_3 - \hat{J}_3 \hat{V}_1 = -i\hbar \hat{V}_2. \quad (9.55)$$

Repeating this process for $i = 2$ and $i = 3$ gives, similarly,

$$\hat{V}_2 \hat{J}_2 - \hat{J}_2 \hat{V}_2 = 0, \quad \hat{V}_2 \hat{J}_3 - \hat{J}_3 \hat{V}_2 = i\hbar \hat{V}_1, \quad \hat{V}_2 \hat{J}_1 - \hat{J}_1 \hat{V}_2 = -i\hbar \hat{V}_3, \quad (9.56)$$

$$\hat{V}_3 \hat{J}_2 - \hat{J}_2 \hat{V}_3 = 0, \quad \hat{V}_3 \hat{J}_3 - \hat{J}_3 \hat{V}_3 = i\hbar \hat{V}_1, \quad \hat{V}_3 \hat{J}_1 - \hat{J}_1 \hat{V}_3 = -i\hbar \hat{V}_2. \quad (9.57)$$

Altogether, for *any* vector operator \hat{V} , the equations above can be written compactly as

$$[\hat{J}_i, \hat{V}_j] = i\hbar \epsilon_{ijk} \hat{V}_k$$

(9.58)

Equivalently, using the antisymmetry properties of ϵ_{ijk} , this is

$$[\hat{V}_i, \hat{J}_j] = i\hbar \epsilon_{ijk} \hat{V}_k \quad (9.59)$$

Explicitly, the non-zero commutation relations between the components of the operators $\hat{\mathbf{J}}$ and $\hat{\mathbf{V}}$ can be summarised as

$$\begin{aligned} [\hat{J}_1, \hat{V}_2] &= i\hbar \hat{V}_3, & [\hat{J}_2, \hat{V}_3] &= i\hbar \hat{V}_1, & [\hat{J}_3, \hat{V}_1] &= i\hbar \hat{V}_2, \\ [\hat{J}_1, \hat{V}_3] &= -i\hbar \hat{V}_2, & [\hat{J}_2, \hat{V}_1] &= -i\hbar \hat{V}_3, & [\hat{J}_3, \hat{V}_2] &= -i\hbar \hat{V}_1. \end{aligned} \quad (9.60)$$

Eq. (9.58) can also serve as the *definition* of a vector operator $\hat{\mathbf{V}}$ with respect to an angular momentum operator $\hat{\mathbf{J}}$.

9.3.2.1 Spherical Components of a Vector Operator

The *spherical components* \hat{V}_{+1} , \hat{V}_{-1} , and \hat{V}_0 of a vector operator are defined as

$$\boxed{\hat{V}_{+1} \equiv -\frac{1}{\sqrt{2}}(\hat{V}_1 + i\hat{V}_2), \quad \hat{V}_{-1} \equiv \frac{1}{\sqrt{2}}(\hat{V}_1 - i\hat{V}_2), \quad \hat{V}_0 \equiv \hat{V}_3.} \quad (9.61)$$

Note that the spherical component \hat{V}_{+1} is distinct from the Cartesian component $\hat{V}_1 = \hat{V}_x$ and from the ladder operator $\hat{V}_+ = \hat{V}_1 + i\hat{V}_2$.

The commutator of the spherical component \hat{V}_{+1} of $\hat{\mathbf{V}}$ with the component $\hat{J}_3 = \hat{J}_z$ of $\hat{\mathbf{J}}$ can be evaluated as

$$[\hat{J}_3, \hat{V}_{+1}] = -\frac{1}{\sqrt{2}}[\hat{J}_3, \hat{V}_1] - \frac{1}{\sqrt{2}}[\hat{J}_3, \hat{V}_2] = -\frac{1}{\sqrt{2}}(i\hbar\hat{V}_2) - \frac{i}{\sqrt{2}}(-i\hbar\hat{V}_1) = \hbar\hat{V}_{+1}. \quad (9.62)$$

For \hat{V}_{-1} we similarly obtain

$$[\hat{J}_3, \hat{V}_{-1}] = -\hbar\hat{V}_{-1}, \quad (9.63)$$

while for \hat{V}_0 we have simply

$$[\hat{J}_3, \hat{V}_0] = [\hat{J}_3, \hat{V}_3] = 0. \quad (9.64)$$

The above relations can be summarised as

$$\boxed{[\hat{J}_3, \hat{V}_m] = \hbar m \hat{V}_m, \quad m = \{0, \pm 1\}.} \quad (9.65)$$

The spherical component \hat{V}_{+1} commutes with the ladder operator $\hat{J}_+ = \hat{J}_1 + i\hat{J}_2$,

$$[\hat{J}_+, \hat{V}_{+1}] = -\frac{1}{\sqrt{2}}[\hat{J}_1 + i\hat{J}_2, \hat{V}_1 + i\hat{V}_2] = -\frac{1}{\sqrt{2}}(i(i\hbar\hat{V}_3) + i(-i\hbar\hat{V}_3)) = 0, \quad (9.66)$$

while for \hat{V}_{-1} and \hat{V}_0 , the commutators with \hat{J}_+ are

$$[\hat{J}_+, \hat{V}_{-1}] = \frac{1}{\sqrt{2}}[\hat{J}_1 + i\hat{J}_2, \hat{V}_1 - i\hat{V}_2] = \frac{1}{\sqrt{2}}(-i(i\hbar\hat{V}_3) + i(-i\hbar\hat{V}_3)) = \sqrt{2}\hbar\hat{V}_3 = \sqrt{2}\hbar\hat{V}_0, \quad (9.67)$$

and

$$[\hat{J}_+, \hat{V}_0] = [\hat{J}_1 + i\hat{J}_2, \hat{V}_3] = -i\hbar\hat{V}_2 + i(i\hbar\hat{V}_1) = \sqrt{2}\hbar\hat{V}_{+1}. \quad (9.68)$$

For the ladder operator $\hat{J}_- = \hat{J}_1 - i\hat{J}_2$, the equivalent expressions are

$$[\hat{J}_-, \hat{V}_{+1}] = \sqrt{2}\hbar\hat{V}_0, \quad [\hat{J}_-, \hat{V}_{-1}] = 0, \quad [\hat{J}_-, \hat{V}_0] = \sqrt{2}\hbar\hat{V}_{-1}. \quad (9.69)$$

All the above commutation relations involving the ladder operators \hat{J}_\pm can be summarised as

$$[\hat{J}_\pm, \hat{V}_m] = \hbar\sqrt{2 - m(m \pm 1)}\hat{V}_{m\pm 1}, \quad m = \{0, \pm 1\}, \quad (9.70)$$

or, equivalently, as

$$\boxed{[\hat{J}_\pm, \hat{V}_m] = \hbar\sqrt{j(j+1) - m(m \pm 1)}\hat{V}_{m\pm 1}, \quad j \equiv 1, m = \{0, \pm 1\}.} \quad (9.71)$$

9.3.2.2 Properties of Vector Operators

If the vector operator $\hat{\mathbf{V}}$ corresponds to an observable quantity, then its Cartesian components \hat{V}_i must all be Hermitian. It then follows directly from Eq. (9.61) that the spherical components \hat{V}_m of $\hat{\mathbf{V}}$ have Hermitian conjugates given by

$$\hat{V}_m^\dagger = (-1)^m \hat{V}_{-m}. \quad (9.72)$$

Inverting Eq. (9.61), the Cartesian components \hat{V}_i are given in terms of the spherical components \hat{V}_m as

$$\hat{V}_1 = \frac{1}{\sqrt{2}} (\hat{V}_{-1} - \hat{V}_{+1}), \quad \hat{V}_2 = \frac{i}{\sqrt{2}} (\hat{V}_{-1} + \hat{V}_{+1}), \quad \hat{V}_3 = \hat{V}_0. \quad (9.73)$$

The commutation relations between the scalar product $\hat{\mathbf{V}} \cdot \hat{\mathbf{W}} = \hat{V}_1 \hat{W}_1 + \hat{V}_2 \hat{W}_2 + \hat{V}_3 \hat{W}_3$ of two vector operators $\hat{\mathbf{V}}$ and $\hat{\mathbf{W}}$ and the component \hat{J}_1 of $\hat{\mathbf{J}}$ can then be evaluated as

$$[\hat{J}_1, \hat{V}_1 \hat{W}_1] = [\hat{J}_1, \hat{V}_1] \hat{W}_1 + \hat{V}_1 [\hat{J}_1, \hat{W}_1] = 0, \quad (9.74)$$

$$[\hat{J}_1, \hat{V}_2 \hat{W}_2] = [\hat{J}_1, \hat{V}_2] \hat{W}_2 + \hat{V}_2 [\hat{J}_1, \hat{W}_2] = i\hbar (\hat{V}_3 \hat{W}_2 + \hat{V}_2 \hat{W}_3), \quad (9.75)$$

$$[\hat{J}_1, \hat{V}_3 \hat{W}_3] = [\hat{J}_1, \hat{V}_3] \hat{W}_3 + \hat{V}_3 [\hat{J}_1, \hat{W}_3] = -i\hbar (\hat{V}_2 \hat{W}_3 + \hat{V}_3 \hat{W}_2). \quad (9.76)$$

Summing the above equations shows that \hat{J}_1 commutes with the scalar product $\hat{\mathbf{V}} \cdot \hat{\mathbf{W}}$:

$$[\hat{J}_1, \hat{\mathbf{V}} \cdot \hat{\mathbf{W}}] = 0. \quad (9.77)$$

Since the choice of the component \hat{J}_1 was arbitrary, this relation must be true for all components of $\hat{\mathbf{J}}$, not just \hat{J}_1 :

$$[\hat{\mathbf{J}}, \hat{\mathbf{V}} \cdot \hat{\mathbf{W}}] = 0. \quad (9.78)$$

Hence the scalar product $\hat{\mathbf{V}} \cdot \hat{\mathbf{W}}$ of two vector operators (for example, the spin-orbit coupling product $\hat{\mathbf{L}} \cdot \hat{\mathbf{S}}$) is a scalar operator. In particular, taking $\hat{\mathbf{W}} = \hat{\mathbf{V}}$, we also obtain

$$[\hat{\mathbf{J}}, \hat{\mathbf{V}}^2] = 0. \quad (9.79)$$

The scalar product $\hat{\mathbf{V}} \cdot \hat{\mathbf{W}}$ can be expressed in terms of spherical components as

$$\hat{\mathbf{V}} \cdot \hat{\mathbf{W}} = \hat{V}_1 \hat{W}_1 + \hat{V}_2 \hat{W}_2 + \hat{V}_3 \hat{W}_3 \quad (9.80)$$

$$= \frac{1}{2} (\hat{V}_{-1} - \hat{V}_{+1}) (\hat{W}_{-1} - \hat{W}_{+1}) - \frac{1}{2} (\hat{V}_{-1} + \hat{V}_{+1}) (\hat{W}_{-1} + \hat{W}_{+1}) + \hat{V}_0 \hat{W}_0 \quad (9.81)$$

$$= -\hat{V}_{+1} \hat{W}_{-1} - \hat{V}_{-1} \hat{W}_{+1} + \hat{V}_0 \hat{W}_0. \quad (9.82)$$

Similar arguments show that the vector product $\hat{\mathbf{V}} \times \hat{\mathbf{W}}$ of two vector operators $\hat{\mathbf{V}}$ and $\hat{\mathbf{W}}$ is also a vector operator.

9.4 The Wigner-Eckart Theorem for Scalar Operators

Let \hat{K} be a scalar operator with respect to an angular momentum operator $\hat{\mathbf{J}}$:

$$[\hat{\mathbf{J}}, \hat{K}] = 0. \quad (9.83)$$

Then the Wigner-Eckart theorem states that matrix elements of \hat{K} taken between total angular momentum eigenstates $|\alpha''j''m''\rangle$ and $|\alpha'j'm'\rangle$ must be of the form

$$\langle \alpha''j''m'' | \hat{K} | \alpha'j'm' \rangle = C(\alpha''\alpha'; j') \delta_{j''j'} \delta_{m''m'}, \quad (9.84)$$

where $C(\alpha''\alpha'; j')$ is a complex constant known as the reduced matrix element which is independent of the quantum numbers m'' and m' . The quantities α'' and α' collectively label all other quantum numbers needed to uniquely identify the angular momentum eigenstates involved.

For completeness, the proof of Eq. (9.84) is provided in the Appendix A.3.

The Wigner-Eckart theorem shows that matrix elements between total angular momentum eigenstates for rotationally invariant (scalar) operators do not depend on the quantum number m . This is intuitively reasonable; the angular momentum eigenstates $|jm\rangle$ are defined with respect to a particular choice of quantisation axis, conventionally taken to be the z -axis. A different choice of axis, the x -axis for example, would result in a different set of eigenstates, $|jm_x\rangle$, which are linear combinations of the conventional z -axis states $|jm\rangle$. For a rotationally invariant operator, all spatial directions are equivalent. Hence the matrix elements of \hat{K} cannot depend on a particular choice of quantisation axis; hence they cannot depend on the quantum number m (or m_x or \dots) whose definition depends on the choice of axis.

The reduced matrix element $C(\alpha''\alpha'; j')$ is usually written in Dirac-style notation as

$$C(\alpha''\alpha'; j') \equiv \langle \alpha''j'' | \hat{K} | \alpha'j' \rangle. \quad (9.85)$$

Despite its name, the reduced matrix element, $\langle \alpha''j'' | \hat{K} | \alpha'j' \rangle$, is *not* a matrix element; it is a common, constant component of a family of $(2j+1)(2j'+1)$ matrix elements.

The product of Kronecker- δ factors in Eq. (9.84) can be expressed as a single Clebsch-Gordan coefficient. To see this, consider the (trivial) angular momentum combination $0 \otimes j'' = j''$. The connection between the total (coupled) and uncoupled eigenstates is simply

$$|j''m''\rangle = |00\rangle \otimes |j''m''\rangle \equiv |00\rangle |j''m''\rangle. \quad (9.86)$$

Comparing with the expansion

$$|j''m''\rangle = \sum_{j'm'} |00\rangle |j'm'\rangle \langle 00; j'm' | j''m'' \rangle \quad (9.87)$$

then shows that the relevant Clebsch-Gordan coefficients are either zero or unity:

$$\langle 00; j'm' | j''m'' \rangle = \delta_{j''j'} \delta_{m''m'}. \quad (9.88)$$

Eq. (9.85) and (9.88) allow the Wigner-Eckart theorem for scalar operators, Eq. (9.84), to be written in the form

$$\langle \alpha''j''m'' | \hat{K} | \alpha'j'm' \rangle = \langle \alpha''j'' | \hat{K} | \alpha'j' \rangle \langle 00; j'm' | j''m'' \rangle. \quad (9.89)$$

The reason for writing the Wigner-Eckart theorem in this way will become clear once we have also considered the equivalent result for *vector* operators.

9.4.1 Consequences of the Wigner-Eckart Theorem (for Scalars)

Consider the Hamiltonian operator \hat{H} for an isolated system. Since there can be no preferred spatial direction for an isolated system, the expectation values of \hat{H} (the energy eigenvalues) cannot depend on the spatial orientation of the system. Hence \hat{H} must be a scalar operator, commuting with (all components of) the total angular momentum operator $\hat{\mathbf{J}}$:

$$[\hat{J}_i, \hat{H}] = 0, \quad [\hat{\mathbf{J}}, \hat{H}] = 0, \quad [\hat{\mathbf{J}}^2, \hat{H}] = 0. \quad (9.90)$$

Since \hat{H} and $\hat{\mathbf{J}}$ commute, it follows from Ehrenfest's theorem that the expectation values $\langle \hat{J}_i \rangle$ of each component of $\hat{\mathbf{J}}$ are constant:

$$\frac{d}{dt} \langle \psi(t) | \hat{\mathbf{J}} | \psi(t) \rangle = 0. \quad (9.91)$$

Thus, for an isolated system, total angular momentum is conserved; this is a direct consequence of rotational symmetry.

Since \hat{H} and $\hat{\mathbf{J}}$ commute, they possess a simultaneous set of eigenstates $|\alpha jm\rangle$, where “ α ” represents all other quantum numbers needed to uniquely identify a particular energy eigenstate of \hat{H} . From the Wigner-Eckart theorem, the matrix elements of \hat{H} between these eigenstates must be of the form

$$\langle \alpha'' j'' m'' | \hat{H} | \alpha' j' m' \rangle = \langle \alpha'' j'' | \hat{H} | \alpha' j' \rangle \delta_{j'' j'} \delta_{m'' m'}, \quad (9.92)$$

where the reduced matrix element $\langle \alpha'' j'' | \hat{H} | \alpha' j' \rangle$ is a constant, independent of the quantum number m . In particular, the expectation values of \hat{H} are

$$\langle \alpha jm | \hat{H} | \alpha jm \rangle = \langle \alpha j | \hat{H} | \alpha j \rangle. \quad (9.93)$$

Hence the energy eigenvalues for an isolated system can depend only on α and j , and must be independent of m . The energy levels therefore have a degeneracy $2j + 1$ with respect to angular momentum. Further degeneracy is possible if the Hamiltonian \hat{H} possesses additional symmetry properties involving the quantum numbers α .

This illustrates a general principle, that degeneracies in quantum mechanics are hardly ever accidental, but indicate the existence of an underlying symmetry (in this case, rotational invariance).

9.5 The Wigner-Eckart Theorem for Vector Operators

Consider an operator $\hat{\mathbf{V}}$ which is a vector operator with respect to an angular momentum operator $\hat{\mathbf{J}}$, and thus satisfies Eq. (9.58):

$$[\hat{J}_i, \hat{V}_j] = i\hbar\epsilon_{ijk}\hat{V}_k. \quad (9.94)$$

Then the Wigner-Eckart theorem states that the matrix elements of $\hat{\mathbf{V}}$ between eigenstates $|\alpha' j' m'\rangle$ and $|\alpha'' j'' m''\rangle$ of $\hat{\mathbf{J}}$ must be of the form

$$\langle \alpha'' j'' m'' | \hat{V}_m | \alpha' j' m' \rangle = \langle \alpha'' j'' | \hat{\mathbf{V}} | \alpha' j' \rangle \langle 1m; j' m' | j'' m'' \rangle. \quad (9.95)$$

where the quantum number m takes the values $m = \pm 1, 0$, and \hat{V}_m is a spherical component of the operator $\hat{\mathbf{V}}$. The reduced matrix element $\langle \alpha''j'' | \hat{\mathbf{V}} | \alpha'j' \rangle$ is a complex constant which is independent of the quantum numbers m , m' and m'' . The final factor, $\langle 1m; j'm' | j''m'' \rangle$, is a Clebsch-Gordan coefficient which can be obtained by considering the angular momentum combination $j'' = 1 \otimes j' = j', j' = \pm 1$.

For completeness, the proof of Eq. (9.95) is provided in the Appendix A.3.

The definition of the quantum numbers m , m' and m'' depends on the arbitrary choice of quantisation axis. The Wigner-Eckart theorem shows that the dependence of vector operator matrix elements on this arbitrary choice of axis is carried entirely by Clebsch-Gordan coefficients; these are independent of the particular vector operator, $\hat{\mathbf{V}}$, being considered.

The reduced matrix element $\langle \alpha''j'' | \hat{\mathbf{V}} | \alpha'j' \rangle$ is a common, constant component of a set of $3(2j' + 1)(2j'' + 1)$ matrix elements, $\langle \alpha''j''m'' | \hat{V}_m | \alpha'j'm' \rangle$. To determine the reduced matrix element, only *one* of the matrix elements belonging to this set needs to be evaluated explicitly. Once this has been done, all other matrix elements of $\hat{\mathbf{V}}$ are determined straightforwardly by looking up the appropriate Clebsch-Gordan coefficients.

The Wigner-Eckart theorem allows various general properties of the matrix elements of vector operators (positions, linear momenta, angular momenta, ...) to be obtained. For example, it leads directly to selection rules in atomic transitions, to general expressions for atomic energy level shifts in electric and magnetic fields (the Stark and Zeeman effects), as well as to many applications involving the scattering or decay of nuclei and elementary particles.

9.5.1 Selection Rules for Vector Operator Matrix Elements

For the case $j' = j'' = 0$, and hence also $m' = m'' = 0$, the Clebsch-Gordan coefficient $\langle 1m; j'm' | j''m'' \rangle$ vanishes for all m :

$$\langle 1m; 00 | 00 \rangle = 0, \quad m = 0, \pm 1. \quad (9.96)$$

The Wigner-Eckart theorem, Eq. (9.95), then shows that the matrix elements of any vector operator $\hat{\mathbf{V}}$ taken between eigenstates of zero angular momentum must vanish:

$$\langle \alpha''00 | \hat{V}_m | \alpha'00 \rangle = 0. \quad (9.97)$$

Equivalently, this can be written as

$\langle \alpha''00 | \hat{\mathbf{V}} | \alpha'00 \rangle = 0.$

(9.98)

For all other cases, with $j' > 0$ or $j'' > 0$, the Clebsch-Gordan coefficient $\langle 1m; j'm' | j''m'' \rangle$ vanishes unless

$$j'' = \{j', j' \pm 1\}, \quad m'' = m + m', \quad m = \{0, \pm 1\} \quad (9.99)$$

The conditions in Eq. (9.96) and (9.99) for a non-zero Clebsch-Gordan coefficient can be summarised as the *selection rules*

$$|\Delta j| = \{0, 1\}; \quad j' + j'' \geq 1; \quad |\Delta m| = \{0, 1\}. \quad (9.100)$$

The Wigner-Eckart theorem shows that the matrix element $\langle 1m; j'm' | j''m'' \rangle$ will vanish if these selection rules are not satisfied.

Satisfying the above selection rules does not, on its own, guarantee that the matrix element will be non-zero. The Clebsch-Gordan coefficient may anyway still vanish, for example $\langle 10; 10 | 10 \rangle = 0$, or the operator $\hat{\mathbf{V}}$ may possess additional symmetry properties which lead to additional selection rules.

9.5.2 The Landé Projection Formula

The angular momentum operator $\hat{\mathbf{J}}$ is itself a vector operator, and so must satisfy the Wigner-Eckart theorem. Therefore, matrix elements of the spherical components \hat{J}_{+1} , \hat{J}_{-1} , \hat{J}_0 of $\hat{\mathbf{J}}$ taken between angular momentum eigenstates must be of the form

$$\langle \alpha''j''m'' | \hat{J}_m | \alpha'j'm' \rangle = \langle \alpha''j'' | \hat{\mathbf{J}} | \alpha'j' \rangle \langle 1m; j'm' | j''m'' \rangle, \quad (9.101)$$

where the reduced matrix element $\langle \alpha''j'' | \hat{\mathbf{J}} | \alpha'j' \rangle$ is an overall constant which is independent of m , m' and m'' . In particular, taking $\alpha'' = \alpha'$ and $j'' = j'$, we have

$$\langle \alpha'j'm'' | \hat{J}_m | \alpha'j'm' \rangle = \langle \alpha'j' | \hat{\mathbf{J}} | \alpha'j' \rangle \langle 1m; j'm' | j'm'' \rangle. \quad (9.102)$$

Similarly, for any vector operator $\hat{\mathbf{V}}$, the Wigner-Eckart theorem gives

$$\langle \alpha'j'm'' | \hat{V}_m | \alpha'j'm' \rangle = \langle \alpha'j' | \hat{\mathbf{V}} | \alpha'j' \rangle \langle 1m; j'm' | j'm'' \rangle. \quad (9.103)$$

Eliminating the Clebsch-Gordan coefficient between Eq. (9.102) and (9.103), we obtain

$$\langle \alpha'j'm'' | \hat{V}_m | \alpha'j'm' \rangle = \frac{\langle \alpha'j' | \hat{\mathbf{V}} | \alpha'j' \rangle}{\langle \alpha'j' | \hat{\mathbf{J}} | \alpha'j' \rangle} \langle \alpha'j'm'' | \hat{J}_m | \alpha'j'm' \rangle. \quad (9.104)$$

This can equivalently be written as

$$\langle \alpha'j'm'' | \hat{\mathbf{V}} | \alpha'j'm' \rangle = \frac{\langle \alpha'j' | \hat{\mathbf{V}} | \alpha'j' \rangle}{\langle \alpha'j' | \hat{\mathbf{J}} | \alpha'j' \rangle} \langle \alpha'j'm'' | \hat{\mathbf{J}} | \alpha'j'm' \rangle. \quad (9.105)$$

This expression no longer explicitly involves the quantum number $m = \pm 1, 0$. Thus we can simplify the notation by dropping one prime throughout, and write

$$\langle \alpha jm' | \hat{\mathbf{V}} | \alpha jm \rangle = g_{\alpha j} \langle \alpha jm' | \hat{\mathbf{J}} | \alpha jm \rangle, \quad (9.106)$$

where $g_{\alpha j}$, the ratio of reduced matrix elements, is defined as

$$g_{\alpha j} \equiv \frac{\langle \alpha j | \hat{\mathbf{V}} | \alpha j \rangle}{\langle \alpha j | \hat{\mathbf{J}} | \alpha j \rangle}, \quad (9.107)$$

and is a constant which is independent of m and m' . Eq. (9.106) can be considered as holding for each of the *spherical* components of $\hat{\mathbf{V}}$ and $\hat{\mathbf{J}}$, or equivalently, for each of the *Cartesian* components of $\hat{\mathbf{V}}$ and $\hat{\mathbf{J}}$, with a common value of the constant $g_{\alpha j}$.

The expectation value $\langle \alpha jm | \hat{\mathbf{V}} \cdot \hat{\mathbf{J}} | \alpha jm \rangle$ of the scalar product operator $\hat{\mathbf{V}} \cdot \hat{\mathbf{J}}$ is

$$\langle \alpha jm | \hat{\mathbf{V}} \cdot \hat{\mathbf{J}} | \alpha jm \rangle = \sum_k \langle \alpha jm | \hat{V}_k \hat{J}_k | \alpha jm \rangle. \quad (9.108)$$

The eigenstates $|\alpha jm\rangle$ of the total angular momentum operator $\hat{\mathbf{J}}$ satisfy the completeness relation

$$\sum_{\alpha', j', m'} |\alpha' j' m'\rangle \langle \alpha' j' m'| = \hat{\mathbb{I}}. \quad (9.109)$$

Using this completeness relation to insert a factor of identity into the matrix element of Eq. (9.108) then gives

$$\langle \alpha jm | \hat{\mathbf{V}} \cdot \hat{\mathbf{J}} | \alpha jm \rangle = \sum_{\alpha', j', m'} \langle \alpha jm | \hat{\mathbf{V}} | \alpha' j' m' \rangle \cdot \langle \alpha' j' m' | \hat{\mathbf{J}} | \alpha jm \rangle. \quad (9.110)$$

The matrix elements $\langle \alpha' j' m' | \hat{\mathbf{J}} | \alpha jm \rangle$ vanish unless $\alpha' = \alpha$ and $j' = j$:

$$\langle \alpha' j' m' | \hat{\mathbf{J}} | \alpha jm \rangle = \langle \alpha jm' | \hat{\mathbf{J}} | \alpha jm \rangle \delta_{\alpha' \alpha} \delta_{j' j}. \quad (9.111)$$

Hence the summations over α' and j' can be trivially carried out, leaving

$$\langle \alpha jm | \hat{\mathbf{V}} \cdot \hat{\mathbf{J}} | \alpha jm \rangle = \sum_{m'} \langle \alpha jm | \hat{\mathbf{V}} | \alpha jm' \rangle \cdot \langle \alpha jm' | \hat{\mathbf{J}} | \alpha jm \rangle. \quad (9.112)$$

Substituting from Eq. (9.106) then gives

$$\langle \alpha jm | \hat{\mathbf{V}} \cdot \hat{\mathbf{J}} | \alpha jm \rangle = g_{\alpha j} \sum_{m'} \langle \alpha jm | \hat{\mathbf{J}} | \alpha jm' \rangle \cdot \langle \alpha jm' | \hat{\mathbf{J}} | \alpha jm \rangle. \quad (9.113)$$

The sum over j' can now be reinstated, giving

$$\langle \alpha jm | \hat{\mathbf{V}} \cdot \hat{\mathbf{J}} | \alpha jm \rangle = g_{\alpha j} \sum_{j', m'} \langle \alpha jm | \hat{\mathbf{J}} | \alpha j' m' \rangle \cdot \langle \alpha j' m' | \hat{\mathbf{J}} | \alpha jm \rangle. \quad (9.114)$$

Invoking completeness again, the summation over j' and m' just gives a factor of identity. Hence we have

$$\langle \alpha jm | \hat{\mathbf{V}} \cdot \hat{\mathbf{J}} | \alpha jm \rangle = g_{\alpha j} \langle \alpha jm | \hat{\mathbf{J}}^2 | \alpha jm \rangle \quad (9.115)$$

$$= g_{\alpha j} j(j+1)\hbar^2. \quad (9.116)$$

Provided $j > 0$, the constant $g_{\alpha j}$ is therefore given by

$$g_{\alpha j} = \frac{\langle \alpha jm | \hat{\mathbf{V}} \cdot \hat{\mathbf{J}} | \alpha jm \rangle}{j(j+1)\hbar^2}. \quad (9.117)$$

Although the quantum number m appears inside the matrix element on the right-hand side of Eq. (9.117), the constant $g_{\alpha j}$ is in fact independent of m , as is evident from its original definition as a ratio of reduced matrix elements in Eq. (9.107).

Substituting for $g_{\alpha j}$ in Eq. (9.106) gives (for $j > 0$) the *Landé projection formula*

$$\langle \alpha jm' | \hat{\mathbf{V}} | \alpha jm \rangle = \frac{\langle \alpha jm | \hat{\mathbf{V}} \cdot \hat{\mathbf{J}} | \alpha jm \rangle}{j(j+1)\hbar^2} \langle \alpha jm' | \hat{\mathbf{J}} | \alpha jm \rangle. \quad (9.118)$$

The z -component of the projection formula is

$$\langle \alpha jm' | \hat{V}_z | \alpha jm \rangle = g_{\alpha j} m e \alpha jm' \hat{J}_z \alpha jm = g_{\alpha j} m \hbar \delta_{m'm}. \quad (9.119)$$

Taking $m' = m$ then gives the expectation values of \hat{V}_z as

$$\langle \alpha jm | \hat{\mathbf{V}} | \alpha jm \rangle = g_{\alpha j} m \hbar. \quad (9.120)$$

Finally, for the case $j = 0$, Eq. (9.98) gives simply

$$\langle \alpha 00 | \hat{\mathbf{V}} | \alpha 00 \rangle = 0. \quad (9.121)$$

Thus in Eq. (9.106), for $j = 0$, the Landé g -factor vanishes:

$$g_{\alpha 0} = 0. \quad (9.122)$$

9.6 Magnetic Dipole Moments

The Hamiltonian operator describing the motion of a particle of charge q and mass m in a magnetic field \mathbf{B} includes an interaction which can be described in terms of an orbital magnetic dipole moment operator $\hat{\mu}_L$ as

$$\hat{H}_B = -\hat{\mu}_L \cdot \mathbf{B}; \quad \hat{\mu}_L = \frac{q}{2m} \hat{\mathbf{L}}, \quad (9.123)$$

where $\hat{\mathbf{L}} = \hat{\mathbf{r}} \times \hat{\mathbf{p}} = -i\hbar\hat{\mathbf{r}} \times \nabla$ is the orbital angular momentum operator for the particle. Similarly, a spin-half particle is predicted by the Dirac equation to possess an intrinsic magnetic dipole moment $\hat{\mu}_S$ which interacts with a magnetic field as

$$\hat{H}_B = -\hat{\mu}_S \cdot \mathbf{B}; \quad \hat{\mu}_S = \frac{q}{m} \hat{\mathbf{S}}, \quad (9.124)$$

where $\hat{\mathbf{S}}$ is the spin operator for the particle. Thus, both orbital and spin magnetic moments have interactions with a magnetic field which are generically of the form

$$\hat{H} = -\hat{\mu} \cdot \mathbf{B}; \quad \hat{\mu} = \gamma \hat{\mathbf{J}}, \quad (9.125)$$

where $\hat{\mathbf{J}}$ is an angular momentum operator, and γ is a constant known as the *gyromagnetic ratio*.

The (scalar) magnetic dipole moment associated with a (vector) magnetic dipole operator $\hat{\mu}$ is defined as

$$\mu = \langle jj | \hat{\mu}_z | jj \rangle. \quad (9.126)$$

This is the expectation value of the operator $\hat{\mu}_z$ when the particle is in its “most spin-up” state, $|jj\rangle \equiv |j, m_j = j\rangle$, for which the angular momentum quantum number m_j (defined with respect to the z -axis) takes its maximum possible value.

The state $|jj\rangle$ is an eigenstate of \hat{J}_z with eigenvalue $m_j\hbar = j\hbar$:

$$\hat{J}_z |jj\rangle = j\hbar |jj\rangle. \quad (9.127)$$

The magnetic moment μ is therefore related to the gyromagnetic ratio γ as

$$\mu = \gamma \langle jj | \hat{J}_z | jj \rangle = \gamma j\hbar \quad (9.128)$$

9.6.1 g -Factors

It is often convenient to express magnetic moments in terms of a common overall scale, typically either the Bohr magneton $\mu_B \equiv e\hbar/2m_e$ or the nuclear magneton $\mu_N \equiv e\hbar/2m_p$. In atomic physics, for example, since the magnetic moment of the electron is $\mu_e \approx -\mu_B$, the Bohr magneton μ_B is a convenient scale and it is common to set

$$\mu = -gj\mu_B, \quad (9.129)$$

where the constant g is referred to simply as the *g-factor*. In nuclear physics, the nuclear magneton μ_N is a more convenient scale, and the *g-factor* is defined using

$$\mu = gj\mu_N. \quad (9.130)$$

The signs in the expressions above reflect the most common sign conventions used in each area. Using Eq. (9.128), the connections between the gyromagnetic ratio and the *g*-factor for each choice of scale are seen to be

$$\gamma = -g \frac{\mu_B}{\hbar}; \quad \gamma = g \frac{\mu_N}{\hbar}. \quad (9.131)$$

The magnetic moment of the electron, for example, with spin $j = 1/2$, can be expressed in units of the Bohr magneton as

$$\mu_e = -\frac{1}{2}g_e\mu_B, \quad (9.132)$$

where the *g*-factor for the electron is predicted by the Dirac equation to be $g_e = 2$, while the measured value is $g_e \approx 2.0023$. The equivalent expression for the magnetic moment of the proton is

$$\mu_p = -\frac{1}{2}g_p\mu_B, \quad (9.133)$$

where the measured *g*-factor is $g_p \approx -0.0031$. In nuclear and particle physics applications, the proton magnetic moment is more conveniently expressed in terms of μ_N rather than μ_B , as

$$\mu_p = \frac{1}{2}g_p\mu_N, \quad (9.134)$$

In this case, the measured *g*-factor becomes $g_p \approx 5.586$. Similarly, the magnetic moment of the neutron is expressed as

$$\mu_n = \frac{1}{2}g_n\mu_N, \quad (9.135)$$

with $g_n \approx -3.826$.

	spin J, I	"nuclear"		"atomic"	
		μ/μ_N	g_J	μ/μ_N	g_I
e ⁻	1/2			-1.0011	-2.0023
e ⁺	1/2			+1.0011	-2.0023
p	1/2	+2.793	+5.586	+0.00152	-0.00304
n	1/2	-1.913	-3.826	-0.00104	+0.00208
⁴ He	0	0	0	0	0
⁶ Li	1	+0.822	+0.822	+0.00045	-0.00045
⁷ Li	3/2	+3.256	+2.171	+0.00177	-0.00118
¹² C	0	0	0	0	0
⁸⁵ Rb	5/2	+1.353	+0.541	+0.00074	-0.00029
⁸⁷ Rb	3/2	+2.750	+1.833	+0.00150	-0.00100
¹⁰⁷ Ag	1/2	-0.114	-0.228	-0.00006	+0.00012
¹⁰⁹ Ag	1/2	-0.131	-0.262	-0.00007	+0.00014

Table 9.1: Examples of the magnetic moments and g -factors of various particles and nuclei, expressed in terms of both μ_B and μ_N .

9.6.2 Combining magnetic moments

It is often the case that we need to determine the combined effects of two or more magnetic dipole moment interactions, for example the combined effect of the orbital and spin magnetic dipole moments of a particle, or the combined effect of the spin magnetic dipole moments of two spin-half particles.

Thus, in general, we are led to consider the combined effect of two magnetic dipole moment operators, $\hat{\mu}_1 = \gamma_1 \hat{\mathbf{J}}_1$ and $\hat{\mu}_2 = \gamma_2 \hat{\mathbf{J}}_2$. In a magnetic field \mathbf{B} , the Hamiltonian contains the sum of the two interaction terms

$$\hat{H}_B = -\hat{\mu}_1 \cdot \mathbf{B} - \hat{\mu}_2 \cdot \mathbf{B}. \quad (9.136)$$

The overall interaction with the magnetic field is therefore

$$\hat{H}_B = -\hat{\mu} \cdot \mathbf{B}, \quad (9.137)$$

where $\hat{\mu}$ is the total magnetic moment operator:

$$\hat{\mu} = \hat{\mu}_1 + \hat{\mu}_2 = \gamma_1 \hat{\mathbf{J}}_1 + \gamma_2 \hat{\mathbf{J}}_2 \quad (9.138)$$

If we choose to orient the z -axis along the magnetic field direction,

$$\mathbf{B} = (0, 0, B_z), \quad (9.139)$$

then the overall interaction becomes

$$\hat{H}_B = -\hat{\mu}_z B_z, \quad \hat{\mu}_z = \gamma_1 \hat{J}_{1z} + \gamma_2 \hat{J}_{2z}. \quad (9.140)$$

In the uncoupled basis of states,

$$|\alpha j_1 m_1 j_2 m_2\rangle = |\alpha j_1 m_1\rangle \otimes |\alpha j_2 m_2\rangle, \quad (9.141)$$

finding the matrix elements of $\hat{\mu}_z$ is straightforward. We have

$$\hat{J}_{1z} |\alpha j_1 m_1 j_2 m_2\rangle = m_1 \hbar |\alpha j_1 m_1 j_2 m_2\rangle, \quad (9.142)$$

$$\hat{J}_{2z} |\alpha j_1 m_1 j_2 m_2\rangle = m_2 \hbar |\alpha j_1 m_1 j_2 m_2\rangle, \quad (9.143)$$

and hence

$$\boxed{\langle \alpha j_1 m'_1 j_2 m'_2 | \hat{\mu}_z | \alpha j_1 m_1 j_2 m_2 \rangle = (\gamma_1 m_1 + \gamma_2 m_2) \hbar \delta_{m'_1 m_1} \delta_{m'_2 m_2}.} \quad (9.144)$$

Thus, in the uncoupled basis, for given j_1 and j_2 , the matrix representation of $\hat{\mu}_z$ is diagonal with respect to the quantum numbers m_1 and m_2 .

This is the case that applies, for example, to the Zeeman effect at high magnetic field.

However, in other cases, for example the Zeeman effect at low magnetic field, or the determination of the magnetic moment of the proton in terms of its constituent quark magnetic moments, the matrix elements of $\hat{\mu}_z$ must be determined in the *coupled* basis $|\alpha jm\rangle$ of total angular momentum eigenstates, rather than in the uncoupled basis $|\alpha j_1 m_1 j_2 m_2\rangle$.

In the coupled basis, if the two gyromagnetic ratios involved are the same, $\gamma_1 = \gamma_2$, finding the matrix elements is trivial. We have

$$\hat{\mu} = \gamma \hat{\mathbf{J}}, \quad (9.145)$$

where $\hat{\mathbf{J}}$ is the total angular momentum of the combined system,

$$\hat{\mathbf{J}} = \hat{\mathbf{J}}_1 + \hat{\mathbf{J}}_2, \quad (9.146)$$

and where

$$\gamma = \gamma_1 + \gamma_2. \quad (9.147)$$

The matrix elements of $\hat{\mu}_z$ in the coupled basis are then given simply by

$$\langle \alpha jm' | \hat{\mu}_z | \alpha jm \rangle = \gamma m \hbar \delta_{m'm}. \quad (9.148)$$

However, in the more general case that the gyromagnetic ratios are different, $\gamma_1 \neq \gamma_2$, the operator $\hat{\mu}$ is *not* directly proportional to the total angular momentum $\hat{\mathbf{J}}$:

$$\hat{\mu} \neq \gamma \hat{\mathbf{J}}. \quad (9.149)$$

In this case, finding the matrix elements of $\hat{\mu}_z$ between total angular momentum eigenstates is no longer straightforward. However, as we now show, the Landé projection formula allows proportionality between the operators $\hat{\mu}$ and $\hat{\mathbf{J}}$ to be *effectively* restored for states of given total angular momentum j , with a constant of proportionality γ_j which depends on j (and j_1 and j_2).

9.6.2.1 The Landé g-Factor

Since $\hat{\mu}$ is a vector operator, the projection formula of Eq. (9.118) (obtained from the Wigner-Eckart theorem) is applicable. This gives the matrix elements of the operator $\hat{\mu}$ between eigenstates of total angular momentum $\hat{\mathbf{J}}$ as

$$\langle \alpha jm' | \hat{\mu} | \alpha jm \rangle = \gamma_{\alpha j} \langle \alpha jm' | \hat{\mathbf{J}} | \alpha jm \rangle, \quad (9.150)$$

where the constant $\gamma_{\alpha j}$ is independent of m and m' . The z component of this equation gives

$$\langle \alpha jm' | \hat{\mu}_z | \alpha jm \rangle = \gamma_{\alpha j} \langle \alpha jm' | \hat{J}_z | \alpha jm \rangle = \gamma_{\alpha j} m \hbar \delta_{m'm}. \quad (9.151)$$

Thus, for a given total angular momentum j , the matrix representation is diagonal with respect to m . However, we still need to find $\gamma_{\alpha j}$ in terms of γ_1 and γ_2 .

For $j > 0$, the effective gyromagnetic ratio $\gamma_{\alpha j}$ is given by Eq. (9.117) as

$$\gamma_{\alpha j} = \frac{\langle \alpha jm | \hat{\mu} \cdot \hat{\mathbf{J}} | \alpha jm \rangle}{j(j+1)\hbar^2}. \quad (9.152)$$

The numerator in the above expression is

$$\langle \alpha jm | \hat{\mu} \cdot \hat{\mathbf{J}} | \alpha jm \rangle = \gamma_1 \langle \alpha jm | \hat{\mathbf{J}}_1 \cdot \hat{\mathbf{J}} | \alpha jm \rangle + \gamma_2 \langle \alpha jm | \hat{\mathbf{J}}_2 \cdot \hat{\mathbf{J}} | \alpha jm \rangle. \quad (9.153)$$

Squaring the relation $\hat{\mathbf{J}}_2 = \hat{\mathbf{J}} - \hat{\mathbf{J}}_1$ gives the scalar product $\hat{\mathbf{J}}_1 \cdot \hat{\mathbf{J}}$ as

$$\hat{\mathbf{J}}_1 \cdot \hat{\mathbf{J}} = \frac{1}{2} (\hat{\mathbf{J}}^2 - \hat{\mathbf{J}}_1^2 - \hat{\mathbf{J}}_2^2). \quad (9.154)$$

For given j_1 and j_2 , the state $|\alpha jm\rangle$ is an eigenstate of $\hat{\mathbf{J}}_1^2$ and $\hat{\mathbf{J}}_2^2$ (as well as of $\hat{\mathbf{J}}^2$) with

$$\hat{\mathbf{J}}_1^2 |\alpha jm\rangle = j_1(j_1+1)\hbar^2 |\alpha jm\rangle, \quad \hat{\mathbf{J}}_2^2 |\alpha jm\rangle = j_2(j_2+1)\hbar^2 |\alpha jm\rangle. \quad (9.155)$$

Hence the expectation value of $\hat{\mathbf{J}}_1 \cdot \hat{\mathbf{J}}$ is

$$\langle \alpha jm | \hat{\mathbf{J}}_1 \cdot \hat{\mathbf{J}} | \alpha jm \rangle = \frac{\hbar^2}{2} [j(j+1) + j_1(j_1+1) - j_2(j_2+1)]. \quad (9.156)$$

The equivalent result for $\hat{\mathbf{J}}_2 \cdot \hat{\mathbf{J}}$ is obtained by interchanging j_1 and j_2 :

$$\langle \alpha jm | \hat{\mathbf{J}}_2 \cdot \hat{\mathbf{J}} | \alpha jm \rangle = \frac{\hbar^2}{2} [j(j+1) + j_2(j_2+1) - j_1(j_1+1)]. \quad (9.157)$$

Combining the various equations above, we obtain $\gamma_{\alpha j}$ as

$$\gamma_j = \gamma_1 \frac{j(j+1) + j_1(j_1+1) - j_2(j_2+1)}{2j(j+1)} + \gamma_2 \frac{j(j+1) + j_2(j_2+1) - j_1(j_1+1)}{2j(j+1)}. \quad (9.158)$$

As anticipated, the constant $\gamma_{\alpha j}$ is independent of the quantum number m . Eq. (9.158) shows that it is also independent of the quantum numbers α , so we can rename $\gamma_{\alpha j} \rightarrow \gamma_j$.

Eq. (9.150) can now be written as

$$\langle \alpha jm' | \hat{\mu} | \alpha jm \rangle = \gamma_j \langle \alpha jm' | \hat{\mathbf{J}} | \alpha jm \rangle. \quad (9.159)$$

From Eq. (9.138), we have therefore shown that, for given values of j , j_1 and j_2 , matrix elements of magnetic dipole moment operators combine as

$$\langle \alpha jm' | \gamma_1 \hat{\mathbf{J}}_1 + \gamma_2 \hat{\mathbf{J}}_2 | \alpha jm \rangle = \gamma_j \langle \alpha jm' | \hat{\mathbf{J}} | \alpha jm \rangle. \quad (9.160)$$

Thus, when a composite particle or system built from j_1 and j_2 is known to be in an eigenstate with total angular momentum quantum number j , the interaction of the composite system with a magnetic field \mathbf{B} , Eq. (9.136), is *effectively* described by the Hamiltonian

$$\hat{H}_B = -\gamma_j \hat{\mathbf{J}} \cdot \mathbf{B}. \quad (9.161)$$

Equivalently, in an angular momentum state j , the total magnetic moment operator of Eq. (9.138) is effectively given by

$$\hat{\mu}_j = \gamma_j \hat{\mathbf{J}}. \quad (9.162)$$

The magnetic moments can be expressed in terms of g -factors rather than gyromagnetic ratios. For example, using the Bohr magneton as a common overall scale, Eq. (9.131), we can introduce g -factors for each magnetic dipole as

$$\gamma_j = -g_j \frac{\mu_B}{\hbar}, \quad \gamma_1 = -g_1 \frac{\mu_B}{\hbar}, \quad \gamma_2 = -g_2 \frac{\mu_B}{\hbar}. \quad (9.163)$$

Eq. (9.158) retains the same form; we obtain

$$g_j = g_1 \frac{j(j+1) + j_1(j_1+1) - j_2(j_2+1)}{2j(j+1)} + g_2 \frac{j(j+1) + j_2(j_2+1) - j_1(j_1+1)}{2j(j+1)}.$$

(9.164)

Eq. (9.158) and (9.164) apply to the case $j > 0$. For the case $j = 0$, Eq. (9.98) gives the matrix elements of $\hat{\mu}$ as simply

$$\langle \alpha 00 | \hat{\mu} | \alpha 00 \rangle = 0.$$

(9.165)

Thus a composite system in a state with zero total angular momentum has no net dipole interaction with a magnetic field \mathbf{B} ; we effectively have $\gamma = 0$.

9.6.2.2 Particular Cases

For the case that the two g -factors are equal, $g_1 = g_2$, Eq. (9.164) reduces simply to

$$g_j = g_1 = g_2, \quad (9.166)$$

as expected.

For the case $g_1 = 1$ and $g_2 = 2$, Eq. (9.164) simplifies as

$$g_j = \frac{3}{2} - \frac{1}{2} \frac{j_1(j_1+1) - j_2(j_2+1)}{j(j+1)}. \quad (9.167)$$

This case is relevant to combining the orbital ($g_\ell = 1$) and spin ($g_e = 2$) magnetic dipole moments of atomic electrons, for example. In this context, the equation above would typically be written as

$$g_J = \frac{3}{2} - \frac{1}{2} \frac{L(L+1) - S(S+1)}{J(J+1)}. \quad (9.168)$$

This applies both to a single electron, in which case $S = 1/2$, or to an N -electron system, in which case S could take any integer or half-integer value up to a maximum of $S = N/2$.

For the case $j_2 = 1/2$, and hence $j = j_1 \pm (1/2)$, Eq. (9.164) simplifies as

$$g_j = g_1 \pm \frac{g_2 - g_1}{2j_1 + 1}, \quad (j = j_1 \pm 1/2). \quad (9.169)$$

This case is relevant to combining the orbital and spin magnetic moments of a single spin-half particle. For an electron, with $j = \ell \otimes s = \ell \pm (1/2)$, and with $g_1 = g_\ell = 1$, $g_2 = g_e \approx 2$, Eq. (9.169) further simplifies as

$$g_j = 1 \pm \frac{1}{2\ell + 1}, \quad (j = \ell \pm 1/2). \quad (9.170)$$

Similarly, for a proton, we have $g_\ell = 1$ and $g_2 = g_p = 5.586$, and hence

$$g_j = 1 \pm \frac{4.586}{2\ell + 1}, \quad (j = \ell \pm 1/2). \quad (9.171)$$

For a neutron, we have $g_\ell = 0$ and $g_2 = g_n = -3.826$, and hence

$$g_j = \mp \frac{3.826}{2\ell + 1}, \quad (j = \ell \pm 1/2). \quad (9.172)$$

Eq. (9.171) and (9.172) have application in the shell model of nuclear structure.

CHAPTER 10

Identical Particles

Until now, most of our focus has been on the quantum mechanical behaviour of *individual* particles, or problems which can be “factorised” into independent single-particle systems.¹ However, most physical systems of interest involve the interaction of large numbers of particles; electrons in a solid, atoms in a gas, etc. In classical mechanics, particles are always distinguishable in the sense that, at least formally, their “trajectories” through phase space can be traced and their identity disclosed. However, in quantum mechanics, the intrinsic uncertainty in position, embodied in Heisenberg’s principle, demands a careful and separate consideration of distinguishable and indistinguishable particles. In the present section, we will consider how to formulate the wavefunction of many-particle systems, and address some of the (sometimes striking and often counter-intuitive) implications of particle indistinguishability.

10.1 Quantum Statistics

Consider then two indistinguishable particles confined to a box in one-dimension. Here, by indistinguishable, we mean that the particles can not be discriminated by some internal quantum number. For example, we might have two electrons of the same spin. The normalised two-particle wavefunction $\psi(x_1, x_2)$, which gives the probability $|\psi(x_1, x_2)|^2 dx_1 dx_2$ of finding simultaneously one particle in the interval x_1 to $x_1 + dx_1$ and another between x_2 to $x_2 + dx_2$, only makes sense if $|\psi(x_1, x_2)|^2 = |\psi(x_2, x_1)|^2$, since we can’t know which of the two indistinguishable particles we are finding where. It follows from this that the wavefunction can exhibit two (and, generically, only two) possible symmetries under exchange: $\psi(x_1, x_2) = \psi(x_2, x_1)$ or $\psi(x_1, x_2) = -\psi(x_2, x_1)$.² If two identical particles have a symmetric wavefunction in some state, particles of that type always have symmetric wavefunctions, and are called **bosons**. (If in some other state they had an antisymmetric wavefunction, then a linear superposition of those states would be neither symmetric nor antisymmetric, and so could not satisfy the relation $|\psi(x_1, x_2)|^2 = |\psi(x_2, x_1)|^2$.) Similarly, particles having antisymmetric wavefunctions are called **fermions**.

To construct wavefunctions for three or more fermions, let first suppose that the particles do not interact with each other, and are confined by a spin independent potential, such as the Coulomb field of a nucleus. In this case, the Hamiltonian will be symmetric

¹For example, the treatment of the hydrogen atom involves the separation of the system into centre of mass and relative motion. Each could be referred to an effective single-particle dynamics.

²We could in principle have $\psi(x_1, x_2) = e^{i\alpha}\psi(x_2, x_1)$, with α a constant phase. However, in this case we would not recover the original wavefunction on exchanging the particles twice. Curiously, in some two-dimensional theories used to describe the fractional quantum Hall effect, there exist collective excitations of the electron system – called **anyons** – that do have this kind of property. Such anyonic systems have been proposed as a strong candidate for the realisation of quantum computation. However, all ordinary “fundamental” particles are either bosons or fermions.

in the fermion degrees of freedom,

$$\hat{H} = \frac{\hat{\mathbf{p}}_1^2}{2m} + \frac{\hat{\mathbf{p}}_2^2}{2m} + \frac{\hat{\mathbf{p}}_3^2}{2m} + \cdots V(\mathbf{r}_1) + V(\mathbf{r}_2) + V(\mathbf{r}_3) + \cdots , \quad (10.1)$$

and the solutions of the Schrödinger equation will be products of eigenfunctions of the single-particle Hamiltonian $\hat{H}_s = \hat{\mathbf{p}}^2/2m + V(\mathbf{r})$. However, single products such as $\psi_a(1)\psi_b(2)\psi_c(3)$ do not have the required antisymmetry property under the exchange of any two particles. (Here a, b, c, \dots label the single-particle eigenstates of \hat{H}_s , and $1, 2, 3, \dots$ denote *both* space and spin coordinates of single particles, i.e. 1 stands for (\mathbf{r}_1, s_1) , etc.)

We could achieve the necessary antisymmetrisation for particles 1 and 2 by subtracting the same product wavefunction with the particles 1 and 2 interchanged, i.e. $\psi_a(1)\psi_b(2)\psi_c(3) \mapsto (\psi_a(1)\psi_b(2) - \psi_a(2)\psi_b(1))\psi_c(3)$, ignoring the overall normalisation for now. However, the wavefunction needs to be antisymmetrised with respect to *all* possible particle exchanges. So, for 3 particles, we must add together all $3!$ permutations of 1, 2, 3 in the state a, b, c , with a factor -1 for each particle exchange necessary to get to a particular ordering from the original ordering of 1 in a , 2 in b , and 3 in c . In fact, such a sum over permutations is precisely the *definition* of the determinant. So, with the appropriate normalisation factor:

$$\psi_{abc}(1, 2, 3) = \frac{1}{\sqrt{3!}} \begin{vmatrix} \psi_a(1) & \psi_b(1) & \psi_c(1) \\ \psi_a(2) & \psi_b(2) & \psi_c(2) \\ \psi_a(3) & \psi_b(3) & \psi_c(3) \end{vmatrix} \quad (10.2)$$

The determinantal form makes clear the antisymmetry of the wavefunction with respect to exchanging any two of the particles, since exchanging two rows of a determinant multiplies it by -1 . We also see from the determinantal form that the three states a, b, c *must* all be different, for otherwise two columns would be identical, and the determinant would be zero. This is just the manifestation of Pauli's exclusion principle: no two fermions can be in the same state.

Although these determinantal wavefunctions (known as **Slater determinants**), involving superpositions of single-particle states, are only strictly correct for non-interacting fermions, they provide a useful platform to describe electrons in atoms (or in a metal), with the electron-electron repulsion approximated by a single-particle potential. For example, the Coulomb field in an atom, as seen by the outer electrons, is partially shielded by the inner electrons, and a suitable $V(\mathbf{r})$ can be constructed self-consistently, by computing the single-particle eigenstates and finding their associated charge densities. (We shall discuss this "Hartree" approximation later in the course, when we study multielectron atoms.)

In the bosonic system, the corresponding many-particle wavefunction must be symmetric under particle exchange. We can obtain such a state by expanding all of the contributing terms from the Slater determinant and setting all of the signs to be positive. In other words, the bosonic wave function describes the uniform (equal phase) superposition of all possible permutations of product states. In this case, there is no requirement that the orbitals be different. (There is no exclusion principle for bosons.) One important example is the groundstate wavefunction of a collection of non-interacting gas of bosons, in which all bosons occupy the lowest energy single particle wavefunction, $\psi_0(\mathbf{r})$. The many-particle state (for N indistinguishable bosons) is

$$\Psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) = \prod_{i=1}^N \psi_0(\mathbf{r}_i) \quad (10.3)$$

This is the wavefunction for a **Bose-Einstein condensate**, with all particles condensed in the state $\psi_0(\mathbf{r})$. It describes a non-interacting gas of (spinless) bosons at zero temperature: i.e. the many-particle groundstate.

10.2 Space and spin wavefunctions

Although the methodology for constructing a basis of many-particle states outlined above is generic, it is not particularly convenient when the Hamiltonian is spin-independent. In this case we can simplify the structure of the wavefunction by factorising the spin and spatial components. Suppose we have two electrons (i.e. fermions) in some spin-independent potential $V(\mathbf{r})$. We know that the two-electron wavefunction must be antisymmetric under particle exchange. Since the Hamiltonian has no spin-dependence, we must be able to construct a set of common eigenstates of the Hamiltonian, the total spin, and the z -component of the total spin. For two electrons, there are four basis states in the spin space, the $S = 0$ spin singlet state, $|\chi_{S=0, S_z=0}\rangle = \frac{1}{\sqrt{2}}(|\uparrow_1\downarrow_2\rangle - |\downarrow_1\uparrow_2\rangle)$, and the three $S = 1$ spin triplet states,

$$|\chi_{11}\rangle = |\uparrow_1\uparrow_2\rangle, \quad |\chi_{10}\rangle = \frac{1}{\sqrt{2}}(|\uparrow_1\downarrow_2\rangle + |\downarrow_1\uparrow_2\rangle), \quad |\chi_{1,-1}\rangle = |\downarrow_1\downarrow_2\rangle. \quad (10.4)$$

Here the first arrow in the ket refers to the spin of particle 1, the second to particle 2.

It is evident that the spin singlet wavefunction is antisymmetric under the exchange of two particles, while the spin triplet wavefunction is symmetric. For a general state, the total wavefunction for the two electrons in a common eigenstate of $\hat{\mathbf{S}}^2$, \hat{S}_z and the Hamiltonian \hat{H} then has the form:

$$\Psi(\mathbf{r}_1, s_1; \mathbf{r}_2, s_2) = \psi(\mathbf{r}_1, \mathbf{r}_2)\chi(s_1, s_2), \quad (10.5)$$

where $\chi(s_1, s_2) = \langle s_1, s_2 | \chi \rangle$. For two electron degrees of freedom, the total wavefunction, ψ , must be antisymmetric under exchange. It follows that a pair of electrons in the spin *singlet* state must have a *symmetric* spatial wavefunction, $\psi(\mathbf{r}_1, \mathbf{r}_2)$, whereas electrons in the spin *triplet* states must have an *antisymmetric* spatial wavefunction. Before discussing the physical consequences of this symmetry, let us mention how this scheme generalises to more particles.

Unfortunately, in seeking a factorisation of the Slater determinant into a product of spin and spatial components for three electrons, things become more challenging. There are now $2^3 = 8$ basis states in the spin space. Four of these are accounted for by the spin $3/2$ state with $S_z = 3/2, 1/2, -1/2, -3/2$. Since all spins are aligned, this is evidently a symmetric state, so must be multiplied by an antisymmetric spatial wavefunction, itself a determinant. So far so good. But the other four states involve two pairs of total spin $1/2$ states built up of a singlet and an unpaired spin. They are orthogonal to the symmetric spin $3/2$ state, so they can't be symmetric. But they can't be antisymmetric either, since in each such state, two of the spins must be pointing in the same direction! An example of such a state is presented by $|\chi\rangle = |\uparrow\rangle \otimes \frac{1}{\sqrt{2}}(|\uparrow_2\downarrow_3\rangle - |\downarrow_2\uparrow_3\rangle)$. Evidently, this must be multiplied by a spatial wavefunction symmetric in 2 and 3. But to recover a total wave

function with *overall* antisymmetry it is necessary to add more terms:

$$\Psi(1, 2, 3) = \chi(s_1, s_2, s_3)\psi(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3) + \chi(s_2, s_3, s_1)\psi(\mathbf{r}_2, \mathbf{r}_3, \mathbf{r}_1) + \chi(s_3, s_1, s_2)\psi(\mathbf{r}_3, \mathbf{r}_1, \mathbf{r}_2). \quad (10.6)$$

Requiring the spatial wavefunction $\psi(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3)$ to be symmetric in 2, 3 is sufficient to guarantee the overall antisymmetry of the total wavefunction Ψ . For more than three electrons, similar considerations hold. The mixed symmetries of the spatial wavefunctions and the spin wavefunctions which together make a totally antisymmetric wavefunction are quite complex, and are described by **Young diagrams** (or **tableaux**).

10.3 Physical Consequences of Particle Statistics

The overall antisymmetry demanded by the many-fermion wavefunction has important physical implications. In particular, it determines the magnetic properties of atoms. The magnetic moment of the electron is aligned with its spin, and even though the spin variables do not appear in the Hamiltonian, the energy of the eigenstates depends on the relative spin orientation. This arises from the electrostatic repulsion between electrons. In the spatially antisymmetric state, the probability of electrons coinciding at the same position necessarily vanishes. Moreover, the nodal structure demanded by the antisymmetry places the electrons further apart on average than in the spatially symmetric state. Therefore, the electrostatic repulsion raises the energy of the spatially symmetric state above that of the spatially antisymmetric state. It therefore follows that the lower energy state has the electron spins pointing in the same direction. This argument is still valid for more than two electrons, and leads to **Hund's rule** for the magnetisation of incompletely filled inner shells of electrons in transition metal and rare earths atoms. This is the first step in understanding ferromagnetism (see 10.4.3).

A gas of hydrogen molecules provides another manifestation of wavefunction antisymmetry. In particular, the specific heat depends sensitively on whether the two protons (spin 1/2) in H₂ have their spins parallel or antiparallel, even though that alignment involves only a very tiny interaction energy. If the proton spins occupy a spin singlet configuration, the molecule is called **parahydrogen** while the triplet states are called **orthohydrogen**. These two distinct gases are remarkably stable – in the absence of magnetic impurities, para-ortho transitions take weeks.

The actual energy of interaction of the proton spins is of course completely negligible in the specific heat. The important contributions to the specific heat are the usual kinetic energy term, and the rotational energy of the molecule. This is where the overall (space × spin) antisymmetric wavefunction for the protons plays a role. Recall that the parity of a state with rotational angular momentum ℓ is $(-1)^\ell$. Therefore, parahydrogen, with an antisymmetric proton spin wavefunction, must have a symmetric proton *spatial* wavefunction, and so can only have even values of the rotational angular momentum. Orthohydrogen can only have odd values. The energy of the rotational level with angular momentum ℓ is $E_{\text{tot}}^\ell = 2\hbar\ell(\ell+1)/I$, where I denotes the moment of inertia of the molecule. So the two kinds of hydrogen gas have different sets of rotational energy levels, and consequently different specific heats.

10.3.0.1 Example: Excited State Spectrum of Helium

As a final example, and one that will feed into our discussion of multielectron atoms in the next chapter, let us consider the implications of particle statistics for the excited state spectrum of Helium. After Hydrogen, Helium is the simplest atom having two protons and two neutrons in the nucleus ($Z = 2$), and two bound electrons. As a complex many-body system, we have seen already that the Schrödinger equation is analytically intractable and must be treated perturbatively. Previously, in Chapter 7, we have used the ground state properties of Helium as a vehicle to practice perturbation theory. In the absence of direct electron-electron interaction, the Hamiltonian

$$\hat{H}_0 = \sum_{n=1}^2 \left[\frac{\hat{\mathbf{p}}_n^2}{2m} + V(\mathbf{r}_n) \right], \quad V(\mathbf{r}_n) = -\frac{1}{4\pi\epsilon_0} \frac{Ze^2}{r}, \quad (10.7)$$

is separable and the wavefunction can be expressed through the states of the hydrogen atom, ψ_{nlm} . In this approximation, the ground state wavefunction involves both electrons occupying the 1s state leading to an antisymmetric spin singlet wavefunction for the spin degrees of freedom, $|\Psi_{\text{g.s.}}\rangle = (|100\rangle \otimes |100\rangle) \otimes |\chi_{00}\rangle$. In Chapter 7, we made use of both the perturbative series expansion and the variational method to determine how the ground state energy is perturbed by the repulsive electron-electron interaction,

$$\hat{H}_1 = \frac{1}{4\pi\epsilon_0} \frac{e^2}{|\mathbf{r}_1 - \mathbf{r}_2|}. \quad (10.8)$$

Now let us consider the implications of particle statistics on the spectrum of the lowest excited states.

From the symmetry perspective, the ground state wavefunction belongs to the class of states with symmetric spatial wavefunctions, and antisymmetric spin (singlet) wavefunctions. These states are known as **parahelium**. In the absence of electronelectron interaction, the first excited states are degenerate and have the form,

$$|\psi_p\rangle = \frac{1}{\sqrt{2}} (|100\rangle \otimes |2\ell m\rangle + |2\ell m\rangle \otimes |100\rangle) \otimes |\chi_{100}\rangle. \quad (10.9)$$

The second class of states involve an antisymmetric spatial wavefunction, and symmetric (triplet) spin wavefunction. These states are known as **orthohelium**. Once again, in the absence of electron-electron interaction, the first excited states are degenerate and have the form,

$$|\psi_o\rangle = \frac{1}{\sqrt{2}} (|100\rangle \otimes |2\ell m\rangle - |2\ell m\rangle \otimes |100\rangle) \otimes |\chi_{1S_z}\rangle. \quad (10.10)$$

The perturbative shift in the ground state energy has already been calculated within the framework of first order perturbation theory. Let us now consider the shift in the excited states. Despite the degeneracy, since the off-diagonal matrix elements vanish, we can make use of the first order of perturbation theory to compute the shift. In doing so, we obtain

$$\Delta E_{nl}^{\text{p.o.}} = \frac{1}{2} \frac{e^2}{4\pi\epsilon_0} \int d^3 r_1 d^3 r_2 |\psi_{100}(\mathbf{r}_1)\psi_{n\ell 0}(\mathbf{r}_2) \pm \psi_{n\ell 0}(\mathbf{r}_1)\psi_{100}(\mathbf{r}_2)|^2 \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|}, \quad (10.11)$$

with the plus sign refers to parahelium and the minus to orthohelium. Since the matrix element is independent of m , the $m = 0$ value considered here applies to all values of m . Rearranging this equation, we thus obtain $\Delta E_{n\ell}^{\text{p.o.}} = J_{n\ell} \pm K_{n\ell}$ where

$$J_{n\ell} = \frac{e^2}{4\pi\epsilon_0} \int d^3r_1 d^3r_2 \frac{|\psi_{100}(\mathbf{r}_1)|^2 |\psi_{n\ell 0}(\mathbf{r}_2)|^2}{|\mathbf{r}_1 - \mathbf{r}_2|}, \quad (10.12)$$

$$K_{n\ell} = \frac{e^2}{4\pi\epsilon_0} \int d^3r_1 d^3r_2 \frac{\psi_{100}^*(\mathbf{r}_1) \psi_{n\ell 0}^*(\mathbf{r}_2) \psi_{100}(\mathbf{r}_2) \psi_{n\ell 0}(\mathbf{r}_1)}{|\mathbf{r}_1 - \mathbf{r}_2|}. \quad (10.13)$$

Physically, the term $J_{n\ell}$ represents the electrostatic interaction energy associated with the two charge distributions $|\psi_{100}(\mathbf{r}_1)|^2$ and $|\psi_{n\ell 0}(\mathbf{r}_2)|^2$, and it is clearly positive. By contrast, the **exchange term**, which derives from the antisymmetry of the wavefunction, leads to a shift with opposite signs for ortho and para states. In fact, one may show that, in the present case, $K_{n\ell}$ is positive leading to a positive energy shift for parahelium and a negative shift for orthohelium. Moreover, noting that

$$2\mathbf{S}_1 \cdot \mathbf{S}_2 = (\mathbf{S}_1 + \mathbf{S}_2)^2 - \mathbf{S}_1^2 - \mathbf{S}_2^2 = \hbar^2 \left(S(S+1) - 2 \times \frac{3}{4} \right) = \hbar^2 \begin{cases} 1/2 & \text{triplet} \\ 3/2 & \text{singlet} \end{cases} \quad (10.14)$$

the energy shift can be written as

$$\Delta E_{n\ell}^{\text{p.o.}} = J_{n\ell} - \frac{1}{2} \left(1 + \frac{4}{\hbar^2} \mathbf{S}_1 \cdot \mathbf{S}_2 \right) K_{n\ell}. \quad (10.15)$$

This result shows that the electron-electron interaction leads to an effective **ferromagnetic** interaction between spins – i.e. the spins want to be aligned.

In addition to the large energy shift between the singlet and triplet states, electric dipole decay selection rules $\Delta\ell = \pm 1$, $\Delta s = 0$ cause decays from triplet to singlet states (or vice-versa) to be suppressed by a large factor (compared to decays from singlet to singlet or from triplet to triplet). This caused early researchers to think that there were two separate kinds of Helium. The diagrams in Fig. 10.1 shows the levels for parahelium (singlet) and for orthohelium (triplet) along with the dominant decay modes.

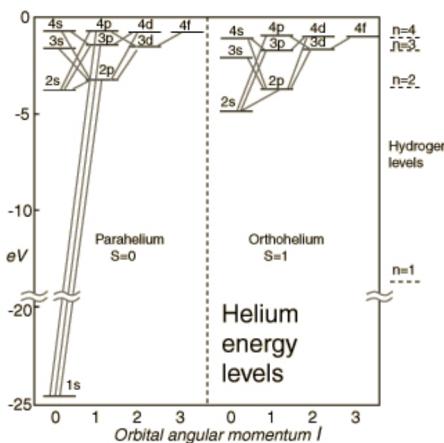


Fig. 10.1: Energy level diagram for ortho- and parahelium showing the first order shift in energies due to the Coulomb repulsion of electrons. Here we assume that one of the electrons stays close to the ground state of the unperturbed Hamiltonian.

10.4 Many-Body Systems

An important and recurring example of a many-body system is provided by the problem of free (i.e. non-interacting) non-relativistic quantum particles in a closed system – a box. The many-body Hamiltonian is then given simply by the kinetic energy

$$\hat{H}_0 = \sum_{i=1}^N \frac{\hat{\mathbf{p}}_i^2}{2m}, \quad (10.16)$$

where $\hat{\mathbf{p}}_i = -i\hbar\nabla_i$ and m denotes the particle mass. If we take the dimensions of the box to be L^d , and the boundary conditions to be periodic³ the normalised eigenstates of the single-particle Hamiltonian are simply given by plane waves, $\psi_{\mathbf{k}}(\mathbf{r}) = \langle \mathbf{r} | \mathbf{k} \rangle = \frac{1}{L^{d/2}} e^{i\mathbf{k}\cdot\mathbf{r}}$, with wavevectors taking discrete values,⁴

$$\mathbf{k} = \frac{2\pi}{L}(n_1, n_2, \dots, n_d), \quad n_i \text{ integer.} \quad (10.17)$$

To address the quantum mechanics of the system, we start with fermions.

10.4.1 Non-interacting Fermi Gas

In the (**spinless**) **Fermi system**, Pauli exclusion inhibits multiple occupancy of single-particle states. In this case, the many-body ground state wavefunction is obtained by filling states sequentially up to the **Fermi energy**, $E_F = \hbar^2 k_F^2 / 2m$, where the **Fermi wavevector**, k_F , is fixed by the number of particles. All the plane wave states $\psi_{\mathbf{k}}$ with energies lower than E_F are filled, while all states with energies larger than E_F remain empty. Since each state is associated with a k -space volume $(2\pi/L)^d$ (see Fig. 10.2), in the three-dimensional system, the total number of occupied states is given by $N = \left(\frac{L}{2\pi}\right)^3 \frac{4}{3}\pi k_F^3$, i.e. defining the mean particle density $\bar{n} \equiv N/L^3 = k_F^3/6\pi^2$,

$$E_F = \frac{\hbar^2}{2m} \left(6\pi^2 \bar{n}\right)^{2/3}. \quad (10.18)$$

The **density of states** per unit volume,

$$g(E) = \frac{1}{L^3} \frac{dN}{dE} = \frac{1}{6\pi^2} \frac{d}{dE} \left(\frac{2mE}{\hbar} \right)^{3/2} = \frac{1}{4\pi^2} \left(\frac{2m}{\hbar^2} \right)^{3/2} E^{1/2}. \quad (10.19)$$

We can also integrate to obtain the total energy density of all the fermions,

$$\frac{E_{\text{tot}}}{L^3} = \frac{1}{L^3} \int_0^{k_F} \frac{4\pi k^2}{(2\pi/L)^3} \frac{\hbar^2 k^2}{2m} = \frac{\hbar^2}{20\pi^2 m} (6\pi^2 \bar{n})^{5/3} = \frac{3}{5} \bar{n} E_F. \quad (10.20)$$

³It may seem odd to consider such an unphysical geometry – in reality, we are invariably dealing with a closed system in which the boundary conditions translate to “hard walls” – think of electrons in a metallic sample. Here, we have taken the boundary conditions to be periodic since it leads to a slightly more simple mathematical formulation. We could equally well consider closed boundary conditions, but we would have to separately discriminate between “even and odd” states and sum them accordingly. Ultimately, we would arrive to the same conclusions!

⁴The quantisation condition follows from the periodic boundary condition, $\phi[\mathbf{r} + L(m_x \hat{\mathbf{x}} + m_y \hat{\mathbf{y}} + m_z \hat{\mathbf{z}})] = \phi(\mathbf{r})$, where $\mathbf{m} = (m_x, m_y, m_z)$ denotes an arbitrary vector of integers.

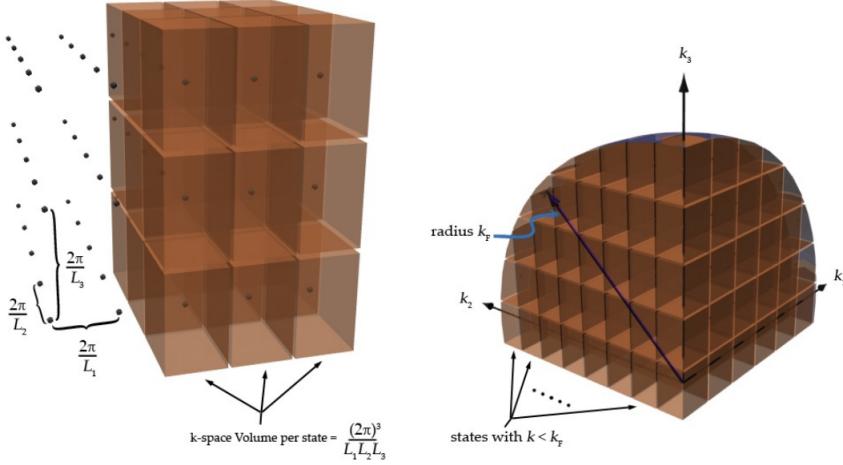


Fig. 10.2: (Left) Schematic showing the phase space volume associated with each plane wave state in a Fermi gas. (Right) Schematic showing the state occupancy of a filled Fermi sea.

Therefore the energy per particle is

$$\frac{E_{\text{tot}}}{\bar{n}} = \frac{E_{\text{tot}}}{\bar{n}L^3} = \frac{3}{5}E_F. \quad (10.21)$$

The groundstate wavefunction is the Slater determinant

$$\Psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_n) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \phi_{\mathbf{k}_1}(\mathbf{r}_1) & \phi_{\mathbf{k}_2}(\mathbf{r}_1) & \cdots & \phi_{\mathbf{k}_N}(\mathbf{r}_1) \\ \phi_{\mathbf{k}_1}(\mathbf{r}_2) & \phi_{\mathbf{k}_2}(\mathbf{r}_2) & \cdots & \phi_{\mathbf{k}_N}(\mathbf{r}_2) \\ \vdots & \vdots & & \vdots \\ \phi_{\mathbf{k}_1}(\mathbf{r}_N) & \phi_{\mathbf{k}_2}(\mathbf{r}_N) & \cdots & \phi_{\mathbf{k}_N}(\mathbf{r}_N) \end{vmatrix} \quad (10.22)$$

formed using the plane wave states with wavevectors $|\mathbf{k}_i| \leq k_F$ (i.e. those of the filled Fermi sea).

10.4.2 Interacting Fermi Gas

For realistic metals or (doped) semiconductors, it is unreasonable to treat the electrons as non-interacting fermions. After all, they have electric charge e , so interact via the Coulomb interactions. As a result, in addition to the kinetic energy (10.16) in the Hamiltonian one should add an interaction⁵

$$\hat{V}_1 = \frac{1}{2} \sum_{i \neq j} \frac{1}{4\pi\epsilon_0} \frac{e^2}{|\mathbf{r}_i - \mathbf{r}_j|} + \sum_i V_{\text{bg}}(\mathbf{r}_i), \quad (10.23)$$

The first term describes the electron-electron repulsion. In the second term, $V_{\text{bg}}(\mathbf{r})$ is the average (attractive) background potential due to the charged ions; this is required in order to make the system overall charge neutral.

⁵The electrons move in the background potential of the ionic cores of the atoms. This is a periodic potential acting on each electron, and leads to a more complicated band structure: with a new “effective mass” m^* and with band gaps at the Brillouin zone boundaries, *etc.* This can have important quantitative effects on the properties of the metal. Here we focus on the *qualitative* effects of electron-electron interactions. Therefore, for simplicity, we neglect this periodic potential.

Let us estimate how large this interaction energy is. For a gas with density \bar{n} , the typical interparticle spacing is $\bar{a} \equiv \bar{n}^{-1/3}$. Therefore, by dimensional analysis, the mean interaction energy per particle must be of order

$$\frac{\langle \hat{V}_1 \rangle}{N} \sim \frac{e^2}{4\pi\epsilon_0\bar{a}} \quad (10.24)$$

(We shall discuss the size of the prefactor below.) On the other hand, we derived the kinetic energy per particle for the groundstate of the spinless Fermi gas,

$$\frac{\langle \hat{H}_0 \rangle}{N} = \frac{3}{5}E_F \sim \frac{\hbar^2}{m\bar{a}^2} \quad (10.25)$$

where we have used $E_F \sim \hbar^2/(m\bar{a}^2)$.

The ratio of the mean interaction energy to the mean kinetic energy is

$$\frac{\langle \hat{V}_1 \rangle}{\langle \hat{H}_0 \rangle} \sim \frac{e^2/(4\pi\epsilon_0\bar{a})}{\hbar^2/(m\bar{a}^2)} = \frac{\bar{a}}{a_0} \quad (10.26)$$

where $a_0 = \frac{4\pi\hbar^2\epsilon_0}{me^2}$ is the Bohr radius. We can understand the properties of the interacting Fermi gas in two simple limiting cases.

1. **Weak interactions:** For $\bar{a} \ll a_0$ the typical interaction energy is small compared to the kinetic energy. In this case, one can treat the interactions as a perturbative correction to the groundstate of the kinetic energy. As discussed above, this groundstate is the filled Fermi sea. For weak interactions, the perturbative corrections are small and the interacting electron gas has the same qualitative properties as the non-interacting (ideal) Fermi gas. This case is typical for “good” metals such as copper.
2. **Strong interactions:** For $\bar{a} \gg a_0$ the typical interaction energy is large compared to the kinetic energy. In this case, one should treat the kinetic energy as a perturbative correction to the groundstate of the interaction potential (10.23). The interaction potential is minimised by keeping the electrons as far apart from each other as possible. At a fixed average density \bar{n} , this is achieved by placing the electrons at points located at the sites of a regular lattice. This unperturbed wavefunction is therefore a *crystal of electrons*. Provided the kinetic energy is small ($\bar{a} \gg a_0$), the perturbation by \hat{H}_0 acts only to move the electrons a small distance from the sites of the crystalline lattice. The integrity of the lattice remains, as a so-called **Wigner crystal**. This limit of strong interactions is *not* achieved in typical metallic systems for which the density is too high. However, in two-dimensional electron gases (formed both in doped semiconductor devices and on the surface of liquid helium) there is evidence for the appearance of a 2D Wigner crystal in experiment.

The two limiting cases described above are simple to understand, since we can use perturbation theory. However, many metallic systems are in an intermediate regime, with $\bar{a} \simeq a_0$, where perturbation theory in either \hat{V}_1 or \hat{H}_0 is inappropriate. This case of intermediate coupling poses a very difficult challenge to theory, and few results are

known with certainty. One technique that one can always fall back on is the **variational** approach. One can “guess” what possible phases of the electron system might appear, and construct trial wavefunctions for each of these phases. By comparing the expectation values of the total energy for these different trial wavefunctions, one obtains an indication of which of them best describes the groundstate and therefore what phase the electron system may be in. For example, this technique has been used to determine at what (numerical) value of \bar{a}/a_0 the transition between the Wigner crystal and Fermi-gas states occurs.

Another important example of the application of variational approach is in understanding **ferromagnetism** of an itinerant spin-1/2 Fermi gas. In this case, the two representative states to consider are: (1) a filled Fermi sea of spin-polarised electrons, say with all spins “up”; and (2) a filled Fermi sea of both spin components. By fixing the occupation numbers to be filled Fermi seas, these states have no remaining variational parameters. A more general class of wavefunctions is provided by “Hartree-Fock” theory, in which the orbitals in the Slater determinants are viewed as variational parameters.

1. Case (1), the spin-polarised Fermi sea, was discussed in detail above. For a density \bar{n} the states are filled out to the Fermi wavevector $k_f = (6\pi\bar{n})^{1/3}$. A direct calculation of the mean energy per particle in this spin-polarised Fermi gas gives

$$\frac{E_{\text{tot}}^{\text{pol.}}}{N} \equiv \frac{\langle \hat{H}_0 \rangle + \langle \hat{V}_1 \rangle}{N} = \frac{3}{5} \frac{\hbar^2}{2m\bar{a}^2} (6\pi^2)^{2/3} - \frac{3(6\pi^2)^{1/3}}{4\pi} \frac{e^2}{4\pi\epsilon_0\bar{a}} \quad (10.27)$$

The prefactor of the interaction term arises from a straightforward but lengthy calculation (not shown here!). The negative sign of the interaction term indicates that the energy is lowered by the electrons residing in the material, and thereby neutralising the background charge density of the ions. The overall size of the energy reduction is influenced by the extent to which the electrons repel each other.

2. Case (2), the spin-unpolarised Fermi gas, follows similar considerations. Note now that a total density \bar{n} will consist of density $\bar{n}/2$ of spin-up electrons and a density $\bar{n}/2$ of spin-down electrons. Since there is no Pauli exclusion between these two spin states, each spin-state can fill its own Fermi sea. The resulting Fermi wavevector (for both spin-up and spin-down) is therefore $K'_F = (6\pi\bar{n}/2)^{1/3} = k_F/2^{1/3}$. A direct calculation of the mean energy per particle in this spin-unpolarised Fermi gas leads to

$$\frac{E_{\text{tot}}^{\text{unpol.}}}{N} \equiv \frac{\langle \hat{H}_0 \rangle + \langle \hat{V}_1 \rangle}{N} = \frac{3}{5} \frac{\hbar^2}{2m\bar{a}^2} (3\pi^2)^{2/3} - \frac{3(6\pi^2)^{1/3}}{4\pi 2^{1/3}} \frac{e^2}{4\pi\epsilon_0\bar{a}} \quad (10.28)$$

Compared to the polarised case, the kinetic energy is smaller. This is due to the fact that, since both spin states are occupied, the electrons are not forced up to such a high Fermi energy as in the polarised case. However, note also that the reduction in the energy due to the interactions is smaller. That is, the total interaction energy of the electrons is larger for the spin-unpolarised case than for the spin-polarised case. For the polarised Fermi gas, the wavefunction must vanish when any two electrons approach each other. This effectively keeps the electrons far apart from each other, and reduces the electron-electron repulsion. For the unpolarised Fermi gas, the wavefunction must vanish when any two particles of the same spin approach each other. However, particles of opposite spin can be located nearby in space, and can contribute a larger Coulomb energy.

For high density (small \bar{a}), the kinetic energy dominates and favours the unpolarised Fermi gas. For low density (large \bar{a}), the interaction energy dominates and favours the polarised Fermi gas. A comparison of the total energies of the two variational states shows that the transition between these is at

$$\frac{\bar{a}}{a_0} = \frac{2\pi}{5}(3\pi^2)^{1/3}(1 + 2^{1/3}) \simeq 8.8. \quad (10.29)$$

This provides a simple variational estimate for the condition **ferromagnetism** in the interacting Fermi gas. This competition is responsible for the ferromagnetism of iron and other metallic magnets. (For even larger values of \bar{a}/a_0 there should be a transition from the polarised Fermi gas into a Wigner crystal.)

10.4.3 Non-interacting Bose gas

For the case of non-interacting bosons, confined in the same box, the nature of the groundstate is even simpler. There is no exclusion principle, so the groundstate is obtained by putting all bosons in the lowest energy single particle state, $\phi_{\mathbf{k}=0}(\mathbf{r})$. This single particle state has vanishing kinetic energy. Thus, the many-particle groundstate has energy

$$E_{\text{tot}} = 0. \quad (10.30)$$

The groundstate wavefunction is

$$\Psi(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N) = \prod_{i=1}^N \phi_{\mathbf{k}=0}(\mathbf{r}_i) \quad (10.31)$$

10.4.4 Interacting Bose Gas

For the BECs formed in cold atomic gases, there are residual interactions between the particles. To a good approximation, these can be represented by contact interactions, with

$$\hat{V}_1 = \frac{1}{2} \sum_{i \neq j} g \delta(\mathbf{r}_i - \mathbf{r}_j). \quad (10.32)$$

A simple variational approach for the interacting Bose gas is provided by the **Gross-Pitaevskii** approximation. One assumes that the groundstate is a condensate, Eq. (10.3), and treats the (normalised) condensate wavefunction $\psi(\mathbf{r})$ as the variational function. The mean energy per particle (for $N \gg 1$) is

$$\frac{\langle \hat{H}_0 + \hat{V}_1 \rangle}{N} = \int \left[\frac{\hbar^2}{2m} |\nabla \psi|^2 + V_{\text{ext}}(\mathbf{r}) |\psi|^2 \right] d^3 \mathbf{r} + \frac{1}{2} g N \int |\psi|^4 d^3 \mathbf{r} \quad (10.33)$$

for the BEC in an external trapping potential $V_{\text{ext}}(\mathbf{r})$. The optimal condensate wavefunction $\psi(\mathbf{r})$ is found by minimising the energy per particle, subject to the normalisation $\langle \psi | \psi \rangle = 1$. This constraint can be imposed by introducing a Lagrange multiplier, $-E \int |\psi|^2 d^3 \mathbf{r}$, with minimisation over ψ then leading to

$$-\frac{\hbar^2}{2m} \nabla^2 \psi(\mathbf{r}) + [V_{\text{ext}}(\mathbf{r}) + gN|\psi|^2] \psi(\mathbf{r}) = E\psi(\mathbf{r}) \quad (10.34)$$

Since $\psi(\mathbf{r})$ is normalised, the mean particle density is $n(\mathbf{r}) = N|\psi|^2$. The interaction term can be viewed as contributing to an effective potential to $V_{\text{total}}(\mathbf{r}) = V_{\text{ext}}(\mathbf{r}) + gn(\mathbf{r})$ in which $n(\mathbf{r})$ is a function of $\psi(\mathbf{r})$ and must therefore be self-consistently determined.

CHAPTER 11

From Molecules to Solids

In the previous section, we studied the quantum mechanics of multi-electron atoms – the subject of atomic physics. In this section, we will begin to explore how these concepts get translated into systems with many atoms, from simple molecular structures to the solid state. As with atomic physics, the subjects of molecular and solid state physics represent fields in their own right and it would be fanciful to imagine that we could do more than touch on the basic principles. Nevertheless, in establishing the foundations of these subjects, we will see two things: firstly, by organising the hierarchy of approximation schemes, much can be understood about the seemingly complex quantum dynamics of many-particle systems. Secondly, we will find that constraints imposed by symmetries (such as translation) allow a simple phenomenology to emerge from complex solid state systems. We begin our discussion with the molecular system.

A molecule consists of electrons moving in the complicated potential set up by all the constituent electric charges. Even in classical mechanics, it would be extremely difficult to solve the equations of motion of the internal molecular degrees of freedom. Fortunately, for most purposes, we can treat the motion of the electrons and nuclei separately, due to their very different masses. As the forces on a nucleus are similar in magnitude to those that act on an electron, so the electrons and nuclei must have comparable momenta. Therefore the nuclei are typically moving much more slowly than the electrons. In studying the motion of the electrons, we can therefore treat the nuclei as being “nailed down” in fixed positions. Conversely, in studying the nuclear motion (vibrations and rotations of the molecule) we can assume, as a first approximation, that the electrons adjust instantly to changes in the *molecular conformation* defined by the positions of the nuclei. This picture forms the basis of the **Born-Oppenheimer approximation**.

In quantum mechanics, the wavefunction of a molecule, $\Psi(\{\mathbf{r}_n\}, \{\mathbf{R}_N\})$, is a function of the positions of all the electrons and nuclei, and the Hamiltonian has the form

$$\hat{H} = \sum_n \frac{\hat{\mathbf{p}}_n^2}{2m_e} + \sum_N \frac{\hat{\mathbf{p}}_N^2}{2m_N} + V(\{\mathbf{r}_n\}, \{\mathbf{R}_N\}), \quad (11.1)$$

where, as usual, the momentum operators, $\hat{\mathbf{p}}_n^2$ and $\hat{\mathbf{p}}_N^2$, act only on the corresponding coordinates. Here we have labelled electrons by lower case letters $n = 1, 2, \dots$ to distinguish them from nuclei which are denoted by capitals, $N = a, b, \dots$. The statement that the electrons and nuclei have comparable momenta translates into the fact that $\nabla_n^2 \Psi$ and $\nabla_N^2 \Psi$ will be comparable. Therefore the second term in the Hamiltonian above, the sum over nuclear kinetic energies, can be neglected as a first approximation when we solve for the dependence of Ψ on the electron position vectors $\{\mathbf{r}_n\}$.

In the Born-Oppenheimer approximation, the time-independent Schrödinger equation for the electronic motion is therefore given by

$$\left[-\sum_n \frac{\hbar^2 \nabla_n^2}{2m_e} + V(\{\mathbf{r}_n\}, \{\mathbf{R}_N\}) \right] \psi_k(\{\mathbf{r}_n\}, \{\mathbf{R}_N\}) = E_k(\{\mathbf{R}_N\}) \psi_k(\{\mathbf{r}_n\}, \{\mathbf{R}_N\}), \quad (11.2)$$

where the eigenfunctions ψ_k , with $k = 0, 1, 2, \dots$, describe the electronic ground state and excited states with the nuclei nailed down at positions $\mathbf{R}_a, \mathbf{R}_b, \dots$; E_k are the corresponding energy levels. Notice that the nuclear positions appear as parameters in ψ_k and E_k . As the molecular conformation is varied by changing $\mathbf{R}_a, \mathbf{R}_b, \dots$, the ground state energy E_0 follows a curve called the molecular potential energy curve and the minimum of this curve defines the equilibrium conformation of the molecule.

11.1 The H_2^+ Ion

The simplest system that exhibits molecular properties is the hydrogen ion H_2^+ , which consists of two protons with positions $\mathbf{R}_a, \mathbf{R}_b$ and one electron at \mathbf{r} . With the potential energy

$$V(\mathbf{r}, \mathbf{R}_a, \mathbf{R}_b) = \frac{e^2}{4\pi\epsilon_0} \left(\frac{1}{|\mathbf{R}_a - \mathbf{R}_b|} - \frac{1}{|\mathbf{r} - \mathbf{R}_a|} - \frac{1}{|\mathbf{r} - \mathbf{R}_b|} \right), \quad (11.3)$$

the Schrödinger equation takes the form,

$$\left[-\frac{\hbar^2 \nabla_r^2}{2m_e} + V(\mathbf{r}, \mathbf{R}_a, \mathbf{R}_b) \right] \psi(\mathbf{r}; \mathbf{R}_a, \mathbf{R}_b) = E\psi(\mathbf{r}; \mathbf{R}_a, \mathbf{R}_b). \quad (11.4)$$

Although this equation can be solved exactly using elliptical polar coordinates, it will be more instructive for our purposes to seek an approximate method of solution. Since there is no obvious parameter in which to develop a perturbative expansion, we will instead follow a variational route to explore the low energy states of system.

If the electron is close to one of the protons, one would expect the other proton to have a small influence on its dynamics, and that the wavefunction in this region would be close to that of a hydrogen atomic orbital. Therefore, in seeking the ground state of the H_2^+ ion, we may take a trial wavefunction that is a linear combination of the ground state ($1s$) wavefunctions centred on the two protons,

$$\psi(\mathbf{r}; \mathbf{R}_a, \mathbf{R}_b) = \alpha\psi_a(\mathbf{r}; \mathbf{R}_a) + \beta\psi_b(\mathbf{r}; \mathbf{R}_b), \quad (11.5)$$

where $\psi_{a,b} = (\pi a_o^3)^{-1/2} \exp[-|\mathbf{r} - \mathbf{R}_{a,b}|/a_0]$, represents the corresponding hydrogenic wavefunction with a_0 the atomic Bohr radius. In this case, the coefficients α and β can be taken as real. The variational expression to be minimised in order to estimate the ground state energy is given by

$$\langle E \rangle = \frac{\langle \psi | \hat{H} | \psi \rangle}{\langle \psi | \psi \rangle} = \frac{\alpha^2 H_{aa} + \beta^2 H_{bb} + 2\alpha\beta H_{ab}}{\alpha^2 + \beta^2 + 2\alpha\beta S} \quad (11.6)$$

where $H_{aa} = \langle \psi_a | \hat{H} | \psi_a \rangle = H_{bb} = \langle \psi_b | \hat{H} | \psi_b \rangle$, and $H_{ab} = \langle \psi_a | \hat{H} | \psi_b \rangle = \langle \psi_b | \hat{H} | \psi_a \rangle$. Note that the matrix elements $H_{aa} = H_{bb}$ because \hat{H} is symmetric with respect to \mathbf{R}_a and \mathbf{R}_b . Moreover, since ψ_a and ψ_b are not orthogonal we have to introduce the **overlap integral**, $S = \langle \psi_a | \psi_b \rangle$, which measures the overlap between the two atomic wavefunctions. In fact we can simplify this expression further, because the potential is symmetric about the midpoint between the two protons. The wavefunction must therefore be either symmetric or antisymmetric, $\alpha = \pm\beta$, and hence,

$$E_0 \leq \langle E \rangle = \frac{H_{aa} \pm H_{bb}}{1 \pm S}. \quad (11.7)$$

The matrix elements in this expression can be evaluated in closed form, though the calculation is rather tedious.¹

We have, therefore, found two possible wavefunctions for the H_2^+ ion,

$$\psi_g = \frac{\psi_a + \psi_b}{\sqrt{2(1+S)}}, \quad \psi_u = \frac{\psi_a - \psi_b}{\sqrt{2(1-S)}}, \quad (11.10)$$

with energies $E_g = \frac{H_{aa}+H_{bb}}{1+S}$, $E_u = \frac{H_{aa}-H_{bb}}{1-S}$. The subscript g refers to the term *gerade* (German for even) used in molecular physics to denote a state that is even under the operation of inverting the electronic wavefunction through the centre of symmetry of the molecule, without changing the positions of the nuclei. Such an inversion changes $\mathbf{r} \rightarrow \mathbf{R}_a + \mathbf{R}_b - \mathbf{r}$, which interchanges ψ_a and ψ_b . Note that this is not the same as parity inversion, which would also affect the nuclear coordinates. The *ungerade* (odd) state is denoted by subscript u . Note that ψ_g and ψ_u are orthogonal, even though ψ_a and ψ_b are not. In fact, ψ_g and ψ_u are just the orthonormal states that diagonalise the Hamiltonian, if we limit ourselves to linear combinations of ψ_a and ψ_b . In chemistry, ψ_g and ψ_u are called **molecular orbitals** and the assumption that they are linear combinations of atomic stationary states is called the **linear combination of atomic orbitals** (LCAO) approximation.

The state ψ_g has the lower energy, while ψ_u represents an excited state of the molecular ion. Physically, the reason for this is that, in the ψ_g state, the two atomic wavefunctions interfere constructively in the region between the protons, giving an enhanced electron density in the region where the electron is attracted strongly by both protons, which serves to screen the two protons from each other. Conversely, in ψ_u we have destructive interference in the region between the protons. If we plot E_g and E_u as functions of the nuclear separation $R = |\mathbf{R}_a - \mathbf{R}_b|$, the results are as shown in Fig. 11.1. For both curves, we have plotted $E + \text{Ry}$ since $-\text{Ry}$ is the ground state energy of the hydrogen atom. The curve of $E_g + \text{Ry}$ in the LCAO approximation has a minimum value of -1.8eV at $R = R_0 = 130\text{pm}$, which is the predicted equilibrium nuclear separation. The predicted energy required to dissociate the molecular ion into a hydrogen atom and a free proton is thus 1.8eV . The curve of $E_u + \text{Ry}$ does not have a minimum, suggesting that the odd wavefunction ψ_u does not correspond to a bound molecular state.

As the variational method provides an upper limit on the ground state energy, it is no surprise that the true molecular potential energy curve (shown dotted) lies below the LCAO one. The true values of the equilibrium separation and dissociation energy are 106pm and 2.8eV respectively. Clearly the LCAO wavefunction ψ_g is not a very accurate approximation to the true ground state. We could improve it by introducing further variational parameters or additional atomic orbitals. For example, when R becomes very small the true wavefunction should approach that of a He^+ ion; we could include such a

¹We simply quote the results here:

$$S = \left(1 + \frac{R}{a_0} + \frac{R^2}{3a_0^2}\right)e^{-R/a_0}, \quad H_{aa} = -\text{Ry} + \frac{e^2}{4\pi\epsilon_0}\left(1 + \frac{R}{a_0}\right)e^{-2R/a_0}, \quad (11.8)$$

$$H_{ab} = \frac{e^2}{4\pi\epsilon_0 a_0}\left(1 + \frac{R}{a_0}\right)e^{-R/a_0} + S\left(-\text{Ry} + \frac{e^2}{4\pi\epsilon_0}\right), \quad (11.9)$$

where $R = |\mathbf{R}_a - \mathbf{R}_b|$ and Ry is the Rydberg constant, the binding energy of the Hydrogen atom in its ground state.

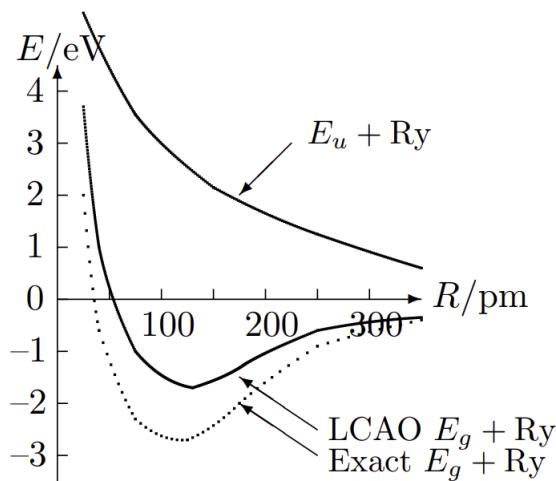


Fig. 11.1: Molecular potential energy curves for H_2^+ ion.

term in the trial function. We could also include an **effective charge parameter** in ψ_a and ψ_b , which is equivalent to replacing the Bohr radius a_0 by a free parameter.

Although not very reliable quantitatively, the LCAO wavefunction ψ_g does however exhibit a number of important features of the true ground state: (i) it is even (g) with respect to inversion of the electron wavefunction; (ii) there is constructive interference which leads to an enhanced probability of finding the electron in the region between the two nuclei. These features will be important when we come to discuss **bonding**. Since the odd u -states are orthogonal to the even g -states, and the true ground state is a g -state, the curve of E_u actually represents an upper limit on the energy of the lowest u -state. Thus the fact that the curve has no minimum does not really prove that there are no bound u -states; but this does turn out to be the case. The LCAO wavefunction ψ_u shows the characteristic feature of an **anti-bonding** state: there is destructive interference in the region between the two nuclei, so that the electron is actually forced out of the region of overlap of the two atomic wavefunctions.

At this stage, it is helpful to introduce some notation to **label the molecular orbitals**. Although the wavefunctions $\psi_{g,u}$ that we have been discussing are formed in the LCAO approximation from linear combinations of atomic $1s$ ($n = 1, \ell = 0$) states, with no orbital angular momentum, they do not themselves necessarily have zero orbital angular momentum. An $\ell = 0$ state must be proportional to Y_{00} , i.e., it must have no dependence on θ and ϕ , giving an isotropic probability distribution. But these states are certainly not isotropic: they have a ‘dumbbell’ shape, concentrated around the two protons. They do not have unique electronic orbital angular momentum because the operator $\hat{\mathbf{L}}^2$ for the electron does not commute with the Hamiltonian on account of the non-central terms $1/|\mathbf{r} - \mathbf{R}_a|$ and $1/|\mathbf{r} - \mathbf{R}_b|$ in the potential.

The only component of $\hat{\mathbf{L}}$ that does commute with these terms is \hat{L}_z , provided we choose the z -axis parallel to the internuclear axis $\mathbf{R}_a - \mathbf{R}_b$. Therefore instead of classifying the states as s, p, d, \dots orbitals according to whether $\ell = 0, 1, 2, \dots$, we call them $\sigma, \pi, \delta, \dots$ orbitals according to whether $\Lambda = 0, 1, 2, \dots$, where $\Lambda = |m_\ell|$. A subscript u or g denotes whether the state is even or odd under inversion; this notation can be applied to all *homonuclear* diatomic molecules, in which the potential is symmetric about the median

plane of the molecule. Thus the ground state of the hydrogen molecular ion is σ_g and the corresponding odd state is σ_u^* , where the star signifies an antibonding orbital. In the LCAO approximation used above, these molecular orbitals are linear combinations of $1s$ atomic orbitals and so they can be written as $1s\sigma_g$ and $1s\sigma_u^*$. To get some insight into how this notation applies, it is helpful to refer to figure 11.2.

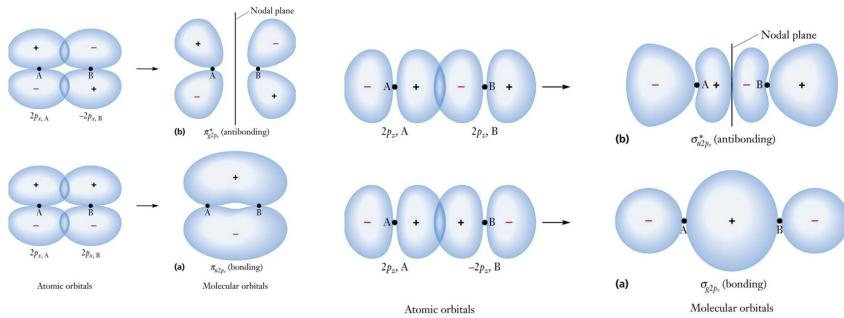


Fig. 11.2: Example notation for the labelling of homonuclear diatomic molecular orbitals.

11.2 The H_2 molecule

We now turn to consider the hydrogen molecule. One might imagine that this would be a trivial extension of the H_2^+ ion, but in fact several new features arise when we consider this simple molecule. For two electrons with positions $\mathbf{r}_{1,2}$ and two protons at $\mathbf{R}_{a,b}$, in the Born-Oppenheimer approximation, the Hamiltonian is given by

$$\hat{H} = -\frac{\hbar^2}{2m_e} \left(\nabla_{r_1}^2 \nabla_{r_2}^2 \right) + \frac{e^2}{4\pi\epsilon_0} \left[\frac{1}{r_{ab}} + \frac{1}{r_{12}} - \frac{1}{r_{1a}} - \frac{1}{r_{1b}} - \frac{1}{r_{2a}} - \frac{1}{r_{2b}} \right], \quad (11.11)$$

where $r_{1b} = |\mathbf{r}_1 - \mathbf{R}_b|$, etc. This is just the sum of two Hamiltonians for H_2^+ ions, plus the additional term $\frac{e^2}{4\pi\epsilon_0} \left(\frac{1}{r_{12}} - \frac{1}{r_{ab}} \right)$. It is plausible that the expectation values of $1/r_{12}$ and $1/r_{ab}$ will be comparable and therefore the extra term can be treated as a perturbation. Thus, as a first approximation we neglect it and assign each electron to one of the H_2^+ molecular orbitals defined above. There are four ways of filling the two orbitals σ_g and σ_u^* , which we can represent by

$$\psi_g(\mathbf{r}_1)\psi_g(\mathbf{r}_2), \quad \psi_g(\mathbf{r}_1)\psi_{u^*}(\mathbf{r}_2), \quad \psi_{u^*}(\mathbf{r}_1)\psi_g(\mathbf{r}_2), \quad \psi_{u^*}(\mathbf{r}_1)\psi_{u^*}(\mathbf{r}_2). \quad (11.12)$$

Of these, we would expect the first to be the ground state.

However, at this stage, we have given no consideration to the constraints imposed by particle statistics. In fact, since the electrons are identical fermions, the wavefunction must be antisymmetric with respect to their interchange. Taking into account the spin degree of freedom, for both electrons to occupy the bonding σ_g molecular orbital, they must have opposite spins and occupy the singlet spin state, $X_{0,0} = \frac{1}{\sqrt{2}}(\chi_+(1)\chi_-(2) - \chi_+(1)\chi_-(2))$.

If we calculate the energy of the state $\psi_g(\mathbf{r}_1)\psi_g(\mathbf{r}_2)X_{0,0}$ as a function of the nuclear separation R , the minimum ground state energy occurs at $R_0 = 85\text{pm}$ and corresponds to a binding energy of 2.7eV . The true molecule is smaller and more tightly bound. Allowing

Approximation	B.E./eV	R_0/pm
Experiment	4.75	74.1
MO fixed charge	2.70	85.2
MO variable charge	3.49	73.0
VB variable charge	3.68	74.6
Variable λ and charge	4.03	75.7
13 variable parameters	4.72	74.1

Table 11.1: Binding energy and equilibrium nuclear separation of the hydrogen molecule in various approximations.

for more variation in the atomic orbitals, in the form of a variable effective charge, gives an equilibrium value of R_0 much closer to experiment, but a binding energy that is still not high enough (see Table 11.1). The reason is that the σ_g^2 configuration alone is not a very good representation of the ground state. To understand why, consider the following.

If we multiply out the spatial part of the σ_g^2 wavefunction in the LCAO approximation, we see that it has a rather strange form,

$$\psi_g(\mathbf{r}_1)\psi_g(\mathbf{r}_2) \propto [\psi_a(\mathbf{r}_1)\psi_b(\mathbf{r}_2) + \psi_b(\mathbf{r}_1)\psi_a(\mathbf{r}_2)] + [\psi_a(\mathbf{r}_1)\psi_a(\mathbf{r}_2) + \psi_b(\mathbf{r}_1)\psi_b(\mathbf{r}_2)]. \quad (11.13)$$

The terms in the first square bracket correspond to the two electrons being shared between the two hydrogen atoms. This is the **covalent bonding** picture of the bound state. In the other square bracket, however, both electrons are assigned to the same atom, corresponding to **ionic bonding**. Since all the terms have equal coefficients, the ionic and covalent contributions are equal, which seems rather constraining, if not implausible. For example, it means that, when the two protons are pulled apart, the system is just as likely to be found to consist of an H^+ and an H^- ion as two neutral atoms.

If we go to the pure **valence bonding** (VB) approximation and drop the ionic part of the wavefunction altogether, we find that the predicted binding energy and nuclear separation are both improved (see Table 11.1). Including a small parameter λ for the amplitude of the ionic component, i.e., taking

$$\psi^{\text{VB}} \propto [\psi_a(\mathbf{r}_1)\psi_b(\mathbf{r}_2) + \psi_b(\mathbf{r}_1)\psi_a(\mathbf{r}_2)] + \lambda[\psi_a(\mathbf{r}_1)\psi_a(\mathbf{r}_2) + \psi_b(\mathbf{r}_1)\psi_b(\mathbf{r}_2)]. \quad (11.14)$$

we find that the variational method gives an optimal value of λ of about 1/6, meaning only about a 3% probability of finding the ionic configuration. Still, even with this refinement, the agreement with experiment is far from perfect. However, this doesn't mean quantum mechanics is failing; by taking enough free parameters in the trial function an excellent result can be obtained, as shown in the Table 11.1.

As mentioned above, there are four possible ways of putting two electrons into the $1s\sigma_g$ and $1s\sigma_u^*$ molecular orbitals. From these four two-electron states we can make three states that are symmetric under interchange of the positions of the electrons, all of which need to be combined with the antisymmetric spin state $X_{0,0}$,

$${}^1\Sigma_g : X_{0,0}\sigma_g(\mathbf{r}_1)\sigma_g(\mathbf{r}_2) \quad (11.15)$$

$${}^1\Sigma_u : X_{0,0}[\sigma_g(\mathbf{r}_1)\sigma_u^*(\mathbf{r}_2) + \sigma_u^*(\mathbf{r}_1)\sigma_g(\mathbf{r}_2)]/\sqrt{2} \quad (11.16)$$

$${}^1\Sigma_g : X_{0,0}\sigma_u^*(\mathbf{r}_1)\sigma_u^*(\mathbf{r}_2). \quad (11.17)$$

In addition we can make a triplet of states from an antisymmetric spatial wavefunction and the symmetric triplet spin states X_{1,m_S} ($m_S = 0, \pm 1$),

$$^3\Sigma_u : X_{0,0}[\sigma_g(\mathbf{r}_1)\sigma_u^*(\mathbf{r}_2) - \sigma_u^*(\mathbf{r}_1)\sigma_g(\mathbf{r}_2)]/\sqrt{2}. \quad (11.18)$$

We have introduced to the left of these equations a new notation, called the molecular term, to describe the overall quantum numbers of the molecule, all of which must be good quantum numbers because they correspond to operators which commute with the molecular Hamiltonian. It derives from a historic spectroscopic notation used in atomic physics. The term is written $^{2S+1}\Lambda_{u/g}$. The prefix $2S + 1$ denotes the multiplicity of the total spin state of the electrons, hence 1 for a singlet and 3 for a triplet state. The central greek capital letter represents the magnitude of the total \hat{L}_z quantum number, $\Lambda = 0, 1, 2, \dots$, being represented by $\Sigma, \Pi, \Delta, \dots$. In the case of the molecular orbitals above based on the 1s atomic orbitals, all the wavefunctions clearly have zero orbital angular momentum about the internuclear axis and hence they are all Σ states. The g or u suffix means even or odd under inversion, and is only meaningful for homonuclear molecules. Notice that since $(\pm 1)^2 = +1$ we get a g -state by combining two g or two u -states, and a u -state by combining a g and a u -state.

In terms of the molecular orbital approach, the VB approximation implies that the ground state is not simply the σ_g^2 configuration but rather a mixture of the two $^1\Sigma_g$ states,

$$\psi^{\text{VB}} \propto (1 + \lambda)(1 + S)\sigma_g(\mathbf{r}_1)\sigma_g(\mathbf{r}_2) - (1 - \lambda)(1 - S)\sigma_u^*(\mathbf{r}_1)\sigma_u^*(\mathbf{r}_2). \quad (11.19)$$

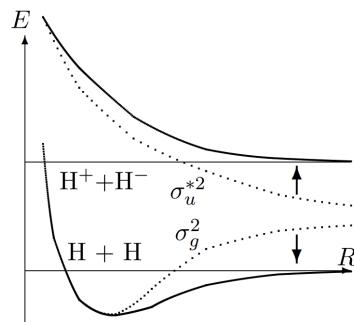


Fig. 11.3: Molecular potential energy and configuration mixing in H_2 .

Thus there is **configuration mixing** in the hydrogen molecule. The two states can mix because they have the same overall quantum numbers. At large nuclear separations the energy eigenstates are clearly separated into those that are almost a pair of neutral atoms and those consisting of an H^+ and an H^- ion. In the case of the u -states, we can see by expanding the spatial wave functions that $^3\Sigma_u$ is of the former type and $^1\Sigma_u$ of the latter. Since these configurations have different resultant electron spins they do not mix significantly (just like singlet and triplet states of atomic helium). Also the u - and g -states are prevented from mixing by their different inversion symmetry. However, the two $^1\Sigma_g$ configurations have the same electron spin and symmetry and can mix to give the above covalent-bonded ground state and an orthogonal excited state that is more ionic.

11.3 From molecules to solids

With these basic principles in hand, we could go on to discuss the orbital and electronic structure of more complicated molecules. In doing so, we would sink deeper into the realm of quantum chemistry. Instead, we will use these simple ideas of molecular bonding to develop a surprisingly versatile and faithful description of the ordered solid state. With *ca.* 10^{23} nuclei and electrons now involved such an enterprise seems foolhardy. However, we will see that, by exploiting symmetries, we can capture many of the basic principles of the solid state.

Let us then consider the electronic structure of an ordered crystalline array of equivalent atoms. (The consideration of more complicated periodic structures would not present conceptual challenges, but would bring unnecessary complications to the discussion.) Once again, we can draw upon the Born-Oppenheimer approximation and focus solely on the motion of electrons around an otherwise ordered array of nuclei. Even then, we are confronted with a many-particle Hamiltonian of apparently great complexity,

$$\hat{H} = - \sum_n \frac{\hbar^2 \nabla_n^2}{2m_e} - \sum_{n,N} \frac{1}{4\pi\epsilon_0} \frac{Ze^2}{r_{nN}} + \sum_{m < n} \frac{1}{4\pi\epsilon_0} \frac{e^2}{r_{mn}}. \quad (11.20)$$

Here the second term represents the interaction of the electrons with the constituent nuclei, and the third term involves the electron-electron interaction. In a physical system, we would have to take into account the influence of further relativistic corrections which would introduce additional spin-orbit couplings.

To address the properties of such a complex interacting system, we will have to draw upon many of the insights developed previously. To begin, it is helpful to partition the electrons into those which are bound to the core and those which are able to escape the potential of the individual atomic nuclei and propagate “freely” through the lattice. The electrons which are tightly bound to the nuclei screen the nuclear charge leading to a modified nuclear potential, $V_{\text{eff}}(r)$. Focussing on those electrons which are free, the effective Hamiltonian can be written as, $\hat{H} \simeq \sum_n \hat{H}_n + \sum_{m,n} \frac{e^2}{4\pi\epsilon_0} \frac{1}{r_{mn}}$, where

$$\hat{H}_n = - \frac{\hbar^2 \nabla_n^2}{2m_e} + V_{\text{eff}}(r) \quad (11.21)$$

represents the single-particle Hamiltonian experienced by each of the electrons – i.e. \hat{H}_n describes the motion of an electron moving in a periodic lattice potential, $V_{\text{eff}}(\mathbf{r}) = V_{\text{eff}}(\mathbf{r} + \mathbf{R})$ with \mathbf{R} belonging to the set of periodic lattice vectors.

Despite engineering this approximation, we are still confronted by a challenging many-particle problem. Firstly, the problem remains coupled through the electron-electron Coulomb interaction. Secondly, the electrons move in a periodic potential. However, if we assume that the electrons remain mobile – the jargon is **itinerant** – and free to propagate through the lattice, they screen each other and diminish the effect of the Coulomb interaction.² Therefore, in these circumstances, we can proceed by neglecting the Coulomb interaction altogether. Of course, we still have to contend with the constraints placed by Pauli exclusion and wavefunction antisymmetry. But we are now in the realm of

²This screening is effective for $\bar{a} \ll a_0$, in the notation of section 10.4.3.

the molecular orbital theory, and can proceed analogously using the variational LCAO approach.

In particular, we can build a trial wavefunction by combining orbital states of the single ion, $V_{\text{ion}}(r)$, where $V_{\text{eff}}(\mathbf{r}) = \sum_{\mathbf{R}} V_{\text{ion}}(\mathbf{r} - \mathbf{R})$. As with the hydrogen molecule, the Hamiltonian for the individual nuclei, $\hat{H}_0 = -\frac{\hbar^2 \nabla^2}{2m_e} + V_{\text{ion}}(\mathbf{r})$ are associated with a set of atomic orbitals, ψ_q , characterised by a set of quantum numbers, q . In the atomic limit, when the atoms are far-separated, these will mirror the simple hydrogenic states. To find the variational ground state of the system, we can then build a trial state from a linear combination of these atomic orbitals. Taking only the lowest orbital, $q = 0$, into account we have,

$$\Psi(\mathbf{r}) = \sum_{\mathbf{R}} \alpha_{\mathbf{R}} \psi(\mathbf{r} - \mathbf{R}), \quad (11.22)$$

where, as before, $\alpha_{\mathbf{R}}$ represent the set of variational coefficients, one for each site (and, in principle, each atomic orbital if we had taken more than one!).

Once again, we can construct the variational state energy,

$$\langle E \rangle = \frac{\langle \Psi | \hat{H}_{1e} | \Psi \rangle}{\langle \Psi | \Psi \rangle} = \frac{\sum_{\mathbf{R}, \mathbf{R}'} \alpha_{\mathbf{R}}^* H_{\mathbf{R}, \mathbf{R}'} \alpha_{\mathbf{R}'}}{\sum_{\mathbf{R}, \mathbf{R}'} \alpha_{\mathbf{R}}^* S_{\mathbf{R}, \mathbf{R}'} \alpha_{\mathbf{R}'},} \quad (11.23)$$

where, as before, $H_{\mathbf{R}, \mathbf{R}'} = \int d^d r \psi^*(\mathbf{r} - \mathbf{R}) \hat{H}_{1e} \psi(\mathbf{r} - \mathbf{R}')$ denote the matrix elements of the orbital wavefunction on the Hamiltonian $\hat{H}_{1e} = -\frac{\hbar^2 \nabla^2}{2m_e} + V_{\text{eff}}(\mathbf{r})$ and $S_{\mathbf{R}, \mathbf{R}'} = \int d^d r \psi^*(\mathbf{r} - \mathbf{R}) \psi(\mathbf{r} - \mathbf{R}')$ represent the overlap integrals. Then, varying the energy with respect to $\alpha_{\mathbf{R}}^*$, we find that the coefficients obey the **secular equation**

$$\sum_{\mathbf{R}'} [H_{\mathbf{R}, \mathbf{R}'} - ES_{\mathbf{R}, \mathbf{R}'}] \alpha_{\mathbf{R}'} = 0. \quad (11.24)$$

Note that, if the basis functions were orthogonal, this would just be an eigenvalue equation.

To develop this idea, let us first see how the method relates back to the problem of H_2^+ : In this case, the secular equation translates to the 2×2 matrix equation,

$$\begin{pmatrix} H_{aa} - E & H_{ab} - ES \\ H_{ab} - ES & H_{bb} - E \end{pmatrix} \boldsymbol{\alpha} = \mathbf{0}. \quad (11.25)$$

where the notation and symmetries of matrix elements follow from section 11.1. As a result, we find that the states $\boldsymbol{\alpha}$ divide into even and odd states as expected.

Now let us consider a one-dimensional periodic lattice system. If we assume that the atoms are well-separated, it is evident that both the overlap integrals and matrix elements will decay exponentially fast with orbital separation. The dominant contribution to the energy will then derive from matrix elements coupling only neighbouring states. In this case, the secular equation separates into the coupled sequence of equations,

$$(\varepsilon - E) \alpha_n - (t + ES)(\alpha_{n+1} + \alpha_{n-1}) = 0, \quad (11.26)$$

for each n , where $H_{nn} = \varepsilon$ denotes the atomic orbital energy, $H_{n,n+1} = H_{n+1,n} = -t$ is the matrix element between neighbouring states, $S_{n,n} = 1$ and $S_{n,n+1} = S_{n+1,n} = S$. All other matrix elements are exponentially small. Here we consider a system with N lattice

sites and impose the periodic boundary condition, $\alpha_{n+N} = \alpha + n$. The periodicity of the secular equation suggests a solution of the form $\alpha_n = \frac{1}{\sqrt{N}} e^{ikna}$, where $k = 2\pi m/Na$ denote a discrete set of N wavevectors with m integer lying in the range $-N/2 < m \leq N/2$. (The size of this range, $k_{max} - k_{min}$, is fixed by the magnitude of the **reciprocal lattice vector** for the one-dimensional lattice, $2\pi/a$.) Substitution confirms that this is a solution with energies,

$$E_k = \frac{\epsilon - 2t \cot ka}{1 + 2S \cot ka}. \quad (11.27)$$

The wavevector k parameterises a **band** of states.

Then, according to the LCAO approximation, for a single electron, distributed throughout the system, the lowest energy state is predicted to be the uniform state $\alpha_n = \frac{1}{\sqrt{N}}$ with the energy $E_0 = \frac{\epsilon - 2t}{1 + 2S}$. (Here we have assumed that the matrix element, t , is positive – in the atomic limit, consider why this is a sound assumption.) However, if we suppose that each atom contributes a non-zero fraction of electrons to the system, we must consider the influence of Pauli exclusion and particle statistics.

Since the electrons are fermions, each state k can host two electrons in a total spin singlet configuration. The lowest energy state is then obtained by adding electron pairs sequentially into states of increasing k until all electrons are accounted for. If the maximum k value, known as the Fermi wavevector, k_F , lies within the band of states, elementary excitations of the electrons cost vanishingly small energy leading to **metallic** behaviour. On the other hand, if the system is stoichiometric, with each atom contributing an integer number of electrons, the Fermi wavevector may lie at a **band gap** between two different band of states. In this case, the system has a gap to excitations and the material forms a **band insulator**.

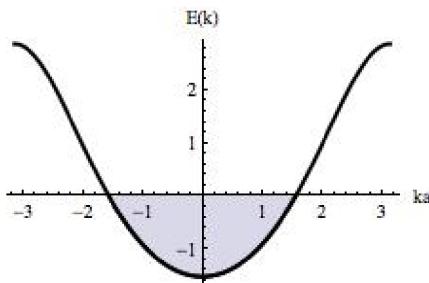


Fig. 11.4: One dimensional band structure for E_k for $S = 0.3$, $t = 1$ and $\epsilon = 0$. Here we have shown schematically a band-filling with $k_F a = \pi/2$.

To add flesh to these ideas, let us then consider a simple, but prominent problem from the realm of quantum condensed matter physics. In recent years, there has been great interest in the properties of **graphene**, a single layer of graphite. (These developments were recognised by the award of the 2010 Nobel prize in physics to Geim and Novoselov.) Remarkably, high quality single crystals of graphene can be obtained by running graphite – a pencil! — over an adhesive layer. The resulting electron states of the single layer compound have been of enormous interest to physicists. To understand why, let us implement the LCAO technology to explore the valence electron structure of graphene.

Graphene forms a periodic two-dimensional honeycomb lattice structure with two

atoms per each unit cell. With an electron configuration $1s^22s^22p^2$, the two $1s$ electrons are bound tightly to the nucleus. The remaining $2s$ electrons **hybridise** with one of the p orbitals to form three sp^2 hybridised orbitals. These three orbitals form the basis of a strong covalent bond of σ orbitals that constitute the honeycomb lattice. The remaining electron, which occupies the out-of-plane p_z orbital, is then capable of forming an itinerant band of electron states. It is this band which we now address.

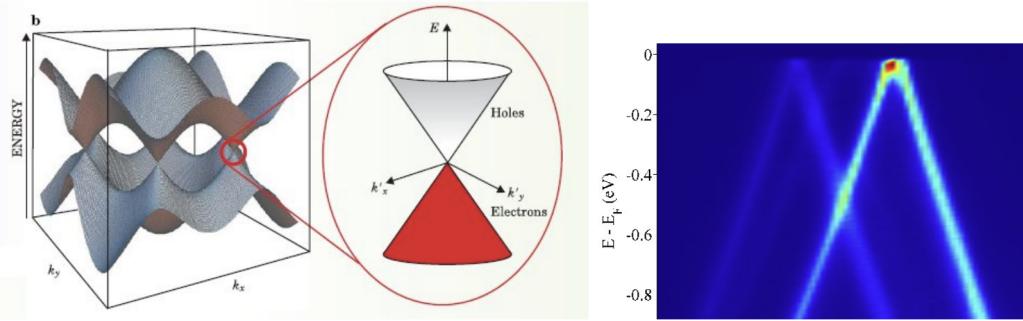


Fig. 11.5: Left: dispersion relation of graphene E_k obtained with the LCAO approximation. Notice that near the centre of the band, the dispersion becomes point-like. Around these points, the dispersion takes form of a linear (Dirac) cone. Right: band structure of a sample of multilayer epitaxial graphene. Linear bands emerge from the K -points in the band structure. Three “cones” can be seen from three layers of graphene in the MEG sample.

Once again, let us suppose that the wavefunction of this band involves the basis of single p_z orbital states localised to each lattice site,

$$\psi(\mathbf{r}) = \sum_{\mathbf{R}} [\alpha_{\mathbf{R}} \psi_1(\mathbf{r} - \mathbf{R}) + \beta_{\mathbf{R}} \psi_2(\mathbf{r} - \mathbf{R})]. \quad (11.28)$$

Here the p_z orbital wavefunction ψ_1 is centred on one of the atoms in the unit cell, and ψ_2 is centred on the other. Once again, taking into account matrix elements involving only nearest neighbours, the trial wavefunction translates to the secular equation,

$$(\varepsilon - E)\alpha_{\mathbf{R}} - (t + ES)(\beta_{\mathbf{R}} + \beta_{\mathbf{R}-\mathbf{a}_1} + \beta_{\mathbf{R}-\mathbf{a}_2}) = 0, \quad (11.29)$$

$$(\varepsilon - E)\beta_{\mathbf{R}} - (t + ES)(\alpha_{\mathbf{R}} + \alpha_{\mathbf{R}-\mathbf{a}_1} + \alpha_{\mathbf{R}-\mathbf{a}_2}) = 0, \quad (11.30)$$

where the lattice vectors $\mathbf{a}_1 = (\sqrt{3}/2, 1/2)a$ and $\mathbf{a}_2 = (\sqrt{3}/2, -1/2)a$, with a the lattice spacing, are shown in figure 11.6. Note that the off-diagonal matrix elements involve only couplings between atoms on different sublattices. Once again, we can make the ansatz that the solutions are of the form of plane waves with $\alpha_{\mathbf{R}} = \frac{\alpha_{\mathbf{k}}}{\sqrt{N}} e^{i\mathbf{k}\cdot\mathbf{R}}$ and $\beta_{\mathbf{R}} = \frac{\beta_{\mathbf{k}}}{\sqrt{N}} e^{i\mathbf{k}\cdot\mathbf{R}}$. Notice that, in this case, we must allow for different relative weights, $\alpha_{\mathbf{k}}$ and $\beta_{\mathbf{k}}$. Substituting, we find that this ansatz is consistent if

$$(\varepsilon - E)\alpha_{\mathbf{k}} - (t + ES)f_{\mathbf{k}}\beta_{\mathbf{k}} = 0 \quad (11.31)$$

$$(\varepsilon - E)\beta_{\mathbf{k}} - (t + ES)f_{\mathbf{k}}^*\alpha_{\mathbf{k}} = 0 \quad (11.32)$$

where $f_{\mathbf{k}} = 1 + 2e^{-i\sqrt{3}k_x a/2} \cos(k_y a/2)$. Although this equation can be solved straightforwardly, it takes a particularly simple form when the overlap integral between neighbouring sites, S , is neglected. In this case, one obtains,

$$E_{\mathbf{k}} = \varepsilon \pm |f_{\mathbf{k}}|t. \quad (11.33)$$

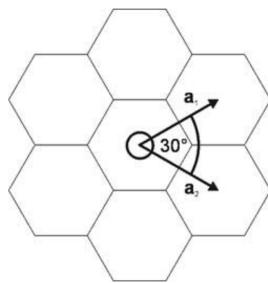


Fig. 11.6: Lattice vectors in periodic graphene lattice.

The corresponding band structure is shown right as a function of \mathbf{k} . In particular, one may note that, at the band centre, the dispersion relation for the electrons becomes point-like (see Fig. 11.5).

In the half-filled system, where each carbon atom contributes a single electron to the band, the Fermi level lies precisely at the centre of the band where the dispersion, $E_{\mathbf{k}}$ is point like. Doping electrons into (or removing electrons from) the system results in (two copies) of a linear dispersion, $E_{\mathbf{k}} \simeq c|\mathbf{k}|$, where c is a constant (velocity). Such a linear dispersion relation is the hallmark of a relativistic particle (cf. a photon). Of course, in this case, the electrons are not moving at relativistic velocities. But their properties will mirror the behaviour of relativistic particles. And herein lies the interest that these materials have drawn.

Finally, let us comment on the influence of the electron-electron interaction effects that were neglected in our treatment above. In principle, we could adopt a Hartree or Hartree-Fock scheme to address the effects of electron interaction in a perturbative manner. Indeed, such a programme would lead to interaction corrections which would modify the electronic band structure derived above. It is a surprising yet robust feature of Fermi systems that the properties of the non-interacting ground state wavefunction remain qualitatively correct over an unreasonably wide range of interaction strengths.³ This rigidity can be ascribed to the constraints on the nodal structure of the wavefunction imposed by particle statistics. However, in some cases, the manifestations of electron interactions translate to striking modifications in the observable properties. We have already discussed the possibility of electron-electron interactions leading to Wigner crystallisation or to itinerant ferromagnetism; other effects include interaction-driven electron localisation (the “Mott transition”), local moment formation, quantum Hall fluids, and superconductivity. Such phases, which by their nature lie outside any perturbative scheme built around the non-interacting ground state, form the field of quantum condensed matter and solid state physics.

³This feature is embodied in the theory of the electron liquid known as **Landau’s Fermi liquid theory**.

CHAPTER 12

Time-Dependent Perturbation Theory

So far, we have focused largely on the quantum mechanics of systems in which the Hamiltonian is time-independent. In such cases, the time dependence of a wavepacket can be developed through the time-evolution operator, $\hat{U} = e^{-i\hat{H}t/\hbar}$, or, when cast in terms of the eigenstates of the Hamiltonian, $\hat{H}|n\rangle = E_n|n\rangle$, as $|\psi(t)\rangle = e^{-i\hat{H}t/\hbar}|\psi(0)\rangle = \sum_n e^{-iE_nt/\hbar}c_n(0)|n\rangle$. Although this framework provides access to any closed quantum mechanical system, it does not describe interaction with an external environment such as that imposed by an external electromagnetic field. In such cases, it is more convenient to describe the induced interactions of a small isolated system, \hat{H}_0 , through a time-dependent interaction $V(t)$. Examples include the problem of magnetic resonance describing the interaction of a quantum mechanical spin with an external time-dependent magnetic field, or the response of an atom to an external electromagnetic field. In the following, we will develop a formalism to treat time-dependent perturbations.

12.1 Time-Dependent Potentials: General Formalism

Consider then the Hamiltonian $\hat{H} = \hat{H}_0\hat{V}(t)$, where all time-dependence enters through the potential $\hat{V}(t)$. In the **Schrödinger representation**, the dynamics of the system are specified by the time-dependent wavefunction, $|\psi(t)\rangle_S$ through the Schrödinger equation $i\hbar|\psi(t)\rangle_S = \hat{H}|\psi(t)\rangle_S$. However, in many cases, and in particular with the current application, it is convenient to work in the **Interaction representation**,¹ defined by

$$|\psi(t)\rangle_I = e^{i\hat{H}_0t/\hbar}|\psi(t)\rangle_S, \quad (12.1)$$

where $|\psi(0)\rangle_I = |\psi(0)\rangle_S$. With this definition, one may show that the wavefunction obeys the equation of motion

$$i\hbar\partial_t|\psi(t)\rangle_I = \hat{V}_I(t)|\psi(t)\rangle_I, \quad (12.2)$$

where $\hat{V}_I(t) = e^{i\hat{H}_0t/\hbar}\hat{V}e^{-i\hat{H}_0t/\hbar}$. Then, if we form the eigenfunction expansion, $|\psi(t)\rangle_I = \sum_n c(t)|n\rangle$, and contract the equation of motion with a general state, $\langle n|$, we obtain

$$i\hbar\dot{c}(t) = \sum_n V_{mn}(t)e^{i\omega_{mn}t}c_n(t), \quad (12.3)$$

where the matrix elements $V_{mn}(t) = \langle m|\hat{V}(t)|n\rangle$, and $\omega_{mn} = (E_m - E_n)/\hbar = \omega_{nm}$. To develop some intuition for the action of a time-dependent potential, it is useful to consider first a periodically-driven two-level system where the dynamical equations can be solved exactly.

The **two-level system** plays a special place in the modern development of quantum theory. In particular, it provides a platform to encode the simplest quantum logic gate,

¹Note how this definition differs from that of the **Heisenberg representation**, $|\psi\rangle_H = e^{i\hat{H}t/\hbar}|\psi(t)\rangle_S$ in which all time-dependence is transferred into the operators

the **qubit**. A classical computer has a memory made up of bits, where each bit holds either a one or a zero. A quantum computer maintains a sequence of qubits. A single qubit can hold a one, a zero, or, crucially, any quantum superposition of these. Moreover, a pair of qubits can be in any quantum superposition of four states, and three qubits in any superposition of eight. In general a quantum computer with n qubits can be in an arbitrary superposition of up to 2^n different states simultaneously (this compares to a normal computer that can only be in one of these 2^n states at any one time). A quantum computer operates by manipulating those qubits with a fixed sequence of quantum logic gates. The sequence of gates to be applied is called a quantum algorithm.

An example of an implementation of qubits for a quantum computer could start with the use of particles with two spin states: $|\downarrow\rangle$ and $|\uparrow\rangle$, or $|0\rangle$ and $|1\rangle$. In fact any system possessing an observable quantity A which is conserved under time evolution and such that A has at least two discrete and sufficiently spaced consecutive eigenvalues is a suitable candidate for implementing a qubit. This is true because any such system can be mapped onto an effective spin-1/2 system.

12.1.1 Dynamics of a Driven Two-Level System

Let us consider a two-state system with

$$\hat{H}_0 = \begin{pmatrix} E_1 & 0 \\ 0 & E_2 \end{pmatrix}, \quad \hat{V}(t) = \begin{pmatrix} 0 & \delta e^{i\omega t} \\ \delta e^{-i\omega t} & 0 \end{pmatrix}. \quad (12.4)$$

Specifying the wavefunction by the two-component vector, $\mathbf{c}(t) = (c_1(t), c_2(t))$, Eq. (12.3) translates to the equation of motion

$$i\hbar\partial_t \mathbf{c} = \delta \begin{pmatrix} 0 & e^{i(\omega-\omega_{21})t} \\ e^{-i(\omega-\omega_{21})t} & 0 \end{pmatrix} \quad (12.5)$$

where $\omega_{21} = (E_2 - E_1)/\hbar$. With the initial condition $c_1(0) = 1$, and $c_2(0) = 0$, this equation has the solution,

$$|c_2(t)|^2 = \frac{\delta^2}{\delta^2 + \hbar^2(\omega - \omega_{21})^2/4} \sin^2 \Omega t, \quad |c_1(t)|^2 = 1 - |c_2(t)|^2, \quad (12.6)$$

where $\Omega = ((\delta/\hbar) + (\omega - \omega_{21})^2/4)^{1/2}$ is known as the **Rabi frequency**. The solution, which varies periodically in time, describes the transfer of probability from state 1 to state 2 and back. The maximum probability of occupying state 2 is a Lorentzian function of frequency ω , with

$$|c_2(t)|_{\max}^2 = \frac{\delta^2}{\delta^2 + \hbar^2(\omega - \omega_{21})^2/4}, \quad (12.7)$$

taking the value of unity at resonance, $\omega = \omega_{21}$.

The dynamics of the driven two-level system finds practical application in the **Ammonia maser**: The ammonia molecule NH₃ has a pyramidal structure with an orientation characterised by the position of the “lone-pair” of electrons sited on the nitrogen atom. At low temperature, the molecule can occupy two possible states, $|A\rangle$ and $|S\rangle$, involving symmetric (S) or antisymmetric (A) atomic configurations, separated by a small energy

splitting, ΔE . (More precisely, along the axis of three-fold rotational symmetry, the effective potential energy of the nitrogen atom takes the form of a double-well. The tunneling of the nitrogen atom through the double well leads to the symmetric and antisymmetric combination of states.) In a time-dependent uniform electric field the molecules experience a potential $V = \boldsymbol{\mu}_d \cdot \mathbf{E}$, where $\mathbf{E} = E\hat{\mathbf{e}}_z \cos \omega t$, and $\boldsymbol{\mu}_d$ denotes the electric dipole moment. Since $\boldsymbol{\mu}_d$ is odd under parity transformation, $P\boldsymbol{\mu}_dP = -\boldsymbol{\mu}_d$, and as $P|A\rangle = -|A\rangle$ and $P|S\rangle = |S\rangle$, the matrix elements of the electric dipole moment are off-diagonal: $\langle S|\boldsymbol{\mu}_d|S\rangle = \langle A|\boldsymbol{\mu}_d|A\rangle = 0$ and $\langle S|\boldsymbol{\mu}_d|A\rangle = \langle A|\boldsymbol{\mu}_d|S\rangle \neq 0$.

If we start with all of the molecules in the symmetric ground state, we have shown above that the action of an oscillating field for a particular time can drive a collection of molecules from their ground state into the antisymmetric first excited state. The ammonia maser works by sending a stream of ammonia molecules, traveling at known velocity, down a tube having an oscillating field for a definite length, so the molecules emerging at the other end are all (or almost all, depending on the precision of in going velocity, etc.) in the first excited state. Application of a small amount of electromagnetic radiation of the same frequency to the outgoing molecules will cause some to decay, generating intense radiation and therefore a much shorter period for all to decay, emitting coherent radiation.

12.1.2 Paramagnetic Resonance

An important example of a dynamical two-level system is provided by a spin-1/2 nucleus under the influence of external magnetic fields. In Section 5.4 we analysed the effect of a uniform magnetic field, and showed that this leads to spin precession. We found that the spin precession frequency is independent of the angle of the spin with respect to the field direction. Consider then how this looks in a frame of reference which is itself rotating with angular velocity ω about the z -axis. Let us specify the magnetic field $\mathbf{B}_0 = B_0\hat{\mathbf{z}}$, since we'll soon be adding another component. In the rotating frame, the observed precession frequency is $\omega_r = -\gamma(\mathbf{B}_0 + \boldsymbol{\omega}/\gamma)$, so there is a different effective field $(\mathbf{B}_0 + \boldsymbol{\omega}/\gamma)$ in the rotating frame. If the frame rotates exactly at the precession frequency, $\boldsymbol{\omega} = \boldsymbol{\omega}_0 = -\gamma\mathbf{B}_0$, spins pointing in any direction will remain at rest in that frame – there is no effective field at all.

Suppose we now add a small rotating magnetic field with angular frequency ω in the xy plane, so the total magnetic field is

$$\mathbf{B} = B_0\hat{\mathbf{z}} + B_1(\hat{\mathbf{r}}_x \cos \omega t - \hat{\mathbf{r}}_y \sin \omega t). \quad (12.8)$$

The effective magnetic field in the frame rotating with the same frequency ω as the small added field is then given by

$$\mathbf{B}_r = (B_0 + \omega/\gamma)\hat{\mathbf{z}} + B_1\hat{\mathbf{e}}_x. \quad (12.9)$$

Now, if we tune the angular frequency of the small rotating field so that it exactly matches the precession frequency in the original static magnetic field, $\boldsymbol{\omega} = \boldsymbol{\omega}_0 = -\gamma\mathbf{B}_0$, all the magnetic moment will see in the rotating frame is the small field in the x -direction! It will therefore precess about the x -direction at the slow angular speed γB_1 . This matching of the small field rotation frequency with the large field spin precession frequency is the “resonance”.

If the spins are lined up preferentially in the z -direction by the static field, and the small resonant oscillating field is switched on for a time such that $\gamma B_1 t = \pi/2$, the spins will be preferentially in the y -direction in the rotating frame, so in the lab they will be rotating in the xy plane, and a coil will pick up an a.c. signal from the induced e.m.f.

12.1.2.1 Nuclear Magnetic Resonance

Nuclear magnetic resonance is an important tool in chemical analysis. As the name implies, the method uses the spin magnetic moments of nuclei (particularly hydrogen) and resonant excitation. **Magnetic resonance imaging** uses the same basic principle to get an image (of the inside of a body for example). In basic NMR, a strong static B field is applied. A spin 1/2 proton in a hydrogen nucleus then has two energy eigenstates. After some time, most of the protons fall into the lower of the two states. We now use an electromagnetic wave (RF pulse) to excite some of the protons back into the higher energy state. The proton's magnetic moment interacts with the oscillating B field of the EM wave through the Hamiltonian

$$\hat{H} = -\boldsymbol{\mu} \cdot \mathbf{B} = \frac{g_p e}{4m_p c} \mathbf{S} \cdot \mathbf{B} = \frac{g_p e \hbar}{4m_p c} \boldsymbol{\sigma} \cdot \mathbf{B} = \frac{g_p}{2} \mu_N \boldsymbol{\sigma} \cdot \mathbf{B}, \quad (12.10)$$

where the gyromagnetic ratio of the proton is about +5.6. The magnetic moment is $2.79\mu_N$ (nuclear magnetons). Different nuclei will have different gyromagnetic ratios giving more degrees of freedom with which to work. The strong static B field is chosen to lie in the z -direction and an oscillating B -field (EM wave) is applied in the xy plane. For an oscillating field that is *circularly* polarised, with $\mathbf{B}(t) = B_+(\cos \omega t, \sin \omega t, 0)$, the dynamics of the two level system are identical to (12.5), with $\omega_1 - \omega_2 = g_p \mu_N B_z / (\hbar)$ (this is the Larmor frequency ω_0) and $\delta = g_p \mu_N B_+ / (2\hbar)$. A *linearly* polarised wave, with $\mathbf{B} = B_x(\cos \omega t, 0, 0)$, may be decomposed as two circularly polarised waves with $B_{\pm} = B_x/2$ (\pm refer to the handedness of the circular polarisations). Provided $\delta \ll \omega_0$, only the circular polarisation that follows the precessing magnetic moment has any significant effect (the other circular polarisation provides a force that oscillates as $e^{\pm i(\omega+\omega_0)t}$ and averages to zero). In the **rotating wave approximation**, we can again model the system according to (12.5), with $\omega_1 - \omega_2 = g_p \mu_N B_z / (\hbar) = \omega_0$ and $\delta = g_p \mu_N (4)$. Note that the resonance condition $\omega = \omega_0$ is exactly the condition that ensures that the energy of a photon of the EM field $E = \hbar\omega$ is equal to the energy difference between the two spin states $\Delta E \hbar\omega_0$. The conservation of energy must be satisfied well enough to get a significant transition rate. In NMR, we observe the transitions back to the lower energy state. These emit EM radiation at the same frequency and we can detect it after the stronger input pulse ends (or by more complex methods).

NMR is a powerful tool in chemical analysis because the molecular field adds to the external B field so that the resonant frequency depends on the molecule as well as the nucleus. We can learn about molecular fields or just use NMR to see what molecules are present in a sample. In MRI, we typically concentrate on the one nucleus like hydrogen. We can put a gradient in B_z so that only a thin slice of the material has ω tuned to the resonant frequency. Therefore we can excite transitions to the higher energy state in only a slice of the sample. If we vary (in the orthogonal direction!) the B field during the decay, we can recover 3d images.

12.2 Time-dependent perturbation theory

We now turn to consider a generic time-dependent Hamiltonian for which an analytical solution is unavailable – sadly the typical situation! In this case, we must turn to a perturbative analysis, looking for an expansion of the basis coefficients $c_n(t)$ in powers of the interaction,

$$c_n(t) = c^{(0)} + c^{(1)}(t) + c^{(2)}(t) + \dots, \quad (12.11)$$

where $c^{(m)} \sim \mathcal{O}(V^m)$ and $c^{(0)}$ is some (time-independent) initial state. The programme to complete this series expansion is straightforward but technical.

In the interaction representation, the state $|\psi(t)\rangle_I$ can be related to an initial state $|\psi(t_0)\rangle_I$ through the time-evolution operator, $U_I(t, t_0)$, i.e. $|\psi(t)\rangle_I = U_I(t, t_0)|\psi(t_0)\rangle_I$. Since this is true for any initial state $|\psi(t_0)\rangle_I$, from Eq. (12.2), we must have

$$i\hbar\partial_t U_I(t, t_0) = V_I(t, t_0)U_I(t, t_0), \quad (12.12)$$

with the boundary condition $U_I(t_0, t_0) = \hat{\mathbb{I}}$. Integrating this equation from t_0 to t , formally we obtain,

$$U_I(t, t_0) = \hat{\mathbb{I}} - \frac{i}{\hbar} \int_{t_0}^t dt' V_I(t') U_I(t', t_0). \quad (12.13)$$

This result provides a *self-consistent* equation for $U_I(t, t_0)$, i.e. if we take this expression and substitute $U_I(t', t_0)$ under the integrand, we obtain

$$U_I(t, t_0) = \hat{\mathbb{I}} - \frac{i}{\hbar} \int_{t_0}^t dt' V_I(t') + \left(-\frac{i}{\hbar}\right)^2 \int_{t_0}^t dt' V_I(t') \int_{t_0}^{t'} dt'' V_I(t'') U_I(t'', t_0) \quad (12.14)$$

Iterating this procedure, we thus obtain

$$U_I(t, t_0) = \sum_{n=0}^{\infty} \left(-\frac{i}{\hbar}\right)^2 \int_{t_0}^{t_1} dt_1 \cdots \int_{t_0}^{t_{n-1}} dt_n V_I(t_1) V_I(t_2) \cdots V_I(t_n), \quad (12.15)$$

where the term $n = 0$ translates to $\hat{\mathbb{I}}$. Note that the operators $V_I(t)$ are organised in a time-ordered sequence, with $t_0 \leq t_n \leq t_{n-1} \leq \dots \leq t_1 \leq t$. With this understanding, we can write this expression more compactly as

$$U_I(t, t_0) = T \left[e^{-\frac{i}{\hbar} \int_{t_0}^t dt' V_I(t')} \right], \quad (12.16)$$

where “ T ” denotes the time-ordering operator and its action is understood by Eq. (12.15).

If a system is prepared in an initial state, $|i\rangle$ at time t_0 , at a subsequent time, t , the system will be in a final state,

$$|i, t_0, t\rangle = U_I(t_0, t)|i\rangle = \sum_i |n\rangle \overbrace{\langle n|U_I(t_0, t)|i\rangle}^{c_n(t)}. \quad (12.17)$$

Making use of Eq. (12.15), and the resolution of identity, $\sum_m |m\rangle\langle m| = \hat{\mathbb{I}}$, we obtain

$$c_n(t) = \delta_{ni} \overbrace{-\frac{i}{\hbar} \int_{t_0}^t dt' \langle n|V_I(t')|i\rangle}^{c_n^{(1)}} \overbrace{-\frac{1}{\hbar^2} \int_{t_0}^t dt' \int_{t_0}^{t'} dt'' \sum_m \langle n|V_I(t')|m\rangle \langle m|V_I(t'')|i\rangle}^{c_n^{(2)}} + \dots \quad (12.18)$$

Recalling that $V_I = e^{i\hat{H}_0 t/\hbar} V E^{-i\hat{H}_0 t/\hbar}$, we thus find that

$$c_n^{(1)}(t) = -\frac{i}{\hbar} \int_{t_0}^t e^{i\omega_{ni} t'} V_{ni}(t'), \quad (12.19)$$

$$c_n^{(2)}(t) = -\frac{1}{\hbar^2} \sum_n \int_{t_0}^t \int_{t_0}^{t'} dt' e^{i\omega_{nm} t' + i\omega_{mi} t''} V_{nm}(t') V_{mi}(t''), \quad (12.20)$$

where $V_{nm}(t) = \langle n | V(t) | m \rangle$ and $\omega_{nm} = (E_n - E_m)/\hbar$, etc. In particular, the probability of effecting a transition from state $|i\rangle$ to state $|n\rangle$ for $n \neq i$ is given by $P_{i \rightarrow n} = |c_n(t)|^2 = |c_n^{(1)}(t) + c_n^{(2)}(t) + \dots|^2$.

12.2.0.1 Example: The Kicked Oscillator

Suppose a simple harmonic oscillator is prepared in its ground state $|0\rangle$ at time $t = -\infty$. If it is perturbed by a small time-dependent potential $V(t) = -eExe^{-t^2/\tau^2}$, what is the probability of finding it in the first excited state, $|1\rangle$, at $t = +\infty$?

Working to first order of perturbation theory, the probability is given by $P_{0 \rightarrow 1} \simeq |c_1^{(1)}|^2$ where $c_1^{(1)} = -\frac{i}{\hbar} \int_{t_0}^t dt' e^{i\omega_{10} t'} V_{10}(t')$, with $V_{10}(t') = -eExe^{-t'^2/\tau^2}$ and $\omega_{10} = \omega$. Using the ladder operator formalism, with $|1\rangle = a^\dagger |0\rangle$ and $x = \sqrt{\frac{\hbar}{2m\omega}}(a + a^\dagger)$, we have $\langle 1 | x | 0 \rangle \sqrt{\frac{\hbar}{2m\omega}}$. Therefore, making use of the identity $\int_{-\infty}^{\infty} dt' \exp[i\omega t' - t'^2/\tau^2] = \sqrt{\pi}\tau \exp[-\omega^2\tau^2/4]$, we obtain the transition amplitude, $c_1^{(1)}(t \rightarrow \infty) = ieE\tau \sqrt{\frac{\pi}{2m\hbar\omega}} e^{-\omega^2\tau^2/4}$. As a result, we obtain the transition probability, $P_{0 \rightarrow 1} \simeq (eE\tau)^2 (\pi/2m\hbar\omega) e^{-\omega^2\tau^2/4}$. Note that the probability is maximised for $\tau \sim 1/\omega$.

12.3 “Sudden” Perturbation

To further explore the time-dependent perturbation theory, we turn now to consider the action of fast or “sudden” perturbations. Here we define sudden as a perturbation in which the switch from one time-independent Hamiltonian \hat{H}_0 to another \hat{H}'_0 takes place over a time much shorter than any natural period of the system. If the system is initially in an eigenstate $|n\rangle$ of \hat{H}_0 , its time evolution following the switch will follow that of \hat{H}'_0 , i.e. one simply has to expand the initial state as a sum over the eigenstates of \hat{H}'_0 , $|n\rangle = \sum_{n'} |n'\rangle \langle n'|n\rangle$. The non-trivial part of the problem lies in establishing that the change is sudden enough. This is achieved by estimating the actual time taken for the Hamiltonian to change, and the periods of motion associated with the state $|n\rangle$ and with its transitions to neighboring states.

12.3.1 Harmonic Perturbations: Fermi’s Golden Rule

Let us then consider a system prepared in an initial state $|i\rangle$ and perturbed by a periodic harmonic potential $\hat{V}(t) = \hat{V}e^{-i\omega t}$ which is abruptly switched on at time $t = 0$. This could

represent an atom perturbed by an external oscillating electric field, such as an incident light wave. What is the probability that, at some later time t , the system lies in state $|f\rangle$?

From Eq. (12.19), to first order in perturbation theory, we have

$$c_f^{(1)}(t) = -\frac{i}{\hbar} \int_0^t dt' \langle f | \hat{V} | i \rangle e^{i(\omega_{fi} - \omega)t'} = -\frac{i}{\hbar} \langle f | \hat{V} | i \rangle \frac{e^{i(\omega_{fi} - \omega)t'} - 1}{i(\omega_{fi} - \omega)}. \quad (12.21)$$

The probability of effecting the transition after a time t is therefore given by

$$P_{i \rightarrow f}(t) \simeq |c_f^{(1)}(t)|^2 = \frac{1}{\hbar^2} |\langle f | \hat{V} | i \rangle|^2 \left(\frac{\sin((\omega_{fi} - \omega)/2)}{(\omega_{fi} - \omega)/2} \right)^2. \quad (12.22)$$

Setting $\alpha = (\omega_{fi} - \omega)/2$, the probability takes the form $\sin^2(\alpha t)/\alpha^2$ with a peak at $\alpha = 0$, with maximum value t^2 and width of order $1/t$ giving a total weight of order t . The function has more peaks positioned at $\alpha t = (n + 1/2)\pi$. These are bounded by the denominator at $1/\alpha^2$. For large t their contribution comes from a range of order $1/t$ also, and as $t \rightarrow \infty$ the function tends towards a δ -function centred at the origin, but multiplied by t , i.e. the likelihood of transition is proportional to time elapsed. We should therefore divide by t to get the transition rate.

Finally, with the normalisation, $\int_{-\infty}^{\infty} d\alpha \left(\frac{\sin \alpha t}{\alpha} \right)^2 = \pi t$, we may effect the replacement, $\lim_{t \rightarrow \infty} \frac{1}{t} \left(\frac{\sin \alpha t}{\alpha} \right)^2 = \pi \delta(\alpha) = 2\pi \delta(2\alpha)$ leading to the following expression for the transition rate,

$$\Gamma_{i \rightarrow f} = \lim_{t \rightarrow \infty} \frac{P_{i \rightarrow f}(t)}{t} = \frac{2\pi}{\hbar^2} |\langle f | \hat{V} | i \rangle|^2 \delta(\omega_{fi} - \omega). \quad (12.23)$$

This expression is known as **Fermi’s Golden Rule**. It is often written with the delta-function applying to the energy, as

$$\Gamma_{i \rightarrow f}(t) = \frac{2\pi}{\hbar^2} |\langle f | \hat{V} | i \rangle|^2 \delta(E_f - E_i - \hbar\omega).$$

(12.24)

One might worry that, in the long time limit, we found that the probability of transition is in fact diverging — so how can we justify the use of perturbation theory? For a transition with $\omega_{fi} \neq \omega$, the “long time” limit is reached when $t \gg 1/(\omega_{fi} - \omega)$, a value that can still be very short compared with the mean transition time, which depends on the matrix element. In fact, Fermi’s Rule agrees extremely well with experiment when applied to atomic systems.

From the expression for the Golden rule (12.24) we see that, for transitions to occur, and to satisfy energy conservation:

1. the final states must exist over a continuous energy range to match $\Delta E = \hbar\omega$ for fixed perturbation frequency ω , or
2. the perturbation must cover a sufficiently wide spectrum of frequency so that a discrete transition with a fixed $\Delta E = \hbar\omega$ is possible.

For two discrete states, since $|V_{fi}|^2 = |V_{if}|^2$, we have the semiclassical result $P_{i \rightarrow f} = P_{f \rightarrow i}$ — a statement of **detailed balance**.

12.3.1.1 Alternative Derivation of the Golden Rule

When light falls on an atom, the full periodic potential is not suddenly imposed on an atomic time scale, but builds up over many cycles (of the atom and of the light). If we assume that $\hat{V}(t) = e^{\varepsilon t} \hat{V} e^{-i\omega t}$, with ε very small, \hat{V} is switched on very gradually in the past, and we are looking at times much smaller than $1/\varepsilon$. We can then take the initial time to be $-\infty$, that is,

$$c_f^{(1)}(t) = -\frac{i}{\hbar} \int_{-\infty}^t \langle f | \hat{V} | i \rangle e^{i(\omega_{fi} - \omega - i\varepsilon)t'} dt' = \frac{1}{\hbar} \frac{e^{i(\omega_{fi} - \omega - i\varepsilon)t}}{(\omega_{fi} - \omega - i\varepsilon)} \langle f | \hat{V} | i \rangle, \quad (12.25)$$

i.e. $|c_f(t)|^2 = \frac{1}{\hbar^2} \frac{e^{2\varepsilon t}}{(\omega_{fi} - \omega)^2 + \varepsilon^2} |\langle f | \hat{V} | i \rangle|^2$. Applied to the transition rate $\frac{d}{dt} |c_f^{(1)}(t)|^2$, the identity $\lim_{\varepsilon \rightarrow 0} \frac{2\varepsilon}{(\omega_{fi} - \omega)^2 + \varepsilon^2} \rightarrow 2\pi\delta(\omega_{fi} - \omega)$ leads to the Golden Rule.

12.3.2 Harmonic Perturbations: Second-Order Transitions

Although first order perturbation theory is often sufficient to describe transition probabilities, sometimes the first order matrix element, $\langle f | \hat{V} | i \rangle$ is identically zero due to symmetry (e.g. under parity, or through some selection rule, etc.), but other matrix elements are non-zero. In such cases, the transition may be accomplished by an indirect route. We can estimate the transition probabilities by turning to the second order of perturbation theory (12.19),

$$c_f^{(2)}(t) = -\frac{1}{\hbar^2} \sum_m \int_{t_0}^t dt' \int_{t_0}^{t'} dt'' e^{i\omega_{fm}t'' + i\omega_{mi}t''} V_{fm}(t') V_{mi}(t''). \quad (12.26)$$

If, as above, we suppose that a harmonic potential perturbation is gradually switched on, $\hat{V}(t) = e^{\varepsilon t} \hat{V} e^{-i\omega t}$, with the initial time $t_0 \rightarrow \infty$, we have

$$c_f^{(2)}(t) = -\frac{1}{\hbar^2} \sum_m \langle f | \hat{V} | m \rangle \langle m | \hat{V} | i \rangle \int_{-\infty}^t dt' \int_{-\infty}^{t'} dt'' e^{i(\omega_{fm} - \omega - i\varepsilon)t'} e^{i(\omega_{mi} - \omega - i\varepsilon)t''}. \quad (12.27)$$

The integrals are straightforward, and yield

$$c_n^{(2)} = -\frac{1}{\hbar^2} e^{i(\omega_{fi} - 2\omega)t} \frac{e^{2\varepsilon t}}{(\omega_{fi} - 2\omega - 2i\varepsilon)} \sum_m \frac{\langle f | \hat{V} | m \rangle \langle m | \hat{V} | i \rangle}{\omega_m - \omega_i - \omega - i\varepsilon}. \quad (12.28)$$

Then, following our discussion above, we obtain the transition rate:

$$\frac{d}{dt} |c_n^{(2)}|^2 = \frac{2\pi}{\hbar^2} \left| \sum_m \frac{\langle f | \hat{V} | m \rangle \langle m | \hat{V} | i \rangle}{\omega_m - \omega_i - \omega - i\varepsilon} \right|^2 \delta(\omega_{fi} - 2\omega). \quad (12.29)$$

This is a transition in which the system gains energy $2\hbar\omega$ from the harmonic perturbation, i.e. two “photons” are absorbed in the transition, the first taking the system to the intermediate energy ω_m , which is short-lived and therefore not well defined in energy – indeed there is no energy conservation requirement for the virtual transition into this state, only between initial and final states. Of course, if an atom in an arbitrary state is exposed to monochromatic light, other second order processes in which two photons are emitted, or one is absorbed and one emitted (in either order) are also possible.

CHAPTER 13

Scattering Theory

Almost everything we know about nuclei and elementary particles has been discovered in scattering experiments, from Rutherford's surprise at finding that atoms have their mass and positive charge concentrated in almost point-like nuclei, to the more recent discoveries, on a far smaller length scale, that protons and neutrons are themselves made up of apparently point-like quarks. More generally, the methods that we have to probe the properties of condensed matter systems rely fundamentally on the notion of scattering. In this section, we will provide a brief introduction to the concepts and methodology of scattering theory

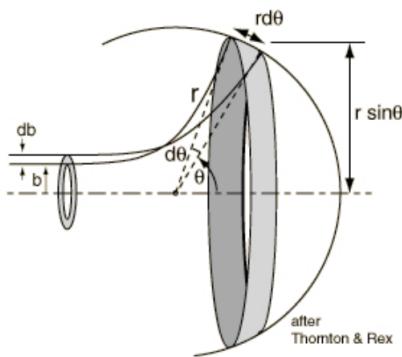


Fig. 13.1: Notation and labelling for classical scattering phenomena.

As preparation for the quantum mechanical scattering problem, let us first consider the classical problem. This will allow us to develop some elementary concepts of scattering theory, and to introduce some notation. In a classical scattering experiment, one considers particles of energy $E = \frac{1}{2}mv_0^2$ (mass m and asymptotic speed v_0), incident upon a target with a central potential $V(r)$. For a repulsive potential, particles are scattered through an angle θ (see Fig. 13.1). The **scattering cross-section**, σ , can be inferred from the number of particles dn scattered into some element of solid angle, $d\Omega$, at angle (θ, ϕ) , i.e. for an incident flux j_i (number of particles per unit time per unit area), $dn = j_i \sigma d\Omega$. The total cross-section is then obtained as $\sigma_T = \int d\Omega \sigma(\theta, \phi) = \int_0^\pi \sin \theta d\theta \int_0^{2\pi} d\phi \sigma(\theta, \phi)$. The angle of deflection of the beam depends on the impact parameter, b (see Fig. 13.1). We therefore have that $dn = j_i b db d\phi = j_i \sigma \sin \theta d\theta d\phi$ and

$$\boxed{\sigma(\theta, \phi) = \frac{b}{\sin \theta} \frac{db}{d\theta}.} \quad (13.1)$$

13.0.0.1 Example: Classic Coulomb Scattering

Let us consider then the case of classical Coulomb scattering from a repulsive potential $V(r) = \frac{\kappa}{r}$ where $\kappa > 0$. From classical physics, we know that the particle will follow a

hyperbolic trajectory with

$$r = \frac{L^2}{m\kappa(e \cos \varphi - 1)}, \quad (13.2)$$

where $\mathbf{r}(r\phi)$ parameterises the relative coordinates of the particle and target, and $e = \left(1 + \frac{2EL^2}{\kappa^2 m}\right)^{1/2} > 1$ denotes the eccentricity. Since the potential is central, the angular momentum L is conserved and can be fixed asymptotically by the condition $L = mv_0 b$. To obtain the scattering angle, θ , we can use the relation above to find the limiting angle, $\cos \varphi_0 = 1/e$, where $\varphi_0 = (\pi - \theta)/2$. We therefore have $\tan \theta/2 = \cot \phi_0 = 1/\sqrt{e^2 - 1} = \left(\frac{m\kappa^2}{2EL^2}\right)^{1/2} = \frac{\kappa}{2Eb}$. Then, from this relation, we obtain the cross-section

$$\sigma(\theta, \phi) = \frac{b}{\sin \theta} \frac{db}{d\theta} = \frac{\kappa^2}{16E^2} \frac{1}{\sin^4 \theta/2}, \quad (13.3)$$

known as the **Rutherford formula**.

13.1 Basics

Let us now turn to the quantum mechanical problem of a beam of particles incident upon a target. The potential of the target, $V(r)$, might represent that experienced by a fast electron striking an atom, or an α particle colliding with a nucleus. As in the classical problem, the basic scenario involves directing a stream or flux of particles, all at the same energy, at a target and detect how many particles are deflected into a battery of detectors which measure angles of deflection. In principle, if we assume that all the in-going particles are represented by wavepackets of the same shape and size, our challenge is to solve the full time-dependent Schrödinger equation for such a wavepacket,

$$i\hbar\partial_t\Psi(\mathbf{r}, t) = \left[-\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{r})\right]\Psi(\mathbf{r}, t), \quad (13.4)$$

and find the probability amplitudes for out-going waves in different directions at some later time after scattering has taken place. However, if the incident beam of particles is switched on for times very long as compared with the time a particle would take to cross the interaction region, steady-state conditions apply. Moreover, if we assume that the wavepacket has a well-defined energy (and hence momentum), so it is many wavelengths long, and we may consider it a plane wave. Setting $\Psi(\mathbf{r}, t) = \psi(\mathbf{r})e^{-iEt/\hbar}$, we may therefore look for solutions $\psi(\mathbf{r})$ of the time-*independent* Schrödinger equation,

$$E\psi(\mathbf{r}) = \left[-\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{r})\right]\psi(\mathbf{r}), \quad (13.5)$$

subject to the boundary condition that the incoming component of the wavefunction is a plane wave, $e^{i\mathbf{k}\cdot\mathbf{x}}$. Here $E = \mathbf{p}^2/2m = \hbar^2\mathbf{k}^2/2m$ denotes the energy of the incoming particles while their flux is given by

$$\mathbf{j} = -i\frac{\hbar}{2m}(\psi^*\nabla\psi - \psi\nabla\psi^*) = \frac{\hbar\mathbf{k}}{m}. \quad (13.6)$$

In the one-dimensional geometry, the impact of a plane wave with the localised target resulted in a portion of the wave being reflected and a portion transmitted through the

potential region. From energy conservation, we may deduce that both components of the outgoing scattered wave are plane waves with wavevector $\pm k$, while the influence of the potential is encoded in the amplitude of the reflected and transmitted beams, and a potential phase shift. Both amplitudes and phase shifts are then determined by solving the time-independent Schrödinger equation subject to the boundary conditions which ensure energy and flux conservation. In the three-dimensional system, the phenomenology is similar: In this case, the wavefunction well outside the localised target region will involve a superposition of the incident plane wave and the scattered wave

$$\psi(\mathbf{r}) \simeq e^{i\mathbf{k}\cdot\mathbf{r}} + f(\theta, \phi) \frac{e^{ikr}}{r}, \quad (13.7)$$

where the function $f(\theta, \phi)$ records the relative amplitude and phase of the scattered components along the direction (θ, ϕ) relative to the incident beam.

The corresponding asymptotic flux is then given by

$$\mathbf{j} = \frac{\hbar}{m} \Im \left\{ \left[e^{i\mathbf{k}\cdot\mathbf{r}} + f(\theta, \phi) \frac{e^{ikr}}{r} \right]^* \nabla \left[e^{i\mathbf{k}\cdot\mathbf{r}} + f(\theta, \phi) \frac{e^{ikr}}{r} \right] \right\}. \quad (13.8)$$

In general, an expansion then leads to a formidable collection of contributing terms. However, for most of these contributions, there remains an exponential factor, $e^{\pm ikr(1-\cos\theta)}$ where θ denotes the angle between \mathbf{k} and \mathbf{r} . For $r \rightarrow \infty$, the small angular integration implied by any physical measurement leads to a fast oscillation of this factor. As a result, such terms are strongly suppressed and can be neglected. Retaining only those terms where the phase cancellation is complete, we obtain

$$\mathbf{j} = \frac{\hbar\mathbf{k}}{m} + \frac{\hbar k}{m} \hat{\mathbf{e}}_r \frac{|f(\theta)|^2}{r^2} + \mathcal{O}(1/r^3). \quad (13.9)$$

The first term represents the incident flux, while the remainder describes the radial flux of scattered particles. In particular, the number of particles crossing the area that subtends a solid angle $d\Omega$ at the origin (the target) is given by

$$\mathbf{j} \cdot \hat{\mathbf{e}}_r dA = \frac{\hbar k}{m} \frac{|f(\theta)|^2}{r^2} r^2 d\Omega + \mathcal{O}(1/r). \quad (13.10)$$

The terms of order $1/r$ can be dropped, as they are negligible in the asymptotic limit.

The **differential cross-section** is defined as

$$\frac{d\sigma}{d\Omega} = \frac{\text{rate of scattering into } d\Omega}{\text{incident flux} \times d\Omega} \quad (13.11)$$

Thus, it is the ratio of the scattered flux $\mathbf{j} \cdot \hat{\mathbf{e}}_r$ times dA to the incident flux $(\hbar k/m)$ times $d\Omega$, giving

$$\frac{d\sigma}{d\Omega} = |f(\theta)|^2.$$

(13.12)

The total cross-section is then given by $\sigma_T = \int d\sigma = \int |f(\theta)|^2 d\Omega$.

13.2 The Born Approximation

We shall consider only the case where the scattering is weak. We can then derive the scattering properties perturbatively.

We view the scattering as a time-dependent process, in which an incident plane wave with wavevector \mathbf{k} is scattered into an outgoing wave $\mathbf{k}' \neq \mathbf{k}$ by the potential $V(r)$. We calculate the rate of these transitions, employing Fermi's golden rule in the limit of zero frequency (since the potential is time-independent). We work with normalised plane wave states, defined on a system of volume L^3 with periodic boundary conditions.

Consider a particle initially in a plane wave state $|\mathbf{k}\rangle$, with position representation

$$\langle \mathbf{r} | \mathbf{k} \rangle = \frac{1}{L^{3/2}} e^{i\mathbf{k} \cdot \mathbf{r}}. \quad (13.13)$$

Note that the state is normalised to unity. Therefore a particle incident in state $|\mathbf{k}\rangle$ has flux $(\hbar\mathbf{k}/m)(1/L^3)$.

By Fermi's golden rule, the rate at which the particle is scattered into states $|\mathbf{k}'\rangle$ within a small solid angle $d\Omega$ is

$$\Gamma_{\mathbf{k} \rightarrow \mathbf{k}' \in d\Omega} = \frac{2\pi}{\hbar} \sum_{\mathbf{k}' \in d\Omega} \left| \langle \mathbf{k}' | \hat{V} | \mathbf{k} \rangle \right|^2 \delta(E_{k'} - E_k) \quad (13.14)$$

where $E_k = \frac{\hbar^2 |\mathbf{k}|^2}{2m}$. The matrix element is

$$\langle \mathbf{k}' | \hat{V} | \mathbf{k} \rangle = \frac{1}{L^3} \int d^3 r V(\mathbf{r}) e^{i(\mathbf{k}-\mathbf{k}') \cdot \mathbf{r}} \equiv \frac{1}{L^3} \tilde{V}(\mathbf{k}) \quad (13.15)$$

which defines \tilde{V} and the **wavevector transfer** $\mathbf{q} \equiv \mathbf{k}' - \mathbf{k}$. The expression for the rate becomes

$$\Gamma_{\mathbf{k} \rightarrow \mathbf{k}' \in d\Omega} = \frac{2\pi}{\hbar} \frac{1}{L^6} \sum_{\mathbf{k}' \in d\Omega} \left| \tilde{V}(\mathbf{k}) \right|^2 \delta(E_{k'} - E_k) \quad (13.16)$$

$$= \frac{2\pi}{\hbar} \frac{1}{L^6} \frac{L^3}{(2\pi)^3} d\Omega \int_0^\infty k'^2 dk' \left| \tilde{V}(\mathbf{k}) \right|^2 \delta \left[\frac{\hbar^2}{2m} (k'^2 - k^2) \right] \quad (13.17)$$

$$= \frac{2\pi}{\hbar} \frac{1}{L^3} \frac{1}{(2\pi)^3} d\Omega k^2 \frac{m}{\hbar^2 k} \left| \tilde{V}(\mathbf{k}) \right|^2 \quad (13.18)$$

with the constraint from the delta-function that $|\mathbf{k}'| = |\mathbf{k}|$. As a result of this constraint, the momentum transfer \mathbf{q} can be written as $|\mathbf{q}| = 2|\mathbf{k}| \sin(\theta/2)$ where θ is the scattering angle.

The differential cross-section is

$$\frac{d\sigma}{d\Omega} = \frac{\text{rate of scattering into } d\Omega}{\text{incident flux } \times d\Omega} \quad (13.19)$$

$$= \frac{\Gamma_{\mathbf{k} \rightarrow \mathbf{k}' \in d\Omega}}{(\hbar k/m)(1/L^3) d\Omega} \quad (13.20)$$

$$= \frac{1}{(2\pi)^2} \frac{mk}{\hbar^3} \frac{1}{L^3} d\Omega \left| \tilde{V}(\mathbf{k}) \right|^2 \times \frac{mL^3}{\hbar k d\Omega} \quad (13.21)$$

$$= \left| \frac{m}{2\pi\hbar} \tilde{V}(\mathbf{k}) \right|^2 \quad (13.22)$$

Thus, we obtain the **Born approximation** expression for the scattering cross section

$$\boxed{\frac{d\sigma}{d\Omega} = \left| \frac{m}{2\pi\hbar^2} \int d^3r V(\mathbf{r}) e^{-i(\mathbf{q}) \cdot \mathbf{r}} \right|^2} \quad (13.23)$$

The scattering cross-section is related to the Fourier transform of the scattering potential.

Since the potential is time-independent one can also derive the scattering cross-section within the Born approximation using standard time-independent perturbation theory. One calculates corrections to the wavefunction of a particle in state $|\mathbf{k}\rangle$, that are first order in \hat{V} , and that involve the other states $|\mathbf{k}'\rangle$. However, note that the scattering involves coupling to degenerate states ($E_{k'} = E_k$). As a result, the perturbative correction has a singular amplitude. When integrating over k' , careful handling of this singularity is required. The physically relevant solution is found by imposing the boundary condition that the correction to the wavefunction involves only *outward* propagating waves.

CHAPTER 14

Radiative Transitions

Previously, we have addressed the quantum theory of atoms coupled to a timeindependent classical electromagnetic field, cf. our discussion of the Zeeman and Stark effects. However, to develop a complete quantum mechanical description of light-matter interaction, we have to address both the quantum theory of the electromagnetic field and the coupling of light to matter. In the following section, we will address both of these issues in turn. Our motivation for developing such a consistent theory is that it will (a) provide us with a platform to address the problem of radiative transitions in atoms and (b) it forms the basis of the field of **quantum optics**.

14.1 Coupling of Matter to the Electromagnetic Field

Let us then consider the Hamiltonian of a single-electron atom subject to a time-dependent external electromagnetic field,

$$\hat{H}_{\text{atom}} = \frac{1}{2m}(\hat{\mathbf{p}} + e\mathbf{A}(\hat{\mathbf{r}}, t))^2 - e\phi(\hat{\mathbf{r}}, t) + V(\hat{\mathbf{r}}). \quad (14.1)$$

Here $V(\mathbf{r})$ denotes the binding potential associated with the atomic nucleus. To keep our discussion of a complex problem as simple as possible, we focus on the single electron system. However, a generalisation of the methodology to multi-electron atoms would not present significant challenges. Expanding the kinetic energy, the atomic Hamiltonian can be recast as $\hat{H}_{\text{atom}} = \hat{H}_0 + \hat{H}_{\text{para}} + \hat{H}_{\text{dia}}$, where

$$\hat{H}_0 = \frac{\hat{\mathbf{p}}^2}{2m} + V(\hat{\mathbf{r}}), \quad (14.2)$$

denotes the usual non-interacting Hamiltonian of the isolated atom,

$$\boxed{\hat{H}_{\text{para}}(t) = \frac{e}{m}\hat{\mathbf{A}}(t) \cdot \hat{\mathbf{p}}}, \quad (14.3)$$

represents the time-dependent paramagnetic term arising from the coupling of the electron to the electromagnetic field, and $\hat{H}_{\text{dia}} = (e\mathbf{A})^2/2m$ represents the diamagnetic term. Since we will be interested in the absorption and emission of single photons, we shall be able to neglect the influence of the diamagnetic term which presents only a small perturbation in the atomic system.

14.1.1 Quantum Fields

As we shall see in section 15.2, the quantum Hamiltonian for the electromagnetic field can be expressed as,

$$\boxed{\hat{H}_{\text{rad}} = \sum_{\mathbf{k}, \lambda=1,2} \hbar\omega_{\mathbf{k}} \left(a_{\mathbf{k}\lambda}^\dagger a_{\mathbf{k}\lambda} + \frac{1}{2} \right)}, \quad (14.4)$$

where the operators $a_{\mathbf{k}\lambda}^\dagger$ and $a_{\mathbf{k}\lambda}$ create and annihilate photons with wavevector \mathbf{k} and polarisation λ , and $\omega_{\mathbf{k}} = c|\mathbf{k}|$. These ladder operators obey the (bosonic) commutation relations, $[a_{\mathbf{k}\lambda}, a_{\mathbf{k}'\lambda'}^\dagger] = \delta_{\mathbf{k},\mathbf{k}'}\delta_{\lambda,\lambda'}$, with $[a_{\mathbf{k}\lambda}, a_{\mathbf{k}'\lambda'}] = [a_{\mathbf{k}\lambda}^\dagger, a_{\mathbf{k}'\lambda'}^\dagger] = 0$, and act on photon number states as

$$a_{\mathbf{k}\lambda} |n_{\mathbf{k}\lambda}\rangle = \sqrt{n_{\mathbf{k}\lambda}} |n_{\mathbf{k}\lambda} - 1\rangle \quad (14.5)$$

$$a_{\mathbf{k}\lambda}^\dagger |n_{\mathbf{k}\lambda}\rangle = \sqrt{n_{\mathbf{k}\lambda} + 1} |n_{\mathbf{k}\lambda} + 1\rangle. \quad (14.6)$$

Here $|n_{\mathbf{k}\lambda}\rangle$ represents a photon number state with $n_{\mathbf{k},\lambda}$ photons in the mode $(\mathbf{k}\lambda)$. In terms of these field operators, the vector potential in the Heisenberg representation can be expanded as

$$\hat{\mathbf{A}}(\mathbf{r}, t) = \sum_{\mathbf{k}, \lambda} \sqrt{\frac{\hbar}{2\epsilon_0\omega_{\mathbf{k}}V}} (\hat{\mathbf{e}}_{\mathbf{k}\lambda} a_{\mathbf{k}\lambda} e^{i(\mathbf{k}\cdot\mathbf{r} - \omega_{\mathbf{k}}t)} + \hat{\mathbf{e}}_{\mathbf{k}\lambda}^* a_{\mathbf{k}\lambda}^\dagger e^{-i(\mathbf{k}\cdot\mathbf{r} - \omega_{\mathbf{k}}t)}). \quad (14.7)$$

Taken together, $\hat{H} = \hat{H}_0 + \hat{H}_{\text{rad}} + \hat{H}_{\text{para}}(t)$ specify the full quantum mechanical Hamiltonian of the atom light system.

14.1.2 Spontaneous Emission

With this background, let us now consider the probability for an atom, initially in a state $|i\rangle$ to make a transition to a state $|f\rangle$ leading to the emission of a photon of wavevector \mathbf{k} and polarisation λ – a process of spontaneous emission. If we suppose that the radiation field is initially prepared in the vacuum state, $|0\rangle$, then the final state involves one photon, $|\mathbf{k}\lambda\rangle = a_{\mathbf{k}\lambda}^\dagger |0\rangle$. Therefore, making use of Fermi's Golden rule (12.24), with the perturbation $\hat{H}_{\text{para}} = \frac{e}{m} \hat{\mathbf{A}}(t) \cdot \hat{\mathbf{p}}$, we have the transition probability

$$\Gamma_{i \rightarrow f}(t) = \frac{2\pi}{\hbar^2} \left| \langle f | \otimes \langle \mathbf{k}\lambda | \hat{H}_{\text{para}} | i \rangle \otimes | 0 \rangle \right|^2 \delta(\omega_{fi} - \omega_{\mathbf{k}}) \quad (14.8)$$

where $\omega_{fi} = (E_i - E_f)/\hbar$. Then substituting the field operator expansion of $\hat{\mathbf{A}}$, we have

$$\Gamma_{i \rightarrow f, \mathbf{k}\lambda} = \frac{2\pi}{\hbar} \left| \langle f | \otimes \langle 0 | a_{\mathbf{k}\lambda} \frac{e}{m} \sqrt{\frac{\hbar}{2\epsilon_0\omega_{\mathbf{k}}V}} \hat{\mathbf{e}}_{\mathbf{k}\lambda}^* a_{\mathbf{k}\lambda}^\dagger e^{-i\mathbf{k}\cdot\mathbf{r}} \cdot \hat{\mathbf{p}} | i \rangle \otimes | 0 \rangle \right|^2 \delta(\omega_{fi} - \omega_{\mathbf{k}}). \quad (14.9)$$

As a result, we finally obtain

$$\Gamma_{i \rightarrow f, \mathbf{k}\lambda} = \frac{2\pi}{\hbar} \left| \langle f | \frac{e}{m} \sqrt{\frac{\hbar}{2\epsilon_0\omega_{\mathbf{k}}V}} \hat{\mathbf{e}}_{\mathbf{k}\lambda}^* e^{-i\mathbf{k}\cdot\mathbf{r}} \cdot \hat{\mathbf{p}} | i \rangle \right|^2 \delta(E_i - E_f - \hbar\omega_{\mathbf{k}}). \quad (14.10)$$

To determine the transition rate, we have to analyse matrix elements of the form $\langle f | e^{-i\mathbf{k}\cdot\mathbf{r}} \hat{\mathbf{e}}_{\mathbf{k}\lambda}^* \cdot \hat{\mathbf{p}} | i \rangle$. Let us begin by estimating its magnitude. For a typical atomic state, $\langle \hat{\mathbf{e}}_{\mathbf{k}\lambda}^* \cdot \hat{\mathbf{p}} \rangle \simeq p \simeq Zmc\alpha$, where we have included a general nuclear charge, Z . But what about the exponential factor? With $r \simeq \hbar/p \simeq \hbar/Zmc\alpha$, and $\omega_{\mathbf{k}} = c|\mathbf{k}| \simeq \frac{p^2}{2m}$ (for electronic transitions), we have

$$\mathbf{k} \cdot \mathbf{r} \simeq \frac{\omega_{\mathbf{k}} \hbar}{c p} \simeq \frac{\hbar p}{mc} \simeq Z\alpha. \quad (14.11)$$

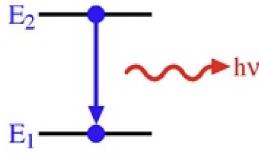


Fig. 14.1: Schematic showing spontaneous emission from an initial state at energy $E_i = E_2$ to a final state at energy $E_f = E_1$.

This means that, for $Z\alpha \ll 1$, we can expand the exponential as a power series in $\mathbf{k} \cdot \mathbf{r}$ with the lowest terms being dominant. Taking the zeroth order term, and making use of the operator identity, $\frac{\dot{\mathbf{P}}}{m} = \frac{i}{\hbar} [\hat{H}_0, \mathbf{r}]$, the matrix element may be written as

$$\langle f | \hat{\mathbf{e}}_{\mathbf{k}\lambda}^* \cdot \hat{\mathbf{p}} | i \rangle = m \hat{\mathbf{e}}_{\mathbf{k}\lambda}^* \cdot \langle f | \frac{i}{\hbar} [\hat{H}_0, \mathbf{r}] | i \rangle = im \frac{E_f - E_i}{\hbar} \hat{\mathbf{e}}_{\mathbf{k}\lambda}^* \cdot \langle f | \mathbf{r} | i \rangle = -im\omega_{\mathbf{k}} \langle f | \hat{\mathbf{e}}_{\mathbf{k}\lambda}^* \cdot \mathbf{r} | i \rangle. \quad (14.12)$$

This result, which emerges from the leading approximation in $Z\alpha$, is known as the **electric dipole approximation**. Effectively, we have set

$$\frac{e}{m} \hat{\mathbf{A}}(\mathbf{r}, t) \cdot \hat{\mathbf{p}} \simeq e \hat{\mathbf{E}}(\mathbf{r}, t) \hat{\mathbf{r}},$$

(14.13)

translating to the potential energy of a dipole, with moment $\mathbf{d} = -e\mathbf{r}$, in an oscillating electric field.

14.1.3 Absorption and Stimulated Emission

Let us now consider the absorption of a photon with wave number \mathbf{k} , and polarisation λ . If we assume that, in the initial state, there are $n_{\mathbf{k}\lambda}$ photons in state $(\mathbf{k}\lambda)$ then, after the transition, there will be $n_{\mathbf{k}\lambda} - 1$. Then, if the initial state of the atom is $|i\rangle$ and the final state is $|f\rangle$, the transition amplitude involves the matrix element,

$$\begin{aligned} \langle f | \otimes \langle (n_{\mathbf{k}\lambda} - 1) | \hat{H}_{\text{para}} | i \rangle \otimes | n_{\mathbf{k}\lambda} \rangle \\ = \langle f | \otimes \langle (n_{\mathbf{k}\lambda} - 1) | \frac{e}{m} \sqrt{\frac{\hbar}{2\epsilon_0\omega_{\mathbf{k}}V}} \hat{\mathbf{e}}_{\mathbf{k}\lambda} a_{\mathbf{k}\lambda} e^{i\mathbf{k}\cdot\mathbf{r}} \cdot \hat{\mathbf{p}} | i \rangle \otimes | n_{\mathbf{k}\lambda} \rangle \end{aligned} \quad (14.14)$$

Then, using the relation $a_{\mathbf{k}\lambda} | n_{\mathbf{k}\lambda} \rangle = \sqrt{n_{\mathbf{k}\lambda}} | n_{\mathbf{k}\lambda} - 1 \rangle$,

$$\langle f | \otimes \langle (n_{\mathbf{k}\lambda} - 1) | \hat{H}_{\text{para}} | i \rangle \otimes | n_{\mathbf{k}\lambda} \rangle = \langle f | \frac{e}{m} \sqrt{\frac{\hbar n_{\mathbf{k}\lambda}}{2\epsilon_0\omega_{\mathbf{k}}V}} e^{i\mathbf{k}\cdot\mathbf{r}} \hat{\mathbf{e}}_{\mathbf{k}\lambda} \cdot \hat{\mathbf{p}} | i \rangle. \quad (14.15)$$

As a result, using Fermi's Golden rule, we obtain the transition amplitude,

$$\Gamma_{i \rightarrow f, \mathbf{k}\lambda} = \frac{2\pi}{\hbar} \left| \langle f | \frac{e}{m} \sqrt{\frac{\hbar n_{\mathbf{k}\lambda}}{2\epsilon_0\omega_{\mathbf{k}}V}} e^{i\mathbf{k}\cdot\mathbf{r}} \hat{\mathbf{e}}_{\mathbf{k}\lambda} \cdot \hat{\mathbf{p}} | i \rangle \right|^2 \delta(E_f - E_i - \hbar\omega_{\mathbf{k}}). \quad (14.16)$$

In particular, we find that the absorption rate *increases* linearly with photon number, $n_{\mathbf{k}\lambda}$.

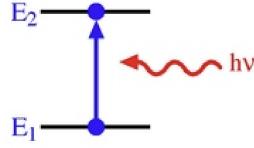


Fig. 14.2: Schematic showing absorption from an initial state at energy $E_i = E_1$ to a final state at energy $E_f = E_2$.

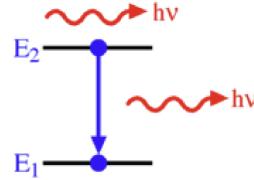


Fig. 14.3: Schematic showing the stimulated emission from an initial state at energy $E_i = E_2$ to a final state at energy $E_f = E_1$.

Similarly, if we now consider the emission process in which there are already $n_{\mathbf{k}\lambda}$ photons in the initial state, we have the revised transition rate,

$$\Gamma_{i \rightarrow f, \mathbf{k}\lambda} = \frac{2\pi}{\hbar} \left| \langle f | \frac{e}{m} \sqrt{\frac{\hbar(n_{\mathbf{k}\lambda} + 1)}{2\epsilon_0 \omega_{\mathbf{k}} V}} e^{-i\mathbf{k} \cdot \mathbf{r}} \hat{\mathbf{e}}_{\mathbf{k}\lambda}^* \cdot \hat{\mathbf{p}} | i \rangle \right|^2 \delta(E_f - E_i - \hbar\omega_{\mathbf{k}}). \quad (14.17)$$

This enhancement of the transition rate by the photon occupancy is known as **stimulated emission**.

Altogether, in the dipole approximation, we have the transition rates,

$$\Gamma_{i \rightarrow f, \mathbf{k}\lambda} = \frac{\pi \omega_{\mathbf{k}}}{\epsilon_0 V} |\langle 2 | \hat{\mathbf{e}}_{\mathbf{k}\lambda} \cdot \mathbf{d} | 1 \rangle|^2 \begin{cases} n_{\mathbf{k}\lambda} \delta(E_f - E_i - \hbar\omega_{\mathbf{k}}) & \text{absorption} \\ (n_{\mathbf{k}\lambda} + 1) \delta(E_i - E_f - \hbar\omega_{\mathbf{k}}) & \text{emission} \end{cases} \quad (14.18)$$

where $|1\rangle$ and $|2\rangle$ are the states with energies E_1 and E_2 (with $E_2 > E_1$). If there are no photons present initially, this expression reduces to that obtained from spontaneous emission. The $n_{\mathbf{k}\lambda}$ -independent component of the expression for absorption and emission coincide, an equality known as **detailed balance**.

If we are interested in the total rate, $d\Gamma_{\lambda}$ at which photons of polarisation λ are scattered into the solid angle $d\Omega$, we must compute $dR_{\lambda} = \sum_{\mathbf{k} \in d\Omega} \Gamma_{i \rightarrow f, \mathbf{k}\lambda}$. Since, in the elemental volume $d^3k = k^2 dk d\Omega$, there are $d^3k V / (2\pi)^3$ states, we may set $\sum_{\mathbf{k}} = \frac{V}{(2\pi)^3} \int k^2 dk d\Omega$. Finally, if we assume that the photon occupation of state $(\mathbf{k}\lambda)$ is isotropic, dependent only on $|\mathbf{k}|$, we find that the integrated transition rate per unit solid angle is given by $\frac{dR_{\lambda}}{d\Omega} = V \int \frac{k^2 dk}{(2\pi)^3} \Gamma_{i \rightarrow f, \mathbf{k}\lambda}$ from which we obtain

$$\frac{dR_{\lambda}}{d\Omega} = \frac{1}{4\pi\epsilon_0} \frac{\omega^3}{2\pi\hbar c^3} |\langle 2 | \hat{\mathbf{e}}_{\mathbf{k}\lambda} \cdot \mathbf{d} | 1 \rangle|^2 \begin{cases} n_{\lambda}(\omega) & \text{absorption} \\ n_{\lambda}(\omega) + 1 & \text{emission} \end{cases} \quad (14.19)$$

Here, in carrying out the integral, we have used the relation $\omega_{\mathbf{k}} = c|\mathbf{k}|$ and $\hbar\omega = |E_f - E_i|$. For a thermal distribution of photons, with the energy density specified by the Planck

formula,

$$u(\omega) = \frac{\hbar\omega^3}{\pi^2 c^3} \bar{n}_\lambda(\omega), \quad \bar{n}_\lambda(\omega) = \frac{1}{e^{\hbar\omega/k_B T} - 1}, \quad (14.20)$$

this equates to an absorption / stimulated emission rate,

$$\frac{dR_\lambda}{d\Omega} = \frac{1}{4\epsilon_0} \frac{1}{2\hbar^2} |\langle 2|\hat{e}_{k\lambda} \cdot \mathbf{d}|1\rangle|^2 u(\omega).$$

(14.21)

From these expressions, we can obtain the power loss as $P_\lambda = \hbar\omega R_\lambda$. Before discussing the selection rules implied by the form of the dipolar coupling, it is first helpful to digress and discuss connections of this result to a famous result due to Einstein.

14.1.4 Einstein's A and B Coefficients

In fact, the frequency dependence of the spontaneous emission rate can be inferred without invoking quantum field theoretic methods by means of an ingenious argument due to Einstein which showed that the stimulated and spontaneous transitions must be related. Consider an ensemble of atoms exposed to a black-body radiation field at temperature T . Let us consider transitions between two states $|\psi_j\rangle$ and $|\psi_i\rangle$, with $E_k - E_j = \hbar\omega$. Suppose the numbers of atoms in the two states are n_j and n_k . The possible transitions and their rates per atom are given by:

	absorption	$j \rightarrow k$	$B_{j \rightarrow k} u(\omega)$
	stimulated emission	$k \rightarrow j$	$B_{k \rightarrow j} u(\omega)$
	spontaneous emission	$k \rightarrow j$	$A_{k \rightarrow j}(\omega)$

(14.22)

where $u(\omega)$, the energy density of radiation per unit ω . A and B are known as Einstein's A and B coefficients, and, as we have seen, are properties of the atomic states concerned.

Now, in thermodynamic equilibrium, the rates must balance, so that

$$n_k [A_{k \rightarrow j}(\omega) + B_{k \rightarrow j} u(\omega)] = n_j B_{j \rightarrow k} u(\omega). \quad (14.23)$$

At the same time, the relative populations of the two states (assumed non-degenerate for simplicity), are given by a Boltzmann factor

$$\frac{n_j}{n_k} = \frac{e^{-E_j/k_B T}}{e^{-E_k/k_B T}} = e^{\hbar\omega/k_B T}. \quad (14.24)$$

Thus we have:

$$A_{k \rightarrow j}(\omega) = [B_{j \rightarrow k} e^{\hbar\omega/k_B T} - B_{k \rightarrow j}] u(\omega) \quad (14.25)$$

For a black-body, the energy density $u(\omega)$ is just given by Planck's formula, $u(\omega) = \frac{\hbar\omega^3}{\pi^2 c^3} \frac{1}{e^{\hbar\omega/k_B T} - 1}$. The $A_{k \rightarrow j}$ coefficient in Eq. (14.25) certainly cannot depend on temperature, so T must cancel on the right hand side. Hence,

$$B_{k \rightarrow j} = B_{j \rightarrow k}, \quad A_{k \rightarrow j}(\omega) = B_{k \rightarrow j} \frac{\hbar\omega^3}{\pi^2 c^3}. \quad (14.26)$$

So, the A and B coefficients are related, and if we can calculate the B coefficient for stimulated emission from Fermi's golden rule, we can infer A , and *vice versa*.

14.2 Selection Rules

It is clear from the formulae for the transition rates that no transition, either spontaneous or stimulated, will occur between the states $|i\rangle$ and $|f\rangle$ unless at least one component of the dipole transition matrix element $\langle f|\hat{\mathbf{d}}|i\rangle$ is non-zero. It is often possible to show that the matrix elements are zero for certain pairs of states. If so, the transition is not allowed (at least in the electric dipole approximation), and the results can often be summarised in terms of **simple selection** rules governing the allowed changes in quantum numbers in transitions.

Since the dipole operator $\hat{\mathbf{d}} = q\hat{\mathbf{r}}$ for the atom¹ changes sign under parity ($\mathbf{r} \rightarrow -\mathbf{r}$), the matrix element $\langle f|\hat{\mathbf{d}}|i\rangle$ will trivially vanish if the states $|f\rangle$ and $|i\rangle$ have the same parity.² Therefore, **the parity of the wavefunction must change in an electric dipole transition.**

To determine further restrictions on the possible transitions induced by electric dipole coupling, we consider first the case in which spin-orbit interactions are absent. Then the energy eigenstates are also eigenstates of the electron spin $\hat{\mathbf{S}}^2$ and its z -component \hat{S}_z , with eigenvalues $s(s+1)\hbar^2$ and $m_s\hbar$.³ Since the dipole operator only acts on the spatial part of the wavefunction, it commutes with the spin: $[\hat{\mathbf{S}}^2, \hat{\mathbf{d}}] = [\hat{S}_z, \hat{\mathbf{d}}] = 0$. Taking the matrix element of the relation $\hat{S}_z\hat{\mathbf{d}} = \hat{\mathbf{d}}\hat{S}_z$ gives $\langle f|\hat{S}_z\hat{\mathbf{d}}|i\rangle = \langle f|\hat{\mathbf{d}}\hat{S}_z|i\rangle$ which leads to $\hbar(m_s^f - m_s^i)\langle f|\hat{\mathbf{d}}|i\rangle = 0$ where $\hbar m_s^{i/f}$ is the eigenvalue of \hat{S}_z for the initial/final state. Hence if the matrix element $\langle f|\hat{\mathbf{d}}|i\rangle$ is to be non-zero, we must have $\Delta m_s = (m_s^f - m_s^i) = 0$. A similar argument holds for the change in total spin, and we find the selection rule

$$\Delta s = 0, \quad \Delta m_s = 0 \quad (14.27)$$

The spin state is not altered in an electric dipole transition.

Let us now consider the selection rules for the orbital angular momenta. From the operator identity, $[\hat{L}_i, r_j] = i\hbar\epsilon_{ijk}r_k$, it follows that

$$[\hat{L}_z, \hat{z}] = 0, \quad [\hat{L}_z, \hat{x} \pm i\hat{y}] = \pm(\hat{x} \pm i\hat{y})\hbar, \quad (14.28)$$

We therefore obtain the relation,

$$\langle \ell', m'_\ell | [\hat{L}_z, \hat{z}] | \ell, m_\ell \rangle = (m'_\ell - m_\ell)\hbar \langle \ell', m'_\ell | \hat{z} | \ell, m_\ell \rangle = 0. \quad (14.29)$$

Similarly, since $\langle \ell', m'_\ell | [\hat{L}_z, \hat{x} \pm i\hat{y}] | \ell, m_\ell \rangle = \pm\hbar \langle \ell', m'_\ell | \hat{x} \pm i\hat{y} | \ell, m_\ell \rangle$, it follows that

$$(m'_\ell - m_\ell \mp 1) \langle \ell', m'_\ell | \hat{x} \pm i\hat{y} | \ell, m_\ell \rangle = 0. \quad (14.30)$$

Therefore, to get non-zero component of the dipole matrix element, we require.

$$\Delta m_\ell = 0, \pm 1. \quad (14.31)$$

¹More generally, for an atom with N electrons we should consider $\mathbf{d} = q \sum_{i=1}^N \mathbf{r}_i$. The conclusions remain unchanged.

²This result may be found in an operator approach, by noting that $\hat{P}\hat{\mathbf{d}} = -\hat{\mathbf{d}}\hat{P}$ where \hat{P} is the parity operator. Hence $\langle f|\hat{P}\hat{\mathbf{d}}|i\rangle = -\langle f|\hat{\mathbf{d}}\hat{P}|i\rangle$ and so $(P_f + P_i)\langle f|\hat{\mathbf{d}}|i\rangle = 0$ where $P_{i,f}$ are the parities (± 1) of the initial/final states. A non-zero matrix element $\langle f|\hat{\mathbf{d}}|i\rangle \neq 0$ requires $P_f = -P_i$, i.e. that the parity changes.

³We consider the possibility of a multi-electron atom so the total spin s need not just be $s = 1/2$.

Similarly, using the operator identity $[\hat{\mathbf{L}}^2, [\hat{\mathbf{L}}^2, \hat{\mathbf{r}}]] = 2\hbar^2(\hat{\mathbf{r}}\hat{\mathbf{L}}^2 + \hat{\mathbf{L}}^2\hat{\mathbf{r}})$, we have

$$\begin{aligned}\langle \ell', m'_\ell | [\hat{\mathbf{L}}^2, \hat{\mathbf{r}}] | \ell, m_\ell \rangle &= [\ell'(\ell' + 1) - \ell(\ell + 1)]^2 \langle \ell', m'_\ell | \hat{\mathbf{r}} | \ell, m_\ell \rangle \\ &= 2[\ell'(\ell' + 1) + \ell(\ell + 1)] \langle \ell', m'_\ell | \hat{\mathbf{r}} | \ell, m_\ell \rangle\end{aligned}\quad (14.32)$$

i.e. $(\ell + \ell')(\ell + \ell' + 2)[(\ell' - \ell)^2 - 1] \langle \ell', m'_\ell | \hat{\mathbf{r}} | \ell, m_\ell \rangle = 0$. Since $\ell, \ell' \geq 0$, we can conclude that, to effect an electric dipole transition, we must have

$$\Delta\ell = \pm 1 \quad (14.33)$$

One may summarise the selection rules for ℓ and m_ℓ by saying that the photon carries off (or brings in, in an absorption transition) one unit of angular momentum. It should be noted, however, that these rules were derived for the specific case of an electric dipole transition of the system. It is possible, though much less likely in the case of an atom, for the electromagnetic field to interact with some other observable such as the magnetic dipole moment or the electric quadrupole moment. In such transitions the selection rules are different. For example, the magnetic dipole operator is $\hat{\mu} = -\mu_B \hat{\mathbf{L}}/\hbar$ (or $-2\mu_B \hat{\mathbf{S}}/\hbar$ for the spin) and since the angular momentum does not change sign under the parity transformation, there is no change of parity in a magnetic dipole transition. To avoid confusion, we shall continue to confine the discussion to electric dipole transitions, which are responsible for the prominent lines in atomic spectra.

For transitions with $\Delta m_\ell = 0$, the dipole matrix element $\langle f | \hat{\mathbf{d}} | i \rangle \sim \hat{\mathbf{e}}_z$, so the only component of polarisation is along the z -direction. Similarly, for electric dipole transitions with $m' = m \pm 1$, $\langle \ell', m'_\ell | x \mp iy | \ell, m_\ell \rangle \neq 0$, and $\langle f | \hat{\mathbf{d}} | i \rangle \sim (1, \pm i, 0)$ lies in the $x-y$ plane. In this case, if the wavevector of photon lies along z , the emitted light is circularly polarised with a polarisation which depends on helicity. Conversely, if the wavevector lies in xy plane, the emitted light is linearly polarised, while in general the polarisation is elliptical.

We conclude by discussing how the selection rules are affected by spin-orbit coupling. In the most general case (i.e. of strong spin-orbit interaction), the stationary states are labelled by their parity and by the quantum numbers j, m_j associated with the total angular momentum (squared), $\hat{\mathbf{J}}^2$, and its z -component \hat{J}_z . In this case, the selection rules can be inferred by looking for the conditions for non-zero matrix elements $\langle j', m'_j | \hat{\mathbf{r}} | j, m_j \rangle$. One finds that, for dipole transitions to take place, we require that

- Parity must change;
- $\Delta j = \pm 1, 0$ (but $0 \rightarrow 0$ is not allowed) and $m_j = \pm 1, 0$.

Atomic states are always eigenstates of parity and of total angular momentum so these selection rules can be regarded as absolutely valid in electric dipole transitions. It should be emphasised again, though, that the electric dipole approximation is an approximation, and higher order processes may occur, albeit at a slower rate, and have their own selection rules.

Even in the presence of spin-orbit coupling, if this coupling is sufficiently weak that the system is described by ideal LS coupling, then additional selection rules apply. Specifically, we also require:

1. $\Delta s = 0$
2. $\Delta\ell = \pm 1, 0$ (but $0 \rightarrow 0$ is not allowed)
3. and $\Delta\ell_i = \pm 1$ if only electron i is involved in the transition.

These follow from the fact that, in LS coupling, the states are eigenstates also of total spin $\hat{\mathbf{S}}^2$ and total orbital angular momentum $\hat{\mathbf{L}}^2$. Since the dipole operator does not operate on the spin part of the wavefunction, then the conditions (14.27) also apply. Combining this and the absolute rules relating to j , imply the rules for ℓ and m_ℓ . The rule for $\Delta\ell_i$ follows from the parity change rule, since the parity of the atom is the product of the parities of the separate electron wavefunctions, given by $(-1)^{\ell_i}$. However, since LS coupling is only an approximation, these rules should themselves be regarded as approximate.

As another example of selection rules, consider a charged particle moving in a one-dimensional harmonic potential. The wavefunctions are characterised by the quantum number n , so the state $|n\rangle$ corresponds to energy $(n + 1/2)\hbar\omega$. An oscillating electric field in the x -direction can induce transitions between states $|n\rangle$ and $|n'\rangle$, governed by matrix elements of the form $\langle n'|x|n\rangle$. These can be evaluated by use of the ladder operators \hat{a} and \hat{a}^\dagger . Since $x = \sqrt{\frac{\hbar}{2m\omega}}(\hat{a} + \hat{a}^\dagger)$, the matrix element becomes

$$\langle n'|x|n\rangle = \sqrt{\frac{\hbar}{2m\omega}} \left(\sqrt{n+1} \langle n'|n+1\rangle + \sqrt{n} \langle n'|n-1\rangle \right) \quad (14.34)$$

and therefore vanishes unless $n = n' \pm 1$. Hence the selection rule, in the electric dipole approximation, is $\Delta n = \pm 1$.

14.3 Lasers

Finally, to close this section, we will consider an important application of light matter interaction – the laser. The laser is the light source which enables modern spectroscopy. The term “laser” is an acronym for “light amplification by stimulated emission of radiation”. However, a laser not only amplifies light, but it acts as a special kind of light source which is characterised by a number of properties:

Monochromaticity: The emission of the laser generally corresponds to just one of the atomic transitions of the gain medium (in contrast to discharge lamps, which emit on all transitions). The spectral line width can be much smaller than that of the atomic transition. This is because the emission is affected by the optical cavity. In certain cases, the laser can be made to operate on just one of the modes of the cavity. Since the Q of the cavity⁴ is generally rather large, the mode is usually much narrower than the atomic transition, and the spectral line width is orders of magnitude smaller than the atomic transition.

Coherence: In discussing the coherence of an optical beam, we must distinguish between spatial and temporal coherence – laser beams have a high degree of both. **Spatial**

⁴Recall that the Q -factor is approximately the number of oscillations required for a freely oscillating system’s energy to fall by a factor of $1/e^{2\pi}$ of its original energy.

Source	$\Delta\nu$ (Hz)	t_{coh} (s)	ℓ_{coh} (m)
Na discharge lamp (D-lines at 589nm)	5×10^{11}	2×10^{-12}	6×10^{-4}
Multi-mode HeNe laser (632.8nm line)	1.5×10^9	6×10^{-10}	0.2
Single-mode HeNe laser (632.8nm line)	1×10^6	1×10^{-6}	300

Table 14.1: Typical values of the coherence length for a number of light sources.

coherence refers to whether there are irregularities in the optical phase in a cross-sectional slice of the beam. **Temporal coherence** refers to the time duration over which the phase of the beam is well defined. In general, the temporal coherence time, t_{coh} is given as the reciprocal of the spectral line width, $\Delta\nu$. Thus the coherence length ℓ_{coh} is given by,

$$\ell_{\text{coh}} = ct_{\text{coh}} = \frac{c}{\Delta\nu}. \quad (14.35)$$

Typical values of the coherence length for a number of light sources are given in table 14.1. These figures explain why it is much easier to conduct interference experiments with a laser than with a discharge lamp. If the path difference exceeds ℓ_{coh} you will not get interference fringes, because the light is incoherent.

Brightness: The brightness of lasers arises from two factors. First of all, the fact that the light is emitted in a well-defined beam means that the power per unit area is very high, even though the total amount of power can be rather low. Then we must consider that all the energy is concentrated within the narrow spectrum of the active atomic transition. This means that the spectral brightness (i.e. the intensity in the beam divided by the width of the emission line) is even higher in comparison with a white light source like a light bulb. For example, the spectral brightness of a 1mW laser beam could easily be millions of times greater than that of a 100W light bulb.

Ultra-short pulse generation: In some cases, lasers can be made to operate in pulses. The time duration of the pulses t_p is linked to the spectral band width of the laser light $\Delta\nu$ by the “uncertainty” product $\Delta t\Delta\nu \sim 1$. This follows from taking the Fourier transform of a pulse of duration t_p . As an example, the bandwidth of the 632.8nm line in the HeNe laser is 1.5GHz (see table 14.1), so that the shortest pulses that a HeNe laser can produce would be 0.67ns long. This is not particularly short by modern standards. Dye lasers typically have gain bandwidths greater than 1013Hz, and can be used to generate pulses shorter than 100fs. This is achieved by a technique called “mode-locking”. These short pulsed lasers are very useful for studying fast processes in physics, chemistry and biology

14.3.1 Operating Principles of a Laser

Light amplification is achieved by stimulated emission. Ordinary optical materials do not amplify light. Instead, they tend to absorb or scatter the light, so that the light intensity out of the medium is less than the intensity that went in. To get amplification you have to drive the material into a non-equilibrium state by pumping energy into it. Positive

optical feedback is achieved by inserting the amplifying medium inside a resonant cavity. Light in the cavity passes through the gain medium and is amplified. It then bounces off the end mirrors and passes through the gain medium again, getting amplified further. This process repeats itself until a stable equilibrium condition is achieved when the total round trip gain balances all the losses in the cavity.

The losses in the cavity fall into two categories: useful, and useless. The useful loss comes from the output coupling. One of the mirrors (called the “output coupler”) has reflectivity less than unity, and allows some of the light oscillating around the cavity to be transmitted as the output of the laser. The value of the transmission is chosen to maximise the output power. If the transmission is too low, very little of the light inside the cavity can escape, and thus we get very little output power. On the other hand, if the transmission is too high, there may not be enough gain to sustain oscillation, and there would be no output power. The optimum value is somewhere between these two extremes. Useless losses arise from absorption in the optical components (including the laser medium), scattering, and the imperfect reflectivity of the other mirror (the “high reflector”).

In general we expect the gain to increase as we pump more energy into the laser medium. At low pump powers, the gain will be small, and there will be insufficient gain to reach the oscillation condition. The laser will not start to oscillate until there is enough gain to overcome all the losses. This implies that the laser will have a threshold in terms of the pump power.

14.3.2 Gain Mechanism

Laser operation relies upon the phenomenon of stimulated emission. In a gas of atoms in thermal equilibrium, the population of lower levels will always be greater than the population of upper levels. Therefore, if a light beam is incident on the medium, there will always be more upward transitions due to absorption than downward transitions due to stimulated emission. Hence there will be net absorption, and the intensity of the beam will diminish on progressing through the medium.

To amplify the beam, we require that the rate of stimulated emission exceeds the rate of absorption. If the light beam is sufficiently intense that we can ignore spontaneous emission, and the levels are non-degenerate, this implies that the number of atoms in some upper level, N_2 , must exceed that of the lower level N_1 . This is a highly non-equilibrium situation, and is called **population inversion**. Once we have population inversion, we have a mechanism for generating gain in the laser medium. The art of making a laser operate is to work out how to get population inversion for the relevant transition.

To develop a theory of the laser threshold, we can consider separately the rate equations for the photon and atomic excitation. Starting with photons, let us consider excitations created by the transitions between just two levels of the atom – a lower level 1, and an excited state 2. If the dipole matrix elements, W , between the two levels are independent of position and frequency, the net downwards transition rate is given by

$$W(N_2(n+1) - N_1n) \quad (14.36)$$

where n denotes the total number of photons in the cavity, and $N_{1,2}$ is the number of atoms in states 1, 2. The first term represents the contribution from stimulated and spontaneous emission, while the latter is associated with absorption. Taking into account photon loss from the leaky cavity, the rate of change of photon number is therefore given by

$$\dot{n} = DWn + N_2W - \frac{n}{\tau_{\text{ph}}}, \quad (14.37)$$

where $D = N_2 - N_1$ represents the population imbalance and $1/\tau_{\text{ph}}$ is the photon loss rate. This equation shows that the gain in a laser medium is directly proportional to the degree of population inversion. Laser operation will occur when there is enough gain to overcome the losses in the cavity. This implies that a minimum amount of population inversion must be obtained before the laser will oscillate.

To achieve population inversion atoms must be “pumped” into the upper level 2. This can be achieved by a variety of techniques: Lasers are classified as being either three-level or four-level systems. In the following, we will consider the case of a three-level laser, although four-level lasers are more common. Examples of four-level lasers include Helium Neon or Nd:YAG. In a four-level laser, the levels comprise the ground state (0), the two lasing levels (1 and 2), and a fourth level (3) which is used as part of the pumping mechanism. In the three-level system, such as the first laser, ruby, level 1 is the ground state, and pumping is achieved by exciting atoms to level 3 with a bright flash lamp or by an electrical discharge, and then allowing them to decay rapidly to level 2. In this case, the corresponding rate equations for the populations of levels 1 and 2 can be written as

$$\dot{N}_2 \simeq -\omega_{21}N_2 + \omega_{12}N_1 - (N_2 - N_1)Wn \simeq -\dot{N}_1 \quad (14.38)$$

where ω_{12} , ω_{21} denote the “effective” transition rates between states 1 and 2 due to the pumping via the third state, and we have dropped the small contribution from spontaneous emission. From this equation, we can deduce that $N_1 + N_2 = N$, a constant, i.e. the decay from state 3 is so rapid that its population is always negligible. In this case, we obtain

$$\dot{D} = \frac{D_0 - D}{T} - 2DWn. \quad (14.39)$$

where $D_0(\omega_{12} - \omega_{21})/(\omega_{12} + \omega_{21})$ denotes the **unsaturated inversion** (i.e. the degree of population inversion that would exist if there were no photons in the cavity, $n = 0$) and $1/T = \omega_{12} + \omega_{21}$.

In steady-state, $\dot{n} = \dot{D} = 0$, and Eq. (14.39) translates to a population imbalance,

$$D \equiv N_2 - N_1 = \frac{D_0}{1 + 2TWn}. \quad (14.40)$$

From this result, we find the steady state photon number is given by

$$n = \frac{D_0W - 1/\tau_{\text{ph}}}{2TW/\tau_{\text{ph}}}. \quad (14.41)$$

When $D_0 > 1/W\tau_{\text{ph}}$, the **laser threshold condition**, there is a rapid increase in the number of photons in the cavity and the system starts lasing.

Although this analysis addresses the threshold conditions, it does not provide any insight into the coherence properties of the radiation field. In fact, one may show that the radiation field generated by the laser cavity forms a **coherent** state. The properties of this state will be discussed further when we discuss quantum states of light fields.

CHAPTER 15

Field Theory: From Phonons to Photons

In our survey of quantum mechanics so far, it has been possible to work with a discrete representation in which we index individual constituent particles. However, when the “elementary excitations” of the system involve the coherent *collective* motion of many individual discrete particle degrees of freedom – such as the wave-like atomic vibrations of an ordered elastic solid, or where discrete underlying classical particles can not even be identified – such as the electromagnetic field, such a representation is inconvenient or even inaccessible. In such cases, it is useful to turn to a continuum formulation of quantum mechanics. In the following, we will develop these fundamental ideas on the background of the simplest continuum theory: lattice vibrations of the atomic chain. As we will see, this study will provide a platform to investigate the quantum mechanics of the electromagnetic field – the subject of quantum electrodynamics – which also paves the way to the development of quantum field theory of relativistic particles.

15.1 Quantisation of the Classical Atomic Chain

As a simplified model of an ordered (one-dimensional) crystal, let us consider a chain of point particles each of mass m (atoms) which are elastically connected by springs with spring constant k_s (chemical bonds) (see Fig. 15.1). Although our target will be to construct a *quantum* theory of the elementary vibrational excitations, it is helpful to begin our analysis by reviewing the classical properties of the system.

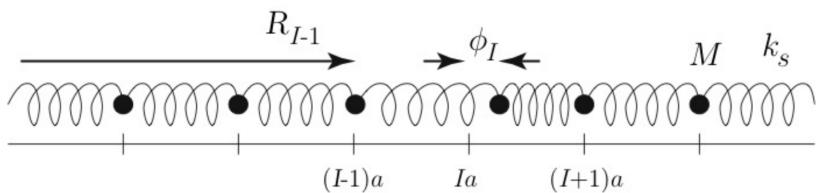


Fig. 15.1: Toy model of a one-dimensional solid: a chain of point-like particles each of mass m coupled elastically by springs with spring constant k_s .

15.1.1 Classical Chain

For reasons that will become clear, it is instructive to consider the **Lagrangian formulation** of the problem. For the N -atom chain, the classical Lagrangian is given by

$$L = T - V = \sum_{n=1}^N \left[\frac{1}{2} m \dot{x}_n^2 - \frac{1}{2} k_s (x_{n+1} - x_n - a)^2 \right]. \quad (15.1)$$

where the first term accounts for the kinetic energy of the particles whilst the second describes their coupling.¹ For convenience, we adopt periodic boundary conditions such that $x_{N+1} \equiv Na + x_1$ where a denotes the “natural” equilibrium lattice spacing. Anticipating that the effect of lattice vibrations on the solid is weak (i.e. long-range atomic order is maintained) we will assume that (a) the n^{th} atom has its equilibrium position at $\bar{x}_n \equiv na$, and (b) that the deviation from the equilibrium position remains small ($|x_n(t) - \bar{x}_n| \ll a$), i.e. the integrity of the solid is maintained. With $x_n(t) = \bar{x}_n + \phi_n(t)$ ($\psi_{N+1} = \phi_1$) the Lagrangian (15.1) then takes the form

$$L = \sum_{n=1}^N \left[\frac{1}{2} m \dot{\phi}_n^2 - \frac{1}{2} (\phi_{n+1} - \phi_n)^2 \right]. \quad (15.2)$$

Now, typically, we are not concerned with the behaviour of a given system on ‘atomic’ length scales. (For such purposes, our model is in any case much too primitive!) Rather, we are interested in **universal** features, i.e. experimentally observable behaviour, common to a wide range of physical systems, that manifests itself on macroscopic length scales where the detailed form of the model is inessential. For example, we might wish to study the specific heat of the solid in the limit of infinitely many atoms (or at least a macroscopically large number, $\mathcal{O}(10^{23})$). Under these conditions, microscopic models can usually be substantially simplified. In particular it is often permissible to subject a discrete lattice model to a **continuum approximation**, i.e. to neglect the discreteness of the microscopic entities of the system and to describe it in terms of effective continuum degrees of freedom. In the present case, taking a continuum limit amounts to describing the lattice displacements ϕ_n in terms of *smooth functions*, $\phi(x)$ of a continuous variable x (see Fig. 15.2). Clearly such a description makes sense only if relative fluctuations on atomic scales are weak (for otherwise the smoothness condition would be violated).

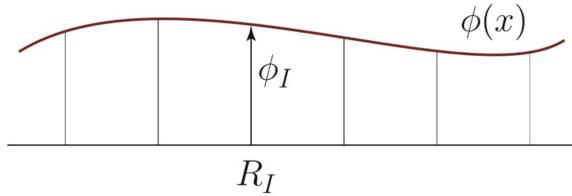


Fig. 15.2: Description of the lattice displacements ϕ_n in terms of *smooth functions*, $\phi(x)$ of a continuous variable x

Introducing continuum degrees of freedom, $\phi(x)$, and applying a first order Taylor expansion,² we can define

$$\phi_n(x) \rightarrow \phi(x)|_{x=na}, \quad \phi_{n+1} - \phi_n \rightarrow a\partial_x \phi(x)|_{x=na}, \quad \sum_{n=1}^N \rightarrow \frac{1}{a} \int_0^L dx, \quad (15.3)$$

where $L = Na$ (not to be confused with the Lagrangian itself!) denotes the total length of the chain. Expressed in terms of the new degrees of freedom, the continuum limit of

¹In real solids, the inter-atomic potential is, of course, more complex than our quadratic approximation. Yet, for “weak coupling”, the harmonic contribution dominates (cf. our discussion of molecular vibrations). For the sake of simplicity we, therefore, neglect the effects caused by higher order contributions.

²Indeed, for reasons that will become clear, higher order contributions to the Taylor expansion do not contribute to the low-energy properties of the system where the continuum approximation is valid.

the Lagrangian then reads $L[\phi] = \int_0^L dx \mathcal{L}(\phi, \dot{\phi})$, where

$$\mathcal{L}(\phi, \dot{\phi}) = \frac{1}{2}\rho\dot{\phi}^2 - \frac{1}{2}\kappa_s a(\partial_x\phi)^2, \quad (15.4)$$

denotes the Lagrangian density, $\rho = m/a$ denotes the mass per unit length and $\kappa_s = k_s/a$. The corresponding **classical action** is given by

$$S[\phi] = \int dx L[\phi]. \quad (15.5)$$

Thus, we have succeeded in trading the N -point particle description in for one involving *continuous* degrees of freedom, $\phi(x)$, a (**classical**) **field**. The dynamics of the latter are specified by “functionals” $L[\phi]$ and $S[\phi]$ which represent the continuum generalisations of the discrete classical Lagrangian and action, respectively.³ However, although we have achieved a continuum formulation, we have yet to extract concrete physical information from the action. To do so, we need to derive equations of motion. At first sight, it may not be entirely clear what is meant by ‘equations of motion’ in the context of an infinite dimensional model. The answer to this question lies in Hamilton’s extremal principle of classical mechanics:

- **Hamilton’s extremal principle:** Suppose that the dynamics of a classical point particle with coordinate $x(t)$ is described by the classical Lagrangian $L(x, \dot{x})$, and action $S[x] = \int dx L(x, \dot{x})$. Hamilton’s extremal principle states that the configurations $x(t)$ that are actually realised are those that extremise the action. This means that, for any smooth curve $y(t)$,

$$\lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} (S[x + \epsilon y] - S[x]) = 0, \quad (15.6)$$

i.e. to first order in ϵ , the action has to remain invariant. Applying this condition, one finds that it is fulfilled if and only if $x(t)$ obeys the **Euler-Lagrange equation** of motion,

$$\frac{d}{dt}(\partial_{\dot{x}}L) - \partial_x L = 0. \quad (15.7)$$

Now, in Eq. (15.5), we are dealing with a system of infinitely many degrees of freedom, $\phi(x, t)$. Yet Hamilton’s principle is general, and we may see what happens if (15.5) is subjected to an extremal principle analogous to Eq. (15.6). To do so, we must effect the substitution $\phi(x, t) \rightarrow \phi(x, t) + \epsilon\eta(x, t)$ into Eq. (15.5) and demand that the contribution first order in ϵ vanishes. When applied to the specific Lagrangian (15.4), a substitution of the of the ‘varied’ field leads to

$$S[\phi + \epsilon\eta] = S[\phi] + \epsilon \int dt \int_0^L dx (\rho\dot{\phi}\dot{\eta} - \kappa_s a^2 \partial_x\phi \partial_x\eta) + \mathcal{O}(\epsilon^2). \quad (15.8)$$

Integrating by parts (with respect to time for the first term under the integral, and space in the second) and demanding that the contribution linear in ϵ vanishes, one obtains

³In the mathematics and physics literature, mappings of functions into the real or complex numbers are generally called **functionals**. The argument of a functional is commonly indicated in rectangular brackets $[\bullet]$. For example, in this case, S maps the ‘functions’ $\partial_x\phi(x, t)$ and $\phi(x, t)$ to the real number $S[\phi]$.

$\int dt \int_0^L dx (\rho \ddot{\phi} - \kappa_s a^2 \partial_x^2 \phi) \eta = 0$. (Notice that the boundary terms associated with both t and x vanish identically.⁴ Now, since η was defined to be any arbitrary smooth function, the integral above can only vanish if the term in parentheses is globally vanishing. Thus the equation of motion takes the form of a **wave equation**,

$$\rho \ddot{\phi} = \kappa_s a^2 \partial_x^2 \phi. \quad (15.9)$$

The solutions of Eq. (15.9) have the general form $\phi_+(x + vt) + \phi_-(x - vt)$ where $v = a\sqrt{\kappa_s/\rho}$, and ϕ_{\pm} are arbitrary smooth functions of their argument. From this we can deduce that the basic low energy **elementary excitations** of our model are lattice vibrations propagating as **sound waves** to the left or right at a constant velocity v . The trivial behaviour of our model is of course a direct consequence of its simplistic definition – no dissipation, dispersion or other non-trivial ingredients. Adding these refinements leads to the general classical theory of lattice vibrations. With this background, let us now turn to the consider the quantisation of the quantum mechanical chain.

15.1.2 Quantum Chain

In addressing the quantum description, the first question to ask is a conceptual one: is there a general methodology to quantise models of the form described by the atomic chain (15.4)? Indeed, there is a standard procedure to quantise continuum theories which closely resembles the quantisation of point mechanics. The first step is to introduce canonical momenta conjugate to the continuum degrees of freedom (coordinates), ϕ , which will later be used to introduce canonical commutation relations. The natural generalisation of the definition $p_n \equiv \partial_{x_n} L$ of point mechanics to a continuum suggests setting

$$\pi = \partial_{\dot{\phi}} \mathcal{L}(\dot{\phi}, \phi). \quad (15.10)$$

In common with $\phi(x, t)$, the **canonical momentum**, $\pi(x, t)$, is a continuum degree of freedom. At each space point it may take an independent value. From the Lagrangian, we can define the Hamiltonian, $H[\phi, \pi] \equiv \int dx \mathcal{H}(\phi, \pi)$, where $\mathcal{H}(\phi, \pi) \equiv \pi \dot{\phi} - \mathcal{L}(\dot{\phi}, \phi)$ represents the **Hamiltonian density**. Applied to the atomic chain (15.4), the canonical momentum $\pi = \rho \dot{\phi}$ and $\mathcal{H}(\phi, \pi) = \frac{1}{2} \rho \pi^2 + \frac{1}{2} \kappa_s a^2 (\partial_x \phi)^2$.

In this form, the Hamiltonian can be quantised according to the following rules: (a) promote the fields $\phi(x)$ and $\pi(x)$ to operators: $\phi \mapsto \hat{\phi}$, $\pi \mapsto \hat{\pi}$, and (b) generalise the canonical commutation relations of one-particle quantum mechanics, $[\hat{p}_m, x_n] = -i\hbar \delta_{mn}$, according to the relation⁵

$$[\hat{\pi}(x), \hat{\phi}(x')] = -i\hbar \delta(x - x'). \quad (15.11)$$

⁴If we assume that the function ϕ already obeys the boundary conditions, we must have $\eta(0, t) = \eta(L, t) = \eta(x, 0) = \eta(x, T) = 0$.

⁵Note that the dimensionality of both the quantum and classical continuum fields is compatible with the dimensionality of the Dirac δ -function, $[\delta(x - x_0)] = [\text{Length}]^{-1}$.

Operator-valued functions like $\hat{\phi}$ and $\hat{\phi}$ are generally referred to as **quantum fields**. Employing these definitions, we obtain the quantum Hamiltonian density

$$\boxed{\hat{H}(\hat{\phi}, \hat{\pi}) = \frac{1}{2\rho}\hat{\pi}^2 + \frac{\kappa_s a^2}{2}(\partial_x \hat{\phi})^2.} \quad (15.12)$$

The Hamiltonian represents a quantum field theoretical *formulation* of the problem but not yet a *solution*. To address the quantum properties of the system, it is helpful now to switch to a Fourier representation. As with any function, operator-valued functions can be represented in a variety of different ways. In particular they can be subjected to Fourier expansion,

$$\begin{cases} \hat{\phi}_k \\ \hat{\pi}_k \end{cases} \equiv \frac{1}{L^{1/2}} \int_0^L dx e^{\{\mp ikx\}} \begin{cases} \hat{\phi}(x) \\ \hat{\pi}(x) \end{cases}, \quad \begin{cases} \hat{\phi}(x) \\ \hat{\pi}(x) \end{cases} = \frac{1}{L^{1/2}} \sum_k e^{\{\pm ikx\}} \begin{cases} \hat{\phi}_k \\ \hat{\pi}_k \end{cases} \quad (15.13)$$

where \sum_k represents the sum over all Fourier coefficients indexed by quantised wavevectors $k = 2\pi m/L$, m integer. Note that, since the classical field $\phi(x)$ is *real*, the quantum field $\hat{\phi}(x)$ is *Hermitian*, i.e. $\hat{\phi}_k = \hat{\phi}_{-k}^\dagger$ (and similarly for $\hat{\pi}_k$). In the Fourier representation, the transformed field operators obey the canonical commutation relations,

$$\boxed{[\hat{\pi}_k, \hat{\phi}_{k'}] = -i\hbar\delta_{kk'}} \quad (15.14)$$

When expressed in the Fourier representation, making use of the identity

$$\int_0^L dx (\partial \hat{\phi})^2 = \sum_{k,k'} (ik\hat{\phi}_k)(ik'\hat{\phi}_{k'}) \overbrace{\frac{1}{L} \int_0^L dx e^{i(k+k')x}}^{\delta_{k+k',0}} = \sum_k k^2 \hat{\phi}_k \hat{\phi}_{-k}, \quad (15.15)$$

together with the parallel relation for $\int_0^L dx \hat{\pi}^2$, the Hamiltonian assumes the “near diagonal” form,

$$\hat{H} = \sum_k \left[\frac{1}{2\rho} \hat{\pi}_k \hat{\pi}_{-k} + \frac{1}{2} \rho \omega_k^2 \hat{\phi}_k \hat{\phi}_{-k} \right], \quad (15.16)$$

where $\omega_k = v|k|$, and $v = a(\kappa_s/\rho)^{1/2}$ denotes the classical sound wave velocity. In this form, the Hamiltonian can be identified as nothing more than a superposition of independent quantum harmonic oscillators. The only difference between (15.16) and the canonical form of an oscillator Hamiltonian $\hat{H} = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2$ is the presence of sub-indices k and $-k$ (a consequence of $\hat{\phi}^\dagger = \hat{\phi}_{-k}$). As we will show shortly, this difference is inessential. This result is actually not difficult to understand (see Fig. 15.3): Classically, the system supports a discrete set of wave-like excitations, each indexed by a wave number $k = 2\pi m/L$. Within the quantum picture, each of these excitations is described by an oscillator Hamiltonian with a k -dependent frequency. However, it is important not to confuse the atomic constituents, also oscillators (albeit they coupled), with the independent *collective* oscillator modes described by \hat{H} .

The description above, albeit perfectly valid, still suffers from a deficiency: Our analysis amounts to explicitly describing the effective low energy excitations of the system (the waves) in terms of their microscopic constituents (the atoms). Indeed the different contributions to \hat{H} keep track of details of the microscopic oscillator dynamics of individual

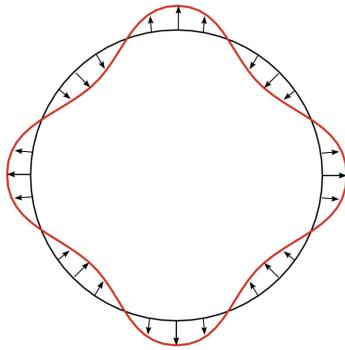


Fig. 15.3: Illustration of the quantum system's discrete wave-like excitations, each indexed by a wave number $k = 2\pi m/L$. These excitations are depicted as independent collective oscillator modes with k -dependent frequencies, distinct from the coupled atomic constituents also behaving as oscillators."

k -modes. However, it would be much more desirable to develop a picture where the relevant excitations of the system, the waves, appear as fundamental units, without explicit account of underlying microscopic details. (As with hydrodynamics, information is encoded in terms of collective density variables rather than through individual molecules.) To understand how this programme can be achieved let us recall the properties of the quantum harmonic oscillator.

In quantum mechanics, the **harmonic oscillator** has the status of a single-particle problem. However, the fact that the energy levels, $\epsilon_n = \hbar\omega(n + 1/2)$, are *equidistant* suggests an alternative interpretation: One can think of a given energy state ϵ_n as an accumulation of n elementary entities, or quasi-particles, each having energy $\hbar\omega$. What can be said about the features of these new objects? First, they are structureless, i.e. the only 'quantum number' identifying the quasi-particles is their energy $\hbar\omega$ (otherwise n -particle states formed of the quasi-particles would not be equidistant). This implies that the quasi-particles must be *bosons*. (The same state $\hbar\omega$ can be occupied by more than one particle—see Fig. 15.4.) This idea can be formulated in quantitative terms by employing the formalism of ladder operators in which the operators \hat{p} and \hat{x} are traded for the pair of Hermitian adjoint operators $a \equiv \sqrt{\frac{m\omega}{2\hbar}}(\hat{x} + \frac{i}{m\omega}\hat{p})$, $a^\dagger \equiv \sqrt{\frac{m\omega}{2\hbar}}(\hat{x} - \frac{i}{m\omega}\hat{p})$. Up to a factor of i , the transformation $(\hat{x}, \hat{p}) \rightarrow (a, a^\dagger)$ is canonical, i.e. the new operators obey the canonical commutation relation, $[a, a^\dagger] = 1$. More importantly, in the a -representation, the Hamiltonian takes the simple form, $\hat{H} = \hbar\omega(a^\dagger a + 1/2)$, as can be checked by direct substitution. The complete hierarchy of higher energy states can be generated by setting $|n\rangle \equiv \frac{1}{\sqrt{n!}}(a^\dagger)^n |0\rangle$.

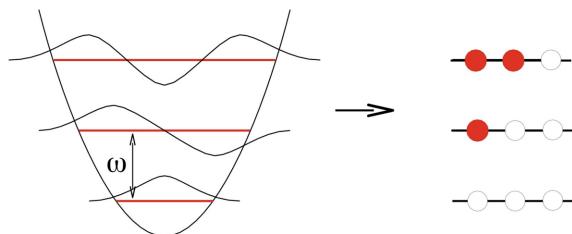


Fig. 15.4: Harmonic oscillator energy state ϵ_0 , as an accumulation of n elementary entities, or quasi-particles, each having energy $\hbar\omega$. These quasi-particles must be bosons since the same state $\hbar\omega$ can be occupied by more than one particle.

While the a -representation provides another way of constructing eigenstates of the quantum harmonic oscillator, its real advantage is that it naturally affords a many-particle interpretation. Temporarily forgetting about the original definition of the oscillator, we can declare $|0\rangle$ to be a ‘vacuum’ state, i.e. a state with no particles present. $a^\dagger|0\rangle$ then represents a state with a single featureless particle (the operator a^\dagger does not carry any quantum number labels) of energy $\hbar\omega$. Similarly, $(a^\dagger)^n|0\rangle$ is considered as a many-body state with n particles, i.e. within the new picture, a^\dagger is an operator that creates particles. The total energy of these states is given by $\hbar\omega \times (\text{occupation number})$. Indeed, it is straightforward to verify that $a^\dagger a|n\rangle = n|n\rangle$, i.e. the Hamiltonian basically counts the number of particles. While, at first sight, this may look unfamiliar, the new interpretation is internally consistent. Moreover, it fulfils our objective: it allows an interpretation of the excited states of the harmonic oscillator as a superposition of independent structureless entities.

With this background, we may return to the harmonic atomic chain (15.16) and, inspired by the ladder operator formalism, define⁶

$$a_k \equiv \sqrt{\frac{m\omega_k}{2\hbar}} \left(\hat{\phi}_k + \frac{i}{m\omega_k} \hat{\pi}_{-k} \right), \quad a_k^\dagger \equiv \sqrt{\frac{m\omega_k}{2\hbar}} \left(\hat{\phi}_{-k} - \frac{i}{m\omega_k} \hat{\pi}_k \right). \quad (15.17)$$

With this definition, one finds that the ladder operators obey the commutation relations

$$[a_k, a_{k'}^\dagger] = \delta_{kk'}, \quad [a_k, a_{k'}] = [a_k^\dagger, a_{k'}^\dagger] = 0, \quad (15.18)$$

and the Hamiltonian assumes the diagonal form

$$\hat{H} = \sum_k \hbar\omega_k \left(a_k^\dagger a_k + \frac{1}{2} \right).$$

(15.19)

Eqs. (15.18) and (15.19) represent the final result of our analysis: The low-lying elementary excitations of the discrete atomic chain are described by oscillator wave-like modes – known as **phonons** – each characterised by a wavevector k and a linear dispersion, $\omega_k = c|k|$. A generic state of the system is given by

$$|\{n_k\} = (n_1, n_2, \dots)\rangle = \frac{1}{\sqrt{\prod_i n_i!}} \left(a_{k_1}^\dagger \right)^{n_1} \left(a_{k_2}^\dagger \right)^{n_2} \cdots |0\rangle. \quad (15.20)$$

The representation derived above illustrates the capacity to think about quantum problems in different complementary “pictures”, a principle that finds innumerable applications. The existence of different interpretations of a given system is by no means heretic but, rather, is consistent with the spirit of quantum mechanics. Indeed, it is one of the prime principles of quantum theories that there is no such thing as ‘the real system’ which underpins the phenomenology. The only thing that matters is observable phenomena. For example, the ‘fictitious’ quasi-particle states of the harmonic chain, the phonons, *behave* as ‘real’ particles, i.e. they have dynamics, can interact, be detected experimentally, etc. From a quantum point of view there is actually no fundamental difference between these objects and ‘real’ particles.

This completes our discussion of the classical and quantum field theory of the harmonic atomic chain. In this example, we have seen how we can effect a quantum formulation of

⁶As for the consistency of these definitions, recall that $\hat{\phi}_k^\dagger = \hat{\phi}_{-k}$ and $\hat{\pi}_k^\dagger = \hat{\pi}_{-k}$. Under these conditions the second of the definitions below indeed follows from the first upon taking the Hermitian conjugate.

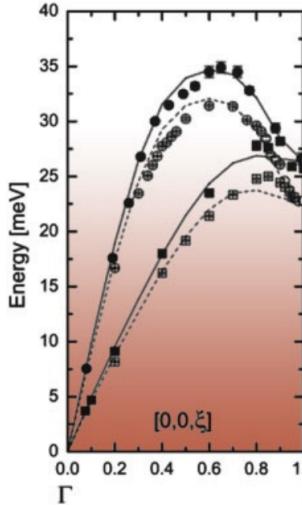


Fig. 15.5: The figure shows a typical measured phonon dispersion of an ordered crystalline solid obtained by neutron scattering. The x -axis indexes wavenumbers along a lattice direction (specified in units of π/a). Three generic aspects are visible: (1) near $k = 0$, the dispersion is, as expected, linear. (2) The several branches are associated with different “polarisations” of the lattice fluctuations. (3) For wavelengths comparable to the lattice spacing, $k \sim \pi/a$, non-universal features specific to the particular material become visible.

a continuum system. Using the insights obtained in this example, we now turn to consider the quantisation of the electromagnetic field.

15.2 Quantum electrodynamics

In common with the continuous formulation of the atomic chain, in vacua, the electromagnetic (EM) field satisfies a wave equation. The generality of the procedure outlined above suggests that the quantisation of the EM field might proceed in an entirely analogous manner. However, there are a number of practical differences that make quantisation a slightly more difficult enterprise: Firstly, the vector character of the vector potential \mathbf{A} , alongside relativistic covariance, gives the problem a non-trivial internal geometry. Moreover, the gauge freedom of the vector potential introduces redundant degrees of freedom whose removal on the quantum level is not straightforward. To circumvent a lengthy discussion of these issues, we will not address the problem of EM field quantisation in all its detail. On the other hand, the photon field plays a much too important role in all branches of physics for us to drop the problem altogether. We will therefore aim at an intermediate exposition, largely insensitive to the problems outlined above but sufficiently general to illustrate the main principles. As with the harmonic chain, to prepare the way, we begin by developing the classical field theory of the EM field.

15.2.1 Classical Theory of the Electromagnetic Field

In vacuum, the Lagrangian density of the EM field is given by $\mathcal{L} = -\frac{1}{4\mu_0} F_{\mu\nu} F^{\mu\nu}$ (summation convention implied) where $\mu_0 = 4\pi \times 10^{-7} \text{ H m}^{-1}$ denotes the **vacuum permeability**,

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu = \begin{pmatrix} 0 & -E_x/c & -E_y/c & -E_z/c \\ E_x/c & 0 & B_z & -B_y \\ E_y/c & -B_z & 0 & B_x \\ E_z/c & B_y & -B_x & 0 \end{pmatrix}_{\mu\nu} \quad (15.21)$$

denotes the EM field tensor, $\mathbf{E} = \dot{\mathbf{A}}$ is the electric field, and $\mathbf{B} = \nabla \times \mathbf{A}$ is the magnetic field. As a first step towards quantisation, we must specify a gauge. In the absence of charge, a particularly convenient choice is the Coulomb gauge, $\nabla \cdot \mathbf{A} = 0$, with the scalar component $\phi = 0$.⁷ Using these gauge conditions, one may verify that the classical Lagrangian assumes the form,

$$L[\mathbf{A}(\mathbf{x}, t)] = \frac{1}{2\mu_0} \int d^3x \left[\frac{1}{c^2} \dot{\mathbf{A}}^2 - (\nabla \times \mathbf{A})^2 \right]. \quad (15.22)$$

The corresponding classical Euler-Lagrange equations of motion, $\partial_\mu F_{\mu\nu} = 0$, translate to the wave equation

$$\frac{1}{c^2} \ddot{\mathbf{A}} = \nabla^2 \mathbf{A} \quad (15.23)$$

The structural similarity between the EM field and the continuous formulation of the harmonic chain is clear. By analogy with our discussion above, we should now switch to the Fourier representation and quantise the classical field. However, in contrast to our analysis of the chain, we are now dealing (i) with the full three-dimensional Laplacian acting upon (ii) the vector field \mathbf{A} that is (iii) subject to the constraint $\nabla \cdot \mathbf{A} = 0$. It is these aspects which lead to the complications outlined above.

We can circumvent these difficulties by considering cases where the geometry of the system reduces the complexity of the eigenvalue problem while still retaining the key conceptual aspects of the problem. This restriction is less artificial than it might appear. For example, just as the field ϕ in the classical atomic chain can be expanded in Fourier harmonics, in long waveguides, the EM vector potential can be expanded in solutions of the eigenvalue equation⁸

$$-\nabla^2 \mathbf{u}_k(\mathbf{x}) = \lambda_k \mathbf{u}_k(\mathbf{x}), \quad (15.24)$$

where k denotes a discrete *one-dimensional* index, and the vector-valued functions \mathbf{u}_k are real and orthonormalised, $\int d^3x \mathbf{u}_k \cdot \mathbf{u}_k = k' = \delta_{kk'}$. The dependence of the eigenvalues λ_k on k depends on details of the geometry and need not be specified for the moment.

- An **electrodynamic waveguide** is a quasi one-dimensional cavity with metallic boundaries (see Fig. 15.6). The practical advantage of waveguides is that they

⁷Keep in mind that, once a gauge has been set, we cannot expect further results to display “gauge invariance.”

⁸More precisely, one should say that Eq. (15.24) defines the set of eigenfunctions relevant for the *low-energy* dynamics of the waveguide. More complex eigenfunctions of the Laplace operator exist but they involve much higher energy

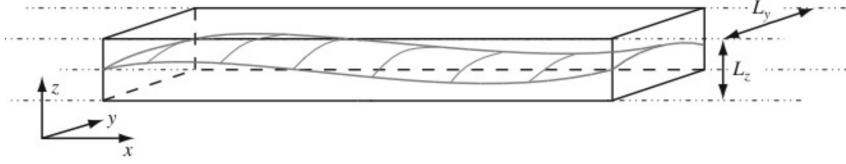


Fig. 15.6: EM waveguide with rectangular cross-section. The structure of the eigenmodes of the EM field is determined by boundary conditions at the walls of the cavity.

are good at confining EM waves. At large frequencies, where the wavelengths are of order meters or less, radiation loss in conventional conductors is high. In these frequency domains, hollow conductors provide the only practical way of transmitting radiation. EM field propagation inside a waveguide is constrained by boundary conditions. Assuming the walls of the system to be perfectly conducting,

$$\mathbf{E}_{\parallel}(\mathbf{x}_b) = 0, \quad \mathbf{B}_{\perp}(\mathbf{x}_b) = 0 \quad (15.25)$$

where \mathbf{x}_b parameterise points on the boundary of the system, and $\mathbf{E}_{\parallel}(\mathbf{B}_{\perp})$ is the parallel (perpendicular) component of the electric (magnetic) field. Applied to the problem at hand, let us consider a long cavity with uniform rectangular cross-section $L_y \times L_z$. To conveniently represent the Lagrangian of the system, we wish to express the vector potential in terms of eigenfunctions \mathbf{u}_k that are consistent with the boundary conditions (15.25). A complete set of functions fulfilling this condition is given by

$$\mathbf{u}_k = \mathcal{N}_k \begin{pmatrix} c_1 \cos(k_x x) \sin(k_y y) \sin(k_z z) \\ c_2 \sin(k_x x) \cos(k_y y) \sin(k_z z) \\ c_3 \sin(k_x x) \sin(k_y y) \cos(k_z z) \end{pmatrix}. \quad (15.26)$$

Here $k_i = n_i \pi / L_i$, with $i = x, y, z$ and n_i is integer, \mathcal{N}_k is a factor normalising \mathbf{u}_k to unit modulus, and the coefficients c_i are subject to the condition $c_1 k_x + c_2 k_y + c_3 k_z = 0$ (reflecting the gauge choice $\nabla \cdot \mathbf{A} = 0$). Indeed, it is straightforward to verify that a general superposition of the type $\mathbf{A}(\mathbf{x}, t) \equiv \sum_k \alpha_k(t) \mathbf{u}_k(\mathbf{x})$, $\alpha_k(t) \in \mathbb{R}$, is divergenceless, and generates an EM field compatible with (15.25). Substitution of \mathbf{u}_k into (15.24) identifies the eigenvalues as $\lambda_k = k_x^2 + k_y^2 + k_z^2$.

- In the physics and electronic engineering literature, eigenfunctions of the Laplace operator in a quasi-one-dimensional geometry are commonly described as **modes**. As we will see shortly, the energy of a mode (i.e. the Hamiltonian evaluated on a specific mode configuration) grows with λ_k . In cases where one is interested in the low-energy dynamics of the EM field, only configurations with small λ_k are relevant. For example, let us consider a massively anisotropic waveguide with $L_z < L_y \ll L_x$. In this case the modes with smallest λ_k are those with $k_z = 0$, $k_y = \pi / L_y$, and $k_x \equiv k \ll L_{z,y}^{-1}$. With this choice,

$$\mathbf{u}_k = \frac{2}{\sqrt{V}} \sin(\pi y / L_y) \sin(k_x) \hat{\mathbf{e}}_z, \quad \lambda_k = k^2 + \left(\frac{\pi}{L_y} \right)^2, \quad (15.27)$$

and a scalar index k suffices to label both eigenvalues and eigenfunctions \mathbf{u}_k . A caricature of the spatial structure of the functions \mathbf{u}_k is shown in Fig. 15.6.

Returning to the problem posed by (15.22) and (15.24), one can expand the vector potential in terms of eigenfunctions \mathbf{u}_k as $\mathbf{A}(\mathbf{x}, t) = \sum_k \alpha_k(t) \mathbf{u}_k(\mathbf{x})$, where the sum runs

over all allowed values of the index parameter k . (In a waveguide of length L , $k = \pi n/L$ with n integer.) Substituting this expansion into (15.22), and using the normalisation properties of \mathbf{u}_k , we obtain the Lagrangian,

$$L[\dot{\alpha}, \alpha] = \frac{1}{2\mu_0} \sum_k \left[\frac{1}{c^2} \dot{\alpha}_k^2 - \lambda_k \alpha_k^2 \right], \quad (15.28)$$

i.e. a decoupled representation where the system is described in terms of independent dynamical systems with coordinates α_k . From this point on, the quantisation procedure mirrors that of the atomic chain

15.2.2 Quantum Field Theory of the Electromagnetic Field

To achieve the electromagnetic field quantisation, we first define the canonical momenta through the relation,

$$\pi_k = \partial_{\dot{\alpha}_k} \mathcal{L} = \epsilon_0 \dot{\alpha}_k, \quad (15.29)$$

where $\epsilon_0 = 1/\mu_0 c^2$ denotes the **vacuum permittivity**, which leads to the classical Hamiltonian $H = \sum_k \left(\frac{1}{2\epsilon_0} \pi_k^2 + \frac{1}{2} \epsilon_0 c^2 \lambda_k \alpha_k^2 \right)$. Next we quantise the theory by promoting fields to operators $\alpha_k \rightarrow \hat{\alpha}_k$ and $\pi_k \rightarrow \hat{\pi}_k$, and declare the canonical commutation relations $[\hat{\pi}_k, \hat{\alpha}_{k'}] = -i\hbar\delta_{kk'}$. The quantum Hamiltonian operator, again of harmonic oscillator type, then reads

$$\hat{H} = \sum_k \left[\frac{\hat{\pi}_k^2}{2\epsilon_0} + \frac{1}{2} \epsilon_0 \omega_k^2 \hat{\alpha}_k^2 \right]. \quad (15.30)$$

where $\omega_k^2 = c^2 \lambda_k$.

Then, guided by the analysis of the atomic chain, we now introduce the ladder operators,

$$a_k = \sqrt{\frac{\epsilon_0 \omega_k}{2\hbar}} \left(\hat{\alpha}_k + \frac{i}{\epsilon_0 \omega_k} \hat{\pi}_k \right), \quad a_k^\dagger = \sqrt{\frac{\epsilon_0 \omega_k}{2\hbar}} \left(\hat{\alpha}_k - \frac{i}{\epsilon_0 \omega_k} \hat{\pi}_k \right), \quad (15.31)$$

whereupon the Hamiltonian assumes the now familiar form

$$\hat{H} = \sum_k \hbar \omega_k \left(a_k^\dagger a_k + \frac{1}{2} \right). \quad (15.32)$$

For the specific problem of the first excited mode in a waveguide of width L_y , $\hbar \omega_k = c[k^2 + (\pi/L_y)^2]^{1/2}$. Eq. (15.32) represents our final result for the quantum Hamiltonian of the EM waveguide. Physically, the quantum excitations described by this Hamiltonian are the **photons** of the EM field. The unfamiliar appearance of the dispersion ω_k is a peculiarity of the waveguide. In the limit of large longitudinal wave numbers $k \gg L_y^{-1}$, the dispersion approaches $\omega_k \simeq c|k|$, i.e. the familiar linear (relativistic) dispersion of the photon field.

With the analysis of the waveguide complete, let us go back and consider the quantisation of the full three-dimensional system. For the waveguide, we have found that the vector potential can be expanded in modes of the cavity as $\hat{\mathbf{A}}(\mathbf{x}) = \sum_k \hat{\alpha}_k \mathbf{u}_k$ where, rearranging the expressions for the ladder operators, $\hat{\alpha}_k = \sqrt{\frac{\hbar}{2\epsilon_0 \omega_k}} (a_k + a_k^\dagger)$. More generally,

in a fully three-dimensional cavity, one may show that⁹

$$\hat{\mathbf{A}}(\mathbf{x}) = \sum_{\mathbf{k}\lambda=1,2} \sqrt{\frac{\hbar}{2\epsilon_0\omega_{\mathbf{k}}V}} [\hat{\mathbf{e}}_{\mathbf{k}\lambda} a_{\mathbf{k}\lambda} e^{i\mathbf{k}\cdot\mathbf{x}} + \hat{\mathbf{e}}_{\mathbf{k}\lambda}^* a_{\mathbf{k}\lambda}^\dagger e^{-i\mathbf{k}\cdot\mathbf{x}}], \quad (15.33)$$

where V denotes the volume of the system, $\omega_{\mathbf{k}} = c|\mathbf{k}|$, and the two sets of polarisation vectors, $\hat{\mathbf{e}}_{\mathbf{k}\lambda}$, are in general complex and normalised to unity, $\hat{\mathbf{e}}_{\mathbf{k}\lambda}^* \cdot \hat{\mathbf{e}}_{\mathbf{k}\lambda} = 1$. To ensure that the vector potential satisfies the Coulomb gauge condition, we require that $\hat{\mathbf{e}}_{\mathbf{k}\lambda} \cdot \mathbf{k} = \hat{\mathbf{e}}_{\mathbf{k}\lambda}^* \cdot \mathbf{k} = 0$, i.e. the two polarisation vectors are orthogonal to the wave vector. Two real vectors, $\hat{\mathbf{e}}_{\mathbf{k}\lambda}$ correspond to two linear polarisations while, for circular polarisation, the vectors are complex. It is also convenient to assume that the two polarisation vectors are mutually orthogonal, $\hat{\mathbf{e}}_{\mathbf{k}\lambda} \cdot \hat{\mathbf{e}}_{\mathbf{k}\nu} = \delta_{\lambda\nu}$. The corresponding operators obey the commutation relations,

$$[a_{\mathbf{k}\lambda}, a_{\mathbf{k}'\lambda'}^\dagger] = \delta_{\mathbf{k},\mathbf{k}'}, \quad [a_{\mathbf{k}\lambda}, a_{\mathbf{k}'\lambda'}] = 0 = [a_{\mathbf{k}\lambda}^\dagger, a_{\mathbf{k}'\lambda'}^\dagger]. \quad (15.34)$$

With these definitions, the Hamiltonian then takes the familiar form

$$\hat{H} = \sum_{\mathbf{k}\lambda} \hbar\omega_{\mathbf{k}} (a_{\mathbf{k}\lambda}^\dagger a_{\mathbf{k}\lambda} + \frac{1}{2}), \quad (15.35)$$

while, defining the vacuum, $|\Omega\rangle$, the eigenstates involve photon number states,

$$|\{n_{\mathbf{k}\lambda}\}\rangle \equiv |\cdots, n_{\mathbf{k}\lambda}, \cdots\rangle = \prod_{\mathbf{k}\lambda} \frac{(a_{\mathbf{k}\lambda}^\dagger)^{n_{\mathbf{k}\lambda}}}{\sqrt{n_{\mathbf{k}\lambda}!}} |\Omega\rangle. \quad (15.36)$$

Finally, in practical applications (including our study of radiative transitions in atoms), it is convenient to transfer the time-dependence to the operators by turning to the Heisenberg representation. In this representation, the field operators obey the Heisenberg equations of motion,

$$\dot{a}_{\mathbf{k}\lambda} = \frac{i}{\hbar} [\hat{H}, a_{\mathbf{k}\lambda}] = -i\omega_{\mathbf{k}} a_{\mathbf{k}\lambda}. \quad (15.37)$$

Integrating, we have $a_{\mathbf{k}\lambda}(t) = a_{\mathbf{k}\lambda}(0)e^{-i\omega_{\mathbf{k}}t}$, which translates to the relation

$$\hat{\mathbf{A}}(\mathbf{x}, t) = \sum_{\mathbf{k}\lambda=1,2} \sqrt{\frac{\hbar}{2\epsilon_0\omega_{\mathbf{k}}V}} [\hat{\mathbf{e}}_{\mathbf{k}\lambda} a_{\mathbf{k}\lambda} e^{i(\mathbf{k}\cdot\mathbf{x}-\omega_{\mathbf{k}}t)} + \hat{\mathbf{e}}_{\mathbf{k}\lambda}^* a_{\mathbf{k}\lambda}^\dagger e^{-i(\mathbf{k}\cdot\mathbf{x}-\omega_{\mathbf{k}}t)}]. \quad (15.38)$$

From the vector potential, we can obtain the electric field and magnetic fields, via $\mathbf{E} = -\dot{\mathbf{A}}$ and $\mathbf{B} = \nabla \times \mathbf{A}$, noting that we have chosen a gauge in which the scalar potential vanished, $\phi = 0$. In terms of the field operators, these are

$$\hat{\mathbf{E}}(\mathbf{x}, t) = i \sum_{\mathbf{k}\lambda=1,2} \sqrt{\frac{\hbar\omega_{\mathbf{k}}}{2\epsilon_0 V}} [\hat{\mathbf{e}}_{\mathbf{k}\lambda} a_{\mathbf{k}\lambda} e^{i(\mathbf{k}\cdot\mathbf{x}-\omega_{\mathbf{k}}t)} - \hat{\mathbf{e}}_{\mathbf{k}\lambda}^* a_{\mathbf{k}\lambda}^\dagger e^{-i(\mathbf{k}\cdot\mathbf{x}-\omega_{\mathbf{k}}t)}], \quad (15.39)$$

$$\hat{\mathbf{B}}(\mathbf{x}, t) = i \sum_{\mathbf{k}\lambda=1,2} \sqrt{\frac{\hbar\omega_{\mathbf{k}}\mu_0}{2V}} [\hat{\mathbf{k}} \times \hat{\mathbf{e}}_{\mathbf{k}\lambda} a_{\mathbf{k}\lambda} e^{i(\mathbf{k}\cdot\mathbf{x}-\omega_{\mathbf{k}}t)} + \hat{\mathbf{k}} \times \hat{\mathbf{e}}_{\mathbf{k}\lambda}^* a_{\mathbf{k}\lambda}^\dagger e^{-i(\mathbf{k}\cdot\mathbf{x}-\omega_{\mathbf{k}}t)}]. \quad (15.40)$$

⁹In the infinite system, the mode sum becomes replaced by an integral, $\sum_k \rightarrow \frac{V}{(2\pi)^3} \int d^3k$.

15.2.3 Fock States

The energy eigenstates of the EM field (15.36) are referred to as “number” or “Fock” states. For these states, the expectation value of the electric and magnetic fields both vanish

$$\langle \hat{\mathbf{E}} \rangle = \langle \hat{\mathbf{B}} \rangle = 0. \quad (15.41)$$

These results follow from the facts that: (i) $\hat{\mathbf{A}}$ is linear in the operators \hat{a} and \hat{a}^\dagger , and therefore so too are $\hat{\mathbf{E}}$ and $\hat{\mathbf{B}}$; and (ii) for any Fock state $|n\rangle$,

$$\langle n | \hat{a} | n \rangle = \langle n | \hat{a}^\dagger | n \rangle = 0. \quad (15.42)$$

The expectation values of the *squares* of the electric and magnetic fields do not vanish for Fock states. Direct calculations show that for the Fock state (15.36) the mean square electric and magnetic fields at a point \mathbf{r} are

$$\left\langle \left| \hat{\mathbf{E}}(\mathbf{r}) \right|^2 \right\rangle = \sum_{\mathbf{k},\lambda} \frac{\hbar\omega_{\mathbf{k}}}{2\epsilon_0 V} \left(n_{\mathbf{k},\lambda} + \frac{1}{2} \right), \quad (15.43)$$

$$\left\langle \left| \hat{\mathbf{B}}(\mathbf{r}) \right|^2 \right\rangle = \sum_{\mathbf{k},\lambda} \frac{\hbar\omega_{\mathbf{k}}\mu_0}{2V} \left(n_{\mathbf{k},\lambda} + \frac{1}{2} \right). \quad (15.44)$$

These results are consistent with the expected total energy in the EM field of these Fock states

$$\langle U \rangle = \int \left[\frac{\epsilon_0}{2} \left\langle \left| \hat{\mathbf{E}}(\mathbf{r}) \right|^2 \right\rangle + \frac{1}{2\mu_0} \left\langle \left| \hat{\mathbf{B}}(\mathbf{r}) \right|^2 \right\rangle \right] d^3r = \sum_{\mathbf{k},\lambda} \hbar\omega_{\mathbf{k}} \left(n_{\mathbf{k},\lambda} + \frac{1}{2} \right). \quad (15.45)$$

Note that the vacuum state, with all $\{n_{\mathbf{k},\lambda} = 0\}$, has a non-zero energy density. This is a manifestation of the zero-point motion of the EM fields. The vacuum energy density is a (formally divergent) background energy density predicted by the theory. These zero-point fluctuations have a number of observable physical consequences. Indeed, the spontaneous emission described in section 14.1.2 can be viewed as arising from zero-point motion of the fields. Another dramatic consequence is given by the **Casimir effect**. In its simplest setting, this refers to the force between two parallel perfectly conducting planes that arises from the change in the zero-point energy due to the boundary conditions on the plates.

15.2.4 Coherent States

Coherent states – also called Glauber states – of light are relevant to much of modern optical physics. They are the “most classical” form of light field. They are generated from classical sources, and they also describe the state of light in a laser cavity. To discuss the properties of coherent states, we focus on a single mode with frequency ω and ladder operators \hat{a}^\dagger .

A coherent state $|\alpha\rangle$ is defined by the condition

$$\hat{a} |\alpha\rangle = \alpha |\alpha\rangle. \quad (15.46)$$

That is, it is an eigenstate of the destruction operator, with eigenvalue α . Since \hat{a} is not Hermitian, the eigenvalue α is, in general, complex.

The coherent state may be written

$$|\alpha\rangle = \frac{1}{\mathcal{N}_\alpha} \exp(\alpha\hat{a}^\dagger) |0\rangle = \frac{1}{\mathcal{N}_\alpha} \sum_{n=0}^{\infty} \frac{(\alpha\hat{a}^\dagger)^n}{n!} |0\rangle \quad (15.47)$$

where $|0\rangle$ is the oscillator vacuum and \mathcal{N}_α is a normalisation factor.

To find the normalisation, consider

$$\langle\alpha|\alpha\rangle = \frac{1}{\mathcal{N}_\alpha^2} \sum_{n,n'} \langle 0| \frac{(\alpha^*\hat{a})^{n'}}{n'!} \frac{(\alpha\hat{a}^\dagger)^n}{n!} |0\rangle \quad (15.48)$$

$$= \frac{1}{\mathcal{N}_\alpha^2} \sum_n \frac{(\alpha^*\alpha)^n}{n!} \quad (15.49)$$

$$= \frac{1}{\mathcal{N}_\alpha^2} e^{|\alpha|^2}. \quad (15.50)$$

Thus, a normalised state is obtained for $\mathcal{N}_\alpha = \exp(|\alpha|^2/2)$, and

$$|\alpha\rangle = \exp(-|\alpha|^2/2) \exp(\alpha\hat{a}^\dagger) |0\rangle. \quad (15.51)$$

Since the coherent state is not an energy eigenstate, expectation values of observables are, in general, time-dependent. In the Heisenberg representation, where the time-dependence is in the operators, the electric field (for this mode) is

$$\hat{\mathcal{E}}(\mathbf{r}, t) = i\sqrt{\frac{\hbar\omega}{2\epsilon_0 V}} [-\hat{a}e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)} + \hat{a}^\dagger e^{-i(\mathbf{k}\cdot\mathbf{r}-\omega t)}]. \quad (15.52)$$

Writing α in polar form

$$\alpha = |\alpha|e^{i\theta}, \quad (15.53)$$

one finds

$$\langle\alpha|\hat{\mathcal{E}}(\mathbf{r}, t)|\alpha\rangle = i\sqrt{\frac{\hbar\omega}{2\epsilon_0 V}} [-|\alpha|e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t+\theta)} + |\alpha|e^{-i(\mathbf{k}\cdot\mathbf{r}-\omega t+\theta)}] \quad (15.54)$$

$$= 2|\alpha|\sqrt{\frac{\hbar\omega}{2\epsilon_0 V}} \sin[\omega t - \mathbf{k}\cdot\mathbf{r} - \theta]. \quad (15.55)$$

This looks like a classical field, with amplitude $\propto |\alpha|$ and phase θ .

Furthermore, a calculation of the fluctuations of $\hat{\mathcal{E}}(\mathbf{r}, t)$ leads to

$$\langle(\Delta\mathcal{E})^2\rangle = \frac{\hbar\omega}{2\epsilon_0 V} \quad (15.56)$$

which is identical to the fluctuations for the vacuum state.

The coherent-state is the “most classical” state of the oscillator in the sense that it exhibits the same time dependence as a classical field, while having only the noise of the vacuum.

15.2.5 Non-Classical Light

Since the EM field is a quantum system, all states of the EM field are in some sense non-classical. Of course, in practice it may be difficult to prepare or detect quantum effects in a light field. For example, classical sources typically generate coherent states (described above). If the amplitude of the field is large (or our sensitivity to measuring the field is low), then the quantum fluctuations of the field in the coherent state will be undetectable.

In modern optical physics – particularly in cavities with small mode volume V where even a single photon can generate a sizeable electric field – techniques are very advanced, and it is routine to generate and to detect non-classical effects of light. “Non-classical” in this context is used to denote a state of the light field that is *not* a coherent state. As a definition this is rather vague, qualifying all but this special coherent state as non-classical. There are, however, a few central examples of non-classical states. These are referred to as **squeezed states** of light. “Squeezing” refers to the reduction in the uncertainty in one particular observable. (Owing to the uncertainty relations, this will lead to an increase in the uncertainty of the variable conjugate to this.)

15.2.5.1 Squeezed Light

We have already come across what is perhaps the most non-classical of all such states: the Fock state, for which the EM field is time-independent, with vanishing expectation value. This can be described as an amplitude, or “number”, squeezed state, having the minimum possible uncertainty in photon number. (By contrast there is a conjugate variable, which is loosely defined as the phase of the wave, which is completely uncertain.)

Another form of squeezed light is obtained by “quadrature” squeezing. The quadrature operators are defined by

$$\hat{X}_1 = \frac{1}{2}(\hat{a} + \hat{a}^\dagger) \quad (15.57)$$

$$\hat{X}_2 = \frac{1}{2i}(\hat{a} - \hat{a}^\dagger) \quad (15.58)$$

For a simple harmonic oscillator, they can be viewed as (proportional to) the position and momentum operators.

For a single mode of the EM field, the electric field may be written

$$\hat{\mathcal{E}} = (\hat{a}e^{-i\omega t} + \hat{a}^\dagger e^{+i\omega t}). \quad (15.59)$$

(For simplicity, we neglect spatial dependence and polarisation dependence, and have rescaled the electric field by the vacuum amplitude $\sqrt{\hbar\omega/(2\epsilon_0 V)}$.) In terms of the quadrature operators, this becomes

$$\hat{\mathcal{E}} = 2[\hat{X}_1 \cos(\omega t) + \hat{X}_2 \sin(\omega t)]. \quad (15.60)$$

Thus, X_1 and X_2 are the field amplitudes for the components that oscillate $\pi/2$ out of phase with each other.¹⁰

¹⁰These two components are in quadrature, hence the name.

The quadrature operators satisfy the commutation relation

$$[\hat{X}_1, \hat{X}_2] = \frac{i}{2} \quad (15.61)$$

from which it follows that

$$\Delta X_1 \Delta X_2 \geq \frac{1}{4}. \quad (15.62)$$

For the number states $|n\rangle$

$$\Delta X_1 = \Delta X_2 = \frac{\sqrt{(2n+1)}}{2}. \quad (15.63)$$

Of these, only the vacuum $|0\rangle$ is a minimum uncertainty state. The coherent states $|\alpha\rangle$ have the (remarkable!) property that

$$\Delta X_1 = \Delta X_2 = \frac{1}{2} \quad (15.64)$$

independent of α . Thus, they are minimum uncertainty states. Since $\Delta X_1 = \Delta X_2$, at all times during their time evolution, the uncertainty in the electric field

$$(\Delta \mathcal{E})^2 = (2\Delta X_1)^2 \cos^2(\omega t) + (2\Delta X_2)^2 \sin^2(\omega t) \quad (15.65)$$

is constant and equal to the uncertainty of the vacuum.

The “quadrature squeezed” states are minimum uncertainty states, $\Delta X_1 \Delta X_2 = 1/4$, for which $\Delta X_1 \neq \Delta X_2$. Thus, either ΔX_1 or ΔX_2 is less than $1/2$. Now, the uncertainty in the electric field¹¹

$$(\Delta \mathcal{E})^2 = (2\Delta X_1)^2 \cos^2(\omega t) + (2\Delta X_2)^2 \sin^2(\omega t) \quad (15.66)$$

varies in time. At certain times in the cycle, the uncertainty in the electric field is *smaller than the uncertainty in the vacuum*. (If $\Delta X < 1/2$, then the uncertainty is below its vacuum value for $t = \pi/\omega \times \text{integer}$).

Quadrature squeezed light can be generated using nonlinear optical elements. That there is less “noise” in one of the quadratures than in a coherent state, or in the vacuum state, has possible technological applications, particularly in precision measurements and the detection of weak signals. One example is in the detection of gravity waves with large-scale interferometers as in LIGO and VIRGO.

¹¹We assume that $\langle \hat{X}_1 \hat{X}_2 + \hat{X}_2 \hat{X}_1 \rangle = 2 \langle \hat{X}_1 \rangle \langle \hat{X}_2 \rangle$, which can always be achieved for quadrature squeezed light by introducing an offset in time.

APPENDIX A

Appendix

A.1 Additional Proofs

A.1.1 Probability Current

Consider the expression

$$i\hbar \frac{\partial}{\partial t} \int_{\mathcal{V}} P(\mathbf{r}, t) dV = i\hbar \frac{\partial}{\partial t} \int_{\mathcal{V}} \Psi^*(\mathbf{r}, t) \Psi(\mathbf{r}, t) dV; \quad (\text{A.1})$$

apart from the factor $i\hbar$, this is the rate of change of the probability of finding the particle in a closed region \mathcal{V} :

$$i\hbar \frac{\partial}{\partial t} \int_{\mathcal{V}} P(\mathbf{r}, t) dV = i\hbar \int_{\mathcal{V}} \left[\Psi^*(\mathbf{r}, t) \frac{\partial \Psi(\mathbf{r}, t)}{\partial t} + \frac{\partial \Psi^*(\mathbf{r}, t)}{\partial t} \Psi(\mathbf{r}, t) \right] dV \quad (\text{A.2})$$

$$= \int_{\mathcal{V}} \left[\Psi^*(\mathbf{r}, t) \hat{H} \Psi(\mathbf{r}, t) - (\hat{H} \Psi^*(\mathbf{r}, t)) \Psi(\mathbf{r}, t) \right] dV \quad (\text{A.3})$$

$$= -\frac{\hbar^2}{2m} \int_{\mathcal{V}} \left[\Psi^*(\mathbf{r}, t) \nabla^2 \Psi(\mathbf{r}, t) - (\nabla^2 \Psi^*(\mathbf{r}, t)) \Psi(\mathbf{r}, t) \right] dV \quad (\text{A.4})$$

$$= -\frac{\hbar^2}{2m} \int_{\mathcal{V}} \nabla \cdot [\Psi^*(\mathbf{r}, t) \nabla \Psi(\mathbf{r}, t) - \Psi(\mathbf{r}, t) \nabla \Psi^*(\mathbf{r}, t)] dV \quad (\text{A.5})$$

$$= -\frac{\hbar^2}{2m} \oint_{\mathcal{S}} [\Psi^*(\mathbf{r}, t) \nabla \Psi(\mathbf{r}, t) - \Psi(\mathbf{r}, t) \nabla \Psi^*(\mathbf{r}, t)] \cdot d\mathbf{S} \quad (\text{A.6})$$

where the third line follows from Eq. (1.20), and the fourth line follows from applying the Hamiltonian, Eq. (2.4), and noting that the terms involving the potential cancel. The fifth line follows by applying the vector differential identity $\nabla \cdot (f\mathbf{F}) = \mathbf{F} \cdot \nabla f + f \nabla \cdot \mathbf{F}$, which is valid for any scalar function f and vector field \mathbf{F} . The sixth line comes from applying the divergence theorem. Thus,

$$\frac{\partial}{\partial t} \int_{\mathcal{V}} P(\mathbf{r}, t) dV = - \oint_{\mathcal{S}} \mathbf{J}(\mathbf{r}, t) \cdot d\mathbf{S} \quad (\text{A.7})$$

where

$$\mathbf{J}(\mathbf{r}, t) = \frac{\hbar}{2mi} [\Psi^*(\mathbf{r}, t) \nabla \Psi(\mathbf{r}, t) - \Psi(\mathbf{r}, t) \nabla \Psi^*(\mathbf{r}, t)] \quad (\text{A.8})$$

which is a vector field. Eq. (A.7) shows that the time rate of increase of the probability within some closed region is equal to the negative surface integral of the outward normal component of some flux over the surface bounding the region. We can interpret $\mathbf{J}(\mathbf{r}, t)$ as a *probability current*.

Eq. (A.7) must be true for all volumes, and so

$$\frac{\partial P}{\partial t} + \nabla \cdot \mathbf{J} = 0, \quad (\text{A.9})$$

which is the *continuity equation* for the probability.

Notice that Eq. (A.8) can be written as

$$\boxed{\mathbf{J}(\mathbf{r}, t) = \Re \left\{ \Psi^*(\mathbf{r}, t) \frac{\hbar}{im} \Psi(\mathbf{r}, t) \right\}}, \quad (\text{A.10})$$

and so $\mathbf{J}(\mathbf{r}, t)$ is a real vector field, which can be plotted as a function of space and time.

A.1.2 Gaussian Integral

The Gaussian integral, also known as the Euler–Poisson integral, is the integral of the Gaussian function $f(x) = e^{-x^2}$ over the entire real line. It is used to compute the normalising constant of the normal distribution and the same integral with finite limits is closely related to both the error function and the cumulative distribution function of the normal distribution. In physics this type of integral appears frequently, for example, to find the probability density of the ground state of the harmonic oscillator. This integral is also used in the path integral formulation, to find the propagator of the harmonic oscillator, and in statistical mechanics, to find its partition function. Although no elementary function exists for the error function, the Gaussian integral can be solved analytically.

A standard way to compute the Gaussian integral is by, with us first defining

$$I = \int_{-\infty}^{\infty} dx e^{-ax^2}, \quad (\text{A.11})$$

considering I^2 where we integrate over both x and y so that

$$I^2 = \int_{-\infty}^{\infty} dx e^{-ax^2} \int_{-\infty}^{\infty} dy e^{-ay^2} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} dx dy e^{-a(x^2+y^2)}. \quad (\text{A.12})$$

Transforming to plane polar coordinates then gives

$$\begin{aligned} I^2 &= 2\pi \int_0^{\infty} r dr e^{-ar^2} \\ &= \pi \int_0^{\infty} d(r^2) e^{-ar^2} \\ &= \pi \left[-\frac{1}{a} e^{-ar^2} \right]_0^{\infty} = \frac{\pi}{a} \end{aligned} \quad (\text{A.13})$$

Now just take the square root of Eq. (A.13),

$$\boxed{\int_{-\infty}^{\infty} dx e^{-ax^2} = \sqrt{\frac{\pi}{a}}}. \quad (\text{A.14})$$

Another useful result can be obtained by considering the partial derivative of Eq. (A.14) with respect to the parameter a ,

$$\frac{\partial}{\partial a} \int_{-\infty}^{\infty} dx e^{-ax^2} = \frac{\partial}{\partial a} \sqrt{\frac{\pi}{a}}, \quad (\text{A.15})$$

$$\int_{-\infty}^{\infty} dx x^2 e^{-ax^2} = \frac{1}{2a} \sqrt{\frac{\pi}{a}}. \quad (\text{A.16})$$

A.1.2.1 Relation to the Gamma Function

The integrand in Eq. (A.14) is an even function,

$$\int_{-\infty}^{\infty} dx e^{-x^2} = 2 \int_0^{\infty} dx e^{-x^2} \quad (\text{A.17})$$

Thus, after the change of variable $x = \sqrt{t}$, this turns into the Euler integral

$$2 \int_0^{\infty} dx e^{-x^2} = 2 \int_0^{\infty} dt \frac{1}{2} e^{-t} t^{-\frac{1}{2}} = \Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}. \quad (\text{A.18})$$

where $\Gamma(z) = \int_0^{\infty} dt t^{z-1} e^{-t}$ is the *gamma function*. This shows why the factorial of a half-integer is a rational multiple of $\sqrt{\pi}$.

A.1.2.2 Useful Integrals of Similar Form

$$\int_0^{\infty} dx x^{2n} e^{-\frac{x^2}{a^2}} = \sqrt{\pi} \frac{a^{2n+1} (2n-1)!!}{2^{n+1}}, \quad (\text{A.19})$$

$$\int_0^{\infty} dx x^{2n+1} e^{-\frac{x^2}{a^2}} = \frac{n!}{2} a^{2n+2}, \quad (\text{A.20})$$

$$\int_0^{\infty} dx x^{2n} e^{-bx^2} = \frac{(2n-1)!!}{b^n 2^{n+1}} \sqrt{\frac{\pi}{b}}, \quad (\text{A.21})$$

$$\int_0^{\infty} dx x^{2n+1} e^{-bx^2} = \frac{n!}{2b^{n+1}}, \quad (\text{A.22})$$

$$\int_0^{\infty} dx x^n e^{-x^2} = \frac{\Gamma\left(\frac{n+1}{2}\right)}{2b^{\frac{n+1}{2}}}, \quad (\text{A.23})$$

where n is an integer. An easy way to derive these is by differentiating under the integral sign.

$$\begin{aligned} \int_{-\infty}^{\infty} dx x^{2n} e^{-\alpha x^2} &= (-1)^n \int_{-\infty}^{\infty} dx \frac{\partial^n}{\partial \alpha^n} e^{-\alpha x^2} \\ &= (-1)^n \frac{\partial^n}{\partial \alpha^n} \int_{-\infty}^{\infty} e^{-\alpha x^2} \\ &= \sqrt{\pi} (-1)^n \frac{\partial^n}{\partial \alpha^n} \alpha^{-\frac{1}{2}} \\ &= \sqrt{\frac{\pi}{\alpha}} \frac{(2n-1)!!}{(2\alpha)^n} \end{aligned} \quad (\text{A.24})$$

One could also integrate by parts and find a recurrence relation to solve this. In a more general form, we have

$$\int_0^{\infty} dx x^n e^{-\alpha x^2} = \begin{cases} \frac{\Gamma\left(\frac{n+1}{2}\right)}{2\alpha^{\frac{n+1}{2}}} & (n > -1, a > 0) \\ \frac{(2k-1)!!}{\alpha^k 2^{k+1}} \sqrt{\frac{\pi}{\alpha}} & (n = 2k, k \text{ integer}, a > 0) \\ \frac{k!}{2\alpha^{k+1}} & (n = 2k+1, k \text{ integer}, a > 0) \end{cases}. \quad (\text{A.25})$$

A.2 Symmetry Transformations

We can define a *symmetry* as a transformation that leaves a system unchanged. When a measurement of an observable \hat{A} is carried out on a quantum system which is initially in a state $|\psi\rangle$, the probabilities of the various possible measurement outcomes are given by

$$P(\psi \rightarrow \phi_i) = |\langle \phi_i | \psi \rangle|^2 \quad (\text{A.26})$$

where the $|\phi_i\rangle$ are the eigenstates of \hat{A} . A *symmetry transformation* can be implemented by introducing a group of *unitary operators* \hat{U} which transforms all the states of the system such that

$$|\psi\rangle \rightarrow |\psi'\rangle = \hat{U}|\psi\rangle, \quad |\phi\rangle \rightarrow |\phi'\rangle = \hat{U}|\phi\rangle. \quad (\text{A.27})$$

The probabilities of the various possible measurement outcomes are left unchanged (invariant) under the symmetry operation if, for all states $|\phi_i\rangle$, we have

$$|\langle \phi'_i | \psi' \rangle|^2 = |\langle \phi_i | \psi \rangle|^2. \quad (\text{A.28})$$

A theorem due to Wigner states that the condition above can be satisfied in two distinct ways. The first possibility is that the overlaps themselves remain invariant,

$$\langle \phi'_i | \psi' \rangle = \langle \phi_i | \psi \rangle. \quad (\text{A.29})$$

In this case, we have

$$\langle \phi'_i | \psi' \rangle = \langle \hat{U} \phi_i | \hat{U} \psi \rangle = \langle \phi_i | \hat{U}^\dagger \hat{U} | \psi \rangle = \langle \phi_i | \psi \rangle, \quad (\text{A.30})$$

and the operator \hat{U} must be linear and unitary:

$$\hat{U} | \lambda \phi_i \rangle = \lambda | \phi'_i \rangle, \quad \hat{U}^\dagger \hat{U} = \hat{\mathbb{I}}. \quad (\text{A.31})$$

The second possibility is that the overlaps pick up a complex conjugation,

$$\langle \phi'_i | \psi' \rangle = \langle \phi_i | \psi \rangle^*. \quad (\text{A.32})$$

in which case the operator \hat{U} is antilinear and antiunitary:

$$\hat{U} | \lambda \phi_i \rangle = \lambda^* | \phi'_i \rangle, \quad \hat{U}^* \hat{U} = \hat{\mathbb{I}}. \quad (\text{A.33})$$

The symmetries we shall consider are all covered by the first possibility; they are described by linear, unitary operators \hat{U} . An important example of a symmetry requiring an antiunitary, rather than unitary, operator \hat{U} is time-reversal invariance, $t \rightarrow -t$.

Matrix elements of an observable \hat{H} transform as

$$\langle \psi | \hat{A} | \phi \rangle \rightarrow \langle \psi' | \hat{A} | \phi' \rangle = \langle \hat{U} \psi | \hat{A} \hat{U} | \phi \rangle = \langle \psi | \hat{U}^\dagger \hat{A} \hat{U} | \phi \rangle. \quad (\text{A.34})$$

Hence the effect of a symmetry transformation on a matrix element can be obtained by transforming the observable as

$$\hat{A} \rightarrow \hat{U}^\dagger \hat{A} \hat{U}, \quad (\text{A.35})$$

while leaving the states unchanged.

Symmetries can be classified as *discrete* or *continuous* depending on the transformation that generates them. The parity (or spatial inversion, $\mathbf{r} \rightarrow -\mathbf{r}$) symmetry is an example of discrete symmetry. The symmetry under spatial translations, $\mathbf{r} \rightarrow \mathbf{r} + \mathbf{a}$, is an example of continuous symmetry.

For continuous symmetries, such as spatial translations or rotations, the symmetry operation can be taken arbitrarily close to the identity operation (for example, a rotation through an infinitesimal angle). For such symmetries, the operator \hat{U} can be taken arbitrarily close to the identity operator $\hat{\mathbb{I}}$, in which case it must take the form

$$\hat{U}(\epsilon) = \hat{\mathbb{I}} + i\epsilon\hat{T} + \mathcal{O}(\epsilon^2), \quad (\text{A.36})$$

where ϵ is a small arbitrary real number, and \hat{T} is the *generator* of the symmetry transformation. Since the identity operator $\hat{\mathbb{I}}$ is unitary, continuous symmetries must always be described by operators \hat{U} which are unitary, rather than antiunitary. Imposing this condition we then have

$$\hat{U}^\dagger(\epsilon)\hat{U}(\epsilon) = (\hat{\mathbb{I}} - i\epsilon\hat{T}^\dagger + \mathcal{O}(\epsilon^2))(\hat{\mathbb{I}} + i\epsilon\hat{T} + \mathcal{O}(\epsilon^2)) = \hat{\mathbb{I}} + i\epsilon(\hat{T} - \hat{T}^\dagger) + \mathcal{O}(\epsilon^2), \quad (\text{A.37})$$

from which it follows that the operator \hat{T} must be Hermitian:

$$\hat{T} = \hat{T}^\dagger. \quad (\text{A.38})$$

(The factor of i was introduced in Eq. (A.36) for convenience, to ensure that \hat{T} is Hermitian rather than anti-Hermitian). The operator \hat{T} is known as the *generator* of the symmetry. In general, a symmetry will be described by a set of parameters ϵ_k , and as a result, the symmetry will possess a set of generators \hat{T}_k .

Any finite transformation \hat{U} can be built up from an infinite number of infinitesimal transformations. To see this, take $\epsilon = \theta/N$, where θ is a finite, N independent parameter, and apply N successive infinitesimal transformations, ϵ . This corresponds to applying the operator

$$\left(\hat{\mathbb{I}} + \frac{i\theta\hat{T}}{N}\right)^N = \hat{\mathbb{I}} + N\left(\frac{i\theta\hat{T}}{N}\right) + \frac{1}{2!}N(N-1)\left(\frac{i\theta\hat{T}}{N}\right)^2 + \frac{1}{3!}N(N-1)(N-2)\left(\frac{i\theta\hat{T}}{N}\right)^3 + \dots \quad (\text{A.39})$$

Taking the limit $N \rightarrow \infty$ (and hence $\epsilon \rightarrow 0$) then gives

$$\left(\hat{\mathbb{I}} + \frac{i\theta\hat{T}}{N}\right)^N \rightarrow \exp(i\theta\hat{T}). \quad (\text{A.40})$$

Thus, for finite θ , the symmetry operator \hat{U} can be written as

$$\hat{U}(\theta) = \exp(i\theta\hat{T}). \quad (\text{A.41})$$

Under a symmetry transformation, a matrix element of an observable \hat{A} transforms as

$$\langle \phi | \hat{A} | \psi \rangle \rightarrow \langle \phi' | \hat{A} | \psi' \rangle = \langle \hat{A}\phi | \hat{A} | \hat{U}\psi \rangle = \langle \phi | \hat{U}^\dagger \hat{A} \hat{U} | \psi \rangle. \quad (\text{A.42})$$

Hence the transformation properties of matrix elements can be found by subjecting operators to a *similarity transform*

$$\hat{A} \rightarrow \hat{U}^\dagger \hat{A} \hat{U}, \quad (\text{A.43})$$

while leaving the states of the system unaltered.

For an infinitesimal transformation, the similarity transformation is

$$\hat{A} \rightarrow (\hat{\mathbb{I}} - i\epsilon \hat{T}^\dagger + \mathcal{O}(\epsilon^2)) \hat{A} (\hat{\mathbb{I}} + i\epsilon \hat{T} + \mathcal{O}(\epsilon^2)), \quad (\text{A.44})$$

and hence

$$\hat{A} \rightarrow \hat{A} - i\epsilon [\hat{T}, \hat{A}] + \mathcal{O}(\epsilon^2). \quad (\text{A.45})$$

In the particular case that the magnitudes of all matrix elements of an observable \hat{A} are left unaltered by the symmetry transformation, we have

$$|\langle \phi | \hat{A} | \psi \rangle|^2 = |\langle \phi | \hat{U}^\dagger \hat{A} \hat{U} | \psi \rangle|^2. \quad (\text{A.46})$$

For this to hold for all possible choices of the states $|\phi\rangle$ and $|\psi\rangle$, the operator \hat{A} must satisfy the condition

$$\hat{A} = \hat{U}^\dagger \hat{A} \hat{U}. \quad (\text{A.47})$$

Since \hat{U} is unitary, this implies that \hat{A} must commute with \hat{U} , and hence also with the generator \hat{T} :

$$[\hat{A}, \hat{U}] = 0, \quad [\hat{A}, \hat{T}] = 0. \quad (\text{A.48})$$

In particular, if all matrix elements of the Hamiltonian \hat{H} of a system are unchanged by a symmetry transformation, then we must have

$$[\hat{H}, \hat{U}] = 0, \quad [\hat{H}, \hat{T}] = 0. \quad (\text{A.49})$$

Ehrenfest's theorem then immediately implies that, for any state of the system, the expectation value of the generator \hat{T} must remain constant:

$$\frac{d}{dt} \langle \psi | \hat{T} | \psi \rangle = 0. \quad (\text{A.50})$$

The generator \hat{T} is therefore *conserved*; its expectation values remain constant. Thus invariance of a system under a symmetry operation leads directly to a corresponding conservation law.

For an eigenstates of \hat{A} , $\hat{A} |\psi\rangle = a |\psi\rangle$, we have

$$\hat{A} |\psi'\rangle = \hat{A} \hat{U} |\psi\rangle = \hat{U} \hat{A} |\psi\rangle = a \hat{U} |\psi\rangle = a |\psi'\rangle. \quad (\text{A.51})$$

Provided the state $|\psi'\rangle$ is a distinct state to $|\psi\rangle$, the eigenvalue a must then be degenerate

symmetries \Leftrightarrow degeneracies.

A.2.1 Time Translations

The time-translation operator $\hat{U}(\tau)$ shifts the origin of the time coordinate by an amount τ ,

$$\hat{U}(\tau) |\psi(t)\rangle = |\psi(t + \tau)\rangle. \quad (\text{A.52})$$

If all probabilities are preserved by the time translation,

$$|\langle\psi(t + \tau)|\phi(t + \tau)\rangle|^2 = |\langle\psi(t)|\phi(t)\rangle|^2, \quad (\text{A.53})$$

then the operator $\hat{U}(\tau)$ is unitary, $\hat{U}^\dagger(\tau)\hat{U}(\tau) = \hat{\mathbb{I}}$.

For an infinitesimal time translation, $\hat{U}(\tau)$ must also be of the form

$$\hat{U}(\tau) = \hat{\mathbb{I}} - \frac{i}{\hbar}\hat{H}\tau + \mathcal{O}(\tau^2), \quad (\text{A.54})$$

where \hat{H} is the generator, which, by including the factor $1/\hbar$, we have chosen to have units of energy.

A finite time translation is then represented by

$$\hat{U}(\tau) = e^{-i\hat{H}\tau/\hbar}, \quad |\psi(t + \tau)\rangle = e^{-i\hat{H}\tau/\hbar} |\psi(t)\rangle. \quad (\text{A.55})$$

Choosing $t = 0$, and relabelling $\tau \rightarrow t$, all states must have a time dependence given by

$$|\psi(t)\rangle = e^{-i\hat{H}t/\hbar} |\psi(0)\rangle. \quad (\text{A.56})$$

Hence all states must satisfy the differential equation

$$\frac{\partial}{\partial t} |\psi(t)\rangle = -\frac{i}{\hbar}\hat{H} |\psi(t)\rangle. \quad (\text{A.57})$$

The Hamiltonian \hat{H} is seen to be the generator of time translations. The time-dependent Schrödinger equation is a consequence of invariance under time translations.

A.2.2 Spatial Translations

A spatial translation of a system through a vector displacement \mathbf{a} changes the expectation value of the position \mathbf{r}_n of the n 'th particle for any state $|\psi\rangle$ the system such that

$$\langle\psi'|\mathbf{r}_n|\psi'\rangle = \langle\psi|\mathbf{r}_n|\psi\rangle + \mathbf{a}. \quad (\text{A.58})$$

Introducing the symmetry operator $\hat{U}(\mathbf{a})$ defined such that

$$|\psi\rangle \rightarrow |\psi'\rangle = \hat{U}(\mathbf{a}) |\psi\rangle, \quad (\text{A.59})$$

then gives

$$\langle\psi|\hat{U}^\dagger(\mathbf{a})\mathbf{r}_n\hat{U}(\mathbf{a})|\psi\rangle = \langle\psi|\mathbf{r}_n|\psi\rangle + \underbrace{\langle\psi|\mathbf{a}|\psi\rangle}_{\mathbf{a}}. \quad (\text{A.60})$$

For this relation to hold for all possible choices of the state $|\psi\rangle$ then requires that

$$\hat{U}^\dagger(\mathbf{a})\mathbf{r}_n\hat{U}(\mathbf{a}) = \mathbf{r}_n + \mathbf{a}. \quad (\text{A.61})$$

If the probabilities of all measurement outcomes are left unchanged by the spatial translation, then $\hat{U}(\mathbf{a})$ must be unitary:

$$\hat{U}^\dagger(\mathbf{a})\hat{U}(\mathbf{a}) = \hat{\mathbb{I}}. \quad (\text{A.62})$$

For an infinitesimal translation, the operator $\hat{U}(\mathbf{a})$ must take the form

$$\hat{U}(\mathbf{a}) = \hat{\mathbb{I}} - \frac{i}{\hbar}(\hat{\mathbf{P}} \cdot \mathbf{a}) + \mathcal{O}(a^2), \quad (\text{A.63})$$

where $\hat{\mathbf{P}} = (\hat{P}_x, \hat{P}_y, \hat{P}_z)$ is a vector containing the generators \hat{P}_k of spatial translations. The factor $1/\hbar$ has been introduced in order that the operator $\hat{\mathbf{P}}$ has dimensions of momentum. Substituting Eq. (A.63) into Eq. (A.61) gives

$$\left(\hat{\mathbb{I}} - \frac{i}{\hbar}(\hat{\mathbf{P}} \cdot \mathbf{a}) + \mathcal{O}(a^2)\right)\hat{\mathbf{r}}_n\left(\hat{\mathbb{I}} - \frac{i}{\hbar}(\hat{\mathbf{P}} \cdot \mathbf{a}) + \mathcal{O}(a^2)\right) = \hat{\mathbf{r}}_n + \mathbf{a}. \quad (\text{A.64})$$

Multiplying out, and keeping only terms up to $\mathcal{O}(a)$, we obtain the commutation relation

$$\frac{i}{\hbar}[(\hat{\mathbf{P}} \cdot \mathbf{a}), \hat{\mathbf{r}}_n] = \mathbf{a}. \quad (\text{A.65})$$

The k^{th} component of the above equation is

$$\frac{i}{\hbar}\left[\sum_{j=1}^3(\hat{P}_j a_j), \hat{r}_{nk}\right] = a_k, \quad (\text{A.66})$$

or, equivalently

$$\frac{i}{\hbar}\sum_{j=1}^3[\hat{P}_j, \hat{r}_{nk}]a_j = \sum_{j=1}^3\delta_{jk}a_j. \quad (\text{A.67})$$

For this relation to hold for all possible translations \mathbf{a} , the coefficients of a_j on each side of the equation must be equal:

$$[\hat{r}_{n,k}, \hat{P}_j] = i\hbar\delta_{jk}. \quad (\text{A.68})$$

The commutation relation of the operator $\hat{\mathbf{P}}$ with the position operator $\hat{\mathbf{r}}_n$ is the same for all particles in the system; thus the generator \mathbf{P} must be associated with the system as a whole and we therefore interpret $\hat{\mathbf{P}}$ as the *total* momentum operator for the system.

The order in which two distinct spatial translations defined by vectors \mathbf{a} and \mathbf{b} are applied to a system is immaterial. The corresponding symmetry operators $\hat{U}(\mathbf{a})$ and $\hat{U}(\mathbf{b})$ therefore commute with each other:

$$\hat{U}(\mathbf{b})\hat{U}(\mathbf{a}) = \hat{U}(\mathbf{a})\hat{U}(\mathbf{b}). \quad (\text{A.69})$$

If \mathbf{a} and \mathbf{b} are both infinitesimal translations, this becomes

$$\left(\hat{\mathbb{I}} - \frac{i}{\hbar}(\hat{\mathbf{P}} \cdot \mathbf{b}) + \dots\right)\left(\hat{\mathbb{I}} - \frac{i}{\hbar}(\hat{\mathbf{P}} \cdot \mathbf{a}) + \dots\right) = \left(\hat{\mathbb{I}} - \frac{i}{\hbar}(\hat{\mathbf{P}} \cdot \mathbf{a}) + \dots\right)\left(\hat{\mathbb{I}} - \frac{i}{\hbar}(\hat{\mathbf{P}} \cdot \mathbf{b}) + \dots\right). \quad (\text{A.70})$$

Multiplying out, we obtain

$$(\hat{\mathbf{P}} \cdot \mathbf{b})(\hat{\mathbf{P}} \cdot \mathbf{a}) = (\hat{\mathbf{P}} \cdot \mathbf{a})(\hat{\mathbf{P}} \cdot \mathbf{b}) \quad (\text{A.71})$$

$$\implies \sum_{j,k} \hat{P}_k \hat{P}_j b_k a_j = \sum_{j,k} \hat{P}_j \hat{P}_k a_j b_k. \quad (\text{A.72})$$

Equating the coefficients of $a_j b_k$ on both sides then shows that the components \hat{P}_x , \hat{P}_y , \hat{P}_z of the total momentum operator $\hat{\mathbf{P}}$ must mutually commute

$$[\hat{P}_i, \hat{P}_j] = 0. \quad (\text{A.73})$$

If the expectation values of an operator \hat{A} are invariant under translations for all possible states of the system,

$$\langle \phi' | \hat{A} | \psi' \rangle = \langle \phi | \hat{A} | \psi \rangle, \quad (\text{A.74})$$

then we must have

$$\hat{U}^\dagger(\mathbf{a}) \hat{A} \hat{U}(\mathbf{a}) = \hat{A}. \quad (\text{A.75})$$

Since $\hat{U}(\mathbf{a})$ is unitary, this implies that \hat{A} commutes with $\hat{U}(\mathbf{a})$:

$$[\hat{A}, \hat{U}(\mathbf{a})] = 0. \quad (\text{A.76})$$

Taking $\hat{U}(\mathbf{a})$ to correspond to an infinitesimal translation, as in Eq. (A.63), then gives

$$[\hat{A}, \hat{\mathbf{P}} \cdot \mathbf{a}] = 0. \quad (\text{A.77})$$

For this to hold for all possible translations \mathbf{a} , the operator \hat{A} must commute with the generators of the translation:

$$[\hat{A}, \hat{P}_i] = 0, \quad [\hat{A}, \hat{\mathbf{P}}] = 0. \quad (\text{A.78})$$

The Hamiltonian for an *isolated* system must be invariant under spatial translations (the energy eigenvalues can not depend on where the system happens to be located)

$$\langle \psi' | \hat{H} | \phi' \rangle = \langle \hat{U}(\mathbf{a}) \psi | \hat{H} | \hat{U}(\mathbf{a}) \phi \rangle = \langle \psi | \hat{U}^\dagger(\mathbf{a}) \hat{H} \hat{U}(\mathbf{a}) | \phi \rangle = \langle \psi | \phi \rangle. \quad (\text{A.79})$$

Hence we must have, for all operators $\hat{U}(\mathbf{a})$,

$$\hat{U}^\dagger(\mathbf{a}) \hat{H} \hat{U}(\mathbf{a}) = \hat{H}, \quad [\hat{H}, \hat{U}(\mathbf{a})] = 0. \quad (\text{A.80})$$

In particular, considering an infinitesimal spatial translation, we must have

$$[\hat{H}, \hat{P}_j] = 0. \quad (\text{A.81})$$

From Ehrenfest's theorem, the expectation values, $\langle \hat{P}_x \rangle$, $\langle \hat{P}_y \rangle$, $\langle \hat{P}_z \rangle$, of each component of the total linear momentum are therefore conserved:

$$\frac{d}{dt} \langle \psi | \hat{P}_j | \psi \rangle = 0. \quad (\text{A.82})$$

We thus establish that conservation of total linear momentum for an isolated system is a consequence of invariance under spatial translations.

Consider, for example, a one-dimensional system described by states $\psi(x)$ which depend on the spatial coordinate x . Applying a spatial translation transforms the wavefunction $\psi(x)$ as

$$\psi(x) \rightarrow \psi'(x) = \psi(x - a) \equiv \hat{U}(a)\psi(x). \quad (\text{A.83})$$

For the case of an infinitesimal translation, the transformed wavefunction is

$$\psi(x - a) = \psi(x) - a \frac{\partial\psi(x)}{\partial x} + \dots = \left(\hat{\mathbb{I}} - a \frac{\partial}{\partial x} + \dots \right) \psi(x). \quad (\text{A.84})$$

Hence, we can identify

$$\hat{P}_x = -i\hbar \frac{\partial}{\partial x}. \quad (\text{A.85})$$

Thus the form $\hat{\mathbf{P}} = -i\hbar\nabla$ is coherent with the fact that the momentum operator is the generator of spatial translations.

A.2.3 Spatial Rotations

A spatial rotation is a real, linear transformation of spatial three-vectors $\mathbf{r} = (r_x, r_y, r_z) = (r_1, r_2, r_3)$ of the form

$$r_i \rightarrow r'_i = \sum_{j=1}^3 R_{ij} r_j \quad (\text{A.86})$$

that leaves all three-vector scalar products $\mathbf{x} \cdot \mathbf{y}$ invariant:

$$\sum_i x'_i y'_i = \sum_i \left(\sum_j R_{ij} x_j \right) \left(\sum_k R_{ik} y_k \right) = \sum_i x_i y_i. \quad (\text{A.87})$$

Writing the right-hand side of the above equation as

$$\sum_i x_i y_i = \sum_i \left(\sum_j \delta_{ij} x_j \right) \left(\sum_k \delta_{ik} y_k \right), \quad (\text{A.88})$$

and equating the coefficients of $x_j y_k$ on both sides then gives

$$\sum_i R_{ij} R_{ik} = \delta_{jk}. \quad (\text{A.89})$$

In matrix notation, this is

$$R^T R = \mathbb{I}, \quad (\text{A.90})$$

where R and R^T are matrices which contain the elements

$$[R]_{ij} = [R^T]_{ji} = R_{ij}. \quad (\text{A.91})$$

Using the identities $\det(R^T R) = \det(R^T) \det(R)$ and $\det(R^T) = \det(R)$, Eq. (A.90) then implies that the determinant of R is

$$\det(R) = \pm 1. \quad (\text{A.92})$$

Transformations with $\det(R) = +1$ are spatial rotations; transformations with $\det(R) = -1$ are spatial inversions.

For an infinitesimal rotation, we can write the matrix R in the form

$$R = \mathbb{I} + \omega + \mathcal{O}(\omega^2); \quad R_{ij} = \delta_{ij} + \omega_{ij} + \mathcal{O}(\omega^2), \quad (\text{A.93})$$

where the ω_{ij} are infinitesimal parameters. The condition in Eq. (A.90) then gives

$$\mathbb{I} = [\mathbb{I} + \omega^T + \mathcal{O}(\omega^2)][\mathbb{I} + \omega + \mathcal{O}(\omega^2)] = \mathbb{I} + \omega^T + \omega + \mathcal{O}(\omega^2). \quad (\text{A.94})$$

Hence the matrix ω must satisfy the constraint

$$\omega^T = -\omega, \quad \omega_{ji} = -\omega_{ij}, \quad (\text{A.95})$$

and so is of the form

$$\omega = \begin{pmatrix} 0 & \omega_{12} & \omega_{13} \\ -\omega_{12} & 0 & \omega_{23} \\ -\omega_{13} & -\omega_{23} & 0 \end{pmatrix}. \quad (\text{A.96})$$

Ω has only *three* independent parameters that define the direction of the axis of rotation and the angle of rotation. To find the explicit form of the matrix ω for a general infinitesimal rotation, we begin by considering a rotation through a finite angle ϕ about the z -axis. The rotation matrix R_z corresponding to such a rotation is

$$R_z = \begin{pmatrix} \cos \phi & \sin \phi & 0 \\ -\sin \phi & \cos \phi & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad (\text{A.97})$$

This can be considered to be either an *active* rotation of the position vector \mathbf{r} , keeping the coordinate axes fixed, or, equivalently, a *passive* rotation of the coordinate axes, keeping all positions (the apparatus) fixed.

The sign convention in Eq. (A.97) corresponds to the left-hand rule (right-hand rule) for active (passive) rotations: when the thumb is oriented along the rotation axis (the $+z$ direction), the direction of the fingers gives the sense of the rotation for positive $\phi > 0$.

When the angle ϕ is infinitesimal, the matrix R_z becomes, to leading order,

$$R_z = \begin{pmatrix} 1 & \phi & 0 \\ -\phi & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \mathcal{O}(\phi^2), \quad \omega = \begin{pmatrix} 0 & \phi & 0 \\ -\phi & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \quad (\text{A.98})$$

To find the corresponding matrices for rotations about the x and y axes, cycle the coordinates as $(x, y, z) \rightarrow (y, z, x) \rightarrow (z, x, y)$:

For example, the matrix element $[R_z]_{xy}$ (equal to $\sin \phi$) for a rotation about the z -axis becomes the matrix element $[R_y]_{zx}$ for a rotation about the y -axis. This gives

$$R_x = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \phi & \sin \phi \\ 0 & -\sin \phi & \cos \phi \end{pmatrix}, \quad R_y = \begin{pmatrix} \cos \phi & 0 & -\sin \phi \\ 0 & 1 & 0 \\ \sin \phi & 0 & \cos \phi \end{pmatrix}. \quad (\text{A.99})$$

For a general infinitesimal rotation, combining three matrices of the forms given in Eq. (A.98) and (A.99) then gives

$$\omega = \begin{pmatrix} 0 & \omega_3 & -\omega_2 \\ -\omega_3 & 0 & \omega_1 \\ \omega_2 & -\omega_1 & 0 \end{pmatrix} \quad (\text{A.100})$$

This corresponds to a rotation through an infinitesimal angle $|\omega|$ about an axis oriented in the direction of the vector $\omega = (\omega_1, \omega_2, \omega_3)$.

For two successive rotations R_1 and R_2 , satisfying $R_1^T R_1 = \mathbb{I}$ and $R_2^T R_2 = \mathbb{I}$, we have

$$(R_2 R_1)^T (R_2 R_1) = (R_1)^T (R_2)^T (R_2 R_1) = R_1^T R_1 = \mathbb{I}. \quad (\text{A.101})$$

Thus the matrix product $R_2 R_1$ satisfies the condition in Eq. (A.90), showing that two successive rotations correspond to a single, overall spatial rotation.

Since, in general, matrices do not commute, neither do rotations; applying R_1 followed by R_2 will in general give a different overall rotation to applying R_2 followed by R_1 .

A.3 The Wigner-Eckart Theorem

Proofs of the Wigner-Eckart theorem for the particular cases of a scalar operator, \hat{K} , and a vector operator, $\hat{\mathbf{V}}$, are presented below.

A.3.1 Scalar and Vector Operators

The Wigner-Eckart theorem dictates the form of matrix elements such as $\langle \alpha'' j'' m'' | \hat{K} | \alpha' j' m' \rangle$ and $\langle \alpha'' j'' m'' | \hat{\mathbf{V}} | \alpha' j' m' \rangle$ between eigenstates $|\alpha jm\rangle$ of an angular momentum operator $\hat{\mathbf{J}}$. In this context, the operator $\hat{\mathbf{J}}$ is any operator whose components $(\hat{J}_1, \hat{J}_2, \hat{J}_3)$ satisfy the commutation relations

$$[\hat{J}_i, \hat{J}_j] = i\hbar\epsilon_{ijk}\hat{J}_k. \quad (\text{A.102})$$

An operator \hat{K} is a scalar operator with respect to $\hat{\mathbf{J}}$ if it commutes with all components of $\hat{\mathbf{J}}$:

$$[\hat{J}_i, \hat{K}] = 0; \quad [\hat{\mathbf{J}}, \hat{K}] = 0; \quad [\hat{\mathbf{J}}^2, \hat{K}] = 0. \quad (\text{A.103})$$

An operator $\hat{\mathbf{V}} = (\hat{V}_1, \hat{V}_2, \hat{V}_3)$ is a vector operator with respect to $\hat{\mathbf{J}}$ if its components satisfy the commutation relations

$$[\hat{J}_i, \hat{V}_j] = i\hbar \sum_k \epsilon_{ijk} \hat{V}_k. \quad (\text{A.104})$$

Eq. (A.102)–(A.104) are the only requirements which the operators $\hat{\mathbf{J}}$, \hat{K} and $\hat{\mathbf{V}}$ must satisfy in order for the Wigner-Eckart theorem to be applicable.

A.3.2 Angular Momentum Relations

The operators $\hat{\mathbf{J}}^2$ and $\hat{J}_z = \hat{J}_3$ possess a set of simultaneous eigenstates, $|\alpha jm\rangle$, such that

$$\hat{\mathbf{J}}^2 |\alpha jm\rangle = j(j+1)\hbar^2 |\alpha jm\rangle, \quad \hat{J}_z |\alpha jm\rangle = m\hbar |\alpha jm\rangle, \quad (\text{A.105})$$

where j and m are the usual angular momentum quantum numbers, and α collectively labels all other quantum numbers needed to uniquely identify the states. The ladder operators $\hat{J}_{\pm} \equiv \hat{J}_1 \pm i\hat{J}_2$ acting on the state $|\alpha jm\rangle$ raise or lower the index m by ± 1 :

$$\hat{J}_{\pm} |\alpha jm\rangle = \hbar \sqrt{j(j+1) - m(m \pm 1)} |\alpha j, m \pm 1\rangle. \quad (\text{A.106})$$

Noting that $(\hat{J}_{\pm})^\dagger = \hat{J}_{\mp}$, the conjugate of this relation is

$$\langle \alpha jm | \hat{J}_{\mp} = \hbar \sqrt{j(j+1) - m(m \pm 1)} \langle \alpha j, m \pm 1|. \quad (\text{A.107})$$

Writing this equation in terms of \hat{J}_{\pm} rather than \hat{J}_{\mp} (i.e. interchanging all “ \pm ” and “ \mp ” signs), we obtain

$$\langle \alpha jm | \hat{J}_{\pm} = \hbar \sqrt{j(j+1) - m(m \mp 1)} \langle \alpha j, m \mp 1|. \quad (\text{A.108})$$

A.3.3 The Wigner-Eckart Theorem for Scalar Operators

Since the scalar operator \hat{K} commutes with \hat{J}_z , we have, for any pair of total angular momentum eigenstates $|\alpha''j''m''\rangle$ and $|\alpha'j'm'\rangle$,

$$0 = \langle \alpha''j''m'' | [\hat{J}_z, \hat{K}] | \alpha'j'm' \rangle \quad (\text{A.109})$$

$$= \hbar(m'' - m') \langle \alpha''j''m'' | \hat{K} | \alpha'j'm' \rangle. \quad (\text{A.110})$$

Therefore, if $m'' \neq m'$, the matrix element $\langle \alpha''j''m'' | \hat{K} | \alpha'j'm' \rangle$ must vanish. Similarly, since \hat{K} commutes with $\hat{\mathbf{J}}^2$, we have

$$0 = \langle \alpha''j''m'' | [\hat{\mathbf{J}}^2, \hat{K}] | \alpha'j'm' \rangle = \hbar^2 [j''(j''+1) - j'(j'+1)] \langle \alpha''j''m'' | \hat{K} | \alpha'j'm' \rangle. \quad (\text{A.111})$$

Hence, if $j'' \neq j'$, the matrix element $\langle \alpha''j''m'' | \hat{K} | \alpha'j'm' \rangle$ must vanish. Overall therefore, we must have

$$\langle \alpha''j''m'' | \hat{K} | \alpha'j'm' \rangle = \delta_{j''j'} \delta_{m''m'} C(\alpha''\alpha'; j'm'), \quad (\text{A.112})$$

where $C(\alpha''\alpha'; j'm')$ is a constant. For $j'' = j'$ and $m'' = m'$, the equation above becomes

$$\langle \alpha''j'm' | \hat{K} | \alpha'j'm' \rangle = C(\alpha''\alpha'; j'm'). \quad (\text{A.113})$$

Since \hat{K} commutes with the ladder operators \hat{J}_{\pm} , we have

$$\langle \alpha''j''m'' | \hat{J}_{\pm} \hat{K} | \alpha'j'm' \rangle = \langle \alpha''j''m'' | \hat{K} \hat{J}_{\pm} | \alpha'j'm' \rangle. \quad (\text{A.114})$$

Using Eq. (A.106) and (A.108) then gives

$$\begin{aligned} & \sqrt{j''(j''+1) - m''(m'' \mp 1)} \langle \alpha''j'', m'' \mp 1 | \hat{K} | \alpha'j'm' \rangle \\ &= \sqrt{j'(j'+1) - m'(m' \pm 1)} \langle \alpha''j''m'' | \hat{K} | \alpha'j', m' \pm 1 \rangle. \end{aligned} \quad (\text{A.115})$$

For $j'' = j'$ and $m'' = m' \pm 1$, the square-root factors on either side of the equation above are equal to each other; cancelling these factors then leaves

$$\langle \alpha'' j' m' | \hat{K} | \alpha' j' m' \rangle = \langle \alpha'' j', m' \pm 1 | \hat{K} | \alpha' j', m' \pm 1 \rangle. \quad (\text{A.116})$$

Comparing with Eq. (A.113) then gives

$$C(\alpha'' \alpha'; j' m') = C(\alpha'' \alpha'; j', m' \pm 1). \quad (\text{A.117})$$

Thus the constants $C(\alpha'' \alpha'; j' m')$ are, in fact, independent of the quantum number m' , and can be written as $C(\alpha'' \alpha'; j')$. Eq. (A.112) then gives the *Wigner-Eckart theorem* for scalar operators,

$\langle \alpha'' j'' m'' | \hat{K} | \alpha' j' m' \rangle = C(\alpha'' \alpha'; j') \delta_{j'' j'} \delta_{m'' m'}.$

(A.118)

as in Eq. (9.84).

A.3.4 The Wigner-Eckart Theorem for Vector Operators

The commutation relations of the vector operator $\hat{\mathbf{V}}$ with the ladder operators \hat{J}_\pm were obtained in Eq. (9.71):

$$[\hat{J}_\pm, \hat{V}_m] = \hbar \sqrt{j(j+1) - m(m \pm 1)} \hat{V}_{m \pm 1}, \quad (j \equiv 1, m = 0, \pm 1). \quad (\text{A.119})$$

Taking matrix elements of Eq. (9.71) between eigenstates $|\alpha' j' m'\rangle$ and $|\alpha'' j'' m''\rangle$ of the total angular momentum operator $\hat{\mathbf{J}}$ gives

$$\langle \alpha'' j'' m'' | (\hat{J}_\pm \hat{K}_m - \hat{K}_m \hat{J}_\pm) | \alpha' j' m' \rangle = \hbar \sqrt{j(j+1) - m(m \pm 1)} \langle \alpha'' j'' m'' | \hat{K}_{m \pm 1} | \alpha' j' m' \rangle, \quad (\text{A.120})$$

where

$$j \equiv 1, \quad m = 0, \pm 1. \quad (\text{A.121})$$

From Eq. (A.106) and (A.108), the ladder operators act on the states $\langle \alpha'' j'' m'' |$ and $|\alpha' j' m'\rangle$ as

$$\langle \alpha'' j'' m'' | \hat{J}_\pm = \hbar \sqrt{j''(j''+1) - m''(m'' \mp 1)} \langle \alpha'' j'', m'' \mp 1 |, \quad (\text{A.122})$$

$$\hat{J}_\pm |\alpha' j' m'\rangle = \hbar \sqrt{j'(j'+1) - m'(m' \pm 1)} |\alpha' j', m' \pm 1\rangle. \quad (\text{A.123})$$

Substituting the above into the left-hand side of Eq. (A.120) and rearranging, we obtain the following relation between spherical component matrix elements:

$$\begin{aligned} & \sqrt{j''(j''+1) - m''(m'' \mp 1)} \langle \alpha'' j'', m'' \mp 1 | \hat{V}_m | \alpha' j' m' \rangle \\ &= \sqrt{j(j+1) - m(m \pm 1)} \langle \alpha'' j'' m'' | \hat{V}_{m \pm 1} | \alpha' j' m' \rangle \\ &+ \sqrt{j'(j'+1) - m'(m' \pm 1)} \langle \alpha'' j'' m'' | \hat{V}_m | \alpha' j', m' \pm 1 \rangle. \end{aligned} \quad (\text{A.124})$$

This relation connects matrix elements with common values of j' and j'' throughout, but with varying m , m' and m'' .

Now consider the addition of two angular momentum operators $\hat{\mathbf{J}}$ and $\hat{\mathbf{J}'}$ to form a combined angular momentum $\hat{\mathbf{J}''}$:

$$|jm; j'm'\rangle \equiv |jm\rangle \otimes |j'm'\rangle. \quad (\text{A.125})$$

The action of the ladder operator $\hat{J}_{\pm}'' = \hat{J}_{\pm} + \hat{J}'_{\pm}$ on these product states is

$$\begin{aligned} \hat{J}_{\pm}'' |jm; j'm'\rangle &= \hbar \sqrt{j(j+1) - m(m \pm 1)} |j, m \pm 1; j'm'\rangle \\ &\quad + \hbar \sqrt{j'(j'+1) - m'(m' \pm 1)} |jm; j'm', \pm 1\rangle. \end{aligned} \quad (\text{A.126})$$

Premultiplying by an eigenstate $\langle j''m''|$ of $\hat{\mathbf{J}}''$ then gives

$$\begin{aligned} \langle j''m''| \hat{J}_{\pm}'' |jm; j'm'\rangle &= \hbar \sqrt{j(j+1) - m(m \pm 1)} \langle j''m''|j, m \pm 1; j'm'\rangle \\ &\quad + \hbar \sqrt{j'(j'+1) - m'(m' \pm 1)} \langle j''m''|jm; j'm', \pm 1\rangle. \end{aligned} \quad (\text{A.127})$$

Taking the complex conjugate of the above equation, and using $(\hat{J}_{\pm}'')^\dagger = \hat{J}_{\mp}''$, we obtain

$$\begin{aligned} \langle jm; j'm'| \hat{J}_{\mp}'' |j''m''\rangle &= \hbar \sqrt{j(j+1) - m(m \pm 1)} \langle j, m \pm 1; j'm'|j''m''\rangle \\ &\quad + \hbar \sqrt{j'(j'+1) - m'(m' \pm 1)} \langle jm; j'm', \pm 1|j''m''\rangle. \end{aligned} \quad (\text{A.128})$$

The matrix element on the left-hand side can be evaluated using

$$\hat{J}_{\mp}'' |j''m''\rangle = \hbar \sqrt{j''(j''+1) - m''(m'' \mp 1)} |j'', m'' \pm 1\rangle. \quad (\text{A.129})$$

Hence we obtain the following relation between Clebsch-Gordan coefficients of fixed j , j' and j'' , but varying m , m' and m'' :

$$\begin{aligned} \sqrt{j''(j''+1) - m''(m'' \mp 1)} \langle jm; j'm'|j'', m'' \pm 1\rangle \\ &= \hbar \sqrt{j(j+1) - m(m \pm 1)} \langle j, m \pm 1; j'm'|j''m''\rangle \\ &\quad + \hbar \sqrt{j'(j'+1) - m'(m' \pm 1)} \langle jm; j'm', \pm 1|j''m''\rangle. \end{aligned} \quad (\text{A.130})$$

This relation is *identical in form* to Eq. (A.124), under the replacements

$$\langle \alpha''j''m''|\hat{V}_{m\pm 1}|\alpha'j'm'\rangle \longleftrightarrow \langle j, m \pm 1; j'm'|j''m''\rangle. \quad (\text{A.131})$$

Thus Eq. (A.124) and (A.130) represent two identical sets of linear, homogenous equations of the form

$$\sum_j A_{ij} x_j = 0, \quad \sum_j A_{ij} y_j = 0. \quad (\text{A.132})$$

In the first set of equations, the quantities x_j are matrix elements $\langle \alpha''j''m''|\hat{V}_{m\pm 1}|\alpha'j'm'\rangle$ of $\hat{\mathbf{V}}$. In the second set, the y_j are Clebsch-Gordan coefficients $\langle j, m \pm 1; j'm'|j''m''\rangle$. The matrix A_{ij} contains all the factors of the form $\sqrt{j(j+1) - m(m \pm 1)}$, and is common to both sets of equations.

The solution to linear equations of the form given in Eq. (A.132) is unique up to an overall constant. Hence the solutions x_j and y_j must be directly proportional to each other:

$$\langle \alpha''j''m''|\hat{V}_{m\pm 1}|\alpha'j'm'\rangle = C(\alpha''\alpha'; j''j') \langle j, m \pm 1; j'm'|j''m''\rangle. \quad (\text{A.133})$$

The (complex) constant of proportionality, $C(\dots)$, must be independent of m , m' and m'' because these parameters vary from element to element within the matrix A_{ij} . On the other hand, the parameters α'', α', j'' and j' are common to all the elements A_{ij} .

Replacing $m \pm 1$ by m throughout Eq. (A.133), we obtain the Wigner-Eckart theorem for a vector operator $\hat{\mathbf{V}}$:

$$\langle \alpha'' j'' m'' | \hat{V}_m | \alpha' j' m' \rangle = C(\alpha'' \alpha'; j'' j') \langle jm; j' m' | j'' m'' \rangle. \quad (\text{A.134})$$

The constant $C(\alpha'' \alpha'; j'' j')$, the reduced matrix element, is usually written in Dirac-style notation as $\langle \alpha'' j'' | \hat{\mathbf{V}} | \alpha' j' \rangle$, giving the Wigner-Eckart Theorem in standard form as

$$\boxed{\langle \alpha'' j'' m'' | \hat{V}_m | \alpha' j' m' \rangle = \langle \alpha'' j'' | \hat{\mathbf{V}} | \alpha' j' \rangle \langle jm; j' m' | j'' m'' \rangle}, \quad (\text{A.135})$$

as in Eq. (9.95). In the equation above, the quantum number j is, by definition, $j \equiv 1$, and the quantum number m takes the values $m = \pm 1, 0$.

A.4 Virial Theorem

The virial theorem is an important theorem for a system of moving particles both in classical physics and quantum physics. The virial Theorem is useful when considering a collection of many particles and has a special importance to central-force motion. For a general system of mass points with position vectors \mathbf{r}_i and applied forces \mathbf{F}_i , consider the scalar product G

$$G \equiv \sum_i \mathbf{p}_i \cdot \mathbf{r}_i, \quad (\text{A.136})$$

where i sums over all particles. The time derivative of G is

$$\frac{dG}{dt} = \sum_i \mathbf{p}_i \cdot \dot{\mathbf{r}}_i + \sum_i \dot{\mathbf{p}}_i \cdot \mathbf{r}_i. \quad (\text{A.137})$$

However,

$$\sum_i \mathbf{p}_i \cdot \dot{\mathbf{r}}_i = \sum_i m \dot{\mathbf{r}}_i \cdot \dot{\mathbf{r}}_i = \sum_i mv_i^2 = 2T. \quad (\text{A.138})$$

Also, since

$$\sum_i \dot{\mathbf{p}}_i \cdot \mathbf{r}_i = \sum_i \mathbf{F}_i \cdot \mathbf{r}_i. \quad (\text{A.139})$$

Thus,

$$\frac{dG}{dt} = 2T + \sum_i \mathbf{F}_i \cdot \mathbf{r}_i. \quad (\text{A.140})$$

The time average over a period τ is

$$\frac{1}{\tau} \int_0^\tau \frac{dG}{dt} dt = \frac{G(\tau) - G(0)}{\tau} = \langle 2T \rangle + \left\langle \sum_i \mathbf{F}_i \cdot \mathbf{r}_i \right\rangle. \quad (\text{A.141})$$

where the $\langle \rangle$ brackets refer to the time average. Note that if the motion is periodic and the chosen time τ equals a multiple of the period, then $\frac{G(\tau) - G(0)}{\tau} = 0$. Even if the motion

is not periodic, if the constraints and velocities of all the particles remain finite, then there is an upper bound to G . This implies that choosing $\tau \rightarrow \infty$ means that $\frac{G(\tau) - G(0)}{\tau} \rightarrow 0$. In both cases the left-hand side of the equation tends to zero giving the virial theorem:

$$\langle T \rangle = -\frac{1}{2} \left\langle \sum_i \mathbf{F}_i \cdot \mathbf{r}_i \right\rangle. \quad (\text{A.142})$$

The right-hand side of this equation is called the virial of the system. For a single particle subject to a conservative central force $\mathbf{F} = -\nabla U$ the virial theorem equals

$$\langle T \rangle = \frac{1}{2} \langle \nabla U \cdot \mathbf{r} \rangle = \frac{1}{2} \left\langle r \frac{\partial U}{\partial r} \right\rangle. \quad (\text{A.143})$$

If the potential is of the form $U = kr^{n+1}$ that is, $\mathbf{F} = -k(n+1)r^n$, then $r \frac{\partial U}{\partial r} = (n+1)U$. Thus for a single particle in a central potential $U = kr^{n+1}$ the virial theorem reduces to

$$\langle T \rangle = \frac{n+1}{2} \langle U \rangle. \quad (\text{A.144})$$

The following two special cases are of considerable importance in physics.

- **Hooke's Law:** Note that for a linear restoring force $n = 1$ then

$$\langle T \rangle = +\langle U \rangle \quad (\text{A.145})$$

Which should be familiar, for simple harmonic motion where the average kinetic and potential energies are the same and both equal half of the total energy.

- **Inverse-Square Law:** The other interesting case is for the inverse square law $n = -2$ where

$$\langle T \rangle = -\frac{1}{2} \langle U \rangle \quad (\text{A.146})$$

The virial theorem is useful for solving problems in that knowing the exponent n of the field makes it possible to write down directly the average total energy in the field. For example, for

$$\langle E \rangle = \langle T \rangle + \langle U \rangle \quad (\text{A.147})$$

$$= -\frac{1}{2} \langle U \rangle + \langle U \rangle = \frac{1}{2} \langle U \rangle \quad (\text{A.148})$$

This occurs for the Bohr model of the hydrogen atom where the kinetic energy of the bound electron is half of the potential energy. The same result occurs for planetary motion in the solar system.

A.5 Density of States

A.5.1 2D Density of States

Consider a particle confined to the 2D square region

$$0 < x < L, \quad 0 < y < L. \quad (\text{A.149})$$

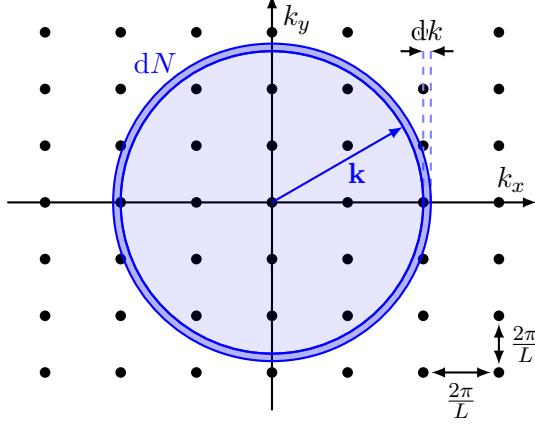


Fig. A.1: Quantised locations of allowed \mathbf{k} vectors in k -space, showing the annulus containing a number of available states dN between k and $k + dk$, where $k = |\mathbf{k}|$.

If we impose “hard wall” boundary conditions, $\Psi(\mathbf{r}) = 0$ on the sides of the square region, the allowed \mathbf{k} vectors are

$$\mathbf{k} = (k_x, k_y) = \left(\frac{\pi n_x}{L}, \frac{\pi n_y}{L} \right), \quad n_x, n_y = 1, 2, 3, \dots . \quad (\text{A.150})$$

The vector \mathbf{k} is restricted to lie in the positive quadrant,

$$k_x > 0, \quad k_y > 0, \quad (\text{A.151})$$

on a grid of points with spacing

$$\Delta k_x = \Delta k_y = \frac{\pi}{L}. \quad (\text{A.152})$$

In this case, the number of states dN with wavenumber between k and $k + dk$ is obtained by counting states within one quarter of an annulus

$$dN = \frac{1}{4} \times \frac{2\pi k dk}{(\pi/L)^2} = \frac{kL^2}{2\pi} dk \quad (k = |\mathbf{k}|). \quad (\text{A.153})$$

Alternatively, we can restrict the available states by imposing *periodic* boundary conditions in the x and y -directions

$$\psi(x + L, y) = \psi(x, y + L) = \psi(x, y). \quad (\text{A.154})$$

In this case, the allowed states are separated in 2D \mathbf{k} -space by intervals

$$\Delta k_x = \Delta k_y = 2\pi/L, \quad (\text{A.155})$$

and \mathbf{k} is no longer restricted to a single quadrant (k_x and k_y can be of either sign). This gives the same number of states dN as for hard-wall boundary conditions

$$dN = \frac{2\pi k dk}{(2\pi/L)^2} = \frac{kL^2}{2\pi} dk. \quad (\text{A.156})$$

For a non-relativistic particle of mass m , the density of states can be converted from wavenumber k to energy E using

$$E = \frac{\hbar^2 k^2}{2m} \quad (\text{A.157})$$

$$\frac{dN}{dE} = \frac{dN}{dk} \frac{dk}{dE} = \frac{kL^2}{2\pi} \frac{m}{\hbar^2 k} = \frac{L^2 m}{2\pi \hbar^2}. \quad (\text{A.158})$$

The density of states per unit area, $g(E) = dN/dE / L^2$, is therefore

$$g(E) = \frac{m}{2\pi \hbar^2}. \quad (\text{A.159})$$

Thus, in 2D, $g(E)$ is in fact a constant, independent of energy.

The density of states above applies to spin-zero particles. For spin-half, for example, an extra factor of two is needed to account for the spin degrees of freedom

$$g(E) = \frac{m}{\pi \hbar^2}. \quad (\text{A.160})$$

A.5.2 3D Density of States

In 3D, we calculate the density of states $g(E)$ by considering a particle confined within a cubical box of side L .

Imposing hard-wall boundary conditions $\Psi(\mathbf{r}) = 0$ on the sides of the box restricts the allowed \mathbf{k} vector components to be

$$\mathbf{k} = \left(\frac{\pi n_x}{L}, \frac{\pi n_y}{L}, \frac{\pi n_z}{L} \right), \quad n_x, n_y, n_z = 1, 2, 3, \dots. \quad (\text{A.161})$$

The vector \mathbf{k} is restricted to lie in the positive octant,

$$k_x > 0, \quad k_y > 0, \quad k_z > 0, \quad (\text{A.162})$$

on a grid of points with grid spacing

$$\Delta k_x = \Delta k_y = \Delta k_z = \frac{\pi}{L}. \quad (\text{A.163})$$

The number of available states dN between k and $k + dk$ is obtained by counting states within one octant of a spherical shell

$$dN = \frac{1}{8} \times \frac{4\pi k^2 dk}{(\pi/L)^3} = \frac{k^2 L^3}{2\pi^2} dk \quad (k = |\mathbf{k}|). \quad (\text{A.164})$$

The same result can be obtained by imposing periodic instead of hard-wall boundary conditions

$$dN = \frac{4\pi k^2 dk}{(2\pi/L)^3} = \frac{k^2 L^3}{2\pi^2} dk. \quad (\text{A.165})$$

For a non-relativistic particle of mass m , the number of states per unit energy is

$$\frac{dN}{dE} = \frac{dN}{dk} \frac{dk}{dE} = \frac{k^2 L^3}{2\pi^2} \frac{m}{\hbar^2 k} = \frac{L^3 m k}{2\pi^2 \hbar^2}. \quad (\text{A.166})$$

The density of states per unit volume is therefore proportional to \sqrt{E}

$$g(E) = \frac{mk}{2\pi^2 \hbar^2}, \quad k = \frac{\sqrt{2mE}}{\hbar}. \quad (\text{A.167})$$

For a (relativistic) *massless* particle such as a photon, with energy

$$E = \hbar\omega(\mathbf{k}) = \hbar c k, \quad (\text{A.168})$$

the conversion from k to E becomes

$$\frac{dN}{dE} = \frac{dN}{dk} \frac{dk}{dE} = \frac{k^2 L^3}{2\pi^2} \frac{1}{\hbar c} = \frac{L^3}{2\pi^2 (\hbar c)^3} E^2. \quad (\text{A.169})$$

The density of states per unit volume is therefore proportional to E^2

$$g(E) = \frac{E^2}{2\pi^2 (\hbar c)^3}. \quad (\text{A.170})$$

Additional factors may be needed to take spin degrees of freedom into account (an extra factor of two for photons, for example).

For particles whose direction of motion is restricted to lie within a solid angle $\Delta\Omega$, an extra factor $\Delta\Omega/4\pi$ is needed. For a non-relativistic particle of mass m , the density of states (per unit volume, per spin degree of freedom) within $\Delta\Omega$ is

$$g(E) = \frac{mk}{2\pi^2 \hbar^2} \times \frac{\Delta\Omega}{4\pi} = \frac{mk}{8\pi^3 \hbar^2} \Delta\Omega, \quad k = \frac{\sqrt{2mE}}{\hbar}. \quad (\text{A.171})$$

For a massless particle, we obtain instead

$$g(E) = \frac{E^2}{2\pi^2 (\hbar c)^3} \times \frac{\Delta\Omega}{4\pi} = \frac{E^2}{(2\pi \hbar c)^3} \Delta\Omega. \quad (\text{A.172})$$

A.6 Landau Levels from a Gauge Independent Approach

As we have seen, the appropriate Hamiltonian describing this system is

$$\hat{H} = \frac{1}{2m} \left[\hat{\mathbf{p}} - \frac{q}{c} \mathbf{A} \right]^2 = \frac{1}{2m} \hat{\mathbf{\Pi}}^2, \quad (\text{A.173})$$

where $\hat{\mathbf{\Pi}}$ is a vector operator, hence $\hat{\mathbf{\Pi}}^2 = \hat{\Pi}_x^2 + \hat{\Pi}_y^2 + \hat{\Pi}_z^2$. The components of $\hat{\mathbf{\Pi}}$ do *not* commute, e.g.

$$[\hat{\Pi}_x, \hat{\Pi}_y] = [\hat{p}_x - qA_x, \hat{p}_y - qA_y] = -q[A_x, \hat{p}_y] - q[\hat{p}_x, A_y] \neq 0. \quad (\text{A.174})$$

Using the identity $[f(\mathbf{r}), \hat{p}] = i\hbar\nabla f(\mathbf{r})$,

$$[A_x, \hat{p}_y] = i\hbar \frac{\partial A_x}{\partial y}, \quad [A_y, \hat{p}_x] = i\hbar \frac{\partial A_y}{\partial x} \quad (\text{A.175})$$

$$\implies [\hat{\Pi}_x, \hat{\Pi}_y] = -iq\hbar \left(\frac{\partial A_x}{\partial y} - \frac{\partial A_y}{\partial x} \right) = -iq\hbar B_z. \quad (\text{A.176})$$

Let's consider a uniform magnetic field $\mathbf{B} = (0, 0, B)$. Introduce non-Hermitian operators \hat{a} and \hat{a}^\dagger defined as

$$\hat{a} = \sqrt{\frac{1}{2q\hbar B}} (\hat{\Pi}_x + i\hat{\Pi}_y), \quad \hat{a}^\dagger = \sqrt{\frac{1}{2q\hbar B}} (\hat{\Pi}_x - i\hat{\Pi}_y). \quad (\text{A.177})$$

Their commutator is

$$[\hat{a}, \hat{a}^\dagger] = -\frac{i}{q\hbar B} [\hat{\Pi}_x, \hat{\Pi}_x] = 1. \quad (\text{A.178})$$

The operator product $\hat{a}^\dagger \hat{a}$ is

$$\begin{aligned} \hat{a}^\dagger \hat{a} &= \frac{1}{2q\hbar B} (\hat{\Pi}_x - i\hat{\Pi}_y)(\hat{\Pi}_x + i\hat{\Pi}_y) \\ &= \frac{1}{2q\hbar B} (\hat{\Pi}_x^2 + \hat{\Pi}_y^2 + i[\hat{\Pi}_x, \hat{\Pi}_y]) \\ &= \frac{1}{2q\hbar B} (\hat{\Pi}^2 - \hat{\Pi}_z^2) - \frac{1}{2}. \end{aligned} \quad (\text{A.179})$$

Hence, we can write the Hamiltonian in Eq.(A.173) as

$$\hat{H} = \hbar\omega_c \left(\hat{a}^\dagger \hat{a} + \frac{1}{2} \right) + \frac{1}{2m} \hat{\Pi}_z^2, \quad (\text{A.180})$$

where $\omega_c = |q|B/m$. The first term in this Hamiltonian corresponds to a quantum harmonic oscillator with frequency ω_c and energy eigenvalues

$$E_n = \hbar\omega_c \left(\nu + \frac{1}{2} \right), \quad n = 0, 1, 2, \dots \quad (\text{A.181})$$

Hence, the $\hat{\Pi}_x$ and $\hat{\Pi}_y$ terms in the Hamiltonian, transverse to the magnetic field \mathbf{B} , contribute quantised *Landau levels* with constant energy spacing

$$\Delta E = \hbar\omega_c = \frac{\hbar|q|B}{m} \quad (\text{A.182})$$

A.7 The Coulomb Gauge

The Hamiltonian describing a non-relativistic charged particle moving in an EM field is

$$\hat{H} = \frac{1}{2m} (\hat{\mathbf{p}} - q\mathbf{A}(\mathbf{r}, t))^2 + q\phi(\mathbf{r}, t). \quad (\text{A.183})$$

Expanding out the first term explicitly, Schrödinger's equation can be written as

$$\hat{H}\psi = \frac{1}{2m}(-i\hbar\nabla - q\mathbf{A}) \cdot (-i\hbar\nabla - q\mathbf{A})\psi + q\phi\psi = i\hbar\frac{\partial\psi}{\partial t}. \quad (\text{A.184})$$

Using the identity

$$\nabla \cdot (\mathbf{A}\psi) = \mathbf{A} \cdot (\nabla\psi) + (\nabla \cdot \mathbf{A})\psi, \quad (\text{A.185})$$

the Hamiltonian (A.184) can be written as

$$\hat{H}\psi = \frac{1}{2m}[-\hbar^2\nabla^2\psi + i\hbar q(2\mathbf{A} \cdot \nabla\psi + \psi\nabla \cdot \mathbf{A}) + q^2\mathbf{A}^2\psi] + q\phi\psi. \quad (\text{A.186})$$

It is often convenient to work in a particular gauge¹, a common choice being the Coulomb gauge

$$\boxed{\nabla \cdot \mathbf{A} = 0.} \quad (\text{A.187})$$

The Coulomb gauge must always exist: to see this, suppose we are working initially in a gauge with

$$\nabla \cdot \mathbf{A} \neq 0. \quad (\text{A.188})$$

Applying a gauge transformation

$$\mathbf{A} \rightarrow \mathbf{A}' = \mathbf{A} + \nabla f, \quad \nabla \cdot \mathbf{A} \rightarrow \nabla \cdot \mathbf{A}' = \nabla \cdot \mathbf{A} + \nabla^2 f \quad (\text{A.189})$$

with $f(\mathbf{r}, t)$ chosen to be a solution of $\nabla^2 f = -(\nabla \cdot \mathbf{A})$ then gives

$$\nabla \cdot \mathbf{A}' = 0. \quad (\text{A.190})$$

In the Coulomb gauge, the Hamiltonian of equation (A.186) simplifies as

$$\boxed{\hat{H} = \frac{1}{2m}[-\hbar^2\nabla^2 + 2i\hbar q\mathbf{A} \cdot \nabla + q^2\mathbf{A}^2] + q\phi.} \quad (\text{A.191})$$

A.7.1 The Symmetric Gauge

If the magnetic field \mathbf{B} is stationary and uniform, a convenient choice of Coulomb gauge is the *symmetric gauge*, defined by

$$\boxed{\mathbf{A}(\mathbf{r}) = -\frac{1}{2}\mathbf{r} \times \mathbf{B}.} \quad (\text{A.192})$$

To check this is valid, orient the z -axis along the \mathbf{B} field direction so that

$$\mathbf{B} = (0, 0, B_z), \quad \mathbf{A}(\mathbf{r}) = -\frac{1}{2}\mathbf{r} \times \mathbf{B} = \frac{1}{2}B_z(-y, x, 0). \quad (\text{A.193})$$

For B_z constant, we then find, as required

$$\nabla \times \mathbf{A} = \mathbf{B}, \quad \nabla \cdot \mathbf{A} = 0. \quad (\text{A.194})$$

There is nothing special about the z -direction, so equation (A.192) must be valid for any uniform magnetic field \mathbf{B} , oriented in any direction.

¹Observable quantities are independent of this choice.

In the symmetric gauge, and for \mathbf{B} directed along z , the operator $\mathbf{A} \cdot \nabla$ in equation (A.191) is

$$\mathbf{A} \cdot \nabla = \frac{1}{2} B_z \left(-y \frac{\partial}{\partial x} + x \frac{\partial}{\partial y} \right). \quad (\text{A.195})$$

Comparing with

$$\hat{L}_z = (\hat{\mathbf{r}} \times \hat{\mathbf{p}})_z = -i\hbar(\hat{\mathbf{r}} \times \nabla)_z = -i\hbar \left(x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x} \right), \quad (\text{A.196})$$

and using $B_z \hat{L}_z = \mathbf{B} \cdot \hat{\mathbf{L}}$ then gives

$$\mathbf{A} \cdot \nabla = \frac{i}{2\hbar} \mathbf{B} \cdot \hat{\mathbf{L}} \quad (\hat{\mathbf{L}} = \hat{\mathbf{r}} \times \hat{\mathbf{p}}). \quad (\text{A.197})$$

This must hold for *any* uniform \mathbf{B} , since the choice of z -axis is arbitrary.

In the symmetric gauge, the factor A^2 in equation (A.191) is

$$A^2 = \frac{1}{4}(\mathbf{r} \times \mathbf{B})^2 = \frac{1}{4} \left[r^2 B^2 - (\mathbf{r} \cdot \mathbf{B})^2 \right]. \quad (\text{A.198})$$

Substituting from equations (A.197) and (A.198), the Hamiltonian (A.191) for a uniform magnetic field \mathbf{B} , in the symmetric gauge, is

$$\hat{H} = \frac{\hat{\mathbf{p}}^2}{2m} - \frac{q}{2m}(\hat{\mathbf{L}} \cdot \mathbf{B}) + \frac{q^2}{8m} \left(B^2 r^2 - (\mathbf{B} \cdot \mathbf{r})^2 \right) + q\phi(\mathbf{r}).$$

(A.199)

For an electron confined within an atom, we can estimate

$$\langle \hat{L}_z \rangle \sim \hbar, \quad \langle r \rangle \sim a_0 \approx 0.5 \times 10^{-10} \text{ m}. \quad (\text{A.200})$$

Hence, for an atom in an external magnetic field \mathbf{B} , the ratio of quadratic and linear B -field terms in equation (A.199) is of order

$$\frac{e^2 B^2 a_0^2 / 8m_e}{eB\hbar / 2m_e} = \frac{eBa_0^2}{4\hbar} \approx 1.1 \times 10^{-6} \times (B/\text{T}). \quad (\text{A.201})$$

Hence, for atomic electrons ($q = -e$), we can often neglect the quadratic (B^2) terms in equation (A.199), leaving just the linear magnetic field term

$$\hat{H}_{\mathbf{B}} = \frac{e}{2m_e} \hat{\mathbf{L}} \cdot \mathbf{B}. \quad (\text{A.202})$$

For a hydrogen atom in an external magnetic field \mathbf{B} , for example, the Hamiltonian can be taken, to good approximation, to be

$$\hat{H} = \frac{\hat{\mathbf{p}}^2}{2m} - \frac{e^2}{4\pi\epsilon_0 r} + \frac{e}{2m_e}(\hat{\mathbf{L}} \cdot \mathbf{B}). \quad (\text{A.203})$$

APPENDIX B

Appendix: Equations

B.1 Wave Mechanics and the Schrödinger Equation 1

- Time-dependent Schrödinger equation (1.20)

$$i\hbar \frac{\partial \Psi}{\partial t} = \hat{H}\Psi(\mathbf{r}, t). \quad (\text{B.1})$$

B.2 Quantum Mechanics in One Dimension 2

- Hamiltonian (2.4)

$$\hat{H} = -\frac{\hbar^2 \nabla^2}{2m} + V(\mathbf{r}). \quad (\text{B.2})$$

- Particle flux (2.6)

$$\mathbf{j}(\mathbf{r}, t) = -\frac{i\hbar}{2m} [\Psi^*(\mathbf{r}, t) \nabla \Psi(\mathbf{r}, t) - \Psi(\mathbf{r}, t) \nabla \Psi^*(\mathbf{r}, t)]. \quad (\text{B.3})$$

- Wavefunction for propagating plane wave (2.7)

$$\Psi(x, t) = A e^{i(kx - \omega t)}. \quad (\text{B.4})$$

B.3 Operator Methods and Formalism in Quantum Mechanics 3

- Inner Product (3.2)

$$\langle \phi | \psi \rangle \equiv \int_{-\infty}^{\infty} dx \phi^*(x) \psi(x). \quad (\text{B.5})$$

- Operator Eigenstates (3.26)

$$\hat{A} |\psi\rangle = a |\psi\rangle. \quad (\text{B.6})$$

- Expectation Value of an Operator (3.28)

$$\langle A \rangle = \langle \psi | \hat{A} | \psi \rangle = \int_{-\infty}^{\infty} dx \phi^*(x) \hat{A} \psi(x). \quad (\text{B.7})$$

- Commutator (3.33)

$$[\hat{A}, \hat{B}] \equiv \hat{A}\hat{B} - \hat{B}\hat{A}. \quad (\text{B.8})$$

- Resolution of the Identity (3.41)

$$\hat{\mathbb{I}} = \sum_i |i\rangle\langle i|. \quad (\text{B.9})$$

- Probability of Measuring an Eigenvalue of an Operator (3.75)

$$\text{Prob}(\lambda_n) = |a_n|^2. \quad (\text{B.10})$$

- Time-Evolution Operator (3.89)

$$\hat{U} = e^{-i\hat{H}t/\hbar}. \quad (\text{B.11})$$

- Canonical Commutation Relations (3.94)

$$[\hat{x}_i, \hat{x}_j] = [\hat{p}_i, \hat{p}_j] = 0 \quad \text{and} \quad [\hat{x}_i, \hat{p}_j] = i\hbar\delta_{ij}. \quad (\text{B.12})$$

- Generalised Uncertainty Principle (3.113)

$$\Delta A \Delta B \geq \left| \frac{1}{2} \langle [\hat{A}, \hat{B}] \rangle \right|. \quad (\text{B.13})$$

- Heisenberg's Uncertainty Principle (3.119)

$$\Delta p \Delta x \geq \frac{\hbar}{2}. \quad (\text{B.14})$$

- Ehrenfest's Theorem (3.123)

$$\frac{d}{dt} \langle \hat{A} \rangle = \frac{i}{\hbar} \langle [\hat{H}, \hat{A}] \rangle. \quad (\text{B.15})$$

- The Heisenberg Equation (3.129)

$$i\hbar \frac{d\hat{B}}{dt} = [\hat{B}(t), \hat{H}]. \quad (\text{B.16})$$

- Quantum Harmonic Oscillator Hamiltonian (3.131)

$$\hat{H} = \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega^2\hat{x}^2. \quad (\text{B.17})$$

- Ladder Operators (3.132)

$$a = \sqrt{\frac{m\omega}{2\hbar}} \left(\hat{x} + i\frac{\hat{p}}{m\omega} \right), \quad a^\dagger = \sqrt{\frac{m\omega}{2\hbar}} \left(\hat{x} - i\frac{\hat{p}}{m\omega} \right). \quad (\text{B.18})$$

- Ladder Operator Commutator (3.134)

$$[a, a^\dagger] \equiv aa^\dagger - a^\dagger a = 1. \quad (\text{B.19})$$

- State $|n\rangle$ raised from ground state (3.142)

$$|n\rangle = \frac{1}{\sqrt{n!}} (a^\dagger)^n |0\rangle. \quad (\text{B.20})$$

- Ladder Operator action on Eigenstates (3.143)

$$a^\dagger |n\rangle = \sqrt{n+1} |n+1\rangle, \quad a |n\rangle = \sqrt{n} |n-1\rangle. \quad (\text{B.21})$$

B.4 Quantum Mechanics in More Than One Dimension 4

B.5 Spin 5

B.6 Motion in a Magnetic Field 6

B.7 Approximation Methods for Stationary States 7

B.8 Symmetry in Quantum Mechanics 9

B.9 Identical Particles 10

B.10 Atomic Structure 8

B.11 From Molecules to Solids 11

B.12 Time-Dependent Perturbation Theory 12

B.13 Scattering Theory 13

B.14 Radiative Transitions 14

B.15 Field Theory: From Phonons to Photons 15