ETH Zürich · Institute for Theoretical Physics

Quantum Mechanics I and II

Lecture script

Babis Anastasiou

Institute for Theoretical Physics, ETH Zurich
8093 Zurich, Switzerland
babis@phys.ethz.ch

Contents

I Quantum Mechanics I 1

Preface 2

1 Quantum behaviour 3
1.1 Atomic mechanics . . . 3
1.2 Basic two-slit experiments . . . 3
1.2.1 An experiment with bullets . . . 3
1.2.2 An experiment with waves . . . 4
1.2.3 An experiment with electrons . . . 5
1.3 Watching electrons . . . 5
1.4 Back to the bullet experiment. . . . 7
1.5 First principles of quantum mechanics . . . 7
1.6 The laws for combining amplitudes . . . 8
1.6.1 Revisiting the two-slit experiment . . . 9
1.6.2 A subatomic journey from Zurich to China? . . . 10

2 Quantum measurement and quantum states 11
2.1 Stern–Gerlach experiment . . . 11
2.2 The Stern–Gerlach filter . . . 12
2.2.1 Experimental setup . . . 12
2.2.2 Successive aligned Stern–Gerlach filters . . . 13
2.2.3 Successive rotated Stern–Gerlach filters . . . 14
2.2.4 Three filters . . . 15
2.3 Base states . . . 16
2.4 States and vector spaces . . . 17

3 Quantum mechanics and linear algebra 20
3.1 Ket-space . . . 20
3.2 Operators . . . 20
3.3 Dual space . . . 21
3.4 Inner product . . . 21
3.5 Properties of operators and their duals . . . 21
3.6 Hermitian operators . . . 23
3.7 Matrix representation of states and operators . . . 25
3.8 Compatible and incompatible observables . . . 27
3.9 Expectation value and uncertainty of Hermitian operators . . . 29
3.10 The uncertainty principle. . . . 30
3.11 Change of basis . . . 32
3.11.1 Transformation matrix . . . 33
3.11.2 Trace of operators . . . 34
3.12 Eigenstates and eigenvalues . . . 35
3.13 Unitarily equivalent observables . . . 35

4 Time evolution 37
4.1 Time evolution in classical mechanics . . . 37

4.4 Schrödinger equation . . . 45
4.5 Eigenstates of the Hamiltonian operator . . . 46

5 Two-state systems 48
5.1 Spin precession . . . 48
5.2 A generic two-state system . . . 50
5.3 The ammonia molecule . . . 52
5.4 The ammonia molecule inside an electric field . . . 55
5.5 Ammonia in an oscillating electric field . . . 56
5.5.1 Transitions at resonance . . . 58
5.5.2 Microwave amplification with stimulated emission of radiation . . . 59
5.5.3 Transitions off resonance . . . 59

6 Position and momentum 61
6.1 Continuous spectra . . . 61
6.2 Position operator and eigenstates . . . 63
6.3 Momentum operator and eigenstates . . . 64
6.4 Translation operator . . . 65
6.5 Representation of momentum in position space . . . 67
6.6 Eigenstates of the momentum operator . . . 68
6.6.1 Position and momentum wave functions . . . 68

7 The harmonic oscillator 70
7.1 Quantisation . . . 70
7.2 Time evolution . . . 74
7.3 Coherent states . . . 76

8 Schrödinger's wave equation 80
8.1 Probability density and probability current . . . 81
8.2 Quantised energy levels . . . 82
8.3 Semiclassical approximation . . . 83
8.4 Linear potential . . . 87

9 Particle in a constant electromagnetic field 89
9.1 Wave function for a particle in an electromagnetic field . . . 90
9.2 Aharonov–Bohm effect . . . 91

10 Symmetries in quantum mechanics 94
10.1 Continuous symmetry transformations . . . 96
10.2 Lie algebra and generators . . . 97
10.3 Symmetry and degeneracy . . . 100
10.4 Rotations and translations . . . 100
10.4.1 Generators . . . 101
10.4.2 Lie algebra . . . 103

11 Representations of angular momentum 104
11.1 Eigenstates of angular momentum . . . 104
11.2 Spin-$\frac{1}{2}$ representation of angular momentum . . . 106
11.3 Orbital angular momentum . . . 109
11.3.1 Spherical coordinates . . . 109

11.3.2 Orbital angular momentum operators and the Laplacian . . . 110
11.3.3 Spherical harmonics . . . 111
11.3.4 Space representations of rotations . . . 115
11.4 Potentials with spherical symmetry . . . 116
11.5 The hydrogen atom . . . 118
11.5.1 $SO(4)$ symmetry of the hydrogen atom . . . 121

12 Addition of angular momenta **122**
12.1 Addition of angular momenta . . . 122
12.2 Application: the hydrogen atom . . . 126
12.3 Wigner–Eckart theorem . . . 127

13 Discrete symmetries **130**
13.1 Parity . . . 130
13.2 Time reversal . . . 132
13.2.1 Antilinear and antiunitary operators . . . 134
13.2.2 Position, momentum and wave functions . . . 134
13.2.3 Angular momentum . . . 135
13.2.4 Time reversal of angular momentum eigenstates . . . 136
13.2.5 Time-reversal invariance and its consequences . . . 137

II Quantum Mechanics II **139**

Preface to Part II **140**

14 Perturbation Theory **141**
14.1 Perturbation Theory for Time-Independent Potentials with No Degeneracy . . . 141
14.1.1 Zeroth- and First-Order Perturbative Corrections . . . 142
14.1.2 Second-Order Perturbative Correction . . . 145
14.1.3 Perturbed Harmonic Oscillator . . . 146
14.1.4 The Variational Method (An Alternative) . . . 148
14.2 Perturbation Theory for Degenerate Energy Levels . . . 150
14.2.1 Summary of Degenerate Perturbation Theory . . . 155
14.2.2 Example of Degenerate Perturbation Theory . . . 156
14.2.3 Example: Stark Effect for the $n = 2$ Energy Levels . . . 159
14.3 Applications of Time-Independent Perturbation Theory . . . 160
14.3.1 Example 1: The van der Waals Force . . . 161
14.3.2 Example 2: Relativistic Kinetic Energy . . . 164
14.4 Time Evolution with Time-Dependent Hamiltonians . . . 167
14.4.1 General Formalism . . . 167
14.4.2 Time Evolution for Time-Dependent Hamiltonians . . . 168
14.4.3 Time Ordering . . . 169
14.5 Perturbation Theory for Time-Dependent Potentials . . . 171
14.5.1 The Interaction Picture . . . 172
14.5.2 Transition Amplitudes . . . 173
14.5.3 Sudden Constant Perturbation . . . 174
14.5.4 Fermi's Golden Rule . . . 176
14.5.5 Exercises: Golden Rule for a Lorentzian Density of States . . . 179
14.5.6 Harmonic Perturbation . . . 182
14.5.7 Ionisation . . . 183

15 Theory of Elastic Scattering 188
15.1 An Introduction to Elastic Scattering Experiments . . . 188
15.2 The Lippmann–Schwinger Equation . . . 189
15.3 Green's Functions . . . 192
15.4 Large-Distance Expansion . . . 194
15.5 Cross-Sections . . . 195
15.6 Born Approximation . . . 198
15.6.1 Example: The Yukawa Potential . . . 199
15.7 Partial Wave Expansion . . . 200
15.7.1 The Optical Theorem . . . 204
15.7.2 Application: Partial-Wave Expansion of a Plane Wave . . . 205
15.8 Born Approximation and Scattering Phases . . . 206

16 Radiation and Matter Interaction 208
16.1 Maxwell Equations and Electromagnetic Radiation . . . 208
16.2 The Classical Electromagnetic Interaction with Matter . . . 211
16.2.1 Photon Absorption . . . 214
16.3 Matrix Elements in the Dipole Approximation . . . 215
16.4 Quantisation of the Radiation Field . . . 218
16.4.1 Canonical Quantisation . . . 218
16.4.2 Field Operators . . . 220
16.4.3 Photon Emission and Absorption . . . 221
16.4.4 Spontaneous Emission and the Lifetime of the $2p$ State of Hydrogen . . . 222
16.4.5 Einstein Coefficients and Planck's Law . . . 223

17 Identical Particles 225
17.1 Introduction . . . 225
17.2 Permutation Operator . . . 226
17.3 Symmetrisation Postulate . . . 228
17.3.1 Pauli Exclusion Principle . . . 229
17.4 Two-Electron Systems . . . 229
17.5 Scattering of Identical Particles . . . 231
17.6 System of N Fermions . . . 232
17.6.1 Free Particles in a Box . . . 233

18 Multi-Electron Atoms 236
18.1 The Helium Atom . . . 236
18.1.1 Ground State . . . 237
18.1.2 Higher Energy Levels . . . 239
18.2 Atoms with More than Two Electrons . . . 240
18.2.1 The Hartree Approximation . . . 240
18.2.2 The Hartree–Fock Approximation . . . 243

19 Quantised Fields 244
19.1 Bosons from Quantised Fields . . . 244
19.2 Fermions from Quantised Fields . . . 251
19.3 Field Quantisation: Number of Particles and Total Energy . . . 253
19.3.1 Particle Density Operator . . . 253
19.3.2 Energy and Energy Density (Hamiltonian) Operators . . . 254
19.4 Non-Relativistic Quantum Field Theory of Interacting Fields . . . 255

19.5 Field Hamiltonian in Momentum Space . . . 259
19.5.1 Creation Operators as Fourier Transforms of the Field . . . 260
19.6 Perturbation (Field) Theory . . . 261
19.6.1 Leading Order in Perturbation Theory . . . 262
19.6.2 Next-to-Leading Order . . . 263
19.6.3 Transition Rate . . . 265

A Spherical Potentials and Spherical Bessel Functions 267
A.1 Spherical Hankel Functions, Bound States and Phase Shifts . . . 269

B Explicit Calculation of the Green's Function for Scattering 272

C An Alternative Derivation of the Lippmann–Schwinger Equation 274

D Functional Derivatives 276

PART I

Quantum Mechanics I

Preface

This script is based on Refs. [1–3]. These books are excellent, and students are encouraged to study them as a first priority. The purpose of the present notes, which contain some personalised material, is to support the author's lecture presentations. Students are of course welcome to read the script if they find it useful. However, the script should not replace the study of the recommended and other literature.

1

Quantum behaviour

1.1 Atomic mechanics

For a long time, light was thought to behave like waves and electrons were thought to behave like particles. There are phenomena which violate these main "rules": light and electrons behave neither like particles nor like waves. However, light, electrons, protons and all subatomic particles do behave in the same way. This common subatomic behaviour is very different from our intuition of large objects. We will demonstrate the difference between quantum and classical behaviour with a "two-slit" thought experiment, first with macroscopic bullets, second with waves and third with subatomic particles such as electrons.

1.2 Basic two-slit experiments

1.2.1 An experiment with bullets

We have a gun firing very hard, indestructible bullets towards a wall. The wall has two slits (holes) which are close to each other and big enough for the bullets to pass through (Fig. 1.1).

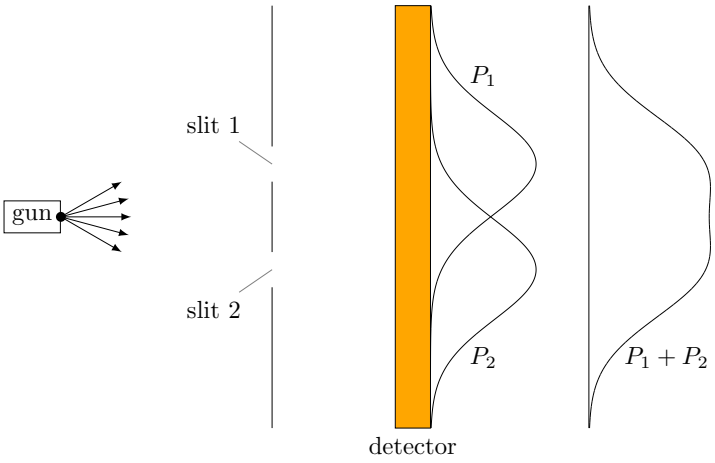


Figure 1.1. A two-slit experiment with bullets.

Behind the wall there is a "detector", a material which can stop the bullets. We can take a look at it afterwards and see how the bullets are distributed in space after they have gone

through the holes. We perform the experiment in three stages:

- First, we leave "Hole 1" open and cover up the second hole. We find that the bullets are distributed according to P_1, a Gaussian centred on Hole 1.
- Then, we leave "Hole 2" open while we cover up "Hole 1". We find a similar Gaussian P_2 centred on Hole 2.
- Finally, we leave both holes open. We find the distribution

$$P_{12} = P_1 + P_2 \,. \tag{1.1}$$

The probability that a bullet goes through the holes is the sum of the probabilities that it passes either through hole 1 or through hole 2.

1.2.2 An experiment with waves

We now perform a different experiment, in which we create a spherical wave at some distance in front of the wall with the two holes (Fig. 1.2). The detector measures the

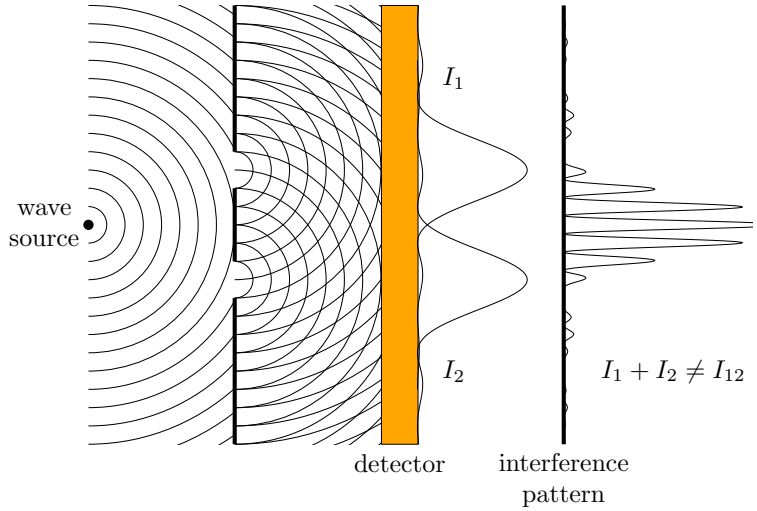


Figure 1.2. A two-slit experiment with waves.

intensity of the wave arriving at it, which is proportional to the square of the amplitude:

$$I \sim |A|^2 \,.$$

We perform the experiment in three stages:

- First, we leave "Hole 1" open and cover up the second hole. We find that the intensity I_1 is a function centred on Hole 1.
- Then, we leave "Hole 2" open while we cover up "Hole 1". We find a similar intensity function I_2 centred on Hole 2.
- Finally, we leave both holes open. We find a distribution $I_{12} \neq I_1 + I_2$.

There is a simple explanation of why the intensity I_{12} is not equal to the sum of the intensities measured when one of the two holes is covered. The amplitude of the wave is given by the sum of the amplitudes,

$$A_{12} = A_1 + A_2 ,$$

where A_1 and A_2 can be positive or negative. The total intensity,

$$\begin{aligned} I_{12} &\sim |A_{12}|^2 \\ &= I_1 + I_2 + 2\sqrt{|I_1|\,|I_2|}\cos\delta, \end{aligned} \tag{1.2}$$

includes an interference term, giving the characteristic interference pattern of Fig. 1.2.
Remark. The amplitudes are best taken to be complex, $A_{1,2} = |A_{1,2}|\, e^{i\varphi_{1,2}}$, and $\delta = \varphi_1 - \varphi_2$ is their phase difference (set by the path-length difference from the two holes to the detector point); "positive or negative" is the special case $\delta = 0, \pi$. Cf. Feynman Lectures III, Sec. 1-4.

1.2.3 An experiment with electrons

We now shoot electrons at the wall. What happens is pictured in Fig. 1.3.

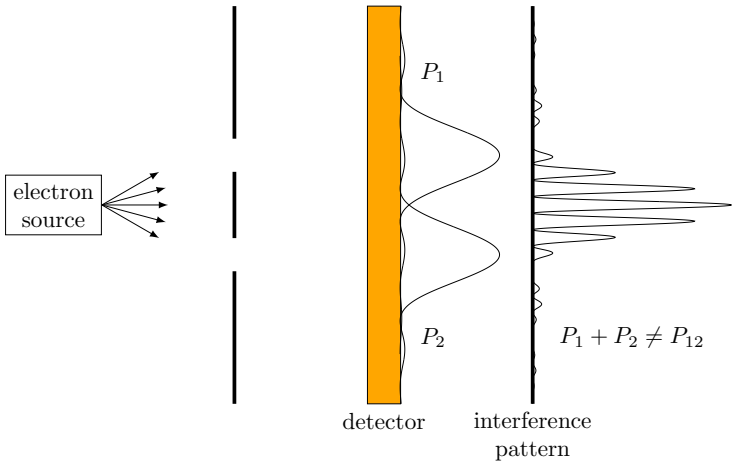


Figure 1.3. A two-slit experiment with electrons, showing wave behaviour.

This is a surprising result. If an electron went through either hole 1 or hole 2, we would have found $P_{12} = P_1 + P_2$; instead, $P_{12} \neq P_1 + P_2$. Are electrons waves? No. We always detect one "full" electron at a time at the detector. For example, we can reconstruct the full energy of an electron as it dissipates it in the detector, or, if the detector is inside a magnetic field, the electron track has the curvature of a particle with one unit of charge.

So, the electron arrives at the detector as a particle, and the probability of arrival is distributed like the intensity of a wave.

1.3 Watching electrons

Let us put a light source behind the two holes. Deflected light from the vicinity of hole 1 or hole 2 can tell us where the electron passed through. What happens? Indeed, we can see a flash coming from the hole every time that an electron passes through. However, we now find the probability pattern of Fig. 1.4. The interference pattern is destroyed when we can verify experimentally that the electrons go through either hole 1 or hole 2. If we switch off the light, the interference leading to $P_{12} \neq P_1 + P_2$ is restored.

Let us now try to understand how the light affects our measurement. Before going ahead with modifying the light characteristics, we equip our detector with a sound system. Each time an electron hits the detector we also hear a click.

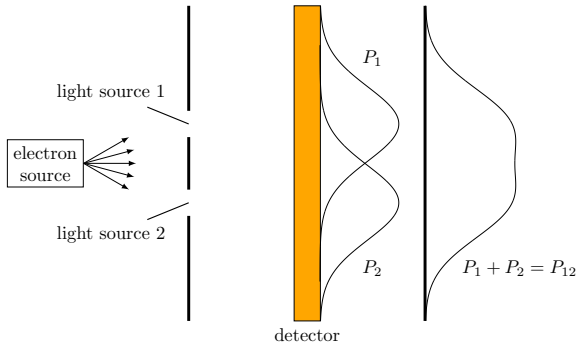


Figure 1.4. A two-slit experiment with electrons. In this experiment, we observe which hole the electrons pass through.

What happens with a dimmer light? We observe two types of events:

i) events where a flash comes from hole 1 or hole 2 and, immediately after, we hear a click coming from the detector;

ii) events where we only hear a click from the detector without seeing any flash.

We plot the probability distribution for each of the above categories of events. For the events where we have seen a flash, we find a distribution as in Fig. 1.4 with no interference. For the events where we have not seen a flash but we have heard the click, we get a probability distribution as in Fig. 1.3 with an interference pattern. The interference pattern is destroyed if we are able to observe the hole through which the electrons passed.

The intensity of the light is not its only physical characteristic. When we dim it, we reduce the amplitude of the electromagnetic wave. This means that there is a smaller density of "light particles" (photons) and therefore some electrons can pass through the holes without colliding with any of them. It is then that we only hear a click at the moment the electron arrives at the detector without having seen a flash from a hole before. But we can also tune the wavelength of the light. There is a simple relation between the momentum of the photons and the wavelength,

$$p = \frac{h}{\lambda},$$

where h is Planck's constant. Photons are less energetic and the light is "gentler" if we increase the wavelength, using infrared light or even radio waves.

So, let us gradually increase the wavelength without changing the intensity. We find that at the beginning there is no change in the type of probability distribution for the electrons arriving at the detector. As long as the wavelength is shorter than the distance between the two holes, we can tell whether a flash came from one or the other hole, and we find no interference pattern:

$$P_{12}(\lambda) = P_1(\lambda) + P_2(\lambda), \qquad \lambda \sim \text{small}.$$

If we increase the wavelength to a size comparable to the distance between the two holes, we find something dramatic. The flash becomes fuzzy and it is not possible to tell anymore whether it comes from the first or the second hole. We then find that the interference is restored:

$$P_{12}(\lambda) \neq P_1(\lambda) + P_2(\lambda), \qquad \lambda \sim \text{hole distance or larger}.$$

In conclusion, there is no configuration of our apparatus for which we can determine which hole the electrons went through without destroying the interference pattern.

1.4 Back to the bullet experiment

If the laws of quantum mechanics are universal, holding for light particles, electrons, protons and macroscopic objects, why then did we not observe an interference pattern in the experiment with the bullets? The reason is that the wavelength of a large object is

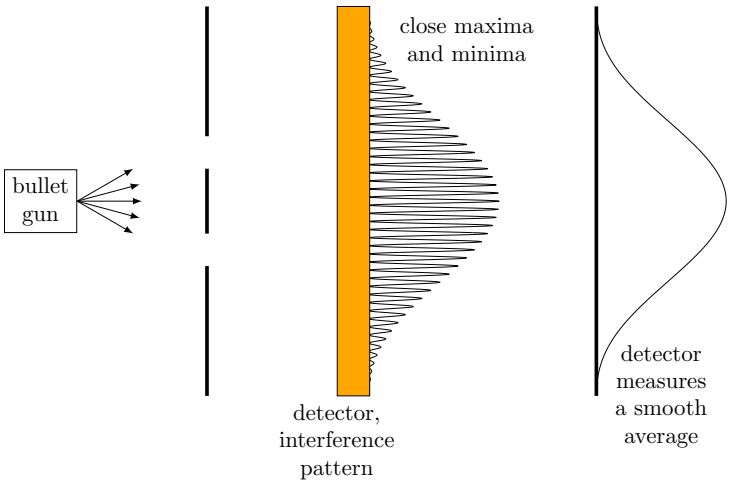


Figure 1.5. A two-slit experiment with bullets: the interference pattern is too fine to be resolved.

very short. The minima and maxima of the interference pattern are very dense and cannot be resolved by our detectors. Instead, we measure a smooth average over a distance of several wavelengths, as in Fig. 1.5.

1.5 First principles of quantum mechanics

Let us define an "ideal" experiment as one in which all initial and final conditions are completely specified. An event is a set of such initial and final conditions. (For example, a bullet leaves the gun, arrives at the detector and nothing else happens.)

i) The probability of an event in an ideal experiment is given by the square of the absolute value of a complex number, which is called the probability amplitude,

$$P = \left|\phi\right|^2, \quad \begin{cases} \phi \equiv \text{probability amplitude} \\ P \equiv \text{probability} \end{cases} \tag{1.3}$$

ii) If an event can occur in several ways, there is interference:

$$\begin{aligned} \phi &= \phi_1 + \phi_2 \\ P &= \left|\phi_1 + \phi_2\right|^2 . \end{aligned} \tag{1.4}$$

iii) If an experiment can determine whether one or another alternative happened, the interference is destroyed:

$$P = P_1 + P_2. \tag{1.5}$$

Note that we compute probabilities. We cannot know what an electron does at any given instant.

1.6 The laws for combining amplitudes

This is a good point to formalise and develop further the principles of Section 1.5. We will denote the probability amplitude of an event using the "bra-ket" notation of Dirac:

$$\langle \text{final condition} | \text{initial condition} \rangle .$$

For example, we can write the amplitude for an electron leaving the source s and arriving at the point x of the detector as

$$\langle \text{arrives at } x | \text{leaves from } s \rangle$$

or, in short,

$$\langle x|s\rangle ,$$

and the corresponding probability is

$$P = |\langle x|s\rangle|^2 .$$

According to our second principle, if an event can occur in more than one way, then we must add the amplitudes for all alternatives and we have interference. For example, in our two-slit experiment, where the electron could pass through either of the slits, we write

$$\langle x|s\rangle = \langle x|s\rangle_{\substack{\text{through}\\ \text{slit 1}}} + \langle x|s\rangle_{\substack{\text{through}\\ \text{slit 2}}} . \tag{1.6}$$

We now introduce a new rule for amplitudes (which leads to the third principle of the previous section). If an event can be expressed as a sequence of other events, then the amplitude factorises. For example, we can write

$$\left\langle \begin{matrix}\text{arrives}\\ \text{at } x\end{matrix} \middle| \begin{matrix}\text{leaves}\\ \text{source } s\end{matrix} \right\rangle_{\substack{\text{through}\\ \text{slit 1}}} = \left\langle \begin{matrix}\text{arrives}\\ \text{at } x\end{matrix} \middle| \begin{matrix}\text{leaves}\\ \text{slit 1}\end{matrix} \right\rangle \left\langle \begin{matrix}\text{arrives}\\ \text{at slit 1}\end{matrix} \middle| \begin{matrix}\text{leaves}\\ \text{source } s\end{matrix} \right\rangle$$

or, in short,

$$\langle x|s\rangle_{\substack{\text{through}\\ \text{slit 1}}} = \langle x|1\rangle \langle 1|s\rangle .$$

Then,

$$\langle x|s\rangle = \langle x|1\rangle \langle 1|s\rangle + \langle x|2\rangle \langle 2|s\rangle .$$

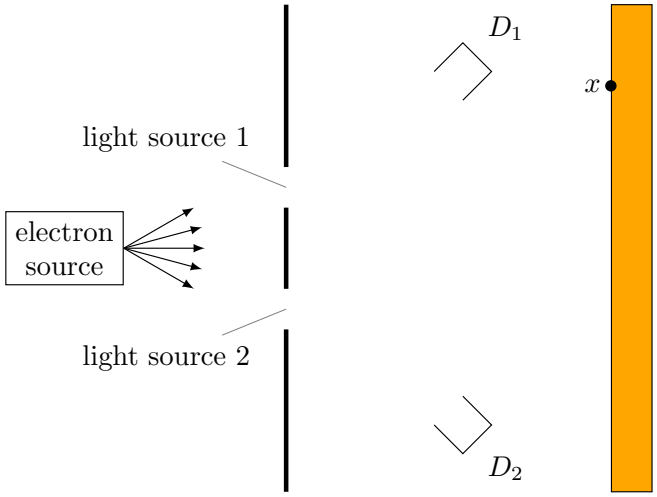


Figure 1.6. A two-slit experiment with electrons, with photon detectors used in an attempt to see which slit they go through.

1.6.1 Revisiting the two-slit experiment

Let us now revisit the two-slit experiment in which we fire electrons at the slits and use a light source behind them in an attempt to see which slit is chosen by an electron before it arrives at the detector. For the purpose of "seeing" the flashes, we place two photon detectors, D_1 to the left of the left hole and D_2 to the right of the right hole (Fig. 1.6). Let us focus on the events seen by the detector D_1. The amplitude is

$$\begin{aligned}\mathcal{M}_1 &= \left\langle \begin{matrix}\text{electron arrives at } x \\ \text{photon strikes } D_1\end{matrix} \middle| \begin{matrix}\text{electron leaves} \\ \text{source } s\end{matrix} \right\rangle \\ &= \langle x|1\rangle\, A_{1\to 1}(\lambda)\, \langle 1|s\rangle + \langle x|2\rangle\, A_{2\to 1}(\lambda)\, \langle 2|s\rangle\,.\end{aligned} \tag{1.7}$$

Here $A_{1\to 1}(\lambda)$ is the amplitude that a photon strikes an electron in hole 1 and is deflected into the detector D_1. Similarly, $A_{2\to 1}(\lambda)$ is the amplitude that a photon strikes an electron in hole 2 and is deflected into the detector D_1. The detector D_1 is placed closer to hole 1, and for small wavelengths $A_{2\to 1}$ is much smaller than $A_{1\to 1}$. The interference term in the probability $P_1 = |\mathcal{M}_1|^2$ vanishes for $A_{2\to 1} = 0$. Instead, for a large wavelength λ the photon cannot be localised precisely enough and $A_{2\to 1} \sim A_{1\to 1}$; then, there is interference.

Similarly, we can write an expression for the amplitude that the photon scattered from the electron strikes the detector D_2:

$$\begin{aligned}\mathcal{M}_2 &= \left\langle \begin{matrix}\text{electron arrives at } x \\ \text{photon strikes } D_2\end{matrix} \middle| \begin{matrix}\text{electron leaves} \\ \text{source } s\end{matrix} \right\rangle \\ &= \langle x|1\rangle\, A_{1\to 2}(\lambda)\, \langle 1|s\rangle + \langle x|2\rangle\, A_{2\to 2}(\lambda)\, \langle 2|s\rangle\,.\end{aligned} \tag{1.8}$$

What is then the probability for the photon to strike either D_1 or D_2? Notice that we now ask for the combined probability of two different events with different final conditions. We should not combine the two amplitudes into a common amplitude: $\mathcal{M} \neq \mathcal{M}_1 + \mathcal{M}_2$. Our rule of combining amplitudes applies to alternative ways of a single event (where the initial and final conditions are the same for all alternatives). If the events differ, we obtain

their combined probability by adding up the probabilities of the separate events. In our specific case,

$$\begin{aligned} P\left(\text{light in } D_1 \text{ or } D_2\right) &= P\left(\text{light in } D_1\right) + P\left(\text{light in } D_2\right) \\ &= |\mathcal{M}_1|^2 + |\mathcal{M}_2|^2 \,. \end{aligned}$$

1.6.2 A subatomic journey from Zurich to China?

Our product rule for amplitudes can puzzle a physical intuition trained on macroscopic physics. Think of a particle at a position $\mathbf{x}_1$ which transitions to a position $\mathbf{x}_2$. The corresponding amplitude is

$$\langle \mathbf{x}_2 | \mathbf{x}_1 \rangle \,.$$

Let $\mathbf{y}$ be an intermediate position in the journey of the particle. According to our rules for combining amplitudes we must have:

$$\langle \mathbf{x}_2 | \mathbf{x}_1 \rangle_{\text{through } \mathbf{y}} = \langle \mathbf{x}_2 | \mathbf{y} \rangle \langle \mathbf{y} | \mathbf{x}_1 \rangle \,.$$

However, if we have no knowledge of which point exactly the particle passed through, we must sum over all positions:

$$\langle \mathbf{x}_2 | \mathbf{x}_1 \rangle = \sum_{\mathbf{y}} \langle \mathbf{x}_2 | \mathbf{y} \rangle \langle \mathbf{y} | \mathbf{x}_1 \rangle \,.$$

In the sum, we are supposed to include all positions $\mathbf{y}$ no matter how far they reside from $\mathbf{x}_1, \mathbf{x}_2$ or how improbable they appear to us from our knowledge of classical physics.
Remark. If we put up a wall with a single open hole at $\mathbf{y}$, this *fixes* the intermediate point, and then only the product formula holds. The sum applies when nothing selects $\mathbf{y}$; $\mathbf{y}$ then runs over all points of an intermediate surface (or over all of space at an intermediate time), and for a continuum the sum is an integral. This is the starting point of Feynman's path integral (Feynman Lectures III, Sec. 3-1).

It turns out that amplitudes for classically improbable transitions are suppressed. Up to some normalisation, the transition amplitude of a free particle (no forces are exerted on it) of momentum $\mathbf{p}$ from a position $\mathbf{x}$ to a position $\mathbf{y}$ is

$$\mathcal{M}_{\mathbf{x}\to\mathbf{y}} \sim \frac{e^{-\frac{i}{\hbar}\mathbf{p}\cdot(\mathbf{x}-\mathbf{y})}}{|\mathbf{x}-\mathbf{y}|}, \tag{1.9}$$

where $\mathbf{p}$ is the classical momentum of the particle and $\hbar = \frac{h}{2\pi}$ is the reduced Planck constant. We note that the amplitude is suppressed at large distances and that the motion along the direction of the momentum produces a maximum phase in the exponential.
Remark. More precisely, the amplitude is the outgoing spherical wave $e^{ipr_{12}/\hbar}/r_{12}$ with $p = |\mathbf{p}|$ fixed by the energy and $r_{12} = |\mathbf{x}-\mathbf{y}|$ (Feynman Lectures III, Eq. (3.7)); combining the plane-wave phase $\mathbf{p}\cdot(\mathbf{y}-\mathbf{x})$ with the $1/r$ factor, as above, is a heuristic. Also, the $1/r$ decay is only a mild suppression: classically improbable paths are suppressed mainly because neighbouring paths have rapidly varying phases and cancel (stationary phase), as made precise in the path-integral formulation.

Probability amplitudes depend in general on both space and time. The time evolution of the amplitudes is determined by the Schrödinger equation, which we shall meet later.

2

Quantum measurement and quantum states

In this chapter we shall explore further the basic rules of probability amplitudes and introduce quantum states.

2.1 Stern–Gerlach experiment

Let us think of a magnetic dipole (a current loop) localised in a small volume, with a magnetic moment $\boldsymbol{\mu}$. The potential energy when the current is inside a magnetic field is

$$U = -\boldsymbol{\mu} \cdot \mathbf{B} \,. \tag{2.1}$$

Inside an inhomogeneous magnetic field a force is exerted on the magnetic dipole, driving it towards lower energy. The force is

$$\mathbf{F} = -\boldsymbol{\nabla} U = \boldsymbol{\nabla} \left(\boldsymbol{\mu} \cdot \mathbf{B} \right) \,. \tag{2.2}$$

For a magnetic field which varies, for example, in the z-direction, there is a force exerted on the magnetic dipole in the same direction,

$$F_z = \mu_z \frac{\partial B_z}{\partial z}. \tag{2.3}$$

Remark. This is an approximation: $\boldsymbol{\nabla} \cdot \mathbf{B} = 0$ forbids a field varying only along z, so there are also terms $\mu_x \partial B_x / \partial z$ etc.; they average to zero because the moment precesses rapidly about the strong field B_z (Griffiths, *Introduction to QM*, Sec. 4.4.2; Sakurai–Napolitano, Sec. 1.1).

In classical physics, we can attempt to describe atoms as dipoles, i.e. currents which are localised in a tiny space. We now perform a conceptually very simple experiment, the Stern–Gerlach experiment. We produce a beam of atoms and direct it towards a magnet with an inhomogeneous magnetic field. In classical physics we expect μ_z to be a continuous variable. The effect of the magnetic field gradient should then be to spread out the beam along the z-direction, as in Fig. 2.1. This is not what really happens. Instead, we find that the beam of atoms is split into separate beams, as in Fig. 2.2. The number of beams into which the original beam is split depends on the kind of atom, through a property called "spin".[1]. Atoms with spin-$\frac{1}{2}$ split into two beams, atoms with spin-1 split into three beams and so on; in general, atoms with spin-j split into $2j + 1$ beams.

[1] It is not yet time to describe spin. We will only say for now that it is an intrinsic angular momentum of subatomic particles which can be detected even when these particles are motionless.

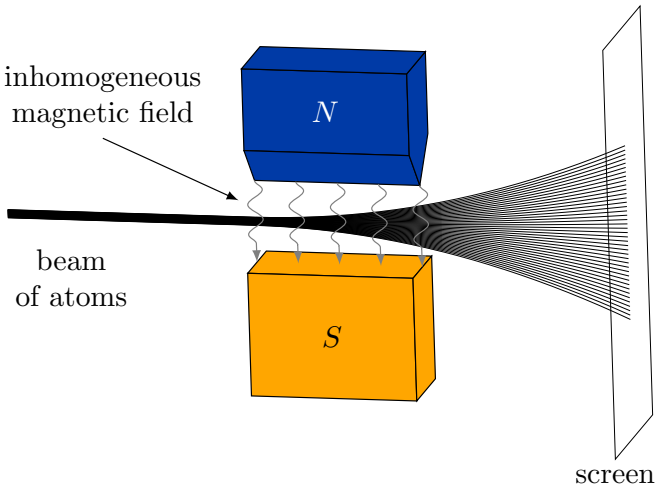


Figure 2.1. Stern–Gerlach experiment: expectation from classical electrodynamics, with atoms described as magnetic dipoles. (*This is not what happens in reality!*)

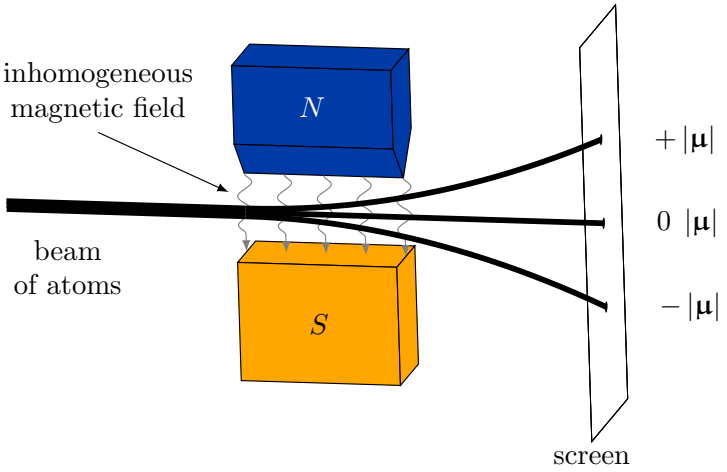


Figure 2.2. Stern–Gerlach experiment: the beam splits into a number of separated beams. The magnetic moments of the atoms assume discrete values.

2.2 The Stern–Gerlach filter

2.2.1 Experimental setup

In the following, we shall take a beam of atoms with spin-1 and examine experimentally the three separated beams which emerge from a Stern–Gerlach apparatus. We construct a Stern–Gerlach filter (Fig. 2.3). This filter takes a beam of atoms and passes it first through a magnetic field with a gradient in a certain direction, as in the Stern–Gerlach experiment, splitting the beam into three. The split beams are channelled through gates, each of which we can open or close at will. Afterwards, magnets reunite the beams which pass through the gates and let them come out from an exit point as a single beam.

The filter serves the purpose of selecting atoms of a certain state out of the Stern–Gerlach experiment. We can produce three types of "pure" beams with our Stern–Gerlach filter:

i) atom beams in the state $|\hat{z}, +\rangle$, where we close the lower two gates and allow only the

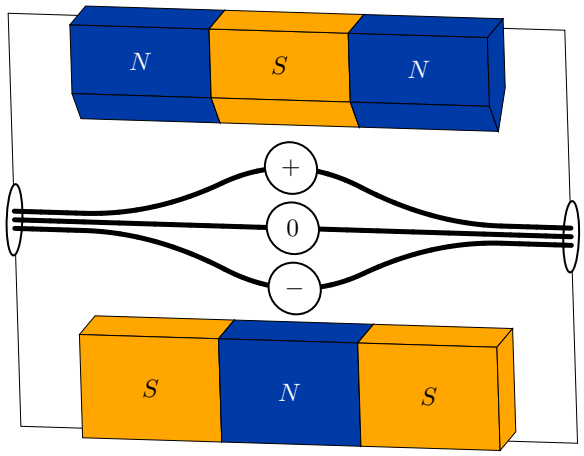


Figure 2.3. Stern–Gerlach filter for spin-1 atoms with gates +, 0 and −.

"upper" beam to go through;

ii) atom beams in the state $|\hat{z}, 0\rangle$, where we block the upper and lower beams and let the undeflected central beam go through;

iii) atom beams in the state $|\hat{z}, -\rangle$, where we close the upper two gates and allow only the "lower" beam to go through.

The unit vector $\hat{z}$ denotes the orientation of the magnetic field gradient in the Stern–Gerlach apparatus.

2.2.2 Successive aligned Stern–Gerlach filters

We shall now perform experiments passing atom beams through two successive Stern–Gerlach filters (Filter 1 and Filter 2), both having the same magnetic field gradient.

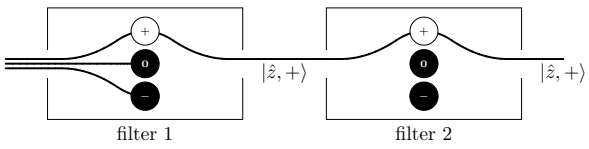


Figure 2.4. Two aligned Stern–Gerlach filters for spin-1 atoms.

In a first experiment (Fig. 2.4), we block the middle and lower gates of both filters. Filter 1 gives a beam of atoms in the state $|\hat{z}, +\rangle$. We observe that the *full* beam passes through the second filter, which is also designed to select the same $|\hat{z}, +\rangle$ state. We then conclude that the probability amplitude for the transition from a state $|\hat{z}, +\rangle$ to the same state $|\hat{z}, +\rangle$ is unity,[2]

$$\langle \hat{z}, +|\hat{z}, +\rangle = 1. \tag{2.4}$$

In a second experiment (Fig. 2.5), we close the middle and lower gates of Filter 1 and the upper and lower gates of Filter 2. Filter 1 selects atoms in the state $|\hat{z}, +\rangle$ and Filter 2 selects atoms in the state $|\hat{z}, 0\rangle$. We observe that no beam passes through the second filter.

[2] Up to a phase $\exp(i\alpha)$, which drops out when computing the modulus squared of the amplitude.

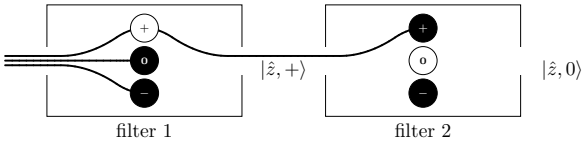


Figure 2.5. Two aligned Stern–Gerlach filters for spin-1 atoms.

Therefore, the probability amplitude for the transition from the state $|\hat{z},+\rangle$ to the state $|\hat{z},0\rangle$ vanishes,

$$\langle \hat{z},0|\hat{z},+\rangle = 0. \tag{2.5}$$

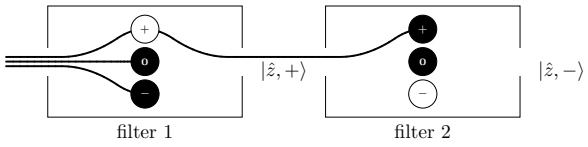


Figure 2.6. Two aligned Stern–Gerlach filters for spin-1 atoms.

The same result holds for transitions from $|\hat{z},+\rangle$ to $|\hat{z},-\rangle$ (Fig. 2.6),

$$\langle \hat{z},-|\hat{z},+\rangle = 0. \tag{2.6}$$

The following equations summarise our results for the transitions among the "pure" states $|\hat{z},+\rangle, |\hat{z},-\rangle, |\hat{z},0\rangle$:

$$\langle \hat{z},+|\hat{z},+\rangle = \langle \hat{z},0|\hat{z},0\rangle = \langle \hat{z},-|\hat{z},-\rangle \ = 1 \tag{2.7a}$$

and

$$\langle \hat{z},+|\hat{z},0\rangle = \langle \hat{z},0|\hat{z},+\rangle = \langle \hat{z},+|\hat{z},-\rangle \ = 0\,, \tag{2.7b}$$

$$\langle \hat{z},-|\hat{z},+\rangle = \langle \hat{z},0|\hat{z},-\rangle = \langle \hat{z},-|\hat{z},0\rangle \ = 0\,. \tag{2.7c}$$

2.2.3 Successive rotated Stern–Gerlach filters

We now perform a more interesting experiment with two Stern–Gerlach filters in a row, where the second filter is rotated by an angle θ. Specifically, in Filter 1 the magnetic field $\mathbf{B}$ varies along the $\hat{z}$ direction and in Filter 2 $\mathbf{B}$ varies along the $\hat{n}$ direction, with $\hat{n}\cdot\hat{z} = \cos\theta$ (Fig. 2.7). The first filter splits a beam of spin-1 atoms along the $\hat{z}$-axis and can select atoms

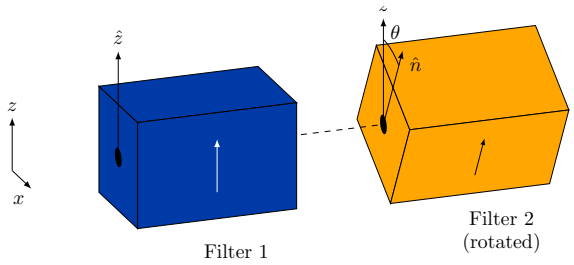


Figure 2.7. Two Stern–Gerlach filters with their directions of $\mathbf{B}$-field variation rotated by an angle θ.

in the "pure" states

$$|\hat{z},+\rangle, |\hat{z},-\rangle, |\hat{z},0\rangle\,.$$

The second filter splits a beam of spin-1 atoms along the $\hat{n}$-axis and can select atoms in the "pure" states

$$|\hat{n},+\rangle\,,|\hat{n},-\rangle\,,|\hat{n},0\rangle\,.$$

In our sequential experiments, we can find the probability that an atom transitions from a definite $|\hat{z},\{+,0,-\}\rangle$ state to a definite $|\hat{n},\{+,0,-\}\rangle$ state. We find that all such transitions are indeed possible, and therefore the probability amplitudes

$$\langle\hat{n},a|\hat{z},b\rangle\,,\quad a,b\in\{+,0,-\}$$

are different from zero. **Remark.** This holds for a generic angle $0<\theta<\pi$, not for every θ. For spin 1 the amplitudes are the Wigner functions $d^1_{m',m}(\theta)$, e.g. $\langle\hat{n},0|\hat{z},0\rangle=\cos\theta$, $\langle\hat{n},\pm|\hat{z},0\rangle=\mp\sin\theta/\sqrt{2}$ (up to phases), so the $0\to0$ transition is forbidden at $\theta=\pi/2$ (Sakurai–Napolitano, Sec. 3.5).

We also observe that a transition from a $|\hat{z},a\rangle$ state to some state $|\hat{n},b\rangle$ always takes place with 100% certainty. This gives

$$\left|\langle\hat{n},+|\hat{z},+\rangle\right|^2+\left|\langle\hat{n},0|\hat{z},+\rangle\right|^2+\left|\langle\hat{n},-|\hat{z},+\rangle\right|^2=1\,, \tag{2.8a}$$

$$\left|\langle\hat{n},+|\hat{z},-\rangle\right|^2+\left|\langle\hat{n},0|\hat{z},-\rangle\right|^2+\left|\langle\hat{n},-|\hat{z},-\rangle\right|^2=1\,, \tag{2.8b}$$

$$\left|\langle\hat{n},+|\hat{z},0\rangle\right|^2+\left|\langle\hat{n},0|\hat{z},0\rangle\right|^2+\left|\langle\hat{n},-|\hat{z},0\rangle\right|^2=1\,. \tag{2.8c}$$

2.2.4 Three filters

We now consider an experiment with three Stern–Gerlach filters (Fig. 2.8). The first and the third apparatus have a direction of **B**-field variation along the positive $\hat{z}$ axis. The second apparatus is rotated by an angle θ, with a **B**-field variation pointing along $\hat{n}$. We arrange

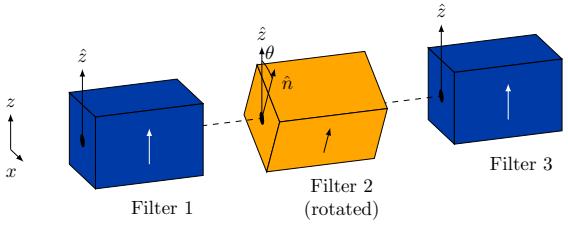


Figure 2.8. Three Stern–Gerlach filters in a row. The first and the third apparatus have a direction of **B**-field variation along the positive $\hat{z}$ axis. The second apparatus is rotated by an angle θ, with a **B**-field variation pointing along $\hat{n}$.

that the first filter sets the atoms in the state $|\hat{z},+\rangle$ and that the second filter puts them in the state $|\hat{n},0\rangle$. What happens when the atoms pass through the third filter? Do the atoms remember that they were previously in a $|\hat{z},+\rangle$ state? No! The atoms can transition to any of the three states $|\hat{z},\{+,0,-\}\rangle$ despite the fact that they have once been made to be in a pure $|\hat{z},+\rangle$ state.

Indeed, we can easily show that the fraction of atoms that ends up in any of the $|\hat{z},\{+,0,-\}\rangle$ states after the last filter is independent of the transition through the first filter. The

amplitudes for the transition through the second and third filters from the state $|\hat{z},+\rangle$ to, say, either a $|\hat{z},+\rangle$ or a $|\hat{z},-\rangle$ state are

$$\langle \hat{z},+|\hat{n},0\rangle \langle \hat{n},0|\hat{z},+\rangle$$

and

$$\langle \hat{z},-|\hat{n},0\rangle \langle \hat{n},0|\hat{z},+\rangle$$

respectively. The ratio of probabilities for the two transitions is

$$\frac{\left|\left\langle \hat{z},+\middle|\hat{n},0\right\rangle \left\langle \hat{n},0\middle|\hat{z},+\right\rangle\right|^2}{\left|\left\langle \hat{z},-\middle|\hat{n},0\right\rangle \left\langle \hat{n},0\middle|\hat{z},+\right\rangle\right|^2} = \frac{\left|\left\langle \hat{z},+\middle|\hat{n},0\right\rangle\right|^2}{\left|\left\langle \hat{z},-\middle|\hat{n},0\right\rangle\right|^2}$$

and does not depend on the state of the beam prior to the second filter.

2.3 Base states

The previous results illustrate one of the basic principles of quantum mechanics: atomic systems can be decomposed through a filtering process into *base states*. The evolution of the system in any of these states is independent of the past and depends solely on the nature of the base state. The base states depend on the filtering process. For example, the states $|\hat{z},\{+,0,-\}\rangle$ are one set of base states and the states $|\hat{n},\{+,0,-\}\rangle$ are another.

Let us now go back to our three-filter experiment and do the following:

i) open only the $\hat{z}$-"+" gate in the first filter;

ii) open only the $\hat{n}$-"0" gate in the second filter;

iii) open only the $\hat{z}$-"−" gate in the third filter.

If the beam exiting the first filter has N atoms, there will be

$$N \times |\langle \hat{z},-|\hat{n},0\rangle \langle \hat{n},0|\hat{z},+\rangle|^2$$

atoms exiting the last filter.

Now, let us open all the gates in the second filter. How many atoms will go through this time? The result of this experiment is very intriguing: no atom exits our apparatus! A filter with all its gates open has no effect, since no base state is selected. We can then write

$$\sum_{\text{all } i} \left\langle \hat{z},-\middle|i\right\rangle \left\langle i\middle|\hat{z},+\right\rangle = 0, \quad i \in \left|\hat{n},\{+,0,-\}\right\rangle. \tag{2.9}$$

To verify that this is a generic property of a "wide open" filter, we repeat the experiment with a middle Stern–Gerlach filter oriented in a direction $\hat{r}$ completely different from those of the other two filters. If the first filter is set to prepare the atoms in a state $\left|\phi\right\rangle$ and the third filter is set to prepare the atoms in a state $|\chi\rangle$, we find that the open second filter does not affect the transition from $\left|\phi\right\rangle$ to $|\chi\rangle$:

$$\left\langle \chi\middle|\phi\right\rangle = \sum_{\text{all } i} \langle \chi|i\rangle \left\langle i\middle|\phi\right\rangle . \tag{2.10}$$

Let us list here the properties of base states:

i) If a system is in a base state, then its future evolution is independent of the past.

ii) Base states satisfy Eq. (2.10).

iii) Base states are completely different from each other:

$$\langle i|j\rangle = \delta_{ij}. \tag{2.11}$$

Base states are not unique; they depend on the filtering method. For example, a Stern–Gerlach filter in the $\hat{n}$ direction and a Stern–Gerlach filter in a different $\hat{r}$ direction yield a different set of base states.

Let us now compare Eq. (2.10) and Eq. (2.8a). In order for both of them to be valid, we require one more rule, for the complex conjugation of probability amplitudes, namely

$$\langle \phi|\chi\rangle^* = \langle \chi|\phi\rangle . \tag{2.12}$$

2.4 States and vector spaces

We now take a bold step. We postulate that atomic states live in a vector space. We will also assume that the pure states filtered with the aid of a physical measurement, such as angular momentum in the $\hat{z}$ direction, form an orthonormal basis of that vector space. In our example, we can write

$$|\hat{z},+\rangle = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad |\hat{z},0\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad |\hat{z},-\rangle = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \tag{2.13}$$

A generic state will be expressed as a superposition of basis states,

$$|\phi\rangle = \phi_1 |\hat{z},+\rangle + \phi_2 |\hat{z},0\rangle + \phi_3 |\hat{z},-\rangle = \begin{pmatrix} \phi_1 \\ \phi_2 \\ \phi_3 \end{pmatrix}. \tag{2.14}$$

We will also equip the vector space with an inner product, which we will physically identify with a transition amplitude. For

$$|\phi\rangle = \begin{pmatrix} \phi_1 \\ \phi_2 \\ \phi_3 \end{pmatrix}, \quad |\chi\rangle = \begin{pmatrix} \chi_1 \\ \chi_2 \\ \chi_3 \end{pmatrix}, \tag{2.15}$$

we define

$$\langle \phi|\chi\rangle = \phi_1^* \chi_1 + \phi_2^* \chi_2 + \phi_3^* \chi_3. \tag{2.16}$$

Equivalently, we can define dual vectors, "bras" in Dirac's notation, as

$$\langle \phi| = (\phi_1^*, \phi_2^*, \phi_3^*) , \tag{2.17}$$

and we can think of the inner product as the multiplication of a dual "bra" vector and an ordinary "ket" vector. It will be more practical for us to think of the inner product in this equivalent way.

The above assumptions of a vector space and an inner product are consistent with the physical properties that we inferred from our experiments with the Stern–Gerlach apparatus. For example, the property

$$\langle \phi|\chi\rangle = \phi_1^* \chi_1 + \phi_2^* \chi_2 + \phi_3^* \chi_3 = \left(\chi_1^* \phi_1 + \chi_2^* \phi_2 + \chi_3^* \phi_3\right)^* = \langle \chi|\phi\rangle^* \tag{2.18}$$

emerges naturally. Let us now focus on the "completeness" identity,

$$\langle\phi|\chi\rangle = \sum_i \langle\phi|i\rangle\langle i|\chi\rangle = \langle\phi|\left(\sum_i |i\rangle\langle i|\right)|\chi\rangle\,, \tag{2.19}$$

with $|i\rangle \in \{|\hat{z},+\rangle\,, |\hat{z},0\rangle\,, |\hat{z},-\rangle\}$. From the vector representation of the base states, namely

$$\sum_i |i\rangle\langle i| = \begin{pmatrix}1\\0\\0\end{pmatrix}(1,0,0) + \begin{pmatrix}0\\1\\0\end{pmatrix}(0,1,0) + \begin{pmatrix}0\\0\\1\end{pmatrix}(0,0,1) = 1_{3\times3}, \tag{2.20}$$

we naturally derive this result too.

Recall the values of the angular momentum component L_z that we observed experimentally for each element of our basis of states. We have

$$\begin{pmatrix}1\\0\\0\end{pmatrix} \to \hbar, \quad \begin{pmatrix}0\\1\\0\end{pmatrix} \to 0, \quad \begin{pmatrix}0\\0\\1\end{pmatrix} \to -\hbar. \tag{2.21}$$

We can define a matrix

$$\hat{L}_z \equiv \hbar\begin{pmatrix}1&0&0\\0&0&0\\0&0&-1\end{pmatrix}, \tag{2.22}$$

with the property that the basis states are its eigenvectors and the corresponding eigenvalues are the values of the angular momentum L_z measured in each state. Namely,

$$\hat{L}_z|\hat{z},+\rangle = \hbar|\hat{z},+\rangle\,, \tag{2.23}$$
$$\hat{L}_z|\hat{z},0\rangle = 0|\hat{z},0\rangle\,, \tag{2.24}$$
$$\hat{L}_z|\hat{z},-\rangle = -\hbar|\hat{z},-\rangle\,. \tag{2.25}$$

We pass a beam of atoms, all prepared identically in the same state $|\psi\rangle$, through a Stern–Gerlach apparatus. We measure the following average value of the angular momentum L_z:

$$\langle L_z\rangle = \left|\langle\hat{z},+|\psi\rangle\right|^2\hbar + \left|\langle\hat{z},0|\psi\rangle\right|^2 0 + \left|\langle\hat{z},-|\psi\rangle\right|^2(-\hbar). \tag{2.26}$$

Using that $\langle\phi|\chi\rangle = \langle\chi|\phi\rangle^*$ and that the matrix $\hat{L}_z$ has the set $|\hat{z},\{+,0,-\}\rangle$ as eigenvectors, we cast the above equation as

$$\begin{aligned}\langle L_z\rangle &= \langle\psi|\hat{z},+\rangle\langle\hat{z},+|\psi\rangle\hbar + \langle\psi|\hat{z},0\rangle\langle\hat{z},0|\psi\rangle 0 + \langle\psi|\hat{z},-\rangle\langle\hat{z},-|\psi\rangle(-\hbar)\\ &= \langle\psi|\hat{L}_z|\hat{z},+\rangle\langle\hat{z},+|\psi\rangle + \langle\psi|\hat{L}_z|\hat{z},0\rangle\langle\hat{z},0|\psi\rangle + \langle\psi|\hat{L}_z|\hat{z},-\rangle\langle\hat{z},-|\psi\rangle\\ &= \langle\psi|\left[\hat{L}_z\left(\sum_{i\in\{+,0,-\}}|\hat{z},i\rangle\langle\hat{z},i|\right)\right]|\psi\rangle\,.\end{aligned} \tag{2.27}$$

Now, we can use

$$\sum_{i\in\{+,0,-\}}|\hat{z},i\rangle\langle\hat{z},i| = 1, \tag{2.28}$$

and we obtain that the average value of the physical measurement for our observable, the angular momentum component L_z, is given by

$$\langle L_z \rangle = \left\langle \psi \right| \hat{L}_z \left| \psi \right\rangle . \tag{2.29}$$

Similarly, if we start from the experimental statement that the sum of the probabilities for all the transitions

$$\left| \psi \right\rangle \longrightarrow |\hat{z}, i\rangle , \quad i \in \{+, 0, -\}$$

is unity, we can conclude that

$$\left\langle \psi | \psi \right\rangle = 1. \tag{2.30}$$

The matrix operator $\hat{L}_z$ is a "generator" of a symmetry transformation. The connection to symmetry will be explored in more depth in subsequent chapters. Here, we will just give a first glimpse of it. With a standard linear-algebra computation, we can compute the exponential of $\hat{L}_z$, weighted by a parameter θ, as

$$U(\theta) \equiv \exp\left(-i \frac{\hat{L}_z}{\hbar} \theta\right) = \begin{pmatrix} e^{-i\theta} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & e^{i\theta} \end{pmatrix}. \tag{2.31}$$

The parameter θ is arbitrary for now, but we will be able to relate it to a rotation angle. Let us now define a new transformed state

$$\left| \psi' \right\rangle = U(\theta) \left| \psi \right\rangle . \tag{2.32}$$

The average value of L_z in the transformed state is

$$\langle L_z \rangle' = \left\langle \psi' \right| \hat{L}_z \left| \psi' \right\rangle = \left\langle \psi \right| U^\dagger(\theta) \, \hat{L}_z \, U(\theta) \left| \psi \right\rangle . \tag{2.33}$$

We can easily show that

$$U^\dagger(\theta) \, \hat{L}_z \, U(\theta) = \hat{L}_z , \tag{2.34}$$

from which we conclude that

$$\langle L_z \rangle' = \langle L_z \rangle . \tag{2.35}$$

Classically, we expect that a vector remains unchanged if we perform a rotation by an angle θ around its axis. The matrix $U(\theta)$ encodes, as we will learn in detail later, the action of this rotational symmetry on the space of quantum states. The classical result also holds in quantum mechanics, but for the average values of measurements of physical observables.
Remark. More precisely, the classical statement that matches the computation is that the component along the rotation axis of *any* vector is unchanged by a rotation about that axis; here $\langle L_z \rangle$ is invariant for every state $\left| \psi \right\rangle$, not only for angular momenta pointing along $\hat{z}$. Under the same $U(\theta)$, $\langle L_x \rangle$ and $\left\langle L_y \right\rangle$ rotate into each other.

3

Quantum mechanics and linear algebra

In quantum mechanics we are interested in computing the probability amplitudes $\langle b|a\rangle$ for transitions from an initial state $|a\rangle$ to a final state $|b\rangle$. In this chapter, we will postulate that such amplitudes are the inner products of vectors in a space of physical states,

$$\langle b|a\rangle \leftrightarrow \mathbf{b}\cdot\mathbf{a}.$$

Mathematically, the space of quantum states is a vector space of a special kind, known as a Hilbert space. All physical information about a physical system (e.g. a particle, an atom, a system of many particles) is encoded in these vectors in a Hilbert space. In what follows, we will describe the properties of the Hilbert space of quantum states.

3.1 Ket-space

Physical states are represented by vectors in a complex vector space. We call such a state vector a *ket*. Kets possess complete information about the physical system. The dimensionality of the ket space is equal to the number of base states which we can obtain with a filtering experiment, such as our Stern–Gerlach experiment.

i) Two kets can be added, yielding a new ket for another physical state:

$$|\alpha\rangle + |\beta\rangle = |\gamma\rangle\,. \tag{3.1}$$

ii) The multiplication of a ket by a complex number yields a new ket which, however, corresponds to the same physical state as the original:

$$c\,|\chi\rangle \text{ and } |\chi\rangle \rightarrow \text{same physical state}, \quad c \neq 0. \tag{3.2}$$

3.2 Operators

The state of a physical system can change during an experiment or by letting the system evolve with time. We represent changes in physical states mathematically by the action of operators on kets:

$$A\,|\psi\rangle = |\phi\rangle \qquad (\text{Operator}\,|\text{state}\rangle = |\text{new state}\rangle). \tag{3.3}$$

For a given operator A, there are some special states, $|i\rangle$, which are not changed, up to a multiplicative factor λ_i, by the action of the operator,

$$A\,|i\rangle = \lambda_i\,|i\rangle\,. \tag{3.4}$$

The values λ_i are called the eigenvalues of the operator. The eigenstates of a physical operator are postulated to correspond to base states: any other physical state $|\phi\rangle$ can be written as a superposition of the base states,

$$\forall |\phi\rangle \ \exists \{c_i\} : \quad |\phi\rangle = \sum_i c_i |i\rangle .$$

3.3 Dual space

We define a dual space of "bra" states $\langle\phi|$,

$$|\phi\rangle \leftrightarrow \langle\phi| . \tag{3.5}$$

By dual we mean that all information about the bra $\langle\phi|$ is already encoded in the ket $|\phi\rangle$ and we only need to use it for a different purpose. We can think of kets as states of an "initial condition" and bras as states of a "final condition". The duality means that all states are eligible as starting or ending states of an experiment. For a general superposition of ket states, the bra dual reads

$$c_1 |a_1\rangle + \ldots + c_n |a_n\rangle \leftrightarrow c_1^* \langle a_1| + \ldots + c_n^* \langle a_n| , \tag{3.6}$$

where we conjugate the multiplicative factors in front of every bra component.

3.4 Inner product

The dual bra space serves to define probability amplitudes as inner products:

$$\langle a|b\rangle \quad \text{(inner product)}. \tag{3.7}$$

We postulate that

$$\langle a|b\rangle = \langle b|a\rangle^* , \tag{3.8}$$

and that the "square" of a state is positive definite:

$$\langle\phi|\phi\rangle \geq 0, \quad \forall |\phi\rangle , \qquad \langle\phi|\phi\rangle = 0 \Leftrightarrow |\phi\rangle = 0. \tag{3.9}$$

We call $\sqrt{\langle\phi|\phi\rangle}$ the norm of the state $|\phi\rangle$. We can normalise all states to unit norm without altering the physics,

$$|\phi\rangle \to |\tilde{\phi}\rangle = \frac{|\phi\rangle}{\sqrt{\langle\phi|\phi\rangle}}, \tag{3.10}$$

where $\langle\tilde{\phi}|\tilde{\phi}\rangle = 1$.

3.5 Properties of operators and their duals

i) Two operators are said to be equal, $X = Y$, if

$$X |\phi\rangle = Y |\phi\rangle , \quad \forall |\phi\rangle . \tag{3.11}$$

ii) An operator is zero, $X = 0$, if

$$X |\phi\rangle = 0, \quad \forall |\phi\rangle . \tag{3.12}$$

iii) Operators can be added together, with the properties

$$X + Y = Y + X\,, \tag{3.13}$$

$$X + (Y + Z) = (X + Y) + Z = X + Y + Z\,. \tag{3.14}$$

iv) Operators can be multiplied together, with the property

$$\begin{aligned} X(YZ) &= (XY)Z \\ &= XYZ\,. \end{aligned} \tag{3.15}$$

v) However, the multiplication order is important and, in general,

$$XY \neq YX.$$

The dual of an operator acting on a ket, $X\left|\phi\right\rangle$, involves in general a different operator,

$$X\left|\phi\right\rangle \leftrightarrow \left\langle\phi\right| X^\dagger, \tag{3.16}$$

with

$$X^\dagger \neq X.$$

$X^\dagger$ is called the "Hermitian adjoint" of X. Operators with the special property that $X^\dagger = X$ are called Hermitian operators.

The dual of a product of operators is

$$(XY)^\dagger = Y^\dagger X^\dagger. \tag{3.17}$$

Proof. Let us write

$$Y\left|a\right\rangle = \left|b\right\rangle,$$

where the dual is

$$\left\langle b\right| = \left\langle a\right| Y^\dagger\,.$$

Then

$$XY\left|a\right\rangle = X(Y\left|a\right\rangle) = X\left|b\right\rangle \leftrightarrow \left\langle b\right| X^\dagger = \left\langle a\right| Y^\dagger X^\dagger,$$

which proves our assertion.

We can define an outer product $\left|b\right\rangle\left\langle a\right|$, which is an operator turning a generic state $\left|\phi\right\rangle$ into the state $\left|b\right\rangle$:

$$\Big(\left|b\right\rangle\left\langle a\right|\Big)\left|\phi\right\rangle = \left|b\right\rangle\Big(\left\langle a|\phi\right\rangle\Big) = \Big(\left\langle a|\phi\right\rangle\Big)\left|b\right\rangle, \tag{3.18a}$$

$$\big(\text{operator}\big)\left|\text{state}\right\rangle = \ldots = \big(\text{number}\big)\left|\text{new state}\right\rangle. \tag{3.18b}$$

The dual of an outer product is also an outer product,

$$\Big(\left|\phi\right\rangle\left\langle\psi\right|\Big)^\dagger = \left|\psi\right\rangle\left\langle\phi\right|. \tag{3.19}$$

Indeed,

$$\Big(\left|\phi\right\rangle\left\langle\psi\right|\Big)\left|a\right\rangle = \left|\phi\right\rangle\Big(\left\langle\psi|a\right\rangle\Big) \leftrightarrow \left\langle\phi\right|\Big(\left\langle\psi|a\right\rangle\Big)^* = \Big(\left\langle a|\psi\right\rangle\Big)\left\langle\phi\right| = \left\langle a\right|\Big(\left|\psi\right\rangle\left\langle\phi\right|\Big).$$

3.6 Hermitian operators

For a general operator X, we can prove that

$$\langle a| X |b\rangle = \langle b| X^\dagger |a\rangle^* . \tag{3.20}$$

Proof.

$$\left\langle a\right| X \left|b\right\rangle = \langle a| \left(X \left|b\right\rangle\right) = \left(\left(\left\langle b\right| X^\dagger\right) |a\rangle\right)^* = \left\langle b\right| X^\dagger \left|a\right\rangle^* .$$

Hermitian operators, $H^\dagger = H$, have the special property that

$$\langle a| H |b\rangle = \langle b| H |a\rangle^* .$$

We can prove the following:

i) The eigenvalues of a Hermitian operator are real.

Proof. For a Hermitian operator H with eigenstates $|i\rangle$,

$$H |i\rangle = \lambda_i |i\rangle .$$

The dual of the above equation is

$$\langle i| H = \lambda_i^* \langle i| ,$$

where we have exploited the Hermiticity of H ($H^\dagger = H$). Multiplying the first equation by $\langle i|$ from the left and the second equation by $|i\rangle$ from the right, we obtain (for a normalised eigenstate)

$$\langle i| H |i\rangle = \lambda_i = \lambda_i^* ,$$

which proves that the eigenvalue λ_i is real.

ii) The eigenstates of a Hermitian operator with non-degenerate eigenvalues are orthogonal.

Proof. Following the same reasoning as above, we can easily show that for two eigenstates $|i\rangle , \left|j\right\rangle$ of a Hermitian operator H, we can write the quantity $\left\langle i\right| H \left|j\right\rangle$ in two alternative ways:

$$\left\langle i\right| H \left|j\right\rangle = \lambda_i \left\langle i|j\right\rangle = \lambda_j \left\langle i|j\right\rangle ,$$

which leads to

$$(\lambda_i - \lambda_j) \left\langle i|j\right\rangle = 0.$$

For non-degenerate eigenvalues, i.e.

$$\lambda_i \neq \lambda_j, \quad \forall |i\rangle \neq \left|j\right\rangle$$

we conclude that

$$\left\langle i|j\right\rangle = 0, \quad \forall |i\rangle \neq \left|j\right\rangle$$

and the eigenstates are orthogonal.

We typically normalise eigenstates of Hermitian operators so that they are orthonormal:

$$\langle i|j\rangle = \delta_{ij}. \tag{3.21}$$

Remark. For degenerate eigenvalues the proof above gives no orthogonality, but within each degenerate eigenspace one can always choose an orthonormal basis (Gram–Schmidt), so a Hermitian operator in finite dimension always admits an orthonormal eigenbasis. Likewise, the theorem below on commuting operators extends to the degenerate case: B is then block-diagonal and can be diagonalised within each degenerate block of A (Sakurai–Napolitano, Sec. 1.4).

We postulate that a Hermitian operator A which corresponds to a physical observable has a "complete" set of eigenstates $\{|i\rangle\}$. Every other physical state can be written as a superposition:

$$\forall |\phi\rangle \;\; \exists\{c_i\} \;:\; |\phi\rangle = \sum_i c_i |i\rangle . \tag{3.22}$$

The coefficients c_i are the transition amplitudes from the state $|\phi\rangle$ to the eigenstate $|i\rangle$:

$$c_i = \langle i|\phi\rangle . \tag{3.23}$$

Indeed,

$$\sum_j c_j |j\rangle = |\phi\rangle$$

$$\rightsquigarrow \sum_j c_j \langle i|j\rangle = \langle i|\phi\rangle$$

$$\rightsquigarrow \sum_j c_j \delta_{ij} = \langle i|\phi\rangle$$

$$\rightsquigarrow c_i = \langle i|\phi\rangle .$$

From Eq. (3.22) and Eq. (3.23) we obtain that for every state $|\phi\rangle$,

$$|\phi\rangle = \sum_i \left(\langle i|\phi\rangle\right) |i\rangle , \tag{3.24}$$

which we can rearrange into

$$|\phi\rangle = \left(\sum_i |i\rangle\langle i|\right) |\phi\rangle . \tag{3.25}$$

Thus, for base states $|i\rangle$ which are eigenstates of a non-degenerate Hermitian operator, we have

$$1 = \sum_{\text{all } i} |i\rangle\langle i| . \tag{3.26}$$

For a normalised state $|\phi\rangle$ we have

$$1 = \langle\phi|\phi\rangle = \sum_i \langle\phi|i\rangle \langle i|\phi\rangle \tag{3.27}$$

$$\rightsquigarrow 1 = \sum_i \left|\langle i|\phi\rangle\right|^2 . \tag{3.28}$$

This is consistent with the probabilistic interpretation of the inner product, which associates the probability $P\left(|\phi\rangle \to |i\rangle\right)$ for a transition from a state $|\phi\rangle$ to an eigenstate with the modulus squared of the inner product,

$$P\left(|\phi\rangle \to |i\rangle\right) = \left|\langle i|\phi\rangle\right|^2 . \tag{3.29}$$

In a filtering experiment (corresponding to the operator A with eigenstates $|i\rangle$) of a quantum mechanical system (such as an atom in a state $|\phi\rangle$), the probability that the system passes through one of the filters (the state $|\phi\rangle$ collapses to one of the eigenstates $|i\rangle$) is one:

$$\sum_{\text{all } i} P\left(|\phi\rangle \to |i\rangle\right) = 1 . \tag{3.30}$$

The operator $\Lambda_i \equiv |i\rangle\langle i|$ projects a general state onto the eigenstate $|i\rangle$ of the Hermitian operator A. Indeed,

$$\begin{aligned} \Lambda_i |\phi\rangle &= \left(|i\rangle\langle i|\right)|\phi\rangle = \Big(\langle i|\phi\rangle\Big)|i\rangle \\ &= \Big[\text{amplitude}(|\phi\rangle \to |i\rangle)\Big] \times |i\rangle . \end{aligned} \tag{3.31}$$

This operator has the defining property of a projector,

$$\begin{aligned} \Lambda_i\Lambda_j &= \left(|i\rangle\langle i|\right)\Big(|j\rangle\langle j|\Big) = |i\rangle\langle i|j\rangle\langle j| \\ &= |i\rangle\,\delta_{ij}\langle j| = |i\rangle\langle i|\,\delta_{ij} \\ &= \Lambda_i\delta_{ij} . \end{aligned}$$

For $i = j$, we find that $\Lambda_i^2 = \Lambda_i$, which tells us that filtering a quantum system onto a pure state $|i\rangle$ repeatedly does not induce any further change to the system after the first filtering. For $i \neq j$, we find that $\Lambda_i\Lambda_j = 0$. A system which is already filtered onto a pure state $|i\rangle$ cannot transition directly into a different pure state of the same observable.

3.7 Matrix representation of states and operators

Consider a Hermitian operator A with a set of eigenstates $\{|i\rangle\}$, $i = 1 \ldots N$. These satisfy the orthonormality condition $\langle j|i\rangle = \delta_{ij}$. We can represent these eigenstates as vectors,

$$|1\rangle \doteq \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad |2\rangle \doteq \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \quad \cdots \quad , \quad |N\rangle \doteq \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}. \tag{3.32a}$$

We represent the dual bra eigenstates as

$$\langle 1| \doteq (1,0,\ldots,0)\,, \quad \langle 2| \doteq (0,1,\ldots,0)\,, \quad \cdots \quad , \quad \langle N| \doteq (0,0,\ldots,1)\,. \tag{3.32b}$$

The above representations of the bra and ket eigenstates are consistent with their orthonormality condition. For example,

$$\langle 1|2\rangle \doteq (1,0,\ldots,0)\begin{pmatrix}0\\1\\\vdots\\0\end{pmatrix} = 0$$

and

$$\langle 2|2\rangle \doteq (0,1,\ldots,0)\begin{pmatrix}0\\1\\\vdots\\0\end{pmatrix} = 1$$

etc.

A general state $\big|\phi\big\rangle$ must satisfy

$$\big|\phi\big\rangle = \sum_i \big\langle i\big|\phi\big\rangle\,|i\rangle\,. \tag{3.33}$$

This can be represented as

$$\big|\phi\big\rangle \doteq \big\langle 1\big|\phi\big\rangle\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix} + \big\langle 2\big|\phi\big\rangle\begin{pmatrix}0\\1\\\vdots\\0\end{pmatrix} + \ldots + \big\langle N\big|\phi\big\rangle\begin{pmatrix}0\\0\\\vdots\\1\end{pmatrix} \tag{3.34}$$

yielding,

$$\big|\phi\big\rangle \doteq \begin{pmatrix}\big\langle 1\big|\phi\big\rangle\\\big\langle 2\big|\phi\big\rangle\\\vdots\\\big\langle N\big|\phi\big\rangle\end{pmatrix}.$$

For a bra state $\big\langle\phi\big|$ we write

$$\big\langle\phi\big| = \big\langle\phi\big|\sum_i |i\rangle\langle i| = \sum_i \big\langle\phi\big|i\big\rangle\,\langle i| = \sum_i \big\langle i\big|\phi\big\rangle^*\,\langle i|\,. \tag{3.35}$$

This is represented as

$$\big\langle\phi\big| \doteq \Big(\big\langle 1\big|\phi\big\rangle^*, \big\langle 2\big|\phi\big\rangle^*, \ldots, \big\langle N\big|\phi\big\rangle^*\Big)\,. \tag{3.36}$$

The inner product of two general states is

$$\langle a|b\rangle = \langle a|\left(\sum_i |i\rangle\langle i|\right)|b\rangle = \sum_i \langle a|i\rangle\,\langle i|b\rangle = \sum_i \langle i|a\rangle^*\,\langle i|b\rangle\,. \tag{3.37}$$

This is consistent with the result that we obtain by using our representation,

$$\langle a|b\rangle \doteq (\langle 1|a\rangle^*, \langle 2|a\rangle^*, \ldots, \langle N|a\rangle^*) \begin{pmatrix} \langle 1|b\rangle \\ \langle 2|b\rangle \\ \vdots \\ \langle N|b\rangle \end{pmatrix}. \tag{3.38}$$

A general operator X can be written as

$$X = \left(\sum_i |i\rangle\langle i|\right) X \left(\sum_j |j\rangle\langle j|\right) = \sum_{ij} |i\rangle \langle i|X|j\rangle \langle j|. \tag{3.39}$$

The operator is represented as an $N \times N$ matrix,

$$X \doteq \begin{pmatrix} \langle 1|X|1\rangle & \langle 1|X|2\rangle & \ldots & \langle 1|X|N\rangle \\ \langle 2|X|1\rangle & \langle 2|X|2\rangle & \ldots & \langle 2|X|N\rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle N|X|1\rangle & \langle N|X|2\rangle & \ldots & \langle N|X|N\rangle \end{pmatrix}. \tag{3.40}$$

For the case of the outer product operator, we obtain the following representation:

$$|a\rangle\langle b| \doteq \begin{pmatrix} \langle 1|a\rangle\langle 1|b\rangle^* & \langle 1|a\rangle\langle 2|b\rangle^* & \ldots & \langle 1|a\rangle\langle N|b\rangle^* \\ \langle 2|a\rangle\langle 1|b\rangle^* & \langle 2|a\rangle\langle 2|b\rangle^* & \ldots & \langle 2|a\rangle\langle N|b\rangle^* \\ \vdots & \vdots & \ddots & \vdots \\ \langle N|a\rangle\langle 1|b\rangle^* & \langle N|a\rangle\langle 2|b\rangle^* & \ldots & \langle N|a\rangle\langle N|b\rangle^* \end{pmatrix}. \tag{3.41}$$

3.8 Compatible and incompatible observables

Two Hermitian operators which commute with each other,

$$[A, B] \equiv AB - BA = 0, \tag{3.42}$$

are called compatible. If they do not commute, $[A, B] \neq 0$, they are incompatible. We shall prove a couple of theorems which elucidate the meaning of this terminology.

Theorem 3.1. *Let A, B be two Hermitian operators with $[A, B] = 0$, where A has a spectrum of eigenstates $|i\rangle$ with non-degenerate eigenvalues,*

$$A|i\rangle = \lambda_i |i\rangle.$$

Then

i) B is a diagonal matrix in the representation of the $|i\rangle$ basis;

ii) the set of $|i\rangle$ states is also a set of eigenstates of the B operator.

Proof. For any two eigenstates $|i\rangle, |j\rangle$ of A we have

$$\begin{aligned} 0 &= [A, B] \\ \rightsquigarrow 0 &= \langle i|AB - BA|j\rangle = \langle i|AB|j\rangle - \langle i|BA|j\rangle \\ &= (\lambda_i - \lambda_j)\langle i|B|j\rangle .^{1} \end{aligned}$$

For $|i\rangle \neq |j\rangle$, given that the eigenvalues are not degenerate, we have $\lambda_i \neq \lambda_j$. Thus, it must be that $\langle i| B |j\rangle = 0$. Only the diagonal elements, $i = j$, are allowed to be different from zero. We write

$$\langle i| B |j\rangle = \delta_{ij} \langle i| B |i\rangle , \tag{3.43}$$

which is the first statement of our theorem.

The B operator can be written as

$$\begin{aligned} B &= \left(\sum_i |i\rangle\langle i| \right) B \left(\sum_j |j\rangle\langle j| \right) \\ &= \sum_{ij} |i\rangle \langle i| B |j\rangle \langle j| \\ &= \sum_{ij} |i\rangle \delta_{ij} \langle j| \left(\langle i| B |i\rangle \right) \\ &= \sum_i \left(\langle i| B |i\rangle \right) |i\rangle\langle i|. \end{aligned}$$

Acting on an eigenstate $|m\rangle$ of the A operator, we have

$$\begin{aligned} B |m\rangle &= \sum_i \left(\langle i| B |i\rangle \right) |i\rangle\langle i| |m\rangle \\ &= \sum_i \left(\langle i| B |i\rangle \right) |i\rangle \delta_{im} \\ &= \left(\langle m| B |m\rangle \right) |m\rangle , \end{aligned}$$

which proves that the eigenstates $|m\rangle$ of A are also eigenstates of the commuting operator B.

Theorem 3.2. *If two operators do not commute, $[A, B] \neq 0$, their common eigenstates do not form a complete set.*

Proof. We can prove the above by assuming the opposite, i.e. that the common eigenstates of A, B, denoted by $|a, b\rangle$ and satisfying

$$\begin{aligned} A |a, b\rangle &= a |a, b\rangle , \\ B |a, b\rangle &= b |a, b\rangle , \end{aligned}$$

form a complete set. Then a general state can be written as

$$|\phi\rangle = \sum_{a,b} c_{ab} |a, b\rangle . \tag{3.44}$$

Acting with the commutator on an arbitrary state $|\phi\rangle$, we obtain

$$[A, B] |\phi\rangle = \sum_{a,b} c_{ab} (AB - BA) |a, b\rangle = \sum_{a,b} c_{ab} (ab - ba) |a, b\rangle = 0. \tag{3.45}$$

From the above we conclude that the commutator vanishes, $[A, B] = 0$, which contradicts our hypothesis.

[1] $\langle i| AB - BA |j\rangle$ can be seen as $\langle i| (AB - BA) |j\rangle$.

Remark. The theorem does not say that incompatible observables have *no* common eigenstate. For example $[L_x, L_y] = i\hbar L_z \neq 0$, yet every state with $\ell = 0$ is a simultaneous eigenstate of L_x and L_y (eigenvalue 0). What fails is only the existence of a complete common eigenbasis (Sakurai–Napolitano, Sec. 1.4).

3.9 Expectation value and uncertainty of Hermitian operators (measurements)

Consider a Hermitian operator A which corresponds to a physical observable, such as energy, momentum, position, spin magnetic moment, etc. We associate the average value measured in an experiment (which is repeated infinitely many times) for the quantity corresponding to A with the expectation value of the operator with respect to the system's state $|\phi\rangle$:

$$\langle A \rangle = \langle \phi | A | \phi \rangle . \tag{3.46}$$

Let us use the eigenstates $|i\rangle$ of the operator A as a basis of physical states. These have eigenvalues λ_i,

$$A |i\rangle = \lambda_i |i\rangle . \tag{3.47}$$

The expectation value of A can be written as

$$\langle A \rangle = \langle \phi | \left(\sum_i |i\rangle\langle i| \right) A \left(\sum_j |j\rangle\langle j| \right) |\phi\rangle , \tag{3.48}$$

which yields

$$\langle A \rangle = \sum_i \lambda_i \left| \langle i|\phi\rangle \right|^2 . \tag{3.49}$$

We can rewrite this expression as

$$\langle A \rangle = \sum_i \lambda_i \mathrm{Prob}(|\phi\rangle \to |i\rangle) . \tag{3.50}$$

The expectation value is a sum over all possible eigenvalues weighted by the probability that the state of the system collapses to the corresponding eigenstate. If $|\phi\rangle$ is itself an eigenstate of A, e.g. $|\phi\rangle = |j\rangle$, the expectation value of the measurement is simply the eigenvalue λ_j:

$$\langle A \rangle = \langle j | A | j \rangle = \sum_i \lambda_i \left| \langle i|j\rangle \right|^2 = \sum_i \lambda_i \delta_{ij} = \lambda_j . \tag{3.51}$$

In addition to the average of measurements of an observable in a quantum mechanical system, we can compute the uncertainty in these measurements. This is defined as

$$\left\langle (\Delta A)^2 \right\rangle \equiv \left\langle (A - \langle A \rangle \mathbf{1})^2 \right\rangle . \tag{3.52}$$

[2] Equivalently,

$$\left\langle (\Delta A)^2 \right\rangle = \left\langle A^2 \right\rangle - \langle A \rangle^2 , \tag{3.53}$$

[2] From now on we will drop the **1**.

which we can easily prove:

$$\begin{aligned}\left\langle \left(A - \langle A\rangle\right)^2 \right\rangle &= \left\langle A^2 \right\rangle - 2\left\langle A\langle A\rangle\right\rangle + \langle A\rangle^2 = \left\langle A^2\right\rangle - 2\langle A\rangle^2 + \langle A\rangle^2 \\ &= \left\langle A^2\right\rangle - \langle A\rangle^2 ,\end{aligned} \tag{3.54}$$

where

$$\begin{aligned}\left\langle A\langle A\rangle\right\rangle &= \langle\phi| A \underbrace{\langle\phi| A |\phi\rangle}_{\text{number}} |\phi\rangle = \langle\phi| A |\phi\rangle \langle\phi| A |\phi\rangle = \left(\langle\phi| A |\phi\rangle\right)^2 \\ &= \langle A\rangle^2 .\end{aligned} \tag{3.55}$$

Notice that the uncertainty for a system in an eigenstate of A is zero,

$$\left\langle (\Delta A)^2 \right\rangle = \langle i| A^2 |i\rangle - \langle i| A |i\rangle^2 = \lambda_i^2 - \lambda_i^2 = 0. \tag{3.56}$$

3.10 The uncertainty principle

In this section, we shall derive Heisenberg's uncertainty principle for any pair of incompatible (non-commuting) Hermitian (physical) operators A, B.

We start with

$$\begin{aligned}\left\langle (A - \langle A\rangle)(B - \langle B\rangle)\right\rangle &= \left\langle AB - A\langle B\rangle - \langle A\rangle B + \langle A\rangle\langle B\rangle\right\rangle \\ &= \langle AB\rangle - \langle A\rangle\langle B\rangle \\ &= \frac{1}{2}\left\langle \left[A, B\right]\right\rangle + \frac{1}{2}\langle AB + BA\rangle - \langle A\rangle\langle B\rangle \\ &= \frac{1}{2}\left\langle \left[A, B\right]\right\rangle + \frac{1}{2}\left\langle \left(A - \langle A\rangle\right)\left(B - \langle B\rangle\right) + \left(B - \langle B\rangle\right)\left(A - \langle A\rangle\right)\right\rangle \\ &= \frac{1}{2}\langle [A, B]\rangle + \frac{1}{2}\langle \{A - \langle A\rangle, B - \langle B\rangle\}\rangle ,\end{aligned} \tag{3.57}$$

where the anti-commutator is defined as

$$\{X, Y\} \equiv XY + YX. \tag{3.58}$$

The anti-commutator of Hermitian operators is also a Hermitian operator. Indeed,

$$\{X, Y\}^\dagger = (XY)^\dagger + (YX)^\dagger = Y^\dagger X^\dagger + X^\dagger Y^\dagger = YX + XY = XY + YX = \{X, Y\} .$$

The commutator of two Hermitian operators, on the other hand, is anti-Hermitian.[3] Indeed,

$$[A, B]^\dagger = [B^\dagger, A^\dagger] = [B, A] = -[A, B].$$

Now we can prove an important theorem:

[3] An operator Z is anti-Hermitian if its adjoint is $Z^\dagger = -Z$.

Theorem 3.3. *The expectation value of a Hermitian operator is real, while the expectation value of an anti-Hermitian operator is imaginary.*

Proof. Indeed, the real and imaginary parts of the expectation value for an operator A are

$$\Re\left(\langle A\rangle\right) = \frac{\langle A\rangle + \langle A\rangle^*}{2} = \frac{1}{2}\left\langle\phi\right| A + A^\dagger \left|\phi\right\rangle, \tag{3.59}$$

$$\Im\left(\langle A\rangle\right) = \frac{\langle A\rangle - \langle A\rangle^*}{2i} = \frac{1}{2i}\left\langle\phi\right| A - A^\dagger \left|\phi\right\rangle. \tag{3.60}$$

The real part vanishes if A is anti-Hermitian, $A^\dagger = -A$, while the imaginary part vanishes if A is Hermitian, $A^\dagger = A$.

On the right-hand side of Eq. (3.57), the first term is then purely imaginary, while the second term is real. Both give independent, positive contributions to the modulus squared of the left-hand side, and we can write

$$\left|\left\langle (A - \langle A\rangle)(B - \langle B\rangle)\right\rangle\right|^2 = \left|\frac{1}{2}\left\langle [A, B]\right\rangle\right|^2 + \left|\frac{1}{2}\left\langle \{A - \langle A\rangle, B - \langle B\rangle\}\right\rangle\right|^2. \tag{3.61}$$

This leads to the inequality

$$\left|\langle (A - \langle A\rangle)(B - \langle B\rangle)\rangle\right|^2 \geq \left|\frac{1}{2}\langle [A, B]\rangle\right|^2. \tag{3.62}$$

Now we shall use the Schwarz inequality,

$$\left|\left\langle a\middle|b\right\rangle\right|^2 \leq \langle a|a\rangle \left\langle b\middle|b\right\rangle, \tag{3.63}$$

which we can prove easily by applying the postulate $\left\langle\phi\middle|\phi\right\rangle \geq 0$ to

$$\left|\phi\right\rangle = |a\rangle - \frac{\langle b|a\rangle}{\langle b|b\rangle}|b\rangle.$$

Applying Schwarz's inequality to

$$|a\rangle = (A - \langle A\rangle)\left|\phi\right\rangle$$

and

$$|b\rangle = (B - \langle B\rangle)\left|\phi\right\rangle,$$

we obtain

$$\left|\langle (A - \langle A\rangle)(B - \langle B\rangle)\rangle\right|^2 \leq \left\langle (A - \langle A\rangle)^2\right\rangle \left\langle (B - \langle B\rangle)^2\right\rangle. \tag{3.64}$$

Combining the inequalities of Eqs. (3.62) and (3.64), we obtain the "uncertainty principle":

$$\left\langle (\Delta A)^2\right\rangle \left\langle (\Delta B)^2\right\rangle \geq \frac{1}{4}\left|\langle [A, B]\rangle\right|^2. \tag{3.65}$$

Remark. This is the Robertson relation (1929). Keeping the anticommutator term instead of dropping it gives the sharper Robertson–Schrödinger inequality $\left\langle (\Delta A)^2\right\rangle \left\langle (\Delta B)^2\right\rangle \geq \frac{1}{4}\left|\langle [A, B]\rangle\right|^2 + \frac{1}{4}\left\langle \{A - \langle A\rangle, B - \langle B\rangle\}\right\rangle^2$. It bounds the spreads of *separate* measurements on identically prepared systems; it is not by itself a statement about the disturbance caused by measuring A on a later measurement of B.

3.11 Change of basis

Incompatible operators offer different sets of base kets to describe the states of a quantum mechanical system (furnishing a different "representation" for each one of such operators).

Given two sets of base kets, there is a unitary operator which can transform general states and operators from one basis to the other. Consider two Hermitian operators A, B and their corresponding sets of eigenstates $\{|a_i\rangle\}, \{|b_i\rangle\}$, where

$$A|a_i\rangle = a_i|a_i\rangle\,, \tag{3.66}$$
$$B|b_i\rangle = b_i|b_i\rangle\,, \tag{3.67}$$

and

$$\langle a_i|a_j\rangle = \langle b_i|b_j\rangle = \delta_{ij}. \tag{3.68}$$

There is a unitary operator U which transforms a state $|a_i\rangle$ into a state $|b_i\rangle$:

$$|b_i\rangle = U|a_i\rangle\,, \tag{3.69}$$

with

$$UU^\dagger = U^\dagger U = 1. \tag{3.70}$$

Explicitly, we can easily verify that

$$U = \sum_k |b_k\rangle\langle a_k|\,. \tag{3.71}$$

Indeed,

$$\begin{aligned} U|a_i\rangle &= \sum_k |b_k\rangle\langle a_k|a_i\rangle = \sum_k |b_k\rangle\,\delta_{ki} \\ &= |b_i\rangle\,. \end{aligned} \tag{3.72}$$

Similarly,

$$\begin{aligned} U^\dagger|b_i\rangle &= \sum_k \left(|b_k\rangle\langle a_k|\right)^\dagger |b_i\rangle = \sum_k \left(|a_k\rangle\langle b_k|\right)|b_i\rangle = \sum_k |a_k\rangle\,\delta_{ki} \\ &= |a_i\rangle\,. \end{aligned} \tag{3.73}$$

Finally,

$$\begin{aligned} U^\dagger U &= \left(\sum_j |a_j\rangle\langle b_j|\right)\left(\sum_k |b_k\rangle\langle a_k|\right) = \sum_{jk} |a_j\rangle\langle b_j|b_k\rangle\langle a_k| \\ &= \sum_{jk} |a_j\rangle\,\delta_{jk}\langle a_k| = \sum_j |a_j\rangle\langle a_j| = 1\,, \end{aligned} \tag{3.74}$$

and

$$\begin{aligned} UU^\dagger &= \left(\sum_j |b_j\rangle\langle a_j|\right)\left(\sum_k |a_k\rangle\langle b_k|\right) = \sum_{jk} |b_j\rangle\langle a_j|a_k\rangle\langle b_k| \\ &= \sum_{jk} |b_j\rangle\,\delta_{jk}\langle b_k| = \sum_j |b_j\rangle\langle b_j| = 1\,. \end{aligned} \tag{3.75}$$

3.11.1 Transformation matrix

The transformation matrix

$$U = \sum_k |b_k\rangle\langle a_k| \tag{3.76}$$

is represented in the basis of $\{|a_i\rangle\}$ as

$$\langle a_j| U |a_i\rangle = \sum_k \langle a_j|b_k\rangle \langle a_k|a_i\rangle = \sum_k \langle a_j|b_k\rangle \delta_{ki} = \langle a_j|b_i\rangle . \tag{3.77}$$

An arbitrary state $|\phi\rangle$ is written as

$$|\phi\rangle = \sum_k c_k |a_k\rangle = \sum_k \langle a_k|\phi\rangle |a_k\rangle \doteq \begin{pmatrix} \langle a_1|\phi\rangle \\ \langle a_2|\phi\rangle \\ \vdots \\ \langle a_N|\phi\rangle \end{pmatrix}, \tag{3.78}$$

so in the $\{|a_i\rangle\}$ basis it is represented as a column vector with elements $\langle a_i|\phi\rangle$. Similarly, in the $\{|b_i\rangle\}$ basis it is represented as a column vector with elements $\langle b_i|\phi\rangle$:

$$|\phi\rangle = \sum_k d_k |b_k\rangle = \sum_k \langle b_k|\phi\rangle |b_k\rangle \doteq \begin{pmatrix} \langle b_1|\phi\rangle \\ \langle b_2|\phi\rangle \\ \vdots \\ \langle b_N|\phi\rangle \end{pmatrix}. \tag{3.79}$$

The two column vectors are related as follows:

$$\begin{aligned} \langle b_i|\phi\rangle &= \sum_k \langle b_i|a_k\rangle \langle a_k|\phi\rangle \\ &= \sum_k \langle a_i| U^\dagger |a_k\rangle \langle a_k|\phi\rangle . \end{aligned} \tag{3.80}$$

In matrix form,

$$\underbrace{\begin{pmatrix} \vdots \\ \langle b_i|\phi\rangle \\ \vdots \end{pmatrix}}_{\text{new basis}} = \underbrace{\begin{pmatrix} \ddots & \vdots & \iddots \\ \dots & \langle a_i| U^\dagger |a_k\rangle & \dots \\ \iddots & \vdots & \ddots \end{pmatrix}}_{\langle a_i|U^\dagger|a_k\rangle \text{ at } i\text{th row and } k\text{th column}} \underbrace{\begin{pmatrix} \vdots \\ \langle a_k|\phi\rangle \\ \vdots \end{pmatrix}}_{\text{old basis}}. \tag{3.81}$$

Let us now look at changing representations for operators. In the two bases, $\{|a_i\rangle\}$ and $\{|b_i\rangle\}$, an operator X is represented by the matrices

$$\langle a_j| X |a_i\rangle \quad \text{and} \quad \langle b_j| X |b_i\rangle ,$$

respectively. We write:

$$\begin{aligned}\langle b_j| X |b_i\rangle &= \sum_{k\ell} \langle b_j|a_k\rangle \langle a_k| X |a_\ell\rangle \langle a_\ell|b_i\rangle \\ &= \sum_{k\ell} \langle a_j| U^\dagger |a_k\rangle \langle a_k| X |a_\ell\rangle \langle a_\ell| U |a_i\rangle ,\end{aligned} \tag{3.82}$$

which, in matrix notation, is written as[4]

$$\underbrace{\begin{pmatrix} \langle b_j| X |b_i\rangle \end{pmatrix}}_{\text{new basis}} = \begin{pmatrix} \langle a_j| U^\dagger |a_k\rangle \end{pmatrix} \underbrace{\begin{pmatrix} \langle a_k| X |a_\ell\rangle \end{pmatrix}}_{\text{old basis}} \begin{pmatrix} \langle a_\ell| U |a_i\rangle \end{pmatrix}. \tag{3.83}$$

3.11.2 Trace of operators

The trace of an operator is defined as

$$\mathrm{tr}(X) = \sum_i \langle a_i| X |a_i\rangle , \tag{3.84}$$

and it is independent of the representation. Indeed, for two representations $\{|a_i\rangle\}$, $\{|b_i\rangle\}$, we have:

$$\begin{aligned}\sum_i \langle a_i| X |a_i\rangle &= \sum_{ijk} \langle a_i|b_j\rangle \langle b_j| X |b_k\rangle \langle b_k|a_i\rangle \\ &= \sum_{jk} \langle b_k| \underbrace{\left(\sum_i |a_i\rangle\langle a_i| \right)}_{1} |b_j\rangle \langle b_j| X |b_k\rangle \\ &= \sum_{jk} \langle b_k|b_j\rangle \langle b_j| X |b_k\rangle \\ &= \sum_{jk} \delta_{jk} \langle b_j| X |b_k\rangle \\ &= \sum_i \langle b_i| X |b_i\rangle .\end{aligned}$$

It is easy to prove the familiar properties of traces, such as

$$\mathrm{tr}(XY) = \mathrm{tr}(YX). \tag{3.85}$$

We also note that the trace of an outer product is an inner product,

$$\mathrm{tr}\left(|c\rangle\langle b|\right) = \sum_i \langle a_i|c\rangle \langle b|a_i\rangle = \langle b| \left(\sum_i |a_i\rangle\langle a_i| \right) |c\rangle = \langle b|c\rangle . \tag{3.86}$$

[4] With the same conventions as in Eq. (3.81).

3.12 Eigenstates and eigenvalues

Suppose that we know the representation $\langle a_i | X | a_j \rangle$ of an operator X in a basis $\{|a_i\rangle\}$. We would like to compute the eigenstates of X in the same representation. We have

$$\begin{aligned} X|b_i\rangle &= b_i |b_i\rangle \\ \rightsquigarrow X \sum_l |a_l\rangle\langle a_l| b_i\rangle &= b_i |b_i\rangle \\ \rightsquigarrow \langle a_\lambda | X \sum_l |a_l\rangle\langle a_l| b_i\rangle &= b_i \langle a_\lambda | b_i\rangle \\ \rightsquigarrow \sum_l \langle a_\lambda | X |a_l\rangle \langle a_l | b_i\rangle &= b_i \langle a_\lambda | b_i\rangle \end{aligned}$$

which, in matrix notation, is cast as

$$\begin{pmatrix} \langle a_1|X|a_1\rangle & \dots & \langle a_1|X|a_N\rangle \\ \vdots & \ddots & \vdots \\ \langle a_N|X|a_1\rangle & \dots & \langle a_N|X|a_N\rangle \end{pmatrix} \begin{pmatrix} \langle a_1|b_i\rangle \\ \vdots \\ \langle a_N|b_i\rangle \end{pmatrix} = b_i \begin{pmatrix} \langle a_1|b_i\rangle \\ \vdots \\ \langle a_N|b_i\rangle \end{pmatrix}. \tag{3.87}$$

We can find the eigenvalues b_i and eigenstates

$$\begin{pmatrix} \langle a_1|b_i\rangle \\ \vdots \\ \langle a_N|b_i\rangle \end{pmatrix} \tag{3.88}$$

as the eigenvalues and eigenvectors of the matrix on the left-hand side of Eq. (3.87). The eigenvalues satisfy

$$\det(X - b_i \mathbf{1}) = 0. \tag{3.89}$$

Knowing the eigenvalues, we can construct the eigenstates as usual in linear algebra, by substituting b_i explicitly in Eq. (3.87) and solving for the $\langle a_l|b_i\rangle$'s.

3.13 Unitarily equivalent observables

Two operators A, B are equivalent if they can be related by a unitary transformation,

$$B = UAU^\dagger, \qquad U^\dagger = U^{-1}. \tag{3.90}$$

These operators have the same eigenvalues and their eigenstates are related by the same unitary transformation:

$$\begin{aligned} A|a\rangle &= a|a\rangle \\ \rightsquigarrow A \underbrace{U^\dagger U}_{1} |a\rangle &= a \underbrace{U^\dagger U}_{1} |a\rangle \\ \rightsquigarrow \left(UAU^\dagger\right)(U|a\rangle) &= aUU^\dagger U|a\rangle \\ \rightsquigarrow B(U|a\rangle) &= a(U|a\rangle). \end{aligned} \tag{3.91}$$

The operator B satisfies an eigenvalue equation,

$$B\,|b\rangle = b\,|b\rangle\,.$$

Comparing with the above, we conclude that the eigenstates of the operator B are

$$|b\rangle = U\,|a\rangle\,,$$

and the eigenvalues are

$$b = a\,.$$

4

Time evolution

Before we introduce time evolution in quantum mechanics, it will be useful to recall how we describe the dynamics of simple mechanical systems in classical physics. In quantum mechanics, we will postulate principles that extend this picture of classical mechanics.

4.1 Time evolution in classical mechanics

Consider, for simplicity, a one-dimensional mechanical system whose dynamical behaviour is encoded in a Lagrangian

$$L\left(x(t), \dot{x}(t)\right).$$

The Lagrangian depends on the position and the velocity, which are functions of time. We will focus on energy conserving systems, in which the Lagrangian acquires all of its time dependence through these functions and has no other explicit time dependence,

$$\frac{\partial L}{\partial t} = 0. \tag{4.1}$$

The time evolution of the position and the velocity obeys the principle of least action. This states that the action is stationary,

$$\delta S = 0\,, \tag{4.2}$$

under small variations of the physical trajectory

$$x(t) \longrightarrow x(t) + \delta x(t). \tag{4.3}$$

Eq. (4.2) leads to the Euler–Lagrange differential equation

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{x}} - \frac{\partial L}{\partial x} = 0. \tag{4.4}$$

For example, a particle moving in one dimension under the influence of a force potential $V(x)$ has a Lagrangian

$$L = \frac{1}{2}m\dot{x}^2 - V(x), \tag{4.5}$$

and the corresponding Euler–Lagrange equation yields Newton's law,

$$m\ddot{x} + \frac{\partial V}{\partial x} = 0. \tag{4.6}$$

An alternative description of the same dynamics can be obtained through the Hamiltonian formalism. We define first the canonical momentum,

$$p = \frac{\partial L}{\partial \dot{x}}, \tag{4.7}$$

which, in general, is a function of both the position and velocity,

$$p = p\left(x(t), \dot{x}(t)\right). \tag{4.8}$$

We now define the Hamiltonian

$$H(x, p) = p\dot{x} - L, \tag{4.9}$$

and pick the pair (x, p) as our independent variables, "inverting" Eq. (4.8), and considering the velocity as a function of position and momentum,

$$\dot{x} = \dot{x}\left(x, p\right). \tag{4.10}$$

Differentiating the Hamiltonian with respect to the position, we have

$$\frac{\partial H}{\partial x} = \ldots = -\dot{p}, \tag{4.11}$$

while differentiating with respect to the momentum we arrive at

$$\frac{\partial H}{\partial p} = \ldots = \dot{x}. \tag{4.12}$$

Therefore, as an alternative to the second-order Euler–Lagrange differential equation, we can describe the dynamical system with the pair of first-order differential equations of Hamilton,

$$\dot{p} = -\frac{\partial H}{\partial x}, \quad \dot{x} = \frac{\partial H}{\partial p}. \tag{4.13}$$

Let us now consider a generic physical quantity,

$$A = A(x, p, t). \tag{4.14}$$

Differentiating with respect to time, we have

$$\frac{dA}{dt} = \frac{\partial A}{\partial p}\dot{p} + \frac{\partial A}{\partial x}\dot{x} + \frac{\partial A}{\partial t}. \tag{4.15}$$

Substituting Hamilton's equations, Eq. (4.13), we obtain

$$\frac{dA}{dt} = \{A, H\} + \frac{\partial A}{\partial t}, \tag{4.16}$$

where the Poisson bracket for two functions $f(x, p)$ and $g(x, p)$ is given by

$$\{f, g\} \equiv \frac{\partial f}{\partial x}\frac{\partial g}{\partial p} - \frac{\partial f}{\partial p}\frac{\partial g}{\partial x}. \tag{4.17}$$

For example, the canonical momentum for a particle in a potential, corresponding to the Lagrangian of Eq. (4.5), is

$$p = \frac{\partial L}{\partial \dot{x}} = m\dot{x}, \tag{4.18}$$

which we invert to read

$$\dot{x} = \frac{p}{m}. \tag{4.19}$$

The Hamiltonian is then computed to be

$$H = p\dot{x} - L = \frac{p^2}{2m} + V(x), \tag{4.20}$$

and Hamilton's equations are

$$\dot{x} = \frac{p}{m}\,, \quad \dot{p} = -\frac{\partial V}{\partial x}, \tag{4.21}$$

and the general evolution equation reads

$$\frac{dA}{dt} = \left\{A, \frac{p^2}{2m} + V(x)\right\} + \frac{\partial A}{\partial t}\,. \tag{4.22}$$

We will soon study the theoretical implications of this time-evolution equation for generic physical quantities A and of the Poisson brackets.

4.1.1 Properties of Poisson brackets

We can derive the following properties of Poisson brackets from their definition in Eq. (4.17).

- Antisymmetry,

$$\{A, B\} = -\{B, A\}\,, \tag{4.23}$$

- Linearity,

$$\{A, B + C\} = \{A, B\} + \{A, C\}\,, \tag{4.24}$$

- Product rule,

$$\{A, B\,C\} = B\ \{A, C\} + \{A, B\}\ C, \tag{4.25}$$

- Action on position and momentum,

$$\{x, p\} = 1\,, \quad \{x, x\} = 0\,, \quad \{p, p\} = 0\,. \tag{4.26}$$

In addition, Poisson brackets satisfy the Jacobi identity,

$$\{A, \{B, C\}\} + \{B, \{C, A\}\} + \{C, \{A, B\}\} = 0. \tag{4.27}$$

Conversely, for the class of functions of position and momentum which have a series representation,

$$f(x,p) = \sum_{n,m=0}^{\infty} c_{nm} x^n p^m, \tag{4.28}$$

the properties of Eqs. (4.23)–(4.26) lead back to the definition of Eq. (4.17). The proof is somewhat lengthy but relies on simple induction.

We first prove inductively that

$$\{x, x^m\} = 0, \quad \{p, p^m\} = 0, \tag{4.29}$$

and

$$\{x, p^m\} = m p^{m-1} = \frac{\partial}{\partial p} p^m, \quad \{x^m, p\} = m x^{m-1} = \frac{\partial}{\partial x} x^m. \tag{4.30}$$

Then, it follows that for a generic function $f(x,p)$, as in Eq. (4.28), the Poisson bracket with respect to position (momentum) is equivalent to the derivative with respect to momentum (position),

$$\{x, f\} = \frac{\partial f}{\partial p}, \quad \{f, p\} = \frac{\partial f}{\partial x}. \tag{4.31}$$

Next, we can show inductively that

$$\begin{aligned} \{x^n, f(x,p)\} &= n x^{n-1} \frac{\partial f}{\partial p} = \frac{\partial x^n}{\partial x} \frac{\partial f}{\partial p}, \\ \{f(x,p), p^n\} &= n p^{n-1} \frac{\partial f}{\partial x} = \frac{\partial p^n}{\partial p} \frac{\partial f}{\partial x}. \end{aligned} \tag{4.32}$$

In Eq. (4.32) and in the antisymmetry, Eq. (4.23), linearity, Eq. (4.24), and product rule, Eq. (4.25), of the Poisson brackets we have all the ingredients needed to return to Eq. (4.17):

$$\begin{aligned} \{f(x,p), g(x,p)\} &= \left\{ \sum_{nm} c_{nm} x^n p^m, g(x,p) \right\} = \sum_{nm} c_{nm} \{x^n p^m, g\} \\ &= \sum_{nm} c_{nm}\, x^n \{p^m, g\} + \sum_{nm} c_{nm} \{x^n, g\}\, p^m \\ &= \sum_{nm} c_{nm} \left[-x^n \{g, p^m\} + \{x^n, g\}\, p^m \right] \\ &= \sum_{nm} c_{nm} \left[-x^n \frac{\partial p^m}{\partial p} \frac{\partial g}{\partial x} + \frac{\partial x^n}{\partial x} p^m \frac{\partial g}{\partial p} \right] \\ &= \frac{\partial f}{\partial x} \frac{\partial g}{\partial p} - \frac{\partial f}{\partial p} \frac{\partial g}{\partial x}. \end{aligned} \tag{4.33}$$

4.1.2 A way to think of classical time evolution

We have thus shown that the definition of Poisson brackets in Eq. (4.17) is equivalent to the properties of Eqs. (4.23)–(4.26). We can then choose to formulate time evolution as follows.

The dynamics of a physical system is encoded in its Hamiltonian,

$$H(x_i, p_i),$$

which depends on space coordinates x_i and canonical momenta p_i. A physical quantity $A(x_i, p_i, t)$ evolves in time according to

$$\frac{dA}{dt} = \{A, H\} + \frac{\partial A}{\partial t}, \tag{4.34}$$

and the Poisson brackets are defined to satisfy

$$\{x_i, x_j\} = \{p_i, p_j\} = 0, \quad \{x_i, p_j\} = \delta_{ij}, \tag{4.35}$$

as well as antisymmetry, linearity, and the product rule of Eqs. (4.23)–(4.25). These statements are all that one needs to describe the dynamical evolution of a classical system.

4.2 Time evolution in quantum mechanics

Physical systems are in quantum states $|\psi\rangle$ of a Hilbert space. Physical quantities, such as the position and momentum, are Hermitian operators acting on this space,

$$\hat{x}_i\,, \hat{p}_i\,, \ldots$$

The mean value of experimental measurements of an observable quantity $O(x_i, p_i)$, which is represented by an operator $\hat{O}$ in Hilbert space, is given by

$$\langle O \rangle = \langle \psi | \hat{O}(\hat{x}_i, \hat{p}_i) | \psi \rangle\,. \tag{4.36}$$

Quantum time evolution has a form very similar to that of classical evolution, as presented in Subsection 4.1.2. A central object in classical evolution is the Poisson bracket of position and momentum. What is the analogous object for the corresponding quantum mechanical operators? Dirac observed that the commutator of two operators,

$$[A, B] \equiv AB - BA,$$

possesses algebraic properties analogous to those of the classical Poisson brackets, i.e.

$$[A, B] = -[B, A]\,, \tag{4.37}$$
$$[A, B + C] = [A, B] + [A, C]\,, \tag{4.38}$$
$$[A, BC] = B\,[A, C] + [A, B]\,C\,, \tag{4.39}$$
$$[A, [B, C]] + [B, [C, A]] + [C, [A, B]] = 0\,. \tag{4.40}$$

In classical mechanics, the Poisson bracket of a coordinate and its canonical momentum is a constant (unity). We will postulate the same for the commutator of their quantum mechanical operators,

$$\left[\hat{x}, \hat{p}\right] = \text{constant}.$$

The constant must then be imaginary, given that position and momentum are Hermitian,

$$\text{constant}^* = \left[\hat{x}, \hat{p}\right]^\dagger = \left[\hat{p}^\dagger, \hat{x}^\dagger\right] = -\left[\hat{x}, \hat{p}\right] = -\text{constant}. \tag{4.41}$$

In analogy to classical mechanics, it is natural to postulate the following commutation relations for positions and momenta,

$$\left[\hat{x}_i, \hat{x}_j\right] = \left[\hat{p}_i, \hat{p}_j\right] = 0\,, \quad \left[\hat{x}_i, \hat{p}_j\right] = i\hbar\delta_{ij}\,. \tag{4.42}$$

Comparing Eq. (4.42) with the Poisson-bracket "postulates" of classical mechanics in Eq. (4.35), we are led to identify the quantum mechanical analogue of the Poisson bracket with the commutator,

$$\{A, B\} \to \frac{1}{i\hbar}\left[\hat{A}, \hat{B}\right]. \tag{4.43}$$

Inspired by this analogy and Eq. (4.34), it seems natural to postulate the following time evolution of the operators of physical quantities, $\hat{O}\left(\hat{x}_i, \hat{p}_i\right)$, in quantum mechanics,

$$\frac{d\hat{O}}{dt} = \frac{1}{i\hbar}\left[\hat{O}, \hat{H}\right], \tag{4.44}$$

where $\hat{H}$ is the Hamiltonian operator. Eq. (4.44) has the solution,

$$\hat{O}(t) = e^{\frac{i}{\hbar}Ht}\hat{O}_0 e^{-\frac{i}{\hbar}Ht}. \tag{4.45}$$

4.2.1 Ehrenfest theorem

In the above, in complete analogy to their classical mechanics counterparts, we have endowed quantum mechanical operators with a time dependence. However, quantum operators themselves are not measured experimentally. What is measured are their expectation values in quantum states of the Hilbert space. For such a state, $\left|\psi_0\right\rangle$, the average value is

$$\left\langle \hat{O}(t) \right\rangle = \left\langle \psi_0 \right| \hat{O}(t) \left| \psi_0 \right\rangle. \tag{4.46}$$

The time evolution of the measured average value of the observable satisfies the differential equation

$$\begin{aligned} \frac{d}{dt}\left\langle \hat{O}(t) \right\rangle &= \left\langle \psi_0 \right| \frac{d}{dt}\hat{O}(t) \left| \psi_0 \right\rangle \\ &= \frac{1}{i\hbar}\left\langle \psi_0 \right| \left[\hat{O}, \hat{H}\right] \left| \psi_0 \right\rangle . \end{aligned} \tag{4.47}$$

Let us concretely consider the Hamiltonian of a particle in a potential,

$$H = \frac{\hat{p}^2}{2M} + V\left(\hat{x}\right), \tag{4.48}$$

and specialise to $\hat{O} = \hat{x}$ and $\hat{O} = \hat{p}$. Then, from Eq. (4.47) and the commutation relations

$$[\hat{x}(t), \hat{x}(t)] = \left[\hat{p}(t), \hat{p}(t)\right] = 0, \quad \left[\hat{x}(t), \hat{p}(t)\right] = i\hbar, \tag{4.49}$$

we obtain (**Exercise**),

$$\frac{d}{dt}\left\langle \hat{x}(t) \right\rangle = \frac{\left\langle \hat{p}(t) \right\rangle}{M}, \tag{4.50}$$

$$\frac{d}{dt}\left\langle \hat{p}(t) \right\rangle = -\left\langle V'\left(\hat{x}(t)\right) \right\rangle. \tag{4.51}$$

These equations are analogous to Hamilton's equations of classical mechanics. As in classical mechanics, we can combine them into a second-order differential equation, yielding

$$M\frac{d^2}{dt^2}\langle\hat{x}(t)\rangle = -\langle V'(\hat{x}(t))\rangle\,. \tag{4.52}$$

The result of Eq. (4.52) is Ehrenfest's theorem and generalises Newton's law to quantum mechanics.

While Eq. (4.52) is similar in form to the classical law of Newton, its solution for the quantum average value of the trajectory, $\langle\hat{x}(t)\rangle$, can be qualitatively and quantitatively very different from the classical trajectory. This is because, in general,

$$\langle V'(\hat{x})\rangle \neq V'(\langle\hat{x}\rangle)\,. \tag{4.53}$$

4.2.2 Heisenberg and Schrödinger picture

So far, we have endowed the physical quantum operators with time evolution while we took the quantum states $|\psi_0\rangle$ to be time independent. This description of time evolution is known as the *Heisenberg picture*, but it is not unique. Eq. (4.46) allows further possibilities. The Heisenberg picture will be more convenient for studying time evolution in quantum field theory. An equally useful description, which is commonly used, is known as the Schrödinger picture.

In the Schrödinger picture, operators are taken to be constant in time and time evolution is carried by the states. Indeed, we can rearrange the right-hand side of Eq. (4.46) shifting time evolution from the operator to the quantum states,

$$\langle\psi_0|\,\hat{O}(t)\,|\psi_0\rangle = \langle\psi(t)|\,\hat{O}_0\,|\psi(t)\rangle\,, \tag{4.54}$$

where

$$|\psi(t)\rangle = e^{-\frac{i}{\hbar}Ht}\,|\psi_0\rangle\,. \tag{4.55}$$

This shift of picture does not affect the prediction for the physically measured value of the observable; the Heisenberg and Schrödinger pictures are physically equivalent. It is easy to see, by differentiating the above equation with respect to time, that the time-dependent states (Schrödinger picture) satisfy Schrödinger's equation,

$$i\hbar\frac{\partial}{\partial t}|\psi(t)\rangle = \hat{H}\,|\psi(t)\rangle\,. \tag{4.56}$$

4.3 Time evolution in the Schrödinger picture

In the previous sections, we followed Dirac's reasoning and guessed the time evolution of a class of observable operators by extending the classical mechanics formalism of Poisson brackets. Here, we will revisit time evolution in the Schrödinger picture, in which the time evolution is carried by the states and not by the operators. We will realise that we can infer fundamental properties of time evolution from the fact that quantum states are vectors of a Hilbert space. Our description of time evolution needs to be consistent with the mathematical language of kets, bras and operators upon which we have formulated

the principles of quantum mechanics. This language is very restrictive and leads to an elegant and austere formalism for time evolution. Consider a state which evolves as

$$\left|\phi, t_0\right\rangle \rightarrow \left|\phi, t\right\rangle, \qquad t \geq t_0. \tag{4.57}$$

We introduce a time evolution operator and relate the two states at times t_0 and t as

$$\left|\phi, t\right\rangle = U(t - t_0)\left|\phi, t_0\right\rangle. \tag{4.58}$$

So far, we have only defined the purpose and the name of the time evolution operator $U(t - t_0)$. We can easily guess some of its properties.

1. Let us consider a normalised state $\left|\phi, t_0\right\rangle$, satisfying

$$\left\langle\phi, t_0\middle|\phi, t_0\right\rangle = 1. \tag{4.59}$$

   We require that time evolution does not destroy the normalisation of the state,

$$\left\langle\phi, t\middle|\phi, t\right\rangle = 1. \tag{4.60}$$

   This is satisfied if the time evolution operator is unitary,

$$U^\dagger(t - t_0)U(t - t_0) = 1, \quad t > t_0\,. \tag{4.61}$$

   We will take Eq. (4.61) to hold as one of our starting assumptions.

2. We also expect that a time evolution $t_0 \rightarrow t_1 > t_0$ followed by a time evolution $t_1 \rightarrow t_2 > t_1$ is equivalent to a time evolution $t_0 \rightarrow t_2$,

$$U(t_2 - t_1)U(t_1 - t_0) = U(t_2 - t_0). \tag{4.62}$$

3. Finally, we anticipate that

$$\lim_{t \rightarrow t_0} U(t - t_0) = 1. \tag{4.63}$$

For small time intervals, all of the above are satisfied if

$$U(\Delta t) = 1 - i\Omega\Delta t + O((\Delta t)^2), \tag{4.64}$$

$$\Omega^\dagger = \Omega \tag{4.65}$$

.

with

We postulate that

$$\Omega = \frac{H}{\hbar}, \tag{4.66}$$

where H is the Hamiltonian operator. This is inspired by classical mechanics, as in our earlier discussion.

4.4 Schrödinger equation

We start from the composition property for a time translation in the interval

$$(t_0 = 0, t + \Delta t) \tag{4.67}$$

with Δt being infinitesimal.

$$\begin{aligned} U(t+\Delta t) &= U(\Delta t)U(t) \\ \rightsquigarrow U(t+\Delta t) &= \left(1 - i\frac{H}{\hbar}\Delta t\right) U(t) \\ \rightsquigarrow U(t) + \Delta t \frac{\partial}{\partial t} U(t) &= \left(1 - i\frac{H}{\hbar}\Delta t\right) U(t) \\ \overset{\cdot i\hbar/\Delta t}{\rightsquigarrow}\ i\hbar\frac{\partial}{\partial t} U(t) &= HU(t). \end{aligned} \tag{4.68}$$

Acting with the operators on the left- and right-hand sides of the last equation on a state at the starting time $t_0 = 0$, we find

$$i\hbar\frac{\partial}{\partial t} U(t) \left|\phi, t_0 = 0\right\rangle = HU(t) \left|\phi, t_0 = 0\right\rangle . \tag{4.69}$$

Recalling that

$$U(t) \left|\phi, t_0 = 0\right\rangle = \left|\phi, t\right\rangle , \tag{4.70}$$

we obtain Schrödinger's equation,

$$i\hbar\frac{\partial}{\partial t} \left|\phi, t\right\rangle = H \left|\phi, t\right\rangle . \tag{4.71}$$

We see that the only axiom in Schrödinger's equation is to identify the operator H with the Hamiltonian. The form of the equation and the Hermiticity of H are derived from demanding states to be vectors in Hilbert space and that time evolution is a unitary operator, Eq. (4.61). Notice also that our derivation is based on examining time evolution in infinitesimally small intervals, within which the Hamiltonian can always be assumed to be constant.[1] Thus, we can take Schrödinger's equation to be valid for time dependent Hamiltonians $H(t)$. **Remark.** For a time-dependent $H(t)$ the evolution operator is no longer a function of $t - t_0$ alone; one must write $U(t, t_0)$ with the composition law $U(t_2, t_1)U(t_1, t_0) = U(t_2, t_0)$, and the derivation above goes through with $U(t + \Delta t, t_0) = \left(1 - \frac{i}{\hbar}H(t)\Delta t\right) U(t, t_0)$. The properties listed for $U(t - t_0)$ assume time-translation invariance, i.e. a constant H (Sakurai, Sec. 2.1.2).

In the familiar case of a constant Hamiltonian, $H(t) = H$, the solution of Schrödinger's equation is

$$\left|\phi, t\right\rangle = e^{-\frac{i}{\hbar}(t-t_0)\cdot H} \left|\phi, t_0\right\rangle , \tag{4.72}$$

[1] If not, make Δt smaller.

which we also found in the Schrödinger picture above. For a time-dependent Hamiltonian with

$$[H(t_1), H(t_2)] = 0, \quad \forall\, t_1, t_2 \in \underbrace{[t_0, t]}_{\text{interval}}, \tag{4.73}$$

we can prove that

$$\left|\phi, t\right\rangle = e^{-\frac{i}{\hbar}\int_{t_0}^{t} \mathrm{d}t' H(t')} \left|\phi, t_0\right\rangle . \tag{4.74}$$

The exponential of an operator X, such as the Hamiltonian above, is defined through its series expansion,

$$e^X = 1 + X + \frac{X^2}{2!} + \ldots \tag{4.75}$$

The generic solution of Schrödinger's equation can be formally written as a "time-ordered exponential", which we will present in Part II of this script and which is also central in quantum field theory.

4.5 Eigenstates of the Hamiltonian operator and time evolution

Consider a constant Hamiltonian operator H with eigenstates $|n\rangle$ and eigenvalues E_n,

$$H\left|n\right\rangle = E_n \left|n\right\rangle . \tag{4.76}$$

A general state at an initial time t_0 is written as

$$\left|\psi, t_0\right\rangle = \left(\sum_n \left|n\right\rangle\!\left\langle n\right|\right) \left|\psi, t_0\right\rangle \ . \tag{4.77}$$

Acting with the time evolution operator we obtain

$$\begin{aligned} U(t-t_0)\left|\psi, t_0\right\rangle &= U(t-t_0)\left(\sum_n \left|n\right\rangle\!\left\langle n\right|\right)\left|\psi, t_0\right\rangle \\ \rightsquigarrow \left|\psi, t\right\rangle &= e^{-\frac{i}{\hbar}(t-t_0)\cdot H}\left(\sum_n \left|n\right\rangle\!\left\langle n\right|\right)\left|\psi, t_0\right\rangle \\ \rightsquigarrow \left|\psi, t\right\rangle &= \sum_n \left|n\right\rangle e^{-\frac{i}{\hbar}(t-t_0)\cdot E_n}\left\langle n\middle|\psi, t_0\right\rangle . \end{aligned} \tag{4.78}$$

For the special situation where the initial state is also an eigenstate,

$$\left|\psi, t_0\right\rangle = \left|m\right\rangle ,$$

we obtain

$$\begin{aligned} \left|\psi, t\right\rangle &= \sum_n \left|n\right\rangle e^{-\frac{i}{\hbar}(t-t_0)\cdot E_n}\left\langle n|m\right\rangle \\ \rightsquigarrow \left|\psi, t\right\rangle &= \sum_n \left|n\right\rangle e^{-\frac{i}{\hbar}(t-t_0)\cdot E_n}\delta_{nm} \\ \\ \rightsquigarrow \left|\psi, t\right\rangle &= \left|m\right\rangle e^{-\frac{i}{\hbar}(t-t_0)\cdot E_m} . \end{aligned} \tag{4.79}$$

The time-evolved state is the original eigenstate up to a phase factor, which is unimportant for physics. **Once a physical system is in an eigenstate of the Hamiltonian, it will always remain in this eigenstate.**

Assuming again that the system is in an eigenstate of the Hamiltonian, the expectation value of a physical observable A at a later time t will be

$$\begin{aligned}\langle A\rangle_t &= \left\langle \phi, t\right| A \left|\phi, t\right\rangle = \langle m| e^{+\frac{i}{\hbar}(t-t_0)\cdot H} A e^{-\frac{i}{\hbar}(t-t_0)\cdot H} |m\rangle \\ &= \langle m| e^{+\frac{i}{\hbar}(t-t_0)\cdot E_m} A e^{-\frac{i}{\hbar}(t-t_0)\cdot E_m} |m\rangle \\ &= \langle m| A |m\rangle \ . \end{aligned} \tag{4.80}$$

Therefore, **expectation values of physical observables do not change with time if the system is in an eigenstate of the Hamiltonian.**

Otherwise, we anticipate that expectation values "oscillate" among various values with time. For an initial state (at $t_0 = 0$)

$$\left|\phi\right\rangle = \sum_m |m\rangle \left\langle m\middle|\phi\right\rangle ,$$

the expectation value at a later time becomes

$$\begin{aligned}\langle A\rangle_t &= \left\langle \phi, t\right| A \left|\phi, t\right\rangle \\ &= \left\langle \phi\right| e^{+\frac{i}{\hbar}(t-t_0)\cdot H} A e^{-\frac{i}{\hbar}(t-t_0)\cdot H} \left|\phi\right\rangle \\ &= \sum_{n,m} \left\langle \phi\middle|m\right\rangle \langle m| e^{+\frac{i}{\hbar}(t-t_0)\cdot H} A e^{-\frac{i}{\hbar}(t-t_0)\cdot H} |n\rangle \left\langle n\middle|\phi\right\rangle \\ &= \sum_{n,m} \left\langle \phi\middle|m\right\rangle \langle m| e^{+\frac{i}{\hbar}(t-t_0)\cdot E_m} A e^{-\frac{i}{\hbar}(t-t_0)\cdot E_n} |n\rangle \left\langle n\middle|\phi\right\rangle \\ &= \sum_{n,m} e^{-i(t-t_0)\cdot\omega_{nm}} \left\langle m\middle|\phi\right\rangle^* \left\langle n\middle|\phi\right\rangle \langle m| A |n\rangle , \end{aligned} \tag{4.81}$$

where the "oscillation" frequencies are

$$\omega_{nm} = \frac{E_n - E_m}{\hbar}. \tag{4.82}$$

5

Two-state systems

In this chapter, we will study the evolution with time of simple quantum mechanical systems. The simplest situation occurs when the Hilbert space has a basis with only two states. For example, a particle with spin-$\frac{1}{2}$ can be found in two states, a spin "up" and a spin "down" state. We will see that all two-state systems can be described with the same formalism.

5.1 Spin precession

Consider an electron, which is a spin-$\frac{1}{2}$ particle, inside a magnetic field. The potential energy of a classical magnetic dipole is given by

$$U = -\boldsymbol{\mu} \cdot \mathbf{B},$$

where $\boldsymbol{\mu}$ is the magnetic dipole moment, proportional to the angular momentum of the rotating electric charge. Inspired by this, we define a quantum mechanical Hamiltonian operator

$$H = -\left(\frac{e}{m_e c}\right) \mathbf{S} \cdot \mathbf{B}, \tag{5.1}$$

where $\mathbf{S}$ is the spin operator and $\mathbf{B}$ is the magnetic field. **Remark.** A classical rotating charge has $\boldsymbol{\mu} = \frac{e}{2m_e c}\mathbf{L}$; the coefficient $\frac{e}{m_e c}$ used here corresponds to the electron gyromagnetic factor $g = 2$ (Dirac theory; $g - 2 \approx 0.0023$ from QED), i.e. $\boldsymbol{\mu} = g\frac{e}{2m_e c}\mathbf{S}$, Gaussian units, $e = -|e|$ for the electron (Sakurai, Sec. 2.1.4). The classical formula therefore fixes only the form of H; the factor of 2 is an experimental (and relativistic) input. Let us take the magnetic field to be along the $\hat{z}$ direction,

$$\mathbf{B} = B\hat{z}.$$

Thus,

$$H = -\left(\frac{eB}{m_e c}\right) S_z \equiv \omega S_z, \tag{5.2}$$

with

$$\omega \equiv \frac{|e|B}{m_e c}. \tag{5.3}$$

The Hamiltonian and the operator of spin in the z-direction, S_z, commute. Therefore, they must have a common set of eigenstates. The spin operator has two eigenstates,

$$S_z \left|\pm\right\rangle = \pm\frac{\hbar}{2} \left|\pm\right\rangle, \tag{5.4}$$

which are also energy eigenstates with eigenvalues

$$H|\pm\rangle = E_\pm|\pm\rangle, \quad E_\pm = \pm\frac{\hbar\omega}{2}. \tag{5.5}$$

The time evolution operator is

$$U(t) = e^{-\frac{iHt}{\hbar}} = e^{-\frac{i\omega t S_z}{\hbar}}. \tag{5.6}$$

Acting on an initial state,

$$|\phi\rangle = c_+|+\rangle + c_-|-\rangle, \tag{5.7}$$

we obtain

$$|\phi, t\rangle = c_+ e^{-i\omega t/2}|+\rangle + c_- e^{+i\omega t/2}|-\rangle. \tag{5.8}$$

The coefficients $c_\pm$ are such that

$$\langle\phi|\phi\rangle = 1, \tag{5.9}$$

$$\rightsquigarrow |c_+|^2 + |c_-|^2 = 1. \tag{5.10}$$

Note that there is a state $|\psi\rangle$ orthogonal to $|\phi\rangle$,

$$|\psi\rangle = c_-^*|+\rangle - c_+^*|-\rangle, \tag{5.11}$$

such that

$$\langle\phi|\psi\rangle = c_+^* c_-^* - c_-^* c_+^* = 0.$$

The probability that the electron is found again in the same state $|\phi\rangle$ after some time t is given by

$$P_1(t) = \left|\langle\phi|\phi, t\rangle\right|^2 = \ldots = 1 - 4|c_+|^2|c_-|^2 \sin^2\left(\frac{\omega t}{2}\right). \tag{5.12}$$

If we set $|c_+| = \sin\theta$ and, thus, $|c_-| = \cos\theta$, the probability can be written as

$$P_1(t) = \left|\langle\phi|\phi, t\rangle\right|^2 = \ldots = 1 - \sin^2\left(\frac{\omega t}{2}\right)\sin^2(2\theta). \tag{5.13}$$

Notice that for $\theta = 0, \frac{\pi}{2}$, which corresponds to $|\phi\rangle$ being an eigenstate $|\pm\rangle$ of the Hamiltonian, the probability that the electron remains in the same state is 100% at all times. For all other states, $\theta \neq 0, \frac{\pi}{2}$, the probability oscillates with time. There exist, however, periodic times,

$$t = n\frac{2\pi}{\omega}, \quad n = 0, 1, 2, \ldots$$

for which the probability to measure the system in the original, arbitrarily chosen, state is always one,

$$P_1\left(t = \frac{2\pi}{\omega}\right) = 1.$$

The probability to transition to the orthogonal state $|\psi\rangle$ after some time t is

$$P_2(t) = \left|\langle\psi|\phi, t\rangle\right|^2 = \ldots = \sin^2\left(\frac{\omega t}{2}\right)\sin^2(2\theta). \tag{5.14}$$

Both probabilities are shown in Fig. 5.1. The total probability is conserved,

$$P_1 + P_2 = 1.$$

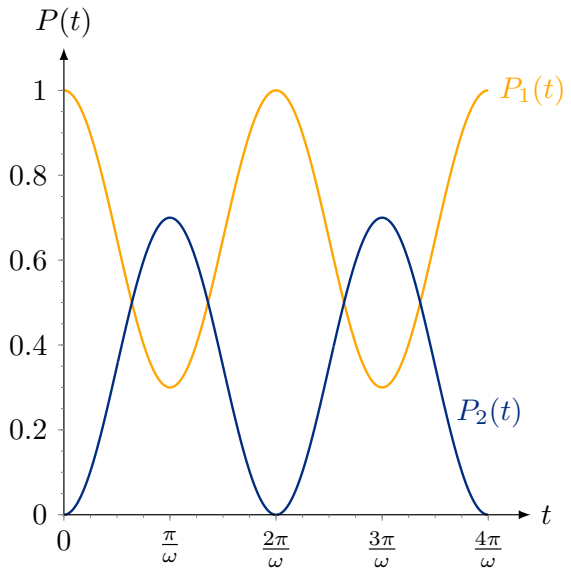


Figure 5.1. The probability P_1 of finding the system again in its initial state and the probability P_2 of a transition to the orthogonal state, as functions of the time t.

5.2 A generic two-state system

Consider a system with only two base states,

$$|1\rangle \doteq \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad |2\rangle \doteq \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

A generic state can be written, in this basis, as

$$\begin{aligned} |\psi\rangle &= |1\rangle\langle 1|\psi\rangle + |2\rangle\langle 2|\psi\rangle \\ &\doteq \begin{pmatrix} \langle 1|\psi\rangle \\ \langle 2|\psi\rangle \end{pmatrix}. \end{aligned} \tag{5.15}$$

The action of the Hamiltonian on a generic state is

$$H|\psi\rangle = \sum_{i,j} |i\rangle\langle i|H|j\rangle\langle j|\psi\rangle \,, \tag{5.16}$$

and it is represented as

$$H|\psi\rangle \doteq \begin{pmatrix} \langle 1|H|1\rangle & \langle 1|H|2\rangle \\ \langle 2|H|1\rangle & \langle 2|H|2\rangle \end{pmatrix} \begin{pmatrix} \langle 1|\psi\rangle \\ \langle 2|\psi\rangle \end{pmatrix}. \tag{5.17}$$

The Hamiltonian matrix of the system, as represented in this basis, has the general form

$$H \doteq \begin{pmatrix} H_{11} & H_{12} \\ H_{12}^* & H_{22} \end{pmatrix}, \tag{5.18}$$

where

$$H_{ij} = H_{ji}^* = \langle i|H|j\rangle \,,$$

so that

$$H^\dagger = H.$$

Notice that the diagonal elements must be real, $H_{11}^* = H_{11}$, $H_{22}^* = H_{22}$, while the off-diagonal elements can be complex.

We shall assume for now that the Hamiltonian is independent of time,

$$H(t) = H, \quad \forall\, t\,. \tag{5.19}$$

The Hamiltonian has two energy eigenstates with eigenvalues E satisfying

$$\det \begin{pmatrix} H_{11} - E & H_{12} \\ H_{12}^* & H_{22} - E \end{pmatrix} = 0\,. \tag{5.20}$$

This yields the solutions

$$E_\pm = \frac{H_{11} + H_{22} \pm \sqrt{\Delta}}{2}, \quad \Delta = (H_{11} - H_{22})^2 + 4|H_{12}|^2. \tag{5.21}$$

The energy eigenstates, $|E_\pm\rangle$, satisfy

$$(H - E_\pm \mathbf{1})\,|E_\pm\rangle = 0\,. \tag{5.22}$$

We find (exercise) that

$$|E_\pm\rangle = \frac{1}{\sqrt{|H_{12}|^2 + (E_\pm - H_{11})^2}} \begin{pmatrix} H_{12} \\ E_\pm - H_{11} \end{pmatrix}, \tag{5.23}$$

where we have normalised so that $\langle E_\pm | E_\pm \rangle = 1$. This formula fails when $H_{12} = 0$ (the vector for $E_\pm = H_{11}$ vanishes); then H is already diagonal and the eigenstates are $|1\rangle, |2\rangle$. An equivalent choice is $(E_\pm - H_{22},\, H_{12}^*)^T$. The eigenstates also satisfy (exercise)

$$\langle E_\pm | E_\mp \rangle = 0.$$

Having found the eigenvalues and eigenstates of the Hamiltonian, we can simplify the study of time evolution by changing to a new basis of our Hilbert space whose elements are the Hamiltonian eigenstates,

$$|E_+\rangle \doteq \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad |E_-\rangle \doteq \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

In the new basis, the Hamiltonian is represented as

$$H \doteq \begin{pmatrix} \langle E_+|\, H\, |E_+\rangle & \langle E_+|\, H\, |E_-\rangle \\ \langle E_-|\, H\, |E_+\rangle & \langle E_-|\, H\, |E_-\rangle \end{pmatrix} \doteq \begin{pmatrix} E_+ & 0 \\ 0 & E_- \end{pmatrix}. \tag{5.24}$$

The Hamiltonian takes the form of a diagonal matrix. The time evolution operator is now computed by exponentiating this diagonal matrix. With standard linear algebra, we find that

$$U(t) = \exp\left(-\frac{iHt}{\hbar}\right) \doteq \begin{pmatrix} \exp\left(-\frac{iE_+t}{\hbar}\right) & 0 \\ 0 & \exp\left(-\frac{iE_-t}{\hbar}\right) \end{pmatrix}. \tag{5.25}$$

It is easy to verify that the evolution operator does not change the energy eigenstates, except for a phase:

$$U(t)\left|E_\pm\right\rangle = \exp\left(-\frac{iE_\pm t}{\hbar}\right)\left|E_\pm\right\rangle. \tag{5.26}$$

To obtain the representation of $U(t)$ in the $\{|1\rangle, |2\rangle\}$ basis, we use that

$$\langle i_1|U(t)|i_2\rangle = \sum_{kl}\langle i_1|e_k\rangle\langle e_k|U(t)|e_l\rangle\langle e_l|i_2\rangle, \quad i_{1,2}\in\{1,2\}, e_{1,2}\in\{E_+,E_-\}. \tag{5.27}$$

In matrix notation, we have

$$\overbrace{\begin{pmatrix}\langle 1|U(t)|1\rangle & \langle 1|U(t)|2\rangle \\ \langle 2|U(t)|1\rangle & \langle 2|U(t)|2\rangle\end{pmatrix}}^{\substack{\text{time evolution operator}\\ \text{in } \{|1\rangle,|2\rangle\} \text{ basis}}} = \underbrace{\begin{pmatrix}\langle 1|E_+\rangle & \langle 1|E_-\rangle \\ \langle 2|E_+\rangle & \langle 2|E_-\rangle\end{pmatrix}}_{\substack{\text{basis transformation}\\ \{|E_+\rangle,|E_-\rangle\}\to\{|1\rangle,|2\rangle\}}}\underbrace{\begin{pmatrix}\langle E_+|U(t)|E_+\rangle & \langle E_+|U(t)|E_-\rangle \\ \langle E_-|U(t)|E_+\rangle & \langle E_-|U(t)|E_-\rangle\end{pmatrix}}_{\substack{\text{time evolution operator}\\ \text{in } \{|E_+\rangle,|E_-\rangle\} \text{ basis}}}\underbrace{\begin{pmatrix}\langle E_+|1\rangle & \langle E_+|2\rangle \\ \langle E_-|1\rangle & \langle E_-|2\rangle\end{pmatrix}}_{\substack{\text{basis transformation}\\ \{|1\rangle,|2\rangle\}\to\{|E_+\rangle,|E_-\rangle\}}}. \tag{5.28}$$

5.3 The ammonia molecule

In this section, we will describe a simplified model of the ammonia molecule, NH_3. The molecule consists of three hydrogen atoms and one nitrogen atom and can be depicted as in Fig. 5.2. Let us make the simplifying assumption that the collective motion of the molecule is not important and focus only on the spinning motion of the molecule around its axis. The molecule can spin in two different ways, with the nitrogen atom either "above" the plane of the hydrogen atoms or "below" that plane. We denote the two states corresponding to the two types of rotations as $|1\rangle, |2\rangle$. Ignoring all other motions, we can consider these kets as base kets:

$$|1\rangle \doteq \begin{pmatrix}1\\0\end{pmatrix}, \quad |2\rangle \doteq \begin{pmatrix}0\\1\end{pmatrix}. \tag{5.29}$$

We assume that the system is completely symmetric under exchanging the states $|1\rangle, |2\rangle$. Specifically, the expectation value for the energy in the two states is the same:

$$\langle 1|H|1\rangle = \langle 2|H|2\rangle = E_0. \tag{5.30}$$

We also allow for a probability that, with time, the nitrogen atom can push through the hydrogen plane and flip its state:

$$\langle 1|H|2\rangle = \langle 2|H|1\rangle = -A. \tag{5.31}$$

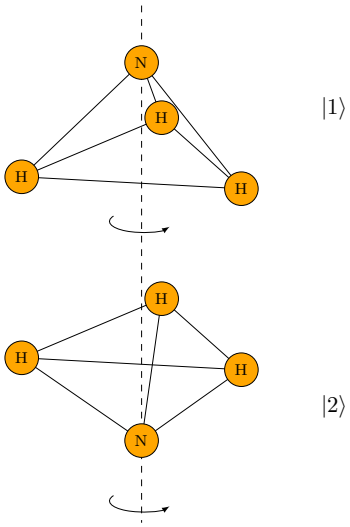


Figure 5.2. The rotation of the molecule of ammonia around its axis can be described as a two-state system.

Physically, this is quantum tunnelling of the nitrogen atom through the potential barrier formed by the hydrogen plane, not a classical push; classically, the molecule would stay in $|1\rangle$ or $|2\rangle$. The tunnelling amplitude A is exponentially small in the barrier, giving the inversion splitting $2A \approx 10^{-4}\,\text{eV}$, i.e. a frequency $\nu \approx 24\,\text{GHz}$ (Feynman Lectures, Vol. III, Chs. 8–9). The Hamiltonian of the system can be written as

$$H \doteq \begin{pmatrix} E_0 & -A \\ -A & E_0 \end{pmatrix}. \tag{5.32}$$

The energy eigenvalues are

$$E_\pm = E_0 \pm A, \tag{5.33}$$

and the corresponding eigenstates are

$$|+\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad |-\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}. \tag{5.34}$$

The time evolution operator is a function of the Hamiltonian:

$$U(t) = e^{-\frac{iHt}{\hbar}}. \tag{5.35}$$

Exercise 5.1. *Show that the matrix $U(t)$ is a polynomial in H with no quadratic or higher-order terms.*

In the $\{|+\rangle, |-\rangle\}$ basis,

$$U(t) \doteq \begin{pmatrix} \langle +|U(t)|+\rangle & \langle +|U(t)|-\rangle \\ \langle -|U(t)|+\rangle & \langle -|U(t)|-\rangle \end{pmatrix} \tag{5.36}$$

$$= \begin{pmatrix} \exp\left(-\frac{i(E_0+A)t}{\hbar}\right) & 0 \\ 0 & \exp\left(-\frac{i(E_0-A)t}{\hbar}\right) \end{pmatrix}. \tag{5.37}$$

We now compute the transformation matrix,

$$|b\rangle\langle a|, \quad |a\rangle \in \{|+\rangle, |-\rangle\}, \quad |b\rangle \in \{|1\rangle, |2\rangle\}.$$

The time evolution operator in the $\{|1\rangle, |2\rangle\}$ basis is

$$U(t) \doteq \exp\left(-\frac{iE_0 t}{\hbar}\right)\begin{pmatrix} \cos\left(\frac{At}{\hbar}\right) & i\sin\left(\frac{At}{\hbar}\right) \\ i\sin\left(\frac{At}{\hbar}\right) & \cos\left(\frac{At}{\hbar}\right) \end{pmatrix}. \tag{5.38}$$

Unsurprisingly, if we start at $t = 0$ from any state other than an energy eigenstate, the system changes with time. A general state $\big|\psi, t_0 = 0\big\rangle$ with unit norm, $\big\langle\psi, t_0\big|\psi, t_0\big\rangle = 1$,

$$\big|\psi, t_0 = 0\big\rangle \doteq e^{i\phi}\begin{pmatrix}\cos\theta \\ \sin\theta\end{pmatrix}, \tag{5.39}$$

evolves as

$$\big|\psi, t\big\rangle \doteq e^{i\left(\phi - E_0 t/\hbar\right)}\begin{pmatrix} \cos\theta\cos\left(\frac{At}{\hbar}\right) + i\sin\theta\sin\left(\frac{At}{\hbar}\right) \\ \sin\theta\cos\left(\frac{At}{\hbar}\right) + i\cos\theta\sin\left(\frac{At}{\hbar}\right) \end{pmatrix}. \tag{5.40}$$

The probability P_1 that the ammonia molecule is found in the same state $\big|\psi, t_0 = 0\big\rangle$ after some time t has elapsed is

$$P_1 = \Big|\big\langle\psi, t_0 = 0\big|\psi, t\big\rangle\Big|^2 = 1 - \sin^2\left(\frac{At}{\hbar}\right)\cos^2(2\theta). \tag{5.41}$$

We notice that for times

$$t = \frac{n\pi\hbar}{A}, \quad n = 0, 1, 2, \ldots \tag{5.42}$$

the molecule is certainly ($P_1 = 1$) back in the original state. For values of θ corresponding to the energy eigenstates, $\theta = \pm\frac{\pi}{4}$, the molecule remains in the original state at all times.

At $t_0 = 0$ we can find a state $|\xi, t_0 = 0\rangle$ which is orthogonal to $\big|\psi, t_0 = 0\big\rangle$, satisfying

$$\big\langle\psi, t_0 = 0\big|\xi, t_0 = 0\big\rangle = 0.$$

Namely,

$$|\xi, t_0 = 0\rangle \doteq e^{i\phi}\begin{pmatrix}\sin\theta \\ -\cos\theta\end{pmatrix}. \tag{5.43}$$

The probability P_2 that the ammonia molecule transitions after some time t from the state $\big|\psi, t_0 = 0\big\rangle$ to its orthogonal state $|\xi, t_0 = 0\rangle$ is

$$P_2 = \Big|\big\langle\xi, t_0 = 0\big|\psi, t\big\rangle\Big|^2 = \sin^2\left(\frac{At}{\hbar}\right)\cos^2(2\theta), \tag{5.44}$$

and we find that

$$P_1 + P_2 = 1.$$

The time evolution is completely analogous to the evolution of a spin-1/2 particle inside a magnetic field which we have studied earlier.

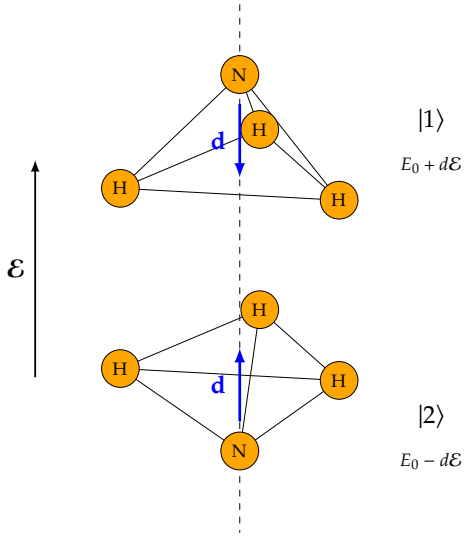


Figure 5.3. The molecule inside an electric field $\boldsymbol{\mathcal{E}}$ (pointing up). The electric dipole moment $\mathbf{d}$ has a vertical direction from the nitrogen atom to the plane of the hydrogen atoms. In $|1\rangle$ it is antiparallel to the field (energy $E_0 + d\mathcal{E}$), in $|2\rangle$ parallel to it (energy $E_0 - d\mathcal{E}$).

5.4 The ammonia molecule inside an electric field

The distribution of the electric charges in the ammonia molecule is not spatially symmetric; the molecule has a non-zero electric dipole moment with a vertical direction from the nitrogen atom (the negative end) to the hydrogen plane (the positive end). When the molecule is placed inside an electric field $\boldsymbol{\mathcal{E}}$, the energy of the molecule in the states $|1\rangle$, $|2\rangle$ is not the same anymore, since in one state the dipole moment d is in the direction of the electric field and in the other against it. More precisely, with the potential energy $-\mathbf{d}\cdot\boldsymbol{\mathcal{E}}$ and $\boldsymbol{\mathcal{E}}$ as in Fig. 5.3, the dipole is antiparallel to the field in $|1\rangle$ (energy raised by $d\mathcal{E}$) and parallel to it in $|2\rangle$ (energy lowered by $d\mathcal{E}$). We then have the Hamiltonian

$$H \doteq \begin{pmatrix} E_0 + d\mathcal{E} & -A \\ -A & E_0 - d\mathcal{E} \end{pmatrix}. \tag{5.45}$$

The energy eigenvalues are

$$E_\pm = E_0 \pm \sqrt{A^2 + d^2\mathcal{E}^2}. \tag{5.46}$$

For weak electric fields, $d\mathcal{E} \ll A$, we can approximate[1]

$$E_\pm \approx E_0 \pm A \pm \frac{d^2\mathcal{E}^2}{2A}. \tag{5.47}$$

In an inhomogeneous electric field, the molecules will move to the regions of the field where they can minimise their energy. The force acting on a molecule in a $|\pm\rangle$ state is,

$$\mathbf{F}_\pm = -\boldsymbol{\nabla}E_\pm = \mp\frac{d^2}{2A}\boldsymbol{\nabla}\mathcal{E}^2\,. \tag{5.48}$$

The electric force acts in opposite directions on molecules in the two energy eigenstates and separates them (Fig. 5.4).

[1] Second-order Taylor expansion.

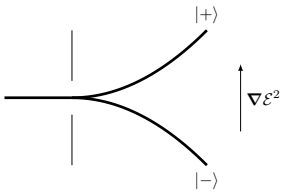


Figure 5.4. The force on ammonia molecules inside an inhomogeneous electric field is opposite for the two energy eigenstates.

5.5 Ammonia in an oscillating electric field

Let us put the ammonia molecule in a time-varying electric field,

$$\mathcal{E}\cos(\omega t).$$

The Hamiltonian in the basis of rotation states $\{|1\rangle, |2\rangle\}$ is

$$H(t) \doteq \begin{pmatrix} E_0 + d\mathcal{E}\cos(\omega t) & -A \\ -A & E_0 - d\mathcal{E}\cos(\omega t) \end{pmatrix}. \tag{5.49}$$

Notice (exercise) that the Hamiltonian at one time t_1 does not commute with the Hamiltonian at a later time t_2:

$$[H(t_1), H(t_2)] \neq 0\,.$$

We now need to solve Schrödinger's equation,

$$i\hbar\frac{\partial}{\partial t}\left|\psi, t\right\rangle = H(t)\left|\psi, t\right\rangle \tag{5.50}$$

in a more complicated situation where the Hamiltonian varies with time. We can exploit the fact that we already know how to solve Schrödinger's equation for a vanishing electric field. For small electric fields, the "bulk" of the evolution should be determined by this static solution. For

$$H_0 = H(t)|_{\mathcal{E}=0} = \begin{pmatrix} E_0 & -A \\ -A & E_0 \end{pmatrix}, \tag{5.51}$$

we found two energy eigenstates, $\{|+\rangle, |-\rangle\}$ with eigenvalues $E_+ = E_0 + |A|, E_- = E_0 - |A|$. It will be most convenient to bring the H_0 Hamiltonian (what we envisage to be the dominant part of the Hamiltonian) into diagonal form, by working with the energy eigenstates of H_0 as a basis of our Hilbert space.

To change basis, we need the transformation matrix

$$V = \begin{pmatrix} \langle +|1\rangle & \langle -|1\rangle \\ \langle +|2\rangle & \langle -|2\rangle \end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}. \tag{5.52}$$

We have that

$$\begin{pmatrix} |+\rangle \\ |-\rangle \end{pmatrix} = V^T \begin{pmatrix} |1\rangle \\ |2\rangle \end{pmatrix} = \begin{pmatrix} \frac{|1\rangle - |2\rangle}{\sqrt{2}} \\ \frac{|1\rangle + |2\rangle}{\sqrt{2}} \end{pmatrix}, \tag{5.53}$$

and, because $V^T V = 1$, we also have that

$$\begin{pmatrix} |1\rangle \\ |2\rangle \end{pmatrix} = V \begin{pmatrix} |+\rangle \\ |-\rangle \end{pmatrix} = \begin{pmatrix} \frac{|+\rangle + |-\rangle}{\sqrt{2}} \\ \frac{|-\rangle - |+\rangle}{\sqrt{2}} \end{pmatrix}. \tag{5.54}$$

In the $\{|+\rangle, |-\rangle\}$ basis, we have

$$H_0 = \begin{pmatrix} E_0 + |A| & 0 \\ 0 & E_0 - |A| \end{pmatrix}. \tag{5.55}$$

The full Hamiltonian $H(t)$ is not diagonal in the new basis, as $|+\rangle, |-\rangle$ are only eigenstates of H_0. Indeed, we find that in the $\{|+\rangle, |-\rangle\}$ basis,

$$H(t) = \begin{pmatrix} E_0 + |A| & d\mathcal{E}\cos(\omega t) \\ d\mathcal{E}\cos(\omega t) & E_0 - |A| \end{pmatrix}. \tag{5.56}$$

Notice that the non-diagonal elements of the Hamiltonian are indeed small for a small electric field.

Without loss of generality, and without making any approximation yet, we write

$$\left|\psi, t\right\rangle = \exp\left(-\frac{iH_0 t}{\hbar}\right)\left|\phi, t\right\rangle, \tag{5.57}$$

where, in the convenient $\{|+\rangle, |-\rangle\}$ basis, the evolution operator for a free ammonia molecule is also diagonal,

$$\exp\left(-\frac{iH_0 t}{\hbar}\right) = \exp\left(-\frac{iE_0 t}{\hbar}\right)\begin{pmatrix} e^{-\frac{i\omega_0 t}{2}} & 0 \\ 0 & e^{+\frac{i\omega_0 t}{2}} \end{pmatrix}, \tag{5.58}$$

with

$$\omega_0 = \frac{2|A|}{\hbar}. \tag{5.59}$$

With our substitution of Eq. (5.57), Schrödinger's equation becomes

$$i\hbar\frac{\partial}{\partial t}\left|\phi, t\right\rangle = H_I(t)\left|\phi, t\right\rangle, \tag{5.60}$$

with

$$\begin{aligned} H_I(t) &= \exp\left(+\frac{iH_0 t}{\hbar}\right)[H(t) - H_0]\exp\left(-\frac{iH_0 t}{\hbar}\right) \\ &= d\mathcal{E}\cos(\omega t)\begin{pmatrix} 0 & e^{+i\omega_0 t} \\ e^{-i\omega_0 t} & 0 \end{pmatrix}. \end{aligned} \tag{5.61}$$

In components, Schrödinger's equation becomes

$$i\hbar\frac{\partial}{\partial t}\begin{pmatrix}\phi_+(t)\\ \phi_-(t)\end{pmatrix} = \frac{d\mathcal{E}}{2}\begin{pmatrix} 0 & e^{+i(\omega_0-\omega)t} + e^{+i(\omega_0+\omega)t} \\ e^{-i(\omega_0-\omega)t} + e^{-i(\omega_0+\omega)t} & 0 \end{pmatrix}\begin{pmatrix}\phi_+(t)\\ \phi_-(t)\end{pmatrix}, \tag{5.62}$$

where

$$\phi_\pm(t) \equiv \left\langle \pm\middle|\phi, t\right\rangle. \tag{5.63}$$

Now, we will assume that the electric field is small ($d\mathcal{E} \ll \hbar\omega_0 = 2|A|$). The $\left|\phi, t\right\rangle$ ket should not change very fast in comparison to the rapidly oscillating exponential terms

with frequency $\omega + \omega_0$. These oscillations average to zero and do not contribute much to the change of the $\phi_\pm$ coefficients. We make an approximation and ignore them. The exponentials with frequency $\omega - \omega_0$ are also rapidly oscillating unless the two frequencies are very close to each other, which is what we will arrange by tuning our electric field. We write

$$i\hbar \frac{\partial}{\partial t} \begin{pmatrix} \phi_+(t) \\ \phi_-(t) \end{pmatrix} \approx \frac{d\mathcal{E}}{2} \begin{pmatrix} 0 & e^{+i(\omega_0-\omega)t} \\ e^{-i(\omega_0-\omega)t} & 0 \end{pmatrix} \begin{pmatrix} \phi_+(t) \\ \phi_-(t) \end{pmatrix}. \tag{5.64}$$

5.5.1 Transitions at resonance

Let us consider the situation where the frequency of the electric field is tuned to $\omega = \omega_0$. Then,

$$i\hbar \frac{\partial}{\partial t} \begin{pmatrix} \phi_+(t) \\ \phi_-(t) \end{pmatrix} \approx \frac{d\mathcal{E}}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \phi_+(t) \\ \phi_-(t) \end{pmatrix} = \hbar\Omega_\mathcal{E} \begin{pmatrix} \phi_-(t) \\ \phi_+(t) \end{pmatrix}, \tag{5.65}$$

with

$$\Omega_\mathcal{E} = \frac{d\mathcal{E}}{2\hbar}. \tag{5.66}$$

These equations have the same form as our familiar equations for the canonical position and canonical momentum of a simple (classical) harmonic oscillator. The solution of our system of differential equations is

$$\begin{pmatrix} \phi_+(t) \\ \phi_-(t) \end{pmatrix} = \begin{pmatrix} \cos(\Omega_\mathcal{E} t) & -i\sin(\Omega_\mathcal{E} t) \\ -i\sin(\Omega_\mathcal{E} t) & \cos(\Omega_\mathcal{E} t) \end{pmatrix} \begin{pmatrix} \phi_+(0) \\ \phi_-(0) \end{pmatrix}. \tag{5.67}$$

We can now calculate the time dependence of the $|\psi, t\rangle$ state from

$$\begin{aligned} \begin{pmatrix} \psi_+(t) \\ \psi_-(t) \end{pmatrix} &= \exp\left(-\frac{iH_0 t}{\hbar}\right) \begin{pmatrix} \phi_+(t) \\ \phi_-(t) \end{pmatrix} \\ &= \exp\left(-\frac{iE_0 t}{\hbar}\right) \begin{pmatrix} e^{-\frac{i\omega_0 t}{2}} & 0 \\ 0 & e^{+\frac{i\omega_0 t}{2}} \end{pmatrix} \begin{pmatrix} \cos(\Omega_\mathcal{E} t) & -i\sin(\Omega_\mathcal{E} t) \\ -i\sin(\Omega_\mathcal{E} t) & \cos(\Omega_\mathcal{E} t) \end{pmatrix} \begin{pmatrix} \phi_+(0) \\ \phi_-(0) \end{pmatrix} \\ &= \exp\left(-\frac{iE_0 t}{\hbar}\right) \begin{pmatrix} \cos(\Omega_\mathcal{E} t)e^{-\frac{i\omega_0 t}{2}} & -i\sin(\Omega_\mathcal{E} t)e^{-\frac{i\omega_0 t}{2}} \\ -i\sin(\Omega_\mathcal{E} t)e^{+\frac{i\omega_0 t}{2}} & \cos(\Omega_\mathcal{E} t)e^{+\frac{i\omega_0 t}{2}} \end{pmatrix} \begin{pmatrix} \psi_+(0) \\ \psi_-(0) \end{pmatrix}. \end{aligned} \tag{5.68}$$

In the last step, we have used that $|\psi, t = 0\rangle = |\phi, t = 0\rangle$.

Let us assume that at $t = 0$ we start with a molecule in the state of highest energy,

$$|\psi, t = 0\rangle = |+\rangle. \tag{5.69}$$

Then, after time t the state will be

$$\begin{aligned} \begin{pmatrix} \psi_+(t) \\ \psi_-(t) \end{pmatrix} &= \exp\left(-\frac{iE_0 t}{\hbar}\right) \begin{pmatrix} \cos(\Omega_\mathcal{E} t)e^{-\frac{i\omega_0 t}{2}} & -i\sin(\Omega_\mathcal{E} t)e^{-\frac{i\omega_0 t}{2}} \\ -i\sin(\Omega_\mathcal{E} t)e^{+\frac{i\omega_0 t}{2}} & \cos(\Omega_\mathcal{E} t)e^{+\frac{i\omega_0 t}{2}} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ &= \exp\left(-\frac{iE_0 t}{\hbar}\right) \begin{pmatrix} \cos(\Omega_\mathcal{E} t)e^{-\frac{i\omega_0 t}{2}} \\ -i\sin(\Omega_\mathcal{E} t)e^{+\frac{i\omega_0 t}{2}} \end{pmatrix}. \end{aligned} \tag{5.70}$$

The probability that the molecule transitions to the lower energy state is

$$P\left(|+\rangle \to |-\rangle\right) = |\psi_-(t)|^2 = \sin^2(\Omega_\mathcal{E} t). \tag{5.71}$$

This is periodic: at certain times, controlled by the frequency $\Omega_{\mathcal{E}}$, the molecule has certainly made the transition to the lower energy state. Similarly, the probability for the molecule to be found again in the state $|+\rangle$ is

$$P\left(|+\rangle \to |+\rangle\right) = |\psi_+(t)|^2 = \cos^2(\Omega_{\mathcal{E}} t). \tag{5.72}$$

5.5.2 Microwave amplification with stimulated emission of radiation

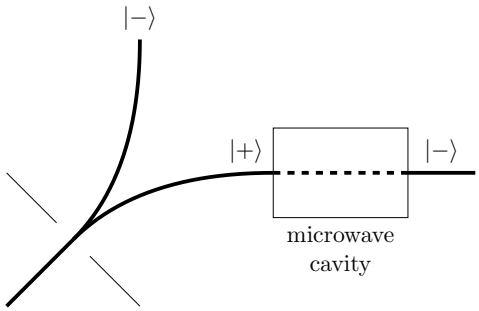


Figure 5.5. The ammonia maser.

We now discuss the ammonia maser (Fig. 5.5), which is based on the quantum behaviour of the two-state molecule. By means of a static inhomogeneous electric field in an apparatus of Stern–Gerlach type, we can separate molecules in the two energy eigenstates. Molecules in the upper state are directed into a cavity with an oscillating electric field having the resonance frequency $\omega = \omega_0 = 2A/\hbar$. After a time $T = \frac{\pi}{2\Omega_{\mathcal{E}}}$ the molecules will transition to the lower energy eigenstate, losing energy equal to the difference of the two energy eigenvalues. This energy feeds back to the oscillating electric field. To prevent the molecules from transitioning back to the original state, thus recapturing the released energy, we arrange the length of the cavity such that the molecules exit the cavity after a time T.

5.5.3 Transitions off resonance

We now solve the differential equation for very small times t and $\omega \sim \omega_0$,

$$i\hbar \frac{\partial}{\partial t} \begin{pmatrix} \phi_+(t) \\ \phi_-(t) \end{pmatrix} \approx \frac{d\mathcal{E}}{2} \begin{pmatrix} 0 & e^{+i(\omega_0-\omega)t} \\ e^{-i(\omega_0-\omega)t} & 0 \end{pmatrix} \begin{pmatrix} \phi_+(t) \\ \phi_-(t) \end{pmatrix}. \tag{5.73}$$

Assuming that we start from a $|+\rangle$ state corresponding to $\phi_+(0) = 1, \phi_-(0) = 0$, for very small times the coefficient $\phi_+(t)$ will remain very close to 1. Then, we have

$$i\hbar \frac{\partial}{\partial t} \phi_-(t) = \frac{d\mathcal{E}}{2} e^{-i(\omega_0-\omega)t}, \tag{5.74}$$

which yields

$$\phi_-(t) = \frac{d\mathcal{E}}{2\hbar} \frac{1 - e^{+i(\omega-\omega_0)t}}{\omega - \omega_0}. \tag{5.75}$$

The probability for a transition is

$$P\left(|+\rangle \to |-\rangle\right) = \left[\frac{d\mathcal{E}}{\hbar}\right]^2 \frac{\sin^2\left[(\omega-\omega_0)t/2\right]}{\left[\omega-\omega_0\right]^2}. \tag{5.76}$$

This function is sharply peaked around $\omega = \omega_0$. Unless the frequency of the field is tuned very precisely to ω_0, a transition practically does not occur.

Exercise 5.2. *Plot the transition probability as a function of ω. Prove that*

$$\int_{-\infty}^{+\infty} \mathrm{d}x \, \frac{\sin^2 x}{x^2} = \pi, \tag{5.77}$$

and use this to compute the area under the peak.

6

Position and momentum

6.1 Continuous spectra

So far, we have restricted our discussions to quantum mechanical systems which can be described by a discrete basis of states. We will now discuss operators for physical quantities with a continuous spectrum of eigenvalues. For these operators and their eigenstates, we need some generalisations. Primary examples are the quantum operators for the position and momentum of particles, which have a real-valued continuous spectrum of eigenvalues.

As we have already developed a formalism of quantum mechanics for operators with discrete eigenvalues, we can generalise the same formalism to continuous spectra with a limiting procedure. For simplicity, we will describe the method assuming a one-dimensional, rather than three-dimensional, space of positions described by a single continuous variable x. The values of x are eigenvalues of an operator $\hat{x}$,

$$\hat{x}\ |x\rangle = x\ |x\rangle\ . \tag{6.1}$$

We will approximate the space by a set of equally spaced discrete points x_i, separated from their nearest neighbours by a distance a. We can obtain the physical continuous space by increasing the density of discrete points, i.e. by taking the limit of zero spacing, $a \to 0$. Before we take this limit, the eigenvalues of the position operator are discrete,

$$\hat{x}\ |x_i\rangle = x_i\ |x_i\rangle\ , \quad i = 1, 2, \ldots\ . \tag{6.2}$$

In accordance with the principles of quantum mechanics that we have introduced in the discrete case, we take the position eigenstates to be orthogonal. We write the following orthogonality condition,

$$\langle x_i | x_j \rangle = \frac{1}{a}\,\delta_{ij}\ . \tag{6.3}$$

Anticipating the continuous limit, $a \to 0$, and that a state $|x_i\rangle$ stands for an infinite number of continuous states in its vicinity, in the above we have normalised the states $|x_i\rangle$ to the density $1/a$ of states, rather than to one. A general state $|\psi\rangle$ can be written as a superposition of discrete position eigenstates,

$$|\psi\rangle = \sum_i c_i\ |x_i\rangle\ . \tag{6.4}$$

Multiplying by a position bra $\langle x_j|$, we obtain

$$\begin{aligned} \langle x_j|\psi\rangle &= \sum_i c_i \langle x_j|x_i\rangle = \sum_i c_i \frac{1}{a}\delta_{ij} = \frac{c_j}{a} \\ &\rightsquigarrow c_j = a\langle x_j|\psi\rangle . \end{aligned} \tag{6.5}$$

Thus,

$$|\psi\rangle = \sum_i a\,|x_i\rangle\langle x_i|\psi\rangle . \tag{6.6}$$

We can now take the continuous limit of zero spacing, $a \to 0$, in Eq. (6.6). The sum over an infinite number of points becomes an integral,[1]

$$\lim_{a\to 0}\sum_i a = \int dx . \tag{6.7}$$

We then have

$$|\psi\rangle = \int dx\,|x\rangle\langle x|\psi\rangle . \tag{6.8}$$

As $|\psi\rangle$ can be any state, we then have the completeness relation for continuous spectra,

$$\int dx\,|x\rangle\langle x| = \mathbf{1} . \tag{6.9}$$

The probability that a particle is measured at the position corresponding to the state $|x_i\rangle$, in the discretised space, is given by[2]

$$P\left(|\psi\rangle \to |x_i\rangle\right) = \left|\frac{\langle x_i|\psi\rangle}{\sqrt{\langle x_i|x_i\rangle}\sqrt{\langle\psi|\psi\rangle}}\right|^2 = \frac{|\langle x_i|\psi\rangle|^2}{\langle x_i|x_i\rangle} = a\,|\langle x_i|\psi\rangle|^2 . \tag{6.10}$$

You can check that the sum of all such probabilities is one,

$$\begin{aligned} \sum_i P\left(|\psi\rangle \to |x_i\rangle\right) &= \sum_i a\,|\langle x_i|\psi\rangle|^2 = \ldots = \langle\psi|\left(\sum_i a\,|x_i\rangle\langle x_i|\psi\rangle\right) \\ &= \langle\psi|\psi\rangle = 1 , \end{aligned} \tag{6.11}$$

as it should. We can now take the continuous limit of the above result,

$$\begin{aligned} 1 &= \langle\psi|\psi\rangle = \lim_{a\to 0}\sum_i a\,|\langle x_i|\psi\rangle|^2 \\ &= \int dx\,|\langle x|\psi\rangle|^2 . \end{aligned} \tag{6.12}$$

We found that the sum of all probabilities is an integral over a probability density,

$$\rho(x) \equiv |\langle x|\psi\rangle|^2 . \tag{6.13}$$

[1] Recall the definition of the Riemann integral.

[2] Note that it would be wrong to write $P\left(|\psi\rangle \to |x_i\rangle\right) = |\langle x_i|\psi\rangle|^2$, as the ket $|x_i\rangle$ is not normalised to unity, $\langle x_i|x_i\rangle = 1/a \neq 1$.

The probability for a particle to be located in a region of length dx around a point x is

$$dP(x) = \rho(x)dx\,. \tag{6.14}$$

The amplitude $\langle x|\psi\rangle$ is the *wavefunction* of the particle.

Let us now return to the orthogonality condition of position states of Eq. (6.3). Multiplying by a and summing over i, we have

$$\sum_i a\,\langle x_i|x_j\rangle = \sum_i \delta_{ij} = 1. \tag{6.15}$$

In the continuum limit, the above turns into

$$\int dx\,\langle x|x'\rangle = 1\,. \tag{6.16}$$

From the above, we conclude that

$$\langle x|x'\rangle = \delta\,(x - x')\,. \tag{6.17}$$

The analysis we made above is generic and it applies to general operators with a continuous spectrum of eigenvalues. We summarise below our conclusions, highlighting the correspondence to the discrete case. For a Hermitian operator Ξ with eigenstates $|\xi\rangle$ and eigenvalues ξ,

$$\Xi\,|\xi\rangle = \xi\,|\xi\rangle\,, \tag{6.18}$$

where ξ is a continuous variable, we have

Discrete	Continuous
Kronecker δ_{ij}	Dirac $\delta(\xi - \xi')$
$\langle i|j\rangle = \delta_{ij}$	$\langle\xi|\xi'\rangle = \delta(\xi - \xi')$
$\|\psi\rangle = \sum_i |i\rangle\langle i|\psi\rangle$	$\|\psi\rangle = \int \mathrm{d}\xi\ |\xi\rangle\langle\xi|\psi\rangle$
$\sum_i P\left(\|\psi\rangle \to |\xi_i\rangle\right) = 1$	$\int d\xi\, dP(\xi) = \int d\xi\ \left\|\langle\xi|\psi\rangle\right\|^2 = 1$

6.2 Position operator and eigenstates

Let us now discuss position states in a three-dimensional space. We define Hermitian position operators $\hat{x}, \hat{y}, \hat{z}$ which commute with each other:

$$\left[\hat{x}, \hat{y}\right] = \left[\hat{y}, \hat{z}\right] = [\hat{z}, \hat{x}] = 0. \tag{6.19}$$

The vanishing commutators correspond to the physical requirement that all position coordinates of a particle can, in principle, be measured simultaneously without an uncertainty (in the sense of a quantum uncertainty principle). These operators have a common set of eigenstates, $|x, y, z\rangle$, with

$$\begin{aligned} \hat{x}\,|x, y, z\rangle &= x\,|x, y, z\rangle \\ \hat{y}\,|x, y, z\rangle &= y\,|x, y, z\rangle \\ \hat{z}\,|x, y, z\rangle &= z\,|x, y, z\rangle\,. \end{aligned} \tag{6.20}$$

We identify the eigenvalues x, y, z with the positions in the x-, y- and z-directions, respectively. Employing a compact notation, we can write

$$|\mathbf{r}\rangle \equiv \big|x, y, z\big\rangle \,. \tag{6.21}$$

Also, when there is no risk of confusion between the position operators and their eigenstates, we may drop the ˆ notation for the operators.

The eigenstates of the position operator,

$$\hat{\mathbf{r}}\,|\mathbf{r}\rangle = \mathbf{r}\,|\mathbf{r}\rangle \,,$$

are orthonormal,

$$\langle \mathbf{r}_1|\mathbf{r}_2\rangle = \delta^{(3)}(\mathbf{r}_1 - \mathbf{r}_2),$$

and form a complete basis, such that every one-particle state can be written as a superposition of position eigenstates,

$$\big|\phi\big\rangle = \int \mathrm{d}^3\mathbf{r}\; |\mathbf{r}\rangle \,\big\langle \mathbf{r}\big|\phi\big\rangle \;. \tag{6.22}$$

An experimental measurement of the position will collapse a general state $\big|\phi\big\rangle$ to a state $|\mathbf{r}\rangle$:

$$\big|\phi\big\rangle \longrightarrow |\mathbf{r}\rangle \,,$$

with a probability density amplitude $\big\langle \mathbf{r}\big|\phi\big\rangle$. This amplitude is the wave function. The probability for a transition to an infinitesimal volume $\mathrm{d}^3\mathbf{r}$ around the position $\mathbf{r}$ is

$$P\left(\big|\phi\big\rangle \to (\ \mathrm{d}^3\mathbf{r} \text{ around } \mathbf{r})\right) = \mathrm{d}^3\mathbf{r}\,\Big|\big\langle \mathbf{r}\big|\phi\big\rangle\Big|^2 \;. \tag{6.23}$$

A description of quantum states in terms of wave-functions corresponds to a representation of quantum states in the basis of eigenstates of the position operator. Although important, the position operator and the wave-function representation are no more special than other operators and the representations of Hilbert space states that they furnish. However, traditionally, many of the results of quantum mechanics have been worked out in the position representation first. For example, the scalar product of two quantum states is

$$\begin{aligned} \langle a|b\rangle &= \int \mathrm{d}^3\mathbf{r}\; \langle a|\mathbf{r}\rangle\,\langle \mathbf{r}|b\rangle \\ \rightsquigarrow \langle a|b\rangle &= \int \mathrm{d}^3\mathbf{r}\; \psi_a^*(\mathbf{r})\psi_b(\mathbf{r}), \end{aligned} \tag{6.24}$$

with the wave functions $\psi_{a,b}(\mathbf{r}) \equiv \langle \mathbf{r}|a, b\rangle$.

6.3 Momentum operator and eigenstates

As for the position operators, we have Hermitian operators for the components of the momentum which commute with each other,

$$\big[\hat{p}_x, \hat{p}_y\big] = \big[\hat{p}_y, \hat{p}_z\big] = \big[\hat{p}_z, \hat{p}_x\big] = 0. \tag{6.25}$$

The vanishing commutators originate from the physical requirement that all momentum components of a particle can, in principle, be measured simultaneously without an

uncertainty (in the sense of a quantum uncertainty principle). These operators have a common set of eigenstates, $|p_x, p_y, p_z\rangle$, with

$$\begin{aligned} \hat{p}_x |p_x, p_y, p_z\rangle &= p_x |p_x, p_y, p_z\rangle \\ \hat{p}_y |p_x, p_y, p_z\rangle &= p_y |p_x, p_y, p_z\rangle \\ \hat{p}_z |p_x, p_y, p_z\rangle &= p_z |p_x, p_y, p_z\rangle \,. \end{aligned} \tag{6.26}$$

We identify the eigenvalues p_x, p_y, p_z with the momenta in the x-, y- and z-directions, respectively. Employing a compact notation, we can write

$$|\mathbf{p}\rangle \equiv |p_x, p_y, p_z\rangle \,. \tag{6.27}$$

The three eigenvalue equations of Eq. (6.26) take the compact form,

$$\hat{\mathbf{p}} |\mathbf{p}\rangle = \mathbf{p} |\mathbf{p}\rangle \,. \tag{6.28}$$

As we have seen in Section 4.2, the link between quantum and classical mechanics requires that the position and momentum components in a certain direction do not commute. We repeat here the complete set of position and momentum commutators,

$$[\hat{x}_i, \hat{x}_j] = [\hat{p}_i, \hat{p}_j] = 0 \,, \quad [\hat{x}_i, \hat{p}_j] = i\hbar \delta_{ij} \,. \tag{6.29}$$

The above commutation relations imply, from the uncertainty principle of Eq. (3.65), that the position and the momentum components in a certain direction, for a quantum system in a state $|\psi\rangle$, cannot both be sharply defined:

$$(\Delta x)\,(\Delta p_x) \geq \frac{\left|\langle\psi| \,[x, p_x]\, |\psi\rangle\right|}{2} = \frac{\left|\langle\psi|\psi\rangle \, i\hbar\right|}{2} = \frac{\hbar}{2} \,. \tag{6.30}$$

Remark. The Robertson relation constrains the spreads $\Delta x, \Delta p_x$ of repeated measurements on an ensemble identically prepared in $|\psi\rangle$ (each quantity measured on separate copies); it does not limit the precision of a single position or momentum measurement. A precise position measurement is possible, but leaves the system in a state with large Δp_x (Sakurai, Sec. 1.4; Ballentine, Sec. 8.4).

6.4 Translation operator

The position-momentum commutation relations of Eq. (6.29) endow the quantum mechanical operator of momentum with the property of generating space translations on the Hilbert space of quantum systems. To see that, let us define an operator

$$\hat{P}(\Delta\mathbf{x}) = 1 - \frac{i}{\hbar} \hat{\mathbf{p}} \cdot \Delta\mathbf{x} \,, \tag{6.31}$$

where $\Delta\mathbf{x}$ is a vector corresponding to a space interval. The commutator of the operator $\hat{P}$ and a position operator is

$$\begin{aligned} \left[\hat{x}, \hat{P}(\Delta\mathbf{x})\right] &= -\frac{i}{\hbar} \left[\hat{x}, \hat{\mathbf{p}} \cdot \Delta\mathbf{x}\right] \\ &= -\frac{i}{\hbar} \left(\left[\hat{x}, \hat{p}_x\right] \Delta x + \left[\hat{x}, \hat{p}_y\right] \Delta y + \left[\hat{x}, \hat{p}_z\right] \Delta z\right) \\ &= \Delta x \,. \end{aligned} \tag{6.32}$$

Collectively, for the commutators with all three position operators, we obtain

$$\left[\hat{\mathbf{r}}, \hat{P}(\Delta\mathbf{x})\right] = \Delta\mathbf{x}. \tag{6.33}$$

Let us now consider the state

$$|\mathbf{x}'\rangle \equiv \hat{P}(\Delta\mathbf{x})\,|\mathbf{x}\rangle\,. \tag{6.34}$$

Probing it with a position operator, we have

$$\begin{aligned}
\hat{\mathbf{r}}\,|\mathbf{x}'\rangle &= \hat{\mathbf{r}}\hat{P}(\Delta\mathbf{x})\,|\mathbf{x}\rangle \\
&= \left[\hat{\mathbf{r}}, \hat{P}(\Delta\mathbf{x})\right]|\mathbf{x}\rangle + \hat{P}(\Delta\mathbf{x})\hat{\mathbf{r}}\,|\mathbf{x}\rangle \\
&= \Delta\mathbf{x}\;|\mathbf{x}\rangle + \hat{P}(\Delta\mathbf{x})\,\mathbf{x}\;|\mathbf{x}\rangle \\
&= (\mathbf{x}+\Delta\mathbf{x})\;\hat{P}(\Delta\mathbf{x})\;|\mathbf{x}\rangle + \Delta\mathbf{x}\frac{i}{\hbar}\hat{\mathbf{p}}\cdot\Delta\mathbf{x}\;|\mathbf{x}\rangle\;.
\end{aligned} \tag{6.35}$$

Let us now specialise to small intervals, $\Delta\mathbf{x} = \delta\mathbf{r} \to 0$. Then, the second term in the last equality is of order $O\left((\delta r)^2\right)$ and can be ignored. Thus, for infinitesimal δr, we write

$$\hat{\mathbf{r}}\,\left(\hat{P}(\delta\mathbf{r})\,|\mathbf{x}\rangle\right) = (\mathbf{x}+\delta\mathbf{r})\,\left(\hat{P}(\delta\mathbf{r})\;|\mathbf{x}\rangle\right). \tag{6.36}$$

This is an eigenvalue equation for the position operator, and we can identify $\hat{P}(\delta\mathbf{r})\;|\mathbf{x}\rangle$ as the position eigenstate with eigenvalue $(\mathbf{x}+\delta\mathbf{r})$. We write

$$\hat{P}(\delta\mathbf{r})\;|\mathbf{x}\rangle = |\mathbf{x}+\delta\mathbf{r}\rangle\;. \tag{6.37}$$

In summary, the operator

$$P(\delta\mathbf{r}) = 1 - \frac{i\mathbf{p}\cdot\delta\mathbf{r}}{\hbar}\,, \tag{6.38}$$

where $\mathbf{p}$ is the momentum operator, is a space translation operator for small intervals $\delta\mathbf{r}$. To perform a large translation $\mathbf{r}_B - \mathbf{r}_A$ from a point A to a point B we divide the translation into infinitesimally small steps,

$$\frac{\mathbf{r}_B - \mathbf{r}_A}{N} = \delta\mathbf{r}, \quad N \to \infty, \tag{6.39}$$

and perform an infinite number of very small translations $\delta\mathbf{r}$ in succession:

$$\begin{aligned}
|\mathbf{x}+\mathbf{r}_B-\mathbf{r}_A\rangle &= P(\delta\mathbf{r})^N\,|\mathbf{x}\rangle \\
&= \lim_{N\to\infty}\left(1 - \frac{i\mathbf{p}\cdot\delta\mathbf{r}}{\hbar}\right)^N|\mathbf{x}\rangle \\
&= \lim_{N\to\infty}\left(1 - \frac{i\mathbf{p}\cdot(\mathbf{r}_B-\mathbf{r}_A)}{N\hbar}\right)^N|\mathbf{x}\rangle
\end{aligned}$$

which leads to

$$|\mathbf{x}+\mathbf{r}_B-\mathbf{r}_A\rangle = U(\mathbf{r}_B-\mathbf{r}_A)\,|\mathbf{x}\rangle\,, \tag{6.40}$$

with

$$U(\mathbf{r}) = e^{-\frac{i}{\hbar}\mathbf{p}\cdot\mathbf{r}}\,. \tag{6.41}$$

The exponential of an operator is defined as usual through its Taylor series expansion to all orders,

$$e^X = \sum_{n=0}^{\infty} \frac{X^n}{n!} .$$

Due to the Hermiticity of the momentum operator, we have

$$U^\dagger(\mathbf{r})U(\mathbf{r}) = \mathbf{1} , \tag{6.42}$$

and the physically intuitive result that

$$U^{-1}(\mathbf{r}) = U(-\mathbf{r}) . \tag{6.43}$$

6.5 Representation of momentum in position space

Let us act with a small translation operator on an arbitrary state $|a\rangle$,

$$\begin{aligned}\left(1 - \frac{i\hat{\mathbf{p}}\cdot\delta\mathbf{r}}{\hbar}\right)|a\rangle &= \left(1 - \frac{i\hat{\mathbf{p}}\cdot\delta\mathbf{r}}{\hbar}\right)\int \mathrm{d}^3\mathbf{x}\ |\mathbf{x}\rangle\langle\mathbf{x}|a\rangle \\ &= \int \mathrm{d}^3\mathbf{x}\ |\mathbf{x}+\delta\mathbf{r}\rangle\langle\mathbf{x}|a\rangle \\ &= \int \mathrm{d}^3\mathbf{x}\ |\mathbf{x}\rangle\langle\mathbf{x}-\delta\mathbf{r}|a\rangle\end{aligned}$$

For small displacements, we can expand up to linear terms in $\delta\mathbf{r}$:

$$\langle\mathbf{x}-\delta\mathbf{r}|a\rangle = \langle\mathbf{x}|a\rangle - \delta\mathbf{r}\cdot\nabla\langle\mathbf{x}|a\rangle$$

Substituting above, we have

$$\begin{aligned}\left(1 - \frac{i\hat{\mathbf{p}}\cdot\delta\mathbf{r}}{\hbar}\right)|a\rangle &= \int \mathrm{d}^3\mathbf{x}\ |\mathbf{x}\rangle\langle\mathbf{x}|a\rangle - \int \mathrm{d}^3\mathbf{x}\ |\mathbf{x}\rangle(\delta\mathbf{r}\cdot\nabla)\langle\mathbf{x}|a\rangle \\ &= |a\rangle - \int \mathrm{d}^3\mathbf{x}\ |\mathbf{x}\rangle(\delta\mathbf{r}\cdot\nabla)\langle\mathbf{x}|a\rangle\end{aligned}$$

Comparing the two sides of the equation, we have

$$\hat{\mathbf{p}}\,|a\rangle = \int \mathrm{d}^3\mathbf{x}\ |\mathbf{x}\rangle[-i\hbar\nabla]\langle\mathbf{x}|a\rangle \tag{6.44}$$

and

$$\begin{aligned}\langle\mathbf{r}|\,\hat{\mathbf{p}}\,|a\rangle &= \int \mathrm{d}^3\mathbf{x}\ \langle\mathbf{r}|\mathbf{x}\rangle[-i\hbar\nabla]\langle\mathbf{x}|a\rangle \\ &= \int \mathrm{d}^3\mathbf{x}\ \delta(\mathbf{r}-\mathbf{x})[-i\hbar\nabla]\langle\mathbf{x}|a\rangle\end{aligned} \tag{6.45}$$

which leads to

$$\langle\mathbf{r}|\,\hat{\mathbf{p}}\,|a\rangle = [-i\hbar\nabla_{\mathbf{r}}]\langle\mathbf{r}|a\rangle . \tag{6.46}$$

Finally, we obtain the representation of the momentum operator in the position-ket basis by setting $|a\rangle = |\mathbf{r}'\rangle$, a position eigenstate,

$$\langle\mathbf{r}|\,\hat{\mathbf{p}}\,|\mathbf{r}'\rangle = [-i\hbar\nabla_{\mathbf{r}}]\langle\mathbf{r}|\mathbf{r}'\rangle , \tag{6.47}$$

we find

$$\langle\mathbf{r}|\,\hat{\mathbf{p}}\,|\mathbf{r}'\rangle = [-i\hbar\nabla_{\mathbf{r}}]\,\delta(\mathbf{r}-\mathbf{r}') . \tag{6.48}$$

6.6 Eigenstates of the momentum operator

The eigenstates of the momentum operator satisfy

$$\hat{\mathbf{p}}\,|\mathbf{p}\rangle = \mathbf{p}\,|\mathbf{p}\rangle\,, \tag{6.49}$$

with

$$\langle \mathbf{p}_1|\mathbf{p}_2\rangle = \delta^{(3)}\left(\mathbf{p}_1 - \mathbf{p}_2\right). \tag{6.50}$$

The probability amplitude that a particle with a momentum $\mathbf{p}$ is found at a position $\mathbf{x}$ is

$$\langle \mathbf{x}|\mathbf{p}\rangle\,.$$

On the one hand,

$$\langle \mathbf{x}|\,\hat{\mathbf{p}}\,|\mathbf{p}\rangle = \mathbf{p}\,\langle \mathbf{x}|\mathbf{p}\rangle\,,$$

but also, from the previous section,

$$\langle \mathbf{x}|\,\hat{\mathbf{p}}\,|\mathbf{p}\rangle = -i\hbar\nabla\,\langle \mathbf{x}|\mathbf{p}\rangle\,.$$

We have then derived the following differential equation for the sought probability amplitude:

$$-i\hbar\nabla\,\langle \mathbf{x}|\mathbf{p}\rangle = \mathbf{p}\,\langle \mathbf{x}|\mathbf{p}\rangle\,, \tag{6.51}$$

which has the solution

$$\langle \mathbf{x}|\mathbf{p}\rangle = N e^{\frac{i\mathbf{p}\cdot\mathbf{x}}{\hbar}}. \tag{6.52}$$

We compute the normalisation constant as follows:

$$\begin{aligned}
\delta^{(3)}(\mathbf{x}-\mathbf{y}) = \langle \mathbf{x}|\mathbf{y}\rangle &= \int \mathrm{d}^3\mathbf{p}\;\langle \mathbf{x}|\mathbf{p}\rangle\,\langle \mathbf{p}|\mathbf{y}\rangle = \int \mathrm{d}^3\mathbf{p}\;\langle \mathbf{x}|\mathbf{p}\rangle\,\langle \mathbf{y}|\mathbf{p}\rangle^* \\
&= |N|^2 \int \mathrm{d}^3\mathbf{p}\; e^{i\frac{\mathbf{p}}{\hbar}(\mathbf{x}-\mathbf{y})} = |N|^2\,(2\pi\hbar)^3\,\delta^{(3)}(\mathbf{x}-\mathbf{y}) \\
\rightsquigarrow |N| &= \frac{1}{(2\pi\hbar)^{3/2}}.
\end{aligned}$$

Thus, we can write

$$\langle \mathbf{x}|\mathbf{p}\rangle = \frac{1}{(2\pi\hbar)^{3/2}} e^{\frac{i\mathbf{p}\cdot\mathbf{x}}{\hbar}}. \tag{6.53}$$

6.6.1 Position and momentum wave functions

Let us assume that a particle is in a state $|\phi\rangle$. The probability amplitudes for finding the particle at a position $\mathbf{x}$ or with a momentum $\mathbf{p}$ are given by the position and momentum wave functions,

$$\langle \mathbf{x}|\phi\rangle\,, \quad \langle \mathbf{p}|\phi\rangle\,,$$

respectively. The two amplitudes are related by a Fourier transformation. Indeed,

$$\begin{aligned}
\langle \mathbf{x}|\phi\rangle &= \int_{-\infty}^{+\infty} \mathrm{d}^3\mathbf{p}\;\langle \mathbf{x}|\mathbf{p}\rangle\,\langle \mathbf{p}|\phi\rangle \\
&= \int_{-\infty}^{+\infty} \frac{\mathrm{d}^3\mathbf{p}}{(2\pi\hbar)^{(3/2)}} e^{i\frac{\mathbf{p}}{\hbar}\cdot\mathbf{x}}\,\langle \mathbf{p}|\phi\rangle\,.
\end{aligned} \tag{6.54}$$

Similarly,

$$\begin{aligned}\langle \mathbf{p}|\phi\rangle &= \int_{-\infty}^{+\infty} \mathrm{d}^3\mathbf{x}\ \langle \mathbf{p}|\mathbf{x}\rangle \langle \mathbf{x}|\phi\rangle \\ &= \int_{-\infty}^{+\infty} \frac{\mathrm{d}^3\mathbf{x}}{(2\pi\hbar)^{(3/2)}} e^{-i\frac{\mathbf{x}}{\hbar}\cdot\mathbf{p}} \langle \mathbf{x}|\phi\rangle\,.\end{aligned} \tag{6.55}$$

7

The harmonic oscillator

A very simple system in quantum mechanics, which we can solve exactly, is the harmonic oscillator in one space dimension. The Hamiltonian of the oscillator in classical physics is

$$H = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2, \tag{7.1}$$

where m is the mass of the oscillating particle and ω the angular frequency of the oscillation. For convenience, we shall set

$$m = \hbar = \omega = 1$$

in this chapter. It will always be easy to recover the full dependence on these parameters with dimensional analysis.

7.1 Quantisation

We quantise the harmonic oscillator by promoting the coordinate x and momentum p to operators, satisfying the commutation relation

$$\left[x, p\right] = i. \tag{7.2}$$

The Hamiltonian can be cast as

$$H = \frac{(x - ip)}{\sqrt{2}}\frac{(x + ip)}{\sqrt{2}} - i\frac{xp - px}{2}, \tag{7.3}$$

and using the commutation relation above,

$$H = a^\dagger a + \frac{1}{2}, \tag{7.4}$$

with

$$a = \frac{x + ip}{\sqrt{2}}, \quad a^\dagger = \frac{x - ip}{\sqrt{2}}. \tag{7.5}$$

For reasons that will become clear later, the $a, a^\dagger$ operators are called *annihilation* and *creation* operators, respectively. Their commutator is

$$\left[a, a^\dagger\right] = \left[\frac{x + ip}{\sqrt{2}}, \frac{x - ip}{\sqrt{2}}\right] = \frac{1}{2}\left(\left[x, -ip\right] + \left[ip, x\right]\right) = -i\left[x, p\right] = 1.$$

Exercise 7.1. *Prove that for any operators A, B, C,*

$$[AB, C] = [A, C]\, B + A\, [B, C]\,.$$

We define the so-called *number operator*

$$N \equiv a^\dagger a. \tag{7.6}$$

The commutators of the number operator and the creation or annihilation operators are

$$\left[N, a^\dagger\right] = \left[a^\dagger a, a^\dagger\right] = a^\dagger \left[a, a^\dagger\right] = a^\dagger \tag{7.7}$$

and

$$[N, a] = \left[a^\dagger a, a\right] = \left[a^\dagger, a\right] a = -a. \tag{7.8}$$

The Hamiltonian $H = N + \frac{1}{2}$ and the number operator have common eigenstates $|n\rangle$, since they commute.

If $|n\rangle$ is such an eigenstate with an eigenvalue n,

$$N\,|n\rangle = n\,|n\rangle\,, \tag{7.9}$$

the eigenvalue n cannot be negative. Indeed,

$$0 \le |a\,|n\rangle|^2 = \langle n|\, a^\dagger a\,|n\rangle = \langle n|\, N\,|n\rangle = n. \tag{7.10}$$

Also, the states $a^\dagger\,|n\rangle$ and $a\,|n\rangle$ are eigenstates of N with eigenvalues $n+1$ and $n-1$ respectively:

$$\begin{aligned} N\left(a^\dagger\,|n\rangle\right) &= \left(\left[N, a^\dagger\right] + a^\dagger N\right)|n\rangle = \left(a^\dagger + a^\dagger n\right)|n\rangle \\ &= (n+1)\left(a^\dagger\,|n\rangle\right), \end{aligned} \tag{7.11a}$$

$$\begin{aligned} N\,(a\,|n\rangle) &= ([N, a] + aN)\,|n\rangle = (-a + an)\,|n\rangle \\ &= (n-1)\,(a\,|n\rangle)\,. \end{aligned} \tag{7.11b}$$

Thus,

$$a\,|n\rangle = c_n\,|n-1\rangle\,, \quad a^\dagger\,|n\rangle = d_n\,|n+1\rangle\,.$$

We can fix the normalisation factors c_n, d_n by requiring that the states $|n\rangle$ have unit norm:

$$\begin{aligned} 1 = \langle n+1|n+1\rangle &= \frac{\langle n|\, aa^\dagger\,|n\rangle}{|d_n|^2} = \frac{\langle n|\,[a, a^\dagger] + a^\dagger a\,|n\rangle}{|d_n|^2} \\ &= \frac{\langle n|\, 1 + N\,|n\rangle}{|d_n|^2} = \frac{1+n}{|d_n|^2}. \end{aligned} \tag{7.12a}$$

Similarly,

$$\begin{aligned} 1 = \langle n-1|n-1\rangle &= \frac{\langle n|\, a^\dagger a\,|n\rangle}{|c_n|^2} = \frac{\langle n|\, N\,|n\rangle}{|c_n|^2} \\ &= \frac{n}{|c_n|^2}\,. \end{aligned} \tag{7.12b}$$

We can choose $c_n = \sqrt{n}$ and $d_n = \sqrt{n+1}$. With these choices, we have

$$a^\dagger |n\rangle = \sqrt{n+1}\,|n+1\rangle\,, \tag{7.13a}$$

and

$$a\,|n\rangle = \sqrt{n}\,|n-1\rangle\,. \tag{7.13b}$$

Notice that for non-integer n, the repeated application of the last equation leads eventually to an eigenstate with a negative eigenvalue, which is inconsistent with the requirement $n \geq 0$. The only allowed values of n are non-negative integers. Notice that the coefficient on the right-hand side of Eq. (7.13b) vanishes for $n = 0$, thus preventing us from obtaining a state $|-1\rangle$ with a negative eigenvalue. The spectrum of eigenvalues for the number operator and the Hamiltonian is discrete (quantised), with

$$n = 0, 1, 2, 3, \ldots,$$

and energy eigenvalues

$$E_n = n + \frac{1}{2}. \tag{7.14}$$

The eigenstates can all be generated from the "ground" state $|0\rangle$ by a repeated application of the creation operator:

$$\begin{aligned} |1\rangle &= a^\dagger |0\rangle \\ |2\rangle &= \frac{a^{\dagger 2}}{\sqrt{2!}} |0\rangle \\ &\vdots \\ |n\rangle &= \frac{a^{\dagger n}}{\sqrt{n!}} |0\rangle \\ &\vdots \end{aligned}$$

Let us focus on the ground state, for which we shall compute the expectation values

$$\langle x\rangle\,, \langle p\rangle\,, \langle x^2\rangle\,, \langle p^2\rangle\,.$$

We recall that

$$a = \frac{x + ip}{\sqrt{2}}, \quad a^\dagger = \frac{x - ip}{\sqrt{2}},$$

which, by inverting, yield

$$x = \frac{a + a^\dagger}{\sqrt{2}}, \quad p = \frac{a - a^\dagger}{\sqrt{2}i}. \tag{7.15}$$

We shall need that

$$a\,|0\rangle = 0, \quad \langle 0|\,a^\dagger = 0,$$

and

$$aa^\dagger |0\rangle = \underbrace{\left[a, a^\dagger\right]}_{1} |0\rangle + a^\dagger \underbrace{a\,|0\rangle}_{0} = |0\rangle\,.$$

It is then easy to compute that

$$\langle 0|\, x\, |0\rangle = \langle 0|\, p\, |0\rangle = 0, \tag{7.16}$$

and

$$\langle 0|\, x^2\, |0\rangle = \langle 0|\, p^2\, |0\rangle = \frac{1}{2}. \tag{7.17}$$

In the ground state, the product of the uncertainties in position and momentum is therefore minimal:

$$\left\langle (\Delta x)^2 \right\rangle \left\langle (\Delta p)^2 \right\rangle = \frac{1}{4}.$$

Exercise 7.2. *Show that for the $|n\rangle$ state, the product of uncertainties is*

$$\left\langle (\Delta x)^2 \right\rangle_n \left\langle (\Delta p)^2 \right\rangle_n = \left(n + \frac{1}{2}\right)^2. \tag{7.18}$$

We shall now determine the wave function in the ground state:

$$\psi_0(x) \equiv \langle x|0\rangle\,,$$

where

$$\hat{x}\,|x\rangle = x\,|x\rangle\,,$$

and

$$a\,|0\rangle = 0.$$

We have

$$\begin{aligned} 0 = \langle x|\, a\, |0\rangle = \langle x|\, \frac{\hat{x} + i\hat{p}}{\sqrt{2}}\, |0\rangle = \frac{x}{\sqrt{2}} \langle x|0\rangle + \frac{i}{\sqrt{2}} \langle x|\, \hat{p}\, |0\rangle \\ = \frac{x}{\sqrt{2}} \langle x|0\rangle + \frac{i}{\sqrt{2}} \left(-i \frac{\mathrm{d}}{\mathrm{d}x}\right) \langle x|0\rangle\,. \end{aligned} \tag{7.19}$$

We have thus derived the following differential equation for the wave function in the ground state:

$$\left(x + \frac{\mathrm{d}}{\mathrm{d}x}\right) \psi_0(x) = 0\,. \tag{7.20}$$

The solution of the above is

$$\psi_0(x) = A e^{-\frac{x^2}{2}}\,. \tag{7.21}$$

We can fix the normalisation by requiring

$$\begin{aligned} 1 = \langle 0|0\rangle = \int_{-\infty}^{\infty} \mathrm{d}x\ \langle 0|x\rangle \langle x|0\rangle = \int_{-\infty}^{\infty} \mathrm{d}x\ \left|\psi_0(x)\right|^2 \\ = |A|^2 \int_{-\infty}^{\infty} \mathrm{d}x\, e^{-x^2} \\ \rightsquigarrow |A| = \left(\int_{-\infty}^{\infty} \mathrm{d}x\, e^{-x^2} \right)^{-1/2}. \end{aligned} \tag{7.22}$$

The integral is

$$I = \int_{-\infty}^{\infty} \mathrm{d}x\, e^{-x^2} = \sqrt{\pi}.$$

We obtain this result as follows:

$$I^2 = \left(\int_{-\infty}^{\infty} \mathrm{d}x\, e^{-x^2}\right)^2 = \int_{-\infty}^{\infty} \mathrm{d}x\mathrm{d}y\, e^{-x^2-y^2}\,, \tag{7.23}$$

substituting $x = r\cos\theta, y = r\sin\theta$ gives

$$\begin{aligned} I^2 &= \int_0^{\infty} \mathrm{d}r\, r e^{-r^2} \int_0^{2\pi} \mathrm{d}\theta = 2\pi \int_0^{\infty} \mathrm{d}r\, r e^{-r^2} = 2\pi \int_0^{\infty} \mathrm{d}r\, \frac{-1}{2}\frac{\mathrm{d}}{\mathrm{d}r} e^{-r^2} \\ &= -\pi\, e^{-r^2}\Big|_{r=0}^{\infty} = \pi\,. \end{aligned} \tag{7.24}$$

We thus have that

$$\langle x|0\rangle = \frac{1}{\pi^{1/4}} e^{-x^2/2}. \tag{7.25}$$

We compute the wave functions of all energy eigenstates recursively,

$$\begin{aligned} \langle x|n\rangle &= \langle x|\, \frac{a^\dagger}{\sqrt{n}}\, |n-1\rangle \\ &= \frac{1}{\sqrt{n}} \int \mathrm{d}y\, \left\langle x\right| a^\dagger \left|y\right\rangle \left\langle y|n-1\right\rangle \\ &= \frac{1}{\sqrt{2n}} \int \mathrm{d}y\, \left\langle x\right| x - ip \left|y\right\rangle \left\langle y|n-1\right\rangle \\ &= \frac{1}{\sqrt{2n}} \left(x - \frac{\mathrm{d}}{\mathrm{d}x}\right) \langle x|n-1\rangle\,. \end{aligned} \tag{7.26}$$

Iterating, we obtain

$$\langle x|n\rangle = \frac{1}{\sqrt{2^n n!}} \left(x - \frac{\mathrm{d}}{\mathrm{d}x}\right)^n \langle x|0\rangle\,. \tag{7.27}$$

The solutions for $n = 0, 1, 2, 3$ can be seen in Fig. 7.1.

7.2 Time evolution

We now consider the position and momentum operators in the Heisenberg picture. They satisfy the Heisenberg equations of motion (recall that $\hbar = 1$),

$$\begin{aligned} i\frac{\partial}{\partial t} x &= [x, H] = \left[x, \frac{x^2+p^2}{2}\right] = \left[x, \frac{p^2}{2}\right] = [x,p]\,\frac{p}{2} + \frac{p}{2}\,[x,p] \\ &= ip\,. \end{aligned} \tag{7.28}$$

Similarly,

$$\begin{aligned} i\frac{\partial}{\partial t} p &= [p, H] = \left[p, \frac{x^2+p^2}{2}\right] = \left[p, \frac{x^2}{2}\right] = [p,x]\,\frac{x}{2} + \frac{x}{2}\,[p,x] \\ &= -ix\,, \end{aligned} \tag{7.29}$$

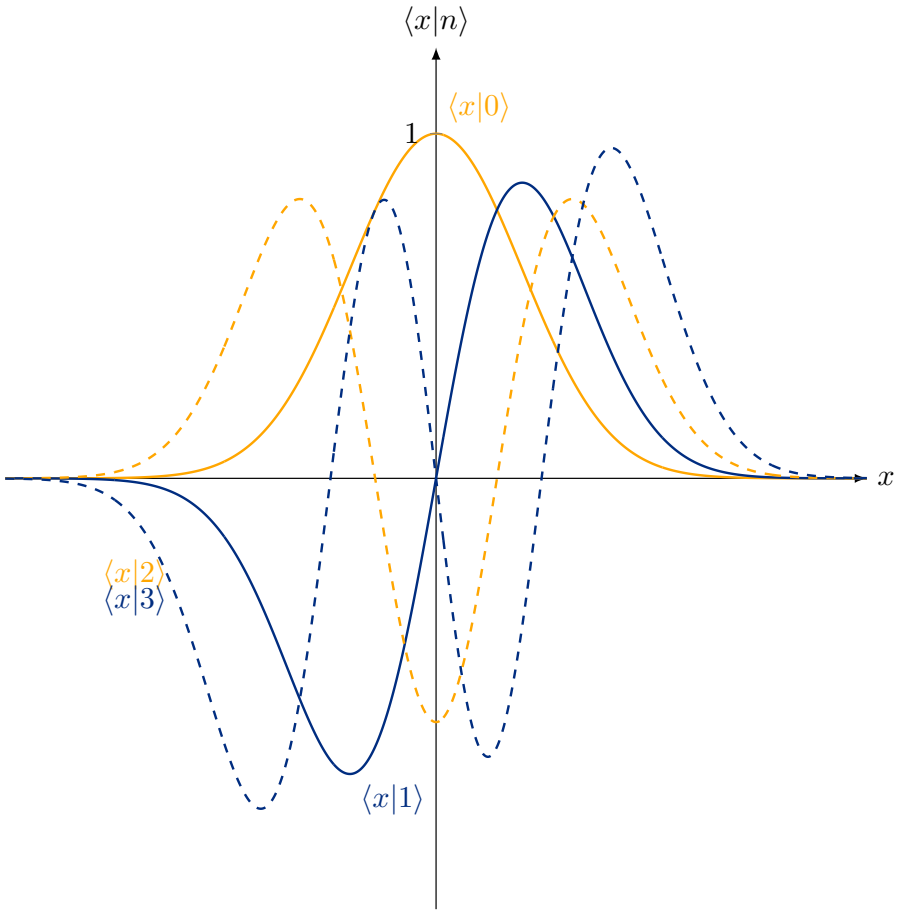


Figure 7.1. The wave functions of the energy levels $n = 0, 1, 2, 3$ of the quantum harmonic oscillator.

where we have used

$$[A, BC] = [A, B]\, C + B\, [A, C]\,. \tag{7.30}$$

Thus,

$$\frac{\partial x(t)}{\partial t} = p(t)\,, \qquad \frac{\partial p(t)}{\partial t} = -x(t)\,. \tag{7.31}$$

The solution of these equations is

$$x(t) = x(0)\cos t + p(0)\sin t \tag{7.32a}$$

$$p(t) = -x(0)\sin t + p(0)\cos t \tag{7.32b}$$

The above are all operator equations. Notice that for any energy eigenstate, the expectation values of the momentum and position operators vanish. Indeed,

$$\langle n|\, a^\dagger\, |n\rangle = \langle n|\, a\, |n\rangle = 0 \tag{7.33}$$

and

$$x(0) = \frac{a + a^\dagger}{\sqrt{2}}\,, \quad p(0) = \frac{a - a^\dagger}{\sqrt{2}i} \tag{7.34}$$

and therefore,

$$\langle n|\, x(t)\, |n\rangle = \langle n|\, p(t)\, |n\rangle = 0\,. \tag{7.35}$$

Unlike the position and momentum of a classical harmonic oscillator, the expectation values of their analogue quantum operators do not oscillate in an energy eigenstate.

7.3 Coherent states

Are there quantum states in which the expectation values of the position and momentum operators oscillate as in the classical system? Yes. These states are called coherent states. Consider a state $|\lambda\rangle$ which is an eigenstate of the annihilation operator a:

$$a\, |\lambda\rangle = \lambda\, |\lambda\rangle\,. \tag{7.36}$$

We can write a coherent state as a superposition of energy eigenstates:

$$|\lambda\rangle = \sum_{n=0}^{\infty} f(n)\, |n\rangle\,. \tag{7.37}$$

Then,

$$\begin{aligned} a\, |\lambda\rangle &= \sum_{n=1}^{\infty} f(n)\sqrt{n}\, |n-1\rangle \\ &= \sum_{n=0}^{\infty} f(n+1)\sqrt{n+1}\, |n\rangle\,. \end{aligned}$$

Substituting into Eq. (7.36), we find

$$\sum_{n=0}^{\infty}\left[f(n+1)\sqrt{n+1}-\lambda f(n)\right]|n\rangle=0. \tag{7.38}$$

Since the set of $|n\rangle$ eigenstates is a complete basis of Hilbert space, we must have the recurrence relation

$$f(\ell+1)=f(\ell)\frac{\lambda}{\sqrt{1+\ell}},\quad \ell=0,1,2,\ldots. \tag{7.39}$$

We can solve this recurrence relation easily:

$$\begin{aligned}\prod_{\ell=0}^{n-1}f(\ell+1)&=\prod_{\ell=0}^{n-1}f(\ell)\frac{\lambda}{\sqrt{1+\ell}}\\ \rightsquigarrow f(n)\left(\prod_{\ell=1}^{n-1}f(\ell)\right)&=f(0)\left(\prod_{\ell=1}^{n-1}f(\ell)\right)\frac{\lambda^n}{\sqrt{n!}}\\ \rightsquigarrow f(n)&=f(0)\frac{\lambda^n}{\sqrt{n!}}.\end{aligned} \tag{7.40}$$

Therefore, the coherent state $|\lambda\rangle$ is

$$|\lambda\rangle=f(0)\sum_{n=0}^{\infty}\frac{\lambda^n}{\sqrt{n!}}|n\rangle. \tag{7.41}$$

We fix the constant $f(0)$ by normalising the state,

$$\begin{aligned}1=\langle\lambda|\lambda\rangle&=\left|f(0)\right|^2\sum_{n=0}^{\infty}\sum_{m=0}^{\infty}\frac{\lambda^n}{\sqrt{n!}}\frac{(\lambda^*)^m}{\sqrt{m!}}\langle m|n\rangle\\ &=\left|f(0)\right|^2\sum_{n=0}^{\infty}\sum_{m=0}^{\infty}\frac{\lambda^n}{\sqrt{n!}}\frac{(\lambda^*)^m}{\sqrt{m!}}\delta_{nm}\\ &=\left|f(0)\right|^2\sum_{n=0}^{\infty}\frac{(|\lambda|^2)^n}{n!}\\ &=\left|f(0)\right|^2 e^{|\lambda|^2}\\ \rightsquigarrow f(0)&=e^{-|\lambda|^2/2}\end{aligned} \tag{7.42}$$

Therefore, the normalised coherent state takes the form (up to a choice of an overall phase)

$$|\lambda\rangle=e^{-|\lambda|^2/2}\sum_{n=0}^{\infty}\frac{\lambda^n}{\sqrt{n!}}|n\rangle. \tag{7.43}$$

The probability that a particle in a coherent state $|\lambda\rangle$ is measured to have an energy $E_n=n+\frac{1}{2}$ is

$$P\left(|\lambda\rangle\to|n\rangle\right)=|\langle n|\lambda\rangle|^2=e^{-|\lambda|^2}\frac{\left(|\lambda|^2\right)^n}{n!}, \tag{7.44}$$

which is a Poisson distribution with mean value

$$\begin{aligned}\langle n\rangle &= \langle\lambda| N |\lambda\rangle \\ &= |\lambda|^2 .\end{aligned} \tag{7.45}$$

Notice that for $\lambda = 0$ in Eq. (7.43) the coherent state becomes the ground state (first energy eigenstate) of the harmonic oscillator:

$$|\lambda = 0\rangle = |n = 0\rangle . \tag{7.46}$$

We can work further on Eq. (7.43),

$$\begin{aligned}|\lambda\rangle &= e^{-|\lambda|^2/2} \sum_{n=0}^{\infty} \frac{\lambda^n}{\sqrt{n!}} |n\rangle \\ &= e^{-|\lambda|^2/2} \sum_{n=0}^{\infty} \frac{\lambda^n}{\sqrt{n!}} \frac{(a^\dagger)^n}{\sqrt{n!}} |0\rangle \\ &= e^{-|\lambda|^2/2} \left(\sum_{n=0}^{\infty} \frac{(\lambda a^\dagger)^n}{n!} \right) |0\rangle ,\end{aligned}$$

leading to

$$|\lambda\rangle = e^{-|\lambda|^2/2} e^{\lambda a^\dagger} |0\rangle . \tag{7.47}$$

The expectation value of the energy in a coherent state is

$$\begin{aligned}\langle\lambda| H |\lambda\rangle &= \langle\lambda| a^\dagger a |\lambda\rangle + \langle\lambda| \frac{1}{2} |\lambda\rangle \\ &= |a |\lambda\rangle|^2 + \frac{1}{2} \\ &= |\lambda |\lambda\rangle|^2 + \frac{1}{2} \\ \rightsquigarrow \langle\lambda| H |\lambda\rangle &= |\lambda|^2 + \frac{1}{2}.\end{aligned} \tag{7.48}$$

In what follows, x and p denote the Heisenberg-picture operators $x(t)$, $p(t)$ of the previous section, expanded with $x(t) = x(0)\cos t + p(0)\sin t$; equivalently, one may evolve the state, $|\lambda\rangle \to \left|\lambda e^{-it}\right\rangle$ up to a phase (Sakurai, Sec. 2.3). This applies also to $\langle\lambda| p |\lambda\rangle$, $\langle\lambda| x^2 |\lambda\rangle$ and $\langle\lambda| p^2 |\lambda\rangle$ below. The expectation value of the position operator in a coherent state is

$$\begin{aligned}\langle\lambda| x |\lambda\rangle &= \langle\lambda| \frac{a + a^\dagger}{\sqrt{2}} |\lambda\rangle \cos t + \langle\lambda| \frac{a - a^\dagger}{i\sqrt{2}} |\lambda\rangle \sin t \\ &= \frac{\lambda + \lambda^*}{\sqrt{2}} \cos t + \frac{\lambda - \lambda^*}{i\sqrt{2}} \sin t .\end{aligned} \tag{7.49}$$

For

$$\lambda = |\lambda| e^{i\theta} = |\lambda| (\cos\theta + i \sin\theta) ,$$

we obtain

$$\langle\lambda| x |\lambda\rangle = \sqrt{2} |\lambda| \cos(t - \theta). \tag{7.50}$$

The expectation value of the position in a coherent state oscillates exactly as the position in a classical harmonic oscillator, with amplitude $\sqrt{2}\,|\lambda|$. Performing a similar computation, we find for the expectation value of the momentum

$$\langle\lambda|\,p\,|\lambda\rangle = -\sqrt{2}\,|\lambda|\sin(t-\theta). \tag{7.51}$$

Carrying on to the second powers of the position and momentum operators we find (**exercise**) the following results:

$$\langle\lambda|\,x^2\,|\lambda\rangle = \frac{1}{2} + 2\,|\lambda|^2\cos^2(t-\theta), \tag{7.52}$$

$$\langle\lambda|\,p^2\,|\lambda\rangle = \frac{1}{2} + 2\,|\lambda|^2\sin^2(t-\theta). \tag{7.53}$$

For the uncertainties in position and momentum, we have

$$\left\langle(\Delta x)^2\right\rangle = \langle\lambda|\,x^2\,|\lambda\rangle - \langle\lambda|\,x\,|\lambda\rangle^2 = \frac{1}{2} \tag{7.54}$$

and

$$\left\langle(\Delta p)^2\right\rangle = \langle\lambda|\,p^2\,|\lambda\rangle - \langle\lambda|\,p\,|\lambda\rangle^2 = \frac{1}{2}. \tag{7.55}$$

The product of the uncertainties in position and momentum for a particle in a coherent state is therefore exactly the minimum allowed by Heisenberg's uncertainty principle:

$$\left\langle(\Delta x)^2\right\rangle\left\langle(\Delta p)^2\right\rangle = \frac{1}{4}. \tag{7.56}$$

8

Schrödinger's wave equation

Consider a particle in a quantum state $|\psi, t\rangle$ and subject to the Hamiltonian

$$H = \frac{\mathbf{p}^2}{2m} + V(\mathbf{r}), \tag{8.1}$$

where $\mathbf{p}, \mathbf{r}$ are the momentum and position operators. The time evolution of the particle's state is determined by Schrödinger's equation,

$$i\hbar \frac{\partial}{\partial t} |\psi, t\rangle = H |\psi, t\rangle . \tag{8.2}$$

The traditional *Schrödinger wave equation* is the application of the above to the position representation of the state $|\psi, t\rangle$,

$$\psi(\mathbf{r}, t) = \langle \mathbf{r} | \psi, t \rangle ,$$

called the wave function. We have

$$\begin{aligned} &\langle \mathbf{r} | \, i\hbar \frac{\partial}{\partial t} |\psi, t\rangle = \langle \mathbf{r} | H |\psi, t\rangle \\ \rightsquigarrow \; &i\hbar \frac{\partial}{\partial t} \psi(\mathbf{r}, t) = \langle \mathbf{r} | \frac{\mathbf{p}^2}{2m} + V(\mathbf{r}) |\psi, t\rangle \\ \rightsquigarrow \; &i\hbar \frac{\partial}{\partial t} \psi(\mathbf{r}, t) = \left[-\frac{\hbar^2}{2m} \nabla^2 + V(\mathbf{r}) \right] \psi(\mathbf{r}, t) . \end{aligned} \tag{8.3}$$

Eq. (8.3) is Schrödinger's wave equation.

For time-independent Hamiltonians, the time evolution of energy eigenstates is particularly simple. If

$$H |\psi_E, t = 0\rangle = E |\psi_E, t = 0\rangle , \tag{8.4}$$

then

$$\begin{aligned} |\psi_E, t\rangle &= e^{-iEt/\hbar} |\psi_E, t = 0\rangle \\ \rightsquigarrow \langle \mathbf{r} | \psi_E, t \rangle &= e^{-iEt/\hbar} \langle \mathbf{r} | \psi_E, t = 0 \rangle \\ \rightsquigarrow \psi_E(\mathbf{r}, t) &= e^{-iEt/\hbar} \psi_E(\mathbf{r}). \end{aligned} \tag{8.5}$$

Substituting into Eq. (8.3), we obtain

$$\left[-\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{r})\right]\psi_E(\mathbf{r}) = E\psi_E(\mathbf{r}). \tag{8.6}$$

This is Schrödinger's time-independent wave equation.

Exercise 8.1. *Solve Eq. (8.6) for the wave functions of the energy eigenstates of the harmonic oscillator.*

8.1 Probability density and probability current

Consider a particle in a state $\left|\psi, t\right\rangle$ at a time t. The probability amplitude that the particle is found at a position $\mathbf{r}$ is the wave function,

$$\psi(\mathbf{r}, t) = \left\langle \mathbf{r}\middle|\psi, t\right\rangle.$$

From the normalisation of the state we have

$$\begin{aligned} 1 = \left\langle \psi, t\middle|\psi, t\right\rangle &= \int \mathrm{d}^3\mathbf{r}\, \left\langle \psi, t\middle|\mathbf{r}\right\rangle \left\langle \mathbf{r}\middle|\psi, t\right\rangle \\ &= \int \mathrm{d}^3\mathbf{r}\, \left|\psi(\mathbf{r}, t)\right|^2. \end{aligned} \tag{8.7}$$

We can interpret the quantity

$$\rho(\mathbf{r}, t) = \left|\psi(\mathbf{r}, t)\right|^2, \tag{8.8}$$

as the probability density for the particle to be found at the position $\mathbf{r}$.

From Schrödinger's wave equation, Eq. (8.3), we can derive

$$i\hbar\psi^*\partial_t\psi = -\frac{\hbar^2}{2m}\psi^*\nabla^2\psi + V(\mathbf{r})\left|\psi\right|^2, \tag{8.9}$$

$$i\hbar\psi\partial_t\psi^* = \frac{\hbar^2}{2m}\psi\nabla^2\psi^* - V(\mathbf{r})\left|\psi\right|^2. \tag{8.10}$$

Adding the two equations, we obtain

$$\begin{aligned} \partial_t(\psi^*\psi) &= -\frac{\hbar}{2mi}\left(\psi^*\nabla^2\psi - \psi\nabla^2\psi^*\right) \\ &= -\nabla\cdot\left[\frac{\hbar}{2mi}\left(\psi^*\nabla\psi - \psi\nabla\psi^*\right)\right]. \end{aligned} \tag{8.11}$$

We have derived a *continuity equation* for the probability density $\rho = \left|\psi\right|^2$,

$$\frac{\partial\rho}{\partial t} + \nabla\cdot\mathbf{j} = 0, \tag{8.12}$$

with the probability current

$$\mathbf{j} = \frac{\hbar}{2mi}\left[\psi^*\nabla\psi - \psi\nabla\psi^*\right]. \tag{8.13}$$

Let us now write

$$\psi(\mathbf{r},t) = \sqrt{\rho(\mathbf{r},t)}\cdot e^{iF(\mathbf{r},t)/\hbar}, \tag{8.14}$$

where the function $F(\mathbf{r},t)$ is real. Then,

$$\psi^*\nabla\psi - \psi\nabla\psi^* = \frac{2i}{\hbar}\rho\nabla F. \tag{8.15}$$

The probability current is then

$$\mathbf{j}(\mathbf{r},t) = \rho(\mathbf{r},t)\frac{\nabla F(\mathbf{r},t)}{m}. \tag{8.16}$$

We can regard $\nabla F/m$ as a "velocity".

Integrating the probability current over all space we find that it is equal to the expectation value of the momentum divided by the mass:

$$\begin{aligned}\int \mathrm{d}^3\mathbf{r}\,\mathbf{j}(\mathbf{r},t) &= \int \mathrm{d}^3\mathbf{r}\,\frac{1}{2m}\left[\psi^*(-i\hbar\nabla)\psi + \psi(i\hbar\nabla)\psi^*\right]\\ &= \int \mathrm{d}^3\mathbf{r}\,\frac{1}{2m}\left[\langle\psi,t|\mathbf{r}\rangle\langle\mathbf{r}|\,\mathbf{p}\,|\psi,t\rangle + \langle\psi,t|\,\mathbf{p}\,|\mathbf{r}\rangle\langle\mathbf{r}|\psi,t\rangle\right]\\ &= \frac{1}{m}\langle\psi,t|\,\mathbf{p}\,|\psi,t\rangle\,.\end{aligned} \tag{8.17}$$

8.2 Quantised energy levels

One of the biggest triumphs of quantum mechanics is the prediction of bound states, where particles are restricted to small regions of space (atoms, molecules) and have a discrete energy spectrum. In this section, we shall try to understand these phenomena qualitatively from the wave equation of Schrödinger. Consider a potential well (i.e. a potential which

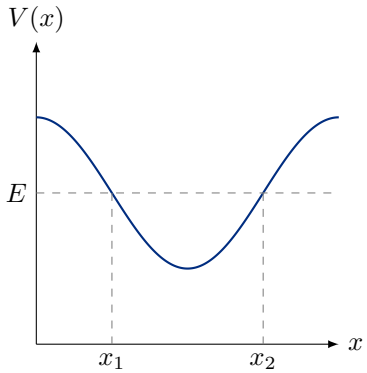
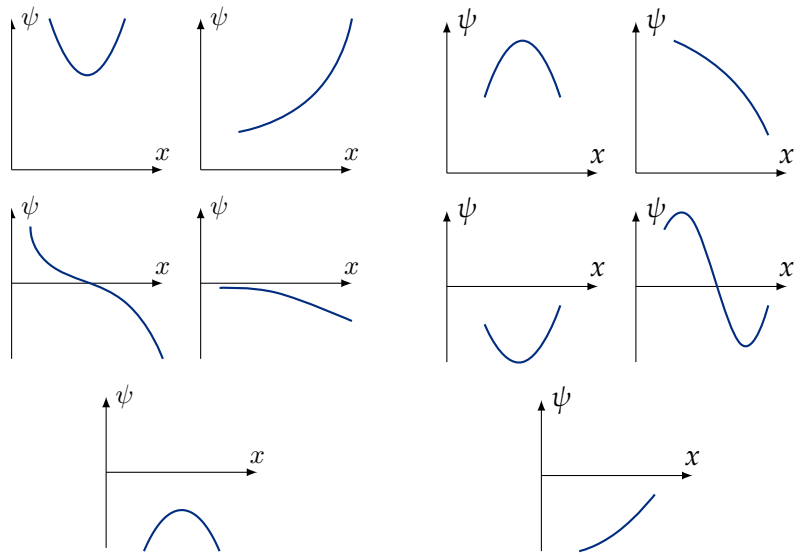


Figure 8.1. A potential well.

can trap a particle, as in Fig. 8.1) in one space dimension. Schrödinger's equation can be written in the form

$$\frac{\psi''}{\psi} = \frac{2m\,(V(x)-E)}{\hbar^2}\,. \tag{8.18}$$

The second derivative of the wave function is proportional to the difference between the potential and the energy of the particle. Geometrically, the second derivative represents the rate of change of the slope of the wave function.



(a) In regions of x where the potential energy is higher than the energy of the particle, the wave function is concave *away* from the x-axis.

(b) Otherwise it is concave *towards* the axis.

Figure 8.2. Examples of Schrödinger wave functions.

i) If $V(x) > E$, then ψ'' and ψ must have the *same* sign. In that case, the wave function is concave *away* from the axis (as depicted in the left panel of Fig. 8.2).

ii) If $V(x) < E$, then ψ'' and ψ must have opposite signs. In that case, the wave function is concave *towards* the axis (as depicted in the right panel of Fig. 8.2).

Let us apply the above qualitative principles to a bound state, where

$$\psi(-\infty) = 0\,. \tag{8.19}$$

Drawing wave functions with the above rules, starting from $x = -\infty$ and guessing the wave function all the way up to $x = +\infty$, we realise that it is not easy to end up with a wave function which vanishes at $x = +\infty$ (see Fig. 8.3).

In order to get a physical solution with finite probabilities, it is necessary to tune the energy very precisely. Only for selected values of the energy are we able to get physical solutions like the one in Fig. 8.4. The difficulty of matching a physical wave function over the full range of space explains the quantisation of energy. As we increase the energy level of the bound particle, the wave function in the region $V(x) < E$ oscillates more strongly and over a larger range. Higher energy levels lead to more crossings of the wave function with the x-axis (Fig. 8.5).

8.3 Semiclassical approximation

In this section we shall develop an approximation method in order to compute the wave function of a particle in a static system. This means that we demand that the probability of finding the particle in a certain region is the same irrespective of the time at which we

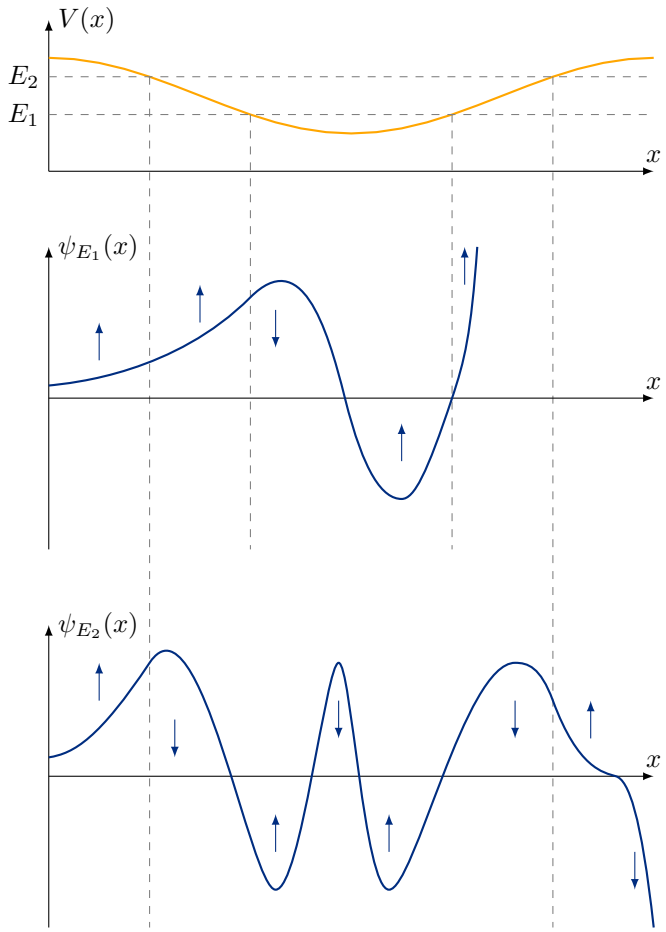


Figure 8.3. A few failed attempts to arrive at a vanishing wave function as $x \to +\infty$, starting from a physical behaviour at $x = -\infty$.

perform our experiment. For such a system, we must have

$$\frac{\partial \rho}{\partial t} = 0. \tag{8.20}$$

The continuity equation, Eq. (8.12), gives

$$\nabla \cdot \mathbf{j} = 0. \tag{8.21}$$

For simplicity, we will restrict our discussion to one-dimensional systems. Writing the wave function in the form

$$\psi(x) = \sqrt{\rho(x)} e^{iS(x)/\hbar}, \tag{8.22}$$

with a probability current

$$j(x) = \rho(x) \frac{\frac{\partial}{\partial x} S(x)}{m}, \tag{8.23}$$

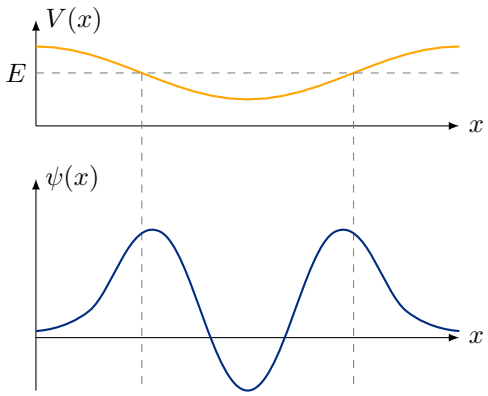


Figure 8.4. Only for selected (quantised) values of the energy are we able to match a physical behaviour of the wave function at $x \to +\infty$ with a physical behaviour at $x = -\infty$.

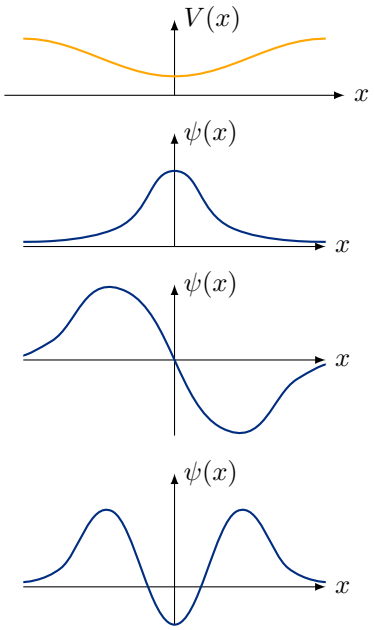


Figure 8.5. Higher energy levels correspond to wave functions with more crossings with the x-axis.

we have from Eq. (8.21) that

$$\begin{aligned} \frac{\partial}{\partial x}\left(\rho\frac{\partial S}{\partial x}\right) &= 0 \\ \rightsquigarrow \rho S' &= \text{const} \\ \rightsquigarrow \rho &= \frac{\text{const}}{S'}\,. \end{aligned} \tag{8.24}$$

Let us assume that $\hbar$ is a small parameter ($\hbar \to 0$ corresponds to the classical limit) and expand:

$$S = S_0 + \hbar S_1 + \hbar^2 S_2 + \dots\,. \tag{8.25}$$

Notice that the derivatives of the wave function are proportional to the wave function itself (as follows by differentiating Eq. (8.22)):

$$\psi' = \psi\left[\frac{\sqrt{\rho}\,'}{\sqrt{\rho}} + i\frac{S'}{\hbar}\right], \tag{8.26}$$

$$\psi'' = \psi\left[\left(\frac{\sqrt{\rho}'}{\sqrt{\rho}} + i\frac{S'}{\hbar}\right)^2 + \left(\frac{\sqrt{\rho}'}{\sqrt{\rho}}\right)' + i\frac{S''}{\hbar}\right]. \tag{8.27}$$

Substituting these, with the expansion of Eq. (8.25), into Eq. (8.18) and keeping the leading-order term in $\hbar$ (semiclassical approximation), we obtain the following differential equation:

$$\left(S_0'\right)^2 = -2m\left(V(x) - E\right). \tag{8.28}$$

For the regions of x where $V(x) > E$ we have

$$S_0'(x) = \pm i\sqrt{2m(V(x) - E)}, \tag{8.29}$$

and, thus,

$$\psi_{(V(x)>E)} \approx \frac{\text{const}}{\left[2m\left(V(x) - E\right)\right]^{\frac{1}{4}}} \exp\left(\pm\int_{x_0}^{x} \mathrm{d}y\, \frac{\sqrt{2m\left(V(y) - E\right)}}{\hbar}\right). \tag{8.30}$$

Similarly, for $V(x) < E$, we find an approximate solution:

$$S_0'(x) = \pm\sqrt{2m(E - V(x))} \tag{8.31}$$

and, thus,

$$\psi_{(V(x)<E)} \approx \frac{\text{const}}{[2m(E - V(x))]^{\frac{1}{4}}} \exp\left(\pm i\int_{x_0}^{x} \mathrm{d}y\, \frac{\sqrt{2m\left(E - V(y)\right)}}{\hbar}\right). \tag{8.32}$$

Remark. Strictly, the prefactor $1/\sqrt{S_0'}$ is the next order, $\hbar S_1$, of Eq. (8.25), with $S_1' = \frac{i}{2}S_0''/S_0'$; the argument via $\rho S' = \text{const}$ reproduces it only where S_0 is real ($V < E$). For $V > E$, S_0 is imaginary, the decomposition into a real phase and $\sqrt{\rho}$ does not apply (a real decaying ψ carries no current), and the amplitude $|2m(V - E)|^{-1/4}$ follows from the S_1 equation (Sakurai Sec. 2.4; Griffiths Ch. 9).

Obviously, neither of the two solutions is a good approximation where the potential is close to the energy, $V(x) \sim E$. In these regions, we can find an explicit solution by expanding the potential around the point $x = x_{sp}$ at which $V(x_{sp}) = E$. We write

$$\begin{aligned} V(x) &\approx V(x_{sp}) + (x - x_{sp})V'(x_{sp}) + \ldots \\ &= E + (x - x_{sp})V'(x_{sp}) + \ldots. \end{aligned} \tag{8.33}$$

The Hamiltonian becomes

$$H \approx \frac{\hat{p}^2}{2m} + E + (\hat{x} - x_{sp})V'(x_{sp}), \tag{8.34}$$

and we need to solve the eigenvalue problem

$$H\left|\psi_E\right\rangle = E\left|\psi_E\right\rangle$$
$$\rightsquigarrow \left[\frac{\hat{p}^2}{2m} + (\hat{x} - x_{sp})V'(x_{sp})\right]\left|\psi_E\right\rangle = 0\,. \tag{8.35}$$

The linear potential problem can be solved exactly, as we shall see in the next section.

Having approximate solutions for the three distinct regions $V(x) < E$, $V(x) > E$ and $V(x) \approx E$, we can construct an approximate solution for all values of x. The solutions are required to agree in the regions where more than one approximation is valid. This leads to "matching" conditions which, for bound states, lead to quantisation conditions for the energy levels.

8.4 Linear potential

In this section we shall solve Schrödinger's energy eigenstate problem for a linear potential:

$$\left[\frac{\hat{p}^2}{2m} + a\hat{x} + b\right]\left|\psi\right\rangle = 0. \tag{8.36}$$

It is easier to solve this problem in momentum space. Recall that

$$\left\langle x\middle|p\right\rangle = \frac{1}{\sqrt{2\pi\hbar}}e^{ipx/\hbar}. \tag{8.37}$$

It is easy to verify that this amplitude satisfies the correct momentum eigenstate equation:

$$\begin{aligned}\left\langle x\middle|\,\hat{p}\,\middle|p\right\rangle &= -i\hbar\frac{\partial}{\partial x}\left\langle x\middle|p\right\rangle = -i\hbar\frac{\partial}{\partial x}\frac{1}{\sqrt{2\pi\hbar}}e^{ipx/\hbar}\\ &= p\frac{1}{\sqrt{2\pi\hbar}}e^{ipx/\hbar} = p\left\langle x\middle|p\right\rangle.\end{aligned} \tag{8.38}$$

For two momentum states, $\left|p\right\rangle, \left|q\right\rangle$, we have

$$\begin{aligned}\left\langle p\middle|\,\hat{x}\,\middle|q\right\rangle &= \int \mathrm{d}y\ \left\langle p\middle|\,\hat{x}\,\middle|y\right\rangle\left\langle y\middle|q\right\rangle = \int \mathrm{d}y\ y\left\langle p\middle|y\right\rangle\left\langle y\middle|q\right\rangle\\ &= \int \mathrm{d}y\ \left\langle p\middle|y\right\rangle\left(-i\hbar\frac{\partial}{\partial q}\right)\frac{1}{\sqrt{2\pi\hbar}}e^{iqy/\hbar}\\ &= \left(-i\hbar\frac{\partial}{\partial q}\right)\int \mathrm{d}y\ \left\langle p\middle|y\right\rangle\left\langle y\middle|q\right\rangle\\ &= \left(-i\hbar\frac{\partial}{\partial q}\right)\left\langle p\middle|q\right\rangle = \left(-i\hbar\frac{\partial}{\partial q}\right)\delta(p-q).\end{aligned} \tag{8.39}$$

Also,

$$\begin{aligned}\left\langle p\middle|\,\hat{x}\,\middle|\psi\right\rangle &= \int \mathrm{d}q\left\langle p\middle|\,\hat{x}\,\middle|q\right\rangle\left\langle q\middle|\psi\right\rangle\\ &= +i\hbar\frac{\partial}{\partial p}\int \mathrm{d}q\,\delta(q-p)\left\langle q\middle|\psi\right\rangle\\ \rightsquigarrow \left\langle p\middle|\,\hat{x}\,\middle|\psi\right\rangle &= +i\hbar\frac{\partial}{\partial p}\left\langle p\middle|\psi\right\rangle.\end{aligned} \tag{8.40}$$

Thus,

$$\begin{aligned}
&\left\langle p\right| \frac{\hat{p}^2}{2m} + a\hat{x} + b \left|\psi\right\rangle = 0 \\
&\quad \rightsquigarrow i\hbar a \frac{\partial}{\partial p} \left\langle p|\psi\right\rangle = -\left(\frac{p^2}{2m} + b\right) \left\langle p|\psi\right\rangle \\
&\quad\quad \rightsquigarrow \left\langle p|\psi\right\rangle = N \exp\left(-\frac{\frac{p^3}{6m} + pb}{i\hbar a}\right).
\end{aligned} \tag{8.41}$$

The wave function in x-space can be found as

$$\begin{aligned}
\left\langle x|\psi\right\rangle &= \int_{-\infty}^{\infty} \mathrm{d}p \left\langle x|p\right\rangle \left\langle p|\psi\right\rangle \\
&= N \int_{-\infty}^{\infty} \frac{\mathrm{d}p}{\sqrt{2\pi\hbar}} e^{ipx/\hbar} \exp\left(-\frac{\frac{p^3}{6m} + pb}{i\hbar a}\right).
\end{aligned} \tag{8.42}$$

The solution is an Airy function. The Airy function is defined through its integral representation,

$$\mathrm{Ai}(z) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \mathrm{d}t \, \exp\left[i\left(\frac{t^3}{3} + zt\right)\right], \tag{8.43}$$

and can be computed numerically for all values of z. **Remark.** Rescaling $p = (2m\hbar a)^{1/3} t$ gives $\left\langle x|\psi\right\rangle \propto \mathrm{Ai}((2m/\hbar^2 a^2)^{1/3}(ax+b))$, which decays for $ax + b > 0$, i.e. in the classically forbidden region $V - E > 0$, and oscillates for $ax + b < 0$, as it must; indeed, $\mathrm{Ai}''(z) = z\,\mathrm{Ai}(z)$ (Sakurai Sec. 2.5; Landau–Lifshitz QM §24).

9

Particle in a constant electromagnetic field

The Hamiltonian for a charged particle in a time-independent electromagnetic field is

$$H = \frac{\left(\mathbf{p} - q\mathbf{A}\right)^2}{2m} + q\Phi, \tag{9.1}$$

where $\mathbf{A} = \mathbf{A}(\mathbf{r})$ and $\Phi = \Phi(\mathbf{r})$ are the vector and scalar potentials, respectively.

Working in the Heisenberg picture, we find that the velocity operator is (**exercise**)

$$\frac{\mathrm{d}r_i}{\mathrm{d}t} = -\frac{i}{\hbar}\left[r_i, H\right] = \frac{p_i - qA_i}{m}, \tag{9.2}$$

and the operator analogue of the classical momentum is then

$$\mathbf{\Pi} = m\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t} = \mathbf{p} - q\mathbf{A}. \tag{9.3}$$

We call $\mathbf{\Pi}$ the kinetic momentum and $\mathbf{p}$ the canonical momentum. For the commutator of the kinetic momenta we find

$$\left[\Pi_i, \Pi_j\right] = i\hbar q\left(\partial_i A_j - \partial_j A_i\right). \tag{9.4}$$

Recalling that the magnetic field is given by

$$B_i = \epsilon_{ijk}\partial_j A_k, \tag{9.5}$$

we have

$$\epsilon_{ijk}B_k = \left(\partial_i A_j - \partial_j A_i\right). \tag{9.6}$$

We then write

$$\left[\Pi_i, \Pi_j\right] = i\hbar q\epsilon_{ijk}B_k. \tag{9.7}$$

We now compute the commutator

$$\begin{aligned}\left[\Pi_i, H\right] &= \left[\Pi_i, \frac{\Pi_j\Pi_j}{2m} + q\Phi\right] \\ &= i\hbar q\left(-\partial_i\Phi + \epsilon_{ijk}\frac{1}{2m}\left[\Pi_j B_k - B_j\Pi_k\right]\right).\end{aligned}$$

We then obtain, for the quantum operator corresponding to the classical force,

$$m\frac{\mathrm{d}^2\mathbf{r}}{\mathrm{d}t^2} = \frac{1}{i\hbar}\left[m\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t}, H\right] = q\left[-\nabla\Phi + \frac{1}{2}\left(\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t}\times\mathbf{B} - \mathbf{B}\times\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t}\right)\right]. \tag{9.8}$$

9.1 Wave function for a particle in an electromagnetic field

The time evolution of a state is determined by Schrödinger's equation, which in the case of a particle in a constant electromagnetic field takes the form

$$i\hbar \frac{\mathrm{d}}{\mathrm{d}t} \left|\psi, t\right\rangle = \left(\frac{\Pi^2}{2m} + q\Phi\right) \left|\psi, t\right\rangle . \tag{9.9}$$

For the wave function

$$\psi(\mathbf{r}, t) = \left\langle \mathbf{r} \middle| \psi, t \right\rangle ,$$

we obtain

$$i\hbar \frac{\partial \psi(\mathbf{r}, t)}{\partial t} = \left[-\frac{\hbar^2}{2m} \left(\nabla - \frac{i}{\hbar} qA \right) \left(\nabla - \frac{i}{\hbar} qA \right) + q\Phi \right] \psi(\mathbf{r}, t). \tag{9.10}$$

We define a "covariant" gradient operator,

$$\nabla_c \equiv \nabla - \frac{i}{\hbar} qA. \tag{9.11}$$

We then rewrite Schrödinger's equation as

$$i\hbar \frac{\partial \psi(\mathbf{r}, t)}{\partial t} = \left[-\frac{\hbar^2}{2m} \nabla_c^2 + q\Phi \right] \psi(\mathbf{r}, t). \tag{9.12}$$

Maxwell's equations are invariant under gauge transformations, which in the case of static fields take the form

$$\mathbf{A}(\mathbf{r}) \longrightarrow \mathbf{A}(\mathbf{r}) + \nabla \Lambda(\mathbf{r}). \tag{9.13}$$

Under this transformation, Schrödinger's equation becomes

$$i\hbar \frac{\partial \psi_\Lambda(\mathbf{r}, t)}{\partial t} = \left[-\frac{\hbar^2}{2m} \left(\nabla_c - \frac{iq}{\hbar} \nabla \Lambda \right)^2 + q\Phi \right] \psi_\Lambda(\mathbf{r}, t). \tag{9.14}$$

We would now like to compare the solutions of the Schrödinger equation before and after the gauge transformation. We can easily verify that if $\psi(\mathbf{r}, t)$ is a solution of Eq. (9.12), then the function

$$\psi_\Lambda(\mathbf{r}, t) = e^{\frac{iq}{\hbar}\Lambda(\mathbf{r})} \psi(\mathbf{r}, t) \tag{9.15}$$

is also a solution of Eq. (9.14). Indeed,

$$\left(\nabla - \frac{iq\mathbf{A}}{\hbar} - \frac{iq\nabla\Lambda}{\hbar} \right) e^{\frac{iq}{\hbar}\Lambda(\mathbf{r})} = e^{\frac{iq}{\hbar}\Lambda(\mathbf{r})} \left(\nabla - \frac{iq\mathbf{A}}{\hbar} \right), \tag{9.16}$$

and

$$\left(\nabla_c - i\frac{q}{\hbar} \nabla\Lambda \right)^2 e^{\frac{iq}{\hbar}\Lambda(\mathbf{r})} \psi(\mathbf{r}, t) = e^{\frac{iq}{\hbar}\Lambda(\mathbf{r})} \nabla_c^2 \psi(\mathbf{r}, t). \tag{9.17}$$

Therefore, a gauge transformation results in the wave function being multiplied by an overall phase.

We shall now revisit our interpretation of the modulus squared of the wave function as a probability density,

$$\rho(\mathbf{r}, t) = \left|\psi(\mathbf{r}, t)\right|^2 .$$

Taking the time derivative, we have

$$\begin{aligned}\frac{\partial\rho}{\partial t} &= \psi^*\left(\partial_t\psi\right) + \left(\partial_t\psi^*\right)\psi \\ &= i\frac{\hbar}{2m}\left(\psi^*\nabla_c^2\psi - \psi\left(\nabla_c^2\psi\right)^*\right) \\ &= -\nabla\cdot\mathbf{J},\end{aligned} \tag{9.18}$$

where

$$\mathbf{J} = -i\frac{\hbar}{2m}\left[\psi^*\nabla_c\psi - \psi\nabla_c^*\psi^*\right], \tag{9.19}$$

and

$$\nabla_c^* = \nabla + \frac{i}{\hbar}qA\,. \tag{9.20}$$

Contrast this current with the one in the absence of the electromagnetic field. It is consistent with the substitution

$$\nabla \to \nabla_c,$$

or, equivalently, replacing the canonical momentum operator with the kinetic momentum operator,

$$\mathbf{p} \to \mathbf{\Pi}.$$

9.2 Aharonov–Bohm effect

Consider a particle that at $t = 0$ is found at the position $\mathbf{a}$, i.e. in the state $|\mathbf{a}\rangle$. After some time t, the state will have evolved according to Schrödinger's equation,

$$i\hbar\partial_t\,|\mathbf{a},t\rangle = \left[\frac{\Pi^2}{2m} + q\Phi\right]|\mathbf{a},t\rangle\,. \tag{9.21}$$

The probability amplitude that the particle is found at a position $\mathbf{r}$ after time t is then

$$\langle\mathbf{r}|\mathbf{a},t\rangle\,,$$

and it satisfies

$$i\hbar\partial_t\,\langle\mathbf{r}|\mathbf{a},t\rangle = \left[-\frac{\hbar^2}{2m}\nabla_c^2 + q\Phi\right]\langle\mathbf{r}|\mathbf{a},t\rangle\,. \tag{9.22}$$

Let us assume that we can compute the amplitude for the same transition,

$$\langle\mathbf{r}|\mathbf{a},t\rangle_0\,,$$

in the absence of a vector potential ($\mathbf{A} = 0 \rightsquigarrow \nabla_c = \nabla$). Then, the amplitude for the transition with $\mathbf{A} \neq 0$ is

$$\langle\mathbf{r}|\mathbf{a},t\rangle = \langle\mathbf{r}|\mathbf{a},t\rangle_0 \exp\left(\frac{iq}{\hbar}\int_{\mathbf{a}}^{\mathbf{r}}\mathrm{d}\mathbf{l}\cdot\mathbf{A}\right). \tag{9.23}$$

Remark. Eq. (9.23) requires $\mathbf{B} = \nabla\times\mathbf{A} = 0$ in the region where the particle propagates: only then is $\int_{\mathbf{a}}^{\mathbf{r}}\mathrm{d}\mathbf{l}\cdot\mathbf{A}$ independent of the path (locally $\mathbf{A} = \nabla\chi$, a pure gauge), and the formula is

the gauge transformation above with $\Lambda = \chi$. In a non-simply-connected field-free region (the Aharonov–Bohm geometry) it holds separately for each homotopy class of paths, which is how it is used below. The "multiplication rule" of the exercise below holds for these phase factors; for the full amplitudes one must integrate over the intermediate point, $\int \mathrm{d}^3 b \, \langle c|b\rangle \langle b|a\rangle$ (Sakurai Sec. 2.7).

Exercise 9.1. *Prove that the above is indeed a solution of Eq. (9.22) and that it satisfies the amplitude "multiplication rule" for two successive transitions:*

$$\mathcal{M}(a \to b)\mathcal{M}(b \to c) = \mathcal{M}(a \to c),$$

where $\mathcal{M}$ denotes

$$\mathcal{M}\left(x \to y\right) = \left\langle y \middle| x \right\rangle .$$

Now consider a two-slit experiment in which there is a magnetic field $\mathbf{B}$ between the two slits, inside the wall. The field is contained inside the wall, and the particles cannot penetrate it. The transition amplitude from the shooting point to a point on the screen behind the wall is the sum of the amplitude for the particle to go through slit 1 (path Γ_1) and the amplitude to go through slit 2 (path Γ_2).

$$\mathcal{M} = \mathcal{M}_{\Gamma_1} + \mathcal{M}_{\Gamma_2} . \tag{9.24}$$

According to Eq. (9.23), this is equivalent to

$$\mathcal{M} = \mathcal{M}_1 \exp\left(\frac{iq}{\hbar}\int_{\Gamma_1} \mathrm{d}\mathbf{l}\cdot\mathbf{A}\right) + \mathcal{M}_2 \exp\left(\frac{iq}{\hbar}\int_{\Gamma_2} \mathrm{d}\mathbf{l}\cdot\mathbf{A}\right), \tag{9.25}$$

where $\mathcal{M}_1$ and $\mathcal{M}_2$ are the corresponding amplitudes in the absence of the vector potential.

The probability of this transition is

$$\begin{aligned} P &= \left|\mathcal{M}\right|^2 \\ &= \left|\mathcal{M}_1\right|^2 + \left|\mathcal{M}_2\right|^2 + 2\Re\left(\mathcal{M}_1^*\mathcal{M}_2 \exp\left(\frac{iq}{\hbar}\oint \mathrm{d}\mathbf{l}\cdot\mathbf{A}\right)\right). \end{aligned} \tag{9.26a}$$

The interference term now depends on the flux of the magnetic field enclosed by the two alternative paths (Fig. 9.1),

$$\Phi_B = \oint \mathrm{d}\mathbf{l}\cdot\mathbf{A} = \int_{S(\Gamma_1+\Gamma_2)} \mathrm{d}\mathbf{S}\cdot(\nabla\times\mathbf{A}) = \int_{S(\Gamma_1+\Gamma_2)} \mathrm{d}\mathbf{S}\cdot\mathbf{B}, \tag{9.26b}$$

$$\rightsquigarrow P = \left|\mathcal{M}_1\right|^2 + \left|\mathcal{M}_2\right|^2 + 2\Re\left(\mathcal{M}_1^*\mathcal{M}_2 \exp\left(\frac{iq}{\hbar}\Phi_B\right)\right). \tag{9.26c}$$

It is remarkable that the path of the particles lies outside the range of the magnetic field. Nevertheless, the interference pattern is affected by it. This is the Aharonov–Bohm effect.

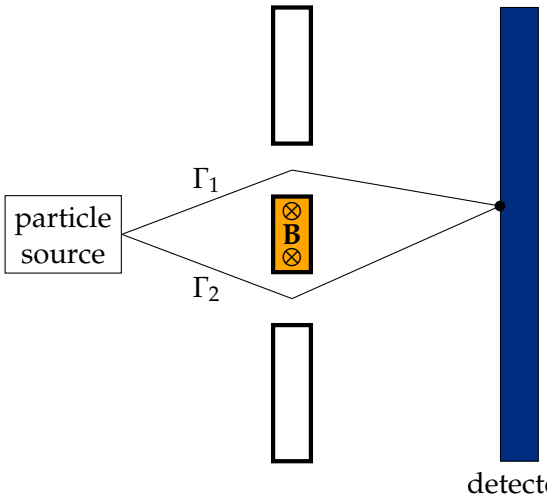


Figure 9.1. Sketch of the set-up for the Aharonov–Bohm effect. The **B**-field is only within the wall and is enclosed by the closed path $\Gamma_1 + \Gamma_2$.

10

Symmetries in quantum mechanics

A symmetry changes our way of looking at a physical system without changing the physics of it. In quantum mechanics, symmetry transformations alter quantum states,

$$|\psi\rangle \xrightarrow{T} |\psi'\rangle\,, \quad |\phi\rangle \xrightarrow{T} |\phi'\rangle\,, \tag{10.1}$$

without changing probabilities,

$$\left|\langle\phi'|\psi'\rangle\right|^2 = \left|\langle\phi|\psi\rangle\right|^2. \tag{10.2}$$

States can be represented as vectors in Hilbert space. If $\{|a_i\rangle\}$ is a set of base kets, we can write

$$|\psi\rangle = \sum_i c_i\,|a_i\rangle\,,$$

and we represent $|\psi\rangle$ as a Hilbert space vector

$$|\psi\rangle \doteq \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_N \end{pmatrix}.$$

Under the symmetry transformation T, the vector of $|\psi\rangle$ in Hilbert space transforms as

$$\begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_N \end{pmatrix} \xrightarrow{T} \begin{pmatrix} c_1' \\ c_2' \\ \vdots \\ c_N' \end{pmatrix} = \begin{pmatrix} U_{11}(T) & \dots & U_{1N}(T) \\ U_{21}(T) & \dots & U_{2N}(T) \\ \vdots & \ddots & \vdots \\ U_{N1}(T) & \dots & U_{NN}(T) \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_N \end{pmatrix},$$

or, briefly,

$$\underbrace{|\psi'\rangle}_{\text{Hilbert space vector}} \doteq \overbrace{U(T)}^{\text{matrix}}\ \underbrace{|\psi\rangle}_{\text{Hilbert space vector}}\,, \tag{10.3}$$

where the kets in the above equations are the vector representations of the states in Hilbert space and not the states themselves. The matrix $U(T)$ changes for other representations of the states, with sets of base states different from $\{|a_i\rangle\}$.

To preserve the probabilities under symmetry transformations, according to Wigner's theorem, we have two options:

i) $U(T)$ is linear and unitary,

$$U(a\left|\psi\right\rangle + b\left|\phi\right\rangle) = aU\left|\psi\right\rangle + bU\left|\phi\right\rangle$$
$$\left\langle\psi\right|U^\dagger U\left|\phi\right\rangle = \left\langle\psi|\phi\right\rangle,$$

ii) $U(T)$ is anti-linear and anti-unitary,

$$U(a\left|\psi\right\rangle + b\left|\phi\right\rangle) = a^*U\left|\psi\right\rangle + b^*U\left|\phi\right\rangle$$
$$\left\langle\psi\right|U^\dagger U\left|\phi\right\rangle = \left\langle\phi|\psi\right\rangle.$$

We shall not prove Wigner's theorem here.[1] We will restrict ourselves to verifying the converse statement. Suppose, for example, that U is anti-linear and anti-unitary. For two states,

$$\left|\phi\right\rangle, \quad \left|\psi\right\rangle = a_1\left|\psi_1\right\rangle + a_2\left|\psi_2\right\rangle, \tag{10.4}$$

the probability that $\left|\psi\right\rangle$ collapses into $\left|\phi\right\rangle$ is transformed into

$$\begin{aligned}
\left|\left\langle\phi'|\psi'\right\rangle\right|^2 &= \left|\left\langle\phi\right|U^\dagger U\left|\psi\right\rangle\right|^2 \\
&= \left|\left\langle\phi\right|U^\dagger U\left(a_1\left|\psi_1\right\rangle + a_2\left|\psi_2\right\rangle\right)\right|^2 \\
&= \left|\left\langle\phi\right|U^\dagger\left(a_1^*U\left|\psi_1\right\rangle + a_2^*U\left|\psi_2\right\rangle\right)\right|^2 \\
&= \left|a_1^*\left\langle\phi\right|U^\dagger U\left|\psi_1\right\rangle + a_2^*\left\langle\phi\right|U^\dagger U\left|\psi_2\right\rangle\right|^2 \\
&= \left|a_1^*\left\langle\psi_1|\phi\right\rangle + a_2^*\left\langle\psi_2|\phi\right\rangle\right|^2 \\
&= \left|\left(a_1^*\left\langle\psi_1\right| + a_2^*\left\langle\psi_2\right|\right)\left|\phi\right\rangle\right|^2 \\
&= \left|\left\langle\psi|\phi\right\rangle\right|^2 \\
&= \left|\left\langle\phi|\psi\right\rangle\right|^2.
\end{aligned}$$

Similarly, if U is linear and unitary,

$$\begin{aligned}
\left|\left\langle\phi'|\psi'\right\rangle\right|^2 &= \left|\left\langle\phi\right|U^\dagger U\left|\psi\right\rangle\right|^2 \\
&= \left|\left\langle\phi\right|U^\dagger U\left(a_1\left|\psi_1\right\rangle + a_2\left|\psi_2\right\rangle\right)\right|^2 \\
&= \left|\left\langle\phi\right|U^\dagger\left(a_1U\left|\psi_1\right\rangle + a_2U\left|\psi_2\right\rangle\right)\right|^2 \\
&= \left|a_1\left\langle\phi|\psi_1\right\rangle + a_2\left\langle\phi|\psi_2\right\rangle\right|^2 \\
&= \left|\left\langle\phi\right|\left(a_1\left|\psi_1\right\rangle + a_2\left|\psi_2\right\rangle\right)\right|^2 \\
&= \left|\left\langle\phi|\psi\right\rangle\right|^2.
\end{aligned}$$

[1] A proof can be found in S. Weinberg, *The Quantum Theory of Fields*, Vol. I.

10.1 Continuous symmetry transformations

Notice that the unit operator is unitary and linear, not anti-unitary and anti-linear. Some symmetry transformations $T(\theta^a)$, such as rotations, translations and Lorentz boosts, are connected smoothly to the identity via continuous parameters (angles, velocities, etc.),

$$T \equiv T(\theta^a), \quad \theta^a \equiv \begin{cases} \theta_1, \theta_2, \theta_3 & \text{angles} \\ \beta_x, \beta_y, \beta_z & \text{boosts} \\ \varepsilon_1, \varepsilon_2, \varepsilon_3 & \text{displacements} \end{cases} . \tag{10.5}$$

Such transformations, which are continuously connected to the identity (no transformation), must have a unitary and linear representation,

$$\left|\psi'\right\rangle = U\left(T(\theta^a)\right)\left|\psi\right\rangle . \tag{10.6}$$

We choose the parameters θ^a of these continuous transformations so that when the parameters of the transformation vanish, $\theta^a = 0$, the system is unchanged:

$$U\left(T(\theta^a = 0)\right) = 1. \tag{10.7}$$

Let us restrict ourselves to unitary and linear transformations $T \equiv T(\theta^a)$ and study their representation $U(T)$. The product of two such transformations is also a transformation,

$$T_3 = T_2 T_1 .$$

A state $\left|\psi\right\rangle$ transforms as

$$\left|\psi\right\rangle \to \left|\psi'\right\rangle = U(T_2)U(T_1)\left|\psi\right\rangle ,$$

under the T_1 transformation followed by the T_2 transformation. Transforming under $T_3 = T_2 T_1$, we obtain

$$\left|\psi\right\rangle \to \left|\psi''\right\rangle = U(T_3)\left|\psi\right\rangle = U(T_2 T_1)\left|\psi\right\rangle .$$

The states $\left|\psi'\right\rangle$ and $\left|\psi''\right\rangle$ must be equivalent, describing the same physics. Therefore, they can only differ by a phase:

$$U(T_2)U(T_1)\left|\psi\right\rangle = e^{i\phi_\psi(T_2,T_1)}U(T_2T_1)\left|\psi\right\rangle . \tag{10.8}$$

We can show that the phase $\phi_\psi(T_2, T_1)$ is independent of the state $\left|\psi\right\rangle$. Take a state $\left|\psi\right\rangle$ which is a superposition of two linearly independent states,

$$\left|\psi\right\rangle = \left|\psi_A\right\rangle + \left|\psi_B\right\rangle .$$

Then,

$$\begin{aligned} e^{i\phi_{AB}}U(T_2T_1)\left(\left|\psi_A\right\rangle + \left|\psi_B\right\rangle\right) &= U(T_2)U(T_1)\left(\left|\psi_A\right\rangle + \left|\psi_B\right\rangle\right) \\ &= U(T_2)U(T_1)\left|\psi_A\right\rangle + U(T_2)U(T_1)\left|\psi_B\right\rangle \\ &= e^{i\phi_A}U(T_2T_1)\left|\psi_A\right\rangle + e^{i\phi_B}U(T_2T_1)\left|\psi_B\right\rangle . \end{aligned} \tag{10.9}$$

Multiplying both sides by $U^{-1}(T_2T_1)$, we obtain

$$\begin{aligned} e^{i\phi_{AB}}\left(\left|\psi_A\right\rangle + \left|\psi_B\right\rangle\right) &= e^{i\phi_A}\left|\psi_A\right\rangle + e^{i\phi_B}\left|\psi_B\right\rangle \\ \rightsquigarrow 0 &= \left(e^{i\phi_{AB}} - e^{i\phi_A}\right)\left|\psi_A\right\rangle + \left(e^{i\phi_{AB}} - e^{i\phi_B}\right)\left|\psi_B\right\rangle \\ \rightsquigarrow e^{i\phi_{AB}} &= e^{i\phi_A} = e^{i\phi_B}. \end{aligned} \tag{10.10}$$

Thus, the phase is the same for all states, and we can write

$$U(T_2)U(T_1) = e^{i\phi(T_2,T_1)}U(T_2T_1). \tag{10.11}$$

For $\phi(T_2, T_1) = 0$, the matrix $U(T)$ furnishes a representation of the symmetry transformation T. For general $\phi(T_2, T_1) \neq 0$, $U(T)$ is a so-called *projective* representation of the symmetry transformations. It can be shown that we can always enlarge the symmetry group so that projective representations turn into representations with zero phase. Let us consider the product of the representations of three symmetry transformations:

$$\begin{aligned} U(T_3)U(T_2)U(T_1) &= U(T_3)\left[U(T_2)U(T_1)\right] \\ &= \left[U(T_3)U(T_2)\right]U(T_1) \\ \rightsquigarrow U(T_3)e^{i\phi(T_2,T_1)}U(T_2T_1) &= e^{i\phi(T_3,T_2)}U(T_3T_2)U(T_1) \\ \rightsquigarrow e^{i\phi(T_2,T_1)+i\phi(T_3,T_2T_1)}U(T_3T_2T_1) &= e^{i\phi(T_3,T_2)+i\phi(T_3T_2,T_1)}U(T_3T_2T_1) \\ \rightsquigarrow \phi(T_2,T_1) + \phi(T_3,T_2T_1) &= \phi(T_3,T_2) + \phi(T_3T_2,T_1). \end{aligned} \tag{10.12}$$

This is satisfied if we write the phase ϕ in terms of a function χ as

$$\phi(T_a, T_b) = \chi(T_aT_b) - \chi(T_a) - \chi(T_b). \tag{10.13}$$

Remark. This form is sufficient but not necessary for the consistency condition above. When it holds, the phase is removed simply by re-phasing, $\bar{U}(T) = U(T)e^{i\chi(T)}$ below, and no enlargement of the group is involved. Phases that cannot be removed this way are the genuinely projective cases; they arise either from the topology of the group (spin 1/2 for rotations, cured by passing to the covering group $SU(2)$) or from central charges in the Lie algebra (the mass in the Galilei group), cured by adding central generators (S. Weinberg, *The Quantum Theory of Fields*, Vol. I, Sec. 2.7 and App. 2B).

Defining the re-phased operators

$$\bar{U}(T) = U(T)e^{i\chi(T)}, \tag{10.14}$$

we see that $\bar{U}(T)$ furnishes a representation (with zero phase),

$$\begin{aligned} \bar{U}(T_b)\bar{U}(T_a) &= U(T_b)U(T_a)e^{i\left(\chi(T_a)+\chi(T_b)\right)} \\ &= U(T_bT_a)e^{i\left(\phi(T_b,T_a)+\chi(T_a)+\chi(T_b)\right)} \\ &= \bar{U}(T_bT_a)e^{i\left(\phi(T_b,T_a)+\chi(T_a)+\chi(T_b)-\chi(T_bT_a)\right)} \\ &= \bar{U}(T_bT_a)\,. \end{aligned} \tag{10.15}$$

For the symmetries of interest, we shall assume that they have representations with zero phase, or that they have been enlarged so that this is achieved.

10.2 Lie algebra and generators

We consider a continuous symmetry transformation

$$T(\theta^a)$$

parametrised by a set of continuous parameters

$$\{\theta^a\}, \quad a = 1 \dots N.$$

The symmetry transformations form a Lie group, and the product of two such transformations is also a symmetry transformation belonging to the group:

$$T(\theta_1^a)T(\theta_2^a) = T\left(f^a(\theta_1^a, \theta_2^a)\right). \tag{10.16}$$

We arrange that T corresponds to the unit element of the symmetry group if all values of the parameters are zero, $\theta^a = 0$:

$$T(\theta^a = 0) = 1. \tag{10.17}$$

The product rule of Eq. (10.16) and our definition of the unit transformation in Eq. (10.17) yield

$$\begin{aligned} f^a(0, \theta^a) &= f^a(\theta^a, 0) \\ &= \theta^a . \end{aligned} \tag{10.18}$$

Restricting ourselves to small transformations around the unity transformation, we perform a Taylor expansion around $\theta^a = 0$:

$$\begin{aligned} f^a(\theta_1^a, \theta_2^a) = f^a(0,0) &+ \frac{\partial f^a}{\partial \theta_1^b}\theta_1^b + \frac{\partial f^a}{\partial \theta_2^b}\theta_2^b + \frac{1}{2}\frac{\partial^2 f^a}{\partial \theta_1^b \partial \theta_1^c}\theta_1^b\theta_1^c \\ &+ \frac{1}{2}\frac{\partial^2 f^a}{\partial \theta_2^b \partial \theta_2^c}\theta_2^b\theta_2^c + \frac{\partial^2 f^a}{\partial \theta_1^b \partial \theta_2^c}\theta_1^b\theta_2^c + \dots, \end{aligned} \tag{10.19}$$

with $b, c = 1, \dots, N$ and the Einstein summation convention. Imposing the requirement of Eq. (10.18) leads to the form

$$f^a(\theta_1^a, \theta_2^a) = \theta_1^a + \theta_2^a + f_{bc}^a \theta_1^b \theta_2^c + \dots, \tag{10.20}$$

with f_{bc}^a the second-order coefficients of the Taylor expansion of f^a; here b, c are dummy indices which are summed over.

For an ordinary representation $U(T)$ of the symmetry group,

$$U\left(T(\theta_1^a)\right) U\left(T(\theta_2^a)\right) = U\left(T\left(f^a(\theta_1^a, \theta_2^a)\right)\right). \tag{10.21}$$

For small parameters θ^a, we can expand the representation of the transformation as follows:

$$U\left(T(\theta^a)\right) = 1 + i\theta^a t_a + \frac{1}{2}\theta^b\theta^c t_{bc} + \dots, \tag{10.22}$$

where $1, t_a, t_{bc}$ are matrices of the same dimensionality as $U\left(T(\theta^a)\right)$ and the matrix t_{bc} is symmetric in b and c:

$$t_{bc} = t_{cb}\,. \tag{10.23}$$

Expanding the terms in Eq. (10.21) using Eq. (10.22), we have

$$\left(1 + i\theta_1^a t_a + \frac{1}{2}\theta_1^b\theta_1^c t_{bc} + \ldots\right)\left(1 + i\theta_2^a t_a + \frac{1}{2}\theta_2^b\theta_2^c t_{bc} + \ldots\right)$$
$$= 1 + i\left(\theta_1^a + \theta_2^a + f_{bc}^a\theta_1^b\theta_2^c + \ldots\right)t_a + \frac{1}{2}\left(\theta_1^b + \theta_2^b + \ldots\right)\left(\theta_1^c + \theta_2^c + \ldots\right)t_{bc}\,,$$

where we have neglected terms of third and higher order in the θ^a. Matching the powers of θ_1^a, θ_2^a, we find that

$$t_{bc} = -t_b t_c - i f_{bc}^a t_a. \tag{10.24}$$

From the symmetry of t_{bc}, we obtain

$$0 = t_{bc} - t_{cb} = -(t_b t_c - t_c t_b) - i(f_{bc}^a - f_{cb}^a)t_a. \tag{10.25}$$

This leads to the commutation relation for the matrices t_a,

$$[t_b, t_c] = iC_{bc}^a t_a, \tag{10.26}$$

where the constants

$$C_{bc}^a = f_{cb}^a - f_{bc}^a \tag{10.27}$$

are antisymmetric under $b \leftrightarrow c$ and are called the *structure constants* of the Lie group. Observe that the structure constants do not depend on the specific representation $U(T(\theta^a))$. Instead, the matrices t_a are specific to the representation and have its dimensionality. These matrices are known as the *generators* of the representation. Knowing the generators, we can construct the representation of an arbitrary symmetry transformation by using the product property of Eq. (10.21) to assemble large symmetry transformations from many small ones, for which the expansion of Eq. (10.22) is valid. The common commutation relation of Eq. (10.26), satisfied by the generators of all representations of the symmetry group, is known as a *Lie algebra*.

Requiring that the representation of a symmetry transformation be unitary leads to the conclusion that the generators of the transformation are Hermitian operators:

$$UU^\dagger = 1 \rightsquigarrow t_a^\dagger = t_a. \tag{10.28}$$

Therefore, generators are good candidates for physical observables.

In the special case of "Abelian" symmetry groups,

$$f^a(\theta_1^a, \theta_2^a) = \theta_1^a + \theta_2^a\,,$$

the generators commute:

$$[t_b, t_c] = 0. \tag{10.29}$$

A representation for general values of the parameters of the symmetry transformation can then be found as

$$\begin{aligned} U\left(T(\theta^a)\right) &= \lim_{N\to\infty} U\left(T\left(\frac{\theta^a}{N}\right)\right)^N \\ &= \lim_{N\to\infty}\left(1 + i\frac{\theta^a}{N}t_a\right)^N \\ &= \exp\left(i\theta^a t_a\right). \end{aligned} \tag{10.30}$$

10.3 Symmetry and degeneracy

A symmetry transformation T should change a system in the same way irrespective of the time at which we perform the transformation. Therefore, the following two states must be the same:

$$U(T)\left(e^{-iHt/\hbar}\left|\psi\right\rangle\right) = e^{-iHt/\hbar}\left(U(T)\left|\psi\right\rangle\right). \tag{10.31}$$

The state on the left-hand side corresponds to performing the transformation after the system has evolved for a time t. The state on the right-hand side corresponds to performing the transformation at $t = 0$ and then evolving for a time t. Since the above equation is valid for any state $\left|\psi\right\rangle$, the time-evolution operator and the representation of the symmetry must commute:

$$\left[U(T), e^{-iHt/\hbar}\right] = 0. \tag{10.32}$$

Equivalently, the Hamiltonian and the symmetry representations also commute:

$$[U(T), H] = 0. \tag{10.33}$$

For continuous symmetry transformations, where

$$U(T) = 1 - i\theta^a t_a + \ldots,$$

the generators of the representation commute with the Hamiltonian:

$$[t_a, H] = 0. \tag{10.34}$$

Then, the Hamiltonian and the generators have a common set of eigenstates.

Symmetries are associated with a degeneracy in the energy spectrum of a physical system. Consider an energy eigenstate,

$$H\left|E\right\rangle = E\left|E\right\rangle. \tag{10.35}$$

All states,

$$U(T)\left|E\right\rangle,$$

are also eigenstates of the Hamiltonian with the same energy eigenvalue E. Indeed,

$$H\left(U(T)\left|E\right\rangle\right) = U(T)H\left|E\right\rangle = U(T)E\left|E\right\rangle = E\left(U(T)\left|E\right\rangle\right). \tag{10.36}$$

10.4 Rotations and translations

In this section, we shall study symmetry transformations corresponding to changes of reference frame. They express the requirement that the laws of quantum mechanics should be the same irrespective of where we perform our experiment or how we have set up our coordinate system. We shall restrict ourselves to Euclidean changes of reference frame, ignoring special-relativistic effects. These transformations preserve distances,

$$\mathbf{r} = (r_1, r_2, r_3) \longrightarrow \mathbf{r}' = (r_1', r_2', r_3'),$$

with

$$|\mathbf{r}_A - \mathbf{r}_B| = \left|\mathbf{r}_A' - \mathbf{r}_B'\right|. \tag{10.37}$$

One can prove that such distance-preserving transformations $\mathbf{r} \longrightarrow \mathbf{r}'$ must *necessarily* be linear,

$$r_i' = R_{ij} r_j + a_i\,. \tag{10.38}$$

Notice that the constant (coordinate-independent) vector a_i displaces the centre of the coordinate system and the constant (coordinate-independent) matrix R_{ij} rotates the coordinates. Substituting into the condition $\delta_{k\ell} = \delta_{ij}\,\partial r'_i/\partial r_k\;\partial r'_j/\partial r_\ell$, which expresses the invariance of distances between infinitesimally close points, we find

$$\delta_{k\ell} = \delta_{ij} R_{ik} R_{j\ell}. \tag{10.39}$$

We write this in matrix notation as

$$RR^T = 1, \quad R^T_{ij} = R_{ji}. \tag{10.40}$$

Taking the determinant of the above, we have

$$\det(RR^T) = 1 \quad \rightsquigarrow (\det R)^2 = 1 \quad \rightsquigarrow \det R \neq 0\,, \tag{10.41}$$

which proves that the inverse transformation R^{-1} exists. We have

$$\begin{aligned} RR^T &= 1 \\ \rightsquigarrow R^{-1}RR^T &= R^{-1} \\ \rightsquigarrow R^{-1}_{ij} &= R^T_{ij} = R_{ji}\,. \end{aligned} \tag{10.42}$$

Making two successive coordinate transformations, we have

$$\begin{aligned} r''_i &= R_{ij} r'_j + a_i \\ &= R_{ij}\left(\bar{R}_{j\ell} r_\ell + \bar{a}_j\right) + a_i \\ &= \left(R\bar{R}\right)_{i\ell} r_\ell + [R\bar{a} + a]_i\,. \end{aligned} \tag{10.43}$$

Thus, we have the following multiplication rule for the group of coordinate transformations:

$$T(R, a)T(\bar{R}, \bar{a}) = T(R\bar{R}, R\bar{a} + a)\,. \tag{10.44}$$

Notice that rotations alone and translations alone form subgroups:

$$T(R, 0)T(\bar{R}, 0) = T(R\bar{R}, 0)\,, \tag{10.45}$$

and

$$T(1, a)T(1, \bar{a}) = T(1, \bar{a} + a)\,. \tag{10.46}$$

10.4.1 Generators

Consider an infinitesimal transformation,

$$T(\delta_{ij} + \omega_{ij}, \epsilon_i), \quad \omega_{ij}, \epsilon_i \in \mathbb{R}\,.$$

From Eq. (10.39), keeping only linear terms in ω, we have

$$\begin{aligned} \delta_{k\ell} &= \delta_{ij} R_{ik} R_{j\ell} \\ \rightsquigarrow \omega_{k\ell} &= -\omega_{\ell k}\,. \end{aligned} \tag{10.47a}$$

Let us now look at the transformation of vectors in Hilbert space corresponding to quantum states in a generic representation of the symmetry group:

$$\left|\psi\right\rangle \to \left|\psi'\right\rangle = U(\omega,\epsilon)\left|\psi\right\rangle . \tag{10.48}$$

Expanding the representation U in the small parameters of the transformation, we have

$$U(\omega,\epsilon) = 1 + \frac{i}{2}\omega_{ij}J^{ij} - i\epsilon_\rho P^\rho + \dots . \tag{10.49}$$

For $U(\omega,\epsilon)$ to be unitary, we need

$$UU^\dagger = 1 \rightsquigarrow (J^{k\ell})^\dagger = J^{k\ell},\ (P^\rho)^\dagger = P^\rho. \tag{10.50}$$

Since $\omega_{ij} = -\omega_{ji}$, we can take

$$J^{k\ell} = -J^{\ell k}. \tag{10.51}$$

We have already identified the *generator*

$$\hbar P^\rho \equiv p^\rho \tag{10.52}$$

as the *operator of momentum*. Similarly, we shall identify the *generators of rotations*

$$\hbar J^{k\ell} \tag{10.53}$$

as *operators of angular momentum*.

How do the expectation values of the generators $J^{k\ell}, P^\rho$ themselves transform under a coordinate transformation? Consider the product

$$U(R,a)U(1+\omega,\epsilon)U^{-1}(R,a) = U\left(R(1+\omega)R^{-1}, R\epsilon - R\omega R^{-1}a\right), \tag{10.54}$$

where

$$U^{-1}(R,a) = U\left(R^{-1}, -R^{-1}a\right), \tag{10.55}$$

and the parameters ω,ϵ of the small transformation are not related to the large transformation parameters R,a. Expanding the left- and right-hand sides of the above in ω,ϵ, we have

$$\begin{aligned} &U(R,a)U(1+\omega,\epsilon)U^{-1}(R,a) \\ &\quad = U\left(1 + \frac{i}{2}\omega_{k\ell}J^{k\ell} - i\epsilon_\rho P^\rho\right)U^{-1} &(10.56)\\ &\quad = 1 + \frac{i}{2}\omega_{k\ell}UJ^{k\ell}U^{-1} - i\epsilon_\rho UP^\rho U^{-1}, &(10.57)\end{aligned}$$

and

$$\begin{aligned} &U\left(R(1+\omega)R^{-1}, R\epsilon - R\omega R^{-1}a\right) \\ &\quad = 1 + \frac{i}{2}\left(R\omega R^{-1}\right)_{k\ell}J^{k\ell} - i\left(R\epsilon - R\omega R^{-1}a\right)_\rho P^\rho \\ &\quad = 1 + \frac{i}{2}\omega_{k\ell}R^{-1}_{km}R^{-1}_{\ell n}\left(J^{mn} - a^m P^n + a^n P^m\right) - i\epsilon_\rho R^{-1}_{\rho k}P^k. &(10.58)\end{aligned}$$

Comparing the two, we obtain

$$U(R,a)J^{k\ell}U^{-1}(R,a) = R^{-1}_{km}R^{-1}_{\ell n}\left(J^{mn} - a^m P^n + a^n P^m\right) \tag{10.59}$$

and

$$U(R,a)P^{\rho}U^{-1}(R,a) = R^{-1}_{\rho m}P^m. \tag{10.60}$$

From the above, we can also show that (**exercise**):

$$U^{-1}(R,a)J^{k\ell}U(R,a) = R_{km}R_{\ell n}J^{mn} + a^k R_{\ell n}P^n - a^\ell R_{kn}P^n\,, \tag{10.61}$$

and

$$U^{-1}(R,a)P^{\rho}U(R,a) = R_{\rho m}P^m. \tag{10.62}$$

The expectation values of the generators transform as follows under a coordinate symmetry transformation:

$$\left\langle\psi\right|P^\rho\left|\psi\right\rangle \to \left\langle\psi\right|U^{-1}P^\rho U\left|\psi\right\rangle = R_{\rho m}\left\langle\psi\right|P^m\left|\psi\right\rangle, \tag{10.63}$$

and

$$\begin{aligned}\left\langle\psi\right|J^{k\ell}\left|\psi\right\rangle &\to \left\langle\psi\right|U^{-1}J^{k\ell}U\left|\psi\right\rangle \\ &= R_{km}R_{\ell n}\left\langle\psi\right|J^{mn}\left|\psi\right\rangle + a^k R_{\ell n}\left\langle\psi\right|P^n\left|\psi\right\rangle - a^\ell R_{kn}\left\langle\psi\right|P^n\left|\psi\right\rangle.\end{aligned} \tag{10.64}$$

Therefore, the expectation value of P^ρ transforms as a classical vector, and the expectation value of $J^{k\ell}$ transforms as a classical tensor under rotations and translations.

10.4.2 Lie algebra

We now further assume that the transformation $T(R,a)$ is also a small transformation,

$$T(R_{ij},a_i) = T(\delta_{ij}+\omega'_{ij},\epsilon'_i), \tag{10.65}$$

and expand Eqs. (10.61)–(10.62) in ω'_{ij},ϵ'_i. Matching the coefficients of the expansion, we find the following Lie algebra for coordinate symmetry transformations:

$$i\left[J^{ij},J^{k\ell}\right] = \left[\delta_{i\ell}J^{jk} - \delta_{ik}J^{j\ell} + \delta_{jk}J^{i\ell} - \delta_{j\ell}J^{ik}\right], \tag{10.66a}$$

$$i\left[P^m,J^{rs}\right] = \delta_{mr}P^s - \delta_{ms}P^r\,, \tag{10.66b}$$

$$\left[P^r,P^s\right] = 0\,. \tag{10.66c}$$

We identify the momentum and angular momentum operators as

$$\mathbf{p} \equiv (p_1,p_2,p_3) \equiv \hbar\left(P^1,P^2,P^3\right) \tag{10.67}$$

and

$$\mathbf{J} \equiv (J_1,J_2,J_3) \equiv \hbar\left(J^{23},J^{31},J^{12}\right) \tag{10.68}$$

respectively. In terms of these operators, the Lie algebra takes the form

$$\left[J_i,J_j\right] = i\hbar\epsilon_{ijk}J_k\,, \tag{10.69a}$$

$$\left[J_i,p_j\right] = i\hbar\epsilon_{ijk}p_k\,, \tag{10.69b}$$

$$\left[p_i,p_j\right] = 0\,. \tag{10.69c}$$

11

Representations of angular momentum

In the previous chapter, we identified the generators of rotations with the angular momentum operators. They satisfy the commutation relations,

$$\left[J_i, J_j\right] = i\hbar\epsilon_{ijk}J_k. \tag{11.1}$$

11.1 Eigenstates of angular momentum

We observe that the square of the angular momentum vector,

$$J^2 \equiv J_k J_k = J_1^2 + J_2^2 + J_3^2, \tag{11.2}$$

commutes with the generators J_i:

$$\begin{aligned}\left[J^2, J_i\right] &= \left[J_k J_k, J_i\right] = J_k\left[J_k, J_i\right] + \left[J_k, J_i\right]J_k \\ &= i\hbar\epsilon_{ki\ell}\left(J_k J_\ell + J_\ell J_k\right) = 0.\end{aligned} \tag{11.3}$$

Thus we can find common eigenstates $\left|f, m\right\rangle$[1] of J^2 and one of the generators J_i, say J_3:

$$J^2\left|f, m\right\rangle = \hbar^2 f\left|f, m\right\rangle, \tag{11.4a}$$

$$J_3\left|f, m\right\rangle = \hbar m\left|f, m\right\rangle. \tag{11.4b}$$

We have

$$\begin{aligned}0 \le &\left|J_1\left|f, m\right\rangle\right|^2 + \left|J_2\left|f, m\right\rangle\right|^2 + \left|J_3\left|f, m\right\rangle\right|^2 \\ &= \left\langle f, m\right|J_1^2 + J_2^2 + J_3^2\left|f, m\right\rangle = \left\langle f, m\right|J^2\left|f, m\right\rangle \\ &= \hbar^2 f\left\langle f, m\right|f, m\right\rangle \\ \rightsquigarrow\ & f \ge 0.\end{aligned} \tag{11.5}$$

Also, the eigenvalue number m of the third component J_3 is bounded:

$$\begin{aligned}0 \le \left\langle f, m\right|J_1^2 + J_2^2\left|f, m\right\rangle &= \left\langle f, m\right|J^2 - J_3^2\left|f, m\right\rangle = \hbar^2(f - m^2) \\ \rightsquigarrow\ m^2 \le f, \quad \text{i.e.} \quad &-\sqrt{f} \le m \le \sqrt{f}.\end{aligned} \tag{11.6}$$

[1] We shall relabel these states later.

We can construct linear combinations of the first two components of angular momentum,

$$J_\pm = J_1 \pm iJ_2, \tag{11.7}$$

which are *ladder* operators:

$$[J_3, J_\pm] = \pm\hbar J_\pm. \tag{11.8}$$

Then

$$\begin{aligned} J_3\left(J_\pm \left|f, m\right\rangle\right) &= \left([J_3, J_\pm] + J_\pm J_3\right)\left|f, m\right\rangle \\ &= \left(\pm\hbar J_\pm + J_\pm J_3\right)\left|f, m\right\rangle \\ &= \left(\pm\hbar J_\pm + \hbar m J_\pm\right)\left|f, m\right\rangle \\ &= \hbar(m \pm 1)\left(J_\pm \left|f, m\right\rangle\right). \end{aligned} \tag{11.9}$$

Thus, the state $J_\pm \left|f, m\right\rangle$ is also an eigenstate of J_3 with eigenvalue $\hbar\,(m \pm 1)$. We may write

$$\begin{aligned} &J_\pm \left|f, m\right\rangle = c_\pm \left|f, m \pm 1\right\rangle, \\ &\rightsquigarrow |c_\pm|^2 = \left\langle f, m\right| J_\mp J_\pm \left|f, m\right\rangle \\ &\qquad = \left\langle f, m\right| J_1^2 + J_2^2 \pm i\,[J_1, J_2]\left|f, m\right\rangle \\ &\qquad = \left\langle f, m\right| J^2 - J_3^2 \mp \hbar J_3 \left|f, m\right\rangle, \\ &\rightsquigarrow |c_\pm|^2 = \hbar^2\left(f - m(m \pm 1)\right). \end{aligned} \tag{11.10a}$$

We then normalise as

$$J_\pm \left|f, m\right\rangle = \hbar\sqrt{f - m(m \pm 1)}\left|f, m \pm 1\right\rangle. \tag{11.10b}$$

Let j be the maximum value of m. Then,

$$J_+ \left|f, j\right\rangle = 0, \tag{11.11}$$

$$\rightsquigarrow f - j(j+1) = 0, \tag{11.12}$$

$$\rightsquigarrow f = j(j+1). \tag{11.13}$$

Now, assume that $j' < j$ is the minimum value of m. Then

$$J_- \left|f, j'\right\rangle = 0, \tag{11.14a}$$

$$\rightsquigarrow j(j+1) - j'(j'-1) = 0, \tag{11.14b}$$

$$\rightsquigarrow j' = -j. \tag{11.14c}$$

The ladder operators span the allowed range $-j \le m \le j$ in integer steps. Thus the number $2j$ must be an integer, and therefore

$$j = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \frac{5}{2}, \ldots. \tag{11.15}$$

We shall relabel the eigenstates $\left|f,m\right\rangle$ of J^2, J_3 with j instead of f. The corresponding eigenvalue equations for the angular momentum operators are:

$$J^2\left|j,m\right\rangle = \hbar^2 j(j+1)\left|j,m\right\rangle, \tag{11.16a}$$

$$J_3\left|j,m\right\rangle = \hbar m\left|j,m\right\rangle, \tag{11.16b}$$

where we have found that the eigenvalue numbers j, m are quantised. For the ladder operators we have found

$$J_\pm\left|j,m\right\rangle = \hbar\sqrt{j(j+1)-m(m\pm1)}\left|j,m\pm1\right\rangle. \tag{11.16c}$$

11.2 Spin-$\frac{1}{2}$ representation of angular momentum

The simplest representation of the angular momentum operators J_i is the spin-$\frac{1}{2}$ representation, where the operators act on the states of a two-state system. For the spin-$\frac{1}{2}$ representation of angular momentum, we have a basis of two angular momentum eigenstates,

$$\left|\frac{1}{2},\pm\frac{1}{2}\right\rangle, \tag{11.17}$$

satisfying

$$J^2\left|\frac{1}{2},\pm\frac{1}{2}\right\rangle = \hbar^2\frac{3}{4}\left|\frac{1}{2},\pm\frac{1}{2}\right\rangle, \tag{11.18a}$$

$$J_3\left|\frac{1}{2},\pm\frac{1}{2}\right\rangle = \pm\hbar\frac{1}{2}\left|\frac{1}{2},\pm\frac{1}{2}\right\rangle, \tag{11.18b}$$

$$J_+\left|\frac{1}{2},\frac{1}{2}\right\rangle = 0, \qquad J_+\left|\frac{1}{2},-\frac{1}{2}\right\rangle = \hbar\left|\frac{1}{2},\frac{1}{2}\right\rangle, \tag{11.19}$$

$$J_-\left|\frac{1}{2},\frac{1}{2}\right\rangle = \hbar\left|\frac{1}{2},-\frac{1}{2}\right\rangle, \qquad J_-\left|\frac{1}{2},-\frac{1}{2}\right\rangle = 0. \tag{11.20}$$

In the representation where

$$\left|\frac{1}{2},\frac{1}{2}\right\rangle \doteq \begin{pmatrix}1\\0\end{pmatrix}, \qquad \left|\frac{1}{2},-\frac{1}{2}\right\rangle \doteq \begin{pmatrix}0\\1\end{pmatrix} \tag{11.21}$$

we find

$$\begin{aligned} J_3 &\doteq \begin{pmatrix} \left\langle\frac{1}{2},\frac{1}{2}\right|J_3\left|\frac{1}{2},\frac{1}{2}\right\rangle & \left\langle\frac{1}{2},\frac{1}{2}\right|J_3\left|\frac{1}{2},-\frac{1}{2}\right\rangle \\ \left\langle\frac{1}{2},-\frac{1}{2}\right|J_3\left|\frac{1}{2},\frac{1}{2}\right\rangle & \left\langle\frac{1}{2},-\frac{1}{2}\right|J_3\left|\frac{1}{2},-\frac{1}{2}\right\rangle \end{pmatrix} \\ &= \frac{\hbar}{2}\begin{pmatrix}1 & 0\\0 & -1\end{pmatrix}. \end{aligned} \tag{11.22a}$$

Similarly,

$$J_- \doteq \hbar\begin{pmatrix}0 & 0\\1 & 0\end{pmatrix}, \tag{11.22b}$$

and

$$J_+ \doteq \hbar \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}. \tag{11.22c}$$

The ladder operators are $J_\pm = J_1 \pm iJ_2$. From the above we obtain

$$J_1 \doteq \frac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad J_2 \doteq \frac{\hbar}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}. \tag{11.23}$$

The matrices

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{11.24}$$

are known as the *Pauli matrices*.

We can verify with an explicit calculation (**exercise**) that

$$\left[\sigma_i, \sigma_j\right] = 2i\epsilon_{ijk}\sigma_k. \tag{11.25}$$

Therefore, the representation of angular momentum as

$$J_i \doteq \hbar \frac{\sigma_i}{2}$$

satisfies the anticipated Lie algebra of rotations.

Parenthetically, we note the following property of the Pauli matrices:

$$\{\sigma_i, \sigma_j\} \equiv \sigma_i\sigma_j + \sigma_j\sigma_i = 2\delta_{ij}\mathbf{1}_{2\times 2}. \tag{11.26}$$

Combining Eq. (11.25) and Eq. (11.26), we obtain the product of two Pauli matrices,

$$\sigma_i\sigma_j = \delta_{ij}\mathbf{1}_{2\times 2} + i\,\epsilon_{ijk}\sigma_k. \tag{11.27}$$

We shall now use the property of the angular momentum operators as generators of rotations. For a rotation by a small angle θ around the z-axis, the representation of the transformation reads

$$\begin{aligned} U(\theta) &= \mathbf{1}_{2\times 2} + \frac{i}{2}\left(\omega_{12}J_{12} + \omega_{21}J_{21}\right) \\ &= \mathbf{1}_{2\times 2} + \theta i\frac{J_3}{\hbar} = \mathbf{1}_{2\times 2} + i\frac{\theta\sigma_3}{2}, \end{aligned} \tag{11.28}$$

where we have used that

$$\omega_{12} = -\omega_{21} = \theta, \quad J_{12} = -J_{21} = \frac{J_3}{\hbar} = \frac{\sigma_3}{2}.$$

For a large angle θ, we have

$$U(\theta) = \lim_{N\to\infty}\left[U\left(\frac{\theta}{N}\right)\right]^N = \exp\left(\frac{i\sigma_3\theta}{2}\right). \tag{11.29}$$

Remark. Sign convention. With $\omega_{12} = \theta$ the coordinates transform as $x_1 \to x_1 + \theta x_2$, $x_2 \to x_2 - \theta x_1$, i.e. a rotation by $-\theta$ in the right-handed sense; hence $U = \exp(+i\theta J_3/\hbar)$

here, whereas Sakurai (Eq. 3.1.15) writes $\exp(-iJ_3\phi/\hbar)$ for a counter-clockwise rotation by ϕ. All physics (e.g. $U(2\pi) = -1$) is unaffected, but the signs of phases differ from those in most textbooks.

The exponential $e^{a\sigma_k}$, where σ_k ($k = 1, 2$ or 3) is a Pauli matrix with the property

$$\sigma_k^2 = \mathbf{1}_{2\times 2}\,,$$

is given by

$$\begin{aligned}
\exp(a\sigma_k) &= \sum_{n=0}^{\infty} \frac{(a\sigma_k)^n}{n!} \\
&= \sum_{n=0}^{\infty} \frac{(a\sigma_k)^{2n}}{(2n)!} + \sum_{n=0}^{\infty} \frac{(a\sigma_k)^{2n+1}}{(2n+1)!} \\
&= 1_{2\times 2} \sum_{n=0}^{\infty} \frac{a^{2n}}{(2n)!} + \sigma_k \sum_{n=0}^{\infty} \frac{a^{2n+1}}{(2n+1)!} \\
&= \mathbf{1}_{2\times 2} \frac{e^a + e^{-a}}{2} + \sigma_k \frac{e^a - e^{-a}}{2}.
\end{aligned} \tag{11.30}$$

Thus,

$$\begin{aligned}
U(\theta) &= \exp\left(\frac{i\sigma_3\theta}{2}\right) = \mathbf{1}_{2\times 2} \cos\left(\frac{\theta}{2}\right) + i\sigma_3 \sin\left(\frac{\theta}{2}\right) \\
&= \begin{pmatrix} \cos\left(\frac{\theta}{2}\right) + i\sin\left(\frac{\theta}{2}\right) & 0 \\ 0 & \cos\left(\frac{\theta}{2}\right) - i\sin\left(\frac{\theta}{2}\right) \end{pmatrix} \\
&= \begin{pmatrix} \exp\left(i\frac{\theta}{2}\right) & 0 \\ 0 & \exp\left(-i\frac{\theta}{2}\right) \end{pmatrix}.
\end{aligned} \tag{11.31}$$

A general state of a two-state system,

$$\left|\psi\right\rangle \doteq \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}, \tag{11.32}$$

under a rotation by θ around the z-axis transforms as

$$\left|\psi\right\rangle \to \left|\psi'\right\rangle = U(\theta)\left|\psi\right\rangle \doteq \begin{pmatrix} e^{i\theta/2}\psi_1 \\ e^{-i\theta/2}\psi_2 \end{pmatrix}. \tag{11.33}$$

Notice that a rotation by $\theta = 2\pi$ does not bring the system back to the original state, against common intuition:

$$U(2\pi)\left|\psi\right\rangle = -\left|\psi\right\rangle. \tag{11.34}$$

For spin-$\frac{1}{2}$ systems, a rotation by 4π is required in order to return to the original state.

We can generalise our discussion to a rotation by an angle θ around any axis $\hat{n}$. We can prove (**exercise**) that the corresponding spin-$\frac{1}{2}$ representation is:

$$U(\hat{n}, \theta) = \exp\left(i\frac{\boldsymbol{\sigma}\cdot\hat{n}}{2}\theta\right), \tag{11.35}$$

which can also be written as

$$U(\hat{n},\theta) = \mathbf{1}_{2\times 2}\cos\frac{\theta}{2} + i\boldsymbol{\sigma}\cdot\hat{n}\sin\frac{\theta}{2}, \tag{11.36}$$

where

$$\boldsymbol{\sigma} = (\sigma_1,\sigma_2,\sigma_3)\,. \tag{11.37}$$

To prove the above, the following relations are useful:

$$\boldsymbol{\sigma}\cdot\mathbf{a} = \begin{pmatrix} a_3 & a_1 - ia_2 \\ a_1 + ia_2 & -a_3 \end{pmatrix}, \tag{11.38a}$$

$$(\boldsymbol{\sigma}\cdot\mathbf{a})^2 = \mathbf{a}^2, \tag{11.38b}$$

and

$$(\boldsymbol{\sigma}\cdot\mathbf{a})(\boldsymbol{\sigma}\cdot\mathbf{b}) = \mathbf{a}\cdot\mathbf{b} + i(\mathbf{a}\times\mathbf{b})\cdot\boldsymbol{\sigma}. \tag{11.38c}$$

11.3 Orbital angular momentum

In this section we study the space representation of the eigenstates of the angular momentum operators $L_3, L^2 = \sum_i L_i^2$, where we have defined

$$L_i = \epsilon_{ijk} x_j p_k, \tag{11.39}$$

with x_i, p_j being the position and momentum operators. It is straightforward to prove that the operators $\frac{L_i}{\hbar}$ satisfy the Lie algebra of the generators of rotations,

$$\left[\frac{L_i}{\hbar}, \frac{L_j}{\hbar}\right] = i\epsilon_{ijk}\frac{L_k}{\hbar}. \tag{11.40}$$

According to the general theory of the generators of rotations, we anticipate that we can find common eigenstates $|\ell, m\rangle$ for $L^2 = \sum_i L_i^2$ and one of the generators L_i with

$$L^2|\ell,m\rangle = \hbar^2\ell(\ell+1)|\ell,m\rangle$$

and

$$L_3|\ell,m\rangle = \hbar m|\ell,m\rangle\,.$$

11.3.1 Spherical coordinates

It will be useful to analyse angular momentum using spherical coordinates, where

$$x_1 = r\sin\theta\sin\phi\,, \tag{11.41a}$$

$$x_2 = r\sin\theta\cos\phi\,, \tag{11.41b}$$

$$x_3 = r\cos\theta\,. \tag{11.41c}$$

Remark. We use an unusual convention: here ϕ is measured from the x_2-axis ($x_1 = r\sin\theta\sin\phi$, $x_2 = r\sin\theta\cos\phi$), so the frame (r,θ,ϕ) is left-handed and $L_3 = +i\hbar\,\partial_\phi$, $Y_\ell^m \propto e^{-im\phi}$. With the standard azimuth $\phi_s = \pi/2 - \phi$ (Sakurai, Jackson) one has

$L_3 = -i\hbar\partial_{\phi_s}$, and the final result, Eq. (11.90), coincides exactly with the standard Condon–Shortley $Y_\ell^m(\theta, \phi_s)$.

For the differentials, we have

$$\partial_r = \sum_i \frac{\partial x_i}{\partial r}\partial_i \,, \tag{11.42a}$$

$$\partial_\theta = \sum_i \frac{\partial x_i}{\partial \theta}\partial_i \,, \tag{11.42b}$$

$$\partial_\phi = \sum_i \frac{\partial x_i}{\partial \phi}\partial_i \,, \tag{11.42c}$$

which in matrix notation gives

$$\begin{pmatrix}\partial_r\\ \partial_\theta\\ \partial_\phi\end{pmatrix} = \begin{pmatrix}\sin\theta\sin\phi & \sin\theta\cos\phi & \cos\theta\\ r\cos\theta\sin\phi & r\cos\theta\cos\phi & -r\sin\theta\\ r\sin\theta\cos\phi & -r\sin\theta\sin\phi & 0\end{pmatrix}\begin{pmatrix}\partial_1\\ \partial_2\\ \partial_3\end{pmatrix}. \tag{11.43}$$

Inverting, we obtain

$$\begin{pmatrix}\partial_1\\ \partial_2\\ \partial_3\end{pmatrix} = \begin{pmatrix}\sin\theta\sin\phi & \frac{\cos\theta\sin\phi}{r} & \frac{\cos\phi}{r\sin\theta}\\ \sin\theta\cos\phi & \frac{\cos\phi\cos\theta}{r} & -\frac{\sin\phi}{r\sin\theta}\\ \cos\theta & -\frac{\sin\theta}{r} & 0\end{pmatrix}\begin{pmatrix}\partial_r\\ \partial_\theta\\ \partial_\phi\end{pmatrix}. \tag{11.44}$$

We then calculate (with $\cot\theta = \frac{\cos\theta}{\sin\theta}$)

$$x_2\partial_3 - x_3\partial_2 = \cot\theta\sin\phi\,\partial_\phi - \cos\phi\,\partial_\theta \,, \tag{11.45a}$$

$$x_3\partial_1 - x_1\partial_3 = \cot\theta\cos\phi\,\partial_\phi + \sin\phi\,\partial_\theta \,, \tag{11.45b}$$

$$x_1\partial_2 - x_2\partial_1 = -\partial_\phi \,. \tag{11.45c}$$

The Laplacian operator is

$$\begin{aligned}\nabla^2 &= \sum_i \partial_i^2\\ &= \frac{1}{r^2}\left[\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right) + \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right].\end{aligned} \tag{11.46}$$

11.3.2 Orbital angular momentum operators and the Laplacian

Consider the action of an angular momentum operator on a general state $|a\rangle$,

$$L_i\,|a\rangle = \epsilon_{ijk}\hat{x}_j\hat{p}_k\,|a\rangle \,. \tag{11.47}$$

Acting with a position bra $\langle\mathbf{r}| = \langle x_1, x_2, x_3|$, we have

$$\begin{aligned}\langle\mathbf{r}|\,L_i\,|a\rangle &= \epsilon_{ijk}\,\langle\mathbf{r}|\,\hat{x}_j\hat{p}_k\,|a\rangle\\ &= -i\hbar\epsilon_{ijk}x_j\partial_k\,\langle\mathbf{r}|a\rangle \,.\end{aligned} \tag{11.48}$$

For example,

$$\begin{aligned}\langle \mathbf{r}| L_3 |a\rangle &= \langle \mathbf{r}| \hat{x}_1 \hat{p}_2 - \hat{x}_2 \hat{p}_1 |a\rangle \\ &= -i\hbar(x_1\partial_2 - x_2\partial_1)\langle \mathbf{r}|a\rangle .\end{aligned} \tag{11.49}$$

Using the spherical coordinates introduced above, we obtain

$$\langle \mathbf{r}| L_3 |a\rangle = i\hbar \left(+\partial_\phi\right) \langle \mathbf{r}|a\rangle . \tag{11.50}$$

Similarly,

$$\langle \mathbf{r}| L_1 |a\rangle = -i\hbar \left(\cot\theta \sin\phi\, \partial_\phi - \cos\phi\, \partial_\theta\right) \langle \mathbf{r}|a\rangle , \tag{11.51}$$

and

$$\langle \mathbf{r}| L_2 |a\rangle = -i\hbar \left(\cot\theta \cos\phi\, \partial_\phi + \sin\phi\, \partial_\theta\right) \langle \mathbf{r}|a\rangle . \tag{11.52}$$

For the *ladder operators*,

$$L_\pm = L_1 \pm iL_2 , \tag{11.53}$$

we obtain

$$\langle \mathbf{r}| L_\pm |a\rangle = i\hbar \left(e^{\mp i\phi} \left[\partial_\theta \mp i \cot\theta\, \partial_\phi\right]\right) \langle \mathbf{r}|a\rangle . \tag{11.54}$$

We can write

$$L^2 = L_3^2 + \frac{1}{2}\left(L_+L_- + L_-L_+\right). \tag{11.55}$$

After a little algebra, we arrive at

$$\langle \mathbf{r}| L^2 |a\rangle = -\hbar^2 \left[\frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2} + \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right)\right] \langle \mathbf{r}|a\rangle . \tag{11.56}$$

We notice that the representation of the operator L^2 in position space yields the angular part of the Laplacian operator, which, in turn, represents the square of the momentum. Explicitly,

$$\begin{aligned}\langle \mathbf{r}| p^2 |a\rangle &= -\hbar^2\nabla^2 \langle \mathbf{r}|a\rangle \\ &= -\hbar^2 \left[\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right)\right] \langle \mathbf{r}|a\rangle + \frac{1}{r^2}\langle \mathbf{r}| L^2 |a\rangle .\end{aligned} \tag{11.57}$$

11.3.3 Spherical harmonics

The eigenstates $|\ell, m\rangle$ of L^2, L_3 satisfy

$$i\hbar\partial_\phi \langle \mathbf{r}|\ell, m\rangle = \hbar m \langle \mathbf{r}|\ell, m\rangle , \tag{11.58}$$

and admit the general solution

$$\langle \mathbf{r}|\ell, m\rangle = \eta(r, \theta)e^{-im\phi}. \tag{11.59}$$

For the maximum value ℓ of m, we have

$$\langle \mathbf{r}| L_+ |\ell, \ell\rangle = 0, \tag{11.60}$$

which yields the differential equation

$$e^{-i\phi}\left[\partial_\theta - i\cot\theta\,\partial_\phi\right]\eta(r,\theta)e^{-i\ell\phi} = 0 \tag{11.61}$$

$$\rightsquigarrow \left[\partial_\theta - \ell\cot\theta\right]\eta(r,\theta) = 0\,. \tag{11.62}$$

This admits the solution

$$\eta(r,\theta) = \psi(r)(\sin\theta)^\ell. \tag{11.63}$$

We then have

$$\langle \mathbf{r}|\ell,\ell\rangle = \psi(r)Y_\ell^\ell(\theta,\phi), \tag{11.64}$$

where

$$Y_\ell^\ell(\theta,\phi) = c_{\ell\ell}e^{-i\ell\phi}(\sin\theta)^\ell \tag{11.65}$$

contains the angular dependence on the variables (θ,ϕ) of $\langle\mathbf{r}|\ell,\ell\rangle$. This function is a so-called *spherical harmonic*. We fix the normalisation constant as follows:

$$\begin{aligned} 1 &= \langle\ell,\ell|\ell,\ell\rangle \\ &= \int \mathrm{d}^3\mathbf{r}\,\langle\ell,\ell|\mathbf{r}\rangle\langle\mathbf{r}|\ell,\ell\rangle \\ &= \int_0^\infty \mathrm{d}r\,r^2\left|\psi(r)\right|^2\int\mathrm{d}\Omega\,\left|Y_\ell^\ell\right|^2. \end{aligned} \tag{11.66}$$

We require that the angular integral and the radial integral are separately normalised to unity:

$$\int_0^\infty \mathrm{d}r\,r^2\left|\psi(r)\right|^2 = 1 \tag{11.67}$$

and

$$\int\mathrm{d}\Omega\,\left|Y_\ell^\ell\right|^2 = 1\,. \tag{11.68}$$

The angular normalisation condition yields

$$|c_{\ell\ell}|^2\int_0^{2\pi}\mathrm{d}\phi\int_0^\pi\mathrm{d}\theta\,\sin\theta\left(\sin^2\theta\right)^\ell = 1$$

$$\rightsquigarrow 2\pi B\left(\ell+1,\frac{1}{2}\right) = |c_{\ell\ell}|^{-2}, \tag{11.69}$$

where $B(x,y) = \Gamma(x)\Gamma(y)/\Gamma(x+y)$ is Euler's Beta function, so that

$$\sqrt{4\pi\frac{4^\ell(\ell!)^2}{(2\ell+1)!}} = |c_{\ell\ell}|^{-1}\,. \tag{11.70}$$

We can obtain the spherical harmonics for $|\ell, m\rangle$ with lower values $m < \ell$ by means of the ladder operators. Acting with the L_- operator on a general $|\ell, m\rangle$ state, we obtain

$$\langle \mathbf{r}|\ell, m-1\rangle = \frac{\langle \mathbf{r}| L_- |\ell, m\rangle}{\hbar\sqrt{(\ell+m)(\ell+1-m)}}, \tag{11.71}$$

or equivalently,

$$\langle \mathbf{r}|\ell, m-1\rangle = i\hbar e^{i\phi}\left(\partial_\theta + i\cot\theta\,\partial_\phi\right)\frac{\langle \mathbf{r}|\ell, m\rangle}{\sqrt{(\ell+m)(\ell+1-m)}\hbar}. \tag{11.72}$$

It is obvious that the application of this operator does not mix the radial and angular parts in producing $\langle \mathbf{r}|\ell, m\rangle$ for $m \neq \ell$. So, we can decompose

$$\langle \mathbf{r}|\ell, m\rangle = \psi_{\ell m}(r) Y_\ell^m(\theta, \phi). \tag{11.73}$$

The spherical harmonics Y_ℓ^m depend purely on the direction of the vector $\mathbf{r}$. We can think of them as

$$\langle \hat{\mathbf{r}}|\ell, m\rangle = Y_\ell^m(\theta, \phi), \quad \hat{\mathbf{r}} = \frac{\mathbf{r}}{|\mathbf{r}|}. \tag{11.74}$$

For the spherical harmonics $Y_\ell^m(\theta, \phi)$, we obtain

$$Y_\ell^{m-1}(\theta, \phi) = i\left(e^{i\phi}\left[\partial_\theta + i\cot\theta\,\partial_\phi\right]\right)\frac{Y_\ell^m(\theta, \phi)}{\sqrt{(\ell+m)(\ell+1-m)}}. \tag{11.75}$$

According to our general theory for the eigenvalues of the generators of rotations, we anticipate that the maximum value ℓ of m is an integer or a half-integer and that $m = -\ell, -\ell+1, \ldots, \ell-1, \ell$. Here we shall find a little surprise. While the operators $\mathbf{L} = \mathbf{r} \times \mathbf{p}$ furnish a representation of rotations in Hilbert space, their eigenvalues need not take all the values permitted by the Lie algebra. In fact, for orbital angular momentum only integer values of ℓ, m are allowed. Assume that $\ell = \frac{1}{2}$. Then, we find that

$$Y_{\frac{1}{2}}^{\frac{1}{2}}(\theta, \phi) = c_{\frac{1}{2}\frac{1}{2}} e^{-i\phi/2}\sqrt{\sin\theta}. \tag{11.76}$$

Applying the lowering operator, we obtain

$$\begin{aligned} Y_{\frac{1}{2}}^{-\frac{1}{2}}(\theta, \phi) &= i\left(e^{i\phi}\left[\partial_\theta + i\cot\theta\,\partial_\phi\right]\right) c_{\frac{1}{2}\frac{1}{2}} e^{-i\phi/2}\sqrt{\sin\theta} \\ &= i c_{\frac{1}{2}\frac{1}{2}} e^{i\phi/2}\frac{\cos\theta}{\sqrt{\sin\theta}}. \end{aligned} \tag{11.77}$$

For a consistent spherical harmonic, applying the lowering operator once more should yield zero; it does not. Notice also that the derived spherical harmonic gives an infinite value for $\theta = 0$.

Spherical harmonics are therefore only defined for integer values of the angular momentum index, $\ell = 0, 1, 2, 3, \ldots$. Starting from a generic $m = \ell$ value, we can obtain the spherical

harmonic Y_ℓ^m by applying Eq. (11.75) $\ell - m$ times. Let us denote the differential operator of Eq. (11.75) as

$$D_- \equiv i\left(e^{i\phi}\left[\partial_\theta + i\cot\theta\,\partial_\phi\right]\right) \tag{11.78}$$

and define

$$\omega_\ell \equiv e^{-i\ell\phi}\sin^\ell\theta. \tag{11.79}$$

We find that

$$\omega_\ell = e^{-i\phi}\sin\theta\,\omega_{\ell-1}\,, \tag{11.80}$$

$$D_-\omega_\ell = (2i\ell\cos\theta)\omega_{\ell-1}\,, \tag{11.81}$$

$$D_-\left(f\left(\theta\right)\right) = -ie^{i\phi}\sin\theta\frac{\mathrm{d}}{\mathrm{d}\left(\cos\theta\right)}f(\theta)\,, \tag{11.82}$$

$$\frac{\mathrm{d}}{\mathrm{d}\left(\cos\theta\right)}\left[(\sin\theta)^{2\ell}\,f(\theta)\right] = -\left(\sin\theta\right)^{2(\ell-1)}\left(2\ell\cos\theta f(\theta) - \sin^2\theta\frac{\mathrm{d}}{\mathrm{d}\left(\cos\theta\right)}f(\theta)\right), \tag{11.83}$$

and that

$$\begin{aligned} D_-\left(f(\theta)\omega_\ell\right) &= f\left(\theta\right)D_-\omega_\ell + \omega_\ell D_-\left(f(\theta)\right) \\ &= \omega_{\ell-1}\frac{-i}{\left(\sin\theta\right)^{2(\ell-1)}}\frac{\mathrm{d}}{\mathrm{d}\cos\theta}\left[(\sin\theta)^{2\ell}f(\theta)\right], \end{aligned} \tag{11.84}$$

where $f(\theta)$ is an arbitrary function of the polar angle θ. It is now straightforward to compute the repeated application of the differential operator on ω_ℓ. For example,

$$\begin{aligned} D_-(D_-\omega_\ell) &= 2i\ell D_-(\underbrace{\cos\theta}_{f(\theta)}\,\omega_{\ell-1}) \\ &= \omega_{\ell-2}\frac{(-i)^2}{\left(\sin\theta\right)^{2(\ell-2)}}\frac{\mathrm{d}^2}{\mathrm{d}\left(\cos\theta\right)^2}\left[(\sin\theta)^{2\ell}\right], \end{aligned} \tag{11.85}$$

and in general,

$$D_-{}^n\omega_\ell = \omega_{\ell-n}\frac{(-i)^n}{\left(\sin\theta\right)^{2(\ell-n)}}\frac{\mathrm{d}^n}{\mathrm{d}\left(\cos\theta\right)^n}\left[(\sin\theta)^{2\ell}\right]. \tag{11.86}$$

We are now ready to determine the spherical harmonics Y_ℓ^m for positive values of m, using the recurrence identity of Eq. (11.75) on

$$Y_\ell^\ell\left(\theta,\phi\right) = c_{\ell\ell}\omega_\ell \tag{11.87}$$

$\ell - m$ times, we find

$$Y_\ell^m(\theta,\phi) = (-i)^\ell c_{\ell\ell}\sqrt{\frac{(\ell+m)!}{(2\ell)!\,(\ell-m)!}}\frac{e^{-im(\phi-\frac{\pi}{2})}}{\left(\sin\theta\right)^m}\frac{\mathrm{d}^{\ell-m}}{\mathrm{d}\left(\cos\theta\right)^{\ell-m}}(\sin\theta)^{2\ell}. \tag{11.88}$$

The normalisation $c_{\ell\ell}$ can be determined up to a phase. We conventionally choose this phase to be

$$c_{\ell\ell} = |c_{\ell\ell}|\,(-i)^\ell. \tag{11.89}$$

With this normalisation, we obtain

$$Y_\ell^m(\theta,\phi) = \sqrt{\frac{(2\ell+1)}{4\pi}\frac{(\ell+m)!}{(\ell-m)!}}\frac{e^{-im(\phi-\frac{\pi}{2})}}{(\sin\theta)^m}\frac{(-1)^\ell}{2^\ell\ell!}\frac{\mathrm{d}^{\ell-m}}{\mathrm{d}\,(\cos\theta)^{\ell-m}}(\sin\theta)^{2\ell}. \tag{11.90}$$

For $m = 0$, the derivatives above are related to the Legendre polynomials, defined as

$$P_\ell(x) = \frac{(-1)^\ell}{2^\ell\ell!}\frac{\mathrm{d}^\ell}{\mathrm{d}x^\ell}(1-x^2)^\ell. \tag{11.91}$$

As one can easily check, they satisfy the differential equation

$$\frac{\mathrm{d}}{\mathrm{d}x}\left((1-x^2)\frac{\mathrm{d}}{\mathrm{d}x}P_\ell(x)\right) + \ell(\ell+1)P_\ell(x) = 0. \tag{11.92}$$

For $x = \cos\theta$, we have

$$\frac{\mathrm{d}}{\mathrm{d}\cos\theta}\sin^2\theta\frac{\mathrm{d}}{\mathrm{d}\cos\theta}P_\ell(\cos\theta) + \ell(\ell+1)P_\ell(\cos\theta) = 0, \tag{11.93}$$

or, equivalently,

$$\frac{1}{\sin\theta}\frac{\mathrm{d}}{\mathrm{d}\theta}\sin\theta\frac{\mathrm{d}}{\mathrm{d}\theta}P_\ell(\cos\theta) + \ell(\ell+1)P_\ell(\cos\theta) = 0. \tag{11.94}$$

For $m = 0$, the spherical harmonics become

$$Y_\ell^0(\theta,\phi) = \sqrt{\frac{(2\ell+1)}{4\pi}}P_\ell(\cos\theta). \tag{11.95}$$

To obtain the spherical harmonics for negative values of m, we have to repeat a completely analogous study, starting from $Y_\ell^{-\ell}$ as derived from

$$L_-\,|\ell,-\ell\rangle = 0,$$

and using that

$$L_+\,|\ell,m\rangle = \hbar\sqrt{(\ell-m)(\ell+m+1)}\,|\ell,m+1\rangle\,.$$

We find (**exercise**)

$$Y_\ell^{-m}(\theta,\phi) = (-1)^m\left[Y_\ell^m(\theta,\phi)\right]^*. \tag{11.96}$$

11.3.4 Space representations of rotations

An arbitrary rotation of coordinates $x_i \to x_i' = R_{ij}x_j$ is represented by a unitary operator

$$\left|\psi\right\rangle \to U(R)\left|\psi\right\rangle, \tag{11.97}$$

where the operator $U(R)$ is a function of the generators J_i with $i = 1,2,3$. Since the square of the angular momentum commutes with all generators J_i, we should also have that

$$\left[J^2, U(R)\right] = 0. \tag{11.98}$$

Thus, the states $|\ell,m\rangle$ and $U(R)\,|\ell,m\rangle$ have the same eigenvalue of the J^2 operator. The state $U(R)\,|\ell,m\rangle$ can only be a superposition of $|\ell,m'\rangle$ states with the same value of ℓ: in a general decomposition,

$$U(R)\,|\ell,m\rangle = \sum_{\ell'm'} c_{\ell'm'}\,|\ell',m'\rangle\,, \tag{11.99}$$

only the terms with $\ell' = \ell$ are present. We can write

$$U(R)\,|\ell, m\rangle = \sum_{m'} c_{m'}\,|\ell, m'\rangle\,. \tag{11.100}$$

Physically, this result is anticipated, since a rotation should not change the total angular momentum.

Consider a position state $|\hat{z}\rangle$, corresponding to the unit vector in the z-direction. A position state $|\hat{r}\rangle$ with

$$\hat{r} = (\sin\theta\sin\phi, \sin\theta\cos\phi, \cos\theta), \tag{11.101}$$

can be produced by a rotation,

$$|\hat{r}\rangle = U(\theta, \phi)\,|\hat{z}\rangle\,. \tag{11.102}$$

We then have

$$\begin{aligned} |\hat{r}\rangle &= \sum_{\ell', m'} U(\theta,\phi)|\ell', m'\rangle\langle\ell', m'|\,|\hat{z}\rangle \\ \rightsquigarrow \langle\ell, m|\hat{r}\rangle &= \sum_{\ell', m'} \langle\ell, m|\,U(\theta,\phi)\,|\ell', m'\rangle\,\langle\ell', m'|\hat{z}\rangle \\ \rightsquigarrow \langle\ell, m|\hat{r}\rangle &= \sum_{m'} \langle\ell, m|\,U(\theta,\phi)\,|\ell, m'\rangle\,\langle\ell, m'|\hat{z}\rangle\,. \end{aligned} \tag{11.103}$$

The matrix element $\langle\hat{z}|\ell, m\rangle$ is zero unless $m = 0$. Indeed,

$$L_3\,|\hat{z}\rangle = (\hat{x}_1\hat{p}_2 - \hat{x}_2\hat{p}_1)\,|\hat{z}\rangle = 0,$$

since the position operators $\hat{x}_{1,2}$ yield a zero eigenvalue for a position ket in the orthogonal z-direction. Also,

$$0 = \langle\hat{z}|\,L_3\,|\ell, m\rangle = \hbar m\,\langle\hat{z}|\ell, m\rangle\,.$$

For $m \neq 0$, we must have that $\langle\hat{z}|\ell, m\rangle = 0$. Therefore,

$$\begin{aligned} \langle\hat{z}|\ell, m\rangle &= \langle\hat{z}|\ell, 0\rangle\,\delta_{0m} = Y_\ell^0(\theta = 0, \phi)\delta_{0m} \\ &= \sqrt{\frac{(2\ell+1)}{4\pi}}\,P_\ell(\cos\theta)|_{\theta=0}\,\delta_{0m} \\ \rightsquigarrow \langle\hat{z}|\ell, m\rangle &= \sqrt{\frac{(2\ell+1)}{4\pi}}\delta_{0m}. \end{aligned}$$

Substituting in Eq. (11.103), we obtain

$$Y_\ell^m(\theta,\phi)^* = \langle\ell, m|\,U(\theta,\phi)\,|\ell, 0\rangle\sqrt{\frac{(2\ell+1)}{4\pi}}. \tag{11.104}$$

11.4 Potentials with spherical symmetry

We have proven that systems with spherical symmetry have a Hamiltonian which commutes with the generators of rotations. The common eigenstates of J^2 and J_3 should also be eigenstates of the Hamiltonian, satisfying

$$H\,|E, \ell, m\rangle = E\,|E, \ell, m\rangle\,, \tag{11.105a}$$

$$L^2\,|E, \ell, m\rangle = \hbar^2\ell(\ell+1)\,|E, \ell, m\rangle\,, \tag{11.105b}$$

$$L_3\,|E, \ell, m\rangle = \hbar m\,|E, \ell, m\rangle\,. \tag{11.105c}$$

The Hamiltonian of a particle inside a spherically symmetric potential,

$$H = \frac{p^2}{2M} + V(r), \tag{11.106}$$

where the potential $V(r)$ is a function only of the radial distance r, is invariant under rotations. Indeed, we can easily compute

$$\left[r_i, L_j\right] = i\hbar\epsilon_{ijk} r_k \,, \tag{11.107a}$$

$$\left[p_i, L_j\right] = i\hbar\epsilon_{ijk} p_k \,, \tag{11.107b}$$

$$\left[p^2, L_i\right] = 0 \,, \tag{11.107c}$$

$$\left[r^2, L_i\right] = 0 \,, \tag{11.107d}$$

from which we conclude that

$$[H, L_i] = 0. \tag{11.108}$$

Recall that

$$\langle \mathbf{r} | E, \ell, m \rangle = \psi_{E,\ell,m}(r) Y_\ell^m(\theta, \phi) \,, \tag{11.109}$$

and, from Eq. (11.57), we also obtain

$$\langle \mathbf{r} | \, p^2 \, | E, \ell, m \rangle = -\hbar^2 \left[\frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) - \frac{\ell(\ell+1)}{r^2} \right] \psi_{E,\ell,m}(r) Y_\ell^m(\theta, \phi). \tag{11.110}$$

From Eq. (11.105a), we obtain a differential equation for the radial part of the wave function,

$$\left[\frac{1}{r^2} \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) - \frac{\ell(\ell+1)}{r^2} - \frac{2M(V(r) - E)}{\hbar^2} \right] \psi_{E,\ell,m}(r) = 0. \tag{11.111}$$

Notice that the "radial" wave function does not depend on the eigenvalue m of L_3,

$$\psi_{E,\ell,m}(r) \equiv \psi_{E,\ell}(r) \,. \tag{11.112}$$

We write

$$\psi_{E,\ell}(r) = \frac{R(r)}{r} \,, \tag{11.113}$$

which yields

$$0 = R''(r) - \left[\frac{\ell(\ell+1)}{r^2} + \frac{2M(V(r) - E)}{\hbar^2} \right] R(r) \,. \tag{11.114}$$

Exercise 11.1. *Solve this equation in the case of a free particle, $V(r) = 0$.*

Exercise 11.2. *Solve this equation for a spherical well $V(r) = V_0 \Theta(r < r_0)$.*

11.5 The hydrogen atom

We are now in a position to discuss a simplified version of the hydrogen atom, where we ignore the spin of the electron. The potential is given by

$$V(r) = -\frac{Ze^2}{r}. \tag{11.115}$$

Exercise 11.3. *Solve Eq. (11.114) for the hydrogen atom potential.*

In this section, we shall determine algebraically the energy spectrum of the hydrogen atom. Consider the operator

$$\mathbf{R} = -\frac{Ze^2\mathbf{r}}{r} + \frac{1}{2m}\left(\mathbf{p}\times\mathbf{L} - \mathbf{L}\times\mathbf{p}\right), \tag{11.116}$$

where

$$\mathbf{L} = \mathbf{r}\times\mathbf{p}.$$

This is the quantum-mechanical analogue of the Runge–Lenz vector, a conserved quantity in classical mechanics for the Coulomb potential. Note that

$$\left[r_i, L^2\right] = i\hbar\epsilon_{ijk}(L_j r_k - r_j L_k) \tag{11.117a}$$

$$\left[p_i, L^2\right] = i\hbar\epsilon_{ijk}(L_j p_k - p_j L_k) \tag{11.117b}$$

or, equivalently,

$$\left[\mathbf{r}, L^2\right] = i\hbar(\mathbf{L}\times\mathbf{r} - \mathbf{r}\times\mathbf{L}) \tag{11.118a}$$

$$\left[\mathbf{p}, L^2\right] = i\hbar(\mathbf{L}\times\mathbf{p} - \mathbf{p}\times\mathbf{L}) \tag{11.118b}$$

We can then rewrite

$$R_i = -\frac{Ze^2 r_i}{r} + \frac{\left[L^2, p_i\right]}{2mi\hbar}. \tag{11.119}$$

From the above, the vector operator R_i can easily be shown to be Hermitian,

$$R_i^\dagger = R_i.$$

As anticipated from our experience with classical mechanics, it commutes with the Hamiltonian (**exercise**),

$$[H, R_i] = 0, \tag{11.120}$$

with

$$H = \frac{p^2}{2m} - \frac{Ze^2}{r}. \tag{11.121}$$

The Runge–Lenz vector is orthogonal to the angular momentum:

$$\mathbf{L}\cdot\mathbf{R} = \mathbf{R}\cdot\mathbf{L} = 0. \tag{11.122}$$

Also, we can prove the following identities (**exercise**):

$$\mathbf{L} \times \mathbf{p} = -\mathbf{p} \times \mathbf{L} + 2i\hbar\mathbf{p}\,, \tag{11.123a}$$

$$\begin{aligned} \mathbf{r} \cdot (\mathbf{p} \times \mathbf{L}) &= r_i \epsilon_{ijk} p_j L_k \\ &= (\epsilon_{kij} r_i p_j) L_k \\ &= L_k L_k \\ &= \mathbf{L}^2, \end{aligned} \tag{11.123b}$$

$$\begin{aligned} (\mathbf{p} \times \mathbf{L}) \cdot \mathbf{r} &= \epsilon_{ijk} p_j L_k r_i \\ &= \epsilon_{ijk} p_j \left([L_k, r_i] + r_i L_k\right) \\ &= i\hbar \epsilon_{kij} \epsilon_{ki\ell} p_j r_\ell + L^2 \\ &= L^2 + i\hbar 2\delta_{j\ell} p_j r_\ell \\ &= L^2 + 2i\hbar \mathbf{p} \cdot \mathbf{r}\,. \end{aligned} \tag{11.123c}$$

Similarly,

$$(\mathbf{p} \times \mathbf{L})^2 = p^2 L^2\,, \tag{11.124a}$$

$$\mathbf{p} \cdot (\mathbf{p} \times \mathbf{L}) = 0\,, \tag{11.124b}$$

$$(\mathbf{p} \times \mathbf{L}) \cdot \mathbf{p} = 2i\hbar p^2\,, \tag{11.124c}$$

and

$$\left[\frac{r_i}{r}, p_i\right] = \frac{2i\hbar}{r}. \tag{11.125}$$

For the square of the Runge–Lenz vector we find

$$\begin{aligned} R^2 &= \left(-\frac{Ze^2\mathbf{r}}{r} + \frac{\mathbf{p} \times \mathbf{L}}{m} - \frac{i\hbar}{m}\mathbf{p}\right)^2 \\ &= \dots \\ &= Z^2 e^4 + \frac{2}{m}\left(\frac{p^2}{2m} - \frac{Ze^2}{r}\right)\left(\hbar^2 + L^2\right). \end{aligned} \tag{11.126}$$

Thus,

$$R^2 = Z^2 e^4 + \frac{2H}{m}\left(\hbar^2 + L^2\right). \tag{11.127}$$

We can then determine the eigenvalues of the hydrogen-atom Hamiltonian if we know the eigenvalues of R^2. We shall do this by means of an algebraic method.

We find (**exercise**) the following commutation relations:

$$\left[R_i, R_j\right] = \frac{-2H}{m} i\hbar \epsilon_{ijk} L_k\,, \tag{11.128a}$$

$$\left[L_i, R_j\right] = i\hbar \epsilon_{ijk} R_k\,. \tag{11.128b}$$

We also recall

$$\left[L_i, L_j\right] = i\hbar\epsilon_{ijk}L_k\,. \tag{11.128c}$$

We now define

$$A_i^{\pm} = \frac{1}{2}\left(L_i \pm \sqrt{\frac{m}{-2H}}R_i\right). \tag{11.129}$$

The new operators satisfy the algebra

$$\left[A_i^{\pm}, A_j^{\pm}\right] = i\hbar\epsilon_{ijk}A_k^{\pm} \tag{11.130a}$$

and

$$\left[A_i^{\pm}, A_j^{\mp}\right] = 0\,, \tag{11.130b}$$

which is our familiar algebra of angular momentum. Notice that the operators $A^{\pm}$ are Hermitian if the Hamiltonian has negative eigenvalues, i.e. for negative energies. For the squares $(A^{\pm})^2$, we find that

$$\begin{aligned}(A^{\pm})^2 &= \frac{1}{4}\left[L^2 - \frac{m}{2H}R^2\right],\\ &= \frac{1}{4}\left[L^2 - \frac{m}{2H}\left(Z^2e^4 + \frac{2H}{m}\left(\hbar^2 + L^2\right)\right)\right]\\ \rightsquigarrow (A^{\pm})^2 &= -\frac{\hbar^2}{4} - \frac{m}{8H}Z^2e^4.\end{aligned} \tag{11.131}$$

According to the Lie algebra for the $A^{\pm}$ operators, the eigenvalues of $(A^{\pm})^2$ are

$$\hbar^2 j(j+1), \quad j = 0, \frac{1}{2}, 1, \frac{3}{2}, \ldots. \tag{11.132}$$

Thus, for the energy eigenvalues E, we must have

$$\hbar^2 j(j+1) = -\frac{\hbar^2}{4} - \frac{m}{8E_j}Z^2e^4, \tag{11.133}$$

or, equivalently,

$$E_n = -Z^2\frac{me^4}{2\hbar^2}\frac{1}{n^2}, \tag{11.134}$$

with

$$n = 2j + 1 = 1, 2, \ldots. \tag{11.135}$$

Notice that we have obtained only negative eigenvalues. For states with positive energy eigenvalues, the operators $A_i^{\pm}$ are not Hermitian and our derivation of the selection rules based on the Lie algebra does not go through.

Both $(A^{\pm})^2$ operators have a $(2j+1) = n$-fold degeneracy, which is revealed by the eigenvalues of the $A_3^{\pm}$ operators. Since the A^+ and A^- operators are independent, an energy level of the Hamiltonian has a $(2j+1)^2 = n^2$-fold degeneracy.

11.5.1 $SO(4)$ symmetry of the hydrogen atom

The Lie algebra for the group of rotations $SO(N)$ in N dimensions is given by

$$[J_{ab}, J_{cd}] = -i\hbar\left(\delta_{ad}J_{bc} - \delta_{ac}J_{bd} + \delta_{bc}J_{ad} - \delta_{bd}J_{ac}\right), \tag{11.136}$$

with $a, b, c, d = 1 \ldots N$ and $J_{ab} = -J_{ba}$. Obviously, this group contains rotations in three dimensions as a subgroup. For $J_{12} \equiv L_3, J_{23} = L_1, J_{31} = L_2$, the commutation relation above takes the familiar form

$$\left[L_i, L_j\right] = i\hbar\epsilon_{ijk}L_k, \tag{11.137a}$$

where $i, j, k = 1, 2, 3$. For the remaining three generators, J_{14}, J_{24}, J_{34}, the commutation relations yield

$$\left[J_{i4}, J_{j4}\right] = i\hbar\epsilon_{ijk}L_k, \quad i, j = 1 \ldots 3, \tag{11.137b}$$

and

$$\left[L_i, J_{j4}\right] = i\hbar\epsilon_{ijk}J_{k4}. \tag{11.137c}$$

Setting

$$R_i \equiv \sqrt{\frac{-2H}{m}}J_{i4}, \tag{11.138}$$

we recognise the Lie algebra of Eqs. (11.128a)–(11.128c). The hydrogen atom has an $SO(4)$ symmetry, which is a larger symmetry than rotation invariance in three dimensions, $SO(3)$.

12

Addition of angular momenta

A particle, such as an electron in an atom, can have more than one type of angular momentum, for example spin and orbital angular momentum. These two types will be described by *independent* generators of angular momenta living in different representations: J_a, J_b with

$$\left[J_{ai}, J_{bj}\right] = 0, \quad i, j = 1 \ldots 3 . \tag{12.1a}$$

As is required for angular-momentum generators, each set satisfies the same Lie algebra:

$$\left[J_{ai}, J_{aj}\right] = i\hbar\epsilon_{ijk} J_{ak} , \tag{12.1b}$$

$$\left[J_{bi}, J_{bj}\right] = i\hbar\epsilon_{ijk} J_{bk} . \tag{12.1c}$$

Therefore, there exist common eigenstates of the operators $J_a^2, J_b^2, J_{a3}, J_{b3}$, with eigenvalues

$$J_a^2 \left|j_a, m_a, j_b, m_b\right\rangle = \hbar^2 j_a(j_a+1) \left|j_a, m_a, j_b, m_b\right\rangle , \tag{12.2a}$$

$$J_b^2 \left|j_a, m_a, j_b, m_b\right\rangle = \hbar^2 j_b(j_b+1) \left|j_a, m_a, j_b, m_b\right\rangle , \tag{12.2b}$$

$$J_{a3} \left|j_a, m_a, j_b, m_b\right\rangle = \hbar m_a \left|j_a, m_a, j_b, m_b\right\rangle , \tag{12.2c}$$

$$J_{b3} \left|j_a, m_a, j_b, m_b\right\rangle = \hbar m_b \left|j_a, m_a, j_b, m_b\right\rangle . \tag{12.2d}$$

12.1 Addition of angular momenta

The sum of the two angular momenta,

$$J_i = J_{ai} + J_{bi}, \tag{12.3}$$

satisfies the same Lie algebra,

$$\left[J_i, J_j\right] = i\hbar\epsilon_{ijk} J_k . \tag{12.4}$$

Notice that the operators J_a^2 and J_b^2 commute with all the components of the total angular momentum:

$$\left[J_a^2, J_i\right] = \left[J_b^2, J_i\right] = 0 . \tag{12.5}$$

We can then find common eigenstates of the operators J^2, J_3, J_a^2, J_b^2,

$$\left|j, m, j_a, j_b\right\rangle ,$$

with

$$J_a^2 \left|j, m, j_a, j_b\right\rangle = \hbar^2 j_a(j_a+1) \left|j, m, j_a, j_b\right\rangle , \tag{12.6a}$$

$$J_b^2 \left|j, m, j_a, j_b\right\rangle = \hbar^2 j_b(j_b+1) \left|j, m, j_a, j_b\right\rangle , \tag{12.6b}$$

$$J^2 \left|j, m, j_a, j_b\right\rangle = \hbar^2 j(j+1) \left|j, m, j_a, j_b\right\rangle , \tag{12.6c}$$

$$J_3 \left|j, m, j_a, j_b\right\rangle = \hbar m \left|j, m, j_a, j_b\right\rangle . \tag{12.6d}$$

The states $\left|j, m, j_a, j_b\right\rangle$ can be written as a linear superposition of the states $\left|j_a, m_a, j_b, m_b\right\rangle$:

$$\left|j, m, j_a, j_b\right\rangle = \sum_{m_a, m_b} \left|j_a, m_a, j_b, m_b\right\rangle \left\langle j_a, m_a, j_b, m_b \middle| j, m, j_a, j_b\right\rangle . \tag{12.7}$$

The coefficients

$$\left\langle j_a, m_a, j_b, m_b \middle| j, m, j_a, j_b\right\rangle$$

are known as *Clebsch–Gordan coefficients*.

From Eq. (12.7), it is easy to infer that

$$\sum_{m_a, m_b} \left|\left\langle j_a, m_a, j_b, m_b \middle| j, m, j_a, j_b\right\rangle\right|^2 = 1 . \tag{12.8}$$

From the definition of the total angular momentum, we have

$$\begin{aligned} 0 &= \left\langle j_a, m_a, j_b, m_b\right| J_3 - J_{a3} - J_{b3} \left|j, m, j_a, j_b\right\rangle \\ &= (m - m_a - m_b) \left\langle j_a, m_a, j_b, m_b \middle| j, m, j_a, j_b\right\rangle . \end{aligned} \tag{12.9}$$

This leads to the conclusion that

$$\left\langle j_a, m_a, j_b, m_b \middle| j, m, j_a, j_b\right\rangle = \left\langle j_a, m_a, j_b, m_b \middle| j, m, j_a, j_b\right\rangle \delta_{m, m_a+m_b} . \tag{12.10}$$

Finally, we recall the raising and lowering operators of the eigenvalues m, m_a and m_b, respectively,

$$J_\pm = J_1 \pm i J_2 , \tag{12.11a}$$

$$J_{a\pm} = J_{a1} \pm i J_{a2} , \tag{12.11b}$$

$$J_{b\pm} = J_{b1} \pm i J_{b2} , \tag{12.11c}$$

with

$$J_\pm = J_{a\pm} + J_{b\pm} . \tag{12.11d}$$

From the last equation we obtain

$$\begin{aligned} \left\langle j_a, m_a, j_b, m_b\right| J_\pm \left|j, m, j_a, j_b\right\rangle = &\left\langle j_a, m_a, j_b, m_b\right| J_{a\pm} \left|j, m, j_a, j_b\right\rangle \\ &+ \left\langle j_a, m_a, j_b, m_b\right| J_{b\pm} \left|j, m, j_a, j_b\right\rangle , \end{aligned} \tag{12.12}$$

which yields

$$\begin{aligned}&\left\langle j_a, m_a, j_b, m_b \middle| j, m \pm 1, j_a, j_b \right\rangle \\ &= \sqrt{\frac{j_a(j_a+1) - m_a(m_a \mp 1)}{j(j+1) - m(m \pm 1)}} \left\langle j_a, m_a \mp 1, j_b, m_b \middle| j, m, j_a, j_b \right\rangle \\ &\quad + \sqrt{\frac{j_b(j_b+1) - m_b(m_b \mp 1)}{j(j+1) - m(m \pm 1)}} \left\langle j_a, m_a, j_b, m_b \mp 1 \middle| j, m, j_a, j_b \right\rangle .\end{aligned} \tag{12.13}$$

The maximum value of j is

$$j_{\max} = j_a + j_b$$

and there is only one such state,

$$\left| j = j_a + j_b, m = j_a + j_b, j_a, j_b \right\rangle = \left| j_a, m_a = j_a, j_b, m_b = j_b \right\rangle , \quad (\textbf{exercise}) .$$

Therefore,

$$\left\langle j_a, j_a, j_b, j_b \middle| j, m, j_a, j_b \right\rangle = \delta_{j, j_a+j_b} \delta_{m, j_a+j_b} . \tag{12.14}$$

Acting with J_- on both sides, we obtain a state with $j = j_a + j_b$ and $m = j_a + j_b - 1$,

$$\left| j_a + j_b, j_a + j_b - 1, j_a, j_b \right\rangle = \sqrt{\frac{j_a}{j_a + j_b}} \left| j_a, j_a - 1, j_b, j_b \right\rangle + \sqrt{\frac{j_b}{j_a + j_b}} \left| j_a, j_a, j_b, j_b - 1 \right\rangle . \tag{12.15}$$

There is only one more state with $m = j_a + j_b - 1$, and it must have $j = j_a + j_b - 1$. Since it has a different quantum number j, it must be orthogonal to the previous state:

$$0 = \left\langle j_a + j_b, j_a + j_b - 1, j_a, j_b \middle| j_a + j_b - 1, j_a + j_b - 1, j_a, j_b \right\rangle . \tag{12.16}$$

This condition allows us to determine the state (up to a phase):

$$\left| j_a + j_b - 1, j_a + j_b - 1, j_a, j_b \right\rangle = \sqrt{\frac{j_b}{j_a + j_b}} \left| j_a, j_a - 1, j_b, j_b \right\rangle - \sqrt{\frac{j_a}{j_a + j_b}} \left| j_a, j_a, j_b, j_b - 1 \right\rangle . \tag{12.17}$$

Remark. The overall sign chosen here is opposite to the standard Condon–Shortley convention, which requires $\left\langle j_a, j_a, j_b, j - j_a \middle| j, j, j_a, j_b \right\rangle > 0$ and is used in most tables of Clebsch–Gordan coefficients (e.g. by the Particle Data Group and in Sakurai's textbook). In that convention $\left| j_a + j_b - 1, j_a + j_b - 1 \right\rangle = -\sqrt{j_b/(j_a + j_b)} \left| j_a, j_a - 1, j_b, j_b \right\rangle + \sqrt{j_a/(j_a + j_b)} \left| j_a, j_a, j_b, j_b - 1 \right\rangle$. For $j_a = j_b = \frac{1}{2}$ the singlet is then $(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle)/\sqrt{2}$. Both choices are legitimate, but coefficients taken from such tables must not be combined with the sign convention used here.

We can repeat this procedure until the ranges of m_a and m_b are exhausted and no new state can be created. We find $(2j_a + 1)(2j_b + 1)$ states in total, and the quantum numbers of the total angular momentum take the discrete values $j = \left| j_a - j_b \right| , \ldots , j_a + j_b$ and $m = -j, \ldots , j$.

The above conventions lead to real-valued Clebsch–Gordan coefficients. Clebsch–Gordan coefficients are the coefficients for the linear transformation from a complete set of states

$$\left|j_1, m_1, j_2, m_2\right\rangle$$

to another complete set

$$\left|j, m, j_1, j_2\right\rangle .$$

Consider in general two complete sets of states:

$$\{\left|\psi_i\right\rangle\} \quad \text{and} \quad \{\left|\phi_i\right\rangle\}$$

with

$$\left\langle\phi_i\middle|\phi_j\right\rangle = \left\langle\psi_i\middle|\psi_j\right\rangle = \delta_{ij} ,$$

and the sets are related by a linear transformation with real coefficients:

$$\left|\phi_i\right\rangle = \sum_j C_{ij} \left|\psi_j\right\rangle . \tag{12.18}$$

Then

$$\begin{aligned} \delta_{kl} &= \left\langle\phi_k\middle|\phi_l\right\rangle \\ &= \sum_{jm} C_{km} C_{lj} \left\langle\psi_j\middle|\psi_m\right\rangle \\ &= \sum_j C_{kj} C_{lj} \\ &= \sum_j C_{kj} C^T_{jl} , \end{aligned} \tag{12.19}$$

which leads to

$$\det C = \pm 1$$

and

$$C^{-1} = C^T .$$

It follows that

$$\begin{aligned} C^T C &= 1 \\ \rightsquigarrow \sum_j C_{ji} C_{jk} &= \delta_{ik} . \end{aligned}$$

We can then invert Eq. (12.18) to obtain

$$\left|\psi_i\right\rangle = \sum_j C_{ji} \left|\phi_j\right\rangle .$$

Applying the above to the Clebsch–Gordan coefficients, we derive

$$\sum_{j,m} \left\langle j_1, m_a, j_2, m_b\middle|j, m, j_1, j_2\right\rangle \left\langle j_1, m_c, j_2, m_d\middle|j, m, j_1, j_2\right\rangle = \delta_{m_a,m_c} \delta_{m_b,m_d} , \tag{12.20}$$

and we can also invert Eq. (12.7) to write

$$\left|j_1, m_1, j_2, m_2\right\rangle = \sum_{j,m} \left|j, m, j_1, j_2\right\rangle \left\langle j_1, m_1, j_2, m_2\middle|j, m, j_1, j_2\right\rangle . \tag{12.21}$$

12.2 Application: the hydrogen atom

The hydrogen atom has an n^2 degeneracy, as we found earlier. The general solution for the wave function is of the form

$$\psi(\mathbf{r}) = \frac{R_{n\ell}(r)}{r} Y_\ell^m(\theta, \phi). \tag{12.22}$$

In this form, the degeneracy is due to the fact that the energy levels depend only on n and not on ℓ, m. Given that the number of degenerate states is finite, the value of ℓ must be bounded:

$$\ell = 0, \ldots, \ell_{\max}.$$

The number of degenerate states is then

$$n^2 = \sum_{\ell=0}^{\ell_{\max}} (2\ell + 1) = (\ell_{\max} + 1)^2. \tag{12.23}$$

From the above we conclude that the range of ℓ is

$$\ell = 0, \ldots, n - 1. \tag{12.24}$$

Remark. The counting argument assumes that each $\ell \leq \ell_{\max}$ occurs exactly once at a given n, which we have not shown. This follows from the $SO(4)$ algebra of the previous chapter: $\mathbf{L} = \mathbf{A}^+ + \mathbf{A}^-$ with $(A^+)^2 = (A^-)^2 = \hbar^2 j(j+1)$, so by the addition rule above $\ell = 0, 1, \ldots, 2j = n - 1$, each once (Gottfried–Yan, Sec. 5.3).

The total angular momentum of the electron in the hydrogen atom is the sum of its spin and its orbital angular momentum,

$$J = L + S. \tag{12.25}$$

In the ground state, $n = 1$, the orbital angular momentum is zero, and the only possibility for j is $j = \frac{1}{2}$. A general hydrogen state is denoted by

$$n\ell_j,$$

where $\ell = 0, 1, 2, 3, 4, \ldots, n - 1$ is represented by the letters

$$s, p, d, f, g, \text{ and then alphabetically.}$$

The hydrogen states are

$$1s_{\frac{1}{2}}, 2p_{\frac{3}{2}}, 2p_{\frac{1}{2}}, 2s_{\frac{1}{2}}, 3d_{\frac{5}{2}}, 3d_{\frac{3}{2}}, 3p_{\frac{3}{2}}, 3p_{\frac{1}{2}}, 3s_{\frac{1}{2}}, \ldots . \tag{12.26}$$

The interaction $\mathbf{L} \cdot \mathbf{S}$ splits the energy levels with the same n, ℓ but different j. This splitting is known as the fine structure of the hydrogen atom. An even finer splitting, the hyperfine splitting, is found by combining the spin of the electron with the spin of the proton; the corresponding interaction is due to the magnetic field of the proton.

12.3 Wigner–Eckart theorem

A set of $(2j+1)$ operators O_j^m is said to have spin j if

$$\left[J_3, O_j^m\right] = \hbar m O_j^m \tag{12.27}$$

and

$$\left[J_\pm, O_j^m\right] = \hbar\sqrt{j(j+1) - m(m\pm 1)}\, O_j^{m\pm 1}\,. \tag{12.28}$$

Notice that

$$\left\langle j, m'\right| J_3 \left|j, m\right\rangle = \hbar m \delta_{m',m} \tag{12.29}$$

and

$$\left\langle j, m'\right| J_\pm \left|j, m\right\rangle = \hbar\sqrt{j(j+1) - m(m\pm 1)}\,\delta_{m',m\pm 1}\,. \tag{12.30}$$

The coefficients are expressed in terms of the $(2j+1)$-dimensional representation of the angular momentum generators. We cast the above algebra in the form

$$\begin{aligned}\left[J_3, O_j^m\right] &= \sum_{m'} \left[J_3, O_j^m\right] \\ &= \sum_{m'} O_j^{m'} \left\langle j, m'\right| J_3 \left|j, m\right\rangle,\end{aligned} \tag{12.31a}$$

$$\left[J_\pm, O_j^m\right] = \sum_{m'} O_j^{m'} \left\langle j, m'\right| J_\pm \left|j, m\right\rangle, \tag{12.31b}$$

or, in general,

$$\left[J_i, O_j^m\right] = \sum_{m'} O_j^{m'} \left\langle j, m'\right| J_i \left|j, m\right\rangle. \tag{12.31c}$$

For example, a scalar (spin-0) operator, for which

$$\langle 0, 0| J_i |0, 0\rangle = 0, \quad \forall\, i = 1, 2, 3\,, \tag{12.32}$$

commutes with the angular momentum operators,

$$\left[J_i, O_0^0\right] = 0. \tag{12.33}$$

A "vector" operator $V_m \equiv O_1^m$ satisfies

$$\left[J_i, V_j\right] = i\hbar\epsilon_{ijk} V_k. \tag{12.34}$$

We can define the spherical components of such a vector as

$$V^{+1} = -\frac{V_1 + iV_2}{\sqrt{2}}\,, \tag{12.35a}$$

$$V^{-1} = \frac{V_1 - iV_2}{\sqrt{2}}\,, \tag{12.35b}$$

$$V^0 = V_3\,. \tag{12.35c}$$

Exercise 12.1. *Show that*

$$[J_3, V^m] = \hbar m V^m ,$$
$$[J_\pm, V^m] = \hbar\sqrt{1(1+1) - m(m\pm 1)}V^{m\pm 1} .$$

Exercise 12.2. *Prove that a spherical harmonic $Y_\ell^m(\hat{r})$, where $\hat{r}$ is treated as an operator, is an operator of spin ℓ in the above sense.*

Consider the state

$$\left|\Omega_{j_1 j_2}^{m_1 m_2}\right\rangle = O_{j_1}^{m_1}\left|a, j_2, m_2\right\rangle . \tag{12.36}$$

Acting on it with a rotation generator J_i, we obtain

$$\begin{aligned} J_i\left|\Omega_{j_1 j_2}^{m_1 m_2}\right\rangle &= \left[J_i, O_{j_1}^{m_1}\right]\left|a, j_2, m_2\right\rangle + O_{j_1}^{m_1} J_i\left|a, j_2, m_2\right\rangle \\ &= \sum_{m'}\left\langle j_1, m'\right| J_i\left|j_1, m_1\right\rangle\left|\Omega_{j_1 j_2}^{m' m_2}\right\rangle + \sum_{m'}\left\langle j_2, m'\right| J_i\left|j_2, m_2\right\rangle\left|\Omega_{j_1 j_2}^{m_1 m'}\right\rangle . \end{aligned} \tag{12.37}$$

We can compare this with the action of the rotation generators on a state which is a common eigenstate of two types of angular momentum in the j-th and j'-th representations,

$$\begin{aligned} &J_i\left|j_1, m_1, j_2, m_2\right\rangle \\ &= J_i^{(j)}\left|j_1, m_1, j_2, m_2\right\rangle + J_i^{(j')}\left|j_1, m_1, j_2, m_2\right\rangle \\ &= \sum_{m'}\left\langle j_2, m'\right| J_i\left|j_2, m_2\right\rangle\left|j_1, m_1, j_2, m'\right\rangle + \sum_{m'}\left\langle j_1, m'\right| J_i\left|j_1, m_1\right\rangle\left|j_1, m', j_2, m_2\right\rangle . \end{aligned} \tag{12.38}$$

We observe from Eqs. (12.37) and (12.38) that the state $\left|\Omega_{j_1 j_2}^{m_1 m_2}\right\rangle$ transforms as a common eigenstate of two different types of angular momentum. We can then change basis from $\left|\Omega_{j_1, j_2}^{m_1, m_2}\right\rangle$ to a basis $\left|\Omega_{j_1, j_2; j}^{m}\right\rangle$ with quantum numbers j, m, j_1, j_2. The coefficients of the linear transformation are the familiar Clebsch–Gordan coefficients, as in Eq. (12.21). We write

$$O_{j_1}^{m_1}\left|a, j_2, m_2\right\rangle = \sum_{jm}\left\langle j_1, m_1, j_2, m_2 | j, m, j_1, j_2\right\rangle\left|\Omega_{j_1, j_2; j}^{m}\right\rangle . \tag{12.39}$$

From the above, we have

$$\begin{aligned} &\left\langle a', j_3, m_3\right| O_{j_1}^{m_1}\left|a, j_2, m_2\right\rangle \\ &\qquad = \sum_{jm}\left\langle j_1, m_1, j_2, m_2 | j, m, j_1, j_2\right\rangle\left\langle a', j_3, m_3\middle|\Omega_{j_1, j_2; j}^{m}\right\rangle , \end{aligned} \tag{12.40a}$$

and, by using orthonormality,

$$\left\langle a', j, m\right| O_{j_1}^{m_1}\left|a, j_2, m_2\right\rangle = \left\langle j_1, m_1, j_2, m_2 | j, m, j_1, j_2\right\rangle\left\langle a', j, m\middle|\Omega_{j_1, j_2; j}^{m}\right\rangle . \tag{12.40b}$$

The matrix element

$$\left\langle a', j, m\middle|\Omega_{j_1, j_2; j}^{m}\right\rangle$$

is independent of the quantum number m. Indeed, for two states with the same j, m quantum numbers we have

$$\begin{aligned}\left\langle a, j, m \pm 1 \middle| b, j, m \pm 1 \right\rangle &= \left[\hbar^2 \left(j(j+1) - m(m \pm 1)\right)\right]^{-1} \left\langle a, j, m \middle| J_\mp J_\pm \middle| b, j, m \right\rangle \\ &= \left\langle a, j, m \middle| b, j, m \right\rangle .\end{aligned}$$

Therefore, their scalar product is independent of the quantum number m. We have thus proven the *Wigner–Eckart theorem*.

Theorem 12.1 (Wigner–Eckart theorem). *The matrix elements of operators of definite spin satisfy*

$$\left\langle a', j_3, m_3 \middle| O^{m_2}_{j_2} \middle| a, j_1, m_1 \right\rangle = \left\langle j_2, m_2, j_1, m_1 \middle| j_3, m_3, j_2, j_1 \right\rangle \langle a' | \| O \| | a \rangle ,$$

where $\langle a' | \| O \| | a \rangle$ is known as the reduced matrix element; it is independent of the quantum numbers m_1, m_2, m_3.

Remark. The reduced matrix element depends on j_1, j_2, j_3 (and on a, a'), but not on m_1, m_2, m_3. The theorem holds with Condon–Shortley Clebsch–Gordan coefficients and the tensor-operator definition above; a different sign convention for the Clebsch–Gordan coefficients, such as the one adopted earlier in this chapter, changes only the sign of the reduced matrix element. Sakurai (Eq. 3.11.31) uses the order $\left\langle j_1 m_1; j_2 m_2 \middle| j_3 m_3 \right\rangle$ and an extra factor $1/\sqrt{2j_1 + 1}$, which only redefines $\langle a' | \| O \| | a \rangle$. As an example, one can verify explicitly that the ratio $\langle \ell_3 m_3 | \, Y^q_1 \, | \ell_1 m_1 \rangle \, / \left\langle 1, q, \ell_1, m_1 \middle| \ell_3, m_3 \right\rangle$ is independent of m_1, q and m_3.

The coefficient

$$\left\langle j_2, m_2, j_1, m_1 \middle| j_3, m_3, j_2, j_1 \right\rangle$$

is a *Clebsch–Gordan coefficient*. The Wigner–Eckart theorem is of paramount importance for understanding atomic transitions through the absorption and emission of light.

13

Discrete symmetries

In this chapter, we shall discuss symmetry transformations which cannot be connected to the identity by varying a continuous parameter. Examples of such symmetries are *space inversion* (mirror symmetry or parity) and *time reversal*.

13.1 Parity

Under a parity transformation in classical physics, the position vector changes sign,

$$\mathbf{r} \longrightarrow -\mathbf{r}\,. \tag{13.1}$$

We would like to preserve this property for the expectation value of the position operator,

$$\langle\psi|\,\hat{\mathbf{r}}\,|\psi\rangle \longrightarrow -\langle\psi|\,\hat{\mathbf{r}}\,|\psi\rangle\,. \tag{13.2}$$

Under a parity transformation, a state changes as

$$|\psi\rangle \longrightarrow |\psi'\rangle = \Pi\,|\psi\rangle\,, \tag{13.3}$$

which suggests that we can achieve our requirement of Eq. (13.2) if

$$\Pi^\dagger \hat{\mathbf{r}} \Pi = -\hat{\mathbf{r}}. \tag{13.4}$$

According to the above,

$$\begin{aligned} \hat{\mathbf{r}}\,(\Pi\,|\mathbf{r}\rangle) &= -\Pi\,(\hat{\mathbf{r}}\,|\mathbf{r}\rangle) \\ &= -\mathbf{r}(\Pi\,|\mathbf{r}\rangle)\,. \end{aligned} \tag{13.5}$$

Therefore, the state $\Pi\,|\mathbf{r}\rangle$ describes the same physics as the state $|-\mathbf{r}\rangle$. We then have

$$\Pi\,|\mathbf{r}\rangle = e^{i\delta}\,|-\mathbf{r}\rangle\,. \tag{13.6}$$

By convention, we choose the phase to be zero:

$$\Pi\,|\mathbf{r}\rangle = |-\mathbf{r}\rangle\,. \tag{13.7}$$

Acting twice on a position state, the parity operator has no effect:

$$\begin{aligned} \Pi^2\,|\mathbf{r}\rangle &= \Pi\,|-\mathbf{r}\rangle \\ &= |\mathbf{r}\rangle\,. \end{aligned} \tag{13.8}$$

Given that the position eigenkets form a complete basis, we have

$$\Pi^2 = 1\,. \tag{13.9}$$

Thus, the representation Π of the parity transformation is unitary and hermitian, satisfying

$$\Pi^\dagger = \Pi^{-1} = \Pi\,. \tag{13.10}$$

Let us consider the space-translation operator

$$U(\Delta\mathbf{r}) = 1 - i\frac{\mathbf{p}\cdot\Delta\mathbf{r}}{\hbar}$$

acting on a position eigenket $|\mathbf{r}\rangle$. We have

$$\Pi U(\Delta\mathbf{r})\,|\mathbf{r}\rangle = |-\mathbf{r} - \Delta\mathbf{r}\rangle\,. \tag{13.11}$$

Also,

$$U(-\Delta\mathbf{r})\Pi\,|\mathbf{r}\rangle = |-\mathbf{r} - \Delta\mathbf{r}\rangle\,. \tag{13.12}$$

Equating the left-hand sides of the last two equations, we have

$$\Pi U(\Delta\mathbf{r}) = U\,(-\Delta\mathbf{r})\,\Pi\,, \tag{13.13}$$

$$\rightsquigarrow \Pi U(\Delta\mathbf{r})\Pi = U(-\Delta\mathbf{r})\,. \tag{13.14}$$

This leads to the following transformation of the momentum operator under parity:

$$\Pi\,\hat{\mathbf{p}}\,\Pi = -\hat{\mathbf{p}}\,. \tag{13.15}$$

We can repeat this analysis for angular momentum, the generator of rotations. An infinitesimal rotation is represented by

$$U(\omega_{ij}) = 1 + \frac{i}{2}\omega_{ij}J^{ij}\,, \tag{13.16}$$

and

$$U(\omega_{ij})\,|r_i\rangle = \left|r_i + \omega_{ij}r_j\right\rangle. \tag{13.17}$$

We find that

$$\begin{aligned}\Pi U(\omega_{ij})\Pi\,|r_i\rangle &= \Pi U(\omega_{ij})\,|-r_i\rangle \\ &= \Pi\left|-r_i - \omega_{ij}r_j\right\rangle \\ &= \left|r_i + \omega_{ij}r_j\right\rangle \\ &= U(\omega_{ij})\,|r_i\rangle\,.\end{aligned} \tag{13.18}$$

Therefore,

$$\Pi U(\omega_{ij})\Pi = U(\omega_{ij})\,. \tag{13.19}$$

For the angular momentum operators, the above equation implies that

$$\Pi\,J_i\Pi = J_i\,, \tag{13.20}$$

where we recall that $\hbar J^{12} = J_3, \hbar J^{23} = J_1$ and $\hbar J^{31} = J_2$.

For a system which is symmetric under parity, the energy eigenstates are also parity eigenstates.

Remark. This holds for non-degenerate energy eigenstates; for degenerate levels, one can only choose a basis of simultaneous eigenstates of H and Π. For example, the free-particle state $e^{i\mathbf{k}\cdot\mathbf{r}}$ is an energy eigenstate but not a parity eigenstate (Sakurai, Sec. 4.2).

Let us consider the eigenstates of the parity operator,

$$\Pi|\pi\rangle = \pi|\pi\rangle\,. \tag{13.21}$$

The parity operator satisfies

$$\Pi^2 = 1\,. \tag{13.22}$$

The same identity must be fulfilled by the eigenvalues. Therefore, there are two types of parity eigenstates, $|\pi\rangle = |\pm\rangle$, with eigenvalues ± 1:

$$\Pi|\pm\rangle = \pm|\pm\rangle\,. \tag{13.23}$$

Multiplying by a position bra, we find

$$\langle\mathbf{r}|\Pi|\pm\rangle = \pm\langle\mathbf{r}|\pm\rangle\,, \tag{13.24}$$

$$\langle -\mathbf{r}|\pm\rangle = \pm\langle\mathbf{r}|\pm\rangle\,. \tag{13.25}$$

As we have remarked above, in a parity-conserving system the energy eigenstates can be chosen to be parity eigenstates. Their wave functions,

$$\psi_\pm(\mathbf{r}) \equiv \langle\mathbf{r}|\pm\rangle\,,$$

are therefore either even or odd under $\mathbf{r} \to -\mathbf{r}$:

$$\psi_\pm(\mathbf{r}) = \pm\psi_\pm(-\mathbf{r})\,. \tag{13.26}$$

13.2 Time reversal

Let us consider a system which is symmetric under time reversal,

$$t \to -t\,.$$

In this section, we shall explore the properties of the representation Θ of time reversal on quantum states:

$$|\psi\rangle \to \Theta|\psi\rangle\,.$$

Classically, time reversal is the "motion picture run backwards": a particle which passes through the point $\mathbf{r}$ with momentum $\mathbf{p}$ is replaced by a particle which passes through the same point with momentum $-\mathbf{p}$, and retraces its trajectory. In the following, $|\psi, -t\rangle$ denotes the time-reversed state $\Theta|\psi, t\rangle$ in this sense. The time-translation operator for infinitesimal time intervals is

$$U(\delta t) = 1 - i\frac{H}{\hbar}\delta t\,. \tag{13.27}$$

Acting on a generic state $\left|\psi, t\right\rangle$, we have

$$\begin{aligned}\Theta^{-1}U(-\delta t)\Theta\left|\psi, t\right\rangle &= \Theta^{-1}U(-\delta t)\left|\psi, -t\right\rangle \\ &= \Theta^{-1}\left|\psi, -t-\delta t\right\rangle \\ &= \left|\psi, t+\delta t\right\rangle \\ &= U(\delta t)\left|\psi, t\right\rangle .\end{aligned} \tag{13.28}$$

Therefore,

$$\Theta^{-1}U(-\delta t)\Theta = U(\delta t) . \tag{13.29}$$

Substituting the expression for the operator of an infinitesimally small time evolution, we have

$$\Theta^{-1}(-iH)\Theta = iH. \tag{13.30}$$

The representation Θ cannot be unitary and linear. If it were, we would derive

$$\Theta^{-1}H\Theta = -H \tag{13.31}$$

and, equivalently,

$$\Theta H = -H\Theta. \tag{13.32}$$

Assume that there is a state $|E\rangle$ which is an eigenstate of the Hamiltonian:

$$H\left|E\right\rangle = E\left|E\right\rangle . \tag{13.33}$$

Then, for the state $\Theta\left|E\right\rangle$, we find that it is also an energy eigenstate, but with the energy eigenvalue $-E$:

$$\begin{aligned}H\left(\Theta\left|E\right\rangle\right) &= -\Theta\left(H\left|E\right\rangle\right) \\ &= -E\left(\Theta\left|E\right\rangle\right) .\end{aligned} \tag{13.34}$$

This contradicts observations: for free particles, it would erroneously predict negative energies. The spectrum of H would have to be symmetric under $E \to -E$, and there would be no stable ground state. However, the spectrum of a free particle, $H = \mathbf{p}^2/2m$, is $[0, \infty)$, and, more generally, the energy of a physical system is bounded from below. According to Wigner's theorem, which we discussed in the chapter on symmetries in quantum mechanics, a transformation which preserves all transition probabilities is represented either by a linear unitary or by an antilinear antiunitary operator. The first option has failed.

The above problem is solved if Θ is antilinear and antiunitary. Then the factor i is complex conjugated when it is moved through Θ, $\Theta^{-1}\, i\, \Theta = -i$, and Eq. (13.30) gives

$$\begin{aligned}&\Theta^{-1}(-iH)\Theta = iH \\ \rightsquigarrow\ &i\Theta^{-1}H\Theta = iH \\ \rightsquigarrow\ &\Theta^{-1}H\Theta = H .\end{aligned} \tag{13.35}$$

Equivalently,

$$\Theta H = H\Theta . \tag{13.36}$$

Notice that this is a statement about the Hamiltonian of a particular system: it holds if the system is symmetric under time reversal. We will see below that this is not the case, for example, for a charged particle in an external magnetic field.

13.2.1 Antilinear and antiunitary operators

Let us collect the salient properties of the operators that we will need. An operator Θ is **antilinear** if

$$\Theta\left(a\left|\psi\right\rangle + b\left|\phi\right\rangle\right) = a^*\,\Theta\left|\psi\right\rangle + b^*\,\Theta\left|\phi\right\rangle. \tag{13.37}$$

It is **antiunitary** if, in addition, it is invertible and, denoting $\left|\Theta\psi\right\rangle \equiv \Theta\left|\psi\right\rangle$,

$$\left\langle\Theta\phi\middle|\Theta\psi\right\rangle = \left\langle\psi\middle|\phi\right\rangle = \left\langle\phi\middle|\psi\right\rangle^*. \tag{13.38}$$

Transition probabilities $\left|\left\langle\phi\middle|\psi\right\rangle\right|^2$ are then preserved, as they should be. Notice that we avoid writing $\Theta^\dagger$ or letting Θ act to the left on a bra; Θ always acts on the ket to its right.

The simplest antiunitary operator is the complex conjugation K with respect to a chosen basis $\{|n\rangle\}$:

$$K\sum_n c_n\left|n\right\rangle = \sum_n c_n^*\left|n\right\rangle, \qquad K^2 = 1. \tag{13.39}$$

It is important to notice that K depends on the basis: in the position basis it maps $\psi(\mathbf{r}) \to \psi^*(\mathbf{r})$, and in the momentum basis $\tilde\psi(\mathbf{p}) \to \tilde\psi^*(\mathbf{p})$, which are different operations. If Θ is antiunitary, then $U \equiv \Theta K$ is linear (the product of two antilinear operators) and unitary. Therefore, every antiunitary operator can be written as

$$\Theta = UK, \qquad U \text{ unitary}. \tag{13.40}$$

Matrix elements. Let O be a linear operator. Using Eq. (13.38) with $\left|\phi\right\rangle = O\left|b\right\rangle$ and $\left|\psi\right\rangle = \left|a\right\rangle$, and $\Theta O\left|b\right\rangle = \left(\Theta O\Theta^{-1}\right)\Theta\left|b\right\rangle$, we find

$$\left\langle a\right|O\left|b\right\rangle = \left\langle\Theta O b\middle|\Theta a\right\rangle = \left\langle\Theta b\right|\left(\Theta O\Theta^{-1}\right)^\dagger\left|\Theta a\right\rangle. \tag{13.41}$$

The operator $\Theta O\Theta^{-1}$ is linear, so its hermitian conjugate is the usual one. If O is hermitian, so is $\Theta O\Theta^{-1}$ (**exercise**), and

$$\left\langle a\right|O\left|b\right\rangle = \left\langle\Theta b\right|\Theta O\Theta^{-1}\left|\Theta a\right\rangle, \qquad O = O^\dagger. \tag{13.42}$$

The initial and final states are interchanged, as we expect when we run the film backwards. For $a = b = \psi$, the expectation value of a hermitian operator in the time-reversed state $\left|\psi'\right\rangle = \Theta\left|\psi\right\rangle$ is

$$\left\langle\psi'\right|O\left|\psi'\right\rangle = \left\langle\psi\right|\Theta^{-1}O\,\Theta\left|\psi\right\rangle. \tag{13.43}$$

In this section, we shall write the transformation of operators as $\Theta^{-1}O\,\Theta$. For the operators that we encounter, $\Theta^{-1}O\,\Theta = \pm O$ with a **real** sign. Multiplying by Θ from the left and Θ^{-1} from the right, this is the same as $\Theta O\,\Theta^{-1} = \pm O$; both forms can be used.

13.2.2 Position, momentum and wave functions

Classically, time reversal acts as

$$\mathbf{r} \to \mathbf{r}, \tag{13.44a}$$

$$\mathbf{p} \to -\mathbf{p}. \tag{13.44b}$$

The classical transformation rules should also hold for the expectation values of the corresponding quantum operators. Indeed, by Ehrenfest's theorem, $\frac{d}{dt}\langle\mathbf{r}\rangle = \langle\mathbf{p}\rangle/m$. If the

film runs backwards, the position is the same at each instant, but its rate of change flips sign. Thus, we must have

$$\Theta^{-1}\hat{\mathbf{r}}\Theta = \hat{\mathbf{r}} \tag{13.45}$$

and

$$\Theta^{-1}\hat{\mathbf{p}}\Theta = -\hat{\mathbf{p}}\,. \tag{13.46}$$

This is consistent with the commutation relations:

$$\begin{aligned}
\Theta^{-1}\left[r_i, p_j\right]\Theta &= \Theta^{-1} i\hbar\delta_{ij}\Theta \\
\left[\Theta^{-1} r_i \Theta, \Theta^{-1} p_j \Theta\right] &= -i\hbar\delta_{ij} \\
\left[r_i, -p_j\right] &= -i\hbar\delta_{ij} \\
\left[r_i, p_j\right] &= i\hbar\delta_{ij}\,.
\end{aligned} \tag{13.47}$$

It is very important to notice that the antilinearity of Θ has been used in the second line, $\Theta^{-1} i\hbar\,\Theta = -i\hbar$. For a linear Θ, the left-hand side would be $[r_i, -p_j] = -i\hbar\delta_{ij}$ and the right-hand side $+i\hbar\delta_{ij}$, which is a contradiction. The canonical commutation relations alone tell us, once again, that time reversal must be antilinear.

From Eq. (13.45), $\hat{\mathbf{r}}\,(\Theta\,|\mathbf{r}\rangle) = \Theta\hat{\mathbf{r}}\,|\mathbf{r}\rangle = \mathbf{r}\,(\Theta\,|\mathbf{r}\rangle)$, where we used that the eigenvalue $\mathbf{r}$ is real. Thus $\Theta\,|\mathbf{r}\rangle$ is a position eigenstate with the same eigenvalue, and we can choose its phase such that

$$\Theta\,|\mathbf{r}\rangle = |\mathbf{r}\rangle\,. \tag{13.48}$$

For a spinless particle, Θ is then the complex conjugation K in the position basis. For a generic state,

$$\Theta\,|\psi\rangle = \Theta\int d^3\mathbf{r}\,\psi(\mathbf{r})\,|\mathbf{r}\rangle = \int d^3\mathbf{r}\,\psi^*(\mathbf{r})\,|\mathbf{r}\rangle\,, \qquad \text{i.e.} \qquad \psi(\mathbf{r}) \to \psi^*(\mathbf{r}). \tag{13.49}$$

This is familiar! If $\psi(\mathbf{r}, t)$ solves the Schrödinger equation

$$i\hbar\partial_t\psi = \left[-\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{r})\right]\psi$$

with a real potential, then $\psi^*(\mathbf{r}, -t)$ is also a solution. For a momentum eigenstate, with $\langle\mathbf{r}|\mathbf{p}\rangle = e^{i\mathbf{p}\cdot\mathbf{r}/\hbar}/(2\pi\hbar)^{3/2}$,

$$\Theta\,|\mathbf{p}\rangle = \int d^3\mathbf{r}\,\frac{e^{-i\mathbf{p}\cdot\mathbf{r}/\hbar}}{(2\pi\hbar)^{3/2}}\,|\mathbf{r}\rangle = |-\mathbf{p}\rangle\,, \tag{13.50}$$

in agreement with Eq. (13.46). A plane wave travelling to the right becomes a plane wave travelling to the left.

13.2.3 Angular momentum

For orbital angular momentum the transformation follows immediately from Eqs. (13.45)–(13.46):

$$\Theta^{-1}\mathbf{L}\,\Theta = \Theta^{-1}\hat{\mathbf{r}}\,\Theta \times \Theta^{-1}\hat{\mathbf{p}}\,\Theta = -\mathbf{L}. \tag{13.51}$$

For spin, there is no such classical expression. Let us see what we can learn from the Lie algebra. The Lie algebra of angular momentum is

$$\left[J_i, J_j\right] = i\hbar\epsilon_{ijk}J_k\,. \tag{13.52}$$

This leads to

$$\left[-\Theta^{-1}J_i\Theta, -\Theta^{-1}J_j\Theta\right] = i\hbar\epsilon_{ijk}\left(-\Theta^{-1}J_k\Theta\right), \tag{13.53}$$

which is consistent with

$$\Theta^{-1}J_i\Theta = -J_i\,. \tag{13.54}$$

However, the algebra alone does not imply Eq. (13.54): it only tells us that the operators $-\Theta^{-1}J_i\Theta$ satisfy the same commutation relations as J_i. For example, $-\Theta^{-1}J_i\Theta = R_{ij}J_j$ with any rotation matrix R would also be consistent with it. The argument which fixes the transformation uses that $\mathbf{J}$ is the generator of rotations, and that time reversal commutes with rotations: a rotated film run backwards is the same as the film run backwards and then rotated. For a spinless particle we can verify this directly. With $U(R)\,|\mathbf{r}\rangle = |R\mathbf{r}\rangle$ and Eq. (13.48),

$$\Theta\,U(R)\,\Theta^{-1}\,|\mathbf{r}\rangle = \Theta\,U(R)\,|\mathbf{r}\rangle = \Theta\,|R\mathbf{r}\rangle = |R\mathbf{r}\rangle = U(R)\,|\mathbf{r}\rangle\,, \tag{13.55}$$

and, since the position eigenstates form a complete basis,

$$\Theta\,U(R)\,\Theta^{-1} = U(R). \tag{13.56}$$

For particles with spin, we **require** Eq. (13.56): time reversal should not depend on the orientation of our coordinate axes. Now we insert an infinitesimal rotation, $U(\omega_{ij}) = 1 + \frac{i}{2}\omega_{ij}J^{ij}$, with real parameters ω_{ij}. Because Θ is antilinear,

$$\Theta\,U(\omega_{ij})\,\Theta^{-1} = 1 - \frac{i}{2}\omega_{ij}\,\Theta J^{ij}\Theta^{-1} \overset{!}{=} 1 + \frac{i}{2}\omega_{ij}J^{ij} \quad\rightsquigarrow\quad \Theta J^{ij}\Theta^{-1} = -J^{ij}. \tag{13.57}$$

Thus $\Theta\mathbf{J}\,\Theta^{-1} = -\mathbf{J}$ or, equivalently, Eq. (13.54). Notice that a linear Θ commuting with rotations would give $+\mathbf{J}$ instead. Angular momentum is odd under time reversal, like the classical angular momentum $\mathbf{r}\times\mathbf{p}$: a spinning top spins in the opposite sense when the film runs backwards.

> **Exercise 13.1.** *Show that a Hermitian O gives a Hermitian $\Theta O\Theta^{-1}$, and that $\Theta J_\pm\Theta^{-1} = -J_\mp$, where $J_\pm = J_1 \pm iJ_2$.*

13.2.4 Time reversal of angular momentum eigenstates

Consider the eigenstates $|j,m\rangle$ of J^2 and J_3, with the ladder operators of Chapter 11 (Condon–Shortley phases),

$$J_\pm\,|j,m\rangle = \hbar\sqrt{j(j+1) - m(m\pm1)}\;|j,m\pm1\rangle\,.$$

From $\Theta J_3\Theta^{-1} = -J_3$ and $\Theta J^2\Theta^{-1} = J^2$,

$$J_3\left(\Theta\,|j,m\rangle\right) = -\Theta J_3\,|j,m\rangle = -\hbar m\left(\Theta\,|j,m\rangle\right), \tag{13.58}$$

so that $\Theta\,|j,m\rangle = c_m\,|j,-m\rangle$. The phases c_m are related through $\Theta J_+ = -J_-\Theta$. Acting on $|j,m\rangle$, and noticing that the square roots are real and equal on the two sides, we find $c_{m+1} = -c_m$. Choosing $c_j = 1$, we obtain

$$\Theta\,|j,m\rangle = (-1)^{j-m}\,|j,-m\rangle\,. \tag{13.59}$$

This is our phase convention. Other books use $i^{2m}\left|j,-m\right\rangle$, which differs by the overall phase i^{2j}. Acting twice,

$$\Theta^2\left|j,m\right\rangle = (-1)^{j-m}\,\Theta\left|j,-m\right\rangle = (-1)^{j-m}(-1)^{j+m}\left|j,m\right\rangle \quad\rightsquigarrow\quad \Theta^2 = (-1)^{2j}. \tag{13.60}$$

Notice that Θ^2 does not depend on the phase convention: if we replace $\Theta \to e^{i\alpha}\Theta$, then $\left(e^{i\alpha}\Theta\right)^2 = e^{i\alpha}e^{-i\alpha}\Theta^2 = \Theta^2$. For integer j (e.g. orbital angular momentum) $\Theta^2 = 1$; for half-integer j, $\Theta^2 = -1$. For a system of many particles, $\Theta^2 = +1$ (-1) for an even (odd) number of spin-$\frac{1}{2}$ particles.

Spin $\frac{1}{2}$. In the basis $|\uparrow\rangle = \left|\frac{1}{2},\frac{1}{2}\right\rangle$, $|\downarrow\rangle = \left|\frac{1}{2},-\frac{1}{2}\right\rangle$, Eq. (13.59) reads $\Theta|\uparrow\rangle = |\downarrow\rangle$, $\Theta|\downarrow\rangle = -|\uparrow\rangle$, i.e. $\Theta = UK$ with

$$U = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = -i\sigma_2, \qquad \Theta = -i\sigma_2 K. \tag{13.61}$$

We can verify directly that $\Theta\sigma_k\Theta^{-1} = \sigma_2\sigma_k^*\sigma_2 = -\sigma_k$ for $k = 1,2,3$, since σ_1, σ_3 are real and σ_2 is imaginary. Also,

$$\Theta^2 = (-i\sigma_2)K(-i\sigma_2)K = (-i\sigma_2)(-i\sigma_2)^* = -\sigma_2^2 = -1. \tag{13.62}$$

A spin-$\frac{1}{2}$ state returns to minus itself when time is reversed twice, in the same way as it does after a rotation by 2π.

> **Exercise 13.2.** *Show that $\Theta = e^{-i\pi J_2/\hbar}K$, with K the complex conjugation in the basis $\left|j,m\right\rangle$, reproduces Eq. (13.59). Verify it explicitly for $j = \frac{1}{2}$ (where it gives Eq. (13.61)) and for $j = 1$, using the matrices J_1, J_2, J_3 of spin 1. Check that $\Theta\mathbf{J}\,\Theta^{-1} = -\mathbf{J}$ and $\Theta^2 = +1$ for $j = 1$. (Hint: J_1 and J_3 are real and J_2 is imaginary in this basis.)*

13.2.5 Time-reversal invariance and its consequences

A system is symmetric under time reversal if Eq. (13.36) holds, $[\Theta, H] = 0$. For example, $H = \frac{\mathbf{p}^2}{2m} + V(\mathbf{r})$ with a real potential is time-reversal invariant, and so is a spin-orbit interaction $f(r)\,\mathbf{L}\cdot\mathbf{S}$, as a product of two odd operators. A particle in an external magnetic field is **not**: the term $-\gamma\,\mathbf{S}\cdot\mathbf{B}$ changes sign, because we reverse the spin but not the currents which produce $\mathbf{B}$. If $[\Theta, H] = 0$ and $H|E\rangle = E|E\rangle$, then

$$H\left(\Theta|E\rangle\right) = \Theta H|E\rangle = E\left(\Theta|E\rangle\right). \tag{13.63}$$

In contrast to our failed attempt with a linear Θ, the time-reversed state has the **same** energy. There are three important consequences.

Real wave functions. For a spinless particle, $\Theta = K$ in the position basis. If the level E is non-degenerate, $\Theta|E\rangle$ must be the same state up to a phase, $\psi_E^*(\mathbf{r}) = e^{i\alpha}\psi_E(\mathbf{r})$. Then $e^{i\alpha/2}\psi_E(\mathbf{r})$ is real: the wave function of a non-degenerate energy eigenstate can be chosen real. For a degenerate level, we can always choose a basis of real wave functions, $\psi + \psi^*$ and $i(\psi - \psi^*)$. This is why the bound-state wave functions of one-dimensional problems, which are non-degenerate, came out real. A real wave function carries no probability current, $\mathbf{j} \propto \mathrm{Im}\left(\psi^*\nabla\psi\right) = 0$.

Kramers degeneracy. Let $\Theta^2 = -1$, i.e. an odd number of spin-$\frac{1}{2}$ particles, and $[\Theta, H] = 0$. Assume that $\Theta|E\rangle = c\,|E\rangle$. Then

$$-|E\rangle = \Theta^2|E\rangle = \Theta\,c\,|E\rangle = c^*\,\Theta|E\rangle = |c|^2\,|E\rangle, \tag{13.64}$$

which is impossible. Thus $\Theta|E\rangle$ is a different state with the same energy. In fact, it is orthogonal to $|E\rangle$: from Eq. (13.38),

$$\langle E|\Theta E\rangle = \langle \Theta\Theta E|\Theta E\rangle = -\langle E|\Theta E\rangle = 0. \tag{13.65}$$

Every energy level of such a system is at least twofold degenerate. This is known as **Kramers degeneracy**. It holds for an electron in an arbitrarily complicated electric field (e.g. in a crystal), which can break all the rotational symmetry. Only a magnetic field, which breaks time-reversal invariance, can lift it.

No permanent electric dipole moment. Consider a state $|\alpha; j, m\rangle$ of a rotation-invariant and time-reversal-invariant system, where α labels the energy level and the level has no degeneracy other than the $2j+1$ values of m. The electric dipole operator $\mathbf{d} = \sum_a q_a \hat{\mathbf{r}}_a$ is even under time reversal, $\Theta \mathbf{d}\, \Theta^{-1} = \mathbf{d}$, while $\mathbf{J}$ is odd. The state $\Theta|\alpha; j, m\rangle$ has the same energy and $J_3 = -\hbar m$, so $\Theta|\alpha; j, m\rangle = \pm|\alpha; j, -m\rangle$ (the phase is irrelevant here). From Eq. (13.42) with $a = b$,

$$\langle \alpha; j, m|\, d_3\, |\alpha; j, m\rangle = \langle \alpha; j, -m|\, d_3\, |\alpha; j, -m\rangle\,. \tag{13.66}$$

On the other hand, a rotation by π around the 1-axis maps $|\alpha; j, m\rangle$ to $|\alpha; j, -m\rangle$ up to a phase and $d_3 \to -d_3$, so that $\langle \alpha; j, -m|\, d_3\, |\alpha; j, -m\rangle = -\langle \alpha; j, m|\, d_3\, |\alpha; j, m\rangle$. Therefore $\langle d_3\rangle = 0$ (and $\langle d_{1,2}\rangle = 0$, since $d_{1,2}$ change m). In other words, by the Wigner–Eckart theorem (Section 12.3) a permanent dipole moment would have to be proportional to $\mathbf{J}$, $\langle \mathbf{d}\,\rangle = c\,\langle \mathbf{J}\,\rangle$, and this is incompatible with time-reversal invariance, since $\mathbf{d}$ is even and $\mathbf{J}$ is odd. (A parity-symmetric system has no permanent dipole moment either, since $\mathbf{d}$ is odd under parity.) The linear Stark effect of the excited states of hydrogen is not a counter-example: there the $2s$ and $2p$ levels are degenerate.

This makes the search for a permanent electric dipole moment of the neutron a very sensitive test of time-reversal symmetry. No such moment has been found: the present limit, from an experiment at PSI, is $|d_n| < 1.8 \times 10^{-26}\, e$ cm. For comparison, the size of the neutron is about 10^{-13} cm. Time reversal is known to be violated in weak interactions (in the decays of K and B mesons), but the predicted neutron dipole moment from this source is many orders of magnitude below the present limit.

PART II

Quantum Mechanics II

Preface to Part II

The second part of the script covers approximation methods (time-independent and time-dependent perturbation theory, the variational method), the theory of elastic scattering, the interaction of radiation with matter, identical particles, multi-electron atoms and an introduction to quantised fields. It builds directly on Part I, and the same recommendation applies: the script accompanies the lectures and does not replace the study of the textbooks.

Acknowledgements

The notes of Part II were transcribed from my handwritten lecture notes and typeset by Maximilian Krahn and Somesh Vijayananda. Deniz Coskun contributed the appendix “An alternative derivation of the Lippmann–Schwinger equation” (Appendix C), and Robin Reding and Ophélie Rivière sent many corrections. I thank them all.

14

Perturbation Theory

Perturbation theory is a mathematical method used to find approximate solutions to a problem, by starting from an exact solution of a related but simpler problem. The essence of the method is to split the problem into a solvable part and a perturbation, which is assumed to be small.

Perturbation theory originates in the celestial mechanics of the 18th and 19th centuries, where it was developed in order to analyse many-body problems. As we know from classical mechanics, a two-body problem is exactly solvable. However, once a third mass is introduced, Newton's equations of motion do not, in general, have a closed-form solution. By assuming that the mass of the third body is small, however, we can find some intriguing results. One such result is the existence of the Lagrangian points. Indeed, in the system of the Sun and Jupiter, a large number of asteroids, of much smaller mass, are found at the stable Lagrangian points.

In quantum mechanics, perturbation theory goes back to the early 20th century, when Paul Dirac developed it in order to formulate quantum electrodynamics. Over the course of the last century it was further developed by many people, such as Werner Heisenberg, Enrico Fermi and Richard Feynman. Nowadays, its applications are numerous, including nuclear and particle physics, quantum field theory and condensed matter physics. In this chapter, we will look at the fundamentals of perturbation theory and study some of its applications.

14.1 Perturbation Theory for Time-Independent Potentials with No Degeneracy

The physics of a quantum mechanical system is governed by the Schrödinger equation, which is a linear partial differential equation. In Part I we saw that, in simple cases, the Schrödinger equation can be solved exactly by expanding the state in the eigenstates of the Hamiltonian. However, there are infinitely many Hamiltonians for which no exact solution is known. How do we treat them? This is an important question, since most systems we encounter in nature do not have an exactly solvable Hamiltonian. As we cannot solve them exactly, we need a method that approximates the solutions to good accuracy. Perturbation theory is such an approximation: we assume that the Hamiltonian is dominated by an exactly solvable term, to which a small term that we cannot solve exactly is added. The formalism we develop in this chapter is an expansion around the solvable Hamiltonian.

Let us consider a system with Hamiltonian H, eigenvalues E_n and eigenstates $|n\rangle$,

$$H|n\rangle = E_n|n\rangle\,, \tag{14.1}$$

where the Hamiltonian H is the sum of an exactly solvable Hamiltonian H_0 and a small perturbation of order $\lambda \ll 1$:

$$H = H_0 + \lambda V. \tag{14.2}$$

Since H_0 is exactly solvable, for $\lambda = 0$ we can solve the Schrödinger equation and find the energy eigenstates and eigenvalues

$$H_0 \left|n^{(0)}\right\rangle = E_n^{(0)} \left|n^{(0)}\right\rangle. \tag{14.3}$$

In order to describe the complete system we have to determine E_n and $|n\rangle$. We seek them in perturbation theory, writing $|n\rangle$ and E_n as power series in the small parameter λ:

$$\begin{aligned} |n\rangle &= \left|n^{(0)}\right\rangle + \lambda \left|n^{(1)}\right\rangle + \lambda^2 \left|n^{(2)}\right\rangle + \dots \\ E_n &= E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \dots . \end{aligned} \tag{14.4}$$

By substituting Eq. (14.4) into Eq. (14.1), we obtain

$$\begin{aligned} & H|n\rangle = E_n |n\rangle \\ \leftrightarrow\ & (H_0 + \lambda V)\left(\left|n^{(0)}\right\rangle + \lambda \left|n^{(1)}\right\rangle + \lambda^2 \left|n^{(2)}\right\rangle + \dots\right) \\ =\ & \left(E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \dots\right)\left(\left|n^{(0)}\right\rangle + \lambda \left|n^{(1)}\right\rangle + \lambda^2 \left|n^{(2)}\right\rangle + \dots\right). \end{aligned} \tag{14.5}$$

Grouping together terms of the same order in λ, we find

$$\begin{aligned} 0 = & \left[H_0 - E_n^{(0)}\right] \left|n^{(0)}\right\rangle \\ & + \lambda \left[H_0 \left|n^{(1)}\right\rangle + V \left|n^{(0)}\right\rangle - E_n^{(0)} \left|n^{(1)}\right\rangle - E_n^{(1)} \left|n^{(0)}\right\rangle\right] \\ & + \lambda^2 \left[H_0 \left|n^{(2)}\right\rangle + V \left|n^{(1)}\right\rangle - E_n^{(0)} \left|n^{(2)}\right\rangle - E_n^{(1)} \left|n^{(1)}\right\rangle - E_n^{(2)} \left|n^{(0)}\right\rangle\right] \\ & + O\left(\lambda^3\right). \end{aligned} \tag{14.6}$$

Observe the structure of the coefficients in the perturbative series in Eq. (14.6). At each order m, the coefficient of the λ^m term contains pairs $\left(E_n^{(i)}, \left|n^{(i)}\right\rangle\right)$ which have already appeared at lower orders. Only one new pair, $\left(E_n^{(m)}, \left|n^{(m)}\right\rangle\right)$, appears for the first time at order m. Thus we can determine the sequence of $\left(E_n^{(i)}, \left|n^{(i)}\right\rangle\right)$ iteratively, demanding successively that each coefficient of λ^i is separately zero.

14.1.1 Zeroth- and First-Order Perturbative Corrections

At leading order, we have

$$O\left(\lambda^0\right): \qquad \left(H_0 - E_n^{(0)}\right) \left|n^{(0)}\right\rangle = 0, \tag{14.7}$$

which is the Schrödinger equation for the unperturbed system. In order to apply perturbation theory, we must first have solved the unperturbed system. At first order we have

$$O\left(\lambda^1\right): \qquad H_0 \left|n^{(1)}\right\rangle + V \left|n^{(0)}\right\rangle - E_n^{(0)} \left|n^{(1)}\right\rangle - E_n^{(1)} \left|n^{(0)}\right\rangle = 0. \tag{14.8}$$

We multiply Eq. (14.8) by the bra $\langle m^{(0)}|$ from the left to find

$$\begin{aligned} 0 = &\langle m^{(0)}| H_0 |n^{(1)}\rangle + \langle m^{(0)}| V |n^{(0)}\rangle \\ &- \langle m^{(0)}| E_n^{(0)} |n^{(1)}\rangle - \langle m^{(0)}| E_n^{(1)} |n^{(0)}\rangle . \end{aligned} \tag{14.9}$$

We use normalised and mutually orthogonal eigenvectors of the unperturbed Hamiltonian H_0:

$$\langle m^{(0)}|n^{(0)}\rangle = \delta_{nm}. \tag{14.10}$$

With this property Eq. (14.9) simplifies to

$$\boxed{E_n^{(1)}\delta_{nm} = \langle m^{(0)}| V |n^{(0)}\rangle + \left(E_m^{(0)} - E_n^{(0)}\right)\langle m^{(0)}|n^{(1)}\rangle} . \tag{14.11}$$

This is the basic equation of first-order perturbation theory.

For $n = m$, we obtain the first-order correction to the energy levels:

$$\boxed{E_n^{(1)} = \langle n^{(0)}| V |n^{(0)}\rangle} . \tag{14.12}$$

This means that the first perturbative correction, $\lambda E_n^{(1)}$, to the energy level $E_n^{(0)}$ is given by the expectation value of the perturbation λV in the unperturbed state.

We can make further use of Eq. (14.11). We focus on the case in which the unperturbed Hamiltonian H_0 has no degeneracies in the energy spectrum:

$$E_n \neq E_m, \qquad \text{if} \quad n \neq m. \tag{14.13}$$

Then we can compute the overlap $\langle m^{(0)}|n^{(1)}\rangle$. For $n \neq m$, Eq. (14.11) implies

$$\langle m^{(0)}|n^{(1)}\rangle = \frac{\langle m^{(0)}| V |n^{(0)}\rangle}{E_n^{(0)} - E_m^{(0)}}. \tag{14.14}$$

This yields the perturbative correction $|n^{(1)}\rangle$ to the energy eigenstates. Inserting a complete set of unperturbed states, we have

$$\begin{aligned} |n^{(1)}\rangle = \mathbb{1}\,|n^{(1)}\rangle &= \left(\sum_m |m^{(0)}\rangle\langle m^{(0)}|\right)|n^{(1)}\rangle \\ \Longrightarrow\ |n^{(1)}\rangle &= |n^{(0)}\rangle\langle n^{(0)}|n^{(1)}\rangle + \sum_{m\neq n} |m^{(0)}\rangle\langle m^{(0)}|n^{(1)}\rangle . \end{aligned} \tag{14.15}$$

Finally we find

$$|n^{(1)}\rangle = |n^{(0)}\rangle\langle n^{(0)}|n^{(1)}\rangle + \sum_{m\neq n} |m^{(0)}\rangle \frac{\langle m^{(0)}| V |n^{(0)}\rangle}{E_n^{(0)} - E_m^{(0)}}. \tag{14.16}$$

In Eq. (14.16) we have expressed $|n^{(1)}\rangle$ in terms of the unperturbed kets $|m^{(0)}\rangle$. However, Eq. (14.16) still contains $\langle n^{(0)}|n^{(1)}\rangle$, which we have not yet determined. It is fixed by

imposing a normalisation condition on the state $|n\rangle$. We demand

$$\begin{aligned} 1 = \langle n|n\rangle = & \left(\langle n^{(0)}| + \lambda \langle n^{(1)}| + \lambda^2 \langle n^{(2)}| + \dots\right) \\ & \cdot \left(|n^{(0)}\rangle + \lambda |n^{(1)}\rangle + \lambda^2 |n^{(2)}\rangle + \dots\right) \\ = & \langle n^{(0)}|n^{(0)}\rangle + \lambda \left[\langle n^{(0)}|n^{(1)}\rangle + \langle n^{(1)}|n^{(0)}\rangle\right] + O\left(\lambda^2\right) \\ = & 1 + \lambda \left[\langle n^{(0)}|n^{(1)}\rangle + \langle n^{(1)}|n^{(0)}\rangle\right] + O\left(\lambda^2\right). \end{aligned} \tag{14.17}$$

This implies

$$\lambda \left[\langle n^{(0)}|n^{(1)}\rangle + \langle n^{(1)}|n^{(0)}\rangle\right] + O\left(\lambda^2\right) = 0. \tag{14.18}$$

The state is correctly normalised at this order if we impose the condition

$$\boxed{\langle n^{(0)}|n^{(1)}\rangle = 0.} \tag{14.19}$$

Combining our results we have

$$\boxed{|n^{(1)}\rangle = \sum_{m\neq n} |m^{(0)}\rangle \frac{\langle m^{(0)}| V |n^{(0)}\rangle}{E_n^{(0)} - E_m^{(0)}}}. \tag{14.20}$$

In summary, at first order in perturbation theory we have

$$\begin{aligned} |n\rangle = & \ |n^{(0)}\rangle + \lambda |n^{(1)}\rangle + O\left(\lambda^2\right) \\ E_n = & \ E_n^{(0)} + \lambda E_n^{(1)} + O\left(\lambda^2\right) \end{aligned} \tag{14.21}$$

with

$$E_n^{(1)} = \langle n^{(0)}| V |n^{(0)}\rangle \tag{14.22}$$

and

$$|n^{(1)}\rangle = \sum_{m\neq n} |m^{(0)}\rangle \frac{\langle m^{(0)}| V |n^{(0)}\rangle}{E_n^{(0)} - E_m^{(0)}}, \tag{14.23}$$

for the Hamiltonian system

$$(H_0 + \lambda V)|n\rangle = E_n |n\rangle. \tag{14.24}$$

It is convenient to generalise Eq. (14.19) to

$$\begin{aligned} & \langle n^{(0)}|n\rangle = 1 \\ \Longrightarrow \ & \langle n^{(0)}|n^{(1)}\rangle = \langle n^{(0)}|n^{(2)}\rangle = \dots = \langle n^{(0)}|n^{(m)}\rangle = 0. \end{aligned} \tag{14.25}$$

As we shall see, this can be made compatible with $\langle n|n\rangle = 1$ at higher orders in λ by a multiplicative correction factor.

14.1.2 Second-Order Perturbative Correction

It is not hard to extend our analysis in order to compute the corrections of $O\left(\lambda^2\right)$ to the energy and the eigenstate. The terms of order λ^2 in Eq. (14.6) give

$$\begin{aligned} O\left(\lambda^2\right): \qquad 0 = H_0 \left|n^{(2)}\right\rangle + V \left|n^{(1)}\right\rangle - E_n^{(2)} \left|n^{(0)}\right\rangle \\ - E_n^{(1)} \left|n^{(1)}\right\rangle - E_n^{(0)} \left|n^{(2)}\right\rangle . \end{aligned} \tag{14.26}$$

Rearranging, we have

$$E_n^{(2)} \left|n^{(0)}\right\rangle = \left(H_0 - E_n^{(0)}\right) \left|n^{(2)}\right\rangle + V \left|n^{(1)}\right\rangle - E_n^{(1)} \left|n^{(1)}\right\rangle . \tag{14.27}$$

As for the first-order correction, we multiply Eq. (14.27) by the bra $\langle k^{(0)}|$:

$$E_n^{(2)} \delta_{kn} = \left(E_k^{(0)} - E_n^{(0)}\right) \langle k^{(0)}|n^{(2)}\rangle + \langle k^{(0)}| V \left|n^{(1)}\right\rangle - E_n^{(1)} \langle k^{(0)}|n^{(1)}\rangle . \tag{14.28}$$

For $k = n$, Eq. (14.28) gives

$$\begin{aligned} E_n^{(2)} &= \langle n^{(0)}| V \left|n^{(1)}\right\rangle - \cancel{E_n^{(1)} \langle n^{(0)}|n^{(1)}\rangle} \\ &= \langle n^{(0)}|V| \sum_{m\neq n} |m^{(0)}\rangle \langle m^{(0)}|n^{(1)}\rangle + \cancel{\langle n^{(0)}|V|n^{(0)}\rangle \langle n^{(0)}|n^{(1)}\rangle} \\ &= \langle n^{(0)}|V| \sum_{m\neq n} |m^{(0)}\rangle \langle m^{(0)}|n^{(1)}\rangle , \end{aligned} \tag{14.29}$$

where the cancelled terms vanish because of the condition $\langle n^{(0)}|n^{(1)}\rangle = 0$ of Eq. (14.19). Using the result for $|n^{(1)}\rangle$ in Eq. (14.20), we find

$$E_n^{(2)} = \sum_{m\neq n} \langle n^{(0)}|V|m^{(0)}\rangle \frac{\langle m^{(0)}|V|n^{(0)}\rangle}{E_n^{(0)} - E_m^{(0)}}, \tag{14.30}$$

and thus

$$\boxed{E_n^{(2)} = \sum_{m\neq n} \frac{\left|\langle m^{(0)}|V|n^{(0)}\rangle\right|^2}{E_n^{(0)} - E_m^{(0)}}}, \tag{14.31}$$

which is the second-order correction to the energy.

For the state correction $|n^{(2)}\rangle$ we have

$$\begin{aligned} |n^{(2)}\rangle &= \sum_{k\neq n} |k^{(0)}\rangle \langle k^{(0)}|n^{(2)}\rangle + \cancel{|n^{(0)}\rangle \langle n^{(0)}|n^{(2)}\rangle} \\ &= \sum_{k\neq n} |k^{(0)}\rangle \langle k^{(0)}|n^{(2)}\rangle , \end{aligned} \tag{14.32}$$

where again the cancelled term vanishes because of the condition (14.25). We can compute $\langle k^{(0)}|n^{(2)}\rangle$ from Eq. (14.28).

Let us now revisit our normalisation condition $\langle n^{(0)}|n\rangle = 1$ of Eq. (14.25). We would like it to be compatible with

$$\langle n|n\rangle = 1, \tag{14.33}$$

through the order in λ at which we are computing. However,

$$\begin{aligned} 1 &= \langle n|n\rangle \\ &= \left(\langle n^{(0)}| + \lambda \langle n^{(1)}| + \lambda^2 \langle n^{(2)}| + \dots\right) \cdot \left(|n^{(0)}\rangle + \lambda |n^{(1)}\rangle + \lambda^2 |n^{(2)}\rangle + \dots\right) \\ &= 1 + \lambda^2 \langle n^{(1)}|n^{(1)}\rangle + \dots, \end{aligned} \tag{14.34}$$

where we have imposed the condition $\langle n^{(0)}|n^{(m)}\rangle = 0$ of Eq. (14.25). As expected, our naive normalisation condition is not sufficient beyond first order in perturbation theory. We therefore define

$$|n\rangle = Z_\lambda^{-1}\left(|n^{(0)}\rangle + \lambda |n^{(1)}\rangle + \lambda^2 |n^{(2)}\rangle + \dots\right). \tag{14.35}$$

For this renormalised state, Eq. (14.34) becomes

$$\begin{aligned} 1 &= Z_\lambda^{-2}\left(1 + \lambda^2 \langle n^{(1)}|n^{(1)}\rangle + \dots\right) \\ \implies Z_\lambda &= \sqrt{1 + \lambda^2 \langle n^{(1)}|n^{(1)}\rangle} + O\left(\lambda^3\right) \\ \implies Z_\lambda &= 1 + \frac{\lambda^2}{2}\langle n^{(1)}|n^{(1)}\rangle + O\left(\lambda^3\right), \end{aligned} \tag{14.36}$$

where in the final step we used the Taylor expansion of the square root.

Having derived the perturbative corrections to the eigenstates and eigenvalues, let us now look at an example of how perturbation theory is applied.

14.1.3 Perturbed Harmonic Oscillator

Let us consider a perturbation of a harmonic oscillator with mass m and angular frequency ω,

$$H = \underbrace{\frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2}_{H_0} + \underbrace{\lambda V(x)}_{\text{perturbation}}. \tag{14.37}$$

Recall the algebraic solution of the unperturbed system. We define the creation and annihilation operators

$$\begin{aligned} a^\dagger &= \sqrt{\frac{m\omega}{2\hbar}}\left(x - \frac{ip}{m\omega}\right), \\ a &= \sqrt{\frac{m\omega}{2\hbar}}\left(x + \frac{ip}{m\omega}\right), \end{aligned} \tag{14.38}$$

with the commutation relation and the number operator

$$\left[a, a^\dagger\right] = 1, \qquad N = a^\dagger a. \tag{14.39}$$

We can therefore rewrite the Hamiltonian as

$$H_0 = \hbar\omega\left(N + \frac{1}{2}\right). \tag{14.40}$$

The eigenstates of the Hamiltonian are the eigenstates of the number operator,

$$N|n\rangle = n|n\rangle. \tag{14.41}$$

The creation and annihilation operators map an eigenstate to the eigenstate of the next higher or lower energy level:

$$a^\dagger |n\rangle = \sqrt{n+1}\,|n+1\rangle\,, \qquad a\,|n\rangle = \sqrt{n}\,|n-1\rangle\,. \tag{14.42}$$

From Eq. (14.40) we see that the energy eigenvalues are $E_n = \left(n+\frac{1}{2}\right)\hbar\omega$. Thus the zeroth-order eigenstates and energy eigenvalues satisfy

$$E_n^{(0)} = \left(n+\frac{1}{2}\right)\hbar\omega, \quad N\,|n^{(0)}\rangle = n^{(0)}\,|n^{(0)}\rangle\,. \tag{14.43}$$

We now come back to the perturbation. As an example, we take

$$\lambda V(x) = \lambda x, \tag{14.44}$$

and we express the position operator in terms of the creation and annihilation operators,

$$x = \sqrt{\frac{\hbar}{2m\omega}}\left(a^\dagger + a\right). \tag{14.45}$$

We require the matrix elements

$$\begin{aligned}\langle m|V|n\rangle &= \langle m|x|n\rangle \\ &= \sqrt{\frac{\hbar}{2m\omega}}\,\langle m|a+a^\dagger|n\rangle \\ &= \sqrt{\frac{\hbar}{2m\omega}}\{\langle m|n-1\rangle\sqrt{n} + \langle m|n+1\rangle\sqrt{n+1}\}.\end{aligned} \tag{14.46}$$

At first order in perturbation theory we have

$$\begin{aligned}E_n^{(1)} &= \langle n|V|n\rangle \\ &= \sqrt{\frac{\hbar}{2m\omega}}\left(\sqrt{n}\,\delta_{n-1,n} + \sqrt{n+1}\,\delta_{n+1,n}\right) = 0.\end{aligned} \tag{14.47}$$

Note that $\delta_{n-1,n} = \delta_{n,n+1} = 0$ for all n. There is no first-order correction to the energy levels. Let us continue to the next order:

$$\begin{aligned}E_n^{(2)} &= \sum_{m\neq n}\frac{|\langle m|V|n\rangle|^2}{E_n - E_m} \\ &= \frac{|\langle n-1|V|n\rangle|^2}{E_n - E_{n-1}} + \frac{|\langle n+1|V|n\rangle|^2}{E_n - E_{n+1}} \\ &= \left(\sqrt{\frac{\hbar}{2m\omega}}\right)^2 \cdot \left\{\frac{n}{\left[\left(n+\frac{1}{2}\right) - \left((n-1)+\frac{1}{2}\right)\right]\hbar\omega}\right. \\ &\quad \left. + \frac{n+1}{\left[\left(n+\frac{1}{2}\right) - \left((n+1)+\frac{1}{2}\right)\right]\hbar\omega}\right\} = -\frac{1}{2m\omega^2}.\end{aligned} \tag{14.48}$$

In the second step, we used the fact that

$$V\,|n\rangle = x\,|n\rangle = \sqrt{\frac{\hbar}{2m\omega}}\left(\sqrt{n}\,|n-1\rangle + \sqrt{n+1}\,|n+1\rangle\right).$$

Thus most terms drop out because of the orthogonality of the eigenstates ($\langle m|n\rangle = \delta_{mn}$). Therefore,

$$\begin{aligned} E_n &= E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \dots \\ &= \hbar\omega\left(n + \frac{1}{2}\right) - \frac{\lambda^2}{2m\omega^2} + \mathcal{O}\left(\lambda^3\right). \end{aligned} \tag{14.49}$$

This example serves only as an illustration, since the problem can be solved exactly, without resorting to perturbation theory:

$$\begin{aligned} H &= \frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2 + \lambda x \\ &= \frac{p^2}{2m} + \frac{1}{2}m\omega^2\left(x + \frac{\lambda}{m\omega^2}\right)^2 - \frac{\lambda^2}{2m\omega^2} \\ &= \frac{p^2}{2m} + \frac{1}{2}m\omega^2 y^2 \underbrace{- \frac{\lambda^2}{2m\omega^2}}_{\substack{\text{shift of} \\ \text{energy levels}}}, \end{aligned} \tag{14.50}$$

with $y = x + \lambda/(m\omega^2)$. The shift $-\lambda^2/(2m\omega^2)$ of all energy levels agrees with the perturbative result (14.49), which is in fact exact in this case.

14.1.4 The Variational Method (An Alternative)

Before we study the intricacies of perturbation theory, it is useful to learn about an alternative method. So far we have assumed that we can identify an exactly solvable term H_0 which dominates the Hamiltonian. Often, however, no such solvable dominant term exists. In this situation, we can still constrain the spectrum by putting an upper bound on the ground-state energy of the Hamiltonian.

We consider a Hamiltonian H with eigenstates $|k\rangle$,

$$H|k\rangle = E_k|k\rangle. \tag{14.51}$$

Let us consider a state $|\psi\rangle$ which is a linear combination of the $|k\rangle$:

$$|\psi\rangle = \sum_k c_k|k\rangle. \tag{14.52}$$

In the equation above, we know neither the kets $|k\rangle$ nor the coefficients c_k; we only require that such an expansion exists. Now consider

$$\begin{aligned} \langle\psi|H|\psi\rangle &= \langle\psi|H\left(\sum_k |k\rangle\langle k|\right)|\psi\rangle \\ &= \sum_k \langle\psi|k\rangle\langle k|\psi\rangle E_k \\ &= \sum_k E_k\left|\langle\psi|k\rangle\right|^2 \\ &= \sum_k (E_k - E_0)\left|\langle\psi|k\rangle\right|^2 + \sum_k E_0\left|\langle\psi|k\rangle\right|^2. \end{aligned} \tag{14.53}$$

Given that E_0 is the ground-state energy, we have

$$E_k - E_0 \geq 0, \tag{14.54}$$

and thus

$$\begin{aligned} \langle\psi|H|\psi\rangle &\geq E_0 \sum_k \left|\langle\psi|k\rangle\right|^2 \\ \implies \langle\psi|H|\psi\rangle &\geq E_0 \langle\psi|\psi\rangle . \end{aligned} \tag{14.55}$$

Finally, we have

$$\boxed{E_0 \leq \frac{\langle\psi|H|\psi\rangle}{\langle\psi|\psi\rangle}} . \tag{14.56}$$

The expectation value of the Hamiltonian in any state $|\psi\rangle$ that can be expanded in the eigenstates of the Hamiltonian provides an upper bound on the true ground-state energy. By analysing the potential and its boundary conditions, we can make educated guesses for states $|\psi\rangle$ which closely resemble the ground state. The better our guess, the lower the value of the right-hand side of Eq. (14.56), and the closer it is to the true ground-state energy.

As an example, let us consider the infinite well potential

$$V(x) = \begin{cases} \infty, & |x| > 1 \\ 0, & |x| \leq 1 \end{cases} , \tag{14.57}$$

and the Hamiltonian

$$H = \frac{p^2}{2} + V(x). \tag{14.58}$$

The problem can be solved exactly. The ground state and the ground-state energy are

$$\begin{aligned} \langle x|0\rangle &= \cos\left(\frac{\pi x}{2}\right), \\ \text{and} \qquad E_0 &= \frac{\pi^2}{8}. \end{aligned} \tag{14.59}$$

Let us estimate this result with the ideas we have just explained. We choose a function $\psi(x)$ as an Ansatz, using our imagination and intuition. This function is not, however, completely arbitrary. It must satisfy the boundary conditions of the problem; otherwise it cannot be written as a superposition of energy eigenstates, as our theorem requires. The simplest function which satisfies

$$\psi(-1) = \psi(1) = 0, \tag{14.60}$$

is given by

$$\psi_{\text{trial}}(x) = \left(1 - x^2\right). \tag{14.61}$$

We find

$$\begin{aligned} E_{\text{trial}} &= \frac{\langle\psi_{\text{trial}}|H|\psi_{\text{trial}}\rangle}{\langle\psi_{\text{trial}}|\psi_{\text{trial}}\rangle} \\ &= \frac{5}{4}, \end{aligned} \tag{14.62}$$

and thus

$$\begin{aligned} E_{\text{trial}} &= \frac{5}{4}\frac{1}{E_0}\cdot E_0 = \frac{10}{\pi^2}E_0 \\ &= 1.013212\, E_0, \end{aligned} \tag{14.63}$$

which is really close!

One can improve on such a result in a rather systematic way. Let us consider a new trial function which depends on a parameter λ,

$$\psi(x) = \left(1 - x^2\right)\left(1 + \lambda x^2\right). \tag{14.64}$$

We find

$$\begin{aligned} E_{\text{trial}} &= \frac{3}{4}\frac{11\lambda^2 + 14\lambda + 35}{21 + 6\lambda + \lambda^2} \\ \Longrightarrow\ E_{\text{trial}} &= \frac{6}{\pi^2}\frac{11\lambda^2 + 14\lambda + 35}{21 + 6\lambda + \lambda^2}E_0. \end{aligned} \tag{14.65}$$

We can vary the parameter λ in search of the minimum of E_{trial}. We readily find that E_{trial} has a minimum at

$$\lambda = \lambda_0 = -\frac{49}{13} + \frac{4\sqrt{133}}{13} \tag{14.66}$$

and takes the value

$$E_{\text{trial}} = 1.000015 \cdot E_0, \tag{14.67}$$

which is a truly good approximation. We will see further applications of the variational method in Chapter 18.

14.2 Perturbation Theory for Degenerate Energy Levels

In our derivation of the perturbative corrections we assumed that the unperturbed energy levels are all different. However, systems often have degeneracies, for example due to rotational symmetry.

One example is the Stark effect, i.e. an atom in an electric field. In the hydrogen atom, the $2s$ state and the three $2p$ states all have the same energy at zeroth order. If we turn on a static electric field, two of the four states receive energy corrections, and new spectral lines appear in the emission spectrum. As the four states are degenerate, our previous results cannot be used. Another example will be discussed in Sec. 14.3.2, where the states of the hydrogen atom with the same principal quantum number n split due to relativistic corrections.

We therefore revisit the derivation without assuming that

$$E_n^{(0)} \neq E_m^{(0)}, \qquad \forall n \neq m. \tag{14.68}$$

Once again we consider an exactly solvable Hamiltonian H_0 and a perturbation δH (later we will set $\delta H = \lambda V$, as before). We know the eigenvalues and eigenstates of H_0,

$$H_0 \left|m^{(0)}\right\rangle = E_m^{(0)} \left|m^{(0)}\right\rangle. \tag{14.69}$$

The eigenstates of the complete Hamiltonian satisfy

$$(H_0 + \delta H)\,|n\rangle = \left(E_n^{(0)} + \delta E_n\right)|n\rangle\,. \tag{14.70}$$

If we multiply this equation by the bra $\langle m^{(0)}|$, we obtain

$$\langle m^{(0)}|H_0 + \delta H|n\rangle = \left(E_n^{(0)} + \delta E_n\right)\langle m^{(0)}|n\rangle\,, \tag{14.71}$$

which we can also write as

$$\boxed{\left(E_n^{(0)} - E_m^{(0)} + \delta E_n\right)\langle m^{(0)}|n\rangle = \langle m^{(0)}|\delta H|n\rangle\,.} \tag{14.72}$$

At this point we would normally expand $|n\rangle$ and δE_n in a perturbative series. Since we may now have $E_n^{(0)} = E_m^{(0)}$ for some $n \neq m$, we postpone the expansion. First we write

$$\boxed{\langle m^{(0)}|n\rangle = \frac{\langle m^{(0)}|\delta H|n\rangle}{E_n^{(0)} - E_m^{(0)} + \delta E_n}.} \tag{14.73}$$

We also have

$$\begin{aligned} |n\rangle &= \mathbb{1}\cdot|n\rangle \\ &= \left(\sum_m |m^{(0)}\rangle\langle m^{(0)}|\right)|n\rangle \\ &= \sum_m |m^{(0)}\rangle\langle m^{(0)}|n\rangle\,. \end{aligned} \tag{14.74}$$

Substituting Eq. (14.73) into Eq. (14.74), we obtain

$$\boxed{|n\rangle = \sum_m |m^{(0)}\rangle\,\frac{\langle m^{(0)}|\delta H|n\rangle}{\Delta_{nm} + \delta E_n},} \tag{14.75}$$

where we defined

$$\Delta_{nm} := E_n^{(0)} - E_m^{(0)} = -\Delta_{mn}. \tag{14.76}$$

The sum includes terms with $\Delta_{mn} = 0$, i.e. $E_m^{(0)} = E_n^{(0)}$. These correspond to the states $|m^{(0)}\rangle$ which are degenerate with $|n^{(0)}\rangle$. It also includes terms from states which are not degenerate with $|n^{(0)}\rangle$, with $\Delta_{mn} \neq 0$, i.e. $E_m^{(0)} \neq E_n^{(0)}$. Symbolically, we write

$$\sum_m = \sum_{m\in D_n} + \sum_{m\notin D_n}\,, \tag{14.77}$$

where D_n denotes the set of states $|m^{(0)}\rangle$ which are degenerate with $|n^{(0)}\rangle$. Therefore, we have

$$\begin{aligned} |n\rangle = &\sum_{m\in D_n} |m^{(0)}\rangle\,\frac{\langle m^{(0)}|\delta H|n\rangle}{\delta E_n} \qquad (A) \\ &+ \sum_{m\notin D_n} |m^{(0)}\rangle\,\frac{\langle m^{(0)}|\delta H|n\rangle}{\Delta_{nm} + \delta E_n} \qquad (B). \end{aligned} \tag{14.78}$$

Note that so far we have made no approximation whatsoever. We now perform a perturbative expansion, setting

$$\delta H = \lambda V. \tag{14.79}$$

As in the non-degenerate case (Sec. 14.1), we expand

$$\begin{aligned} |n\rangle &= |n^{(0)}\rangle + \lambda\,|n^{(1)}\rangle + \lambda^2\,|n^{(2)}\rangle + \ldots \\ \delta E_n &= \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \ldots. \end{aligned} \tag{14.80}$$

The first term in Eq. (14.78) becomes

$$\begin{aligned} (A) &= \sum_{m\in D_n} |m^{(0)}\rangle \frac{\langle m^{(0)}|\,\lambda V|\left(|n^{(0)}\rangle + \lambda\,|n^{(1)}\rangle + \ldots\right)}{\lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \ldots} \\ &= \sum_{m\in D_n} |m^{(0)}\rangle \frac{\lambda}{\lambda} \frac{\langle m^{(0)}|\,V|\left(|n^{(0)}\rangle + \lambda\,|n^{(1)}\rangle + \ldots\right)}{E_n^{(1)} + \lambda E_n^{(2)} + \ldots} \\ &= \sum_{m\in D_n} |m^{(0)}\rangle \frac{\langle m^{(0)}|\,V|\left(|n^{(0)}\rangle + \lambda\,|n^{(1)}\rangle + \ldots\right)}{E_n^{(1)}} \\ &\quad \times \frac{1}{1 + \lambda \frac{E_n^{(2)}}{E_n^{(1)}} + \ldots} \\ &= \sum_{m\in D_n} |m^{(0)}\rangle \frac{\langle m^{(0)}|\,V|\left(|n^{(0)}\rangle + \lambda\,|n^{(1)}\rangle + \ldots\right)}{E_n^{(1)}} \\ &\quad \times \left(1 - \lambda \frac{E_n^{(2)}}{E_n^{(1)}} + \ldots\right) \\ &= \sum_{m\in D_n} |m^{(0)}\rangle \left\{ \frac{\langle m^{(0)}|V|n^{(0)}\rangle}{E_n^{(1)}} + \lambda \left[\frac{\langle m^{(0)}|V|n^{(1)}\rangle}{E_n^{(1)}} - \frac{\langle m^{(0)}|V|n^{(0)}\rangle}{E_n^{(1)}} \frac{E_n^{(2)}}{E_n^{(1)}} \right] \right. \\ &\quad \left. + O(\lambda^2) \right\}. \end{aligned} \tag{14.81}$$

Here we used the Taylor expansion

$$\frac{1}{1+x} = 1 - x + x^2 - x^3 + \ldots. \tag{14.82}$$

Notice that, unlike before, we have a term of order $O(\lambda^0)$ due to the degeneracy. For the second term, we obtain

$$\begin{aligned} (B) &= \sum_{m\notin D_n} |m^{(0)}\rangle \frac{\langle m^{(0)}|\,\lambda V|\left(|n^{(0)}\rangle + \lambda\,|n^{(1)}\rangle + \ldots\right)}{\Delta_{nm} + \lambda E_n^{(1)} + \ldots} \\ &= \sum_{m\notin D_n} |m^{(0)}\rangle \frac{\langle m^{(0)}|\,\lambda V|\left(|n^{(0)}\rangle + \lambda\,|n^{(1)}\rangle + \ldots\right)}{\Delta_{nm}} \frac{1}{1 + \lambda \frac{E_n^{(1)}}{\Delta_{nm}} + \ldots} \\ &= \sum_{m\notin D_n} |m^{(0)}\rangle \frac{\langle m^{(0)}|\,\lambda V|\left(|n^{(0)}\rangle + \lambda\,|n^{(1)}\rangle + \ldots\right)}{\Delta_{nm}} \left(1 - \lambda \frac{E_n^{(1)}}{\Delta_{nm}} + \ldots\right) \\ &= \sum_{m\notin D_n} |m^{(0)}\rangle \frac{\langle m^{(0)}|V|n^{(0)}\rangle}{\Delta_{nm}} \lambda + O(\lambda^2). \end{aligned} \tag{14.83}$$

Putting everything together, we have

$$\begin{aligned}
&|n^{(0)}\rangle + \lambda\,|n^{(1)}\rangle + O(\lambda^2) \\
&= \sum_{m\in D_n} |m^{(0)}\rangle \frac{\langle m^{(0)}|V|n^{(0)}\rangle}{E_n^{(1)}} \\
&\quad + \lambda\left\{ \sum_{m\notin D_n} |m^{(0)}\rangle \frac{\langle m^{(0)}|V|n^{(0)}\rangle}{\Delta_{nm}} \right. \\
&\quad \left. + \sum_{m\in D_n} |m^{(0)}\rangle \left[\frac{\langle m^{(0)}|V|n^{(1)}\rangle}{E_n^{(1)}} - \frac{\langle m^{(0)}|V|n^{(0)}\rangle}{E_n^{(1)}} \frac{E_n^{(2)}}{E_n^{(1)}} \right] \right\} \\
&\quad + O(\lambda^2).
\end{aligned} \tag{14.84}$$

As before, we equate the coefficients of each power of λ on the two sides of the equation. At leading order,

$$|n^{(0)}\rangle = \sum_{m\in D_n} |m^{(0)}\rangle \frac{\langle m^{(0)}|V|n^{(0)}\rangle}{E_n^{(1)}}. \tag{14.85}$$

We multiply by $\langle l^{(0)}|$:

$$\langle l^{(0)}|n^{(0)}\rangle = \sum_{m\in D_n} \langle l^{(0)}|m^{(0)}\rangle \frac{\langle m^{(0)}|V|n^{(0)}\rangle}{E_n^{(1)}}. \tag{14.86}$$

If $l \in D_n$, we obtain

$$\delta_{ln} = \sum_{m\in D_n} \delta_{lm} \frac{\langle m^{(0)}|V|n^{(0)}\rangle}{E_n^{(1)}}. \tag{14.87}$$

From this it follows that

$$\boxed{E_n^{(1)}\delta_{ln} = \langle l^{(0)}|V|n^{(0)}\rangle}. \tag{14.88}$$

Eq. (14.88) requires that the matrix $\langle l^{(0)}|V|n^{(0)}\rangle$, with $l,\ n \in D_n$, is diagonal. When H_0 has degenerate states, the basis of eigenstates of H_0 is not unique: any basis of a degenerate subspace consists of eigenstates with the same energy $E_n^{(0)}$.

However, we should keep in mind that here we solve the more complicated Hamiltonian

$$H = H_0 + \delta H. \tag{14.89}$$

The perturbation δH removes this arbitrariness (at least in the non-trivial case $[H_0, \delta H] \neq 0$). Among all possible choices of the degenerate states $|n^{(0)}\rangle$, we need to find the basis in which the perturbation is diagonal within each degenerate subspace, as required by Eq. (14.88).

We now examine the $O(\lambda)$ coefficient in Eq. (14.84):

$$\begin{aligned}
|n^{(1)}\rangle = &-\frac{E_n^{(2)}}{E_n^{(1)}} \sum_{m\in D_n} |m^{(0)}\rangle \frac{\langle m^{(0)}|V|n^{(0)}\rangle}{E_n^{(1)}} \\
&+ \sum_{m\in D_n} |m^{(0)}\rangle \frac{\langle m^{(0)}|V|n^{(1)}\rangle}{E_n^{(1)}} \\
&+ \sum_{m\notin D_n} |m^{(0)}\rangle \frac{\langle m^{(0)}|V|n^{(0)}\rangle}{\Delta_{nm}}.
\end{aligned} \tag{14.90}$$

By inserting Eq. (14.88), we obtain

$$\boxed{\begin{aligned}|n^{(1)}\rangle = &-\frac{E_n^{(2)}}{E_n^{(1)}}|n^{(0)}\rangle \\ &+\sum_{m\in D_n}|m^{(0)}\rangle\frac{\langle m^{(0)}|V|n^{(1)}\rangle}{E_n^{(1)}} \\ &+\sum_{m\notin D_n}|m^{(0)}\rangle\frac{\langle m^{(0)}|V|n^{(0)}\rangle}{\Delta_{nm}}\end{aligned}}. \tag{14.91}$$

We multiply by $\langle l^{(0)}|$:

$$\begin{aligned}\langle l^{(0)}|n^{(1)}\rangle = &-\frac{E_n^{(2)}}{E_n^{(1)}}\langle l^{(0)}|n^{(0)}\rangle \\ &+\sum_{m\in D_n}\langle l^{(0)}|m^{(0)}\rangle\frac{\langle m^{(0)}|V|n^{(1)}\rangle}{E_n^{(1)}} \\ &+\sum_{m\notin D_n}\langle l^{(0)}|m^{(0)}\rangle\frac{\langle m^{(0)}|V|n^{(0)}\rangle}{\Delta_{nm}}.\end{aligned} \tag{14.92}$$

If $l \notin D_n$, using $\langle l^{(0)}|n^{(0)}\rangle = 0$ and $\langle l^{(0)}|m^{(0)}\rangle = \delta_{lm}$, we obtain

$$\boxed{\langle l^{(0)}|n^{(1)}\rangle = \frac{\langle l^{(0)}|V|n^{(0)}\rangle}{\Delta_{nl}}, \quad l \notin D_n}. \tag{14.93}$$

For $l \in D_n$, the derivation is longer:

$$\begin{aligned}\langle l^{(0)}|n^{(1)}\rangle &= -\frac{E_n^{(2)}}{E_n^{(1)}}\langle l^{(0)}|n^{(0)}\rangle + \sum_{m\in D_n}\langle l^{(0)}|m^{(0)}\rangle\frac{\langle m^{(0)}|V|n^{(1)}\rangle}{E_n^{(1)}} \\ &= -\frac{E_n^{(2)}}{E_n^{(1)}}\delta_{ln} + \frac{\langle l^{(0)}|V|n^{(1)}\rangle}{E_n^{(1)}}.\end{aligned} \tag{14.94}$$

Using the fact that $|n^{(1)}\rangle = \sum_k |k^{(0)}\rangle\langle k^{(0)}|n^{(1)}\rangle$, we can rewrite this as

$$\langle l^{(0)}|n^{(1)}\rangle = -\frac{E_n^{(2)}}{E_n^{(1)}}\delta_{ln} + \frac{\langle l^{(0)}|\,V|\sum_k|k^{(0)}\rangle\langle k^{(0)}|n^{(1)}\rangle}{E_n^{(1)}}. \tag{14.95}$$

As before, we split the sum into $k \in D_n$ and $k \notin D_n$:

$$\begin{aligned}\langle l^{(0)}|n^{(1)}\rangle = -\frac{E_n^{(2)}}{E_n^{(1)}}\delta_{ln} &+ \sum_{k\in D_n}\frac{\langle l^{(0)}|\,V\,|k^{(0)}\rangle}{E_n^{(1)}}\langle k^{(0)}|n^{(1)}\rangle \\ &+ \sum_{k\notin D_n}\frac{\langle l^{(0)}|\,V\,|k^{(0)}\rangle}{E_n^{(1)}}\langle k^{(0)}|n^{(1)}\rangle\,.\end{aligned} \tag{14.96}$$

Using Eq. (14.88), we can rewrite the first sum:

$$\langle l^{(0)}|n^{(1)}\rangle = -\frac{E_n^{(2)}}{E_n^{(1)}}\delta_{ln} + \sum_{k\in D_n}\delta_{kl}\frac{E_l^{(1)}}{E_n^{(1)}}\langle k^{(0)}|n^{(1)}\rangle + \sum_{k\notin D_n}\frac{\langle l^{(0)}|\,V\,|k^{(0)}\rangle}{E_n^{(1)}}\langle k^{(0)}|n^{(1)}\rangle\,. \tag{14.97}$$

Using Eq. (14.93), we can rewrite the second sum:

$$\langle l^{(0)}|n^{(1)}\rangle = -\frac{E_n^{(2)}}{E_n^{(1)}}\delta_{ln} + \sum_{k\in D_n}\delta_{kl}\frac{E_l^{(1)}}{E_n^{(1)}}\langle k^{(0)}|n^{(1)}\rangle + \sum_{k\notin D_n}\frac{\langle l^{(0)}|\,V\,|k^{(0)}\rangle}{E_n^{(1)}}\frac{\langle k^{(0)}|V|n^{(0)}\rangle}{\Delta_{nk}}. \tag{14.98}$$

The first sum simplifies, and we find

$$\boxed{\langle l^{(0)}|n^{(1)}\rangle = -\frac{E_n^{(2)}}{E_n^{(1)}}\delta_{ln} + \frac{E_l^{(1)}}{E_n^{(1)}}\langle l^{(0)}|n^{(1)}\rangle + \sum_{k\notin D_n}\frac{\langle l^{(0)}|\,V\,|k^{(0)}\rangle}{E_n^{(1)}}\frac{\langle k^{(0)}|V|n^{(0)}\rangle}{\Delta_{nk}}, \quad l\in D_n} \tag{14.99}$$

Rearranging the equation, we obtain

$$\boxed{E_n^{(2)}\delta_{nl} = \langle l^{(0)}|n^{(1)}\rangle\left(E_l^{(1)} - E_n^{(1)}\right) + \sum_{k\notin D_n}\frac{\langle l^{(0)}|\,V\,|k^{(0)}\rangle\langle k^{(0)}|V|n^{(0)}\rangle}{\Delta_{nk}}.} \tag{14.100}$$

For $l = n$, we obtain

$$\boxed{E_n^{(2)} = \sum_{k\notin D_n}\frac{\left|\langle n^{(0)}|\,V\,|k^{(0)}\rangle\right|^2}{\Delta_{nk}}.} \tag{14.101}$$

14.2.1 Summary of Degenerate Perturbation Theory

In the case of degeneracy, we need to diagonalise the perturbation within each degenerate subspace,

$$E_n^{(1)}\delta_{ln} = \langle l^{(0)}|V|n^{(0)}\rangle\,. \tag{14.102}$$

Once we have found the states $\{|n^{(0)}\rangle\}$ which satisfy this condition, we can construct a perturbative expansion which is very similar to the non-degenerate case. The first- and second-order corrections to the energy levels are given by

$$\boxed{E_n^{(1)}\delta_{ln} = \langle l^{(0)}|V|n^{(0)}\rangle} \tag{14.103}$$

and

$$\boxed{E_n^{(2)}\delta_{nl} = \langle l^{(0)}|n^{(1)}\rangle \left(E_l^{(1)} - E_n^{(1)}\right) + \sum_{k\notin D_n} \frac{\langle l^{(0)}|\, V\, |k^{(0)}\rangle \langle k^{(0)}|V|n^{(0)}\rangle}{\Delta_{nk}},} \tag{14.104}$$

where the degenerate states are excluded from the sum.

What is $|n^{(1)}\rangle$? We decompose

$$\begin{aligned} |n^{(1)}\rangle &= \sum_{k\in D_n} |k^{(0)}\rangle \langle k^{(0)}|n^{(1)}\rangle + \sum_{k\notin D_n} |k^{(0)}\rangle \langle k^{(0)}|n^{(1)}\rangle \\ &= |n^{(0)}\rangle \langle n^{(0)}|n^{(1)}\rangle + \sum_{\substack{k\in D_n \\ k\neq n}} |k^{(0)}\rangle \langle k^{(0)}|n^{(1)}\rangle + \sum_{k\notin D_n} |k^{(0)}\rangle \langle k^{(0)}|n^{(1)}\rangle\,. \end{aligned} \tag{14.105}$$

The three terms can be calculated as follows. From Eq. (14.93) we know

$$\langle k^{(0)}|n^{(1)}\rangle = \frac{\langle k^{(0)}|V|n^{(0)}\rangle}{\Delta_{nk}}, \quad k \notin D_n. \tag{14.106}$$

For $k \in D_n$, $k \neq n$, rearranging Eq. (14.100) gives

$$\langle k^{(0)}|n^{(1)}\rangle = \frac{1}{E_n^{(1)} - E_k^{(1)}} \sum_{m\notin D_n} \frac{\langle k^{(0)}|\, V\, |m^{(0)}\rangle \langle m^{(0)}|V|n^{(0)}\rangle}{\Delta_{nm}}. \tag{14.107}$$

Remark: Eq. (14.107) requires $E_k^{(1)} \neq E_n^{(1)}$, i.e. that the first-order perturbation lifts the degeneracy completely. If some $E_k^{(1)} = E_n^{(1)}$ with $k \neq n$, the degeneracy persists at first order and Eq. (14.100) instead becomes a condition on the choice of basis within that sub-subspace (one has to diagonalise the second-order matrix $\sum_{m\notin D_n} V_{km}V_{mn}/\Delta_{nm}$); see, e.g., Sakurai, *Modern Quantum Mechanics*, Sec. 5.2.

Finally, from the normalisation condition, we know

$$\langle n^{(0)}|n^{(1)}\rangle = \langle n^{(0)}|n^{(2)}\rangle = \ldots = 0. \tag{14.108}$$

14.2.2 Example of Degenerate Perturbation Theory

Consider the Hamiltonian

$$\hat{H} = \begin{pmatrix} |1\rangle & |2\rangle & |3\rangle \end{pmatrix} \underbrace{\begin{pmatrix} E & \lambda & 0 \\ \lambda & E & 0 \\ 0 & 0 & E' \end{pmatrix}}_{H} \begin{pmatrix} \langle 1| \\ \langle 2| \\ \langle 3| \end{pmatrix}. \tag{14.109}$$

Before we apply perturbation theory, we solve this example exactly:

$$\begin{aligned} \det(H - \omega\mathbb{1}) &= \det \begin{pmatrix} E-\omega & \lambda & 0 \\ \lambda & E-\omega & 0 \\ 0 & 0 & E'-\omega \end{pmatrix} \\ &= (E'-\omega)\left[(E-\omega)^2 - \lambda^2\right] = 0. \end{aligned} \tag{14.110}$$

The three energy levels are

$$\omega = E \pm \lambda,\ E'. \tag{14.111}$$

Let us now apply, just for the fun of it, perturbation theory. We separate

$$\hat{H} = \hat{H}_0 + \lambda \hat{V}, \tag{14.112}$$

with

$$H_0 = H|_{\lambda=0} = \begin{pmatrix} E & 0 & 0 \\ 0 & E & 0 \\ 0 & 0 & E' \end{pmatrix} \tag{14.113}$$

and

$$V = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}. \tag{14.114}$$

The states $|1\rangle$ and $|2\rangle$ are degenerate:

$$\begin{aligned} \hat{H}_0 |1\rangle &= E |1\rangle \\ \hat{H}_0 |2\rangle &= E |2\rangle \\ \hat{H}_0 |3\rangle &= E' |3\rangle . \end{aligned} \tag{14.115}$$

Therefore, we must apply degenerate perturbation theory. We seek the states $|\tilde{1}\rangle$, $|\tilde{2}\rangle$, $|\tilde{3}\rangle$ which diagonalise the full Hamiltonian,

$$\begin{aligned} \left(\hat{H}_0 + \lambda \hat{V}\right) |\tilde{1}\rangle &= \tilde{E}_1 |\tilde{1}\rangle \\ \left(\hat{H}_0 + \lambda \hat{V}\right) |\tilde{2}\rangle &= \tilde{E}_2 |\tilde{2}\rangle \\ \left(\hat{H}_0 + \lambda \hat{V}\right) |\tilde{3}\rangle &= \tilde{E}_3 |\tilde{3}\rangle . \end{aligned} \tag{14.116}$$

We write the expansions

$$\begin{aligned} |\tilde{1}\rangle &= |1^{(0)}\rangle + \lambda |1^{(1)}\rangle + \lambda^2 |1^{(2)}\rangle + ... \\ |\tilde{2}\rangle &= |2^{(0)}\rangle + \lambda |2^{(1)}\rangle + \lambda^2 |2^{(2)}\rangle + ... \\ |\tilde{3}\rangle &= |3^{(0)}\rangle + \lambda |3^{(1)}\rangle + \lambda^2 |3^{(2)}\rangle + ... \end{aligned} \tag{14.117}$$

and

$$\begin{aligned} \tilde{E}_1 &= E + \lambda E_1^{(1)} + \lambda^2 E_1^{(2)} + ... \\ \tilde{E}_2 &= E + \lambda E_2^{(1)} + \lambda^2 E_2^{(2)} + ... \\ \tilde{E}_3 &= E' + \lambda E_3^{(1)} + \lambda^2 E_3^{(2)} + ... \end{aligned} \tag{14.118}$$

Since $|1\rangle$ and $|2\rangle$ are degenerate, we **must not** simply identify $|1^{(0)}\rangle$ and $|2^{(0)}\rangle$ with them. Only for the non-degenerate state do we know that

$$|3^{(0)}\rangle = |3\rangle . \tag{14.119}$$

For the other two states we must diagonalise the matrix

$$V_{nm} = \begin{pmatrix} \langle 1|\hat{V}|1\rangle & \langle 1|\hat{V}|2\rangle \\ \langle 2|\hat{V}|1\rangle & \langle 2|\hat{V}|2\rangle \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{14.120}$$

We calculate its eigenvalues:

$$\begin{aligned} \det\left(V_{mn} - \omega\delta_{mn}\right) = \det\begin{pmatrix} -\omega & 1 \\ 1 & -\omega \end{pmatrix} = \omega^2 - 1 \\ \implies \omega = \pm 1. \end{aligned} \tag{14.121}$$

For the eigenvectors we have

$$\begin{aligned} \begin{pmatrix} -\omega & 1 \\ 1 & -\omega \end{pmatrix}\begin{pmatrix} \alpha \\ \beta \end{pmatrix} = 0 \\ \implies \begin{pmatrix} \beta - \alpha\omega \\ \alpha - \beta\omega \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \end{aligned} \tag{14.122}$$

For the two cases we have

$$\begin{aligned} \omega = 1 \quad &\implies \quad \alpha = \beta = \frac{1}{\sqrt{2}} \\ \omega = -1 \quad &\implies \quad \alpha = -\beta = \frac{1}{\sqrt{2}}. \end{aligned} \tag{14.123}$$

Therefore,

$$\begin{aligned} |1^{(0)}\rangle &= \frac{|1\rangle - |2\rangle}{\sqrt{2}} \\ |2^{(0)}\rangle &= \frac{|1\rangle + |2\rangle}{\sqrt{2}} \\ |3^{(0)}\rangle &= |3\rangle\,. \end{aligned} \tag{14.124}$$

Now that we have diagonalised the perturbation in the degenerate subspace, we can proceed with the formulae of non-degenerate perturbation theory (omitting degenerate states from the sums):

$$E_3^{(1)} = \langle 3|V|3\rangle = 0. \tag{14.125}$$

We have already found the first-order corrections $E_1^{(1)}$ and $E_2^{(1)}$ by diagonalising V_{nm}. They are

$$\begin{aligned} E_1^{(1)} &= -1 \\ E_2^{(1)} &= +1. \end{aligned} \tag{14.126}$$

Explicitly,

$$\begin{aligned} E_1^{(1)} &= \langle 1^{(0)}|\hat{V}|1^{(0)}\rangle \\ &= \langle 1^{(0)}|\begin{pmatrix} |1\rangle & |2\rangle & |3\rangle \end{pmatrix}\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} \langle 1| \\ \langle 2| \\ \langle 3| \end{pmatrix}|1^{(0)}\rangle \\ &= \langle 1^{(0)}|\begin{pmatrix} |1^{(0)}\rangle & |2^{(0)}\rangle & |3\rangle \end{pmatrix}\begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} \langle 1^{(0)}| \\ \langle 2^{(0)}| \\ \langle 3| \end{pmatrix}|1^{(0)}\rangle \\ &= \begin{pmatrix} 1 & 0 & 0 \end{pmatrix}\begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = -1. \end{aligned} \tag{14.127}$$

Analogously, we obtain

$$E_2^{(1)} = \langle 2^{(0)}|\hat{V}|2^{(0)}\rangle = \ldots = 1. \tag{14.128}$$

Therefore,

$$\begin{aligned} \tilde{E}_1 &= E - \lambda \quad + O(\lambda^2) \\ \tilde{E}_2 &= E + \lambda \quad + O(\lambda^2) \\ \tilde{E}_3 &= E' \qquad + O(\lambda^2). \end{aligned} \tag{14.129}$$

From the exact result, we anticipate that the $O(\lambda^2)$ terms vanish. Indeed,

$$\begin{aligned} E_1^{(2)} &= \sum_{k \notin D_1} \frac{\left|\langle 1^{(0)}|V|k^{(0)}\rangle\right|^2}{E - E_k} \\ &= \frac{\left|\langle 1^{(0)}|V|3\rangle\right|^2}{E - E'} = 0. \end{aligned} \tag{14.130}$$

Similarly,

$$\begin{aligned} E_2^{(2)} &= \sum_{k \notin D_2} \frac{\left|\langle 2^{(0)}|V|k^{(0)}\rangle\right|^2}{E - E_k} \\ &= \frac{\left|\langle 2^{(0)}|V|3\rangle\right|^2}{E - E'} = 0 \end{aligned} \tag{14.131}$$

and

$$\begin{aligned} E_3^{(2)} &= \sum_{k \notin D_3} \frac{\left|\langle 3|V|k^{(0)}\rangle\right|^2}{E - E_k} \\ &= \frac{\left|\langle 3|V|1^{(0)}\rangle\right|^2}{E' - E} + \frac{\left|\langle 3|V|2^{(0)}\rangle\right|^2}{E' - E} = 0. \end{aligned} \tag{14.132}$$

We have not shown that the terms of $O(\lambda^3)$ and higher vanish. Nevertheless, if we carried perturbation theory to higher orders, we would find that all higher-order energy corrections are zero, in agreement with the exact result.

14.2.3 Example: Stark Effect for the $n = 2$ Energy Levels

We place a hydrogen atom in a constant electric field $\mathcal{E}$ pointing in the z-direction. The perturbation is

$$\delta H = e\mathcal{E}\, z. \tag{14.133}$$

For $n = 2$ there are four degenerate states $|n, l, m\rangle$:

$$n = 2: \quad \begin{cases} l = 0: & m = 0, \\ l = 1: & m = -1,\ 0,\ 1. \end{cases} \tag{14.134}$$

We need the matrix elements

$$e\mathcal{E}\ \langle n', l', m'|\hat{z}|n, l, m\rangle\,. \tag{14.135}$$

Notice that $[\delta H, L_z] = 0$, so that

$$\begin{aligned} 0 &= \langle m'|[\delta H, L_z]\,|m\rangle \\ &= \langle m'|\delta H L_z - L_z \delta H|m\rangle \\ &= (m - m')\,\langle m'|\delta H|m\rangle\,. \end{aligned} \tag{14.136}$$

$$\boxed{\text{For } m' \neq m: \quad \langle m'|\delta H|m\rangle = 0.} \tag{14.137}$$

The matrix elements

$$\langle 2,0,0|\delta H|2,1,0\rangle \quad \text{and} \quad \langle 2,1,0|\delta H|2,0,0\rangle \tag{14.138}$$

are, so far, unconstrained. What about the diagonal elements? We examine $\langle 2,l,m|\delta H|2,l,m\rangle$ under parity; the spherical harmonics have parity $(-1)^l$, while $\delta H \equiv H'$ is odd:

$$\begin{aligned} \langle Y_{lm}|H'|Y_{lm}\rangle &= \langle Y_{lm}|(-1)^l(-1)H'(-1)^l|Y_{lm}\rangle \\ &= (-1)^{2l+1}\,\langle Y_{lm}|H'|Y_{lm}\rangle \\ &= -\langle Y_{lm}|H'|Y_{lm}\rangle = 0. \end{aligned} \tag{14.139}$$

In the 4×4 degenerate subspace (basis $|2,0,0\rangle, |2,1,0\rangle, |2,1,-1\rangle, |2,1,1\rangle$), the perturbation is therefore represented by the matrix

$$\begin{pmatrix} 0 & \delta & 0 & 0 \\ \delta & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \qquad \delta \equiv \langle 2,0,0|\delta H|2,1,0\rangle\,. \tag{14.140}$$

The perturbation partially lifts the degeneracy: the fourfold degenerate level splits into three levels. Indeed, with the method of Sec. 14.2.2: the 2×2 block is the matrix of Eq. (14.120) times δ, so the correct zeroth-order states are $(|2,0,0\rangle \pm |2,1,0\rangle)/\sqrt{2}$ with first-order shifts $\pm\delta$, while $|2,1,\pm1\rangle$ are not shifted at first order (they remain degenerate). With the wave functions

$$\psi_{200} = \frac{1}{\sqrt{32\pi a_0^3}}\left(2 - \frac{r}{a_0}\right)e^{-r/2a_0}, \qquad \psi_{210} = \frac{1}{\sqrt{32\pi a_0^3}}\,\frac{r}{a_0}\,e^{-r/2a_0}\cos\theta,$$

one finds $\langle 2,0,0|z|2,1,0\rangle = -3a_0$, i.e. $\delta = -3e\mathcal{E}a_0$: the shifts $\mp 3e\mathcal{E}a_0$ are linear in the field (linear Stark effect), possible only because of the l-degeneracy of hydrogen; see, e.g., Sakurai, *Modern Quantum Mechanics*, Sec. 5.2.

14.3 Applications of Time-Independent Perturbation Theory

In this section we consider two applications of time-independent perturbation theory. In both examples, symmetries will play an important role.

14.3.1 Example 1: The van der Waals Force

We now look at the van der Waals force between two atoms. We derive it from a Hamiltonian consisting of the Hamiltonians of the two independent atoms and the potential of the interaction between them. At large enough distances, we can treat this interaction potential as a perturbation. We will obtain a negative energy correction which depends on the distance between the two atoms; this gives rise to the van der Waals force. We consider two hydrogen atoms in their ground state at a distance r, much larger than their radii (see Fig. 14.1):

$$|\mathbf{r}| \gg |\mathbf{r}_1|\,,\ |\mathbf{r}_2|\,. \tag{14.141}$$

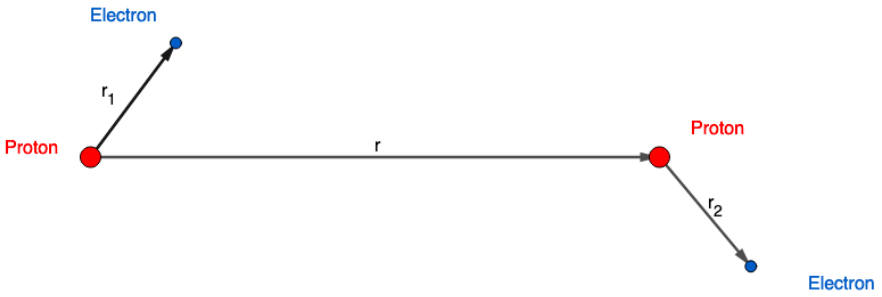


Figure 14.1. Two hydrogen atoms at a distance $\mathbf{r}$; $\mathbf{r}_1$ and $\mathbf{r}_2$ are the positions of the electrons relative to their nuclei.

The Hamiltonian of the system is

$$H = H_0 + V\,, \tag{14.142}$$

with

$$H_0 = -\frac{\hbar^2}{2m}\nabla_1^2 - \frac{\hbar^2}{2m}\nabla_2^2 - \frac{e^2}{|\mathbf{r}_1|} - \frac{e^2}{|\mathbf{r}_2|} \tag{14.143}$$

and

$$\begin{aligned} V = {} & \frac{e^2}{|\mathbf{r}|} && \text{proton-proton interaction} \\ & + \frac{e^2}{|\mathbf{r} - \mathbf{r}_1 + \mathbf{r}_2|} && \text{electron-electron interaction} \\ & - \frac{e^2}{|\mathbf{r} - \mathbf{r}_1|} && \text{electron-proton interaction} \\ & - \frac{e^2}{|\mathbf{r} + \mathbf{r}_2|} && \text{electron-proton interaction.} \end{aligned} \tag{14.144}$$

Here, the electron-proton interaction is the interaction between the proton of one hydrogen atom and the electron of the other.

H_0 is the sum of the Hamiltonians of two isolated hydrogen atoms, and V contains the interactions between the particles of different atoms. We shall treat V as a small perturbation. Our expansion parameter is $1/|\mathbf{r}|$, and before we compute the energy levels

we expand the potential V in $1/|\mathbf{r}|$. We have

$$\begin{aligned}\frac{1}{|\mathbf{r}+\delta\mathbf{r}|} &= \frac{1}{\sqrt{(\mathbf{r}+\delta\mathbf{r})^2}} = \frac{1}{\left(|\mathbf{r}|^2 + 2\mathbf{r}\cdot\delta\mathbf{r} + |\delta\mathbf{r}|^2\right)^{1/2}} \\ &= \frac{1}{|\mathbf{r}|}\cdot\frac{1}{\left(1 + 2\frac{\mathbf{r}}{|\mathbf{r}|}\cdot\frac{\delta\mathbf{r}}{|\mathbf{r}|} + \left(\frac{|\delta\mathbf{r}|}{|\mathbf{r}|}\right)^2\right)^{1/2}} \\ &= \frac{1}{|\mathbf{r}|}\cdot\left(1 - \frac{\mathbf{r}}{|\mathbf{r}|}\cdot\frac{\delta\mathbf{r}}{|\mathbf{r}|} + \frac{1}{2|\mathbf{r}|^2}\cdot\left[3\left(\frac{\mathbf{r}}{|\mathbf{r}|}\cdot\delta\mathbf{r}\right)^2 - |\delta\mathbf{r}|^2\right]\right) + O\left(\frac{1}{|\mathbf{r}|^4}\right) \\ &= \frac{1}{|\mathbf{r}|}\cdot\left(1 - \hat{r}\cdot\frac{\delta\mathbf{r}}{|\mathbf{r}|} + \frac{1}{2|\mathbf{r}|^2}\cdot\left[3(\hat{r}\cdot\delta\mathbf{r})^2 - |\delta\mathbf{r}|^2\right]\right) + O\left(\frac{1}{|\mathbf{r}|^4}\right),\end{aligned} \tag{14.145}$$

where in the last step we used $\hat{r} = \mathbf{r}/|\mathbf{r}|$. Applying this expansion to Eq. (14.144), we find that the terms of order $1/|\mathbf{r}|$ and $1/|\mathbf{r}|^2$ in the potential cancel, and we obtain

$$V = \frac{e^2}{|\mathbf{r}|^3}\left[\mathbf{r}_1\cdot\mathbf{r}_2 - 3(\hat{r}\cdot\mathbf{r}_1)(\hat{r}\cdot\mathbf{r}_2)\right] + O\left(\frac{1}{|\mathbf{r}|^4}\right). \tag{14.146}$$

We can choose a frame of reference with the z-axis along $\mathbf{r}$,

$$\hat{z} = \hat{r}. \tag{14.147}$$

Then

$$V = \frac{e^2}{|\mathbf{r}|^3}\left[x_1x_2 + y_1y_2 - 2z_1z_2\right]. \tag{14.148}$$

We are interested in the corrections to the energy of the two atoms in their ground state. The ground state of the unperturbed Hamiltonian H_0 is

$$|0\rangle_{(0)} = |1,0,0\rangle_{(1)}\cdot|1,0,0\rangle_{(2)}, \tag{14.149}$$

where the subscripts $_{(1)}$ and $_{(2)}$ label the usual eigenkets $|n,l,m\rangle$ of the two independent hydrogen atoms. Recall that $|1,0,0\rangle$ is invariant under a parity transformation, which changes the coordinates from

$$(x,y,z) \quad \text{to} \quad (-x,-y,-z). \tag{14.150}$$

So

$$|1,0,0\rangle \overset{\text{parity}}{\longrightarrow} |1,0,0\rangle. \tag{14.151}$$

The ground-state expectation value of the perturbation is

$$\begin{aligned}{}_{(0)}\langle 0|V|0\rangle_{(0)} = {} & {}_{(1)}\langle 1,0,0|\,{}_{(2)}\langle 1,0,0|V|1,0,0\rangle_{(1)}\,|1,0,0\rangle_{(2)} \\ = \frac{e^2}{|\mathbf{r}|^3}\Big\{ & {}_{(1)}\langle 1,0,0|x_1|1,0,0\rangle_{(1)}\,{}_{(2)}\langle 1,0,0|x_2|1,0,0\rangle_{(2)} \\ & + {}_{(1)}\langle 1,0,0|y_1|1,0,0\rangle_{(1)}\,{}_{(2)}\langle 1,0,0|y_2|1,0,0\rangle_{(2)} \\ & - 2\,{}_{(1)}\langle 1,0,0|z_1|1,0,0\rangle_{(1)}\,{}_{(2)}\langle 1,0,0|z_2|1,0,0\rangle_{(2)}\Big\}.\end{aligned} \tag{14.152}$$

This is zero, since all terms vanish, as we can show using parity. Take as an example

$$ {}_{(1)}\langle 1,0,0|y_1\,|1,0,0\rangle_{(1)}\,, \tag{14.153} $$

which is a three-dimensional integral over the coordinates of the electron of the first atom. Performing the parity transformation

$$ (x_1, y_1, z_1) \quad \longrightarrow \quad (-x_1, -y_1, -z_1) \tag{14.154} $$

and using the invariance of $|1,0,0\rangle$ under this transformation, we obtain

$$ \begin{aligned} {}_{(1)}\langle 1,0,0|y_1\,|1,0,0\rangle_{(1)} &= {}_{(1)}\langle 1,0,0|(-y_1)\,|1,0,0\rangle_{(1)} \\ &= -\,{}_{(1)}\langle 1,0,0|y_1\,|1,0,0\rangle_{(1)}\,. \end{aligned} \tag{14.155} $$

A number which is equal to minus itself must vanish. Therefore,

$$ {}_{(1)}\langle 1,0,0|y_1\,|1,0,0\rangle_{(1)} = 0. \tag{14.156} $$

Analogously, all other terms in Eq. (14.152) vanish. Therefore, at first order in perturbation theory the correction to the ground-state energy is

$$ {}_{(0)}\langle 0|V\,|0\rangle_{(0)} = 0. \tag{14.157} $$

The ground-state energy is corrected only at second order in perturbation theory. The second-order correction is given by the familiar formula

$$ \Delta E^{(2)} = \sum_{k \neq \substack{\text{ground}\\ \text{level}}} \frac{\left|\langle k^{(0)}|\,V\,|0^{(0)}\rangle\right|^2}{E_{\substack{\text{ground}\\ \text{level}}} - E_k^{(0)}}, \tag{14.158} $$

where once again $|0^{(0)}\rangle$ is the ground state of the two independent atoms and the $|k^{(0)}\rangle$,

$$ \begin{aligned} |0^{(0)}\rangle &= |1,0,0\rangle\,|1,0,0\rangle \\ |k^{(0)}\rangle &= |n_1,l_1,m_1\rangle\,|n_2,l_2,m_2\rangle\,, \end{aligned} \tag{14.159} $$

are the excited states constructed from the states of the two independent atoms. It is important that

$$ E_{\substack{\text{ground}\\ \text{level}}} < E_k^{(0)} \quad \forall k \neq {}^{\text{ground}}_{\text{level}}. \tag{14.160} $$

Also, the numerators in the sum are always positive. From the form of the potential, we readily infer that

$$ \Delta E^{(2)} = -\frac{e^4}{|\mathbf{r}|^6} \cdot (\text{positive quantity}). \tag{14.161} $$

Thus, the closer we bring the two atoms, the lower the energy of the total system. *The force is therefore attractive*. We should recall the assumption

$$ |\mathbf{r}| \gg |\mathbf{r}_1|\,, |\mathbf{r}_2|\,. \tag{14.162} $$

When we bring the two atoms too close together, perturbation theory can no longer be used and our conclusion of an attractive force is no longer justified. **Remark:** Explicitly, $\Delta E^{(2)} = -C_6/|\mathbf{r}|^6$ with $C_6 \approx 6.50\, e^2 a_0^5$ for two ground-state hydrogen atoms (Pauling–Beach). The sum over $|k^{(0)}\rangle$ includes the continuum (ionised) states of each atom. The $1/r^6$ law holds for $a_0 \ll r \ll \lambda_{\text{transition}}$; at larger distances retardation gives the Casimir–Polder $1/r^7$ law. See, e.g., Sakurai, *Modern Quantum Mechanics*, Sec. 5.2 (problem on the van der Waals interaction).

14.3.2 Example 2: Relativistic Kinetic Energy

The energy levels of the hydrogen atom are degenerate: they depend only on the principal quantum number n. Relativistic corrections lift this degeneracy. **Remark:** The p^4 term alone lifts the l-degeneracy, but it is only one of three corrections of the same order $\alpha^2 E_n$: together with spin–orbit coupling and the Darwin term the fine-structure shift is $\Delta E_{nj} = -\frac{\alpha^2 E_n}{n^2}\left(\frac{3}{4} - \frac{n}{j+1/2}\right)$ with sign conventions as in Eq. (14.190), which depends only on n and j (e.g. $2s_{1/2}$ and $2p_{1/2}$ remain degenerate until the Lamb shift). See, e.g., Sakurai, *Modern Quantum Mechanics*, Sec. 5.3, or Griffiths, *Introduction to Quantum Mechanics*, Sec. 6.3.

So far, we have considered non-relativistic Hamiltonians of the form

$$H = \frac{p^2}{2m} + U. \tag{14.163}$$

In special relativity, the energy of a free particle is

$$\begin{aligned} E =& \sqrt{m^2c^4 + p^2c^2} = mc^2\sqrt{1 + \frac{p^2}{m^2c^2}} \\ \simeq& mc^2\left[1 + \frac{p^2}{2m^2c^2} - \frac{p^4}{8m^4c^4} + \ldots\right]. \end{aligned} \tag{14.164}$$

This is the total energy without the potential energy, i.e. the kinetic energy plus the rest energy:

$$E \simeq mc^2 + \frac{p^2}{2m} - \frac{p^4}{8m^3c^2} + \ldots. \tag{14.165}$$

A proper synthesis of quantum mechanics and special relativity requires quantum field theory and "second quantisation". However, we can do quite a lot with this expansion of the relativistic energy in powers of $1/c$.

The constant term mc^2 is a common shift of all energy levels. Since we are interested in energy differences, we can drop it from the Hamiltonian. The corrected Hamiltonian then reads

$$\boxed{H = \frac{p^2}{2m} + U - \frac{p^4}{8m^3c^2}.} \tag{14.166}$$

The new relativistic term has small but measurable effects on the spectral lines of atoms. Let us consider the hydrogen atom with relativistic kinetic energy corrections treated as a perturbation:

$$H = H_0 + V, \quad H_0 = \frac{p^2}{2m} - \frac{e^2}{r}, \quad V = -\frac{p^4}{8m^3c^2}. \tag{14.167}$$

Note that U in Eq. (14.166) is the actual potential energy and has nothing to do with V in Eq. (14.167), which is the perturbation.

How are the energy levels

$$E_n = -\frac{mc^2\alpha^2}{2n^2}, \quad n = 1, 2, 3, \ldots, \tag{14.168}$$

with

$$\alpha = \frac{e^2}{\hbar c} = \frac{1}{137.036}, \tag{14.169}$$

affected by relativity? For a generic energy level n, there are several degenerate states $|n,l,m\rangle$, with

$$\begin{aligned} L^2 |n,l,m\rangle &= l(l+1) |n,l,m\rangle \\ L_z |n,l,m\rangle &= m |n,l,m\rangle , \end{aligned} \tag{14.170}$$

where

$$\begin{aligned} l &= 0, 1, ..., n-1, \\ m &= -l, ..., 0, ..., l, \end{aligned} \tag{14.171}$$

and L^2, L_z are the usual angular momentum operators. Due to the degeneracies, we ought to use *degenerate perturbation theory*. This requires finding linear combinations of the degenerate states $|n,l,m\rangle$ which diagonalise the perturbation V. We can easily prove that the states $|n,l,m\rangle$ already diagonalise the perturbation,

$$\langle n,l,m| V |n,l',m'\rangle \propto \delta_{ll'}\delta_{mm'}, \tag{14.172}$$

within the space of degenerate states with energy E_n. As a starting point of the proof, we show that

$$\boxed{\begin{aligned} \underline{\text{Exercise:}} \quad & \left[L^2, V\right] = 0 \\ \text{and} \quad & [L_z, V] = 0 \end{aligned}} \tag{14.173}$$

Once this is proven, we have

$$\begin{aligned} 0 &= \langle n,l,m| \left[L^2, V\right] |n,l',m'\rangle \\ &= \langle n,l,m| L^2 V |n,l',m'\rangle - \langle n,l,m| V L^2 |n,l',m'\rangle \\ &= l(l+1) \langle n,l,m| V |n,l',m'\rangle - l'(l'+1) \langle n,l,m| V |n,l',m'\rangle \\ \longrightarrow 0 &= [l(l+1) - l'(l'+1)] \langle n,l,m| V |n,l',m'\rangle . \end{aligned} \tag{14.174}$$

Thus, we obtain

$$\langle n,l,m| V |n,l',m'\rangle = 0, \quad \text{if } l \neq l'. \tag{14.175}$$

Similarly,

$$\begin{aligned} 0 &= \langle n,l,m| [L_z, V] |n,l',m'\rangle \\ &= \langle n,l,m| L_z V |n,l',m'\rangle - \langle n,l,m| V L_z |n,l',m'\rangle \\ &= m \langle n,l,m| V |n,l',m'\rangle - m' \langle n,l,m| V |n,l',m'\rangle \\ \longrightarrow 0 &= [m - m'] \langle n,l,m| V |n,l',m'\rangle . \end{aligned} \tag{14.176}$$

Thus, we obtain

$$\langle n,l,m| V |n,l',m'\rangle = 0, \quad \text{if } m \neq m', \tag{14.177}$$

or, combining the two results,

$$\langle n,l,m| V |n,l',m'\rangle \propto \delta_{ll'}\delta_{mm'}, \tag{14.178}$$

which completes the proof that the perturbation is already diagonal in the basis $|n,l,m\rangle$ of each degenerate subspace. We are now free to compute the corrections to the energy levels at first order in perturbation theory,

$$E_{nlm} = E_n + \Delta E_{nlm}, \tag{14.179}$$

with

$$\Delta E_{nlm} = \langle n,l,m|\, V\, |n,l,m\rangle\,. \tag{14.180}$$

Inserting the perturbation V from Eq. (14.167), we obtain

$$\begin{aligned}
\Delta E_{nlm} &= \langle n,l,m| - \frac{p^4}{8m^3c^2}\,|n,l,m\rangle = -\frac{1}{2mc^2}\,\langle n,l,m|\left(\frac{p^2}{2m}\right)^2|n,l,m\rangle \\
&= -\frac{1}{2mc^2}\,\langle n,l,m|\left(H_0 + \frac{e^2}{r}\right)^2|n,l,m\rangle \\
&= -\frac{1}{2mc^2}\,\langle n,l,m|\, H_0^2 + H_0\left(\frac{e^2}{r}\right) + \left(\frac{e^2}{r}\right)H_0 + \frac{e^4}{r^2}\,|n,l,m\rangle \\
&= -\frac{1}{2mc^2}\left(E_n^2 + 2E_n e^2\left\langle\frac{1}{r}\right\rangle + e^4\left\langle\frac{1}{r^2}\right\rangle\right).
\end{aligned} \tag{14.181}$$

We need to compute the expectation values

$$\left\langle\frac{1}{r}\right\rangle \quad\text{and}\quad \left\langle\frac{1}{r^2}\right\rangle. \tag{14.182}$$

For this we can use a trick.

Theorem: Consider a Hamiltonian H_λ which depends on a parameter λ, with a normalised eigenstate $|\psi_\lambda\rangle$ and eigenvalue E_λ. Then

$$\begin{aligned}
\frac{\partial E_\lambda}{\partial\lambda} &= \frac{\partial}{\partial\lambda}\,\langle\psi_\lambda|\, H_\lambda\,|\psi_\lambda\rangle \\
&= \left(\frac{\partial}{\partial\lambda}\,\langle\psi_\lambda|\right)\left(H_\lambda\,|\psi_\lambda\rangle\right) + \langle\psi_\lambda|\,\frac{\partial H_\lambda}{\partial\lambda}\,|\psi_\lambda\rangle + \left(\langle\psi_\lambda|\, H_\lambda\right)\left(\frac{\partial}{\partial\lambda}\,|\psi_\lambda\rangle\right) \\
&= E_\lambda\left\{\left(\frac{\partial}{\partial\lambda}\,\langle\psi_\lambda|\right)|\psi_\lambda\rangle + \langle\psi_\lambda|\left(\frac{\partial}{\partial\lambda}\,|\psi_\lambda\rangle\right)\right\} + \langle\psi_\lambda|\,\frac{\partial H_\lambda}{\partial\lambda}\,|\psi_\lambda\rangle \\
&= \underbrace{E_\lambda\,\frac{\partial}{\partial\lambda}\underbrace{\langle\psi_\lambda|\psi_\lambda\rangle}_{1}}_{0} + \langle\psi_\lambda|\,\frac{\partial H_\lambda}{\partial\lambda}\,|\psi_\lambda\rangle\,.
\end{aligned} \tag{14.183}$$

We thus obtain the Hellmann–Feynman theorem,

$$\boxed{\langle\psi_\lambda|\,\frac{\partial H_\lambda}{\partial\lambda}\,|\psi_\lambda\rangle = \frac{\partial E_\lambda}{\partial\lambda}}\,. \tag{14.184}$$

We now apply the theorem to the expectation values in Eq. (14.182), treating e^2 as the parameter:

$$\left\langle\frac{1}{r}\right\rangle = -\left\langle\frac{\partial H_0}{\partial e^2}\right\rangle = -\frac{\partial}{\partial e^2}\underbrace{E_n}_{E_n\propto e^4} = -\frac{1}{e^2}2E_n. \tag{14.185}$$

The calculation of $\left\langle\frac{1}{r^2}\right\rangle$ proceeds as follows. Acting on states of angular momentum l, the Hamiltonian is

$$H_0 = \frac{p^2}{2m} - \frac{e^2}{r} = -\frac{\hbar^2}{2m}\left(\frac{1}{r^2}\frac{\partial}{\partial r}r^2\frac{\partial}{\partial r}\right) + \frac{\hbar^2 l(l+1)}{2mr^2} - \frac{e^2}{r}. \tag{14.186}$$

Treating l as a continuous parameter, we have

$$\frac{\partial H_0}{\partial l} = \frac{\hbar^2(2l+1)}{2mr^2}. \tag{14.187}$$

Since $n = n_r + l + 1$, with n_r the radial quantum number held fixed, $dn/dl = 1$. We can therefore write our expectation value as follows:

$$\begin{aligned}\left\langle \frac{1}{r^2} \right\rangle &= \frac{2m}{\hbar^2(2l+1)} \left\langle \frac{\partial H_0}{\partial l} \right\rangle = \frac{2m}{\hbar^2(2l+1)} \frac{\partial E_n}{\partial l} = \frac{2m}{\hbar^2(2l+1)} \frac{\partial E_n}{\partial n} \underbrace{\frac{\partial n}{\partial l}}_{1} \\ &= \frac{2mc^2}{(2l+1)e^4} \frac{e^4}{\hbar^2 c^2} \frac{\partial E_n}{\partial n} = \frac{2mc^2\alpha^2}{(2l+1)e^4} \frac{\partial E_n}{\partial n} = \frac{2mc^2\alpha^2}{(2l+1)e^4} \left(-\frac{2}{n}\right) E_n \\ &= \frac{E_n^2}{e^4} \left(\frac{-4mc^2\alpha^2/(2l+1)n}{-mc^2\alpha^2/2n^2} \right) = \frac{E_n^2}{e^4} \frac{8n}{2l+1}.\end{aligned} \tag{14.188}$$

In total,

$$\begin{aligned}\Delta E_{nlm} &= -\frac{1}{2mc^2} \left(E_n^2 + 2E_n e^2 \left\langle \frac{1}{r} \right\rangle + e^4 \left\langle \frac{1}{r^2} \right\rangle \right) \\ &= -\frac{1}{2mc^2} \left(E_n^2 - 4E_n^2 + \frac{8n}{2l+1} E_n^2 \right) \\ &= \frac{E_n^2}{2mc^2} \left(3 - \frac{8n}{2l+1} \right).\end{aligned} \tag{14.189}$$

It follows that

$$\frac{\Delta E_{nlm}}{E_n} = \frac{E_n}{2mc^2} \left(3 - \frac{8n}{2l+1} \right) = \alpha^2 \left(\frac{2}{n(2l+1)} - \frac{3}{4n^2} \right) \sim \alpha^2 \approx 5 \times 10^{-5}. \tag{14.190}$$

Therefore, the relativistic correction is a very small deviation from the non-relativistic result. Still, it has been measured with very good precision.

14.4 Time Evolution with Time-Dependent Hamiltonians

So far we have studied potentials which do not change with time. We shall now consider time-dependent Hamiltonians. Before we embark on a study of perturbative methods for time-dependent Hamiltonians, it is very instructive to take a second look at the problem of time evolution.

14.4.1 General Formalism

In order to get familiar with the formalism used in this chapter, we look at a very simple example. Let us consider the differential equation

$$\dot{q} = q. \tag{14.191}$$

From calculus we know that the solution of this simple differential equation is

$$q(t) = q_0 e^t, \tag{14.192}$$

where $q_0 \equiv q(0)$. We can, however, also separate the variables and integrate both sides:

$$\begin{aligned} \frac{\mathrm{d}q}{\mathrm{d}t} &= q, \\ \implies \mathrm{d}q &= q(t)\mathrm{d}t, \\ \implies q(t) &= q_0 + \int_0^t dt_1\, q(t_1). \end{aligned} \tag{14.193}$$

The unknown function $q(t)$ appears under the integral on the right-hand side. By inserting this expression for $q(t_1)$ into the integral, we get

$$q(t) = q_0 + \int_0^t dt_1\, q_0 + \int_0^t \int_0^{t_1} dt_1 dt_2\, q(t_2), \tag{14.194}$$

where $t_1 < t$. Iterating this process n times, we obtain

$$\begin{aligned} q(t) = &\left(q_0 + \sum_{i=1}^{n} \int_0^t dt_1 \int_0^{t_1} dt_2 \ldots \int_0^{t_{i-1}} dt_i\, q_0 \right) \\ &+ \int_0^t dt_1 \int_0^{t_1} dt_2 \ldots \int_0^{t_n} dt_{n+1}\, q(t_{n+1}). \end{aligned} \tag{14.195}$$

Exercise

Show that $\int_0^t \int_0^{t_1} \ldots \int_0^{t_{n-1}} dt_1 \ldots dt_n\, q_0 = \frac{1}{n!} t^n q_0.$ (14.196)

Using Eq. (14.196) and continuing the substitution indefinitely, we can rewrite Eq. (14.195) as

$$q(t) = q_0 + q_0 t + q_0 \frac{t^2}{2!} + \cdots + q_0 \frac{t^n}{n!} + \cdots, \tag{14.197}$$

that is,

$$q(t) = \sum_{n=0}^{\infty} q_0 \frac{t^n}{n!}. \tag{14.198}$$

Thus we have recovered the familiar result

$$q(t) = q_0 e^t. \tag{14.199}$$

Time-dependent perturbation theory follows the same principle. We solve the Schrödinger equation, a first-order differential equation, with a series of nested integrals. The new difficulty is that the coefficient of the equation is now an operator, the time-dependent Hamiltonian $H(t)$, and operators at different times do not commute in general. We can overcome this difficulty, and as a result we will find one of the most important results of perturbation theory: Fermi's golden rule.

14.4.2 Time Evolution for Time-Dependent Hamiltonians

Consider a general Hamiltonian $H(t)$ which depends explicitly on time. The Schrödinger equation,

$$i\hbar \frac{\partial}{\partial t} |\psi, t\rangle = H(t) |\psi, t\rangle, \tag{14.200}$$

can be cast into an integral form:

$$\boxed{|\psi, t\rangle = |\psi, t_0\rangle + \frac{1}{i\hbar}\int_{t_0}^{t} dt_1\, H(t_1)\,|\psi, t_1\rangle} \tag{14.201}$$

By inserting Eq. (14.201) into Eq. (14.200), we can verify that Eq. (14.201) satisfies the Schrödinger equation:

$$\begin{aligned} i\hbar\frac{\partial}{\partial t}|\psi, t\rangle &= i\hbar\left[\frac{\partial}{\partial t}|\psi, t_0\rangle + \frac{1}{i\hbar}\frac{\partial}{\partial t}\int_{t_0}^{t} dt_1\, H(t_1)\,|\psi, t_1\rangle\right] \\ &= i\hbar\left[\frac{1}{i\hbar}H(t)\,|\psi, t\rangle\right] = H(t)\,|\psi, t\rangle\,. \end{aligned} \tag{14.202}$$

Furthermore, Eq. (14.201) satisfies the correct initial condition at $t = t_0$,

$$|\psi, t_0\rangle = |\psi, t_0\rangle + \frac{1}{i\hbar}\int_{t_0}^{t_0} dt_1\, H(t_1)\,|\psi, t_1\rangle = |\psi, t_0\rangle\,. \tag{14.203}$$

Equation (14.201) is not a solution of the Schrödinger equation. Rather, it rewrites the differential equation as an integral equation; notice that $|\psi, t\rangle$ appears on both sides. We can now substitute Eq. (14.201) into itself iteratively,

$$|\psi, t\rangle = |\psi, t_0\rangle + \frac{1}{i\hbar}\int_{t_0}^{t} dt_1\, H(t_1)\cdot\left(|\psi, t_0\rangle + \frac{1}{i\hbar}\int_{t_0}^{t_1} dt_2\, H(t_2)\,|\psi, t_2\rangle\right), \tag{14.204}$$

where $t > t_1 > t_2 > t_0$, as can be seen from the limits of the integrals. We may repeat this substitution many times. After n iterations, we obtain

$$\begin{aligned} |\psi, t\rangle = \Bigg[1 &+ \frac{1}{i\hbar}\int_{t_0}^{t} dt_1\, H(t_1) + \frac{1}{(i\hbar)^2}\int_{t_0}^{t} dt_1\int_{t_0}^{t_1} dt_2\, H(t_1)H(t_2) + \ldots \\ &+ \frac{1}{(i\hbar)^n}\int_{t_0}^{t} dt_1\int_{t_0}^{t_1} dt_2 \ldots \int_{t_0}^{t_{n-1}} dt_n\, H(t_1)H(t_2)\ldots H(t_n)\Bigg]\,|\psi, t_0\rangle \\ &+ \frac{1}{(i\hbar)^{n+1}}\int_{t_0}^{t} dt_1\int_{t_0}^{t_1} dt_2 \ldots \int_{t_0}^{t_n} dt_{n+1}\, H(t_1)H(t_2)\ldots H(t_{n+1})\,|\psi, t_{n+1}\rangle\,. \end{aligned} \tag{14.205}$$

Iterating indefinitely, we obtain

$$|\psi, t\rangle = \left[1 + \sum_{n=1}^{\infty}\frac{1}{(i\hbar)^n}\int_{t_0}^{t} dt_1\int_{t_0}^{t_1} dt_2 \ldots \int_{t_0}^{t_{n-1}} dt_n\, H(t_1)H(t_2)\ldots H(t_n)\right]|\psi, t_0\rangle\,. \tag{14.206}$$

14.4.3 Time Ordering

Let us consider the second-order term in the series above,

$$I = \int_{t_0}^{t} dt_1\int_{t_0}^{t_1} dt_2\, H(t_1)H(t_2). \tag{14.207}$$

We can extend both integrations to the interval $[t_0, t]$ by introducing the Heaviside step function

$$\theta(t_1 - t_2) = \begin{cases} 1, & \text{if } t_1 > t_2, \\ 0, & \text{otherwise.} \end{cases} \tag{14.208}$$

We have

$$I = \int_{t_0}^{t} dt_1 \int_{t_0}^{t} dt_2 \, H(t_1)H(t_2)\theta(t_1 - t_2). \tag{14.209}$$

We now rename the integration variables,

$$\begin{aligned} t_1 &\to t_2, \\ t_2 &\to t_1. \end{aligned} \tag{14.210}$$

The integral now reads

$$I = \int_{t_0}^{t} dt_2 \int_{t_0}^{t} dt_1 \, H(t_2)H(t_1)\theta(t_2 - t_1). \tag{14.211}$$

Adding Eqs. (14.209) and (14.211), we obtain

$$2I = \int_{t_0}^{t} dt_1 \int_{t_0}^{t} dt_2 \, \{H(t_1)H(t_2)\theta(t_1 - t_2) + H(t_2)H(t_1)\theta(t_2 - t_1)\}. \tag{14.212}$$

We now introduce the time-ordering symbol

$$\mathcal{T}\{A(t_1)B(t_2)\} = \begin{cases} A(t_1)B(t_2), & \text{if } t_1 > t_2, \\ B(t_2)A(t_1), & \text{if } t_2 > t_1, \end{cases} \tag{14.213}$$

which always places the operator at the later time to the left. Then we can write Eq. (14.212) in a compact way:

$$2I = \int_{t_0}^{t} dt_1 \int_{t_0}^{t} dt_2 \, \mathcal{T}\{H(t_1)H(t_2)\}. \tag{14.214}$$

Thus the integral in Eq. (14.207) can be written with the time-ordering symbol:

$$\int_{t_0}^{t} dt_1 \int_{t_0}^{t_1} dt_2 \, H(t_1)H(t_2) = \frac{1}{2} \int_{t_0}^{t} dt_1 \int_{t_0}^{t} dt_2 \, \mathcal{T}\{H(t_1)H(t_2)\}. \tag{14.215}$$

The time-ordering symbol can be generalised to more than two operators,

$$\mathcal{T}\{A(t_1)B(t_2)\Gamma(t_3)\} = \begin{cases} A(t_1)B(t_2)\Gamma(t_3), & \text{if } t_1 > t_2 > t_3, \\ \Gamma(t_3)A(t_1)B(t_2), & \text{if } t_3 > t_1 > t_2, \\ \quad \dots & \quad \dots, \\ \quad \dots & \quad \dots, \end{cases} \tag{14.216}$$

ordering them from later to earlier times (from left to right).

Exercise

Prove the relation

$$\begin{aligned} &\int_{t_0}^{t} dt_1 \int_{t_0}^{t_1} dt_2 \dots \int_{t_0}^{t_{n-1}} dt_n \, H(t_1)H(t_2) \dots H(t_n) \\ =&\frac{1}{n!} \int_{t_0}^{t} dt_1 \dots dt_n \, \mathcal{T}\{H(t_1)H(t_2) \dots H(t_n)\} \end{aligned} \tag{14.217}$$

With the time-ordering symbol and Eq. (14.217), we can rewrite Eq. (14.206) as

$$|\psi, t\rangle = \left[1 + \sum_{n=1}^{\infty} \frac{1}{(i\hbar)^n} \frac{1}{n!} \int_{t_0}^{t} dt_1 \ldots dt_n \, \mathcal{T}\{H(t_1)H(t_2) \ldots H(t_n)\}\right] |\psi, t_0\rangle \,. \tag{14.218}$$

Were it not for the time ordering, the sum would be an exponential. We write it as

$$|\psi, t\rangle = \mathcal{T}\left\{e^{-\frac{i}{\hbar}\int_{t_0}^{t} dt' \, H(t')}\right\} |\psi, t_0\rangle \,, \tag{14.219}$$

where the time-ordered exponential is defined as

$$\mathcal{T}\left\{e^{\int_{t_0}^{t} dt' \, A(t')}\right\} \equiv 1 + \sum_{n=1}^{\infty} \int_{t_0}^{t} dt_1 \ldots \int_{t_0}^{t} dt_n \, \frac{\mathcal{T}\{A(t_1) \ldots A(t_n)\}}{n!}. \tag{14.220}$$

Eq. (14.219) describes the evolution of a quantum-mechanical system with a general time-dependent Hamiltonian.

Exercise

Prove using Eq. (14.219) that

$$|\psi, t\rangle = e^{-\frac{i}{\hbar}H(t-t_0)} |\psi, t_0\rangle$$

if $\dfrac{\partial H}{\partial t} = 0$. (14.221)

Exercise

Prove that

$$|\psi, t\rangle = \mathcal{T}\left\{e^{-\frac{i}{\hbar}\int_{t_0}^{t} dt' \, H(t')}\right\} |\psi, t_0\rangle \tag{14.222}$$

is a solution of the Schrödinger equation.

14.5 Perturbation Theory for Time-Dependent Potentials

We can now use what we have learnt about time evolution with time-dependent Hamiltonians to develop a perturbation theory. Let us consider a Hamiltonian

$$H(t) = H_0 + V(t), \tag{14.223}$$

where $V(t)$ is a time-dependent perturbation and H_0 is an exactly solvable time-independent Hamiltonian. The Schrödinger equation for this Hamiltonian is

$$i\hbar \frac{\partial}{\partial t} |\psi, t\rangle = [H_0 + V(t)] |\psi, t\rangle \,. \tag{14.224}$$

For $V(t) = 0$ we have

$$|\psi, t\rangle = U_0(t, t_0) |\psi, t_0\rangle \,, \tag{14.225}$$

with

$$U_0(t, t_0) = \mathcal{T}\left\{e^{-\frac{i}{\hbar}\int_{t_0}^{t} dt' \, H_0}\right\}, \tag{14.226}$$

which is the solution we found in Sec. 14.4.3. Since H_0 is time-independent, we can drop the time ordering and obtain the usual time evolution for time-independent Hamiltonians:

$$U_0(t,t_0) = e^{-\frac{i}{\hbar}H_0\int_{t_0}^{t}dt'} = e^{-\frac{i}{\hbar}H_0(t-t_0)}. \tag{14.227}$$

Since $V(t)$ is a small perturbation, we expect that the time evolution of our system is dominated by the time-independent part H_0. This leads us to an Ansatz for the state $|\psi,t\rangle$ at time t which displays explicitly the evolution factor due to the constant part of the Hamiltonian:

$$|\psi,t\rangle = e^{-\frac{i}{\hbar}H_0(t-t_0)}\,|\psi_I,t\rangle\,. \tag{14.228}$$

How do we interpret this Ansatz? The idea is that the as yet undetermined state $|\psi_I,t\rangle$ captures the effect of the interaction $V(t)$. We substitute this expression into Eq. (14.224) and evaluate the two sides separately:

$$\begin{aligned}\text{LHS} &= i\hbar\frac{\partial}{\partial t}\left[e^{-\frac{i}{\hbar}H_0(t-t_0)}\,|\psi_I,t\rangle\right],\\ &= H_0 e^{-\frac{i}{\hbar}H_0(t-t_0)}\,|\psi_I,t\rangle + i\hbar e^{-\frac{i}{\hbar}H_0(t-t_0)}\frac{\partial}{\partial t}\,|\psi_I,t\rangle\,,\\ &= H_0\,|\psi,t\rangle + e^{-\frac{i}{\hbar}H_0(t-t_0)}\,i\hbar\frac{\partial}{\partial t}\,|\psi_I,t\rangle\,.\end{aligned} \tag{14.229}$$

$$\begin{aligned}\text{RHS} &= H_0\,|\psi,t\rangle + V(t)\,|\psi,t\rangle\,,\\ &= H_0\,|\psi,t\rangle + V(t)e^{-\frac{i}{\hbar}H_0(t-t_0)}\,|\psi_I,t\rangle\,.\end{aligned} \tag{14.230}$$

Equating the two sides, we find

$$i\hbar\frac{\partial}{\partial t}\,|\psi_I,t\rangle = V_I(t)\,|\psi_I,t\rangle\,, \tag{14.231}$$

with

$$V_I(t) = e^{\frac{i}{\hbar}H_0(t-t_0)}\,V(t)\,e^{-\frac{i}{\hbar}H_0(t-t_0)}. \tag{14.232}$$

14.5.1 The Interaction Picture

Eq. (14.232) looks very similar to the time evolution of operators in the Heisenberg picture, which we know from Part I. However, H_0 is not the complete Hamiltonian, so this is an "incomplete Heisenberg picture". This defines a new picture, the interaction picture. In the interaction picture, operators evolve as

$$\underbrace{A_I(t)}_{\substack{\text{Interaction}\\ \text{Picture}}} = e^{\frac{i}{\hbar}H_0(t-t_0)}\ \underbrace{A_S(t)}_{\substack{\text{Schrödinger}\\ \text{Picture}}}\ e^{-\frac{i}{\hbar}H_0(t-t_0)}. \tag{14.233}$$

Similarly, states in the interaction picture are related to states in the Schrödinger picture by

$$|\psi_I,t\rangle = e^{\frac{i}{\hbar}H_0(t-t_0)}\,|\psi_S,t\rangle\,. \tag{14.234}$$

Note that the state $|\psi_S,t\rangle$ itself evolves, as it is written in the Schrödinger picture.

The interaction picture is intermediate between the Heisenberg and Schrödinger pictures: both operators and states depend on time. Operators evolve with H_0, the constant part of the Hamiltonian, while states evolve with $V(t)$ only. For $H = H_0$ and $V(t) = 0$ the

interaction picture coincides with the Heisenberg picture while for $H = V(t)$ and $H_0 = 0$ it is the same as the Schrödinger picture.

We now return to Eq. (14.231). Its general solution is

$$|\psi_I, t\rangle = \mathcal{T}\left\{e^{-\frac{i}{\hbar}\int_{t_0}^{t} dt'\, V_I(t')}\right\} |\psi_I, t_0\rangle\,, \tag{14.235}$$

since $V_I(t)$ plays the role of a Hamiltonian in Eq. (14.231). For general potentials $V_I(t)$ this formula is not particularly useful, since all terms of the infinite series of time-ordered integrals may be important. If, however, $V_I(t)$ is a small perturbation, we expect the first few terms of the series to be a good approximation. We write

$$|\psi_I, t\rangle = U_I(t, t_0)\, |\psi_I, t_0\rangle \tag{14.236}$$

and expand $U_I(t, t_0)$:

$$U_I(t, t_0) = 1 - \frac{i}{\hbar}\int_{t_0}^{t} dt_1\, V_I(t_1) - \frac{1}{\hbar^2}\int_{t_0}^{t} dt_1 \int_{t_0}^{t_1} dt_2\, V_I(t_1)V_I(t_2) + \mathcal{O}(V_I^3). \tag{14.237}$$

This equation is our basic tool for time-dependent perturbation theory.

14.5.2 Transition Amplitudes

Let us now look at some applications of time-dependent perturbation theory. Our first example is transition amplitudes. We have a Hamiltonian which consists of a constant part H_0 and a time-dependent potential $V(t)$, which will be treated perturbatively. The perturbation starts acting at time $t = t_0$. Thus our Hamiltonian reads

$$H = H_0 + V(t)\theta(t - t_0). \tag{14.238}$$

Additionally, we assume that the system is in an energy eigenstate before the perturbation acts:

$$\begin{aligned} |\psi, t_0\rangle &= |i, t_0\rangle \equiv |i\rangle\,, \\ H_0\, |\psi, t_0\rangle &= H_0\, |i\rangle = E_i\, |i\rangle\,. \end{aligned} \tag{14.239}$$

What is the probability that at a later time $t > t_0$ the system is in a given eigenstate $|f\rangle$ of H_0? We need the probability amplitude for this transition, which we call the transition amplitude:

$$\mathcal{M}_{i\to f} = \langle f, t|\psi, t\rangle\,, \tag{14.240}$$

where $|\psi, t\rangle$ denotes the state of the system at time t. Since $|f, t\rangle$ is an eigenstate of the unperturbed Hamiltonian, it satisfies

$$H_0\, |f, t\rangle = E_f\, |f, t\rangle\,, \qquad |f, t\rangle = e^{-\frac{i}{\hbar}H_0(t-t_0)}\, |f, t_0\rangle\,. \tag{14.241}$$

We insert the second relation into Eq. (14.240):

$$\begin{aligned} \mathcal{M}_{i\to f} &= \langle f, t|\psi, t\rangle \\ &= \langle f, t_0|\, e^{\frac{i}{\hbar}H_0(t-t_0)}\, |\psi, t\rangle \\ &= \langle f|\psi_I, t\rangle\,, \end{aligned} \tag{14.242}$$

where $|f\rangle = |f, t_0\rangle$. Inserting the time evolution (14.236) of $|\psi_I, t\rangle$ into Eq. (14.242), we have

$$\begin{aligned} \mathcal{M}_{i\to f} &= \langle f|\psi_I, t\rangle \\ &= \langle f|U_I(t, t_0)|\psi_I, t_0\rangle\,. \end{aligned} \tag{14.243}$$

Furthermore, we can relate $|\psi_I, t_0\rangle$ to $|\psi, t_0\rangle$:

$$\begin{aligned} |\psi_I, t_0\rangle &= e^{\frac{i}{\hbar}H_0(t_0-t_0)}\,|\psi, t_0\rangle \\ &= e^0\,|\psi, t_0\rangle \\ &= |i\rangle\,. \end{aligned} \tag{14.244}$$

Thus our final expression for the transition amplitude reads

$$\boxed{\mathcal{M}_{i\to f} = \langle f|U_I(t, t_0)|i\rangle}\,. \tag{14.245}$$

So far we have not made any approximation. Let us now see what happens at first order in perturbation theory:

$$\begin{aligned} \mathcal{M}_{i\to f} &= \langle f|\left[1 - \frac{i}{\hbar}\int_{t_0}^{t} dt'\, V_I(t')\right]|i\rangle + \mathcal{O}(V_I(t)^2) \\ &= \langle f|i\rangle - \frac{i}{\hbar}\int_{t_0}^{t} dt'\, \langle f|V_I(t')|i\rangle + \mathcal{O}(V_I(t)^2). \end{aligned} \tag{14.246}$$

We can simplify the matrix element $\langle f|V_I(t')|i\rangle$:

$$\begin{aligned} \langle f|V_I(t')|i\rangle &= \langle f|\, e^{\frac{i}{\hbar}H_0(t'-t_0)}V(t')e^{-\frac{i}{\hbar}H_0(t'-t_0)}\,|i\rangle \\ &= e^{\frac{i}{\hbar}E_f(t'-t_0)}\,\langle f|V(t')|i\rangle\, e^{-\frac{i}{\hbar}E_i(t'-t_0)} \\ &= e^{\frac{i}{\hbar}(E_f-E_i)(t'-t_0)}\,\langle f|V(t')|i\rangle\,. \end{aligned} \tag{14.247}$$

The transition amplitude then becomes

$$\boxed{\mathcal{M}_{i\to f} = \delta_{if} - \frac{i}{\hbar}\int_{t_0}^{t} dt'\, e^{i\omega_{fi}(t'-t_0)}\,\langle f|V(t')|i\rangle}\,, \tag{14.248}$$

where $\omega_{fi} = (E_f - E_i)/\hbar$.

14.5.3 Sudden Constant Perturbation

Let us consider a Hamiltonian with a constant part and a constant perturbation which is switched on at time $t_0 = 0$. Our Hamiltonian takes the form

$$H = H_0 + V\theta(t). \tag{14.249}$$

We assume that the system is initially in an eigenstate of H_0:

$$|\psi, t_0 = 0\rangle = |i\rangle\,, \qquad H_0\,|i\rangle = E_i\,|i\rangle\,. \tag{14.250}$$

The transition amplitude for finding the system in an eigenstate $|f\rangle$ at a later time t is

$$\begin{aligned}\mathcal{M}_{i\to f} &= \delta_{if} - \frac{i}{\hbar}\int_{t_0}^{t} dt'\, e^{i\omega_{fi}(t'-t_0)}\langle f|V(t')|i\rangle \\ &= \delta_{if} - \frac{i}{\hbar}\left[\int_{t_0}^{t} dt'\, e^{i\omega_{fi}(t'-t_0)}\right]\langle f|V|i\rangle \\ &= \delta_{if} + \frac{\langle f|V|i\rangle}{\hbar\omega_{fi}}\left(1 - e^{i\omega_{fi}(t-t_0)}\right).\end{aligned} \tag{14.251}$$

We want to examine the case of a true transition, in which the final state is different from the initial state:

$$|f\rangle \neq |i\rangle\,. \tag{14.252}$$

The probability of this transition is the modulus squared of the transition amplitude:

$$\begin{aligned}P_{i\to f} &= \left|\mathcal{M}_{i\to f}\right|^2 \\ &= \left|\delta_{if} + \frac{\langle f|V|i\rangle}{\hbar\omega_{fi}}\left(1 - e^{i\omega_{fi}(t-t_0)}\right)\right|^2 \\ &= \frac{4\left|\langle f|V|i\rangle\right|^2}{\left(\hbar\omega_{fi}\right)^2}\sin^2\left(\frac{\omega_{fi}\,(t-t_0)}{2}\right).\end{aligned} \tag{14.253}$$

We see that at first order in time-dependent perturbation theory the transition probability vanishes for

$$\frac{\omega_{fi}\,(t-t_0)}{2} = n\pi, \qquad \text{for } n = 1, 2, \ldots \tag{14.254}$$

We now take a closer look at the case where the difference between the energies of the initial and final states is small. To this end, let us plot the function

$$\frac{4\sin^2\left(\frac{\omega_{fi}(t-t_0)}{2}\right)}{\omega_{fi}^2}. \tag{14.255}$$

Fig. 14.2 shows the plot, with $\Delta t = t - t_0$. For $\omega_{fi} \to 0$ the function approaches Δt^2, and the distance between the two zeros closest to the origin, at $\omega_{fi} = \pm 2\pi/\Delta t$, is proportional to $1/\Delta t$. For $\omega_{fi} \to 0$ the probability is

$$P_{i\to f} = \frac{\left|\langle f|V|i\rangle\right|^2 \Delta t^2}{\hbar^2}. \tag{14.256}$$

How do we interpret this result? The bulk of the area under the curve lies between these two zeros. We can show this by computing the total area under the curve of Eq. (14.255) and comparing it with the area between $\pm 2\pi/\Delta t$:

$$\int_{-\infty}^{\infty} \frac{4\sin^2\left(\frac{\omega_{fi}(t-t_0)}{2}\right)}{\omega_{fi}^2} d\omega_{fi} = 2\pi\Delta t, \tag{14.257}$$

and

$$\int_{-\frac{2\pi}{\Delta t}}^{\frac{2\pi}{\Delta t}} \frac{4\sin^2\left(\frac{\omega_{fi}(t-t_0)}{2}\right)}{\omega_{fi}^2} d\omega_{fi} \approx 5.673\Delta t. \tag{14.258}$$

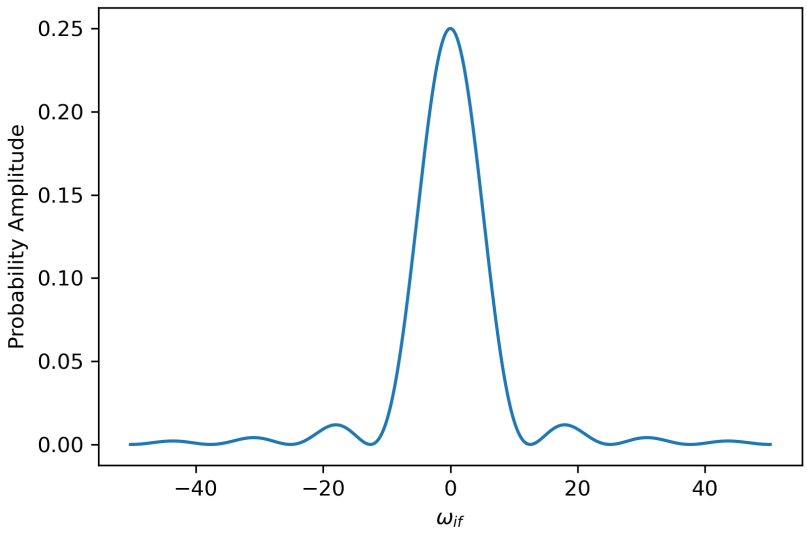


Figure 14.2. The function of Eq. (14.255) as a function of ω_{fi}, for fixed $\Delta t = t - t_0$.

Thus $\approx$ 90% of the area lies between the two zeros. This means that the transition probability is appreciable only for states that satisfy

$$\Delta t \sim \frac{2\pi}{\omega_{fi}} = \frac{2\pi\hbar}{\left|E_f - E_i\right|}. \tag{14.259}$$

or, up to a factor of order one,

$$\Delta t \Delta E \sim \hbar. \tag{14.260}$$

The transition probabilities are thus controlled by Planck's constant $\hbar$ and by the duration of the perturbation. For short perturbations, transitions to a wider range of energies are possible.

14.5.4 Fermi's Golden Rule

We continue the study of the transition probability of Eq. (14.253) for

$$H = H_0 + \theta(t)V, \quad \frac{\partial V}{\partial t} = 0. \tag{14.261}$$

We are interested in the probability of a transition from the state $|i\rangle$ to any different state $|f\rangle$. As we shall see, under some reasonable assumptions we can draw general conclusions about how the state $|i\rangle$ decays due to the perturbation.

This probability is the sum of the probabilities of all possible transitions,

$$P_{i\to\text{ANY}} = \sum_{\substack{f \\ |f\rangle\neq|i\rangle}} P_{i\to f}. \tag{14.262}$$

We consider the general case where the final states $|f\rangle$ form a continuous spectrum. The sum above is then replaced by an integral. If the final states are distributed in energy with density $\rho(E_f)$, we write

$$\sum_f \quad \longrightarrow \quad \int_{-\infty}^{\infty} dE_f\, \rho(E_f), \tag{14.263}$$

where the density $\rho(E_f)$ is finite and positive. We normalise it to one,

$$\int_{-\infty}^{\infty} dE_f\, \rho(E_f) = 1. \tag{14.264}$$

Note that with this normalisation $\rho(E)$ has dimensions of inverse energy. **Remark:** Normalising $\int \rho = 1$ treats ρ as a probability distribution of final energies. For a genuine density of states (number of states per unit energy, e.g. $\rho(E) = Vmp/(2\pi^2\hbar^3)$ for a free particle in a box of volume V, summed over directions) the integral is not 1 and in general diverges; only $\rho(E_i)$ enters the golden rule (14.274), so the normalisation is a convention of this simple setting and plays no role in the result; see, e.g., Sakurai, *Modern Quantum Mechanics*, Sec. 5.6. Using Eq. (14.253), the transition probability becomes

$$P_{i\to\mathrm{ANY}} = \int dE_f\, \rho(E_f) \frac{4\left|\langle f|V|i\rangle\right|^2}{\hbar^2} \frac{\sin^2\left(\frac{\omega_{fi}\Delta t}{2}\right)}{\omega_{fi}^2}. \tag{14.265}$$

It is useful to compute the transition rate

$$\Gamma_i \equiv \frac{dP_{i\to\mathrm{ANY}}}{dt}, \tag{14.266}$$

which measures how fast the state $|i\rangle$ is depopulated at a given time $t > t_0$. We obtain

$$\Gamma_i = \int dE_f\, \rho(E_f) \frac{2\left|\langle f|V|i\rangle\right|^2}{\hbar^2} \frac{\sin\left(\omega_{fi}\Delta t\right)}{\omega_{fi}}, \tag{14.267}$$

where we used $\frac{d}{dt}\sin^2(t/2) = \sin(t/2)\cos(t/2) = \frac{\sin(t)}{2}$. Let us manipulate Eq. (14.267) further by changing the integration variable from E_f to the dimensionless variable λ,

$$E_f = \lambda \cdot E_i. \tag{14.268}$$

Thus we have

$$\Gamma_i = \int d\lambda\, E_i \rho(\lambda E_i) \frac{2\left|\langle f|V|i\rangle\right|^2}{\hbar} \frac{\sin\left((\lambda-1)\frac{E_i\Delta t}{\hbar}\right)}{E_i(\lambda-1)}. \tag{14.269}$$

We now assume that $E_i\Delta t/\hbar \gg 1$, i.e. $\Delta t \gg \hbar/E_i$. This means that we wait long enough for the energy–time uncertainty to be small. Using the limit of Exercise 2 below, we obtain

$$\Gamma_i = \int d\lambda\, E_i \rho(\lambda E_i) \frac{2\left|\langle f|V|i\rangle\right|^2}{\hbar} \frac{\pi}{E_i} \delta(\lambda - 1). \tag{14.270}$$

Changing the variable back from λ to E_f, we have

$$\Gamma_i = \frac{2\pi}{\hbar} \int dE_f\, \rho(E_f) \left|\langle f|V|i\rangle\right|^2 \delta(E_f - E_i). \tag{14.271}$$

The delta function selects the states $|f\rangle$ for which the energy of the system is conserved. Often the squared matrix elements

$$\left|\langle f|V|i\rangle\right|^2 \tag{14.272}$$

are similar for all states $|f\rangle$ with the same energy E_f. If this is the case, we can replace them by an average value,

$$\left|\langle f|V|i\rangle\right|^2 \quad\longrightarrow\quad \overline{\left|\langle f|V|i\rangle\right|}^2. \tag{14.273}$$

Then, we obtain a rather simple result:

$$\boxed{\Gamma_i = \frac{2\pi}{\hbar}\overline{\left|\langle f|V|i\rangle\right|}^2 \rho(E_i).} \tag{14.274}$$

This equation is called Fermi's golden rule; it was first derived by Dirac. Alternatively, one often writes

$$\Gamma_i = \int dE_f \rho(E_f)\Gamma_{i\to f}, \tag{14.275}$$

with

$$\boxed{\Gamma_{i\to f} = \frac{2\pi}{\hbar}\overline{\left|\langle f|V|i\rangle\right|}^2 \delta(E_f - E_i).} \tag{14.276}$$

We should always keep in mind the approximations that we have made in deriving Eq. (14.276):

(i) $\left|\langle f|V|i\rangle\right|^2 \quad\longrightarrow\quad \overline{\left|\langle f|V|i\rangle\right|}^2$;

(ii) $\Delta t \gg \hbar/E_i$.

Remark: The substitution $E_f = \lambda E_i$ singles out E_i as the energy scale, which is not physical (energies are defined up to a constant; for $E_i < 0$ the limit (14.270) even flips sign, and for $E_i = 0$ it fails). The precise form of condition (ii) is $\Delta t \gg \hbar/\delta E$, where δE is the energy range over which $\rho(E_f)\left|\langle f|V|i\rangle\right|^2$ varies appreciably around E_i, together with $\Gamma_i \Delta t \ll 1$ for first-order perturbation theory to hold. Exercise 1 below illustrates this: there $\delta E = 1$ (the width of ρ) and $\Gamma_i(t)$ approaches the golden-rule value as e^{-t}. See, e.g., Sakurai, *Modern Quantum Mechanics*, Sec. 5.6, or Cohen-Tannoudji, Diu and Laloë, *Quantum Mechanics*, Ch. XIII.

We should now examine the range of validity of our perturbative computation. We have found in Eq. (14.274) that the transition rate Γ_i becomes constant after a time $\Delta t \sim \hbar/E_i$ has elapsed. Therefore, the transition probability at $t \gg \Delta t$ is

$$P_{i\to\mathrm{ANY}} = \int_{t_0}^{t} dt'\, \Gamma_i(t') \overset{t\gg\Delta t}{\sim} \Gamma_i t, \tag{14.277}$$

where we used that Γ_i is constant for $t \gg \Delta t$. There is therefore a time $t = t_{\mathrm{NP}}$ after which the transition probability would exceed one. Clearly, our perturbative approach breaks down at times larger than the inverse of the transition rate:

$$P_{i\to\mathrm{ANY}} = \Gamma_i t_{\mathrm{NP}} = 1 \quad\longrightarrow\quad t_{\mathrm{NP}} = \frac{1}{\Gamma_i}. \tag{14.278}$$

In practice we can do better by applying perturbation theory to very short time intervals and combining the effects cleverly into an exponential decay law for the number N of systems remaining in the initial state,

$$N = N_0 e^{-\Gamma_0 t}, \tag{14.279}$$

with $\Gamma_0 \equiv \Gamma_i$. The exponential law follows because the depopulation rate per remaining system is constant (Weisskopf–Wigner).

Intuitively, Fermi's golden rule says that the rate of transitions out of a state $|i\rangle$ is proportional to the squared strength $\left|\langle f|V|i\rangle\right|^2$ of the interaction between the initial and final states and to the density $\rho(E_i)$ of final states with the same energy as the initial state. The stronger the coupling and the more final states are available, the faster the initial state decays.

Furthermore, Fermi's golden rule applies to systems that undergo a large number of transitions, and it provides a statistical description of transition rates in terms of the energy of the system and the properties of its environment. The rule plays a crucial role in understanding the behaviour of atoms, molecules and other quantum systems in a variety of contexts, including quantum optics, condensed matter physics and nuclear physics.

14.5.5 Exercises: Golden Rule for a Lorentzian Density of States

Exercise 1. Consider a system with Hamiltonian

$$H = H_0 + \theta(t)\, V\,, \tag{14.280}$$

with

$$H_0\,|i\rangle = |i\rangle \quad \text{for the state } |i\rangle \qquad \text{and} \qquad V\,|f\rangle = \lambda\,|i\rangle \quad \forall\ |f\rangle \neq |i\rangle\,. \tag{14.281}$$

The states $|f\rangle$ have the energy density

$$\rho(E_f) = \frac{1}{\pi}\,\frac{1}{1+E_f^2}. \tag{14.282}$$

(i) Show that $\int_{-\infty}^{+\infty} \rho(E_f)\, dE_f = 1$.

(ii) Compute the transition rate for arbitrary t.

(iii) When does the "golden rule" become a good approximation?

Exercise 2. Prove that

$$\lim_{x\to\infty} \frac{\sin\left(x(\omega-\omega_0)\right)}{\omega-\omega_0} = \pi\,\delta(\omega-\omega_0). \tag{14.283}$$

Solution 1

(i) We consider the contour $C = C_L + C_R$, where C_L runs along the real axis from $-R$ to R and C_R is the semicircle of radius R in the upper half-plane. The integrand of $\frac{1}{\pi}\int_C dz\,\frac{1}{1+z^2}$ has poles at $z = \pm i$, of which only $z = i$ lies inside the contour. We have

$$\oint_C \frac{dz}{1+z^2} = 2\pi i \operatorname*{Res}_{z=i}\left(\frac{1}{1+z^2}\right) = 2\pi i\left(\operatorname*{Res}_{z=i}\frac{1}{(z-i)(z+i)}\right) = \pi. \tag{14.284}$$

Since

$$\int_C = \underbrace{\int_{C_R}}_{\substack{\text{vanishes}\\ \text{at } R\to\infty}} + \int_{C_L} \quad\Longrightarrow\quad \int_{-\infty}^{+\infty} dx\,\frac{1}{1+x^2} = \pi, \tag{14.285}$$

as we wanted (so that $\int \rho\, dE_f = \pi/\pi = 1$).

(ii) The transition rate is (Eq. (14.267) with $t_0 = 0$, $\Delta t = t$):

$$\begin{aligned}\Gamma_i &= \frac{2}{\hbar}\int dE_f\, \rho(E_f)\, \left|\langle f|V|i\rangle\right|^2 \frac{\sin\left(\frac{(E_f-E_i)t}{\hbar}\right)}{E_f - E_i} \\ &= \frac{2}{\hbar}\int dE_f\, \rho(E_f)\, |\lambda|^2 \left|\langle i|i\rangle\right|^2 \frac{\sin\left((E_f-1)t/\hbar\right)}{E_f - 1},\end{aligned} \tag{14.286}$$

and we set $\hbar = 1$ for convenience:

$$\Gamma_i = \frac{2|\lambda|^2}{\pi}\int \frac{dE}{1+E^2}\,\frac{\sin\left((E-1)t\right)}{E-1}. \tag{14.287}$$

We choose a contour $C = C_{L_1} + C_r + C_{L_2} + C_R$ consisting of the segment C_{L_1} along the real axis from $1-R$ to $1-r$, the small semicircle C_r above the point $E = 1$ from $1-r$ to $1+r$, the segment C_{L_2} from $1+r$ to $1+R$, and the large semicircle C_R in the upper half-plane back to $1-R$. Of the poles at $z = \pm i$, only $z = i$ lies inside C. We write

$$\Gamma_i = \frac{2|\lambda|^2}{\pi}\,\mathrm{Im}\,(I_z) \tag{14.288}$$

with

$$I_z = I_C - I_{C_r} - I_{C_R}\Big|_{\substack{r\to 0\\ R\to\infty}}, \qquad I_X = \int_X dz\, f(z), \qquad f(z) = \frac{e^{i(z-1)t}}{(z-1)(1+z^2)}. \tag{14.289}$$

I_C is computed with the residue theorem:

$$I_C = 2\pi i \operatorname*{Res}_{z=i}\left[\frac{e^{i(z-1)t}}{(i+z)(z-i)(z-1)}\right] = \pi\,\frac{e^{-(1+i)t}}{i-1}. \tag{14.290}$$

$$I_{C_R} = \int_{C_R} dz\, f(z) \xrightarrow{z=1+Re^{i\theta}} \xrightarrow{R\to\infty} i\int_0^\pi \frac{d\theta}{1+(Re^{i\theta}+1)^2}\, e^{iRe^{i\theta}t} \xrightarrow{R\to\infty} 0, \tag{14.291}$$

and finally, with $z = 1 + re^{i\theta}$,

$$I_{C_r} = i\int_\pi^0 d\theta\; \frac{e^{i(z-1)t}}{z^2+1}\bigg|_{r=0} = -\frac{i\pi}{2}. \tag{14.292}$$

Therefore,

$$I_z = \pi\left[\frac{e^{-t}\,(\cos t - i\sin t)\,(-1-i)}{2}\right] + \frac{i\pi}{2} \tag{14.293}$$

$$\Longrightarrow \quad \Gamma_i = \frac{2|\lambda|^2}{\pi}\,\mathrm{Im}\, I_z = |\lambda|^2\left[1 + e^{-t}\,(\sin t - \cos t)\right]. \tag{14.294}$$

(iii) The golden rule gives

$$\Gamma_i = \Gamma_i(t)|_{t=\infty} \tag{14.295}$$

(indeed, Eq. (14.274) gives $2\pi|\lambda|^2\,\rho(E_i = 1) = |\lambda|^2$) and the difference is

$$\frac{|\Gamma_i - \Gamma_i(t)|_{t=\infty}|}{\Gamma_i(t)|_{t=\infty}} = e^{-t}\,|\sin t - \cos t|. \tag{14.296}$$

We observe an exponential suppression: the rate very quickly approaches the asymptotic value given by the golden rule.

In units with $\hbar$ restored, t stands for $t/(\hbar/\Delta)$, where $\Delta = 1$ is the width of $\rho(E)$ (the scale over which $\rho\left|\langle f|V|i\rangle\right|^2$ varies); the golden rule is good for $t \gg \hbar/\Delta$, and first-order perturbation theory requires in addition $\Gamma_i t \ll 1$, i.e. $|\lambda|^2 \ll \Delta$.

Solution 2

We consider the integral

$$\begin{aligned} I &= \int_{-\infty}^{+\infty} d\omega\, \rho(\omega)\, \frac{\sin\left((\omega-\omega_0)x\right)}{\omega-\omega_0}, \qquad x \gg 0 \\ &= \mathrm{Im}\left(\underbrace{\int_{-\infty}^{+\infty} d\omega\, \rho(\omega)\, \frac{e^{ix(\omega-\omega_0)}}{\omega-\omega_0}}_{I_z} \right) \equiv \mathrm{Im}\left(I_z\right). \end{aligned} \tag{14.297}$$

We define a contour consisting of the real axis from $\omega_0 - R$ to $\omega_0 + R$, indented by a small semicircle C_r of radius r above ω_0, and closed by the large semicircle C_R in the upper half-plane, and we consider the limits $R \to \infty$ and $r \to 0$. In this limit, we have

$$\begin{aligned} I_z + I_{C_r} + I_{C_R} &= 2\pi i \left[\sum_k \left(\operatorname*{Res}_{\omega=\omega_k} \rho(\omega) \right) \frac{e^{ix(\omega_k-\omega_0)}}{\omega_k-\omega_0} \right] \\ &= 2\pi i \left[\sum_k \left(\operatorname*{Res}_{\omega=\omega_k} \rho(\omega) \right) \frac{e^{-x\,\mathrm{Im}\,\omega_k}\, e^{ix(\mathrm{Re}\,\omega_k-\omega_0)}}{\omega_k-\omega_0} \right]. \end{aligned} \tag{14.298}$$

Since $x \gg 0$ ($x \to \infty$) and $\mathrm{Im}\,\omega_k > 0$, where ω_k are the pole positions of $\rho(\omega)$ in the upper half plane, the right-hand side vanishes asymptotically ($x \to \infty$). Therefore,

$$I_z + I_{C_r} + I_{C_R} = 0, \qquad \text{as } x \to \infty. \tag{14.299}$$

The integral over the large semicircle is

$$I_{C_R} = \int_{C_R} d\omega\, \frac{e^{ix(\omega-\omega_0)}}{\omega-\omega_0}\, \rho(\omega) \overset{\omega=\omega_0+Re^{i\theta}}{=} i \int_0^{\pi} d\theta\, e^{ixRe^{i\theta}}\, \rho\left(\omega_0 + Re^{i\theta}\right). \tag{14.300}$$

If $\rho(Re^{i\theta})$ vanishes for $R \to \infty$, as required for a proper density, then

$$\boxed{I_{C_R} = 0.} \tag{14.301}$$

Then

$$\begin{aligned} I_z = -I_{C_r} &= -i \int_{\pi}^{0} d\theta\, e^{ixre^{i\theta}}\, \rho\left(\omega_0 + re^{i\theta}\right) \\ &\simeq -i \int_{\pi}^{0} d\theta\, \left[1 + O(r)\right] \left[\rho(\omega_0) + r\rho'(\omega_0) + \dots\right] \\ &\overset{r\to 0}{=} +i\pi\rho(\omega_0). \end{aligned} \tag{14.302}$$

Thus,

$$I_z = \int_{-\infty}^{+\infty} d\omega\, \frac{e^{ix(\omega-\omega_0)}}{\omega-\omega_0}\, \rho(\omega) = i\pi\rho(\omega_0) + O\left(\frac{1}{x}\right). \tag{14.303}$$

Taking the imaginary part, $I = \mathrm{Im}(I_z)$ (for real ρ), we find

$$\lim_{x\to\infty} \int_{-\infty}^{+\infty} d\omega\, \frac{\sin\left(x(\omega-\omega_0)\right)}{\omega-\omega_0}\, \rho(\omega) = \pi\rho(\omega_0), \tag{14.304}$$

and hence

$$\boxed{\lim_{x\to\infty} \frac{\sin[x(\omega-\omega_0)]}{\omega-\omega_0} = \pi\,\delta(\omega-\omega_0).} \tag{14.305}$$

Remark: The proof assumes that ρ is real on the real axis and extends to a meromorphic function in the upper half plane that vanishes at infinity (true for the Lorentzian of Exercise 1); I_z is understood as a principal-value integral, which is why the indentation C_r appears. The result holds for any smooth integrable test function: write $\rho(\omega) = \rho(\omega_0) + [\rho(\omega) - \rho(\omega_0)]$; the first term gives $\rho(\omega_0)\int_{-\infty}^{\infty} du\ \sin(u)/u = \pi\rho(\omega_0)$ and the second vanishes as $x \to \infty$ by the Riemann–Lebesgue lemma. The limit is used in Eq. (14.270) with $x = E_i\Delta t/\hbar$.

14.5.6 Harmonic Perturbation

As a further example, let us look at the perturbed Hamiltonian

$$H = H_0 + \theta(t)V(t) \quad \text{with} \quad V(t) = Ve^{i\omega t} + V^\dagger e^{-i\omega t}. \tag{14.306}$$

The operators V, $V^\dagger$ are time-independent, but they may depend on position, momentum, spin, angular momentum or other operators. The transition amplitude is

$$M_{i\to f} = \delta_{if} - \frac{i}{\hbar}\int_0^t dt'\, e^{i\omega_{fi}t'}\,\langle f|V(t')|i\rangle\,. \tag{14.307}$$

Let us compute one of the two integrals for $|f\rangle \neq |i\rangle$:

$$\int_0^t dt' e^{i(\omega_{fi}+\omega)t'}\langle f|V|i\rangle = i\,\frac{1-e^{i(\omega+\omega_{fi})t}}{\omega+\omega_{fi}}\,\langle f|V|i\rangle\,.$$

Considering both terms, we get

$$M_{i\to f} = \frac{1}{\hbar}\left[\frac{1-e^{i(\omega+\omega_{fi})t}}{\omega+\omega_{fi}}\,\langle f|V|i\rangle + \frac{1-e^{i(-\omega+\omega_{fi})t}}{-\omega+\omega_{fi}}\,\langle f|V^\dagger|i\rangle\right]. \tag{14.308}$$

The coefficients

$$\frac{1-e^{i(\omega+\omega_{fi})t}}{\omega+\omega_{fi}} \quad \text{and} \quad \frac{1-e^{i(-\omega+\omega_{fi})t}}{-\omega+\omega_{fi}}$$

are maximal for

$$\omega_{fi} = -\omega \quad \text{and} \quad \omega_{fi} = +\omega, \tag{14.309}$$

respectively. Since $\omega_{fi} = (E_f - E_i)/\hbar$, these cases correspond to

$$E_f = E_i - \hbar\omega \text{ and } E_f = E_i + \hbar\omega. \tag{14.310}$$

Repeating the steps of the previous section, we obtain

$$\frac{dP_{i\to\{f\}}}{dt} = \int dE_f\,\rho(E_f)\Gamma_{i\to f}, \tag{14.311}$$

with

$$\begin{aligned}\Gamma_{i\to f} = {}& \frac{2\pi}{\hbar}\overline{\left|\langle f|V|i\rangle\right|^2}\delta(E_f - E_i + \hbar\omega) \\ & + \frac{2\pi}{\hbar}\overline{\left|\langle f|V^\dagger|i\rangle\right|^2}\delta(E_f - E_i - \hbar\omega).\end{aligned} \tag{14.312}$$

The potential we describe here has the time dependence of an electromagnetic wave. Thus, we can use the above formulae to describe the interaction of radiation with atoms. We defer this discussion until we have learnt how to quantise the electromagnetic field and introduced photons.

But already now we can observe that an "incoming" electromagnetic wave with frequency ω can excite an atom from an energy level E_i to an energy level $E_i + \hbar\omega$,

$$\underset{\text{"photon absorption"}}{\underline{\text{Excitation:}}} \qquad E_i \to E_f = E_i + \hbar\omega. \tag{14.313}$$

An "outgoing" electromagnetic wave with frequency ω can be produced by the de-excitation of an atom,

$$\underset{\text{"photon emission"}}{\underline{\text{De-excitation:}}} \qquad E_i \to E_f = E_i - \hbar\omega. \tag{14.314}$$

14.5.7 Ionisation

Consider hydrogen atoms in their ground state $|0\rangle \equiv |1,0,0\rangle$. The wave function of a hydrogen atom in its ground state is

$$\psi_0(\mathbf{x}) = \frac{1}{\sqrt{\pi a_{\text{Bohr}}^3}} e^{-r/a_{\text{Bohr}}}, \tag{14.315}$$

with

$$a_{\text{Bohr}} = \frac{\hbar^2}{m_e e^2} = 0.529 \times 10^{-10}\,\text{m}. \tag{14.316}$$

Let us put the atoms in an electromagnetic field,

$$\mathbf{E}(t) = \mathbf{E} e^{-i\omega t} + \mathbf{E}^* e^{i\omega t}. \tag{14.317}$$

We ignore the magnetic field because its effect is suppressed by v/c. Furthermore, we ignore the spatial variation of the field, assuming a wavelength much larger than the Bohr radius. The potential energy due to the field is then

$$V(t) = e\mathbf{E}(t) \cdot \mathbf{x} = e\,(\mathbf{E} \cdot \mathbf{x})\, e^{-i\omega t} + e\,(\mathbf{E}^* \cdot \mathbf{x})\, e^{i\omega t}. \tag{14.318}$$

According to Fermi's golden rule, the rate at which the ground state of the atom is depopulated is

$$\Gamma_i = \frac{1}{\hbar} \sum_f 2\pi \left| \langle f | e\mathbf{E} \cdot \mathbf{x} | \psi_0 \rangle \right|^2 \cdot \delta(E_f - E_0 - \hbar\omega). \tag{14.319}$$

The first term of Eq. (14.312) (corresponding to de-excitation in Eq. (14.314)) does not contribute, as the atoms are in the ground state. We specialise to final states in which the atom is ionised and its electron is free (or sufficiently close to free). There is a continuum of free electron states with definite momentum,

$$H_{\text{free}} |\mathbf{p}\rangle = \frac{\mathbf{p}^2}{2m} |\mathbf{p}\rangle, \tag{14.320}$$

with

$$\langle \mathbf{x} | \mathbf{p} \rangle = \frac{1}{(2\pi\hbar)^{3/2}} e^{+i\frac{\mathbf{p}\cdot\mathbf{x}}{\hbar}} \tag{14.321}$$

and

$$\begin{aligned}\langle \mathbf{p}|\mathbf{p}'\rangle = \int d^3\mathbf{x}\, \langle \mathbf{p}|\mathbf{x}\rangle \langle \mathbf{x}|\mathbf{p}'\rangle &= \frac{1}{(2\pi\hbar)^3}\int d^3\mathbf{x}\, e^{-i\frac{\mathbf{x}\cdot(\mathbf{p}-\mathbf{p}')}{\hbar}} \\ &= \frac{1}{(2\pi\hbar)^3}(2\pi)^3\delta^{(3)}\left(\frac{\mathbf{p}-\mathbf{p}'}{\hbar}\right) = \delta^{(3)}(\mathbf{p}-\mathbf{p}').\end{aligned} \tag{14.322}$$

We now have to "sum" over all possible states $|\mathbf{p}\rangle$.

The Density of Final States

With the normalisation of Eq. (14.322) the completeness relation of the free states reads $\int d^3\mathbf{p}\ |\mathbf{p}\rangle\langle\mathbf{p}| = 1$, and the sum over final states becomes an integral over momenta. Let us write it in spherical coordinates, with Ω the direction of $\mathbf{p}$, and trade the length $p = |\mathbf{p}|$ for the energy $E_f = p^2/(2m)$, $dE_f = p\, dp/m$:

$$\sum_f \longrightarrow \int d^3\mathbf{p} = \int p^2 dp\, d\Omega = \int dE_f\, d\Omega\, \rho(E_f), \quad \rho(E_f) = m\, p = m\sqrt{2mE_f}. \tag{14.323}$$

This is the density of states of Eq. (14.271), now per unit solid angle of the outgoing electron. Notice that it depends on our choice of normalisation of the states. If we put the atom in a large box of volume L^3 with periodic boundary conditions, the momenta are discrete, $\mathbf{p} = 2\pi\hbar\,\mathbf{n}/L$, the states are normalised to one, $\langle\mathbf{x}|\mathbf{p}\rangle = L^{-3/2}e^{i\mathbf{p}\cdot\mathbf{x}/\hbar}$, and $\sum_{\mathbf{n}} \to (L/2\pi\hbar)^3\int d^3\mathbf{p}$. The density of states acquires a factor $L^3/(2\pi\hbar)^3$ and the squared matrix element a factor $(2\pi\hbar)^3/L^3$. Physical rates do not depend on L, as they should.

Let us take a linearly polarised wave, $\mathbf{E} = \mathcal{E}\,\boldsymbol{\epsilon}$, with $\mathcal{E}$ real and $\boldsymbol{\epsilon}$ a real unit vector. The ground-state energy is $E_0 = -I$, with $I = e^2/(2a_{\text{Bohr}}) = 13.6\,\text{eV}$ the ionisation energy. Inserting Eq. (14.323) into Eq. (14.319), the delta function is used up by the energy integration and we obtain the rate for the electron to be emitted into the solid angle $d\Omega$,

$$\frac{d\Gamma}{d\Omega} = \frac{2\pi}{\hbar}\, m\, p\, e^2\mathcal{E}^2\left|\langle\mathbf{p}|\boldsymbol{\epsilon}\cdot\mathbf{x}|\psi_0\rangle\right|^2, \qquad \frac{p^2}{2m} = \hbar\omega - I. \tag{14.324}$$

The atom is ionised only if $\hbar\omega > I$, i.e. for ultraviolet light with wavelength shorter than 91 nm. This is the **photoelectric effect** on a single atom.

Plane Waves as Final States

We have been a little too quick. The plane waves $|\mathbf{p}\rangle$ are eigenstates of the free Hamiltonian, but the ionised electron still moves in the Coulomb field of the proton. The true final states are eigenstates of $H_0 = \mathbf{p}^2/(2m) - e^2/r$ with positive energy, the Coulomb scattering states. When can we replace them by plane waves? The matrix element is dominated by distances $r \lesssim a_{\text{Bohr}}$, where the ground state lives. There the potential energy is of the order of I. If the kinetic energy of the outgoing electron is much larger,

$$\frac{p^2}{2m} \gg I = 13.6\,\text{eV} \quad \rightsquigarrow \quad k\, a_{\text{Bohr}} \gg 1, \qquad \mathbf{k} \equiv \frac{\mathbf{p}}{\hbar}, \tag{14.325}$$

the Coulomb potential distorts the plane wave only little. This is the same approximation as the **Born approximation** of Section 15.6: the plane wave is the zeroth-order term of the Lippmann–Schwinger equation for the outgoing electron.

However, there is a subtlety. For exact eigenstates of H_0 we can write the dipole matrix element in two equivalent ways. Indeed, from $[\mathbf{x}, H_0] = i\hbar\,\mathbf{p}/m$ (we will use the same identity for the matrix elements of radiative transitions in Eq. (16.49)) we find

$$\langle f|\mathbf{x}|\psi_0\rangle = \frac{-i}{m\,\omega_{f0}}\,\langle f|\mathbf{p}|\psi_0\rangle\,, \qquad \omega_{f0} = \frac{E_f - E_0}{\hbar} = \omega. \tag{14.326}$$

A plane wave is not an eigenstate of H_0, and once we approximate $\langle f|$ by $\langle \mathbf{p}|$ the two sides of Eq. (14.326) are no longer equal. The plane waves are not even orthogonal to the bound state, $\langle \mathbf{p}|\psi_0\rangle \neq 0$. Which form should we trust? The position operator weighs the large distances, where the long-range Coulomb potential distorts the outgoing wave most. The momentum operator weighs the short distances, where the electron must have taken its large momentum from the proton. We therefore use the momentum (*velocity*) form, which is also what the coupling $\mathbf{A}\cdot\mathbf{p}$ of Chapter 16 gives directly. We will confirm this choice at the end against the exact result.

Exercise: *Repeat the calculation below with the position (length) form,*

$$\langle \mathbf{p}|\mathbf{x}|\psi_0\rangle = i\hbar\nabla_{\mathbf{p}}\,\tilde{\psi}_0(\mathbf{p}),$$

and show that it gives an amplitude twice as large as the velocity form, i.e. a cross section four times larger, at all energies.

The Matrix Element

In the velocity form we need

$$\langle \mathbf{p}|\hat{\mathbf{p}}|\psi_0\rangle = \mathbf{p}\,\langle \mathbf{p}|\psi_0\rangle = \mathbf{p}\,\tilde{\psi}_0(\mathbf{p}), \qquad \tilde{\psi}_0(\mathbf{p}) = \int \frac{d^3\mathbf{x}}{(2\pi\hbar)^{3/2}}\,e^{-i\mathbf{k}\cdot\mathbf{x}}\,\psi_0(\mathbf{x}), \tag{14.327}$$

which is the Fourier transform of the ground state, i.e. its wave function in momentum space. Choosing the z-axis along $\mathbf{k}$, the angular integration is elementary, $\int d\Omega_{\mathbf{x}}\,e^{-ikr\cos\theta} = 4\pi\sin(kr)/(kr)$, and we have

$$\begin{aligned}
\tilde{\psi}_0(\mathbf{p}) &= \frac{1}{(2\pi\hbar)^{3/2}}\,\frac{1}{\sqrt{\pi a_{\text{Bohr}}^3}}\,\frac{4\pi}{k}\int_0^\infty dr\,r\sin(kr)\,e^{-r/a_{\text{Bohr}}} \\
&= \frac{1}{(2\pi\hbar)^{3/2}}\,\frac{1}{\sqrt{\pi a_{\text{Bohr}}^3}}\,\frac{4\pi}{k}\,\frac{2k\,a_{\text{Bohr}}^3}{\left(1+k^2a_{\text{Bohr}}^2\right)^2} \\
&= \frac{2\sqrt{2}}{\pi}\left(\frac{a_{\text{Bohr}}}{\hbar}\right)^{3/2}\frac{1}{\left(1+k^2a_{\text{Bohr}}^2\right)^2},
\end{aligned} \tag{14.328}$$

where we have used $\int_0^\infty dr\,r\sin(kr)e^{-r/a} = \text{Im}\left[(1/a - ik)^{-2}\right]$. The momentum distribution of the electron in the ground state falls off as k^{-8}. A fast electron is found in the atom only with a tiny probability, and this will make ionisation by energetic photons rare.

Exercise: *Verify that* $\int d^3\mathbf{p}\,|\tilde{\psi}_0(\mathbf{p})|^2 = 1$, *and compute* $\langle \mathbf{p}^2\rangle$ *with* $\tilde{\psi}_0$. *Compare with the virial theorem,* $\langle \mathbf{p}^2\rangle/(2m) = -E_0$.

Angular Distribution and Cross Section

Inserting Eqs. (14.326)–(14.328) into Eq. (14.324) we find

$$\frac{d\Gamma}{d\Omega} = \frac{2\pi}{\hbar}\,\frac{e^2\mathcal{E}^2}{m\,\omega^2}\,p\,(\boldsymbol{\epsilon}\cdot\mathbf{p})^2\left|\tilde{\psi}_0(\mathbf{p})\right|^2 . \tag{14.329}$$

As in Section 16.2.1, it is natural to divide the rate by the flux of the incoming photons. The real electric field of the wave is $2\mathcal{E}\boldsymbol{\epsilon}\cos\omega t$, its time-averaged energy flux (Poynting vector) is $S = \frac{c}{4\pi}\overline{\mathbf{E}(t)^2} = c\,\mathcal{E}^2/(2\pi)$, and the photon flux is $S/(\hbar\omega)$. The differential cross section, $d\sigma/d\Omega = \hbar\omega\,(d\Gamma/d\Omega)/S$, becomes

$$\frac{d\sigma}{d\Omega} = 32\,\frac{e^2}{mc\,\omega}\,a_{\text{Bohr}}^3\,\frac{k\,(\boldsymbol{\epsilon}\cdot\mathbf{k})^2}{\left(1+k^2a_{\text{Bohr}}^2\right)^4}. \tag{14.330}$$

This is the dipole cross section of Eq. (16.51), with the sum over n replaced by the integral over the continuum and the matrix element of $\mathbf{x}$ traded for that of $\mathbf{p}$ with Eq. (14.326). With $\boldsymbol{\epsilon}\cdot\mathbf{k} = k\cos\theta$, the electrons are emitted preferentially along the polarisation of the light, with a $\cos^2\theta$ distribution, and never perpendicular to it. They do **not** prefer the direction of the incoming photon! The electric field shakes the electron along $\boldsymbol{\epsilon}$, and the electron leaves in this direction.[1]

The angular integral is

$$\int d\Omega\,(\boldsymbol{\epsilon}\cdot\mathbf{k})^2 = \frac{4\pi k^2}{3}.$$

Using $e^2 = \alpha\hbar c$, $I = \hbar^2/(2ma_{\text{Bohr}}^2)$ and energy conservation, $\hbar\omega = I + \hbar^2k^2/(2m) = I\,(1+k^2a_{\text{Bohr}}^2)$, we obtain the total cross section

$$\begin{aligned}\sigma(\omega) &= \frac{128\pi}{3}\,\frac{\alpha\hbar}{m\,\omega}\,\frac{(ka_{\text{Bohr}})^3}{\left(1+k^2a_{\text{Bohr}}^2\right)^4} = \frac{256\pi}{3}\,\alpha\,a_{\text{Bohr}}^2\,\frac{(ka_{\text{Bohr}})^3}{\left(1+k^2a_{\text{Bohr}}^2\right)^5}\\ &= \frac{256\pi}{3}\,\alpha\,a_{\text{Bohr}}^2\left(\frac{I}{\hbar\omega}\right)^{7/2}\left(1-\frac{I}{\hbar\omega}\right)^{3/2}.\end{aligned} \tag{14.331}$$

Our approximation is valid for $\hbar\omega \gg I$, where the last factor is one. The final result reads

$$\sigma(\omega) \simeq \frac{256\pi}{3}\,\alpha\,a_{\text{Bohr}}^2\left(\frac{I}{\hbar\omega}\right)^{7/2} = 5.5\times10^{-17}\,\text{cm}^2\left(\frac{13.6\,\text{eV}}{\hbar\omega}\right)^{7/2}, \qquad \hbar\omega \gg I. \tag{14.332}$$

The cross section is of the order of the geometric size of the atom, $\pi a_{\text{Bohr}}^2 = 8.8\times10^{-17}\,\text{cm}^2$, times α and times the steep suppression $(\omega_0/\omega)^{7/2}$, with $\hbar\omega_0 = I$. For $\hbar\omega = 1\,\text{keV}$ it is already only $1.6\times10^{-23}\,\text{cm}^2$. Physically, the photon must transfer its large momentum to an electron which hardly has any high-momentum components in the atom, Eq. (14.328). For a hydrogen-like ion with nuclear charge Z we replace $a_{\text{Bohr}} \to a_{\text{Bohr}}/Z$ and $I \to Z^2I$, and $\sigma \propto Z^5$ at fixed frequency. This is why X-rays are absorbed mostly by the inner electrons of heavy atoms.

[1] Beyond the dipole approximation, i.e. keeping $e^{i\mathbf{k}_\gamma\cdot\mathbf{x}}$ in the coupling of Chapter 16, $\mathbf{k}$ in the denominator of Eq. (14.330) is replaced by $\mathbf{k}-\mathbf{k}_\gamma$, and the distribution is tilted forward, towards the direction $\mathbf{k}_\gamma$ of the photon.

How high can we go in frequency? We have used the dipole approximation, which requires that the photon momentum is small compared to the momentum of the electron, $\hbar\omega/c \ll \hbar k$, i.e. $\hbar\omega \ll 2mc^2 \approx 1\,\mathrm{MeV}$, and a non-relativistic electron. Eq. (14.332) holds therefore in the window $13.6\,\mathrm{eV} \ll \hbar\omega \ll 100\,\mathrm{keV}$.

Remark: The Threshold

Near threshold, $\hbar\omega \to I$, Eq. (14.331) vanishes as $(\hbar\omega - I)^{3/2}$. This is wrong. A slow electron is strongly attracted by the proton, and the plane wave is a bad approximation for it. The calculation with the exact Coulomb scattering states of the continuum can be done in closed form (Stobbe, 1930; see Bethe and Salpeter, *Quantum Mechanics of One- and Two-Electron Atoms*, or Landau-Lifshitz Vol. IV). With $\eta = 1/(k a_{\mathrm{Bohr}}) = \sqrt{I/(\hbar\omega - I)}$ it reads

$$\sigma(\omega) = \frac{2^9\pi^2}{3}\,\alpha\, a_{\mathrm{Bohr}}^2 \left(\frac{I}{\hbar\omega}\right)^4 \frac{e^{-4\eta\,\mathrm{arccot}\,\eta}}{1 - e^{-2\pi\eta}}. \tag{14.333}$$

At threshold, $\eta \to \infty$ and $\eta\,\mathrm{arccot}\,\eta \to 1$, so that the cross section is finite and maximal,

$$\sigma(\omega_0) = \frac{2^9\pi^2}{3\,\mathrm{e}^4}\,\alpha\, a_{\mathrm{Bohr}}^2 \approx 0.225\, a_{\mathrm{Bohr}}^2 \approx 6.3 \times 10^{-18}\,\mathrm{cm}^2, \tag{14.334}$$

where $\mathrm{e} = 2.718\ldots$ is Euler's number. For $\eta \to 0$ we have $\mathrm{arccot}\,\eta \to \pi/2$ and $1 - e^{-2\pi\eta} \to 2\pi\eta$, and Eq. (14.333) reduces exactly to Eq. (14.332). This confirms our choice of the velocity form; the length form would be wrong by a factor of four. Notice, however, that the approach is slow: the Born result of Eq. (14.331) overestimates the exact cross section by a factor $\approx 1 + \pi\eta$, i.e. by 10% at $\hbar\omega \approx 1000\,I$ and by a factor of two at $\hbar\omega = 10\,I$. The parameter $\eta = e^2/(\hbar v)$, with v the velocity of the outgoing electron, measures the strength of the Coulomb interaction compared to the kinetic energy, and the Born approximation is an expansion in it.

15

Theory of Elastic Scattering

Scattering is a process in which moving particles interact, either with one another or with a potential, so that their trajectories are changed. Scattering experiments use this phenomenon to determine properties of the particles and of their interactions. A famous example is Rutherford's scattering of alpha particles off gold nuclei. It led to the discovery that almost all of the mass of an atom is concentrated in a nucleus that occupies a tiny fraction of the atomic volume. Other examples are the optical phenomena of Rayleigh scattering and Thomson scattering.

In this chapter we analyse elastic scattering, a process in which two particles collide and bounce off each other without any loss of the total kinetic energy. We shall not consider inelastic scattering.

15.1 An Introduction to Elastic Scattering Experiments

Scattering experiments are the best method we know to reveal the properties of matter and of particle interactions at short distances. Let us consider a particle (in a beam of particles) which moves towards a target at the centre of our coordinate system, as in Fig. 15.1.

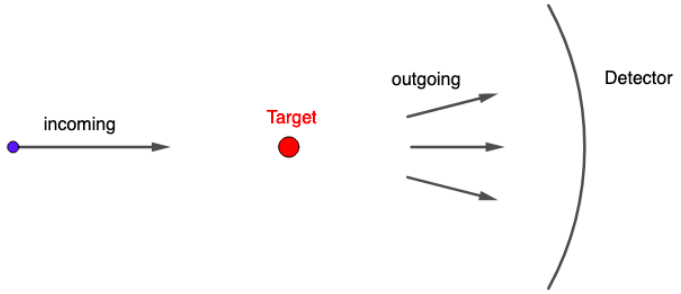


Figure 15.1. Set-up of an elastic scattering experiment: an incoming beam is scattered by a target at the origin.

The Hamiltonian of the system is

$$H = \frac{\mathbf{p}^2}{2M} + V(\mathbf{r}), \tag{15.1}$$

where $\frac{\mathbf{p}^2}{2M}$ is the kinetic energy of the particle and $V(\mathbf{r})$ describes the interaction with the target. Let us assume that the target is small. It is then reasonable to assume that, far away

from the target, its effect on the incoming particle is negligible. Formally, we require

$$V(\mathbf{r}) \to 0 \quad \text{as} \quad r \to \infty. \tag{15.2}$$

With the Hamiltonian of Eq. (15.1), elastic scattering becomes a problem of time evolution of a quantum-mechanical system, governed by the Schrödinger equation. Let us denote a general state of the system by $|\psi, t\rangle$. The Schrödinger equation for this state is

$$i\hbar \frac{\partial}{\partial t} |\psi, t\rangle = H |\psi, t\rangle . \tag{15.3}$$

We can expand states in the basis of energy eigenstates of the Hamiltonian, $H|E\rangle = E|E\rangle$. The time evolution of such an eigenstate is $|E, t\rangle = e^{-i\frac{E}{\hbar}t} |E\rangle$.

The most general solution is a wave packet, a superposition of energy eigenstates,

$$|\psi, t\rangle = \sum_{|E\rangle} g(E) e^{-i\frac{E}{\hbar}t} |E\rangle , \tag{15.4}$$

where the sum is over all possible energy eigenstates and $g(E)$ are arbitrary coefficients. Recall that, in general, the spectrum of the Hamiltonian has both a discrete and a continuous part. The sum over all energy eigenstates, denoted by "$\sum_{|E\rangle}$", thus consists of a sum over the discrete spectrum (bound states) and an integral over the continuum of energies. We keep writing it as a sum to avoid clutter in the notation.

We denote the kinetic-energy term of the Hamiltonian by

$$H_0 \equiv \frac{\mathbf{p}^2}{2M}. \tag{15.5}$$

The eigenvalue problem for the energy eigenstates can then be cast into the form

$$(E - H_0)|E\rangle = V|E\rangle . \tag{15.6}$$

This form of the eigenvalue problem is very useful: it allows us to treat the problem perturbatively for most target potentials $V(\mathbf{r})$.

15.2 The Lippmann–Schwinger Equation

To treat the eigenvalue problem of Eq. (15.6) perturbatively, we need an expression that can be expanded in powers of V. Before deriving it, we establish some notation.

Eq. (15.6) contains three operators, $\hat{E}$, $\hat{H}_0$ and $\hat{V}$. To solve for $|E\rangle$, we would like to invert the operator $(\hat{E} - \hat{H}_0)$. However, as we show below, this operator has at least one vanishing eigenvalue (a zero mode), so that $\det(\hat{E} - \hat{H}_0) = 0$ and it cannot be inverted.

The Lippmann–Schwinger equation circumvents this problem by introducing an infinitesimal quantity that makes the operator invertible. In the rest of the chapter we do not denote operators with a hat, unless this is necessary.

The momentum operator has a continuous spectrum of eigenvalues $\mathbf{q}$, with eigenstates $|\mathbf{p}\rangle$:

$$\hat{\mathbf{p}} |\mathbf{p}\rangle = \mathbf{q} |\mathbf{p}\rangle . \tag{15.7}$$

Since H_0 is proportional to the square of the momentum operator, it has a continuous spectrum of eigenvalues $\mathbf{q}^2/2M$ with the same eigenstates $|\mathbf{p}\rangle$:

$$H_0 = \frac{\mathbf{p}^2}{2M} \quad \text{and} \quad H_0 |\mathbf{p}\rangle = \frac{\mathbf{q}^2}{2M} |\mathbf{p}\rangle . \tag{15.8}$$

Therefore, the operator $(E - H_0)$ has a zero eigenvalue,

$$E - \frac{\mathbf{q}^2}{2M} = 0 \quad \text{for} \quad |\mathbf{q}| = \sqrt{2ME}, \tag{15.9}$$

so that it is singular and cannot be inverted. To circumvent this issue, we consider the modified operator

$$(E - H_0 + i\delta) , \tag{15.10}$$

where δ is an infinitesimal positive real number ($\delta \to 0^+$). This modified operator can be inverted, since its eigenvalues acquire an imaginary part, while the eigenvalues of H_0 are real.

The eigenvalue problem of Eq. (15.6) now reads

$$(E - H_0 + i\delta) |E\rangle = V |E\rangle . \tag{15.11}$$

Since the operator on the left-hand side of Eq. (15.11) is invertible, we can multiply the equation by its inverse from the left and obtain

$$|E\rangle = (E - H_0 + i\delta)^{-1} V |E\rangle . \tag{15.12}$$

We also have to consider the equation in the absence of a scattering potential V. In this case, the energy eigenstates are eigenstates of the kinetic-energy Hamiltonian, $H_0 |E'\rangle = E' |E'\rangle$. We denote these free eigenstates by $|E_0\rangle$. Eq. (15.11) then reads

$$(E - H_0 + i\delta) |E\rangle = 0. \tag{15.13}$$

For $\delta \to 0^+$, its solution is a free eigenstate,

$$|E\rangle = |E_0\rangle . \tag{15.14}$$

We must remember that the eigenvalue problem of the Schrödinger equation is a differential equation. Its solutions therefore consist of a *particular* and a *homogeneous* part. Inverting the differential operator, as in Eq. (15.12), gives the particular solution only; the homogeneous solution is the one of Eq. (15.14). Combining Eqs. (15.12) and (15.14), we obtain a general expression for the energy eigenstate,

Lippmann–Schwinger equation

$$|E\rangle = |E_0\rangle + (E - H_0 + i\delta)^{-1} V |E\rangle . \tag{15.15}$$

This is the Lippmann–Schwinger equation. It was derived jointly by the theoretical physicists Bernard Lippmann and Julian Schwinger in 1950.

The Lippmann–Schwinger equation gives a general expression for the solution of a scattering problem. We now check that it makes sense physically.

Any state $|\psi, t\rangle$ can be expressed as a superposition of energy eigenstates,

$$|\psi, t\rangle = \sum_{|E\rangle} g(E) e^{-iEt/\hbar} |E\rangle . \tag{15.16}$$

Inserting Eq. (15.15) into Eq. (15.16), we obtain

$$\begin{aligned} |\psi, t\rangle &= \sum_{|E\rangle} g(E) e^{-iEt/\hbar} |E_0\rangle + \sum_{|E\rangle} g(E) e^{-iEt/\hbar} (E - H_0 + i\delta)^{-1} V |E\rangle \\ &= \sum_{|E\rangle} g(E) e^{-iEt/\hbar} |E_0\rangle + \sum_{|E\rangle} g(E) e^{-iEt/\hbar} \underbrace{\int dE_0' \, |E_0'\rangle \langle E_0'|}_{1} (E - H_0 + i\delta)^{-1} V |E\rangle . \end{aligned} \tag{15.17}$$

In the last step we have inserted an integral over the projectors onto the eigenstates of the free Hamiltonian H_0, which is the identity operator (completeness of the free eigenstates). We further assume that the relevant states $|E\rangle$ lie in the continuous spectrum and write the sum as an integral. We obtain

$$|\psi, t\rangle = \sum_{|E\rangle} g(E) e^{-iEt/\hbar} |E_0\rangle + \int dE \int dE_0' \, g(E) \frac{e^{-iEt/\hbar} |E_0'\rangle \langle E_0'| V |E\rangle}{E - E_0' + i\delta} . \tag{15.18}$$

Here we used that H_0 is diagonal in the basis $|E_0'\rangle$, with $H_0 |E_0'\rangle = E_0' |E_0'\rangle$. Let us assume that $g(E)$ is sharply peaked around some energy well above zero (Fig. 15.2).

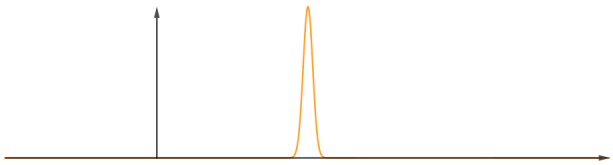


Figure 15.2. The energy distribution $g(E)$ of the incoming wave packet, sharply peaked around a positive energy.

This assumption is motivated by the fact that the particles in the beam can be prepared experimentally with a small energy spread. We can then extend the integration over E to

$$\int_{-\infty}^{+\infty} dE . \tag{15.19}$$

Let us evaluate the integral in the limit $t \to -\infty$, that is, at early times, before the particle has interacted with the target. The calculation will confirm this physical picture, as a consequence of our choice $\delta \to 0^+$ (rather than $\delta \to 0^-$). Indeed, for $t \to -\infty$ we can close the integration contour in the upper half of the complex plane (Fig. 15.3).

The contour does not enclose any pole, since the only pole, at $E = E_0' - i\delta$, lies in the lower half plane. Thus the integral vanishes. We now consider the wave packet built in the same way as in Eq. (15.16) from the non-interacting states $|E_0\rangle$,

$$|E_0\rangle_g = \sum_{|E\rangle} g(E) e^{-iEt/\hbar} |E_0\rangle, \tag{15.20}$$

and evaluate the difference of the two wave packets,

$$|\psi, t\rangle - |E_0\rangle_g = \sum_{|E\rangle} g(E) e^{-iEt/\hbar} (|E\rangle - |E_0\rangle). \tag{15.21}$$

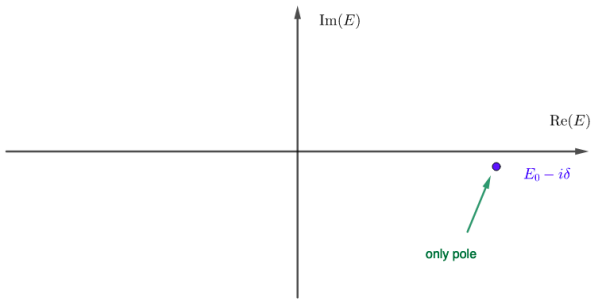


Figure 15.3. The pole of the integrand at $E = E_0' - i\delta$ in the complex E plane. For $t \to -\infty$ the contour is closed in the upper half plane and encloses no pole.

By the reasoning above, this difference vanishes at early times. Symbolically, we write

$$|E\rangle = |E_0\rangle + (E - H_0 + i\delta)^{-1} V |E\rangle \overset{t\to-\infty}{\to} |E_0\rangle . \tag{15.22}$$

Eq. (15.22) is a shorthand: a stationary state $|E\rangle$ has no time dependence; the statement is about the wave packets: $|\psi, t\rangle - |E_0\rangle_g \to 0$ for $t \to -\infty$, because the second term of Eq. (15.18) vanishes by the contour argument above. With $-i\delta$ instead one would obtain packets that become free for $t \to +\infty$ (incoming-wave solutions). See Sakurai, Modern QM, Ch. 7 (Ch. 6 in the 2nd ed.), and Gottfried–Yan, Ch. 8. The $i\delta$ prescription with $\delta \to 0^+$, which we introduced to invert the operator $(E - H_0)$, thus guarantees the physical requirement that the particle is free at early times, before the scattering takes place, when it has not yet interacted with the scattering potential.

15.3 Green's Functions

To make further progress, we write the Lippmann–Schwinger equation in position space. It then becomes a linear integral equation, which we solve by finding the corresponding Green's function. We multiply Eq. (15.15) by a position bra from the left:

$$\langle \mathbf{x}|E\rangle = \langle \mathbf{x}|E_0\rangle + \langle \mathbf{x}|\frac{1}{E - H_0 + i\delta} V|E\rangle , \tag{15.23}$$

which implies

$$\psi_E(\mathbf{x}) = \psi_E^{(0)}(\mathbf{x}) + \int \mathrm{d}^3\mathbf{x}' \, \langle \mathbf{x}|\frac{1}{E - H_0 + i\delta} V(\hat{\mathbf{x}})|\mathbf{x}'\rangle \langle \mathbf{x}'|E\rangle . \tag{15.24}$$

Here we have inserted the identity operator

$$\int \mathrm{d}^3\mathbf{x}' \, |\mathbf{x}'\rangle \langle \mathbf{x}'| = \mathbb{1}. \tag{15.25}$$

We now define the Green's function

$$G_\delta(\mathbf{x} - \mathbf{x}') = \langle \mathbf{x}|\frac{1}{E - H_0 + i\delta}|\mathbf{x}'\rangle , \tag{15.26}$$

which we will compute shortly. We also recall that

$$V(\hat{\mathbf{x}}) \, |\mathbf{x}\rangle = V(\mathbf{x}) \, |\mathbf{x}\rangle . \tag{15.27}$$

Using Eqs. (15.26) and (15.27), we can cast the Lippmann–Schwinger equation into the form

$$\boxed{\psi_E(\mathbf{x}) = \psi_E^{(0)}(\mathbf{x}) + \int \mathrm{d}^3\mathbf{x}'\, G_\delta(\mathbf{x}-\mathbf{x}')V(\mathbf{x}')\psi_E(\mathbf{x}').} \tag{15.28}$$

We now compute the Green's function defined in Eq. (15.26); it is also crucial to check that the limit $\delta \to 0^+$ yields a physical state. Using $\hat{H}_0 = \frac{\hat{\mathbf{p}}^2}{2M}$, we have

$$\begin{aligned} G_\delta(\mathbf{x}-\mathbf{x}') &= \left\langle \mathbf{x} \middle| \frac{1}{E - \frac{\hat{\mathbf{p}}^2}{2M} + i\delta} \middle| \mathbf{x}' \right\rangle \\ &= \langle \mathbf{x} | \int \mathrm{d}^3\mathbf{k}\, |\mathbf{k}\rangle \left\langle \mathbf{k} \middle| \frac{1}{E - \frac{\hat{\mathbf{p}}^2}{2M} + i\delta} \middle| \mathbf{x}' \right\rangle, \end{aligned} \tag{15.29}$$

where we have introduced the momentum eigenkets

$$\hat{\mathbf{p}}\,|\mathbf{k}\rangle = \mathbf{k}\,|\mathbf{k}\rangle\,. \tag{15.30}$$

The Green's function simplifies to

$$G_\delta(\mathbf{x}-\mathbf{x}') = \int \mathrm{d}^3\mathbf{k} \frac{\langle \mathbf{x}|\mathbf{k}\rangle \langle \mathbf{k}|\mathbf{x}'\rangle}{E - \frac{\mathbf{k}^2}{2M} + i\delta}. \tag{15.31}$$

We recall that the inner product of a position bra and a momentum ket is

$$\langle \mathbf{x}|\mathbf{k}\rangle = \frac{1}{\sqrt{2\pi\hbar}^3} e^{\frac{i}{\hbar}\mathbf{k}\cdot\mathbf{x}}. \tag{15.32}$$

We thus obtain the integral representation of the Green's function

$$G_\delta(\mathbf{x}-\mathbf{x}') = \int_{-\infty}^{\infty} \frac{\mathrm{d}^3\mathbf{k}}{(2\pi\hbar)^3} \frac{e^{\frac{i}{\hbar}\mathbf{k}\cdot(\mathbf{x}-\mathbf{x}')}}{E - \frac{\mathbf{k}^2}{2M} + i\delta}. \tag{15.33}$$

The result of this integral in the limit $\delta \to 0$ is

$$G(\mathbf{x}-\mathbf{x}') = -\frac{2M}{\hbar^2}\frac{1}{4\pi}\frac{e^{\frac{i}{\hbar}\sqrt{2ME}|\mathbf{x}-\mathbf{x}'|}}{|\mathbf{x}-\mathbf{x}'|}, \tag{15.34}$$

as we show in Appendix B. Substituting Eq. (15.34) into Eq. (15.28), we obtain the Lippmann–Schwinger equation in position space,

$$\psi_E(\mathbf{x}) = \psi_E^{(0)}(\mathbf{x}) - \frac{2M}{\hbar^2}\int \frac{\mathrm{d}^3\mathbf{x}'}{4\pi} \frac{e^{\frac{i}{\hbar}\sqrt{2ME}|\mathbf{x}-\mathbf{x}'|}}{|\mathbf{x}-\mathbf{x}'|} V(\mathbf{x}')\psi_E(\mathbf{x}'). \tag{15.35}$$

The boundary condition that the particle is free at $t \to -\infty$ determines $\psi_E^{(0)}(\mathbf{x})$ to be a plane wave,

$$\psi_E^{(0)}(\mathbf{x}) = \langle \mathbf{x}|\mathbf{k}\rangle = \frac{e^{\frac{i}{\hbar}\mathbf{k}\cdot\mathbf{x}}}{\sqrt{2\pi\hbar}^3}. \tag{15.36}$$

Finally, we have

$$\psi_E(\mathbf{x}) = \frac{1}{\sqrt{2\pi\hbar}^3} \cdot e^{\frac{i}{\hbar}\mathbf{k}\cdot\mathbf{x}} - \frac{2M}{\hbar^2} \int \frac{\mathrm{d}^3\mathbf{x}'}{4\pi} \frac{e^{\frac{i}{\hbar}|\mathbf{k}||\mathbf{x}-\mathbf{x}'|}}{|\mathbf{x}-\mathbf{x}'|} V(\mathbf{x}')\psi_E(\mathbf{x}'). \tag{15.37}$$

15.4 Large-Distance Expansion

Eq. (15.37) gives an expression for the wave function of the energy eigenstate. The expression is implicit, since $\psi_E(\mathbf{x})$ also appears in the integral on the right-hand side, which makes it difficult to calculate the wave function explicitly. Nevertheless, the expression is very useful. In a scattering experiment, the detectors are placed far away from the target, so we can take the large-distance limit $|\mathbf{x}| \to \infty$ (Fig. 15.4).

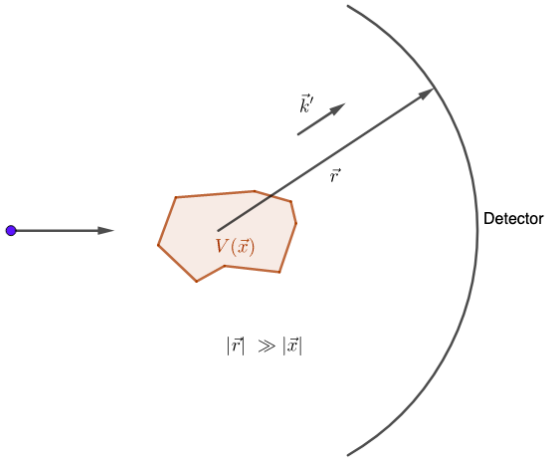


Figure 15.4. The detector at position $\mathbf{x}$ is far away from the region where the potential acts, $|\mathbf{x}| \gg |\mathbf{x}'|$.

For $|\mathbf{x}| \gg |\mathbf{x}'|$ we can expand the distance $|\mathbf{x}-\mathbf{x}'|$ appearing in Eq. (15.37):

$$\begin{aligned} |\mathbf{x}-\mathbf{x}'| &= \left(\mathbf{x}^2 + \mathbf{x}'^2 - 2\mathbf{x}\cdot\mathbf{x}'\right)^{\frac{1}{2}} \\ &= |\mathbf{x}| \left(1 - 2\frac{\mathbf{x}\cdot\mathbf{x}'}{|\mathbf{x}|^2} + \frac{\mathbf{x}'^2}{\mathbf{x}^2}\right)^{\frac{1}{2}} \\ &\approx |\mathbf{x}| \left(1 - \frac{\mathbf{e}_r\cdot\mathbf{x}'}{|\mathbf{x}|} + O\left(\frac{|\mathbf{x}'|^2}{|\mathbf{x}|^2}\right)\right). \end{aligned} \tag{15.38}$$

Expanding Eq. (15.37) in powers of $1/|\mathbf{x}|$, we obtain

$$\psi_E(\mathbf{x}) = \frac{1}{\sqrt{2\pi\hbar}^3} \left\{ e^{\frac{i}{\hbar}\mathbf{k}\cdot\mathbf{x}} + \frac{e^{\frac{i}{\hbar}|\mathbf{k}||\mathbf{x}|}}{|\mathbf{x}|} f(\mathbf{k},\mathbf{k}') \right\}, \tag{15.39}$$

where the function $f(\mathbf{k}, \mathbf{k}')$, the *scattering amplitude*, is

Scattering amplitude

$$f(\mathbf{k}, \mathbf{k}') = -\frac{2M}{\hbar^2} \int \frac{d^3\mathbf{x}'}{4\pi} e^{-\frac{i}{\hbar}\mathbf{k}'\cdot\mathbf{x}'} V(\mathbf{x}') \, (2\pi\hbar)^{3/2} \, \psi_E(\mathbf{x}') \tag{15.40}$$

with

$$\mathbf{k}' = |\mathbf{k}| \, \mathbf{e}_r = |\mathbf{k}| \, \frac{\mathbf{x}}{|\mathbf{x}|}. \tag{15.41}$$

The scattering amplitude is directly related to the cross-section, which can be measured experimentally. We introduce the cross-section in the next section. In what follows, we keep working at large distances.

15.5 Cross-Sections

In scattering experiments, we measure the number of particles passing through a surface element of a detector in a given time interval. We can also determine the properties of the incoming beam before the scattering. In particular, we can measure the flux $\mathbf{j}_{\text{incoming}}$ of particles moving towards the target (Fig. 15.5).

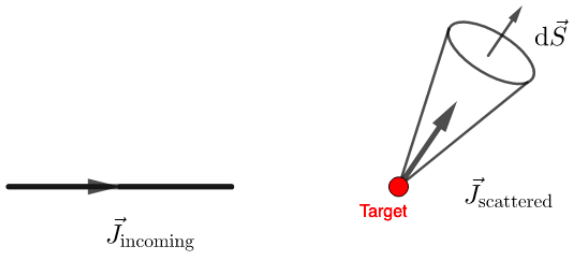


Figure 15.5. Incoming flux and the flux scattered through a surface element $d\mathbf{S}$ of the detector.

The *cross-section* is defined as the ratio of the rate of particles passing through a surface element $\mathrm{d}\mathbf{S}$ to the incoming flux,

$$d\sigma \equiv \frac{\mathbf{j}_{\text{scattered}} \cdot d\mathbf{S}}{\left|\mathbf{j}_{\text{incoming}}\right|}. \tag{15.42}$$

Intuitively, it measures the fraction of particles that are scattered through a given surface element. Using $d\mathbf{S} = \mathbf{e}_r r^2 d\Omega$, we obtain

$$d\sigma = \frac{\left(\mathbf{j}_{\text{scattered}} \cdot \mathbf{e}_r\right) r^2 d\Omega}{\left|\mathbf{j}_{\text{incoming}}\right|}. \tag{15.43}$$

The *differential cross-section* is the cross-section per solid angle,

$$\frac{d\sigma}{d\Omega} = r^2 \frac{\mathbf{j}_{\text{scattered}} \cdot \mathbf{e}_r}{\left|\mathbf{j}_{\text{incoming}}\right|}. \tag{15.44}$$

To determine $\mathbf{j}_{\text{scattered}}$ and $\mathbf{j}_{\text{incoming}}$, we differentiate the probability density with respect to time:

$$\begin{aligned}\frac{\partial}{\partial t}\rho(\mathbf{x}',t) &= \frac{\partial}{\partial t}\left|\psi(\mathbf{x}',t)\right|^2 = \frac{\partial}{\partial t}\left(\psi(\mathbf{x}',t)\psi^*(\mathbf{x}',t)\right) = \psi\frac{\partial\psi^*}{\partial t} + \psi^*\frac{\partial\psi}{\partial t} \\ &= \psi\frac{\left[-\frac{\hbar^2}{2M}\Delta\psi^* + V\psi^*\right]}{-i\hbar} + \psi^*\frac{\left[-\frac{\hbar^2}{2M}\Delta\psi + V\psi\right]}{i\hbar} = \underset{\text{(exercise)}}{\ldots} \\ &= -\nabla\cdot\text{Re}\left(\psi^*\frac{(-i\hbar\nabla)}{M}\psi\right).\end{aligned} \tag{15.45}$$

This is the continuity equation

$$\frac{\partial\rho}{\partial t} + \nabla\cdot\mathbf{j} = 0, \tag{15.46}$$

with the probability current

$$\mathbf{j} = \text{Re}\left(\psi^*\frac{(-i\hbar\nabla)}{M}\psi\right) = \frac{1}{M}\text{Re}\left(\langle\psi|\mathbf{x}\rangle\langle\mathbf{x}|\hat{\mathbf{p}}|\psi\rangle\right). \tag{15.47}$$

The quantity $\mathbf{j}$ is the total probability flux. In the absence of a scattering potential, $V = 0$, the total probability flux is the incoming flux,

$$\mathbf{j}_{\text{incoming}} = \mathbf{j}|_{V=0}\,. \tag{15.48}$$

Since $\mathbf{j}$ is the total probability flux, the scattered flux is

$$\mathbf{j}_{\text{scattered}} = \mathbf{j} - \mathbf{j}_{\text{incoming}}. \tag{15.49}$$

Using Eq. (15.47) and setting $M = 1$, we calculate the incoming flux of a free particle,

$$\mathbf{j}_{\text{incoming}} = \text{Re}\left(\frac{e^{-i\mathbf{k}\cdot\mathbf{x}/\hbar}}{(2\pi\hbar)^{3/2}}(-i\hbar\nabla)\frac{e^{+i\mathbf{k}\cdot\mathbf{x}/\hbar}}{(2\pi\hbar)^{3/2}}\right) = \frac{\mathbf{k}}{(2\pi\hbar)^3}. \tag{15.50}$$

To calculate $\mathbf{j}_{\text{scattered}}$, we first need $\mathbf{j}$, which is given by

$$\begin{aligned}\mathbf{j} &= \text{Re}\left(\psi^*(-i\hbar\nabla)\psi\right) \\ &= \text{Re}\left(\frac{1}{\sqrt{2\pi\hbar}^3}\left\{e^{-\frac{i}{\hbar}\mathbf{k}\cdot\mathbf{x}} + \frac{e^{-i|\mathbf{k}||\mathbf{x}|}}{|\mathbf{x}|}f^*(\mathbf{k},\mathbf{k}')\right\}\right. \\ &\qquad \left.(-i\hbar\nabla)\frac{1}{\sqrt{2\pi\hbar}^3}\left\{e^{\frac{i}{\hbar}\mathbf{k}\cdot\mathbf{x}} + \frac{e^{i|\mathbf{k}||\mathbf{x}|}}{|\mathbf{x}|}f(\mathbf{k},\mathbf{k}')\right\}\right),\end{aligned} \tag{15.51}$$

where we inserted Eq. (15.39) into Eq. (15.47). We first compute

$$-i\hbar\nabla\psi = -\frac{i\hbar\nabla}{\sqrt{2\pi\hbar}^3}\left(e^{\frac{i}{\hbar}\mathbf{k}\cdot\mathbf{x}} + \frac{e^{i|\mathbf{k}|r}}{r}f(\mathbf{k},\mathbf{k}')\right). \tag{15.52}$$

The scattering amplitude depends only on the scattering angle, through $\mathbf{e}_r$, and not on the distance. To compute the derivative, we use spherical coordinates, in which

$$\nabla = \frac{\partial}{\partial r}\mathbf{e}_r + \frac{1}{r}\hat{L}(\theta,\phi), \tag{15.53}$$

where $\hat{L}$ contains the angular derivatives and does not depend on r. Since we are interested only in the terms that dominate at large distances, we can approximate

$$\nabla = \frac{\partial}{\partial r}\mathbf{e}_r + O\left(\frac{1}{r}\right). \tag{15.54}$$

With

$$\psi^* = \frac{1}{(2\pi\hbar)^{3/2}}\left[e^{-i\mathbf{k}\cdot\mathbf{x}/\hbar} + \frac{e^{-i|\mathbf{k}|r/\hbar}}{r} f^*(\mathbf{k},\mathbf{k}')\right] \tag{15.55}$$

we find

$$(-i\hbar\nabla)\,\psi = \frac{1}{(2\pi\hbar)^{3/2}}\left[\mathbf{k}e^{i\mathbf{k}\cdot\mathbf{x}/\hbar} + \mathbf{k}'\frac{e^{+i|\mathbf{k}|r/\hbar}}{r} f(\mathbf{k},\mathbf{k}') + \ldots\right] \tag{15.56}$$

where $\mathbf{k}' = |\mathbf{k}|\,\mathbf{e}_r$. Combining Eqs. (15.55) and (15.56) and taking the real part gives the total probability flux $\mathbf{j}$. We first combine the two expressions,

$$\begin{aligned}\psi^*\,(-i\hbar\nabla)\,\psi = \frac{1}{(2\pi\hbar)^3}\Bigg[&\mathbf{k} + \mathbf{k}'\frac{\left|f(\mathbf{k},\mathbf{k}')\right|^2}{r^2} \\ &+ \frac{e^{-i(\mathbf{k}\cdot\mathbf{x}-|\mathbf{k}|r)/\hbar}\mathbf{k}'\,f(\mathbf{k},\mathbf{k}')}{r} \\ &+ \frac{e^{i(\mathbf{k}\cdot\mathbf{x}-|\mathbf{k}|r)/\hbar}\mathbf{k}\,f^*(\mathbf{k},\mathbf{k}')}{r} + \ldots\Bigg]\end{aligned} \tag{15.57}$$

Here we have omitted terms which are suppressed as $O\left(1/r^3\right)$. The cross terms in the second and third lines contain the phase

$$e^{i(\mathbf{k}\cdot\mathbf{x}-|\mathbf{k}|r)/\hbar} = e^{-i|\mathbf{k}|r(1-\cos\theta)/\hbar}. \tag{15.58}$$

which oscillates rapidly as a function of the angle at large distances. For "wide" scattering angles, $\theta \neq 0$, the contribution of these terms to the flux integrated over a range of angles (even a small one) cancels. Ignoring them (the remaining terms are real), the total probability flux reduces to

$$\mathbf{j} = \frac{1}{(2\pi\hbar)^3}\left[\mathbf{k} + \mathbf{k}'\frac{\left|f(\mathbf{k},\mathbf{k}')\right|^2}{r^2} + \ldots\right]. \tag{15.59}$$

The interference terms dropped here do not average away in the forward direction $\theta \simeq 0$, where the phase in Eq. (15.58) is stationary. Integrated over a small cone around $\theta = 0$ they remove flux from the beam (the "shadow" of the target), and flux conservation then yields exactly the optical theorem of Sec. 15.7.1, $\sigma = \frac{4\pi\hbar}{|\mathbf{k}|}\,\mathrm{Im}\,f(\theta=0)$ in the present units. See Sakurai, Modern QM, Ch. 7 (Ch. 6 in the 2nd ed.); Gottfried–Yan, Ch. 8.

We can now calculate the scattered flux,

$$\mathbf{j}_{\text{scattered}} \approx \mathbf{j} - \mathbf{j}_{\text{incoming}}\Big|_{\substack{\text{wide angles}\\ \text{large distances}}} = \frac{1}{(2\pi\hbar)^3}\mathbf{k}'\frac{\left|f(\mathbf{k},\mathbf{k}')\right|^2}{r^2}. \tag{15.60}$$

To obtain the differential cross-section from Eq. (15.44), we calculate

$$r^2 \mathbf{j}_{\text{scattered}} \cdot \mathbf{e}_r = \frac{1}{(2\pi\hbar)^3} |\mathbf{k}| \left| f(\mathbf{k}, \mathbf{k}') \right|^2 . \tag{15.61}$$

Finally, we obtain

$$\boxed{\frac{d\sigma}{d\Omega} = \left| f(\mathbf{k}, \mathbf{k}') \right|^2 .} \tag{15.62}$$

15.6 Born Approximation

The Lippmann–Schwinger equation for the wave function, Eq. (15.37), and hence the scattering amplitude of Eq. (15.40), can be hard to evaluate analytically. An approximate solution can be found with perturbation theory. In this approach, we express the scattering amplitude without reference to the unknown wave function $\psi_E(\mathbf{x}')$, as a function of the momentum transfer $\mathbf{q} = \mathbf{k}' - \mathbf{k}$. First, recall that

$$\psi_E(\mathbf{x}) = \frac{1}{(\sqrt{2\pi\hbar})^3} \left\{ e^{\frac{i}{\hbar}\mathbf{k}\cdot\mathbf{x}} + \frac{e^{\frac{i}{\hbar}|\mathbf{k}||\mathbf{x}|}}{|\mathbf{x}|} f(\mathbf{k}, \mathbf{k}') \right\}, \tag{15.63}$$

with the scattering amplitude

$$f(\mathbf{k}, \mathbf{k}') = -\frac{2M}{\hbar^2} \int \frac{d^3\mathbf{x}'}{4\pi} e^{-\frac{i}{\hbar}\mathbf{k}'\cdot\mathbf{x}'} V(\mathbf{x}') \, (2\pi\hbar)^{3/2} \, \psi_E(\mathbf{x}'). \tag{15.64}$$

Inserting Eq. (15.63) into Eq. (15.64), we obtain an expansion of $f(\mathbf{k}, \mathbf{k}')$ in powers of V, which we truncate at leading order:

$$\begin{aligned} f(\mathbf{k}, \mathbf{k}') &= -\frac{2M}{\hbar^2} \int \frac{d^3\mathbf{x}'}{4\pi} e^{-\frac{i}{\hbar}\mathbf{k}'\cdot\mathbf{x}'} V(\mathbf{x}') \, (2\pi\hbar)^{3/2} \left[\frac{1}{\sqrt{2\pi\hbar}^3} \left\{ e^{\frac{i}{\hbar}\mathbf{k}\cdot\mathbf{x}'} + \frac{e^{\frac{i}{\hbar}|\mathbf{k}||\mathbf{x}'|}}{|\mathbf{x}'|} f(\mathbf{k}, \mathbf{k}') \right\} \right] \\ &= -\frac{2M}{\hbar^2} \int \frac{d^3\mathbf{x}'}{4\pi} e^{-\frac{i}{\hbar}\mathbf{k}'\cdot\mathbf{x}'} V(\mathbf{x}') \left\{ e^{\frac{i}{\hbar}\mathbf{k}\cdot\mathbf{x}'} + \frac{e^{\frac{i}{\hbar}|\mathbf{k}||\mathbf{x}'|}}{|\mathbf{x}'|} f(\mathbf{k}, \mathbf{k}') \right\} \\ &\approx -\frac{2M}{\hbar^2} \int \frac{d^3\mathbf{x}'}{4\pi} e^{-\frac{i}{\hbar}\mathbf{q}\cdot\mathbf{x}'} V(\mathbf{x}'), \end{aligned} \tag{15.65}$$

where $\mathbf{q} = \mathbf{k}' - \mathbf{k}$ is the momentum transferred in the scattering process.

Inserting the large-distance form (15.63) under the integral is only heuristic, since $\mathbf{x}'$ runs over the range of V, where $|\mathbf{x}'|$ is not large. The clean argument is to iterate the exact equation (15.37): $\psi_E = \psi_E^{(0)} + O(V)$, so to first order in V one replaces $\psi_E(\mathbf{x}') \to \psi_E^{(0)}(\mathbf{x}')$ in Eq. (15.64), which gives the last line. See Sakurai, Modern QM, Ch. 7 (Ch. 6 in the 2nd ed.).

This expression for the scattering amplitude is known as the *Born approximation*, after the theoretical physicist Max Born, who proposed it in 1926. In this approximation, the scattering amplitude is simply the Fourier transform of the scattering potential.

Let us now express the magnitude of the momentum transfer $\mathbf{q}$ in terms of the scattering angle θ, shown in Fig. 15.6.

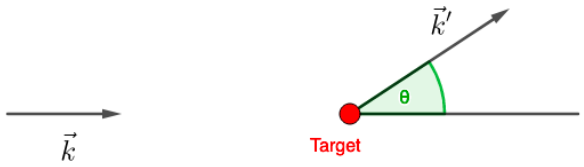


Figure 15.6. Incoming momentum $\mathbf{k}$, outgoing momentum $\mathbf{k}'$ and scattering angle θ.

We have

$$\begin{aligned} q = |\mathbf{q}| &= |\mathbf{k} - \mathbf{k}'| = |\mathbf{k}|\,|\mathbf{e}_z - \mathbf{e}_r| \\ &= |\mathbf{k}|\sqrt{2 - 2\cos\theta} = |\mathbf{k}|\sqrt{4\sin^2\left(\frac{\theta}{2}\right)}. \end{aligned} \tag{15.66}$$

Hence, we obtain a simple expression for the momentum transfer,

$$\boxed{q = 2\,|\mathbf{k}|\sin\frac{\theta}{2}.} \tag{15.67}$$

For scattering off small targets, the potential can often be assumed to be spherically symmetric,

$$V(\mathbf{x}) = V(r). \tag{15.68}$$

The scattering amplitude in the Born approximation becomes

$$\begin{aligned} f(q) &= -\frac{M}{2\pi\hbar^2}\int d^3\mathbf{x}\, e^{-\frac{i}{\hbar}\mathbf{q}\cdot\mathbf{x}} V(r) \\ &= -\frac{M}{2\pi\hbar^2}\int dr r^2 d\cos\theta d\phi\, e^{-\frac{i}{\hbar}qr\cos\theta} V(r) \\ &= -\frac{M}{\hbar^2}\int dr\, r^2 V(r)\left[\frac{e^{-\frac{i}{\hbar}qr\cos\theta}}{-\frac{i}{\hbar}qr}\right]_{\cos\theta=-1}^{\cos\theta=1} \\ &= -\frac{iM}{\hbar q}\int_0^\infty dr\ (rV(r))\left(e^{-\frac{i}{\hbar}qr} - e^{\frac{i}{\hbar}qr}\right). \end{aligned} \tag{15.69}$$

Thus, the scattering amplitude in the Born approximation is

$$\boxed{f(q) = -\frac{2M}{\hbar}\int_0^\infty \mathrm{d}r\,(rV(r))\,\frac{\sin\left(\frac{qr}{\hbar}\right)}{q}.} \tag{15.70}$$

15.6.1 Example: The Yukawa Potential

An important example is the Yukawa potential, which describes interactions mediated by the exchange of a massive particle. Unlike the Coulomb potential, it contains an exponentially decaying factor:

$$V(r) = \frac{V_0}{r}e^{-\frac{m_\phi c}{\hbar}r}. \tag{15.71}$$

The Yukawa form arises from the exchange of any massive boson of mass m_ϕ, with range $\hbar/m_\phi c$. Yukawa introduced it in 1935 for the nuclear force mediated by the pion. The exchange of the weak gauge bosons $W^\pm$ and Z gives the same form, with a range of $\sim 10^{-3}$ fm. Here M remains the mass of the scattered particle.

The scattering amplitude is

$$f(q) = -\frac{2M}{\hbar q} V_0 \int_0^\infty dr\, \sin\left(\frac{qr}{\hbar}\right) e^{-\frac{m_\phi c}{\hbar} r}. \tag{15.72}$$

Upon integration, we find

$$f(q) = -\frac{2MV_0}{q^2 + m_\phi^2 c^2}. \tag{15.73}$$

The differential cross-section is the squared modulus of the amplitude,

$$\frac{d\sigma}{d\Omega} = \left|f(q)\right|^2 = \frac{4M^2V_0^2}{\left|q^2 + m_\phi^2 c^2\right|^2}. \tag{15.74}$$

Written in terms of momenta, Eq. (15.74) contains no explicit Planck constant $\hbar$. One should not conclude from this that the Born approximation reproduces the classical cross-section in general: $\hbar$ hides in $m_\phi c = \hbar/(\text{range})$, and the classical cross-section of a Yukawa potential is *not* Eq. (15.74). The coincidence of the Born, the exact quantum-mechanical and the classical results holds only for the pure Coulomb potential (Rutherford formula, below). The Born approximation is a weak-potential (high-energy) approximation, not a classical limit. See Sakurai, Modern QM, Ch. 7 (Ch. 6 in the 2nd ed.); Gottfried–Yan, Ch. 8.

The result of Eq. (15.74) can be used to obtain the cross-section for the Coulomb potential in the limit $m_\phi \to 0$. This entails the assumption that the Coulomb potential falls off faster than $1/r$ at extremely large distances, so that the integral of Eq. (15.72) converges. With this assumption and with $V_0 = Z_1 Z_2 e^2$, where $Z_1 e$ and $Z_2 e$ are the charges of the two particles, we obtain the Rutherford cross-section

$$\left(\frac{d\sigma}{d\Omega}\right)_{\text{Rutherford}} = \frac{M^2 Z_1^2 Z_2^2 e^4}{4\,|\mathbf{k}|^4} \frac{1}{\sin^4\frac{\theta}{2}} = \left(\frac{Z_1 Z_2 e^2}{4E}\right)^2 \frac{1}{\sin^4\frac{\theta}{2}}, \quad E = \frac{|\mathbf{k}|^2}{2M}. \tag{15.75}$$

To summarise: the Yukawa potential vanishes fast enough for $r \to \infty$ that the large-r region gives a negligible contribution to the integral of Eq. (15.72). For the Coulomb potential, $m_\phi = 0$, the integral does not converge. We can, however, assume that physical mechanisms (such as screening) suppress the Coulomb potential at extremely large distances, and take the limit $m_\phi \to 0$ after performing the integration in Eq. (15.72):

$$\frac{d\sigma}{d\Omega} \overset{m_\phi \to 0}{=} \frac{4M^2V_0^2}{q^4}. \tag{15.76}$$

15.7 Partial Wave Expansion

At the beginning of this chapter, we cast the problem of elastic scattering into the form of the Schrödinger equation. In this section we analyse a spherically symmetric scattering potential. We are interested in the behaviour of the wave function in the limit $r \to \infty$, where

it is the sum of a free wave and a scattered wave, as given by the Lippmann–Schwinger equation. We shall use spherical Bessel functions (Appendix A) to express the scattering amplitude.

The Schrödinger equation for a spherically symmetric potential reads

$$\left[-\frac{\hbar^2}{2M}\nabla^2 + V(r)\right]\psi(r,\theta,\phi) = E\psi(r,\theta,\phi). \tag{15.77}$$

Its solutions are

$$\psi_{\ell m}(r,\theta,\phi) \equiv R_\ell(r)Y_\ell^m(\theta,\phi) \equiv \frac{U_\ell(r)}{r}Y_\ell^m(\theta,\phi), \tag{15.78}$$

where $Y_\ell^m(\theta,\phi)$ are the spherical harmonics and the radial functions satisfy

$$U_\ell''(r) - \left[\frac{2M}{\hbar^2}V(r) + \frac{\ell(\ell+1)}{r^2} - k^2\right]U_\ell(r) = 0, \tag{15.79}$$

with

$$k^2 \equiv \frac{2ME}{\hbar^2}. \tag{15.80}$$

For short-range scattering potentials, $V(r)$ vanishes as r goes to infinity,

$$\lim_{r\to\infty} V(r) = 0. \tag{15.81}$$

so that at large distances the radial equation simplifies to

$$U_\ell''(r) + k^2U_\ell(r) = 0, \quad r \to \infty, \tag{15.82}$$

which has the solutions

$$U_\ell(r) = A_\ell \sin(kr + \tilde{\delta}_\ell), \quad r \to \infty. \tag{15.83}$$

The phase $\tilde{\delta}_\ell \equiv \tilde{\delta}_\ell(k)$ is a function of k. For $r \to \infty$, the complete wave function is

$$\psi_{\ell m}(r,\theta,\phi) = \frac{U_\ell(r)}{r}Y_\ell^m(\theta,\phi) = \frac{A_\ell}{r}\sin(kr + \tilde{\delta}_\ell)Y_\ell^m(\theta,\phi). \tag{15.84}$$

Let us recall the set-up of a scattering experiment, shown in Fig. 15.7.

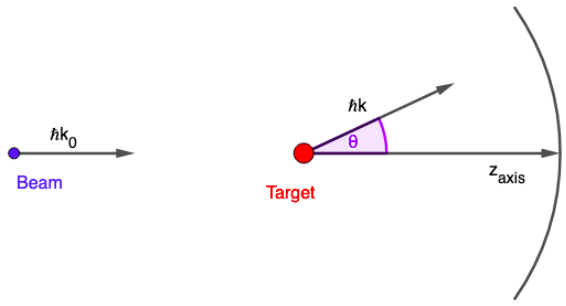


Figure 15.7. Scattering geometry: the beam, with momentum $\hbar\mathbf{k}_0$, moves along the z-axis; a particle scattered with momentum $\hbar\mathbf{k}$ forms the angle θ with it.

The problem is symmetric under rotations about the beam axis, so the solutions are independent of the azimuthal angle ϕ, that is, they have $m = 0$. The solutions then take the form

$$\psi_\ell(r,\theta) = \frac{U_\ell(r)}{r} Y_\ell^0(\theta,\phi), \tag{15.85}$$

where the spherical harmonics reduce to Legendre polynomials,

$$Y_\ell^0(\theta,\phi) = \sqrt{\frac{2\ell+1}{4\pi}} P_\ell(\cos\theta). \tag{15.86}$$

Thus, for $r \to \infty$, the solutions behave as

$$\boxed{\psi_\ell(r,\theta) \overset{r\to\infty}{\sim} \frac{A_\ell}{r} \sin(kr + \tilde{\delta}_\ell) P_\ell(\cos\theta),} \tag{15.87}$$

where the factor $\sqrt{(2\ell+1)/4\pi}$ has been absorbed into the constant A_ℓ. The Legendre polynomials satisfy the differential equation

$$(1-x^2)\frac{\partial^2}{\partial x^2} P_\ell(x) - 2x\frac{\partial}{\partial x} P_\ell(x) + \ell(\ell+1) P_\ell(x) = 0. \tag{15.88}$$

The Legendre polynomial $P_\ell(x)$ has degree ℓ. As expected for eigenfunctions of a Hermitian operator, the square of the angular momentum, they are orthogonal,

$$\int_{-1}^{1} d\cos\theta \, P_\ell(\cos\theta) P_{\ell'}(\cos\theta) = \frac{2\delta_{\ell\ell'}}{2\ell+1}. \tag{15.89}$$

Exercise

Verify the above starting from $\langle \ell, m=0 | \ell', m=0 \rangle = \delta_{\ell\ell'}$. (15.90)

Furthermore, the Legendre polynomials form a basis in which we can expand any regular function $f(\cos\theta)$,

$$f(\cos\theta) = \sum_\ell c_\ell P_\ell(\cos\theta). \tag{15.91}$$

Exercise

Argue physically why the above is true. (15.92)

Let us return to the large-distance behaviour of the wave function, Eq. (15.87), which we derived directly from the Schrödinger equation in the limit $r \to \infty$. We compare it with the prediction of the Lippmann–Schwinger equation, Eq. (15.39),

$$\psi(r,\theta) \overset{r\to\infty}{\sim} e^{ikr\cos\theta} + \frac{e^{ikr}}{r} f(k,\cos\theta). \tag{15.93}$$

Although Eqs. (15.87) and (15.93) look different, we can cast Eq. (15.93) into the form of Eq. (15.87). This gives a relation between the scattering amplitude and the phases $\tilde{\delta}_\ell$. We write

$$e^{ikz} = e^{ikr\cos\theta} \tag{15.94}$$

as a superposition of Legendre polynomials,

$$e^{ikr\cos\theta} = \sum_{\ell=0}^{\infty} i^\ell (2\ell+1) J_\ell(kr) P_\ell(\cos\theta), \tag{15.95}$$

which we prove in Sec. 15.7.2. Each coefficient is proportional to a spherical Bessel function (see Appendix A),

$$J_\ell(\rho) = (-\rho)^\ell \left[\frac{1}{\rho}\frac{\partial}{\partial\rho}\right]^\ell J_0(\rho), \tag{15.96}$$

with

$$J_0(\rho) = \frac{\sin\rho}{\rho}. \tag{15.97}$$

The behaviour of $J_\ell(\rho)$ for $\rho \to 0$ is

$$J_\ell(\rho) \overset{\rho\to 0}{\sim} \frac{\rho^\ell}{1\cdot 3\cdot 5\cdot \ldots \cdot (2\ell+1)} = \frac{\rho^\ell}{\frac{(2\ell+1)!}{2^\ell \ell!}}. \tag{15.98}$$

and its asymptotic behaviour for $\rho \to \infty$ is

$$J_\ell(\rho) \overset{\rho\to\infty}{\sim} \frac{1}{\rho}\sin\left(\rho - \frac{\pi\ell}{2}\right). \tag{15.99}$$

Applying Eq. (15.99) to Eq. (15.95), we obtain for $r \to \infty$

$$\begin{aligned} e^{ikr\cos\theta} &\overset{r\to\infty}{\approx} \sum_{\ell=0}^{\infty} i^\ell (2\ell+1)\frac{\sin(kr-\pi\ell/2)}{kr} P_\ell(\cos\theta) \\ &= \sum_{\ell=0}^{\infty} i^\ell \frac{(2\ell+1)}{2ik}\left[\frac{e^{i(kr-\pi\ell/2)} - e^{-i(kr-\pi\ell/2)}}{r}\right] P_\ell(\cos\theta) \\ &= \sum_{\ell=0}^{\infty} i^\ell e^{-i\pi\ell/2}\frac{(2\ell+1)}{2ik}\left[\frac{e^{ikr} - e^{-i(kr-\pi\ell)}}{r}\right] P_\ell(\cos\theta) \\ &= \sum_{\ell=0}^{\infty} \frac{(2\ell+1)}{2ik}\left[\frac{e^{ikr} - e^{-i(kr-\pi\ell)}}{r}\right] P_\ell(\cos\theta). \end{aligned} \tag{15.100}$$

We expand the scattering amplitude $f(k,\theta)$ of Eq. (15.93) in the same basis of Legendre polynomials,

$$f(k,\theta) = \frac{1}{k}\sum_{\ell=0}^{\infty} a_\ell(k)(2\ell+1)P_\ell(\cos\theta). \tag{15.101}$$

The Lippmann–Schwinger form of the wave function then becomes

$$\begin{aligned} \psi(r) &\overset{r\to\infty}{\sim} e^{ikr\cos\theta} + \frac{f(k,\theta)}{r}e^{ikr} \\ &= \sum_{\ell=0}^{\infty} \frac{P_\ell(\cos\theta)(2\ell+1)}{kr}\left\{\frac{1}{2i}\left(e^{ikr} - e^{-i(kr-\pi\ell)}\right) + e^{ikr}a_\ell(k)\right\} \\ &= \sum_{\ell=0}^{\infty} \frac{P_\ell(\cos\theta)(2\ell+1)}{2ikr}\left\{e^{ikr}\left(1+2ia_\ell(k)\right) - e^{-i(kr-\pi\ell)}\right\}. \end{aligned} \tag{15.102}$$

Comparing Eqs. (15.102) and (15.87) (the ratios of the coefficients of e^{ikr} and e^{-ikr} must agree, $1 + 2ia_\ell = e^{2i\tilde{\delta}_\ell + i\pi\ell}$), we find

$$\boxed{a_\ell(k) = e^{i\delta_\ell(k)} \sin\delta_\ell(k), \quad \delta_\ell(k) = \tilde{\delta}_\ell(k) + \frac{\ell\pi}{2}.} \tag{15.103}$$

We thus obtain the partial-wave expansion of the scattering amplitude,

$$\boxed{f(k,\theta) = \frac{1}{k}\sum_{\ell=0}^{\infty}(2\ell+1)P_\ell(\cos\theta)e^{i\delta_\ell(k)}\sin\delta_\ell(k).} \tag{15.104}$$

The phase shifts $\delta_\ell(k)$ are obtained by comparing this expression with the scattering amplitude of Eq. (15.40). Often one uses the scattering amplitude in the Born approximation, Eq. (15.70) (see Sec. 15.8). The reader may find it useful to compare this derivation with the expansion of the full wave function in spherical Bessel functions in Sakurai [4].

15.7.1 The Optical Theorem

The optical theorem relates the total cross-section to the scattering amplitude at $\theta = 0$. We integrate the differential cross-section using Eq. (15.104):

$$\begin{aligned}\sigma_{\text{SCAT}} &\equiv \int \frac{d\sigma_{\text{SCAT}}}{d\Omega}d\Omega = \int d\Omega\,\left|f(k,\theta)\right|^2 \\ &= \frac{2\pi}{k^2}\int d\cos\theta\,\sum_{\ell,\ell'}(2\ell+1)P_\ell(\cos\theta)(2\ell'+1)P_{\ell'}(\cos\theta)\cdot e^{i(\delta_\ell-\delta_{\ell'})}\sin\delta_\ell\sin\delta_{\ell'} \\ &= \frac{4\pi}{k^2}\sum_{\ell=0}^{\infty}(2\ell+1)\sin^2\delta_\ell.\end{aligned} \tag{15.105}$$

The forward scattering amplitude, at $\theta = 0$, is

$$f(k,\theta=0) = \frac{1}{k}\sum_{\ell=0}^{\infty}(2\ell+1)\underbrace{P_\ell(\cos 0)}_{1}e^{i\delta_\ell}\sin\delta_\ell. \tag{15.106}$$

This implies

$$\text{Im}\,f(k,\theta=0) = \frac{1}{k}\sum_{\ell=0}^{\infty}(2\ell+1)\sin^2\delta_\ell. \tag{15.107}$$

Comparing this expression with Eq. (15.105), we obtain

$$\boxed{\sigma_{\text{SCAT}} \equiv \int \frac{\partial\sigma_{\text{SCAT}}}{\partial\Omega}d\Omega = \frac{4\pi}{k}\text{Im}\,f(k,\theta=0).} \tag{15.108}$$

This relation between the total cross-section and the forward scattering amplitude is known as the *optical theorem*.

15.7.2 Application: Partial-Wave Expansion of a Plane Wave

As an example, we consider a free particle. We already know the solution, but the example allows us to expand the free-particle wave in spherical Bessel functions. The Schrödinger equation for a free particle is

$$\left(\nabla^2 + k^2\right)\psi(\mathbf{x}') = 0, \tag{15.109}$$

with $k^2 = 2ME/\hbar^2$. The solution is a plane wave,

$$\psi(\mathbf{x}') = e^{i\mathbf{k}\cdot\mathbf{x}'}. \tag{15.110}$$

If we take $\mathbf{k} = k\mathbf{e}_z$, this becomes

$$\psi(r,\theta,\phi) = e^{ikr\cos\theta}. \tag{15.111}$$

Let us consider Eq. (15.109) in spherical coordinates. Since the solution is regular at $r \to 0$, we can write

$$\psi(r,\theta,\phi) = e^{ikr\cos\theta} = \sum_{\ell,m} c_{\ell,m} J_\ell(rk) Y_\ell^m(\theta,\phi). \tag{15.112}$$

The spherical Neumann functions y_ℓ do not appear, since they are singular at $r \to 0$, as shown in Appendix A. As $\psi(r,\theta,\phi)$ is independent of ϕ, this simplifies to

$$\psi(r,\theta) = e^{ikr\cos\theta} = \sum_{\ell} A_\ell J_\ell(rk) P_\ell(\cos\theta). \tag{15.113}$$

To determine the coefficients, we use the orthogonality of the Legendre polynomials. Multiplying both sides by $P_{\ell'}(\cos\theta)$ and integrating over $\cos\theta$, we obtain

$$\sum_{\ell} A_\ell J_\ell(rk) \int_{-1}^{1} P_\ell(\cos\theta) P_{\ell'}(\cos\theta)\, d\cos\theta = \int_{-1}^{1} \psi(r,\theta) P_{\ell'}(\cos\theta)\, d\cos\theta. \tag{15.114}$$

The left-hand side of Eq. (15.114) is

$$\sum_{\ell} A_\ell J_\ell(rk) \frac{2\delta_{\ell\ell'}}{2\ell'+1} = \frac{2}{2\ell'+1} A_{\ell'} J_{\ell'}(rk). \tag{15.115}$$

For the right-hand side of Eq. (15.114), we expand the exponential (renaming $\ell' \to \ell$ and $z = \cos\theta$):

$$\begin{aligned} \int_{-1}^{1} dz\, e^{ikrz} P_\ell(z) &= \int_{-1}^{1} dz \left[(*) + \frac{(ikrz)^\ell}{\ell!} + \frac{(ikrz)^{\ell+1}}{(\ell+1)!} + \ldots\right] P_\ell(z) \\ &= \frac{i^\ell (kr)^\ell}{\ell!} \int_{-1}^{1} dz\, z^\ell P_\ell(z) \; + O\left(r^{\ell+1}\right), \end{aligned} \tag{15.116}$$

where $(*)$ denotes the terms of degree lower than ℓ. They do not contribute to the integral because of orthogonality:

$$\int_{-1}^{1} dz \left(\sum_{i=0}^{n} a_i z^i\right) P_\ell(z) \overset{n<\ell}{=} 0. \tag{15.117}$$

Recall that the Legendre polynomial P_ℓ has degree ℓ. We can therefore expand the monomial z^ℓ as

$$z^\ell = \sum_{n=0}^{\ell} c_n^{(\ell)} P_n(z). \tag{15.118}$$

Thus the integral in Eq. (15.116) is

$$\int_{-1}^{1} dz\, z^\ell P_\ell(z) = \int_{-1}^{1} dz \left(\sum_{n=0}^{\ell} c_n^{(\ell)} P_n(z) \right) P_\ell(z) = \frac{2}{2\ell+1} c_\ell^{(\ell)}. \tag{15.119}$$

Inserting Eqs. (15.115), (15.116) and (15.119) into Eq. (15.114), we obtain

$$\frac{2}{2\ell+1} A_\ell J_\ell(rk) = i^\ell \frac{(rk)^\ell}{\ell!} \frac{2}{2\ell+1} c_\ell^{(\ell)} + O\left(r^{\ell+1}\right). \tag{15.120}$$

Cancelling common factors and inserting the small-r behaviour of $J_\ell(rk)$ from Eqs. (A.17) and (A.18) into Eq. (15.120), we obtain

$$A_\ell \cdot \left(\frac{(rk)^\ell 2^\ell \ell!}{(2\ell+1)!} \right) + O\left(r^{\ell+1}\right) = i^\ell \frac{(rk)^\ell}{\ell!} c_\ell^{(\ell)}. \tag{15.121}$$

Solving Eq. (15.121) for A_ℓ, we obtain

$$\boxed{A_\ell = i^\ell c_\ell^{(\ell)} \frac{(2\ell+1)!}{2^\ell (\ell!)^2} = i^\ell (2\ell+1).} \tag{15.122}$$

In the last step we used that the leading coefficient of $P_\ell(z)$ is $\frac{(2\ell)!}{2^\ell(\ell!)^2}$, so that $c_\ell^{(\ell)} = \frac{2^\ell(\ell!)^2}{(2\ell)!}$. Thus the free-particle wave function of Eq. (15.113) is (for all r; the limit $r \to 0$ was only used to determine the constants A_ℓ)

$$\psi(r,\theta) = e^{ikr\cos\theta} = \sum_\ell i^\ell (2\ell+1) J_\ell(rk) P_\ell(\cos\theta). \tag{15.123}$$

15.8 Born Approximation and Scattering Phases

The scattering amplitude in the Born approximation is simply the Fourier transform of the potential,

$$f(k, \mathbf{k}') = -\frac{M}{2\pi\hbar^2} \int d^3\mathbf{y}\, V(|\mathbf{y}|)\, e^{i\mathbf{k}\cdot\mathbf{y}}\, e^{-i\mathbf{k}'\cdot\mathbf{y}}. \tag{15.124}$$

(This is Eq. (15.65), with $\mathbf{k}$, $\mathbf{k}'$ now wave vectors, i.e. momenta divided by $\hbar$, and $|\mathbf{k}'| = |\mathbf{k}| = k$.)

Generalising Eq. (15.123), we have

$$e^{i\mathbf{k}\cdot\mathbf{x}} = 4\pi \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} i^\ell J_\ell(kr)\, Y_{\ell m}(\Omega_k)^*\, Y_{\ell m}(\Omega_x) \tag{15.125}$$

and

$$P_\ell(\cos\theta) = \frac{4\pi}{2\ell+1} \sum_{m=-\ell}^{\ell} Y_{\ell m}(\Omega_k)^*\, Y_{\ell m}(\Omega_{k'}), \tag{15.126}$$

where θ is the angle between the two directions Ω_k and $\Omega_{k'}$ (addition theorem of spherical harmonics).

Substituting these expansions into Eq. (15.124) and performing the angular integrations, (the integral over Ω_y of $Y_{\ell m}(\Omega_y)\, Y^*_{\ell' m'}(\Omega_y)$ gives $\delta_{\ell\ell'}\delta_{mm'}$, the phases $i^\ell(-i)^\ell = 1$ cancel, and the sum over m gives $\frac{2\ell+1}{4\pi} P_\ell(\cos\theta)$) we obtain

$$f(k,\theta) = -\frac{2M}{\hbar^2} \sum_{\ell=0}^{\infty} (2\ell+1)\, P_\ell(\cos\theta) \int_0^\infty dr\, r^2\, V(r)\, J_\ell(kr)^2 .$$

We compare this result with the general form of the partial-wave expansion, Eq. (15.104), which for small potentials reads

$$f(k,\theta) = \sum_{\ell=0}^{\infty} \frac{2\ell+1}{k} \underbrace{e^{i\delta_\ell} \sin\delta_\ell}_{\simeq \delta_\ell} P_\ell(\cos\theta), \tag{15.127}$$

and find

$$\boxed{\delta_\ell \simeq -\frac{2Mk}{\hbar^2} \int_0^\infty dr\, r^2\, V(r)\, J_\ell(kr)^2 .} \tag{15.128}$$

For short-range potentials, the integral is dominated by the region of small r, where we can approximate

$$J_\ell(kr) \simeq \frac{(kr)^\ell}{(2\ell+1)!!} . \tag{15.129}$$

The contributions of large ℓ are therefore suppressed. This is very important: only the first few partial waves contribute to the scattering.

Eq. (15.128) is the Born phase shift (Sakurai, Modern QM, Ch. 7; Ch. 6 in the 2nd ed.); summed over ℓ, it reproduces the Born amplitude of the Yukawa potential, Eq. (15.73). It requires $|\delta_\ell| \ll 1$, so the linearised amplitude does not satisfy unitarity exactly (the optical theorem holds only at the next order in V). More precisely, for a potential of range a the integral runs over $r \lesssim a$, and $J_\ell(kr) \simeq (kr)^\ell/(2\ell+1)!!$ is tiny there as long as $ka \ll \ell$; hence only partial waves with $\ell \lesssim ka$ (classically, impact parameter $b = \ell/k \lesssim a$) contribute, and at low energy, $ka \ll 1$, only $\ell = 0$ (s-wave) scattering survives.

16

Radiation and Matter Interaction

Atoms emit and absorb light only at sharp, characteristic frequencies.[1] The spectral lines of hydrogen were measured with great precision long before quantum mechanics existed, and explaining them was one of its first triumphs. The interaction of electromagnetic radiation with matter was, in fact, central at the very beginning of quantum theory. Planck introduced the quanta of energy in order to understand the spectrum of the radiation which is emitted and absorbed by the walls of a hot cavity, the black body. Einstein showed a few years later that Planck's spectrum follows if atoms can absorb light, emit light when they are "stimulated" by light, and, in addition, emit light **spontaneously**, even when no light is around.

This last process should puzzle us. Consider a hydrogen atom in the excited state $2p$, in an empty box. The state $|2, 1, m\rangle$ is an eigenstate of the hydrogen Hamiltonian and, according to everything we have learned so far, it is stationary. It should therefore remain excited forever. Experimentally, it decays to the ground state after about 1.6 ns, emitting a photon. Why does an excited atom decay at all?

In this chapter we will first review the salient features of classical electrodynamics and decompose the radiation field into independent modes. We will then couple an atom to a classical radiation field and compute, with the golden rule of Chapter 14, the rates for absorption and stimulated emission. We will see that symmetries, rotations and parity, forbid most of the transitions which energy conservation alone would allow. Finally, we will quantise the electromagnetic field. Photons will then emerge as the quanta of harmonic oscillators, and spontaneous emission will appear as a prediction of the theory. We will compute the lifetime of the $2p$ state of hydrogen and, as a reward, re-derive Planck's law.

16.1 Maxwell Equations and Electromagnetic Radiation

The classical theory of electromagnetism is summarised by the **equations of Maxwell**. In this chapter we will use Gaussian units, which are the units of all the atomic physics formulae of this part of the script (the hydrogen potential is $-e^2/r$ and the fine structure constant is $\alpha = e^2/(\hbar c)$). In these units the electric and the magnetic field have the same

[1] The presentation in this chapter is partly based on lecture notes of M. Gaberdiel.

dimensions.[2] The Maxwell equations read:

$$\begin{aligned} \nabla \cdot \mathbf{B} &= 0, & \nabla \times \mathbf{E} + \frac{1}{c}\frac{\partial \mathbf{B}}{\partial t} &= 0, \\ \nabla \cdot \mathbf{E} &= 4\pi\rho, & \nabla \times \mathbf{B} - \frac{1}{c}\frac{\partial \mathbf{E}}{\partial t} &= \frac{4\pi}{c}\mathbf{j}, \end{aligned} \tag{16.1}$$

where ρ is the charge density and $\mathbf{j}$ the current density.

The two homogeneous equations (the first line) can be solved once and for all by introducing potentials. A divergence-free field is a curl,

$$\nabla \cdot \mathbf{B} = 0 \quad \rightsquigarrow \quad \mathbf{B} = \nabla \times \mathbf{A}, \tag{16.2}$$

where $\mathbf{A}$ is the **vector potential**. Substituting into Faraday's law, we find

$$\nabla \times \left(\mathbf{E} + \frac{1}{c}\frac{\partial \mathbf{A}}{\partial t} \right) = 0 \quad \rightsquigarrow \quad \mathbf{E} = -\nabla\phi - \frac{1}{c}\frac{\partial \mathbf{A}}{\partial t}, \tag{16.3}$$

since a curl-free field is a gradient. ϕ is the **scalar potential**. Both fields are now expressed in terms of four functions, $(\phi, \mathbf{A})$, instead of six.

We shall work with a microscopic description of matter, where all charges (electrons and nuclei) are taken into account explicitly. Therefore, we do not need a macroscopic description of dielectric and magnetic polarisation. Substituting the potentials into the inhomogeneous Maxwell equations, we obtain:

$$4\pi\rho = -\nabla^2\phi - \frac{1}{c}\frac{\partial}{\partial t}\nabla \cdot \mathbf{A}, \tag{16.4}$$

$$\frac{4\pi}{c}\mathbf{j} = \nabla(\nabla \cdot \mathbf{A}) - \nabla^2\mathbf{A} + \frac{1}{c}\frac{\partial}{\partial t}\nabla\phi + \frac{1}{c^2}\frac{\partial^2 \mathbf{A}}{\partial t^2}. \tag{16.5}$$

The potentials are not unique. The **gauge transformation**

$$\mathbf{A} \to \mathbf{A}' = \mathbf{A} + \nabla\chi, \qquad \phi \to \phi' = \phi - \frac{1}{c}\frac{\partial \chi}{\partial t}, \tag{16.6}$$

with an arbitrary function $\chi(\mathbf{x}, t)$, leaves $\mathbf{E}$ and $\mathbf{B}$ unchanged (**exercise**). We can use this freedom to impose a condition on the potentials. For the problems of this chapter the most convenient choice is the **Coulomb gauge**,

$$\nabla \cdot \mathbf{A} = 0. \tag{16.7}$$

Exercise: *Show that, starting from arbitrary potentials, we can always find a function χ such that $\nabla \cdot \mathbf{A}' = 0$.*

In the Coulomb gauge, Eq. (16.4) becomes the Poisson equation of electrostatics, $\nabla^2\phi = -4\pi\rho$, with the solution

$$\phi(\mathbf{x}, t) = \int d^3\mathbf{y}\, \frac{\rho(\mathbf{y}, t)}{|\mathbf{x} - \mathbf{y}|}. \tag{16.8}$$

For point charges Q_j at the positions $\mathbf{x}_j(t)$, i.e. $\rho(\mathbf{x}, t) = \sum_j Q_j\, \delta^{(3)}\left(\mathbf{x} - \mathbf{x}_j(t)\right)$, we obtain

$$\phi(\mathbf{x}, t) = \sum_j \frac{Q_j}{|\mathbf{x} - \mathbf{x}_j(t)|}. \tag{16.9}$$

[2] In the Electrodynamics course we use SI units. To translate, we replace $e^2 \to e^2/(4\pi\epsilon_0)$; all physical results of this chapter, written in terms of $\alpha = e^2/(\hbar c)$, are the same in both systems.

This is a remarkable result. The scalar potential in the Coulomb gauge is the instantaneous Coulomb potential of the charges. It depends only on their positions at the same time t and it carries no independent degrees of freedom. For an atom, it is simply the Coulomb attraction to the nucleus and the Coulomb repulsion among the electrons, which are already part of the atomic Hamiltonian.

The radiation is then contained entirely in the vector potential. Eq. (16.5) becomes a wave equation with a source:

$$\Box \mathbf{A} \equiv \left(\frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \nabla^2 \right) \mathbf{A} = \frac{4\pi}{c} \mathbf{j} - \frac{1}{c} \frac{\partial}{\partial t} \nabla \phi. \tag{16.10}$$

Exercise: *Using charge conservation, $\partial_t \rho + \nabla \cdot \mathbf{j} = 0$, show that the right-hand side of Eq. (16.10) is divergence-free. Only the "transverse" part of the current is a source of* $\mathbf{A}$.

Let us now consider a region of space without charges. There, the vector potential satisfies

$$\Box \mathbf{A} = 0, \qquad \nabla \cdot \mathbf{A} = 0. \tag{16.11}$$

It is convenient to place the system in a large cubic box of length L and to impose periodic boundary conditions. At the end of our calculations we will take $L \to \infty$. The allowed wave vectors are

$$\mathbf{k} = \frac{2\pi}{L} \left(n_x, n_y, n_z \right), \qquad n_x, n_y, n_z \in \mathbb{Z}. \tag{16.12}$$

We expand the vector potential in plane waves. The wave equation requires that each plane wave oscillates with the frequency $\omega_k = c\,|\mathbf{k}|$, and we can write the general real solution as

$$\begin{aligned} \mathbf{A}(\mathbf{x}, t) \;=\; & \frac{1}{\sqrt{L^3}} \sum_{\mathbf{k},\lambda} \Big\{ \alpha(\mathbf{k}, \lambda)\, \mathbf{e}(\mathbf{k}, \lambda)\, e^{i(\mathbf{k}\cdot\mathbf{x} - \omega_k t)} \\ & + \alpha^*(\mathbf{k}, \lambda)\, \mathbf{e}^*(\mathbf{k}, \lambda)\, e^{-i(\mathbf{k}\cdot\mathbf{x} - \omega_k t)} \Big\}. \end{aligned} \tag{16.13}$$

The complex numbers $\alpha(\mathbf{k}, \lambda)$ are the amplitudes of the waves and $\mathbf{e}(\mathbf{k}, \lambda)$ are **polarisation vectors**. How many polarisations do we have for a given $\mathbf{k}$? The Coulomb gauge condition gives:

$$\nabla \cdot \mathbf{A} = \frac{i}{\sqrt{L^3}} \sum_{\mathbf{k},\lambda} \left\{ \alpha\, \mathbf{k} \cdot \mathbf{e}\, e^{i(\mathbf{k}\cdot\mathbf{x} - \omega_k t)} - \alpha^*\, \mathbf{k} \cdot \mathbf{e}^*\, e^{-i(\mathbf{k}\cdot\mathbf{x} - \omega_k t)} \right\} = 0, \tag{16.14}$$

which gives

$$\mathbf{k} \cdot \mathbf{e}(\mathbf{k}, \lambda) = 0, \tag{16.15}$$

since plane waves with different $\mathbf{k}$ are linearly independent. The polarisation vectors lie in the plane perpendicular to the direction of propagation. Electromagnetic waves are **transverse** and there are only two independent polarisations, $\lambda = 1, 2$. Without loss of generality we choose them orthonormal,

$$\mathbf{e}^*(\mathbf{k}, \lambda) \cdot \mathbf{e}(\mathbf{k}, \lambda') = \delta_{\lambda\lambda'}, \tag{16.16}$$

for example two real unit vectors (linear polarisation) or the combinations $(\mathbf{e}_1 \pm i\mathbf{e}_2)/\sqrt{2}$ (circular polarisation).

Exercise: *Show that $\sum_\lambda e_i(\mathbf{k}, \lambda)\, e_j^*(\mathbf{k}, \lambda) = \delta_{ij} - k_i k_j / \mathbf{k}^2$. We will need this identity later.*

The electric and magnetic fields of the wave are obtained by differentiation. For $\phi = 0$ we have $\mathbf{E} = -\frac{1}{c}\partial_t \mathbf{A}$ and $\mathbf{B} = \nabla \times \mathbf{A}$. Abbreviating $\alpha \equiv \alpha(\mathbf{k}, \lambda)$, $\mathbf{e} \equiv \mathbf{e}(\mathbf{k}, \lambda)$ and $\varphi_k \equiv \mathbf{k} \cdot \mathbf{x} - \omega_k t$, we find:

$$\begin{aligned} \mathbf{E}(\mathbf{x}, t) &= \frac{1}{\sqrt{L^3}} \sum_{\mathbf{k},\lambda} \frac{i\omega_k}{c} \left\{ \alpha\, \mathbf{e}\, e^{i\varphi_k} - \alpha^*\, \mathbf{e}^*\, e^{-i\varphi_k} \right\}, \\ \mathbf{B}(\mathbf{x}, t) &= \frac{i}{\sqrt{L^3}} \sum_{\mathbf{k},\lambda} \left\{ \alpha\, (\mathbf{k} \times \mathbf{e})\, e^{i\varphi_k} - \alpha^*\, (\mathbf{k} \times \mathbf{e}^*)\, e^{-i\varphi_k} \right\}. \end{aligned} \tag{16.17}$$

The energy stored in the radiation field is

$$E_{\text{rad}} = \int_{L^3} d^3\mathbf{x}\, \frac{\mathbf{E}^2 + \mathbf{B}^2}{8\pi}. \tag{16.18}$$

Substituting Eq. (16.17) and performing the space integrals with

$$\int_{L^3} d^3\mathbf{x}\, e^{i(\mathbf{k}-\mathbf{k}')\cdot\mathbf{x}} = L^3 \delta_{\mathbf{k}\mathbf{k}'},$$

we find that all modes decouple:

$$E_{\text{rad}} = \sum_{\mathbf{k},\lambda} \frac{\omega_k^2}{2\pi c^2} \left|\alpha(\mathbf{k}, \lambda)\right|^2. \tag{16.19}$$

Exercise: *Prove Eq.* (16.19). *Notice that the terms proportional to* $\alpha(\mathbf{k})\,\alpha(-\mathbf{k})$, *which oscillate in time as* $e^{-2i\omega_k t}$, *cancel between* $\mathbf{E}^2$ *and* $\mathbf{B}^2$. *The energy is time-independent, as it should be.*

So far everything is classical. We know, however, that light comes in quanta, the **photons**. A photon in the mode $(\mathbf{k}, \lambda)$ carries the energy $\hbar\omega_k$ and the momentum $\hbar\mathbf{k}$. If the mode $(\mathbf{k}, \lambda)$ is occupied by $N(\mathbf{k}, \lambda)$ photons, the energy of the radiation is

$$E_{\text{rad}} = \sum_{\mathbf{k},\lambda} N(\mathbf{k}, \lambda)\hbar\omega_k. \tag{16.20}$$

Comparing with Eq. (16.19), we can assign a number of photons to the classical amplitude of each mode:

$$N(\mathbf{k}, \lambda) = \frac{\omega_k}{2\pi\hbar c^2} \left|\alpha(\mathbf{k}, \lambda)\right|^2. \tag{16.21}$$

At this stage, this is a heuristic identification and nothing forces N to be an integer. We will justify it properly in Section 16.4.

16.2 The Classical Electromagnetic Interaction with Matter

We now place an atom in a classical radiation field. Let us start with a single electron of charge e and mass m, which is bound by a potential $U(\mathbf{x})$, for example the Coulomb potential of the nucleus. In an electromagnetic field the momentum $\mathbf{p}$ is replaced by $\mathbf{p} - \frac{e}{c}\mathbf{A}$, and the Schrödinger equation reads

$$i\hbar \frac{\partial}{\partial t}\psi(\mathbf{x}, t) = \left[\frac{1}{2m} \left(-i\hbar\nabla - \frac{e}{c}\mathbf{A}(\mathbf{x}, t) \right)^2 + e\phi(\mathbf{x}, t) + U(\mathbf{x}) \right] \psi(\mathbf{x}, t). \tag{16.22}$$

The potentials are not unique. What happens to the wave function under the gauge transformation of Eq. (16.6)? The physics cannot change, and indeed, if ψ solves Eq. (16.22) with the potentials $(\phi, \mathbf{A})$, then

$$\psi'(\mathbf{x}, t) = e^{\frac{ie}{\hbar c}\chi(\mathbf{x},t)}\, \psi(\mathbf{x}, t) \tag{16.23}$$

solves the same equation with the potentials $(\phi', \mathbf{A}')$.
Exercise: *Prove it. Hint: show first that*

$$\left(-i\hbar\nabla - \frac{e}{c}\mathbf{A}'\right)\psi' = e^{\frac{ie}{\hbar c}\chi}\left(-i\hbar\nabla - \frac{e}{c}\mathbf{A}\right)\psi.$$

Since χ is real, the probability density $|\psi'|^2 = |\psi|^2$ is gauge invariant. A gauge transformation of the potentials is accompanied by a local change of the phase of the wave function.

Let us expand the square in Eq. (16.22). We separate the Hamiltonian into the Hamiltonian of the atom,

$$H_0 = \frac{\hat{\mathbf{p}}^2}{2m} + U(\hat{\mathbf{x}}), \tag{16.24}$$

where $\hat{\mathbf{p}}$ and $\hat{\mathbf{x}}$ are the momentum and position operators, and an interaction term:

$$H_{\text{int}} = -\frac{e}{2mc}\left[\hat{\mathbf{p}}\cdot\mathbf{A}(\hat{\mathbf{x}}) + \mathbf{A}(\hat{\mathbf{x}})\cdot\hat{\mathbf{p}}\right] + \frac{e^2}{2mc^2}\mathbf{A}^2(\hat{\mathbf{x}}) + e\phi(\hat{\mathbf{x}}). \tag{16.25}$$

The Coulomb gauge simplifies this expression. Acting on any wave function, we have

$$\nabla\cdot\left(\mathbf{A}\psi\right) = (\nabla\cdot\mathbf{A})\,\psi + \mathbf{A}\cdot\nabla\psi = \mathbf{A}\cdot\nabla\psi \quad \rightsquigarrow \quad \hat{\mathbf{p}}\cdot\mathbf{A}(\hat{\mathbf{x}}) = \mathbf{A}(\hat{\mathbf{x}})\cdot\hat{\mathbf{p}}. \tag{16.26}$$

The two orderings of $\hat{\mathbf{p}}$ and $\mathbf{A}$ coincide. Moreover, as we have seen in Eq. (16.9), the scalar potential in the Coulomb gauge is the instantaneous Coulomb potential of the charges. The Coulomb potential of the atomic charges is already part of U, and for an atom in a radiation field without additional external charges we can set $\phi = 0$. Then,

$$H_{\text{int}} = -\frac{e}{mc}\,\mathbf{A}(\hat{\mathbf{x}})\cdot\hat{\mathbf{p}} + \frac{e^2}{2mc^2}\mathbf{A}^2(\hat{\mathbf{x}}). \tag{16.27}$$

For an atom with many electrons, we simply sum this expression over all electrons. We neglect here the interaction of the electron spin with the magnetic field; for the transitions that we will study it is much weaker than the term linear in $\hat{\mathbf{p}}$.

It is useful to write the interaction in terms of a current density and a charge density operator,

$$\hat{\mathbf{J}}(\mathbf{x}) = \frac{1}{2m}\left[\hat{\mathbf{p}}\,\delta^{(3)}(\mathbf{x}-\hat{\mathbf{x}}) + \delta^{(3)}(\mathbf{x}-\hat{\mathbf{x}})\,\hat{\mathbf{p}}\right], \qquad \hat{\rho}(\mathbf{x}) = \delta^{(3)}(\mathbf{x}-\hat{\mathbf{x}}), \tag{16.28}$$

where $\mathbf{x}$ is an ordinary position vector and $\hat{\mathbf{x}}$ is the position operator. Notice that $\hat{\mathbf{J}}$ is a current of probability (a velocity times a density); the electric current is $e\hat{\mathbf{J}}$.
Exercise: *Show that the matrix element* $\left\langle\psi\right|\hat{\mathbf{J}}(\mathbf{x})\left|\psi\right\rangle$ *is the probability current* $\frac{\hbar}{2mi}\left(\psi^*\nabla\psi - \psi\nabla\psi^*\right)$ *of Part I.*
We can then write:

$$H_{\text{int}} = \int d^3\mathbf{x}\left[-\frac{e}{c}\,\hat{\mathbf{J}}(\mathbf{x})\cdot\mathbf{A}(\mathbf{x}, t) + \frac{e^2}{2mc^2}\,\mathbf{A}^2(\mathbf{x}, t)\,\hat{\rho}(\mathbf{x})\right]. \tag{16.29}$$

The coupling of the electron to the radiation is weak. The strength of the interaction is governed by the fine structure constant $\alpha = e^2/(\hbar c) \approx 1/137$, and we can treat $H_{\rm int}$ as a perturbation. In first-order perturbation theory, the term linear in $\mathbf{A}$ changes the number of photons by one, and we keep only this term. The $\mathbf{A}^2$ term is of order e^2; it contributes to processes with two photons, such as the scattering of light on atoms, which we will not discuss. Thus,

$$H_{\rm int} \approx -\frac{e}{c}\int d^3\mathbf{x}\, \hat{\mathbf{j}}(\mathbf{x})\cdot\mathbf{A}(\mathbf{x},t). \tag{16.30}$$

Substituting the mode expansion of Eq. (16.13), the space integral becomes a Fourier transform of the current:

$$\begin{aligned} H_{\rm int} &= -\frac{e}{c\sqrt{L^3}}\sum_{\mathbf{k},\lambda}\Big\{\alpha(\mathbf{k},\lambda)\,\hat{\mathbf{j}}(-\mathbf{k})\cdot\mathbf{e}(\mathbf{k},\lambda)\,e^{-i\omega_k t} \\ &\qquad + \alpha^*(\mathbf{k},\lambda)\,\hat{\mathbf{j}}(\mathbf{k})\cdot\mathbf{e}^*(\mathbf{k},\lambda)\,e^{i\omega_k t}\Big\}\,, \end{aligned} \tag{16.31}$$

with

$$\hat{\mathbf{j}}(\mathbf{k}) \equiv \int d^3\mathbf{x}\, e^{-i\mathbf{k}\cdot\mathbf{x}}\,\hat{\mathbf{j}}(\mathbf{x}) = \frac{1}{2}\left[\frac{\hat{\mathbf{p}}}{m}\,e^{-i\mathbf{k}\cdot\hat{\mathbf{x}}} + e^{-i\mathbf{k}\cdot\hat{\mathbf{x}}}\,\frac{\hat{\mathbf{p}}}{m}\right]. \tag{16.32}$$

Notice that $\hat{\mathbf{j}}(\mathbf{k})^\dagger = \hat{\mathbf{j}}(-\mathbf{k})$, so that the two terms in Eq. (16.31) are Hermitian conjugates of each other, and $H_{\rm int}$ is Hermitian.

The Hamiltonian of Eq. (16.31) is a sum of harmonic perturbations, exactly of the type that we studied in Section 14.5.6. Let us consider a single mode $(\mathbf{k},\lambda)$ and an atom which is initially in its ground state $|0\rangle$ with energy E_0. The term with $e^{-i\omega_k t}$ can bring the atom to an excited state $|n\rangle$ with energy $E_n = E_0 + \hbar\omega_k$: the atom **absorbs** energy from the wave. The golden rule gives the transition rate:

$$\Gamma_{0\to n} = \frac{2\pi}{\hbar}\,\frac{e^2}{c^2L^3}\,|\alpha(\mathbf{k},\lambda)|^2\left|\langle n|\,\hat{\mathbf{j}}(-\mathbf{k})\cdot\mathbf{e}(\mathbf{k},\lambda)\,|0\rangle\right|^2\delta\left(E_n - E_0 - \hbar\omega_k\right). \tag{16.33}$$

This is the absorption rate due to one mode. Real light, from the sun or from a lamp, is a superposition of very many modes with random phases. We assume that the radiation is **incoherent**, i.e. that interference terms between different modes average to zero. Then the rates of the individual modes simply add up:

$$\Gamma_{0\to n} = \frac{1}{L^3}\sum_{\mathbf{k},\lambda}\frac{2\pi e^2}{\hbar c^2}\,|\alpha(\mathbf{k},\lambda)|^2\left|\langle n|\,\hat{\mathbf{j}}(-\mathbf{k})\cdot\mathbf{e}(\mathbf{k},\lambda)\,|0\rangle\right|^2\delta\left(E_n - E_0 - \hbar\omega_k\right). \tag{16.34}$$

For a large box the allowed wave vectors become dense, and we can replace the sum by an integral:

$$\frac{1}{L^3}\sum_{\mathbf{k}} \to \int\frac{d^3\mathbf{k}}{(2\pi)^3} = \int\frac{dk\,k^2}{(2\pi)^3}\,d\Omega_k = \int\frac{d\omega\,\omega^2}{(2\pi c)^3}\,d\Omega_k. \tag{16.35}$$

Using $\delta(E_n - E_0 - \hbar\omega) = \frac{1}{\hbar}\delta(\omega - \omega_{n0})$, with

$$\omega_{n0} \equiv \frac{E_n - E_0}{\hbar}, \tag{16.36}$$

the frequency integral is trivial, and we obtain the absorption rate for radiation coming from the solid angle $d\Omega_k$:

$$d\Gamma_{0\to n} = \frac{2\pi e^2}{\hbar^2c^2}\,\frac{\omega_{n0}^2}{(2\pi c)^3}\sum_{\lambda}|\alpha(\mathbf{k},\lambda)|^2\left|\langle n|\,\hat{\mathbf{j}}(-\mathbf{k})\cdot\mathbf{e}(\mathbf{k},\lambda)\,|0\rangle\right|^2 d\Omega_k, \tag{16.37}$$

where now $|\mathbf{k}| = \omega_{n0}/c$.

The reverse process is also possible. An atom in the excited state $|n\rangle$ can make a transition to the ground state, giving its energy to the wave. This is caused by the term with $e^{+i\omega_k t}$ in Eq. (16.31). Repeating the same steps, we find

$$d\Gamma_{n\to 0} = \frac{2\pi e^2}{\hbar^2 c^2}\frac{\omega_{n0}^2}{(2\pi c)^3}\sum_\lambda |\alpha(\mathbf{k},\lambda)|^2 \left|\langle 0|\,\hat{\mathbf{J}}(\mathbf{k})\cdot\mathbf{e}^*(\mathbf{k},\lambda)\,|n\rangle\right|^2 d\Omega_k. \tag{16.38}$$

This process is called **induced** or **stimulated emission**. Since $\hat{\mathbf{J}}(\mathbf{k})\cdot\mathbf{e}^* = \left(\hat{\mathbf{J}}(-\mathbf{k})\cdot\mathbf{e}\right)^\dagger$, the two matrix elements are complex conjugates of each other, and we find that the rates for absorption and stimulated emission are identical:

$$d\Gamma_{n\to 0} = d\Gamma_{0\to n}. \tag{16.39}$$

Notice what this result also says. If there is no radiation, $\alpha = 0$, the excited atom stays excited forever. Within the framework of this section there is no spontaneous emission. We will come back to this embarrassment in Section 16.4.

16.2.1 Photon Absorption

In the above calculation the radiation was a classical wave. Without yet quantising the field, we can express the absorption rate in terms of the number of photons of Eq. (16.21), i.e. by substituting $|\alpha|^2 = 2\pi\hbar c^2 N/\omega_k$ into Eq. (16.33). We find:

$$\Gamma_{0\to n;\mathbf{k},\lambda} = \frac{4\pi^2 e^2}{\omega_k}\frac{N(\mathbf{k},\lambda)}{L^3}\left|\langle n|\,\hat{\mathbf{J}}(-\mathbf{k})\cdot\mathbf{e}(\mathbf{k},\lambda)\,|0\rangle\right|^2 \delta\left(E_n - E_0 - \hbar\omega_k\right). \tag{16.40}$$

The rate is proportional to the density of photons, N/L^3, in the mode. It is now natural to think of the absorption as a collision of a photon with the atom, which disappears in the process, and to characterise it, as in scattering theory (Section 15.5), by a **cross-section**. The flux of photons with wave vector $\mathbf{k}$ and polarisation λ is their density times their velocity,

$$j_{\text{photon}} = \frac{N(\mathbf{k},\lambda)}{L^3}\,c. \tag{16.41}$$

The absorption cross-section for photons of frequency ω and polarisation λ is the total absorption rate divided by the flux:

$$\begin{aligned}\sigma_{\text{abs},\lambda}(\omega) &= \sum_n \frac{\Gamma_{0\to n;\mathbf{k},\lambda}}{j_{\text{photon}}} \\ &= \frac{4\pi^2 e^2}{\omega\, c}\sum_n \left|\langle n|\,\hat{\mathbf{J}}(-\mathbf{k})\cdot\mathbf{e}(\mathbf{k},\lambda)\,|0\rangle\right|^2 \delta\left(E_n - E_0 - \hbar\omega\right).\end{aligned} \tag{16.42}$$

For unpolarised light we average over the two polarisations, $\sigma_{\text{abs}} = \frac{1}{2}\sum_\lambda \sigma_{\text{abs},\lambda}$.
Exercise: *Verify that $\sigma_{\text{abs},\lambda}$ has the dimensions of an area.*
To compute the cross-section we need, of course, the matrix element

$$\langle n|\,\hat{\mathbf{J}}(-\mathbf{k})\cdot\mathbf{e}(\mathbf{k},\lambda)\,|0\rangle \tag{16.43}$$

for the atom of interest. This is the subject of the next section.

16.3 Matrix Elements in the Dipole Approximation

Energy conservation alone would allow an atom to absorb any photon whose energy matches the difference of two energy levels. Look, however, at the spectrum of hydrogen. Many lines that are allowed by energy conservation are absent or extremely weak. For atoms, and for any system with a central potential, rotational symmetry restricts radiative transitions much further. We will find **selection rules**, which physically are a consequence of the conservation of angular momentum of the total system of the atom and the photon.

Let us look at the matrix element for a transition $|0\rangle \to |n\rangle$ in an atom with several electrons, labelled by j:

$$\langle n|\,\hat{\mathbf{J}}(-\mathbf{k})\cdot\mathbf{e}(\mathbf{k},\lambda)\,|0\rangle\,, \qquad \hat{\mathbf{J}}(-\mathbf{k}) = \sum_j \left[\frac{\hat{\mathbf{p}}_j}{2m}\, e^{i\mathbf{k}\cdot\hat{\mathbf{x}}_j} + e^{i\mathbf{k}\cdot\hat{\mathbf{x}}_j}\,\frac{\hat{\mathbf{p}}_j}{2m}\right]. \tag{16.44}$$

Let us Taylor expand the exponential:

$$\hat{\mathbf{J}}(-\mathbf{k}) = \sum_j \left\{\frac{\hat{\mathbf{p}}_j}{2m}\left(1 + i\mathbf{k}\cdot\hat{\mathbf{x}}_j + \ldots\right) + \left(1 + i\mathbf{k}\cdot\hat{\mathbf{x}}_j + \ldots\right)\frac{\hat{\mathbf{p}}_j}{2m}\right\}. \tag{16.45}$$

How large is $\mathbf{k}\cdot\hat{\mathbf{x}}_j$? The wave functions of the electrons extend over the size of the atom, about 1 Å, while the wavelength of visible light is several thousand Ångström. Let us estimate this for hydrogen. A photon from a transition between the levels $n = 2$ and $n = 1$ has the energy $\hbar\omega = \frac{3}{8}\alpha^2 mc^2$, and the Bohr radius is $a_0 = \hbar/(\alpha mc)$. Therefore,

$$ka_0 = \frac{\omega a_0}{c} = \frac{3}{8}\alpha \approx 3\times 10^{-3}. \tag{16.46}$$

The wavelength of the photon is much larger than the atom, and the electromagnetic wave is practically constant over the atom. It is then an excellent approximation to keep only the first term of the expansion:

$$\hat{\mathbf{J}}(-\mathbf{k}) \approx \sum_j \frac{\hat{\mathbf{p}}_j}{m} + O\left(\frac{r_{\text{atom}}}{\lambda_{\text{photon}}/2\pi}\right). \tag{16.47}$$

This is called the **electric dipole approximation**.

We can trade the momentum for the position operator. For a Hamiltonian H_0 without velocity-dependent potentials, we have

$$\left[\hat{\mathbf{x}}_j, H_0\right] = \left[\hat{\mathbf{x}}_j, \frac{\hat{\mathbf{p}}_j^2}{2m}\right] = \frac{i\hbar}{m}\,\hat{\mathbf{p}}_j, \tag{16.48}$$

and therefore

$$\begin{aligned}\langle n|\sum_j \frac{\hat{\mathbf{p}}_j}{m}|0\rangle &= \frac{1}{i\hbar}\sum_j \langle n|\,\hat{\mathbf{x}}_j H_0 - H_0\hat{\mathbf{x}}_j\,|0\rangle \\ &= \frac{E_0 - E_n}{i\hbar}\sum_j \langle n|\,\hat{\mathbf{x}}_j\,|0\rangle = i\,\omega_{n0}\sum_j \langle n|\,\hat{\mathbf{x}}_j\,|0\rangle\,.\end{aligned} \tag{16.49}$$

The matrix element of the current is reduced to the matrix element of the **electric dipole operator**, $\hat{\mathbf{d}} = e\sum_j \hat{\mathbf{x}}_j$. The name has a simple origin. Inserting Eq. (16.49) into the

absorption term of Eq. (16.31), with $\mathbf{k}\cdot\hat{\mathbf{x}}_j \to 0$, and recalling that the electric field of the wave is $\mathbf{E} = -\frac{1}{c}\partial_t \mathbf{A}$, we find that at resonance, $\omega_k = \omega_{n0}$,

$$\langle n|\, H_{\rm int}\, |0\rangle = \langle n|\, -\hat{\mathbf{d}}\cdot\mathbf{E}(\mathbf{0},t)\, |0\rangle\,, \tag{16.50}$$

which is the energy of an electric dipole in a homogeneous electric field.
Exercise: *Verify Eq. (16.50), keeping only the absorption term of* $\mathbf{E}$*, the one proportional to* $e^{-i\omega_k t}$.

Inserting Eq. (16.49) into Eq. (16.42), the absorption cross-section in the dipole approximation becomes

$$\sigma_{{\rm abs},\lambda}(\omega) = \frac{4\pi^2 e^2}{c}\,\omega \sum_n \left|\mathbf{e}(\mathbf{k},\lambda)\cdot\sum_j \langle n|\,\hat{\mathbf{x}}_j\,|0\rangle\right|^2 \delta\left(E_n - E_0 - \hbar\omega\right). \tag{16.51}$$

Exercise: *For a single electron and a polarisation* $\mathbf{e}$ *along the x-direction, prove the Thomas–Reiche–Kuhn sum rule* $\sum_n \frac{2m\,\omega_{n0}}{\hbar}\,|\langle n|\,\hat{x}\,|0\rangle|^2 = 1$ *by computing* $\langle 0|\,[\hat{x},[H_0,\hat{x}]]\,|0\rangle$ *in two ways. Then show that the absorption cross-section integrated over all frequencies is* $\int_0^\infty d\omega\,\sigma_{\rm abs}(\omega) = 2\pi^2 e^2/(mc)$*, independently of the details of the atom.*

We are now in a position to prove that $\langle n|\,\hat{\mathbf{x}}\,|0\rangle$ vanishes unless special conditions are met. For simplicity, we consider the case where only one electron takes part in the transition. In a central potential the energy eigenstates can be chosen as $|n,\ell,m\rangle$, where ℓ labels the representation of the rotation group and m is the magnetic quantum number. Physical intuition tells us that the matrix element, which originates from the interaction of an electron with a photon, should vanish if the total angular momentum of the "photon-electron" system is not conserved. Let us see how this comes about.

We need the relation of the position operator $\hat{\mathbf{x}} = (\hat{x}_1,\hat{x}_2,\hat{x}_3)$ to angular momentum. With $\hat{L}_k = \varepsilon_{ijk}\hat{x}_i\hat{p}_j$ we have:

$$\left[\hat{L}_k,\hat{x}_l\right] = \varepsilon_{ijk}\hat{x}_i\left[\hat{p}_j,\hat{x}_l\right] = -i\hbar\,\varepsilon_{ijk}\hat{x}_i\delta_{jl} = i\hbar\,\varepsilon_{ikl}\,\hat{x}_i. \tag{16.52}$$

Let us now define the "spherical" components of the position operator,

$$V_0 = \hat{x}_3, \qquad V_{+1} = -\frac{1}{\sqrt{2}}\left(\hat{x}_1 + i\hat{x}_2\right), \qquad V_{-1} = \frac{1}{\sqrt{2}}\left(\hat{x}_1 - i\hat{x}_2\right). \tag{16.53}$$

Up to a factor $\sqrt{4\pi/3}\,r$, these are the spherical harmonics Y_{1q}. Using Eq. (16.52), we can prove that

$$\left[L_z, V_q\right] = \hbar\, q\, V_q, \qquad \left[L_\pm, V_q\right] = \hbar\sqrt{(1\mp q)(2\pm q)}\,V_{q\pm1}, \qquad L_\pm \equiv L_x \pm iL_y. \tag{16.54}$$

Exercise: *Prove Eq. (16.54) for* $q = 0,\pm1$.
Operators with these commutation relations are called **spherical tensors of rank** 1 (vector operators). The momentum $\hat{\mathbf{p}}$ and the angular momentum $\hat{\mathbf{L}}$ itself are other examples.

We consider the two states of the transition,

$$|0\rangle = |n_1,\ell_1,m_1\rangle\,, \qquad |n\rangle = |n_2,\ell_2,m_2\rangle\,, \tag{16.55}$$

and the matrix elements of the spherical components V_q. The commutation relations of Eq. (16.54) put severe constraints on them. Sandwiching the first relation between the two states, we find:

$$\begin{aligned} 0 &= \langle n_2,\ell_2,m_2|\left[L_z,V_q\right] - \hbar q V_q\,|n_1,\ell_1,m_1\rangle \\ &= \langle n_2,\ell_2,m_2|\,L_z V_q - V_q L_z - \hbar q V_q\,|n_1,\ell_1,m_1\rangle \\ &= \hbar\left(m_2 - m_1 - q\right)\langle n_2,\ell_2,m_2|\,V_q\,|n_1,\ell_1,m_1\rangle\,. \end{aligned} \tag{16.56}$$

The matrix element vanishes, unless

$$\Delta m \equiv m_2 - m_1 = q = 0, \pm 1. \tag{16.57}$$

The polarisation decides which q contributes. Light which is linearly polarised along the z-axis, $\mathbf{e} \cdot \hat{\mathbf{x}} = V_0$, induces transitions with $\Delta m = 0$, while circularly polarised light propagating along the z-axis induces $\Delta m = +1$ or $\Delta m = -1$ (**exercise**).

Let us now sandwich the second relation of Eq. (16.54). Using $L_\pm |\ell, m\rangle = \hbar\sqrt{(\ell \mp m)(\ell \pm m + 1)}\,|\ell, m \pm 1\rangle$ and its Hermitian conjugate, we obtain:

$$\begin{aligned}
&\sqrt{(\ell_2 \pm m_2)(\ell_2 \mp m_2 + 1)}\ \langle n_2, \ell_2, m_2 \mp 1|\, V_q\, |n_1, \ell_1, m_1\rangle \\
&\quad = \sqrt{(\ell_1 \mp m_1)(\ell_1 \pm m_1 + 1)}\ \langle n_2, \ell_2, m_2|\, V_q\, |n_1, \ell_1, m_1 \pm 1\rangle \\
&\quad + \sqrt{(1 \mp q)(2 \pm q)}\ \langle n_2, \ell_2, m_2|\, V_{q\pm 1}\, |n_1, \ell_1, m_1\rangle\,.
\end{aligned} \tag{16.58}$$

These are recurrence relations among matrix elements with different magnetic quantum numbers. We have encountered exactly the same recurrence relations when we added two angular momenta, $\mathbf{J} = \mathbf{J}_A + \mathbf{J}_B$, in Chapter 12. The Clebsch–Gordan coefficients

$$\langle J_A, J_B; m_A, m_B | J_A, J_B; J, m\rangle\,, \tag{16.59}$$

which change the basis from the eigenstates of $\mathbf{J}_A^2, \mathbf{J}_B^2, J_{A,z}, J_{B,z}$ to the eigenstates of $\mathbf{J}_A^2, \mathbf{J}_B^2, \mathbf{J}^2, J_z$, satisfy the same recursion, obtained by acting with $J_\pm = J_{A,\pm} + J_{B,\pm}$, if we identify

$$J_A = \ell_1, \quad m_A = m_1, \qquad J_B = 1, \quad m_B = q, \qquad J = \ell_2, \quad m = m_2. \tag{16.60}$$

A recursion determines its solution up to an overall factor. Therefore, the matrix element is proportional to a Clebsch–Gordan coefficient,

$$\langle n_2, \ell_2, m_2|\, V_q\, |n_1, \ell_1, m_1\rangle = \langle \ell_1, 1; m_1, q | \ell_1, 1; \ell_2, m_2\rangle\ \langle n_2, \ell_2|\, \|V\|\, |n_1, \ell_1\rangle\,, \tag{16.61}$$

where the "reduced matrix element" $\langle n_2, \ell_2|\, \|V\|\, |n_1, \ell_1\rangle$ does not depend on m_1, m_2, q. This is a special case of the Wigner–Eckart theorem, which we have proven in Section 12.3 of Part I. All the dependence on the orientation of the atom is in a Clebsch–Gordan coefficient. The latter vanishes unless

$$|\ell_1 - 1| \le \ell_2 \le \ell_1 + 1, \qquad m_2 = m_1 + q, \qquad |q| \le 1. \tag{16.62}$$

The matrix element of a radiative transition is non-vanishing only if the angular momentum of the electron, (ℓ_1, m_1), and the angular momentum of the "photon", $(1, q)$, can be added up to (ℓ_2, m_2)! Notice that for $\ell_1 = 0$ we must have $\ell_2 = 1$. A transition from $\ell_1 = 0$ to $\ell_2 = 0$ is impossible: an atom in an s ground state is excited by absorbing a photon always to a p state.

We obtain a more stringent selection rule from parity. Under a parity transformation (Section 13.1), $\Pi \hat{\mathbf{x}} \Pi^\dagger = -\hat{\mathbf{x}}$, and the eigenstates of a central potential transform as the spherical harmonics,

$$\Pi\, |n, \ell, m\rangle = (-1)^\ell\, |n, \ell, m\rangle\,. \tag{16.63}$$

Inserting $\Pi^\dagger \Pi = 1$ on both sides of the position operator, and abbreviating $|1\rangle \equiv |n_1, \ell_1, m_1\rangle$, $|2\rangle \equiv |n_2, \ell_2, m_2\rangle$, we find:

$$\langle 2|\, \hat{\mathbf{x}}\, |1\rangle = \langle 2|\, \Pi^\dagger \left(\Pi \hat{\mathbf{x}} \Pi^\dagger\right) \Pi\, |1\rangle = -(-1)^{\ell_1 + \ell_2}\, \langle 2|\, \hat{\mathbf{x}}\, |1\rangle\,. \tag{16.64}$$

The matrix element survives only if $(-1)^{\ell_1+\ell_2+1} = 1$, i.e. if the two states have opposite parity. Together with Eq. (16.62), this leaves

$$\Delta\ell \equiv \ell_2 - \ell_1 = \pm 1, \qquad \Delta m = 0, \pm 1. \tag{16.65}$$

These are the **selection rules** for electric dipole transitions. They reflect the constraints of rotational symmetry and parity. Transitions are not allowed if the angular momenta of the photon and the electron do not add up together.

Exercise: *Compute the angular integrals* $\int d\Omega\, Y^*_{\ell_2 m_2} \frac{V_q}{r} Y_{\ell_1 m_1}$ *for* $\ell_1, \ell_2 \le 2$ *and verify Eqs.* (16.61) *and* (16.65).

Notice that the selection rules hold in the dipole approximation. They do not forbid the "forbidden" transitions completely, but make them very rare. The higher terms of the expansion in Eq. (16.45) (magnetic dipole, electric quadrupole, ...) are suppressed by powers of $k\, r_{\text{atom}} \sim \alpha$. The $2s$ state of hydrogen, for example, cannot decay to $1s$ by emitting one dipole photon. It decays mostly by emitting two photons, with a lifetime of about 0.1 s, which is eight orders of magnitude longer than the lifetime of the $2p$ state.

16.4 Quantisation of the Radiation Field

We have decomposed the radiation field into modes $(\mathbf{k}, \lambda)$ and we have associated a number of photons $N(\mathbf{k}, \lambda)$ with each mode. A state of the radiation with many photons would then be described by the occupation numbers of all modes,

$$|N(\mathbf{k}_1, \lambda_1), N(\mathbf{k}_2, \lambda_2), \ldots\rangle\,, \qquad E_{\text{ph}} = \sum_j N(\mathbf{k}_j, \lambda_j)\hbar c \left|\mathbf{k}_j\right|. \tag{16.66}$$

Our treatment so far has two shortcomings. First, we have not explained how the radiation field is quantised into photons; we have only imposed it by hand, through Eq. (16.21). Second, we have paid no attention to what happens to the radiation itself in a transition, and we have found no spontaneous emission. In this section we will develop the formalism for the quantisation of the electromagnetic field and give a precise meaning to the emission and absorption of photons.

16.4.1 Canonical Quantisation

We return to the free field, which satisfies the wave equation $\Box\mathbf{A} = 0$ in the Coulomb gauge. We write the solution of Eq. (16.13) as

$$\mathbf{A}(\mathbf{x}, t) = \frac{1}{\sqrt{L^3}} \sum_{\mathbf{k},\lambda} \left\{ q_{\mathbf{k}\lambda}(t)\, \mathbf{e}(\mathbf{k}, \lambda)\, e^{i\mathbf{k}\cdot\mathbf{x}} + q^*_{\mathbf{k}\lambda}(t)\, \mathbf{e}^*(\mathbf{k}, \lambda)\, e^{-i\mathbf{k}\cdot\mathbf{x}} \right\}, \tag{16.67}$$

with

$$q_{\mathbf{k}\lambda}(t) \equiv \alpha(\mathbf{k}, \lambda)\, e^{-i\omega_k t}. \tag{16.68}$$

Observe that

$$\frac{d^2}{dt^2} q_{\mathbf{k}\lambda} + \omega_k^2\, q_{\mathbf{k}\lambda} = 0. \tag{16.69}$$

The field $\mathbf{A}(\mathbf{x}, t)$ is a sum of simple harmonic oscillators, one for every mode. Let us express the energy of the field in terms of the variables $q_{\mathbf{k}\lambda}$. The fields are

$$\begin{aligned} \mathbf{E}(\mathbf{x}, t) &= \frac{1}{\sqrt{L^3}} \sum_{\mathbf{k},\lambda} \frac{i\omega_k}{c} \left\{ q_{\mathbf{k}\lambda}\, \mathbf{e}\, e^{i\mathbf{k}\cdot\mathbf{x}} - q^*_{\mathbf{k}\lambda}\, \mathbf{e}^*\, e^{-i\mathbf{k}\cdot\mathbf{x}} \right\}, \\ \mathbf{B}(\mathbf{x}, t) &= \frac{i}{\sqrt{L^3}} \sum_{\mathbf{k},\lambda} \left\{ q_{\mathbf{k}\lambda}\, (\mathbf{k}\times\mathbf{e})\, e^{i\mathbf{k}\cdot\mathbf{x}} - q^*_{\mathbf{k}\lambda}\, (\mathbf{k}\times\mathbf{e}^*)\, e^{-i\mathbf{k}\cdot\mathbf{x}} \right\}, \end{aligned} \tag{16.70}$$

and, as in Eq. (16.19), we find

$$E_{\rm SF} = \int d^3\mathbf{x}\, \frac{\mathbf{E}^2 + \mathbf{B}^2}{8\pi} = \sum_{\mathbf{k},\lambda} \frac{\omega_k^2}{4\pi c^2} \left\{ q_{\mathbf{k}\lambda} q^*_{\mathbf{k}\lambda} + q^*_{\mathbf{k}\lambda} q_{\mathbf{k}\lambda} \right\}. \tag{16.71}$$

It is natural to rescale the variables so that the energy of each mode is measured in units of $\hbar\omega_k$. We define

$$a_{\mathbf{k}\lambda}(t) \equiv \sqrt{\frac{\omega_k}{2\pi\hbar c^2}}\, q_{\mathbf{k}\lambda}(t), \tag{16.72}$$

and the energy becomes

$$E_{\rm SF} = \sum_{\mathbf{k},\lambda} \frac{\hbar\omega_k}{2} \left\{ a_{\mathbf{k}\lambda} a^*_{\mathbf{k}\lambda} + a^*_{\mathbf{k}\lambda} a_{\mathbf{k}\lambda} \right\}. \tag{16.73}$$

We have stopped short of writing $aa^* = a^*a = |a|^2$, which is certainly true for a classical field. For a classical field, $|a_{\mathbf{k}\lambda}|^2$ is the number $N(\mathbf{k},\lambda)$ of Eq. (16.21). Experimentally, however, we know that the energy of each mode comes in units of $\hbar\omega_k$, while $|a|^2$ can take any real value. As long as a and a^* are ordinary complex numbers, the energy of Eq. (16.73) is not quantised.

Eq. (16.73) is, however, very inspiring. It is exactly the form of the energy of a harmonic oscillator, written in terms of its ladder operators (Chapter 7). We thus do something very drastic, which goes under the name of **second quantisation**, or canonical quantisation of the field:

1. $a_{\mathbf{k}\lambda}$ and $a^*_{\mathbf{k}\lambda}$ are not numbers but operators, $\hat{a}_{\mathbf{k}\lambda}$ and $\hat{a}^\dagger_{\mathbf{k}\lambda}$.
2. They obey the commutation relations

$$\left[\hat{a}_{\mathbf{k}\lambda}, \hat{a}^\dagger_{\mathbf{k}'\lambda'}\right] = \delta_{\mathbf{k}\mathbf{k}'}\delta_{\lambda\lambda'}, \qquad \left[\hat{a}_{\mathbf{k}\lambda}, \hat{a}_{\mathbf{k}'\lambda'}\right] = \left[\hat{a}^\dagger_{\mathbf{k}\lambda}, \hat{a}^\dagger_{\mathbf{k}'\lambda'}\right] = 0. \tag{16.74}$$

So far in quantum mechanics we have postulated commutation relations for the position and momentum operators of particles. We now go a step beyond, and promote the classical electromagnetic field itself into a superposition of operators.

Exercise: *Define, for each mode, the Hermitian operators* $\hat{Q} = \sqrt{\hbar/(2\omega_k)}\,(\hat{a} + \hat{a}^\dagger)$ *and* $\hat{P} = i\sqrt{\hbar\omega_k/2}\,(\hat{a}^\dagger - \hat{a})$. *Show that* $\left[\hat{Q}, \hat{P}\right] = i\hbar$ *and that the energy of the mode is* $\frac{1}{2}(\hat{P}^2 + \omega_k^2\hat{Q}^2)$. *Each mode of the field is a harmonic oscillator of unit mass, quantised exactly as in Part I.*

The energy is promoted to the Hamiltonian operator of the electromagnetic field. Using the commutation relations, we find:

$$\begin{aligned} \hat{H}_{\rm SF} &= \sum_{\mathbf{k},\lambda} \frac{\hbar\omega_k}{2} \left\{ \hat{a}_{\mathbf{k}\lambda} \hat{a}^\dagger_{\mathbf{k}\lambda} + \hat{a}^\dagger_{\mathbf{k}\lambda} \hat{a}_{\mathbf{k}\lambda} \right\} \\ &= \sum_{\mathbf{k},\lambda} \hbar\omega_k \left\{ \hat{N}(\mathbf{k},\lambda) + \frac{1}{2} \right\}, \qquad \hat{N}(\mathbf{k},\lambda) \equiv \hat{a}^\dagger_{\mathbf{k}\lambda} \hat{a}_{\mathbf{k}\lambda}. \end{aligned} \tag{16.75}$$

The electromagnetic field is a collection of independent, quantum mechanical harmonic oscillators, one for each mode. From Part I we know everything about them. The

eigenvalues of the number operator $\hat{N}(\mathbf{k},\lambda)$ are the non-negative integers $N_{\mathbf{k}\lambda} = 0, 1, 2, \ldots$, and the energy of each mode is discrete,

$$E_{\mathbf{k}\lambda} = \hbar\omega_k \left(N_{\mathbf{k}\lambda} + \frac{1}{2} \right). \tag{16.76}$$

We interpret the eigenvalue $N_{\mathbf{k}\lambda}$ as the number of photons in the mode $(\mathbf{k},\lambda)$. Photons are the quanta of the oscillators of the field. Notice that the "$\frac{1}{2}$" gives a zero-point energy to every mode, and the sum over infinitely many modes diverges. Only energy differences are measured in our experiments, and we can subtract this constant from $\hat{H}_{\rm SF}$.[3]

Once we have transformed the problem into a collection of quantised harmonic oscillators, it is very easy to "give life" to photons. We assume the existence of a state without any photons in any mode, the **vacuum**,

$$|0\rangle\,, \qquad \hat{a}_{\mathbf{k}\lambda}\,|0\rangle = 0 \quad \text{for all } (\mathbf{k},\lambda). \tag{16.77}$$

A one-photon state with momentum $\hbar\mathbf{k}$ and polarisation λ is

$$|1(\mathbf{k},\lambda)\rangle = \hat{a}^\dagger_{\mathbf{k}\lambda}\,|0\rangle\,, \tag{16.78}$$

and successive applications of the **creation operators** $\hat{a}^\dagger_{\mathbf{k}\lambda}$ accumulate more photons. A general multi-photon state is

$$|N_{\mathbf{k}_1\lambda_1}, N_{\mathbf{k}_2\lambda_2}, \ldots\rangle = \frac{\left(\hat{a}^\dagger_{\mathbf{k}_1\lambda_1}\right)^{N_{\mathbf{k}_1\lambda_1}} \left(\hat{a}^\dagger_{\mathbf{k}_2\lambda_2}\right)^{N_{\mathbf{k}_2\lambda_2}} \ldots}{\sqrt{N_{\mathbf{k}_1\lambda_1}!\; N_{\mathbf{k}_2\lambda_2}! \cdots}}\,|0\rangle\,. \tag{16.79}$$

The annihilation and creation operators act on it as on the states of a harmonic oscillator. Writing only the occupation number $N \equiv N_{\mathbf{k}\lambda}$ of the mode on which they act, we have

$$\begin{aligned} \hat{a}_{\mathbf{k}\lambda}\,|\ldots, N, \ldots\rangle &= \sqrt{N}\;|\ldots, N-1, \ldots\rangle\,, \\ \hat{a}^\dagger_{\mathbf{k}\lambda}\,|\ldots, N, \ldots\rangle &= \sqrt{N+1}\;|\ldots, N+1, \ldots\rangle\,. \end{aligned} \tag{16.80}$$

Photons with the same $(\mathbf{k},\lambda)$ are indistinguishable: the state is specified by how many photons occupy each mode, and not by which photon is where. The photons are bosons; we will come back to this in Chapters 17 and 19.

16.4.2 Field Operators

In this language the vector potential is a linear combination of the operators $\hat{a}$ and $\hat{a}^\dagger$. Inverting Eq. (16.72) at $t = 0$, we find the field operator

$$\hat{\mathbf{A}}(\mathbf{x}) = \frac{1}{\sqrt{L^3}} \sum_{\mathbf{k},\lambda} \sqrt{\frac{2\pi\hbar c^2}{\omega_k}} \left[\hat{a}_{\mathbf{k}\lambda}\,\mathbf{e}(\mathbf{k},\lambda)\, e^{i\mathbf{k}\cdot\mathbf{x}} + \hat{a}^\dagger_{\mathbf{k}\lambda}\,\mathbf{e}^*(\mathbf{k},\lambda)\, e^{-i\mathbf{k}\cdot\mathbf{x}} \right]. \tag{16.81}$$

[3] The zero-point energy is not entirely without physical consequences. Its dependence on the boundary conditions gives rise to the Casimir force between two neutral conducting plates; see the course on Quantum Field Theory.

The time dependence $e^{\mp i\omega_k t}$ of the classical solution is now carried by the states, through the Hamiltonian $\hat{H}_{\rm SF}$.
Exercise: *Show that in the Heisenberg picture*

$$\hat{a}_{\mathbf{k}\lambda}(t) = e^{i\hat{H}_{\rm SF}t/\hbar}\,\hat{a}_{\mathbf{k}\lambda}\,e^{-i\hat{H}_{\rm SF}t/\hbar} = \hat{a}_{\mathbf{k}\lambda}\,e^{-i\omega_k t}\,,$$

and that $\hat{\mathbf{A}}(\mathbf{x},t)$ *satisfies the wave equation.*

We can now revisit the transition rates. The full system consists of the atom and the radiation, with the Hamiltonian $H = H_0 + \hat{H}_{\rm SF} + H_{\rm int}$. How does the interaction term change when we promote $\mathbf{A}$ to an operator? It becomes

$$\begin{aligned} H_{\rm int} &= -\frac{e}{c}\int d^3\mathbf{x}\,\hat{\mathbf{j}}(\mathbf{x})\cdot\hat{\mathbf{A}}(\mathbf{x}) + O(e^2) \\ &= -\frac{e}{\sqrt{L^3}}\sum_{\mathbf{k},\lambda}\sqrt{\frac{2\pi\hbar}{\omega_k}}\left[\hat{\mathbf{j}}(-\mathbf{k})\cdot\mathbf{e}(\mathbf{k},\lambda)\,\hat{a}_{\mathbf{k}\lambda}\right. \\ &\qquad\left. + \hat{\mathbf{j}}(\mathbf{k})\cdot\mathbf{e}^*(\mathbf{k},\lambda)\,\hat{a}^\dagger_{\mathbf{k}\lambda}\right]. \end{aligned} \tag{16.82}$$

It acts both on the atom, through $\hat{\mathbf{j}}$, and on the radiation, through $\hat{a}$ and $\hat{a}^\dagger$. The first term absorbs a photon; the second term creates one.

16.4.3 Photon Emission and Absorption

Let us consider an atom in its ground state $|0\rangle$, in a radiation field with $N_{\mathbf{k}\lambda}$ photons in the mode $(\mathbf{k},\lambda)$. In an absorption the atom is excited to $|n\rangle$ and one photon disappears:

$$\underbrace{|0\rangle}_{\text{atom}}\otimes\underbrace{|\ldots,N_{\mathbf{k}\lambda},\ldots\rangle}_{\text{photons}} \;\to\; |n\rangle\otimes|\ldots,N_{\mathbf{k}\lambda}-1,\ldots\rangle\,. \tag{16.83}$$

The absorbed photon must have the correct energy, $\hbar\omega_k = E_n - E_0$. The matrix element of the interaction between the initial and the final state is

$$\begin{aligned} M_{0\to n} &= \left(\langle n|\otimes\langle\ldots,N_{\mathbf{k}\lambda}-1,\ldots|\right) H_{\rm int}\left(|0\rangle\otimes|\ldots,N_{\mathbf{k}\lambda},\ldots\rangle\right) \\ &= -e\sqrt{\frac{2\pi\hbar}{\omega_k L^3}}\,\langle n|\hat{\mathbf{j}}(-\mathbf{k})\cdot\mathbf{e}(\mathbf{k},\lambda)\,|0\rangle \\ &\quad\times\underbrace{\left\langle\ldots,N_{\mathbf{k}\lambda}-1,\ldots\right|\hat{a}_{\mathbf{k}\lambda}\left|\ldots,N_{\mathbf{k}\lambda},\ldots\right\rangle}_{\sqrt{N_{\mathbf{k}\lambda}}}. \end{aligned} \tag{16.84}$$

The golden rule, $\Gamma = \frac{2\pi}{\hbar}|M|^2\delta(E_f - E_i)$, gives the absorption rate:

$$\Gamma_{0\to n;\mathbf{k}\lambda} = \frac{4\pi^2e^2}{\omega_k L^3}\left|\langle n|\hat{\mathbf{j}}(-\mathbf{k})\cdot\mathbf{e}(\mathbf{k},\lambda)\,|0\rangle\right|^2 N_{\mathbf{k}\lambda}\;\delta\left(E_n - E_0 - \hbar\omega_k\right). \tag{16.85}$$

This is exactly the result of Eq. (16.40), which we had obtained with a classical field and the heuristic photon number of Eq. (16.21).

Similarly, we compute the matrix element for the emission of a photon, which is the reverse process:

$$\begin{aligned} M_{n\to 0} &= \left(\langle 0|\otimes\langle\ldots,N_{\mathbf{k}\lambda}+1,\ldots|\right) H_{\text{int}}\left(|n\rangle\otimes|\ldots,N_{\mathbf{k}\lambda},\ldots\rangle\right) \\ &= -e\sqrt{\frac{2\pi\hbar}{\omega_k L^3}}\,\langle 0|\,\hat{\mathbf{J}}(\mathbf{k})\cdot\mathbf{e}^*(\mathbf{k},\lambda)\,|n\rangle \\ &\quad\times\underbrace{\left\langle\ldots,N_{\mathbf{k}\lambda}+1,\ldots\right|\hat{a}^\dagger_{\mathbf{k}\lambda}\left|\ldots,N_{\mathbf{k}\lambda},\ldots\right\rangle}_{\sqrt{N_{\mathbf{k}\lambda}+1}}, \end{aligned} \tag{16.86}$$

and the emission rate

$$\Gamma_{n\to 0;\mathbf{k}\lambda} = \frac{4\pi^2 e^2}{\omega_k L^3}\left|\langle 0|\,\hat{\mathbf{J}}(\mathbf{k})\cdot\mathbf{e}^*(\mathbf{k},\lambda)\,|n\rangle\right|^2\left(N_{\mathbf{k}\lambda}+1\right)\delta\left(E_n-E_0-\hbar\omega_k\right). \tag{16.87}$$

With a classical field we had found that emission and absorption rates are equal, Eq. (16.39). After quantising the field this is no longer true. The atomic matrix elements are the same, but the photon matrix elements differ, and the ratio of the rates is

$$\frac{\Gamma_{n\to 0;\mathbf{k}\lambda}}{\Gamma_{0\to n;\mathbf{k}\lambda}} = \frac{N_{\mathbf{k}\lambda}+1}{N_{\mathbf{k}\lambda}}. \tag{16.88}$$

For a large number of photons the ratio is very close to one, and the classical description is an excellent approximation. The two theories differ for a small number of photons. Consider an empty mode, $N_{\mathbf{k}\lambda}=0$. Both theories predict, correctly, that nothing can be absorbed. The classical theory also predicts that an excited atom remains excited, since the emission rate is zero. Here the quantised theory makes a remarkable prediction. The emission rate is proportional to $N_{\mathbf{k}\lambda}+1$ and does not vanish: an excited atom decays to the ground state even in the absence of any radiation. This is **spontaneous emission**, and it is of course what happens in nature. We can now answer the question at the beginning of this chapter. The state $|2,1,m\rangle\otimes|0\rangle$ is not an eigenstate of the full Hamiltonian $H_0+\hat{H}_{\text{SF}}+H_{\text{int}}$, because H_{int} connects it to the states $|1,0,0\rangle\otimes|1(\mathbf{k},\lambda)\rangle$. True quantum mechanics requires the quantisation of the electromagnetic field.

The "+1" in Eq. (16.88) has another important consequence. The emission into a mode which is already occupied by N photons is enhanced by the factor $N+1$, and the emitted photon has exactly the same $(\mathbf{k},\lambda)$ as the photons that stimulated it. This is the principle of the laser.

16.4.4 Spontaneous Emission and the Lifetime of the $2p$ State of Hydrogen

Let us compute the rate of spontaneous emission. We set $N_{\mathbf{k}\lambda}=0$ in Eq. (16.87) and sum over all modes, since the photon can be emitted in any direction and with any polarisation. Using Eq. (16.35), we have

$$\Gamma^{\text{spont}}_{n\to 0} = \sum_\lambda\int\frac{d\omega\,\omega^2}{(2\pi c)^3}\,d\Omega_k\,\frac{4\pi^2e^2}{\omega}\left|\langle 0|\,\hat{\mathbf{J}}(\mathbf{k})\cdot\mathbf{e}^*(\mathbf{k},\lambda)\,|n\rangle\right|^2\frac{1}{\hbar}\delta\left(\omega-\omega_{n0}\right). \tag{16.89}$$

In the dipole approximation, as in Eq. (16.49), we have $\langle 0|\,\hat{\mathbf{p}}/m\,|n\rangle = -i\omega_{n0}\,\mathbf{x}_{0n}$, with $\mathbf{x}_{0n}\equiv\langle 0|\,\hat{\mathbf{x}}\,|n\rangle$. We then find the angular distribution of the emitted photons,

$$\frac{d\Gamma^{\text{spont}}_{n\to 0}}{d\Omega_k} = \frac{e^2\omega^3_{n0}}{2\pi\hbar c^3}\sum_\lambda\left|\mathbf{e}^*(\mathbf{k},\lambda)\cdot\mathbf{x}_{0n}\right|^2. \tag{16.90}$$

The sum over polarisations is done with the identity $\sum_\lambda e_i^* e_j = \delta_{ij} - \hat{k}_i \hat{k}_j$, and the remaining angular integral with $\int d\Omega_k \, \hat{k}_i \hat{k}_j = \frac{4\pi}{3}\delta_{ij}$. Therefore,

$$\sum_\lambda \int d\Omega_k \, |\mathbf{e}^*(\mathbf{k},\lambda)\cdot \mathbf{x}_{0n}|^2 = \left(4\pi - \frac{4\pi}{3}\right)|\mathbf{x}_{0n}|^2 = \frac{8\pi}{3}\,|\mathbf{x}_{0n}|^2 \,,$$

and we obtain:

$$\Gamma^{\text{spont}}_{n\to 0} = \frac{4}{3}\,\frac{e^2\omega_{n0}^3}{\hbar c^3}\,|\mathbf{x}_{0n}|^2 = \frac{4}{3}\,\alpha\,\frac{\omega_{n0}^3}{c^2}\,|\langle 0|\,\hat{\mathbf{x}}\,|n\rangle|^2 \,. \tag{16.91}$$

The rate grows with the third power of the frequency. Excited states of atoms, which emit visible or ultraviolet light, live for nanoseconds; excited states of molecules which emit in the infrared or microwave region live much longer.

Let us apply this to the transition $2p \to 1s$ of hydrogen, the Lyman-α line. The frequency is $\omega = \frac{3}{8}\alpha^2 mc^2/\hbar$. The matrix element is computed with the wave functions $\psi_{100} = R_{10}(r)Y_{00}$ and $\psi_{21m} = R_{21}(r)Y_{1m}$, where $R_{10} = 2a_0^{-3/2}e^{-r/a_0}$ and $R_{21} = (24)^{-1/2}a_0^{-3/2}(r/a_0)e^{-r/(2a_0)}$. The radial integral is

$$\int_0^\infty dr\, r^3 R_{10}(r) R_{21}(r) = \frac{128\sqrt{6}}{243}\,a_0, \tag{16.92}$$

and the angular part, by the Wigner–Eckart theorem, gives a factor $1/3$ for every m (the atom does not know about our choice of the z-axis). We find

$$|\langle 1,0,0|\,\hat{\mathbf{x}}\,|2,1,m\rangle|^2 = \frac{2^{15}}{3^{10}}\,a_0^2 \approx 0.555\,a_0^2. \tag{16.93}$$

Exercise: *Verify the last two equations.*
Substituting into Eq. (16.91), with $a_0 = \hbar/(\alpha mc)$, all the numerical factors combine into a remarkably simple result:

$$\Gamma(2p\to 1s) = \left(\frac{2}{3}\right)^8 \alpha^5\,\frac{mc^2}{\hbar} \approx 6.3\times 10^8\,\text{s}^{-1}, \qquad \tau_{2p} = \frac{1}{\Gamma} \approx 1.6\,\text{ns}. \tag{16.94}$$

This agrees with the measured lifetime. The $2p$ state can decay only to $1s$; the decay to $2s$ is allowed by the selection rules, but the two levels are degenerate (up to the tiny Lamb shift) and the factor ω^3 suppresses this decay to irrelevance. Notice also the power α^5: the decay rate is tiny compared to the frequency of the emitted light, $\Gamma/\omega_{n0} = \frac{8}{3}\left(\frac{2}{3}\right)^8\alpha^3 \approx 4\times 10^{-8}$. The excited state oscillates millions of times before it decays, which justifies our use of first-order perturbation theory and of the golden rule. The finite lifetime gives the spectral line a natural width $\hbar\Gamma \approx 4\times 10^{-7}$ eV.

16.4.5 Einstein Coefficients and Planck's Law

We have come full circle. Consider atoms in a cavity filled with thermal radiation at temperature T. The radiation is isotropic and unpolarised, and the occupation number $N(\omega)$ of a mode depends only on its frequency. The energy density of the radiation per unit frequency is obtained by counting the modes, two polarisations times $4\pi k^2 dk/(2\pi)^3$ per unit volume:

$$u(\omega)\,d\omega = \hbar\omega\, N(\omega)\,\frac{\omega^2 d\omega}{\pi^2 c^3}. \tag{16.95}$$

Summing Eqs. (16.85) and (16.87) over all modes, exactly as for the spontaneous emission, we find for the rates of absorption, stimulated emission and spontaneous emission between two levels:

$$\Gamma_{\text{abs}} = \Gamma_{\text{stim}} = A\, N(\omega_{n0}) = B\, u(\omega_{n0}), \qquad \Gamma_{\text{spont}} = A, \tag{16.96}$$

where A is the spontaneous rate of Eq. (16.91) and

$$B = \frac{4\pi^2 e^2}{3\hbar^2} |\mathbf{x}_{0n}|^2, \qquad \frac{A}{B} = \frac{\hbar\omega_{n0}^3}{\pi^2 c^3}. \tag{16.97}$$

These are the **Einstein coefficients**. Einstein postulated them in 1916, together with the relation between A and B; here we have derived them.
Exercise: *Derive Eq.* (16.96).
In thermal equilibrium the number of atoms that make a transition up per unit time must equal the number of those that make a transition down. The populations of the two levels are in the ratio of the Boltzmann factors, $P_n/P_0 = e^{-\hbar\omega_{n0}/k_B T}$ (for simplicity we take two non-degenerate levels). Then,

$$P_0\, A\, N = P_n\, A\, (N+1) \quad \rightsquigarrow \quad N(\omega) = \frac{1}{e^{\hbar\omega/k_B T} - 1}, \tag{16.98}$$

and the energy density of the radiation is

$$u(\omega) = \frac{\hbar\omega^3}{\pi^2 c^3} \frac{1}{e^{\hbar\omega/k_B T} - 1}. \tag{16.99}$$

This is Planck's law of black-body radiation! The "+1" of spontaneous emission is precisely what is needed to obtain Planck's spectrum. Without it, equilibrium would require $P_0 = P_n$ for any amount of radiation, which is impossible at finite temperature. The problem with which quantum mechanics began is solved by the quantisation of the electromagnetic field.

17

Identical Particles

17.1 Introduction

In classical mechanics, particles with identical properties (mass, charge, etc.) can be distinguished by their trajectories in space-time. In quantum mechanics, such a distinction is not possible. Take as an example the scattering of two identical particles that start from positions (1) and (2) and are detected some time later at positions (3) and (4), as in Fig. 17.1. We cannot tell whether, after the scattering, particle (1) arrives at (3) and particle (2) at

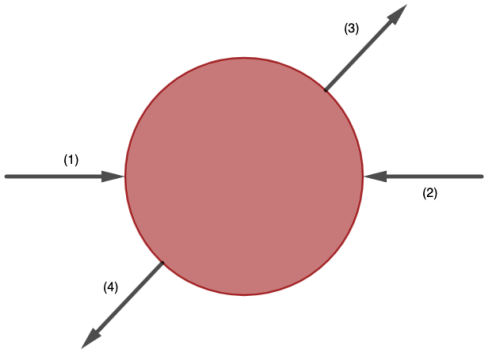


Figure 17.1. Scattering of two identical particles, starting at positions (1) and (2) and detected at positions (3) and (4).

(4), or particle (1) arrives at (4) and particle (2) at (3). The two possibilities are sketched in Fig. 17.2. When the two incoming particles come close to each other, the uncertainty

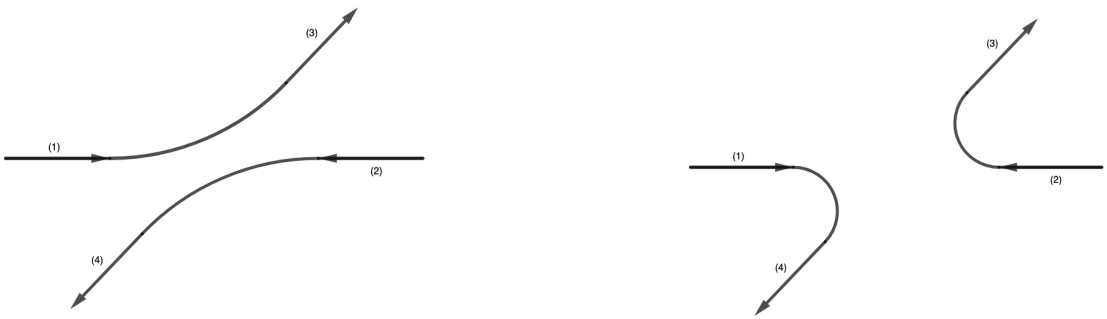


Figure 17.2. The two indistinguishable histories of the scattering process of Fig. 17.1.

principle prevents us from measuring their positions precisely enough to tell them apart.

Let us consider two identical non-interacting particles, where (1) is in the state $|\psi'\rangle$ and (2) in the state $|\psi''\rangle$. A state for the combined system is denoted by

$$|1\rangle \equiv |\psi'\rangle\,|\psi''\rangle\,. \tag{17.1}$$

This is a common shorthand for the tensor product of the two one-particle states: $|\psi'\rangle\,|\psi''\rangle = |\psi'\rangle \otimes |\psi''\rangle$. Since we cannot distinguish particles (1) and (2), the state in which the two particles have interchanged wave functions,

$$|2\rangle \equiv |\psi''\rangle\,|\psi'\rangle\,, \tag{17.2}$$

must yield the same results as $|1\rangle$ for all possible measurements. Therefore, $|1\rangle$ and $|2\rangle$ are degenerate. Nevertheless, the two states are in general quite different. To see this, we calculate their inner product, which factorises into a product of one-particle inner products:

$$\langle 2|1\rangle = \langle\psi''|\psi'\rangle\,\langle\psi'|\psi''\rangle = \left|\langle\psi''|\psi'\rangle\right|^2. \tag{17.3}$$

Note that

$$\langle 2|1\rangle = 0, \quad \text{if } |\psi'\rangle \text{ and } |\psi''\rangle \text{ are orthogonal.} \tag{17.4}$$

More generally, any superposition of $|1\rangle$ and $|2\rangle$,

$$|n\rangle = \alpha\left(|\psi'\rangle\,|\psi''\rangle\right) + \beta\left(|\psi''\rangle\,|\psi'\rangle\right). \tag{17.5}$$

describes the same physical situation. These states are degenerate for all α, β, and they all yield identical results. This degeneracy is called the *exchange degeneracy*. For a single isolated particle, we could determine the state by measuring a complete set of commuting observables. This is no longer possible for systems of two or more identical particles, since no measurement can remove the exchange degeneracy.

17.2 Permutation Operator

We now define an operator that maps a state onto the state with the identical particles interchanged. To this end, we define the *permutation operator* P_{12}, which exchanges particles (1) and (2), and analogously P_{21}:

$$\begin{aligned} P_{12}\left(|\psi'\rangle\,|\psi''\rangle\right) &= |\psi''\rangle\,|\psi'\rangle\,, \\ P_{21}\left(|\psi''\rangle\,|\psi'\rangle\right) &= |\psi'\rangle\,|\psi''\rangle\,. \end{aligned} \tag{17.6}$$

Formally, this is an operator on a tensor product. Let us denote the one-particle Hilbert space by $\mathcal{H}$. Since the two particles are identical, the Hilbert space of the total system is $\mathcal{H}\otimes\mathcal{H}$, and the permutation operator acts on it as

$$P_{12} : \mathcal{H}\otimes\mathcal{H} \to \mathcal{H}\otimes\mathcal{H}, \qquad |\psi'\rangle\,|\psi''\rangle \mapsto |\psi''\rangle\,|\psi'\rangle\,. \tag{17.7}$$

Exchanging the two particles twice gives back the original state, and it does not matter in which order we name them. Hence

$$P_{12} = P_{21} \quad \text{and} \quad P_{12}^2 = \mathbb{1}. \tag{17.8}$$

We are now interested in the eigenstates of the permutation operator, which, as we will see, are of great importance because of their (anti)symmetry. An eigenstate $|n\rangle$ of P_{12} fulfils

$$P_{12}\,|n\rangle = \lambda\,|n\rangle\,. \tag{17.9}$$

Using Eq. (17.8), we deduce that the eigenvalue is ±1:

$$|n\rangle = \mathbb{1}\,|n\rangle = P_{12}^2\,|n\rangle = \lambda^2\,|n\rangle \quad \Longrightarrow \quad \lambda = \pm 1. \tag{17.10}$$

To find the eigenstates, we need states that satisfy

$$P_{12}\,|n\rangle = \pm\,|n\rangle\,. \tag{17.11}$$

Inserting the general state of Eq. (17.5) into Eq. (17.11), we obtain

$$P_{12}\left(\alpha\left(|\psi'\rangle\,|\psi''\rangle\right) + \beta\left(|\psi''\rangle\,|\psi'\rangle\right)\right) = \pm\left[\alpha\left(|\psi'\rangle\,|\psi''\rangle\right) + \beta\left(|\psi''\rangle\,|\psi'\rangle\right)\right]. \tag{17.12}$$

Applying P_{12} on the left-hand side, we find

$$\beta\left(|\psi'\rangle\,|\psi''\rangle\right) + \alpha\left(|\psi''\rangle\,|\psi'\rangle\right) = \pm\left[\alpha\left(|\psi'\rangle\,|\psi''\rangle\right) + \beta\left(|\psi''\rangle\,|\psi'\rangle\right)\right], \tag{17.13}$$

and thus

$$\beta = \pm\alpha. \tag{17.14}$$

The two normalised eigenstates of P_{12} are

$$|\psi'\psi''\rangle_\pm = \frac{1}{\sqrt{2}}\left[|\psi'\rangle\,|\psi''\rangle \pm |\psi''\rangle\,|\psi'\rangle\right], \tag{17.15}$$

with eigenvalues +1 and −1, respectively (for $\langle\psi'|\psi''\rangle = 0$). We can extend the definition of the permutation operator to systems of more than two identical particles. We denote by P_{ij} the permutation operator that exchanges particles i and j:

$$P_{ij}\left(|\psi^1\rangle\,|\psi^2\rangle\cdots|\psi^i\rangle\cdots|\psi^j\rangle\cdots|\psi^N\rangle\right) = |\psi^1\rangle\,|\psi^2\rangle\cdots|\psi^j\rangle\cdots|\psi^i\rangle\cdots|\psi^N\rangle\,. \tag{17.16}$$

As for two particles, $P_{ij}^2 = \mathbb{1}$. However, different permutation operators do not commute in general:

$$\left[P_{ij}, P_{kl}\right] \neq 0. \tag{17.17}$$

Let us come back to two identical particles and consider their Hamiltonian. Our goal is to show that the permutation operator is a symmetry of the system. The Hamiltonian contains three potentials, an external potential for each particle and a pair potential:

$$H = \frac{p_1^2}{2m} + \frac{p_2^2}{2m} + V_{\text{ext}}(\mathbf{x}_1) + V_{\text{ext}}(\mathbf{x}_2) + V_{\text{pair}}(|\mathbf{x}_1 - \mathbf{x}_2|). \tag{17.18}$$

Note that, being identical, the two particles have the same mass and feel the same potential V_{ext}. Furthermore, they exert a mutual force on each other, which (if we ignore spin–spin interactions) depends only on their distance.

The Hamiltonian is built out of the operators $\hat{\mathbf{p}}_1, \hat{\mathbf{x}}_1, \hat{\mathbf{p}}_2, \hat{\mathbf{x}}_2$ of the two identical particles. We now argue that the permutation operator P_{12} commutes with the Hamiltonian. A heuristic argument goes as follows. Consider an eigenstate $|1,2\rangle_K$ of the Hamiltonian,

$$H\,|1,2\rangle_K = E_K\,|1,2\rangle_K\,. \tag{17.19}$$

The exchange degeneracy requires that the states $|1,2\rangle_K$ and $|2,1\rangle_K$ have the same energy:

$$H\,|2,1\rangle_K = E_K\,|2,1\rangle_K\,. \tag{17.20}$$

Applying the commutator of the permutation operator and the Hamiltonian to the first eigenstate, we obtain

$$\begin{aligned}[P_{12}, H]\,|1,2\rangle_K &= P_{12}H\,|1,2\rangle_K - HP_{12}\,|1,2\rangle_K \\ &= P_{12}E_K\,|1,2\rangle_K - H\,|2,1\rangle_K \\ &= E_K\,|2,1\rangle_K - H\,|2,1\rangle_K \\ &= (E_K - H)\,|2,1\rangle_K\,.\end{aligned} \tag{17.21}$$

Inserting Eq. (17.20), we obtain

$$[P_{12}, H]\,|1,2\rangle_K = 0. \tag{17.22}$$

Since, by the same argument, Eq. (17.22) also holds for $|2,1\rangle_K$, we conclude that

$$[P_{12}, H] = 0. \tag{17.23}$$

This argument is not a proof, since Eq. (17.20) and $P_{12}\,|1,2\rangle_K = |2,1\rangle_K$ already assume what is to be shown. The direct proof is that $P_{12}\,\mathbf{x}_1 P_{12}^{-1} = \mathbf{x}_2$, $P_{12}\,\mathbf{p}_1 P_{12}^{-1} = \mathbf{p}_2$ (and vice versa), so $P_{12}HP_{12}^{-1} = H$ because H in Eq. (17.18) is symmetric under $1 \leftrightarrow 2$ (Sakurai, Sec. 7.1). The permutation operator is therefore a symmetry, and its eigenvalue is a constant of motion. In general, a system of N identical particles that is in an eigenstate of P_{ij} remains in this eigenstate forever.

A priori, there is no reason for a system to be in an eigenstate of the permutation operators. From what we know so far, we can only conclude that the system is in general in a superposition of eigenstates of the permutation operators, and no measurement can single out one of them. Nature, however, turns out to be much more selective.

17.3 Symmetrisation Postulate

Physical systems with an arbitrary number N of identical particles fall into only two categories: under the exchange of any pair of the N particles, the wave function is either *totally symmetric* or *totally antisymmetric*.

Particles which form symmetric systems are called *bosons*, after the Indian polymath Satyendra Nath Bose, who described their statistics in a letter to Einstein in 1924. They obey *Bose–Einstein* statistics. Particles which form antisymmetric states are called *fermions*, after Enrico Fermi. They obey *Fermi–Dirac* statistics. Thus,

$$\begin{aligned}P_{ij}\,|N \text{ identical fermions}\rangle &= -|N \text{ identical fermions}\rangle\,, \\ P_{ij}\,|N \text{ identical bosons}\rangle &= +|N \text{ identical bosons}\rangle\,.\end{aligned} \tag{17.24}$$

In nature we observe that particles are either bosons or fermions. We cannot prove this within quantum mechanics, but since the observation has never been falsified, we postulate it. The spin of a particle determines whether it is a boson or a fermion: particles with half-integer spin ($s = n + 1/2$, $n \in \mathbb{N}$) are fermions, and particles with integer spin ($s = n$, $n \in \mathbb{N}$) are bosons.

Prime examples of fermions are the leptons and the quarks. The leptons are the electron, the muon, the tau and the three neutrinos; the quarks are the up, down, strange, charm, bottom and top quarks. All of them are elementary particles with spin 1/2. Examples of bosons are the "gauge bosons": the photon, the gluons, and the $W^{\pm}$ and Z bosons. They are elementary particles with spin 1.

This relation between spin and statistics is called the *spin–statistics connection*. Non-relativistic quantum mechanics offers no explanation for it; it can only be derived in relativistic quantum field theory. The spin–statistics connection also holds for composite particles. For example, a ^{4}He nucleus is a boson, while a ^{3}He nucleus is a fermion.

17.3.1 Pauli Exclusion Principle

Pauli's exclusion principle states that two identical fermions can never occupy the same quantum state. As an example, let us consider two electrons (1) and (2), which occupy states $|\psi^1\rangle$ and $|\psi^2\rangle$. As electrons are fermions, the two-electron system is necessarily in an antisymmetric state. The corresponding eigenstate is given by

$$\frac{1}{\sqrt{2}}\left[|\psi^1\rangle \otimes |\psi^2\rangle - |\psi^2\rangle \otimes |\psi^1\rangle\right]. \tag{17.25}$$

If both electrons are in the same state, $|\psi^1\rangle = |\psi^2\rangle$, this expression vanishes. Therefore, the two electrons cannot be in the same state.

Bosons, on the contrary, can occupy the same state. At very low temperatures, particles occupy the lowest energy levels. In a system of bosons, such as ^{4}He atoms, a macroscopic fraction of the particles then falls into the same one-particle ground state, forming a *Bose–Einstein condensate*.

17.4 Two-Electron Systems

We now consider a system of two electrons. According to the previous discussion, its wave function is antisymmetric under the exchange of the two electrons:

$$\psi(1,2) = -\psi(2,1). \tag{17.26}$$

The total spin of the system is the sum of the spins of the two electrons,

$$\mathbf{S}_{\text{tot}} = \mathbf{S}_1 + \mathbf{S}_2. \tag{17.27}$$

We consider the case where the square of the total spin commutes with the Hamiltonian,

$$\left[\mathbf{S}_{\text{tot}}^2, H\right] = 0. \tag{17.28}$$

In this case, we can write the wave function ψ as the product of a spatial part and a spin part,

$$\psi(1,2) = \phi(\mathbf{x}_1, \mathbf{x}_2)\chi(1,2). \tag{17.29}$$

Strictly speaking, $[\mathbf{S}_{\text{tot}}^2, H] = 0$ alone only guarantees eigenstates of definite total spin S, $\psi = \sum_m \phi_m(\mathbf{x}_1, \mathbf{x}_2)\chi_{Sm}$. The product form (17.29) holds when H is spin independent (then also $[\mathbf{S}_{\text{tot}}, H] = 0$ and S_z can be fixed), which is the case considered here. We find $\chi(1,2)$ with the addition rules for angular momenta. We write $|\uparrow\rangle$ for a spin parallel to the z-axis and $|\downarrow\rangle$ for a spin antiparallel to it. The spin states are the triplet with total spin 1,

$$\chi_{\text{triplet}} = \begin{cases} |\uparrow\rangle \otimes |\uparrow\rangle \\ \frac{1}{\sqrt{2}}\left(|\uparrow\rangle \otimes |\downarrow\rangle + |\downarrow\rangle \otimes |\uparrow\rangle\right) \\ |\downarrow\rangle \otimes |\downarrow\rangle \end{cases} \tag{17.30}$$

and the singlet with total spin 0,

$$\chi_{\text{singlet}} = \frac{1}{\sqrt{2}} \left(|\uparrow\rangle \otimes |\downarrow\rangle - |\downarrow\rangle \otimes |\uparrow\rangle \right). \tag{17.31}$$

The triplet states are symmetric under the exchange of the two electrons. Since the total wave function is antisymmetric, a two-electron system in a spin-triplet state must have an antisymmetric spatial wave function,

$$\phi(\mathbf{x}_1, \mathbf{x}_2) = -\phi(\mathbf{x}_2, \mathbf{x}_1), \quad \text{for spin-triplet.} \tag{17.32}$$

Analogously, a two-electron system in a spin-singlet state must have a symmetric spatial wave function,

$$\phi(\mathbf{x}_1, \mathbf{x}_2) = \phi(\mathbf{x}_2, \mathbf{x}_1), \quad \text{for spin-singlet.} \tag{17.33}$$

So far, when studying a single electron, we have ignored the constraints on its wave function due to the existence of other identical particles. As we have just seen, however, such constraints exist, and formally they apply even if the closest identical particle is at the edge of the universe. It is worth examining this question in detail.

Let us consider two electrons which are far from each other, so that their Coulomb interaction is negligible,

$$V_{\text{pair}}(|\mathbf{x}_1 - \mathbf{x}_2|) \simeq 0. \tag{17.34}$$

Neglecting the pair interaction and the spin, the Schrödinger equation for the spatial wave function is

$$\left[-\frac{\hbar^2}{2m}\nabla_1^2 - \frac{\hbar^2}{2m}\nabla_2^2 + V_{\text{ext}}(\mathbf{x}_1) + V_{\text{ext}}(\mathbf{x}_2) \right] \phi = E\phi. \tag{17.35}$$

To solve this differential equation, we separate the wave function into two factors, each of which depends on only one of the two particles,

$$\phi(\mathbf{x}_1, \mathbf{x}_2) = \omega_A(\mathbf{x}_1)\omega_B(\mathbf{x}_2). \tag{17.36}$$

Inserting this ansatz into Eq. (17.35), we obtain two independent differential equations,

$$\begin{aligned} &\left[-\frac{\hbar^2}{2m}\nabla_1^2 + V_{\text{ext}}(\mathbf{x}_1) \right] \omega_A(\mathbf{x}_1) = E_A \omega_A(\mathbf{x}_1), \\ &\left[-\frac{\hbar^2}{2m}\nabla_2^2 + V_{\text{ext}}(\mathbf{x}_2) \right] \omega_B(\mathbf{x}_2) = E_B \omega_B(\mathbf{x}_2), \quad E = E_A + E_B. \end{aligned} \tag{17.37}$$

Depending on the total-spin state, only the symmetric or the antisymmetric combination is physical. The spatial wave function is thus

$$\phi_\pm(\mathbf{x}_1, \mathbf{x}_2) = \frac{1}{\sqrt{2}} \left[\omega_A(\mathbf{x}_1)\omega_B(\mathbf{x}_2) \pm \omega_B(\mathbf{x}_1)\omega_A(\mathbf{x}_2) \right]. \tag{17.38}$$

The probability of finding an electron in the volume $d^3\mathbf{x}_1$ and the other electron in the volume $d^3\mathbf{x}_2$ is

$$\begin{aligned} dP_\pm(\mathbf{x}_1, \mathbf{x}_2) &\equiv d^3\mathbf{x}_1\, d^3\mathbf{x}_2 \left| \phi_\pm(\mathbf{x}_1, \mathbf{x}_2) \right|^2 \\ &= \frac{1}{2} \Big[\left(d^3\mathbf{x}_1 \left| \omega_A(\mathbf{x}_1) \right|^2 \right) \left(d^3\mathbf{x}_2 \left| \omega_B(\mathbf{x}_2) \right|^2 \right) \\ &\quad + \left(d^3\mathbf{x}_1 \left| \omega_B(\mathbf{x}_1) \right|^2 \right) \left(d^3\mathbf{x}_2 \left| \omega_A(\mathbf{x}_2) \right|^2 \right) \Big] \\ &\quad \pm \text{Re} \Big[\left(d^3\mathbf{x}_1 \omega_A(\mathbf{x}_1) \omega_B^*(\mathbf{x}_1) \right) \left(d^3\mathbf{x}_2 \omega_A^*(\mathbf{x}_2) \omega_B(\mathbf{x}_2) \right) \Big]. \end{aligned} \tag{17.39}$$

The last term in Eq. (17.39) is called the *exchange density*. The first two terms are mapped into each other by the exchange of variables $\mathbf{x}_1 \leftrightarrow \mathbf{x}_2$. Since the two particles are identical, they give the same contribution, and we can rewrite Eq. (17.39) as[1]

$$\begin{aligned} dP_\pm(\mathbf{x}_1, \mathbf{x}_2) = & \left(d^3\mathbf{x}_1 \, |\omega_A(\mathbf{x}_1)|^2\right) \left(d^3\mathbf{x}_2 \, |\omega_B(\mathbf{x}_2)|^2\right) \\ & \pm \mathrm{Re}\Big[\left(d^3\mathbf{x}_1 \omega_A(\mathbf{x}_1)\omega_B^*(\mathbf{x}_1)\right) \left(d^3\mathbf{x}_2 \omega_A^*(\mathbf{x}_2)\omega_B(\mathbf{x}_2)\right) \Big]. \end{aligned} \tag{17.40}$$

Eq. (17.40) consists of two terms, the "classical probability" and the exchange density. The classical probability is the probability that two independent particles are found in the volumes $d^3\mathbf{x}_1$ and $d^3\mathbf{x}_2$.

If we integrate over the position of one of the particles, the classical term gives simply the probability density of the other particle. The exchange density, however, does not disappear in general, since it involves the integral of the product of the two one-electron wave functions.

Let us look at this integral,

$$\int d^3\mathbf{x} \, \omega_A(\mathbf{x})\omega_B^*(\mathbf{x}), \quad \mathbf{x} = \mathbf{x}_1, \mathbf{x}_2. \tag{17.41}$$

We consider the case where $\omega_A(\mathbf{x})$ is negligible away from a point P_1 and $\omega_B(\mathbf{x})$ is negligible away from a point P_2, as shown in Fig. 17.3. If P_1 and P_2 are well separated, the product

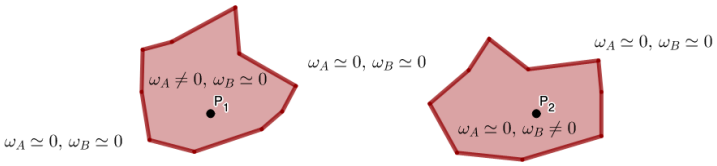


Figure 17.3. Two distant identical particles, each of which can be studied in isolation.

$\omega_A(\mathbf{x})\omega_B^*(\mathbf{x})$ is negligible at all points, including P_1 and P_2. We conclude that exchange effects can be ignored if the wave functions of identical particles do not overlap significantly: we can safely study electrons in isolation as long as they are truly isolated.

17.5 Scattering of Identical Particles

Consider the scattering of two identical spin-0 particles (Fig. 17.4). In Eq. (15.39) we derived the large-distance form of the wave function. The boundary conditions, together with the requirement that the wave function is symmetric under the exchange of the two particles, motivate the ansatz

$$e^{i\mathbf{k}\cdot\mathbf{x}} + e^{-i\mathbf{k}\cdot\mathbf{x}} + \frac{e^{ikr}}{r}\left[f(\theta) + f(\pi - \theta)\right], \tag{17.42}$$

where $\mathbf{x} = \mathbf{x}_1 - \mathbf{x}_2$ is the relative position vector for the two particles. The differential

[1] The two direct terms are not equal as functions of $(\mathbf{x}_1, \mathbf{x}_2)$; they are mapped into each other by $\mathbf{x}_1 \leftrightarrow \mathbf{x}_2$. Eq. (17.40) is therefore an equality only for symmetric quantities (integrals of $dP_\pm$ against functions symmetric in $\mathbf{x}_1, \mathbf{x}_2$), or when $\mathbf{x}_1$ is restricted to the region where ω_A is supported and $\mathbf{x}_2$ to that of ω_B (and the probability for "one electron in $d^3\mathbf{x}_1$, the other in $d^3\mathbf{x}_2$" then carries a factor 2 relative to $|\phi_\pm|^2$).

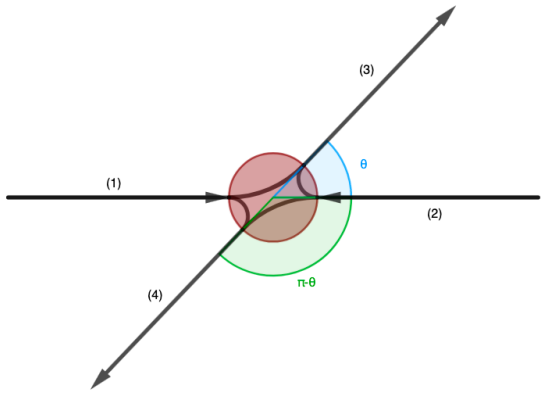


Figure 17.4. Scattering of two identical particles in the centre-of-mass frame: scattering angle θ and the exchanged configuration $\pi - \theta$.

cross-section is the squared modulus of the coefficient of e^{ikr}/r,

$$\begin{aligned}\frac{d\sigma}{d\Omega}\Big(\substack{\text{0-spin}\\\text{scattering}}\Big) &= \left|f(\theta) + f(\pi-\theta)\right|^2 \\ &= \left|f(\theta)\right|^2 + \left|f(\pi-\theta)\right|^2 + 2\mathrm{Re}\left[f(\theta)f^*(\pi-\theta)\right].\end{aligned} \tag{17.43}$$

At $\theta = \pi/2$, the cross-section is

$$\frac{d\sigma}{d\Omega}\Big(\substack{\text{0-spin}\\\text{scattering}}\Big)\left(\frac{\pi}{2}\right) = 4\left|f\left(\frac{\pi}{2}\right)\right|^2. \tag{17.44}$$

For the scattering of two spin-1/2 particles with unpolarised beams, spin-singlet scattering is described by a symmetric spatial wave function and spin-triplet scattering by an antisymmetric one. On average, the two particles are in a spin-triplet state 3/4 of the time and in a spin-singlet state 1/4 of the time. The cross-section is

$$\begin{aligned}\frac{d\sigma}{d\Omega}\Big(\substack{\text{(1/2-spin)-(1/2-spin)}\\\text{scattering}}\Big) &= \frac{3}{4}\left|f(\theta) - f(\pi-\theta)\right|^2 + \frac{1}{4}\left|f(\theta) + f(\pi-\theta)\right|^2 \\ &= \left|f(\theta)\right|^2 + \left|f(\pi-\theta)\right|^2 - \mathrm{Re}\left[f(\theta)f^*(\pi-\theta)\right].\end{aligned} \tag{17.45}$$

At $\theta = \pi/2$, the interference term is destructive,

$$\frac{d\sigma}{d\Omega}\Big(\substack{\text{(1/2-spin)-(1/2-spin)}\\\text{scattering}}\Big)(\theta = \pi/2) = \left|f(\pi/2)\right|^2. \tag{17.46}$$

This characteristic behaviour is indeed observed in experiments.

17.6 System of N Fermions

We consider $N \geq 2$ non-interacting identical fermions. We can build a totally antisymmetric state from the states of the individual fermions. For $N = 2$ we had

$$\psi(1,2) = \frac{1}{\sqrt{2}}\left(\psi_{a_1}(1)\psi_{a_2}(2) - \psi_{a_1}(2)\psi_{a_2}(1)\right) = \frac{1}{\sqrt{2}}\det\begin{pmatrix}\psi_{a_1}(1) & \psi_{a_1}(2)\\ \psi_{a_2}(1) & \psi_{a_2}(2)\end{pmatrix}. \tag{17.47}$$

The total wave function is built from as many one-particle wave functions as there are particles; each of them is a different solution of the one-particle Schrödinger equation. The generalisation to $N \geq 2$ particles is the *Slater determinant*

$$\psi(1,2,\ldots,N) = \frac{1}{\sqrt{N!}} \det \begin{pmatrix} \psi_{a_1}(1) & \cdots & \psi_{a_1}(N) \\ \vdots & \ddots & \vdots \\ \psi_{a_N}(1) & \cdots & \psi_{a_N}(N) \end{pmatrix}. \tag{17.48}$$

17.6.1 Free Particles in a Box

Let us consider N particles in a box of volume L^3 (Fig. 17.5). Assuming that the particles

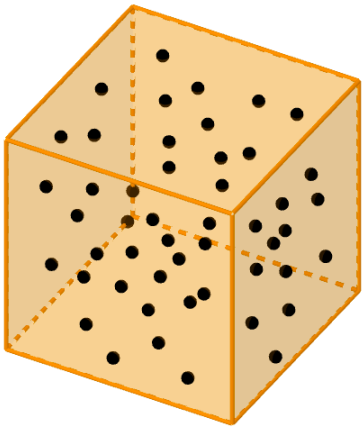


Figure 17.5. N particles in a cubic box of side L.

are free, the one-particle wave functions are

$$\psi_P = \frac{1}{L^{3/2}} e^{i\mathbf{p}\cdot\mathbf{x}/\hbar}. \tag{17.49}$$

with energy

$$E = \frac{\mathbf{p}^2}{2m}. \tag{17.50}$$

We impose periodic boundary conditions, $\psi(\mathbf{x} + L\,\hat{\mathbf{e}}_i) = \psi(\mathbf{x})$. The momentum spectrum is then discrete,

$$\mathbf{p}_i = \frac{2\pi\hbar}{L}(n_1, n_2, n_3), \quad n_i \in \mathbb{Z}. \tag{17.51}$$

There is an interesting difference between a box of fermions and a box of bosons. Bosons can all occupy the same state. We can put all of them in the state with $\mathbf{p} = 0$, and the ground-state energy of the bosonic system is $E_{\text{bosons}} = 0$.

Fermions cannot be in the same state. The ground state of the system is obtained by filling the one-particle states with the lowest momenta. For N particles, the occupied states all lie within a sphere in momentum space, the *Fermi sphere*, whose radius P_F is the *Fermi momentum* (Fig. 17.6):

$$|\mathbf{p}_i| \leq P_F. \tag{17.52}$$

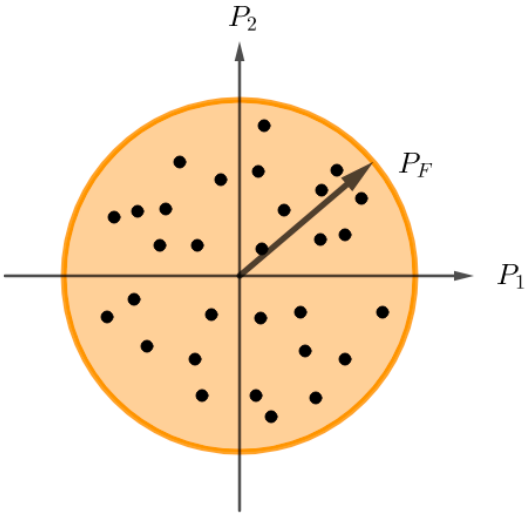


Figure 17.6. Two-dimensional sketch of the three-dimensional Fermi sphere.

To calculate the number of states within the Fermi sphere (for electrons), we first calculate the number of allowed momenta $n_{\max}$ along one direction. We divide the Fermi momentum by the spacing between two momentum states, Eq. (17.51), and obtain

$$n_{\max} = \frac{LP_F}{2\pi\hbar}. \tag{17.53}$$

The total number of momenta within the sphere follows from the volume of the sphere, and for each momentum there are two spin states. Thus the total number of states within the Fermi sphere is

$$N = 2 \cdot \frac{4\pi}{3} n_{\max}^3 = \frac{\left(\frac{P_F}{\hbar}\right)^3}{3\pi^2} \cdot L^3. \tag{17.54}$$

The number density is then

$$n \equiv \frac{N}{L^3} = \frac{\left(\frac{P_F}{\hbar}\right)^3}{3\pi^2}. \tag{17.55}$$

The energy of the ground state is the sum of the energies of all N fermions. For large L the spectrum is almost continuous, and we can replace the sum by an integral over $d^3\mathbf{n}$, where $\mathbf{n} = (n_1, n_2, n_3)$ is the vector of Eq. (17.51):

$$E_{\text{ground}} = 2 \sum_i \frac{\mathbf{p}_i^2}{2m} = 2 \int d^3\mathbf{n}\, \frac{\mathbf{p}^2}{2m}. \tag{17.56}$$

With Eq. (17.51) we change the integration variable to $\mathbf{p}$:

$$E_{\text{ground}} = 2 \int_{|\mathbf{p}| \le P_F} d^3\mathbf{p}\, \frac{L^3}{(2\pi\hbar)^3} \frac{\mathbf{p}^2}{2m} = 8\pi \int_0^{P_F} dp\, \frac{L^3}{(2\pi\hbar)^3} \frac{p^4}{2m} = \frac{L^3}{\pi^2\hbar^3} \frac{P_F^5}{10m}. \tag{17.57}$$

Using Eq. (17.54), we can express the ground-state energy in terms of the number of particles N:

$$E_{\text{ground}} = 3N \frac{P_F^2}{10m}. \tag{17.58}$$

In terms of the Fermi energy

$$E_F \equiv \frac{P_F^2}{2m}, \tag{17.59}$$

we can rewrite Eq. (17.58) as

$$E_{\text{ground}} = \frac{3}{5} N E_F. \tag{17.60}$$

18

Multi-Electron Atoms

Having introduced identical particles in the previous chapter, we now ask what impact they have on the structure of matter. Atoms are complicated quantum systems with many identical electrons, and multi-electron systems cannot be solved exactly. To describe them, we need a variety of approximation methods together with a careful implementation of the Pauli exclusion principle.

18.1 The Helium Atom

We start with the simplest multi-electron atom, helium. Since the nucleons are tightly bound, we can consider them to be localised in a *nucleus*, and we describe the helium atom as a system of two electrons and a nucleus (Fig. 18.1). The Hamiltonian consists of the kinetic energies of the two electrons, the potential energy of each electron in the field of the nucleus, and the potential energy between the two electrons. We neglect the kinetic energy of the nucleus, which is small compared to the other terms:

$$H = \frac{p_1^2}{2m} + \frac{p_2^2}{2m} - \frac{Ze^2}{r_1} - \frac{Ze^2}{r_2} + \frac{e^2}{r_{12}}, \quad Z_{\text{He}} = 2. \tag{18.1}$$

Since it does not depend on spin, the Hamiltonian commutes with the total spin operator of the two electrons. Hence, its eigenstates can be chosen to be either spin-triplet or spin-singlet states. The spatial part of the wave function is found by solving the Schrödinger equation with the Hamiltonian of Eq. (18.1).

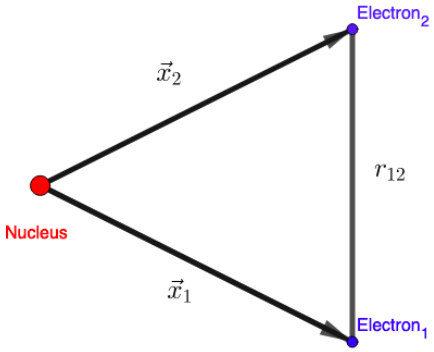


Figure 18.1. The helium atom: two electrons bound to a nucleus of charge Ze, with $Z = 2$.

This equation cannot be solved analytically. Fortunately, in Chapter 14 we developed two approximation methods, perturbation theory and the variational method, which we can now apply.

From Part I we know that the Schrödinger equation can be solved exactly for the hydrogen atom (and for hydrogen-like ions of charge Ze). A system of two independent hydrogen-like electrons is therefore also exactly solvable. Our strategy for the helium atom is thus to split the Hamiltonian into an exactly solvable part and a perturbation, and then to perform a perturbative expansion:

$$H = \underbrace{\frac{p_1^2}{2m} - \frac{Ze^2}{r_1}}_{\text{1st hydrogen-like atom}} + \underbrace{\frac{p_2^2}{2m} - \frac{Ze^2}{r_2}}_{\text{2nd hydrogen-like atom}} + \underbrace{\frac{e^2}{r_{12}}}_{\text{perturbation}} . \tag{18.2}$$

18.1.1 Ground State

To perform the perturbative expansion, we first have to solve the unperturbed system. Since the unperturbed Hamiltonian is the sum of two hydrogen-like Hamiltonians, we can write the wave function of the total system as the product of two spatial wave functions, one for each electron, and a spin part, which describes the spin of the entire system:

$$\psi = \psi_{n\ell m}(\mathbf{x}_1)\psi_{n'\ell'm'}(\mathbf{x}_2)\chi . \tag{18.3}$$

In the ground state, both electrons occupy the same spatial wave function, the one with the lowest energy. The spatial part of the unperturbed wave function is then symmetric under the exchange of the two electrons, as Eq. (18.3) shows. Thus, the spin part must be antisymmetric, that is, a singlet:

$$\psi_{\text{ground}} = \psi_{100}(\mathbf{x}_1)\psi_{100}(\mathbf{x}_2)\chi_{\text{singlet}} . \tag{18.4}$$

The spatial wave function is

$$\psi_{100} = \frac{1}{\sqrt{\pi}} \left(\frac{Z}{a_B}\right)^{3/2} e^{-\frac{Zr}{a_B}} , \tag{18.5}$$

where

$$a_B = \frac{\hbar^2}{m_e e^2} = 0.529 \tag{18.6}$$

is the Bohr radius (also denoted a_0 below). The energy is given by

$$E_{100} = Z^2 \left(-\frac{e^2}{2a_0}\right) \quad \Longrightarrow \quad E_{\text{ground}} = 2E_{100} = -108.8\,\text{eV} . \tag{18.7}$$

This first approximation lies about 38% below the measured value. To improve our estimate, we compute the first-order perturbative correction,

$$\Delta E^{(1)} = \left\langle \frac{e^2}{r_{12}} \right\rangle , \tag{18.8}$$

where we have used Eq. (14.12). The expectation value is

$$\begin{aligned}\Delta E^{(1)} &= \int d^3\mathbf{x}_1\, d^3\mathbf{x}_2\, \left|\psi_{100}(\mathbf{x}_1)\right|^2 \frac{e^2}{|\mathbf{x}_1-\mathbf{x}_2|} \left|\psi_{100}(\mathbf{x}_2)\right|^2 \\ &= \left(\frac{Z}{a_B}\right)^6 \frac{e^2}{\pi^2} \int_0^\infty dr_1\, r_1^2 e^{-\frac{2Zr_1}{a_B}} \cdot \int_0^\infty dr_2\, r_2^2 e^{-\frac{2Zr_2}{a_B}} \\ &\quad \cdot \int d\Omega_1 d\Omega_2 \frac{1}{|\mathbf{x}_1-\mathbf{x}_2|}.\end{aligned} \tag{18.9}$$

The result is

$$\Delta E^{(1)} = Z\frac{5}{4}\frac{e^2}{2a_0}. \tag{18.10}$$

The first-order correction raises the ground-state energy:

$$E_{\text{ground}} = 2E_{100} + \Delta E^{(1)} = \left(-2Z^2 + \frac{5}{4}Z\right)\frac{e^2}{2a_0} = -74.8\,\text{eV}. \tag{18.11}$$

This result should be compared with the measured value

$$E_{\text{ground, exp.}} = -79.0\,\text{eV}. \tag{18.12}$$

The first-order correction thus raises the ground-state energy by about 30%, and the estimate is now within about 5% of the measured value. Higher-order corrections improve the result further.

A better approximation is obtained with the variational method, the second approximation method of Chapter 14. As variational parameter we take the charge number appearing in the wave function, which we call Z_{eff}. The physical motivation for this ansatz is that each electron does not feel the full charge of the nucleus, because the other electron partially screens it. We use the same wave function as before, with Z replaced by the parameter Z_{eff}:

$$\psi_{100} = \frac{1}{\sqrt{\pi}}\left(\frac{Z_{\text{eff}}}{a_B}\right)^{3/2} e^{-\frac{Z_{\text{eff}}r}{a_B}}. \tag{18.13}$$

The expectation value of the Hamiltonian is

$$\langle\psi|H|\psi\rangle = 2\left(Z_{\text{eff}}^2 - 2Z_{\text{eff}}Z + \frac{5}{8}Z_{\text{eff}}\right)\frac{e^2}{2a_0}. \tag{18.14}$$

It is easy to see that the expectation value is minimal for

$$Z_{\text{eff}} = Z - \frac{5}{16} = 1.6875. \tag{18.15}$$

Finally, we obtain the energy

$$\langle\psi|H|\psi\rangle = -77.5\,\text{eV}, \tag{18.16}$$

which is closer to the measured value than the first-order perturbative result. With more elaborate trial wave functions, the variational method yields the energy to excellent precision.

18.1.2 Higher Energy Levels

An interesting phenomenon occurs for excited states, in which one of the two electrons is in the ground state ($1s$) and the other one is in a higher state $|\psi_{n\ell m}\rangle$. If the total spin state is a singlet, the spatial part is symmetric and the wave function at leading order in perturbation theory is

$$\psi(\mathbf{x}_1, \mathbf{x}_2) = \frac{1}{\sqrt{2}} \left[\psi_{100}(\mathbf{x}_1)\psi_{n\ell m}(\mathbf{x}_2) + \psi_{n\ell m}(\mathbf{x}_1)\psi_{100}(\mathbf{x}_2)\right] \chi_{\text{singlet}}. \tag{18.17}$$

If the spin state is a triplet, the spatial part must be antisymmetric and thus the wave function at leading order is given by

$$\psi(\mathbf{x}_1, \mathbf{x}_2) = \frac{1}{\sqrt{2}} \left[\psi_{100}(\mathbf{x}_1)\psi_{n\ell m}(\mathbf{x}_2) - \psi_{n\ell m}(\mathbf{x}_1)\psi_{100}(\mathbf{x}_2)\right] \chi_{\text{triplet}}. \tag{18.18}$$

The energy of such a state is

$$E = E_{100} + E_{n\ell m} + \Delta E, \tag{18.19}$$

where the energy correction ΔE is computed in perturbation theory. At first order, it is the expectation value of the perturbation,

$$\Delta E = \left\langle \frac{e^2}{r_{12}} \right\rangle. \tag{18.20}$$

The expectation value can be expressed in terms of two integrals, the *direct* integral I and the *exchange* integral J:

$$\left\langle \frac{e^2}{r_{12}} \right\rangle = I \pm J, \tag{18.21}$$

where

$$\begin{aligned} I &= \int d^3\mathbf{x}_1 \int d^3\mathbf{x}_2 \left|\psi_{100}(\mathbf{x}_1)\right|^2 \left|\psi_{nlm}(\mathbf{x}_2)\right|^2 \frac{e^2}{r_{12}}, \\ J &= \int d^3\mathbf{x}_1 \int d^3\mathbf{x}_2 \, \psi_{100}(\mathbf{x}_1)\, \psi_{nlm}(\mathbf{x}_2) \frac{e^2}{r_{12}} \psi^*_{100}(\mathbf{x}_2)\, \psi^*_{nlm}(\mathbf{x}_1). \end{aligned} \tag{18.22}$$

The plus sign holds for the spin singlet and the minus sign for the spin triplet. The integral I is obviously positive, and one can show that J is positive as well. Therefore, the energy of the spin-singlet state is higher than the energy of the spin-triplet state.

The physical reason for this separation is that in the singlet case, the spatial function is symmetric and the two electrons tend to come closer together, so that the electrostatic repulsion is larger. In the triplet case, the spatial wave function is antisymmetric and the electrons tend to avoid each other; the electrostatic repulsion is smaller.

This result is remarkable: although the Hamiltonian does not depend on spin, the energy eigenvalue depends on the spin state. The reason is the Pauli exclusion principle, since a symmetric (antisymmetric) spin state forces the spatial part of the wave function to be antisymmetric (symmetric). Spin-singlet states of helium are called *parahelium*, and spin-triplet states *orthohelium*. As we saw in Sec. 18.1.1, the ground state can only be parahelium, since its spatial wave function is necessarily symmetric.

The singlet–triplet splitting $2J$ acts like an effective spin–spin coupling: within the $1s\,n\ell$ configuration the first-order energy can be written $I - \frac{1}{2}J\,(1 + 4\,\mathbf{S}_1 \cdot \mathbf{S}_2/\hbar^2)$ (Dirac's exchange operator), which explains the "spin dependence" of the levels of a spin-independent H (Gottfried–Yan, Sec. 6.4; Sakurai, Sec. 7.4).

18.2 Atoms with More than Two Electrons

Having studied the helium atom, we now turn to an atom with N electrons and nuclear charge number Z. As we saw for the helium ground state, perturbation theory is tedious, and the amount of calculation grows with each order. For atoms with more than two electrons the variational method is therefore commonly used. Note, however, that the variational method requires an ansatz, whereas perturbation theory is a complete recipe. In this section we look at two methods, the Hartree approximation and the Hartree–Fock approximation, named after the English physicist D. R. Hartree and the Soviet physicist V. A. Fock.

The Hamiltonian of the atom is

$$H = \sum_{i=1}^{N} \left(\frac{\mathbf{p}_i^2}{2m} - \frac{Ze^2}{r_i} \right) + \sum_{i>j} \frac{e^2}{|\mathbf{x}_i - \mathbf{x}_j|}. \tag{18.23}$$

The first sum contains the kinetic energy of each electron and its interaction with the nucleus; the second sum is the interaction between the electrons. The system is described by an N-electron wave function

$$\psi(1, 2, \ldots, N), \tag{18.24}$$

which is an eigenstate of the Hamiltonian:

$$H\psi(1, 2, \ldots, N) = E\psi(1, 2, \ldots, N). \tag{18.25}$$

We now have to make an approximation in order to develop an ansatz for the variational method. We motivate it with the following physical picture. Each electron is subject to

(i) the potential of the nucleus, and

(ii) a collective potential due to all other electrons, which act on it in a "uniform" way.

In this approximation we do not treat the individual electron–electron interactions separately, but combine them into a collective effect. Physically, we argue that for a large number of electrons the potential they generate does not change much if a single electron is removed. Therefore, by symmetry, this potential is the same for every electron.[1]

18.2.1 The Hartree Approximation

The Hartree approximation was formulated in 1927, one year after the publication of the Schrödinger equation. It does not respect the antisymmetry required by the Pauli exclusion principle and is therefore fundamentally flawed. Nevertheless, it was historically a widely accepted method because of its simplicity and low computational cost. Moreover, more accurate methods require only slight modifications of the Hartree approximation. We therefore derive the Hartree method first.

We start by writing the wave function as a product of one-particle wave functions,

$$\psi(1, 2, \ldots, N) = \varphi_1(1)\varphi_2(2)\ldots\varphi_N(N). \tag{18.26}$$

[1] "Uniform" should not be read as spatially constant (a constant potential would have no effect). What is meant is a smooth, averaged (mean-field) potential with the same functional form for every electron; in the central-field approximation it is taken spherically symmetric, $V(r)$, interpolating between $-Ze^2/r$ at small r and $-e^2/r$ at large r.

Each one-particle wave function is in turn the product of a spatial wave function and a spin wave function,

$$\varphi_i(i) = \phi_i(\mathbf{x}_i)\chi_i(m_{s_i}). \tag{18.27}$$

This ansatz is clearly not antisymmetric and therefore violates the Pauli exclusion principle. We shall see a simple remedy for this problem in the next subsection. For now, we only demand that the wave functions ϕ_i are distinct and orthogonal. They should also be normalised:

$$\int d^3\mathbf{x}_i \; \left|\phi_i(\mathbf{x}_i)\right|^2 = 1. \tag{18.28}$$

The Hamiltonian expectation value is

$$\begin{aligned}
\langle H\rangle &\equiv \langle\psi(1,2,\ldots,N)|H|\psi(1,2,\ldots,N)\rangle \\
&= \sum_i \langle\prod_k \varphi_k(k)|\left(\frac{\hat{\mathbf{p}}_i^2}{2m} - \frac{Ze^2}{\hat{r}_i}\right) + \sum_{j>i}\frac{e^2}{|\hat{\mathbf{x}}_i - \hat{\mathbf{x}}_j|}|\prod_\ell \varphi_\ell(\ell)\rangle \\
&= \sum_i \langle\varphi_i(i)|\frac{\hat{\mathbf{p}}_i^2}{2m} - \frac{Ze^2}{\hat{r}_i}|\varphi_i(i)\rangle + \sum_{j>i}\langle\varphi_j(j)\varphi_i(i)|\frac{e^2}{|\hat{\mathbf{x}}_i - \hat{\mathbf{x}}_j|}|\varphi_j(j)\varphi_i(i)\rangle \\
&= \sum_i \int d^3\mathbf{x}\,\phi_i^*(\mathbf{x})\left[-\frac{\hbar^2}{2m}\Delta - \frac{Ze^2}{|\mathbf{x}|}\right]\phi_i(\mathbf{x}) + \sum_{j>i}\int d^3\mathbf{x}\,d^3\mathbf{y}\,\phi_i^*(\mathbf{x})\phi_j^*(\mathbf{y})\frac{e^2}{|\mathbf{x}-\mathbf{y}|}\phi_i(\mathbf{x})\phi_j(\mathbf{y}).
\end{aligned} \tag{18.29}$$

In the second step (third line) we used that $\varphi_k(k)$ is not affected by the operators $\hat{\mathbf{p}}_i$, $\hat{r}_i$ and $\hat{\mathbf{x}}_i$ for $i \neq k$, so that this factor drops out because of the normalisation. In the third step (fourth line) we dropped the spin part, which is not affected by the Hamiltonian and also drops out because of the normalisation.

The variational method requires the functions $\phi_i(\mathbf{x})$ to be chosen such that the right-hand side of Eq. (18.29) is minimal. In addition, the $\phi_i(\mathbf{x})$ must obey the normalisation constraint of Eq. (18.28). We take this constraint into account with Lagrange multipliers, which we call ε_i. We thus search for the functions $\phi_k(\mathbf{x})$, $k = 1,\ldots,N$, which minimise the functional

$$\begin{aligned}
\left\langle\tilde{H}\right\rangle &= \langle H\rangle + \sum_i \varepsilon_i\left(1 - \int d^3\mathbf{x}\,\left|\phi_i(\mathbf{x})\right|^2\right) \\
&= \sum_i \varepsilon_i + \int d^3\mathbf{x}\,\phi_i^*(\mathbf{x})\left[-\frac{\hbar^2}{2m}\Delta - \frac{Ze^2}{|\mathbf{x}|} - \varepsilon_i\right]\phi_i(\mathbf{x}) \\
&\quad + \sum_{j>i}\int d^3\mathbf{x}\,d^3\mathbf{y}\,\phi_i^*(\mathbf{x})\phi_j^*(\mathbf{y})\frac{e^2}{|\mathbf{x}-\mathbf{y}|}\phi_i(\mathbf{x})\phi_j(\mathbf{y}).
\end{aligned} \tag{18.30}$$

To minimise $\left\langle\tilde{H}\right\rangle$, we set for all k

$$0 = \frac{\delta\left\langle\tilde{H}\right\rangle}{\delta\phi_k^*(\mathbf{z})}. \tag{18.31}$$

Functional derivatives are explained in Appendix D. Differentiating the part of Eq. (18.30) that is summed only over i, we obtain

$$\sum_i \int d^3\mathbf{x}\,\delta_{ik}\delta(\mathbf{x}-\mathbf{z})\cdot\left[-\frac{\hbar^2}{2m}\Delta - \frac{Ze^2}{|\mathbf{x}|} - \varepsilon_i\right]\phi_i(\mathbf{x}). \tag{18.32}$$

Differentiating the second part of Eq. (18.30), we obtain

$$\begin{aligned}&\sum_{i>j}\int d^3\mathbf{x}\, d^3\mathbf{y}\, \delta_{ik}\delta(\mathbf{x}-\mathbf{z})\phi_j^*(\mathbf{y})\frac{e^2}{|\mathbf{x}-\mathbf{y}|}\phi_i(\mathbf{x})\phi_j(\mathbf{y})\\ &+\sum_{i>j}\int d^3\mathbf{x}\, d^3\mathbf{y}\, \delta_{jk}\delta(\mathbf{y}-\mathbf{z})\phi_i^*(\mathbf{x})\frac{e^2}{|\mathbf{x}-\mathbf{y}|}\phi_i(\mathbf{x})\phi_j(\mathbf{y}).\end{aligned} \tag{18.33}$$

Adding the two contributions, simplifying, and setting the sum to zero, we obtain

$$\boxed{\left[-\frac{\hbar^2}{2m}\Delta-\frac{Ze^2}{r_i}+V_i(\mathbf{x}_i)\right]\phi_i(\mathbf{x}_i)=\varepsilon_i\phi_i(\mathbf{x}_i),} \tag{18.34}$$

with

$$\boxed{V_i(\mathbf{x}_i)=\sum_{j\neq i}\int d^3\mathbf{x}_j\,\frac{e^2}{|\mathbf{x}_i-\mathbf{x}_j|}\left|\phi_j(\mathbf{x}_j)\right|^2.} \tag{18.35}$$

These are the Hartree equations for the wave functions $\phi_i(\mathbf{x})$. They can be solved by an iterative numerical procedure:

1. We make a rough guess for the wave functions $\phi_i(\mathbf{x})$ and calculate the potentials $V_i(\mathbf{x})$ numerically from Eq. (18.35).
2. With these potentials, we solve Eq. (18.34) for new wave functions $\phi_i^{(1)}(\mathbf{x})$, which will differ from our first guess.
3. We repeat the first two steps with the new wave functions $\phi_i^{(1)}(\mathbf{x})$, and continue until the changes in the wave functions become negligible.

To determine the coefficients ε_i, we multiply Eq. (18.34) by $\phi_i^*(\mathbf{x})$ and integrate over all space:

$$\int d^3\mathbf{x}\,\phi_i^*(\mathbf{x})\left[-\frac{\hbar^2}{2m}\Delta-\frac{Ze^2}{r}+V_i(\mathbf{x})\right]\phi_i(\mathbf{x})=\varepsilon_i\int d^3\mathbf{x}\left|\phi_i(\mathbf{x})\right|^2. \tag{18.36}$$

Integrating the kinetic term by parts, we obtain

$$\varepsilon_i=\int d^3\mathbf{x}\left[\frac{\hbar^2}{2m}\left|\nabla\phi_i(\mathbf{x})\right|^2-\frac{Ze^2}{r}\left|\phi_i(\mathbf{x})\right|^2\right]+\sum_{j\neq i}\int d^3\mathbf{x}\, d^3\mathbf{y}\,\left|\phi_i(\mathbf{x})\right|^2\left|\phi_j(\mathbf{y})\right|^2\frac{e^2}{|\mathbf{x}-\mathbf{y}|}, \tag{18.37}$$

where we used the identity

$$\phi^*\Delta\phi=\nabla\cdot(\phi^*\nabla\phi)-\left|\nabla\phi\right|^2 \tag{18.38}$$

and the normalisation of ϕ_i. The quantity $-\varepsilon_i$ is the "ionisation energy" needed to remove the electron occupying the state ϕ_i.[2] The total energy is

$$\begin{aligned}E=\langle H\rangle&=\left\langle\tilde{H}\right\rangle-\sum_i\varepsilon_i\left(\int d^3\mathbf{x}\left|\phi_i(\mathbf{x})\right|^2-1\right)=\left\langle\tilde{H}\right\rangle\\ &=\sum_i\varepsilon_i-\sum_{i<j}\int d^3\mathbf{x}\, d^3\mathbf{y}\,\frac{e^2}{|\mathbf{x}-\mathbf{y}|}\left|\phi_i(\mathbf{x})\right|^2\left|\phi_j(\mathbf{y})\right|^2.\end{aligned} \tag{18.39}$$

[2] This is only approximate (Koopmans' theorem): removing an electron changes the self-consistent potential felt by all the others, and this orbital relaxation is neglected.

18.2.2 The Hartree–Fock Approximation

As mentioned above, the Hartree method violates the Pauli exclusion principle. We can cure this by changing the ansatz for the wave function. In the Hartree–Fock method, the wave function is given by a *Slater determinant*,

$$\psi(1,2,\ldots,N) = \frac{1}{\sqrt{N!}} \det \begin{pmatrix} \psi_1(1) & \cdots & \psi_1(N) \\ \vdots & \ddots & \vdots \\ \psi_N(1) & \cdots & \psi_N(N) \end{pmatrix}. \tag{18.40}$$

The antisymmetry of the wave function under the exchange of two particles follows from the antisymmetry of the determinant under the exchange of two columns. The antisymmetrised self-consistent field was introduced by Fock and by Slater in 1930, and Hartree and Hartree (1935) recast it in a form practical for numerical computation. The method nevertheless became popular only in the second half of the 20th century, since solving the equations requires substantial computing power; the main difficulty is the non-local exchange term in Eq. (18.41) below.

Repeating the steps of the Hartree method, we find that the energy is minimal for

$$\begin{aligned} &\left[-\frac{\hbar^2}{2m}\Delta - \frac{Ze^2}{r_i}\right]\phi_i(\mathbf{x}) + \sum_j \int d^3\mathbf{y}\, \frac{e^2}{|\mathbf{x}-\mathbf{y}|}\phi_j^*(\mathbf{y}) \\ &\cdot \left[\phi_j(\mathbf{y})\phi_i(\mathbf{x}) - \phi_j(\mathbf{x})\phi_i(\mathbf{y})\delta_{m_{s_i} m_{s_j}}\right] = \varepsilon_i \phi_i(\mathbf{x}). \end{aligned} \tag{18.41}$$

Here

$$\begin{aligned} &\sum_j \phi_j^*(\mathbf{y}) \left[\phi_j(\mathbf{y})\phi_i(\mathbf{x}) - \phi_j(\mathbf{x})\phi_i(\mathbf{y})\delta_{m_{s_i} m_{s_j}}\right] \\ &= \underbrace{\sum_{j\neq i} \left|\phi_j(\mathbf{y})\right|^2 \phi_i(\mathbf{x}) + \left|\phi_i(\mathbf{y})\right|^2 \phi_i(\mathbf{x})}_{\text{Hartree}} \underbrace{- \sum_j \phi_j^*(\mathbf{y})\phi_j(\mathbf{x})\phi_i(\mathbf{y})\delta_{m_{s_i} m_{s_j}}}_{\text{new term (Fock)}}. \end{aligned} \tag{18.42}$$

The additional Fock term, also called the exchange term, improves the description of the ground state. The $j = i$ term of the Fock sum cancels the self-interaction $\left|\phi_i(\mathbf{y})\right|^2 \phi_i(\mathbf{x})$ contained in the Hartree part, so Eq. (18.41) is free of self-interaction (in Eq. (18.35), the self-interaction was excluded by hand). For the helium ground state ($1s^2$, opposite spins) Hartree and Hartree–Fock coincide.

19

Quantised Fields

In Chapter 16 we decomposed the electromagnetic field into an infinite set of harmonic oscillators. We quantised each of these oscillators and thereby discretised the energy of the electromagnetic field, and we associated each oscillator with a photon mode. We concluded that photons can be described as the excitations of the quantised electromagnetic field. An interesting question then arises: do all particles have underlying quantum fields, whose excitations give a complete description of them? The answer to this question is yes.

Many of the great physicists of the 20th century worked on field quantisation. The first half of the century produced quantum electrodynamics, the first quantum field theory (QFT). Later, QFT became the language of the Standard Model of particle physics, which is currently our best description of the fundamental particles. The study of QFT is therefore a cornerstone of every theoretical physicist's education.

One may ask why we need a field description of particles at all. The reason is that it is needed to describe particle interactions, in particular processes in which particles are created or destroyed. Throughout this course we have analysed interacting particles and encountered considerable difficulties time and again. QFT provides a framework in which such interactions can be described. This chapter gives an introduction to QFT and motivates the field approach to quantum mechanics.

19.1 Bosons from Quantised Fields

We want to develop a formalism that describes many-particle systems. This new formalism must not contradict our existing results. Since most of these were derived for one-particle systems, the constraint is simple: for a one-particle system, the new formalism must reproduce all previous results.

We start by defining an operator-valued function $\hat{f}(\mathbf{x}, t)$ of space and time, called the *field* or *field operator*. We claim that the field operator contains all the information about the particles of the system it describes. Furthermore, we propose that the field satisfies the same dynamical equation as the wave function of a particle. In the non-relativistic field theory, we thus postulate that $\hat{f}(\mathbf{x}, t)$ satisfies the Schrödinger equation

$$i\hbar \frac{\partial \hat{f}(\mathbf{x}, t)}{\partial t} = \underbrace{\left(-\frac{\hbar^2}{2m}\Delta + V(\mathbf{x}) \right)}_{\text{Hamiltonian}} \hat{f}(\mathbf{x}, t). \tag{19.1}$$

We should not confuse the field $\hat{f}(\mathbf{x}, t)$ with a wave function. Although both have the same dynamics, the field has many more properties than a wave function.

Let us assume that the eigenfunctions of the Hamiltonian are known,

$$\left[-\frac{\hbar^2}{2m}\Delta + V(\mathbf{x})\right]\psi_n(\mathbf{x}) = E_n\psi_n(\mathbf{x}). \tag{19.2}$$

Then the general solution for $\hat{f}(\mathbf{x}, t)$ is

$$\boxed{\hat{f}(\mathbf{x}, t) = \sum_n \hat{c}_n e^{-iE_n t/\hbar}\psi_n(\mathbf{x}).} \tag{19.3}$$

The field $\hat{f}(\mathbf{x}, t)$ differs from a wave function in that the coefficients $\hat{c}_n$ are operators and not numbers. We define

$$\hat{a}_n(t) \equiv \hat{c}_n e^{-iE_n t/\hbar}, \tag{19.4}$$

where the index n labels the energy level. Thus

$$\hat{f}(\mathbf{x}, t) = \sum_n \hat{a}_n(t)\psi_n(\mathbf{x}). \tag{19.5}$$

Inserting Eq. (19.5) into Eq. (19.1) and using Eq. (19.2), we obtain

$$\frac{d^2}{dt^2}\hat{a}_n(t) = -\frac{E_n^2}{\hbar^2}\hat{a}_n(t). \tag{19.6}$$

We can rewrite this as

$$\frac{d^2}{dt^2}\hat{a}_n(t) + \omega_n^2\hat{a}_n(t) = 0, \tag{19.7}$$

with

$$E_n = \hbar\omega_n. \tag{19.8}$$

This is the equation of motion of a harmonic oscillator with frequency ω_n. We have thus shown that a field which satisfies the Schrödinger equation can be decomposed into an infinite set of harmonic oscillators, one for each mode ψ_n.

We interpret $\hat{a}_n$ and $\hat{a}_n^\dagger$ as the annihilation and creation operators of a particle with energy E_n. We postulate the existence of a vacuum state

$$|\text{vacuum}\rangle = |0\rangle, \tag{19.9}$$

which contains no particles. The operators $\hat{a}_n$ annihilate the vacuum state,

$$\hat{a}_n|0\rangle = 0, \tag{19.10}$$

and

$$\langle 0|\hat{a}_n^\dagger = (\hat{a}_n|0\rangle)^\dagger = 0. \tag{19.11}$$

We also postulate that the creation and annihilation operators satisfy the commutation relations of independent harmonic oscillators,

$$\begin{aligned}[\hat{a}_n, \hat{a}_m] = \left[\hat{a}_n^\dagger, \hat{a}_m^\dagger\right] &= 0, \\ \left[\hat{a}_n, \hat{a}_m^\dagger\right] &= \delta_{nm}.\end{aligned} \tag{19.12}$$

In addition, we postulate that the state

$$|n\rangle = \hat{a}_n^\dagger |0\rangle \tag{19.13}$$

is a one-particle state with energy E_n.

To understand the meaning of the field operator better, we calculate the action of the conjugate field operator $\hat{f}^\dagger(\mathbf{x}, t)$ on the vacuum state:

$$\begin{aligned}
\hat{f}^\dagger(\mathbf{x}, t)\,|0\rangle &= \underbrace{\left(\sum_n |n\rangle\langle n|\right)}_{1} \hat{f}^\dagger(\mathbf{x}, t)\,|0\rangle \\
&= \sum_n |n\rangle\langle 0|\,\hat{a}_n \hat{f}^\dagger(\mathbf{x}, t)\,|0\rangle \\
&= \sum_n |n\rangle\langle 0|\,\hat{a}_n \left(\sum_m \hat{a}_m^\dagger \psi_m^*(\mathbf{x})\right)|0\rangle \\
&= \sum_{n,m} |n\rangle\langle 0|\,\hat{a}_n \hat{a}_m^\dagger\,|0\rangle\,\psi_m^*(\mathbf{x}).
\end{aligned} \tag{19.14}$$

Here we used Eq. (19.5) at $t = 0$, where $\hat{a}_n(0) = \hat{c}_n$, and only the one-particle states $|n\rangle$ contribute to the sum over intermediate states. The matrix element in the sum follows from the commutation relations (19.12):

$$\begin{aligned}
\langle 0|\,\hat{a}_n \hat{a}_m^\dagger\,|0\rangle &= \langle 0|\left[\hat{a}_n, \hat{a}_m^\dagger\right] + \hat{a}_m^\dagger \hat{a}_n\,|0\rangle \\
&= \langle 0|\,\delta_{nm}\,|0\rangle = \delta_{nm}\langle 0|0\rangle = \delta_{nm}.
\end{aligned} \tag{19.15}$$

Here we used that $\hat{a}_n$ annihilates the vacuum and that the vacuum is normalised. Inserting Eq. (19.15) into Eq. (19.14), we obtain

$$\hat{f}^\dagger(\mathbf{x}, t)\,|0\rangle = \sum_{n,m} |n\rangle\,\delta_{nm}\psi_m^*(\mathbf{x}) = \sum_n |n\rangle\,\psi_n^*(\mathbf{x}) = \sum_n |n\rangle\langle n|\mathbf{x}\rangle = |\mathbf{x}\rangle\,. \tag{19.16}$$

Thus,

$$\boxed{\hat{f}^\dagger(\mathbf{x}, t)\,|0\rangle = |\mathbf{x}\rangle\,.} \tag{19.17}$$

Strictly speaking, Eqs. (19.14)–(19.17) hold at $t = 0$. Keeping the time dependence of Eq. (19.4) one finds $\hat{f}^\dagger(\mathbf{x}, t)\,|0\rangle = \sum_n e^{iE_n t/\hbar}\psi_n^*(\mathbf{x})\,|n\rangle = e^{i\hat{H}t/\hbar}\,|\mathbf{x}\rangle$, i.e. the Heisenberg-picture position eigenstate; the same remark applies to Eq. (19.22) below.

Acting on the vacuum, the conjugate field operator $\hat{f}^\dagger(\mathbf{x}, t)$ of a given type of particle therefore creates a particle localised at $\mathbf{x}$.

We can now define multi-particle states with the conjugate field operator. A state with N particles is created by acting repeatedly with the conjugate field operator on the vacuum:

$$|\mathbf{x}_1, \mathbf{x}_2, \ldots \mathbf{x}_N\rangle \equiv \frac{1}{\sqrt{N!}} \hat{f}^\dagger(\mathbf{x}_1)\hat{f}^\dagger(\mathbf{x}_2)\ldots\hat{f}^\dagger(\mathbf{x}_N)\,|0\rangle\,, \tag{19.18}$$

where the factor $1/\sqrt{N!}$ is required for the normalisation,

$$\int d^3\mathbf{x}_1 \cdots d^3\mathbf{x}_N\ |\mathbf{x}_1, \mathbf{x}_2, \ldots \mathbf{x}_N\rangle\langle \mathbf{x}_1, \mathbf{x}_2, \ldots \mathbf{x}_N| = \mathbb{1}_N\,. \tag{19.19}$$

Position eigenstates cannot be normalised to one: with Eq. (19.50) one finds $\langle \mathbf{x}_1, \ldots, \mathbf{x}_N | \mathbf{y}_1, \ldots, \mathbf{y}_N \rangle = \frac{1}{N!} \sum_P (\pm 1)^P \prod_i \delta^{(3)}(\mathbf{x}_i - \mathbf{y}_{P(i)})$ (the sign $(-1)^P$ for the fermions of Sec. 19.2). The factor $1/\sqrt{N!}$ is precisely what makes the integral over $|\mathbf{x}_1, \ldots\rangle \langle \mathbf{x}_1, \ldots|$ the unit operator $\mathbb{1}_N$ on the (anti)symmetric N-particle space, the analogue of $\int d^3\mathbf{x}\, |\mathbf{x}\rangle \langle \mathbf{x}| = 1$ (cf. Sakurai–Napolitano, Sec. 7.5). Here the conjugate field operators are taken at a fixed time; we do not yet consider the time evolution of the system and suppress the time argument.

We can now check that the interpretation of $\hat{a}_n^\dagger |0\rangle$ as a state with energy E_n is consistent:

$$\begin{aligned} \hat{a}_n^\dagger |0\rangle &= \underbrace{\int d^3\mathbf{x}\, |\mathbf{x}\rangle \langle \mathbf{x}|}_{1} \hat{a}_n^\dagger |0\rangle \\ &= \int d^3\mathbf{x}\, |\mathbf{x}\rangle \langle 0| \hat{f}(\mathbf{x}, t) \hat{a}_n^\dagger |0\rangle \\ &= \int d^3\mathbf{x}\, |\mathbf{x}\rangle \langle 0| \sum_m \hat{a}_m \psi_m(\mathbf{x}) \hat{a}_n^\dagger |0\rangle \\ &= \int d^3\mathbf{x}\, |\mathbf{x}\rangle \sum_m \psi_m(\mathbf{x}) \langle 0| \hat{a}_m \hat{a}_n^\dagger |0\rangle . \end{aligned} \tag{19.20}$$

In the second step we used Eq. (19.17). With Eq. (19.15) we obtain

$$\begin{aligned} \hat{a}_n^\dagger |0\rangle &= \int d^3\mathbf{x}\, |\mathbf{x}\rangle \sum_m \psi_m(\mathbf{x}) \delta_{mn} \\ &= \int d^3\mathbf{x}\, |\mathbf{x}\rangle \psi_n(\mathbf{x}) = \int d^3\mathbf{x}\, |\mathbf{x}\rangle \langle \mathbf{x}| n\rangle = |n\rangle . \end{aligned} \tag{19.21}$$

Thus, the postulate of Eq. (19.13) is consistent with our field theory. So far, we have postulated the existence of one-particle states created from the vacuum, and we have shown that the conjugate field operator creates a localised particle when acting on the vacuum. In summary,

$$\boxed{\begin{aligned} |n\rangle &= \hat{a}_n^\dagger |0\rangle , \quad E_n = \hbar\omega_n , \\ |\mathbf{x}\rangle &= \hat{f}^\dagger(\mathbf{x}, t) |0\rangle . \end{aligned}} \tag{19.22}$$

We now define the number operator

$$\hat{N}_n \equiv \hat{a}_n^\dagger \hat{a}_n . \tag{19.23}$$

Physically, the number operator counts how many particles with energy E_n a state contains. To show this, we need some commutation relations. We compute

$$\left[\hat{N}_n, \hat{a}_n^\dagger\right] = \left[\hat{a}_n^\dagger \hat{a}_n, \hat{a}_n^\dagger\right] = \hat{a}_n^\dagger \underbrace{\left[\hat{a}_n, \hat{a}_n^\dagger\right]}_{1} + \underbrace{\left[\hat{a}_n^\dagger, \hat{a}_n^\dagger\right]}_{0} \hat{a}_n = \hat{a}_n^\dagger . \tag{19.24}$$

Repeating the same steps for the annihilation operator, we find the two commutation relations

$$\boxed{\left[\hat{N}_n, \hat{a}_n^\dagger\right] = \hat{a}_n^\dagger , \qquad \left[\hat{N}_n, \hat{a}_n\right] = -\hat{a}_n .} \tag{19.25}$$

Let us now consider the repeated action of the creation operator on the vacuum state:

$$(\hat{a}_n^\dagger)^m \left|0\right\rangle . \tag{19.26}$$

To find the action of the number operator on this state, we need the commutator of $\hat{N}_n$ with $(\hat{a}_n^\dagger)^m$. We want to show that

$$\left[\hat{N}_n, (\hat{a}_n^\dagger)^m\right] = m(\hat{a}_n^\dagger)^m . \tag{19.27}$$

We prove this relation by induction. The case $m = 1$ is the commutation relation of Eq. (19.25). For the induction step we assume

$$\left[\hat{N}_n, (\hat{a}_n^\dagger)^{m-1}\right] = (m-1)(\hat{a}_n^\dagger)^{m-1} . \tag{19.28}$$

Then we have

$$\begin{aligned} \left[\hat{N}_n, (\hat{a}_n^\dagger)^m\right] &= \left[\hat{N}_n, \hat{a}_n^\dagger(\hat{a}_n^\dagger)^{m-1}\right] \\ &= \hat{a}_n^\dagger \left[\hat{N}_n, (\hat{a}_n^\dagger)^{m-1}\right] + \left[\hat{N}_n, \hat{a}_n^\dagger\right](\hat{a}_n^\dagger)^{m-1} \\ &= \hat{a}_n^\dagger (m-1)(\hat{a}_n^\dagger)^{m-1} + \hat{a}_n^\dagger(\hat{a}_n^\dagger)^{m-1} \\ &= m(\hat{a}_n^\dagger)^m , \end{aligned} \tag{19.29}$$

which proves the proposition. We can now calculate the action of the number operator on the state of Eq. (19.26):

$$\begin{aligned} \hat{N}_n(\hat{a}_n^\dagger)^m \left|0\right\rangle &= \left\{ \left[\hat{N}_n, (\hat{a}_n^\dagger)^m\right] + (\hat{a}_n^\dagger)^m \hat{N}_n \right\} \left|0\right\rangle \\ &= m(\hat{a}_n^\dagger)^m \left|0\right\rangle , \end{aligned} \tag{19.30}$$

where in the second step we used the commutation relation just derived and the fact that the right-most factor of the number operator is an annihilation operator, which annihilates the vacuum. We have thus shown that the number operator $\hat{N}_n$ counts the number of particles with energy E_n in a state. We simplify the notation for such states by writing

$$\left|N_n\right\rangle \equiv \frac{(\hat{a}_n^\dagger)^{N_n}}{\sqrt{N_n!}} \left|0\right\rangle , \tag{19.31}$$

where we have now included a normalisation constant. We can extend this notation to multi-particle states with several occupied energy levels. A state with N_i particles in the energy level E_i is denoted by

$$\left|N_1, N_2, \ldots\right\rangle . \tag{19.32}$$

As an example, consider the state

$$\left|1, 0, 2\right\rangle . \tag{19.33}$$

This state contains one particle in the energy level E_1, no particle in the level E_2 and two particles in the level E_3. In terms of creation operators, such states read

$$\left|N_1, N_2, \ldots N_n\right\rangle = \frac{(\hat{a}_1^\dagger)^{N_1} \ldots (\hat{a}_n^\dagger)^{N_n}}{\sqrt{N_1! \ldots N_n!}} \left|0\right\rangle . \tag{19.34}$$

We can also relate the state with N_n particles in the energy level E_n to the states with $N_n \pm 1$ particles in the same level:

$$\left|N_n + 1\right\rangle = \frac{(\hat{a}_n^\dagger)^{N_n+1}}{\sqrt{(N_n+1)!}} \left|0\right\rangle = \frac{(\hat{a}_n^\dagger)}{\sqrt{N_n+1}} \frac{(\hat{a}_n^\dagger)^{N_n}}{\sqrt{(N_n)!}} \left|0\right\rangle = \frac{\hat{a}_n^\dagger}{\sqrt{N_n+1}} \left|N_n\right\rangle . \tag{19.35}$$

For the other direction, we need the commutation relation of the following exercise.

Exercise

Show that $\left[\hat{a}_n, (\hat{a}_n^\dagger)^m\right] = m(\hat{a}_n^\dagger)^{m-1}$ (19.36)

With it, we find directly

$$\begin{aligned}\hat{a}_n |N_n\rangle &= \frac{\hat{a}_n(\hat{a}_n^\dagger)^{N_n}}{\sqrt{(N_n)!}}|0\rangle = \frac{\left[\hat{a}_n, (\hat{a}_n^\dagger)^{N_n}\right] + (\hat{a}_n^\dagger)^{N_n}\hat{a}_n}{\sqrt{(N_n)!}}|0\rangle \\ &= \frac{N_n(\hat{a}_n^\dagger)^{N_n-1}}{\sqrt{(N_n)!}}|0\rangle = \sqrt{N_n}\,|N_n-1\rangle\,.\end{aligned} \tag{19.37}$$

In summary, the creation operator increases the number of particles by one, whereas the annihilation operator reduces it by one:

$$\boxed{\begin{aligned}\hat{a}_n^\dagger |N_n\rangle &= \sqrt{N_n+1}\,|N_n+1\rangle\,, \\ \hat{a}_n |N_n\rangle &= \sqrt{N_n}\,|N_n-1\rangle\,.\end{aligned}} \tag{19.38}$$

The way we have quantised the field operator gives rise to multi-particle systems of *identical bosons*. We show this for a two-particle system with one particle in the energy level n (with energy E_n) and the other in the level m (with energy E_m). The wave function of this state is

$$\begin{aligned}\psi_{nm}(\mathbf{x}_1,\mathbf{x}_2) \equiv \langle \mathbf{x}_1,\mathbf{x}_2|n,m\rangle &= \frac{1}{\sqrt{2}}\langle 0|\,\hat{f}(\mathbf{x}_2)\hat{f}(\mathbf{x}_1)\hat{a}_n^\dagger\hat{a}_m^\dagger\,|0\rangle \\ &= \frac{1}{\sqrt{2}}\sum_{k,\ell}\psi_k(\mathbf{x}_2)\psi_\ell(\mathbf{x}_1)\,\langle 0|\,\hat{a}_k\hat{a}_\ell\hat{a}_n^\dagger\hat{a}_m^\dagger\,|0\rangle\,.\end{aligned} \tag{19.39}$$

The factor $1/\sqrt{2}$ comes from the normalisation of the two-particle position state, Eq. (19.18), and in the second step we inserted the mode expansion of the field operators. We simplify the expression as in Eqs. (19.14) and (19.15). The vacuum expectation value in the sum is

$$\begin{aligned}\langle 0|\,\hat{a}_k\hat{a}_\ell\hat{a}_n^\dagger\hat{a}_m^\dagger\,|0\rangle &= \langle 0|\,\hat{a}_k\left[\hat{a}_\ell,\hat{a}_n^\dagger\right]\hat{a}_m^\dagger + \hat{a}_k\hat{a}_n^\dagger\hat{a}_\ell\hat{a}_m^\dagger\,|0\rangle \\ &= \langle 0|\,\hat{a}_k\hat{a}_m^\dagger\delta_{\ell n} + \left(\left[\hat{a}_k,\hat{a}_n^\dagger\right] + \hat{a}_n^\dagger\hat{a}_k\right)\left(\left[\hat{a}_\ell,\hat{a}_m^\dagger\right] + \hat{a}_m^\dagger\hat{a}_\ell\right)|0\rangle \\ &= \langle 0|\left(\left[\hat{a}_k,\hat{a}_m^\dagger\right] + \hat{a}_m^\dagger\hat{a}_k\right)\delta_{\ell n} + \delta_{kn}\delta_{\ell m}\,|0\rangle \\ &= \delta_{km}\delta_{\ell n} + \delta_{kn}\delta_{\ell m},\end{aligned} \tag{19.40}$$

where we used that any annihilation operator acting on the vacuum $|0\rangle$ gives zero and that the vacuum is normalised. Inserting this result into Eq. (19.39), we obtain

$$\boxed{\psi_{nm}(\mathbf{x}_1,\mathbf{x}_2) = \frac{1}{\sqrt{2}}\left[\psi_n(\mathbf{x}_1)\psi_m(\mathbf{x}_2) + \psi_n(\mathbf{x}_2)\psi_m(\mathbf{x}_1)\right].} \tag{19.41}$$

As expected for bosons, the wave function is symmetric. Eq. (19.41) assumes $n \neq m$. For $n = m$ the state $\hat{a}_n^\dagger\hat{a}_n^\dagger|0\rangle$ has norm $\sqrt{2}$; the normalised state $|N_n = 2\rangle$ of Eq. (19.34) has the wave function $\psi_n(\mathbf{x}_1)\psi_n(\mathbf{x}_2)$.

Let us summarise our derivation of bosonic multi-particle states from quantised fields.

(i) We postulate the existence of an operator-valued quantum field satisfying the Schrödinger equation,

$$i\hbar\frac{\partial\hat{f}(\mathbf{x},t)}{\partial t}=\left[-\frac{\hbar^2}{2m}\Delta+V(\mathbf{x})\right]\hat{f}(\mathbf{x},t). \tag{19.42}$$

(ii) We express the field as an infinite collection of quantised harmonic oscillators,

$$\hat{f}(\mathbf{x},t)=\sum_n \hat{a}_n(t)\psi_n(\mathbf{x}), \tag{19.43}$$

with

$$\left[-\frac{\hbar^2}{2m}\Delta+V(\mathbf{x})\right]\psi_n(\mathbf{x})=E_n\psi_n(\mathbf{x}) \tag{19.44}$$

and the commutation relations

$$\begin{aligned}[\hat{a}_n,\hat{a}_m]=\left[\hat{a}_n^\dagger,\hat{a}_m^\dagger\right]&=0,\\ \left[\hat{a}_n,\hat{a}_m^\dagger\right]&=\delta_{nm}.\end{aligned} \tag{19.45}$$

(iii) Multi-particle states with definite particle positions are

$$|\mathbf{x}_1,\mathbf{x}_2,\ldots,\mathbf{x}_N\rangle=\frac{\hat{f}^\dagger(\mathbf{x}_1)\ldots\hat{f}^\dagger(\mathbf{x}_N)}{\sqrt{N!}}|0\rangle\,, \tag{19.46}$$

where $|0\rangle$ is the vacuum state.

(iv) Multi-particle states with definite occupation numbers of the energy levels are

$$|N_1,\ldots N_n\rangle=\frac{(\hat{a}_1^\dagger)^{N_1}\ldots(\hat{a}_n^\dagger)^{N_n}}{\sqrt{N_1!\ldots N_n!}}|0\rangle\,. \tag{19.47}$$

As an aside, we note that our route via the creation and annihilation operators is rather indirect. Equivalently, we could have imposed commutation relations on the field and its hermitian conjugate. Indeed,

$$\begin{aligned}\left[\hat{f}(\mathbf{x},t),\hat{f}^\dagger(\mathbf{x}',t)\right]&=\left[\sum_m\hat{a}_m\psi_m(\mathbf{x}),\sum_n\hat{a}_n^\dagger\psi_n^*(\mathbf{x}')\right]\\ &=\sum_{n,m}\psi_m(\mathbf{x})\psi_n^*(\mathbf{x}')\left[\hat{a}_m,\hat{a}_n^\dagger\right]\\ &=\sum_n\psi_n(\mathbf{x})\psi_n^*(\mathbf{x}')=\sum_n\langle\mathbf{x}|n\rangle\langle n|\mathbf{x}'\rangle\\ &=\langle\mathbf{x}|\mathbf{x}'\rangle=\delta^{(3)}(\mathbf{x}-\mathbf{x}').\end{aligned} \tag{19.48}$$

Analogously, we obtain

$$\left[\hat{f}(\mathbf{x},t),\hat{f}(\mathbf{x}',t)\right]=\left[\hat{f}^\dagger(\mathbf{x},t),\hat{f}^\dagger(\mathbf{x}',t)\right]=0. \tag{19.49}$$

We could therefore have replaced the quantisation conditions on the creation and annihilation operators by the equivalent conditions

$$\boxed{\begin{aligned}\left[\hat{f}(\mathbf{x},t),\hat{f}(\mathbf{x}',t)\right]=\left[\hat{f}^\dagger(\mathbf{x},t),\hat{f}^\dagger(\mathbf{x}',t)\right]&=0,\\ \left[\hat{f}(\mathbf{x},t),\hat{f}^\dagger(\mathbf{x}',t)\right]&=\delta^{(3)}(\mathbf{x}-\mathbf{x}').\end{aligned}} \tag{19.50}$$

19.2 Fermions from Quantised Fields

Field quantisation allows us to describe states with many bosons. We would also like to describe fermions with the same formalism. This requires a small modification of the bosonic derivation: the introduction of the *anticommutator*. We start again by postulating the existence of a field which satisfies the same Schrödinger equation as the particles it describes,

$$i\hbar \frac{\partial \hat{f}(\mathbf{x},t)}{\partial t} = \left[-\frac{\hbar^2}{2m}\Delta + V(\mathbf{x})\right] \hat{f}(\mathbf{x},t). \tag{19.51}$$

As in the boson case, the general solution for the field is

$$\hat{f}(\mathbf{x},t) = \sum_n \hat{c}_n e^{-iE_n t/\hbar} \psi_n(\mathbf{x}), \tag{19.52}$$

where the functions $\psi_n(\mathbf{x})$ are the familiar energy eigenfunctions,

$$\left[-\frac{\hbar^2}{2m}\Delta + V(\mathbf{x})\right] \psi_n(\mathbf{x}) = E_n \psi_n(\mathbf{x}). \tag{19.53}$$

We also define an annihilation operator

$$\hat{a}_n = \hat{c}_n e^{-iE_n t/\hbar}, \tag{19.54}$$

and rewrite the field operator as

$$\hat{f}(\mathbf{x},t) = \sum_n \hat{a}_n \psi_n(\mathbf{x}). \tag{19.55}$$

The annihilation operators annihilate the vacuum,

$$\hat{a}_n |0\rangle = 0. \tag{19.56}$$

In addition, we postulate that the state

$$|n\rangle \equiv \hat{a}_n^\dagger |0\rangle \tag{19.57}$$

is a one-particle state with energy E_n.

So far, all steps are identical to those for the bosonic field. To obtain fermionic multi-particle states, we quantise with *anticommutators* instead of commutators. We demand that the creation and annihilation operators fulfil the anticommutation relations

$$\begin{aligned} \{\hat{a}_n, \hat{a}_m\} = \left\{\hat{a}_n^\dagger, \hat{a}_m^\dagger\right\} &= 0, \\ \left\{\hat{a}_n, \hat{a}_m^\dagger\right\} &= \delta_{nm}, \end{aligned} \tag{19.58}$$

where the anticommutator of two operators is defined as

$$\left\{\hat{A}, \hat{B}\right\} = \hat{A}\hat{B} + \hat{B}\hat{A}. \tag{19.59}$$

As for bosons, we expect that the conjugate field operator creates a particle localised at the position $\mathbf{x}$,

$$\hat{f}^\dagger(\mathbf{x},t) |0\rangle = |\mathbf{x}\rangle . \tag{19.60}$$

This follows from a straightforward calculation:

$$\begin{aligned}\hat{f}^\dagger(\mathbf{x},t)\,|0\rangle &= \underbrace{\left(\sum_n |n\rangle\langle n|\right)}_{1}\hat{f}^\dagger(\mathbf{x},t)\,|0\rangle \\ &= \sum_n |n\rangle\langle 0|\,\hat{a}_n\hat{f}^\dagger(\mathbf{x},t)\,|0\rangle \\ &= \sum_n |n\rangle\langle 0|\,\hat{a}_n\left(\sum_m \hat{a}_m^\dagger\psi_m^*(\mathbf{x})\right)|0\rangle \\ &= \sum_{n,m} |n\rangle\langle 0|\,\hat{a}_n\hat{a}_m^\dagger\,|0\rangle\,\psi_m^*(\mathbf{x}).\end{aligned} \tag{19.61}$$

The matrix element in the sum follows from the anticommutation relations (19.58):

$$\begin{aligned}\langle 0|\,\hat{a}_n\hat{a}_m^\dagger\,|0\rangle &= \langle 0|\left\{\hat{a}_n,\hat{a}_m^\dagger\right\} - \hat{a}_m^\dagger\hat{a}_n\,|0\rangle \\ &= \langle 0|\,\delta_{nm}\,|0\rangle = \delta_{nm}\langle 0|0\rangle = \delta_{nm},\end{aligned} \tag{19.62}$$

where we used that $\hat{a}_n$ annihilates the vacuum and that the vacuum is normalised. Inserting Eq. (19.62) into Eq. (19.61), we obtain

$$\hat{f}^\dagger(\mathbf{x},t)\,|0\rangle = \sum_{n,m}|n\rangle\,\delta_{nm}\psi_m^*(\mathbf{x}) = \sum_n |n\rangle\,\psi_n^*(\mathbf{x}) = \sum_n |n\rangle\langle n|\mathbf{x}\rangle = |\mathbf{x}\rangle\,. \tag{19.63}$$

We see that quantisation with anticommutators leads to the same description of one-particle energy and position eigenstates as in the case of bosons. The results of Eqs. (19.20)–(19.38) for the bosonic field can be carried over to the fermionic field with the anticommutation relations, but with important modifications. The operator $\hat{N}_n = \hat{a}_n^\dagger\hat{a}_n$ still satisfies Eq. (19.25), but now $(\hat{a}_n^\dagger)^2 = \frac{1}{2}\left\{\hat{a}_n^\dagger,\hat{a}_n^\dagger\right\} = 0$, so that $N_n \in \{0,1\}$: this is the Pauli principle. Consequently, the commutator of the exercise above fails for $m = 2$, and $\hat{a}_n^\dagger\,|1_n\rangle = 0$ instead of $\sqrt{2}\,|2_n\rangle$. Moreover, in multi-mode states $\hat{a}_n^\dagger$, $\hat{a}_n$ pick up the sign $(-1)^{\sum_{j<n}N_j}$ from the ordering in Eq. (19.34) (Sakurai–Napolitano, Sec. 7.5).

For multi-particle systems, the two formalisms differ. Let us consider the two-particle state

$$|n,m\rangle \equiv \hat{a}_n^\dagger\hat{a}_m^\dagger\,|0\rangle\,, \tag{19.64}$$

with one particle in the energy level E_n and the other in the level E_m. Its wave function is

$$\psi_{nm}(\mathbf{x}_1,\mathbf{x}_2) = \langle\mathbf{x}_1,\mathbf{x}_2|n,m\rangle\,, \tag{19.65}$$

where the state $|\mathbf{x}_1,\mathbf{x}_2\rangle$ with two particles at the positions $\mathbf{x}_1$ and $\mathbf{x}_2$ is obtained from the field operators and the vacuum,

$$|\mathbf{x}_1,\mathbf{x}_2\rangle = \frac{1}{\sqrt{2}}\hat{f}^\dagger(\mathbf{x}_1,t)\hat{f}^\dagger(\mathbf{x}_2,t)\,|0\rangle\,. \tag{19.66}$$

We calculate the wave function ψ_{nm} by inserting Eqs. (19.66) and (19.64) into Eq. (19.65):

$$\begin{aligned}\psi_{nm}(\mathbf{x}_1,\mathbf{x}_2) &= \frac{1}{\sqrt{2}}\langle 0|\hat{f}(\mathbf{x}_2,t)\hat{f}(\mathbf{x}_1,t)\hat{a}_n^\dagger\hat{a}_m^\dagger|0\rangle \\ &= \frac{1}{\sqrt{2}}\langle 0|\left(\sum_k \hat{a}_k\psi_k(\mathbf{x}_2)\right)\left(\sum_\ell \hat{a}_\ell\psi_\ell(\mathbf{x}_1)\right)\hat{a}_n^\dagger\hat{a}_m^\dagger|0\rangle \\ &= \frac{1}{\sqrt{2}}\sum_{k,\ell}\psi_k(\mathbf{x}_2)\psi_\ell(\mathbf{x}_1)\,\langle 0|\hat{a}_k\hat{a}_\ell\hat{a}_n^\dagger\hat{a}_m^\dagger|0\rangle\,.\end{aligned} \tag{19.67}$$

To evaluate the sum, we first compute the vacuum expectation value inside it, as we did for the bosonic field:

$$\begin{aligned}\langle 0|\,\hat{a}_k\hat{a}_\ell\hat{a}_n^\dagger\hat{a}_m^\dagger\,|0\rangle &= \langle 0|\,\hat{a}_k\left\{\hat{a}_\ell,\hat{a}_n^\dagger\right\}\hat{a}_m^\dagger - \hat{a}_k\hat{a}_n^\dagger\hat{a}_\ell\hat{a}_m^\dagger\,|0\rangle \\ &= \langle 0|\,\hat{a}_k\hat{a}_m^\dagger\delta_{\ell n} - \left(\left\{\hat{a}_k,\hat{a}_n^\dagger\right\} - \hat{a}_n^\dagger\hat{a}_k\right)\left(\left\{\hat{a}_\ell,\hat{a}_m^\dagger\right\} - \hat{a}_m^\dagger\hat{a}_\ell\right)|0\rangle \\ &= \langle 0|\left(\left\{\hat{a}_k,\hat{a}_m^\dagger\right\} - \hat{a}_m^\dagger\hat{a}_k\right)\delta_{\ell n} - \delta_{kn}\delta_{\ell m}\,|0\rangle \\ &= \delta_{km}\delta_{\ell n} - \delta_{kn}\delta_{\ell m},\end{aligned} \tag{19.68}$$

where we used that any annihilation operator acting on the vacuum gives zero, and the normalisation of the vacuum. Substituting this result into Eq. (19.67), we obtain

$$\psi_{nm}(\mathbf{x}_1,\mathbf{x}_2) = \frac{1}{\sqrt{2}}\left[\psi_n(\mathbf{x}_1)\psi_m(\mathbf{x}_2) - \psi_n(\mathbf{x}_2)\psi_m(\mathbf{x}_1)\right]. \tag{19.69}$$

Thus, field quantisation with anticommutators instead of commutators describes multi-particle fermion states. To summarise, the quantisation of bosonic and fermionic multi-particle systems differs only in the choice of commutation (bosons) or anticommutation (fermions) relations. Commutators give rise to symmetric multi-particle states, as expected for bosons; anticommutators give rise to antisymmetric multi-particle states, as expected for fermions.

19.3 Field Quantisation: Number of Particles and Total Energy

The material of this and the following sections goes beyond what was covered in the lectures.

19.3.1 Particle Density Operator

In quantum mechanics, the probability density for finding a particle at the position $\mathbf{x}$ is

$$\rho(\mathbf{x},t) \overset{\text{QM}}{=} \left|\psi(\mathbf{x},t)\right|^2. \tag{19.70}$$

In the formalism of field quantisation, we can associate the probability density directly with the field operator. Consider the analogue of Eq. (19.70) in which the wave function is replaced by the field operator,

$$\hat{\rho}(\mathbf{x},t) \equiv \hat{f}^\dagger(\mathbf{x},t)\hat{f}(\mathbf{x},t). \tag{19.71}$$

Substituting the mode expansion

$$\hat{f}(\mathbf{x},t) = \sum_n \hat{a}_n\psi_n(\mathbf{x}) = \sum_n \hat{a}_n\,\langle\mathbf{x}|n\rangle \tag{19.72}$$

into Eq. (19.71), we obtain

$$\hat{\rho}(\mathbf{x},t) = \sum_{n,m}\hat{a}_n^\dagger\hat{a}_m\,\langle n|\mathbf{x}\rangle\,\langle\mathbf{x}|m\rangle\,. \tag{19.73}$$

Integrating $\hat{\rho}(\mathbf{x},t)$ over all space,

$$\begin{aligned}\int d^3\mathbf{x}\,\hat{\rho}(\mathbf{x},t) &= \sum_{n,m}\hat{a}_n^\dagger\hat{a}_m\,\langle n|\underbrace{\int d^3\mathbf{x}\;|\mathbf{x}\rangle\,\langle\mathbf{x}|}_{1}|m\rangle = \sum_{n,m}\hat{a}_n^\dagger\hat{a}_m\,\langle n|m\rangle \\ &= \sum_{n,m}\hat{a}_n^\dagger\hat{a}_m\delta_{m,n} = \sum_n\hat{a}_n^\dagger\hat{a}_n = \sum_n\hat{N}_n,\end{aligned} \tag{19.74}$$

we obtain the sum of the number operators of all energy levels. Let us consider a state

$$|N_1, N_2, ..., N_m\rangle, \tag{19.75}$$

in which N_i of a total of N_{tot} particles are in the energy level E_i. Using Eq. (19.74), the expectation value of the integrated density operator $\hat{\rho}(\mathbf{x}, t)$ is

$$\begin{aligned}\int d^3\mathbf{x}\, \langle N_1, ...N_m|\, \hat{\rho}(\mathbf{x}, t)\, |N_1, ..., N_m\rangle &= \langle N_1, ...N_m| \sum_{i=1}^{m} \hat{N}_i\, |N_1, ..., N_m\rangle \\ &= \sum_{i=1}^{m} N_i\, \langle N_1, ...N_m|N_1, ..., N_m\rangle \\ &= \sum_{i=1}^{m} N_i = N_{\text{tot}}.\end{aligned} \tag{19.76}$$

19.3.2 Energy and Energy Density (Hamiltonian) Operators

In analogy with the Hamiltonian of quantum mechanics, we construct the operator

$$\hat{H}_F \equiv \int d^3\mathbf{x}\, \hat{f}^\dagger(\mathbf{x}, t)\hat{H}\hat{f}(\mathbf{x}, t), \tag{19.77}$$

where the subscript "F" stands for "field" and $\hat{H}$ is the quantum-mechanical Hamiltonian

$$\hat{H} = -\frac{\hbar^2}{2m}\Delta + V(\mathbf{x}). \tag{19.78}$$

Inserting Eq. (19.72) into Eq. (19.77), we obtain

$$\begin{aligned}\hat{H}_F &= \int d^3\mathbf{x} \left(\sum_n \hat{a}_n^\dagger\, \langle n|\mathbf{x}\rangle\right) \hat{H} \left(\sum_m \hat{a}_m\, \langle \mathbf{x}|m\rangle\right) \\ &= \sum_{n,m} \hat{a}_n^\dagger \hat{a}_m E_m\, \langle n| \underbrace{\int d^3\mathbf{x}\, |\mathbf{x}\rangle\, \langle \mathbf{x}|}_{1} |m\rangle \\ &= \sum_{n,m} \hat{a}_n^\dagger \hat{a}_m E_m \delta_{nm}.\end{aligned} \tag{19.79}$$

Thus we have

$$\boxed{\hat{H}_F = \sum_n E_n \hat{N}_n.} \tag{19.80}$$

The operator $\hat{H}_F$ is the Hamiltonian, or energy operator, of the field. In the state

$$|N_1, ..., N_m\rangle, \tag{19.81}$$

the expectation value of $\hat{H}_F$ is

$$\begin{aligned}\langle N_1, ..., N_m|\, \hat{H}_F\, |N_1, ..., N_m\rangle &= \sum_n E_n\, \langle N_1, ..., N_m|\, \hat{N}_n\, |N_1, ..., N_m\rangle \\ &= \sum_n E_n N_n = E_{\text{tot}},\end{aligned} \tag{19.82}$$

which is the total energy of the particles populating the various energy levels. The energy density operator $\hat{\mathcal{H}}_F$ is defined by

$$\hat{H}_F = \int d^3\mathbf{x}\, \hat{\mathcal{H}}_F. \tag{19.83}$$

Comparing with Eq. (19.77), we see that the energy density operator is

$$\begin{aligned} \hat{\mathcal{H}}_F &= \hat{f}^\dagger(\mathbf{x},t)\hat{H}\hat{f}(\mathbf{x},t) = \hat{f}^\dagger(\mathbf{x},t)\left[-\frac{\hbar^2}{2m}\Delta + V(\mathbf{x})\right]\hat{f}(\mathbf{x},t) \\ &= -\frac{\hbar^2}{2m}\nabla\left(\hat{f}^\dagger(\mathbf{x},t)\nabla\hat{f}(\mathbf{x},t)\right) + \frac{\hbar^2}{2m}\left|\nabla\hat{f}(\mathbf{x},t)\right|^2 + V(\mathbf{x})\left|\hat{f}(\mathbf{x},t)\right|^2. \end{aligned} \tag{19.84}$$

The first of the three terms is a total divergence and does not contribute to the total energy after integration in Eq. (19.83). We therefore redefine the field Hamiltonian density as (for operators, $\left|\nabla\hat{f}\right|^2$ and $\left|\hat{f}\right|^2$ stand for $(\nabla\hat{f}^\dagger)\cdot(\nabla\hat{f})$ and $\hat{f}^\dagger\hat{f}$, with $\hat{f}^\dagger$ to the left; the surface term vanishes for fields that decay at infinity):

$$\boxed{\hat{\mathcal{H}}_F \equiv \frac{\hbar^2}{2m}\left|\nabla\hat{f}(\mathbf{x},t)\right|^2 + V(\mathbf{x})\left|\hat{f}(\mathbf{x},t)\right|^2.} \tag{19.85}$$

19.4 Non-Relativistic Quantum Field Theory of Interacting Fields

So far, we have described systems of particles which do not interact with each other. We now extend the field Hamiltonian so that it incorporates interactions between the particles:

$$\begin{aligned} \hat{H}_F = &\int d^3\mathbf{x}\, \hat{f}^\dagger(\mathbf{x},t)\hat{H}_0\hat{f}(\mathbf{x},t) \\ &+ \frac{1}{2}\int d^3\mathbf{x}\, d^3\mathbf{x}'\, \hat{f}^\dagger(\mathbf{x},t)\hat{f}^\dagger(\mathbf{x}',t)V(\mathbf{x},\mathbf{x}')\hat{f}(\mathbf{x}',t)\hat{f}(\mathbf{x},t), \end{aligned} \tag{19.86}$$

with

$$\hat{H}_0 = -\frac{\hbar^2}{2m}\Delta + V(\mathbf{x}). \tag{19.87}$$

The first term in Eq. (19.86) describes independent particles in a common external potential; the second term, containing $V(\mathbf{x},\mathbf{x}')$, describes their mutual interactions. Let us examine whether this ansatz reproduces the known results of quantum mechanics. For brevity, we treat commutator quantisation (bosons) and anticommutator quantisation (fermions) simultaneously. We introduce the symbol

$$\left[\hat{A},\hat{B}\right]_\pm = \hat{A}\hat{B} \pm \hat{B}\hat{A}, \tag{19.88}$$

so that

$$\left[\hat{A},\hat{B}\right]_- = \left[\hat{A},\hat{B}\right] \quad \text{and} \quad \left[\hat{A},\hat{B}\right]_+ = \left\{\hat{A},\hat{B}\right\}. \tag{19.89}$$

It is now convenient to impose the (anti)commutation relations on the field operators rather than on the operators $\hat{a}$ and $\hat{a}^\dagger$:

$$\begin{aligned} \left[\hat{f}(\mathbf{x},t),\hat{f}(\mathbf{x}',t)\right]_\pm = \left[\hat{f}^\dagger(\mathbf{x},t),\hat{f}^\dagger(\mathbf{x}',t)\right]_\pm = 0, \\ \left[\hat{f}(\mathbf{x},t),\hat{f}^\dagger(\mathbf{x}',t)\right]_\pm = \delta^{(3)}(\mathbf{x}-\mathbf{x}'). \end{aligned} \tag{19.90}$$

Here the upper sign holds for fermions and the lower sign for bosons; for a given field, only one of the two sets of relations is imposed. For a multi-particle system with mutual interactions, we cannot assign an independent potential to each particle, and it is difficult to solve a Schrödinger equation for the field. We now postulate that a state of many particles at definite positions,

$$|\mathbf{x}_1, ..., \mathbf{x}_N\rangle\,, \tag{19.91}$$

is created by the fields acting on the vacuum state:

$$|\mathbf{x}_1, ..., \mathbf{x}_N\rangle = \frac{\hat{f}^\dagger(\mathbf{x}_1, t)\ldots\hat{f}^\dagger(\mathbf{x}_N, t)}{\sqrt{N!}}\,|0\rangle\,. \tag{19.92}$$

A general N-particle state is a linear combination of the states of Eq. (19.92):

$$|\phi_N, t\rangle \equiv \int d^3\mathbf{x}_1 \ldots \int d^3\mathbf{x}_N\, \phi_N(\mathbf{x}_1, \ldots, \mathbf{x}_N) \cdot \frac{\hat{f}^\dagger(\mathbf{x}_1, t)\ldots\hat{f}^\dagger(\mathbf{x}_N, t)}{\sqrt{N!}}\,|0\rangle\,, \tag{19.93}$$

where $\phi_N(\mathbf{x}_1, \ldots, \mathbf{x}_N)$ is a wave function whose meaning we discuss below. The probability of finding the N particles at the positions $\mathbf{x}_1, \ldots, \mathbf{x}_N$ is

$$\begin{aligned} dP(\mathbf{x}_1, \ldots, \mathbf{x}_N) &= d^3\mathbf{x}_1 \ldots d^3\mathbf{x}_N \left|\langle \mathbf{x}_1, \ldots, \mathbf{x}_N|\phi_N, t\rangle\right|^2 \\ &= d^3\mathbf{x}_1 \ldots d^3\mathbf{x}_N \left|\phi_N(\mathbf{x}_1, \ldots, \mathbf{x}_N)\right|^2. \end{aligned} \tag{19.94}$$

The last equality requires ϕ_N to be symmetric (bosons) or antisymmetric (fermions); in general $\langle \mathbf{x}_1, \ldots, \mathbf{x}_N|\phi_N, t\rangle$ is the (anti)symmetrised part of ϕ_N, since only that part survives in Eq. (19.93). Without loss of generality we take ϕ_N (anti)symmetric from now on. Strictly, dP is the probability to find one particle in each volume element $d^3\mathbf{x}_i$ with the labelling fixed; summing over the $N!$ relabellings of distinct volume elements it is $N!\; dP$. We now demonstrate that the field-theoretical description with $\hat{H}_F$ is equivalent to the familiar quantum-mechanical picture. We prove that if a multi-particle state $|\phi_N, t\rangle$ is an eigenstate of the field Hamiltonian $\hat{H}_F$ with energy eigenvalue E_N,

$$\boxed{\hat{H}_F\,|\phi_N, t\rangle = E_N\,|\phi_N, t\rangle} \tag{19.95}$$

then the wave function $\phi_N(\mathbf{x}_1, \ldots, \mathbf{x}_N)$ is an eigenfunction of the usual quantum-mechanical Hamiltonian with the same energy E_N,

$$\boxed{\hat{H}_{\mathrm{QM}}\phi_N(\mathbf{x}_1, \ldots, \mathbf{x}_N) = E_N\phi_N(\mathbf{x}_1, \ldots, \mathbf{x}_N),} \tag{19.96}$$

where the quantum-mechanical Hamiltonian of the multi-particle system is

$$\hat{H}_{\mathrm{QM}} = \sum_{i=1}^{N}\left[-\frac{\hbar^2}{2m}\Delta_i + V(\mathbf{x}_i)\right] + \frac{1}{2}\sum_{i\neq j} V(\mathbf{x}_i, \mathbf{x}_j). \tag{19.97}$$

Proof: We decompose $\hat{H}_F$ of Eq. (19.86) as

$$\hat{H}_F = \hat{H}_{0,F} + \hat{V}_F, \tag{19.98}$$

with

$$\begin{aligned}\hat{H}_{0,F} &= \int d^3\mathbf{x}\, \hat{f}^\dagger(\mathbf{x},t)\hat{H}_0\hat{f}(\mathbf{x},t),\\ \hat{V}_F &= \frac{1}{2}\int d^3\mathbf{x}\, d^3\mathbf{x}'\; \hat{f}^\dagger(\mathbf{x},t)\hat{f}^\dagger(\mathbf{x}',t)V(\mathbf{x},\mathbf{x}')\hat{f}(\mathbf{x}',t)\hat{f}(\mathbf{x},t).\end{aligned} \tag{19.99}$$

We consider the case $N = 2$; the general case is proven analogously. We have

$$\begin{aligned}\hat{H}_{0,F}\,|\phi_2,t\rangle &= \int d^3\mathbf{x}\, \hat{f}^\dagger(\mathbf{x},t)\hat{H}_0\hat{f}(\mathbf{x},t)\,|\phi_2,t\rangle\\ &= \int d^3\mathbf{x}\int d^3\mathbf{x}_1\int d^3\mathbf{x}_2\, \hat{f}^\dagger(\mathbf{x},t)\hat{H}_0\hat{f}(\mathbf{x},t)\phi_2(\mathbf{x}_1,\mathbf{x}_2)\cdot\frac{\hat{f}^\dagger(\mathbf{x}_1,t)\hat{f}^\dagger(\mathbf{x}_2,t)}{\sqrt{2!}}\,|0\rangle\\ &= \int d^3\mathbf{x}\int d^3\mathbf{x}_1\int d^3\mathbf{x}_2\, \hat{f}^\dagger(\mathbf{x},t)\left[-\frac{\hbar^2}{2m}\Delta + V(\mathbf{x})\right]\hat{f}(\mathbf{x},t)\\ &\quad\cdot\frac{\hat{f}^\dagger(\mathbf{x}_1,t)\hat{f}^\dagger(\mathbf{x}_2,t)}{\sqrt{2}}\,|0\rangle\cdot\phi_2(\mathbf{x}_1,\mathbf{x}_2).\end{aligned} \tag{19.100}$$

Here we used Eqs. (19.93) and (19.87). The (anti)commutation relations (19.90) give

$$\hat{f}(\mathbf{x},t)\hat{f}^\dagger(\mathbf{x}_i,t) = \delta^{(3)}(\mathbf{x}-\mathbf{x}_i) \mp \hat{f}^\dagger(\mathbf{x}_i,t)\hat{f}(\mathbf{x},t). \tag{19.101}$$

We use this relation to the right of the operator $\hat{H}_0$ in Eq. (19.100):

$$\begin{aligned}\hat{f}(\mathbf{x},t)\hat{f}^\dagger(\mathbf{x}_1,t)\hat{f}^\dagger(\mathbf{x}_2,t)\,|0\rangle &= \delta^{(3)}(\mathbf{x}-\mathbf{x}_1)\hat{f}^\dagger(\mathbf{x}_2,t)\,|0\rangle \mp \hat{f}^\dagger(\mathbf{x}_1,t)\hat{f}(\mathbf{x},t)\hat{f}^\dagger(\mathbf{x}_2,t)\,|0\rangle\\ &= \left(\delta^{(3)}(\mathbf{x}-\mathbf{x}_1)\hat{f}^\dagger(\mathbf{x}_2,t) \mp \hat{f}^\dagger(\mathbf{x}_1,t)\delta^{(3)}(\mathbf{x}-\mathbf{x}_2)\right)|0\rangle\\ &\quad + (\mp)^2\hat{f}^\dagger(\mathbf{x}_1,t)\hat{f}^\dagger(\mathbf{x}_2,t)\underbrace{\hat{f}(\mathbf{x},t)\,|0\rangle}_{0}.\end{aligned} \tag{19.102}$$

Inserting Eq. (19.102) into Eq. (19.100), integrating over $\mathbf{x}$ and moving the Laplacian from the δ-function onto ϕ_2 by partial integration in $\mathbf{x}_1$ (or $\mathbf{x}_2$), we obtain

$$\begin{aligned}\hat{H}_{0,F}\,|\phi_2,t\rangle &= \int d^3\mathbf{x}d^3\mathbf{x}_1d^3\mathbf{x}_2\hat{f}^\dagger(\mathbf{x},t)\left[-\frac{\hbar^2}{2m}\Delta + V(\mathbf{x})\right]\\ &\quad\cdot\frac{1}{\sqrt{2}}\left\{\delta^{(3)}(\mathbf{x}-\mathbf{x}_1)f^\dagger(\mathbf{x}_2,t) \mp f^\dagger(\mathbf{x}_1,t)\delta^{(3)}(\mathbf{x}-\mathbf{x}_2)\right\}|0\rangle\cdot\phi_2(\mathbf{x}_1,\mathbf{x}_2)\\ &= \int d^3\mathbf{x}_1d^3\mathbf{x}_2\frac{1}{\sqrt{2}}\left[f^\dagger(\mathbf{x}_1,t)\left[-\frac{\hbar^2}{2m}\Delta_1 + V(\mathbf{x}_1)\right]f^\dagger(\mathbf{x}_2,t)\,|0\rangle\cdot\phi_2(\mathbf{x}_1,\mathbf{x}_2)\right.\\ &\quad\left. \mp f^\dagger(\mathbf{x}_2,t)\left[-\frac{\hbar^2}{2m}\Delta_2 + V(\mathbf{x}_2)\right]f^\dagger(\mathbf{x}_1,t)\,|0\rangle\cdot\phi_2(\mathbf{x}_1,\mathbf{x}_2)\right]\\ &= \int d^3\mathbf{x}_1d^3\mathbf{x}_2\frac{f^\dagger(\mathbf{x}_1,t)f^\dagger(\mathbf{x}_2,t)}{\sqrt{2}}\,|0\rangle\left[-\frac{\hbar^2}{2m}\Delta_1 + V(\mathbf{x}_1)\right]\phi_2(\mathbf{x}_1,\mathbf{x}_2)\\ &\quad \mp\int d^3\mathbf{x}_1d^3\mathbf{x}_2\frac{f^\dagger(\mathbf{x}_2,t)f^\dagger(\mathbf{x}_1,t)}{\sqrt{2}}\,|0\rangle\left[-\frac{\hbar^2}{2m}\Delta_2 + V(\mathbf{x}_2)\right]\phi_2(\mathbf{x}_1,\mathbf{x}_2)\\ &= \int d^3\mathbf{x}_1d^3\mathbf{x}_2\frac{f^\dagger(\mathbf{x}_1,t)f^\dagger(\mathbf{x}_2,t)}{\sqrt{2}}\,|0\rangle\sum_{i=1}^{2}\left[-\frac{\hbar^2}{2m}\Delta_i + V(\mathbf{x}_i)\right]\phi_2(\mathbf{x}_1,\mathbf{x}_2).\end{aligned} \tag{19.103}$$

We perform analogous steps for the interaction term,

$$\hat{V}_F |\phi_2, t\rangle = \frac{1}{2} \int d^3\mathbf{x} \int d^3\mathbf{x}' \int d^3\mathbf{x}_1 \int d^3\mathbf{x}_2 \\ \hat{f}^\dagger(\mathbf{x}, t)\hat{f}^\dagger(\mathbf{x}', t)V(\mathbf{x}, \mathbf{x}')\hat{f}(\mathbf{x}', t)\hat{f}(\mathbf{x}, t) \cdot \frac{\hat{f}^\dagger(\mathbf{x}_1, t)\hat{f}^\dagger(\mathbf{x}_2, t)}{\sqrt{2}} |0\rangle \ \phi_2(\mathbf{x}_1, \mathbf{x}_2). \tag{19.104}$$

One can show that

$$\hat{f}(\mathbf{x}', t)\hat{f}(\mathbf{x}, t)\hat{f}^\dagger(\mathbf{x}_1, t)\hat{f}^\dagger(\mathbf{x}_2, t) |0\rangle = \left(\delta^{(3)}(\mathbf{x} - \mathbf{x}_1)\delta^{(3)}(\mathbf{x}' - \mathbf{x}_2) \mp \delta^{(3)}(\mathbf{x}' - \mathbf{x}_1)\delta^{(3)}(\mathbf{x} - \mathbf{x}_2)\right) |0\rangle . \tag{19.105}$$

Inserting Eq. (19.105) into Eq. (19.104), we obtain

$$\hat{V}_F |\phi_2, t\rangle = \int d^3\mathbf{x}_1 \int d^3\mathbf{x}_2 \frac{\hat{f}^\dagger(\mathbf{x}_1, t)\hat{f}^\dagger(\mathbf{x}_2, t)}{\sqrt{2}} |0\rangle \\ \cdot \left[\frac{V(\mathbf{x}_1, \mathbf{x}_2)\phi_2(\mathbf{x}_1, \mathbf{x}_2) + V(\mathbf{x}_2, \mathbf{x}_1)\phi_2(\mathbf{x}_1, \mathbf{x}_2)}{2}\right]. \tag{19.106}$$

For two particles, Eq. (19.95) reads

$$\left(\hat{H}_F - E_2\right) |\phi_2, t\rangle = 0. \tag{19.107}$$

With Eqs. (19.103) and (19.106), this becomes

$$\int d^3\mathbf{x}_1 \int d^3\mathbf{x}_2 \left(\frac{\hat{f}^\dagger(\mathbf{x}_1, t)\hat{f}^\dagger(\mathbf{x}_2, t)}{\sqrt{2}} |0\rangle\right) \left\{\left[-\frac{\hbar^2}{2m}\Delta_1 + V(\mathbf{x}_1)\right] \right. \\ \left. + \left[-\frac{\hbar^2}{2m}\Delta_2 + V(\mathbf{x}_2)\right] + \frac{1}{2}\left[V(\mathbf{x}_1, \mathbf{x}_2) + V(\mathbf{x}_2, \mathbf{x}_1)\right] - E_2\right\} \phi_2(\mathbf{x}_1, \mathbf{x}_2) = 0. \tag{19.108}$$

In the integrand of Eq. (19.108) we recognise the Schrödinger equation of a system of two interacting particles. Note that we need to assume

$$V(\mathbf{x}_1, \mathbf{x}_2) = V(\mathbf{x}_2, \mathbf{x}_1), \tag{19.109}$$

which is the case for identical particles. Since the operators $\hat{f}^\dagger(\mathbf{x}_1, t)$ and $\hat{f}^\dagger(\mathbf{x}_2, t)$ are non-zero, we conclude that Eq. (19.108) holds if (and, conversely, Eq. (19.108) implies it: since $\int \hat{f}^\dagger(\mathbf{x}_1)\hat{f}^\dagger(\mathbf{x}_2) |0\rangle \ g(\mathbf{x}_1, \mathbf{x}_2)$ projects g onto its (anti)symmetric part, and the curly bracket in Eq. (19.108) is symmetric under $1 \leftrightarrow 2$, the function $\{\dots\}\phi_2$ has the same symmetry as ϕ_2 and must itself vanish)

$$\left\{\left[-\frac{\hbar^2}{2m}\Delta_1 + V(\mathbf{x}_1)\right] + \left[-\frac{\hbar^2}{2m}\Delta_2 + V(\mathbf{x}_2)\right] + V(\mathbf{x}_1, \mathbf{x}_2) - E_2\right\} \phi_2(\mathbf{x}_1, \mathbf{x}_2) = 0. \tag{19.110}$$

The generalisation to an arbitrary number of particles follows the same steps. In conclusion, multi-particle systems in non-relativistic quantum mechanics can be described by two equivalent methods:

(i) field quantisation, with field states and the field Hamiltonian of Eq. (19.86), Eq. (19.95);

(ii) the Schrödinger equation with the quantum-mechanical Hamiltonian of Eq. (19.97).

The field-theory method (i) is very compact and mathematically elegant. In addition, it is the only one of the two methods that can be extended to a quantum theory consistent with special relativity. The non-relativistic field theory of this chapter is of course not itself relativistic. The point is that a relativistic quantum theory must allow particle creation and annihilation (particle number is not conserved once $E \gtrsim mc^2$), which a wave function for a fixed number N of particles cannot describe, whereas quantised fields can; together with locality and causality this makes quantum field theory the framework for relativistic quantum mechanics (Weinberg, *The Quantum Theory of Fields* I, Ch. 1 and 5).

19.5 Field Hamiltonian in Momentum Space

Field quantisation offers many advantages over the wave-function description of quantum mechanics. It is often very powerful to work in momentum space. Our starting point is the field Hamiltonian

$$\begin{aligned}\hat{H}_F &= \int d^3\mathbf{x}\, \hat{f}^\dagger(\mathbf{x},t)\left[-\frac{\hbar^2}{2m}\Delta + V_{\text{ext}}(\mathbf{x})\right]\hat{f}(\mathbf{x},t) \\ &\quad + \frac{1}{2}\int d^3\mathbf{x}\int d^3\mathbf{x}'\, \hat{f}^\dagger(\mathbf{x},t)\hat{f}^\dagger(\mathbf{x}',t)V(\mathbf{x},\mathbf{x}')\hat{f}(\mathbf{x}',t)\hat{f}(\mathbf{x},t) \\ &\equiv \hat{H}_{0,F} + \frac{1}{2}\hat{V}_F.\end{aligned} \tag{19.111}$$

From here on, $\hat{V}_F$ denotes the double integral *without* the factor $\frac{1}{2}$, unlike Eq. (19.99); the full Hamiltonian is $\hat{H}_{0,F} + \frac{1}{2}\hat{V}_F$. We assume that there is no external potential,

$$V_{\text{ext}}(\mathbf{x}) = 0, \tag{19.112}$$

and that the particles interact only with each other, through $V(\mathbf{x}_1,\mathbf{x}_2)$. By exchange symmetry and translation and rotation invariance, we can assume that

$$V(\mathbf{x}_1,\mathbf{x}_2) = V(\mathbf{x}_2,\mathbf{x}_1) = V(|\mathbf{x}_1-\mathbf{x}_2|). \tag{19.113}$$

We now Fourier transform the field $\hat{f}(\mathbf{x},t)$,

$$\hat{f}(\mathbf{x},t) = \int_{-\infty}^{+\infty}\frac{d^3\mathbf{k}}{(2\pi)^{3/2}}e^{-i\mathbf{k}\cdot\mathbf{x}}\hat{b}(\mathbf{k},t), \tag{19.114}$$

and the potential $V(\mathbf{x}_1,\mathbf{x}_2)$,

$$V(\mathbf{x}_1,\mathbf{x}_2) = V(\mathbf{x}_1-\mathbf{x}_2) = \int_{-\infty}^{+\infty}\frac{d^3\mathbf{q}}{(2\pi)^{3/2}}e^{-i\mathbf{q}\cdot(\mathbf{x}_1-\mathbf{x}_2)}\bar{V}(q). \tag{19.115}$$

The symmetry

$$V(\mathbf{x}_1-\mathbf{x}_2) = V(\mathbf{x}_2-\mathbf{x}_1) \tag{19.116}$$

of the potential implies the symmetry

$$\bar{V}(\mathbf{q}) = \bar{V}(-\mathbf{q}) \tag{19.117}$$

of its Fourier transform. Inserting Eqs. (19.114) and (19.115) into Eq. (19.111) with $V_{\text{ext}} = 0$, we obtain for the free part

$$\begin{aligned}\hat{H}_{0,F} &= \int d^3\mathbf{x} \left(\int \frac{d^3\mathbf{k}_1}{(2\pi)^{3/2}} e^{i\mathbf{k}_1\cdot\mathbf{x}} \hat{b}^\dagger(\mathbf{k}_1, t) \right) \left[-\frac{\hbar^2}{2m}\Delta \right] \left(\int \frac{d^3\mathbf{k}_2}{(2\pi)^{3/2}} e^{-i\mathbf{k}_2\cdot\mathbf{x}} \hat{b}(\mathbf{k}_2, t) \right) \\ &= \int\int \frac{d^3\mathbf{k}_1\, d^3\mathbf{k}_2}{(2\pi)^3} \frac{\hbar^2 \mathbf{k}_2^{\,2}}{2m} \hat{b}^\dagger(\mathbf{k}_1, t)\hat{b}(\mathbf{k}_2, t) \underbrace{\int d^3\mathbf{x}\, e^{i(\mathbf{k}_1-\mathbf{k}_2)\cdot\mathbf{x}}}_{(2\pi)^3\delta^{(3)}(\mathbf{k}_1-\mathbf{k}_2)} \\ &= \int\int d^3\mathbf{k}_1\, d^3\mathbf{k}_2 \frac{\hbar^2 \mathbf{k}_2^{\,2}}{2m} \hat{b}^\dagger(\mathbf{k}_1, t)\hat{b}(\mathbf{k}_2, t)\delta^{(3)}(\mathbf{k}_1 - \mathbf{k}_2).\end{aligned} \tag{19.118}$$

Thus, we have

$$\boxed{\hat{H}_{0,F} = \int d^3\mathbf{k} \frac{\hbar^2\mathbf{k}^2}{2m} \hat{b}^\dagger(\mathbf{k}, t)\hat{b}(\mathbf{k}, t).} \tag{19.119}$$

Similarly, we derive an expression for the interaction term $\hat{V}_F$:

$$\begin{aligned}\hat{V}_F &= \int \frac{d^3\mathbf{q}}{(2\pi)^{3/2}} \bar{V}(q) \int \frac{d^3\mathbf{x}\, d^3\mathbf{x}'\, d^3\mathbf{k}_1\, d^3\mathbf{k}_2\, d^3\mathbf{k}_3\, d^3\mathbf{k}_4}{(2\pi)^3(2\pi)^3} \\ &\quad \cdot e^{i\mathbf{k}_1\cdot\mathbf{x}} e^{i\mathbf{k}_2\cdot\mathbf{x}'} e^{-i\mathbf{q}\cdot(\mathbf{x}-\mathbf{x}')} e^{-i\mathbf{k}_3\cdot\mathbf{x}} e^{-i\mathbf{k}_4\cdot\mathbf{x}'} \hat{b}^\dagger(\mathbf{k}_1, t)\hat{b}^\dagger(\mathbf{k}_2, t)\hat{b}(\mathbf{k}_4, t)\hat{b}(\mathbf{k}_3, t) \\ &= \int \frac{d^3\mathbf{q}}{(2\pi)^{3/2}} \bar{V}(q) \int d^3\mathbf{k}_1\, d^3\mathbf{k}_2\, d^3\mathbf{k}_3\, d^3\mathbf{k}_4 \left(\int \frac{d^3\mathbf{x}}{(2\pi)^3} e^{i(\mathbf{k}_1-\mathbf{q}-\mathbf{k}_3)\cdot\mathbf{x}} \right) \\ &\quad \cdot \left(\int \frac{d^3\mathbf{x}'}{(2\pi)^3} e^{i(\mathbf{k}_2+\mathbf{q}-\mathbf{k}_4)\cdot\mathbf{x}'} \right) \hat{b}^\dagger(\mathbf{k}_1, t)\hat{b}^\dagger(\mathbf{k}_2, t)\hat{b}(\mathbf{k}_4, t)\hat{b}(\mathbf{k}_3, t) \\ &= \int \frac{d^3\mathbf{q}}{(2\pi)^{3/2}} \bar{V}(q) \int d^3\mathbf{k}_1\, d^3\mathbf{k}_2\, d^3\mathbf{k}_3\, d^3\mathbf{k}_4\, \delta^{(3)}(\mathbf{k}_1 - \mathbf{q} - \mathbf{k}_3)\, \delta^{(3)}(\mathbf{k}_2 + \mathbf{q} - \mathbf{k}_4) \\ &\quad \cdot \hat{b}^\dagger(\mathbf{k}_1, t)\hat{b}^\dagger(\mathbf{k}_2, t)\hat{b}(\mathbf{k}_4, t)\hat{b}(\mathbf{k}_3, t).\end{aligned} \tag{19.120}$$

Thus, we have

$$\boxed{\hat{V}_F = \int \frac{d^3\mathbf{q}}{(2\pi)^{3/2}} \bar{V}(q) \int\int d^3\mathbf{k}_1\, d^3\mathbf{k}_2\, \hat{b}^\dagger(\mathbf{k}_1, t)\hat{b}^\dagger(\mathbf{k}_2, t)\hat{b}(\mathbf{k}_2 + \mathbf{q}, t)\hat{b}(\mathbf{k}_1 - \mathbf{q}, t).} \tag{19.121}$$

We have thus written the field Hamiltonian

$$\hat{H}_F = \hat{H}_{0,F} + \frac{1}{2}\hat{V}_F \tag{19.122}$$

as integrals in momentum space, Eqs. (19.119) and (19.121).

19.5.1 Creation Operators as Fourier Transforms of the Field

The operators $\hat{b}$ and $\hat{b}^\dagger$, which appear in the Fourier representation of the field operators $\hat{f}$ and $\hat{f}^\dagger$, are annihilation and creation operators. From the (anti)commutation relation

$$\left[\hat{f}(\mathbf{x}, t), \hat{f}^\dagger(\mathbf{x}', t)\right]_\pm = \delta^{(3)}(\mathbf{x} - \mathbf{x}'), \tag{19.123}$$

we can derive (anti)commutation relations for $\hat{b}$ and $\hat{b}^\dagger$:

$$\begin{aligned}\delta^{(3)}(\mathbf{x}-\mathbf{x}') &= \left[\int \frac{d^3\mathbf{k}}{(2\pi)^{3/2}} e^{-i\mathbf{k}\cdot\mathbf{x}}\hat{b}(\mathbf{k},t), \int \frac{d^3\mathbf{k}'}{(2\pi)^{3/2}} e^{i\mathbf{k}'\cdot\mathbf{x}'}\hat{b}^\dagger(\mathbf{k}',t)\right]_\pm \\ &= \int \frac{d^3\mathbf{k}\, d^3\mathbf{k}'}{(2\pi)^3} e^{-i(\mathbf{k}\cdot\mathbf{x}-\mathbf{k}'\cdot\mathbf{x}')} \left[\hat{b}(\mathbf{k},t), \hat{b}^\dagger(\mathbf{k}',t)\right]_\pm .\end{aligned} \tag{19.124}$$

This is satisfied if

$$\boxed{\left[\hat{b}(\mathbf{k},t), \hat{b}^\dagger(\mathbf{k}',t)\right]_\pm = \delta^{(3)}(\mathbf{k}-\mathbf{k}').} \tag{19.125}$$

Furthermore, from

$$\left[\hat{f}(\mathbf{x},t), \hat{f}(\mathbf{x}',t)\right]_\pm = \left[\hat{f}^\dagger(\mathbf{x},t), \hat{f}^\dagger(\mathbf{x}',t)\right]_\pm = 0, \tag{19.126}$$

we obtain

$$\boxed{\left[\hat{b}(\mathbf{k},t), \hat{b}(\mathbf{k}',t)\right]_\pm = \left[\hat{b}^\dagger(\mathbf{k},t), \hat{b}^\dagger(\mathbf{k}',t)\right]_\pm = 0.} \tag{19.127}$$

Finally, the equations

$$\hat{f}(\mathbf{x},t)\,|0\rangle = 0, \quad \hat{f}^\dagger(\mathbf{x},t)\,|0\rangle = |\mathbf{x}\rangle \tag{19.128}$$

yield

$$\boxed{\hat{b}(\mathbf{k},t)\,|0\rangle = 0, \quad \hat{b}^\dagger(\mathbf{k},t)\,|0\rangle = |\mathbf{k}\rangle\,.} \tag{19.129}$$

Thus, the operator $\hat{b}^\dagger(\mathbf{k},t)$ acting on the vacuum creates a particle with definite momentum. With the sign convention of Eq. (19.114), $\hat{b}^\dagger(\mathbf{k})\,|0\rangle = \int \frac{d^3\mathbf{x}}{(2\pi)^{3/2}}\, e^{-i\mathbf{k}\cdot\mathbf{x}}\,|\mathbf{x}\rangle$, i.e. $\langle\mathbf{x}|\mathbf{k}\rangle = e^{-i\mathbf{k}\cdot\mathbf{x}}/(2\pi)^{3/2}$: the particle has momentum $-\hbar\mathbf{k}$, energy $\hbar^2\mathbf{k}^2/2m$. The more common convention, matching $\langle\mathbf{x}|\mathbf{p}\rangle \propto e^{+i\mathbf{p}\cdot\mathbf{x}/\hbar}$ and used in quantum field theory courses, is $\hat{f} = \int d^3\mathbf{k}\, e^{+i\mathbf{k}\cdot\mathbf{x}}\,\hat{b}(\mathbf{k})/(2\pi)^{3/2}$, for which $\hat{b}^\dagger(\mathbf{k})$ creates momentum $+\hbar\mathbf{k}$. Since all momenta are flipped together and $\bar{V}(\mathbf{q}) = \bar{V}(-\mathbf{q})$, none of the results below depend on this choice.

Exercise: Derive Eq. (19.129).

19.6 Perturbation (Field) Theory

With the results of the previous section, perturbation theory can be formulated in a simple manner. As in quantum mechanics, perturbation theory requires a small expansion parameter. If the particles interact weakly with each other and most of their energy is kinetic,

$$|V(\mathbf{q})| \ll \frac{\hbar^2\mathbf{k}^2}{2m}, \tag{19.130}$$

we can expand in powers of $\bar{V}(\mathbf{q})$. (Eq. (19.130) is meant qualitatively: $\bar{V}(\mathbf{q})$ has dimension energy×volume. A dimensionally correct criterion is the validity of the Born approximation of Chapter 15, e.g. $m\left|\bar{V}\right|/(\hbar^2 a) \ll 1$ for a potential of range a, up to factors of 2π.) Specifically, we solve the eigenvalue problem of the field Hamiltonian,

$$\hat{H}_F\,|\phi,t\rangle = E_{\text{field}}\,|\phi,t\rangle\,, \tag{19.131}$$

where

$$\hat{H}_F = \hat{H}_{0,F} + \frac{1}{2}\hat{V}_F, \tag{19.132}$$

with

$$\hat{H}_{0,F} = \int d^3\mathbf{k}\, \frac{\hbar^2\mathbf{k}^2}{2m} \hat{b}^\dagger_{\mathbf{k}} \hat{b}_{\mathbf{k}} \tag{19.133}$$

and (after $\mathbf{q} \to -\mathbf{q}$ in Eq. (19.121))

$$\hat{V}_F = \int\int\int d^3\mathbf{k}_1\, d^3\mathbf{k}_2\, \frac{d^3\mathbf{q}}{(2\pi)^{3/2}} \bar{V}(q) \hat{b}^\dagger_{\mathbf{k}_1} \hat{b}^\dagger_{\mathbf{k}_2} \hat{b}_{\mathbf{k}_2-\mathbf{q}}\, \hat{b}_{\mathbf{k}_1+\mathbf{q}}. \tag{19.134}$$

19.6.1 Leading Order in Perturbation Theory

At leading order, we set $\hat{V}_F = 0$ and have

$$\hat{H}_{0,F} \left|\phi^{(0)}, t\right\rangle = E^{(0)}_{\text{field}} \left|\phi^{(0)}, t\right\rangle. \tag{19.135}$$

Inserting Eq. (19.133) into Eq. (19.135), we find

$$\boxed{\int d^3\mathbf{k}\, \frac{\hbar^2\mathbf{k}^2}{2m} \hat{b}^\dagger_{\mathbf{k}} \hat{b}_{\mathbf{k}} \left|\phi^{(0)}, t\right\rangle = E^{(0)}_{\text{field}} \left|\phi^{(0)}, t\right\rangle.} \tag{19.136}$$

The solutions of this equation are states of many free, non-interacting particles with definite momenta. Obviously, the state with zero particles, the vacuum $|0\rangle$, is a solution of Eq. (19.136) with zero energy:

$$\hat{H}_{0,F} |0\rangle = \int d^3\mathbf{k}\, \frac{\hbar^2\mathbf{k}^2}{2m} \hat{b}^\dagger_{\mathbf{k}} \hat{b}_{\mathbf{k}} |0\rangle = 0 \cdot |0\rangle. \tag{19.137}$$

The simplest non-trivial solution is

$$\left|\phi^{(0)}, t\right\rangle = \hat{b}^\dagger_{\mathbf{p}_1} |0\rangle \equiv |\mathbf{p}_1\rangle. \tag{19.138}$$

Indeed,

$$\begin{aligned}
\hat{H}_{0,F} |\mathbf{p}_1\rangle &= \int d^3\mathbf{k}\, \frac{\hbar^2\mathbf{k}^2}{2m} \hat{b}^\dagger_{\mathbf{k}} \hat{b}_{\mathbf{k}} \left(\hat{b}^\dagger_{\mathbf{p}_1} |0\rangle\right) \\
&= \int d^3\mathbf{k}\, \frac{\hbar^2\mathbf{k}^2}{2m} \hat{b}^\dagger_{\mathbf{k}} \left(\underbrace{\left[\hat{b}_{\mathbf{k}}, \hat{b}^\dagger_{\mathbf{p}_1}\right]_\pm}_{\delta^{(3)}(\mathbf{k}-\mathbf{p}_1)} \mp \hat{b}^\dagger_{\mathbf{p}_1} \hat{b}_{\mathbf{k}} \right) |0\rangle \\
&= \frac{\hbar^2\mathbf{p}_1^2}{2m} \left(\hat{b}^\dagger_{\mathbf{p}_1} |0\rangle\right) = \frac{\hbar^2\mathbf{p}_1^2}{2m} |\mathbf{p}_1\rangle.
\end{aligned} \tag{19.139}$$

where we used $\hat{b}_{\mathbf{k}} |0\rangle = 0$. This result is easily extended to multi-particle states:

Exercise: Show that the state

$$|\mathbf{p}_1, \ldots, \mathbf{p}_N\rangle = \hat{b}^\dagger_{\mathbf{p}_1} \ldots \hat{b}^\dagger_{\mathbf{p}_N} |0\rangle$$

is an eigenstate of the free field Hamiltonian:

$$\hat{H}_{0,F} |\mathbf{p}_1, \ldots, \mathbf{p}_N\rangle = E^{(0)} |\mathbf{p}_1, \ldots, \mathbf{p}_N\rangle$$

with energy

$$E^{(0)} = \sum_{i=1}^{N} \frac{\hbar^2\mathbf{p}_i^2}{2m}.$$

(19.140)

19.6.2 Next-to-Leading Order

We now solve the eigenvalue problem at next-to-leading order, taking into account the interaction potential of the identical particles,

$$V(\mathbf{x}_1, \mathbf{x}_2) = \int \frac{d^3\mathbf{q}}{(2\pi)^{3/2}} e^{-i\mathbf{q}\cdot(\mathbf{x}_1-\mathbf{x}_2)} \bar{V}(\mathbf{q}). \tag{19.141}$$

We start by studying the action of the corresponding field operator,

$$\hat{V}_F = \int \frac{d^3\mathbf{q}}{(2\pi)^{3/2}} \bar{V}(q) \int\int d^3\mathbf{k}_1\, d^3\mathbf{k}_2\, \hat{b}^\dagger_{\mathbf{k}_1} \hat{b}^\dagger_{\mathbf{k}_2} \hat{b}_{\mathbf{k}_2-\mathbf{q}}\, \hat{b}_{\mathbf{k}_1+\mathbf{q}}, \tag{19.142}$$

on the eigenstates $|\mathbf{p}_1, \ldots, \mathbf{p}_N\rangle$ of the free Hamiltonian $\hat{H}_{0,F}$.

Exercise: Show that

$$\begin{aligned}
&\text{(i)} && \hat{V}_F\,|\mathbf{p}_1\rangle = && 0 \\
&\text{(ii)} && \langle \mathbf{p}_1|\mathbf{p}_2\rangle = && \delta^{(3)}(\mathbf{p}_1-\mathbf{p}_2) \\
&\text{(iii)} && \langle \mathbf{k}_1, \mathbf{k}_2|\mathbf{p}_1, \mathbf{p}_2\rangle = && \delta^{(3)}(\mathbf{k}_1-\mathbf{p}_1)\delta^{(3)}(\mathbf{k}_2-\mathbf{p}_2) \mp \delta^{(3)}(\mathbf{k}_1-\mathbf{p}_2)\delta^{(3)}(\mathbf{k}_2-\mathbf{p}_1) \\
&\text{(iv)} && \hat{V}_F\,|\mathbf{p}_1, \mathbf{p}_2\rangle = && \int \frac{d^3\mathbf{q}}{(2\pi)^{3/2}} \bar{V}(q) \left[|\mathbf{p}_1-\mathbf{q}, \mathbf{p}_2+\mathbf{q}\rangle \mp |\mathbf{p}_2-\mathbf{q}, \mathbf{p}_1+\mathbf{q}\rangle\right].
\end{aligned} \tag{19.143}$$

Here the upper signs hold for fermions and the lower signs for bosons. We now employ time-dependent perturbation theory. At a time t_0, the field is in the free state $|\text{initial}\rangle$, with

$$\hat{H}_{0,F}\,|\text{initial}\rangle = E^{(0)}_{\text{initial}}\,|\text{initial}\rangle\,. \tag{19.144}$$

We compute the probability of a transition to a different free state $|\text{final}\rangle$ at a time $t > t_0$, with

$$\hat{H}_{0,F}\,|\text{final}\rangle = E^{(0)}_{\text{final}}\,|\text{final}\rangle\,. \tag{19.145}$$

According to Eq. (14.248), the transition amplitude is

$$M_{i\to f} = \langle\text{final}|\text{initial}\rangle - \frac{i}{\hbar}\int_{t_0}^{t} dt'\, e^{i\omega_{fi}(t'-t_0)}\, \langle\text{final}|\frac{\hat{V}_F}{2}|\text{initial}\rangle + O(\hat{V}_F^2), \tag{19.146}$$

where

$$\hbar\omega_{fi} = E_{\text{final}} - E_{\text{initial}}. \tag{19.147}$$

As a concrete example, we consider the transition amplitude for

$$\begin{aligned}
|\text{initial}\rangle &= |\mathbf{p}_1, \mathbf{p}_2\rangle\,, \\
|\text{final}\rangle &= |\mathbf{p}_3, \mathbf{p}_4\rangle\,,
\end{aligned} \tag{19.148}$$

with

$$\{\mathbf{p}_3, \mathbf{p}_4\} \neq \{\mathbf{p}_1, \mathbf{p}_2\} \tag{19.149}$$

(as sets of momenta). Using (iii) of Eq. (19.143), we obtain

$$\langle\text{final}|\text{initial}\rangle = \langle \mathbf{p}_3, \mathbf{p}_4|\mathbf{p}_1, \mathbf{p}_2\rangle = \delta^{(3)}(\mathbf{p}_3-\mathbf{p}_1)\delta^{(3)}(\mathbf{p}_4-\mathbf{p}_2) \mp \delta^{(3)}(\mathbf{p}_3-\mathbf{p}_2)\delta^{(3)}(\mathbf{p}_4-\mathbf{p}_1) = 0. \tag{19.150}$$

We now compute the matrix element in the integral of Eq. (19.146) with (iv) of Eq. (19.143):

$$\langle \text{final}|\frac{\hat{V}_F}{2}|\text{initial}\rangle = \langle \mathbf{p}_3, \mathbf{p}_4| \int \frac{d^3\mathbf{q}}{(2\pi)^{3/2}} \frac{\bar{V}(\mathbf{q})}{2} \left[|\mathbf{p}_1 - \mathbf{q}, \mathbf{p}_2 + \mathbf{q}\rangle \mp |\mathbf{p}_2 - \mathbf{q}, \mathbf{p}_1 + \mathbf{q}\rangle\right]. \quad (19.151)$$

The second term can be rewritten with the (anti)commutation relations of the creation operators,

$$\mp |\mathbf{p}_2 - \mathbf{q}, \mathbf{p}_1 + \mathbf{q}\rangle = \mp \hat{b}^\dagger_{\mathbf{p}_2-\mathbf{q}} \hat{b}^\dagger_{\mathbf{p}_1+\mathbf{q}} |0\rangle = (\mp)^2 \hat{b}^\dagger_{\mathbf{p}_1+\mathbf{q}} \hat{b}^\dagger_{\mathbf{p}_2-\mathbf{q}} |0\rangle = |\mathbf{p}_1 + \mathbf{q}, \mathbf{p}_2 - \mathbf{q}\rangle. \quad (19.152)$$

Inserting this relation into Eq. (19.151), we obtain

$$\langle \text{final}|\frac{\hat{V}_F}{2}|\text{initial}\rangle = \langle \mathbf{p}_3, \mathbf{p}_4| \int \frac{d^3\mathbf{q}}{(2\pi)^{3/2}} \frac{\bar{V}(\mathbf{q})}{2} \left[|\mathbf{p}_1 - \mathbf{q}, \mathbf{p}_2 + \mathbf{q}\rangle + |\mathbf{p}_1 + \mathbf{q}, \mathbf{p}_2 - \mathbf{q}\rangle\right]. \quad (19.153)$$

Substituting $\mathbf{q} \to -\mathbf{q}$ in the second term and using the symmetry

$$\bar{V}(\mathbf{q}) = \bar{V}(-\mathbf{q}), \quad (19.154)$$

we obtain

$$\langle \text{final}|\frac{\hat{V}_F}{2}|\text{initial}\rangle = \langle \mathbf{p}_3, \mathbf{p}_4| \int \frac{d^3\mathbf{q}}{(2\pi)^{3/2}} \left[\frac{\bar{V}(\mathbf{q}) + \bar{V}(-\mathbf{q})}{2}\right] \left[|\mathbf{p}_1 - \mathbf{q}, \mathbf{p}_2 + \mathbf{q}\rangle\right]. \quad (19.155)$$

Thus,

$$\boxed{\langle \mathbf{p}_3, \mathbf{p}_4|\frac{\hat{V}_F}{2}|\mathbf{p}_1, \mathbf{p}_2\rangle = \int \frac{d^3\mathbf{q}}{(2\pi)^{3/2}} \bar{V}(\mathbf{q}) \langle \mathbf{p}_3, \mathbf{p}_4|\mathbf{p}_1 - \mathbf{q}, \mathbf{p}_2 + \mathbf{q}\rangle.} \quad (19.156)$$

Using (iii) of Eq. (19.143), we find

$$\begin{aligned}\langle \mathbf{p}_3, \mathbf{p}_4|\frac{\hat{V}_F}{2}|\mathbf{p}_1, \mathbf{p}_2\rangle &= \int \frac{d^3\mathbf{q}}{(2\pi)^{3/2}} \bar{V}(\mathbf{q}) \Big[\delta^{(3)}(\mathbf{p}_3 - \mathbf{p}_1 + \mathbf{q})\delta^{(3)}(\mathbf{p}_4 - \mathbf{p}_2 - \mathbf{q}) \\ &\quad \mp \delta^{(3)}(\mathbf{p}_3 - \mathbf{p}_2 - \mathbf{q})\delta^{(3)}(\mathbf{p}_4 - \mathbf{p}_1 + \mathbf{q})\Big] \\ &= \frac{1}{(2\pi)^{3/2}} \left[\bar{V}(\mathbf{p}_3 - \mathbf{p}_1) \mp \bar{V}(\mathbf{p}_3 - \mathbf{p}_2)\right] \delta^{(3)}(\mathbf{p}_3 + \mathbf{p}_4 - \mathbf{p}_1 - \mathbf{p}_2).\end{aligned} \quad (19.157)$$

The transition amplitude becomes

$$\boxed{M_{fi} = -\frac{i}{\hbar} \int_{t_0}^{t} dt' \, e^{i\omega_{fi}(t'-t_0)} \delta^{(3)}(\mathbf{p}_3 + \mathbf{p}_4 - \mathbf{p}_1 - \mathbf{p}_2) \frac{1}{(2\pi)^{3/2}} \left[\bar{V}(\mathbf{p}_3 - \mathbf{p}_1) \mp \bar{V}(\mathbf{p}_3 - \mathbf{p}_2)\right].} \quad (19.158)$$

We observe that:

- the total momentum of the particles is conserved, because of the δ-function;
- the two terms interfere destructively for fermions and constructively for bosons. (The two terms are the direct amplitude, momentum transfer $\mathbf{p}_3 - \mathbf{p}_1$, and the exchange amplitude, momentum transfer $\mathbf{p}_3 - \mathbf{p}_2 = \mathbf{p}_1 - \mathbf{p}_4$; up to the factors of (2π) this is the Born amplitude for identical particles, cf. the symmetrised scattering amplitude $f(\theta) \pm f(\pi - \theta)$ of Chapter 17.)

19.6.3 Transition Rate

From Eq. (19.158) we obtain the transition probability

$$P_{i\to f} \equiv \left|M_{i\to f}\right|^2 = \frac{1}{\hbar^2}\frac{\left|\bar{V}(\mathbf{p}_3-\mathbf{p}_1)\mp\bar{V}(\mathbf{p}_3-\mathbf{p}_2)\right|^2}{(2\pi)^3}\delta^{(3)}(\mathbf{p}_3+\mathbf{p}_4-\mathbf{p}_1-\mathbf{p}_2)^2\left|\int_{t_0}^{t}dt'\,e^{i\omega_{fi}(t'-t_0)}\right|^2. \tag{19.159}$$

The time integral is familiar from Eq. (14.253) and leads to Fermi's golden rule. Repeating the steps of Secs. 14.5.3 and 14.5.4, we obtain the transition rate

$$\Gamma_{i\to f} \equiv \frac{dP_{i\to f}}{dt} = \frac{2\pi}{\hbar}\delta(E_{\text{final}}-E_{\text{initial}})\frac{\left|\bar{V}(\mathbf{p}_3-\mathbf{p}_1)\mp\bar{V}(\mathbf{p}_3-\mathbf{p}_2)\right|^2}{(2\pi)^3}\delta^{(3)}(\mathbf{p}_3+\mathbf{p}_4-\mathbf{p}_1-\mathbf{p}_2)^2, \tag{19.160}$$

with

$$\begin{aligned} E_{\text{final}} &= E_3+E_4 = \frac{\hbar^2\mathbf{p}_3^2}{2m}+\frac{\hbar^2\mathbf{p}_4^2}{2m}, \\ E_{\text{initial}} &= E_1+E_2 = \frac{\hbar^2\mathbf{p}_1^2}{2m}+\frac{\hbar^2\mathbf{p}_2^2}{2m}. \end{aligned} \tag{19.161}$$

Note the square of the δ-function in Eq. (19.160). We can rewrite it as

$$\delta^{(3)}(\mathbf{p}_3+\mathbf{p}_4-\mathbf{p}_1-\mathbf{p}_2)^2 = \delta^{(3)}(\mathbf{p}_3+\mathbf{p}_4-\mathbf{p}_1-\mathbf{p}_2)\delta^{(3)}(0). \tag{19.162}$$

We thus find an explicit infinity, $\delta^{(3)}(0)=\infty$, which seems to indicate a physically meaningless result. While $\delta^{(3)}(\mathbf{p}_3+\mathbf{p}_4-\mathbf{p}_1-\mathbf{p}_2)$ gives a finite result when we integrate over a continuum of final states $|\mathbf{p}_3,\mathbf{p}_4\rangle$ with some density $\rho(\mathbf{p}_3,\mathbf{p}_4)$, the factor $\delta^{(3)}(0)$ cannot be removed by integration.

This infinity is due to the infinite volume in which the field $\hat{f}(\mathbf{x},t)$ is defined. Physically, a field description of a scattering process includes all possible scatterings, which may occur anywhere in space. Intuitively, if the probability of a scattering in a small region around a point is finite, then the probability of a scattering anywhere in an infinite space is infinite. In other words, the momentum eigenstates describe particles spread over all of space, and pairs of particles can scatter everywhere.

What we are interested in is the probability of a scattering in the region on which we focus our observation, so we should divide out the volume factor. The δ-function has the integral representation

$$\delta^{(3)}(\mathbf{k}) = \int_{-\infty}^{\infty}\frac{d^3\mathbf{x}}{(2\pi)^3}e^{i\mathbf{k}\cdot\mathbf{x}}. \tag{19.163}$$

Setting $\mathbf{k}=0$, we find

$$\delta^{(3)}(0) = \int_{-\infty}^{\infty}\frac{d^3\mathbf{x}}{(2\pi)^3}e^{i0} = \frac{\text{Volume}}{(2\pi)^3}. \tag{19.164}$$

Thus, instead of the transition rate of Eq. (19.160), we consider the finite transition rate per unit volume,

$$w_{i\to f} \equiv \frac{\Gamma_{i\to f}}{\text{Volume}} = \frac{1}{\text{Volume}}\frac{dP_{i\to f}}{dt}. \tag{19.165}$$

Inserting Eqs. (19.160)–(19.164) into Eq. (19.165), we obtain

$$w_{i\to f} = \frac{2\pi}{\hbar}\frac{\left|\bar{V}(\mathbf{p}_3-\mathbf{p}_1) \mp \bar{V}(\mathbf{p}_3-\mathbf{p}_2)\right|^2}{(2\pi)^6}\delta^{(3)}(\mathbf{p}_3+\mathbf{p}_4-\mathbf{p}_1-\mathbf{p}_2) \cdot \delta\left(\frac{\hbar^2}{2m}(\mathbf{p}_3^2+\mathbf{p}_4^2-\mathbf{p}_1^2-\mathbf{p}_2^2)\right). \tag{19.166}$$

A

Spherical Potentials and Spherical Bessel Functions

Consider a spherically symmetric potential,

$$V(r) = \begin{cases} -V_0, & r < a, \\ 0, & r > a, \end{cases} \quad V_0 > 0. \tag{A.1}$$

Because of the spherical symmetry, we solve the Schrödinger equation in spherical coordinates. We first consider "scattering solutions" with positive energy $E > 0$. They have the form

$$\psi(r, \theta, \phi) = R(r) Y_\ell^m(\theta, \phi). \tag{A.2}$$

For the radial part, the Schrödinger equation yields

$$\left[\frac{d^2}{dr^2} + \frac{2}{r} \frac{d}{dr} - \frac{\ell(\ell+1)}{r^2} + k^2 \right] R(r) = 0, \tag{A.3}$$

with

$$k^2 = \begin{cases} \frac{2m(E+V_0)}{\hbar^2}, & r < a \\ \frac{2mE}{\hbar^2}, & r > a. \end{cases} \tag{A.4}$$

We first focus on the region $r < a$. We rescale $r = \rho/k$, and Eq. (A.3) becomes

$$\left[\frac{d^2}{d\rho^2} + \frac{2}{\rho} \frac{d}{d\rho} - \frac{\ell(\ell+1)}{\rho^2} + 1 \right] R(\rho) = 0. \tag{A.5}$$

To solve this equation, we write the solution in the form

$$R_\ell(\rho) = \rho^\ell X_\ell(\rho). \tag{A.6}$$

Eq. (A.5) then becomes

$$\boxed{X_\ell'' + \frac{2(\ell+1)}{\rho} X_\ell' + X_\ell = 0.} \tag{A.7}$$

For $\ell = 0$,

$$X_0'' + \frac{2}{\rho} X_0' + X_0 = 0, \tag{A.8}$$

we find two solutions,

$$X_0 = \frac{\sin\rho}{\rho}, \tag{A.9}$$

which is regular at $\rho = 0$, and

$$X_0 = -\frac{\cos\rho}{\rho}, \tag{A.10}$$

which is singular at $\rho = 0$. We could have multiplied the solutions (A.9) and (A.10) by any constant, since the differential equation is linear. In Eq. (A.9) we chose the factor 1, which gives the zeroth spherical Bessel function. In Eq. (A.10) we chose a minus sign, which gives the zeroth spherical Neumann function, conventionally defined with a minus sign. In general, any linear combination of Eqs. (A.9) and (A.10) solves Eq. (A.8).

Suppose that $X_\ell(\rho)$ solves Eq. (A.7) for a given ℓ. Then one can show that the function

$$X_{\ell+1} = \frac{1}{\rho}\frac{d}{d\rho}X_\ell \tag{A.11}$$

solves the differential equation with ℓ replaced by $\ell + 1$.

Exercise: Verify this statement.

We thus know how to build the solutions for any integer ℓ from the solution for $\ell = 0$:

$$X_\ell = \left(\frac{1}{\rho}\frac{d}{d\rho}\right)^\ell X_0. \tag{A.12}$$

Spherical Bessel functions: The regular solutions are the spherical Bessel functions, defined as

$$J_\ell(\rho) = (-\rho)^\ell \left(\frac{1}{\rho}\frac{d}{d\rho}\right)^\ell \left(\frac{\sin\rho}{\rho}\right). \tag{A.13}$$

Explicitly,

$$J_0(\rho) = \frac{\sin\rho}{\rho}, \quad J_1(\rho) = -\rho\frac{1}{\rho}\frac{d}{d\rho}\frac{\sin\rho}{\rho} = \frac{\sin\rho - \rho\cos\rho}{\rho^2}, \quad \text{etc.} \tag{A.14}$$

Their behaviour for $\rho \to 0$ is

$$\begin{aligned} J_0(\rho) &\overset{\rho\to 0}{=} 1 + O(\rho^2) \\ J_1(\rho) &\overset{\rho\to 0}{=} \left.\frac{\sin\rho - \rho\cos\rho}{\rho^2}\right|_{\rho\to 0} = \frac{\rho - \frac{1}{6}\rho^3 - \rho + \frac{1}{2}\rho^3 + O(\rho^5)}{\rho^2} = \frac{\rho}{3} + O(\rho^3), \end{aligned} \tag{A.15}$$

and for $\rho \to \infty$ it is

$$\begin{aligned} J_0(\rho) &\overset{\rho\to\infty}{\sim} \frac{\sin\rho}{\rho} \\ J_1(\rho) &\overset{\rho\to\infty}{\sim} \left.\frac{\sin\rho}{\rho^2} - \frac{\cos\rho}{\rho}\right|_{\rho\to\infty} = -\frac{\cos\rho}{\rho} = \frac{1}{\rho}\sin\left(\rho - \frac{\pi}{2}\right). \end{aligned} \tag{A.16}$$

Exercise: Show that

$$J_\ell(\rho) \overset{\rho\to 0}{\sim} \frac{\rho^\ell}{1\cdot 3\cdot 5\cdot \ldots \cdot (2\ell+1)}$$
$$\text{and } J_\ell(\rho) \overset{\rho\to\infty}{\sim} \frac{1}{\rho}\sin\left(\rho - \frac{\pi\ell}{2}\right). \tag{A.17}$$

Exercise: Show that

$$\frac{\rho^\ell}{1\cdot 3\cdot 5\cdot \ldots \cdot (2\ell+1)} = \frac{\rho^\ell 2^\ell \ell!}{(2\ell+1)!}. \tag{A.18}$$

Spherical Neumann functions: Analogously, the singular solutions (more precisely, the functions $(-\rho)^\ell X_\ell$, which solve Eq. (A.5)) are the spherical Neumann functions, defined as

$$y_\ell(\rho) = -(-\rho)^\ell \left(\frac{1}{\rho}\frac{d}{d\rho}\right)^\ell \left(\frac{\cos\rho}{\rho}\right). \tag{A.19}$$

Explicitly,

$$y_0(\rho) = -\frac{\cos\rho}{\rho}, \quad y_1(\rho) = \rho\frac{1}{\rho}\frac{d}{d\rho}\frac{\cos\rho}{\rho} = \frac{-\cos\rho - \rho\sin\rho}{\rho^2}, \quad \text{etc.} \tag{A.20}$$

Their behaviour for $\rho \to 0$ is

$$\begin{aligned} y_0(\rho) &\overset{\rho\to 0}{=} -\frac{1}{\rho} + O(\rho) \\ y_1(\rho) &\overset{\rho\to 0}{\sim} \left.\frac{-\cos\rho - \rho\sin\rho}{\rho^2}\right|_{\rho\to 0} = \frac{-1 + \frac{1}{2}\rho^2 - \rho^2 + O(\rho^4)}{\rho^2} = -\frac{1}{\rho^2} + O(\rho^0), \end{aligned} \tag{A.21}$$

and for $\rho \to \infty$ it is

$$\begin{aligned} y_0(\rho) &\overset{\rho\to\infty}{\sim} -\frac{\cos\rho}{\rho} \\ y_1(\rho) &\overset{\rho\to\infty}{\sim} \left.-\frac{\cos\rho}{\rho^2} - \frac{\sin\rho}{\rho}\right|_{\rho\to\infty} = -\frac{\sin\rho}{\rho} = -\frac{1}{\rho}\cos\left(\rho - \frac{\pi}{2}\right). \end{aligned} \tag{A.22}$$

A.1 Spherical Hankel Functions, Bound States and Phase Shifts

Exercise: Prove that for $\rho \to \infty$

$$J_\ell(\rho) \simeq \frac{1}{\rho}\sin\left(\rho - \frac{\pi\ell}{2}\right) \quad \text{and} \quad y_\ell(\rho) \simeq -\frac{1}{\rho}\cos\left(\rho - \frac{\pi\ell}{2}\right). \tag{A.23}$$

We also define the spherical Hankel functions

$$h_\ell^{(1)}(\rho) = J_\ell(\rho) + i\, y_\ell(\rho) \quad \text{and} \quad h_\ell^{(2)}(\rho) = h_\ell^{(1)}(\rho)^*. \tag{A.24}$$

(The complex conjugation refers to real ρ.) Asymptotically, for $\rho \to \infty$,

$$h_\ell^{(1)}(\rho) \simeq -\frac{i}{\rho}\, e^{i(\rho - \frac{\ell\pi}{2})}. \tag{A.25}$$

For example, $h_0^{(1)}(\rho) = -i\, e^{i\rho}/\rho$: $h_\ell^{(1)}$ is an outgoing spherical wave, and for imaginary argument $\rho = i\kappa r$ it decays as $e^{-\kappa r}$, which is what we need for bound states.

Bound states

We now consider bound states, with energy

$$-V_0 < E < 0. \tag{A.26}$$

For $r < a$ the radial equation reads

$$\left[\frac{d^2}{dr^2} + \frac{2}{r}\frac{d}{dr} - \frac{\ell(\ell+1)}{r^2} + q^2\right] R(r) = 0, \qquad q^2 = \frac{2m(V_0+E)}{\hbar^2}, \tag{A.27}$$

and for $r > a$

$$\left[\frac{d^2}{dr^2} + \frac{2}{r}\frac{d}{dr} - \frac{\ell(\ell+1)}{r^2} - \kappa^2\right] R(r) = 0, \qquad \kappa^2 = \frac{2m(-E)}{\hbar^2}. \tag{A.28}$$

The solution that is regular at the centre of the potential is

$$R(r) = A\, j_\ell(qr), \qquad 0 \le r \le a, \tag{A.29}$$

and outside the well

$$R(r) = B\, h_\ell^{(1)}(i\kappa r), \qquad r > a \tag{A.30}$$

which vanishes for $r \to \infty$. At $r = a$, the wave function and its derivative must be continuous:

$$A\, j_\ell(qa) = B\, h_\ell^{(1)}(i\kappa a) \tag{A}$$

$$A\, q\, j_\ell'(qa) = B\, i\kappa\, h_\ell^{(1)\prime}(i\kappa a) \tag{B}$$

$$\frac{\text{(B)}}{\text{(A)}} \;\Rightarrow\; q\,\left.\frac{d\ln j_\ell}{d\rho}\right|_{\rho=qa} = i\kappa\,\left.\frac{d\ln h_\ell^{(1)}}{d\rho}\right|_{\rho=i\kappa a}. \tag{A.31}$$

For example, for $\ell = 0$ we have

$$R(r) = \begin{cases} A\,\dfrac{\sin(qr)}{r}, & r < a \\[2mm] B\,\dfrac{e^{-\kappa r}}{r}, & r > a. \end{cases} \tag{A.32}$$

The continuity conditions give

$$\boxed{\tan(qa) = -\frac{\hbar q}{\sqrt{2m\,|E|}} = -\frac{q}{\kappa}.} \tag{A.33}$$

With $\xi = qa$, $\eta = \kappa a$ it reads $\eta = -\xi\cot\xi$, to be solved together with $\xi^2 + \eta^2 = 2mV_0a^2/\hbar^2$; a solution with $\eta > 0$ needs $\xi > \pi/2$, so the three-dimensional well has an s-wave bound state only if $V_0 > \pi^2\hbar^2/(8ma^2)$ (unlike the one-dimensional well, which always binds). These are the odd states of the one-dimensional well, since $u = rR$ must vanish at $r = 0$. See Gottfried–Yan, Ch. 8; Sakurai, Modern QM, Ch. 7 (Ch. 6 in the 2nd ed.).

Scattering solution

For $E > 0$, we have

$$R_\ell(r) = \begin{cases} A\, j_\ell(qr), & r < a \\ B\, j_\ell(kr) + C\, y_\ell(kr), & r > a, \end{cases} \tag{A.34}$$

with

$$k = \frac{\sqrt{2mE}}{\hbar} \quad \text{and} \quad q = \frac{\sqrt{2m(E+V_0)}}{\hbar}.$$

The continuity conditions at $r = a$ give

$$q\, \frac{dj_\ell/d\rho}{j_\ell}\bigg|_{\rho=qa} = k \left[\frac{B\, \frac{dj_\ell}{d\rho} + C\, \frac{dy_\ell}{d\rho}}{B\, j_\ell + C\, y_\ell} \right]_{\rho=ka} \quad \rightsquigarrow \quad \frac{C}{B} = \dots \tag{A.35}$$

Explicitly, with $\gamma_\ell = q\, j_\ell'(qa)/j_\ell(qa)$: $\frac{C}{B} = -\frac{k\, j_\ell'(ka) - \gamma_\ell\, j_\ell(ka)}{k\, y_\ell'(ka) - \gamma_\ell\, y_\ell(ka)}$.

Asymptotically, for $r \to \infty$,

$$R_\ell(r) \simeq \frac{B}{kr} \left[\sin\left(kr - \frac{\ell\pi}{2}\right) - \frac{C}{B} \cos\left(kr - \frac{\ell\pi}{2}\right) \right], \tag{A.36}$$

or

$$R_\ell(r) \simeq \frac{B}{\cos\delta_\ell(k)}\, \frac{1}{kr}\, \sin\left(kr - \frac{\ell\pi}{2} + \delta_\ell(k)\right), \tag{A.37}$$

with

$$\boxed{\tan\delta_\ell(k) = -\frac{C}{B}.} \tag{A.38}$$

Eqs. (A.36)–(A.38) follow from Eq. (A.23) and $\sin(x+\delta)/\cos\delta = \sin x + \tan\delta\cos x$. Thus $\tan\delta_\ell = \frac{k\, j_\ell'(ka) - \gamma_\ell\, j_\ell(ka)}{k\, y_\ell'(ka) - \gamma_\ell\, y_\ell(ka)}$ (Sakurai, Modern QM, Ch. 7 (Ch. 6 in the 2nd ed.)); for $\ell = 0$ this reduces to $\tan(ka + \delta_0) = \frac{k}{q}\tan(qa)$. The phase δ_ℓ is the one of Eq. (15.103); for weak V_0 it agrees with the Born phase shift, Eq. (15.128).

B

Explicit Calculation of the Green's Function for Scattering

Our goal is to find an explicit expression for the Green's function in Eq. (15.33),

$$G_\delta(\mathbf{x}-\mathbf{x}') = \int_{-\infty}^{\infty} \frac{\mathrm{d}^3\mathbf{k}}{(2\pi\hbar)^3} \frac{e^{\frac{i}{\hbar}\mathbf{k}\cdot(\mathbf{x}-\mathbf{x}')}}{E - \frac{\mathbf{k}^2}{2M} + i\delta}. \tag{B.1}$$

We introduce spherical coordinates in momentum space,

$$d^3\mathbf{k} = k^2 dk\, d\cos\theta\, d\phi, \quad k \equiv |\mathbf{k}|\,. \tag{B.2}$$

The angular integrations can be carried out immediately:

$$\begin{aligned} G_\delta(\mathbf{x}-\mathbf{x}') &= -\frac{4\pi i\hbar}{(2\pi\hbar)^3}\frac{2M}{|\mathbf{x}-\mathbf{x}'|}\int_0^\infty \frac{dk\,k}{2}\frac{e^{\frac{ik|\mathbf{X}-\mathbf{X}'|}{\hbar}} - e^{-\frac{ik|\mathbf{X}-\mathbf{X}'|}{\hbar}}}{2ME - |k|^2 + i\delta} \\ &= -\frac{2\pi i\hbar}{(2\pi\hbar)^3}\frac{2M}{|\mathbf{x}-\mathbf{x}'|}\int_{-\infty}^\infty dk k \frac{e^{\frac{ik|\mathbf{X}-\mathbf{X}'|}{\hbar}}}{2ME - k^2 + i\delta} \\ &= \frac{i}{(2\pi\hbar)^2}\frac{2M}{|\mathbf{x}-\mathbf{x}'|}\int_{-\infty}^\infty dk k \frac{e^{\frac{ik|\mathbf{X}-\mathbf{X}'|}{\hbar}}}{k^2 - (2ME + i\delta)}. \end{aligned} \tag{B.3}$$

Here δ stands for $2M\delta$ (still an infinitesimal positive number), and the second line uses that the integrand of the first line is even in k, so that the $e^{-ik|\mathbf{X}-\mathbf{X}'|/\hbar}$ term equals the $k<0$ part of the $e^{+ik|\mathbf{X}-\mathbf{X}'|/\hbar}$ term. The integral can be evaluated with the residue theorem. The integrand has two poles, at

$$k = \pm\sqrt{2ME + i\delta}. \tag{B.4}$$

The integral is therefore given by the residue at the pole with positive imaginary part. (Since $|\mathbf{x}-\mathbf{x}'|>0$, $e^{ik|\mathbf{X}-\mathbf{X}'|/\hbar}$ decays for $\mathrm{Im}\,k>0$; by Jordan's lemma the contour can be closed by a large semicircle in the upper half plane, which encloses only $k_0 = +\sqrt{2ME+i\delta}$.) We find

$$\begin{aligned} G_\delta(\mathbf{x}-\mathbf{x}') &= \frac{i}{(2\pi\hbar)^2}\frac{2M}{|\mathbf{x}-\mathbf{x}'|}\cdot 2\pi i\,\mathrm{Res}\left(f(k), \sqrt{2ME+i\delta}\right), \\ &\text{with} \quad f(k) = \frac{k\,e^{\frac{ik|\mathbf{X}-\mathbf{X}'|}{\hbar}}}{k^2 - (2ME+i\delta)}. \end{aligned} \tag{B.5}$$

Evaluating the residue yields

$$\begin{aligned}\operatorname{Res}\left(f(k), \sqrt{2ME+i\delta}\right) &= \lim_{k\to\sqrt{2ME+i\delta}} (k-\sqrt{2ME+i\delta}) f(k) \\ &= + \lim_{k\to\sqrt{2ME+i\delta}} \frac{k e^{\frac{ik|\mathbf{x}-\mathbf{x}'|}{\hbar}}}{k+\sqrt{2ME+i\delta}} \\ &= + \frac{k e^{\frac{i\sqrt{2ME+i\delta}|\mathbf{x}-\mathbf{x}'|}{\hbar}}}{2k} = \frac{e^{\frac{i\sqrt{2ME+i\delta}|\mathbf{x}-\mathbf{x}'|}{\hbar}}}{2}.\end{aligned} \tag{B.6}$$

Inserting Eq. (B.6) into Eq. (B.5), we obtain

$$G_\delta(\mathbf{x}-\mathbf{x}') = -\frac{1}{2\pi\hbar^2}\frac{2M}{|\mathbf{x}-\mathbf{x}'|}\frac{e^{\frac{i\sqrt{2ME+i\delta}|\mathbf{x}-\mathbf{x}'|}{\hbar}}}{2}. \tag{B.7}$$

Performing the limit $\delta \to 0$, we finally have

$$G(\mathbf{x}-\mathbf{x}') = -\frac{2M}{\hbar^2}\frac{1}{4\pi}\frac{e^{\frac{i}{\hbar}\sqrt{2ME}|\mathbf{x}-\mathbf{x}'|}}{|\mathbf{x}-\mathbf{x}'|}. \tag{B.8}$$

C

An Alternative Derivation of the Lippmann–Schwinger Equation

In this appendix we present an alternative derivation of the Lippmann–Schwinger equation. It is not part of the lectures, but some readers may find this way of thinking helpful for understanding the result.

In what follows, ψ denotes the wave function, and all operators act in position space. Our goal is to solve the eigenvalue problem $(H_0 + V)\psi = E\psi$, where $H_0 = -\frac{\hbar^2}{2M}\Delta$ and V is an arbitrary potential. Since this problem cannot be solved explicitly in general, we reformulate it in an equivalent form in which our knowledge of H_0 can be used to approximate the solutions of the original problem.

We first state the main result in its most general form, then prove it, and finally apply it to arrive at Eq. (15.37).

Theorem (Lippmann–Schwinger). Let $\psi, V : \mathbb{R}^n \to \mathbb{C}$ be sufficiently regular, $E \in \mathbb{R}$, let $(E - H_0)$ be a linear differential operator with constant coefficients[1], f_E its fundamental solution[2] and $h_E \in \ker(E - H_0)$. Then

$$\boxed{\psi = h_E + f_E * (V\psi) \quad \Rightarrow \quad (H_0 + V)\psi = E\psi,} \tag{C.1}$$

where $V\psi$ is the pointwise product of V and ψ, and $*$ denotes the convolution. The left-hand side of the implication already has the general form of Eq. (15.37).

Proof. We apply $(E - H_0)$ to both sides of the left-hand side of the implication:

$$\begin{aligned}[E - H_0]\psi &= [E - H_0]\left(h_E + f_E * (V\psi)\right) = [E - H_0]\left(f_E * (V\psi)\right) \\ &= \left([E - H_0]f_E\right) * (V\psi) = \delta * (V\psi) = V\psi. \qquad \square\end{aligned}$$

Our only task is therefore to find the fundamental solution of $[E - H_0] = E + \frac{\hbar^2}{2M}\Delta$, that is, to solve

$$\left[E + \frac{\hbar^2}{2M}\Delta\right] f_E = \delta \quad \Longleftrightarrow \quad \left[\frac{2ME}{\hbar^2} + \Delta\right] f_E = \frac{2M}{\hbar^2}\delta. \tag{C.2}$$

[1] That is, $(E - H_0) = \sum_{|\alpha| \le N} a_\alpha \partial^\alpha$.
[2] That is, $(E - H_0) f_E = \delta$.

This can be done by taking the Fourier transform of both sides and evaluating the resulting integral with the residue theorem, as in Appendix B. Here we present an alternative method.

First, we notice that if g_α solves $(\Delta + \alpha^2)g_\alpha = \delta$, then $f_E = \frac{2M}{\hbar^2} g_\alpha$ with $\alpha = \frac{\sqrt{2ME}}{\hbar}$ solves Eq. (C.2). This removes the constants from the calculation; we reinstate them at the end. We thus need the fundamental solution of the Helmholtz operator $(\Delta + \alpha^2)$ in three dimensions,

$$(\Delta + \alpha^2)g_\alpha = \delta. \tag{C.3}$$

Using the spherical symmetry of the problem, $g_\alpha(\mathbf{x}) = g_\alpha(r)$ with $r = |\mathbf{x}|$, we have for $\mathbf{x} \neq 0$

$$(\Delta + \alpha^2)g_\alpha = \frac{1}{r}\partial_r^2(r g_\alpha) + \alpha^2 g_\alpha = 0, \tag{C.4}$$

or, equivalently,

$$\partial_r^2(r g_\alpha) + \alpha^2(r g_\alpha) = 0 \quad \Rightarrow \quad g_\alpha(\mathbf{x}) = C\frac{e^{\pm i\alpha|\mathbf{x}|}}{|\mathbf{x}|}. \tag{C.5}$$

The delta function on the right-hand side of Eq. (C.3) fixes the constant C through the condition $\lim_{\varepsilon\to 0}\int_{B_\varepsilon(0)} d^3\mathbf{x}\,\Delta g_\alpha = 1$. In this way we "measure" the discontinuity of the first derivative of g_α at $\mathbf{x} = 0$. With the divergence theorem we find

$$\lim_{\varepsilon\to 0}\int_{B_\varepsilon(0)} d^3\mathbf{x}\,\Delta g_\alpha = \lim_{\varepsilon\to 0}(\partial_r g_\alpha)\int_{\partial B_\varepsilon(0)} dA = -4\pi C, \tag{C.6}$$

and thus $C = -\frac{1}{4\pi}$. Putting everything together, we obtain

$$g_\alpha = -\frac{e^{\pm i\alpha|\mathbf{x}|}}{4\pi\,|\mathbf{x}|} \quad \Rightarrow \quad f_E = -\frac{2M}{\hbar^2}\frac{e^{\pm i\frac{\sqrt{2ME}}{\hbar}|\mathbf{x}|}}{4\pi\,|\mathbf{x}|}. \tag{C.7}$$

We now apply the theorem with $h_E = \frac{e^{\frac{i}{\hbar}\mathbf{k}\cdot\mathbf{x}}}{\sqrt{2\pi\hbar}^3} \in \ker(E - H_0)$, where $|\mathbf{k}| = \sqrt{2ME}$, and write out the convolution:

$$\boxed{\psi(\mathbf{x}) = \frac{e^{\frac{i}{\hbar}\mathbf{k}\cdot\mathbf{x}}}{\sqrt{2\pi\hbar}^3} - \frac{2M}{\hbar^2}\int\frac{d^3\mathbf{y}}{4\pi}\frac{e^{\pm\frac{i}{\hbar}|\mathbf{k}||\mathbf{x}-\mathbf{y}|}}{|\mathbf{x}-\mathbf{y}|}(V\psi)(\mathbf{y}).} \tag{C.8}$$

With the upper sign, this is Eq. (15.37). The theorem itself does not fix the sign: both $e^{\pm i\alpha|\mathbf{x}|}/|\mathbf{x}|$ are fundamental solutions (as is any combination of them). The outgoing choice, the upper sign, is a boundary condition; it is the one encoded by the $+i\delta$ prescription of Sec. 15.2 (a free particle at $t \to -\infty$). The lower sign gives the incoming-wave solution. Note also that the theorem only proves the implication "⇒", which is all that is needed here.

D

Functional Derivatives

The derivative of a functional $G\left[f_i\right]$ of a set of functions f_i is defined by

$$\frac{\delta G[f_j]}{\delta f_i(\mathbf{x})} = \lim_{\varepsilon \to 0} \left(\frac{G[f_j(\mathbf{y}) + \varepsilon \delta_{ij} \delta(\mathbf{y} - \mathbf{x})] - G[f_j(\mathbf{y})]}{\varepsilon} \right). \tag{D.1}$$

For example,

$$\begin{aligned} \frac{\delta \int d^3\mathbf{x}\, f_p(\mathbf{x})}{\delta f_i(\mathbf{x})} &= \frac{\delta}{\delta f_i(\mathbf{x})} \left(\int d^3\mathbf{w}\, f_p(\mathbf{w}) \right) \\ &= \lim_{\varepsilon \to 0} \frac{\int d^3\mathbf{w}\, \left[f_p(\mathbf{w}) + \varepsilon \delta_{pi} \delta(\mathbf{w} - \mathbf{x}) \right] - \int d^3\mathbf{w}\, f_p(\mathbf{w})}{\varepsilon} \\ &= \delta_{pi} \int d^3\mathbf{w} \delta(\mathbf{w} - \mathbf{x}) = \delta_{pi}. \end{aligned} \tag{D.2}$$

As with normal derivatives, a chain rule can be used. Its basic building block follows directly from the definition, applied to the functional $G = f_i(\mathbf{x})$ and differentiated with respect to $f_j(\mathbf{y})$:

$$\frac{\delta f_i(\mathbf{x})}{\delta f_j(\mathbf{y})} = \delta_{ij} \delta(\mathbf{x} - \mathbf{y}). \tag{D.3}$$

Bibliography

[1] J. J. Sakurai. *Modern Quantum Mechanics*.

[2] R. P. Feynman. *The Feynman Lectures on Physics, Vol. III*.

[3] S. Weinberg. *Lectures on Quantum Mechanics*.

[4] J. J. Sakurai and Jim Napolitano. *Modern Quantum Mechanics*, chapter 6, pages 397–398. Addison-Wesley, 3rd edition, 2020.