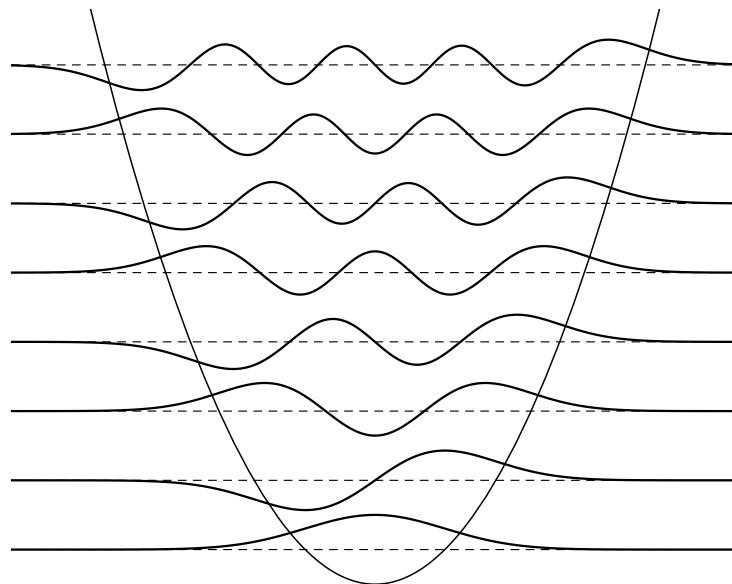


# PHYSICS 714

## QUANTUM MECHANICS B

Department of Physics - Stellenbosch University



2022

## Prerequisites:

It is assumed that you have a good understanding of basic wave mechanics and its probabilistic interpretation. You should also be familiar with the standard examples: spin-1/2, a free particle, a particle in a box, the simple harmonic oscillator and the hydrogen atom. We will occasionally make use of probability currents and the continuity equation. Time-independent perturbation theory is used in some of the examples. The third-year textbook by Griffiths is a good resource for refreshing your memory. Appendix C contains information on some of these topics.

## References:

- J.J. Sakurai, *Modern Quantum Mechanics*
- R. Shankar, *Principles of Quantum Mechanics*
- L.E. Ballentine, *Quantum Mechanics: A Modern Development*  
(Good for sorting out some mathematical details.)
- D.J. Griffiths, *Introduction to Quantum Mechanics*

*Please report any errors or typos.*

# Contents

<b>1 Mathematical Background</b>	<b>1</b>
1.1 Basic definitions . . . . .	1
1.2 Properties of Hilbert spaces . . . . .	4
1.2.1 Examples of Hilbert spaces . . . . .	5
1.2.2 Example: The SHO basis for $L^2(\mathbb{R})$ . . . . .	6
1.3 Linear operators . . . . .	7
1.3.1 Definition . . . . .	7
1.3.2 The matrix representation . . . . .	7
1.3.3 The adjoint . . . . .	9
1.3.4 Functions of operators or matrices . . . . .	9
1.3.5 Eigenvectors and eigenvalues . . . . .	10
1.4 Dirac's bra-ket notation . . . . .	10
1.4.1 Definitions . . . . .	10
1.4.2 The correspondence between bras and kets . . . . .	10
1.4.3 Operators in bra-ket notation . . . . .	11
1.4.4 Eigenvalue equations in bra-ket notation . . . . .	12
1.4.5 The outer product . . . . .	12
1.4.6 Bra-ket notation the easy way . . . . .	13
1.5 Tensor product spaces . . . . .	13
1.5.1 Definition . . . . .	13
1.5.2 Operators on tensor product spaces . . . . .	14
1.6 Hermitian operators . . . . .	14
1.6.1 Basic properties . . . . .	15
1.6.2 Commuting Hermitian operators . . . . .	15
1.6.3 Complete sets of commuting observables . . . . .	16
1.6.4 Example: Quantum numbers for hydrogen . . . . .	16
1.7 Unitary operators . . . . .	16
1.8 Representations in terms of orthogonal bases . . . . .	17
1.8.1 Discrete bases . . . . .	17
1.8.2 Continuous bases . . . . .	18
1.9 The Dirac delta function . . . . .	18
1.9.1 Definition . . . . .	18
1.9.2 Properties of the Dirac delta function . . . . .	19
1.9.3 Representation of the Dirac delta function . . . . .	20
1.10 Commutator identities . . . . .	21

<b>2 Fundamental Postulates</b>	<b>23</b>
2.1 Fundamental Postulates I	23
2.2 Exercise: Repeated spin measurements	24
2.3 Classical uncertainty and quantum mechanics	24
2.4 The density matrix of an ensemble	26
2.4.1 Background	26
2.4.2 Construction of the density matrix	26
2.4.3 Spin-1/2 example (continued)	27
2.4.4 Non-unique representations	28
2.5 Fundamental Postulates II	28
2.5.1 The eigenvalues of $\hat{\rho}$	29
2.6 Pure and mixed states	29
2.6.1 Pure states	29
2.6.2 Mixed states	30
2.7 The von Neumann entropy	30
2.8 Density matrix of the canonical ensemble	31
<b>3 Recovering Wave Mechanics</b>	<b>32</b>
3.1 The position operator	32
3.2 Momentum and spatial translations	32
3.3 Commutation relations	34
3.4 Finite translations	35
3.5 The coordinate representation	35
3.5.1 States in the coordinate representation	36
3.5.2 Operators in the coordinate representation	36
3.5.3 The time-independent Schrödinger equation	37
3.6 The momentum representation	37
3.6.1 Definition	37
3.6.2 Transforming between coordinate and momentum space wave functions	38
3.7 Active and passive transformations	39
3.7.1 Equivalence of the active and passive pictures	40
<b>4 Quantum Dynamics</b>	<b>41</b>
4.1 The time evolution operator	41
4.2 The time-dependent Schrödinger equation	42
4.3 The time evolution of states and observables	42
4.3.1 Example: Spin precession in the Schrödinger picture	43
4.3.2 The Schrödinger and Heisenberg pictures	44
<b>5 Angular Momentum</b>	<b>46</b>
5.1 Introduction	46
5.2 The rotation group $SO(3)$	46
5.3 Rotation matrices	47
5.3.1 Generators	47
5.3.2 Finite rotations	47
5.3.3 Commutation relations and group multiplication	48
5.4 Rotations in quantum mechanics	49

5.5	Orbital angular momentum . . . . .	50
5.6	Eigenvalues and matrix elements of angular momentum operators . . . . .	52
5.6.1	Ladder operators . . . . .	52
5.6.2	Eigenvalues . . . . .	53
5.6.3	Notation . . . . .	53
5.6.4	Matrix elements . . . . .	54
5.6.5	Reducible and irreducible representations . . . . .	54
5.7	Matrix elements of rotation operators . . . . .	55
5.7.1	Euler angles . . . . .	55
5.7.2	Calculating the matrix elements . . . . .	56
5.8	Spin-1/2 rotations . . . . .	56
5.9	Orbital angular momentum in the position representation . . . . .	58
5.9.1	Wave functions . . . . .	58
5.9.2	Observables . . . . .	59
5.10	Angular momentum addition . . . . .	60
5.10.1	Background . . . . .	60
5.10.2	Rotations of composite systems . . . . .	60
5.10.3	Eigenstates of the total angular momentum . . . . .	61
5.10.4	Selection rules for Clebsch-Gordon coefficients . . . . .	62
5.10.5	Constructing the coupled basis . . . . .	62
5.11	Multi-component wave functions . . . . .	65
5.12	Scalar, vector and tensor operators . . . . .	66
5.12.1	Background . . . . .	66
5.12.2	Definitions . . . . .	66
5.12.3	Matrix elements of tensor operators . . . . .	68
5.12.4	Exercise . . . . .	69
<b>6</b>	<b>Identical Particles</b>	<b>70</b>
6.1	Introduction . . . . .	70
6.2	Indistinguishability and observables . . . . .	70
6.3	Symmetrization . . . . .	71
<b>7</b>	<b>Time-Dependent Perturbation Theory</b>	<b>74</b>
7.1	Background . . . . .	74
7.2	Reminder: Quantum dynamics . . . . .	75
7.3	The interaction picture . . . . .	75
7.3.1	Definition . . . . .	75
7.3.2	The dynamics of a two-state system . . . . .	76
7.4	Perturbation theory . . . . .	78
7.4.1	The Dyson series . . . . .	78
7.4.2	Transition probabilities from perturbation theory . . . . .	79
7.5	Applications . . . . .	80
7.5.1	Constant perturbation . . . . .	80
7.5.2	Example: Fermi's golden rule and scattering in 2D . . . . .	83
7.5.3	Harmonic perturbation . . . . .	85

<b>8 Scattering Theory</b>	<b>87</b>
8.1 Background . . . . .	87
8.2 Classical scattering . . . . .	87
8.2.1 The scattering angle . . . . .	87
8.2.2 The total scattering cross section . . . . .	88
8.2.3 Differential scattering cross section . . . . .	89
8.2.4 Example: Hard-sphere scattering . . . . .	90
8.3 Quantum Scattering - Introduction . . . . .	90
8.4 The Lippmann-Schwinger equation . . . . .	92
8.5 The scattering amplitude and differential cross section . . . . .	94
8.6 The Born approximation . . . . .	96
8.6.1 Validity of the Born approximation . . . . .	96
8.6.2 Applications of the Born approximation . . . . .	98
8.7 Partial wave expansion . . . . .	99
8.7.1 Background . . . . .	99
8.7.2 Some additional comments . . . . .	100
8.7.3 Asymptotic form of $\hat{H}$ eigenstates . . . . .	100
8.7.4 Asymptotics of the incident plane wave . . . . .	102
8.7.5 Asymptotics of the outgoing scattered wave . . . . .	103
8.7.6 Phase shifts and the cross section . . . . .	104
8.7.7 Calculating the phase shifts . . . . .	104
8.7.8 Application: Hard-sphere scattering I . . . . .	105
8.7.9 Application: Hard-sphere scattering II . . . . .	106
8.7.10 Application: Scattering off a finite spherical well . . . . .	108
8.7.11 Why $s$ -wave scattering dominates at low energies . . . . .	110
<b>A Dimensional Analysis Basics</b>	<b>112</b>
<b>B Time-Independent Perturbation Theory</b>	<b>114</b>
B.1 Introduction . . . . .	114
B.2 Non-Degenerate Perturbation Theory . . . . .	114
B.3 Degenerate Perturbation Theory . . . . .	115
<b>C Elementary Systems</b>	<b>117</b>
C.1 The harmonic oscillator . . . . .	117
C.2 The infinite square well . . . . .	117
C.3 Spin-1/2 summary . . . . .	118
<b>D Special Functions</b>	<b>119</b>
<b>E Physics 254 Notes</b>	<b>120</b>
E.1 Introduction . . . . .	122
E.2 The Basic Postulates . . . . .	122
E.2.1 Postulate 1 . . . . .	122
E.2.2 Postulate 2a . . . . .	123
E.2.3 Definitions . . . . .	124
E.2.4 Postulate 2b . . . . .	124
E.3 Final Remarks . . . . .	125

<b>F Hydrogen-Like Wave Functions</b>	<b>126</b>
<b>G Angular Momentum Related Coefficients</b>	<b>128</b>
G.1 Ladder operator coefficients . . . . .	128
<b>H Further Information</b>	<b>130</b>

# Chapter 1

## Mathematical Background

### 1.1 Basic definitions

#### Vector Space

A complex vector space  $V$  is a set of elements, called vectors, which is closed under addition and scalar multiplication:

$$\text{If } \phi, \psi \in V \text{ and } a, b \in \mathbb{C} \text{ then } a\phi + b\psi \in V.$$

The operations of addition and scalar multiplication are

1. Distributive:  $a(\phi + \psi) = a\phi + a\psi$  and  $(a + b)\phi = a\phi + b\phi$
2. Associative:  $a(b\phi) = (ab)\phi$  and  $\phi + (\psi + \chi) = (\phi + \psi) + \chi$
3. Commutative:  $\phi + \psi = \psi + \phi$

There exists a zero vector, denoted by 0, such that  $\phi + 0 = \phi$  for all  $\phi \in V$ .

#### Linearly independent vectors

A set of vectors  $\{\phi_1, \phi_2, \phi_3, \dots\}$  is linearly independent if

$$\sum_n c_n \phi_n = 0 \text{ implies that } c_n = 0 \text{ for all } n. \quad (1.1)$$

This means that the **only** linear combination of  $\phi_n$ s that produces the zero vector is the trivial one with all the coefficients equal to zero.

#### Dimension of a vector space

The maximum number of linearly independent vectors in any set is called the dimension of the vector space. The dimension can generally be finite, countably infinite or uncountably infinite.

## Basis

A maximal set  $\{\phi_1, \phi_2, \phi_3, \dots\}$  of linearly independent vectors in  $V$  is called a basis of the vector space. These vectors **span** the vector space and so any  $\psi \in V$  can be expressed as a unique linear combination of the basis vectors:

$$\psi = \sum_n c_n \phi_n. \quad (1.2)$$

The number of vectors in any basis is again the dimension of the vector space.

## Inner product

An inner product on a vector space  $V$  is a mapping  $(\cdot, \cdot) : V \times V \rightarrow \mathbb{C}$  with the properties:

1.  $(\phi, \psi)^* = (\psi, \phi)$
2.  $(\phi, c_1\psi_1 + c_2\psi_2) = c_1(\phi, \psi_1) + c_2(\phi, \psi_2)$     (**Linear in second argument**)
3.  $(\phi, \phi) \geq 0$
4.  $(\phi, \phi) = 0$  if and only if  $\phi = 0$

It follows from (1) and (2) that  $(c_1\psi_1 + c_2\psi_2, \phi) = c_1^*(\psi_1, \phi) + c_2^*(\psi_2, \phi)$  which shows that the inner product is **anti-linear in the first argument**.

A vector space on which an inner product has been defined is called an inner product space.

## Normed space

A normed space is a vector space  $V$  together with a function  $\|\cdot\| : V \rightarrow \mathbb{R}$ , called the norm, which satisfies:

1.  $\|\phi\| \geq 0$
2.  $\|\phi\| = 0$  if and only if  $\phi = 0$
3.  $\|c\phi\| = |c| \|\phi\|$  for all  $c \in \mathbb{C}$
4.  $\|\phi + \psi\| \leq \|\phi\| + \|\psi\|$

## Metric space

A metric space is a set  $S$  together with a function  $d(\cdot, \cdot) : S \times S \rightarrow \mathbb{R}$ , called the metric, which satisfies:

1.  $d(x, y) \geq 0$
2.  $d(x, y) = 0$  if and only if  $x = y$
3.  $d(x, y) = d(y, x)$
4.  $d(x, y) \leq d(x, z) + d(z, y)$     (**Triangle inequality**)

## The inner product induced norm and metric

If  $V$  is an inner product space we can define a natural norm and metric on  $V$  by

$$\|\phi\| \equiv \sqrt{(\phi, \phi)} \quad \text{and} \quad d(\phi, \psi) \equiv \|\phi - \psi\|. \quad (1.3)$$

The inner product satisfies the Schwarz inequality

$$|(\phi, \psi)|^2 \leq (\phi, \phi)(\psi, \psi) \quad \text{or equivalently} \quad |(\phi, \psi)| \leq \|\phi\| \|\psi\|. \quad (1.4)$$

If  $(\phi, \psi) = 0$  we say that  $\psi$  and  $\phi$  are **orthogonal**. A vector  $\phi$  is **normalised** if  $\|\phi\| = 1$ .

## Orthogonal bases

A basis  $\{\phi_1, \phi_2, \phi_3, \dots\}$  of an inner product space is orthogonal if  $(\phi_n, \phi_m) = \delta_{n,m}$  for all  $n$  and  $m$ .<sup>1</sup>

## Convergence in a metric space

A sequence  $\{x_n : n = 1, 2, 3, \dots\}$  of points in a metric space  $S$  converges to a point  $x \in S$  if for any  $\epsilon > 0$  there exists an  $N_\epsilon$  such that  $d(x_n, x) \leq \epsilon$  for all  $n \geq N_\epsilon$ .

## Cauchy sequences

A sequence  $\{x_n\}$  of points in a metric space  $S$  is a Cauchy sequence if for any  $\epsilon > 0$  there exists an  $N_\epsilon$  such that  $d(x_n, x_m) \leq \epsilon$  for all  $n, m \geq N_\epsilon$ .

## Discussion

Roughly speaking, the elements of a convergent sequence tend towards a particular point in the metric space, while the elements of a Cauchy sequence tend towards one other. It is not difficult to show that any convergent sequence is also a Cauchy sequence. However, the reverse is not always true. Consider the metric space  $S = (0, 1]$  with metric  $d(x, y) = |x - y|$ . The sequence  $x_n = 1/n$  is Cauchy, since  $d(x_n, x_m) = |x_n - x_m| < x_n + x_m$  and we can make  $x_n + x_m$  arbitrarily small by taking  $n$  and  $m$  to be sufficiently large. However, this sequence does not converge to any point in  $S$ . If we consider the same sequence in the space  $\bar{S} = [0, 1]$  we see that it indeed converges to  $0 \in \bar{S}$ . In fact, in  $\bar{S}$  any Cauchy sequence will be convergent, which is not true in  $S$ . This leads to the definition below.

## Complete metric spaces

A metric space  $S$  is complete if every Cauchy sequence in the space converges to a point in  $S$ , i.e.  $S$  must contain the limit points of all its Cauchy sequences.

---

<sup>1</sup>Since these basis vectors are also normalised this is actually an *orthonormal* basis. We will only deal with normalised basis vectors, and so orthogonal and orthonormal will be treated as being synonymous.

## Hilbert space

A Hilbert space  $\mathcal{H}$  is a complete inner product space. Every Cauchy sequence of vectors in  $\mathcal{H}$  is therefore convergent.

## Linear functionals

A linear functional  $F$  on a vector space  $V$  is a function from  $V$  to  $\mathbb{C}$  such that  $F(a\phi + b\psi) = aF(\phi) + bF(\psi)$  for all vectors  $\phi, \psi \in V$  and scalars  $a, b \in \mathbb{C}$ .

## Dual space

The set of all linear functionals on a vector space  $V$  is denoted by  $V^*$  and is called the dual space of  $V$ . The dual space can be made into a vector space by defining the linear combination of two functionals  $F_1$  and  $F_2$  by  $(aF_1 + bF_2)(\psi) \equiv aF_1(\psi) + bF_2(\psi)$ .

## Isomorphisms and isometries

Let  $V$  and  $V'$  be two vector spaces. A bijective (i.e. injective and surjective) linear function  $\hat{T} : V \rightarrow V'$  is called a (vector space) isomorphism. If such a function exists the vector spaces  $V$  and  $V'$  are said to be isomorphic. Since  $\hat{T}$  preserves scalar multiplication and vector addition this implies that  $V$  and  $V'$  are mathematically identical.

Suppose  $\mathcal{H}$  and  $\mathcal{H}'$  are Hilbert spaces which are isomorphic as vector spaces through the function  $\hat{T}$ . If  $\hat{T}$  also preserves the inner product, i.e.  $(\hat{T}(\phi), \hat{T}(\psi)) = (\phi, \psi)$ , we call  $\hat{T}$  an isometry. In this case  $\mathcal{H}$  and  $\mathcal{H}'$  are isomorphic as Hilbert spaces.

## 1.2 Properties of Hilbert spaces

The following is of fundamental importance:

**The Riesz-Fischer Theorem:** *Every separable Hilbert space  $\mathcal{H}$  possesses a countable orthogonal basis.*

We usually denote such a basis by  $\{\phi_n\}$ , which is shorthand for  $\{\phi_n : n \in \mathcal{I}\}$  where  $\mathcal{I}$  is some countable index set. Since we can label the vectors as we please the precise nature of  $\mathcal{I}$  is usually left vague. By definition the elements of the orthogonal basis satisfy

$$(\phi_n, \phi_m) = \delta_{n,m}. \quad (1.5)$$

Any vector  $\psi \in \mathcal{H}$  can be expressed as

$$\psi = \sum_n c_n \phi_n \quad (1.6)$$

where the expansion coefficients can be determined using the inner product:

$$(\phi_m, \psi) = \sum_n c_n (\phi_m, \phi_n) = c_m. \quad (1.7)$$

All the Hilbert spaces we will consider are separable. Another important result is:

**Riesz's Theorem:** For every linear functional  $F \in \mathcal{H}^*$  there exists a unique vector  $f \in \mathcal{H}$  such that

$$F(\psi) = (f, \psi) \quad \text{for all } \psi \in \mathcal{H}. \quad (1.8)$$

This establishes a one-to-one correspondence between  $\mathcal{H}^*$  and  $\mathcal{H}$ , i.e. between the Hilbert space and its dual.

**Proof:** Given a vector  $f \in \mathcal{H}$  we can clearly define a linear functional by  $F(\psi) = (f, \psi)$ . We need to show that *all* linear functionals are of this form. Suppose  $F \in \mathcal{H}^*$  is given. We can construct the required vector  $f$  as follows. By the Riesz-Fischer theorem there exists an orthogonal basis  $\{\phi_n\}$  for  $\mathcal{H}$ . If we set  $f = \sum_m F^*(\phi_m)\phi_m$  then

$$(f, \phi_n) = \sum_m F(\phi_m)(\phi_m, \phi_n) = F(\phi_n) \quad (1.9)$$

for any basis vector  $\phi_n$ . Since this holds for all the basis vectors it will, by linearity, hold for any  $\psi \in \mathcal{H}$ . ■

### 1.2.1 Examples of Hilbert spaces

1. Any finite dimensional inner product space is complete and therefore a Hilbert space. If the space is complex and  $D$ -dimensional then it is isometric to  $\mathbb{C}^D$ , the space of  $D$ -dimensional column vectors with the Euclidean inner product. See section 1.3.2 for details.
2. The sequence space

$$\ell^2 = \{(x_1, x_2, x_3, \dots) : x_i \in \mathbb{C} \text{ and } \sum_{n=1}^{\infty} |x_n|^2 < \infty\} \quad (1.10)$$

with inner product  $(x, y) = \sum_{n=1}^{\infty} x_n^* y_n$ .

3. The space of periodic functions with period  $L$ . The inner product is

$$(f, g) = \int_0^L dx f^*(x)g(x). \quad (1.11)$$

The standard choice of basis for this space is the Fourier basis:

$$\phi_n(x) = \frac{1}{\sqrt{L}} e^{2\pi i n x / L} \quad n \in \mathbb{Z}. \quad (1.12)$$

4. The space  $L^2(\mathbb{R}^n)$  of square integrable functions on  $\mathbb{R}^n$ . The inner product is

$$(f, g) = \int_{\mathbb{R}^n} d\vec{x} f^*(\vec{x})g(\vec{x}). \quad (1.13)$$

This is the Hilbert space of wave functions in  $n$ -dimensions.

It might appear that there exists a wide variety of different Hilbert spaces. Surprisingly, this is not the case. In fact, *any* separable Hilbert space  $\mathcal{H}$  is isometric to either  $\mathbb{C}^D$  or  $\ell^2$ , depending on its dimensionality. This is a direct consequence of the Riesz-Fischer theorem which guarantees the existence of an orthogonal basis  $\{\phi_n\}$ . The isometry then simply amounts to mapping each vector  $\psi = \sum_n c_n \phi_n$  onto the sequence or column vector formed by the expansion coefficients  $\{c_n\}$ . This underpins the matrix representation presented in section 1.3.2.

### 1.2.2 Example: The SHO basis for $L^2(\mathbb{R})$

In appropriate dimensionless coordinates the eigenfunctions of the simple harmonic oscillator Hamiltonian are

$$\phi_n(x) = \frac{1}{\sqrt{2^n n! \pi^{\frac{1}{4}}}} H_n(x) e^{-x^2/2} \quad \text{with} \quad n = 0, 1, 2, \dots \quad (1.14)$$

where  $H_n(x)$  is the  $n$ th order Hermite polynomial. These functions form an orthogonal basis for  $L^2(\mathbb{R})$ , and so any other element can be expressed as a linear combination of these. Such a linear combination will generally contain an infinite number of non-zero terms. We can still produce a very good approximation of a function by using a sufficiently large, but finite, linear combination of basis elements. This is illustrated in figure 1.2.

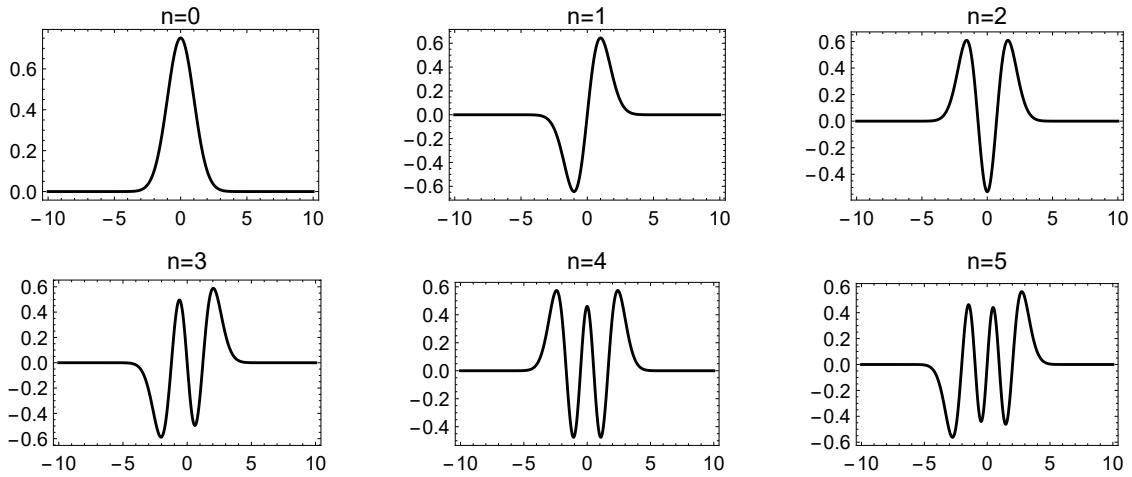


Figure 1.1: The first six basis functions  $\phi_n(x)$ .

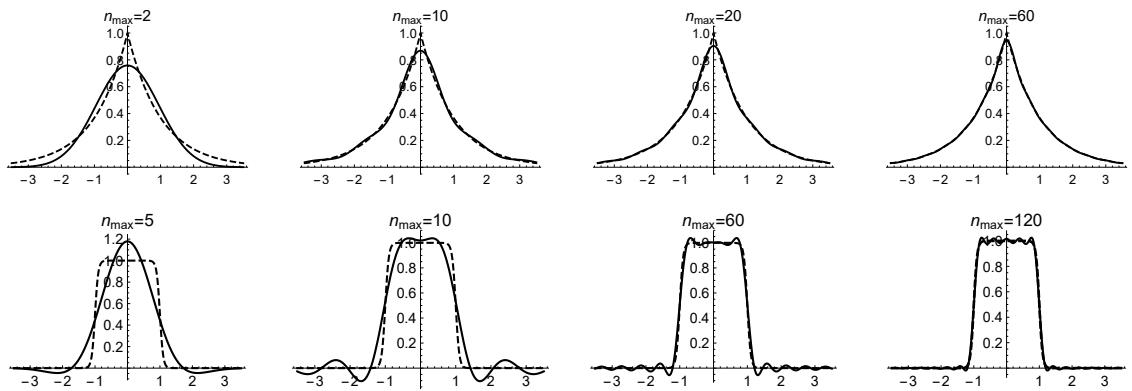


Figure 1.2: Approximations of functions (shown dashed) using linear combinations of the SHO basis elements up to  $n = n_{max}$ . For the top row the function to be approximated is  $f(x) = e^{-|x|}$ , while for the bottom row  $f(x) = (\tanh[10(1-|x|)]+1)/2$ .

## 1.3 Linear operators

### 1.3.1 Definition

A linear operator  $\hat{A}$  from  $\mathcal{H}$  to  $\mathcal{H}$  is a linear function which maps vectors onto vectors. We use the notation  $\hat{A}\phi \equiv \hat{A}(\phi)$ . The linearity condition then reads

$$\hat{A}(a\phi + b\psi) = a\hat{A}\phi + b\hat{A}\psi \quad (1.15)$$

for all  $\phi, \psi \in \mathcal{H}$  and  $a, b \in \mathbb{C}$ . Natural definitions for scalar multiplication, operator multiplication and operator addition are given by

$$(c\hat{A})\psi \equiv c(\hat{A}\psi) \quad (\hat{A}\hat{B})\psi \equiv \hat{A}(\hat{B}\psi) \quad (\hat{A} + \hat{B})\psi \equiv \hat{A}\psi + \hat{B}\psi. \quad (1.16)$$

Note that operator multiplication is associative, i.e.  $(\hat{A}\hat{B})\hat{C} = \hat{A}(\hat{B}\hat{C})$ , and so we can just write  $\hat{A}\hat{B}\hat{C}$  without fear of ambiguity. **Note that operator multiplication is not commutative and so  $\hat{A}\hat{B}$  and  $\hat{B}\hat{A}$  need not be equal.** We define the **commutator** of two operators  $\hat{A}$  and  $\hat{B}$  as the operator

$$[\hat{A}, \hat{B}] \equiv \hat{A}\hat{B} - \hat{B}\hat{A}. \quad (1.17)$$

The identity operator  $\hat{I}$  is defined by the property that  $\hat{I}\psi = \psi$  for all  $\psi \in \mathcal{H}$ . The inverse of an operator  $\hat{A}$ , if it exists, is denoted  $\hat{A}^{-1}$  and satisfies  $\hat{A}\hat{A}^{-1} = \hat{A}^{-1}\hat{A} = \hat{I}$ .

### 1.3.2 The matrix representation

The matrix representation is a means of representing abstract vectors and operators in terms of column vectors and matrices. It is a very useful computational tool, especially for computer-based calculations.

Let  $\mathcal{H}$  be a  $D$ -dimensional Hilbert space with orthogonal basis  $\{\phi_n : n = 1, 2, 3, \dots\}$ , and consider two arbitrary vectors  $\phi = \sum_n c_n \phi_n$  and  $\psi = \sum_n b_n \phi_n$ . We now define a mapping between  $\mathcal{H}$  and  $\mathbb{C}^D$  by<sup>2</sup>

$$\phi = \sum_n c_n \phi_n \iff c = \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ \vdots \end{bmatrix} \quad \text{and} \quad \psi = \sum_n b_n \phi_n \iff b = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ \vdots \end{bmatrix} \quad (1.18)$$

Not only is this mapping one-to-one, but it is also a vector space isomorphism since it clearly preserves both scalar multiplication and vector addition. Now suppose there exists an operator  $\hat{A}$  such that  $\hat{A}\phi = \psi$ . To see what this implies about the column vectors  $c$  and  $b$  we first take the inner product with a basis state  $\phi_m$  on both sides of  $\hat{A}\phi = \psi$  and then expand both  $\phi$  and  $\psi$  in the  $\{\phi_n\}$  basis. This produces

$$\sum_n c_n (\phi_m, \hat{A}\phi_n) = \sum_n b_n (\phi_m, \phi_n) = b_m. \quad (1.19)$$

---

<sup>2</sup>If  $D = \infty$  the column vectors will be infinitely long and should be identified with elements of  $\ell^2$ .

Now set  $A_{m,n} = (\phi_m, \hat{A}\phi_n)$ , which is called the matrix element<sup>3</sup> of  $\hat{A}$  between  $\phi_m$  and  $\phi_n$ . In this notation we have

$$\sum_n A_{m,n} c_n = b_m \quad (1.20)$$

which we immediately recognise as a matrix equation. The operator  $\hat{A}$  can therefore be represented by the matrix

$$A = \begin{bmatrix} A_{1,1} & A_{1,2} & A_{1,3} & \cdots \\ A_{2,1} & A_{2,2} & A_{2,3} & \cdots \\ A_{3,1} & A_{3,2} & A_{3,3} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \quad (1.21)$$

The abstract equation  $\hat{A}\phi = \psi$  has now been translated into the more concrete matrix-and-column-vector expression  $Ac = b$ . It should be noted that the matrix representation is basis-dependent. There is no “wrong” choice of basis, but some might be more convenient to use than others.

The mapping of  $\phi$  and  $\psi$  to  $c$  and  $b$  also preserves the inner product since

$$(\phi, \psi) = \sum_n c_n^* b_n = c^\dagger b \quad (1.22)$$

where  $\dagger$  denotes the conjugate transpose, i.e.  $c^\dagger = (c^*)^T$ . The right hand side is just the standard inner product on  $\mathbb{C}^D$ , i.e. the complex dot-product. This mapping of vectors onto column vectors is therefore an isometry: it preserves both the vector space structure as well as the inner product. *In this sense  $\mathcal{H}$  and  $\mathbb{C}^D$  are mathematically identical.*

From Reisz’s theorem we know that there exists a one-to-one correspondence between elements of the Hilbert space  $\mathcal{H}$  and its dual  $\mathcal{H}^*$ . Suppose  $F_\phi$  is the linear functional associated with the vector  $\phi$ , i.e.

$$F_\phi(\psi) = (\phi, \psi) = c^\dagger b. \quad (1.23)$$

It is sensible to represent  $F_\phi$  by the row vector  $c^\dagger$ , i.e. by the conjugate transpose of the column vector representing  $\phi$ :

$$F_\phi \iff c^\dagger = [c_1^* \ c_2^* \ c_3^* \dots] \quad (1.24)$$

In this representation the linear functional acts on the column vectors through regular matrix multiplication.

We end this section with an important definition. The **trace** of a linear operator  $\hat{A}$  is the sum of the diagonal elements of its matrix representation. We write

$$\text{tr}(\hat{A}) = \sum_n A_{n,n} = \sum_n (\phi_n, \hat{A}\phi_n) \quad (1.25)$$

where  $\{\phi_n\}$  is a basis of  $\mathcal{H}$ . Note that this sum might not exist if  $\mathcal{H}$  is infinite dimensional. It is important to remember that the value of the trace is independent of the choice of basis in which it is calculated.

---

<sup>3</sup>A linear operator can be represented by a matrix for any choice of basis. However, it is only for orthogonal bases that the matrix elements are given by this simple expression.

### 1.3.3 The adjoint

The adjoint of an operator  $\hat{A}$  on  $\mathcal{H}$  is denoted by  $\hat{A}^\dagger$  and satisfies

$$(\phi, \hat{A}\psi) = (\hat{A}^\dagger\phi, \psi) \quad \text{for all } \phi, \psi \in \mathcal{H}. \quad (1.26)$$

It follows from this defining property that

$$(\hat{a}\hat{A})^\dagger = a^*\hat{A}^\dagger \quad (1.27)$$

$$(\hat{A}^\dagger)^\dagger = \hat{A} \quad (1.28)$$

The following important observation is easy to prove: *If  $A$  is the matrix representation of  $\hat{A}$ , then  $A^\dagger = (A^*)^T$  is the matrix representation of  $\hat{A}^\dagger$ , i.e. on the matrix level the adjoint is just the conjugate transpose.*

### 1.3.4 Functions of operators or matrices

All statements regarding operators in this section are also applicable to matrices. Let  $\hat{A}$  be an operator and  $f(x)$  a function with a power series representation  $f(x) = \sum_n a_n x^n$ . We then define the operator  $f(\hat{A})$  by

$$f(\hat{A}) = \sum_n a_n \hat{A}^n. \quad (1.29)$$

The most prominent example of this is the exponential of an operator:

$$e^{\hat{A}} = \sum_{n=0}^{\infty} \frac{\hat{A}^n}{n!}. \quad (1.30)$$

The adjoint of  $f(\hat{A})$  is

$$(f(\hat{A}))^\dagger = \sum_n a_n^* (\hat{A}^\dagger)^n = f^*(\hat{A}^\dagger). \quad (1.31)$$

If an expression contains only one operator the usual rules of algebra and calculus apply. For example,

$$e^{a\hat{A}} e^{b\hat{A}} = e^{(a+b)\hat{A}} \quad \text{and} \quad \frac{d}{d\lambda} e^{\lambda\hat{A}} = \hat{A} e^{\lambda\hat{A}} = e^{\lambda\hat{A}} \hat{A}. \quad (1.32)$$

However, we must be extremely careful when manipulating expressions involving two or more operators that do not commute. For example, if  $[\hat{A}, \hat{B}] \neq 0$  then

$$e^{\hat{A}} e^{\hat{B}} \neq e^{\hat{A} + \hat{B}} \quad \text{and} \quad e^{-\hat{A}} \hat{B} e^{+\hat{A}} \neq \hat{B}. \quad (1.33)$$

When taking derivatives of expressions involving non-commuting operators we must always preserve the ordering, if this is at all possible. For example

$$\frac{d}{d\lambda} [e^{\lambda\hat{A}} e^{\lambda\hat{B}}] = \hat{A} e^{\lambda\hat{A}} e^{\lambda\hat{B}} + e^{\lambda\hat{A}} \hat{B} e^{\lambda\hat{B}}. \quad (1.34)$$

Note the position of  $\hat{B}$  in the second term; it must stand to the *right* of the  $e^{\lambda\hat{A}}$  factor. Even more problematic are derivatives such as

$$\frac{d}{d\lambda} e^{\hat{A} + \lambda\hat{B}} \quad (1.35)$$

for which there is no simple expression available when  $[\hat{A}, \hat{B}] \neq 0$ .

### 1.3.5 Eigenvectors and eigenvalues

A non-zero vector  $\phi \in \mathcal{H}$  is an eigenvector of an operator  $\hat{A}$  if

$$\hat{A}\phi = \lambda\phi \quad (1.36)$$

where  $\lambda \in \mathbb{C}$  is the corresponding eigenvalue. Since  $\hat{A}^n\phi = \lambda^n\phi$  it follows that for any function of  $\hat{A}$  we have

$$f(\hat{A})\phi = f(\lambda)\phi. \quad (1.37)$$

The set of eigenvalues of an operator or matrix is called its **spectrum**.

## 1.4 Dirac's bra-ket notation

### 1.4.1 Definitions

We now introduce a new notation which will allow expressions involving vectors, operators and functionals to be manipulated with ease. From now on we will often refer to elements of the Hilbert space as states rather than vectors. We define our new notation as follows:

- A vector  $\psi \in \mathcal{H}$  is now denoted by the **ket**  $|\psi\rangle \in \mathcal{H}$ .
- The linear functional  $F_\phi \in \mathcal{H}^*$  associated with the vector  $\phi \in \mathcal{H}$  through Reisz's theorem is now denoted by the **bra**  $\langle\phi| \in \mathcal{H}^*$ .

The action of the bra  $\langle\phi|$  on the ket  $|\psi\rangle$  is

$$\langle\phi| \text{ acting on } |\psi\rangle \equiv \underbrace{\langle\phi|\psi\rangle}_{\text{bra}(c)ket} \equiv F_\phi(\psi) = (\phi, \psi) \quad (1.38)$$

which also provides the new notation for the inner product. It is worth noting that

$$\langle\phi|\psi\rangle = (\phi, \psi) = (\psi, \phi)^* = \langle\psi|\phi\rangle^*. \quad (1.39)$$

At this stage we still label kets by the vectors that they represent. This allows us to write

$$a|\phi_1\rangle + b|\phi_2\rangle = |a\phi_1 + b\phi_2\rangle. \quad (1.40)$$

*Note that the identity above is really just notation, and a trivial consequence of the fact that we are labelling kets by the vectors they represent. Once we introduce other labelling schemes expressions like the one above will obviously no longer make sense!*

### 1.4.2 The correspondence between bras and kets

Reisz's theorem can now be formulated as the statement that for each ket  $|\phi\rangle$  there exists a unique bra  $\langle\phi|$ . Not only is this correspondence one-to-one, it is also anti-linear, i.e.

$$a|\phi_1\rangle + b|\phi_2\rangle \iff a^*\langle\phi_1| + b^*\langle\phi_2|. \quad (1.41)$$

To prove this we first write  $a|\phi_1\rangle + b|\phi_2\rangle = |a\phi_1 + b\phi_2\rangle$  for which the corresponding bra is  $\langle a\phi_1 + b\phi_2|$ . Now consider the action of this bra on an arbitrary ket  $|\psi\rangle$ :

$$\langle a\phi_1 + b\phi_2|\psi\rangle = (a\phi_1 + b\phi_2, \psi) = a^*(\phi_1, \psi) + b^*(\phi_2, \psi) \quad (1.42)$$

$$= a^*\langle\phi_1|\psi\rangle + b^*\langle\phi_2|\psi\rangle \quad (1.43)$$

$$= (a^*\langle\phi_1| + b^*\langle\phi_2|)|\psi\rangle \quad (1.44)$$

and so

$$\langle a\phi_1 + b\phi_2| = a^*\langle\phi_1| + b^*\langle\phi_2|. \quad (1.45)$$

### 1.4.3 Operators in bra-ket notation

We now want to define how operators act on bras and kets. Since kets just represent vectors it follows that if  $\hat{A}\phi = \psi$  then

$$\hat{A}|\phi\rangle = |\hat{A}\phi\rangle = |\psi\rangle. \quad (1.46)$$

Next we want to define the action of an operator on a bra. In particular, we want to give meaning to the combination  $\langle\phi|\hat{A}$ . Since  $\hat{A}$  maps vectors onto vectors while  $\langle\phi|$  maps vectors onto scalars the combination  $\langle\phi|\hat{A}$ , when read from right to left, can be interpreted as a linear mapping of vectors onto scalars, i.e. a linear functional. This motivates the following definition:

*We denote by  $\langle\phi|\hat{A}$  the linear functional which acts on a ket  $|\psi\rangle$  according to*

$$(\langle\phi|\hat{A})|\psi\rangle \equiv \langle\phi|(\hat{A}|\psi\rangle). \quad (1.47)$$

Since, by definition, the placement of parentheses is now irrelevant we can just write  $\langle\phi|\hat{A}|\psi\rangle$  to mean the expression above. Now, according to Reisz's theorem, if  $\langle\phi|\hat{A}$  is a linear functional it must be the bra of some ket  $|\phi'\rangle$ . In other words, there must exist a  $|\phi'\rangle$  such that  $\langle\phi|\hat{A} = \langle\phi'|\$ . The question is how  $|\phi'\rangle$  and  $|\phi\rangle$ , or  $\phi'$  and  $\phi$ , are related. To see this we note that

$$(\langle\phi|\hat{A})|\psi\rangle = (\phi, \hat{A}\psi) = (\hat{A}^\dagger\phi, \psi) = \langle\hat{A}^\dagger\phi|\psi\rangle \quad (1.48)$$

for all  $|\psi\rangle \in \mathcal{H}$ . It follows that

$$\langle\phi|\hat{A} = \langle\hat{A}^\dagger\phi| \quad \text{or equivalently} \quad \langle\phi|\hat{B}^\dagger = \langle\hat{B}\phi| \quad (1.49)$$

for any operators  $\hat{A}$  and  $\hat{B}$ . To put this into words: “When an operator  $\hat{A}$  acts to the left on the bra  $\langle\phi|$  corresponding to the ket  $|\phi\rangle$  it produces the bra  $\langle\hat{A}^\dagger\phi|$  corresponding to the ket  $|\hat{A}^\dagger\phi\rangle$  which results from letting  $\hat{A}^\dagger$  act (to the right) on  $|\phi\rangle$ ”.

$$\begin{array}{ccc} |\phi\rangle & \xrightarrow{\hat{A}^\dagger \text{ a.t.t.r.}} & \hat{A}^\dagger|\phi\rangle = |\hat{A}^\dagger\phi\rangle \\ \uparrow & & \downarrow \\ \langle\phi| & \xrightarrow{\hat{A} \text{ a.t.t.l.}} & \langle\phi|\hat{A} = \langle\hat{A}^\dagger\phi| \end{array}$$

It follows that the four expressions

$$\langle \hat{A}^\dagger \phi | \psi \rangle = (\langle \phi | \hat{A}) | \psi \rangle = \langle \phi | (\hat{A} | \psi \rangle) = \langle \phi | \hat{A} \psi \rangle \quad (1.50)$$

are all equivalent ways of writing the matrix element  $\langle \phi | \hat{A} | \psi \rangle = (\phi, \hat{A} \psi)$ . These different forms can be extremely useful. For example, we might know a lot more about  $\hat{A}^\dagger \phi$  than we do about  $\hat{A} \psi$ , or vice versa. *Henceforth we will always denote matrix elements using this new notation  $\langle \phi | \hat{A} | \psi \rangle$ .* A simple but extremely important result that follows from this discussion is that

$$\langle \phi | \hat{A} | \psi \rangle^* = \langle \psi | \hat{A}^\dagger | \phi \rangle. \quad (1.51)$$

In particular, this proves the earlier assertion that the adjoint just translates into the conjugate transpose on the level of the matrix representation.

#### 1.4.4 Eigenvalue equations in bra-ket notation

Suppose  $\hat{A}\phi = a\phi$ . In terms of kets this implies that

$$\hat{A}|\phi\rangle = a|\phi\rangle \quad \text{and} \quad f(\hat{A})|\phi\rangle = f(a)|\phi\rangle. \quad (1.52)$$

We also have

$$\langle \phi | \hat{A}^\dagger = \langle \hat{A} \phi | = \langle a \phi | = a^* \langle \phi | = \langle \phi | a^* \quad (1.53)$$

and so

$$\langle \phi | (f(\hat{A}))^\dagger = \langle \phi | f^*(\hat{A}^\dagger) = \langle \phi | f^*(a^*). \quad (1.54)$$

In particular, if  $|\phi\rangle$  is an eigenket of  $\hat{A}$  (acting to the right) with eigenvalue  $a$  then  $\langle \phi |$  is an eigenbra of  $\hat{A}^\dagger$  (acting to the left) with eigenvalue  $a^*$ .

#### 1.4.5 The outer product

The outer product of a bra  $\langle \psi |$  and a ket  $|\phi\rangle$  is a linear operator denoted by  $|\phi\rangle\langle \psi |$  and defined by its action on an arbitrary ket  $|\lambda\rangle$  as

$$(|\phi\rangle\langle \psi |) |\lambda\rangle \equiv |\phi\rangle\langle \psi | \lambda\rangle = \langle \psi | \lambda\rangle |\phi\rangle. \quad (1.55)$$

The adjoint of  $|\phi\rangle\langle \psi |$  can be found by considering the complex conjugate of an arbitrary matrix element:

$$[\langle \chi | (|\phi\rangle\langle \psi |) | \lambda \rangle]^* = (\langle \chi | \phi \rangle \langle \psi | \lambda \rangle)^* = \langle \lambda | \psi \rangle \langle \phi | \chi \rangle = \langle \lambda | (|\psi\rangle\langle \phi |) | \chi \rangle. \quad (1.56)$$

However, (1.51) implies that

$$[\langle \chi | (|\phi\rangle\langle \psi |) | \lambda \rangle]^* = \langle \lambda | (|\phi\rangle\langle \psi |)^\dagger | \chi \rangle \quad (1.57)$$

and so  $\langle \lambda | (|\phi\rangle\langle \psi |)^\dagger | \chi \rangle = \langle \lambda | (|\psi\rangle\langle \phi |) | \chi \rangle$  for arbitrary  $|\lambda\rangle$  and  $|\chi\rangle$ . If all the matrix elements of two operators are equal then the operators themselves must be equal, and therefore

$$(|\phi\rangle\langle \psi |)^\dagger = |\psi\rangle\langle \phi|. \quad (1.58)$$

### 1.4.6 Bra-ket notation the easy way

We have derived a number of important identities involving combinations of bras, kets, operators and scalars. In hindsight all these identities can be obtained using a simple recipe. First note that all these mathematical objects have natural “dual” partners associated with them, in the sense that

1. For each ket  $|\phi\rangle$  there exists a unique bra  $\langle\phi|$ , and vice versa.
2. Each operator  $\hat{A}$  has a unique adjoint  $\hat{A}^\dagger$ .
3. Each scalar  $z$  has a unique complex conjugate  $z^*$ .

Let us now *define* the  $\dagger$  operation on kets, bras and scalars as<sup>4</sup>

$$(|\phi\rangle)^\dagger \equiv \langle\phi| \quad ((\langle\phi|)^\dagger \equiv |\phi\rangle) \quad (z)^\dagger \equiv z^*. \quad (1.59)$$

The following statement is the result the various definitions we have introduced in previous sections.

*Suppose  $A$ ,  $B$ ,  $C$  and  $D$  are bras, kets, operators or scalars. Then*

$$(A + B)^\dagger = A^\dagger + B^\dagger \quad \text{and} \quad (CD)^\dagger = D^\dagger C^\dagger. \quad (1.60)$$

This means that any expression can be turned into its dual by first reversing the symbols in each term and then replacing each one by its own dual. In particular, this procedure allows us to easily translate between equivalent ket and bra equations. You should convince yourself that this procedure is compatible with the definitions and results of the previous sections.

In light of the isometry between  $\mathcal{H}$  and  $\mathbb{C}^D$  this recipe is easy to understand. We have simply translated the notion of the conjugate transpose on the level of the matrix representation into a more abstract language.

## 1.5 Tensor product spaces

The tensor product allows us to construct a Hilbert space by combining two other Hilbert spaces in a particular way.

### 1.5.1 Definition

Let  $\mathcal{H}_1$  and  $\mathcal{H}_2$  be Hilbert spaces with orthogonal bases  $\{|a_i\rangle : i = 1, 2, 3, \dots\}$  and  $\{|b_i\rangle : i = 1, 2, 3, \dots\}$  respectively. We now define the tensor product of  $\mathcal{H}_1$  and  $\mathcal{H}_2$  as the Hilbert space formed by all linear combinations of pairs of these basis states:

$$\mathcal{H}_1 \otimes \mathcal{H}_2 = \text{span}\{|a_n\rangle \otimes |b_m\rangle : n, m = 1, 2, 3, \dots\} \quad (1.61)$$

---

<sup>4</sup>We introduce this notation to make the algorithm we will use easier to state and apply. It is not completely standard, but some books do use it.

A general element of  $\mathcal{H}_1 \otimes \mathcal{H}_2$  therefore has the form

$$|\phi\rangle = \sum_{n,m} c_{n,m} |a_n\rangle \otimes |b_m\rangle. \quad (1.62)$$

We will often abbreviate  $|a_n\rangle \otimes |b_m\rangle$  by  $|a_n, b_m\rangle$ . The inner product on  $\mathcal{H}_1 \otimes \mathcal{H}_2$  is defined by

$$\langle a_n, b_m | a_{n'}, b_{m'} \rangle \equiv \langle a_n | a_{n'} \rangle \langle b_m | b_{m'} \rangle \quad (1.63)$$

and it follows that  $\{|a_n, b_m\rangle\}$  is an orthogonal basis for  $\mathcal{H}_1 \otimes \mathcal{H}_2$ . It is clear that  $\dim(\mathcal{H}_1 \otimes \mathcal{H}_2) = \dim(\mathcal{H}_1)\dim(\mathcal{H}_2)$ .

We also require that the tensor product acts linearly in the sense that

$$(c_1|\alpha\rangle + c_2|\beta\rangle) \otimes |\gamma\rangle = c_1|\alpha\rangle \otimes |\gamma\rangle + c_2|\beta\rangle \otimes |\gamma\rangle \quad (1.64)$$

$$|\alpha\rangle \otimes (c_1|\beta\rangle + c_2|\gamma\rangle) = c_1|\alpha\rangle \otimes |\beta\rangle + c_2|\alpha\rangle \otimes |\gamma\rangle \quad (1.65)$$

in which case

$$|\alpha\rangle \otimes 0 = 0 \otimes |\beta\rangle = 0. \quad (1.66)$$

### 1.5.2 Operators on tensor product spaces

Suppose  $\hat{A}_1$  and  $\hat{A}_2$  are operators acting on  $\mathcal{H}_1$  and  $\mathcal{H}_2$  respectively. We define the operator  $\hat{A}_1 \otimes \hat{A}_2$  on  $\mathcal{H}_1 \otimes \mathcal{H}_2$  by

$$\hat{A}_1 \otimes \hat{A}_2(|\alpha\rangle \otimes |\beta\rangle) \equiv (\hat{A}_1|\alpha\rangle) \otimes (\hat{A}_2|\beta\rangle). \quad (1.67)$$

Abusing notation slightly we can also define  $\hat{A}_1$  and  $\hat{A}_2$  directly on  $\mathcal{H}_1 \otimes \mathcal{H}_2$  by writing:

$$\hat{A}_1 \equiv \hat{A}_1 \otimes \hat{I} \quad \text{and} \quad \hat{A}_2 \equiv \hat{I} \otimes \hat{A}_2. \quad (1.68)$$

When  $\hat{A}_1$  ( $\hat{A}_2$ ) acts on states in  $\mathcal{H}_1 \otimes \mathcal{H}_2$  it therefore only affects the first (second) factor in a product state. As a result  $\hat{A}_1$  and  $\hat{A}_2$  always commute:

$$[\hat{A}_1, \hat{A}_2] = [\hat{A}_1 \otimes \hat{I}, \hat{I} \otimes \hat{A}_2] = \hat{A}_1 \otimes \hat{A}_2 - \hat{A}_1 \otimes \hat{A}_2 = 0. \quad (1.69)$$

The following properties should now be obvious:

$$(\hat{A}_1 \otimes \hat{A}_2)(\hat{B}_1 \otimes \hat{B}_2) = \hat{A}_1 \hat{B}_1 \otimes \hat{A}_2 \hat{B}_2 \quad (1.70)$$

$$(c_1 \hat{A} + c_2 \hat{B}) \otimes \hat{C} = c_1 \hat{A} \otimes \hat{C} + c_2 \hat{B} \otimes \hat{C} \quad (1.71)$$

$$\hat{C} \otimes (c_1 \hat{A} + c_2 \hat{B}) = c_1 \hat{C} \otimes \hat{A} + c_2 \hat{C} \otimes \hat{B} \quad (1.72)$$

$$\hat{I} \otimes \hat{I} = \hat{I} \quad (1.73)$$

## 1.6 Hermitian operators

An operator  $\hat{A}$  on  $\mathcal{H}$  is said to be Hermitian if  $\hat{A}^\dagger = \hat{A}$ . Operators of this type play a central role in quantum mechanics since they represent physical observables. In this section we investigate the properties of Hermitian operators that make them suitable for this task. All these results are clearly also applicable to Hermitian matrices.

### 1.6.1 Basic properties

**Property 1:** All the diagonal matrix elements of a Hermitian operator are real.

**Proof:** The result in (1.51) immediately implies that  $\langle \phi | \hat{A} | \phi \rangle$  is real.

**Property 2:** All the eigenvalues of a Hermitian operator are real.

**Proof:** Suppose  $\hat{A}|\phi\rangle = a|\phi\rangle$ . Then  $\langle \phi | \hat{A} | \phi \rangle = a\langle \phi | \phi \rangle$  and since both  $\langle \phi | \hat{A} | \phi \rangle$  and  $\langle \phi | \phi \rangle$  are real, so is the eigenvalue  $a$ .

**Property 3:** Eigenstates of a Hermitian operator with different eigenvalues are orthogonal.

**Proof:** Suppose  $\hat{A}|\phi_1\rangle = a_1|\phi_1\rangle$  and  $\hat{A}|\phi_2\rangle = a_2|\phi_2\rangle$ . We then have

$$\langle \phi_2 | \hat{A} | \phi_1 \rangle = a_1 \langle \phi_2 | \phi_1 \rangle. \quad (1.74)$$

However,  $\hat{A}|\phi_2\rangle = a_2|\phi_2\rangle$  implies that  $\langle \phi_2 | \hat{A} = \langle \phi_2 | a_2$  and so  $\langle \phi_2 | \hat{A} | \phi_1 \rangle = a_2 \langle \phi_2 | \phi_1 \rangle$ . It follows that

$$(a_1 - a_2) \langle \phi_2 | \phi_1 \rangle = 0 \quad (1.75)$$

and so if  $a_1 \neq a_2$  the two eigenstates must be orthogonal. ■

**Property 4:** The eigenstates of a Hermitian operator (can be chosen to) form a complete orthogonal basis for the Hilbert space.

Note that while eigenstates with different eigenvalues are always orthogonal this is not automatically true for eigenstates sharing the same eigenvalue, i.e. for states which are degenerate. However, we can always choose an orthogonal basis within each eigenspace.

### 1.6.2 Commuting Hermitian operators

The following result is of fundamental importance in quantum mechanics.

**Theorem:** If  $\hat{A}$  and  $\hat{B}$  are commuting Hermitian operators on  $\mathcal{H}$  then there exists an orthogonal basis for  $\mathcal{H}$  consisting of **simultaneous** eigenstates of  $\hat{A}$  and  $\hat{B}$ .

**Proof:** Let  $\hat{A}|\phi_n\rangle = a_n|\phi_n\rangle$  denote the eigenstates and eigenvalues of  $\hat{A}$ . We denote by  $\Lambda_a = \text{span}\{|\phi_n\rangle : a_n = a\}$  the eigenspace of  $\hat{A}$  corresponding to the eigenvalue  $a$ . Now consider an arbitrary ket  $|a\rangle \in \Lambda_a$  and note that

$$\hat{A}(\hat{B}|a\rangle) = \hat{A}\hat{B}|a\rangle = \hat{B}\hat{A}|a\rangle = a(\hat{B}|a\rangle). \quad (1.76)$$

It follows that  $\hat{B}|a\rangle$  is itself an eigenstate of  $\hat{A}$  with eigenvalue  $a$  and so must be an element of  $\Lambda_a$ . The operator  $\hat{B}$  therefore leaves each eigenspaces of  $\hat{A}$  invariant. For each eigenspace  $\Lambda_a$  we now consider the restriction  $\hat{B}|_{\Lambda_a} : \Lambda_a \rightarrow \Lambda_a$  and calculate its eigenstates. These states can be chosen to be orthogonal and are automatically eigenstates of  $\hat{A}$  as well. Combining the states from the various eigenspaces then produces an orthogonal basis for  $\mathcal{H}$  of simultaneous eigenstates of  $\hat{A}$  and  $\hat{B}$ . ■

**Exercise:** Show that if there exists a basis of simultaneous eigenstates of  $\hat{A}$  and  $\hat{B}$  then the two operators necessarily commute.

### 1.6.3 Complete sets of commuting observables

In quantum mechanics physical observables are represented by Hermitian operators. A set of commuting Hermitian operators  $\{\hat{A}_1, \hat{A}_2, \dots\}$  is said to be *complete* if the operators' eigenvalues together specify a unique simultaneous eigenstate. This defines an orthogonal basis of which the elements are labelled by their eigenvalues with respect to  $\{\hat{A}_1, \hat{A}_2, \dots\}$ . In this context the eigenvalues are known as **quantum numbers** and determine the values of the physical observables that the operators represent. *This is one of the most important ideas in quantum mechanics, and you need to understand it well.*

### 1.6.4 Example: Quantum numbers for hydrogen

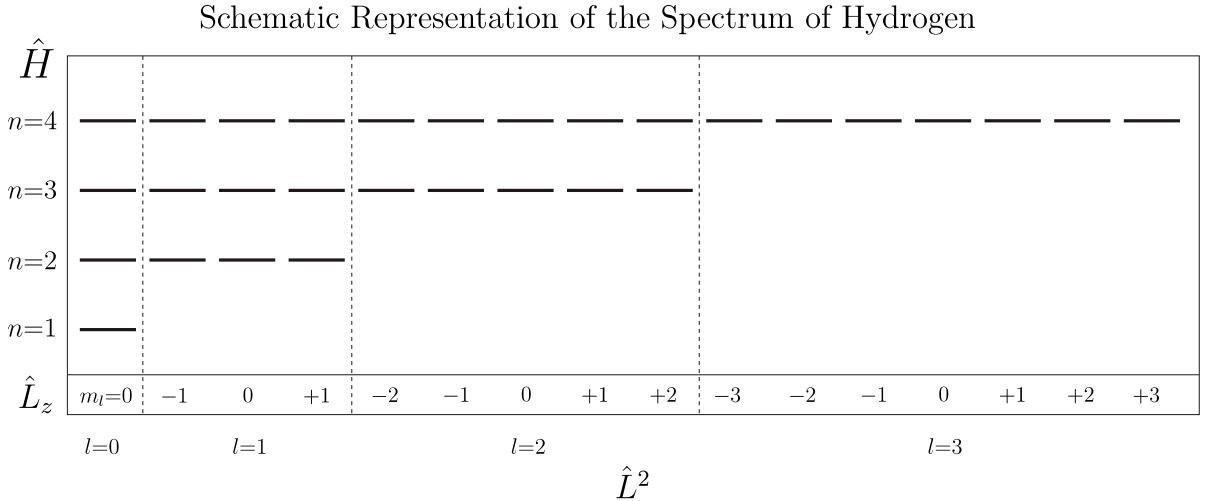
For hydrogen, the three operators  $\hat{H}$ ,  $\hat{L}^2$  and  $\hat{L}_z$  form a complete set of commuting observables. (We are ignoring the electron's spin for the moment.) The basis of simultaneous eigenstates will be labelled by the three quantum numbers  $n$ ,  $l$  and  $m_l$ . These are not the eigenvalues themselves, but they do determine the eigenvalues in a simple way. We have

$$\hat{H}|n, l, m_l\rangle = E_n|n, l, m_l\rangle = \frac{E_1}{n^2}|n, l, m_l\rangle \quad n = 1, 2, 3, \dots \quad (1.77)$$

$$\hat{L}^2|n, l, m_l\rangle = l(l+1)\hbar^2|n, l, m_l\rangle \quad l = 0, 1, \dots, n-1 \quad (1.78)$$

$$\hat{L}_z|n, l, m_l\rangle = m_l\hbar|n, l, m_l\rangle \quad m_l = -l, \dots, +l \quad (1.79)$$

where  $E_1$  is the ground state energy. Including the electron's spin would introduce another observable  $\hat{S}_z$  with corresponding quantum number  $m_s = \pm 1/2$ .



## 1.7 Unitary operators

An operator  $\hat{U}$  on  $\mathcal{H}$  is said to be unitary if  $\hat{U}^\dagger = \hat{U}^{-1}$ . These operators preserve the inner product in the sense that

$$\langle \hat{U}\phi | \hat{U}\psi \rangle = \langle \phi | \hat{U}^\dagger \hat{U} | \psi \rangle = \langle \phi | \psi \rangle. \quad (1.80)$$

In particular, they preserve both orthogonality and normalisation and can therefore be used to map one orthogonal basis onto another.

Unitary operators are also often used to transform other operators. If  $\hat{A}$  and  $\hat{B}$  are operators related by  $\hat{A} = \hat{U}^\dagger \hat{B} \hat{U}$  they are said to be *unitarily equivalent*. The mathematical properties of two unitarily equivalent operators are essentially identical since they amount to the same linear mapping just expressed in different bases. An important result in this regard is the following:

**Theorem:** *Two unitarily equivalent operators have exactly the same eigenvalues and their eigenstates are related by  $\hat{U}$  and  $\hat{U}^\dagger$ .*

**Proof:** Exercise.

We note that unitary transformations also preserve the Hermiticity of operators, since if  $\hat{A}^\dagger = \hat{A}$  then  $(\hat{U}^\dagger \hat{A} \hat{U})^\dagger = \hat{U}^\dagger \hat{A} \hat{U}$ .

## 1.8 Representations in terms of orthogonal bases

### 1.8.1 Discrete bases

Let  $\{|\phi_n\rangle\}$  be a discrete orthogonal basis for  $\mathcal{H}$ . By definition

$$\langle \phi_n | \phi_m \rangle = \delta_{n,m}, \quad (1.81)$$

and we know that any  $|\psi\rangle \in \mathcal{H}$  can be written as  $|\psi\rangle = \sum_n c_n |\phi_n\rangle$  where  $c_n = \langle \phi_n | \psi \rangle$ . Therefore

$$|\psi\rangle = \sum_n |\phi_n\rangle \langle \phi_n | \psi \rangle = \left( \sum_n |\phi_n\rangle \langle \phi_n| \right) |\psi\rangle \quad (1.82)$$

holds for all  $|\psi\rangle \in \mathcal{H}$ , which is only possible if

$$\hat{I} = \sum_n |\phi_n\rangle \langle \phi_n|. \quad (1.83)$$

This is called the **resolution of the identity** in terms of the  $\{|\phi_n\rangle\}$  basis. It is a very important and useful result. A similar representation can be found for Hermitian operators. Suppose  $\hat{A}$  is Hermitian with eigenstates and eigenvalues given by  $\hat{A}|\phi_n\rangle = a_n|\phi_n\rangle$ . Since these eigenstates form an orthogonal basis for  $\mathcal{H}$  we see that

$$\hat{A} = \hat{A}\hat{I} = \sum_n \hat{A}|\phi_n\rangle \langle \phi_n| = \sum_n a_n |\phi_n\rangle \langle \phi_n|. \quad (1.84)$$

This is the **spectral representation** of  $\hat{A}$ . Similarly, for any function of  $\hat{A}$ , we have

$$f(\hat{A}) = \sum_n f(a_n) |\phi_n\rangle \langle \phi_n|. \quad (1.85)$$

In fact, this allows us to *define*  $f(\hat{A})$  for functions which do not have Taylor series representations.

### 1.8.2 Continuous bases

Up to now we have only considered countable bases in which the elements are labelled with a discrete index. However, there are operators in quantum mechanics which have continuous sets of eigenvalues and of which the eigenstates form a continuous basis.

Suppose  $\hat{Q}$  is a Hermitian operator with a continuous set of non-degenerate eigenvalues. In the examples we consider this set is usually just  $\mathbb{R}$  and so

$$\hat{Q}|q\rangle = q|q\rangle \quad \text{for} \quad q \in \mathbb{R}. \quad (1.86)$$

Just as in the discrete case it holds that  $\langle q|q'\rangle = 0$  if  $q \neq q'$ , but it is not so obvious what  $\langle q|q\rangle$  should be. We will fix this normalisation by requiring that

$$\hat{I} = \int dq |q\rangle\langle q| \quad (1.87)$$

is a valid resolution of the identity. It must then hold that

$$\hat{I}|q'\rangle = \int dq |q\rangle\langle q|q'\rangle = |q'\rangle \quad (1.88)$$

which clearly requires that

$$\langle q|q'\rangle = \delta(q - q'). \quad (1.89)$$

Continuous orthogonal basis states are therefore normalised to the Dirac delta function, and not the Kronecker delta function as in the discrete case. See section 1.9 for more information regarding the Dirac delta function.

The spectral representation of  $\hat{Q}$  is given by

$$\hat{Q} = \int dq q|q\rangle\langle q|. \quad (1.90)$$

Any state  $|\psi\rangle$  can be expanded in this basis as

$$|\psi\rangle = \int dq \langle q|\psi\rangle |q\rangle. \quad (1.91)$$

## 1.9 The Dirac delta function

### 1.9.1 Definition

The defining property of the Dirac delta function  $\delta(x)$  is that

$$\int_{-\infty}^{\infty} dx f(x)\delta(x) = f(0) \quad (1.92)$$

for any continuous function  $f(x)$ . In particular,  $\int_{-\infty}^{\infty} dx \delta(x) = 1$ . The Dirac delta function is often described informally as an “*infinitely narrow, infinitely high peak with unit area*”. Clearly this is not a function in the strict sense. There are a number of ways in which the Dirac delta function can be defined formally, but we will not need

this level of rigour here. It is sufficient to imagine  $\delta(x)$  as the result of a limit process. Let us consider an example. We define a sequence of block functions

$$d_n(x) = \begin{cases} n & \text{if } |x| \leq (2n)^{-1} \\ 0 & \text{otherwise} \end{cases} \quad (1.93)$$

for which  $\int_{-\infty}^{\infty} dx d_n(x) = 1$  for all  $n \geq 1$ . Note that this sequence is certainly not convergent. Now suppose  $f(x)$  is continuous and consider the limit of the sequence of integrals  $I_n = \int_{-\infty}^{\infty} dx f(x)d_n(x)$ :

$$\lim_{n \rightarrow \infty} I_n = \lim_{n \rightarrow \infty} \int_{-1/2n}^{+1/2n} dx f(x)n = f(0) \lim_{n \rightarrow \infty} \int_{-1/2n}^{+1/2n} dx n = f(0) \quad (1.94)$$

where the mean value theorem has been used. We can express this as

$$\lim_{n \rightarrow \infty} \int_{-\infty}^{\infty} dx f(x)d_n(x) = \int_{-\infty}^{\infty} dx f(x)\delta(x) \quad (1.95)$$

or by using the short hand

$$\lim_{n \rightarrow \infty} d_n(x) = \delta(x). \quad (1.96)$$

Of course, the equation above is not really a statement about the functions themselves since the limit does not exist and  $\delta(x)$  is not a regular function. Instead, it should be understood as a statement about the *integrals* of the two sides of the equation. In fact, all expressions involving Dirac delta functions should be interpreted in this way; they are shorthand for expressing the limit of some integral and not the integral of a (non-existent) limit.

The  $n$ -dimensional delta function is defined by

$$\delta(\vec{x}) = \delta(x_1)\delta(x_2)\cdots\delta(x_n) \quad (1.97)$$

where  $\vec{x} = (x_1, x_2, \dots, x_n)$ .

### 1.9.2 Properties of the Dirac delta function

The following properties are easy to prove using the definition in (1.92) and appropriate changes of variables.

1.  $\delta(x) = \delta(-x)$
2.  $\delta(ax) = \delta(x)/|a|$
3. The derivative of the delta function is defined by

$$\int_{-\infty}^{\infty} dx f(x)\delta'(x) = -f'(0) \quad (1.98)$$

To apply the limit process interpretation to this case a smoother version of  $d_n(x)$  is clearly required. One candidate is the Gaussian representation defined in the next section.

4. Let  $\{x_n\}$  denote the set of zeros of  $f(x)$ . Then

$$\delta(f(x)) = \sum_n \frac{\delta(x - x_n)}{|f'(x_n)|} \quad (1.99)$$

### 1.9.3 Representation of the Dirac delta function

There are numerous ways of representing the delta function as the limit of a function sequence or as an infinite series of functions. As explained in section (1.9.1) these expressions should really be understood as statements about integrals.

Some commonly used representations are:

1.  $\delta(x) = \lim_{n \rightarrow \infty} d_n(x)$  with  $d_n(x)$  the block functions defined in (1.93).

2.  $\delta(x) = \lim_{\epsilon \rightarrow 0} g_\epsilon(x)$  with  $g_\epsilon(x) = \frac{1}{\sqrt{\pi\epsilon^2}} e^{-x^2/\epsilon^2}$

3.  $\delta(x) = \lim_{a \rightarrow 0} \frac{a}{\pi(x^2 + a^2)}$

4. In terms of the step function

$$\Theta(x) = \begin{cases} 1 & x > 0 \\ 0 & x < 0 \end{cases} \quad (1.100)$$

we have  $\Theta(x) = \int_{-\infty}^x dy \delta(y)$  and  $d\Theta(x)/dx = \delta(x)$ .

5. If  $\mathcal{H}$  is a function space on the interval  $[a, b]$  with inner product  $(f, g) = \int_a^b dx f^*(x)g(x)$  and  $\{\phi_n(x)\}$  is an orthogonal basis for  $\mathcal{H}$  then

$$\delta(x - y) = \sum_n \phi_n(x) \phi_n^*(y). \quad (1.101)$$

This expression is called a **completeness relation**. This delta function satisfies

$$\int_a^b dx \delta(x - x') f(x) = f(x') \quad (1.102)$$

for  $f \in \mathcal{H}$  and  $x' \in (a, b)$ .

### Fourier series representation

For the function space of periodic functions with period  $L$  on  $[0, L]$  we have the orthogonal Fourier basis

$$\phi_n(x) = \frac{1}{\sqrt{L}} e^{2\pi i n x / L} \quad n \in \mathbb{Z}. \quad (1.103)$$

Inserting this into the completeness relation produces

$$\delta(x - y) = \frac{1}{L} \sum_{n=-\infty}^{+\infty} e^{2\pi i n (x-y) / L} \quad (1.104)$$

as a delta function representation for this function space.

## Fourier transform representation

Now consider the space of square integrable functions  $L^2(\mathbb{R})$ . It is well known that the set of Fourier modes

$$\phi_k(x) = \frac{1}{\sqrt{2\pi}} e^{ikx} \quad \text{with} \quad k \in \mathbb{R} \quad (1.105)$$

provides an orthogonal basis for this space. (In wave mechanics this is just the plane wave basis, i.e. eigenstates of the momentum operator.) The corresponding completeness relation is

$$\delta(x - y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} dk e^{ik(x-y)}. \quad (1.106)$$

*This is an extremely important result and should be memorized.* From this you can show that the basis elements are normalised such that  $(\phi_k, \phi_{k'}) = \delta(k - k')$ , which is just as we expect for a continuous orthogonal basis. For later reference, the  $n$ -dimensional version of (1.106) is

$$\delta(\vec{x} - \vec{y}) = \frac{1}{(2\pi)^n} \int d\vec{k} e^{i\vec{k}\cdot(\vec{x}-\vec{y})}. \quad (1.107)$$

It is enlightening to see precisely how (1.106) encodes the fact that the set  $\{\phi_k(x)\}$  constitutes an orthogonal basis for  $L^2(\mathbb{R})$ . Consider an arbitrary  $f \in L^2(\mathbb{R})$ . Using (1.106) we can express  $f(x)$  as

$$f(x) = \int_{-\infty}^{\infty} dy \delta(x - y) f(y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dk e^{ikx} \underbrace{\left[ \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dy f(y) e^{-iky} \right]}_{f(k)=(\phi_k,f)} \quad (1.108)$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dk e^{ikx} f(k) \quad (1.109)$$

Here  $f(k)$  is the Fourier transform of  $f(x)$ . Note that these are two completely different functions; they are distinguished only by the symbols we use for their arguments! The inverse transformation is

$$f(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dx e^{-ikx} f(x) \quad (1.110)$$

## 1.10 Commutator identities

**The following identities must be memorized.** Here  $a$ ,  $b$  and  $c$  are scalars.

1. Antisymmetry:  $[\hat{A}, \hat{B}] = -[\hat{B}, \hat{A}]$
2. Linearity:  $[\hat{A}, b\hat{B} + c\hat{C}] = b[\hat{A}, \hat{B}] + c[\hat{A}, \hat{C}]$  and  $[a\hat{A} + b\hat{B}, \hat{C}] = a[\hat{A}, \hat{C}] + b[\hat{B}, \hat{C}]$
3. Leibniz rule:  $[\hat{A}, \hat{B}\hat{C}] = \hat{B}[\hat{A}, \hat{C}] + [\hat{A}, \hat{B}]\hat{C}$  and  $[\hat{A}\hat{B}, \hat{C}] = \hat{A}[\hat{B}, \hat{C}] + [\hat{A}, \hat{C}]\hat{B}$
4. Jacobi identity:  $[\hat{A}, [\hat{B}, \hat{C}]] + [\hat{C}, [\hat{A}, \hat{B}]] + [\hat{B}, [\hat{C}, \hat{A}]] = 0$ .
5. If  $[\hat{A}, \hat{B}] = c\hat{I}$  with  $c$  a scalar then  $[\hat{A}, \hat{B}^n] = cn\hat{B}^{n-1}$  and  $[\hat{A}, f(\hat{B})] = cf'(\hat{B})$ .

6. The following is known as the Hadamard lemma, but it is also often called the Baker-Campbell-Hausdorff (BCH) formula:

$$e^{-\hat{A}} \hat{B} e^{\hat{A}} = \hat{B} + [\hat{B}, \hat{A}] + \frac{1}{2!} [[\hat{B}, \hat{A}], \hat{A}] + \frac{1}{3!} [[[[\hat{B}, \hat{A}], \hat{A}], \hat{A}] + \dots \quad (1.111)$$

The proof involves first using induction to show that

$$\frac{d^n}{dx^n} [e^{-x\hat{A}} \hat{B} e^{x\hat{A}}] = e^{-x\hat{A}} \underbrace{[\dots [[[\hat{B}, \hat{A}], \hat{A}], \hat{A}], \dots, \hat{A}]}_{n \text{ } \hat{A}\text{s in total}} e^{x\hat{A}}. \quad (1.112)$$

From this a Taylor series expansion for  $e^{-x\hat{A}} \hat{B} e^{x\hat{A}}$  in  $x$  can be found.

# Chapter 2

## Fundamental Postulates

In this chapter we introduce the postulates that underpin our conceptual framework of quantum mechanics. Two versions of these postulates will be considered. The first applies to the familiar case of a system in a definite state. The second, more general, set of postulates is based on the density matrix formalism which unifies statistical and quantum mechanics.

### 2.1 Fundamental Postulates I

**You should memorize everything in this section.**

Consider a system of which the true state is known. We then postulate the following:

1. *The state of a quantum system is represented by a normalised vector  $|\psi\rangle$  belonging to some Hilbert space  $\mathcal{H}$ .*
2. *To each classical observable  $A$ , there corresponds a Hermitian operator  $\hat{A}$  acting on  $\mathcal{H}$ .*
3. *A measurement of an observable  $A$  always yields an eigenvalue of  $\hat{A}$ .*
4. *If the system is in a state  $|\psi\rangle$  the probability that a measurement of  $A$  will yield a particular non-degenerate eigenvalue  $a_n$  is*

$$P(a_n) = |\langle a_n | \psi \rangle|^2$$

*where  $|a_n\rangle$  is the corresponding eigenstate. Immediately after the measurement the system will be in the  $|a_n\rangle$  state.*

**Remarks:**

1. According to the fourth postulate a measurement brings about an abrupt change in the state of the system. This process is irreversible and commonly referred to as the “collapse of the wave function”.
2. The fourth postulate also implies that measurements in quantum mechanics are inherently non-deterministic. Even complete knowledge of the system’s state does not generally allow us to predict the outcome of a single measurement. At best we can predict with what probability a certain outcome will occur or what the

average value of a repeated measurement will be. **Despite these limitations, quantum mechanics, and quantum field theory in particular, is by far the most accurate theory ever formulated in science.**

3. Expanding the state  $|\psi\rangle$  in the basis of  $\hat{A}$  eigenstates as  $|\psi\rangle = \sum_n c_n |a_n\rangle$  reveals that  $P(a_n) = |c_n|^2$ .
4. If  $|\psi\rangle$  is normalised then so is the probability distribution  $P(a_n)$ :

$$\sum_n P(a_n) = \sum_n |\langle a_n | \psi \rangle|^2 = \sum_n \langle \psi | a_n \rangle \langle a_n | \psi \rangle = \langle \psi | \psi \rangle = 1.$$

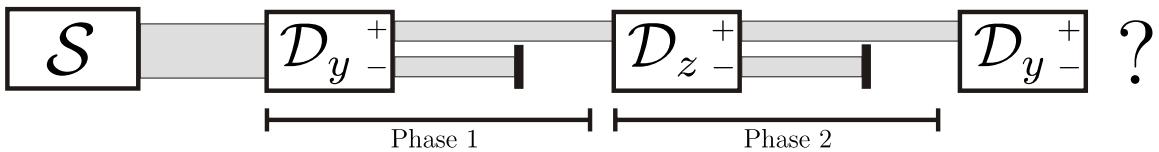
5. The expectation value of a measurement can be expressed as a diagonal matrix element:

$$\langle A \rangle = \sum_n P(a_n) a_n = \langle \psi | \sum_n a_n | a_n \rangle \langle a_n | \psi \rangle = \langle \psi | \hat{A} | \psi \rangle.$$

6. Two kets which differ only by a (global) phase represent exactly the same physical state, i.e. the physical content of  $|\psi\rangle$  and  $e^{i\theta}|\psi\rangle$  are identical.

## 2.2 Exercise: Repeated spin measurements

In the experiment shown below a source  $\mathcal{S}$  generates a beam of spin-1/2 particles all in the  $|x, +\rangle$  state. The beam first enters the detector  $\mathcal{D}_y$  which measures the  $S_y$  component of each particle's spin. Depending on the result of the measurement, either  $+\hbar/2$  or  $-\hbar/2$ , the particle will exit at the top or the bottom of the detector respectively. The bottom beam is blocked while the top beam enters a second detector  $\mathcal{D}_z$  which performs the same operation but with respect to the  $S_z$  spin component. The bottom beam is again blocked while the top beam enters the final  $\mathcal{D}_y$  detector. *The particles' spin do not undergo any time evolution between measurements.*



What do you make of the following statement?

*“In phase 1 all the particles with negative  $S_y$  components are removed from the beam. Similarly, in phase two all particles with negative  $S_z$  components are filtered out. Only particles with positive  $S_z$  and  $S_y$  therefore enter the final detector and so a measurement of  $S_y$  is sure to produce  $+\hbar/2$ . All the particles that enter the final detector will therefore exit at the top.”*

## 2.3 Classical uncertainty and quantum mechanics

To motivate the introduction of a second, more general, set of postulates we will consider an example which illustrates the distinction between regular statistical uncertainty

and quantum non-determinism.

Suppose a spin-1/2 particle has been prepared in such a way that it is either in the  $|+\rangle$  or  $|-\rangle$  state with equal probability. **We are ignorant about the true state of the particle.** In fact, imagine we have an ensemble of such particles all of which have been prepared in this way, but are otherwise independent of each other. A large number of measurements of the observables  $S_z$  and  $S_x$  are now performed on members of the ensemble. The average values of these measurements should be well approximated by the theoretical expectation values. We can calculate these expectation values as follows:

- The expectation value of  $S_z$  is 0 since  $S_z$  is either  $+1/2$  or  $-1/2$  with equal probability. (We set  $\hbar = 1$  in this section.)
- The expectation value of  $S_x$  is also 0 since

$$|\langle x, +|+\rangle|^2 = |\langle x, +|-\rangle|^2 = |\langle x, -|+\rangle|^2 = |\langle x, -|-\rangle|^2 = 1/2 \quad (2.1)$$

and so  $S_x$  is either  $+1/2$  or  $-1/2$  with equal probability.

Note that these calculations involve two different types of probabilities, one classical and the other quantum mechanical. We now ask the following: *How do we incorporate the statistical probabilities which reflect our ignorance about the true state of the particle into the formalism of quantum mechanics?*

One guess would be to assign to each particle the state

$$|\psi\rangle = \frac{1}{\sqrt{2}} (|+\rangle + |-\rangle).$$

This seems plausible at first, since the probability of obtaining  $S_z = +1/2$  in a measurement of  $S_z$  is indeed  $|\langle +|\psi\rangle|^2 = 1/2$ , and similar for  $S_z = -1/2$ . However, this state does not produce the correct expectation value of  $S_x$ :

$$\langle S_z \rangle = \langle \psi | \hat{S}_z | \psi \rangle = \frac{1}{4} [ \begin{array}{cc} 1 & 1 \end{array} ] \left[ \begin{array}{cc} 1 & 0 \\ 0 & -1 \end{array} \right] \left[ \begin{array}{c} 1 \\ 1 \end{array} \right] = 0 \quad (\text{Correct})$$

$$\langle S_x \rangle = \langle \psi | \hat{S}_x | \psi \rangle = \frac{1}{4} [ \begin{array}{cc} 1 & 1 \end{array} ] \left[ \begin{array}{cc} 0 & 1 \\ 1 & 0 \end{array} \right] \left[ \begin{array}{c} 1 \\ 1 \end{array} \right] = \frac{1}{2} \quad (\text{Wrong})$$

This mismatch implies that:

*“The particle is in either the  $|+\rangle$  or  $|-\rangle$  state with equal probability.”*

**is NOT the same statement as**

*“The particle is in the  $|\psi\rangle = \frac{1}{\sqrt{2}} (|+\rangle + |-\rangle)$  state.”*

In fact, these two statements describe completely different situations. In the first case the statistical probabilities reflect our uncertainty, i.e. lack of knowledge, regarding the true state of the system. This is just regular *statistical* mechanics. In the second case the state of the particle is known exactly. However, despite knowing “everything”, we

are still not able to predict the outcome of a single measurement. This is due to the probabilistic nature of quantum mechanical measurements. *Make sure you understand this point well.*

The density matrix formalism will combine the statistical probabilities that reflect our lack of knowledge regarding the true state of the system and the inherently probabilistic nature of quantum mechanical measurements into a single framework. This will form the basis of the second set of fundamental postulates.

## 2.4 The density matrix of an ensemble

### 2.4.1 Background

In this section we will construct the density matrix associated with a particular ensemble. This serves to motive the more abstract treatment of the density matrix in later sections where there is no mention of a particular ensemble.

Let us consider a system of which the true state is unknown and for which we can only attach probabilities to the different possible states. As in statistical mechanics we imagine an ensemble of identically prepared systems characterised by this probability distribution. A simple example is a beam of unpolarised photons emitted by a light source. To predict the polarization state of a single photon would require knowing its entire history and interactions with all the components of the source. This is clearly an impossible task. At best we can attempt to describe the beam *statistically* as a 50/50 mixture of horizontally and vertically polarized photons.

### 2.4.2 Construction of the density matrix

Consider a system which is known to be in a state  $|\phi_n\rangle$  with probability  $\rho_n$ . Here  $n$  is some label. The list  $\{(\rho_n, |\phi_n\rangle)\}$  of possible states and corresponding probabilities defines the ensemble. Note that, apart from normalization, we place no other restrictions on the possible states  $\{|\phi_n\rangle\}$ ; they need not be orthogonal or even linearly independent. Now consider an observable  $\hat{A}$  with eigenvalues and eigenstates  $\hat{A}|a_n\rangle = a_n|a_n\rangle$  and let

$$P_n^m = |\langle a_n | \phi_m \rangle|^2 \quad (2.2)$$

denote the probability of obtaining  $a_n$  in a measurement of  $A$  when the system is in the state  $|\phi_m\rangle$ . If we pick an element of the ensemble at random and measure  $A$  then the probability of obtaining  $a_n$  will be

$$P(a_n) = \sum_m \rho_m P_n^m = \sum_m \rho_m \langle a_n | \phi_m \rangle \langle \phi_m | a_n \rangle = \langle a_n | \left[ \sum_m \rho_m |\phi_m\rangle \langle \phi_m| \right] |a_n\rangle \quad (2.3)$$

$$= \langle a_n | \hat{\rho} | a_n \rangle \quad (2.4)$$

where we have introduced the density matrix as

$$\hat{\rho} = \sum_m \rho_m |\phi_m\rangle \langle \phi_m|. \quad (2.5)$$

Note that this is not necessarily the spectral representation of  $\hat{\rho}$ . (Why not?) The ensemble average of  $A$  can now be expressed in terms of  $\hat{\rho}$  as

$$\langle A \rangle_{EA} = \sum_n P(a_n) a_n = \sum_n a_n \langle a_n | \hat{\rho} | a_n \rangle = \sum_n \langle a_n | \hat{\rho} \hat{A} | a_n \rangle = \text{tr}(\hat{\rho} \hat{A}). \quad (2.6)$$

Clearly all the information regarding measurements on elements of the ensemble is encoded in the density matrix. It contains both the “statistical mechanics” probabilities  $\rho_m$  as well as the quantum mechanical probabilities  $P_n^m$ . The density matrix can therefore be regarded as describing the state of the ensemble itself, and is also often called the state operator.

The form of the density matrix in (2.5) implies that it has the following properties:

1.  $\hat{\rho}$  is Hermitian
2.  $\text{tr}(\hat{\rho}) = 1$
3.  $\hat{\rho}$  is positive semi-definite, i.e.  $\langle \psi | \hat{\rho} | \psi \rangle \geq 0$  for all  $|\psi\rangle \in \mathcal{H}$

In the present setting these properties are a direct result of our construction of  $\hat{\rho}$ . In the second set of postulates we will reverse this argument and postulate the existence of an operator with exactly these properties.

### 2.4.3 Spin-1/2 example (continued)

Let us now treat the example in section 2.3 using the density matrix formalism.

The statement

*“The particle is in either the  $|+\rangle$  or  $|-\rangle$  state with equal probability.”*

corresponds to the density matrix

$$\hat{\rho}_1 = \frac{1}{2}|+\rangle\langle+| + \frac{1}{2}|-\rangle\langle-| = \frac{1}{2} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad (2.7)$$

while the statement

*“The particle is in the  $|\psi\rangle = \frac{1}{\sqrt{2}}(|+\rangle + |-\rangle)$  state.”*

corresponds to the density matrix

$$\hat{\rho}_2 = |\psi\rangle\langle\psi| = \frac{1}{2}|+\rangle\langle+| + \frac{1}{2}|-\rangle\langle-| + \frac{1}{2}|-\rangle\langle+| + \frac{1}{2}|+\rangle\langle-| = \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \quad (2.8)$$

The difference between these two cases is now apparent. Calculating the ensemble averages we find

$$\langle S_z \rangle_1 = \text{tr}(\hat{\rho}_1 \hat{S}_z) = 0 \quad \langle S_x \rangle_1 = \text{tr}(\hat{\rho}_1 \hat{S}_x) = 0 \quad (2.9)$$

$$\langle S_z \rangle_2 = \text{tr}(\hat{\rho}_2 \hat{S}_z) = 0 \quad \langle S_x \rangle_2 = \text{tr}(\hat{\rho}_2 \hat{S}_x) = \frac{1}{2} \quad (2.10)$$

which agrees with our previous results.

#### 2.4.4 Non-unique representations

We have seen that with a particular ensemble  $\{(\rho_n, |\phi_n\rangle)\}$  there is associated a unique density matrix  $\hat{\rho} = \sum_n \rho_n |\phi_n\rangle\langle\phi_n|$ . The reverse is not true: in general there are infinitely many different ensembles associated with the same density matrix. The reason for this is that the representation of  $\hat{\rho}$  as a convex combination<sup>1</sup> of outer products is generally *not* unique; i.e.  $\hat{\rho}$  can also be written as  $\hat{\rho} = \sum_n \tilde{\rho}_n |\tilde{\phi}_n\rangle\langle\tilde{\phi}_n|$  with  $\tilde{\rho}_n$  and  $|\tilde{\phi}_n\rangle$  different from  $\rho_n$  and  $|\phi_n\rangle$ . For example, consider a beam of electrons with random spin polarizations. If we view this beam as a 50/50 mix of electrons with up and down spins in the  $z$ -direction we obtain the density matrix

$$\hat{\rho}_z = \frac{1}{2}|+\rangle\langle+| + \frac{1}{2}|-\rangle\langle-| = \frac{1}{2} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}. \quad (2.11)$$

If instead we treat the beam as a 50/50 mix of electrons with up and down spins in the  $y$ -direction we obtain

$$\hat{\rho}_y = \frac{1}{2}|y,+\rangle\langle y,+| + \frac{1}{2}|y,-\rangle\langle y,-| = \frac{1}{2}\hat{I} = \frac{1}{2} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad (2.12)$$

which is identical to  $\hat{\rho}_z$ . These two ensembles therefore correspond to the same density matrix. Since the density matrix encodes *all* the information relevant to actual physical measurements these two ensembles are completely indistinguishable from each other. In fact, it is now clear that the density matrix is actually the more fundamental object since the descriptions of the ensembles in terms of  $\{(1/2, |+\rangle), (1/2, |-\rangle)\}$  and  $\{(1/2, |y,+\rangle), (1/2, |y,-\rangle)\}$  clearly contain information which is not physically relevant and cannot be probed by any experiment. This suggests that we should reformulate our fundamental postulates to make the density matrix the central focus, without any reference to a set of possible states and associated probabilities.

## 2.5 Fundamental Postulates II

We will now use the intuition gained in the previous sections to formulate a more general set of postulates in which the existence of the density matrix and its properties is taken as a basic assumption:

1. *To each classical observable  $A$  there corresponds a Hermitian operator  $\hat{A}$  acting on a Hilbert space  $\mathcal{H}$ . A measurement of  $A$  always yields an eigenvalue of  $\hat{A}$ . Following the measurement the system is in the corresponding eigenstate of  $\hat{A}$ .*
2. *To each state of the system (or ensemble) there corresponds a unique state operator  $\hat{\rho}$  called the density matrix. It has the following properties:*
  - (a) *It is Hermitian:  $\hat{\rho}^\dagger = \hat{\rho}$ .*
  - (b) *It is positive semi-definite:  $\langle\psi|\hat{\rho}|\psi\rangle \geq 0$  for all  $|\psi\rangle \in \mathcal{H}$ .*
  - (c)  *$\text{tr}(\hat{\rho}) = 1$ .*

---

<sup>1</sup>A linear combination with real, non-negative coefficients which add up to one.

3. For a system, or ensemble of systems, described by  $\hat{\rho}$  the probability of obtaining a particular non-degenerate eigenvalue  $a_n$  of  $\hat{A}$  in a measurement of  $A$  is

$$P(a_n) = \langle a_n | \hat{\rho} | a_n \rangle$$

where  $|a_n\rangle$  is the corresponding eigenstate. The ensemble average of  $A$  is  $\langle A \rangle_{EA} = \text{tr}(\hat{\rho}\hat{A})$ .

Note that the set of all density matrices is convex, i.e. any convex combination of density matrices is again a density matrix.

**Exercise:** Prove that  $P(a_n) = \langle a_n | \hat{\rho} | a_n \rangle$  has all the properties required of a probability distribution.

### 2.5.1 The eigenvalues of $\hat{\rho}$

Since  $\hat{\rho}$  is Hermitian it has a spectral representation of the form

$$\hat{\rho} = \sum_n \rho_n |n\rangle \langle n| \quad (2.13)$$

where  $\hat{\rho}|n\rangle = \rho_n |n\rangle$ . It is straightforward to show that the eigenvalues satisfy

$$\rho_n \in \mathbb{R} \quad \rho_n \geq 0 \quad \sum_n \rho_n = 1 \quad 0 \leq \rho_n \leq 1. \quad (2.14)$$

## 2.6 Pure and mixed states

In this section we investigate the two fundamental classes of density matrices.

### 2.6.1 Pure states

A system (or ensemble) is in a pure state if the density matrix is a single outer product:

$$\hat{\rho} = |\psi\rangle \langle \psi|. \quad (2.15)$$

Two important results regarding pure states are:

**Theorem:**  $\hat{\rho}$  is a pure state if and only if  $\hat{\rho}^2 = \hat{\rho}$ .

**Theorem:** A pure state can never be written as a non-trivial convex combination of other pure states, i.e. its representation is unique.

The second theorem allows us to unambiguously associate a specific state in the Hilbert space with a pure state. If  $\hat{\rho} = |\psi\rangle \langle \psi|$  we know with certainty that the system is in the  $|\psi\rangle$  state<sup>2</sup>. In the ensemble picture this implies that each member of the ensemble is in the state  $|\psi\rangle$  with a probability of one. Since the representation of  $\hat{\rho}$  is unique the

---

<sup>2</sup>Since  $|\psi\rangle$  and  $e^{i\theta}|\psi\rangle$  correspond to the same density matrix  $\hat{\rho} = |\psi\rangle \langle \psi|$  this mapping is only one-to-one up to a phase.

ambiguity discussed in section 2.4.4 is absent here. For a pure state the expressions for probabilities and ensemble averages reduce to those of the first set of postulates:

$$P(a_n) = \langle a_n | \hat{\rho} | a_n \rangle = \langle a_n | \psi \rangle \langle \psi | a_n \rangle = |\langle a_n | \psi \rangle|^2 \quad (2.16)$$

$$\langle A \rangle_{EA} = \text{tr}(\hat{A}\hat{\rho}) = \sum_n \langle \phi_n | \hat{A}\hat{\rho} | \phi_n \rangle = \sum_n \langle \phi_n | \hat{A} | \psi \rangle \langle \psi | \phi_n \rangle = \langle \psi | \hat{A} | \psi \rangle \quad (2.17)$$

Here  $\{|\phi_n\rangle\}$  is an arbitrary orthogonal basis. If the system is in a pure state we are clearly free to describe it using either the density matrix  $|\psi\rangle\langle\psi|$  or the ket  $|\psi\rangle$ . This choice then dictates which set of postulates are to be used for calculations.

### 2.6.2 Mixed states

Mixed states are non-trivial convex combinations of pure states. In other words, the convex combination  $\hat{\rho} = \sum_n c_n |\phi_n\rangle\langle\phi_n|$  will be a mixed state unless all but one of the  $c_n$  coefficients are zero. A mixed state can therefore never be expressed as a single outer product. In fact, the representation of a mixed state as a convex combination of pure states is **never** unique. This is precisely the issue highlighted in section 2.4.4.

## 2.7 The von Neumann entropy

The von Neumann entropy of a density matrix  $\hat{\rho}$  is defined as

$$S = -\text{tr}(\hat{\rho} \ln(\hat{\rho})). \quad (2.18)$$

Roughly speaking, it is an indication of our level of uncertainty regarding the true state of the system. This can be seen by considering the density matrices which minimize or maximise the entropy. First we express  $S$  in terms of the eigenvalues of  $\hat{\rho}$  as

$$S = -\sum_{n=1}^D \rho_n \ln(\rho_n) \quad (2.19)$$

where  $D$  is the dimension of the Hilbert space. We take  $D$  to be finite. By convention  $0 \ln(0) \equiv \lim_{x \rightarrow 0} x \ln(x) = 0$ . The lowest possible value of  $S$  is clearly zero, and this occurs when one of the eigenvalues equals one and the rest are zero. In other words, the entropy is at a minimum when the system is in a pure state; exactly as we would expect.

Now consider the opposite limit. By using Lagrange multipliers to maximise  $S$  under the constraints in (2.14) it can be shown that  $S$  is maximal when all the eigenvalues are equal. This implies that the density matrix is simply  $D^{-1}\hat{I}$ , i.e. a multiple of the identity operator. To understand the consequences of this, consider a measurement of an observable  $A$  for which  $\hat{A}|a_n\rangle = a_n|a_n\rangle$ . The probability of obtaining  $a_n$  is then  $P(a_n) = \langle a_n | D^{-1}\hat{I} | a_n \rangle = D^{-1}$ . This is just a uniform distribution where all the possible outcomes are equally likely. This clearly corresponds to a state of maximum uncertainty.

## 2.8 Density matrix of the canonical ensemble

In statistical mechanics you encountered the canonical ensemble which describes a system at a fixed temperature, volume and number of particles. If the system's Hamiltonian is  $\hat{H}$  with  $\hat{H}|\phi_n\rangle = E_n|\phi_n\rangle$  the probability of finding the system in the  $|\phi_n\rangle$  state is

$$P(\phi_n) = \frac{1}{Z} e^{-E_n/kT} \quad \text{with} \quad Z = \sum_n e^{-E_n/kT}. \quad (2.20)$$

The density matrix describing this ensemble is therefore

$$\hat{\rho} = \sum_n P(\phi_n)|\phi_n\rangle\langle\phi_n| = \frac{1}{Z} \sum_n e^{-E_n/kT} |\phi_n\rangle\langle\phi_n| = \frac{1}{Z} e^{-\hat{H}/kT}. \quad (2.21)$$

Since  $\text{tr}(\hat{\rho}) = 1$  we see that the partition function is just

$$Z = \text{tr}(e^{-\hat{H}/kT}). \quad (2.22)$$

Note that this expression is basis-independent, i.e. we can calculate  $Z$  without knowing the eigenvalues of the Hamiltonian. There exists sophisticated methods for calculating this trace.

# Chapter 3

## Recovering Wave Mechanics

In this chapter we will demonstrate how the familiar wave mechanics picture of quantum mechanics is contained within the more general abstract formalism developed in the preceding two chapters. We will also encounter the first example of observables acting as generators of transformations on the Hilbert space.

### 3.1 The position operator

We associate with a point particle a position operator

$$\hat{\vec{x}} = (\hat{x}_1, \hat{x}_2, \hat{x}_3) = (\hat{x}, \hat{y}, \hat{z}) \quad (3.1)$$

which has three Hermitian components corresponding to the particle's Cartesian coordinates. We postulate that these components commute with each other as

$$[\hat{x}_i, \hat{x}_j] = 0 \quad i, j = 1, 2, 3. \quad (3.2)$$

This allows us to construct a basis of simultaneous eigenstates of  $\hat{x}_1$ ,  $\hat{x}_2$  and  $\hat{x}_3$ . Such an eigenstate describes a particle with a well-defined position, i.e. that is localised at a particular point in space. These states will be labelled by their eigenvalues:

$$\hat{x}_i |\vec{x}\rangle = x_i |\vec{x}\rangle \quad \text{or} \quad \hat{\vec{x}} |\vec{x}\rangle = \vec{x} |\vec{x}\rangle \quad \text{with } \vec{x} \in \mathbb{R}^3. \quad (3.3)$$

These eigenstates provide a continuous orthogonal basis for the Hilbert space and so we have

$$\langle \vec{x}' | \vec{x} \rangle = \delta(\vec{x}' - \vec{x}) \quad \text{and} \quad \int d\vec{x} |\vec{x}\rangle \langle \vec{x}| = \hat{I}. \quad (3.4)$$

Note that the eigenstates of  $\hat{\vec{x}}$ , like those of  $\hat{\vec{p}}$ , are not normalisable. However, we can still use these states as a basis for constructing proper normalisable states.

### 3.2 Momentum and spatial translations

Next we consider the operators corresponding to a particle's linear momentum. Of central importance in this regard are the **commutation relations** between the position and momentum operators. We will begin by investigating how **spatial translations** are realised on the Hilbert space. This might seem like a peculiar starting point, but

as we will see there exists a fundamental connection between physical observables and transformations of space and time. *In particular, the commutation relations between physical observables are determined by the algebraic properties of these transformations.* By studying spatial translations we will gain knowledge about the momentum operators, just as studying rotations will teach us something about angular momentum.

In quantum mechanics spatial translations (as well as rotations and time evolution) are represented by unitary operators which act on the Hilbert space. The operator  $\hat{T}(\Delta\vec{x})$  corresponding to a translation by  $\Delta\vec{x}$  is defined by

$$\hat{T}(\Delta\vec{x})|\vec{x}\rangle = |\vec{x} + \Delta\vec{x}\rangle \quad (3.5)$$

where  $\vec{x}, \Delta\vec{x} \in \mathbb{R}^3$ . It follows that  $\hat{T}(\Delta\vec{x})$  satisfies:

1.  $\hat{T}^\dagger(\Delta\vec{x})\hat{T}(\Delta\vec{x}) = \hat{I}$  (Unitarity)
2.  $\hat{T}(\Delta\vec{x})\hat{T}(\Delta\vec{y}) = \hat{T}(\Delta\vec{x} + \Delta\vec{y})$  (Composition rule)
3.  $\hat{T}(\vec{0}) = \hat{I}$
4.  $\hat{T}^\dagger(\Delta\vec{x}) = \hat{T}^{-1}(\Delta\vec{x}) = \hat{T}(-\Delta\vec{x})$

Note that the set of all translation operators, just like the set of translations itself, is an Abelian group. A very useful trick for studying these operators is to first consider an *infinitesimal* translation by  $\Delta\vec{x}$ . Expanding  $\hat{T}(\Delta\vec{x})$  to linear order yields

$$\hat{T}(\Delta\vec{x}) = \hat{T}(\vec{0}) + \vec{\nabla}\hat{T}(\vec{x})\Big|_{\vec{x}=0} \cdot \Delta\vec{x} = \hat{I} + \hat{\vec{G}} \cdot \Delta\vec{x} \quad (3.6)$$

where  $\hat{\vec{G}} = (\hat{G}_1, \hat{G}_2, \hat{G}_3)$ . Since  $\hat{T}^\dagger(\Delta\vec{x}) = \hat{T}(-\Delta\vec{x})$  each  $\hat{G}_i$  must be anti-Hermitian. What does  $\hat{\vec{G}}$  represent physically? It is clearly closely related to the position operator, since it is responsible for generating infinitesimal changes in the particle's position. What follows next can be motivated in a number of ways, but we will not spend time doing so here. Instead, we will accept this as a fundamental postulate and investigate its consequences.

**Postulate:**  $\hat{\vec{G}}$  is proportional to the momentum operator  $\hat{\vec{p}} = (\hat{p}_1, \hat{p}_2, \hat{p}_3) = (\hat{p}_x, \hat{p}_y, \hat{p}_z)$ . In particular,  $\hat{\vec{G}} = -i\hat{\vec{p}}/\hbar$ . We say that the momentum operators are the **generators of translations** on the Hilbert space.

For infinitesimal  $\Delta\vec{x}$  we therefore have

$$\hat{T}(\Delta\vec{x}) = \hat{I} - i\hat{\vec{p}} \cdot \Delta\vec{x}/\hbar. \quad (3.7)$$

Note that we have not made any assumptions regarding commutation relations involving the momentum operators. It will soon be clear that all such commutation relations follow directly from the algebraic properties of the translation operators. This is a very powerful idea, and one which will form the basis of our study of angular momentum in chapter 5.

### 3.3 Commutation relations

We first consider the commutator of the position and translation operators. Applying these operators to a position eigenstate in different orders reveals that

$$\hat{\vec{x}} \hat{T}(\Delta\vec{x}) |\vec{x}\rangle = \hat{\vec{x}} |\vec{x} + \Delta\vec{x}\rangle = (\vec{x} + \Delta\vec{x}) |\vec{x} + \Delta\vec{x}\rangle = (\vec{x} + \Delta\vec{x}) \hat{T}(\Delta\vec{x}) |\vec{x}\rangle \quad (3.8)$$

$$\hat{T}(\Delta\vec{x}) \hat{\vec{x}} |\vec{x}\rangle = \vec{x} \hat{T}(\Delta\vec{x}) |\vec{x}\rangle \quad (3.9)$$

and therefore

$$[\hat{\vec{x}}, \hat{T}(\Delta\vec{x})] |\vec{x}\rangle = \Delta\vec{x} \hat{T}(\Delta\vec{x}) |\vec{x}\rangle. \quad (3.10)$$

This *vector (ket) equality* holds for any choice of  $|\vec{x}\rangle$ , and since these states form a basis the expression above implies the *operator equality*

$$[\hat{\vec{x}}, \hat{T}(\Delta\vec{x})] = \Delta\vec{x} \hat{T}(\Delta\vec{x}). \quad (3.11)$$

Now suppose  $\Delta\vec{x}$  is infinitesimal and expand both sides of this equation up to linear order. After some simplifications we find

$$\sum_{l=1}^3 [\hat{x}_j, \hat{p}_l] \Delta x_l = i\hbar \Delta x_j \quad j = 1, 2, 3 \quad (3.12)$$

which must hold for *any* infinitesimal displacement  $\Delta\vec{x}$ . This is only possible if

$$[\hat{x}_j, \hat{p}_l] = i\hbar \delta_{j,l} \quad j, l = 1, 2, 3 \quad (3.13)$$

which is exactly what we would have expected.

Now consider the commutation relations of the momentum operators amongst themselves. First note that any two spatial translations commute:

$$\hat{T}(\Delta\vec{x}) \hat{T}(\Delta\vec{y}) = \hat{T}(\Delta\vec{x} + \Delta\vec{y}) = \hat{T}(\Delta\vec{y}) \hat{T}(\Delta\vec{x}). \quad (3.14)$$

Plugging in the infinitesimal form  $\hat{T}(\Delta\vec{x}) = \hat{I} - i\hat{\vec{p}} \cdot \Delta\vec{x}/\hbar$  quickly reveals that this requires the generators to commute as well:

$$[\hat{p}_i, \hat{p}_j] = 0 \quad i, j = 1, 2, 3. \quad (3.15)$$

In summary, the **canonical commutation relations** are:

$$[\hat{p}_j, \hat{p}_k] = 0 \quad (3.16)$$

$$[\hat{x}_j, \hat{x}_k] = 0 \quad (3.17)$$

$$[\hat{x}_j, \hat{p}_k] = i\hbar \delta_{j,k} \quad (3.18)$$

where  $j, k = 1, 2, 3$ .

## 3.4 Finite translations

For infinitesimal  $\Delta\vec{x}$  the translation operator is given by  $\hat{T}(\Delta\vec{x}) = \hat{I} - i\hat{\vec{p}} \cdot \Delta\vec{x}/\hbar$ . We now want to derive an expression which is also valid for finite translations. The main idea here is that a finite translation can be regarded as the composition of infinitely many, infinitesimal translations. To be precise, suppose  $\vec{a}$  is a finite translation and let  $\Delta\vec{x} = \vec{a}/N$  with  $N$  an integer; i.e. we are dividing  $\vec{a}$  into  $N$  steps. It follows that

$$|\vec{x} + \vec{a}\rangle = \lim_{N \rightarrow \infty} [\hat{T}(\Delta\vec{x})]^N |\vec{x}\rangle = \lim_{N \rightarrow \infty} \left[ \hat{I} - \frac{i\vec{a} \cdot \hat{\vec{p}}}{N\hbar} \right]^N |\vec{x}\rangle = e^{-i\vec{a} \cdot \hat{\vec{p}}/\hbar} |\vec{x}\rangle. \quad (3.19)$$

The translation operator is therefore the *exponential of a linear combination of the generators*:

$$\hat{T}(\vec{a}) = e^{-i\vec{a} \cdot \hat{\vec{p}}/\hbar} \quad (3.20)$$

We will encounter other examples of this form in later chapters.

It is interesting to note that this result can also be derived directly from the commutation relations in (3.18). To do so we need to show that

$$\hat{x}_i e^{-i\vec{a} \cdot \hat{\vec{p}}/\hbar} |\vec{x}\rangle = (x_i + a_i) e^{-i\vec{a} \cdot \hat{\vec{p}}/\hbar} |\vec{x}\rangle \quad (3.21)$$

since this implies  $e^{-i\vec{a} \cdot \hat{\vec{p}}/\hbar} |\vec{x}\rangle = |\vec{x} + \vec{a}\rangle$ . We can prove (3.21) as follows:

$$\hat{x}_i e^{-i\vec{a} \cdot \hat{\vec{p}}/\hbar} |\vec{x}\rangle = \left[ e^{-i\vec{a} \cdot \hat{\vec{p}}/\hbar} \hat{x}_i + [\hat{x}_i, e^{-i\vec{a} \cdot \hat{\vec{p}}/\hbar}] \right] |\vec{x}\rangle \quad (3.22)$$

$$= \left[ e^{-i\vec{a} \cdot \hat{\vec{p}}/\hbar} x_i + i\hbar \frac{\partial e^{-i\vec{a} \cdot \hat{\vec{p}}/\hbar}}{\partial \hat{p}_i} \right] |\vec{x}\rangle \quad (3.23)$$

$$= \left[ e^{-i\vec{a} \cdot \hat{\vec{p}}/\hbar} x_i + e^{-i\vec{a} \cdot \hat{\vec{p}}/\hbar} a_i \right] |\vec{x}\rangle \quad (3.24)$$

$$= (x_i + a_i) e^{-i\vec{a} \cdot \hat{\vec{p}}/\hbar} |\vec{x}\rangle \quad (3.25)$$

*Make sure you understand every step of this derivation.*

## 3.5 The coordinate representation

Let us first recall some basic facts regarding the matrix representation introduced in section 1.3.2. This construction involved choosing a *discrete* orthogonal basis  $\{|\phi_n\rangle\}$  for the Hilbert space and then representing states by the column vectors formed using their expansion coefficients in this basis. Operators are represented by matrices acting on these column vectors. We will now repeat this procedure using the *continuous* basis  $\{|\vec{x}\rangle : \vec{x} \in \mathbb{R}^3\}$  of position eigenstates.<sup>1</sup>

---

<sup>1</sup>We are ignoring internal degrees of freedom for the moment. See section 5.11 for a generalization which incorporates spin.

### 3.5.1 States in the coordinate representation

Using the resolution of the identity  $\int d\vec{x}|\vec{x}\rangle\langle\vec{x}| = \hat{I}$  we can express any state  $|\psi\rangle$  as

$$|\psi\rangle = \hat{I}|\psi\rangle = \int d\vec{x}|\vec{x}\rangle\langle\vec{x}|\psi\rangle \quad (3.26)$$

where  $\langle\vec{x}|\psi\rangle$  plays the role of the expansion coefficient. Since  $\langle\vec{x}|\psi\rangle$  now depends on the continuous label  $\vec{x}$  it should be treated as a complex valued function on  $\mathbb{R}^3$ . *In fact, it is nothing but the wave function of the state  $|\psi\rangle$  in the coordinate representation:*

$$\psi(\vec{x}) \equiv \langle\vec{x}|\psi\rangle. \quad (3.27)$$

We call this the coordinate space (or real space) wave function of  $|\psi\rangle$  and attach to  $|\psi(\vec{x})|^2$  the usual probabilistic interpretation of the probability density associated with position measurements. To be clear, the probability of finding the particle in some volume  $\mathcal{V} \subseteq \mathbb{R}^3$  is

$$P(\mathcal{V}) = \int_{\mathcal{V}} d\vec{x} |\psi(\vec{x})|^2. \quad (3.28)$$

All the quantities we encountered in the bra-ket formalism now have familiar coordinate space counterparts. The inner product between two states  $|\phi\rangle$  and  $|\psi\rangle$  can be expressed in terms of their wave functions as

$$\langle\phi|\psi\rangle = \int d\vec{x} \langle\phi|\vec{x}\rangle\langle\vec{x}|\psi\rangle = \int d\vec{x} \phi^*(\vec{x})\psi(\vec{x}) \quad (3.29)$$

which is just the regular  $L^2(\mathbb{R}^3)$  inner product. Matrix elements can be written as

$$\langle\phi|\hat{A}|\psi\rangle = \int d\vec{x} d\vec{x}' \phi^*(\vec{x})\langle\vec{x}|\hat{A}|\vec{x}'\rangle\psi(\vec{x}'). \quad (3.30)$$

If  $\hat{A} = A(\hat{\vec{x}})$  then  $\langle\vec{x}|\hat{A}|\vec{x}'\rangle = A(\vec{x})\langle\vec{x}|\vec{x}'\rangle = A(\vec{x})\delta(\vec{x} - \vec{x}')$  and we can simplify the expression above to the more familiar form

$$\langle\phi|\hat{A}|\psi\rangle = \int d\vec{x} \phi^*(\vec{x})A(\vec{x})\psi(\vec{x}). \quad (3.31)$$

### 3.5.2 Operators in the coordinate representation

In this representation linear operators become *differential operators* which act on the wave functions. We define the coordinate space representation of an operator  $\hat{A}$  by

$$\hat{A}\psi(\vec{x}) \equiv \langle\vec{x}|\hat{A}|\psi\rangle. \quad (3.32)$$

For the position operator we have

$$\hat{x}_i\psi(\vec{x}) = \langle\vec{x}|\hat{x}_i|\psi\rangle = \langle\vec{x}|x_i|\psi\rangle = x_i\psi(\vec{x}). \quad (3.33)$$

The *operator*  $\hat{x}_i$  therefore acts on a wave function by multiplying it by the *variable*  $x_i$ . *Do not confuse these two symbols.* The function  $f(x) = 7x$  and the number 7 are two different things!

To calculate  $\langle \vec{x} | \hat{\vec{p}} | \psi \rangle$  we first use the translation operator to express  $|\vec{x}\rangle$  as

$$e^{-i\vec{x} \cdot \hat{\vec{p}}/\hbar} |\vec{0}\rangle = |\vec{x}\rangle. \quad (3.34)$$

It follows that

$$\langle \vec{x} | \hat{\vec{p}} | \psi \rangle = \langle \vec{0} | e^{+i\vec{x} \cdot \hat{\vec{p}}/\hbar} \hat{\vec{p}} | \psi \rangle = -i\hbar \vec{\nabla} \langle \vec{0} | e^{+i\vec{x} \cdot \hat{\vec{p}}/\hbar} | \psi \rangle = -i\hbar \vec{\nabla} \langle \vec{x} | \psi \rangle = -i\hbar \vec{\nabla} \psi(\vec{x}) \quad (3.35)$$

and so

$$\hat{\vec{p}} = -i\hbar \vec{\nabla} = -i\hbar \left( \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right). \quad (3.36)$$

We also see that

$$\langle \phi | A(\hat{\vec{p}}) | \psi \rangle = \int d\vec{x} \phi^*(\vec{x}) A(-i\hbar \vec{\nabla}) \psi(\vec{x}). \quad (3.37)$$

Finally, let us check that the translation operator  $\hat{T}(\vec{a})$  acts on  $\psi(\vec{x})$  in the expected way. We have that

$$\hat{T}(\vec{a}) \psi(\vec{x}) = \langle \vec{x} | \hat{T}(\vec{a}) | \psi \rangle = \langle \vec{x} - \vec{a} | \psi \rangle = \psi(\vec{x} - \vec{a}) \quad (3.38)$$

where the unitarity of  $\hat{T}(\vec{a})$  has been used. This is the expected result: shifting the argument of  $\psi$  by  $-\vec{a}$  indeed translates the wave function in space by  $+\vec{a}$ .

### 3.5.3 The time-independent Schrödinger equation

Let us use these results to find the coordinate space representation of the eigenvalue equation  $\hat{H}|\psi\rangle = E|\psi\rangle$  for a Hamiltonian of the form

$$\hat{H} = \frac{\hat{\vec{p}}^2}{2m} + V(\hat{\vec{x}}). \quad (3.39)$$

All this requires is the matrix element  $\langle \vec{x} | \hat{H} | \psi \rangle$  which follows directly from

$$\langle \vec{x} | \hat{\vec{p}}^2 | \psi \rangle = -\hbar^2 \vec{\nabla}^2 \psi(\vec{x}) \quad \text{and} \quad \langle \vec{x} | V(\hat{\vec{x}}) | \psi \rangle = V(\vec{x}) \psi(\vec{x}). \quad (3.40)$$

Applying  $\langle \vec{x} |$  to both sides of  $\hat{H}|\psi\rangle = E|\psi\rangle$  then produces

$$\left[ -\frac{\hbar^2}{2m} \vec{\nabla}^2 + V(\vec{x}) \right] \psi(\vec{x}) = E \psi(\vec{x}) \quad (3.41)$$

which is the familiar time-independent Schrödinger equation.

## 3.6 The momentum representation

### 3.6.1 Definition

It should be clear that there is a wave function representation associated with any choice of continuous orthogonal basis for the Hilbert space. Consider the basis of momentum eigenstates which satisfy

$$\hat{\vec{p}}|\vec{p}\rangle = \vec{p}|\vec{p}\rangle \quad \langle \vec{p}'|\vec{p}\rangle = \delta(\vec{p} - \vec{p}') \quad \int d\vec{p} |\vec{p}\rangle \langle \vec{p}| = \hat{I}. \quad (3.42)$$

The momentum space wave function of  $|\psi\rangle$  is  $\psi(\vec{p}) = \langle \vec{p} | \psi \rangle$  and the operators  $\hat{p}_i$  and  $\hat{x}_i$  are represented by

$$\hat{p}_i \psi(\vec{p}) = p_i \psi(\vec{p}) \quad \text{and} \quad \hat{x}_i \psi(\vec{p}) = +i\hbar \frac{\partial}{\partial p_i} \psi(\vec{p}). \quad (3.43)$$

Note that in this case  $\hat{p}_i$  amounts to multiplication while  $\hat{x}_i$  is a derivative.

**Exercise:** Derive (3.43).

### 3.6.2 Transforming between coordinate and momentum space wave functions

The wave functions  $\psi(\vec{x})$  and  $\psi(\vec{p})$  represent the same state in two different bases and are therefore related by a basis transformation. To find this transformation we first calculate  $\langle \vec{x} | \vec{p} \rangle$ , i.e. the coordinate space wave function of a momentum eigenstate. In coordinate space the defining eigenvalue equation  $\hat{p} |\vec{p}\rangle = \vec{p} |\vec{p}\rangle$  becomes

$$-i\hbar \vec{\nabla}_x \langle \vec{x} | \vec{p} \rangle = \vec{p} \langle \vec{x} | \vec{p} \rangle \quad (3.44)$$

which has solution

$$\langle \vec{x} | \vec{p} \rangle = c(\vec{p}) e^{i\vec{x} \cdot \vec{p} / \hbar}. \quad (3.45)$$

The function  $c(\vec{p})$  is fixed by the normalisation condition of the momentum eigenstates as

$$\delta(\vec{p} - \vec{p}') = \langle \vec{p}' | \vec{p} \rangle = \int d\vec{x} \langle \vec{p}' | \vec{x} \rangle \langle \vec{x} | \vec{p} \rangle = c^*(\vec{p}') c(\vec{p}) \int d\vec{x} e^{i\vec{x} \cdot (\vec{p} - \vec{p}') / \hbar} \quad (3.46)$$

$$= c^*(\vec{p}') c(\vec{p}) (2\pi\hbar)^n \delta(\vec{p} - \vec{p}') \quad (3.47)$$

where the Fourier representation of the Dirac delta function from (1.107) has been used. Here  $n$  is the number of spatial dimensions. It follows that  $c^*(\vec{p}') c(\vec{p}) (2\pi\hbar)^n = 1$  and we choose  $c(\vec{p}) = (2\pi\hbar)^{-n/2}$ . The coordinate space wave function of a momentum eigenstate is therefore

$$\langle \vec{x} | \vec{p} \rangle = \frac{1}{(2\pi\hbar)^{n/2}} e^{i\vec{x} \cdot \vec{p} / \hbar}. \quad (3.48)$$

This is just an ordinary plane wave; exactly what we expect from wave mechanics.

We can now establish the connection between  $\psi(\vec{x})$  and  $\psi(\vec{p})$ . Using the resolution of the identity in terms of momentum eigenstates we see that

$$\psi(\vec{x}) = \langle \vec{x} | \psi \rangle = \int d\vec{p} \langle \vec{x} | \vec{p} \rangle \langle \vec{p} | \psi \rangle = \int d\vec{p} \langle \vec{x} | \vec{p} \rangle \psi(\vec{p}) \quad (3.49)$$

and so the two wave functions are related by an integral transform with kernel  $\langle \vec{x} | \vec{p} \rangle$ . The kernel of the inverse transformation is  $\langle \vec{p} | \vec{x} \rangle = \langle \vec{x} | \vec{p} \rangle^*$ .

In summary, we have shown that

$$\psi(\vec{x}) = \frac{1}{(2\pi\hbar)^{n/2}} \int d\vec{p} e^{+i\vec{x} \cdot \vec{p} / \hbar} \psi(\vec{p}) \quad (3.50)$$

$$\psi(\vec{p}) = \frac{1}{(2\pi\hbar)^{n/2}} \int d\vec{x} e^{-i\vec{x} \cdot \vec{p} / \hbar} \psi(\vec{x}) \quad (3.51)$$

and so the coordinate and momentum wave functions are Fourier transforms of each other.

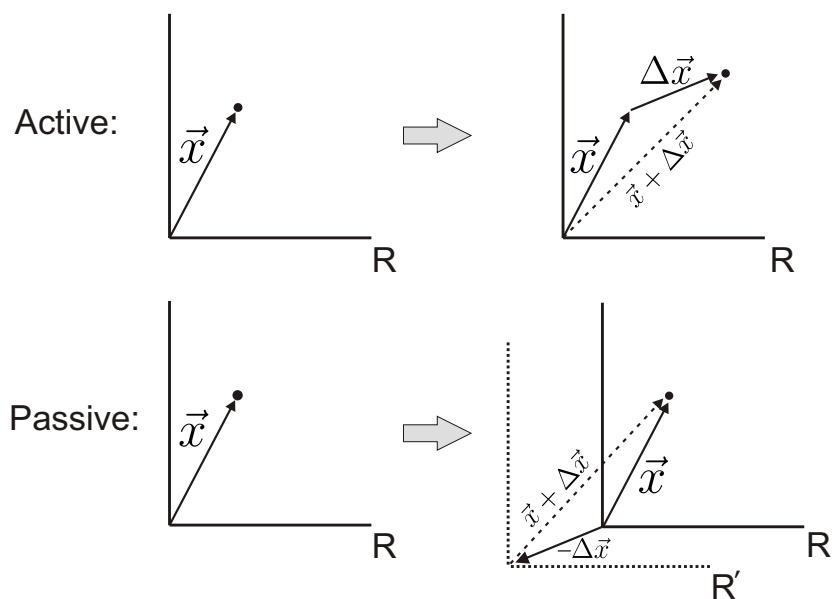
### 3.7 Active and passive transformations

We have seen that translations are represented by unitary operators acting on the Hilbert space. Indeed, the same is true of rotations and even time-evolution.

There are generally two ways in which these transformations can be applied:

- **Active:** The system itself is affected by an active transformation. For example, it may be rotated or translated in physical space, or undergo time evolution. In this case the associated unitary operator is applied directly to the state vector. The state of the system therefore changes under an active transformation.
- **Passive:** This corresponds to a change of coordinate system or reference frame. The state of the system remains unchanged. The associated unitary operator is applied to the observables to transform them to the new reference frame.

Let us illustrate these ideas using translations.



First consider an active translation which displaces the particle by  $\Delta\vec{x}$ . When applied to a position eigenstate this results in

$$|\vec{x}\rangle \implies \hat{T}(\Delta\vec{x})|\vec{x}\rangle = |\vec{x} + \Delta\vec{x}\rangle. \quad (3.52)$$

Now consider a passive transformation in which the reference frame itself is shifted by  $-\Delta\vec{x}$ . The state of the system is not affected by this, but the observables are. In fact, if  $\hat{A}$  is an observable in the original  $R$  frame it becomes

$$\hat{A}' = \hat{T}^\dagger(\Delta\vec{x})\hat{A}\hat{T}(\Delta\vec{x}) \quad (3.53)$$

in the  $R'$  frame. To see that this is correct, consider the case where  $\hat{A} = \hat{\vec{x}}$ . We found in (3.11) that  $[\hat{\vec{x}}, \hat{T}(\Delta\vec{x})] = \Delta\vec{x}\hat{T}(\Delta\vec{x})$  and therefore

$$\hat{\vec{x}}' = \hat{\vec{x}} + \Delta\vec{x}. \quad (3.54)$$

This is certainly what we expect geometrically. If the system was in the eigenstate  $|\vec{x}\rangle$  of  $\hat{\vec{x}}$  before the passive transformation it remains in this state. A measurement of  $\hat{\vec{x}}$  would then produce  $\vec{x}$  while a measurement of  $\hat{\vec{x}}'$  will yield  $\vec{x} + \Delta\vec{x}$ . This just reflects the obvious fact that a particle localised at position  $\vec{x}$  in reference frame  $R$  is localised at position  $\vec{x} + \Delta\vec{x}$  in  $R'$ .

### 3.7.1 Equivalence of the active and passive pictures

In the previous example we compared an active translation of the system by  $\Delta\vec{x}$  to a passive translation of the reference frame by  $-\Delta\vec{x}$ . Our intuition suggests that these two cases are not really different at all. In fact, we expect that an active transformation can always be reinterpreted as the inverse transformation applied passively to the reference frame. Let us prove this in general.

We begin with a system in a state  $|\phi\rangle$  and consider an observable  $\hat{A}$  with  $\hat{A}|a\rangle = a|a\rangle$  in the  $R$  reference frame. Let  $\hat{U}$  denote an arbitrary unitary transformation. Now compare the following two ways of applying  $\hat{U}$ :

- **Active:** Apply  $\hat{U}$  to the state itself:  $|\phi\rangle \implies |\phi'\rangle = \hat{U}|\phi\rangle$
- **Passive:** Apply the inverse transformation to the reference frame. The observable becomes  $\hat{A} \rightarrow \hat{A}' = \hat{U}^\dagger \hat{A} \hat{U}$ . The eigenvalue equation  $\hat{A}|a\rangle = a|a\rangle$  now becomes  $\hat{A}'|a'\rangle = a|a'\rangle$  where  $|a'\rangle = \hat{U}^\dagger|a\rangle$ .

We need to show that all the relevant physical quantities, i.e. probabilities and expectation values, are the same for these two cases. For the expectation value of  $A$  this requires that  $\langle\phi'|\hat{A}|\phi'\rangle = \langle\phi|\hat{A}'|\phi\rangle$ , which is indeed true:

$$\langle\phi'|\hat{A}|\phi'\rangle = (\langle\phi|\hat{U}^\dagger)\hat{A}(\hat{U}|\phi\rangle) = \langle\phi|(\hat{U}^\dagger\hat{A}\hat{U})|\phi\rangle = \langle\phi|\hat{A}'|\phi\rangle. \quad (3.55)$$

Now we must check that the probabilities

- $|\langle a|\phi'\rangle|^2$  = “the probability that a measurement of  $A$  will produce  $a$  when the system is in the  $|\phi'\rangle$  state”
- $|\langle a|\phi\rangle|^2$  = “the probability that measurement of  $A'$  will produce  $a$  when the system is in the  $|\phi\rangle$  state”

are equal. This follows directly from the fact that

$$|\langle a|\phi'\rangle|^2 = |\langle a|(\hat{U}|\phi\rangle)|^2 = |\langle(a|\hat{U})|\phi\rangle|^2 = |\langle a|\phi\rangle|^2. \quad (3.56)$$

*In fact, on the level of probabilities and expectation values the difference between the passive and active pictures simply amount to the placement of parentheses!*

Finally we note that applying the *same* transformation to both the states and observables has absolutely no effect. For example, if  $|\phi'\rangle = \hat{U}|\phi\rangle$  and  $\hat{A}' = \hat{U}\hat{A}\hat{U}^\dagger$  then

$$\langle\phi'|\hat{A}'|\phi'\rangle = \langle\phi|\hat{U}^\dagger\hat{U}\hat{A}\hat{U}^\dagger\hat{U}|\phi\rangle = \langle\phi|\hat{A}|\phi\rangle. \quad (3.57)$$

In a similar way you can show that inner products and probabilities are also unaffected.

# Chapter 4

## Quantum Dynamics

### 4.1 The time evolution operator

This chapter is concerned with the time evolution of quantum systems. The central question is: “*If a system is in a state  $|\psi\rangle$  at time  $t_0$ , what is its state at a later time  $t > t_0$ ?*” We will use the following notation for the time evolved state:

$$|\psi, t, t_0\rangle \equiv \text{“The system’s state at time } t \text{ if it was in the state } |\psi\rangle \text{ at time } t_0.”$$

Time evolution is realised on the Hilbert space through the time evolution operator  $\hat{U}(t, t_0)$  which defined by

$$|\psi, t, t_0\rangle = \hat{U}(t, t_0)|\psi\rangle \quad (4.1)$$

and has the following properties:

1.  $\hat{U}(t, t_0)$  is unitary
2.  $\hat{U}(t_0, t_0) = \hat{I}$
3.  $\hat{U}(t_2, t_1)\hat{U}(t_1, t_0) = \hat{U}(t_2, t_0)$

In analogy with our treatment of spatial translations we want to identify the generator of time evolution and associate it with a particular physical observable. To this end, consider an infinitesimal  $\delta t$  and expand  $\hat{U}(t_0 + \delta t, t_0)$  to linear order as:

$$\hat{U}(t_0 + \delta t, t_0) = \hat{I} + \hat{G}(t_0)\delta t. \quad (4.2)$$

The operator  $\hat{G}(t_0)$  must be anti-Hermitian for  $\hat{U}(t_0 + \delta t, t_0)$  to be unitary and may depend on the starting time  $t_0$ . The physical significance of  $\hat{G}(t_0)$  is contained in the following fundamental postulate:

**Postulate:** *The operator  $\hat{G}(t_0)$  is given by  $\hat{G}(t_0) = -i\hat{H}(t_0)/\hbar$  where  $\hat{H}(t_0)$  is the system Hamiltonian, i.e. the operator corresponding to the system’s energy at time  $t_0$ . We say that the Hamiltonian acts as the **generator of time evolution**.*

## 4.2 The time-dependent Schrödinger equation

We can use the infinitesimal form of the time evolution operator to derive a differential equation for  $\hat{U}(t, t_0)$ . First note that

$$\hat{U}(t + \delta t, t_0) = \hat{U}(t + \delta t, t)\hat{U}(t, t_0) = (\hat{I} - i\hat{H}(t)\delta t/\hbar)\hat{U}(t, t_0). \quad (4.3)$$

Rearranging this expression and sending  $\delta t \rightarrow 0$  produces

$$i\hbar \frac{d\hat{U}(t, t_0)}{dt} = \hat{H}(t)\hat{U}(t, t_0) \quad \text{with} \quad \hat{U}(t_0, t_0) = \hat{I}. \quad (4.4)$$

If we let both sides of this equation act on a state  $|\psi\rangle$  we obtain the **time-dependent Schrödinger equation**

$$i\hbar \frac{d}{dt} |\psi, t, t_0\rangle = \hat{H}(t)|\psi, t, t_0\rangle. \quad (4.5)$$

Regarding the solvability of (4.4) there are three possible scenarios:

1. If the Hamiltonian is time-independent, i.e.  $\hat{H}(t) = \hat{H}$ , then (4.4) can be solved immediately to produce

$$\hat{U}(t, t_0) = e^{-i\hat{H}(t-t_0)/\hbar}. \quad (4.6)$$

Clearly  $\hat{U}(t, t_0)$  only depends on the difference between the initial and final times. In this case we usually choose  $t_0 = 0$  and write

$$\hat{U}(t) = e^{-i\hat{H}t/\hbar}. \quad (4.7)$$

2. Suppose the Hamiltonian depends on time, but still commutes with itself at different times, i.e.  $[\hat{H}(t), \hat{H}(t')] = 0$  for all  $t$  and  $t'$ . Since all the operators appearing in (4.4) commute the differential equation can be treated as a regular scalar equation to produce

$$\hat{U}(t, t_0) = \exp \left[ -\frac{i}{\hbar} \int_{t_0}^t dt' \hat{H}(t') \right]. \quad (4.8)$$

3. Suppose the Hamiltonian is time-dependent but does not commute with itself at different times. This is the case in many practical applications. Unfortunately it is generally impossible to find analytic solutions for  $\hat{U}(t, t_0)$  under these conditions. Chapter 7 is dedicated to developing a systematic perturbative treatment of this problem.

## 4.3 The time evolution of states and observables

In this section we consider time-independent Hamiltonians and set  $t_0 = 0$ . The time evolution operator is therefore just  $\hat{U}(t) = \exp[-i\hat{H}t/\hbar]$ . Since the latter is a function of  $\hat{H}$  we expect that time evolution should be particularly simple when described in

the basis of  $\hat{H}$  eigenstates. Let  $\hat{H}|E_n\rangle = E_n|E_n\rangle$  define the eigenstates and eigenvalues of  $\hat{H}$ . Any initial state can be expanded as  $|\psi\rangle = \sum_n c_n|E_n\rangle$ , in which case

$$|\psi, t\rangle = \hat{U}(t)|\psi\rangle = \sum_n c_n e^{-i\hat{H}t/\hbar}|E_n\rangle = \sum_n e^{-iE_nt/\hbar}c_n|E_n\rangle. \quad (4.9)$$

Each term in the expansion therefore gains a time and energy dependent phase. We may also regard the expansion coefficients themselves as being time-dependent and define

$$c_n(t) \equiv e^{-iE_nt/\hbar}c_n. \quad (4.10)$$

Time-dependent expectation values and probabilities are defined in the obvious way as

$$\langle A\rangle(t) = \langle\psi, t|\hat{A}|\psi, t\rangle \quad \text{and} \quad P(a, t) = |\langle a|\psi, t\rangle|^2 \quad (4.11)$$

where  $\hat{A}|a\rangle = a|a\rangle$ . Also note that from

$$i\hbar\frac{d\hat{U}(t)}{dt} = \hat{H}(t)\hat{U}(t) \quad \text{and} \quad i\hbar\frac{d\hat{U}^\dagger(t)}{dt} = -\hat{U}^\dagger(t)\hat{H}(t) \quad (4.12)$$

we can derive the following differential equation for  $\langle A\rangle(t)$ :

$$i\hbar\frac{d}{dt}\langle A\rangle(t) = \langle\psi, t|[\hat{A}, \hat{H}]|\psi, t\rangle \quad \text{with} \quad \langle A\rangle(0) = \langle\psi|\hat{A}|\psi\rangle. \quad (4.13)$$

It is sometimes possible to solve this equation without having to determine the time-dependence of  $|\psi, t\rangle$  in full.

## Stationary states

The time evolution of an energy eigenstate is given by

$$|E_n, t\rangle = e^{-iE_nt/\hbar}|E_n\rangle. \quad (4.14)$$

This is clearly completely trivial since it just amounts to a changing global phase. All expectation values and probabilities associated with  $|E_n, t\rangle$  is therefore time-independent:

$$\langle A\rangle(t) = \langle E_n, t|\hat{A}|E_n, t\rangle = \langle E_n|\hat{A}|E_n\rangle = \langle A\rangle(0) \quad (4.15)$$

$$P(a, t) = |\langle a|E_n, t\rangle|^2 = |\langle a|E_n\rangle|^2 = P(a, 0) \quad (4.16)$$

If a system has a well-defined energy, i.e. if it is in an energy eigenstate, it does not undergo any dynamics. All its physical properties are time-independent. For this reason these states are referred to as stationary states.

### 4.3.1 Example: Spin precession in the Schrödinger picture

Appendix C.3 contains a summary of the basic facts regarding spin-1/2. It is well known that an electron possesses a magnetic dipole moment  $\vec{\mu} = \frac{e}{mc}\vec{S}$  where  $e < 0$  is its charge and  $\vec{S}$  its spin. The interaction energy of the dipole moment with an external magnetic field  $\vec{B}$  is  $E = -\vec{\mu} \cdot \vec{B}$ . On the quantum level these observables are represented by Hermitian operators:

$$\vec{\mu} \rightarrow \hat{\vec{\mu}} = \frac{e}{mc}\hat{\vec{S}} = \frac{e}{mc}(\hat{S}_x, \hat{S}_y, \hat{S}_z) \quad (4.17)$$

$$E \rightarrow \hat{H} = -\hat{\vec{\mu}} \cdot \vec{B}. \quad (4.18)$$

If we take the magnetic field to be orientated in the positive  $z$ -direction the Hamiltonian becomes

$$\hat{H} = -\frac{eB}{mc}\hat{S}_z \equiv \omega\hat{S}_z \quad (4.19)$$

where  $\omega > 0$  is a frequency with units of  $s^{-1}$ . The eigenstates of  $\hat{H}$  are clearly just  $|z, \pm\rangle$  with eigenvalues  $E_{\pm} = \pm\hbar\omega/2$ . Now consider the following question: “*What is the probability of obtaining  $\pm\hbar/2$  in a measurement of  $S_x$  at time  $t$  if the system was in the state  $|x, +\rangle$  at  $t_0 = 0$ ?*” We often abuse terminology and simply phrase this as “*What is the probability of finding the system in the state  $|x, \pm\rangle$  at time  $t$ ?*” Since we know the eigenstates of  $\hat{H}$  the first step is to express the starting state in this basis:

$$|x, +\rangle = \frac{1}{\sqrt{2}} [ |z, +\rangle + |z, -\rangle ]. \quad (4.20)$$

Applying the time evolution operator yields

$$|x, +, t\rangle = \frac{1}{\sqrt{2}} [ e^{-i\omega t/2} |z, +\rangle + e^{+i\omega t/2} |z, -\rangle ] \quad (4.21)$$

and so the probability in question is

$$P(S_x = \pm\hbar/2) = |\langle x, \pm | x, +, t \rangle|^2 \quad (4.22)$$

$$= \frac{1}{4} |(\langle z, + | \pm \langle z, - |) (e^{-i\omega t/2} |z, +\rangle + e^{+i\omega t/2} |z, -\rangle)|^2 \quad (4.23)$$

$$= \frac{1}{4} |e^{-i\omega t/2} \pm e^{+i\omega t/2}|^2 \quad (4.24)$$

where we used the orthogonality relation  $\langle z, \sigma' | z, \sigma \rangle = \delta_{\sigma', \sigma}$ . This simplifies to

$$P(S_x = +\hbar/2) = \cos^2(\omega t/2) \quad \text{and} \quad P(S_x = -\hbar/2) = \sin^2(\omega t/2). \quad (4.25)$$

Next we consider the expectation values of the components of  $\vec{S}$ . As with all spin-1/2 calculations we can either use the matrix representation or stick with the bra-ket formalism. *It is important that you understand both these approaches well.* Either way, it is not difficult to show that

$$\langle S_x \rangle(t) = \langle x, +, t | \hat{S}_x | x, +, t \rangle = \frac{\hbar}{2} \cos(\omega t) \quad \langle S_y \rangle(t) = \frac{\hbar}{2} \sin(\omega t) \quad \langle S_z \rangle(t) = 0.$$

It follows that the *expectation value* of the spin  $\vec{S}$  lies in the  $x - y$  plane and precesses around the  $z$ -axis with an angular frequency of  $\omega$ . Note that this is **not** a statement about  $\vec{S}$  itself, since, due to the uncertainty principle, we cannot assign a well-defined value to more than one component of  $\vec{S}$  at a time.

### 4.3.2 The Schrödinger and Heisenberg pictures

This section is applicable to any Hamiltonian  $\hat{H}(t)$ . For convenience we set  $t_0 = 0$ . Thus far our description of quantum dynamics has been phrased in terms of time-dependent states. This is known as the Schrödinger picture. An equivalent description in terms of time-dependent observables is given by the Heisenberg picture. These two descriptions correspond to active and passive pictures of time evolution, as defined in

section 3.7. Of course, the physical content of the theory cannot depend on which picture we use. In fact, we can use the invariance of physical quantities to “derive” these two pictures as follows.

First consider a time-dependent expectation value which can be written in two equivalent ways

$$\langle \phi, t | \hat{A} | \phi, t \rangle = \underbrace{\langle \phi | \hat{U}^\dagger(t)}_{\text{state bra}} \underbrace{\hat{A}}_{\text{obs.}} \underbrace{\hat{U}(t) | \phi \rangle}_{\text{state ket}} = \underbrace{\langle \phi |}_{\text{state bra}} \underbrace{\hat{U}^\dagger(t) \hat{A} \hat{U}(t)}_{\hat{A}_H(t)} \underbrace{| \phi \rangle}_{\text{state ket}}. \quad (4.26)$$

The first case corresponds to the Schrödinger (active) picture in which the state evolves in time while observables are fixed. The second case is the Heisenberg (passive) description in which the state is time-independent while observables are transformed by  $\hat{U}(t)$ . We use  $\hat{A}_H(t) = \hat{U}^\dagger(t) \hat{A} \hat{U}(t)$  to denote the observable  $\hat{A}$  in the Heisenberg picture. It is insightful to consider probabilities in the same manner. Let  $\hat{A}|a\rangle = a|a\rangle$ . We then have two ways of writing  $P(a, t)$ :

$$P(a, t) = |\langle a | \phi, t \rangle|^2 = |\langle a | \underbrace{\hat{U}(t)}_{\text{state}} | \phi \rangle|^2 = |\langle a | \hat{U}(t) | \underbrace{\phi}_{\text{state}} \rangle|^2. \quad (4.27)$$

Note that in the active case  $\langle a |$  is the bra corresponding to the eigenstate  $|a\rangle$  of  $\hat{A}$  with eigenvalue  $a$ . In the passive description  $\langle a | \hat{U}(t)$  is the bra of the eigenstate  $\hat{U}^\dagger(t)|a\rangle$  of  $\hat{A}_H(t)$  with eigenvalue  $a$ . To summarize:

$$\text{Schrödinger Picture:} \quad \hat{A}_S = \hat{A} \quad |\phi, t\rangle_S = \hat{U}(t)|\phi\rangle \quad (4.28)$$

$$\text{Heisenberg Picture:} \quad \hat{A}_H(t) = \hat{U}^\dagger(t) \hat{A} \hat{U}(t) \quad |\phi\rangle_H = |\phi\rangle \quad (4.29)$$

We usually drop all the subscripts except on  $\hat{A}_H(t)$ . The time-dependent Schrödinger equation (4.5) provides an equation of motion for states in the Schrödinger picture. The Heisenberg equation of motion for an observable follows from (4.12) as

$$\frac{d\hat{A}_H(t)}{dt} = \frac{1}{i\hbar} [\hat{A}_H(t), \hat{H}_H(t)] + \hat{U}^\dagger(t) \frac{\partial \hat{A}}{\partial t} \hat{U}(t). \quad (4.30)$$

The final term is required if  $\hat{A}$  has some intrinsic time dependence. Usually the Hamiltonian is a function of other observables, for example  $\hat{H} = H(\hat{x}, \hat{p}, \hat{S})$ . In this case  $\hat{H}_H(t) = H(\hat{x}_H(t), \hat{p}_H(t), \hat{S}_H(t))$ . The significance of having all the operators in the Heisenberg picture is that we use the standard commutation relations to perform calculations. To be clear, if  $[\hat{A}, \hat{B}] = \hat{C}$  then it is easy to see that  $[\hat{A}_H(t), \hat{B}_H(t)] = \hat{C}_H(t)$ , and so evolving the operators in time does not affect their commutation relations.

The Heisenberg equation of motion also reveals the close connection between the Heisenberg picture and the Hamilton formulation of classical mechanics. For a classical observable  $A$  with no explicit time dependence the equation of motion can be expressed as

$$\frac{dA}{dt} = \{A, H\}_{PB} \quad (4.31)$$

where  $\{\cdot, \cdot\}_{PB}$  is the Poisson bracket. In this setting the transition from classical to quantum mechanics therefore amounts to the replacement

$$\{A, H\}_{PB} \implies \frac{1}{i\hbar} [\hat{A}_H, \hat{H}_H] \quad (4.32)$$

# Chapter 5

## Angular Momentum

### 5.1 Introduction

In this chapter we develop the general quantum mechanical theory of angular momentum. This formalism will allow for the description of both orbital angular momentum as well as spin degrees of freedom. Just as the momentum operators generate translations and the Hamiltonian generates time evolution we expect that rotations are generated by the angular momentum operators. This is indeed the case, but to construct these operators in general requires a thorough understanding of rotations and how they are realised on the Hilbert space. Of central importance in this regard are the algebraic properties of the rotation operators which are inherited directly from the rotation group itself. We will therefore begin with a careful study of this group.

### 5.2 The rotation group $SO(3)$

With each rotation in three dimensions there is associated a  $3 \times 3$  rotation matrix which is real, orthogonal and has a determinant of  $+1$ . It is easy to check that the set of such matrices forms a group<sup>1</sup> under matrix multiplication. This group is known as  $SO(3)$ .<sup>2</sup> Since rotations in three dimensions do not commute this group is not Abelian. To be clear, if  $R_1$  and  $R_2$  are rotation matrices in  $SO(3)$  then so are  $R_1R_2$  and  $R_2R_1$ , but generally  $R_1R_2 \neq R_2R_1$ . This fact has far-reaching consequences on the quantum mechanical level. As we will see, since rotations don't commute the generators of rotations, i.e. the angular momentum operators, don't commute either, and so it is not possible to construct states for which all three components of the angular momentum

---

<sup>1</sup>A group is a set  $G$  together with an operation  $\cdot$  which combines two elements  $a, b \in G$  into another element  $a \cdot b \in G$ . The set  $G$  and operation  $\cdot$  must satisfy the following group axioms:

1. **Closure:** If  $a, b \in G$  then  $a \cdot b \in G$ .
2. **Associativity:** For all  $a, b, c \in G$  it holds that  $a \cdot (b \cdot c) = (a \cdot b) \cdot c$ .
3. **Identity Element:** There exists a unique element  $I \in G$  such that  $I \cdot a = a = a \cdot I$  for all  $a \in G$ .
4. **Inverse:** For each  $a \in G$  there exists a unique  $a^{-1} \in G$  such that  $a \cdot a^{-1} = I = a^{-1} \cdot a$ .

<sup>2</sup>Here  $S$  stands for “special” which reflects the fact that the determinants equal  $+1$ . What kind of transformations does  $O(3)$  represent?

are well-defined. In contrast, spatial translations do commute and consequently so do the three generators  $(\hat{p}_1, \hat{p}_2, \hat{p}_3)$ . We will continue this discussion in detail later. For now, take note of the following: *There is a fundamental connection between the group structure of transformations groups (rotations, translations, Lorentz transformations) and the commutation relations between quantum mechanical observables.*

## 5.3 Rotation matrices

### 5.3.1 Generators

Let us investigate the general structure of rotation matrices in more detail. If  $R_z(\phi)$  denotes a rotation around the  $z$ -axis by an angle  $\phi$  then it is easy to verify that

$$R_z(\phi) = \begin{bmatrix} \cos(\phi) & -\sin(\phi) & 0 \\ \sin(\phi) & \cos(\phi) & 0 \\ 0 & 0 & 1 \end{bmatrix}. \quad (5.1)$$

It is again useful to consider infinitesimal transformations; in this case a rotation by an infinitesimal angle  $\delta\phi$ . Up to linear order we see that

$$R_z(\delta\phi) = \begin{bmatrix} 1 & -\delta\phi & 0 \\ \delta\phi & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = (I + \delta\phi T_z) \quad (5.2)$$

where  $I$  is the  $3 \times 3$  identity matrix and we call  $T_z$  the *generator of rotations* around the  $z$ -axis.  $T_x$  and  $T_y$  can be found in the same way. Altogether we have

$$T_x = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} \quad T_y = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix} \quad T_z = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad (5.3)$$

Note that these three matrices span the space of all anti-symmetric  $3 \times 3$  matrices.

### 5.3.2 Finite rotations

Now consider a finite rotation by an angle  $\phi$ . Since rotations around a single axis are additive we see that

$$R_x(\phi) = \lim_{N \rightarrow \infty} [R_x(\phi/N)]^N = \lim_{N \rightarrow \infty} [I + (\phi/N)T_x]^N = e^{\phi T_x} \quad (5.4)$$

and similar for  $R_y(\phi)$  and  $R_z(\phi)$ . A rotation by  $\phi$  around an axis specified by the unit vector  $\hat{n} = (n_x, n_y, n_z)$  should be of the form

$$R_{\hat{n}}(\phi) = e^{\phi T_{\hat{n}}} \quad (5.5)$$

where  $T_{\hat{n}}$  is the corresponding generator. Since  $T_{\hat{n}}$  is antisymmetric it can be expressed as a linear combination of  $T_x$ ,  $T_y$  and  $T_z$ . In other words, there exists a vector  $\vec{m} = (m_x, m_y, m_z)$  such that

$$T_{\hat{n}} = \vec{m} \cdot \vec{T} = m_x T_x + m_y T_y + m_z T_z = \begin{bmatrix} 0 & -m_z & m_y \\ m_z & 0 & -m_x \\ -m_y & m_x & 0 \end{bmatrix}. \quad (5.6)$$

We want to determine the relation between  $\vec{m}$  and  $\hat{n}$ . Since  $R_{\hat{n}}(\phi)\hat{n} = \hat{n}$  for all  $\phi$  it must hold that  $T_{\hat{n}}\hat{n} = 0$ . You can easily check that

$$T_{\hat{n}}\hat{n} = (\vec{m} \cdot \vec{T})\hat{n} = \vec{m} \times \hat{n} \quad (5.7)$$

and so  $T_{\hat{n}}\hat{n} = 0$  implies that  $\vec{m} = \lambda\hat{n}$  for some constant  $\lambda$ . This fixes the form of  $\hat{T}_{\hat{n}}$  to be  $\hat{T}_{\hat{n}} = \lambda\hat{n} \cdot \vec{T}$  with  $\lambda$  still unknown. However,  $\hat{T}_{\hat{n}}$  must reduce to  $T_{x,y,z}$  when  $\hat{n} = \hat{x}, \hat{y}, \hat{z}$  and this implies that  $\lambda = 1$ .

The matrix for a rotation by an angle  $\phi$  around a general axis  $\hat{n}$  can therefore be expressed in terms of the generators as

$$R_{\hat{n}}(\phi) = e^{\phi\hat{n}\cdot\vec{T}}. \quad (5.8)$$

### 5.3.3 Commutation relations and group multiplication

We know that the rotation matrices form a group under multiplication, and so the product of any two such matrices is again a rotation matrix. In fact,  $SO(3)$  is a particular type of group called a **Lie group**. The generators (or linear combinations of them) don't form a group in the same way, but they do form another extremely important algebraic structure. To see this, calculate the commutators between the three generators:

$$[T_x, T_y] = T_z \quad [T_z, T_x] = T_y \quad [T_y, T_z] = T_x \quad (5.9)$$

or in general

$$[T_i, T_j] = \epsilon_{ijk} T_k \quad (5.10)$$

where  $\epsilon_{ijk}$  is the Levi-Civita symbol:

$$\epsilon_{ijk} = \begin{cases} 0 & \text{if any two indices are equal} \\ 1 & \text{for even permutations of } 123 \\ -1 & \text{for odd permutations of } 123 \end{cases} \quad (5.11)$$

Now define  $so(3) = \text{span}_{\mathbb{R}}\{T_x, T_y, T_z\}$  as the real vector space spanned by  $T_x$ ,  $T_y$  and  $T_z$ . This is nothing but the vector space of  $3 \times 3$  antisymmetric real matrices. Unlike  $SO(3)$  the set  $so(3)$  is not closed under multiplication, but according to (5.10) it is closed under *commutation*. We say that  $so(3)$  is the **Lie algebra** corresponding to the Lie group  $SO(3)$ . These two algebraic structures are very closely related and studying one reveals a lot about the other. In fact, the Lie group completely determines the corresponding Lie algebra, but the reverse is not always true: different Lie groups may share the same Lie algebra. This fact will have important implications later on.

Let us investigate the link between the multiplication rule of  $SO(3)$  and the commutation relations of  $so(3)$  more closely. Consider two arbitrary rotations

$$R_{\hat{n}_1}(\phi_1) = e^{\phi_1\hat{n}_1\cdot\vec{T}} \quad \text{and} \quad R_{\hat{n}_2}(\phi_2) = e^{\phi_2\hat{n}_2\cdot\vec{T}}. \quad (5.12)$$

Since we are dealing with a group there must exist a rotation  $R_{\hat{n}_3}(\phi_3)$  such that

$$R_{\hat{n}_1}(\phi_1)R_{\hat{n}_2}(\phi_2) = e^{\phi_1\hat{n}_1\cdot\vec{T}}e^{\phi_2\hat{n}_2\cdot\vec{T}} = e^{\phi_3\hat{n}_3\cdot\vec{T}} = R_{\hat{n}_3}(\phi_3). \quad (5.13)$$

Here  $\phi_3 = \phi_3(\phi_{1,2}, \hat{n}_{1,2})$  and  $\hat{n}_3 = \hat{n}_3(\phi_{1,2}, \hat{n}_{1,2})$  are functions which serve as the definition of the group multiplication rule. We will now show that these functions, and therefore the group structure itself, are completely determined by the commutation relations in (5.10). Clearly we need a way of combining exponentials of matrices into a single exponential, i.e. we need to determine how  $Z$  depends on  $A$  and  $B$  in

$$e^A e^B = e^Z. \quad (5.14)$$

Of course, if  $A$  and  $B$  don't commute it is not simply a case of  $Z = A + B$ . The general expression for  $Z$  is an infinite series of nested commutators known as the Baker-Campbell-Hausdorff (BCH) formula<sup>3</sup>. The first few terms are:

$$Z = A + B + \frac{1}{2}[A, B] + \frac{1}{12}([[A, B], B] + [A, [A, B]]) + \frac{1}{24}[[[B, A], A], B] + \dots \quad (5.15)$$

Now, if  $A, B \in so(3)$  then so are  $[A, B]$ ,  $[[A, B], B]$ ,  $[A, [A, B]]$  and so on. In fact, it is clear that  $Z$  itself will be an element of  $so(3)$ . Of course, this *must* be the case for (5.13) to hold at all. The fact that  $so(3) = \text{span}_R\{T_x, T_y, T_z\}$  is closed under commutation is therefore no coincidence; it is *required* if the group is to be closed under multiplication. Furthermore, the commutation relations between the  $T_i$ s completely determine  $\phi_3$  and  $\hat{n}_3$ , and therefore also the group structure itself. *The commutation relations in (5.10) therefore encode, in a very compact way, all the algebraic properties of rotations in three dimensions.*

## 5.4 Rotations in quantum mechanics

The results of the previous section has provided insight into the structure of rotation *matrices*. Of course, these matrices act on regular vectors in  $\mathbb{R}^3$  and certainly not on the elements of a quantum Hilbert space. To describe how the quantum state of a system transforms under a rotation requires a way of representing the rotation as a unitary operator acting on the Hilbert space. In other words, we need to construct operators that capture the fundamental properties of rotations despite not being rotation matrices themselves. In mathematical terms we are asking for a **representation of the Lie group**  $SO(3)$  in terms of unitary operators. To be precise, we seek a mapping  $\hat{D}$  which associates with each rotation  $R \in SO(3)$  a unitary operator  $\hat{D}(R)$  acting on a Hilbert space  $\mathcal{H}$ . Since the algebraic properties of rotations are contained in the group multiplication rule of  $SO(3)$  we require that this is preserved by the mapping  $\hat{D}$ . This means that  $\hat{D}$  must be a **group homomorphism**:

$$\hat{D}(R_1 R_2) = \hat{D}(R_1) \hat{D}(R_2) \quad \text{for all } R_1, R_2 \in SO(3). \quad (5.16)$$

It follows that  $\hat{D}(I) = \hat{I}$  where  $I$  is the  $3 \times 3$  identity matrix and  $\hat{I}$  the identity operator on  $\mathcal{H}$ . We also see that

$$\hat{D}(RR^{-1}) = \hat{I} = \hat{D}(R)\hat{D}(R^{-1}) \implies \hat{D}(R^{-1}) = (\hat{D}(R))^{-1} = (\hat{D}(R))^\dagger. \quad (5.17)$$

Now consider an infinitesimal rotation  $R_{\hat{n}}(\delta\phi)$ . The corresponding operator depends on the three infinitesimal parameters  $(\delta\phi n_x, \delta\phi n_y, \delta\phi n_z)$  and can be expanded to linear order as

$$\hat{D}(R_{\hat{n}}(\delta\phi)) = \hat{I} + \delta\phi \hat{n} \cdot \hat{\vec{G}} \quad (5.18)$$

---

<sup>3</sup>There are a number of related formulas which we just call “BCH formulas” in general.

where  $\hat{\vec{G}} = (\hat{G}_1, \hat{G}_2, \hat{G}_3)$ . The same procedure as in (5.4) now leads to an expression for a finite rotation in the familiar form

$$\hat{D}(R_{\hat{n}}(\phi)) = e^{\phi \hat{n} \cdot \hat{\vec{G}}} \quad (5.19)$$

which is the analog of  $R_{\hat{n}}(\phi) = e^{\phi \hat{n} \cdot \vec{T}}$ . Whereas the  $T_i$  matrices were antisymmetric the  $\hat{G}_i$  operators must clearly be anti-hermitian. We now ask the following: *What algebraic properties must the  $\hat{G}_i$  operators have to ensure that (5.16) is satisfied, i.e. that the multiplication rule for the operators matches that of the rotations they represent?* From the discussion in section 5.3.3 the answer should be clear: *The  $\hat{G}_i$  operators must satisfy exactly the same commutation relations as the  $T_i$  matrices, i.e.  $[\hat{G}_i, \hat{G}_j] = \epsilon_{ijk} \hat{G}_k$ .* In mathematical terms this means that the  $\hat{G}_i$  operators form a **representation of the Lie algebra  $so(3)$** . To bring things into their final forms we introduce the Hermitian operator  $\hat{\vec{J}} = i\hbar \hat{\vec{G}}$  of which the components satisfy

$$[\hat{J}_i, \hat{J}_j] = i\hbar \epsilon_{ijk} \hat{J}_k. \quad (5.20)$$

The rotation operator becomes

$$\hat{D}(R_{\hat{n}}(\phi)) = e^{-i\phi \hat{n} \cdot \hat{\vec{J}}/\hbar}. \quad (5.21)$$

As we have done twice before we now postulate that  $\hat{\vec{J}}$  is in fact physical observable:

**Postulate:**  $\hat{\vec{J}} = (\hat{J}_1, \hat{J}_2, \hat{J}_3) = (\hat{J}_x, \hat{J}_y, \hat{J}_z)$  is the angular momentum operator. Its three components are the **generators of rotations** on the Hilbert space.

To summarize, we have found that if the operators  $\hat{D}(R)$  satisfy (5.16) then they are of the exponential form given in (5.21) where the generators satisfy the commutation relations in (5.20). In other words, every representation of the Lie group  $SO(3)$  induces a representation of the Lie algebra  $so(3)$ . One may ask whether the reverse is also true, i.e. if we construct operators  $\{\hat{J}_1, \hat{J}_2, \hat{J}_3\}$  that satisfy (5.20) will inserting them into (5.21) produce operators that satisfy (5.16)? Or, even more compactly, does every representation of  $so(3)$  produce, upon exponentiation, a representation of  $SO(3)$ ? This is a subtle question, and the answer is not yet obvious. What is certainly true is that if *all* the representations of  $so(3)$  is known we can construct *all* possible  $SO(3)$  representations via the exponential formula (5.20). This is the approach we will follow in sections 5.6 and 5.7.

## 5.5 Orbital angular momentum

Before proceeding with the general construction we first consider a special case for which there exists a very natural representation of  $SO(3)$ . Consider the Hilbert space  $\mathcal{H} = \text{span}\{|\vec{x}\rangle : \vec{x} \in \mathbb{R}^3\}$  associated with the spatial degrees of freedom of a single particle. This is just the Hilbert space on which the position and momentum operators act. Since  $|\vec{x}\rangle$  represents a particle localized at the point  $\vec{x} \in \mathbb{R}^3$  the natural representation of rotations on this space is given by

$$\hat{D}(R)|\vec{x}\rangle = |R\vec{x}\rangle. \quad (5.22)$$

We note that

$$\hat{D}(R_1)\hat{D}(R_2)|\vec{x}\rangle = \hat{D}(R_1)|R_2\vec{x}\rangle = |R_1R_2\vec{x}\rangle = \hat{D}(R_1R_2)|\vec{x}\rangle \quad (5.23)$$

holds for any  $|\vec{x}\rangle$  basis state, and so, on the operator level,  $\hat{D}(R_1)\hat{D}(R_2) = \hat{D}(R_1R_2)$ . It follows that  $\hat{D}$  is a homomorphism and a valid representation of  $SO(3)$  on this Hilbert space. Furthermore,  $\hat{D}(R_{\hat{n}}(\phi))$  must be of the form

$$\hat{D}(R_{\hat{n}}(\phi)) = e^{-i\phi\hat{n}\cdot\hat{J}/\hbar}. \quad (5.24)$$

We want to find an expression for the angular momentum operator  $\hat{\vec{J}}$  associated with this particular representation. To do so we consider an infinitesimal rotation by  $\delta\phi$  around the  $\hat{n}$  axis. When applied to the position vector  $\vec{x}$  this rotation produces

$$R_{\hat{n}}(\delta\phi)\vec{x} = \vec{x} + \delta\phi(\hat{n}\cdot\vec{T})\vec{x} = \vec{x} + \delta\phi(\hat{n}\times\vec{x}) \equiv \vec{x} + \delta\vec{x} \quad (5.25)$$

which resembles an infinitesimal translation, albeit one that depends on the starting position. On the quantum state level we have

$$\hat{D}(R_{\hat{n}}(\delta\phi))|\vec{x}\rangle = |\vec{x} + \delta\vec{x}\rangle = \hat{T}(\delta\vec{x})|\vec{x}\rangle = (\hat{I} - i\hat{\vec{p}}\cdot\delta\vec{x}/\hbar)|\vec{x}\rangle \quad (5.26)$$

which simplifies to

$$\hat{D}(R_{\hat{n}}(\delta\phi))|\vec{x}\rangle = (\hat{I} - i\delta\phi\hat{\vec{p}}\cdot(\hat{n}\times\vec{x})/\hbar)|\vec{x}\rangle \quad (5.27)$$

$$= (\hat{I} - i\delta\phi\hat{n}\cdot(\vec{x}\times\hat{\vec{p}})/\hbar)|\vec{x}\rangle \quad (5.28)$$

$$= (\hat{I} - i\delta\phi\hat{n}\cdot(\hat{\vec{x}}\times\hat{\vec{p}})/\hbar)|\vec{x}\rangle. \quad (5.29)$$

Since this holds for all position eigenstates we can conclude that

$$\hat{D}(R_{\hat{n}}(\delta\phi)) = (\hat{I} - i\delta\phi\hat{n}\cdot(\hat{\vec{x}}\times\hat{\vec{p}})/\hbar). \quad (5.30)$$

From (5.24) we also have  $\hat{D}(R_{\hat{n}}(\delta\phi)) = (\hat{I} - i\delta\phi\hat{n}\cdot\hat{\vec{J}}/\hbar)$  and so

$$\hat{\vec{J}} = \hat{\vec{x}} \times \hat{\vec{p}}, \quad (5.31)$$

which is precisely the well known orbital angular momentum operator  $\hat{\vec{L}} = \hat{\vec{x}} \times \hat{\vec{p}}$ . Since  $\hat{D}(R)$  is a representation of  $SO(3)$  these operators must satisfy

$$[\hat{L}_i, \hat{L}_j] = i\hbar\epsilon_{ijk}\hat{L}_k. \quad (5.32)$$

You can also verify this independently by using the known position-momentum commutation relations.

It is important to understand that, although everything has worked as expected here, this is not the whole story. The  $\hat{\vec{L}}$  operators only generate rotations of spatial degrees of freedom, and not internal degrees of freedom such as spin. The latter requires a more general formalism of rotations and angular momentum which we will develop in the next section.

## 5.6 Eigenvalues and matrix elements of angular momentum operators

In this section we will derive, using *only* the commutation relations  $[\hat{J}_i, \hat{J}_j] = i\hbar\epsilon_{ijk}\hat{J}_k$ , the general form of the matrix representation of angular momentum operators in the basis of  $\hat{J}_z$  and  $\hat{J}^2$  eigenstates. This provides a classification scheme for all the distinct ways of constructing angular momentum operators, and therefore also for all the possible representations of  $SO(3)$ .

First we identify an appropriate basis. Since the angular momentum operators don't commute there is no hope of finding a basis of simultaneous eigenstates of all three, or even any two. We therefore pick one component, traditionally  $\hat{J}_z$ , and seek another Hermitian operator which commutes with it. One such operator is

$$\hat{J}^2 = \hat{\vec{J}} \cdot \hat{\vec{J}} = \hat{J}_x^2 + \hat{J}_y^2 + \hat{J}_z^2 \quad (5.33)$$

which satisfies

$$[\hat{J}^2, \hat{J}_x] = [\hat{J}^2, \hat{J}_y] = [\hat{J}^2, \hat{J}_z] = 0. \quad (5.34)$$

The operator  $\hat{J}^2$  represents the square of the total angular momentum. In other words,  $\hat{J}^2$  given an indication of the *magnitude* of the angular momentum, and  $\hat{J}_z$  of its *orientation*. We label the simultaneous eigenstates of  $\hat{J}^2$  and  $\hat{J}_z$  by their corresponding eigenvalues as

$$\hat{J}^2|a, b\rangle = a|a, b\rangle \quad \text{and} \quad \hat{J}_z|a, b\rangle = b|a, b\rangle. \quad (5.35)$$

For the moment we assume that these states are completely specified by their eigenvalues with respect to  $\hat{J}^2$  and  $\hat{J}_z$ . This is generally *not* the case and other “non-angular momentum” labels are typically required.

### 5.6.1 Ladder operators

We now turn to the question of what the possible values of the eigenvalues  $a$  and  $b$  are. For this purpose we introduce the ladder operators  $\hat{J}_+$  and  $\hat{J}_-$  as

$$\hat{J}_{\pm} = \hat{J}_x \pm i\hat{J}_y \quad \text{for which} \quad \hat{J}_{\pm}^\dagger = \hat{J}_{\mp}. \quad (5.36)$$

These operators satisfy

$$[\hat{J}_+, \hat{J}_-] = 2\hbar\hat{J}_z \quad [\hat{J}_z, \hat{J}_{\pm}] = \pm\hbar\hat{J}_{\pm} \quad [\hat{J}^2, \hat{J}_{\pm}] = 0 \quad (5.37)$$

as well as

$$\hat{J}^2 = (\hat{J}_+ \hat{J}_- + \hat{J}_- \hat{J}_+)/2 + \hat{J}_z^2 = \hat{J}_- \hat{J}_+ + \hbar\hat{J}_z + \hat{J}_z^2 = \hat{J}_+ \hat{J}_- - \hbar\hat{J}_z + \hat{J}_z^2. \quad (5.38)$$

To see what effect  $\hat{J}_{\pm}$  has on a  $|a, b\rangle$  state consider the expression

$$\hat{J}_z \hat{J}_{\pm}|a, b\rangle = (\hat{J}_{\pm} \hat{J}_z + [\hat{J}_z, \hat{J}_{\pm}])|a, b\rangle = (\hat{J}_{\pm} b \pm \hbar\hat{J}_{\pm})|a, b\rangle = (b \pm \hbar)\hat{J}_{\pm}|a, b\rangle \quad (5.39)$$

from which it is clear that  $\hat{J}_{\pm}|a, b\rangle$  is also an eigenstate of  $\hat{J}_z$ , but with an eigenvalue of  $(b \pm \hbar)$ . Similarly

$$\hat{J}^2 \hat{J}_{\pm}|a, b\rangle = \hat{J}_{\pm} \hat{J}^2|a, b\rangle = a \hat{J}_{\pm}|a, b\rangle \quad (5.40)$$

and so  $\hat{J}_\pm|a, b\rangle$  is an  $\hat{J}^2$  eigenstate with eigenvalue  $a$ . Since these eigenvalues are assumed to specify the state uniquely it follows that

$$\hat{J}_\pm|a, b\rangle \propto |a, b \pm \hbar\rangle. \quad (5.41)$$

The  $\hat{J}_\pm$  operators therefore move us along a ladder of states, all with the same  $\hat{J}^2$  eigenvalue but with  $\hat{J}_z$  eigenvalues differing by integer multiples of  $\hbar$ . Now consider the range of the  $\hat{J}_z$  eigenvalues. Classically we expect  $|b|$  to be bounded by  $\sqrt{a}$ , since the magnitude of a vector cannot be less than the magnitude of one of its components. To prove this on the quantum level we note that

$$a - b^2 = \langle a, b | \hat{J}^2 - \hat{J}_z^2 | a, b \rangle = \langle a, b | \hat{J}_x^2 + \hat{J}_y^2 | a, b \rangle = \langle a, b | \hat{J}_x^\dagger \hat{J}_x | a, b \rangle + \langle a, b | \hat{J}_y^\dagger \hat{J}_y | a, b \rangle \geq 0 \quad (5.42)$$

and therefore  $-\sqrt{a} \leq b \leq \sqrt{a}$ . It follows that there must exist an eigenvalue  $b_{max}$  of  $\hat{J}_z$  such that  $\hat{J}_+|a, b_{max}\rangle = 0$ . Furthermore, we expect  $b_{max}$  to be determined by  $a$ . To establish this connection note that

$$0 = \hat{J}_- \hat{J}_+ |a, b_{max}\rangle = (\hat{J}^2 - \hbar \hat{J}_z - \hat{J}_z^2) |a, b_{max}\rangle = (a - \hbar b_{max} - b_{max}^2) |a, b_{max}\rangle \quad (5.43)$$

and so  $a = b_{max}(b_{max} + \hbar)$ . By the same argument there exists a  $b_{min}$  such that  $\hat{J}_-|a, b_{min}\rangle = 0$ , in which case  $a = b_{min}(b_{min} - \hbar)$ . Combining these two expressions for  $a$  produces

$$b_{max}(b_{max} + \hbar) = b_{min}(b_{min} - \hbar) \implies b_{min} = -b_{max}. \quad (5.44)$$

Since we can ladder from  $|a, b_{min}\rangle$  up to  $|a, b_{max}\rangle$  using  $\hat{J}_+$  there must exist a integer  $n$  such that  $b_{max} = b_{min} + n\hbar = -b_{max} + n\hbar$  and so  $b_{max} = n\hbar/2$ .

### 5.6.2 Eigenvalues

We now switch to a more convenient labelling in terms of  $m \equiv b/\hbar$  and  $j \equiv n/2$ . Here  $j$  takes integer and half-integer values while  $m$  runs from  $-j$  up to  $+j$  in integer steps. To summarize:

$$\hat{J}^2|j, m\rangle = j(j+1)\hbar^2|j, m\rangle \quad j \in \{0, \frac{1}{2}, 1, \frac{3}{2}, 2, \dots\} \quad (5.45)$$

$$\hat{J}_z|j, m\rangle = m\hbar|j, m\rangle \quad m = j, j-1, \dots, -j+1, -j \quad (5.46)$$

$$\langle j', m' | j, m \rangle = \delta_{j,j'} \delta_{m,m'} \quad (\text{Orthogonality}) \quad (5.47)$$

Note that we have only determined the *allowed* values of  $j$ . It is possible that, in a particular construction of the angular momentum operators, only a subset of these values are present or that the same  $j$  appears multiple times. See section 5.6.5.

### 5.6.3 Notation

When dealing with orbital angular momentum we may use the notation

$$\hat{\vec{L}} = \hat{\vec{J}} \quad l = j \quad m_l = m. \quad (5.48)$$

For spin angular momentum we use

$$\hat{\vec{S}} = \hat{\vec{J}} \quad s = j \quad m_s = m. \quad (5.49)$$

Of course, results derived for a general  $\hat{\vec{J}}$  are applicable to both these two cases.

### 5.6.4 Matrix elements

The matrix elements of  $\hat{J}^2$  and  $\hat{J}_z$  follow directly from the eigenvalue equations as

$$\langle j', m' | \hat{J}^2 | j, m \rangle = j(j+1)\hbar^2 \delta_{j',j} \delta_{m',m} \quad \langle j', m' | \hat{J}_z | j, m \rangle = m\hbar \delta_{j',j} \delta_{m',m} \quad (5.50)$$

To find the matrix elements of the ladder operators we first need to determine the proportionality constants  $c_{\pm}$  in  $\hat{J}_{\pm}|j, m\rangle = c_{\pm}|j, m \pm 1\rangle$ . Taking the normed squared of both sides yields

$$|c_{\pm}|^2 = \langle j, m | \hat{J}_{\mp} \hat{J}_{\pm} | j, m \rangle = \langle j, m | \hat{J}^2 \mp \hbar \hat{J}_z - \hat{J}_z^2 | j, m \rangle = (j(j+1) - m(m \pm 1))\hbar^2 \quad (5.51)$$

where (5.38) has been used. By convention  $c_{\pm}$  is taken to be real and positive:

$$c_{\pm}(j, m) = \sqrt{j(j+1) - m(m \pm 1)}\hbar. \quad (5.52)$$

To summarize, we have

$$\hat{J}_{\pm}|j, m\rangle = c_{\pm}(j, m)|j, m \pm 1\rangle \quad (5.53)$$

and so

$$\langle j', m' | \hat{J}_{\pm} | j, m \rangle = c_{\pm}(j, m) \delta_{j',j} \delta_{m',m \pm 1}. \quad (5.54)$$

From these results the matrix elements of  $\hat{J}_x$  and  $\hat{J}_y$  can be calculated easily.

**Exercise:** Derive the matrix representation of the spin-1/2 operators as it appears in section C.3.

### 5.6.5 Reducible and irreducible representations

We set out to classify all the representations of  $so(3)$ , i.e. all the ways of constructing Hermitian operators  $\hat{J}_{x,y,z}$  that satisfy  $[\hat{J}_i, \hat{J}_j] = i\hbar\epsilon_{ijk}\hat{J}_k$ . The results of the previous section provide us with such a classification scheme. This can be seen as follows. Imagine you are given a set of operators  $\hat{J}_{x,y,z}$  that satisfy the required commutation relations. Construct the corresponding  $\hat{J}^2$  and find the simultaneous eigenstates of  $\hat{J}_z$  and  $\hat{J}^2$  which are known to form an orthogonal basis for the Hilbert space. Group these eigenstates by their  $j$ -values and then order them according to their  $m$ -values. Now use this basis to construct the matrix representations of  $\hat{J}_{x,y,z}$ . It should be clear from the expressions for the matrix elements in section 5.6.4 that the matrix representations of these operators will be **block-diagonal** with each block corresponding to a specific value of  $j$ . There are two possibilities:

- The matrices consist of a single block, i.e. only one value of  $j$  appears and the matrix is  $(2j+1) \times (2j+1)$  dimensional. In this case the representation is said to be **irreducible**. *To each value of  $j$  there corresponds a unique irreducible representation.*
- The matrices consist of more than one block, and so different values of  $j$  are present or the same value occurs multiple times. Such a representation is said to be **reducible**.

Note that in the reducible case the  $(2j + 1)$  basis states associated with a particular block span a subspace which is invariant under  $\hat{J}_{x,y,z}$ . In fact, the restriction of  $\hat{J}_{x,y,z}$  to such a subspaces is just the irreducible representation labelled by  $j$ . In other words, each block appearing in the matrix representation of  $\hat{J}_{x,y,z}$  is just the matrix representation of some irreducible representation. The irreducible representations therefore act as the basic “building blocks” from which all other representations are constructed.

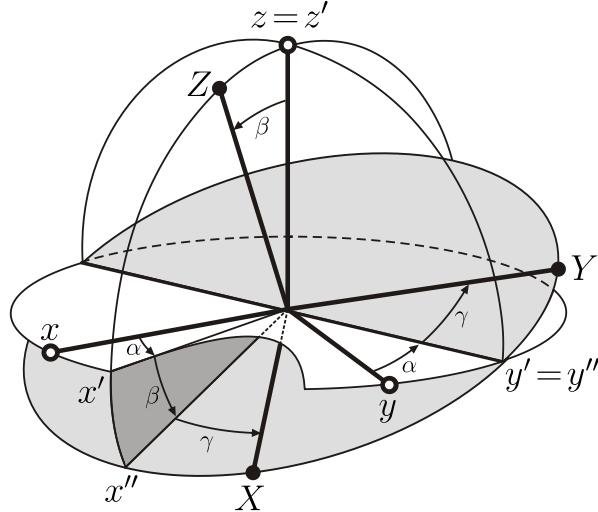


Figure 5.1: The Euler angle parametrisation of a rotation.

## 5.7 Matrix elements of rotation operators

We have classified all the representations of the angular momentum operators by explicitly calculating the matrix elements of the irreducible representations in the basis of  $\hat{J}^2$  and  $\hat{J}_z$  eigenstates. This approach can also be applied to the corresponding rotation operators, which are exponential functions of the angular momentum operators.

### 5.7.1 Euler angles

The axis-angle parametrization is not well-suited to the calculation of matrix elements. A more convenient choice is the set of Euler angles  $(\alpha, \beta, \gamma)$  in terms of which a general rotation can be decomposed into three steps as:

- (1) Rotate anticlockwise around the  $z$ -axis by  $\alpha$ .  $[xyz \rightarrow x'y'z']$  ( $z' = z$ )
- (2) Rotate anticlockwise around the  $y'$ -axis by  $\beta$ .  $[x'y'z' \rightarrow x''y''z'']$  ( $y'' = y'$ )
- (3) Rotate anticlockwise around the  $z''$ -axis by  $\gamma$ .  $[x''y''z'' \rightarrow XYZ]$  ( $Z = z''$ )

The rotation matrix can be expressed as<sup>4</sup>

$$R(\alpha, \beta, \gamma) = R_{z''}(\gamma)R_{y'}(\beta)R_z(\alpha) = R_z(\alpha)R_y(\beta)R_z(\gamma) \quad (5.55)$$

and so the corresponding operator is

$$\hat{D}(R(\alpha, \beta, \gamma)) = e^{-i\alpha\hat{J}_z/\hbar}e^{-i\beta\hat{J}_y/\hbar}e^{-i\gamma\hat{J}_z/\hbar}. \quad (5.56)$$

---

<sup>4</sup>See Sakurai for details.

### 5.7.2 Calculating the matrix elements

Since  $\hat{J}_x$ ,  $\hat{J}_y$  and  $\hat{J}_z$  commute with  $\hat{J}^2$  the same is true of *any* rotation operator, i.e.

$$[\hat{D}(R), \hat{J}^2] = 0. \quad (5.57)$$

This implies that  $\hat{D}(R)$  cannot change the  $j$ -value of a state. To be precise,  $\hat{D}(R)$  leaves the eigenspaces of  $\hat{J}^2$  invariant and so if  $|\phi\rangle$  is an eigenstate of  $\hat{J}^2$  with eigenvalue  $j(j+1)\hbar^2$  then so is  $\hat{D}(R)|\phi\rangle$ . This just reflects the fact that a rotation cannot change the magnitude of the angular momentum, but only its orientation. All the matrix elements of  $\hat{D}(R)$  between states with different  $j$ -values therefore vanish, and so  $\hat{D}(R)$  has exactly the same block-diagonal structure as we encountered in the matrix representations of the angular momentum operators. Each block is labelled by some value of  $j$  and corresponds to an irreducible representation of the rotation group. Inside each block the matrix elements are

$$D_{m',m}^j(R) = \langle j, m' | \hat{D}(R) | j, m \rangle \quad (5.58)$$

where  $D_{m',m}^j(R)$  is known as the Wigner  $D$ -function. These matrix elements characterise how the basis states transform under rotations:

$$\hat{D}(R)|j, m\rangle = \left[ \sum_{m'=-j}^{+j} |j, m'\rangle \langle j, m'| \right] \hat{D}(R)|j, m\rangle = \sum_{m'=-j}^{+j} D_{m',m}^j(R)|j, m'\rangle. \quad (5.59)$$

Again we see that a rotation cannot affect the  $j$ -label of a state, but it does mix up states with different  $m$ -labels, i.e. it changes the angular momentum's orientation but not its magnitude. To calculate the  $D$ -functions we parametrize the rotation operator as in (5.56) to obtain

$$D_{m',m}^j(\alpha, \beta, \gamma) = \langle j, m' | \hat{D}(R(\alpha, \beta, \gamma)) | j, m \rangle = e^{-i\alpha m'} e^{-i\gamma m} \langle j, m' | e^{-i\beta \hat{J}_y/\hbar} | j, m \rangle \quad (5.60)$$

$$= e^{-i\alpha m'} e^{-i\gamma m} d_{m',m}^j(\beta). \quad (5.61)$$

Here  $d_{m',m}^j(\beta)$  is known as the Wigner small- $d$  function for which Sakurai gives a complicated general expression. For small values of  $j$  we will exploit special properties of the angular momentum operators to calculate this function directly.

## 5.8 Spin-1/2 rotations

The matrix representation of the angular momentum operators in the  $j = 1/2$  representation is

$$\hat{J}_i = \hat{S}_i = \frac{\hbar}{2} \hat{\sigma}_i \quad (5.62)$$

where

$$\hat{\sigma}_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad \hat{\sigma}_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \quad \hat{\sigma}_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad (5.63)$$

are the Pauli spin matrices. It is easily seen that

$$\{\hat{\sigma}_i, \hat{\sigma}_j\} = 2I\delta_{i,j} \quad \text{and} \quad (\hat{n} \cdot \hat{\vec{\sigma}})^2 = I \quad (5.64)$$

where  $\hat{n}$  is any unit vector. These two properties allow the expression for the rotation operator to be simplified as

$$\hat{D}(R_{\hat{n}}(\phi)) = e^{-i\phi\hat{n}\cdot\hat{J}/\hbar} = e^{-i\phi\hat{n}\cdot\hat{\sigma}/2} = \cos(-\phi\hat{n}\cdot\hat{\sigma}/2) + i\sin(-\phi\hat{n}\cdot\hat{\sigma}/2) \quad (5.65)$$

$$= \cos(\phi/2) - i\sin(\phi/2)(\hat{n}\cdot\hat{\sigma}) \quad (5.66)$$

Let us see what this expression reduces to when we set  $\phi$  equal to 0,  $2\pi$  or  $4\pi$ . Note that all three these angles correspond to the same rotation matrix, namely the identity. Substituting these angles into (5.66) produces

$$\hat{D}(R_{\hat{n}}(0)) = \hat{I} \quad \hat{D}(R_{\hat{n}}(2\pi)) = -\hat{I} \quad \hat{D}(R_{\hat{n}}(4\pi)) = \hat{I}. \quad (5.67)$$

*Apparently a spin-1/2 state is not invariant under a  $2\pi$  rotation, but instead undergoes a change of sign!* This is both counter-intuitive and mathematically puzzling. Since  $R_{\hat{n}}(0)$  and  $R_{\hat{n}}(2\pi)$  are exactly the same element of  $SO(3)$  it appears that this representation associates two *different* operators with the same rotation. For  $j = 1/2$ , in fact for all half-integer  $j$ , the representation we constructed for  $so(3)$  therefore does not produce a  $SO(3)$  representation which conforms to the requirements set out in section 5.4. Put differently, the angular momentum operators we constructed obey the correct commutation relations, but when we insert them into the exponential formula we get rotation operators which behave very strangely. We are now confronted with two questions, one physical and the other mathematical. On the physics side, one may ask whether this minus sign is actually “physically real”, i.e. experimentally observable, or just a mathematical anomaly. If we would rotate a single spin-1/2 particle by  $2\pi$  degrees then the resulting minus sign amounts to a global phase which is not physically observable. In other words,  $|\psi\rangle$  and  $-|\psi\rangle$  represent exactly the same physical state and the minus sign is irrelevant. However, consider the experiment shown in the figure.<sup>5</sup> A spin-polarized beam of neutrons is split into two sub-beams. The top beam passed through a region in which a magnetic field is present and this produces a rotation of each neutron by a specific angle. The two beams are then combined and allowed to interfere. This interference will depend on the rotation angle and can be probed by comparing the detection rates of neutrons in the two detectors. If the neutrons in the top beam are rotated by  $2\pi$  degrees the change in sign amounts to a phase difference which affects how the two beams interfere. This has been verified experimentally. Furthermore, this effect is absent for a  $4\pi$  rotation. It is an experimental fact that for spin-1/2 degrees of freedom a rotation by  $\phi = 2\pi$  is not equivalent to one by  $\phi = 0$  or  $\phi = 4\pi$ . The rotation operator derived in (5.66) is indeed correct. However, this implies that the rotation group  $SO(3)$  does not describe the rotations of spin-1/2 degrees of freedom. (In retrospect, it is not obvious that  $SO(3)$ , which represent rotations of regular vectors in  $\mathbb{R}^3$ , should be applicable to purely quantum mechanical degrees of freedom such as spin.) Despite this we have arrived at the correct result, but now it is not obvious why everything worked out. What are the  $\hat{D}(R)$  operators in (5.66) if they are not representations of  $SO(3)$ ? To answer this, remember that we constructed the representations of the group  $SO(3)$  by first constructing representations of the algebra  $so(3)$ . Whereas every Lie group has a unique algebra associated with it the

---

<sup>5</sup>For details see “The neutron interferometer as a device for illustrating the strange behaviour of quantum systems”, D Greenberger, Reviews of Modern Physics, Volume 55, Issue 4, October 1983, pp.875-905

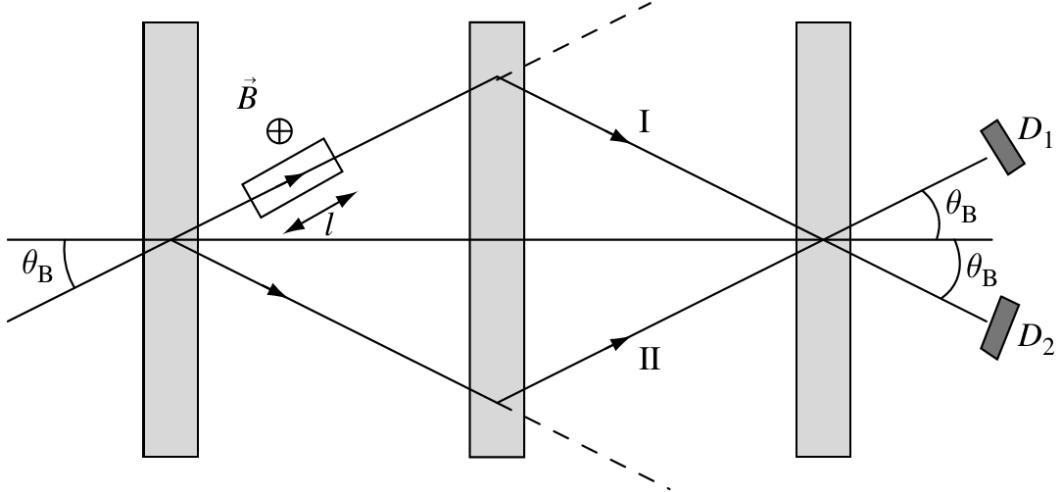


Figure 5.2: Neutron interferometry experiment.

inverse is not true; there may be many different groups all with the same Lie algebra. If another group shares the same algebra as  $SO(3)$  then any representation of  $so(3)$  is automatically also one of this other group's algebra. One such group is  $SU(2)$ , the group of complex,  $2 \times 2$  unitary matrices with a determinant of one. The Lie algebra  $su(2)$  is identical to  $so(3)$ . (In the sense that the commutation relations are the same.) Any representation of  $so(3)$  is therefore also one of  $su(2)$ . In fact, plugging such a representation into the exponential formula  $\hat{D}(R_{\hat{n}}(\phi)) = \exp[-i\phi\hat{n} \cdot \hat{\vec{J}}/\hbar]$  always yields a representation of  $SU(2)$ , but only produces one of  $SO(3)$  for integer values of  $j$ . We will show this in the next section. The rotation of spin degrees of freedom is therefore not described by  $SO(3)$ , but rather by  $SU(2)$ . Luckily, since their Lie algebras are the same we have already found all the representations of  $su(2)$  and thereby also of  $SU(2)$  and  $SO(3)$

## 5.9 Orbital angular momentum in the position representation

### 5.9.1 Wave functions

In section 5.5 we constructed a representation of  $SO(3)$  on the  $\hat{\vec{x}} - \hat{\vec{p}}$  Hilbert space by defining  $\hat{D}(R)|\vec{x}\rangle = |R\vec{x}\rangle$ . We then showed that

$$\hat{D}(R_{\hat{n}}(\phi)) = e^{-i\phi\hat{n} \cdot \hat{\vec{L}}/\hbar} \quad \text{with} \quad \hat{\vec{L}} = \hat{\vec{x}} \times \hat{\vec{p}}. \quad (5.68)$$

Since  $\hat{\vec{L}}$  is a particular representation of the angular momentum algebra there must exist a basis of simultaneous eigenstates of  $\hat{L}^2$  and  $\hat{L}_z$  such that

$$\hat{L}^2|l, m\rangle = l(l+1)\hbar^2|l, m\rangle \quad (5.69)$$

$$\hat{L}_z|l, m\rangle = m\hbar|l, m\rangle \quad m = -l, \dots, +l \quad (5.70)$$

Note that an additional label, which we suppress for the moment, is required to label these states uniquely. We want to determine which values of  $l$  appear in this representation. To do so we investigate how the coordinate space wave function of a  $|l, m\rangle$  state transforms under a  $2\pi$ -rotation around the  $z$ -axis. Since  $R_z(2\pi)\vec{x} = \vec{x}$  for all  $\vec{x} \in \mathbb{R}^3$  we see that

$$\langle \vec{x}|l, m\rangle = \langle R_z(2\pi)\vec{x}|l, m\rangle = \langle \vec{x}|e^{+i2\pi\hat{L}_z/\hbar}|l, m\rangle = e^{2\pi im}\langle \vec{x}|l, m\rangle. \quad (5.71)$$

Unless  $e^{2\pi im} = 1$  the wave function, and therefore the state itself, must be zero. It follows that  $l$  can take only integer values. The general form of  $\langle \vec{x}|l, m\rangle$  is then

$$\langle \vec{x}|\alpha, l, m\rangle = f_{\alpha,l}(r)Y_m^l(\theta, \phi) \quad \text{for } l = 0, 1, 2, \dots \quad (5.72)$$

where  $f_{\alpha,l}(r)$  is a radial wave function and  $Y_m^l(\theta, \phi)$  a spherical harmonic. The additional “non-angular momentum” label  $\alpha$  only affects the radial dependence of the wave function. This should be well known to you from your study of the hydrogen atom. The spherical harmonics satisfy the orthogonality relation

$$\int_0^{2\pi} d\phi \int_0^\pi d\theta \sin \theta [Y_m^l(\theta, \phi)]^* Y_{m'}^{l'}(\theta, \phi) = \delta_{m,m'}\delta_{l,l'}. \quad (5.73)$$

A detailed definition of these functions can be found in appendix D.

### 5.9.2 Observables

In the coordinate representation the orbital angular momentum operator is

$$\hat{\vec{L}} = \hat{\vec{x}} \times \hat{\vec{p}} = -i\hbar(\vec{x} \times \vec{\nabla}). \quad (5.74)$$

In Cartesian coordinates its components are

$$\hat{L}_x = -i\hbar \left[ y \frac{\partial}{\partial z} - z \frac{\partial}{\partial y} \right] \quad \hat{L}_y = -i\hbar \left[ z \frac{\partial}{\partial x} - x \frac{\partial}{\partial z} \right] \quad \hat{L}_z = -i\hbar \left[ x \frac{\partial}{\partial y} - y \frac{\partial}{\partial x} \right] \quad (5.75)$$

while in spherical coordinates we have

$$\hat{L}_x = -i\hbar \left[ -\sin \phi \frac{\partial}{\partial \theta} - \cot \theta \cos \phi \frac{\partial}{\partial \phi} \right] \quad (5.76)$$

$$\hat{L}_y = -i\hbar \left[ \cos \phi \frac{\partial}{\partial \theta} - \cot \theta \sin \phi \frac{\partial}{\partial \phi} \right] \quad (5.77)$$

$$\hat{L}_z = -i\hbar \frac{\partial}{\partial \phi} \quad (5.78)$$

$$\hat{L}^2 = -\hbar^2 \left[ \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2} + \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left( \sin \theta \frac{\partial}{\partial \theta} \right) \right] \quad (5.79)$$

Of course, the eigenfunctions of  $\hat{L}_z$  and  $\hat{L}^2$  are the spherical harmonics:

$$\hat{L}^2 Y_m^l(\theta, \phi) = l(l+1)\hbar^2 Y_m^l(\theta, \phi) \quad l = 0, 1, 2, \dots \quad (5.80)$$

$$\hat{L}_z Y_m^l(\theta, \phi) = m\hbar Y_m^l(\theta, \phi) \quad m = -l, \dots, +l \quad (5.81)$$

## 5.10 Angular momentum addition

### 5.10.1 Background

**Note:** Make sure you have studied section 1.5 on tensor product spaces.

Consider a composite system consisting of two subsystems with Hilbert spaces  $\mathcal{H}_1$  and  $\mathcal{H}_2$  and angular momentum operators  $\hat{\vec{J}}_1$  and  $\hat{\vec{J}}_2$ . The precise nature of these subsystems is unimportant at this stage; they may be individual particles or collections of particles themselves. In fact,  $\mathcal{H}_1$  and  $\mathcal{H}_2$  may also describe the spin and spatial degrees of freedom of a single particle. The Hilbert space of the composite system is  $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$  and the angular momentum operators can be defined directly on  $\mathcal{H}$  through the identifications  $\hat{\vec{J}}_1 \equiv \hat{\vec{J}}_1 \otimes \hat{I}$  and  $\hat{\vec{J}}_2 \equiv \hat{I} \otimes \hat{\vec{J}}_2$ . Our goal is to understand how these two angular momenta can be coupled, i.e. “added together”, to produce a well-defined *total* angular momentum. Mathematically this amounts to finding the eigenstates and eigenvalues of the operator representing the square of the total angular momentum.

### 5.10.2 Rotations of composite systems

We claim that  $\hat{\vec{J}} = \hat{\vec{J}}_1 + \hat{\vec{J}}_2$  represents the total angular momentum of the composite system. To prove this we will again make use of our identification of physical observables as the generators of transformations on the Hilbert space. This leads us to consider rotations of the composite system. The generator of these rotation should, by definition, be the total angular momentum.

Let  $\hat{D}_1(R)$  and  $\hat{D}_2(R)$  denote the rotation operators for the two subsystems. The natural definition for the rotation operator of the composite system is

$$\hat{D}(R) = \hat{D}_1(R) \otimes \hat{D}_2(R). \quad (5.82)$$

It is clear that, like  $\hat{D}_1(R)$  and  $\hat{D}_2(R)$ ,  $\hat{D}(R)$  is a representations of  $SO(3)$  and so there is a representation of the angular momentum algebra associated with it. To find the latter we expand the infinitesimal rotation operator  $\hat{D}(R_{\hat{n}}(\delta\phi))$  as

$$\hat{D}(R_{\hat{n}}(\delta\phi)) = e^{-i\delta\phi \hat{n} \cdot \hat{\vec{J}}_1/\hbar} \otimes e^{-i\delta\phi \hat{n} \cdot \hat{\vec{J}}_2/\hbar} \quad (5.83)$$

$$= (\hat{I}_1 - i\delta\phi \hat{n} \cdot \hat{\vec{J}}_1/\hbar) \otimes (\hat{I}_2 - i\delta\phi \hat{n} \cdot \hat{\vec{J}}_2/\hbar) + \mathcal{O}(\delta\phi^2) \quad (5.84)$$

$$= \hat{I} - i\delta\phi \hat{n} \cdot [\hat{\vec{J}}_1 + \hat{\vec{J}}_2]/\hbar + \mathcal{O}(\delta\phi^2) \quad (5.85)$$

from which we can immediately conclude that  $\hat{\vec{J}} = \hat{\vec{J}}_1 + \hat{\vec{J}}_2$  is the generator of rotations for the composite system and therefore also the total angular momentum. For future reference we note that

$$\hat{J}^2 = \hat{\vec{J}} \cdot \hat{\vec{J}} = \hat{J}_1^2 + \hat{J}_2^2 + 2\hat{J}_1 \cdot \hat{J}_2. \quad (5.86)$$

### 5.10.3 Eigenstates of the total angular momentum

We will fix the magnitudes of the subsystem's angular momenta to  $j_1$  and  $j_2$  and consider only the corresponding subspaces, i.e.

$$\mathcal{H}_1 = \text{span}\{|j_1, m_1\rangle : m_1 = -j_1, \dots, +j_1\} \quad (5.87)$$

$$\mathcal{H}_2 = \text{span}\{|j_2, m_2\rangle : m_2 = -j_2, \dots, +j_2\} \quad (5.88)$$

and so

$$\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2 = \text{span}\{|j_1, m_1; j_2, m_2\rangle : m_{1,2} = -j_{1,2}, \dots, +j_{1,2}\}. \quad (5.89)$$

Note that  $\dim(\mathcal{H}_i) = 2j_i + 1$  and  $\dim(\mathcal{H}) = (2j_1 + 1)(2j_2 + 1)$ . Since the values of  $j_1$  and  $j_2$  are fixed *all* the states in  $\mathcal{H}$  are simultaneous eigenstates of  $\hat{J}_1^2$  and  $\hat{J}_2^2$ . There are two sets of commuting operators acting on  $\mathcal{H}$ :

$$\textbf{Set 1: } \{\hat{J}_1^2, \hat{J}_2^2, \hat{J}_{1z}, \hat{J}_{2z}\} \quad \text{and} \quad \textbf{Set 2: } \{\hat{J}_1^2, \hat{J}_2^2, \hat{J}^2, \hat{J}_z\}. \quad (5.90)$$

Each of these sets defines a particular orthogonal basis of simultaneous eigenstates which are labelled by a set of quantum numbers. See section 1.6.3. The first produces the *uncoupled* or tensor product basis:

$$\mathcal{B}_u = \{|j_1, m_1; j_2, m_2\rangle : m_{1,2} = -j_{1,2}, \dots, +j_{1,2}\} \quad (5.91)$$

for which

$$\hat{J}_1^2 |j_1, m_1; j_2, m_2\rangle = j_1(j_1 + 1)\hbar^2 |j_1, m_1; j_2, m_2\rangle \quad (5.92)$$

$$\hat{J}_2^2 |j_1, m_1; j_2, m_2\rangle = j_2(j_2 + 1)\hbar^2 |j_1, m_1; j_2, m_2\rangle \quad (5.93)$$

$$\hat{J}_{1z} |j_1, m_1; j_2, m_2\rangle = m_1 \hbar |j_1, m_1; j_2, m_2\rangle \quad (5.94)$$

$$\hat{J}_{2z} |j_1, m_1; j_2, m_2\rangle = m_2 \hbar |j_1, m_1; j_2, m_2\rangle \quad (5.95)$$

The second set defines the *coupled* basis which consists of eigenstates of the square and the  $z$ -component of the total angular momentum:

$$\mathcal{B}_c = \{|j_1, j_2; j, m\rangle : j = ? \quad m = -j, \dots, +j\} \quad (5.96)$$

for which

$$\hat{J}_1^2 |j_1, j_2; j, m\rangle = j_1(j_1 + 1)\hbar^2 |j_1, j_2; j, m\rangle \quad (5.97)$$

$$\hat{J}_2^2 |j_1, j_2; j, m\rangle = j_2(j_2 + 1)\hbar^2 |j_1, j_2; j, m\rangle \quad (5.98)$$

$$\hat{J}^2 |j_1, j_2; j, m\rangle = j(j + 1)\hbar^2 |j_1, j_2; j, m\rangle \quad (5.99)$$

$$\hat{J}_z |j_1, j_2; j, m\rangle = m \hbar |j_1, j_2; j, m\rangle \quad (5.100)$$

The possible values of  $j$  are yet to be determined. We note that since  $\hat{J}_z = \hat{J}_{1z} + \hat{J}_{2z}$  the elements of  $\mathcal{B}_u$  are also eigenstates of  $\hat{J}_z$ . However, whereas the uncoupled states are eigenstates of  $\hat{J}_{1z}$  and  $\hat{J}_{2z}$  individually, the coupled states are only eigenstates of the sum  $\hat{J}_z = \hat{J}_{1z} + \hat{J}_{2z}$ . Since  $\mathcal{B}_u$  and  $\mathcal{B}_c$  are two different bases for the same Hilbert space we can express one in terms of the other as

$$|j_1, j_2; j, m\rangle = \sum_{m_1, m_2} \langle j_1, m_1; j_2, m_2 | j_1, j_2; j, m \rangle |j_1, m_1; j_2, m_2\rangle \quad (5.101)$$

where  $\langle j_1, m_1; j_2, m_2 | j_1, j_2; j, m \rangle$  is known as a Clebsch-Gordon coefficient (CGC). In other words, the CGCs are just the expansion coefficients of the coupled basis in terms of the uncoupled one. Since  $j_1$  and  $j_2$  are fixed we can use the more compact notation:

$$|j, m\rangle_c = \sum_{m_1, m_2} {}_u\langle m_1, m_2 | j, m \rangle_c |m_1, m_2\rangle_u \quad (5.102)$$

### 5.10.4 Selection rules for Clebsch-Gordon coefficients

We will show that CGCs exhibit the following important property:

#### Selection rules for CGCs:

If  $m \neq m_1 + m_2$  or  $j \notin \{|j_1 - j_2|, |j_1 - j_2| + 1, \dots, j_1 + j_2\}$  then

$${}_u\langle m_1, m_2 | j, m \rangle_c = \langle j_1, m_1; j_2, m_2 | j_1, j_2; j, m \rangle = 0. \quad (5.103)$$

The first condition says that: “To construct  $|j, m\rangle_c$  we only need uncoupled states  $|m_1, m_2\rangle_u$  for which  $m_1 + m_2 = m$ .” To prove this just note that  $\hat{J}_z |j, m\rangle_c = m\hbar |j, m\rangle_c$  and  $\hat{J}_z |m_1, m_2\rangle_u = (m_1 + m_2)\hbar |m_1, m_2\rangle_u$  and so  ${}_u\langle m_1, m_2 | j, m \rangle_c$  is zero whenever the eigenvalues are different. The second condition states that: “The possible values of  $j$ , which determines the total angular momentum, run from  $|j_1 - j_2|$  up to  $j_1 + j_2$  in integer steps.” To prove this condition requires more effort. We first provide an algorithm for constructing the coupled basis states in terms of the uncoupled ones. Of course, this is equivalent to calculating the CGCs.

### 5.10.5 Constructing the coupled basis

The following simple result will be useful:

#### The top step theorem (TST)

If  $|\phi\rangle \in \mathcal{H}$  is such that  $\hat{J}_+ |\phi\rangle = 0$  and  $\hat{J}_z |\phi\rangle = m\hbar |\phi\rangle$  when  $|\phi\rangle = |jj\rangle_c$  with  $j = m$ .

In other words, if  $|\phi\rangle$  is an eigenstate of  $\hat{J}_z$  which sits at the top of the ladder, meaning that  $\hat{J}_+ |\phi\rangle = 0$ , then  $|\phi\rangle$  is automatically an eigenstate of  $\hat{J}^2$ , i.e. a coupled basis state. To prove this just note that, since  $\hat{J}_+ |\phi\rangle = 0$  and  $\hat{J}_z |\phi\rangle = m\hbar |\phi\rangle$ , we have

$$\hat{J}^2 |\phi\rangle = (\hat{J}_- \hat{J}_+ + \hbar \hat{J}_z + \hat{J}_z^2) |\phi\rangle = m(m+1)\hbar^2 |\phi\rangle \quad (5.104)$$

and so  $|\phi\rangle = |j, j\rangle_c$  with  $j = m$ .

#### Construction algorithm

1. First group the uncoupled basis states according to the value of  $m_1 + m_2$ , i.e. their eigenvalues with respect to  $\hat{J}_z = \hat{J}_{1z} + \hat{J}_{2z}$ . This is shown in Table I. The states in each row form a basis for the corresponding eigenspace of  $\hat{J}_z$ . Since  ${}_u\langle m_1, m_2 | j, m \rangle_c = 0$  when  $m_1 + m_2 \neq m$  it follows that each coupled basis state is an element of one of these eigenspaces.
2. The highest possible value of  $m = m_1 + m_2$  is  $m_{max} = j_1 + j_2$ . This eigenspace is one-dimensional with basis state  $|j_1, j_2\rangle_u$ . By the TST this is precisely the

coupled state  $|j_1 + j_2, j_1 + j_2\rangle_c$  which also has the largest possible value of  $j$ , namely  $j_{max} = j_1 + j_2$ . Now apply  $\hat{J}_-$  repeatedly to produce all the other coupled states with  $j = j_{max}$ . In this way we obtain the first column in Table II.

3. The  $m = m_{max} - 1$  eigenspace is two dimensional. It contains  $|j_{max}, j_{max} - 1\rangle_c$ , which has already been found, and one other state  $|\alpha\rangle$  orthogonal to it. This state can clearly not be a  $j = j_{max}$  state as well, and so  $\hat{J}_+|\alpha\rangle = 0$ . According to the TST this implies that  $|\alpha\rangle$  is the coupled state  $|j_{max} - 1, j_{max} - 1\rangle_c$ , i.e. the top state in the  $j = j_{max} - 1$  ladder. Again apply  $\hat{J}_-$  repeatedly to produce all the  $j = j_{max} - 1$  states.
4. Continue in this way to construct all the coupled basis states for  $j$  ranging from  $j_{max} = j_1 + j_2$  down to some  $j_{min}$ . The process terminates automatically once a full set of coupled states have been constructed in each  $m$ -eigenspace.

**Note:** The states in each column of Table II are unique only up to a global phase. We adopt the convention that  ${}_u\langle j_1, j - j_1 | j, j \rangle_c$  is chosen to be real and positive.

Based on this algorithm the following is clear: *The possible values of  $j$  range from  $j_{max} = j_1 + j_2$  down to some  $j_{min}$  in integer steps. Each value of  $j$  in this range occurs exactly once.* All that remains is to determine  $j_{min}$ . To do so, remember that the number of coupled basis states must equal the dimension of  $\mathcal{H}$ , which is  $(2j_1+1)(2j_2+1)$ . Adding up the dimensions of the  $\hat{J}^2$  eigenspaces gives

$$\sum_{j=j_{min}}^{j_1+j_2} (2j+1) = (j_1 + j_2 + 1)^2 - j_{min}^2. \quad (5.105)$$

For this to equal  $(2j_1+1)(2j_2+1)$  requires that

$$j_{min}^2 = (j_1 + j_2 + 1)^2 - (2j_1+1)(2j_2+1) = (j_1 - j_2)^2 \quad (5.106)$$

and therefore  $j_{min} = |j_1 - j_2|$ . This concludes the derivation of the selection rules for the CGCs.

Table I

Basis for the $m$ -eigenspace of $\hat{J}_z = \hat{J}_{z,1} + \hat{J}_{z,2}$	
$m = m_{max} = j_1 + j_2$	$ j_1, j_2\rangle_u$
$m = j_1 + j_2 - 1$	$ j_1 - 1, j_2\rangle_u \quad  j_1, j_2 - 1\rangle_u$
$m = j_1 + j_2 - 2$	$ j_1 - 2, j_2\rangle_u \quad  j_1 - 1, j_2 - 1\rangle_u \quad  j_1, j_2 - 2\rangle_u$
$\vdots$	$\vdots$
$m = -j_1 - j_2 + 2$	$ -j_1 + 2, -j_2\rangle_u \quad  -j_1 + 1, -j_2 + 1\rangle_u \quad  -j_1, -j_2 + 2\rangle_u$
$m = -j_1 - j_2 + 1$	$ -j_1 + 1, -j_2\rangle_u \quad  -j_1, -j_2 + 1\rangle_u$
$m = m_{min} = -j_1 - j_2$	$ -j_1, -j_2\rangle_u$

Table II

	$j = j_{max} = j_1 + j_2$	$j = j_1 + j_2 - 1$	$j = j_1 + j_2 - 2$	$\dots$	$j = j_{min}$
$m = m_{max} = j_1 + j_2$	$ j_1 + j_2, j_1 + j_2\rangle_c =  j_1, j_2\rangle_u$				
$m = j_1 + j_2 - 1$	$ j_1 + j_2, j_1 + j_2 - 1\rangle_c$	$ j_1 + j_2 - 1, j_1 + j_2 - 1\rangle_c$			
$m = j_1 + j_2 - 2$	$ j_1 + j_2, j_1 + j_2 - 2\rangle_c$	$ j_1 + j_2 - 1, j_1 + j_2 - 2\rangle_c$	$ j_1 + j_2 - 2, j_1 + j_2 - 2\rangle_c$	$\dots$	
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\dots$	
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\dots$	$ j_{min}, j_{min}\rangle_c$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\dots$	$\vdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\dots$	$ j_{min}, -j_{min}\rangle_c$
$m = -j_1 - j_2 + 2$	$ j_1 + j_2, -j_1 - j_2 + 2\rangle_c$	$ j_1 + j_2 - 1, -j_1 - j_2 + 2\rangle_c$	$ j_1 + j_2 - 2, -j_1 - j_2 + 2\rangle_c$	$\dots$	
$m = -j_1 - j_2 + 1$	$ j_1 + j_2, -j_1 - j_2 + 1\rangle_c$	$ j_1 + j_2 - 1, -j_1 - j_2 + 1\rangle_c$			
$m = m_{min} = -j_1 - j_2$	$ j_1 + j_2, -j_1 - j_2\rangle_c =  -j_1, -j_2\rangle_u$				
	dim = $2(j_1 + j_2) + 1$	dim = $2(j_1 + j_2 - 1) + 1$	dim = $2(j_1 + j_2 - 2) + 1$	$\dots$	dim = $2j_{min} + 1$

Spin-1/2

Basis for the $m$ -eigenspace of $\hat{J}_z = \hat{J}_{z,1} + \hat{J}_{z,2}$	
$m = m_{max} = 1$	$ +1/2, +1/2\rangle_u$
$m = 0$	$ +1/2, -1/2\rangle_u \quad  -1/2, +1/2\rangle_u$
$m = m_{min} = -1$	$ -1/2, -1/2\rangle_u$

	$j = j_{max} = 1$	$j = j_{min} = 0$
$m = m_{max} = 1$	$ 1, 1\rangle_c =  1/2, 1/2\rangle_u$	
$m = 0$	$ 1, 0\rangle_c = ( 1/2, -1/2\rangle_u +  -1/2, 1/2\rangle_u)/\sqrt{2}$	$ 0, 0\rangle_c = ( 1/2, -1/2\rangle_u -  -1/2, 1/2\rangle_u)/\sqrt{2}$
$m = m_{min} = -1$	$ 1, -1\rangle_c =  -1/2, -1/2\rangle_u$	

## 5.11 Multi-component wave functions

Thus far we have always considered the spatial and spin degrees of freedom of particles separately. For a single spin-1/2 particle the corresponding Hilbert spaces are  $\mathcal{H}_{xp} = \text{span}\{|\vec{x}\rangle : \vec{x} \in \mathbb{R}^3\}$  and  $\mathcal{H}_{spin} = \text{span}\{|\pm\rangle = |z, \pm\rangle\}$ . The Hilbert space describing both degrees of freedom is therefore

$$\mathcal{H} = \mathcal{H}_{xp} \otimes \mathcal{H}_{spin} = \text{span}\{|\vec{x}\rangle \otimes |\sigma\rangle = |\vec{x}, \sigma\rangle : \vec{x} \in \mathbb{R}^3, \sigma = \pm\}. \quad (5.107)$$

We can resolve the identity on  $\mathcal{H}$  as

$$\hat{I} = \sum_{\sigma=\pm} \int d\vec{x} |\vec{x}, \sigma\rangle \langle \vec{x}, \sigma| \quad (5.108)$$

and operators acting on  $\mathcal{H}_{xp}$  and  $\mathcal{H}_{spin}$  can be defined on  $\mathcal{H}$  in the usual way:

$$A(\hat{\vec{x}}, \hat{\vec{p}})|\vec{x}, \sigma\rangle = (A(\hat{\vec{x}}, \hat{\vec{p}})|\vec{x}\rangle) \otimes |\sigma\rangle \quad (5.109)$$

$$B(\hat{\vec{S}})|\vec{x}, \sigma\rangle = |\vec{x}\rangle \otimes (B(\hat{\vec{S}})|\sigma\rangle) \quad (5.110)$$

The total angular momentum is  $\hat{\vec{J}} = \hat{\vec{L}} + \hat{\vec{S}}$ . This operator generates rotations of the particle's spatial and spin degrees of freedom simultaneously

Note that there are two distinct spatial wave functions associated with a state  $|\phi\rangle$ , namely  $\phi_{\pm}(\vec{x}) = \langle \vec{x}, \pm | \phi \rangle$ . Using these two functions we form the 2-component wave function

$$\Phi(\vec{x}) = \begin{bmatrix} \phi_+(\vec{x}) \\ \phi_-(\vec{x}) \end{bmatrix} \quad (5.111)$$

which represents  $|\phi\rangle$  in the  $\{|\vec{x}, \sigma\rangle\}$  basis. The inner product between two states  $|\phi\rangle$  and  $|\tilde{\phi}\rangle$  can be expressed as

$$\langle \tilde{\phi} | \phi \rangle = \int d\vec{x} \tilde{\Phi}^\dagger(\vec{x}) \Phi(\vec{x}). \quad (5.112)$$

The orbital angular momentum operators now become differential operators which act on the components of  $\Phi(\vec{x})$  as

$$\hat{L}_i \begin{bmatrix} \phi_+(\vec{x}) \\ \phi_-(\vec{x}) \end{bmatrix} = \begin{bmatrix} \hat{L}_i \phi_+(\vec{x}) \\ \hat{L}_i \phi_-(\vec{x}) \end{bmatrix}. \quad (5.113)$$

The spin operators are represented by the usual  $2 \times 2$  matrices which act on  $\Phi(\vec{x})$  through regular matrix multiplication. In other words,  $\hat{\vec{L}}$  is only concerned with the  $\vec{x}$ -dependence of  $\Phi(\vec{x})$  while  $\hat{\vec{S}}$  only sees its column vector structure. A simultaneous eigenstate of  $\hat{L}^2$ ,  $\hat{S}^2$ ,  $\hat{L}_z$  and  $\hat{S}_z$  can be represented as

$$|\alpha, l, m, s, \sigma\rangle \iff f_\alpha(r) Y_m^l(\theta, \phi) \chi_\sigma^{(1/2)} \quad (5.114)$$

where

$$\chi_+^{(1/2)} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \chi_-^{(1/2)} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad (5.115)$$

represent  $|z, +\rangle$  and  $|z, -\rangle$ . The rotation operator  $\hat{D}(R_{\hat{n}}(\theta)) = e^{-i\theta\hat{n}\cdot\hat{J}/\hbar}$  acts on  $\Phi(\vec{x})$  as

$$e^{-i\theta\hat{n}\cdot\hat{J}/\hbar} \begin{bmatrix} \phi_+(\vec{x}) \\ \phi_-(\vec{x}) \end{bmatrix} = e^{-i\theta\hat{n}\cdot\hat{S}/\hbar} \begin{bmatrix} e^{-i\theta\hat{n}\cdot\hat{L}/\hbar}\phi_+(\vec{x}) \\ e^{-i\theta\hat{n}\cdot\hat{L}/\hbar}\phi_-(\vec{x}) \end{bmatrix} = e^{-i\theta\hat{n}\cdot\hat{S}/\hbar} \begin{bmatrix} \phi_+(R^{-1}\vec{x}) \\ \phi_-(R^{-1}\vec{x}) \end{bmatrix} \quad (5.116)$$

where (5.66) can be used to express  $e^{-i\theta\hat{n}\cdot\hat{S}/\hbar}$  in terms of the Pauli matrices.

## 5.12 Scalar, vector and tensor operators

### 5.12.1 Background

To set the scene we first summarize some basic facts regarding the transformation properties of spatial vectors and kets under rotations. The elements of  $SO(3)$  are rotation matrices which act on regular vectors in three dimensions. Specifically, if  $R \in SO(3)$  and  $\vec{x} \in \mathbb{R}^3$  then if  $\vec{x}' = R\vec{x}$  we have  $x'_i = \sum_j R_{ij}x_j$ . On the Hilbert space rotations are realised in terms of operators  $\hat{D}(R)$  which acts on a  $|j, m\rangle$  state to produce  $\hat{D}(R)|j, m\rangle = \sum_{m'=-j}^j D_{m',m}^j(R)|j, m'\rangle$ . Simply put,  $R$  mixes up the three components of  $\vec{x}$  with coefficients  $R_{ij}$  while  $\hat{D}(R)$  mixes up the  $(2j+1)$  states  $\{|j, m\rangle : m = -j, \dots, j\}$  with coefficients  $D_{m',m}^j(R)$ .

In this section we consider the transformation properties of *operators*, and multi-component operators in particular, under rotations. In reality we are still dealing with representations of  $SO(3)$ , except that the vectors on which the rotation operators act are now no longer kets but operators themselves. On a mathematical level this distinction is little more than a change of notation. We therefore expect that the transformation properties of operators should mimic those of spatial vectors and kets very closely.

### 5.12.2 Definitions

We now identify three classes of multi-component operators based on the way they transform under rotations.

#### Scalar operators

An operator  $\hat{S}$  is called a scalar operator if it is invariant under rotations, i.e. if

$$\hat{D}(R)\hat{S}\hat{D}^\dagger(R) = \hat{D}^\dagger(R)\hat{S}\hat{D}(R) = \hat{S} \quad \text{for all } R \in SO(3). \quad (5.117)$$

This is equivalent to requiring that  $[\hat{S}, \hat{D}(R)] = 0$  for all  $R \in SO(3)$ , which implies that

$$[\hat{S}, \hat{J}_i] = 0. \quad (5.118)$$

A system of which the Hamiltonian is a scalar operator is said to exhibit *spherical symmetry*. Well-known examples include the hydrogen atom and the isotropic simple harmonic oscillator. In fact, any Hamiltonian of the form

$$\hat{H} = \frac{\hat{\vec{p}}^2}{2m} + V(\hat{r}) \quad (5.119)$$

is clearly a scalar operator. The most important consequence of spherical symmetry is that the eigenstates of  $\hat{H}$  can be chosen to be eigenstates of  $\hat{J}^2$  and  $\hat{J}_z$  as well.

## Vector operators

A vector operator  $\hat{\vec{V}} = (\hat{V}_1, \hat{V}_2, \hat{V}_3)$  has three components that transform like the components of a regular  $\mathbb{R}^3$  vector:

$$\hat{D}^\dagger(R)\hat{V}_i\hat{D}(R) = \sum_j R_{ij}\hat{V}_j \quad \text{for all } R \in SO(3). \quad (5.120)$$

This condition can be recast in the more convenient commutator form

$$[\hat{V}_i, \hat{J}_j] = \epsilon_{ijk}i\hbar\hat{V}_k. \quad (5.121)$$

Examples of vector operators include  $\hat{\vec{x}}$ ,  $\hat{\vec{p}}$ ,  $\hat{\vec{S}}$  and  $\hat{\vec{L}}$ .

**Exercise:** Derive (5.121) from (5.120) by considering an infinitesimal rotation.

## Spherical tensor operators

A spherical tensor operator  $\hat{T}^k$  of rank  $k$  has  $2k+1$  components  $\{\hat{T}_q^k : q = -k, \dots, k\}$  which transform like the states  $\{|j=k, m=q\rangle : q = -k, \dots, k\}$  under rotations:

$$\hat{D}(R)\hat{T}_q^k\hat{D}^\dagger(R) = \sum_{q'=-k}^k D_{q',q}^k(R)\hat{T}_{q'}^k. \quad (5.122)$$

Again you can check that this condition is equivalent to the commutation relations

$$[\hat{J}_z, \hat{T}_q^k] = q\hbar\hat{T}_q^k \quad \text{and} \quad [\hat{J}_\pm, \hat{T}_q^k] = \sqrt{k(k+1)-q(q\pm 1)}\hbar\hat{T}_{q\pm 1}^k. \quad (5.123)$$

We will now show that *spherical tensor operators include both scalar and vector operators as special cases*.

- Consider a rank-0 tensor operator with a single component  $\hat{T}_0^0$ . From (5.122) we have

$$\hat{D}(R)\hat{T}_0^0\hat{D}^\dagger(R) = D_{0,0}^0(R)\hat{T}_0^0 = \hat{T}_0^0 \quad (5.124)$$

for all  $R \in SO(3)$ . *Clearly scalar operators are just rank-0 tensor operators.*

- Now consider a vector operator  $\hat{\vec{V}} = (\hat{V}_1, \hat{V}_2, \hat{V}_3) = (\hat{V}_x, \hat{V}_y, \hat{V}_z)$ . Here  $\hat{V}_{x,y,z}$  are the *Cartesian components* of  $\hat{\vec{V}}$ . The *spherical components* of  $\hat{\vec{V}}$  are denoted  $(\hat{V}_{-1}^1, \hat{V}_0^1, \hat{V}_{+1}^1)$  and are defined by

$$\hat{V}_{\pm 1}^1 = \frac{\mp 1}{\sqrt{2}} [\hat{V}_x \pm i\hat{V}_y] \quad \text{and} \quad \hat{V}_0^1 = \hat{V}_z. \quad (5.125)$$

It is straightforward to verify that  $(\hat{V}_{-1}^1, \hat{V}_0^1, \hat{V}_{+1}^1)$  satisfies (5.123) and is a rank-1 tensor operator. *Vector operators are therefore just rank-1 tensor operators expressed in a particular basis, namely spherical rather than Cartesian.*

### 5.12.3 Matrix elements of tensor operators

In this section we consider the matrix elements of tensor operators between eigenstates of  $\hat{J}^2$  and  $\hat{J}_z$ , i.e. matrix elements of the general form  $\langle \alpha', j', m' | \hat{T}_q^k | \alpha, j, m \rangle$  where  $\alpha'$  and  $\alpha$  are additional labels required to specify the states uniquely. The main result in this regard is the **Wigner-Eckart theorem** which states that such a matrix element can be factorized as

$$\langle \alpha', j', m' | \hat{T}_q^k | \alpha, j, m \rangle = \frac{\langle \alpha', j' | \hat{T}^k | \alpha, j \rangle}{\sqrt{2j+1}} {}_u \langle jm; kq | jk; j'm' \rangle_c \quad (5.126)$$

where  $\langle \alpha', j' | \hat{T}^k | \alpha, j \rangle$  is the so-called **reduced matrix element** and  ${}_u \langle jm; kq | jk; j'm' \rangle_c$  is the Clebsch-Gordon coefficient for coupling angular momenta  $j_1 = j$  and  $j_2 = k$  to a state with total angular momentum  $j_{tot} = j'$  and  $m_{tot} = m'$ . See Sakurai for a proof. To understand what this theorem says we must take note of how the various labels appear on the right of (5.126). The reduced matrix element  $\langle \alpha', j' | \hat{T}^k | \alpha, j \rangle$  is not a matrix element in the usual sense, but instead a kind of proportionality factor. It depends on  $\alpha', j', \alpha$  and  $j$  as well as the tensor operator *as a whole*. It is independent of the “orientation” labels  $m, m'$  and  $q$ . The Clebsch-Gordon coefficient depends only on the angular momentum labels and not the tensor operator itself.

The two main applications of this theorem are as follows. First, suppose all the labels except  $\{m', q, m\}$  are fixed, and denote the matrix element by  $\mathcal{M}(m', q, m)$ . Calculating this matrix element for all possible values of  $\{m', q, m\}$  would involve  $(2j'+1)(2k+1)(2j+1)$  separate calculations. However, from (5.126) it is clear that once the reduced matrix element is known any  $\mathcal{M}(m', q, m)$  can be found from a simple calculation of the Clebsch-Gordon coefficient. It follows that it is only necessary to calculate  $\mathcal{M}(m', q, m)$  once for a specific choice of  $\{m', q, m\}$  since, provided the result is non-zero, this would determine the reduced matrix element uniquely.

The second important application relates to selection rules, which are conditions under which the matrix element  $\langle \alpha', j', m' | \hat{T}_q^k | \alpha, j, m \rangle$  is guaranteed to be zero. This is often used to identify forbidden transitions in atomic and nuclear physics. We have:

**Selection rules for tensor operators:** *The matrix element  $\langle \alpha', j', m' | \hat{T}_q^k | \alpha, j, m \rangle$  is zero whenever  $m' \neq m + q$  or  $j' \notin \{|j - k|, |j - k| + 1, \dots, j + k\}$ .*

The proof of this statement is immediate from the selection rules for the Clebsch-Gordon coefficients found in section 5.10.4. It is clear that spherical tensor operators act as generalized ladder operators which can change both the orientation as well as the magnitude of a state’s angular momentum. This can also be understood by noting that the transformation properties of  $\hat{T}_q^k | \alpha, j, m \rangle$  and  $|k, q \rangle \otimes | \alpha, j, m \rangle$  are identical. Applying  $\hat{T}_q^k$  to  $| \alpha, j, m \rangle$  is therefore akin to coupling an angular momentum of  $k$  to one of  $j$ , and in this sense the appearance of the Clebsch-Gordon coefficients are not surprising.

### 5.12.4 Exercise

The five operators

$$\hat{T}_0^2 = (3\hat{z}^2 - \hat{r}^2)/\sqrt{6} \quad \hat{T}_{\pm 1}^2 = \mp\hat{z}(\hat{x} \pm i\hat{y}) \quad \hat{T}_{\pm 2}^2 = (\hat{x} \pm i\hat{y})^2/2 \quad (5.127)$$

constitute the components of a rank-2 tensor operator.

Suppose the matrix element

$$Q = \langle \alpha, j, j | (3\hat{z}^2 - \hat{r}^2) | \alpha, j, j \rangle \quad (5.128)$$

is known. Use the Wigner-Eckart theorem to find an expression for the matrix element

$$\langle \alpha, j, m' | (\hat{x}^2 - \hat{y}^2) | \alpha, j, j \rangle \quad (5.129)$$

in terms of  $Q$  and Clebsch-Gordan coefficients.

# Chapter 6

## Identical Particles

### 6.1 Introduction

In this chapter we will give a very brief overview of some results relating to systems of identical particles. You will explore the formalism of many-body quantum mechanics in more detail in later courses. This presentation mainly follows the book by Ballentine.

Consider a system consisting of  $N$  *identical* particles each with Hilbert space

$$\mathcal{H}_1 = \text{span}\{|\phi_n\rangle : n = 1, 2, 3, \dots\}. \quad (6.1)$$

In what follows we will abbreviate  $|\phi_n\rangle \equiv |n\rangle$ . The Hilbert space of the  $N$ -particle system is then the tensor product of  $N$  copies of  $\mathcal{H}_1$ , namely

$$\mathcal{H}_N = \underbrace{\mathcal{H}_1 \otimes \mathcal{H}_1 \otimes \mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_1}_{N \text{ times}} = \text{span}\{ |n_1 n_2 n_3 \cdots n_N\rangle : n_i = 1, 2, 3, \dots\} \quad (6.2)$$

We define the permutation (or exchange) operators  $\hat{P}_{ij}$  with  $i, j \in \{1, 2, \dots, N\}$  and  $i \neq j$  by

$$\hat{P}_{ij}|n_1 n_2 \dots n_i \dots n_j \dots n_N\rangle = |n_1 n_2 \dots n_j \dots n_i \dots n_N\rangle. \quad (6.3)$$

In other words,  $\hat{P}_{ij}$  exchanges the states of particles  $i$  and  $j$ . Note that  $\hat{P}_{ij}^2 = \hat{I}$  and so  $\hat{P}_{ij}$  is both unitary and Hermitian:

$$\hat{P}_{ij}^\dagger = \hat{P}_{ij}^{-1} = \hat{P}_{ij}. \quad (6.4)$$

The only possible eigenvalues of  $\hat{P}_{ij}$  are therefore  $+1$  and  $-1$ .

### 6.2 Indistinguishability and observables

We now introduce the first of two fundamental principles:

#### Principle 1: The Indistinguishability of Particles

*The interchange of two identical particles can have no physically observable effect.*

This implies that the states  $|\psi\rangle$  and  $\hat{P}_{ij}|\psi\rangle$  cannot be distinguished through *any* experimental observation whatsoever. In particular, the expectation value of any  $N$ -particle observable  $A$  must remain unchanged when two identical particles are exchanged. This requires that

$$\langle\psi|\hat{A}|\psi\rangle = \langle\psi|\hat{P}_{ij}^\dagger\hat{A}\hat{P}_{ij}|\psi\rangle \quad \text{for all } |\psi\rangle \in \mathcal{H}_N. \quad (6.5)$$

This is only possible if  $\hat{A} = \hat{P}_{ij}^\dagger\hat{A}\hat{P}_{ij}$  or equivalently, if  $[\hat{A}, \hat{P}_{ij}] = 0$ , and so we conclude that *all physical observables must be permutation invariant*. Let us consider what this implies for the general form of an  $N$ -particle observable. Suppose  $\hat{\mathcal{O}}_i$  is an observable associated with particle  $i$ . By taking the basis of  $\mathcal{H}_1$  to be eigenstates of  $\hat{\mathcal{O}}$  it is easy to see that

$$\hat{P}_{ij}^\dagger\hat{\mathcal{O}}_i\hat{P}_{ij} = \hat{\mathcal{O}}_j. \quad (6.6)$$

A general  $N$ -particle observable  $\hat{A}(\hat{x}_1, \dots, \hat{x}_N, \hat{p}_1, \dots, \hat{p}_N, \hat{S}_1, \dots, \hat{S}_N)$  must therefore be invariant under the exchange  $\hat{x}_i \leftrightarrow \hat{x}_j$ ,  $\hat{p}_i \leftrightarrow \hat{p}_j$  and  $\hat{S}_i \leftrightarrow \hat{S}_j$ . For example, consider the generic form of the Hamiltonian for an  $N$ -particle system in which the particles move in an external potential and interact pairwise:

$$\hat{H} = \sum_i \frac{\hat{p}_i^2}{2m} + \sum_i V_{ext}(\hat{x}_i) + \sum_{i < j} V_{int}(|\hat{x}_i - \hat{x}_j|). \quad (6.7)$$

**Note:** Some textbooks claim that Principle 1 forces  $|\psi\rangle$  to be an eigenstate of the permutation operator, i.e. that  $\hat{P}_{ij}|\psi\rangle = \pm|\psi\rangle$ . *This is not correct.* Rather, the indistinguishability of identical particles places a restriction on the physical observables, namely that they must be permutation invariant. In order to make a statement about which  $N$ -particle states are physically allowed requires a second, stronger postulate.

### 6.3 Symmetrization

We began our investigation of angular momentum by considering different ways of realising rotations as operators acting on the Hilbert space. In mathematical terms this amounted to finding all the various representations of the rotation group  $SO(3)$ . Based on this we could then label states according to how they transformed under rotations. We now find ourselves in a similar position. The group in question is the symmetry group  $S_N$  corresponding to all the different ways of permuting  $N$  objects. The representation on the Hilbert space is provided by the permutation operators. We can again label states according to the manner in which they transform under this representation. However, a full treatment of this problem requires some sophisticated tools from representation theory which we will not introduce here. Fortunately it turns out that only states with very simple transformation properties are actually physically relevant. To illustrate these ideas we consider the following simple example. Suppose we have three particles of which one is in state  $\alpha$ , one in  $\beta$  and one in  $\gamma$ . The six basis states of  $\mathcal{H}_3$  that correspond to this situation are  $\{|\alpha, \beta, \gamma\rangle, |\alpha, \gamma, \beta\rangle, |\beta, \alpha, \gamma\rangle, |\beta, \gamma, \alpha\rangle, |\gamma, \alpha, \beta\rangle, |\gamma, \beta, \alpha\rangle\}$ . The permutation operators clearly leave the subspace spanned by these six states invariant. In fact, this space actually contains four subspaces, each of which is invariant under the permutation

operators:<sup>1</sup>

$$S_1 = \text{span}\{|\alpha, \beta, \gamma\rangle + |\beta, \alpha, \gamma\rangle + |\alpha, \gamma, \beta\rangle + |\gamma, \beta, \alpha\rangle + |\gamma, \alpha, \beta\rangle + |\beta, \gamma, \alpha\rangle\}$$

$$S_2 = \text{span}\{|\alpha, \beta, \gamma\rangle - |\beta, \alpha, \gamma\rangle - |\alpha, \gamma, \beta\rangle - |\gamma, \beta, \alpha\rangle + |\gamma, \alpha, \beta\rangle + |\beta, \gamma, \alpha\rangle\}$$

$$\begin{aligned} S_3 = \text{span}\{ & 2|\alpha, \beta, \gamma\rangle + 2|\beta, \alpha, \gamma\rangle - |\alpha, \gamma, \beta\rangle - |\gamma, \beta, \alpha\rangle - |\gamma, \alpha, \beta\rangle - |\beta, \gamma, \alpha\rangle, \\ & 0 + 0 - |\alpha, \gamma, \beta\rangle + |\gamma, \beta, \alpha\rangle + |\gamma, \alpha, \beta\rangle - |\beta, \gamma, \alpha\rangle \} \end{aligned}$$

$$\begin{aligned} S_4 = \text{span}\{ & 2|\alpha, \beta, \gamma\rangle - 2|\beta, \alpha, \gamma\rangle + |\alpha, \gamma, \beta\rangle + |\gamma, \beta, \alpha\rangle - |\gamma, \alpha, \beta\rangle - |\beta, \gamma, \alpha\rangle, \\ & 0 + 0 - |\alpha, \gamma, \beta\rangle + |\gamma, \beta, \alpha\rangle - |\gamma, \alpha, \beta\rangle + |\beta, \gamma, \alpha\rangle \} \end{aligned}$$

We see that  $S_1$  is one-dimensional and that its basis state is completely *symmetric* since it is invariant under any permutation of particles. The basis state of  $S_2$  is completely *antisymmetric* since exchanging two particles produces a minus sign. The remaining two subspaces do not have a single definite symmetry with respect to particle exchange and are said to be of a mixed symmetry type. The general picture is therefore as follows: The  $N$ -particle Hilbert space  $\mathcal{H}_N$  will contain a subspace of completely symmetric states, one of completely anti-symmetric states, as well as numerous subspaces of mixed symmetry. As yet there is nothing to suggest that some of these symmetries might be unphysical. For this we need the following fundamental principle:

### Principle 2: The Symmetrization Postulate

- *Physical states are of one of only two possible symmetry types: either completely symmetric or completely antisymmetric.*
- *Particles with integer spin are called bosons and have symmetric states, i.e. states which are completely invariant under any permutation of particles.*
- *Particles with half-integer spin are called fermions and have anti-symmetric states, i.e. states for which the exchange of two particles produces a minus sign.*

**Remark:** We will treat this as an experimental fact. It can be proven in quantum field theory where it is known as the *Spin-Statistics theorem*.

If  $N = 2$  the Hilbert space is  $\mathcal{H}_2 = \text{span}\{|n_1, n_2\rangle : n_1, n_2 = 1, 2, 3, \dots\}$ . For  $n_1 \neq n_2$  the pair  $|n_1, n_2\rangle$  and  $|n_2, n_1\rangle$  can be used to construct one completely symmetric and one completely antisymmetric state:

$$|n_1, n_2\rangle_S = \frac{1}{\sqrt{2}}(|n_1, n_2\rangle + |n_2, n_1\rangle) \quad P_{12}|n_1, n_2\rangle_S = |n_1, n_2\rangle_S \quad (6.8)$$

$$|n_1, n_2\rangle_A = \frac{1}{\sqrt{2}}(|n_1, n_2\rangle - |n_2, n_1\rangle) \quad P_{12}|n_1, n_2\rangle_A = -|n_1, n_2\rangle_A \quad (6.9)$$

If  $n_1 = n_2$  then  $|n_1, n_1\rangle_S = |n_1, n_1\rangle$ , but  $|n_1, n_1\rangle_A = 0$  which is not a physical state. This is true in general: attempting to construct an anti-symmetric state in which two

---

<sup>1</sup>In the language introduced earlier we can say that each subspace carries an irreducible representation of the symmetry group.

particles occupy the same single particle state always produces the zero vector. The antisymmetry of fermionic many-body states therefore results in the following fundamental principle:

**Pauli exclusion principle:** *Two fermions cannot occupy the same single particle state.*

# Chapter 7

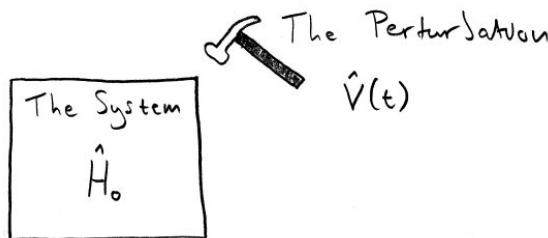
## Time-Dependent Perturbation Theory

### 7.1 Background

Many practical applications of quantum mechanics require an understanding of the time evolution of systems with Hamiltonians of the form

$$\hat{H}(t) = \hat{H}_0 + \hat{V}(t). \quad (7.1)$$

Here  $\hat{H}_0$  is the time-independent Hamiltonian of the unperturbed system and  $\hat{V}(t)$  is a potentially time-dependent perturbation. In most applications of time-dependent perturbation theory (TDPT) the unperturbed Hamiltonian  $\hat{H}_0$  is that of the physical system of interest, while the perturbation is some form of “external influence”. We want to study the response of the system to this external influence. In particular, we will investigate the transitions (between  $\hat{H}_0$  eigenstates) that the perturbation can induce. See technical diagram below.



For example, consider the interaction of a hydrogen atom with a (classical) electromagnetic field. In this case  $\hat{H}_0$  is the usual atomic Hamiltonian while  $\hat{V}(t)$  describes the interaction of the atom with the oscillating field. If  $\hat{V}(t)$  is time-dependent then solving for the time evolution of a particular state is generally very difficult since the time evolution operator is no longer a simple exponential. Even cases where  $\hat{V}(t) = \hat{V}$  is independent of time may still be problematic if the eigenstates and eigenvalues of the perturbed Hamiltonian are not known. This highlights the need for a systematic perturbative formalism of time evolution.

In what follows we assume that the eigenstates and eigenvalues of  $\hat{H}_0$  are known. We use the notation

$$\hat{H}_0|n\rangle = E_n|n\rangle \quad (7.2)$$

where  $n$  is a general label.

## 7.2 Reminder: Quantum dynamics

Suppose  $\hat{V}(t) = 0$ . The time evolution operator is then the usual exponential

$$\hat{U}_0(t, t_0) = e^{-i(t-t_0)\hat{H}_0/\hbar}. \quad (7.3)$$

If the system starts out in the state  $|\alpha\rangle = \sum_n \tilde{c}_n |n\rangle$  at time  $t_0$  its state at time  $t$  will be

$$|\alpha, t, t_0\rangle = \hat{U}_0(t, t_0)|\alpha\rangle = \sum_n e^{-i(t-t_0)E_n/\hbar} \tilde{c}_n |n\rangle. \quad (7.4)$$

Time evolution in the basis of energy eigenstates is therefore very simple: each expansion coefficient is just modified by a time and energy dependent phase factor. When  $\hat{V}(t) \neq 0$  things are much more complicated. We can still expand  $|\alpha, t, t_0\rangle$  in the  $\hat{H}_0$ -basis as

$$|\alpha, t, t_0\rangle = \hat{U}(t, t_0)|\alpha\rangle = \sum_n \tilde{c}_n(t) |n\rangle, \quad (7.5)$$

but now the expansion coefficients  $\tilde{c}_n(t)$  evolve in some complicated way and are no longer just  $\tilde{c}_n$  times a phase.

The probability of finding the system<sup>1</sup> in an eigenstate  $|n\rangle$  of  $\hat{H}_0$  at time  $t$  is

$$|\langle n | \alpha, t, t_0 \rangle|^2 = |\langle n | \hat{U}(t, t_0) | \alpha \rangle|^2 = |\tilde{c}_n(t)|^2. \quad (7.6)$$

These *transition probabilities* are the central physical quantities in this chapter.

## 7.3 The interaction picture

### 7.3.1 Definition

Time-dependent perturbation theory is most conveniently formulated in the interaction picture, which is a hybrid of the Schrödinger and Heisenberg pictures encountered earlier. Let  $|\alpha, t, t_0\rangle$  be a state and  $\hat{A}$  an observable in the Schrödinger picture.  $\hat{A}$  is assumed to have no explicit time dependence. We define their interaction picture counterparts as<sup>2</sup>

$$|\alpha, t, t_0\rangle_I \equiv e^{it\hat{H}_0/\hbar} |\alpha, t, t_0\rangle \equiv \sum_n c_n(t) |n\rangle \quad (7.7)$$

$$\hat{A}_I(t) \equiv e^{it\hat{H}_0/\hbar} \hat{A} e^{-it\hat{H}_0/\hbar} \quad (7.8)$$

---

<sup>1</sup>This is a standard abuse of terminology. “*Finding in  $|n\rangle$* ”=“*Performing a measurement of the unperturbed energy and obtaining a result of  $E_n$* ”

<sup>2</sup>Here we could also apply the transformation  $e^{i(t-t_0)\hat{H}_0/\hbar}$  instead of  $e^{it\hat{H}_0/\hbar}$ . The former is conceptually clearer, as it is nothing but the inverse of the time evolution operator associated with the unperturbed Hamiltonian. This convention is sometimes used in more advanced textbooks. The only difference between the two conventions will be some constant phases which eventually cancel out. We will stick to using  $e^{it\hat{H}_0/\hbar}$ , as its makes the intermediate expressions a bit simpler. *Always make sure that you know which picture, and which conventions, you are using.* This is particularly important in the contexts of quantum field theory and the path integral formalism.

which are *both* time-dependent when  $\hat{V}(t) \neq 0$ . Here  $\{c_n(t)\}$  are the state's expansion coefficients in the interaction picture. They are related to the  $\tilde{c}_n(t)$  coefficients of the Schrödinger picture by

$$c_n(t) = e^{iE_n t/\hbar} \tilde{c}_n(t). \quad (7.9)$$

Of course,  $|c_n(t)|^2 = |\tilde{c}_n(t)|^2$  and so knowing  $c_n(t)$  is sufficient for calculating the transition probabilities.

Roughly speaking, the purpose of the interaction picture is to eliminate the “trivial” part of the state’s time evolution due to  $\hat{H}_0$ . This allows us to focus on how the perturbation affects the dynamics. In fact, in the interaction picture the time evolution of states is governed by  $\hat{V}_I(t)$  while that of observables is governed by  $\hat{H}_{0,I}(t)$ . The equations of motion are seen to be

$$i\hbar \frac{\partial}{\partial t} |\alpha, t, t_0\rangle_I = \hat{V}_I(t) |\alpha, t, t_0\rangle_I \quad (7.10)$$

$$i\hbar \frac{\partial \hat{A}_I(t)}{\partial t} = [\hat{A}_I(t), \hat{H}_{0,I}(t)]. \quad (7.11)$$

Note that if  $\hat{V}(t) = 0$  then  $|\alpha, t, t_0\rangle_I$  is time-independent. Applying  $\langle n |$  to both sides of (7.10) yields a coupled set of differential equations for the expansion coefficients  $c_n(t)$ :

$$i\hbar \frac{dc_n(t)}{dt} = \sum_m V_{nm}(t) e^{i\omega_{nm}t} c_m(t). \quad (7.12)$$

Here  $V_{nm}(t) = \langle n | \hat{V}(t) | m \rangle$  and  $\omega_{nm} = (E_n - E_m)/\hbar$ . These equations provide a complete description of the state’s time evolution but are generally not exactly solvable and only of limited practical value. A very important special case which *is* analytically solvable is that of a perturbed two-state system. We consider this example next.

	Schrödinger	Heisenberg	Interaction
State	$ \psi, t\rangle_S = U(t) \psi, 0\rangle$	$ \psi, 0\rangle$	$ \psi, t\rangle_I = U_0^\dagger(t) \psi, t\rangle_S$
Density Matrix	$\hat{\rho}_S(t) = U(t)\hat{\rho}(0)U^\dagger(t)$	$\hat{\rho}(0)$	$\hat{\rho}_I(t) = U_0^\dagger(t)\hat{\rho}_S(t)U_0(t)$
Observable	$\hat{A}_S$	$\hat{A}_H(t) = U^\dagger(t)\hat{A}_S U(t)$	$\hat{A}_I(t) = U_0^\dagger(t)\hat{A}_S U_0(t)$

Table 7.1: A summary of the three pictures. Here  $t_0 = 0$  and  $U_0(t) = \exp[-i\hat{H}_0 t/\hbar]$ .

### 7.3.2 The dynamics of a two-state system

Consider a system with a two-dimensional Hilbert space spanned by the orthogonal basis  $\{|1\rangle, |2\rangle\}$ . We take the unperturbed Hamiltonian to be

$$\hat{H}_0 = E_1|1\rangle\langle 1| + E_2|2\rangle\langle 2|. \quad (7.13)$$

Here  $E_2 > E_1$  and so  $|1\rangle$  and  $|2\rangle$  are the (unperturbed) ground and excited states respectively. We will consider a harmonic perturbation with frequency  $\omega$  of the form

$$\hat{V}(t) = \gamma e^{i\omega t}|1\rangle\langle 2| + \gamma e^{-i\omega t}|2\rangle\langle 1|. \quad (7.14)$$

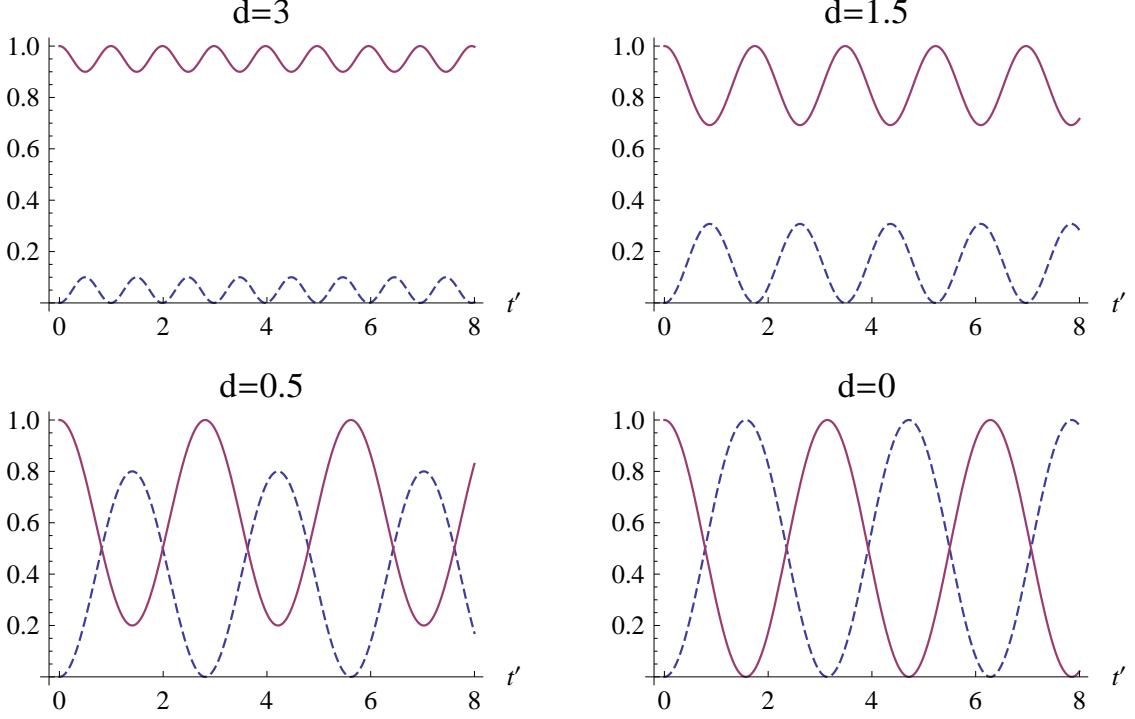


Figure 7.1: The probabilities  $|c_1(t')|^2$  (solid) and  $|c_2(t')|^2$  (dashed) as functions of  $t' = \gamma t/\hbar$  for different values of the dimensionless detuning parameter  $d = |\omega - \omega_{21}| \hbar / (2\gamma)$ . When  $d = 0$  the system is at resonance.

The relevant matrix elements are  $V_{12}(t) = V_{21}^*(t) = \gamma \exp[i\omega t]$  and  $V_{11} = V_{22} = 0$ . Suppose the system was in the ground state  $|1\rangle$  at time  $t_0 = 0$ . At time  $t$  we then have, in the interaction picture,

$$|1, t\rangle_I = c_1(t)|1\rangle + c_2(t)|2\rangle \quad (7.15)$$

with  $c_1(0) = 1$  and  $c_2(0) = 0$ . The coupled differential equations in (7.12) now read

$$i\hbar\dot{c}_1(t) = V_{12}(t)e^{i\omega_{12}t}c_2(t) = \gamma e^{i(\omega-\omega_{21})t}c_2(t) \quad (7.16)$$

$$i\hbar\dot{c}_2(t) = V_{21}(t)e^{i\omega_{21}t}c_1(t) = \gamma e^{-i(\omega-\omega_{21})t}c_1(t) \quad (7.17)$$

where the dot denotes a time derivative. These equations can be decoupled easily to produce the second-order differential equation

$$\ddot{c}_2(t) + a\dot{c}_2(t) + bc_2(t) = 0 \quad (7.18)$$

with  $a = i(\omega - \omega_{21})$  and  $b = \gamma^2/\hbar^2$ . The general solution to this equation is

$$c_2(t) = e^{-at/2} [C_1 \sin(\lambda t/2) + C_2 \cos(\lambda t/2)] \quad (7.19)$$

where  $\lambda^2 = 4b - a^2 = 4\gamma^2/\hbar^2 + (\omega - \omega_{21})^2$  and  $C_{1,2}$  are integration constants. To fix the latter we must use the initial conditions. Specifically,  $c_2(0) = 0$  implies that  $C_2 = 0$  while  $c_1(0) = 1$  requires that  $C_1 = 2\gamma/(i\hbar\lambda)$ . Therefore,  $c_2(t)$  is given by

$$c_2(t) = \frac{2\gamma}{i\hbar\lambda} \sin(\lambda t/2) e^{-i(\omega-\omega_{21})t/2}. \quad (7.20)$$

The probability of finding the system in the ground state ( $|c_1(t)|^2$ ) or the excited state ( $|c_2(t)|^2$ ) are

$$|c_2(t)|^2 = \Omega^2 \sin^2 \left[ \frac{\gamma t}{\Omega \hbar} \right] \quad \text{and} \quad |c_1(t)|^2 = 1 - |c_2(t)|^2 \quad (7.21)$$

where

$$\Omega^2 = \frac{\gamma^2/\hbar^2}{\gamma^2/\hbar^2 + (\omega - \omega_{21})^2/4}. \quad (7.22)$$

We see that  $|c_2(t)|^2$  oscillates between 0 and  $\Omega^2$ . This cycle of *stimulated emission and absorption* is known as the Rabi cycle. Clearly  $\Omega^2$  is maximal at **resonance**, i.e. when the driving frequency  $\omega$  of the perturbation matches the system's natural transition frequency  $\omega_{21}$ . See figure 7.1. This is a fundamental result with applications in a wide range of fields, including laser physics, quantum optics, nuclear magnetic resonance and quantum computing.

## 7.4 Perturbation theory

In this section we develop a perturbative approximation scheme for treating the time evolution of perturbed systems. This formalism will allow us to express transition probabilities as power series in the perturbation  $\hat{V}(t)$ .

### 7.4.1 The Dyson series

We begin by deriving a perturbative expansion for the time evolution operator in the interaction picture. First we rewrite the basic relation

$$|\alpha, t, t_0\rangle = \hat{U}(t, t_0)|\alpha, t_0, t_0\rangle \quad (7.23)$$

in the interaction picture as

$$|\alpha, t, t_0\rangle_I = \hat{U}_I(t, t_0)|\alpha, t_0, t_0\rangle_I \quad (7.24)$$

where

$$\hat{U}_I(t, t_0) \equiv e^{+i\hat{H}_0 t/\hbar} \hat{U}(t, t_0) e^{-i\hat{H}_0 t_0/\hbar} \quad (7.25)$$

is the time evolution operator in the interaction picture. It is easy to show that  $\hat{U}_I(t, t_0)$  satisfies

$$i\hbar \frac{\partial}{\partial t} \hat{U}_I(t, t_0) = \hat{V}_I(t) \hat{U}_I(t, t_0) \quad \text{with} \quad \hat{U}_I(t_0, t_0) = \hat{I}. \quad (7.26)$$

Integrating with respect to  $t$  on both sides of this differential equation produces the implicit integral equation

$$\hat{U}_I(t, t_0) = \hat{I} - \frac{i}{\hbar} \int_{t_0}^t dt' \hat{V}_I(t') \hat{U}_I(t', t_0). \quad (7.27)$$

This form allows for a systematic iterative solution in which each iteration introduces higher order corrections:

- **Zeroth order:**

$$\hat{U}_I^{(0)}(t, t_0) = \hat{I} \quad (7.28)$$

- **First (linear) order:** Insert  $\hat{U}_I^{(0)}$  into the RHS of the integral equation:

$$\hat{U}_I^{(1)}(t, t_0) = \hat{I} - \frac{i}{\hbar} \int_{t_0}^t dt' \hat{V}_I(t') \hat{U}_I^{(0)}(t, t_0) = \hat{I} - \frac{i}{\hbar} \int_{t_0}^t dt' \hat{V}_I(t') \quad (7.29)$$

- **Second (quadratic) order:** Insert  $\hat{U}_I^{(1)}$  into the RHS of the integral equation:

$$\hat{U}_I^{(2)}(t, t_0) = \hat{I} - \frac{i}{\hbar} \int_{t_0}^t dt' \hat{V}_I(t') + \left( \frac{i}{\hbar} \right)^2 \int_{t_0}^t dt' \int_{t_0}^{t'} dt'' \hat{V}_I(t') \hat{V}_I(t'') \quad (7.30)$$

This series  $\{\hat{U}_I^{(n)}(t, t_0)\}$  of progressively higher order approximations to  $\hat{U}_I(t, t_0)$  is called the Dyson series. If the series converges we have

$$\lim_{n \rightarrow \infty} \hat{U}_I^{(n)}(t, t_0) = \hat{U}_I(t, t_0). \quad (7.31)$$

#### 7.4.2 Transition probabilities from perturbation theory

The perturbative expansion for  $\hat{U}_I(t, t_0)$  allows us to derive an analogous expansion for the transition probabilities between eigenstates of the unperturbed Hamiltonian. First we remind ourselves of the central question: *If the system is in the state  $|n\rangle$  (with  $\hat{H}_0|n\rangle = E_n|n\rangle$ ) at  $t = t_0$ , what is the probability of finding it in the state  $|m\rangle$  (with  $\hat{H}_0|m\rangle = E_m|m\rangle$ ) at a later time  $t$ ?* This is precisely the transition probability

$$P_{n \rightarrow m} = |\langle m | \hat{U}(t, t_0) | n \rangle|^2. \quad (7.32)$$

Clearly

$$\langle m | \hat{U}(t, t_0) | n \rangle = e^{i(E_n t_0 - E_m t)/\hbar} \langle m | \hat{U}_I(t, t_0) | n \rangle \quad (7.33)$$

and we denote the matrix element of  $\hat{U}_I(t, t_0)$  by

$$c_{n \rightarrow m} = \langle m | \hat{U}_I(t, t_0) | n \rangle. \quad (7.34)$$

The **transition amplitude** is

$$A_{n \rightarrow m} = \langle m | \hat{U}(t, t_0) | n \rangle = e^{i(E_n t_0 - E_m t)/\hbar} c_{n \rightarrow m} \quad (7.35)$$

while the **transition probability** is

$$P_{n \rightarrow m} = |A_{n \rightarrow m}|^2 = |c_{n \rightarrow m}|^2. \quad (7.36)$$

The matrix elements of the perturbation (in the interaction picture) are

$$\langle m | \hat{V}_I(t) | n \rangle = e^{i(E_m - E_n)t/\hbar} \langle m | \hat{V}(t) | n \rangle = e^{i\omega_{mn}t} V_{mn}(t) \quad (7.37)$$

where

$$V_{mn}(t) = \langle m | \hat{V}(t) | n \rangle \quad \text{and} \quad \omega_{mn} = (E_m - E_n)/\hbar. \quad (7.38)$$

Let  $c_{n \rightarrow m}^{(l)} = \langle m | \hat{U}_I^{(l)}(t, t_0) | n \rangle$  denote the  $l$ 'th order approximation to  $c_{n \rightarrow m}$ . From (7.28), (7.29) and (7.30) it follows that

$$c_{n \rightarrow m}^{(0)} = \delta_{n,m} \quad (7.39)$$

$$c_{n \rightarrow m}^{(1)} = \delta_{n,m} - \frac{i}{\hbar} \int_{t_0}^t dt' e^{i\omega_{mn}t'} V_{mn}(t') \quad (7.40)$$

$$c_{n \rightarrow m}^{(2)} = \delta_{n,m} - \frac{i}{\hbar} \int_{t_0}^t dt' e^{i\omega_{mn}t'} V_{mn}(t') + \left(\frac{i}{\hbar}\right)^2 \sum_k \int_{t_0}^t dt' \int_{t_0}^{t'} dt'' e^{i\omega_{mk}t'} e^{i\omega_{kn}t''} V_{mk}(t') V_{kn}(t'')$$

and so on. **These expressions are the central results that we will use in the remainder of the chapter.** From the structure of the expression for  $c_{n \rightarrow m}^{(1)}$  we see that at linear order the perturbation can only induce transitions between states that it “connects” directly, i.e. for which the matrix element  $\langle m | \hat{V}(t) | n \rangle$  is non-zero. In particular, this implies that selection rules for matrix elements can be used to formulate selection rules for transitions, i.e. criteria for identifying forbidden transitions. Note that this, and other such statements, hold at a particular order in the perturbative calculation. At higher orders a larger set of transitions will be possible, although these will occur with a lower probability if  $\hat{V}(t)$  is indeed “small”. For example,  $c_{n \rightarrow m}^{(k>1)}$  will contain a contribution from the matrix element

$$\langle m | \hat{V}_I(t') \hat{V}_I(t'') | n \rangle = \sum_k \langle m | \hat{V}_I(t') | k \rangle \langle k | \hat{V}_I(t'') | n \rangle, \quad (7.41)$$

which could allow for a transition from the state  $|n\rangle$  to the state  $|m\rangle$  via the intermediate state  $|k\rangle$ . This would be a *higher order process*. Two-photon absorption is an example of this.

## 7.5 Applications

### 7.5.1 Constant perturbation

Consider the case where a time-independent perturbation is suddenly switched on at  $t = 0$ :

$$\hat{V}(t) = \begin{cases} 0 & t < 0 \\ \hat{V} & t > 0 \end{cases} \quad (7.42)$$

Prior to  $t = 0$  the system is in the  $\hat{H}_0$  eigenstate  $|n\rangle$ . First we calculate the probability that the system will make a transition to a *different* state  $|m\rangle$  at a later time  $t$ . To linear order we have from (7.40) that

$$c_{n \rightarrow m}^{(1)} = \frac{V_{mn}}{\hbar\omega_{mn}} [1 - e^{i\omega_{mn}t}] = \frac{-2iV_{mn}}{\hbar} \frac{e^{i\omega_{mn}t/2} \sin(\omega_{mn}t/2)}{\omega_{mn}} \quad (7.43)$$

$$P_{n \rightarrow m}^{(1)} = |c_{n \rightarrow m}^{(1)}|^2 = \frac{4|V_{mn}|^2}{\hbar^2} \frac{\sin^2(\omega_{mn}t/2)}{\omega_{mn}^2} \quad (7.44)$$

Figure 7.2 shows  $P_{n \rightarrow m}^{(1)}$  as a function of the transition frequency  $\omega_{mn}$  at a fixed time  $t$ . From the shape of the curve it is clear that the transition probability  $P_{n \rightarrow m}^{(1)}$  is only appreciable for final states with energies  $E_m$  such that  $|\omega_{mn}| \lesssim 2\pi/t$ , i.e. in the range

$E_n - 2\pi\hbar/t \lesssim E_m \lesssim E_n + 2\pi\hbar/t$ . If we measure the system's energy (meaning  $H_0$ ) a time  $\Delta t = t - t_0 = t$  after switching on the perturbation then the spread in the measured energies will be  $\Delta E \sim 2\pi\hbar/\Delta t$ . This suggests that

$$\Delta E \Delta t \sim \hbar \quad (7.45)$$

which is one instance of the **time-energy uncertainty relation**. Note that this result is *not* on the same level as the Heisenberg uncertainty relations since time is just a parameter and not an observable. We reach the following important conclusion:

*For a constant perturbation “energy conserving” transitions dominate at long times.*

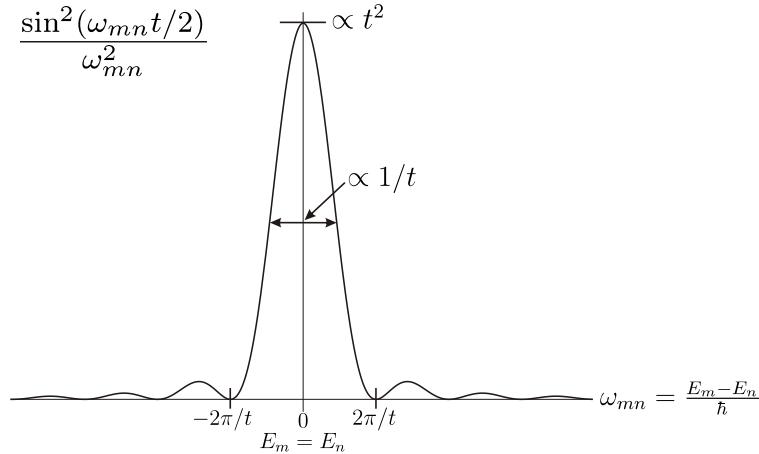
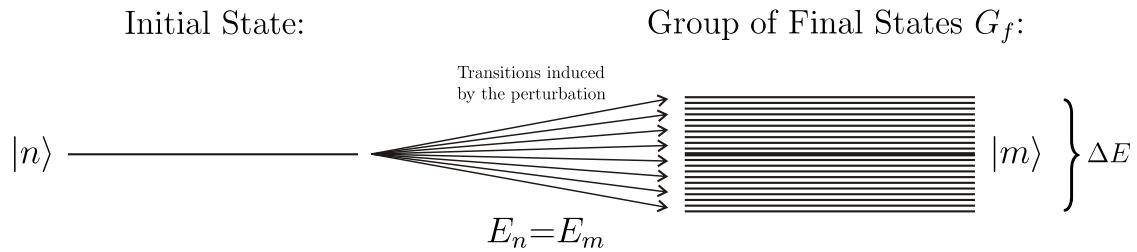


Figure 7.2:  $\sin^2(\omega_{mn}t/2)/\omega_{mn}^2$  as a function of  $\omega_{mn}$ .

### Fermi’s golden rule for constant perturbations

**Note:** The derivation of Fermi’s golden rule is somewhat technical, and contains a number of assumptions about the system at hand. There are two versions of this derivation, the more detailed version at the start of this section, and the alternative short-cut version at the end. It can be difficult to keep track of all the moving parts in this derivation, so working through the example in section 7.5.2 in parallel should be useful. Despite all the assumptions that enter here, Fermi’s golden rule turns out to be a very useful and widely applicable result. The examples in the homework should help to make things clearer. A good textbook reference for this is Cohen-Tannoudji, Chapter XIII, Section C.

In practice the possible final states often form part of a continuum, and we are then interested in the probability of finding the system in one of a group of possible final states all with energies close to  $E_n$ .



Let  $G_f$  denote this group of final states which have energies  $E_f$  in the range

$$E_m - \Delta E/2 \leq E_f \leq E_m + \Delta E/2 \quad (7.46)$$

where  $E_m = E_n$  and  $|m\rangle (\neq |n\rangle)$  is a representative state in  $G_f$ . Here  $\Delta E$  defines a small fixed energy interval around  $E_m = E_n$ . The precise choice of  $\Delta E$  will not play a role in the final result.<sup>3</sup> Also note that  $G_f$  will generally *not* contain all states with energies in the range (7.46), but only a particular subset.

We now define

$$P_{n \rightarrow [m]}(t) = \text{The probability of finding the system in a state } |f\rangle \in G_f \text{ at time } t,$$

which can be calculated by “summing” over all the final states in  $G_f$ . Since these states form a continuum (or can be approximated as such) this can be done using an integral with an appropriate density of states:

$$P_{n \rightarrow [m]}^{(1)}(t) = \int_{E_m - \Delta E/2}^{E_m + \Delta E/2} dE_f \rho_f(E_f) \frac{4|V_{fn}|^2}{|E_f - E_n|^2} \sin^2 \left[ \frac{(E_f - E_n)t}{2\hbar} \right]. \quad (7.47)$$

Here  $\rho_f(E_f)$  is the density of final states:

$$\rho_f(E_f) dE_f = \text{The number of states in } G_f \text{ with energies in an interval } dE_f \text{ around } E_f.$$

Note that  $\rho_f(E_f)$  is generally *not* the total density of states. Also note that this integral expression assumes that the matrix element  $V_{fn}$  is approximately constant for all states in  $G_f$  with the same energy, i.e. we can treat  $V_{fn}$  as a function of  $E_f$  rather than the state label  $f$ . From figure 7.2 we see that at large  $t$  the integrand is sharply peaked at  $E_f = E_m$ . We will assume that the functions  $\rho_f(E_f)$  and  $|V_{fn}|^2$  vary slowly over the width of this peak and may therefore be approximated by their values at  $E_f = E_m$ . This produces

$$P_{n \rightarrow [m]}^{(1)}(t) = 4\rho_f(E_m)|V_{mn}|^2 \int_{E_m - \Delta E/2}^{E_m + \Delta E/2} dE_f \frac{1}{|E_f - E_n|^2} \sin^2 \left[ \frac{(E_f - E_n)t}{2\hbar} \right]. \quad (7.48)$$

Furthermore, once the entire peak falls inside the integration domain we may extend the bounds to  $\pm\infty$ . Finally we use the fact that  $\int_{-\infty}^{\infty} dx \sin^2(ax)/x^2 = a\pi$  to obtain

$$P_{n \rightarrow [m]}^{(1)}(t) = \frac{2\pi}{\hbar} t \rho_f(E_m) |V_{mn}|^2. \quad (7.49)$$

The transition rate is then a constant and given by

$$\omega_{n \rightarrow [m]}^{(1)} = \frac{dP_{n \rightarrow [m]}^{(1)}(t)}{dt} = \frac{2\pi}{\hbar} \rho_f(E_m) |V_{mn}|^2. \quad (7.50)$$

This is **Fermi’s golden rule**.

---

<sup>3</sup>It is *not* the same  $\Delta E$  as in the time-energy uncertainty relation (7.45).

### Alternative Derivation:

These results can also be obtained using a short-cut where the long-time limit is taken before integrating over the group of final states. Since

$$\lim_{a \rightarrow \infty} \frac{\sin^2(ax)}{ax^2} = \pi\delta(x) \quad (7.51)$$

it follows from (7.44) that as  $t \rightarrow \infty$  we have

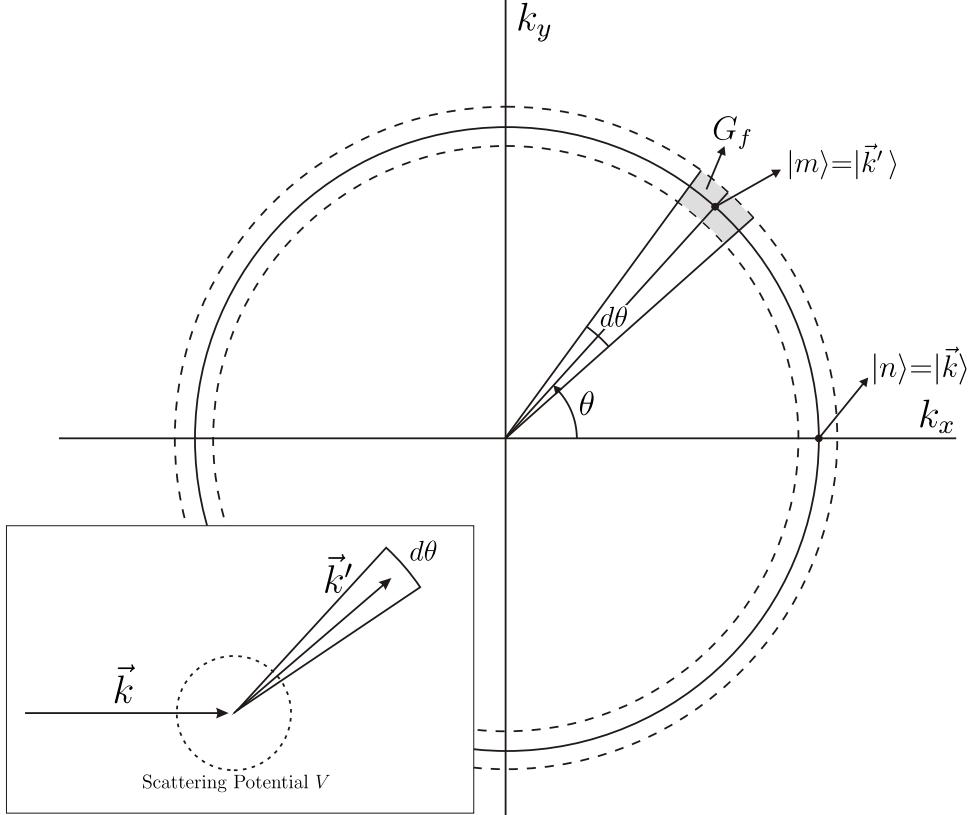
$$P_{n \rightarrow f}^{(1)}(t) \rightarrow \frac{2\pi}{\hbar} t |V_{fn}|^2 \delta(E_f - E_n), \quad (7.52)$$

where the delta-function enforces energy conservation at long times. Note that the expression above is no longer a valid probability. It now only makes sense when integrated over the final state energy  $E_f$  with an appropriate density of states:

$$P_{n \rightarrow [m]}^{(1)}(t) = \int dE_f \rho_f(E_f) P_{n \rightarrow f}^{(1)}(t) = \int dE_f \rho_f(E_f) \frac{2\pi}{\hbar} t |V_{fn}|^2 \delta(E_f - E_n) = \frac{2\pi}{\hbar} t \rho_f(E_m) |V_{mn}|^2.$$

This is precisely the result we obtained previously.

### 7.5.2 Example: Fermi's golden rule and scattering in 2D



In this section we will derive the scattering rate and scattering cross section for a particle undergoing elastic scattering in two dimensions. We will revisit this topic in much more detail in chapter 8. Consider a free particle with mass  $m$ , which is initially in the momentum eigenstate  $|\vec{k}\rangle$  with  $\vec{p} = \hbar\vec{k}$ . In scattering theory we usually use  $\vec{k}$  rather

than  $\vec{p}$  as a label. At time  $t = 0$  a constant short range scattering potential  $\hat{V} = V(\hat{\vec{x}})$  is turned on. As a result, the particle is *scattered* from its initial momentum state  $|\vec{k}\rangle$  to a superposition of momentum eigenstates. We want to calculate the associated transition rate for scattering from  $|\vec{k}\rangle$  to a final momentum eigenstate  $|\vec{k}_f\rangle$  corresponding to a wave vector  $\vec{k}_f$  within an infinitesimal angle  $d\theta$  around a specific  $\vec{k}'$ . We will take  $|\vec{k}\rangle$  to lie along the  $x$ -direction. In the scattering experiment this means that the incoming beam of particles approaches the target along the  $x$ -axis. In the notation of the previous sections we have

$$\hat{H}_0 = \frac{\hat{\vec{p}}^2}{2m} \quad |n\rangle = |\vec{k} = (k, 0)\rangle \quad |m\rangle = |\vec{k}'\rangle \quad (7.53)$$

and

$$E_k = \frac{\hbar^2 k^2}{2m} \quad |\vec{k}| = |\vec{k}'| = k = k' \quad E_k = E_{k'} = E_n = E_m. \quad (7.54)$$

The constant potential produces *elastic scattering*, and so the initial and final energies are the same. The group of final states  $G_f$  corresponds to the grey region in the figure. The two dashed lines define the energy interval  $\Delta E$ , corresponding to a momentum range  $\Delta p_f = \hbar \Delta k_f$ , which does not play any role in the actual calculation.

It will be convenient to impose *periodic box boundary conditions*, where we imagine the system as being contained in a square with side-length  $L$ . This discretises the momentum as

$$\vec{p} = \hbar \vec{k} = \frac{2\pi\hbar}{L}(n_x, n_y) \quad \text{with} \quad n_x, n_y \in \mathbb{Z}. \quad (7.55)$$

However, since we can take  $L$  to be arbitrarily large, we can treat  $\vec{k}$  as being essentially continuous. The final results will not depend on the precise value of  $L$ .<sup>4</sup> The (normalised!) plane wave states will then be

$$\langle \vec{x} | \vec{k} \rangle = \frac{1}{L} e^{i\vec{k} \cdot \vec{x}}. \quad (7.56)$$

To apply Fermi's golden rule we also require the density of final states  $\rho_f(E_f)$ . This can be derived in the usual way, by first writing

$$\sum_{(n_x, n_y) \in G_f} = \int_{G_f} dn_x dn_y = \left(\frac{L}{2\pi}\right)^2 \int_{G_f} dk_x dk_y = \left(\frac{L}{2\pi}\right)^2 \int_{G_f} dk_f d\theta_f k_f. \quad (7.57)$$

Since  $d\theta$  is infinitesimal the integral over  $\theta_f$  just contributes a factor of  $d\theta$ . Changing variables to  $E_f = \hbar^2 k_f^2 / (2m)$  then leads to

$$\left(\frac{L}{2\pi}\right)^2 d\theta \int_{\Delta k_f} dk_f k_f = \int_{\Delta E} dE_f \left[ \left(\frac{L}{2\pi}\right)^2 \frac{m}{\hbar^2} d\theta \right] \quad (7.58)$$

from which we identify

$$\rho_f(E_f) = \left(\frac{L}{2\pi}\right)^2 \frac{m}{\hbar^2} d\theta. \quad (7.59)$$

---

<sup>4</sup>We can drop the box boundary conditions and work in an infinite space from the beginning. However, this requires the use of non-normalisable plane waves, which makes some of the intermediate expressions tricky to interpret. The final result will be the same.

In particular, we note that the density of states in 2D is a constant, independent of  $E_f$ . Fermi's golden rule predicts the transition rate

$$d\omega_{n \rightarrow [m]}^{(1)} = d\omega_{\vec{k} \rightarrow [\vec{k}']}^{(1)} = \frac{2\pi}{\hbar} \left( \frac{L}{2\pi} \right)^2 \frac{m}{\hbar^2} |\langle \vec{k}' | \hat{V} | \vec{k} \rangle|^2 d\theta. \quad (7.60)$$

In the scattering context the real quantity of interest is the transition rate per unit angle ( $d\omega_{\vec{k} \rightarrow [\vec{k}']}^{(1)}/d\theta$ ) expressed as a fraction of the incident particle flux density  $j_{inc} = \hbar k/(mL^2)$ . This is known as the *differential scattering cross section* (or probably *scattering length* in 2D):

$$\frac{d\sigma}{d\theta} = \frac{m^2 L^4}{2\pi \hbar^4 k} |\langle \vec{k}' | \hat{V} | \vec{k} \rangle|^2. \quad (7.61)$$

Note that the factor of  $1/L^2$  in the matrix element cancels the  $L^4$  in the pre-factor. The implicit  $L$  dependence in the domain of the spatial integral which defines the matrix element also falls away once  $L$  is larger than the range of the potential. Overall, the cross section is therefore independent of  $L$ . The result we derived here is known as the linear order Born approximation. We will derive this again, in three dimensions, via a different route in chapter 8.

### 7.5.3 Harmonic perturbation

Consider a harmonic perturbation which oscillates with a well-defined frequency  $\omega > 0$ . The most general form of such a perturbation is

$$\hat{V}(t) = \hat{\mathcal{V}} e^{i\omega t} + \hat{\mathcal{V}}^\dagger e^{-i\omega t} \quad (7.62)$$

with  $\hat{\mathcal{V}}$  a time-independent operator that need not be Hermitian. Using (7.40) we find that

$$\begin{aligned} c_{n \rightarrow m}^{(1)} &= \frac{[1 - e^{i(\omega_{mn} + \omega)t}]}{\hbar(\omega_{mn} + \omega)} \mathcal{V}_{mn} + \frac{[1 - e^{i(\omega_{mn} - \omega)t}]}{\hbar(\omega_{mn} - \omega)} \mathcal{V}_{nm}^* \\ &= \frac{-2i}{\hbar} \left[ \frac{e^{i(\omega_{mn} + \omega)t/2} \sin[(\omega_{mn} + \omega)t/2]}{\omega_{mn} + \omega} \mathcal{V}_{mn} + \frac{e^{i(\omega_{mn} - \omega)t/2} \sin[(\omega_{mn} - \omega)t/2]}{\omega_{mn} - \omega} \mathcal{V}_{nm}^* \right]. \end{aligned} \quad (7.63)$$

Note that the two terms above closely resemble (7.43) but with  $\omega_{mn}$  shifted by either  $+\omega$  or  $-\omega$ . From figure 7.2 it follows that at large  $t$  the first term is sharply peaked at  $\omega_{mn} = -\omega$  and the second at  $\omega_{mn} = +\omega$ . Since the overlap between these peaks are negligible the transition probability is just

$$P_{n \rightarrow m}^{(1)} = |c_{n \rightarrow m}^{(1)}|^2 = \frac{4}{\hbar^2} \left[ \frac{\sin^2[(\omega_{mn} + \omega)t/2]}{(\omega_{mn} + \omega)^2} |\mathcal{V}_{mn}|^2 + \frac{\sin^2[(\omega_{mn} - \omega)t/2]}{(\omega_{mn} - \omega)^2} |\mathcal{V}_{nm}|^2 \right]. \quad (7.64)$$

The first term of this expression is dominant when  $\omega_{mn} \approx -\omega$ , i.e. when the energy of the final state  $E_m$  is approximately  $\hbar\omega$  lower than that of the initial state. This corresponds to the process of **stimulated emission**. Similarly, the second term dominates when the final state energy is  $\hbar\omega$  higher than that of the initial state and therefore describes the process of **absorption**. This brings us to the following important conclusion:

*For a harmonic perturbation with frequency  $\omega$  transitions corresponding to the absorption and emission of energy quanta of  $\hbar\omega$  dominate. These transitions are said to be resonant with the driving frequency of the potential.*

## Fermi's golden rule for harmonic perturbations

Here we use the short-cut method for deriving the golden rule by first taking the large- $t$  limit in (7.64) and then using (7.51) to obtain

$$P_{n \rightarrow m}^{(1)} = \frac{2\pi}{\hbar} t [|\mathcal{V}_{mn}|^2 \delta(E_n - \hbar\omega - E_m) + |\mathcal{V}_{nm}|^2 \delta(E_n + \hbar\omega - E_m)] \quad (7.65)$$

$$= P_{n \rightarrow m}^{(1),em} + P_{n \rightarrow m}^{(1),abs}. \quad (7.66)$$

Note how the delta-functions fix the final state energy at precisely  $E_m = E_n \pm \hbar\omega$ . The transition rates are again constant and given by

$$\omega_{n \rightarrow m}^{(1),em/abs} = \frac{dP_{n \rightarrow m}^{(1),em/abs}(t)}{dt}. \quad (7.67)$$

Also note that

$$\omega_{n \rightarrow m}^{(1),abs} = \omega_{m \rightarrow n}^{(1),em}, \quad (7.68)$$

which says that the transition rate from  $|n\rangle$  to  $|m\rangle$  via absorption matches the transition rate from  $|m\rangle$  to  $|n\rangle$  via stimulated emission. This is known as **detailed balance**.

To obtain Fermi's golden rule we integrate over a group of final states around either  $E_m = E_n + \hbar\omega$  or  $E_m = E_n - \hbar\omega$ . This produces

$$P_{n \rightarrow [m]}^{(1),em} = \frac{2\pi t}{\hbar} |\mathcal{V}_{mn}|^2 \rho_f(E_m = E_n - \hbar\omega) \quad (7.69)$$

$$P_{n \rightarrow [m]}^{(1),abs} = \frac{2\pi t}{\hbar} |\mathcal{V}_{nm}|^2 \rho_f(E_m = E_n + \hbar\omega) \quad (7.70)$$

for the two processes respectively.

# Chapter 8

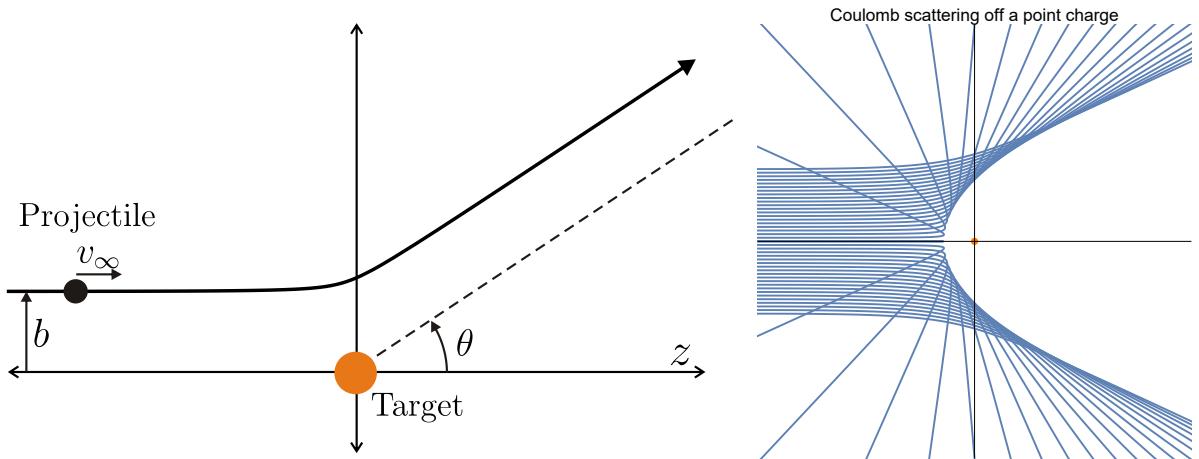
## Scattering Theory

### 8.1 Background

Scattering experiments provide a wealth of information about the properties of matter and its interactions on atomic scales and below. A variety of such experiments are performed almost daily at accelerator facilities such as CERN in Geneva or iThemba LABS just outside Stellenbosch. An enormous amount of theoretical and experiment work continues to be dedicated to predicting, measuring, and analysing the results of these experiments. In this chapter we will develop a formalism for the simplest possible version of such a scattering process. We start with a brief recap of classical scattering theory, before investigating the quantum mechanical description.

### 8.2 Classical scattering

#### 8.2.1 The scattering angle



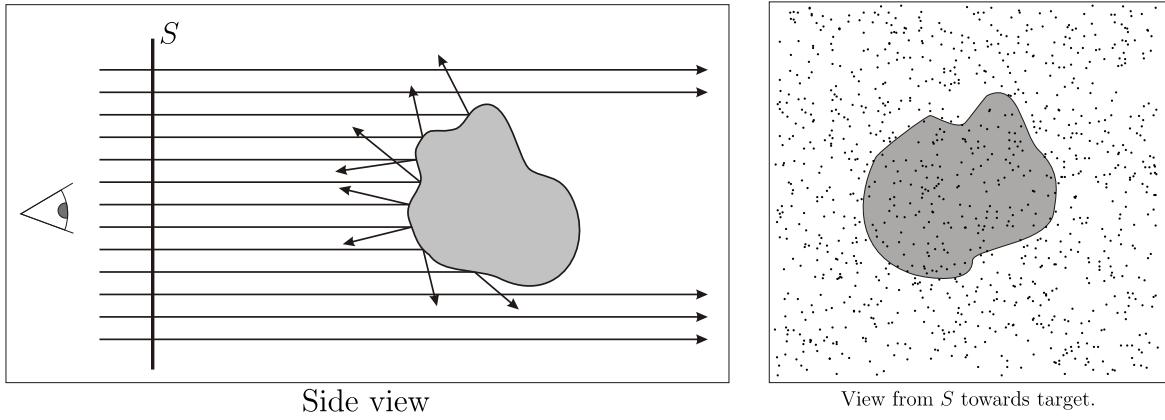
Consider a projectile particle with mass  $m$  impinging on a target at the origin. The two interact via a central potential  $V(r)$ . The particle starts out infinitely far away, moving with a speed of  $v_\infty$  along the  $z$ -direction. Its initial distance  $b$  from the  $z$ -axis is called the **impact parameter**. The particle interacts with the potential and is scattered into the direction indicated by the **scattering angle**  $\theta$ , which is a function of the impact parameter  $b$ . In fact, all the relevant information about the scattering

process is contained in this function. By combining energy and angular momentum conservation it can be shown that<sup>1</sup>

$$\theta(b) = \pi - \int_{r_{ctp}}^{\infty} dr \frac{2L}{r^2 \sqrt{2m[E - V_{eff}(r)]}} \quad (8.1)$$

where  $E = mv_\infty^2/2$ ,  $L = mbv_\infty$ , and  $V_{eff}(r) = V(r) + L^2/(2mr^2)$  is the effective radial potential. The integral's lower bound  $r_{ctp}$  is the classical turning point of the radial motion, i.e. the minimum distance from the origin that the particle reaches. It is the largest solution of the equation  $V_{eff}(r) = E$ .

### 8.2.2 The total scattering cross section



Now consider a beam of projectile particles with a range of different impact parameters. The beam has a uniform particle number density of  $n$ . The **incident particle flux density**  $j_{inc} = nv_\infty$  is the number of particles per second, per unit area, crossing the surface  $S$ . For a solid target each particle will either bounce off the surface, or miss the target entirely. The **scattering rate** is then the number of particles hitting the target per second. This is simply the product of the incident particle flux density  $j_{inc}$  and the **geometric cross-sectional area** of the target  $A_{cs}$ . The latter is a measure of how large the target appears to the particles in the beam. The **total scattering cross section** is the ratio of the scattering rate and the incident particle flux density:

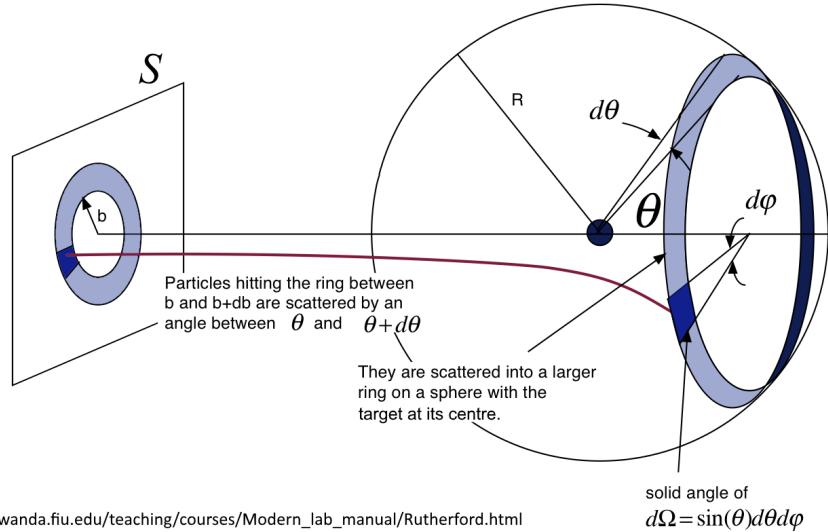
$$\sigma_{tot} = \frac{\text{Particles scattered per second}}{\text{Incident particles per second per unit area}} = \frac{A_{cs}j_{inc}}{j_{inc}} = A_{cs}. \quad (8.2)$$

For the scattering of classical point particles off a solid target there is therefore no difference between the total scattering cross section and the geometrical cross section. More generally, the total scattering cross section provides an *effective* measure of how large (as an area) the target appears to the projectile particles. In quantum mechanical scattering, wave effects such as interference and diffraction can result in the total scattering cross section being very different from the geometrical cross section.

---

<sup>1</sup>See the book of Landau and Lifshitz on mechanics, or for a more detailed account, Friedrich's book on scattering theory.

### 8.2.3 Differential scattering cross section



The **differential cross section** provides a more detailed notion of the effective target area, one which incorporates information about the direction of the scattering. Consider the setup in the figure, where we imagine a large sphere with radius  $R$  centred at the target. On the sphere's surface there is an infinitesimal patch at position  $(\theta, \phi)$ , which is subtended by the solid angle  $d\Omega = \sin(\theta)d\theta d\phi$ . The scattered particles that move through this patch must all have travelled through some specific region on the surface  $S$ . The surface area of this region on  $S$  is denoted  $d\sigma = d\sigma(\theta, \phi)$ . In other words, for a particle to be scattered into the solid angle  $d\Omega$ , it needs to pass through a particular region with surface area  $d\sigma$  on  $S$ . Equivalently, we can think of  $d\sigma$  as a more detailed version of  $\sigma_{tot}$  in (8.2):

$$d\sigma = \frac{\text{number of particles scattered into } d\Omega \text{ per second}}{\text{incident particle flux density}}. \quad (8.3)$$

We need to determine the precise connection between  $d\sigma$  and  $d\Omega$ . We first make the simplifying assumption that  $\theta(b)$  is an invertible function, i.e. that for each scattering angle  $\theta$  there is only one corresponding impact parameter  $b$ . All the particles passing through the sphere between  $\theta$  and  $\theta + d\theta$  then originally passed through a ring on  $S$  corresponding to impact parameters between  $b$  and  $b + db$ . Restricting the azimuthal angle to lie between  $\phi$  and  $\phi + d\phi$  then picks out the patch on the sphere and a corresponding segment with area  $d\sigma = b db d\phi$  on  $S$ . We can rewrite  $d\sigma$  as

$$d\sigma = b db d\phi = b \left| \frac{db}{d\theta} \right| d\theta d\phi = \frac{b}{\sin \theta} \left| \frac{db}{d\theta} \right| d\Omega. \quad (8.4)$$

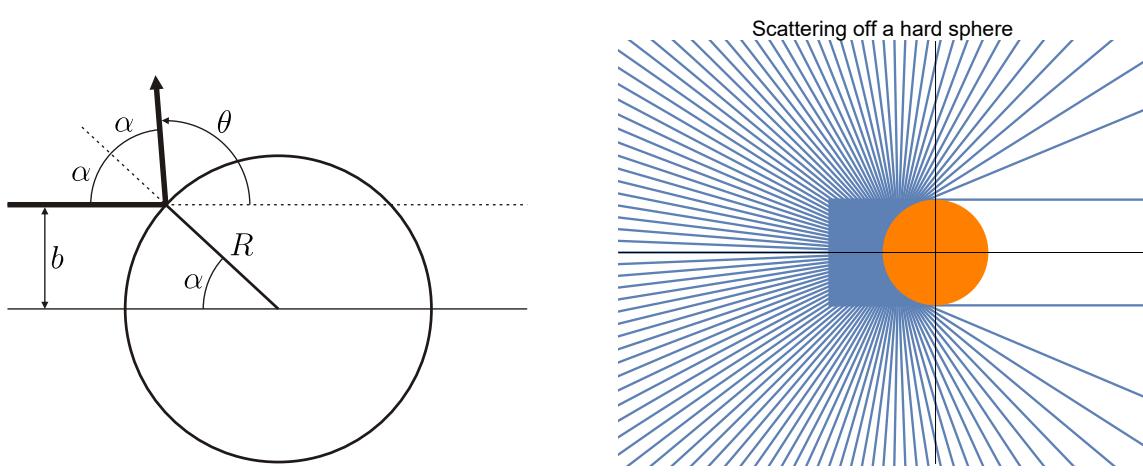
The absolute value is included to ensure that all the relevant factors remain positive. The differential cross section is now defined as

$$\frac{d\sigma}{d\Omega} = \frac{b}{\sin \theta} \left| \frac{db}{d\theta} \right|, \quad (8.5)$$

which has the units of area per unit solid angle. It is generally a function of  $\theta$  and  $\phi$ , but for a central potential it can only depend on  $\theta$ . Integrating the differential cross section over all directions gives the total cross section

$$\sigma_{tot} = \int \frac{d\sigma}{d\Omega} d\Omega. \quad (8.6)$$

#### 8.2.4 Example: Hard-sphere scattering



Suppose the target is a hard sphere with radius  $R$ . In this case we can determine the scattering angle  $\theta(b)$  using geometry alone, without the need for expression (8.1). The particle will bounce off the surface with equal incident and reflection angles (specular reflection). We can then read off from the figure that

$$\theta(b) = \pi - 2\alpha = \pi - 2 \arcsin(b/R) = 2 \arccos(b/R) \quad \text{for} \quad b \in [0, R]. \quad (8.7)$$

For  $b > R$  we have  $\theta(b) = 0$ . Clearly  $\theta(b)$  is an invertible function for impact parameters  $b \in [0, R]$ , and we can write

$$b(\theta) = R \cos(\theta/2). \quad (8.8)$$

The differential cross section is then

$$\frac{d\sigma}{d\Omega} = \frac{b}{\sin \theta} \left| \frac{db}{d\theta} \right| = \frac{R^2}{4}. \quad (8.9)$$

This is an *isotropic* cross section since it is constant, i.e. independent of direction. The total cross section is

$$\sigma_{tot} = \int \frac{d\sigma}{d\Omega} d\Omega = \pi R^2, \quad (8.10)$$

which is, as expected, just the geometric cross section.

### 8.3 Quantum Scattering - Introduction

Consider the setup depicted in figure 8.1 in which a projectile particle impinges on a stationary target at the origin. The two interact via a short-range potential  $V(\vec{r})$  generated by the target. As a result of this interaction the incident particle is scattered and registered by a detector  $D$  situated far away from the target and outside the range of  $V(\vec{r})$ . Now suppose we have a uniform beam of incident particles corresponding to a constant incident particle flux density (number of particles per second per unit area) across the surface  $S$ . We are interested in the rate at which particles are scattered into a small solid angle  $d\Omega$ , expressed as a fraction of the incident particle flux density. The differential cross section  $d\sigma(\theta, \phi)/d\Omega$  is defined as

$$\frac{d\sigma(\theta, \phi)}{d\Omega} d\Omega = \frac{\text{number of particles scattered into } d\Omega \text{ per second}}{\text{incident particle flux density}}. \quad (8.11)$$

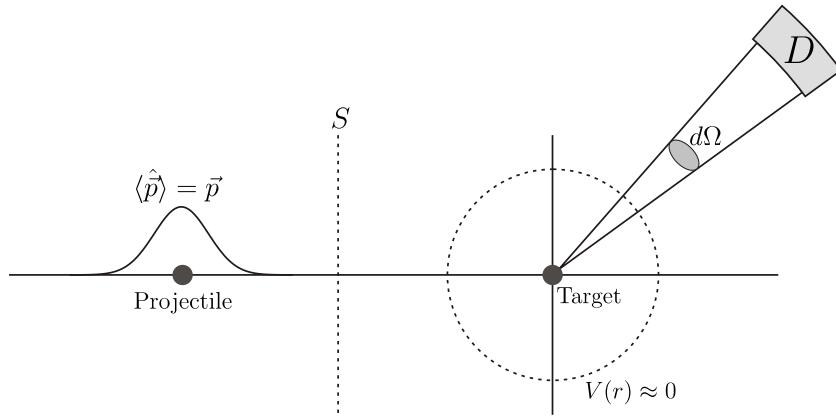


Figure 8.1: Scattering in the time-dependent picture

Calculating the differential cross section using a time-dependent wave-packet picture is technically very difficult. Luckily, a careful analysis reveals that the cross section depends almost entirely on the average momentum  $\langle \hat{p} \rangle$  of the projectile and *not* on the detailed shape of its wave function. This observation allows us to trade the complicated time-dependent description of scattering for a much simpler time-independent formalism in terms of particular eigenstates of the projectile's Hamiltonian. These eigenstates turn out to be superpositions of an incoming plane wave and outgoing scattered wave. Since these are stationary states the spatial probability density is time-independent. We can therefore imagine the incoming plane wave as describing a continuous flow of probability towards the target while the scattered wave component carries probability outward. A comparison of the incoming and outgoing probability fluxes will then yield the differential cross section. In this sense the figure below is a more accurate representation of our formalism.

**Note:** Here we are studying what is known as *potential scattering*, where the target is taken to be stationary and immobile. In particular, any recoil in the target due to the collision is ignored. A more realistic approach would be to consider a proper two-body collision, and to then analyse the problem using relative and centre-of-mass coordinates. See the books by Robinett and Shankar for details.

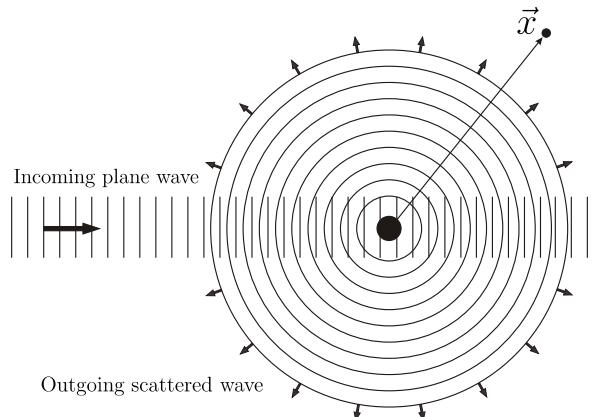


Figure 8.2: A time-independent view of scattering in terms of a stationary state which is a superposition of an incoming plane wave and an outgoing scattered wave.

## 8.4 The Lippmann-Schwinger equation

The full Hamiltonian of the incident particle is

$$\hat{H} = \hat{H}_0 + \hat{V} = \frac{\hat{p}^2}{2m} + V(\hat{x}) \quad (8.12)$$

where  $\hat{H}_0$  is the kinetic term and  $V$  the short-range scattering potential generated by the target. We seek eigenstates of  $\hat{H}$  which satisfy *scattering boundary conditions* and are of the form

$$|\psi\rangle = |\text{incoming plane wave}\rangle + |\text{outgoing scattered wave}\rangle. \quad (8.13)$$

Let  $|\phi\rangle$  denote the incoming plane wave which satisfies  $\hat{H}_0|\phi\rangle = E|\phi\rangle$ . Of course,  $|\phi\rangle$  will not be an eigenstate of the full Hamiltonian  $\hat{H}$ , just like  $|\psi\rangle$  will not be a simple plane wave everywhere in space. We require the following:

1.  $|\psi\rangle$  must satisfy  $\hat{H}|\psi\rangle = (\hat{H}_0 + \hat{V})|\psi\rangle = E|\psi\rangle$ . Note that  $|\psi\rangle$  and  $|\phi\rangle$  share the same energy  $E$ , but with respect to two different Hamiltonians. This reflects the elastic nature of the scattering process.
2.  $|\psi\rangle$  should reduce to  $|\phi\rangle$  when  $\hat{V} = 0$ .

The Lippmann-Schwinger equation is an implicit equation for  $|\psi\rangle$  which captures these properties. To derive it we first rewrite the eigenvalue equation for  $|\psi\rangle$  as  $(E - \hat{H}_0)|\psi\rangle = \hat{V}|\psi\rangle$  and then multiply on both sides by  $(E - \hat{H}_0)^{-1}$ . We will say more about this inverse shortly. The eigenvalue equation now reads

$$|\psi\rangle = \frac{1}{E - \hat{H}_0} \hat{V}|\psi\rangle. \quad (8.14)$$

However, this still does not reflect the fact that  $|\psi\rangle$  should reduce to  $|\phi\rangle$  when  $\hat{V} = 0$ . We can fix this by adding  $|\phi\rangle$  on the right:

$$|\psi\rangle = \frac{1}{E - \hat{H}_0} \hat{V}|\psi\rangle + |\phi\rangle. \quad (8.15)$$

Any  $|\psi\rangle$  satisfying this equation clearly reduces to  $|\phi\rangle$  when  $\hat{V} = 0$ . Furthermore, multiplying by  $(E - \hat{H}_0)$  on both sides and using  $\hat{H}_0|\phi\rangle = E|\phi\rangle$  reveals that such a  $|\psi\rangle$  will *still* satisfy  $\hat{H}|\psi\rangle = (\hat{H}_0 + \hat{V})|\psi\rangle = E|\psi\rangle$ . Solutions to (8.15) therefore exhibit the two required properties above.

Let us now return to the inverse  $(E - \hat{H}_0)^{-1}$  and clarify some technicalities.<sup>2</sup> First of all, note that since  $(E - \hat{H}_0)|\phi\rangle = 0$ , the operator  $E - \hat{H}_0$  has a zero eigenvalue. It is well-known that a finite dimensional matrix with a zero eigenvalue is not invertible. However, here we are dealing with an operator acting on an infinite dimensional vector space, and which has a continuous set of eigenstates and eigenvalues. The zero eigenvalue is an element of this continuous set. This turns out to be less severe of

---

<sup>2</sup>The notes <http://bohr.physics.berkeley.edu/classes/221/1112/notes/greensfuns.pdf> contain a very readable discussion of these concepts.

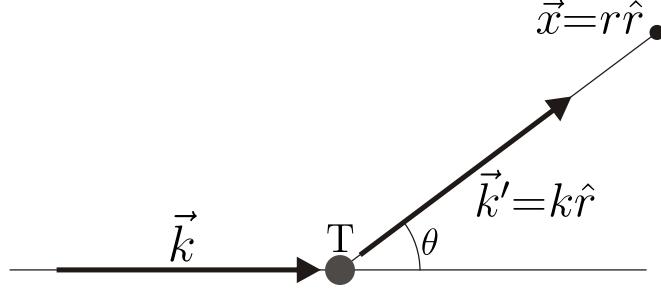


Figure 8.3: Interpretation of  $\vec{k}'$  as the scattered particle's outgoing momentum.

a problem than in the finite dimensional matrix case, where the eigenvalues form a discrete set. In fact, the operator  $(E - \hat{H}_0)$  is invertible, but due to its zero eigenvalue this inverse is not unique. We need to choose the correct inverse to ensure that the solutions of the Lippmann-Schwinger equation have the physical properties of states which describe scattering. This amounts to imposing a second scattering boundary condition. Luckily, selecting the correct inverse turns out to be quite simple: we just replace  $(E - \hat{H}_0)^{-1}$  by  $(E - \hat{H}_0 \pm i\epsilon)^{-1}$  where  $\epsilon$  is real and positive. Since  $E - \hat{H}_0 \pm i\epsilon$  has no zero eigenvalues its inverse is unique. At the end of the calculation  $\epsilon$  will be sent to zero, and the choice of sign will be determined by the boundary conditions later on. This brings us to

$$|\psi^\pm\rangle = \frac{1}{E - \hat{H}_0 \pm i\epsilon} \hat{V} |\psi^\pm\rangle + |\phi\rangle, \quad (8.16)$$

which is known as the Lippmann-Schwinger equation. It is straightforward to rewrite this in coordinate space as

$$\psi^\pm(\vec{x}) = \phi(\vec{x}) + \int d\vec{x}' \langle \vec{x} | (E - \hat{H}_0 \pm i\epsilon)^{-1} | \vec{x}' \rangle V(\vec{x}') \psi^\pm(\vec{x}'). \quad (8.17)$$

Henceforth we choose  $|\phi\rangle = |\vec{k}\rangle$  with  $\vec{p} = \hbar\vec{k}$  the momentum of the incoming plane wave. The energy is then  $E = \hbar^2 k^2 / (2m)$  while

$$\phi(\vec{x}) = \frac{1}{(2\pi)^{3/2}} e^{i\vec{k}\cdot\vec{x}}. \quad (8.18)$$

The matrix element  $\langle \vec{x} | (E - \hat{H}_0 \pm i\epsilon)^{-1} | \vec{x}' \rangle$  can be calculated using contour integral methods. This leads to

$$\psi^\pm(\vec{x}) = \phi(\vec{x}) - \frac{2m}{\hbar^2} \int d\vec{x}' \frac{e^{\pm ik|\vec{x}-\vec{x}'|}}{4\pi|\vec{x}-\vec{x}'|} V(\vec{x}') \psi^\pm(\vec{x}'). \quad (8.19)$$

Note that the  $\epsilon \rightarrow 0^+$  limit has been taken but that the choice of sign remains. This result is still exact and valid for all  $\vec{x}$ . Calculating the cross section only requires knowledge of  $\psi^\pm(\vec{x})$  at points far away from the target. In the large- $|\vec{x}|$  limit the fact that the potential is short-ranged allows the integrand of (8.19) to be simplified using

$$\frac{e^{\pm ik|\vec{x}-\vec{x}'|}}{|\vec{x}-\vec{x}'|} \approx \frac{e^{\pm ikr} e^{\mp i\vec{k}'\cdot\vec{x}'}}{r} \quad (8.20)$$

where

$$r = |\vec{x}|, \quad \hat{r} = \vec{x}/r \quad \text{and} \quad \vec{k}' = k\hat{r}. \quad (8.21)$$

Here  $\vec{k}'$  is a wave vector with magnitude  $k$  pointing towards  $\vec{x}$ . Note that  $\vec{k}'$  has nothing to do with the integration variable  $\vec{x}'$ . In the large- $r$  limit equation (8.19) becomes

$$\psi^\pm(\vec{x}) = \frac{1}{(2\pi)^{3/2}} \left[ e^{i\vec{k}\cdot\vec{x}} + \frac{e^{\pm ikr}}{r} f(\vec{k}', \vec{k}) \right] \quad (\mathbf{r} \rightarrow \infty). \quad (8.22)$$

where

$$f(\vec{k}', \vec{k}) = -\frac{1}{4\pi} (2\pi)^3 \frac{2m}{\hbar^2} \int d\vec{x}' \frac{e^{\mp i\vec{k}'\cdot\vec{x}'}}{(2\pi)^{3/2}} V(\vec{x}') \psi^\pm(\vec{x}') \quad (8.23)$$

is the **scattering amplitude**. In (8.22) the  $e^{i\vec{k}\cdot\vec{x}}$  term clearly represents the incoming plane wave. The second term accounts for the scattered wave and is the product of a spherical wave  $e^{\pm ikr}/r$  with the scattering amplitude  $f(\vec{k}', \vec{k})$ . The latter depends on  $\vec{k}$  as well as  $\hat{r}$ , but *not* on  $r = |\vec{x}|$  itself. At this point we note that for  $\psi(x)$  to satisfy scattering boundary conditions the spherical wave must have a momentum directed away from the target. *This requires choosing the top sign everywhere.* Henceforth we write  $\psi(x)$  for  $\psi^+(x)$ . The scattering amplitude now reads

$$f(\vec{k}', \vec{k}) = -\frac{1}{4\pi} (2\pi)^3 \frac{2m}{\hbar^2} \int d\vec{x}' \frac{e^{-i\vec{k}'\cdot\vec{x}'}}{(2\pi)^{3/2}} V(\vec{x}') \psi(\vec{x}') = -\frac{1}{4\pi} (2\pi)^3 \frac{2m}{\hbar^2} \langle \vec{k}' | \hat{V} | \psi \rangle. \quad (8.24)$$

To interpret this expression, imagine detecting the scattered particle at position  $\vec{x}$ . The magnitude of the particle's momentum would be  $p = \hbar k$  and since  $\vec{x}$  points in the  $\hat{r}$  direction it is natural to interpret  $\vec{k}' = k\hat{r}$  as the particle's final or outgoing momentum. The scattering amplitude can therefore conveniently be considered a function of the incoming and outgoing (or initial and final) momenta  $\vec{k}$  and  $\vec{k}'$ . See figure 8.3. (Of course, until we detect the particle somewhere we cannot really talk about a well-defined final momentum. The outgoing scattered wave is a superposition of plane waves propagating in various different directions.)

## 8.5 The scattering amplitude and differential cross section

Next we relate the differential cross section  $d\sigma(\theta, \phi)/d\Omega$  as defined in (8.11) to the scattering amplitude  $f(\vec{k}', \vec{k})$ . Recall that the probability density  $\rho(\vec{x}, t) = |\Psi(\vec{x}, t)|^2$  associated with a state  $\Psi(\vec{x}, t)$  obeys the continuity equation

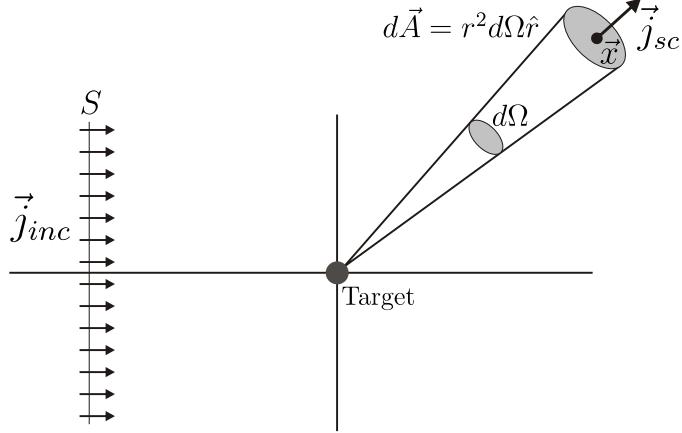
$$\frac{\partial \rho}{\partial t} + \vec{\nabla} \cdot \vec{j} = 0 \quad (8.25)$$

where

$$\vec{j}(\vec{x}, t) = \frac{\hbar}{m} \text{Im} \left[ \Psi^*(\vec{x}, t) \vec{\nabla} \Psi(\vec{x}, t) \right] \quad (8.26)$$

is the probability flux density. The probability flux through an infinitesimal flat surface with vector area  $d\vec{A}$  is therefore  $\vec{j} \cdot d\vec{A}$ . We expect the ratio of particle fluxes that define the cross section to be equal to the corresponding ratio of probability fluxes:

$$\frac{d\sigma(\theta, \phi)}{d\Omega} d\Omega = \frac{\text{probability flux through the surface subtended by } d\Omega \text{ at } \vec{x}}{\text{incident probability flux density through the surface } S}. \quad (8.27)$$



Next we write (8.22) as  $\psi(\vec{x}) = \psi_{inc}(\vec{x}) + \psi_{sc}(\vec{x})$  where  $\psi_{inc}(\vec{x})$  is the incident plane wave and  $\psi_{sc}(\vec{x})$  the outgoing scattered wave. The associated probability flux densities are found to be

$$\vec{j}_{inc} = \frac{\hbar}{(2\pi)^3 m} \vec{k} \quad \text{and} \quad \vec{j}_{sc}(\vec{x}) = \frac{\hbar k}{(2\pi)^3 m} \frac{|f(\vec{k}', \vec{k})|^2}{r^2} \hat{r} \quad (\mathbf{r} \rightarrow \infty) \quad (8.28)$$

where  $\vec{k}' = k\hat{r}$ . The expression for  $\vec{j}_{sc}(\vec{x})$  was derived in the large- $r$  limit. As expected  $\vec{j}_{inc}(\vec{x})$  is uniform and directed along  $\vec{k}$  while  $\vec{j}_{sc}(\vec{x})$  is directed radially outward. We conclude that

$$\frac{d\sigma(\theta, \phi)}{d\Omega} d\Omega = \frac{|\vec{j}_{sc}(\vec{x})| r^2 d\Omega}{|\vec{j}_{inc}|} = |f(\vec{k}', \vec{k})|^2 d\Omega \quad (8.29)$$

and so

$$\frac{d\sigma(\theta, \phi)}{d\Omega} = |f(\vec{k}', \vec{k})|^2 \quad (8.30)$$

where  $\vec{k}' = k\hat{r}$  contains the  $(\theta, \phi)$ -dependence. The relative rate at which particles are scattered *in any direction* is given by the total cross section

$$\sigma_{tot} = \int d\Omega \frac{d\sigma}{d\Omega} = \int d\theta d\phi \sin(\theta) |f(\vec{k}', \vec{k})|^2. \quad (8.31)$$

Remarkably, the total cross section is related to the scattering amplitude in the forward direction by

$$\sigma_{tot} = \frac{4\pi}{k} \text{Im}[f(\theta = 0)]. \quad (8.32)$$

This result is known as the **Optical theorem**. We will prove it for spherically symmetric potentials using the partial wave expansion in a later section.

### Summary:

We have developed a time-independent description of scattering in terms of eigenstates of  $\hat{H} = \hat{H}_0 + \hat{V}$  which are superpositions of an incoming plane wave and an outgoing scattered wave. In the large- $r$  limit these states were found to be of the form

$$\psi(\vec{x}) = \frac{1}{(2\pi)^{3/2}} \left[ e^{i\vec{k}\cdot\vec{x}} + \frac{e^{ikr}}{r} f(\vec{k}', \vec{k}) \right] \quad \text{with} \quad f(\vec{k}', \vec{k}) = -\frac{1}{4\pi} (2\pi)^3 \frac{2m}{\hbar^2} \langle \vec{k}' | \hat{V} | \psi \rangle. \quad (8.33)$$

Here  $f(\vec{k}', \vec{k})$  is the scattering amplitude which determines the cross section directly via (8.30). Note that this is still an *implicit* equation for  $\psi(\vec{x})$ . Our focus now shifts to finding solutions, whether approximate or exact, for the scattering amplitude. Two standard methods for doing so, namely the Born approximation and the partial wave expansion, are considered next.

## 8.6 The Born approximation

In the same way as we constructed the Dyson series for  $\hat{U}_I(t, t_0)$  in the previous chapter we will now use (8.33) to derive a perturbation series for  $\psi(\vec{x})$  in orders of the scattering potential  $V(\vec{x})$ . Of course, this requires that the potential is in some sense “weak”. We will investigate this restriction later. Proceeding as before we find that at zeroth order

$$\psi^{(0)}(\vec{x}) = \frac{1}{(2\pi)^{3/2}} e^{i\vec{k}\cdot\vec{x}} = \langle \vec{x} | \vec{k} \rangle \quad (8.34)$$

and therefore  $f^{(0)}(\vec{k}', \vec{k}) = 0$ . From this the linear-order result follows as

$$\psi^{(1)}(\vec{x}) = \frac{1}{(2\pi)^{3/2}} \left[ e^{i\vec{k}\cdot\vec{x}} + \frac{e^{ikr}}{r} f^{(1)}(\vec{k}', \vec{k}) \right] \quad \text{with} \quad f^{(1)}(\vec{k}', \vec{k}) = -\frac{1}{4\pi} (2\pi)^3 \frac{2m}{\hbar^2} \langle \vec{k}' | \hat{V} | k \rangle. \quad (8.35)$$

To this order the scattering amplitude is therefore given by the matrix element of the scattering potential between the incoming and outgoing momentum states. Explicitly,

$$f^{(1)}(\vec{k}', \vec{k}) = -\frac{1}{4\pi} \frac{2m}{\hbar^2} \int d\vec{x}' e^{i(\vec{k}-\vec{k}')\cdot\vec{x}'} V(\vec{x}') \quad (8.36)$$

which is just the Fourier transform of the scattering potential with respect to  $\vec{q} = \vec{k} - \vec{k}'$ , the momentum transferred to the target. If the potential is spherically symmetric we can choose the  $\hat{z}$ -axis of the  $\vec{x}'$ -coordinate system to lie along  $\vec{q}$  and then perform the angular integrals to produce

$$f^{(1)}(\vec{k}', \vec{k}) = -\frac{2m}{\hbar^2} \frac{1}{q} \int_0^\infty dr r V(r) \sin(qr) = f^{(1)}(q) \quad (8.37)$$

where  $q = |\vec{k} - \vec{k}'| = 2k \sin(\theta/2)$  and  $\theta$  is the deflection angle between the incoming and outgoing momenta. See figure 8.3.

When comparing (8.36) to (7.40) we see that both expressions are Fourier transforms, one with respect to space and the other to time. This is again an illustration of how the relationship between space and momentum is analogous to that between time and energy. In particular, it shows that how a potential oscillates in space determines how it affects the momentum of a particle, while how it oscillates in time determines how it affects the particle’s energy.

### 8.6.1 Validity of the Born approximation

Although the Born approximation is both simple to state and apply, formulating sufficient conditions under which it is valid is no easy task. In fact, even the question of

whether the Born series converges is non-trivial, since the cross section is generally not an analytic function of the scattering potential's effective strength. Here we only state one necessary condition for the Born approximation to hold which reflects the notion that the potential must, in some sense, be weak.

To obtain  $f^{(1)}(\vec{k}', \vec{k})$  in (8.36) we replaced  $|\psi\rangle$  by  $|\vec{k}\rangle$  in (8.33), that is, we approximated the full wave function  $\psi(\vec{x}) = \psi_{inc}(\vec{x}) + \psi_{sc}(\vec{x})$  by its plane wave component  $\psi_{inc}(\vec{x})$ . This is valid when the distortion of the incident plane wave by the scattering potential is not too great and  $\psi_{sc}(\vec{x})$  is small compared to  $\psi_{inc}(\vec{x})$ . A necessary condition for this is

$$\left| \frac{\psi_{sc}^{(1)}(\vec{x})}{\psi_{inc}(\vec{x})} \right| \ll 1. \quad (8.38)$$

Generally we expect  $V(\vec{x})$  to have a maximum at  $\vec{x} = 0$  and that the distortion in the incoming state will be greatest there. It should then be sufficient to require that

$$\left| \frac{\psi_{sc}^{(1)}(0)}{\psi_{inc}(0)} \right| = \left| \frac{1}{4\pi} \frac{2m}{\hbar^2} \int d\vec{x}' \frac{e^{ikr'}}{r'} V(\vec{x}') e^{i\vec{k}\cdot\vec{x}'} \right| \ll 1. \quad (8.39)$$

Here  $\psi_{sc}^{(1)}(0)$  follows from replacing  $\psi(\vec{x}')$  by  $\phi(\vec{x}')$  in the second term of (8.19). If  $V(\vec{x})$  is spherically symmetric this simplifies to

$$\left| \frac{\psi_{sc}^{(1)}(0)}{\psi_{inc}(0)} \right| = \left| \frac{2m}{\hbar^2} \frac{1}{k} \int_0^\infty dr r e^{ikr} V(r) \sin(kr) \right| \ll 1. \quad (8.40)$$

This condition is always met when  $k$  is sufficiently large, i.e. if the energy of the incident particle is so high that it is hardly affected by the potential. Let us therefore consider the low-energy limit where  $k \rightarrow 0$ . The constraint now becomes

$$\frac{2m}{\hbar^2} \left| \int_0^\infty dr r V(r) \right| \ll 1. \quad (8.41)$$

This clearly amounts to a restriction on the effective strength of the potential. However, as this ad hoc derivation suggests, it is often much simpler to make approximations than to quantify their range of validity.

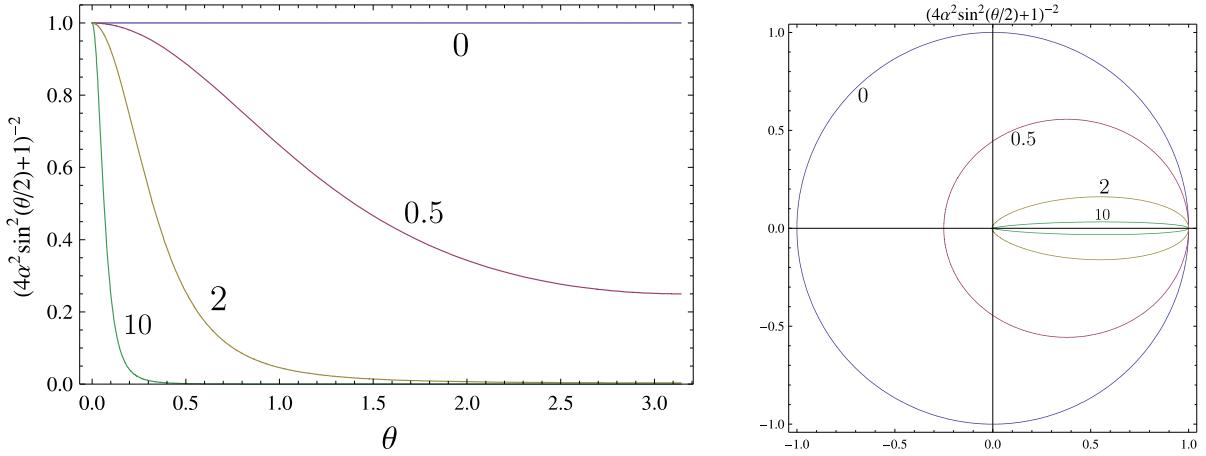


Figure 8.4: Shape of the differential cross section for the Yukawa potential as a function of  $\theta$  for  $k/\mu = 0, 0.5, 2, 10$ . The figure on the right is a polar plot.

### 8.6.2 Applications of the Born approximation

#### The Yukawa potential

The Yukawa potential is defined as

$$V(r) = \frac{V_0 e^{-\mu r}}{\mu r} \quad (8.42)$$

where  $V_0$  and  $\mu$  are parameters which control the potential's strength and range respectively. The linear-order Born approximation gives

$$f^{(1)}(q) = \frac{-2mV_0}{\mu\hbar^2 q} \int_0^\infty dr e^{-\mu r} \sin(qr) \quad (8.43)$$

$$= \frac{-2mV_0}{\mu\hbar^2 q} \int_0^\infty dr e^{-\mu r} \frac{1}{2i} [e^{iqr} - e^{-iqr}] \quad (8.44)$$

$$= \frac{-2mV_0}{\mu\hbar^2} \frac{1}{q^2 + \mu^2} \quad (8.45)$$

where  $q = |\vec{k} - \vec{k}'| = 2k \sin(\theta/2)$ . The differential cross section is

$$\frac{d\sigma}{d\Omega} = \left[ \frac{2mV_0}{\mu\hbar^2} \right]^2 \frac{1}{(4k^2 \sin^2(\theta/2) + \mu^2)^2}. \quad (8.46)$$

In the low energy limit where  $k \ll \mu$  the angular dependence is very weak and the cross section becomes approximately isotropic. This is clearly reflected in figure 8.4. As the energy increases the differential cross section becomes increasingly focussed in the forward direction around  $\theta = 0$ , since at these energies the projectiles are hardly deflected by the potential. These high and low energy features of the cross section are quite generic and common to a variety of short-range potentials.

## Coulomb scattering

Suppose the target and projectile have charges  $Z_1 e$  and  $Z_2 e$  and interact via the Coulomb potential

$$V(r) = \frac{Z_1 Z_2 e^2}{r}. \quad (8.47)$$

If we insert this potential directly into expression (8.37) for the Born approximation we are confronted with a non-convergent integral. This reflects the fact that the Coulomb potential is long-ranged and cannot be treated using our formalism. To address this problem a much more sophisticated approach is required which we will not consider here. However, if we are lucky we may be able to extract results for the Coulomb potential by considering an appropriate limit of the results found in the Yukawa case. From the form of the Yukawa potential in (8.42) it is clear that taking  $\mu, V_0 \rightarrow 0$  while keeping  $V_0/\mu$  fixed at  $Z_1 Z_2 e^2$  produces the Coulomb potential. Taking the same limit in the cross section (8.46) yields

$$\frac{d\sigma}{d\Omega} = \left[ \frac{Z_1 Z_2 e^2}{E} \right]^2 \frac{1}{16 \sin^4(\theta/2)} \quad (8.48)$$

with  $E = \hbar^2 k^2 / (2m) = p^2 / (2m)$  the energy of the projectile. This is the famous **Rutherford cross section**. Amazingly, this result turns out to be *exact* for both the classical and quantum scattering problem. Note that in the Rutherford cross section we can absorb  $\hbar$  into the energy  $E = \hbar^2 k^2 / (2m)$ , leaving no parameters which are inherently quantum mechanical in nature. This makes the exact match with the classical result possible. For the Yukawa cross section this is not possible. You can introduce  $E = \hbar^2 k^2 / (2m)$  into that expression, but there will always be an  $\hbar$  or a  $k$  stuck somewhere. Of course, classically there is no notion of  $\hbar$ , or of a wave number associated with the projectile.

Also note that the total Rutherford cross section is infinite as a result of the divergent  $1/\sin^4(\theta/2)$  factor in the differential cross section. This is a consequence of the essentially infinite range of the potential, which results in all particles undergoing scattering, regardless of their impact parameters. In this sense the target appears infinitely large to the projectiles. For the short-range Yukawa potential the total cross section is clearly finite.

## 8.7 Partial wave expansion

### 8.7.1 Background

The partial wave expansion is a non-perturbative technique for investigating scattering by spherically symmetric potentials. We will assume that the incoming plane wave travels along the  $z$ -direction with wave vector  $\vec{k} = k\hat{z}$ . Both this plane wave and the scattering potential are therefore cylindrically symmetric around the  $\hat{z}$ -axis and so the scattered wave will also exhibit this symmetry. All the wave functions considered in this section are therefore independent of the azimuthal angle  $\phi$ . As before we assume that the potential is short ranged and that beyond a certain distance  $R$  the projectile particle may be considered to be free.

### 8.7.2 Some additional comments

This section can be intimidating at first. In short, the strategy will be as follows:

1. We are interested in eigenstates of  $\hat{H}$  of the form  $\psi(\vec{x}) = \psi_{inc}(\vec{x}) + \psi_{sc}(\vec{x})$ , i.e. superpositions of an incoming plane wave and outgoing scattered wave. We are going to expand both sides as series of partial waves (which are  $\hat{L}^2$ ,  $\hat{L}_z$  eigenstates). We are interested in the large- $r$  behaviour of these wave functions.
2. We then compare the expansions on the two sides to try and relate the scattering amplitude (contained in  $\psi_{sc}(\vec{x})$ ) to the potential  $V(r)$  (which imprints itself on  $\psi(\vec{x})$  via the radial wave equation).

The main assumptions here are that the potential is spherically symmetric and that it is either finite ranged (i.e. zero beyond some distance  $R$ ), or short-ranged. The latter term is not always clearly defined in the literature. It is certainly sufficient that  $V(r)$  tends to zero exponentially quickly as  $r \rightarrow \infty$ , which is what we will assume. Power law behaviour is much more tricky, and requires a more detailed analysis. However, we will see that some of our results still work provided that  $V(r)$  tends to zero like  $1/r^2$  or faster. To set out the steps more clearly, we will assume that there is a distance  $R$  beyond which we can approximate the potential as zero, and where the solution of the radial wave equation is therefore a linear combination of the Bessel and Neumann functions. This is not strictly necessary, but doing so allows us to obtain the large- $r$  behaviour of the wave functions in two clear steps, by first looking at radii beyond  $R$ , and then using the asymptotic  $r \rightarrow \infty$  forms of the Bessel and Neumann functions.

### 8.7.3 Asymptotic form of $\hat{H}$ eigenstates

Since the Hamiltonian  $\hat{H}$  is spherically symmetric it commutes with the generators of rotations, i.e. the angular momentum operators. The eigenstates of  $\hat{H}$  may therefore be chosen to also be eigenstates of  $\hat{L}^2$  and  $\hat{L}_z$ . The wave function of such a state has the form

$$\psi_{Elm}(\vec{x}) = R_{El}(r)Y_m^l(\theta\phi) \quad (8.49)$$

where  $E = \hbar^2 k^2 / (2\mu)$  is the energy eigenvalue<sup>3</sup>. The radial component  $R_{El}(r)$  must satisfy the radial wave equation

$$\left[ \frac{-\hbar^2}{2\mu} \left[ \frac{1}{r^2} \frac{d}{dr} r^2 \frac{d}{dr} - \frac{l(l+1)}{r^2} \right] + V(r) \right] R_{El}(r) = E R_{El}(r). \quad (8.50)$$

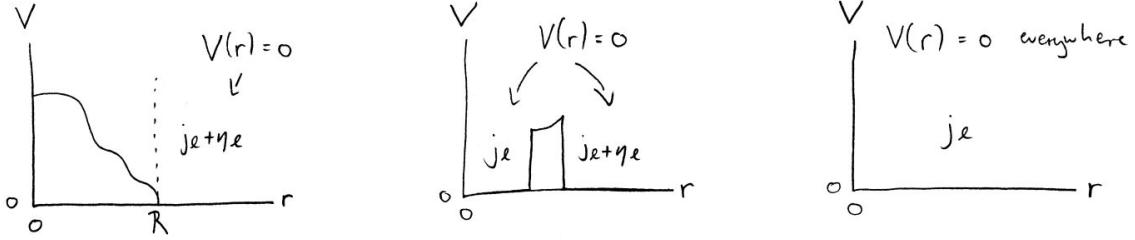
In a region where  $V(r) = 0$  the general solution to this equation is

$$R_{El}(r) = A_l j_l(kr) + B_l \eta_l(kr) \quad (8.51)$$

with  $j_l$  and  $\eta_l$  the spherical Bessel and Neumann functions respectively. For details see section D of the appendix. Note that the Neumann function is singular at the origin. *If the region in which (8.51) is applicable includes the origin we must therefore set  $B_l = 0$ .* This is the case for a free particle for which  $V(r) = 0$  everywhere. See figure below.

---

<sup>3</sup>In this section  $\mu$  denotes the projectile particle's mass.



A general eigenstate of  $\hat{H}$  will be a superposition of the form

$$\psi(\vec{x}) = \sum_{l,m} c_{lm} \psi_{Elm}(\vec{x}) = \sum_{l,m} c_{lm} R_{El}(r) Y_m^l(\theta, \phi). \quad (8.52)$$

Since  $\psi(\vec{x})$  is cylindrically symmetric, i.e.  $\phi$ -independent, only the  $m = 0$  terms contribute:

$$\psi(\vec{x}) = \sum_{l=0}^{\infty} c_l \psi_{El0}(\vec{x}) = \sum_{l=0}^{\infty} c_l R_{El}(r) Y_0^l(\theta, \phi). \quad (8.53)$$

For  $r > R$  we may approximate  $V(r) \approx 0$  and this then becomes

$$\psi(\vec{x}) = \sum_{l=0}^{\infty} (i)^l [A_l j_l(kr) + B_l \eta_l(kr)] P_l(\cos \theta) \quad (\mathbf{r} > \mathbf{R}). \quad (8.54)$$

The ratio  $B_l/A_l$  will turn out to be very important. The asymptotic behaviour of  $\psi(\vec{x})$  at large  $r$  is given by

$$\psi(\vec{x}) = \sum_{l=0}^{\infty} \frac{P_l(\cos \theta)}{2ki} \left[ (A_l - iB_l) \frac{e^{ikr}}{r} - (A_l + iB_l) \frac{e^{-ikr} e^{i\pi l}}{r} \right] \quad (\mathbf{r} \rightarrow \infty). \quad (8.55)$$

We see that  $\psi(\vec{x})$  is a linear combination of incoming and outgoing spherical waves, one for each value of  $l$ . Since  $\psi(\vec{x})$  is a stationary state the probability density  $|\psi(\vec{x})|^2$  must be time-independent. This requires that the incoming and outgoing fluxes cancel. The amplitudes of the incoming and outgoing spherical waves must therefore have the same magnitude, and so they can only differ by a phase:

$$e^{2i\delta_l} \equiv \frac{A_l - iB_l}{A_l + iB_l} = \frac{1 - iB_l/A_l}{1 + iB_l/A_l}. \quad (8.56)$$

Here  $\delta_l$  is known as the **phase shift** of the  $l$ 'th partial wave. Once  $B_l/A_l$  is known we can calculate  $\delta_l$  and, as we will see, also the scattering amplitude.

A better understanding of the phase shift can be gained by comparing the asymptotic form in (8.55) to that of a free particle. Consider the  $l$ 'th partial wave in (8.55). The radial part can be written as

$$R_{El}(r) \propto \frac{\sin(kr - l\pi/2 + \delta_l)}{kr} \quad (\text{General } \hat{V}) \quad (\mathbf{r} \rightarrow \infty). \quad (8.57)$$

up to  $r$ -independent constants. This is just a sine wave with a  $1/r$  dampening factor. In the case of a free particle we must set  $B_l = 0$  and so  $\delta_l = 0$ . This implies that

$$R_{El}(r) \propto \frac{\sin(kr - l\pi/2)}{kr} \quad (\hat{V} = 0) \quad (\mathbf{r} \rightarrow \infty). \quad (8.58)$$

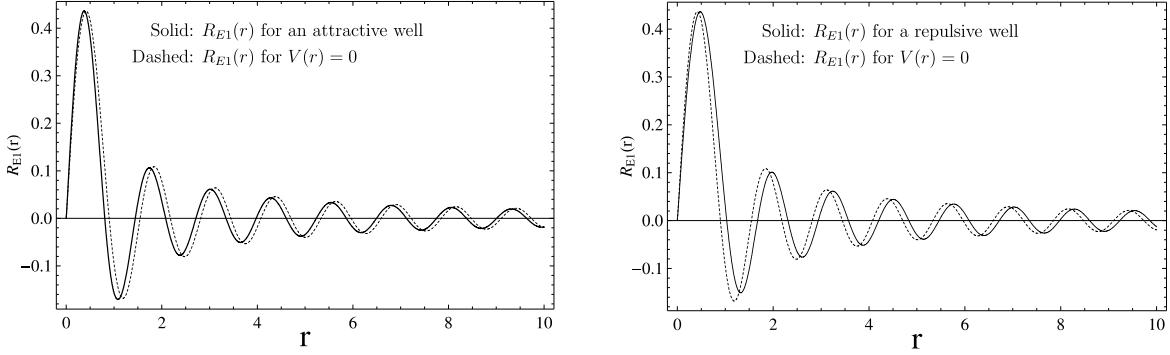


Figure 8.5: Solutions of the radial wave equation for an attractive and repulsive spherical well with  $V(r) = 0$  at  $r > R$  and  $V(r) = \pm V_0$  at  $r \leq R$ . Note that at  $r > R$  the presence of the potential simply introduces a phase shift relative to the free particle solution. Here  $R = 1$ .

*Amazingly, the entire effect of the scattering potential is to introduce a phase shift in the radial part of the wave function far away from the target. All the information about the potential relevant to the scattering process is contained in these phase shifts.* Generally one finds that if the potential is weakly repulsive then  $\delta_l < 0$  and the radial wave function is pushed slightly outward. For a weakly attractive potential  $\delta_l > 0$  and the wave function is pulled inward. See figure 8.5.

We are interested in those eigenstates of  $\hat{H}$  which satisfy scattering boundary conditions and are of the form

$$\psi(\vec{x}) = \psi_{inc}(\vec{x}) + \psi_{sc}(\vec{x}) = \frac{1}{(2\pi)^{3/2}} e^{i\vec{k}\cdot\vec{x}} + \frac{1}{(2\pi)^{3/2}} \frac{e^{ikr}}{r} f(k, \theta) \quad (\mathbf{r} \rightarrow \infty). \quad (8.59)$$

The general large- $r$  asymptotic form of  $\psi(\vec{x})$  has already been found in (8.55), and we next derive the asymptotic forms of  $\psi_{inc}(\vec{x})$  and  $\psi_{sc}(\vec{x})$ . Inserting these into (8.59) will then lead to the desired connection between the phase shifts and the scattering amplitude.

### 8.7.4 Asymptotics of the incident plane wave

Earlier we encountered the free particle states

$$\psi_{Elm}(\vec{x}) = j_l(kr) Y_m^l(\theta, \phi). \quad (8.60)$$

However, the plane wave states

$$\psi_{\vec{k}}(\vec{x}) = \frac{1}{(2\pi)^{3/2}} e^{i\vec{k}\cdot\vec{x}} \quad (8.61)$$

are *also* eigenstates of the free particle Hamiltonian but look nothing like  $\psi_{Elm}(\vec{x})$ . How is this possible? Remember that the free particle Hamiltonian is highly degenerate: there are many different states with exactly the same energy. Within these degenerate subspaces we may choose the basis in any way we want. Furthermore, the free particle Hamiltonian is invariant under both rotations and translations: it commutes with

the linear *and* angular momentum operators. If we choose our basis states to be simultaneous eigenstates of  $\hat{H}_0$ ,  $\hat{L}^2$  and  $\hat{L}_z$  we are led to the  $\psi_{Elm}(\vec{x})$  states. If we instead construct eigenstates of  $\hat{H}_0$  and  $\hat{\vec{p}} = (\hat{p}_x, \hat{p}_y, \hat{p}_z)$  we obtain the plane wave states. Of course, these are just two different bases for the same Hilbert space, and so we must be able to express one basis in terms of the other. In particular, we can expand the incident place wave  $\psi_{inc}(\vec{x})$  in terms of partial waves as

$$\psi_{inc}(\vec{x}) = \frac{1}{(2\pi)^{3/2}} e^{i\vec{k}\cdot\vec{x}} = \frac{1}{(2\pi)^{3/2}} \sum_{l,m} 4\pi(i)^l [Y_m^l(\theta_k, \phi_k)]^* Y_m^l(\theta, \phi) j_l(kr). \quad (8.62)$$

This result is still completely general. Using the fact that  $\vec{k} = k\hat{z}$  simplifies this to

$$\psi_{inc}(\vec{x}) = \frac{1}{(2\pi)^{3/2}} e^{ikr \cos \theta} = \frac{1}{(2\pi)^{3/2}} \sum_l (2l+1)(i)^l j_l(kr) P_l(\cos \theta) \quad (8.63)$$

and in the large- $r$  limit we then find

$$\psi_{inc}(\vec{x}) = \frac{1}{(2\pi)^{3/2}} \sum_{l=0}^{\infty} (2l+1) \frac{P_l(\cos \theta)}{2ki} \left[ \frac{e^{ikr}}{r} - \frac{e^{-ikr} e^{i\pi l}}{r} \right] \quad (\mathbf{r} \rightarrow \infty). \quad (8.64)$$

This is again just a complicated linear superposition of incoming and outgoing spherical waves.

### 8.7.5 Asymptotics of the outgoing scattered wave

First consider an arbitrary function  $g(\theta, \phi)$  defined on the surface of a sphere. The spherical harmonics form an orthogonal basis for the space of such functions and so  $g(\theta, \phi)$  can be expressed as

$$g(\theta, \phi) = \sum_{lm} c_{lm} Y_m^l(\theta, \phi). \quad (8.65)$$

When  $g(\theta, \phi)$  is  $\phi$ -independent this expansion simplifies to

$$g(\theta) = \sum_l c_l Y_0^l(\theta, \phi) = \sum_l (2l+1) c'_l P_l(\cos \theta). \quad (8.66)$$

Now consider the scattered wave

$$\psi_{sc}(\vec{x}) = \frac{1}{(2\pi)^{3/2}} \frac{e^{ikr}}{r} f(\vec{k}', \vec{k}) \quad (8.67)$$

where  $\vec{k} = k\hat{z}$ . Cylindrical symmetry implies that  $f(\vec{k}', \vec{k})$  is independent of  $\phi$  and so it may be expanded in spherical harmonics as

$$f(\vec{k}', \vec{k}) = f(k, \theta) = \sum_{l=0}^{\infty} (2l+1) a_l(k) P_l(\cos \theta). \quad (8.68)$$

We need to find these  $a_l(k)$  expansion coefficients since they determine the scattering amplitude and therefore also the cross section. We will relate them to the phase shifts

in the next section.

To conclude, the asymptotic form of  $\psi_{sc}(\vec{x})$  is given by

$$\psi_{sc}(\vec{x}) = \frac{1}{(2\pi)^{3/2}} \frac{e^{ikr}}{r} f(k, \theta) = \frac{1}{(2\pi)^{3/2}} \sum_{l=0}^{\infty} (2l+1) \frac{P_l(\cos \theta)}{2ki} [2kia_l(k)] \frac{e^{ikr}}{r} \quad (\mathbf{r} \rightarrow \infty). \quad (8.69)$$

Of course, in this case there are only outgoing spherical waves present.

### 8.7.6 Phase shifts and the cross section

We have found asymptotic forms for both sides of  $\psi(\vec{x}) = \psi_{inc}(\vec{x}) + \psi_{sc}(\vec{x})$  as

$$\psi(\vec{x}) = \sum_{l=0}^{\infty} \frac{P_l(\cos \theta)}{2ki} \left[ (A_l - iB_l) \frac{e^{ikr}}{r} - (A_l + iB_l) \frac{e^{-ikr} e^{i\pi l}}{r} \right] \quad (8.70)$$

$$\psi_{inc}(\vec{x}) + \psi_{sc}(\vec{x}) = \frac{1}{(2\pi)^{3/2}} \sum_{l=0}^{\infty} (2l+1) \frac{P_l(\cos \theta)}{2ki} \left[ \underbrace{(1 + 2kia_l(k))}_{\exp[2i\delta_l]} \frac{e^{ikr}}{r} - \frac{e^{-ikr} e^{i\pi l}}{r} \right] \quad (8.71)$$

Clearly the ratios of the amplitudes of the incoming and outgoing spherical waves in (8.70) and (8.71) must be equal, and so

$$1 + 2kia_l(k) = \frac{A_l - iB_l}{A_l + iB_l} = e^{2i\delta_l} \quad (8.72)$$

where  $\delta_l$  is the phase shift introduced in (8.56). This implies that

$$a_l(k) = \frac{\sin(\delta_l) e^{i\delta_l}}{k} \quad \text{and} \quad \tan(\delta_l) = -\frac{B_l}{A_l}. \quad (8.73)$$

Keep in mind that  $\delta_l$  is a function of the energy  $E = \hbar^2 k^2 / (2\mu)$ . Once we have the ratio  $B_l/A_l$  both the phase shift  $\delta_l$  and the expansion coefficient  $a_l(k)$  of the scattering amplitude can be calculated. The differential cross section is then given by

$$\frac{d\sigma}{d\Omega} = |f(k, \theta)|^2 = \sum_{l,l'=0}^{\infty} (2l+1)(2l'+1) a_l^*(k) a_{l'}(k) P_l(\cos \theta) P_{l'}(\cos \theta) \quad (8.74)$$

while the total cross section reads

$$\sigma_{tot} = \int d\Omega |f(k, \theta)|^2 = \sum_l 4\pi (2l+1) |a_l(k)|^2. \quad (8.75)$$

**Exercise:** Prove the optical theorem:  $\sigma_{tot} = \frac{4\pi}{k} \text{Im}[f(\theta = 0)]$ .

### 8.7.7 Calculating the phase shifts

Determining the phase shifts requires knowledge of the asymptotic ( $r \rightarrow \infty$ ) form of the solutions of the radial wave equation. There are two scenarios that we will typically have to deal with.

### Case I: $V(r) = 0$ for $r > R$

This case involves a potential with a finite range  $R$  such that  $V(r) = 0$  for  $r > R$ . This divides space into an interior ( $r < R$ ) and exterior ( $r > R$ ) region. In the exterior the solution to the radial wave equation is just

$$R_{El}^>(r) = A_l j_l(kr) + B_l \eta_l(kr). \quad (8.76)$$

The solution in the interior,  $R_{El}^<(r)$ , of course depends on the details of the potential, and only in special cases can it be found analytically. Suppose for now that  $R_{El}^<(r)$  is known. How do we use this knowledge to determine the ratio  $B_\ell/A_\ell$  needed to calculate the phase shifts? Remember that the radial wave function is always continuous and will also be smooth (i.e. have a continuous derivative) wherever the potential is finite. If  $V(R) < \infty$  then the interior and exterior solutions  $R_{El}^<(r)$  and  $R_{El}^>(r)$  must match up smoothly at  $r = R$ . The matching conditions are

$$\begin{aligned} R_{El}^<(R) &= R_{El}^>(R) = A_l j_l(kR) + B_l \eta_l(kR), \\ \frac{dR_{El}^<(r)}{dr} \Big|_{r=R} &= \frac{dR_{El}^>(r)}{dr} \Big|_{r=R} = k A_l j'_l(kR) + k B_l \eta'_l(kR). \end{aligned} \quad (8.77)$$

*The details of the potential are therefore communicated to the exterior wave function via these matching conditions at  $r = R$ .* From the set of equations above we can solve for the ratio  $B_l/A_l$ , and then calculate the phase shift using

$$\tan(\delta_l) = -\frac{B_l}{A_l}. \quad (8.78)$$

### Case II: $V(r)$ is not zero beyond some radius $R$

In this case we need to solve the radial wave equation in all space and extract the phase shift from the asymptotic form of the solution:

$$R_{El}(r) \propto \left[ e^{2i\delta_l} \frac{e^{ikr}}{r} - \frac{e^{-ikr} e^{i\pi l}}{r} \right]. \quad (8.79)$$

Of course, we only expect this asymptotic form to arise if the potential tends to zero sufficiently quickly.

#### 8.7.8 Application: Hard-sphere scattering I

The hard-sphere potential with range  $R$  is

$$V(r) = \begin{cases} \infty & r \leq R \\ 0 & r > R \end{cases} \quad (8.80)$$

Clearly there is no hope of treating this potential using the Born approximation. A non-perturbative approach such as the partial wave expansion is required. We see that for  $r \leq R$  the radial wave function is simply zero, while at  $r > R$  it will be that of a free particle. We have

$$R_{El}(r) = \begin{cases} R_{El}^<(r) = 0 & r \leq R \\ R_{El}^>(r) = A_l j_l(kr) + B_l \eta_l(kr) & r > R \end{cases} \quad (8.81)$$

This wave function must be continuous at  $r = R$ , but it need not be smooth since the potential is not finite at this point. We therefore have only one matching condition:

$$R_{El}^>(R) = R_{El}^<(R) \implies A_l j_l(kR) + B_l \eta_l(kR) = 0. \quad (8.82)$$

It follows that

$$\tan(\delta_l) = \frac{-B_l}{A_l} = \frac{j_l(kR)}{\eta_l(kR)} \implies \delta_l = \arctan \left[ \frac{j_l(kR)}{\eta_l(kR)} \right]. \quad (8.83)$$

Inserting this result into (8.73) and (8.74) produces an exact expression for the cross section as an infinite series.

### Low energy scattering

The partial wave expansion is particularly useful for treating low energy scattering. As we shall see, at low energies **s-wave scattering**, corresponding to  $l = 0$ , dominates the cross section. In this context “low-energy” should be understood to mean that the wavelength of the incoming plane wave greatly exceeds the range of the scattering potential. Since  $k = 2\pi/\lambda$  this implies that  $kR \ll 1$  is the appropriate low-energy condition. Using the asymptotic forms of  $j_l$  and  $\eta_l$  from appendix D we find that

$$\tan(\delta_l) \approx \frac{-(kR)^{(2l+1)}}{(2l+1)!!(2l-1)!!} \quad \text{when} \quad kR \ll 1. \quad (8.84)$$

The phase shift  $\delta_l$  and expansion coefficient  $a_l(k)$  both clearly decrease very rapidly with increasing  $l$ . It should therefore be sufficient to consider only the *s*-wave contribution to the cross section. Since  $j_0 = \sin(x)/x$  and  $\eta_0 = -\cos(x)/x$  we find that

$$\frac{d\sigma}{d\Omega} = \frac{\sin^2(\delta_0)}{k^2} \approx R^2. \quad (8.85)$$

The cross section is again found to be isotropic at low energies.

### 8.7.9 Application: Hard-sphere scattering II

Let us analyse the results for hard-sphere scattering in a bit more detail, and also compare with the classical results of section 8.2.4. At low energies we found that  $\sigma_{tot} = 4\pi R^2$ , which is four times the classical/geometric result. At these energies  $\lambda \gg R$ , and so any wave packet we construct using these plane waves will be much larger than the target itself. We therefore cannot expect agreement with the classical point particle picture. In particular, the incoming plane wave will be strongly diffracted by the target potential.

In the spirit of the correspondence principle we would expect to find better agreement with the classical result at high energies, where  $kR \gg 1$ . Here we will need to include contributions to the cross section from the higher angular momentum channels, at least up to around  $l \approx kR$ . It is found that as  $kR \rightarrow \infty$  the total cross section tends to  $\sigma_{tot} = 2\pi R^2$ , which is still *twice* the classical result. This is shown in figure 8.6. This is not so easy to understand. In short, this extra  $\pi R^2$  contribution to the cross section is due to the “shadow flux” required to produce the shadow behind the target.

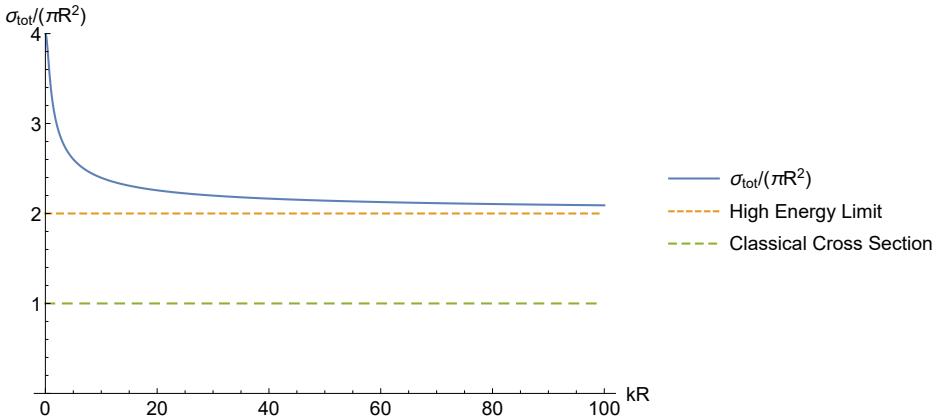


Figure 8.6: The total cross section for hard-sphere scattering, as a function of  $kR$ , together with the high energy and classical result.

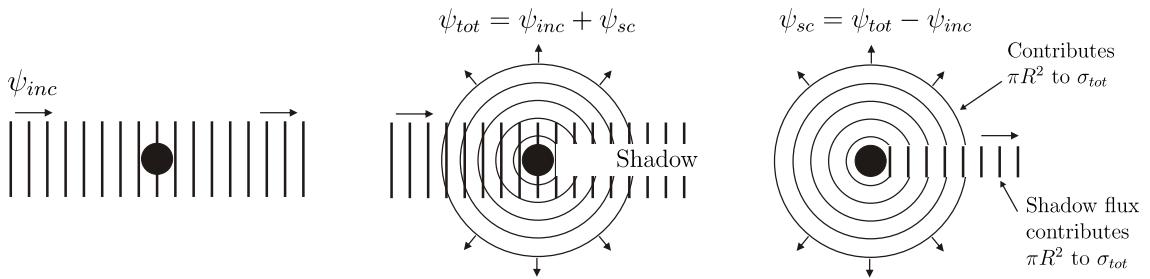


Figure 8.7: Cartoon depicting the various wave functions relevant to high energy hard-sphere scattering.

Consider figure 8.7. Since we are looking at very short wave lengths the diffraction of the wave fronts passing the target is minimal. The left figure shows the wave fronts of the incoming plane wave, which is present both in front and behind the hard-sphere target. The next figure shows the full wave function  $\psi_{tot}$ , with a shadow region behind the target. Subtracting  $\psi_{inc}$  from  $\psi_{tot}$  produces the scattered wave  $\psi_{sc}$  in the final figure, which includes a forward scattering flux required to cancel, via destructive interference,  $\psi_{inc}$  in order to generate the shadow behind the target. This forward scattering flux contributes  $\pi R^2$  to the total cross section, while scattering at non-zero angles contributes the other  $\pi R^2$ . The claims made above can be verified by calculating the scattering amplitude numerically using

$$f(\vec{k}', \vec{k}) = f(k, \theta) = \sum_{l=0}^{\infty} (2l+1) a_l(k) P_l(\cos \theta). \quad (8.86)$$

Figure 8.8 shows the results for the cross section

$$\frac{d\sigma}{d\Omega} = |f(k, \theta)|^2 \quad (8.87)$$

for various values of  $kR$ . As expected, the low energy limit yields an isotropic cross section of  $\frac{d\sigma}{d\Omega} = R^2$ . As we increase  $kR$  there is clearly a trend towards stronger scattering in the forward (small  $\theta$ ) direction. At high energies the cross section at

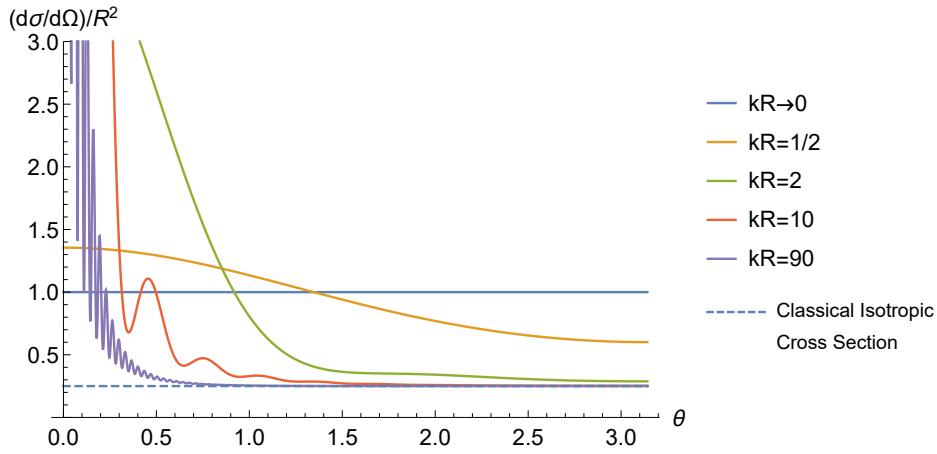


Figure 8.8: The differential cross section for hard-sphere scattering, shown for different energies. On the far right of the plot the curves correspond to increasing energy from top to bottom.

non-zero angles approaches a constant, i.e. again isotropic, value of  $\frac{d\sigma}{d\Omega} = R^2/4$ , which matches the classical result. However, we also observe the formation of a peak in the differential cross section at small  $\theta$ , corresponding to scattering increasingly focussed in the forward direction at high energies. This is the shadow flux contribution. In the high energy limit we can therefore separate the contributions to the total cross section into two parts: scattering at non-zero angles which contributes the classical result  $\pi R^2$ , and scattering at  $\theta = 0$  (or infinitesimally close to it) which contributes another  $\pi R^2$ . Of course, if you were measuring the total cross section experimentally you would probably not count particles scattered at  $\theta = 0$  to have been scattered at all, and so you would conclude that the total cross section is indeed simply  $\pi R^2$ .

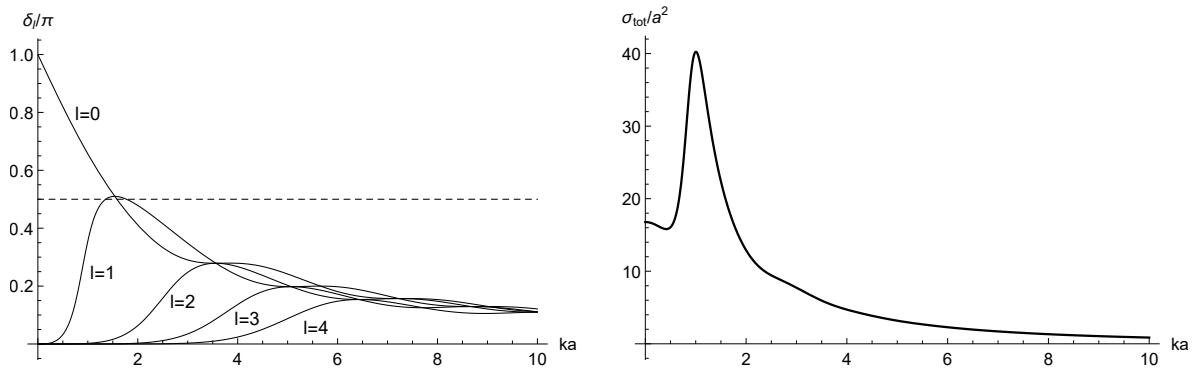


Figure 8.9: Phase shifts and the total cross section as functions of  $ka$ . Here the well depth is  $V_0 = 7.5\hbar^2/(2\mu a^2)$ .

### 8.7.10 Application: Scattering off a finite spherical well

Consider scattering by the attractive spherical well

$$V(r) = \begin{cases} -V_0 & r \leq a \\ 0 & r > a \end{cases} \quad (8.88)$$

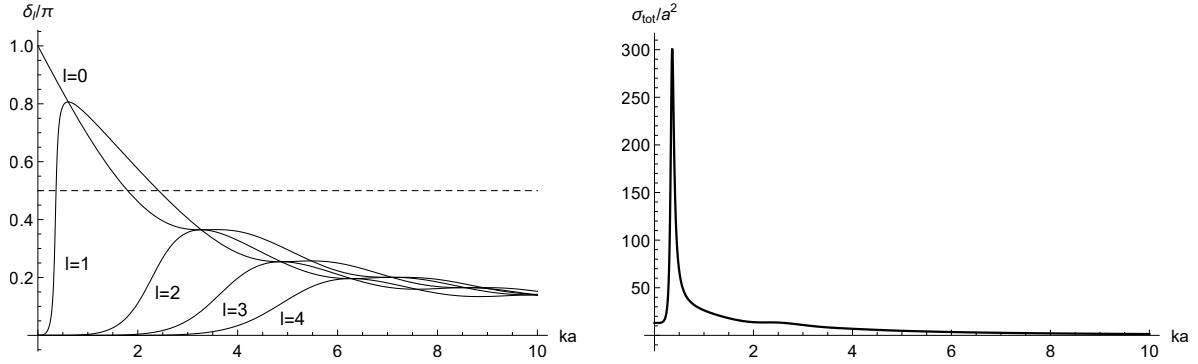


Figure 8.10: Phase shifts and the total cross section as functions of  $ka$ . Here the well depth is  $V_0 = 9.5\hbar^2/(2\mu a^2)$ .

with range  $a$  and depth  $V_0$ . Figures 8.9 and 8.10 show the phase shifts and total cross sections as functions of  $ka$  for two different well depths. Remember that phase shifts differing by integer multiples of  $\pi$  contribute identically to the scattering since  $a_l(k)$  is  $\pi$ -periodic in  $\delta_l$ . A phase shift of  $\delta_l = \pi$  is therefore equivalent to one of  $\delta_l = 0$ . We see that the phase shifts for the  $l \geq 1$  partial waves are very small at low energies ( $ka \ll 1$ ), and these partial waves therefore contribute very little to the cross section. For the  $l = 0$  partial wave the phase shift also starts at zero (or equivalently at  $\pi$ ), but changes *linearly* with increasing momentum. This linear dependence ensures a non-zero  $a_0(k) = \sin(\delta_0)e^{i\delta_0}/k$  even at  $k = 0$ . At low energies, *s*-wave scattering therefore makes the dominant contribution to the cross section. For a particular  $l$ , the largest contribution to the cross section occurs when  $\delta_l = \pi/2$ . In the two figures we indeed see pronounced peaks in the cross section at momenta where  $\delta_1$  passes through  $\pi/2$ . These sharp peaks are known as scattering **resonances**. In figure 8.10 the resonance is much narrower than in 8.9. In short, these resonances correspond to quasi-bound states, where the particle gets temporarily trapped by the potential before tunnelling out and escaping. The lifetime of this quasi-bound state is related to the width of the resonance. The narrower the resonance the longer the particle remains trapped.

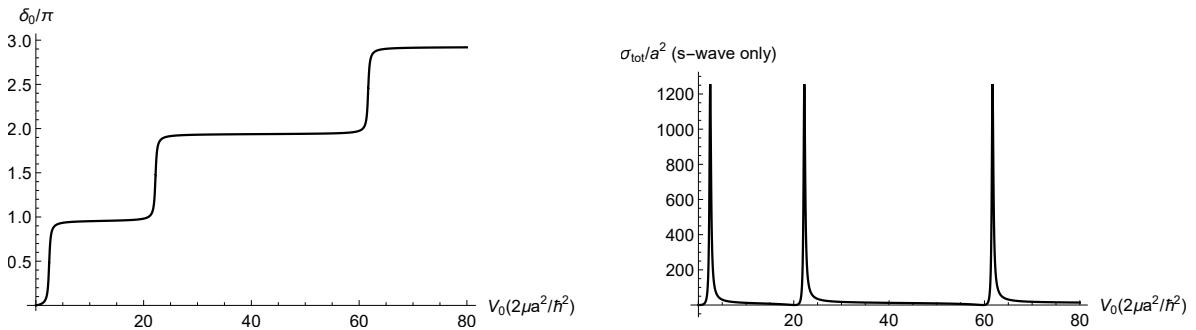


Figure 8.11: The *s*-wave phase shift and cross section as functions of the well depth for low energy scattering at  $ka = 0.1$ .

Another type of resonance is illustrated in figure 8.11. Here we consider low energy scattering with  $ka = 0.1$ , and so only the *s*-wave contribution is included. We monitor the phase shift  $\delta_0$  and the total cross section as the well depth  $V_0$  increases. We

observe very sharp peaks in the total cross section at particular well depths, which again correspond to  $\delta_0$  passing through a half-integer multiple of  $\pi$ . These resonances coincide with the formation of new bound states in the well. As the well gets deeper these states form at the top of the well, just below zero energy, and then move downwards. Although the projectiles enter at low positive energies, the scattering is clearly very sensitive to the proximity of a bound state near the top of the well. By keeping track of the phase shift we can therefore determine the number of bound states the potential accommodates. This general result is known as **Levinson's theorem**.

### 8.7.11 Why $s$ -wave scattering dominates at low energies

It has been noted several times that at low energies the scattering cross section is dominated by the  $l = 0$   $s$ -wave contribution. In this section we consider some simple classical, semi-classical and quantum mechanical arguments for why this is the case.

#### Classical

Here the argument is simple. If the projectile has a low energy  $E = p^2/(2m)$ , but a high angular momentum  $L = pb$ , the impact parameter  $b$  must be large. If  $b$  is larger than the range  $R$  of the potential then no scattering will occur. More concretely, since  $L = pb$  we need  $b = L/p < R$ , or  $L < pR$ , for scattering to occur. At low energies, i.e. low  $p$ , this requires a sufficiently low  $L$  as well.

#### Semi-classical

Here we use the quantum mechanical relations  $p = \hbar k$  and  $L^2 = l(l+1)\hbar^2$ . Approximating  $L \approx l\hbar$  results in the classical condition above becoming  $b = l/k < R$ , or  $l < kR$ . At low energies  $kR \ll 1$ , and so to observe scattering here we need  $l \ll 1$ , i.e.  $l = 0$ .

#### Quantum mechanical

Here we need to investigate the wave functions themselves. First consider a free particle with energy  $E = \hbar^2 k^2/(2m)$  and angular momentum  $l$ . Its radial wave function is then  $j_l(kr)$ . Figure 8.12 shows the function  $j_l(x)$  for various values of  $l$ . We see that for  $l > 0$  the function  $j_l(x)$  rapidly tends to zero as  $x$  decreases below  $l$ . This shows that the particle is most likely to be found at radii  $r$  beyond  $l/k$ , in line with the previous discussion. Now imagine switching on a potential with a finite range  $R$ . What effect will this have on the radial wave function? If  $R < l/k$  then the potential is non-zero only in the region where  $j_l(kr)$  almost vanishes completely. The addition of the potential will therefore have virtually no effect on the radial wave function, and therefore induce no phase shift or scattering. It is only  $j_0(kr)$ , which is non-zero even when  $kr \ll 1$ , which will be modified and undergo a phase shift due to the presence of the potential.

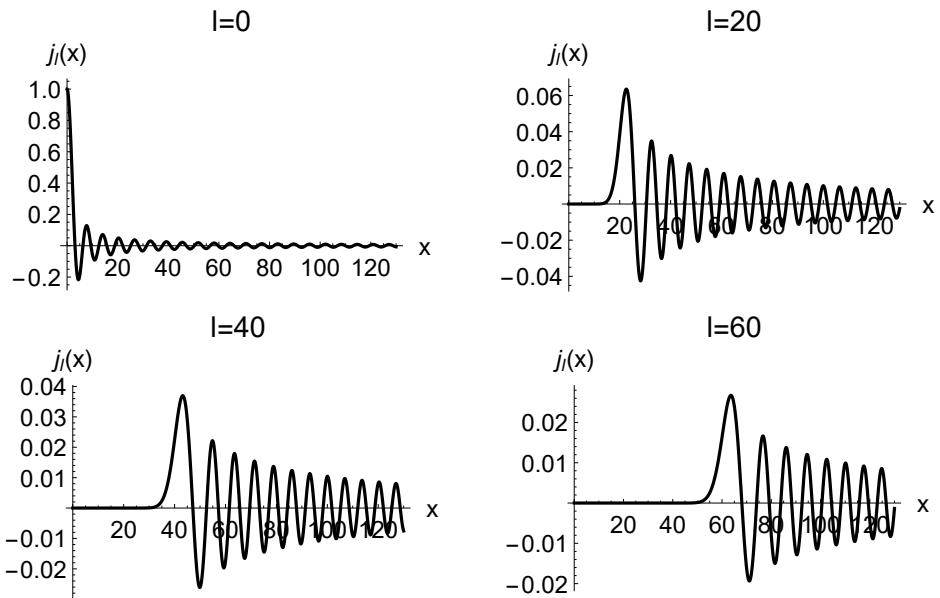


Figure 8.12: The spherical Bessel function for various values of  $l$ .

# Appendix A: Dimensional Analysis Basics

## Introduction:

A **dimension** is a type of physical quantity, while a **unit** is a means of assigning a numerical value to a physical quantity. The table lists six of the primary dimensions and some of the secondary dimensions derived from them. We denote the dimension of a quantity  $r$  by  $[r]$ . **Dimensionless** quantities are assigned a dimension of 1. The ratio of two quantities with the same dimension is dimensionless. In particular, angles are dimensionless, despite having units of radians or degrees.

For an expression to make logical sense it must be dimensionally consistent: we cannot add, subtract, equate or compare quantities with different dimensions. Of course, numerical correctness requires unit consistency as well: we cannot add 12 cm to 2 m to produce 14 in mystery units. Note that the argument of any trigonometric, exponential or logarithmic function *must* be dimensionless. Expressions such as  $\sin(2 \text{ meters})$  or  $e^{3 \text{ seconds}}$  are *meaningless*. This is obvious from the Taylor series definitions of these functions which involve the sum of *different* powers of their arguments.

In practice, dimensional analysis provides us with:

- A consistency check on our results, and a way of detecting simple errors easily.
- A means of identifying the relevant scales in a problem in terms of the fundamental parameters that appear in it.
- A means of determining how a quantity should depend on the dimensionful parameters of a problem.

## Example:

Consider a puck with mass  $m_p$  moving on a flat frictionless table. The puck is moving at a constant speed  $v$  in a circle around a point to which it is attached by a light piece of string. The length of the string, and the radius of the puck's circular path, is  $r$ . *What would happen to the tension in the string if we double  $v$ ,  $m_p$  and  $r$ ?*

Consider the dimensions of the relevant quantities:

$$[r] = L \quad [m_d] = M \quad [v] = L/t \quad [T] = ML/t^2$$

We need to combine  $r$ ,  $m_d$  and  $v$  into an expression for  $T$ . Dimensional consistency requires that the expression has the form

$$T = k v^a m_d^b r^c$$

with  $k$ ,  $a$ ,  $b$  and  $c$  dimensionless constants. Equating the dimensions on the two sides leads to

$$[T] = [v]^a [m_d]^b [r]^c \longrightarrow \frac{ML}{t^2} = \left(\frac{L}{t}\right)^a M^b L^c.$$

To match the dimensions we must set  $a = 2$ ,  $b = 1$  and  $c = -1$ . The tension is therefore given by  $T = kv^2m_d/r$ . Of course, we cannot obtain any information about

the dimensionless constant  $k$  from this approach. Doubling  $v$ ,  $m_p$  and  $r$  will therefore result in the tension increasing by a factor of four.

### Final Remarks:

Dimensional analysis is a very useful and versatile skill. It is highly recommended that you integrate it into your thinking. For further discussion and some interesting examples, see the paper “*Dimensional analysis as the other language of physics*” by RW Robinett.

Primary Dimensions	Symbol	SI Unit
Mass	M	kg
Length	L	m
Time	t	s
Electric Current	I	A
Temperature	T	K

Secondary Dimensions	i.t.o. PD	SI Unit
Velocity and Speed	$\frac{L}{t}$	$\frac{m}{s}$
Acceleration	$\frac{L}{t^2}$	$\frac{m}{s^2}$
Force	$\frac{ML}{t^2}$	$\frac{kg \cdot m}{s^2} = N$
Energy and Work	$\frac{ML^2}{t^2}$	$\frac{kg \cdot m^2}{s^2} = J$
Electric Charge	It	$A \cdot s = C$
Electric Field Strength	$\frac{ML}{It^3}$	$\frac{kg \cdot m}{A \cdot s^3} = \frac{N}{C}$
Angle	1	Radians
:	:	:

# Appendix B: Time-Independent Perturbation Theory

## B.1 Introduction

Consider the **perturbed Hamiltonian**

$$\hat{H} = \hat{H}_0 + \lambda \hat{H}_1 \quad (\text{B.1})$$

where  $\hat{H}_0$  is the **unperturbed Hamiltonian** and  $\hat{H}_1$  is the **perturbation**. The parameter  $\lambda$  determines the strength of the perturbation. In some cases  $\lambda$  is a real physical quantity, while in others it serves as a book-keeping device for keeping track of terms in a series expansion. We always assume that the unperturbed Hamiltonian  $\hat{H}_0$  is completely solvable, i.e. that its eigenstates and eigenvalues are known exactly. The goal of time-independent perturbation theory is to derive approximate results for the eigenstates and eigenvalues of  $\hat{H}$  as power series in  $\lambda$ . The approach we will follow depends on whether the unperturbed Hamiltonian  $\hat{H}_0$  is degenerate.

## B.2 Non-Degenerate Perturbation Theory

We denote the eigenstates and eigenvalues of  $\hat{H}_0$  by

$$\hat{H}_0 |n\rangle = \epsilon_n |n\rangle \quad (\text{B.2})$$

where  $n$  is a discrete label. Since  $\hat{H}_0$  is non-degenerate the eigenvalues are distinct; so if  $n \neq m$  then  $\epsilon_n \neq \epsilon_m$ . We are interested in the eigenstates and eigenvalues of the perturbed Hamiltonian:

$$\hat{H} |\phi_n\rangle = (\hat{H}_0 + \lambda \hat{H}_1) |\phi_n\rangle = E_n |\phi_n\rangle. \quad (\text{B.3})$$

Both  $E_n = E_n(\lambda)$  and  $|\phi_n\rangle = |\phi_n, \lambda\rangle$  are functions of  $\lambda$ , and we will assume that these vary sufficiently smoothly with  $\lambda$  to permit power series expansions. This allows us to track how the eigenstates and eigenvalues of  $\hat{H}$  change as the perturbation is turned on, and so we can use the same set of  $n$ -labels for the eigenstates and eigenvalues of both  $\hat{H}_0$  and  $\hat{H}$ . We are interested in the power series

$$E_n = E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \dots \quad (\text{B.4})$$

$$|\phi_n\rangle = |\phi_n^{(0)}\rangle + \lambda |\phi_n^{(1)}\rangle + \lambda^2 |\phi_n^{(2)}\rangle + \dots \quad (\text{B.5})$$

and now need to find expressions for the various terms in these expansions. Since  $E_n$  should reduce to the unperturbed energy  $\epsilon_n$  as  $\lambda \rightarrow 0$  we can identify

$$E_n^{(0)} = \epsilon_n. \quad (\text{B.6})$$

Also, since  $\hat{H}_0$  is non-degenerate, we have

$$|\phi_n^{(0)}\rangle = |n\rangle \quad (\text{B.7})$$

since  $|n\rangle$  is the only eigenstate of  $\hat{H}_0$  with energy  $\epsilon_n$ . If  $\hat{H}_0$  was degenerate this argument would not work! It is a standard derivation to show that the linear order correction to the energy is given by

$$E_n^{(1)} = \langle n | \hat{H}_1 | n \rangle, \quad (\text{B.8})$$

which is simply the expectation value of the perturbation with respect to the unperturbed eigenstate  $|n\rangle$ . The quadratic order correction is

$$E_n^{(2)} = \sum_{m \neq n} \frac{|\langle n | \hat{H}_1 | m \rangle|^2}{\epsilon_n - \epsilon_m}. \quad (\text{B.9})$$

The linear order correction to  $|\phi_n\rangle$  is given by

$$|\phi_n^{(1)}\rangle = \sum_{m \neq n} \frac{\langle m | \hat{H}_1 | n \rangle}{\epsilon_n - \epsilon_m} |m\rangle. \quad (\text{B.10})$$

### B.3 Degenerate Perturbation Theory

If the unperturbed Hamiltonian is degenerate then the results stated above no longer apply. This is obvious from the fact that denominators in the sums for  $E_n^{(2)}$  and  $|\phi_n^{(1)}\rangle$  will vanish when two of the unperturbed energies are equal. To treat this case we need to return to the start of our derivation. We first introduce a more appropriate labelling of the unperturbed eigenstates of  $\hat{H}_0$  as

$$\hat{H}_0 |n, r\rangle = \epsilon_n |n, r\rangle \quad (\text{B.11})$$

where, for a fixed  $n$ ,  $r = 1, 2, \dots, d_n$ . Here  $n$  labels the energy *levels* of  $\hat{H}_0$ . Level  $n$  then has energy  $\epsilon_n$  and is  $d_n$ -fold degenerate. The label  $r$  distinguishes states within a single level. The eigenspace of  $\hat{H}_0$  corresponding to  $\epsilon_n$  is

$$\mathcal{E}_n = \text{span} \{ |n, r\rangle : r = 1, 2, \dots, d_n \}. \quad (\text{B.12})$$

The two power series now read

$$E_{n,r} = E_{n,r}^{(0)} + \lambda E_{n,r}^{(1)} + \lambda^2 E_{n,r}^{(2)} + \dots \quad (\text{B.13})$$

$$|\phi_{n,r}\rangle = |\phi_{n,r}^{(0)}\rangle + \lambda |\phi_{n,r}^{(1)}\rangle + \lambda^2 |\phi_{n,r}^{(2)}\rangle + \dots \quad (\text{B.14})$$

Again  $E_{n,r}$  should reduce to  $\epsilon_n$  as  $\lambda \rightarrow 0$ , and so  $E_{n,r}^{(0)} = \epsilon_n$ . Similarly,  $|\phi_{n,r}^{(0)}\rangle$  should be an eigenstate of  $\hat{H}_0$  with eigenvalue  $\epsilon_0$ , i.e. an element of the eigenspace  $\mathcal{E}_n$ , though not necessarily  $|n, r\rangle$  itself. In general we can write

$$|\phi_{n,r}^{(0)}\rangle = \sum_{r'=1}^{d_n} c_{r,r'} |n, r'\rangle. \quad (\text{B.15})$$

The set of states  $\{|\phi_{n,r}^{(0)}\rangle : r = 1, 2, \dots, d_n\}$  define a new basis for the  $\mathcal{E}_n$  eigenspace, one which is usually different from the original  $\{|n, r\rangle : r = 1, 2, \dots, d_n\}$  basis. It can be shown that this new basis is the one that diagonalises the perturbation  $\hat{H}_1$  in the  $\mathcal{E}_n$  eigenspace, i.e. for which the  $d_n \times d_n$  matrix with elements  $\langle \phi_{n,r}^{(0)} | \hat{H}_1 | \phi_{n,r'}^{(0)} \rangle$  is

diagonal. To find these basis states we must therefore determine the eigenvectors of the matrix  $W$  with elements  $W_{r,r'} = \langle n, r | \hat{H}_1 | n, r' \rangle$ . These eigenvectors are the desired  $|\phi_{n,r}^{(0)}\rangle$  states expressed in the original unperturbed basis, i.e. column vectors of the form  $\vec{c}_r = [c_{r,1}, c_{r,2}, \dots, c_{r,d_n}]^T$ . The eigenvector equation is therefore

$$W \vec{c}_r = E_{n,r}^{(1)} \vec{c}_r \quad \text{for } r = 1, \dots, d_n \quad (\text{B.16})$$

where the eigenvalues are the linear order corrections to the energy. Expressions for the higher-order corrections can also be derived, but these are quite complicated. The results we have obtained for  $|\phi_{n,r}^{(0)}\rangle$  and  $E_{n,r}^{(1)}$  already cover a wide range of applications.

To linear order the perturbed energies are  $E_{n,r} = \epsilon_n + \lambda E_{n,r}^{(1)}$ , where the linear order correction typically produces a splitting of the unperturbed energy levels. This suggests that the presence of the perturbation will typically reduce, or possibly completely eliminate, the degeneracy present in the unperturbed Hamiltonian.

# Appendix C: Elementary Systems

## C.1 The harmonic oscillator

The quantum Hamiltonian is

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

We will define the annihilation operator  $\hat{a}$  and creation operator  $\hat{a}^\dagger$  as

$$\hat{a} = \alpha\hat{x} + i\beta\hat{p} \quad \text{and} \quad \hat{a}^\dagger = \alpha\hat{x} - i\beta\hat{p} \quad (\text{C.2})$$

where

$$\alpha = \sqrt{\frac{m\omega}{2\hbar}} \quad \text{and} \quad \beta = \frac{1}{\sqrt{2\hbar m\omega}}. \quad (\text{C.3})$$

The Hamiltonian can then be expressed as

$$\hat{H} = \hbar\omega \left( \hat{a}^\dagger\hat{a} + \frac{1}{2} \right). \quad (\text{C.4})$$

The eigenstates of  $\hat{H}$  are therefore those of the counting (or number) operator  $\hat{a}^\dagger\hat{a}$ . These are denoted by  $|n\rangle$  with  $n = 0, 1, 2, \dots$  and satisfy  $\hat{a}^\dagger\hat{a}|n\rangle = n|n\rangle$ . The corresponding eigenvalues of  $\hat{H}$  are

$$E_n = \hbar\omega(n + 1/2). \quad (\text{C.5})$$

The annihilation and creation operators satisfy

$$[\hat{a}, \hat{a}^\dagger] = 1 \quad (\text{C.6})$$

and act as ladder operators between the different eigenstates:

$$\hat{a}^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle \quad (\text{C.7})$$

$$\hat{a}|n\rangle = \sqrt{n}|n-1\rangle. \quad (\text{C.8})$$

It follows that

$$(\hat{a}^\dagger)^n|0\rangle = \sqrt{n!}|n\rangle. \quad (\text{C.9})$$

In particular, note that  $\hat{a}|0\rangle = 0|0\rangle = 0$  and so the ground state  $|0\rangle$  is an eigenstate of  $\hat{a}$  with eigenvalue zero.

## C.2 The infinite square well

The quantum Hamiltonian is

$$\hat{H} = \frac{\hat{p}^2}{2m} + V(\hat{x}) \quad (\text{C.10})$$

where  $V(x)$  is zero for  $0 \leq x \leq L$  and infinitely large elsewhere. The eigenstates are denoted by  $|n\rangle$  for  $n = 1, 2, 3, \dots$  with corresponding energy eigenvalues

$$E_n = \frac{n^2\pi^2\hbar^2}{2mL^2}. \quad (\text{C.11})$$

The wave functions of the eigenstates are

$$\phi_n(x) = \langle x|n\rangle = \sqrt{\frac{2}{L}} \sin\left(\frac{n\pi x}{L}\right). \quad (\text{C.12})$$

### C.3 Spin-1/2 summary

The eigenstates and eigenvalues of  $\hat{S}_{x,y,z}$  are

$$\hat{S}_x|x,+\rangle = \frac{\hbar}{2}|x,+\rangle \quad \hat{S}_x|x,-\rangle = -\frac{\hbar}{2}|x,-\rangle$$

$$\hat{S}_y|y,+\rangle = \frac{\hbar}{2}|y,+\rangle \quad \hat{S}_y|y,-\rangle = -\frac{\hbar}{2}|y,-\rangle$$

$$\hat{S}_z|z,+\rangle = \frac{\hbar}{2}|z,+\rangle \quad \hat{S}_z|z,-\rangle = -\frac{\hbar}{2}|z,-\rangle$$

We abbreviate  $|\pm\rangle \equiv |z,\pm\rangle$ . In this basis the matrix representations of  $\hat{S}_{x,y,z}$  are

$$\hat{S}_x = \frac{\hbar}{2} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = \frac{\hbar}{2} \hat{\sigma}_x = \frac{\hbar}{2} (|+\rangle\langle-| + |-\rangle\langle+|)$$

$$\hat{S}_y = \frac{\hbar}{2} \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} = \frac{\hbar}{2} \hat{\sigma}_y = \frac{\hbar}{2} (-i|+\rangle\langle-| + i|-\rangle\langle+|)$$

$$\hat{S}_z = \frac{\hbar}{2} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = \frac{\hbar}{2} \hat{\sigma}_z = \frac{\hbar}{2} (|+\rangle\langle+| - |-\rangle\langle-|)$$

with eigenstates

$$|x,+\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \frac{1}{\sqrt{2}}(|+\rangle + |-\rangle) \quad |x,-\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \frac{1}{\sqrt{2}}(|+\rangle - |-\rangle)$$

$$|y,+\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ i \end{bmatrix} = \frac{1}{\sqrt{2}}(|+\rangle + i|-\rangle) \quad |y,-\rangle = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -i \end{bmatrix} = \frac{1}{\sqrt{2}}(|+\rangle - i|-\rangle)$$

$$|z,+\rangle = \begin{bmatrix} 1 \\ 0 \end{bmatrix} = |+\rangle \quad |z,-\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix} = |-\rangle$$

**Commutation relations:**  $[\hat{S}_i, \hat{S}_j] = i\hbar\epsilon_{ijk}\hat{S}_k$

# Appendix D: Special Functions

**Legendre polynomials:**

$$P_l(x) = \frac{1}{2^l l!} \left( \frac{d}{dx} \right)^l (x^2 - 1)^l \quad (\text{D.1})$$

**Associated Legendre functions:**

$$P_l^m(x) = (1 - x^2)^{|m|/2} \left( \frac{d}{dx} \right)^{|m|} P_l(x) \quad (\text{D.2})$$

Note that  $P_l^m(1) = \delta_{m0}$ .

**Orthogonality of Legendre polynomials:**

$$\int_{-1}^1 dx P_l(x) P_{l'}(x) = \int_0^\pi d\theta \sin \theta P_l(\cos \theta) P_{l'}(\cos \theta) = \frac{2}{2l+1} \delta_{ll'} \quad (\text{D.3})$$

**Spherical harmonics: (with  $m \geq 0$ )**

$$Y_m^l(\theta, \phi) = (-1)^m \left[ \frac{(2l+1)}{4\pi} \frac{(l-m)!}{(l+m)!} \right]^{1/2} P_l^m(\cos \theta) e^{im\phi} \quad (\text{D.4})$$

$$Y_{-m}^l(\theta, \phi) = (-1)^m [Y_m^l(\theta, \phi)]^* \quad (\text{D.5})$$

**Spherical Bessel function of order  $l$ :**

$$j_l(x) = (-x)^l \left( \frac{1}{x} \frac{d}{dx} \right)^l \left( \frac{\sin x}{x} \right) \quad (\text{D.6})$$

**Spherical Neumann function of order  $l$ :**

$$\eta_l(x) = -(-x)^l \left( \frac{1}{x} \frac{d}{dx} \right)^l \left( \frac{\cos x}{x} \right) \quad (\text{D.7})$$

**Special cases:**

$$j_0(x) = \frac{\sin x}{x} \quad j_1(x) = \frac{\sin x}{x^2} - \frac{\cos x}{x} \quad (\text{D.8})$$

$$\eta_0(x) = -\frac{\cos x}{x} \quad \eta_1(x) = -\frac{\cos x}{x^2} - \frac{\sin x}{x} \quad (\text{D.9})$$

**Asymptotic forms for  $x \ll 1$ :**

$$j_l(x) \approx \frac{x^l}{(2l+1)!!} \quad \eta_l(x) \approx -\frac{(2l-1)!!}{x^{l+1}} \quad (\text{D.10})$$

**Asymptotic forms for  $x \rightarrow \infty$ :**

$$j_l(x) \approx \frac{\sin(x - l\pi/2)}{x} = \frac{e^{i(x-l\pi/2)} - e^{-i(x-l\pi/2)}}{2ix} \quad (\text{D.11})$$

$$\eta_l(x) \approx -\frac{\cos(x - l\pi/2)}{x} = -\frac{e^{i(x-l\pi/2)} + e^{-i(x-l\pi/2)}}{2x} \quad (\text{D.12})$$

# Appendix E: Physics 254 Notes

## WAVE MECHANICS SUMMARY

---

### Wave Functions

---

**Inner Product:**

$$(\psi, \phi) = \int_{-\infty}^{+\infty} dx \psi^*(x)\phi(x) \quad (\text{E.1})$$

**Normalisation:**

$$(\psi, \psi) = \int_{-\infty}^{+\infty} dx \psi^*(x)\psi(x) = 1 \quad (\text{E.2})$$

**Orthogonality:**

$$(\psi, \phi) = 0 \quad (\text{E.3})$$

**Transforming between Position and Momentum Space:**

$$\tilde{\psi}(p) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} dx \psi(x)e^{-ipx/\hbar} \quad \psi(x) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} dp \tilde{\psi}(p)e^{ipx/\hbar} \quad (\text{E.4})$$

---

### Operators

---

**Linearity:**

$$\hat{A}(\alpha\phi(x) + \beta\psi(x)) = \alpha\hat{A}\phi(x) + \beta\hat{A}\psi(x) \quad (\text{E.5})$$

**The Commutator:**

$$[\hat{A}, \hat{B}] = \hat{A}\hat{B} - \hat{B}\hat{A} \quad (\text{E.6})$$

**The Adjoint:**

$$(\psi, \hat{A}\phi) = (\hat{A}^\dagger\psi, \phi) \quad \text{and} \quad (\hat{A}\psi, \phi) = (\psi, \hat{A}^\dagger\phi) \quad (\text{E.7})$$

**Hermiticity:**

$$\hat{A} = \hat{A}^\dagger \quad (\text{E.8})$$

**Eigenstates and Eigenvalues:**

$$\hat{A}\phi(x) = \lambda\phi(x) \quad (\text{E.9})$$

---

### Observables

---

**Position and Momentum:**

$$\hat{x}\psi(x) = x\psi(x) \quad \hat{p}\psi(x) = \frac{\hbar}{i} \frac{d}{dx}\psi(x) \quad [\hat{x}, \hat{p}] = i\hbar \quad (\text{E.10})$$

**Hamiltonian:**

$$\hat{H} = \hat{K} + \hat{V} = \frac{\hat{p}^2}{2m} + V(\hat{x}) \quad \hat{H}\psi(x) = \left[ -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + V(x) \right] \psi(x) \quad (\text{E.11})$$

**Momentum Eigenstates - Plane waves:**

$$\hat{p}\phi_p(x) = p\phi_p(x) \quad \phi_p(x) = \frac{1}{\sqrt{2\pi\hbar}} e^{ipx/\hbar} \quad (\text{E.12})$$

**Energy Eigenstates - Time-Independent Schrödinger Equation (TISE)**

$$\hat{H}\phi(x) = E\phi(x) \quad -\frac{\hbar^2}{2m} \frac{d^2\phi(x)}{dx^2} + V(x)\phi(x) = E\phi(x) \quad (\text{E.13})$$

---

## Measurements

---

**Position Measurements:**

$$\rho(x) = |\psi(x)|^2 \quad P(x \in [a, b]) = \int_a^b dx |\psi(x)|^2 \quad (\text{E.14})$$

**Momentum Measurements:**

$$\rho(p) = |\tilde{\psi}(p)|^2 \quad P(p \in [a, b]) = \int_a^b dp |\tilde{\psi}(p)|^2 \quad (\text{E.15})$$

**General Observable:**

$$\hat{\mathcal{O}}\phi_\lambda(x) = \lambda\phi_\lambda(x) \quad (\text{E.16})$$

$$\rho(\lambda) \text{ or } P(\lambda) = |(\phi_\lambda, \psi)|^2 = \left| \int_{-\infty}^{+\infty} dx \phi_\lambda^*(x) \psi(x) \right|^2 \quad (\text{E.17})$$

**Expectation Values:**

$$\langle \mathcal{O} \rangle = (\psi, \hat{\mathcal{O}}\psi) = \int_{-\infty}^{+\infty} dx \psi^*(x) \hat{\mathcal{O}}\psi(x) \quad (\text{E.18})$$

$$\langle x \rangle = \int_{-\infty}^{+\infty} dx \psi^*(x) x \psi(x) \quad \langle p \rangle = \int_{-\infty}^{+\infty} dx \psi^*(x) \frac{\hbar}{i} \frac{d}{dx} \psi(x) \quad (\text{E.19})$$

**Uncertainties:**

$$\sigma_{\mathcal{O}} = \sqrt{\langle (\mathcal{O} - \langle \mathcal{O} \rangle)^2 \rangle} = \sqrt{\langle \mathcal{O}^2 \rangle - \langle \mathcal{O} \rangle^2} \quad (\text{E.20})$$

**Position-Momentum Uncertainty Principle:**

$$\sigma_x \sigma_p \geq \hbar/2 \quad (\text{E.21})$$

---

## Dynamics

---

**Time-Dependent Schrödinger Equation (TDSE):**

$$i\hbar \frac{\partial}{\partial t} \psi(x, t) = \hat{H}\psi(x, t) \quad i\hbar \frac{\partial \psi(x, t)}{\partial t} = -\frac{\hbar^2}{2m} \frac{\partial^2 \psi(x, t)}{\partial x^2} + V(x)\psi(x, t) \quad (\text{E.22})$$

**General Solution to the TDSE:**

$$\psi(x, t) = \sum_n c_n e^{-iE_n t/\hbar} \phi_n(x) \quad \text{where} \quad \hat{H}\phi_n(x) = E_n \phi_n(x) \quad (\text{E.23})$$

$$c_n = (\phi_n, \psi|_{t=0}) = \int_{-\infty}^{+\infty} dx \phi_n^*(x) \psi(x, 0) \quad (\text{E.24})$$

**Free Particle Wave Packet:**

$$\psi(x, t) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{+\infty} dp \tilde{\psi}(p, t) e^{ipx/\hbar} \quad (\text{E.25})$$

$$\tilde{\psi}(p, t) = \tilde{\psi}(p, 0) e^{-iE_p t/\hbar} \quad \text{with} \quad E_p = \frac{p^2}{2m} \quad (\text{E.26})$$

# BASIC QUANTUM MECHANICS VIA SPIN- $1/2$

---

## E.1 Introduction

Consider a single electron. In addition to its spatial degrees of freedom, i.e. its position and momentum, the electron also possesses an *internal* degree of freedom called spin. Spin is a type of *intrinsic angular momentum*, which is not associated with motion or rotation in space like the orbital angular momentum  $\vec{L} = \vec{r} \times \vec{p}$ . Instead, spin is a purely quantum mechanical degree of freedom which has no analogue in classical mechanics. Despite its name, one should therefore not think of spin as the result of the electron “spinning” around some axis like a tiny sphere. Spin provides the simplest setting for getting to grips with the surprising and unfamiliar laws of quantum theory. Here we focus exclusively on the electron’s spin, and leave the description of its position and momentum for later.

Like its orbital angular momentum, the electron’s spin has three components,  $S_x$ ,  $S_y$ , and  $S_z$ . It is tempting to think of the spin as a vector  $\vec{S} = (S_x, S_y, S_z)$ , but this can be misleading since quantum mechanics does not allow all three, or even any two, of the components to have well-defined values simultaneously.

## E.2 The Basic Postulates

### E.2.1 Postulate 1

*The quantum state of the electron’s spin is represented by a normalised state vector  $\psi \in \mathcal{H} = \mathbb{C}^2$ , i.e. a two component column vector with complex entries and a norm of one.*

We can identify two states in  $\mathcal{H}$  which will play an important role. These are

$$\chi_+ = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{and} \quad \chi_- = \begin{bmatrix} 0 \\ 1 \end{bmatrix}. \quad (\text{E.27})$$

We call  $\chi_+$  the *spin-up* state. In this state the  $S_z$  component of the spin equals  $+\hbar/2$  and so it points, roughly speaking, “upwards” along the positive  $z$ -direction. Here

$$\hbar = 1.0546 \times 10^{-34} \text{ J} \cdot \text{s} \quad (\text{E.28})$$

is the reduced Planck constant. We need to be careful with the wording, since we are not making, and *cannot* make, any statements about the values of the other two components. Similarly, we call  $\chi_-$  the *spin-down* state. In this state  $S_z$  equals  $-\hbar/2$ . You can check that  $\hbar$  indeed has the dimension of angular momentum. We have now encountered the first surprising feature of quantum mechanics, namely that the values of physical quantities can be *quantised*, i.e. that they can only take certain discrete values. Classically, the  $z$ -component of orbital angular momentum  $\vec{L}$  can take any value between  $-L$  and  $+L$ , depending on how  $\vec{L}$  is orientated in space. In contrast, there are only two possible values for any spin component, namely  $+\hbar/2$  or  $-\hbar/2$ . This fact is illustrated experimentally in the Stern-Gerlach experiment.

A general state  $\psi \in \mathcal{H}$  can be expressed as a linear combination of  $\chi_+$  and  $\chi_-$  as

$$\psi = \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \alpha\chi_+ + \beta\chi_- \quad (\text{E.29})$$

where  $\alpha$  and  $\beta$  are complex numbers. Therefore  $\psi$  is a *linear superposition* of the spin-up and spin-down states.

**Remark 1:** The mathematical description of spin-1/2 is identical to that of a qubit, the basic component of a quantum computer. Indeed, it is via the spin of particles that qubits are often realised experimentally. You can think of the spin-up and spin-down states as representing a qubit with a value of 1 or 0 respectively. In a classical computer these are the only two options for the state of a bit. In a quantum computer arbitrary superpositions of the 1 and 0 states are possible.

**Remark 2:** Remember that the inner product for  $\mathbb{C}^2$  is the regular complex dot product. If  $u = [u_1, u_2]^T$  and  $v = [v_1, v_2]^T$  then  $(u, v) = u \cdot v = u_1^*v_1 + u_2^*v_2$ . We can also write this as  $(u, v) = u^\dagger v$  where  $u^\dagger = [u_1^*, u_2^*]$  is the *conjugate transpose* of  $u$ . The requirement that  $\psi = [\alpha, \beta]^T$  is normalised therefore amounts to  $|\psi|^2 = (\psi, \psi) = |\alpha|^2 + |\beta|^2 = 1$ .

### E.2.2 Postulate 2a

If the electron is in the state  $\psi = \alpha\chi_+ + \beta\chi_-$  and we measure  $S_z$  we will obtain  $+\hbar/2$  with a probability of  $|(\chi_+, \psi)|^2 = |\alpha|^2$  or  $-\hbar/2$  with a probability of  $|(\chi_-, \psi)|^2 = |\beta|^2$ .

There is a lot to digest in this statement. *First of all, it reveals the fundamentally non-deterministic nature of quantum mechanics, in that the same initial conditions do not necessarily lead to the same experimental results.* Two identical experiments, performed on particles prepared in exactly the same state are not guaranteed to produce the same outcome. This is in stark contrast to classical mechanics where, once the initial state of the system is specified, its future behaviour can in principle be predicted with certainty. Even having full knowledge of the state  $\psi$  does not enable us to predict the outcome of a single measurement. At best we can calculate the *probability* of obtaining a certain outcome. It is crucial to understand that this is not a result of a lack of knowledge on our part. The state vector  $\psi$  contains *all there is to know* about the system. In particular, the + sign in the superposition  $\psi = \alpha\chi_+ + \beta\chi_-$  is not an “or” statement. We are *not* saying that, prior to the measurement, the particle was in the spin-up state with a probability of  $|\alpha|^2$  or in the spin-down state with a probability of  $|\beta|^2$ . The particle was in neither of these states prior to the measurement; it was in the  $\psi = \alpha\chi_+ + \beta\chi_-$  state. The superposition therefore does not reflect any uncertainty about the “true state” of the particle; it *is* the true state of the particle. Prior to the measurement the  $z$ -component of the spin simply does not have a well-defined value. (Unless of course either  $\alpha$  or  $\beta$  was zero.)

Now things get even stranger: *After the measurement the particle will be in the state corresponding to the value of  $S_z$  that was obtained in the measurement. The state of the particle therefore collapses from the superposition  $\psi = \alpha\chi_+ + \beta\chi_-$  to either  $\chi_+$  or  $\chi_-$ .*

Performing a measurement on a system therefore changes its quantum state in an irreversible way. Whereas prior to the measurement  $S_z$  may not have had a well-defined value, it will have a definite value after the measurement. *The act of measuring  $S_z$  is what gives it a definite value.* If we perform a follow-up measurement of  $S_z$  we are certain to obtain the

same result again.

**Remark 1:** We call  $\alpha$  the *probability amplitude* for obtaining  $+\hbar/2$  in a measurement of  $S_z$ . The square of its absolute value is the corresponding probability. Similar for  $\beta$ .

**Remark 2:** When we add the probabilities of obtaining  $+\hbar/2$  and  $-\hbar/2$  we should get one, i.e.  $|\alpha|^2 + |\beta|^2 = 1$ . *This is guaranteed by the fact that  $\psi$  is normalised.*

### E.2.3 Definitions

A complex matrix  $M$  is *Hermitian* if  $M^\dagger = M$  where  $M^\dagger = (M^*)^T$  is the conjugate transpose of  $M$ . An eigenvector of  $M$  is a *non-zero* vector  $v$  such that  $Mv = \lambda v$  with the scalar  $\lambda$  being the corresponding eigenvalue. It is known that the eigenvalues of a Hermitian matrix are real and that its eigenvectors form an orthogonal basis for the relevant vector space.

### E.2.4 Postulate 2b

The previous section dealt with measurements of  $S_z$ . What about measurements of the other two components? Here we must first clarify how physical observables, i.e. quantities we can measure, are represented in quantum mechanics.

*Physical observables, such as the three spin components, are represented by  $2 \times 2$  Hermitian matrices which act on  $\mathcal{H} = \mathbb{C}^2$  via regular matrix multiplication.*

The matrices representing the three spin components are

$$\hat{S}_x = \frac{\hbar}{2} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad \hat{S}_y = \frac{\hbar}{2} \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \quad \hat{S}_z = \frac{\hbar}{2} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad (\text{E.30})$$

Of course, the origin of these matrices is a mystery at the moment. In later courses you will see where they come from. We can now formulate a more general statement regarding the measurement of an observable.

*Suppose the electron is a spin state  $\psi$  and we measure an observable  $\mathcal{O}$  represented by the matrix  $\hat{\mathcal{O}}$ . The result of the measurement will be one of the eigenvalues of  $\hat{\mathcal{O}}$ . The probability of obtaining a specific eigenvalue  $\lambda$  is given by  $|(\phi, \psi)|^2$  where  $\phi$  is the corresponding normalised eigenvector, i.e.  $\hat{\mathcal{O}}\phi = \lambda\phi$ . Following the measurement the particle will be in the  $\phi$  state.*

We call  $(\phi, \psi)$  the *probability amplitude* of obtaining the eigenvalue  $\lambda$ . Again, the corresponding probability is the square of the absolute value of the amplitude. We see that the larger in magnitude the overlap between the system state  $\psi$  and the eigenvector  $\phi$  the greater the probability of obtaining the corresponding eigenvalue. In particular, if the system starts out in one of the eigenvectors of  $\hat{\mathcal{O}}$  then we will get the corresponding eigenvalue with certainty. (Make sure you understand why this is the case. The orthogonality of the eigenvectors is important here.) *The eigenvectors of  $\hat{\mathcal{O}}$  therefore represent states in which the observable  $\mathcal{O}$  has a well-defined value.*

Of course, we should check that this general postulate is compatible with our earlier statements regarding measurements of  $S_z$ . From the form of  $\hat{S}_z$  above this is clearly the case, since the eigenvectors of  $\hat{S}_z$  are just  $\chi_+$  and  $\chi_-$  with eigenvalues  $+\hbar/2$  and  $-\hbar/2$  respectively. This agrees with the earlier discussion. In the states  $\chi_+$  and  $\chi_-$  the  $S_z$  component of the

spin has well-defined values corresponding to the two eigenvalues of the  $\hat{S}_z$  matrix.

It will be useful to know the eigenvectors and eigenvalues of  $\hat{S}_x$  and  $\hat{S}_y$ . We find that for each of these matrices the eigenvalues are again  $+\hbar/2$  and  $-\hbar/2$ . This is really unavoidable, since the possible values of a spin component cannot depend on what we call the direction in which we are measuring the spin. No matter the direction, the only possible results are  $+\hbar/2$  and  $-\hbar/2$ . Let  $\chi_{\pm}^{(x)}$  denote the two eigenvectors of  $\hat{S}_x$  and  $\chi_{\pm}^{(y)}$  the two eigenvectors of  $\hat{S}_y$ . It is easy to show that

$$\begin{aligned}\chi_{+}^{(x)} &= \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \frac{1}{\sqrt{2}}(\chi_{+} + \chi_{-}) & \chi_{-}^{(x)} &= \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \frac{1}{\sqrt{2}}(\chi_{+} - \chi_{-}) \\ \chi_{+}^{(y)} &= \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ i \end{bmatrix} = \frac{1}{\sqrt{2}}(\chi_{+} + i\chi_{-}) & \chi_{-}^{(y)} &= \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -i \end{bmatrix} = \frac{1}{\sqrt{2}}(\chi_{+} - i\chi_{-}) \\ \chi_{+}^{(z)} &= \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \chi_{+} & \chi_{-}^{(z)} &= \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \chi_{-}\end{aligned}$$

We note that  $\chi_{+}^{(x)}$  and  $\chi_{-}^{(x)}$  are non-trivial linear combinations of  $\chi_{+}$  and  $\chi_{-}$ . This implies that while  $\chi_{\pm}^{(x)}$  are states with well-defined values of  $S_x$ , they are *not* states with well-defined values of  $S_z$ . Similarly,  $\chi_{\pm}$  are non-trivial linear combinations of  $\chi_{\pm}^{(x)}$ .

*It is therefore impossible for the electron to have well-defined  $S_x$  and  $S_z$  components simultaneously. We say that there is an uncertainty principle between these two observables.*

The same is true for any other pair of spin components. It is only possible for *one* of the three components of the electron's spin to have a well-defined value at a particular time. The underlying reason for this is that the matrices that represent the spin components do not commute, i.e.  $\hat{S}_x\hat{S}_y \neq \hat{S}_y\hat{S}_x$  and so on. We say that the spin components are *incompatible observables*. We will see that the same is true of the electron's position and momentum.

**Remark:** Using “*uncertainty*” in this context is somewhat misleading, since it creates the impression that there is some lack of knowledge about the “true state” of the system on the part of the observer. This is not the case. The state vector  $\psi$  is a complete description of the electron's spin state. *If we know  $\psi$  then we are not uncertain about anything.* However, even with this full knowledge of the system's state we are still not able to predict the outcome of a single measurement. We should rather speak about an *indeterminacy principle*, since we are not really uncertain about the values of  $S_x$  and  $S_y$ , they simply have no well-defined values until a measurement is performed.

### E.3 Final Remarks

The quantum mechanical description of the electron's spatial degrees of freedom, i.e. its position and momentum, is known as *wave mechanics*. The basic principles of the theory, particularly its linear algebra structure, remains exactly the same as in the spin formalism we have developed here. We simply trade the state (column) vectors for wave functions and the matrices that represented the observables for differential operators that act on these wave functions.

# Appendix F: Hydrogen-Like Wave Functions

## Angular Momentum

$$L^2 = \hbar^2(l^2 + l)$$

$$L_z = \hbar m$$

$$S^2 = \frac{3}{4}\hbar^2$$

$$S_z = \hbar m_s$$

## Quantum Numbers

$$(n, l, m, m_s)$$

$$n = 1, 2, 3, \dots$$

$$l = 0, 1, 2, 3, \dots, n-1$$

$$m = -l, -l+1, \dots, 0, \dots, l$$

$$m_s = \pm \frac{1}{2}$$

## Energy

$$E_n = -\frac{k^2 Z^2 e^4 \mu}{2\hbar^2 n^2} = -\frac{(13.6 \text{ eV}) Z^2}{n^2}$$

## Characteristic radius:

$$a = \frac{\hbar^2}{kZe^2\mu} = \frac{a_0}{Z} = \frac{0.0529 \text{ nm}}{Z}$$

$$\langle r \rangle \sim an^2$$

## Wave function (ignoring spin):

$$\psi_{n,l,m}(r, \theta, \phi) = R_{n,l}(r)Y_{l,m}(\theta, \phi)$$

## Spherical Harmonics

$$Y_{0,0} = \frac{1}{\sqrt{4\pi}}$$

$$Y_{1,\pm 1} = \mp \sqrt{\frac{3}{8\pi}} \sin \theta e^{\pm i\phi}$$

$$Y_{1,0} = \sqrt{\frac{3}{4\pi}} \cos \theta$$

$$Y_{2,\pm 2} = \sqrt{\frac{15}{32\pi}} \sin^2 \theta e^{\pm 2i\phi}$$

$$Y_{2,\pm 1} = \mp \sqrt{\frac{15}{8\pi}} \sin \theta \cos \theta e^{\pm i\phi}$$

$$Y_{2,0} = \sqrt{\frac{5}{16\pi}} (3 \cos^2 \theta - 1)$$

$$Y_{3,\pm 3} = \mp \sqrt{\frac{35}{64\pi}} \sin^3 \theta e^{\pm 3i\phi}$$

$$Y_{3,\pm 2} = \sqrt{\frac{105}{32\pi}} \sin^2 \theta \cos \theta e^{\pm 2i\phi}$$

$$Y_{3,\pm 1} = \mp \sqrt{\frac{21}{64\pi}} \sin \theta (5 \cos^2 \theta - 1) e^{\pm i\phi}$$

$$Y_{3,0} = \sqrt{\frac{7}{16\pi}} (5 \cos^3 \theta - 3 \cos \theta)$$

## Radial Wave Functions

$$R_{1,0} = \frac{2}{\sqrt{a^3}} e^{-r/a}$$

$$R_{2,0} = \frac{1}{\sqrt{2a^3}} \left(1 - \frac{r}{2a}\right) e^{-r/2a}$$

$$R_{2,1} = \frac{r}{2\sqrt{6a^5}} e^{-r/2a}$$

$$R_{3,0} = \frac{2}{3\sqrt{3a^3}} \left(1 - \frac{2r}{3a} + \frac{2r^2}{27a^2}\right) e^{-r/3a}$$

$$R_{3,1} = \frac{4\sqrt{2}r}{27\sqrt{3a^5}} \left(1 - \frac{r}{6a}\right) e^{-r/3a}$$

$$R_{3,2} = \frac{2\sqrt{2}r^2}{81\sqrt{15a^7}} e^{-r/3a}$$

$$R_{4,0} = \frac{1}{4\sqrt{a^3}} \left(1 - \frac{3r}{4a} + \frac{r^2}{8a^2} - \frac{r^3}{192a^3}\right) e^{-r/4a}$$

$$R_{4,1} = \frac{\sqrt{5}r}{16\sqrt{3a^5}} \left(1 - \frac{r}{4a} + \frac{r^2}{80a^2}\right) e^{-r/4a}$$

$$R_{4,2} = \frac{r^2}{320\sqrt{a^7}} \left(1 - \frac{r}{12a}\right) e^{-r/4a}$$

$$R_{4,3} = \frac{r^3}{768\sqrt{35a^9}} e^{-r/4a}$$

**TABLE 2.3 Hydrogen Atom Wave Functions: Angular Functions**

Angular Factors				Real Wave Functions			
	Related to Angular Momentum	Functions of $\theta$	In Polar Coordinates	In Cartesian Coordinates	Shapes	Label	
$l$	$m_l$	$\Phi$	$\Theta$	$\Theta\Phi(\theta, \phi)$	$\Theta\Phi(x, y, z)$		
0(s)	0	$\frac{1}{\sqrt{2\pi}}$	$\frac{1}{\sqrt{2}}$		$\frac{1}{2\sqrt{\pi}}$	$\frac{1}{2\sqrt{\pi}}$	 s
1(p)	0	$\frac{1}{\sqrt{2\pi}}$	$\frac{\sqrt{6}}{2} \cos \theta$		$\frac{1}{2}\sqrt{\frac{3}{\pi}} \cos \theta$	$\frac{1}{2}\sqrt{\frac{3}{\pi}} \frac{z}{r}$	 p <sub>z</sub>
	+1	$\frac{1}{\sqrt{2\pi}} e^{i\phi}$	$\frac{\sqrt{3}}{2} \sin \theta$		$\frac{1}{2}\sqrt{\frac{3}{\pi}} \sin \theta \cos \phi$	$\frac{1}{2}\sqrt{\frac{3}{\pi}} \frac{x}{r}$	 p <sub>x</sub>
	-1	$\frac{1}{\sqrt{2\pi}} e^{-i\phi}$	$\frac{\sqrt{3}}{2} \sin \theta$		$\frac{1}{2}\sqrt{\frac{3}{\pi}} \sin \theta \sin \phi$	$\frac{1}{2}\sqrt{\frac{3}{\pi}} \frac{y}{r}$	 p <sub>y</sub>
2(d)	0	$\frac{1}{\sqrt{2\pi}}$	$\frac{1}{2}\sqrt{\frac{5}{2}} (3\cos^2 \theta - 1)$		$\frac{1}{4}\sqrt{\frac{5}{\pi}} (3\cos^2 \theta - 1)$	$\frac{1}{4}\sqrt{\frac{5}{\pi}} \frac{(2z^2 - x^2 - y^2)}{r^2}$	 d <sub>z</sub>
	+1	$\frac{1}{\sqrt{2\pi}} e^{i\phi}$	$\frac{\sqrt{15}}{2} \cos \theta \sin \theta$		$\frac{1}{2}\sqrt{\frac{15}{\pi}} \cos \theta \sin \theta \cos \phi$	$\frac{1}{2}\sqrt{\frac{15}{\pi}} \frac{xz}{r^2}$	 d <sub>xz</sub>
	-1	$\frac{1}{\sqrt{2\pi}} e^{-i\phi}$	$\frac{\sqrt{15}}{2} \cos \theta \sin \theta$		$\frac{1}{2}\sqrt{\frac{15}{\pi}} \cos \theta \sin \theta \sin \phi$	$\frac{1}{2}\sqrt{\frac{15}{\pi}} \frac{yz}{r^2}$	 d <sub>yz</sub>
	+2	$\frac{1}{\sqrt{2\pi}} e^{2i\phi}$	$\frac{\sqrt{15}}{4} \sin^2 \theta$		$\frac{1}{4}\sqrt{\frac{15}{\pi}} \sin^2 \theta \cos 2\phi$	$\frac{1}{4}\sqrt{\frac{15}{\pi}} \frac{(x^2 - y^2)}{r^2}$	 d <sub>x<sup>2</sup>-y<sup>2</sup></sub>
	-2	$\frac{1}{\sqrt{2\pi}} e^{-2i\phi}$	$\frac{\sqrt{15}}{4} \sin^2 \theta$		$\frac{1}{4}\sqrt{\frac{15}{\pi}} \sin^2 \theta \sin 2\phi$	$\frac{1}{4}\sqrt{\frac{15}{\pi}} \frac{xy}{r^2}$	 d <sub>xy</sub>

Source: Hydrogen Atom Wave Functions: Angular Functions, *Physical Chemistry*, 5th ed., Gordon Barrow (c) 1988. McGraw-Hill Companies, Inc.

NOTE: The relations  $(e^{i\phi} - e^{-i\phi})/(2i) = \sin \phi$  and  $(e^{i\phi} + e^{-i\phi})/2 = \cos \phi$  can be used to convert the exponential imaginary functions to real trigonometric functions, combining the two orbitals with  $m_l = \pm 1$  to give two orbitals with  $\sin \phi$  and  $\cos \phi$ . In a similar fashion, the orbitals with  $m_l = \pm 2$  result in real functions with  $\cos^2 \phi$  and  $\sin^2 \phi$ . These functions have then been converted to Cartesian form by using the functions  $x = r \sin \theta \cos \phi$ ,  $y = r \sin \theta \sin \phi$ , and  $z = r \cos \theta$ .

# Appendix G: Angular Momentum Related Coefficients

## G.1 Ladder operator coefficients

$c_+^2(j, m)$		m								
		4	3	2	1	0	-1	-2	-3	-4
j	0				0					
	1				0	2	2			
	2			0	4	6	6	4		
	3		0	6	10	12	12	10	6	
	4	0	8	14	18	20	20	18	14	8

$c_+^2(j, m)$		m							
		7/2	5/2	3/2	1/2	-1/2	-3/2	-5/2	-7/2
j	1/2				0	1			
	3/2			0	3	4	3		
	5/2		0	5	8	9	8	5	
	7/2	0	7	12	15	16	15	12	7

$c_-^2(j, m)$		m								
		4	3	2	1	0	-1	-2	-3	-4
j	0				0					
	1				2	2	0			
	2			4	6	6	4	0		
	3		6	10	12	12	10	6	0	
	4	8	14	18	20	20	18	14	8	0

$c_-^2(j, m)$		m							
		7/2	5/2	3/2	1/2	-1/2	-3/2	-5/2	-7/2
j	1/2				1	0			
	3/2			3	4	3	0		
	5/2		5	8	9	8	5	0	
	7/2	7	12	15	16	15	12	7	0

### 34. CLEBSCH-GORDAN COEFFICIENTS, SPHERICAL HARMONICS, AND $d$ FUNCTIONS

Note: A square-root sign is to be understood over *every* coefficient, e.g., for  $-8/15$  read  $-\sqrt{8/15}$ .

$$1/2 \times 1/2 \begin{matrix} 1 \\ +1/2+1/2 \\ -1/2+1/2 \\ -1/2-1/2 \end{matrix} \begin{matrix} 1 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{matrix}$$

$$Y_1^0 = \sqrt{\frac{3}{4\pi}} \cos \theta$$

$$2 \times 1/2 \begin{matrix} 5/2 \\ +5/2 \\ +1/2 \\ +1/2 \end{matrix} \begin{matrix} 5/2 & 3/2 \\ 1 & +3/2+3/2 \\ 1 & 1 \\ 1 & 1 \end{matrix}$$

$J$	$J$	...
$M$	$M$	...
$m_1$	$m_2$	
$m_1$	$m_2$	Coefficients
$\vdots$	$\vdots$	
$\vdots$	$\vdots$	

$$1 \times 1/2 \begin{matrix} 3/2 \\ +3/2 \\ +1/2 \\ 0+1/2 \end{matrix} \begin{matrix} 3/2 & 1/2 \\ 1 & +1/2+1/2 \\ 1/3 & 2/3 \\ 2/3 & -1/3 \end{matrix} \begin{matrix} 3/2 & 1/2 \\ -1/2 & -1/2 \end{matrix}$$

$$Y_2^0 = \sqrt{\frac{5}{4\pi}} \left( \frac{3}{2} \cos^2 \theta - \frac{1}{2} \right)$$

$$Y_2^1 = -\sqrt{\frac{15}{8\pi}} \sin \theta \cos \theta e^{i\phi}$$

$$3/2 \times 1/2 \begin{matrix} 2 \\ +3/2 \\ +1/2 \end{matrix} \begin{matrix} 2 & 1 \\ 1 & +1 \\ 1/4 & 3/4 \\ 3/4 & -1/4 \end{matrix} \begin{matrix} 2 & 1 \\ 0 & 0 \\ 1/2 & 1/2 \\ -1/2 & -1/2 \end{matrix}$$

$$1 \times 1 \begin{matrix} 2 \\ +2 \\ +1+1 \end{matrix} \begin{matrix} 2 & 1 \\ 0+1 \\ 2/3 & -1/3 \end{matrix} \begin{matrix} -1/2 & 1 \\ 1 & 1 \end{matrix}$$

$$Y_2^2 = \frac{1}{4} \sqrt{\frac{15}{2\pi}} \sin^2 \theta e^{2i\phi}$$

$$3/2 \times 1 \begin{matrix} 5/2 \\ +5/2 \\ +3/2+1 \end{matrix} \begin{matrix} 5/2 & 3/2 \\ 1 & +3/2+3/2 \\ 0 & 2/5 \\ 3/5 & -2/5 \end{matrix} \begin{matrix} 5/2 & 3/2 & 1/2 \\ 1/10 & 2/5 & 1/2 \\ 3/5 & 1/15 & -1/3 \\ -1/2+1/10 & -8/15 & 1/6 \end{matrix}$$

$$Y_\ell^{-m} = (-1)^m Y_\ell^m$$

$$d_{m,0}^\ell = \sqrt{\frac{4\pi}{2\ell+1}} Y_\ell^m e^{-im\phi}$$

$$\langle j_1 j_2 m_1 m_2 | j_1 j_2 JM \rangle \\ = (-1)^{J-j_1-j_2} \langle j_2 j_1 m_2 m_1 | j_2 j_1 JM \rangle$$

$$d_{m',m}^j = (-1)^{m-m'} d_{m,m'}^j = d_{-m,-m'}^j$$

$$2 \times 2 \begin{matrix} 4 \\ +4 \\ +2+2 \\ +2+2 \end{matrix} \begin{matrix} 4 & 3 \\ 1 & +3+3 \\ 1/2 & 1/2 \\ 1/2-1/2 \end{matrix} \begin{matrix} 4 & 3 & 2 \\ 2 & +2 & +2 \\ 0 & 0+2 \\ 3/14 & -1/2 & 2/7 \end{matrix}$$

$$3/2 \times 3/2 \begin{matrix} 3 \\ +3 \\ +3/2+3/2 \end{matrix} \begin{matrix} 3 & 2 \\ 1 & +2+2 \\ 1/2 & 1/2 \\ 1/2+1/2 & 1/2+1/2 \end{matrix} \begin{matrix} 3 & 2 & 1 \\ 1/5 & 1/2 & 3/10 \\ 3/5 & 0 & -2/5 \\ -1/2+3/2 & 1/5-1/2 & 3/10 \end{matrix}$$

$$d_{0,0}^1 = \cos \theta$$

$$d_{1/2,1/2}^{1/2} = \cos \frac{\theta}{2}$$

$$d_{1,1}^1 = \frac{1+\cos \theta}{2}$$

$$d_{1/2,-1/2}^{1/2} = -\sin \frac{\theta}{2}$$

$$d_{1,0}^1 = -\frac{\sin \theta}{\sqrt{2}}$$

$$d_{1,-1}^1 = \frac{1-\cos \theta}{2}$$

$$2 \times 2 \begin{matrix} 4 \\ +4 \\ +2+2 \\ +2+2 \end{matrix} \begin{matrix} 4 & 3 & 2 \\ 1 & +3+3 \\ 1/2 & 1/2 \\ 1/2-1/2 \end{matrix} \begin{matrix} 4 & 3 & 2 & 1 \\ 1 & +1 & +1 & +1 \\ 1/14 & 3/10 & 3/7 & 1/5 \\ 3/7 & 1/5-1/14 & -3/10 & 0 \\ 3/7 & -1/5-1/14 & 3/10 & 0+1 \\ 1/14-3/10 & 3/7 & -1/5 & -1+3/2 \end{matrix}$$

$$2 \times 3/2 \begin{matrix} 7/2 \\ +7/2 \\ +2+3/2 \\ +2+3/2 \end{matrix} \begin{matrix} 7/2 & 5/2 \\ 1 & +5/2+5/2 \\ 3/7 & 4/7 \\ 4/7-3/7 & +3/2+3/2 \end{matrix} \begin{matrix} 7/2 & 5/2 & 2/5 \\ 1/7 & 16/35 & 2/5 \\ 4/7 & 1/35-2/5 & 1/5 \\ 0+3/2 & 2/7-18/35 & 1/5 \end{matrix}$$

$$2 \times 3/2 \begin{matrix} 7/2 \\ +7/2 \\ +2+3/2 \\ +2+3/2 \end{matrix} \begin{matrix} 7/2 & 5/2 & 2/5 \\ 1/7 & 16/35 & 2/5 \\ 4/7 & 1/35-2/5 & 1/5 \\ 0+3/2 & 2/7-18/35 & 1/5 \end{matrix} \begin{matrix} 7/2 & 5/2 & 3/2 & 1/2 \\ 1/20 & 1/4 & 9/20 & 1/4 \\ 9/20 & 1/4-1/20 & 1/4 & 1/20-1/4 \\ -1/2+1/20 & -1/4-1/20 & 1/4 & -1/20-1/4 \end{matrix}$$

$$2 \times 3/2 \begin{matrix} 7/2 \\ +7/2 \\ +2+3/2 \\ +2+3/2 \end{matrix} \begin{matrix} 7/2 & 5/2 & 3/2 & 1/2 \\ 1/20 & 1/4 & 9/20 & 1/4 \\ 9/20 & 1/4-1/20 & 1/4 & -1/20-1/4 \\ -1/2+1/20 & -1/4-1/20 & 1/4 & -1/20-1/4 \end{matrix} \begin{matrix} 3 & 2 & 1 \\ -1 & -1 & -1 \end{matrix}$$

$$2 \times 3/2 \begin{matrix} 7/2 \\ +7/2 \\ +2+3/2 \\ +2+3/2 \end{matrix} \begin{matrix} 7/2 & 5/2 & 3/2 & 1/2 \\ 1/20 & 1/4 & 9/20 & 1/4 \\ 9/20 & 1/4-1/20 & 1/4 & -1/20-1/4 \\ -1/2+1/20 & -1/4-1/20 & 1/4 & -1/20-1/4 \end{matrix} \begin{matrix} 3 & 2 & 1 \\ -1 & -1 & -1 \end{matrix}$$

$$2 \times 3/2 \begin{matrix} 7/2 \\ +7/2 \\ +2+3/2 \\ +2+3/2 \end{matrix} \begin{matrix} 7/2 & 5/2 & 3/2 & 1/2 \\ 1/20 & 1/4 & 9/20 & 1/4 \\ 9/20 & 1/4-1/20 & 1/4 & -1/20-1/4 \\ -1/2+1/20 & -1/4-1/20 & 1/4 & -1/20-1/4 \end{matrix} \begin{matrix} 3 & 2 & 1 \\ -1 & -1 & -1 \end{matrix}$$

$$2 \times 3/2 \begin{matrix} 7/2 \\ +7/2 \\ +2+3/2 \\ +2+3/2 \end{matrix} \begin{matrix} 7/2 & 5/2 & 3/2 & 1/2 \\ 1/20 & 1/4 & 9/20 & 1/4 \\ 9/20 & 1/4-1/20 & 1/4 & -1/20-1/4 \\ -1/2+1/20 & -1/4-1/20 & 1/4 & -1/20-1/4 \end{matrix} \begin{matrix} 3 & 2 & 1 \\ -1 & -1 & -1 \end{matrix}$$

$$d_{3/2,3/2}^{3/2} = \frac{1+\cos \theta}{2} \cos \frac{\theta}{2}$$

$$d_{3/2,1/2}^{3/2} = -\sqrt{3} \frac{1+\cos \theta}{2} \sin \frac{\theta}{2}$$

$$d_{3/2,-1/2}^{3/2} = \sqrt{3} \frac{1-\cos \theta}{2} \cos \frac{\theta}{2}$$

$$d_{3/2,-3/2}^{3/2} = -\frac{1-\cos \theta}{2} \sin \frac{\theta}{2}$$

$$d_{1/2,1/2}^{3/2} = \frac{3\cos \theta - 1}{2} \cos \frac{\theta}{2}$$

$$d_{1/2,-1/2}^{3/2} = -\frac{3\cos \theta + 1}{2} \sin \frac{\theta}{2}$$

$$d_{2,2}^2 = \left( \frac{1+\cos \theta}{2} \right)^2$$

$$d_{2,1}^2 = -\frac{1+\cos \theta}{2} \sin \theta$$

$$d_{2,0}^2 = \frac{\sqrt{6}}{4} \sin^2 \theta$$

$$d_{2,-1}^2 = -\frac{1-\cos \theta}{2} \sin \theta$$

$$d_{2,-2}^2 = \left( \frac{1-\cos \theta}{2} \right)^2$$

$$d_{1,2}^2 = \frac{1+\cos \theta}{2} (2\cos \theta - 1)$$

$$d_{1,0}^2 = -\sqrt{\frac{3}{2}} \sin \theta \cos \theta$$

$$d_{1,-1}^2 = \frac{1-\cos \theta}{2} (2\cos \theta + 1)$$

$$d_{0,0}^2 = \left( \frac{3}{2} \cos^2 \theta - \frac{1}{2} \right)$$

Figure 34.1: The sign convention is that of Wigner (*Group Theory*, Academic Press, New York, 1959), also used by Condon and Shortley (*The Theory of Atomic Spectra*, Cambridge Univ. Press, New York, 1953), Rose (*Elementary Theory of Angular Momentum*, Wiley, New York, 1957), and Cohen (*Tables of the Clebsch-Gordan Coefficients*, North American Rockwell Science Center, Thousand Oaks, Calif., 1974). The coefficients here have been calculated using computer programs written independently by Cohen and at LBNL.

# Appendix H: Further Information

**Periodic Table of the Elements**

1 <b>H</b> Hydrogen 1.008	2 <b>He</b> Helium 4.003	3 <b>Li</b> Lithium 6.941	4 <b>Be</b> Beryllium 9.012	5 <b>B</b> Boron 10.811	6 <b>C</b> Carbon 12.011	7 <b>N</b> Nitrogen 14.007	8 <b>O</b> Oxygen 15.999	9 <b>F</b> Fluorine 18.998	10 <b>Ne</b> Neon 20.180
11 <b>Na</b> Sodium 22.990	12 <b>Mg</b> Magnesium 24.305	13 <b>Al</b> Aluminum 26.982	14 <b>Si</b> Silicon 28.086	15 <b>P</b> Phosphorus 30.974	16 <b>S</b> Sulfur 32.066	17 <b>Cl</b> Chlorine 35.453	18 <b>Ar</b> Argon 39.948	19 <b>K</b> Potassium 39.098	20 <b>Ca</b> Calcium 40.078
21 <b>Sc</b> Scandium 44.956	22 <b>Ti</b> Titanium 47.88	23 <b>V</b> Vanadium 50.942	24 <b>Cr</b> Chromium 51.996	25 <b>Mn</b> Manganese 54.938	26 <b>Fe</b> Iron 55.933	27 <b>Co</b> Cobalt 58.933	28 <b>Ni</b> Nickel 58.693	29 <b>Cu</b> Copper 63.546	30 <b>Zn</b> Zinc 65.39
31 <b>Ga</b> Gallium 69.732	32 <b>Ge</b> Germanium 72.61	33 <b>As</b> Arsenic 74.922	34 <b>Se</b> Selenium 78.09	35 <b>Br</b> Bromine 79.904	36 <b>Kr</b> Krypton 84.80	37 <b>Rb</b> Rubidium 84.468	38 <b>Sr</b> Strontium 87.62	39 <b>Y</b> Yttrium 88.906	40 <b>Nb</b> Niobium 92.906
41 <b>Tc</b> Technetium 98.907	42 <b>Mo</b> Molybdenum 95.94	43 <b>Ru</b> Ruthenium 101.07	44 <b>Rh</b> Rhodium 102.906	45 <b>Pd</b> Palladium 106.442	46 <b>Ag</b> Silver 107.868	47 <b>Cd</b> Cadmium 112.411	48 <b>In</b> Indium 114.818	49 <b>Sn</b> Tin 118.71	50 <b>Sb</b> Antimony 121.760
51 <b>Tl</b> Thallium 121.760	52 <b>Pb</b> Lead 127.6	53 <b>Bi</b> Bismuth 126.904	54 <b>Xe</b> Xenon 131.29	55 <b>Cs</b> Cesium 132.905	56 <b>Ba</b> Barium 137.327	57-71 <b>La</b> Lanthanum 138.906	72 <b>Hf</b> Hafnium 178.49	73 <b>Ta</b> Tantalum 180.948	74 <b>W</b> Tungsten 183.85
75 <b>Re</b> Rhenium 186.207	76 <b>Os</b> Osmium 190.23	77 <b>Ir</b> Iridium 192.22	78 <b>Pt</b> Platinum 195.08	79 <b>Au</b> Gold 196.967	80 <b>Hg</b> Mercury 202.59	81 <b>Tl</b> Thallium 204.383	82 <b>Pb</b> Lead 207.2	83 <b>Bi</b> Bismuth 208.960	84 <b>Po</b> Polonium [208.962]
85 <b>Rn</b> Radon 222.018	86 <b>At</b> Astatine 223.867	87 <b>Fr</b> Francium 223.020	88 <b>Ra</b> Radium 226.025	89-103 <b>Rf</b> Rutherfordium [261]	104 <b>Ds</b> Dubnium [262]	105 <b>Rg</b> Roentgenium [269]	106 <b>Sg</b> Seaborgium [266]	107 <b>Bh</b> Bohrium [254]	108 <b>Mt</b> Meitnerium [268]
109 <b>Hs</b> Hassium [289]	110 <b>Ds</b> Darmstadtium [269]	111 <b>Rg</b> Roentgenium [272]	112 <b>Cn</b> Copernicium [277]	113 <b>Uut</b> Ununtrium unknown	114 <b>Fl</b> Flerovium [289]	115 <b>Uup</b> Ununpentium unknown	116 <b>Lv</b> Livermorium [298]	117 <b>Uus</b> Ununseptium unknown	118 <b>Uuo</b> Ununoctium unknown
Lanthanide Series	Actinide Series	57 <b>La</b> Lanthanum 138.906	58 <b>Ce</b> Cerium 140.115	59 <b>Pr</b> Praseodymium 140.908	60 <b>Nd</b> Neodymium 144.24	61 <b>Pm</b> Promethium 144.913	62 <b>Sm</b> Samarium 150.36	63 <b>Eu</b> Europium 151.966	64 <b>Gd</b> Gadolinium 157.25
89 <b>Ac</b> Actinium 227.028	90 <b>Th</b> Thorium 232.038	91 <b>Pa</b> Protactinium 231.036	92 <b>U</b> Uranium 238.029	93 <b>Np</b> Neptunium 237.048	94 <b>Pu</b> Plutonium 244.064	95 <b>Am</b> Americium 243.061	96 <b>Cm</b> Curium 247.070	97 <b>Bk</b> Berkelium 247.070	98 <b>Cf</b> Californium 251.080
99 <b>E</b> Einsteinium [254]	100 <b>Fm</b> Fermium 257.095	101 <b>Md</b> Mendelevium 258.1	102 <b>No</b> Nobelium 259.101	103 <b>Lr</b> Lawrencium [262]	71 <b>Lu</b> Lutetium 174.967	70 <b>Yb</b> Ytterbium 173.04	69 <b>Tm</b> Thulium 168.934	68 <b>Er</b> Erbium 167.26	67 <b>Ho</b> Holmium 164.930

© 2013 Todd Helmenstine  
chemistry4abouth.com

## Fundamental Constants

Planck's constant:  $\hbar = 1.05457 \times 10^{-34} \text{ Js}$

Speed of light:  $c = 2.99792 \times 10^8 \text{ m/s}$

Mass of electron:  $m_e = 9.10938 \times 10^{-31} \text{ kg}$

Mass of proton:  $m_p = 1.67262 \times 10^{-27} \text{ kg}$

Charge of proton:  $e = 1.60218 \times 10^{-19} \text{ C}$

Charge of electron:  $-e = -1.60218 \times 10^{-19} \text{ C}$

Permittivity of space:  $\epsilon_0 = 8.85419 \times 10^{-12} \text{ C}^2/\text{J m}$

Boltzmann constant:  $k_B = 1.38065 \times 10^{-23} \text{ J/K}$

## Hydrogen Atom

Fine structure constant:  $\alpha = \frac{e^2}{4\pi\epsilon_0\hbar c} = 1/137.036$

Bohr radius:  $a = \frac{4\pi\epsilon_0\hbar^2}{m_e e^2} = \frac{\hbar}{\alpha m_e c} = 5.29177 \times 10^{-11} \text{ m}$

Bohr energies:  $E_n = -\frac{m_e e^4}{2(4\pi\epsilon_0)^2\hbar^2 n^2} = \frac{E_1}{n^2} (n = 1, 2, 3, \dots)$

Binding energy:  $-E_1 = \frac{\hbar^2}{2m_e a^2} = \frac{\alpha^2 m_e c^2}{2} = 13.6057 \text{ eV}$

Ground state:  $\psi_0 = \frac{1}{\sqrt{\pi a^3}} e^{-r/a}$

Rydberg formula:  $\frac{1}{\lambda} = R \left( \frac{1}{n_f^2} - \frac{1}{n_i^2} \right)$

Rydberg constant:  $R = -\frac{E_1}{2\pi\hbar c} = 1.09737 \times 10^7 / \text{m}$