

PHY472

Advanced Quantum Mechanics

Pieter Kok, The University of Sheffield.

August 2015

PHY472: Lecture Topics

1. Linear vector spaces
2. Operators and the spectral decomposition
3. Observables, projectors and time evolution
4. Tensor product spaces
5. The postulates of quantum mechanics I
6. The postulates of quantum mechanics II
7. The Schrödinger and Heisenberg picture
8. Mixed states and the density matrix
9. Perfect and imperfect measurements
10. Composite systems and entanglement
11. Quantum teleportation
12. Open quantum systems and completely positive maps
13. Orbital angular momentum
14. Spin angular momentum
15. Total angular momentum
16. Identical particles
17. Bose-Einstein and Fermi-Dirac statistics
18. Spin waves in solids
19. An atom in a cavity
20. Quantum field theory
21. Revision lecture (week before the exam)

PHY472: Advanced Quantum Mechanics

Contents

1	Linear Vector Spaces and Hilbert Space	7
1.1	Linear vector spaces	7
1.2	Operators in Hilbert space	8
1.3	Hermitian and unitary operators	11
1.4	Projection operators and tensor products	12
1.5	The trace and determinant of an operator	15
2	The Postulates of Quantum Mechanics	18
3	Schrödinger and Heisenberg Pictures	24
4	Mixed States and the Density Operator	29
4.1	Mixed states	29
4.2	Decoherence	31
4.3	Imperfect measurements	32
5	Composite Systems and Entanglement	37
5.1	Composite systems	37
5.2	Entanglement	37
5.3	Quantum teleportation	41
6	Evolution of Open Quantum Systems	44
6.1	The Lindblad equation	44
6.2	Positive and completely positive maps	45
7	Orbital Angular Momentum and Spin	47
7.1	Orbital angular momentum	47
7.2	Spin	49
7.3	Total angular momentum	52
7.4	Composite systems with angular momentum	53
8	Identical Particles	57
8.1	Symmetric and anti-symmetric states	57
8.2	Creation and annihilation operators	58
8.3	Observables based on creation and annihilation operators	62
8.4	Bose-Einstein and Fermi-Dirac statistics	63
9	Many-Body Problems in Quantum Mechanics	66
9.1	Interacting electrons in atomic shells	66
9.2	Spin waves in solids	68
9.3	An atom in a cavity	71
9.4	Outlook: quantum field theory	73

Suggested Further Reading

1. **Quantum Processes, Systems, and Information**, by Schumacher and Westmoreland, Cambridge University Press (2010). This is an excellent book, and should be your first choice for additional material. It has everything up to many-body quantum mechanics.
2. **Quantum Information and Quantum Computation**, by Nielsen and Chuang, Cambridge University Press (2000). This is the current standard work on quantum information theory. It has a comprehensive introduction to quantum mechanics along the lines treated here, but in more depth. The book is from 2000, which means that several important recent topics are not covered.
3. **Introductory Quantum Optics**, by Gerry and Knight, Cambridge University Press (2005). This is a very accessible introduction to the quantum theory of light.
4. **Quantum Field Theory**, by Lewis Ryder, Cambridge University Press (1996). This is a quite advanced introduction to relativistic quantum field theory.

I would like to thank the students who have used these lecture notes in previous years and spotted typos or errors (notably Tom Bullock). Their efforts have greatly improved the readability of these notes.

1 Linear Vector Spaces and Hilbert Space

The modern version of quantum mechanics was formulated in 1932 by John von Neumann in his famous book *Mathematical Foundations of Quantum Mechanics*, and it unifies Schrödingers wave theory with the matrix mechanics of Heisenberg, Born, and Jordan. The theory is framed in terms of linear vector spaces, so the first couple of lectures we have to remind ourselves of the relevant mathematics.

1.1 Linear vector spaces

Consider a set of *vectors*, denoted by $|\psi\rangle$, $|\phi\rangle$, etc., and the complex numbers a , b , c , etc. A *linear vector space* \mathcal{V} is a mathematical structure of vectors and numbers that obeys the following rules:

1. $|\psi\rangle + |\phi\rangle = |\phi\rangle + |\psi\rangle$ (commutativity),
2. $|\psi\rangle + (|\phi\rangle + |\chi\rangle) = (|\psi\rangle + |\phi\rangle) + |\chi\rangle$ (associativity),
3. $a(|\psi\rangle + |\phi\rangle) = a|\psi\rangle + a|\phi\rangle$ (linearity),
4. $(a + b)|\psi\rangle = a|\psi\rangle + b|\psi\rangle$ (linearity),
5. $a(b|\phi\rangle) = (ab)|\phi\rangle$.

There is also a *null vector* 0 such that $|\psi\rangle + 0 = |\psi\rangle$, and for every $|\psi\rangle$ there is a vector $|\phi\rangle$ such that $|\psi\rangle + |\phi\rangle = 0$.

For each vector $|\phi\rangle$ there is a *dual vector* $\langle\phi|$, and the set of dual vectors also form a linear vector space \mathcal{V}^* . There is an inner product between vectors from \mathcal{V} and \mathcal{V}^* denoted by $\langle\psi|\phi\rangle$. The inner product has the following properties:

1. $\langle\psi|\phi\rangle = \langle\phi|\psi\rangle^*$,
2. $\langle\psi|\psi\rangle \geq 0$,
3. $\langle\psi|\psi\rangle = 0 \Leftrightarrow |\psi\rangle = 0$,
4. $|\psi\rangle = c_1|\psi_1\rangle + c_2|\psi_2\rangle \Rightarrow \langle\phi|\psi\rangle = c_1\langle\phi|\psi_1\rangle + c_2\langle\phi|\psi_2\rangle$,
5. $\|\phi\| \equiv \sqrt{\langle\phi|\phi\rangle}$ is the *norm* of $|\phi\rangle$.

If $\|\phi\| = 1$, the vector $|\phi\rangle$ is a *unit* vector. The set of unit vectors $\{e^{i\varphi}|\psi\rangle\}$ with $\varphi \in [0, 2\pi)$ form a so-called *ray* in the linear vector space. A linear vector space that has a norm $\|\cdot\|$ (there are many different ways we can define a norm) is called a *Hilbert space*. We will always assume that the linear vector spaces are Hilbert spaces.

For linear vector spaces with an inner product we can derive the *Cauchy-Schwarz* inequality, also known as the Schwarz inequality:

$$|\langle\phi|\psi\rangle|^2 \leq \langle\psi|\psi\rangle\langle\phi|\phi\rangle. \quad (1.1)$$

This is a very important relation, since it requires only the inner product structure. Relations that are based on this inequality, such as the Heisenberg uncertainty relation between observables, therefore have a very general validity.

If two vectors have an inner product equal to zero, then these vectors are called orthogonal. This is the definition of orthogonality. When these vectors are also unit vectors, they are called *orthonormal*. A set of vectors $|\phi_1\rangle, |\phi_2\rangle, \dots |\phi_N\rangle$ are linearly independent if

$$\sum_j a_j |\phi_j\rangle = 0 \quad (1.2)$$

implies that all $a_j = 0$. The maximum number of linearly independent vectors in \mathcal{V} is the *dimension* of \mathcal{V} . Orthonormal vectors form a complete orthonormal basis for \mathcal{V} if *any* vector can be written as

$$|\psi\rangle = \sum_{k=1}^N c_k |\phi_k\rangle, \quad (1.3)$$

and $\langle\phi_j|\phi_k\rangle = \delta_{jk}$. We can take the inner product of $|\psi\rangle$ with any of the basis vectors $|\phi_j\rangle$ to obtain

$$\langle\phi_j|\psi\rangle = \sum_{k=1}^N c_k \langle\phi_j|\phi_k\rangle = \sum_{k=1}^N c_k \delta_{jk} = c_j. \quad (1.4)$$

Substitute this back into the expansion of $|\psi\rangle$, and we find

$$|\psi\rangle = \sum_{k=1}^N |\phi_k\rangle \langle\phi_k|\psi\rangle. \quad (1.5)$$

Therefore $\sum_k |\phi_k\rangle \langle\phi_k|$ must act like the *identity*. In fact, this gives us an important clue that operators of states must take the general form of sums over objects like $|\phi\rangle \langle\chi|$.

1.2 Operators in Hilbert space

The objects $|\psi\rangle$ are vectors in a Hilbert space. We can imagine applying rotations of the vectors, rescaling, permutations of vectors in a basis, and so on. These are described mathematically as *operators*, and we denote them by capital letters A, B, C , etc. In general we write

$$A |\phi\rangle = |\psi\rangle, \quad (1.6)$$

for some $|\phi\rangle, |\psi\rangle \in \mathcal{V}$. It is important to remember that operators act on *all* the vectors in Hilbert space. Let $\{|\phi_j\rangle\}_j$ be an orthonormal basis. We can calculate the inner product between the vectors $|\phi_j\rangle$ and $A |\phi_k\rangle$:

$$\langle\phi_j|(A |\phi_k\rangle) = \langle\phi_j|A |\phi_k\rangle \equiv A_{jk}. \quad (1.7)$$

The two indices indicate that operators are *matrices*.

As an example, consider two vectors, written as two-dimensional column vectors

$$|\phi_1\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |\phi_2\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad (1.8)$$

and suppose that

$$A = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}. \quad (1.9)$$

We calculate

$$A_{11} = \langle \phi_1 | A | \phi_1 \rangle = (1, 0) \cdot \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = (1, 0) \cdot \begin{pmatrix} 2 \\ 0 \end{pmatrix} = 2. \quad (1.10)$$

Similarly, we can calculate that $A_{22} = 3$, and $A_{12} = A_{21} = 0$ (check this). We therefore have that $A | \phi_1 \rangle = 2 | \phi_1 \rangle$ and $A | \phi_2 \rangle = 3 | \phi_2 \rangle$.

Complex numbers a have complex conjugates a^* and vectors $|\psi\rangle$ have dual vectors $\langle\phi|$. Is there an equivalent for operators? The answer is yes, and it is called the *adjoint*, or *Hermitian conjugate*, and is denoted by a dagger \dagger . The natural way to define it is according to the rule

$$\langle\psi|A|\phi\rangle^* = \langle\phi|A^\dagger|\psi\rangle, \quad (1.11)$$

for any $|\phi\rangle$ and $|\psi\rangle$. In matrix notation, and given an orthonormal basis $\{|\phi_j\rangle\}_j$, this becomes

$$\langle\phi_j|A|\phi_k\rangle^* = A_{jk}^* = \langle\phi_k|A^\dagger|\phi_j\rangle = A_{kj}^\dagger. \quad (1.12)$$

So the matrix representation of the adjoint A^\dagger is the transpose and the complex conjugate of the matrix A , as given by $(A^\dagger)_{jk} = A_{kj}^*$. The adjoint has the following properties:

1. $(cA)^\dagger = c^* A^\dagger$,
2. $(AB)^\dagger = B^\dagger A^\dagger$,
3. $(|\phi\rangle)^\dagger = \langle\phi|$.

Note the order of the operators in 2: AB is generally not the same as BA . The difference between the two is called the *commutator*, denoted by

$$[A, B] = AB - BA. \quad (1.13)$$

For example, we can choose

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad (1.14)$$

which leads to

$$\begin{aligned} [A, B] &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} - \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} - \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 & -2 \\ 2 & 0 \end{pmatrix} \neq 0. \end{aligned} \quad (1.15)$$

Many, but not all, operators have an *inverse*. Let $A|\phi\rangle = |\psi\rangle$ and $B|\psi\rangle = |\phi\rangle$. Then we have

$$BA|\phi\rangle = |\phi\rangle \quad \text{and} \quad AB|\psi\rangle = |\psi\rangle. \quad (1.16)$$

If Eq. (1.16) holds true for all $|\phi\rangle$ and $|\psi\rangle$, then B is the inverse of A , and we write $B = A^{-1}$. An operator that has an inverse is called *invertible*. Another important property that an operator may possess is *positivity*. An operator is positive if

$$\langle\phi|A|\phi\rangle \geq 0 \quad \text{for all } |\phi\rangle. \quad (1.17)$$

We also write this as $A \geq 0$.

From the matrix representation of operators you can easily see that the operators themselves form a linear vector space:

1. $A + B = B + A$,
2. $A + (B + C) = (A + B) + C$,
3. $a(A + B) = aA + aB$,
4. $(a + b)A = aA + bA$,
5. $a(bA) = (ab)A$.

We can also define the operator norm $\|A\|$ according to

$$\|A\| = \sqrt{\text{Tr}(A^\dagger A)} \equiv \sqrt{\sum_{ij} A_{ij}^* A_{ji}}, \quad (1.18)$$

which means that the linear vector space of operators is again a Hilbert space. The symbol $\text{Tr}(\cdot)$ denotes the *trace* of an operator, and we will return to this special operator property later in this section.

Every operator has a set of vectors for which

$$A|a_j\rangle = a_j|a_j\rangle, \quad \text{with } a_j \in \mathbb{C}. \quad (1.19)$$

This is called the eigenvalue equation (or eigenequation) for A , and the vectors $|a_j\rangle$ are the eigenvectors. The complex numbers a_j are eigenvalues. In the basis of eigenvectors, the matrix representation of A becomes

$$A = \begin{pmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_N \end{pmatrix} \quad (1.20)$$

When some of the a_j s are the same, we speak of *degenerate* eigenvalues. When there are n identical eigenvalues, we have n -fold degeneracy. The eigenvectors corresponding to this eigenvalue then span an n -dimensional subspace of the vector space. We will return to subspaces shortly, when we introduce projection operators.

For any orthonormal basis $\{|\phi_j\rangle\}_j$ we have

$$\langle\phi_j|A|\phi_k\rangle = A_{jk}, \quad (1.21)$$

which can be written in the form

$$A = \sum_{jk} A_{jk} |\phi_j\rangle\langle\phi_k|. \quad (1.22)$$

For the special case where $|\phi_j\rangle = |a_j\rangle$ this reduces to

$$A = \sum_j a_j |a_j\rangle\langle a_j|. \quad (1.23)$$

This is the *spectral decomposition* of A . When all a_j are equal, we have complete degeneracy over the full vector space, and the operator becomes proportional (up to a factor a_j) to the identity \mathbb{I} .

Note that this is *independent* of the basis $\{|a_j\rangle\}$. As a consequence, for any orthonormal basis $\{|\phi_j\rangle\}$ we have

$$\mathbb{I} = \sum_j |\phi_j\rangle\langle\phi_j|. \quad (1.24)$$

This is the *completeness relation*, and we will use this many times in our calculations.

Lemma: If two non-degenerate operators commute ($[A, B] = 0$), then they have a common set of eigenvectors.

Proof: Let $A = \sum_k a_k |a_k\rangle\langle a_k|$ and $B = \sum_{jk} B_{jk} |a_j\rangle\langle a_k|$. We can choose this without loss of generality: we write both operators in the eigenbasis of A . Furthermore, $[A, B] = 0$ implies that $AB = BA$.

$$\begin{aligned} AB &= \sum_{klm} a_k B_{lm} |a_k\rangle\langle a_l| \langle a_m| = \sum_{lm} a_l B_{lm} |a_l\rangle\langle a_m| \\ BA &= \sum_{klm} a_k B_{lm} |a_l\rangle\langle a_m| \langle a_k| = \sum_{lm} a_m B_{lm} |a_l\rangle\langle a_m| \end{aligned} \quad (1.25)$$

Therefore

$$[A, B] = \sum_{lm} (a_l - a_m) B_{lm} |a_l\rangle\langle a_m| = 0. \quad (1.26)$$

If $a_l \neq a_m$ for $l \neq m$, then $B_{lm} = 0$, and $B_{lm} \propto \delta_{lm}$. Therefore $\{|a_j\rangle\}$ is an eigenbasis for B . \square

The proof of the converse is left as an exercise. It turns out that this is also true when A and/or B are degenerate.

1.3 Hermitian and unitary operators

Next, we will consider two special types of operators, namely Hermitian and unitary operators. An operator A is Hermitian if and only if $A^\dagger = A$.

Lemma: An operator is Hermitian if and only if it has real eigenvalues: $A^\dagger = A \Leftrightarrow a_j \in \mathbb{R}$.

Proof: The eigenvalue equation of A implies that

$$A |a_j\rangle = a_j |a_j\rangle \Rightarrow \langle a_j| A^\dagger = a_j^* \langle a_j|, \quad (1.27)$$

which means that $\langle a_j| A |a_j\rangle = a_j$ and $\langle a_j| A^\dagger |a_j\rangle = a_j^*$. It is now straightforward to show that $A = A^\dagger$ implies $a_j = a_j^*$, or $a_j \in \mathbb{R}$. Conversely, $a_j \in \mathbb{R}$ implies $a_j = a_j^*$, and

$$\langle a_j| A |a_j\rangle = \langle a_j| A^\dagger |a_j\rangle. \quad (1.28)$$

Let $|\psi\rangle = \sum_k c_k |a_k\rangle$. Then

$$\begin{aligned} \langle \psi| A |\psi\rangle &= \sum_j |c_j|^2 \langle a_j| A |a_j\rangle = \sum_j |c_j|^2 \langle a_j| A^\dagger |a_j\rangle = \sum_j |c_j|^2 \langle a_j| A^\dagger |a_j\rangle \\ &= \langle \psi| A^\dagger |\psi\rangle \end{aligned} \quad (1.29)$$

for all $|\psi\rangle$, and therefore $A = A^\dagger$. \square

Next, we construct the exponent of an operator A according to $U = \exp(icA)$. We have included the complex number c for completeness. At first sight, you may wonder what it means to take the exponent of an operator. Recall, however, that the exponent has a power expansion:

$$U = \exp(icA) = \sum_{n=0}^{\infty} \frac{(ic)^n}{n!} A^n. \quad (1.30)$$

The n^{th} power of an operator is straightforward: just multiply A n times with itself. The expression in Eq. (1.30) is then well defined, and the exponent is taken as an abbreviation of the power expansion. In general, we can construct any function of operators, as long as we can define the function in terms of a power expansion:

$$f(A) = \sum_{n=0}^{\infty} f_n A^n. \quad (1.31)$$

This can also be extended to functions of multiple operators, but now we have to be very careful in the case where these operators do not commute. Familiar rules for combining normal functions no longer apply (see exercise 4b).

We can construct the adjoint of the operator U according to

$$U^\dagger = \left(\sum_{n=0}^{\infty} \frac{(ic)^n}{n!} A^n \right)^\dagger = \sum_{n=0}^{\infty} \frac{(-ic^*)^n}{n!} A^{\dagger n} = \exp(-ic^* A^\dagger). \quad (1.32)$$

In the special case where $A = A^\dagger$ and c is real, we calculate

$$UU^\dagger = \exp(icA)\exp(-ic^* A^\dagger) = \exp(icA)\exp(-icA) = \exp[ic(A - A)] = \mathbb{I}, \quad (1.33)$$

since A commutes with itself. Similarly, $U^\dagger U = \mathbb{I}$. Therefore, $U^\dagger = U^{-1}$, and an operator with this property is called *unitary*. Each unitary operator can be generated by a Hermitian (self-adjoint) operator A and a real number c . A is called the *generator* of U . We often write $U = U_A(c)$. Unitary operators are basis transformations.

1.4 Projection operators and tensor products

We can combine two linear vector spaces \mathcal{U} and \mathcal{V} into a new linear vector space $\mathcal{W} = \mathcal{U} \oplus \mathcal{V}$. The symbol \oplus is called the *direct sum*. The dimension of \mathcal{W} is the sum of the dimensions of \mathcal{U} and \mathcal{V} :

$$\dim \mathcal{W} = \dim \mathcal{U} + \dim \mathcal{V}. \quad (1.34)$$

A vector in \mathcal{W} can be written as

$$|\Psi\rangle_{\mathcal{W}} = |\psi\rangle_{\mathcal{U}} + |\phi\rangle_{\mathcal{V}}, \quad (1.35)$$

where $|\psi\rangle_{\mathcal{U}}$ and $|\phi\rangle_{\mathcal{V}}$ are typically *not* normalized (i.e., they are not unit vectors). The spaces \mathcal{U} and \mathcal{V} are so-called *subspaces* of \mathcal{W} .

As an example, consider the three-dimensional Euclidean space spanned by the Cartesian axes x , y , and z . The xy -plane is a two-dimensional subspace of the full space, and the z -axis is a one-dimensional subspace. Any three-dimensional form can be *projected* onto the xy -plane by setting the z component to zero. Similarly, we can project onto the z -axis by setting the x and y coordinates to zero. A projector is therefore associated with a subspace. It acts on a vector in the full space, and forces all components to zero, except those of the subspace it projects onto.

The formal definition of a projector $P_{\mathcal{U}}$ on \mathcal{U} is given by

$$P_{\mathcal{U}} |\Psi\rangle_{\mathcal{W}} = |\psi\rangle_{\mathcal{U}}. \quad (1.36)$$

This is equivalent to requiring that $P_{\mathcal{U}}^2 = P_{\mathcal{U}}$, or $P_{\mathcal{U}}$ is *idempotent*. One-dimensional projectors can be written as

$$P_j = |\phi_j\rangle\langle\phi_j|. \quad (1.37)$$

Two projectors P_1 and P_2 are orthogonal is $P_1P_2 = 0$. If $P_1P_2 = 0$, then $P_1 + P_2$ is another projector:

$$(P_1 + P_2)^2 = P_1^2 + P_1P_2 + P_2P_1 + P_2^2 = P_1^2 + P_2^2 = P_1 + P_2. \quad (1.38)$$

When P_1 and P_2 commute but are non-orthogonal (i.e., they overlap), the general projector onto their combined subspace is

$$P_{1+2} = P_1 + P_2 - P_1P_2. \quad (1.39)$$

(Prove this.) The orthocomplement of P is $\mathbb{I} - P$, which is also a projector:

$$P(\mathbb{I} - P) = P - P^2 = P - P = 0 \quad \text{and} \quad (\mathbb{I} - P)^2 = \mathbb{I} - 2P + P^2 = \mathbb{I} - P. \quad (1.40)$$

Another way to combine two vector spaces \mathcal{U} and \mathcal{V} is via the tensor product: $\mathcal{W} = \mathcal{U} \otimes \mathcal{V}$, where the symbol \otimes is called the *direct product* or *tensor product*. The dimension of the space \mathcal{W} is then

$$\dim \mathcal{W} = \dim \mathcal{U} \cdot \dim \mathcal{V}. \quad (1.41)$$

Let $|\psi\rangle \in \mathcal{U}$ and $|\phi\rangle \in \mathcal{V}$. Then

$$|\psi\rangle \otimes |\phi\rangle \in \mathcal{W} = \mathcal{U} \otimes \mathcal{V}. \quad (1.42)$$

If $|\psi\rangle = \sum_j a_j |\psi_j\rangle$ and $|\phi\rangle = \sum_j b_j |\phi_j\rangle$, then the tensor product of these vectors can be written as

$$|\psi\rangle \otimes |\phi\rangle = \sum_{jk} a_j b_k |\psi_j\rangle \otimes |\phi_k\rangle = \sum_{jk} a_j b_k |\psi_j\rangle |\phi_k\rangle = \sum_{jk} a_j b_k |\psi_j, \phi_k\rangle, \quad (1.43)$$

where we introduced convenient abbreviations for the tensor product notation. The inner product of two vectors that are tensor products is

$$(\langle\psi_1| \otimes \langle\phi_1|)(|\psi_2\rangle \otimes |\phi_2\rangle) = \langle\psi_1|\psi_2\rangle \langle\phi_1|\phi_2\rangle. \quad (1.44)$$

Operators also obey the tensor product structure, with

$$(A \otimes B)|\psi\rangle \otimes |\phi\rangle = (A|\psi\rangle) \otimes (B|\phi\rangle), \quad (1.45)$$

and

$$(A \otimes B)(C \otimes D)|\psi\rangle \otimes |\phi\rangle = (AC|\psi\rangle) \otimes (BD|\phi\rangle). \quad (1.46)$$

General rules for tensor products of operators are

1. $A \otimes 0 = 0$ and $0 \otimes B = 0$,

2. $\mathbb{I} \otimes \mathbb{I} = \mathbb{I}$,
3. $(A_1 + A_2) \otimes B = A_1 \otimes B + A_2 \otimes B$,
4. $aA \otimes bB = (ab)A \otimes B$,
5. $(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$,
6. $(A \otimes B)^\dagger = A^\dagger \otimes B^\dagger$.

Note that the last rule preserves the order of the operators. In other words, operators always act on their own space. Often, it is understood implicitly which operator acts on which subspace, and we will write $A \otimes \mathbb{I} = A$ and $\mathbb{I} \otimes B = B$. Alternatively, we can add subscripts to the operator, e.g., $A_{\mathcal{U}}$ and $B_{\mathcal{V}}$.

As a practical example, consider two two-dimensional operators

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} \quad (1.47)$$

with respect to some orthonormal bases $\{|a_1\rangle, |a_2\rangle\}$ and $\{|b_1\rangle, |b_2\rangle\}$ for A and B , respectively (not necessarily eigenbases). The question is now: what is the matrix representation of $A \otimes B$? Since the dimension of the new vector space is the product of the dimensions of the two vector spaces, we have $\dim \mathcal{W} = 2 \cdot 2 = 4$. A natural basis for $A \otimes B$ is then given by $\{|a_j, b_k\rangle\}_{jk}$, with $j, k = 1, 2$, or

$$|a_1\rangle|b_1\rangle, \quad |a_1\rangle|b_2\rangle, \quad |a_2\rangle|b_1\rangle, \quad |a_2\rangle|b_2\rangle. \quad (1.48)$$

We can construct the matrix representation of $A \otimes B$ by applying this operator to the basis vectors in Eq. (1.48), using

$$A|a_j\rangle = A_{1j}|a_1\rangle + A_{2j}|a_2\rangle \quad \text{and} \quad B|a_k\rangle = B_{1k}|b_1\rangle + B_{2k}|b_2\rangle, \quad (1.49)$$

which leads to

$$\begin{aligned} A \otimes B |a_1, b_1\rangle &= (A_{11}|a_1\rangle + A_{21}|a_2\rangle)(B_{11}|b_1\rangle + B_{21}|b_2\rangle) \\ A \otimes B |a_1, b_2\rangle &= (A_{11}|a_1\rangle + A_{21}|a_2\rangle)(B_{12}|b_1\rangle + B_{22}|b_2\rangle) \\ A \otimes B |a_2, b_1\rangle &= (A_{12}|a_1\rangle + A_{22}|a_2\rangle)(B_{11}|b_1\rangle + B_{21}|b_2\rangle) \\ A \otimes B |a_2, b_2\rangle &= (A_{12}|a_1\rangle + A_{22}|a_2\rangle)(B_{12}|b_1\rangle + B_{22}|b_2\rangle) \end{aligned} \quad (1.50)$$

Looking at the first line of Eq. (1.50), the basis vector $|a_1, b_1\rangle$ gets mapped to all basis vectors:

$$A \otimes B |a_1, b_1\rangle = A_{11}B_{11}|a_1, b_1\rangle + A_{11}B_{21}|a_1, b_2\rangle + A_{21}B_{11}|a_2, b_1\rangle + A_{21}B_{21}|a_2, b_2\rangle. \quad (1.51)$$

Combining this into matrix form leads to

$$A \otimes B = \begin{pmatrix} A_{11}B_{11} & A_{11}B_{12} & A_{12}B_{11} & A_{12}B_{12} \\ A_{11}B_{21} & A_{11}B_{22} & A_{12}B_{21} & A_{12}B_{22} \\ A_{21}B_{11} & A_{21}B_{12} & A_{22}B_{11} & A_{22}B_{12} \\ A_{21}B_{21} & A_{21}B_{22} & A_{22}B_{21} & A_{22}B_{22} \end{pmatrix} = \begin{pmatrix} A_{11}B & A_{12}B \\ A_{21}B & A_{22}B \end{pmatrix}. \quad (1.52)$$

Recall that this is dependent on the basis that we have chosen. In particular, $A \otimes B$ may be diagonalized in some other basis.

1.5 The trace and determinant of an operator

There are two special functions of operators that play a key role in the theory of linear vector spaces. They are the *trace* and the *determinant* of an operator, denoted by $\text{Tr}(A)$ and $\det(A)$, respectively. While the trace and determinant are most conveniently evaluated in matrix representation, they are independent of the chosen basis.

When we defined the norm of an operator, we introduced the trace. It is evaluated by adding the diagonal elements of the matrix representation of the operator:

$$\text{Tr}(A) = \sum_j \langle \phi_j | A | \phi_j \rangle, \quad (1.53)$$

where $\{|\phi_j\rangle\}_j$ is *any* orthonormal basis. This independence means that the trace is an invariant property of the operator. Moreover, the trace has the following important properties:

1. If $A = A^\dagger$, then $\text{Tr}(A)$ is real,
2. $\text{Tr}(aA) = a\text{Tr}(A)$,
3. $\text{Tr}(A + B) = \text{Tr}(A) + \text{Tr}(B)$,
4. $\text{Tr}(AB) = \text{Tr}(BA)$ (the "cyclic property").

The first property follows immediately when we evaluate the trace in the diagonal basis, where it becomes a sum over real eigenvalues. The second and third properties convey the *linearity* of the trace. The fourth property is extremely useful, and can be shown as follows:

$$\begin{aligned} \text{Tr}(AB) &= \sum_j \langle \phi_j | AB | \phi_j \rangle = \sum_{jk} \langle \phi_j | A | \psi_k \rangle \langle \psi_k | B | \phi_j \rangle \\ &= \sum_{jk} \langle \psi_k | B | \phi_j \rangle \langle \phi_j | A | \psi_k \rangle = \sum_k \langle \psi_k | BA | \psi_k \rangle \\ &= \text{Tr}(BA). \end{aligned} \quad (1.54)$$

This derivation also demonstrates the usefulness of inserting a resolution of the identity in strategic places. In the cyclic property, the operators A and B may be products of two operators, which then leads to

$$\text{Tr}(ABC) = \text{Tr}(BCA) = \text{Tr}(CAB). \quad (1.55)$$

Any cyclic (even) permutation of operators under a trace gives rise to the same value of the trace as the original operator ordering.

Finally, we construct the *partial* trace of an operator that lives on a tensor product space. Suppose that $A \otimes B$ is an operator in the Hilbert space $\mathcal{H}_1 \otimes \mathcal{H}_2$. We can trace out Hilbert space \mathcal{H}_1 , denoted by $\text{Tr}_1(\cdot)$:

$$\text{Tr}_1(A \otimes B) = \text{Tr}(A)B, \quad \text{or equivalently} \quad \text{Tr}_1(A_1 B_2) = \text{Tr}(A_1)B_2. \quad (1.56)$$

Taking the partial trace has the effect of removing the entire Hilbert space \mathcal{H}_1 from the description. It reduces the total vector space. The partial trace always carries an index, which determines which space is traced over.

The determinant of a 2×2 matrix is given by

$$\det(A) = \det \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} = A_{11}A_{22} - A_{12}A_{21}. \quad (1.57)$$

The determinant of higher-dimensional matrices can be defined recursively as follows: The top-left element of an $n \times n$ matrix defines an $(n-1) \times (n-1)$ matrix by removing the top row and the left column. Similarly, any other element in the left column defines an $(n-1) \times (n-1)$ matrix by removing the left column and the row of the element we chose. The determinant of the $n \times n$ matrix is then given by the top-left element times the determinant of the remaining $(n-1) \times (n-1)$ matrix, *minus* the product of the second element down in the left column and the remaining $(n-1) \times (n-1)$ matrix, *plus* the third element times the remaining matrix, etc.

The determinant of the product of matrices is equal to the product of the determinants of the matrices:

$$\det(AB) = \det(A)\det(B). \quad (1.58)$$

Moreover, if A is an invertible matrix, then we have

$$\det(A^{-1}) = \det(A)^{-1}. \quad (1.59)$$

This leads to an important relation between similar matrices $A = X^{-1}BX$:

$$\begin{aligned} \det(A) &= \det(X^{-1}BX) = \det(X^{-1})\det(B)\det(X) \\ &= \det(X)^{-1}\det(B)\det(X) = \det(B). \end{aligned} \quad (1.60)$$

In particular, this means that the determinant is independent of the basis in which the matrix is written, which means that it is an intrinsic property of the operator associated with that matrix.

Finally, here's a fun relation between the trace and the determinant of an operator:

$$\det[\exp(A)] = \exp[\text{Tr}(A)]. \quad (1.61)$$

Exercises

1. Vectors and matrices:

- Are the following three vectors linearly dependent or independent: $a = (2, 3, -1)$, $b = (0, 1, 2)$, and $c = (0, 0, -5)$?
- Consider the vectors $|\psi\rangle = 3i|\phi_1\rangle - 7i|\phi_2\rangle$ and $|\chi\rangle = |\phi_1\rangle + 2|\phi_2\rangle$, with $\{|\phi_i\rangle\}$ an orthonormal basis. Calculate the inner product between $|\psi\rangle$ and $|\chi\rangle$, and show that they satisfy the Cauchy-Schwarz inequality.
- Consider the two matrices

$$A = \begin{pmatrix} 0 & i & 2 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 2 & i & 0 \\ 3 & 1 & 5 \\ 0 & -i & -2 \end{pmatrix}. \quad (1.62)$$

Calculate $A^{-1}B$ and BA^{-1} . Are they equal?

- Calculate $A \otimes B$ and $B \otimes A$, where $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ and $B = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.

2. Operators:

- Which of these operators are Hermitian: $A + A^\dagger$, $i(A + A^\dagger)$, $i(A - A^\dagger)$, and $A^\dagger A$?
- Prove that a shared eigenbasis for two operators A and B implies that $[A, B] = 0$.

- (c) Let U be a transformation matrix that maps one complete orthonormal basis to another. Show that U is unitary.
 - (d) How many real parameters completely determine a $d \times d$ unitary matrix?
3. Properties of the trace and the determinant:
- (a) Calculate the trace and the determinant of the matrices A and B in exercise 1c.
 - (b) Show that the expectation value of A can be written as $\text{Tr}(|\psi\rangle\langle\psi|A)$.
 - (c) Prove that the trace is independent of the basis.
4. Commutator identities.
- (a) Let $F(t) = e^{At}e^{Bt}$. Calculate dF/dt and use $[e^{At}, B] = (e^{At}Be^{-At} - B)e^{At}$ to simplify your result.
 - (b) Let $G(t) = e^{At+Bt+f(t)H}$. Show by calculating dG/dt , and setting $dF/dt = dG/dt$ at $t = 1$, that the following operator identity

$$e^A e^B = e^{A+B+\frac{1}{2}[A,B]}, \quad (1.63)$$

holds if A and B both commute with $[A, B]$. Hint: use the Hadamard lemma

$$e^{At} B e^{-At} = B + \frac{t}{1!}[A, B] + \frac{t^2}{2!}[A, [A, B]] + \dots \quad (1.64)$$

- (c) Show that the commutator of two Hermitian operators is anti-Hermitian ($A^\dagger = -A$).
- (d) Prove the commutator analog of the Jacobi identity

$$[A, [B, C]] + [B, [C, A]] + [C, [A, B]] = 0. \quad (1.65)$$

2 The Postulates of Quantum Mechanics

The entire structure of quantum mechanics (including its relativistic extension) can be formulated in terms of states and operations in Hilbert space. We need rules that map the physical quantities such as states, observables, and measurements to the mathematical structure of vector spaces, vectors and operators. There are several ways in which this can be done, and here we summarize these rules in terms of five postulates.

Postulate 1: A physical system is described by a Hilbert space \mathcal{H} , and the *state* of the system is represented by a *ray* with norm 1 in \mathcal{H} .

There are a number of important aspects to this postulate. First, the fact that states are rays, rather than vectors means that an overall phase $e^{i\varphi}$ of the state does not have any physically observable consequences, and $e^{i\varphi}|\psi\rangle$ represents the same state as $|\psi\rangle$. Second, the state contains *all* information about the system. In particular, there are no hidden variables in this standard formulation of quantum mechanics. Finally, the dimension of \mathcal{H} may be infinite, which is the case, for example, when \mathcal{H} is the space of square-integrable functions.

As an example of this postulate, consider a two-level quantum system (a *qubit*). This system can be described by two orthonormal states $|0\rangle$ and $|1\rangle$. Due to linearity of Hilbert space, the superposition $\alpha|0\rangle + \beta|1\rangle$ is again a state of the system if it has norm 1, or

$$(\alpha^* \langle 0| + \beta^* \langle 1|)(\alpha|0\rangle + \beta|1\rangle) = 1 \quad \text{or} \quad |\alpha|^2 + |\beta|^2 = 1. \quad (2.1)$$

This is called the *superposition principle*: any normalised superposition of valid quantum states is again a valid quantum state. It is a direct consequence of the linearity of the vector space, and as we shall see later, this principle has some bizarre consequences that have been corroborated in many experiments.

Postulate 2: Every physical observable A corresponds to a self-adjoint (Hermitian¹) operator \hat{A} whose eigenvectors form a complete basis.

We use a hat to distinguish between the observable and the operator, but usually this distinction is not necessary. In these notes, we will use hats only when there is a danger of confusion.

As an example, take the operator X :

$$X|0\rangle = |1\rangle \quad \text{and} \quad X|1\rangle = |0\rangle. \quad (2.2)$$

This operator can be interpreted as a bit flip of a qubit. In matrix notation the state vectors can be written as

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad (2.3)$$

which means that X is written as

$$X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad (2.4)$$

with eigenvalues ± 1 . The eigenstates of X are

$$|\pm\rangle = \frac{|0\rangle \pm |1\rangle}{\sqrt{2}}. \quad (2.5)$$

These states form an orthonormal basis.

¹In Hilbert spaces of infinite dimensionality, there are subtle differences between self-adjoint and Hermitian operators. We ignore these subtleties here, because we will be mostly dealing with finite-dimensional spaces.

Postulate 3: The eigenvalues of A are the possible measurement outcomes, and the probability of finding the outcome a_j in a measurement is given by the *Born rule*:

$$p(a_j) = |\langle a_j | \psi \rangle|^2, \quad (2.6)$$

where $|\psi\rangle$ is the state of the system, and $|a_j\rangle$ is the eigenvector associated with the eigenvalue a_j via $A|a_j\rangle = a_j|a_j\rangle$. If a_j is m -fold degenerate, then

$$p(a_j) = \sum_{l=1}^m |\langle a_j^{(l)} | \psi \rangle|^2, \quad (2.7)$$

where the $|a_j^{(l)}\rangle$ span the m -fold degenerate subspace.

The expectation value of A with respect to the state of the system $|\psi\rangle$ is denoted by $\langle A \rangle$, and evaluated as

$$\langle A \rangle = \langle \psi | A | \psi \rangle = \langle \psi | \left(\sum_j a_j |a_j\rangle \langle a_j| \right) | \psi \rangle = \sum_j p(a_j) a_j. \quad (2.8)$$

This is the weighted average of the measurement outcomes. The spread of the measurement outcomes (or the *uncertainty*) is given by the variance

$$(\Delta A)^2 = \langle (A - \langle A \rangle)^2 \rangle = \langle A^2 \rangle - \langle A \rangle^2. \quad (2.9)$$

So far we mainly dealt with discrete systems on finite-dimensional Hilbert spaces. But what about continuous systems, such as a particle in a box, or a harmonic oscillator? We can still write the spectral decomposition of an operator A but the sum must be replaced by an integral:

$$A = \int da f_A(a) |a\rangle \langle a|, \quad (2.10)$$

where $|a\rangle$ is an eigenstate of A . Typically, there are problems with the normalization of $|a\rangle$, which is related to the impossibility of preparing a system in *exactly* the state $|a\rangle$. We will not explore these subtleties further in this course, but you should be aware that they exist. The expectation value of A is

$$\langle A \rangle = \langle \psi | A | \psi \rangle = \int da f_A(a) \langle \psi | a \rangle \langle a | \psi \rangle \equiv \int da f_A(a) |\psi(a)|^2, \quad (2.11)$$

where we defined the *wave function* $\psi(a) = \langle a | \psi \rangle$, and $|\psi(a)|^2$ is properly interpreted as the probability *density* that you remember from second-year quantum mechanics.

The probability of finding the eigenvalue of an operator A in the interval a and $a + da$ given the state $|\psi\rangle$ is

$$\langle \psi | (|a\rangle \langle a| da) | \psi \rangle \equiv dp(a), \quad (2.12)$$

since both sides must be infinitesimal. We therefore find that

$$\frac{dp(a)}{da} = |\psi(a)|^2. \quad (2.13)$$

Postulate 4: The dynamics of quantum systems is governed by unitary transformations.

We can write the state of a system at time t as $|\psi(t)\rangle$, and at some time $t_0 < t$ as $|\psi(t_0)\rangle$. The fourth postulate tells us that there is a unitary operator $U(t, t_0)$ that transforms the state at time t_0 to the state at time t :

$$|\psi(t)\rangle = U(t, t_0) |\psi(t_0)\rangle. \quad (2.14)$$

Since the evolution from time t to t is denoted by $U(t, t)$ and must be equal to the identity, we deduce that U depends only on time differences: $U(t, t_0) = U(t - t_0)$, and $U(0) = \mathbb{I}$.

As an example, let $U(t)$ be generated by a Hermitian operator A according to

$$U(t) = \exp\left(-\frac{i}{\hbar} A t\right). \quad (2.15)$$

The argument of the exponential must be dimensionless, so A must be proportional to \hbar times an angular frequency (in other words, an energy). Suppose that $|\psi(t)\rangle$ is the state of a qubit, and that $A = \hbar\omega X$. If $|\psi(0)\rangle = |0\rangle$ we want to calculate the state of the system at time t . We can write

$$|\psi(t)\rangle = U(t) |\psi(0)\rangle = \exp(-i\omega t X) |0\rangle = \sum_{n=0}^{\infty} \frac{(-i\omega t)^n}{n!} X^n. \quad (2.16)$$

Observe that $X^2 = \mathbb{I}$, so we can separate the power series into even and odd values of n :

$$|\psi(t)\rangle = \sum_{n=0}^{\infty} \frac{(-i\omega t)^{2n}}{(2n)!} |0\rangle + \sum_{n=0}^{\infty} \frac{(-i\omega t)^{2n+1}}{(2n+1)!} X |0\rangle = \cos(\omega t) |0\rangle - i \sin(\omega t) |1\rangle. \quad (2.17)$$

In other words, the state oscillates between $|0\rangle$ and $|1\rangle$.

The fourth postulate also leads to the Schrödinger equation. Let's take the infinitesimal form of Eq. (2.14):

$$|\psi(t + dt)\rangle = U(dt) |\psi(t)\rangle. \quad (2.18)$$

We require that $U(dt)$ is generated by some Hermitian operator H :

$$U(dt) = \exp\left(-\frac{i}{\hbar} H dt\right). \quad (2.19)$$

H must have the dimensions of energy, so we identify it with the energy operator, or the *Hamiltonian*. We can now take a Taylor expansion of $|\psi(t + dt)\rangle$ to first order in dt :

$$|\psi(t + dt)\rangle = |\psi(t)\rangle + dt \frac{d}{dt} |\psi(t)\rangle + \dots, \quad (2.20)$$

and we expand the unitary operator to first order in dt as well:

$$U(dt) = 1 - \frac{i}{\hbar} H dt + \dots \quad (2.21)$$

We combine this into

$$|\psi(t)\rangle + dt \frac{d}{dt} |\psi(t)\rangle = \left(1 - \frac{i}{\hbar} H dt\right) |\psi(t)\rangle, \quad (2.22)$$

which can be recast into the Schrödinger equation:

$$i\hbar \frac{d}{dt} |\psi(t)\rangle = H |\psi(t)\rangle. \quad (2.23)$$

Therefore, the Schrödinger equation follows directly from the postulates!

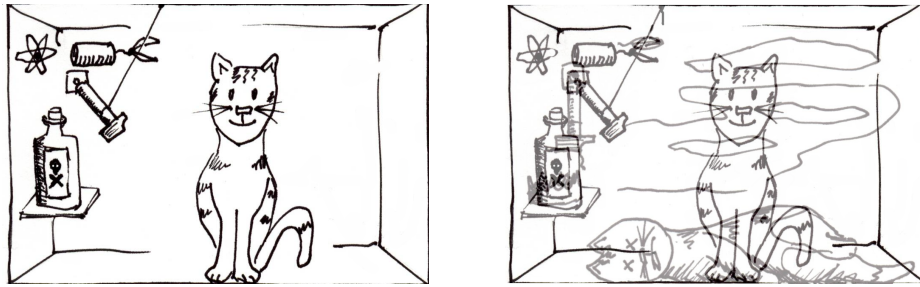


Figure 1: Schrödinger's Cat.

Postulate 5: If a measurement of an observable A yields an eigenvalue a_j , then immediately after the measurement, the system is in the eigenstate $|a_j\rangle$ corresponding to the eigenvalue.

This is the infamous *projection postulate*, so named because a measurement “projects” the system to the eigenstate corresponding to the measured value. This postulate has as observable consequence that a second measurement immediately after the first will also find the outcome a_j . Each measurement outcome a_j corresponds to a projection operator P_j on the subspace spanned by the eigenvector(s) belonging to a_j . A (perfect) measurement can be described by applying a projector to the state, and renormalize:

$$|\psi\rangle \rightarrow \frac{P_j |\psi\rangle}{\|P_j |\psi\rangle\|}. \quad (2.24)$$

This also works for degenerate eigenvalues.

We have established earlier that the expectation value of A can be written as a trace:

$$\langle A \rangle = \text{Tr}(|\psi\rangle\langle\psi| A). \quad (2.25)$$

Now instead of the full operator A , we calculate the trace of $P_j = |a_j\rangle\langle a_j|$:

$$\langle P_j \rangle = \text{Tr}(|\psi\rangle\langle\psi| P_j) = \text{Tr}(|\psi\rangle\langle\psi| a_j \langle a_j|) = |\langle a_j | \psi \rangle|^2 = p(a_j). \quad (2.26)$$

So we can calculate the probability of a measurement outcome by taking the expectation value of the projection operator that corresponds to the eigenstate of the measurement outcome. This is one of the basic calculations in quantum mechanics that you should be able to do.

The projection postulate is somewhat problematic for the interpretation of quantum mechanics, because it leads to the so-called *measurement problem*: Why does a measurement induce a non-unitary evolution of the system? After all, the measurement apparatus can also be described quantum mechanically² and then the system plus the measurement apparatus evolves unitarily. But then we must invoke a new device that measures the combined system and measurement apparatus. However, this in turn can be described quantum mechanically, and so on.

On the other hand, we do see definite measurement outcomes when we do experiments, so at some level the projection postulate is necessary, and somewhere there must be a “collapse of the wave function”. Schrödinger already struggled with this question, and came up with his famous

²This is something most people require from a fundamental theory: quantum mechanics should not just break down for macroscopic objects. Indeed, experimental evidence of macroscopic superpositions has been found in the form of “cat states”.

thought experiment about a cat in a box with a poison-filled vial attached to a Geiger counter monitoring a radioactive atom (see Fig. 1). When the atom decays, it will trigger the Geiger counter, which in turn causes the release of the poison killing the cat. When we do not look inside the box (more precisely: when no information about the atom-counter-vial-cat system escapes from the box), the entire system is in a quantum superposition. However, when we open the box, we do find the cat either dead or alive. One solution of the problem seems to be that the quantum state represents our knowledge of the system, and that looking inside the box merely updates our information about the atom, counter, vial and the cat. So nothing “collapses” except our own state of mind.

However, this cannot be the entire story, because quantum mechanics clearly is not just about our opinions of cats and decaying atoms. In particular, if we prepare an electron in a spin “up” state $|\uparrow\rangle$, then whenever we measure the spin along the z -direction we will find the measurement outcome “up”, no matter what we think about electrons and quantum mechanics. So there seems to be some physical property associated with the electron that determines the measurement outcome and is described by the quantum state.

Various interpretations of quantum mechanics attempt to address these (and other) issues. The original interpretation of quantum mechanics was mainly put forward by Niels Bohr, and is called the Copenhagen interpretation. Broadly speaking, it says that the quantum state is a convenient fiction, used to calculate the results of measurement outcomes, and that the system cannot be considered separate from the measurement apparatus. Alternatively, there are interpretations of quantum mechanics, such as the Ghirardi-Rimini-Weber interpretation, that do ascribe some kind of reality to the state of the system, in which case a physical mechanism for the collapse of the wave function must be given. Many of these interpretations can be classified as *hidden variable* theories, which postulate that there is a deeper physical reality described by some “hidden variables” that we must average over. This in turn explains the probabilistic nature of quantum mechanics. The problem with such theories is that these hidden variables must be quite weird: they can change instantly depending on events lightyears away³, thus violating Einstein’s theory of special relativity. Many physicists do not like this aspect of hidden variable theories.

Alternatively, quantum mechanics can be interpreted in terms of “many worlds”: the Many Worlds interpretation states that there is one state vector for the entire universe, and that each measurement splits the universe into different branches corresponding to the different measurement outcomes. It is attractive since it seems to be a philosophically *consistent* interpretation, and while it has been acquiring a growing number of supporters over recent years⁴, a lot of physicists have a deep aversion to the idea of parallel universes.

Finally, there is the epistemic interpretation, which is very similar to the Copenhagen interpretation in that it treats the quantum state to a large extent as a measure of our knowledge of the quantum system (and the measurement apparatus). At the same time, it *denies a deeper underlying reality* (i.e., no hidden variables). The attractive feature of this interpretation is that it requires a minimal amount of fuss, and fits naturally with current research in quantum information theory. The downside is that you have to abandon simple scientific realism that allows you to talk about the properties of electrons and photons, and many physicists are not prepared to do that.

As you can see, quantum mechanics forces us to abandon some deeply held (classical) convictions about Nature. Depending on your preference, you may be drawn to one or other interpreta-

³... even though the averaging over the hidden variables means you can never signal faster than light.

⁴There seems to be some evidence that the Many Worlds interpretation fits well with the latest cosmological models based on string theory.

tion. It is currently not known which interpretation is the correct one.

Exercises

1. Calculate the eigenvalues and the eigenstates of the bit flip operator X , and show that the eigenstates form an orthonormal basis. Calculate the expectation value of X for $|\psi\rangle = 1/\sqrt{3}|0\rangle + i\sqrt{2/3}|1\rangle$.
2. Show that the variance of A vanishes when $|\psi\rangle$ is an eigenstate of A .
3. Prove that an operator is Hermitian if and only if it has real eigenvalues.
4. Show that a qubit in an unknown state $|\psi\rangle$ cannot be copied. This is the *no-cloning theorem*. Hint: start with a state $|\psi\rangle|i\rangle$ for some initial state $|i\rangle$, and require that for $|\psi\rangle = |0\rangle$ and $|\psi\rangle = |1\rangle$ the cloning procedure is a unitary transformation $|0\rangle|i\rangle \rightarrow |0\rangle|0\rangle$ and $|1\rangle|i\rangle \rightarrow |1\rangle|1\rangle$.
5. The uncertainty principle.

- (a) Use the Cauchy-Schwarz inequality to derive the following relation between non-commuting observables A and B :

$$(\Delta A)^2(\Delta B)^2 \geq \frac{1}{4} |\langle[A, B]\rangle|^2. \quad (2.27)$$

Hint: define $|f\rangle = (A - \langle A \rangle)|\psi\rangle$ and $|g\rangle = i(B - \langle B \rangle)|\psi\rangle$, and use that $|\langle f|g\rangle| \geq \frac{1}{2}|\langle f|g\rangle + \langle g|f\rangle|$.

- (b) Show that this reduces to Heisenberg's uncertainty relation when A and B are canonically conjugate observables, for example position and momentum.
- (c) Does this method work for deriving the uncertainty principle between energy and time?
6. Consider the Hamiltonian H and the state $|\psi\rangle$ given by

$$H = E \begin{pmatrix} 0 & i & 0 \\ -i & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} \quad \text{and} \quad |\psi\rangle = \frac{1}{\sqrt{5}} \begin{pmatrix} 1-i \\ 1-i \\ 1 \end{pmatrix}. \quad (2.28)$$

where E is a constant with dimensions of energy. Calculate the energy eigenvalues and the expectation value of the Hamiltonian.

7. Show that the momentum and the total energy can be measured simultaneously only when the potential is constant everywhere. What does a constant potential mean in terms of the dynamics of a particle?

3 Schrödinger and Heisenberg Pictures

So far we have assumed that the quantum states $|\psi(t)\rangle$ describing the system carry the time dependence. However, this is not the only way to keep track of the time evolution. Since all physically observed quantities are expectation values, we can write

$$\begin{aligned}\langle A \rangle &= \text{Tr} [|\psi(t)\rangle \langle \psi(t)| A] = \text{Tr} [U(t) |\psi(0)\rangle \langle \psi(0)| U^\dagger(t) A] \\ &= \text{Tr} [|\psi(0)\rangle \langle \psi(0)| U^\dagger(t) A U(t)] \quad (\text{cyclic property}) \\ &\equiv \text{Tr} [|\psi(0)\rangle \langle \psi(0)| A(t)],\end{aligned}\tag{3.1}$$

where we defined the time-varying operator $A(t) = U^\dagger(t) A U(t)$. Clearly, we can keep track of the time evolution in the *operators*!

- Schrödinger picture: Keep track of the time evolution in the states,
- Heisenberg picture: Keep track of the time evolution in the operators.

We can label the states and operators “S” and “H” depending on the picture. For example,

$$|\psi_H\rangle = |\psi_S(0)\rangle \quad \text{and} \quad A_H(t) = U^\dagger(t) A_S U(t).\tag{3.2}$$

The time evolution for states is given by the Schrödinger equation, so we want a corresponding “Heisenberg equation” for the operators. First, we observe that

$$U(t) = \exp\left(-\frac{i}{\hbar} H t\right),\tag{3.3}$$

such that

$$\frac{d}{dt} U(t) = -\frac{i}{\hbar} H U(t).\tag{3.4}$$

Next, we calculate the time derivative of $\langle A \rangle$:

$$\frac{d}{dt} \text{Tr} [|\psi_S(t)\rangle \langle \psi_S(t)| A_S] = \frac{d}{dt} \langle \psi_S(t) | A_S | \psi_S(t) \rangle = \frac{d}{dt} \langle \psi_H | A_H(t) | \psi_H \rangle.\tag{3.5}$$

The last equation follows from Eq. (3.1). We can now calculate

$$\begin{aligned}\frac{d}{dt} \langle \psi_S(t) | A_S | \psi_S(t) \rangle &= \frac{d}{dt} \langle \psi_S(0) | U^\dagger(t) A_S U(t) | \psi_S(0) \rangle \\ &= \langle \psi_S(0) | \left[\dot{U}^\dagger(t) A_S U(t) + U^\dagger(t) \dot{A}_S U(t) + U^\dagger(t) A_S \dot{U}(t) \right] | \psi_S(0) \rangle \\ &= \langle \psi_H | \left[\frac{i}{\hbar} H A_H(t) - \frac{i}{\hbar} A_H(t) H + \frac{\partial A_H(t)}{\partial t} \right] | \psi_H \rangle \\ &= -\frac{i}{\hbar} \langle \psi_H | [A_H(t), H] | \psi_H \rangle + \langle \psi_H | \frac{\partial A_H(t)}{\partial t} | \psi_H \rangle \\ &= \langle \psi_H | \frac{dA_H(t)}{dt} | \psi_H \rangle.\end{aligned}\tag{3.6}$$

Since this must be true for all $|\psi_H\rangle$, this is an operator identity:

$$\frac{dA_H(t)}{dt} = -\frac{i}{\hbar} [A_H(t), H] + \frac{\partial A_H(t)}{\partial t}.\tag{3.7}$$

This is the Heisenberg equation. Note the difference between the “straight d ” and the “curly ∂ ” in the time derivative and the partial time derivative, respectively. The partial derivative deals only with the *explicit* time dependence of the operator. In many cases (such as position and momentum) this is zero.

We have seen that both the Schrödinger and the Heisenberg equation follows from Von Neumann’s Hilbert space formalism of quantum mechanics. Consequently, we have proved that this formalism properly unifies both Schrödingers wave mechanics, and Heisenberg, Born, and Jordans matrix mechanics.

As an example, consider a qubit with time evolution determined by the Hamiltonian $H = \frac{1}{2}\hbar\omega Z$, with $Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. This may be a spin in a magnetic field, for example, such that $\omega = -eB/mc$. We want to calculate the time evolution of the operator $X_H(t)$. Since we work in the Heisenberg picture alone, we will omit the subscript H . First, we evaluate the commutator in the Heisenberg equation

$$i\hbar \frac{1}{2} \frac{dX}{dt} = \frac{1}{2} [X, H] = -i\hbar \frac{\omega}{2} Y, \quad (3.8)$$

where we defined $Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$. So now we must know the time evolution of Y as well:

$$i\hbar \frac{1}{2} \frac{dY}{dt} = \frac{1}{2} [Y, H] = i\hbar \frac{\omega}{2} X. \quad (3.9)$$

These are two coupled linear equations, which are relatively easy to solve:

$$\dot{X} = -\omega Y \quad \text{and} \quad \dot{Y} = \omega X \quad \text{and} \quad \dot{Z} = 0. \quad (3.10)$$

We can define two new operators $S_{\pm} = X \pm iY$, and obtain

$$\dot{S}_{\pm} = -\omega Y \pm i\omega X = \pm i\omega S_{\pm}. \quad (3.11)$$

Solving these two equations yields $S_{\pm}(t) = S_{\pm}(0)e^{\pm i\omega t}$, and this leads to

$$\begin{aligned} X(t) &= \frac{S_+(t) + S_-(t)}{2} = \frac{S_+(0)e^{i\omega t} + S_-(0)e^{-i\omega t}}{2} \\ &= \frac{1}{2} \left[X(0)e^{i\omega t} + iY(0)e^{i\omega t} + X(0)e^{-i\omega t} - iY(0)e^{-i\omega t} \right] \\ &= X(0)\cos(\omega t) - Y(0)\sin(\omega t). \end{aligned} \quad (3.12)$$

You are asked to show that $Y(t) = Y(0)\cos(\omega t) + X(0)\sin(\omega t)$ in exercise 3.

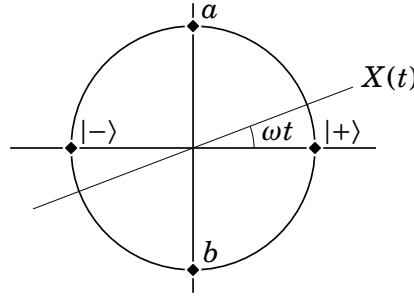
We now take $|\psi_H\rangle = |0\rangle$ and $X(0) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, $Y(0) = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$. The expectation value of $X(t)$ is then readily calculated to be

$$\langle 0|X(t)|0\rangle = \cos(\omega t) \langle 0|X(0)|0\rangle - \sin(\omega t) \langle 0|Y(0)|0\rangle = 0. \quad (3.13)$$

Alternatively, when $|\psi_H\rangle = |\pm\rangle$, we find

$$\langle +|X(t)|+\rangle = \cos(\omega t) \quad \text{and} \quad \langle +|Y(t)|+\rangle = \sin(\omega t). \quad (3.14)$$

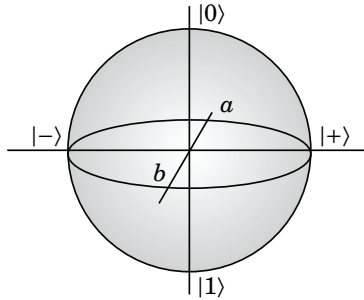
This is a circular motion in time:



The eigenstate of $X(\pi/2)$ is point a , and the eigenstate of $X(-\pi/2)$ is point b . Furthermore, $X(\pm\pi/2) = \mp Y(0)$, and the states at point a and b are therefore the eigenstates of Y :

$$|\psi_a\rangle = \frac{|0\rangle - i|1\rangle}{\sqrt{2}} \quad \text{and} \quad |\psi_b\rangle = \frac{|0\rangle + i|1\rangle}{\sqrt{2}}. \quad (3.15)$$

A natural question to ask is where the states $|0\rangle$ and $|1\rangle$ fit in this picture. These are the eigenstates of the operator Z , which we used to generate the unitary time evolution. Clearly the states on the circle never become either $|0\rangle$ or $|1\rangle$, so we need to add another dimension:



This is called the Bloch sphere, and operators are represented by straight lines through the origin. The axis of rotation for the straight lines that rotate with time is determined by the Hamiltonian. In the above case the Hamiltonian was proportional to Z , which means that the straight lines rotate around the axis through the eigensates of Z , which are $|0\rangle$ and $|1\rangle$.

Exercises

1. Show that for the Hamiltonian $H_S = H_H$.
2. The harmonic oscillator has the energy eigenvalue equation $H|n\rangle = \hbar\omega(n + \frac{1}{2})|n\rangle$.
 - (a) The classical solution of the harmonic oscillator is given by

$$|\alpha\rangle = e^{-\frac{1}{2}|\alpha|^2} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} |n\rangle, \quad (3.16)$$

in the limit of $|\alpha| \gg 1$. Show that $|\alpha\rangle$ is a properly normalized state for any $\alpha \in \mathbb{C}$.

- (b) Calculate the time-evolved state $|\alpha(t)\rangle$.
- (c) We introduce the ladder operators $\hat{a}|n\rangle = \sqrt{n}|n-1\rangle$ and $\hat{a}^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle$. Show that the number operator defined by $\hat{n}|n\rangle = n|n\rangle$ can be written as $\hat{n} = \hat{a}^\dagger \hat{a}$.

- (d) Write the coherent state $|\alpha\rangle$ as a superposition of ladder operators acting on the ground state $|0\rangle$.
- (e) Note that the ground state is time-independent ($U(t)|0\rangle = |0\rangle$). Calculate the time evolution of the ladder operators.
- (f) Calculate the position $\hat{q} = (\hat{a} + \hat{a}^\dagger)/2$ and momentum $\hat{p} = -i(\hat{a} - \hat{a}^\dagger)/2$ of the harmonic oscillator in the Heisenberg picture. Can you identify the classical harmonic motion?
3. Let A be an operator given by $A = a_0\mathbb{I} + a_x X + a_y Y + a_z Z$. Calculate the matrix $A(t)$ given the Hamiltonian $H = \frac{1}{2}\hbar\omega Z$, and show that A is Hermitian when the a_μ are real.
4. The interaction picture.
- (a) Let the Hamiltonian of a system be given by $H = H_0 + V$, with $H_0 = p^2/2m$. Using $|\psi(t)\rangle_I = U_0^\dagger(t)|\psi(t)\rangle_S$ with $U_0(t) = \exp(-iH_0 t/\hbar)$, calculate the time dependence of an operator in the interaction picture $A_I(t)$.
- (b) Defining $H_I(t) = U_0^\dagger(t)VU_0(t)$, show that

$$i\hbar \frac{d}{dt} |\psi(t)\rangle_I = H_I(t) |\psi(t)\rangle_I. \quad (3.17)$$

Is H_I identical to H_H and H_S ?

5. The time operator in quantum mechanics.

- (a) Let $H|\psi\rangle = E|\psi\rangle$, and assume the existence of a *time operator* conjugate to H , i.e., $[H, T] = i\hbar$. Show that

$$H e^{i\omega T} |\psi\rangle = (E - \hbar\omega) e^{i\omega T} |\psi\rangle. \quad (3.18)$$

- (b) Given that $\omega \in \mathbb{R}$, calculate the spectrum of H .
- (c) The energy of a system must be bounded from below in order to avoid infinite decay to ever lower energy states. What does this mean for T ?
6. Consider a three-level atom with two (degenerate) low-lying states $|0\rangle$ and $|1\rangle$ with zero energy, and a high level $|e\rangle$ (the “excited” state) with energy $\hbar\omega$. The low levels are coupled to the excited level by optical fields $\Omega_0 \cos \omega_0 t$ and $\Omega_1 \cos \omega_1 t$, respectively.

- (a) Give the (time-dependent) Hamiltonian H for the system.
- (b) The time dependence in H is difficult to deal with, so we must transform to the rotating frame via some unitary transformation $U(t)$. Show that

$$H' = U(t) H U^\dagger(t) - i\hbar U \frac{dU^\dagger}{dt}.$$

You can use the Schrödinger equation with $|\psi\rangle = U^\dagger |\psi'\rangle$.

- (c) Calculate H' if $U(t)$ is given by

$$U(t) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & e^{-i(\omega_0 - \omega_1)t} & 0 \\ 0 & 0 & e^{-i\omega_0 t} \end{pmatrix}.$$

Why can we ignore the remaining time dependence in H' ? This is called the *Rotating Wave Approximation*.

- (d) Calculate the $\lambda = 0$ eigenstate of H' in the case where $\omega_0 = \omega_1$.
- (e) Design a way to bring the atom from the state $|0\rangle$ to $|1\rangle$ without ever populating the state $|e\rangle$.

4 Mixed States and the Density Operator

So far we have considered states as vectors in Hilbert space that, according to the first postulate, contain all the information about the system. In reality, however, we very rarely have complete information about a system. For example, the system may have interacted with its environment, which introduces some uncertainty in our knowledge of the state of the system. The question is therefore how we describe systems given *incomplete* information. Much of contemporary research in quantum mechanics is about gaining full control over the quantum system (meaning to minimize the interaction with the environment). This includes the field of quantum information and computation. The concept of incomplete information is therefore central to modern quantum mechanics.

4.1 Mixed states

First, we recall some properties of the trace:

- $\text{Tr}(aA) = a\text{Tr}(A)$,
- $\text{Tr}(A + B) = \text{Tr}(A) + \text{Tr}(B)$.

Also remember that we can write the expectation value of A as

$$\langle A \rangle = \text{Tr}(|\psi\rangle\langle\psi|A), \quad (4.1)$$

where $|\psi\rangle$ is the state of the system. It tells us everything there is to know about the system. But what if we don't know everything?

As an example, consider that Alice prepares a qubit in the state $|0\rangle$ or in the state $|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$ depending on the outcome of a balanced (50:50) coin toss. How does Bob describe the state before any measurement? First, we *cannot* say that the state is $\frac{1}{2}|0\rangle + \frac{1}{2}|+\rangle$, because this is not normalized!

The key to the solution is to observe that the *expectation values* must behave correctly. The expectation value $\langle A \rangle$ is the average of the eigenvalues of A for a given state. If the state is itself a statistical mixture (as in the example above), then the expectation values must also be averaged. So for the example, we require that for any A

$$\begin{aligned} \langle A \rangle &= \frac{1}{2}\langle A \rangle_0 + \frac{1}{2}\langle A \rangle_+ = \frac{1}{2}\text{Tr}(|0\rangle\langle 0|A) + \frac{1}{2}\text{Tr}(|+\rangle\langle +|A) \\ &= \text{Tr}\left[\left(\frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|+\rangle\langle +|\right)A\right] \\ &\equiv \text{Tr}(\rho A), \end{aligned} \quad (4.2)$$

where we defined

$$\rho = \frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|+\rangle\langle +|. \quad (4.3)$$

The statistical mixture is therefore properly described by an *operator*, rather than a simple vector. We can generalize this as

$$\rho = \sum_k p_k |\psi_k\rangle\langle\psi_k|, \quad (4.4)$$

where the p_k are probabilities that sum up to one ($\sum_k p_k = 1$) and the $|\psi_k\rangle$ are normalized states (not necessarily complete or orthogonal). Since ρ acts as a weight, or a density, in the expectation value, we call it the *density operator*. We can diagonalize ρ to find the spectral decomposition

$$\rho = \sum_j \lambda_j |\lambda_j\rangle\langle\lambda_j|, \quad (4.5)$$

where $\{|\lambda_j\rangle\}_j$ forms a complete orthonormal basis, $0 \leq \lambda_j \leq 1$, and $\sum_j \lambda_j = 1$. We can also show that ρ is a positive operator:

$$\begin{aligned} \langle\psi|\rho|\psi\rangle &= \sum_{jk} c_j^* c_k \langle\lambda_j|\rho|\lambda_k\rangle = \sum_{jk} c_j^* c_k \langle\lambda_j|\sum_l \lambda_l |\lambda_l\rangle\langle\lambda_l|\lambda_k\rangle \\ &= \sum_{jkl} c_j^* c_k \lambda_l \langle\lambda_j|\lambda_l\rangle\langle\lambda_l|\lambda_k\rangle = \sum_{jkl} c_j^* c_k \lambda_l \delta_{jl} \delta_{lk} \\ &= \sum_l \lambda_l |c_l|^2 \\ &\geq 0. \end{aligned} \quad (4.6)$$

In general, an operator ρ is a valid density operator if and only if it has the following three properties:

1. $\rho^\dagger = \rho$,
2. $\text{Tr}(\rho) = 1$,
3. $\rho \geq 0$.

The density operator is a generalization of the state of a quantum system when we have incomplete information. In the special case where one of the $p_j = 1$ and the others are zero, the density operator becomes the projector $|\psi_j\rangle\langle\psi_j|$. In other words, it is completely determined by the state vector $|\psi_j\rangle$. We call these *pure states*. The statistical mixture of pure states giving rise to the density operator is called a *mixed state*.

The unitary evolution of the density operator can be derived directly from the Schrödinger equation $i\hbar\partial_t|\psi\rangle = H|\psi\rangle$:

$$\begin{aligned} i\hbar \frac{d\rho}{dt} &= i\hbar \frac{d}{dt} \sum_j p_j |\psi_j\rangle\langle\psi_j| \\ &= i\hbar \sum_j \left\{ \frac{dp_j}{dt} |\psi_j\rangle\langle\psi_j| + p_j \left[\left(\frac{d}{dt} |\psi_j\rangle \right) \langle\psi_j| + |\psi_j\rangle \left(\frac{d}{dt} \langle\psi_j| \right) \right] \right\} \\ &= i\hbar \frac{\partial\rho}{\partial t} + H\rho - \rho H \\ &= [H, \rho] + i\hbar \frac{\partial\rho}{\partial t}. \end{aligned} \quad (4.7)$$

This agrees with the Heisenberg equation for operators, and it is sometimes known as the *Von Neumann equation*. In most problems the probabilities p_j have no explicit time-dependence, and $\partial_t \rho = 0$.

4.2 Decoherence

The density operator allows us to consider the phenomenon of *decoherence*. Consider the pure state $|+\rangle$. In matrix notation with respect to the basis $\{|+\rangle, |-\rangle\}$, this can be written as

$$\rho = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}. \quad (4.8)$$

The trace is 1, and one of the eigenvalues is 1, as required for a pure state. We can also write the density operator in the basis $\{|0\rangle, |1\rangle\}$:

$$\rho = |+\rangle\langle+| = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} \times \frac{1}{\sqrt{2}} (1, 1) = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}. \quad (4.9)$$

Notice how the outer product (as opposed to the inner product) of two vectors creates a matrix representation of the corresponding projection operator.

Let the time evolution of $|+\rangle$ be given by

$$|+\rangle = \frac{|0\rangle + |1\rangle}{\sqrt{2}} \rightarrow \frac{|0\rangle + e^{i\omega t}|1\rangle}{\sqrt{2}}. \quad (4.10)$$

The corresponding density operator becomes

$$\rho(t) = \frac{1}{2} \begin{pmatrix} 1 & e^{i\omega t} \\ e^{-i\omega t} & 1 \end{pmatrix}. \quad (4.11)$$

The “population” in the state $|+\rangle$ is given by the expectation value

$$\langle+|\rho(t)|+\rangle = \frac{1}{2} + \frac{1}{2} \cos(\omega t). \quad (4.12)$$

This oscillation is due to the off-diagonal elements of $\rho(t)$, and it is called the *coherence* of the system (see Fig. 2). The state is pure at any time t . In real physical systems the coherence often decays exponentially at a rate γ , and the density matrix can be written as

$$\rho(t) = \frac{1}{2} \begin{pmatrix} 1 & e^{i\omega t - \gamma t} \\ e^{-i\omega t - \gamma t} & 1 \end{pmatrix}. \quad (4.13)$$

The population in the state $|+\rangle$ decays accordingly as

$$\langle+|\rho(t)|+\rangle = \frac{1}{2} + \frac{e^{-\gamma t} \cos(\omega t)}{2}. \quad (4.14)$$

This is called *decoherence* of the system, and the value of γ depends on the physical mechanism that leads to the decoherence.

The decoherence described above is just one particular type, and is called *dephasing*. Another important decoherence mechanism is relaxation to the ground state. If the state $|1\rangle$ has a larger energy than $|0\rangle$ there may be processes such as spontaneous emission that drive the system to the ground state. Combining these two decay processes, we can write the density operator as

$$\rho(t) = \frac{1}{2} \begin{pmatrix} 2 - e^{-\gamma_1 t} & e^{i\omega t - \gamma_2 t} \\ e^{-i\omega t - \gamma_2 t} & e^{-\gamma_1 t} \end{pmatrix}. \quad (4.15)$$

The study of decoherence is currently one of the most important research areas in quantum physics.

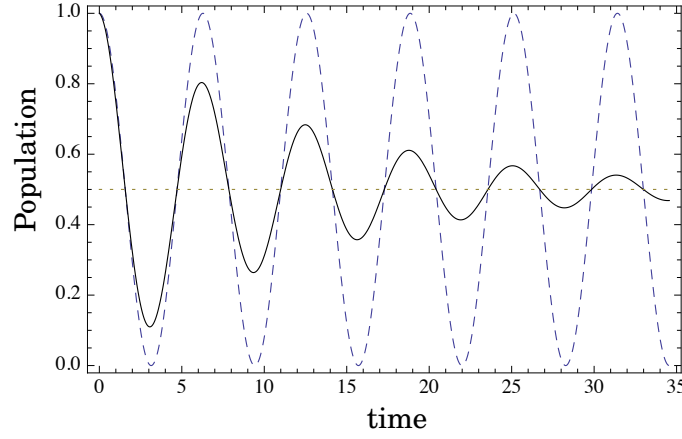


Figure 2: Population in the state $|+\rangle$ with decoherence (solid curve) and without (dashed curve).

4.3 Imperfect measurements

Postulates 3 and 5 determine what happens when a quantum system is subjected to a measurement. In particular, these postulates concern *ideal* measurements. In practice, however, we often have to deal with imperfect measurements that include noise. As a simple example, consider a photodetector: not every photon that hits the detector will result in a detector “click”, which tells us that there indeed was a photon. How can we describe situations like these?

First, let’s recap ideal measurements. We can ask the question what will be the outcome of a *single* measurement of the observable A . We know from Postulate 3 that the outcome m must be an eigenvalue a_m of A . If the spectral decomposition of A is given by

$$A = \sum_m a_m |m\rangle \langle m|, \quad (4.16)$$

then the probability of finding measurement outcome m is given by the Born rule

$$p(m) = |\langle m|\psi\rangle|^2 = \text{Tr}(P_m |\psi\rangle \langle \psi|) \equiv \langle P_m \rangle, \quad (4.17)$$

where we introduced the operator $P_m = |m\rangle \langle m|$. One key interpretation of $p(m)$ is as the expectation value of the operator P_m associated with measurement outcome m .

When a measurement does not destroy the system, the state of the system must be updated after the measurement to reflect the fact that another measurement of A immediately following the first will yield outcome m with certainty. The update rule is given by Eq. (2.24)

$$|\psi\rangle \rightarrow \frac{P_m |\psi\rangle}{\sqrt{\langle \psi | P_m^\dagger P_m | \psi \rangle}}, \quad (4.18)$$

where the square root in the denominator is included to ensure proper normalization of the state after the measurement. However, this form is not so easily generalized to measurements yielding incomplete information, so instead we will write (using $P_m^2 = P_m = P_m^\dagger$)

$$|\psi\rangle \langle \psi| \rightarrow \frac{P_m |\psi\rangle \langle \psi| P_m^\dagger}{\text{Tr}(P_m |\psi\rangle \langle \psi| P_m^\dagger)}. \quad (4.19)$$

From a physical perspective this notation is preferable, since the unobservable global phase of $\langle m|\psi\rangle$ no longer enters the description. Up to this point, both state preparation and measurement were assumed to be ideal. How must this formalism of ideal measurements be modified in order to take into account measurements that yield only partial information about the system? First, we must find the probability for measurement outcome m , and second, we have to formulate the update rule for the state after the measurement.

When we talk about imperfect measurements, we must have some way of knowing (or suspecting) how the measurement apparatus fails. For example, we may suspect that a photon hitting a photodetector has only a finite *probability* of triggering a detector “click”. We therefore describe our measurement device with a general probability distribution $\{q_j(m)\}$, where $q_j(m)$ is the probability that the measurement outcome m in the detector is triggered by a system in the state $|\psi_j\rangle$. The probabilities $q_j(m)$ can be found by modelling the physical aspects of the measurement apparatus. The accuracy of this model can then be determined by experiment. The probability of measurement outcome m for the ideal case is given in Eq. (4.17) by $\langle P_m \rangle$. When the detector is not ideal, the probability of finding outcome m is given by the weighted average over all possible expectation values $\langle P_j \rangle$ that can lead to m :

$$\begin{aligned} p(m) &= \sum_j q_j(m) \langle P_j \rangle = \sum_j q_j(m) \text{Tr}(P_j |\psi\rangle \langle \psi|) \\ &= \text{Tr} \left[\sum_j q_j(m) P_j |\psi\rangle \langle \psi| \right] \equiv \text{Tr}(E_m |\psi\rangle \langle \psi|) \\ &= \langle E_m \rangle, \end{aligned} \tag{4.20}$$

where we defined the operator E_m associated with outcome m as

$$E_m = \sum_j q_j(m) P_j. \tag{4.21}$$

Each possible measurement outcome m has an associated operator E_m , the expectation value of which is the probability of obtaining m in the measurement. The set of E_m is called a *Positive Operator-Valued Measure* (POVM). While the total number of ideal measurement outcomes, modelled by P_m , must be identical to the dimension of the Hilbert space (ignoring the technical complication of degeneracy), this is no longer the case for the POVM described by E_m ; there can be more or fewer measurement outcomes, depending on the physical details of the measurement apparatus. For example, the measurement of an electron spin in a Stern-Gerlach apparatus can have outcomes “up”, “down”, or “failed measurement”. The first two are determined by the position of the fluorescence spot on the screen, and the last may be the situation where the electron fails to produce a spot on the screen. Here, the number of possible measurement outcomes is greater than the number of eigenstates of the spin operator. Similarly, photodetectors in Geiger mode have only two possible outcomes, namely a “click” or “no click” depending on whether the detector registered photons or not. However, the number of eigenstates of the intensity operator (the photon number states) is infinite.

Similar to the density operator, the POVM elements E_m have three key properties. First, the $p(m)$ are probabilities and therefore real, so for all states $|\psi\rangle$ we have

$$\langle E_m \rangle^* = \langle E_m \rangle \iff \langle \psi | E_m | \psi \rangle = \langle \psi | E_m^\dagger | \psi \rangle, \tag{4.22}$$

and E_m is therefore Hermitian. Second, since $\sum_m p(m) = 1$ we have

$$\sum_m p(m) = \sum_m \langle \psi | E_m | \psi \rangle = \langle \psi | \sum_m E_m | \psi \rangle = 1, \tag{4.23}$$

for all $|\psi\rangle$, and therefore

$$\sum_m E_m = \mathbb{I}, \quad (4.24)$$

where \mathbb{I} is the identity operator. Finally, since $p(m) = \langle \psi | E_m | \psi \rangle \geq 0$ for all $|\psi\rangle$, the POVM element E_m is a positive operator. Note the close analogy of the properties of the POVM and the density operator. In particular, just as in the case of density operators, POVMs are *defined* by these three properties.

The second question about generalized measurements is how the measurement outcomes should be used to update the quantum state of the system. The rule for ideal von Neumann measurements is given in Eq. (4.19), which can be generalized immediately to density operators using the techniques presented above. This yields

$$\rho \rightarrow \frac{P_m \rho P_m^\dagger}{\text{Tr}(P_m \rho P_m^\dagger)}. \quad (4.25)$$

What if we have instead a measurement apparatus described by a POVM? Consistency with the Born rule in Eq. (4.20) requires that we again replace P_m , associated with measurement outcome m , with a probability distribution over all P_j :

$$\rho \rightarrow \sum_j q_j(m) \frac{P_j \rho P_j^\dagger}{\text{Tr}[\sum_k q_k(m) P_k \rho P_k^\dagger]} = \sum_j q_j(m) \frac{P_j \rho P_j^\dagger}{\text{Tr}(E_m \rho)}, \quad (4.26)$$

where we adjusted the normalization factor to maintain $\text{Tr}(\rho) = 1$. We also used that the POVM element E_m in Eq. (4.21) can be written as

$$E_m = \sum_j q_j(m) P_j^\dagger P_j. \quad (4.27)$$

If we rescale the P_j by a factor $\sqrt{q_j(m)}$, we obtain the standard form of the POVM

$$E_m = \sum_j A_{jm}^\dagger A_{jm}, \quad (4.28)$$

where $A_{jm} = \sqrt{q_j(m)} P_j$ are the so-called the Kraus operators. Consequently, the update rule can be written as

$$\rho \rightarrow \frac{\sum_j A_{jm} \rho A_{jm}^\dagger}{\text{Tr}(E_m \rho)}, \quad (4.29)$$

which generally yields a mixed state (described by a density operator) after an incomplete measurement.

Finally, we can consider the case of some nonunitary evolution *without* a measurement (e.g., when the system interacts with its environment). We can model this purely formally by removing the index of the measurement outcome m from the description (since there are no measurement outcomes). The most general evolution then takes the form

$$\rho \rightarrow \rho' = \sum_k A_k \rho A_k^\dagger, \quad (4.30)$$

and the question is what form the Kraus operators A_k take. We will return to this in section 6.

Exercises

1. The density matrix.

- Show that $\frac{1}{2}|0\rangle + \frac{1}{2}|+\rangle$ is not a properly normalized state.
- Show that $\text{Tr}(\rho) = 1$ with ρ given by Eq. (4.3), and then prove that any density operator has unit trace and is Hermitian.
- Show that density operators are *convex*, i.e., that $\rho = w_1\rho_1 + w_2\rho_2$ with $w_1 + w_2 = 1$ ($w_1, w_2 \geq 0$), and ρ_1, ρ_2 again density operators.
- Calculate the expectation value of A using the two representations of ρ in terms of p_i and the spectral decomposition. What is the difference in the physical interpretation of p_j and λ_j ?

2. Using the identity $\langle x|A|\psi\rangle = A\psi(x)$, and the resolution of the identity $\int dx |x\rangle\langle x| = \mathbb{I}$, calculate the expectation value for an operator A , given a mixed state of wave functions.

3. Calculate P^2 with P given by

$$P = \begin{pmatrix} a \\ b \end{pmatrix} (a^*, b^*) = \begin{pmatrix} |a|^2 & ab^* \\ a^*b & |b|^2 \end{pmatrix} \quad \text{and} \quad |a|^2 + |b|^2 = 1. \quad (4.31)$$

4. Calculate the eigenvalues of the density matrix in Eq. (4.15), and show that $\gamma_1 \leq 2\gamma_2$. Hint: it is difficult to derive the inequality directly, so you should try to demonstrate that $\gamma_1 \geq 2\gamma_2$ leads to a contradiction.

5. A system with energy eigenstates $|E_n\rangle$ is in thermal equilibrium with a heat bath at temperature T . The probability for the system to be in state E_n is proportional to $e^{-E_n/kT}$.

- Write the Hamiltonian of the system in terms of E_n and $|E_n\rangle$.
- Give the normalized density operator ρ for the system as a function of the Hamiltonian.
- We identify the normalization of ρ with the partition function \mathcal{Z} . Calculate the average energy directly and via

$$\langle E \rangle = -\frac{\partial \ln \mathcal{Z}}{\partial \beta} \quad (4.32)$$

of the system, where $\beta = 1/(kT)$.

- What is the entropy $S = k \ln \mathcal{Z}$ if the system is a harmonic oscillator? Comment on the limit $T \rightarrow 0$.

6. Lossy photodetectors.

- The state of a beam of light can be written in the photon number basis $|n\rangle$ as $|\psi\rangle = \sum_n c_n |n\rangle$. What are the possible measurement outcomes for a perfect photon detector? Calculate the probabilities of the measurement outcome using projection operators.
- Suppose that the detector can only tell the difference between the presence and absence of photons (a so-called “bucket detector”). How do we calculate the probability of finding the measurement outcomes?

- (c) Real bucket detectors have a finite efficiency η , which means that each photon has a probability η of triggering the detector. Calculate the probabilities of the measurement outcomes.
 - (d) What other possible imperfections do realistic bucket detectors have?
7. An electron with spin state $|\psi\rangle = \alpha|\uparrow\rangle + \beta|\downarrow\rangle$ and $|\alpha|^2 + |\beta|^2 = 1$ is sent through a Stern-Gerlach apparatus to measure the spin in the z direction (i.e., the $|\uparrow\rangle, |\downarrow\rangle$ basis).
- (a) What are the possible measurement outcomes? If the position of the electron is measured by an induction loop rather than a screen, what is the state of the electron immediately after the measurement?
 - (b) Suppose that with probability p the induction loops fail to give a current signifying the presence of an electron. What are the three possible measurement outcomes? Give the POVM.
 - (c) Calculate the probabilities of the measurement outcomes, and the state of the electron immediately after the measurement (for all possible outcomes).

5 Composite Systems and Entanglement

5.1 Composite systems

Suppose we have two systems, described by Hilbert spaces \mathcal{H}_1 and \mathcal{H}_2 , respectively. We can choose orthonormal bases for each system:

$$\mathcal{H}_1 : \{|\phi_1\rangle, |\phi_2\rangle, \dots, |\phi_N\rangle\} \quad \text{and} \quad \mathcal{H}_2 : \{|\psi_1\rangle, |\psi_2\rangle, \dots, |\psi_M\rangle\}. \quad (5.1)$$

The respective dimensions of \mathcal{H}_1 and \mathcal{H}_2 are N and M . We can construct $N \times M$ basis states for the composite system via $|\phi_j\rangle$ and $|\psi_k\rangle$. This implies that the total Hilbert space of the composite system can be spanned by the tensor product

$$\{|\phi_j\rangle \otimes |\psi_k\rangle\}_{jk} \quad \text{on} \quad \mathcal{H}_{1+2} = \mathcal{H}_1 \otimes \mathcal{H}_2. \quad (5.2)$$

An arbitrary pure state on \mathcal{H}_{1+2} can be written as

$$|\Psi\rangle = \sum_{jk} c_{jk} |\phi_j\rangle \otimes |\psi_k\rangle \equiv \sum_{jk} c_{jk} |\phi_j, \psi_k\rangle. \quad (5.3)$$

For example, the system of two qubits can be written on the basis $\{|0,0\rangle, |0,1\rangle, |1,0\rangle, |1,1\rangle\}$. If system 1 is in state $|\phi\rangle$ and system 2 is in state $|\psi\rangle$, the partial trace over system 2 yields

$$\text{Tr}_2(|\phi, \psi\rangle \langle \phi, \psi|) = \text{Tr}_2(|\phi\rangle \langle \phi| \otimes |\psi\rangle \langle \psi|) = |\phi\rangle \langle \phi| \text{Tr}(|\psi\rangle \langle \psi|) = |\phi\rangle \langle \phi|, \quad (5.4)$$

since the trace over any density operator is 1. We have now *lost* system 2 from our description! Therefore, taking the partial trace without inserting any other operators is the mathematical version of forgetting about it. This is a very useful feature: you often do not want to deal with every possible system you are interested in. For example, if system 1 is a qubit, and system two is a very large environment the partial trace allows you to “trace out the environment”.

However, tracing out the environment will not always leave you with a pure state as in Eq. (5.4). If the system has interacted with the environment, taking the partial trace generally leaves you with a mixed state. This is due to *entanglement* between the system and its environment.

5.2 Entanglement

Consider the following experiment: Alice and Bob each blindly draw a marble from a vase that contains one black and one white marble. Let’s call the state of the white marble $|0\rangle$ and the state of the black marble $|1\rangle$. If we describe this classical experiment quantum mechanically (we can always do this, because classical physics is contained in quantum physics), then there are two possible states, $|1,0\rangle$ and $|0,1\rangle$. Since blind drawing is a statistical procedure, the state of the marbles held by Alice and Bob is the mixed state

$$\rho = \frac{1}{2} |0,1\rangle \langle 0,1| + \frac{1}{2} |1,0\rangle \langle 1,0|. \quad (5.5)$$

From Alice’s perspective, the state of her marble is obtained by tracing over Bob’s marble:

$$\rho_A = \text{Tr}_B(\rho) = \frac{1}{2} |0\rangle \langle 0| + \frac{1}{2} |1\rangle \langle 1|. \quad (5.6)$$

This is what we expect: Alice has a 50:50 probability of finding “white” or “black” when she looks at her marble (i.e., when she measures the colour of the marble).

Next, consider what the state of Bob's marble is when Alice finds a white marble. Just from the setup we know that Bob's marble must be black, because there was only one white and one black marble in the vase. Let's see if we can reproduce this in our quantum mechanical description. Finding a white marble can be described mathematically by a projection operator $|0\rangle\langle 0|$ (see Eq. (2.24)). We need to include this operator in the trace over Alice's marble's Hilbert space:

$$\rho_B = \frac{\text{Tr}_A(|0\rangle\langle 0|_A \rho)}{\text{Tr}(|0\rangle\langle 0|_A \rho)} = |1\rangle\langle 1|, \quad (5.7)$$

which we set out to prove: if Alice finds that when she sees that her marble is white, she describes the state of Bob's marble as black. Based on the setup of this experiment, Alice knows instantaneously what the state of Bob's marble is as soon as she looks at her own marble. There is nothing spooky about this; it just shows that the marbles held by Alice and Bob are *correlated*.

Next, consider a second experiment: By some procedure, the details of which are not important right now, Alice and Bob each hold a two-level system (a qubit) in the *pure* state

$$|\Psi\rangle_{AB} = \frac{|0, 1\rangle + |1, 0\rangle}{\sqrt{2}}. \quad (5.8)$$

Since $|1, 0\rangle$ and $|0, 1\rangle$ are valid quantum states, by virtue of the first postulate of quantum mechanics $|\Psi\rangle_{AB}$ is also a valid quantum mechanical state. It is not difficult to see that these systems are also correlated in the states $|0\rangle$ and $|1\rangle$: When Alice finds the value “0”, Bob must find the value “1”, and vice versa. We can write this state as a density operator

$$\begin{aligned} \rho &= \frac{1}{2} (|0, 1\rangle + |1, 0\rangle)(\langle 0, 1| + \langle 1, 0|) \\ &= \frac{1}{2} (|0, 1\rangle\langle 0, 1| + |0, 1\rangle\langle 1, 0| + |1, 0\rangle\langle 0, 1| + |1, 0\rangle\langle 1, 0|). \end{aligned} \quad (5.9)$$

Notice the two extra terms with respect to Eq. (5.5). If Alice now traces out Bob's system, she finds that the state of her marble is

$$\rho_A = \text{Tr}_B(\rho) = \frac{1}{2} |0\rangle\langle 0| + \frac{1}{2} |1\rangle\langle 1|. \quad (5.10)$$

In other words, even though the total system was in a pure state, the subsystem held by Alice (and Bob, check this) is mixed! We can try to put the two states back together:

$$\begin{aligned} \rho_A \otimes \rho_B &= \left(\frac{1}{2} |0\rangle\langle 0| + \frac{1}{2} |1\rangle\langle 1| \right) \otimes \left(\frac{1}{2} |0\rangle\langle 0| + \frac{1}{2} |1\rangle\langle 1| \right) \\ &= \frac{1}{4} (|0, 0\rangle\langle 0, 0| + |0, 1\rangle\langle 0, 1| + |1, 0\rangle\langle 1, 0| + |1, 1\rangle\langle 1, 1|), \end{aligned} \quad (5.11)$$

but this is not the state we started out with! It is also a mixed state, instead of the pure state we started with. Since mixed states mean incomplete knowledge, there must be some information in the combined system that *does not reside in the subsystems alone*! This is called *entanglement*.

Entanglement arises because states like $(|0, 1\rangle + |1, 0\rangle)/\sqrt{2}$ cannot be written as a tensor product of two pure states $|\psi\rangle \otimes |\phi\rangle$. These latter states are called *separable*. In general a state is separable if and only if it can be written as

$$\rho = \sum_j p_j \rho_j^{(A)} \otimes \rho_j^{(B)}. \quad (5.12)$$

Classical correlations such as the black and white marbles above fall into the category of separable states.

So far, we have considered the quantum states in the basis $\{|0\rangle, |1\rangle\}$. However, we can also describe the same system in the rotated basis $\{|\pm\rangle\}$ according to

$$|0\rangle = \frac{|+\rangle + |-\rangle}{\sqrt{2}} \quad \text{and} \quad |1\rangle = \frac{|+\rangle - |-\rangle}{\sqrt{2}}. \quad (5.13)$$

The entangled state $|\Psi\rangle_{AB}$ can then be written as

$$\frac{|0, 1\rangle + |1, 0\rangle}{\sqrt{2}} = \frac{|+, +\rangle - |-, -\rangle}{\sqrt{2}}, \quad (5.14)$$

which means that we have again perfect correlations between the two systems with respect to the states $|+\rangle$ and $|-\rangle$. Let's do the same for the state ρ in Eq. (5.5) for classically correlated marbles. After a bit of algebra, we find that

$$\begin{aligned} \rho = \frac{1}{4} & (|++\rangle\langle++| + |+-\rangle\langle+-| + |-+\rangle\langle-+| + |--\rangle\langle--| \\ & - |++\rangle\langle--| - |--\rangle\langle++| - |+-\rangle\langle-+| - |-+\rangle\langle+-|). \end{aligned} \quad (5.15)$$

Now there are *no correlations* in the conjugate basis $\{|\pm\rangle\}$, which you can check by calculating the conditional probabilities of Bob's state given Alice's measurement outcomes. This is another key difference between classically correlated states and entangled states. A good interpretation of entanglement is that entangled systems exhibit correlations that are *stronger* than classical correlations. We will shortly see how these stronger correlations can be used in information processing.

We have seen that operators, just like states, can be combined into tensor products:

$$A \otimes B |\phi\rangle \otimes |\psi\rangle = A |\phi\rangle \otimes B |\psi\rangle. \quad (5.16)$$

And just like states, some operators cannot be written as $A \otimes B$:

$$C = \sum_k A_k \otimes B_k. \quad (5.17)$$

This is the most general expression of an operator in the Hilbert space $\mathcal{H}_1 \otimes \mathcal{H}_2$. In Dirac notation this becomes

$$C = \sum_{jklm} \phi_{jklm} |\phi_j\rangle\langle\phi_k| \otimes |\phi_l\rangle\langle\phi_m| = \sum_{jklm} \phi_{jklm} |\phi_j, \phi_l\rangle\langle\phi_k, \phi_m|. \quad (5.18)$$

As an example, the Bell operator is diagonal on the Bell basis:

$$|\Phi^\pm\rangle = \frac{|0, 0\rangle \pm |1, 1\rangle}{\sqrt{2}} \quad \text{and} \quad |\Psi^\pm\rangle = \frac{|0, 1\rangle \pm |1, 0\rangle}{\sqrt{2}}. \quad (5.19)$$

The eigenvalues of the Bell operator are not important, as long as they are not degenerate (why?). A measurement of the Bell operator projects onto an eigenstate of the operator, which is an entangled state. Consequently, we cannot implement such composite measurements by measuring each subsystem individually, because those individual measurements would project onto pure states of the subsystems. And we have seen that the subsystems of pure entangled states are mixed states.

A particularly useful technique when dealing with two systems is the so-called *Schmidt decomposition*. In general, we can write any pure state over two systems as a superposition of basis states:

$$|\Psi\rangle = \sum_{j=1}^{d_A} \sum_{k=1}^{d_B} c_{jk} |\phi_j\rangle_A |\psi_k\rangle_B, \quad (5.20)$$

where d_A and d_B are the dimensions of the Hilbert spaces of system A and B , respectively, and we order the systems such that $d_B \geq d_A$. It turns out that we can always simplify this description and write $|\Psi\rangle$ as a *single* sum over basis states. We state it as a theorem:

Theorem: Let $|\Psi\rangle$ be a pure state of two systems, A and B with Hilbert spaces \mathcal{H}_A and \mathcal{H}_B of dimension d_A and $d_B \geq d_A$, respectively. There exist orthonormal basis vectors $|a_j\rangle_A$ for system A and $|b_j\rangle_B$ for system B such that

$$|\Psi\rangle = \sum_j \lambda_j |a_j\rangle_A |b_j\rangle_B, \quad (5.21)$$

with real, positive Schmidt coefficients λ_j , and $\sum_j \lambda_j^2 = 1$. This decomposition is *unique*, and the sum runs at most to d_A , the dimension of the *smallest* Hilbert space. Traditionally, we order the Schmidt coefficients in descending order: $\lambda_1 \geq \lambda_2 \geq \dots$. The total number of non-zero λ_i is the *Schmidt number*.

The proof can be found in many graduate texts on quantum mechanics and quantum information theory.

Given the Schmidt decomposition for a bi-partite system, we can immediately write down the reduced density matrices for the sub-systems:

$$\rho_A = \text{Tr}_B(|\Psi\rangle\langle\Psi|) = \sum_j \lambda_j^2 |a_j\rangle_A \langle a_j|, \quad (5.22)$$

and

$$\rho_B = \text{Tr}_A(|\Psi\rangle\langle\Psi|) = \sum_j \lambda_j^2 |b_j\rangle_B \langle b_j|. \quad (5.23)$$

The basis states $|a_j\rangle_A$ and $|b_j\rangle_B$ may have completely different physical meanings; here we care only that the states of the decomposition can be labelled with a single index, as opposed to two indices.

Conversely, when we have a single system in a mixed state

$$\rho = \sum_j p_j |a_j\rangle \langle a_j|, \quad (5.24)$$

we can always construct a pure state $|\Psi\rangle$ that obeys ($\lambda_j = \sqrt{p_j}$)

$$|\Psi\rangle = \sum_j \lambda_j |a_j, b_j\rangle, \quad (5.25)$$

By virtue of the Schmidt decomposition. The state $|\Psi\rangle$ is called the *purification* of ρ . Since many theorems are easier to prove for pure states than for mixed states, purifications can make our work load significantly lighter.

When there is more than one non-zero λ_j in Eq. (5.25), the state $|\Psi\rangle$ is clearly entangled: there is no alternative choice of λ_j due to the uniqueness of the Schmidt decomposition that would result in $\lambda'_1 = 1$ and all others zero. Moreover, the more uniform the values of λ_j , the more the state is entangled. One possible measure for the amount of entanglement in $|\Psi\rangle$ is the *Shannon entropy* H

$$H = -\sum_j \lambda_j^2 \log_2 \lambda_j^2. \quad (5.26)$$

This is identical to the von Neumann entropy S of the reduced density matrix ρ of $|\Psi\rangle$ given in Eq. (5.24):

$$S(\rho) = -\text{Tr}(\rho \log_2 \rho). \quad (5.27)$$

Both entropies are measured in classical bits.

How do we find the Schmidt decomposition? Consider the state $|\Psi\rangle$ from Eq. (5.20). The (not necessarily square) matrix C with elements c_{jk} needs to be transformed into a single array of numbers λ_j . This is achieved by applying the singular-value decomposition:

$$c_{jk} = \sum_i u_{ji} d_{ii} v_{ik}, \quad (5.28)$$

where u_{ji} and v_{ik} are elements of unitary matrices U and V , respectively, and d_{ii} is a diagonal matrix with singular values λ_i . The vectors in the Schmidt decomposition become

$$|a_i\rangle = \sum_j u_{ji} |\phi_j\rangle \quad \text{and} \quad |b_i\rangle = \sum_k v_{ik} |\psi_k\rangle. \quad (5.29)$$

This is probably a good time to remind ourselves about the singular-value decomposition. All we need to do is find U and V , the rest is just matrix multiplication. To find U , we diagonalize CC^\dagger and find its eigenvectors. These form the columns of U . Similarly, we diagonalize $C^\dagger C$ and arrange the eigenvectors in columns to find V . If C is an $n \times m$ matrix, U should be $n \times n$ and V should be $m \times m$.

5.3 Quantum teleportation

Probably the most extraordinary use of the quantum correlations present in entanglement is *quantum teleportation*. Alice and Bob share two entangled qubits, labelled 2 (held by Alice) and 3 (held by Bob), in the state $(|0,0\rangle_{23} + |1,1\rangle_{23})/\sqrt{2}$. In addition, Alice holds a qubit in the state

$$|\psi\rangle_1 = \alpha|0\rangle_1 + \beta|1\rangle_1. \quad (5.30)$$

The object of quantum teleportation is to transfer the *state* of qubit 1 to qubit 3, without either Alice or Bob gaining *any* information about α or β . To make things extra hard, the three qubits must not change places (so Alice cannot take qubit 1 and bring it to Bob).

Classically, this is an impossible task: we cannot extract enough information about α and β with a single measurement to reproduce $|\psi\rangle$ faithfully, otherwise we could violate the no-cloning theorem. However, in quantum mechanics it can be done (without violating no-cloning). Write the total state as

$$|\chi\rangle = |\psi\rangle_1 |\Phi^+\rangle_{23} = \frac{1}{\sqrt{2}} (\alpha|000\rangle + \alpha|011\rangle + \beta|100\rangle + \beta|111\rangle). \quad (5.31)$$

Alice now performs a Bell measurement on her two qubits (1 and 2), which project them onto one of the Bell states $|\Phi^\pm\rangle_{12}$ or $|\Psi^\pm\rangle_{12}$. We write $|00\rangle$, $|01\rangle$, $|10\rangle$, and $|11\rangle$ in the Bell basis:

$$\begin{aligned} |00\rangle &= \frac{|\Phi^+\rangle + |\Phi^-\rangle}{\sqrt{2}} \\ |01\rangle &= \frac{|\Psi^+\rangle + |\Psi^-\rangle}{\sqrt{2}} \\ |10\rangle &= \frac{|\Psi^+\rangle - |\Psi^-\rangle}{\sqrt{2}} \\ |11\rangle &= \frac{|\Phi^+\rangle - |\Phi^-\rangle}{\sqrt{2}}. \end{aligned} \quad (5.32)$$

We can use these substitutions to write the state $|\chi\rangle$ before the measurement as

$$\begin{aligned} |\chi\rangle &= \frac{1}{2} [\alpha |\Phi^+\rangle_{12} |0\rangle_3 + \alpha |\Phi^-\rangle_{12} |0\rangle_3 + \alpha |\Psi^+\rangle_{12} |1\rangle_3 + \alpha |\Psi^-\rangle_{12} |1\rangle_3 \\ &\quad + \beta |\Psi^+\rangle_{12} |0\rangle_3 - \beta |\Psi^-\rangle_{12} |0\rangle_3 + \beta |\Phi^+\rangle_{12} |1\rangle_3 - \beta |\Phi^-\rangle_{12} |1\rangle_3] \\ &= \frac{1}{2} [|\Phi^+\rangle(\alpha|0\rangle + \beta|1\rangle) + |\Phi^-\rangle(\alpha|0\rangle - \beta|1\rangle) \\ &\quad + |\Psi^+\rangle(\beta|0\rangle + \alpha|1\rangle) + |\Psi^-\rangle(\beta|0\rangle - \alpha|1\rangle)]. \end{aligned} \quad (5.33)$$

Alice finds one of four possible outcomes:

$$\begin{aligned} \Phi^+ : \quad \text{Tr}_{12}(|\Phi^+\rangle\langle\Phi^+| |\chi\rangle\langle\chi|) &\rightarrow |\psi\rangle_3 = \alpha|0\rangle + \beta|1\rangle, \\ \Phi^- : \quad \text{Tr}_{12}(|\Phi^-\rangle\langle\Phi^-| |\chi\rangle\langle\chi|) &\rightarrow |\psi\rangle_3 = \alpha|0\rangle - \beta|1\rangle, \\ \Psi^+ : \quad \text{Tr}_{12}(|\Psi^+\rangle\langle\Psi^+| |\chi\rangle\langle\chi|) &\rightarrow |\psi\rangle_3 = \alpha|1\rangle + \beta|0\rangle, \\ \Psi^- : \quad \text{Tr}_{12}(|\Psi^-\rangle\langle\Psi^-| |\chi\rangle\langle\chi|) &\rightarrow |\psi\rangle_3 = \alpha|1\rangle - \beta|0\rangle. \end{aligned} \quad (5.34)$$

From these outcomes, it is clear that the state held by Bob is different for the different measurement outcomes of Alice's Bell measurement. Let this sink in for a moment: After setting up the entangled state between Alice and Bob, who may be literally light years apart, Bob has done *absolutely nothing* to his qubit, yet its state is different depending on Alice's measurement outcome. This suggests that there is some instantaneous communication taking place, possibly violating causality!

In order to turn the state of Bob's qubit into the original state, Alice needs to send the measurement outcome to Bob. This will take two classical bits, because there are four outcomes. The correction operators that Bob need to apply are as follows:

$$\Phi^+ : \mathbb{I}, \quad \Phi^- : Z, \quad \Psi^+ : X, \quad \Psi^- : ZX. \quad (5.35)$$

So in each case Bob needs to do something different to his qubit. To appreciate how remarkable this protocol is, here are some of its relevant properties:

1. No matter is transported, only the *state* of the system;
2. neither Alice nor Bob learns *anything* about α or β ;
3. any attempt to use quantum teleportation for signaling faster than light is *futile*!

Exercises

1. Derive Eq. (5.15), and show that $\rho_B = \mathbb{I}/2$ when Alice's qubit is projected onto $|+\rangle$.
2. Quantum teleportation. Write the Bell states as

$$|\psi_{nm}\rangle = (|0, 0 \oplus n\rangle + (-1)^m |1, 1 \oplus n\rangle) / \sqrt{2},$$

where \oplus denotes addition modulo 2 and $n, m = 0, 1$.

- (a) Write $|\psi_{nm}\rangle$ in terms of $|\psi_{00}\rangle$ and the Pauli operators X and Z acting on the second qubit.
- (b) Using the shared Bell state $|\psi_{nm}\rangle$ between Alice and Bob, and the two-bit measurement outcome (j, k) for Alice's Bell measurement, determine the correction operator for Bob.
- (c) We now generalize to N -dimensional systems. We define the N^2 entangled states

$$|\psi_{nm}\rangle = \frac{1}{\sqrt{N}} \sum_{j=0}^{N-1} e^{2\pi i j n / N} |j, j \oplus m\rangle,$$

where $n \oplus m = n + m \pmod{N}$. Prove that this is an orthonormal basis.

- (d) Give the teleportation protocol for the N -dimensional Hilbert space.
 - (e) What is Bob's state before he learns Alice's measurement outcome?
3. Imperfect measurements.
 - (a) A two-qubit system (held by Alice and Bob) is in the anti-symmetric Bell state $|\Psi^-\rangle$. Calculate the state of Bob's qubit if Alice measures her qubit in the state $|0\rangle$. Hint: write the measurement procedure as a partial trace over Alice's qubit.
 - (b) Now Alice's measurement is imperfect, and when her apparatus indicates "0", there was actually a small probability p that the qubit was projected onto the state $|1\rangle$. What is Bob's state?
 - (c) If Alice's (imperfect) apparatus has only two measurement outcomes, what will Bob's state be if she finds outcome "1"?

6 Evolution of Open Quantum Systems

We have considered mixed states, where the experimenter has incomplete information about the state preparation procedure, and we have also seen that mixing arises in a system when it is entangled with another system. The combined system can still be pure, but the subsystem has become mixed. This phenomenon arises often when we want to describe systems that have some interaction with their environment. The interaction creates entanglement, and the system taken by itself evolves from a pure state to a mixed state. Such a system is called “open”, since it can leak quantum information to the environment. The theory of open quantum systems revolves around the so-called *Lindblad equation*.

6.1 The Lindblad equation

Next, we will derive the Lindblad equation, which is the direct extension of the Heisenberg equation for the density operator, i.e., the mixed state of a system. We have seen in Eq. (4.30) that formally, we can write the evolution of a density operator as a mathematical *map* \mathcal{E} , such that the density operator ρ transforms into

$$\rho \rightarrow \rho' = \mathcal{E}(\rho) \equiv \sum_k A_k \rho A_k^\dagger, \quad (6.1)$$

where the A_k are the Kraus operators. Requiring that ρ' is again a density operator ($\text{Tr}(\rho') = 1$) leads to the restriction $\sum_k A_k^\dagger A_k = \mathbb{I}$.

We want to describe an infinitesimal evolution of ρ , in order to give the continuum evolution later on. We therefore have that

$$\rho' = \rho + \delta\rho = \sum_k A_k \rho A_k^\dagger. \quad (6.2)$$

Since $\delta\rho$ is very small, one of the Kraus operators must be close to the identity. Without loss of generality we choose this to be A_0 , and then we can write

$$A_0 = \mathbb{I} + (L_0 - iK)\delta t \quad \text{and} \quad A_k = L_k \sqrt{\delta t}, \quad (6.3)$$

where we introduced the Hermitian operators L_0 and K , and the remaining L_k are not necessarily Hermitian. We could have written $A_0 = \mathbb{I} + L_0 \delta t$ and keep L_0 general (non-Hermitian as well), but it will be useful later on to explicitly decompose it into Hermitian parts. We can now write

$$\begin{aligned} A_0 \rho A_0^\dagger &= \rho + [(L_0 - iK)\rho + \rho(L_0 + iK)] \delta t + O(\delta t^2) \\ A_k \rho A_k^\dagger &= L_k \rho L_k^\dagger \delta t. \end{aligned} \quad (6.4)$$

We can substitute this into Eq. (6.2), to obtain up to first order in δt

$$\delta\rho = \left[(L_0 \rho + \rho L_0) - i(K\rho - \rho K) + \sum_{k \neq 0} L_k \rho L_k^\dagger \right] \delta t. \quad (6.5)$$

We now give the continuum evolution by dividing by δt and taking the limit $\delta t \rightarrow dt$:

$$\frac{d\rho}{dt} = -i[K, \rho] + \{L_0, \rho\} + \sum_{k \neq 0} L_k \rho L_k^\dagger, \quad (6.6)$$

where $\{A, B\} = AB + BA$ is the anti-commutator of A and B . We are almost there, but we must determine what the different terms mean. Suppose we consider the free evolution of the system.

Eq. (6.6) must then reduce to the Heisenberg equation for the density operator ρ in Eq. (4.7), and we see that all L_k including L_0 are zero, and K is proportional to the Hamiltonian $K = H/\hbar$. Again from the general property that $\text{Tr}(\rho) = 1$ we have

$$\text{Tr}\left(\frac{d\rho}{dt}\right) = 0 \rightarrow L_0 = -\frac{1}{2} \sum_{k \neq 0} L_k^\dagger L_k. \quad (6.7)$$

This finally leads to the Lindblad equation

$$\frac{d\rho}{dt} = \frac{1}{i\hbar}[H, \rho] + \frac{1}{2} \sum_k \left(2L_k \rho L_k^\dagger - \{L_k^\dagger L_k, \rho\} \right). \quad (6.8)$$

The operators L_k are chosen such that they model the relevant physical processes. This may sound vague, but in practice it will be quite clear. For example, modelling a transition $|1\rangle \rightarrow |0\rangle$ without keeping track of where the energy is going or coming from will require a single Lindblad operator

$$L = \gamma |0\rangle \langle 1|, \quad (6.9)$$

where γ is a real parameter indicating the strength of the transition. This can model both decay and excitations.

6.2 Positive and completely positive maps

We considered the evolution of the density operator under a family of Kraus operators in Eq. (4.30):

$$\rho \rightarrow \rho' = \mathcal{E}(\rho) = \sum_k A_k \rho A_k^\dagger, \quad (6.10)$$

where $\sum_k A_k^\dagger A_k = \mathbb{I}$ (that is, \mathcal{E} is trace-preserving). When \mathcal{E} transforms any positive operator into another positive operator, we call it a *positive map*. We may be tempted to conclude that all positive maps correspond to physically allowed transformations. After all, it maps density operators to density operators. Unfortunately, Nature (or Mathematics?) is not that tidy.

Consider the transpose of the density operator $\rho \rightarrow \rho^T$, which acts according to

$$\rho = \sum_{ij} \rho_{ij} |i\rangle \langle j| \rightarrow \rho^T = \sum_{ij} \rho_{ji} |i\rangle \langle j|. \quad (6.11)$$

You can verify immediately that the trace is preserved in this operation (check this!), and ρ^T is again a positive operator since the eigenvalues are identical to those of ρ . For example, consider the qubit state $(|0\rangle + i|1\rangle)/\sqrt{2}$. The density operator and its transpose are

$$\rho = \frac{1}{2} \begin{pmatrix} 1 & -i \\ i & 1 \end{pmatrix} \quad \text{and} \quad \rho^T = \frac{1}{2} \begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix}. \quad (6.12)$$

The transpose therefore corresponds to the state $(|0\rangle - i|1\rangle)/\sqrt{2}$. Now consider that the qubit is part of an entangled state $(|00\rangle + |11\rangle)/\sqrt{2}$. The density operator is given by

$$\rho = \frac{1}{2} (|00\rangle \langle 00| + |00\rangle \langle 11| + |11\rangle \langle 00| + |11\rangle \langle 11|), \quad (6.13)$$

and the *partial* transpose on the first qubit is

$$\rho^T = \frac{1}{2} (|00\rangle \langle 00| + |10\rangle \langle 01| + |01\rangle \langle 10| + |11\rangle \langle 11|). \quad (6.14)$$

The eigenvalues of ρ are all positive, but ρ^T has a negative eigenvalue! So ρ^T cannot be a density operator. Consequently, it is *not* correct to say that positive maps correspond to physical processes. We need to put another restriction on maps.

From the example of the partial transpose, we can deduce that maps must not only be positive for the system S that they act on, but also positive on larger systems that include S as a subsystem. When this is the case, we call the map *completely positive*. There is a very important theorem in mathematics, called *Kraus' Representation Theorem*, which states that maps of the form in Eq. (6.10) with the restriction that $\sum_k A_k^\dagger A_k = \mathbb{I}$ is a completely positive map, and moreover, that any completely positive map can be expressed in this form.

Exercises

1. (a) Show that $\sum_k A_k^\dagger A_k = \mathbb{I}$,
 (b) prove that any non-Hermitian square matrix can be written as $A + iB$, with A and B Hermitian,
 (c) prove that $L_0 = -\frac{1}{2} \sum_{k \neq 0} L_k^\dagger L_k$.
2. Consider a two-level system ($|0\rangle, |1\rangle$) that has a dephasing process, modelled by the Lindblad operators $L_1 = \gamma |0\rangle\langle 1|$ and $L_2 = \gamma |1\rangle\langle 0|$.
 (a) write down the Lindblad equation (choose $H = 0$ for simplicity).
 (b) Calculate the evolution of the pure states $|0\rangle$ and $|+\rangle$ at $t = 0$. Hint: write the density matrix in the Pauli matrix basis $\{\mathbb{I}, X, Y, Z\}$. What can you say about the equilibrium state of the system?
 (c) Calculate and plot the entropy $S(\rho)$ of the state $\rho(t)$ as a function of γ and t .
3. Calculate the eigenvalues of ρ^T in Eq. (6.14).

7 Orbital Angular Momentum and Spin

Angular momentum plays an important role in quantum mechanics, not only as the orbital angular momentum of electrons orbiting the central potentials of nuclei, but also as the intrinsic magnetic moment of particles, known as spin, and even as isospin in high-energy particle physics.

7.1 Orbital angular momentum

From classical physics we know that the orbital angular momentum of a particle is given by the cross product of its position and momentum

$$\mathbf{L} = \mathbf{r} \times \mathbf{p} \quad \text{or} \quad L_i = \epsilon_{ijk} r_j p_k, \quad (7.1)$$

where we used Einstein's summation convention for the indices. In quantum mechanics, we can find the operator for orbital angular momentum by promoting the position and momentum observables to operators. The resulting orbital angular momentum operator turns out to be rather complicated, due to a combination of the cross product and the fact that position and momentum do not commute. As a result, the components of orbital momentum do not commute with each other. When we use $[r_j, p_k] = i\hbar\delta_{jk}$, the commutation relation for the components of \mathbf{L} becomes

$$[L_i, L_j] = i\hbar\epsilon_{ijk}L_k. \quad (7.2)$$

A set of relations like this is called an *algebra*, and the algebra here is called *closed* since we can take the commutator of any two elements L_i and L_j , and express it in terms of another element L_k . Another (simpler) closed algebra is $[x, p_x] = i\hbar$ and $[x, \mathbb{I}] = [p_x, \mathbb{I}] = 0$.

Since the components of angular momentum do not commute, we cannot find simultaneous eigenstates for L_x , L_y , and L_z . We will choose one of them, traditionally denoted by L_z , and construct its eigenstates. It turns out that there is another operator, a function of all L_i s, that commutes with any component L_j , namely $\mathbf{L}^2 = L_x^2 + L_y^2 + L_z^2$. This operator is unique, in that there is no other operator that differs from \mathbf{L}^2 in a nontrivial way and still commutes with all L_i s. We can now construct simultaneous eigenvectors for L_z and \mathbf{L}^2 .

Since we are looking for simultaneous eigenvectors for the square of the angular momentum *and* the z -component, we expect that the eigenvectors will be determined by two quantum numbers, l , and m . First, and without any prior knowledge, we can formally write down the eigenvalue equation for L_z as

$$L_z |l, m\rangle = m\hbar |l, m\rangle, \quad (7.3)$$

where m is some real number, and we included \hbar to fit the dimensions of angular momentum. We will now proceed with the derivation of the eigenvalue equation for \mathbf{L}^2 , and determine the possible values for l and m .

From the definition of \mathbf{L}^2 , we have $\mathbf{L}^2 - L_z^2 = L_x^2 + L_y^2$, and

$$\langle l, m | \mathbf{L}^2 - L_z^2 | l, m \rangle = \langle l, m | L_x^2 + L_y^2 | l, m \rangle \geq 0. \quad (7.4)$$

The spectrum of L_z is therefore bounded by

$$l \leq m \leq l \quad (7.5)$$

for some value of l . We derive the eigenvalues of \mathbf{L}^2 given these restrictions. First, we define the ladder operators

$$L_{\pm} = L_x \pm iL_y \quad \text{with} \quad L_- = L_+^\dagger. \quad (7.6)$$

The commutation relations with L_z and \mathbf{L}^2 are

$$[L_z, L_{\pm}] = \pm \hbar L_{\pm}, \quad [L_+, L_-] = 2\hbar L_z, \quad [L_{\pm}, \mathbf{L}^2] = 0. \quad (7.7)$$

Next, we calculate $L_z(L_+|l, m\rangle)$:

$$\begin{aligned} L_z(L_+|l, m\rangle) &= (L_+L_z + [L_z, L_+])|l, m\rangle = m\hbar L_+|l, m\rangle + \hbar L_+|l, m\rangle \\ &= (m+1)\hbar L_+|l, m\rangle. \end{aligned} \quad (7.8)$$

Therefore $L_+|l, m\rangle \propto |l, m+1\rangle$. By similar reasoning we find that $L_-|l, m\rangle \propto |l, m-1\rangle$. Since we already determined that $-l \leq m \leq l$, we must also require that

$$L_+|l, l\rangle = 0 \quad \text{and} \quad L_-|l, -l\rangle = 0. \quad (7.9)$$

Counting the states between $-l$ and $+l$ in steps of one, we find that there are $2l+1$ different eigenstates for L_z . Since $2l+1$ is a positive integer, l must be a half-integer ($l = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \dots$). Later we will restrict this further to $l = 0, 1, 2, \dots$

The next step towards finding the eigenvalues of \mathbf{L}^2 is to calculate the following identity:

$$L_-L_+ = (L_x - iL_y)(L_x + iL_y) = L_x^2 + L_y^2 + i[L_x, L_y] = \mathbf{L}^2 - L_z^2 - \hbar L_z. \quad (7.10)$$

We can then evaluate

$$L_-L_+|l, l\rangle = 0 \Rightarrow (\mathbf{L}^2 - L_z^2 - \hbar L_z)|l, l\rangle = \mathbf{L}^2|l, l\rangle - (l^2 + l)\hbar^2|l, l\rangle = 0. \quad (7.11)$$

It is left as an exercise (see exercise 1b) to show that

$$\mathbf{L}^2|l, m\rangle = l(l+1)\hbar^2|l, m\rangle. \quad (7.12)$$

We now have derived the eigenvalues for L_z and \mathbf{L}^2 .

One aspect of our algebraic treatment of angular momentum we still have to determine is the matrix elements of the ladder operators. We again use the relation between L_{\pm} , and L_z and \mathbf{L}^2 :

$$\langle l, m | L_-L_+ | l, m \rangle = \sum_{j=-l}^l \langle l, m | L_- | l, j \rangle \langle l, j | L_+ | l, m \rangle. \quad (7.13)$$

Both sides can be rewritten as

$$\langle l, m | \mathbf{L}^2 - L_z^2 - \hbar L_z | l, m \rangle = \langle l, m | L_- | l, m+1 \rangle \langle l, m+1 | L_+ | l, m \rangle, \quad (7.14)$$

where on the right-hand-side we used that only the $m+1$ -term survives. This leads to

$$[l(l+1) - m(m+1)]\hbar^2 = |\langle l, m+1 | L_+ | l, m \rangle|^2. \quad (7.15)$$

The ladder operators then act as

$$L_+|l, m\rangle = \hbar\sqrt{l(l+1) - m(m+1)}|l, m+1\rangle, \quad (7.16)$$

and

$$L_-|l, m\rangle = \hbar\sqrt{l(l+1) - m(m-1)}|l, m-1\rangle. \quad (7.17)$$

We have seen that the angular momentum L is quantized, and that this gives rise to a discrete state space parameterized by the quantum numbers l and m . However, we still have to restrict the values of l further, as mentioned above. We cannot do this using only the algebraic approach (i.e., using the commutation relations for L_i), and we have to consider the spatial properties of angular momentum. To this end, we write L_i as

$$L_i = -i\hbar\epsilon_{ijk}\left(x_j\frac{\partial}{\partial x_k}\right), \quad (7.18)$$

which follows directly from the promotion of \mathbf{r} and \mathbf{p} in Eq. (7.1) to quantum mechanical operators. In spherical coordinates,

$$r = \sqrt{x^2 + y^2 + z^2}, \quad \phi = \arctan\left(\frac{y}{x}\right), \quad \theta = \arctan\left(\frac{\sqrt{x^2 + y^2}}{z}\right), \quad (7.19)$$

the angular momentum operators can be written as

$$\begin{aligned} L_x &= -i\hbar\left(-\sin\phi\frac{\partial}{\partial\theta} - \cot\theta\cos\phi\frac{\partial}{\partial\phi}\right), \\ L_y &= -i\hbar\left(\cos\phi\frac{\partial}{\partial\theta} - \cot\theta\sin\phi\frac{\partial}{\partial\phi}\right), \\ L_z &= -i\hbar\frac{\partial}{\partial\phi}, \\ \mathbf{L}^2 &= -\hbar^2\left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\phi^2}\right]. \end{aligned} \quad (7.20)$$

The eigenvalue equation for L_z then becomes

$$L_z\psi(r, \theta, \phi) = -i\hbar\frac{\partial}{\partial\phi}\psi(r, \theta, \phi) = m\hbar\psi(r, \theta, \phi) \quad (7.21)$$

We can solve this differential equation to find that

$$\psi(r, \theta, \phi) = \zeta(r, \theta)e^{im\phi}. \quad (7.22)$$

A spatial rotation over 2π must return the wave function to its original value, because $\psi(r, \theta, \phi)$ must have a unique value at each point in space. This leads to $\psi(r, \theta, \phi + 2\pi) = \psi(r, \theta, \phi)$ and

$$e^{im(\phi+2\pi)} = e^{im\phi}, \quad \text{or} \quad e^{2\pi im} = 1. \quad (7.23)$$

This means that m is an integer, which in turn means that l must be an integer also.

7.2 Spin

For orbital angular momentum we found that $2l + 1$ must be an integer, and moreover the spatial properties of the wave function force l to be an integer as well. However, we can also construct states with half-integer l , but this must then be an *internal* degree of freedom. This is called spin angular momentum, or spin for short. We will show later in the course that the spin observable is interpreted as an intrinsic magnetic moment of a system.

To describe spin, we switch from \mathbf{L} to \mathbf{S} , which is no longer related to \mathbf{r} and \mathbf{p} . The commutation relations between the components S_i are the same as for L_i ,

$$[S_i, S_j] = i\hbar\epsilon_{ijk}S_k, \quad (7.24)$$

so \mathbf{S} and \mathbf{L} obey the same algebra. The commutation relations between \mathbf{S} and \mathbf{L} , \mathbf{r} , and \mathbf{p} vanish:

$$[S_i, L_j] = [S_i, r_j] = [S_i, p_j] = 0. \quad (7.25)$$

Therefore, spin generates a whole new vector space, since it commutes with observables that themselves do not commute (like r_j and p_j), and it is independent of the spatial degrees of freedom.

Since the commutation relations for \mathbf{S} (its algebra) are the same as for \mathbf{L} , we can immediately copy the algebraic structure of the eigenstates and eigenvalues:

$$\begin{aligned} S_z |s, m_s\rangle &= m_s \hbar |s, m_s\rangle, \quad \text{with } s = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \dots \\ \mathbf{S}^2 |s, m_s\rangle &= s(s+1) \hbar^2 |s, m_s\rangle. \end{aligned} \quad (7.26)$$

When $s = \frac{1}{2}$, the system has two levels (a qubit) with spin eigenstates $|\frac{1}{2}, \frac{1}{2}\rangle$ and $|\frac{1}{2}, -\frac{1}{2}\rangle$. We often write $m_s = +\frac{1}{2} = \uparrow$ (“up”) and $m_s = -\frac{1}{2} = \downarrow$ (“down”), which finds its origin in the measurement outcomes of electron spin in a Stern-Gerlach apparatus.

Now that we have introduced a whole new vector space related to spin, how do we write the wave function of a particle with spin? Without spin, the wave function is a normal single-valued function $\psi(\mathbf{r}, t) = \langle \mathbf{r} | \psi(t) \rangle$ of space and time coordinates. Now we have to add the spin degree of freedom. For each spin (\uparrow or \downarrow when $s = \frac{1}{2}$), we have a wave function $\psi_{\uparrow}(\mathbf{r}, t)$ for the particle with spin up, and $\psi_{\downarrow}(\mathbf{r}, t)$ for the particle with spin down. We can write this as a vector:

$$\psi(\mathbf{r}, t) = \begin{pmatrix} \psi_{\uparrow}(\mathbf{r}, t) \\ \psi_{\downarrow}(\mathbf{r}, t) \end{pmatrix}. \quad (7.27)$$

The spin degree of freedom generates a vector space, after all. The vector ψ is called a *spinor*.

Expectation values are evaluated in the usual way, but now we have to sum over the spin degree of freedom, as well as integrate over space. For example, the probability of finding a particle with spin up in a region Ω of space is given by

$$p(\uparrow, \Omega) = \sum_{m_s=\uparrow, \downarrow} \int_{\Omega} d\mathbf{r} \delta_{m_s, \uparrow} |\psi_{m_s}(\mathbf{r}, t)|^2 = \int_{\Omega} d\mathbf{r} |\psi_{\uparrow}(\mathbf{r}, t)|^2, \quad (7.28)$$

and the expectation value of finding a particle with *any* spin in a region Ω of space is given by

$$p(\Omega) = \sum_{m_s=\uparrow, \downarrow} \int_{\Omega} d\mathbf{r} |\psi_{m_s}(\mathbf{r}, t)|^2. \quad (7.29)$$

The normalization of the spinor is such that

$$\sum_{m_s=\uparrow, \downarrow} \int_V d\mathbf{r} |\psi_{m_s}(\mathbf{r}, t)|^2 = 1, \quad (7.30)$$

where V is the entire space available to the particle (this may be the entire universe, or the volume of a box with impenetrable walls, etc.).

If spin is represented by $(2s+1)$ -dimensional spinors (vectors), then spin transformations (operators) are represented by $(2s+1) \times (2s+1)$ matrices. In the two-dimensional case, we have by construction:

$$S_z |\uparrow\rangle = \frac{\hbar}{2} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad S_z |\downarrow\rangle = -\frac{\hbar}{2} \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad (7.31)$$

which means that the matrix representation of S_z is given by

$$S_z = \frac{\hbar}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (7.32)$$

Next, the ladder operators act according to

$$S_+ |\uparrow\rangle = 0, \quad S_+ |\downarrow\rangle = \hbar |\uparrow\rangle, \quad S_- |\uparrow\rangle = \hbar |\downarrow\rangle, \quad S_- |\downarrow\rangle = 0, \quad (7.33)$$

which leads to the matrix representation

$$S_+ = \hbar \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad S_- = \hbar \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}. \quad (7.34)$$

From $S_{\pm} = S_x \pm iS_y$ we can then deduce that

$$S_x = \frac{\hbar}{2} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \text{and} \quad S_y = \frac{\hbar}{2} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}. \quad (7.35)$$

We often define $S_i \equiv \frac{1}{2}\hbar\sigma_i$, where σ_i are the so-called *Pauli matrices*. Previously, we have called these matrices X , Y , and Z . The commutation relations of the Pauli matrices are

$$[\sigma_i, \sigma_j] = 2i\epsilon_{ijk}\sigma_k \quad \text{or} \quad \left[\frac{\sigma_i}{2}, \frac{\sigma_j}{2} \right] = i\epsilon_{ijk} \frac{\sigma_k}{2}. \quad (7.36)$$

Other important properties of the Pauli matrices are

$$\{\sigma_i, \sigma_j\} \equiv \sigma_i\sigma_j + \sigma_j\sigma_i = 2\delta_{ij}\mathbb{I} \quad (\text{anti-commutator}). \quad (7.37)$$

They are both Hermitian and unitary, and the square of the Pauli matrices is the identity: $\sigma_i^2 = \mathbb{I}$. Moreover, they obey an “orthogonality” relation

$$\frac{1}{2}\text{Tr}(\sigma_i\sigma_j) = \delta_{ij}. \quad (7.38)$$

The proof of this statement is as follows:

$$\sigma_i\sigma_j = \sigma_i\sigma_j + \sigma_j\sigma_i - \sigma_j\sigma_i = \{\sigma_i, \sigma_j\} - \sigma_j\sigma_i = 2\delta_{ij}\mathbb{I} - \sigma_j\sigma_i. \quad (7.39)$$

Taking the trace then yields

$$\text{Tr}(2\delta_{ij}\mathbb{I} - \sigma_j\sigma_i) = \text{Tr}(\sigma_i\sigma_j) = \text{Tr}(\sigma_j\sigma_i), \quad (7.40)$$

or (using $\text{Tr}(\mathbb{I}) = 2$)

$$2\text{Tr}(\sigma_j\sigma_i) = 4\delta_{ij}, \quad (7.41)$$

which proves Eq. (7.38). If we define $\sigma_0 \equiv \mathbb{I}$, we can extend this proof to the four-dimensional case

$$\frac{1}{2}\text{Tr}(\sigma_\mu\sigma_\nu) = \delta_{\mu\nu}, \quad (7.42)$$

with $\mu, \nu = 0, 1, 2, 3$.

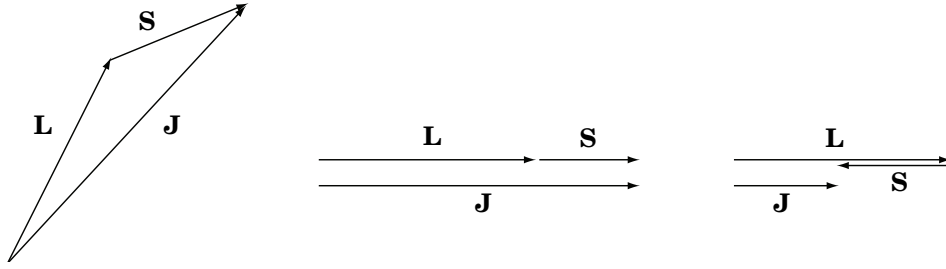


Figure 3: Addition of angular momentum.

We can then write any 2×2 matrix as a sum over the two-dimensional Pauli operators:

$$A = \sum_{\mu} a_{\mu} \sigma_{\mu}, \quad (7.43)$$

since

$$\frac{1}{2} \text{Tr}(A \sigma_{\nu}) = \frac{1}{2} \text{Tr} \left(\sum_{\mu} a_{\mu} \sigma_{\mu} \sigma_{\nu} \right) = \frac{1}{2} \sum_{\mu} a_{\mu} \text{Tr}(\sigma_{\mu} \sigma_{\nu}) = \sum_{\mu} a_{\mu} \delta_{\mu\nu} = a_{\nu}. \quad (7.44)$$

The Pauli matrices and the identity matrix form a *basis* for the 2×2 matrices, and we can write

$$A = a_0 \mathbb{I} + \mathbf{a} \cdot \boldsymbol{\sigma} = \begin{pmatrix} a_0 + a_z & a_x - i a_y \\ a_x + i a_y & a_0 - a_z \end{pmatrix}, \quad (7.45)$$

where we used the notation $\boldsymbol{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$.

7.3 Total angular momentum

In general, a particle may have both spin and orbital angular momentum. Since \mathbf{L} and \mathbf{S} have the same dimensions, we can ask what is the *total* angular momentum \mathbf{J} of the particle. We write this as

$$\mathbf{J} = \mathbf{L} + \mathbf{S} \equiv \mathbf{L} \otimes \mathbb{I} + \mathbb{I} \otimes \mathbf{S}, \quad (7.46)$$

which emphasizes that orbital and spin angular momentum are described in distinct Hilbert spaces.

Since $[L_i, S_j] = 0$, we have

$$\begin{aligned} [J_i, J_j] &= [L_i + S_i, L_j + S_j] = [L_i, L_j] + [S_i, S_j] \\ &= i\hbar \epsilon_{ijk} L_k + i\hbar \epsilon_{ijk} S_k = i\hbar \epsilon_{ijk} (L_k + S_k) \\ &= i\hbar \epsilon_{ijk} J_k. \end{aligned} \quad (7.47)$$

In other words, \mathbf{J} obeys the same algebra as \mathbf{L} and \mathbf{S} , and we can immediately carry over the structure of the eigenvalues and eigenvectors from \mathbf{L} and \mathbf{S} .

In addition, \mathbf{L} and \mathbf{S} must be added as vectors. However, only one of the components of the total angular momentum can be sharp (i.e., having a definite value). Recall that l and s are *magnitudes* of the orbital and spin angular momentum, respectively. We can determine the

extremal values of \mathbf{J} , denoted by $\pm j$, by adding and subtracting the spin from the orbital angular momentum, as shown in Figure 3:

$$|l - s| \leq j \leq l + s. \quad (7.48)$$

For example, when $l = 1$ and $s = \frac{1}{2}$, the possible values of j are $j = \frac{1}{2}$ and $j = \frac{3}{2}$.

The commuting operators for \mathbf{J} are, first of all, \mathbf{J}^2 and J_z as we expect from the algebra, but also the operators \mathbf{L}^2 and \mathbf{S}^2 . You may think that S_z and L_z also commute with these operators, but that it not the case:

$$[\mathbf{J}^2, L_z] = [(\mathbf{L} + \mathbf{S})^2, L_z] = [\mathbf{L}^2 + 2\mathbf{L} \cdot \mathbf{S} + \mathbf{S}^2, L_z] = 2[\mathbf{L}, L_z] \cdot \mathbf{S} \neq 0. \quad (7.49)$$

We can construct a full basis for total angular momentum in terms of \mathbf{J}^2 and J_z , as before:

$$\mathbf{J}^2 |j, m_j\rangle = \hbar^2 j(j+1) |j, m_j\rangle \quad \text{and} \quad J_z |j, m_j\rangle = m_j \hbar |j, m_j\rangle. \quad (7.50)$$

Alternatively, we can construct spin and orbital angular momentum eigenstates directly as a tensor product of the eigenstates

$$\mathbf{L}^2 |l, m\rangle |s, m_s\rangle = \hbar^2 l(l+1) |l, m\rangle |s, m_s\rangle \quad \text{and} \quad L_z |l, m\rangle |s, m_s\rangle = m \hbar |l, m\rangle |s, m_s\rangle, \quad (7.51)$$

and

$$\mathbf{S}^2 |l, m\rangle |s, m_s\rangle = \hbar^2 s(s+1) |l, m\rangle |s, m_s\rangle \quad \text{and} \quad S_z |l, m\rangle |s, m_s\rangle = m_s \hbar |l, m\rangle |s, m_s\rangle. \quad (7.52)$$

Since the L_z and S_z do not commute with \mathbf{J}^2 , the states $|j, m_j\rangle$ are *not the same* as the states $|l, m\rangle |s, m_s\rangle$.

7.4 Composite systems with angular momentum

Now consider two systems, 1 and 2, with total angular momentum \mathbf{J}_1 and \mathbf{J}_2 , respectively. The total angular momentum is again additive, and given by

$$\mathbf{J} = \mathbf{J}_1 + \mathbf{J}_2 \equiv \mathbf{J}_1 \otimes \mathbb{I} + \mathbb{I} \otimes \mathbf{J}_2. \quad (7.53)$$

Completely analogous to the addition of spin and orbital angular momentum, we can construct the commuting operators \mathbf{J}^2 , J_z , \mathbf{J}_1^2 , and \mathbf{J}_2^2 , but not J_{1z} and J_{2z} . Again, we construct two natural bases for the total angular momentum of the composite system, namely

$$\{|j, m\rangle\} \quad \text{or} \quad \{|j_1, m_1\rangle \otimes |j_2, m_2\rangle\} \equiv \{|j_1, j_2, m_1, m_2\rangle\}. \quad (7.54)$$

We want to know how the two bases relate to each other, because sometimes we wish to talk about the angular momentum of the composite system, and at other times we are interested in the angular momentum of the subsystems. Since the second basis (as well as the first) in Eq. (7.54) forms a complete orthonormal basis, we can write

$$|j, m\rangle = \sum_{m_1, m_2} |j_1, j_2, m_1, m_2\rangle \langle j_1, j_2, m_1, m_2 | j, m \rangle. \quad (7.55)$$

The amplitudes $\langle j_1, j_2, m_1, m_2 | j, m \rangle$ are called *Clebsch-Gordan* coefficients, and we will now present a general procedure for calculating them.

Let's first consider a simple example of two spin- $\frac{1}{2}$ systems, such as two electrons. The spin basis for each electron is given by $|\frac{1}{2}, \frac{1}{2}\rangle = |\uparrow\rangle$ and $|\frac{1}{2}, -\frac{1}{2}\rangle = |\downarrow\rangle$. The spin basis for the two electrons is therefore

$$|j_1, j_2, m_1, m_2\rangle \in \{|\uparrow, \uparrow\rangle, |\uparrow, \downarrow\rangle, |\downarrow, \uparrow\rangle, |\downarrow, \downarrow\rangle\}. \quad (7.56)$$

The total spin is given by $j = \frac{1}{2} + \frac{1}{2} = 1$ and $j = \frac{1}{2} - \frac{1}{2} = 0$, so the four basis states for total angular momentum are

$$|j, m\rangle \in \{|1, 1\rangle, |1, 0\rangle, |1, -1\rangle, |0, 0\rangle\}. \quad (7.57)$$

The latter state is the eigenstate for $j = 0$. The maximum total angular momentum state $|1, 1\rangle$ can occur only when the two electron spins are parallel, and we therefore have

$$|1, 1\rangle = |\uparrow, \uparrow\rangle = |\frac{1}{2}, \frac{1}{2}\rangle \otimes |\frac{1}{2}, \frac{1}{2}\rangle. \quad (7.58)$$

To find the expansion of the other total angular momentum eigenstates in terms of spin eigenstates we employ the following trick: use that $J_{\pm} = J_{1\pm} + J_{2\pm}$. We can then apply J_{\pm} to the state $|1, 1\rangle$, and $J_{1\pm} + J_{2\pm}$ to the state $|\frac{1}{2}, \frac{1}{2}\rangle \otimes |\frac{1}{2}, \frac{1}{2}\rangle$. This yields

$$J_- |1, 1\rangle = \hbar \sqrt{j(j+1) - m(m-1)} |1, 0\rangle = \hbar \sqrt{2} |1, 0\rangle. \quad (7.59)$$

Similarly, we calculate

$$J_{1-} |\frac{1}{2}, \frac{1}{2}\rangle = \hbar \sqrt{\frac{1}{2}(\frac{3}{2}) - \frac{1}{2}(-\frac{1}{2})} |\frac{1}{2}, -\frac{1}{2}\rangle = \hbar |\frac{1}{2}, -\frac{1}{2}\rangle, \quad (7.60)$$

and a similar result for J_{2-} . Therefore, we find that

$$\hbar \sqrt{2} |1, 0\rangle = \hbar |\uparrow, \downarrow\rangle + \hbar |\downarrow, \uparrow\rangle \implies |1, 0\rangle = \frac{|\uparrow, \downarrow\rangle + |\downarrow, \uparrow\rangle}{\sqrt{2}}. \quad (7.61)$$

Applying J_- again yields

$$|1, -1\rangle = |\downarrow, \downarrow\rangle. \quad (7.62)$$

This agrees with the construction of adding parallel spins. The three total angular momentum states

$$\begin{aligned} |1, 1\rangle &= |\uparrow, \uparrow\rangle, \\ |1, 0\rangle &= \frac{1}{\sqrt{2}} (|\uparrow, \downarrow\rangle + |\downarrow, \uparrow\rangle), \\ |1, -1\rangle &= |\downarrow, \downarrow\rangle \end{aligned} \quad (7.63)$$

form a so-called *triplet* of states with $j = 1$. We now have to find the final state corresponding to $j = 0, m = 0$. The easiest way to find it at this point is to require orthonormality of the four basis states, and this gives us the *singlet* state

$$|0, 0\rangle = \frac{1}{\sqrt{2}} (|\uparrow, \downarrow\rangle - |\downarrow, \uparrow\rangle). \quad (7.64)$$

The singlet state has zero total angular momentum, and it is therefore invariant under rotations.

In general, this procedure of finding the Clebsch-Gordan coefficients results in *multiplets* of constant j . In the case of two spins, we have a tensor product of two two-dimensional spaces, which are decomposed in two subspaces of dimension 3 (the triplet) and 1 (the singlet), respectively. We write this symbolically as

$$2 \otimes 2 = 3 \oplus 1. \quad (7.65)$$

If we had combined a spin 1 particle with a spin $\frac{1}{2}$ particle, the largest multiplet would have been due to $j = 1 + \frac{1}{2} = \frac{3}{2}$, which is a 4-dimensional subspace, and the smallest subspace is due to $j = 1 - \frac{1}{2} = \frac{1}{2}$, which is a two-dimensional subspace:

$$3 \otimes 2 = 4 \oplus 2. \quad (7.66)$$

In general, the total angular momentum of two systems with angular momentum k and l is decomposed into multiplets according to the following rule ($k \geq l$):

$$(2k+1) \otimes (2l+1) = [2(k+l)+1] \oplus [2(k+l)-1] \oplus \dots \oplus [2(k-l)+1], \quad (7.67)$$

or in terms of the dimensions of the subspaces ($n \geq m$):

$$n \otimes m = (n+m-1) \oplus (n+m-3) \oplus \dots \oplus (n-m+1). \quad (7.68)$$

Exercises

1. Angular momentum algebra.

- Prove the algebra given in Eq. (7.2). Also show that $[L^2, L_i] = 0$, and verify the commutation relations in Eq. (7.7).
- Show that $L^2 |l, m\rangle = l(l+1)\hbar^2 |l, m\rangle$. Use the fact that $[L_-, L^2] = 0$.

2. Pauli matrices.

- Check that the matrix representation of the spin- $\frac{1}{2}$ operators obey the commutation relations.
- Calculate the matrix representation of the Pauli matrices for $s = 1$.
- Prove that $\exp[-i\boldsymbol{\theta} \cdot \boldsymbol{\sigma}]$ is a 2×2 unitary matrix.

3. Isospin \mathbf{I} describes certain particle families called *multiplets*, and the components of the isospin obey the commutation relations $[I_i, I_j] = i\epsilon_{ijk} I_k$.

- What is the relation between spin and isospin?
- Organize the nucleons (proton and neutron), the pions (π^+ , π^0 , and π^-), and the delta baryons (Δ^{++} , Δ^+ , Δ^0 , and Δ^-) into multiplets. You will have to determine their isospin quantum number.
- Give all possible decay channels of the delta baryons into pions and nucleons (use charge and baryon number conservation).
- Calculate the relative decay ratios of Δ^+ and Δ^0 into the different channels.

4. A simple atom has orbital and spin angular momentum, and the Hamiltonian for the atom contains a spin-orbit coupling term $H_{so} = g\hbar \mathbf{L} \cdot \mathbf{S}$, where $g\hbar$ is the coupling strength.

- (a) Are orbital and spin angular momentum good quantum numbers for this system? What about total angular momentum?
- (b) Use first-order perturbation theory to calculate the energy shift due to the spin-orbit coupling term.
- (c) Calculate the transition matrix elements of H_{so} in the basis $\{|l, m; s, m_s\rangle\}$.

5. Multiplets.

- (a) A spin $\frac{3}{2}$ particle and a spin 2 particle form a composite system. How many multiplets are there, and what is the dimension of the largest multiplet?
- (b) How many multiplets do two systems with equal angular momentum have?

8 Identical Particles

We have so far looked at the quantum mechanical description of a few particles with spin in the previous section, and particles that exhibit entanglement in section 5. In all these cases, we assumed that the individual particles could be distinguished from each other. For example, the two-electron state $(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle)/\sqrt{2}$ assumes that we have two electrons, one held “over here”, and the other “over there”, and we can talk meaningfully about their respective spins. The tensor product structure of our Hilbert space is a manifestation of our ability to label our particles unambiguously.

However, what happens when we place the two electrons inside a sealed box? The wave functions of the electrons will quickly start to overlap. Since the electrons are *identical* particles, which according to basic quantum mechanics do not have well-defined paths, we cannot keep track of which electron is which inside the box. Not even in principle.

8.1 Symmetric and anti-symmetric states

The indistinguishability of identical particles means that we have to adjust our quantum mechanical description of these objects. There are two ways of doing this, namely via a modification of the allowed states and via a restructuring of the observables⁵. In this section we consider the restricted state space, and in the next we will be considering the new observables.

First of all, since the total number of particles is an observable quantity (for example by measuring the total charge in the box), we can give the particles an artificial labelling. The wave functions of the two particles are then given by $|\psi(\mathbf{r}_1)\rangle_1$ for particle 1 at position \mathbf{r}_1 , and $|\phi(\mathbf{r}_2)\rangle_2$ for particle 2 at position \mathbf{r}_2 . Since we can swap the positions of the particle without observable consequences, we find that there are two states that denote the same physical situation:

$$|\psi(\mathbf{r}_1), \phi(\mathbf{r}_2)\rangle_{12} \quad \text{and} \quad |\psi(\mathbf{r}_2), \phi(\mathbf{r}_1)\rangle_{12}. \quad (8.1)$$

However, we wish that each physically distinct situation has exactly *one* quantum state. Since there is no preference for either state, we can denote the physical situation of identical particles at position \mathbf{r}_1 and \mathbf{r}_2 by the quantum state that is an *equal weight* over these two possibilities:

$$|\Psi(\mathbf{r}_1, \mathbf{r}_2)\rangle_{12} = \frac{|\psi(\mathbf{r}_1), \phi(\mathbf{r}_2)\rangle_{12} + e^{i\varphi} |\psi(\mathbf{r}_2), \phi(\mathbf{r}_1)\rangle_{12}}{\sqrt{2}}. \quad (8.2)$$

You can verify that swapping the positions \mathbf{r}_1 and \mathbf{r}_2 of the indistinguishable particles incurs only a global (unobservable) phase. The question is now how we should choose φ .

Suppose that the two identical particles in the box are electrons. We know from Pauli’s exclusion principle that the two electrons cannot be in the same state. Therefore, when $\phi = \psi$, the state in Eq. (8.2) should naturally disappear:

$$|\Psi(\mathbf{r}_1, \mathbf{r}_2)\rangle_{12} = \frac{|\psi(\mathbf{r}_1), \psi(\mathbf{r}_2)\rangle_{12} + e^{i\varphi} |\psi(\mathbf{r}_2), \psi(\mathbf{r}_1)\rangle_{12}}{\sqrt{2}} = 0, \quad (8.3)$$

which means that for particles obeying Pauli’s exclusion principle we must choose $e^{i\varphi} = -1$. The quantum state of the two particles is *anti-symmetric*.

What about particles that do not obey Pauli’s exclusion principle? These must be restricted to states that are orthogonal to the anti-symmetric states. In other words, they must be in states

⁵This is sometimes called *second quantisation*. This is a misnomer, since quantisation occurs only once, when observables are promoted to operators.

that are *symmetric* under the exchange of two particles. For the two identical particles in a box, we therefore choose the value $e^{i\varphi} = +1$, which makes the state orthogonal to the anti-symmetric state. The two possibilities for combining two identical particles are therefore

$$\begin{aligned} |\Psi_S(\mathbf{r}_1, \mathbf{r}_2)\rangle &= \frac{|\psi(\mathbf{r}_1), \phi(\mathbf{r}_2)\rangle + |\psi(\mathbf{r}_2), \phi(\mathbf{r}_1)\rangle}{\sqrt{2}}, \\ |\Psi_A(\mathbf{r}_1, \mathbf{r}_2)\rangle &= \frac{|\psi(\mathbf{r}_1), \phi(\mathbf{r}_2)\rangle - |\psi(\mathbf{r}_2), \phi(\mathbf{r}_1)\rangle}{\sqrt{2}}. \end{aligned} \quad (8.4)$$

These states include both the internal degrees of freedom, such as spin, and the external degrees of freedom. So two electrons can still be in the state $|\uparrow\uparrow\rangle$, as long as their spatial wave function is anti-symmetric. The particles that are in a symmetric overall quantum state are bosons, while the particles in an overall anti-symmetric state are fermions.

We can extend this to N particles in a fairly straightforward manner. For bosons, we sum over all possible permutations of \mathbf{r}_1 to \mathbf{r}_N :

$$|\Psi_S(\mathbf{r}_1, \dots, \mathbf{r}_N)\rangle = \frac{1}{\sqrt{N!}} \sum_{\text{perm}(\mathbf{r}_1, \dots, \mathbf{r}_N)} |\psi_1(\mathbf{r}_1), \dots, \psi_N(\mathbf{r}_N)\rangle. \quad (8.5)$$

For fermions, the odd permutations pick up a relative minus sign:

$$|\Psi_A(\mathbf{r}_1, \dots, \mathbf{r}_N)\rangle = \frac{1}{\sqrt{N!}} \sum_{\text{even}} |\psi_1(\mathbf{r}_1), \dots, \psi_N(\mathbf{r}_N)\rangle - \frac{1}{\sqrt{N!}} \sum_{\text{odd}} |\psi_1(\mathbf{r}_1), \dots, \psi_N(\mathbf{r}_N)\rangle. \quad (8.6)$$

This can be written compactly as the so-called *Slater determinant*

$$\Psi_A(\mathbf{r}_1, \dots, \mathbf{r}_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \psi_1(\mathbf{r}_1) & \psi_1(\mathbf{r}_2) & \dots & \psi_1(\mathbf{r}_N) \\ \psi_2(\mathbf{r}_1) & \psi_2(\mathbf{r}_2) & \dots & \psi_2(\mathbf{r}_N) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_N(\mathbf{r}_1) & \psi_N(\mathbf{r}_2) & \dots & \psi_N(\mathbf{r}_N) \end{vmatrix}, \quad (8.7)$$

where we removed the kets for notational convenience. The N particles in the state $|\Psi_A(\mathbf{r}_1, \dots, \mathbf{r}_N)\rangle$ automatically obey the Pauli exclusion principle.

8.2 Creation and annihilation operators

The second, and particularly powerful way to implement the description of identical particles is via creation and annihilation operators. To see how this description arises, consider some single-particle Hermitian operator A with eigenvalues a_j . On physical grounds, and regardless of distinguishability, we require that n_j particles in the eigenstate $|a_j\rangle$ of A must have a total physical value $n_j \times a_j$ for the observable A . We can repeat this for all eigenvalues a_j , and obtain a potentially infinite set of basis vectors

$$|n_1, n_2, n_3, \dots\rangle,$$

for all integer values of n_j , including zero. You should convince yourself that this exhausts all the possible ways any number of particles can be distributed over the eigenvalues a_j . The spectrum of A can be bounded or unbounded, and discrete or continuous. It may even be degenerate. For simplicity we consider here an unbounded, non-degenerate discrete spectrum.

A special state is given by

$$|\emptyset\rangle = |0, 0, 0, \dots\rangle, \quad (8.8)$$

which indicates the state of no particles, or the *vacuum*. The numbers n_j are called the occupation number, and any physical state can be written as a superposition of these states:

$$|\Psi\rangle = \sum_{n_1, n_2, n_3, \dots=0}^{\infty} c_{n_1, n_2, n_3, \dots} |n_1, n_2, n_3, \dots\rangle. \quad (8.9)$$

The basis states $|n_1, n_2, n_3, \dots\rangle$ span a linear vector space called a *Fock space* \mathcal{F} . It is the direct sum of the Hilbert spaces for zero particles \mathcal{H}_0 , one particle \mathcal{H}_1 , two particles, etc.:

$$\mathcal{F} = \mathcal{H}_0 \oplus \mathcal{H}_1 \oplus \mathcal{H}_2 \oplus \mathcal{H}_3 \oplus \dots \quad (8.10)$$

Since $|\Psi\rangle$ is now a superposition over different particle numbers, we require operators that change the particle number. These are the creation and annihilation operators, \hat{a}^\dagger and \hat{a} respectively. Up to a proportionality constant that we will determine later, the action of these operators is defined by

$$\begin{aligned} \hat{a}_j^\dagger |n_1, n_2, \dots, n_j, \dots\rangle &\propto |n_1, n_2, \dots, n_j + 1, \dots\rangle, \\ \hat{a}_j |n_1, n_2, \dots, n_j, \dots\rangle &\propto |n_1, n_2, \dots, n_j - 1, \dots\rangle. \end{aligned} \quad (8.11)$$

So the operator \hat{a}_j^\dagger creates a particle in a state with eigenvalue a_j , and the operator \hat{a}_j removes a particle in a state with eigenvalue a_j . These operators are each others' Hermitian adjoint, since removing a particle is the time reversal of adding a particle. Clearly, when an annihilation operator attempts to remove particles that are not there, the result must be zero:

$$\hat{a}_j |n_1, n_2, \dots, n_j = 0, \dots\rangle = 0. \quad (8.12)$$

The vacuum is then defined as the state that gives zero when acted on by any annihilation operator: $\hat{a}_j |\emptyset\rangle = 0$ for any j . Notice how we have so far sidestepped the problem of particle swapping; we exclusively used aspects of the *total* particle number.

What are the basic properties of these creation and annihilation operators? In particular, we are interested in their commutation relations. We will now derive these properties from what we have determined so far. First, note that we can create two particles with eigenvalues a_i and a_j in the system in any order, and the only difference this can make is in the normalisation of the state:

$$\hat{a}_i^\dagger \hat{a}_j^\dagger |\Psi\rangle = \lambda \hat{a}_j^\dagger \hat{a}_i^\dagger |\Psi\rangle, \quad (8.13)$$

where λ is some complex number. Since state $|\Psi\rangle$ is certainly not zero, we require that

$$\hat{a}_k^\dagger \hat{a}_l^\dagger - \lambda \hat{a}_l^\dagger \hat{a}_k^\dagger = 0. \quad (8.14)$$

Since k and l are just dummy variables, we equally have

$$\hat{a}_l^\dagger \hat{a}_k^\dagger - \lambda \hat{a}_k^\dagger \hat{a}_l^\dagger = 0. \quad (8.15)$$

We now substitute Eq. (8.15) into Eq. (8.14) to eliminate $\hat{a}_l^\dagger \hat{a}_k^\dagger$. This leads to

$$(1 - \lambda^2) \hat{a}_k^\dagger \hat{a}_l^\dagger = 0, \quad (8.16)$$

and therefore

$$\lambda = \pm 1. \quad (8.17)$$

The relation between different creation operators can thus take two forms. They can obey a commutation relation when $\lambda = +1$:

$$\hat{a}_l^\dagger \hat{a}_k^\dagger - \hat{a}_k^\dagger \hat{a}_l^\dagger = [\hat{a}_l^\dagger, \hat{a}_k^\dagger] = 0, \quad (8.18)$$

or they can obey an anti-commutation relation when $\lambda = -1$:

$$\hat{a}_l^\dagger \hat{a}_k^\dagger + \hat{a}_k^\dagger \hat{a}_l^\dagger = \{\hat{a}_l^\dagger, \hat{a}_k^\dagger\} = 0, \quad (8.19)$$

While creating the particles in different temporal order is not the same as swapping two particles, it should not come as a surprise that there are two possible situations (the commutation relation and the anti-commutation relation). We encountered two possibilities in our previous approach as well, where we found that many-particle states are either symmetric or anti-symmetric. In fact, creation operators that obey the commutation relation produce symmetric states, while creation operators that obey the anti-commutation relation produce anti-symmetric states. We also see that the creation operators described by the anti-commutation relations naturally obey Pauli's exclusion principle. Suppose that we wish to create two identical particles in the same eigenstate $|a_j\rangle$. The anti-commutation relations say that $\{\hat{a}_j^\dagger, \hat{a}_j^\dagger\} = 0$, so

$$\hat{a}_j^{\dagger 2} = 0. \quad (8.20)$$

Any higher powers of \hat{a}_j^\dagger will also be zero, and we can create at most one particle in the state $|a_j\rangle$.

Taking the adjoint of the commutation relations for the creation operators gives us the corresponding relations for the annihilation operators

$$\hat{a}_l \hat{a}_k - \hat{a}_k \hat{a}_l = [\hat{a}_l, \hat{a}_k] = 0, \quad (8.21)$$

or

$$\hat{a}_l \hat{a}_k + \hat{a}_k \hat{a}_l = \{\hat{a}_l, \hat{a}_k\} = 0. \quad (8.22)$$

The remaining question is now what the (anti-) commutation relations are for products of creation and annihilation operators.

We proceed along similar lines as before. Consider the operators \hat{a}_j and \hat{a}_k^\dagger with $j \neq k$, and apply them in different orders to a state $|\Psi\rangle$.

$$\hat{a}_i \hat{a}_j^\dagger |\Psi\rangle = \mu \hat{a}_j^\dagger \hat{a}_i |\Psi\rangle. \quad (8.23)$$

The same argumentation as before leads to $\mu = \pm 1$. For different j and k we therefore find

$$[\hat{a}_j, \hat{a}_k^\dagger] = 0 \quad \text{or} \quad \{\hat{a}_j, \hat{a}_k^\dagger\} = 0. \quad (8.24)$$

Now let's consider the case $j = k$. For the special case where $|\Psi\rangle = |\emptyset\rangle$, we find

$$(\hat{a}_j \hat{a}_k^\dagger - \mu \hat{a}_j^\dagger \hat{a}_k) |\emptyset\rangle = \hat{a}_j \hat{a}_k^\dagger |\emptyset\rangle = \delta_{jk} |\emptyset\rangle, \quad (8.25)$$

based on the property that $\hat{a}_j |\emptyset\rangle = 0$. When $l = k$,

$$(\hat{a}_k \hat{a}_k^\dagger - \mu \hat{a}_k^\dagger \hat{a}_k) |\emptyset\rangle = |\emptyset\rangle, \quad (8.26)$$

we find for the two possible values of μ

$$\hat{a}_k \hat{a}_k^\dagger - \hat{a}_k^\dagger \hat{a}_k = 1 \quad \text{or} \quad \hat{a}_k \hat{a}_k^\dagger + \hat{a}_k^\dagger \hat{a}_k = 1 \quad (8.27)$$

which is equivalent to

$$[\hat{a}_k, \hat{a}_k^\dagger] = 1 \quad \text{or} \quad \{\hat{a}_k, \hat{a}_k^\dagger\} = 1. \quad (8.28)$$

To summarise, we have two sets of algebras for the creation and annihilation operators. The algebra in terms of the commutation relations is given by

$$[\hat{a}_k, \hat{a}_l] = [\hat{a}_k^\dagger, \hat{a}_l^\dagger] = 0 \quad \text{and} \quad [\hat{a}_k, \hat{a}_l^\dagger] = \delta_{kl}. \quad (8.29)$$

This algebra describes particles that obey Bose-Einstein statistics, or bosons. The algebra in terms of anti-commutation relations is given by

$$\{\hat{a}_k, \hat{a}_l\} = \{\hat{a}_k^\dagger, \hat{a}_l^\dagger\} = 0 \quad \text{and} \quad \{\hat{a}_k, \hat{a}_l^\dagger\} = \delta_{kl}. \quad (8.30)$$

This algebra describes particles that obey Fermi-Dirac statistics, or fermions.

Finally, we have to determine the constant of proportionality for the creation and annihilation operators. We have already required that $\hat{a}_j \hat{a}_k^\dagger |\emptyset\rangle = \delta_{jk} |\emptyset\rangle$. To determine the rest, we consider a new observable that gives us the total number of particles in the system. We denote this observable by \hat{n} , and we see that it must be additive over all particle numbers for the different eigenvalues of A :

$$\hat{n} = \sum_j \hat{n}_j, \quad (8.31)$$

where \hat{n}_j is the number of particles in the eigenstate $|a_j\rangle$. The total number of particles does not change if we consider a different observable (although the distribution typically will), so this relation is also true when we count the particles in the states $|b_j\rangle$. Pretty much the only way we can achieve this is to choose

$$\hat{n} = \sum_j \hat{n}_j = \sum_j \hat{a}_j^\dagger \hat{a}_j = \sum_j \hat{b}_j^\dagger \hat{b}_j. \quad (8.32)$$

For the case of n_j particles in state $|a_j\rangle$ this gives

$$\hat{a}_j^\dagger \hat{a}_j |n_j\rangle = n_j |n_j\rangle, \quad (8.33)$$

where we ignored the particles in other states $|a_k\rangle$ with $k \neq j$ for brevity. For the Bose-Einstein case this leads to the relations

$$\hat{a}_j |n_j\rangle = \sqrt{n_j} |n_j - 1\rangle \quad \text{and} \quad \hat{a}_j^\dagger |n_j\rangle = \sqrt{n_j + 1} |n_j + 1\rangle. \quad (8.34)$$

For Fermi-Dirac statistics, the action of the creation and annihilation operators on number states becomes

$$\begin{aligned} \hat{a}_j |0\rangle_j &= 0 & \text{and} & & \hat{a}_j^\dagger |0\rangle_j &= e^{-i\alpha} |1\rangle_j, \\ \hat{a}_j |1\rangle_j &= e^{i\alpha} |0\rangle_j & \text{and} & & \hat{a}_j^\dagger |1\rangle_j &= 0. \end{aligned} \quad (8.35)$$

The phase factor $e^{i\alpha}$ can be chosen ± 1 .

8.3 Observables based on creation and annihilation operators

So far, we have considered only the basis states of many particles for a single observable A . What about other observables, in particular those that do not commute with A ? We can make a similar construction. Suppose an observable B has eigenvalues b_j . We can construct creation and annihilation operators \hat{b}_j^\dagger and \hat{b}_j that act according to

$$\begin{aligned}\hat{b}_j^\dagger |m_1, m_2, \dots, m_j, \dots\rangle &= \sqrt{m_j + 1} |m_1, m_2, \dots, m_j + 1, \dots\rangle, \\ \hat{b}_j |m_1, m_2, \dots, m_j, \dots\rangle &= \sqrt{m_j} |m_1, m_2, \dots, m_j - 1, \dots\rangle.\end{aligned}\quad (8.36)$$

where m_j is the number of particles with value b_j . Typically, the basis states of two observables are related via a single unitary transformation $|b_j\rangle = U|a_j\rangle$ for all j . How does this relate the creation and annihilation operators?

To answer this, let's look at the single particle states. We can write the single-particle eigenstates $|a_j\rangle$ and $|b_j\rangle$ as

$$|a_j\rangle = \hat{a}_j^\dagger |\emptyset\rangle \quad \text{and} \quad |b_j\rangle = \hat{b}_j^\dagger |\emptyset\rangle. \quad (8.37)$$

We assume that U does not change the vacuum⁶, so $U|\emptyset\rangle = |\emptyset\rangle$. This means that we can relate the two eigenstates via

$$|b_j\rangle = U|a_j\rangle = U\hat{a}_j^\dagger |\emptyset\rangle = U\hat{a}_j^\dagger (U^\dagger U) |\emptyset\rangle = U\hat{a}_j^\dagger U^\dagger |\emptyset\rangle = \hat{b}_j^\dagger |\emptyset\rangle, \quad (8.38)$$

where we have strategically inserted the identity $\mathbb{I} = U^\dagger U$. This leads to the operator transformation

$$\hat{b}_j^\dagger = U\hat{a}_j^\dagger U^\dagger. \quad (8.39)$$

The Hermitian adjoint is easily calculated as $\hat{b}_j = U\hat{a}_j U^\dagger$. It is left as an exercise for you to prove that

$$\hat{b}_j^\dagger = \sum_k u_{jk} \hat{a}_k^\dagger \quad \text{and} \quad \hat{b}_j = \sum_k u_{kj}^* \hat{a}_k, \quad (8.40)$$

where $u_{jk} = \langle a_k | b_j \rangle$.

How do we construct operators using the creation and annihilation operators? Suppose that a one-particle observable H has eigenvalues E_j and eigenstates $|j\rangle$. This can be, for example the Hamiltonian of the system, which ensures that the physical values of the particles (the eigenvalues) are additive. The operator for many identical particles then becomes

$$H = \sum_j E_j \hat{n}_j = \sum_j E_j \hat{a}_j^\dagger \hat{a}_j, \quad (8.41)$$

which transforms according to Eq. (8.40). More generally, the operator may not be written in the eigenbasis $|n_1, n_2, \dots\rangle$, in which case it has the form

$$H = \sum_{ij} H_{ij} \hat{b}_i^\dagger \hat{b}_j, \quad (8.42)$$

where H_{ij} are matrix elements. The creation and annihilation operators \hat{a}_j^\dagger and \hat{a}_j diagonalise H , and are sometimes called *normal modes*. The reason for this is that the creation and annihilation operators for bosons obey the same mathematical rules as the raising and lowering operators for the harmonic oscillator. The index j then denotes different oscillators. A system of coupled oscillators can be decomposed into normal modes, which are themselves isolated harmonic oscillators.

⁶This is a natural assumption when we are confined to the single particle Hilbert space, but there are general unitary transformations for which this does not hold, such as the transformation to an accelerated frame.

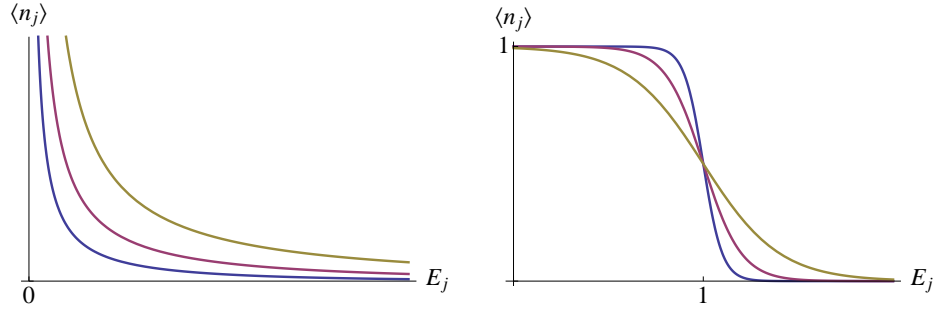


Figure 4: Left: Bose-Einstein distribution for different temperatures ($\mu = 0$). The lower the temperature, the more particles occupy the low energy states. Right: Fermi-Dirac distribution for different temperatures and $\mu = 1$. The fermions will not occupy energy states with numbers higher than 1, and therefore higher energies are necessarily populated. The energy values E_j form a continuum on the horizontal axis.

8.4 Bose-Einstein and Fermi-Dirac statistics

Finally, in this section we will derive the Bose-Einstein and Fermi-Dirac statistics. In particular, we are interested in the thermal equilibrium for a large number of (non-interacting) identical particles with some energy spectrum E_j , which may be continuous.

Since the number of particles is not fixed, we are dealing with the Grand Canonical Ensemble. Its partition function Ξ is given by

$$\Xi = \text{Tr} \left[e^{\mu\beta\hat{n} - \beta H} \right], \quad (8.43)$$

where H is the many-body Hamiltonian, $\beta = 1/k_B T$ and μ is the chemical potential. The average number of particles with single particle energy E_j is then given by

$$\langle n_j \rangle = -\frac{1}{\beta} \frac{\partial \ln \Xi}{\partial E_j}. \quad (8.44)$$

For the simple case where $H = \sum_j E_j \hat{n}_j$ and the creation and annihilation operators obey the commutator algebra, the exponent can be written as

$$\exp \left[\beta \sum_j (\mu - E_j) \hat{a}_j^\dagger \hat{a}_j \right] = \bigotimes_j \sum_{n_j=0}^{\infty} e^{\beta(\mu - E_j)n_j} |n_j\rangle \langle n_j|, \quad (8.45)$$

and the trace becomes

$$\Xi = \prod_j \frac{1}{1 - e^{\beta(\mu - E_j)}}. \quad (8.46)$$

The average photon number for energy E_j is

$$\langle n_j \rangle = -\frac{1}{\beta} \frac{\partial \ln \Xi}{\partial E_j} = -\frac{1}{\beta \Xi} \frac{\partial \Xi}{\partial E_j} = \frac{1}{e^{-\beta(\mu - E_j)} - 1}. \quad (8.47)$$

This is the Bose-Einstein distribution for particles with energy E_j . It is shown for increasing E_j in Fig. 4 on the left.

Alternatively, if the creation and annihilation operators obey the anti-commutation relations, the sum over n_j in Eq. (8.45) runs not from 0 to ∞ , but over 0 and 1. The partition function of the grand canonical ensemble then becomes

$$\Xi = \prod_j \left[1 + e^{\beta(\mu - E_j)} \right], \quad (8.48)$$

and the average number of particles with energy E_j becomes

$$\langle n_j \rangle = -\frac{1}{\beta \Xi} \frac{\partial \Xi}{\partial \hbar \omega_j} = \frac{1}{e^{-\beta(\mu - E_j)} + 1}. \quad (8.49)$$

This is the Fermi-Dirac statistics for these particles, and it is shown in Fig. 4 on the right. The chemical potential is the highest occupied energy at zero temperature, and in solid state physics this is called the Fermi level. Note the sign difference in the denominator with respect to the Bose-Einstein statistics.

Exercises

1. Calculate the Slater determinant for three electrons and show that no two electrons can be in the same state.
2. Particle statistics.
 - (a) What is the probability of finding n bosons with energy E_j in a thermal state?
 - (b) What is the probability of finding n fermions with energy E_j in a thermal state?
3. Consider a system of (non-interacting) identical bosons with a discrete energy spectrum and a ground state energy E_0 . Furthermore, the chemical potential starts out lower than the ground state energy $\mu < E_0$.
 - (a) Calculate $\langle n_0 \rangle$ and increase the chemical potential to $\mu \rightarrow E_0$ (e.g., by lowering the temperature). What happens when μ passes E_0 ?
 - (b) What is the behaviour of $\langle n_{\text{thermal}} \rangle \equiv \sum_{j=1}^{\infty} \langle n_j \rangle$ as $\mu \rightarrow E_0$? Sketch both $\langle n_0 \rangle$ and $\langle n_{\text{thermal}} \rangle$ as a function of μ . What is the fraction of particles in the ground state at $\mu = E_0$?
 - (c) What physical process does this describe?
4. The process $U = \exp(r \hat{a}_1^\dagger \hat{a}_2^\dagger - r^* \hat{a}_1 \hat{a}_2)$ with $r \in \mathbb{C}$ creates particles in two systems, 1 and 2, when applied to the vacuum state $|\Psi\rangle = U|\emptyset\rangle$.

- (a) Show that the bosonic operators $\hat{a}_1^\dagger \hat{a}_2^\dagger$ and $\hat{a}_1 \hat{a}_2$ obey the algebra

$$[K_-, K_+] = 2K_0 \quad \text{and} \quad [K_0, K_\pm] = \pm K_\pm,$$

with $K_+ = K_-^\dagger$.

- (b) For operators obeying the algebra in (a) we can write

$$e^{rK_+ - r^*K_-} = \exp\left[\frac{r}{|r|} \tanh|r| K_+\right] \exp\left[-2\ln(\cosh|r|) K_0\right] \times \exp\left[-\frac{r^*}{|r|} \tanh|r| K_-\right]. \quad (8.50)$$

Calculate the state $|\Psi\rangle$ of the two systems.

- (c) The amount of entanglement between two systems can be measured by the entropy $S(r)$ of the reduced density matrix $\rho_1 = \text{Tr}_2[\rho]$ for one of the systems. Calculate $S(r) = -\text{Tr}[\rho_1 \ln \rho_1]$.
- (d) What is the probability of finding n particles in system 1?

9 Many-Body Problems in Quantum Mechanics

In this final section we study a variety of problems in many-body quantum mechanics. First, we introduce the Hartree-Fock method for taking into account the effect of electron-electron interactions in atoms. Next, we describe spin waves in magnetic materials using the Heisenberg model. Third, we describe the behaviour of an atom interacting with photons in a cavity, and introduce the Jaynes-Cummings Hamiltonian. And finally, we take a brief look at the basic ideas behind quantum field theory.

9.1 Interacting electrons in atomic shells

Previously, you have encountered the Schrödinger equation for a particle in a central potential, which can be interpreted as an electron bound to a proton in a hydrogen atom. We found that the energy levels are quantised with quantum number n . In addition, the spin and orbital angular momentum is quantised with quantum numbers m_s , l , and m , respectively (we assume that $s = \frac{1}{2}$ since we are considering electrons). We can denote the set of quantum numbers n , m_s , l , and m by greek indices α , β , \dots , and the wave-functions $u_\alpha(\mathbf{r})$ then form a complete orthonormal basis for the bound electron.

It is tempting to keep this complete orthonormal basis for other atoms as well, and assume that the ground state of an N -electron atom is the tensor product of the N lowest energy eigenstates, appropriately anti-symmetrized via the Slater determinant. Indeed, the periodic table is based on this assumption. However, this ignores the fact that the electrons interact with each other, and the ground state of a many-electron atom is different. The Hartree-Fock method is designed to take this into account. It is a constrained variational approach, in which the trial state that is optimised over is forced to be a Slater determinant in order to keep the correct particle statistics. In this section we present the Hartree-Fock method, and arrive at the Hartree-Fock equations, which can be solved iteratively. We follow the derivation given by Bransden and Joachain (*Physics of Atoms and Molecules*, 1983 pp. 320-339).

First, we specify the Hamiltonian. Using the notation $r_i = |\mathbf{r}_i|$ for the distance of the i^{th} electron from the nucleus, and $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$ for the distance between electrons i and j , we find that

$$H = H_1 + H_2 = - \sum_{i=1}^N \left(\frac{\hbar^2}{2m} \nabla_i^2 + \frac{Z}{r_i} \right) + \sum_{i < j=1}^N \frac{1}{r_{ij}}, \quad (9.1)$$

where we used units in which $e/4\pi\epsilon_0 = 1$ and the Hamiltonian is divided into the single electron Hamiltonian (H_1) and the inter-electron Hamiltonian (H_2). We choose as a normalised trial quantum state $\Psi(\mathbf{r}_1, \dots, \mathbf{r}_N)$ the Slater determinant

$$\Psi(\mathbf{r}_1, \dots, \mathbf{r}_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} u_\alpha(\mathbf{r}_1) & u_\beta(\mathbf{r}_1) & \dots & u_\nu(\mathbf{r}_1) \\ u_\alpha(\mathbf{r}_2) & u_\beta(\mathbf{r}_2) & \dots & u_\nu(\mathbf{r}_2) \\ \vdots & \vdots & \ddots & \vdots \\ u_\alpha(\mathbf{r}_N) & u_\beta(\mathbf{r}_N) & \dots & u_\nu(\mathbf{r}_N) \end{vmatrix}, \quad (9.2)$$

and calculate the expectation value of the Hamiltonian H . This must be larger or equal to the ground state E_0

$$\langle \Psi | H | \Psi \rangle \geq E_0, \quad (9.3)$$

and varying the trial state then allows us to minimise the expectation value. This will get us close to the ground state energy.

Since the Slater determinant is a rather large expression, it will save us quite a bit of writing if we introduce the anti-symmetrisation operator \mathcal{A} , such that

$$\Psi(\mathbf{r}_1, \dots, \mathbf{r}_N) = \sqrt{N!} \mathcal{A} u_\alpha(\mathbf{r}_1) u_\beta(\mathbf{r}_2) \dots u_\nu(\mathbf{r}_N) \equiv \sqrt{N!} \mathcal{A} \Phi_H, \quad (9.4)$$

where $\Phi_H(\mathbf{r}_1, \dots, \mathbf{r}_N)$ is the Hartree wave function. The operator \mathcal{A} can then be written as a sum over permutations P of the labels $\alpha, \beta, \dots, \nu$:

$$\mathcal{A} = \frac{1}{N!} \sum_P (-1)^P P, \quad (9.5)$$

and \mathcal{A} is a projection operator: $\mathcal{A}^2 = \mathcal{A} = \mathcal{A}^\dagger$. Both H_1 and H_2 commute with \mathcal{A} .

Next, we calculate the expectation values $\langle \Psi | H_1 | \Psi \rangle$ and $\langle \Psi | H_2 | \Psi \rangle$. Since $[H_1, \mathcal{A}] = 0$, we can write

$$\langle \Psi | H_1 | \Psi \rangle = N! \langle \Phi_H | \mathcal{A} H_1 \mathcal{A} | \Phi_H \rangle = N! \langle \Phi_H | H_1 \mathcal{A}^2 | \Phi_H \rangle = N! \langle \Phi_H | H_1 \mathcal{A} | \Phi_H \rangle. \quad (9.6)$$

A permutation of the labels $\alpha, \beta, \dots, \nu$ leads to an orthonormal state and H_1 is a sum over one-electron Hamiltonians. We can therefore write this as

$$\begin{aligned} \langle \Psi | H_1 | \Psi \rangle &= N! \langle \Phi_H | H_1 \mathcal{A} | \Phi_H \rangle = \sum_P (-1)^P \langle \Phi_H | H_1 P | \Phi_H \rangle = \sum_{i=1}^N \langle \Phi_H | \hat{h}_i | \Phi_H \rangle \\ &= \sum_\alpha \langle u_\alpha(\mathbf{r}_i) | \hat{h}_i | u_\alpha(\mathbf{r}_i) \rangle \equiv \sum_\alpha I_\alpha \end{aligned} \quad (9.7)$$

where \hat{h}_i is the single electron Hamiltonian

$$\hat{h}_i = -\frac{\hbar^2}{2m} \nabla_i^2 - \frac{Z}{r_i}. \quad (9.8)$$

Next, we calculate the expectation value $\langle \Psi | H_2 | \Psi \rangle$ of the two-electron interaction Hamiltonians. Using the same reasoning as in Eq. (9.6), we find that

$$\langle \Psi | H_2 | \Psi \rangle = N! \langle \Phi_H | H_2 \mathcal{A} | \Phi_H \rangle. \quad (9.9)$$

Substituting the explicit form of \mathcal{A} , we find

$$\langle \Psi | H_2 | \Psi \rangle = \sum_{i < j} \sum_P (-1)^P \left\langle \Phi_H \left| \frac{P}{r_{ij}} \right| \Phi_H \right\rangle = \sum_{i < j} \left\langle \Phi_H \left| \frac{1 - P_{ij}}{r_{ij}} \right| \Phi_H \right\rangle, \quad (9.10)$$

where P_{ij} is the exchange operator of electrons i and j . This expression allows us to write

$$\begin{aligned} \langle \Psi | H_2 | \Psi \rangle &= \frac{1}{2} \sum_{\alpha, \beta} \left[\left\langle u_\alpha(\mathbf{r}_i) u_\beta(\mathbf{r}_j) \left| \frac{1}{r_{ij}} \right| u_\alpha(\mathbf{r}_i) u_\beta(\mathbf{r}_j) \right\rangle \right. \\ &\quad \left. - \left\langle u_\alpha(\mathbf{r}_i) u_\beta(\mathbf{r}_j) \left| \frac{1}{r_{ij}} \right| u_\beta(\mathbf{r}_i) u_\alpha(\mathbf{r}_j) \right\rangle \right]. \end{aligned} \quad (9.11)$$

Note the swap of α and β in the last ket. This expectation value consists of two terms, namely the direct term

$$J_{\alpha\beta} \equiv \left\langle u_\alpha(\mathbf{r}_i) u_\beta(\mathbf{r}_j) \left| \frac{1}{r_{ij}} \right| u_\alpha(\mathbf{r}_i) u_\beta(\mathbf{r}_j) \right\rangle, \quad (9.12)$$

and the exchange term

$$K_{\alpha\beta} \equiv \left\langle u_{\alpha}(\mathbf{r}_i) u_{\beta}(\mathbf{r}_j) \left| \frac{1}{r_{ij}} \right| u_{\beta}(\mathbf{r}_i) u_{\alpha}(\mathbf{r}_j) \right\rangle. \quad (9.13)$$

The total expectation value therefore becomes

$$\langle \Psi | H | \Psi \rangle = \sum_{\alpha} I_{\alpha} + \frac{1}{2} \sum_{\alpha, \beta} (J_{\alpha\beta} - K_{\alpha\beta}). \quad (9.14)$$

The matrix elements $J_{\alpha\beta}$ and $K_{\alpha\beta}$ are real and symmetric in α and β .

The second step towards the Hartree-Fock method is to find the minimum of $E \equiv \langle \Psi | H | \Psi \rangle$ by varying the $u_{\alpha}(\mathbf{r}_i)$. This means finding $\delta E = 0$. However, we must keep the functions $u_{\alpha}(\mathbf{r}_i)$ orthonormal to each other, and this imposes N^2 constraints. We can incorporate these constraints in the variational procedure by including Lagrange multipliers $\epsilon_{\alpha\beta}$, and the variational equation becomes

$$\delta E - \sum_{\alpha, \beta} \epsilon_{\alpha\beta} \delta \langle u_{\alpha}(\mathbf{r}) | u_{\beta}(\mathbf{r}) \rangle = 0. \quad (9.15)$$

There is no explicit reference to electron positions \mathbf{r}_i in $\langle u_{\alpha}(\mathbf{r}) | u_{\beta}(\mathbf{r}) \rangle$ since we are only interested in its orthonormality properties. The Lagrange multipliers $\epsilon_{\alpha\beta}$ form the elements of a Hermitian matrix.

The variational approach ultimately leads to a set of N coupled equations:

$$\begin{aligned} E_{\alpha} u_{\alpha}(\mathbf{r}_i) = & \hat{h}_i u_{\alpha}(\mathbf{r}_i) + \sum_{\beta} \langle u_{\beta}(\mathbf{r}_j) | r_{ij}^{-1} | u_{\beta}(\mathbf{r}_j) \rangle u_{\alpha}(\mathbf{r}_i) \\ & - \sum_{\beta} \langle u_{\beta}(\mathbf{r}_j) | r_{ij}^{-1} | u_{\alpha}(\mathbf{r}_j) \rangle u_{\beta}(\mathbf{r}_i), \end{aligned} \quad (9.16)$$

which are known as the *Hartree-Fock equations*. These can be solved by iteration up to any desired precision.

9.2 Spin waves in solids

Consider a system of spins with a nearest-neighbour interaction. For a uniform interaction in all directions, this is described by the Hamiltonian

$$H = \pm J \sum_{(i,j)} \mathbf{S}_i \cdot \mathbf{S}_j = \pm J \sum_{(i,j)} S_{z,i} S_{z,j} + \frac{1}{2} (S_{+,i} S_{-,j} + S_{-,i} S_{+,j}). \quad (9.17)$$

where $J > 0$ is the coupling strength between the spins, $\sum_{(i,j)}$ is the sum over all neighbouring pairs, and $S_{\pm} = S_x \pm iS_y$. The physics described by this Hamiltonian is known as the *Heisenberg model*. The sign of the coupling (here made explicit) determines whether the spins want to align up in parallel ($-J$) or antiparallel ($+J$). The former situation describes ferromagnets, while the latter describes anti-ferromagnets. The spin operators for different sites ($i \neq j$) commute with each other, while the spin operators at the same site ($i = j$) obey the spin algebra of Eq. (7.24).

Both systems have a well-defined ground state. For the ferromagnet this is the tensor product of the ground state of each individual spin. We are interested in the behaviour of the excitations with respect to this ground state. Due to the large degeneracy in the system (all the spins are of the same species with the same coupling J) the excitations act as identical quasi-particles. Consequently we can describe them using creation and annihilation operators. It turns out that

they behave like bosons. You can think of an excitation as a higher spin value at some site that propagates to its neighbours due to the interaction. This is called a *spin wave*.

Suppose that the spins are aligned in the positive z -direction (so we consider $-J$), and S_z has the maximum eigenvalue s . When the spin is lowered by \hbar , this creates an excitation in the system, because the spin is no longer lined up. So the z -component of the spin at site j is given by

$$S_{z,j} = s - \hat{a}_j^\dagger \hat{a}_j, \quad (9.18)$$

where $\hat{a}_j^\dagger \hat{a}_j$ is the operator for the number of excitations at site j . Since S_\pm raise and lower the eigenvalue of S_z , we expect that $S_+ \propto \hat{a}$ and $S_- \propto \hat{a}^\dagger$. When we insist on the commutation relation $[S_+, S_-] = 2S_z$, they become

$$S_{+,j} = \left(2s - \hat{a}_j^\dagger \hat{a}_j\right)^{\frac{1}{2}} \hat{a}_j \quad \text{and} \quad S_{-,j} = \left(2s - \hat{a}_j^\dagger \hat{a}_j\right)^{\frac{1}{2}} \hat{a}_j^\dagger. \quad (9.19)$$

This is known as the *Holstein-Primakoff* transformation.

For small numbers, the operators S_\pm can be approximated as

$$S_{+,j} \simeq \sqrt{2s} \hat{a}_j \quad \text{and} \quad S_{-,j} \simeq \sqrt{2s} \hat{a}_j^\dagger. \quad (9.20)$$

This allows us to write the Heisenberg Hamiltonian of Eq. (9.17) with $-J$ to lowest order as

$$H = -J \sum_{(i,j)} \left[s^2 + s \left(\hat{a}_i^\dagger \hat{a}_j + \hat{a}_i \hat{a}_j^\dagger - \hat{a}_i^\dagger \hat{a}_i - \hat{a}_j^\dagger \hat{a}_j \right) \right]. \quad (9.21)$$

For a simple cubic lattice of side L , lattice constant a and total number of spins $N = (L/a)^3$ we expect the spin waves to have wave vectors

$$\mathbf{k} = \frac{2\pi}{L}(m, n, o) \quad \text{with} \quad m, n, o \in \mathbb{N}, \quad (9.22)$$

and $1 \leq m, n, o \leq L$. The spin sites must now be indicated by a vector \mathbf{r} instead of a single number j , and the Fourier transformation of $\hat{a}_\mathbf{r}^\dagger$ and $\hat{a}_\mathbf{r}$ is given by

$$\hat{a}_\mathbf{r} = \frac{1}{\sqrt{N}} \sum_{\mathbf{k}} e^{-i\mathbf{k} \cdot \mathbf{r}} \hat{a}_\mathbf{k} \quad \text{and} \quad \hat{a}_\mathbf{r}^\dagger = \frac{1}{\sqrt{N}} \sum_{\mathbf{k}} e^{i\mathbf{k} \cdot \mathbf{r}} \hat{a}_\mathbf{k}^\dagger, \quad (9.23)$$

which transforms the Heisenberg Hamiltonian to

$$\begin{aligned} H &= -3Js^2N - \frac{Js}{N} \sum_{\mathbf{r}, \mathbf{d}} \sum_{\mathbf{k}, \mathbf{k}'} e^{i\mathbf{r}(\mathbf{k}-\mathbf{k}')} \left(e^{i\mathbf{d} \cdot \mathbf{q}} - 1 \right) \hat{a}_\mathbf{k}^\dagger \hat{a}_{\mathbf{k}'} \\ &= -3Js^2N - \frac{Js}{N} \sum_{\mathbf{k}} \epsilon(\mathbf{k}) \hat{a}_\mathbf{k}^\dagger \hat{a}_\mathbf{k}, \end{aligned} \quad (9.24)$$

where \mathbf{r} is the position of a lattice site, \mathbf{d} is the vector from a site to its nearest neighbours, which takes care of the sum over nearest neighbours. This is a diagonal Hamiltonian with eigen-energies

$$\epsilon(\mathbf{k}) = 2Js(3 - \cos k_x a - \cos k_y a - \cos k_z a). \quad (9.25)$$

This is the dispersion relation for the spin waves, and to lowest order ($\cos x \simeq 1 - \frac{1}{2}x^2$) it is quadratic:

$$\epsilon(\mathbf{k}) = Js a^2 k^2. \quad (9.26)$$

Spin waves are important when we want to manipulate magnetic properties with high frequency, such as in microwave devices. They carry energy, and are therefore a mechanism for dissipation.

For the case of anti-ferromagnets ($+J$), the ground state is harder to find. Consider an anti-ferromagnet that is again a simple cubic lattice with alternating spin $\pm s$ and lattice constant a . We can think of this lattice as two sub-lattices with constant spin, and redefine the spins on the $-s$ sub-lattice according to

$$S_x \rightarrow -S_x, \quad S_y \rightarrow S_y, \quad \text{and} \quad S_z \rightarrow -S_z. \quad (9.27)$$

These operators still obey the commutation relations of spin (which $\mathbf{S} \rightarrow -\mathbf{S}$ would not), and the Heisenberg Hamiltonian becomes

$$H = -J \sum_{(i,j)} S_{z,i} S_{z,j} + \frac{1}{2} (S_{+,i} S_{+,j} + S_{-,i} S_{-,j}). \quad (9.28)$$

When we apply the Holstein-Primakoff transformation to this Hamiltonian, to first order we obtain

$$H = -J \sum_{(i,j)} \left[s^2 + s \left(\hat{a}_i^\dagger \hat{a}_i + \hat{b}_j^\dagger \hat{b}_j + \hat{a}_i \hat{b}_j + \hat{a}_i^\dagger \hat{b}_j^\dagger \right) \right], \quad (9.29)$$

where \hat{a}_i^\dagger and \hat{a}_i are the creation and annihilation operators for the spin $+s$ sub-lattice, and \hat{b}_j^\dagger and \hat{b}_j are the creation and annihilation operators for the original spin $-s$ sub-lattice. After the Fourier transform of the creation and annihilation operators we get

$$H = -3Js^2N + 3Js \sum_{\mathbf{k}} \left[\hat{a}_{\mathbf{k}}^\dagger \hat{a}_{\mathbf{k}} + \hat{b}_{-\mathbf{k}}^\dagger \hat{b}_{-\mathbf{k}} + f(\mathbf{k}) \left(\hat{a}_{\mathbf{k}} \hat{b}_{-\mathbf{k}} + \hat{a}_{\mathbf{k}}^\dagger \hat{b}_{-\mathbf{k}}^\dagger \right) \right], \quad (9.30)$$

where $f(\mathbf{k}) = \frac{1}{3}(\cos k_x a + \cos k_y a + \cos k_z a)$.

To find the ground state we must diagonalise H so that it is a sum over number operators. This will involve mixing creation and annihilation operators. This is a unitary transformation that can be written as

$$\hat{a}_{\mathbf{k}} = u_{\mathbf{k}} \hat{c}_{\mathbf{k}} - v_{\mathbf{k}} \hat{d}_{-\mathbf{k}}^\dagger \quad \text{and} \quad \hat{b}_{-\mathbf{k}} = u_{\mathbf{k}} \hat{d}_{-\mathbf{k}} - v_{\mathbf{k}} \hat{c}_{\mathbf{k}}^\dagger. \quad (9.31)$$

This leads to the Hamiltonian

$$H = -3Js(s+1)N + \sum_{\mathbf{k}} \epsilon(\mathbf{k}) \left(\hat{c}_{\mathbf{k}}^\dagger \hat{c}_{\mathbf{k}} + \hat{d}_{-\mathbf{k}}^\dagger \hat{d}_{-\mathbf{k}} + 1 \right), \quad (9.32)$$

with the spin wave energy

$$\epsilon(\mathbf{k}) = 3Js \left(1 - f(\mathbf{k})^2 \right)^{\frac{1}{2}}. \quad (9.33)$$

For small \mathbf{k} the dispersion relation of the spin wave is linear in the wave vector, $\epsilon(\mathbf{k}) \simeq \sqrt{3}Jsa k$, which means that the spin waves behave markedly different in ferromagnets and anti-ferromagnets.

9.3 An atom in a cavity

Consider a two-level atom with bare energy eigenstates $|g\rangle$ and $|e\rangle$ and energy splitting $\hbar\omega_0$. The free Hamiltonian H_0 of the atom is given by

$$H_0 = \frac{1}{2}\hbar\omega_0 \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (9.34)$$

The atom interacts with an electromagnetic wave, and the interaction is approximately the coupling between the dipole moment $\hat{\mathbf{d}}$ associated with the $|g\rangle \leftrightarrow |e\rangle$ transition and the electric field \mathbf{E} , leading to the interaction Hamiltonian

$$H_{\text{int}} = -\hat{\mathbf{d}} \cdot \mathbf{E}, \quad (9.35)$$

where $\hat{\mathbf{d}} = -e\hat{\mathbf{r}}$, and \mathbf{E} is the classical, complex-valued electric field at the position of the atom. For an atom at the position $\mathbf{r} = 0$, the electric field can be written as

$$\mathbf{E} = \mathcal{E}_0 \epsilon e^{i\omega t} + \mathcal{E}_0 \epsilon^* e^{-i\omega t}, \quad (9.36)$$

where \mathcal{E}_0 is the real amplitude of the electric field ϵ is the polarisation vector of the wave. The off-diagonal matrix elements of H_{int} are given by

$$\langle e|H_{\text{int}}|g\rangle = e\mathcal{E}_0 \langle e|\hat{\mathbf{r}}|g\rangle \cdot \epsilon e^{i\omega t} + \text{c.c.}, \quad (9.37)$$

and c.c. denotes the complex conjugate. Since $\hat{\mathbf{r}}$ has odd parity, the diagonal matrix elements $\langle g|H_{\text{int}}|g\rangle$ and $\langle e|H_{\text{int}}|e\rangle$ vanish. When we define $\mathbf{r}_{eg} \equiv \langle e|\hat{\mathbf{r}}|g\rangle$, the total Hamiltonian becomes

$$H = \begin{pmatrix} \frac{1}{2}\hbar\omega_0 & e\mathcal{E}_0 \mathbf{r}_{eg}^* \cdot (\epsilon e^{i\omega t} + \epsilon^* e^{-i\omega t}) \\ e\mathcal{E}_0 \mathbf{r}_{eg} \cdot (\epsilon e^{i\omega t} + \epsilon^* e^{-i\omega t}) & -\frac{1}{2}\hbar\omega_0 \end{pmatrix}. \quad (9.38)$$

Using the Rotating Wave Approximation (see exercise 9.1), this Hamiltonian can be written as

$$H = \frac{\hbar}{2} \begin{pmatrix} \nu & \Omega^* \\ \Omega & -\nu \end{pmatrix}, \quad (9.39)$$

where we made the substitution

$$\nu = \omega - \omega_0 \quad \text{and} \quad \Omega = \frac{2e\mathcal{E}_0}{\hbar} \mathbf{r}_{eg} \cdot \epsilon. \quad (9.40)$$

We can use the standard matrix techniques in quantum mechanics to solve for the eigenvalues, the eigenstates, and the time evolution of the atom.

Next, we consider the situation where atom is placed inside a cavity of volume V , and the electric field in the cavity has angular frequency ω with wave vector k propagating in the z -direction. Assume that the length of the cavity is a multiple of $\lambda/2$, such that ω is a resonant cavity mode. The field is very weak, so that the classical description of \mathbf{E} is no longer sufficient. In particular, the field is made of photons, i.e., identical bosons. Consequently, we need to express \mathbf{E} in terms of bosonic creation and annihilation operators \hat{a}^\dagger and \hat{a} , which create photons of frequency ω . Since the intensity of the field is proportional to both \mathcal{E}_0^2 and $\hat{a}^\dagger \hat{a}$, we expect the electric field to be proportional to the creation and annihilation operators themselves. Hermiticity requires that it is proportional to the sum of the creation and annihilation operators. Furthermore, the electric field is a transverse standing wave cavity mode and must vanish at the mirrors due to the boundary

conditions imposed by Maxwell's equations. The spatial amplitude variation therefore includes a factor $\sin kz$. For linear polarisation the operator form of \mathbf{E} then becomes

$$\hat{\mathbf{E}}(z, t) = \epsilon \sqrt{\frac{\hbar\omega}{\epsilon_0 V}} \left(\hat{a} e^{-i\omega t} + \hat{a}^\dagger e^{i\omega t} \right) \sin kz, \quad (9.41)$$

where we assumed that ϵ is real⁷. The creation and annihilation operators are thus the *amplitude* operators of the field. Note that by the analogy with the harmonic oscillator, the electric field operator acts as a position operator of a particle in a harmonic potential well characterised by ω .

We again consider the dipole approximation of the atom in the field, and the Hamiltonian is written as

$$H_{\text{int}} = -\hat{\mathbf{d}} \cdot \hat{\mathbf{E}} = e \hat{\mathbf{r}} \cdot \epsilon \sqrt{\frac{\hbar\omega}{\epsilon_0 V}} \sin kz \left(\hat{a} e^{-i\omega t} + \hat{a}^\dagger e^{i\omega t} \right). \quad (9.42)$$

The operator $\hat{\mathbf{r}}$ can be written as

$$\hat{\mathbf{r}} = \mathbf{r}_{eg} |e\rangle \langle g| + \mathbf{r}_{eg}^* |g\rangle \langle e|, \quad (9.43)$$

and for notational simplicity, we define the coupling constant g as

$$g = e \mathbf{r}_{eg} \cdot \epsilon \sqrt{\frac{\hbar\omega}{\epsilon_0 V}} \sin kz. \quad (9.44)$$

We again calculate the matrix elements of H_{int} as before, but this time we write the operator in terms of $|g\rangle \langle e|$ and $|e\rangle \langle g|$:

$$H_{\text{int}} = g |e\rangle \langle g| \left(\hat{a} e^{-i\omega t} + \hat{a}^\dagger e^{i\omega t} \right) + g^* |g\rangle \langle e| \left(\hat{a} e^{-i\omega t} + \hat{a}^\dagger e^{i\omega t} \right). \quad (9.45)$$

It is convenient to express $|g\rangle \langle e|$ and $|e\rangle \langle g|$ in terms of the two-level raising and lowering operators σ_+ and σ_- :

$$\sigma_+ = |e\rangle \langle g| \quad \text{and} \quad \sigma_- = |g\rangle \langle e|, \quad (9.46)$$

with the commutation relation

$$[\sigma_+, \sigma_-] = 2\sigma_3 \quad \text{with} \quad \sigma_3 = |g\rangle \langle g| - |e\rangle \langle e|. \quad (9.47)$$

The interaction Hamiltonian becomes

$$H_{\text{int}} = g \sigma_+ \left(\hat{a} e^{-i\omega t} + \hat{a}^\dagger e^{i\omega t} \right) + g^* \sigma_- \left(\hat{a} e^{-i\omega t} + \hat{a}^\dagger e^{i\omega t} \right). \quad (9.48)$$

Including the free Hamiltonian for the field and the two-level atom, this becomes in the Rotating Wave Approximation

$$H_{\text{JC}} = \frac{1}{2} \hbar \omega_0 \sigma_3 + \hbar \omega \hat{a}^\dagger \hat{a} + g \sigma_+ \hat{a} + g^* \sigma_- \hat{a}^\dagger, \quad (9.49)$$

This is the Jaynes-Cummings Hamiltonian for a two-level atom with energy splitting $\hbar\omega_0$ interacting with a cavity mode of frequency ω . To achieve the strongest coupling g , the volume of the cavity should be small, and the atom should sit at an anti-node of the field.

⁷This is true for linear polarisation. For elliptical polarisation ϵ will be complex. The subsequent derivation will be slightly modified (with more terms in H_{int}), but no extra technical or conceptual difficulties arise.

Quantities are conserved when they commute with the Hamiltonian. We can identify two observables that satisfy this requirement, namely the number of electrons

$$\hat{P}_e = |g\rangle\langle g| + |e\rangle\langle e| = \mathbb{I}, \quad (9.50)$$

and the total number of excitations

$$\hat{N}_e = \hat{a}^\dagger \hat{a} + |e\rangle\langle e|. \quad (9.51)$$

This means that the Hamiltonian will not couple states with different total excitations.

In a real system, the cavity will not be perfect, and the excited state of the atom will suffer from spontaneous emission into modes other than the cavity mode. This can be modelled by a Lindblad equation for the joint state ρ of the atom and the cavity mode. The Lindblad operator for a leaky cavity is proportional to the annihilation operator \hat{a} , with a constant of proportionality $\sqrt{\kappa}$ that denotes the leakage rate. The spontaneous emission of the atom is modelled by the Lindblad operator $\sqrt{\gamma}\sigma_-$. The Lindblad equation then becomes

$$\frac{d\rho}{dt} = \frac{1}{i\hbar} [H_{JC}, \rho] + \gamma \sigma_- \rho \sigma_+ - \frac{\gamma}{2} \{\sigma_+ \sigma_-, \rho\} + \kappa \hat{a} \rho \hat{a}^\dagger - \frac{\kappa}{2} \{\hat{a}^\dagger \hat{a}, \rho\}. \quad (9.52)$$

The research field of *cavity quantum electrodynamics* (or cavity QED) is devoted in a large part to solving this equation.

9.4 Outlook: quantum field theory

We have surreptitiously introduced the basic elements of non-relativistic quantum field theory. Consider again the Heisenberg model, where we described a lattice of spins with nearest-neighbour interactions. If we take the limit of the lattice constant $a \rightarrow 0$ we end up with a continuum of creation and annihilation operators for each point in space. This is a *field*.

Traditionally we construct a quantum field theory from harmonic oscillators at each point in space. To this end, we characterise a classical harmonic oscillator with mass M in terms of its displacement q and velocity \dot{q} . The equations of motion for the classical harmonic oscillator are the Euler-Lagrange equations

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}} - \frac{\partial L}{\partial q} = 0, \quad (9.53)$$

where L is the Lagrangian

$$L = \frac{1}{2} M \dot{q}^2 - \frac{1}{2} K q^2, \quad (9.54)$$

and K can be thought of as a spring constant. Substituting this L into Eq. (9.53) yields the familiar differential equations for the harmonic oscillator $\ddot{q} + \Omega^2 q = 0$, with $\Omega^2 = K/M$.

Next, we arrange N particles in a one-dimensional lattice of length L and lattice constant a , where $L = Na$. Each particle's displacement is coupled to the displacement of its nearest neighbours by a spring with constant K . The equations of motion of this set of coupled particles is given by

$$\ddot{q}_n = \Omega^2 [(q_{n+1} - q_n) + (q_{n-1} - q_n)]. \quad (9.55)$$

We take the limit of $a \rightarrow 0$ and $N \rightarrow \infty$ while keeping $L = Na$ fixed. Our variable $q_n(t)$ then becomes a *field* $u(x, t)$, and it takes only a few lines of algebra to show that

$$\ddot{u}(x, t) = a^2 \Omega^2 u''(x, t) = v^2 u''(x, t) \quad \text{with} \quad \lim_{a \rightarrow 0} a \Omega = v. \quad (9.56)$$

This is a wave equation, and v is the velocity of the wave. We can generalise this immediately to three dimensions:

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} - \frac{1}{v^2} \frac{\partial^2 u}{\partial t^2} = 0. \quad (9.57)$$

In quantum field theory, we consider only these waves as the field excitations, and ignore the underlying particle structure we used to arrive at this result. We have already done something similar when we considered the spin waves in the Heisenberg model. Note also that the finite speed of propagation of waves means that we can make the theory Lorentz invariant when $v = c$, the speed of light, and Eq. (9.57) becomes $\partial_\mu \partial^\mu u = 0$.

The wave equation is typically derived from a Lagrangian L , or in the case of a field theory, the Lagrangian *density* \mathcal{L} . A massless scalar field is described by the Lorentz-invariant Lagrangian density

$$\mathcal{L} = \frac{1}{2} \left[\left(\frac{\partial \phi}{\partial x} \right)^2 + \left(\frac{\partial \phi}{\partial y} \right)^2 + \left(\frac{\partial \phi}{\partial z} \right)^2 - \frac{1}{c^2} \left(\frac{\partial \phi}{\partial t} \right)^2 \right] = \frac{1}{2} (\partial_\mu \phi)(\partial^\mu \phi), \quad (9.58)$$

where for technical reasons we redefined $\phi = u/\sqrt{a}$. The dispersion relation for such a field is $c^2 k^2 = \omega^2$, with k the wave number and ω the frequency of the wave. Similarly, a massive field is described by

$$\mathcal{L} = \frac{1}{2} (\partial_\mu \phi)(\partial^\mu \phi) - \frac{1}{2} m^2 \phi^2. \quad (9.59)$$

The Euler-Lagrange equation for this Lagrangian density is the so-called *Klein-Gordon* equation

$$\left(\partial_\mu \partial^\mu + \frac{m^2 c^2}{\hbar^2} \right) \phi = 0. \quad (9.60)$$

The mass term leads to a new dispersion relation

$$c^2 \hbar^2 k^2 - \hbar^2 \omega^2 + m^2 c^4 = 0, \quad (9.61)$$

and the group velocity for wave packets is

$$v_g = \frac{d\omega}{dk} = \frac{c}{\sqrt{1 + \mu^2}} \quad \text{with} \quad \mu = \frac{mc}{\hbar k}. \quad (9.62)$$

For relativistic particles the momentum $\hbar k$ is much larger than the rest mass mc , and therefore v_g approaches c .

We can solve the Klein-Gordon equation formally by writing

$$\phi(\mathbf{r}, t) = \int \frac{d\mathbf{k}}{\sqrt{2(2\pi)^3 \omega_{\mathbf{k}}}} \left(a(\mathbf{k}) e^{i\mathbf{k} \cdot \mathbf{r} - i\omega t} + a^*(\mathbf{k}) e^{-i\mathbf{k} \cdot \mathbf{r} + i\omega t} \right), \quad (9.63)$$

where $a_{\mathbf{k}}$ is the complex amplitude of a wave with wave vector \mathbf{k} . The field is essentially a superposition of (non-interacting) eigenmodes labelled by \mathbf{k} , and we call this a *free field*. We can

introduce interactions between the waves in the field by adding higher-order terms to \mathcal{L} with coupling constants ν, λ, \dots

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \phi)(\partial^\mu \phi) - \frac{1}{2}m^2 \phi^2 - \frac{\nu}{3!}\phi^3 - \frac{\lambda}{4!}\phi^4 - \dots \quad (9.64)$$

The solution to the Klein-Gordon equation is now no longer the correct solution to the new equations of motion, but when the interaction is reasonably weak we can use the solutions $\phi(\mathbf{r}, t)$ of the free field as a starting point in a perturbation expansion.

So far, everything in this section has been a classical treatment. In order to extend the theory to quantum mechanics we have to *quantise* the field. We achieve this by promoting the amplitudes in ϕ (and therefore ϕ itself) to operators that obey commutation of anti-commutation relations

$$\hat{\phi}(\mathbf{r}, t) = \int \frac{d\mathbf{k}}{\sqrt{2(2\pi)^3 \omega_{\mathbf{k}}}} \left(\hat{a}(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{r} - i\omega_{\mathbf{k}}t} + \hat{a}^\dagger(\mathbf{k}) e^{-i\mathbf{k}\cdot\mathbf{r} + i\omega_{\mathbf{k}}t} \right). \quad (9.65)$$

These are the creation and annihilation operators for excitations of the field. For the Klein-Gordon equation they obey the bosonic commutation relations

$$[\hat{a}(\mathbf{k}), \hat{a}(\mathbf{k}')] = [\hat{a}^\dagger(\mathbf{k}), \hat{a}^\dagger(\mathbf{k}')] = 0 \quad \text{and} \quad [\hat{a}(\mathbf{k}), \hat{a}^\dagger(\mathbf{k}')] = \delta^3(\mathbf{k} - \mathbf{k}'). \quad (9.66)$$

The state of the field can then be written as a superposition of Fock states. The field $\hat{\phi}$ has now become an observable.

In quantum field theory, the excitations of the field are interpreted as particles. All fundamental particles like quarks, electrons, photons, and the Higgs boson are excitations of a corresponding field. So the excitations of the Higgs field are Higgs bosons, and the excitations of the electromagnetic field are photons. Spin- $\frac{1}{2}$ particles obey the Dirac equation

$$(i\hbar\gamma^\mu\partial_\mu - mc)\hat{\psi} = 0, \quad (9.67)$$

where the γ^μ are 4×4 matrices

$$\gamma^0 = \begin{pmatrix} \mathbb{I} & 0 \\ 0 & -\mathbb{I} \end{pmatrix}, \quad \gamma^1 = \begin{pmatrix} 0 & \sigma_x \\ -\sigma_x & 0 \end{pmatrix}, \quad \gamma^2 = \begin{pmatrix} 0 & \sigma_y \\ -\sigma_y & 0 \end{pmatrix}, \quad \gamma^3 = \begin{pmatrix} 0 & \sigma_z \\ -\sigma_z & 0 \end{pmatrix}, \quad (9.68)$$

and ψ is a four-dimensional vector field called a *spinor* field. The solution to the free Dirac field can be written as

$$\hat{\psi}(\mathbf{r}, t) = \sum_s \int \frac{d\mathbf{k}}{\sqrt{2(2\pi)^3 \omega_{\mathbf{k}}}} \left[\hat{b}_s(\mathbf{k}) u_s(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{r} - i\omega_{\mathbf{k}}t} + \hat{d}_s^\dagger(\mathbf{k}) v_s(\mathbf{k}) e^{-i\mathbf{k}\cdot\mathbf{r} + i\omega_{\mathbf{k}}t} \right], \quad (9.69)$$

where $s = \pm \frac{1}{2}\hbar$ is the spin value, and $u_s(\mathbf{k})$ and $v_s(\mathbf{k})$ are two spinors carrying the spin component of the field

$$u_s(\mathbf{k}) = \mathcal{N} \begin{pmatrix} \chi_s \\ \frac{\hbar c \boldsymbol{\sigma} \cdot \mathbf{k}}{\hbar \omega + mc^2} \chi_s \end{pmatrix} \quad \text{and} \quad v_s(\mathbf{k}) = \mathcal{N} \begin{pmatrix} -\frac{\hbar c \boldsymbol{\sigma} \cdot \mathbf{k}}{\hbar \omega + mc^2} \chi_s \\ \chi_s \end{pmatrix}, \quad (9.70)$$

where \mathcal{N} is a normalisation constant, $\boldsymbol{\sigma}$ is a vector of Pauli matrices, and

$$\chi_+ = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad \chi_- = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \quad (9.71)$$

The creation and annihilation operators $\hat{b}_s(\mathbf{k})$ and $\hat{d}_s^\dagger(\mathbf{k})$ obey anti-commutation relations:

$$\begin{aligned} \{\hat{b}_s(\mathbf{k}), \hat{b}_r(\mathbf{k}')\} &= \{\hat{b}_s^\dagger(\mathbf{k}), \hat{b}_r^\dagger(\mathbf{k}')\} = 0 & \text{and} & & \{\hat{b}_s(\mathbf{k}), \hat{b}_r^\dagger(\mathbf{k}')\} &= \delta_{rs} \delta^3(\mathbf{k} - \mathbf{k}'), \\ \{\hat{d}_s(\mathbf{k}), \hat{d}_r(\mathbf{k}')\} &= \{\hat{d}_s^\dagger(\mathbf{k}), \hat{d}_r^\dagger(\mathbf{k}')\} = 0 & \text{and} & & \{\hat{d}_s(\mathbf{k}), \hat{d}_r^\dagger(\mathbf{k}')\} &= \delta_{rs} \delta^3(\mathbf{k} - \mathbf{k}'), \\ \{\hat{b}_s(\mathbf{k}), \hat{d}_r(\mathbf{k}')\} &= \{\hat{b}_s^\dagger(\mathbf{k}), \hat{d}_r^\dagger(\mathbf{k}')\} = 0 & \text{and} & & \{\hat{b}_s(\mathbf{k}), \hat{d}_r^\dagger(\mathbf{k}')\} &= \{\hat{d}_s(\mathbf{k}), \hat{b}_r^\dagger(\mathbf{k}')\} = 0. \end{aligned} \quad (9.72)$$

This means that the excitations of the Dirac field are fermions with spin $\frac{1}{2}$, such as the electron. You see immediately that $\hat{\psi}$ is not Hermitian due to the appearance of \hat{d}^\dagger . This means that $\hat{\psi}$ is not an observable and we cannot think of the Dirac field as a quantised version of a classical observable field. There is no classical analog to the Dirac field. This is a consequence of the fact that the anti-particle of the Dirac excitations are not the same as the particle itself. E.g., the positron is different from the electron. Anti-particles are a quintessentially quantum mechanical phenomenon.

There is of course a lot more to quantum field theory than this. For example, the techniques for doing the perturbation expansion of interacting fields leads to Feynman diagrams, and renormalisation theory must be employed to deal with the infinities that crop up in the perturbation theory. Furthermore, one has to choose the right Lagrangian density, and principles such as gauge invariance and CPT invariance are imposed to constrain the possible choices. This leads ultimately to the extraordinary successful Standard Model of particle physics. It is the most fundamental theory of Nature that we have, and it is tested to unprecedented accuracy.

Exercises

1. The Hamiltonian for a two-level atom in the presence of an electromagnetic wave, as given in Eq. (9.38) depends on the time t . This makes it difficult to solve the Schrödinger equation, so in this exercise we will get rid of the time dependence by applying the Rotating Wave Approximation.

- (a) If $|\psi'(t)\rangle = U(t)|\psi(t)\rangle$, find the Schrödinger equation for $|\psi'(t)\rangle$. What is the new Hamiltonian?
- (b) Choose $U(t)$ such that

$$U(t) = \begin{pmatrix} e^{i\omega t/2} & 0 \\ 0 & e^{-i\omega t/2} \end{pmatrix},$$

and H is given by Eq. (9.38). Show that the new Hamiltonian is given by Eq. (9.39).

2. A two-level atom is placed in a perfect cavity with an electromagnetic field of frequency ω .
 - (a) Show that the Jaynes-Cummings Hamiltonian can be written as a direct sum of 2×2 matrices H_n , and specify H_n .
 - (b) Diagonalize H_n to find the energy values of the system, and calculate the eigenstates.
 - (c) At $t = 0$, the system is in the state $|\psi(0)\rangle = |e, n\rangle$. Calculate the state $|\psi(t)\rangle$ when the light is on resonance with the atomic transition ($\omega = \omega_0$).
 - (d) Calculate the amount of entanglement between the atom and the cavity field. Use the relative entropy as the entanglement measure.
 - (e) How long does the atom need to reside in the cavity in order to achieve maximum entanglement?