Notes on Quantum Mechanics

Module 2 of Refresher course conducted by Indian Academies of Sciences at
Loyola College, Chennai, Tamil Nadu, May 11-23, 2015

Govind S. Krishnaswami, Chennai Mathematical Institute

- These are very brief and incomplete notes for lectures at the above Refresher Course. Please let me know (govind@cmi.ac.in) of any comments or corrections. 7 Apr, 2020.

Contents

1 References

- Here are some books on quantum mechanics.

1. E Wichmann, Quantum Physics, Berkeley Physics Course, Vol 4.
2. R P Feynman, R B Leighton and M Sands, Feynman Lectures on Physics, Vol 3.
3. H S Mani and G K Mehta, Introduction to Modern Physics.
4. B Dutta-Roy, Elements of Quantum Mechanics.
5. J J Sakurai and J J Napolitano, Modern Quantum Mechanics, 2nd Ed.
6. L I Schiff, Quantum Mechanics, 3rd Ed. McGraw-Hill (1968).
7. P J E Peebles, Quantum Mechanics, Princeton Univ Press (1992).
8. B H Bransden and C J Joachain, Quantum Mechanics, 2nd Ed., Pearson (2000).
9. D J Griffiths, Introduction to Quantum Mechancis.
10. R L Liboff, Introductory Quantum Mechancis.
11. R Shankar, Principles of Quantum Mechanics, 2nd Ed. Plenum Press (1994).
12. P M Mathews and K Venkatesan, A textbook of quantum mechanics, Tata McGraw-Hill (1977).
13. P A M Dirac, The principles of quantum mechanics, 4th Ed., Oxford (1958).
14. W Heitler, Elementary wave mechanics with applications to Quantum Chemistry.
15. L D Landau and E M Lifshitz, Quantum mechanics: Non-relativistic theory, 3rd Edition.

2 Formalisms of classical and quantum mechanics

2.1 Setup of classical mechanics

- The set of possible instantaneous locations of a classical particle is called its configuration space. This is usually three dimensional Euclidean space $\mathbb{R}^3$. The number of coordinates needed to specify the instantaneous configuration of a system is the number of degrees of freedom. A system consisting of a pair of particles has 6 degrees of freedom $x_1, y_1, z_1, x_2, y_2, z_2$, its configuration space is $\mathbb{R}^3 \times \mathbb{R}^3 = \mathbb{R}^6$. A particle attached to a fixed support by a rod of fixed length has two degrees of freedom, its configuration space is a sphere. The configuration space and number of degrees of freedom are kinematical notions. They do not depend on the nature of forces between the particles.

• If the forces acting on/between the particles are known, then we may determine the dynamical time evolution of the system by solving Newton's equations for the trajectories. For one particle, $m\ddot{\mathbf{r}} = \mathbf{F}$. Newton's equations are second order in time, they require two sets of initial conditions, the initial positions $\mathbf{r}(0)$ and initial velocities $\dot{\mathbf{r}}(0)$. In other words the initial coordinates $\mathbf{r}(0)$ and initial momenta $\mathbf{p}(0) = m\dot{\mathbf{r}}(0)$ determine the future trajectory. We say that the instantaneous **state** of the system is specified by giving the coordinates and momenta of all the particles. The set of possible instantaneous states of a system is its **phase space**. For a particle moving along a line, its phase space is the $x - p$ phase plane. Newton's equations may be formulated as Hamilton's equations for the time evolution of coordinates and momenta

$$\dot{x} = \frac{\partial H}{\partial p} \quad \text{and} \quad \dot{p} = -\frac{\partial H}{\partial x}. \tag{1}$$

For a particle in a potential $H(x,p) = \frac{p^2}{2m} + V(x)$ and Hamilton's equations are a pair of first order equations

$$\dot{x} = \frac{p}{m} \quad \text{and} \quad \dot{p} = -\frac{dV}{dx}, \tag{2}$$

which may be written as a single second order equation expressing Newton's second law $m\ddot{x} = -V'(x)$. The curve in phase space $(x(t), p(t))$ is called the phase trajectory. Draw the phase portrait for a free particle as well as for a simple harmonic oscillator, indicating the direction of trajectories. A dynamical variable that is constant along trajectories is called a constant of motion. Its value may differ from trajectory to trajectory. The hamiltonian $H = T + V$ is a conserved quantity for conservative systems (i.e. where the force is the negative gradient of a scalar potential).

• Dynamical variables like angular momentum and the hamiltonian are functions of the basic dynamical variables position and momentum. In general, any real function of position and momentum is called an observable. Observables are simply real-valued functions on phase space. They must be real since observables are physical quantities that may be measured.

2.2 Framework and postulates of quantum mechanics

2.2.1 Hilbert Space of states of a quantum system

• States of a quantum system are vectors in a linear space ("vector space") called a complex Hilbert space $\mathcal{H}$. For a particle moving on a line, its configuration space is $\mathbb{R}^1$, parametrized by one coordinate x. Its quantum state space $\mathcal{H} = L^2(\mathbb{R})$ is the space of square-integrable functions $\psi(x)$ on the classical configuration space. ψ is called the state function or state vector or wave function of the particle.

• By Born's probability postulate, $|\psi(x)|^2 dx$ is interpreted as the probability of finding the particle between x and $x + dx$. Since the total probability of the particle being somewhere should be one, we normalize the wave function $\int_0^\infty |\psi(x)|^2 dx = 1$. This is why we restrict to square-integrable wave functions. $\psi(x)$ itself is called a **probability amplitude**, its square is a probability density.

• Unlike the classical space of states (phase space) which can be a non-linear manifold (e.g. if a particle is constrained to move on a circle), the quantum Hilbert space is always a linear space. The sum of two states $\psi + \phi$ is a possible state and so is a complex multiple $c\psi$ of any state.

This is the principle of linear superposition of states, used to explain the interference of matter waves in the double slit experiment.

- A complex Hilbert space $\mathcal{H}$ is a vector space over the complex numbers. It is a space of ket vectors $|\psi\rangle$ closed under linear superposition. If $|\psi\rangle$ and $|\chi\rangle$ are state vectors, then so is $\alpha|\psi\rangle + \beta|\chi\rangle$, for any $\alpha, \beta \in \mathbb{C}$. A simple example is the two dimensional complex vector space of spin states of a spin half particle which are usually denoted as column vectors $|\psi\rangle = \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}$. Notably, the space of states of a quantum system is a complex, rather than a real vector space.

- The quantum state space is equipped with an inner or dot product. The inner product of a pair of vectors ψ, χ is denoted $\langle\psi|\chi\rangle$. For the spin-$\frac{1}{2}$ Hilbert space, the inner product is

$$\langle\psi|\chi\rangle = \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}^\dagger \begin{pmatrix} \chi_1 \\ \chi_2 \end{pmatrix} = \begin{pmatrix} \psi_1^* & \psi_2^* \end{pmatrix} \begin{pmatrix} \chi_1 \\ \chi_2 \end{pmatrix} = \psi^\dagger \chi = \psi_1^* \chi_1 + \psi_2^* \chi_2. \quad (3)$$

$\psi^\dagger$ is called the hermitian adjoint, it is the complex conjugate transpose, it is a row vector. So associated with a vector space of colummn/ket vectors there is a 'dual' space of row/bra vectors, the adjoints of the kets $|\psi\rangle^\dagger = \langle\psi| = \begin{pmatrix} \psi_1^* & \psi_2^* \end{pmatrix}$. The inner product may also be regarded as producing a complex number from a bra and a ket vector $\langle\psi|\chi\rangle$. However, the inner product of a non-zero vector with itself is always a positive real number $\langle\psi|\psi\rangle > 0$, it is called the square of the length of the vector.

- Another example is n-dimensional complex vector space $\mathbb{C}^n$ with the inner product $\langle u|v\rangle = \sum_i u_i^* v_i$. The Hilbert space of a particle moving on a line is $L^2(\mathbb{R})$ with $\langle f|g\rangle = \int_{-\infty}^{\infty} f^*(x) g(x)\, dx$.

- From these examples (keep $\langle u|v\rangle = u_i^* v_i$ in mind) we abstract the basic properties of the inner product (these are its defining properties in an axiomatic approach)

$$\langle \alpha u|v\rangle = \alpha^* \langle u|v\rangle, \quad \langle u|\beta v\rangle = \beta\langle u|v\rangle, \quad \langle u+v|w\rangle = \langle u|w\rangle + \langle v|w\rangle, \quad \langle u|v\rangle^* = \langle v|u\rangle. \quad (4)$$

$\langle u|v\rangle$ is linear in the second vector v and anti-linear in the first vector u on account of complex conjugation of the components of the first vector.

- The norm/length of a vector is $||v|| = \sqrt{\langle v|v\rangle}$. The norm of a vector is unchanged upon multiplying by a phase $e^{i\alpha}$. If $\langle u|v\rangle = 0$ then the vectors are orthogonal.

- Two state vectors that differ by multiplication by a non-zero complex number $\psi_2(x) = c\psi_1(x)$ represent the **same physical state**. We often work with unit norm states.

- A basis for the Hilbert space is a set of vectors $|e_i\rangle$ such that any vector $|v\rangle$ may be expressed as a linear combination of $|e_i\rangle$ in a unique way. The number of basis vectors is the dimension of the vector space. The standard basis vectors of the two dimensional spin Hilbert space $\mathbb{C}^2$ are

$$|e_1\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |e_2\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \text{so} \quad \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = v_1|e_1\rangle + v_2|e_2\rangle. \quad (5)$$

The coefficients v_i in the expansion $|v\rangle = \sum_i v_i|e_i\rangle$ are called the components of $|v\rangle$. The components of the adjoint are the complex conjugates: $\langle v| = \sum_i \langle e_i| v_i^*$. [We will often drop the summation symbol and assume repeated indices are summed.] E.g. the adjoints of the basis vectors are row bra-vectors

$$\langle e_1| = e_1^\dagger = \begin{pmatrix} 1 & 0 \end{pmatrix}, \quad \langle e_2| = e_2^\dagger = \begin{pmatrix} 0 & 1 \end{pmatrix}. \quad (6)$$

$\mathbb{C}^n$ is an n-dimensional vector space. The state space of a particle moving on a line, $L^2(\mathbb{R})$ is infinite dimensional, it is called a function space. It is intuitively clear that this is an infinite dimensional space since the values of the function $\psi(x)$ at each $x \in \mathbb{R}$ can be freely specified (subject to normalizability). x here plays the role of the index $i = 1, 2$ in the two dimensional spin-half vector space $\mathbb{C}^2$. A possible basis for a function space is the set of monomials $\{1, x, x^2, x^3, x^4, \cdots\}$. Indeed, any function ψ that has a Taylor series around $x = 0$ admits a expression as a linear combination of these. The coefficients are the derivatives of ψ at $x = 0$:

$$\psi(x) = \psi(0) + \psi'(0)x + \frac{1}{2}\psi''(0)x^2 + \frac{1}{3!}\psi'''(0)x^3 + \cdots \tag{7}$$

However this basis of monomials is a bit inconvenient. In particular, the basis vectors are not orthogonal, in fact they are not even normalizable with respect to the above L^2 inner product. A more convenient basis for $L^2(\mathbb{R})$ consists of the energy eigenstates of the harmonic oscillator $|n\rangle$, which we will study in more detail later.

• It is often convenient to work with an orthonormal basis, i.e., a basis of vectors $|e_i\rangle$ which are pairwise orthogonal and each normalized to have unit norm, $\langle e_i|e_j\rangle = \delta_{ij}$. The standard basis $|e_i\rangle$ for $\mathbb{C}^n$ with components $|e_i\rangle_j = \delta_{ij}$ is orthonormal with respect to the usual inner product $\langle u|v\rangle = \sum_i u_i^* v_j$.

• A set of orthonormal vectors is said to be a complete orthonormal set if it forms a basis for the vector space, i.e., if we may write any vector as a linear combination.

2.2.2 Observables and hermitian linear operators on states

An observable A in quantum mechanics (e.g. hamiltonian, position, momentum, angular momentum, spin, magnetic moment) is a hermitian (self-adjoint) linear operator on the Hilbert space of states $\mathcal{H}$. Hermiticity is the quantum analogue of classical observables being real-valued functions. We will see that a hermitian operator has real eigenvalues, which are possible results when A is measured. To define a hermitian operator, we first note that a linear operator on a vector space takes vectors to vectors in a linear way: $A(a|\psi\rangle + b|\chi\rangle) = aA|\psi\rangle + bA|\chi\rangle$. When A acts on a vector $|v\rangle$ it produces a new ket vector $A|v\rangle$ which is also denoted $|Av\rangle$.

• A linear operator is an abstract concept, whose concrete realisation is a matrix. A linear operator on $\mathbb{C}^2$ is simply a 2×2 matrix, once we choose a basis to represent it. For example, the Pauli matrices $\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ are linear operators represented as matrices in the standard basis for $\mathbb{C}^2$.

• If $|e_j\rangle$ are a basis for $\mathcal{H}$, then a linear operator A is determined by how it acts on the basis vectors. Since A is linear $A|e_j\rangle$ must be a linear combination of the basis vectors themselves

$$A|e_j\rangle = \sum_k |e_k\rangle A_{kj}. \tag{8}$$

A_{kj} are the components of A in this basis, they may be written as entries in a matrix, with A_{kj} occupying the slot in the k^{th} row and j^{th} column. The vector that makes up the first column A_{k1} is the 'image' of e_1 (i.e. coefficients in the linear combination appearing in $A|e_1\rangle$), the second column A_{k2} is the image of e_2 and so on.

- If the basis e_i is orthonormal $\langle e_i|e_j\rangle = \delta_{ij}$, then we have

$$\langle e_i|A|e_j\rangle = \sum_k \langle e_i|e_k\rangle A_{kj} = \sum_k \delta_{ik} A_{kj} = A_{ij}. \tag{9}$$

We say that A_{ij} are the matrix elements of A between the o.n. basis states e_i and e_j.

- A matrix A is hermitian if it equals its own complex conjugate transpose. The latter is called its adjoint $A^\dagger = (A^*)^t$. So A is hermitian if $A = A^\dagger$, i.e., if it is self-adjoint. In terms of matrix entries, $A^*_{ij} = A_{ji}$. In particular, the diagonal entries of a hermitian matrix are real, while the off diagonal entries are complex conjugates of each other. The Pauli matrices are hermitian. Note that the adjoint of a product is the product of adjoints in the opposite order. $(AB)^\dagger = B^\dagger A^\dagger$ and that $(A|\psi\rangle)^\dagger = \langle\psi|A^\dagger$. We also denote $A|\psi\rangle = |A\psi\rangle$, so that $|A\psi\rangle^\dagger = \langle A\psi|$.

- The concept of hermiticity makes sense for a linear operator, even if we have not represented it explicitly as a matrix by choosing a basis. To explain the concept, we need the idea of matrix elements between states. If u, v are a pair of states, then $\langle u|A|v\rangle$ is called the **matrix element** of A between the states u and v. To know an operator is to know its matrix elements.

- The **adjoint** of A is the operator $A^\dagger$ defined via its matrix elements $\langle u|A^\dagger|v\rangle = \langle Au|v\rangle$. So if we know the matrix elements of A, then we may find the matrix elements of $A^\dagger$. A linear operator is **hermitian** if $\langle u|Av\rangle = \langle Au|v\rangle$ for all states $u, v \in \mathcal{H}$. A hermitian operator is also called symmetric by mathematicians since it does not matter whether A is written on the left or on the right.

- Now let us see how this abstract definition of hermiticity reduces to the formula $A_{ij} = A^*_{ji}$ for hermitian matrices. We must equate the matrix elements of A and those of $A^\dagger$. Let e_i be an orthonormal basis, then the matrix element of A between the states e_i and e_j is just A_{ij}, as is seen by taking the inner product of the above equation with e_i

$$\langle e_i|A|e_j\rangle = \sum_k \langle e_i|e_k\rangle A_{kj} = \sum_k \delta_{ik} A_{kj} = A_{ij}. \tag{10}$$

On the other hand, what are the matrix elements of $A^\dagger$? By the definition of the adjoint,

$$\langle e_i|A^\dagger e_j\rangle = \langle A e_i|e_j\rangle = \langle e_j|A e_i\rangle^* = (A_{ji})^* \tag{11}$$

So a linear operator is self-adjoint if its matrix elements in an orthonormal basis satisfy $A_{ij} = (A_{ji})^*$.

2.2.3 Outer products of vectors and completeness relation

- **Outer products of vectors:** Consider the vector space $\mathbb{C}^n$ with standard basis $|e_i\rangle$. Just as we may 'dot' row and column n-vectors to get a scalar inner product, we may also form their 'outer' product (column times a row), to get an $n \times n$ matrix. For $n = 2$ show that

$$|e_1\rangle\langle e_1| = e_1 e_1^\dagger = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad |e_2\rangle\langle e_2| = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \quad |e_1\rangle\langle e_2| = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad |e_2\rangle\langle e_1| = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}. \tag{12}$$

More generally, check that $|e_i\rangle\langle e_j|$ is a matrix with a 1 in the ij-entry and 0's elsewhere. From this we see that a matrix whose entries are A_{ij} in the i^{th} row and j^{th} column, can be expressed as

$$A = \sum_{ij} A_{ij} |e_i\rangle\langle e_j| \tag{13}$$

Now let us use this expression to find how a matrix acts on a vector $v = v_k|e_k\rangle$. We get

$$Av = \sum_{ij} A_{ij}\ |e_i\rangle\langle e_j|v_k|e_k\rangle = A_{ij}v_k|e_i\rangle\langle e_j|e_k\rangle = A_{ij}v_k\delta_{jk}|e_i\rangle = A_{ik}v_k|e_i\rangle. \tag{14}$$

So the i^{th} component of Av is $\sum_k A_{ik}v_k$.

• In particular, the identity operator I, may be expressed as

$$I = \sum_i |e_i\rangle\langle e_i| = \sum_{ij} \delta_{ij}|e_i\rangle\langle e_j| \tag{15}$$

The identity operator has the components δ_{ij} in any basis since it takes every vector to itself. This 'resolution' of the identity operator as a sum of outer products of a set of orthonormal basis vectors is called the **completeness relation**. It is quite useful in many physical problems and calculations. E.g. the energy eigenstates of the Harmonic oscillator form a complete orthonormal set and satisfy the above completeness relation. Coherent states for the harmonic oscillator also satisfy a completeness relation even though they are not orthogonal and are in fact an over-complete set.

• An anti-hermitian operator is one that satisfies $A^\dagger = -A$. A unitary operator is one whose inverse is its adjoint, $UU^\dagger = U^\dagger U = I$. It is clear that the identity I is hermitian as well as unitary. If A is anti-hermitian, then iA is hermitian since $(iA)^\dagger = A^\dagger i^\dagger = -A(-i) = A$.

2.2.4 Hermiticity of position and momentum operators

• Physically interesting examples of hermitian operators for a particle with one degree of freedom moving on a line include the position operator $\hat{x}\psi(x) = x\psi(x)$, and momentum operator $\hat{p}\psi(x) = -i\hbar\psi'(x)$. Check that $\hat{x}^\dagger = \hat{x}$ and $\hat{d} = \frac{\partial}{\partial x}$ is anti-hermitian. We must show $\langle f|\hat{x}g\rangle = \langle\hat{x}f|g\rangle$ for any two states f, g. This is seen as follows:

$$\langle f|\hat{x}g\rangle = \int f^*(x)xg(x)\,dx = \int (xf(x))^*g(x)\,dx = \langle\hat{x}f|g\rangle. \tag{16}$$

Showing hermiticity of $\hat{p} = -i\hbar\frac{\partial}{\partial x}$ requires integration by parts. Let us show that $\hat{d} = \frac{\partial}{\partial x}$ is anti-hermitian, from which it will follow that $\hat{p} = -i\hbar\hat{d}$ is hermitian. Let us denote complex conjugate of f by $\bar{f}$ here for convenience

$$\langle f|\hat{d}g\rangle = \int \bar{f}(x)g'(x)dx = -\int \bar{f}'(x)g(x)\,dx + \left[\bar{f}g\right]_{-\infty}^{\infty} = -\langle\hat{d}f|g\rangle. \tag{17}$$

Here we assumed f, g vanish at $\pm\infty$, which is the case for square-integrable functions. Boundary conditions play an important role in determining the hermiticity of momentum. If we have a particle moving on a finite interval $[a, b]$ (as in a square well), then

$$\langle f|\hat{d}g\rangle = \int_a^b \bar{f}(x)g'(x)dx = -\int_a^b \bar{f}'(x)g(x)\,dx + \left[\bar{f}g\right]_a^b = -\langle\hat{d}f|g\rangle + \left[\bar{f}g\right]_a^b. \tag{18}$$

For $\hat{d}$ to be anti-hermitian, the boundary term must vanish. This happens, for instance, if the functions vanish at the end points ($f(a) = f(b) = 0$, as in an infinite square well) or satisfy 'periodic boundary conditions' $f(a) = f(b)$.

• Of particular importance is the concept of **expectation value** of an observable A in a state ψ, which is defined as the normalzed diagonal matrix element of A in the state ψ

$$\langle A \rangle_\psi = \frac{\langle \psi | A \psi \rangle}{\langle \psi | \psi \rangle} \tag{19}$$

The expectation value of a hermitian operator in any state is a real number. For, by hermiticity, and $\langle u|v\rangle = \langle v|u\rangle^*$, we have

$$\langle \psi | A\psi \rangle = \langle A\psi | \psi \rangle = \langle \psi | A\psi \rangle^* \tag{20}$$

In other words, the diagonal matrix element of A is equal to its own complex conjugate. We are familiar with this: the diagonal entries of a hermitian matrix in an orthonormal basis $\langle e_i|A|e_i\rangle = A_{ii}$ are real.

• It follows from the reality of expectation values of a hermitian operator that the eigenvalues of a hermitian operator are also real. In fact, the eigenvalues are simply the expectation values in the corresponding eigenstates.

2.2.5 Commutators of operators

• Multiplication of matrices/operators is in general not commutative $AB \neq BA$ (in general). The amount by which they fail to commute is called the commutator $[A, B] = AB - BA$. Any operator commutes with itself or any power of itself $[A, A^n] = A^{n+1} - A^{n+1} = 0$. On the other hand, check that $xp - px = [x, p] = i\hbar I$ by acting on a state $\psi(x)$:

$$xp\psi = -i\hbar x\psi'(x), \quad \text{while} \quad px\psi = -i\hbar\frac{\partial}{\partial x}(x\psi) = -i\hbar x\psi'(x) - i\hbar\psi(x) \quad \Rightarrow \quad [x,p]\psi = i\hbar\psi. \tag{21}$$

• x and p are said to be canonically conjugate observables. In QM, the commutator plays the role that the Poisson bracket plays in CM. Just as the Poisson bracket $\{f, g\}$ of two observables is another observable, $\frac{1}{i\hbar}[A, B]$ is again an observable (i.e., hermitian) if A, B are hermitian. To show this it suffices to check that $[A, B]$ is anti-hermitian if A and B are hermitian.

$$([A,B])^\dagger = (AB - BA)^\dagger = B^\dagger A^\dagger - A^\dagger B^\dagger = BA - AB = -[A,B]. \tag{22}$$

An important property of the commutator is the product or Leibnitz rule, check that

$$[A, BC] = [A, B]C + B[A, C]. \tag{23}$$

• In three dimensions, we have three coordinate and momentum operators x, y, z and $p_x = -i\hbar\frac{\partial}{\partial x}, p_y = -i\hbar\frac{\partial}{\partial y}, p_z = -i\hbar\frac{\partial}{\partial z}$. It is easily seen that the momenta commute with each other and the coordinates commute among themselves, more over $[x, p_x] = i\hbar$ while $[x, p_y] = 0$ etc. These so-called **Heisenberg canonical commutation relations** may be summarised as $[x_i, p_j] = i\hbar\,\delta_{ij}$.

2.2.6 Eigenvalue problem for hermitian operators

- The **eigenvalue problem** for a linear operator (hermitian or not) is the equation $A|\psi\rangle = \lambda|\psi\rangle$. A non-zero vector $|\psi\rangle \neq 0$ that satisfies this equation for some complex number λ is called an eigenvector of A with eigenvalue λ. Taking the adjoint of the eigenvalue equation we also have

$$(A|\psi\rangle)^\dagger = \langle\psi|A^\dagger = \lambda^*\langle\psi| \tag{24}$$

So if $|\psi\rangle$ is an eigen-ket of A with eigenvalue λ, then $\langle\psi|$ is an eigen-bra of A with eigenvalue λ^*. In particular, if $A = A^\dagger$ is hermitian, then $\langle\psi|A^\dagger = \langle\psi|A = \lambda^*\langle\psi|$. In other words, if $|\psi\rangle$ is an eigen-ket of A, then $\langle\psi|$ is an eigen-bra of A with eigenvalue λ^*. We will soon show that λ is real if A is hermitian.

- The eigenstate of the position operator $\hat{x}$ with eigenvalue x' is denoted $|x'\rangle$, i.e., $\hat{x}|x'\rangle = x'|x'\rangle$. We will see that measurement of the position of a particle that is in state $|x'\rangle$ is guaranteed to give the value x'. The 'position-space' or 'coordinate-space' wave function of any state $|\psi\rangle$ is defined as the inner product $\langle x|\psi\rangle = \psi(x)$. It follows that $\psi^*(x) = \langle\psi|x\rangle$.

- Similarly, the eigenvalue problem for momentum is $\hat{p}|k\rangle = \hbar k|k\rangle$. It is conventional to write the momentum eigenvalue in terms of wave number as $\hbar k$. We will see that $|k\rangle$ is a state in which a measurement of the particle's momentum will give $\hbar k$. The momentum space wave function of a particle in state $|\psi\rangle$ is defined as $\tilde{\psi}(k) = \langle k|\psi\rangle$. $\tilde{\psi}$ is pronounced 'psi-tilde'.

- Here are some **useful facts about hermitian matrices/operators**:

1. The eigenvalues of a hermitian operator are real. This is because the eigenvalues of a hermitian operator are simply the expectation values in the corresponding eigenstates

$$A|\psi\rangle = \lambda|\psi\rangle \quad \Rightarrow \quad \langle\psi|A|\psi\rangle = \langle\psi|\lambda\psi\rangle = \lambda\langle\psi|\psi\rangle \quad \Rightarrow \quad \lambda = \frac{\langle\psi|A|\psi\rangle}{\langle\psi|\psi\rangle}. \tag{25}$$

2. Eigenvectors $|\chi\rangle, |\psi\rangle$ corresponding to distinct (necessarily real) eigenvalues $\mu \neq \lambda$ are orthogonal. To see this, we calculate $\langle\chi|A\psi\rangle$ in two ways using hermiticity and reality of eigenvalues and subtract.

$$\langle\chi|A\psi\rangle = \lambda\langle\chi|\psi\rangle \quad \text{and} \quad \langle\chi|A\psi\rangle = \langle A\chi|\psi\rangle = \langle\psi|A\chi\rangle^* = \mu^*\langle\psi|\chi\rangle^* = \mu\langle\chi|\psi\rangle. \tag{26}$$

   Thus $(\lambda - \mu)\langle\chi|\psi\rangle = 0$. Since $\lambda \neq \mu$ we must have $\langle\chi|\psi\rangle = 0$, i.e., eigenvectors corresponding to distinct eigenvalues are orthogonal.

3. It can be shown that a hermitian operator can be diagonalised by a unitary transformation $U^\dagger H U = \Lambda$ where Λ is a diagonal matrix with eigenvalues along the diagonal. Moreover, the eigenvectors of a hermitian operator can be chosen to form a complete orthonormal basis for $\mathcal{H}$

$$A|\psi_i\rangle = \lambda_i|\psi_i\rangle, \quad \langle\psi_i|\psi_j\rangle = \delta_{ij}, \quad \sum_i |\psi_i\rangle\langle\psi_i| = I, \quad . \tag{27}$$

   Furthermore, two hermitian operators which commute can be simultaneously diagonalised. In other words, there is a basis of common eigenvectors in which both are diagonal. And if they do not commute, as in the case of $[x,p] = i\hbar I$, they cannot be simultaneously diagonalised. Operators that commute are said to be compatible, we will see that they can be simultaneously measured.

4. The eigenvalue problem for the momentum operator is $\hat{p}|k\rangle = \hbar k|k\rangle$. The position space eigenfunction $\langle x|k\rangle$ of the momentum operator is a plane wave. $\hat{p}\psi(x) = \hbar k\psi(x)$ becomes $-i\hbar\psi' = \hbar k\psi$ or $\psi = Ae^{ikx}$. We will choose $A = 1$. In other words $\langle x|k\rangle = e^{ikx}$ and so $\langle k|x\rangle = e^{-ikx}$.

5. The position-space or coordinate-space eigenfunctions of the position operator are delta-functions. Let's see why. The eigenvalue problem is

$$\hat{x}\ \psi(x) = x\psi(x) = \lambda\psi(x) \quad \text{where} \quad \lambda \quad \text{is a constant.} \tag{28}$$

The only way this can be satisfied for all x is for $\psi(x)$ to vanish at all $x \neq \lambda$. Now if $\psi(x)$ were to vanish at $x = \lambda$ as well, then it would be the zero function and not qualify as a non-trivial eigenvector. The value of $\psi(x)$ at $x = \lambda$ can either be finite or $\psi(\lambda) = \pm\infty$. If $|\psi(\lambda)| < \infty$, then the state will have zero norm and cannot describe a particle that can be found somewhere. So ψ must be infinite at $x = \lambda$. In fact, $\psi(x)$ is proportional to the Dirac delta function. It is normalized so that $\psi(x) = \delta(x - \lambda)$. It is conventional to denote the position eigenvalue by x' rather than λ. So $\delta(x - x')$ is an eigenfunction of the position operator with eigenvalue x', it is a function of x that is zero every where except at x'. Think of it as a limit of functions that are sharply peaked at $x = x'$. Thus the coordinate space wave function of the eigenstate $|x'\rangle$ of $\hat{x}$ is $\langle x|x'\rangle = \delta(x - x')$. Now if we have two position eigenstates $|x'\rangle$ and $|x''\rangle$, then their coordinate space wave functions are $\langle x|x'\rangle = \delta(x - x')$ and $\langle x|x''\rangle = \delta(x - x'')$. Their inner product is

$$\langle x''|x'\rangle = \int \delta(x - x'')\delta(x - x')\ dx = \delta(x' - x''). \tag{29}$$

So position eigenstates are orthogonal and 'delta-normalized'. They form a complete set in the sense that they satisfy a completeness relation

$$\int dx\ |x\rangle\langle x| = I. \tag{30}$$

To see this, take the matrix elements of the LHS between coordinate basis states $|x'\rangle$ and $|x''\rangle$

$$\int dx\ \langle x'|x\rangle\langle x|x''\rangle = \int dx\ \delta(x - x')\delta(x - x'') = \delta(x' - x''). \tag{31}$$

On the other hand, the matrix elements of the identity are also the same $\langle x'|I|x''\rangle = \langle x'|x''\rangle = \delta(x' - x'')$. Since $\int dx\ |x\rangle\langle x|$ and I have the same matrix elements, they are equal.

- Similarly, momentum eigenstates form a complete set

$$\int \frac{dk}{2\pi}|k\rangle\langle k| = I. \tag{32}$$

Check this by evaluating the matrix elements between position basis states $|x'\rangle$ and $|x''\rangle$. On the rhs we get $\langle x'|I|x''\rangle = \delta(x' - x'')$. On the lhs we get the same using the Fourier representation of the delta function

$$\int_{-\infty}^{\infty} \frac{dk}{2\pi}\ \langle x'|k\rangle\langle k|x''\rangle = \int_{-\infty}^{\infty} \frac{dk}{2\pi}\ e^{ikx'}e^{-ikx''} = \delta(x' - x''). \tag{33}$$

How do we get the last equality? If $x' = x''$ then we are integrating the function 1, and the answer should be infinite, and indeed $\delta(0) = \infty$. On the other hand, when $x' \neq x''$, then we have

$$\int \frac{dk}{2\pi}[\cos(k(x'-x'')) + i\sin(k(x'-x''))] = 0 \tag{34}$$

Since the average value of both the sine and cosine functions is zero.

• Momentum eigenstates with distinct wave numbers are orthogonal (as we expect for the eigenstates of a hermitian operator)

$$\langle k'|k''\rangle = \int dx\, \langle k'|x\rangle\langle x|k''\rangle = \int dx\, e^{-ik'x}e^{ik''x} = 2\pi\,\delta(k'-k''). \tag{35}$$

6. Among hermitian operators, the positive operators are particularly interesting physically. A hermitian operator is positive (or non-negative) if its expectation value in every state is non-negative $\langle\psi|A|\psi\rangle \geq 0$, for all $\psi \in \mathcal{H}$. Since eigenvalues are simply the expectation values in eigenstates, we see that positive operators have non-negative eigenvalues. If A is any linear operator, then we check that $A^\dagger A$ and $AA^\dagger$ are both hermitian and positive operators.

   $$\text{E.g.}\quad (AA^\dagger)^\dagger = A^{\dagger\dagger}A^\dagger = AA^\dagger. \tag{36}$$

   To check positivity, we work out the expectation value in any state:

   $$\langle\psi|A^\dagger A|\psi\rangle = \langle A\psi|A\psi\rangle = ||A\psi||^2 \geq 0 \quad \text{and} \quad \langle\psi|AA^\dagger|\psi\rangle = \langle A^\dagger\psi|A^\dagger\psi\rangle = ||A^\dagger\psi||^2 \geq 0. \tag{37}$$

   An example is kinetic energy $T = \frac{1}{2m}p^2 = \frac{1}{2m}p^\dagger p = \frac{1}{2m}pp^\dagger$, since $p = p^\dagger$ is hermitian. So we may conclude that the energy eigenvalues of a free particle must all be non-negative.

2.2.7 Measured value of observables in states and interpretation of expectation values

• Measurement of an observable A in state ψ of unit norm produces a real number that is one of the eigenvalues of A. Born's probability postulate: Suppose we have several identically prepared systems in the same state ψ and we measure the value of A in each system and collect the values. Then the frequency of occurrence of the measured value λ is $p_\lambda = |\langle\psi_\lambda|\psi\rangle|^2$ where ψ_λ is the unit norm eigenstate corresponding to the eigenvalue λ.

• The expectation value of an observable A in a state ψ is the mean value obtained when A is measured on many copies of the system prepared in the same state ψ. How do we see this? Each measurement gives a (possibly different) eigenvalue λ with probability p_λ. So the mean measured value is a sum over the eigenvalues of A (counted with multiplicity)

$$\sum_\lambda p_\lambda\lambda = \sum_\lambda \lambda|\langle\psi|\psi_\lambda\rangle|^2 = \sum_\lambda \lambda\langle\psi|\psi_\lambda\rangle\langle\psi|\psi_\lambda\rangle^* = \sum_\lambda \lambda\langle\psi|\psi_\lambda\rangle\langle\psi_\lambda|\psi\rangle = \sum_\lambda\langle\psi A|\psi_\lambda\rangle\langle\psi_\lambda|\psi\rangle = \langle\psi|A|\psi\rangle. \tag{38}$$

We used the eigenvalue equation and completeness of the eigenvectors $\sum_\lambda |\psi_\lambda\rangle\langle\psi_\lambda| = I$.

• **Physical interpretation** of $\langle x|k'\rangle = e^{ik'x}$ and $\langle x|x'\rangle = \delta(x-x')$ in the context of probability of results of measurements. Suppose a particle is in a position eigenstate $|x'\rangle$. Then its coordinate space wave function is $\langle x|x'\rangle = \delta(x-x')$. Now suppose we make a measurement of its position. Then the probability of getting the value x is $p_x \propto |\langle x|x'\rangle|^2$. Notice that $p_x = 0$ for $x \neq x'$. So if we measure the position of a particle known to be in the position eigenstate $|x'\rangle$, then the only value of position that can result is x' itself.

• Suppose a particle is in a position eigenstate $|x'\rangle$. Then its momentum space wave function is $\langle k|x'\rangle = e^{-ikx'}$. Suppose we make a measurement of its momentum. Then the probability of getting the value $\hbar k$ is $p_k \propto |\langle k|x'\rangle|^2 = |e^{ikx}|^2 = 1$. In other words, all momenta are equally probable. This makes physical sense in light of the Heisenberg uncertainty principle. If the particle is in a position eigenstate, then its position is known with perfect accuracy. So we would expect its momentum to be maximally uncertain. And indeed, what we find is that all possible momenta are equally likely, so we have no knowledge as to what the result of a momentum measurement may give.

• After measuring an observable A and getting the eigenvalue λ, the state of the system 'collapses' from state ψ to eigenstate ψ_λ corresponding to the eigenvalue λ ($A|\psi_\lambda\rangle = \lambda|\psi_\lambda\rangle$).

• **Reproducibility of measurements:** If A is measured again, soon after a previous measurement of A, then the same value λ will be obtained and the system will remain in the same eigenstate of A. If a system is in an eigenstate $|\psi_0\rangle$ of energy, then we know in advance that measurement of energy will result only in the eigenvalue E_0 and that the state will not change after the measurement.

• If two observables (hermitian operators A, B) commute, they have common eigenvectors and are simultaneously diagonalisable. We say they are simultaneously measurable or compatible. What this means is that if A has been measured, and a value a obtained, then a measurement of B will not affect the eigenstate $|\psi_a\rangle$ of A to which the system had collapsed. This is because $|\psi_a\rangle$ is an eigenstate of B as well. An immediate measurement of B will certainly result in the eigenvalue of B corresponding to the eigenvector ψ_a. A subsequent measurement of A will again result in the value a. It is in this sense that A and B can be simultaneously measured.

• Let us indicate why commuting observables have common eigenfunctions. Suppose A is hermitian and has eigenvalues λ_i (assumed non-degenerate) with corresponding eigenfunctions ψ_i, so $A\psi_i = \lambda_i\psi_i$. Non-degeneracy means that each eigenspace is one dimensional. Now suppose B commutes with A. Then consider $B(A\psi)$, we evaluate it in two ways. On the one hand, $B(A\psi_i) = \lambda_i B\psi_i$. On the other, $BA\psi_i = AB\psi_i$. Thus $A(B\psi_i) = \lambda_i(B\psi_i)$. In other words, both ψ_i and $B\psi_i$ are eigenfunctions of A with the same eigenvalue. Since the eigenspaces of A are assumed one dimensional $B\psi_i$ and ψ_i must be linearly dependent, i.e. multiples of eachother: $B\psi_i = \mu_i\psi_i$. In other words we have shown that an eigenfunction of A is also an eigenfunction of B!

• It is worth noting that measurement of an observable in a state ψ is a complicated process that is still not well-understood, and is certainly not the multiplication of the operator A with the state vector ψ.

2.2.8 Relation between wave function in position and momentum space

• The wave function is a complete specification of the state of a quantum mechanical system, just as giving the position and momentum of a particle completely specifies its classical state. For a particle moving in 3-space, the coordinate space wave function is $\psi(x, y, z; t)$. For a system of n particles, the coordinate space wave function is a function of the three coordinates of each of the n particles $\psi(\vec{r}_1, \vec{r}_2, \cdots \vec{r}_n; t)$. In other words, the coordinate space wave function is a (time-dependent) function on the classical configuration space of the system.

• We have seen that the position space wave function of a state $|\psi\rangle$ is defined as $\psi(x) = \langle x|\psi\rangle$. Let us denote a momentum eigenstate with momentum eigenvalue $p = \hbar k$ by $|k\rangle$, where k is the wave number. Then the momentum space wave function of the same state $|\psi\rangle$ is $\tilde{\psi}(k) = \langle k|\psi\rangle$.

The point is that $|\psi\rangle$ is an abstract state vector. We can study it ('represent it') via its components in any basis. In particular, we may look at its components $\langle x|\psi\rangle = \psi(x)$ in the basis of position eigenstates or its components $\langle k|\psi\rangle = \tilde{\psi}(k)$ in the basis of momentum eigenstates. Let us see how $\psi(x)$ is related to $\tilde{\psi}(k)$.

• Now inserting a complete set of momentum eigenstates and using $\langle x|k\rangle = e^{ikx}$,

$$\psi(x) = \langle x|\psi\rangle = \int \frac{dk}{2\pi}\langle x|k\rangle\langle k|\psi\rangle = \int \frac{dk}{2\pi} e^{ikx}\tilde{\psi}(k) \tag{39}$$

So the position space wave function is the inverse-Fourier transform of the momentum space wave function. Similarly, we have the Fourier transform

$$\tilde{\psi}(k) = \int dx e^{-ikx}\psi(x). \tag{40}$$

• $\psi(x)$ and $\tilde{\psi}(k)$ are to be compared with the state of a classical mechanical system, which is given by a *simultaneous* specification of coordinates **and** momenta. In the quantum theory, ψ cannot depend on both the coordinates and momenta (in an arbitrary manner). This is related to the uncertainty principle.

• The absolute square of the wave function $|\psi(x,t)|^2 = \psi^*(x,t)\psi(x,t)$ gives the probability density for finding the particle at location x at time t. Similarly, $|\tilde{\psi}(k,t)|^2 \frac{dk}{2\pi}$ is the probability of finding the particle in momentum interval $[k, k+dk]$ at time t.

2.2.9 Schrödinger time evolution of states

• When left to itself, the state of the system evolves according to the Schrödinger equation $i\hbar\frac{\partial}{\partial t}|\psi(t)\rangle = H|\psi(t)\rangle$. H is the hermitian hamiltonian. Given the initial state $|\psi(0)\rangle$, the SE determines the state at subsequent times, just as Hamilton's equations $\dot{x} = \frac{\partial H}{\partial p}, \dot{p} = -\frac{\partial H}{\partial x}$ do in classical mechanics.

• In the position basis, the SE is

$$i\hbar\frac{\partial}{\partial t}\langle x|\psi(t)\rangle = \langle x|H\psi(t)\rangle \quad \text{or} \quad i\hbar\frac{\partial\psi(x,t)}{\partial t} = (H\psi)(x,t) \tag{41}$$

For a particle in a potential $(H\psi)(x,t) = -\frac{\hbar^2}{2m}\frac{\partial^2\psi(x,t)}{\partial x^2} + V(x)\psi(x,t)$, and we get

$$i\hbar\frac{\partial\psi(x,t)}{\partial t} = -\frac{\hbar^2}{2m}\frac{\partial^2\psi}{\partial x^2} + V(x)\psi(x). \tag{42}$$

The SE is a linear PDE, first order in time and second order in space derivatives of the unknown ψ. Contrast this with Newton's equation which in general is a system of non-linear ODEs for $x_i(t)$.

• We often need to work with the adjoint of the Schrodinger equation, which is obtained using $H = H^\dagger$

$$-i\hbar\frac{\partial}{\partial t}\langle\psi(t)| = \langle\psi(t)|H. \tag{43}$$

In the coordinate basis, the adjoint of the SE reads

$$-i\hbar\frac{\partial}{\partial t}\langle\psi(t)|x\rangle = \langle\psi(t)|H|x\rangle = \langle H\psi|x\rangle = \langle x|H\psi\rangle^* \quad \Rightarrow \quad -i\hbar\frac{\partial}{\partial t}\psi^*(x,t) = ((H\psi)(x))^* \tag{44}$$

or $-i\hbar\frac{\partial}{\partial t}\psi^*(x,t) = -\frac{\hbar^2}{2m}\frac{\partial^2\psi^*(x)}{\partial x^2} + V(x)\psi^*(x)$ for a particle in a real potential $V(x)$. So in the coordinate basis, the adjoint of the SE is just its complex conjugate.

2.2.10 Separation of variables, stationary states, time-independent Schrodinger equation

- The problem of time-evolution is to solve the Schrodinger equation $i\hbar\frac{\partial|\Psi(t)\rangle}{\partial t} = \hat{H}|\Psi(t)\rangle$ given the initial state $\Psi(t=0)\rangle$. For a particle in a potential $V(x)$, the SE is a LINEAR partial differential equation for the unknown function $\Psi(x,t) = \langle x|\Psi(t)\rangle$.

$$i\hbar\frac{\partial\Psi(x,t)}{\partial t} = (H\Psi)(x,t) = -\frac{\hbar^2}{2m}\frac{\partial^2\Psi(x,t)}{\partial t} + V(x)\Psi(x,t). \tag{45}$$

To solve it we use the method of separation of variables. We look for separable solutions in the form of a product $\Psi(x,t) = \psi(x)T(t)$. Now, not every solution of the SE is such a product. But due to the linearity of the equation linear combinations of solutions are again solutions. The idea is to find sufficiently many separable solutions so that every solution can be written as a linear combination of separable solutions. Putting the 'ansatz' (guess) $\Psi(x,t) = \psi(x)T(t)$ into the equation, we get

$$i\hbar\dot{T}(t)\psi(x) = T(t)(H\psi)(x). \tag{46}$$

Dividing by $T\psi$ we get

$$i\hbar\frac{\dot{T}(t)}{T(t)} = \frac{(H\psi)(x)}{\psi(x)} = E. \tag{47}$$

LHS depends only on time while the RHS depends only on position, provided H is not explicitly time-dependent. The only way these can be equal is for both to equal the same constant, say E, so-named, since it turns out to have the physical meaning of energy eigenvalue. Now we have two separate equations. The one for $T(t)$ has the solution $T(t) = c\exp(-iEt/\hbar)$. The other equation

$$(H\psi)(x) = E\psi(x) \quad \text{or} \quad \langle x|H|\psi\rangle = E\langle x|\psi\rangle \quad \text{or} \quad H|\psi\rangle = E|\psi\rangle \tag{48}$$

is simply the eigenvalue equation for the hamiltonian operator. It is also called the time-independent Schrodinger eigenvalue equation. It typically has lots of solutions, namely all the eigenstates $|\psi_n\rangle$ of the hamiltonian, with their corresponding energy eigenvalues E_n. As for any hermitian operator, we can take these $|\psi_n\rangle$ to be orthonormal. Thus the separable solutions of the Schrodinger equation are

$$\Psi_n(x,t) = c_n\psi_n(x)e^{-iE_nt/\hbar}. \tag{49}$$

where ψ_n are eigenstates of the hamiltonian. These separable solutions are called stationary states since the probability density in these states $P(x,t) = |\Psi(x,t)|^2 = |c_n|^2|\psi_n(x)|^2$ are independent of time. Stationary states have the simplest possible time dependence of all solutions of the Schrodinger equation, i.e., sinusoidal or harmonic time dependence.

- Now the general solution of the SE is got by taking a linear combination of stationary states

$$\Psi(x,t) = \sum_n c_n\psi_n(x)e^{-iE_nt/\hbar}. \tag{50}$$

To find the solution of the initial value problem, we must choose the c_n so that the initial state is $|\Psi(0)\rangle$. In other words, we must have

$$\sum_n c_n|\psi_n\rangle = |\Psi(0)\rangle \tag{51}$$

To find the c_n we take the inner product with $|\psi_m\rangle$, and use orthogonality of energy eigenstates

$$\sum_n c_n\langle\psi_m|\psi_n\rangle = \sum_n \delta_{mn}c_n = c_m = \langle\psi_m|\Psi(0)\rangle \quad \Rightarrow \quad c_m = \int \psi_m^*(x)\Psi(x,0)\,dx. \tag{52}$$

Thus we have solved the initial value problem for the Schrodinger equation.

2.2.11 Conserved probability density and current

- The absolute square of the wave function $|\psi(x,t)|^2 = \psi^*(x,t)\psi(x,t)$ gives the probability density for finding the particle at location x at time t. Suppose n copies of a system are prepared in the same quantum mechanical state $\psi(x)$. (For example, we could have a hydrogen atom in its ground state in each of 100 different boxes) Then a measurement of the position of each particle (at the same time) gives a (possibly) different result (this is an experimental fact). Born's statistical interpretation of the wave function is that, as $n \to \infty$, the distribution of position measurements approaches the probability density $|\psi(x,t)^2|$.

- To qualify as a probability density, the total probability of finding the particle anywhere must be one. In other words, we need $||\psi||^2 = \int dx\ |\psi(x,t)|^2 = 1$. However, there could be a problem. For consistency, the total probability of finding the particle *somewhere* must remain equal to one at all times, total probability must be conserved. This is indeed the case, as is checked using the Schrödinger equation and its adjoint

$$i\hbar\frac{\partial}{\partial t}\langle\psi|\psi\rangle = i\hbar\langle\psi|H\psi\rangle - i\hbar\langle\psi H|\psi\rangle = 0. \tag{53}$$

In other words, if the wave function is normalized to one initially ($t = 0$), then it continues to have norm one in the future. This is called global conservation of probability. But it is not merely the total probability that is conserved. Probability cannot jump from one place to another, it flows continuously like a fluid. There is a local conservation of probability just like for mass in a fluid. The rate of increase of mass of fluid in a box is equal to the inward flux of fluid across the walls of the box (provided there isn't a source/sink of fluid inside the box). The probability density $|\psi(x,t)|^2$ satisfies a continuity equation with an associated probability current. Consider a particle in a potential

$$\begin{aligned} i\hbar\,\partial_t(\psi^*\psi) = i\hbar\,(\psi_t^*\psi + \psi^*\psi_t) &= \left(\frac{\hbar^2}{2m}\psi^{*\prime\prime} - V\psi^*\right)\psi + \psi^*\left(-\frac{\hbar^2}{2m}\psi'' + V\psi\right) \\ &= \frac{\hbar^2}{2m}\left[\psi^{*\prime\prime}\psi - \psi^*\psi''\right] = \frac{\hbar^2}{2m}\partial_x\left(\psi^{*\prime}\psi - \psi^*\psi'\right) \end{aligned} \tag{54}$$

Let $P(x,t) = |\psi(x,t)|^2$ and define the probability current density

$$j(x,t) = \frac{\hbar}{2mi}\left(\psi^*\psi' - \psi^{*\prime}\psi\right), \quad \text{then} \quad \partial_t P(x,t) + \partial_x j(x,t) = 0. \tag{55}$$

The last equation is called the law of local conservation of probability (in differential form) or a continuity equation. To interpret this formula we consider how the probability for the particle to be in an interval $[x_0, x_1]$ changes with time. So integrate $\partial_t P + \partial_x j = 0$ over this interval at a fixed time t to get the law of local conservation of probability in integral form:

$$\partial_t \int_{x_0}^{x_1} P(x)\, dx + \int_{x_0}^{x_1} \frac{\partial j(x)}{\partial x}\, dx = 0 \quad \Rightarrow \quad \partial_t \int_{x_0}^{x_1} P(x)\, dx = j(x_0) - j(x_1) \tag{56}$$

by the fundamental theorem of calculus. This equation says the rate of increase of probability in $[x_0, x_1]$ equals the probability current flowing in at x_0 minus that flowing out at x_1.

• All of this also works in three dimensions. The rate of increase of probability in a region (volume) Ω must equal the inward flux of probability across the surface $\partial\Omega$ that borders Ω.

$$\begin{aligned} P(\vec r, t) = \psi^*(\vec r, t)\psi(\vec r, t), \qquad & \vec j = \frac{\hbar}{2mi}\left[\psi^* \left(\nabla\psi\right) - \left(\nabla\psi^*\right)\psi\right] = \frac{\hbar}{m}\Im\psi^*\nabla\psi \\ \partial_t P(\vec r, t) + \nabla\cdot\vec j(x,t) \quad = \quad & 0, \quad \text{i.e.} \quad \frac{\partial\rho}{\partial t} + \frac{\partial j_1}{\partial x} + \frac{\partial j_2}{\partial y} + \frac{\partial j_3}{\partial z} = 0. \\ \partial_t \int_\Omega P(\vec r, t) d^3r + \int_\Omega d^3r \nabla\cdot\vec j = 0 \quad \text{or} \quad & \partial_t \int P(\vec r, t) d^3 r = -\int_{\partial\Omega} \vec j\cdot d\vec S. \end{aligned} \tag{57}$$

$d\vec S$ is the outward pointing area element on the bounding surface $\partial\Omega$. It says that the rate of increase of probability in a region must equal the inward flux of probability current across the surface of the region. We used the divergence theorem to write the volume integral of a divergence as a surface integral.

2.3 Ehrenfest's theorem

• The expectation values $\langle x\rangle, \langle p\rangle, \langle E\rangle$ etc are functions of time (space has been integrated over). The average position and momentum of an electron will depend on time in a way governed by the Schrödinger equation. According to Ehrenfest's theorem, these expectation values evolve as do the corresponding classical variables, whose evolution is given by Newton's/Hamilton's equations! E.g. $\frac{d\langle x\rangle}{dt} = \frac{\langle p\rangle}{m}$, so the average position evolves in the same way as given by the first of Hamilton's equations. To see this and related results, we first derive a general equation for the time evolution of expectation value of an observable A in a unit-norm state that evolves via the SE

$$i\hbar\frac{\partial}{\partial t}\langle\psi|A|\psi\rangle = -\langle\psi|HA|\psi\rangle + \langle\psi|AH|\psi\rangle = \langle\psi|[A,H]|\psi\rangle. \tag{58}$$

• Putting $A = H$ and using $[H, H] = 0$ shows that the average energy (expectation value of hamiltonian) is constant $\frac{\partial\langle\hat H\rangle}{\partial t} = 0$. This is the analogue of the classical constancy of energy along a trajectory.

• Taking $A = p$ we find the time evolution of mean momentum for a particle subject to the hamiltonian $H = \frac{p^2}{2m} + V$. Show that

$$[p, H] = [p, V] = -i\hbar V' \tag{59}$$

Thus we have

$$\frac{\partial\langle p\rangle}{\partial t} = \langle -V'\rangle. \tag{60}$$

Thus Newton's second law (or the second of Hamilton's equations) $\dot{p} = -V'(x)$ continues to hold in quantum mechanics, but in the sense of expectation values. The average momentum evolves as though it is a classical variable subject to an 'average force'!

- If $A = x$, then $[x, H] = [x, \frac{p^2}{2m}] = \frac{i\hbar p}{m}$. So

$$\frac{\partial \langle x \rangle}{\partial t} = \left\langle \frac{p}{m} \right\rangle. \tag{61}$$

This is the first of Hamilton's equations $\dot{x} = \frac{\partial H}{\partial p} = \frac{p}{m}$, but now in the sense of expectation values.

- So if the electron is in the initial state $\psi(x, t = 0)$, Schrödinger's equation tells us how the state evolves in time. We have used this to determine the motion of the average position of the electron and found that it is related to the average momentum in the same way as the actual position and momentum of a particle are related by Hamilton's equation of classical mechanics. To the extent that the expectation value of x provides an approximate position for a localized electron wave packet, we see that the quantum mechanical motion of the wave-packet mimics the classical motion of a particle. However, the wave packet typically spreads out in time, and ceases to be well-described by merely its mean position. This reduces the utility of the Ehrenfest result in determining where a quantum particle may be found at later times, based purely on its classical motion.

2.4 Generators of transformations, symmetries and conservation laws

2.4.1 Generators of space translations

- What does it mean to say that momentum $\hat{p}$ generates translations, angular momentum generates rotations and the hamiltonian generates time translations?

- Consider first a system with one degree of freedom so that the wave function or potential is a function of one variable $f(x)$. Recall that $\hat{p} = -i\hbar \frac{\partial}{\partial x}$. Under a translation $x \mapsto x + \delta a$, the function becomes

$$f(x+\delta a) = f(x) + \delta a\, f'(x) + \frac{1}{2}(\delta a)^2 f''(x) + \cdots = f(x) - \delta a\, \frac{1}{i\hbar}(-i\hbar)\frac{\partial}{\partial x} f(x) + \cdots = \left(I + \frac{i\hat{p}\,\delta a}{\hbar} + \cdots \right) f.$$

It is in this sense that momentum generates an infinitesimal translation. More generally, for a finite translation one gets $f(x + a) = e^{ipa/\hbar} f(x)$, as we will see below. Physically, f could be a wave function or a potential. If $pf = 0$ we say that f is translation invariant since $f(x + \delta x) = f(x)$. For example, a constant potential $V(x) = V_0$ is translation invariant. Or a potential $V(x, y)$ is translation invariant in the z-direction. We say it is annihilated by p_z, since $p_z V(x, y) = 0$.

- In 3d, under an infinitesimal translation $\mathbf{r} \to \mathbf{r} + \delta\mathbf{a}$, a function changes to $f(\mathbf{r} + \delta\mathbf{a})$. Using the Taylor expansion

$$f(\mathbf{r} + \delta\mathbf{a}) = f(\mathbf{r}) + \delta\mathbf{a} \cdot \nabla f(\vec{r}) + \frac{1}{2}\delta a_i\, \delta a_j\, \partial_i \partial_j f(\mathbf{r}) + \cdots \tag{62}$$

We can write this in terms of momentum

$$f(\mathbf{r} + \delta a) \approx (1 + \frac{i}{\hbar}\delta\mathbf{a} \cdot \hat{\mathbf{p}}) f(\mathbf{r}) \tag{63}$$

And for a finite translation by $\mathbf{a} = n\,\delta\mathbf{a}$ we apply the infinitesimal translation n times and let $n \to \infty$ holding $\mathbf{a}$ fixed

$$f(\mathbf{r}+\mathbf{a}) = \lim_{n\to\infty}\left(I + \frac{i}{\hbar}\frac{1}{n}\mathbf{a}\cdot\hat{\mathbf{p}}\right)^n f(\mathbf{r}) = e^{\frac{i}{\hbar}\mathbf{a}\cdot\hat{\mathbf{p}}} f(\mathbf{x}). \tag{64}$$

We say that $\hat{\mathbf{p}}$ is the infinitesimal generator of translations. If we expand the exponential we will recover the Taylor expansion.

2.4.2 Generator of time translation and time evolution operator

• To find the generator of time translations, we observe that the Schrödinger equation $i\hbar\dot{\psi} = H\psi$ may be solved for short times to write

$$\psi(t+\delta t) \approx \psi(t) - \frac{iH\delta t}{\hbar}\psi(t) \quad \Rightarrow \quad \psi(t+\delta t) = \left[1 - \frac{iH\delta t}{\hbar}\right]\psi(t) \tag{65}$$

In this sense, the hamiltonian generates infinitesimal time-translations of the wave function. As above, a finite time-translation is given by composing a succession of infinitesimal time-evolutions $|\psi(t)\rangle = e^{-i\hat{H}t/\hbar}|\psi(0)\rangle$. The operator $U(0,t) = e^{-iHt/\hbar}$ is called the time evolution operator. It is unitary if the hamiltonian is hermitian, since then $U^\dagger U = e^{iHt/\hbar}e^{-iHt/\hbar} = I$. We may write the time evolution operator as an infinite exponential series

$$U(0,t) = I - \frac{iHt}{\hbar} + \left(\frac{-it}{\hbar}\right)^2 \frac{1}{2!}H^2 + \left(\frac{-it}{\hbar}\right)^3 \frac{1}{3!}H^3 + \ldots \tag{66}$$

• Though U is not an observable, it is still an important operator in quantum mechanics. It is easy to find its eigenstates and eigenvalues. Suppose the eigenstates of the hamiltonian are $H|\psi_n\rangle = E_n|\psi_n\rangle$. Then one finds that the energy eigenstates are also eigenstates of the time evolution operator and that

$$U|\psi_n\rangle = e^{-iE_nt/\hbar}|\psi_n\rangle. \tag{67}$$

The eigenvalues of U are complex numbers of unit magnitude $e^{-iE_nt/\hbar}$. The eigenvalues of the time evolution operator give the time evolution of stationary states $\Psi_n(t) = \psi_n e^{-iE_nt/\hbar}$.

2.4.3 Angular momentum as generator of rotations

• The small change in $\vec{r}$ upon making a rotation by small angle $\delta\phi$ counter clockwise about the axis $\delta\vec{\phi}$ is $\delta\vec{r} = \delta\vec{\phi}\times\vec{r}$.

$$f(\vec{r} + \delta\vec{\phi}\times\vec{r}) \approx f(\vec{r}) + \left(\delta\vec{\phi}\times\vec{r}\right)\cdot\vec{\nabla}f(\vec{r}) \tag{68}$$

Now $(A\times B)\cdot C = A\cdot(B\times C)$ since both of them measure the volume of a parallelepiped defined by the vectors A, B, C. Thus using $p = -i\hbar\nabla$,

$$f(\vec{r} + \delta\vec{\phi}\times\vec{r}) \approx f(\vec{r}) + \frac{i}{\hbar}\delta\vec{\phi}\cdot(r\times p)\,f(\vec{r}) = \left(I + \frac{i}{\hbar}\delta\vec{\phi}\cdot\vec{L}\right)f(\vec{r}). \tag{69}$$

So angular momentum generates rotations. For a finite rotation counter clockwise about axis $\vec{\phi}$ by angle $|\vec{\phi}|$ we have

$$f(\vec{r}) \mapsto e^{\frac{i}{\hbar}\vec{\phi}\cdot\vec{L}} f(\vec{r}). \tag{70}$$

Here $\mathbf{L}\cdot\phi = L_x\phi_x + L_y\phi_y + L_z\phi_z$. We say that L_x generates a rotation about the x axis, L_y generates a rotation about the y axis and L_z generates a counter clockwise rotation about the z.

• Now if $V(r)$ is a spherically symmetric (central) potential, then we know it is unaffected by rotations. It follows that $\mathbf{L}V(r) = 0$. In other words, $L_i V(r) = 0$, or the angular momentum operators annihilate the potential. It follows that $L^2 = \sum_i L_i^2$ also annihilates a spherically symmetric potential. Check that this implies $[L^2, V(r)] = 0$. This result will be used in studying motion in a central potential, where one shows L^2 commutes with the hamiltonian (say of the hydrogen atom), so that L^2 and H are simultaneous observables.

2.4.4 Symmetries and conservation laws

• An observable A (not explicitly dependent on time) is said to generate a symmetry of the hamiltonian H if it commutes with the hamiltonian $[A,H] = 0$. A symmetry leads to a conserved quantity in the following sense: the expectation value of A in any state (evolving via the Schrödinger equation) is independent of time. Let us see why. Recall from Ehrenfest's theorem, that

$$i\hbar\frac{\partial}{\partial t}\langle\psi|A|\psi\rangle = \langle\psi|[A,H]|\psi\rangle. \tag{71}$$

In addition, we know that time evolution does not change the norm of a state, $\partial_t\langle\psi|\psi\rangle = 0$. So if $[A,H] = 0$, then the expectation value $\langle A\rangle_\psi = \frac{\langle\psi|A|\psi\rangle}{\langle\psi|\psi\rangle}$ is time-independent.

• The classical version of this relation between symmetries and conservation laws is called Noether's theorem.

• E.g., consider a particle moving in a potential $V(\mathbf{r})$. If the potential is independent of the z-coordinate, then $V(\mathbf{r}+a\hat{z}) = V(\mathbf{r})$ so that $\hat{z}\cdot\nabla V = 0$ or $p_z(V) = 0$. It follows that $[p_z, V] = 0$ and since $[p_z, \frac{\mathbf{p}^2}{2m}] = 0$, $[p_z, H] = 0$. So if the potential is z-translation-invariant then $\langle\psi|p_z|\psi\rangle$ is conserved in time.

• Similarly, if the potential is invariant under rotations (a central potential), then each component of angular momentum is conserved in the sense of expectation values $\partial_t\langle\psi|\mathbf{L}|\psi\rangle = 0$.

• Another consequence of a symmetry ($[A,H] = 0$), is that if the initial state $\psi(0)$ is an eigenstate of A with eigenvalue a, then the system remains in that eigenstate and measurement of A at any time will result in the value a. Indeed, suppose the initial state is $\psi(0)$, then we know that $|\psi(t)\rangle = e^{-iHt/\hbar}|\psi(0)\rangle$. If $[A,H] = 0$, it is easily checked that $|\psi(t)\rangle$ is an eigenvector of A

$$A|\psi(t)\rangle = Ae^{-iHt/\hbar}|\psi(0)\rangle = e^{-iHt/\hbar}A|\psi(0)\rangle = e^{-iHt/\hbar}a|\psi(0)\rangle = a|\psi(t)\rangle. \tag{72}$$

• If A is a symmetry of the hamiltonian H, then A and H are commuting observables $[A,H] = 0$ and therefore have a common basis of eigenvectors where both operators are diagonal[1]. It follows that energy eigenstates can be labelled by energy as well as by the eigenvalue of the symmetry A. This is particularly important when there are several energy levels that are degenerate. The degenerate levels can often be distinguished by the eigenvalue of a symmetry

---

[1] Suppose $A\psi = \lambda\psi$ and for simplicity assume that λ is a non-degenerate eigenvalue of A. Then $A(H\psi) = HA\psi = \lambda(H\psi)$ so that $H\psi$ is also an eigenvector of A with the same eigenvalue λ. As this we assumed λ was a nondegenerate eigenvalue, we must have $H\psi = \mu\psi$ for some μ. So ψ is a common eigenvector of both A and H.

generator. For example, the hamiltonian of the hydrogen atom commutes with L_z as well as with L^2, which are a set of three commuting operators. So energy levels can be labelled with the eigenvalues $\hbar m$ and $\hbar l(l+1)$ of L_z and L^2 in addition to the energy eigenvalue E_n which depends only on the principal quantum number.

3 Simple Harmonic oscillator

• Small oscillations of a spring about equilibrium are governed by Hooke's restoring force $-kx$ or potential energy $\frac{1}{2}kx^2$. Classical eq. $\ddot{x} = -kx$ has solution $x(t) = A\cos(\omega t) + B\sin(\omega t)$ where the angular freq. $\omega = \sqrt{k/m}$. So the hamiltonian is $H = p^2/2m + \frac{1}{2}m\omega^2 x^2$. Larger oscillations often require corrections to Hooke's law, and the inclusion of anharmonic forces. Hooke's law is called the linearized (since force is linear in displacement and equation of motion is linear) or harmonic approximation.

• Approximation applies to oscillations of many systems about a point of equilibrium. For example, for small amplitude motion in neighborhood of a minimum x_0 of a potential $V(x) = V(x_0) + \frac{1}{2}V''(x_0)(x-x_0)^2 + \ldots$, as long as the second derivative of the potential is non-zero at x_0. The omitted higher order terms in the potential are 'anharmonic' terms.

• Classically, for fixed energy, motion lies between the turning points, where the particle comes instantaneously to rest. The turning points are the solutions of $E = V(x_0) = \frac{1}{2}kx_0^2$. All classical trajectories are bound, the particle cannot escape to $|x| \to \infty$ for any fixed energy, unlike the free particle. Newton's equation $m\ddot{x} = -kx$ has solutions $x(t) = x_0\cos(\omega t + \delta)$ where the amplitude x_0 and phase δ are determined by initial conditions.

• Quantum mechanically, to study the dynamics we reduce the time-dependent Schrodinger equation to the time-independent one for energy eigenstates by separation of variables. The Schrodinger eigenvalue problem is

$$-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\psi(x) + \frac{1}{2}m\omega^2 x^2\psi(x) = E\psi(x). \tag{73}$$

3.1 Algebraic approach via creation and annihilation (ladder) operators

• We will find all the energy eigenstates of the SHO by solving $H\psi = E\psi$. Rather than solving this second order differential equation directly, we follow a different approach. The ground state will be found by solving a *first order* differential equation. Then we will create all the excited states by suitably modifying the ground state. This may seem an odd way to proceed. But the method works in many quantum mechanical systems and is justified by its success.

• It is convenient to work in dimensionless variables. For example we may express all energies as multiples of $\frac{1}{2}\hbar\omega$. $\kappa = \sqrt{m\omega/\hbar}$ has dimensions of L^{-1}, so we may define dimensionless position and momentum variables by

$$\xi = \kappa x, \quad \mathtt{p} = \frac{p}{\hbar\kappa} = -i\frac{\partial}{\partial \xi}, \quad \Rightarrow \quad H = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2 = \frac{\hbar\omega}{2}\left(\xi^2 + \mathtt{p}^2\right) \tag{74}$$

$\xi, \mathtt{p}$ are dimensionless. Being the sum of two squares, the hamiltonian is a positive function on phase space. Now, to the extent possible, we wish to factorize the hamiltonian as the product of an operator and its adjoint, which would automatically be a positive operator. To this end, define the dimensionless annihilation (lowering) and creation (raising) operators

$$a = \frac{\xi + i\mathtt{p}}{\sqrt{2}}, \qquad a^\dagger = \frac{\xi - i\mathtt{p}}{\sqrt{2}} \quad \Rightarrow \quad a^\dagger a = \frac{1}{2}\left(\xi^2 + \mathtt{p}^2 + i[\xi, \mathtt{p}]\right) = \frac{1}{2}\left(\xi^2 + \mathtt{p}^2 - 1\right) \tag{75}$$

Here we used $[\xi, \mathtt{p}] = i$. Thus, up to an additive constant, the hamiltonian has been factorized:

$$H = \hbar\omega\left(a^\dagger a + \frac{1}{2}\right) = \hbar\omega\left(a a^\dagger - \frac{1}{2}\right), \qquad a = \frac{1}{\sqrt{2}}\left(\xi + \frac{\partial}{\partial\xi}\right), \quad a^\dagger = \frac{1}{\sqrt{2}}\left(\xi - \frac{\partial}{\partial\xi}\right). \tag{76}$$

- $a, a^\dagger$ do not commute, their commutator is $[a, a^\dagger] = aa^\dagger - a^\dagger a = 1$. In terms of x and p

$$a^\dagger = \frac{1}{\sqrt{2}}\left(\kappa x - \frac{ip}{\hbar\kappa}\right) = \frac{1}{\sqrt{2}}\left(\kappa x - \frac{1}{\kappa}\frac{\partial}{\partial x}\right), \qquad a = \frac{1}{\sqrt{2}}\left(\kappa x + \frac{1}{\kappa}\frac{\partial}{\partial x}\right) \tag{77}$$

- Now suppose ψ is an energy eigenstate $H\psi = E\psi$. The main virtue of the creation and annihilation operators is that they act on ψ to produce energy eigenstates with higher ($E + \hbar\omega$) and lower ($E - \hbar\omega$) energies. Using $[a, a^\dagger] = 1$ we find

$$H\left(a^\dagger\psi\right) = (E + \hbar\omega)\left(a^\dagger\psi\right), \qquad H(a\psi) = (E - \hbar\omega)(a\psi). \tag{78}$$

Now by repeated application of a, we can produce states with negative energy. However, we already saw that $\langle H\rangle \geq 0$. A way around this problem is that by repeated application of a we must eventually reach a state ψ_0 that is annihilated by a: $a\psi_0 = 0$. ψ_0 is the ground state. By repeated application of $a^\dagger$ on ψ_0 we get all the excited states. Repeated application of a on ψ_0 gives the zero vector.

- $a\psi_0(x) = 0$ leads to the first order differential equation $(\xi + \frac{\partial}{\partial\xi})\psi_0(\xi) = 0$ or

$$\xi\psi_0 = -\psi_0' \quad \Rightarrow \quad \log\psi_0 = -\frac{-\xi^2}{2} + c \quad \Rightarrow \psi_0 = Ae^{-\xi^2/2} \quad \Rightarrow \quad \psi_0(x) = Ae^{-\kappa^2 x^2/2}, \quad A = \frac{\sqrt{\kappa}}{\pi^{1/4}}$$

The g.s. energy is $\frac{1}{2}\hbar\omega$. Why? $H\psi_0 = \hbar\omega(a^\dagger a + \frac{1}{2})\psi_0 = \frac{1}{2}\hbar\omega\psi_0$.

- The classical ground state is the static solution of Hamilton's equations, $x(t) = 0, p(t) = 0$, where the particle is eternally at rest at the equilibrium point. The quantum g.s. energy $\frac{1}{2}\hbar\omega$ is not zero, it is called the zero point energy. Classically, the ground state wave function has zero width, it is concentrated at the point $x = 0$. Let us indicate why the ground state wave function cannot be too narrow or too broad. The g.s. energy is the expectation value of the hamiltonian in the normalised ground state. Integrate by parts to show that it can be written as

$$\langle H\rangle = \frac{\hbar^2}{2m}\int_{-\infty}^{\infty} |\psi'(x)|^2\, dx + \frac{1}{2}m\omega^2 \int x^2 |\psi(x)|^2\, dx. \tag{79}$$

- We introduced $\kappa = \sqrt{\frac{m\omega}{\hbar}}$ which is a constant with dimensions of inverse length. κ controls the rate of decay of the position space probability distribution. The larger κ is, the less probable it is to find the particle far from the point of equilibrium $x = 0$. $\kappa^2 = \frac{\sqrt{mk}}{\hbar}$ is large if the force constant or mass is large, both of which prevent the particle from going far from $x = 0$. This is true both quantum mechanically as well as classically. But unlike classically, in QM, the particle in the g.s. has a non-zero probability to be found beyond the classical turning points. The probability for a particle in the ground state to be found beyond the classical turning points x_0 is

$$P(|x| > x_0) = 2 \int_{x_0}^{\infty} |\psi_0(x)|^2 \, dx \tag{80}$$

This can be evaluated using the error function, it is a very small number for macroscopic m, ω and E. Now if $\psi(x)$ is very wide, then the mean potential energy would become large, since large values of x contribute more to the potential energy. On the other hand, if $\psi(x)$ is very narrow, it must also rise and fall very steeply, resulting in large values of $\psi'(x)$, which would make the mean kinetic energy very large. The ground state wave function must strike a balance between these two competing tendencies, the gaussian with width $1/\kappa$ is neither too broad nor too narrow and results in a minimum possible energy expectation value.

- *All* the excited states are obtained as $\psi_n(x) = A_n (a^\dagger)^n \psi_0(x)$, for $n = 1, 2, 3, \ldots$ for appropriate normalization constants shown below to be $A_n = \frac{1}{\sqrt{n!}}$ to ensure $||\psi_n|| = 1$.

- Since $a^\dagger \psi_0$ is an eigenstate with energy $\hbar\omega$ more than that of ψ_0, the energy level obtained by n-fold application of $a^\dagger$ is $E_n = \frac{1}{2}\hbar\omega + n\hbar\omega = \hbar\omega\left(n + \frac{1}{2}\right)$. This is the energy spectrum of the SHO.

- Define the number operator $\hat{N} = a^\dagger a$, then $\hat{H} = \hbar\omega\left(\hat{N} + \frac{1}{2}\right)$. So $\hat{N}\psi_n = n\psi_n$ for $n = 0, 1, 2, \cdots$.

- To find normalization constants A_n so that $||\psi_n|| = 1$ for $n = 1, 2, 3, \cdots$. We first note that though ψ_0 may have norm one, $a^\dagger \psi_0$ need not have norm one, so let us write

$$a^\dagger \psi_n = c_n \psi_{n+1}, \quad \text{and} \quad a\psi_n = d_n \psi_{n-1} \tag{81}$$

To find d_n, we observe that (use the fact that a and $a^\dagger$ are adjoints)

$$1 = \langle \psi_{n-1} | \psi_{n-1} \rangle = \langle \frac{a\psi_n}{d_n} | \frac{a\psi_n}{d_n} \rangle \quad \Rightarrow \quad |d_n|^2 = \langle \psi_n | a^\dagger a \psi_n \rangle = \langle \psi_n | \hat{N} \psi_n \rangle = n ||\psi_n||^2 = n. \tag{82}$$

Thus $d_n = \sqrt{n}$. Similarly we find that $c_n = \sqrt{n+1}$. So

$$a^\dagger \psi_0 = c_0 \psi_1 = \psi_1. \quad a^\dagger \psi_1 = \sqrt{2}\psi_2 \ \Rightarrow \psi_2 = \frac{1}{\sqrt{2!}} a^\dagger a^\dagger \psi_0, \ \ldots \ \Rightarrow \psi_n = \frac{1}{\sqrt{n!}} (a^\dagger)^n \psi_0. \tag{83}$$

- For example, the normalized 1st excited state is ($A_1 = 1$)

$$\psi_1 = A_1 a^\dagger \psi_0 = \frac{A_1 A}{\sqrt{2}} \left(\xi - \frac{\partial}{\partial \xi}\right) e^{-\frac{\xi^2}{2}} = A\sqrt{2}\xi e^{-\frac{\xi^2}{2}} = \sqrt{2}\kappa \frac{\sqrt{\kappa}}{\pi^{1/4}} x e^{-\frac{\kappa^2 x^2}{2}} = \sqrt{\frac{2m\omega}{\hbar}} \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} x e^{-\frac{m\omega x^2}{2\hbar}}$$

In particular, the first excited state wave function is odd, it has one node ($\psi_1(0) = 0$).

- The excited states are $\psi_n = A_n \frac{1}{\sqrt{2^n}} \left(\xi - \frac{\partial}{\partial \xi}\right)^n \psi_0 \propto H_n(\xi) e^{-\xi^2/2}$ for some polynomials $H_n(\xi)$, which turn out to be the Hermite polynomials. It is clear that H_n is even in ξ (parity even) for n even and odd in ξ (parity odd) for n odd since it involves $\left(\xi - \frac{\partial}{\partial \xi}\right)^n$. The particle has zero probability density to be found at the equilibrium point $x = 0$ in all the odd stationary states! In general, the n^{th} excited state has n nodes. Precisely,

$$\psi_n(\xi) = \sqrt{\frac{\kappa}{\sqrt{\pi}}} \frac{1}{\sqrt{2^n n!}} H_n(\xi) e^{-\xi^2/2} \tag{84}$$

The $\sqrt{2^n}$ is because Hermite polynomials are conventionally normalized to have leading coefficient 2^n. The first few Hermite polynomials are

$$H_0 = 1, \quad H_1 = 2\xi, \quad H_2 = 4\xi^2 - 2, \quad H_3 = 8\xi^3 - 12\xi. \tag{85}$$

• Let us define the parity operator $\mathbb{P} f(x) = f(-x)$. It may be seen to be a hermitian operator on $\mathrm{L}^2(\mathbb{R})$

$$\langle f|\mathbb{P}|g\rangle = \int_{-\infty}^{\infty} f^*(x) g(-x)\, dx = \int_{-\infty}^{\infty} f^*(-x) g(x)\, dx = \langle \mathbb{P} f|g\rangle. \tag{86}$$

$\mathbb{P}^2 = I$ so the eigenvalues of parity are ± 1. The eigenfunctions of parity with eigenvalue one are the even functions ($\mathbb{P} f(x) = f(x)$ called even parity functions) and the eigenfunctions with eigenvalue -1 are the odd functions, $\mathbb{P} f(x) = f(-x) = -f(x)$. Since $H = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + \frac{1}{2} m\omega^2 x^2$ is unchanged under $x \to -x$ one can show that $[\mathbb{P}, H] = 0$. Indeed,

$$\mathbb{P}(H\psi) = -\frac{\hbar^2}{2m}\frac{\partial^2 \psi(-x)}{\partial x^2} + \frac{1}{2} m\omega^2 x^2 \psi(-x) = H(\mathbb{P}\psi) \tag{87}$$

So parity is a symmetry, and its expectation value in any state is preserved in time. Moreover $\mathbb{P}$ and H are simultaneously diagonalisable. So energy eigenfunctions can be labelled by the energy as well as parity eigenvalue. The n^{th} state has energy eigenvalue $E_n = \hbar\omega(n + \frac{1}{2})$ and parity eigenvalue $(-1)^n$.

• The energy eigenstates are orthogonal. For $m, n \neq 0$, we calculate $\langle \psi_n|\psi_m\rangle$ in two different ways, by inserting the number operator $N/n = a^\dagger a/n$ which acts as the identity on the n^{th} eigenstate $N\psi_n = n\psi_n$:

$$\langle \psi_n|\psi_m\rangle = \frac{1}{m}\langle \psi_n|a^\dagger a \psi_m\rangle = \frac{1}{m}\langle a\psi_n|a\psi_m\rangle \tag{88}$$

On the other hand,

$$\langle \psi_n|\psi_m\rangle = \frac{1}{n}\langle a^\dagger a\psi_n|\psi_m\rangle = \frac{1}{n}\langle a\psi_n|a\psi_m\rangle \tag{89}$$

To be equal, $m = n$ or if $m \neq n$, $\langle \psi_n|\psi_m\rangle = 0$. Combining with normalization, $\langle \psi_n|\psi_m\rangle = \delta_{mn}$, for $m, n \neq 0$ (they cannot be zero since we divided by them). Now, if one of the states is the ground state, then $\langle 0|n\rangle = \frac{1}{\sqrt{n!}}\langle 0|(a^\dagger)^n|0\rangle = \langle a^n \psi_0|\psi_0\rangle = 0$ since $a\psi_0 = 0$. So all excited states are orthogonal to the ground state. Of course, the g.s. has been normalised to 1, so it follows that $\langle m|n\rangle = \delta_{mn}$ for all $m, n = 0, 1, 2, 3, \ldots$.

• The higher excited states are increasingly oscillatory between the classical turning points. But they decay exponentially to zero outside the classical turning points. As the quantum number n grows, the probability density in the n^{th} excited state begins to approach (in a coarse-grained sense) a classical probability density. By the latter we mean: suppose we have several springs with the same force constant and same energy (large compared to $\hbar\omega$), but let them start oscillating at different times. Then at a common time, we measure the location of each classical oscillator and plot the distribution of positions. We are likely to find fewer oscillators near their points of equilibrium and more near their turning points. This is because the oscillator slows down as it approaches the turning point and speeds up near the equilibrium point. The classical probability distribution is the one to which $|\psi_n(x)|^2$ tends (provided we do not look too close to see the rapid oscillations). The appearance of classical behaviour in the limit of large quantum numbers is one manifestation of Bohr's correspondence principle.

• Find the energy eigenstates and energy levels for a 2d simple harmonic oscillator with $H = \frac{p_x^2}{2m} + \frac{p_y^2}{2m} + \frac{1}{2}m\omega^2 x^2 + \frac{1}{2}m\omega^2 y^2$. Find the degeneracies of the lowest 4 energy levels.

4 Heisenberg uncertainty principle (inequality)

• Given an observable A and a unit norm state $|\psi\rangle$, we have the variance of A in the state ψ (or the square of the standard deviation or simply the square of the uncertainty of A)

$$(\Delta A)^2 = \langle\psi|(A - \langle A\rangle)^2|\psi\rangle = \langle A^2\rangle - \langle A\rangle^2 \tag{90}$$

The uncertainty in A measures the spread/width of the distribution of possible measured values of A in the state $|\psi\rangle$. It depends *both* on A and $|\psi\rangle$. If ψ is an eigenstate of A with eigenvalue a, then the uncertainty of A is zero. We say that A takes a definite value a in an eigenstate. We say that A has quantum fluctuations in the state ψ if $\langle A^2\rangle \neq \langle A\rangle^2$.

• Suppose ψ is a unit norm state, then the Heisenberg uncertainty inequality is $\Delta x\Delta p \geq \frac{1}{2}\hbar$. It says that if you prepare a large number of copies of a system in the same state ψ, and make measurements of position on half of them and momentum on the other half, the product of standard deviations in the measurements of position and momentum is bounded below by $\hbar/2$.

• An extreme case: if ψ is a position eigenstate $|x_0\rangle$. In such a state, the uncertainty in x is zero, a measurement of position always results in the value x_0. However, the uncertainty in momentum is infinite in a position eigenstate, all values of momentum are equally likely.

• The ground state ψ_0 of the SHO is a minimum uncertainty state. $\Delta x\Delta p = \hbar/2$ in this state. Check this statement.

• To show this we define an uncertainty functional U in a unit norm state ψ for a pair of observables A, B with $[A,B] = iC$. Later we will specialize to $A = x, B = p, C = \hbar I$.

$$U(\psi) = (\Delta A)^2(\Delta B)^2 = \langle\psi|(A-\bar{A})^2|\psi\rangle\langle\psi|(B-\bar{B})^2|\psi\rangle = \langle\alpha|\alpha\rangle\langle\beta|\beta\rangle \tag{91}$$

where $|\alpha\rangle = (A-\bar{A})|\psi\rangle \equiv \delta A|\psi\rangle$ and $|\beta\rangle = (B-\bar{B})|\psi\rangle = \delta B|\psi\rangle$. By the Cauchy-Schwarz inequality,

$$U \geq |\langle\alpha|\beta\rangle|^2 = |\langle\psi|\ \delta A\ \delta B\ |\psi\rangle|^2 \tag{92}$$

We bring in the commutator and the anticommutator via

$$\delta A\delta B = \frac{1}{2}[\delta A, \delta B] + \frac{1}{2}\{\ \delta A, \delta B\ \} = \frac{1}{2}iC + \frac{1}{2}\{\ \delta A, \delta B\ \}. \tag{93}$$

Now C is hermitian as is $\{\ \delta A, \delta B\ \}$. It follows that $\frac{1}{2}\langle iC\rangle$ is purely imaginary and $\frac{1}{2}\langle\{\ \delta A, \delta B\ \}\rangle$ is real. So the absolute square of the sum is just the sum of the squares of the imaginary and real parts:

$$U \geq |\frac{i}{2}\langle C\rangle + \frac{1}{2}\langle\{\ \delta A, \delta B\ \}\rangle|^2 = \frac{1}{4}\langle C\rangle_\psi^2 + \frac{1}{4}\langle\psi|\{\delta A, \delta B\}|\psi\rangle^2. \tag{94}$$

The second term is ≥ 0. So we get

$$(\Delta A)^2(\Delta B)^2 = U \geq \frac{1}{4}\langle C\rangle_\psi^2. \tag{95}$$

Specializing to $A = x, B = p, C = \hbar I$ we get the Heisenberg uncertainty inequality $\Delta x\Delta p \geq \hbar/2$.

5 Angular momentum

5.1 Basic properties of angular momentum

- Classically angular momentum $\mathbf{L} = \mathbf{r} \times \mathbf{p}$. Its cartesian components are

$$L_x = yp_z - zp_y, \quad L_y = zp_x - xp_z, \quad L_z = zp_y - yp_x \qquad \text{Cyclic:} \quad x \to y \to z \to x \tag{96}$$

We could also write them in terms of the Levi-Civita tensor ϵ_{ijk} which is anti-symmetric under interchange of any pair of indices.

$$L_i = \epsilon_{ijk} r_j p_k, \quad \epsilon_{123} = +1, \epsilon_{132} = -1, \ldots . \tag{97}$$

- We saw that each component of angular momentum generates a symmetry and is a conserved quantity (in the sense of expectation values), for a particle in a central potential $V(r)$.

- In QM, the angular momentum operators are

$$L_x = \hat{y}\hat{p}_z - \hat{z}\hat{p}_y = -i\hbar\left(y\frac{\partial}{\partial z} - z\frac{\partial}{\partial y}\right), \cdots \tag{98}$$

Note that y and p_z commute so it does not matter whether we write yp_z or $p_z y$ etc. Angular momentum has the dimensions of $\hbar$. As a vector $\hat{\mathbf{p}} = -i\hbar(\partial_x, \partial_y, \partial_z) = -i\hbar\nabla$. So $L = -i\hbar\mathbf{r}\times\nabla$. Check that the angular momentum operators L_x, L_y, L_z are hermitian: $L_x^\dagger = L_x$ etc. This is because y and p_z, for instance, are hermitian and commute. It follows that $\vec{L}$ is also hermitian.

- The square of total angular momentum L^2 is defined as the square of the length of the vector $\vec{L} = (L_x, L_y, L_z)$. $L^2 = L_x^2 + L_y^2 + L_z^2$. L^2 is also hermitian.

- Angular momentum commutation relations. Using the commutators of x, y, z, p_x, p_y, p_z we show

$$[L_x, L_y] = i\hbar L_z, \quad [L_y, L_z] = i\hbar L_x, \quad [L_z, L_x] = i\hbar L_y \quad \text{cyclic.} \tag{99}$$

The components of angular momentum aren't simultaneously diagonalizable:

$$(\Delta L_x)^2(\Delta L_y)^2 \geq -\frac{1}{4}\langle [L_x, L_y]\rangle^2 = \frac{\hbar^2}{4}\langle L_z\rangle^2 \quad \Rightarrow \quad \Delta L_x \Delta L_y \geq \frac{\hbar}{2}|\langle L_z\rangle|. \tag{100}$$

L_x and L_y can't take definite values in the same state, except if $\langle L_z\rangle = 0$ in that state.

- However, check that $[L^2, \vec{L}] = 0$. In other words, each component of angular momentum L_x, L_y, L_z commutes with the square of the total angular momentum. E.g., L^2 and L_z may be simultaneously diagonalized, but not L^2, L_z and L_x.

- We may also express angular momentum in spherical polar coordinates

$$z = r\cos\theta, \quad x = r\sin\theta\cos\phi, \quad y = r\sin\theta\sin\phi. \tag{101}$$

Calculate derivatives by the chain rule to get

$$L_z = -i\hbar\frac{\partial}{\partial\phi}, \quad L_x = i\hbar\left(\sin\phi\frac{\partial}{\partial\theta} + \cot\theta\cos\phi\frac{\partial}{\partial\phi}\right), \quad L_y = i\hbar\left(-\cos\phi\frac{\partial}{\partial\theta} + \cot\theta\sin\phi\frac{\partial}{\partial\phi}\right) \tag{102}$$

For example, $x_\phi = -y, y_\phi = x, z_\phi = 0$ where subscripts denote partial derivatives. So

$$-i\hbar\partial_\phi = -i\hbar\left(\frac{\partial x}{\partial\phi}\frac{\partial}{\partial x} + \frac{\partial y}{\partial\phi}\frac{\partial}{\partial y} + \frac{\partial z}{\partial\phi}\frac{\partial}{\partial z}\right) = -i\hbar\left(-y\partial_x + x\partial_y\right) = L_z. \tag{103}$$

- We may also express L^2 in spherical coordinates

$$L^2\psi = -\hbar^2 \left[\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\psi}{\partial\theta}\right) + \frac{1}{\sin^2\theta}\frac{\partial^2\psi}{\partial\phi^2}\right]. \tag{104}$$

- It is easy to solve the eigenvalue problem for $L_z = -i\hbar\partial_\phi$. L_z has the same dimensions as $\hbar$. So let us write the eigenvalue as $\hbar m$ where m is dimensionless: $L_z f = \hbar m f$. The eigenfunctions are $f(\phi) = Ae^{im\phi}$. What are the eigenvalues m? $\phi = 0$ and $\phi = 2\pi$ represent the same physical point. We require the wave function to be single-valued, i.e. $f(0) = f(2\pi)$. This forces m to be an integer $0, \pm 1, \pm 2, \ldots$. So the eigenvalues of L_z are integers in units of $\hbar$.

5.2 Eigenvalues of L^2 and L_z by ladder operators

- We seek simultaneous eigenstates ψ of L^2 and L_z with eigenvalues $\lambda\hbar^2$ and $\hbar m$

$$L^2\psi = \hbar^2\lambda\psi, \quad L_z\psi = \hbar m\psi \tag{105}$$

To find the allowed values of λ, m we use the ladder operator method. Define $L_\pm = L_x \pm iL_y$. The virtue of $L_\pm$ is that if ψ is one such simultaneous eigenstate, then so are $L_\pm\psi$. Indeed, we check using the commutation relations $[L^2, L_\pm] = 0$ and $[L_z, L_\pm] = \pm\hbar L_\pm$ that

$$L^2\left(L_\pm\psi\right) = \hbar^2\lambda(L_\pm\psi), \quad L_z\left(L_\pm\psi\right) = \hbar(m\pm 1)\left(L_\pm\psi\right). \tag{106}$$

So $L_\pm$ raises/lowers the eigenvalue of L_z by $\hbar$ while leaving the eigenvalue of L^2 unchanged. However, by positivity of L_x^2, L_y^2, L_z^2 and $L^2 = \vec{L}\cdot\vec{L}$ it follows that the eigenvalues of L_z for *fixed* λ cannot get too big or too small. So starting with a common eigenfunction ψ and repeatedly applying L_+, there must be a state (ψ_h) with highest eigenvalue of L_z (say, $\hbar l$), which is annihilated by L_+

$$L_z\psi_h = \hbar l\psi_h, \quad L_+\psi_h = 0. \tag{107}$$

But ψ_h must also be an eigenfunction of L^2. To see how L^2 acts on ψ_h we express it in terms of $L_\pm$

$$L^2 = L_-L_+ + L_z^2 + \hbar L_z \quad\Rightarrow\quad L^2\psi_h = \left(\hbar^2 l^2 + \hbar^2 l\right)\psi_h = \hbar^2 l(l+1)\psi_h \tag{108}$$

So if $\hbar l$ is the largest eigenvalue of L_z, then the corresponding eigenvalue of L^2 must be $\hbar^2 l(l+1)$. But what are the allowed values of l? To find out, we note that there is also a state with lowest eigenvalue of L_z which is annihilated by L_-

$$L_z\psi_{low} = \hbar l'\psi_{low}, \quad L_-\psi_{low} = 0, \qquad l' \leq l \tag{109}$$

But ψ_{low} must also be an eigenfunction of L^2. Using

$$L^2 = L_+L_- + L_z^2 - \hbar L_z \quad\Rightarrow\quad L^2\psi_{low} = \left(\hbar^2 l'^2 - \hbar^2 l'\right)\psi_{low} = \hbar^2 l'(l'-1)\psi_{low} \tag{110}$$

However, raising and lowering does not change the eigenvalue of L^2, so

$$l(l+1) = l'(l'-1), \qquad l \geq l'. \tag{111}$$

The solutions are $l' = -l$ and $l' = l+1$. The latter solution is disallowed since it would imply $l' > l$. Thus $l' = -l$. So for fixed λ, i.e., fixed l, the eigenvalues of L_z go in integer steps from l to $-l$. But this is possible only if $l - (-l)$ is itself an integer, i.e., if l is half an integer

$$l = 0, \pm\frac{1}{2}, \pm 1, \pm\frac{3}{2}, \pm 2 \cdots. \tag{112}$$

And for each such value of l, the possible eigenvalues of L_z are $\hbar(-l,-l+1,\cdots,l-l,l)$. However, there is a small problem with the non-integer values of l. We already saw in the last section that the eigenfunctions of L_z are $Ae^{im\phi}$ with eigenvalues $\hbar m$ where m is an integer. If m was half an odd integer, then $\Phi(\phi) \neq \Phi(\phi+2\pi)$ would not be single-valued. Only the integer values $l = 0,1,2,\ldots$ lead to simultaneous eigenfunctions of L^2 and L_z. To summarize, the spectrum of L^2 consists of the numbers $\hbar^2 l(l+1)$ where $l = 0,1,2,\ldots$. For each such l, the spectrum of L_z consists of the numbers $\hbar m$ where m is an integer satisfying $-l \leq m \leq l$. So the degeneracy of each eigenvalue of L^2 is $2l+1$. The simultaneous eigenstates are denoted $|lm\rangle$, they are orthonormal $\langle lm|l'm'\rangle = \delta_{mm'}\delta_{ll'}$.

• The Spherical harmonics $Y_{lm}(\theta,\phi) = \langle\theta\phi|lm\rangle$ are the simultaneous eigenfunctions of L^2 and L_z expressed in the spherical polar coordinate basis. We will work out some of them later.

5.3 Rigid Body

• A molecule can be crudely modelled as a rigid body if we are interested in its rotational spectrum. Consider an isotropic rigid body, (i.e., one whose principal moments of inertia are equal) e.g., a spherical ball free to rotate about its center, which is held fixed, or a cube shaped molecule or a tetrahedral molecule. Its rotational kinetic energy is $H = \frac{L^2}{2I}$ where I is its moment of inertia about *any* axis passing through its center. Classically, its rotational energy is any positive number, depending on its angular velocity $\omega = L/I$. For an isotropic rigid body, in quantum mechanics, H, L^2 and L_z commute and are simultaneously diagonal in the basis of spherical harmonics. The spectrum of energies is discrete, $E_{lm} = \frac{l(l+1)\hbar^2}{2I}$ where the angular momentum quantum number $l = 0,1,2,\ldots$. The corresponding eigenfunctions are the spherical harmonics $Y_{lm}(\theta,\phi)$. The energies are independent of the magnetic quantum number m. This is because of spherical symmetry. All $2l+1$ states Y_{lm}, for $m = -l,\ldots,l$, irrespective of their z-component of angular momentum ($L_z Y_{lm} = \hbar m Y_{lm}$) are degenerate in energy. This is due to the spherical symmetry of the hamiltonian. L_x, L_y and L_z are symmetries of H, they commute with it. The independence of E on m means the rotational energy of the body does not depend on its orientation (i.e., independent of the projection of the angular momentum on the z-axis). In general, symmetries imply degeneracies. And when symmetries are broken, the degeneracy is lost.

• More generally, the rotational kinetic energy of a rigid body free to rotate about a point is $H = \frac{L_1^2}{2I_1} + \frac{L_2^2}{2I_2} + \frac{L_3^2}{2I_3}$ where I_1, I_2, I_3 are the principal moments of inertia about the three principal axes of inertia of the body. For an isotropic body, $I_1 = I_2 = I_3$.

• The next simplest case is a top, a rigid body with an axis of symmetry, which we can take as the third principal axis. Then $I_1 = I_2 = I$ and the hamiltonian becomes

$$H = \frac{L_1^2 + L_2^2}{2I} + \frac{L_3^2}{2I_3} = \frac{L^2 - L_3^2}{2I} + \frac{L_3^2}{2I_3}. \tag{113}$$

Written this way, we see that $[H, L^2] = 0$ and $[H, L_3] = 0$, so the hamiltonian is again diagonal in the basis of spherical harmonics. The energy spectrum is

$$E_{lm} = \frac{\hbar^2 l(l+1) - \hbar^2 m^2}{2I} + \frac{\hbar^2 m^2}{2I_3}. \tag{114}$$

Due to lack of isotropy, energy levels now depend on the magnetic quantum number m. In a sense, the energy depends on the orientation of the rigid body. The $2l+1$-fold degeneracy of energy in the m quantum number is lost. We still have a 2-fold $m \to -m$ degeneracy in the energy spectrum, corresponding to the symmetry under reflection $z \to -z$ in the x-y plane. The isotropic case is obtained by putting $I_3 = I$.

5.4 Coordinate representation of spherical harmonics for $l = 1$

• In the preceding section we did not require the explicit functional forms of the spherical harmonics $Y_{lm}(\theta, \phi)$. But suppose we wanted to know the angular regions where an electron in the state $|lm\rangle$ is likely to be found. Then we need to know $Y_{lm}(\theta, \phi) = \langle \theta, \phi | lm \rangle$. As an illustration, we focus on p-wave states, i.e. those with $l = 1$, there are three linearly independent such states. Let us obtain Y_{1m} using the algebraic method of raising and lowering operators. We already know the ϕ dependence of Y_{lm}. As it is an eigenfunction of L_z, $Y_{lm}(\theta, \phi) = e^{im\phi} P_{lm}(\theta)$. $P_{lm}(\theta)$ are called the associated Legendre functions since they satisfy the associated Legendre differential equation. But we can find them using the raising and lowering operators. We begin by finding the state with top-most value of m, Y_{ll}. By definition, $L_+ Y_{ll} = 0$, so if we denote $Y_{ll}(\theta, \phi) = \psi(\theta, \phi)$ whose ϕ dependence is known, then

$$L_+ \psi = \hbar e^{i\phi} \left(\partial_\theta + i \cot\theta \partial_\phi \right) \psi = 0 \quad \Rightarrow \quad \psi_\theta + i \cot\theta \psi_\phi = 0 \quad \Rightarrow \quad \psi_\theta - l \cot\theta \psi = 0. \quad (115)$$

We find $\psi = N_{ll} e^{il\phi} \sin^l \theta$ for a normalization constant N_{ll}. The normalization constant is fixed by $\int Y_{lm}^*(\theta, \phi) Y_{lm}(\theta, \phi) \sin\theta d\theta \, d\phi = 1$.

• In particular, $Y_{11} \propto e^{i\phi} \sin\theta$. The angular probability distribution $|Y_{11}|^2 \propto \sin^2\theta$ is peaked around the equator $\theta = \pi/2$ while the particle is less likely to be found near the poles. To get the next state, we apply $L_- = L_+^\dagger$. Do this!

$$Y_{10} \propto L_- e^{i\phi} \sin\theta = -\hbar e^{-i\phi} \left(\partial_\theta - i \cot\theta \partial_\phi \right) e^{i\phi} \sin\theta = -2\hbar \cos\theta \quad \Rightarrow \quad Y_{10} = N_{10} \cos\theta. \quad (116)$$

Thus $|Y_{10}|^2 \propto \cos^2\theta$, and the particle is more likely to be found near the poles in this case. We also notice that $Y_{11} \propto e^{i\phi} \sin\theta$ and $Y_{10} \propto \cos\theta$ are orthogonal, as we would expect of eigenstates of L_z with distinct eigenvalues.

• To get $Y_{1,-1}$ we can either apply L_- to Y_{10} or find it by solving $L_- Y_{1,-1} = 0$. Check that both give the same answer!

5.5 Visualization of angular momentum and location in states Y_{lm}

• We observe that the largest eigenvalue of L_z among states with square of total angular momentum $\hbar^2 l(l+1) > 0$ is $\hbar l$. But $l < \sqrt{l(l+1)}$ for $l > 0$. This means in a state of definite L^2 and L_z, measured values of L_z can never equal the (non-zero) total angular momentum. This is different from the classical situation where all the angular momentum can be pointing in (say) the z direction. Quantum mechanically there cannot be any simultaneous eigenstate where the eigenvalues of L_x, L_y are zero and that of L_z and L equal (and non-zero). Such a state would violate $\Delta L_x \Delta L_y \geq \frac{\hbar}{2} |\langle L_z \rangle|$.

• Classically, suppose we are in a stationary state of angular momentum, i.e., one where $\vec{L}$ points in a fixed direction with fixed magnitude over time. For example, we can be in a classical state where $L_z = 10^5 \hbar, L_y = 0, L_x = 0$. We can visualize this in terms of a rigid body that is rotating with constant angular speed about an axis pointing along $\hat{z}$. Quantum mechanically, the stationary states may be taken as simultaneous eigenstates $\psi = Y_{lm}$ of L^2 and L_z. In such a state we can imagine the angular momentum vector as having length equal to the square-root of the eigenvalue of L^2, i.e., $\sqrt{\hbar^2 l(l+1)}$. Moreover, this vector has projection on the z-axis of $\hbar m$. So we know its angle with the z-axis. But that is as much as we can say. We cannot unambiguously specify its projections on x or y axes since L_x and L_y do not have definite values in this state. To visualize the 'angular momentum vector' in this state, we can think of it as the cone of vectors from the origin whose lengths are $\sqrt{\hbar^2 l(l+1)}$ and projection $\hbar m$ on the z-axis. This hand-waving visualization can be a useful aid to memory. There are as many vectors on the cone with x-component of angular momentum equal to $\hbar m_x$ as $-\hbar m_x$. So by symmetry we would expect the expectation value of L_x in the state Y_{lm} to vanish, as it does. It is important to realize that this cone *does not* tell us where the particle is likely to be found, it only gives some crude information on the likely values that may be obtained upon a measurement of various components of angular momentum.

- To visualize where the particle is likely to be found if its angular position is measured in state Y_{lm}, we must plot the probability density $|Y_{lm}(\theta,\phi)|^2$ on the surface of a unit sphere in polar coordinates $r = 1, \theta, \phi$. For example, in the S-wave state $Y_{00} = \frac{1}{\sqrt{4\pi}}$, this probability density is constant, indicating that the particle is equally likely to be found at all angular locations. More generally, $Y_{lm}(\theta,\phi) \propto e^{im\phi} P_{lm}(\cos\theta)$. So $|Y_{lm}|^2 \propto |P_{lm}(\theta)|^2$. Thus the angular probability distribution is azimuthally symmetric (independent of ϕ). For example, $|Y_{10}|^2 \propto \cos^2\theta$. So in this state, the particle is more likely to be found near the north pole ($\theta = 0$) or south pole ($\theta = \pi$), than along the equator $\theta = \pi/2$. Polar plots of $P_{lm}(\theta)$ are given in many text books.

5.6 Matrix elements of $L_\pm$ in the $|lm\rangle$ basis

- Let $|l,m\rangle$ be the normalized simultaneous eigenstates of L^2 and L_z, $\langle lm|l'm'\rangle = \delta_{ll'}\delta_{mm'}$.
- We already know that

$$\begin{aligned} L^2|lm\rangle &= \hbar^2 l(l+1)|lm\rangle \quad \text{and} \quad L_z|lm\rangle = \hbar m|lm\rangle \\ \text{So} \quad \langle lm|L^2|l'm'\rangle &= \hbar^2 l(l+1)\delta_{ll'}\delta_{mm'} \quad \text{and} \quad \langle lm|L_z|l'm'\rangle = \hbar m\delta_{ll'}\delta_{mm'}. \end{aligned} \tag{117}$$

- Moreover, $L_\pm = L_x \pm iL_y$ raise and lower the value of m by one, so for some dimensionless constants $C^\pm_{lm}$ we must have

$$L_\pm|lm\rangle = \hbar C^\pm_{lm}|l,m\pm 1\rangle. \tag{118}$$

- To find C^-, we use the identity $L^2 = L_+L_- + L_z^2 - \hbar L_z$ to evaluate the diagonal matrix elements $\langle lm|L^2|lm\rangle$ in two different ways. We get

$$\begin{aligned} \hbar^2 l(l+1) &= \langle L_- lm|L_- lm\rangle + \langle lm|L_z^2 - \hbar L_z|lm\rangle = \hbar^2|C^-_{l,m}|^2 + \hbar^2 m^2 - \hbar^2 m \\ \Rightarrow &\quad |C^-_{lm}|^2 = l(l+1) - m(m-1) \quad \Rightarrow \quad |C^-_{lm}|^2 = l(l+1) - m(m-1) \end{aligned} \tag{119}$$

Similarly[2], using $L^2 = L_-L_+ + L_z^2 + \hbar L_z$ we get $|C^+_{lm}|^2 = l(l+1) - m(m+1)$. Thus for some phases $e^{i\phi_{lm}}$, we have

$$C^-_{lm} = \sqrt{l(l+1) - m(m-1)}\,e^{-i\phi_{lm}} \quad \text{and} \quad C^+_{lm} = \sqrt{l(l+1) - m(m+1)}\,e^{i\phi_{lm}}. \tag{121}$$

---

[2] Alternatively, we may use hermiticity $L_+^\dagger = L_-$, to get $C^-_{lm} = (C^+_{l,m-1})^*$. To see this, we note that

$$\langle l'm'|L_-|lm\rangle = C^-_{lm}\delta_{ll'}\delta_{m',m-1} = \langle L_+ l'm'|lm\rangle = \langle lm|L_+|l'm'\rangle^* = C^{+*}_{l'm'}\delta_{ll'}\delta_{m,m'+1} = C^{+*}_{l,m-1}\delta_{ll'}\delta_{m',m-1} \tag{120}$$

Since these must be equal for all l, l', m, m' we have $C^-_{lm} = (C^+_{l,m-1})^*$.

It is possible to absorb these phases into the wave functions and take $C^{\pm}$ to be real.

$$C^{\pm}_{lm} = \sqrt{l(l+1) - m(m \pm 1)} \tag{122}$$

Thus the matrix elements of $L_{\pm}$ are

$$\langle l'm'|L_{\pm}|lm\rangle = \hbar\sqrt{l(l+1) - m(m \pm 1)}\ \delta_{l'l}\delta_{m',m\pm 1}. \tag{123}$$

Using these we can easily get the matrix elements of L_x and L_y.

5.6.1 E.g.: Matrix representation of angular momenta for $l = 1$

- Let us illustrate with the $l = 1$ subspace which is $2l + 1 = 3$ dimensional and is spanned by the orthonormal spherical harmonics $Y_{1m}(\theta, \phi)$ for $m = 1, 0, -1$ which we denote $|m\rangle$ and represent by the column vectors

$$|1\rangle = \begin{pmatrix}1\\0\\0\end{pmatrix}, \quad |0\rangle = \begin{pmatrix}0\\1\\0\end{pmatrix}, \quad |-1\rangle = \begin{pmatrix}0\\0\\1\end{pmatrix}. \tag{124}$$

In this basis $L_z|m\rangle = \hbar m|m\rangle$ and $L^2|m\rangle = l(l+1)\hbar^2|m\rangle$ are represented by the diagonal matrices

$$L_z = \hbar\begin{pmatrix}1 & 0 & 0\\0 & 0 & 0\\0 & 0 & -1\end{pmatrix} \quad \text{and} \quad L^2 = 2\hbar^2\begin{pmatrix}1 & 0 & 0\\0 & 1 & 0\\0 & 0 & 1\end{pmatrix}. \tag{125}$$

$L_{\pm}$ have the action $L_{\pm}|m\rangle = \sqrt{2 - m(m \pm 1)}\ \hbar\ |m \pm 1\rangle$ and are adjoints of each other, so

$$\begin{pmatrix}L_+|1\rangle = 0\\ L_+|0\rangle = \sqrt{2}\hbar|1\rangle\\ L_+|-1\rangle = \sqrt{2}\hbar|0\rangle\end{pmatrix} \Rightarrow L_+ = \sqrt{2}\hbar\begin{pmatrix}0 & 1 & 0\\0 & 0 & 1\\0 & 0 & 0\end{pmatrix} \quad \text{and} \quad L_- = \sqrt{2}\hbar\begin{pmatrix}0 & 0 & 0\\1 & 0 & 0\\0 & 1 & 0\end{pmatrix}. \tag{126}$$

Notice that $L_{\pm}$ are strictly upper/lower triangular and not hermitian. They are not observables. But using them we find the matrices for L_x and L_y

$$L_x = \frac{L_+ + L_-}{2} = \frac{\hbar}{\sqrt{2}}\begin{pmatrix}0 & 1 & 0\\1 & 0 & 1\\0 & 1 & 0\end{pmatrix}, \quad L_y = \frac{L_+ - L_-}{2i} = \frac{\hbar}{\sqrt{2}i}\begin{pmatrix}0 & 1 & 0\\-1 & 0 & 1\\0 & -1 & 0\end{pmatrix}. \tag{127}$$

- The eigenvalues of L_z are $\hbar, 0, -\hbar$. Since there is nothing special about the z-direction we should expect the eigenvalues of L_y and L_z to also be the same. Check if this is so!
- However, L_x is not diagonal in the eigenbasis of L_z. Suppose an atom in the $l = 1$ state is in the $m = 0$ eigenstate $|10\rangle$ or $Y_{10}(\theta, \phi)$ of L_z. A measurement of L_x is made. What are the probabilities of various outcomes of the measurement?
- Measurement of L_x can result in any one of the eigenvalues $\hbar m_x = \hbar, 0, -\hbar$, after which the state collapses to the corresponding eigenstate X_+, X_0, X_-, which in this case are (show it!)

$$X_+ = \frac{1}{2}\begin{pmatrix}1\\\sqrt{2}\\1\end{pmatrix}, \quad X_0 = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\0\\-1\end{pmatrix}, \quad X_- = \frac{1}{2}\begin{pmatrix}1\\-\sqrt{2}\\1\end{pmatrix}. \tag{128}$$

The measurement postulate further tells us that the probability of each outcome is the absolute square of the projection of the initial state $|0\rangle = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$ on the final eigenstate. We find

$$\text{prob}(l_x = 0) = |\langle X_0 | 0 \rangle|^2 = 0, \quad \text{prob}(l_x = \hbar) = |\langle X_+ | 0 \rangle|^2 = \frac{1}{2}, \quad \text{prob}(l_x = -\hbar) = |\langle X_- | 0 \rangle|^2 = \frac{1}{2}. \tag{129}$$

Interestingly, this means measurement of L_x for a particle in the state $L_z = 0$ cannot result in the value zero while the values $\pm\hbar$ occur with equal probability. So if a beam of particles in the $l = 1, m = 0\rangle$ state enters a Stern-Gerlach apparatus with an inhomogeneous magnetic field component B_x that couples to the L_x component of angular momentum via the magnetic moment, then two beams will emerge, corresponding to the $l_x = \pm\hbar$ values. Different values of L_x feel different magnetic dipole forces.

6 Spin angular momentum

- Spin was initially discovered (by the work of many physicists, culminating in the work of Goudsmit and Uhlenbeck) through a detailed study of atomic spectral lines. Certain forbidden/additional spectral lines were seen (e.g. where one line was expected, two closely spaced lines were seen). To account for these, a new 'spin' quantum number $m_s = \pm\frac{1}{2}$ was introduced, in addition to the n, l, m quantum numbers used to label the energy levels of the hydrogen atom. It is interesting to read the translation of a lecture by Goudsmit on the history of the discovery of electron spin, see `http://www.lorentz.leidenuniv.nl/history/spin/goudsmit.html`

- Subsequently, a theory of spin was developed (among others by Pauli), wherein it is treated as another type of angular momentum. The initial analogy was with the spinning of the Earth about its axis, while orbital angular momentum was compared with the motion of the center of mass of the Earth around the sun. However, no such mechanical model for the spin of the electron has been successful. The electron is point-like to current experimental accuracy and even extended particles like the proton (which are not point-like), have intrinsic spin which has not been explained by mechanical analogies.

- In classical E & M we learn that a charged particle going round in a loop produces a current loop which has a magnetic dipole moment. The magnetic moment is proportional to the orbital angular momentum of the particle. Thus we expect magnetic moments to be proportional to angular momentum.

- Quantum mechanical spin is regarded as an 'intrinsic spin' represented by a vector observable $\vec{S} = (S_x, S_y, S_z)$. Intrinsic means not of '$\mathbf{r} \times \mathbf{p}$' origin, i.e., not having to do with movement in three dimensional space. In fact, a massive particle at rest can have a magnetic moment! This magnetic moment must necessarily be of non-orbital origin (since $\vec{p} = 0$ for the particle at rest). Such a magnetic moment is ascribed to its spin. The spin observables are postulated to satisfy the same commutation relations as angular momentum

$$[S_i, S_j] = i\hbar\epsilon_{ijk}S_k \tag{130}$$

Unlike L_i, S_i are not expressible in terms of position and momentum, spin is a new degree of freedom.

• To describe spin in the quantum theory, we need to represent S_i as hermitian operators acting on some Hilbert space of spin states. In fact, we have already learned how to do this in the context of orbital angular momentum. $S^2 = S_x^2 + S_y^2 + S_z^2$ and S_z are simultaneously diagonalizable and and their common eigenbasis is denoted $|s, m\rangle$.

$$S^2|sm\rangle = s(s+1)\hbar^2|sm\rangle, \qquad S_z|sm\rangle = \hbar m|sm\rangle. \tag{131}$$

$S_\pm|sm\rangle = (S_x \pm iS_y)|sm\rangle$ are found to be eigenvectors of S^2 with the same eigenvalue $s(s+1)\hbar^2$, but eigenvectors of S_z with eigenvalue raised/lowered by $\hbar$, $S_\pm|sm\rangle \propto |s, m\pm 1\rangle$. The largest and smallest possible values of m must be s and $-s$ and since S_+ raises m in steps of 1, $2s = N$ for some non-negative integer $N = 0, 1, 2, \ldots$. Thus

$$s = 0, \frac{1}{2}, 1, \frac{3}{2}, 2, \frac{5}{2}, \ldots \quad \text{and} \quad m = -s, -s+1, \ldots, s-1, s. \tag{132}$$

However, unlike for orbital angular momentum where $L_z = -i\hbar\frac{\partial}{\partial\phi}$, S_z is not a differential operator in the coordinates of the particle. There is no restriction on the allowed values of m being integral rather than half integral. Indeed, in nature, s is found to take on both integer and half odd-integer values. s is referred to as the spin of the particle. Unlike the orbital angular momentum quantum number l, a given type of particle (say electron) has a fixed value of spin s, irrespective of its state of motion or location.

• Electrons, protons, neutrons, muons, neutrinos, quarks all have spin $s = \frac{1}{2}$. Mesons like pions and kaons have spin zero, as does the proposed Higgs particle. Rho mesons, photons, gluons and the weak gauge bosons $W^\pm, Z^0$ have spin 1, Delta baryons have spin 3/2, gravitons would have spin two. Atoms as a whole can also have non-zero spin arising collectively from the spins of the electrons protons and neutrons. There are particles with higher spin as well, but we will focus on spin $s = \frac{1}{2}$, which is relevant to electrons.

• For each value of spin s, we need to find a representation of the spin observables S_x, S_y, S_z as hermitian operators on some Hilbert space. The relevant Hilbert spaces are $2s+1$ dimensional. This representation may be built up from the lowest state $|s, -s\rangle$ by repeatedly applying S_+.

6.1 Spin half: Pauli spin matrices and spinor wave functions

Here the Hilbert space of spin states is $2s+1 = 2$ dimensional and has basis $|\frac{1}{2}, \frac{1}{2}\rangle$ and $|\frac{1}{2}, -\frac{1}{2}\rangle$ which are also denoted $|\uparrow\rangle$ and $|\downarrow\rangle$. The general spin state of a spin half particle is $a|\uparrow\rangle + b|\downarrow\rangle$. In this basis,

$$|\uparrow\rangle = \begin{pmatrix}1\\0\end{pmatrix}, \qquad |\downarrow\rangle = \begin{pmatrix}0\\1\end{pmatrix} \quad \Rightarrow \quad S^2 = \frac{3\hbar^2}{4}\begin{pmatrix}1 & 0\\0 & 1\end{pmatrix}, \quad S_z = \frac{\hbar}{2}\begin{pmatrix}1 & 0\\0 & -1\end{pmatrix} \tag{133}$$

To find the matrix representation of S_x and S_y we recall that

$$S_\pm|sm\rangle = \hbar\sqrt{s(s+1) - m(m\pm 1)}\;|s, m\pm 1\rangle. \tag{134}$$

Thus $S_+|\uparrow\rangle = 0, S_+|\downarrow\rangle = \hbar|\uparrow\rangle$ and

$$S_+ = \hbar\begin{pmatrix}0 & 1\\0 & 0\end{pmatrix}, S_- = \hbar\begin{pmatrix}0 & 0\\1 & 0\end{pmatrix} \quad \Rightarrow \quad S_x = \frac{S_+ + S_-}{2} = \frac{\hbar}{2}\begin{pmatrix}0 & 1\\1 & 0\end{pmatrix}, \quad \text{and} \quad S_y = \frac{S_+ - S_-}{2i} = \frac{\hbar}{2}\begin{pmatrix}0 & -i\\i & 0\end{pmatrix}.$$

The traceless hermitian matrices $\sigma_i = \frac{2}{\hbar} S_i$ are called Pauli matrices.

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{135}$$

Check that they satisfy

$$[\sigma_i, \sigma_j] = 2\sqrt{-1}\epsilon_{ijk}\sigma_k, \quad \text{and} \quad \sigma_i\sigma_j = \delta_{ij} + i\epsilon_{ijk}\sigma_k. \tag{136}$$

If S_z is measured in the above general state $a|\uparrow\rangle + b|\downarrow\rangle$ then the possible values obtained are $\pm\frac{1}{2}\hbar$ with probabilities given by the absolute squares of the projections of the normalized initial state with the corresponding normalized eigenvectors of S_z. These probabilities are $\frac{|a|^2}{|a|^2+|b^2|}$ and $\frac{|b^2|}{|a|^2+|b^2|}$. On the other hand, if we measure S_x, the possible values obtained are again $\pm\frac{1}{2}\hbar$ but with probabilities $\frac{1}{2}|a+b|^2$ and $\frac{1}{2}|a-b|^2$, where we assume the state has been normalized so that $|a|^2 + |b|^2 = 1$.

• The hydrogen atom hamiltonian commutes not just with L^2, L_z but also S^2 and S_z. All these operators commute pairwise. So we can label stationary states by their eigenvalues, n, l, m and also s, m_s. Of course $s = \frac{1}{2}$ is fixed for electrons and the energy eigenvalues are independent of l, m, m_s. So accounting for the two possible spin projections of an electron, the degeneracy of the hydrogen energy levels is $2n^2$.

7 Addition of angular momenta

7.1 State space and observables for two particle systems

So far we dealt with 1-particle systems (electron in an atom, spin in a magnetic field e.t.c.). Now suppose we have a system of two particles, each of whose space of quantum states are the Hilbert spaces $\mathcal{H}_1$ and $\mathcal{H}_2$. For example, $\mathcal{H}_1$ could be specified by a basis, i.e., $\mathcal{H}_1$ is the linear span of a collection of basis vectors, which may be the energy eigenstates of an SHO if we have a particle in one dimension. Similarly, we specify $\mathcal{H}_2$ as the linear span of a set of basis vectors. Then the space of states of the combined system is the so-called tensor/direct product Hilbert space $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$. If ψ_n and ϕ_m are a basis for $\mathcal{H}_1$ and $\mathcal{H}_2$ respectively, then a basis for $\mathcal{H}_1 \otimes \mathcal{H}_2$ is given by the vectors $\psi_n \otimes \phi_m$. So any state of the combined system is a linear combination of tensor products of basis states of the individual systems.

• This definition is imposed on us by the superposition principle of quantum mechanics. In classical mechanics, the state of a system of 2 particles is given by specifying the state of each particle (x_1, p_1, x_2, p_2). Analogously in quantum mechanics we might expect the state of the system to be specified by giving the states of each particle, say $|\psi_1\rangle|\phi_2\rangle$. However, the superposition principle says that these cannot be all the states, since a linear combination of states is also an allowed state. This directly leads to the conclusion that arbitrary linear combinations of product states are allowed states of the combined system. Now the tensor product is merely a notation for writing these product states. The definition of tensor product Hilbert space is merely kinematical, it does not presuppose any particular interaction between the particles.

• E.g., if we have two spin half particles, a basis for their combined spinorial Hilbert space is given by

$$|\uparrow\rangle \otimes |\uparrow\rangle \equiv |\uparrow\uparrow\rangle, \quad |\uparrow\rangle \otimes |\downarrow\rangle \equiv |\uparrow\downarrow\rangle, \quad |\downarrow\rangle \otimes |\uparrow\rangle \equiv |\downarrow\uparrow\rangle, \quad |\downarrow\rangle \otimes |\downarrow\rangle \equiv |\downarrow\downarrow\rangle. \tag{137}$$

We sometimes abbreviate by dropping the $\otimes$ symbol as above. The dimension of the tensor product Hilbert space is the product of dimensions.

• Observables for the two particle system are (hermitian) operators acting on the tensor product Hilbert space. Examples of such operators are tensor products of operators acting on each particle separately. For example, if $\vec{S}_1$ and $\vec{S}_2$ are the spin operators for each, then we have the total spin operator

$$\vec{S} = \vec{S}_1 \otimes 1 + 1 \otimes \vec{S}_2 \tag{138}$$

The identity operator for the combined system is $1 \otimes 1$. We often drop the $\otimes$ symbol and suppress the identity operators and write $\vec{S} = \vec{S}_1 + \vec{S}_2$. The z-component of the first spin is $S_{1z} \otimes 1$. The total z-component of spin is $S_z = S_{1z} \otimes 1 + 1 \otimes S_{2z} \equiv S_{1z} + S_{2z}$. A tensor product operator $A \otimes B$ acts on a basis state of the tensor product Hilbert space by

$$(A \otimes B)(u \otimes v) = (Au) \otimes (Bv) \tag{139}$$

The composition of a pair of tensor product operators is given by

$$(A \otimes B)(C \otimes D) = AC \otimes BD \tag{140}$$

This is what you get if you keep in mind that C acts on the state of the first particle followed by A, again on the first particle. Similarly, D acts on the state of the second particle followed by the action of B. For example, the square of total spin works out to be

$$S^2 = \left(\vec{S}_1 + \vec{S}_2\right)^2 = S_1^2 \otimes 1 + 1 \otimes S_2^2 + 2\vec{S}_1 \overset{\otimes}{\cdot} \vec{S}_2. \tag{141}$$

Make sense of this formula and work out how it acts on the above-mentioned tensor product states.

• The tensor product carries a natural inner product induced by the inner products on the individual Hilbert spaces. On basis states, the inner product is

$$\langle u \otimes v | w \otimes z \rangle = \langle u | w \rangle \langle v | z \rangle. \tag{142}$$

For example, $\langle \uparrow \otimes \uparrow | \uparrow \otimes \downarrow \rangle = \langle \uparrow | \uparrow \rangle \langle \uparrow | \downarrow \rangle = 0$.

7.2 Addition of two spin-half angular momenta

The sum of two spin operators $\vec{S} = \vec{S}_1 + \vec{S}_2$ satisfies the same angular momentum algebra as each of the individual spins. Check that

$$[S_x, S_y] = [S_{1x} + S_{2x}, S_{1y} + S_{2y}] = i\hbar(S_{1z} + S_{2z}) = i\hbar S_z, \quad \text{e.t.c.} \tag{143}$$

In particular we can carry over all our results on the matrix representation of spin observables which depend only on these commutation relations. In particular, $[S^2, S_z] = 0$ and we have raising and lowering operators $S_\pm = S_{1\pm} + S_{2\pm}$ which raise or lower the eigenvalue of S_z by $\hbar$.

We may denote the simultaneous eigenstates of S^2 and S_z by $|sm\rangle$ and as before we must have highest/lowest weight states $|ss\rangle, |s,-s\rangle$ that are annihilated by $S_\pm$. Of course, there could be more than one highest weight state labelled by different values of s. From each highest weight state we can build a representation of spin operators by acting repeatedly with S_-. Now we want to know what possible values s can take given that $\vec{S} = \vec{S}_1 + \vec{S}_2$ is the sum of two spin half observables. For each value of s, the allowed values of m will be as before, running from $-s$ to s in steps of one.

• Since $|sm\rangle$ is a state of a two-spin system, it must be expressible as a linear combination of tensor products of states of single spins. There are four tensor product basis vectors

$$|\uparrow\uparrow\rangle, |\uparrow\downarrow\rangle, |\downarrow\uparrow\rangle, |\downarrow\downarrow\rangle \tag{144}$$

We find that all of these are eigenstates of $S_z = S_{1z} + S_{2z}$ which just adds the z-components of the two spins. The corresponding values of m are $1, 0, 0, -1$. So it would appear that the value of s is either 1 or 0, as we will see.

• Now $S_+|\uparrow\uparrow\rangle = 0$. So we can build one representation by successively applying $S_- = S_{1-} + S_{2-}$ to $|\uparrow\uparrow\rangle$

$$S_-|\uparrow\uparrow\rangle = \hbar|\downarrow\uparrow\rangle + \hbar|\uparrow\downarrow\rangle; \qquad S_-\left(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle\right) = 2\hbar|\downarrow\downarrow\rangle, \qquad S_-|\downarrow\downarrow\rangle = 0. \tag{145}$$

These three basis states are found to be eigenstates of S^2 with eigenvalue $2\hbar^2$ (show this, using the formulae $S^2 = S_+S_- + S_z^2 - \hbar S_z = S_-S_+ + S_z^2 + \hbar S_z$). Thus we have a spin $s = 1$ three dimensional 'triplet' representation on the states

$$|11\rangle = |\uparrow\uparrow\rangle, \qquad |10\rangle = \frac{1}{\sqrt{2}}\left(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle\right), \qquad |1,-1\rangle = |\downarrow\downarrow\rangle. \tag{146}$$

The remaining orthogonal state with $m = 0$,

$$|00\rangle = \frac{1}{\sqrt{2}}\left(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle\right) \tag{147}$$

is annihilated by both S_- and S_+ and therefore furnishes a 1-dimensional representation of total spin. It is shown to be an eigenstate of S^2 and S_z with eigenvalues 0 and 0. Thus it corresponds to $s = 0, m_s = 0$ and is called the singlet state $|00\rangle$.

• To summarize, the total spin of a composite system consisting of two spin half particles can be either $s = 0$ or $s = 1$. This is expressed in the formula $\frac{1}{2} \otimes \frac{1}{2} = 1 \oplus 0$. The four dimensional Hilbert space of the composite system splits into the *direct sum* of a one dimensional subspace spanned by the singlet state with $s = 0$ and a three dimensional subspace spanned by the triplet states with $s = 1$. Notice that the spin singlet state is anti-symmetric under exchange of particles while the spin triplet states are all symmetric under exchange of particles; this will be exploited when we discuss the Pauli principle.

• The Hilbert space of the combined system has two interesting bases. The first is the 'un-coupled' basis of tensor product states $\uparrow\uparrow, \uparrow\downarrow, \downarrow\uparrow, \downarrow\downarrow$ which are common eigenstates of S_{1z}, S_{2z}, S_1^2 and S_2^2, which are all simultaneously diagonalizable. (In fact S_1^2 and S_2^2 are just proportional to the identity). The un-coupled basis states may be denoted $|s_1, m_1, s_2, m_2\rangle$.

• Then there is the 'coupled' basis of eigenstates of S^2, S_z (and S_1^2 and S_2^2) which we could denote $|s, m, s_1, s_2\rangle$. Here $s_1 = s_2 = \frac{1}{2}$ are constant so we usually suppress them. The coupled basis states consist of the triplet and singlet states.

• We can express the basis vectors of the coupled basis as linear combinations of the uncoupled basis vectors and vice versa.

$$|s\ m\ s_1\ s_2\rangle = \sum_{m_1,m_2} C^{s,m}_{m_1,m_2} |s_1\ m_1\rangle \otimes |s_2\ m_2\rangle. \qquad (148)$$

The coefficients in these linear combinations are called Clebsch-Gordan coefficients. We have obtained a few of them above.

7.3 Addition of two or more angular momenta

• More generally, we can combine two spins or angular momenta l_1 and l_2. The result is that the angular mometum l of the combination takes each value from $|l_1 - l_2|$ to $l_1 + l_2$ in integer steps. This is expressed in the formula

$$l_1 \otimes l_2 = (l_1 + l_2) \oplus (l_1 + l_2 - 1) \oplus \cdots \oplus |l_1 - l_2|. \qquad (149)$$

For example, if we form a composite system from two spin one particles, the combined system again behaves like a spin system, but with spin either 2 or 1 or 0. The Clebsch-Gordan coefficients for combining angular momenta are tabulated in various quantum mechanics books. To combine three spins, we first combine two of them and then the third and so on.

• This quantum mechanical procedure for finding the resultant spin is to be contrasted with the vector addition formula for combining angular momenta in classical mechanics. However, there is some similarity. Classically, the resultant angular momentum $\vec{L} = \vec{L}_1 + \vec{L}_2$ can take a maximum magnitude equal to $|\vec{L}_1| + |\vec{L}_2|$ when the angular momenta are parallel and a minimal value $||\vec{L}_1| - |\vec{L}_2||$ when the angular momenta are anti-parallel. These maximal and minimal values are analogous to the quantum mechanical result that the maximum angular momentum quantum number is $l_1 + l_2$ and minimal value is $|l_1 - l_2|$. Of course, classically every value in between is also a possibility depending on the relative orientations of the two angular momentum vectors. Quantum mechanically, only certain intermediate values of l are allowed, in integer steps! The classical limit is obtained in the limit of large quantum numbers $l(l+1)\hbar^2 \gg \hbar^2, l_1, l_2 \gg 1$. In this limit, $\sqrt{\hbar^2 l(l+1)} \approx \hbar l \to |\vec{L}|$ e.t.c. Moreover the size of the small integer steps in units of $\hbar$ is very small compared to the angular momentum and a continuum of values is approximately obtained.

• More generally we may combine any number of spins/angular momenta $s_1 \otimes s_2 \otimes \ldots \otimes s_n$. Then there is a unique (up to normalization) highest weight state (annihilated by S_+) ψ_{hws} with maximal spin projections $S_z|s_1, \cdots, s_n\rangle = \hbar(s_1 + \cdots + s_n)|s_1, \cdots, s_n\rangle$. Applying S_- to this state produces an eigenstate of S_z with eigenvalue $\hbar(s_1 + \cdots + s_n - 1)$ which is $\hbar$ less than maximal. This is the symmetric state

$$\psi_S \propto |s_1 - 1, s_2, \cdots, s_n\rangle + |s_1, s_2 - 1, \cdots, s_n\rangle + \cdots + |s_1, s_2, \cdots, s_n - 1\rangle \qquad (150)$$

We may successively apply S_- to this to produce a spin $s_1 + \cdots + s_n$ representation. However the space of states with spin projection $\hbar$ less than maximal is n dimensional with basis

$$|s_1 - 1, s_2, \cdots, s_n\rangle, \quad |s_1, s_2 - 1, \cdots, s_n\rangle, \ \ldots, \ |s_1, s_2, \cdots, s_n - 1\rangle. \qquad (151)$$

Any state ψ in this subspace that is orthogonal to ψ_S is automatically annihilated by S_+ and furnishes a highest weight state for a new spin $s_1 + \cdots + s_n - 1$ representation. This procedure is then repeated. Let us see why $S_+\psi = 0$. Since $S_z(S_+\psi) = \hbar(s_1 + \cdots + s_n)S_+\psi$, and the space of states with maximal spin projection is one dimensional we must have for some constant c

$$S_+\psi = c\psi_{hws} \tag{152}$$

We will show that $c = 0$ if $\psi \perp \psi_S$. Taking an inner product

$$\langle\psi_{hws}|S_+|\psi\rangle = \langle\psi_{hws}|c\psi_{hws}\rangle = c \quad\Rightarrow\quad c = \langle S_-\psi_{hws}|\psi\rangle = \langle\psi_S|\psi\rangle = 0. \tag{153}$$

The possible resultant spins of a composite system of three spin half particles (e.g. the three electrons in a Lithium atom) can be worked out by combining two of the spins first and then the third

$$\frac{1}{2}\otimes(\frac{1}{2}\otimes\frac{1}{2}) = \frac{1}{2}\otimes(0\oplus 1) = \frac{1}{2}\oplus\frac{1}{2}\otimes 1 = \frac{1}{2}\oplus\frac{1}{2}\oplus\frac{3}{2}. \tag{154}$$

So a system of three spin half particles can behave as a spin half particle in two different ways and also as a spin $3/2$ particle. The $\otimes$ can be read as 'and', and $\oplus$ as 'or', they are the direct product (tensor product) and direct sum, respectively. So when reading $\frac{1}{2}\otimes\frac{1}{2}$ we say that if a system consists of a spin half particle AND another spin half particle, then the composite system behaves either as a spin zero system OR as a spin one system.

8 Multiparticle systems: bosons & fermions

8.1 Identical and indistinguishable particles

In classical physics we can distinguish a pair of tennis balls since they may not have exactly the same mass or number of fibres. Classical objects may be identified/labelled by their intrinsic properties detected by their different interactions with various measuring devices (like light (for color) or smell or mass). By intrinsic properties we mean charge, mass, shape or inertia tensor etc, not the position or speed of the particle. Even if the masses, shapes and inertia tensors were the same to the accuracy of our devices (i.e., the balls seem identical) we could still distinguish the two tennis balls by keeping track of their historical trajectories $(x(t), p(t))$, ball A was the one that left the factory at 10am on 23/4/10 etc while ball B had a slightly different trajectory (after all we have never found two tennis balls with coincident trajectories even for a short duration).

• The situation with objects needing a quantum mechanical treatment is somewhat different. No experiment so far has been able to distinguish one electron from another, they all have the same mass, charge, spin, magnetic moment and interactions with light, atoms etc. What is more, the possibility of distinguishing two electrons based on their semi-classical trajectories may seem to work approximately for some time, but can easily fail. The electron wave packets, if they come quite close, would overlap, interfere and we would not be able to unambiguously label the electrons when the wave packets separate.

• This is a particularly severe problem when we are dealing with atoms with several electrons. The electrons do not have well-defined orbits, their wave functions overlap and it has not been possible to keep track of which electron is which as time progresses. There are many other examples of identical particles in nature including photons in a black body cavity, neutrons in a neutron star, Rubidium atoms in a laser trap etc.

• Contrast this with the situation with planets in the solar system where we can distinguish them by their color/mass/temperature/support for life/historical trajectories etc. So quantum mechanically electrons are identical and indistinguishable. Of course, not all particles are identical and indistinguishable. We *can* tell a proton apart from an electron by its mass.

• The state of a multi-particle system is a ray in the tensor product Hilbert space. Let us consider two identical particles for simplicity. Since the particles are identical, their individual Hilbert spaces of states can be taken to be the same $\mathcal{H}_1 \cong \mathcal{H}_2$. If ϕ_a is a basis for the individual Hilbert space, a vector in the tensor product $\mathcal{H}_1 \otimes \mathcal{H}_2$ is of the form

$$\psi(x_1, x_2) = \sum_{a,b} c_{ab}\ \phi_a(x_1)\ \otimes\ \phi_b(x_2) \tag{155}$$

Define the permutation or exchange operator (not to be confused with parity) $P\psi(x_1, x_2) = \psi(x_2, x_1)$ which exchanges the coordinates (and spins if we were keeping track of the spins). Check that this implies that P is hermitian. To implement the identical and indistinguishable nature of the particles, we postulate that the vector ψ and the vector $P\psi$ must represent the same physical state. This implies that

$$\psi(x_1, x_2) = e^{i\theta}\psi(x_2, x_1) \tag{156}$$

for some phase $e^{i\theta}$. We will find that $e^{i\theta} = \pm 1$ are the allowed values, corresponding to bosons and fermions respectively. In particular,

$$P\psi(x_1, x_2) = \psi(x_2, x_1) = e^{-i\theta}\psi(x_,x_2). \tag{157}$$

From the definition of P we also see that $P^2 = I$ is the identity. Thus we have

$$P^2\psi(x_1, x_2) = \psi(x_1, x_2) \quad \text{and} \quad P^2\psi(x_1, x_2) = e^{-2i\theta}\psi(x_1, x_2). \tag{158}$$

Comparing we find $e^{-2i\theta} = 1$ or $e^{i\theta} = \pm 1$, which means there are only two possibilities

$$\psi(x_1, x_2) = \pm\psi(x_2, x_1) \tag{159}$$

and we say that wavefunctions of a system of two (though it is true also for more than two) identical particles must be either symmetric or antisymmetric under exchange. Note that this applies to all states, not just energy eigenstates!

• In addition we postulate that exchange of particles must be a symmetry of the dynamics, i.e., classically the forces and energy must not be affected by a permutation of the particles (if this weren't true we could distinguish the particles). This means the classical hamiltonian is symmetric

$$H(x_1, p_1; x_2, p_2) = H(x_2, p_2; x_1, p_1) \tag{160}$$

In the quantum theory this means $[P, H] = 0$. Thus, P and H are simultaneously diagonalizable. Energy eigenstates may therefore be chosen to be eigenstates of the exchange operator and must be either symmetric or anti-symmetric functions of the coordinates.

• The symmetric and anti-symmetric functions describe different types of particles. Fermionic particles are those with anti-symmetric wave functions while bosonic particles are those with symmetric wave functions. Electrons in an atom need to be described by anti-symmetric wave functions to match the measured atomic spectra and periodic table. They are fermions. Photons and many isotopes of alkali metal atoms need to be described by symmetric wave functions to match their experimental properties such as Bose condensation. They are bosons. Remarkably, it is found that particles with integer spin $s = 0, 1, 2, \ldots$ are bosons and those with half-odd-integer spin $s = \frac{1}{2}, \frac{3}{2}, \ldots$ are fermions.

• What is more, $[H, P] = 0$ implies that if the initial state is an eigenstate of P, then it will remain an eigenstate of P under Schrödinger evolution with the same eigenvalue. So a system of identical bosons remains this way at all times. We may build multi-particle wave functions by taking products of single particle orbitals and either symmetrizing or anti-symmetrizing for bosons and fermions respectively.

• Examples of symmetrized and antisymmetrized wave functions for two particles are

$$\psi_S(x_1, x_2) = \phi_a(x_1)\phi_b(x_2) + \phi_a(x_2)\phi_b(x_1), \qquad \psi_A(x_1, x_2) = \phi_a(x_1)\phi_b(x_2) - \phi_a(x_2)\phi_b(x_1). \tag{161}$$

We remark that this anti-symmetric state could also be written as a 'Slater' determinant.

$$\psi_A(x_1, x_2) = \det \begin{pmatrix} \phi_a(x_1) & \phi_a(x_2) \\ \phi_b(x_1) & \phi_b(x_2) \end{pmatrix} \tag{162}$$

Written this way, antisymmetry is equivalent to the statement that the determinant changes sign under a permutation of two columns. This way of writing anti-symmetric states is particularly useful when there are more than two particles.

• Of course, these are just the simplest examples of (anti-)symmetrized wave functions built out of a pair of single particle wavefunctions. More generally we could take linear combinations of (anti-)symmetric wavefunctions to produce new (anti-)symmetric wave functions. E.g. we may build a fermionic state from four one particle wavefunctions

$$\psi_A(x_1, x_2) = N_1 \det \begin{pmatrix} \phi_a(x_1) & \phi_a(x_2) \\ \phi_b(x_1) & \phi_b(x_2) \end{pmatrix} + N_2 \det \begin{pmatrix} \phi_c(x_1) & \phi_c(x_2) \\ \phi_d(x_1) & \phi_d(x_2) \end{pmatrix} \tag{163}$$

• Symmetric and anti-symmetric wave functions have rather different physical features. For example, the probability amplitude to find both particles at the same location $x_1 = x_2 = x$ in an anti-symmetrized state is zero $\psi_A(x, x) = 0$, though it can be non-zero in a symmetrized state. Alternatively $\psi_A = 0$ if $\phi_a = \phi_b = \phi$. This is loosely stated as 'two identical fermions (e.g. electrons) *cannot* occupy the same 1-particle state ϕ'. On the other hand, two identical bosons (e.g. photons) *can* occupy the same state. We use this while obtaining the distribution of photons in a black body cavity.

• This property of electrons had been postulated by Pauli (even before the development of Schrodinger's equation), in an attempt to explain atomic spectra and the periodic table. Pauli's exclusion principle stated that there was at most one electron per state in an atom. The states could be labelled by the same quantum numbers as appearing in the hydrogen atom n, l, m and an extra quantum number with two possible values. The latter quantum number was subsequently identified as the electron spin projection m_s. Of course, the hydrogen atom stationary state wave functions do not account for electron-electron repulsion, and cannot just be taken over to construct the wave function of a many-electron atom. However, there are ways to do this approximately. In any case, the Pauli-exclusion principle in the form 'multi-electron wave functions must be anti-symmetric' under exchange of quantum numbers (coordinates, spins) of any pair of electrons, has stood the test of time.

• For a system of N identical particles, we define the exchange operators which permute the quantum numbers of i^{th} and j^{th} particle (coordinates and spin projections)

$$P_{ij}\psi(\cdots,x_i,\cdots,x_j,\cdots) = \psi(\cdots,x_j,\cdots,x_i,\cdots), \qquad \text{for } 1 \le i,j \le N. \tag{164}$$

Then each P_{ij} must commute with the hamiltonian. Every energy eigenstate $\psi(x_1,\cdots x_n)$ can be taken as either symmetric or antisymmetric under *every* such permutation. Symmetric states describe bosons and anti-symmetric ones describe fermions.

• For three identical particles we have three exchange operators P_{12}, P_{23}, P_{31}. Note that $P_{12} = P_{21}$ etc. The action of a permutation operator on a state vector $\psi(x_1,x_2,x_3)$ must produce a state vector that differs from the original one at most by a phase: $P_{ij}\psi = e^{\mathrm{i}\theta_{ij}}\psi$. Each permutation operator P_{ij} squares to the identity $P_{ij}P_{ij} = I$, proceeding as before, we find $e^{i\theta_{ij}} = \pm 1$.

• In other words, every state of a system of identical and indistinguishable particles must be an eigenstate of all the $P_{ij}'s$ with eigenvalues ± 1. Note that the permutation operators pairwise commute $P_{ij}P_{kl} = e^{i(\theta_{ij}+\theta_{kl})} = P_{kl}P_{ij}$. So they are simultaneously diagonalizable. Now a question arises whether a quantum mechanical system of identical and indistinguishable particles can be in a state where some of the P_{ij}'s have eigenvalue one and some other permutations have eigenvalue minus one. This is not allowed since if there were such a state where, say, $P_{12}\psi = \psi$ and $P_{13}\psi = -\psi$ then we could use these eigenvalues to distinguish between particle 2 and particle 3, violating the indistinguishability. So the only allowed states of identical particles are those that are symmetric under *every* exchange and those that are anti-symmetric under *every* exchange. Moreover, since permutations of particles must be symmetries of the dynamics, we must also have $[P_{ij}, H] = 0$ for all i,j. Thus, the states which are symmetric under every exchange are preserved under time evolution; they are called systems of bosons. The anti-symmetric states are also preserved under time evolution and are called systems of fermions.

• An example of a fermionic state of N particles is given by the Slater determinant, which is constructed using N one particle wave functions ('orbitals'). For $N = 3$ let ϕ_a, ϕ_b, ϕ_c be three one-particle wave functions ('orbitals')

$$\psi_A(x_1,x_2,x_3) = \det\begin{pmatrix} \phi_a(x_1) & \phi_a(x_2) & \phi_a(x_3) \\ \phi_b(x_1) & \phi_b(x_2) & \phi_b(x_3) \\ \phi_c(x_1) & \phi_c(x_2) & \phi_c(x_3) \end{pmatrix} \tag{165}$$

To build an example of a bosonic state of three identical particles we only need a single 1-particle orbital $\psi_S(x_1,x_2,x_3) = \phi(x_1)\phi(x_2)\phi(x_3)$.

• We may also consider what happens to the state vector of a system of identical particles under permutations that are not simple exchanges of particles. An example is a cyclic permutation of the particles. However, every permutation can be built by composing exchanges, so these new permutations do not contain new information. One finds that wave functions of systems of fermions are anti-symmetric under odd-permutations, such as pairwise exchanges. Under even permutations such as $P_{12}P_{23}$, wave functions of systems of fermions are unchanged. Wave functions of systems of bosons are unchanged under all permutations, irrespective of whether they are even or odd.

8.2 Two spin half electrons in Helium

Consider a pair of spin half fermions, whose orbital motion is not coupled to the spin degrees of freedom, e.g. if the hamiltonian operator is the identity in spin space. This is true for the two electrons in Helium in a non-relativistic treatment where the hamiltonian does not involve any spin-orbit coupling terms:

$$H = \frac{p_1^2}{2m} + \frac{p_2^2}{2m} - \frac{2e^2}{4\pi\epsilon_0 r_1} - \frac{2e^2}{4\pi\epsilon_0 r_2} + \frac{e^2}{4\pi\epsilon_0 |\vec{r}_1 - \vec{r}_2|}. \tag{166}$$

H is symmetric under exchange of the degrees of freedom of the two particles. So it commutes with the exchange operator and we seek simultaneous eigenstates of H and P. We write the total wave function as a product of a spatial and a spin part[3]

$$\psi(x_1, m_1; x_2, m_2) = \phi(x_1, x_2)\chi(m_1, m_2) \tag{167}$$

$\phi(x_1, x_2)$ is $\propto$ to the probability amplitude for the first particle to be at x_1 and second to be at x_2 irrespective of their spin states. Similarly, $\chi(m_1, m_2)$ is $\propto$ the probability amplitude for the spin projection of the first particle to be $\hbar m_1$ and that of the second particle to be $\hbar m_2$, irrespective of their positions[4]. Here m_1 and m_2 only take the values $\pm\frac{1}{2}$ while x_1, x_2 are any points in 3d space. In general,

$$\chi = a \uparrow\uparrow + b \uparrow\downarrow + c \downarrow\uparrow + d \downarrow\downarrow \quad \text{where} \quad \chi\left(\frac{1}{2}, \frac{1}{2}\right) = a, \quad \chi\left(\frac{1}{2}, -\frac{1}{2}\right) = b, \ldots. \tag{168}$$

The Pauli principle of anti-symmetrization states that the total wavefunction must be anti-symmetric under exchange of both coordinate and spin degrees of freedom

$$\psi(x_1, m_1; x_2, m_2) = -\psi(x_2, m_2; x_1, m_1) \tag{169}$$

Now let us look for other operators that commute with both H and P. $H \propto I$ in spin space, so it commutes with the combined spin operators of the two particles S^2 and S_z as well as with S_1^2 and S_2^2 which are each proportional to the identity. All these four spin observables also commute with P. On the other hand, S_{1z} commutes with H but not with P. So in particular, P is *not* diagonal in the uncoupled basis of spin states, but it is diagonal in the coupled basis.

- So we look for energy eigenstates that are simultaneous eigenstates of P, S^2 and S_z. Then there are only two ways to make ψ anti-symmetric: ϕ and χ are symmetric and anti-symmetric respectively or vice-versa.

$$^1\psi = \phi_S(x_1, x_2)\chi_A(m_1, m_2) \quad \text{or} \quad ^3\psi = \phi_A(x_1, x_2)\chi_S(m_1, m_2) \tag{170}$$

We have seen that there is only one anti-symmetric 2 particle spin state, the singlet $\chi_A = \frac{1}{\sqrt{2}}(\uparrow\downarrow - \downarrow\uparrow)$ while there are a triplet of symmetric spin states χ_S. In the case of Helium, the spin singlet states are called parahelium and the spin triplet ones orthohelium. Orthohelium states are triply degenerate since the hamiltonian is diagonal in spin space.

---

[3] If the orbital and spin degrees of freedom were coupled, we would need linear combinations of such product wave functions

[4] If the spin and coordinate degrees of freedom were coupled, the amplitudes for spin projections could depend on position.

• Often we may approximate ϕ by an (anti-)symmetrized combination of single particle wave functions ϕ_1 and ϕ_2 (in general we will need linear combinations of these):

$$\phi_A(x_1,x_2) = \frac{1}{\sqrt{2}}\left(\phi_1(x_1)\phi_2(x_2) - \phi_1(x_2)\phi_2(x_1)\right) \quad \text{and} \quad \phi_S(x_1,x_2) = \frac{1}{\sqrt{2}}\left(\phi_1(x_1)\phi_2(x_2) + \phi_1(x_2)\phi_2(x_1)\right)$$

In the Helium atom, these two single particle wave functions ϕ_1, ϕ_2 could be taken as two stationary state wavefunctions of a Hydrogenic atom with nuclear charge $Z = 2$[5]. By making different choices for ϕ_1 and ϕ_2 we may search for a state ψ which minimizes the expectation value of the Helium hamiltonian. It turns out that the ground state of the Helium atom is a spin singlet (parahelium) state $\phi_S \chi_A$ where ϕ_1 and ϕ_2 are both approximately ground state wave functions of hydrogenic atoms. This is why we say Helium has the electronic configuration $1s^2$. The first excited state is a triplet of degenerate orthohelium states etc. To find the eigenfunctions (only $\phi(r_1,r_2)$ remains to be found) we must solve the Schrodinger eigenvalue problem. This is quite hard and one resorts to approximations such as perturbation theory and the variational method. In both approaches, one needs to calculate the expectation value of the the helium hamiltonian. The variational principle states that the ground state energy is the minimal value of $\langle\psi|H|\psi\rangle$ among all unit norm states.

8.3 Coulomb and exchange interactions

To find the ground state of helium, we need to minimize the expectation value of H among all anti-symmetric two particle states. The symmetry /antisymmetry of $\phi_{S,A}$ imply some interesting qualitative features for the expectation value of the interaction energy arising from electrostatic repulsion of electrons. The interaction part of the helium hamiltonian can be written

$$V_I = G(r_1 - r_2) = \frac{e^2}{4\pi\epsilon_0 |r_1 - r_2|} \tag{171}$$

Since electrons are identical, $G(r_1 - r_2) = G(r_2 - r_1)$. As discussed in the last section, we are interested in spatial wave functions that are built from a pair of (normalized) one-particle wave functions

$$\phi_{S,A} = \frac{1}{\sqrt{2}}\left[\phi_a(r_1)\phi_b(r_2) \pm \phi_a(r_2)\phi_b(r_1)\right] \tag{172}$$

Let us calculate the expectation value of V_I in the states $\phi_{S,A}$ Show that you get

$$\langle V_I\rangle = \int d^3r_1 d^3r_2 G(r_1 - r_2)\left[|\phi_a(r_1)|^2|\phi_b(r_2)|^2 \pm \Re\, \phi_a^*(r_1)\phi_b^*(r_2)\phi_a(r_2)\phi_b(r_1)\right] = V_C \pm V_E \tag{173}$$

The first term V_C is called the Coulomb (or direct) interaction energy while the second term V_E is the exchange interaction energy. Of course, both arise from the Coulomb repulsion of electrons. While the direct Coulomb interaction can be classically interpreted as the energy of a pair of charge clouds, the exchange term has no such simple classical interpretation. Indeed, it is a quantum mechanical effect due to the identical nature of particles and (anti-)symmetry of wave functions. Moreover, the exchange interaction contributes with a different sign according as the spatial wave function is symmetric or anti-symmetric (spin singlet or spin triplet state). The Coulomb energy V_C is clearly positive. In cases where the exchange term is also positive

---

[5] If we ignore the electron-electron repulsion, each electron is in a hydrogenic atom with $Z = 2$.

$V_E \geq 0$ (as is the case for Helium) we can conclude that the interaction energy $\langle V_I \rangle$ is greater for ϕ_S than for ϕ_A. So (other things being equal) the spin singlet states ${}^1\chi_A$ will have a higher energy than the spin triplet states ${}^3\chi_S$. This is seen in Helium, where typically the spin triplet orthohelium states have a lower energy than the corresponding spin singlet paraheluim states constructed from the same pair of single particle states ϕ_a, ϕ_b. There is one notable exception to this rule: the ground state of helium is a spin singlet parahelium state with $\phi_a = \phi_b$. There is no anti-symmetric wave function that can be constructed if $\phi_a = \phi_b$.

8.4 Electronic configuration of light elements and periodic table

• Now let us briefly apply the ideas of identical and indistinguishable particles and Pauli's principle to the n-electrons in an atom, governed by the hamiltonian

$$H = \sum_{i=1}^{N} \left(\frac{p_i^2}{2m} - \frac{Ze^2}{4\pi\epsilon_o |\vec{r}_i|} \right) + \sum_{1 \leq i < j \leq N} \frac{e^2}{4\pi\epsilon_o |\vec{r}_i - \vec{r}_j|} \tag{174}$$

To determine and understand the ground (and excited states) of the atom is a very challenging problem, which occupied and continues to occupy many physicists and chemists. To find the ground state, we may (as a first approximation) ignore the inter-electron repulsion. In this approximation, it would seem that each electron 'sees' only the nucleus in a hydrogenic atom with nuclear charge Z. But there is a crucial difference. The N-electron wave function $\psi(\vec{r}_1, m_1; \cdots, \vec{r}_N, m_N)$ depends on the positions x_i and spin projections m_i of each of the electrons. By Pauli's principle, it must be antisymmetric under exchange of any pair of electrons. A simple way of constructing such an antisymmetric wave function is to choose n one-particle wave functions or orbitals $\phi_1, \cdots, \phi_n$ and form their Slater determinant, e.g.,

$$\psi(\vec{r}_1, \vec{r}_2, \vec{r}_3) = A \det \begin{pmatrix} \phi_1(r_1) & \phi_1(r_2) & \phi_1(r_3) \\ \phi_2(r_1) & \phi_2(r_2) & \phi_2(r_3) \\ \phi_3(r_1) & \phi_3(r_2) & \phi_3(r_3) \end{pmatrix} \tag{175}$$

We suppress the spin projections m_i so r_i refers to $\vec{r}_i, m_i$. Now the orbitals must be chosen carefully to ensure that the energy is minimized (this is the statement of the variational principle). An example of an orbital is the ground state wave function of hydrogen.

$$\phi_1(r) = \langle r | n = 1, l = 0, m_l = 0, s = \frac{1}{2}, m_s = \frac{1}{2} \rangle \propto e^{-r/a_o} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{176}$$

We could equally well have chosen $m_s = -\frac{1}{2}$. The ground state electronic configuration of hydrogen is called $1s^1$, one electron in the $n = 1$ shell with zero angular momentum ($s = 0$). For Helium, we select a second hydrogenic orbital with as little energy as possible $|\phi_2\rangle = |n = 1, l = 0, m_l = 0, s = \frac{1}{2}, m_s = -\frac{1}{2}\rangle$, different from ϕ_1 to implement Pauli's principle. ϕ_1 and ϕ_2 are anti-symmetrized via a 2×2 Slater determinant, resulting in

$$\psi = \phi_{n=1,l=0,m=0}(r_1)\phi_{n=1,l=0,m=0}(r_2)\frac{1}{\sqrt{2}}(\uparrow\downarrow - \downarrow\uparrow) \sim e^{-(r_1+r_2)/a}(\uparrow\downarrow - \downarrow\uparrow). \tag{177}$$

This wave function is symmetric under exchange of spatial variables but anti-symmetric under exchange of spins, it is the singlet spin-zero state obtained from combining two spin halves. It is called para helium. The electronic configuration of Helium is $1s^2$, two S-wave electrons in the $K(n=1)$ shell. Helium has a filled K-shell.

- For Lithium, we could select the next orbital as $|\phi_3\rangle = |n=2, l=0, m_l=0, m_s=\frac{1}{2}\rangle$ resulting in $1s^2 2s^1$. Proceeding in this way, we get $1s^2 2s^2 2p^6$ for Neon and $1s^2 2s^2 2p^6 3s^2 3p^6$ for Argon. But the configurations of potassium Ne $4s^1$ and magnesium Ne $4s^2$ are somewhat anomalous, the $4s$ shell is filled before the $3d$ shell, as it is energetically favourable to do so, when one includes the effect of inter-electron repulsion. Let us crudely indicate why. In general, the $n=4$ shell has a higher energy than the $n=3$ shell, since the hydrogenic energies are $E \sim -\frac{1}{n^2}$. On the other hand, in a higher angular momentum orbital (e.g. $l=2$ d-orbital), the electron probability density is concentrated farther from the nucleus than in a lower angular momentum orbital (say s-orbital, $l=0$). This is seen from the $\psi \sim r^l$ behaviour of hydrogenic wave functions, they are more strongly suppressed for small r if l is larger. In a multi-electron atom, like potassium, the nuclear charge is partly screened by the inner shell electrons ('Neon core'), and it is energetically favourable for the valence electron of Potassium to be more tightly bound to the nucleus by occupying the $4s$ orbital rather than the $3d$ orbital. For further discussion of the periodic table, see the books by Griffiths, Liboff, Schiff, Heitler etc mentioned in the list of reference books and the books cited therein.

9 Description of an ensemble of quantum mechanical states by a density matrix

9.1 Pure and mixed ensembles in classical mechanics

The state of a classical system of particles is given by specifying the positions and momenta of all the particles. This is the most that classical mechanics allows us to freely specify. Such a state of a classical mechanical system is called a 'pure' state, for emphasis. Suppose we prepare several copies of a particle (or system) in the same pure state. Such an ensemble is called a pure ensemble, i.e., an ensemble where each constituent is in the same pure state. This could be a set of identical billiard balls all at rest at height 1m above the floor.

Often we do not know all the coordinates and momenta of particles even in an isolated container of gas, owing to the difficulty of measuring all of these even at the initial instant of time (though it is not forbidden classically). On other occasions, the gas may be in contact with an environment. Owing to the interaction of the system with its very complicated environment, it is not possible to pin down the state of the gas precisely. In such situations, it is hopeless to try to describe the state of a gas by a pure state (even if it is, strictly speaking, in a pure state).

However, even in such situations, we may be able to say that among all possible pure states of the gas, some are more likely than others. For example, if the gas is in equilibrium in a room maintained at a fixed temperature, we can say that the gas is more likely to be in a state where the molecules are uniformly distributed over the room than all located in one corner. Similarly, the velocities of the gas molecules are more likely to follow the Maxwell-Boltzmann distribution than one where all the velocities are equal and pointing downwards.

In such circumstances, we describe the properties of a box of gas by considering an ensemble of several chambers of gas, each in a possibly different pure state, with different pure states occurring with different relative probabilities. Then we hope to describe the properties of the original gas box by an ensemble average of the properties in each of the pure states in the various chambers. Such an ensemble of pure states is called a mixed ensemble. Loosely, we say that the gas in the room is in a mixed state. What we mean is that on average, its properties are well approximated by an ensemble average. Thus, a mixed ensemble is a way of saying which pure states are more likely and which are less likely. So a mixed ensemble is an ensemble of pure states with a specification of their relative likelihoods.

Operationally, the specification of a mixed ensemble should allow us to calculate the ensemble average values of observables. For example, we should be able to calculate the average pressure on the walls by averaging over the pressure in each of the possible pure states in the ensemble, weighted by their likelihoods. For a gas in equilibrium at temperature T, the relevant mixed ensemble is the so-called thermal ensemble of Boltzmann that says pure states of high energy are exponentially less likely than pure states of lower energy.

9.2 Pure and mixed ensembles in quantum mechanics

• The most that quantum mechanics allows us to specify about the state of a system is its wave function $\psi(x_S)$ where x_S denotes the degrees of freedom of the system. A quantum system with a wave function is said to be in a (pure) state.

• The measurement postulate implies (among other things) that if we prepare several copies of a system in the same state ψ and make measurements of the observable A, then on average we get the expectation value of A in the state ψ. Such a collection of identically prepared states is called a pure ensemble, where all the states is the ensemble are the same (pure) state.

• More generally, we may consider a so-called mixed ensemble, which contains several copies of the system in possibly different pure states.

• E.g. suppose we pass an unpolarized beam of electrons through a Stern-Gerlach-like apparatus to produce two beams of spin polarized electrons. The electrons in each output beam are in pure spin states. One beam has electrons exclusively in the state $\uparrow$ (i.e. with S_z eigenvalue $\hbar/2$) and the other beam has electrons exclusively in the state $\downarrow$. Each of these beams is separately a pure ensemble.

• Now we combine the two outputs to form a single beam of electrons. In this new beam, all we know is that 50% of the electrons are polarized spin up and 50% polarized spin down. We can regard this combined beam as a mixed ensemble. If this combined beam is fairly weak, and we isolate the individual electrons, we would not know the state of the individual electrons. Any given electron is equally likely to be in one of two pure spin states $\uparrow$ and $\downarrow$. Unlike an electron which is in the state $\uparrow + \downarrow$, this electron *does not* have a definitely known wave function and is said to be in a 'mixed spin state'. Here we are focusing on the spin wave function and not the coordinate degrees of freedom of the electrons.

• Note that the initial unpolarized beam is itself an example of a mixed ensemble. An unpolarized beam is one where there is no preferred direction for the spin vector. For instance it means that if we measure S_z on the electrons in the beam, we will get $\hbar/2$ for half the electrons and $-\hbar/2$ for the other half. It also means that if we measure S_y we will get $\hbar/2$ for half the electrons and $-\hbar/2$ for the other half. The same is true for measurements of S_x on the electrons in such an unpolarized beam.

• Another way in which the concept of a pure state may fail to be adequate is if we have a quantum system in contact with an environment whose degrees of freedom are denoted x_E. Then the system+environment may have a wave function $\psi(x_S, x_E)$. But it is rarely the case that we can factorize this into a system wave function and an environment wave function $\psi_S(x_S)\psi_E(x_E)$ at all times. In such situations, the system does not have a wave function of its own. It is said to be in a mixed state in which we know less about the system (treated as an entity by itself) than is otherwise allowed by quantum mechanics.

• Just as in classical mechanics, we should be able to calculate the average values of observables in any ensemble. QM already tells us how to do this for a pure ensemble. Now consider a mixed ensemble where the normalized pure states ψ_i occur with relative probabilities p_i, with $\sum_i p_i = 1$ where the sum runs over the pure states present in the ensemble. In the next section we will show how to calculate average values of observables in mixed ensembles.

• For example in the above combined beam there are only two distinct pure states in the ensemble, $\uparrow$ and $\downarrow$ with equal populations. The pure states in an ensemble need not be orthogonal, nor do they have to form a basis for the Hilbert space. For example we may have a mixed ensemble of electrons where a third each are in the states $\uparrow$, $\downarrow$ and $\uparrow + \downarrow$. This example also shows that the number of distinct pure states in an ensemble can exceed the dimension of the Hilbert space.

• Irrespective of the type of ensemble, when we measure an observable A on a constituent of an ensemble, we get some eigenvalue of the observable. In a pure ensemble of normalized pure states ψ, we know moreover that we get the eigenvalue a_n with probability $|\langle\phi_n|\psi\rangle|^2$ and the system collapses to the normalized eigenstate ϕ_n of A corresponding to the eigenvalue a_n. This is the famous indeterminacy of quantum mechanics. In a mixed ensemble, there is a *further indeterminacy* beyond that intrinsic to quantum mechanics. This is because there are several different pure states present in the ensemble. When we measure A for a constituent of a mixed ensemble, we get an eigenvalue a_n but the probability of getting this eigenvalue depends on which pure state the constituent of the ensemble happened to be in. The probability of getting a_n can be written as an ensemble average of the probability in each pure state

$$P(a_n) = \text{probability of getting } a_n = \sum_{i \,\in\, \text{ensemble}} p_i |\langle\phi_n|\psi_i\rangle|^2 \tag{178}$$

Check that these probabilities add up to one.

9.2.1 Why the density matrix?

• There is no single wave function that describes all the constituents of a mixed ensemble. But we can still describe such an ensemble by use of a density matrix.

• Suppose we want to predict the average value resulting from measurements of observable A on states in a mixed ensemble consisting of normalized pure states ψ_i occurring with relative frequencies p_i. Quantum mechanics plus ensemble averaging says that it is

$$\langle A\rangle = \sum_i p_i \langle\psi_i|A|\psi_i\rangle \tag{179}$$

Let us insert the identity $I = \sum_n |n\rangle\langle n|$ where $|n\rangle$ is any orthonormal basis for the Hilbert space. Then

$$\langle A\rangle = \sum_i p_i \sum_{n,m} \langle\psi_i|n\rangle\langle n|A|m\rangle\langle m|\psi_i\rangle = \sum_{mn} A_{nm}\left(\sum_i p_i \langle m|\psi_i\rangle\langle\psi_i|n\rangle\right) \equiv \sum_{mn} A_{nm}\rho_{mn} = \text{ tr } A\rho. \tag{180}$$

The quantity within parenthesis is independent of the observable and has been defined to be the density matrix ρ

$$\rho = \sum_i \; p_i \; |\psi_i\rangle\langle\psi_i| = \sum_i p_i \psi_i \psi_i^\dagger \quad \text{and its matrix elements are} \quad \rho_{mn} = \sum_i p_i \langle m|\psi_i\rangle\langle\psi_i|n\rangle \tag{181}$$

The density matrix is a hermitian operator $\rho^\dagger = \sum_i p_i \left(\psi_i\psi_i^\dagger\right)^\dagger = \rho$. It has trace equal to one

$$\text{tr } \rho = \sum_n \rho_{nn} = \sum_i p_i \text{ tr } \psi_i\psi_i^\dagger = \sum_i p_i \psi^\dagger\psi = \sum_i p_i = 1 \tag{182}$$

since the states ψ_i were assumed to be of norm one.

9.2.2 Pure ensemble density matrices

• In particular, the density matrix of a pure ensemble is got by restricting the sum to a single pure state ψ which occurs with probability $p = 1$. So the density matrix of a pure state ψ is $\rho = |\psi\rangle\langle\psi|$. In other words, the density matrix of a pure state is the outer product of the normalized state vector with itself or the ratio of outer to inner products

$$\rho_\psi^{pure} = \frac{|\psi\rangle\langle\psi|}{\langle\psi|\psi\rangle} = \frac{\psi\psi^\dagger}{\psi^\dagger\psi} = \frac{\psi\psi^\dagger}{||\psi||^2} \tag{183}$$

Here we can regard ψ as a column vector and $\psi^\dagger$ as a row vector.

• The density matrix for a particle in a unit norm pure state $\psi(x)$ is just the projection operator to the one-dimensional sub-space spanned by ψ, i.e., $\rho = \psi\psi^\dagger = |\psi\rangle\langle\psi|$. We may represent ρ in any orthonormal basis, (say the basis of eigenstates $|n\rangle$ of some observable A like energy, $A|n\rangle = \lambda_n|n\rangle$)

$$\rho_{mn} = \langle m|\psi\rangle\langle\psi|n\rangle = c_m c_n^* \tag{184}$$

where $|\psi\rangle = \sum_n c_n|n\rangle$ and c_n are the components of ψ in the basis $|n\rangle$. In particular, the diagonal elements $\rho_{nn} = |c_n|^2$ give the probabilities that upon a measurement of the observable A, the value λ_n is obtained and the system collapses to the state $|n\rangle$.

• In a pure state, we know the projections c_n of ψ on each basis vector. In a mixed state, the c_n's are not known with certainty. The normalization of ψ is the condition $\psi^\dagger\psi = 1$ or

$$1 = \langle\psi|\psi\rangle = \sum_n \langle\psi|n\rangle\langle n|\psi\rangle = \sum_n c_n^* c_n = \sum_n \rho_{nn} = \text{ tr } \rho. \tag{185}$$

• The fact that a state is pure implies that ρ is a projection operator $\rho^2 = \psi\psi^\dagger\psi\psi^\dagger = \psi\psi^\dagger = \rho$. In fact the density matrix of a pure state is a rank-1 projection operator[6]. In particular, ρ has

---

[6] A projection to the subspace spanned by the columns of a matrix A is given by $P_A = A(A^\dagger A)^{-1}A^\dagger$. The rank of a projection is the dimension of the subspace to which it projects. An introduction to projection matrices was given in the MP1 lecture notes.

one eigenvalue equal to one and all remaining eigenvalues equal to zero. The 1-eigenspace of ρ is spanned by the pure state ψ. In fact, a pure state can be defined as one whose density matrix is a rank one projection (that projects to the pure state!).

• In addition, for a pure state since $\rho^2 = \rho$ we also have $\text{tr}\,\rho^2 = \text{tr}\,\rho = 1$.

• We will see that mixed ensemble density matrices aren't projections and aren't rank one, moreover $\text{tr}\,\rho^2 < 1$ for a mixed ensemble.

9.2.3 More on mixed ensemble density matrices

• A mixed ensemble can be thought of as an ensemble of pure states, so that the c_n's are not precisely known but take different values for the various pure states in the ensemble. The elements of the density matrix may be obtained as an ensemble average (this ensemble average is different from the quantum mechanical averaging that we do to find the expectation value of an observable in a given pure state)

$$\rho_{mn} = \overline{c_m c_n^*} \tag{186}$$

Alternatively, the ensemble average can be thought of as an average over the density matrices of each of the pure states in the ensemble weighted by their probabilities (this is how we introduced the concept)

$$\rho = \sum_{i\,\in\,\text{ensemble}} p_i |\psi_i\rangle\langle\psi_i| \quad \text{with} \quad \sum_{i\,\in\,\text{ensemble}} p_i = 1. \tag{187}$$

Here $p_i \geq 0$ is the probability of occurrence of the pure state $|\psi_i\rangle$ in the ensemble.

• For example, suppose we have a beam of electrons with an isotropic distribution of spin projections (no preferred direction). Then we can describe the spin state of an electron in such an ensemble using a mixed state density matrix. Working in the basis in which S_z is diagonal, both the pure state density matrices $|\uparrow\rangle\langle\uparrow|$ and $|\downarrow\rangle\langle\downarrow|$ are equally likely. Thus the density matrix is

$$\rho = \frac{1}{2}|\uparrow\rangle\langle\uparrow| + \frac{1}{2}|\downarrow\rangle\langle\downarrow| = \frac{1}{2}\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix} + \frac{1}{2}\begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix} = \frac{1}{2}\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix} \tag{188}$$

Calculate the expectation values of the various components of spin and show that they vanish for this mixed state.

• The same beam with isotropic spin projections can also be described as a mixed ensemble where there are equal numbers of electrons with $S_x = \hbar/2$ as there are electrons with $S_x = -\hbar/2$. If $X_\pm$ are the normalized eigenvectors of S_x corresponding to these two eigenvalues, then the density matrix in the basis where S_x is diagonal is given by

$$\rho = \frac{1}{2}X_+ X_+^\dagger + \frac{1}{2}X_- X_-^\dagger \tag{189}$$

By expressing $X_\pm$ in terms of $\uparrow$ and $\downarrow$ show that this density matrix is the same as the above density matrix in the basis where S_z is diagonal.

• The density matrix of the mixed ensemble mentioned earlier with a third each of the population in the states $\uparrow, \downarrow$ and $\uparrow + \downarrow$ is

$$\rho_3 = \frac{1}{3}\uparrow\uparrow^\dagger + \frac{1}{3}\downarrow\downarrow^\dagger + \frac{1}{3}\frac{1}{\sqrt{2}}(\uparrow + \downarrow)\frac{1}{\sqrt{2}}(\uparrow^\dagger + \downarrow^\dagger). \tag{190}$$

• Example: Suppose we have a system in thermal equilibrium at temperature T. Then we may describe its properties using a thermal ensemble with (mixed) thermal density matrix. In the basis of energy eigenstates $H|n\rangle = E_n|n\rangle$, the thermal density matrix is

$$\rho = A \sum_{n=0}^{\infty} e^{-E_n/kT} |n\rangle\langle n| \quad \text{where the normalization constant} \quad A^{-1} = \sum_n e^{-E_n/kT} \tag{191}$$

is fixed by the requirement that $\text{tr}\ \rho = 1$. A thermal density matrix (in the energy basis) corresponds to an ensemble where each pure state $|n\rangle$ of definite energy is present with a probability proportional to the Boltzmann factor $e^{-E_n/kT}$. High energy pure states occur less frequently in the ensemble compared to lower energy pure states. The thermal density matrix at temperature T is $\rho = Ae^{-H/kT}$. The diagonal entries of the density matrix in the energy basis are the relative occupation numbers of the various energy levels of the system at that temperature. So a gas of hydrogen atoms in equilibrium at temperature T would contain more atoms in the ground state than in highly excited states.

• For a mixed ensemble, ρ is neither a rank-1 matrix nor can it be a projection matrix.

• If a density matrix is a projection operator, then it must be the density matrix of a pure ensemble. Proof: Suppose that a density matrix is a projection $\rho^2 = \rho$. So it projects to a subspace $V \subseteq \mathcal{H}$. Take an o.n. basis $\psi_1, \cdots \psi_n$ for V. Then $\rho = \psi_0\psi_0^\dagger + \cdots + \psi_n\psi_n^\dagger$. But then $1 = \text{tr}\ \rho = n$, this implies $n = 1$. So $\rho = \psi_1\psi_1^\dagger$ must be a pure density matrix and necessarily of rank one.

• If ρ is rank 1, does it have to be pure? Ans: Yes. Why? If ρ is rank one, then $\rho = ab^\dagger$ for some vectors a and b. But $\rho^\dagger = \rho$ so $ab^\dagger = ba^\dagger$. Moreover $\text{tr}\ \rho = 1$, so $b^\dagger a = 1$, so $b = bb^\dagger a^\dagger = |b|^2 a$. Therefore b is parallel to a. Now define $c = |b|a$, then $\rho = cc^\dagger$, and therefore, ρ is a pure state density matrix.

• Let us see these facts in an example. Suppose a mixed state is formed from an ensemble of *orthonormal pure states* $|n\rangle$

$$\rho = \sum_{n\ \in\ \text{ensemble}} P_n |n\rangle\langle n|, \qquad \langle n|m\rangle = \delta_{mn}. \tag{192}$$

Then check that $\rho^2 = \sum_n P_n^2 |n\rangle\langle n|$. But for a mixed state, $P_n < 1$ for n in the ensemble, so $P_n^2 < P_n$ for every non-zero P_n. Thus $\rho^2 \neq \rho$ for this mixed state. The rank of a mixed state density matrix in the above example is the number of orthonormal pure states in the ensemble for which $P_n \neq 0$. This is always two or more for a mixed ensemble, and equal to one only for a pure ensemble.

• $\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|$ with $0 \leq p_i \leq 1$ is a non-negative operator. Here ψ_i are a collection of pure states that are present in the ensemble. They need not be a basis and need not be orthogonal. To show positivity, we find the expectation value of ρ in any state ϕ

$$\langle\phi|\rho|\phi\rangle = \sum_i p_i \langle\phi|\psi_i\rangle\langle\psi_i|\phi\rangle = \sum_i p_i |\langle\phi|\psi_i\rangle|^2 \geq 0 \tag{193}$$

So all expectation values of ρ are non-negative. So all eigenvalues of ρ are non-negative.

• Since ρ is hermitian we can go to an orthonormal basis $|n\rangle$ where it is diagonal.

$$\rho = \sum_n \rho_n |n\rangle\langle n| \tag{194}$$

Since ρ is a non-negative operator, its eigenvalues ρ_n must be non-negative. And moreover, $\text{tr}\ \rho = \sum_n \rho_n = 1$. So $0 \leq \rho_n \leq 1$.

- It is easy to compute ρ^2 in its eigen-basis and we get

$$\rho^2 = \sum_n \rho_n^2 |n\rangle\langle n| \tag{195}$$

Since $0 \leq \rho_n \leq 1$, we must have $\rho_n^2 \leq \rho_n$. Moreover, in a mixed state, at least one of these inequalities is strict $\rho_j^2 < \rho_j$. It follows that $\rho^2 \neq \rho$ for a mixed ensemble. So $\rho^2 = \rho$ iff ρ is a pure state density matrix. In particular, $\text{tr}\ \rho^2 < \text{tr}\ \rho = 1$ for a mixed ensemble. So $\text{tr}\ \rho^2 < 1$ for a mixed ensemble and $\text{tr}\ \rho^2 = 1$ only for a pure ensemble. So $\text{tr}\ \rho^2 = 1$ is a test of purity.

- von Neumann entropy is another interesting quantity. It can also be used as another test of purity.

$$S = -k\ \text{tr}\ \rho \log \rho \tag{196}$$

where k is Boltzmann's constant. It is easiest to compute S in the basis where $\rho = \sum_n \rho_n |n\rangle\langle n|$ is diagonal and we get

$$S = -k \sum_n \rho_n \log \rho_n \tag{197}$$

From $0 \leq \rho_n \leq 1$ it follows that $S \geq 0$.

- We see that $S = 0$ for a pure ensemble since for a pure state $\rho = (10|00)$ is a block matrix with only one non-zero entry in any basis whose first basis vector is the pure state wave function. In fact $S = 0$ if and only if the state is pure. For an impure state at-least two ρ_n's must be non-zero and less than one so $S > 0$ for a mixed ensemble. So $S = 0$ is a test of purity.

- It is interesting to find a density matrix of maximal entropy for a system with a finite dimensional Hilbert space. Working in the basis where ρ is diagonal, we must maximize S while holding $\sum_n \rho_n = 1$. Thus we must extremize the quantity

$$T = -k \sum_n \rho_n \log \rho_n + \lambda \left(\sum_n \rho_n - 1 \right) \tag{198}$$

Imposing $\frac{\partial T}{\partial \rho_m} = 0$ for all m we get

$$\frac{\partial T}{\partial \rho_m} = -k \sum_n \left(\delta_{mn} \log \rho_n + \frac{\rho_n}{\rho_n} \delta_{mn} \right) + \lambda \sum_n \delta_{mn} = 0 \quad \Rightarrow \quad \log \rho_m = \frac{\lambda}{k} - 1. \tag{199}$$

Thus all the ρ_m must be equal. Since $\sum_n \rho_n = 1$ in the case of a finite dimensional Hilbert space, this implies

$$\rho_m = \frac{1}{\dim \mathcal{H}} \equiv \frac{1}{d}. \tag{200}$$

Thus the maximum value of the entropy is

$$S_{\max} = k \log d. \tag{201}$$

9.2.4 Time evolution of density matrix

• Finally, the time-evolution of a pure state can be expressed in terms of its density matrix

$$i\hbar\dot{\rho} = i\hbar\dot{\psi}\psi^\dagger + i\hbar\psi\dot{\psi}^\dagger = H\psi\psi^\dagger + \psi\psi^\dagger H = [H,\rho]. \tag{202}$$

Alternatively, we can work in a basis

$$\begin{aligned} i\hbar\partial_t|\psi\rangle &= H|\psi\rangle \;\Rightarrow\; i\hbar\partial_t\langle n|\psi\rangle = i\hbar\partial_t c_n = \langle n|H|\psi\rangle = \sum_m \langle n|H|m\rangle\langle m|\psi\rangle = \sum_m H_{nm}c_m \\ \Rightarrow\; i\hbar\partial_t\rho_{mn} &= i\hbar(c_m\partial_t c_n^* + c_n^*\partial_t c_m) = \sum_p c_n^* H_{mp}c_p - \sum_p c_m H_{pn} c_p^* = \sum_p H_{mp}\rho_{pn} - H_{pn}\rho_{mp} = [H,\rho]_{mn} \end{aligned}$$

So the Schrödinger equation for a pure state becomes $i\hbar\frac{\partial\rho}{\partial t} = [H,\rho]$.

• Since each of the pure states in the ensemble evolve according to the Schrodinger equation, the ensemble average mixed density matrix $\rho = \sum_i p_i\rho_i$ also evolves in the same way

$$i\hbar\frac{\partial\rho}{\partial t} = i\hbar\sum_i p_i\frac{\partial\rho_i}{\partial t} = \sum_i p_i[H,\rho_i] = \left[H, \sum_i p_i\rho_i\right] = [H,\rho]. \tag{203}$$

• We recognize that this evolution equation for the density matrix is the quantum mechanical analog of the Liouville equation for the evolution of the phase space density. $\frac{1}{i\hbar}[H,\rho]$ is replaced by the Poisson bracket $\{H,\rho\}$ in classical mechanics.

• The evolution equation for the density matrix also bears a resemblance to the Heisenberg equation of motion of quantum mechanics.