

Contents

1	Introductory Remarks	3
2	Review of classical mechanics of point particles	5
2.1	Newton's force law, state of a classical particle	5
2.2	Energy, Angular momentum, conserved quantities	6
2.3	Hamiltonian and Hamilton's equations	7
2.4	Principle of extremal action and Lagrangian	9
2.5	Hamiltonian from Lagrangian	12
3	Basic concepts of waves	13
3.1	Wave packet and its width	15
3.2	Motion of a wave packet and group velocity	16
4	Quanta of light: particle-like properties of radiation	18
4.1	Black body radiation and Planck hypothesis	19
4.2	Photoelectric Effect	20
4.3	Compton Scattering Effect	22
5	Quantum and wave properties of matter (electrons, molecules and atoms)	24
5.1	Discrete atomic spectra and stability of atom	24
5.2	Bohr model of the atom	25
5.3	Franck-Hertz experiment	27
5.4	Raman scattering	27
5.5	de Broglie hypothesis on matter waves	28
5.6	Electron Diffraction: Davisson-Germer-Thompson experiment	30
5.7	Double slit interference experiment: wave function as probability amplitude	31
5.8	Wave-Particle Duality and Complementarity Principle	34
6	Schrödinger equation, formulation and postulates of quantum mechanics	34
6.1	Obtaining Schrödinger's equation from matter wave hypothesis	34
6.2	Wavefunction, probability density and current	36
6.3	Superposition principle and vector space of states	38
6.4	Dirac Delta function	39
6.5	Average (expectation) values of observables in a state	41
6.6	Ehrenfest's theorem	42
6.7	Hilbert space of states, Dirac bra-ket notation	43
6.8	Position eigenstates	44
6.9	Momentum eigenstates	46
6.10	Heisenberg commutation relation and uncertainty	47
6.11	Hermiticity of observables $\hat{x}, \hat{p}, \hat{H}$	48
6.12	Collapse of the wavefunction/measurement and probability postulate	49
6.13	Energy eigenstates: Time-independent Schrödinger equation	49
6.14	Summary of postulates of quantum mechanics	51

6.15	Unitarity of Schrödinger time evolution	52
7	Square well, Free particle and Harmonic oscillator	53
7.1	Particle in a box : Infinite square well	53
7.2	Free particle and gaussian wave packet	57
7.3	Harmonic oscillator	61
7.3.1	Variational Approach	61
7.3.2	Algebraic approach via creation and annihilation (ladder) operators	63
7.3.3	Direct solution of Schrödinger equation for SHO	65
8	Some general properties of one dimensional quantum mechanical models	67
8.1	Parity	67
8.2	Ehrenfest theorem and evolution of expectation values	68
8.3	Symmetries and conservation laws: Translation invariance	69
8.4	Energy-time uncertainty relation	70
8.5	Schrödinger vs Heisenberg pictures and equations of motion	71
8.6	Absence of degenerate bound states in one dimension	73
8.7	Complete set of commuting observables	74
9	Models with bound and scattering states, tunneling	75
9.1	Bound state of the delta function potential	75
9.2	Reflection and Transmission coefficients	76
9.3	Scattering states for a delta function potential	77
9.4	Finite square well	78
9.5	Scattering against a rectangular barrier/well and tunneling	80
10	Some simple quantum mechanical models in 2 dimensions	82
10.1	Is every energy eigenfunction separable (factorizable as $\psi(x, y) = X(x)Y(y)$)?	82
11	Quantum mechanics in three dimensions: central potentials	83
11.1	Free particle in three dimensions	83
11.2	Basic properties of angular momentum	84
11.2.1	Angular momentum and rotations	86
11.3	Free particle hamiltonian as a hermitian operator in spherical coordinates	87
11.3.1	Positivity of free particle hamiltonian, radial and angular kinetic energies	88
11.4	Conservation of angular momentum for a particle in a central potential	89
11.5	Separation of variables for particle in central potential	89
11.6	Spectrum of L^2 and L_z : spherical harmonics	90
11.7	Eigenvalues of L^2 and L_z by ladder operators	92
11.8	Free particle radial wave function	93
11.9	Radial eigenfunctions for particle in a spherical box/well	94
11.10	Radial eigenfunctions for Hydrogen atom	95

1 Introductory Remarks

- These notes (<http://www.cmi.ac.in/~govind/teaching/qm1-s11>) are very sketchy and are no substitute either for text books or attendance and taking notes at lectures. There are many books on quantum mechanics, some are listed here

“Modern Physics” phenomena necessitating quantum mechanics

- Arthur Beiser, Concepts of Modern Physics
- H. S. Mani and G. K. Mehta, Introduction to Modern Physics

Introductory quantum mechanics text books

- D. J. Griffiths, Introduction to quantum mechanics
- R. Liboff, Introductory quantum mechanics
- Mathews, P. M & Venkatesan, K., A Textbook of quantum mechanics.
- L. I. Schiff, Quantum mechanics
- R. Shankar, Principles of quantum mechanics
- E. Merzbacher, Quantum mechanics

Quantum Mechanics at a deeper/more advanced level or by the masters

- P. A. M. Dirac, The principles of quantum mechanics
- D. Bohm, Quantum Theory
- Landau and Lifshitz, Quantum mechanics Non-relativistic theory
- Pauli, Wave mechanics (and Pauli, General principles of quantum mechanics, trans. Achuthan and Venkatesan)
- R. P. Feynman, Lectures in Physics, Vol 3
- Esposito, Marmo & Sudarshan, From Classical to Quantum Mechanics.
- J. Schwinger (ed. Berthold-Georg Englert), Quantum Mechanics: Symbolism of Atomic Measurements
- Wilson & Pauling, Introduction to quantum mechanics: with applications to chemistry.
- Sakurai, Modern quantum mechanics
- Albert Messiah, Quantum mechanics
- Claude Cohen-Tannoudji, Bernard Diu, Frank Laloe, Quantum Mechanics

- Quantum mechanics is the best available framework with which we currently describe, understand and predict properties of many physical systems. This is especially so for microscopic systems (e.g. atoms, molecules, nuclei, protons) but also applies to larger systems (aggregates of atoms) like the properties of a ferromagnet, metal/semi-conductor or super-conductor or astrophysical systems such as a neutron star.

- Quantum mechanics is a collection of principles and rules that physicists reluctantly adopted when it was found that classical physics (based on Newtonian mechanics, Maxwell electrodynamics and thermodynamics) failed to correctly account for certain aspects of atomic systems. (eg. discrete atomic spectra, stability of the atom, Photoelectric effect, black body radiation and diffraction of electrons.)

- A key conundrum was that electrons, atoms and light displayed both particle-like and wave-like properties in different situations. This could not be accommodated classically. Quantum mechanics essentially solved this problem in a radical and unexpected manner, which we are still getting to grips with.

- The discovery of Quantum Mechanics was the work of many experimental and theoretical physicists in the early part of the 20th century. The theorists Schrödinger, Heisenberg, Born, Pauli and Dirac were prominent among those who clarified the basic principles of quantum

mechanics by around 1928¹. However, this came in the midst of a series of experimental discoveries and partial explanations, some of which are mentioned in this time line

1800-1887 Young, Faraday, Maxwell, Hertz wave nature of light and electromagnetic radiation.

late 1800s Discrete emission spectrum of hydrogen and other gases, Balmer formula 1885.

1897 J J Thomson discovery of particle-like electron: showed atom is not indivisible

1901 Blackbody radiation: Planck's and Einstein's quantum hypothesis (interpreted by Einstein as a packet of light)

1905 Einstein's explanation of photoelectric effect via photon

1911 Rutherford discovery of heavy small positively charged nucleus via α (He_2^4) scattering

1913 Bohr model of the atom

1922 Compton scattering (γ - e^-) confirms particle nature of photon

1924 Pauli exclusion principle and stability of matter

1925 de Broglie hypothesis on wave nature of electron

1927 Davisson-Germer-Thomson experiment on electron diffraction

1928 Krishnan & Raman: inelastic scattering of light by molecules

- However, despite the development of quantum mechanics, classical mechanics remains an adequate framework for describing many phenomena such as projectile, planetary and fluid motion where 'quantum effects' are small.

- It is possible to obtain the equations of classical mechanics as an approximation to those of quantum mechanics in special circumstances (e.g. when the *action* is large compared to Planck's constant $h \approx 6 \times 10^{-34} Js$, to be clarified later). This is analogous to how ray optics is an approximation to wave optics when the wavelength of light is small compared to the other length scales involved.

- So quantum mechanics can be regarded as modifying Newton's equations. However, often these modifications are so great that it is necessary to use an entirely new vocabulary and set of principles and not rely too much on the framework of Newtonian mechanics. It is necessary to develop a new intuition for the quantum behavior of atomic systems, which is far removed from our daily experience.

- The principles of quantum mechanics have been used to model numerous physical systems (atoms, light, molecules, Bose-Einstein condensation etc) and in all cases, its predictions have been in line with experimental findings.

- However, there are some questions which quantum mechanics does not answer (e.g. trajectory of an electron in Hydrogen atom). Classical mechanics provides answers to such questions, but they turn out to be wrong when applied to atomic physics. On the other hand, the predictions of quantum mechanics are often of a statistical nature (what is the

¹Some mathematicians (e.g., Weyl and von Neumann) also contributed slightly later

average distance of the electron from the proton in the hydrogen atom if the experiment is repeated many times?).

- This restriction ('indeterminism' vis-a-vis classical mechanics) on the types of questions that one can ask and hope to answer in quantum mechanics has troubled many physicists. But no attempt to improve upon quantum mechanics has so far succeeded.
- There are still many physical situations where the implications of quantum mechanics have not been adequately understood due to a lack of experimental, conceptual or mathematical tools (e.g. quantum effects in sub-atomic physics or gravity, mesoscopic systems and the problem of measurement). So there is still a lot to do!
- Quantum mechanics provides conceptual and technical tools to construct new devices (e.g. scanning tunneling microscopes, lasers, nuclear reactors, x-ray and PET scans and perhaps even quantum computers). The harnessing of quantum effects for engineering novel materials is a thriving area.
- Quantum mechanics is the basis of atomic and molecular chemistry.
- Newton's equations are a system of non-linear ordinary differential equations, while the Schrödinger equation of quantum mechanics is a linear partial differential equation. So there are aspects of quantum mechanics that are apparently simpler than Newtonian mechanics.
- Quantum mechanics introduces a new way of thinking which has been used to solve or provide insight into several outstanding mathematical problems (e.g. group representations: of semi-simple groups in the early days of quantum mechanics and more recently current algebras and the orbit method; classification of knots using quantum Chern-Simons theory. Several areas of mathematics have grown/benefitted from attempts to understand aspects of quantum mechanics (e.g. probability from the Feynman path integral, functional analysis from the stability of the atom). Indeed, most major branches of mathematics have been enriched by ideas from quantum mechanics (algebra: operator algebras, deformation quantization, quantum integrable systems; geometry: gauge theory, supersymmetric quantum mechanics, non-commutative geometry; analysis: functional analysis from quantum field theory, number theory & arithmetic: quantum chaos and random matrix theory, perturbative quantum field theory, dynamical systems: renormalization)
- Our primary concern in this course is to learn what quantum mechanics is, why it was introduced, its relation to classical mechanics and how to use it to understand atomic and other physical systems. Quantum mechanics has many interesting and counter intuitive phenomena such as tunneling through a barrier and superconductivity, which make it exciting to learn. The basic formulation of quantum mechanics and its application to simple systems can be made accessible to students with some background in classical mechanics, calculus and linear algebra.

2 Review of classical mechanics of point particles

2.1 Newton's force law, state of a classical particle

- A point particle moving along a wire in the shape of a line or circle has one *degree of freedom*, namely its position (coordinate) along the wire. A point-like molecule in the air has three degrees of freedom, e.g., its cartesian coordinates with respect to a chosen set of axes.

N such molecules have $3N$ degrees of freedom. The set of possible locations of the molecules forms the *configuration space* of the system. For a particle on a plane, the configuration space is \mathbb{R}^2 , two-dimensional Euclidean space.

- The zeroth law of classical (Newtonian) mechanics can be regarded as saying that the trajectory $x(t)$ of a particle is a (twice) differentiable function of time. This is a law that applies to tennis balls, planets etc. But it fails for Brownian motion (movement of pollen grains in water). It also fails in quantum mechanics. To the extent that one can discuss paths of particles in quantum mechanics, they need not be differentiable functions of time. Indeed, the typical path contributing to the motion of a free particle is only one-half differentiable in quantum mechanics.

- Newton's force law says that the trajectory $\vec{r}(t)$ of a particle is determined by the force acting on it

$$m\ddot{\vec{r}} = \vec{F} \quad \text{or} \quad \dot{\vec{p}} = \vec{F}, \quad \text{or} \quad m\ddot{x}_i = F_i. \quad (1)$$

Here the momentum $\vec{p} = m\vec{v} = m\dot{\vec{r}}$. We also denote $x_1 = x, x_2 = y, x_3 = z$, so x_i are the coordinates. Sometimes the coordinates are denoted q_i instead of x_i .

- In many interesting cases, the force can be written as the negative gradient of a potential energy²

$$\vec{F} = -\nabla V = -\frac{\partial V}{\partial x}\hat{x} + \frac{\partial V}{\partial y}\hat{y} + \frac{\partial V}{\partial z}\hat{z} \quad (2)$$

E.g. $V = mgz$ for the gravitational potential energy and so $\vec{F} = -mg\hat{z}$ points downwards. In this case, Newton's second law is

$$\dot{\vec{p}} = -\nabla V \quad \text{or} \quad m\ddot{x}_i = -\frac{\partial V}{\partial x_i}. \quad (3)$$

- Since Newton's equation is second order in time, to solve it we need two sets of initial conditions, the initial position vector and velocity vector $(\vec{r}, \dot{\vec{r}})$ or the initial position and momentum.

- So the *state* of the particle (and its future trajectory) is specified by giving its instantaneous position *and* momentum. The set of possible instantaneous states of the particle is called its phase space. For a particle moving in three dimensional space, its configuration space is \mathbb{R}^3 and its phase space is \mathbb{R}^6

2.2 Energy, Angular momentum, conserved quantities

- The energy of a particle moving in three dimensional space (say under the gravitational force) is a sum of kinetic and potential energies

$$E = T + V = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) + V(x, y, z) = \frac{1}{2m}(p_x^2 + p_y^2 + p_z^2) + V(x, y, z) \quad (4)$$

$\dot{x} = \frac{dx}{dt}$ is the x -component of velocity. We should think of the energy as a function of coordinates and momenta $E(x, y, z, p_x, p_y, p_z)$. So energy is a function on the phase space. Thought of this way, it is also called the hamiltonian $H = E$.

²Force points in the direction of greatest decrease in potential energy

- Newton's equation implies that the energy is a constant of motion³

$$\dot{E} = m \sum_i \dot{x}_i \ddot{x}_i + \sum_i \frac{\partial V}{\partial x_i} \dot{x}_i = 0. \quad (5)$$

We say the energy is conserved in time or that it is a conserved quantity.

- Conserved quantities are useful. They help us solve Newton's equation for the trajectory. For example, in the case of one degree of freedom, we can integrate once and get an (implicit) expression for $x(t)$:

$$E = \frac{1}{2}m\dot{x}^2 + V(x) \Rightarrow \frac{dx}{dt} = \sqrt{\frac{2}{m}(E - V(x))} \Rightarrow t - t_0 = \int_{x_0}^x \frac{dx'}{\sqrt{\frac{2}{m}(E - V(x'))}} \quad (6)$$

- If there is no force, then each of the components of momentum is conserved, since $\dot{\vec{p}} = \vec{F} = 0$ (this is Newton's first law). If the force only acts downwards, then the horizontal components of momentum p_x, p_y are conserved.
- The angular momentum of a particle about a fixed point (origin) is $\vec{L} = \vec{r} \times \vec{p}$, where \vec{r} is the position vector of the particle from the origin. In components

$$L_x = yp_z - zp_y, \quad L_y = zp_x - xp_z, \quad L_z = xp_y - yp_x. \quad (7)$$

- Newton's force law then implies that the rate of change of angular momentum is the torque:

$$\dot{\vec{L}} = \dot{\vec{r}} \times \vec{p} + \vec{r} \times \dot{\vec{p}} = \frac{1}{m} \vec{p} \times \vec{p} + \vec{r} \times \vec{F} = \vec{r} \times \vec{F} \equiv \vec{\tau} \quad (8)$$

For example, if we have a projectile moving under the vertical gravitational force, then the torque must be in the horizontal plane. So the vertical component of angular momentum $L_z = xp_y - yp_x$ must be conserved. Since p_x and p_y are also conserved, we conclude that the trajectory $(x, y, z)(t)$ must be such that its projection on the horizontal plane is a straight line $L_z = xp_y - yp_x$. Again, knowledge of conserved quantities allowed us to clarify the trajectory.

2.3 Hamiltonian and Hamilton's equations

- Newton's equation of motion $F = ma$ can usefully be recast in **Hamiltonian form**, using the Hamiltonian or energy

$$H(x, p) = \frac{p^2}{2m} + V(x) \quad (9)$$

First write $m\ddot{x} = \frac{dV}{dx}$ as the pair of equations: $p = m\dot{x}$, (which is merely the definition of momentum) and $\dot{p} = -\frac{\partial V}{\partial x}$. Then observe that these equations could also be written as

$$\dot{x} = \frac{\partial H}{\partial p} = \frac{p}{m}, \quad \text{and} \quad \dot{p} = -\frac{\partial H}{\partial x} = -\frac{dV}{dx} \quad (10)$$

³This is the case for conservative forces which are those that may be expressed as the gradient of a potential.

We say that p is the momentum *conjugate* to x . For several degrees of freedom, Hamilton's equations take the form

$$\dot{x}_i = \frac{\partial H}{\partial p_i} \quad \text{and} \quad \dot{p}_i = -\frac{\partial H}{\partial x_i} \quad (11)$$

Hamilton's equations are 1st order in time. So there are twice as many of them compared to Newton's equations. Hamilton's equations treat position and momentum on an even footing (except for a sign).

- Moreover, it is easier to write down Hamilton's equations, since we only need to know one function (the energy). By contrast, to write Newton's equations we need to know the force, which is a vector.

- Hamilton's equations take the same form in all coordinate systems⁴. By contrast, Newton's equations $m\ddot{x}_i = -\partial_i V$ look quite different in curvilinear coordinates.

- Let us illustrate this with the example of a particle free to move on a plane without any force acting on it. Newton's equations are then $m\ddot{x} = 0$ and $m\ddot{y} = 0$ and the energy is $E = \frac{m}{2}(\dot{x}^2 + \dot{y}^2)$. The coordinates are x, y and the conjugate momenta are $p_x = m\dot{x}$ and $p_y = m\dot{y}$.

- Now let us transform to polar coordinates $x = r \cos \phi$ and $y = r \sin \phi$. Then the components of velocity are $v = (\dot{x}, \dot{y})$

$$\dot{x} = \dot{r} \cos \phi - r \sin \phi \dot{\phi}, \quad \dot{y} = \dot{r} \sin \phi + r \cos \phi \dot{\phi}. \quad (12)$$

So $\dot{x}^2 + \dot{y}^2 = \dot{r}^2 + r^2 \dot{\phi}^2$. So the energy is

$$E = H = \frac{m}{2} (\dot{r}^2 + r^2 \dot{\phi}^2) \quad (13)$$

Now we define the momentum conjugate to r as the component of linear momentum in the $\hat{r} = (\cos \phi, \sin \phi)$ direction.

$$p_r = \frac{\vec{p} \cdot \vec{r}}{r} = \hat{r} \cdot \vec{p} = m \hat{r} \cdot \vec{v} = m \dot{r} \quad (14)$$

We define the momentum conjugate to ϕ as the angular momentum corresponding to a displacement of ϕ holding r fixed. The moment arm is r and the velocity corresponding to a shift in ϕ is $r\dot{\phi}$ so $p_\phi = r m r \dot{\phi}$. In other words, p_ϕ is just the angular momentum about the vertical axis $L_z = x p_y - y p_x = m(x\dot{y} - y\dot{x}) = m r^2 \dot{\phi}$.

- Note that even though we call p_ϕ the momentum conjugate to ϕ , it does not have dimensions of momentum, it actually has dimensions of angular momentum. In the same way we will refer to r, ϕ as coordinates, though they don't both have dimensions of length.

- The energy is now expressed in polar coordinates and momenta

$$E = H = \frac{p_r^2}{2m} + \frac{p_\phi^2}{2mr^2} \quad (15)$$

- Now Newton's equations $\ddot{x} = \ddot{y} = 0$ can be expressed in polar coordinates. One can check that

$$m\ddot{r} = m r \dot{\phi}^2 \quad (16)$$

⁴And in all systems of coordinates and momenta that are related by canonical transformations.

Notice that Newton's equation in polar coordinates has a different form to that in cartesian coordinates. If we had naively regarded $m\ddot{r}$ as an acceleration and observed that there is no force on the particle, we would have got a wrong equation. Though there is no force on the particle, neither \ddot{r} nor $\ddot{\phi}$ is zero. Of course, we sometimes say that $mr\dot{\phi}^2$ is a centripetal force, but this is not really a force, it is just another term in the acceleration due to the curvilinear coordinates.

- On the other hand, let us write down Hamilton's equations, which we claimed take the same form in all coordinate systems

$$\dot{\phi} = \frac{\partial H}{\partial p_\phi} = \frac{p_\phi}{mr^2}, \quad \dot{r} = \frac{\partial H}{\partial p_r} = p_r/m, \quad \dot{p}_\phi = -\frac{\partial H}{\partial \phi} = 0, \quad \dot{p}_r = -\frac{\partial H}{\partial r} = \frac{p_\phi^2}{mr^3} \quad (17)$$

- Let us check if these equations agree with Newton's equations. The first two of Hamilton's equations just reproduce the definitions of p_r and p_ϕ .

- The third equation says that the angular momentum $p_\phi = L_z$ is conserved.

- The hamiltonian formulation allows us to identify conserved quantities easily. A cyclic coordinate q is one that does not appear in the hamiltonian so that $\frac{\partial H}{\partial q} = 0$. Its conjugate momentum is then automatically conserved $\dot{p} = -\frac{\partial H}{\partial q} = 0$. In our case, ϕ does not appear in the hamiltonian, so the conjugate momentum (angular momentum) is conserved. This was not quite obvious from Newton's equations, though we knew it.

- The last of hamilton's equations says $m\ddot{r} = mr\dot{\phi}^2$, reproducing Newton's force law.

- So we have checked that in this case, Hamilton's equations $\dot{x}_i = \frac{\partial H}{\partial p_i}$ and $\dot{p}_i = -\frac{\partial H}{\partial x_i}$ take the same form in Cartesian as well as polar coordinates. In other words we can choose $x_i = (x, y)$ or (r, ϕ) in the preceding expressions.

- Our main reason for rewriting Newton's equations in Hamiltonian form, is that they allow a very simple passage to the Heisenberg equations of quantum mechanics.

- Furthermore, the Hamiltonian plays a central role in the Schrödinger equation of quantum mechanics.

- However, the problem of identifying the appropriate momenta conjugate to the coordinates is not solved. In the above example, we could motivate the definition of p_r, p_ϕ based on simple arguments. But there is a systematic procedure to find the conjugate momenta. This is called the Lagrangian formulation, which we turn to next. Needless to say, the Lagrangian formulation does a lot more than give a prescription to find the momenta conjugate to a chosen set of coordinates.

2.4 Principle of extremal action and Lagrangian

- Let us define the Lagrange function (*Lagrangian*) $L(q, \dot{q}) = T - V(q) = \frac{1}{2}m\dot{q}^2 - V(q)$. Then we observe that Lagrange's equation

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{q}} = \frac{\partial L}{\partial q} \quad \Rightarrow \quad m\ddot{q} = -\frac{\partial V}{\partial q} \quad (18)$$

reproduces Newton's equation. Here we denote coordinates by q rather than x to emphasize that it need not be a cartesian coordinate. Let us briefly describe how Lagrange's equations arise.

- The Lagrangian formulation is a useful alternative formulation of Newton's equations. Here we consider the problem of determining the classical trajectory that a particle must take if it was at x_i at t_i and x_f at t_f . Instead of specifying the initial velocity, we give the initial and final positions at these times. Which among all the paths that connect these points, is the one that solves Newton's equation? We seek a variational principle, called the principle of extremal action (sometimes called least action) that should determine the classical trajectory.
- The idea is to find an *action functional* $S[x]$, which is a function of paths connecting (t_i, x_i) and (t_f, x_f) such that the classical trajectory is a local extremum of $S[x(t)]$. For each path, $S[x(t)]$ is just a (real) number. In other words, if we modify the classical trajectory $x_{cl}(t) + \delta x(t)$, then the action $S[x]$ must remain stationary to 1st order in $\delta x(t)$.
- Let us mention a practical use of such a principle of least action. Suppose we want to find the trajectories of a complicated mechanical system (bodies interacting by some forces). This is usually very difficult to do exactly, since we need to solve Newton's equations, which are non-linear differential equations. But suppose several research workers A, B, C have by their own methods found what they believe to be good approximate solutions for the trajectory $x^A(t), x^B(t), x^C(t)$. How do we judge which among these proposed paths is the best? A variational principle allows you to pick the path that is closest to the true trajectory. Indeed, all we have to do is to calculate the actions for each of the paths $S[x^A], S[x^B], \dots$, and pick the path for which the action is minimal⁵. Note that none of the paths given by A, B, C may be exact solutions of Newton's equations, yet the action allows us to assign a 'figure of merit' to each of the paths and pick the best one.
- The *action* that does this job is defined as (for simplicity, for one degree of freedom)

$$S[x] = \int_{t_i}^{t_f} L(x, \dot{x}) dt \quad (19)$$

where L is the *Lagrangian*. For the simple case discussed above,

$$L(x, \dot{x}) = T - V = \frac{1}{2}m\dot{x}^2 - V(x) \quad (20)$$

Using the calculus of variations, one finds that the condition for $S[x]$ to be extremal is

$$\frac{\partial L}{\partial x} = \frac{d}{dt} \frac{\partial L}{\partial \dot{x}} \quad (21)$$

For n degrees of freedom, one gets a system of n equations, called the Euler-Lagrange equations or Lagrange's equations or simply equations of motion

$$\frac{\partial L}{\partial x_i} = \frac{d}{dt} \frac{\partial L}{\partial \dot{x}_i}, \quad i = 1, 2, \dots, n. \quad (22)$$

Moreover, the momentum conjugate to the coordinate x_i is defined as

$$p_i = \frac{\partial L}{\partial \dot{x}_i} \quad (23)$$

⁵We are oversimplifying, the trajectory that solves Newton's equations is the one for which the action is extremal, not necessarily minimal. To practically implement this procedure the paths given by A, B, C should be related through a continuous variation of a parameter, then we would graphically plot the action as a function of A, B, C in increasing order of this parameter and look for signs of an extremum to select the best path among the options.

Now suppose we have a particle moving in the potential $V(x_i)$. For the above Lagrangian, we can check that in cartesian coordinates, p_i are the usual momenta and Lagrange's equations are just Newton's equations

$$p_i = \frac{\partial L}{\partial \dot{x}_i} = m\dot{x}_i, \quad \text{and} \quad \dot{p}_i = -\frac{\partial V}{\partial x_i}. \quad (24)$$

The Lagrangian scheme has the advantage of generating all the equations of motion from a single function. Moreover, the Euler-Lagrange equations may be literally carried over to any coordinate system, so x_i need not be Cartesian coordinates. To emphasize this, one often calls the x_i generalized coordinates and denotes them by q_i instead of x_i . The momentum conjugate to q_i is conventionally denoted $p_i = \frac{\partial L}{\partial \dot{q}_i}$. There are as many generalized coordinates q_i as there are degrees of freedom. So for a pair of particles in a room, there would be six generalized coordinates q_1, \dots, q_6 .

- A coordinate q that does not appear in the Lagrangian $\frac{\partial L}{\partial q} = 0$, is called a cyclic coordinate. The corresponding momentum is conserved since $\dot{p} = \frac{d}{dt} \frac{\partial L}{\partial \dot{q}} = \frac{\partial L}{\partial q} = 0$.
- It is important to bear in mind that the Lagrangian $L(q, \dot{q})$ is a function of the coordinates q and velocities \dot{q} , and that the momentum p is a derived concept.
- The example of a harmonic oscillator (particle connected to a spring of force constant k). Here the potential $V(x) = \frac{1}{2}kx^2$ leads to the restoring force $-kx$, so

$$L = \frac{1}{2}m\dot{x}^2 - \frac{1}{2}kx^2 \quad (25)$$

and Lagrange's equation reproduces Newton's equation

$$m\ddot{x} = -kx. \quad (26)$$

- For the free particle on a plane that we studied, in polar coordinates

$$L = T = \frac{m}{2}(\dot{r}^2 + r^2\dot{\theta}^2) \quad (27)$$

The momenta conjugate to (r, θ) are

$$p_r = \frac{\partial L}{\partial \dot{r}} = m\dot{r}, \quad p_\theta = \frac{\partial L}{\partial \dot{\theta}} = mr^2\dot{\theta} \quad (28)$$

They coincide with the radial component of linear momentum and the angular momentum as we had argued earlier. Moreover, Lagrange's equations are

$$\dot{p}_r = m\ddot{r} = \frac{\partial L}{\partial r} = mr\dot{\theta}^2 \quad (29)$$

which we see is the balance of radial acceleration and centripetal 'force' and

$$\dot{p}_\theta = \frac{d}{dt}(mr^2\dot{\theta}) = \frac{\partial L}{\partial \theta} = 0 \quad \Rightarrow \quad mr\ddot{\theta} = -2m\dot{r}\dot{\theta} = 0. \quad (30)$$

which states the conservation of angular momentum, and involves the so-called Coriolis term when written out.

2.5 Hamiltonian from Lagrangian

- The Lagrangian $L(q, \dot{q})$ that generates the equations of motion, is a function of coordinates and velocities. On the other hand, in the Hamiltonian $H(q, p)$ is a function of coordinates and momenta.

- In the simple case we considered, $L = T - V$ while $H = T + V$, so they are not the same function! In general, $H(q, p)$ is obtained from $L(q, \dot{q})$ via a *Legendre transformation*.

- Notice that the definition of momenta $p = \frac{\partial L}{\partial \dot{q}}$ is in fact the condition for the function $p\dot{q} - L(q, \dot{q})$ to be extremal with respect to \dot{q} . Now we may invert the relation $p = \frac{\partial L(q, \dot{q})}{\partial \dot{q}}$ and solve for $\dot{q}(p, q)$ as a function of p and q . The hamiltonian is defined as the result of substituting this expression for \dot{q} in $p\dot{q} - L(q, \dot{q})$. In other words, $H(q, p)$ is the extremal value of $p\dot{q} - L(q, \dot{q})$

$$H(q, p) = \text{ext}_{\dot{q}}(p\dot{q} - L(q, \dot{q})) \quad (31)$$

- Let us check that this reproduces the energy of a particle in a potential V

$$L = T - V = \frac{1}{2}m\dot{q}^2 - V(q), p = \frac{\partial L}{\partial \dot{q}} = m\dot{q} \Rightarrow \dot{q} = p/m \quad (32)$$

For this \dot{q} ,

$$H = p\dot{q} - L = p\frac{p}{m} - \frac{1}{2}m\frac{p^2}{m^2} + V(q) = \frac{p^2}{2m} + V. \quad (33)$$

which we recognize as the energy.

- Another way of viewing the Legendre transformation is to look at the differential of the Lagrangian.

$$dL(q, \dot{q}) = \frac{\partial L}{\partial q} dq + \frac{\partial L}{\partial \dot{q}} d\dot{q} = \dot{p} dq + p d\dot{q} \quad (34)$$

where we used the definition of momentum and Lagrange's equation:

$$p = \frac{\partial L}{\partial \dot{q}}, \quad \text{and} \quad \dot{p} = \frac{\partial L}{\partial q}. \quad (35)$$

Now we want the independent variable to be p instead of q . So define $H = p\dot{q} - L$ and compute its differential

$$dH = p d\dot{q} + \dot{q} dp - \dot{p} dq - p d\dot{q} = \dot{q} dp - \dot{p} dq \quad (36)$$

This makes it clear that the independent variables in H are q and p and moreover, that

$$\dot{q} = \frac{\partial H}{\partial p}, \quad \text{and} \quad \dot{p} = -\frac{\partial H}{\partial q}. \quad (37)$$

These we recognise as Hamilton's equations!

- The Lagrangian can be re-obtained from $H(q, p)$ by by an (inverse) Legendre transform

$$L(q, \dot{q}) = \text{ext}_p [p\dot{q} - H(q, p)] \quad (38)$$

3 Basic concepts of waves

- A particle is a localized object characterized by its position and momentum. Given its current state (q, p) or (q, \dot{q}) , Newton's equation tells us how to determine its trajectory.
- A wave on the other hand is a disturbance (say a wave in a guitar string or a sound wave in air) that is typically spread out and described by a *wavefunction* $\psi(x, t)$ which specifies the *amplitude* of the disturbance at each spatial point x at time t . ψ could be the height of a Veena string about its undisturbed position, or the excess air pressure or in the case of electromagnetic waves, it could be a component of the electric field.
- A simple example of a wave is a traveling wave, i.e., one that maintains its shape as it propagates. A wave traveling along the x-axis can be described by

$$\psi(x, t) = f(x - ct) \quad (39)$$

This describes a disturbance with profile $f(x)$ moving to the right at a speed $c > 0$. If $c < 0$ it moves to the left.

- As a specific example, consider a traveling wave with a sinusoidal profile, conventionally written

$$\psi(x, t) = A \sin(kx - \omega t) = A \sin \left[k \left(x - \frac{\omega}{k} t \right) \right] = A \Im e^{ikx - i\omega t} \quad (40)$$

For $\omega/k > 0$, such a wave travels to the right at a speed $c = \frac{\omega}{k}$, called the phase speed. This may represent an electromagnetic wave, in which case c is the speed of light. k (sometimes $|k|$) is called the wave number and $|\omega|$ is called the angular frequency. We will assume $\omega, k > 0$ in most of what follows.

- Written this way, A is called the amplitude of the wave, it is actually the maximum amplitude of the disturbance ψ .
- ω describes how often the wave oscillates at (any) fixed point x . The time period of such oscillations is $T = \frac{2\pi}{\omega}$. The frequency of such oscillations is defined as $\nu = \frac{\omega}{2\pi}$ and has dimensions of inverse time (*hertz* = s^{-1}).
- The wave number k (with dimensions of inverse length) controls 'how many waves' there are (how many times the profile repeats itself) in a spatial interval of length 2π at (any) fixed time. For e.g. $\sin x$ has wave number one and $\sin 2x$ has wave number two since the wave repeats twice within a distance of 2π . In terms of the wave length, $k = \frac{2\pi}{\lambda}$. The wave length is the minimal spatial distance over which the profile repeats itself, at any fixed time.
- We can also have waves in more than one dimension. For example, in three dimensions, the wave function (e.g. excess pressure), depends on x, y, z and time. A traveling sinusoidal plane wave is for example

$$\psi(x, y, z, t) = A \sin(k_x x + k_y y + k_z z - \omega t) = A \sin(\vec{k} \cdot \vec{r} - \omega t) \quad (41)$$

Here $\vec{k} = (k_x, k_y, k_z)$ is called the wave vector, the vector whose components are the wave numbers. This wave travels in the direction \vec{k} at a speed given by $c = \frac{\omega}{k}$ where $k = |\vec{k}|$ is the magnitude of the wave vector. This is called a plane wave since the wave fronts (set of points with a common amplitude at a given time) are planes perpendicular to the wave vector. To see this, take $\vec{k} = k\hat{x}$ for instance. Then the amplitude $\psi(\vec{r}, t) = A \sin(kx - \omega t)$ is independent of y, z . Thus, all points on the $y - z$ plane $x = 0$ have the same amplitude

A when $t = 0$. In fact the amplitude is the same on any plane parallel to the $y - z$ plane at a distance $x = \frac{2n\pi}{k}$. Thus the wave fronts are planes perpendicular to the wave vector. This wave travels in the \hat{x} direction, i.e., in the direction of the wave vector \vec{k} . To get a plane wave traveling in any other direction, we just need to rotate (and re-scale) this particular \vec{k} .

- A spherical wave has wave fronts that are concentric spheres centered about a point (say the source of the explosion or the destination of the implosion). For example, if $r = \sqrt{x^2 + y^2 + z^2}$, then

$$\psi(x, y, z, t) = \frac{A \sin(kr - \omega t)}{r} \quad (42)$$

is an outgoing spherical wave with wave number k , angular frequency ω and speed $c = \omega/k > 0$. Notice that the amplitude $A(r) = A/r$ depends only on the distance from the center. This particular choice of $A(r)$ ensures that the total intensity over a spherical surface is independent of the radius of the sphere, which is necessary for conservation of energy.

- Not all waves need travel. A standing wave or stationary wave is what is observed in a stretched guitar string held down at its end points $x = 0, 2\pi$. For example

$$\psi(x, t) = A \sin(\omega t) \sin x \quad (43)$$

describes a standing wave. The spatial profile is always $\propto \sin x$, just that the coefficient of proportionality oscillates at a frequency ω .

- Standing waves can arise through a superposition of a pair of left and right moving waves of the same frequency. For example, suppose we make a right moving disturbance $\sin(kx - \omega t)$ in a string can reflect off the end point to produce a left moving wave $\sin(kx + \omega t)$. The two can interfere (add up) and produce a standing wave $2 \sin(kx) \cos(\omega t)$.
- The analogue of Newton's equation for a wave is called a wave equation. One of the simplest wave equations is the one describing small amplitude vibrations of a stretched string

$$\frac{\partial^2 \psi}{\partial x^2} = \frac{1}{c^2} \frac{\partial^2 \psi}{\partial t^2} \quad (44)$$

In this case, the analogue of position and velocity are ψ and $\dot{\psi}$. Given these at $t = 0$ the wave equation may be solved to determine $\psi(x, t)$ at future times. Notice that this wave equation is a linear (small amplitude), homogeneous (no external driving force) second order partial differential equation.

- For example, if $\frac{\omega^2}{k^2} = c^2$, then the traveling wave

$$\psi(x, t) = A e^{i(kx - \omega t + \phi)} \quad (45)$$

is a (complex) solution of the wave equation for any real constants A and ϕ . ϕ is called the initial phase while A is the maximum amplitude. Thus we may say that the above wave equation describes waves propagating at the speed c .

- Since the wave equation is a linear equation, we may add (superpose) two solutions and get another solution.
- When $\psi(x, t) = u(x, t) + iv(x, t)$ is a complex solution to the wave equation (44), its real and imaginary parts u, v are also solutions of the same wave equation (44). This is because the imaginary unit i does not appear explicitly in (44). This wave equation does not

couple the real and imaginary parts. So we are free to work either with real solutions, or for convenience, with complex solutions. In particular, the real and imaginary parts (with cosine and sine profile) of the above complex solution are also solutions of the wave equation and are appropriate for describing real disturbances like heights of strings or acoustic pressure disturbances.

- The Schrödinger equation of quantum mechanics for a free particle is also a wave equation, but first order in time

$$i\hbar \frac{\partial \psi(x,t)}{\partial t} = -\frac{\hbar^2}{2m} \frac{\partial^2 \psi(x,t)}{\partial x^2} \quad (46)$$

We will motivate and study this equation later. Here we observe that the imaginary unit explicitly appears in (46). This implies that (46) necessarily couples the real and imaginary parts of ψ . So ψ cannot be taken to be real if it has to be a non-zero solution of the Schrödinger equation. In this context, $\psi(x,t)$ represents a complex-valued probability amplitude, rather than the height of a string or electric field.

3.1 Wave packet and its width

- A monochromatic wave $\psi = Ae^{i(kx-\omega t)}$ has a definite wavelength $\lambda = 2\pi/k$ and travels at a fixed speed $c = \omega/k$. It extends over the whole of space (x), indeed its intensity $I = |\psi(x,t)|^2 = |A|^2$ is a constant. So it cannot be used to model a localized pulse (such as a light pulse).

- Recall (from Young's double slit experiment) that the interference of two light waves with slightly different phases can produce an interference pattern of intensities that is peaked in a region of constructive interference and dies down elsewhere through a sequence of bright and dark fringes.

- This suggests that we can build a wave in the shape of a localized pulse by superposing monochromatic waves of slightly different wavelengths and frequencies. The idea is that they will constructively interfere in a small region (where the pulse is located) and destructively interfere elsewhere. Such a pulse is called a wave packet.

- Let us first construct a static wave packet (not moving). For this we superpose a bunch of waves with wave numbers in a small neighborhood of k , i.e., in the interval $k - \delta k, k + \delta k$. Here $\delta k \ll k$

$$\psi(x) = \int_{k-\delta k}^{k+\delta k} e^{ik'x} \frac{dk'}{2\pi} = \left[\frac{e^{ik'x}}{2\pi ix} \right]_{k-\delta k}^{k+\delta k} = \frac{e^{ikx} (e^{i\delta kx} - e^{-i\delta kx})}{2\pi ix} = \frac{2e^{ikx} \sin(\delta kx)}{2\pi x} \quad (47)$$

We see that the $\frac{1}{x}$ factor ensures that the amplitude decays far from $x = 0$. This is a wave packet located near $x = 0$. Since $\delta\lambda = 2\pi/\delta k \gg 2\pi/k = \lambda$, the $\sin(\delta kx)$ provides a long wavelength oscillatory envelope (modulation) to a short wavelength oscillation of e^{ikx} . Draw a diagram of the real part.

- This wave packet has a width in wave-number space of twice δk . Its width in physical x -space is the region δx over which the intensity

$$I = |\psi(x)|^2 = \frac{4 \sin^2(\delta kx)}{x^2} \quad (48)$$

is significant. This corresponds to the interval between the first positive and negative nodes of $\sin^2(\delta kx)$, since for larger x , the $1/x$ factor will dampen the oscillations. Thus $\delta x \sim \frac{2\pi}{\delta k}$.

- Thus, the width of a wave packet in position and wave number spaces are reciprocally related. To make the wave packet concentrated in a small region near $x = 0$, we will need δx small. To achieve this, we will have to take δk large. As a result, we need a large range of wave numbers to synthesize a sharply localized wave packet. A sharply localized wave packet begins to look like a particle, since it can be located precisely. However, its monochromatic-wave-like character is lost, since we cannot assign to it a particular wavelength, but rather needed a whole range of wavelengths $[\frac{2\pi}{k-\delta k}, \frac{2\pi}{k+\delta k}]$ to construct it.

- Conversely, the more monochromatic a wave packet is (δk small), the more spread out it is ($\delta x \sim \frac{1}{\delta k}$ large) in physical x -space.

- This observation will eventually lead to the uncertainty principle $\Delta x \Delta p \geq \hbar/2$ of quantum mechanics that relates the accuracy with which position and momentum of a quantum mechanical object may be simultaneously specified. The crucial missing link is to relate the wave number to momentum. This is a connection that is classically absent, but de Broglie postulated that quantum mechanically, $p = \hbar k$. We will explain this later.

- We have secretly introduced the (inverse) Fourier transform. Let us write (47) as

$$\psi(x) = \int_{-\infty}^{\infty} \frac{dk}{2\pi} \tilde{\psi}(k) e^{ikx}. \quad (49)$$

where $\tilde{\psi}(k)$ is 1 on the interval $[k - \delta k, k + \delta k]$ and zero otherwise. More generally $\tilde{\psi}(k)$ may be any function of wave numbers that decays sufficiently fast as $|k| \rightarrow \infty$. We have merely taken a linear combination of monochromatic waves (Fourier modes) e^{ikx} with amplitudes $\tilde{\psi}(k)$. $\tilde{\psi}(k)$ is called the amplitude of the k^{th} Fourier mode. $\psi(x)$ is called the inverse Fourier transform of $\tilde{\psi}(k)$. The function $\tilde{\psi}(k)$ is called the Fourier transform of $\psi(x)$:

$$\tilde{\psi}(k) = \int_{-\infty}^{\infty} \psi(x) e^{-ikx} dx \quad (50)$$

- The Fourier transform $\psi \mapsto \tilde{\psi}$ is a linear transformation on the vector space of (complex-valued) square-integrable functions on the real line $L^2(\mathbb{R}, \mathbb{C})$. If $\psi(x)$ is square-integrable,

$$\text{square of } L^2 \text{ norm of } \psi = \|\psi\|^2 = \int dx |\psi(x)|^2 < \infty, \quad (51)$$

the integral defining the Fourier transform makes sense (actually, it is enough for ψ to be absolutely integrable). We will return to Fourier transforms later.

3.2 Motion of a wave packet and group velocity

- We constructed (synthesized) a wave packet by superposing monochromatic waves with a range of wave numbers. To see how a wave packet evolves, we restore the time dependence of the component waves.

- If the component monochromatic waves all travelled at the same (phase) speed c , then the wave packet as a whole would travel at that speed:

$$\psi(x, t) = \int_{k-\delta k}^{k+\delta k} e^{ik'(x-ct)} dk' = \left[\frac{e^{ik'(x-ct)}}{i(x-ct)} \right]_{k-\delta k}^{k+\delta k} = \frac{2e^{ik(x-ct)} \sin(\delta k(x-ct))}{x-ct} \quad (52)$$

This represents a traveling wave shaped like a localized pulse located near $x = ct$ at any instant t . Notice that the pulse maintains its profile while moving to the right.

- More generally, different wavelengths could travel at different speeds. This is indeed the case for light traveling through a glass prism or other such *dispersive* medium where the refractive index depends on wavelength.
- In a dispersive medium, different wavelength components of a wave packet travel at different speeds. So dispersion tends to change the shape (broaden out) a wave packet with time. This is evident in the fanning out of a pulse of light upon passing through a prism.

- Though the component sinusoidal waves of a packet may travel at different speeds, the packet as a whole may have a well defined speed. This could be defined as the speed at which the center of the wave packet propagates.

- For simplicity let us try to construct a wave packet by superposing a pair of sinusoidal waves with nearby wave numbers and frequencies

$$\psi_1(x, t) = A \sin(kx - \omega t) \quad \text{and} \quad \psi_2(x, t) = \sin((k + \delta k)x - (\omega + \delta \omega)t) \quad (53)$$

In this case, the two waves travel at different phase speeds $c_1 = \omega/k$ and $c_2 = \frac{\omega + \delta \omega}{k + \delta k}$. What is the speed of the combined wave $\psi = \psi_1 + \psi_2$?

$$\psi = \psi_1 + \psi_2 = A \sin\left(\frac{x(2k + \delta k) - t(2\omega + \delta \omega)}{2}\right) \cos\left(\frac{x \delta k - t \delta \omega}{2}\right) \quad (54)$$

We assume that $\delta k \ll k$ and $\delta \omega \ll \omega$. So we may approximate $2k + \delta k \approx 2k$ etc.

$$\psi(x, t) \approx A \sin(kx - \omega t) \cos\left(\frac{x \delta k - t \delta \omega}{2}\right) \quad (55)$$

So the combined wave is the original wave multiplied by $\cos\left(\frac{x \delta k - t \delta \omega}{2}\right)$ which is a low frequency, low wave number (larger time period and larger wave length) modulation.

- Plot $\psi(x, t)$ and observe that it consists of a sequence of wave groups (wave packets) whose envelope is defined by the modulating wave.

- The velocity of the wave packets is the phase velocity of the modulating wave (envelope), the so-called group velocity $c_g = \frac{\delta \omega}{\delta k}$. The group velocity is the speed at which energy (\propto square of the amplitude) is carried by the wave packet.

- This concept of group velocity can be adapted to describe the speed of other wave packets. To do so we need the concept of dispersion relation.

- We can express a sinusoidal wave passing through a dispersive medium as

$$\psi = A e^{i(kx - \omega(k)t)} \quad (56)$$

Here $\omega = \omega(k)$ expresses the frequency as a function of wave number (wavelength). It is called the dispersion relation and gives the phase velocity $c(k) = \frac{\omega(k)}{k}$ as a function of wave number.

- For example, let us consider the DISPERSIVE LINEAR WAVE EQUATION (for some positive constants c, σ with dimensions of speed and area)

$$\frac{1}{c} \partial_t \psi + \partial_x \psi + \sigma \partial_{xxx} \psi = 0 \quad (57)$$

- The above sinusoidal wave form is a solution provided

$$-\frac{i}{c}\omega(k) + ik - i\sigma k^3 = 0 \quad (58)$$

In other words, the dispersion relation for waves in this medium is

$$\omega(k) = c(k - \sigma k^3) \quad (59)$$

The phase speed of a sinusoidal wave of wavelength $\lambda = 2\pi/k$ would be

$$c_p(k) = \frac{\omega(k)}{k} = \frac{c(k - \sigma k^3)}{k} = c - c\sigma k^2 \quad (60)$$

- Let us synthesize a wave packet by superposing several of these sinusoidal waves

$$\psi(x, t) = \int \frac{dk'}{2\pi} \tilde{\psi}(k') e^{i(k'x - \omega(k')t)} \quad (61)$$

Notice that we have combined these waves with possibly different amplitudes $\tilde{\psi}(k)$. $\tilde{\psi}(k)$ specifies how much of each wavelength component to take in the wave packet. In our earlier example, $\tilde{\psi}(k)$ was the indicator function (characteristic function) of the interval $[k - \delta k, k + \delta k]$.

$$\tilde{\psi}(k) = \theta(k - \delta k \leq k \leq k + \delta k) \quad (62)$$

which is 1 for k in this interval and zero for k outside. In wave number space, this wave packet would be centered at wavenumber k and have a width $2\delta k$.

- In this case, a reasonable definition of group velocity is

$$c_g = \left. \frac{d\omega(k')}{dk'} \right|_{k'=k} = c - 3c\sigma k^2 \quad (63)$$

- This would be the speed with which the wave packet $\psi(x, t)$ as a whole travels. Notice that the group speed is less than the phase speed for every k , as is often the case for waves in dispersive media.
- If we think of a wave packet as approximating a pulse or bullet or particle, then its group velocity is the velocity of the pulse/bullet/particle!

4 Quanta of light: particle-like properties of radiation

- Light displays interference as established in Young's double slit experiment. This was explained using the wave theory of light, but could not be accounted for using the corpuscular model. Other optical phenomena like reflection and refraction could also be explained using Maxwell's theory of electromagnetic waves.
- So it came as a surprise that several phenomena discovered around the turn of the twentieth century, could not be explained using the wave nature of light but admitted a simple explanation based on the assumption that light came in packets called photons (Planck's quanta of light), each with a definite frequency ν and energy $h\nu$. According to this hypothesis, monochromatic light consisted of a train of n photons having a total energy of $nh\nu$ with $n = 0, 1, 2, \dots$. There was no such thing as half a photon. Chief among these phenomena are the black body spectrum, photo-electric effect and Compton scattering.

4.1 Black body radiation and Planck hypothesis

- The Blackbody radiation problem concerns the frequency distribution of radiant energy in a cavity whose walls are maintained at a fixed temperature T .
- Hot iron radiates visible light. As the metal is heated, the colour changes from red through orange, yellow, blue etc. Even a colder object radiates, but we can't see the infrared light coming out. In fact, at any temperature, objects radiate a whole range of frequencies, it is just that our eyes only perceive the most intense (energetic) frequency (provided it lies in the visible range). The distribution of radiant energies as a function of frequency is called the spectral distribution of emitted radiation.
- We will be interested in bodies in thermal equilibrium with their surroundings, so that they can be assigned a temperature. In this case, the body must emit energy at the same rate as it absorbs, in order to remain at the same temperature.
- A black body is an idealized object that absorbs all radiation incident on it, irrespective of wavelength. For example, a blackbody can be modeled by a small hole in a hollow box kept in an oven. The oven allows us to control temperature. Any radiation that enters the hole gets absorbed by the box after repeated internal reflections against the walls. However, the walls of the box also radiate, and some of this radiation escapes through the hole. We are interested in the spectral energy distribution of this radiation. The advantage of using this idealized blackbody is that the spectral radiance is universal, i.e., independent of the material used to make the box and its shape.
- The walls of the cavity are in equilibrium with the interior at a temperature T , which is controlled by the oven. When the blackbody emission is analyzed, it is found to include all wavelengths, but is peaked around a wavelength that depends on T . Hotter bodies radiate maximally at shorter wavelengths (higher frequencies). This is Wien's law $\lambda_{max}T = \text{constant}$. Moreover, the total energy radiated increases with temperature.
- Let $u(\nu)d\nu$ be the radiant energy per unit volume in the frequency interval $[\nu, \nu + d\nu]$. $u(\nu)$ is the energy per unit volume per unit frequency interval. When measured in the late 1800s and plotted $u(\nu)$ increases quadratically at low frequencies, reaches a peak at $\nu_{max} = c/\lambda_{max}$ and then decays rapidly with growing frequency.
- Rayleigh (1900) and Raleigh and Jeans (1905) derived a formula for $u(\nu)$ based on classical electrodynamics and thermodynamics. They assumed that the atoms in the cavity walls oscillate and emit electromagnetic waves that must fit into the cavity (standing waves). They counted the number of standing waves per frequency interval and used equipartition at temperature T to assign a mean energy kT per standing wave at equilibrium to get

$$u(\nu) = \frac{8\pi\nu^2 kT}{c^3}, \quad \text{where } k = 1.38 \times 10^{-23} \text{ J/K} = 86 \mu\text{eV/K} = \text{Boltzmann's constant.} \quad (64)$$

Check that $u(\nu)$ has the correct dimensions.

- Though the Raleigh-Jeans law gives the correct spectral radiance at low frequencies (quadratic in ν) it grows without bound at large frequencies (UV), in marked contrast with the measurements. Moreover, it is conceptually flawed since it predicts an infinite total energy density across all frequencies. This was called the ultraviolet catastrophe and represented a dramatic failing of classical physics.

- Around the same time (1900) Planck found a formula that fit the experimental measurements of spectral energy density at all frequencies

$$u(\nu) = \frac{8\pi\nu^2}{c^3} \frac{h\nu}{e^{h\nu/kT} - 1}, \quad \text{where } h = 6.626 \times 10^{-34} \text{ Js} = 4.1 \times 10^{-15} \text{ eVs} = \text{Planck's constant} \quad (65)$$

This formula reproduces the Raleigh-Jeans law at low frequencies but decays exponentially at large frequencies, ensuring a finite total energy density.

- Planck's constant h has dimensions of action, or energy times time or length times momentum and is called the quantum of action.
- To explain his formula, Planck assumed that the energy of the oscillators in the walls of the cavity were restricted to multiples of $h\nu$. Though the essential idea of quantized energies was correct, a proper explanation of Planck's radiation law needed additional ideas on how to count quantum objects (quantum statistics). This was provided by S N Bose only in 1924.
- To explain this formula, one has to postulate that light of frequency ν cannot have an arbitrary energy, but can only have energies in packets of size $h\nu$. n photons would have energy $nh\nu$. These packets of light are not electromagnetic waves, but photons⁶. Moreover, the average number of photons of energy $h\nu$ is given by the Bose-Einstein distribution $(e^{h\nu/kT} - 1)^{-1}$. So the average energy of the photons is $h\nu (e^{h\nu/kT} - 1)^{-1}$, rather than just kT as classical equipartition suggests. Multiplying by the number of standing waves (number of photons), one gets Planck's black body spectrum.

4.2 Photoelectric Effect

- The photo electric effect (discovered by Hertz and others and explained by Einstein) concerns the ejection of electrons (photo-electrons) from the surface of a metal when irradiated by (UV) light.
- The main problem was that the energy of the ejected electrons was independent of the intensity of light, and only depended on its frequency, a fact that could not be explained classically.
- The experiment consisted of irradiating a metal surface (photocathode) with monochromatic light (single frequency). A clever trick was used to determine the maximal energy of the liberated photoelectrons. A collector is held at a negative voltage compared to the photocathode, so the negatively charged electrons have to cross the potential barrier to arrive at the collector and register a current.
- As the collector voltage is lowered (making it harder for the electrons to reach it), the photocurrent decreases and vanishes at a certain stopping potential V_s (a few volts). eV_s is then the energy of the most energetic photoelectrons. Merely increasing the intensity of the light did not restore the photocurrent.
- So the stopping potential is independent of the intensity (energy crossing unit area per unit time) of light. However, is found to grow linearly with the frequency $V_s \sim a\nu + b$. Moreover, the slope a was independent of the metal that was irradiated, while the intercept b was a

⁶However, a large collection of photons mimics an electromagnetic wave. Light of appreciable intensity has such a large number of photons n that the energy in the wave is very large compared to $h\nu$, the discreteness in steps of $h\nu$ becomes imperceptible.

property of the material. Classically, it was expected that increasing the intensity would increase the energy transmitted to the electrons, this was at odds with the constancy of the stopping potential.

- The photo current was proportional to the light intensity (while holding the frequency fixed): more electrons were ejected with intense light.
- There was almost no time lag ($< 10^{-9} s$) between switching on the light and the emission of photoelectrons. Classical estimates indicated a time-lag of a second or more.
- Einstein explained this in 1905 by postulating that light of frequency ν consists of discrete lumps (photons) each with energy $h\nu$ moving at the speed of light. $h = 4.135 \times 10^{-15} eV s$ is Planck's constant introduced by Planck to explain the blackbody spectrum. So n photons in a beam would have energy $nh\nu$. Increasing the intensity of the beam just increases the number of photons n without changing the energy of any single photon. Moreover, he hypothesized that the photo electric emission was caused by a single photon knocking out a single electron from the metal, with the energy transmitted in the collision almost instantaneously.
- The energy that an electron can gain is then $h\nu$, since the likelihood of multiple photon collisions is very small. Now, an electron that receives this energy has to overcome the attractive forces of the ions in the metal, and in favorable cases, would make it out of the metal. Its energy upon emerging from the metal would depend on details such as how deep inside the metal the electron was, how much energy it initially had etc. However, the minimum energy required to extract an electron from a metal is called its work function Φ (sometimes denoted W). Thus, the maximum energy that a photoelectron can have is $h\nu - \Phi$. Recall that eV_s is the maximum energy of the ejected electrons. Thus

$$V_s = \frac{h\nu}{e} - \frac{\Phi}{e} \quad (66)$$

We see the linear dependence of the stopping potential on frequency, with a universal slope h/e independent of the metal. Moreover, the intercept $\frac{\Phi}{e}$ does depend on the metal through the work function.

- Increasing the intensity of the light only increases the number of photons, which correspondingly increases the number of photoelectrons and thereby the photocurrent, it does not change the maximal energy of the photoelectrons.
- The slope of stopping potential as a function of frequency is a way of measuring Planck's constant, this was done by Millikan in 1916. It is approximately

$$h = 6.626 \times 10^{-34} J s = 4.135 \times 10^{-15} eV s. \quad (67)$$

Planck's constant can be regarded as an order of magnitude of quantum effects in quantities having the dimensions of action (energy \times time). Planck's constant is very small on the time (hours) and energy scales (kilowatt hours) we usually encounter.

- The photoelectric effect indicates that the transfer of energy from light (radiation) to electrons happens discontinuously (in the amount $h\nu$, this quantum of light is absorbed), and almost instantaneously. Classically electromagnetic theory predicts that the transfer of energy from electromagnetic wave to the electron is a much slower and continuous process till the electron has accumulated the necessary energy to escape from the metal. This contradicts the experimental finding.

4.3 Compton Scattering Effect

- further reading, see Beiser, Mani & Mehta, Bohm.
- The Compton effect indicates the particle-like (photon) behavior of light (x-rays and γ -rays). It was discovered by Compton in 1923.
- Compton scattering is the scattering of light of a fixed wavelength λ (usually X-rays or γ rays) with an electron (in an atom or in matter). The outgoing light makes an angle θ (with respect to the incoming beam), and also has a greater wavelength to λ' . We could say that the light has lost some of its energy to the electron. Meanwhile, the electron is ejected from the atom/material with a certain kinetic energy and at an angle ϕ to the incoming beam of light. In the experiment, the wave length of incoming X-rays and the wavelength of outgoing X-rays and the angle θ at which they come out can be measured using a spectrometer. The energy and direction ϕ of the recoil electron is not as easily measured.
- Classical electromagnetic theory predicts that by changing the intensity of the incoming light (or by changing the duration of exposure to light), one can change the wavelength of outgoing X-rays at a fixed angle. Experimentally, the shift (increase) in the wavelength $\lambda' - \lambda$ depends on the scattering angle in a specific manner, independent of light intensity and exposure time. At a given angle, the outgoing X-rays have a specific wavelength. This was explained by Compton by assuming that light behaves like a particle (photon) with energy $E = h\nu$ and scatters with the electron in a manner similar to the elastic collision of a mustard seed with a golf ball. He derived the formula

$$\lambda' - \lambda = \frac{h}{mc} (1 - \cos \theta) \quad (68)$$

According to this formula, the shift in wavelength at a fixed angle is a universal constant (totally independent of the light intensity, in contrast to classical expectations).

- In the quantum theory this process is described as follows: the incoming quantum of light (photon) is absorbed by the electron. The electron then emits another photon (with a different wavelength) and also emerges with some kinetic energy. The wavelength increases since the photon has lost energy to the electron. The formula says that the shift in wavelength is maximal ($= 2h/mc$) for photons that are backscattered ($\theta = \pi$) and half as much for photons that are scattered at right angles ($\theta = \pi/2$). The forward scattered photons $\theta = 0$ suffer no wavelength shift. This prediction was confirmed by Compton himself, using a spectrometer to measure the wavelength of X-rays coming out in various directions. In practice, photons are scattered in all directions. However, one can check whether the photons coming out at a particular angle θ have an elongated wavelength as given by Compton's formula. This was found to be the case and represented strong evidence in support of the particle-like character of the photon. It also confirmed the hypothesis that the energy in monochromatic light beams comes in discrete packets of size $h\nu$ and that the momentum of these quanta of light is of magnitude $h\nu/c$ and as a vector, $\frac{h\vec{k}}{2\pi}$ where \vec{k} is the wave vector of the corresponding electromagnetic wave.
- Compton's derivation uses conservation of energy and momentum. However, he treats the electron motion after collision relativistically since the photons come in with a relatively high energy (1 MeV or more), exceeding the rest energy of the electron (.511 MeV). The energy of the electron due to its motion in the atom is small compared to the X-ray energy and can be ignored.

- Initially, the electron is assumed to be at rest (in any case, its speed in an atom is quite small compared to the speed it acquires in the collision), so it has a rest energy mc^2 ($m = 511\text{KeV}$ is its rest mass). Conservation of energy says that

$$h\nu + mc^2 = h\nu' + \text{KE} + mc^2 \quad (69)$$

Here KE is the kinetic energy of the scattered electron.

- Let p be the magnitude of the momentum of the outgoing electron. The momentum of the incoming photon is $h\nu/c$. Momentum conservation in the direction of the incoming X-rays gives

$$\frac{h\nu}{c} = \frac{h\nu'}{c} \cos \theta + p \cos \phi \quad (70)$$

Momentum conservation in the direction perpendicular to the incoming beam (and in the plane defined by the incoming & outgoing photons and the electron) gives

$$0 = \frac{h\nu'}{c} \sin \theta - p \sin \phi \quad (71)$$

We eliminate ϕ by adding the squares of the momentum conservation equations

$$p^2 c^2 = (h\nu - h\nu' \cos \theta)^2 + h^2 \nu'^2 \sin^2 \theta \quad (72)$$

Now we will eliminate p using the fact that the kinetic energy of the electron is $\text{KE} = h\nu - h\nu'$. Recall that the total energy of the electron is a sum of its rest-energy and kinetic energy $E = mc^2 + \text{KE}$. On the other hand, the energy of the electron is related to its momentum by⁷

$$E^2 - p^2 c^2 = m^2 c^4 \Rightarrow p^2 c^2 = \text{KE}^2 + 2mc^2 \text{KE}, \quad \text{KE} = h(\nu - \nu') \quad (73)$$

This allows us to eliminate p and get the Compton shift in wavelength

$$h^2 \nu \nu' - h^2 \nu \nu' \cos \theta = mc^2 (\nu - \nu') \Rightarrow \frac{\nu - \nu'}{\nu \nu'} = \frac{h}{mc^2} (1 - \cos \theta) \Rightarrow \lambda' - \lambda = \frac{h}{mc} (1 - \cos \theta). \quad (74)$$

The shift in wavelength is expressed in terms of the easily measured photon deflection angle.

- The coefficient $\frac{h}{mc}$ is a property of the electron, called the Compton wavelength of the electron. It is the wavelength of a photon that has the same energy as the rest energy of the electron

$$E_e = m_e c^2 = E_\gamma = h\nu = \frac{hc}{\lambda} \Rightarrow \lambda_{\text{Compton}} = \frac{hc}{m_e c^2} = \frac{h}{mc} \quad (75)$$

The Compton shift in the photon wavelength is \propto to the Compton wavelength of the electron.

- Note that the definition of Compton wavelength involves the fundamental constants of both relativity and quantum mechanics. The Compton wavelength of an electron is much smaller than atomic dimensions.

$$\lambda_{e\text{-Compton}} = \frac{h}{mc} = \frac{4.1 \times 10^{-15} \text{ eV s}}{511000 \text{ eV}/c} = 2.4 \times 10^{-12} \text{ m} \quad (76)$$

⁷The lab-frame energy of a particle of rest mass $m > 0$ moving at a speed v with respect to the lab is γmc^2 , its momentum is of size $p = \gamma mv$ where $\gamma = 1/\sqrt{1 - v^2/c^2}$. Thus $E^2 - p^2 c^2 = m^2 c^4$ irrespective of the speed. Taking $m \rightarrow 0$ gives $E = pc$ for a massless particle. We could also take $m \rightarrow 0$ and $v \rightarrow c$ directly in the formulae $E = \gamma mc^2, p = \gamma mv$, they would say that E and p can be arbitrary positive quantities in the limit, subject only to one constraint, that E/p equals c in the limit. This is indeed the case with photons, they can have arbitrary energy and momentum as long as $E/p = c$.

- The Compton wavelength of the electron is *not* the size of the electron. In situations where the electron can be described as a particle, the electron is point-like to current precision. Experimentally, the size of an electron is known to be smaller than $10^{-19}m$.
- Despite Einstein's success in explaining the photo electric effect and Compton's success in predicting the Compton shift, it was not possible to say what would happen when one photon hits one particular electron. All that could be explained (in Compton scattering) was the angular distribution of scattered photons (if many were incident), not where any particular photon would go⁸. This is an indication that the laws of quantum mechanics primarily allow statistical predictions, quite unlike those of classical mechanics. We say *primarily*, since the laws of QM guarantee conservation of momentum and energy in *each* collision (not just on average). The deterministic laws of continuous classical evolution are replaced by quantum laws of how probability (amplitudes) evolve. But classical laws that are valid even in discontinuous processes (like energy conservation in collisions) continue to be valid in quantum mechanics.

5 Quantum and wave properties of matter (electrons, molecules and atoms)

5.1 Discrete atomic spectra and stability of atom

- Matter in condensed form (solids and liquids) emit radiation at all temperatures (e.g. the radiation emitted by the walls of a cavity). The spectrum of this radiation in the case of a blackbody in equilibrium at temperature T was explained by Planck and Einstein based on the hypothesis that energy of the radiation oscillators comes in packets of size $h\nu$. In this case, there is a (nearly) continuous distribution of wavelengths with varying intensities. Remarkably, the distribution was largely independent of the material and indicates that the radiation is the collective effect of many interacting atoms in the walls of the cavity.
- On the other hand, low density gases of pure elements (e.g. hydrogen, sodium, neon) also emit light when a voltage is passed through them. When this light is passed through a spectrometer, only certain lines in the spectrum are illuminated. These wavelengths constitute the characteristic line spectrum of the element and differ from element to element (unlike the blackbody spectrum). The same is true of the absorption spectrum. Elements absorb light of certain characteristic frequencies. For example, when star light passes through planetary atmospheres (or even the gas around the star), the received light is found to be deficient in specific wavelengths. Light of these wavelengths has been absorbed by the atoms in the planetary atmosphere. The emission/absorption spectra provide information on the chemical composition of the gas. The characteristic line spectrum is a property of individual atoms, so there is no mention of temperature. Temperature is a useful concept only when we have many atoms/molecules in equilibrium.
- Moreover, Balmer fit a simple formula (1885) for lines in the visible part of the hydrogen emission spectrum

$$\frac{1}{\lambda} = R \left(\frac{1}{2^2} - \frac{1}{n^2} \right), \quad R \sim 10^7/m \quad \text{is Rydberg's constant} \quad (77)$$

⁸Here we only found the relation between the wave length shift and angle by simple application of conservation of energy and momentum. QM also predicts the probability distribution of electrons with respect to the angle of scattering

- There was no classical explanation for such discrete spectra of emission/absorption of radiation from atoms⁹. After Rutherford discovered (1911) that the atom is made of a heavy small nucleus surrounded by a large cloud of electrons, he suggested a planetary model of the atom.
- Classical radiation theory predicts that a charged electron moving in an orbit around a nucleus must emit radiation due to its acceleration, however, this radiation does not have a discrete spectrum that matches the characteristic line spectrum measured. Even more problematic is that the electron loses energy by radiating (at progressively higher frequencies) and spirals into the nucleus. This made it impossible to account for the stability of atoms using classical mechanics and classical electrodynamics.

5.2 Bohr model of the atom

- The collection of ideas and heuristic rules surrounding the Bohr model is sometimes called the old quantum theory. It was the best available framework for atomic physics till the development of quantum mechanics in 1925-27.
- Bohr (1913) modified Rutherford's planetary model to accommodate for the stability of the hydrogen atom and discreteness in its emission spectrum.
- The success of his model was due to its simplicity and ability to explain the Balmer formula (and other related ones) for hydrogen spectral lines, as well as calculate the Rydberg constant in terms of other known constants.
- In order to do this, Bohr (and also Sommerfeld) had to postulate a quantization rule for angular momentum. The resulting spectrum for hydrogen turned out to be the same as that implied (later on) by quantum mechanics.
- The weakness of Bohr's model is that it was somewhat ad hoc (basically postulated the stability of the atom and wished away certain classical trajectories because they led to unseen emission lines) and could not be applied to atoms with more than one electron. Moreover, even for the hydrogen atom, the Bohr model has difficulty in explaining its spherical symmetry (and zero angular momentum in the ground state) since it is based on a planetary orbit picture.
- Bohr postulated that electrons in a hydrogen atom can only exist in states (*levels*) with certain discrete (*quantized*) energies E_n for $n = 1, 2, 3, \dots$ (since the radiation emitted/absorbed by atoms was also discrete). In the simplest case, these could be thought of as circular electron orbits around the nucleus. These levels were called stationary, in that an electron could remain in such a level without radiating. He postulated that these levels were characterized by having certain quantized values of angular momentum.

$$L_n = n \frac{h}{2\pi} \equiv n\hbar, \quad n = 1, 2, 3, \dots, \quad \hbar \approx 1.05 \times 10^{-34} \text{ Js.} \quad (78)$$

- Associated to such an orbit is a de Broglie wave, which has to be a stationary/standing wave in order not to destructively interfere with itself. The Bohr quantization condition is

⁹Interestingly, the frequencies are not just multiples of a fundamental frequency. Classically, a charged particle that is kept (by some outside agency) in a fixed periodic orbit around an oppositely charged nucleus, emits radiation in a fundamental frequency (reciprocal of the time period) or possibly in certain multiples thereof. However, in the absence of the agent, the particle will fall into the nucleus

then the condition that n whole wavelengths fit around the orbit of the electron.

- Bohr originally obtained his quantization condition only for large quantum numbers n , using the correspondence principle. This was the requirement that the quantum theory must be well approximated by the classical theory in the limit of large quantum numbers (e.g. when a large number of light quanta are present or when the angular momentum is large compared to \hbar). It was somewhat fortuitous that the quantization condition led to the correct energy levels even down to $n = 1$. But we have not followed the ‘correspondence principle’ line of argumentation here.

- Let us consider an electron of charge e in a planar circular orbit around a proton (assumed infinitely heavy, it is about 2000 times as heavy as an electron) at a distance r . The Coulomb attraction results in an inward force of magnitude $\frac{ke^2}{r^2}$, where e is the proton charge ($-e$ is the electron charge) and $k^{-1} = 4\pi\epsilon_0$.

- To get the allowed energy levels of the hydrogen atom we write

$$E = \frac{1}{2}mv^2 - \frac{ke^2}{r} \quad (79)$$

For a circular orbit at radius r with speed v , the angular momentum is orthogonal to the plane of the orbit and $L = r \times p = mvr$. Moreover, the centripetal force must be provided by the electrostatic Coulomb force

$$\frac{mv^2}{r} = \frac{ke^2}{r^2} \quad \Rightarrow \quad L^2 = mke^2r \quad (80)$$

For the allowed values of angular momentum $L_n = n\hbar$, we have the allowed radii and speeds

$$r_n = \frac{n^2\hbar^2}{mke^2}, \quad v_n = \frac{L_n}{mr_n} = \frac{ke^2}{n\hbar} \quad (81)$$

The corresponding allowed energies are for $n = 1, 2, 3, \dots$

$$E_n = \frac{1}{2}m \frac{k^2e^4}{n^2\hbar^2} - \frac{ke^2mke^2}{n^2\hbar^2} = -\frac{1}{2} \frac{mk^2e^4}{n^2\hbar^2} = -\frac{\mathbb{R}}{n^2} \quad \mathbb{R} = -13.6 \text{ eV}. \quad (82)$$

\mathbb{R} is the Rydberg energy. E_n grow with n and accumulate at $E_\infty = 0$. One Rydberg is the energy required to ionize (take the electron far from the proton) a Hydrogen atom from its ground state.

- Once the energy levels were quantized, it was easy to see why the emission spectrum from hydrogen was discrete. Bohr postulated that light of frequency ν given by $h\nu = E_i - E_f$ was emitted when the electron made a transition from an initial level i to a final level f (or absorbed if $E_f > E_i$). Thus the wavelength of the emitted light is

$$\frac{1}{\lambda} = \frac{\mathbb{R}}{hc} \left(\frac{1}{n_f^2} - \frac{1}{n_i^2} \right) = -R \left(\frac{1}{n_f^2} - \frac{1}{n_i^2} \right). \quad (83)$$

This reproduces Balmer’s formula for the lines in the visible emission spectrum of hydrogen if we take $n_f = 2$. Other choices of n_f correspond to the Lyman, Paschen and other series of spectral lines.

- The Rydberg energy may be expressed in terms of other well-known constants

$$\mathbb{R} = \frac{mk^2e^4}{2\hbar^2} = \frac{1}{2}mc^2 \left(\frac{ke^2}{\hbar c} \right)^2 = \frac{1}{2}mc^2\alpha^2 \approx 13.6eV \quad (84)$$

The combination $\alpha = \frac{ke^2}{\hbar c}$ is a dimensionless number measuring the strength of the electromagnetic force, called the fine structure constant. mc^2 is the rest energy of the electron.

- It is interesting to work out the time period of a Bohr orbit, $T_n = \frac{2\pi r_n}{v_n} = \frac{2\pi\hbar^3 n^3}{mk^2e^4}$, so that the energies E_n are related to the angular frequencies $\omega_n = \frac{2\pi}{T_n}$ by $E_n = \frac{1}{2}n\hbar\omega_n$. We will later study the corresponding relation of energy to angular frequency of oscillations for free particles and particles under linear restoring forces etc.

5.3 Franck-Hertz experiment

- The experiment by Franck and Hertz in 1914 provided experimental support for Bohr's model of the atom (with discrete energy levels). They showed that atoms could absorb energy only in discrete lumps, not just from light, but also from other particles such as electrons.
- Consider a chamber with a gas of atoms (hydrogen or mercury vapour). Bohr's model says that there is a gap in energy ΔE between the ground state of the atom and its first excited state. Now suppose a current of electrons enters the gas of atoms with energy E_i per electron. If $E_i < \Delta E$, the electron cannot excite the atom to its next excited state and therefore any collision has to be elastic, the electron retains its energy E_i . But if $E_i = \Delta E$, the collision can be inelastic and the electron loses its energy to the atom, which is excited to the next level. The remaining electron has lost its kinetic energy and can easily be absorbed by an electrode and cannot contribute to the electron current. Similarly, if the electron energy is $E_i = 2\Delta E$, the electron can have two inelastic collisions with a pair of atoms in the gas and thereby lose its kinetic energy. In this manner we expect the electron current through the gas to sharply decrease at precisely integer multiples of ΔE . This is what Franck and Hertz observed in a gas of mercury vapour kept in a chamber with three electrodes. The electrons entered at the cathode, and were accelerated towards a grid held at a higher voltage. The difference between the grid voltage and cathode voltage is the accelerating voltage that the electrons acquire. Further down, there was an anode at a slightly lower voltage compared to the grid. Electrons that lost their energy to the mercury atoms in inelastic collisions were captured by the grid and did not make it to the anode to contribute to the current measured at the anode. As the accelerating voltage was increased, the anode current also increased, except for sharp drops in the current at integer multiples of ΔE . There were also other drops in the anode current when the accelerating voltage was adequate to excite the atom to states other than the first excited state. Moreover, the excited atoms returned to their ground states by emitting photons. The energies of these photons were just right to match the energies of electrons accelerated by the potential at which the anode current dropped. This provided independent confirmation (independent of emission/absorption spectra) of the discreteness of atomic energy levels.

5.4 Raman scattering

- Closer to home, inelastic scattering of light by molecules (say in a liquid or gas) was discovered by K S Krishnan and C V Raman in 1928.

- Molecules are made of several atoms bound together to form the analogue of a rigid body (like a dumbbell for a diatomic molecule or a tetrahedron etc).
- Monochromatic (single wavelength) light (around the visible, near IR or near UV range) is scattered off a molecule. Most of the scattered light has the same frequency as the incident light – this is elastic scattering (Raleigh scattering).
- A small fraction of the light is inelastically scattered (Raman effect). Inelastic means the energy (and frequency ν_e) of the emitted light is different from that of the incident light (ν_i).
- Moreover, the ‘Raman shift’ in the frequency $\nu_e - \nu_i$ was found to take on only a few discrete values, which are characteristic of the molecule.
- Classically it isn’t possible to explain these discrete Raman shifts in frequency of the incident light.
- Quantum mechanical explanation: a molecule can only be in one of several discrete ‘energy levels’. These may be rotational or vibrational or electronic energy levels. So the rotational and vibrational motions of molecules also occur at discrete energy levels just like the electronic energy levels of individual atoms. Moreover, it is often the case that the vibrational and rotational energy levels are more easily excited as they are lower lying than the electronic excitations.
- When a photon is incident on a molecule, it ‘excites’ the molecule from energy E_0 , to a temporary/virtual state of energy E_{temp} . Then the molecule relaxes back to a final energy E_1 (typically a vibrational or rotational excitation), which may be different from E_0 , and emits a photon. The energy gained by the molecule is $E_1 - E_0$ which corresponds to the shift in the frequency of light $h(\nu_i - \nu_e)$.
- Most of the emitted light is of course unchanged in wavelength (elastic scattering). Some of the light was red-shifted (energy lost to the atom), and these cases are called Stokes lines, where the atom relaxes to a state of higher energy than the ground state where it started from. Occasionally, an atom that was already in an excited state would absorb a photon and then emit a photon while relaxing to the ground state. In this case, the out going light is of a higher frequency than the incoming light. This is called the anti-Stokes line.
- Thus the Raman shift in frequency encodes the differences in the energy levels of the molecule. It is possible to identify a material using the Raman shifts, they serve as a fingerprint. This method is called Raman spectroscopy and is very widely used in chemistry and physics.

5.5 de Broglie hypothesis on matter waves

- We have seen that light (radiation) can display both wave like (interference) and particle-like (photons - photo electric effect) properties. However, the wave-like properties of light are pronounced only on length scales comparable to the wavelength of the light. Otherwise light appears to travel like a beam of bullets.
- In 1924 de Broglie suggested that electrons, (which had been discovered by J J Thompson through their particle-like properties) could also display wave-like properties. Moreover, it was suggested that the typical wavelength of electrons studied so far was rather small, which is why its wave-like properties had not been noticed. However, a careful experiment should be able to detect the diffraction of electrons around edges and interference between coherent

electron beams.

- He suggested that all forms of matter, not just electrons, displayed both wave and particle characteristics. He proposed that a particle of momentum p had associated with it a ‘matter-wave’ of wave-length λ , such that

$$\lambda = \frac{h}{p}, \quad \text{where } h \text{ is Planck's constant} \quad (85)$$

Or, if we define the ‘wave number’ $k = 2\pi/\lambda$, then

$$p = \frac{hk}{2\pi} = \hbar k, \quad \text{where } \hbar = \frac{h}{2\pi}. \quad (86)$$

- One motivation for the de Broglie formula is that the same is true for photons:

$$E = pc = h\nu \Rightarrow p = \frac{h\nu}{c} = \frac{h}{\lambda} \quad (87)$$

- We can also motivate the de Broglie hypothesis using the Bohr quantization rule for angular momentum $L = n\hbar$ for the allowed orbits of electrons around the atom. For a circular orbit at radius r and momentum p , $L = rp$, so $2\pi r = n\frac{h}{p}$. If we now interpret $\frac{h}{p}$ as the wave length of the electron wave, following de Broglie, then the Bohr quantization rule says that there are an integer number of wavelengths around the circumference of the orbit. In other words, The electron wave is a standing wave (“stationary state”) that does not “destructively interfere with itself”.

- Another way of arriving at the de Broglie hypothesis is to demand that the wave packet that describes a matter particle (e.g. electron) must have a group speed equal to the speed of the particle (at least in the classical limit). Let us see how to implement this idea for a free non-relativistic particle of mass m and momentum p . We need $v_g = \frac{\partial\omega}{\partial k} = p/m$. Now by definition, $\omega = 2\pi\nu$. For photons, $E = h\nu$ and by analogy we suppose the same is true for free matter particles as well. In other words $E = \hbar\omega$. Thus $\frac{p}{m} = \frac{1}{\hbar} \frac{\partial E}{\partial k}$. For a free particle, $E = p^2/2m$, so that

$$\frac{p}{m} = \frac{1}{\hbar m} p \frac{\partial p}{\partial k} \Rightarrow \frac{\partial p}{\partial k} = \hbar \Rightarrow p = \hbar k = \frac{h}{\lambda}. \quad (88)$$

Thus the de Broglie relation $p = \hbar k$ relates the wave vector of a matter wave to the momentum of the matter particle in such a way that the group speed of the wave packet is equal to the velocity of the particle $v_g = p/m$.

- de Broglie originally obtained the relation $\lambda = h/p$ for the wavelength of a matter wave associated to a particle of momentum p by considering a possibly relativistic free particle. It is instructive to calculate the phase and group speeds of the matter wave packet of a (possibly) relativistic particle. Above we replace E by γmc^2 and p by γmv where $\gamma = 1/\sqrt{1 - v^2/c^2}$. Then¹⁰

$$\omega = \frac{\gamma mc^2}{\hbar} = \frac{mc^2}{\hbar\sqrt{1 - v^2/c^2}} \approx \frac{mc^2}{\hbar} + \frac{mv^2}{2\hbar}, \quad k = \frac{\gamma mv}{\hbar} = \frac{mv}{\hbar\sqrt{1 - v^2/c^2}} \approx \frac{mv}{\hbar} + \frac{mv^3}{2\hbar c^2}. \quad (89)$$

¹⁰We do not need the approximate formulae.

So the group and phase speeds are (using $\frac{\partial \gamma}{\partial v} = v\gamma^3/c^2$, $\hbar \frac{\partial \omega}{\partial v} = mv\gamma^3$ and $\hbar \frac{\partial k}{\partial v} = \gamma m + mv^2\gamma^3/c^2$)

$$v_g = \frac{\partial \omega}{\partial k} = \frac{\partial \omega}{\partial v} \frac{\partial v}{\partial k} = \frac{mv\gamma^3}{\gamma m + m \frac{v^2}{c^2} \gamma^3} = \frac{v}{\frac{v^2}{c^2} + \frac{1}{\gamma^2}} = v. \quad (90)$$

$$v_p = \frac{\omega}{k} = \frac{\gamma mc^2}{\hbar} \frac{\hbar}{\gamma mv} = \frac{c^2}{v} = \frac{c^2}{v_g} \quad (91)$$

The phase speed of the de Broglie wave of a massive relativistic particle exceeds the speed of light. Nevertheless, the group speed $v_g = v$ (which is the speed at which the particle moves, and transmits information and energy) is less than the speed of light.

- The de Broglie hypothesis that electrons should display wave-like characteristics was spectacularly confirmed within a few years by Davisson & Germer and Thomson who showed that electrons are diffracted by a crystal in the same way as light is.
- Indeed, all forms of matter and radiation (even atoms and molecules), display both wave and particle characteristics. This is sometimes called wave-particle duality.
- In retrospect, we can say that electrons are neither classical waves nor particles, they just behave quantum mechanically. In certain circumstances, their properties may be usefully described using particle language and in other instances, using wave ideas.

5.6 Electron Diffraction: Davisson-Germer-Thompson experiment

- Bragg diffraction (1913) concerns reflection of x-rays of wavelength λ against a crystalline solid (say Nickel). The crystal is made of atoms arranged in a regular array. There are parallel planes in the crystal on which the atoms lie, called the crystal planes, separated by a distance d .
- The Braggs (Lawrence and Henry, father and son) found that for the reflected beam intensity to be strong, the angle of incidence must equal the angle of reflection¹¹. But this alone was not enough to observe a peak in the diffraction, because the waves reflected from different atoms (e.g. atoms in the first and second planes) could destructively interfere even if the angle of incidence matched the angle of reflection. For constructive interference, the difference in path length for waves reflected off the first and second plane would have to be an integer multiple of the wavelength.
- The reflected beam displays an interference pattern of high and low intensities, with peaks occurring at specific angles θ (with respect to the plane of the crystal) given by Bragg's formula for a crystal whose planes are separated by a distance d

$$n\lambda = 2d \sin \theta, \quad n = 1, 2, 3, \dots \quad (92)$$

In this case, the reflected waves arrive in phase and constructively interfere.

- If the wave length of the X-rays is fixed and the crystal has been chosen, then d, λ are fixed. Then Bragg's formula says that there is a peak in the reflected intensity precisely if the angle of the crystal (w.r.to the incident beam) is θ_1 where $2d \sin \theta_1 = \lambda$. There is a second Bragg

¹¹Photons emerge at other angles too, but their intensity is substantially less due to non-constructive interference.

peak (of lower intensity) at θ_2 where $2d \sin \theta_2 = 2\lambda$. The Bragg peaks for higher values of n are also present but decrease in intensity.

- In the Davisson-Germer-Thomson experiment (1927), instead of light, electrons of energy E (about 54 eV) were reflected off a Nickel crystal, and an interference pattern was observed in the reflected beam. Moreover, the peak intensity was observed at just the angle predicted by Bragg's law, provided one assigned a wavelength to electrons given by de Broglie's formula $\lambda = h/p$, where $E = p^2/2m$.

- Let us find the de Broglie wavelength of electrons accelerated through a potential difference of 54 Volts, i.e. electrons with 54 eV energy. Such electrons have a momentum given by

$$E = 54eV = \frac{p^2}{2m} = \frac{p^2}{2 \cdot 511KeV/c^2} \Rightarrow p = 7429eV/c. \quad (93)$$

The corresponding wavelength is

$$\lambda = \frac{h}{p} = \frac{4.135 \times 10^{-15} eV s}{7429eV/c} = \frac{4.135 \cdot 10^{-15} eV s \times 3 \cdot 10^8 m/s}{7429 eV} = 1.67 \times 10^{-10} m \quad (94)$$

The distance between planes in the Nickel crystal is $d = .91$ Angstroms (a fact known from x-ray diffraction). So according to the de Broglie hypothesis, the first electron diffraction peak should be seen at an angle of $\sin \theta = 1.67/1.82$ or $\theta = 66.5$ degrees, and this was indeed the case.

- The de Broglie wavelength of heavier objects (e.g. a tennis ball, say 60g) moving at speeds we are familiar with (100 km/h) is incredibly small compared to human length scales. The wave-like aspects (interference and diffraction) of tennis balls have not yet been detected.

$$\lambda_{dB} = \frac{h}{p} = \frac{6.6 \cdot 10^{-34} Js}{.06 \times 10^3 / 36 kg m/s} = \frac{6.6 \times 10^{-34} \times 36m}{60} = 3.96 \times 10^{-34} m. \quad (95)$$

To detect the first diffraction maximum $2d \sin \theta = \lambda$ we will need a crystal with a spacing on the order of $10^{-34}m$, which is not available. If the de Broglie wavelength is very small compared to the crystal plane spacing $\lambda \ll d$, then the diffraction peaks are all concentrated at grazing incidence $\theta_n \approx 0$ and very hard to detect. By decreasing the mass of the particle, we increase the de Broglie wavelength to be comparable to that of atomic crystals (or other diffraction gratings) so that diffractive effects can be observed (in particular the diffraction peaks are well separated).

5.7 Double slit interference experiment: wave function as probability amplitude

- Young's double slit experiment (1803) established the wave nature of light.
- The corresponding experiment with water waves, bullets and electrons is very instructive in clarifying the difference between the behaviour of classical particles and waves on the one hand, and quantum mechanical objects on the other hand. (See Feynman lectures vol 3.)
- We have a source directed rightwards towards a detection screen that absorbs what ever hits it. But along the way, there is a wall parallel to the screen with two holes/slits S_1 and S_2 . The slits are a distance a apart and for simplicity, equally spaced on either side of the perpendicular from the source to the screen. The screen is a distance d behind the wall. The

screen is equipped with detectors. We will focus on the two-dimensional plane that is visible when the apparatus is looked at side-on.

- **EXPERIMENT WITH PARTICLES:** First, suppose the source produces a stream of classical particles (say bullets or small balls) fired randomly at all angles in the general direction of the screen. The width of the slits is several times the size of the bullets. Initially only slit S_1 is open. We count the number of bullets arriving at each point x on the screen in unit time. This is called the intensity $I_1(x)$. After a lot of bullets arrive, $I_1(x)$ is peaked at a point on the screen close to S_1 and decays in either direction up and downwards. Similarly the intensity distribution of bullets through S_2 is $I_2(x)$. Now suppose both slits are open. The intensity distribution is found to be $I(x) = I_1(x) + I_2(x)$. This is understandable in classical mechanics since each bullet that reaches the screen has a precise trajectory that takes it either through S_1 or S_2 . So the number of particles arriving at x in unit time is the sum of those that come through S_1 and those that come through S_2 . To avoid the possible collisions of bullets, they could be sent with a short gap between successive bullets. From the relation $I(x) = I_1(x) + I_2(x)$ we say that the bullets (classical particles) do not interfere with each other. Moreover, the particles arrive individually at the screen, the particles are lumpy and at most one detector records an arrival at any instant.

- **EXPERIMENT WITH WAVES:** Next we repeat the experiment with classical waves (eg water waves by immersing the apparatus in a lake, or an intense beam of monochromatic electromagnetic waves). We will assume the waves are described by an amplitude $\psi(x, t)$, which could be the height of the water wave or an appropriate component of the electric field. The wave intensity at any point is defined as the square of the amplitude $|\psi|^2$. The detector detects the intensity of the wave (or the energy, which is proportional to the intensity). We can regard S_1 and S_2 as virtual sources for spherical waves $\psi_{1,2} = \frac{\sin\{kr_{1,2}-\omega t\}}{r_{1,2}}$ where r_1 and r_2 are the distances from S_1 and S_2 . First, slit S_1 alone is kept open and the time average $I_1(x) = |\psi_1|^2$ is measured. Approximately the same distribution $I_1(x)$ as for bullets is obtained. The difference is that there is no lumpiness in the detection of intensity, it is a fairly smooth function of x and time. Similarly, when S_2 alone is open, $I_2 = |\psi_2|^2$. When both slits are open, the detected intensity is found not to be $I_1 + I_2$ but $I(x) = |\psi_1 + \psi_2|^2$. We say that the waves through the two slits have interfered. A sequence of bands of high and low intensity are detected on either side of a central maximum. The locations of the maxima along the screen are given by points whose distances r_1, r_2 from the slits differ by an integer multiple of the wavelength $\lambda = 2\pi/k$. Let us calculate the intensity at a point \vec{r} on the screen

$$\begin{aligned}
 I(\vec{r}, t) &= |\psi_1 + \psi_2|^2 = \frac{\sin^2(kr_1 - \omega t)}{r_1^2} + \frac{\sin^2(kr_2 - \omega t)}{r_2^2} + \frac{2 \sin(kr_1 - \omega t) \sin(kr_2 - \omega t)}{r_1 r_2} \\
 &= I_1 + I_2 + \frac{1}{r_1 r_2} [\cos(k(r_1 - r_2)) - \cos(k(r_1 + r_2) - 2\omega t)] \tag{96}
 \end{aligned}$$

This follows from trigonometry $\sin A \sin B = \frac{1}{2}(\cos(A - B) - \cos(A + B))$. Now let us average the intensity over time. The last term averages to zero, but the first three survive. In the first two terms, the square of the sine function averages to one-half. The third term is time independent and is responsible for the interference pattern

$$\langle I(r_1, r_2) \rangle_{\text{time average}} = \langle I_1 \rangle + \langle I_2 \rangle + \frac{1}{r_1 r_2} \cos(k(r_1 - r_2)) \tag{97}$$

We see that the intensity falls off far from the slits (r_1, r_2 large). In addition, the time-average intensity is maximal (interference peak) if $k(r_1 - r_2) = 2n\pi$, with $n \in \mathbb{Z}$ which is where the cosine is one. In other words, $r_1 - r_2 = n\lambda$: the path difference for waves through the two slits to reach the point on the screen must be an integer multiple of the wavelength. The interference minima occur if $k(r_1 - r_2)$ is an odd multiple of π . For example, the n^{th} interference maximum is at a height $h_n \approx \frac{n\lambda d}{a}$ above and below the central maximum. Here we suppose the distance between the slits is a and the distance between the slits and the detection screen is $d \gg a$ ¹².

- **EXPERIMENT LOW INTENSITY LIGHT - PHOTONS (OR ELECTRONS):** When the intensity of light is reduced sufficiently, the intensity of light received by the detectors becomes lumpy and particle-like. The energy arrives in quanta of light (photons) each with the same energy $E = h\nu$ and same momentum $p = \hbar k$, corresponding to massless particles. So photons behave like particles. Nevertheless, the intensity profile that develops after many of these photons have hit the screen with both slits open matches the interference pattern characteristic to waves, $I(x) = |\psi_1 + \psi_2|^2$. In particular, fewer photons arrive at an interference minimum with both slits open than with a single slit open! So despite arriving individually, we cannot just add the intensities I_1 and I_2 to find $I(x)$. Despite its lumpiness, we cannot ascribe to a photon a trajectory that took it either through S_1 or S_2 , as we could do for bullets. Indeed, if we try to find out which slit the photon came through before it hits the screen (by positioning a detector each near each slit), then the interference pattern is lost, and the intensity returns to that for classical bullets $I(x) = I_1(x) + I_2(x)$.

- What is more, since $I(x)$ matches the intensity for waves, we suppose that there must be some wave function $\psi(x)$ (associated with the electrons or the photons) whose absolute square is the intensity. By the way we have defined it, $I(x, t)dx$ is the probability of photons arriving in an interval $[x, x+dx]$ on the screen¹³. Of course, to build up reasonable statistics, the time over which we average must be large compared to the time between the arrival of successive photons on the screen. We call the wave function $\psi(x, t)$ a probability amplitude, its absolute square is a probability density. Note that this wave function is not the electric or magnetic field in the case of photons. Electromagnetic waves predicted by Maxwell's equations are purely classical waves. The probability amplitude (wave function) $\psi(x, t)$ is a new quantum mechanical concept introduced to understand the interference of photons or electrons or whatever else (e.g. atoms or molecules) the source is producing. For matter particles, the probability amplitude is to be identified with the matter wave introduced by de Broglie. For example, $\psi_1(x, t)$ is called the amplitude for the electron to reach x at time t through slit S_1 . The idea of interpreting the wave function of matter waves as a probability amplitude is due to Max Born (1926-27), however, that came after the formulation of Schrödinger's equation.

¹²If θ is the angle of elevation of the interference peak from the horizontal at the midpoint of the slits, then $h/d = \tan \theta \approx (r_1 - r_2)/a$ from which $h_n \approx n\lambda d/a$

¹³Our previous definition was not quite a probability since we just counted the number of bullets/photons arriving in the region covered by a detector per unit time, so one must divide by the total number of bullets arriving anywhere per unit time to make it a probability

5.8 Wave-Particle Duality and Complementarity Principle

- We saw that both light (radiation) and matter (electrons, atoms) display both wave-like and particle-like properties in different situations. This was called wave-particle duality.
- However, in a given experimental situation, if the wave properties are detected, then it is not possible to simultaneously describe the situation using discrete particle-like lumps. This complementary (mutually exclusive) manifestation of wave and particle properties was called the complementarity principle by Neils Bohr. For example, we cannot explain the discrete lumpy arrivals of single photons at the detector screen as the arrival of a continuous wave. On the other hand, in the same double slit experiment, the interference pattern that develops is attributed to wave like behavior but cannot be explained using classical particle trajectories.

6 Schrödinger equation, formulation and postulates of quantum mechanics

6.1 Obtaining Schrödinger's equation from matter wave hypothesis

- According to de Broglie's hypothesis, there is a matter wave associated to a massive particle. Suppose the matter wave is described by the amplitude $\psi(x, t)$. A natural question is how such a matter wave evolves in time. Does it satisfy a wave equation? This is indeed the case, it satisfies the Schrödinger wave equation (obtained by him in 1926). Recall that matter waves are dispersive and cannot satisfy the simplest wave equation $c^{-2}\ddot{\psi} = \psi''$ for oscillating strings or sound waves.
- We will use the results of the de Broglie hypothesis to infer the time evolution of the wave function of a free particle and then find which equation the wave function must satisfy¹⁴
- Suppose at $t = 0$, the wave function of an electron is a wave packet $\psi(x)$. Being a wave packet, we write it as a superposition of several monochromatic waves of different wave lengths

$$\psi(x, t = 0) = \int_{-\infty}^{\infty} \frac{dk}{2\pi} \tilde{\psi}(k) e^{ikx}. \quad (98)$$

The de Broglie hypothesis (by analogy with photons) implied that the angular frequency of such a matter wave packet was given in terms of the particle energy by (using $p = \hbar k$)

$$\omega = \frac{E}{\hbar} = \frac{p^2}{2m\hbar} = \frac{\hbar k^2}{2m} \quad (99)$$

This means each wavelength mode must oscillate in time according to $e^{ikx - i\omega(k)t} = e^{ikx - i\frac{\hbar k^2}{2m}t}$. Fourier synthesizing these oscillations,

$$\psi(x, t) = \int \frac{dk}{2\pi} \tilde{\psi}(k) e^{ikx - i\frac{\hbar k^2}{2m}t} \quad (100)$$

Let us denote $\frac{dk}{2\pi} = [dk]$. Differentiating under the integral,

$$\dot{\psi}(x, t) = -\frac{i\hbar}{2m} \int [dk] k^2 \tilde{\psi}(k) e^{ikx - i\omega(k)t}$$

¹⁴This is basically a reverse engineering of the equation from the knowledge of a solution on physical grounds.

$$\psi''(x, t) = - \int [dk] k^2 \tilde{\psi}(k) e^{ikx - i\omega(k)t}. \quad (101)$$

Thus, the wave function of a free non-relativistic particle of mass m satisfies the linear partial differential Schrödinger wave equation

$$i\hbar \frac{\partial \psi}{\partial t} = - \frac{\hbar^2}{2m} \frac{\partial^2 \psi}{\partial x^2} \quad (102)$$

• We notice that the effect of differentiating with respect to time in $\int [dk] \tilde{\psi}(k) e^{i(kx - \omega(k)t)}$ is to multiply the integrand in Fourier space by $-i\omega(k)$. Differentiating with respect to x has the effect of multiplying the integrand in Fourier space by ik . So we will say that $\frac{\partial}{\partial t}$ is represented in Fourier space by multiplication by $-i\omega(k)$ and $\frac{\partial}{\partial x}$ is represented in Fourier space by multiplying by ik . Along with the de Broglie relations $E = \hbar\omega, p = \hbar k$ for a free particle, this leads to the correspondence¹⁵

$$\begin{aligned} \frac{\partial}{\partial t} &\leftrightarrow -i\omega(k) = -iE/\hbar \Rightarrow i\hbar \frac{\partial}{\partial t} \leftrightarrow E. \\ \frac{\partial}{\partial x} &\leftrightarrow ik = ip/\hbar \Rightarrow -i\hbar \frac{\partial}{\partial x} \leftrightarrow p. \end{aligned} \quad (103)$$

Let us introduce differential operators \hat{E} and \hat{p} that represent the energy and momentum:

$$\hat{E} = i\hbar \frac{\partial}{\partial t}, \quad \hat{p} = -i\hbar \frac{\partial}{\partial x} \quad (104)$$

Then we see that the Schrödinger equation can be written as

$$\hat{E}\psi(x, t) = \frac{\hat{p}^2}{2m}\psi(x, t) \quad (105)$$

We notice that the operator on the right corresponds to the classical kinetic energy of the particle. Now, the generalization of the Schrödinger equation to a particle that is moving under the influence of force derived from a potential $V(x)$ is quite natural

$$\hat{E}\psi(x, t) = \left(\frac{\hat{p}^2}{2m} + V(x) \right) \psi(x, t) \quad (106)$$

Here $V(x)\psi(x, t)$ just denotes the multiplication of functions. To be more precise, we could introduce another operator \hat{x} which acts on functions of x via multiplication by x and write

$$\hat{E}\psi(x, t) = \left(\frac{\hat{p}^2}{2m} + V(\hat{x}) \right) \psi(x, t). \quad (107)$$

The operator on the right that corresponds to the energy of the particle is called the hamiltonian operator in the quantum theory

$$\hat{H} = \frac{\hat{p}^2}{2m} + V(\hat{x}), \quad \text{so} \quad \hat{E} \psi(x, t) = \hat{H} \psi(x, t). \quad (108)$$

¹⁵ \leftrightarrow is to be read as *is represented in Fourier space as*

Explicitly, the Schrödinger equation for a non-relativistic particle moving in one dimension under the influence of a potential $V(x)$ is the linear PDE

$$i\hbar \frac{\partial \psi}{\partial t} = -\frac{\hbar^2}{2m} \frac{\partial^2 \psi}{\partial x^2} + V(x)\psi(x). \quad (109)$$

- More generally, if the particle was moving in three dimensional space,

$$i\hbar \frac{\partial \psi(x, y, z, t)}{\partial t} = -\frac{\hbar^2}{2m} \left(\frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} + \frac{\partial^2 \psi}{\partial z^2} \right) + V(x, y, z)\psi(x, y, z, t). \quad (110)$$

The Schrödinger equation was obtained using the matter wave hypothesis and extended to include forces by analogy with classical mechanics. Whether it is a correct physical law of time evolution is to be settled by giving $\psi(x)$ a physical meaning and then comparing its predictions with experiments (aside from checking its internal mathematical consistency). To date, predictions based on the Schrödinger equation have been in line with all experimental tests.

- What we saw above for momentum, energy and position extends to other classical *observables* (quantities like the various components of angular momentum, L_x, L_y, L_z, L^2). In the quantum theory, each is replaced by an operator that acts on the wave function. Energy and momentum were represented by differential operators when acting on $\psi(x)$ (e.g. $\hat{p}_y \psi(x, y, z, t) = -i\hbar \frac{\partial \psi}{\partial y}$), while position is represented by a multiplication operator $\hat{x}\psi(x) = x\psi(x)$.

6.2 Wavefunction, probability density and current

- We say that the state of a quantum mechanical system is specified by the wave function ψ . For this reason the wavefunction is also called the state function. In fact, the wave function contains the maximum amount of information that can be specified about the state of a quantum mechanical system.
- For a particle moving in 3-space, the wave function is $\psi(x, y, z; t)$. For a system of n particles, the wave function is a function of three coordinates of each of the n particles $\psi(\vec{r}_1, \vec{r}_2, \dots, \vec{r}_n; t)$. In other words, the wave function is a (possibly time-dependent) function on the classical configuration space of the system.
- A useful alternate specification of the state of a quantum system is via the Fourier transform of the wave function, which is a function of the momenta (or wavenumbers) rather than position. For a system with one degree of freedom,

$$\tilde{\psi}(k) = \int dx e^{-ikx} \psi(x) \quad \text{and} \quad \psi(x) = \int \frac{dk}{2\pi} e^{ikx} \tilde{\psi}(k). \quad (111)$$

Writing $p = \hbar k$ we have

$$\tilde{\psi}(p/\hbar) = \int dx e^{-ipx/\hbar} \psi(x) \quad \text{and} \quad \psi(x) = \int \frac{dp}{2\pi\hbar} e^{ipx/\hbar} \tilde{\psi}(p/\hbar). \quad (112)$$

- This is to be compared with the state of a classical mechanical system being 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, to be seen later.

- As indicated, according to Born, 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, $|\tilde{\psi}(p, t)|^2 \frac{dp}{2\pi}$ is the probability of finding the particle in momentum interval $[p, p + dp]$ at time t .

- Suppose n identical particles are prepared in the same quantum mechanical state $\psi(x)$. Then a measurement of the position of each particle gives a (possibly) different result (this is an experimental fact). Born's statistical interpretation is that as $n \rightarrow \infty$, the distribution of position measurements approaches the probability density $|\psi(x)|^2$.

- To qualify as a probability density, the total probability of finding the particle anywhere must be one. In other words, we need $\int dx |\psi(x, t)|^2 = 1$. The quantity $\int dx |\psi(x, t)|^2 = \|\psi\|^2$ is called the square of the norm of ψ , and if this is finite we say that ψ is square-integrable.

- The Schrödinger equation is linear in ψ , so if $\psi(x, t)$ is a solution, so is any multiple $A\psi(x, t)$, $A \in \mathbb{C}$. If ψ is square-integrable, the normalization constant A is chosen¹⁶ so that $\|\psi\|^2 = 1$.

- However, there could be a problem. For consistency, the total probability of finding the particle *somewhere* must remain equal to one at all times. This is indeed the case, as is checked using the Schrödinger equation

$$\begin{aligned} i\hbar \partial_t \int_{-\infty}^{\infty} \psi^* \psi \, dx &= i\hbar \int (\psi_t^* \psi + \psi^* \psi_t) \, dx \\ &= \int \left(\frac{\hbar^2}{2m} \psi^{*''} - V \psi^* \right) \psi \, dx + \int \psi^* \left(-\frac{\hbar^2}{2m} \psi'' + V \psi \right) \, dx \\ &= \frac{\hbar^2}{2m} \int_{-\infty}^{\infty} [\psi^{*''} \psi - \psi^* \psi''] \, dx = \frac{\hbar^2}{2m} \int_{-\infty}^{\infty} \partial_x (\psi^{*'} \psi - \psi^* \psi') \, dx = \text{0} \end{aligned} \quad (113)$$

For the last equality, we assumed the wave function vanishes sufficiently fast at $\pm\infty$, which was necessary for the total probability to be finite in the first place. The same holds for more than one degree of freedom. 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. It asserts that the total probability of the particle being found anywhere remains one for all time.

- In fact more can be said. Probability flows like a fluid, it can't jump from one location to another, it must flow like a current. 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). Let $P(x, t) = |\psi(x, t)|^2$ and based on the previous calculation, define the probability current density

$$j(x, t) = \frac{\hbar}{2mi} (\psi^* \psi' - \psi^{*'} \psi), \quad \text{then} \quad \partial_t P(x, t) + \partial_x j(x, t) = 0. \quad (114)$$

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

¹⁶Of course, only the modulus $|A|$ is determined by normalization. We are free to multiply a wave function of norm one by a phase $e^{i\alpha}$, $\alpha \in \mathbb{R}$ with out changing the norm.

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:

$$\partial_t \int_{x_0}^{x_1} P(x) dx + \int_{x_0}^{x_1} \frac{\partial j(x)}{\partial x} dx = 0 \quad \Rightarrow \quad \partial_t \int_{x_0}^{x_1} P(x) dx = j(x_0) - j(x_1) = 0 \quad (115)$$

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 .

$$\begin{aligned} P(\vec{r}, t) &= \psi^*(\vec{r}, t)\psi(\vec{r}, t), & \vec{j} &= \frac{\hbar}{2mi} [\psi^* (\nabla\psi) - (\nabla\psi^*) \psi] = \frac{\hbar}{m} \Im\psi^* \nabla\psi \\ \partial_t P(\vec{r}, t) + \nabla \cdot \vec{j}(x, t) &= 0, & \text{i.e.} & \quad \frac{\partial \rho}{\partial t} + \frac{\partial j_1}{\partial x} + \frac{\partial j_2}{\partial y} + \frac{\partial j_3}{\partial z} = 0. \\ \partial_t \int_{\Omega} P(\vec{r}, t) d^3r + \int_{\Omega} d^3r \nabla \cdot \vec{j} &= 0 & \text{or} & \quad \partial_t \int_{\Omega} P(\vec{r}, t) d^3r + \int_{\partial\Omega} \vec{j} \cdot d\vec{S} = 0. \end{aligned} \quad (116)$$

$d\vec{S}$ is the outward pointing area element on the bounding surface $\partial\Omega$. We used the divergence/Gauss/Stokes theorem to write the volume integral of a divergence as a surface integral.

6.3 Superposition principle and vector space of states

- The intensity pattern from the double slit experiment indicates that interference results from the superposition of wave functions. So it is natural to add probability amplitudes. Indeed this is elevated to a principle of quantum mechanics, the superposition principle. If an electron can be in a state with amplitude $\psi_1(x)$ and also one with amplitude $\psi_2(x)$, then it can also be in a state with amplitude $\psi_1(x) + \psi_2(x)$. Moreover, if an electron can be in a state $\psi_1(x)$, it can also be in a state $A\psi_1(x)$ for any complex constant A , indeed the two states are the same insofar as they lead to the same probability distributions for electron positions. So it is natural to form linear combinations of state functions with complex coefficients, this is the superposition principle.

- We also see that the superposition principle is mathematically consistent with the Schrödinger equation. Indeed, if by a state we mean a particular solution of the Schrödinger equation, then on account of its linearity, linear combinations of solutions are still solutions of the Schrödinger equation.

- In other words, the space of states of a quantum mechanical system is a linear (vector) space. The possible wavefunctions $\psi(x, t)$ at any fixed time t are elements of a vector space: we can add vectors and multiply them by scalars (complex numbers). The Schrödinger equation tells us how the state changes in time from an initial vector $\psi(x, t = 0)$. Indeed, once the initial state $\psi(x, 0)$ has been specified, the Schrödinger equation uniquely determines the state $\psi(x, t)$ at all subsequent times.

- However, the vector space of wavefunctions $\psi(x)$ of a quantum mechanical particle in one dimension is infinite dimensional. It is the space V of complex valued functions of one real variable x . You can think of the components of the vector ψ in the ‘position-basis’ as the

values of the function $\psi(x')$ at each of the real positions x' . More formally,

$$\psi(x) = \int dx \delta(x - x') \psi(x') \quad (117)$$

- Here we regard the Dirac delta ‘functions’ (distributions) $D_{x'}(x) = \delta(x - x')$ for various values of x' as forming a basis for the above vector space V . We can think of $D_{x'}(x)$ as a function of x that is concentrated at $x = x'$. We will discuss the Dirac delta function shortly.
- If the state vector was expressed in momentum space $\tilde{\psi}(k)$, then the components of this vector in the momentum basis are the values $\tilde{\psi}(k)$

$$\tilde{\psi}(k) = \int \frac{dk'}{2\pi} 2\pi \delta(k - k') \tilde{\psi}(k') \quad (118)$$

6.4 Dirac Delta function

- The Dirac delta function/distribution $\delta(x)$ is to be thought of as a function that is concentrated at the point $x = 0$ (infinite at $x = 0$ and zero elsewhere). Similarly, $\delta(x - x')$ as a function of x is concentrated at the point $x = x'$.
- We can think of the Delta function $\delta(x)$ as the limit $a \rightarrow 0$ of a sequence of functions $\delta_a(x)$ that get taller and more concentrated in smaller neighborhoods of $x = 0$ as $a \rightarrow 0$. However, the sequence of functions must be chosen carefully so that the area under the graph of $\delta_a(x)$ is one (or at least approaches one as $a \rightarrow 0$). For example, we could take the wall-like functions

$$w_a(x) = \frac{1}{a} \theta\left(-\frac{a}{2} \leq x \leq \frac{a}{2}\right), \quad \delta(x) = \lim_{a \rightarrow 0} w_a(x). \quad (119)$$

Here $\theta(\cdot)$ is the indicator function of the interval. We notice that $\int dx w_a(x) = 1$ for each $a > 0$. The limit of this sequence of functions as $a \rightarrow 0$ is the Dirac delta function.

- Another model for $\delta(x)$ is as a limit of gaussians of increasing height and decreasing width

$$g_a(x) = \frac{1}{a\sqrt{\pi}} e^{-x^2/a^2}, \quad \delta(x) = \lim_{a \rightarrow 0} g_a(x). \quad (120)$$

The factors have been chosen so that the area under the graph of $g_a(x)$ is always one.

- Yet another model uses the sequence of functions $\frac{1}{\pi x} \sin(x/w)$ of decreasing width w

$$\delta(x) = \lim_{w \rightarrow 0} \frac{\sin(x/w)}{\pi x} = \lim_{w \rightarrow 0} \frac{1}{w} \frac{\sin(x/w)}{\pi(x/w)} \quad (121)$$

In this case too, (e.g. by contour integration) the area under the curve is one for all widths w ¹⁷

$$\int_{-\infty}^{\infty} \frac{\sin(x/w)}{\pi x} dx = \frac{1}{\pi} \int \frac{\sin y}{y} dy = \frac{1}{\pi} [\text{Si}(\infty) - \text{Si}(-\infty)] = 1. \quad (122)$$

- The delta function is defined by its behavior upon integration with respect to a ‘test’ function $\psi(x)$ (e.g. a continuous function or a smooth function that decays sufficiently fast

¹⁷ $\text{Si}(x) = \int_0^x \frac{\sin(y)dy}{y}$ is the Sine integral function. It asymptotes to $\pm\pi/2$ as $x \rightarrow \pm\infty$.

at $\pm\infty$)

$$\text{Definition: } \int dx \psi(x) \delta(x) = \psi(0) \quad \text{for all test functions } \psi(x) \quad (123)$$

In particular, $\int dx \delta(x) = 1$ so it encloses unit area. This is why we imposed the unit area condition on each of the above models for the Dirac δ -function.

- Any proposed property of the delta function is considered true if it is true upon integrating with respect to any test function. For example, the delta function is even $\delta(-x) = \delta(x)$. Using $x \rightarrow -x$ change of variable of integration,

$$\int \delta(x)\psi(x) = \psi(0) = \int \delta(-x)\psi(x) \quad (124)$$

- The usual methods of manipulating integrals, such as change of integration variable and integration by parts etc are used to derive additional properties of the delta function.
- For example $\delta'(x)$ is defined by using the rules of integration by parts

$$\int dx \delta'(x) f(x) = - \int dx \delta(x) f'(x) = -f'(0). \quad (125)$$

$\delta'(x)$ is an odd function.

- The delta function can be regarded as the derivative of the Heaviside step function (i.e. this formula is valid upon integrating with respect to any test function)

$$\frac{d}{dx}\theta(x > 0) = \delta(x) \quad (126)$$

- $\delta(ax) = \frac{1}{|a|}\delta(x)$ as is checked by the change of integration variable $y = ax$.
- A particularly useful representation of the delta function¹⁸ is as a Fourier integral¹⁹

$$\int \frac{dk}{2\pi} e^{ikx} = \delta(x) \quad (127)$$

Let us establish this representation of the delta function using the $\sin(x/w)/\pi x$ model.

$$\int_{-\infty}^{\infty} [dk] e^{ikx} = \lim_{\Lambda \rightarrow \infty} \int_{-\Lambda}^{\Lambda} e^{ikx} [dk] = \lim_{\Lambda \rightarrow \infty} \int_{-\Lambda}^{\Lambda} \cos(kx) [dk] = \lim_{\Lambda \rightarrow \infty} \frac{2 \sin(\Lambda x)}{2\pi x} = \lim_{w \rightarrow 0} \frac{\sin(x/w)}{\pi x} = \delta(x).$$

- We can use this representation of the delta function to establish the Fourier inversion formula. Suppose $f(x) = \int [dk] e^{ikx} \tilde{f}(k)$. Then multiplying by e^{-ilx} and integrating with respect to x , we get

$$\int f(x) e^{-ilx} dx = \int [dk] dx \tilde{f}(k) e^{i(k-l)x} = \int [dk] \tilde{f}(k) 2\pi \delta(k-l) = \tilde{f}(l) \quad (128)$$

Thus (relabeling $k \leftrightarrow l$) we have the inverse Fourier transform $\tilde{f}(k) = \int dx e^{-ikx} f(x)$.

¹⁸We saw why this is true in the lecture. The waves destructively interfere to give zero, due to the cancellation of phases for any $x \neq 0$. $x = 0$ is the only point where all the waves constructively interfere to produce an infinite integral.

¹⁹And by a relabeling of variables, we also have $\int dx e^{ikx} = 2\pi \delta(k)$

- We use this representation of the delta function to check that the Fourier transform preserves the norm of a function.

$$\int |\psi(x)|^2 dx = \int |\tilde{\psi}(k)|^2 \frac{dk}{2\pi} \quad (129)$$

Proof (exercise: Start with the RHS and express $\psi(k)$ and $\psi^*(k)$ as Fourier integrals). In particular if ψ is normalized to one in position space then it is also normalized to one in momentum(k)-space.

6.5 Average (expectation) values of observables in a state

- If we make position measurements on a large sample of electrons, all prepared in the same state $\psi(x)$, then we get different values in each measurement²⁰. However, the distribution of the position measurements is found to follow the probability density $P(x) = |\psi(x)|^2$.
- Thus, the average position that is measured (for an electron that is in the normalized state $\psi(x)$) is given by the *expectation value*

$$\langle x \rangle_\psi = \int_{-\infty}^{\infty} x |\psi(x)|^2 dx = \int \psi^*(x) x \psi(x) dx \quad (130)$$

Similarly, for a particle in three dimensions, the average y -coordinate is given by the expectation value $\langle y \rangle = (\int y |\psi(x, y, z)|^2 dy) / (\int |\psi(y)|^2 dy)$.

- If the electron is in the normalized state $\tilde{\psi}(k)$, ($\int [dk] |\tilde{\psi}(k)|^2 = 1$) then its average momentum is

$$\langle p \rangle = \langle \hbar k \rangle_{\tilde{\psi}} = \int_{-\infty}^{\infty} \hbar k |\tilde{\psi}(k)|^2 \frac{dk}{2\pi}. \quad (131)$$

- But what if the electron is in the state $\psi(x)$ of norm one? What is its average momentum? It turns out that (using Fourier transforms) we can re-write the above formula as

$$\langle p \rangle_\psi = \int \psi^*(x) \left(-i\hbar \frac{\partial}{\partial x} \right) \psi(x) dx \quad (132)$$

To see this, we write $\psi^*(x)$ and $\psi'(x)$ as Fourier integrals

$$-i\hbar \int dx \psi^*(x) \psi'(x) = \hbar \int [dk dl] \tilde{\psi}^*(l) \tilde{\psi}(k) k \int dx e^{i(k-l)x} = \int [dk] |\tilde{\psi}(k)|^2 \hbar k = \langle p \rangle. \quad (133)$$

- The moral of the story is that to compute the expectation value of an observable A which is represented on wave functions $\psi(x)$ by the (differential) operator \hat{A} , we must evaluate the integral

$$\langle A \rangle_\psi = \frac{\int \psi^*(x) \hat{A} \psi(x) dx}{\int |\psi(x)|^2 dx}. \quad (134)$$

²⁰This is *not* the same as repeatedly measuring the position of one electron. If we repeat the same position measurement in quick succession on the same electron, we get the same position as on the first occasion. On the other hand, if we measure the position of an electron in the state $\psi(x)$ once, we are more likely to get a value of x for which $|\psi(x)|^2$ is large than a value of x for which $|\psi(x)|^2$ is small. Repeating the measurement (soon) on an electron that has been found to be at x just confirms that it is at x .

We will explain this formula in the next section. Based on these two examples, we can get a formula for the average energy (expectation value of hamiltonian in a state of norm one)

$$\langle E \rangle = \int dx \psi^*(x) \left(\frac{\hat{p}^2}{2m} + V(\hat{x}) \right) \psi(x) = \int dx \psi^*(x) \left(-\frac{\hbar^2}{2m} \psi''(x) + V(x)\psi(x) \right) \quad (135)$$

As a consequence of Schrödinger's equation, we have another formula for the average energy

$$\langle E \rangle = \int \psi^*(x, t) \left(i\hbar \frac{\partial}{\partial t} \right) \psi(x, t) dx, \quad \text{if } \psi \text{ satisfies the Schrödinger equation.} \quad (136)$$

6.6 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!

For example $\frac{d\langle x \rangle}{dt} = \frac{\langle p \rangle}{m}$, so the average position evolves in the same way as given by the first of Hamilton's equations. This is checked by a small calculation

$$\begin{aligned} i\hbar \frac{\partial \langle x \rangle}{\partial t} &= i\hbar \frac{\partial}{\partial t} \int \psi^*(x, t) x \psi(x, t) dx = i\hbar \int (\dot{\psi}^* x \psi + \psi^* x \dot{\psi}) dx \\ &= \frac{\hbar^2}{2m} \int (\psi^{*''} x \psi - \psi^* x \psi'') dx = \frac{\hbar^2}{2m} \int 2\psi^* \psi' dx \end{aligned} \quad (137)$$

We have used the Schrödinger equation (the terms involving the potential cancel out) and then integrated by parts twice assuming $\psi(x)$ vanishes at $x \rightarrow \pm\infty$ sufficiently fast. We conclude that

$$\frac{d\langle x \rangle}{dt} = \frac{1}{m} \int \psi^* \left(-i\hbar \frac{\partial}{\partial x} \right) \psi dx = \frac{\langle p \rangle}{m}. \quad (138)$$

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 on its classical motion.

- By a similar calculation (exercise) we can find the time evolution of the mean momentum. It turns out to be related to the expectation value of the force (negative gradient of the potential):

- Another such calculation 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.

6.7 Hilbert space of states, Dirac bra-ket notation

• To understand the formula for the expectation value of an observable in a normalized state, $\langle A \rangle = \int \psi^*(x) \hat{A} \psi(x) dx$, we return to the observation that the state space is a vector space. But the space of wave functions $\psi(x)$ it is more than just a vector space, it has an inner product (dot product), which makes it a Hilbert space $\mathcal{H} = L^2(R)$

$$\langle \psi | \phi \rangle = \int \psi^*(x) \phi(x) dx \quad (139)$$

Notice that upon taking the complex conjugate, $\langle \psi | \phi \rangle^* = \langle \phi | \psi \rangle$. In particular, $\langle \psi | \psi \rangle = \int |\psi(x)|^2 dx$.

• Think of the ket-vector $|\psi\rangle$ as a column vector and the bra-vector $\langle\psi|$ as its adjoint (its complex-conjugate transposed, a row vector) $\langle\psi| = |\psi\rangle^\dagger$.

• The ket-vectors span the Hilbert space \mathcal{H} of column vectors. We say that the bra-vectors span the dual Hilbert space \mathcal{H}^* of row-vectors.

• Given an orthonormal basis $|e_i\rangle_{i=1}^\infty$ for the Hilbert space of ket-vectors, we also have a dual basis $\langle e_i|$ for the space of bra-vectors defined by $\langle e_i | e_j \rangle = \delta_{ij}$. As $\dim \mathcal{H} = \infty$ we need infinitely many basis vectors.

• Given an orthonormal basis of unit vectors $|e_i\rangle$, $\|e_i\| = 1$, the components of a ket-vector $|\psi\rangle$ in this basis are the inner products $\langle e_i | \psi \rangle$. This is to say $|\psi\rangle = \sum_i |e_i\rangle \langle e_i | \psi \rangle$.

• Similarly, the components of $\langle\phi|$ in the dual basis are the inner products $\langle\phi| e_i\rangle$. This is to say $\langle\phi| = \sum_i \langle\phi| e_i\rangle \langle e_i|$.

• A useful basis for ket-vectors is the position basis $|x'\rangle$ ²¹. $|x'\rangle$ is the vector that is localized at the position x' . This basis vector has components $\delta(x - x')$ in the position basis, i.e.,

$$\langle x | x' \rangle = \delta(x - x'). \quad (140)$$

So $|x'\rangle$ and $|x\rangle$ are orthogonal if $x \neq x'$. But they are not unit vectors, $\| |x\rangle \| = \infty$, we say they are delta-normalized. In the position basis, the components of a general ket-vector $|\psi\rangle$ are $\psi(x)$:

$$\langle x | \psi \rangle = \psi(x) \quad (141)$$

Taking the complex conjugate, the components of the bra-vector $\langle\psi|$ in the position basis are $\psi^*(x)$:

$$\langle \psi | x \rangle = \psi^*(x) \quad (142)$$

Moreover, any function can be written as a linear combination of delta functions $\phi(x) = \int \delta(x - x') \phi(x') dx'$. This means the position basis is complete, any vector can be expanded in terms of them. Let us write this formula in bra-ket notation. For all $|\phi\rangle$ and x we have

$$\langle x | \phi \rangle = \int \langle x | x' \rangle \langle x' | \phi \rangle dx' = \langle x | \left(\int |x'\rangle \langle x'| dx' \right) | \phi \rangle \quad (143)$$

²¹Strictly speaking, the position ‘basis vectors’ $|x'\rangle$ are not even elements of \mathcal{H} , since they have infinite norm. In fact, they are better thought of as elements of the dual space of linear functionals on \mathcal{H} . So strictly speaking, the ‘position basis’ is not a basis at all! Nevertheless they are a very useful concept. We must learn to work with them without worrying about technicalities at this stage.

In other words,

$$\int |x'\rangle\langle x'| dx' = I = \text{identity operator} \quad (144)$$

This is called the completeness relation. It is the analogue of the statement that the sum of outer products of the standard basis vectors of \mathbb{R}^n with themselves is the identity matrix $\sum_{i=1}^n e_i e_i^T = I$.

- Now just as matrices can act on column vectors to the right and produce other column vectors, operators like $\hat{x}, \hat{p}, \hat{p}^2, \hat{H}$ can act on ket-vectors to produce new ket-vectors. For example, in the position basis, the components of $\hat{x}|\phi\rangle$ are $x\phi(x)$ i.e.,

$$\langle x|\hat{x}|\phi\rangle = \langle x|\hat{x}\phi\rangle = x\phi(x). \quad (145)$$

- Similarly, $\hat{p}|\phi\rangle$ has the components $-i\hbar\frac{\partial\phi}{\partial x}$ in the position basis i.e.,

$$\langle x|\hat{p}|\phi\rangle = \langle x|\hat{p}\phi\rangle = -i\hbar\frac{\partial\phi}{\partial x}. \quad (146)$$

We could also consider the component of $\hat{x}|\phi\rangle$ in the direction of $|\phi\rangle$ itself, i.e.,

$$\frac{\langle\phi|\hat{x}|\phi\rangle}{\langle\phi|\phi\rangle} = \frac{\int\phi^*(x)x\phi(x)dx}{\int\phi^*(x)\phi(x)dx} \quad (147)$$

- Terminology from linear algebra: the matrix elements of a linear transformation $\hat{A}: V \rightarrow V$ in the orthonormal basis e_i are $A_{ij} = \langle e_i|\hat{A}|e_j\rangle$. In particular, the diagonal matrix element in the state e_i is $A_{ii} = \langle e_i|\hat{A}|e_i\rangle$ (no sum on i).

- Thus the average (expectation) value of \hat{x} in the state ϕ is the same as the diagonal matrix element of \hat{x} in the normalized state $|\phi\rangle$.

$$\langle x\rangle = \frac{\int\phi^*(x)x\phi(x)dx}{\int\phi^*(x)\phi(x)dx} = \frac{\langle\phi|\hat{x}|\phi\rangle}{\langle\phi|\phi\rangle} \quad (148)$$

More generally, the expectation value of any observable \hat{A} in the state $|\phi\rangle$ is $\langle A\rangle = \frac{\langle\phi|\hat{A}|\phi\rangle}{\langle\phi|\phi\rangle}$.

6.8 Position eigenstates

- An electron is prepared in some state $|\psi\rangle$ with the position space wave function $\psi(x)$. Suppose the position of the electron is measured and the value $x = x'$ is obtained²². It is found that if the measurement is repeated soon after, the same value of position is obtained. Repeating the measurement of position on the same electron, reveals the same location. We interpret this fact by saying that after the first measurement of position, the state of the electron has changed ('collapsed') from $|\psi\rangle$ to a new state with a definite value of position, i.e., one where the probability density is concentrated at $x = x'$ so that a measurement of position gives x' with 100% certainty. But what is a state with a definite value of position $x = x'$? It is one that is sharply localized at $x = x'$. We have already met such a state, the position basis state $|x'\rangle$. Such position basis states have a special relationship with the position operator, they are its eigenstates. This motivates our discussion of eigenstates.

²²Of course, other values of x could also have been obtained, the possible values are distributed according to $|\psi(x)|^2$

- Let us first look for a function $f_{x'}(x)$ that is an eigenfunction of \hat{x} with eigenvalue x'

$$x f_{x'}(x) = x' f_{x'}(x) \quad \text{for all } x \quad \Rightarrow (x - x') f_{x'}(x) = 0 \quad \forall x \quad (149)$$

This means $f_{x'}(x)$ must vanish at all values of x except possibly at $x = x'$. If it vanished at x' as well, it would be the zero function and not an eigenfunction. If it had a finite and non-zero value at $x = x'$, $f_{x'}(x) = a$, then it would still be a vector of zero norm, indeed $\langle f_{x'} | f_{x''} \rangle = 0$ for all x', x'' . So these candidate eigenfunctions are orthogonal but all have norm zero. So in a sense, a must be infinite if we want the norm to be non-zero. There is unfortunately no such function in the ordinary sense of the word, but the Dirac delta function will do, so we take $f_{x'}(x) = \delta(x - x')$. Then we immediately see that $\langle f_{x'} | f_{x''} \rangle = \int \delta(x - x') \delta(x - x'') dx = \delta(x' - x'')$. So the eigenstates of position are $f_{x'}(x) = \delta(x - x')$. These eigenfunctions are orthogonal, but do not have finite norm, their inner product is delta-normalized $\langle f_{x'} | f_{x''} \rangle = \delta(x' - x'')$. In terms of the position basis states introduced earlier, $\langle x'' | f_{x'} \rangle = \delta(x'' - x')$. So $f_{x'}$ is the same as the position basis vector $|x'\rangle$.

- Alternatively, for any vector $|\phi\rangle$ the components of $\hat{x}|\phi\rangle$ in the position basis are $x\phi(x)$, i.e.,

$$\langle x | \hat{x} | \phi \rangle = x \phi(x) = x \times \text{'component of } |\phi\rangle \text{ in direction of } |x\rangle \text{' } \quad (150)$$

In particular, taking $|\phi\rangle = |x'\rangle$, a position basis vector,

$$\langle x | \hat{x} | x' \rangle = x \times \text{'component of } |x'\rangle \text{ in direction of } |x\rangle \text{' } = x \delta(x - x') = x' \delta(x - x') \quad (151)$$

But we also know that $\langle x | x' \rangle = \delta(x - x')$. So the components of $\hat{x}|x'\rangle$ and the components of $|x'\rangle$ in the position basis are the same upto a proportionality factor equal to x' . So the vectors themselves must be proportional:

$$\hat{x}|x'\rangle = x'|x'\rangle \quad (152)$$

So the position basis state $|x'\rangle$ is an eigenstate of the position operator \hat{x} with eigenvalue x' .

- We can say that the position operator is diagonal in the position basis

$$\langle x' | \hat{x} | x'' \rangle = x' \delta(x' - x'') \quad (153)$$

- In particular, if an electron is prepared in the position eigenstate $|x'\rangle$, then a measurement of the position of the electron must yield the value x' with certainty!

• Postulate (*collapse of wave function*): If we measure the position of a particle, then after the measurement, the particle's wavefunction (state) collapses into an eigenfunction (eigenstate) of the position operator, say $|x'\rangle$. The value of position obtained is the corresponding eigenvalue x' . If a measurement of position is repeated quickly²³, the state of the particle remains the same eigenstate $|x'\rangle$ and one gets the same value x' .

- In particular, the possible values of position that one can obtain in a position measurement are the eigenvalues of the position operator (this happens to be all real values of x' , so this is no restriction).

²³We say quickly, for if we wait for some time, then the state of the particle will change according to the Schrodinger equation and we are no longer guaranteed to measure the same position. The condition of quickly repeating the measurement was also mentioned in the experimental discussion that motivated this postulate in the first place.

6.9 Momentum eigenstates

Now let us look for the eigenstates of the momentum operator $\hat{p} = -i\hbar\frac{\partial}{\partial x}$. Let $\hat{k} = -i\frac{\partial}{\partial x}$ so that $\hat{p} = \hbar\hat{k}$. Then the eigenvalue problem for \hat{k} is

$$\hat{k}\psi(x) = k\psi(x) \Rightarrow -i\psi' = k\psi \Rightarrow \psi(x) = Ae^{ikx} \quad (154)$$

Let us choose the normalization $A = 1$ (this is to ensure the δ -normalization of momentum eigenstates, see below). Thus the eigenstate of momentum \hat{p} with eigenvalue $\hbar k$ is a plane wave e^{ikx} . Let us denote the ket-vector which is an eigenstate of \hat{k} with eigenvalue k as $|k\rangle$, i.e., $\hat{k}|k\rangle = k|k\rangle$. So our result can be stated as: the components of $|k\rangle$ in the position basis are e^{ikx} :

$$\langle x|k\rangle = e^{ikx} \quad \text{and taking the complex conjugate,} \quad \langle k|x\rangle = e^{-ikx} \quad (155)$$

- The momentum eigenstates are orthogonal and δ -normalized just like the position eigenstates. To see this we use the completeness of position eigenstates

$$\langle k'|k\rangle = \int \langle k'|x\rangle \langle x|k\rangle dx = \int e^{-ik'x} e^{ikx} dx = 2\pi\delta(k - k') \quad (156)$$

The momentum eigenstates $|k\rangle$ are the momentum-space analogue of the position basis $|x\rangle$.

- The momentum operator is diagonal in the momentum basis

$$\langle k'|\hat{p}|k\rangle = \hbar k \langle k'|k\rangle = \hbar k 2\pi \delta(k - k') \quad (157)$$

- If we measure the momentum of an electron that is in the momentum eigenstate $|k'\rangle$ (with eigenvalue k'), then we are sure to get the value $\hbar k$.
- In particular, the expectation value of momentum in a momentum eigenstate is

$$\frac{\langle k'|\hat{p}|k'\rangle}{\langle k'|k'\rangle} = \hbar k' \quad (158)$$

- We can expand any function of momentum as a linear combination of $2\pi\delta(k - k')$

$$\tilde{\psi}(k) = \int [dk'] 2\pi\delta(k - k') \tilde{\psi}(k') \quad (159)$$

$\tilde{\psi}(k)$ are the components of the vector ψ in the momentum basis

$$\langle k|\psi\rangle = \tilde{\psi}(k). \quad (160)$$

So in bra-ket notation we have for any vector $|\psi\rangle$ and any k

$$\langle k|\psi\rangle = \int_{-\infty}^{\infty} \langle k|k'\rangle \langle k'|\psi\rangle \frac{dk'}{2\pi} \quad (161)$$

It follows that the momentum eigenstates are complete

$$\int_{-\infty}^{\infty} |k'\rangle \langle k'| \frac{dk'}{2\pi} = I = \text{identity operator} \quad (162)$$

6.10 Heisenberg commutation relation and uncertainty

- Consider the relationship between the position and momentum eigenstates²⁴

$$\delta(x - x') = \int [dk] e^{ik(x-x')} \Leftrightarrow \langle x|x' \rangle = \int \langle x|k \rangle \langle k|x' \rangle [dk] \quad \text{i.e. } |x' \rangle = \int |k \rangle \langle k|x' \rangle [dk] \quad (163)$$

In other words, a position eigenstate is a linear combination of all the momentum eigenstates with coefficients $e^{ik(x-x')}$ of unit modulus. A position eigenstate is completely localized in position space, but completely delocalized in momentum space. So we would expect measurements of the position of electrons prepared in momentum eigenstates to give all position values with equal probability (This is indeed what is experimentally found)! On the other hand, the measurement of the momentum of a particle that is prepared in a momentum eigenstate $|k'\rangle$ is sure to give $\hbar k'$. This is a symptom of the uncertainty principle. If the momentum is precisely determined, the position cannot and vice versa.

- Another aspect is that the position operator and momentum operator are not simultaneously diagonalizable. They are diagonal in different bases. This implies that they cannot commute. Indeed let us calculate the commutator. We get (I is the identity operator)

$$[\hat{x}, \hat{p}] = \hat{x}\hat{p} - \hat{p}\hat{x} = i\hbar I \quad (164)$$

- A consequence of this non-vanishing commutator is the uncertainty principle (proof postponed). It is possible to show using the Cauchy-Schwarz inequality, that for position and momentum x and p ,

$$(\Delta x)^2 (\Delta p)^2 \geq -\frac{1}{4} \langle [x, p] \rangle_\psi^2 \quad (165)$$

which implies Heisenberg's uncertainty principle

$$\Delta x \Delta p \geq \frac{\hbar}{2}. \quad (166)$$

Here $(\Delta x)^2 = \langle (\hat{x} - \langle \hat{x} \rangle)^2 \rangle$ is the variance of the position in the state ψ and Δp is the standard deviation of \hat{p} in the same state. Roughly, it says that the position and momentum of a quantum mechanical particle in a state ψ cannot both be simultaneously measured with arbitrary precision.

- A way to interpret the uncertainty principle: Suppose we prepare several electrons in the same state ψ and try to make position measurements on half of them and momentum measurements on the other half. Then the product of the spread (standard deviation) in the position measurements and spread in momentum measurements is bounded below by $\hbar/2$.
- Another interpretation: Suppose we make a position measurement (with some accuracy), then the wave function collapses to one that is sharply peaked about the measured value of position. Such a wave packet has a very large spread of wave numbers and therefore a momentum measurement can lead to widely different values.
- Of course, Planck's constant is a rather small unit of action, so we can get reasonably accurate simultaneous position and momentum measurements for macroscopic bodies.
- Similarly, the operation of multiplication by time t and the energy operator $\hat{E} = i\hbar \frac{\partial}{\partial t}$ do not commute, $[t, \hat{E}] = -i\hbar I$. But we cannot derive $\Delta E \Delta t \geq \hbar/2$ in the same way as above. The time-energy uncertainty relation is of a different sort and will be discussed later.

²⁴ $\delta(x) = \int [dk] e^{ikx} = \int [dk] e^{-ikx}$ is an even function of x . All integrals are from $-\infty$ to ∞ unless otherwise specified.

6.11 Hermiticity of observables $\hat{x}, \hat{p}, \hat{H}$

- The expectation values of x, p, H are the average position, momentum and energy of a quantum mechanical particle in a state ψ . To correspond to the average values measured in experiments, these numbers must be real. Let \hat{A} be the operator representing an observable, i.e. a quantity that can be measured. Its expected value is $\langle A \rangle = \langle \psi | \hat{A} | \psi \rangle$. We must ensure that observables are such that their expectation values (in every state) are real.

- If a quantum system is prepared in an eigenstate of \hat{A} , with eigenvalue a , then a measurement of A will give a with certainty. So a special case of the above requirement is that our observables must have real eigenvalues.

- An operator \hat{A} , all of whose expectation values (with respect to a given inner product) are real ($\langle \phi | \hat{A} | \phi \rangle \in \mathbb{R}$ for all ϕ) is called hermitian (with respect to the relevant inner product). We will see that $\hat{x}, \hat{p}, \hat{H}$ are all hermitian with respect to the standard L^2 inner product.

- By extension, we postulate that all quantum observables be represented by hermitian operators.

- There is an equivalent definition of hermiticity which is often more useful²⁵. An operator \hat{A} is hermitian²⁶ if for every pair of vectors ψ, ϕ

$$\langle \psi | \hat{A} \phi \rangle = \langle \hat{A} \psi | \phi \rangle \quad (167)$$

Note that this can also be written $\langle \psi | \hat{A} \phi \rangle = \langle \phi | \hat{A} \psi \rangle^*$ or $\hat{A}_{\psi\phi} = \hat{A}_{\phi\psi}^*$. The matrix elements of \hat{A} (between any pair of states) must be symmetric after complex conjugation.

- For an operator A we can define the adjoint A^\dagger as the operator whose matrix elements are

$$\langle \psi | A^\dagger \phi \rangle = \langle A \psi | \phi \rangle = \langle \phi | A \psi \rangle^*, \quad \text{or} \quad (A^\dagger)_{\psi\phi} = (A_{\phi\psi})^*. \quad (168)$$

So a (formally) self-adjoint operator $A^\dagger = A$ is hermitian²⁷. It also follows that the adjoint of a product of operators is $(AB)^\dagger = B^\dagger A^\dagger$. In particular this means $(A^2)^\dagger = (A^\dagger)^2$.

- We can check that with this equivalent definition, the expectation values of a hermitian operator are automatically real

$$\langle A \rangle = \langle \psi | \hat{A} \psi \rangle = \langle A \psi | \psi \rangle = \langle \psi | A \psi \rangle^* = \langle A \rangle^* \quad (169)$$

- In particular, the eigenvalues of hermitian operators are automatically real.

- We check that $\hat{x}, \hat{p}, \hat{p}^2, V(\hat{x}), \hat{H}$ are all hermitian. For example, $x_{\phi\phi} = \int |\phi|^2 x dx = (x_{\phi\phi})^*$. Or $\langle \psi | \hat{x} \phi \rangle = \int x \psi^* \phi dx = \langle \hat{x} \psi | \phi \rangle$. Hermiticity of \hat{p} is checked, by an integration by parts, assuming ψ vanishes at $\pm\infty$.

- The square of a hermitian operator is also hermitian $(A^2)^\dagger = A^\dagger A^\dagger = A^2$. So the kinetic energy $\hat{p}^2/2m$ is hermitian. So is the potential energy operator $V(\hat{x})$ for any real potential $V(x)$. So the hamiltonian for a particle in the potential $V(x)$ is hermitian.

²⁵To get this definition, put $\psi = u + v$ and $\psi = u + iv$ in $\langle \psi | \hat{A} \psi \rangle = \langle \hat{A} \psi | \psi \rangle$ simplify and add the resulting equations.

²⁶Hermitian operators are often called symmetric operators in the mathematical literature.

²⁷The concepts of self-adjointness and hermiticity are the same for operators on a finite-dimensional vector space. In the infinite dimensional case, to be self-adjoint, an operator A must be hermitian and A and its adjoint A^\dagger must satisfy certain conditions on their domains of definition

- An important property of hermitian operators is that their eigenfunctions can be chosen to form an orthonormal basis for the Hilbert space. So we can expand any state $|\psi\rangle$ as a linear combination of eigenvectors. For example, if the energy spectrum is discrete and $|\phi_n\rangle_{n=0}^{\infty}$ are the eigenstates of the hamiltonian, then $|\psi\rangle = \sum_n c_n |\phi_n\rangle$.

6.12 Collapse of the wavefunction/measurement and probability postulate

- The collapse of wave function postulate is now extended from position measurements to the measurement of other observables. If we measure an observable \hat{A} (e.g. energy/momentum/position etc) then the value we get must be an eigenvalue a of A . The state of the particle collapses to the eigenstate of \hat{A} corresponding to the eigenvalue a . All quick subsequent measurements of A will give the same value a , and will leave the particle in the same eigenstate.

- In particular, if we measure an observable A in an eigenstate $|\phi_a\rangle$ of that observable ($\hat{A}|\phi_a\rangle = a|\phi_a\rangle$), we will get the corresponding eigenvalue a with 100 % certainty.

- In fact more can be said. *Probability interpretation postulate* (refined version of Born's probability interpretation): Suppose the particle is in the normalized state $|\psi\rangle$ and we measure the observable A . Then the measured value must be an eigenvalue a of A and the measurement leaves the system in the corresponding eigenstate $|\phi_a\rangle$. The probability p_a of getting the particular eigenvalue a is the absolute square of the projection of $|\psi\rangle$ on the eigenstate $|\phi_a\rangle$. In other words, $p_a = |\langle\phi_a|\psi\rangle|^2 = |c_a|^2$, where $c_a = \langle\phi_a|\psi\rangle$ and we assumed that $\|\phi_a\| = 1$. In other words, the probability of getting eigenvalue a is equal to the absolute square of the component of $|\psi\rangle$ in the direction of the eigenvector $|\phi_a\rangle$. This is reasonable. If the state $|\psi\rangle$ had zero projection on the eigenvector ϕ_a , then we can be sure that a measurement of A in the state ψ will not result in the value a . Let us check that the probabilities add up. The probability of getting any eigenvalue should be one. We will assume that the eigenstates of A form an orthonormal basis so that the state $|\psi\rangle$ can be written as a linear combination of them with coefficients c_a given earlier:

$$|\psi\rangle = \sum_a c_a |\phi_a\rangle \Rightarrow \langle\psi|\psi\rangle = 1 = \sum_a c_a \langle\psi|\phi_a\rangle = \sum_a c_a c_a^* = \sum_a p_a \quad (170)$$

In fact, this is strictly true only if a is a discrete eigenvalue of \hat{A} (e.g. if A is the hydrogen hamiltonian, then a would be one of the discrete bound state energies). If a is part of the continuous spectrum of A (e.g. a is the energy of an unbound electron from an ionized hydrogen atom), then the probability of the measurement falling in the range $[a, a + da]$ is given by $|\langle\phi_a|\psi\rangle|^2 da$ and we must replace the sums above by integrals over the continuous spectrum.

6.13 Energy eigenstates: Time-independent Schrödinger equation

- To understand the time evolution, we must solve the Schrödinger equation $i\hbar \frac{\partial\Psi}{\partial t} = \hat{H}\Psi$ where $\hat{H} = -\frac{\hbar^2}{2m} \frac{\partial^2\Psi}{\partial x^2} + V(x)\Psi(x)$. The initial condition is that $\Psi(x, t = 0) = \phi(x)$.

- This linear partial differential equation can be simplified by separation of variables. Suppose $\Psi(x, t) = \psi(x) T(t)$ is the product of a function of space and one of time. Of course, only very special solutions of Schrödinger's equation will have this form. But we will be able to synthesise the general solution as a linear combination of such product functions. Upon

separation of variables,

$$i\hbar\psi(x)\dot{T}(t) = T(t)\hat{H}\psi(x) \quad \Rightarrow \quad i\hbar\frac{\dot{T}(t)}{T(t)} = \frac{\hat{H}\psi(x)}{\psi(x)} \equiv E \quad (171)$$

Here we observe that the lhs is a function of time alone while the rhs is a function of x alone. So the only way these two functions could be equal for all x and t is for each of them to be equal to the same constant, which we call E .

- $T(t)$ is easily obtained: $T(t) = Ce^{-iEt/\hbar}$ where C is a (possibly complex) integration const.

- The spatial factor $\psi(x)$ is the solution of the time-independent Schrödinger equation $\hat{H}\psi = E\psi$.

$$-\frac{\hbar^2}{2m}\frac{\partial^2\psi(x,t)}{\partial x^2} + V(x)\psi(x) = E\psi(x) \quad (172)$$

This is the eigenvalue problem for the hamiltonian operator \hat{H} with eigenvalue E which we interpret as energy (by the measurement postulate).

- The Hamiltonian typically has an infinite number of eigenstates²⁸, called the energy eigenstates $\hat{H}\psi_n(x) = E_n\psi_n(x)$ corresponding to the energy eigenvalues (levels) E_n , the possible results of a measurement of energy. n could be a discrete or a continuous index.

- If the particle is initially in the n^{th} eigenstate of the hamiltonian $\hat{H}\psi_n = E_n\psi_n$, $\Psi(x, t = 0) = \psi_n(x)$, then the above factorization works

$$\Psi(x, t) = \psi_n(x)e^{-iE_nt/\hbar}. \quad (173)$$

The time-dependence is oscillatory and called harmonic time-dependence. It is periodic with angular frequency $\omega_n = E_n/\hbar$ and time period $T_n = 2\pi/\omega_n$.

- The wave function $\psi(x)$ is not directly experimentally observed. We only measure values of observables like position, energy etc.

- If $\psi_1(x)$ is the state of a quantum mechanical particle, then $\psi_2(x) = c\psi_1(x)$ represents the same state for any complex number c (which is independent of x). The reason for this identification is that the expectation value of any observable is the same in both the states. The constant c cancels:

$$\langle A \rangle_{\psi_2} = \frac{\langle \psi_2 | A | \psi_2 \rangle}{\langle \psi_2 | \psi_2 \rangle} = \frac{|c|^2 \langle \psi_1 | A | \psi_1 \rangle}{|c|^2 \langle \psi_1 | \psi_1 \rangle} = \langle A \rangle_{\psi_1} \quad (174)$$

So the state of a particle is given by a ray in a Hilbert space. A ray is the equivalence class of vectors that differ by multiplicative rescaling, e.g. the above vectors ψ_1, ψ_2 correspond to the same ray.

- Eigenstates ψ_n evolve in time by a multiplicative (time-dependent, but x -independent) phase $e^{-i\omega_n t}$. So the particle remains in the same state (n^{th} eigenstate) for all time. For this

²⁸If the Hilbert space is infinite dimensional, like $L^2(\mathbb{R})$ for a particle moving on a line, then the Hamiltonian has infinitely many linearly independent eigenstates. If the Hilbert space is finite dimensional, then H can have only a finite number of linearly independent eigenstates, as for the spin degrees of freedom of a particle. But in all cases, the eigenstates of a self-adjoint Hamiltonian form a basis for the vector space and can be taken to be orthonormal.

reason, the eigenstates are called stationary states. In particular, the probability distribution of possible positions of the particle do not change in time, if the particle starts out in an eigenstate of energy: $|\Psi_n(x, 0)|^2 = |\Psi_n(x, t)|^2$. The same is true also of the probability distribution of any other observable, say momentum $|\tilde{\Psi}_n(k, 0)|^2 = |\tilde{\Psi}_n(k, t)|^2$: Why so? Similarly, the distribution of energies also remains the same, initially, the particle was in a state of definite energy $E = E_n$ (for discrete energy levels, probability distribution of energies is the Kronecker- δ $\rho(E) = \delta_{E, E_n}$) and subsequently remains in a state of definite energy $E = E_n$.

- On account of linearity of the SE, we can superpose solutions to get new solutions. In particular, if the initial state is a linear sum of two eigenstates, $c_1\psi_1(x) + c_2\psi_2(x)$, then each eigenstate evolves independently at its own frequency

$$\Psi(x, t) = c_1\psi_1(x)e^{-i\omega_1 t} + c_2\psi_2(x)e^{-i\omega_2 t}. \quad (175)$$

- Note that such a $\Psi(x, t)$ need not be periodic any more, it will be periodic only if the two frequencies (or equivalently time periods T_1, T_2) are in rational ratio $\frac{T_1}{T_2} = \frac{p}{q}$ where p, q are natural numbers.
- More generally, given any initial state $\Psi(x, t = 0)$, we may expand it in the basis of energy eigenstates²⁹ $|\Psi(t = 0)\rangle = \sum_n c_n |\psi_n\rangle$. Then each eigenstate evolves independently so that

$$\Psi(x, t) = \sum_n c_n e^{-iE_n t/\hbar} \psi_n(x) \quad (176)$$

This is the general solution to the initial value problem for the Schrödinger equation (SE).

- Given the initial state $\Psi(t = 0)$ the Schrödinger equation determines uniquely the subsequent state of the system. However, we do not prove the uniqueness here.
- However, to find this general solution, we need the coefficients c_n and also the eigenfunctions $\psi_n(x)$ and corresponding eigenvalues E_n of the hamiltonian operator. c_n is easily obtained, it is the component of $\Psi(t = 0)$ in the direction of the energy eigenstate $\psi_n(x)$

$$c_n = \frac{\langle \psi_n | \Psi(t = 0) \rangle}{\langle \psi_n | \psi_n \rangle} \quad (177)$$

- The problem of solving the Schrödinger initial value problem has now been reduced to finding the energy eigenstates and eigenvalues, i.e., to the solution of the time-independent Schrödinger equation (172). We will see how this is done for various simple choices of potential $V(x)$. However, in general it is quite difficult to find the spectrum and eigenfunctions of the Hamiltonian and one has to resort to approximation methods to extract physically interesting information.

6.14 Summary of postulates of quantum mechanics

- The states of a quantum system are vectors in a Hilbert space \mathcal{H} . For a system of particles, \mathcal{H} is the space of square-integrable functions on the classical configuration space. Two state

²⁹This is always possible on account of the energy eigenstates forming an orthonormal basis for the state space

vectors that differ by a multiplicative complex constant represent the same physical state. So more precisely, the space of states of a quantum system are rays in a Hilbert space.

- Observables (such as the hamiltonian) are hermitian (more precisely self-adjoint) operators on \mathcal{H} .
- Time evolution of a state is given by Schrödinger's equation.
- Measurement of an observable A in a state ψ (of norm one) produces a real number that is one of the eigenvalues of A . The probability of getting eigenvalue a is equal to the absolute square of the projection of ψ on the corresponding normalized eigenstate ϕ_a , $p_a = |\langle \phi_a | \psi \rangle|^2$.

6.15 Unitarity of Schrödinger time evolution

- We saw that Schrödinger evolution preserves (in time) the total probability of finding the particle somewhere, i.e., it preserves the norm of a state.
- More is true, time evolution also preserves the inner product of any pair of states. Suppose ψ and ϕ evolve according to the SE. Then $\langle \psi(t) | \phi(t) \rangle = \langle \psi(0) | \phi(0) \rangle$ for all time t . In other words $\frac{\partial \langle \psi(t) | \phi(t) \rangle}{\partial t} = 0$. This has the geometric meaning that the length of a state vector and the angle between any pair of state vectors remains the same in time. To see this we just differentiate and use the SE for a particle in a real potential $V(x)$, assuming the wave functions vanish at $x = \pm\infty$

$$i\hbar \frac{\partial \langle \psi | \phi \rangle}{\partial t} = i\hbar \int \frac{\partial(\psi^* \phi)}{\partial t} dx = \frac{\hbar^2}{2m} \int_{-\infty}^{\infty} (\psi^{*''} \phi - \psi^* \phi'') dx = \frac{\hbar^2}{2m} \int_{-\infty}^{\infty} \frac{\partial}{\partial x} (\psi^{*'} \phi - \psi^* \phi') dx = 0$$

- Now let us consider the time evolution as a transformation from (any) initial state vector $|\psi(0)\rangle$ to the final state vector $|\psi(t)\rangle$. This transformation is linear in the initial state, on account of the linearity of the SE. So we can represent it by a linear operator $U(t)$ on the Hilbert space \mathcal{H} :

$$|\psi(t)\rangle = U(t)|\psi(0)\rangle \quad \text{and taking the adjoint,} \quad \langle \psi(t) | = \langle \psi(0) | U(t)^\dagger. \quad (178)$$

Now let us see what the preservation of inner products means for the time-evolution operator.

$$\langle \psi(0) | \phi(0) \rangle = \langle \psi(t) | \phi(t) \rangle = \langle \psi(0) | U^\dagger U \phi(0) \rangle \quad (179)$$

Thus $\langle \psi(0) | \phi(0) \rangle = \langle \psi(0) | U^\dagger U \phi(0) \rangle$. Since this is true for *any* pair of initial state vectors in the Hilbert space \mathcal{H} , it follows that $U(t)^\dagger U(t) = I$ is the identity operator, U is unitary. We say that Schrödinger time evolution is unitary: it preserves inner products of state vectors (i.e., lengths of state vectors and angles between state vectors).

- The time evolution operator (in most cases, an exception being when the hamiltonian is zero!) is not hermitian, so it is not an observable. Indeed, the eigenvalues of a unitary operator are complex numbers of unit modulus. Nevertheless, $U(t)$ is obviously important since it contains the solution of the Schrödinger equation. So it would be nice to have an expression for it.
- $U(t)$ is an operator, we can express it in any basis. Finding an expression for an operator means saying how it acts on the vectors in any basis for the Hilbert space. In the position basis $\langle x' | U(t) | x'' \rangle = U(t)_{x', x''}$. But finding the matrix entries of $U(t)$ in the position basis

is not easy. $U(t)$ is more simply expressed in the basis of energy eigenstates $\psi_n(x)$ ³⁰. This is because the energy eigenstates evolve very simply in time $\Psi_n(t) = e^{-iE_n t/\hbar}\psi_n(x)$ where $\psi_n(x) = \Psi_n(x, t=0)$ is the initial state, i.e.,

$$|\Psi_n(t)\rangle = U(t) |\Psi_n(0)\rangle = e^{-iE_n t/\hbar} |\psi_n\rangle \quad (180)$$

In other words, eigenstates of the Hamiltonian are also eigenstates of the time-evolution operator. So the time evolution operator is diagonal in the basis of energy eigenstates with diagonal entries $e^{-i\omega_n t}$

$$\langle \psi_n | U(t) | \psi_m \rangle = e^{-iE_n t/\hbar} \langle \psi_n | \psi_m \rangle = e^{-iE_n t/\hbar} \delta_{nm} \quad (181)$$

Equivalently, the time evolution operator is (check that it has the correct matrix elements)

$$\hat{U}(t) = e^{-i\hat{H}t/\hbar}. \quad (182)$$

The exponential of an operator is defined as $e^{\hat{A}} = \sum_{n=0}^{\infty} \frac{\hat{A}^n}{n!}$.

7 Square well, Free particle and Harmonic oscillator

- We seek intuition for how a quantum mechanical particle moving in one dimension behaves under the influence of simple potentials $V(x)$, by solving the Schrödinger equation.

7.1 Particle in a box : Infinite square well

- This is the quantum analogue of the classical problem of a particle in a box (e.g. a billiard ball constrained to move inside a frictionless table, where it feels no force except at the boundaries, which it cannot penetrate and just reflects off.). We consider the one-dimensional case, particle constrained to move within a segment of length L . The potential is $V(x) = 0$ for $0 < x < l$ and $V(x) = V_0 = \infty$ for $x \leq 0$ or $x \geq L$. This is called an infinite well as $V(x)$ is infinite outside a segment.
- As outlined above, given any initial state $\psi(x, t=0)$ we will expand it in the basis of energy eigenstates $\psi(x, t=0) = \sum_n c_n \psi_n(x)$ and then use the harmonic time evolution of the eigenstates to find how our initial state develops.

$$\psi(x, t) = \sum_n c_n \psi_n(x) e^{-iE_n t/\hbar}, \quad c_n = \int_0^L \psi_n^*(x) \psi(x, t=0) dx \quad (183)$$

So the main problem is to find the energy eigenstates ψ_n and energy eigenvalues E_n by solving

$$-\frac{\hbar^2}{2m} \psi''(x) = E \psi(x), \quad 0 < x < L. \quad (184)$$

³⁰Which we shall take to be orthonormal without loss of generality, and discrete for simplicity in this discussion.

- In addition to this, we must specify the boundary conditions, how does $\psi(x)$ behave at $x = 0, L$? The particle cannot be outside the interval³¹, so $\psi(x) = 0$ for $x < 0$ or $x > L$.
- We will postulate that the wave function must be continuous at the edge of the well³². A physical reason for this: $|\psi(x)|^2$ gives the probability density of finding the particle at x . If $\psi(x)$ was discontinuous at x , we could not unambiguously specify³³ the probability density for the particle to be found at x . Moreover, if ψ was discontinuous at x , then ψ' would be infinite at x and there can be difficulties in defining the probability current density $j(x)$. So we will assume $\psi(x)$ is continuous everywhere and in particular $\psi(0) = \psi(L) = 0$ since the probability density vanishes outside the well³⁴.
- If $\psi(x)$ is continuous and the potential $V(x)$ is finite (it could have finite discontinuities), then we can show that the gradient $\psi'(x)$ must also be finite. This is by integrating the SE between $x - \epsilon$ and $x + \epsilon$. We get

$$-\frac{\hbar^2}{2m} (\psi'(x + \epsilon) - \psi'(x - \epsilon)) = \int_{x-\epsilon}^{x+\epsilon} (E - V(x))\psi(x). \quad (185)$$

Since $\psi(x)$ is cts (and in particular finite) and $V(x)$ is finite (though it may be discontinuous), the integral on the rhs is the area under a (possibly discontinuous curve). But in the limit $\epsilon \rightarrow 0$, this area must vanish so we have $\psi'(x^+) = \psi'(x^-)$, i.e., ψ' is continuous at x .

- But if the potential has an infinite discontinuity, ψ' can have a (finite) discontinuity. For example, let us consider the vicinity of the left edge of the well. We have already argued that $\psi(x < 0) \equiv 0$, from which it follows that $\psi'(x < 0) \equiv 0$. Then integrating the SE gives

$$-\frac{\hbar^2}{2m}\psi'(\epsilon) = E \int_0^\epsilon \psi(x)dx + \int_0^\epsilon V(x)\psi(x)dx. \quad (186)$$

In the limit $\epsilon \rightarrow 0^+$, the first term on the rhs vanishes, since $\psi(x)$ is cts and the area under it will vanish as $\epsilon \rightarrow 0$. But $\int_0^\epsilon V(x)\psi(x)dx$ can take any value since $V(0) = \infty$ and $\psi(0) = 0$. This indicates that $\psi'(0^+)$ can take any value and is not restricted. We will see below that $\psi'(0^+)$ indeed takes several different values for the various stationary states of the system and if we do not restrict to stationary states, $\psi'(0^+)$ can take any finite value.

- The above second order ODE with constant coefficients

$$\psi''(x) = -k^2\psi(x), \quad k^2 = \frac{2mE}{\hbar^2} \quad (187)$$

has general solution consisting of a linear combination $A \sin(kx) + B \cos(kx)$. $\psi(0) = 0$ implies that $B = 0$ and $\psi(L) = 0$ implies that $k = \frac{n\pi}{L}$ for $n \in \mathbb{Z}$. In particular, k must be

³¹It is of course physically reasonable (classically) that the particle cannot be outside the well due to the infinite inward force at $x = 0, L$ and the infinite potential energy required to keep the particle outside the well. We can also see that $\psi(x) = 0$ is a solution of the time independent SE outside the well: $-\frac{\hbar^2}{2m}\psi''(x) = (E - V_0)\psi(x)$, provided we interpret $V_0\psi(x) = 0$ for $V_0 = \infty$ and $\psi(x) = 0$. This is reasonable if regarded as a limit. If V_0 were large (more than E) but finite, $\psi(x)$ would vanish exponentially fast $\psi(x) \propto \exp -\frac{2m(V_0-E)|x|}{\hbar^2}$ for x outside the well. So though $V_0 \rightarrow \infty$, $\psi(x) \rightarrow 0$ much faster and the product vanishes outside the well. This is the justification for taking $\psi(x) = 0$ outside the well.

³²Of course, ψ must be continuous and indeed twice differentiable to define ψ'' . But the issue is more delicate when $V(x)$ has an infinite discontinuity. At such points the ψ' could have a jump discontinuity.

³³Limit from the left or from the right would give different probability densities.

³⁴This can be experimentally tested. We could measure the distribution of positions of particles in a quantum dot and see that the distribution goes to zero at the boundary.

real, or $E > 0$. There are no non-zero solutions satisfying the boundary conditions if $E < 0$, in which case we have a linear combination of hyperbolic sines and cosines. Of course, this is physically reasonable, since we expect the classical energy of the particle at x to be at least as much as the potential energy $V(x)$, which is zero. So $E < 0$ is forbidden both classically and quantum mechanically for a particle in the square well. $n = 0$ gives a solution with zero norm (particle nowhere). So the eigenstates are $\psi_n(x) = A_n \sin(\frac{n\pi x}{L})$. Taking $A_n = \sqrt{2/L}$ ensures that the particle is somewhere, $\|\psi_n\| = 1$. We see that n and $-n$ give linearly dependent solutions so we may restrict to $n = 1, 2, 3, \dots$. Thus

$$\psi_n(x) = \sqrt{\frac{2}{L}} \sin\left(\frac{n\pi x}{L}\right), \quad E_n = \frac{n^2 \pi^2 \hbar^2}{2mL^2}, \quad n = 1, 2, 3, \dots \quad (188)$$

Plot the first few eigenstates. $n = 1$ is the ground state, $n = 2$ the first excited state etc. E_n are the only allowed values of energy, so the energy eigenvalues are infinite in number but discrete. Notice that $\psi'_n(0^+) = \sqrt{\frac{2}{L}} \frac{n\pi}{L}$ is non-zero.

- The stationary states are orthonormal, $\langle \psi_n | \psi_m \rangle = \delta_{nm}$ as $2 \sin A \sin B = \cos(A - B) - \cos(A + B)$. It follows that $\langle \psi_n | \hat{H} | \psi_m \rangle = \delta_{nm} E_n$. We sometimes denote the energy eigenstates $|\psi_n\rangle$ by $|n\rangle$.

- They are complete (this is a result from Fourier series), we can expand any state ψ as a linear combination of the energy eigenstates $\psi(x) = \sqrt{\frac{2}{L}} \sum_{n=1}^{\infty} c_n \sin(n\pi x/L)$ where $c_n = \langle \psi_n | \psi \rangle$. So

$$\sum_n |\psi_n\rangle \langle \psi_n| = I \quad (189)$$

- The ground state has no nodes (zeros) in the interval $(0, L)$, the first excited state has one zero, and so on: the n^{th} excited state $n - 1$ nodes. This is a general feature. Excited state wave functions have more nodes, they oscillate more often. n is called a quantum number. It labels the stationary states.

- As n grows, the highly energetic stationary state wave functions are highly oscillatory, $n/2$ wave lengths fit into the box. The wave length $\lambda_n = \frac{2\pi}{k_n} = \frac{2L}{n}$ becomes smaller and it becomes harder to detect the wave nature of the quantum particle. The probability of finding the particle in any small interval $(x - \delta x, x + \delta x)$ becomes nearly $2\delta x/L$, independent of x . This is what we see classically. If we have a free particle of given energy (so fixed speed) bouncing between two walls, it spends an equal amount of time in each such sub-interval of length $2\delta x$ in the region $[0, L]$. The classical limit is approached in the limit of large quantum number n , this is Bohr's correspondence principle.

- The ground state is even about $x = \frac{L}{2}$, the first excited state is odd about the mid point. $\psi_{2n}(x)$ are even and $\psi_{2n+1}(x)$ are odd about $x = L/2$ for $n \in \mathbb{N}$. This can be seen to be a consequence of a **parity** symmetry $x \mapsto L/2 - x$ of the hamiltonian.

- **Digression:** If the potential $V(x)$ is an even function (or even about some point $x_0 = L/2$ in this case), then we may always take the energy eigenfunctions $\psi_n(x)$ to be either even or odd (about x_0). Let us show this in the case $x_0 = 0$ for simplicity. Suppose $\psi(x)$ is an energy eigenstate with eigenvalue E , then

$$-\frac{\hbar^2}{2m} \psi''(x) + V(x)\psi(x) = E\psi(x) \quad (190)$$

We will now see that $\phi(x) = \psi(-x)$ is also an eigenfunction of \hat{H} with the same eigenvalue E . Indeed $\phi'(x) = -\psi'(-x)$ and so $\phi''(x) = \psi''(-x)$ ³⁵. Also $\phi(x)V(x) = \psi(-x)V(-x)$. So (for the 2nd equality, we take $x \rightarrow -x$ in the time independent SE above)

$$\hat{H}\phi(x) = -\frac{\hbar^2}{2m}\psi''(-x) + V(-x)\psi(-x) = E\psi(-x) = E\phi(x) \quad (191)$$

Thus $\hat{H}\phi = E\phi$ and $\phi(x) = \psi(-x)$ is also an eigenfunction of \hat{H} with the same energy. In particular, the even and odd combinations $\psi(x) + \psi(-x)$ and $\psi(x) - \psi(-x)$ are also eigenstates with the same energy E . In other words, we have shown that for an even potential, the energy eigenfunctions can be taken to be either even or odd. Actually, it may seem as if we have shown that there is always both an even and an odd eigenfunction for each energy. This is not the case. Most often, one of the above two combinations is the zero function, and therefore does not represent a linearly independent eigenfunction. In fact we will see later that there cannot be two linearly independent bound state eigenfunctions with the same energy in one dimension.

- Since $\psi_n(x)$ are even or odd about $x = L/2$, the probability density $|\psi_n(x)|^2$ is always even about $x = L/2$. We conclude that the expectation value of x is $L/2$ for all energy eigenstates. Thus, in the n^{th} eigenstate, the particle is on average at the midpoint. Note that this does not mean that $\langle x \rangle_\psi = 0$ for *any* state ψ , but rather, only for energy eigenstates. For example, we could prepare the particle in an initial state that is localized near $x = L/4$.
- Since ψ_n are stationary states, $\langle x \rangle_{\psi_n(t)}$ are time independent. So $\frac{\partial \langle x \rangle_{\psi_n(t)}}{\partial t} = 0$. It follows from Ehrenfest's theorem that $\langle p \rangle_{\psi_n} = m \frac{\partial \langle x \rangle}{\partial t} = 0$ at all times in a stationary state.
- The average energy in the n^{th} excited state is of course $\langle \psi_n | \hat{H} | \psi_n \rangle = E_n$.
- Now given any initial state $\Psi(x, 0) = \psi(x)$, we expand it in the energy basis $\psi(x) = \sum_n c_n \psi_n(x)$ where $c_n = \langle \psi_n | \psi \rangle$. Its time evolution is

$$\Psi(x, t) = \sum_{n=1}^{\infty} c_n \psi_n(x) e^{-iE_n t/\hbar} \quad \text{where} \quad E_n = \frac{n^2 \pi^2 \hbar^2}{2mL^2}. \quad (192)$$

In other words $|\Psi(t)\rangle = \sum_n c_n(t) |\psi_n\rangle$, where $c_n(t) = c_n e^{-iE_n t/\hbar}$.

- We may use this and the completeness of energy eigenstates to find the expectation value of energy:

$$\langle \Psi(t) | \hat{H} | \Psi(t) \rangle = \sum_{m,n} \langle \Psi(t) | \psi_n \rangle \langle \psi_n | \hat{H} | \psi_m \rangle \langle \psi_m | \Psi(t) \rangle = \sum_{n,m} c_n^*(t) E_n \delta_{nm} c_m(t) = \sum_n |c_n(t)|^2 E_n = \sum_n |c_n|^2 E_n$$

Since $|c_n(t)|^2 = |c_n|^2 e^{-iE_n t/\hbar} e^{iE_n t/\hbar} = |c_n|^2$ is time independent and E_n are time-independent, the average energy in the state $\Psi(t)$ is constant in time. This is the quantum version of the law of conservation of energy.

- We can estimate the ground state (g.s.) energy using the uncertainty principle. $\Delta p \geq \frac{\hbar}{2\Delta x}$. Since the particle is confined in the well, $\Delta x \leq L/2$. This implies $\Delta p \geq \frac{\hbar}{2} \frac{2}{L} = \frac{\hbar}{L}$. But since the average momentum is zero in the ground state, $(\Delta p)^2 = \langle p^2 \rangle$. So $\langle \hat{H} \rangle_1 = \langle \frac{p^2}{2m} \rangle_1 \geq \frac{\hbar^2}{2mL^2}$.

³⁵Primes denote derivatives w. r. to the specified argument of the function: $\phi'(x) = \partial_x \phi(x)$ and $\psi'(-x) = \partial_{-x} \psi(-x)$.

This is not too bad an estimate (it is only an inequality, a lower bound on the g.s. energy). The true g.s. energy $E_1 = \frac{\pi^2 \hbar^2}{2mL^2}$ is about ten times ($\pi^2 \approx 10$) this estimate.

- It is instructive to find the uncertainty product $\Delta x \Delta p$ for the stationary states. $\langle x \rangle_n = \frac{L}{2}$, $\langle p \rangle_n = 0$,

$$\langle x^2 \rangle_n = \int x^2 |\psi_n(x)|^2 dx = L^2 \left(\frac{1}{3} - \frac{1}{2n^2 \pi^2} \right), \quad \langle p^2 \rangle_n = 2m \langle \hat{H} \rangle_n = \frac{n^2 \pi^2 \hbar^2}{L^2}. \quad (193)$$

E.g., for $n = 1$, $\Delta x = \sqrt{\langle x^2 \rangle} \approx 0.53L$ is roughly half the width of the interval. The uncertainty product grows with n , but the uncertainty inequality is always satisfied

$$\Delta x \Delta p = \frac{\hbar}{2} \sqrt{\frac{n^2 \pi^2}{3} - 2} > \frac{\hbar}{2}. \quad (194)$$

The minimal uncertainty stationary state is the ground state for which $\Delta x \Delta p \approx 0.57 \hbar > \hbar/2$. The highly excited stationary states have a large value for the uncertainty product.

- We could try to find an eigenfunction $f(x)$ of momentum $\hat{p} = -i\hbar \frac{\partial}{\partial x}$ with eigenvalue $\hbar k$ for the particle in an infinite square well. But unfortunately, \hat{p} has no eigenfunctions that satisfy the boundary conditions $f(0) = f(L) = 0$. Indeed

$$-i\hbar f'(x) = \hbar k f(x) \Rightarrow f(x) = c e^{ikx} \quad (195)$$

But if $c e^{ikx}$ must vanish at $x = 0$ then $c = 0$. So \hat{p} has no (non-trivial) eigenfunctions. So what happens if we measure the momentum of the particle? The measurement postulate says that the wave function should collapse to a momentum eigenstate! The problem is that \hat{p} is not really an observable in the strict sense. Though it is hermitian on wave functions that vanish at the boundary, it is not self-adjoint. A physical interpretation is as follows. Since the particle is confined to a box, we already know a lot about the position of the particle, i.e. the experimenter has checked and told us that the particle is not outside the box. Given this partial knowledge of position, it is no longer possible to make arbitrarily accurate measurements of momentum. In other words, if we try to measure the momentum, we will get some value p_o with a certain accuracy. If we try to measure the momentum again at the next instant of time, we won't get the same value p_o . We interpret this by saying that after the first measurement, the wave function did not collapse to a momentum eigenstate. Indeed, there were no momentum eigenstates available to collapse to. Note that though \hat{p} has no eigenfunctions, $\hat{p}^2 = 2m\hat{H}$ has a complete set of eigenfunctions, the energy eigenstates ψ_n for the particle in a box.

7.2 Free particle and gaussian wave packet

- Consider a particle free to move on the infinite real line, with hamiltonian $H = p^2/2m$. The time-independent SE is³⁶ (here $k' = \sqrt{2mE/\hbar^2} \geq 0$ is the positive square root)

$$\psi''(x) = -\frac{2mE}{\hbar^2} \psi(x) = -k'^2 \psi(x) \quad (196)$$

³⁶Here we assume that the energy eigenvalue $E > 0$ since otherwise, the particle will have an ever growing amplitude of being found at larger and larger values of $|x|$, which is physically inappropriate to describe one or even a stream of particles.

There are two linearly independent solutions $e^{ik'x}$ and $e^{-ik'x}$ for any $k' > 0$. So we have two linearly independent energy eigenstates for each energy eigenvalue $E = \hbar^2 k'^2 / 2m > 0$. The ground state is the limiting case $E = 0$. The g.s. is non-degenerate, and corresponds to the constant eigenfunction $\psi(x) = 1$ ³⁷

- However, the position space probability distribution in any of these eigenstates (e.g. $|Ae^{ik'x}|^2$) is spread out over all of x -space. None of these eigenfunctions (nor any non-trivial linear combination) is square integrable. The energy eigenstates do not represent localized particles.

- Time dependence of any vector in the eigenspace corresponding to energy $E = \hbar^2 k'^2 / 2m = \hbar\omega$ is

$$(Ae^{ik'x} + Be^{-ik'x})e^{-iEt/\hbar} = Ae^{i(k'x - \omega t)} + Be^{-i(k'x + \omega t)} = e^{i(kx - \frac{\hbar k^2}{2m}t)}(A_{k>0} + B_{k<0}) \quad (197)$$

This is a linear combination of two traveling waves, a right-moving one and a left-moving one. In the last equality, we introduced a new wave vector variable $k = k'$ for right-moving waves and $k = -k'$ for left-moving waves.

- To summarize, the energy eigenstates $|k\rangle$ are the plane waves e^{ikx} for all $k \in \mathbb{R}$. The eigenstates labeled by k and $-k$ are degenerate in energy. The energy eigenstates are also eigenstates of momentum $-i\hbar\partial_x e^{ikx} = \hbar k e^{ikx}$, with eigenvalue $\hbar k$. States with momentum $k > 0$ move to the right and those with $k < 0$ move to the left.

- Energy eigenstates $|k\rangle$ are not localized in position. They do not represent wave packets. They do not have finite norm, though they are delta-normalized $\langle k'|k\rangle = 2\pi\delta(k - k')$, so energy eigenfunctions e^{ikx} do not have a probability interpretation. But we can draw an analogy with a fluid by computing the ‘probability’ current density $j(x, t) = \frac{\hbar}{2mi}(\psi^*\psi' - \psi'^*\psi)$ for the stationary state $\psi(x, t) = Ae^{i(kx - \omega(k)t)}$. We get $j(x, t) = |A|^2 \frac{\hbar k}{m} = |A|^2 v$ where $v = p/m$ is the corresponding classical velocity and $P(x, t) = |\psi(x, t)|^2 = |A|^2$ is the ‘probability’ density. This is akin to ρv for the mass density current in a fluid flow. So energy eigenstates can be interpreted as an always-present constant stream of free particles ($|A|^2 dx$ particles in the interval dx). For $k > 0$ enter from $x = -\infty$ and exit at $x = \infty$.

- Such non-normalizable wave functions with oscillatory $e^{\pm ikx}$ behavior as $|x| \rightarrow \infty$ are called *scattering states*. They correspond to particle trajectories that escape to infinity in classical mechanics. Though not normalizable, $\int_a^b |\psi(x)|^2$ is finite on any finite interval $[a, b]$. Though the total probability to be between $\pm\infty$ is infinite, we can still speak of relative probabilities. For example, in the state Ae^{ikx} with $k > 0$, the relative probability that a particle coming in from $-\infty$ scatters out to $+\infty$ is $\frac{|A|^2}{|A|^2} = 1$ while the relative probability for it to go back to $-\infty$ is $\frac{0}{|A|^2} = 0$.

- On the other hand, *bound states* are represented by normalizable wave functions that decay as $|x| \rightarrow \pm\infty$. Bound states correspond to classical particle trajectories that do not escape to infinity. All the eigenstates of the free particle hamiltonian are scattering states.

- Now we’d like to describe the evolution of an initial state that represents a particle, i.e., a localized wave packet $\psi(x)$ with finite norm. It cannot be an energy eigenstate, but may be expressed as a linear combination of energy eigenstates (same as momentum eigenstates)

³⁷The linear solution $\psi(x) = Ax + B, A \neq 0$ for $E = 0$ is disallowed for the same reasons.

which evolve via $e^{-i\omega t}$

$$\psi(x, t) = \int_{-\infty}^{\infty} [dk] \tilde{\psi}(k) e^{i(kx - \frac{\hbar k^2}{2m} t)}. \quad (198)$$

- A particularly useful wave packet is the gaussian one corresponding to the initial state

$$\psi(x) = A e^{-\frac{x^2}{4a^2}}, \quad A = \frac{1}{\sqrt{a}(2\pi)^{1/4}}, \quad \|\psi\| = 1, \quad |\psi(x)|^2 = \frac{1}{a\sqrt{2\pi}} e^{-x^2/2a^2} \quad (199)$$

$\langle x \rangle_\psi = 0$, so this packet is localized near $x = 0$. So $\psi(x) = A e^{-(x-x_0)^2/4a^2}$ is located near $\langle x \rangle = x_0$. The width of the packet is $\Delta x = \sqrt{\langle x^2 \rangle} = a$. This is a state of zero mean momentum $\langle p \rangle_\psi$ as the integrand $\psi^*(x)(-i\hbar\partial_x)\psi(x)$ is odd.

- To find the time evolution of this Gaussian wave packet, we write it in the energy basis

$$\tilde{\psi}(k) = \int dx \psi(x) e^{-ikx} = \int A e^{-\left(\frac{x^2}{4a^2} + ikx\right)} dx = 2aA\sqrt{\pi} e^{-a^2 k^2} = 2\sqrt{a} \left(\frac{\pi}{2}\right)^{1/4} e^{-a^2 k^2}. \quad (200)$$

The integral is done by *completing the square*, the change of variable $y = \frac{x}{2a} + ika$ and using $\int e^{-y^2} = \sqrt{\pi}$. The Fourier transform of the gaussian is again a gaussian.

- The width in momentum space is $\Delta p = \sqrt{\langle p^2 \rangle} = \frac{\hbar}{2a}$. We see that the Gaussian wave function minimizes the uncertainty product $\Delta x \Delta p = a \frac{\hbar}{2a} = \frac{\hbar}{2}$.

- Time evolution is simple in the energy basis

$$\psi(x, t) = \int [dk] \tilde{\psi}(k) e^{-iEt/\hbar} e^{ikx} = 2aA\sqrt{\pi} \int [dk] e^{-[k^2(a^2 + \frac{i\hbar t}{2m}) - ikx]} \quad (201)$$

This is again a Gaussian integral done by completing the square $l = k\sqrt{a^2 + \frac{i\hbar t}{2m}} - \frac{1}{2} \frac{ikx}{\sqrt{a^2 + \frac{i\hbar t}{2m}}}$.

We get

$$\psi(x, t) = \frac{1}{(2\pi)^{1/4} \sqrt{a + \frac{i\hbar t}{2ma}}} \exp \left\{ -\frac{x^2}{4 \left(a^2 + \frac{i\hbar t}{2m}\right)} \right\} \quad (202)$$

The probability density at time t is

$$|\psi(x, t)|^2 = \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{a^2 + \frac{\hbar^2 t^2}{4m^2 a^2}}} \exp \left\{ \frac{-x^2}{2 \left(a^2 + \frac{\hbar^2 t^2}{4m^2 a^2}\right)} \right\} = \frac{1}{a(t)\sqrt{2\pi}} e^{-x^2/2a(t)^2} \quad (203)$$

Here $a(t) \equiv \sqrt{a^2 + \frac{\hbar^2 t^2}{4m^2 a^2}}$. Width of the Gaussian $\sqrt{\langle x^2 \rangle} = a(t)$ grows with time. It remains centered at $\langle x \rangle = 0$. This is an indication of the dispersive behavior of de Broglie matter waves, the wave packet spreads out as its component plane waves travel at different phase speeds $c(k) = \hbar k^2/2m$.

- How fast does the wave packet disperse? We can write the width as

$$a(t) = a\sqrt{1 + \frac{\hbar^2 t^2}{4m^2 a^4}} = a\sqrt{1 + \frac{t^2}{\tau^2}}, \quad \tau = \frac{2ma^2}{\hbar}. \quad (204)$$

τ has dimensions of time and gives the rate of broadening. For $t \ll \tau$ there is not much broadening. For example, if we make a measurement of position with accuracy a , the wave

function collapses to a packet of width a . A subsequent measurement (after time t) of position will yield roughly the same position as long as long as $t \ll \tau$. If we wait too long $t \gg \tau$ to make the next measurement, the wave packet broadens significantly (by a factor of $\sqrt{1 + t^2/\tau^2}$), and we are no longer guaranteed to get roughly the same position.

- For example, suppose we know the position of the center of a tennis ball of mass 60g to within an accuracy of $a \sim 1\text{mm}$. If we model the tennis ball as a wave packet with a width equal to the above accuracy, then $\tau = 1.8 \times 10^{26}\text{s}$. So it takes a very long time for the quantum mechanical broadening of the tennis ball wave packet to become significant. In other words, we will get the same position even if we wait several centuries between successive measurements of the position of a tennis ball (that was initially at rest and was acted upon by no forces).

- The uncertainty product $\Delta x \Delta p$ remains equal to $\hbar/2$ at all times, since all that changes is the width $a(t)$, and $\Delta x \Delta p = \hbar/2$ was independent of the width a .

- The (expectation value of) energy of the gaussian wave packet at $t = 0$ is

$$\langle H \rangle_{t=0} = \frac{\langle p^2 \rangle}{2m} = \frac{(\Delta p)^2}{2m} = \frac{\hbar^2}{8ma^2}. \quad (205)$$

$\langle H \rangle$ is constant in time, as can be explicitly checked most easily in momentum space, where $\tilde{\psi}(k, t) = 2aA\sqrt{\pi}e^{-k^2(a^2 + \frac{\hbar t}{2m})}$

$$\langle H \rangle_t = \int [dk] |\tilde{\psi}(k, t)|^2 \frac{\hbar^2 k^2}{2m} = \frac{\hbar^2 a^2 A^2}{m} \int dk k^2 e^{-2a^2 k^2} = \frac{\hbar^2}{8ma^2}. \quad (206)$$

- This is a general feature, the expectation value of energy in any state is constant under Schrödinger evolution, provided the hamiltonian is hermitian and does not depend explicitly on time. To see this we note that $i\hbar\dot{\psi} = H\psi$ and $-i\hbar\dot{\psi}^* = (H\psi)^*$, so that

$$\begin{aligned} i\hbar\partial_t \int \psi^* H\psi &= i\hbar \int (\dot{\psi}^* H\psi + \psi^* H\dot{\psi}) dx = \int (-(H\psi)^* H\psi + \psi^* H H\psi) dx \\ &= -\langle H\psi | H\psi \rangle + \langle \psi | H H\psi \rangle = 0. \end{aligned} \quad (207)$$

- So far, our wave packet represented a particle that was on average at rest. To get a gaussian wave packet with non-zero mean momentum $\langle p \rangle = \hbar k_0$, we merely have to center the gaussian at k_0 in momentum space $\tilde{\psi}(k) = 2aA\sqrt{\pi}e^{-a^2(k-k_0)^2}$ so that $\langle \hbar\hat{k} \rangle = \hbar k_0$. This corresponds to the wave packet

$$\psi(x) = \int [dk] e^{ikx} 2Aa\sqrt{\pi}e^{-a^2(k-k_0)^2} = e^{ik_0x} \int [dl] e^{ilx} 2Aa\sqrt{\pi}e^{-a^2 l^2} = A e^{ik_0x} e^{-x^2/4a^2}. \quad (208)$$

Check directly that $\langle p \rangle = \hbar k_0$ by observing that $\psi^* \psi' = i\psi^* k_0 \psi_0 e^{ik_0x} + \psi^* e^{ik_0x} \psi_0'$, where $\psi = \psi_0 e^{ik_0x}$. The second term does not contribute to $\langle \hat{p} \rangle$ as it is odd and the first gives $\langle p \rangle = \hbar k_0$.

- The gaussian wave packet with non-zero mean momentum also has minimal uncertainty product $\Delta x \Delta p = \hbar/2$. $\Delta x = a$ is unaffected by the phase e^{ik_0x} . $\langle p \rangle = \hbar k_0$. $\langle p^2 \rangle = \hbar^2 k_0^2 + \frac{\hbar^2}{4a^2}$ is most easily evaluated in k -space. Thus $\langle p^2 \rangle - \langle p \rangle^2 = \frac{\hbar^2}{4a^2}$ is independent of k_0 . So $\Delta p = \hbar/2a$ and $\Delta x \Delta p = \hbar/2$.

7.3 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 $H = p^2/2m + \frac{1}{2}m\omega^2 x^2$. Large oscillations require corrections to Hooke's law. 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 $V(x) = V(x_0) + \frac{1}{2}V''(x_0)(x - x_0)^2$, as long as the second derivative of the potential is non-zero at x_0 .

- Classically, for fixed energy, motion lies between the turning points which are the solutions of $E = V(x) = \frac{1}{2}kx^2$. All classical trajectories are bound, particle cannot escape to $|x| \rightarrow \infty$ for any fixed energy, unlike the free particle.

- Quantum mechanically, the time-independent SE for energy eigenstates is

$$-\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} \psi(x) + \frac{1}{2}m\omega^2 x^2 \psi(x) = E\psi(x). \quad (209)$$

7.3.1 Variational Approach

- To begin with, we want to find the ground state, the eigenstate with lowest energy eigenvalue. Interestingly, there is an alternate characterization of the ground state among all normalizable states (not just among all eigenstates). It is the one which minimizes the expectation value $\langle H \rangle_\psi$.

- We first notice that the kinetic energy expectation value may be rewritten in a manifestly positive form after an integration by parts assuming $\psi(x) \rightarrow 0$ as $|x| \rightarrow \infty$

$$\langle T \rangle = -\frac{\hbar^2}{2m|\langle \psi | \psi \rangle|^2} \int \psi^* \psi'' dx = \frac{\hbar^2}{2m|\langle \psi | \psi \rangle|^2} \int \psi'(x)^* \psi'(x) dx = \frac{\hbar^2}{2m|\langle \psi | \psi \rangle|^2} \int |\psi'(x)|^2 dx \geq 0. \quad (210)$$

$V(x) = \frac{1}{2}m\omega^2 x^2$ is also non-negative, so the expectation value of energy in any state is ≥ 0

$$E(\psi) = \frac{1}{\langle \psi | \psi \rangle} \left[\int \left(\frac{\hbar^2}{2m} |\psi'(x)|^2 + \frac{1}{2}m\omega^2 x^2 |\psi(x)|^2 \right) dx \right] \geq 0. \quad (211)$$

The energy eigenvalues are just expectation values of the hamiltonian in the eigenstates. So it follows that the energy levels E_n (assumed discrete for convenience, we'll show later that the energy spectrum of the SHO is indeed discrete) must also be positive. So let us order them as $0 \leq E_0 < E_1 < E_2 < \dots$.

- Given any unit-norm state ψ we may write it as a linear combination of the orthonormal energy eigenstates $|\psi\rangle = \sum_n c_n |\psi_n\rangle$ and thereby express the expectation value of the hamiltonian as

$$\langle H \rangle_\psi = \sum |c_n|^2 E_n, \quad \sum_n |c_n|^2 = 1. \quad (212)$$

To minimize $\langle H \rangle$, we must select $|c_0|^2 = 1$ and $|c_{n>0}|^2 = 0$ ³⁸. In other words, the state with

³⁸Though it may seem reasonable, we have not *proved* that the choice $|c_n|^2 = \delta_{n,0}$ minimizes $\langle H \rangle$. This

the least energy eigenvalue (g.s.) is the one which minimizes $\langle H \rangle_\psi$!³⁹.

$$E_0 = \min_{\|\psi\|=1} \langle \psi | H | \psi \rangle \quad (213)$$

• **VARIATIONAL APPROXIMATIONS:** Suppose we want to find the g.s. approximately. It is usually too hard to minimize $E(\psi)$ over all square integrable $\psi(x)$. By minimizing it over a restricted class of wave functions, we will get an upper bound for the ground state energy and the best approximation to the g.s. wave function from that restricted class.

• Let us get a variational estimate for the gs energy and wave function of the harmonic oscillator by taking a gaussian variational ansatz (guess)

$$\psi_a(x) = Ae^{-x^2/4a^2}, \quad A = \frac{1}{\sqrt{a}(2\pi)^{1/4}}, \quad a > 0 \quad (214)$$

A gaussian is reasonable since we are looking for bound states and we expect that the probability of finding the particle far from the point of equilibrium to be very small. The width of the gaussian is our variational parameter. We want to find the width ‘ a ’ for which $\langle H \rangle$ is minimal. If a is large, the potential energy will be large and if a is small, the derivative of ψ will induce a large kinetic energy. So we expect an optimal intermediate value of a for which the energy is minimal.

• Notice that our variational ansatz for the g.s. has no nodes (zeros at finite values of x). This was a feature of the ground state of the particle in a square well. If a wave function has many nodes, it must oscillate and have a large kinetic energy and would not minimize the energy.

• $V(x)$ is real, so the eigenstates can be taken real, that’s why we picked a real $\psi_a(x)$

• $V(x) \propto x^2$ is an even function. So the stationary states must be either even or odd. An odd function would have a node, thus increasing the energy. So our ansatz was chosen as an even function of x .

• We evaluated $\langle p^2 \rangle$ and $\langle x^2 \rangle$ while discussing the gaussian wave packet for the free particle. So,

$$E(a) = \langle H \rangle = \frac{\langle p^2 \rangle}{2m} + \frac{1}{2}m\omega^2 \langle x^2 \rangle = \frac{\hbar^2}{8ma^2} + \frac{1}{2}m\omega^2 a^2 \quad (215)$$

The minimum of $E(a)$ occurs when

$$\frac{\partial E(a)}{\partial a} = m\omega^2 a - \frac{\hbar^2}{4ma^3} = 0 \Rightarrow a^2 = \frac{\hbar}{2m\omega} \quad (216)$$

is a problem in linear programming. Let $s_n = |c_n|^2$, then $\langle H \rangle = \sum_n s_n E_n$ with $\sum s_n = 1$. So $\langle H \rangle$ is a linear function of the variables s_n with non-negative coefficients, subject to the constraints $0 \leq s_n \leq 1$ and $\sum_n s_n = 1$. From linear programming, the minimum of $\langle H \rangle$ must occur at a boundary point of the region of allowed values of s_n . By repeated use of this fact one can show that $\langle H \rangle$ is minimal when $s_0 = 1$ and $s_{n \geq 1} = 0$. For example, if there are only two energy levels, $\langle H \rangle = s_0 E_0 + (1 - s_0) E_1$ is a linear function and is clearly minimal at $s_0 = 1$. More generally, $\frac{\partial \langle H \rangle}{\partial s_n} = E_n \geq 0$ with the slopes growing with n . So to minimize $\langle H \rangle$ we must choose the s_n as small as possible subject to the constraints and this is the above boundary point.

³⁹More generally, consider the energy functional $E(\psi) = \langle H \rangle_\psi$ for *any* state ψ (not just eigenstates). It can be shown that the time-independent SE ($H\psi = E\psi$) is the condition for $\langle H \rangle_\psi$ to be extremal. The extrema of the energy functional are the stationary states of the system! There are usually infinitely many such extrema, ordered by increasing values for the energy functional.

So among the gaussian wave functions, the one with least energy is

$$\psi_o(x) = \left(\frac{m\omega}{\pi\hbar}\right)^{1/4} e^{-\frac{m\omega}{2\hbar}x^2} = \frac{\sqrt{\beta}}{\pi^{1/4}} e^{-\frac{1}{2}\beta^2 x^2}, \quad \beta = \sqrt{\frac{m\omega}{\hbar}}, \quad A = \frac{\sqrt{\beta}}{\pi^{1/4}} \quad (217)$$

Putting this a in $\langle H \rangle$, our variational estimate (upper bound) for the ground state energy is $E_0 = \frac{1}{2}\hbar\omega$. It turns out that this estimate is exact, and the true ground state wave function is as above! So one cannot reduce the energy by choosing some other functional form for our variational ansatz. For e.g. a Lorentzian $\psi_a(x) = \frac{A}{x^2+a^2}$ leads to a larger estimate for the ground state energy.

- We introduced $\beta = \sqrt{\frac{m\omega}{\hbar}}$ which is a constant with dimensions of inverse length. β controls the rate of decay of the position space probability distribution. The larger β is, the less probable it is to find the particle far from the point of equilibrium $x = 0$. $\beta^2 = \frac{\sqrt{mk}}{\hbar}$ 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, so this observation has nothing to do with the size of \hbar .

7.3.2 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 the differential equation directly (which is not as simple as for the free particle), we follow a different approach. The ground state will be found by solving a 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. In the next section we will re-obtain the eigenstates by directly solving $H\psi = E\psi$.

- We change to dimensionless variables using $\beta = \sqrt{m\omega/\hbar}$ which has dimensions of L^{-1} . So let

$$\xi = \beta x, \quad \mathbf{p} = \frac{p}{\hbar\beta} = -i\frac{\partial}{\partial\xi}, \quad \Rightarrow \quad H = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 x^2 = \frac{\hbar\omega}{2} (\xi^2 + \mathbf{p}^2) \quad (218)$$

ξ, \mathbf{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

$$a = \frac{\xi + i\mathbf{p}}{\sqrt{2}}, \quad a^\dagger = \frac{\xi - i\mathbf{p}}{\sqrt{2}} \quad \Rightarrow \quad a^\dagger a = \frac{1}{2} (\xi^2 + \mathbf{p}^2 + i[\xi, \mathbf{p}]) = \frac{1}{2} (\xi^2 + \mathbf{p}^2 - 1) \quad (219)$$

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

$$H = \hbar\omega \left(a^\dagger a + \frac{1}{2} \right) = \hbar\omega \left(a a^\dagger - \frac{1}{2} \right), \quad a = \frac{1}{\sqrt{2}} \left(\xi + \frac{\partial}{\partial\xi} \right), \quad a^\dagger = \frac{1}{\sqrt{2}} \left(\xi - \frac{\partial}{\partial\xi} \right). \quad (220)$$

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

$$a^\dagger = \frac{1}{\sqrt{2}} \left(\beta x - \frac{ip}{\hbar\beta} \right) = \frac{1}{\sqrt{2}} \left(\beta x - \frac{1}{\beta} \frac{\partial}{\partial x} \right), \quad a = \frac{1}{\sqrt{2}} \left(\beta x + \frac{1}{\beta} \frac{\partial}{\partial x} \right) \quad (221)$$

- 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

$$H(a^\dagger\psi) = (E + \hbar\omega)(a^\dagger\psi), \quad H(a\psi) = (E - \hbar\omega)(a\psi). \quad (222)$$

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

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

The exact g.s. wave function matches our variational approximation! The g.s. energy is $\frac{1}{2}\hbar\omega$, same as our variational approximation. Why? $H\psi_0 = \hbar\omega(a^\dagger a + \frac{1}{2})\psi_0 = \frac{1}{2}\hbar\omega\psi_0$.

- So the particle in the g.s. has a non-zero probability to be found beyond the classical turning points.

- All the excited states are obtained as $\psi_n(x) = A_n(a^\dagger)^n\psi_0(x)$, for $n = 1, 2, 3, \dots$ 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 $E_n = \frac{1}{2}\hbar\omega + n\hbar\omega = \hbar\omega(n + \frac{1}{2})$. This is the energy spectrum of the SHO.

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

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

$$a^\dagger\psi_n = c_n\psi_{n+1}, \quad \text{and} \quad a\psi_n = d_n\psi_{n-1} \quad (223)$$

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

$$1 = \langle \psi_{n-1} | \psi_{n-1} \rangle = \langle \frac{a\psi_n}{d_n} | \frac{a\psi_n}{d_n} \rangle \Rightarrow |d_n|^2 = \langle \psi_n | a^\dagger a \psi_n \rangle = \langle \psi_n | \hat{N} \psi_n \rangle = n \|\psi_n\|^2 = n. \quad (224)$$

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

$$a^\dagger\psi_0 = c_0\psi_1 = \psi_1. \quad a^\dagger\psi_1 = \sqrt{2}\psi_2 \Rightarrow \psi_2 = \frac{1}{\sqrt{2!}}a^\dagger a^\dagger\psi_0, \dots \Rightarrow \psi_n = \frac{1}{\sqrt{n!}}(a^\dagger)^n\psi_0. \quad (225)$$

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

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

In particular, the first excited state wave function is odd, it has one node.

- 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. The particle has zero probability density to be found at the equilibrium point $x = 0$ in all the odd stationary states!

- Precisely,

$$\psi_n(\xi) = \sqrt{\frac{\beta}{\sqrt{\pi}}} \frac{1}{\sqrt{2^n n!}} H_n(\xi) e^{-\xi^2/2} \quad (226)$$

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

$$H_0 = 1, \quad H_1 = 2\xi, \quad H_2 = 4\xi^2 - 2, \quad H_3 = 8\xi^3 - 12\xi. \quad (227)$$

- The energy eigenstates are orthogonal. 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 $N\psi_n = n\psi_n$:

$$\langle \psi_n | \psi_m \rangle = \frac{1}{m} \langle \psi_n | a^\dagger a \psi_m \rangle = \frac{1}{m} \langle a \psi_n | a \psi_m \rangle \quad (228)$$

On the other hand,

$$\langle \psi_n | \psi_m \rangle = \frac{1}{n} \langle a^\dagger a \psi_n | \psi_m \rangle = \frac{1}{n} \langle a \psi_n | a \psi_m \rangle \quad (229)$$

To be equal, $m = n$ or if $m \neq n$, $\langle \psi_n | \psi_m \rangle = 0$. Combining with normalization, $\langle \psi_n | \psi_m \rangle = \delta_{mn}$.

- We have already seen that a gaussian wave packet minimizes the uncertainty product. What about the uncertainty product in the n^{th} energy eigenstate?

- 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 behavior in this sense, in the limit of large quantum numbers is called Bohr's correspondence principle.

7.3.3 Direct solution of Schrödinger equation for SHO

- See Schiff or Griffiths for details. So far we indirectly solved the Schrödinger equation for stationary states by reducing it to the 1st order equation $a\psi_0 = 0$. Then we built the excited states by repeatedly applying a^\dagger . Here we will discuss the direct solution of the 2nd order SE for energy eigenstates.

- In dimensionless variables $\xi = \beta x$ and $\frac{\partial}{\partial \xi} = \frac{1}{\beta} \frac{\partial}{\partial x}$,

$$-\frac{\hbar^2}{2m} \partial_x^2 \psi(x) + \frac{1}{2} m \omega^2 x^2 \psi = E \psi \Rightarrow -\psi''(\xi) + \xi^2 \psi(\xi) = \frac{2E}{\hbar \omega} \psi(\xi) \quad (230)$$

Here $\frac{2E}{\hbar \omega} = \frac{E}{E_0} = \mathcal{E}$ is the energy in units of the ground state energy. We seek normalizable solutions of this second order eigenvalue problem. It is not quite straightforward to solve this equation $\psi''(\xi) = (\xi^2 - \mathcal{E})\psi(\xi)$ with variable coefficients.

- The idea is to solve the SE for large ξ^2 (compared to \mathcal{E}), where it essentially reduces to the ground state (where \mathcal{E} is lowest), and then add corrections. For large $\xi^2 \gg \mathcal{E}$ the SE becomes $\psi'' = \xi^2 \psi$. It is natural to try the ground state wave function, $\psi = e^{-\xi^2/2}$ which turns out to be an approximate solution, since $\psi'' = (\xi^2 - 1)e^{-\xi^2/2} \approx \xi^2 e^{-\xi^2/2}$. The linearly independent approximate solution $e^{\xi^2/2}$ is disallowed since it is not normalizable.

- From the ladder operator story, we expect the exact solutions to possibly differ from this asymptotic behavior by polynomial corrections. So put $\psi(\xi) = h(\xi)e^{-\xi^2/2}$, for which $\psi' = (h' - h\xi)e^{-\xi^2/2}$. Then $h(\xi)$ must satisfy Hermite's differential equation

$$h'' - 2h'\xi + (\mathcal{E} - 1)h = 0. \quad (231)$$

Now we look for a polynomial/power series solution $h(\xi) = \sum_j a_j \xi^j$ for which

$$\xi h' = \sum_j j a_j \xi^j, \quad h'' = \sum_j (j+1)(j+2) a_{j+2} \xi^j. \quad (232)$$

So the differential equation for h reduces to the difference equation

$$(j+1)(j+2)a_{j+2} - 2j a_j + (\mathcal{E} - 1)a_j = 0 \quad (233)$$

leading to the recursion relation

$$a_{j+2} = \frac{2j+1-\mathcal{E}}{(j+1)(j+2)} a_j. \quad (234)$$

Given a_0 this determines the even coefficients and given a_1 , the odd coefficients, so $h(\xi) = h_o(\xi) + h_e(\xi)$. But not all such $h(\xi)$ lead to normalizable wave functions $\psi(\xi)$. Indeed, if $h(\xi)$ is not a polynomial but an infinite series, then for large j , $a_{j+2} \approx \frac{2}{j} a_j$. Recalling that for $f_n = \frac{1}{n!}$ we have $f_{n+1} = \frac{1}{n+1} f_n$, we see that for large j , $a_j \approx \frac{c}{(j/2)!}$. In other words,

$$h(\xi) = \sum_j a_j \xi^j \approx c \sum_j \frac{(\xi^2)^{j/2}}{(j/2)!} \approx c e^{\xi^2}. \quad (235)$$

So $\psi(\xi) = h(\xi)e^{-\xi^2/2}$ blows up at infinity. The only way to avoid this is for $h(\xi)$ not to be an infinite series, but just a polynomial (as we found using the ladder operators). In other words a_j must vanish for $j > n$ for some $n = 0, 1, 2, \dots$. From the recursion relation,

$$2n+1-\mathcal{E} = 0 \Rightarrow E = \hbar \omega \left(n + \frac{1}{2} \right), \quad n = 0, 1, 2, \dots \quad (236)$$

It is only for these values of \mathcal{E} that the Schrödinger equation can have normalizable solutions! Thus we re-derived the energy levels of the SHO. However, this condition on \mathcal{E} only terminates

the power series h_e if n is even and h_o if n is odd. So for normalizable ψ , h_e must be identically zero if n is odd and h_o must be zero if n is even. The coefficients follow from the recursion relation. In each case, $h_n(\xi)$ is a polynomial of degree n . Moreover, if by a rescaling, the coefficient a_n is chosen to equal 2^n , then the resulting polynomial is called the Hermite polynomial $H_n(\xi)$.

8 Some general properties of one dimensional quantum mechanical models

8.1 Parity

- The parity operator is defined on position space wave functions by reflection about $x = 0$

$$\mathbb{P}\psi(x) = \psi(-x) \quad (237)$$

In three dimensions, it is defined as $\mathbb{P}\psi(\mathbf{r}) = \psi(-\mathbf{r})$. Applying it twice, we see that $\mathbb{P}^2 = I$ is the identity. The eigenvalues $\mathbb{P}\psi(x) = \lambda\psi(x)$ must satisfy $\lambda^2 = 1$, i.e. $\lambda = \pm 1$. Even functions of x are eigenfunctions of parity with eigenvalue $+1$ and odd functions of x are eigenfunctions corresponding to eigenvalue -1 .

- A state is said to have definite parity if it is an eigenstate of parity, i.e. if its position space wave function is either even or odd. Most states do not have definite parity.
- The g.s. of the SHO $Ae^{-\beta^2 x^2/2}$ has even parity $\mathbb{P}\psi_0(x) = \psi_0(x)$. The excited states alternate in parity, odd and even. Time evolution doesn't change the parity of a stationary state $\psi_n(x)e^{-iE_n t/\hbar}$.
- \mathbb{P} is hermitian on square-integrable functions on any even interval $(-L, L)$ or on $(-\infty, \infty)$ as a change of variable $y = -x$ shows

$$\langle \phi | \mathbb{P} | \psi \rangle = \int_{-L}^L \phi^*(x) \psi(-x) dx = \int_L^{-L} \phi^*(-y) \psi(y) (-dy) = \int_{-L}^L \phi^*(-y) \psi(y) dy = \langle \mathbb{P}\phi | \psi \rangle. \quad (238)$$

As parity is hermitian & has a complete set of eigenfunctions (on $L^2(-L, L)$), parity is an observable.

- We can write the matrix representation of parity in any basis. Suppose we pick the basis of energy eigenstates of the SHO. In this basis P is diagonal. If we order the eigenstates according to $\dots \psi_5, \psi_3, \psi_1, \psi_0, \psi_2, \psi_4, \dots$, then in this basis $\mathbb{P} = \begin{pmatrix} -I & 0 \\ 0 & I \end{pmatrix}$.

- Parity does not commute with position or momentum. So momentum and parity (or position and parity) are not simultaneously measurable. Parity in fact anti-commutes with both position and momentum. Check that $[\mathbb{P}, x]\psi = -x\psi(-x) - x\psi(-x) = -2x\psi(x) \neq 0$ so that $[\mathbb{P}, x]_+\psi = (\mathbb{P}x + x\mathbb{P})\psi = 0$. As for momentum, let D be the operator that acts by differentiation on position space functions. $(Df)(x) = f'(x)$. $\mathbb{P}D + D\mathbb{P} = 0$ or $[\mathbb{P}, D]_+ = 0$. To see this:

$$((\mathbb{P}D)f)(x) = (\mathbb{P}f')(x) = \frac{\partial f(y)}{\partial y} \Big|_{y=-x}; \quad ((D\mathbb{P})f)(x) = \frac{\partial f(-x)}{\partial x} = -\frac{\partial f(y)}{\partial y} \Big|_{y=-x} \quad (239)$$

So $[\mathbb{P}, p]_+ = 0$ as well. We use $p\mathbb{P} + \mathbb{P}p = 0$ to show that \mathbb{P} commutes with p^2 or kinetic energy

$$[\mathbb{P}, p^2] = \mathbb{P}p^2 - p^2\mathbb{P} = -p\mathbb{P}p - p^2\mathbb{P} = p\mathbb{P}p - p^2\mathbb{P} = 0. \quad (240)$$

- Parity commutes with the potential energy operator for an even potential $V(x) = V(-x)$:

$$(\mathbb{P}V - V\mathbb{P})f(x) = V(-x)f(-x) - V(x)f(-x) = 0. \quad (241)$$

Thus parity commutes with the hamiltonian $H = p^2/2m + V$ for even $V(x)$. So we expect parity and energy to be simultaneously diagonalizable and measurable. In other words, eigenstates of energy can be taken to have definite parity: i.e., either even or odd. We saw that energy eigenstates of the SHO are either even or odd. The same was true for the square well (reflection about center of the well). More generally, if $\psi(x)$ is an energy eigenfunction but does not have definite parity, then we can add to it $\psi(-x)$ (which can be seen to also be an energy eigenfunction with the same energy) to get an even parity energy eigenfunction. Similarly, $\psi(x) - \psi(-x)$ has odd parity. Illustrate this with the free particle energy eigenfunctions e^{ikx} and e^{-ikx} . It may seem as if we have shown that there is always both an even and an odd eigenfunction for each energy. This is not the case. Often, one of the above two combinations is the zero function, and therefore does not represent a linearly independent eigenfunction. In fact one can show that bound state energy levels in one-dimensional QM are non-degenerate.

- A measurement of parity can only give the results $+1$ or -1 and the state of the system then collapses to a state whose position space wave function is either even or odd. The average of parity measurements on several copies of the same state is the expectation value of parity in that state. For example, the expected value of parity in the state $\psi = \psi_0 + \psi_1$ which is a linear combination of the ground and first excited states of the SHO is $\langle \psi | \mathbb{P} | \psi \rangle = 0$. This is because ψ_0 has even parity while ψ_1 has odd parity and they are orthogonal, but check it! The avg value of \mathbb{P} measurements in a state ψ tells us to what extent $\psi(x)$ is even or odd. If the avg parity is one, $\psi(x)$ must be even and if the average is -1 , $\psi(x)$ must be odd. But everything in between is also allowed. The expectation value of parity in any state must lie in the interval $[-1, 1]$ since it is a weighted avg of 1 's and -1 's.

- Parity (in three dimensions) is an important symmetry in microscopic physics. It is a symmetry of the electromagnetic and strong interactions and therefore has important consequences for atomic and nuclear/hadronic physics. However, parity is not a symmetry of the forces governing radioactive β decay.

8.2 Ehrenfest theorem and evolution of expectation values

- We have seen examples of Ehrenfest's thm. In any state, the expectation values evolve in time as

$$m \frac{\partial \langle x \rangle}{\partial t} = \langle p \rangle; \quad \frac{\partial \langle p \rangle}{\partial t} = -\langle \frac{\partial V}{\partial x} \rangle; \quad \frac{\partial \langle H \rangle}{\partial t} = 0 \quad (242)$$

What about the evolution of the expectation value of an operator \hat{A} in general? We will allow \hat{A} to possibly depend explicitly on time. From the SE and its adjoint we know (Since $H = H^\dagger$)

$$i\hbar|\dot{\psi}\rangle = H|\psi\rangle, \quad -i\hbar\langle\dot{\psi}| = \langle\psi|H. \quad (243)$$

Using this we find the time evolution of $\langle A \rangle_\psi = \langle \psi(t) | A | \psi(t) \rangle$ ⁴⁰

$$i\hbar \partial_t \langle A \rangle = i\hbar \left(\langle \dot{\psi} | A | \psi \rangle + \langle \psi | \dot{A} | \psi \rangle + \langle \psi | A | \dot{\psi} \rangle \right) = -\langle \psi | HA | \psi \rangle + \langle \psi | \dot{A} | \psi \rangle + \langle \psi | AH | \psi \rangle = \left\langle \frac{\partial \hat{A}}{\partial t} + [A, H] \right\rangle.$$

Here $[A, H] = AH - HA$ is the commutator. If A does not depend explicitly on time (which is most often the case, as for x, p, H, V, L), then

$$i\hbar \frac{\partial \langle A \rangle}{\partial t} = \langle [A, H] \rangle = \langle \psi(t) | [A, H] | \psi(t) \rangle \quad (244)$$

This is an illustration of the fact that the hamiltonian generates time evolution. The time derivative of the expectation value of any observable is given in terms of its commutator with the hamiltonian!

- The same evolution equation holds for any matrix element

$$i\hbar \partial_t \langle \psi | A | \phi \rangle = i\hbar \langle \psi | \partial_t \hat{A} | \phi \rangle + \langle \psi | [A, H] | \phi \rangle. \quad (245)$$

So suppose $\psi(t)$ and $\phi(t)$ evolve according to the SE and suppose \hat{A} does not depend explicitly on time. Then if $[A, H] = 0$ then the matrix elements $A_{\phi\psi}$ are constant in time.

- The above generalization of Ehrenfest's thm tells us when the expectation value of an observable can be constant in time. If A commutes with the hamiltonian, then $\langle A \rangle$ is conserved in *any* state⁴¹.

- For example, $[p, p^2/2m] = 0$. Momentum commutes with the free particle hamiltonian. So $\frac{\partial}{\partial t} \langle p \rangle = 0$, mean momentum is conserved for a free particle in any state.

- $[H, H] = 0$ so the mean energy is conserved (if H is not explicitly time dependent.) in any state.

- If $V(x)$ is even, then $[H, \mathbb{P}] = 0$, where \mathbb{P} is the parity operator. In this case $\langle \psi | \mathbb{P} | \psi \rangle$ is constant in time. Mean parity of any state is conserved in time. We saw this for eigenstates of the SHO, but it is true for any state, e.g. If $\psi = \psi_0 + \psi_1$ is a linear combination of the g.s. and first excited state of the SHO, then $\psi(x)$ is neither even nor odd. Yet $\langle \psi | \mathbb{P} | \psi \rangle$ is zero at all times!

8.3 Symmetries and conservation laws: Translation invariance

- We recall that in classical mechanics, a cyclic coordinate (one that was absent in the hamiltonian) implied that its conjugate momentum was conserved in time. A cyclic coordinate implies that the hamiltonian is independent of that coordinate. We can interpret this as a symmetry. E.g. if the hamiltonian is independent of the position x , it is unchanged under translations $x \mapsto x + a$. This is the case if $V(x) = V_0$ is a constant. Associated to such a translational symmetry there is a conserved quantity, p_x the momentum conjugate to x . This is generally true, symmetries typically lead to conserved quantities.

- Our discussion of parity symmetry for even potentials is an example of the general rule that in QM, observables that commute with the hamiltonian lead to conserved quantities.

⁴⁰This is the expectation value if $\|\psi\| = 1$, but what we derive is true for any diagonal matrix element, not necessarily expectation values.

⁴¹Even if $[A, H]$ is not the zero operator, there could be *some special* (time dependent) state ψ in which the expectation value of $[A, H]$ vanishes, so that $\partial_t \langle A \rangle$ is zero in *that* state. But this is an exceptional situation.

Many such observables correspond to symmetries of the system, such as reflection symmetry (parity), rotation symmetry (angular momentum), spatial translation symmetry (momentum), time translation symmetry (energy) etc. We briefly discuss this connection between symmetries and conservation laws. First let us see how transformations are implemented in the quantum theory.

- Finite evolution of a state by time t was implemented by the time evolution operator

$$|\psi(t)\rangle = e^{-i\hat{H}t/\hbar}|\psi(0)\rangle. \quad (246)$$

We say that H generates time evolution. Infinitesimal time evolution is got by taking $t = \delta t$ small:

$$|\psi(\delta t)\rangle = \left(I - \frac{iH\delta t}{\hbar} \right) |\psi(0)\rangle \quad (247)$$

This is just a restatement of the time dependent Schrodinger equation in finite difference form.

- Similarly recall that infinitesimal translations in x are generated by differentiation:

$$f(x + \delta x) = f(x) + (\delta x) f'(x) + \mathcal{O}(\delta x)^2 = \left(1 + (\delta x) \frac{\partial}{\partial x} \right) f(x) + \mathcal{O}(\delta x)^2. \quad (248)$$

- So it may not come as a surprise that a finite translation in x is generated by p via the exponent

$$e^{iap/\hbar} f(x) = f(x + a) \quad (249)$$

To show this, we write the exponential as a power series, put $p = -i\hbar\partial_x$ and use Taylor's formula

$$e^{iap/\hbar} f(x) = \sum_n \left(\frac{iap}{\hbar} \right)^n \frac{1}{n!} f(x) = \sum_n \frac{a^n}{n!} \frac{\partial^n}{\partial x^n} f(x) = f(x + a) \quad (250)$$

We say that momentum generates spatial-translations and that the hamiltonian generates time-translations.

- Now suppose the potential is independent of x , i.e. it is translationally invariant $V(x + a, y, z) = V(x, y, z)$. Then $H = \vec{p}^2/2m + V$ is independent of x . Classically, x is a cyclic coordinate and p_x is conserved in time. Quantum mechanically, $[H, p_x] = 0$ and it follows that $i\hbar\frac{\partial}{\partial t}\langle p_x \rangle = \langle [p_x, H] \rangle = 0$. So the expectation value of momentum is conserved if the hamiltonian has translation symmetry.

- Similarly, if the hamiltonian does not depend explicitly on time (as is usually the case) we say the system is time-translation invariant. Classically, energy is conserved. In QM,

$$i\hbar\frac{\partial\langle H \rangle}{\partial t} = \langle [H, H] \rangle = 0 \quad (251)$$

The expectation value of H (in any state) is conserved as a consequence of time translation invariance.

8.4 Energy-time uncertainty relation

- If we are in an eigenstate of the hamiltonian, the spread in energy measurements is zero $\Delta E = 0$. Recall that average values of all observables are constant in stationary states. So

when $\Delta E = 0$ it takes an infinite length of time for the mean value of any observable to change. This observation may be extended to an energy-time uncertainty principle.

- The energy-time uncertainty relation $\Delta E \Delta t \gtrsim \hbar/2$ is superficially like the position momentum uncertainty inequality $\Delta x \Delta p \geq \hbar/2$. However, Δt is *not* the standard deviation of measurements of time in the same state! Time is not an observable in the same way as position/momentum or energy. Now Δt may be interpreted as roughly the time it would take the expectation value of any observable A to change by one standard deviation. More specifically, let us define Δt via the equation

$$\Delta A = \left| \frac{\partial \langle A \rangle}{\partial t} \right|_{t=t_0} \Delta t \quad \text{at a fixed time } t_0. \quad (252)$$

For example, if ψ is an eigenstate of H , then $\Delta H = 0$. So we would expect $\Delta t = \infty$. Indeed, all observables have constant expectation values in stationary states, so it would take an infinite amount of time for the expectation value of any observable to change.

- More generally, if ΔE is small, then we would expect it to take a long time for the expectation value of any observable to change appreciably. On the other hand, if the energy of a state is very uncertain, then observables could change quite quickly starting from that state.

- To see this we begin with the result that in any state ψ and for any pair of observables

$$(\Delta A)^2 (\Delta B)^2 \geq -\frac{1}{4} \langle [A, B] \rangle^2 \quad (253)$$

where ΔA is the standard deviation of A in the state ψ . Let us take $B = H$, the hamiltonian and suppose A is not explicitly dependent on time. Then

$$i\hbar \frac{\partial \langle A \rangle}{\partial t} = \langle [A, H] \rangle_\psi \quad (254)$$

It follows that

$$(\Delta A \Delta H)^2 \geq \frac{\hbar^2}{4} \left(\frac{\partial \langle A \rangle}{\partial t} \right)^2 \Rightarrow \frac{\Delta A}{|\partial \langle A \rangle / \partial t|} \Delta H \geq \frac{\hbar}{2} \Rightarrow \Delta E \Delta t \geq \frac{\hbar}{2} \quad (255)$$

- Note that all the quantities that appear in this inequality refer to the same instant of time t_0 . We prepare the particle in a state ψ at a particular time. Then we pick an observable A and hence have the standard deviation of A as well as that of H in state ψ . Moreover, we also have the instantaneous rate of change of $\langle A \rangle_\psi$ at the same time. One must bear these definitions in mind when trying to interpret the above inequality.

8.5 Schrödinger vs Heisenberg pictures and equations of motion

- So far our formulation of QM has been based on the idea that the state of the system ψ evolves in time via the Schrödinger equation. However, the wave function itself is not measured. Rather, when we measure an observable A in a normalized state ψ , we get one of its eigenvalues a with a probability given by the square of the inner product (projection) $|\langle \phi_a | \psi \rangle|^2$. The system collapses to a (normalized) eigenstate ϕ_a where $A\phi_a = a\phi_a$. And if we make several copies of the system in the same state $\psi(t)$, and measure A in each of the copies, the average value obtained is $\langle \psi(t) | A | \psi(t) \rangle$

- We can express this expectation value at time t in terms of the expectation value (of a different operator) at a reference time (say $t = 0$) using the unitary time-evolution operator $U = e^{-iHt}$:

$$\langle \psi(t) | A | \psi(t) \rangle = \langle \psi(0) | e^{iHt/\hbar} A e^{-iHt/\hbar} | \psi(0) \rangle \quad (256)$$

The operator $A_h(t) = e^{iHt/\hbar} A e^{-iHt/\hbar}$ is called the operator A in the Heisenberg picture. The original operator A (sometimes called A_s) is said to be in the Schrodinger picture. Operators in the Heisenberg picture are related to those in the Schrodinger picture via a unitary transformation $A_h(t) = U^\dagger A U$.

- Thus, to calculate the expected value of an observable A at time t in the Heisenberg picture, we must evaluate $\langle \psi(0) | A_h(t) | \psi(0) \rangle$. Since we only need $\psi(0)$, we will say that the state of the system in the Heisenberg picture is $\psi_h = \psi(0)$. We can of course also write $\psi_h = U^\dagger \psi(t)$.
- In the Heisenberg formulation, states do not change in time, but the operators change with time.
- In the Schrödinger formulation, the state of the system evolves in time, while operators do not change with time (except if they are explicitly time-dependent).
- The hamiltonian operator is the same in both pictures $H_h = U^\dagger H U = H$ since the time evolution operator U commutes with H .
- Irrespective of whether we work in the Schrödinger or Heisenberg pictures, physically measurable quantities are the same.
- For e.g., the (measurable) eigenvalues of operators are the same in both pictures. This is because $A_s = A$ and $A_h = U^\dagger A U$ being related by a unitary transformation, share the same spectrum.
- We have already seen that expectation values are the same in both pictures $\langle \psi(t) | A | \psi(t) \rangle = \langle \psi_h | A_h | \psi_h \rangle$.
- In addition, inner products (projections, whose squares give probabilities of measurements) $\langle \phi(t) | \psi(t) \rangle = \langle \phi_h | \psi_h \rangle$ are also the same in both pictures. Here the system is in the Schrödinger state $\psi(t)$; we measure an observable, and get an eigenvalue corresponding to the eigenfunction $\phi(t)$.
- States do not evolve in time, so what replaces the Schrödinger equation in the Heisenberg picture? It is replaced by the Heisenberg equation of motion, which tells us how operators in the Heisenberg picture evolve. Suppose $A_h(t) = U^\dagger A U$ where A is a Schrödinger picture operator (that may have some explicit time dependence), then

$$i\hbar \frac{dA_h}{dt} = i\hbar \dot{U}^\dagger A U + i\hbar U^\dagger \frac{\partial A}{\partial t} U + i\hbar U^\dagger A \dot{U} \quad (257)$$

From $U = e^{-iHt/\hbar}$ and $U^\dagger = e^{-Ht/\hbar}$ we first observe that H, U, U^\dagger all commute with each other (after all, each is a function of H and $[H, H] = 0$). We also find

$$i\hbar \dot{U} = H U, \quad \text{and} \quad i\hbar \dot{U}^\dagger = -H U. \quad (258)$$

Thus the time evolution of A_h is given by the Heisenberg equation of motion

$$i\hbar \frac{dA_h}{dt} = i\hbar \frac{\partial A_h}{\partial t} + [A_h, H]. \quad (259)$$

In particular, if A does not have any explicit time dependence, then $i\hbar\dot{A}_h = [A_h, H]$. Moreover, if $[A, H] = 0$ (which is equivalent to $[A_h, H] = 0$), the Heisenberg operator $A_h(t)$ is a constant of motion. In other words, each of its matrix elements is time-independent.

- For a free particle, the Heisenberg picture momentum is a constant of motion $\dot{p}_h = 0$, since $[p, p^2/2m] = 0$.

8.6 Absence of degenerate bound states in one dimension

- For the SHO and square well we saw that for each bound state energy, there was only one linearly independent energy eigenfunction. In other words, the bound states are non-degenerate. On the other hand, for a free particle, the scattering energy levels $E_k = \frac{\hbar^2 k^2}{2m}$ for $k > 0$ are all doubly degenerate ($k, -k$). This is quite generic. Scattering states can be degenerate, but degenerate bound states tend not to occur in one dimension as we argue below.

- Suppose ψ_1, ψ_2 are normalizable⁴² (bound-state) energy eigenfunctions with the same energy E .

$$-\frac{\hbar^2}{2m}\psi''_{1,2} = (E - V)\psi_{1,2} \quad (260)$$

We wish to show that ψ_1 and ψ_2 are linearly dependent. i.e., $\psi_1(x) = \lambda\psi_2(x)$ so that they represent the same physical state rather than a pair of independent states degenerate in energy.

- Consider the Wronskian of these two solutions

$$W_{12}(x) = \det \begin{pmatrix} \psi_1 & \psi_2 \\ \psi'_1 & \psi'_2 \end{pmatrix} = \psi_1\psi'_2 - \psi'_1\psi_2 \quad (261)$$

The significance of the Wronskian is that it vanishes everywhere if ψ_1 and ψ_2 are linearly dependent. The converse is also true in many cases⁴³, if $W(x) \equiv 0$, then $\psi_1(x)$ and $\psi_2(x)$ are linearly dependent⁴⁴.

- We will show that $W_{12}(x) \equiv 0$. To do so we first show that it is constant and then evaluate it at a convenient value of x .

$$\frac{\hbar^2}{2m}W'_{12}(x) = \frac{\hbar^2}{2m}(\psi_1\psi''_2 - \psi''_1\psi_2) = \psi_1(V - E)\psi_2 + (E - V)\psi_1\psi_2 = 0 \quad (262)$$

⁴²In particular, $\psi(x), \psi'(x) \rightarrow 0$ as $|x| \rightarrow \infty$. If $\psi(x) \sim \frac{1}{x^\alpha}$, $\alpha > .5$, then $\psi' \sim \frac{1}{x^{\alpha+1}}$ decays even faster.

⁴³But there are exceptions (such as if the function fails to be sufficiently differentiable). The Wronskian could be zero everywhere and ψ_1 and ψ_2 may yet be linearly independent. We will assume here that such exceptional cases are excluded. They tend not to occur in the commonly studied quantum mechanics problems. So we assume that $W_{12}(x) = 0$ for all x implies that ψ_1 and ψ_2 are linearly dependent.

⁴⁴A partial indication of why this is true: If $\det \begin{pmatrix} \psi_1 & \psi_2 \\ \psi'_1 & \psi'_2 \end{pmatrix} = 0$ at every x , then the columns must be linearly dependent at every x . So $\psi_2(x) = \lambda(x)\psi_1(x)$ and $\psi'_2(x) = \lambda(x)\psi'_1(x)$. Dividing, $\psi'_2/\psi_2 = \psi'_1/\psi_1$ implies $\log \psi_2 = \log \psi_1 + c$ or $\psi_2(x) = A\psi_1(x)$, i.e. $\lambda(x) = A$ is constant and the two functions are linearly dependent. However, this argument does not always work, we could run into trouble where ψ_1 or ψ_2 vanish, since we divided by them. So this argument works in a region where both are non-vanishing, e.g. in a region between successive common zeros of ψ_1 and ψ_2 . We could have difficulty patching together two regions separated by a common zero of ψ_1 and ψ_2 . $\psi_{1,2}$ may be proportional to each other in each such region but with different proportionality constants. E.g., $\psi_1(x) = x^2 \operatorname{sgn}(x)$ and $\psi_2(x) = -x^2$ are linearly independent, yet their Wronskian vanishes everywhere. Of course, in this artificial example, ψ'_1 is discontinuous at $x = 0$.

Thus $W'(x) = 0$ and the Wronskian is constant. Moreover, $W_{12}(\infty) = 0$ since the bound state wave functions vanish at $x = \infty$. Thus $W_{12}(x) \equiv 0$. So (under suitable conditions), bound state energy eigenfunctions with the same energy must be linearly dependent: degenerate bound states do not occur in one dimension.

- On the other hand, degenerate bound states can and do occur in more than one dimension: the bound state energy levels of the hydrogen atom are degenerate.

8.7 Complete set of commuting observables

- In classical mechanics we could (and needed) to specify both the position and momentum of a particle to specify its state and determine its future evolution. In quantum mechanics a complete specification of the state is given by the wavefunction (upto a multiplicative constant). Instead of specifying the wavefunction, sometimes, it is convenient to identify the state as an eigenstate of appropriate observables, i.e., one where appropriate observables take definite values⁴⁵.

- For example, if we say that an SHO is in the state with energy $E_0 = \frac{1}{2}\hbar\omega$, then we know its wavefunction. On the other hand, if we say a free particle is in a state of energy $E = \hbar^2 k^2 / 2m > 0$, then its state is not completely determined, it could be in any linear combination of e^{ikx} and e^{-ikx} where $k = \sqrt{2mE/\hbar^2}$. To completely specify the state, we could (if possible) additionally specify the momentum $\hbar k$ or $-\hbar k$. This was possible as E and p are compatible observables, they are simultaneously measurable with arbitrary precision and simultaneously diagonalizable as operators, in particular they commute. There is no need to specify any other observable beyond E and p since the two of them serve to determine the state. We say that E, p form a complete set of commuting observables (CSCO). We say that the ‘quantum numbers’ of energy and momentum completely determine the state of a free particle (provided it is in an eigenstate of both energy and momentum). E, p provide labels with which to uniquely specify the state and determine its future evolution. Of course, we could also specify the value of the square of energy in the same state, but it would not provide any additional information about the state. So a CSCO is a maximal set of independent commuting observables (that includes the hamiltonian).

- For the free particle, a state of definite energy $E = \frac{\hbar^2 k^2}{2m}$ is not necessarily one of definite momentum. For e.g. the particle may be in the energy eigenstate $A \cos(kx)$ which is not an eigenstate of momentum. How do we pin down this state? Parity comes to the rescue. It commutes with the free particle energy and the above state is one of even parity. Energy and parity serve to uniquely specify this state. So H, \mathbb{P} form another CSCO for a free particle.

- However, we could not have specified the state of a free particle as one with definite values of E and x as they are not simultaneously specifiable.

- CSCO’s are thus not unique. Different states can require different sets of observables for a complete specification. For example, the eigenvalue of position by itself determines the state of a free particle, provided it is in a position eigenstate. The same can be said about momentum by itself. But we will primarily be interested in complete sets of commuting observables that include the hamiltonian, since we are interested in labeling the energy eigenstates. Quite often, energy by itself serves to uniquely specify an energy level. It is only when energy levels

⁴⁵Of course, this is not always possible, not every state is an eigenstate of some well-known observable.

are degenerate that we need to augment the hamiltonian with other observables to uniquely specify the state.

- In one dimension, bound states tend not to be degenerate in energy, though scattering states can be degenerate. In $d > 1$, degenerate bound states are more common, and tend to be associated with symmetries like rotation invariance or parity.

9 Models with bound and scattering states, tunneling

9.1 Bound state of the delta function potential

- One of the simplest models that displays both bound and scattering states is the attractive delta function potential. Consider a particle in 1d moving in the vicinity of a single point-like attractive force-center ('core') located at $x = 0$. The particle feels no force except when it is very close to $x = 0$ and in this case, there is a strong attractive force towards $x = 0$. A simple model is the idealized potential $V(x) = -g\delta(x)$ for some constant $g > 0$ with dimensions of energy \times length. g is called the coupling strength of the attractive delta function.

- If a particle is fired at such an attractive force center with positive energy, we expect it to get scattered and escape to infinity. On the other hand, if $E < 0$, there could be situations where a particle is bound to the force center and cannot escape. We expect the Schrödinger eigenvalue problem $-\frac{\hbar^2}{2m}\psi'' - g\delta(x)\psi(x) = E\psi(x)$ to support scattering states for $E > 0$ and bound states for $E < 0$. The larger g is, the more strongly would we expect the particle to be bound to the core (in a bound state).

- Consider first the bound (normalizable) states $E = -\frac{\hbar^2\kappa^2}{2m} < 0$. Away from $x = 0$, we just have the free particle SE. $\psi'' = \kappa^2\psi$ whose solution is a linear combination of $e^{\pm\kappa x}$. For the norm to be finite we must have $\psi(x > 0) = Ae^{-\kappa x}$ and $\psi(x < 0) = Be^{\kappa x}$. Now, imposing continuity of the wave function at $x = 0$ we get $A = B$. So $\psi(x) = Ae^{-\kappa|x|}$. The normalization constant is fixed to be $A = \sqrt{\kappa}$ by normalization $\int |A|^2 e^{-2\kappa|x|} dx = 1$. But we haven't yet found the energy eigenvalue E or κ .

- So far, the $-g\delta(x)$ potential has not played any role! What is more, we have only imposed one boundary condition (continuity of ψ at $x = 0$). We need to impose a second boundary condition, evidently at $x = 0$ to account for the attractive force. But what is the boundary condition? It should be something to do with ψ' . To find it, we integrate the SE in a small neighborhood $(-\epsilon, \epsilon)$ of the attractive potential, in anticipation of taking $\epsilon \rightarrow 0$. This is because the integral of ψ'' should give us something involving ψ' .

$$-\frac{\hbar^2}{2m} \int_{-\epsilon}^{\epsilon} \psi''(x) dx - g \int_{-\epsilon}^{\epsilon} \delta(x) \psi(x) dx = E \int_{-\epsilon}^{\epsilon} \psi(x) dx \Rightarrow -\frac{\hbar^2}{2m} \{\psi'(\epsilon) - \psi'(-\epsilon)\} - g\psi(0) = E \int_{-\epsilon}^{\epsilon} \psi(x) dx$$

Since $\psi(x)$ is bounded near $x = 0$, the rhs tends to zero as $\epsilon \rightarrow 0$. Thus we get

$$\psi'(0^+) - \psi'(0^-) = -\frac{2mg}{\hbar^2}\psi(0) \quad (263)$$

We have found that ψ' is not continuous at $x = 0$. It is discontinuous by an amount \propto strength of the attractive delta function. This condition determines $\kappa = \frac{mg}{\hbar^2}$. It follows that there is only one bound state(!) with the wave function and energy

$$\psi(x) = \frac{\sqrt{mg}}{\hbar} \exp\left\{-\frac{mg|x|}{\hbar^2}\right\}, \quad E = -\frac{mg^2}{2\hbar^2}. \quad (264)$$

- As the strength of the attractive force increases (larger g), the wave function is more strongly localized near $x = 0$ and the *binding energy* $-E$ also grows. This is as we would expect. If the sun were heavier, the planets would be more tightly bound to it and have a more negative energy in their bound Keplerian elliptical orbits.
- On the other hand, as $g \rightarrow 0$, the the state is less strongly bound to the center of attraction. Its energy approaches zero. The wave function becomes more and more delocalized and approaches a constant $\psi(x) = Ae^{-\kappa|x|} \rightarrow A$. In other words, we recover the ground state of the free particle.
- It is interesting to see what changes if we had a repulsive potential $V(x) = -g\delta(x)$ where $g < 0$. Then classically, we do not expect any bound orbits. Quantum mechanically, the first change occurs when we impose the discontinuity of ψ' across $x = 0$, we get $\kappa = mg/\hbar^2$. But since $g < 0$, $\kappa < 0$ and $e^{-\kappa|x|}$ is not normalizable. It follows that there are no normalizable bound state energy levels for a repulsive short range $\delta(x)$ -potential.
- Next we look for the scattering states with energy eigenvalues $E = \frac{\hbar^2 k^2}{2m} \geq 0$. Before doing so, we discuss the basic set up of scattering problems.

9.2 Reflection and Transmission coefficients

- Consider the scattering of (say) electrons off a molecule. The scattering centre (molecule) is localized. Far away from the molecule, the electrons feel almost no force. In one dimension, this means the potential due to the molecule $V(x) \rightarrow 0$ as $x \rightarrow \pm\infty$. In such scattering problems, we are primarily interested in understanding the structure of the molecule. The electrons are called test particles, they are used to determine the nature of the molecular forces. We do so by firing in electrons of definite energy ‘from infinity’ in one direction towards the molecule (say from $x = -\infty$). Then, based on how the electrons are scattered, in various directions, we try to infer what $V(x)$ is. Or, if we know $V(x)$, then we can predict how the electrons will scatter.
- We seek the energy eigenstates of the 1-dimensional SE $-(\hbar^2/2m)\psi'' + V(x)\psi(x) = E\psi(x)$ with $V(x) \rightarrow 0$ as $x \rightarrow \pm\infty$. For large $|x|$ the SE reduces to that for a free particle, $\psi''(x) = -\frac{2mE}{\hbar^2}\psi(x)$. If $E < 0$, the wave function would have to decay exponentially at $\pm\infty$ and would not represent a scattering state. So suppose $E \geq 0$, and let $k = \frac{\sqrt{2mE}}{\hbar} \geq 0$. Then asymptotically,

$$\psi(x) \rightarrow Ae^{ikx} + Be^{-ikx} \text{ as } x \rightarrow -\infty, \quad \text{and} \quad \psi(x) \rightarrow Ce^{ikx} + De^{-ikx} \text{ as } x \rightarrow \infty. \quad (265)$$

To physically interpret A, \dots, D we compute the corresponding probability current densities $j = \frac{\hbar}{2mi} (\psi^* \psi' - \psi'^* \psi)$

$$\begin{aligned} j_A &= \frac{\hbar k}{m} |A|^2, \text{ rightward probability current at } x = -\infty \\ j_B &= -\frac{\hbar k}{m} |B|^2, \text{ leftward probability current at } x = -\infty \\ j_C &= \frac{\hbar k}{m} |C|^2, \text{ rightward probability current at } x = \infty \\ j_D &= -\frac{\hbar k}{m} |D|^2, \text{ leftward probability current at } x = \infty \end{aligned} \quad (266)$$

We say that A is the amplitude of the incoming wave from the left, D the amplitude of the incoming wave from the right, B the amplitude of the outgoing wave on the left and C the

amplitude of the outgoing wave on the right. These interpretations are seen to be reasonable if we introduce the harmonic time dependence of energy eigenstates $e^{-i\omega t}$ where $E = \hbar\omega$. Then we have $Ae^{i(kx-\omega t)}$ which represents a right moving wave in the vicinity of $x = -\infty$. Similarly, $Be^{-i(kx+\omega t)}$ represents a left-moving wave at $x = -\infty$, etc.

- For definiteness, suppose particles are sent in from $x = -\infty$ at momentum $\hbar k$, and no particles are sent in from $x = +\infty$. This set up requires $D = 0$. B is called the amplitude of the REFLECTED wave and C is called the amplitude of the TRANSMITTED wave.
- Of particular interest: relative probability current of reflection per unit incoming probability current

$$R = \frac{|j_B|}{|j_A|} = \frac{|B|^2}{|A|^2} = \text{'Reflection coefficient'} \quad (267)$$

Similarly, the transmission coefficient is defined as the relative probability of transmission

$$T = \frac{|j_C|}{|j_A|} = \frac{|C|^2}{|A|^2} = \text{'Transmission coefficient'} \quad (268)$$

One of the goals of scattering problems is to find these reflection and transmission coefficients (for a given potential) since they can be measured experimentally (e.g. experimentally, R is the ratio of number of particles back scattered to the number of particles sent in per unit time). Of course, we are also interested in finding the allowed scattering energy eigenvalues E , just as we wanted to find the bound state energy levels. However, unlike bound states, typically there is no restriction on the allowed energies of scattering scattering states (they form a continuous spectrum for $E > 0$). Nevertheless, R and T will depend on the energy E . E.g. If the energy of the incoming electrons is very high, most are likely to pass through the molecule and get transmitted, only a few may be reflected.

9.3 Scattering states for a delta function potential

- For the attractive $V(x) = -g\delta(x)$ potential, let us study the scattering problem for an incoming probability current from the left. The solutions of the SE in the regions $x < 0$ and $x > 0$ are

$$\psi(x < 0) = Ae^{ikx} + Be^{-ikx}, \quad \psi(x > 0) = Ce^{ikx}. \quad (269)$$

We must now impose continuity of $\psi(x)$ at $x = 0$ and discontinuity of ψ' across $x = 0$ by the amount previously determined. This will allow us to determine B, C in terms of A , leaving $E \geq 0$ undetermined. In other words, there are scattering eigenstates for all non-negative energies.

- In more detail, continuity at $x = 0$ implies $A + B = C$. In addition,

$$\psi'(0^+) - \psi'(0^-) = -\frac{2mg}{\hbar^2}\psi(0) \Rightarrow ik(C - A + B) = -\frac{2mg}{\hbar^2}(A + B) \quad (270)$$

We can solve for B, C in terms of A and the wave number k . In fact it is convenient to introduce a dimensionless measure of the strength of the attractive potential $\gamma = \frac{mg}{\hbar^2 k}$. γ is small if the coupling g is weak or the energy E (or k) is large. We find

$$B = \frac{i\gamma A}{1 - i\gamma}, \quad C = \frac{A}{1 - i\gamma}. \quad (271)$$

In terms of these, the reflection and transmission probabilities are

$$R = \frac{|B|^2}{|A|^2} = \frac{\gamma^2}{1 + \gamma^2}, \quad T = \frac{|C|^2}{|A|^2} = \frac{1}{1 + \gamma^2}, \quad \gamma = \frac{mg}{\hbar^2 k} \quad (272)$$

The particles must either be reflected or transmitted, there is nowhere else for them to go. Indeed, we find that $T + R = 1$. With probability one, the particles are either reflected or transmitted.

- Furthermore, $T \rightarrow 0$ if $k \rightarrow 0$ or $g \rightarrow \infty$. Particles of very low energies are not transmitted. Such particles are reflected $R(\gamma \rightarrow \infty) = 1$. On the other hand, high energy projectiles tend to be transmitted $T(\gamma \rightarrow 0) = 1$. Plot T and R as a function of energy (or γ). What is the critical energy at which there is an equal probability for the particle to be reflected as to be transmitted?

- If we had a repulsive delta-function potential $-g\delta(x)$, $g < 0$, then the sign of γ changes. But the reflection and transmission coefficients are unchanged, being functions only of γ^2 . In particular, despite the repulsive force, the particle injected from $x = -\infty$ has a finite probability to penetrate the potential barrier and emerge on the right. This is the classically forbidden phenomenon of quantum mechanical tunneling. A repulsive delta function potential is like a short-range repulsive barrier. Classically, a particle of any finite energy would have been reflected.

- Tunneling is the phenomenon by which a particle can have a non-zero probability of being found beyond a barrier despite having an energy less than the height of the barrier. If the barrier is very high or very wide (when appropriately compared to the energy of the particle), the transmission probability across the barrier is small.

- Tunneling is a wave phenomenon and does not have a simple explanation in terms of particle trajectories. It has been experimentally observed and used to explain several phenomena. E.g. decay of nuclei by emission of α particles. The alpha particle does not have enough energy to classically escape from the nuclear potential. But it occasionally does so, on account of the non-zero tunneling probability. When two conductors are separated by a thin layer of insulator and a current is injected, there can be a current across the insulator due to tunneling. Scanning tunneling microscopes employ tunneling to permit atomic resolution imaging.

9.4 Finite square well

- The finite potential well corresponds to the potential $V(x) = 0$ for $|x| > a$ and $V(x) = -V_0$ for $|x| \leq L$. For $V_0 > 0$, this is a more realistic version of the infinite square well, and a step closer to a model for the attractive force holding together the constituents of a nucleus or atom. As $V_0 \rightarrow \infty$ we recover the infinite square well. And if $a \rightarrow 0$ as well, $V(x)$ approximates the attractive delta potential. The infinite square well had infinitely many bound states while the delta well had just one. Since the finite square well is in a sense between the two, we might expect it to have finitely many bound states. Since $V(x)$ is even, energy eigenstates can be taken either even or odd. For $V_0 < 0$ the potential represents a repulsive barrier.

- In each of the three regions $x < -a$, $|x| < a$, $x > a$, the potential is constant so each is treated separately. Then we impose continuity of the wave function $\psi(x)$ and its derivative

$\psi'(x)$ at the points $x = \pm a$.

- We begin by looking for stationary states with $E < 0$, which turn out to be bound states. First let us show that the energy eigenvalue $E \geq -V_0$. This is a general feature, the energy eigenvalues must be at least as large as the lowest value of the potential. Integrating by parts, we write the expectation value of the hamiltonian in any normalized state (which in particular vanishes at infinity) as

$$\langle H \rangle = \frac{\hbar^2}{2m} \int |\psi'(x)|^2 dx + \int V(x)\psi(x) dx = \langle T \rangle + \langle V \rangle. \quad (273)$$

It is clear that $\langle T \rangle \geq 0$. On the other hand $\langle V \rangle = -V_0 \int_{-a}^a |\psi(x)|^2 dx$. Now $0 \leq \int_{-a}^a |\psi(x)|^2 dx \leq 1$. So $-V_0 \leq \langle V \rangle \leq 0$. Combining with the positivity of $\langle T \rangle$, we conclude that $\langle H \rangle \geq -V_0$. In particular, energy eigenvalues must be at least as big as $-V_0$.

- For $-V_0 \leq E < 0$, the equation for stationary states outside the well is

$$|x| > a, \quad \psi''(x) = \kappa^2 \psi(x), \quad \text{where } \kappa = \sqrt{-\frac{2mE}{\hbar^2}} > 0 \quad (274)$$

So normalizable solutions are $\psi(x < a) = Ae^{\kappa x}$ and $\psi(x > a) = Be^{-\kappa x}$. In the region of the well,

$$|x| < a \quad \Rightarrow \quad \psi'' = -l^2 \psi(x) \quad \text{where } l^2 = \frac{2m(E + V_0)}{\hbar^2} \quad (275)$$

We showed that $E \geq -V_0$ so that we may take $l \geq 0$. $\psi(x)$ is oscillatory for $|x| < a$ and is conveniently taken as a linear combination of solutions of definite parity $\psi(|x| < a) = C \cos(lx) + D \sin(lx)$.

- Now we must impose the boundary conditions at $x = \pm a$. Without loss of generality we may look for even and odd eigenfunctions. For such functions, imposing the b.c. at $x = a$ automatically imposes them at $x = -a$. Let us consider the even states (so $A = B$ and $D = 0$). Continuity of ψ, ψ' at $x = a$ implies

$$Ae^{-\kappa a} = C \cos(la) \quad \text{and} \quad -A\kappa e^{-\kappa a} = -Cl \sin(la) \quad (276)$$

Dividing, $\kappa = l \tan la$. Both κ and l are functions of E , so this is a non-linear transcendental equation for bound state energies.

- For the odd states $A = -B, C = 0$ and we get the transcendental equation $\kappa = l \cot la$ from the b.c.

$$Ae^{-\kappa a} = D \sin la \quad \text{and} \quad A\kappa e^{-\kappa a} = Dl \cos(la) \quad (277)$$

- l, κ define a point on a circle $l^2 + \kappa^2 = \frac{2mV_0}{\hbar^2}$. Both κ and l have dimensions of inverse length while a is a length scale defining the width of the well. So it makes sense to measure l, κ and the radius of the circle in units of a^{-1} . So we pass to the dimensionless variables $\lambda = al, \kappa a = \sqrt{r^2 - \lambda^2}$ and $r = a \frac{1}{\hbar} \sqrt{2mV_0}$. r measures the strength of the potential, it is large for a wide or deep well.

- Then the equation for even parity energy levels $\kappa a = la \tan la$ becomes $\lambda \tan \lambda = \sqrt{r^2 - \lambda^2}$ or

$$\tan \lambda = \sqrt{\frac{r^2}{\lambda^2} - 1} \quad (278)$$

- We solve this equation graphically by plotting lhs and rhs as functions of λ and looking for points of intersection in the physical region $\lambda \geq 0, r \geq 0$. Draw a picture! There are only finitely many solutions $\lambda_1, \lambda_2, \dots, \lambda_n$, corresponding to finitely many even bound states. Similarly there are finitely many bound states with odd wave functions whose energies are determined by $-\cot \lambda = \sqrt{r^2/\lambda^2 - 1}$. The energy levels alternate in parity.

- The bound state wave functions can be found. E.g., for each even bound state, κ, l, E are fixed, so continuity at a can be used to fix C in terms of A & E : $C = Ae^{-\kappa a} \sec(la)$. Then normalization $\int_{-a}^a |C|^2 \cos^2(lx) dx + 2 \int_a^\infty |A|^2 e^{-2\kappa x} dx = 1$ determines A . A plot of the even bound state wave function shows that it is oscillatory in the region of the well and decays exponentially outside the well.

- An interesting limiting case is $r \rightarrow \infty$ corresponding to a strong/long range potential (e.g. a deep well approaching an infinite square well by keeping a fixed and let $V_0 \rightarrow \infty$). In this case, the points of intersection (for even parity states) lie at

$$\lambda \approx \frac{\pi}{2}, \frac{3\pi}{2}, \frac{5\pi}{2}, \dots \Rightarrow a^2 l^2 \approx \frac{(2n+1)^2 \pi^2}{4} \Rightarrow E_{2n+1} = -V_0 + \frac{(2n+1)^2 \pi^2 \hbar^2}{2m(2a)^2}, \quad n = 0, 1, 2, \dots \quad (279)$$

These correspond to the even parity energy levels (measured with respect to the bottom $-V_0$) of a square-well potential of width $L = 2a$. Similarly, the odd parity energy levels when $r \rightarrow \infty$ are at

$$\lambda \approx \pi, 2\pi, 3\pi, \dots \quad \text{corresponding to} \quad E_{2n} = -V_0 + \frac{(2n)^2 \pi^2 \hbar^2}{2m(2a)^2}, \quad n = 1, 2, 3, \dots \quad (280)$$

Thus we recover the full spectrum of the infinite square well in the limit $r \rightarrow \infty$, provided we measure energies with respect to the bottom of the well.

- Another interesting limiting case is a short-range or weak potential $r = \frac{a\sqrt{2mV_0}}{\hbar} \rightarrow 0$. E.g. we can get the attractive delta potential well by letting $a \rightarrow 0, V_0 \rightarrow \infty$ holding $aV_0 = g$ fixed. Then $r \propto g/\sqrt{V_0} \rightarrow 0$. As r decreases, there are fewer even parity bound states and eventually for $r < \pi/2$, there is only one even parity bound state: estimate its energy by approximately solving the transcendental equation for small λ . For sufficiently small r , there are no odd parity bound states. Interestingly, no matter how weak or narrow the potential may be, there is always at least one (even parity) bound state, as we found for the delta function potential well.

9.5 Scattering against a rectangular barrier/well and tunneling

- We consider scattering from the left against a rectangular barrier of width $2a$ modeled by the repulsive potential $V(x) = V_0 > 0$ for $|x| < a$ and zero otherwise. We seek scattering energy eigenstates $E > 0$.

- For $|x| > a$, the equation for stationary states is $\psi'' = -k^2\psi$ where $\hbar k = \sqrt{2mE} \geq 0$. Thus we have a wave Ae^{ikx} incident from the left, a reflected wave Be^{-ikx} and a transmitted wave Fe^{ikx} .

- For $|x| < a$, $\psi'' = \kappa^2\psi$ where $\hbar^2\kappa^2 = 2m(V - E) \geq 0$. So $\psi(|x| < a) = Ce^{\kappa x} + De^{-\kappa x}$.

- Now we must impose continuity of ψ, ψ' at $x = \pm a$.

$$x = -a \quad Ae^{-ika} + Be^{ika} = Ce^{-\kappa a} + De^{\kappa a}, \quad ikAe^{-ika} - ikBe^{ika} = C\kappa e^{-\kappa a} - D\kappa e^{\kappa a}.$$

$$x = a \quad Ce^{\kappa a} + De^{-\kappa a} = Fe^{ika}, \quad C\kappa e^{\kappa a} - D\kappa e^{-\kappa a} = ikFe^{ika} \quad (281)$$

A, B, C, D, E, F are 6 unknowns. We can choose $A = 1$ to normalize the incoming probability current from the left. The above 4 equations will determine B, C, D, F in terms of the energy $E \geq 0$ which remains arbitrary.

• In more detail, we use the b.c. at $x = a$ to express C, D (which do not appear in the reflection and transmission coefficients) in terms of F

$$C = \frac{F(\kappa + ik)}{2\kappa} e^{(ik - \kappa)a}, \quad D = \frac{F(\kappa - ik)}{2\kappa} e^{(ik + \kappa)a}. \quad (282)$$

Putting these in the b.c. at $x = -a$ we get

$$\begin{aligned} Ae^{-ika} + Be^{ika} &= \frac{F}{\kappa} e^{ika} [\kappa \cosh(2\kappa a) - ik \sinh(2\kappa a)] \\ Ae^{-ika} - Be^{ika} &= \frac{F}{ik} e^{ika} [ik \cosh(2\kappa a) - \kappa \sinh(2\kappa a)]. \end{aligned} \quad (283)$$

Adding we get

$$\frac{A}{F} = \frac{1}{2} e^{2ika} \left[2 \cosh(2\kappa a) + i \sinh(2\kappa a) \frac{(\kappa^2 - k^2)}{\kappa k} \right] \quad (284)$$

The transmission coefficient is the absolute square

$$T^{-1} = |A/F|^2 = \cosh^2(2\kappa a) + \frac{1}{4} \sinh^2(2\kappa a) \frac{(\kappa^2 - k^2)^2}{\kappa^2 k^2} \quad (285)$$

Now $\cosh^2 = 1 + \sinh^2$ and $1 + \frac{(\kappa^2 - k^2)^2}{4\kappa^2 k^2} = \frac{(\kappa^2 + k^2)^2}{4\kappa^2 k^2}$. Moreover, $\kappa^2 = \frac{2m}{\hbar^2}(V_0 - E)$ and $k^2 = \frac{2mE}{\hbar^2}$. So

$$\frac{1}{T} = 1 + \sinh^2(2\kappa a) \left\{ \frac{(\kappa^2 + k^2)^2}{4\kappa^2 k^2} \right\} = 1 + \frac{1}{4} \frac{V_0^2}{E(V_0 - E)} \sinh^2(2\kappa a). \quad (286)$$

This non-zero transmission coefficient is again a quantum tunneling effect. As $E \rightarrow 0$ $T^{-1} \rightarrow \infty$, so there is no transmission for $E = 0$. On the other hand, if $E \rightarrow V_0$, i.e., the energy is just equal to the barrier height, then using $\sinh x \approx x$ for small x ,

$$\frac{1}{T} \rightarrow 1 + \frac{2m}{\hbar^2} V_0 a^2 > 1. \quad (287)$$

So the transmission coefficient is less than one even if the energy of the incoming wave is equal to the height of the barrier.

• We can get T for $E > V_0$ by putting $l^2 = 2m(E - V)/\hbar^2$ and $\kappa = il$ in the above formula and use $\sinh(ix) = i \sin x$:

$$\frac{1}{T} = 1 + \frac{1}{4} \frac{V^2}{E(E - V)} \sin^2(2la), \quad E > V_0. \quad (288)$$

As $E \rightarrow \infty$ we see that $T \rightarrow 1$ and there is complete transmission as we would expect. Interestingly, we also have $T = 1$ when $2la = n\pi$ for $n = 1, 2, \dots$ where l is the wave number in the region of the barrier $|x| < a$. If we call $\lambda = 2\pi/l$ the wave length in the region

of the barrier, then this condition for reflectionless transmission is $n\frac{\lambda}{2} = 2a$, i.e., an integer number of half-wavelengths must fit in the region of the barrier. The barrier is transparent at certain specific energies!

- We can also get the transmission coefficient for scattering over a finite square well by putting $V_0 \rightarrow -V_{well}$. Then

$$\frac{1}{T} = 1 + \frac{1}{4} \frac{V^2}{E(E + V_{well})} \sin^2(2la), \quad l^2 = 2m(E + V_{well})/\hbar^2. \quad (289)$$

Again we have reflectionless transmission if an integer number of half wavelengths fit inside the well. This condition is familiar to us since it also determines the energy levels of the infinite square well.

- Scattering by a finite square well can be used to model the scattering of electrons by atoms. The well represents the attraction between the nucleus and the incident electron. The above transparent scattering at the energy corresponding to $\frac{\lambda}{2} = 2a$ has been observed (Ramsauer effect).

10 Some simple quantum mechanical models in 2 dimensions

- Free particle: Separation of variables. Energy eigenstates are plane waves $e^{i(k_x x + k_y y)}$ for any real k_x, k_y . Energies $E = \hbar^2(k_x^2 + k_y^2)/2m$ labeled by k_x and k_y . Ground state $E = 0$ is non-degenerate. Infinite degeneracy of energy levels for $E > 0$ corresponding to any k_x, k_y lying on a circle of radius $\sqrt{2mE}/\hbar$.
- Two dimensional infinite square well. Separation of variables. Ground state is non-degenerate. Bound excited states can be degenerate in energy unlike in one dimension.
- Parity in two-dimensions.
- Two dimensional Simple Harmonic Oscillator.

10.1 Is every energy eigenfunction separable (factorizable as $\psi(x, y) = X(x)Y(y)$)?

- We assumed a separable form $\psi(x, y) = X(x)Y(y)$ in searching for energy eigenfunctions of the two dimensional free particle, SHO, square well etc. But must every eigenfunction be separable for these problems? In all these cases, the energy eigenfunctions corresponding to non-degenerate energy eigenvalues are in fact separable. Moreover, eigenfunctions corresponding to degenerate eigenvalues *can be taken separable*. The problem with degenerate energy levels is illustrated as follows. If we have two linearly independent separable eigenfunctions, say $X_1(x)Y_2(y)$ and $X_2(x)Y_1(y)$ with the same energy E , then the non-separable states $X_1(x)Y_2(y) \pm X_2(x)Y_1(y)$ are also linearly independent eigenfunctions with the same energy E . So energy eigenfunctions are not necessarily separable, but a basis within each eigenspace can be found, consisting of separable eigenfunctions.

- Let us sketch why this is the case, with the example of a particle in an infinite square well in the region $0 \leq x, y \leq L$. $H = H_x + H_y$, $H_x = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2}$, $H_y = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial y^2}$. We know that H_x and H_y each have a complete orthonormal family of eigenfunctions $\psi_n(x) = \sqrt{\frac{2}{L}} \sin \frac{n\pi x}{L}$ and $\psi_m(y) = \sqrt{\frac{2}{L}} \sin \frac{n\pi y}{L}$ with energies $E_n = \frac{n^2 \pi^2 \hbar^2}{2mL^2}$. Using these we have a basis for the

Hilbert space of square-integrable functions of two variables (vanishing on and outside the well). Any such function $\psi(x, y)$ can be expanded uniquely as

$$\psi(x, y) = \sum_{n, m \geq 1} c_{nm} \psi_n(x) \psi_m(y) \quad (290)$$

Such a function is separable if all but one of the c_{nm} is zero. Now suppose such a function is an eigenfunction of H with energy E . Then we have on the one hand

$$H\psi = E\psi = E \sum_{nm} c_{nm} \psi_n(x) \psi_m(y) \quad (291)$$

and on the other hand,

$$H\psi = \sum_{n, m} c_{nm} (H_x + H_y) \psi_n(x) \psi_m(y) = \sum_{n, m} c_{nm} (E_n + E_m) \psi_n(x) \psi_m(y). \quad (292)$$

Equating the two expressions and noting that $H\psi$ has a unique expansion in the basis of separable functions, we must have

$$Ec_{nm} = (E_n + E_m)c_{nm} \quad \text{for every } n, m \quad (293)$$

Now this is a very peculiar equation, and it is almost impossible to satisfy. For it to hold, either

$$c_{nm} = 0 \quad \text{or} \quad E = E_n + E_m \quad (294)$$

Now for most values of m, n $E_n + E_m$ will not equal E , and all those corresponding c_{nm} must vanish. But there must be some n, m , say $n = 1, m = 2$ for which $E = E_1 + E_2$. Then c_{12} could be non-zero. If there is another way in which $E = E_m + E_n$, then that corresponding c_{nm} could also be non-zero. This is the case if E is a degenerate energy level. In this example, c_{21} could also be non-zero. In other words, the energy eigenfunction $\psi(x, y)$ must be a linear combination of separable functions which are all degenerate in energy. So if E is non-degenerate, then, $\psi(x, y)$ is separable. In the case of degenerate E , the corresponding energy eigenfunctions $\psi(x, y)$ can always be taken separable by choosing only one of the c_{nm} to be non-zero.

11 Quantum mechanics in three dimensions: central potentials

11.1 Free particle in three dimensions

- The free particle hamiltonian in 3d is $H = \vec{p}^2/2m$ where $\vec{p} = (p_x, p_y, p_z) = -i\hbar(\partial_x, \partial_y, \partial_z) = -i\hbar\nabla$. So $H = -\frac{\hbar^2}{2m}\nabla^2$. The stationary states are obtained by separation of variables

$$\psi_{\vec{k}}(x, y, z, t) = e^{i(k_x x + k_y y + k_z z - \omega t)} = e^{i(\vec{k} \cdot \vec{r} - \omega t)}, \quad E = \frac{\hbar^2 k^2}{2m} = \hbar\omega. \quad (295)$$

$\vec{k} = (k_x, k_y, k_z)$ is the wave vector whose squared-length is $k^2 = k_x^2 + k_y^2 + k_z^2$. At any fixed time, the probability amplitude is constant on the planes $\vec{k} \cdot \vec{r} = \text{const} + 2n\pi$. These are planes orthogonal to the wave vector, so the stationary states are plane waves. Points of constant phase travel at the phase speed $c = \omega/k$.

- Each energy eigenvalue $E > 0$ is infinitely degenerate corresponding to points \vec{k} on a sphere of radius $\sqrt{2mE}/\hbar$. Energy eigenstates are orthogonal and delta normalized

$$\int \psi_{\vec{k}}(r)\psi_{\vec{k}'}(r) d\vec{r} = (2\pi)^3 \delta(x-x')\delta(y-y')\delta(z-z') \equiv (2\pi)^3 \delta(\vec{r}-\vec{r}'). \quad (296)$$

They are also complete, we can expand any square-integrable state as a linear combination

$$\psi(\vec{r}) = \int \tilde{\psi}(\vec{k})e^{i\vec{k}\cdot\vec{r}} d^3\vec{k}. \quad (297)$$

The group velocity of a wave packet peaked around $\vec{k} = \vec{k}_*$ is $\vec{v}_g = \left(\frac{\partial\omega}{\partial k_x}, \frac{\partial\omega}{\partial k_y}, \frac{\partial\omega}{\partial k_z} \right)_{|\vec{k}=\vec{k}_*}$.

- $|\psi(x, y, z)|^2 dx dy dz$ is interpreted as the probability of finding the particle in a small volume element $dx dy dz$ near (x, y, z) .

- The free particle hamiltonian commutes with each component of momentum $[H, \vec{p}] = 0$, they are both diagonal in the above basis of plane waves. H, p_x, p_y, p_z form a complete set of commuting observables. Another way of specifying energy eigenstates uses angular momentum.

- We will eventually be interested in the hydrogen atom, which corresponds to an electron moving in the spherically symmetric potential $V(r)$ due to the proton. So it is useful to work in spherical coordinates. Let us first express the classical free particle hamiltonian in spherical coords. Recall that $L^2 = (r \times p)^2 = r^2 p^2 - (r \cdot p)^2$. If we define the radial momentum $p_r = \frac{r \cdot p}{r}$

$$H = \frac{p^2}{2m} = \frac{p_r^2}{2m} + \frac{L^2}{2mr^2} \quad (298)$$

In the quantum theory, H must be represented by a hermitian operator. In particular, we must understand the concept of angular momentum in quantum mechanics, to which we turn now. We will return to the free particle and spherically symmetric potentials subsequently.

11.2 Basic properties of angular momentum

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

$$L_x = yp_z - zp_y, \quad L_y = zp_x - xp_z, \quad L_z = xp_y - yp_x \quad \text{Cyclic: } x \rightarrow y \rightarrow z \rightarrow x \quad (299)$$

We could also write them in terms of the antisymmetric Levi-Civita tensor

$$L_i = \epsilon_{ijk} r_j p_k, \quad \epsilon_{123} = +1, \epsilon_{132} = -1, \dots \quad (300)$$

- In an early assignment we saw that each component of angular momentum is a conserved quantity for a particle in a central potential $V(r)$.

- Angular momentum has the dimensions of \hbar .
- In QM, angular momentum is represented in the position basis by the operators

$$L_x = \hat{y}p_z - \hat{z}p_y = -i\hbar \left(y \frac{\partial}{\partial z} - z \frac{\partial}{\partial y} \right), \dots \quad (301)$$

Note that y and p_z commute so it does not matter whether we write yp_z or zp_y etc.

- 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. 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

$$[L_x, L_y] = i\hbar L_z, \quad [L_y, L_z] = i\hbar L_x, \quad [L_z, L_x] = i\hbar L_y \quad \text{cyclic.} \quad (302)$$

The components of angular momentum aren't simultaneously diagonalizable:

$$(\Delta L_x)^2(\Delta L_y)^2 \geq -\frac{1}{4}\langle [L_x, L_y] \rangle^2 = \frac{\hbar^2}{4}\langle L_z \rangle^2 \Rightarrow \Delta L_x \Delta L_y \geq \frac{\hbar}{2}|\langle L_z \rangle|. \quad (303)$$

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

$$z = r \cos \theta, \quad x = r \sin \theta \cos \phi, \quad y = r \sin \theta \sin \phi. \quad (304)$$

Calculate derivatives by the chain rule to get

$$L_z = -i\hbar \frac{\partial}{\partial \phi}, \quad L_x = i\hbar \left(\sin \phi \frac{\partial}{\partial \theta} + \cot \theta \cos \phi \frac{\partial}{\partial \phi} \right), \quad L_y = i\hbar \left(-\cos \phi \frac{\partial}{\partial \theta} + \cot \theta \sin \phi \frac{\partial}{\partial \phi} \right) \quad (305)$$

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

$$-i\hbar \partial_\phi = -i\hbar \left(\frac{\partial x}{\partial \phi} \frac{\partial}{\partial x} + \frac{\partial y}{\partial \phi} \frac{\partial}{\partial y} + \frac{\partial z}{\partial \phi} \frac{\partial}{\partial z} \right) = -i\hbar(-y\partial_x + x\partial_y) = L_z. \quad (306)$$

- Interesting consequence: an uncertainty relation for measurements of L_ϕ and ϕ on the same state.

$$[\phi, L_z] = -i\hbar \left[\phi, \frac{\partial}{\partial \phi} \right] = i\hbar \Rightarrow \Delta \phi \Delta L_z \geq \frac{\hbar}{2}. \quad (307)$$

More generally, if ϕ is the azimuthal angle about any axis and L_ϕ is the component of angular momentum along that axis, then ϕ and L_ϕ cannot both be measured with arbitrary precision on the same state. This is the angular analogue of the position-linear momentum uncertainty relation.

- Of course, we can also express L^2 in spherical coordinates

$$L^2 \psi = -\hbar^2 \left[\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left(\sin \theta \frac{\partial \psi}{\partial \theta} \right) + \frac{1}{\sin^2 \theta} \frac{\partial^2 \psi}{\partial \phi^2} \right]. \quad (308)$$

11.2.1 Angular momentum and rotations

- Recall that momentum p_x generates translations in x : $f(x + \delta x) \approx f(x) + \frac{i}{\hbar} \delta x p f(x)$ and in particular, p_x annihilates any function that is translation invariant in x : $p_x f(y, z) = -i\hbar \frac{\partial}{\partial x} f(y, z) = 0$. And if a function is unchanged under translations in x , y and z , then it is annihilated by $\vec{p} = -i\hbar \nabla$.

- It turns out that L_x, L_y, L_z generate rotations about the x, y, z axes of 3d space. More generally, a counter clockwise rotation about an axis along the vector \hat{n} is generated by $\hat{n} \cdot \vec{L}$. As motivation, let us first show that L_x annihilates a function symmetric with respect to rotations about the x -axis. Any such function can be expressed as $f(x, y, z) = f(x, \rho)$ where $\rho^2 = y^2 + z^2$, i.e. f is independent of the azimuthal angle around the x -axis. We use $\frac{\partial \rho}{\partial z} = \frac{z}{\rho}, \frac{\partial \rho}{\partial y} = \frac{y}{\rho}, \frac{\partial \rho}{\partial x} = 0$ to find L_x annihilates $f(x, \rho)$:

$$\frac{i}{\hbar} L_x f(x, \rho) = \left(y \frac{\partial}{\partial z} - z \frac{\partial}{\partial y} \right) f(x, \rho) = \left(y \frac{z}{\rho} - z \frac{y}{\rho} \right) \frac{\partial f(x, \rho)}{\partial \rho} = 0. \quad (309)$$

- If a function (e.g. the potential) $V(x, y, z)$ depends only on the radial coordinate, $V = V(r)$ is called rotationally invariant or spherically symmetric. A spherically symmetric function is in particular symmetric under rotations about the x, y or z axis. So a spherically symmetric function should be annihilated by each component of angular momentum. Let us see how L_x acts on $V(r)$. We use $\frac{\partial r}{\partial z} = \frac{z}{r}$ etc to find that L_x annihilates $f(r)$

$$\frac{i}{\hbar} L_x f(r) = \left(y \frac{\partial}{\partial z} - z \frac{\partial}{\partial y} \right) f(r) = y \frac{z}{r} f'(r) - z \frac{y}{r} f'(r) = 0. \quad (310)$$

Similarly, L_y and L_z annihilate any spherically symmetric function $V(r)$.

- This is also a consequence of the formulae (305) for L_x, L_y, L_z . They don't involve the radial coordinate r , so automatically, they annihilate any function $f(r)$ that is independent of θ and ϕ .

- Now we show that \vec{L} generates infinitesimal rotations in 3d. Consider the rotation of a vector \vec{r} by a small angle $\delta\phi$ counter clockwise about an axis. It is convenient to define a vector $\delta\vec{\phi}$ along that axis with a length equal to $\delta\phi$. Then the change in r is orthogonal to both $\delta\vec{\phi}$ and to \vec{r} and is seen to be $\delta\phi \times r$. Thus, the rotated vector is $\vec{r} \rightarrow \vec{r} + \delta\vec{\phi} \times \vec{r}$. Now what happens to a function $f(\vec{r})$ when \vec{r} is rotated slightly? We find out using the multi-variable Taylor expansion

$$f(\vec{r} + \delta\vec{\phi} \times \vec{r}) = f(\vec{r}) + \left(\delta\vec{\phi} \times \vec{r} \right) \cdot \vec{\nabla} f(\vec{r}) \quad (311)$$

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 $p = -i\hbar \nabla$,

$$f(\vec{r} + \delta\vec{\phi} \times \vec{r}) \approx f(\vec{r}) + \frac{i}{\hbar} \delta\vec{\phi} \cdot (r \times p) f(\vec{r}) = \left(I + \frac{i}{\hbar} \delta\vec{\phi} \cdot \vec{L} \right) f(\vec{r}). \quad (312)$$

Comparing with how p generates translations, we say angular momentum generates rotations.

$$f(x + \delta x) \approx \left(I + \frac{i}{\hbar} (\delta x) p \right) f(x). \quad (313)$$

- In light of this, if a function $f(\vec{r}) = f(r)$ is rotationally-invariant (i.e. spherically symmetric), then its change under a small rotation about *any* axis must vanish. In particular, $\vec{L}f(r) = 0$ for any spherically symmetric function as found earlier. A state $|f\rangle$ with a spherically symmetric wave function $\langle \vec{r}|f\rangle = f(r)$ is called an S -wave state. It follows that $L^2f(r)$ is also zero. So L^2 annihilates spherically symmetric functions.

11.3 Free particle hamiltonian as a hermitian operator in spherical coordinates

- Recall that the classical free particle hamiltonian was written as $H = \frac{p_r^2}{2m} + \frac{L^2}{2mr^2}$. How do we make sense of the second term as an operator in the quantum theory? L^2 and $\frac{1}{r^2}$ are both operators. Does it matter whether we write $\frac{1}{r^2}L^2$ or $L^2\frac{1}{r^2}$? No, as we show below.
- An important consequence of the fact that \vec{L} generates rotations and $L^2f(r) = 0$ is that $[L^2, f(r)] = 0$ for any spherically symmetric function $f(r)$. In more detail, we calculate

$$[L^2, f(r)]\psi(\vec{r}) = L^2(f(r)\psi(r, \theta, \phi)) - fL^2\psi(r, \theta, \phi) \quad (314)$$

for any function $\psi(\vec{r})$. Since L^2 by (308) is independent of r , it does not affect $f(r)$ at all, so $L^2(f(r)\psi) = f(r)L^2\psi$. In particular, taking $f(r) = 1/r^2$ we have $L^2\frac{1}{r^2} = \frac{1}{r^2}L^2$. Moreover, L^2/r^2 is a hermitian operator since L^2 and $1/r^2$ are both hermitian and they commute.

- It remains to write radial momentum $\frac{\vec{r}\cdot\vec{p}}{r}$ and p_r^2 as operators in the quantum theory. However, p does not commute with the multiplication operator $\hat{r} = \frac{\vec{r}}{r} = (\frac{x}{r}, \frac{y}{r}, \frac{z}{r})$. So in which order must we write the operators? In addition, suppose we define the quantum mechanical operator

$$\mathbf{p} = \frac{1}{r}\vec{r}\cdot\vec{p} = \frac{1}{r}(xp_x + yp_y + zp_z) \quad (315)$$

This operator is not hermitian! Its adjoint on functions that vanish at infinity is

$$\mathbf{p}^\dagger = (p_x x + p_y y + p_z z)\frac{1}{r} = \vec{p}\cdot\frac{\vec{r}}{r} \quad (316)$$

Classically of course, \mathbf{p} and \mathbf{p}^\dagger are the same. We could just as well have written the classical radial momentum as $\frac{1}{2}(\hat{r}\cdot\vec{p} + \vec{p}\cdot\hat{r})$. In the quantum theory we want to represent radial momentum by a hermitian operator. If A is any operator, $A+A^\dagger$ is automatically hermitian. So we define radial momentum in the quantum theory as the hermitian symmetric combination

$$p_r\psi(x, y, z) = \frac{1}{2}(\hat{r}\cdot\vec{p} + \vec{p}\cdot\hat{r})\psi = \frac{-i\hbar}{2}(\hat{r}\cdot\nabla + \nabla\cdot\hat{r})\psi = -i\hbar\left(\frac{1}{r} + \frac{\partial}{\partial r}\right)\psi = -i\hbar\frac{1}{r}\frac{\partial}{\partial r}(r\psi). \quad (317)$$

Here we used $\hat{r}\cdot\nabla = \frac{x}{r}\partial_x + \frac{y}{r}\partial_y + \frac{z}{r}\partial_z = \frac{\partial}{\partial r}$. This is because $x = r\sin\theta\cos\phi \Rightarrow \frac{\partial x}{\partial r} = \frac{x}{r}$ etc. Similarly, $\nabla\cdot(\hat{r}\psi) = \hat{r}\cdot\nabla\psi + \psi\nabla\cdot\hat{r} = \psi_r + \frac{2\psi}{r}$.

- An interesting corollary is that $-i\hbar\frac{\partial}{\partial r}$ is not hermitian! This is because of the factor $r^2 dr d\Omega$ in the volume element that appears when we compute inner products. Here $d\Omega = \sin\theta d\theta d\phi$ is the angular volume element. Check that $p_r = -i\hbar(\frac{1}{r} + \partial_r)$ is hermitian $(p_r f, g) = (f, p_r g)$ if $f, g \rightarrow 0$ as $r \rightarrow \infty$:

$$\int \left(f_r^* + \frac{f^*}{r}\right) g r^2 dr d\Omega = - \int f^* \left(g_r + \frac{g}{r}\right) r^2 dr d\Omega \quad (318)$$

- Since p_r is hermitian, p_r^2 is also hermitian. Thus the radial kinetic term in $H\psi$ is

$$\frac{p_r^2}{2m}\psi = -\frac{\hbar^2}{2m} \frac{1}{r} \partial_r \left[r \frac{1}{r} \partial_r (r\psi) \right] = -\frac{\hbar^2}{2m} \frac{1}{r} \partial_r^2 (r\psi) = -\frac{\hbar^2}{2m} \frac{1}{r^2} \partial_r (r^2 \psi_r) \quad (319)$$

The last two expressions are just different re-writings of the same thing. So $H = \frac{p_r^2}{2m} + \frac{L^2}{2mr^2}$ acts on $\psi(r, \theta, \phi)$ as the differential operator (subscripts are partial derivatives)

$$H = -\frac{\hbar^2}{2m} \frac{1}{r} (r\psi)_{rr} - \frac{\hbar^2}{2mr^2} \left(\frac{1}{\sin \theta} (\sin \theta \psi_\theta)_\theta + \frac{1}{\sin^2 \theta} \psi_{\phi\phi} \right). \quad (320)$$

We'd get the same result by starting with $H = -\frac{\hbar^2}{2m} (\partial_x^2 + \partial_y^2 + \partial_z^2)$ and just transforming to spherical coordinates. This justifies our symmetric choice $p_r = \frac{1}{2}(\mathbf{p} + \mathbf{p}^\dagger)$ for the radial momentum operator.

11.3.1 Positivity of free particle hamiltonian, radial and angular kinetic energies

- Recall that the energy eigenvalues of the free particle hamiltonian are $E_{\vec{k}} = \hbar^2 k^2 / 2m$ where $k^2 = k_x^2 + k_y^2 + k_z^2$. $E_{\vec{k}} \geq 0$ implies that $\langle H \rangle \geq 0$ in every state $|\psi\rangle$. For, suppose $|\psi\rangle = \int [d^3k] \tilde{\psi}(\vec{k}) e^{i\vec{k}\cdot\vec{r}}$, then

$$\langle \psi | H | \psi \rangle = \int [d^3k] |\tilde{\psi}(\vec{k})|^2 E_{\vec{k}} \geq 0. \quad (321)$$

- If the expectation value of an operator H is positive in every state, we say that the operator $H \geq 0$ is positive⁴⁶. Note that to be positive, every expectation value must in particular be real. So a positive operator is in particular hermitian. So the free particle hamiltonian is a positive operator.

- $A^\dagger A$ and AA^\dagger are both positive for any operator A . To see this, consider the expectation value

$$\langle \psi | A^\dagger A | \psi \rangle = \langle A\psi | A\psi \rangle = \|A\psi\|^2 \geq 0, \quad \langle \psi | AA^\dagger | \psi \rangle = \langle A^\dagger \psi | A^\dagger \psi \rangle = \|A^\dagger \psi\|^2 \geq 0. \quad (322)$$

- We use this to show (once again) that the free particle hamiltonian is positive. For this we recall that each component of momentum is hermitian. So

$$2mH = p_x^2 + p_y^2 + p_z^2 = p_x^\dagger p_x + p_y^\dagger p_y + p_z^\dagger p_z \geq 0. \quad (323)$$

- Likewise, radial kinetic energy is positive on account of hermiticity of radial momentum

$$T_r = \frac{p_r^2}{2m} = \frac{1}{2m} p_r^\dagger p_r \geq 0 \quad (324)$$

- Similarly, the square of total angular momentum $L^2 = \vec{L} \cdot \vec{L} = L_x^\dagger L_x + L_y^\dagger L_y + L_z^\dagger L_z = \vec{L}^\dagger \cdot \vec{L}$ is a positive operator. In particular, $H = \frac{p_r^2}{2m} + \frac{L^2}{2mr^2}$ is a positive operator.

⁴⁶If $\langle H \rangle_\psi > 0$ for every $\psi \neq 0$, then H is strictly-positive. If $\langle H \rangle_\psi \geq 0$ for every ψ then H is positive (or non negative).

11.4 Conservation of angular momentum for a particle in a central potential

• Consider a particle in a central potential $V(r)$. Classically we found that each component of angular momentum about $r = 0$ is conserved in time. As a consequence L^2 is also conserved classically. L^2, L_x, L_y, L_z are all conserved in the quantum theory in the sense that their expectation values in any state $\psi(r, \theta, \phi)$ are constant in time. This is because each of them commutes with the hamiltonian

$$H = \frac{p_r^2}{2m} + \frac{L^2}{2mr^2} + V(r). \quad (325)$$

• E.g. $[L^2, H] = [L^2, \frac{p_r^2}{2m}] + [L^2, \frac{L^2}{2mr^2}] + [L^2, V(r)]$. The first term vanishes since L^2 is a differential operator in θ, ϕ while p_r only involves r , so $L^2 p_r^2 = p_r^2 L^2$. On the other hand, we observed in the last section that L^2 commutes with any spherically symmetric function. It follows that the last two terms are zero. So $[L^2, H] = 0$. Energy and total angular momentum are simultaneously diagonalizable, they are compatible observables.

• $L_z = -i\hbar \frac{\partial}{\partial \phi}$ only involves the angular variables, so it commutes with p_r^2 and $V(r)$. We already observed that $[L_z, L^2] = 0$, so $[L_z, \frac{L^2}{r^2}] = 0$. It follows that $[L_z, H] = 0$. Similarly, we show that $[H, \vec{L}] = 0$, so each component of angular momentum is compatible with the hamiltonian.

• In particular, H, L^2, L_z are a set of commuting observables. It turns out that this is a complete set of commuting observables for a particle in a central potential. However, we cannot add L_x to the CSCO as it is not compatible with L_z . But we are free to *replace* L_z by L_x .

11.5 Separation of variables for particle in central potential

• Now we address the problem of finding the energy eigenstates of a particle in a central potential $V(r)$. We are particularly interested in the attractive electrostatic potential towards the nucleus $V(r) \propto -\frac{1}{r}$ felt by the electrons in an atom. But the free particle and particle in a spherical well etc are also interesting. Due to the spherical symmetry of $V(r)$, it is convenient to work in spherical coordinates. We seek stationary states $\psi(r, \theta, \phi)$ of the hamiltonian $\frac{p_r^2}{2m} + \frac{L^2}{2mr^2} + V$:

$$H\psi = -\frac{\hbar^2}{2m} \frac{1}{r} (r\psi)_{rr} + \frac{L^2}{2mr^2} \psi + V(r)\psi = E\psi \quad (326)$$

H has been separated into radial and angular parts, so we try SOV $\psi(r, \theta, \phi) = R(r)Y(\theta, \phi)$.

$$-\frac{\hbar^2}{2m} \frac{Y}{r} (rR)_{rr} + \frac{R}{2mr^2} L^2 Y + V(r)RY = ERY. \quad (327)$$

To separate radial and angular dependences we divide by RY and multiply by $\frac{2mr^2}{\hbar^2}$

$$\frac{r}{R} (rR)_{rr} + (E - V(r)) \frac{2mr^2}{\hbar^2} = \frac{1}{\hbar^2 Y} L^2 Y. \quad (328)$$

The rhs depends only on θ, ϕ while the lhs depends only on r . So they must both equal a constant, say λ . In particular, the angular wavefunction Y must be an eigenfunction of L^2

with eigenvalue $\hbar^2\lambda$

$$L^2Y(\theta, \phi) = \hbar^2\lambda Y(\theta, \phi). \quad (329)$$

We have already seen that L^2 is a positive operator, so $\lambda \geq 0$. Moreover, the constant function $Y(\theta, \phi) = \frac{1}{\sqrt{4\pi}}$ is an eigenfunction with eigenvalue $\lambda = 0$. The peculiar factor ensures $Y(\theta, \phi)$ has norm one with respect to the inner product on the space of angular wave functions

$$\langle Y|Z \rangle = \int Y^*(\theta, \phi)Z(\theta, \phi) d\Omega, \quad \text{where } d\Omega = \sin\theta d\theta d\phi, \quad \int d\Omega = 4\pi. \quad (330)$$

More generally, we will show that the eigenvalues of L^2 are $l(l+1)\hbar^2$ for $l = 0, 1, 2, \dots$. In anticipation of that, we rename $\lambda = l(l+1)$. Each eigenvalue $l(l+1)\hbar^2$ is $2l+1$ -fold degenerate, corresponding to the distinct eigenvalues $-l, -l+1, \dots, 0, 1, \dots, l-1, l$ of L_z/\hbar , which is simultaneously diagonalizable with L^2 . The common eigenstates of L^2, L_z are denoted $|l, m\rangle$ and are represented by the orthonormal spherical harmonics $Y_{lm}(\theta, \phi) = \langle \theta, \phi | l, m \rangle$. For $l = 0$, $L^2Y = 0$ and Y is a constant. So the $l = 0$ wave function is spherically symmetric $\psi(r, \theta, \phi) = \frac{1}{\sqrt{4\pi}}R(r)$. Energy eigenstates with zero angular momentum $l = 0$ are called S-wave states. The eqn for the radial wavefunction $R(r)$ is

$$\left[-\frac{1}{r} \frac{d^2}{dr^2} r + \frac{l(l+1)}{r^2} + \frac{2mV(r)}{\hbar^2} \right] R = \frac{2mE}{\hbar^2} R. \quad (331)$$

The radial kinetic term may also be written as $-\frac{1}{r^2}(r^2R)'$. To proceed further $V(r)$ must be specified. It is often convenient to introduce the auxiliary radial variable $u(r) = rR(r)$ in terms of which

$$-\frac{\hbar^2}{2m} \frac{d^2u}{dr^2} + \left[\frac{\hbar^2}{2m} \frac{l(l+1)}{r^2} + V(r) \right] u(r) = Eu(r) \quad (332)$$

The radial eigenvalue problem for u is a one dimensional time independent SE, except that it is defined on the positive half line $r > 0$, and the potential $V(r)$ is modified by the centrifugal repulsive term $\frac{\hbar^2}{2m} \frac{l(l+1)}{r^2}$. A particle in a stationary state with greater angular momentum quantum number l should have less probability to be found near $r = 0$, as in classical mechanics.

11.6 Spectrum of L^2 and L_z : spherical harmonics

Separation of the radial and angular parts of the wave function $\psi = RY$ for a particle in a central potential implied that Y must be an eigenfunction of L^2 : $L^2Y(\theta, \phi) = \hbar^2\lambda Y$. In other words,

$$-\left(\frac{1}{\sin\theta} (\sin\theta Y_\theta)_\theta + \frac{1}{\sin^2\theta} Y_{\phi\phi} \right) = \lambda Y \quad (333)$$

We seek eigenfunctions which have a finite norm $\|Y\|^2 = \int |Y(\theta, \phi)|^2 d\Omega$. Since the angular equation is independent of the potential, $Y(\theta, \phi)$ should be finite, continuous and have continuous first derivative everywhere. $Y(\theta, \phi) = \Theta(\theta)\Phi(\phi)$ can be further separated, so that

$$\sin^2\theta \left(\lambda + \frac{1}{\sin\theta} \frac{(\Theta' \sin\theta)'}{\Theta} \right) = -\frac{\Phi''(\phi)}{\Phi(\phi)} = \nu \quad (334)$$

for some constant ν . The equation $\Phi''(\phi) = -\nu\Phi(\phi)$ for the azimuthal angle ϕ is familiar. But we solve it with the requirement that the wave function be continuous (single-valued): $\Phi(0) = \Phi(2n\pi)$ for $n \in \mathbb{Z}$, since $\phi = 2n\pi$ represent the same physical longitude. This implies $\nu \geq 0$ so we write $\nu = m^2$ and so up to a normalization constant, $\Phi(\phi) = e^{im\phi}$ where $m \in \mathbb{Z}$. Interestingly, $\Phi(\phi)$ is also an eigenfunction of $L_z = -i\hbar\frac{\partial}{\partial\phi}$ with eigenvalue $\hbar m$. So any eigenfunction $Y(\theta, \phi)$ of L^2 will automatically be an eigenfunction of L_z !

- Meanwhile the equation for the latitude θ becomes

$$\frac{1}{\sin\theta}(\Theta' \sin\theta)' + \left(\lambda - \frac{m^2}{\sin^2\theta}\right)\Theta(\theta) = 0, \quad 0 \leq \theta \leq \pi. \quad (335)$$

Substitute $x = \cos\theta$ and write $\Theta(\theta) = y(x)$. Then $\frac{\partial}{\partial\theta} = -\sin\theta\frac{\partial}{\partial x}$ or $\Theta'(\theta) = -\sin\theta y'(x)$ and we get

$$(1-x^2)y'' - 2xy' + \left(\lambda - \frac{m^2}{1-x^2}\right)y = 0, \quad -1 \leq x \leq 1 \quad (336)$$

For $m = 0$, (336) is Legendre's equation. For generic values of λ the solutions diverge at $x = \pm 1$ ($\theta = 0, \pi$). Only for $\lambda = l(l+1)$ where $l = 0, 1, 2, 3, \dots$ do we get normalizable solutions (the spectrum of L^2 is obtained using raising and lowering operators in the next section). For $m = 0$ $y_l(x)$ are Legendre's orthogonal polynomials $P_l(x) = P_l(\cos\theta)$

$$P_0 = 1, \quad P_1 = x, \quad P_2 = \frac{1}{2}(3x^2 - 1), \quad P_3 = \frac{1}{2}(5x^3 - 3x), \dots \quad -1 \leq x \leq 1. \quad (337)$$

P_l are polynomials of degree l in $\cos\theta$ and can be obtained using the Rodrigues formula

$$P_l(x) = \frac{1}{2^l l!} \left(\frac{d}{dx}\right)^l (x^2 - 1)^l, \quad l = 0, 1, 2, \dots \quad (338)$$

More generally, the solutions to (336) for the physically acceptable $m = -l, -l+1, \dots, 0, \dots, l-1, l$ are called associated Legendre functions $P_l^m(x)$ which can be obtained from the Legendre polynomials by successive differentiation

$$P_l^m(x) = (1-x^2)^{\frac{1}{2}|m|} \left(\frac{d}{dx}\right)^{|m|} P_l(x), \quad -l \leq m \leq l. \quad (339)$$

Combining, we obtain the eigenfunctions of L^2 , the spherical harmonics

$$Y_{lm}(\theta, \phi) = N_{l,m} P_l^m(\cos\theta) e^{im\phi} \quad (340)$$

The normalization constants $N_{l,m}$ are chosen so that Y_{lm} are orthonormal

$$\int Y_{lm}^*(\theta, \phi) Y_{l'm'}(\theta, \phi) \sin\theta d\theta d\phi = \delta_{ll'} \delta_{mm'}. \quad (341)$$

Summary: The spherical harmonics $Y_{lm}(\theta, \phi)$ are a complete set of simultaneous orthonormal eigenfunctions of L^2 and L_z with the eigenvalues

$$L^2 Y_{lm} = \hbar^2 l(l+1) Y_{lm}, \quad l = 0, 1, 2, \dots \quad L_z Y_{lm} = \hbar m Y_{lm}, \quad m = -l, -l+1, \dots, 0, \dots, l-1, l. \quad (342)$$

$l \rightarrow$ 'azimuthal' or angular momentum quantum number, $m \rightarrow$ 'magnetic' quantum number. The eigenvalues $\hbar^2 l(l+1)$ of L^2 are independent of m . Each eigenvalue of L^2 is $(2l+1)$ -fold degenerate.

11.7 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 $\hbar m$

$$L^2\psi = \hbar^2 \lambda \psi, \quad L_z\psi = \hbar m \psi \quad (343)$$

To find the allowed values of λ, m we use the ladder operator method. Define $L_{\pm} = L_x \pm L_y$. The virtue of L_{\pm} is that if ψ is one such simultaneous eigenstate, then so are $L_{\pm}\psi$. Indeed,

$$L^2(L_{\pm}\psi) = \hbar^2\lambda(L_{\pm}\psi), \quad L_z(L_{\pm}\psi) = \hbar(m \pm 1)(L_{\pm}\psi). \quad (344)$$

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_+

$$L_z\psi_h = \hbar l\psi_h, \quad L_+\psi_h = 0. \quad (345)$$

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}

$$L^2 = L_-L_+ + L_z^2 + \hbar L_z \Rightarrow L^2\psi_h = (\hbar^2 l^2 + \hbar^2 l)\psi_h = \hbar^2 l(l+1)\psi_h \quad (346)$$

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_-

$$L_z\psi_{low} = \hbar l'\psi_{low}, \quad L_-\psi_{low} = 0, \quad l' \leq l \quad (347)$$

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

$$L^2 = L_+L_- + L_z^2 - \hbar L_z \Rightarrow L^2\psi_{low} = (\hbar^2 l'^2 - \hbar^2 l')\psi_{low} = \hbar^2 l'(l'-1)\psi_{low} \quad (348)$$

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

$$l(l+1) = l'(l'-1), \quad l \geq l'. \quad (349)$$

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

$$l = 0, \pm\frac{1}{2}, \pm 1, \pm\frac{3}{2}, \pm 2 \dots \quad (350)$$

And for each such value of l , the possible eigenvalues of L_z are $\hbar(-l, -l+1, \dots, l-l, l)$. However, there is a small problem with the non-integer values of l . We already saw in the last section that the eigenvalues of L_z are $\hbar m$ 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 $l = 0, 1, 2, \dots$ 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 $l = 0, 1, 2, \dots$. 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 $\langle lm|l'm'\rangle = \delta_{mm'}\delta_{ll'}$.

- 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 $\Delta L_x \Delta L_y \geq \frac{\hbar}{2} |\langle L_z \rangle|$.

11.8 Free particle radial wave function

- In the case of a free particle, $V(r) = 0$. It is conventional to write the energy eigenvalue $E = \frac{\hbar^2 k^2}{2m}$ since these were the eigenvalues of H when we solved the problem in Cartesian coordinates⁴⁷. The radial equation becomes

$$\left[-\frac{1}{r} \frac{d^2}{dr^2} r + \frac{l(l+1)}{r^2} \right] R(r) = k^2 R(r) \quad \text{or} \quad R_{rr} + \frac{2}{r} R_r + \left(k^2 - \frac{l(l+1)}{r^2} \right) R = 0. \quad (351)$$

The radial wave function depends on k, l so we call it $R_{kl}(r)$ sometimes. We would at first seek solutions that are normalizable with respect to the inner product

$$\langle R|S \rangle = \int R^*(r) S(r) r^2 dr. \quad (352)$$

However, we do not expect any normalizable solutions, since in cartesian coordinates, the stationary states were plane waves. So we anticipate solutions that represent states with finite probability within any bounded region, but that behave like spherical waves as $r \rightarrow \infty$. The stationary states of energy E we seek now are those with definite total angular momentum. So they must be linear combinations of plane waves with the same energy E but various different directions for linear momentum $\hbar \vec{k}$. We can guess that such a spherical wave would have to be a linear combination of $\frac{\sin kr}{r}$ and $\frac{\cos kr}{r}$ for large enough r , since the volume element has a factor of r^2 built in. In terms of $u = rR$,

$$-u''(r) + \frac{l(l+1)}{r^2} u(r) = k^2 u(r) \quad (353)$$

The S-wave sector $l = 0$ is the simplest, and $u(r) = rR = A \sin kr + B \cos kr$ or

$$R_{k,l=0}(r) = A \frac{\sin kr}{r} + B \frac{\cos kr}{r}, \quad \psi_{k,l=0}(r) = \frac{1}{\sqrt{4\pi}} \left(A \frac{\sin kr}{r} + B \frac{\cos kr}{r} \right) \quad (354)$$

This ψ has finite probability in any bounded region. However, we have derived it assuming $r \neq 0$. It turns out that $\nabla^2 \psi \propto B \delta^3(\vec{r})$ ⁴⁸. But being a free particle eigenfunction, the

⁴⁷The spectrum of H should not depend on which coordinate system we use!

⁴⁸For small r , $\psi(r) \sim \frac{B}{r\sqrt{4\pi}}$. Now $\nabla^2 \frac{1}{r} = -4\pi \delta^3(\vec{r})$. To see this, we first show $\nabla \cdot \frac{\hat{r}}{r^2} = -\frac{\hat{r}}{r^2}$ (using $\partial_x r^{-1} = -x r^{-3}$) which leads to $\nabla \cdot \nabla \frac{1}{r} = 0$ for $r > 0$ (using $\partial_x(xr^{-3}) = r^{-3} - 3x^2 r^{-5}$). So $\nabla^2 \frac{1}{r}$ can be non-vanishing only at $r = 0$, i.e., $\nabla^2 \frac{1}{r} = C \delta^3(\vec{r})$. To find C , we integrate over a sphere of radius r and use Gauss' theorem $\int \nabla \cdot \left(\frac{\hat{r}}{r} \right) \cdot d\vec{S} = \int \nabla^2 \left(\frac{1}{r} \right) dV$. Here dV is the volume element for integration over a solid sphere of radius r while $d\vec{S}$ is the outward area element for integration over the surface of the same sphere, $d\vec{S} = \hat{r} r^2 d\Omega$. Thus $\int \nabla \cdot \left(\frac{\hat{r}}{r} \right) \cdot d\vec{S} = \int -\frac{\hat{r}}{r^2} \cdot \hat{r} r^2 d\Omega = -4\pi$. On the other hand $\int \nabla^2 \left(\frac{1}{r} \right) dV = C$. So $C = -4\pi$.

laplacian of ψ must be $\propto \psi$ every where, including at $r = 0$. But $\psi(\vec{r})$ is not $\propto \delta^3(\vec{r})$. So in order to solve the free particle SE, we need $B = 0$. As a consequence, the probability density approaches a constant as $r \rightarrow 0$.

- More generally, for $l = 0, 1, 2, \dots$, we may reduce the radial equation to the Bessel equation. r, k are the only dimensional quantities in the radial eigenvalue problem. In terms of the dimensionless variable $x = kr$ $\tilde{R}(kr) = R(r)$ we get the spherical Bessel differential equation⁴⁹

$$\tilde{R}''(x) + \frac{2}{x}\tilde{R}'(x) + \left(1 - \frac{l(l+1)}{x^2}\right)\tilde{R}(x) = 0. \quad (355)$$

This equation has two linearly independent solutions for each value of $l = 0, 1, 2, \dots$, the spherical Bessel $j_l(x)$ and spherical Neumann $n_l(x)$ functions. E.g. $j_0(x) = \frac{\sin x}{x}$ and $n_0(x) = -\frac{\cos x}{x}$. However, $n_l(x) \propto x^{-l-1}$ for $x \rightarrow 0$ represent states with infinite probability for the particle to be in any neighborhood of the origin (if $l > 0$). So we discard these solutions. On the other hand, $j_l(x) \propto x^l$ for $x \rightarrow 0$ and are admitted as scattering states.

$$\text{E.g., } j_0(x) = \frac{\sin x}{x}, \quad j_1(x) = \frac{\sin x}{x^2} - \frac{\cos x}{x}, \dots \quad (356)$$

So for each energy eigenvalue $E = \frac{\hbar^2 k^2}{2m} \geq 0$ or $\omega = \frac{E}{\hbar}$ and each value of $l = 0, 1, 2, \dots$ we have one eigenstate. Each energy eigenvalue E is infinitely degenerate. The simplest radial eigenfunction for each energy, is the S-wave (0-angular momentum) radial wave function

$$R_{k,l=0}(r, t) = j_0(kr)e^{-i\omega t} = \frac{\sin kr}{kr}e^{-i\omega t} = \frac{1}{2i} \left(\frac{e^{i(kr-\omega t)}}{kr} - \frac{e^{-i(kr+\omega t)}}{kr} \right) \quad (357)$$

This represents a superposition of an outgoing and incoming spherical wave. Of course, to get the full eigenfunction for $l > 0$ we must multiply R_{kl} by the eigenfunction $Y(\theta, \phi)$ of L^2 . The ground state is non-degenerate, corresponding to the constant ψ with $E = k = 0$. For fixed $E = \hbar^2 k^2 / 2m > 0$, the various stationary states of definite angular momentum quantum number $l = 0, 1, 2, \dots$ are all degenerate in energy. For each value of l there is a further $2l + 1$ -fold degeneracy corresponding to the different eigenvalues of L_z .

11.9 Radial eigenfunctions for particle in a spherical box/well

- The spherical potential well corresponds to the central potential $V(r) = 0$ for $r < a$ and $V(r) = \infty$ for $r > a$. The radial wave function $u(r) = rR(r) = 0$ for $r \geq a$ and for $r \leq a$,

$$-u''(r) + \frac{l(l+1)}{r^2}u(r) = k^2u(r), \quad u(a) = 0, \quad E = \frac{\hbar^2 k^2}{2m} \quad (358)$$

For S-waves $l = 0$, $R(r) = A\frac{\sin(rk)}{r} + B\frac{\cos(kr)}{r}$. As for the free particle, $B = 0$ for the particle to be free at $r = 0$. The b.c. at $r = a$, $\sin(ka) = 0$ implies $ka = \pi, 2\pi, 3\pi, \dots$. Thus the

⁴⁹G. N. Watson wrote a treatise on Bessel functions (1922). The Bessel equation may be solved by substituting a power series and finding the coefficients, but there are other methods. In terms of $u(r) = rR(r)$, and the dimensionless variable $x = kr$, the Bessel equation takes the form $-u''(x) + \frac{l(l+1)}{x^2}u(x) = u(x)$. This equation resembles the SHO eigenvalue problem but with a $1/x^2$ instead of x^2 potential. It can be solved by introducing raising and lowering operators similar to the SHO. The $l = 1, 2, 3, \dots$ spherical Bessel functions $j_l(x)$ are obtained by application of the raising operator to the $l = 0$ 'ground state' $\sin(kr)$. See the discussion in R. Shankar's book based on the method of Infeld.

S -wave eigenfunctions are $\psi(r) = R(r)Y(\theta, \phi) = \sqrt{\frac{2}{a}} \frac{1}{\sqrt{4\pi}} \frac{\sin(n\pi r/a)}{r}$ for $n = 1, 2, 3, \dots$ with energies $E_{n,l=0} = \frac{n^2\pi^2\hbar^2}{2ma^2}$.

- For possibly non-zero angular momentum $L^2 = \hbar^2 l(l+1)$, the free particle radial wavefunctions (spherical Bessel functions) $\frac{j_l(kr)}{r}$ must now satisfy the b.c. $j_l(ka) = 0$. So the energy levels are determined by the zeros of the spherical Bessel functions. $j_l(kr)$ is oscillatory and has an infinite number of zeros for $r > 0$, just like the $\sin(kr)$ function but the zeros are not easily found⁵⁰. If z_{nl} is the n^{th} zero of j_l , then the energy levels are $E_{nl} = \frac{\hbar^2 z_{nl}^2}{2ma^2}$. Unlike scattering states of a free particle, these are bound states. Their energies depend on the angular momentum (azimuthal) quantum number l . The degeneracy of each energy level is equal to the degeneracy $(2l+1)$ of each eigenvalue of L^2 , corresponding to the $2l+1$ different eigenvalues $\hbar m$ of L_z , labeled by $m = -l, -l+1, \dots, 0, \dots, l-1, l$.

11.10 Radial eigenfunctions for Hydrogen atom

- See Schiff, Liboff, Griffiths. See Pauli for a connection to 2d SHO in polar coordinates.
- In the H-atom, the potential felt by the electron due to the electrostatic attraction to the proton is $V(r) = -\frac{e^2}{4\pi\epsilon_0 r}$. The resulting radial eigenvalue problem is

$$-\frac{\hbar^2}{2m} u''(r) + \left[-\frac{e^2}{4\pi\epsilon_0 r} + \frac{\hbar^2 l(l+1)}{2mr^2} \right] u = Eu \quad (359)$$

m is the reduced mass $\frac{1}{m} = \frac{1}{m_e} + \frac{1}{m_p}$ that appears in transforming the two-body problem to center of mass and relative coordinates. The effective potential

$$V_{eff}(r) = -\frac{e^2}{4\pi\epsilon_0 r} + \frac{\hbar^2 l(l+1)}{2mr^2} \quad (360)$$

shows that (for $l \neq 0$), there is a centrifugal barrier that would classically keep the electron away from the proton (in the absence of radiation). But what if the angular momentum is zero? Will the electron fall into the nucleus. Quantum mechanically, the question is whether the atom has a ground state, a stationary state of lowest energy $E_1 > -\infty$. We look for bound states $E = -\frac{\hbar^2 \kappa^2}{2m} < 0$ where $\kappa > 0$:

$$u''(r) + \frac{2me^2}{4\pi\epsilon_0 \hbar^2} \frac{u}{r} - \frac{l(l+1)}{r^2} u = \kappa^2 u \quad (361)$$

We will solve this radial eigenvalue problem by the analytic method followed for the harmonic oscillator. First pass to dimensionless variables, find the asymptotic behavior of $u(r)$ for $r \rightarrow 0, \infty$, then use the Frobenius method to find a series solution. The requirement that the series represent a normalizable wave function will determine the spectrum of Hydrogen.

- In terms of the dimensionless variable $\rho = \kappa r$ and dimensionless measure of the size of the atom (or reciprocal of κ) $\lambda = \frac{me^2}{2\pi\epsilon_0 \hbar^2 \kappa}$ (which plays the role of energy eigenvalue),

$$u''(\rho) = \left(1 - \frac{\lambda}{\rho} + \frac{l(l+1)}{\rho^2} \right) u. \quad (362)$$

⁵⁰For large x $j_l(x) \sim \frac{1}{x} \cos[x - \frac{(l+1)\pi}{2}]$. So for large x , the zeros approach equal separation. For large n , $z_{nl} \approx \frac{\pi}{2}(2n+l+2)$.

• Let us first examine the simplest and most important case S-wave states of zero angular momentum $l = 0$. We expect the g.s. of the atom to have $l = 0$ since higher angular momentum should imply a more energetic atom. The S-wave radial eigenvalue problem is (λ plays the role of eigenvalue)

$$u'' = \left(1 - \frac{\lambda}{\rho}\right) u \quad (363)$$

For large ρ , this becomes $u'' \approx u$ whose normalizable solution is $u(\rho) = e^{-\rho}$. So we expect the radial wave function to decay exponentially.

• For small ρ we make the ansatz $u(\rho) \sim \rho^a$ and find $a(a-1)\rho^{a-2} = \rho^a - \lambda\rho^{a-1} \Rightarrow a(a-1) = 0$, by requiring the coefficient of the dominant term as $\rho \rightarrow 0$ to vanish. But if $u \rightarrow \rho^0$, then $R(r) \rightarrow \frac{1}{r}$ and $\nabla^2\psi \sim \delta^3(r)$ near the origin, which would not be a solution of the Schrodinger equation for a Coulomb potential. So we omit $a = 0$ and conclude that $u(\rho) \sim \rho$ as $\rho \rightarrow 0$. Now we know the asymptotic behavior for both small and large ρ and would expect $u = \rho e^{-\rho}v(\rho)$ to be a good change of variable. It turns out that $u = \rho e^{-\rho}$ is in fact an exact normalizable solution of (363) with $\lambda = 2$ corresponding to the energy eigenvalue

$$\lambda = 2 = \frac{me^2}{2\pi\epsilon_0\hbar^2\kappa} \Rightarrow E_1 = -\frac{\hbar^2\kappa_1^2}{2m} = -\frac{me^4}{2\hbar^2} \frac{1}{(4\pi\epsilon_0)^2} = -1 \text{ Rydberg} \quad (364)$$

We recognize this as Bohr's formula for the ground state energy of hydrogen. We will see that it is in fact the ground state of the hydrogen atom. Remarkably, the atom is stable, the energy cannot be arbitrarily negative even for zero angular momentum! The unnormalized radial wavefunction is

$$R(r) = \frac{1}{r}u(r) \propto \kappa_1 e^{-\kappa_1 r}, \quad \kappa_1 = \frac{me^2}{4\pi\epsilon_0\hbar^2} = a_0^{-1} = (0.53\text{Angstrom})^{-1} \quad (365)$$

So κ_1 is the reciprocal of the Bohr radius a_0 . Combining with the S-wave angular wave function $Y_{00}(\theta, \phi) = \frac{1}{\sqrt{4\pi}}$ we get the normalized g.s. wave function of Hydrogen

$$\psi(r, \theta, \phi) = \frac{1}{\sqrt{\pi a_0^3}} e^{-r/a_0} \quad (366)$$

• **Excited S-wave states:** More generally, we must solve the eigenvalue problem (363) to find the allowed S-wave radial functions and energies λ . Motivated by the above solution and what worked for the SHO, we expect the other eigenfunctions to differ from $\rho e^{-\rho}$ by a polynomial in ρ . Put

$$u(\rho) = \rho e^{-\rho}v(\rho) \Rightarrow \rho v'' + 2(1-\rho)v' + (\lambda-2)v = 0 \quad (367)$$

We look for power series solutions $v(\rho) = \sum_{j \geq 0} c_j \rho^j$ and find the recursion relation

$$c_{j+1} = \left[\frac{2(j+1) - \lambda}{(j+1)(j+2)} \right] c_j \quad (368)$$

Given c_0 , this determines $c_{j \geq 1}$. For the wave function to be normalizable the series must terminate (as with the SHO) say at $j = j_{max}$, i.e., λ must equal an even integer $\lambda = 2n = 2(j_{max} + 1)$. $n = 1, 2, 3 \dots$ is called the principal quantum number. For each value of n , $v(\rho)$

is a polynomial of degree $j_{max} = n - 1$ called the $(n - 1)^{th}$ Laguerre polynomial. For each such value of n , we have $\lambda_n = 2n = \frac{2me^2}{4\pi\epsilon_0\hbar^2\kappa_n}$. So the corresponding energy eigenvalues are

$$E_n = -\frac{\hbar^2\kappa_n^2}{2m} = -\frac{me^4}{2\hbar^2(4\pi\epsilon_0)^2} \frac{1}{n^2} = \frac{E_1}{n^2} \quad (369)$$

So the S-wave eigen-energies are as per Bohr's formula! The eigenfunctions are up to a normalization the products of Laguerre polynomials and an exponential factor $v(\rho)e^{-\rho}$. For e.g., the first excited state ($n = 2$) wave function $\propto \left(1 - \frac{r}{2a_0}\right) e^{-r/2a_0}$.

• For non-zero angular momentum, the asymptotic behavior is $u(\rho) \sim Ae^{-\rho}$ for large ρ and $u(\rho) \sim \rho^{l+1}$ for small ρ . A similar procedure is followed $u = \rho^{l+1}e^{-\rho}v(\rho)$ and a recursion relation derived for the coefficients of the power series $v(\rho) = \sum_j c_j \rho^j$

$$c_{j+1} = \left[\frac{2(j+l+1) - \lambda}{(j+1)(j+2l+2)} \right] c_j \quad (370)$$

For the wave function to be normalizable this must terminate at $j = j_{max}$ and λ is equal to twice the principal quantum number $\lambda_n = 2n = 2(j_{max} + l + 1)$. Notice that we can change the angular momentum quantum number $l = 0, 1, 2 \dots n - 1$ and also change the degree j_{max} of the polynomial $v(\rho)$ without changing the principal quantum number n . The energy eigenvalue only depends on κ or λ or n so the energy levels are again given by (the experimentally verified) Bohr formula:

$$E_n = \frac{E_1}{n^2}, n = 1, 2, 3 \dots, \quad l = 0, 1, 2, \dots n - 1, \quad m = -l, -l + 1, \dots, 0, \dots, l - 1, l \quad (371)$$

E_n depend only on the principal quantum number n and not on l or m . The degeneracy of energy levels is $\sum_{l=0}^{n-1} (2l + 1) = n^2$. However, for non-zero l , the angular part of the wave function $Y_{lm}(\theta, \phi)$ is non-constant. So for states of non-zero angular momentum, the probability density of electron locations is not spherically symmetric.