Mathematical Physics Course Notes

Isaac Holt

February 7, 2023

Contents

1 The action principle **2**
1.1 Calculus of variatons 2
1.2 Configuration space and generalised coordinates 3
1.3 Lagrangians for classical mechanics 4
1.4 Ignorable coordinates and conservation of generalised momenta 4

2 Symmetries, Noether's theorem and conservation laws **5**
2.1 Ordinary symmetries 5
2.2 Energy conservation 7

3 Normal modes **9**
3.1 Canonical kinetic terms 9

4 Hamiltonian Formalism **11**

5 Quantum mechanics introduction **13**

6 The wave function and probabilities **14**
6.1 Examples of wave functions 15
6.2 Collapse of the wave function 16

7 Momentum and Planck's constant **17**
7.1 The Hilbert space 19
7.2 Hermitian operators 22
7.3 The spectrum of a Hermitian operator 25

1 The action principle

1.1 Calculus of variatons

Definition 1.1.1. A **functional** is a map from a set of functions to $\mathbb{R}$, e.g. $f : (\mathbb{R} \to \mathbb{R}) \to \mathbb{R}$.

Definition 1.1.2. Let $y(t)$ be a function with fixed values at endpoints a and b. y is **stationary** for a functional S if

$$\frac{dS(y(t) + \epsilon z(t))}{d\epsilon}\Big|_{\epsilon=0} = 0$$

for every smooth (continuous derivative to every order) $z(t)$ such that $z(a) = z(b) = 0$.

Remark. Functions $y(t)$ may be referred to as **paths** and so functions that satisfy the above definition are referred to as **stationary paths**.

Definition 1.1.3. Let S be an **action functional** (or just **action**). **The action principle** states that the paths described by particles are stationary paths of S.

Mathematically, given a particle moving in one dimension with position given by $x(t)$, for arbitrary smooth small deformations $\delta x(t)$ around the true path $x(t)$ (the path the particle follows):

$$\delta S := S(x + \delta x) - S(x) = O((\delta x)^2)$$

Lemma 1.1.4. (Fundamental lemma of the calculus of variations) Let $f(x)$ be a continuous function in the interval $[a, b]$ such that

$$\int_a^b f(x)g(x)dx = 0$$

for every smooth function $g(x)$ in $[a, b]$ such that $g(a) = g(b) = 0$. Then $f(x) = 0\ \forall x \in [a, b]$.

Definition 1.1.5. Let $L(r, s)$ be a function of two real variables. If a functional S can be expressed as the time integral of L, i.e. if

$$S(x) = \int_{t_0}^{t_1} L(x(t), \dot{x}(t))dt$$

then L is called a **Lagrangian**.

Definition 1.1.6. For a Lagrangian L, the **Euler-Lagrange equation** is given by

$$\frac{\partial L}{\partial x} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right) = 0$$

where

$$\frac{\partial L}{\partial x} = \frac{\partial L(r, s)}{\partial r}\Big|_{(r,s)=(x(t),\dot{x}(t))} \quad \text{and} \quad \frac{\partial L}{\partial \dot{x}} = \frac{\partial L(r, s)}{\partial s}\Big|_{(r,s)=(x(t),\dot{x}(t))}$$

Remark. The Lagrangian L is just an ordinary function of (here) two independent parameters (r, s)t. When constructing the action, L is evaluated at $(r, s) = (x(t), \dot{x}(t))$, but r and s as parameters in the definition of the Lagrangian are independent.

So $\dot{x}$ is treated as a variable which does not depend on x:

$$\frac{\partial x}{\partial \dot{x}} = \frac{\partial \dot{x}}{\partial x} = 0$$

Remark. The Euler-Lagrange equation only applies to one-dimensional cases.

1.2 Configuration space and generalised coordinates

Definition 1.2.1. Configuration space, denoted C, is the set of all possible (in principle) instantaneous configurations for a given a physical system.

Remark. This definition includes positions, but does not include velocities.

Remark. A configuration space must be constructed before a Lagrangian is constructed. The Lagrangian describes the dynamics of this configuration space.

Example 1.2.2. A particle moving in $\mathbb{R}^d$ has configuration space $\mathbb{R}^d$.

Example 1.2.3. N distinct particles moving in $\mathbb{R}^d$ have configuration space $(\mathbb{R}^d)^N = \mathbb{R}^{dN}$. The configuration space would still be $\mathbb{R}^{dN}$ if the particles were electrically charged, as the charge of the particles does not affect their positions, at least initially.

Example 1.2.4. Two distinct particles joined by a rigid rod have configuration space $\mathbb{R}^{2d-1}$. One particle has configuration space $\mathbb{R}^d$ and there are $d-1$ angles that must specified to choose the position of the second particle relative to the other.

Definition 1.2.5. Let S be a physical system with configuration space C. Then S has $\dim(C)$ **degrees of freedom**.

Remark. For every configuration space, any choice of coordinate system is valid, and the Lagrangian formalism holds regardless of this choice.

Definition 1.2.6. For a configuration space C, a set of coordinates in this space is called a set of **generalised coordinates**. Often generalized coordinates are represented with q_i, $i \in \{1, \ldots, \dim(C)\}$ where $\underline{q}$ is the coordinate vector with components q_i.

Example 1.2.7. A particle moving in $\mathbb{R}^2$, with configuration space $\mathbb{R}^2$. We could use Cartesian or polar coordinates to describe the position of the particle in this space (both are equally valid).

Definition 1.2.8. Let C be a configuration space and let $\underline{q}(t) \in C$ be a path. For a Lagrangian function $L(\underline{q}, \dot{\underline{q}})$, the **Euler-Lagrange equations** state that

$$\frac{\partial L}{\partial q_i} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) = 0 \quad \forall i \in \{1, \ldots, \dim(C)\}$$

Remark. The Euler-Lagrange equations are valid in any coordinate system.

Remark. Similarly to the one-dimensional case:

$$\frac{\partial q_i}{\partial \dot{q}_j} = \frac{\partial \dot{q}_i}{\partial q_j} = 0$$

and

$$\frac{\partial q_i}{\partial q_j} = \frac{\partial \dot{q}_i}{\partial \dot{q}_j} = \delta_{ij}$$

1.3 Lagrangians for classical mechanics

Definition 1.3.1. In a system with kinetic energy $T(\underline{q}, \underline{\dot{q}})$ and potential energy $V(\underline{q})$, the Lagrangian that describes the equations of motion in that system is given by

$$L(\underline{q}, \underline{\dot{q}}) = T(\underline{q}, \underline{\dot{q}}) - V(\underline{q})$$

1.4 Ignorable coordinates and conservation of generalised momenta

Definition 1.4.1. Let $\{q_1, \ldots, q_N\}$ be a set of generalised coordinates. A specific coordinates q_i is **ignorable** if the Lagrangian function expressed in these generalised coordinates does not depend on q_i, i.e. if

$$\frac{\partial L}{\partial q_i} = 0$$

Definition 1.4.2. The **generalised momentum** p_i associated with a generalised coordinate q_i is given by

$$p_i := \frac{\partial L}{\partial \dot{q}_i}$$

Proposition 1.4.3. The generalised momentum associated to an ignorable coordinate is conserved.

Proof. From the Euler-Lagrange equation for q_i,

$$0 = \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) - \frac{\partial L}{\partial q_i} = \frac{dp_i}{dt} - 0 = \frac{dp_i}{dt}$$

□

Example 1.4.4. For a free particle moving in d dimensions, in Cartesian coordinates we have

$$L = T - V = \frac{1}{2}m\sum_{i=1}^{d}\dot{x}_i^2$$

so every coordinate is ignorable. The generalised momenta are

$$p_i = \frac{\partial L}{\partial \dot{x}_i} = m\dot{x}_i$$

So here the conservation of generalised momenta is the conservation of the linear momenta.

2 Symmetries, Noether's theorem and conservation laws

2.1 Ordinary symmetries

Definition 2.1.1. For a uniparametric family of smooth maps $\phi(\epsilon) : C \to C$ from configuration space to itself, with $\phi(0)$ the identity map, this family of maps is called a **transformation depending on** ϵ. In any coordinates system this transformation can be written as

$$q_i \to \phi_i(q_1, \ldots, q_N, \epsilon)$$

where the ϕ_i's are a set of $N := \dim(C)$ functions representing the transformation in the coordinate system. The change in velocities is defined as

$$\dot{q}_i \to \frac{d}{dt}\phi_i$$

Remark. q_i' is used to denote $\phi(q_i, \epsilon)$, so often we write $q_i \to q_i' = \ldots$, where $\ldots$ is a function of q_i and ϵ.

Definition 2.1.2. The **generator** of ϕ is

$$\frac{d\phi(\epsilon)}{d\epsilon}\Big|_{\epsilon=0} := \lim_{\epsilon\to 0} \frac{\phi(\epsilon) - \phi(0)}{\epsilon}$$

In any coordinate system,

$$q_i \to \phi_i(\underline{q}, \epsilon) = q_i + \epsilon a_i(\underline{q}) + O(\epsilon^2)$$

where

$$a_i = \frac{\partial \phi_i(\underline{q}, \epsilon)}{\partial \epsilon}\Big|_{\epsilon=0}$$

is a function of the generalised coordinates. Hence the transformation generator is a_i. For the velocities the transformation is

$$\dot{q}_i \to \dot{q}_i + \epsilon a_i(q_1, \ldots, q_N, \dot{q}_1, \ldots, \dot{q}_N) + O(\epsilon^2)$$

where the generator is $\dot{a}_i$.

Example 2.1.3. For a particle moving in two dimensions, the finite transformations given by rotations around the origin, in Cartesian coordinates, are

$$x \to x\cos(\epsilon) - y\sin(\epsilon)$$
$$y \to x\sin(\epsilon) + y\cos(\epsilon)$$

The associated infinitesimal transformations can be derived using the expansions $\sin(\epsilon) = \epsilon + O(\epsilon^3)$ and $\cos(\epsilon) = 1 + O(\epsilon^2)$. Then

$$x \to x - y\epsilon + O(\epsilon^2)$$
$$y \to y + x\epsilon + O(\epsilon^2)$$

Then the generators of the transformation are

$$a_x = -y, \quad a_y = x, \quad \dot{a}_x = -\dot{y}, \quad \dot{a}_y = \dot{x}$$

Lemma 2.1.4. Equations of motion do not change if the Lagrangian is modified by adding the total derivative of a function of coordinates and time, i.e.

$$L \to L + \frac{dF(q_1, \ldots, q_N, t)}{dt}$$

results in the same equations of motion.

Proof. The effect on the action is

$$S = \int_{t_0}^{t_1} L dt \to S' = S + F(q_1(t_1), \ldots, q_N(t_1), t_1) - F(q_1(t_0), \ldots, q_N(t_0), t_0)$$

From the action principle, we must have that δS vanishes to first order in $\delta q_i(t)$, with the q_i's fixed at the path's endpoints. Therefore $F(q_1(t_1), \ldots, q_N(t_1), t_1)$ and $F(q_1(t_0), \ldots, q_N(t_0), t_0)$ are fixed. Hence

$$\begin{aligned}
\delta S' &= S'(\underline{q} + \delta\underline{q}) - S'(\underline{q}) \\
&= (S(\underline{q} + \delta\underline{q}) + F(q_1(t_1), \ldots, q_N(t_1), t_1) - F(q_1(t_0), \ldots, q_N(t_0), t_0)) \\
&\quad - (S(\underline{q}) + F(q_1(t_1), \ldots, q_N(t_1), t_1) - F(q_1(t_0), \ldots, q_N(t_0), t_0)) \\
&= S(\underline{q} + \delta\underline{q}) - S(\underline{q}) = \delta S
\end{aligned}$$

So the variation of the action is not affected, therefore the equations of motion cannot be affected. □

Definition 2.1.5. A transformation $\phi(\epsilon)$ is a **symmetry** if, to first order in ϵ, for some function $F(\underline{q}, t)$, the change in the Lagrangian is a total time derivative of $F(\underline{q}, t)$, i.e.

$$L \to L + \epsilon \frac{dF(\underline{q}, t)}{dt} + O(\epsilon^2)$$

Remark. $F(\underline{q}, t)$ is only defined up to a constant. If some $F(\underline{q}, t)$ satisfies the above equation, then $G(\underline{q}, t) = F(\underline{q}, t) + c$ will also satisfy the equation.

Example 2.1.6. If q_i is an ignorable coordinate, the transformation $q_i \to q_i + c_i$, where c_i is a constant, is a symmetry, because q_i does not appear in the Lagrangian by definition so here $F(\underline{q}, t) = 0$.

Theorem 2.1.7. (Noether's theorem) In a given set of generalised coordinates, let $a_i(\underline{q})$ be the generator of a transformation such that

$$L \to L + \epsilon \frac{dF(\underline{q}, t)}{dt} + O(\epsilon^2)$$

so that it is a symmetry. Let

$$Q := \left(\sum_{i=1}^{N} a_i \frac{\partial L}{\partial \dot{q}_i} \right) - F$$

Then $\frac{dQ}{dt} = 0$, so Q is conserved. Q is called the **Noether charge**.

Proof. Consider the variation of the action under the transformation $q_i \to q_i + \epsilon a_i$. Using the chain rule and then the Euler-Lagrange equations,

$$\begin{aligned}
\delta S &= \int_{t_0}^{t_1} \sum i = 1^N \left(\epsilon a_i \frac{\partial L}{\partial q_i} + \epsilon \dot{a}_i \frac{\partial L}{\partial \dot{q}_i} \right) dt + O(\epsilon^2) \\
&= \int_{t_0}^{t_1} \sum_{i=1}^{N} \left(\epsilon a_i \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \right) + \epsilon \dot{a}_i \frac{\partial L}{\partial \dot{q}_i} \right) + O(\epsilon^2) \\
&= \int_{t_0}^{t_1} \epsilon \frac{d}{dt} \left(\sum_{i=1}^{N} a_i \frac{\partial L}{\partial \dot{q}_i} \right) + O(\epsilon^2) \\
&= \epsilon \left[\sum_{i=1}^{N} a_i \frac{\partial L}{\partial \dot{q}_i} \right]_{t_0}^{t_1} + O(\epsilon^2)
\end{aligned}$$

Note that having used the Euler-Lagrange equations in the second line, the result is only valid along the path satisfying the equations of motion.

Now since the transformation is a symmetry,

$$\begin{aligned}
\delta S &= S(\underline{q} + \delta \underline{q}) - S(\underline{q}) \\
&= \int_{t_0}^{t_1} \left(\left(L + \epsilon \frac{dF}{dt} + O(\epsilon^2) \right) - L \right) \\
&= \epsilon [F]_{t_0}^{t_1} + O(\epsilon^2)
\end{aligned}$$

Equating these two expressions for δS, we see that $Q(t_1) = Q(t_0)$. We didn't specify what t_0 and t_1 were so this equality holds for every t_0 and t_1. So let $t_1 = t_0 + \epsilon$, then

$$Q(t_1) - Q(t_0) = Q(t_0 + \epsilon) - Q(t_0) = \epsilon \frac{dQ}{dt} + O(\epsilon^2) = 0$$

hence $\frac{dQ}{dt} = 0$. □

Example 2.1.8. If a coordinate q_i is ignorable, there is a symmetry generated by $q_i \to q_i + \epsilon$ and leaving the other coordinates constant. So

$$a_k = \delta_{ik} := \begin{cases} 1 & \text{if } i = k \\ 0 & \text{otherwise} \end{cases}$$

The Noether charge is

$$Q = \sum_{k=1}^{N} a_k \frac{\partial L}{\partial \dot{q}_k} = \sum_{k=1}^{N} \delta_{ik} \frac{\partial L}{\partial \dot{q}_k} = \frac{\partial L}{\partial \dot{q}_i}$$

which agrees with Proposition 1.4.3.

2.2 Energy conservation

Definition 2.2.1. Given a Lagrangian that explicitly depends on time, $L(\underline{q}, \underline{\dot{q}}, t)$, the **energy**, E, is defined as

$$E := \left(\sum_{i=1}^{N} \dot{q}_i \frac{\partial L}{\partial \dot{q}_i} \right) - L$$

Theorem 2.2.2. Along a path $\underline{q}(t)$ which satisfies the equations of motion,

$$\frac{dE}{dt} = -\frac{\partial L}{\partial t}$$

Proof.

$$\begin{aligned}\frac{dE}{dt} &= \frac{d}{dt}\left(\left(\sum_{i=1}^{N} \dot{q}_i \frac{\partial L}{\partial \dot{q}_i}\right) - L\right) \\ &= \sum_{i=1}^{N}\left(\ddot{q}_i \frac{\partial L}{\partial \dot{q}_i} + \dot{q}_i \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right)\right) - \frac{dL}{dt} \\ &= \sum_{i=1}^{N}\left(\frac{\partial L}{\partial \dot{q}_i}\ddot{q}_i + \dot{q}_i \frac{\partial L}{\partial q_i}\right) - \frac{dL}{dt}\end{aligned}$$

using the Euler-Lagrange equations. By the chain rule,

$$\frac{dL}{dt} = \sum_{i=1}^{N}\left(\frac{\partial L}{\partial \dot{q}_i} + \frac{\partial L}{\partial q_i}\dot{q}_i\right) + \frac{\partial L}{\partial t}$$

and so substituting this into the expression for $\frac{dE}{dt}$, we get

$$\frac{dE}{dt} = -\frac{\partial L}{\partial t}$$

□

Remark. Note that $\frac{\partial L}{\partial t}$ means the partial derivative of L with respect to t, **keeping $\underline{q}$ and $\underline{\dot{q}}$ fixed**. This is because the Lagrangian is (here) just an ordinary function of three parameters which are unrelated, so t is independent of $\underline{q}$ and $\underline{\dot{q}}$. The parameters only become related when the Lagrangian is used to build the action. (See this remark).

Corollary 2.2.3. Energy is conserved iff the Lagrangian does not depend explicitly on time.

Proof. Energy is conserved iff $\frac{dE}{dt} = 0$, and $\frac{\partial L}{\partial t} = 0$ iff the Lagrangian does not depend explicitly on time. □

3 Normal modes

3.1 Canonical kinetic terms

Definition 3.1.1. Given a Lagrangian $L = T - V$, T is a **canonical** kinetic term if it is of the form

$$T = \frac{1}{2}\sum_{i=1}^{n} \dot{q}_i^2$$

Given a Lagrangian with a canonical kinetic term, assume $\underline{q} = \underline{0}$ is a stationary point of $V(\underline{q})$, so

$$\frac{\partial V}{\partial q_i}\Big|_{\underline{q}=0} = 0 \quad \forall i \in \{1, \ldots, n\}$$

If $\underline{q} = \underline{0}$ is not a stationary point but $\underline{q} = \underline{a}$ for some $\underline{a}$ is, then the new variables defined as $q_i' = q_i - a_i$ are such that the stationary point is at $\underline{q}' = \underline{0}$. This preserves the form of the Lagrangian, so we can assume $\underline{q} = 0$ is a stationary point.

To determine the motion around this extremum, expand $V(\underline{q})$ to second order in $\underline{q}$ to define an approximate Lagrangian:

$$L_{\text{approx}} = \frac{1}{2}\sum_{i=1}^{n} \dot{q}_i^2 - \frac{1}{2}\sum_{i,j} A_{i,j} q_i q_j$$

where

$$A_{i,j} = \frac{\partial^2 V}{\partial q_i \partial q_j}\Big|_{\underline{q}=0}$$

The Euler-Lagrange equations for L_{approx} are given in matrix notation by

$$\ddot{\underline{q}} + A\underline{q} = D_A\underline{q} = \left(\frac{d^2}{dt^2} + A\right)\underline{q} = 0$$

where D_A is defined as $D_A := \frac{d^2}{dt^2} + A$, which is a linear operator ($D_A(\underline{a} + \underline{b}) = D_A\underline{a} + D_A\underline{b}$ and $D_A(c\underline{a}) = cD_A\underline{a}$ for any vectors a and b and any $c \in \mathbb{R}$).

A is an $n \times n$ matrix which is real and symmetric, due to symmetry of second partial derivatives (we assume V has continuous second partial derivatives). So A has n real eigenvaules and eigenvectors. Let the eigenvalues be $\lambda^{(i)}$ and the corresponding eigenvectors be $\underline{v}^{(i)}$, then

$$A\underline{v}^{(i)} = \lambda^{(i)}\underline{v}^{(i)}$$

Define the ansatz (an assumed solution)

$$\underline{q}^{(i)}(t) = f^{(i)}(t)\underline{v}^{(i)}$$

for some function $f^{(i)}(t)$ to be determined. Then

$$\left(\frac{d^2}{dt^2} + A\right)\underline{q}^{(i)}(t) = \left(\frac{d^2}{dt^2} + A\right)f^{(i)}(t)\underline{v}^{(i)} \quad = \underline{v}^{(i)}\left(\frac{d^2}{dt^2} + \lambda^{(i)}\right)f^{(i)}(t) = 0$$

and so since $\underline{v}^{(i)} \neq 0$,

$$\left(\frac{d^2}{dt^2} + \lambda^{(i)}\right)f^{(i)}(t) = 0$$

The solution to this equation is

$$f^{(i)}(t) = \begin{cases} \alpha^{(i)} \cos(\sqrt{\lambda^{(i)}}t) + \beta^{(i)} \sin(\sqrt{\lambda^{(i)}}t) & \text{if } \lambda^{(i)} > 0 \\ C^{(i)}t + D^{(i)} & \text{if } \lambda^{(i)} = 0 \\ \alpha^{(i)} \cosh(\sqrt{-\lambda^{(i)}}t) + \beta^{(i)} \sinh(\sqrt{-\lambda^{(i)}}t) & \text{if } \lambda^{(i)} < 0 \end{cases}$$

where $\alpha^{(i)}, \beta^{(i)}, C^{(i)}, D^{(i)}$ are constants determined by initial conditions.

The behaviour this solution describes depends on the sign of the $\lambda^{(i)}$. If every $\lambda^{(i)}$ is positive, there is a local minimum and there are oscillations around this minimum. If there is a negative eigenvalue, there is exponential behaviour away from the stationary point, which matches with the intuitioin that small perturbations at a maximum will quickly grow. Zero eignvaleues correspond with motion with constant velocity, with no oscillations.

The general solution is

$$\underline{q}(t) = \sum_{i=1}^{N} \underline{v}^{(i)} f^{(i)}(t)$$

Definition 3.1.2. Given an eigenvalue $\lambda^{(i)} > 0$, a **normal mode** is a solution to

$$\underline{q}(t) = \underline{v}^{(i)} \left(\alpha^{(i)} \cos(\sqrt{\lambda^{(i)}}t) + \beta^{(i)} \sin(\sqrt{\lambda^{(i)}}t) \right)$$

Definition 3.1.3. Given an eigenvalue $\lambda^{(i)} = 0$, a **zero mode** is a solution to

$$\underline{q}(t) = \underline{v}^{(i)} \left(C^{(i)}t + D^{(i)} \right)$$

Definition 3.1.4. Given an eigenvalue $\lambda^{(i)} < 0$, an **instability** is a solution to

$$\underline{q}(t) = \underline{v}^{(i)} \left(\alpha^{(i)} \cosh(\sqrt{-\lambda^{(i)}}t) + \beta^{(i)} \sinh(\sqrt{-\lambda^{(i)}}t) \right)$$

Definition 3.1.5. When there are no instabilities, the general solution is the superposition of normal modes for non-zero eigenvalues and zero modes:

$$\underline{q}(t) = \sum_{i=1,\ \lambda^{(i)} \neq 0}^{n} \underline{v}^{(i)} \left(\alpha^{(i)} \cos(\sqrt{\lambda^{(i)}}t) + \beta^{(i)} \sin(\sqrt{\lambda^{(i)}}t) \right) + \sum_{i=1,\ \lambda^{(i)}=0}^{n} \underline{v}^{(i)} \left(C^{(i)}t + D^{(i)} \right)$$

4 Hamiltonian Formalism

Definition 4.0.1. The classical **state** of a system at a given instant in time is a **complete** set of data that fully specifies the future evolution of the system.

Remark. Any set of data that fully fixes future evolution is valid.

Definition 4.0.2. The **phase (or state) space** is the set of all possible states for a system at a given time.

Example 4.0.3. A free particle moving in $\mathbb{R}$. The phase space is $\mathbb{R}^2$ ($\mathbb{R}$ for position, $\mathbb{R}$ for velocity).

Definition 4.0.4. The **Hamiltonian formalism** studies dynamics in a phase space, parameterised by $\underline{q}(t)$ and $\underline{p}(t)$, where $p_i = \frac{\partial L}{\partial \dot{\underline{q}}_i}$, the momentum.

Example 4.0.5. A particle moving in $\mathbb{R}$, with $L(x, \dot{x}) = \frac{1}{2}m\dot{x}^2$.
Then $p_x = \frac{\partial L}{\partial \dot{x}} = m\dot{x}$ so $\dot{x}(x, p_x) = \frac{p_x}{m}$.
In the Hamltonian formalism, $L(x, p_x) = \frac{p_x^2}{2m}$.

Example 4.0.6. A particle moving in $\mathbb{R}^2$ (in polar coordinates).
$L(r, \theta, \dot{r}, \dot{\theta}) = \frac{1}{2}m(\dot{r}^2 + r^2\dot{\theta}^2)$. So $p_r = m\dot{r}$ and $p_\theta = mr^2\dot{\theta}$.
So $\dot{r}(r, \theta, p_r, p_\theta) = \frac{p_r}{m}$, $\dot{\theta}(r, \theta, p_r, p_\theta) = \frac{p_\theta}{mr^2}$.
$L(r, \theta, \dot{r}, \dot{\theta}) = L(r, \theta, p_r, p_\theta) = \frac{1}{2}(\frac{p_r^2}{m} + \frac{p_\theta^2}{mr^2})$.

Definition 4.0.7. Given two functions $f(\underline{q}, \underline{p}, t)$ and $g(\underline{q}, \underline{p}, t)$ in phase space their **Poisson bracket** is:

$$\{f, g\} := \sum_{i=1}^{n} \left(\frac{\partial f}{\partial q_i}\frac{\partial g}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial g}{\partial q_i} \right)$$

where n is the dimension of the configuration space.

Remark. In the Hamiltonian formalism, $\frac{\partial q_i}{\partial p_j} = \frac{\partial p_j}{\partial q_i} = 0$.
Similarly, $\frac{\partial q_i}{\partial q_j} = \frac{\partial p_i}{\partial p_j} = \delta_{i,j}$

Example 4.0.8. Let $f = q_i$, $g = q_j$. $\{q_i, q_j\} = 0$, and $\{p_i, p_j\} = 0$. $\{q_i, p_j\} = \sum_{k=1}^{n} \delta_{i,j}\delta_{j,k} = \delta_{i,j}$.

Definition 4.0.9. Let $\mathbb{F}$ be the set functions from a phase space P to $\mathbb{R}$

Definition 4.0.10. The Hamiltonian flow $\Phi_f^{(s)}$, with $(s) \in \mathbb{R}$, $f \in F$ operator maps $\mathbb{F}$ to $\mathbb{F}$ and is defined as

$$\Phi_f^{(s)}(g) := e^{s\{\cdot, f\}}g := g + s\{g, f\} + \frac{s^2}{2}\{\{g, f\}, f\} + \cdots$$

Remark. The transformation generated by f has generator $a_i = \{q_i, f\}$ where $q_i \to q_i + \epsilon a_i$.
Infinitesimally, $\Phi_f^{(s)}(g) := g + \epsilon\{g, f\} + O(\epsilon^2)$

TODO: properties on poisson bracket

Example 4.0.11. (Rotation in $\mathbb{R}^2$ in Cartesian coordinates) As a guess, choose $f = q_1\dot{q_2} - \dot{q_1}q_2$, the angular momentum.

$L = \frac{1}{2}(\dot{q_1}^2 + \dot{q_2}^2) - V(q_1, q_2)$ so $p_1 = \frac{\partial L}{\partial \dot{q_1}} = \dot{q_1}$ and $p_2 = \frac{\partial L}{\partial \dot{q_2}} = \dot{q_2} \Rightarrow f = q_1p_2 - q_2p_1$.

Then $q_1 \to q_1 + \epsilon\{q_1, f\} + O(\epsilon^2) = q_1 + \epsilon\{q_1, q_1p_2 - q_2p_1\} = q_1 + \epsilon\{q_1, q_1p_2\} - \epsilon\{q_1, q_2p_1\} = q_1 + \epsilon\{q_1, q_1\}p_2 + \epsilon\{q_1, p_2\}q_1 - \epsilon\{q_1, q_2\}p_1 - \epsilon\{q_1, p_1\}q_2 = q_1 - \epsilon q_2$

Similarly, $q_2 \to q_2 + \epsilon q_1$ so $(q_1, q_2) \to (q_1, q_2) + \epsilon((0, -1), (1, 0))(q_1, q_2)$ TODO make into matrices and column vectors.

Definition 4.0.12. The **Hamiltonian** is the energy expressed in Hamiltonian coordinates:

$$H = \sum_{i=1}^{n} q_i(\dot{\underline{q}}, \underline{p})p_i - L(\underline{q}, \dot{\underline{q}}(\underline{q}, \underline{p}))$$

Example 4.0.13. (Harmonic oscillator in one dimension) Let $\frac{1}{2}m\dot{x}^2 - \frac{1}{2}kx^2 \Rightarrow p = m\dot{x} \Rightarrow \dot{x} = \frac{p}{m}$.

$H = \dot{x}p - L = \frac{p^2}{m} - (\frac{1}{2}\frac{p^2}{m} - \frac{1}{2}kx^2) = \frac{1}{2}\frac{p^2}{m} + \frac{1}{2}kx^2$

Theorem 4.0.14. The time evolution of the phase space coordinates $\underline{q}, \underline{p}$ is generated by Hamiltonian flow Φ_H:

$$q_i(t + a) = \Phi_H^{(a)} q_i(t), p_i(t + a) = \Phi_H^{(a)} p_i(t)$$

Infinitesimally, $q_i(t) + \epsilon\dot{q}_i(t) + O(\epsilon^2) = q_i(t + \epsilon) = q_i(t) + \epsilon\{q_i, H\} + O(\epsilon^2) \Leftrightarrow \dot{q}_i = \{q_i, H\} = \frac{\partial H}{\partial p_i}$ and similarly, $\dot{p}_i = \{p_i, H\} = -\frac{\partial H}{\partial q_i}$.

These equations are called **Hamilton's equations**.

Proof. $\frac{\partial H}{\partial q_i}$. TODO: complete this proof, finish rest of notes from lecture. □

Corollary 4.0.15. The time evolution of any function $f(\underline{q}, \underline{p})$ in phase space is generated by Φ_H:

$$\frac{df}{dt} = \{f, H\}$$

If $f(\underline{q}, \underline{p}, t)$ depends explicitly on time then

$$\frac{df}{dt} = \{f, h\} + \frac{\partial f}{\partial t}$$

Proof. $\frac{df}{dt} = \sum_{i=1}^{n} \left(\frac{\partial f}{\partial q_i}\dot{q}_i + \frac{\partial f}{\partial p_i}\dot{p}_i\right) + \frac{\partial f}{\partial t} = \sum_{i=1}^{n} \left(\frac{\partial f}{\partial q_i}\frac{\partial H}{\partial p_i} - \frac{\partial f}{\partial p_i}\frac{\partial H}{\partial q_i}\right) + \frac{\partial f}{\partial t} = \{f, H\} + \frac{\partial f}{\partial t}$. □

5 Quantum mechanics introduction

Definition 5.0.1. The **photo-electric effect** is the phenomena where, when a light source is shone on a metal, electrons are released from the surface of the metal.

Definition 5.0.2. The **classical** description of light would lead to these results:

- The energy of the released electrons depends on the intensity of light I but is independent of the angular frequency of the light, ω.
- Electrons are released for any value of I

Definition 5.0.3. The quantum mechanics prediction of the photoelectric effect is that the energy of the electrons E is independent of I, and not released at all below some value of angular frequency.

This led to the discovery that light consists of packets (photons).

6 The wave function and probabilities

Definition 6.0.1. The **wave function** is a function of position x and time t:

$$\psi : \mathbb{R}^2 \to \mathbb{C}$$

Remark. We require that ψ is continuous, and that $\frac{d\psi(x)}{dx}$ is continuous (ψ must be differentiable), except when the potential energy $V(x)$ is not finite.

Definition 6.0.2. The **probability density** of finding a particle at position x and time t is defined as

$$P(x,t) := |\psi(x,t)|^2$$

The probability of finding the particle at $x \in [a,b]$ is therefore

$$\int_a^b |\psi(x,t)|^2 dx$$

The probability of finding the particle anywhere is must be equal to 1, which means that

$$\int_{-\infty}^{\infty} P(x,t)dx = 1 \quad \forall t \in \mathbb{R}$$

This puts a constraint on ψ.

Definition 6.0.3. A wave function ψ is **square-normalisable** if the integral

$$\int_{-\infty}^{\infty} P(x,t)dx$$

exists.

Definition 6.0.4. A wave function ψ is **normalised** if

$$\int_{-\infty}^{\infty} P(x,t)dx = 1 \quad \forall t \in \mathbb{R}$$

Definition 6.0.5. For a given time t, the **expectation** value of a polynomial function $f(x)$ of position x, is defined as

$$\langle f(x) \rangle := \int_{-\infty}^{-\infty} P(x,t) f(x) dx$$

Remark. The expectation of position, $\langle x \rangle$ is the mean of measurements of the positions of many particles with wave function ψ.

Definition 6.0.6. The **uncertainty** of x is defined as

$$\Delta x = \sqrt{\langle x^2 \rangle - \langle x \rangle^2}$$

6.1 Examples of wave functions

Remark. For now, we will look at wave functions at a fixed time, $\psi(x)$.

Definition 6.1.1. The **Gaussian wave function** is defined as

$$\psi(x) = Ce^{-x^2/4\Delta^2}$$

where $\Delta > 0$ has units of length and C is a normalisation constant.

$$\int_{-\infty}^{\infty} |\psi(x)|^2 = 1$$

so $C = e^{i\theta}(2\pi\Delta^2)^{-1/4}$, with $\theta \in \mathbb{R}$. θ is a free parameter here, so is often set to 0.

So the normalised probability distribution is

$$P(x) = \frac{1}{\sqrt{2\pi\Delta^2}} e^{-x^2/(2\Delta^2)}$$

which is a standard Gaussian probability distribution.

Remark. For the Gaussian wave function:

- $\langle x^{2n+1} \rangle \quad \forall n \in \mathbb{N}_0$ since the integral is an odd function of x over a symmetric interval.
- As Δ is the only quantity with units of length and x has units of length, $\langle x^{2n} \rangle \propto \Delta^{2n} \quad \forall n \in \mathbb{N}_0$.
- $\langle x^2 \rangle = \Delta^2 \implies \Delta x = \Delta$.

Definition 6.1.2. If a particle's position x is confined to $0 < x < L$, particle is said to be in an **infinite potential well**, defined as

$$V(x) = \begin{cases} 0 & \text{if } 0 < x < L \\ \infty & \text{otherwise} \end{cases}$$

To have position $x \leq 0$ or $x \geq L$, the particle would need infinite energy, so the probability of finding it there is 0, therefore the wave function for the particle should vanish in these regions.

Example 6.1.3. A possible wave function for a particle in an infinite potential well is

$$\Psi(x) = \begin{cases} C\sqrt{x(L-x)} & \text{if } 0 < x < L \\ 0 & \text{otherwise} \end{cases}$$

For some constant C. The probability of finding the particle anywhere in $(0, L)$ is 1, so

$$\int_0^L |\psi(x)|^2 dx = |C|^2 \int_0^L x(L-x)dx = |C|^2 \frac{L^3}{6} = 1$$

which implies $C = \sqrt{6/L^3}e^{i\theta}$ for any θ. We choose $\theta = 0$ here.

Remark. Important: multiplying a wave function $\psi(x)$ by $e^{i\theta}$ for any $\theta \in \mathbb{R}$ does not change the probability density function P, which means that measuring position cannot distinguish between $\psi(x)$ and $e^{i\theta}\psi(x)$.

Measuring momentum can distinguish between these two, unless $\theta(x)$ is constant, in which case they cannot be distinguished by any measurement, and they describe the same physical state.

6.2 Collapse of the wave function

Definition 6.2.1. If a particle has a wave function $\psi(x,t)$ for $t < t_0$, then at $t = t_0$, the position x is measured to be $x = x_0$. Then another measurement immediately after the first will be $x = x_0$ with probability 1. This phenomenon is called the **wave function collapse**. Measuring x changes the wave function so that it is very localised around x_0. This is why $\langle x \rangle$ is the average of measurements of many particles with the same wave function, not repeated measurements of the same wave function.

7 Momentum and Planck's constant

Definition 7.0.1. The **momentum operator** $\hat{p}$ is defined as

$$\hat{p}(\psi(x)) := -i\hbar\frac{\partial}{\partial x}(\psi(x))$$

so that

$$\begin{aligned}\delta_\epsilon\psi(x) &:= \psi(x-\epsilon) - \psi(x)\\ &= -\epsilon\frac{\partial}{\partial x}\psi(x) + O(\epsilon^2)\\ &= -\epsilon\frac{i}{\hbar}\hat{p}(\psi(x))\end{aligned}$$

where $\hbar$ is a constant of proportionality.

Definition 7.0.2. In the above definition, $\hbar$ is called the **reduced Planck constant** which has units energy $\times$ time. **Planck's constant** is defined as

$$h = 2\pi\hbar$$

Definition 7.0.3. We define a **position operator** $\hat{x}$ which multiplies a wave function by x:

$$\hat{x}(\psi(x)) = x\psi(x)$$

Definition 7.0.4. The **commutator** of the position and momentum operators $\hat{x}$ and $\hat{p}$ is defined as

$$\begin{aligned}[\hat{x},\hat{p}](\psi(x)) &= \hat{x}(\hat{p}(\psi(x))) - \hat{p}(\hat{x}(\psi(x)))\\ &= \hat{x}\left(-i\hbar\frac{\partial}{\partial x}\psi(x)\right) + i\hbar\frac{\partial}{\partial x}(x\psi(x))\\ &= i\hbar\psi(x)\end{aligned}$$

The relation $[\hat{x},\hat{p}](\psi(x)) = i\hbar\psi(x)$ is called the **canonical commutation relation**.

Remark. This commutator is similar to the Poisson bracket, where $\{x,p\} = 1$ in classical mechanics. So we can transform a classical system into a quantum system by replacing the Poisson bracket $\{,\}$ with $-\frac{i}{\hbar}[,]$. This replacement is called **canonical quantisation**.

Definition 7.0.5. The **expectation value of momentum**, $\langle p\rangle$ is defined as

$$\begin{aligned}\langle p\rangle &= \int_{-\infty}^{\infty}\overline{\psi(x,t)}\hat{p}(\psi(x,t))dx\\ &= -i\hbar\int_{-\infty}^{\infty}\overline{\psi(x,t)}\frac{\partial}{\partial x}\psi(x,t)dx\end{aligned}$$

Remark. Notice the similarity between this definition and the one for the expectation value of position:

$$\begin{aligned}\langle x\rangle &= \int_{-\infty}^{\infty}x|\psi(x,t)|^2dx\\ &= \int_{-\infty}^{\infty}\overline{\psi(x,t)}\hat{x}(\psi(x,t))dx\end{aligned}$$

As with $\langle x\rangle$, $\langle p\rangle$ is interpreted as the average of measurements of the momenta of many particles with the same wave function ψ.

Proposition 7.0.6. $\hbar$ must be a real number (it is not complex).

Proof. The measurement of momentum must be a real number, so $\langle p \rangle \in \mathbb{R}$. Using integration by parts and the fact that $|\psi(x,t)|^2$ vanishes as $x \to \pm\infty$ if ψ is square normalisalbe, the complex conjuate of $\langle p \rangle$ is

$$\begin{aligned}
\overline{\langle p \rangle} &= i\bar{\hbar} \int_{-\infty}^{\infty} \psi(x,t) \frac{\partial}{\partial x} \overline{\psi(x,t)} dx \\
&= -i\bar{\hbar} \int_{-\infty}^{\infty} \overline{\psi(x,t)} \frac{\partial}{\partial x} \psi(x,t) dx + i\hbar [|\psi(x,t)|^2]_{-\infty}^{\infty} \\
&= -i\bar{\hbar} \int_{-\infty}^{\infty} \overline{\psi(x,t)} \frac{\partial}{\partial x} \psi(x,t) dx \\
&= \frac{\bar{\hbar}}{\hbar} \langle p \rangle
\end{aligned}$$

Therefore $\langle p \rangle \in \mathbb{R} \iff \hbar \in \mathbb{R}$, so $\hbar$ must be real. □

Definition 7.0.7. The **expectation value** of any polynomial f of p can be calculated as

$$\begin{aligned}
\langle f(p) \rangle &= \int_{-\infty}^{\infty} \overline{\psi(x,t)} f(\hat{p}) \psi(x,t) dx \\
&= \int_{-\infty}^{\infty} \overline{\psi(x,t)} f\left(-i\hbar \frac{\partial}{\partial x}\right) (\psi(x,t)) dx
\end{aligned}$$

Definition 7.0.8. The **momentum uncertainty**, Δp is defined as

$$\Delta p = \sqrt{\langle p^2 \rangle - \langle p \rangle^2}$$

This quantity describes the spread around $\langle p \rangle$ of measurements of the momenta of many particles with the same wave function ψ.

Example 7.0.9. For the Gaussian wave function, $\langle x \rangle = 0$ and $\langle x^2 \rangle = \Delta^2$ so the uncertainty in position is $\Delta x = \Delta$. The momentum operator applied to ψ gives

$$\begin{aligned}
\hat{p}(\psi(x)) &= \frac{i\hbar}{2\Delta^2} x\psi(x) \\
\hat{p}^2(\psi(x)) &= -\hbar^2 \frac{\delta^2}{\delta x^2} (\psi(x)) \\
&= \frac{\hbar^2}{2\Delta^2} \psi(x) - \frac{\hbar^2}{4\Delta^4} x^2 \psi(x)
\end{aligned}$$

Notice that the momentum operator always gives a polynomial in x multiplied by $\psi(x)$, so we can use the position expectations to calculate the momentum expectations:

$$\begin{aligned}
\langle p \rangle &= \frac{i\hbar}{2\Delta^2} \int_{-\infty}^{\infty} x|\psi(x)|^2 dx = \frac{i\hbar}{2\Delta^2} \langle x \rangle = 0 \\
\langle p^2 \rangle &= \frac{\hbar^2}{2\Delta^2} \int_{-\infty}^{\infty} |\psi(x)|^2 dx - \frac{\hbar^2}{4\Delta^4} \int_{-\infty}^{\infty} x^2 |\psi(x)|^2 dx \\
&= \frac{\hbar^2}{2\Delta^2} - \frac{\hbar^2}{4\Delta^4} \langle x^2 \rangle \\
&= \frac{\hbar^2}{2\Delta^2} - \frac{\hbar^2}{4\Delta^4} \Delta^2 \\
&= \frac{\hbar^2}{4\Delta^2}
\end{aligned}$$

So the uncertainty in momentum is

$$\Delta p = \sqrt{\langle p^2 \rangle - \langle p \rangle^2} = \frac{\hbar}{2\Delta}$$

So the product of the uncertainties of position and momentum is a constant, independent of Δ:

$$\Delta x \Delta p = \frac{\hbar}{2}$$

So as the uncertainty in position or momentum decreases, the other must increase.

Definition 7.0.10. Heisenberg's uncertainty principle states that for every normalised wave function,

$$\Delta x \Delta p \geq \frac{\hbar}{2}$$

Remark. Heisenberg's uncertainty principle shows that there is a limit on the how much the uncertainty in both position and momentum can be reduced.

Definition 7.0.11. A **minimal uncertainty wave function** is a wave function that minimises the quantity $\Delta x \Delta p$, i.e. $\Delta x \Delta p = \frac{\hbar}{2}$.

The Gaussian wave function is an example of this.

7.1 The Hilbert space

Definition 7.1.1. A **Hermitian inner product** on a vector space V is a map $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{C}$ which satisifes the following properties:

1. $\forall v, w \in V, \langle v, w \rangle = \overline{\langle w, v \rangle}$.
2. $\forall v, w_1, w_2 \in V, a_1, a_2 \in \mathbb{C}, \langle v, a_1 w_1 + a_2 w_2 \rangle = a_1 \langle v, w_1 \rangle + a_2 \langle v, w_2 \rangle$.
3. $\forall v, w_1, w_2 \in V, a_1, a_2 \in \mathbb{C}, \langle a_1 w_1 + a_2 w_2, v \rangle = \overline{a_1} \langle w_1, v \rangle + \overline{a_2} \langle w_2, v \rangle$.
4. $\forall v \in V, \langle v, v \rangle \geq 0$ and $\langle v, v \rangle = 0 \iff v = 0$.

Remark. A basis $\{e_j\}$ is an orthonormal basis for an n-dimensional vector space V if

$$\langle e_i, e_j \rangle = \delta_{i,j}$$

This means for every $v \in V$,

$$v = \sum_{i=1}^{n} v_i e_i$$

where the ith component of the vector is $v_i = \langle v, e_i \rangle$. The Hermitian inner product of $v, w \in V$ is therefore

$$\langle v, w \rangle = \sum_{i=1}^{n} \overline{v_i} w_i$$

Definition 7.1.2. The **norm** of a vector $v \in V$ is defined as

$$|v| := \sqrt{\langle v, v \rangle}$$

Remark. With an orthonormal basis, the squared norm of v is simply

$$|v|^2 := \langle v, v \rangle = \sum_{i=1}^{n} |v_i|^2$$

Definition 7.1.3. At a fixed time t, a wave function $\psi : \mathbb{R} \to \mathbb{C}$ is called **square normalisable** if

$$\int_{-\infty}^{\infty} |\psi(x)|^2 dx < \infty$$

which means it has a probabilistic interpretation (as we can multiply it by a constant to make the probability of finding a particle anywhere equal to 1).

Proposition 7.1.4. The set of continuous square-integrable wave functions forms a complex vector space. In particular, for all square-integrable wave functions ψ_1, ψ_2 and for every $a_1, a_2 \in \mathbb{C}$, the wave function

$$a_1\psi_1 + a_2\psi_2$$

is also square-integrable.

Proof. Clearly, if ψ_1 is square-integrable then for every $a_1 \in \mathbb{C}$, $a_1\psi_1$ is square-integrable. So now we must prove $\psi_1 + \psi_2$ is square-integrable. For every $x \in \mathbb{R}$,

$$\begin{aligned}
|\psi_1(x) + \psi_2(x)|^2 &= \langle \psi_1(x) + \psi_2(x), \psi_1(x) + \psi_2(x) \rangle \\
&= \langle \psi_1(x) + \psi_2(x), \psi_1(x) \rangle + \langle \psi_1(x) + \psi_2(x), \psi_2(x) \rangle \\
&= \langle \psi_1(x), \psi_1(x) \rangle + \langle \psi_2(x), \psi_1(x) \rangle + \langle \psi_1(x), \psi_2(x) \rangle + \langle \psi_2(x), \psi_2(x) \rangle \\
&= \langle \psi_1(x), \psi_1(x) \rangle + \overline{\langle \psi_1(x), \psi_2(x) \rangle} + \langle \psi_1(x), \psi_2(x) \rangle + \langle \psi_2(x), \psi_2(x) \rangle \\
&= |\psi_1(x)|^2 + |\psi_2(x)|^2 + 2\mathrm{Re}(\overline{\psi_1(x)}\psi_2(x)) \\
&\leq |\psi_1(x)|^2 + |\psi_2(x)|^2 + 2|\overline{\psi_1(x)}\psi_2(x)| \\
&\leq |\psi_1(x)|^2 + |\psi_2(x)|^2 + 2|\overline{\psi_1(x)}||\psi_2(x)| \\
&= |\psi_1(x)|^2 + |\psi_2(x)|^2 + (|\psi_1(x)|^2 + |\psi_2(x)|^2) - (|\psi_1(x)| + |\psi_2(x)|)^2 \\
&\leq 2|\psi_1(x)|^2 + 2|\psi_2(x)|^2
\end{aligned}$$

Hence,

$$\int_{-\infty}^{\infty} |\psi_1(x) + \psi_2(x)|^2 dx \leq 2\int_{-\infty}^{\infty} |\psi_1(x)|^2 + 2\int_{-\infty}^{\infty} |\psi_2(x)|^2$$

Hence $\psi_1 + \psi_2$ is square-integrable. □

Definition 7.1.5. We define an **inner product** over the set of **wave functions** as

$$\langle \psi_1, \psi_2 \rangle := \int_{-\infty}^{\infty} \overline{\psi_1(x)}\psi_2(x) dx$$

Proposition 7.1.6. The inner product defined above is a Hermitian inner product.

Proof. Properties 1, 2 and 3 are clear from the definition. For property 4,

$$\langle \psi, \psi \rangle := \int_{-\infty}^{\infty} |\psi(x)|^2 dx \geq 0$$

as the integrand $|\psi(x)|^2$ is non-negative. If $\langle \psi, \psi \rangle = 0$, then $|\psi(x)|^2 = 0$ as $|\psi(x)|^2$ is continuous, hence $\psi(x) = 0$. □

Definition 7.1.7. The vector space of wave functions with the Hermitian inner product above form a **Hilbert space**.

Definition 7.1.8. We define an **orthonormal basis for the set of wave functions** to be a set of wave functions $\{\phi_n(x)\}$ such that

$$\langle \phi_m, \phi_n \rangle = \delta_{m,n}$$

and every continuous square-integrable wave function ψ can be expressed uniquely as

$$\psi(x) = \sum_n c_n \phi_n(x)$$

where

$$c_n = \langle \phi_n, \psi \rangle = \int_{-\infty}^{\infty} \overline{\phi_n(x)} \psi(x)$$

This allows us to express the Hermitian inner product in terms of the c_n:

$$\langle \psi_1, \psi_2 \rangle := \int_{-\infty}^{\infty} \overline{\psi_1(x)} \psi_2(x) dx = \sum_n \overline{c_{1,n}} c_{2,n}$$

which gives the squared norm as

$$\langle \psi, \psi \rangle = \int_{-\infty}^{\infty} |\psi(x)|^2 dx = \sum_n |c_n|^2$$

Example 7.1.9. For an infinite potential well in the region $(0, L)$, we can restrict the square-integrable to ones that vanish outside $(0, L)$. For every $n \in \mathbb{N}$, define a basis

$$\phi_n(x) = \sqrt{\frac{2}{L}} \sin\left(\frac{n\pi x}{L}\right)$$

The ϕ_n are orthogonal with respect to the inner product:

$$\begin{aligned}
\langle \phi_m, \phi_n \rangle &= \int_0^L \overline{\phi_m(x)} \phi_n(x) dx \\
&= \frac{2}{L} \sin\left(\frac{m\pi x}{L}\right) \sin\left(\frac{n\pi x}{L}\right) dx \\
&= \frac{1}{L} \int_0^L \left(\cos\left(\frac{(m-n)\pi x}{L}\right) - \cos\left(\frac{(m+n)\pi x}{L}\right) \right) dx \\
&= \delta_{m,n} - \delta_{m,-n} \\
&= \delta_{m,n}
\end{aligned}$$

So any continuous square-integrable wave function in this region has the unique form

$$\phi(x) = \sum_{n=0}^{\infty} c_n \phi_n(x) = \sqrt{\frac{2}{L}} \sum_{n=1}^{\infty} \sin\left(\frac{n\pi x}{L}\right)$$

This is precisely Fourier's theorem, with the Fourier coefficients c_n given by

$$c_n = \langle \phi_n, \psi \rangle = \sqrt{\frac{2}{L}} \int_0^L \sin\left(\frac{n\pi x}{L}\right) \psi(x) dx$$

The norm squared of ψ is

$$\langle \psi, \psi \rangle = \int_0^L |\psi(x)|^2 dx = \sum_{n=1}^{\infty} |c_n|^2$$

which is the statement of Parseval's theorem.

Example 7.1.10. Consider a pyramid wave function in an infinite potential well in $(0, L)$ defined as

$$\psi(x) = \sqrt{\frac{12}{L}} \cdot \begin{cases} \frac{x}{L} & \text{if } 0 \le x \le \frac{L}{2} \\ \frac{L-x}{L} & \text{if } \frac{L}{2} \le x \le L \end{cases}$$

Then the Fourier coefficients are given by

$$\begin{aligned} c_n &= \sqrt{\frac{2}{L}} \int_0^L \sin\left(\frac{n\pi x}{L}\right) \psi(x) dx \\ &= \sqrt{\frac{24}{L^2}} \left(\int_0^{L/2} \frac{x}{L} \sin\left(\frac{n\pi x}{L}\right) dx + \int_{L/2}^L \left(1 - \frac{x}{L}\right) \sin\left(\frac{n\pi x}{L}\right) dx \right) \\ &= \sqrt{\frac{24}{L^2}} (1 - (-1)^n) \int_0^{L/2} \frac{x}{L} \sin\left(\frac{n\pi x}{L}\right) dx \\ &= \sqrt{24}(1 - (-1)^n) \frac{(-1)^{(n+1)/2}}{n^2\pi^2} \\ &= \begin{cases} \frac{\sqrt{96}(-1)^{m+1}}{(2m+1)^2\pi^2} & \text{if } n = 2m+1 \text{ for some } m \in \mathbb{N} \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

We see that

$$\langle \psi, \psi \rangle = \sum_{n=1}^{\infty} |c_n|^2 = \frac{96}{\pi^4} \sum_{m=0}^{\infty} \frac{1}{(2m+1)^4} = 1$$

7.2 Hermitian operators

Definition 7.2.1. Let V be a finite-dimensional complex vector space with Hermitian inner product $\langle \cdot, \cdot \rangle$ and an orthonormal basis $\{e_j\}$. A **linear operator** is a map $A : V \to V$ which satisifes

$$\forall v_1, v_2 \in V,\ \forall a_1, a_2 \in \mathbb{C}, \quad A(a_1 v_1 + a_2 v_2) = a_1 A(v_1) + a_2 A(v_2)$$

Remark. Any linear combination $a_1 A_1 + a_2 A_2$ and composition $A_1 \circ A_2$ of two linear operators A_1 and A_2 is also a linear operator.

Definition 7.2.2. The **matrix elements** of a linear operator A in an orthonormal basis $\{e_j\}$ are defined by

$$A_{i,j} = \langle e_i, A \cdot e_j \rangle$$

Definition 7.2.3. The adjoint $A^\dagger$ of a linear operator A is defined by

$$\forall v_1, v_2 \in V, \quad \langle v_1, A v_2 \rangle = \langle A^\dagger v_1, v_2 \rangle$$

Proposition 7.2.4. $A^\dagger_{i,j} = \overline{A_{j,i}}$, i.e. $A^\dagger$ is the conjugate of the transpose of A.

Proof.

$$A^\dagger_{i,j} = \langle e_i, A^\dagger e_j = \langle A e_i, e_j \rangle = \overline{\langle e_j, A e_i} = \overline{A_{i,j}}$$

□

Proposition 7.2.5. The adjoint operation satisfies

- $(a_1 A_1 + a_2 A_2)^\dagger = \overline{a_1} A_1^\dagger + \overline{a_2} A_2^\dagger$.

- $(A_1 A_2)^\dagger = A_2^\dagger A_1^\dagger$.

Proof. TODO. □

Corollary 7.2.6. For every $n \in \mathbb{N}$ and for every polynomial function f,

$$(A^n)^\dagger = (A^\dagger)^n, \quad f(A)^\dagger = f(A^\dagger)$$

Definition 7.2.7. A **Hermitian operator** A is a linear operator that satisfies

$$A = A^\dagger$$

Equivalently,

$$\forall v_1, v_2 \in V, \quad \langle v_1, Av_2 \rangle = \langle Av_1, v_2 \rangle \quad \text{or} \quad \overline{A_{j,i}} = A_{i,j}$$

hence the matrix of a Hermitian operator is a Hermitian matrix.

Definition 7.2.8. The matrix elements of A, where A is the Hermitian operator that acts on the space of wave functions with the Hermitian inner product defined as before, are given by

$$A_{m,n} := \langle \phi_m, A(\phi_n) \rangle = \int_{-\infty}^{\infty} \overline{\phi_m(x)} A(\phi_n)$$

Definition 7.2.9. The **adjoint** $A^\dagger$ of a **linear differential operator** A is defined by

$$\langle \psi_1, A^\dagger(\psi_2) \rangle = \langle A(\psi_1), \psi_2 \rangle$$

and has the same properties as the adjoint of a linear operator. Its matrix elements are given by

$$A^\dagger_{m,n} = \overline{A_{n,m}}$$

Theorem 7.2.10. The position and momentum operators are Hermitian operators.

Proof. For the position operator, $\hat{x}$:

$$\begin{aligned}
\langle \hat{x}(\psi_1), \psi_2 \rangle &= \int_{-\infty}^{\infty} \overline{x\psi_1(x)} \psi_2(x) dx \\
&= \int_{-\infty}^{\infty} \overline{\psi_1(x)} x \psi_2(x) dx \\
&= \langle \psi_1, \hat{x}(\psi_2)
\end{aligned}$$

For the momentum operator, $\hat{p}$, by integrating parts:

$$\begin{aligned}
\langle \hat{p}(\psi_1), \psi_2 \rangle &= \int_{-\infty}^{\infty} \overline{-i\hbar \frac{\partial \psi_1(x)}{\partial x}} \psi_2(x) dx \\
&= \int_{-\infty}^{\infty} i\hbar \frac{\partial \overline{\psi_1(x)}}{\partial x} \psi_2(x) dx \\
&= \int_{-\infty}^{\infty} \overline{\psi_1(x)} \left(-i\hbar \frac{\partial \psi_2(x)}{\partial x} \right) dx + i\hbar \Big[\overline{\psi_1(x)} \psi_2(x) \Big]_{\infty}^{\infty} \\
&= \langle \psi_1, \hat{p}(\psi_2)
\end{aligned}$$

Since the wave functions must vanish as $|x| \to \infty$. □

Definition 7.2.11. The **Hamiltonian operator** is defined as

$$\hat{H}(\psi(x)) = \frac{\hat{p}^2(\psi(x))}{2m} + V(x) = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}\psi(x) + V(x)$$

and is a Hermitian operator. It relates to measurements of energy.

Example 7.2.12. For an infinite potential well in $[0, L]$, in Example 7.1.9, we defined a basis

$$\phi_n(x) = \sqrt{\frac{2}{L}} \sin\left(\frac{n\pi x}{L}\right)$$

The matrix elements of position are given by

$$\begin{aligned}
x_{m,n} &:= \langle \phi_m, \hat{x}(\phi_n) \rangle \\
&= \int_0^L x\overline{\phi_m(x)}\phi_n(x)dx \\
&= \frac{2}{L}\int_0^L x \sin\left(\frac{m\pi x}{L}\right)\sin\left(\frac{n\pi x}{L}\right)dx \\
&= \frac{1}{L}\int_0^L x\left(\cos\left(\frac{(m-n)\pi x}{L}\right) - \cos\left(\frac{(m+n)\pi x}{L}\right)\right)dx \\
&= \begin{cases} L/2 & \text{if } m = n \\ \frac{4Lmn}{\pi^2(m^2-n^2)}((-1)^{m+n} - 1) & \text{if } m \neq n \end{cases}
\end{aligned}$$

using the substitution $y = \pi x$ l with the fact that for $n \neq 0$,

$$\int_0^\pi y\cos(ny)dy = \frac{(-1)^n - 1}{n^2}$$

The matrix elements $x_{m,n}$ form a Hermitian matrix.

Example 7.2.13. For an infinite potential well in $[0, L]$, with the same basis functions as in the last example, the matrix elements of momentum are given by

$$\begin{aligned}
p_{m,n} &:= \langle \phi_m, \hat{p}(\phi_n) \rangle \\
&= -i\hbar\int_0^L \overline{\phi_m(x)}\frac{\partial\phi_n(x)}{\partial x}dx \\
&= -i\hbar\frac{2}{L}\sin\left(\frac{m\pi x}{L}\right)\frac{n\pi}{L}\cos\left(\frac{n\pi x}{L}\right)dx \\
&= -\frac{i\hbar n\pi}{L^2}\int_0^L\left(\sin\left(\frac{(m+n)\pi x}{L}\right) + \sin\left(\frac{(m-n)\pi x}{L}\right)\right)dx \\
&= \begin{cases} 0 & \text{if } m = n \\ \frac{2i\hbar mn}{L(m^2-n^2)}((-1)^{m+n} - 1) & \text{if } m \neq n \end{cases}
\end{aligned}$$

using the substitution $y = \pi x$ L with the fact that for $n \neq 0$,

$$\int_0^\pi \sin(ny) = \frac{1 - (-1)^n}{n}$$

The matrix elements $p_{m,n}$ form a Hermitian matrix.

Example 7.2.14. For an infinite potential well in $[0, L]$, with the same basis functions as in the last example, the matrix elements of Hamiltonian operator are given by

$$\begin{aligned} H_{m,n} &:= \langle \phi_m, H(\phi_n) \rangle \\ &= -\frac{\hbar^2}{2m} \int_0^L \overline{\phi_m(x)} \frac{\partial^2 \phi_n(x)}{\partial x^2} dx \\ &= E_n \frac{2}{L} \int_0^L \sin\left(\frac{m\pi x}{L}\right) \sin\left(\frac{n\pi x}{L}\right) dx \\ &= E_n \delta_{m,n} \end{aligned}$$

where

$$E_n = \frac{\hbar^2 \pi^2 n^2}{2mL^2}$$

The matrix elements $H_{m,n}$ form a diagonal matrix. The above calculation shows that the ϕ_n are eigenfunctions of the Hamiltonian operator, since

$$\hat{H}(\phi_n(x)) = E_n \phi_n(x)$$

7.3 The spectrum of a Hermitian operator

Definition 7.3.1. A wave function $\psi_a(x)$ is an **eigenfunction** of a Hermitian differential operator A with eigenvalue a if it satisfies

$$A(\psi_a(x)) = a\psi_a(x)$$

Proposition 7.3.2. For a normalised eigenfunction $\psi_a(x)$ of a Hermitian differential operator A with eigenvalue a,

- The expectation value of A^n is a^n for every $n \in \mathbb{N}$.
- The uncertainty of A is 0.

Proof.

- $\langle A^n \rangle = \langle \psi_a, A^n(\psi_a) \rangle = \langle \psi_a, a6n\psi_a \rangle = a^n \langle \psi_a, \psi_a \rangle = a^n$.
- $\Delta A = \sqrt{\langle A^2 \rangle - \langle A \rangle^2} = \sqrt{a^2 - a^2} = 0$.

□

Remark. This means that ψ_a is a wave function such that measurements of A will be a with probability 1.

Theorem 7.3.3. Let A be a Hermitian operator. Then

1. The eigenvalues of A are real, and
2. If two eigenfunctions $\psi_1(x), \psi_2(x)$ of A have distinct eigenvalues $a_1 \neq a_2$, then $\psi_1(x)$ and $\psi_2(x)$ are orthogonal.

Proof. Let $\psi_1(x), \psi_2(x)$ be eigenfunctions of A with eigenvalues a_1 and a_2, so

$$A(\psi_1(x)) = a_1\psi_1(x), \quad A(\psi_2(x)) = a_2\psi_2(x)$$

Then

$$\begin{aligned}\langle\psi_1, A(\psi_2)\rangle &= \langle\psi_1, a_2\psi_2\rangle = a_2\langle\psi_1, \psi_2\rangle \\ \langle A(\psi_2), \psi_1\rangle &= \langle a_1\psi_1, \psi_2\rangle = \overline{a_1}\langle\psi_1, \psi_2\rangle \\ \implies 0 &= (\overline{a_1} - a_2)\langle\psi_1, \psi_2\rangle\end{aligned}$$

If $a_1 = a_2$ and 4 then $(\overline{a_1} - a_1)\langle\psi, \psi\rangle = 0$, so if ψ is non-zero, then $\overline{a_1} = a_1$, so $a_1 = a_2 \in \mathbb{R}$. If $a_1 \neq a_2$ then $(a_1 - a_2)\langle\psi_1, \psi_2\rangle = 0 \implies \langle\psi_1, \psi_2\rangle = 0$ □

Definition 7.3.4. For a Hermitian operator A, a discrete spectrum of eigenvalues $\{a_n\}$ is called **non-degenerate** if there is one linearly independent eigenfunction ϕ_n for each eigenvalue a_n.

Corollary 7.3.5. We can construct an orthonormal basis with normalised eigenfunctions ϕ_n, so that

$$\langle\phi_m, \phi_n\rangle = \delta_{m,n}$$

This means any continuous square-integrable wave function ψ has a unique expansion

$$\psi(x) = \sum_n c_n\phi_n(x)$$

where

$$c_m = \langle\phi_m, \psi\rangle = \sum_n c_n\langle\phi_m, \phi_n\rangle = \sum_n c_n\delta_{m,n}$$

Corollary 7.3.6. The norm of a wave function ψ is

$$\langle\psi, \psi\rangle = \sum_{m,n} \overline{c_m}c_n\langle\phi_m, \phi_n\rangle = \sum_n |c_n|^2$$

and if ψ is normalised, then

$$\langle\psi, \psi\rangle = 1$$

So the $|c_n|^2$ can be interpreted as the probability of a measurement of A being a_n.

Example 7.3.7. In an infinite potential well in $(0, L)$, the ϕ_n defined in Example 7.1.9 eigenfunctions for the Hamiltonian operator $\hat{H}$ with eigenvalues

$$E_n = \frac{\hbar^2\pi^2n^2}{2mL^2}$$

So every wave function ψ can be written as

$$\psi(x) = \sum_{n>0} c_n\phi_n(x)$$

and $|c_n|^2$ is the probability that a measurement of energy will be E_n. These probabilities summing to 1 agrees with Parseval's theorem.

Definition 7.3.8. A Hermitian operator A can have a continuous spectrum of eigenvalues instead a discrete one. Here, we cannot choose eigenfunctions to form a complete orthonormal basis as before. To help with this case, we define the **Dirac delta function**, δ, (which is not actually a function but a distribution). Roughly speaking,

$$\delta(a) = \begin{cases} 0 & \text{if } a \neq 0 \\ \infty & \text{if } a = 0 \end{cases}$$

δ also satisfies

$$\int_{-\infty}^{\infty} \delta(a)da = 1$$

δ can be more precisely defined as the limit of the Gaussian function

$$\delta_\epsilon(a) = \frac{1}{\epsilon\sqrt{\pi}} e^{-a^2/\epsilon^2}$$

as $\epsilon \to 0^+$.

Proposition 7.3.9. (**Properties of the Dirac delta function**)

1. For every continuous function $f(a)$,

$$\int_{-\infty}^{\infty} \delta(a - a')f(a')da = f(a)$$

2. The Dirac delta function is the Fourier transform of 1:

$$\delta(a) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{iaa'} da'$$

   This implies that $\delta(a) = \delta(-a) = \overline{\delta(a)}$.

Remark. Property 1 in the above proposition is the continuous version of $\sum_n \delta_{m,n} f_n = f_m$, so $\delta(a - a')$ is the continuous version of the identity matrix $\delta_{m,n}$.

Proposition 7.3.10. For a Hermitian operator A with a continuous spectrum, there exists a basis of eigenfunctions $\phi_a(x)$ with eigenvalues $a \in \mathbb{R}$ that satisfy

$$\langle \phi_a, \phi_{a'} \rangle = \delta(a - a')$$

..

Corollary 7.3.11. Eigenfunctions $\phi_a(x)$ are therefore not square-normalisable since $\langle \phi_a, \phi_a \rangle = \infty$. However, every square-normalisable wave function can be uniquely expanded as

$$\psi(x) = \int_{-\infty}^{\infty} c(a)\phi_a(x)da$$

where $c(a)$ are complex coefficients that are continuous functions of a. $c(a)$ are given by

$$\begin{aligned} c(a) &= \int_{-\infty}^{\infty} c(a')\delta(a - a')da' \\ &= \int_{-\infty}^{\infty} c(a')\langle \phi_a, \phi_{a'} \rangle da' \\ &= \langle \phi_a, \psi \rangle \end{aligned}$$

which is the continuous version of $\langle \phi_n, \psi \rangle = c_n$.

Corollary 7.3.12. The norm of a wave function ψ can be written as

$$\begin{aligned}\langle\psi,\psi\rangle &= \int_{-\infty}^{\infty} \overline{c(a)}c(a')\langle\phi_a,\phi_{a'}\rangle da da' \\ &= \int_{-\infty}^{\infty} \overline{c(a)}c(a')\delta(a-a')da da' \\ &= \int_{-\infty}^{\infty} |c(a)|^2 da\end{aligned}$$

which is the continuous version of $\langle\psi,\psi\rangle = \sum_n |c_n|^2$.

Note that for a normalised wave function ψ,

$$\langle\psi,\psi\rangle = \int_{-\infty}^{\infty} |c(a)|^2 da = 1$$

so $|c(a)|^2$ can be interpreted as a probability distribution for measurements of A.

Example 7.3.13. The eigenfunctions of the momentum operator $\hat{p} = -i\hbar\frac{\partial}{\partial x}$ are $e^{ipx/\hbar}$ with eigenvalue p. We normalise these eigenfunctions, choosing

$$\phi_p(x) = \frac{1}{\sqrt{2\pi\hbar}} e^{ipx/\hbar}$$

which gives

$$\langle\psi_p,\psi_{p'}\rangle = \frac{1}{2\pi\hbar}\int_{-\infty}^{\infty} e^{i(p-p')x/\hbar}dx = \delta(p-p')$$

So we can write a wave function ψ as

$$\psi(x) = \int_{-\infty}^{\infty} c(p)\phi_p(x)dp = \frac{1}{\sqrt{2\pi\hbar}}\int_{-\infty}^{\infty} c(p)e^{ipx/\hbar}dp$$

which is the result of the Fourier transform.