Basic mathematics for quantum mechanics and other general uses

Andrew Lim

February 2025

Abstract

In this paper, we will go through the mathematics required to get into studying quantum mechanics without overstating the rigor and spending too much time in proofs. This paper is targeted towards people who are simply curious about quantum mechanics and want a more mathematical approach without much prior knowledge on Linear algebra. Readers are expected to understand their basic differential and integral calculus.

Contents

1 Linear Algebra

Linear algebra is the branch of mathematics that studies vectors, vector spaces, and linear transformations. It provides an essential mathematical framework for quantum mechanics, where states, observables, and evolution are all naturally expressed in terms of linear algebra concepts.

1.1 Vectors

Vectors can be interpreted as a list of values, each assigned to a unique dimension. For this sub-section, we will be using 2-dimensional vectors, generalization to higher dimensions will be simple.

Here, vectors are represented by Bras $\langle\psi|$ and Kets $|\phi\rangle$. We can interpret Kets as column matrices and Bras as row matrices, i.e.,

$$|\psi\rangle = \begin{pmatrix}\psi_1\\ \psi_2\end{pmatrix}, \langle\phi| = \begin{pmatrix}\phi_1^* & \phi_2^*\end{pmatrix}.$$

When we add 2 vectors together, we sum their components up:

$$|\psi\rangle = \begin{pmatrix}\psi_1\\ \psi_2\end{pmatrix}, |\phi\rangle = \begin{pmatrix}\phi_1\\ \phi_2\end{pmatrix} \xrightarrow{addition} |\psi\rangle + |\phi\rangle = \begin{pmatrix}\psi_1+\phi_1\\ \psi_2+\phi_2\end{pmatrix}.$$

Inner products are the sum of the products of each component, this can be interpreted as matrix multiplication:

$$|\psi\rangle = \begin{pmatrix}\psi_1\\ \psi_2\end{pmatrix}, \langle\phi| = \begin{pmatrix}\phi_1^* & \phi_2^*\end{pmatrix} \rightarrow \langle\phi|\psi\rangle = \psi_1\phi_1^* + \psi_2\phi_2^*.$$

The magnitude of a vector is defined by the following, where $|\psi|$ is the magnitude of $|\psi\rangle$:

$$|\psi|^2 = \langle\psi|\psi\rangle$$

1.2 Vector Spaces

Definition of a Vector Space:

A vector space V over a field $\mathbb{F}$ (such as $\mathbb{R}$ or $\mathbb{C}$) is a set of objects (vectors) that satisfy the following axioms under vector addition and scalar multiplication:

Vector Addition Axioms

For all vectors $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$:

1. **Closure under addition**: $\mathbf{u} + \mathbf{v} \in V$.

2. **Commutativity**: $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$.
3. **Associativity**: $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$.
4. **Existence of a zero vector**: There exists a vector $\mathbf{0}$ such that $\mathbf{v}+\mathbf{0} = \mathbf{v}$.
5. **Existence of additive inverses**: For every $\mathbf{v}$, there exists $-\mathbf{v}$ such that $\mathbf{v} + (-\mathbf{v}) = \mathbf{0}$.

Scalar Multiplication Axioms

For all $\mathbf{v} \in V$ and scalars $c, d \in \mathbb{F}$:

6. **Closure under scalar multiplication**: $c\mathbf{v} \in V$.
7. **Distributive property**: $c(\mathbf{u} + \mathbf{v}) = c\mathbf{u} + c\mathbf{v}$.
8. **Distributivity over field addition**: $(c + d)\mathbf{v} = c\mathbf{v} + d\mathbf{v}$.
9. **Associativity with scalars**: $c(d\mathbf{v}) = (cd)\mathbf{v}$.
10. **Multiplicative identity**: $1\mathbf{v} = \mathbf{v}$.

If a set satisfies these 10 properties, it is a vector space.

In Quantum mechanics, we will be using Hilbert spaces and it's corresponding dual space denoted by $\mathcal{H}$. It is defined to be an infinite dimensional vector space over $\mathbb{C}$ equipped with a dual space and inner product which use Bras and Kets.

1.3 Bras, Kets and a bit on Dual spaces

Every complex number has their corresponding dual, this is called the complex conjugation. The complex conjugation of any complex number V will be denoted by V^* which is defined by:

$$V = A + Bi \rightarrow V* = A - Bi.$$

We may say that V^* is the dual of V since every complex number V corresponds to a unique complex conjugate.

Similar to complex numbers, every Ket in a Hilbert space has a corresponding dual which are bras. They are related by the procedure that is called a Hermitian conjugate, which is the transpose of the complex conjugate of a vector.

$$|\psi\rangle = \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} \Longleftrightarrow \langle\psi| = \begin{pmatrix} \psi_1^* & \psi_2^* \end{pmatrix}.$$

1.4 Linear Operators

Linear operators will be denoted by the arrow hat symbol $\hat{M}$ and they act on vectors to transform them into other vectors.

$$\hat{M} |\psi\rangle = |\phi\rangle .$$

they can be represented as matrices and they follow the following multiplicative rule :

Let A be a matrix of size $m \times n$ and B be a matrix of size $n \times p$. The product of A and B, denoted as $C = AB$, is a matrix C of size $m \times p$, where the entry c_{ij} in the i-th row and j-th column of C is given by:

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$$

This means that each element c_{ij} is computed as the inner-product of the i-th row of A and the j-th column of B. For quantum mechanics all operators will be square matrices.

Matrices can also be transposed, the formal definition of a matrix transpose on matrix A is:

$$(A^T)_{ij} = A_{ji}, \quad \text{for all } 1 \leq i \leq n, \quad 1 \leq j \leq m.$$

1.5 Eigenvalues and Eigenvectors

A vector $|\psi\rangle$ is said to be an eigenvector of the operator $\hat{H}$ with the eigenvalue λ if and only if:

$$\hat{M} |\psi\rangle = \lambda |\psi\rangle , \lambda \in \mathbb{C}.$$

for an n x n matrix $\hat{M}$, there will be up to n eigenvectors.

1.6 Hermitian conjugation and Operators

As mentioned previously, the Hermitian conjugate of a Ket is a Bra, and vice-versa.
The definition for Hermitian conjugates of operators still hold up for operators, that is:

$$\hat{M}^{\dagger} = (\hat{M}^T)^*.$$

An operator $\hat{M}$ with the property $\hat{M} = \hat{M}^{\dagger}$ is called a Hermitian operator. They have a special property that will be important later:

$$\hat{M} = \hat{M}^{\dagger} \Longrightarrow \langle\psi| \hat{M} = \langle\phi| \Longleftrightarrow \hat{M} |\psi\rangle = |\phi\rangle .$$

Which indirectly leads to the following:

$$\hat{M} = \hat{M}^\dagger \Longrightarrow \hat{M}\,|\psi\rangle = \lambda\,|\psi\rangle \Longleftrightarrow \langle\psi|\,\hat{M} = \langle\psi|\,\lambda^*.$$

Multiplying the left hand side of that equation by $\langle\psi|$ and the right hand side of the equation by $|\psi\rangle$ gives us,

$$\lambda\,\langle\psi|\psi\rangle = \langle\psi|\hat{M}|\psi\rangle\,, \lambda^*\,\langle\psi|\psi\rangle = \langle\psi|\hat{M}|\psi\rangle\,.$$

Since both equations must be true, we may then deduce that the eigenvalue of any hermitian matrix will be real, i.e., $\lambda \in \mathbb{R}$.

Another important principle is that the eigenvectors of a Hermitian matrix form an orthonormal basis. That is, any vector that a Hermitian matrix can create can be expanded as a linear combination of the eigenvectors and that any 2 different eigenvalues will correspond to 2 orthogonal eigenvector. We can also say that 2 eigenvectors can be orthogonal although their eigenvalues are identical.

Proof:
Let $|\lambda_1\rangle$ and $|\lambda_2\rangle$ be eigenvectors of $\hat{L}$ with corresponding eigenvalues λ_1 and λ_2 respectively,

$$\hat{L}\,|\lambda_1\rangle = \lambda_1\,|\lambda_1\rangle$$

$$\hat{L}\,|\lambda_2\rangle = \lambda_1\,|\lambda_2\rangle\,.$$

Turning one of the equations into a Bra equation then gives

$$\langle\lambda_1|\,\hat{L} = \lambda_1\,\langle\lambda_1|$$

$$\hat{L}\,|\lambda_2\rangle = \lambda_1\,|\lambda_2\rangle\,.$$

Multiplying the top equation by $|\lambda_2\rangle$ and the bottom by $\langle\lambda_1|$ then gives us:

$$\langle\lambda_1|\hat{L}|\lambda_2\rangle = \lambda_1\langle\lambda_1|\lambda_2\rangle = \lambda_2\langle\lambda_1|\lambda_2\rangle,$$

which then implies that

$$(\lambda_1 - \lambda_2)\langle\lambda_1|\lambda_2\rangle = 0.$$

Due to this, all different eigenvalues must have orthogonal eigenvectors, however, similar eigenvectors may or may not have orthogonal eigenvalues.

1.7 Commutators

The commutator of 2 operators $\hat{M}$ and $\hat{L}$ is defined to be:

$$[\hat{M}, \hat{L}] = \hat{M}\hat{L} - \hat{L}\hat{M}.$$

Here are some useful commutator results to be used later:

1. $[\hat{A}, \hat{B}] = -[\hat{B}, \hat{A}]$
2. $[\hat{A}, \hat{B} + \hat{C}] = [\hat{A}, \hat{B}] + [\hat{A}, \hat{C}]$
3. $[\hat{A} + \hat{B}, \hat{C}] = [\hat{A}, \hat{C}] + [\hat{B}, \hat{C}]$
4. $[\hat{A}, \hat{B}\hat{C}] = [\hat{A}, \hat{B}]\hat{C} + \hat{B}[\hat{A}, \hat{C}]$
5. $[\hat{A}\hat{B}, \hat{C}] = \hat{A}[\hat{B}, \hat{C}] + [\hat{A}, \hat{C}]$
6. $[\hat{A} - \hat{B}, \hat{C} - \hat{D}] = [\hat{A}, \hat{C}] - [\hat{A}, \hat{D}] - [\hat{B}, \hat{C}] + [\hat{B}, \hat{D}]$
7. $[k\hat{A}, \hat{B}] = [\hat{A}, k\hat{B}] = k[\hat{A}, \hat{B}]$
8. $[\hat{A}, [\hat{B}, \hat{C}]] + [\hat{B}, [\hat{C}, \hat{A}]] + [\hat{C}, [\hat{A}, \hat{B}]] = 0$

The proof will be left as an exercise for the readers.

1.8 Complete set of commuting variables

Let's assume that there is a basis of state vectors $|\lambda_i, \mu_a\rangle$ that are simultaneous eigenvectors for operators $\hat{L}$ and $\hat{M}$, i.e.,

$$\hat{L} |\lambda_i, \mu_a\rangle = \lambda_i |\lambda_i, \mu_a\rangle$$

$$\hat{M} |\lambda_i, \mu_a\rangle = \mu_a |\lambda_i, \mu_a\rangle .$$

For simplicity's sake, we'll leave the subscripts off. In order to have a basis of simultaneous eigenvectors, $\hat{L}$ and $\hat{M}$ must commute. This is because:

$$\hat{L}(\hat{M} |\lambda, \mu\rangle) = \hat{L}\mu |\lambda, \mu\rangle = \lambda\mu |\lambda, \mu\rangle = \hat{M}(\hat{L} |\lambda, \mu\rangle),$$

This implies that $[\hat{L}, \hat{M}] |\lambda, \mu\rangle = 0$ where the 0 here is the zero vector.

Due to this equation we'll find that the condition for a vector to be an eigenvector of any 2 different operators is that they must commute. The complete set of these operators is called the complete set of commuting variables.

1.9 Tensor products, outer products and composite variables

This section will be heavily used when studying entanglement.

1.9.1 Tensor product spaces

The tensor product of 2 vector spaces $\mathbf{C}$ with basis vectors labeled with $\{H, T\}$ and $\mathbf{D}$ with basis vectors labeled with $\{1, 2, 3\}$ outputs another vector space

Tensor products are usually non-commutative, that is, $\mathbf{C} \otimes \mathbf{D} \neq \mathbf{D} \otimes \mathbf{C}$. This is because the basis vectors of $\mathbf{C} \otimes \mathbf{D}$ are $\{|H\rangle \otimes |1\rangle , |H\rangle \otimes |2\rangle , |H\rangle \otimes |3\rangle , |T\rangle \otimes$

$|1\rangle, |T\rangle \otimes |2\rangle, |T\rangle \otimes |3\rangle\}$ but the bases of $\mathbf{D} \otimes \mathbf{C}$ will be $\{|1\rangle \otimes |H\rangle, |2\rangle \otimes |H\rangle, |3\rangle \otimes |H\rangle, |1\rangle \otimes |T\rangle, |2\rangle \otimes |T\rangle, |3\rangle \otimes |T\rangle\}$. We will see why this is an issue when we look at the definition of tensor products for matrices.

1.9.2 Tensor products on operators

Let $\hat{A}$ be an $m \times n$ matrix and $\hat{B}$ be a $p \times q$ matrix. The tensor product of $\hat{A}$ and $\hat{B}$, denoted by $\hat{A} \otimes \hat{B}$, is defined as:

$$\hat{A} \otimes \hat{B} = \begin{bmatrix} a_{11}\hat{B} & a_{12}\hat{B} & \dots & a_{1n}\hat{B} \\ a_{21}\hat{B} & a_{22}\hat{B} & \dots & a_{2n}\hat{B} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}\hat{B} & a_{m2}\hat{B} & \dots & a_{mn}\hat{B} \end{bmatrix}.$$

These composite operators have a special distributive property when acting on tensor product vectors, that is:

$$(\hat{A} \otimes \hat{B})(|\psi\rangle \otimes |\phi\rangle) = \hat{A}|\psi\rangle \otimes \hat{B}|\phi\rangle.$$

1.9.3 Projection operators and outer products

Besides the usual inner product between bras and kets $\langle\psi|\phi\rangle$, we also have ourselves an outer product $|\phi\rangle\langle\psi|$.

The outer product of a ket and a bra will create an operator which acts on a bra or ket by taking the inner product of the ket and the bra in the outer product and vice-versa,

$$(|\phi\rangle\langle\psi|)|a\rangle = \langle\psi|a\rangle|\phi\rangle$$
$$\langle a|(|\phi\rangle\langle\psi|) = \langle a|\phi\rangle\langle\psi|.$$

It can also be defined as the tensor product of the ket and bra, i.e.,

$$|\phi\rangle\langle\psi| = |\phi\rangle \otimes \langle\psi|.$$

If a vector $|\psi\rangle$ has the magnitude 1, then the outer product $|\psi\rangle\langle\psi|$ is known as the projection operator of $|\psi\rangle$.

A non-geometric interpretation of "projection" could be seen as extracting relevant information from a complex data set by filtering out noise or irrelevant components. Or in simpler terms, the projection of $|\phi\rangle$ on $|\psi\rangle$ is the component of $|\phi\rangle$ along $|\psi\rangle$.

Some important things to note about projections are,

- Projection operators are Hermitian

- The vector $|\psi\rangle$ is an eigenvector of the projection operator $|\psi\rangle\langle\psi|$ with the eigenvalue 1.

- Any vector orthogonal to $|\psi\rangle$ is an eigenvector of $|\psi\rangle\langle\psi|$ with eigenvalue 0.

- The trace of an operator is defined as, $Tr(\hat{L}) = \sum_i \langle i|\hat{L}|i\rangle$. and the trace of a projection operator is always 1, since the trace of a matrix is the sum of it's eigenvalues.

- The sum of projection operators in a basis system is equal to the identity operator, this is known as the completeness relation,

$$\sum_i |i\rangle\langle i| = \hat{I}.$$

This results in the formula $\langle\psi|\hat{L}|\psi\rangle = Tr(|\psi\rangle\langle\psi|\hat{L})$. To prove this, we can start by using the definition of the trace,

$$Tr(|\psi\rangle\langle\psi|\hat{L}) = \sum_i \langle i|\psi\rangle\langle\psi|\hat{L}|i\rangle,$$

since these inner products are just numbers we can rewrite it as,

$$Tr(|\psi\rangle\langle\psi|\hat{L}) = \sum_i \langle\psi|\hat{L}|i\rangle\langle i|\psi\rangle.$$

Using the completeness relation we find that $\langle\psi|\hat{L}|\psi\rangle = Tr(|\psi\rangle\langle\psi|\hat{L})$.

2 Some functional analysis

Functional analysis is best described as, an extension to linear algebra. Think, Linear Algebra premium. In here, we'll be dealing with infinite dimensional vectors and start making the relation between vectors and functions.

2.1 Infinite dimensions???

The words "infinite dimensions" may sound daunting at first. However, notice that we have generalized most concepts in the previous section of linear algebra to multiple dimensions. Applications of generalization to infinite dimensions will become more clear in the following sections.

2.2 State vector and the wavefunction

To demonstrate the applications of using infinite dimensions, we'll have to introduce the wavefunction and state vector. This section won't go too deep into what the wavefunction and state vectors are, just the enough for the mathematical concepts to be understood.

The state vector $|\Psi\rangle$, can basically be interpreted as a vector that holds all the data in a system, with each dimension being assigned a "state" (which are represented as basis vectors) and the value of the state vector in the dimension being related to the probability we find the system in the state.

An important point is that for the state vector to have true physical meaning, as we'll see later, it'll have to be normalized i.e., $|\Psi| = 1$.

2.2.1 The wavefunction

The quantum "wavefunction" despite it's name, may not have anything to do with classical waves.

The wavefunction can be defined as the coefficient of each state in the state vector.

To obtain the wavefunction, we must first take a set of basis, for example, let's let our basis be the eigenvectors of the Hermitian operator $\hat{L}$. Expanding the state vector into the superposition of eigenstates (aka eigenvectors) of $\hat{L}$ will give us this:

$$|\Psi\rangle = \sum_i \psi(\lambda_i) |\lambda_i\rangle .$$

The coefficient on the equation above $\psi(\lambda_i)$, is called the wavefunction on the $\hat{L}$ basis and it can be expressed as the inner product,

$$\psi(\lambda) = \langle\lambda|\Psi\rangle.$$

Due to this, we can express the state vector as an infinite dimensional vector which can be turned into a wavefunction which varies with the eigenbasis (basis of eigenvectors) taken.

2.2.2 A short sidenote on functions

After defining the wave function, we can try to generalize the relation between a function and a vector and get used to the functional techniques to be used. The formal definition of a function is :

A function f from set X to a set Y is a subset of $X \times Y$ such that for every $x \in X$, there is exactly one $y \in Y$ for which (x, y) belongs to f.

The $\times$ here is called a cartesian product, an operation between 2 sets with

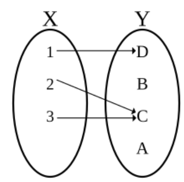

Figure 1: A visualization of functional mapping

the definition $A \times B = \{(a, b) : \forall a \in A, \forall b \in B\}$.

A bad habit of making students study calculus is that they assume that all functions map numbers to numbers or assume they're mostly continuous. However, this is only a special case, functions can take on a domain with discrete values, not necessarily numbers. Functions are very flexible, a function f can map colors to cars, etc.

2.2.3 Vectors as continuous functions

We can think of vectors as functions. By letting each basis vector of a chosen basis become elements of the domain, we may then treat a vector like it's a function that maps each basis vector to a corresponding value, an example of this is if we define a vector to be

$$|\Psi\rangle = \sum_i \lambda_i |i\rangle = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \vdots \\ \lambda_n \\ \vdots \end{bmatrix},$$

we'll find that $\langle i|\Psi\rangle = \lambda_i$ is a function which maps a basis vector $|i\rangle$ to an element in the vector λ_i.

Although basis vectors, like λ_i, seem to be discrete domain elements, that is the set of basis vectors. We can use the fact that our Hilbert space $\mathcal{H}$ is an infinite dimensional vector space. With infinite dimensions, we have infinite elements in our domain. All that is required left to make our domain a set of continuous values is to find a basis to make our function carefully.

A few more important things to note when making vectors into functions with continuous domains:

- Integrals will replace sums

Since our domain is continuous, summation notation must be replaced by integrals. Using the Reimann sum definition of an integral we'll find that technically, sums and integrals are just the same thing, just in a continuous case.

$$\sum_i \iff \int dx.$$

- Dirac delta replaces Kronecker delta

  The Kronecker delta function follows the definition $\delta_{ij} = 1 \iff i = j$ and it's usually used in sums. For example:

  $$\sum_i \delta_{ij} = 1.$$

  For continuous functions, we'll use what's called the Dirac delta function. The Dirac delta function can be defined as,

  $$\delta(x) = \begin{cases} 0 & x \neq 0 \\ \infty & x = 0 \end{cases}$$

  Such that it follows the rule:

  $$\int_{-\infty}^{\infty} dx \delta(x - x_0) f(x) = f(x_0).$$

2.3 The position and momentum eigenbasis

For now, I want you to forget everything you think you know about momentum and position. This section will be heavily used for Particle dynamics.

In quantum mechanics, the momentum and position operators are what are called "observables". These observables need to be Hermitian.

2.3.1 The position eigenbasis

Let us first define the position operator as $\hat{x}$ and give it a corresponding eigenvector $|x\rangle$ with the eigenvalue equation $\hat{x}|x\rangle = x|x\rangle$ where $x \in \mathbb{R}$. Here, the wavefunction in the position basis will simply be the inner product,

$$\langle x|\Psi\rangle = \psi(x).$$

The position basis here is an example of a continuous basis. This is because the eigenvalue of the position operator and the values x can take is continuous as it can take on any value of $\mathbb{R}$.

A good example of using integrals as a sum comes up when we try to take the inner product of the state vector with itself. Given that we know $\psi(x)$, we can easily find that the inner product of $|\Psi\rangle$ with itself is going to be:

$$\langle\Psi|\Psi\rangle = \int_{-\infty}^{\infty} dx\ \psi^*(x)\psi(x).$$

This is true due to the completeness relation, which states that the sum of the projection operators in a complete eigenbasis is equal to the Identity operator, i.e.,

$$\sum_i |i\rangle\langle i| = \hat{I}.$$

The completeness relation here can be found when we turn the wavefunction into an inner product,

$$\int_{-\infty}^{\infty} dx\ \ |x\rangle\langle x| = \hat{I} \to \langle\Psi|\Psi\rangle = \int_{-\infty}^{\infty} dx\ \psi^*(x)\psi(x) = \int_{-\infty}^{\infty} dx\ \langle\Psi|x\rangle\langle x|\Psi\rangle.$$

As mentioned earlier, since the position operator is an observable, we can check whether $\hat{x}$ is an observable by seeing whether it's Hermitian or not. To check this, we simply need to show that for arbitrary vectors,

$$\langle\psi|\hat{x}|\phi\rangle = \langle\phi|\hat{x}|\psi\rangle^*.$$

To do so, we may use the completeness relation and now we need to show that,

$$\int_{-\infty}^{\infty} dx\ \psi^*(x)\hat{x}\phi(x) \Longleftrightarrow \int_{-\infty}^{\infty} dx\ \phi^*(x)\hat{x}\psi(x)$$

Since x has to be a real number, we can easily say that the operators $\hat{x}$ is hermitian.

2.3.2 Momentum

In the position eigenbasis, the momentum operator $\hat{p}$ is defined as

$$\hat{p} = -i\hbar\frac{\partial}{\partial x},$$

we may let $|p\rangle$ be the eigenvector for $\hat{p}$ and get the eigenvalue equation $\hat{p}\,|p\rangle = p\,|p\rangle$. The derivation of this requires an understanding of classical mechanics and, to be more precise, an understanding of how momentum is the generator of changes in position. In mathematics, the term generator appears in semigroup theory, where a generator of a semigroup of operators defines how an operator evolves a function in time.

To demonstrate this, we may consider a function $f(x)$, let ϵ represent an infinitesimal shift in x here. Expanding $f(x+\epsilon)$ into a Taylor series gives:

$$f(x+\epsilon) = f(x) + \epsilon\frac{d}{dx}f(x) + \mathcal{O}(\epsilon^2) + ...$$

Because epsilon is the infinitesimal change in time, we can ignore the higher order terms and stick with,

$$f(x+\epsilon) = f(x) + \epsilon\frac{d}{dx})f(x) = (1+\epsilon\frac{d}{dx})f(x).$$

Since we can generate an infinitesimal shift in the inputted variable x by using a derivative operator, we say that the derivative operator $\frac{d}{dx}$ is the generator of translation in x.

Earlier, it was mentioned that momentum in physics is the generator of translation in position, therefore, $\hat{p} \propto \frac{d}{dx}$. labelling the derivative operator as $\hat{D}$ we can now try check whether $\frac{d}{dx}$ is hermitian using the condition, $\langle\psi|\hat{D}|\phi\rangle = \langle\phi|\hat{D}|\psi\rangle^*$.

To take the inner product we may use the completeness relation again and obtain,

$$\int_{-\infty}^{\infty} dx\ \psi^*(x)\hat{D}\phi(x) = \left(\int_{-\infty}^{\infty} dx\ \phi^*(x)\hat{D}\psi(x)\right)^*.$$

To show that $\hat{D}$ is hermitian we must show that the equation above is true. We can try do so by using integration by parts,

$$\int_{-\infty}^{\infty} dx\ \psi^*(x)\frac{d}{dx}\phi(x) = \psi^*(x)\phi(x)\Big|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} dx\ \phi(x)\frac{d}{dx}\psi^*(x)$$

Due to the normalization condition, we'll find that the first term of the equation vanishes, leaving us with the equation

$$\langle\psi|\hat{D}|\phi\rangle = -\langle\phi|\hat{D}|\psi\rangle^*.$$

This implies that the operator $\hat{D}$ is anti-hermitian, that is $\hat{D}^\dagger = -\hat{D}$. Now because the momentum operator is hermitian, all we need to so to $\hat{D}$ to make it hermitian is to divide by i. which leaves us with a valid momentum operator $\hat{p} = -i\hat{D} = -i\frac{d}{dx}$. However, due to unit conventions, we have to multiply this by $\hbar$ where $\hbar \approx 1.0546 \times 10^{-34}$Js. This is how we result in the equation

$$\hat{p} = -i\hbar\frac{\partial}{\partial x}.$$

2.4 Position Momentum relations

In Hamiltonian mechanics (a branch of classical mechanics), momentum and position are the fundamental variables that describe a physical system's state.

A similar statement can be said in quantum mechanics as well. In this section, we'll cover how our position and momentum operator interacts.

2.4.1 Canonical commutation relation

The commutation relation $[\hat{x}, \hat{p}] = i\hbar$ is called the canonical commutation relation. It can very easily be derived by expanding the commutator and applying it to a wavefunction,

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

simplifying this gives

$$[\hat{x}, \hat{p}]\psi(x) = -i\hbar\left(x\frac{\partial}{\partial x}\psi(x) - \frac{\partial}{\partial x}(x\psi(x))\right) = -i\hbar\psi(x) \Longrightarrow [\hat{x}, \hat{p}] = i\hbar.$$

2.4.2 $\hat{x}$ and $\hat{p}$ in the momentum basis

In order not to take too long repeating a similar process as the previous sections, I will simply tell that the position operator $\hat{x}$ in the momentum basis takes the form:

$$\hat{x} = i\hbar\frac{\partial}{\partial p}$$

This can be found by repeating the steps from the previous sections, except we start with the momentum eigenspace instead of position. The momentum operator will then take the form of a multiplicative operator similar to the position operator in it's own eigenbasis.

2.4.3 Inner product relation $\langle x|p\rangle$

Finding the inner product $\langle x|p\rangle$ could be pretty tricky especially since we don't exactly get a wavefunction to work with. However, we could try to work with both eigenbasis and their operators.

We start with defining the inner products $\langle x|p\rangle$ and $\langle p|x\rangle$ with,

$$\langle x|p\rangle = p(x), \langle x|p\rangle = x(p) = p^*(x)$$

Clearly, since $\hat{x}$ and $\hat{p}$ are both observables, we'll find that,

$$(\langle x|\hat{x})\ |p\rangle = \langle x|\ (\hat{x}|p\rangle)$$

and

$$(\langle x|\hat{p})\ |p\rangle = \langle x|\ (\hat{p}|p\rangle).$$

Turning these equations into functional equations and using the operator's different forms in different basis by using $x(p)$ and $p(x)$ will then yield,

$$p \; p(x) = -i\hbar \frac{\partial}{\partial x} p(x)$$

and

$$x \; x(p) = i\hbar \frac{\partial}{\partial p} x(p)$$

By solving the coupled differential equation we'll find that the solution to $\langle x|p\rangle$ is

$$\langle x|p\rangle = Ae^{\frac{-ipx}{\hbar}},$$

where A is a constant. To solve for A we may use the normalization condition,

$$\langle p|p'\rangle = \delta(p - p')$$

where $|p\rangle$ and $|p'\rangle$ are both basis vectors in the momentum eigenbasis. We may then use the completeness relation with the projection in the $\hat{x}$ basis,

$$\int_{-\infty}^{\infty} dx \; \langle p|x\rangle\langle x|p'\rangle = \int_{-\infty}^{\infty} dx \; |A|^2 e^{\frac{ix(p-p')}{\hbar}} = \delta(p - p')$$

We can take the $|A|^2$ and factor it out of the integral to then get,

$$|A|^2 \int_{-\infty}^{\infty} dx \; e^{\frac{ix(p-p')}{\hbar}} = \delta(p - p').$$

To solve the integral $\int_{-\infty}^{\infty} dx \; e^{\frac{ix(p-p')}{\hbar}}$, we can use the identity,

$$\int_{-\infty}^{\infty} dx \; e^{ikx} = 2\pi\delta(k),$$

in order to find that,

$$\int_{-\infty}^{\infty} dx \; e^{\frac{ix(p-p')}{\hbar}} = 2\pi\hbar\delta(p - p').$$

From here we can deduce that $|A|^2 = \frac{1}{2\pi\hbar}$ and an easy pick for A would simply be the square root of it, therefore, $A = \frac{1}{\sqrt{2\pi\hbar}}$ and we'll find that

$$\langle x|p\rangle = \frac{1}{\sqrt{2\pi\hbar}} e^{\frac{ixp}{\hbar}}.$$

2.5 The Fourier Transform (quantum mechanical interpretation)

In order to switch between the momentum and position basis when doing calculations with your wavefunction, we'll need what's called a Fourier transform.

Here's how it can be done,

$$\langle x|\Psi\rangle = \psi(x), \langle p|\Psi\rangle = \tilde{\psi}(p),$$

we first let $\psi(x)$ be the wavefunction in the position basis and $\tilde{\psi}(p)$ be the wavefunction in the momentum basis. Notice that we added a ˜ to the wavefunction in the momentum basis, this is because these wavefunctions are technically different functions and we are required to label differently due to this.

In order to go from the position basis to the momentum's , we can simply express the momentum wavefunction as $\tilde{\psi}(p) = \langle p|\Psi\rangle$. We then insert the completeness relation by acting the sum of all position projections on the inner product,

$$\langle p|\Psi\rangle = \int_{-\infty}^{\infty} dx \ \langle p|x\rangle\langle x|\Psi\rangle.$$

All we have to do now is to simplify the integral using the inner product relation found previously to obtain,

$$\tilde{\psi}(p) = \langle p|\Psi\rangle = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} dx \ e^{\frac{-ixp}{\hbar}} \psi(x).$$

And that's our Fourier transform from the position to momentum space. To go from the momentum to position space, we can also perform simillar steps but swap the x for the p in order to obtain the following transform,

$$\langle x|\Psi\rangle = \psi(x) = \frac{1}{\sqrt{2\pi\hbar}} \int_{-\infty}^{\infty} dx \ e^{\frac{ixp}{\hbar}} \tilde{\psi}(p).$$