

QUANTUM
Mechanics, Optics, and Information
- A Quick Guide -

Huan Q. Bui

Colby College

PHYSICS & MATHEMATICS
Statistics

Class of 2021

June 27, 2019

Preface

Greetings,

This text is my reading notes from Shankar's *Principles of Quantum Mechanics, Second Edition*, *Introductory Quantum Optics* by , *Optical Coherence & Quantum Optics*. Additional comes from my class notes, my comments/interpretations/solutions, and other books/articles/notes. There are three parts to this text, as the title suggests. The Quantum Mechanics part covers the principles of quantum mechanics. A majority of this part will be my reading notes from Shankar's book. The other two topics parts cover selected phenomena in quantum optics and quantum information theory.

Background in linear algebra will be very helpful. I will try to cover some of the mathematical background, but a lot of familiarity will be assumed.

Enjoy!

Contents

Preface	2
1 QUANTUM MECHANICS	5
1.1 Mathematical Introduction	6
1.1.1 Linear Vector Spaces	6
1.1.2 Inner Product Spaces	7
1.1.3 Dual Spaces and Dirac Notation	8
1.1.4 Subspaces, Sum and Direct Sum of Subspaces	10
1.1.5 Linear Operators	11
1.1.6 Matrix Elements of Linear Operators	12
1.1.7 Matrix Elements of Products of Operators	12
1.1.8 The Adjoint of an Operator	13
1.1.9 Hermitian, Anti-Hermitian, and Unitary Operators	13
1.1.10 Active and Passive Transformation	14
1.1.11 The Eigenvalue Problem	15
1.1.12 Functions of Operators and Related Concepts	16
1.1.13 Generalization to Infinite Dimensions	18
1.2 Review of Classical Mechanics	27
1.2.1 Principle of Least Action & Lagrangian Mechanics	27
1.2.2 The Electromagnetic Lagrangian	28
1.2.3 The two-body problem	29
1.2.4 The Hamiltonian Formalism	30
1.2.5 The Electromagnetic Force in the Hamiltonian Scheme	33
1.2.6 Cyclic Coordinates, Poisson Brackets, and Canonical Transformations	33
1.3 The Postulates –a General Discussion	35
1.3.1 The Postulates	35
1.3.2 Postulates I - III	36
1.3.3 The Schrödinger Equation	36
1.4 Simple Problems in One Dimension	37
1.4.1 The Free Particle	37
1.4.2 The Particle in a Box	37
1.4.3 The Continuity Equation for Probability	37
1.4.4 The Single-Step Potential: A Problem in Scattering	37
1.4.5 The Double-Slit Experiment	37

1.4.6	Some Theorems	37
1.5	The Classical Limit	38
1.6	The Harmonic Oscillator	39
1.7	The Path Integral Formulation of Quantum Theory	40
1.8	The Heisenberg Uncertainty Relation	41
1.9	Systems with N Degrees of Freedom	42
1.10	Symmetries and Their Consequences	43
1.11	Rotational Invariance and Angular Momentum	44
1.12	The Hydrogen Atom	45
1.13	Spin	46
1.14	Additional of Angular Momentum	47
1.15	Variational and WKB Methods	48
1.16	Time-Independent Perturbation Theory	49
1.17	Time-Dependent Perturbation Theory	50
1.18	Scattering Theory	51
1.19	The Dirac Equation	52
1.20	Path Integrals-II	53
2	QUANTUM OPTICS	55
3	QUANTUM INFORMATION	57
4	PROBLEMS & SOLUTIONS	59
4.1	Quantum Mechanics	60
4.2	Quantum Optics	61
4.3	Quantum Information	62

Part 1

QUANTUM MECHANICS

1.1 Mathematical Introduction

1.1.1 Linear Vector Spaces

We should familiar with defining characteristics of linear vector spaces at this point. Here are some important definitions/theorems again:

Definition 1.1.1. A linear vector space \mathbf{V} is a collection of objects called *vectors* for which there exists

1. A definite rule for summing, and
2. A definite rule for scaling, with the following features:
 - Closed under addition: for $x, y \in \mathbf{V}$, $x + y \in \mathbf{V}$.
 - Closed under scalar multiplication: $x \in \mathbf{V}$, then $ax \in \mathbf{V}$ for some scalar a .
 - Scalar multiplication is distributive.
 - Scalar multiplication is associative.
 - Addition is commutative.
 - Addition is associative.
 - There exists a (unique) null element in \mathbf{V} .
 - There exists a (unique) additive inverse.

Vector spaces are defined over some field. The field can be real numbers, complex numbers, or it can also be finite. As for good practice, we will begin to label vectors with Dirac bra-ket notation. So, for instance, $|v\rangle \in \mathbf{V}$ denotes vector $v \in \mathbf{V}$. Basic manipulations of these vectors are intuitive:

1. $|0\rangle$ is unique, and is the null element.
2. $0|V\rangle = |0\rangle$.
3. $|-V\rangle = -|V\rangle$.
4. $|-V\rangle$ is a unique additive inverse of $|V\rangle$.

The reasons for choosing to use the Dirac notation will become clear later on. Another important basic concept is *linear (in)dependence*. Of course, there are a number of equivalent statement for linear independence. We shall just give one here:

Definition 1.1.2. A set of vectors is said to be linearly independent if the only linear relation

$$\sum_{i=1}^n a_i |i\rangle = |0\rangle \tag{1.1}$$

is the trivial one where the components $a_i = 0$ for any i .

The next two basic concepts are *dimension* and *basis*.

Definition 1.1.3. A vector space \mathbf{V} has dimension n if it can accommodate a maximum of n linearly independent vectors. We denote this n -dimensional vector space as \mathbf{V}^n .

We can show that

Theorem 1.1.1. Any vector $|v\rangle \in \mathbf{V}^n$ can be written (uniquely) as a linear combination of any n linearly independent vectors.

Definition 1.1.4. A set of n linearly independent vectors in a n -dimensional space is called a *basis*. So if $|1\rangle, \dots, |n\rangle$ form a basis for \mathbf{V}^n , then any $|v\rangle \in \mathbf{V}$ can be written uniquely as

$$|v\rangle = \sum_{i=1}^n a_i |i\rangle. \quad (1.2)$$

It is nice to remember the following:

$$\boxed{\text{Linear Independence} = \text{Basis} + \text{Span}} \quad (1.3)$$

When a collection of vectors span a vector space \mathbf{V} , it just means that any $|v\rangle \in \mathbf{V}$ can be written as a linear combination of (some of) these vectors.

The algebra of linear combinations is quite intuitive. If $|v\rangle = \sum_i a_i |i\rangle$ and $|w\rangle = \sum_i b_i |i\rangle$ then

1. $|v + w\rangle = \sum_i (a_i + b_i) |i\rangle$.
2. $c|v\rangle = c \sum_i a_i |i\rangle = \sum_i ca_i |i\rangle$.

A linear algebra text will of course provide a much better coverage of these topics.

1.1.2 Inner Product Spaces

A generalization of the familiar dot product is the *inner product* or the *scalar product*. An inner product between two vectors $|v\rangle$ and $|w\rangle$ is denoted $\langle v|w\rangle$. An inner product has to satisfy the following properties:

1. Conjugate symmetry (or skew-symmetry): $\langle v|w\rangle = \langle w|v\rangle^*$.
2. Positive semi-definiteness: $\langle v|v\rangle \geq 0$.
3. Linearity in ket: $\langle v|aw + bz\rangle = a \langle v|w\rangle + b \langle v|z\rangle$.
4. Conjugate-linearity in bra: $\langle av + bz|w\rangle = \bar{a} \langle v|w\rangle + \bar{b} \langle z|w\rangle$.

Definition 1.1.5. An inner product space is a vector space with an inner product.

Definition 1.1.6. $\langle v|w\rangle = 0 \iff |v\rangle \perp |w\rangle$.

Definition 1.1.7. The *norm* (or length) of $|v\rangle$ is defined as

$$\|v\| = \sqrt{\langle v|v\rangle}. \quad (1.4)$$

Unit vectors have unit norm. Unit vectors are said to be *normalized*.

Definition 1.1.8. A set of basis vectors all of unit norm, which are pairwise orthogonal will be called an *orthonormal basis* or ONB.

Let $|v\rangle = \sum_i a_i |i\rangle$ and $|w\rangle = \sum_i b_i |j\rangle$, then

$$\langle v|w\rangle = \sum_i a_i^* b_i \langle i|j\rangle. \quad (1.5)$$

Theorem 1.1.2. Gram-Schmidt: Given a linearly independent basis, we can form linear combinations of the basis vectors to obtain an orthonormal basis.

Suppose that the Gram-Schmidt process gives us an ONB then we have

$$\langle i|j\rangle = \delta_{ij}. \quad (1.6)$$

As a result,

$$\langle v|w\rangle = \sum_i v_i^* w_i. \quad (1.7)$$

Alternatively, we can think this as doing the standard inner products of vectors whose entries are the components of the vectors $|v\rangle$, $|w\rangle$ in the basis:

$$|v\rangle \rightarrow \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \quad |w\rangle \rightarrow \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix} \implies \langle v|w\rangle = \begin{bmatrix} v_1^* & v_2^* & \dots & v_n^* \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix}. \quad (1.8)$$

We can also easily see that

$$\langle v|v\rangle = \sum_i |v_i|^2 \geq 0. \quad (1.9)$$

1.1.3 Dual Spaces and Dirac Notation

Here we deal with some technical details involving the *ket* (the column vectors) and the *bra* (the row vectors). Column vectors are concrete manifestations of an abstract vector $|v\rangle$ in a basis, and we can work backward to go from the column vectors to the kets. We can do a similar thing with the bra vectors - since there's nothing special about writing the entries in a column versus in a row. However, we will do the following. We know that associated with every ket $|v\rangle$ is a column vector. So let its adjoint, which is a row vector, be associated with the bra, called $\langle v|$. Now, we have two vector spaces, the space of kets and the dual space of bras. There is a basis of vectors $|i\rangle$ for expanding kets and a similar basis $\langle i|$ for expanding bras.

Expansion of Vectors in an ONB

It is extremely useful for us to be able to express a vector in an ONB. Suppose we have a vector $|v\rangle$ in an ONB $|i\rangle$. Then, let $|v\rangle$ be written as

$$|v\rangle = \sum_i v_i |i\rangle. \quad (1.10)$$

To find the components v_i , we take the inner product of $|v\rangle$ with $|j\rangle$:

$$\langle j|v\rangle = \sum_i v_i \langle j|i\rangle = \sum_i v_i \delta_{ij} = v_j. \quad (1.11)$$

With this, we can rewrite the vector $|v\rangle$ in the basis $|i\rangle$ as

$$|v\rangle = \sum_i |i\rangle \langle i|v\rangle. \quad (1.12)$$

Adjoint Operations

Here is a few details regarding taking the adjoints of vectors. Suppose that

$$|v\rangle = \sum_i v_i |i\rangle = \sum_i |i\rangle \langle i|v\rangle. \quad (1.13)$$

Then,

$$\langle v| = \sum_i \langle i| v_i^*. \quad (1.14)$$

Now, because $v_i = \langle i|v\rangle$, we have $v_i^* = \langle v|i\rangle$. Thus,

$$\langle v| = \sum_i \langle v|i\rangle \langle i|. \quad (1.15)$$

In plain words, the rule for taking the adjoint is the following. To take the adjoint of an equation involving bras and kets and coefficients, reverse the order of all factors, exchanging bras and kets and complex conjugating all coefficients.

Gram-Schmidt process

Again, the Gram-Schmidt process lets us convert a linearly independent basis into an orthonormal one. For a two-dimensional case, procedure is the following:

1. Rescale the first by its own length, so it becomes a unit vector. This is the first (orthonormal) unit vector.
2. Subtract from the second vector its projection along the first, leaving behind only the part perpendicular to the first. (Such a part will remain since by assumption the vectors are nonparallel).

3. Rescale the left over piece by its own length. We now have the second basis vector: it is orthogonal to the first and of unit length.

In general, let $|I\rangle, |II\rangle, \dots$ be a linearly independent basis. The first vector of the orthonormal basis will be

$$|1\rangle = \frac{|I\rangle}{\| |I\rangle \|}. \quad (1.16)$$

For the second vector in the basis, consider

$$|2'\rangle = |II\rangle - |1\rangle \langle 1|II\rangle. \quad (1.17)$$

We can see that $|2'\rangle$ is orthogonal to $|1\rangle$:

$$\langle 1|2'\rangle = \langle 1|II\rangle - \langle 1|1\rangle \langle 1|II\rangle = 0. \quad (1.18)$$

So dividing $|2'\rangle$ by its norm gives us, $|2\rangle$, the second element in the ONB. To find the third element in the ONB, we have to first make sure it is orthogonal to both $|I\rangle$ and $|II\rangle$, so let us consider

$$|3'\rangle = |III\rangle - |1\rangle \langle 1|III\rangle - |2\rangle \langle 2|III\rangle. \quad (1.19)$$

Once again we have $|3'\rangle$ orthogonal to both $|1\rangle$ and $|2\rangle$. Normalizing $|3'\rangle$ gives us $|3\rangle$, the third element in the ONB. We can now see how this process continues to the last element.

Schwarz and Triangle Inequality

Just two small yet very important details:

Theorem 1.1.3. Schwarz Inequality:

$$|\langle v|w\rangle| \leq \|v\| \|w\| \quad (1.20)$$

Theorem 1.1.4. Triangle Inequality:

$$\|v + w\| \leq \|v\| + \|w\|. \quad (1.21)$$

1.1.4 Subspaces, Sum and Direct Sum of Subspaces

I'm not too happy with the definitions given by Shankar's book. He also uses the notation for direct sum to indicate vector space addition, which is very confusing. Any linear algebra textbook would provide better definitions. For equivalent statements about directness of vector space sums, check out my [Matrix Analysis](#) notes.

1.1.5 Linear Operators

Again, a rigorous definition of an operator can be found in almost any linear algebra textbook. But here, we can simply think of an operator as just some linear transformation from a vector space to itself. Say, if Ω is some operator that sends $|v\rangle$ to $|v'\rangle$, we write

$$\Omega |v\rangle = |v'\rangle. \quad (1.22)$$

By definition, $|v\rangle$ and $|v'\rangle$ are contained in the same vector space. Now, we note that Ω can also act on bras:

$$\langle v| \Omega = \langle v'|. \quad (1.23)$$

But of course the order of writing things is different, and once again, $\langle v|$ and $\langle v'|$ are contained in the same (dual) space.

Next, because Ω is linear, we have the following familiar rules:

$$\Omega \alpha |v_i\rangle = \alpha \Omega |v_i\rangle. \quad (1.24)$$

$$\Omega \{\alpha |v_i\rangle + \beta |v_j\rangle\} = \alpha \Omega |v_i\rangle + \beta \Omega |v_j\rangle. \quad (1.25)$$

$$\langle v_i| \alpha \Omega = \langle v_i| \Omega \alpha \quad (1.26)$$

$$\{\langle v_i| \alpha + \langle v_j| \beta\} \Omega = \alpha \langle v_i| \Omega + \beta \langle v_j| \Omega. \quad (1.27)$$

One of the nice features of linear operators is that the action of an operator is completely determined by what it does to the basis vectors. Suppose

$$|v\rangle = \sum_i v_i |i\rangle \quad (1.28)$$

and

$$\Omega |i\rangle = |i'\rangle, \quad (1.29)$$

then

$$\Omega |v\rangle = \sum_i \Omega v_i |i\rangle = \sum_i v_i \Omega |i\rangle = \sum_i v_i |i'\rangle. \quad (1.30)$$

The next point of interest is *products* of operators. As we might have seen, operators don't always commute. A product of operators applied to a vector just means operators are applied in sequence. The *commutator* of two operators Ω, Λ is defined as

$$\Omega \Lambda - \Lambda \Omega \equiv [\Omega, \Lambda]. \quad (1.31)$$

In general, $[\Omega, \Lambda]$ is not zero. Suppose three operators Ω, Λ, Θ are involved, then we have two useful relations:

$$[\Omega, \Lambda \Theta] = \Lambda [\Omega, \Theta] + [\Omega, \Lambda] \Theta \quad (1.32)$$

$$[\Lambda \Omega, \Theta] = \Lambda [\Omega, \Theta] + [\Lambda, \Theta] \Omega. \quad (1.33)$$

We notice that the form resembles the chain rule in calculus.

1.1.6 Matrix Elements of Linear Operators

One thing we will hear very often in quantum mechanics is the idea of matrix elements. The idea, it turns out, is very simple. Suppose we have a basis $|i\rangle$, and an operator Ω such that

$$\Omega|i\rangle = |i'\rangle. \quad (1.34)$$

Then, for

$$|v\rangle = \sum_i v_i |i\rangle, \quad (1.35)$$

we have

$$\Omega|v\rangle = \Omega \sum_i v_i |i\rangle = \sum_i v_i \Omega|i\rangle = \sum_i v_i |i'\rangle. \quad (1.36)$$

Because we know Ω and $|i\rangle$, $|i'\rangle$ is also known, as in its components in the basis $|j\rangle$ (un-primed) are known:

$$\langle j|i'\rangle = \langle j|\Omega|i\rangle \equiv \Omega_{ji}, \quad (1.37)$$

where the n^2 numbers Ω_{ji} are the matrix elements of Ω in this basis. Now, if

$$\Omega|v\rangle = |v'\rangle \quad (1.38)$$

then the components of the transformed ket $|v'\rangle$ can be expressed in terms of the components of $|v\rangle$ and the matrix elements Ω_{ji} :

$$v'_i = \langle i|v'\rangle = \langle i|\Omega|v\rangle = \langle i|\Omega \sum_j v_j |j\rangle = \sum_j v_j \langle i|\Omega|j\rangle = \sum_j \Omega_{ij} v_j. \quad (1.39)$$

We can see the above equation in matrix form as well:

$$\begin{bmatrix} v'_1 \\ \vdots \\ v'_n \end{bmatrix} = \begin{bmatrix} \langle 1|\Omega|1\rangle & \dots & \langle 1|\Omega|n\rangle \\ \vdots & \ddots & \vdots \\ \langle n|\Omega|1\rangle & \dots & \langle n|\Omega|n\rangle \end{bmatrix} \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}. \quad (1.40)$$

The elements of the first column are simply the components of the first transformed basis vector $|1'\rangle = \Omega|1\rangle$ in the given basis. Likewise, the elements of the j^{th} column represent the image of the j^{th} basis vector after Ω acts on it.

1.1.7 Matrix Elements of Products of Operators

To get the matrix elements of a product of two operators, we do the following. Suppose we have operators Ω and Λ , then

$$(\Omega\Lambda)_{ij} = \langle i|\Omega\Lambda|j\rangle = \langle i|\Omega\Lambda|j\rangle. \quad (1.41)$$

Now, we observe that

$$\mathcal{I} = \sum_k |k\rangle \langle k|. \quad (1.42)$$

So,

$$(\Omega\Lambda)_{ij} = \sum_k \langle i|\Omega|k\rangle \langle k|\Lambda|j\rangle = \sum_k \Omega_{ik}\Lambda_{kj}. \quad (1.43)$$

1.1.8 The Adjoint of an Operator

Recall that for a scalar α

$$\langle \alpha v| = \langle v| \alpha^*, \quad (1.44)$$

then we have a similar thing with operators if

$$\Omega|v\rangle = |v'\rangle \quad (1.45)$$

then

$$\langle \Omega v| = \langle v|\Omega^\dagger, \quad (1.46)$$

where Ω^\dagger is *the* adjoint of Ω . The relationship between Ω^\dagger and Ω can be seen in a basis. We consider the matrix elements of Ω^\dagger in a basis:

$$(\Omega^\dagger)_{ij} = \langle i|\Omega^\dagger|j\rangle = \langle \Omega i|j\rangle = \langle j|\Omega i\rangle^* = \langle j|\Omega|i\rangle^* = \Omega_{ji}^*. \quad (1.47)$$

We see that

$$\Omega_{ij}^\dagger = \Omega_{ji}^*, \quad (1.48)$$

i.e., in matrix form, Ω^\dagger is the conjugate transpose of Ω .

The rule for taking adjoints of equations is rather simple: When a product of operators, bras, kets, and explicit numerical coefficients is encountered, reverse the order of all factors and make the substitution $\Omega \leftrightarrow \Omega^\dagger$, $| \rangle \leftrightarrow \langle |$, $a \leftrightarrow a^*$.

1.1.9 Hermitian, Anti-Hermitian, and Unitary Operators

Definition 1.1.9. An operator Ω is Hermitian $\iff \Omega = \Omega^\dagger$.

Definition 1.1.10. An operator Ω is anti-Hermitian $\iff \Omega = -\Omega^\dagger$.

Shankar's book ignores a bigger class of operators called *normal* operators. Normal operators commute with their adjoints. In a sense, normal operators act *like numbers*. Hermitian (or self-adjoint) operators are a subset of normal operators. So, the number-likeness of normal operators carries over to Hermitian operators and anti-Hermitian operators as well. Hermitian and anti-Hermitian operators are like pure real and pure imaginary numbers. Just as every number

maybe be decomposed into a sum of pure real and pure imaginary parts, it turns out that we can decompose every operator into its Hermitian and anti-Hermitian parts.

$$\Omega = \frac{\Omega + \Omega^\dagger}{2} + \frac{\Omega - \Omega^\dagger}{2}. \quad (1.49)$$

One can verify that the first terms is Hermitian, and the second term is anti-Hermitian.

Definition 1.1.11. An operator \mathcal{U} is unitary $\iff \mathcal{U}\mathcal{U}^\dagger = \mathcal{I}$.

Unitary operators are like complex numbers of unit modulus.

Theorem 1.1.5. Unitary operators preserves the inner product between the vectors they act on.

Proof. Suppose

$$|v'\rangle = \mathcal{U}|v\rangle \quad (1.50)$$

$$|w'\rangle = \mathcal{U}|w\rangle. \quad (1.51)$$

Then

$$\langle v'|w'\rangle = \langle \mathcal{U}v|\mathcal{U}w\rangle = \langle v|\mathcal{U}^\dagger\mathcal{U}|w\rangle = \langle v|w\rangle. \quad (1.52)$$

□

Theorem 1.1.6. The columns (or rows) of a unitary matrix form an ONB.

Proof. Refer to a linear algebra text. The key is to consider an inner product between any two columns/rows. □

1.1.10 Active and Passive Transformation

Suppose all $|v\rangle$ is unitarily transformed to $|v'\rangle$:

$$|v\rangle \rightarrow \mathcal{U}|v\rangle. \quad (1.53)$$

Then under this transformation, the matrix elements of any operator Ω are modified as follows:

$$|v'\rangle\Omega|v\rangle \rightarrow |\mathcal{U}v'\rangle\Omega|\mathcal{U}v\rangle = \langle v'|\mathcal{U}^\dagger\Omega\mathcal{U}|v\rangle. \quad (1.54)$$

It is clear that the same change is equivalent to leaving the vectors alone and subjecting all operators to the change

$$\Omega \rightarrow \mathcal{U}^\dagger\Omega\mathcal{U}. \quad (1.55)$$

Active transformation refers to changing the vectors, while *passive transformation* refers to changing the operators.

1.1.11 The Eigenvalue Problem

I won't say much about what eigenvectors and eigenvalues are because we should be familiar with these concepts at this point. But just to introduce some terminology, each operator has certain kets of its own called *eigenkets*, on which its action is simply that of scaling. So, eigenkets are just a different word for eigenvectors of an operator:

$$\Omega |v\rangle = \omega |v\rangle. \quad (1.56)$$

Shankar's book talks about the characteristic equation and characteristic polynomial. While these are legitimate ways to find eigenvalues and eigenvectors, it is often very difficult. I'd prefer Leo Livshits' and Sheldon Axler's way and use minimal polynomials instead. I would steer away from determinants and characteristic polynomials at this point.

Theorem 1.1.7. Eigenvalues of a Hermitian operator are real.

Proof. Suppose

$$\Omega |w\rangle = a |w\rangle, \quad (1.57)$$

then

$$\langle w | \Omega | w \rangle = a \langle w | w \rangle, \quad (1.58)$$

and thus

$$a^* \langle w | w \rangle = \langle w | \Omega^\dagger | w \rangle = \langle w | \Omega | w \rangle = a \langle w | w \rangle. \quad (1.59)$$

So we have

$$(a - a^*) \langle w | w \rangle = 0. \quad (1.60)$$

Because $|w\rangle$ are eigenkets, they are cannot be the zero vector. This means $a = a^*$. \square

Some might worry about the existence of eigenvalues of Hermitian operators. But worry no more, because Hermitian operators are a subclass of normal operators, which are a subclass of diagonalizable operators. This simply says Hermitian matrices are diagonalizable, and all its eigenvalues are real. But it turns out there is a little bit more to this.

Theorem 1.1.8. For every Hermitian operator Ω , there exists an ONB comprised entirely of the eigenvectors of Ω .

Once again, this should be no surprise if one has studied normal operators. Hermitian operators inherit this property from its normalness. This property of normal operators are called the Spectral Theorem (for normal operators, of course). The proof of all this can be found in many linear algebra texts.

Theorem 1.1.9. The eigenvalues of a unitary operator are complex numbers of unit modulus.

Proof. The key to the proof is using inner products. \square

Simultaneous Diagonalization of Two Hermitian Operators

I would say the topic of simultaneous diagonalizability is covered quite well in Leo Livshits' course and hence in my [Matrix Analysis](#) notes. But here I will just give the most important results.

Theorem 1.1.10. If Ω and Λ are two commuting Hermitian operators, there exists a basis of common eigenvectors that diagonalizes them both.

This result is not too surprising if we have studied simultaneous diagonalizability before. A more general theorem says that

$$\text{Simultaneous diagonalizability} \iff \text{Individual diagonalizability} + \text{Commutativity}. \quad (1.61)$$

It is clear that because all Hermitian operators are diagonalizable, if two Hermitian operators commute, they are simultaneously diagonalizable.

The Propagator

In quantum mechanics (and classical mechanics of course), it is quite common to have some final state vector be obtained from an initial state vector multiplied by some matrix, which is independent of the initial state. We call this matrix the *propagator*.

The central problem in quantum mechanics is finding the state of a quantum system $|\psi\rangle$, which obeys the Schrödinger equation:

$$i\hbar \left| \dot{\psi} \right\rangle = J |\psi\rangle \quad (1.62)$$

where the Hermitian operator H is called the *Hamiltonian*. We will see much more of this as we move on.

1.1.12 Functions of Operators and Related Concepts

In this section, we look at whether it makes sense to define functions of operators. We will only restrict ourselves to functions that can be written as a power series. Consider a series

$$f(x) = \sum_{n=0}^{\infty} a_n x^n \quad (1.63)$$

where x is a scalar. We defined the same function of an operator to be

$$f(\Omega) = \sum_{n=0}^{\infty} a_n \Omega^n. \quad (1.64)$$

Now, this definition only makes sense if we have convergence. Consider this example:

$$e^{\Omega} = \sum_{n=0}^{\infty} \frac{\Omega^n}{n!}, \quad (1.65)$$

where Ω is Hermitian. In the eigenbasis of Ω , Ω is diagonal. This means we can add and/or take powers of Ω by add and/or take powers of the diagonal entries. We can find that

$$e^\Omega = \begin{bmatrix} \sum_{m=0}^{\infty} \frac{\omega_1^m}{m!} & & \\ & \ddots & \\ & & \sum_{m=0}^{\infty} \frac{\omega_n^m}{m!} \end{bmatrix} \quad (1.66)$$

where ω_i are the eigenvalues of Ω . We note that each entry in the expression above converges to e^{ω_i} .

Derivatives of Operators with Respect to Parameters

Now, consider some operator $\Theta(\lambda)$ that depends on a parameter λ . The derivative of Θ with respect to λ is defined to be

$$\frac{d\Theta(\lambda)}{d\lambda} = \lim_{\Delta\lambda \rightarrow 0} \left[\frac{\Theta(\lambda + \Delta\lambda) - \Theta(\lambda)}{\Delta\lambda} \right]. \quad (1.67)$$

If $\Theta(\lambda)$ is written as a matrix, then the matrix of $d\Theta/d\lambda$ is obtained by differentiating the matrix elements of $\Theta(\lambda)$. A case that might be interesting to us is

$$\Theta(\lambda) = e^{\lambda\Omega}. \quad (1.68)$$

It turns out that if Ω is Hermitian or “nice enough” then

$$\frac{d\Theta(\lambda)}{d\lambda} = \Omega e^{\lambda\Omega} = e^{\lambda\Omega} \Omega = \Theta(\lambda) \Omega = \Omega \Theta(\lambda). \quad (1.69)$$

Conversely, if we have

$$\frac{d\Theta(\lambda)}{d\lambda} = \Theta(\lambda) \Omega \quad (1.70)$$

then

$$\Theta(\lambda) = c e^{\lambda\Omega} \quad (1.71)$$

where c is some operator. But we have to be careful that c might not commute with $e^{\lambda\Omega}$.

The business of whether two operators commute or don’t can make things slightly more complicated. If Θ and Ω commute, i.e., $[\Theta, \Omega] = 0$, then the rules of exponentiation carries over very nicely:

$$e^{a\Omega} e^{b\Omega} = e^{(a+b)\Omega} \quad (1.72)$$

$$e^{a\Omega} e^{b\Theta} = e^{a\Omega + b\Theta} \quad (1.73)$$

$$e^{a\Omega} e^{b\Theta} e^{-a\Omega} = e^{b\Theta}. \quad (1.74)$$

If $[\Omega, \Theta] \neq 0$, then the second and third equations no longer hold. Likewise, in differentiating a product, we have to be extra careful:

$$\frac{d}{d\lambda} e^{\lambda\Omega} e^{\lambda\Theta} = \Omega e^{\lambda\Omega} e^{\lambda\Theta} + e^{\lambda\Omega} e^{\lambda\Theta} \Theta. \quad (1.75)$$

While $[\Omega, e^{\lambda\Omega}] = 0$, because Θ and Ω might not commute, we can't bring Ω over to the right of $e^{\lambda\Theta}$.

1.1.13 Generalization to Infinite Dimensions

The Dirac delta function

Consider the ordered n -tuple $\{f_n(x_1), \dots, f_n(x_n)\}$ as components of a ket $|f_n\rangle$ in a vector space $\mathbf{V}^n(\mathbb{R})$:

$$|f_n\rangle \leftrightarrow \begin{bmatrix} f_n(x_1) \\ \vdots \\ f_n(x_n) \end{bmatrix}. \quad (1.76)$$

The basis vectors in this space are:

$$|x_i\rangle \leftrightarrow \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \leftarrow i^{th} \text{ place}. \quad (1.77)$$

The basis vectors satisfy *orthogonality* and *completeness*:

$$\langle x_i | x_j \rangle = \delta_{ij} \quad (1.78)$$

$$\sum_{i=1}^n |x_i\rangle \langle x_i| = \mathcal{I}. \quad (1.79)$$

With this,

$$|f_n\rangle = \sum_{i=1}^n f_n(x_i) |x_i\rangle. \quad (1.80)$$

We next define the inner product in this space:

$$\langle f_n | g_n \rangle = \sum_{i=1}^n f_n(x_i) g_n(x_i). \quad (1.81)$$

The functions f_n and g_n are said to be orthogonal if $\langle f_n | g_n \rangle = 0$. We also have that

$$\langle f_n | f_n \rangle = \sum_{i=1}^n [f_n(x_i)]^2. \quad (1.82)$$

For finite n , nothing “bad” can really happen here. But what if n is infinity? What we need is the redefinition of the inner product for finite n in such a way that as n goes to infinity, we get a smooth limit. A natural choice is

$$\langle f_n | g_n \rangle = \sum_{i=1}^n f_n(x_i) g_n(x_i) \frac{L}{n+1} \quad (1.83)$$

where L is the length of the interval. If we now let n go to infinity, we get

$$\langle f | g \rangle = \int_0^L f(x) g(x) dx \quad (1.84)$$

$$\langle f | f \rangle = \int_0^L f^2(x) dx. \quad (1.85)$$

Now, if we consider complex functions as well in some interval $a \leq x \leq b$, the inner product becomes:

$$\langle f | g \rangle = \int_a^b f^*(x) g(x) dx. \quad (1.86)$$

But what are the basis vectors in this space and are they normalized? We know that

$$\langle x | x' \rangle = 0 \quad (1.87)$$

if $x \neq x'$. But what if $x = x'$? It turns out that we cannot simply require $\langle x | x \rangle = 1$. The best way to see this is to deduce the correct normalization. We start with the completeness relation:

$$\int_a^b |x' \rangle \langle x'| dx' = \mathcal{I}. \quad (1.88)$$

Now, consider this

$$\int_a^b \langle x | x' \rangle \langle x' | f \rangle dx' = \langle x | \mathcal{I} | f \rangle = \langle x | f \rangle. \quad (1.89)$$

This is nothing but the projection of $|f\rangle$ along the basis ket $|x\rangle$, which is just $f(x)$. So, we also have $f(x') = \langle x' | f \rangle$. Let the inner product $\langle x | x' \rangle$ be some unknown function $\delta(x, x')$. Since $\delta(x, x')$ vanishes if $x \neq x'$, we can restrict the

integral to an infinitesimal region near $x' = x$. With these, the equality above gives

$$\int_{x-\epsilon}^{x+\epsilon} \delta(x, x') f(x') dx' = f(x). \quad (1.90)$$

In this infinitesimal region, $f(x)$ can be assumed to be constant, and thus can be pulled out of the integral, leaving

$$f(x) \int_{x-\epsilon}^{x+\epsilon} \delta(x, x') dx' = f(x). \quad (1.91)$$

And so we have

$$\int_{x-\epsilon}^{x+\epsilon} \delta(x, x') dx' = 1. \quad (1.92)$$

Clearly, $\delta(x, x')$ cannot be finite at $x = x'$. It should be infinite in such a way that its integral is 1. Since $\delta(x, x')$ depends only on the difference $x - x'$, we can write it as $\delta(x - x')$. So, the function $\delta(x - x')$ has the properties:

$$\begin{cases} \delta(x - x') = 0, & x \neq x' \\ \int_a^b \delta(x - x') dx' = 1 & a < x < b \end{cases} \quad (1.93)$$

This is called the **Dirac delta function** and it fixes the normalization of the basis vectors:

$$\langle x | x' \rangle = \delta(x - x'). \quad (1.94)$$

The Dirac delta function is “strange” in the sense that its value is either zero or infinite. It’s thus useful to view it as the limit of a Gaussian:



of the form

$$g_{\Delta}(x - x') = \frac{1}{\sqrt{\pi\Delta^2}} e^{-(x-x')^2/\Delta^2}. \quad (1.95)$$

It is clear that the area under the curve is one (one can easily check this). Also, $\Delta \rightarrow 0$, g_Δ becomes closer and closer to $\delta(x - x')$ (the area under the curve is the same, while the width of the peak becomes smaller and smaller).

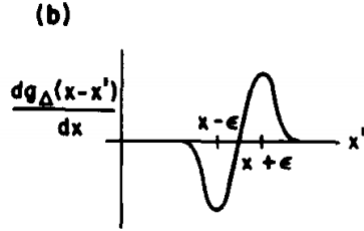
From Gaussian model, we know that the delta function is not only real but also even:

$$\delta(x - x') = \delta(x' - x). \quad (1.96)$$

Next, we consider the derivative of $\delta(x - x')$ with respect to x :

$$\delta'(x - x') = \frac{d}{dx} \delta(x - x') = -\frac{d}{dx'} \delta(x - x'). \quad (1.97)$$

Once again, we consider the Gaussian model. We consider $dg_\Delta(x - x')/dx = -dg_\Delta(x - x')/dx'$ as a function of x' :



As g_Δ shrinks,, each bump at $\pm\epsilon$ will become, up to a scale factor, the δ function, such that

$$\int \delta'(x - x') f(x') dx' \propto f(x + \epsilon) - f(x - \epsilon) = 2\epsilon \left. \frac{df}{dx'} \right|_{x=x'}.$$

The constant of proportionality turns out to be $\epsilon/2$, and so

$$\int \delta'(x - x') f(x') dx' = \left. \frac{df}{dx'} \right|_{x=x'} = \frac{df(x)}{dx}. \quad (1.98)$$

In short, we can describe the δ' function as

$$\delta'(x - x') = \delta(x - x') \frac{d}{dx'}. \quad (1.99)$$

In this way, we can describe higher derivatives of δ :

$$\frac{d^n \delta(x - x')}{dx^n} = \delta(x - x') \frac{d^n}{dx'^n}. \quad (1.100)$$

Next, we will develop an alternative representation of the delta function. Suppose we're given a function $f(x)$. The Fourier transform is given by

$$\mathcal{F}[f](k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-ikx} f(x) dx. \quad (1.101)$$

And the inverse is given by

$$\mathcal{F}^{-1}[f](x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{ikx'} f(k) dk. \quad (1.102)$$

Feeding the inverse formula into the transform formula, we get

$$f(x') = \int_{-\infty}^{\infty} \left(\frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ik(x-x')} dk \right) f(x) dx. \quad (1.103)$$

Comparing this result to (1.90), we see that

$$\boxed{\frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ik(x-x')} dk = \delta(x-x')} \quad (1.104)$$

Operators in Infinite Dimensions

Let us revisit the linear transformation:

$$\Omega |f\rangle = |\tilde{f}\rangle \quad (1.105)$$

in the vector space whose basis is $|x\rangle$.

Let us assume that this action takes place in an infinite-dimensional vector space. Consider the differential operator. We can write the equation above as

$$D_x |f(x)\rangle = |df/dx\rangle. \quad (1.106)$$

What are the matrix elements of the operator D in the $|x\rangle$ basis? To find the matrix elements, we do exactly as before:

$$\langle x | D | f \rangle = \left\langle x \left| \frac{df}{dx} \right\rangle = \frac{df(x)}{dx}, \quad (1.107)$$

followed by

$$\int \langle x | D | x' \rangle \langle x' | D | f \rangle dx' = \frac{df}{dx}. \quad (1.108)$$

We deduce that

$$\langle x | D | x' \rangle = D_{xx'} = \delta'(x-x') = \delta(x-x') \frac{d}{dx'}. \quad (1.109)$$

We notice that D is not Hermitian because if

$$D_{xx'} = D_{x'x}^* \quad (1.110)$$

then

$$D_{xx'} = \delta'(x-x') = D_{x'x}^* = \delta'(x'-x) = -\delta'(x-x'), \quad (1.111)$$

which is obviously not true. But we can convert D to a Hermitian matrix by multiplying it with a purely imaginary number. Consider

$$K = -iD. \quad (1.112)$$

Then

$$K_{x'x}^* = [-i\delta'(x' - x)]^* = i\delta'(x' - x) = -i\delta'(x - x') = K_{xx'}. \quad (1.113)$$

It turns out that this is not enough to make K Hermitian, as we shall show now. Suppose we have $|f\rangle$ and $|g\rangle$ in the function space whose images in the x basis are $f(x)$ and $g(x)$ in the interval $a - b$. If K is Hermitian, we must have that

$$\langle g|K|f\rangle = \langle g|Kf\rangle = \langle Kf|g\rangle^* = \langle f|K^\dagger|g\rangle^* = \langle f|K|g\rangle^*. \quad (1.114)$$

So, we ask

$$\int_a^b \int_a^b \langle g|x\rangle \langle x|K|x'\rangle \langle x'|f\rangle dx dx' \stackrel{?}{=} \left(\int_a^b \int_a^b \langle f|x\rangle \langle x|K|x'\rangle \langle x'|g\rangle dx dx' \right)^*. \quad (1.115)$$

Or, equivalently, we ask that if $K = -iD$, then

$$\int_a^b g^*(x) \left[-i \frac{df(x)}{dx} \right] dx \stackrel{?}{=} \left\{ \int_a^b f^*(x) \left[-i \frac{dg(x)}{dx} \right] dx \right\} = i \int_a^b \frac{dg^*}{dx} f(x) dx. \quad (1.116)$$

Integrating the left hand side by parts gives

$$-ig^*(x)f(x) \Big|_a^b + i \int_a^b \frac{dg^*}{dx} f(x) dx. \quad (1.117)$$

So for equality to hold, we require that

$$\boxed{-ig^*(x)f(x) \Big|_a^b = 0} \quad (1.118)$$

Thus, in contrast to the finite-dimensional case, $K_{xx'} = K_{x'x}^*$ is not a sufficient condition for K to be Hermitian. We must also look at the behavior of the functions at the end points a and b . So what kinds of functions make this work? One set of such functions are the possible configurations $f(x)$ of the string clamped at $x = 0$ and at $x = L$. These functions of course have zero boundary conditions. But this condition (1.118) can also be fulfilled in another way.

Consider functions in 3-dimensional space, parameterized by r, θ, ϕ . Suppose that these functions are single-valued, say, $f(\theta) = f(\theta + 2\pi)$. In the space of these functions, $K = -iD$ is Hermitian. This is very easy to verify since the condition (1.118) is met:

$$-ig^*(x)f(x)\Big|_0^{2\pi} = -ig(2\pi)f(2\pi) + ig^*(0)f(0) = 0. \quad (1.119)$$

In quantum mechanics, we are interested in functions over the full interval $-\infty \leq x \leq \infty$. These functions into two classes: those that vanish at infinity and those that don't. Functions that don't vanish at infinity behave as e^{-ikx} , where k is real. It is clear that if K is sandwiched between two functions of the first class or two functions where one comes from each class, then K is Hermitian, because the boundary terms vanish. But if K is sandwiched between two functions of the second class, then whether K is Hermitian depends on whether

$$e^{ikx}e^{-ik'x}\Big|_{-\infty}^{\infty} \stackrel{?}{=} 0. \quad (1.120)$$

If $k' = k$ then K is Hermitian. If $k' \neq k$ then the answer is unclear because $e^{i(k-k')x}$ oscillates. It turns out that there exists a way of defining a limit for such functions that connect make up their minds: the limit as $|x| \rightarrow \infty$. This limit is defined to be the average over a large interval. According to this prescription, we have, say as $x \rightarrow \infty$:

$$\lim_{x \rightarrow \infty} e^{ikx}e^{-ik'x} = \lim_{L \rightarrow \infty, \Delta \rightarrow \infty} \frac{1}{\Delta} \int_L^{L+\Delta} e^{i(k-k')x} dx = 0 \quad k \neq k'. \quad (1.121)$$

And thus K is Hermitian in this space as well.

Next, we are interested in the eigenvalue problem of K . Let us start with

$$K|k\rangle = k|k\rangle. \quad (1.122)$$

Following the standard procedure,

$$\langle x|K|k\rangle = k\langle x|k\rangle \implies \int \langle x|K|x'\rangle \langle x'|k\rangle dx' = k\psi_k(x) \quad (1.123)$$

$$\implies -i\frac{d}{dx}\psi_k(x) = k\psi_k(x) \quad (1.124)$$

where $\psi_k(x) = \langle x|k\rangle$. This is a very simple differential equation whose solution is

$$\psi_k(x) = Ae^{ikx}. \quad (1.125)$$

Let us chose A to be such that the function is normalized. In this case $A = 1/\sqrt{2\pi}$. And so,

$$|k\rangle \sim \frac{1}{\sqrt{2\pi}}e^{ikx}, \quad (1.126)$$

and

$$\langle k|k'\rangle = \int_{-\infty}^{\infty} \langle k|x\rangle \langle x|k'\rangle dx = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i(k-k')x} dx = \delta(x-x'). \quad (1.127)$$

Now, because K is Hermitian, functions that are expanded in the x basis with components $f(x) = \langle x|f\rangle$ must also have an expansion in the K basis. What are the components in this expansion? We first look at the components in the K basis, starting with $|k\rangle$:

$$f(k) = \langle k|f\rangle = \int_{-\infty}^{\infty} \langle k|x\rangle \langle x|f\rangle dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-ikx} f(x) dx. \quad (1.128)$$

To get back to the x basis, we simply apply the inverse transform:

$$f(x) = \langle x|f\rangle = \int_{-\infty}^{\infty} \langle x|k\rangle \langle k|f\rangle dk = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{ikx} f(k) dk. \quad (1.129)$$

Thus the familiar Fourier transform gives us the passage from one complete basis to another. What about the matrix elements of K in the k basis? It turns out that these elements are trivial:

$$\langle k|K|k'\rangle = k' \langle k|k'\rangle = k' \delta(k-k'). \quad (1.130)$$

So, the k basis is generated by the Hermitian operator K . So what generates the x basis? Let us call this operator X , and that

$$X|x\rangle = x|x\rangle. \quad (1.131)$$

Its matrix elements in the x basis are

$$\langle x'|X|x\rangle = x\delta(x'-x). \quad (1.132)$$

To find its actions on functions, let us define

$$X|f\rangle = |\tilde{f}\rangle. \quad (1.133)$$

And so

$$\langle x|X|f\rangle = \int \langle x|X|x'\rangle \langle x'|f\rangle dx' = xf(x) = \langle x|\tilde{f}\rangle = \tilde{f}(x). \quad (1.134)$$

Therefore,

$$\tilde{f}(x) = xf(x). \quad (1.135)$$

So, X has the effect of multiplying a function f by x :

$$X|f(x)\rangle = |xf(x)\rangle. \quad (1.136)$$

We notice that there is a nice reciprocity between X and K . Let us compute the matrix elements of X in the k basis:

$$\langle k | X | k' \rangle = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-ikx} x e^{ik'x} dx \quad (1.137)$$

$$= i \frac{d}{dk} \left(\frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i(k'-k)x} dx \right) = i \delta'(k - k'). \quad (1.138)$$

So, if $|g(k)\rangle$ is a ket whose image in the k basis is $g(k)$ then

$$X |g(k)\rangle = \left| \frac{idg(k)}{dk} \right\rangle. \quad (1.139)$$

Thus we have the following. In the x basis, X acts as x . In the k basis, X acts as $-id/dx$. On the other hand, in the k basis, K acts as k , and in the x basis as $-id/dk$. Operators with such an interrelationship are said to be **conjugate** of each other. Now, the conjugate operators K and X don't commute. Let us calculate their commutator. Suppose we have some ket $|f\rangle$. Then,

$$X |f\rangle \rightarrow x f(x) \quad (1.140)$$

$$K |f\rangle \rightarrow -i \frac{df(x)}{dx}. \quad (1.141)$$

This is just the definition of these operators. Next,

$$XK |f\rangle \rightarrow -ix \frac{df(x)}{dx} \quad (1.142)$$

$$KX |f\rangle \rightarrow -i \frac{d}{dx} x f(x). \quad (1.143)$$

Thus,

$$[X, K] |f\rangle \rightarrow -ix \frac{df(x)}{dx} + ix \frac{df(x)}{dx} + if \rightarrow i\mathcal{I} |f\rangle. \quad (1.144)$$

So, we have for X and K conjugate of each other,

$$[X, K] = i\mathcal{I}. \quad (1.145)$$

1.2 Review of Classical Mechanics

1.2.1 Principle of Least Action & Lagrangian Mechanics

Suppose we have a particle in a potential $V(x)$. Newton tells us that

$$m \frac{d^2 x}{dt^2} = -\frac{dV}{dx}. \quad (1.146)$$

In general coordinates,

$$m_j \frac{d^2 x_j}{dt^2} = -\frac{\partial V}{\partial x_j}. \quad (1.147)$$

In Lagrangian mechanics, we first define the *Lagrangian*: $\mathcal{L} = T - V$, where T is the kinetic energy and V is the potential energy. This makes $\mathcal{L} = \mathcal{L}(x, \dot{x}, t)$. For each path connecting (x_i, t_i) to (x_f, t_f) , the *action* is given by

$$S[x(t)] = \int_{t_i}^{t_f} \mathcal{L}(x, \dot{x}) dt. \quad (1.148)$$

The classical path which the particle follows is one which minimizes $S[x(t)]$. Variational methods (requiring $\delta S = 0$ and boundary terms to vanish) give us the **Euler-Lagrange equation(s)**:

$$\boxed{\frac{\partial \mathcal{L}}{\partial x(t)} = \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \dot{x}(t)}} \quad (1.149)$$

Details of this derivation can be found in many differential equation textbooks. We can easily show how Newton's Second law emerges from the Euler-Lagrange equation by setting $T = mv^2/2$. In which case, we get

$$\frac{d}{dt}(m\dot{x}) = m\ddot{x} = -\frac{dV}{dx}. \quad (1.150)$$

In general coordinates, we get the same thing:

$$m\ddot{x}_i = -\frac{\partial V}{\partial x_i}. \quad (1.151)$$

Now, we notice that we have assumed the potential V to be velocity-independent. The force of a magnetic field \mathbf{B} on a moving charge is excluded by this restriction ($\mathbf{F} = q\mathbf{v} \times \mathbf{B}$). We will show shortly how to accommodate this force in the Lagrangian formalism. However, this treatment will leave \mathcal{L} no longer in the form $T - V$. So, we will be free from the notion that \mathcal{L} has the form $T - V$, by only requiring that \mathcal{L} gives the correct equations of motion.

Suppose in some generalized coordinates, we have

$$\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{q}_i} \right) = \frac{\partial \mathcal{L}}{\partial q_i}. \quad (1.152)$$

It turns out that the *form* of the Euler-Lagrange equation is invariant under change of coordinates. The Euler-Lagrange equation above can be made to resemble Newton's Second law if one defines a quantity:

$$p_i = \frac{\partial \mathcal{L}}{\partial \dot{q}_i} \quad (1.153)$$

as the **canonical momentum conjugate to q_i** and the quantity

$$F_i = \frac{\partial \mathcal{L}}{\partial q_i} \quad (1.154)$$

as the *generalized force conjugate to q_i* . Note that these quantities are not always linear momentum and force. They can be angular momentum and torque, for instance. In many cases, we can find conservation laws from the Lagrangian, but we won't go into the details for now.

1.2.2 The Electromagnetic Lagrangian

As promised, in this subsection we will incorporate electromagnetism into the Lagrangian formalism. Recall that the force on a charge q due to an electric field \mathbf{E} and a magnetic field \mathbf{B} is given by

$$\mathbf{F} = q \left(\mathbf{E} + \frac{\mathbf{v}}{c} \times \mathbf{B} \right) \quad (1.155)$$

where $\mathbf{v} = \dot{\mathbf{r}}$ is the velocity of the charged particle, and c is the speed of light. It turns out that if we use

$$\mathcal{L}_{EM} = \frac{1}{2} m \mathbf{v} \cdot \mathbf{v} - q\phi + \frac{q}{c} \mathbf{v} \cdot \mathbf{A} \quad (1.156)$$

we get the correct electromagnetic force laws. We note that ϕ and \mathbf{A} are the scalar and the vector potentials related to \mathbf{E} and \mathbf{B} via:

$$\mathbf{E} = -\nabla\phi - \frac{1}{c} \frac{\partial \mathbf{A}}{\partial t} \quad (1.157)$$

and

$$\mathbf{B} = \nabla \times \mathbf{A}. \quad (1.158)$$

With respect to this Lagrangian, the Euler-Lagrange equation is

$$\frac{d}{dt} \left(m\dot{x}_i + \frac{q}{c} \mathbf{A}_i \right) = -q \frac{\partial \phi}{\partial x_i} + \frac{q}{c} \frac{\partial \mathbf{v} \cdot \mathbf{A}}{\partial x_i}. \quad (1.159)$$

We can combine the $i = 1, 2, 3$ equations into one vector equation:

$$\frac{d}{dt} \left(m\mathbf{v} + \frac{q\mathbf{A}}{c} \right) = -q\nabla\phi + \frac{q}{c} \nabla(\mathbf{v} \cdot \mathbf{A}). \quad (1.160)$$

Rewriting this gives

$$\frac{d}{dt}m\mathbf{v} = -q\nabla\phi + \frac{q}{c} \left[-\frac{d\mathbf{A}}{dt} + \nabla(\mathbf{v} \cdot \mathbf{A}) \right] \quad (1.161)$$

The canonical momentum is then

$$\mathbf{p} = \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{x}}} = m\mathbf{v} + \frac{q\mathbf{A}}{c}. \quad (1.162)$$

Now, the total derivative $d\mathbf{A}/dt$ has two parts:

$$\frac{d\mathbf{A}}{dt} = \frac{\partial \mathbf{A}}{\partial t} + (\mathbf{v} \cdot \nabla)\mathbf{A} \quad (1.163)$$

where

$$(\mathbf{v} \cdot \nabla)_i = \frac{dx_i}{dt} \frac{\partial}{\partial x_i}. \quad (1.164)$$

Thus we have

$$\frac{d}{dt}m\mathbf{v} = -q\nabla\phi - \frac{q}{c} \frac{\partial \mathbf{A}}{\partial t} + \frac{q}{c} [\nabla(\mathbf{v} \cdot \mathbf{A}) - (\mathbf{v} \cdot \nabla)\mathbf{A}]. \quad (1.165)$$

Now, we use the identity

$$\mathbf{v} \times (\nabla \times \mathbf{A}) = \nabla(\mathbf{v} \cdot \mathbf{A}) - (\mathbf{v} \cdot \nabla)\mathbf{A} \quad (1.166)$$

to get

$$\frac{d}{dt}m\mathbf{v} = -q\nabla\phi - \frac{q}{c} \frac{\partial \mathbf{A}}{\partial t} + \frac{q}{c} \mathbf{v} \times (\nabla \times \mathbf{A}). \quad (1.167)$$

Using the definition of \mathbf{E} and \mathbf{B} in relation to ϕ and \mathbf{A} we indeed get the correct force law:

$$\frac{d}{dt}m\mathbf{v} = \mathbf{F} = q \left(\mathbf{E} + \frac{\mathbf{v}}{c} \times \mathbf{B} \right) \quad (1.168)$$

1.2.3 The two-body problem

The two-body problem can be solved more elegantly in the *center-of-mass coordinate system* where

$$\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2 \quad (1.169)$$

and

$$\mathbf{r}_{CM} = \frac{m_1\mathbf{r}_1 + m_2\mathbf{r}_2}{m_1 + m_2}. \quad (1.170)$$

The inverse formulas are as follow:

$$\mathbf{r}_1 = \mathbf{r}_{CM} + \frac{m_2 \mathbf{r}}{m_1 + m_2} \quad (1.171)$$

$$\mathbf{r}_2 = \mathbf{r}_{CM} - \frac{m_1 \mathbf{r}}{m_1 + m_2} \quad (1.172)$$

The original Lagrangian is

$$\mathcal{L} = \frac{1}{2}m_1|\dot{\mathbf{r}}_1|^2 + \frac{1}{2}m_2|\dot{\mathbf{r}}_2|^2 - V(\mathbf{r}_1 - \mathbf{r}_2). \quad (1.173)$$

In CM coordinate system, the Lagrangian becomes:

$$\mathcal{L} = \frac{1}{2}(m_1 + m_2)|\dot{\mathbf{r}}_{CM}|^2 + \frac{1}{2}\frac{m_1 m_2}{m_1 + m_2}|\dot{\mathbf{r}}|^2 - V(\mathbf{r}). \quad (1.174)$$

By doing this, we have in a sense “decoupled” the problem:

$$\mathcal{L}(\mathbf{r}, \dot{\mathbf{r}}, \mathbf{r}_{CM}, \dot{\mathbf{r}}_{CM}, t) = \mathcal{L}(\mathbf{r}, \dot{\mathbf{r}}, t) + \mathcal{L}(\mathbf{r}_{CM}, \dot{\mathbf{r}}_{CM}, t). \quad (1.175)$$

The first fictitious particle is the center of mass of the system. The motion of the center of mass is often uninteresting, so we can always go to the center of mass frame, so that the term $\mathcal{L}(\mathbf{r}_{CM}, \dot{\mathbf{r}}_{CM}, t)$ vanishes completely from the total Lagrangian. The second fictitious particle has *reduced mass*:

$$\mu = \frac{m_1 m_2}{m_1 + m_2} \quad (1.176)$$

moves under the potential $V(\mathbf{r})$. Now, we only need to solve this one-body problem.

1.2.4 The Hamiltonian Formalism

Recall the canonical momentum in Lagrangian mechanics:

$$p_i = \frac{\partial \mathcal{L}}{\partial \dot{q}_i}. \quad (1.177)$$

In the Hamiltonian formalism one exchanges the roles of \dot{q} and p : one replaces the Lagrangian $\mathcal{L}(q, \dot{q})$ by a Hamiltonian $\mathcal{H}(q, p)$ which generates the equations of motion, and \dot{q} becomes a derived quantity:

$$\dot{q}_i = \frac{\partial \mathcal{H}}{\partial p_i}. \quad (1.178)$$

But of course the question is, how can we make such a change? It turns out that there exists procedure for effecting such a change, called a *Legendre transformation*. Suppose we have a function $f(x)$ with

$$u(x) = \frac{df(x)}{dx}. \quad (1.179)$$

How do we invert $u(x)$ to get $x(u)$? If we define a function (called the **Legendre transformation**)

$$\boxed{g(u) = x(u)u - f(x(u))} \quad (1.180)$$

then

$$\frac{dg}{du} = \frac{dx}{du}u + x(u) - \frac{df}{dx} \frac{dx}{du} = x(u). \quad (1.181)$$

In going from f to g , we simply exchange the roles of x and u . f and g are called the Legendre transforms of each other. More generally, if $f = f(x_1, \dots, x_n)$, one can eliminate a subset $\{x_i\}$ in favor of the partial derivatives $u_i = \partial f / \partial x_i$ by the transformation

$$g(u_1, \dots, u_j, x_{j+1}, \dots, x_n) = \sum_{i=1}^j u_i x_i - f(x_1, \dots, x_n). \quad (1.182)$$

We can easily check that

$$\frac{\partial g}{\partial u_i} = x_i. \quad (1.183)$$

So, we define the Hamiltonian as

$$\boxed{\mathcal{H}(q, p) = \sum_{i=1}^n p_i \dot{q}_i - \mathcal{L}(q, \dot{q})} \quad (1.184)$$

where the \dot{q} 's are functions of p 's and q 's. Now, we observe that

$$\frac{\partial \mathcal{H}}{\partial p_i} = \frac{\partial}{\partial p_i} \left(\sum_{j=1}^n p_j \dot{q}_j - \mathcal{L}(q, \dot{q}) \right) \quad (1.185)$$

$$= \dot{q}_i + \sum_{j=1}^n p_j \frac{\partial \dot{q}_j}{\partial p_i} - \sum_{j=1}^n \underbrace{\frac{\partial \mathcal{L}}{\partial \dot{q}_j}}_{p_j} \frac{\partial \dot{q}_j}{\partial p_i} \quad (1.186)$$

$$= \dot{q}_i. \quad (1.187)$$

Similarly, we have that

$$\frac{\partial \mathcal{H}}{\partial q_i} = \frac{\partial}{\partial q_i} \left(\sum_{j=1}^n p_j \dot{q}_j - \mathcal{L}(q, \dot{q}) \right) \quad (1.188)$$

$$= \sum_j p_j \frac{\partial \dot{q}_j}{\partial q_i} - \frac{\partial \mathcal{L}}{\partial q_i} - \sum_j \underbrace{\frac{\partial \mathcal{L}}{\partial \dot{q}_j}}_{p_j} \frac{\partial \dot{q}_j}{\partial p_i} \quad (1.189)$$

$$= -\frac{\partial \mathcal{L}}{\partial q_i}. \quad (1.190)$$

We obtain the *Hamilton's canonical equations* by replacing $\partial\mathcal{L}/\partial q_i$ by \dot{p}_i :

$$\boxed{\frac{\partial\mathcal{H}}{\partial p_i} = \dot{q}_i, \quad -\frac{\partial\mathcal{H}}{\partial q_i} = \dot{p}_i} \quad (1.191)$$

We have altogether $2n$ first-order equations in time for a system with n degrees of freedom. Given $(q_i(0), p_i(0))$ it is possible to integrate to find $(q_i(t), p_i(t))$.

Let us take a moment now to compare Lagrangian and Hamiltonian mechanics.

Lagrangian formalism	Hamiltonian formalism
The state of a system with n df's is described by n coordinates (q_1, \dots, q_n) and n velocities $(\dot{q}_1, \dots, \dot{q}_n)$, or in a more compact notation by (q_i, \dot{q}_i) .	The state of a system with n df's is described by n coordinates and n momenta $(q_1, \dots, q_n; p_1, \dots, p_n)$ or, more succinctly, (q, p) .
The state of the system may be represented by a point moving with a definite velocity in an n -dimensional configuration space.	The state of the system may be represented by a point in $2n$ -dimensional phase space with coordinates $(q_1, \dots, q_n; p_1, \dots, p_n)$.
The n coordinates evolve according to n second-order equations.	The $2n$ coordinates and momenta obey $2n$ first-order equations.
For a given \mathcal{L} , several trajectories may pass through a given point in configuration space depending on \dot{q} .	For a given \mathcal{H} only one trajectory passes through a given point in phase space.

So what is \mathcal{H} , physically? We know that \mathcal{L} can be interpreted as $T - V$ if the force is conservative. Let us look at \mathcal{H} for this case. Suppose

$$T = \sum_{i=1}^n \frac{1}{2} m_i \dot{x}_i^2 \quad (1.192)$$

$$p_i = \frac{\partial\mathcal{L}}{\partial\dot{x}_i} = \frac{\partial T}{\partial\dot{x}_i} = m_i \dot{x}_i, \quad (1.193)$$

where the coordinates are Cartesian $q_i = x_i$. This gives

$$\mathcal{H} = \sum_{i=1}^n p_i \dot{x}_i - \mathcal{L}(q, \dot{x}) \quad (1.194)$$

$$= \sum_{i=1}^n m_i \dot{x}_i^2 - \mathcal{L}(q, \dot{x}) \quad (1.195)$$

$$= 2T - \mathcal{L} \quad (1.196)$$

$$= T + V. \quad (1.197)$$

So, we can interpret \mathcal{H} as the total energy (assuming that the force is conservative). Of course, we also assumed that the potential is not velocity-dependent, which does not apply to electromagnetism. In the next section, we will look into electromagnetism in the Hamiltonian scheme.

1.2.5 The Electromagnetic Force in the Hamiltonian Scheme

Recall the Lagrangian for electromagnetism:

$$\mathcal{L}_{EM} = \frac{1}{2}m\mathbf{v} \cdot \mathbf{v} - q\phi + \frac{q}{c}\mathbf{v} \cdot \mathbf{A} \quad (1.198)$$

where

$$\mathbf{p} = m\mathbf{v} + \frac{q\mathbf{A}}{c}. \quad (1.199)$$

In the Hamiltonian scheme, we write

$$\mathcal{H}_{EM} = \mathbf{p} \cdot \mathbf{v} - \mathcal{L}_{EM} \quad (1.200)$$

$$= \left(m\mathbf{v} + \frac{q\mathbf{A}}{c} \right) \cdot \mathbf{v} - \left(\frac{1}{2}m\mathbf{v} \cdot \mathbf{v} - q\phi + \frac{q}{c}\mathbf{v} \cdot \mathbf{A} \right) \quad (1.201)$$

$$= \frac{1}{2}m\mathbf{v} \cdot \mathbf{v} + q\phi \quad (1.202)$$

$$= T + q\phi. \quad (1.203)$$

Wait a minute. What's happened to the vector potential \mathbf{A} ? How can \mathcal{H}_{EM} generate the correct dynamics without knowing \mathbf{A} ? It turns out that T is dependent on \mathbf{p} and \mathbf{A} . Recall that

$$\frac{1}{2}m\mathbf{v} \cdot \mathbf{v} = T = \frac{|\mathbf{p} - q\mathbf{A}/c|^2}{2m}. \quad (1.204)$$

Thus we have

$$\boxed{\mathcal{H}_{EM} = \frac{|\mathbf{p} - q\mathbf{A}/c|^2}{2m} + q\phi} \quad (1.205)$$

1.2.6 Cyclic Coordinates, Poisson Brackets, and Canonical Transformations

Let a Hamiltonian $\mathcal{H}(q, p)$ be given. Suppose that the coordinate q_i is missing in \mathcal{H} , then

$$\dot{p}_i = -\frac{\partial \mathcal{H}}{\partial q_i} = 0, \quad (1.206)$$

which means the canonical momentum p_i is conserved. Now, in many cases there are other quantities that are conserved, such as energy. How do we characterize these in the Hamiltonian formalism? It turns out there is a way to do this.

Suppose $\omega(q, p)$ is some state variable that doesn't depend explicitly on time t . Then we have by the chain rule

$$\frac{d\omega}{dt} = \sum_i \left(\frac{\partial \omega}{\partial q_i} \dot{q}_i + \frac{\partial \omega}{\partial p_i} \dot{p}_i \right) \quad (1.207)$$

$$= \sum_i \left(\frac{\partial \omega}{\partial q_i} \frac{\partial \mathcal{H}}{\partial p_i} - \frac{\partial \omega}{\partial p_i} \frac{\partial \mathcal{H}}{\partial q_i} \right) \quad (1.208)$$

$$\equiv \{\omega, \mathcal{H}\}. \quad (1.209)$$

So, we define the Poisson bracket between two variables $\omega(q, p)$ and $\lambda(q, p)$ to be

$$\boxed{\{\omega, \lambda\} = \sum_i \left(\frac{\partial \omega}{\partial q_i} \frac{\partial \lambda}{\partial p_i} - \frac{\partial \omega}{\partial p_i} \frac{\partial \lambda}{\partial q_i} \right)} \quad (1.210)$$

Let us look at this equation again

$$\frac{d\omega}{dt} = \sum_i \left(\frac{\partial \omega}{\partial q_i} \frac{\partial \mathcal{H}}{\partial p_i} - \frac{\partial \omega}{\partial p_i} \frac{\partial \mathcal{H}}{\partial q_i} \right) = \{\omega, \mathcal{H}\} \quad (1.211)$$

This says that any state variable whose Poisson bracket with \mathcal{H} vanishes is constant in time (i.e., is a conserved quantity).

Poisson brackets have a few important properties and similarities with commutators:

1. $\{\omega, \lambda\} = -\{\lambda, \omega\}$
2. $\{\omega, \lambda + \sigma\} = \{\omega, \lambda\} + \{\omega, \sigma\}$
3. $\{\omega, \lambda\sigma\} = \{\omega, \lambda\}\sigma + \lambda\{\omega, \sigma\}$
4. $\{q_i, q_j\} = \{p_i, p_j\} = 0$
5. $\{q_i, p_j\} = \delta_{ij}$
6. $\dot{q}_i = \{q_i, \mathcal{H}\}$
7. $\dot{p}_i = \{p_i, \mathcal{H}\}$

Please refer to Shankar's book for the sections on *Canonical Transformations*, *Active Transformations*, and *Symmetries and Their Consequences*.

1.3 The Postulates –a General Discussion

1.3.1 The Postulates

Here we introduce the postulates of nonrelativistic quantum mechanics. We consider a system with one df (a particle in one dimension). The straightforward generalization to more particles in higher dimension will be introduced later.

Classical Mechanics	Quantum Mechanics
The state of a particle at any given time is specified by the two variables $x(t)$ and $p(t)$, i.e., as a point in a two-dimensional phase space.	The state of the particle is represented by a vector $ \phi(t)\rangle$ in a Hilbert space.
Every dynamical variable ω is a function of x and p : $\omega = \omega(x, p)$.	<p>The independent variables x and p of classical mechanics are represented by Hermitian operators X and P with the following matrix elements in the eigenbasis of X:</p> $\langle x X x'\rangle = x\delta(x-x') \quad (1.212)$ $\langle x P x'\rangle = -i\hbar\delta'(x-x'). \quad (1.213)$ <p>The operators corresponding to dependent variables $\omega(x, p)$ are given Hermitian operators</p> $\Omega(X, P) = \omega(x \rightarrow X, p \rightarrow P). \quad (1.214)$
If the particle is in a state given by x and p , the measurement of the variable ω will yield a value $\omega(x, p)$. The state will remain unaffected.	If the particle is in a state $ \psi\rangle$, measurement of the variable corresponding to Ω will yield one of the eigenvalues ω with probability $P(\omega) \propto \langle\omega \psi\rangle ^2$. The state of the system will change from $ \psi\rangle$ to $ \omega\rangle$ as a result of the measurement.
The state variables change with time according to Hamilton's equations:	The state vector $ \psi(t)\rangle$ obeys the Schrödinger equation
$\dot{x} = \frac{\partial\mathcal{H}}{\partial p} \quad (1.215)$ $\dot{p} = -\frac{\partial\mathcal{H}}{\partial x} \quad (1.216)$	$i\hbar\frac{d}{dt} \psi(t)\rangle = H \psi(t)\rangle \quad (1.217)$ <p>where $H(X, P) = \mathcal{H}(x \rightarrow X, p \rightarrow P)$ is the quantum Hamiltonian operator and \mathcal{H} is the Hamiltonian for the corresponding classical problem.</p>

1.3.2 Postulates I - III

Collapse of the State Vector

How to Test Quantum Theory

Expectation Value

The Uncertainty

Compatible and Incompatible Variables

The Density Matrix

Generalization to More Degrees of Freedom

1.3.3 The Schrödinger Equation

Setting up

General Approach to the Solution

Choosing a Basis for Solving SE

1.4 Simple Problems in One Dimension

1.4.1 The Free Particle

Time Evolution for the Gaussian Packet

Some General Features of Energy Eigenfunctions

1.4.2 The Particle in a Box

1.4.3 The Continuity Equation for Probability

1.4.4 The Single-Step Potential: A Problem in Scattering

1.4.5 The Double-Slit Experiment

1.4.6 Some Theorems

1.5 The Classical Limit

1.6 The Harmonic Oscillator

1.7 The Path Integral Formulation of Quantum Theory

1.8 The Heisenberg Uncertainty Relation

1.9 Systems with N Degrees of Freedom

1.10 Symmetries and Their Consequences

1.11 Rotational Invariance and Angular Momentum

1.12 The Hydrogen Atom

1.13 Spin

1.14 Additional of Angular Momentum

1.15 Variational and WKB Methods

1.16 Time-Independent Perturbation Theory

1.17 Time-Dependent Perturbation Theory

1.18 Scattering Theory

1.19 The Dirac Equation

1.20 Path Integrals–II

Part 2

QUANTUM OPTICS

Part 3

QUANTUM INFORMATION

Part 4

PROBLEMS & SOLUTIONS

4.1 Quantum Mechanics

4.2 Quantum Optics

4.3 Quantum Information