

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

Huan Q. Bui

Colby College

PHYSICS & MATHEMATICS  
Statistics

Class of 2021

July 7, 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
<b>1 QUANTUM MECHANICS</b>	<b>5</b>
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 Trans- formations . . . . .	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
<b>2</b>	<b>QUANTUM OPTICS</b>	<b>55</b>
2.1	Quantization of the Electromagnetic Field . . . . .	56
2.1.1	Quantization of a Single-mode Field . . . . .	56
2.1.2	Quantum Fluctuations of a Single-mode Field . . . . .	61
2.1.3	Quadrature Operators for a Single-mode Field . . . . .	62
2.1.4	Multimode Fields . . . . .	63
2.2	Emission and Absorption of Radiation by Atoms . . . . .	71
2.2.1	Atom-field interactions . . . . .	71
2.2.2	Interaction of an atom with a classical field . . . . .	73
2.2.3	Interaction of an atom with a quantized field & Spontaneous emission . . . . .	79
2.2.4	The Rabi model . . . . .	84
2.2.5	Fully quantum-mechanical model; the JaynesCummings model . . . . .	86
2.3	Collective Atomic Interactions . . . . .	87
<b>3</b>	<b>QUANTUM INFORMATION</b>	<b>89</b>
<b>4</b>	<b>PROBLEMS &amp; SOLUTIONS</b>	<b>91</b>
4.1	Quantum Mechanics . . . . .	92
4.2	Quantum Optics . . . . .	93
4.3	Quantum Information . . . . .	94

**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  $\Omega$  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  $\Omega$  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  $\Omega$  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  $\Omega, \Lambda$  is defined as

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

In general,  $[\Omega, \Lambda]$  is not zero. Suppose three operators  $\Omega, \Lambda, \Theta$  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  $\Omega$  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  $\Omega$  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  $\Omega_{ji}$  are the matrix elements of  $\Omega$  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  $\Omega_{ji}$ :

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

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  $j^{\text{th}}$  column represent the image of the  $j^{\text{th}}$  basis vector after  $\Omega$  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  $\Omega$  and  $\Lambda$ , 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  $\alpha$

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

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  $\Omega^\dagger$  is *the* adjoint of  $\Omega$ . The relationship between  $\Omega^\dagger$  and  $\Omega$  can be seen in a basis. We consider the matrix elements of  $\Omega^\dagger$  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,  $\Omega^\dagger$  is the conjugate transpose of  $\Omega$ .

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  $\Omega$  is Hermitian  $\iff \Omega = \Omega^\dagger$ .

**Definition 1.1.10.** An operator  $\Omega$  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  $\Omega$  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  $\Omega$ , there exists an ONB comprised entirely of the eigenvectors of  $\Omega$ .

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  $\Omega$  and  $\Lambda$  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  $\Omega$  is Hermitian. In the eigenbasis of  $\Omega$ ,  $\Omega$  is diagonal. This means we can add and/or take powers of  $\Omega$  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  $\omega_i$  are the eigenvalues of  $\Omega$ . We note that each entry in the expression above converges to  $e^{\omega_i}$ .

### Derivatives of Operators with Respect to Parameters

Now, consider some operator  $\Theta(\lambda)$  that depends on a parameter  $\lambda$ . The derivative of  $\Theta$  with respect to  $\lambda$  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  $\Omega$  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  $\Theta$  and  $\Omega$  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  $\Theta$  and  $\Omega$  might not commute, we can't bring  $\Omega$  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,  $\Delta \rightarrow 0$ ,  $g_\Delta$  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).

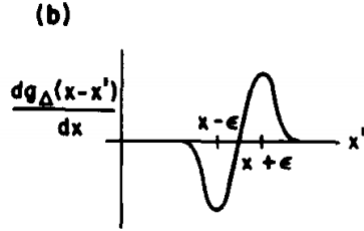
From Gaussian model, we know that the delta function is not only real but also even:

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

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  $g_\Delta$  shrinks,, each bump at  $\pm\epsilon$  will become, up to a scale factor, the  $\delta$  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  $\delta'$  function as

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

In this way, we can describe higher derivatives of  $\delta$ :

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

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

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

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  $r, \theta, \phi$ . 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  $\phi$  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  $\phi$  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 $\omega$ is a function of $x$ and $p$ : $\omega = \omega(x, p)$ .	<p>The independent variables <math>x</math> and <math>p</math> of classical mechanics are represented by Hermitian operators <math>X</math> and <math>P</math> with the following matrix elements in the eigenbasis of <math>X</math>:</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 <math>\omega(x, p)</math> 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 $\omega$ 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 $\Omega$ will yield one of the eigenvalues $\omega$ 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 <math>H(X, P) = \mathcal{H}(x \rightarrow X, p \rightarrow P)</math> is the quantum Hamiltonian operator and <math>\mathcal{H}</math> 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**

## 2.1 Quantization of the Electromagnetic Field

### 2.1.1 Quantization of a Single-mode Field

Consider the following simple yet important scenario: a radiation field confined to a one-dimensional cavity along the  $z$ -axis with perfectly conducting walls at  $z = 0$  and  $z = L$ .



Suppose that the field is a standing wave which vanishes on the boundary and is polarized in the  $x$ -direction. We also assume there is no source of radiation such as a charge or a current. The electric field then can be written as

$$\mathbf{E}(\mathbf{r}, t) = \mathbf{e}_x E_x(z, t) \quad (2.1)$$

where  $\mathbf{e}_x$  is the polarization vector. The Maxwell's equations without source are

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \quad (2.2)$$

$$\nabla \times \mathbf{B} = \mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} \quad (2.3)$$

$$\nabla \cdot \mathbf{E} = 0 \quad (2.4)$$

$$\nabla \cdot \mathbf{B} = 0. \quad (2.5)$$

A single-mode field satisfying Maxwell's equations and the boundary conditions is given by

$$E_x(z, t) = \left( \frac{2\omega^2}{V\epsilon_0} \right)^{1/2} q(t) \sin(kz). \quad (2.6)$$

$\omega$  is the frequency,  $k = \omega/c$  is the wavenumber. The allowed frequencies are

$$\omega_m = c \frac{m\pi}{L}, \quad m = 1, 2, \dots \quad (2.7)$$

$V$  is the effective volume of the cavity and  $q(t)$  is the time-dependent factor having the dimension of length. The magnetic field, from the second Maxwell equation, is

$$\mathbf{B}(\mathbf{r}, t) = \mathbf{e}_y B_y(z, t) \quad (2.8)$$



where

$$B_y(z, t) = \left( \frac{\mu_0 \epsilon_0}{k} \right) \left( \frac{2\omega^2}{V\epsilon_0} \right)^{1/2} \dot{q}(t) \cos(kz). \quad (2.9)$$

It will become clear that  $q(t)$  plays the role of canonical position, and  $\dot{q}(t) = p(t)$  plays the role of canonical momentum. The classical field energy (or the Hamiltonian  $\mathcal{H}$ ) of this single-mode field is then given by

$$\mathcal{H} = \frac{1}{2} \int dV \left[ \epsilon_0 \mathbf{E}^2(\mathbf{r}, t) + \frac{1}{\mu_0} \mathbf{B}^2(\mathbf{r}, t) \right] \quad (2.10)$$

$$= \frac{1}{2} \int dV \left[ \epsilon_0 E_x^2(z, t) + \frac{1}{\mu_0} B_y^2(z, t) \right] \quad (2.11)$$

It turns out that (for  $p(t) = \dot{q}(t)$ ), we have

$$\boxed{\mathcal{H} = \frac{1}{2} (p^2 + \omega^2 q^2)} \quad (2.12)$$

Now, we notice that we have the canonical variables  $q$  and  $p$ . When we promote these variables to operators (so we can quantize the fields), these operators must satisfy the canonical commutation relation

$$[\hat{q}, \hat{p}] = i\hbar \hat{\mathcal{L}} \quad (2.13)$$

We rewrite the fields and the Hamiltonian as operators:

$$\hat{E}_x(z, t) = \left( \frac{2\omega^2}{V\epsilon_0} \right)^{1/2} \hat{q}(t) \sin(kz) \quad (2.14)$$

$$\hat{B}_y(z, t) = \left( \frac{\mu_0 \epsilon_0}{k} \right) \left( \frac{2\omega^2}{V\epsilon_0} \right)^{1/2} \hat{p}(t) \cos(kz). \quad (2.15)$$

$$\hat{\mathcal{H}} = \frac{1}{2} (\hat{p}^2 + \omega^2 \hat{q}^2). \quad (2.16)$$

The operators  $\hat{q}$  and  $\hat{p}$  correspond to observables because they are Hermitian. We can also introduce creation ( $\hat{a}^\dagger$ ) and annihilation ( $\hat{a}$ ) operators through combining  $\hat{q}$  and  $\hat{p}$ :

$$\hat{a} = (2\hbar\omega)^{-1/2} (\omega\hat{q} + i\hat{p}) \quad (2.17)$$

$$\hat{a}^\dagger = (2\hbar\omega)^{-1/2} (\omega\hat{q} - i\hat{p}) \quad (2.18)$$

It is easy to show that the electric and magnetic field operators then become

$$\hat{E}_x(z, t) = \mathcal{E}_0 (\hat{a} + \hat{a}^\dagger) \sin(kz) \quad (2.19)$$

$$\hat{B}_y(z, t) = \mathcal{B}_0 \frac{1}{i} (\hat{a} - \hat{a}^\dagger) \cos(kz) \quad (2.20)$$

where

$$\mathcal{E}_0 = \sqrt{\frac{\hbar\omega}{\epsilon_0 V}}, \quad \mathcal{B}_0 = \left(\frac{\mu_0}{k}\right) \sqrt{\frac{\epsilon_0 \hbar \omega^3}{V}} \quad (2.21)$$

represents respectively the electric and magnetic field “per photon.” We notice that

$$[\hat{a}, \hat{a}^\dagger] = \hat{\mathcal{I}} \quad (2.22)$$

so the Hamiltonian takes the form

$$\boxed{\hat{\mathcal{H}} = \hbar\omega \left( \hat{a}^\dagger \hat{a} + \frac{1}{2} \right)} \quad (2.23)$$

Now, what is the time dependent of these new operators? For an arbitrary operator  $\hat{O}$ , Heisenberg equation reads

$$\frac{d\hat{O}}{dt} = \frac{i}{\hbar} [\hat{H}, \hat{O}]. \quad (2.24)$$

For the annihilation operator, this becomes

$$\frac{d\hat{a}}{dt} = \frac{i}{\hbar} \left[ \hbar\omega \left( \hat{a}^\dagger \hat{a} + \frac{1}{2} \right), \hat{a} \right] = -i\omega \hat{a}. \quad (2.25)$$

Thus,

$$\boxed{\hat{a}(t) = \hat{a}(0)e^{i\omega t}} \quad (2.26)$$

Similarly for the creation operator, we have

$$\boxed{\hat{a}^\dagger(t) = \hat{a}^\dagger(0)e^{i\omega t}} \quad (2.27)$$

Now, let  $|n\rangle$  denote an energy eigenstate of the single mode field corresponding to energy  $E_n$ , then

$$\hat{\mathcal{H}} |n\rangle = \hbar\omega \left( \hat{a}^\dagger \hat{a} + \frac{1}{2} \right) |n\rangle = E_n |n\rangle. \quad (2.28)$$

Applying the creation operator  $\hat{a}^\dagger$  from the left we have

$$\hbar\omega \left( \hat{a}^\dagger \hat{a}^\dagger \hat{a} + \frac{1}{2} \hat{a}^\dagger \right) |n\rangle = E_n \hat{a}^\dagger |n\rangle. \quad (2.29)$$

Using the commutation relation, we can write

$$\hat{a}\hat{a}^\dagger = \hat{\mathcal{I}} + \hat{a}^\dagger \hat{a}. \quad (2.30)$$

Thus,

$$E_n \hat{a}^\dagger |n\rangle = \hbar\omega \left( \hat{a}^\dagger [-\hat{\mathcal{I}} + \hat{a}\hat{a}^\dagger] + \frac{1}{2}\hat{a}^\dagger \right) |n\rangle \quad (2.31)$$

$$= \hbar\omega \left( -\hat{a}^\dagger + \hat{a}^\dagger \hat{a}\hat{a}^\dagger + \frac{1}{2}\hat{a}^\dagger \right) |n\rangle \quad (2.32)$$

$$= \hbar\omega \left( -\hat{\mathcal{I}} + \hat{a}^\dagger \hat{a} + \frac{1}{2} \right) \hat{a}^\dagger |n\rangle. \quad (2.33)$$

This says we have a new eigen-equation:

$$(E_n + \hbar\omega) \hat{a}^\dagger |n\rangle = \hbar\omega \left( \hat{a}^\dagger \hat{a} + \frac{1}{2} \right) \hat{a}^\dagger |n\rangle \quad (2.34)$$

or equivalently,

$$\boxed{\hat{H} [\hat{a}^\dagger |n\rangle] = (E_n + \hbar\omega) [\hat{a}^\dagger |n\rangle]} \quad (2.35)$$

This is the eigenvalue problem for the eigenstate  $\hat{a}^\dagger |n\rangle$  with the energy eigenvalue  $E_n + \hbar\omega$ . It is clear now why  $\hat{a}^\dagger$  is called the creation operator: it creates a “quantum” energy  $\hbar\omega$ . On the other hand, we can also show that

$$\boxed{\hat{H} [\hat{a} |n\rangle] = (E_n - \hbar\omega) [\hat{a} |n\rangle]} \quad (2.36)$$

It is clear that the operator  $\hat{a}$  annihilates a quantum of energy. Now, because for the least energetic eigenstate

$$\hat{a} |0\rangle = 0, \quad (2.37)$$

the Hamiltonian of the annihilation of the least energetic eigenstate is

$$\hat{\mathcal{H}} [\hat{a} |0\rangle] = (E_0 - \hbar\omega) [\hat{a} |0\rangle] = 0. \quad (2.38)$$

Therefore,

$$\hat{\mathcal{H}} |0\rangle = \hbar\omega \left( \hat{a}^\dagger \hat{a} + \frac{1}{2} \right) |0\rangle = \frac{1}{2}\hbar\omega |0\rangle = E_0 |0\rangle. \quad (2.39)$$

Or equivalently, the zero-point energy is

$$E_0 = \hbar\omega \quad (2.40)$$

and the energy eigenvalues  $E_{n+1} = E_n + \hbar\omega$  are

$$\boxed{E_n = \hbar \left( n + \frac{1}{2} \right)} \quad (2.41)$$

Note that this is just like the energy levels for the harmonic oscillator in an infinite well. Now, we observe that

$$\hat{\mathcal{H}}|n\rangle = \hbar\omega \left( \hat{a}^\dagger \hat{a} + \frac{1}{2} \right) |n\rangle = E_n |n\rangle = \hbar\omega \left( n + \frac{1}{2} \right) |n\rangle, \quad (2.42)$$

i.e.,

$$\hat{a}^\dagger \hat{a} |n\rangle = n |n\rangle. \quad (2.43)$$

It makes sense to call  $\hat{a}^\dagger \hat{a}$  the “number operator”  $\hat{n}$ , so that

$$\hat{n} |n\rangle = n |n\rangle. \quad (2.44)$$

Next, the eigenstates must be normalized according to  $\langle n|n\rangle = 1$ . For the state  $\hat{a}|n\rangle$  we have

$$\hat{a} |n\rangle = c_n |n-1\rangle. \quad (2.45)$$

Taking the inner product with itself we have

$$(\langle n| \hat{a}^\dagger) (\hat{a} |n\rangle) = \langle n| \hat{a}^\dagger \hat{a} |n\rangle \quad (2.46)$$

$$= \langle n| \hat{n} |n\rangle \quad (2.47)$$

$$= n \langle n|n\rangle \quad (2.48)$$

$$= \langle n-1| c_n^* c_n |n-1\rangle. \quad (2.49)$$

This says  $n = |c_n|^2$ . So we have

$$\boxed{\hat{a} |n\rangle = \sqrt{n} |n-1\rangle} \quad (2.50)$$

Similarly,

$$\boxed{\hat{a}^\dagger |n\rangle = \sqrt{n+1} |n+1\rangle} \quad (2.51)$$

This says that  $|n\rangle$  can be generated from the  $|0\rangle$  state by applying the creation operator  $\hat{a}^\dagger$  appropriately many times with the appropriate normalization:

$$\boxed{|n\rangle = \frac{(\hat{a}^\dagger)^n}{\sqrt{n!}} |0\rangle} \quad (2.52)$$

Next, because the eigenstates are orthonormal ( $\langle n'|n\rangle = \delta_{n'n}$  and  $\langle n|n\rangle = 1$ ) and form a complete set, i.e.,

$$\sum_{n=0}^{\infty} |n\rangle \langle n| = 1, \quad (2.53)$$

the only non-vanishing matrix elements of the annihilation and creation operators are

$$\langle n-1| \hat{a} |n\rangle = \sqrt{n} \langle n-1|n-1\rangle = \sqrt{n} \quad (2.54)$$

$$\langle n+1| \hat{a}^\dagger |n\rangle = \sqrt{n+1} \langle n+1|n+1\rangle = \sqrt{n+1} \quad (2.55)$$

### 2.1.2 Quantum Fluctuations of a Single-mode Field

Recall the electric field operator

$$\hat{E}_x(z, t) = \mathcal{E}_0(\hat{a} + \hat{a}^\dagger) \sin(kz). \quad (2.56)$$

The mean field is zero:

$$\langle n | \hat{E}_x(z, t) | n \rangle = \mathcal{E}_0 \sin(kz) [\underbrace{\langle n | \hat{a} | n \rangle}_0 + \underbrace{\langle n | \hat{a}^\dagger | n \rangle}_0] = 0 \quad (2.57)$$

by orthonormality. However, the mean of the square of the field is not zero:

$$\langle n | \hat{E}_x^2(z, t) | n \rangle = \mathcal{E}_0^2 \sin^2(kz) \langle n | (\hat{a}^\dagger)^2 + \hat{a}^2 + \hat{a}^\dagger \hat{a} + \hat{a} \hat{a}^\dagger | n \rangle \quad (2.58)$$

$$= \mathcal{E}_0^2 \sin^2(kz) \langle n | 2\hat{a}^\dagger \hat{a} + 1 | n \rangle \quad (2.59)$$

$$= 2\mathcal{E}_0^2 \sin^2(kz) \left( n + \frac{1}{2} \right). \quad (2.60)$$

So, the fluctuations in the field for the state  $|n\rangle$  may be characterized by the variance

$$\text{Var}[\hat{E}_x(z, t)] = \text{E}[\hat{E}_x^2(z, t)] - \text{E}^2[\hat{E}_x(z, t)] \quad (2.61)$$

$$= 2\mathcal{E}_0^2 \sin^2(kz) \left( n + \frac{1}{2} \right). \quad (2.62)$$

And so the uncertainty of the field for the state  $|n\rangle$  is

$$\sigma[\hat{E}_x(z, t)] = \sqrt{2\mathcal{E}_0^2 \sin^2(kz)} \sqrt{n + \frac{1}{2}} \quad (2.63)$$

Now, note that even at  $n = 0$ , the field still has fluctuations. These are called “vacuum fluctuations.” Now the number states  $|n\rangle$  are taken to represent a state of the field containing  $n$  photons. Yet as we have seen, the average field is zero. This is all in accordance with the uncertainty principle because the number operator  $\hat{n} = \hat{a}^\dagger \hat{a}$  does not commute with the electric field:

$$[\hat{n}, \hat{E}_x] = \mathcal{E}_0 \sin(kz) (\hat{a}^\dagger - \hat{a}). \quad (2.64)$$

Thus  $\hat{n}$  and  $\hat{E}_x$  are complementary quantities for which their respective uncertainties obey the inequality

$$\sigma[n] \cdot \sigma[E_x] \geq \frac{1}{2} \mathcal{E}_0 |\sin(kz)| |\langle \hat{a}^\dagger - \hat{a} \rangle| \quad (2.65)$$

This says that if the field is accurately known, then the number of photons is uncertain.

### 2.1.3 Quadrature Operators for a Single-mode Field

When we explicitly include the time dependence of the electric field operator we have

$$\hat{E}_x = \mathcal{E}_0(\hat{a}e^{-i\omega t} + \hat{a}^\dagger e^{i\omega t}) \sin(kz) \quad (2.66)$$

where

$$\hat{a}(0) \equiv \hat{a} \quad \hat{a}^\dagger(0) \equiv \hat{a}^\dagger. \quad (2.67)$$

Next, we define quadrature operators:

$$\hat{X}_1 = \frac{1}{2}(\hat{a} + \hat{a}^\dagger) \quad (2.68)$$

$$\hat{X}_2 = \frac{1}{2i}(\hat{a} - \hat{a}^\dagger). \quad (2.69)$$

With these definitions, we rewrite the time-dependent electric field operator as

$$\hat{E}_x(t) = 2\mathcal{E}_0 \sin(kz)[\hat{X}_1 \cos(\omega t) + \hat{X}_2 \sin(\omega t)]. \quad (2.70)$$

We observe that  $\hat{X}_1$  and  $\hat{X}_2$  are associated with field amplitudes oscillating  $\pi/2$  out of phase. Essentially,  $\hat{X}_1$  and  $\hat{X}_2$  are position and momentum operators, only scaled to be dimensionless. These operators satisfy the commutation relation

$$[\hat{X}_1, \hat{X}_2] = \frac{i}{2}, \quad (2.71)$$

which one can easily check. Next, by definition we have

$$\langle n | \hat{X}_1 | n \rangle = \langle n | \hat{X}_2 | n \rangle = 0. \quad (2.72)$$

However,

$$\langle n | \hat{X}_1^2 | n \rangle = \frac{1}{4} \langle n | \hat{a}^2 + (\hat{a}^\dagger)^2 + \hat{a}^\dagger \hat{a} + \hat{a} \hat{a}^\dagger | n \rangle \quad (2.73)$$

$$= \frac{1}{4} \langle n | 2\hat{a}^\dagger \hat{a} + 1 | n \rangle \quad (2.74)$$

$$= \frac{1}{4}(2n + 1). \quad (2.75)$$

Similarly,

$$\langle n | \hat{X}_2^2 | n \rangle = \frac{1}{4}(2n + 1). \quad (2.76)$$

Thus for a number state, the uncertainties in both quadratures are the same and furthermore the vacuum state ( $n = 0$ ) minimizes the uncertainty product since

$$\text{Var}[\hat{X}_1] = \text{Var}[\hat{X}_2] = \frac{1}{4}. \quad (2.77)$$

From this section, we learn that the quanta of the single-mode cavity field are the excitations of energy in discrete amounts of  $\hbar\omega$ . These quantas (or photons) are not localized particles but rather are spread out over the entire mode volume.

### 2.1.4 Multimode Fields

Now we generalize our results for the single-mode field to multi-mode fields. We again assume that there is no source of radiation. Let the vector potential  $\mathbf{A}(\mathbf{r}, t)$  be given such that

$$\nabla \cdot \mathbf{A}(\mathbf{r}, t) = 0. \quad (2.78)$$

The vector potential gives the electric and magnetic fields by

$$\mathbf{E}(\mathbf{r}, t) = -\frac{\partial \mathbf{A}(\mathbf{r}, t)}{\partial t} \quad (2.79)$$

$$\mathbf{B}(\mathbf{r}, t) = \nabla \times \mathbf{A}(\mathbf{r}, t). \quad (2.80)$$

The vector potential satisfies the wave equation

$$\nabla^2 \mathbf{A}(\mathbf{r}, t) - \frac{1}{c^2} \frac{\partial^2 \mathbf{A}(\mathbf{r}, t)}{\partial t^2} = 0. \quad (2.81)$$

We now assume free space is a cubic cavity with perfectly reflecting walls of length  $L$ . We assume that  $L \rightarrow \infty$ , i.e.,  $L$  is much larger than the wavelength of the field. This allows us to impose boundary conditions:

$$e^{ik_x x} = e^{ik_x(x+L)} \quad (2.82)$$

$$e^{ik_y y} = e^{ik_y(y+L)} \quad (2.83)$$

$$e^{ik_z z} = e^{ik_z(z+L)}. \quad (2.84)$$

It follows that

$$k_x = \left(\frac{2\pi}{L}\right) m_x, \quad m_x = 0, \pm 1, \pm 2, \dots \quad (2.85)$$

$$k_y = \left(\frac{2\pi}{L}\right) m_y, \quad m_y = 0, \pm 1, \pm 2, \dots \quad (2.86)$$

$$k_z = \left(\frac{2\pi}{L}\right) m_z, \quad m_z = 0, \pm 1, \pm 2, \dots \quad (2.87)$$

The wave vector is then

$$\mathbf{k} = \frac{2\pi}{L} \begin{bmatrix} m_x \\ m_y \\ m_z \end{bmatrix}. \quad (2.88)$$

The magnitude of this vector is related to the frequency  $\omega_k$  according to  $k = \omega_k/c$ . A set of integers  $(m_x, m_y, m_z)$  specifies a normal mode of the field (apart from polarization), the number of modes being infinite but denumerable. The total number of modes in the intervals  $\Delta m_x, \Delta m_y, \Delta m_z$  is

$$\Delta m = \Delta m_x \Delta m_y \Delta m_z = 2 \left(\frac{L}{2\pi}\right)^3 \Delta k_x \Delta k_y \Delta k_z \quad (2.89)$$

where the factor of 2 takes into account two independent modes of polarization. With the assumption that  $L$  is very large compared to the wavelengths, this becomes

$$dm = 2 \left( \frac{V}{8\pi^3} \right) dk_x dk_y dk_z. \quad (2.90)$$

In spherical coordinates,

$$\mathbf{k} = k(\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta), \quad (2.91)$$

and

$$dm = 2 \left( \frac{V}{8\pi^3} \right) k^2 dk d\Omega \quad (2.92)$$

where

$$d\Omega = \sin \theta d\theta d\phi. \quad (2.93)$$

With  $k = \omega_k/c$ ,

$$dm = 2 \left( \frac{V}{8\pi^3} \right) \frac{\omega_k^2}{c^3} d\omega_k d\Omega. \quad (2.94)$$

Integrating over all solid angles gives  $\int d\Omega = 4\pi$ , so the number of modes in all directions from  $k \rightarrow k + dk$  is

$$2 \left( \frac{V}{8\pi^3} \right) k^2 \cdot (4\pi) dk = V \frac{k^2}{\pi^2} dk = V \rho_k dk \quad (2.95)$$

where  $\rho_k = k^2/\pi^2$  and  $\rho_k dk$  is the number of modes per unit volume. Integrating in the same fashion but with  $\omega_k$  gives the number of modes in all directions from  $\omega_k \rightarrow \omega_k + d\omega_k$ :

$$V \left( \frac{\omega_k^2}{\pi^2 c^3} \right) d\omega_k = V \rho(\omega_k) d\omega_k \quad (2.96)$$

where now  $\rho(\omega_k)$  is also the mode density.

The vector potential can be written as a superposition on plane waves in the form

$$\mathbf{A}(\mathbf{r}, t) = \sum_{\mathbf{k}, s} \mathbf{e}_{\mathbf{k}s} [A_{\mathbf{k}s}(t) e^{i\mathbf{k} \cdot \mathbf{r}} + A_{\mathbf{k}s}^* e^{-i\mathbf{k} \cdot \mathbf{r}}] \quad (2.97)$$

where  $A_{\mathbf{k}s}$  is the complex amplitude of the field and  $\mathbf{e}_{\mathbf{k}s}$  is a real polarization vector. The sum over  $k$  simply means the sum over the set of integers  $(m_x, m_y, m_z)$  and the sum over  $s$  is the sum over the two independent polarizations. These polarizations must be orthogonal:

$$\mathbf{e}_{\mathbf{k}s} \cdot \mathbf{e}_{\mathbf{k}s'} = 0 \quad (2.98)$$



and the *transversality* condition must be satisfied:

$$\mathbf{k} \cdot \mathbf{e}_{\mathbf{k}s} = 0, \quad (2.99)$$

which is just says that the polarization is orthogonal to the propagation. The polarization vectors  $\mathbf{e}_{\mathbf{k}1}$  and  $\mathbf{e}_{\mathbf{k}2}$  form a right-handed system:

$$\mathbf{e}_{\mathbf{k}1} \times \mathbf{e}_{\mathbf{k}2} = \frac{\mathbf{k}}{|\mathbf{k}|} = \kappa. \quad (2.100)$$

In free space, we make the transition from “discrete” modes to a continuum of modes, thus the sum becomes an integral:

$$\sum_{\mathbf{k}} \rightarrow \frac{V}{\pi^2} \int k^2 dk. \quad (2.101)$$

From (2.78) and (2.81) and  $\omega_k = kc$ , we get

$$\frac{d^2 A_{\mathbf{k}s}}{dt^2} + \omega_k^2 A_{\mathbf{k}s} = 0. \quad (2.102)$$

The solution to this is of course

$$A_{\mathbf{k}s}(t) = A_{\mathbf{k}s}(0)e^{-i\omega_k t} = A_{\mathbf{k}s}e^{-i\omega_k t}. \quad (2.103)$$

From this and (2.79) we can derive the electric and magnetic fields:

$$\mathbf{E}(\mathbf{r}, t) = i \sum_{\mathbf{k}, s} \omega_k \mathbf{e}_{\mathbf{k}s} \left[ A_{\mathbf{k}s} e^{i(\mathbf{k} \cdot \mathbf{r} - \omega_k t)} - A_{\mathbf{k}s}^* e^{-i(\mathbf{k} \cdot \mathbf{r} - \omega_k t)} \right] \quad (2.104)$$

$$\mathbf{B}(\mathbf{r}, t) = \frac{i}{c} \sum_{\mathbf{k}, s} \omega_k (\kappa \times \hat{\mathbf{e}}_{\mathbf{k}s}) \left[ A_{\mathbf{k}s} e^{i(\mathbf{k} \cdot \mathbf{r} - \omega_k t)} - A_{\mathbf{k}s}^* e^{-i(\mathbf{k} \cdot \mathbf{r} - \omega_k t)} \right]. \quad (2.105)$$

The electromagnetic Hamiltonian is given by

$$\mathcal{H} = \frac{1}{2} \int_V \left( \epsilon_0 \mathbf{E} \cdot \mathbf{E} + \frac{1}{\mu_0} \mathbf{B} \cdot \mathbf{B} \right) dV. \quad (2.106)$$

The periodic boundary conditions result in

$$\int_0^L e^{\pm i k_x x} dx = \begin{cases} L, & k_x = 0 \\ 0, & k_x \neq 0 \end{cases} \quad (2.107)$$

same for  $y$  and  $z$ . Collectively we write this as

$$\int_V e^{\pm i(\mathbf{k} - \mathbf{k}') \cdot \mathbf{r}} dV = \delta_{\mathbf{k}\mathbf{k}'} V. \quad (2.108)$$

So the contribution to the Hamiltonian from the electric field is

$$\frac{1}{2} \int_V \epsilon_0 \mathbf{E} \cdot \mathbf{E} dV = \epsilon_0 V \sum_{\mathbf{k}s} \omega_k^2 A_{\mathbf{k}s}(t) A_{\mathbf{k}s}^*(t) \quad (2.109)$$

$$- \frac{1}{2} \epsilon_0 V \underbrace{\sum_{\mathbf{k}ss'} \omega_k^2 \mathbf{e}_{\mathbf{k}s} \cdot \mathbf{e}_{-\mathbf{k}s'} [A_{\mathbf{k}s}(t) A_{\mathbf{k}s'}(t) + A_{\mathbf{k}s}^*(t) A_{-\mathbf{k}s'}^*(t)]}_{R}. \quad (2.110)$$

To obtain the magnetic contribution we use the identity

$$(\mathbf{A} \times \mathbf{B}) \cdot (\mathbf{C} \times \mathbf{D}) = (\mathbf{A} \cdot \mathbf{C})(\mathbf{B} \cdot \mathbf{D}) - (\mathbf{A} \cdot \mathbf{D})(\mathbf{B} \cdot \mathbf{C}). \quad (2.111)$$

This gives

$$(\boldsymbol{\kappa} \times \mathbf{e}_{\mathbf{k}s}) \cdot (\boldsymbol{\kappa} \times \mathbf{e}_{\mathbf{k}s'}) = \delta_{ss'} \quad (2.112)$$

$$(\boldsymbol{\kappa} \times \mathbf{e}_{\mathbf{k}s}) \cdot (-\boldsymbol{\kappa} \times \mathbf{e}_{-\mathbf{k}s'}) = -\mathbf{e}_{\mathbf{k}s} \cdot \mathbf{e}_{-\mathbf{k}s'}. \quad (2.113)$$

With this we get the contribution from the magnetic field

$$\frac{1}{2} \int_V \frac{1}{\mu_0} \mathbf{B} \cdot \mathbf{B} dV = \epsilon_0 V \sum_{\mathbf{k}s} \omega_k^2 A_{\mathbf{k}s}(t) A_{\mathbf{k}s}^*(t) + R. \quad (2.114)$$

Thus the Hamiltonian is

$$\mathcal{H} = 2\epsilon_0 V \sum_{\mathbf{k}s} \omega_k^2 A_{\mathbf{k}s}(t) A_{\mathbf{k}s}^*(t) \quad (2.115)$$

$$= 2\epsilon_0 V \sum_{\mathbf{k}s} \omega_k^2 A_{\mathbf{k}s}(0) e^{-i\omega_k t} A_{\mathbf{k}s}^*(0) e^{i\omega_k t} \quad (2.116)$$

$$= 2\epsilon_0 V \sum_{\mathbf{k}s} \omega_k^2 A_{\mathbf{k}s} A_{\mathbf{k}s}^*. \quad (2.117)$$

To quantize the field, the canonical variables  $q_{\mathbf{k}s}$  and  $p_{\mathbf{k}s}$  must be introduced:

$$A_{\mathbf{k}s} = \frac{1}{2\omega_k \sqrt{\epsilon_0 V}} [\omega_k q_{\mathbf{k}s} + i p_{\mathbf{k}s}] \quad (2.118)$$

$$A_{\mathbf{k}s}^* = \frac{1}{2\omega_k \sqrt{\epsilon_0 V}} [\omega_k q_{\mathbf{k}s} - i p_{\mathbf{k}s}] \quad (2.119)$$

such that when we can re-write the Hamiltonian as

$$\mathcal{H} = \frac{1}{2} \sum_{\mathbf{k}s} (p_{\mathbf{k}s}^2 + \omega_k^2 q_{\mathbf{k}s}^2). \quad (2.120)$$

To quantize the field, we promote these canonical variables to operators:

$$\boxed{\hat{\mathcal{H}} = \frac{1}{2} \sum_{\mathbf{k}s} (\hat{p}_{\mathbf{k}s}^2 + \omega_k^2 \hat{q}_{\mathbf{k}s}^2)} \quad (2.121)$$

where

$$[\hat{q}_{\mathbf{k}s}, \hat{q}_{\mathbf{k}'s'}] = 0 = [\hat{p}_{\mathbf{k}s}, \hat{p}_{\mathbf{k}'s'}] \quad (2.122)$$

$$[\hat{q}_{\mathbf{k}s}, \hat{p}_{\mathbf{k}'s'}] = i\hbar\delta_{\mathbf{k}\mathbf{k}'}\delta_{ss'}. \quad (2.123)$$

The annihilation and creation operators are defined in a similar fashion as in the case of a single-mode field:

$$\hat{a}_{\mathbf{k}s} = (2\hbar\omega_k)^{-1/2}(\omega_k\hat{q}_{\mathbf{k}s} + i\hat{p}_{\mathbf{k}s}) \quad (2.124)$$

$$\hat{a}_{\mathbf{k}s}^\dagger = (2\hbar\omega_k)^{-1/2}(\omega_k\hat{q}_{\mathbf{k}s} - i\hat{p}_{\mathbf{k}s}) \quad (2.125)$$

which satisfy

$$[\hat{a}_{\mathbf{k}s}, \hat{a}_{\mathbf{k}'s'}] = 0 = [\hat{a}_{\mathbf{k}s}^\dagger, \hat{a}_{\mathbf{k}'s'}^\dagger] \quad (2.126)$$

$$[\hat{a}_{\mathbf{k}s}, \hat{a}_{\mathbf{k}'s'}^\dagger] = \delta_{\mathbf{k}\mathbf{k}'}\delta_{ss'}. \quad (2.127)$$

Finally, we define a similar operator for the mode  $\mathbf{k}s$  as the “number” operator before

$$\hat{n}_{\mathbf{k}s} = \hat{a}_{\mathbf{k}s}^\dagger \hat{a}_{\mathbf{k}s}. \quad (2.128)$$

With these, the Hamiltonian (operator) can be written as

$$\hat{\mathcal{H}} = \sum_{\mathbf{k}s} \hbar\omega_k \left( \hat{a}_{\mathbf{k}s}^\dagger \hat{a}_{\mathbf{k}s} + \frac{1}{2} \right) = \sum_{\mathbf{k}s} \omega_k \left( \hat{n}_{\mathbf{k}s} + \frac{1}{2} \right). \quad (2.129)$$

For simplicity, we associate a  $j^{\text{th}}$  mode with  $\mathbf{k}_j s_j$ . We let

$$\hat{a}_{\mathbf{k}_j s_j} \equiv \hat{a}_j \quad (2.130)$$

$$\hat{a}_{\mathbf{k}_j s_j}^\dagger \equiv \hat{a}_j^\dagger \quad (2.131)$$

$$\hat{n}_{\mathbf{k}_j s_j} \equiv \hat{n}_j. \quad (2.132)$$

Now we can write the Hamiltonian as

$$\boxed{\hat{\mathcal{H}} = \sum_j \hbar\omega_j \left( \hat{n}_j + \frac{1}{2} \right)} \quad (2.133)$$

A multimode photon number state is a (tensor) product of the number states  $|n_{\mathbf{k}s}\rangle \equiv |n_j\rangle$  of all the modes which we write as

$$|n_1\rangle |n_2\rangle |n_3\rangle \cdots \equiv |n_1 n_2 n_3 \dots\rangle = |\{n_j\}\rangle. \quad (2.134)$$

This is an eigenstate of the Hamiltonian:

$$\hat{\mathcal{H}} |\{n_j\}\rangle = E |\{n_j\}\rangle. \quad (2.135)$$

We can show that energy eigenvalue is

$$E = \sum_j \hbar \omega_j \left( n_j + \frac{1}{2} \right) \quad (2.136)$$

Because the Hamiltonian is Hermitian (as always) the eigenstates are orthogonal:

$$\langle n_1 n_2 n_3 \dots | n'_1 n'_2 n'_3 \dots \rangle = \delta_{n_1 n'_1} \delta_{n_2 n'_2} \delta_{n_3 n'_3} \dots \quad (2.137)$$

The actions of the annihilation and creation operator are as expected and similar to what we have seen before:

$$\textbf{Annihilation: } \hat{a}_j |n_1, n_2, \dots, n_j, \dots\rangle = \sqrt{n_j} |n_1, n_2, \dots, n_j - 1, \dots\rangle \quad (2.138)$$

$$\textbf{Creation: } \hat{a}_j^\dagger |n_1, n_2, \dots, n_j, \dots\rangle = \sqrt{n_j + 1} |n_1, n_2, \dots, n_j + 1, \dots\rangle \quad (2.139)$$

The vacuum multimode is denoted

$$|\{0\}\rangle = |0_1 0_2 0_3 \dots 0_j \dots\rangle \quad (2.140)$$

for which

$$\hat{a}_j |\{0\}\rangle = 0 \quad \text{for all } j. \quad (2.141)$$

Once again, like in the single-mode case, a multimode number state can be generated by applying the creation operator. Any number state can be written as some creation operator applied to the multimode vacuum:

$$|\{n_j\}\rangle = \prod_j \frac{(\hat{a}_j^\dagger)^{n_j}}{\sqrt{n_j!}} |\{0\}\rangle. \quad (2.142)$$

Now, we notice that the creation/annihilation operators  $\hat{a}_{\mathbf{k}s}/\hat{a}_{\mathbf{k}s}^\dagger$  are very much like the vector potential amplitude  $A_{\mathbf{k}s}$ . Indeed, upon quantization of the field, the amplitudes can be promoted to operators as well:

$$\hat{A}_{\mathbf{k}s} = \left( \frac{\hbar}{2\omega_k \epsilon_0 V} \right)^{\frac{1}{2}} \hat{a}_{\mathbf{k}s}. \quad (2.143)$$

With this, we are able to rewrite the vector potential, electric, and magnetic fields in terms of the creation and annihilation operators:

$$\hat{\mathbf{A}}(\mathbf{r}, t) = \sum_{\mathbf{k}s} \left( \frac{\hbar}{2\omega_k \epsilon_0 V} \right)^{\frac{1}{2}} \mathbf{e}_{\mathbf{k}s} \left[ \hat{a}_{\mathbf{k}s} e^{i(\mathbf{k} \cdot \mathbf{r} - \omega_k t)} + \hat{a}_{\mathbf{k}s}^\dagger e^{-i(\mathbf{k} \cdot \mathbf{r} - \omega_k t)} \right] \quad (2.144)$$

$$\hat{\mathbf{E}}(\mathbf{r}, t) = i \sum_{\mathbf{k}s} \left( \frac{\hbar \omega_k}{2\epsilon_0 V} \right)^{\frac{1}{2}} \mathbf{e}_{\mathbf{k}s} \left[ \hat{a}_{\mathbf{k}s} e^{i(\mathbf{k} \cdot \mathbf{r} - \omega_k t)} - \hat{a}_{\mathbf{k}s}^\dagger e^{-i(\mathbf{k} \cdot \mathbf{r} - \omega_k t)} \right] \quad (2.145)$$

$$\hat{\mathbf{B}}(\mathbf{r}, t) = \frac{i}{c} \sum_{\mathbf{k}s} (\boldsymbol{\kappa} \times \mathbf{e}_{\mathbf{k}s}) \left( \frac{\hbar \omega_k}{2\epsilon_0 V} \right)^{\frac{1}{2}} \mathbf{e}_{\mathbf{k}s} \left[ \hat{a}_{\mathbf{k}s} e^{i(\mathbf{k} \cdot \mathbf{r} - \omega_k t)} - \hat{a}_{\mathbf{k}s}^\dagger e^{-i(\mathbf{k} \cdot \mathbf{r} - \omega_k t)} \right]. \quad (2.146)$$

The annihilation and creation operators are to be understood as Heisenberg picture operators evaluated at  $t = 0$ . The time-dependent annihilation operator for a free field is given by

$$\hat{a}_{\mathbf{k}s}(t) = \hat{a}_{\mathbf{k}s}(0)e^{-i\omega_k t}. \quad (2.147)$$

With this the electric field can be written as

$$\hat{\mathbf{E}}(\mathbf{r}, t) = i \sum_{\mathbf{k}s} \left( \frac{\hbar\omega_k}{2\epsilon_0 V} \right)^{\frac{1}{2}} \mathbf{e}_{\mathbf{k}s} \left[ \hat{a}_{\mathbf{k}s}(t)e^{i\mathbf{k}\cdot\mathbf{r}} - \hat{a}_{\mathbf{k}s}^\dagger(t)e^{-i\mathbf{k}\cdot\mathbf{r}} \right]. \quad (2.148)$$

where the  $e^{\pm i\omega_k t}$  term is absorbed into the operator. Alternatively, the field can be written in a different form:

$$\hat{\mathbf{E}}(\mathbf{r}, t) = \hat{\mathbf{E}}^{(+)}(\mathbf{r}, t) + \hat{\mathbf{E}}^{(-)}(\mathbf{r}, t) \quad (2.149)$$

where

$$\hat{\mathbf{E}}^{(+)}(\mathbf{r}, t) = i \sum_{\mathbf{k}s} \left( \frac{\hbar\omega_k}{2\epsilon_0 V} \right)^{\frac{1}{2}} \mathbf{e}_{\mathbf{k}s} \hat{a}_{\mathbf{k}s}(t) e^{i\mathbf{k}\cdot\mathbf{r}}. \quad (2.150)$$

Of course,

$$\hat{\mathbf{E}}^{(+)}(\mathbf{r}, t) = \left[ \hat{\mathbf{E}}^{(-)}(\mathbf{r}, t) \right]^\dagger. \quad (2.151)$$

Similar expressions can be obtained for the magnetic field. However, because the magnetic field is “weaker” than the electric field by a factor of  $1/c$ , we generally neglect the magnetic field in most matter-field interactions (as most interactions are through the electric dipole).  $\hat{\mathbf{E}}^{(+)}(\mathbf{r}, t)$  is called the *positive frequency part* of the field as it contained all terms oscillate as  $e^{-i\omega t}$  for  $\omega > 0$ . The opposite is for  $\hat{\mathbf{E}}^{(-)}(\mathbf{r}, t)$ . We notice that  $\hat{\mathbf{E}}^{(+)}(\mathbf{r}, t)$  is a collective annihilation operator while the latter is a collective creation operator.

The last observation we make is the following. For a single-mode plane wave field the electric field is

$$\hat{E}(\mathbf{r}, t) = i \left( \frac{\hbar\omega_k}{2\epsilon_0 V} \right)^{\frac{1}{2}} \mathbf{e}_x \left[ \hat{a}e^{i\mathbf{k}\cdot\mathbf{r}-i\omega t} - \hat{a}^\dagger e^{-i\mathbf{k}\cdot\mathbf{r}+i\omega t} \right]. \quad (2.152)$$

In most of quantum optics, the spatial variation of the field over the dimensions of the atomic system may be negligible, i.e., the length characteristic of the size of the atom is often much smaller than a typical wavelength:

$$\frac{\lambda}{2\pi} = \frac{1}{|k|} \gg |\mathbf{r}_{\text{atom}}|. \quad (2.153)$$

With this, we can make an approximation:

$$e^{\pm i\mathbf{k}\cdot\mathbf{r}} \approx 1 \pm i\mathbf{k}\cdot\mathbf{r} \quad (2.154)$$

which says we replace the exponential by unity for  $|\mathbf{r}|$  is negligible. Under this approximation,

$$\hat{\mathbf{E}}(\mathbf{r}, t) \approx \hat{\mathbf{E}}(t) = i \left( \frac{\hbar \omega_k}{2\epsilon_0 V} \right)^{\frac{1}{2}} \mathbf{e}_x [\hat{a} e^{-i\omega t} - \hat{a}^\dagger e^{i\omega t}] \quad (2.155)$$

This is called the “**dipole**” approximation.

## 2.2 Emission and Absorption of Radiation by Atoms

### 2.2.1 Atom-field interactions

Suppose we have an atom in no external field. The Hamiltonian of an electron bound to this atom is given by

$$\hat{\mathcal{H}}_0 = \frac{1}{2m} \hat{\mathbf{P}}^2 + V(r) \quad (2.156)$$

where  $V(r)$  is the Coulomb interaction between the electron and nucleus and  $r = |\mathbf{r}|$  is the distance to the core. In the configuration space representation,

$$\hat{\mathbf{P}} = -i\hbar \nabla \quad (2.157)$$

$$\hat{\mathbf{r}} |\mathbf{r}\rangle = \mathbf{r} |\mathbf{r}\rangle. \quad (2.158)$$

The wave functions are given by

$$\psi(\mathbf{r}) = \langle \mathbf{r} | \psi \rangle. \quad (2.159)$$

We assume that the energy eigenstates  $|k\rangle$  of  $\hat{\mathcal{H}}_0$  satisfy the time-independent Schrödinger equation:

$$\hat{\mathcal{H}}_0 |\psi\rangle_k^{(0)}(\mathbf{r}) = E_k \psi_k^{(0)}(\mathbf{r}) \quad (2.160)$$

where

$$\langle \mathbf{r} | k \rangle = \psi_k^{(0)}(\mathbf{r}) \quad (2.161)$$

are known. In the presence of external fields the Hamiltonian is modified to

$$\hat{\mathcal{H}}(\mathbf{r}, t) = \frac{1}{2m} \left[ \hat{\mathbf{P}} + e\mathbf{A}(\mathbf{r}, t) \right]^2 - e\Phi(\mathbf{r}, t) + V(r) \quad (2.162)$$

where  $\mathbf{A}(\mathbf{r}, t)$  and  $\Phi(\mathbf{r}, t)$  are the vector and scalar potentials, respectively,  $-e$  is the electron charge. The electric and magnetic fields are given by

$$\mathbf{E}(\mathbf{r}, t) = -\nabla\Phi(\mathbf{r}, t) - \frac{\partial\mathbf{A}(\mathbf{r}, t)}{\partial t} \quad (2.163)$$

$$\mathbf{B}(\mathbf{r}, t) = \nabla \times \mathbf{A}(\mathbf{r}, t) \quad (2.164)$$

which we can easily show to be invariant under the following gauge transformations

$$\Phi'(\mathbf{r}, t) = \Phi(\mathbf{r}, t) - \frac{\partial\chi(\mathbf{r}, t)}{\partial t} \quad (2.165)$$

$$\mathbf{A}'(\mathbf{r}, t) = \mathbf{A}(\mathbf{r}, t) + \nabla\chi(\mathbf{r}, t). \quad (2.166)$$

The time-dependent Schrödinger equation is

$$\hat{\mathcal{H}}(\mathbf{r}, t)\psi(\mathbf{r}, t) = i\hbar \frac{\partial \psi(\mathbf{r}, t)}{\partial t}. \quad (2.167)$$

To eventually simplify the form of the atom-field interaction, we define a unitary operator  $\hat{R}$  such that

$$\psi'(\mathbf{r}, t) = \hat{R}\psi(\mathbf{r}, t) \quad (2.168)$$

where

$$\hat{\mathcal{H}}'\psi' = i\hbar \frac{\partial \psi'}{\partial t}. \quad (2.169)$$

We can show that the new Hamiltonian relates to the old Hamiltonian as

$$\hat{\mathcal{H}}' = \hat{R}\hat{\mathcal{H}}\hat{R}^\dagger + i\hbar \frac{\partial \hat{R}}{\partial t} \hat{R}^\dagger. \quad (2.170)$$

Our goal is to choose

$$\hat{R} = e^{-ie\chi(\mathbf{r}, t)/\hbar}, \quad (2.171)$$

so that with  $\hat{\mathbf{P}}^2 = -i\hbar\nabla$ :

$$\hat{\mathcal{H}}' = \frac{1}{2m} \left[ \hat{\mathbf{P}}^2 + e\mathbf{A}' \right]^2 - e\Phi' + V(r), \quad (2.172)$$

where the new vector and scalar potentials  $\mathbf{A}'$  and  $\Phi'$  depend on the gauge field  $\chi$  as (2.165). This form of the Hamiltonian is not desirable because of the interaction terms  $\hat{\mathbf{P}} \cdot \mathbf{A}$  and  $\mathbf{A}^2$ . To remedy this, we can first make a definite choice of gauge, namely the Coulomb gauge, for which  $\Phi = 0$  and  $\nabla \cdot \mathbf{A} = 0$ . Now with  $\mathbf{A}' = \mathbf{A} + \nabla\chi$  and  $\Phi' = \Phi - \partial\chi/\partial t = -\partial\chi/\partial t$ , the new Hamiltonian has the form

$$\hat{\mathcal{H}}' = \frac{1}{2m} \left[ \hat{\mathbf{P}} + e(\mathbf{A} + \nabla\chi) \right]^2 + e\frac{\partial\chi}{\partial t} + V(r). \quad (2.173)$$

The vector potential  $\mathbf{A}$  for no sources near the atom satisfies the wave equation

$$\nabla^2 \mathbf{A} - \frac{1}{c^2} \frac{\partial^2 \mathbf{A}}{\partial t^2} = 0, \quad (2.174)$$

whose solution is simply a wave propagating wave

$$\mathbf{A} = \mathbf{A}_0 e^{i(\mathbf{k} \cdot \mathbf{x} - \omega t)} + c.c. \quad (2.175)$$

where  $|\mathbf{k}| = 2\pi/\lambda$  is the wave vector of the radiation. For  $|\mathbf{r}|$  a typical atomic length ( $10^{-10}$  m) and  $\lambda$  of typical optical wavelengths ( $10^{-7}$  m), the product  $\mathbf{k} \cdot \mathbf{r} \ll 1$  so that over the extent of the atom  $\mathbf{A}$  can be taken to be uniform, i.e.,

$$\mathbf{A}(\mathbf{r}, t) \approx \mathbf{A}(t). \quad (2.176)$$



This is very much like the idea we discussed before regarding the electric field  $\mathbf{E}$  in the quantization of the EM field. This is also called the **dipole approximation**. Now, the Hamiltonian in (2.173) has coupled terms, and we would like to remove  $\nabla\chi$  from the coupling between itself and  $\mathbf{A}$  and between  $\mathbf{A}$  and  $\hat{\mathbf{P}}$ . One very good way to do this is finding  $\chi$  such that

$$\nabla\chi = -\mathbf{A}(\mathbf{r}, t) \approx \mathbf{A}(t). \quad (2.177)$$

Such a  $\chi(\mathbf{r}, t)$  is

$$\chi(\mathbf{r}, t) = -\mathbf{A}(t) \cdot \mathbf{r}. \quad (2.178)$$

With this choice, we have (since  $\nabla\mathbf{r} = \mathbf{1}$ )

$$\nabla\chi(\mathbf{r}, t) = -\mathbf{A}(t) \quad (2.179)$$

$$\frac{\partial\chi}{\partial t}(\mathbf{r}, t) = -\mathbf{r} \cdot \frac{\partial\mathbf{A}(t)}{\partial t} = -\mathbf{r} \cdot \mathbf{E}(t). \quad (2.180)$$

The new Hamiltonian is now “decoupled”:

$$\hat{\mathcal{H}}' = \underbrace{\frac{1}{2m}\hat{\mathbf{P}}^2 + V(r)}_{\text{field-free}} \underbrace{-e\mathbf{r} \cdot \mathbf{E}(t)}_{\text{dipole int.}} \quad (2.181)$$

This new Hamiltonian contains only one interaction term (within the dipole approximation) as opposed to two terms as before.  $-e\mathbf{r}$  is called the **dipole moment**, denoted  $\hat{\mathbf{d}}$ . We can of course promote it to an operator, so that the Hamiltonian in the presence of an external field can be written as

$$\hat{\mathcal{H}} = \hat{\mathcal{H}}_0 - \hat{\mathbf{d}} \cdot \mathbf{E}(t) \quad (2.182)$$

where  $\hat{\mathcal{H}}_0 = \hat{\mathbf{P}}^2/2m + V(r)$  is of course the field-free Hamiltonian.

### 2.2.2 Interaction of an atom with a classical field

So far we have not specified whether the interacting field is classical or quantum mechanical. The derivation we just looked at is valid for both classical and quantum fields. However, we will eventually show that the atom behaves differently in each case.

Let us first look at the the atom-field interaction with a classical field. Suppose the field is given by

$$\mathbf{E}(t) = \mathbf{E}_0 \cos(\omega t) \quad (2.183)$$

where the dipole approximation  $\mathbf{k} \cdot \mathbf{r} \ll 1$  has already been made (so that the field is uniform over the atom). Suppose that  $\omega$  is the frequency of the radiation,

and the field is abruptly turned on at  $t = 0$ . Assume that the initial state of the atom is  $|i\rangle$  where

$$\hat{\mathcal{H}}_0 |i\rangle = E_i |i\rangle. \quad (2.184)$$

For time  $t > 0$ , the state vector  $|\psi\rangle(t)$  is a linear combination of all uncoupled states  $|k\rangle$ :

$$|\psi(t)\rangle = \sum_k C_k(t) e^{-iE_k t/\hbar} |k\rangle \quad (2.185)$$

where

$$\sum_k |C_k(t)|^2 = 1. \quad (2.186)$$

The time-dependent Schrödinger equation is

$$i\hbar \frac{\partial |\psi(t)\rangle}{\partial t} = (\hat{\mathcal{H}}_0 + \hat{\mathcal{H}}^{(I)}) |\psi(t)\rangle \quad (2.187)$$

where

$$\hat{\mathcal{H}}^{(I)} = -\hat{\mathbf{d}} \cdot \mathbf{E}(t). \quad (2.188)$$

Substituting the expansion (2.185) into the time-dependent SE, then multiplying from the left by  $\langle l| e^{-E_l t/\hbar}$  gives

$$i\hbar \langle l| e^{-E_l t/\hbar} \frac{\partial}{\partial t} \sum_k C_k(t) e^{-iE_k t/\hbar} |k\rangle = \langle l| e^{-iE_l t/\hbar} (\hat{\mathcal{H}}_0 + \hat{\mathcal{H}}^{(I)}) \sum_k C_k(t) e^{-iE_k t/\hbar} |k\rangle. \quad (2.189)$$

This simplifies to

$$\dot{C}_l(t) - \frac{i}{\hbar} \sum_k C_k(t) \langle l| \hat{\mathcal{H}}^{(I)} |k\rangle e^{-i(E_l - E_k)t/\hbar}. \quad (2.190)$$

Defining

$$\omega_l = \frac{E_l - E_K}{\hbar} \quad (2.191)$$

as the transition frequencies between levels  $l$  and  $k$  gives a set of couple-first-order differential equations for the amplitudes

$$\dot{C}_l(t) = -\frac{i}{\hbar} \sum_k C_k(t) \langle l| \hat{\mathcal{H}}^{(I)} |k\rangle e^{-i\omega_l t}. \quad (2.192)$$

The initial condition is  $C_i(0) = 1$ :  $|i\rangle$  is the only state initially populated. As time goes forward, the population in  $|i\rangle$  will decrease and move to some unpopulated  $|f\rangle$ . The probability of the transition  $i \rightarrow f$  is given by

$$P_{i \rightarrow f}(t) = C^*(f) C_f(t) = |C_f(t)|^2. \quad (2.193)$$

For simple cases, these probabilities can be solved analytically or even numerically for more complex cases. However, for the “weak” field where  $|\langle f | \mathbf{d} \cdot \hat{\mathbf{E}}_0 | i \rangle|$  is small, we can use the time-dependent perturbation theory to approach the problem. For bookkeeping purposes we shall write the interaction Hamiltonian as  $\lambda \hat{\mathcal{H}}^{(1)}$  where  $0 \leq \lambda \leq 1$ . (At the end of the calculations we often take  $\lambda \rightarrow 1$ .) We then expand the probability amplitude for, say, the state  $|l\rangle$  in the power series

$$C_l(t) = \sum_{n=0}^{\infty} \lambda^n C_l^{(n)}(t). \quad (2.194)$$

Inserting this expansion into (2.192) and equating the like powers of  $\lambda$  we obtain up to second order

$$\dot{C}_l^{(0)} = 0 \quad (2.195)$$

$$\dot{C}_l^{(1)} = -\frac{i}{\hbar} \sum_k C_k^{(0)}(t) \mathcal{H}_{lk}^{(I)}(t) e^{i\omega_{lk}t} \quad (2.196)$$

$$\dot{C}_l^{(2)} = -\frac{i}{\hbar} \sum_k C_k^{(1)}(t) \mathcal{H}_{lk}^{(I)}(t) e^{i\omega_{lk}t} \quad (2.197)$$

where

$$\mathcal{H}_{lk}^{(I)}(t) = \langle l | \hat{\mathcal{H}}^{(I)}(t) | k \rangle = \langle l | -\hat{\mathbf{d}} \cdot \mathbf{E}(t) | k \rangle. \quad (2.198)$$

We observe the pattern relating the  $n^{th}$  order to the  $n+1^{th}$  order:

$$\dot{C}_l^{(n+1)} = -\frac{i}{\hbar} \sum_k C_k^{(n)}(t) \mathcal{H}_{lk}^{(I)}(t) e^{i\omega_{lk}t}. \quad (2.199)$$

The essential assumption underlying the perturbation-theory approach is that the driving field is so weak that the atomic populations change very little. That is, if  $C_i(0) = 1$  and  $C_f(0) = 0$  then for  $t > 0$ , a good approximation is  $C_i(t) \approx 1$  and  $C_f(t) \ll 1$ . Thus, for the expression of  $\dot{C}_l^{(1)}$  the only surviving term is for the state  $k$  to be the initial state  $i$ . This gives

$$\dot{C}_f^{(1)}(t) = -\frac{i}{\hbar} C_i^{(0)}(t) \mathcal{H}_{fi}^{(I)}(t) e^{i\omega_{fi}t}. \quad (2.200)$$

Integrating this gives

$$C_f^{(1)}(t) = -\frac{i}{\hbar} \int_0^t C_i^{(0)}(t') \mathcal{H}_{fi}^{(I)}(t') e^{i\omega_{fi}t'} dt'. \quad (2.201)$$

Following a similar line of reasoning we get for  $\dot{C}_l^{(2)}$

$$\dot{C}_l^{(2)}(t) = -\frac{i}{\hbar} \sum_l \int_0^t C_l^{(1)}(t') \mathcal{H}_{li}^{(I)}(t') e^{i\omega_{li}t'} dt'. \quad (2.202)$$

Now replacing  $C_l^{(1)}$  with (2.201) we get

$$C_l^{(2)}(t) = -\frac{i}{\hbar} \sum_l \int_0^t \left( -\frac{i}{\hbar} \int_0^{t'} C_i^{(0)}(t'') \mathcal{H}_{fi}^{(I)}(t'') e^{i\omega_{fi}t''} dt'' \right) \mathcal{H}_{fl}^{(I)}(t') e^{i\omega_{fl}t'} dt' \quad (2.203)$$

$$= \left( \frac{-i}{\hbar} \right)^2 \sum_l \int_0^t dt' \int_0^{t'} dt'' \left( \mathcal{H}_{fl}^{(I)}(t') e^{i\omega_{fl}t'} \right) \mathcal{H}_{fi}^{(I)}(t'') e^{i\omega_{fi}t''} \quad (2.204)$$

We see that while (2.201) gives the transition probability amplitude from  $|i\rangle \rightarrow |f\rangle$ , (2.203) gives the transition probability amplitude from  $|i\rangle \rightarrow |\{l\}\rangle \rightarrow |f\rangle$ . The total transition probability from  $|i\rangle$  to  $|f\rangle$  is then the square of the sum of the amplitudes:

$$P_{i \rightarrow f}(t) = \left| \sum_{n=0}^{\infty} C_f^{(n)}(t) \right|^2. \quad (2.205)$$

The dipole moment operator  $\hat{\mathbf{d}} = -e\mathbf{r}$  has non-vanishing matrix elements only between states of opposite parity. Thus the first-order correction to the amplitude of the initial state (where  $f = i$ ) vanishes:

$$C_i^{(1)}(t) = -\frac{i}{\hbar} \int_0^t C_i^{(0)}(t') \mathcal{H}_{ii}^{(I)}(t') e^{i\omega_{ii}t'} dt' = 0, \quad (2.206)$$

since

$$\langle i | \hat{\mathcal{H}}^{(I)}(t) | i \rangle = 0. \quad (2.207)$$

Thus, to first order,

$$C_i(t) = C_i^{(0)}(t) = 1. \quad (2.208)$$

Thus, by (2.201)

$$C_f^{(1)}(t) = -\frac{i}{\hbar} \int_0^t C_i^{(0)}(t') \mathcal{H}_{fi}^{(I)}(t') e^{i\omega_{fi}t'} dt' \quad (2.209)$$

$$= -\frac{i}{\hbar} \int_0^t \mathcal{H}_{fi}^{(I)}(t') e^{i\omega_{fi}t'} dt'. \quad (2.210)$$

Now, with

$$\hat{\mathcal{H}}^{(I)}(t) = -\hat{\mathbf{d}} \cdot \mathbf{E}_0 \cos(\omega t) = -\frac{1}{2} \hat{\mathbf{d}} \cdot \mathbf{E}_0 (e^{i\omega t} + e^{-i\omega t}), \quad (2.211)$$

the integral above becomes

$$C_f^{(1)}(t) = \left( -\frac{i}{\hbar} \right) \left( -\frac{1}{2} \right) \langle f | \hat{\mathbf{d}} \cdot \mathbf{E}_0 | i \rangle \int_0^t (e^{i\omega t} + e^{-i\omega t}) e^{i\omega_{fi}t'} dt' \quad (2.212)$$

$$= \boxed{\frac{1}{2\hbar} (\hat{\mathbf{d}} \cdot \mathbf{E}_0)_{fi} \cdot \left\{ \frac{e^{i(\omega + \omega_{fi})t} - 1}{\omega + \omega_{fi}} - \frac{e^{i(\omega - \omega_{fi})t} - 1}{\omega - \omega_{fi}} \right\}} \quad (2.213)$$

If  $\omega \approx \omega_{fi}$ , i.e., the field is near resonance then clearly the second term in the curly brackets dominates. With RWA (rotating wave approximation) we drop the first term. And so the transition probability to first order is

$$P_{i \rightarrow f}^{(1)}(t) = \left| C_f^{(1)}(t) \right|^2 = \frac{\left| (\hat{\mathbf{d}} \cdot \mathbf{E}_0)_{fi} \right|^2}{\hbar^2} \frac{\sin^2(\Delta t/2)}{\Delta^2} \quad (2.214)$$

where  $\Delta = \omega - \omega_{fi}$  is called the detuning and we have used the identity

$$e^{-i\theta-1} = \cos \theta t - i \sin \theta t - 1 \quad (2.215)$$

$$= -2 \sin^2 \frac{\theta t}{2} - 2i \sin \frac{\theta t}{2} \cos \frac{\theta t}{2} \quad (2.216)$$

$$= -2 \sin \frac{\theta t}{2} \left( e^{i\theta t/2} \right) \quad (2.217)$$

so that

$$\left| \frac{-1}{2} (e^{i\Delta t} - 1) \right|^2 = \sin^2 \frac{\Delta t}{2}. \quad (2.218)$$

When  $\Delta \neq 0$ , the transition probability maximizes at

$$\left( P_{i \rightarrow f}^{(1)} \right)_{\max} = \frac{\left| (\hat{\mathbf{d}} \cdot \mathbf{E}_0)_{fi} \right|^2}{\hbar^2} \frac{1}{\Delta^2}. \quad (2.219)$$

On resonance, i.e.,  $\Delta = 0$ ,

$$\left( P_{i \rightarrow f}^{(1)} \right)_{\max} = \frac{\left| (\hat{\mathbf{d}} \cdot \mathbf{E}_0)_{fi} \right|^2}{\hbar^2} t^2. \quad (2.220)$$

Thus we have the following For the perturbation approximation to be valid we



require that  $C_f^{(1)}(t) \ll 1$ , which implies  $P_{i \rightarrow f}^{(1)}(t) \ll 1$ . For  $\Delta \neq 0$ , this places

constraints on both  $\left| \left( \hat{\mathbf{d}} \cdot \mathbf{E}_0 \right)_{fi} \right|$  and  $\Delta$ . On resonance, (2.220) is only valid for a very short time. For the case of the transition probability as a function of  $\Delta$ , the width of the peak is proportional to  $t^{-1}$  while the height is proportional to  $t^2$ . It turns out that the area is proportional to  $t$ :

$$\int_{-\infty}^{\infty} \frac{\sin^2 \Delta t/2}{\Delta^2} d\Delta = \frac{\pi t}{2} \quad (2.221)$$

In the limit where  $\Delta \approx 0$  and  $t \gg 2\pi/\omega_{fi}$ , the function in the integrand may be approximated by the Dirac delta function

$$\lim_{t \rightarrow \infty} \frac{\sin^2 \Delta t/2}{\Delta^2} = \frac{\pi}{2} t \delta(\Delta). \quad (2.222)$$

In this case the transition probability is

$$P_{i \rightarrow f}^{(1)}(t) = \frac{\pi}{2} \left| C_f^{(1)}(t) \right|^2 = \frac{\left| \left( \hat{\mathbf{d}} \cdot \mathbf{E}_0 \right)_{fi} \right|^2}{\hbar^2} t \delta(\Delta). \quad (2.223)$$

Now we define the time-independent transition probability **rate** as

$$W_{i \rightarrow f} = \frac{P_{i \rightarrow f}^{(1)}}{t} = \frac{\pi}{2} \frac{\left| \left( \hat{\mathbf{d}} \cdot \mathbf{E}_0 \right)_{fi} \right|^2}{\hbar^2} \delta(\omega - \omega_{fi}). \quad (2.224)$$

In practice, the driving field will not be monochromatic so that a range of frequencies will have to be summed or integrated over to obtain the total transition rate. If  $[f]$  represents a set of accessible final states, then the transition rate for a monochromatic field is

$$W_{i \rightarrow [f]} = \frac{\pi}{2} \sum_{[f]} \frac{\left| \left( \hat{\mathbf{d}} \cdot \mathbf{E}_0 \right)_{fi} \right|^2}{\hbar^2} \delta(\omega - \omega_{fi}) \quad (2.225)$$

This expression is often famously called Fermi's Golden Rule. Now suppose there is a broad range of frequencies. This means the amplitude of the light is now frequency dependent  $\mathbf{E}_0 \rightarrow \mathbf{E}_0(\omega)$ . Thus the transition probability rate induced by all the frequency components must be

$$\frac{P_{i \rightarrow f}^{(1)}}{t} = \int d\omega \frac{\left| \left( \hat{\mathbf{d}} \cdot \mathbf{E}_0 \right)_{fi} \right|^2}{\hbar^2} \frac{\sin^2(\Delta t/2)}{\Delta^2} \quad (2.226)$$

$$= \frac{1}{\hbar} \int d\omega \frac{\sin^2(\Delta t/2)}{\Delta^2} \underbrace{\left| \langle f | \hat{\mathbf{d}} \cdot \hat{\mathbf{E}}_0(\omega) | i \rangle \right|^2}_{F(\omega)}. \quad (2.227)$$

Now if  $F(\omega)$  varies slowly with  $\omega$  compared to  $\frac{\sin^2(\Delta t/2)}{\Delta^2}$  then we can take  $F(\omega)$  to be its resonant value  $F(\omega_{fi})$ . With this we can take it out of the integral and get

$$W_{i \rightarrow f} = \frac{P_{i \rightarrow f}^{(1)}}{t} = \frac{\pi}{2\hbar^2} F(\omega_{fi}). \quad (2.228)$$

The spread of the frequencies results in the dephasing of the oscillations as shown in the plot of transition probability versus time. This is due to the fact that the light is *incoherent* (lacking the phase relations between the various frequency components). If the atom is driven by coherent light (like a laser), dephasing don't occur and the perturbative time-independent transition rates above generally do not adequately describe the dynamics. We will resolve this issue later.

### 2.2.3 Interaction of an atom with a quantized field & Spontaneous emission

In the previous subsection, we did not make any assumption about the relative position of the states  $i$  and  $f$ , i.e., whether  $E_i$  is greater or less than  $E_f$ . We simply said that so long as  $\mathbf{E}_0 \neq 0$  and  $E_i \neq E_f$ , there is some non-zero transition probability. In this section, we will now show that for the case  $E_i > E_f$ , when the field is quantized, transitions will occur even when no photons are present. This is called *spontaneous emission*. This is one of many differences that appear when we treat the field quantum mechanically versus classically.

Let us revisit the single-mode free space field given by (2.155) after dipole approximation (where we assumed the field is uniform over the atom):

$$\hat{\mathbf{E}}(\mathbf{r}, t) \approx \hat{\mathbf{E}}(t) = i \left( \frac{\hbar \omega_k}{2\epsilon_0 V} \right)^{\frac{1}{2}} \mathbf{e} [\hat{a} e^{-i\omega t} - \hat{a}^\dagger e^{i\omega t}]. \quad (2.229)$$

But also recall that this free-space field is written in the Heisenberg picture (there is a time-dependence when the field operator is expressed in the Heisenberg basis). In the Schrödinger picture, though, the field operator becomes

$$\hat{\mathbf{E}} = i \left( \frac{\hbar \omega_k}{2\epsilon_0 V} \right)^{\frac{1}{2}} \mathbf{e} [\hat{a} - \hat{a}^\dagger]. \quad (2.230)$$

Now, the free-space Hamiltonian  $\hat{\mathcal{H}}_0$  must be

$$\hat{\mathcal{H}}_0 = \hat{\mathcal{H}}_{\text{atom}} + \hat{\mathcal{H}}_{\text{field}} \quad (2.231)$$

where  $\hat{\mathcal{H}}_{\text{atom}}$  is just the free-atom Hamiltonian as before and  $\hat{\mathcal{H}}_{\text{field}}$  is the free-space field Hamiltonian, which we have seen before as well:

$$\hat{\mathcal{H}}_{\text{field}} = \hbar \omega \left( \hat{a}^\dagger \hat{a} + \frac{1}{2} \right). \quad (2.232)$$

Now, because the zero-point energy term does not contribute to the dynamics, we will simply drop it. With this, the interaction Hamiltonian is given by

$$\hat{\mathcal{H}}^{(I)} = -\hat{\mathbf{d}} \cdot \hat{\mathbf{E}} = -i \left( \frac{\hbar \omega_k}{2\epsilon_0 V} \right)^{\frac{1}{2}} (\hat{\mathbf{d}} \cdot \mathbf{e}) (\hat{a} - \hat{a}^\dagger). \quad (2.233)$$

Let us define

$$\boldsymbol{\mathcal{E}}_0 = i \left( \frac{\hbar \omega_k}{2\epsilon_0 V} \right)^{\frac{1}{2}} \mathbf{e}. \quad (2.234)$$

From this,

$$\hat{\mathcal{H}}^{(I)} = -\hat{\mathbf{d}} \cdot \boldsymbol{\mathcal{E}}_0 (\hat{a} - \hat{a}^\dagger). \quad (2.235)$$

Now, because both the atomic and field systems are quantized, the states of the combined system is the product of the states of both the atom and field. Suppose initially we're given the initial state of the atom  $|a\rangle$  and  $n$  photons  $|n\rangle$ , then the initial state of the system is

$$|i\rangle = |a\rangle |n\rangle. \quad (2.236)$$

The perturbation interaction of the quantized field causes a transition to the new state

$$|f_1\rangle = |b\rangle |n-1\rangle \quad (2.237)$$

where  $n-1$  means that one photon has been absorbed by the atom to transition from  $|a\rangle$  to  $|b\rangle$ , or to the state

$$|f_2\rangle = |b\rangle |n+1\rangle \quad (2.238)$$

where  $n+1$  means that one photon has been emitted by the atom to transition from  $|a\rangle$  to  $|b\rangle$ . The energy of these states are

$$\text{for } |i\rangle = |a\rangle |n\rangle, \quad E_i = E_a + n\hbar\omega \quad (2.239)$$

$$\text{for } |f_1\rangle = |b\rangle |n-1\rangle, \quad E_i = E_b + (n-1)\hbar\omega \quad (2.240)$$

$$\text{for } |f_2\rangle = |b\rangle |n+1\rangle, \quad E_i = E_b + (n+1)\hbar\omega \quad (2.241)$$

where  $E_a$  and  $E_b$  are the energies of the atomic states  $|a\rangle$  and  $|b\rangle$ .

We observe that the interaction Hamiltonian (or the Hamiltonian associated with the perturbation) is time-independent. The matrix elements of this interaction can be calculated by finding

$$\text{Absorption: } \langle f_1 | \hat{\mathcal{H}}^{(I)} | i \rangle \quad (2.242)$$

$$\text{Emission: } \langle f_2 | \hat{\mathcal{H}}^{(I)} | i \rangle. \quad (2.243)$$



For absorption, we need a term in the interaction Hamiltonian that contains only the annihilation operator:

$$\textbf{Absorption: } \langle f_1 | \hat{\mathcal{H}}^{(I)} | i \rangle = \langle b, n-1 | -\hat{\mathbf{d}} \cdot \boldsymbol{\mathcal{E}}_0(\hat{a}) | a, n \rangle \quad (2.244)$$

$$= \langle b, n-1 | -\hat{\mathbf{d}} \cdot \boldsymbol{\mathcal{E}}_0 \sqrt{n} | a, n-1 \rangle \quad (2.245)$$

$$= \langle b | -\hat{\mathbf{d}} \cdot \boldsymbol{\mathcal{E}}_0 | a \rangle \sqrt{n} \langle n-1 | n-1 \rangle \quad (2.246)$$

$$= \boxed{-\left(\hat{\mathbf{d}} \cdot \boldsymbol{\mathcal{E}}_0\right)_{ba} \sqrt{n}} \quad (2.247)$$

Similarly, for emission

$$\textbf{Emission: } \langle f_2 | \hat{\mathcal{H}}^{(I)} | i \rangle = \langle b, n+1 | -\hat{\mathbf{d}} \cdot \boldsymbol{\mathcal{E}}_0(-\hat{a}^\dagger) | a, n \rangle \quad (2.248)$$

$$= \langle b, n+1 | \hat{\mathbf{d}} \cdot \boldsymbol{\mathcal{E}}_0 \sqrt{n+1} | a, n+1 \rangle \quad (2.249)$$

$$= \langle b | \hat{\mathbf{d}} \cdot \boldsymbol{\mathcal{E}}_0 | a \rangle \sqrt{n+1} \langle n+1 | n+1 \rangle \quad (2.250)$$

$$= \boxed{\left(\hat{\mathbf{d}} \cdot \boldsymbol{\mathcal{E}}_0\right)_{ba} \sqrt{n+1}} \quad (2.251)$$

where

$$\left(\hat{\mathbf{d}} \cdot \boldsymbol{\mathcal{E}}_0\right)_{ba} = \langle b | \hat{\mathbf{d}} \cdot \boldsymbol{\mathcal{E}}_0 | a \rangle = \langle b | \hat{\mathbf{d}} | a \rangle \cdot \boldsymbol{\mathcal{E}}_0 \equiv \mathbf{d}_{ba} \cdot \boldsymbol{\mathcal{E}}_0 \quad (2.252)$$

is the dipole matrix element between states  $|b\rangle$  and  $|a\rangle$ .

It is worthy to compare this result to the semiclassical picture we saw in the last subsection. For the absorption case, there is nothing new: if no field then no absorption:

$$n = 0 \implies \langle f_1 | \hat{\mathcal{H}}^{(I)} | i \rangle = 0 \quad (2.253)$$

since  $\sqrt{n} = 0$ . However, the emission case has no semiclassical counterpart: transitions may occur even when no photons are present:

$$n = 0 \not\Rightarrow \langle f_2 | \hat{\mathcal{H}}^{(I)} | i \rangle = 0 \quad (2.254)$$

since  $\sqrt{n+1} \neq 0$ . This is called **spontaneous emission**, which cannot be obtained from the semiclassical approach. If  $n > 0$ , the emission of an additional photon is called **stimulated emission**, which is essential for the operation of the “light amplification by stimulated emission of radiation” or LASER.

We might also be interested in the ratio of the absorption/emission rates:

$$\boxed{\frac{\left|\langle f_2 | \hat{\mathcal{H}}^{(I)} | i \rangle\right|^2}{\left|\langle f_1 | \hat{\mathcal{H}}^{(I)} | i \rangle\right|^2} = \frac{n+1}{n}} \quad (2.255)$$

It turns out that we can still use the perturbation method developed in the previous section to calculate the transition amplitudes, provided that we make appropriate modifications to account for the fact that the field is now quantized. The first difference is in the Hamiltonians:

$$\hat{\mathcal{H}}_0 = \hat{\mathcal{H}}_{\text{atom}} \rightarrow \hat{\mathcal{H}}_0 = \hat{\mathcal{H}}_{\text{atom}} + \hat{\mathcal{H}}_{\text{field}} \quad (2.256)$$

$$\hat{\mathcal{H}}^{(I)}(t) = -\hat{\mathbf{d}} \cdot \mathbf{E}(t) \rightarrow \hat{\mathcal{H}}^{(I)}(t) = -\hat{\mathbf{d}} \cdot \boldsymbol{\mathcal{E}}_0 (\hat{a} - \hat{a}^\dagger) \quad (2.257)$$

where of course

$$\hat{\mathcal{H}}_{\text{field}} = \hbar\omega \left( \hat{a}^\dagger \hat{a} + \frac{1}{2} \right) \quad (2.258)$$

where we again leave out the zero-point energy term.

Now, let us assume that the atom has only two levels  $|a\rangle$  and  $|b\rangle$ . The state vector of the system can be written as

$$\begin{aligned} |\psi(t)\rangle = & C_i(t) |a\rangle |n\rangle e^{-i\omega_a t} e^{-in\omega t} \\ & + C_{f_1}(t) |b\rangle |n-1\rangle e^{-i\omega_b t} e^{-i(n-1)\omega t} \\ & + C_{f_2}(t) |b\rangle |n+1\rangle e^{-i\omega_b t} e^{-i(n+1)\omega t} \end{aligned} \quad (2.259)$$

where the first term is the initial state of the system, the second term is corresponds to absorption, and the third term corresponds to emission. We assume that  $|\psi(t)\rangle = |a\rangle |n\rangle$ ,  $C_i(0) = 1$ , and  $C_{f_1}(0) = C_{f_2}(0) = 0$ . Following the perturbative method used before, we can get the first-order correction for the amplitudes  $C_{f_1}$  and  $C_{f_2}$  associated with the atom being in state  $|b\rangle$ :

$$\text{Absorption: } C_{f_1}^{(1)}(t) = -\frac{i}{\hbar} \int_0^t dt' \langle f_1 | \hat{\mathcal{H}}^{(I)} | i \rangle e^{i(\omega_{f_1} - \omega_i)t'} \quad (2.260)$$

$$\text{Emission: } C_{f_2}^{(1)}(t) = -\frac{i}{\hbar} \int_0^t dt' \langle f_2 | \hat{\mathcal{H}}^{(I)} | i \rangle e^{i(\omega_{f_2} - \omega_i)t'}. \quad (2.261)$$

And so the amplitude of the atom being in state  $|b\rangle$  (to first-order of course) regardless of how it got there is the *sum* of these amplitudes:

$$C_f^{(1)}(t) = C_{f_1}^{(1)}(t) + C_{f_2}^{(1)}(t) \quad (2.262)$$

$$= \frac{i}{\hbar} (\hat{\mathbf{d}} \cdot \boldsymbol{\mathcal{E}}_0)_{ba} \int_0^t dt' \sqrt{n} e^{i(\omega_{f_1} - \omega_i)t'} - \sqrt{n+1} e^{i(\omega_{f_2} - \omega_i)t'} \quad (2.263)$$

$$= \frac{i}{\hbar} (\hat{\mathbf{d}} \cdot \boldsymbol{\mathcal{E}}_0)_{ba} \left[ \underbrace{\sqrt{n+1} \cdot \frac{e^{i(\omega + \omega_{ba})t} - 1}{\omega + \omega_{ba}}}_{\text{Emission}} - \underbrace{\sqrt{n} \cdot \frac{e^{i(\omega - \omega_{ba})t} - 1}{\omega - \omega_{ba}}}_{\text{Absorption}} \right] \quad (2.264)$$

If  $n \ll 1$  then  $\sqrt{n+1} \approx \sqrt{n}$ , in which case this result and (2.212) are essentially the same and we get a correspondence between classical and quantum field amplitudes:

$$(\mathbf{E}_0)_{\text{classical}} \leftrightarrow (2i\mathcal{E}_0\sqrt{n})_{\text{quantum}}. \quad (2.265)$$

If  $E_b > E_a$  ( $|b\rangle$  is the excited state) and  $\omega \approx \omega_{ba}$  then by RWA the first term in (2.262) can be dropped as the second term dominates. There is nothing abnormal about this. However, if  $E_a > E_b$  ( $|a\rangle$  is the excited state) and  $\omega \approx -\omega_{ba}$ , then the second term of (2.262) can be dropped (again under RWA). This leaves us with

$$C_f^{(1)}(t) \approx \frac{i}{\hbar} (\hat{\mathbf{d}} \cdot \boldsymbol{\mathcal{E}}_0)_{ba} \underbrace{\left[ \sqrt{n+1} \cdot \frac{e^{i(\omega+\omega_{ba})t} - 1}{\omega + \omega_{ba}} \right]}_{\text{Emission}}. \quad (2.266)$$

We notice that even when  $n = 0$ , this does not vanish. In this case, the transition between  $|a\rangle$  and  $|b\rangle$  occurs through spontaneous emission.

### Field-theoretic derivation of the Planck's distribution

Suppose we have a collection of atoms interacting resonantly with a quantized field of frequency

$$\omega = \frac{E_a - E_b}{\hbar} \quad (2.267)$$

where  $|a\rangle$  and  $|b\rangle$  are atomic eigenstates with  $E_a > E_b$ . Let  $N_a, N_b$  be the populations of atoms in  $|a\rangle$  and  $|b\rangle$  respectively. Let  $W_{\text{emis}}$  and  $W_{\text{abs}}$  be the transition rate due to photon emission and absorption respectively. We have the following set of differential equations:

$$\frac{dN_a}{dt} = -N_a W_{\text{emis}} + N_b W_{\text{abs}} \quad (2.268)$$

$$\frac{dN_b}{dt} = -N_b W_{\text{abs}} + N_a W_{\text{emis}}. \quad (2.269)$$

At thermal equilibrium:

$$\frac{dN_a}{dt} = 0 = \frac{dN_b}{dt} \implies N_a W_{\text{emis}} = N_b W_{\text{abs}}. \quad (2.270)$$

From the relative rates in (2.255), we have

$$\frac{N_b}{N_a} = \frac{W_{\text{emis}}}{W_{\text{abs}}} = \frac{n+1}{n}. \quad (2.271)$$

Now, according to Boltzmann statistics:

$$\frac{N_b}{N_a} = e^{(E_a - E_b)/kT} = e^{\hbar\omega/kT}. \quad (2.272)$$

From (2.271) and (2.272) we have

$$\boxed{n = \frac{1}{e^{\hbar\omega/kT} - 1}} \quad (2.273)$$

There's a slightly different derivation by Einstein (which was done even before quantum electrodynamics was invented), which takes into account as well the distinction between spontaneous and stimulated emission. We won't go into the details of this.

### 2.2.4 The Rabi model

In the previous subsections we have used the perturbation theory approach to calculate transition probabilities. This approach assumes that the population in the initial state is essentially unchanged and the population in the final states are very small. This approach fails in cases of large population transfer, for example when there is a strong, near resonance (with one state and no other) laser field. In this case, the perturbation approach must be abandoned, only two dominant states remain, and the problem can be solved “exactly.” This is the **Rabi model**. We note that the Rabi model is a *semiclassical model*. The full quantum mechanical model is covered in the next subsection.

Suppose our two levels are the ground  $|g\rangle$  and excited state  $|e\rangle$ . The transition frequency is given by

$$\omega_0 = \frac{E_e - E_g}{\hbar} \approx \omega \quad (2.274)$$

where  $\omega$  is the frequency of the driving field. Once again, the interaction Hamiltonian is given by

$$\hat{\mathcal{H}}^{(I)} = -\hat{\mathbf{d}} \cdot \mathbf{E}_0 \cos \omega t \equiv \hat{V}_0 \cos \omega t \quad (2.275)$$

where again the dipole approximation is already made. The state vector is a linear combination of the eigenstates:

$$|\psi(t)\rangle = C_g(t)e^{-iE_g t/\hbar}|g\rangle + C_e(t)e^{-iE_e t/\hbar}|e\rangle. \quad (2.276)$$

The Schrödinger equation

$$i\hbar \frac{\partial |\psi(t)\rangle}{\partial t} = (\hat{\mathcal{H}}_0 + \hat{\mathcal{H}}^{(I)}) |\psi(t)\rangle \quad (2.277)$$

$$= (\hat{\mathcal{H}}_0 + \hat{V}_0 \cos \omega t) |\psi(t)\rangle \quad (2.278)$$

gives a system of differential equations:

$$\begin{bmatrix} \dot{C}_g \\ \dot{C}_e \end{bmatrix} = \begin{bmatrix} 0 & -\frac{i}{\hbar} \langle g | \hat{V}_0 | e \rangle e^{-i\omega_0 t} \\ -\frac{i}{\hbar} \langle g | \hat{V}_0 | e \rangle e^{i\omega_0 t} & 0 \end{bmatrix} \begin{bmatrix} C_g \\ C_e \end{bmatrix} \quad (2.279)$$

where the first equation is obtained by multiplying the SE from the left by  $e^{i\omega_g t} \langle g|$  and the fact that  $\langle i|\hat{V}_0|j\rangle = 0$  when  $i = j$ . Next, we assume that  $\langle g|\hat{V}_0|e\rangle$  is real and  $C_g(0) = 1, C_e(0) = 0$ . Let us call

$$\mathcal{V} = \langle g|\hat{V}_0|e\rangle \quad (2.280)$$

and approximate  $\cos \omega t$  in exponentials and keep only the terms that oscillate at frequency  $\omega - \omega_0$  to get

$$\dot{C}_g = -\frac{i}{2\hbar} \mathcal{V} e^{i(\omega - \omega_0)t} C_e \quad (2.281)$$

$$\dot{C}_e = -\frac{i}{2\hbar} \mathcal{V} e^{-i(\omega - \omega_0)t} C_g. \quad (2.282)$$

Taking the derivative of the second equation, plugging it into the first while applying RWA (eliminating terms with  $\omega + \omega_0$ ) we get

$$\ddot{C}_e + i(\omega - \omega_0)\dot{C}_e + \frac{1}{4} \frac{\mathcal{V}^2}{\hbar^2} C_e = 0. \quad (2.283)$$

This differential equation is quite a classic. The general solution is

$$C_e(t) = A_+ e^{i\lambda_+ t} + A_- e^{i\lambda_- t} \quad (2.284)$$

where with  $\Delta = \omega - \omega_0$ ,

$$\lambda_{\pm} = \frac{1}{2} \left[ \Delta \pm \sqrt{\Delta^2 + \frac{\mathcal{V}^2}{\hbar^2}} \right]. \quad (2.285)$$

From the initial conditions, we also have that

$$A_{\pm} = \mp \frac{1}{2\hbar} \frac{\mathcal{V}}{\sqrt{\Delta^2 + \frac{\mathcal{V}^2}{\hbar^2}}}. \quad (2.286)$$

With this, the solution to the problem is

$$C_e(t) = i \frac{\mathcal{V}}{\Omega_R \hbar} e^{i\Delta t/2} \sin \frac{\Omega_R t}{2} \quad (2.287)$$

$$C_g(t) = e^{i\Delta t/2} \left\{ \cos \frac{\Omega_R t}{2} - i \frac{\Delta}{\Omega_R} \sin \frac{\Omega_R t}{2} \right\} \quad (2.288)$$

where

$$\Omega_R = \sqrt{\Delta^2 + \frac{\mathcal{V}^2}{\hbar^2}} \quad (2.289)$$

is called the **Rabi frequency**, or **Rabi rate**.

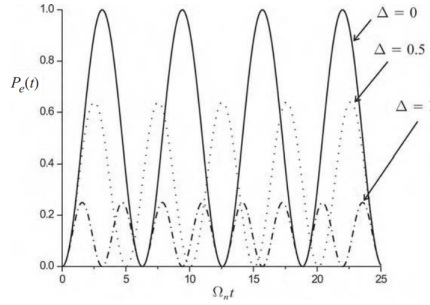
The probability that the atom is in the excited state  $|e\rangle$  is then

$$P_e(t) = |C_e(t)|^2 = \frac{\mathcal{V}^2}{\Omega_R^2 \hbar^2} \sin^2 \frac{\Omega_R t}{2} \quad (2.290)$$

On resonance, this probability is a perfect sine-squared:

$$P_e(t) = \sin^2 \frac{\mathcal{V}t}{2\hbar} \quad (2.291)$$

where at  $t = \pi\hbar/\mathcal{V}$  all the atomic population is transferred to the excited state. The following plot illustrates the population transfer and “oscillation”



The **atomic inversion**  $W(t)$  is defined as the difference in population between the excited and ground states:

$$W(t) = P_e(t) - P_g(t). \quad (2.292)$$

For the case of resonance ( $\Delta = 0$ ) and that initially the atom is in the ground state,

$$W(t) = \sin^2 \frac{\mathcal{V}t}{2\hbar} - \cos^2 \frac{\mathcal{V}t}{2\hbar} = -\cos \frac{\mathcal{V}t}{\hbar} = -\cos \Omega_R(\Delta = 0)t. \quad (2.293)$$

For  $t = \pi\hbar/\mathcal{V}$ ,  $W(\pi\hbar/\mathcal{V}) = 1$ . This kind of population transfer is called the  $\pi$ -pulse. For  $t = \pi\hbar/2\mathcal{V}$ ,  $W(\pi\hbar/2\mathcal{V}) = 0$ , which means the population is shared coherently between the ground and excited states, i.e., the system is in a perfect superposition:

$$|\psi(t)\rangle = \frac{1}{\sqrt{2}} (|g\rangle + i|e\rangle). \quad (2.294)$$

### 2.2.5 Fully quantum-mechanical model; the JaynesCummings model

## 2.3 Collective Atomic Interactions





## Part 3

# QUANTUM INFORMATION



## Part 4

# PROBLEMS & SOLUTIONS

## 4.1 Quantum Mechanics

## 4.2 Quantum Optics

### 4.3 Quantum Information