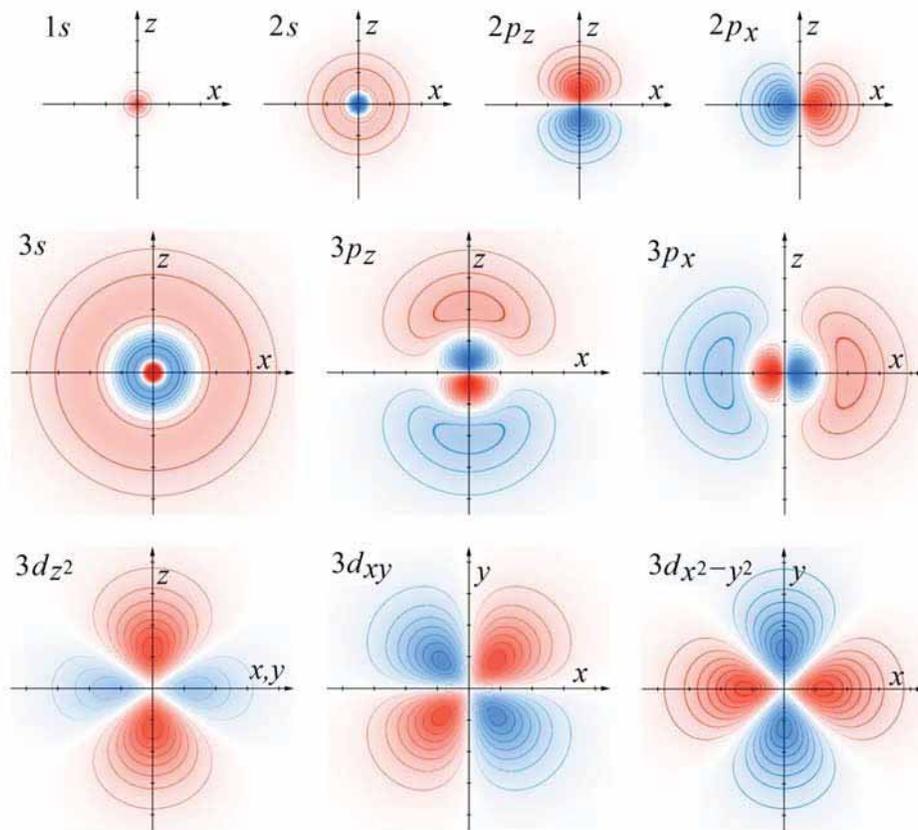


# Intuitive Quantum Mechanics



**C. Julian Chen**

Columbia University in the City of New York

In *A Brief History of Time*, Stephen Hawking said: “Quantum mechanics underlies all of modern science and technology. It governs the behavior of transistors and integrated circuits, and is the basis of modern chemistry and biology”. However, it was difficult to learn. Richard Feynman famously said, “If you think you understand quantum mechanics, you don’t understand quantum mechanics”; because the interpretation of its central subject, wavefunction, is highly controversial.

In 2011, a group of scientists at IBM Zurich Research Laboratory obtained high-resolution images of molecular wavefunctions using a scanning tunneling microscope. Those experiments proved that wavefunction is a directly observable physical reality. It is as concrete as the moons of Jupiter observed by Galileo Galilei, or the cells in biology observed by Robert Hooke. No extra interpretation is needed. Based on experimental facts and following the footsteps of its creator Erwin Schrödinger, the current book makes an intuitive presentation of quantum mechanics as an entry-level college textbook. Color graphics is heavily utilized to visualize wavefunctions. To learn Part I, Elementary Quantum Mechanics, and Part II, Applications to Chemistry and Semiconductor Physics, only high-school physics, chemistry, and advanced-placement calculus are required. After learned Part I and Part II, motivated students can go ahead to read Part III, Complements, to explore and disentangle several advanced topics in quantum mechanics. However, typically, the content in Part III is not a part of an entry-level course.

C. Julian Chen joined IBM T.J. Watson Research Center in 1985 as a Research Staff Member in the Department of Physical Sciences, doing experimental and theoretical research on scanning tunneling microscopy. He authored *Introduction to Scanning Tunneling Microscopy* (Oxford University Press 1993, 2007, 2015), widely recognized as the “Bible” of the field. In 2007, he joined Department of Applied Physics and Applied Mathematics of Columbia University. For ten consecutive years, he teaches a graduate-level course *Physics of Solar Energy*, explaining the quantum physics of solar cells and photosynthesis.

# Contents

|  |           |
|--|-----------|
| <b>Chapter 1: A Review of Classical Physics</b>                | <b>1</b>  |
| 1.1 Classical Mechanics . . . . .                              | 1         |
| 1.1.1 Newton's second law of mechanics . . . . .               | 1         |
| 1.1.2 Conservative Systems and the Energy Integral . . . . .   | 3         |
| 1.1.3 The pendulum . . . . .                                   | 4         |
| 1.1.4 The Kepler problem . . . . .                             | 6         |
| 1.2 Vibration of strings and membranes . . . . .               | 10        |
| 1.2.1 Vibrations of Strings . . . . .                          | 10        |
| 1.2.2 Vibration of membranes: the timpani . . . . .            | 16        |
| 1.3 Light as Electromagnetic Waves . . . . .                   | 22        |
| 1.3.1 Newton's corpuscular theory of light . . . . .           | 22        |
| 1.3.2 Young's double-slit experiment . . . . .                 | 24        |
| 1.3.3 Maxwell's Theory of Electromagnetic Fields . . . . .     | 25        |
| 1.3.4 Fields as a Physical Reality . . . . .                   | 28        |
| 1.3.5 Electromagnetic Waves . . . . .                          | 30        |
| 1.3.6 Polarization of Light . . . . .                          | 32        |
| 1.4 Atomic Spectra . . . . .                                   | 36        |
| <b>Chapter 2: Wave and Quantum</b>                             | <b>41</b> |
| 2.1 Einstein's Energy Quantum of Light . . . . .               | 41        |
| 2.1.1 The photoelectric effect . . . . .                       | 43        |
| 2.1.2 Einstein's law of photoelectrical effect . . . . .       | 45        |
| 2.1.3 Millikan's experimental verification . . . . .           | 45        |
| 2.1.4 Nature of the energy quantum of light . . . . .          | 47        |
| 2.2 Electron as a quantized wave . . . . .                     | 49        |
| 2.2.1 Experimental studies of the cathode ray . . . . .        | 49        |
| 2.2.2 Millikan's oil-drop experiment . . . . .                 | 51        |
| 2.2.3 The de Broglie wave . . . . .                            | 53        |
| <b>Chapter 3: The Schrödinger equation</b>                     | <b>55</b> |
| 3.1 The real Schrödinger equation . . . . .                    | 56        |
| 3.2 Electrons in a one-dimensional box . . . . .               | 58        |
| 3.3 The Dirac notation . . . . .                               | 60        |
| 3.4 Harmonic Oscillator . . . . .                              | 61        |
| 3.4.1 Creation operator and annihilation operator . . . . .    | 61        |
| 3.4.2 Algebraic solution of the Schrödinger equation . . . . . | 62        |

|       |   |    |
|-------|---|----|
| 3.4.3 | Explicit expressions of the wavefunctions . . . . . | 64 |
| 3.5   | The Hydrogen Atom . . . . .                         | 65 |
| 3.5.1 | The Ground State . . . . .                          | 66 |
| 3.5.2 | Spherical harmonics . . . . .                       | 68 |
| 3.5.3 | Energy eigenvalues of excited States . . . . .      | 71 |
| 3.5.4 | Wavefunctions . . . . .                             | 73 |
| 3.5.5 | Degeneracy and Wavefunction Hybridization . . . . . | 76 |

# List of Figures

|      |  |    |
|------|--|----|
| 1.1  | Newton's Principia . . . . .   | 2  |
| 1.2  | The pendulum . . . . .   | 4  |
| 1.3  | Energy conversion of a simple harmonic oscillator . . . . .            | 6  |
| 1.4  | The Kepler Problem . . . . .   | 7  |
| 1.5  | Derivation of the wave equation . . . . .                              | 10 |
| 1.6  | Running waves . . . . .  | 12 |
| 1.7  | Overtone on a string . . . . .   | 14 |
| 1.8  | Overtone on a string . . . . .   | 15 |
| 1.9  | Vibration of a circular membrane . . . . .                             | 16 |
| 1.10 | Bessel functions . . . . .   | 17 |
| 1.11 | Eigenfunctions of the vibration modes . . . . .                        | 19 |
| 1.12 | Nodes in eigenfunctions revealed by Chladni patterns . . . . .         | 20 |
| 1.13 | Linear superposition of degenerate eigenfunctions . . . . .            | 21 |
| 1.14 | Newton's optics and Huyges' Lumiere . . . . .                          | 22 |
| 1.15 | Dispersion of prism and Newton's explanation . . . . .                 | 23 |
| 1.16 | Young's double-slit experiment . . . . .                               | 24 |
| 1.17 | James Clerk Maxwell . . . . .  | 26 |
| 1.18 | Observing and mapping magnetic field . . . . .                         | 29 |
| 1.19 | Observing and mapping electric field . . . . .                         | 30 |
| 1.20 | Electromagnetic wave . . . . .   | 32 |
| 1.21 | The three polarizer experiment: Step 1 . . . . .                       | 34 |
| 1.22 | The three polarizer experiment: Step 2 . . . . .                       | 34 |
| 1.23 | The three polarizer experiment: Step 3 . . . . .                       | 35 |
| 1.24 | The three polarizer experiment: analysis . . . . .                     | 35 |
| 1.25 | A schematic of diffraction grating . . . . .                           | 36 |
| 1.26 | Absorption atomic spectra in the visible range . . . . .               | 37 |
| 1.27 | Emission spectra of hydrogen . . . . .                                 | 38 |
|      |  |    |
| 2.1  | Einstein's paper on energy quantum of light . . . . .                  | 42 |
| 2.2  | Lenard's apparatus for studying photoelectric effect . . . . .         | 43 |
| 2.3  | Result of Millikan's experiment on photoelectric effect . . . . .      | 46 |
| 2.4  | Albert Einstein and Robert Millikan . . . . .                          | 47 |
| 2.5  | Paradox of Dirac's three-polarizer experiment . . . . .                | 48 |
| 2.6  | Paradox of the double-slit experiment . . . . .                        | 49 |
| 2.7  | Schematics of J. J. Thomson's experiment . . . . .                     | 50 |
| 2.8  | Schematics of Millikan's oil-drop experiment . . . . .                 | 52 |
| 2.9  | Louis de Broglie . . . . .   | 53 |
|      |  |    |
| 3.1  | Austrian banknote with a portrait of Schrödinger . . . . .             | 55 |
| 3.2  | First page of Schrödinger's first paper on quantum mechanics . . . . . | 56 |

|      |  |    |
|------|--|----|
| 3.3  | Wavefunctions in a one-dimensional box . . . . .                   | 58 |
| 3.4  | Energy levels and wavefunctions of a harmonic oscillator . . . . . | 65 |
| 3.5  | Hydrogen atom in spherical polar coordinates . . . . .             | 65 |
| 3.6  | Wavefunction of ground-state hydrogen atom . . . . .               | 67 |
| 3.7  | Wavefunctions of excited-states of hydrogen atom . . . . .         | 73 |
| 3.8  | Hydrogen wavefunctions . . . . .                                   | 75 |
| 3.9  | Hybrid <i>sp1</i> wavefunctions . . . . .                          | 76 |
| 3.10 | Hybrid <i>sp2</i> wavefunctions . . . . .                          | 77 |
| 3.11 | Hybrid <i>sp3</i> wavefunctions . . . . .                          | 78 |

# List of Tables

|     |  |    |
|-----|--|----|
| 1.1 | Vibration modes on a string . . . . .              | 14 |
| 1.2 | Zeros of Bessel functions . . . . .                | 18 |
| 1.3 | Quantities in Maxwell's Equations . . . . .        | 27 |
| 1.4 | Wavelengths of hydrogen lines . . . . .            | 37 |
| 2.1 | Stopping Voltage for Photocurrent . . . . .        | 44 |
| 3.1 | Wavefunctions of the harmonic oscillator . . . . . | 63 |
| 3.2 | Spherical Harmonics . . . . .                      | 70 |
| 3.3 | Wavefunctions of the hydrogen atom . . . . .       | 74 |



# Chapter 1

## A Review of Classical Physics

Classical physics dominated physics up to the year of 1900. Quantum mechanics, developed during the first 30 years of the twentieth century, became the most successful and most useful theory of physics. Despite of the importance of quantum mechanics to the understanding of the microscopic world, classical physics is still the basics of our understanding of the macroscopic world. Furthermore, to make an intuitive understanding of quantum mechanics, concepts in classical physics are still essential.

### 1.1 Classical Mechanics

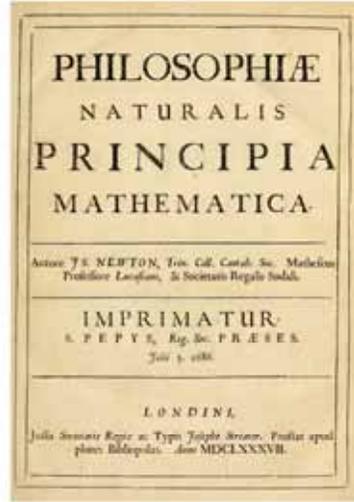
The publication of *Mathematical Principle of Natural Philosophy* by Isaac Newton in 1686 marked the beginning of modern science. In this monumental monograph, among other items, Newton formulated the three laws of mechanics and the law of universal gravity, then explained the Kepler's laws of the motions of planets and moons in the solar system. For several hundreds of years, Newton's mechanics was synonymous to physical science. According to Albert Einstein,

In accordance with Newton's system, physical reality is characterized by concepts of space, time, the material points and force (interaction between material points). Physical events are to be thought of as movements according to the law of material points in space. The material point is the only representative of reality so far as it is subject to change. The concept of material points is obviously due to observable bodies; one conceived of the material point in the analogy of movable bodies by omitting characteristics of extension, form, spatial locality, and all their 'inner' qualities., retaining only inertia, translation, and additional concept of force. ... . All happening was to be conceived of as purely mechanical, that is, merely as motions of material points according to the Newton's laws of motion.

#### 1.1.1 Newton's second law of mechanics

The core of classical mechanics is Newton's second law. A material point is characterized by an intrinsic value of mass  $m$ , a measure of its inertia. At

**Fig. 1.1. Newton's Principia.** In 1686, Isaac Newton published his monumental monograph, *Mathematical Principles of Natural Philosophy*. He defined three laws of mechanics, the law of gravitation, and explained the motion of planets and satellites of the planets among a large number of other subjects. Newton described the material world as composes of material points, each has a well-defined geometrical location and well-defined speed at any given time.



any well-defined time instant  $t$ , a material point has a well-defined position, represented a vector  $\mathbf{r} = (x, y, z)$  in three-dimensional space, and a well-defined velocity  $\mathbf{v}$  as the time derivative of  $\mathbf{r}$ ,

$$\mathbf{v} = \frac{d\mathbf{r}}{dt}. \quad (1.1)$$

And there is a force  $\mathbf{F}$  acting on that material point. Newton's second law states that at any time, the rate of change of its velocity is proportional to the force acting on it and inversely proportional to its mass:

$$\mathbf{F} = m \frac{d\mathbf{v}}{dt}. \quad (1.2)$$

According to Newton's laws, the physical world is *deterministic*: At any time, of the positions and the velocities of the material points and the laws of force are known, the system will evolve precisely according to the second law of mechanics, Eq. 1.2.

By defining a *momentum* as the product of the mass and the velocity of the material point,

$$\mathbf{p} = m\mathbf{v}, \quad (1.3)$$

Newton's second law can be written in a more compact form,

$$\mathbf{F} = \frac{d\mathbf{p}}{dt}. \quad (1.4)$$

Because mass  $m$  is an invariable property of a material point, the validity of Eq. 1.4 is obvious.

### 1.1.2 Conservative Systems and the Energy Integral

In all the problems we treat in this book, the force can be expressed as a gradient of a scalar function of the coordinates, the *potential*  $V(\mathbf{r})$ ,

$$\mathbf{F} = -\nabla V(\mathbf{r}). \quad (1.5)$$

A necessary and sufficient condition for a conservative system is that the curl of force field is zero,

$$\nabla \times \mathbf{F} = 0. \quad (1.6)$$

Newton's second law is then

$$\frac{d\mathbf{p}}{dt} + \nabla V(\mathbf{r}) = 0. \quad (1.7)$$

By multiplying both sides of Eq. 1.7 with  $d\mathbf{r}$ , and integrate from  $\mathbf{r}_1$  to  $\mathbf{r}_2$ , notice that

$$d\mathbf{r} = \frac{d\mathbf{r}}{dt} dt = \frac{1}{m} \mathbf{p} dt, \quad (1.8)$$

we have

$$\int_{\mathbf{r}_1}^{\mathbf{r}_2} \left[ \frac{1}{m} \mathbf{p} \frac{d\mathbf{p}}{dt} dt + dV(\mathbf{r}) \right] = \left[ \frac{1}{2m} \mathbf{p}^2 + V(\mathbf{r}) \right]_{\mathbf{r}_1}^{\mathbf{r}_2}. \quad (1.9)$$

The following identity is obtained:

$$\left[ \frac{1}{2m} \mathbf{p}^2 + V(\mathbf{r}) \right]_{\mathbf{r}=\mathbf{r}_1} = \left[ \frac{1}{2m} \mathbf{p}^2 + V(\mathbf{r}) \right]_{\mathbf{r}=\mathbf{r}_2}. \quad (1.10)$$

The expression in the square bracket is independent of position and time. It is the total energy  $E$  of the system, which is a constant, only depends on the initial condition:

$$\frac{\mathbf{p}^2}{2m} + V(\mathbf{r}) = E. \quad (1.11)$$

The first term is called the *kinetic energy* of the system,

$$T = \frac{\mathbf{p}^2}{2m}. \quad (1.12)$$

Equation 1.11 can be written in a conceptually simpler form, such that the total energy is the sum of kinetic energy and potential energy,

$$T + V = E. \quad (1.13)$$

In the following subsections, we will discuss two cases in detail, both are related to the understanding of quantum mechanics.

### 1.1.3 The pendulum

The pendulum, shown in Fig. 1.2, is a classical prototype of the harmonic oscillator in quantum mechanics. A material particle with mass  $m$  is hung with a flexible string of length  $L$ . A gravitational force  $mg$  is acting on the particle. A component of the gravitational force  $F$  drives the particle to its equilibrium position **b**. Based on elementary geometry, the component force  $F$  can be estimated as follows. The triangle marked light green is similar to the triangle marked yellow. The sides are proportional:

$$\frac{F}{mg} = -\frac{x}{L}. \quad (1.14)$$

The negative sign means that the direction of  $F$  and  $x$  are opposite. If displacement  $x$  is much smaller than length  $L$ , the arc  $\widehat{bc}$  is practically identical to the horizontal line  $x$ . To resolve this problem according to Newton's law, we apply the energy integral, Eq. 1.11 in the previous subsection. The potential energy function, by definition, is

$$V(x) = -\int_0^x F dx = \int_0^x \frac{mg}{L} x dx = \frac{mg}{2L} x^2. \quad (1.15)$$

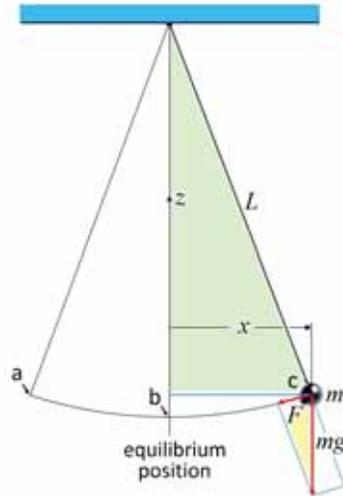
The momentum is a scalar,

$$p = m \frac{dx}{dt}. \quad (1.16)$$

There, the energy integral is

$$\frac{p^2}{2m} + \frac{mg}{2L} x^2 = E. \quad (1.17)$$

**Fig. 1.2. The pendulum.** In 1602, by watching the motion of a chandelier hanging in a cathedral, Galileo Galilei started to study the motions of pendulum. He discovered that the pendulum has a constant period, independent of the amplitude and the angle. The period only depends on the length  $L$  of the string, but independent of the weight  $m$ . In 1657, Christiaan Huygens invented a mechanism to sustain the vibration of the pendulum. The mechanical clock was born. It remained to be the most accurate clock up to early twentieth century.



Using Eq. 1.16, it can be written in a convenient form

$$\left(\frac{dx}{dt}\right)^2 + \frac{g}{L}x^2 = \frac{2}{m}E. \quad (1.18)$$

When the particle moves to the extreme positions,  $x = a$  or  $x = -a$ , the kinetic energy is zero. The potential energy equals the total energy. Using the amplitude parameter  $a$ , Eq. 1.18 can be written as

$$\left(\frac{dx}{dt}\right)^2 = \frac{g}{L}(a^2 - x^2). \quad (1.19)$$

By taking a square root, the equation becomes directly integrable:

$$\int \frac{dx}{\sqrt{a^2 - x^2}} = \sqrt{\frac{g}{L}} \int dt. \quad (1.20)$$

Defining an *angular frequency*

$$\omega \equiv \sqrt{\frac{g}{L}}, \quad (1.21)$$

the integration gives

$$\arcsin \frac{x}{a} = \omega t + \phi, \quad (1.22)$$

where the constant of integration  $\phi$  is a phase angle. In other words,

$$x = a \sin(\omega t + \phi). \quad (1.23)$$

The pendulum makes a simple harmonic oscillation. Because the period of a sine function is  $2\pi$ , the period of the pendulum is

$$\tau = \frac{2\pi}{\omega} = 2\pi \sqrt{\frac{L}{g}}, \quad (1.24)$$

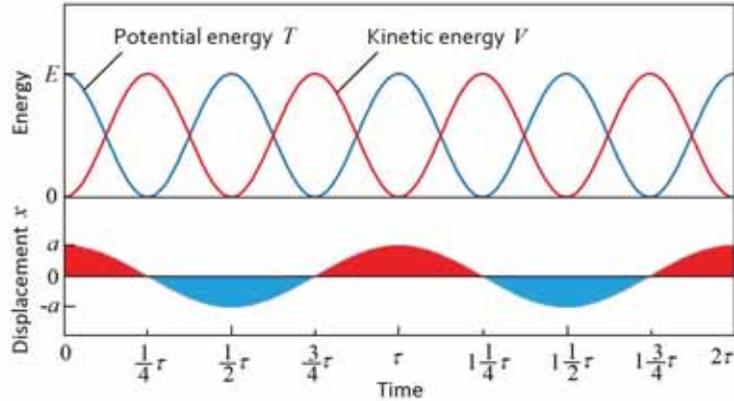
and the frequency is

$$f = \frac{\omega}{2\pi} = \frac{1}{2\pi} \sqrt{\frac{g}{L}}. \quad (1.25)$$

Using angular frequency as a parameter, the energy integral, Eq. 1.17, can be written in a more general form,

$$T + V = \frac{1}{2m}p^2 + \frac{m}{2}\omega^2 x^2 = E. \quad (1.26)$$

Figure 1.3 shows the simple harmonic oscillation of a pendulum, and the process of energy conversion. At  $t = 0$ , the material particle is at its right-most position,  $x = a$ . The potential energy is at its maximum, equals



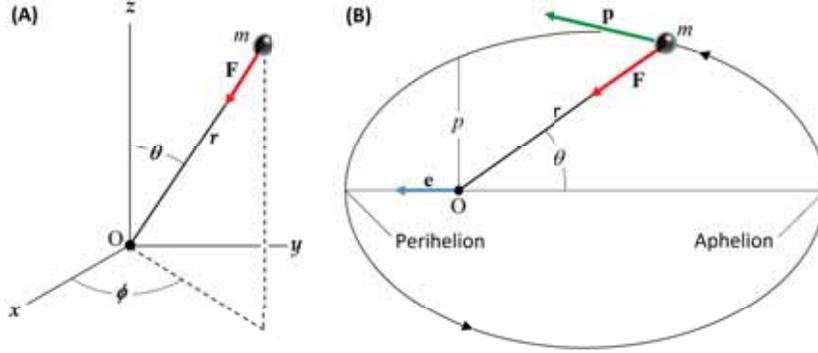
**Fig. 1.3.** Energy conversion of a simple harmonic oscillator. At time 0, the pendulum is at one of its extreme positions, the potential energy is at a maximum, and the kinetic energy is zero. At a quarter of a period  $\tau$ , potential energy converts into kinetic energy. The process goes on.

the entire total energy. The kinetic energy is zero. In a quarter of a period,  $t = \tau/4$ , the particle moves to the equilibrium position  $x = 0$ . At that time, the kinetic energy reaches its maximum, but the potential energy reaches its minimum. Because of inertia, the material particle continues its motion to the left side,  $x < 0$ . After another quarter period, the particle reaches its left most position and stops. The potential energy again reaches its maximum, whereas the kinetic energy becomes zero. In the third quarter of a period, the particle moves in the positive direction, and regains speed. After passing the equilibrium position  $x = 0$ , the particle moves continuously to its original position,  $x = a$ , where the kinetic energy becomes zero and the potential energy reaches its maximum, that is, the total energy  $E$ . Thus the particle completes a full period  $\tau$ .

#### 1.1.4 The Kepler problem

A central problem in Newtonian mechanics is the Kepler problem, where a planet is attracted by the Sun. The greatest achievement of Isaac Newton was the interpretation of Kepler's laws using his laws of mechanics and universal gravitation. It is also a central problem in quantum mechanics as a model of the hydrogen atom. The greatest achievement of Erwin Schrödinger is the explanation of the Rydberg formula using quantum mechanics, see Sections 1.4 and 3.5. The Kepler problem is schematically shown in Fig. 1.4(A), in both rectangular coordinate system and polar coordinate system. The potential function is,

$$V(\mathbf{r}) = \frac{K}{r}, \quad (1.27)$$



**Fig. 1.4. The Kepler Problem.** (A), in Cartesian coordinate system. (B), in polar coordinate system. A planet of mass  $m$  is attracted by the Sun according to Newton's inverse-square law of gravitation. The eccentricity vector  $\mathbf{e}$  always points to the perihelion, where the planet is closest to the Sun. Thus  $\mathbf{e}$  is a constant of motion.

where  $K$  is a constant. According to Eq. 1.3, the force is

$$\mathbf{F} = -\frac{K \mathbf{r}}{r^2 r}. \quad (1.28)$$

The direction of the force is towards the Sun, marked as O, and the magnitude is inversely proportional to the distance of the Sun and the planet  $r$ . Newton's equation is

$$\frac{d\mathbf{p}}{dt} = -\frac{K \mathbf{r}}{r^2 r}. \quad (1.29)$$

Following Eq. 1.11, we can write down the energy integral,

$$\frac{1}{2m} \mathbf{p}^2 + \frac{K}{r} = E, \quad (1.30)$$

where  $m$  is the mass of the planet. Because the potential energy  $V$  only depends on  $r$ , it is natural to use polar coordinates. By definition,

$$\begin{aligned} x &= r \sin \theta \cos \phi, \\ y &= r \sin \theta \sin \phi, \\ z &= r \cos \theta. \end{aligned} \quad (1.31)$$

Using Eq. 1.31, after a simple but lengthy algebraic manipulation, the quantity  $\mathbf{p}^2$  in Eq. 1.31 becomes

$$\begin{aligned} \mathbf{p}^2 &= m^2 \left[ \left( \frac{dx}{dt} \right)^2 + \left( \frac{dy}{dt} \right)^2 + \left( \frac{dz}{dt} \right)^2 \right] \\ &= m^2 \left[ \left( \frac{dr}{dt} \right)^2 + r^2 \left( \frac{d\theta}{dt} \right)^2 + r^2 \sin^2 \theta \left( \frac{d\phi}{dt} \right)^2 \right]. \end{aligned} \quad (1.32)$$

Defining three momentum variables in polar coordinates,

$$p_r = m \frac{dr}{dt}, \quad (1.33)$$

$$p_\theta = mr^2 \frac{d\theta}{dt}, \quad (1.34)$$

and

$$p_\phi = mr^2 \sin^2 \theta \frac{d\phi}{dt}, \quad (1.35)$$

the energy integral, Eq. 1.31 becomes

$$\frac{p_r^2}{2m} + \frac{p_\theta^2}{2mr^2} + \frac{p_\phi^2}{2mr^2 \sin^2 \theta} + \frac{K}{r} = E. \quad (1.36)$$

Because the potential energy only depends on  $r$ , the momenta  $p_\theta$  and  $p_\phi$  are constant. The problem can be resolved analytically.

The most elegant treatment of that problem is by the Hamilton-Jacobi method, which was the basis of the first 1926 paper of Erwin Schrödinger and the phase-integral approach of Richard Feynman, the topic of Chapter 12. An elementary treatment, based on vector algebra, is also interesting. It was the basis of an algebraic solution to the hydrogen atom problem by Wolfgang Pauli, to be presented in Chapter 11. It relies on a vector constant of motion, the *eccentricity vector*  $\mathbf{e}$ ,<sup>1</sup> see Fig 1.4(B).

As a consequence of the inverse square law of gravity, the eccentricity vector  $\mathbf{e}$  is invariant. Here is a proof.

Without losing generality, we assume the planet moves in the  $xy$ -plane. The angular momentum, a vector in the  $z$ -direction, can be written as

$$\mathbf{L} = \mathbf{r} \times \mathbf{p} = m\mathbf{r} \times \frac{d\mathbf{r}}{dt}. \quad (1.37)$$

As a consequence of Eq. 1.29, the angular momentum is a constant,

$$\frac{d\mathbf{L}}{dt} = m \frac{d\mathbf{r}}{dt} \times \frac{d\mathbf{r}}{dt} + \mathbf{r} \times \left( -\frac{K}{r^2} \frac{\mathbf{r}}{r} \right) = 0. \quad (1.38)$$

Consider the time evolution of a vector  $\mathbf{p} \times \mathbf{L}$ . Because  $\mathbf{L}$  is a constant, only the time evolution of  $\mathbf{p}$  has to be counted. Using Eq. 1.29, as well as the following identities,

$$\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) = \mathbf{b}(\mathbf{a} \cdot \mathbf{c}) - \mathbf{c}(\mathbf{a} \cdot \mathbf{b}) \quad (1.39)$$

and

$$\mathbf{r} \cdot \frac{d\mathbf{r}}{dt} = r \frac{dr}{dt}, \quad (1.40)$$

---

<sup>1</sup>In some literature, it was also called the Laplace vector or the Runge-Lenz vector.

we obtain

$$\begin{aligned}
 \frac{d}{dt}(\mathbf{p} \times \mathbf{L}) &= -\frac{K}{r^2} \frac{\mathbf{r}}{r} \times \left( m\mathbf{r} \times \frac{d\mathbf{r}}{dt} \right) \\
 &= \frac{mK}{r^3} \left[ \mathbf{r} \left( \mathbf{r} \cdot \frac{d\mathbf{r}}{dt} \right) - r^2 \frac{d\mathbf{r}}{dt} \right] \\
 &= mK \left[ \frac{1}{r} \frac{d\mathbf{r}}{dt} - \frac{\mathbf{r}}{r^2} \frac{dr}{dt} \right] \\
 &= mK \frac{d}{dt} \left[ \frac{\mathbf{r}}{r} \right].
 \end{aligned} \tag{1.41}$$

Therefore, the eccentricity vector

$$\mathbf{e} \equiv \frac{\mathbf{p} \times \mathbf{L}}{mK} - \frac{\mathbf{r}}{r} \tag{1.42}$$

is a constant. It points to the position of minimum distance  $r$ , called by astronomers as *perihelion*, the closest point to the Sun.

Using the constant vector  $\mathbf{e}$ , the trajectory of Kepler motion can be obtained. Denoting the magnitude of the eccentricity vector as  $e$ , and taking the aphelion point as the origin with  $\theta = 0$ , we have

$$\mathbf{r} \cdot \mathbf{e} = -er \cos \theta. \tag{1.43}$$

On the other hand, from Eq. 1.42,

$$\mathbf{r} \cdot \mathbf{e} = \frac{\mathbf{r} \cdot (\mathbf{p} \times \mathbf{L})}{mK} - r = \frac{\mathbf{L} \cdot (\mathbf{r} \times \mathbf{p})}{mK} - r = \frac{L^2}{mK} - r, \tag{1.44}$$

where  $L$  is the magnitude of angular momentum. The solution is

$$r = \frac{L^2}{mK} \frac{1}{1 - e \cos \theta}. \tag{1.45}$$

Comparing with Fig 1.4, we have the *semilatus rectum*

$$p = \frac{L^2}{mK}, \tag{1.46}$$

the standard formula of the ellipse is obtained

$$r = \frac{p}{1 - e \cos \theta}, \tag{1.47}$$

which explains the origin of the term *eccentricity vector*.

## 1.2 Vibration of strings and membranes

The images and properties of wavefunctions in quantum mechanics are very similar to the acoustic waves in everyday tangible world. The mathematics is almost identical. Therefore, the vibrations of strings and membranes are instructive for an intuitive understanding of wavefunctions.

In quantum mechanics, the concepts of superposition and orthogonality are essential. Those concepts are intuitive and obvious in terms of acoustic waves. By familiarizing with the examples in acoustic waves, similar concepts in quantum mechanics can be easily understood.

### 1.2.1 Vibrations of Strings

The phenomena we are discussing here are related to any string instruments, for example, guitar, violin, cello, and piano. By doing experiments on those string instruments, one can make direct observations.

Figure 1.5 shows a derivation of the wave equation. Consider a small section of a string, between  $x$  and  $x + \Delta x$ . The lateral displacement  $u(x, t)$  is a function of  $x$  and time  $t$ . A tension  $T$  is applied on both sides. As shown in Fig. 1.5, the lateral force acting on the small section is

$$F = T \frac{\partial u}{\partial x}(x + \Delta x) - T \frac{\partial u}{\partial x}(x) \approx T \frac{\partial^2 u}{\partial x^2} \Delta x. \quad (1.48)$$

The mass of the small section is  $\rho \Delta x$ . According to Newton's law,

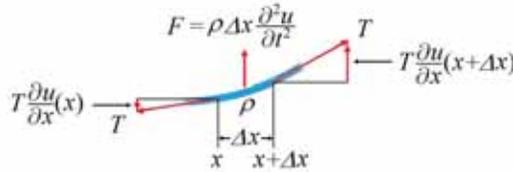
$$F = \rho \Delta x \frac{\partial^2 u}{\partial t^2}. \quad (1.49)$$

Combining Eqs. 1.48 and 1.49, we obtain

$$\frac{\partial^2 u}{\partial t^2} = \frac{T}{\rho} \frac{\partial^2 u}{\partial x^2}. \quad (1.50)$$

By denoting  $v = \sqrt{T/\rho}$ , Eq. 1.50 is brought to a standard form,

$$\frac{\partial^2 u}{\partial t^2} = v^2 \frac{\partial^2 u}{\partial x^2}. \quad (1.51)$$



**Fig. 1.5. Derivation of the wave equation.** The lateral displacement  $u(x, t)$  is a function of  $x$  and time  $t$ . A tension  $T$  is applied on both sides of the small section  $\Delta x$ . The wave equation is a consequence of Newton's law.

Here we show that  $v$  is the velocity of sound. On a string of infinite length, the general solution of the wave equation Eq. 1.51 was obtained by French mathematician and physicist Jean le Rond d'Alembert in 1747,

$$u(x, t) = F(x - vt) + G(x + vt), \quad (1.52)$$

where  $F(x)$  and  $G(x)$  are two independent, arbitrary functions. The solution can be proved by direct substitution. On one hand,

$$\frac{\partial u(x, t)}{\partial t} = v F'(x - vt) - v G'(x + vt), \quad (1.53)$$

thus

$$\frac{\partial^2 u(x, t)}{\partial t^2} = v^2 F''(x - vt) + v^2 G''(x + vt). \quad (1.54)$$

On the other hand,

$$\frac{\partial^2 u(x, t)}{\partial x^2} = F''(x - vt) + G''(x + vt). \quad (1.55)$$

Therefore, the d'Alembert solution satisfies the wave equation, Eq. 1.50. It is a combination of a wave  $F(x - vt)$  propagating in  $+x$  direction at velocity  $v$ , and a wave  $G(x + vt)$  propagating in  $-x$  direction at velocity  $-v$ .

A special case of the d'Alembert solution is the sinusoidal wave. If the frequency is  $f$ , the variation of displacement with time is

$$u(x, t) \sim \sin(2\pi ft + \phi), \quad (1.56)$$

where  $\phi$  is the phase. The factor  $2\pi$  occurs frequently. It is convenient to introduce an *angular frequency* to eliminate it:

$$\omega \equiv 2\pi f. \quad (1.57)$$

Equation 1.56 is simplified to

$$u(x, t) \sim \sin(\omega t + \phi). \quad (1.58)$$

Just as there are two d'Alembert solutions moving into two opposite directions, there could have two sinusoidal waves moving into opposite directions: A wave moving in  $+x$  direction is,

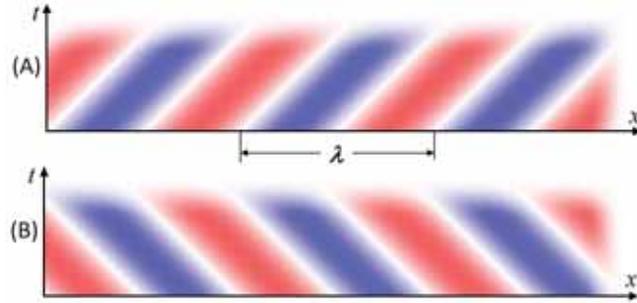
$$u(x, t) = a \sin(kx - \omega t + \phi), \quad (1.59)$$

and a wave moving in  $-x$  direction is,

$$u(x, t) = a \sin(kx + \omega t + \phi), \quad (1.60)$$

where  $k$  is the *wave number*, which has a dimension of  $L^{-1}$ :

$$k = \frac{\omega}{v}. \quad (1.61)$$



**Fig. 1.6. Running waves.** (A) A running wave in positive  $x$  direction, Eq. 1.59. (B) A running wave in negative  $x$  direction, Eq. 1.60. Phase of wave is marked by colors; for example, red as positive, blue as negative. Intensity indicates amplitude. A complete period in length scale  $\lambda$  is a *wavelength*, which contains a positive half-wavelength marked red, and a negative half-wavelength marked blue.

Figure 1.6 shows the waves propagating in  $+x$  direction and in  $-x$  direction. The phases are marked by color. For example, red indicates positive, and blue indicates negative. The intensity of color indicates amplitude. The length of a complete period is a *wavelength*, with a symbol  $\lambda$ . By definition,

$$\lambda = \frac{v}{f} = \frac{2\pi v}{\omega} = \frac{2\pi}{k}. \quad (1.62)$$

### Principle of superposition and interference

The wave equation is linear. If a wave  $f_1(x, t)$  is a solution of a wave equation, Eq. 1.51, and another wave  $f_2(x, t)$  is also a solution, then any linear superposition of the two waves

$$f(x, t) = c_1 f_1(x, t) + c_2 f_2(x, t), \quad (1.63)$$

is also a solution of the same wave equation, where  $c_1$  and  $c_2$  are arbitrary constants. This statement can readily be verified by inserting the expression of the new wave into Eq. 1.51.

The superposition of waves gives rise to the *interference* of waves. Consider two sinusoidal waves of the same amplitude but different phase,

$$u_1(x, t) = a \sin(kx - \omega t + \phi_1), \quad (1.64)$$

and

$$u_2(x, t) = a \sin(kx - \omega t + \phi_2), \quad (1.65)$$

The sum is

$$\begin{aligned}
 u(x, t) &= u_1(x, t) + u_2(x, t), \\
 &= a \sin(kx - \omega t + \phi_1) + a \sin(kx - \omega t + \phi_2) \\
 &= 2a \sin\left(kx - \omega t + \frac{\phi_1 + \phi_2}{2}\right) \cos\left(\frac{\phi_1 - \phi_2}{2}\right).
 \end{aligned} \tag{1.66}$$

The result depends on the phase difference. If the phase difference is an integer multiple of  $2\pi$ , there is a *positive interference*. The amplitude is doubled, thus the power is quadrupled. If the phase difference is an odd integer multiple of  $\pi$ , there is a *negative interference*. The amplitude is zero. The power vanishes.

The interference phenomenon is unique to waves. By combining two beams of particles, the energy of the composite beam is the simple addition of the individual beams. For waves, depending on the relative phase, the energy of the composite beam could be much greater than the simple sum, or can be mutually cancelled, see Section 1.3.2.

The principle of superposition is valid for all linear differential equations, including Schrödinger's equation. Superposition in acoustic phenomena is easily visualized, which is helpful for the understanding of superposition in quantum mechanics.

### Standing wave and the Helmholtz equation

To describe the vibration of a string with both ends fixed at a fixed frequency, the standard way is to write the displacement  $u(x, t)$  as a product of a function of  $x$  and a sinusoidal function of time,

$$u(x, t) = u(x) \sin(\omega t + \phi). \tag{1.67}$$

Insert Eq. 1.67 into Eq. 1.51, we find a differential equation for  $u(x)$ :

$$\frac{d^2 u(x)}{dt^2} = -\frac{\omega^2}{v^2} u(x). \tag{1.68}$$

Using the wave vector  $k$  defined in Eq. 1.61, Eq. 1.68 becomes

$$\frac{d^2 u(x)}{dx^2} = -k^2 u(x), \tag{1.69}$$

which is called a *Helmholtz Equation*.

### Eigenvalues and eigenfunctions

On string instruments, the strings are clamped at both ends, see Fig. 1.7. Let the ends be  $x = 0$  and  $x = L$ . The values of displacement  $u(x)$  must be zero at both ends. The condition limits the values of  $k$  and  $\omega$  in Eqs. 1.68



**Fig. 1.7. Overtones on a string.** The experiment was performed on the string of note E1 on a grand piano. By lightly touching various points on the string and strike it, the overtones of the basic frequency of the string are excited. See Table 1.1 and Fig. 1.8.

and 1.69, and also limits the waveform  $u(x)$ . The allowed values for  $k$  and  $\omega$  under the boundary conditions are called *eigenvalues*, and the allowed waveforms are called *eigenfunctions*. The collection of eigenvalues and eigenfunctions are called *vibration modes* of the system.

First, at  $x = 0$ , the string is fixed. A solution of the Helmholtz equation, Eq. 1.69, that is zero at  $x = 0$  should be

$$u(x) = a \sin kx. \quad (1.70)$$

At  $x = L$ , the string is also fixed. That boundary condition,  $u(x) = 0$  at  $x = L$ , requires the eigenfunctions to be:

$$u(x) = a \sin\left(\frac{n\pi}{L}x\right), \quad n = 1, 2, 3, \dots \quad (1.71)$$

The wavevector eigenvalues are

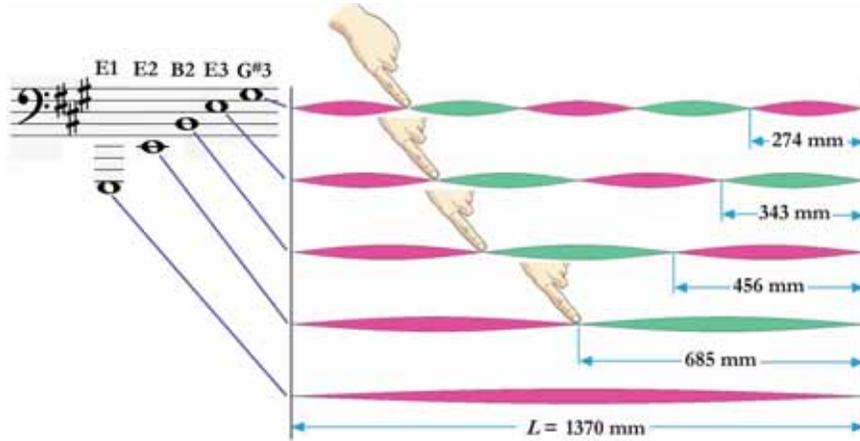
$$k_n = \frac{n\pi}{L}, \quad n = 1, 2, 3, \dots \quad (1.72)$$

Accordingly, the frequency eigenvalues are

$$f_n = \frac{\omega_n}{2\pi} = \frac{vk_n}{2\pi} = \frac{nv}{2L}, \quad n = 1, 2, 3, \dots \quad (1.73)$$

**Table 1.1: Vibration modes on a string**

|                                 |      |      |       |       |       |
|---------------------------------|------|------|-------|-------|-------|
| Length (mm)                     | 1370 | 685  | 456   | 343   | 274   |
| Wave vector ( $\text{m}^{-1}$ ) | 2.29 | 4.57 | 6.88  | 9.17  | 11.46 |
| Frequency (Hz)                  | 41.2 | 82.4 | 123.6 | 164.8 | 206.0 |
| Note name                       | E1   | E2   | B2    | E3    | G#3   |



**Fig. 1.8. Overtones on a string.** By lightly touching various points on a string and strike it with the hammer, the overtones are excited.

The  $n$ -th standing-wave solution is

$$u_n(x, t) = a_n \sin\left(\frac{n\pi}{L}x\right) \sin\left(\frac{nv}{2L}t + \phi\right). \quad (1.74)$$

### Nodes and overtones

Table 1.1 shows the vibration modes of an E1 string on a grand piano with overtones. The total length of the string is 1370 mm. By lightly touching the middle point of the string at 685 mm, the first overtone E2 with frequency 82.4 Hz is excited. By lightly touching the point of one fifth of the string at 274 mm, the fifth overtone G#3 with frequency 206 Hz is excited. The points with no displacement are called a *node*, and the collection of the nodal points are called a *nodal pattern*.

### Orthogonality of eigenfunctions

An important fact of the eigenfunctions is, for vibration modes of different eigenfrequencies, the eigenfunctions are *orthogonal*:

$$\int_0^L u_n(x) u_m(x) dx = 0, \quad n \neq m. \quad (1.75)$$

From Eq. 1.74, it is obvious that unless  $m = n$ , the integral is zero.

Another interesting fact worth noting is, for vibration modes of different eigenfrequencies, the number of nodes are different. In fact, the number of nodes for the five vibration modes are 0, 1, 2, 3, and 4, respectively.

### 1.2.2 Vibration of membranes: the timpani

The vibration modes on a circular membrane show more resemblance to the quantum states in atoms. An example is the timpani, a key percussion instrument in a symphonic orchestra, see Fig. 1.9 (A). A membrane, called a head, is stretched across the opening of a bowl. The tension of the membrane can be adjusted by a number of screws, called tension rods. During playing, the tension, and consequently the frequency, can be temporarily adjusted by a pedal. The typical diameter is 50 to 80 cm. In the 19th century, there were a number of studies on its mechanism, and was described in detail in Lord Rayleigh's classical treatise *Theory of Sound*.

The wave equation of a membrane is similar to that for a string. In Cartesian coordinates, a similar argument would lead to

$$\frac{\partial^2 u}{\partial t^2} = \frac{T}{\rho} \left[ \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \right] = v^2 \left[ \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \right]. \quad (1.76)$$

Here the tension  $T$  is the force per unit length, and the density  $\rho$  is mass per unit area fixed frequency, following the same procedure as the one-dimensional case, using

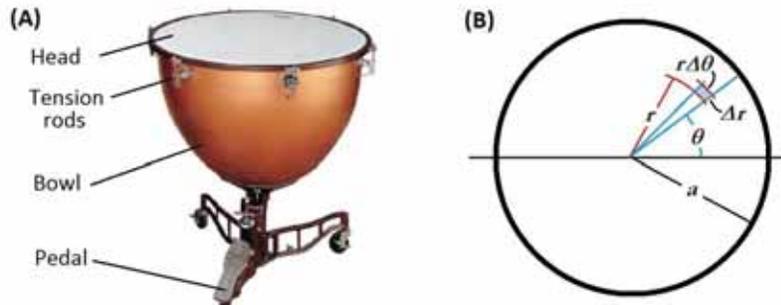
$$u = u(x, y) \sin(\omega t + \phi), \quad (1.77)$$

we find the Helmholtz equation

$$\frac{\partial^2 u(x, y)}{\partial x^2} + \frac{\partial^2 u(x, y)}{\partial y^2} = -k^2 u(x, y), \quad (1.78)$$

Following Eq. 1.61,

$$k = \frac{\omega}{v}. \quad (1.79)$$



**Fig. 1.9. Vibration of a circular membrane.** (A), timpani, a leading percussion instrument in an orchestra. For example, in the first movement of Beethoven's violin concerto and the second movement of his ninth symphony, the timpani takes a prominent role. (B) polar coordinates for the analysis of the circular membrane.

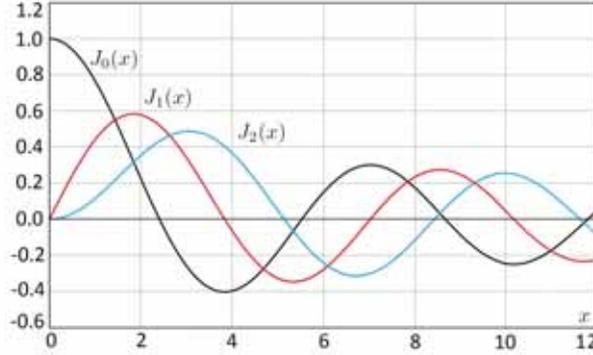


Fig. 1.10. Bessel functions. Values of first three Bessel functions, up to  $x = 12$ .

Apparently, polar coordinates  $r$  and  $\theta$  suit better, see Fig. 1.9 (B). The Helmholtz equation in polar coordinates is

$$\frac{\partial^2 u(r, \theta)}{\partial r^2} + \frac{1}{r} \frac{\partial u(r, \theta)}{\partial r} + \frac{1}{r^2} \frac{\partial^2 u(r, \theta)}{\partial \theta^2} = -k^2 u(r, \theta). \quad (1.80)$$

The standard method of solving Eq. 1.80 is through *separation of variables*. By assuming a solution of  $u(r, \theta)$  as a product of a function only of  $r$  and a function only of  $\theta$ ,

$$u(r, \theta) = R(r) \Theta(\theta), \quad (1.81)$$

insert into Eq. 1.80, after a few simple algebraic moves, we obtain

$$\frac{r^2}{R(r)} \left[ \frac{d^2 R(r)}{dr^2} + \frac{1}{r} \frac{dR(r)}{dr} + k^2 R(r) \right] = -\frac{1}{\Theta(\theta)} \frac{d^2 \Theta(\theta)}{d\theta^2}. \quad (1.82)$$

The left-hand side of the equation only depends on  $r$ . The right-hand side of the equation only depends on  $\theta$ . Therefore, both sides must be a constant  $K$ . From the right-hand side of the equation,

$$\frac{d^2 \Theta(\theta)}{d\theta^2} + K \Theta(\theta) = 0. \quad (1.83)$$

The solution can be a sine function, a cosine function, or exponential function. The boundary condition that the function must be cyclic,

$$\Theta(2\pi) = \Theta(0), \quad (1.84)$$

requires that the function  $\Theta(\theta)$  must be either

$$\Theta(\theta) = \cos(n\theta), \quad n = 0, 1, 2, 3, \dots, \quad (1.85)$$

**Table 1.2: Zeros of Bessel functions**

| Index | $J_0(x)$          | $J_1(x)$          | $J_2(x)$          |
|-------|-------------------|-------------------|-------------------|
| 1     | $x_{01} = 2.4048$ | $x_{11} = 3.8317$ | $x_{21} = 5.1356$ |
| 2     | $x_{02} = 5.5201$ | $x_{12} = 7.0156$ | $x_{22} = 8.4172$ |
| 3     | $x_{03} = 8.6537$ | $x_{13} = 10.174$ | $x_{23} = 11.620$ |

or

$$\Theta(\theta) = \sin(n\theta), \quad n = 1, 2, 3, \dots \quad (1.86)$$

Therefore, the constant in Eq. 1.83 is  $K = n^2$ . The differential equation for the function with  $r$  is then

$$\frac{d^2 R(r)}{dr^2} + \frac{1}{r} \frac{dR(r)}{dr} + \left[ k^2 - \frac{n^2}{r^2} \right] R(r) = 0. \quad (1.87)$$

This is the well-known *Bessel equation*, and the solutions are the *Bessel functions*,

$$R(r) = J_n(kr). \quad (1.88)$$

Mathematical details of the Bessel functions are presented in Appendix A. Figure 1.10 shows the first three Bessel functions. Table 1.2 shows the first three zeros of the first three Bessel functions.

The solution of Eq. 1.87 must satisfy the boundary condition that at the rim of the timpani, where the membrane is fixed. At  $r = a$ ,  $R(r)$  must be zero. This condition fixes the frequency eigenvalues. The vibration pattern of the membrane at a given frequency eigenvalue must conform to the corresponding eigenfunction. Denote the  $m$ -th zero of the  $n$ -th Bessel function be  $x_{nm}$ , the allowed wave vectors must be

$$k_{nm}a = x_{nm}. \quad (1.89)$$

The frequency eigenvalues are

$$f_{nm} = \frac{vk_{nm}}{2\pi} = \frac{vx_{nm}}{2\pi a}, \quad (1.90)$$

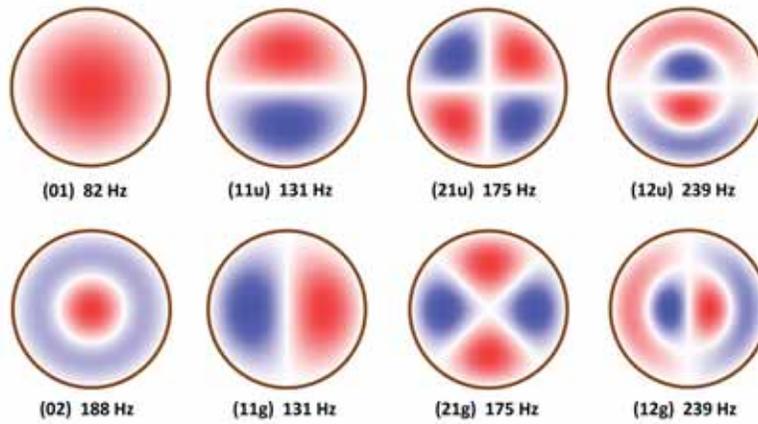
and the eigenfunctions of the vibration modes are either

$$u_{nm}^{(g)}(r, \theta) = J_n\left(\frac{x_{nm}r}{a}\right) \cos(n\theta), \quad n = 0, 1, 2, 3, \dots, \quad (1.91)$$

or

$$u_{nm}^{(u)}(r, \theta) = J_n\left(\frac{x_{nm}r}{a}\right) \sin(n\theta), \quad n = 1, 2, 3, \dots \quad (1.92)$$

Here the notation  $(g)$  indicates that the eigenfunction is symmetric, and  $(u)$  indicates that the eigenfunction is antisymmetric. A notation we will use throughout the book.

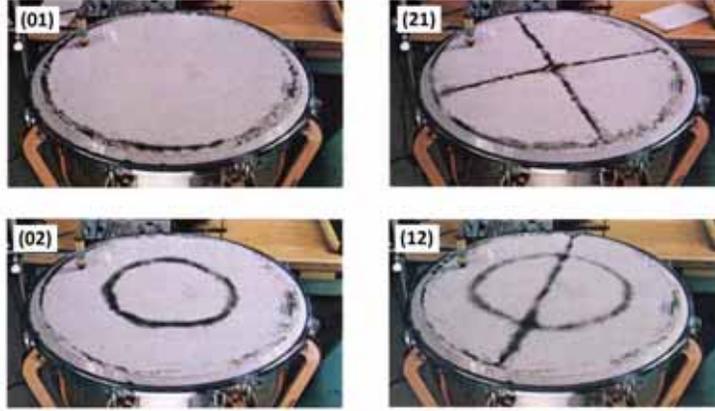


**Fig. 1.11. Eigenfunctions of the vibration modes.** Different color indicates polarity, and the density indicates the magnitude. The places with no or very little vibration are called *nodes*. The nodal pattern can be visualized by spreading powers on the membrane, which is called the *Chladni pattern* according to its discoverer. See Fig. 3.7.

Figure 1.11 shows graphical representations of the eigenfunctions. Different colors indicate different phases, and the density indicates the magnitude. The places with no vibration are indicated by white space, located between regions with different phases of vibration. Similar to the case of a string, those places are called *nodes*. The collection of nodes is called a *nodal pattern*. As shown, the geometry of the nodal patterns determines the nature of the vibration mode, or the nature of the eigenfunctions. In the parentheses in Fig. 1.11, the first digit  $n$  is the order of Bessel function, and the multiplier is  $m$ , the index of zeros of the Bessel function. The letter  $g$  indicates a symmetric angular function, associated with cosine. The letter  $u$  indicates an antisymmetric angular function, associated with sine. With a timpani of about 660 mm in diameter, the fundamental frequency is 82 Hz.

As shown, the frequencies of the higher vibration modes of a timpani are not integer multipliers of the frequency, as in the case of a string, see Table 1.1. Therefore, for a timpani with a bare membrane, the overtones are inharmonic. To make the overtones at least approximately harmonic, different types of *perturbations* are implemented.

The vibration patterns, or the eigenfunctions, can be visualized by spreading powers on the membrane. While the membrane vibrates, only at places with no vibration, that is, the nodes, the powder stays. The nodal pattern of a vibrating membrane can be visualized. That method was invented by German physicist Ernst Chladni in late 18th century and called Chladni patterns. The patterns in Fig. 1.12 were acquired by a group at Northern Illinois University lead by Thomas D. Rossing in the 1980s.



**Fig. 1.12. Nodes in eigenfunctions revealed by Chladni patterns.** By spreading powders on the head of a timpani, during vibration, the powders concentrate at the nodes. The pattern was discovered by German physicist Ernst Chladni, and thus named Chladni pattern. The vibration eigenfunctions are then visualized.

### Orthogonality of eigenfunctions

An interesting fact is that the eigenfunctions of different frequency eigenvalues are *orthogonal*. Two functions  $f_1(r, \theta)$  and  $f_2(r, \theta)$  on an area  $r < a$  are orthogonal means that the following integral is zero,

$$\int_0^a r dr \int_0^{2\pi} d\theta f_1(r, \theta) f_2(r, \theta) = 0. \quad (1.93)$$

The factor  $r$  occurs because the elemental area is  $dr \times r d\theta$ , see Fig. 1.9.

Let us look at the eigenfunctions of Eqs. 1.91 and 1.92. For eigenfunctions with different  $n$ , or different orders of the Bessel function, the integral is zero because of the integration on the angle  $\theta$ . For eigenfunctions of the same  $n$ , denoting

$$f_1(r) = J_n(\kappa_1 r) \quad (1.94)$$

and

$$f_2(r) = J_n(\kappa_2 r). \quad (1.95)$$

Both eigenfunctions satisfy the boundary conditions at  $r = a$ :

$$f_1(a) = J_n(\kappa_1 a) = 0, \quad (1.96)$$

and

$$f_2(a) = J_n(\kappa_2 a) = 0. \quad (1.97)$$

The eigenfunctions  $f_1(r)$  and  $f_2(r)$  should satisfy the Bessel equations,

$$\frac{d}{dr} \left( r \frac{df_1(r)}{dr} \right) - \frac{n^2}{r^2} f_1(r) + \kappa_1^2 r f_1(r) = 0, \quad (1.98)$$

and

$$\frac{d}{dr} \left( r \frac{df_2(r)}{dr} \right) - \frac{n^2}{r^2} f_2(r) + \kappa_2^2 r f_2(r) = 0. \quad (1.99)$$

Multiply Eq. 1.98 by  $f_2(r)$ , multiply Eq. 1.99 by  $f_1(r)$ , subtract one from another, then integrate it over  $(0, a)$ , we obtain

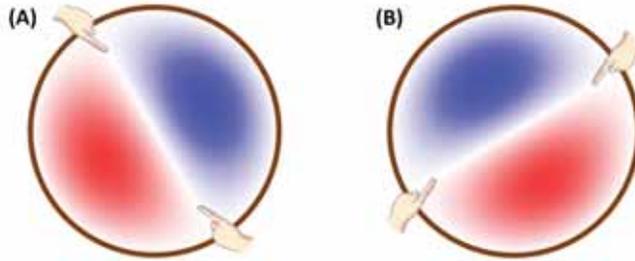
$$(\kappa_1^2 - \kappa_2^2) \int_0^a f_1(r) f_2(r) r dr = r \left( f_1(r) \frac{df_2(r)}{dr} - f_2(r) \frac{df_1(r)}{dr} \right) \Big|_0^a. \quad (1.100)$$

Because of Eqs. 1.96 and 1.97, the right-hand side is zero. Consequently, if  $\kappa_1 \neq \kappa_2$ , the integral must be zero. The two eigenfunctions are orthogonal. In quantum mechanics, such scenario is quite frequent.

### Linear superposition of degenerate eigenfunctions

The symmetric eigenfunctions in Eq. 1.91 and 1.92 with the same order  $n$ , as displayed in the three right-hand rows in Fig. 1.11, have the same frequency eigenvalue. Those pairs of vibration modes are called *degenerate*. This situation is also very frequent in quantum mechanics. Because the two degenerate eigenfunctions satisfy the same Helmholtz equation, any linear combination of the two eigenfunctions is also an eigenfunction with the same frequency eigenvalue. Therefore, by making a *linear superposition* of the eigenfunctions with the same eigenvalue, a new eigenfunction with the same eigenvalue is constructed.

Experimentally, by lightly touching some positions on the membrane and excite the timpani with a frequency source of 131 Hz, vibration modes with nodal lines passing the point of finger touching are generated. Figure 1.13 shows two such vibration patterns by touching the membrane lightly at different places. In quantum mechanics, with the presence of a weak perturbation to a number of degenerate wavefunctions, new wavefunctions with the same energy eigenvalue but well-defined orientations are generated.



**Fig. 1.13.** Linear superposition of degenerate eigenfunctions. By lightly touching some positions on the membrane and excite it, linear superposition of degenerate eigenfunctions with the same frequency eigenvalue can be formed.

### 1.3 Light as Electromagnetic Waves

The corpuscular and wave theories of light, as two completely incompatible theories, coexisted for more than two centuries. In 1690, Christiaan Huygens published *Traité de la Lumière* (A Treatise on Light), systematically expounded his wave theory. In 1704, Isaac Newton published a monograph *Opticks: A Treatise of the Reflexions, Refractions, Inflexions and Colours of Light*, criticized Huygens' wave theory of light, and expounded his corpuscular theory of light. The centuries-old debate is still instructive for the understanding of the modern concept of wave-particle duality.

#### 1.3.1 Newton's corpuscular theory of light

Besides research in mechanics represented by the epoch monograph *Principia*, Isaac Newton made numerous discoveries and inventions in optics. Besides a mathematical genius, he was also a super handy experimentalist. He invented and built the first working reflecting telescope by his own hands in 1668. It remains today the prototype of most of the professional astronomical telescopes as well as amateur telescopes.

A highly influential experimental discovery of Newton is that white light

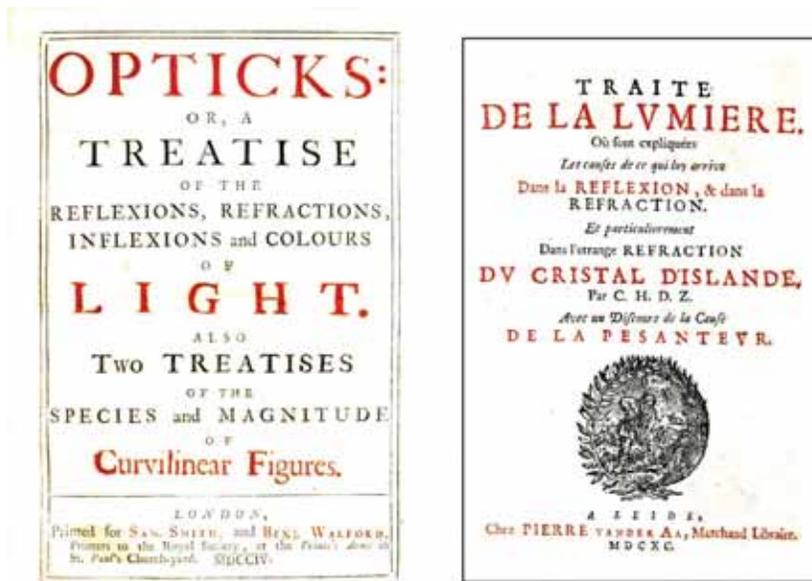
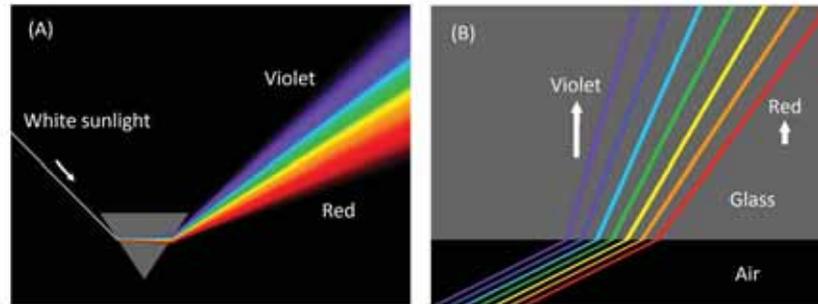


Fig. 1.14. Newton's optics and Huygens' Lumiere. Left, Newton's *Opticks*, published in 1704, systematically presented the corpuscular theory of light. Right, Huygens' *Treatise on Light*, published in 1690, systematically presented the wave theory of light.



**Fig. 1.15. Dispersion of prism and Newton's explanation.** (A), a schematic of the experiments of Isaac Newton, by decomposing the sunlight into colors using a prism. (B), Newton's explanation of dispersion based on his corpuscular theory of light, that the mass of red-light particles is greater than the mass of the violet-light particles, thus the violet light bends more than the red light.

is a mixture of a number of colored light rays, each of the colored rays is original and unchanged during propagation. In 1666, using a glass prism, he decomposed white sunlight into a spectrum, see Fig. 1.15(A). In *Opticks*, Query 29, on pages 533 – 538, he made a detailed explanation of this discovery by the corpuscular theory of light, see Fig. 1.15(B).

Newton's corpuscular theory of light is completely in line with his general mechanical view of the physical world, as summarized by Einstein, quoted in Section 1.1. In Newton's own words, rays of light consist of "very small Bodies emitted from shining Substances". Those material points are structureless, dimensionless, perfectly elastic, characterized only by its geometrical location  $(x, y, z)$  and mass. Its motion is determined only by the forces from other material points in the medium, governed by the law of mechanics. Accordingly, in vacuum, the light particles move in a straight line. In a uniform medium such as glass, because the force from the constituent particles of glass are balanced, light also moves in a straight line. Only at the interface between two different media, the speed changes. As shown in Fig. 1.15(B), as the particles of light enters from vacuum to glass, the material particles constituting glass attracts the particles of light, and the vertical speed of the light particles is increased. Newton assumed that the particles constituting red light are heavier than the particles constituting violet light. At the interface, red light particles gain less vertical speed than violet light particles, thus red light bends less than violet light. He derived Snell's sine law of refraction from that idea mathematically. He further presented evidence that the particles of red light are heavier than the rest, because when red light is absorbed, it generates more heat than for example violet light particles, characterized by Newton as smaller.

Experiments in the 19th century and on overturned Newton's corpuscular theory of light. First, the speed of light in transparent media such as

glass was found *slower* than the speed of light in vacuum. The speed of light in vacuum was found to be the upper limit of any speed. Second, although Einstein's theory of photon assigns a mass to light particles, the photons of the red light have a mass *smaller* than those of violet-light. Especially, Robert Young's double-slit experiment made a definitive proof of the wave nature of light. According to Young, color is associated to wavelength.

### 1.3.2 Young's double-slit experiment

Figure 1.16 shows a schematic of the Young double-slit experiment. From the left side, a plane wave of light with wavelength  $\lambda$  falls on a screen with slits  $S_1$  and  $S_2$ . The distance between the two slits is  $a$ . In the Figure, the phase of light wave is marked by colors; for example, red as positive, and blue as negative. The intensity of colors represents amplitude.

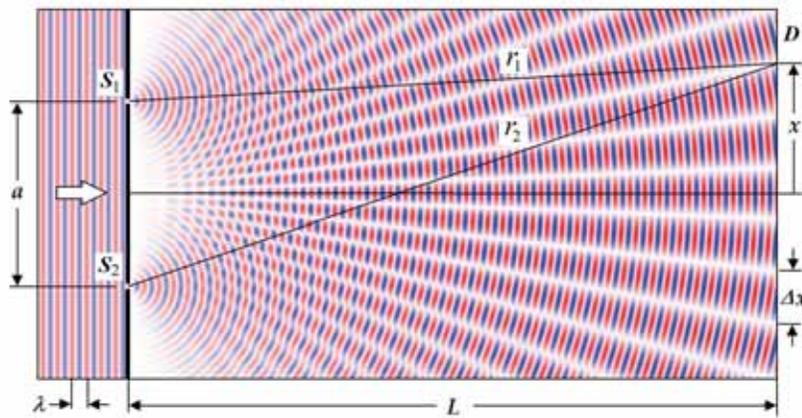
Consider a point  $x$  on the detection screen  $D$ . The distance from point  $x$  to the first slit is

$$r_1 = \sqrt{L^2 + \left(x - \frac{a}{2}\right)^2}. \quad (1.101)$$

and to the second slit is

$$r_2 = \sqrt{L^2 + \left(x + \frac{a}{2}\right)^2}. \quad (1.102)$$

The distance from the slits to the detection screen  $L$  is often much greater



**Fig. 1.16. Young's double-slit experiment.** A plane wave of light with wavelength  $\lambda$  falls on a screen with slits  $S_1$  and  $S_2$ . The distance between the two slits is  $a$ . The two light beams propagate and produce an interference pattern on the detection screen  $D$ . Here the phase of the wave is marked by colors: red is positive, blue is negative. The intensity of color represents amplitude.

that  $x$  and  $a$ . Using Newton's binomial formula,

$$\sqrt{1+u} \cong 1 + \frac{u}{2}, \quad (1.103)$$

A fairly accurate formula of the difference of  $r_2$  and  $r_1$  is obtained

$$r_2 - r_1 = L \left[ \sqrt{1 + \frac{1}{L^2} \left(x + \frac{a}{2}\right)^2} - \sqrt{1 + \frac{1}{L^2} \left(x - \frac{a}{2}\right)^2} \right] \cong \frac{xa}{L}. \quad (1.104)$$

At the point  $x$ , the wave amplitude is the sum of rays from two slits:

$$\begin{aligned} u(x) &= a \sin(kr_1 - \omega t) + a \sin(kr_2 - \omega t) \\ &= 2a \sin\left(k \frac{r_1 + r_2}{2} - \omega t\right) \cos\left(k \frac{r_1 - r_2}{2}\right). \end{aligned} \quad (1.105)$$

According to Eq. 1.62,  $\lambda k = 2\pi$ . If  $r_2 - r_1$  equals an integer multiple of wavelength  $\lambda$ , the cosine factor is  $\pm 1$ . The light intensity reaches a maximum. If  $r_2 - r_1$  equals an odd integer multiple of one half of wavelength  $\lambda$ , the cosine factor is zero. The two rays interfere destructively. The distance between two adjacent intensity minima  $\Delta x$  (see Fig. 1.16) is

$$\Delta x = \frac{\lambda L}{a}. \quad (1.106)$$

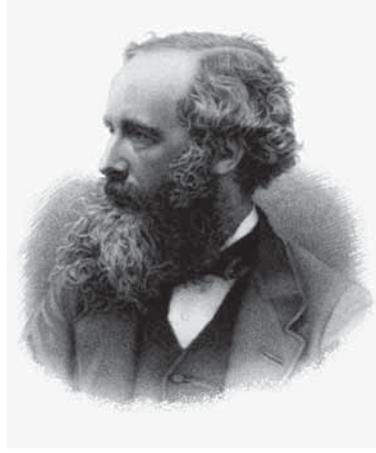
Young's two-slit experiment is essential to the understanding of the wave nature of the electron and other particles. We will repeatedly refer to that experiment in the discussions of wave-particle duality.

### 1.3.3 Maxwell's Theory of Electromagnetic Fields

Up to the middle of the nineteenth century, electromagnetic phenomena and light have been considered as totally independent entities. In 1865, in a monumental paper *A Dynamic Theory of the Electromagnetic Field*, James Clerk Maxwell (Fig 1.17) presented a complete set of partial differential equations explaining electromagnetic phenomena, and inferred that light is an electromagnetic wave. In an article written for the centenary of Maxwell's birth entitled *Maxwell's Influence on the Development of the Conception of Physical Reality*, Einstein said thusly:

Before Maxwell, people thought of physical reality – in so far as it represented events in nature – as material points, whose changes consist only in motions which are subject to total differential equations. After Maxwell, they thought of physical reality as represented by continuous fields, not mechanically explicable, which are subject to partial differential equations. This change in the conception of reality is the most profound and the most fruitful that physics has experienced since Newton.

**Fig. 1.17. James Clerk Maxwell.** Scottish physicist (1831–1879), one of the most influential physicists along with Isaac Newton and Albert Einstein. He developed a set of equations describing electromagnetism, known as the *Maxwell's equations*. In 1865, based on those equations, he predicted the existence of electromagnetic waves and proposed that light is an electromagnetic wave [?]. He also pioneered the kinetic theory of gases, and created a science fiction character *Maxwell's demon*. Portrait courtesy of Smithsonian Museum.



Quantum mechanics represents a further extension of the conceptual breakthrough started by Maxwell. Wavefunctions, the physical reality in quantum mechanics, also subject to partial differential equations.

### Maxwell's equations

In SI unit system, Maxwell's equations are

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\varepsilon_0}, \quad (1.107)$$

$$\nabla \cdot \mathbf{B} = 0, \quad (1.108)$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}, \quad (1.109)$$

$$\nabla \times \mathbf{B} = \varepsilon_0 \mu_0 \frac{\partial \mathbf{E}}{\partial t} + \mu_0 \mathbf{J}. \quad (1.110)$$

Electric current cannot exist in free space. For linear, uniform, isotropic materials, the current density  $\mathbf{J}$  is determined by the electric field intensity  $\mathbf{E}$  through Ohm's law,

$$\mathbf{J} = \sigma \mathbf{E}. \quad (1.111)$$

The names, meanings, and units of the physical quantities in these equations are listed in Table 1.3. For example, the electrical field intensity  $\mathbf{E}$  is related to the force  $\mathbf{F}$  acting on a charge  $q$ ,

$$\mathbf{F} = q\mathbf{E}. \quad (1.112)$$

**Table 1.3: Quantities in Maxwell's Equations**

| Symbol          | Name  | Unit                           | Meaning or Value           |
|-----------------|---|--------------------------------|----------------------------|
| $\mathbf{E}$    | Electric field intensity                          | V/m                            |                            |
| $\mathbf{B}$    | Magnetic field intensity                          | T (tesla)                      | N/A·m                      |
| $\rho$          | Electric charge density                           | C/m <sup>3</sup>               |                            |
| $\mathbf{J}$    | Electric current density                          | A/m <sup>2</sup>               |                            |
| $\varepsilon_0$ | Electric constant<br>(permittivity of free space) | F/m                            | $8.85 \times 10^{-12}$ F/m |
| $\mu_0$         | Magnetic constant<br>(permeability of free space) | H/m                            | $4\pi \times 10^{-7}$ H/m  |
| $\sigma$        | Conductivity                                      | $(\Omega \cdot \text{m})^{-1}$ |                            |

### Vector Potential and Scalar Potential

To treat the electromagnetic field in space, a convenient method is to use the *vector potential*. From Eq. 1.108, it is possible to construct a vector field  $\mathbf{A}$  which satisfies

$$\mathbf{B} = \nabla \times \mathbf{A}. \quad (1.113)$$

Then, Eq. 1.108 is automatically satisfied. Substituting Eq. 1.113 into Eq. 1.109, one obtains

$$\nabla \times \mathbf{E} = -\frac{\partial}{\partial t} \nabla \times \mathbf{A}. \quad (1.114)$$

For any function  $\phi(\mathbf{r})$ ,  $\nabla \times [\nabla\phi(\mathbf{r})] = \mathbf{0}$ . Therefore, it is always possible to set up the vector potential  $\mathbf{A}$  such that

$$\mathbf{E} = -\frac{\partial \mathbf{A}}{\partial t} - \nabla\phi, \quad (1.115)$$

where  $\phi$  is the electrostatic potential arising from the charges. The choice of the vector potential is not unique. By adding a gradient of an arbitrary function to it, values of the electric field and magnetic field do not change. This is called the *gauge invariance* of the vector potential. It is possible to define a vector potential which satisfies the condition

$$\nabla \cdot \mathbf{A} = 0. \quad (1.116)$$

Equation 1.116 is called the *Coulomb gauge*, which is the most convenient gauge to treat nonrelativistic problems of an atomic system and an independent electromagnetic wave. In fact, using Eq. 1.116 and the first Maxwell equation Eq. 1.107, one obtains

$$\nabla^2 \phi = -\frac{\rho}{\varepsilon_0}, \quad (1.117)$$

which means that the scalar potential is generated by the static charges only. It is thus convenient for treating the problems of interactions between the radiation field and atomic systems.

By combining Eqs. 1.107, 1.111, and 1.117, we can derive an expression of the potential energy for a generalized Coulomb's law. Assuming the charge density is spherically symmetric, that it only depends on  $r$ . The symmetry implies that the scalar potential is a function of  $r$  only. Equation 1.117 becomes

$$\frac{1}{r^2} \frac{d}{dr} \left( r^2 \frac{d\phi(r)}{dr} \right) = -\frac{\rho(r)}{\epsilon_0}, \quad (1.118)$$

Because the total charge  $Q(r)$  in a sphere of radius  $r$  is

$$Q(r) = \int_0^r 4\pi r'^2 \rho(r') dr', \quad (1.119)$$

taking a similar integration on the left-hand side of Eq. 1.118 yields

$$\int_0^r 4\pi r'^2 \frac{1}{r'^2} \frac{d}{dr'} \left( r'^2 \frac{d\phi(r')}{dr'} \right) dr' = 4\pi r^2 \frac{d\phi(r)}{dr}. \quad (1.120)$$

Combining Eqs. 1.118, 1.119, and 1.120, we obtain

$$\frac{d\phi(r)}{dr} = \frac{Q(r)}{4\pi\epsilon_0 r^2}. \quad (1.121)$$

Integrating Eq. 1.121 over  $r$ , notice that at  $r = \infty$ ,  $\phi = 0$ , we have

$$\phi(r) = \int_{\infty}^r \frac{Q(r')}{4\pi\epsilon_0 r'^2} dr' = \frac{Q(r)}{4\pi\epsilon_0 r}. \quad (1.122)$$

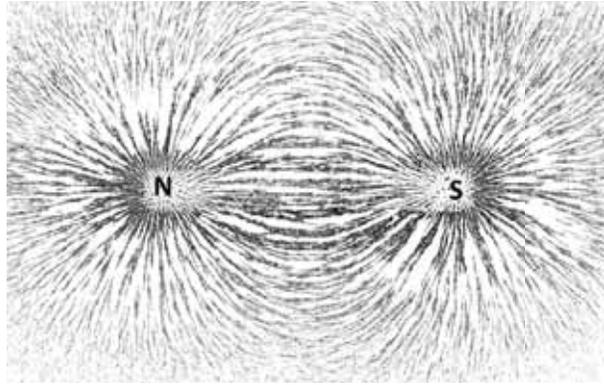
Therefore, for a spherically symmetric charge density function  $\rho(r)$ , the scalar potential at a point  $r$  is the same as an electrical charge  $Q(r)$  concentrated at the origin.

For a hydrogen atom, with a proton of positive charge  $e$  at the origin and an electron with a negative charge  $-e$  at  $r$ , the potential energy is

$$V(r) = -\frac{e^2}{4\pi\epsilon_0 r}. \quad (1.123)$$

### 1.3.4 Fields as a Physical Reality

In Newtonian physics, the physical reality is material particles. The position and momentum of each particle can be observed without significant disturbance. If the position and momentum of a particle is known at a given time, the subsequent values are determined by Newton's laws. Therefore, the Newtonian physics is causal.

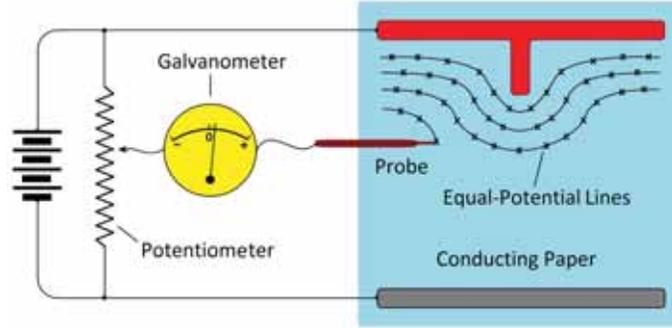


**Fig. 1.18. Observing and mapping magnetic field.** By placing a white paper above a magnet and sprinkle iron filings on it, the iron filings follows the lines of magnetic field and shows a map of the magnetic field. The magnetic field is an observable physical reality, which exists regardless of being measured or not, and only marginally disturbed by the mapping process.

In Maxwellian physics, the physical reality is various fields, including magnetic fields, electrical fields, etc. The observations are not the position and momentum of individual material particles, but rather continuous fields. The methods of observing those fields are different from those in Newtonian physics. However, from a philosophical point of view, there is no difference.

Although not visible by naked eyes, magnetic field can be observed, measured, and mapped by simple experiments. Figure 1.18 shows an experiment easily done in a high-school classroom. By placing a piece of white paper on top of a magnet and sprinkle iron filing over it, the lines of magnetic field are displayed. The experiment was described in Faraday's laboratory notebooks, and motivated Maxwell to represent the fields with continuous functions governed by partial differential equations. More accurate measurements of magnetic fields can be performed using Hall-effect probes, frequently used by scientists and engineers. Note that the magnetic fields around a permanent magnet is an objective physical reality, which exists regardless of being measured or not. The iron filings and the Hall-effect probe does not perturb the magnetic field significantly.

Similarly, the electric field can be observed, measured, and mapped by a simple device. Although not visible by naked eyes, the experiments can also be performed in a high-school classroom, as shown in Fig. 1.19. A battery and a pair of electrodes create a pattern of electrical field on a piece of conducting paper. To find the equal potential lines at a given voltage provided by a potentiometer, a probe with a galvanometer is used. By connecting the points with zero current together, the electrical field is mapped. Again, the electrical field pattern is an objective reality. It exists regardless whether a measurement is conducted or not. The probe, to find



**Fig. 1.19. Observing and mapping electric field.** Using a piece of conducting paper or a tray of salt water, the electric field can be probed and mapped by a probe and a galvanometer. A potentiometer defines or selects a value of potential voltage. When the galvanometer shows no current, the spot of the same potential voltage is identified. The entire electric field can be imaged. Similar to the magnetic field, the electric field is an observable physical reality, which exists regardless of being measured or not, and only marginally disturbed by the measurement process.

the points with no current literally makes no disturbance to the objective physical reality of the electrical field.

The electromagnetic fields are governed by Maxwell's equations, which are first-order in time. Once the electromagnetic fields are known at a given time, the subsequent evolution is completely determined by Maxwell's equations. Therefore, similar to Newtonian physics, Maxwell's physics of electromagnetic fields is completely causal.

### 1.3.5 Electromagnetic Waves

In this section, we study the electromagnetic waves in free space, that is, where the electric charge  $\rho$  and current  $\mathbf{J}$  are zero. Take the curl of Eq. 1.109, then using Eq. 1.110, we have

$$\nabla \times \nabla \times \mathbf{E} = -\frac{\partial}{\partial t} \nabla \times \mathbf{B} = -\varepsilon_0 \mu_0 \frac{\partial^2 \mathbf{E}}{\partial t^2}. \quad (1.124)$$

On the other hand, using the identity

$$\nabla \times \nabla \times \mathbf{E} \equiv \nabla(\nabla \cdot \mathbf{E}) - \nabla^2 \mathbf{E} \quad (1.125)$$

Equation 1.124 becomes

$$\nabla^2 \mathbf{E} = \varepsilon_0 \mu_0 \frac{\partial^2 \mathbf{E}}{\partial t^2}. \quad (1.126)$$

Introducing

$$c = \frac{1}{\sqrt{\varepsilon_0 \mu_0}}, \quad (1.127)$$

Eq. 1.126 becomes

$$\nabla^2 \mathbf{E} = \frac{1}{c^2} \frac{\partial^2 \mathbf{E}}{\partial t^2}, \quad (1.128)$$

which is a wave equation with velocity  $c$ . Using a similar procedure, we can show that the magnetic field intensity satisfy the same wave equation,

$$\nabla^2 \mathbf{B} = \frac{1}{c^2} \frac{\partial^2 \mathbf{B}}{\partial t^2}. \quad (1.129)$$

According to values of  $\varepsilon_0$  and  $\mu_0$  coming from electromagnetic measurements in 1860s, the velocity of electromagnetic waves should be  $3.1 \times 10^8$  m/s. On the other hand, experimental values of the speed of light at that time were  $2.98 \times 10^8$ – $3.15 \times 10^8$  m/s. The difference was within experimental error. Maxwell proposed thusly [?]:

The agreement of the results seems to show that light and magnetism are affections of the same substance, and that light is an electromagnetic disturbance propagated through the field according to electromagnetic laws.

Maxwell's theory of electromagnetic waves was experimentally verified by Heinrich Hertz in 1885. From recent electrical measurements, one finds  $1/\sqrt{\varepsilon_0 \mu_0} = 2.998 \times 10^8$  m/s, which equals the speed of light in vacuum,  $c$ .

### Radiation Power and Poynting Vector

Let us study the energy balance in an electromagnetic field by considering a unit volume with relatively uniform fields. If the current density is  $\mathbf{J}$  and the electric field intensity is  $\mathbf{E}$ , the ohmic energy loss per unit time per unit volume is  $\mathbf{J} \cdot \mathbf{E}$ . Using Eq. 1.110, the expression of energy loss becomes

$$\mathbf{J} \cdot \mathbf{E} = \frac{1}{\mu_0} \mathbf{E} \cdot (\nabla \times \mathbf{B}) - \varepsilon_0 \mathbf{E} \cdot \frac{\partial \mathbf{E}}{\partial t}. \quad (1.130)$$

Using the mathematical identity

$$\mathbf{E} \cdot (\nabla \times \mathbf{B}) = -\nabla \cdot (\mathbf{E} \times \mathbf{B}) + \mathbf{B} \cdot (\nabla \times \mathbf{E}), \quad (1.131)$$

Eq. 1.130 becomes

$$\mathbf{J} \cdot \mathbf{E} = -\nabla \cdot \left( \frac{1}{\mu_0} \mathbf{E} \times \mathbf{B} \right) + \frac{1}{\mu_0} \mathbf{B} \cdot (\nabla \times \mathbf{E}) - \varepsilon_0 \mathbf{E} \cdot \frac{\partial \mathbf{E}}{\partial t}. \quad (1.132)$$

Using Eq. 1.109, Eq. 1.132 becomes

$$\mathbf{J} \cdot \mathbf{E} = -\nabla \cdot \left( \frac{1}{\mu_0} \mathbf{E} \times \mathbf{B} \right) - \frac{\partial}{\partial t} \left( \frac{\varepsilon_0}{2} E^2 + \frac{1}{2\mu_0} B^2 \right). \quad (1.133)$$

The right-hand side of Eq. 1.133 has a straightforward explanation. The energy density of the electromagnetic fields is

$$W = \frac{\epsilon_0}{2} E^2 + \frac{1}{2\mu_0} B^2, \quad (1.134)$$

and the power density of the electromagnetic field per unit area is

$$\mathbf{S} = \frac{1}{\mu_0} \mathbf{E} \times \mathbf{B}. \quad (1.135)$$

The vector  $\mathbf{S}$  represents the power flow of electromagnetic waves, and called the *Poynting vector* after its discoverer.

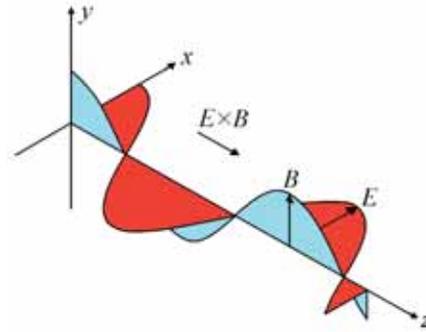
According to Eqs. 1.134 and 1.135, the electromagnetic field is not only a medium to transfer forces among material points. It is a physical reality by itself, possessing energy density and energy flow. According to the Einstein relation,  $E = mc^2$ , the electromagnetic field also has mass density and mass flow. As emphasized by Einstein, continuous field as a physical reality, the Maxwellian point of view, is distinctive from the Newtonian point of view, that only the material points represent physical reality.

### 1.3.6 Polarization of Light

Polarization of light is a phenomenon well known in classical physics. But it is closely related to the basic concepts of quantum physics. A good understanding from the classical-physics point of view will be helpful to understand quantum mechanics.

To make a simple and clear picture, consider a sinusoidal electromagnetic wave with angle frequency  $\omega = 2\pi f$  propagating in the  $z$ -direction. See Fig. 1.20. In general, the electrical field intensity is

$$\mathbf{E} = \mathbf{E}_0 \sin(kz - \omega t), \quad (1.136)$$



**Fig. 1.20. Electromagnetic wave.** The electromagnetic wave is *transverse*, where the intensity vectors  $\mathbf{E}$  and  $\mathbf{B}$  are perpendicular to the direction of propagation. The electric field intensity  $\mathbf{E}$  is perpendicular to the magnetic field intensity  $\mathbf{B}$ . The energy flux vector  $\mathbf{S} = \mu_0^{-1} \mathbf{E} \times \mathbf{B}$  is formed from  $\mathbf{E}$  and  $\mathbf{B}$  by a right-hand rule.

where

$$k = \frac{\omega}{c} \quad (1.137)$$

is the wavevector. Because the field intensities only depends on  $z$ , in space without electrical charge, the first Maxwell equation, Eq. 1.107, becomes

$$\frac{\partial E_z}{\partial z} = 0. \quad (1.138)$$

In other words, the  $z$ -component of the electrical field intensity is a constant, not a wave. The wave is *transverse*, which has  $x$  and  $y$  components only. The  $x$ -polarized wave is

$$E_x(z, t) = E_{x0} \sin(kz - \omega t). \quad (1.139)$$

Using Eq. 1.120, we have

$$B_y(z, t) = -\frac{1}{c} E_x(z, t) = -\frac{1}{c} E_{x0} \sin(kz - \omega t). \quad (1.140)$$

The power density of the wave, according to Eq. 1.135, is

$$S = \frac{1}{c\mu_0} E_{x0}^2 \sin^2(kz - \omega t). \quad (1.141)$$

Because the average of the square of a sine function is  $1/2$ , the average power density is

$$\bar{S} = \frac{1}{2c\mu_0} E_{x0}^2. \quad (1.142)$$

The  $y$ -polarized wave is

$$E_y(z, t) = E_{y0} \sin(kz - \omega t) \quad (1.143)$$

and

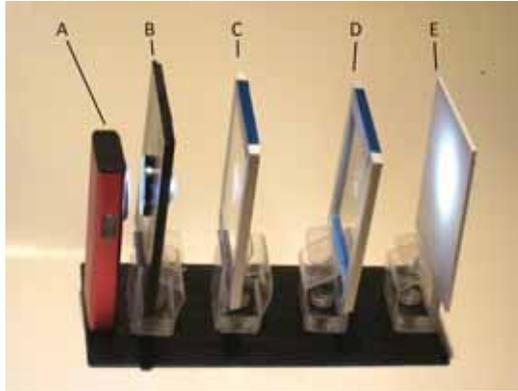
$$B_x(z, t) = \frac{1}{c} E_y(z, t) = \frac{1}{c} E_{y0} \sin(kz - \omega t). \quad (1.144)$$

and the average power density is

$$\bar{S} = \frac{1}{2c\mu_0} E_{y0}^2. \quad (1.145)$$

Typically, for natural light, such as sunlight or from an electrical lamp, the two polarizations are balanced. The magnitude of  $E_{x0}$  and the magnitude of  $E_{y0}$  are roughly equal. By using a lineal polarizer, for example, in the  $x$  direction, the  $y$ -component is blocked. The light comes out from the polarizer is lineally polarized in the  $x$ -direction.

Figures 1.21 to 1.23 show an experimental setup for the demonstration of superposition of polarized light, as described in Chapter 1, Section 2

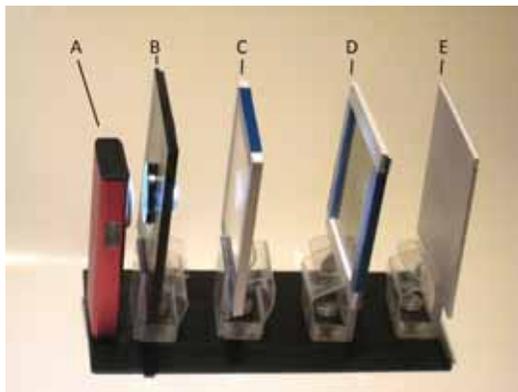


**Fig. 1.21. The three polarizer experiment: Step 1.** If both polarizers (C) and (D) are in  $x$ -direction, all light goes through polarizer (C) passes (D).

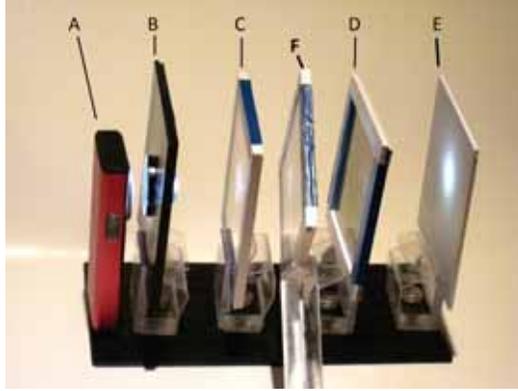
of Dirac's *The Principles of Quantum Mechanics*, sometimes referred to as Dirac's three-polarizer experiment. In all three Figures, (A) is a light source. (B) is a lens. (C) and (D) are linear polarizers, the blue sidebar indicates the direction of polarization. (E) is a screen. All elements are mounted by magnetic and elastic clampers to facilitate adjustment.

In Fig. 1.21, both polarizers (C) and (D) are in  $x$ -direction, all light goes through (C) passes (D). In Fig. 1.22, polarizer (C) is in  $x$ -direction, but polarizer (D) is in  $y$ -direction. All light goes through (C) is blocked by polarized (D). By inserting a third polarizer (F) between (C) and (D) with a tilting angle, as shown in Fig. 1.23, some light can pass all three polarizers.

Figure 1.24 shows an analysis of the role of the tilted polarizer. Assuming that the angle is  $\alpha$ . The insertion of polarizer (F) is equivalent to a



**Fig. 1.22. The three polarizer experiment: Step 2.** If polarizer (C) is in  $x$ -direction, but polarizer (D) is in  $y$ -direction, all light goes through (C) is blocked by polarized (D).



**Fig. 1.23. The three polarizer experiment: Step 3.** By inserting a third polarizer (F) between (C) and (D) with a tilting angle, some light can pass all three polarizers.

coordinate transformation in the  $(x, y)$ -plane. As shown in Fig. 1.24(A), the component of original electrical field intensity, in  $x_0$ -direction, to the new  $x_1$ -direction is

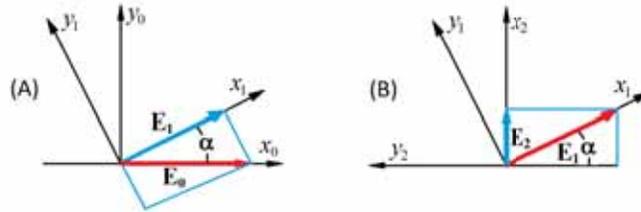
$$E_1 = E_0 \cos \alpha. \quad (1.146)$$

The electrical field intensity on the subsequent polarizer (D) is equivalent to another coordinate transformation in the  $(x, y)$ -plane. As shown in Fig. 1.24(B), the component of the electrical field intensity in the  $x_2$ -direction is

$$E_2 = E_1 \sin \alpha = E_0 \sin \alpha \cos \alpha = \frac{1}{2} \sin 2\alpha. \quad (1.147)$$

Because the power is proportional to the square of the electrical field intensity, the outgoing radiation power intensity  $\bar{S}_2$  is related to the incoming power intensity  $\bar{S}_0$  by

$$\bar{S}_2 = \frac{\bar{S}_0}{4} \sin^2 2\alpha. \quad (1.148)$$



**Fig. 1.24. The three polarizer experiment: analysis.** (A) By inserting a third polarizer with an angle  $\alpha$  shifted from the first polarizer, it is equivalent to do a coordinate transformation from  $(x_0, y_0)$  to  $(x_1, y_1)$ . The projection of the original electrical field intensity  $E_0$  on the transformed  $x_1$  axis is  $E_1 = E_0 \cos \alpha$ . (B) The projection of  $E_1$  to the  $x_2$  axis of the final polarizer is  $E_2 = E_1 \sin \alpha$ .

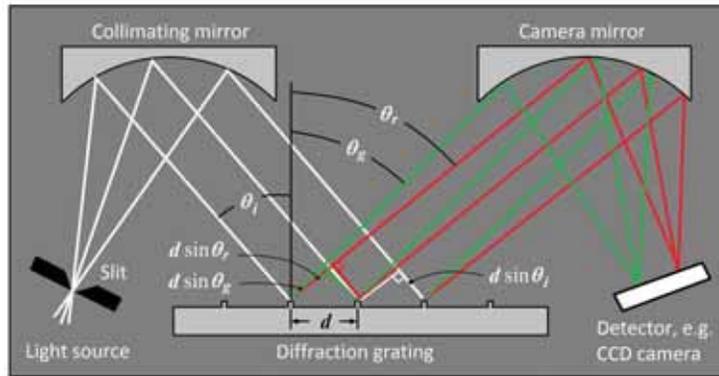
Therefore, if the polarization direction of the third polarizer (F) is either in the  $x$ -direction or in the  $y$ -direction, the output radiation power is zero. At  $\alpha = (2n + 1)\pi/4 = (2n + 1) \times 45^\circ$ , where  $n$  is an integer, the outgoing radiation power density is at maximum. The predicted results can be easily verified by direct experiments.

## 1.4 Atomic Spectra

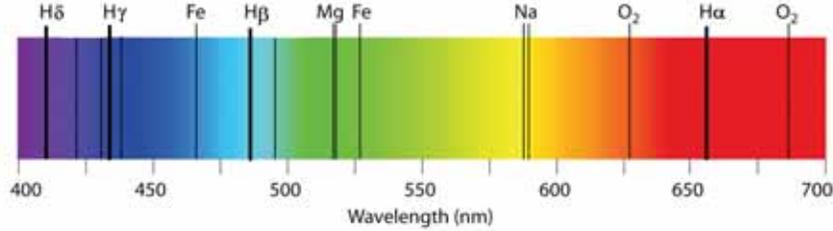
As we have presented in Section 1.3.1, Isaac Newton discovered that by using a prism, sunlight can be dispersed into a variety of colors. That was the starting point of optical spectroscopy. In the 19th century, as a result of the advances in precision machining, a more powerful and accurate device for dispersion of light, *diffraction gratings*, was developed.

Figure 1.25 is a schematic of diffraction grating. A grating is made by engraving a large number of grooves on a flat piece of metal. The spacing of two adjacent grooves  $d$  should be greater than the wavelength of the light of interest, but not excessively greater. A beam of light is incident through a slit on a collimating mirror to become parallel light, falls on the grating. After diffracted by the grating, the parallel light beam is focused by a camera mirror on a detector, typically a CCD camera. Due to interference, light of different wavelength have different condition for maximum intensity. For example, the condition of maximum intensity for red light is that the path difference equals an integer multiple of the wavelength,

$$d(\sin \theta_i - \sin \theta_r) = m\lambda_r, \quad (1.149)$$



**Fig. 1.25. A schematic of diffraction grating.** A beam of light from the source is going through a thin slit, then spread on a collimating mirror to become parallel light, falling on a diffraction grating. After diffraction, the light is focused by a camera mirror onto a detector. Light beams of different wavelength, for example, red light and green light, will hit on different places of the detector.



**Fig. 1.26. Absorption atomic spectra in the visible range.** Absorption lines in the spectrum of solar radiation, originally discovered by Joseph von Fraunhofer in 1814, thus named as the Fraunhofer lines. By comparing with the emission spectra of various elements, the origin of those absorption lines were gradually known. Of special interest are the four strong absorption lines at 656.2 nm, 486.1 nm, 434.0 nm, and 410.0 nm, identified as from the hydrogen atom, and renamed as  $H\alpha$ ,  $H\beta$ ,  $H\gamma$ , and  $H\delta$ . Hydrogen is by far the most abundant element in the atmosphere of the Sun.

where  $\theta_i$  is the incident angle,  $\theta_r$  is the angle of maximum intensity for red light with wavelength  $\lambda_r$ , and  $m$  is an integer of the number wavelengths of the interference path. Likewise, for green light, the condition for maximum intensity is

$$d(\sin \theta_i - \sin \theta_g) = m\lambda_g, \quad (1.150)$$

where  $\theta_r$  is the maximum angle for green light with wavelength  $\lambda_r$ .

The resolution power with a grating with  $N$  grooves of order  $m$  is

$$\frac{\lambda}{\Delta\lambda} = mN, \quad (1.151)$$

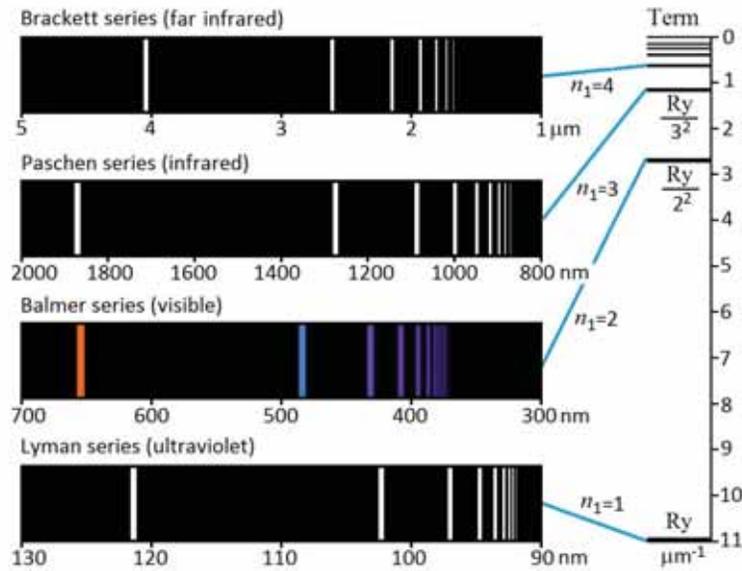
where  $\lambda$  is the wavelength of the light, and  $\Delta\lambda$  is the smallest resolvable wavelength difference. Simply speaking, the more lines on the gratings, the

**Table 1.4: Wavelengths of hydrogen lines**

|                | $n_1 = 1$ | $n_1 = 2$ | $n_1 = 3$ | $n_1 = 4$          |
|----------------|-----------|-----------|-----------|--------------------|
|                | Lyman     | Balmer    | Paschen   | Brackett           |
| $n_2 = 2$      | 121.6 nm  |           |           |                    |
| $n_2 = 3$      | 102.6 nm  | 656.3 nm  |           |                    |
| $n_2 = 4$      | 97.2 nm   | 486.1 nm  | 1875 nm   |                    |
| $n_2 = 5$      | 94.9 nm   | 434.0 nm  | 1282 nm   | 4.05 $\mu\text{m}$ |
| $n_2 = 6$      | 93.8 nm   | 410.2 nm  | 1094 nm   | 2.62 $\mu\text{m}$ |
| $n_2 = 7$      | 93.0 nm   | 397.0 nm  | 1005 nm   | 2.16 $\mu\text{m}$ |
| $n_2 = 8$      | 92.6 nm   | 388.9 nm  | 955 nm    | 1.94 $\mu\text{m}$ |
| $n_2 = \infty$ | 91.2 nm   | 364.6 nm  | 820 nm    | 1.46 $\mu\text{m}$ |

finer its resolution. In late 19th century, a typical gratings has more than 10,000 lines, capable of a resolution power of 10,000 or more.

In 19th century, using the diffraction gratings, a vast field of science and technology, atomic spectroscopy, emerged. Each atom has its characteristic group of spectral lines. Those lines can be either observed as absorption lines or emission lines. In 1814, German physicist Joseph von Fraunhofer discovered a number of absorption lines in the observed spectrum of sunlight. After years of continuing research, thousands of so-called *Fraunhofer lines* were identified. Later, by using flames with known elements as the light source, thousands of emission lines of atoms were discovered. The absorption lines and the emissions lines are correlated one by one. Over the 19th century, optical spectroscopy gradually became one of the most important methods for chemical analysis. A number of elements were discovered first by the observation of spectral lines of unknown origin, including Cs, Rb, and Tl (1860-61); In (1863); Ga (1875); Tm (1870); Nd and Pr (1885); Sm and Ho (1886). An interesting case is helium. Several strong absorption lines were first discovered in the absorption spectrum of the Sun. In 1868, by emission spectroscopy, helium was found as a rare element in the air. The name, helium, was referred to its origin as from the Sun.



**Fig. 1.27. Emission spectra of hydrogen.** The electromagnetic wave is *transverse*, where the intensity vectors  $\mathbf{E}$  and  $\mathbf{B}$  are perpendicular to the direction of propagation. The electric field intensity  $\mathbf{E}$  is perpendicular to the magnetic field intensity  $\mathbf{B}$ . The energy flux vector  $\mathbf{S} = \mu_0^{-1} \mathbf{E} \times \mathbf{B}$  is formed from  $\mathbf{E}$  and  $\mathbf{B}$  by a right-hand rule.

Of particular interest are four Fraunhofer lines, initially observed in the spectrum of sunlight, named C, F, G and h. Later identified as from the hydrogen atom, and renamed  $H\alpha$ ,  $H\beta$ ,  $H\gamma$ , and  $H\delta$  lines. See Fig. 1.26. Because of the high resolution of diffraction gratings, the accuracy of wavelengths of those lines reached six to seven significant digits. In 1885, Swiss mathematician Johann Jakob Balmer (1825 – 1898) found an accurate empirical formula for the four hydrogen lines. In 1889, Swedish physicist Johannes Robert Rydberg (1854 – 1919) extended the Balmer formula to include hydrogen spectral lines in the infrared and ultraviolet regions. The key of their success was to use wavenumber, the inverse of wavelength, to represent the data. Rydberg's formula is

$$\frac{1}{\lambda} = \text{Ry} \left( \frac{1}{n_1^2} - \frac{1}{n_2^2} \right), \quad (1.152)$$

where Ry is the Rydberg constant,  $n_1$  and  $n_2$  are integers. For  $n_1 = 1$ , the ultraviolet lines in the Lyman series are represented. For  $n_1 = 2$ , the visible lines of the Balmer series are represented. See Table 1.4 and Fig. 1.27.

Towards the end of the 19th century, the Rydberg formula was well validated. The Rydberg constant was the most accurately measured physical constant. However, despite many trials, no interpretation of the Rydberg formula was discovered until the rise of the quantum theory.



# Chapter 2

## Wave and Quantum

On November 9, 1922, the Royal Swedish Academy of Sciences announced that the Nobel Prize in Physics 1921 was awarded to Albert Einstein “for his services to Theoretical Physics, and especially for his discovery of the law of the photoelectric effect.”<sup>1</sup>

The announcement was a surprise to the public and news media, as Einstein was already a celebrity because of his theory of relativity. In fact, the Nobel Committee was correct: The paper about the photoelectric effect is the most revolutionary, the most original, and also the most controversial of Einstein’s papers. It started the era of quantum physics, the scientific theory having the most profound impact on human society in the 20th century. The concept of wave-quantum duality, defined in that paper for light, was later generalized to all particles by Louis de Broglie, then became the conceptual basis of quantum mechanics.

### 2.1 Einstein’s Energy Quantum of Light

Figure 2.1 shows the title and a key paragraph of Einstein’s paper, which explained the concept of energy quantum of the electromagnetic wave. Expecting a strong opposition, the title was scrupulously worded: *On a Heuristic Viewpoint Concerning the Production and Transformation of Light*. Notice that the word *propagation* is avoided. The adjective “heuristic”, rarely used in scientific literature, is an attempt to disperse objections. The paragraph showed on Fig. 2.1 said emphatically that the wave theory of light, “has worked well in the representation of purely optical phenomena and will probably never be replaced by another theory”, such as diffraction, reflection, refraction, dispersion, etc. However, the continuous wave theory may “lead to contradictions with experience when it is applied to the phenomena of emission and transformation of light”.

Einstein argued that in order to explain experimental observations during production and transformation of light, energy is always transferred in

---

<sup>1</sup>During the selection process in 1921, the Nobel Committee for Physics decided that none of the year’s nominations met the criteria as outlined in the will of Alfred Nobel. According to the Nobel Foundation’s statutes, the Nobel Prize can in such a case be reserved until the following year, and this statute was then applied. Albert Einstein therefore received his Nobel Prize for 1921 one year later, in 1922.

**8. Über einen  
die Erzeugung und Verwandlung des Lichtes  
betreffenden heuristischen Gesichtspunkt;  
von A. Einstein.**

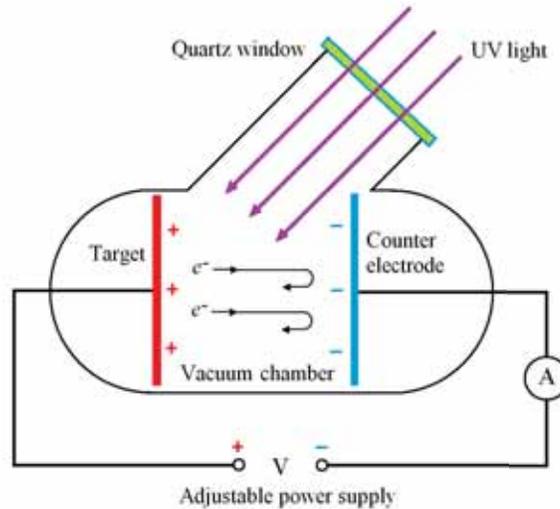
Die mit kontinuierlichen Raumbfunktionen operierende Undulationstheorie des Lichtes hat sich zur Darstellung der rein optischen Phänomene vortrefflich bewährt und wird wohl nie durch eine andere Theorie ersetzt werden. Es ist jedoch im Auge zu behalten, daß sich die optischen Beobachtungen auf zeitliche Mittelwerte, nicht aber auf Momentanwerte beziehen, und es ist trotz der vollständigen Bestätigung der Theorie der Beugung, Reflexion, Brechung, Dispersion etc. durch das Experiment wohl denkbar, daß die mit kontinuierlichen Raumbfunktionen operierende Theorie des Lichtes zu Widersprüchen mit der Erfahrung führt, wenn man sie auf die Erscheinungen der Lichterzeugung und Lichtverwandlung anwendet.

**Fig. 2.1.** Einstein's paper on energy quantum of light. The title and a key paragraph of the paper cited for the 1921 Nobel Prize in Physics, the second paper Einstein published on *Annalen der Physik* in 1905. First, Einstein emphasized the correctness of the wave theory of light. Then, he listed and discussed four experimental observation that led to contradiction with the wave theory of light, and proposed the concept of light quantum. The word "heuristic" was attempted to disperse objections.

an integer multiple of an indivisible unit he called *energy quantum*  $\epsilon$ ,

$$\epsilon = hf, \quad (2.1)$$

here  $h$  is the Planck constant,  $f$  is the frequency. Max Planck proposed this postulation in 1900 as a mathematical trick to explain his blackbody radiation formula, see Section ???. Einstein took it seriously as a basic law of physics. He discussed four cases: blackbody radiation, photofluorescence, photoelectric effect, and ionization of gases by ultraviolet light. The first six sections are about blackbody radiation. Einstein discussed in detail on the validity of the Planck postulation. Especially, Einstein pointed out that by comparing the Planck formula of blackbody radiation with experimental data, the Avogadro constant is found to be  $6.17 \times 10^{23}$ /mole, consistent with the value from other sources. In Section 7, he cited the Stokes's rule of photoluminescence of light that the frequencies of outgoing radiation is always lower than the wavelengths of incoming radiation. This can be explained naturally by assuming that the energy of higher frequency light quanta is always higher than the energy of lower frequency light quanta. Section 8 discussed the photoelectric effect. He made an order of magnitude estimate that the light quanta of ultraviolet light should be greater than the energy to eject an electron from typical metals, or the workfunction of the metals. Section 9 is about the ionization of gases with ultraviolet light. Experimentally, the threshold voltage to ionize a gas is about 10 V, and the



**Fig. 2.2.** Lenard's apparatus for studying photoelectric effect. A quartz window allows the UV light from an electric arc lamp to shine on a target. The voltage between the target and the counter electrode is controlled by an adjustable power supply. An ammeter is used to measure the electric current generated by the UV light, the *photocurrent*. By gradually increasing the voltage (with the polarity as shown), the photocurrent is reduced. The voltage with which the photocurrent becomes zero is recorded as the *stopping voltage* [?].

largest wavelength of light to ionize a gas is  $1.9 \times 10^{-5}$  cm. Using Eq. 1.8, the two thresholds are roughly equivalent.

Attention from the academic community was concentrated to Section 8, where Einstein proposed an experimentally verifiable formula relating the frequency of incoming light and the kinetic energy of outgoing electrons. The formula, referred to the “law of photoelectric effect” by the Nobel Committee, was tested by Robert Millikan with a series of careful experiments for 10 years, and found to be surprisingly accurate.

### 2.1.1 The photoelectric effect

The photoelectric effect was discovered accidentally by Heinrich Hertz in 1887 during experiments to generate electromagnetic waves. Since then, a number of studies have been conducted in an attempt to understand the phenomena. Around 1900, Phillip Lenard did a series of critical studies on the relation of the kinetic energy of ejected electrons with the intensity and wavelength of the impinging light [?]. His results were in direct conflict with the wave theory of light.

Figure 2.2 shows schematically the experimental apparatus of Phillip Lenard. The entire setup was enclosed in a vacuum chamber. An electric

**Table 2.1: Stopping Voltage for Photocurrent**

| Rod material | Driving current (A) | Distance to target (cm) | Photocurrent (pA) | Stopping voltage (V) |
|--------------|---------------------|-------------------------|-------------------|----------------------|
| Carbon       | 28                  | 33.6                    | 276               | -1.07                |
| Carbon       | 20                  | 33.6                    | 174               | -1.12                |
| Carbon       | 28                  | 68                      | 31.7              | -1.10                |
| Carbon       | 8                   | 33.6                    | 4.1               | -1.06                |
| Zinc         | 27                  | 33.6                    | 2180              | -0.85                |
| Zinc         | 27                  | 87.9                    | 319               | -0.86                |

Source: P. Lenard, *Annalen der Physik*, **8**, 167 (1902) [?].

arc lamp, using carbon rods or zinc rods as the electrodes, generates strong UV light. A quartz window allows such UV light to shine on a target made of different metals. The target and a counter electrode are connected to an adjustable power supply. An ammeter is used to measure the electric current generated by the UV light, the *photocurrent*, especially when the voltages between the two electrodes are very small. By gradually increasing the voltage, which tends to reflect the electrons back to the target, the photocurrent is reduced. The voltage with which the photocurrent becomes zero is recorded as the *stopping voltage*.

The stopping voltage is apparently related to the kinetic energy of the electrons ejected from the target:

$$eV = \frac{1}{2}mv^2. \quad (2.2)$$

Understandably, the photocurrent varies with the intensity of light. By changing the magnitude of the current that drives the arc or the distance from the arc lamp to the target, the photocurrent could change by two orders of magnitude: for example, from 4.1 to 276 pA. An unexpected and dramatic effect Lenard observed was that no matter how strong or how weak the light is, and no matter how large or how small the photocurrent is, the stopping voltage does not change; see Table 2.1. The stopping voltage changes only when the material for the electric arc lamp changes. However, for a given type of arc, the stopping voltage stays unchanged.

The effect Lenard observed has no explanation in the framework of the wave theory of light. According to the wave theory of light, the more intense the light is, the more kinetic energy the electrons acquire.

### 2.1.2 Einstein's law of photoelectrical effect

Einstein proposed an explanation that the energy of light can only be transferred as an integer numtiple of energy quantum  $\epsilon$ , solely depending on its frequency  $f$ ,

$$\epsilon = hf, \quad (2.3)$$

where  $h = 6.63 \times 10^{-34}$  J·s is the Planck constant. Einstein also proposed a simple formula that can be tested experimentally. When a photon interacts with an electron in the metal, it transfers the radiation energy  $\epsilon$  to the electron. The electron could escape from the metal by overcoming the *work function*  $W$  of the metal, typically a few electron volts. If the energy of the photon is smaller than the work function of the metal, the electron would stay in the metal. If the energy of the photon is greater than the work function of the metal, then the electron can escape from the metal surface *with an excess kinetic energy*,

$$T = \frac{1}{2}mv^2 = hf - W. \quad (2.4)$$

Einstein took an example of near ultraviolet light,  $\lambda = 0.29 \mu\text{m}$ , and the frequency is  $1.03 \times 10^{15} \text{ s}^{-1}$ . The energy of the photon is  $6.9 \times 10^{-19} \text{ J}$ , or 4.3 eV, in line with the workfunction of the metal  $W$  used by Lenard. The kinetic energy of an escaping electron can be measured by an external voltage, or electric field, to turn it back onto the target. Voltage just enough to cancel the kinetic energy is called the *stopping voltage*,

$$eV_{\text{stop}} = \frac{1}{2}mv^2 = hf - W, \quad (2.5)$$

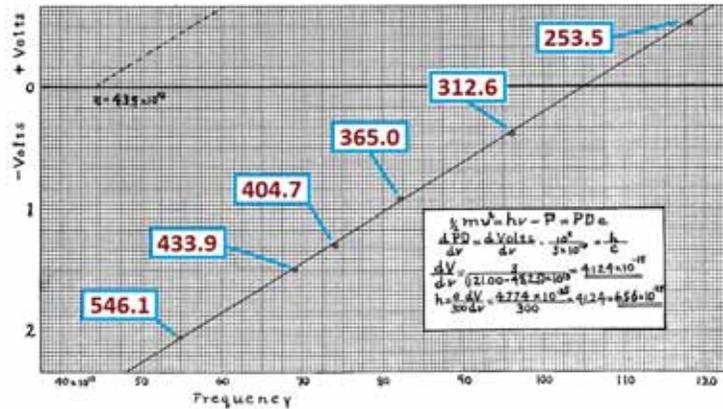
where  $e$  is the electron charge,  $1.60 \times 10^{-19} \text{ C}$ . According to Einstein's theory of light quantum, the stopping voltage is linearly dependent on the frequency of the radiation and *independent of the intensity of light*. The slope should be a universal constant, which provides a direct method to determine the value of the Planck constant  $h$ ,

$$\frac{\Delta V_{\text{stop}}}{\Delta f} = \frac{h}{e}. \quad (2.6)$$

The quantities in Eq. 2.6 are well-defined. It can be verified by direct experiments and to obtain the value of the Planck constant.

### 2.1.3 Millikan's experimental verification

Einstein's theory of photons was rejected by a number of prominent physicists for many years, including Max Planck, Niels Bohr, and notably Robert Millikan. Starting in 1905, for 10 years Millikan tried very hard to experimentally disprove Einstein's theory. Comparing with the primitive experiments



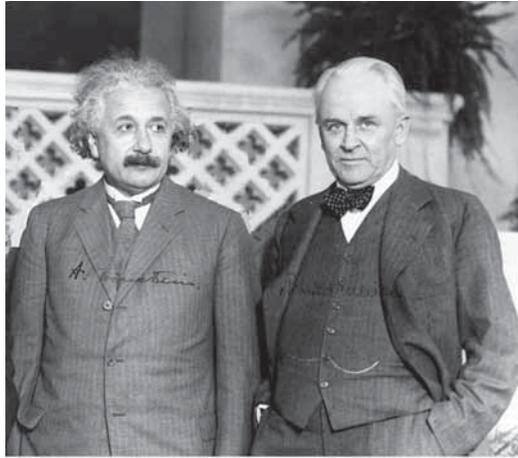
**Fig. 2.3. Result of Millikan's experiment on photoelectric effect.** Figure 3 of Millikan's 1916 *Physical Review* paper on experiment on photoelectric effect. Six spectral lines of a mercury lamp are used, with wavelengths in nm marked with the blue boxes. Planck's constant was directly obtained from the slope of the curve with frequencies of the radiation and the stopping voltages. [?].

of Philip Lenard, Millikan had much more resources to do precision measurements to verify the Planck-Einstein relation  $\epsilon = hf$ . Using six spectral lines from a powerful mercury lamp with wavelengths carefully measured, showing in the blue boxes, Millikan's determination of the Planck's constant was definitive and straightforward. In 1916, Millikan published a long paper on *Physical Review*, entitled *A Direct Photoelectric Determination of Planck's h* [?]. The conclusion reads as follows:

1. Einstein's photoelectric equation has been subject to very searching tests and it appears in every case to predict exactly the observed results.
2. Planck's  $h$  has been photoelectrically determined with a precision of about .5 percent.

An interesting fact is that in the same paper Millikan emphatically rejected Einstein's theory of photons. He said that Einstein's photon hypothesis "may well be called reckless first because an electromagnetic disturbance which remains localized in space seems a violation of the very conception of an electromagnetic disturbance, and second because it flies in the face of the thoroughly established facts of interference." Millikan wrote that Einstein's photoelectric equation, although accurately representing the experimental data, "cannot in my judgment be looked upon at present as resting upon any sort of a satisfactory theoretical foundation [?]."

Millikan's objection was not all unreasonable. Einstein was clear in his 1905 paper that the quantization of radiation energy only occurs at its



**Fig. 2.4. Albert Einstein and Robert Millikan.** Both Einstein and Millikan won a Nobel Prize for their contributions to the photoelectric effect. Photograph taken in 1930 when Robert Millikan invited Albert Einstein to a conference in California. Original photograph courtesy of Smithsonian Museum, slightly cleaned up by the author.

emission and conversion; in other words, only occurs upon interacting with atomic systems. Nevertheless, there is a widespread misunderstanding that light *behaves like Newtonian material points during propagation in the free space*. According to such a popular point of view, a photon can be described as a geometrical point with a well-defined trajectory in space.

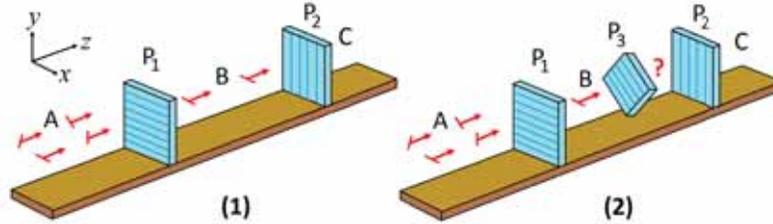
The experiments of Millikan only proved that when light interacts with atomic systems, the amount of radiation energy transfer is always an integer multiple of the light quantum  $\epsilon$ , determined by its frequency  $f$ . It is not a proof that light is a spray of particles during propagation.

In the following subsection, the nature of light quantum is clarified through the analysis of two examples, three-polarizer experiment of Paul Dirac, and the double-slit interference experiment of Thomas Young.

#### 2.1.4 Nature of the energy quantum of light

As emphasized by Einstein in his Nobel-Prize winning 1905 paper, energy quantization occurs only during the emission and transformation of radiation. During propagation, as a wave, light follows the Maxwell theory of electromagnetism, never behaves like a spray of Newtonian material points, as such a view would lead to absurdity.

Consider first the Dirac three-polarizer experiment discussed in Section 1.3.6. By passing a beam of natural light through a polarizer, the outgoing light in compartment B is fully polarized, for example, in  $x$ -direction. By placing a second polarizer in  $y$ -direction, the light is blocked. However,



**Fig. 2.5. Paradox of Dirac's three-polarizer experiment.** The concept that light is composed of individual material points during propagation in free space leads to absurdity, as shown by the three-polarizer experiment. (1) In compartment A, the photons of the natural light are a mixture of different polarizations. After passing polarized P1, all photons are  $x$ -polarized. No such photons can pass the  $y$ -polarizer P2, as expected. (2) By inserting a polarizer P3 in compartment B with a slanted direction, some of the photons can go through. However, there is no mechanism to switch those photons' polarization to  $y$ -direction. Therefore, there is no light comes out in compartment C. To solve this paradox, one should stand by Maxwell's theory that during propagation, light is an electromagnetic wave, not as a spray of particles.

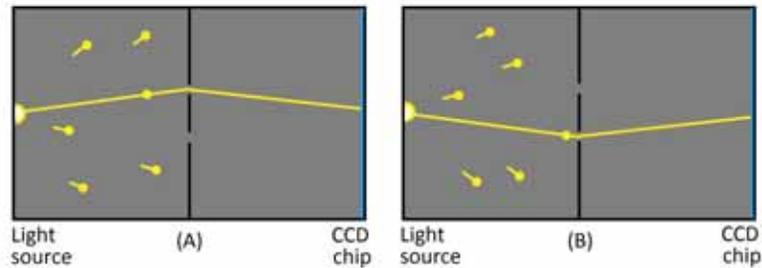
by placing a third polarizer with a slanted polarization angle, light can go through. Maxwell's theory of light provides a perfect quantitative explanation to the results, see Section 1.3.6. Note that Maxwell's theory is strictly linear, thus the result is independent of the intensity of light.

If light is a stream of material points, when the intensity of light is weak, a light beam is composed of widely separated individual photons. After the first polarizer, each photon should be polarized in  $x$ -direction. As shown in Fig. 2.5(1), no such photons can go through the  $y$ -polarizer. From a particle point of view, this is still understandable.

By inserting a third polarizer P3 with a slanted angle between P1 and P2, experimentally, some light goes through into compartment C. The particle view encounters difficulties in its interpretation. Because each photon in compartment B is polarized in  $x$ -direction, with a slanted polarization angle, some of the photons can go through. However, there is no mechanism to switch the polarization of each  $x$ -polarized particle to  $y$ -direction. Therefore, the particle view is in conflict with experimental observations.

The solution to this paradox is to stick to Einstein's original statement, that radiation energy is quantized only during emission and transformation. During the process of propagating, for example, in compartment (B), light is a wave governed by Maxwell's theory, not a spray of particles.

Second, consider the double-slit experiment, as shown in Fig. 2.6. It is known for many decades that even at extremely low intensity, that individual photons are observed on a sensitive light sensor such as a CCD chip, interference still persists. If light is composed of individual particles, each such particle can only go through one of the two slits. Interference can never happen. To resolve this paradox, one should stand by Einstein's original



**Fig. 2.6. Paradox of the double-slit experiment.** The concept that light is composed of individual particles during propagation in free space leads to absurdity, as shown by the double-slit experiment. According to the particle point of view, each photon can only pass through one of the two slits. When the light intensity is very low, the photon passes slit (A) and the photon passes slit (B) are unrelated. No interference pattern can be observed. To resolve this paradox, one should stand by Einstein's statement in his original publication that during propagation, light as an electromagnetic wave, follows Maxwell's theory of light, not as a spray of particles.

statement that during propagation, light follows Maxwell's theory of light as an electromagnetic wave, not as a spray of particles.

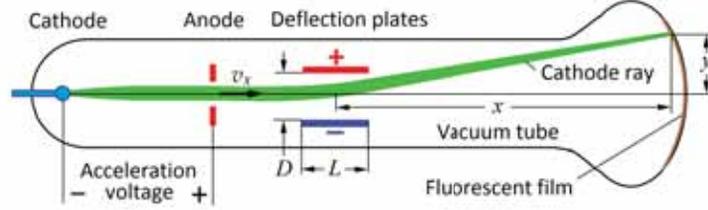
In the advanced quantum theory of light, quantum electrodynamics, outlined in Chapter 11, Einstein's original view is represented precisely by sophisticated mathematics. For example, in compartment B of Fig. 2.5, light is represented by standing waves. Each of the standing waves spreads out in the entire compartment evenly. And yet the energy of each mode of standing wave is quantized. The quantized electromagnetic field still conforms to the rules of superposition as a consequence of Maxwell's equations. However, upon interacting with atomic systems, radiation energy can only be transferred by an integer multiple of the light quantum.

## 2.2 Electron as a quantized wave

### 2.2.1 Experimental studies of the cathode ray

In 1897, J. J. Thomson studied the stream emitted from the cathode (negative electrode) in a vacuum tube, see Fig. 2.7. Naturally, it is called *cathode ray*. Typically, the cathode is heated with a filament to red hot. The cathode ray is then accelerated by a positive anode, then flows through a hole. Although cathode ray is not perceptible by naked eye, when it hits a fluorescent film inside the vacuum tube, visible light is generated.

To study the properties of the cathode ray, Thomson placed a pair of electrodes on the way the cathode ray passes. By applying a voltage on the pair of electrodes, the cathode ray is deflected, indicating as a stream of negatively charged material particles. Assuming each such particle has



**Fig. 2.7. Schematics of J. J. Thomson's experiment.** In a vacuum tube, the cathode (negative electrode), especially when heated, emits a ray of negatively charged material towards the anode (positive electrode). By placing a pair of deflection plates in its path, the cathode ray bends toward the positive plate. The charge-to-mass ratio of the particles can be determined by direct measurements.

an electrical charge  $e$  and mass  $m_e$ , the *charge-to-mass ratio*  $e/m_e$  of the cathode ray can be determined experimentally, see Fig. 2.7.

Before entering to the field of the deflection plates, the particles are accelerated to a velocity in the  $x$ -direction,  $v_x$ . Denote the distance of the two deflecting electrodes as  $D$ , with a deflecting voltage  $V$ , the force acting on each particle is

$$F = \frac{eV}{D}. \quad (2.7)$$

Denote the length of the deflection electrodes as  $L$ , the time the particle passes through is

$$t = \frac{L}{v_x}. \quad (2.8)$$

According to Newton's law, Eq. 1.2, after passing the deflection plates, the particle gains a velocity in the  $y$ -direction,

$$v_y = \frac{Ft}{m_e} = \frac{eVL}{m_e v_x D}. \quad (2.9)$$

By traveling through a distance  $x$  to the fluorescent screen, the  $y$ -displacement of the particle is determined by

$$\frac{y}{x} = \frac{v_y}{v_x} = \frac{eVL}{m_e v_x^2 D}. \quad (2.10)$$

The charge-to-mass ratio is then

$$\frac{e}{m_e} = \frac{v_x^2 D y}{VLx}. \quad (2.11)$$

Through experiments, Thomson found the ratio is

$$\frac{e}{m_e} \approx 1.759 \times 10^{11} \text{ C/Kg}. \quad (2.12)$$

In the 19th century, there was a controversy of whether the cathode ray is a wave or a stream of particles. The interpretation of Thomson's experiment is based on Newtonian mechanics. However, soon it was found that the electron beam is a wave, and the Young's double-slit experiment of interference was observed, especially in the electron microscopes.

In the 20th century, the cathode ray found a widespread application in TV sets, computer displays, and oscilloscopes, called CRT (cathode-ray tubes). Although CRT was largely replaced by the solid-state displays, one application of the cathode ray is still alive and strong, that is the electron microscope. The resolution of the optical microscopes is limited by the wavelength of the visible light, which is a fraction of a micron. The wavelength of electron beams can be much smaller than a nanometer. Nevertheless, in the interpretation of high-resolution electron microscopy, the wave nature of electrons is indispensable.

The double-slit interference experiments for electrons were conducted even with very weak electron beams, where from a naïve point of view the electrons are well separated. The observed interference pattern indicates that in free space, the electrons must be described as waves. In an advanced formalism of quantum mechanics, the path-integral approach of Feynman, outlined in Chapter 12, the interference effect in the double-slit experiment is natural: For an electron to travel from point A to point B in space, all possible paths in the entire open space must be taken into account.

### 2.2.2 Millikan's oil-drop experiment

In 1923, the Nobel Committee announced that Robert Millikan won the Nobel Prize in physics "for his work on the elementary charge of electricity and on the photoelectric effect". Before working on verifying Einstein's law of photoelectric effect, he spent five years (1908-1913) to study the quantization of electric charge.

Figure 2.8 shows a schematics of the experimental setup. A spray gun creates tiny oil drops from the upper chamber, then fall down to the lower chamber through a small hole. The X-ray ionizes the air and generates free electrons. Some of the electrons are attached to some oil drops. A power supply provides a controllable voltage on the electrodes.

Each oil drop is subject to gravity force  $mg$ , pulling downwards, see Fig.2.8(A); and the electrostatic force from the high-voltage electrodes on the electrical charges in the oil drop. If electrical charge is quantized with a minimum unit  $e$ , then the electrical force should depend on the number of electrons in the oil drop, see Fig.2.8(B) and (C). By adjusting the voltage  $V$ , an oil drop could stay immobile when the net force is zero

$$\frac{neV}{D} = mg, \quad (2.13)$$

where  $n$  is the number of electrons in the oil drop,  $V$  is the electrical field intensity and  $D$  is the distance between the two electrodes.

The mass of the oil drop can be determined by turning off the voltage, then the oil drop is subject to gravity and the viscosity force of air. Because the viscosity force is proportional to the radius  $r$  and the mass is proportional to  $r^3$ , the radius of the oil drop can be determined by the steady speed of free fall, and its mass can be calculated.

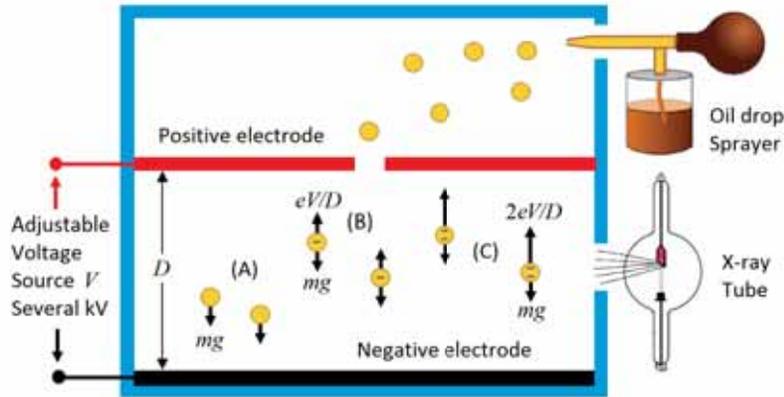
In 1913, Millikan published a paper on *Physical Review*, entitled *On the Elementary Electrical Charge and the Avogadro Constant*. He reported that the observed charge on the oil drop “showed a very exact multiple relationship under all circumstances – a fact which demonstrated very directly the atomic structure of the electric charge”. The value he reported is within 1% to the currently recognized value,

$$e = 1.603 \times 10^{-19} \text{ C.} \quad (2.14)$$

Combining with the value of charge-to-mass ratio, Eq. 2.12, the mass of the electron is

$$m_e = 9.109 \times 10^{-31} \text{ Kg.} \quad (2.15)$$

Because the value of the elementary charge is the Faraday constant divided



**Fig. 2.8.** Schematics of Millikan’s oil-drop experiment. Using an atomizer, drops of oil is formed in the upper chamber of the setup. Some of the oil drops fall through a hole to the lower chamber. The air in the lower chamber can be ionized by X-ray to generate free electrons. The electrons can attached to the oil drops to make it charged. (A), without attached electrons, an oil drop is pulling downwards by gravity force  $mg$ . (B), with one attached electron, the oil drop is pulling upwards by an electric force  $eV/D$ . (C), with two attached electrons, the oil drop is pulling upwards by an electric force  $2eV/D$ . The electric charge can be determined by adjusting the voltage to make the oil drop stationary. The schematics is not on scale, as the radius of the oil drops is only about a micron.

by the Avogadro constant, Millikan's experiment also gave the most accurate value of Avogadro constant at that time.

Millikan's oil-drop experiment proved that electrical charge is quantized: the total electrical charge on any isolated subject must be an integer multiple of the elementary charge  $e$ . It does not mean that an electron is a geometrical point. On the contrary, to reduce the total energy of the system, the charge of a single electron is spread over the entire oil drop.

### 2.2.3 The de Broglie wave

Several years before the wave nature of electrons were observed experimentally, French physicist Louis de Broglie extended Einstein's postulation that light can be both wave and quantum to *all particles*, including the electron, in his 1924 Ph.D. thesis entitled *Recherches sur la théorie des quanta*. Here is a heuristic argument leads to the wave property of electrons.

As shown in Eq. 1.59, a general form of a wave is

$$u(x, t) = a \sin(kx - \omega t + \phi), \quad (2.16)$$

here  $\omega = 2\pi f$  is the angular velocity, and  $k = \omega/v$  is the wavevector. According to Einstein, from Eq. 2.1, for light, the energy quantum  $\epsilon$  is

$$\epsilon = hf = \hbar\omega, \quad (2.17)$$

where  $\hbar$  is a reduced Planck's constant often called the *Dirac constant*

$$\hbar \equiv \frac{h}{2\pi} = 1.054 \times 10^{-34} \text{ J} \cdot \text{s}. \quad (2.18)$$

To describe a wave, a wave vector  $k$  is also required. Intuitively, de Broglie guessed that  $k$  is related to the momentum  $p$  of the particle, and the relation

**Fig. 2.9. Louis de Broglie.** French physicist (1892–1987), proposed a theory of general wave-quantum duality in his 1924 Ph.D. thesis, then won the 1929 Nobel Prize in physics. He was born to a noble family in France and became the 7th duc de Broglie in 1960. In 1942, he was elected as the Perpetual Secretary of the French academy of sciences. After the second world war, he proposed the establishment of multi-national research laboratory, leading to the establishment of European Organization for Nuclear Research (CERN).



can be obtained through a heuristic argument as follows. In general, the momentum  $p$  is defined as in Eq. 1.3,

$$p = mv. \quad (2.19)$$

For a photon with energy  $\epsilon$ , according to Einstein's relation between mass and energy, the mass is

$$m = \frac{\epsilon}{c^2}. \quad (2.20)$$

Because the velocity of the photon is  $c$ , its momentum is

$$p = mc = \frac{\epsilon}{c}. \quad (2.21)$$

Using Eq. 2.17, Eq. 2.21 becomes

$$p = \frac{\hbar\omega}{c}. \quad (2.22)$$

According to Eq. 1.61, the relation between angular frequency and wave vector is through the velocity of the wave  $v$ ,

$$k = \frac{\omega}{v}. \quad (2.23)$$

Because the speed of light is  $c$ , Eq. 2.22 becomes

$$p = \hbar k. \quad (2.24)$$

Equation 2.24 is the second equation correlating a particle with a wave as de Broglie proposed for all particles, including electrons.

Therefore, as proposed by de Broglie, both light and cathode ray – electrons – are waves. For light, as an electromagnetic wave, the energy is quantized. The elementary unit, the energy quantum, is proportional to the frequency. For electrons, as a wave, the electrical charge is quantized. The unit is a universal constant, the elementary charge  $e$ .

According to the historical records, in 1925, Einstein received a preprint of de Broglie's thesis. Einstein highly appreciated the idea, then immediately sent a letter to Schrödinger for his attention. In just a few months, based on the idea of de Broglie, Schrödinger formulated his wave equation of electrons and derived the Rydberg formula for the hydrogen atom. Quantum mechanics in its most productive form was born.

## Chapter 3

### The Schrödinger equation

The year 1905 was Albert Einstein's *annum mirabilis* when he published four papers on *Annalen der Physik* that literally started the modern physics. Similarly, the year 1926 was Erwin Schrödinger's *annum mirabilis* with five papers published on the same journal *Annalen der Physik* that defined non-relativistic quantum mechanics. Thus commented Paul Dirac in 1929: "The underlying physical laws necessary for the mathematical theory of a large part of physics and the whole of chemistry are thus completely known". Those papers, belonging to the defining publications of modern science, are still worth reading. Here is a brief summary of the first two:

The first paper, *Quantization as an Eigenvalue Problem, Part I*, defined a wavefunction  $\psi$ , which is "everywhere real, single-valued, finite, and continuously differentiable up to the second order", as a offspring of the action function  $S$ . Then, a differential equation of  $\psi$  is presented as a variation of the Hamilton-Jacobi equation of the action function  $S$ . By applying that equation to the hydrogen atom, the Rydberg formula was explained.

The second paper, *Quantization as an Eigenvalue Problem, Part II*, presented a parallelism of the relation of quantum mechanics and classical mechanics with the relation of wave optics and geometrical optics. He emphasized that the wave nature of electrons is fundamental, and the particle view is a macroscopic approximation. He also presented an even simpler way to introduce the differential equation of the wavefunction based on de Broglie



**Fig. 3.1.** Austrian banknote with a portrait of Schrödinger. It is a rare honor for a scientist to have a portrait printed on a banknote. Note the large value.

**3. Quantisierung als Eigenwertproblem;  
von E. Schrödinger.**

(Erste Mitteilung.)

§ 1. In dieser Mitteilung möchte ich zunächst an dem einfachsten Fall des (nichtrelativistischen und ungestörten) Wasserstoffatoms zeigen, daß die übliche Quantisierungsvorschrift sich durch eine andere Forderung ersetzen läßt, in der kein Wort von „ganzen Zahlen“ mehr vorkommt. Vielmehr ergibt sich die Ganzzahligkeit auf dieselbe natürliche Art, wie etwa die Ganzzahligkeit der *Knoten*zahl einer schwingenden Saite. Die neue Auffassung ist verallgemeinerungsfähig und rührt, wie ich glaube, sehr tief an das wahre Wesen der Quantenvorschriften.

Die übliche Form der letzteren knüpft an die Hamiltonsche partielle Differentialgleichung an:

$$1) \quad H\left(q, \frac{\partial S}{\partial q}\right) = E.$$

Es wird von dieser Gleichung eine Lösung gesucht, welche sich darstellt als *Summe* von Funktionen je einer einzigen der unabhängigen Variablen  $q$ .

Wir führen nun für  $S$  eine neue unbekannte  $\psi$  ein derart, daß  $\psi$  als ein *Produkt* von eingriffigen Funktionen der einzelnen Koordinaten erscheinen würde. D. h. wir setzen

$$2) \quad S = K \lg \psi.$$

Die Konstante  $K$  muß aus dimensionellen Gründen eingeführt werden, sie hat die Dimension einer *Wirkung*. Damit erhält man

$$1) \quad H\left(q, \frac{K}{\psi} \frac{\partial \psi}{\partial q}\right) = E.$$

**Fig. 3.2.** First page of Schrödinger's first paper on quantum mechanics. The first paragraph said: "In this communication, I would first treat a simple case, the non-relativistic and unperturbed hydrogen atoms. ... Quantization comes up naturally similar to the *nodes* of a vibrating string. I believe that the new understanding can be generalized to gain deep insight to the nature of quantum phenomena". Schrödinger created a wavefunction  $\psi$  as an offspring of the action  $S$ , that is "everywhere real, single-valued, finite, and continuously differentiable up to the second order".

duality. Three further problems were treated: the harmonic oscillator, rigid rotor and non-rigid rotor, all related to molecular physics.

Because for applications in atomic physics, chemistry, molecular biology, and solid-state electronics, time-independent Schrödinger's equation with real wavefunctions are almost always sufficient, we will not present time-dependent phenomena until Part III of the book, Supplements.

### 3.1 The real Schrödinger equation

In the first two 1926 papers of Erwin Schrödinger, a time-independent wave equation with real wavefunctions  $\psi$  was created as a variation of the equation of motion in classical mechanics Eq. 1.11 using the de Broglie relation Eq 2.24. Two basic problems, the hydrogen atom and the harmonic oscillator, were resolved. Here we follow Schrödinger's footsteps.

Consider one-dimensional case first. A time-independent wavefunction  $\psi(x)$  is a standing wave, satisfying the Helmholtz equation, Eq. 1.69:

$$\frac{d^2\psi(x)}{dx^2} = -k^2\psi(x). \quad (3.1)$$

According to Eq. 2.24, de Broglie's relation between momentum  $p$  and wave vector  $k$  is,

$$p = \hbar k. \quad (3.2)$$

Following Eqs. 3.1 and 3.2, the square of momentum is

$$p^2 = -\frac{\hbar^2}{\psi(x)} \frac{d^2\psi(x)}{dx^2}. \quad (3.3)$$

Following Eq. 1.11, for an electron with mass  $m_e$  moving in a potential energy function  $V$ , a key equation in classical mechanics is

$$\frac{p^2}{2m_e} + V(x) = E. \quad (3.4)$$

Insert the expression of  $p^2$  in Eq. 3.3 into Eq. 3.4, multiply both sides by  $\psi(x)$ , a differential equation for wavefunction  $\psi(x)$  is obtained,

$$-\frac{\hbar^2}{2m_e} \frac{d^2\psi(x)}{dx^2} + V(x)\psi(x) = E\psi(x). \quad (3.5)$$

This is the *one-dimensional time-independent Schrödinger equation*.

It straightforward to generalize Eq. 3.5 to three-dimensional space. Similar to Eq. 1.76, the wave equation in three-dimensional space is

$$\frac{\partial^2\psi(\mathbf{r})}{\partial t^2} = v^2\Delta\psi(\mathbf{r}), \quad (3.6)$$

with the Laplace operator  $\Delta$  defined as

$$\Delta \equiv \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}. \quad (3.7)$$

Similar to Eqs. 1.78 and 1.79, the Helmholtz equation is

$$\Delta\psi(\mathbf{r}) = -k^2\psi(\mathbf{r}). \quad (3.8)$$

According to Eq. 3.2, the square of momentum is

$$p^2 = -\frac{\hbar^2}{\psi(\mathbf{r})}\Delta\psi(\mathbf{r}). \quad (3.9)$$

The three-dimensional time-independent Schrödinger equation is then

$$-\frac{\hbar^2}{2m_e}\Delta\psi(\mathbf{r}) + V(\mathbf{r})\psi(\mathbf{r}) = E\psi(\mathbf{r}). \quad (3.10)$$

Equations 3.5 and 3.10 are the Schrödinger equations sufficient for almost all practical problems in the real world, including atomic physics, chemistry, molecular biology, and solid-state electronics.

### 3.2 Electrons in a one-dimensional box

To illustrate the meanings of the Schrödinger equation and wavefunctions, the problem of electrons in a one-dimensional box is studied. Assume the length of the box is  $L$ . Within the box,  $0 < x < L$ , the potential is zero. Schrödinger's equation is

$$-\frac{\hbar^2}{2m_e} \frac{d^2\psi(x)}{dx^2} = E\psi(x). \quad (3.11)$$

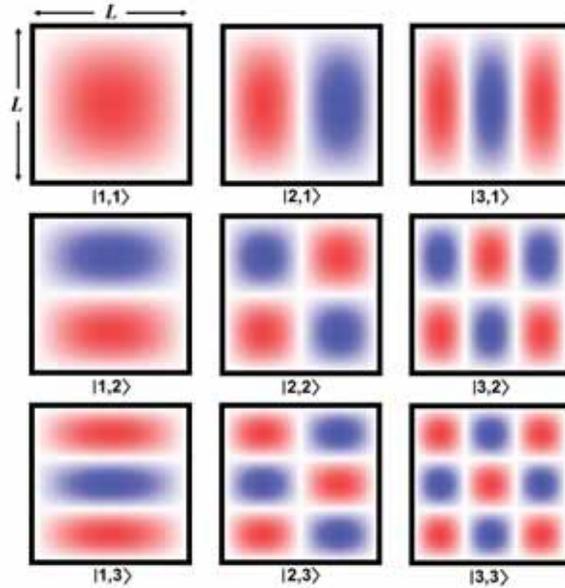
Outside the box, the potential is high. No electrons can exist outside the box. The situation sets up the boundary conditions for the solutions of Eq. 3.11: at the boundaries,  $x = 0$  and  $x = L$ , the wavefunction must be zero.

Introduce a wave vector  $k$  defined as

$$k = \frac{\sqrt{2m_e E}}{\hbar}, \quad (3.12)$$

the Schrödinger's equation, Eq. 3.11 becomes

$$\frac{d^2\psi(x)}{dx^2} = -k^2\psi(x). \quad (3.13)$$



**Fig. 3.3. Wavefunctions in a one-dimensional box.** The wavefunctions are labeled by the two quantum numbers,  $n_x$  and  $n_y$ . For clarity, the Dirac notations are applied, see Section 3.3. The nodal structures in the wavefunctions largely determines their characteristics. Wavefunctions with different  $n_x$  and/or  $n_y$  are orthogonal, see Eq. ??.

The general solution of Eq. 3.13 is

$$\psi(x) = C \sin(kx + \phi), \quad (3.14)$$

where  $C$  is a normalization constant, and  $\phi$  is a phase angle. Because at  $x = 0$ , the wavefunction must be zero, the only possible solution is with  $\phi = 0$ :

$$\psi(x) = C \sin(kx). \quad (3.15)$$

The boundary condition that the wavefunction is zero at the other end of the box,  $x = L$ , requires that

$$k = \frac{n\pi}{L}, \quad (3.16)$$

where  $n$  is an integer.

In classical mechanics, as in Eq. 3.4, energy can take any value. In quantum mechanics, because of condition Eq. 3.16, energy is *quantized*: it can only take one of the discrete values determined by Eqs. 3.12 and 3.16:

$$E_n = \frac{n^2 \hbar^2 \pi^2}{2m_e L^2}, \quad (3.17)$$

where  $n = 1, 2, 3, \dots$  is an integer.

Those allowed values of energy are called the *energy eigenvalues*, adapted from German, the *proper values of energy*.

Because the Schrödinger equation is linear to the wavefunction, the constant  $C$  does not affect the determination of energy eigenvalues. According to Schrödinger, the square of the wavefunction is correlated to the charge density of an electron in the space:

$$\rho = e|\psi|^2. \quad (3.18)$$

Because the total charge of an electron over the space equals to one elementary charge  $e$ , the integral over the entire space must equal to 1. In the current situation,

$$\int_0^L |\psi(x)|^2 dx = 1. \quad (3.19)$$

Because the average value of the square of sine function over any number of half periods is  $1/2$ , for all quantum numbers, the constant is

$$C = \sqrt{\frac{2}{L}}. \quad (3.20)$$

The wavefunctions are

$$\psi_n(x) = \sqrt{\frac{2}{L}} \sin \frac{n\pi x}{L}. \quad (3.21)$$

It is straightforward to show that eigenstates with different quantum numbers  $n$  are orthogonal. In fact, using Eq. 3.21,

$$\int_0^L dx \psi_n(x) \psi_m(x) = \delta_{nm}, \quad (3.22)$$

which is zero only when  $n \neq m$ .

### 3.3 The Dirac notation

The notations of wavefunctions, Eq. 3.21, and the integrals, Eqs. 3.19 and 3.22, occurs very often in quantum mechanics. In the third edition of *Principles of Quantum Mechanics*, Dirac introduced the bra and ket notations, that greatly simplifies mathematical notations in quantum mechanics. In Part I and Part II of this book, wavefunctions are real. A wavefunction can be denoted either by a bra or by a ket. For the case of electrons in a two-dimensional box, it is either

$$\langle n| = \sqrt{\frac{2}{L}} \sin \frac{n\pi x}{L}, \quad (3.23)$$

or

$$|n\rangle = \sqrt{\frac{2}{L}} \sin \frac{n\pi x}{L}. \quad (3.24)$$

A complete bracket represents an integral of two wavefunctions,

$$\langle n|m\rangle = \int_0^L dx \psi_n(x) \psi_m(x). \quad (3.25)$$

The orthogonization and normalization condition in Dirac notation is simply

$$\langle n|m\rangle = \delta_{nm}. \quad (3.26)$$

The Schrödinger equation, Eq. ??, can be written as

$$\hat{H}|\psi\rangle = E|\psi\rangle \quad (3.27)$$

by defining a *Hamiltonian operator* or simply *Hamiltonian*

$$\hat{H} \equiv -\frac{\hbar^2}{2m_e} \frac{d^2}{dx^2}. \quad (3.28)$$

A hat is marked to denote that it is not an ordinary number, but an operator that makes sense only when acting upon a wavefunction.

### 3.4 Harmonic Oscillator

In quantum mechanics, the harmonic oscillator is the most prevailing problem. It describes the oscillation of molecules and solids near its equilibrium point. The electromagnetic wave can be decomposed into a number of simple harmonic oscillators. Using the quantization procedure presented here, the electromagnetic waves can be quantized. It is the basis of quantum electrodynamics, the complete theory of radiation and matter.

From Eqs. 1.26 and 3.5, the Schrödinger equation for a one-dimensional harmonic oscillator is

$$\left(-\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + \frac{m}{2} \omega^2 x^2\right) \psi(x) = E\psi(x). \quad (3.29)$$

By introducing a dimensionless coordinate  $q$  defined as

$$q \equiv \sqrt{\frac{m\omega}{\hbar}} x, \quad (3.30)$$

the Schrödinger equation Eq. 3.29 becomes

$$\frac{1}{2} \left(-\frac{d^2}{dq^2} + q^2\right) \hbar\omega\psi(q) = E\psi(q). \quad (3.31)$$

#### 3.4.1 Creation operator and annihilation operator

As we have presented in previous sections, the basic concept in quantum mechanics is the wavefunction, and it is governed by a partial differential equation, the Schrödinger equation. In some cases, the process of obtaining solutions of partial differential equations can be greatly simplified by using differential operators with algebraic methods. This is especially true for the solution of the harmonic oscillator.

In order to find an algebraic solution of the harmonic oscillator, a pair of operators are introduced: an *annihilation operator*,

$$\hat{a} = \frac{1}{\sqrt{2}} \left(q + \frac{d}{dq}\right), \quad (3.32)$$

and a *creation operator*

$$\hat{a}^\dagger = \frac{1}{\sqrt{2}} \left(q - \frac{d}{dq}\right). \quad (3.33)$$

The meanings of these terms will be clarified soon. Those operators not only greatly simplify the solution of the quantum-mechanical harmonic oscillator problem, but also serve the basis of quantum field theory.

By acting on any function of  $q$ , a simple algebra shows that the two operators satisfy the following *commutation relation*,

$$[\hat{a}, \hat{a}^\dagger] \equiv \hat{a}\hat{a}^\dagger - \hat{a}^\dagger\hat{a} = 1. \quad (3.34)$$

Also by a simple algebra, the Schrödinger equation Eq. 3.31 becomes

$$\left(\hat{a}^\dagger\hat{a} + \frac{1}{2}\right)\hbar\omega\psi(q) = E\psi(q). \quad (3.35)$$

### 3.4.2 Algebraic solution of the Schrödinger equation

In the following, we will show how to utilize Eq. 3.34 to solve the Schrödinger equation Eq. 3.35. First of all, it is sufficient to resolve the following algebraic equation. In Dirac notation (see Section 3.3), it is

$$\hat{a}^\dagger\hat{a}|n\rangle = u_n|n\rangle. \quad (3.36)$$

Here the eigenstates are labeled by a number  $n$  with eigenvalue  $u_n$ . By comparing Eq. 3.35 with Eq. 3.36, the energy eigenvalues are

$$E_n = \left(u_n + \frac{1}{2}\right)\hbar\omega. \quad (3.37)$$

As a consequence of Eq. 3.34, if  $|n\rangle$  is an eigenstate with eigenvalue  $u_n$ , then  $\hat{a}|n\rangle$  is also an eigenstate,

$$\hat{a}^\dagger\hat{a}\hat{a}|n\rangle = (\hat{a}\hat{a}^\dagger - 1)\hat{a}|n\rangle = (u_n - 1)\hat{a}|n\rangle \quad (3.38)$$

with eigenvalue  $u_n - 1$ . Because  $\langle n|\hat{a}^\dagger\hat{a}|n\rangle$  must not be negative, there must be an eigenstate with minimum value,  $u_n = 0$ . For such a state,

$$\hat{a}^\dagger\hat{a}|0\rangle = 0. \quad (3.39)$$

On the other hand, also as a consequence of Eq. 3.34, if  $|n\rangle$  is an eigenstate with eigenvalue  $u_n$ , then  $\hat{a}^\dagger|n\rangle$  is also an eigenstate

$$\hat{a}^\dagger\hat{a}\hat{a}^\dagger|n\rangle = \hat{a}^\dagger(\hat{a}^\dagger\hat{a} + 1)|n\rangle = (u_n + 1)\hat{a}^\dagger|n\rangle \quad (3.40)$$

with eigenvalue  $u_n + 1$ . Starting with the lowest eigenstate  $|0\rangle$ , by applying  $\hat{a}^\dagger$  many times, we have

$$\hat{a}^\dagger\hat{a}(\hat{a}^\dagger)^n|0\rangle = n(\hat{a}^\dagger)^n|0\rangle. \quad (3.41)$$

Because the eigenvalue of the zeroth eigenstate  $|0\rangle$  of the operator  $\hat{a}^\dagger\hat{a}$  is zero, and each time a creation operator  $\hat{a}^\dagger$  is applied, the eigenvalue is added by 1, the eigenvalues of the operator  $\hat{a}^\dagger\hat{a}$  equals to an integer, the

number of times the a creation operator applied. Therefore, except for a numerical constant, it is the  $n$ -th eigenstate of the operator  $\hat{a}^\dagger \hat{a}$ :

$$|n\rangle = C_n (\hat{a}^\dagger)^n |0\rangle. \quad (3.42)$$

The operator  $\hat{a}^\dagger \hat{a}$  is deservably called the *particle number operator*,

$$\hat{N} \equiv \hat{a}^\dagger \hat{a}, \quad (3.43)$$

because its eigenvalue is the number of energy quanta,

$$\hat{N}|n\rangle = n|n\rangle. \quad (3.44)$$

The normalization constant  $C_n$  can be determined as follows. The zeroth-order state is by definition normalized,

$$\langle 0|0\rangle = 1. \quad (3.45)$$

By applying  $\hat{a}^\dagger$  many times, we have

$$\langle 0|(\hat{a})^n (\hat{a}^\dagger)^n |0\rangle = n!. \quad (3.46)$$

Therefore,  $C_n = (n!)^{-1/2}$ , and

$$|n\rangle = \frac{1}{\sqrt{n!}} (\hat{a}^\dagger)^n |0\rangle. \quad (3.47)$$

**Table 3.1: Wavefunctions of the harmonic oscillator**

| State       | Energy                    | wavefunction  |
|-------------|---------------------------|---|
| $ 0\rangle$ | $\frac{1}{2}\hbar\omega$  | $\frac{1}{\sqrt{\sqrt{\pi}}} e^{-q^2/2}$                        |
| $ 1\rangle$ | $\frac{3}{2}\hbar\omega$  | $\frac{1}{\sqrt{2\sqrt{\pi}}} 2qe^{-q^2/2}$                     |
| $ 2\rangle$ | $\frac{5}{2}\hbar\omega$  | $\frac{1}{\sqrt{2\sqrt{\pi}}} (2q^2 - 1) e^{-q^2/2}$            |
| $ 3\rangle$ | $\frac{7}{2}\hbar\omega$  | $\frac{1}{\sqrt{3\sqrt{\pi}}} (2q^3 - 3q) e^{-q^2/2}$           |
| $ 4\rangle$ | $\frac{9}{2}\hbar\omega$  | $\frac{1}{\sqrt{24\sqrt{\pi}}} (4q^4 - 12q^2 + 3) e^{-q^2/2}$   |
| $ 5\rangle$ | $\frac{11}{2}\hbar\omega$ | $\frac{1}{\sqrt{60\sqrt{\pi}}} (4q^5 - 20q^3 + 15q) e^{-q^2/2}$ |

Following Eq. 3.37, the energy eigenvalues of the harmonic oscillator is

$$E_n = \left(n + \frac{1}{2}\right) \hbar\omega. \quad (3.48)$$

The energy level of the harmonic oscillator is thus quantized, with energy quanta  $\hbar\omega$ . The operator  $\hat{a}^\dagger$  adds an energy quanta to the oscillator, thus having the name *creation operator*; the operator  $\hat{a}$  removes an energy quanta from the oscillator, thus having the name *annihilation operator*.

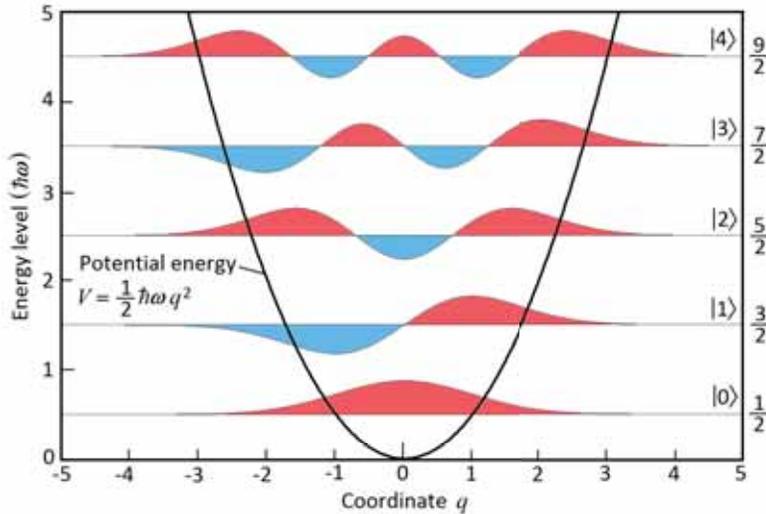
### 3.4.3 Explicit expressions of the wavefunctions

The above algebraic solution provides the simplest approach to find explicit expressions of the wavefunctions. First, from Eqs. 3.39 and 3.32, the zeroth-order wavefunction  $|0\rangle = \psi_0(q)$  should satisfy

$$\left(q + \frac{d}{dq}\right) \psi_0(q) = 0. \quad (3.49)$$

The solution is

$$\psi_0(q) = C_0 \exp\left(-\frac{q^2}{2}\right). \quad (3.50)$$



**Fig. 3.4. Energy levels and wavefunctions of a harmonic oscillator.** The solid curve is the potential energy. The wavefunctions of the first five eigenstates are shown. The red shade indicates positive phase, and blue shade indicates negative phase. The  $y$ -position of the energy eigenvalue is the baseline for the wavefunction.

The normalization constant  $C_0$  can be determined directly by

$$\int_{-\infty}^{\infty} \psi_0^2(q) dq = C_0^2 \int_{-\infty}^{\infty} \exp(-q^2) dq = C_0^2 \sqrt{\pi} = 1, \quad (3.51)$$

which gives

$$C_0 = \pi^{-1/4}. \quad (3.52)$$

All wavefunctions of the harmonic oscillator  $|n\rangle = \psi_n(q)$  can be obtained from Eqs. 3.33 and 3.47 using Eqs. 3.50 and 3.52:

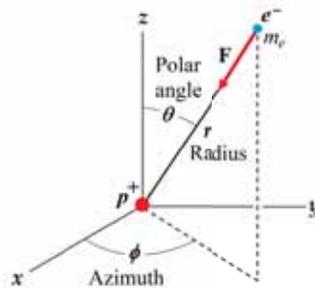
$$\psi_n(q) = \frac{1}{\sqrt{2^n n! \sqrt{\pi}}} \left( q - \frac{d}{dq} \right)^n \exp\left(-\frac{q^2}{2}\right). \quad (3.53)$$

The first few wavefunctions are listed in Table 3.1, and graphically displayed in Fig. 3.4. The solid curve, a parabola, represents the potential energy as a function of the normalized coordinate  $q$ . The horizontal lines represent the energy levels of the eigenstates. For the region inside the potential curve, energy level  $E$  is greater than the potential energy. The wavefunction resembles a sinusoidal wave. Outside the potential curve, the energy level is lower than the potential curve. The wavefunction resembles an exponential function decaying into the barrier. The lowest eigenstate has an energy value of  $\frac{1}{2}\hbar\omega$ . The wavefunction has no node. The energy eigenvalue increases by  $\hbar\omega$  each step, while a new node is added. Those nodal structures make the wavefunctions orthogonal to each other.

### 3.5 The Hydrogen Atom

Hydrogen atom is a central subject of quantum mechanics. The accurate interpretation of the Rydberg formula marked a groundbreaking triumph. Many predictions of relativistic quantum electrodynamics are verified by measurements on the hydrogen atom. The hydrogen wavefunctions are the foundation for the understanding of complex atoms and atomic systems.

**Fig. 3.5. Hydrogen atom in spherical polar coordinates.** The center of the coordinate system is the positively charged proton. The force and potential energy  $V$  only depends on radius  $r$ . Therefore, it is natural to use spherical polar coordinates, with radius  $r$ , polar angle  $\theta$ , and azimuth  $\phi$ .



There are only two real-world systems that the Schrödinger equation has analytic solutions: the hydrogen atom, and the hydrogen molecular ion,  $\text{H}_2^+$ , which is the basis to understand the chemical bond, and how condensed matter is formed. The ground-state wavefunction of hydrogen atom is also the starting point of the perturbation treatment of chemical bonds.

In SI unit system, the potential energy function  $V(r)$  is the Coulomb attractive force from the positively charged proton, see Eq. 1.123,

$$V(\mathbf{r}) = -\frac{e^2}{4\pi\epsilon_0 r}. \quad (3.54)$$

Because the electron mass  $m_e$  is much smaller than the proton mass, to a good approximation, the Schrödinger equation is

$$-\frac{\hbar^2}{2m_e} \left[ \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} \right] \psi - \frac{e^2}{4\pi\epsilon_0 r} \psi = E\psi. \quad (3.55)$$

The potential only depends on  $r$ . It is natural to solve the equation in polar coordinates: radius  $r$ , polar angle  $\theta$ , and azimuth  $\varphi$ , defined as

$$\begin{aligned} x &= r \sin \theta \cos \varphi, \\ y &= r \sin \theta \sin \varphi, \\ z &= r \cos \theta. \end{aligned} \quad (3.56)$$

See Fig. 3.5. Equation 3.55 becomes

$$-\frac{\hbar^2}{2m_e} \left[ \frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial \psi}{\partial r} \right) - \frac{1}{r^2} \mathbf{L}^2 \psi \right] - \frac{e^2}{4\pi\epsilon_0 r} \psi = E\psi. \quad (3.57)$$

where the *angular momentum operator*  $\mathbf{L}^2$  is defined as

$$\mathbf{L}^2 \psi \equiv -\frac{1}{\sin^2 \theta} \left[ \frac{\partial^2 \psi}{\partial \varphi^2} + \sin \theta \frac{\partial}{\partial \theta} \left( \sin \theta \frac{\partial \psi}{\partial \theta} \right) \right]. \quad (3.58)$$

The solutions of Eqs. 3.57 and 3.58 are presented in the following sections.

### 3.5.1 The Ground State

First, we study the ground state, where the wavefunction  $\psi$  only depends on  $r$ . The Schrödinger equation Eq. 3.57 becomes

$$-\frac{\hbar^2}{2m_e} \frac{1}{r^2} \frac{d}{dr} \left( r^2 \frac{d\psi}{dr} \right) - \frac{e^2}{4\pi\epsilon_0 r} \psi = E\psi, \quad (3.59)$$

From an intuitive point of view, since the electron is being attracted by the positively charged proton, the wavefunction should concentrate near

the proton, and decays with distance  $r$ . Therefore, to resolve Schrödinger's equation Eq. 3.59, we use the following trial function

$$\psi = Ce^{-r/a}, \quad (3.60)$$

where  $a$  is a parameter to be determined, and  $C$  is a normalization constant. Insert Eq. 3.60 into Eq. 3.58, eliminate the common factor  $\psi$ , we obtain

$$\frac{\hbar^2}{m_e a r} - \frac{\hbar^2}{2m_e a^2} - \frac{e^2}{4\pi\epsilon_0 r} = E. \quad (3.61)$$

To have a valid solution for all values of  $r$ , the two terms with common factor  $1/r$  must cancel each other. It implies

$$a = \frac{4\pi\epsilon_0\hbar^2}{m_e e^2}. \quad (3.62)$$

Therefore, the trial function Eq. 3.60 is a valid solution of Eq. 3.59, and the decay length is  $a$ , given by Eq. 3.62. The rest of Eq. 3.61 provides an expression of the energy eigenvalue

$$E = -\frac{\hbar^2}{2m_e a^2} = -\frac{m_e e^4}{32\epsilon_0^2 \pi^2 \hbar^2}. \quad (3.63)$$

The parameter  $a$  in Eq. 3.62 equals the *Bohr radius*,

$$a_0 \equiv \frac{4\pi\epsilon_0\hbar^2}{m_e e^2} \approx 52.9 \text{ pm}. \quad (3.64)$$

The absolute value of the energy eigenvalue in Eq. 3.63 is the *Rydberg constant*, which agrees well with experimental findings,

$$\text{Ry} \equiv \frac{m_e e^4}{32\epsilon_0^2 \pi^2 \hbar^2} \approx 13.6 \text{ eV}. \quad (3.65)$$

In order to determine the normalization constant  $C$ , we note that according to Schrödinger, the square of the wavefunction is correlated to the charge density of that single electron in space,

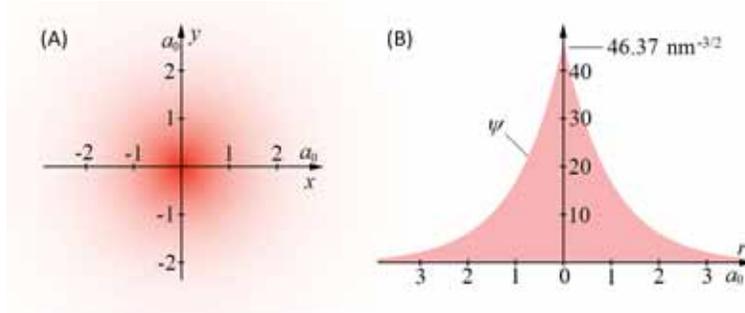
$$\rho = e|\psi|^2, \quad (3.66)$$

requiring that the integration of  $|\psi|^2$  over the entire space is 1,

$$\int_0^\infty 4\pi|\psi|^2 r^2 dr = \int_0^\infty 4\pi C^2 e^{-2r/a_0} r^2 dr = 1, \quad (3.67)$$

which yields

$$C = \frac{1}{\sqrt{\pi a_0^3}}. \quad (3.68)$$



**Fig. 3.6. Wavefunction of ground-state hydrogen atom.** (A) amplitude plot of the wavefunction. The decay constant is  $a_0$ , Eq. 3.64. (B) the amplitude profile. Wavefunction is the form of existence of electrons. The wavefunction of the ground state represents an electron at rest. There is no internal structure, no motion at all.

The ground-state wavefunction of hydrogen atom is then

$$|1s\rangle \equiv \psi_{1s} = \frac{1}{\sqrt{\pi a_0^3}} e^{-r/a_0}. \quad (3.69)$$

The subscript  $1s$  indicates that it is the lowest state and is spherically symmetric. In SI units,  $a_0 = 0.0529 \text{ nm}$ , and the constant  $C$  is  $46.37 \text{ nm}^{-3/2}$ . A density plot and an amplitude contour of the ground-state wavefunction of hydrogen atom are shown in Fig. 3.6. Wavefunction is the form of existence of the electron. The ground state represents an electron at rest. There is no internal structure, and no motion at all. The electrical charge of that electron at rest is smeared around the proton. The charge density of the electron is higher in the immediate vicinity of the proton.

### 3.5.2 Spherical harmonics

To find the solutions for the excited states, the standard method of resolving partial differential equations, separation of variables, similar to that for the timpani in Section 1.2.2 and the square potential well in Section ??, is used. The wavefunction is written as a product of a function of radius  $R(r)$  and a function of angle variables,  $Y(\theta, \varphi)$ , called *spherical harmonics*:

$$\psi(r, \theta, \varphi) = R(r)Y(\theta, \varphi). \quad (3.70)$$

Insert into Eq. 3.57, we find

$$-\frac{\hbar^2}{2m_e} \frac{1}{R(r)} \frac{d}{dr} \left( r^2 \frac{dR(r)}{dr} \right) - \frac{e^2 r}{4\pi\epsilon_0} - Er^2 = -\frac{\hbar^2}{2m_e} \frac{1}{Y(\theta, \varphi)} \mathbf{L}^2 Y(\theta, \varphi). \quad (3.71)$$

The left-hand side of the equation depends only on  $r$ , and the the right-hand side of the equation of the spherical harmonics depends only on angular variables  $\theta$  and  $\varphi$ . Therefore, both sides must be a constant.

In the following, using the differential equation for the function  $Y(\theta, \varphi)$ , Eq. 3.58, we look for the explicit value of constant  $\lambda$ :

$$-\frac{1}{\sin^2 \theta} \frac{\partial^2 Y(\theta, \varphi)}{\partial \varphi^2} - \frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \left( \sin \theta \frac{\partial Y(\theta, \varphi)}{\partial \theta} \right) = \lambda Y(\theta, \varphi). \quad (3.72)$$

The angular variables can be separated by writing  $Y(\theta, \varphi)$  as a product,

$$Y(\theta, \varphi) = \Theta(\theta)\Phi(\varphi). \quad (3.73)$$

Insert Eq. 3.73 into Eq. 3.72, we obtain

$$\frac{\sin^2 \theta}{\Theta} \left[ \frac{1}{\sin \theta} \frac{d}{d\theta} \left( \sin \theta \frac{d\Theta}{d\theta} \right) + \lambda \Theta \right] = -\frac{1}{\Phi} \frac{d^2 \Phi}{d\varphi^2}. \quad (3.74)$$

The left-hand side only depends on  $\theta$ , and the right-hand side only depends on  $\varphi$ . Both sides must be a constant. The function  $\Phi(\varphi)$  should be a sinusoidal function of an integer multiple of  $\varphi$ , either

$$\Phi(\varphi) = \cos m\varphi \quad (3.75)$$

or

$$\Phi(\varphi) = \sin m\varphi, \quad (3.76)$$

to satisfy the ordinary differential equation

$$\frac{d^2 \Phi}{d\varphi^2} + m^2 \Phi = 0. \quad (3.77)$$

The non-negative integer  $m$  is called *azimuthal quantum number*.

The right-hand side of Eq. 3.74 is then

$$\frac{1}{\sin \theta} \frac{d}{d\theta} \left( \sin \theta \frac{d\Theta}{d\theta} \right) - \frac{m^2}{\sin^2 \theta} \Theta + \lambda \Theta = 0, \quad (3.78)$$

To solve Eq. 3.78, we use a trial function

$$\Theta(\theta) = \sin^m \theta. \quad (3.79)$$

Insert Eq. 3.79 into Eq. 3.78, after a short algebra, we obtain

$$-m(m+1)\Theta + \lambda\Theta = 0. \quad (3.80)$$

Therefore, it is a good solution with

$$\lambda = m(m+1). \quad (3.81)$$

Because  $\lambda$  is a parameter for  $\Theta$ , a function of polar angle  $\theta$ , we define a *polar quantum number*  $l$  by

$$\lambda = l(l+1). \quad (3.82)$$

For the case of  $l = m$ , the expression of  $\Theta(\theta)$  in Eq. 3.79 is correct. In general,  $m$  could be smaller than  $l$ . The expressions for a general case is complicated. However, we need only a few more cases, and the expressions are fairly simple. Here is a complete list:

$$\Theta_{10}(\theta) = \cos \theta, \quad l = 1, \quad m = 0; \quad (3.83)$$

$$\Theta_{20}(\theta) = 3 \cos^2 \theta - 1, \quad l = 2, \quad m = 0; \quad (3.84)$$

and

$$\Theta_{21}(\theta) = \cos \theta \sin \theta, \quad l = 2, \quad m = 1. \quad (3.85)$$

The correctness of those solutions can be verified by Eq. 3.78,

$$\frac{1}{\sin \theta} \frac{d}{d\theta} \left( \sin \theta \frac{d\Theta_{lm}(\theta)}{d\theta} \right) - \frac{m^2}{\sin^2 \theta} \Theta_{lm}(\theta) + l(l+1)\Theta_{lm}(\theta) = 0, \quad (3.86)$$

which is left as an exercise.

**Table 3.2: Spherical Harmonics**

| Mathematical notation       | Chemist's Name | Formula in angular variables   | In Cartesian coordinates   |
|-----------------------------|----------------|--|--|
| $Y_{00}(\theta, \varphi)$   | $s$            | $\frac{1}{\sqrt{4\pi}}$  | $\frac{1}{\sqrt{4\pi}}$  |
| $Y_{10}(\theta, \varphi)$   | $p_z$          | $\sqrt{\frac{3}{4\pi}} \cos \theta$  | $\sqrt{\frac{3}{4\pi}} \frac{z}{r}$                                    |
| $Y_{11}^g(\theta, \varphi)$ | $p_x$          | $\sqrt{\frac{3}{4\pi}} \sin \theta \cos \varphi$                               | $\sqrt{\frac{3}{4\pi}} \frac{x}{r}$                                    |
| $Y_{11}^u(\theta, \varphi)$ | $p_y$          | $\sqrt{\frac{3}{4\pi}} \sin \theta \sin \varphi$                               | $\sqrt{\frac{3}{4\pi}} \frac{y}{r}$                                    |
| $Y_{20}(\theta, \varphi)$   | $d_{z^2}$      | $\sqrt{\frac{5}{4\pi}} \left( \frac{3}{2} \cos^2 \theta - \frac{1}{2} \right)$ | $\sqrt{\frac{5}{4\pi}} \left( \frac{3z^2}{2r^2} - \frac{1}{2} \right)$ |
| $Y_{21}^g(\theta, \varphi)$ | $d_{xz}$       | $\sqrt{\frac{15}{4\pi}} \cos \theta \sin \theta \cos \varphi$                  | $\sqrt{\frac{15}{4\pi}} \frac{zx}{r^2}$                                |
| $Y_{21}^u(\theta, \varphi)$ | $d_{yz}$       | $\sqrt{\frac{15}{4\pi}} \cos \theta \sin \theta \sin \varphi$                  | $\sqrt{\frac{15}{4\pi}} \frac{zy}{r^2}$                                |
| $Y_{22}^g(\theta, \varphi)$ | $d_{x^2-y^2}$  | $\sqrt{\frac{15}{16\pi}} \sin^2 \theta \cos 2\varphi$                          | $\sqrt{\frac{15}{16\pi}} \frac{x^2 - y^2}{r^2}$                        |
| $Y_{22}^u(\theta, \varphi)$ | $d_{xy}$       | $\sqrt{\frac{15}{16\pi}} \sin^2 \theta \sin 2\varphi$                          | $\sqrt{\frac{15}{4\pi}} \frac{xy}{r^2}$                                |

For applications in chemistry and molecular biology, expressions of spherical harmonics in Cartesian coordinates are useful. And it looks simpler. Table 3.2 lists the first nine spherical harmonics in real variables, which represents a complete list of spherical harmonics often used in chemistry and molecular biology. In the Table, superscript  $g$  means symmetric versus  $x$ , and superscript  $u$  means antisymmetric versus  $x$ . In most books on mathematical physics and quantum mechanics, complex spherical harmonics are listed. The index  $l$  is sometimes called the *orbital quantum number*, and the index  $m$  is sometimes called *magnetic quantum number*, which implies a complex exponential function. In chemistry and molecular biology, the real spherical harmonics are commonly used. Chemists have special names for those angular dependences, as shown in the second column, Chemist's names. Owing to the origin of those angular dependences, referring to Figure 3.5, the label  $l$  is called the *polar quantum number*, and the label  $m$ , also non-negative, is called the *azimuthal quantum number*. The numerical constant such as  $1/\sqrt{4\pi}$  is to normalize the spherical harmonics,

$$\int_{-\pi/2}^{\pi/2} \cos \theta \, d\theta \int_0^{2\pi} d\phi |Y_{lm}(\theta, \phi)|^2 = 1. \quad (3.87)$$

### 3.5.3 Energy eigenvalues of excited States

By applying the factor  $l(l+1)$  in Eq. 3.86 to replace the right-hand side of Eq. 3.71, its left-hand side becomes

$$-\frac{\hbar^2}{2m_e} \left[ \frac{1}{r^2} \frac{d}{dr} \left( r^2 \frac{dR(r)}{dr} \right) - \frac{l(l+1)R(r)}{r^2} \right] + \frac{e^2}{4\pi\epsilon_0 r} R(r) = ER(r). \quad (3.88)$$

To resolve Eq. 3.88, by intuition, we use the following trial function

$$R(r) = C r^b e^{-r/a}, \quad (3.89)$$

where the parameters  $a$  and  $b$  are to be determined by Eq. 3.88, and  $C$  is a normalization constant. Insert Eq. 3.89 into Eq. 3.88, eliminate the common factor  $R(r)$ , we obtain an equation

$$-\frac{\hbar^2}{2m_e} \left[ \frac{b(b+1)}{r^2} - \frac{l(l+1)}{r^2} + \frac{2(b+1)}{ar} - \frac{1}{a^2} \right] + \frac{e^2}{4\pi\epsilon_0 r} = E. \quad (3.90)$$

To cancel the two terms with  $1/r^2$ , a sufficient condition is

$$b = l. \quad (3.91)$$

To cancel the two terms with  $1/r$ , one must have

$$a = (b+1) \frac{4\pi\epsilon_0 \hbar^2}{m_e e^2} = (b+1)a_0. \quad (3.92)$$

The remaining terms in Eq. 3.90 determine the energy eigenvalue

$$E = -\frac{\hbar^2}{2m_e a^2} = -\frac{\text{Ry}}{(b+1)^2}. \quad (3.93)$$

In all those formulas,  $(b+1)$  occurs. Define an integer label  $n$ ,

$$n = b + 1, \quad (3.94)$$

which is the *principle quantum number*. The wavefunction is

$$\psi = C r^{n-1} e^{-r/na_0} Y_{lm}(\theta, \varphi). \quad (3.95)$$

The energy eigenvalue only depends on the principle quantum number  $n$ , perfectly explains the experimentally discovered Rydberg formula,

$$E_n = -\frac{\text{Ry}}{n^2}. \quad (3.96)$$

The length scale of the wavefunction,  $a = na_0$ , is a useful parameter for further study of the hydrogen wavefunction. It is

$$a = na_0 = \frac{\hbar}{\sqrt{-2m_e E}}. \quad (3.97)$$

Similar to Eq. 3.67, the normalization constant is found to be

$$C = \sqrt{\frac{2^{2n+1}}{\pi(2n+2)!(na_0)^{2n+3}}}. \quad (3.98)$$

In Fig. 3.7, we show some interesting cases. Following Eq. 3.79, for  $m = l$ , the spherical harmonics up to a normalization constant is

$$Y_l^g(\theta, \varphi) \propto \sin^l \theta \cos l\varphi, \quad (3.99)$$

and

$$Y_l^u(\theta, \varphi) \propto \sin^l \theta \sin l\varphi. \quad (3.100)$$

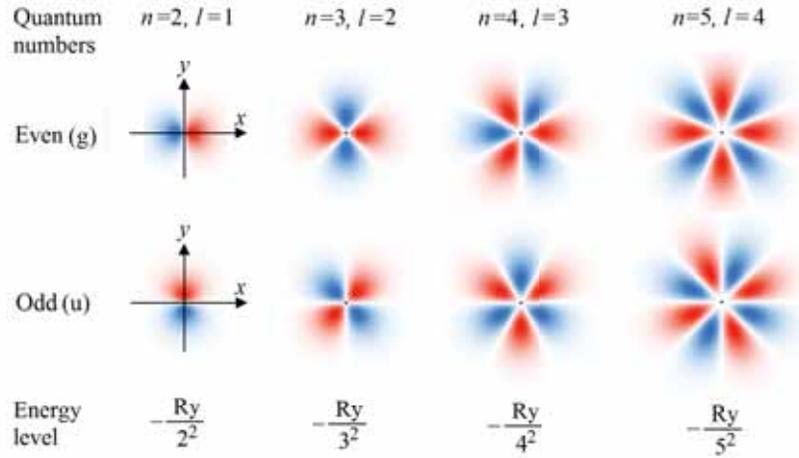
The wavefunctions shown in Fig. 3.7 are on the plane  $z = 0$ , where  $\sin \theta = 1$ . Using Eqs. 3.91 and 3.94,  $l = n - 1$ . The explicit expressions are

$$\psi_g \propto r^{n-1} e^{-r/na_0} \cos(n-1)\varphi, \quad (3.101)$$

and

$$\psi_u \propto r^{n-1} e^{-r/na_0} \sin(n-1)\varphi, \quad (3.102)$$

where the subscript  $g$  and  $u$  indicates even and odd with respect to  $x$ -axis. Those wavefunctions represent standing waves on the  $z = 0$  plane. As remarked by Schrödinger in his first 1926 paper, the energy-level quantization



**Fig. 3.7. Wavefunctions of excited-states of hydrogen atom.** Wavefunctions in Eqs. 3.101 and 3.102, standing waves on the horizontal plane  $z = 0$ . For each principle quantum number, there is an even wavefunction symmetric regarding the  $x$ -axis, and an wavefunction antisymmetric regarding the  $x$ -axis. The energy level, Eq. 3.96, depends only on the principle quantum number  $n$ , explains the Rydberg formula.

in hydrogen atom resembles the nodes of a vibrating string in Section 1.2.1. The electron, attracted by the proton, is trapped in the proton's electrical field. However, an electron is fundamentally a wave. It cannot collapse into the proton. It forms standing waves around the proton. Similar to a vibrating string, different patterns of standing waves with different number of nodes are formed, resulting in different energy levels.

In analogy to Newtonian mechanics, those eigenstates with fixed energy eigenvalues correspond to motions with a constant speed. According to Newton's first law of mechanics, a system will keep its state of motion until an external force acting on it. In quantum mechanics, without external interactions, an eigenstate is stationary. Similar to Newton's second law, external forces such as radiation fields can cause transitions between different eigenstates. Details will be presented in Chapters 10 and 11.

### 3.5.4 Wavefunctions

In the previous Subsection, following a simple mathematical procedure, special solutions of Eq. 3.88 are derived. Due to a high degree of degeneracy, all energy eigenvalues are obtained. Especially, for the cases of  $l = n - 1$  and  $m = l$ , wavefunctions are shown in Figs 3.6 and 3.7. Nevertheless, Eq. 3.88 does have other useful solutions.

Inspired by the special solutions, a dimensionless variable  $\rho$  is introduced

**Table 3.3: Wavefunctions of the hydrogen atom**

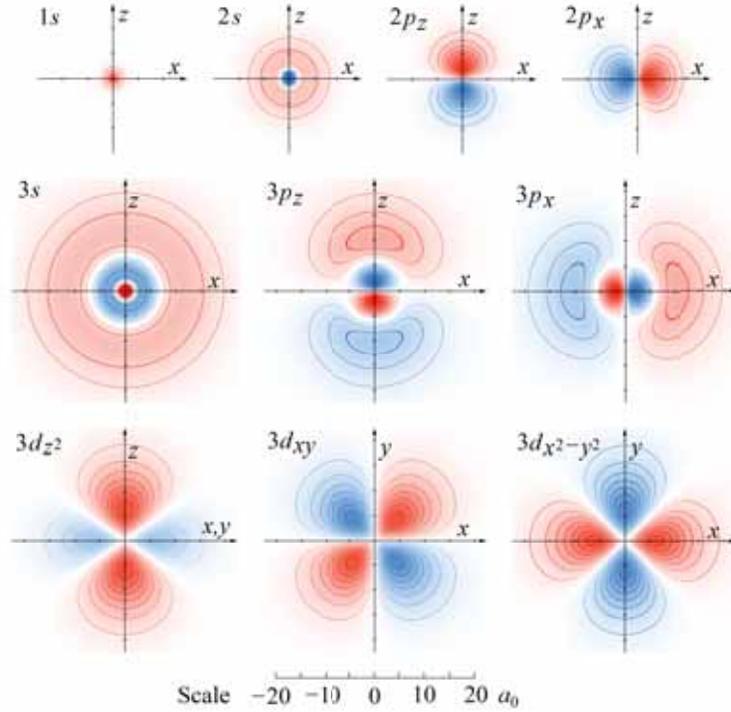
| Name                                      | State          | Wavefunction   | $\bar{r} (a_0)$ |
|---|----------------|--|-----------------|
| 1s  | $ 1s0\rangle$  | $\frac{1}{\sqrt{\pi}} e^{-r}$  | 1.5             |
| 2s  | $ 2s0\rangle$  | $\frac{1}{2\sqrt{2\pi}} (1 - \frac{1}{2}r) e^{-r/2}$                   | 6               |
| 2p <sub>x</sub>                           | $ 2p1u\rangle$ | $\frac{1}{4\sqrt{2\pi}} x e^{-r/2}$                                    | 5               |
| 2p <sub>y</sub>                           | $ 2p1g\rangle$ | $\frac{1}{4\sqrt{2\pi}} y e^{-r/2}$                                    | 5               |
| 2p <sub>z</sub>                           | $ 2p0\rangle$  | $\frac{1}{4\sqrt{2\pi}} z e^{-r/2}$                                    | 5               |
| 3s  | $ 3s0\rangle$  | $\frac{1}{3\sqrt{3\pi}} (1 - \frac{2}{3}r + \frac{2}{27}r^2) e^{-r/3}$ | 13.5            |
| 3p <sub>x</sub>                           | $ 3p1u\rangle$ | $\frac{2}{27\sqrt{6\pi}} (1 - \frac{1}{6}r) x e^{-r/3}$                | 12.5            |
| 3p <sub>y</sub>                           | $ 3p1g\rangle$ | $\frac{2}{27\sqrt{6\pi}} (1 - \frac{1}{6}r) y e^{-r/3}$                | 12.5            |
| 3p <sub>z</sub>                           | $ 3p0\rangle$  | $\frac{2}{27\sqrt{6\pi}} (1 - \frac{1}{6}r) z e^{-r/3}$                | 12.5            |
| 3d <sub>z<sup>2</sup></sub>               | $ 3d0\rangle$  | $\frac{2}{27\sqrt{6\pi}} (z^2 - \frac{1}{3}r^2) e^{-r/3}$              | 10.5            |
| 3d <sub>xz</sub>                          | $ 3d1u\rangle$ | $\frac{2}{27\sqrt{6\pi}} xz e^{-r/3}$                                  | 10.5            |
| 3d <sub>yz</sub>                          | $ 3d1g\rangle$ | $\frac{2}{27\sqrt{6\pi}} yz e^{-r/3}$                                  | 10.5            |
| 3d <sub>xy</sub>                          | $ 3d2u\rangle$ | $\frac{2}{27\sqrt{6\pi}} xy e^{-r/3}$                                  | 10.5            |
| 3d <sub>x<sup>2</sup>-y<sup>2</sup></sub> | $ 3d2g\rangle$ | $\frac{4}{27\sqrt{6\pi}} (x^2 - y^2) e^{-r/3}$                         | 10.5            |

by scaling the radius  $r$  using the length  $a = na_0$  in Eq. 3.97,

$$\rho = \frac{r}{a}. \quad (3.103)$$

Equation 3.88 becomes

$$\frac{1}{\rho^2} \frac{d}{d\rho} \left( \rho^2 \frac{dR(\rho)}{d\rho} \right) + \left[ -1 - \frac{l(l+1)}{\rho^2} + \frac{2n}{\rho} \right] R(\rho) = 0. \quad (3.104)$$



**Fig. 3.8. Hydrogen wavefunctions.** The phase, either positive or negative, is shown in different color. The value is indicated by intensity, but scaled slightly to facilitate visualization. As shown in the last column of Table 3.3, the size of wavefunctions with different principle quantum number  $n$  varies dramatically. All the figures showing here are accurately sized according to the Scale at the bottom of the Figure.

In the process of the algebra, Eq. 3.64 is applied to obtain

$$\frac{m_e a e^2}{2\pi\epsilon_0 \hbar^2} = \frac{a}{a_0} = n. \quad (3.105)$$

Analogous to Eqs. 3.89 and 3.91, we make a substitution

$$R(\rho) = \rho^l e^{-\rho} F(\rho). \quad (3.106)$$

Insert Eq. 3.106 to Eq. 3.104, the differential equation for  $F(\rho)$  is obtained:

$$\rho \frac{d^2 F(\rho)}{d\rho^2} + 2(l+1-\rho) \frac{dF(\rho)}{d\rho} + 2(n-l-1)F(\rho) = 0. \quad (3.107)$$

It is the differential equation for associate Laguerre polynomials, well known for mathematicians for two centuries. The general formula is fairly cumbersome. In condensed-matter physics, chemistry and molecular biology, only

a few more cases are needed. Here is a complete list.

$$F(r) = 1 - \frac{1}{2} \frac{r}{a_0}, \quad n = 2, \quad l = 0; \quad (3.108)$$

$$F(r) = 1 - \frac{2}{3} \frac{r}{a_0} + \frac{2}{27} \left( \frac{r}{a_0} \right)^2, \quad n = 3, \quad l = 0; \quad (3.109)$$

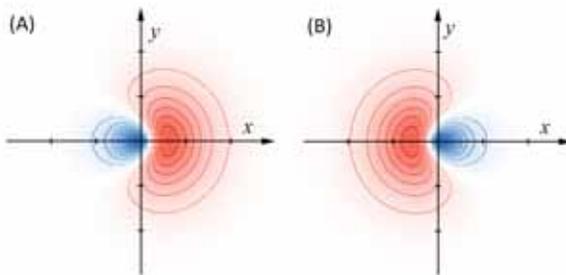
$$F(r) = 1 - \frac{1}{6} \frac{r}{a_0}, \quad n = 3, \quad l = 1. \quad (3.110)$$

Those expressions can be verified by directly inserting those polynomials  $F(r)$  into Eq. 3.107. It is left as an exercise.

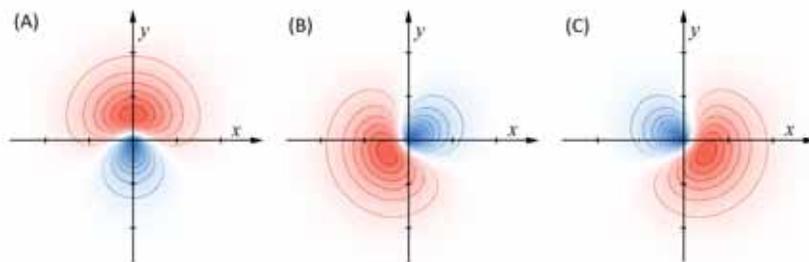
By combining the radial functions in Eq. 3.106, using Eqs. 3.108 through 3.110 and the spherical harmonics in Table 3.2, we found the first nine wavefunctions of the hydrogen atom, listed in Table 3.3, and in Fig. 3.8. In Table 3.2, the first column shows the chemist's name of the wavefunction. The second column is in Dirac notation. The expression of the wavefunction, column 3, is in Cartesian coordinates with Bohr radius  $a_0$  as unit. The last column is the average size of the wavefunction, also in unit of  $a_0$ .

### 3.5.5 Degeneracy and Wavefunction Hybridization

The energy eigenvalues of the hydrogen atom only depends on the principle quantum number  $n$ . For each principle quantum number, there are  $n^2$  different wavefunctions: it is  $n^2$ -fold degenerate. For  $n = 2$ , there are four states,  $|2s\rangle$ ,  $|2px\rangle$ ,  $|2py\rangle$ , and  $|2pz\rangle$ . Because the Schrödinger equation is linear, any linear superposition of wavefunctions with the same energy eigenvalue is also a good wavefunction with the same energy eigenvalue. Especially, an  $s$ -wavefunction can make linear superposition with  $p$ -wavefunctions, to form *hybrid* wavefunctions. This concept is fundamental in chemistry, such



**Fig. 3.9. Hybrid  $sp_1$  wavefunctions.** The phase of the wavefunction is shown by color, positive in red, negative in blue. The equal-value contours are also shown. The main lobe spans a wide angle. It is difficult to show both on the same figure.



**Fig. 3.10. Hybrid  $sp_2$  wavefunctions.** The three  $sp_2$  hybrid wavefunctions are pointing to the three vertices of a regular triangle,  $120^\circ$  apart in the same plane. The phase, positive or negative, is shown by color. The equal-value contours are also shown.

as for carbon and silicon, especially in organic chemistry. Here we show the concept of hybridization using the hydrogen wavefunctions.

Figure 3.9 shows two  $sp_1$  hybrid wavefunctions. (A) is with a positive  $2px$  wavefunction,

$$|2sp_1+\rangle = \frac{1}{\sqrt{2}}|2s\rangle + \frac{1}{\sqrt{2}}|2px\rangle, \quad (3.111)$$

resulting in a wavefunction preferentially concentrated in the  $+x$  direction; and (B) is with a negative  $2px$  wavefunction,

$$|2sp_1-\rangle = \frac{1}{\sqrt{2}}|2s\rangle - \frac{1}{\sqrt{2}}|2px\rangle, \quad (3.112)$$

resulting in a wavefunction preferentially concentrated in the  $-x$  direction. This happens for example in acetylene  $C_2H_2$ , for both the  $\sigma$ -bond between the two carbon atoms, and the two C-H bonds.

Figure 3.10 shows three  $sp_2$  hybrid wavefunctions. The formulas are

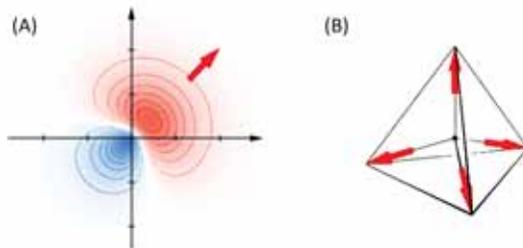
$$|2sp_2u\rangle = \frac{1}{\sqrt{3}}|2s\rangle + \sqrt{\frac{2}{3}}|2py\rangle, \quad (3.113)$$

$$|2sp_2l\rangle = \frac{1}{\sqrt{3}}|2s\rangle - \frac{1}{\sqrt{2}}|2px\rangle - \frac{1}{\sqrt{6}}|2py\rangle, \quad (3.114)$$

and

$$|2sp_2r\rangle = \frac{1}{\sqrt{3}}|2s\rangle + \frac{1}{\sqrt{2}}|2px\rangle - \frac{1}{\sqrt{6}}|2py\rangle, \quad (3.115)$$

respectively. It is the basic structure of grapheme and carbon nanotubes. In ethylene,  $C_2H_4$ , the  $sp_1$  hybrid wavefunctions are the basis of both the  $\sigma$ -bond between the two carbon atoms, and the four C-H bonds.



**Fig. 3.11. Hybrid  $sp^3$  wavefunctions.** (A), one of the four hybrid wavefunctions. The other three have the same shape but different orientation, pointing to the four vertices of a regular tetrahedron. (b), a regular tetrahedron.

A prevailing hybridization is the  $sp^3$  mode, where one  $s$ -wavefunction is linearly superposed with three  $p$ -wavefunctions to form four hybrid wavefunctions pointing to the four vertices of a tetrahedron, see Fig. 3.11. A set of sample formula is

$$\begin{aligned}
 |t_{111}\rangle &= \frac{1}{2} (|2s\rangle + |2px\rangle + |2py\rangle + |2pz\rangle), \\
 |t_{\bar{1}\bar{1}\bar{1}}\rangle &= \frac{1}{2} (|2s\rangle - |2px\rangle + |2py\rangle - |2pz\rangle), \\
 |t_{\bar{1}\bar{1}1}\rangle &= \frac{1}{2} (|2s\rangle - |2px\rangle - |2py\rangle + |2pz\rangle), \\
 |t_{1\bar{1}\bar{1}}\rangle &= \frac{1}{2} (|2s\rangle + |2px\rangle - |2py\rangle - |2pz\rangle).
 \end{aligned}
 \tag{3.116}$$

It is the backbone of all alkanes, including methane, ethane, propane, butane, etc., and the crystalline structures of diamond and silicon.

In some quantum chemistry textbooks, those hybrid wavefunctions are depicted as sharply oriented long and narrow bulbs, to emphasize the directional effect. Often, three or four such long and narrow bulbs are shown in a single drawing. As shown here, the accurate amplitude contours of wavefunctions of  $sp$ -hybridizations span a broad angle. It is difficult to depict three or four such hybrid wavefunctions on a single graph.

By comparing Fig. 3.4 with the vibrational modes of a string, Fig. 1.8, a striking resemblance is found. As an objective reality, wavefunction is of the same nature as the standing waves in classical physics.