AP Courses Review Notes
AP3204
Waves and Optics

WU PEILIN

November 17, 2017

Contents

Chapter 1

Vibration

§1 Free Vibration

1.1 Simple Harmonic Motion

A periodic motion is a motion which repeats itself after regular intervals of time, and the simplest kind of periodic motion is a *simple harmonic motion* in which the displacement varies sinusoidally with time.

To understand simple harmonic motion, we consider a point P rotating on the circumference of a circle of radius a with an angular velocity ω. We can obtain the projection of Point P on x and y axis and taking the initial phases into account:

$$\begin{aligned} x &= a\cos(\omega t + \theta) \\ y &= a\sin(\omega t + \theta), \end{aligned}$$

and the velocity:

$$\begin{aligned} v_x &= \dot{x} = -a\omega\sin(\omega t + \theta) = -\omega x \\ v_y &= \dot{y} = a\omega\cos(\omega t + \theta) = \omega y, \end{aligned}$$

and the acceleration:

$$\begin{aligned} a_x &= \ddot{x} = -a\omega^2\cos(\omega t + \theta) = -\omega^2 x \\ a_y &= \ddot{y} = -a\omega^2\sin(\omega t + \theta) = -\omega^2 y \end{aligned}$$

Here is a crucial observation about the acceleration, we have a second order differential equation:

$$\ddot{\mathbf{r}} = -\omega^2\mathbf{r} \tag{1.1}$$

Let's consider an alternative situation: A mass m attached to a spring with stiffness k. With initial position deviate from the equilibrium position, mass m

will exhibits periodic motion with the net exerted by spring as a *linear restoring force*

$$\mathbf{F} = -k\mathbf{x},$$

and with the *Newton's second law* of motion

$$\mathbf{F} = m\mathbf{a}.$$

Therefore,

$$m\mathbf{a} = \mathbf{F} = -k\mathbf{x}.$$

Thus we have the following equation similar to the equation (1.1)

$$\ddot{\mathbf{x}} + \left(\sqrt{\frac{k}{m}}\right)^2 \mathbf{x} = 0 \tag{1.2}$$

We can see that $\omega = \sqrt{\frac{k}{m}}$. Actually, the original circle is the phase diagram of simple harmonic motion. Moreover, it is important to state that: An object exhibits simple harmonic motion if the net external force acting on it is a linear restoring force.

Following the deduction, we obtain the general solution to simple harmonic motion:

$$x = A\sin\left(\sqrt{\frac{k}{m}}t + \phi\right),$$

where A is the maximum amplitude, k is the stiffness of spring, m is the mass and ϕ

Let's consider the energy of this simple mechanics system.

$$E_{total} = \underbrace{E_0 \sin^2 \omega t}_{\text{Potential Energy}} + \underbrace{E_0 \cos^2 \omega t}_{\text{Kinetic Energy}}$$

According to the *principle of superposition*, any oscillation can be modelled as a composition of multiple or infinite simple harmonic motion components.

1.2 Fourier Theorem

Fourier Theorem[1] states that any periodic function can be expressed as a sum of the sine and cosine functions whose frequencies increase in the ratio of natural numbers.

[1] You can think of that what Fourier Theorem actually states is to decompose original infinite vector space into the linear combination of two sets of orthogonal infinite vector space thus correspondingly, the element naturally following the decomposition. And those two orthogonal vector spaces rely on the orthogonality of triangular functions $f = C\sin(nx)$ and $g = C'\cos(mx)$.

Any periodic function $f(t)$ with period T ($f(t+nT) = f(t)$ with $n \in \mathbb{N}$) can be expanded in the form wit1h $\omega = \frac{2\pi}{T}$:

$$f(t) = \frac{1}{2}a_0 + \sum_{n=1}^{\infty} a_n \cos(n\omega t) + \sum_{n=1}^{\infty} b_n \sin(n\omega t), \tag{1.3}$$

with coefficient series determined by [2]

$$a_n = \frac{2}{T}\int_{t_0}^{t_0+T} f(t)\cos n\omega t\, dt$$
$$b_n = \frac{2}{T}\int_{t_0}^{t_0+T} f(t)\sin n\omega t\, dt$$

1.3 Other Cases of Simple Harmonic Motion

Simple Pendulum

For simple pendulum, the linear restoring force $\mathbf{F} = -mg\boldsymbol{\theta}$, and the acceleration $\mathbf{a} = L\ddot{\boldsymbol{\theta}}$, thus:

$$\mathbf{F} = -mg\boldsymbol{\theta} = mL\ddot{\boldsymbol{\theta}} \quad \Rightarrow \quad \ddot{\boldsymbol{\theta}} + \left(\sqrt{\frac{g}{L}}\right)^2 \boldsymbol{\theta} = 0,$$

which gives:

$$\ddot{\theta} + \left(\sqrt{\frac{g\sin\theta}{L}}\right)^2 \theta = 0.$$

However, this is not simple harmonic motion unless we make approximation that within small range that $\sin\theta = \theta$,

$$\ddot{\theta} + \left(\sqrt{\frac{g}{L}}\right)^2 \theta = 0$$

$$\theta = \Theta\sin(\sqrt{\frac{g}{L}}t + \phi).$$

[2]The coefficients a_n and b_n can be easily determined by using the following properties of the trigonometric functions:

$$\int_{t_0}^{t_0+T} \cos n\omega t \cos \omega m\omega t = \begin{cases} 0 & \text{if } m \neq n \\ \frac{T}{2} & \text{if } m = n \neq 0 \\ T & \text{if } m = n = 0 \end{cases}$$

$$\int_{t_0}^{t_0+T} \sin n\omega t \sin \omega m\omega t = \begin{cases} 0 & \text{if } m \neq n \\ \frac{T}{2} & \text{if } m = n \neq 0 \\ T & \text{if } m = n = 0 \end{cases}$$

$$\int_{t_0}^{t_0+T} \sin n\omega t \cos \omega m\omega t = 0$$

Compound Pendulum

Quite similar to simple pendulum,

$$\ddot{\theta} + \frac{mg}{J}\frac{L}{2}\sin\theta = 0$$

with same approximation $\sin\theta = \theta$, we have $\omega = \sqrt{\frac{mgL}{2J}}$

Torsional Pendulum

The linear restoring force $\mathbf{F} = -c\boldsymbol{\theta}$, followed by $-c\theta = J\ddot{\theta}$,

$$\ddot{\theta} + \frac{c}{J}\theta = 0$$

therefore, $\omega = \sqrt{\frac{c}{J}}$, $\theta = \Theta\sin\left(\sqrt{\frac{c}{J}}t + \phi\right)$

§2 Damped Vibration

For damped vibration, one big difference is that there is an additional first order force $\mathbf{F}_r$, $\mathbf{F} = \mathbf{F}_0 + \mathbf{F}_r = -kx - r\dot{\mathbf{x}}$. Its physical interpretation gives the first order resistance or friction contributing to the continuous energy dissipation and decay of amplitude. Again by Newton's second law, $\mathbf{F} = m\ddot{\mathbf{x}}$,

$$\ddot{\mathbf{x}} + \underbrace{\frac{r}{m}}_{\gamma}\dot{\mathbf{x}} + \underbrace{\frac{k}{m}}_{\omega_0}\mathbf{x} = 0 \tag{1.4}$$

Denote $\beta = \frac{\gamma}{2}$, $\alpha_{1,2} = -\beta \pm \sqrt{\beta^2 - \omega_0^2}$, we have the general solution: $x = Ce^{\alpha_1 t} + C'e^{\alpha_2 t}$.

2.1 Light Damping

If $\beta < \omega_0$ ($\alpha_{1,2} \in \mathbb{C}$), which means that the retarding force is small compared with the restoring force, we called it *light damping*, thus gives an exponential decaying sinusoidal wave $\sin(\omega_f t + \phi)$ enveloped by decaying amplitude $A_0e^{-\beta t}$.

$$x = \underbrace{A_0e^{-\beta t}}_{\text{Amplitude}}\sin(\underbrace{\sqrt{\omega_0^2 - \beta^2}}_{\omega_f}t + \phi)$$

2.2 Critical Damping

If $\beta = \omega_0$ ($\alpha_{1,2} = -\beta$) meaning the retarding force reaches the restoring force in the spring, we called it *critical damping*.

$$x = A_0e^{-\beta t}$$

Noted: Critical damping has the quickest return to equilibrium.

2.3 Heavy Damping

If $\beta > \omega_0$ ($\alpha_{1,2} \in \mathbb{R}$), in this case, the medium is so viscous that the retarding force is greater then the restoring force.

$$x = Ce^{\alpha_{1,2}t}$$

Noted: There is no oscillation but only overdamped curve. And surely it is the most boring one among those three, it just simply return to its original position.

2.4 Quantify Damping

Since three cases have differently behaviours, we want to quantify the quality of damping, etc. Here is some factors to describe decays due to damping

1. Width: $\gamma = \frac{r}{m}$
2. Logarithmic decrement: $\delta = \ln \frac{A_1}{A_2}$
3. Quality factor: $Q = \frac{\omega_0}{\gamma}$ (For a lightly damping system, $Q < 0.5$)

§3 Forced Vibration

Despite the linear restoring force and retarding force, there is also a periodic force angular frequency ω $\mathbf{F}_p = \mathbf{F}_0 \sin \omega t$, following the same deduction,

$$\ddot{\mathbf{x}} + \underbrace{\frac{r}{m}}_{\gamma=2\beta} \dot{\mathbf{x}} + \underbrace{\frac{k}{m}}_{\omega_0} \mathbf{x} = \underbrace{\frac{\mathbf{F}_0}{m}}_{P} \sin \omega t \tag{1.5}$$

We immediately notice that equation (1.4) is the homogeneous equation for equation (1.5), therefore, general solution of equation (1.4) is the contemporary solution of equation (1.5). Here we want to find the particular solution to equation (1.5), $x_p = B \sin(\omega t - \delta)$, where

$$\begin{aligned} \delta &= \arctan\left(\frac{2\beta\omega}{\omega_0^2 - \omega^2}\right) \\ B &= \frac{P}{\sqrt{\left(\omega_0^2 - \omega^2\right)^2 + 4\beta^2\omega^2}} \\ P &= F_0/m \end{aligned}$$

Therefore, the general solution for force damping system:

$$x = \overbrace{\underbrace{A_0 e^{-\beta t}}_{\text{Amplitude}} \sin(\underbrace{\sqrt{\omega_0^2 - \beta^2}}_{\omega_f}\, t + \phi)}^{\text{Damping term}} + \overbrace{\underbrace{\frac{F_0/m}{\sqrt{\left(\omega_0^2 - \omega^2\right)^2 + 4\beta^2\omega^2}}}_{B} \sin(\omega t - \underbrace{\arctan\left(\frac{2\beta\omega}{\omega_0^2 - \omega^2}\right)}_{\delta})}^{\text{Particular Solution}}$$

3.1 Steady State

After damping term goes to zero, we have steady state solution, or physically interpreted as the response to the periodic force.

$$F = F_0 \sin \omega t \quad \rightarrow \quad x = B \sin(\omega t - \delta)$$

3.2 Resonance and Q factor

For lightly damped motion, amplitude B reaches its maximum at resonance. Therefore, we obtain $\omega_r = \sqrt{\omega_0^2 - 2\beta^2}$ is the resonance frequency[3] ($\omega_f = \sqrt{\omega_0^2 - \beta^2}$), and the amplitude for resonance is

$$B_r = \frac{F/m}{2\beta\sqrt{\omega_0^2 - \beta^2}}$$

Furthermore, we can define another evaluation factor $Q = \frac{\omega_0}{\gamma} = \frac{\omega_0}{2\beta}$. The larger the value of Q, the less the dissipative effect and the greater the number of free oscillation for given decrease of amplitude. And the amplitude for resonance with $B_0 = \frac{F_0}{m\omega_0^2} = \frac{F_0}{k}$ is

$$B_r = \frac{F/m}{2\beta\sqrt{\omega_0^2 - \beta^2}} = \frac{\frac{F_0}{m\omega_0^2}}{\frac{2\beta}{\omega_0}\beta\sqrt{1 - \frac{\beta^2}{\omega_0^2}}} = \frac{B_0 Q}{\sqrt{1 - \frac{1}{4Q^2}}}$$

3.3 Power Consumption

Power is given by

$$P = \mathbf{F} \cdot \mathbf{v} = \mathbf{F} \cdot \dot{\mathbf{x}} = F_0 \sin \omega t \frac{d}{dt} \left(B \sin(\omega t - \delta) \right) = F_0 B \omega \sin \omega t \cos(\omega t - \delta) = P(\omega)$$

The mean power within period is given by[4]

$$\bar{P} = \frac{F_0 B \omega}{T} \int_0^T \sin \omega t \cos(\omega t - \delta) dt = \frac{F_0^2 \omega^2 B}{m\left((\omega_0^2 - \omega^2)^2 + 4\omega^2\beta^2\right)} = \bar{P}(\omega)$$

We continue to make observation,

$$\bar{P}(\omega) = \frac{F_0^2 B}{m\left(\frac{(\omega_0^2 - \omega^2)^2}{\omega^2} + 4\beta^2\right)}$$

We get maximun mean power $\bar{P}_{max} = \frac{F_0^2}{4m\beta}$ when $\omega = \omega_0$

[3]For comparison: $\omega_r < \omega_f < \omega_0$

[4]The derivation following some crucial integral and δ:

$$\int_0^T \sin \omega t \cos \omega t dt = 0, \qquad \int_0^T \sin^2 \omega t dt = \frac{T}{2}, \qquad \tan \delta = \frac{2\beta\omega}{\omega_0^2 - \omega^2}$$

In terms of quality factor, we can re-express the mean power as[5]

$$\bar{P} = \frac{F_0^2 \omega^2 B}{m\left((\omega_0^2 - \omega^2)^2 + 4\omega^2\beta^2\right)} = \frac{F_0^2 \omega^2 / 2kQ}{\left(\frac{\omega_0}{\omega} - \frac{\omega}{\omega_0}\right)^2 + 4\beta^2}$$

3.4 Bandwidth

Since the mean power is the a function respect to ω, we can draw $P - \omega$ graph. Within that graph, we further define that the width of the line segment at $\frac{1}{2}\bar{P}_{max}$ is bandwidth γ.[6]

3.5 Principle of Superposition

Principle of superposition[7] states that if two or more driving forces act on a system the response will equal the vector sum of the responses for each force acting alone.

[5]Refer to the previous definition: $k = m\omega_0^2$

[6]This result rises from non-trivial derivation:

$$\bar{P} = \frac{1}{2}\bar{P}_{max} \Rightarrow \left(\frac{\omega_0}{\omega} - \frac{\omega}{\omega_0}\right)^2 + \frac{1}{Q^2} = \frac{2}{Q^2}$$

Making the approximation that $\frac{\omega_0}{\omega} - \frac{\omega}{\omega_0} \approx 2\frac{\omega_0 - \omega}{\omega_0}$, we obtain that $\omega = \omega_0 \pm \frac{\omega_0}{2Q} = \omega_0 \pm \beta$, therefore gives that $\Delta\omega = \gamma$. Moreover, even if you did not make the approximation, the result is the same, because the second term in roots will cancel each out in the last subtraction.

[7]Actually, principle of superposition for multiple periodic force naturally rises from the linear combination of the particular solutions corresponding to different source terms.

Chapter 2

Wave Models

§4 Wave Equation

A wave is a disturbance of a continuous medium that propagates with fixed shape at constant velocity. How would you represent such an object mathematically? In a simple scenario, a wave is generated by shaking one end of a taut string. Let's consider a simple property that the possible mathematical expression of wave should apply. Denote $f(z,t)$ as the displacement of the string at point z at time t and the initial position $g(z) = f(z,0)$

$$f(z,t) = f(z - vt, 0) = g(z - vt)$$

That statement captures (mathematically) the essence of wave motion. It tells us that the function $f(z,t)$, which might have depended on z and t in the very special combination $z - vt$.

Based on that, wave equation gives mathematical description for waves and automatically admits as solution as all function of the form $f(z,t) = g(z - vt)$[1],

$$\underbrace{\nabla^2 \Psi}_{\text{Force}} = \underbrace{\frac{1}{v^2} \frac{\partial^2}{\partial t^2} \Psi}_{\text{Inertia}} \tag{2.1}$$

4.1 Derivation of Wave Function

Transverse Waves

Let's derive wave equation for one-dimensional case. For wave on a string (transverse waves) with linear density ρ and the string tension T (for a small angle θ) as figure 2.1 shows

[1] The most general solution to wave equation is of the form

$$\Psi(z,t) = f(z - vt) + g(z + vt)$$

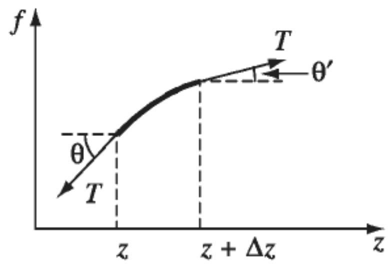


Figure 2.1: String segment

$$\Delta F = T(\sin\theta' - \sin\theta) \sim T(\tan\theta' - \tan\theta) \sim T\left(\left.\frac{\partial f}{\partial z}\right|_{z+\Delta z} - \left.\frac{\partial f}{\partial z}\right|_{z}\right) \sim T\frac{\partial^2 f}{\partial z^2}\Delta z$$

Newton's second law says: $\Delta F = \rho\Delta z\frac{\partial^2 f}{\partial t^2}$, therefore gives the wave equation:

$$\frac{\partial^2 f}{\partial z^2} = \frac{\rho}{T}\frac{\partial^2 f}{\partial t^2} \tag{2.2}$$

and high dimension follow the same derivation.

Notice, in this case, the propagating speed $v = \sqrt{\frac{T}{\rho}}$

For simplicity, we usually consider sinusoidal wave of the form $z = A\sin(kx - \omega t + \phi)$[2]

Longitudinal Waves

First consider the following spring model as shown by figure 2.2, an array of little weights of mass m interconnected with massless springs of length h. The springs have a spring constant of k:

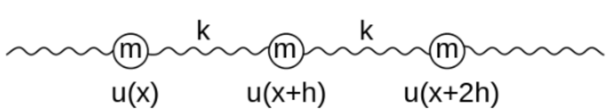


Figure 2.2: Springs system

Here the dependent variable $u(x)$ measures the distance from the equilibrium of the mass situated at x, so that $u(x)$ essentially measures the magnitude of

[2]By convention, we denote $k = \frac{2\pi}{\lambda}$ and called it *angular wave number*, $\omega = \frac{2\pi}{T}$ and called it *angular frequency*, and $\nu = \frac{1}{T}$ as *harmonic frequency*, by this

$$v = \lambda\nu = \frac{\lambda}{T} = \frac{\omega}{k}$$

a disturbance (i.e. strain) that is travelling in an elastic material. The forces exerted on the mass m at the location $x+h$ are:

$$F_{Newton} = m \cdot a(t) = m \cdot \frac{\partial^2}{\partial t^2} u(x+h,t)$$

$$F_{Hooke} = F_{x+2h} - F_x = k\left[u(x+2h,t) - u(x+h,t)\right] - k[u(x+h,t) - u(x,t)]$$

The equation of motion for the weight at the location $x+h$ is given by equating these two forces:

$$\frac{\partial^2}{\partial t^2} u(x+h,t) = \frac{k}{m}[u(x+2h,t) - u(x+h,t) - u(x+h,t) + u(x,t)]$$

where the time-dependence of $u(x)$ has been made explicit.

If the array of weights consists of N weights spaced evenly over the length $L = Nh$ of total mass $M = Nm$, and the total spring constant of the array $K = k/N$ we can write the above equation as:

$$\frac{\partial^2}{\partial t^2} u(x+h,t) = \frac{KL^2}{M} \frac{u(x+2h,t) - 2u(x+h,t) + u(x,t)}{h^2}$$

Taking the limit $N \to \infty$, $h \to \infty$ and assuming smoothness one gets:

$$\frac{\partial^2 u(x,t)}{\partial t^2} = \frac{KL^2}{M} \frac{\partial^2 u(x,t)}{\partial x^2} \tag{2.3}$$

Then, we can continue to apply it on continuous medium, for instance stress pulse in a bar. In the case of a stress pulse propagating through a beam the beam acts much like an infinite number of springs in series and can be taken as an extension of the equation derived for Hooke's law.

A beam of constant cross section made from a linear elastic material has a stiffness K given by $K = \frac{EA}{L}$, where A is the cross sectional area and E is the Young's modulus of the material. The wave equation becomes

$$\frac{\partial^2 u(x,t)}{\partial t^2} = \frac{EAL}{M} \frac{\partial^2 u(x,t)}{\partial x^2}$$

AL is equal to the volume of the beam and therefore $\frac{AL}{M} = \frac{1}{\rho}$ where $\rho$is the density of the material. The wave equation reduces to

$$\frac{\partial^2 u(x,t)}{\partial x^2} = \frac{\rho}{E} \frac{\partial^2 u(x,t)}{\partial t^2} \tag{2.4}$$

The speed of a stress wave in a beam is therefore $v = \sqrt{\frac{E}{\rho}}$.[3]

[3] From this, it is clear that wave speed is fixed by material properties

4.2 Impedance of Stretched String

When a transverse wave propagates in a string, the string offers certain opposition to the propagation of wave. This opposition is known as *impedance*[4] [5] [6] [7].

$$Z = \frac{\text{maximum transverse force}}{\text{maximum transverse velocity}} = \frac{T}{v} = \rho v = \sqrt{T\rho}$$

Derivation of Impedance

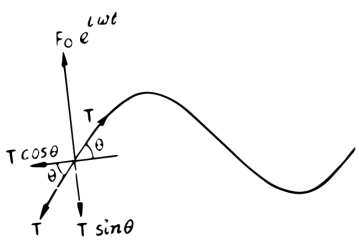


Figure 2.3: Impedance

Here is the derivation of impedance:

$$Z = \frac{\text{maximum transverse force}}{\text{maximum transverse velocity}} = \frac{T}{v}$$

Consider applying an alternating force $F = F_0 e^{i\omega t}$ to a string where F_0 is amplitude of F and ω is the angular frequency of this alternating force. T is he constant force acting throughout the length of string.

As shown left, making approximaion, for a small enough angle θ

$$F_0 e^{i\omega t} = -T\sin\theta \to -T\tan\theta = -T\frac{\partial\psi(z,t)}{\partial z}$$

Since our wave function is $\psi(z,t) = Ae^{i(\omega t-kz)}$, therefore, $\frac{\partial\psi(z,t)}{\partial z} = -ikAe^{i(\omega t-kz)}$. At the end $z = 0$

$$F_0 e^{i\omega t} = -T\left(-ikAe^{i\omega t}\right) \Rightarrow F_0 = ikTA$$

But $k = \frac{\omega}{v}$, therefore, $A = \frac{F_0 v}{iT\omega}$, thus $\psi(z,t) = \frac{F_0 v}{iT\omega}e^{i(\omega t-kz)}$. The transverse velocity is given by

$$V = \dot{\psi} = \frac{F_0 v}{T}e^{i(\omega t-kz)} \Rightarrow V_{max} = \frac{F_0 v}{T}$$

and for the maximum transverse force $F_{max} = F_0$, therefore,

$$Z = \frac{\text{maximum transverse force}}{\text{maximum transverse velocity}} = \frac{F_0 T}{F_0 v} = \frac{T}{v} = \rho v = \sqrt{T\rho}$$

[4] $T = \rho v^2$
[5] Make it clear that the maximum transverse force is not T, but proportional to T
[6] From this you can see that impedance is only depend on the material properties
[7] Mechanics Impedance:
$$F(\omega) = Z(\omega)v(\omega)$$

4.3 Transport of Power

Wave in a stretched string carries energy. And the *transport energy* of a string is given by the *kinetic energy* of string segment of length Δz:

$$\Delta E = \frac{1}{2}\Delta m V_{max}^2 = \frac{1}{2}\rho\omega^2 A^2 \Delta z \quad (V_{max} = A\omega)$$

Since the rate of transfer of energy is given by the product of energy of unit length and velocity of propagation of wave, therefore rate of transfer of energy, or *power*, will be

$$P = \frac{1}{2}\rho\omega^2 A^2 v = \frac{1}{2} Z\omega^2 A^2$$

§5 Wave at Boundary

5.1 Reflection and Transmission

A boundary divides regions of different impedance, an incident wave propagates at boundary will be both reflected and transmitted forming *reflective wave* and *transmitted wave*.

Denote that the incident wave function is $\psi_i = A_i e^{i(\omega t - k_1 z)}$, reflective wave as $\psi = B_r e^{i(\omega t + k_1 z)}$, and transmitted wave as$\psi = A_t e^{i(\omega t - k_2 z)}$ and then $f_i(t - \frac{x}{v_1}) = \psi_i$, $f_r(t + \frac{x}{v_1}) = \psi_r$, and $f_t(t - \frac{x}{v_1}) = \psi_t$.

Derivation of Reflective and Transmission Coefficients

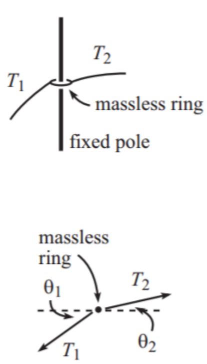


Figure 2.4: Boundary

First, let's consider the boundary condition at $x = 0$ between two different densities. The string is continuous, thus for all t,[8]

$$\psi_L(0,t) = \psi_R(0,t) \Rightarrow f_i(t) + f_r(t) = f_t(t) \tag{2.5}$$

this condition gives that

$$A_i + B_r = A_t \tag{2.6}$$

and the slope is continuous, thus:

$$\frac{\partial \psi_L(z,t)}{\partial z|_{z=0}} = \frac{\partial \psi_R(z,t)}{\partial z|_{z=0}} \Rightarrow -\frac{f_i'(t)}{v_1} + \frac{f_r'(t)}{v_1} = -\frac{f_t'(t)}{v_2} \tag{2.7}$$

or $v_2 f_i(t) - v_2 f_r(t) = v_1 f_t(t)$.

Then, consider the scenario as shown in the left, the net transverse force on the massless ring must also be zero to avoid infinite acceleration, This condition

[8] Denote $\frac{\partial f}{\partial z} = f'$ and $\frac{\partial f}{\partial t} = \dot{f}$

gives $T_1 \sin\theta_1 = T_2 \sin\theta_2$, by approximation $\sin\theta \to \tan\theta$ when θ is small enough,

$$T_1 \left.\frac{\partial \psi_L(z,t)}{\partial z}\right|_{z=0} = T_2 \left.\frac{\partial \psi_R(z,t)}{\partial z}\right|_{z=0}$$

Noted, the tension within string is equal $T_1 = T_2$, then it reduces to the *first boundary condition* (2.5). Then with tension now distinct, consider the *second boundary condition* (2.7), the equation convert into

$$-\frac{f_i'(t)}{v_1}T_1 + \frac{f_r'(t)}{v_1}T_1 = -\frac{f_t'(t)}{v_2}T_2 \Rightarrow -f_i'(t)Z_1 + f_r'(t)Z_1 = -f_t'(t)Z_2$$

therefore,

$$\begin{aligned} -ik_1 A_i e^{i(\omega t - k_1 z)} Z_1 + ik_1 B_r e^{i(\omega t + k_1 z)} Z_1 &= -ik_2 A_t e^{i(\omega t - k_2 z)} Z_2 \\ -k_1 A_i e^{i\omega t} Z_1 + k_1 B_r e^{i\omega t} Z_1 &= -k_2 A_t e^{i\omega t} Z_2 \quad (z=0) \\ -k_1 A_i Z_1 + k_1 B_r Z_1 &= -k_2 A_t Z_2 \end{aligned} \tag{2.8}$$

combining (2.6) and (2.8), we can determine the *reflection coefficient* and *transmission coefficient*

$$R \equiv \frac{B_r}{A_i} = \frac{v_2 - v_1}{v_2 + v_1} = \frac{Z_1 - Z_2}{Z_1 + Z_2} = \frac{\sqrt{\rho_1} - \sqrt{\rho_2}}{\sqrt{\rho_1} + \sqrt{\rho_2}}$$
$$T \equiv \frac{A_t}{A_i} = \frac{2v_2}{v_2 + v_1} = \frac{2Z_1}{Z_1 + Z_2} = \frac{2\sqrt{\rho_1}}{\sqrt{\rho_1} + \sqrt{\rho_2}}$$

Reflection and Transmission of Power

1. *Reflection coefficient of power*

$$\frac{\text{Reflected power}}{\text{Incident power}} = \frac{\frac{1}{2}Z_1\omega^2 B_r^2}{\frac{1}{2}Z_1\omega^2 A_i^2} = R^2 = \left(\frac{Z_1 - Z_2}{Z_1 + Z_2}\right)^2$$

2. *Transmission coefficient of power*

$$\frac{\text{Transmitted power}}{\text{Incident power}} = \frac{\frac{1}{2}Z_2\omega^2 A_t^2}{\frac{1}{2}Z_1\omega^2 A_i^2} = T^2 = \left(\frac{2Z_1}{Z_1 + Z_2}\right)^2$$

5.2 Image Waves

The most general solution to wave equation (2.1) is of the form

$$\Psi(z,t) = f(z - vt) + g(z + vt)$$

its physical interpretation is clear that every solution satisfies wave equation can be decomposed into a wave moving to right $f(z - vt)$ and a wave moving to left $g(z + vt)$.

Applying this idea, we can explain wave reflection at a fixed end quite clearly, please refer to figure 2.5

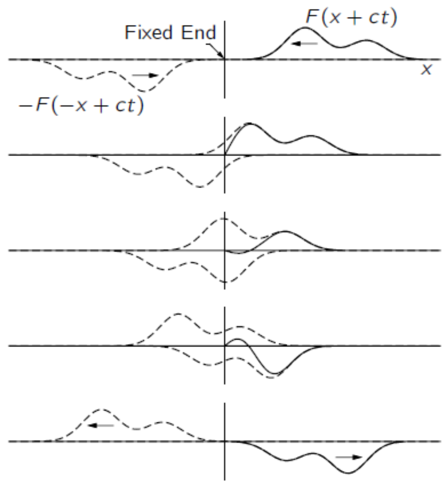


Figure 2.5: Reflection of a wave as a superposition of two travelling waves.

§6 Hygens' Principle

Like all waves, light can be viewed as a series of crests and trough moving away from a source.

A *ray* is an arrow drawn perpendicular to a wave front that points in the direction of the waves motion.

6.1 Huygens' Principle

Huygens' theory is essentially based on a geometrical construction which allows us to determine the shape of the wave front at any time, if the shape of the wave front at an earlier time is known. A *wave front* is the locus of the points which are in the same phase.

Now, according to *Huygens' principle*[9], each point of a wave front is a source of secondary disturbance, and the wavelets emanating from these points spread out in all directions with the speed of the wave. The envelope of these wavelets gives the shape of the new wave front.

[9]Huygens' principle can be seen as a consequence of the homogeneity of space the space is uniform in all locations. Any disturbance created in a sufficiently small region of homogeneous space (or in a homogeneous medium) propagates from that region in all geodesic directions. The waves created by this disturbance, in turn, create disturbances in other regions, and so on. The superposition of all the waves results in the observed pattern of wave propagation.

6.2 Principle of Reversibility

The path of a ray of light through any system is completely reversible.

§7 Wave Behaviours

7.1 Reflection and Transmission

7.2 Refraction

7.3 Interference

Superposition principle states that when two or more waves move in the same linear medium, the net displacement of the medium at any point equals the algebraic sum of the wave functions of the individual waves.

Harmonic Waves

Consider two waves, denoted as $\psi_1(z,t) = A_0 \sin(kz - \omega t)$ and $\psi_2(z,t) = A_0 \sin(kz - \omega t - \phi)$, have same travelling direction with same amplitude and frequency only having phase difference ϕ are called *harmonic waves*. The interference between two harmonic waves

$$\psi(z,t) = \psi_1(z,t) + \psi_2(z,t) = A_0 \sin(kz - \omega t) + A_0 \sin(kz - \omega t - \phi) = 2A_0 \cos\frac{\phi}{2} \sin(kz - \omega t - \frac{\phi}{2})$$

Therefore, we can see that after the interference, the amplitude becomes $A = 2A_0 \cos\frac{\phi}{2}$, and the path difference between the two waves can be determined by $\Delta z = \frac{\phi}{2\pi}\lambda$

Standing Waves

A *standing wave*[10] can be produced in a string due to the interference of two sinusoidal waves with equal amplitude and frequency traveling in opposite directions, denoted as $\psi_i(z,t) = A_0 \sin(kz - \omega t)$ and $\psi_r(z,t) = A_0 \sin(kz + \omega t)$. The interference between these two sinusoidal waves gives the standing waves

$$\psi(z,t) = \psi_i(z,t) + \psi_r(z,t) = A_0 \sin(kz - \omega t) + A_0 \sin(kz + \omega t) = 2A_0 \sin kz \cos \omega t$$

For standing wave, there are two situation that is worthy to consider:

Standing wave in a string fixed as both end First, let's make some definitions. The string has a number of natural patterns of vibration, called *normal modes*. Each normal mode has a *characteristic frequency*. The lowest of these frequencies is called the *fundamental frequency*, which together with the higher

[10] Strictly speaking, standing wave is not a wave because it doesn't satisfy wave equation (2.1).

frequencies form a *harmonic series*. *Antinodes* are points of maximum displacement, while *nodes* are points of zero displacement.

Since both end are fixed, therefore, the total string length must be the multiple of half of the wavelength $L = n\frac{\lambda_n}{2}$, $(n \in \mathbb{N})$, and for the frequency, we have the relation[11] that $\nu_n = \frac{nv}{2L} = \frac{n}{2L}\sqrt{T\rho}$.

Standing waves in air columns (fixed end or open end) Sound sources can be used to produce longitudinal standing waves in air columns.The phase relationship between incident and reflected waves depends on whether or not the reflecting end of the air column is open or closed.

1. an open pipe

   For the wave length, since one end is open, therefore, at both end[12], it has to be an antinode meaning the total pipe length is multiple of the half of the wavelength, $L = n\frac{\lambda_n}{2}$, $(n \in \mathbb{N})$, thus referring to the longitudinal wave equation (2.4), we can derive the frequency $\nu_n = \frac{nv}{2L} = \frac{n}{2L}\sqrt{E/\rho}$

2. a closed pipe

   For the wave length, since one end is closed, therefore, at the end[13], it has to be a node meaning the total pipe length is multiple of the quarter of the wavelength, $L = n\frac{\lambda_n}{4}$, $(n = 2k+1, \text{where } k \in \mathbb{N})$, thus referring to the longitudinal wave equation (2.4), we can derive the frequency $\nu_n = \frac{nv}{4L} = \frac{n}{4L}\sqrt{E/\rho}$

7.4 Beats

Beats are formed by the combination of two waves of equal amplitude but slightly different frequencies travelling in the same direction. Denote the two waves as $\psi_1(z,t) = A_0\cos(\omega_1 t - k_1 z)$ and $\psi_2(z,t) = A_0\cos(\omega_2 t - k_2 z)$. By the superposition principle, we have the interference as shown by 2.6,

$$\begin{aligned}\psi &= \psi_1 + \psi_2 = A_0\cos(\omega_1 t - k_1 z) + A_0\cos(\omega_2 t - k_2 z)\\ &= \underbrace{2A_0\cos\left(\frac{\omega_1-\omega_2}{2}t - \frac{k_1-k_2}{2}z\right)}_{\text{Amplitude modulation}}\underbrace{\cos\left(\frac{\omega_1+\omega_2}{2}t - \frac{k_1+k_2}{2}z\right)}_{\text{Travelling wave part}}\end{aligned}$$

[11]Since $v = \frac{\omega_n}{k_n} = \lambda_n\nu_n = \sqrt{T\rho}$

[12]If you want to generate sound wave, the start node is also an open end thus force to be an antinode.

[13]If you want to generate sound wave, the start node is also an open end thus force to be an antinode.

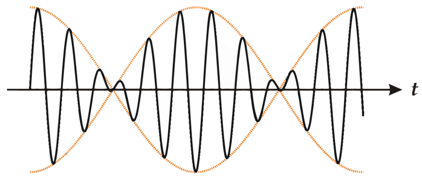


Figure 2.6: Beats

7.5 Wave Packet

Group Velocity and Phase Velocity

Continue with the beats, let's redenote the wave equation that $\psi_1 = A_0 \sin(kz - \omega t)$ and $\psi_2 = A_0 \sin((k + \Delta k)z - (\omega + \Delta\omega)t)$. Their superposition gives the waveform as shown by 2.6,

$$\psi = \psi_1 + \psi_2 = \underbrace{2A_0 \cos\left(\frac{\Delta\omega}{2}t - \frac{\Delta k}{2}z\right)}_{\text{Modulated amplitude}} \underbrace{\sin\left((k + \frac{\Delta k}{2})z - (\omega + \frac{\Delta\omega}{2})t\right)}_{\text{sine wave}}$$

Notice right now, we have two different wave velocity, one is the *phase velocity* v_p defined as the rate at which the phase of the wave propagates in space, and the other is *group velocity* v_g defined as the velocity with which the overall shape of the waves' amplitudes known as the modulation or *envelope* of the wave propagates through space.

$$v_p = \frac{\lambda}{T} = \frac{\omega}{k} \tag{2.9}$$

$$v_g = \frac{\Delta\omega}{\Delta k} = \frac{d\omega(k)}{dk} \tag{2.10}$$

An alternative approach A wave package is the superposition of a bunch of harmonic waves, which can be written as $f(x,t) \sum_\nu a_\nu \cos(2\pi\nu)\left(t - \frac{x}{v(\nu)}\right)$, where ν is the frequency. The position of the center of the package is the coordinate where the amplitude is the maximal, so we can differentiate with ν[14]:

$$\frac{d}{d\nu}[\nu(t - \frac{x}{v(\nu)})] = 0.$$

Simplify, we have $t = x\frac{d}{d\nu}(\frac{\nu}{v})$, thus $V_g = \frac{x}{t} = \frac{d\nu}{d(\frac{1}{\lambda})} = \frac{d\omega}{dk}$[15].

[14] The next maximal is the same as $(x,t) = (0,0)$, as time passing, at that position, all the phase angles of all the components should be the same, just like the scene of $(x,t) = (0,0)$.

[15] This approach is said to be used by Fermi. From: https://www.zhihu.com/question/56142122.

Non-dispersive and Dispersive Waves

If a travelling wave is losing energy, then it called *dispersive waves* and those does not are called *non-dispersive waves*.

In non-dispersive waves[16], all components travel at the same speed: $v_p = \frac{\omega}{k} = \text{constant}$.

In dispersive waves[17], components travel at different speeds: $v_p = \frac{\omega(k)}{k}$, thus we use group velocity $v_g = \frac{d\omega(k)}{dk}$

[16]This is for the wave travelling in vacuum only.

[17]This is true for waves travelling in a medium other than vacuum.

Chapter 3

Sound Waves - Longitudinal Waves

§8 Longitudinal Waves

Please refers to the derivation of longitudinal wave equation (2.4) in §4.1.

$$\frac{\partial^2 u(x,t)}{\partial x^2} = \frac{\rho}{E}\frac{\partial^2 u(x,t)}{\partial t^2} \tag{3.1}$$

where $u(x,t)$ is the particle displacement[1].

Here gives another more directly derivation:

By hooks' law Applying stress = Elastic modulus × Strain to the segment or $\frac{F}{A} = E\frac{\partial u(x,t)}{\partial x}$, thus:

$$\frac{\partial F}{\partial x} = EA\frac{\partial^2 u(x,t)}{\partial x^2}$$

[1]Noted that: by generalization, if we denote that the *particle displacement* as

$$\delta(\mathbf{r},t) = \delta\cos\left(\mathbf{k}\cdot\mathbf{r} - \omega t + \psi_{\delta,0}\right)$$

the *particle velocity* is determined by

$$v(\mathbf{r},t) = \frac{\partial\delta(\mathbf{r},t)}{\partial t} = \omega\delta\cos\left(\mathbf{k}\cdot\mathbf{r} - \omega t + \psi_{\delta,0} + \frac{\pi}{2}\right) = v_0\cos\left(\mathbf{k}\cdot\mathbf{r} - \omega t + \psi_{v,0}\right)$$

and the amplitude of the *acoustic pressure*

$$p(\mathbf{r},t) = -B\frac{\delta(\mathbf{r},t)}{\partial x} = -\rho v^2\frac{\delta(\mathbf{r},t)}{\partial x} = \rho c^2\mathbf{k}_x\delta\cos\left(\mathbf{k}\cdot\mathbf{r} - \omega t + \psi_{\delta,0} + \frac{\pi}{2}\right) = p_0\cos\left(\mathbf{k}\cdot\mathbf{r} - \omega t + \psi_{p,0}\right)$$

and the amplitude of the specific *acoustic impedance* is given by (using Laplace transform)

$$z(\mathbf{r},s) = |\hat{z}(\mathbf{r},s)| = \left|\frac{\hat{p}(\mathbf{r},s)}{\hat{v}(\mathbf{r},s)}\right| = \left|\frac{p_0}{v_0}\frac{\frac{s\cos\psi_{v,0} - \omega\sin\psi_{v,0}}{s^2+\omega^2}}{\frac{s\cos\psi_{p,0} - \omega\sin\psi_{p,0}}{s^2+\omega^2}}\right| = \frac{p_0}{v_0} = \frac{\rho v^2 k_x}{\omega} = \rho v$$

by Newton's second law:

$$F(x+\Delta x) - F(x) = \frac{\partial F}{\partial x}\Delta x = EA\Delta x \frac{\partial^2 u(x,t)}{\partial x^2} = m\ddot{x} = \rho A \Delta x \frac{\partial^2 u(x,t)}{\partial t^2}$$

thus, we obtain $\frac{\partial^2 u(x,t)}{\partial x^2} = \frac{\rho}{E}\frac{\partial^2 u(x,t)}{\partial t^2}$, with wave speed $v = \sqrt{\frac{E}{\rho}}$

§9 Sound Speed

9.1 Sound Speed in Liquid and Gas Media

Sound wave speed v depends on the bulk modulus B and equilibrium density ρ of the medium.

Therefore, in the liquid, the sound wave speed is determined by

$$v = \sqrt{\frac{B}{\rho}} \quad \text{where } B = -\frac{\Delta P}{\Delta V/V}$$

In the gas, the sound wave speed is determined by

$$v = \sqrt{\frac{\gamma P}{\rho}}$$

where $\gamma = \frac{c_p}{c_v}$ is *characteristic parameter* - the ratio of the specific heat at constant pressure to the specific heat at constant volume.

If it is an ideal gas, by applying ideal gas law's corollary $PM = \rho RT$, we can re-express the equation in terms of the absolute temperature T and molar mass M.

$$v = \sqrt{\frac{\gamma RT}{M}}$$

§10 Harmonic Sound Waves

For displacement wave function $u(x,t) = u_0\cos(kx - \omega t)$, its corresponding pressure can be determined by $\Delta P = \Delta P_{max}\sin(kx - \omega t) = -B\frac{\partial u(x,t)}{\partial x} = -Bku_0\sin(kx - \omega t)$, therefore, we have that The displacement and the pressure variation are out of phase by $\frac{\pi}{2}$.

The pressure amplitude is proportional to the displacement amplitude.

$$\Delta P_{max} = Bku_0 = \rho v \omega u_0$$

§11 Sound Wave Intensity and Intensity Level

11.1 Intensity

Sound wave intensity[2] is defined as $\mathbf{I} = p\mathbf{v} = \rho v(\omega A_0)^2$, the multiple of sound pressure and particle velocity.

The average sound intensity during time T is given by $< \mathbf{I} >= \frac{1}{T}\int_0^T p(t)\mathbf{v}(t)dt$, thus for wave as $u(x,t) = u_0 \cos(kx - \omega t)$

$$\begin{aligned} < I > &= \frac{1}{T}\int_0^T p(t)v(t)dt \\ &= \frac{1}{T}\int_0^T \rho v \omega^2 u_0^2 \sin^2(kx - \omega t)dt \\ &= \frac{1}{2}\rho v(\omega u_0)^2 = \frac{p_0^2}{2\rho v} = \frac{p_0^2}{2Z} \end{aligned}$$

11.2 Sound Intensity Level (SIL)

A logarithmic intensity level scale is defined using a *reference intensity* I_0[3]

$$\beta = 10 \log \frac{I}{I_0} \ (dB) \tag{3.2}$$

§12 Doppler Effect

The change in frequency heard by an observer whenever there is relative motion between the source and the observer is called *Doppler effect*.

By conventional assumptions, the observer is in the left hand side the source is in the right hand side. And sign conventional denote right as + and left as − for both observer and source speed.

$$f'_o = f_s \left(\frac{v_{\text{wave speed}} \pm v_{\text{observer speed}}}{v_{\text{wave speed}} \pm v_{\text{source speed}}} \right) \tag{3.3}$$

12.1 Shock Waves

Shock waves are produced when a sound source moves through a medium with a speed v_s which is greater than the wave speed.

$$\sin\theta = \frac{v}{v_s}$$

The shock wave front has a conical shape with a half angle which depends on the Mach number of the source, defined as the ratio $\frac{v_s}{v}$.

[2]Both $\mathbf{I}$ and $\mathbf{v}$ are vectors, which means that both have a direction as well as a magnitude. The direction of sound intensity is the average direction in which energy is flowing.

[3]Threshold of hearing $I_0 = 10^{-12} W/M^2$.

Chapter 4

Light Waves - Electromagnetic Waves

§13 Maxwell's Equations

Maxwell's equations (In general):

$$
\begin{aligned}
\nabla \cdot \mathbf{E} &= \frac{\rho}{\varepsilon_0} && \text{Gauss' Law} && (4.1a)\\
\nabla \cdot \mathbf{B} &= 0 && \text{Gauss' Law } (\mathbf{B} \text{ Fields}) && (4.1b)\\
\nabla \times \mathbf{E} &= -\frac{\partial \mathbf{B}}{\partial t} && \text{Faraday's Law} && (4.1c)\\
\nabla \times \mathbf{B} &= \mu_0 \mathbf{J} + \mu_0 \varepsilon_0 \frac{\partial \mathbf{E}}{\partial t} && \text{Ampere's Law} && (4.1d)
\end{aligned}
$$

Maxwell's equations (In matter):

$$
\begin{aligned}
\nabla \cdot \mathbf{D} &= \rho_f && \text{Gauss' Law} && (4.1e)\\
\nabla \cdot \mathbf{B} &= 0 && \text{Gauss' Law } (\mathbf{B} \text{ Fields}) && (4.1f)\\
\nabla \times \mathbf{E} &= -\frac{\partial \mathbf{B}}{\partial t} && \text{Faraday's Law} && (4.1g)\\
\nabla \times \mathbf{H} &= \mathbf{J}_f + \frac{\partial \mathbf{D}}{\partial t} && \text{Ampere's Law} && (4.1h)
\end{aligned}
$$

13.1 Derivation of Speed of Light

Electric Field Equation

$$
\begin{aligned}
\nabla \times \nabla \times \mathbf{E} = \nabla \times \left(-\frac{\partial \mathbf{B}}{\partial t}\right) &= \frac{\partial}{\partial t}(\nabla \times \mathbf{B}) \\
&= \frac{\partial \mu_0 \mathbf{J}}{\partial t} + \mu_0 \varepsilon_0 \frac{\partial}{\partial t}\frac{\partial \mathbf{E}}{\partial t} \\
\nabla^2 \mathbf{E} &= \mu_0 \varepsilon_0 \frac{\partial^2 \mathbf{E}}{\partial t^2}
\end{aligned}
$$

This is a wave equation, have the electric wave travels at the speed of $c = \sqrt{\mu_0 \varepsilon_0}$

Magnetic Field Equation

$$
\begin{aligned}
\nabla \times \nabla \times \mathbf{B} &= \nabla \times \left(\mu_0 \mathbf{J} + \mu_0 \varepsilon_0 \frac{\partial \mathbf{E}}{\partial t}\right) \\
&= \mu_0 \nabla \times \mathbf{J} + \mu_0 \varepsilon_0 \nabla \times \frac{\partial \mathbf{E}}{\partial t} = -\mu_0 \varepsilon_0 \frac{\partial}{\partial t}(\nabla \times \mathbf{E}) \\
\nabla^2 \mathbf{B} &= -\mu_0 \varepsilon_0 \frac{\partial}{\partial t}\left(-\frac{\partial \mathbf{B}}{\partial t}\right) = \mu_0 \varepsilon_0 \frac{\partial^2 \mathbf{B}}{\partial t^2}
\end{aligned}
$$

This is, again, a wave equation, have the magnetic wave travels at the speed of $c = \frac{1}{\sqrt{\mu_0 \varepsilon_0}}$

13.2 Light as Electromagnetic Waves

For speed of light equals to the phase speed of electric wave as well as magnetic waves suggests a profound nature that light is an electromagnetic waves.

$$
\begin{cases}
\nabla^2 \mathbf{E} = \mu_0 \varepsilon_0 \frac{\partial^2 \mathbf{E}}{\partial t^2} \\
\nabla^2 \mathbf{B} = \mu_0 \varepsilon_0 \frac{\partial^2 \mathbf{B}}{\partial t^2}
\end{cases}
$$

13.3 Conservation Laws

Continuity Equation

Consider the conservation law of charge, formally the charge in a volume Ω is

$$
Q(t) = \int_\Omega \rho(\mathbf{r}, t) d\tau
$$

by the local conservation law,

$$\frac{dQ}{dt} = -\oint_{\S} \mathbf{J} \cdot d\mathbf{a}$$

invoking divergence theorem and $Q(t)$,

$$\int_{\Omega} \frac{\partial \rho(\mathbf{r}, t)}{\partial t} = -\int_{\Omega} \nabla \cdot \mathbf{J} \Rightarrow \frac{\partial}{\partial t}\rho(\mathbf{r}, t) = -\nabla \cdot \mathbf{J}$$

This is the *continuity equation* - the precise mathematical statement of local conservation of charge. The purpose of this chapter is to develop the corresponding equations for local conservation of energy and momentum.

Energy and Energy Density (Poynting's Theorem)

From the electrostatics and magnetostatics, we have that

$$\begin{cases} W_e = \frac{\varepsilon_0}{2} \int E^2 d\tau \\ W_m = \frac{1}{2\mu_0} \int B^2 d\tau \end{cases}$$

Suggesting that $W_{em} = W_e + W_m$, and the total energy in electromagnetic fields, per unit volume, or should called it as *energy density*,is,

$$u = \frac{1}{2}\left(\varepsilon_0 E^2 + \frac{1}{\mu_0} B^2\right)$$

Suppose we have some charge and current configuration which, at time t, produces fields $\mathbf{E}$ and $\mathbf{B}$. In the next instant, dt, the charges move around a bit.

$$\mathbf{F} \cdot d\mathbf{l} = q(\mathbf{E} + \mathbf{v} \cdot \mathbf{B}) \cdot \mathbf{v}dt = q\mathbf{E} \cdot \mathbf{v}dt$$

so the rate at which work is done on all the charges in a volume Ω is

$$\frac{dW}{dt} = \int_{\Omega} \mathbf{E} \cdot \mathbf{J}\ d\tau$$

Thus we can see that $\mathbf{E} \cdot \mathbf{J}$ is the work done per unit time, per unit volume, which is to say, the power delivered per unit volume or *energy flux*.

Then, by using Ampere-Maxwell law, we can eliminate $\mathbf{J}$

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

and by invoking Faraday's law ($\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}$) and product rule $\nabla \cdot (\mathbf{E} \times \mathbf{B}) = \mathbf{B} \cdot (\nabla \times \mathbf{E}) - \mathbf{E} \cdot (\nabla \times \mathbf{B})$,

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

as $\mathbf{B} \cdot \frac{\partial \mathbf{B}}{\partial t} = \frac{1}{2}\frac{\partial}{\partial t}(B^2)$ and $\mathbf{E} \cdot \frac{\partial \mathbf{E}}{\partial t} = \frac{1}{2}\frac{\partial}{\partial t}(E^2)$, so

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

applying the divergence theorem to the second term again, we obtain that:

$$\frac{dW}{dt} = -\frac{d}{dt}\int_\Omega \underbrace{\frac{1}{2}\left(\varepsilon_0 E^2 + \frac{1}{\mu_0}B^2\right)}_{u \text{ energy density}} d\tau - \oint_S \underbrace{\left(\frac{1}{\mu_0}\mathbf{E} \times \mathbf{B}\right)}_{\mathbf{S} \text{ flux density}} \cdot d\mathbf{a}$$

where S is the surface bounding Ω, this is *Poynting's theorem*, it is the "work - energy theorem" of electrodynamics.

Poynting's theorem says, then, that the work done on the charges by the electromagnetic force is equal to the decrease in energy remaining in the fields, less the energy that flowed out through the surface.

The energy per unit time, per unit area, transported by the fields is called the *Poynting vector*:

$$\mathbf{S} \equiv \frac{1}{\mu_0}\mathbf{E} \times \mathbf{B}$$

Therefore, making these definitions, we can simplify Poynting's theorem:

$$\frac{dW}{dt} = -\frac{d}{dt}\int_\Omega u d\tau - \oint_S \mathbf{S} \cdot d\mathbf{a}$$

consider no work is done on the charges in Ω which $dW/dt = 0$, thus

$$\int \frac{\partial u}{\partial t} d\tau = -\oint \mathbf{S} \cdot d\mathbf{a} = -\int (\nabla \cdot \mathbf{S} d\tau)$$

and hence

$$\frac{\partial u}{\partial t} = -\nabla \cdot \mathbf{S}$$

This is the "continuity equation" for energy - u (energy density) plays the role of ρ (charge density), and $\mathbf{S}$ takes the part of $\mathbf{J}$ (current density). It expresses local conservation of electromagnetic energy.

Momentum

Let us consider a linearly polarized electromagnetic wave propagating in the $+z$ direction; we assume the electric field to be along the x direction and the magnetic field along the y direction. The electromagnetic wave is assumed to interact with a charge q; the electric field makes the charge move up and down along the x axis.

Thus the charge acquires a certain velocity in the x direction, and since the magnetic field is along the y axis, a force $\mathbf{F} = q\mathbf{v} \times \mathbf{B}$ acts on the charge q. This force acts along the z axis (along the direction of propagation of the wave) and constitutes what is known as *radiation pressure*. Thus, $\mathbf{F} = qvB\hat{\mathbf{z}}$, since

$B = \mu_0 H = \frac{E}{c}$, we obtain $\mathbf{F} = \frac{qEv}{c}\hat{\mathbf{z}}$. Now qEv represents the work done by the field on the charge per unit time; thus, then $\mathbf{F} = \frac{1}{c}\frac{dU}{dt}\hat{\mathbf{z}}$, where U is total work done by the field.

But the force is equal to the change in momentum per unit time; consequently, the momentum per unit volume associated with the plane wave is given by

$$\mathbf{p} = \frac{U}{c}\hat{\mathbf{z}}$$

and by the Newton's law, the pressure P, applying the divergence theorem:

$$P = \frac{F}{A} = \frac{1}{cA}\frac{\partial U}{\partial t} = \frac{S}{c}$$

13.4 Plane Electromagnetic Wave

Wave Function

The simplest solution is given by plane electromagnetic wave:

$$\begin{cases} \mathbf{E} & = E_{max}\cos(kx - \omega t + \phi) \\ \mathbf{B} & = B_{max}\cos(kx - \omega t + \phi) \end{cases}$$

where $c = \frac{\omega}{k} = \lambda f$.

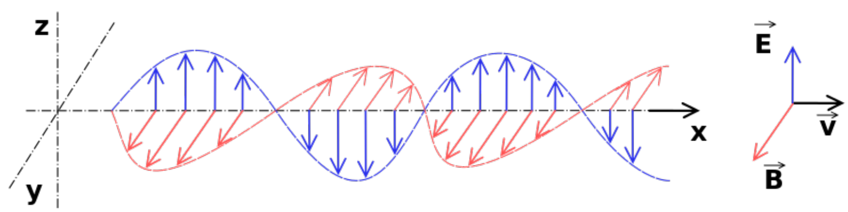


Figure 4.1: Plane electromagnetic wave

For plane electromagnetic wave, there are some basic properties:

- Electromagnetic waves obey the principle of superposition
- $\mathbf{E}$ and $\mathbf{B}$ in the empty space is related by the following equation: $\frac{E}{B} = c$[1]

  Proof. First, we can express the electromagnetic wave as the following form:

$$\mathbf{E}(\mathbf{r}, t) = \mathbf{E}_0\ e^{i(\mathbf{k}\cdot\mathbf{r} - \omega t)}$$
$$\mathbf{B}(\mathbf{r}, t) = \mathbf{E}_0\ e^{i(\mathbf{k}\cdot\mathbf{r} - \omega t)}$$

[1] Refer to the Poynting vector

then we consider the Faraday's law:

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}$$

$$\begin{aligned}
\nabla \times \mathbf{E} &= \nabla \times \left(\mathbf{E}_0 \ e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)} \right) \\
&= i\mathbf{k} \times \mathbf{E}_0 \ e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)} \\
\partial \mathbf{B}(\mathbf{r},t)/\partial t &= -i\omega \mathbf{B}_0 \ e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}
\end{aligned}$$

Thus we can see that:

$$\mathbf{k} \times \mathbf{E}_0 = \omega \ \mathbf{B}_0$$

For uniform plane electromagnetic wave, the relation simplifies itself into the following:

$$E_0 = \frac{\omega}{k} B_0 = cB_0$$

□

Therefore, for the simplest case, the energy per unit volume associated with a plane wave is given by:

$$u = \frac{1}{2}\varepsilon_0 E^2 + \frac{1}{2}\frac{B^2}{\mu_0}$$

Furthermore, for linearly polarized plan wave,

$$u = \frac{1}{2}\varepsilon_0 E^2 \cos^2(kx - \omega t) + \frac{1}{2}\frac{B^2}{\mu_0}\cos^2(kx - \omega t)$$

since $E_0 = \frac{\omega}{k} B_0 = cB_0$, we have that $\frac{1}{2}\varepsilon_0 E^2 = \frac{1}{2}\frac{B^2}{\mu_0}$, meaning the energy associated with the electric field is equal to the energy associated with the magnetic field. If we take the time average of the $\cos^2$ terms, we get $u = \varepsilon_0 E^2 \cos^2(kx - \omega t) = \frac{B^2}{\mu_0}\cos^2(kx - \omega t)$,

$$< u >= \frac{1}{2}\varepsilon_0 E^2 = \frac{1}{2}\frac{B^2}{\mu_0}$$

Further, to obtain the intensity of the beam, we must multiply $< u >$ by the speed of propagation, which will give us the energy crossing a unit area in unit time. Thus, the intensity is given by

$$I = \frac{1}{2}\varepsilon_0 c E_0^2 = \frac{1}{2}\sqrt{\frac{\varepsilon_0}{\mu_0}} E_0^2$$

13.5 Spherical Electromagnetic Waves

Wave Function

For the spherical waves, the wave equation can be rewritten as

$$\frac{\partial^2(ru)}{\partial t^2} = v^2\frac{\partial^2(ru)}{\partial \mathbf{r}^2}$$

therefore, $u(\mathbf{r},t) = Ae^{i(\omega t \pm \mathbf{k}\cdot\mathbf{r})}$, thus it follows the inverse square law:

$$I = |u(\mathbf{r},t)|^2 = \frac{|A|^2}{r^2}$$

Chapter 5

Light Wave Behaviour

§14 Fermats Principle

Fermat's principle states that the ray will correspond to that path for which the time taken is an extremum in comparison to nearby paths.

§15 Reflection and Refraction

15.1 Reflection

Law of reflection states that following the Fermat's principle, the incident angle is equal to the reflected angle. The derivation is shown in the following figure. And same result can also naturally rises with Huygens' principle.

If reflection happens on a smooth surface, then we called it as the *regular reflection*, otherwise, it is *diffused reflection*.

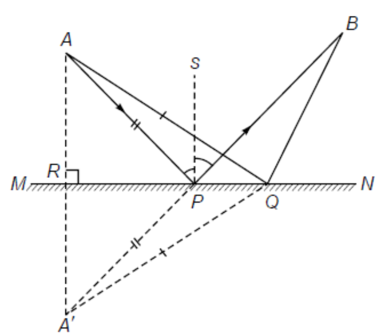


Figure 5.1: Law of Reflection

Total Internal Reflection

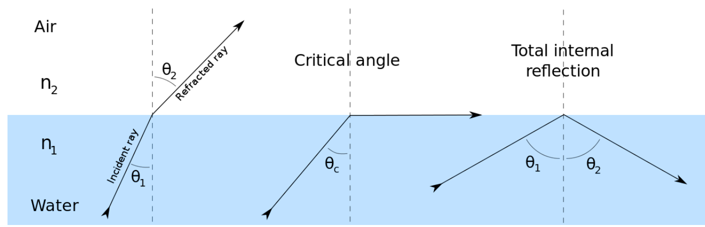


Figure 5.2: Total Internal Reflection

The critical angle: $\theta_c = \arcsin \frac{n_2}{n_1}$

15.2 Refraction

Refraction Index

Refraction index is defined as the ratio between the light speed the wave propagate speed in medium.

$$n = \frac{c}{v}$$

Snell's Law of Refraction

Snell's law of refraction states that the light beam or part of it enters a different medium with changed direction corresponding to both refraction index,

$$n_1 \sin \theta_1 = n_2 \sin \theta_2$$

This result follows naturally from both Fermat's principle and Huygen's principle.

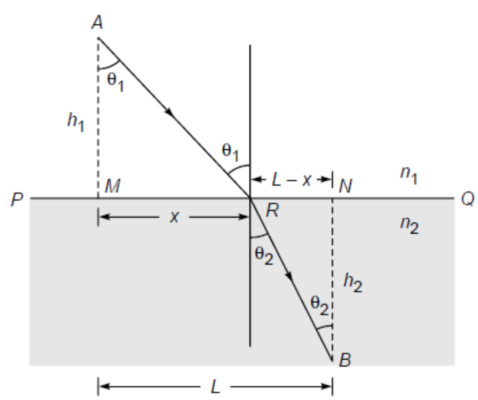


Figure 5.3: Snell's Law of Refraction

Dispersion

An important property of n is that, for a given material, the n varies with the wavelength of the light passing through the material (i.e. Prism) this behaviour is called *dispersion*.

For example, Light from the source is sent through a narrow, adjustable slit to produce a parallel beam. The light then passes through the prism and is dispersed into a spectrum.

$$v = \lambda\nu \Rightarrow \frac{\sin\theta_1}{\sin\theta_2} = \frac{n_2}{n_1} = \frac{v_1}{v_2} = \frac{\lambda_1}{\lambda_2}$$

Therefore, for light beam dispersed by prism, blue light has larger blending angle. This is also true for rainbow that the violet strip is outer side respect to the red strip. For double rainbow, the order is reversed.

15.3 Reflection and Refraction

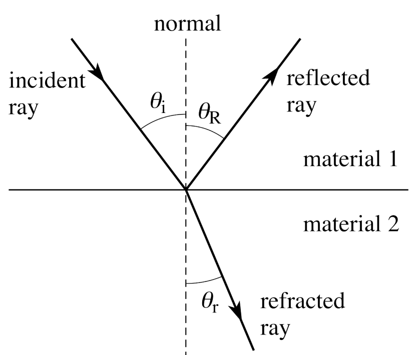


Figure 5.4: Reflection and Refraction

§16 Polarization

Polarization is a property only for transverse waves (not longitudinal).

A beam of light from a normal light source consists of many millions of individual photons and the corresponding vibration axes will be completely random in orientation. The beam as a whole is then *unpolarized*.

When an atom emits a single photon these fluctuations, within the photon, are *plane polarised*.

The vibrations in a transverse wave take place at right angles to the direction of propagation, and so are confined to the plane at right angles to this direction.

16.1 Malus' Law

If a plane-polarized beam strikes a *polaroid* - polarizing sheet whose axis is at an angle θq to the incident polarisation direction, the beam will emerge plane-polarized parallel to the Polaroid axis.

$$I = I_0 \cos^2 \theta$$

Moreover, the average intensity $\bar{I} = \frac{1}{2} I_0$, since the average of $\cos\theta$

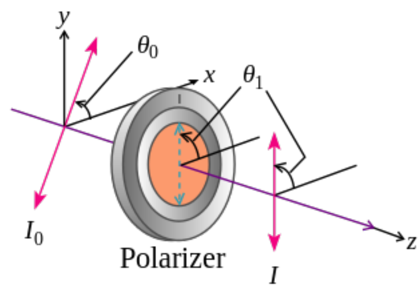


Figure 5.5: Polarizer ($\theta = \theta_1 - \theta_0$)

16.2 Producing Polarized Light

There are several ways to produce polarized light. As we seen before, polarizer is one equipment you can account to. Besides that

Reflection ans Scattering

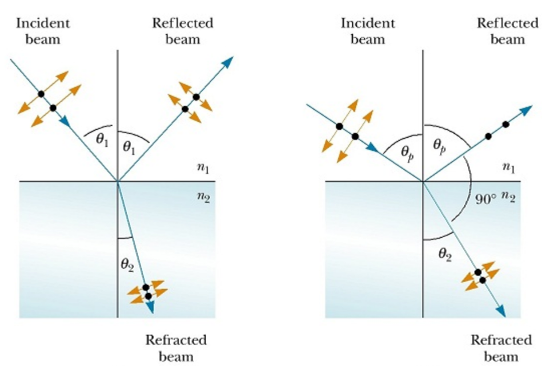


Figure 5.6: Partially Polarization and Completed Polarization

In completed polarization, by *Brewster's law*, we can determine the Brewster angle θ_p

$$\tan\theta_p = \frac{n_2}{n_1}$$

Proof.

$$\theta_p + \theta_2 = \frac{\pi}{2}$$

by the Snell's law, we have

$$n_1 \sin\theta_p = n_2 \sin\theta_2 = n_2 \cos\theta_p$$

thus we have the Brewster's law. □

Double Refraction

In many crystals, such as quartz and calcite, the polarization direction of the incident light affects the speed at which light travels through the material and hence the material's refractive index n, resulting in ordinary (O) ray and extraordinary (E) ray.

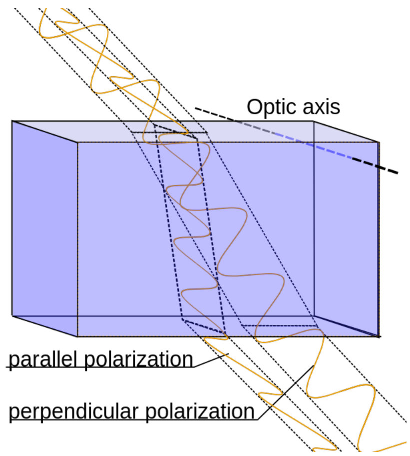


Figure 5.7: Double Refraction - Birefringence

A light beam incident on a crystal will have its two perpendicular components of polarization refracted in different directions.

16.3 Application of Polarized Light

Photoelastic Stress Analysis

Certain materials only become doubly refracting under stress - *photoelasticity*

Optical Rotation

Certain materials rotate the plane of polarization by an amount dependent on their thickness or concentration (i.e. organic chirality molecular) - *optical rotation*

§17 Interference

17.1 Superposition Principle

When two or more waves move in the same linear medium, the net displacement of the medium at any point equals the algebraic sum of the wave functions of the individual waves.

17.2 Light Wave Interference

An interference pattern with light is only observable if it is *stationary*.

A stationary pattern is only obtained if the waves involved are *coherent*[1]

Therefore, in order to get an observable light interference, the following condition must be satisfied:

- maintain a constant phase difference - coherent
- have the same frequency - monochromatic

17.3 Wavefront Division

Young's Double-Slit Experiment Young's double-slit experiment is one method of wavefront division. And the interference pattern can be explained by Huygen's principle.

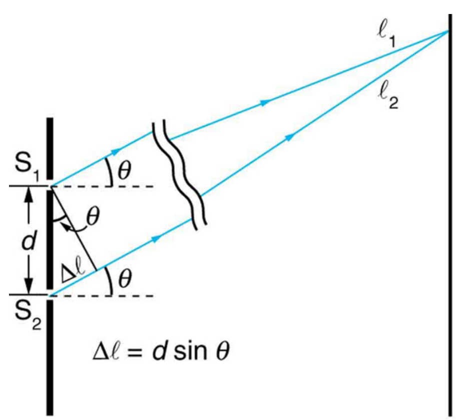


Figure 5.8: Young's Double-Slit Experiment

As the above graph shown, the path difference Δl is given by the angle θ (considered to be small enough), and the slit distance d

[1]Coherent light can be produced be wavefront division (double-slits (Youngs), Lloyds mirror) and wave amplitude division (thin film, air-wedge).

Furthermore, denote L as the distance between the board and the screen and y as the vertical distance from the center of the screen, and D as the slit width.

$$\Delta l \approx \sin\theta$$

$$y = L\tan\theta \approx L\sin\theta$$

The approximation is valid if $L \sim 1m$, $D < 1mm$ and $\lambda < 1\mu m$

For the path difference, the two light rays will have a phase difference respect to each.

$$\phi = \frac{2\pi}{\lambda}\Delta l = \frac{2\pi}{\lambda} d\sin\theta$$

Since the two light rays were divided by the same wavefront, they must have same frequency and amplitude. Therefore, at the point of interference, the two waves can be expressed by:

$$\begin{aligned} E_1 &= E_0 \sin\omega t \\ E_2 &= E_0 \sin(\omega t + \phi) \end{aligned}$$

by the principle of superposition, the resultant wave is:

$$\begin{aligned} E = E_1 + E_2 &= E_0 \sin\omega t + E_0 \sin(\omega t + \phi) \\ &= 2E_0 \cos\frac{\phi}{2} \sin\left(\omega t + \frac{\phi}{2}\right) \end{aligned}$$

And the light intensity will be proportional to the square to the amplitude of the wave.

Let's look at the location of the dark and bright fringes. For the bright fringe, it is the result of the constructive interference, therefore, at the point of interference, two wave are in phase:

$$\Delta l = d\sin\theta = m\lambda \qquad y_{\text{bright}} \approx \frac{m\lambda L}{d} \qquad m \in \mathbb{N}$$

For the dark fringe, it is the result of the destructive interference, therefore, at the point of interference, two wave are π out of phase:

$$\Delta l = d\sin\theta = \left(m + \frac{1}{2}\right)\lambda \qquad y_{\text{dark}} \approx \left(m + \frac{1}{2}\right)\frac{\lambda L}{d} \qquad m \in \mathbb{N}$$

The average intensity can also be given by the phase difference:

$$I_{\text{av}} = I_0 \cos^2\frac{\phi}{2} = I_0 \cos^2\frac{\pi d\sin\theta}{\lambda} = I_0 \cos^2\frac{\pi d y}{\lambda L}$$

Noted: The central fringe is central maximum, it's a bright fringe.

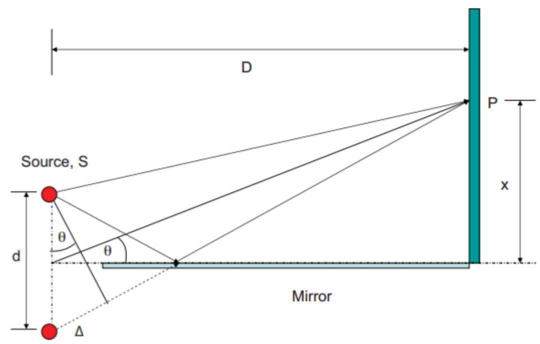


Figure 5.9: Lloyd's Mirror Experiment

Lloyd's Mirror Lloyd's mirror can be considered as one method for the wave-front division. Moreover, in this case, the reflected ray will have a π phase shift. This phenomenon will happen if light get reflected on the interface of two media and back to the one who has smaller reflected index.

Therefore, taking the phase shift at the interface into account, here is the fringe pattern.

For the bright fringe, it is the result of the constructive interference, therefore, at the point of interference, two wave are in phase:

$$\Delta l = d \sin\theta = \left(m + \frac{1}{2}\right)\lambda \qquad y_{\text{bright}} \approx \left(m + \frac{1}{2}\right)\frac{\lambda L}{d} \qquad m \in \mathbb{N}$$

For the dark fringe, it is the result of the destructive interference, therefore, at the point of interference, two wave are π out of phase:

$$\Delta l = d \sin\theta = m\lambda \qquad y_{\text{dark}} \approx \frac{m\lambda L}{d} \qquad m \in \mathbb{N}$$

And we can see that due to the phase change, the central fringe is a central minimum - a dark fringe.

Reflection Phase Change

An electromagnetic wave undergoes a phase change of π upon reflection from a medium that has a higher index of refraction than the one in which the wave is travelling. Why?[2]

This can be shown by the following figure:

[2]It should account for the conservation and continuity of electromagnetic field along any certain direction at the interface. This phase change is actually a sign flip to maintain the continuity of the wave at boundary and the conserved electromagnetic field.

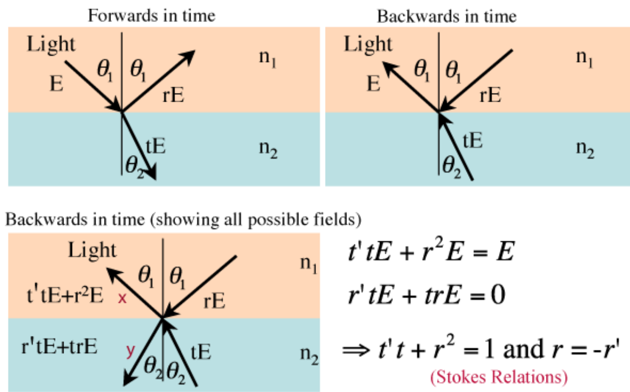


Figure 5.10: Reflection Phase Change - Stokes Relations (Non-Absorption Process)

The most interesting result here is that $r = -r'$. Thus, whatever phase is associated with reflection on one side of the interface, it is π different on the other side of the interface.

17.4 Division of Wave Amplitude

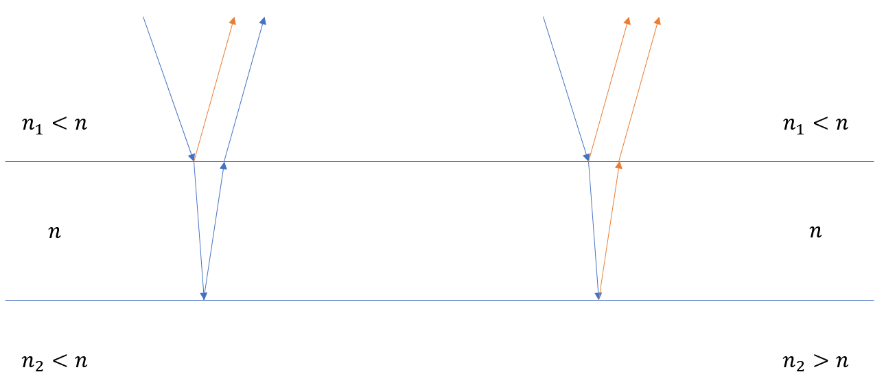


Figure 5.11: Thin Film

Thin Film In this first case, consider the top surface, the direct reflected beam has a reflection phase change and the beam going through the thin film does not, thus if the width of the thin film d is odd multiples of the quarter of the wavelength λ (resulting the path difference is the odd number of half of the wavelength), these two rays will have a constructive interference. And for the case that the width of the thin film is the even multiples of the quarter wavelength (resulting even number of the half wavelength), this two rays will have destructive interference.

Denote the reflected wavelength as λ_n. In a medium of refractive index n the wavelength λ_n is given by (frequency does not change)

$$\lambda_n = \frac{\lambda}{n}$$

and the path difference is given by

$$\Delta l = \frac{\phi}{2\pi}\lambda$$

therefore, for constructive interference:

$$d = \frac{1}{4}(2m+1)\lambda_n \Rightarrow 2nd = \left(m + \frac{1}{2}\right)\lambda \qquad m \in \mathbb{N}$$

and for the destruction interference:

$$d = \frac{1}{4}(2m)\lambda_n \Rightarrow 2nd = m\lambda \qquad m \in \mathbb{N}$$

For the second case, where two rays are all phase shift by π, therefore, the phase changes due to reflection at both the top and bottom surfaces are offsetting.

$$\begin{cases} d = \frac{1}{4}(2m+1)\lambda_n \Rightarrow 2nd = \left(m + \frac{1}{2}\right)\lambda & m \in \mathbb{N} \text{ Construction Interference} \\ d = \frac{1}{4}(2m)\lambda_n \Rightarrow 2nd = m\lambda & m \in \mathbb{N} \text{ Destruction Interference} \end{cases}$$

Fringes of Equal Thickness

Air Wedge This same phenomenon can happen in another experiment - air wedge.

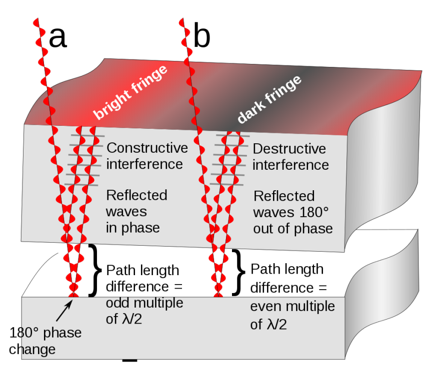
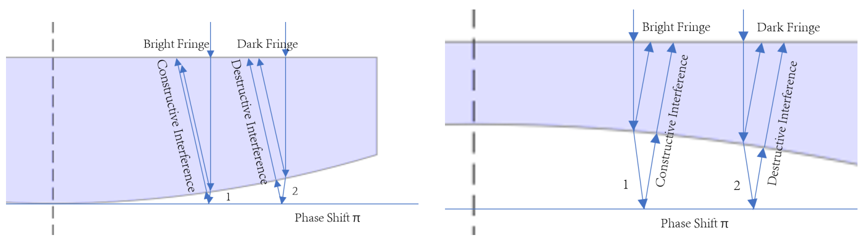


Figure 5.12: Air Wedge

Similarly, the conclusion follows:

$$\begin{cases} d = \frac{1}{4}(2m+1)\lambda_n \Rightarrow 2nd = \left(m+\frac{1}{2}\right)\lambda & m \in \mathbb{N} \text{ Construction Interference - Bright Fringe} \\ d = \frac{1}{4}(2m)\lambda_n \Rightarrow 2nd = m\lambda & m \in \mathbb{N} \text{ Destruction Interference - Dark Fringe} \end{cases}$$



(a) Plano-convex lens (b) Plano-cylindrical lens

Figure 5.13: Netwon's Rings

Newton's Ring *Newton's rings* is a phenomenon in which an interference pattern is created by the reflection of light between two surfaces a spherical surface and an adjacent touching flat surface. Basically, it is same as second case in the thin film interference and the air wedge.

$$\begin{cases} d = \frac{1}{4}(2m+1)\lambda_n \Rightarrow 2nd = \left(m+\frac{1}{2}\right)\lambda & m \in \mathbb{N} \text{ Construction Interference - Bright Fringe} \\ d = \frac{1}{4}(2m)\lambda_n \Rightarrow 2nd = m\lambda & m \in \mathbb{N} \text{ Destruction Interference - Dark Fringe} \end{cases}$$

Moreover, dark fringe has the radii given by:

$$r \approx \sqrt{\frac{m\lambda R}{n}}$$

where n is the reflective index of the glass, and R is the distance of the light source to the lower interface at the fringe. For approximation, we assume that the convex lens is part of spherical lens.

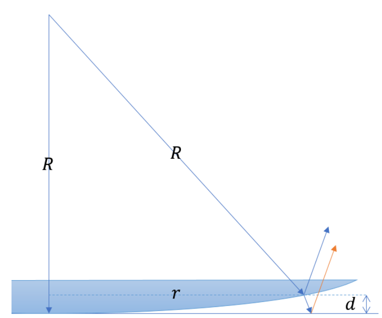


Figure 5.14: Newton's Rings

Proof. For the dark fringe, we have the destruction interference, therefore,

$$2nd = m\lambda \qquad m \in \mathbb{N},$$

Therefore, by the geometrical relation, we have,

$$\begin{aligned} r &\approx \sqrt{R^2 - (R-d)^2} && \text{Spherical lens approximation} \\ &\approx \sqrt{2Rd} && \text{Omit the second order } O(d^2) \\ &= \sqrt{\frac{m\lambda R}{n}} \end{aligned}$$

□

17.5 Phasor Diagram

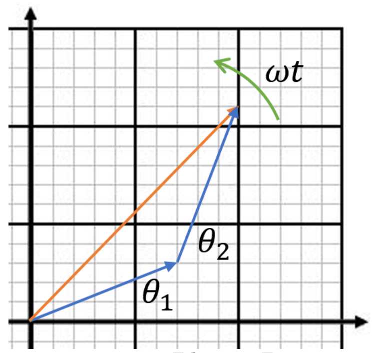


Figure 5.15: Phasor Diagram

Consider several sinusoidal waves with phase angle difference, we can clearly show the phase difference in a phase diagram:

$$\begin{aligned} E_1 &= E_0 \sin(\omega t + \theta_1) \\ E_2 &= E_0 \sin(\omega t + \theta_2) \\ &\vdots \\ E_n &= E_0 \sin(\omega t + \theta_N) \end{aligned}$$

Phasor Addition of Waves

First we can consider the first two sinusoidal waves, as shown by the figure 5.15, with the length of the phasor represents their amplitudes, we can add them like two vectors and obtain the equivalent total phasor, and that phasor represents the interference wave. Moreover, Consider that the phase difference is $\Delta\theta = \theta_2 - \theta_1$, we can calculate:

$$\begin{aligned} E_1 + E_2 &= E_0 \sin(\omega t + \theta_1) + A_0 \sin(\omega t + \theta_1 + \Delta\theta) \\ &= 2E_0 \cos\frac{\Delta\theta}{2} \sin\left(\omega t + \theta_1 + \frac{\Delta\theta}{2}\right) \end{aligned}$$

This result is the same as the double slit interference, thus we can generate this result into multiple slit interference.

17.6 Multiple Slit Interference

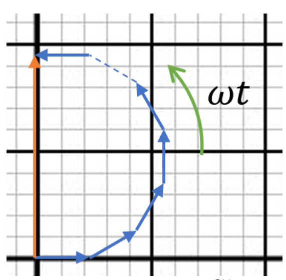


Figure 5.16: N Slits

Consider there are N slits, dividing same wavefront, thus having same frequency and amplitude. Denotes:

$$\begin{aligned} E_1 &= E_0 \sin(\omega t + \theta) \\ E_2 &= E_0 \sin(\omega t + 2\theta) \\ &\vdots \\ E_n &= E_0 \sin(\omega t + N\theta) \end{aligned}$$

In this case, we can get the final results, as

$$E = \sum_{i=1}^{N} E_i = \underbrace{N E_0 \cos\frac{N}{2}\theta}_{\text{Interference amplitude}} \sin\left(\omega t + \frac{N}{2}\theta\right)$$

Thus we can see that the intensity of interference is N^2 as the I_0, and we can also see that the secondary maximum has $N-2$ betweem each primary maximum[3].

17.7 Multiple-Beam Interference (Optional)

If the amplitude-reflection coefficients, r's, for the parallel plate, are not small, the high-order reflected waves, become quite significant, resulting *multiple-beam interference*.

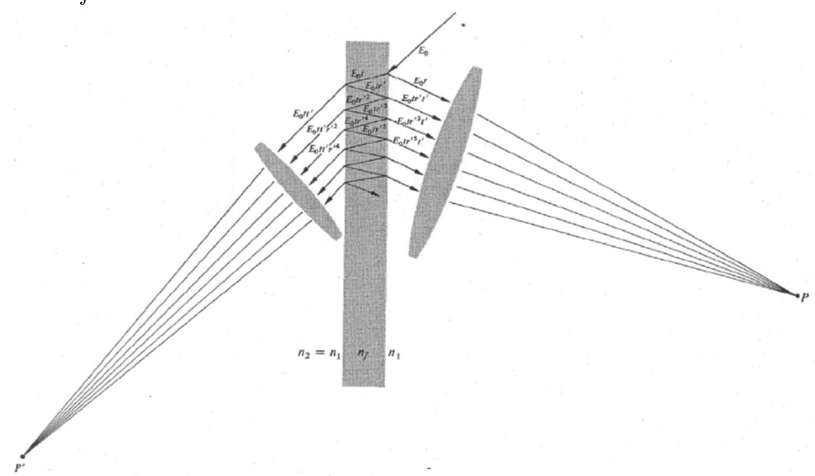


Figure 5.17: Multiple-beam interference from a parallel film

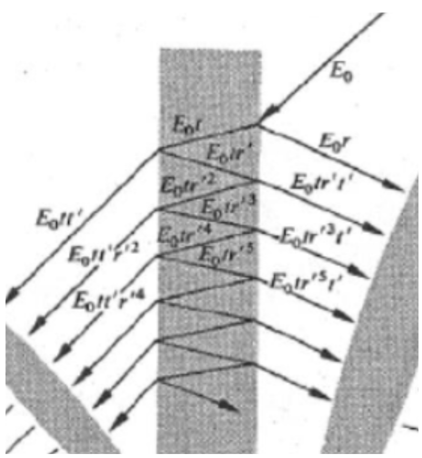


Figure 5.18: Reflection and transmission

First, to simplify the question, we make assumptions:

- The film is non-absorbing.
- $n_1 = n_2$
- Denote amplitude transmission coefficients: t as entering the film and t' as exiting the film.
- Denote amplitude reflection coefficient: t as outside the film and t' as within the film.
- The waves are mutually coherent and will all interfere with the assistance of the lenses.

[3]This question would be so much easier to understand by taking views of the phasor diagram. For N phasor, there are $N-2$ different polygon that cancel out some phasors and leave the rest to interference as secondary maximum.

Thus first, we obtain the scalar amplitudes of the reflected waves as $E_{1r}, E_{2r}, E_{3r}, \cdots$ are respectively, $E_0 r, E_0 t r' t', E_0 t r'^3 t', \cdots$, and the scalar amplitudes of the transmitted waves $E_{1t}, E_{2t}, E_{3t}, \cdots$ are respectively, $E_0 t t', E_0 t r'^2 t', E_0 t r'^4 t', \cdots$.

The phase differences arise from a combination of the optical path-length differences and phase shifts occurring at the various reflections.

Between the adjacent rays, the optical path difference is (where d is the width of the parallel film and θ_i is the incident angle and θ_t is the transmitted angle)

$$\delta l = 2 n_f d \cos\theta_t = \frac{\delta}{k_0}$$

Denotes δ is the phase difference arising from path difference. And 0 or π phase change occurs at each internal reflection ($\theta_i < \theta_c$), thus the optical field at point P are given by:

$$\begin{aligned}
E_{1r} &= E_0 r e^{i(\omega t + \pi)} \\
E_{2r} &= E_0 t r' t' e^{i(\omega t + \delta)} \\
E_{3r} &= E_0 t r'^3 t' e^{i(\omega t + 2\delta)} \\
&\vdots \\
E_{Nr} &= E_0 t r'^{2N-3} t' e^{i(\omega t + (N-1)\delta)}
\end{aligned}$$

Thus the resultant reflected scalar waves is:

$$\begin{aligned}
E_r &= \sum_{k=1}^{N} E_{kr} \\
&= E_0 r e^{i(\omega t + \pi)} + E_0 t r' t' e^{i(\omega t + \delta)} + E_0 t r'^3 t' e^{i(\omega t + 2\delta)} + \cdots + E_0 t r'^{2N-3} t' e^{i(\omega t + (N-1)\delta)} \\
&= E_0 e^{i\omega t} \left(-r + t r' t' e^{-i\delta} \left[\sum_{j=0}^{N-2} \left(r'^2 e^{-i\delta} \right)^j \right] \right) \\
&\overset{N\to\infty}{=} E_0 e^{i\omega t} \left(-r + \frac{t r' t' e^{-i\delta}}{1 - r'^2 e^{-i\delta}} \right)
\end{aligned}$$

In the case of zero absorption (no loss), $r = r'$and $tt' = 1 - r^2$ (Stokes Relations[4]), then

$$E_r = E_0 e^{i\omega t} \, \frac{r(e^{-i\delta} - 1)}{1 - r^2 e^{-i\delta}}$$

Therefore, the reflected flux density at P is then $I_r = E_r E_r^* / 2$.

$$I_r = \frac{E_0^2 r^2 (1 - e^{-i\delta})(1 - e^{i\delta})}{2(1 - r^2 e^{-i\delta})(1 - r^2 e^{i\delta})} = I_i \frac{2r^2(1 - \cos\delta)}{(1 + r^4) - 2r^2 \cos\delta}$$

[4]If you are wonder the sign, I did not flip the sign at the first place :).

where I_i is the incident flux that $I_i = E_0 E_0^*/2 = \frac{E_0^2}{2}$
Similarly, we can find the amplitude and irradiance of the transmitted waves[5]

$$E_t = E_0 e^{i\omega t} \frac{1-r^2}{1-r^2 e^{-i\delta}}$$
$$I_t = I_i \frac{(1-r^2)^2}{(1+r^4) - 2r^2\cos\delta}$$

Let's continuously make analysis for this multi-beam interference:

$$I_r = I_i \frac{2r^2(1-\cos\delta)}{(1+r^4) - 2r^2\cos\delta} \qquad I_{r-max} = I_i \frac{4r^2}{(1+r^2)^2}$$
$$I_t = I_i \frac{(1-r^2)^2}{(1+r^4) - 2r^2\cos\delta} \qquad I_{t-min} = I_i \frac{(1-r^2)^2}{(1+r^2)^2}$$

This gives that: $I_i = I_r + I_t$, coincides with the conservation of irradiance of this given light beam. And we can see that just like the relationship between the kinetic energy and potential energy in spring:

$$\begin{cases} I_i = I_{r-min} + I_{t-max} & \text{when } \cos\delta = 1 \text{ in which } \delta = 2m\pi \\ I_i = I_{r-max} + I_{t-min} & \text{when } \cos\delta = -1 \text{ in which } \delta = (2m+1)\pi \end{cases}$$

17.8 Fabry-Perot Interferometer

Based on the principle of multi beam interferometry, *Fabry-Perot interferometer*[6] consists of two plane glass (or quartz) plates which are coated on one side with a partially reflecting metallic film (of aluminum or silver) of 80% reflectivity.

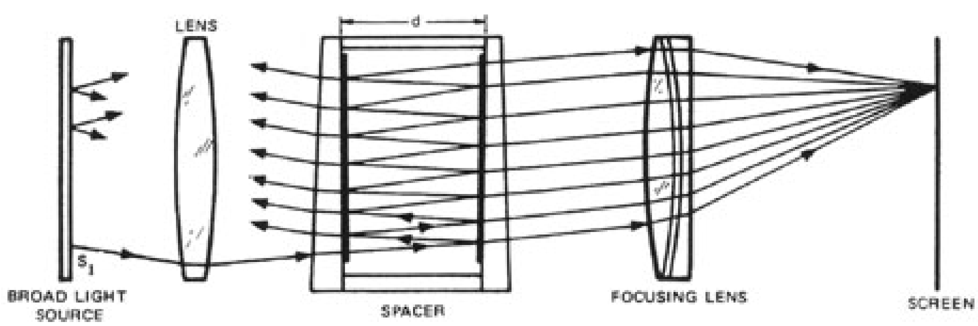


Figure 5.19: Fabry-Perot Interferometer

[5]Proof will be show at appendix.
[6]The Fabry-Perot etalon is very important in laser technology. The optical resonator in most lasers is a Fabry-Perot interferometer.

Light were partially transmitted each time reaches the second surface, resulting in multiple offset beams which can interfere with each other. The large number of interfering rays produces an interferometer with extremely high resolution.

Follow the results in the multi-slit interference, with $tt' = 1-r^2$ and $T+R = 1$ we have that:

$$I_t = I_i \frac{(1-r^2)^2}{(1+r^4) - 2r^2\cos\delta} = I_i \frac{T^2}{1+R^2-2R\cos\delta}$$

Maximum occur when $\cos\delta = 1$ in which $\delta = 2m\pi$,

$$2n_f d\cos\theta_t = \frac{\delta}{k_0} = \frac{2m\pi}{a\pi/\lambda} = m\lambda$$

As $n_f = 1$, we have that $2d\cos\theta_t m\lambda$

§18 Diffraction

The spreading and bending of wave behind obstacles that lie in their path is called *diffraction*.

The diffraction occurs when the split/object is small enough compared with light wave length.

18.1 Interference and Diffraction

Conventionally the term interference is used when considering the intensity pattern produced by superposing a finite number of separated, coherent sources.

The intensity pattern produced by interference between a continuous distribution of coherent sources is usually called *diffraction*.

18.2 Fraunhofer and Fresnel Diffraction

Fraunhofer diffraction - rays reaching a point are parallel. Produced either by having large separations between source, obstacle and screen or by using *lenses* (often the eye lens) to focus the rays.

Fresnel diffraction - observing screen is a finite distance from the slit (or edge) and the light rays are not rendered parallel by a lens.

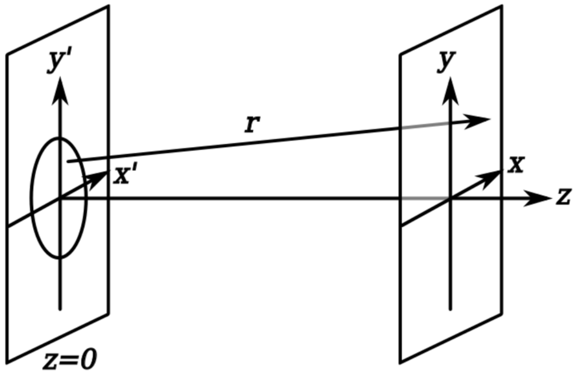


Figure 5.20: Fresnel diffraction

To give you an idea, Fresnel diffraction rises from the Huygen's principle, and its differential equation[7] is the following:

The electric field diffraction pattern at a point (x, y, z) is given by:

$$E(x, y, z) = \frac{1}{i\lambda} \iint_{-\infty}^{+\infty} E(x', y', 0) \frac{e^{ikr}}{r} dx'dy'$$

where: $E(x', y', 0)$ is the aperture, $r = \sqrt{(x - x')^2 + (y - y')^2 + z^2}$.

In contrast the diffraction pattern in the far field region is given by the Fraunhofer diffraction equation.

18.3 Single-Slit Diffraction

In single-slit diffraction, the aperture of the slit can be considered as an array of a large number N of coherent sources, each separated by a small distance Δy.

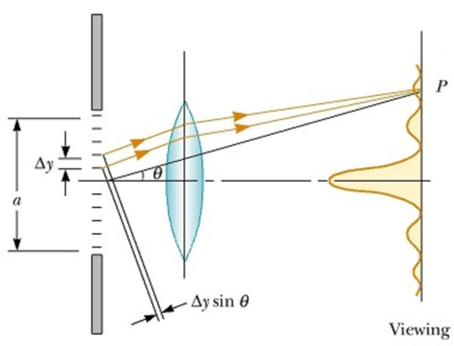


Figure 5.21: Single-Slit Diffraction

[7] Analytical solution of this integral is impossible for all but the simplest diffraction geometries. Therefore, it is usually calculated numerically.

Therefore, by making the assumption that the waves are mutually coherent and will all interfere with assistance of the lens. This reduced to a N-Slits interference, thus we can use the phasor diagram to solve this question, and then take the limits that $N \to \infty$.

Phasor difference from path difference:

$$\Delta\phi = \frac{2\pi}{\lambda}\Delta y \sin\theta \Rightarrow \phi = N\Delta\phi = \frac{2\pi}{\lambda}N\Delta y \sin\theta \overset{N\Delta y = a}{=} \frac{2\pi}{\lambda}a\sin\theta$$

where a is slit width.

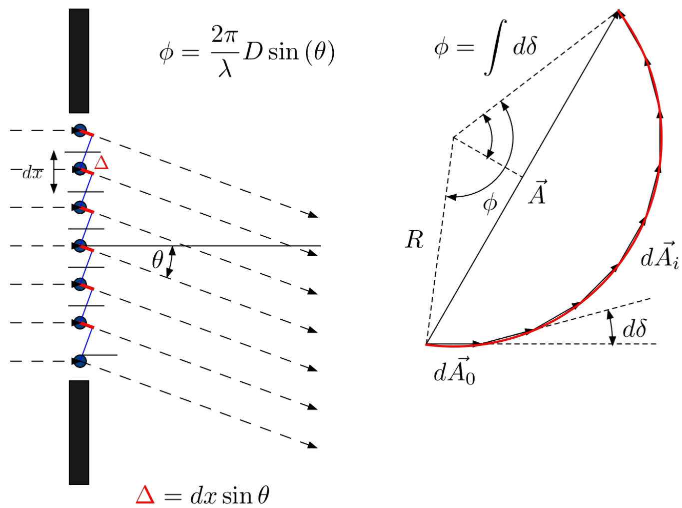


Figure 5.22: Phasor Diagram when $N \to \infty$

The total phase difference ϕ between the first and last phasors will depends on the angle θ which determines the direction to an arbitrary point on the screen.

Thinking it as phasor diagram: As θ increases, the chain of phasors eventually forms the closed path. At greater θ values, the spiral chain of phasors tightens. And for $N \to \infty$, we will have circled phasor diagram, and the chord is the resultant phasor vector.

At the same time, for the irradiance, we have that (where $E_R = \vec{A}$):

$$\sin\frac{\phi}{2} = \frac{E_R}{2R}$$

The arc length is $E_0 = E\phi$, therefore,

$$E_R = 2R\sin\frac{\phi}{2} = E_0\frac{\sin(\phi/2)}{\phi/2}$$

and the irridiance is given by:

$$I_\theta = I_0\left(\frac{\sin\frac{\phi}{2}}{\frac{\phi}{2}}\right)^2$$

And we can determine the location of the maximum and minimum.

Minimum will occur when $\sin\frac{\phi}{2} = 0$ but $\phi \neq 0$, therefore, $\frac{\phi}{2} = \pi, 2\pi, 3\pi, \cdots$, thus the dark fringe will occurs at:

$$\sin\theta = \frac{m\lambda}{a} \qquad m \in \mathbb{N}^+$$

For bright fringe, we have to find that $\phi = 0$ (central maximum) and $\tan\frac{\phi}{2} = \frac{\phi}{2}$

Proof. To obtain the maximum, $\frac{dI}{d\phi} = 0$

$$\frac{dI}{d\phi} = \frac{d}{d\phi} I_0 \left(\frac{\sin\frac{\phi}{2}}{\frac{\phi}{2}}\right)^2 = I_0 \frac{\sin\frac{\phi}{2}\cos\frac{\phi}{2}\cdot\frac{\phi}{2} - \sin^2\frac{\phi}{2}}{(\frac{\phi}{2})^3} = 0$$

Thus it requires that $\tan\frac{\phi}{2} = \frac{\phi}{2}$, where $\phi = \frac{2\pi}{\lambda} a \sin\theta$ □

For this, we have to solve it numerically to find the bright fringe. Here give the first several solutions of $\frac{\phi}{2}$: $0,\ \pm 1.4303\pi,\ \pm 2.4590\pi,\ \pm 3.4707\pi, \cdots$.

Quick Summary:

- Minima will occur when $\sin\frac{\phi}{2} = 0$ but $\phi \neq 0$, thus $\frac{\phi}{2} = \pi, 2\pi, 3\pi, \cdots$ and $\sin\theta = \frac{m\lambda}{a}$
- Maxima will occur when $\phi = 0$ (central maximum) and $\tan\frac{\phi}{2} = \frac{\phi}{2}$. Numerically, $\frac{\phi}{2}$: $0,\ \pm 1.4303\pi,\ \pm 2.4590\pi,\ \pm 3.4707\pi, \cdots$.

18.4 Double-Slit Interference and Diffraction Pattern

As studied, double slits can give interference pattern.

$$I = I_0 \cos^2\frac{\Delta\phi}{2}$$

The interference pattern obtained using slits that are wide enough can produce observable diffraction patterns.

The result gives a superposition of the interference and diffraction patterns.

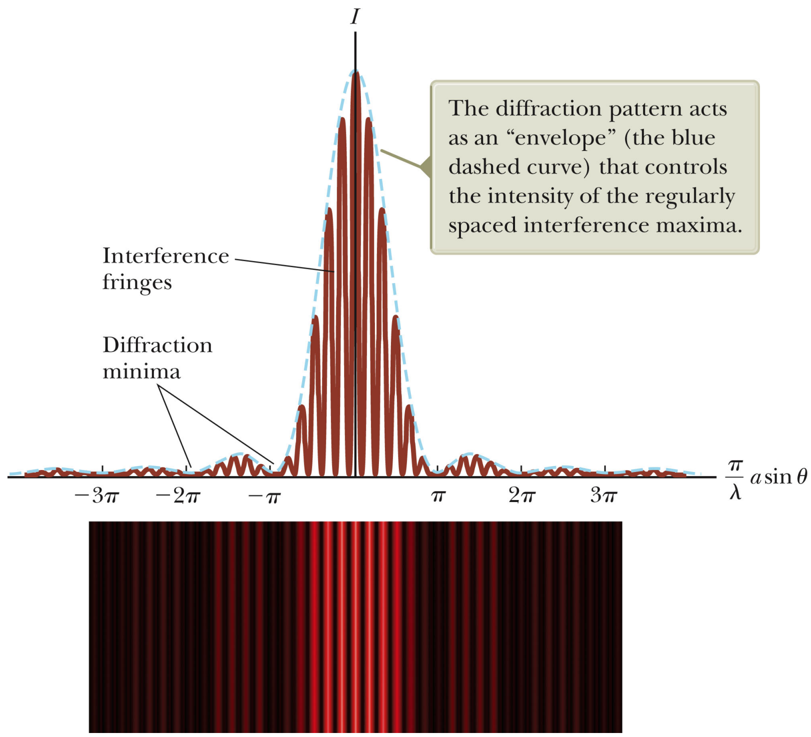


Figure 5.23: Double Slit Interference and Diffraction

Therefore, you can imagine that diffraction curve envelopes the interference, where d is interdistance bewteen two slits, a is the slit width.

$$I = I_0 \underbrace{\cos^2 \frac{\pi d \sin\theta}{\lambda}}_{\text{Interference}} \underbrace{\left[\frac{\sin\left(\frac{\pi}{\lambda} a \sin\theta\right)}{\frac{\pi}{\lambda} a \sin\theta} \right]^2}_{\text{Diffraction(envelope)}}$$

The result of the superposition of the interference and diffraction.
For double slit interference, the maximum

$$d \sin\theta = m\lambda$$

while for diffraction, the minimum

$$a \sin\theta = \lambda$$

therefore, mth $(m = \frac{d}{a})$ maximum will disappear.

18.5 Diffraction and Image Resolution

The ability of a lens to produce distinct images of two close objects is called *resolution*. Diffraction effects limit this ability.

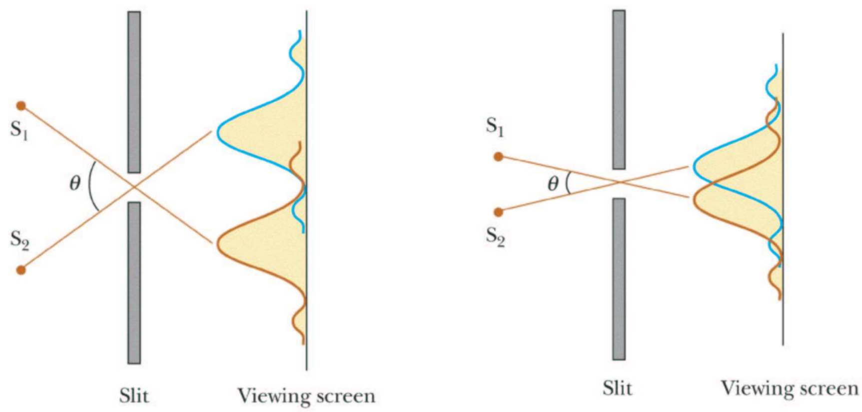


Figure 5.24: Diffraction and Image Resolution

The following showing three definition of the resolutions:

- The *Rayleigh criterion* states that the two images will be just resolvable when the centre of the diffraction disc of one falls directly on the first minimum of the diffraction pattern of the other. For a circular aperture

$$\theta = \frac{1.22\lambda}{D} = \frac{1.22}{2n\sin\theta}$$

  where $D = 2NA$ where NA is *numerical aperture*, and $NA = n\sin\theta$ and n is the index of refraction of the medium being imaged in, and θ is the half-angle subtended by the optical objective lens.

- *Abbe criterion* stipulates that an angular separation cannot be less than the ratio of the wavelength to the aperture diameter.

$$d = \frac{\lambda}{2NA}$$

- *Sparrow's resolution limit* is nearly equivalent to the theoretical diffraction limit of resolution, the wavelength of light divided by the aperture diameter, and about 20% smaller than the Rayleigh limit.

$$d = \frac{0.94\lambda}{2NA}$$

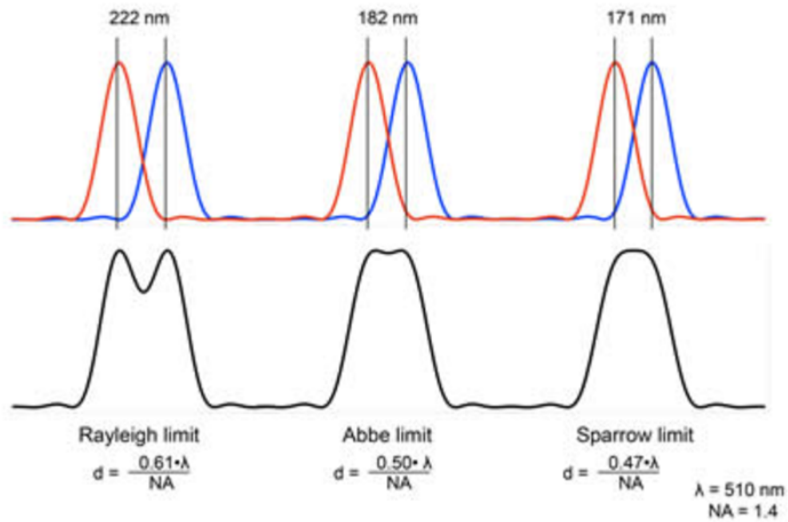


Figure 5.25: Three Criterions

18.6 Grating

Grating is multi-slits interference and diffraction:

$$I = I_0 \underbrace{\left(\frac{\sin\alpha}{\alpha}\right)^2}_{\text{Diffraction}} \underbrace{\left[\frac{\sin N\beta}{\beta}\right]^2}_{\text{Interference}}$$

where:

$$\alpha = \frac{\pi}{\lambda} a \sin\theta$$
$$\beta = \frac{\pi}{\lambda} d \sin\theta$$

and a is the slit width, and the d is the interdistance between two slits.

Interference maxima for: $d\sin\theta = n\lambda$, and diffraction minima for $a\sin\theta = m\lambda$, therefore, just like the double slit diffraction interference pattern, there are some missing order.

Apart from the central maximum, different wavelengths will form images at different angles. The wavelength can be calculated from

$$d\sin\theta_m = m\lambda$$

Images are only found at these defined angles.

If the light consists of a mixture of wavelengths, apart from the central image ($m = 0$), for each order the maxim for different wavelengths will be found at different angles. This forms a spectrum and enables the component wavelengths to be identified.

The number of slits is very large so the number of parallel rays travelling in a direction θ will also be very large. For the light with the same wavelength, except for phase angle equal zero, at all other angles complete darkness occurs.

Resolving Power

The resolving power of a diffraction grating is defined as:

$$R = \frac{\lambda}{\Delta\lambda}$$

where $\lambda = \frac{1}{2}(\lambda_1 + \lambda_2)$ and $\Delta\lambda = \lambda_2 - \lambda_1$.
This depends only on the order m of the pattern and the number of slits N.

$$R = Nm$$

Appendix A

Electromagnetism

§1 Energy and Energy Density

Electromagnetic waves carry energy and can transfer energy to the medium when propagating. The rate of flow of energy in a electromagnetic wave is described by $\mathbf{S}$ defined as: $\mathbf{S} \equiv \mathbf{E} \times \mathbf{H}$

First, consider Faraday's law and Ampere's law in matter,

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \qquad \text{Faraday's Law}$$

$$\nabla \times \mathbf{H} = \mathbf{J}_f + \frac{\partial \mathbf{D}}{\partial t} \qquad \text{Ampere's Law}$$

Consider the divergence of $\mathbf{E} \times \mathbf{H}$,

$$\begin{aligned}
\nabla \cdot (\mathbf{E} \times \mathbf{H}) &= \mathbf{H} \cdot (\nabla \times \mathbf{E}) - \mathbf{E} \cdot (\nabla \times \mathbf{H}) \\
&= -\mathbf{H} \cdot \frac{\partial \mathbf{B}}{\partial t} - \mathbf{J}_f \cdot \mathbf{E} - \mathbf{E} \cdot \frac{\partial \mathbf{D}}{\partial t} \\
&= -\left(\mathbf{H} \cdot \frac{\partial \mathbf{B}}{\partial t} + \mathbf{E} \cdot \frac{\partial \mathbf{D}}{\partial t}\right) - \mathbf{J}_f \cdot \mathbf{E}
\end{aligned}$$

and we know that in a linear materials:

$$\begin{aligned}
\mathbf{H} \cdot \frac{\partial \mathbf{B}}{\partial t} + \mathbf{E} \cdot \frac{\partial \mathbf{D}}{\partial t} &= \mu \mathbf{H} \cdot \frac{\partial \mathbf{H}}{\partial t} + \varepsilon \mathbf{E} \cdot \frac{\partial \mathbf{E}}{\partial t} \\
&= \frac{1}{2}\mu \frac{\partial}{\partial t}(\mathbf{H} \cdot \mathbf{H}) + \frac{1}{2}\varepsilon \frac{\partial}{\partial t}(\mathbf{E} \cdot \mathbf{E}) \\
&= \frac{1}{2}\frac{\partial}{\partial t}(\mathbf{B} \cdot \mathbf{H} + \mathbf{D} \cdot \mathbf{E})
\end{aligned}$$

therefore, we have that $\nabla \cdot \mathbf{S} + \frac{\partial u}{\partial t} = -\mathbf{J} \cdot \mathbf{E}$, where $\mathbf{S} \equiv \mathbf{E} \times \mathbf{H}$ is known as the *Poynting vector*, which has physical interpretation of *energy flow*, and $u = \frac{1}{2}(\mathbf{B} \cdot \mathbf{H} + \mathbf{D} \cdot \mathbf{E})$, which has the physical interpretation of the energy per unit volume - *wave energy density*[1].

[1] This equation is even valid for anisotropic media

Appendix B

Light Wave Behaviour

§2 Multiple-Beam Interference

Prove that the amplitude and irradiance of the transmitted waves are:

$$E_t = E_0 e^{i\omega t} \frac{1 - r^2}{1 - r^2 e^{-i\delta}}$$
$$I_t = I_i \frac{(1 - r^2)^2}{(1 + r^4) - 2r^2 \cos\delta}$$

Proof. With same assumption and notation, we first obtain the scalar amplitudes of the transmitted waves as $E_{1t}, E_{2t}, E_{3t}, \cdots$ are respectively, $E_0tt', E_0tr'^2t', E_0tr'^4t', \cdots$, thus the optical field at point P' are given by:

$$\begin{aligned}
E_{1t} &= E_0tt'e^{i\omega t} \\
E_{2t} &= E_0tr'^2t'e^{i(\omega t+\delta)} \\
E_{3t} &= E_0tr'^4t'e^{i(\omega t+2\delta)} \\
&\vdots \\
E_{Nt} &= E_0tr'^{2N-2}t'e^{i(\omega t+(N-1)\delta)}
\end{aligned}$$

Therefore, the resultant transmitted scalar waves are given by:

$$\begin{aligned}
E_t &= \sum_{k=1}^{N} E_{kt} \\
&= E_0tt'e^{i\omega t} + E_0tr'^2t'e^{i(\omega t+\delta)} + E_0tr'^4t'e^{i(\omega t+2\delta)} + \cdots + E_0tr'^{2N-2}t'e^{i(\omega t+(N-1)\delta)} \\
&= E_0tt'e^{i\omega t}\left(\sum_{j=0}^{N-1}\left[r'^2e^{i\delta}\right]^j\right) \\
&\overset{N\to\infty}{=} E_0tt'e^{i\omega t}\left(\frac{e^{i\delta}}{1 - r'^2e^{i\delta}}\right)
\end{aligned}$$

In case of zero absorption (no loss), $r' = r$ and $tt' = 1 - r^2$ (Stokes Relation[1]), then

$$E_t = E_0 \frac{1 - r^2}{1 - r^2 e^{i\delta}}$$

and the irradiance follows $I_t = \frac{E_t E_t^*}{2}$

$$I_t = I_i \frac{(1 - r^2)^2}{(1 + r^4) - 2r^2 \cos\delta}$$

□

[1] If you are wonder the sign, I did not flip the sign at the first place :).

Index