Optics

Haydn Cheng

February 24, 2025

Contents

1

Geometrical Optics

Imaging systems are devices that aim to produce an image of an object. For this to happen we need that all the rays of light leaving from a single point A in the object arrive at a single point A' in the image, whatever their point of entry M in the device.

Since Fermat's principle tells us that light always takes the fastest path from A to A', this means that the optical path of all rays emerging from A and arriving at A' must be the same and equal to an extrema.

In almost all of geometrical optics problems, we employ the paraxial approximation, which assumes that all light rays are nearly parallel to the optical axis, which implies that the angles formed by the light rays with the optical axis is always small enough that the small angle approximations are valid.

1.1 Fundamentals

1.1.1 Sign Conventions

When discussing spherical mirrors and spherical surfaces, we will see that we can subdivide each of them cases. However, only one equation is needed for one case. This is because we adopt the following sign convention to embed the negative signs into distances:

1. The object distance u is positive if the object point is on the same side as the incident light.

2. The image distance v is positive if the image point is on the same side as the outgoing light (this implies a real object).

3. The radius R is positive if the center of curvature is on the same side as the outgoing light.

An easy way to remember to remember the convention of the object distance and the image distance is that they are positive if the object point and the image point are where they are supposed to be at.

1.1.2 Spherical Reflecting Surfaces

Referring to fig. 1.1,[1] there are two possibilities in which light can reflect at a spherical mirror.

$$① : \tan\alpha = \alpha = \frac{h}{u} \quad,\quad \tan\beta = \beta = 2\theta + \alpha = \frac{h}{v} \quad,\quad \sin\gamma = \gamma = \beta - \theta = \frac{h}{R} = \theta + \alpha \implies \frac{1}{u} + \frac{1}{v} = \frac{2}{R}$$

$$② : \tan\alpha' = \alpha' = \frac{h}{u'} \quad,\quad \tan\beta' = \beta' = 2\theta' - \alpha' = \frac{h}{v'} \quad,\quad \sin\gamma' = \gamma' = \beta' - \theta' = \frac{h}{R} = \theta' - \alpha' \implies \frac{1}{u} - \frac{1}{v'} = -\frac{2}{R}$$
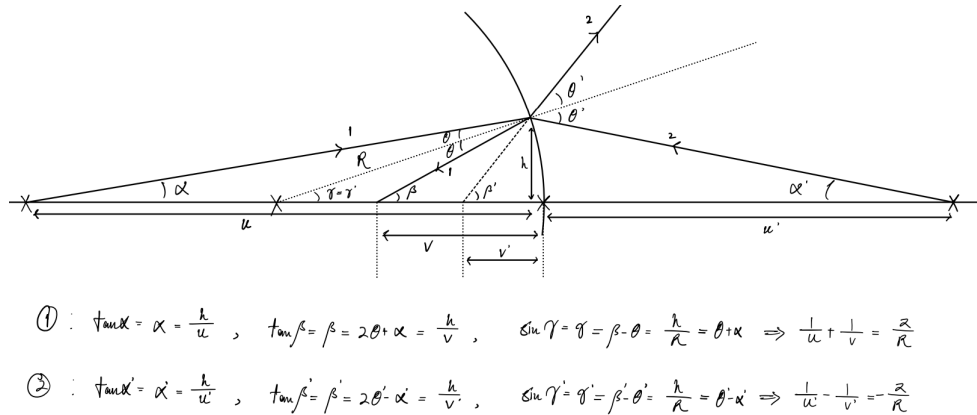
Figure 1.1

The first possibility, labelled as path 1, is that the object (at this stage we assume that it is real) is situated inside the sphere, which creates a real object. The second possibility, labelled as path 2, is that the object is situated outside the sphere, which would create a virtual object.

However, using the sign convention defined above, we can generalized the two equations governing the two cases as

$$\frac{1}{u} + \frac{1}{v} = \frac{2}{R} \equiv \frac{1}{f}. \tag{1.1}$$

The quanitity $\frac{1}{u} + \frac{1}{v}$ of a imaging system is a constant, then the imaging system will focus

The case where the object distance u is negative, which only occurs when there are more than two imaging systems, can be regarded as if u is real but v and R are negative, which corresponds to switching between the first case and the second case mentioned above. Therefore, it is equivalent as if there is a real object situated at the opposite side of the mirror with the same distance from the mirror. This implies that we can simply treat the image point of the first imaging system as the object point of the second imaging system.

1.1.3 Spherical Refracting Surfaces

Referring to fig. 1.2,[2] there are two possibilities[3] in which light can refract at the spherical surfaces.

---

[1]Note that the arrows drawn in the figure may be reversed due to light's reversibility.

[2]See footnote under section 1.1.2.

[3]There are actually four, the other two being if the object is situated at the opposite side of the spherical surface compared to what is shown above. But we have already shown in section 1.1.2 that we can simply add a negative sign to the radius to account for these cases, so we focus our disccusion to the two cases both having their object point inside the sphere.
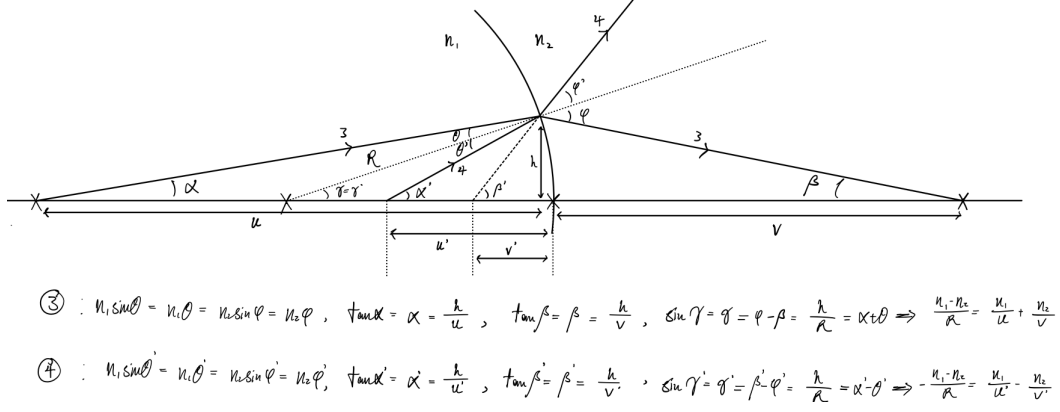
③ : $n_1\sin\theta = n_1\theta = n_2\sin\varphi = n_2\varphi$, $\tan\alpha = \alpha = \frac{h}{u}$, $\tan\beta = \beta = \frac{h}{v}$, $\sin\gamma = \gamma = \varphi - \beta = \frac{h}{R} = \alpha + \theta \Rightarrow \frac{n_1-n_2}{R} = \frac{n_1}{u} + \frac{n_2}{v}$

④ : $n_1\sin\theta' = n_1\theta' = n_2\sin\varphi' = n_2\varphi'$, $\tan\alpha' = \alpha' = \frac{h}{u'}$, $\tan\beta' = \beta' = \frac{h}{v'}$, $\sin\gamma' = \gamma' = \beta' - \varphi' = \frac{h}{R} = \alpha' - \theta' \Rightarrow -\frac{n_1-n_2}{R} = \frac{n_1}{u'} - \frac{n_2}{v'}$

Figure 1.2

The first possibility, labelled as path 3, is that the object (at this stage we assume that it is real) is situated behind the center of curvature, which creates a real object. The second possibility, labelled as path 4, is that the object is situated closer to the surface than the center of curvature, which creates a virtual object.

However, using the sign convention defined above, we can generalized the equations governing the cases as

$$\frac{n_1}{u} + \frac{n_2}{v} = \frac{n_1 - n_2}{R}. \tag{1.2}$$

1.2   Thin Lens

From the two fundamental building blocks, we can construct a lens with two sphereical surfaces. Treating the image of the first surface as the object of the second surface, we can derive the lensmaker's equation

$$\frac{1}{f} = (n - 1)\left(\frac{1}{R_1} - \frac{1}{R_2}\right). \tag{1.3}$$

Suppose our thin lens has thickness $d(z) \ll u, v$, which depends on height z. The paraxial approximation tells us that $z \ll u, v$ and the path length of light travelled inside the lens is simply $d(z)$. Therefore, using Fermat's principle, we have

$$L(z) = \sqrt{\left(u - \frac{d(z)}{2}\right)^2 + z^2} + nd(z) + \sqrt{\left(v - \frac{d(z)}{2}\right)^2 + z^2}$$

$$\approx \left(u - \frac{d(z)}{2}\right)\left(1 + \frac{z^2}{2\left(u - \frac{d(z)}{2}\right)^2}\right) + d(z) + \left(v - \frac{d(z)}{2}\right)\left(1 + \frac{z^2}{2\left(v - \frac{d(z)}{2}\right)^2}\right)$$

$$= u + v + (n-1)d(z) + \frac{z^2}{2}\left(\frac{1}{\left(u - \frac{d(z)}{2}\right)} + \frac{1}{\left(v - \frac{d(z)}{2}\right)}\right)$$

$$\approx u + v + (n-1)d(z) + \frac{z^2}{2}(\frac{1}{u} + \frac{1}{v}) = \text{constant.}$$

$$(1.4)$$

Upon differentiation,

$$\frac{d}{dz}d(z) = -\frac{z}{n-1}(\frac{1}{u} + \frac{1}{v})$$

$$\implies d(z) = d(0) - \frac{z^2}{2(n-1)}\left(\frac{1}{u} + \frac{1}{v}\right).$$

$$(1.5)$$

For the lens to work with arbitrary u, v, we require that

$$\frac{1}{u} + \frac{1}{v} \equiv \frac{1}{f} = \text{constant.}$$

$$(1.6)$$

1.3   Refracting Telescope

1.4   Copmound Microscope

2
Interference and diffraction

To begin, we must be noted that interference and diffraction are not fundamentally different processes, they are both interaction between light as a manifestation of its wave properties. In most contexts, interfrences refer to the interaction between finite number of light waves while diffraciton refer to infinite number of light waves.

For the interaction to be visible, the light waves must be coherent, *i.e.,* has a constant phase difference which is indepdent of time, *i.e.,* has the same frequency. This is acheived either by genearting them with the same plane wave (*i.e.,* spherical wave with distance from the origin far enough), or by a laser. If this is not the case, then the phase differences between the electric fields will vary randomly and average out to no visible pattern.

Since the sources are usually generated by the same plane wave, the electric fields can be assumed to be parallel to each other. Even if this is not the case, unpolarized light (*e.g.,* natural light) can still interfere as long as the lights are coherent, so that the effect of unparallel lights only decrease the intensity of the interference pattern, but the shape would remain unchanged.

For genearl consideration, we consider the interference of two electric fields

$$\mathbf{E}_1 = \mathbf{E}_{01}\cos(\mathbf{k}_1 \cdot \mathbf{r} - \omega t + \epsilon_1) \ \ \text{and} \ \ \mathbf{E}_2 = \mathbf{E}_{02}\cos(\mathbf{k}_2 \cdot \mathbf{r} - \omega t + \epsilon_2). \tag{2.1}$$

The resultant intensity is

$$\begin{aligned} I &= \langle (\mathbf{E}_1 + \mathbf{E}_2) \cdot (\mathbf{E}_1 + \mathbf{E}_2) \rangle = \langle |\mathbf{E}_1|^2 + |\mathbf{E}_2|^2 + 2\mathbf{E}_1 \cdot \mathbf{E}_2 \rangle \\ &= I_1 + I_2 + (\mathbf{E}_{01} \cdot \mathbf{E}_{02}) \langle \cos\left((\mathbf{k}_2 - \mathbf{k}_1) \cdot \mathbf{r} + \varphi_2 - \varphi_1\right) \rangle. \end{aligned} \tag{2.2}$$

When $\mathbf{E}_{01} \parallel \mathbf{E}_{02}$ and $E_0 \equiv E_{01} = E_{02}$, we have

$$I = I_1 + I_2 + \sqrt{I_1 I_2}\langle \cos\alpha \rangle = 4I_0 \cos^2\left(\frac{\alpha}{2}\right) \tag{2.3}$$

where α is the phase difference between two electric fields and at the last step we assumed it to be constant in time.

2.1 Far-field Limit

The far-field limit in optics is the region where the diffracted wavefronts from an aperture or slit can be approximated as planar, which is typically met when $D \gg a^2/\lambda$, where D is the distance from the aperature or slit to the screen and a is the charateristic size of the aperature or the slit.[1] This can be seen from eq. (2.8) since the width of the central maximum is given by $W = 2D\lambda/a$, and we require it to be much greater than a to acheive far-field limit.

As a result, for a certain point on the screen, all the lights arrive approximately in parallel and the electric field has the same amplitude $E_0(\theta) = E_0(0)/\sqrt{\cos\theta}$ for cylindrical wave (or 2D spherical wave),[2] usually we will further neglect this dependency and have $E(\theta) = E(0) \equiv E_0$. Note also that we are not using the electric fields at the aperature or the slit, since due to the singularity at the source one could not obatin the realtionship between the electric fields at the screen and at the source by simple means, except for the simple plane wave case where the electric field amplitude remain constnat througout.

In practice, far-field limit is usually acheived by putting a convex lens between the aperature and the screen, with the screen being the focal plane of the lens. This has the effect of bringing what we would observe at infinity (if there is no lens) closer to the focal plane of the lens. Essentially, what we want to observe is the interference of the parallel beams (at infinity), and the lens change the point of convergence from infinity to a point on the focal plane for easier observation. Illustrated in fig. 2.1, we can see that the parallel rays before the lens is what hit the point P' (not shown in the figure) in an ideal screen situated at infinity. Putting the lens causes the rays to converge to a closer point P while maintaining the phase difference to be unchanged thanks to Fermat's principle.
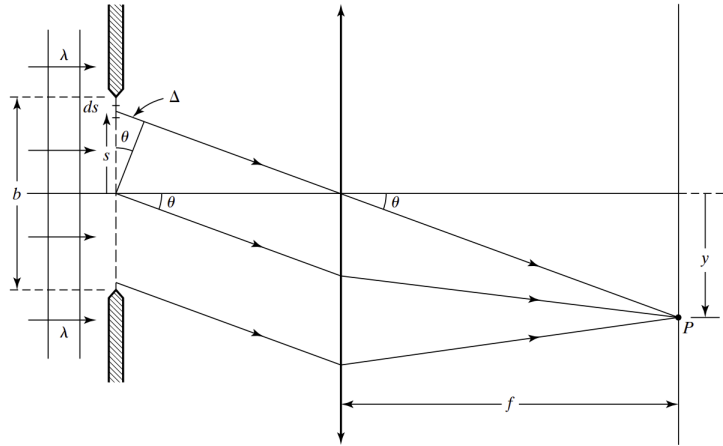

Figure 2.1

---

[1]Aperature and slit are not fundamentally different objects and slit is simply a special extreme case of an aperature, just as interference and diffraction are not fundamentally different processes, diffraction is simply a special name for interference when light passes through an aperature.

[2]Although Huygen's principle says that the secondary wavelets are spherical in nature, the wavefronts of the spherical waves created by a cylindrical source (the slit) is cylindrical in nature, thus the $1/\sqrt{r}$ decaying factor instead of $1/r$. But in doing so we must replace the electric field per unit area to the electric field per unit lenght. Equivalently, the changing of $1/r$ to $1/\sqrt{r}$ is due to one of the two integrations needed in a double integration.

2.1.1 Interference

Double Slits

To find the interference pattern $I(\theta)$, we have

$$E(\theta) = E_0 e^{i(kr_1-\omega t)} + E_0 e^{i(kr_2-\omega t)} = 2E_0 \left(e^{i(k(r_1-r_2)/2)} + e^{-i(k(r_1-r_2)/2)} \right) e^{i(k(r_1+r_2)/2-\omega t)}, \tag{2.4}$$

where E_0 is the electric field produced by the single slit at $x = 0$ on the screen. We have also assumed the parallel of the electric fields, thus used scalar addition, and we have changed $\mathbf{k} \cdot \mathbf{r}$ to just kr since $\mathbf{k}$ and $\mathbf{r}$ are parallel if we use the coordinate systems where the origins are located at the two slits when doing the two dot products. Note that we do not have to use the same coordinate systems throughout, as long as the same coordinate system is used to evaluate one dot product.

$$I(\theta) = \langle |E(y)|^2 \rangle = 4E_0^2 \cos^2 \left(\frac{kd\sin\theta}{2} \right) = 4I_0 \cos^2 \left(\frac{kdx}{2D} \right) = 4I_0 \cos^2 \left(\frac{\alpha}{2} \right), \tag{2.5}$$

where $I_0 \equiv E_0^2$, but $I(0) = (2E_0)^2 = 4I_0$, as expected. The interference pattern are evenly spaced fringes with separation $\Delta x = \lambda D/d$, shown in fig. 2.2.
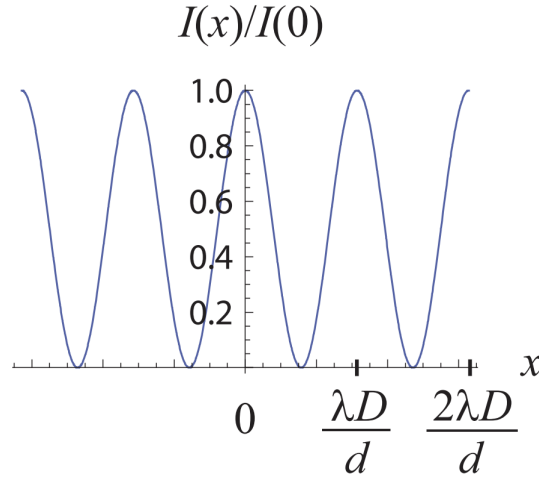

Figure 2.2

N-Slits

We start with finding the electric field

$$E(\theta) = E_0 e^{i(kr_1-\omega t)} \sum_{n=1}^{N} e^{i(k(n-1)d\sin\theta)} = E_0 e^{i(kr_1-\omega t)} \left(\frac{Z^N - 1}{Z - 1} \right)$$

$$= E_0 e^{i(kr_1-\omega t)} \frac{Z^{N/2}(Z^{N/2} - Z^{-N/2})}{Z^{1/2}(Z^{1/2} - Z^{-1/2})} = E_0 e^{i(kr_1-\omega t+(N-1)kd\sin\theta/2)} \left(\frac{\sin\left(\frac{Nkd\sin\theta}{2} \right)}{\sin\left(\frac{kd\sin\theta}{2} \right)} \right), \tag{2.6}$$

where $Z \equiv e^{ikd\sin\theta} = e^{i\alpha}$. So the intensity is

$$I(\theta) = \langle |E(x)|^2 \rangle = I_0 \left(\frac{\sin\left(\frac{Nkd\sin\theta}{2}\right)}{\sin\left(\frac{kd\sin\theta}{2}\right)} \right)^2 = I_0 \left(\frac{\sin\left(\frac{Nkdx}{2D}\right)}{\sin\left(\frac{kdx}{2D}\right)} \right)^2 = I_0 \left(\frac{\sin\left(N\alpha/2\right)}{\sin\left(\alpha/2\right)} \right)^2. \quad (2.7)$$

Here $I(0) = (NE_0)^2 = N^2 I_0$, which can be verified from the above equation by taking the limit as $\alpha \to 0$. Note that although $\theta \ll 1$, kd is not necessarily much smaller than 1, so the expression can not be simplified further.

Therefore, the minima occurs when the numerator is zero while the denominator is not, *i.e.*, $N\alpha/2 = m\pi$, for $m \neq nN$, while the global maximum occurs when both the numerator and the denomiator are zero, *i.e.*, $\alpha/2 = n\pi$, which is when all the lights are in phase. The local minima can be found by differentiation, which gives $\tan(N\alpha/2) = N\tan(\alpha/2)$. For large N, the solutions for α genearlly gives symmetric solutions, and the intensity is the smallest for the local maximum centered between the global maxima, as shown in fig. 2.3a for the $N = 8$.
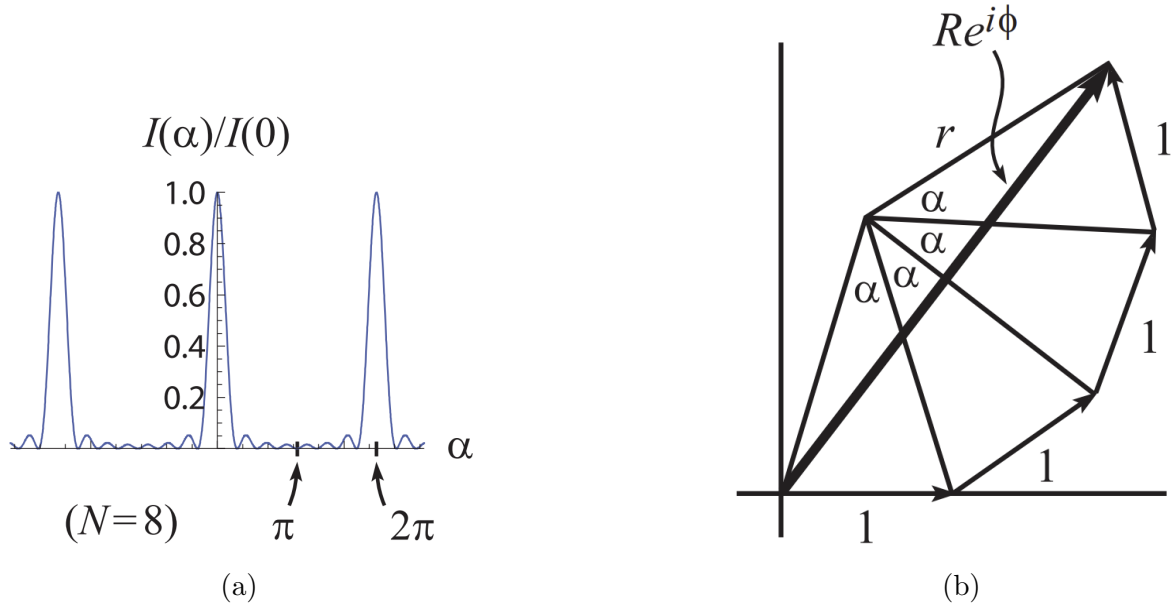

Figure 2.3

To visualize the interference we can make use of the phasor diagram as shown in fig. 2.3b, where the phase difference between each successive electric field is α and the length of each electric field vector is 1. Consider the two types of isosceles triangles in the figure, we have $R = 2r\sin(N\alpha/2)$ and $1 = 2r\sin(\alpha/2)$, which implies $R = \sin(N\alpha/2)/\sin(\alpha/2)$.

This diagram also visualize the pairwise cancellation of electric field when minima occurs at $N\alpha/2 = m\pi$ as shown in fig. 2.4, where the figures are drawn slightly off to make it easier to see the cancellation. The maxima occur approximately between the minima since the overall phase difference is now integer multiple of π instead of 2π, so the resultant electric field vector has the longest length, instead of being cancelled out, as shown in fig. 2.5 for $N = 50$. The maxmima don't occur exactly at the diameters since the circle shrinks as the vectors wrap around futher as α increases so there are competing effect, but this effect is small when N is large as the circle circle hardly changes sizes.

9

$(\alpha=\pi/4)$ $(\alpha=\pi/2)$ $(\alpha=3\pi/4)$ $(\alpha=\pi)$

$(\alpha=5\pi/4)$ $(\alpha=3\pi/2)$ $(\alpha=7\pi/4)$

Figure 2.4

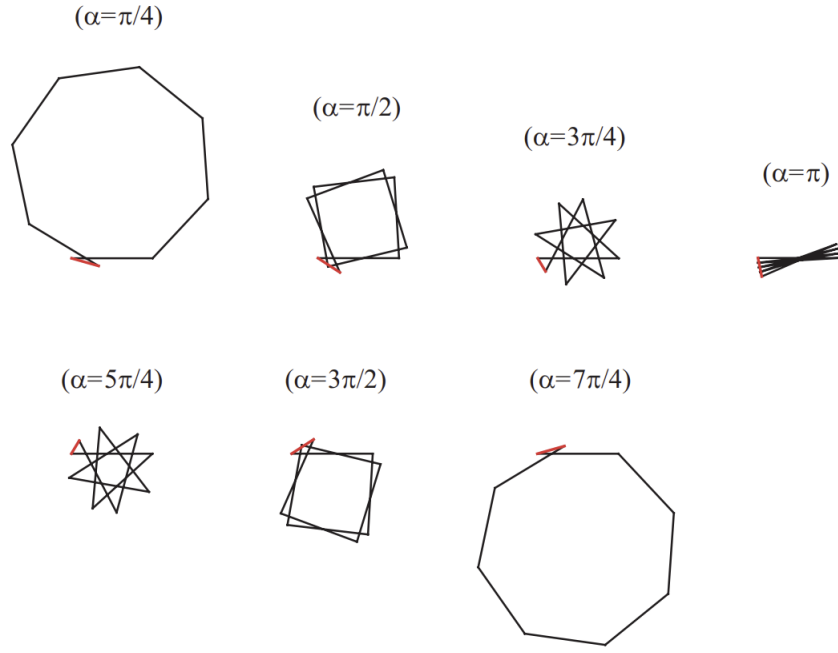$(N=50)$

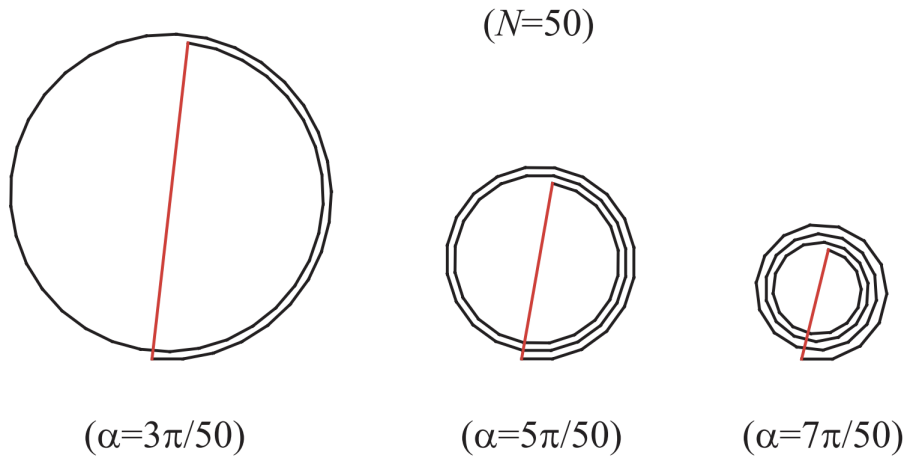$(\alpha=3\pi/50)$ $(\alpha=5\pi/50)$ $(\alpha=7\pi/50)$

Figure 2.5

Diffraction grating is essentially N slits interference with $N \gg 1$, since then the intensity of the local maximum is negligible, since it scales with $1/N^2$, so the maxima are only observed when $\alpha = 2n\pi$, or $\lambda = d\sin\theta$.

2.1.2 Diffraction

Single Slit Diffraction

If the slit width a is not negligible, then for a single slit the intensity can be found by treating every points in the slit as a single slit with slit separation $d = a/N \to 0$ and number of slits $N \to \infty$, so from eq. (2.7)

$$I(\theta) = \frac{I(0)}{N^2}\left(\frac{\sin\left(\frac{Nkdx}{2D}\right)}{\sin\left(\frac{kdx}{2D}\right)}\right)^2 = I(0)\left(\frac{\sin\left(\beta/2\right)}{\beta/2}\right)^2, \quad \beta = ka\sin\theta, \tag{2.8}$$

which as a sine function with a $1/x$ envelope, as shown in fig. 2.6a. Therefore, minima occurs when $\beta/2 = m\pi$, or $\lambda = ma\sin\theta \approx ma\theta$, which is when the electric fields from the top and bottom of the slits are in phase, since then if we divide the slits into two halfs (for $m = 1$) (or four quarters for $m = 2$ etc.), we can always find a corresponding electric field is out of phase for any electric field, achieving pairwise cancellation. The maxima condition is now $\tan(\beta/2) = \beta/2$, which occur roughly halfway between the zeros.

If $a \ll \lambda$, then we effectively have a line source since no matter how the lights tilt no significant path difference can be acheived so no diffraction pattern. If $a \gg \lambda$, then even the slightest tilt of the beam would lead to a path differenece comparable to λ, so the diffraction pattern is very narrow. In any rate, the rule of thumb is that $\theta \approx \lambda/a$ for the first minima.

More rigorously, we can find the intensity profile by summing the electric field by the infinite slits with width dx and taken the averaged modulus squared to find the intensity, we start with

$$E(\theta) = \int_{-\frac{a}{2}}^{\frac{a}{2}}(E_0'dx)e^{-ikx\sin\theta} = \frac{E_0'}{-ik\sin\theta}\left(e^{-ik(a/2)\sin\theta} - e^{ik(a/2)\sin\theta}\right) = E_0'a\left(\frac{\sin\left(\beta/2\right)}{\beta/2}\right), \tag{2.9}$$

where E_0' is the electric field per unit length across the slit produced at $x = 0$ on the screen, averaging the modulus squared give the same result as above.

In fact the electric field can be obtained by the fourier transforming the transimitivity function $T(x)$

$$E(\theta) = E_0'\int_{-\infty}^{+\infty}T(y)e^{-ikx\sin\theta}dx = \tilde{T}(k\sin\theta). \tag{2.10}$$

From the theory of Fourier transform, this means that if T has a large componnet with spatial frequency $k\sin\theta$, then the electric field at angle θ will be larger, which is true as when the spatial frequency of T is $k\sin\theta = 2\pi/(\lambda/\sin\theta)$, then from fig. 2.7a we see that the light generally constructively interfere with one another.

The above results can also be seen from fig. 2.6b, where the digaram is constructed by noting that the phase diffrence between the elctric fields at the top and the bottom of the slit is given by β, then the ratio of intensity is given by the ratio of the length between the straight line and the arc, which gives $I(x)/I(0) = (2r\sin(\beta/2)/r\beta)^2$, which is the same as the above equation.

The minima condition $\beta = 2\pi$ corresponds to the fact that the arc in fig. 2.6b turns into a full circle, so the resultant electric field is zero.
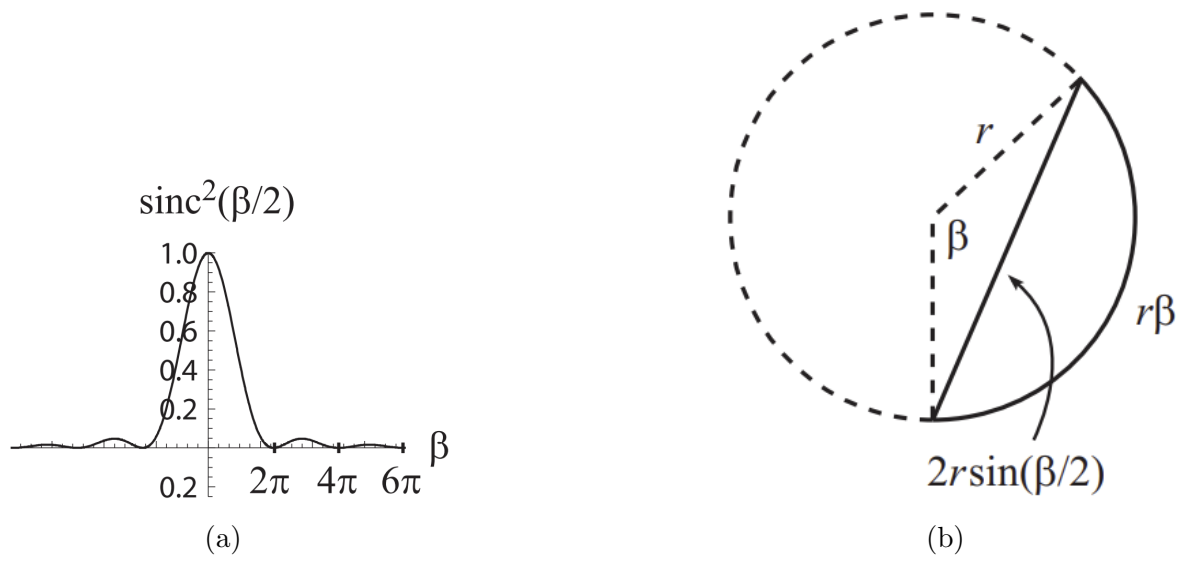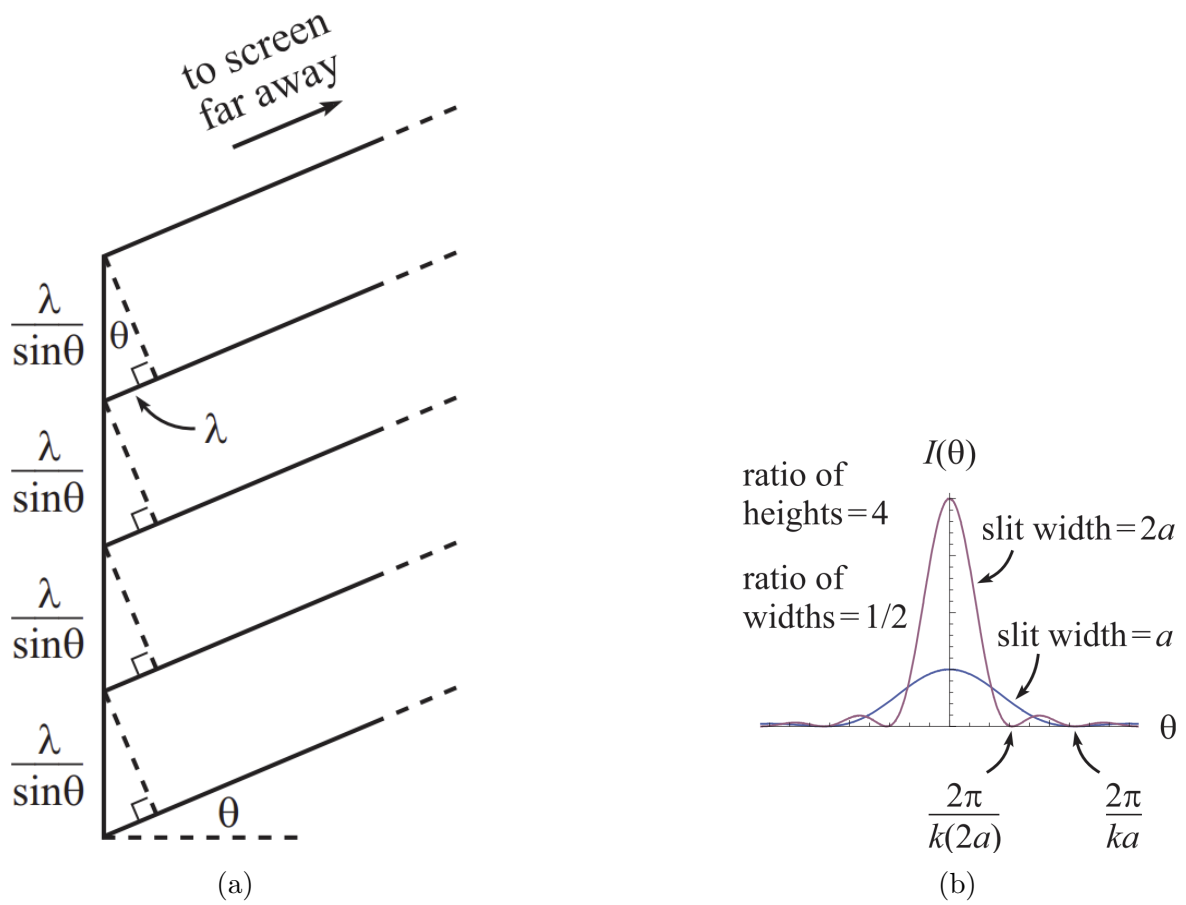
(a)

(b)

Figure 2.6

(a)

(b)

Figure 2.7

Example: Four Times the Light?

Question: It can be easily proved that $I(0) \propto a^2$. This means that if we double a, then $I(0)$ increases by a factor of 4. Does this make sense and does it mean that if we double the width of the slit, then 4 times as much light makes it through?

Solution: The answers are yes and no, respectively. The intensity increases by a factor of 4 since this is what we get mathematically but the energy does not increases by a factor of 4 since intensity is power divided by area, so energy is the area under graph of the intensity profile. So while $I(0)$ quadruple, since the intensity profile is thinner as a gets larger, exactly by a factor of $1/2$, the energy only doubles. The situation is illustrated in fig. 2.7b.

Example: Increasing or Decresing Intensity?

Question: As we make a larger, will the intensity at a particular point on the screen that is reasonable distance off to the side increases or decrease?

Solution: On one hand, increasing a will aloow more light through the slit, so the intensity should increase. But on the other hand, increasing a will make the diffraction pattern narrower, so the intensity should decrease. In turns out that these two effects exactly cancel, so the intensity remained unchanged. For the fact that $I(\theta) \propto I(0) \sin^2(ka\theta/2)/k^2a^2\theta^2$, but since $I(0) \propto a^2$, thaking the average over a few oscillation of θ, we have $\langle I(\theta) \rangle \propto 1/k^2\theta^2$, independent of a.

N-Slits Diffraction

By integrating the contribution of every single points, we obtain the intensity of N-slits diffraction as

$$I(\theta) = I(0) \left(\frac{\sin \alpha/2}{N\alpha/2} \frac{\sin \beta/2}{\beta/2} \right)^2, \tag{2.11}$$

simply the product of the results for the two separate cases we have discussed, so the diffraction effect simply acts as a modulating envelope for the N-slits interference, as shown in fig. 2.8.
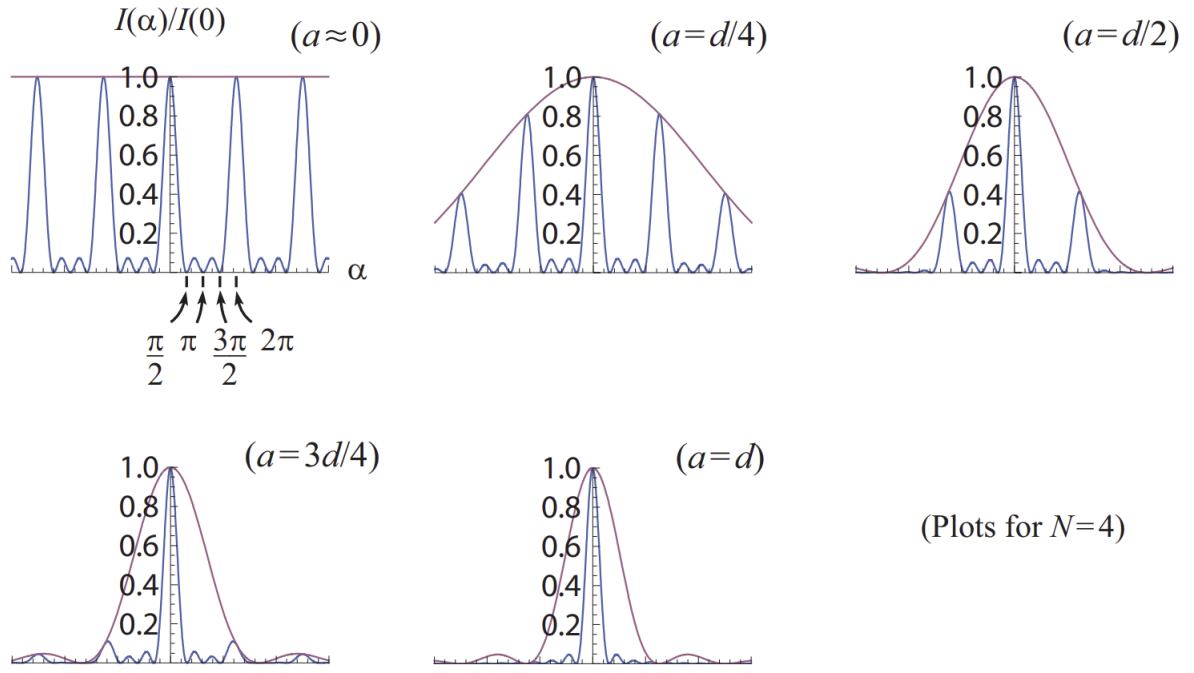
Figure 2.8

Circular Aperature Diffraction

Using the geometry shown in fig. 2.9, the electric field is

$$E(\theta) = \int (E_0'' dA)e^{iks\sin\theta} = 2E_0'' \int_{-R}^{R} e^{iks\sin\theta}\sqrt{R^2 - s^2}ds = \frac{2E_0''\pi J_1(\gamma)}{k\gamma}, \quad \gamma = kR\sin\theta,$$

(2.12)

where E_0'' is the electric field per unit area over the aperature produced at $x = 0$ on the screen and $J_1(\gamma)$ is the first-order Bessel function of the first kind, with the first order term equals to $\gamma/2$.
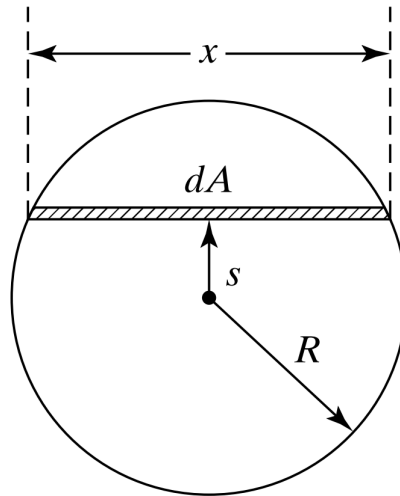

Figure 2.9

The first minimum occurs when $\gamma = 3.832$, or $D \sin \theta = 1.22\lambda$. Rayleigh's criterion for just-resolvable images requires that the central maximum of one image coincedes with the first minimum of the another image, *i.e.,* $\Delta\theta = 1.22\lambda/D$, where D is usually the diameter of the lens if we are viewing two starts through a microscope.

2.2 Optical Interferometry

Before extending our discussion beyond Far-field limit, we will devote this secion to the application of interference of light, maining with optical interferoemters.

2.2.1 Michelson Interferometer

The set up and the equivalent optical system are shown in fig. 2.10, where M_1' is constructed by imaging M_1 with respect to the mirror on the beam splitter BS. Therefore the phase difference between the two beams from the source S is $\delta = 2kd\cos\theta + \pi/2$, and the intensity profile is given by $I = 4I_0 \cos^2(\delta/2)$. So the condition for dark fringes is given by $2d\cos\theta = m\lambda$, where the order of the central dark fringe is given by $m_{\max} = 2d/\lambda$. If we want we could invert the ordering of the fringes by defining $p = m_{\max} - m$, so that the order increase as we get further away from the center, but usually one would not bother to do this as the ordering is arbitrary anyways.

As we decrease d by moving M_1 away, so that M_1' is closer to M_2 (here we assume initially M_1' is in front of M_2 as shown in fig. 2.10), $m_{\max}$ becomes smaller and would goes to zero as $d = 0$, where there is no interference pattern, but then it would gets larger again when M_1 is move further away. In fact, we almost always only concern about the change in m (this is also why the absolute value of m is not important), given by $\Delta m = 2\Delta d/\lambda$.

One application of the Michelson interferometer is to measure the difference in wavelength between two closed spaced components of a spectral line, such as the two yellow sodium lines. Suppose will adjust d to d_1 such that the two circular interference patterns produced by the two lines coincide with each other at $m_1 = m_1' + N$, then as we adjust d we will see a field of uniform brightness when $m_{1/2} = m_{1/2}' + N + 1/2$, and at d_2 we have $m_2 = m_2' + N + 1$, using $m_{\max} = 2d/\lambda$, we have $\Delta\lambda = \lambda^2/2\Delta d$.

Twyman-Green interferometer, shown in fig. 2.11 is a modified version of the Michelson interferometer, which when gives no interference pattern when nothing is placed between it, thus it can be used to test the quality of optical instruments such as prism P or another lens, in which case the mirror M_1 has to be replaced by a spherical mirror that can reflect the refracted rays back along themselves. If there are imperfections in the optical system, then distortions will appear in the interference pattern.

Furthermore, the modified Michelson interfermoeter shown in fig. 2.12 can be used to measure the thickness d of the thin film. Monochromatic light channeled from a light source LS through a fiber optic light pipe LP to a right-angled beam-splitting prism BS, which transmits one beam to a flat mirror M and the other to the film surface. After reflection, each is transmitted by the beam splitter into a microscope MS, where they are allowed to interfere. For normal incidence, bright fringes satisfy $2nt + \pi = m\lambda$, when the film is not present, we have $2n\Delta t = 2d = \lambda\Delta m = (\lambda\Delta x)/x$, which implies $d = (\lambda\Delta x)/2x$.
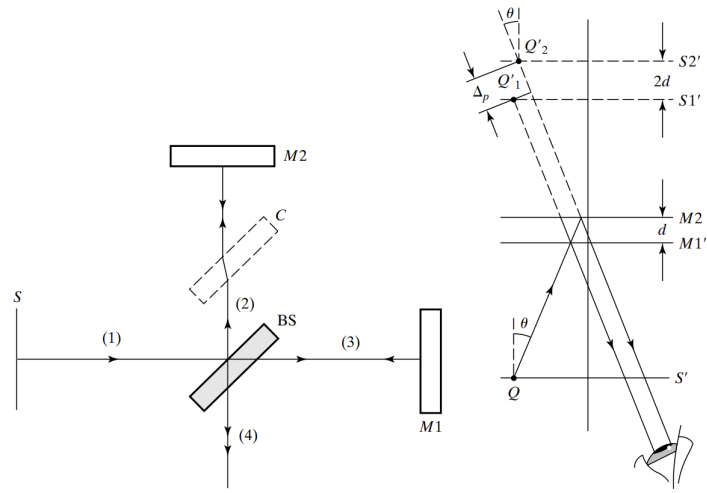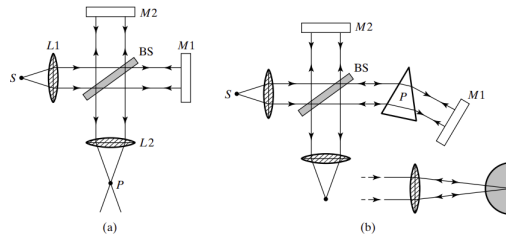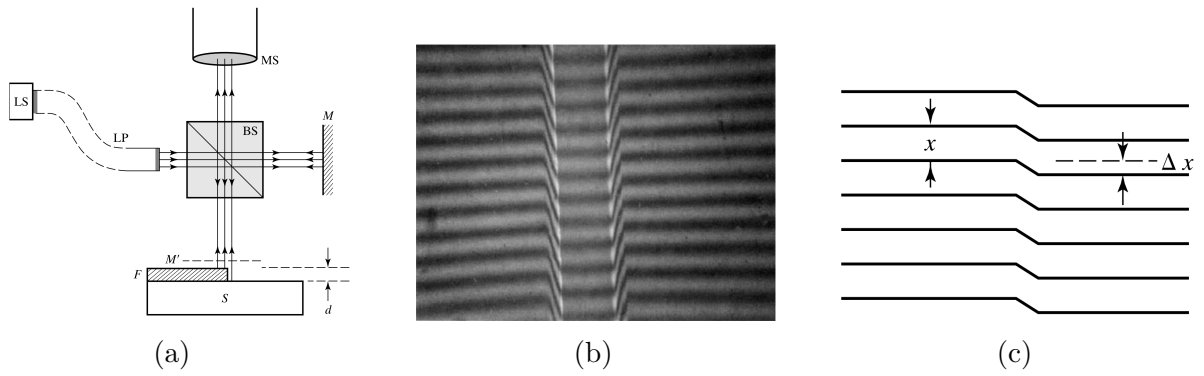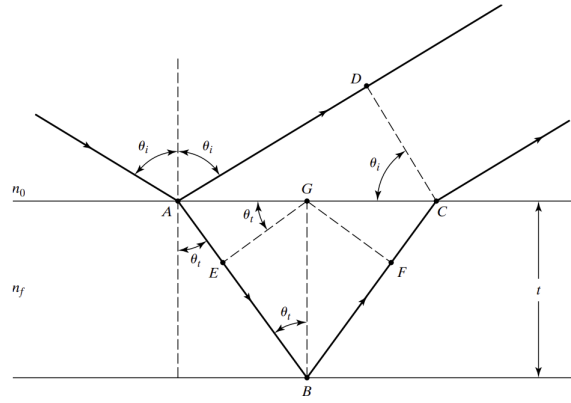
Figure 2.10

Figure 2.11

Figure 2.12

Figure 2.13

2.2.2 Fabry-Perot Interferometer

The two examples below lay the foundation when analysing the fabry-perot interfermoeter.

Example: Thin-film Interference.

Question: Find the condition for constructive interference for the situation shown in fig. 2.13.

Solution: The path difference is

$$\Delta = n_f(AB + BC) - n_0(AD) = \frac{2n_f t}{\cos\theta_t} + \frac{2n_0 t \cos\theta_i}{\sin\theta_t} = 2n_f t \cos\theta_t, \qquad (2.13)$$

where we have used the Snell's law to eliminate θ_i.
So the constructive interference condition is $\Delta + \lambda/2 = m\lambda$.

Example: Multiple-Beam Interference in a Parallel Plate.

Question: Refer to fig. 2.14, find the condition for constructive interference in the transmitted beam.

Solution: At the first mirror, the beam (1) reflects at the mirror-medium interface, thus will not acquire a π phase shift, while the beams (2), (3), *etc.* reflects at the medium-mirror interface, thus will acquire a π phase shift, so we have $r = -r'$.
On the other hand, Stokes' relation gives $tt' = 1 - r^2 = 1 - r'^2$.
Letting $\delta = 2kn_f t \cos\theta_t$ to be the phase difference beteween each successive reflected beam, we have

$$E_1 = (rE_0)e^{i\omega t}, \quad E_2 = (tt'r'E_0)e^{i(\omega t - \delta)}, \quad E_3 = (tt'r'^3 E_0)e^{i(\omega t - 2\delta)} \quad etc., \quad (2.14)$$

where the negative sign in the first reflected beam comes frmo the π phase shift of reflection. Summing the reflected lights, we have

17

$$E_R = \sum_{N=1}^{\infty} E_N = rE_0 e^{i\omega t} + \sum_{N=2}^{\infty} tt' E_0 r'^{(2N-3)} e^{i(\omega t - (N-1)\delta)}$$

$$= E_0 e^{i\omega t} \left(r + \frac{(1-r^2)r' e^{-i\delta}}{1 - r'^2 e^{-i\delta}} \right) = E_0 e^{i\omega t} \left(\frac{r(1 - e^{-i\delta})}{1 - r^2 e^{-i\delta}} \right), \tag{2.15}$$

so the intensity is

$$I_R = \langle |E_R|^2 \rangle = E_R^* E_R = E_0^2 r^2 \left(\frac{e^{-i\omega t}(1 - e^{i\delta})}{1 - r^2 e^{i\delta}} \right) \left(\frac{e^{i\omega t}(1 - e^{-i\delta})}{1 - r^2 e^{-i\delta}} \right)$$

$$= \left(\frac{2r^2(1 - \cos\delta)}{1 + r^4 - 2r^2 \cos\delta} \right) I_i, \quad I_I = E_0^2. \tag{2.16}$$

Similar treatment of the transmitted beams leads to the resultant transmitted intensity

$$I_T = \left(\frac{(1 - r^2)^2}{1 + r^4 - 2r^2 \cos\delta} \right) I_I \implies I_R + I_T = I_I. \tag{2.17}$$

A maximum in transmitted intensity occurs when $\cos\delta = 1$, or $k\delta = 2n_f t \cos\theta_t = m\lambda$, which is when the second reflected beam and all subsequent reflected beams are in phase with one another but exactly out of phase with the first reflected beam due to the π phase shift. Substituting into the reflected and transmitted intensity we found that the reflected beam intensity vanishes and the transmitted beam intensity is the same as the incidence beam.
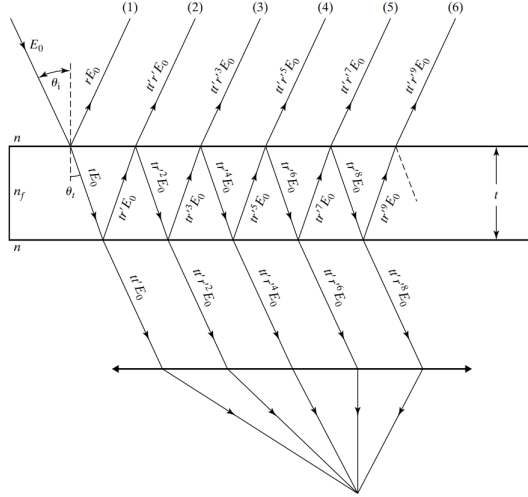

Figure 2.14

2.3 Near-field Limit

In Near-field limit, we can no longer assume that $E(\theta) = E_0$ for all points from the sources, we have to take into account the $1/\sqrt{r}$ dependence in the amplitudes. Moreover, the pathlengths does not thake the nice form $d\sin\theta$, but is more complicated.