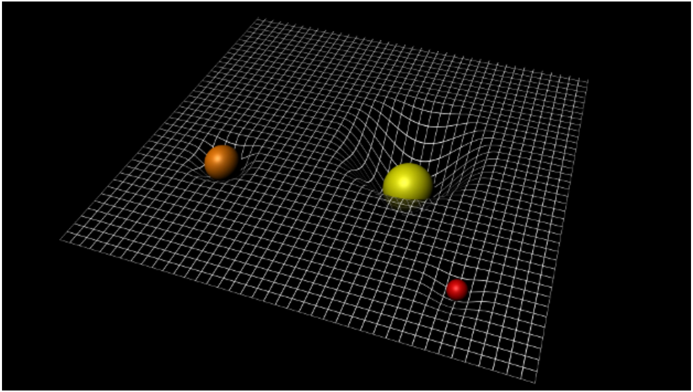
Interference

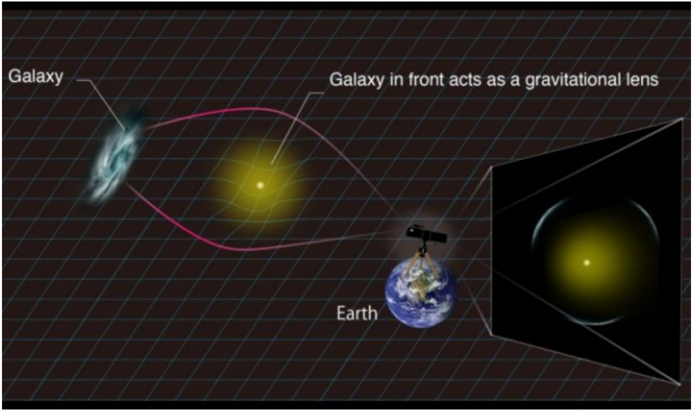
Lih-King Lim (*Zhejiang Univ.*)

Lecture 17

Optics and Gravitation

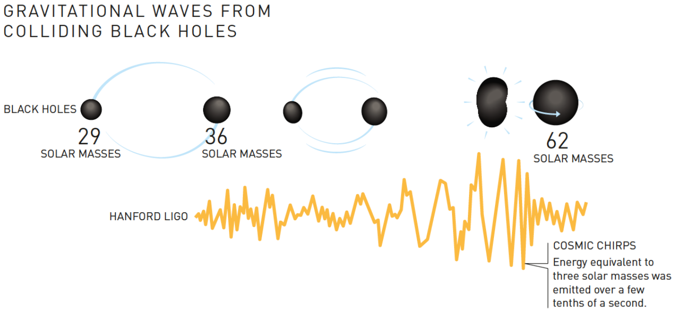


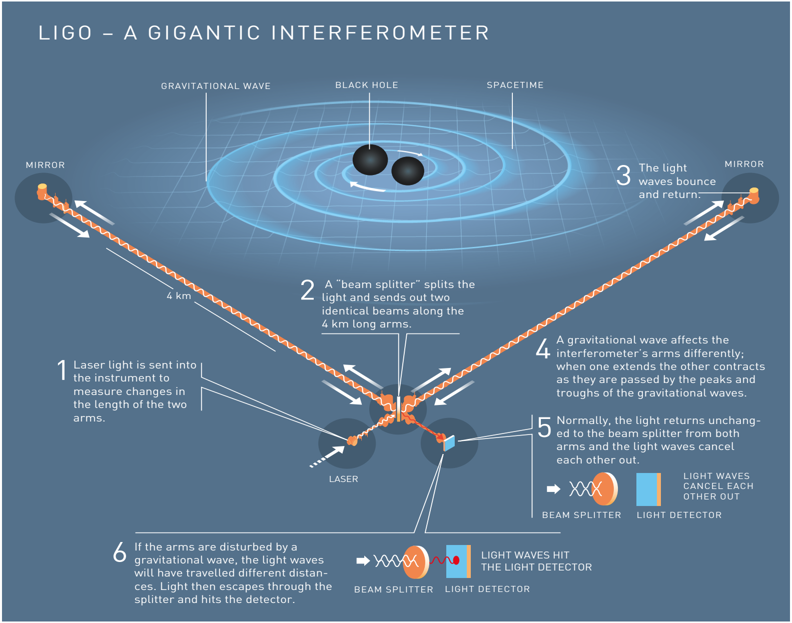
- Space tells matter how to move and matter tells space how to curve. Such distortions of spacetime caused by massive bodies can be measured by light.



- For example, massive galaxies (or black holes) act like special lenses.

- On February 11, 2016, the Advanced LIGO (Laser Interferometer Gravitational-Wave Observatory) team announced that they had directly detected gravitational waves from a pair of black holes merging.




LIGO – A GIGANTIC INTERFEROMETER
GRAVITATIONAL WAVE
BLACK HOLE
SPACETIME
MIRROR
MIRROR
3 The light waves bounce and return.
4 km
2 A "beam splitter" splits the light and sends out two identical beams along the 4 km long arms.
1 Laser light is sent into the instrument to measure changes in the length of the two arms.
4 A gravitational wave affects the interferometer's arms differently; when one extends the other contracts as they are passed by the peaks and troughs of the gravitational waves.
5 Normally, the light returns unchanged to the beam splitter from both arms and the light waves cancel each other out.
LASER
LIGHT WAVES CANCEL EACH OTHER OUT
BEAM SPLITTER
LIGHT DETECTOR
6 If the arms are disturbed by a gravitational wave, the light waves will have travelled different distances. Light then escapes through the splitter and hits the detector.
LIGHT WAVES HIT THE LIGHT DETECTOR
BEAM SPLITTER
LIGHT DETECTOR

Outline

- The Superposition of Waves
- Young's Double-Slit Interference Experiment
- Interference from Thin Films

The Superposition of Waves

- The phenomena of interference, diffraction, and polarization share a common conceptual basis in that they deal with two or more light waves overlap in some region of space.
- We are interested in learning how the specific properties of each constituent wave (amplitude, phase, frequency, etc.) influence the ultimate form of the composite disturbance.

- Recall that each field component of an electromagnetic wave (E_x, E_y, E_z, B_x, B_y, and B_z) satisfies the scalar 3D differential wave equation,

$$\frac{1}{v^2}\frac{\partial^2\psi}{\partial t^2} = \frac{\partial^2\psi}{\partial x^2} + \frac{\partial^2\psi}{\partial y^2} + \frac{\partial^2\psi}{\partial z^2}.$$

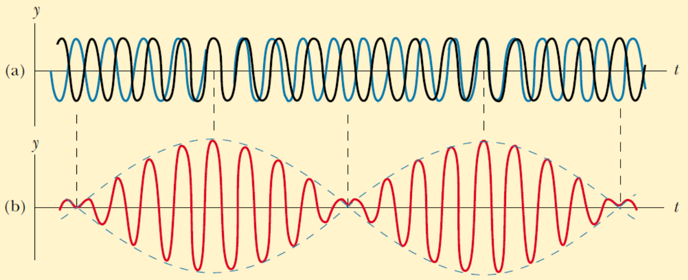
- This equation is *linear*; $\psi(\vec{r}, t)$ and its derivatives appear only to the first power. Consequently, if $\psi_i(\vec{r}, t)$ are solutions, *any linear combination* of them

$$\psi(\vec{r}, t) = \sum_{i=1}^{n} C_i \psi_i(\vec{r}, t)$$

will be a solution as well.

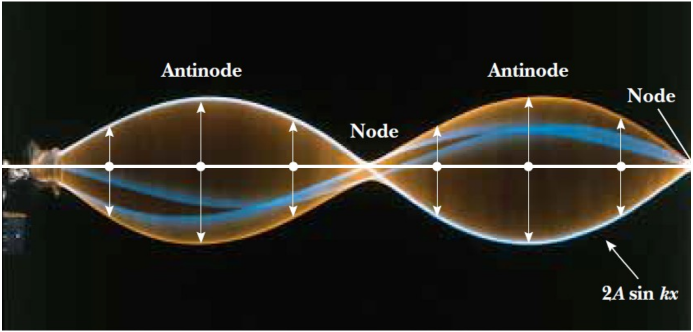
Examples of Superposition

- Beats: Two harmonic waves of different frequency traveling in the same direction.



$$\psi = \psi_1 + \psi_2 = A\cos\omega_1 t + A\cos\omega_2 t, \quad \omega_1 \approx \omega_2$$

- Standing waves: Two harmonic waves of the same frequency propagating in opposite directions.



$$\psi = \psi_1 + \psi_2 = A\cos(kx - \omega t) - A\cos(-kx - \omega t)$$

Trigonometry Formulas

- Sum and difference identities

$$\cos(\alpha+\beta) = \cos\alpha\cos\beta - \sin\alpha\sin\beta$$
$$\cos(\alpha-\beta) = \cos\alpha\cos\beta + \sin\alpha\sin\beta$$

- Sum to product identities

$$\cos\alpha + \cos\beta = 2\cos\frac{\alpha+\beta}{2}\cos\frac{\alpha-\beta}{2}$$
$$\cos\alpha - \cos\beta = -2\sin\frac{\alpha+\beta}{2}\sin\frac{\alpha-\beta}{2}$$

The Algebraic Method of Adding Waves

- There are several equivalent ways of mathematically adding two or more overlapping waves that have the same frequency and wavelength.
- Suppose there are two such waves

$$E_1 = E_{01}\cos(\alpha_1 - \omega t),$$

$$E_2 = E_{02}\cos(\alpha_2 - \omega t),$$

where

$$\alpha_i = kx_i + \phi_i$$

with x_i being the distance from the source s_i of the wave to the point of observation.

- The linear combination of the waves is

$$E \equiv E_0 \cos(\alpha - \omega t) = E_1 + E_2,$$

where (notice $\cos \omega t$ and $\sin \omega t$ are linearly independent)

$$E_0 \cos \alpha = E_{01} \cos \alpha_1 + E_{02} \cos \alpha_2,$$

$$E_0 \sin \alpha = E_{01} \sin \alpha_1 + E_{02} \sin \alpha_2,$$

and

$$E_0^2 = E_{01}^2 + E_{02}^2 + 2E_{01}E_{02} \cos(\alpha_2 - \alpha_1).$$

- The resultant intensity is not simply the sum of the component intensity; there is an additional contribution $2E_{01}E_{02}\cos(\alpha_2 - \alpha_1)$, known as the **interference term**.
- The crucial factor is the **phase difference** between the two interfering waves E_1 and E_2, $\delta \equiv (\alpha_2 - \alpha_1)$.
- The phase difference may arise from a difference in path length traversed by the two waves, as well as a difference in the initial phase angle, i.e.,

$$\delta = \frac{2\pi}{\lambda}(x_2 - x_1) + (\phi_2 - \phi_1).$$

- When $\alpha_2 - \alpha_1 = 2m\pi$ for integer m, we have

$$E_0^2 = E_{01}^2 + E_{02}^2 + 2E_{01}E_{02}.$$

  - In particular, $E_0^2 = 4E_{01}^2$ when $E_{01} = E_{02}$.
  - The two waves **interfere constructively**.

- When $\alpha_2 - \alpha_1 = (2m+1)\pi$ for integer m, we have

$$E_0^2 = E_{01}^2 + E_{02}^2 - 2E_{01}E_{02}.$$

  - In particular, $E_0^2 = 0$ when $E_{01} = E_{02}$.
  - The two waves **interfere destructively**.

The Complex Method

- It is often mathematically convenient to make use of the complex representation when dealing with the superposition of harmonic waves.
- We now redo the calculation of adding two waves. The wave function

$$E_1 = E_{01}\cos(\alpha_1 - \omega t) = E_{01}\cos(kx_1 - \omega t + \phi_1)$$

can be written as

$$\tilde{E}_1 = E_{01}e^{i(\alpha_1 - \omega t)} = E_{01}e^{i\alpha_1}e^{-i\omega t},$$

if we are interested only in the real part.

- Suppose that there are two such overlapping waves having the same frequency and traveling in the positive x-direction. The resultant wave is given by

$$\tilde{E} = E_0 e^{i\alpha} e^{-i\omega t} = \left[E_{01} e^{i\alpha_1} + E_{02} e^{i\alpha_2} \right] e^{-i\omega t},$$

where $E_0 e^{i\alpha}$ is known as the complex amplitude of the composite wave.

- Since $E_0^2 = (E_0 e^{i\alpha})^* (E_0 e^{i\alpha})$, we find

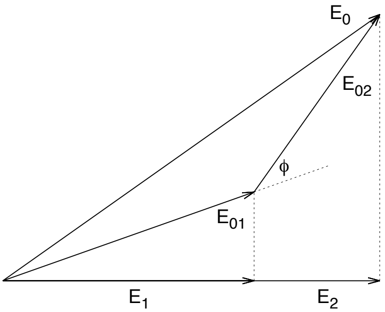
$$\begin{aligned} E_0^2 &= \left[E_{01} e^{-i\alpha_1} + E_{02} e^{-i\alpha_2} \right] \left[E_{01} e^{i\alpha_1} + E_{02} e^{i\alpha_2} \right] \\ &= E_{01}^2 + E_{02}^2 + E_{01} E_{02} \left[e^{i(\alpha_2 - \alpha_1)} + e^{-i(\alpha_2 - \alpha_1)} \right] \\ &= E_{01}^2 + E_{02}^2 + 2 E_{01} E_{02} \cos(\alpha_2 - \alpha_1). \end{aligned}$$

Phasor Addition

- Now we discuss a graphical technique that uses vector additions instead of trigonometric additions.
- We represent a wave, which has an amplitude and a phase, to a vector, known as a **phasor**, in a two-dimensional plane, such that

$$
\begin{aligned}
E_i &= E_{0i}\cos(\alpha_i - \omega t) \\
&= (E_{0i}\cos\alpha_i)\cos\omega t + (E_{0i}\sin\alpha_i)\sin\omega t \\
\Rightarrow \vec{E}_i &= (E_{0i}\cos\alpha_i)\hat{x} + (E_{0i}\sin\alpha_i)\hat{y}.
\end{aligned}
$$

- The algebraic sum, $E = E_1 + E_2$, is the projection on the x axis of the corresponding phasor sum.
- If we denote the phase delay $\phi = \alpha_2 - \alpha_1$ as shown in the **phasor diagram**, the law of cosines applied to the triangle of sides $\vec{E}_{01}$, $\vec{E}_{02}$, and $\vec{E}_0$ yields



$$E_0^2 = E_{01}^2 + E_{02}^2 + 2E_{01}E_{02}\cos(\alpha_2 - \alpha_1).$$

- Understand the equivalence of the following formulas!

$$E_0 \cos\alpha \cos\omega t + E_0 \sin\alpha \sin\omega t = \\ E_{01}\cos\alpha_1 \cos\omega t + E_{01}\sin\alpha_1 \sin\omega t + \\ E_{02}\cos\alpha_2 \cos\omega t + E_{02}\sin\alpha_2 \sin\omega t$$

$$E_0 e^{i\alpha} e^{-i\omega t} = \left[E_{01} e^{i\alpha_1} + E_{02} e^{i\alpha_2}\right] e^{-i\omega t}$$

$$E_0 \cos\alpha \hat{x} + E_0 \sin\alpha \hat{y} = \\ E_{01}\cos\alpha_1 \hat{x} + E_{01}\sin\alpha_1 \hat{y} + E_{02}\cos\alpha_2 \hat{x} + E_{02}\sin\alpha_2 \hat{y}$$

- Note that the three methods all deal with addition in a two-dimensional space.

- The phasor addition adds vectors (with x component and y component) in a two-dimensional real space.
- The complex method is an addition of complex numbers (with real part and imaginary part) in a complex plane.
- The algebraic or trigonometric method is most complex. In fact, it is an addition of functions (linear combination of $\cos \omega t$ and $\sin \omega t$) in a two-dimensional Hilbert space, spanned by $\cos \omega t$ and $\sin \omega t$.

Natural Light

- Now, one may think that light from two fine incandescent wires would interfere.
- This does not occur, because the light is emitted by vast numbers of atoms in the wires, acting randomly and independently for extremely short times — of the order of nanoseconds. The light is said to be **incoherent**.
- As a result, at any given point on the viewing screen, the interference between the waves from the two sources varies rapidly and randomly between fully constructive and fully destructive. The screen is seen as being uniformly illuminated (over the time scale of our observation).

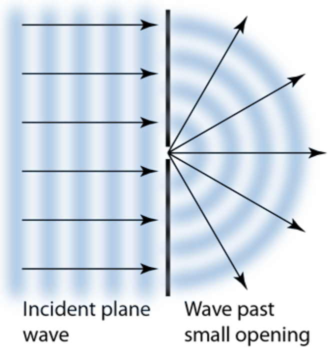
Conditions for Interference

- To observe the interference of the two waves described above, we need the following conditions.
  - Two beams must have (nearly) the same frequency ω. Otherwise, the phase difference is time-dependent. During the detection interval, the interference pattern will be averaged away.
  - The clearest pattern (with maximum contrast) exists when interfering waves have (nearly) equal amplitude. Why?
  - Initial phase difference can exist between sources, as long as it remains constant; the two sources are said to be **coherent**.
- Until the advant of laser, no two individual sources can maintain a constant relative phase long enough for an observable interference pattern.

Huygens' Principle, Again

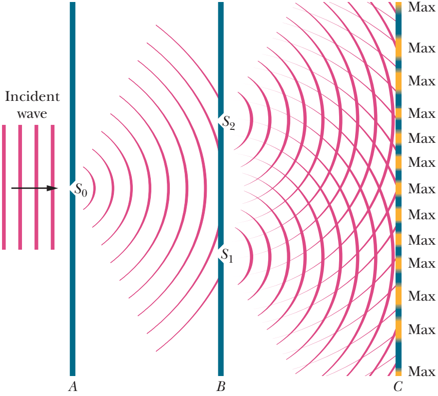
- According to Huygens' principle, each point on a wavefront may be regarded as a source of waves expanding from that point.

- If waves strike a barrier with a small opening, the waves may be seen to expand from the opening. Notice the wavelength is larger than the opening in this case.



Young's Interference Experiment (1801)

- Here is the ingenious arrangement of Thomas Young's double-slit experiment.

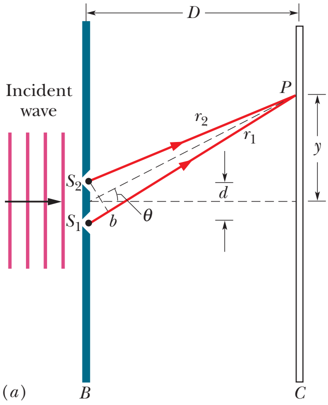


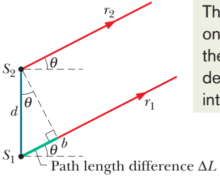
The waves emerging from the two slits overlap and form an interference pattern.

- The slit S_0 in screen A creates a spatially coherent beam that could identically illuminate slits S_1 and S_2 in screen B.
    - Spatial incoherence arises from the fact that a realistic source (like a lightbulb) contains widely spaced point sources that changes phase rapidly and randomly (due to the granular nature of the emission process).
- Nowadays screen A is no longer needed, and plane waves from a laser can provide the spatial coherence the experiment needs.
- Light waves produce fringes in a Young's double-slit interference experiment, but what exactly determines the locations of the fringes?

- The waves are in phase when they pass through the two slits because there they are just portions of the same incident wave.
- However, once they have passed the slits, the two waves must travel different distances to reach point P on screen C.
- The path length difference is, when $d \ll D$,

$$\Delta L = d \sin \theta.$$





The ΔL shifts one wave from the other, which determines the interference.

- In other words, the light leaving the slits is in phase. However, the electric field components of these waves at point P are not in phase and vary with time as

$$E_1 = E_0 \cos(kr_1 - \omega t) = E_0 \cos(kL + \beta - \omega t),$$

$$E_2 = E_0 \cos(kr_2 - \omega t) = E_0 \cos(kL - \beta - \omega t),$$

where the phase difference $[L = (r_1 + r_2)/2 = \sqrt{D^2 + y^2}]$

$$\delta_2 = 2\beta = k\Delta L = \frac{2\pi d}{\lambda} \sin\theta.$$

- The total intensity is thus given by

$$I = 2E_0^2[1 + \cos(2\beta)] = I_{\max} \cos^2 \beta.$$

- Therefore, a bright fringe appears when

$$\Delta L = d \sin\theta = m\lambda,$$

  where m is an integer.
- On the other hand, a dark fringe appears when

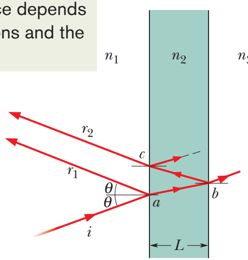
$$\Delta L = d \sin\theta = \left(m + \frac{1}{2}\right)\lambda,$$

  where m is an integer.
- We can then find the angle θ to any fringe and thus use the values of m to label the fringes.

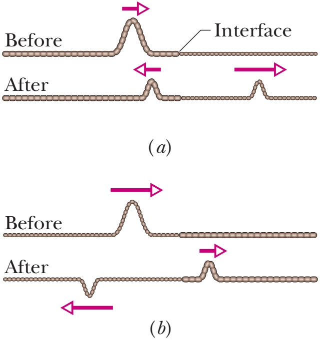
Interference from Thin Films

- Consider a thin transparent film of uniform thickness L and index of refraction n_2, illuminated by bright light of wavelength λ from a distant point source. We assume $n_1 = n_3 = n_{\mathrm{air}}$.
- For simplicity, we also assume that the light rays are almost perpendicular to the film ($\theta \approx 0$). We are interested in whether the film is bright or dark to an observer viewing it almost perpendicularly.

The interference depends on the reflections and the path lengths.



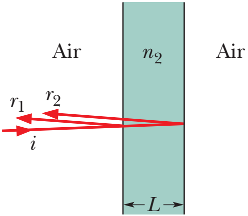
- Reflection at an interface can cause a phase change, depending on the indexes of refraction on the two sides of the interface. There is no phase change for refraction.



- The results for light reflecting off a medium can be summarized as

Reflection	Reflection phase shift
Off lower index	0
Off higher index	0.5 wavelength

- So, reflecting off higher index, ray r_1 has an additional reflection phase shift 0.5 wavelength. There is no such shift for ray r_2.
- In addition, the light waves of rays r_1 and r_2 has a path difference $2L$, which occurs in index n_2. Notice the wavelength in the medium is



$$\lambda_2 = \frac{v_2}{f} = \frac{c}{n_2}\frac{1}{f} = \frac{c}{f}\frac{1}{n_2} = \frac{\lambda}{n_2}.$$

- Therefore, rays are in phase if

$$2L = \left(m + \frac{1}{2}\right)\frac{\lambda}{n_2}, \quad \text{for integer } m.$$

They produce an interference maximum and the nearby region on the film is bright to observers.

- Similarly, if they are exactly out of phase,

$$2L = m\frac{\lambda}{n_2}, \quad \text{for integer } m,$$

they produce an interference minimum and the nearby region is dark, even though it is illuminated.

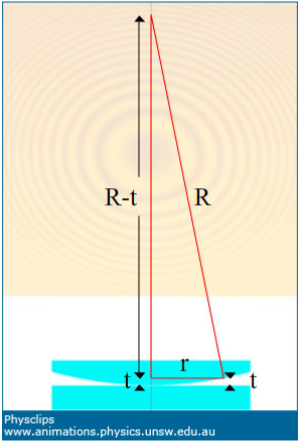
Negligible Film Thickness

- A special situation arises when a film is so thin that L is much less than λ, say, $L < 0.1\lambda$.
- Then the path length difference $2L$ can be neglected, and the phase difference between r_1 and r_2 is due only to reflection phase shifts.
- Thus r_1 and r_2 are exactly out of phase, and thus the film is dark, regardless of the wavelength and intensity of the light.

Newton's Rings

- Newton's rings are interference patterns formed by light incident on the thin film of air between a convex lens and a flat (or between two suitable lenses).

- Now, at what radius do you expect to find dark rings and bright rings?



Summary

- Understand that waves can be superposed. Master at least one of the following equivalent techniques:
  - Addition of trigonometric functions
  - Addition of complex functions
  - Phasor addition
- Understand the interference of light in Young's experiment.
  - How to find the location of bright (dark) fringes?
  - What is the intensity of the interference pattern?
  - Why do we need the single slit before the two slits?
- Generalize the idea of double-slit interference to interference from thin films. Pay attention to the phase shift at one of the surfaces (not the other!).

Reading

Halliday, Resnick & Krane:

- Chapter 41: Interference