

2.11. Wave optics: Interference & polarization



Having examined geometric optics and the ray model of light, we now explore phenomena that reveal light's wave nature. While the ray model is useful for macroscopic interactions, wave effects become significant at smaller scales. We will focus on **interference** and **polarization**: The superposition of light waves leads to interference patterns and the transverse nature of light results in polarization.

2.11.1 Huygens' principle

The **ray model**, while powerful for many optical phenomena, **encounters its limits when dealing with light interacting with objects comparable in size to its wavelength**. Consider a simple yet revealing experiment: allowing sunlight to enter a dark room only through a very small pinhole. If light were solely a stream of particles traveling in straight lines, we would expect to see a sharp, well-defined bright spot on a screen placed opposite the pinhole, essentially a projection of the hole's shape.

However, the reality is strikingly different. Instead of a simple dot, what you will observe is a more complex pattern. This pattern typically consists of a central, brighter spot surrounded by a series of fainter concentric rings or a pattern of alternating bright and dark fringes. This observation directly contradicts the predictions of the ray model and clearly indicates that light does not simply travel in straight lines under these conditions. This spreading of light and the formation of these intricate patterns are direct consequences of light's wave nature, arising from the phenomena of **interference and diffraction**.

To understand and predict these wave-like behaviors of light, we need a more sophisticated framework than geometric optics provides. **Christian Huygens** (1629-1695) was a pioneer in developing such a concept. He proposed a powerful principle that allows us to visualize and analyze the propagation of waves: *Imagine every point of a wavefront as a source of secondary spherical wavelets. These wavelets propagate outwards in all directions at the speed of the wave. The new wavefront at a later time is the tangent surface to all of these secondary wavelets.*

Let's digest this information:

- A **wavefront** is a surface connecting points of a wave that are in phase. Imagine dropping a stone into a still pond. The ripples that spread out are wavefronts; all the points on a single ripple crest are moving up together (in phase). For light, which is an electromagnetic wave, a wavefront connects points where the electric and magnetic fields are oscillating in the same way at a given instant.
- Think of a wavefront as a line (in 2D) or a surface (in 3D) of **many tiny sources**.

- Each of these tiny sources sends out its own little wave, called a **wavelet**, spreading out like tiny ripples.
- **Superposition of these wavelets generates the new wavefront:** The individual wavelets interfere with each other, and their combined effect, through the principle of superposition, determines the shape and position of the wavefront at the next instant (wavefront tangent to all wavelets)
- **Direction of the wave's travel is perpendicular to the wavefront:** Imagine the ripples on the water; the wave moves outwards, and this direction of movement is always at a right angle to the crest of the ripple (the wavefront).

```
interactive(children=(FloatSlider(value=9.0, description='Hole Size', max=10.0, min=0.1),
  Checkbox(value=False...
```

```
<function __main__.simulate_light_through_hole(hole_size, use_wave_model)>
```

2.11.2 Revisiting refraction from a wave perspective

Let's revisit refraction using waves instead of the ray model. Light is traveling from medium 1 to medium 2. We assume the mediums to be **isotropic**, meaning they have the same speed of the wave v in all directions, but have different indices of refraction, n_1 and n_2 . In general, the distance a wave travels in a time interval t is given by $r = vt$. As we learned previously, the speed at which light travels depends on the medium. The speed of light in a vacuum is denoted by c . In a medium with a refractive index n , the speed of light is reduced to $v = \frac{c}{n}$. Therefore, in our two mediums, the speeds are $v_1 = \frac{c}{n_1}$ and $v_2 = \frac{c}{n_2}$. Consequently, the distance light travels in the same time t will be different in both mediums:

$$r_1 = v_1 t \neq r_2 = v_2 t.$$

For cases where the angle of incidence is non-zero ($\theta_1 \neq 0$), refraction occurs, meaning the direction of wave propagation changes. The angle of refraction θ_2 can be found using Snell's law, which we can derive from our wave model.

Consider a wavefront approaching the interface between medium 1 and medium 2 at an angle θ_1 with respect to the normal. Let point A be where one part of the wavefront first hits the second medium. Another part of the same wavefront will reach the interface at a point B after a time t . During this time t , the part of the wavefront that entered medium 2 at point A travels a distance $v_2 t$, while the part still in medium 1 travels a distance $v_1 t$ to reach point B . The distance AB represents the length of the wavefront segment that is just about to cross the interface.

From the geometry, we can see that:

$$\sin \theta_1 = \frac{v_1 t}{AB}; \quad \sin \theta_2 = \frac{v_2 t}{AB}$$

Dividing the first equation by the second, we get:

$$\frac{\sin \theta_1}{\sin \theta_2} = \frac{v_1 t / AB}{v_2 t / AB} = \frac{v_1}{v_2}$$

Substituting the expressions for v_1 and v_2 in terms of the refractive indices:

$$\frac{\sin \theta_1}{\sin \theta_2} = \frac{c/n_1}{c/n_2} = \frac{n_2}{n_1}$$

Rearranging this equation gives us Snell's Law:

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

As we learned before, the **frequency of the wave does not change** when refraction occurs. This is because the frequency is determined by the source of the wave, and it remains constant as the wave passes from one medium to another. However, the **wavelength** λ does change because the speed v changes, and $v = \lambda f \leftrightarrow \lambda = \frac{v}{f}$ with constant f .

Consequently, the ratio of the wavelengths in the two mediums is:

$$\frac{\lambda_1}{\lambda_2} = \frac{v_1/f}{v_2/f} = \frac{v_1}{v_2}$$

And since $\frac{v_1}{v_2} = \frac{n_2}{n_1}$, we have:

$$\frac{\lambda_1}{\lambda_2} = \frac{n_2}{n_1}$$

This means that the wavelength of light is shorter in a medium with a higher refractive index.

If λ represents the wavelength of light in a vacuum (where $n = 1$ and $v = c$), then the wavelength λ_n of light in a medium with refractive index n is given by:

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

This shows that the wavelength of light in a medium is reduced by a factor equal to the refractive index of the medium compared to its wavelength in a vacuum.

2.11.3 Polarization

Electromagnetic (EM) waves are transversal waves and, therefore, can be **polarized**. To understand what plane-polarized waves are (as they oscillate in a plane; sometimes called linearly polarized waves), let's imagine a rope. We could swing the rope to perform oscillations in, for example, the horizontal plane. If we put the rope through a long, but thin slit, the rope would only oscillate if the slit is oriented along the oscillation direction, i.e., horizontally. This concept can be extended to EM waves, i.e., light. Usually light is **unpolarized**, that means, the electric field oscillates randomly in all directions perpendicular to the direction of propagation, with no single, consistent plane of oscillation.

Polarization by absorption (polaroids)

Polaroid sheets, invented by Edwin Land, contain long molecules aligned parallel to each other. We can imagine these molecules to **act as** many parallel slits. The component of the electric field oscillating

parallel to these molecules causes the electrons in the molecules to move, thus transferring energy. Thus, the light oscillating parallel to the molecule is **absorbed**. Light with an electric field oscillating **perpendicular** to the long molecules is not strongly absorbed and passes through the Polaroid. The direction of this transmitted electric field is called the **transmission axis** (or polarization axis) of the Polaroid. Therefore, the transmission axis is **perpendicular** to the orientation of the long molecules.

A Polaroid acts as a polarizer, creating plane-polarized light from unpolarized light by absorbing the component of the electric field that is parallel to the molecules (perpendicular to the transmission axis). The intensity of polarized light transmitted through a Polaroid is given by **Malus's Law**:

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

where θ is the angle between the incident polarization and the transmission axis. For unpolarized light, the transmitted intensity is halved ($I = \frac{1}{2} I_0$).

Polaroids can also act as analyzers to determine if light is polarized and its plane of polarization by observing intensity changes upon rotation. Two crossed polaroids (transmission axes perpendicular) block all light as the light after the first polarizer is perpendicular to the transmission axis of the second polarizer. An interesting effect occurs with three polaroids: even if the first and third are crossed, placing the second at a 45-degree angle between them allows some light to pass.

Polarization by reflection

Unpolarized light can also be **polarized upon reflection** from a nonmetallic surface at non-perpendicular angles. The reflected light becomes **preferentially polarized parallel to the reflecting surface**. Hence, light is preferentially transmitted and refracted with a polarization plane perpendicular to the surface. This is why Polaroid sunglasses, with vertically oriented transmission axes, effectively reduce glare from horizontal surfaces like roads and water by blocking the horizontally polarized reflected light.

The degree of polarization in the reflected beam varies with the angle of incidence, reaching 100% at a specific angle called the **polarizing angle** or **Brewster's angle** (θ_p). This angle is determined by the refractive indices of the two media (n_1 for the incident medium, n_2 for the reflecting medium) according to **Brewster's Law**:

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

If light is incident from air ($n_1 = 1$), this simplifies to $\tan \theta_p = n_2$.

Interestingly, at Brewster's angle, the reflected and refracted rays are perpendicular to each other, meaning the angle of incidence plus the angle of refraction equals 90 degrees ($\theta_p + \theta_r = 90^\circ$).

2.11.4 Scattering in the atmosphere

The blue color of the sky, the redness of sunsets, and the polarization of skylight are all consequences of light scattering by atmospheric molecules. When sunlight strikes a molecule (like nitrogen or oxygen),

the electric field of the light wave causes the molecule's electrons to oscillate. These oscillating charges then re-emit light in all directions.

The intensity of this scattered light is strongest in directions perpendicular to the oscillation of the electrons and zero along the axis of oscillation. For an observer viewing the sky at a 90-degree angle to the direction of sunlight, the scattered light appears plane-polarized because they are observing perpendicular to one of the oscillation components. At other angles, the skylight is only partially polarized.

The amount of scattering is strongly dependent on the wavelength of light, being inversely proportional to the fourth power of the wavelength ($1/\lambda^4$). This means that shorter wavelengths, like blue and violet light, are scattered much more effectively than longer wavelengths, such as red and orange. The blue sky is a result of this preferential scattering of blue light in all directions.

At sunset, sunlight travels through a much longer path in the atmosphere. During this extended passage, most of the blue light is scattered away. The light that reaches our eyes directly from the setting sun, or reflected off clouds near the horizon, is therefore depleted in blue and appears reddish.

This strong wavelength dependence of scattering applies to particles much smaller than the wavelength of light, like air molecules. Clouds, however, are composed of water droplets or ice crystals that are significantly larger than the wavelength of light. These larger particles scatter all colors of light nearly uniformly, which is why clouds typically appear white.

2.11.5 Brightness

While the physical intensity of light is measured in watts per square meter [W/m^2] and the total power output (radiant flux) in watts, perceived brightness considers the visible spectrum and the eye's sensitivity, peaking at 550 nm. **Luminous flux** F_l , measured in lumens [lm], quantifies this visual sensation ($1 \text{ lm} = 1/683 \text{ W of } 555 \text{ nm light}$). **Luminous intensity** I_l is the luminous flux per unit solid angle (steradian), with the unit candela [cd], where $1 \text{ cd} = 1 \text{ lm/sr}$. Finally, **illuminance** E_l measures the luminous flux incident on a surface per unit area, with the unit lumens per square meter [lm/m^2].

2.11.6 Interference: Newton's rings & thin films

Let's simplifying our thought experiment for Newton's rings, by introducing monochromatic light.

Monochromatic light refers to light of a single wavelength or a very narrow range of wavelengths, often produced by lasers or specific filters.

Newton's rings with monochromatic light

When a plano-convex lens with a large radius of curvature is placed on a flat glass surface, a very thin, wedge-shaped air gap is created between them. If this setup is illuminated from above with monochromatic light, an observer will see a series of concentric bright and dark circular rings, known as

Newton's rings. These rings only appear for the part of the air gap, for which the thickness of this air gap is typically only a few multiples of the wavelength of light at its maximum. Newton's rings are an **interference pattern** which arises because light waves are reflected from two surfaces:

- The bottom surface of the lens (glass-air interface)
- The top surface of the flat glass plate (air-glass interface)

Consider a ray of light incident on the setup. The air-glass interface can be ignored in our explanation. The light transmitted into the glass will reach the glass-air interface. Part of the light is reflected and reaches the observer. The other part of the light is transmitted into the air gap and will be reflected at the flat glass surface before reaching the observer (we neglected all other interfaces except the two state above). We assume that, due to the very thin air gap, the part reflected and the one going the "extra mile" though the air gap reach the observe at the same place, i.e. the waves can superimpose, or in other words **interfere** with each other. Compared to the part of the light that is only reflected from the bottom of the lens, and the part which is refracted into the air gap and then reflected from the top of the flat plate, traveled an extra distance of approximately twice the thickness of the air gap at that point. This extra distance is the **path difference** between the two reflected waves.

Before thinking about the individual rays, let's consider the point of contact between the plano-convex lens and the flat glass surface. At this point of contact, the air gap thickness is essentially zero, so the path difference due to the gap is negligible. However, there seems to be no light being emitted from this point. Hence, the waves seem to interfere **destructively** and not **constructively**. What is the reason?

When light reflects from a medium of lower refractive index to one of higher refractive index (like air to glass at the top of the flat plate), **it undergoes a 180° phase shift**, which is equivalent to a path difference of $\lambda/2$. The reflection from the glass-air interface at the bottom of the lens does not incur such a phase shift. Therefore, at the center, the two reflected waves have a phase difference of $\lambda/2$ due to the reflection, leading to **destructive interference and a dark central spot**.

As the distance from the center increases, the thickness (t) of the air gap gradually increases. The **path difference** Δ introduced by the air gap $2t$, combined with the $\frac{\lambda}{2}$ phase shift upon reflection from the flat plate, determines the total path difference

$$\Delta = 2t + \frac{\lambda}{2}$$

Before starting with more complex ray interactions, let consider the central part of the lense, i.e. the point of contact with no air gap ($t = 0$). The path difference will be $\Delta = 0 + \frac{\lambda}{2}$. Therefore, we obtain a **destructive** interference because the path difference is exactly half the wavelength. Consequently, destruction interference would occur whenever the path difference is $\frac{\lambda}{2}, \frac{3\lambda}{2}, \frac{5\lambda}{2}$ which can be generalized as $\frac{\lambda}{2}(2m + 1)$ for $m = 0, 1, 2, \dots$.

Thus, we get **dark rings (destructive interference)**:

$$\Delta = 2t + \frac{\lambda}{2} = \frac{\lambda}{2}(2m + 1)$$

$$2t + \frac{\lambda}{2} = \frac{2\lambda m}{2} + \frac{\lambda}{2}$$

$$2t = \lambda m$$

The same logic can be applied for **constructive** interference. The total path difference must be a multiple of the wavelength λ , 2λ , 3λ which can be generalized as $m\lambda$ for $m = 1, 2, \dots$

Thus, we get **bright rings (constructive interference)**:

$$\Delta = 2t + \frac{\lambda}{2} = m\lambda$$

$$2t = \lambda(m - \frac{1}{2})$$

To summarize:

- **Dark rings (destructive interference):** Occur when this condition is met:
 $2t + \frac{\lambda}{2} = \frac{\lambda}{2}(2m + 1) = (m + \frac{1}{2})\lambda$ with $m = 0, 1, 2, \dots$ as the order of the dark rings.
Rearranging this, we get $2t = \lambda m$.
- **Bright rings (constructive interference):** Occur when this condition is met: $2t + \frac{\lambda}{2} = m\lambda$, with $m = 1, 2, \dots$ as the order of the bright rings. Rearranging this, we get $2t = (m - \frac{1}{2})\lambda$.

Newton's rings with white light

If white light, which consists of a spectrum of wavelengths, is used to illuminate the Newton's rings setup, the observed pattern is different. Instead of alternating bright and dark rings, a series of colored concentric rings is seen. This is because constructive interference occurs for different wavelengths (colors) of light at different thicknesses of the air gap. The rings closest to the center are typically more vibrant and distinct, showing a sequence of colors similar to a rainbow. However, as the distance from the center increases and the order of interference becomes higher, the rings tend to overlap more, and the color contrast becomes less pronounced, eventually fading into a nearly white or blurry pattern.

Interference in thin films

The phenomenon of interference also occurs in other types of thin films, such as soap bubbles, oil films on water, and coatings on various surfaces. Similar to Newton's rings, the colors observed in these films are a result of the **interference of light waves reflected from the two interfaces of the thin film**. We need to consider two waves to understand interference at the thin film: The first one is simply reflected at the first interface before reaching the observer/eye. The second one is refracted at the first interface, subsequently reflected at the second interface, and then being refracted again at the first interface to reach the observer. Within the thin film, the wavelength will be changed (see dispersion) depending on the refraction index of the medium: $\lambda_n = \frac{\lambda}{n}$. The path difference Δ between the two waves w.r.t. wavelength in the thin film λ_n will determine what kind of interference appears.

For near-normal incidence, the path difference between these two reflected waves is approximately $2nt$, where t is its thickness of the film. The occurrence of phase shifts depends on the refractive indices of the materials at each interface. If the light reflects at both interfaces from a lower to a higher refractive index (or vice versa), the phase shifts can be the same (both 0 or both 180°), effectively canceling each other out when considering the conditions for constructive and destructive interference. In such cases, the interference pattern is primarily determined by the path difference of $2nt$ relative to the wavelength of light in the film (λ/n).

For a set-up of a oil film on water, we have $n_{\text{water}} > n_{\text{oil}} > n_{\text{air}}$ and, therefore, no phase shift. In this case, constructive interference occurs for $\Delta = 2t = m\lambda_n$ with $m = 1, 2, \dots$. Consequently, destructive interference occurs for $\Delta = 2t = (2m + 1)\frac{\lambda_n}{2}$ with $m = 0, 1, 2, \dots$.

2.11.7 Interference: Young's double-slit experiment

Thomas Young (1773-1829) conducted a groundbreaking experiment that provided strong evidence for the wave nature of light. He illuminate a screen with two closely spaced, narrow slits, S_1 and S_2 , separated by a distance d . If the distance d is on the same order as the wavelength of the light, interference of waves can be observed. In fact, the waves are **diffracted**, but diffraction will be the topic of the next chapter. In the remainder of this chapter, we consider a discrete number of waves, i.e. mostly only two, to explain the double-slit experiment by interference. A short discussion on interference vs. diffraction will be provided, but (spoilers) it is more a conceptional difference than a true different physical mechanism.

Young's double-slit experiment with monochromatic light

Let's first examine what happens if monochromatic light is shinned onto the double-slit. According to the wave theory of light, the two slits act as **coherent sources**, meaning they **emit waves with the same wavelength and a constant phase relationship**. As the waves emerge from these slits, they diffract and overlap in the region behind the screen.

Consider a point on a viewing screen placed at a distance L behind the double-slit screen. The light waves from S_1 and S_2 travel different distances to reach this point. If we consider a point on the screen at an angle θ relative to the central axis, the **path difference** (Δ) between the waves from the two slits can be approximated as $\Delta = d \sin \theta$, assuming that the distance to the screen L is much larger than the slit separation d .

Interference occurs where these waves overlap. **Constructive interference**, resulting in a bright fringe on the screen, happens when the path difference is an integer multiple of the wavelength λ of the monochromatic light:

$$d \sin \theta = m\lambda, \quad m = 0, 1, 2, \dots$$

Here, m is the order of the bright fringe.

The central bright fringe ($m = 0$) occurs at $\theta = 0$, where the path difference is zero, and the waves from both slits arrive in phase. The first-order bright fringes ($m = 1$) occur at angles where the path difference is one wavelength, and so on. Notice that the distance on the screen between bright fringes is only constant as long as we can approximate $\sin \theta = \theta$.

Destructive interference, resulting in a dark fringe on the screen, occurs when the path difference is a half-integer multiple of the wavelength:

$$d \sin \theta = (m + \frac{1}{2})\lambda, \quad m = 0, 1, 2, \dots$$

In this case, the waves from the two slits arrive exactly out of phase, and their amplitudes cancel out. The viewing screen thus displays a pattern of alternating bright and dark parallel fringes.

Young's double-slit experiment with white light

When white light is used to illuminate the double slits, the interference pattern observed on the screen is more complex. White light is composed of a continuous spectrum of wavelengths. The condition for constructive interference ($d \sin \theta = m\lambda$) depends on the wavelength λ . Therefore, for a given order m (except for $m = 0$), different wavelengths will constructively interfere at different angles θ .

The central fringe ($m = 0$) remains white because the path difference is zero for all wavelengths, so all colors interfere constructively at the center. However, for the first-order fringes ($m = 1$), violet light (shorter wavelength) will have a smaller angle θ for constructive interference compared to red light (longer wavelength). This results in a spectrum of colors appearing in the first-order fringes, with violet being closest to the central fringe and red being furthest. Higher-order fringes will also exhibit spectra, but they will be broader and will overlap more, making the colors less distinct. This observation of colored fringes with white light further supported the idea that different colors of light correspond to different wavelengths. Further, it enables Young to measure the wavelengths of visible light with his experimental set-up.

```
interactive(children=(FloatSlider(value=0.0, description='Intersection Y', max=5.0, min=-5.0,
step=0.2), Output(),
```

```
<function __main__.simulate_double_slit_interference(intersect_y)>
```