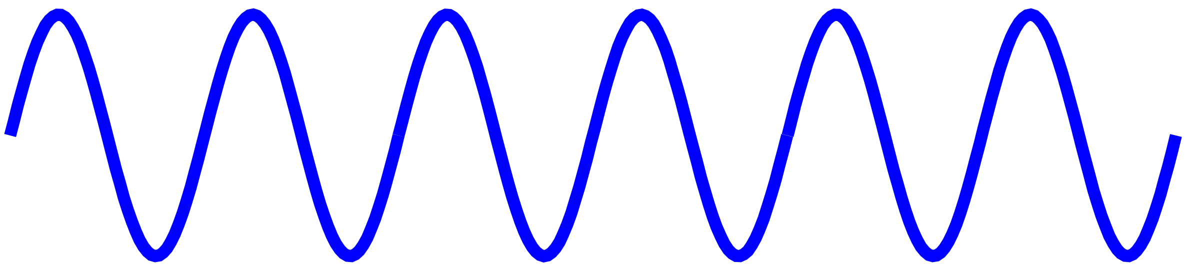
1: PHYS126: Waves and Optics 2004

Waves & Optics

R-D Herzberg



1.1: Overview: Waves

- Classification of waves and their properties
- Wave speed
- Energy in a wave
- The principle of superposition
- Interference
- Reflections and Boundaries
- Standing waves
- Resonance
- Introduction to sound waves
- The decibel scale
- Beats
- The physics of music
- Doppler Effect
- The wave equation

1.2: Some Reference Books

1. **Fundamentals of Physics** Halliday, Resnick, and Walker,
   sixth edition, John Wiley 2001.
2. Berkeley physics course Vol 3: Waves, Frank S. Crawford.
   This book is more math oriented.
3. The Feynman Lectures in Physics (Vol I), Feynman, Leighton, Sands

I will follow Halliday, Resnick and Walker very closely.

Recommended study time: 36 h.

At the end of each lecture are suggestions for exercises and revision.

1.3: Introduction

Classification of waves:

Mechanical waves ⟷ **Light waves**
(need medium) **(needs NO medium)**

sound, water waves, earthquakes, etc

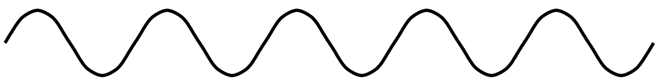
radio, light, X-rays etc

longitudinal — **sound waves**

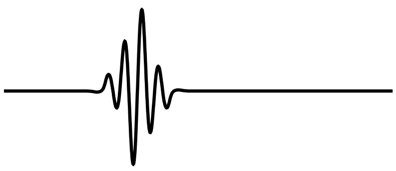
transversal — **water waves**

purely transversal

Periodic waves:

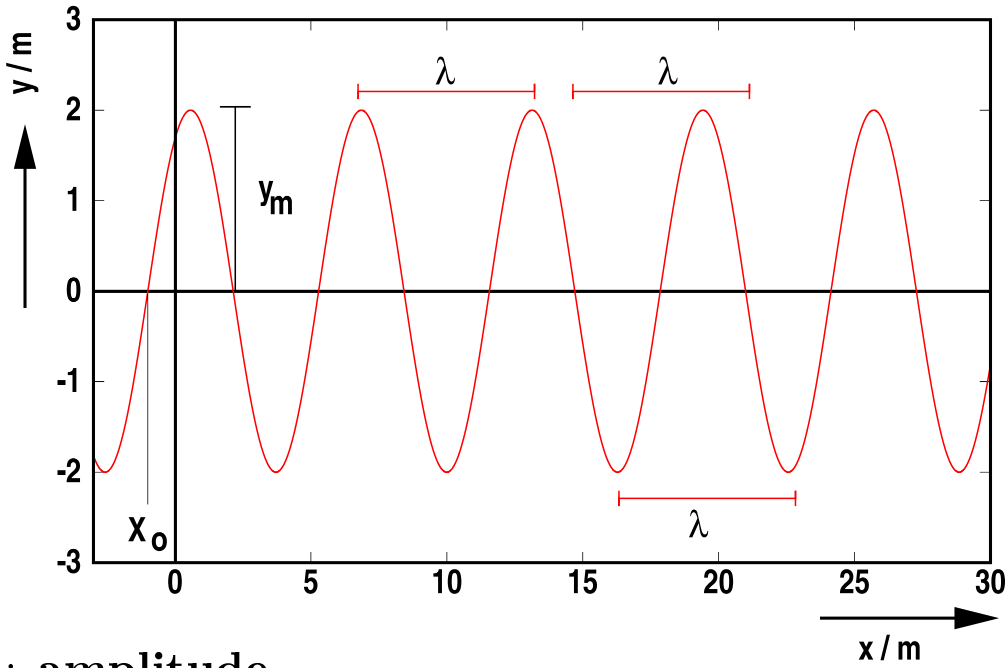


Non-periodic waves (pulses):



In quantum mechanics: matter and probability waves!

Consider a transversal periodic sinusoidal wave travelling along, e.g. a water ripple and take a snapshot of it at a time $t = 0$:



y_m : **amplitude**

λ : **wavelength**

$$y(x,t) = y(x, t = 0) = y_m \sin(k(x - x_0))$$

Substitute $\phi_0 = kx_0$ to give

$$y(x,t) = y(x, t = 0) = y_m \sin(kx - \phi_0)$$

The wave is periodic, therefore we must have

$$y(x + \lambda, 0) = y(x, 0)$$

for all x. Thus:

$$y_m \sin(k(x + \lambda) - \phi_0) = y_m \sin(kx - \phi_0)$$

We know for a sine function:

$$\sin(\alpha + 2\pi) = \sin(\alpha)$$

and therefore $k\lambda = 2\pi$.

$$\boxed{k = \frac{2\pi}{\lambda}}$$

We call k the *angular wave number.* Units:

$$[\lambda] = \mathrm{m}$$

$$[2\pi] = \mathrm{rad}$$

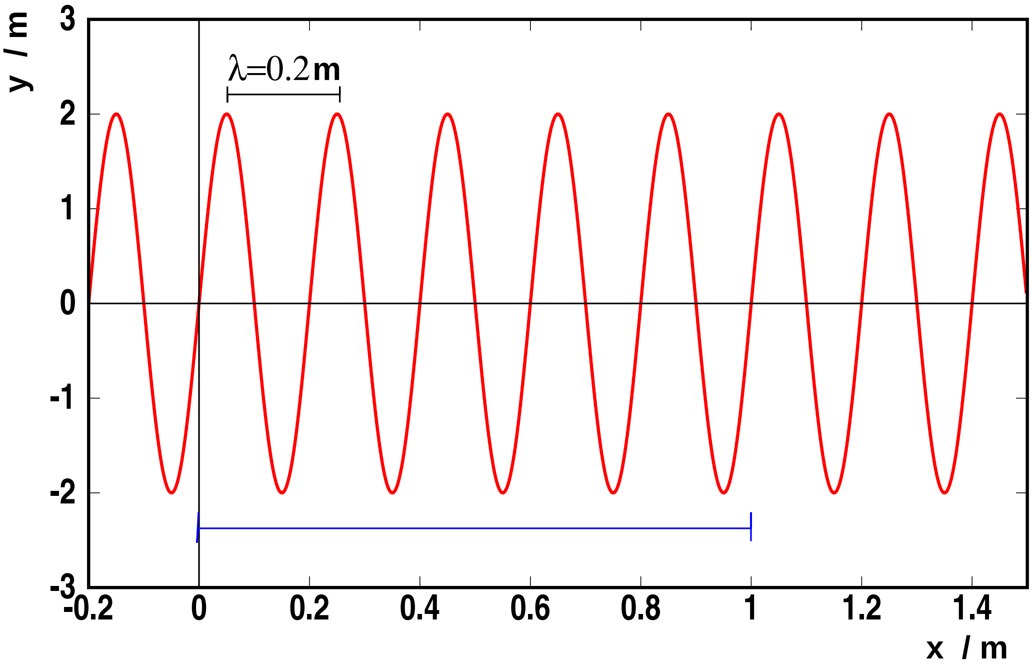
$$[k] = \mathrm{rad/m}$$

It is also useful to define the *wave number* κ:

$$\kappa = 1/\lambda = \frac{k}{2\pi}$$

$$[\kappa] = m^{-1}$$

1.4: Example:



We measure $\lambda = 0.2\,\mathrm{m}$.
Thus $k = 2\pi/\lambda = 31.4\,\mathrm{rad/m}$
and $\kappa = 1/\lambda = 5\,\mathrm{m}^{-1}$.

1.5: Time dependence:

Now look at a given point and let the wave pass by:

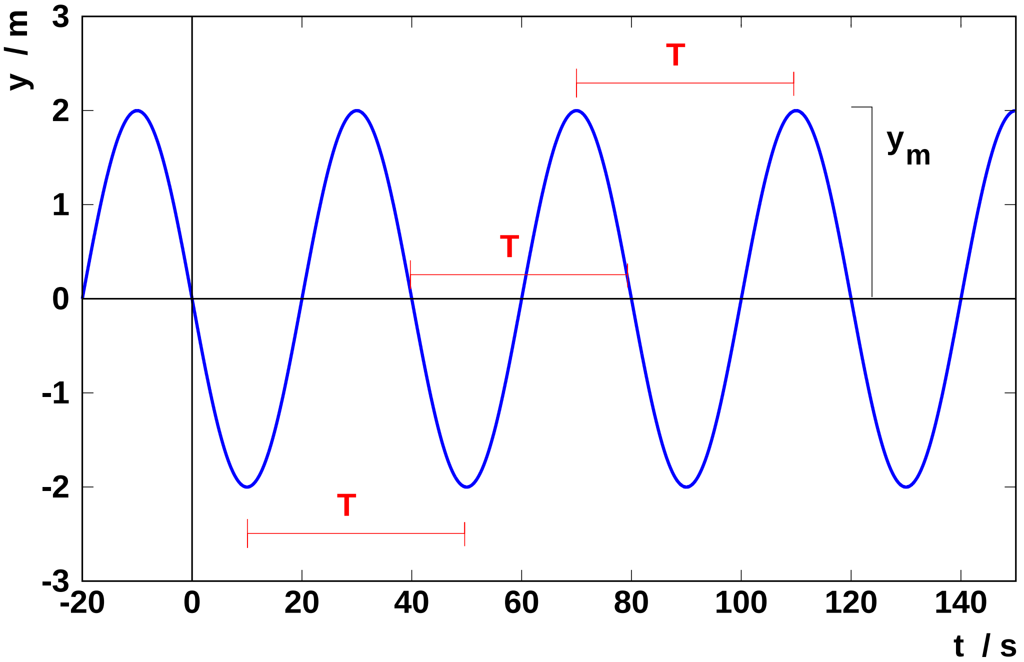
$$y = y(x,t) = y(x_0,t)$$

Take the same wave as before. Two observations:

1. it again is a sine
2. it "goes down" first

$$y = -y_m \sin(\omega t) = y_m \sin(-\omega t)$$

The time it takes from one crest to the next is called the period T.



Again: $\sin(\omega t) = \sin(\omega(t + T))$

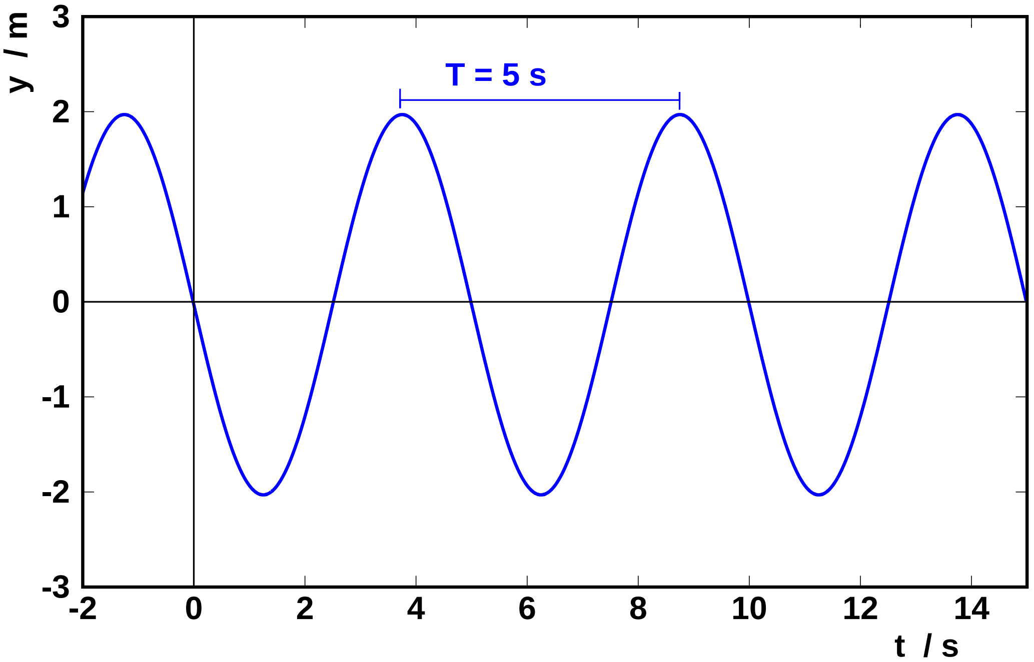
$$\omega T = 2\pi \Longrightarrow \omega = \frac{2\pi}{T}$$

Define $f = 1/T = \frac{\omega}{2\pi}$

Units:

$$\begin{aligned} \text{period}[T] &= \text{s} \\ \text{angular frequency}[\omega] &= \text{rad/s} \\ \text{frequency}[f] &= \text{s}^{-1} \text{ or Hz} \end{aligned}$$

1.6: Example:



We measure $T = 5\,\mathrm{s}$.

Thus $\omega = 2\pi/T = 1.26\,\mathrm{rad/s}$

and $f = 1/T = 0.2\,\mathrm{Hz}$.

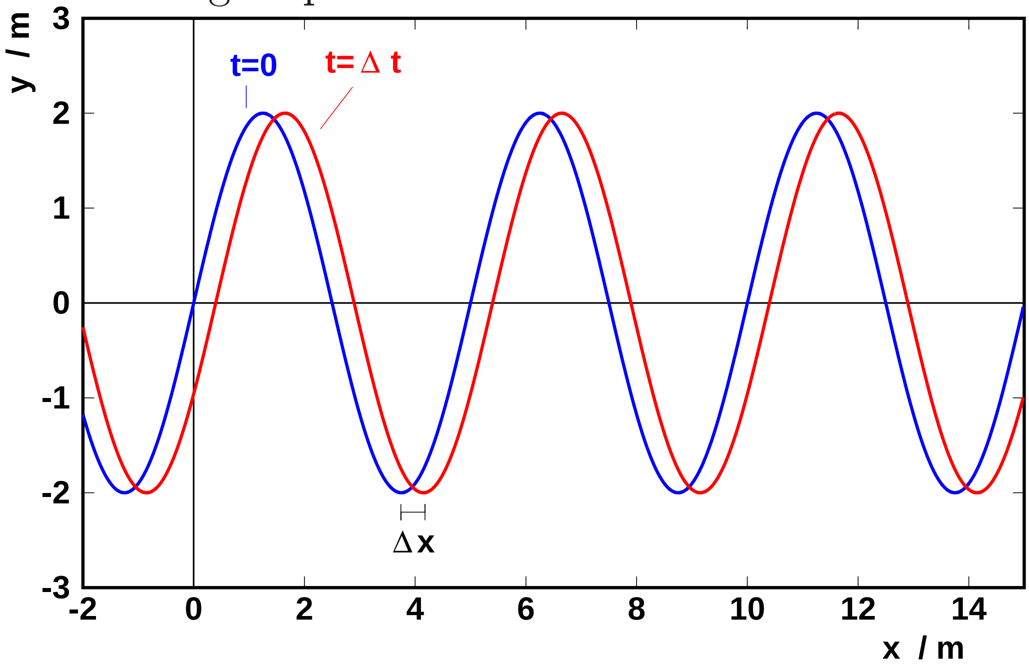
1.7: The complete wave:

How can we combine these things into one description for the entire wave?

$$y(0,t) = y_m \sin(-\omega t)$$

$$y(x,0) = y_m \sin(kx)$$

Look at a single "point on the wave".



The wave moved by Δx in the time Δt. The speed is therefore

$$\boxed{v = \frac{\Delta x}{\Delta t}} \text{ [m/s]}$$

Or:

$$\boxed{v = \frac{\mathrm{d}x}{\mathrm{d}t}} \text{ [m/s]}$$

If the point stays fixed on the wave then the argument of the sine function must remain constant.

If x increases as t increases, then choosing $kx - \omega t$ as the argument of the sin will have the desired effect:

$$\boxed{y(x,t) = y_m \sin(kx - \omega t - \phi_0)}$$

Note the importance of the phase constant ϕ_0. We will see that k depends on the frequency and the medium through which the wave travels. That leaves y_m and ϕ_0 to be determined by initial conditions. This makes sense: *You* decide how big a wave you want to create (y_m) and you also determine which point on the wave you consider as your point of reference (ϕ_0).

1.8: Wave speed:

When we say “the wave moves”, what is actually moving? The only real motion is up and down, not forward!

The position of a point with a certain phase changes with time, and it is that speed that is the wave speed.

The phase of the wave is given by

$$kx - \omega t - \phi_0 = const.$$

Differentiate with respect to t:

$$k\frac{\mathrm{d}\,x}{\mathrm{d}\,t} - \omega = 0 \Longrightarrow \frac{\mathrm{d}\,x}{\mathrm{d}\,t} = \frac{\omega}{k}$$

$$\boxed{v = \frac{\mathrm{d}\,x}{\mathrm{d}\,t} = \frac{\omega}{k}}$$

Check the dimension:

$$\frac{\frac{\mathrm{rad}}{\mathrm{s}}}{\frac{\mathrm{rad}}{\mathrm{m}}} = \frac{m}{s}$$

1.9: Example:

Consider the wave

$$y(x,t) = (0.1\,m)\sin(2x - 4t + \pi/3)$$

What are $y_m, T, \lambda, k, f, \kappa, \omega, \phi_0$?

Easy:

$$\begin{aligned}
y_m &= 0.1\,\mathrm{m} \\
k &= 2\mathrm{rad/m} \\
\omega &= 4\mathrm{rad/s} \\
f &= \omega/2\pi = \frac{4\,\mathrm{rad/s}}{2\pi\,\mathrm{rad}} = 0.636\,\mathrm{Hz} \\
T &= 1/f = 1.571\,\mathrm{s} \\
\lambda &= 2\pi/k = \frac{2\pi\,\mathrm{rad}}{2\,\mathrm{rad/m}} = 3.14\,\mathrm{m} \\
\kappa &= 1/\lambda = 0.318\,\mathrm{m}^{-1} \\
\phi_0 &= -\pi/3
\end{aligned}$$

What is the wave speed?

$$v = \frac{\omega}{k} = \frac{\lambda}{T} = \lambda f = \frac{f}{\kappa} = 2\,\frac{\mathrm{m}}{\mathrm{s}}$$

What if we want to move in the other direction?

we need *decreasing* x with increasing t:

$$kx + \omega t - \phi_0 = const$$
$$v = \frac{\mathrm{d}\,x}{\mathrm{d}\,t} = -\frac{\omega}{k}$$

just as we wanted.

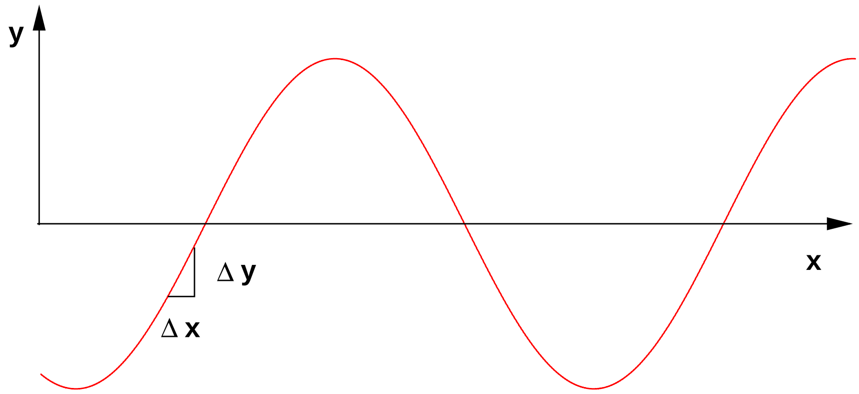
So the two types of sinusoidal waves are:

$$y(x,t) = y_m \sin(kx - \omega t - \phi_0) \Longrightarrow$$
$$y(x,t) = y_m \sin(kx + \omega t - \phi_0) \Longleftarrow$$

1.10: Physical waves

Consider a real physical system, a very long, elastic, massive string in x direction. At the point $x = 0$ we move the string up and down in a periodic sine motion. The string has a mass density $\mu = m/l$ and a tension τ and the wave is described by $y(x,t) = y_m \sin(kx - \omega t)$.

Each element of the string travels up and down periodically.



The kinetic energy for a string element of length Δx is

$$\Delta E_{kin} = \frac{m}{2} v^2 = \frac{\Delta x \mu}{2} \left(\frac{\mathrm{d}\, y}{\mathrm{d}\, t} \right)^2$$

$$\frac{\mathrm{d}\, y}{\mathrm{d}\, t} = y_m \cos(kx - \omega t) \cdot (-\omega)$$

Thus

$$\boxed{\Delta E_{kin} = \frac{\Delta x \mu \omega^2 y_m^2}{2} \cos^2(kx - \omega t)}$$

The potential energy in the stretched part of the string is ΔE_{pot} =(tension in string) · (elongation of string)

$$\begin{aligned}
\Delta E_{pot} &= \tau(\Delta l - \Delta x) \\
\Delta l^2 &= \Delta x^2 + \Delta y^2 \\
\Delta l - \Delta x &= \sqrt{\Delta x^2 + \Delta y^2} - \Delta x \\
\Delta l - \Delta x &= \Delta x(\sqrt{1 + \frac{\Delta y}{\Delta x}^2} - 1)
\end{aligned}$$

For small oscillations we can use the Taylor expansion:

$$\sqrt{1+\alpha^2} - 1 = 1 + \frac{\alpha^2}{2} - \frac{\alpha^4}{8} + \cdots - 1 \simeq \frac{\alpha^2}{2}$$

Finally use $\frac{\Delta y}{\Delta x} = \frac{\mathrm{d}\,y}{\mathrm{d}\,x} = y_m k \cos(kx - \omega t)$ and collect it all:

$$\begin{aligned}
\Delta E_{pot} &= \tau(\Delta l - \Delta x) \\
&= \tau(\Delta x(\sqrt{1 + \left(\frac{\Delta y}{\Delta x}\right)^2} - 1)) \\
&= \tau \Delta x \frac{1}{2} \left(\frac{\Delta y}{\Delta x}\right)^2 \\
&= \frac{\tau y_m^2 k^2 \Delta x}{2} \cos^2(kx - \omega t)
\end{aligned}$$

$$\boxed{\Delta E_{pot} = \frac{\tau y_m^2 k^2 \Delta x}{2} \cos^2(kx - \omega t)}$$

The average energy in the wave can be found by integrating the energy over one period and dividing by that period:

$$\langle E \rangle = \frac{1}{T} \int_{t=0}^{t=T} E(x,t) \mathrm{d}\, t$$

$$\langle \Delta E_{kin} \rangle = \frac{\Delta x \mu \omega^2 y_m^2}{2} \frac{1}{T} \int_0^T \cos^2(kx - \omega t) \mathrm{d}\, t$$

This integral is easy. Since $\sin(kx - \omega t)$ looks exactly like $\cos(kx - \omega t)$, only a little further left, the integral over one period will be the same.

$$\begin{aligned}
A &= \frac{1}{T} \int_0^T \cos^2(kx - \omega t) \mathrm{d}\, t \\
A &= \frac{1}{T} \int_0^T \sin^2(kx - \omega t) \mathrm{d}\, t \\
2A &= \frac{1}{T} \int_0^T (\sin^2(kx - \omega t) + \cos^2(kx - \omega t)) \mathrm{d}\, t \\
2A &= \frac{1}{T} \int_0^T 1 \mathrm{d}\, t \\
A &= \frac{T}{2T} = \frac{1}{2}
\end{aligned}$$

This leaves us with:

$$\langle \Delta E_{kin} \rangle = \frac{\Delta x \mu \omega^2 y_m^2}{4}$$
$$\langle \Delta E_{pot} \rangle = \frac{\Delta x \tau k^2 y_m^2}{4}$$

Recall from the mechanics course:

In every oscillating system, the average kinetic energy equals the average potential energy.

$$\langle \Delta E_{kin} \rangle = \langle \Delta E_{pot} \rangle$$
$$\frac{\Delta x \mu \omega^2 y_m^2}{4} = \frac{\Delta x \tau k^2 y_m^2}{4}$$
$$\mu \omega^2 = \tau k^2 \Longrightarrow \frac{\omega}{k} = \sqrt{\frac{\tau}{\mu}} = v$$

So the speed of a wave on a string with mass density μ and tension τ is given by

$$\boxed{v = \sqrt{\frac{\tau}{\mu}}}$$

1.11: Summary:

transversal sinusoidal waves

$$y(x,t) = y_m \sin(kx \mp \omega t - \Phi_0)$$

"-": Travels in positive x-direction
"+": Travels in negative x-direction

Amplitude y_m
Wavelength λ [m]
angular wave number $k = 2\pi/\lambda$ [rad/m]
wave number $\kappa = 1/\lambda$ [1/m]
Period T [s]
angular frequency ω [rad/s]
frequency f [1/s] or [Hz]
Wave speed $v = \frac{\omega}{k} = \frac{\lambda}{T} = \lambda f [m/s]$

Wave on a string with tension τ and mass density μ:
Wave speed $v = \sqrt{\frac{\tau}{\mu}}$

1.12: Exercises:

Halliday Resnick & Walker:

Reading: HRW Chapter 17 pp 370–382 (Waves I)

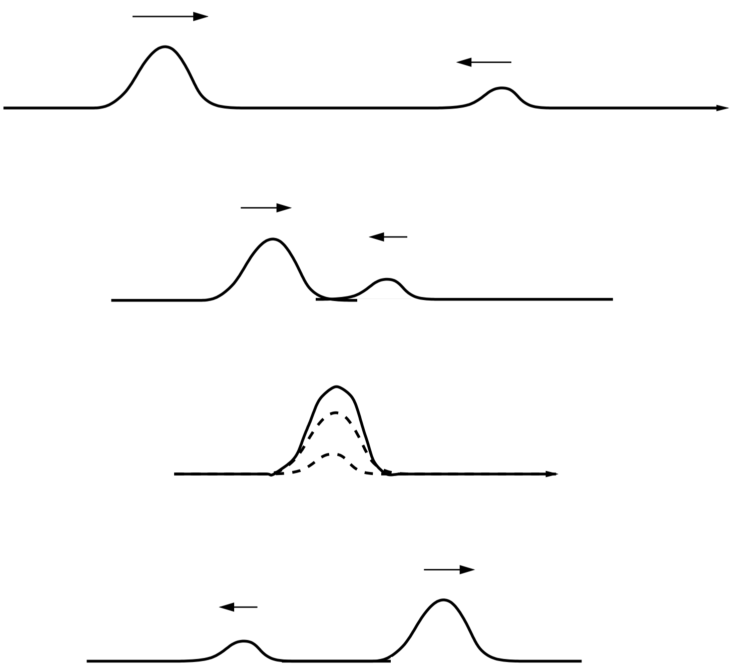
Exercises: p.392 ff.: Q3, Q4, 3E, 4E, 6P, 24E, 25P

1.13: Principle of superposition

If two (or more) waves travel through the same region of space, the net displacement at each point is the sum of the displacements due to the individual waves.

This is the **Principle of Superposition** for waves.

$$y(x,t) = y_1(x,t) + y_2(x,t)$$



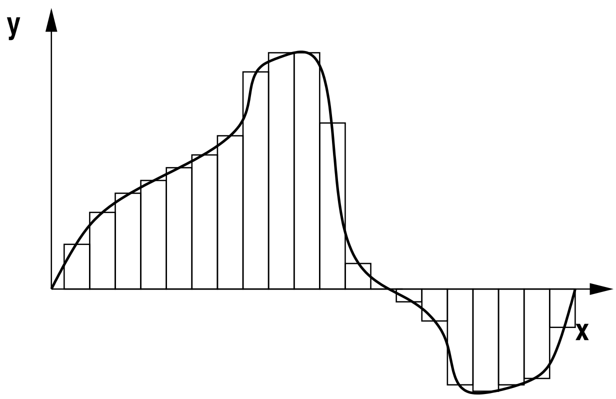
1.14: Superposition of waves

So far we only looked at sinusoidal waves. How general is that?

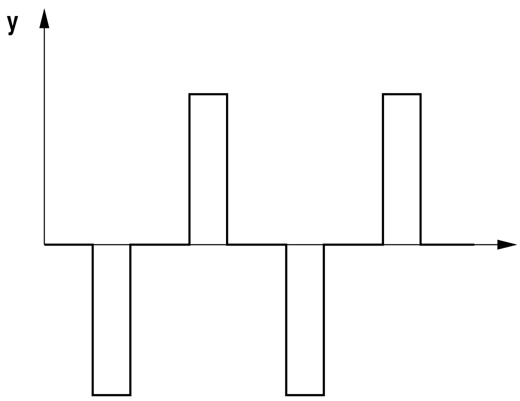
Any function can be treated as a superposition of (infinitely) many sinusoidal waves.

To find these component sinusoidal waves one does a **Fourier analysis**.

We can build up any function as a sequence of rectangular pieces:

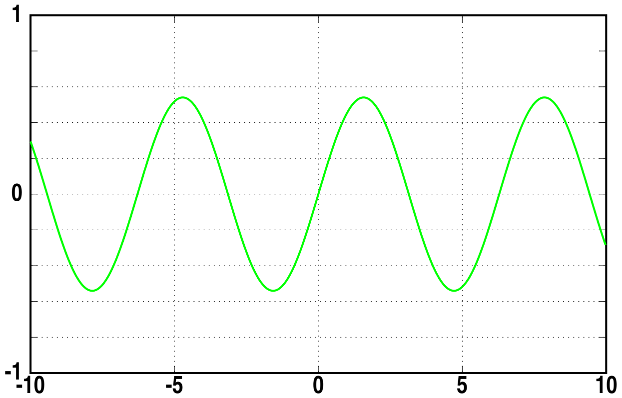


And we can build a rectangular piece out of sine waves:

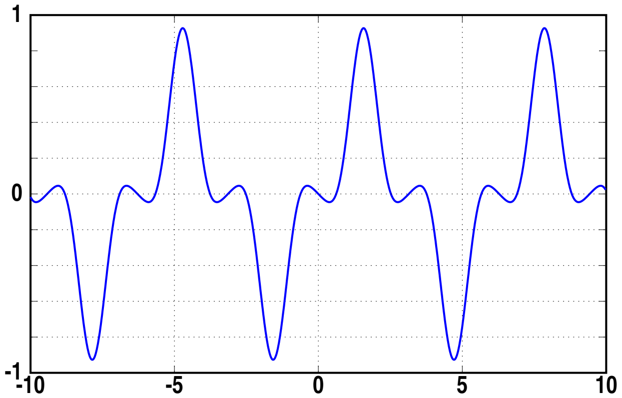


Consider $\sum_{n=0}^{\infty} \frac{\cos(2n+1)}{(2n+1)} \sin((2n+1)x - (2n+1)t)$:

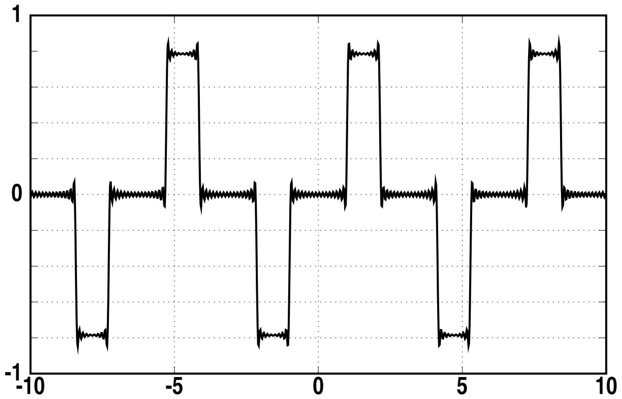
1 Term:



2 Terms:



25 Terms:



1.15: Wave speed

What about the speed of this composite wave?

$$\begin{aligned} y(x,t) &= \cos(1)\sin(x-t) + \frac{\cos(3)}{3}\sin(3x-3t) \\ &+ \frac{\cos(5)}{5}\sin(5x-5t) + \cdots \end{aligned}$$

$$v = \frac{\omega}{k} = \frac{1}{1} = \frac{3}{3} = \frac{5}{5} = 1\text{m/s}$$

All constituent waves travel at the same speed.

1.16: Conclusions

1) We can decompose any periodic or nonperiodic wave or pulse into an (infinite) number of sinusoidal waves.

2) We can observe **Interference**.

Take two waves travelling along a string:

$$y_1(x,t) = y_0 \sin(kx - \omega t)$$

$$y_2(x,t) = y_0 \sin(kx - \omega t - \Phi)$$

We call Φ the phase difference between the two waves: "They are out of phase by Φ".

define $\alpha = kx - \omega t$, $\beta = kx - \omega t - \Phi$.
$(\alpha + \beta)/2 = kx - \omega t - \Phi/2$
$(\alpha - \beta)/2 = +\Phi/2$

Use $\sin(\alpha) + \sin(\beta) = 2\sin((\alpha+\beta)/2)\cos((\alpha-\beta)/2)$

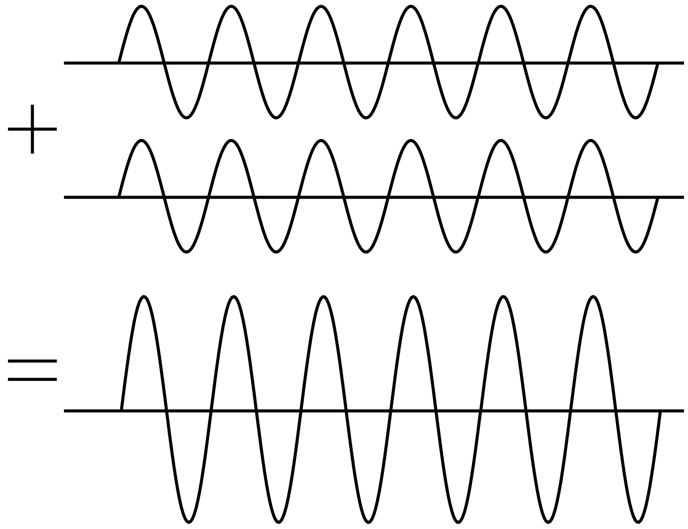
$$\begin{aligned} y(x,t) &= y_1(x,t) + y_2(x,t) \\ &= y_0 \sin(kx - \omega t) + y_0 \sin(kx - \omega t - \Phi) \\ &= 2y_0 \cos(\Phi/2) \sin(kx - \omega t - \Phi/2) \end{aligned}$$

New amplitude $y_m = 2y_0 \cos(\Phi/2)$! The new wave has a different amplitude and a different phase.

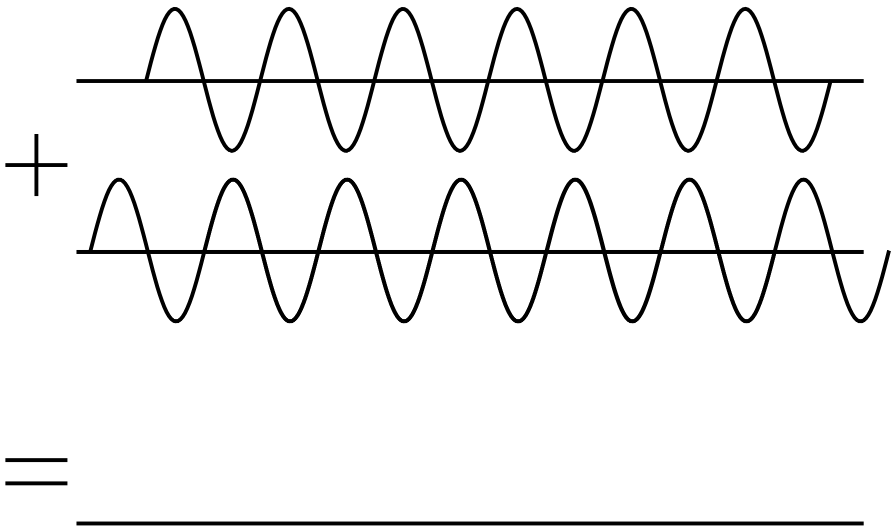
1.17: Interference

We can now distinguish several cases:

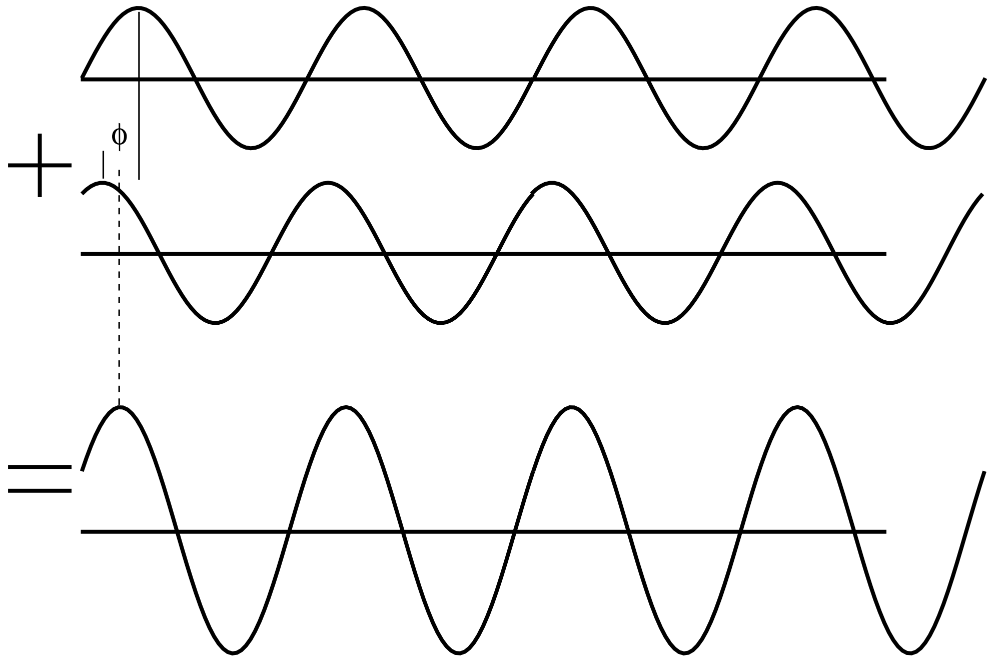
1) $\Phi = 0$: The waves are "in phase", we have $y_m = 2y_0 \cos(0) = 2y_0$.



2) $\Phi = \pi$: The waves are "out of phase", we have $y_m = 2y_0 \cos(\pi/2) = 0$.

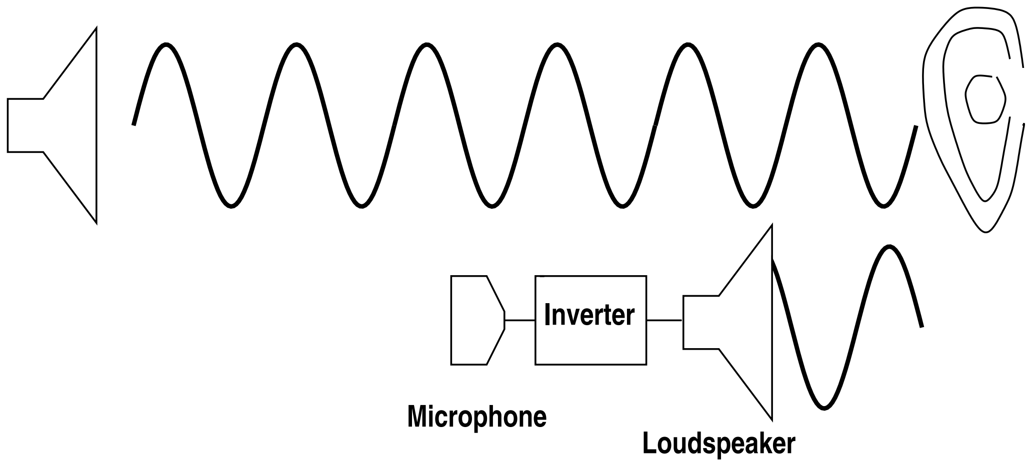


Generally we will have cases in between these extremes:



1.18: Applications of interference

Active earprotection

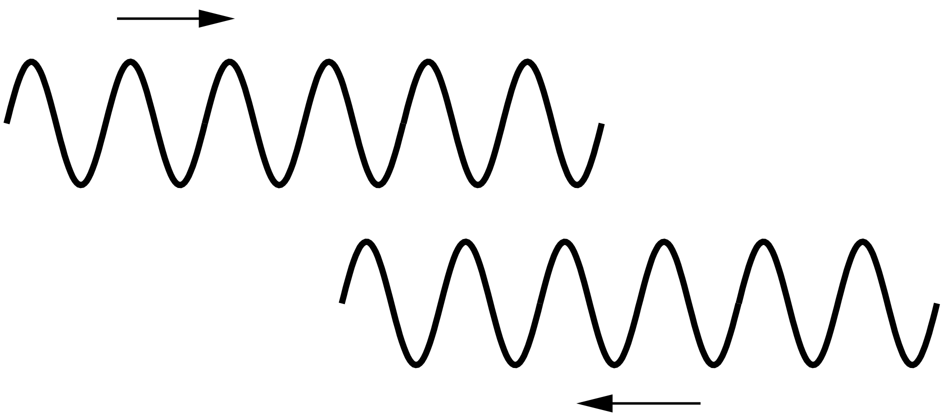


If the microphone is 5 cm away from the ear and the loudspeaker is 2 cm away from the ear then we have 0.1 ms to invert the wave - plenty of time!

1.19: Standing waves I

So far these waves moved in the same direction. What is they move in opposite directions?

$y_1(x,t) = y_0 \sin(kx - \omega t)$
$y_2(x,t) = y_0 \sin(kx + \omega t - \Phi)$

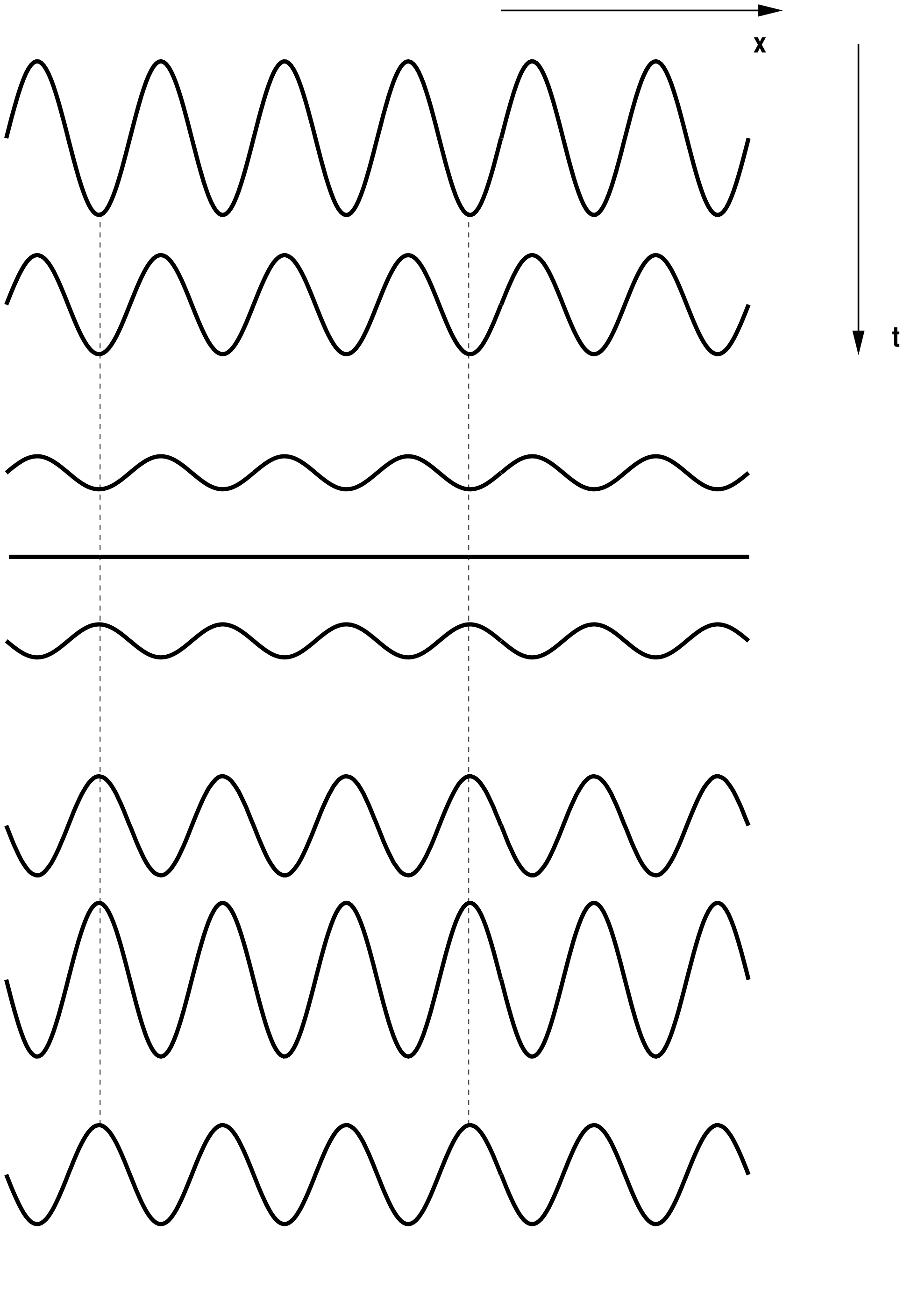


define $\alpha = kx - \omega t$, $\beta = kx + \omega t - \Phi$ and use
$\sin(\alpha) + \sin(\beta) = 2\sin((\alpha+\beta)/2)\cos((\alpha-\beta)/2)$

$y(x,t) = y_1(x,t) + y_2(x,t) =$
$2y_0 \sin(kx - \Phi/2)\cos(\omega t + \Phi/2)$

This is not a travelling wave anymore, it is a **Standing wave**.

1.20: Standing waves II



1.21: Summary

Principle of Superposition: If two (or more) waves travel through the same region of space, the net displacement at each point is the sum of the displacements due to the individual waves.

We can build up any periodic or nonperiodic wave from sinusoidal waves

Interference:
If two sinusoidal waves with the same wavelength travel in the same direction the resulting wave is again a sinusoidal wave with the same wavelength but with a different amplitude and phase:

$$\begin{aligned} y(x,t) &= y_0 \sin(kx - \omega t) + y_0 \sin(kx - \omega t - \Phi) \\ &= 2y_0 \cos(\Phi/2) \sin(kx - \omega t - \Phi/2) \end{aligned}$$

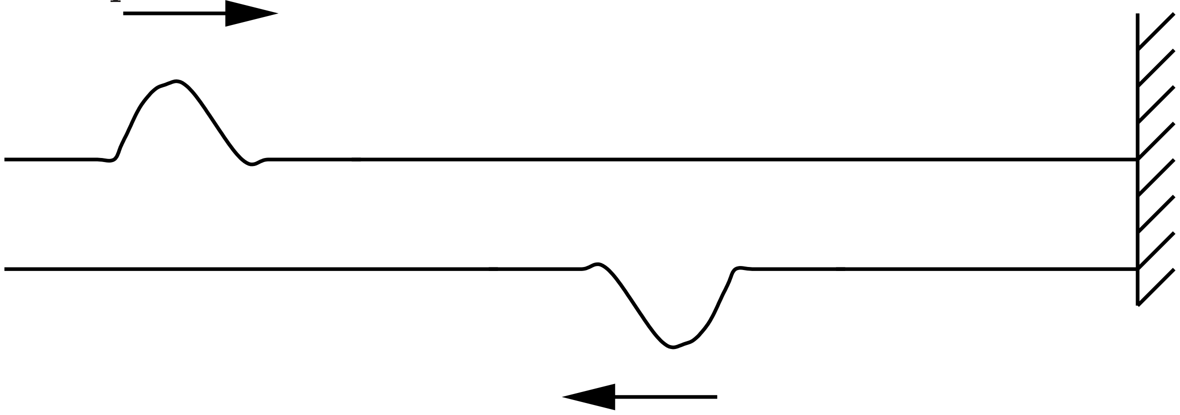
Standing waves:
If two sinusoidal waves with the same wavelength travel in opposite directions the result is a standing wave:

$$\begin{aligned} y(x,t) &= y_0 \sin(kx - \omega t) + y_0 \sin(kx + \omega t - \Phi) \\ &= 2y_0 \sin(kx - \Phi/2) \cos(\omega t + \Phi/2) \end{aligned}$$

1.22: Reflections

What happens if a wave encounters an obstacle?

The puls returns **inverted**!

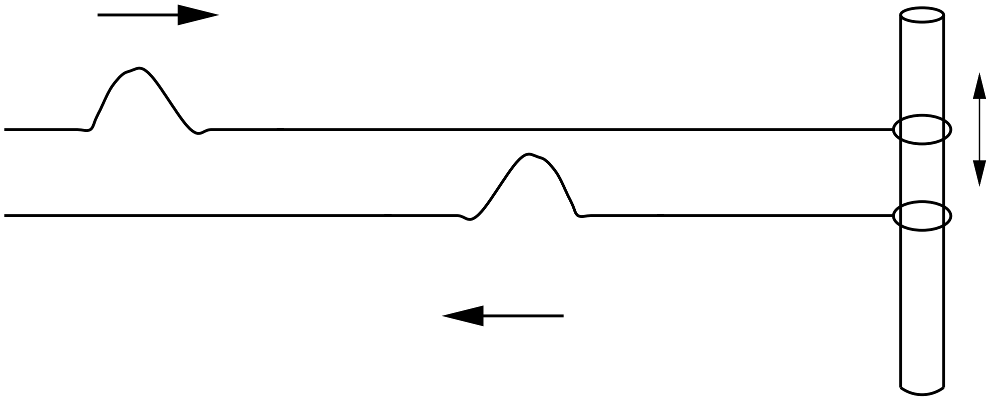


Why?

The string exerts a force on the wall. According to Newtons principle actio = reactio the wall exerts an equal, but **opposite** force on the string: The pulse is inverted and sent back.

We call this type of reflection a "hard" reflection.

The other type of reflection is a "soft" reflection: Consider this string which is attached to a frictionless bearing that slides along a rod:



There is no force on the rod, therefore the rod does not exert any force back on the string.

The end of the string moves, stretching the string. This then generates a restoring force proportional to the elongation of the string. The pulse is sent back, but not inverted.

1.23: Momentum transfer

What is the amount of momentum transferred to the rod/wall in both cases:

1) Hard reflection: The string element Δx closest to the wall has a momentum Δp. As the pulse is reflected from the wall, that momentum is reversed: $\Delta p \to -\Delta p$. The total change of momentum in the string element is $-\Delta p - \Delta p = -2\Delta p$. As momentum is conserved, the wall must have experienced a change in momentum of equal magnitude but opposite direction: $+2\Delta p$.

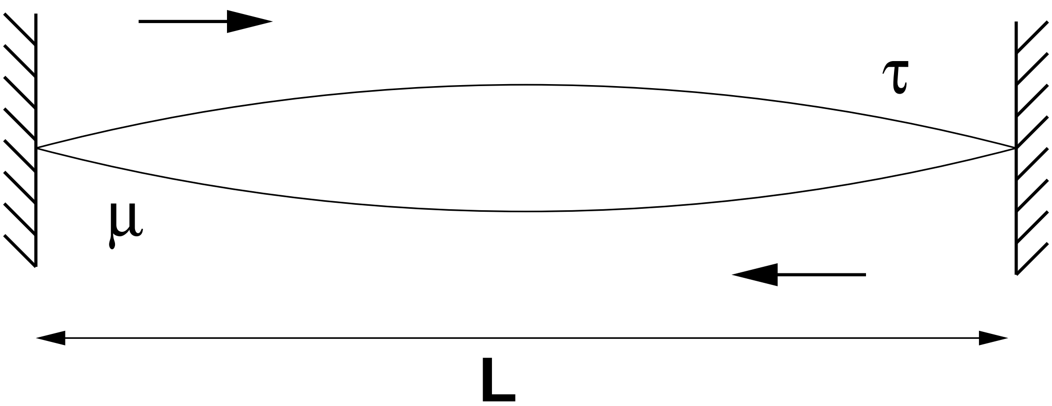
2) Soft reflection: Since the bearing moves frictionless along the rod, the rod never feels any kind of force. Therefore no momentum is transferred to the rod.

This is the difference between "hard" and "soft" reflection: "hard" reflection transfers momentum to the "mirror", "soft" reflection does not.

1.24: Standing waves III

If we want to set up a standing wave, we have to use two waves with the same wavelengths travelling in opposite directions on the same string. The easiest way to create such a thing is to use a wave and trap it between two walls.

Take a string with a length L and a mass density $\mu = m/L$ at a tension τ. If we want to have a standing sinusoidal wave on this string, we must have a node at both ends.

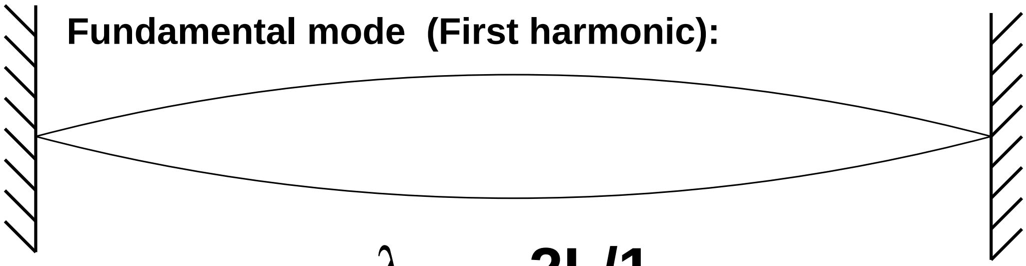


This means, the length of the string only allows certain wavelengths to form standing waves:

$$L_1 = \lambda/2$$

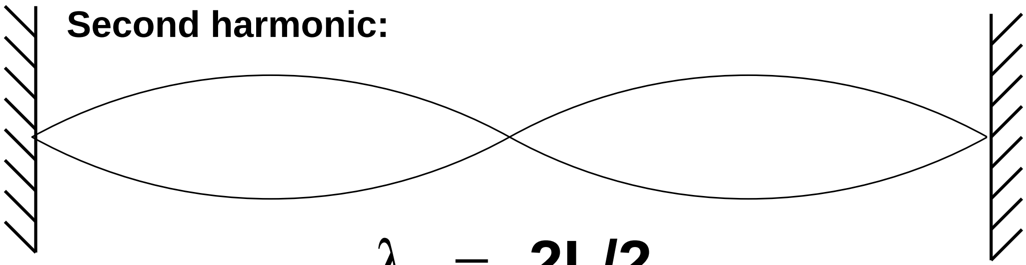
But we can squeeze more nodes onto the string:

Fundamental mode (First harmonic):



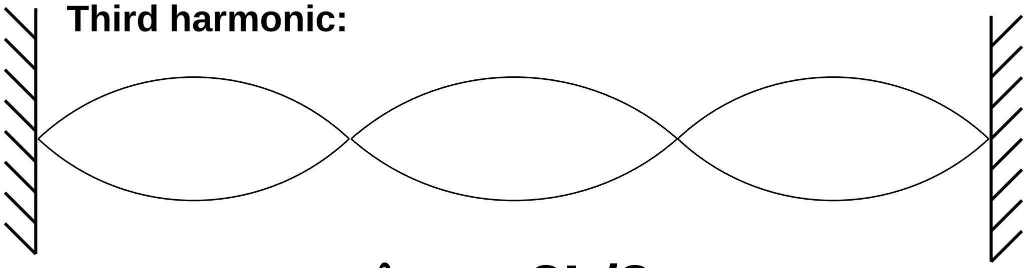
$$\lambda_1 = 2L/1$$

Second harmonic:



$$\lambda_2 = 2L/2$$

Third harmonic:



$$\lambda_3 = 2L/3$$

$$\lambda_n = 2L/n$$

The length of the string determines the wavelength, the tension and mass density determine the wave speed, therefore the frequency is fixed for any mode. Obvious: A guitar string will sound at a given frequency. To change the pitch you must shorten the string (normal play) or change the tension in the string (tuning).

Example: A bass string has a length of 1 m with a mass density of 2.8 g/m. What tension is required to tune the string to a frequency of 55 Hz?

Solution:

$$\lambda = 2L = 2\,\mathrm{m}$$
$$f\lambda = v = \sqrt{\frac{\tau}{\mu}}$$
$$\tau = \mu(f\lambda)^2$$

$$\tau = 2.8 \times 10^{-3}\mathrm{kg/m}(55\mathrm{s}^{-1}2\mathrm{m})^2 = 33.9\mathrm{N}$$

If you have the tension too high by 1 N, how will the pitch change?

Solution:

$$f = v/\lambda = \frac{\sqrt{\tau/\mu}}{\lambda}$$
$$f = \frac{\sqrt{34.9\mathrm{N}/2.8 \times 10^{-3}\mathrm{kg/m}}}{2\mathrm{m}} = 55.8Hz$$

Basses are easy to tune.

1.25: Resonance

Take the string and excite it with a frequency of your choosing: You still will have a wave travelling back and forth, but the backward and forward moving waves will not be in phase, you do not get a nice standing wave.

Only if the frequency approaches the allowed (fundamental and harmonic) frequencies does a standing wave emerge. We say the string is in resonance.

Take the bass string of the previous example: The fundamental mode had a frequency of $f_1 = 55\,\text{Hz}$. What other resonance frequencies are there?

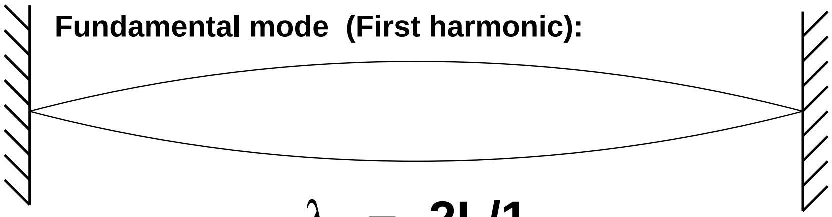
The wave speed is constant for all waves on that string. The wavelengths are given by $\lambda_n = 2L/n$.

$$f_n = v/\lambda_n = \frac{nv}{2L} = nf_1$$

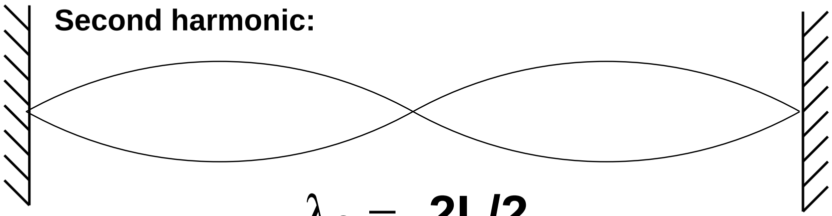
$$f_2 = 2 * 55\text{Hz} = 110\text{Hz}$$

$$f_3 = 3 * 55\text{Hz} = 165\text{Hz}$$

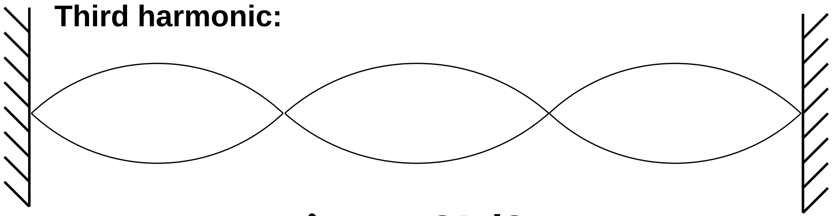
1.26: Summary



$$\lambda_1 = 2L/1$$



$$\lambda_2 = 2L/2$$



$$\lambda_3 = 2L/3$$

$$\lambda_n = 2L/n$$

$\lambda = 2L/n$ where n is the number of antinodes in the standing wave.

The speed of the wave is determined by the tension and the mass density of the string. This together with the length of the string determines the fundamental and harmonic frequencies.

A wave undergoing “hard”reflection returns inverted.

A wave undergoing “soft” reflection returns upright.

1.27: Exercises

Halliday Resnick & Walker:

Reading: HRW p. 382–391 (Chapter 17: Waves I)

Exercises: p. 392 ff.: Q5, Q9, 29P, 33E, 37E, 46P

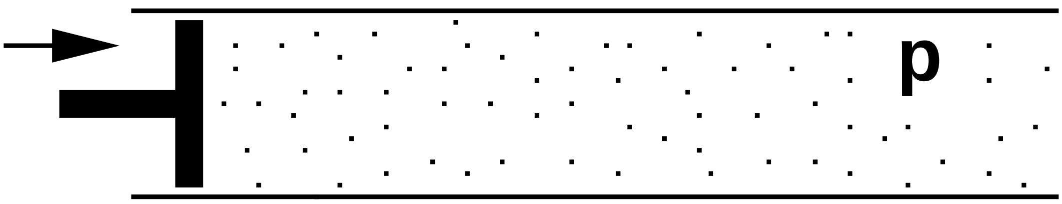
2: Sound

All waves on the string were transversal waves and were possible because the vibrating string can store elastic energy by stretching.

Sound travels through air which is not elastic, so we need a different mechanism for the propagation of sound waves.

Sound waves are longitudinal pressure waves!

Take a tube filled with air at a pressure p:

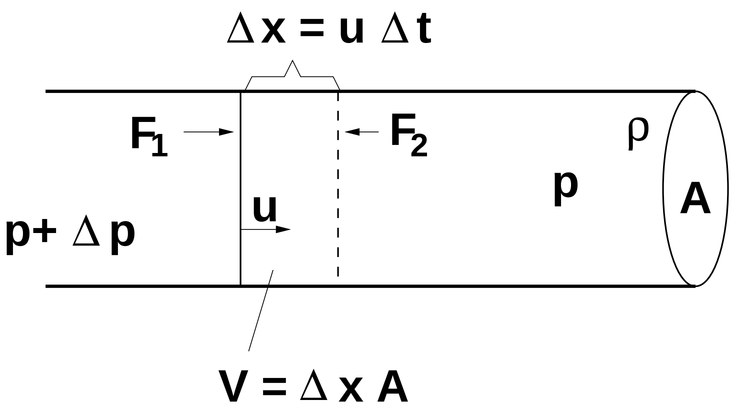


and push on the piston. This will displace the air molecules next to it and increase the pressure. Then air will flow out from this high pressure region pushing on the adjacent molecules - a pressure wave forms.

2.1: Speed of sound

We shall use Newton's 2^{nd} law $F = ma$ to find the speed of sound.

A step in pressure travels through the tube.



The net force on the shaded volume of air is

$$F = F_2 - F_1 = pA - (p + \Delta p)A = -\Delta pA$$

The mass of that volume of air is

$$m = \rho\Delta V$$

The acceleration is

$$a = \frac{\Delta u}{\Delta t} = \frac{u - 0}{\Delta t} = \frac{u}{\Delta t}$$

This gives

$$F = ma \Longleftrightarrow -\Delta pA = \rho\Delta V \frac{u}{\Delta t}$$

To relate the change in pressure to a change in volume we use the **Bulk Modulus** B. It is a material constant and has the dimension of a pressure: Pa (Pascal).

$$\boxed{\Delta p = -B\frac{\Delta V}{V}}$$

We then have

$$-\Delta p A = B\frac{\Delta V}{V}A = \rho \Delta V \frac{u}{\Delta t}$$

$$\frac{BA}{V} = \rho\frac{u}{\Delta t}$$

The Volume we are looking at is $V = A\Delta x = Au\Delta t$

$$\frac{B}{u\Delta t} = \rho\frac{u}{\Delta t}$$

$$\boxed{u = \sqrt{\frac{B}{\rho}}}$$

This looks familiar:

$$\text{speed} = \sqrt{\frac{\text{elastic parameter}}{\text{inertial parameter}}}$$

The elastic parameter stores potential energy,
The inertial parameter stores kinetic energy.

2.2: Speed of sound II

In a medium with bulk modulus B and mass density ρ, the speed of sound is

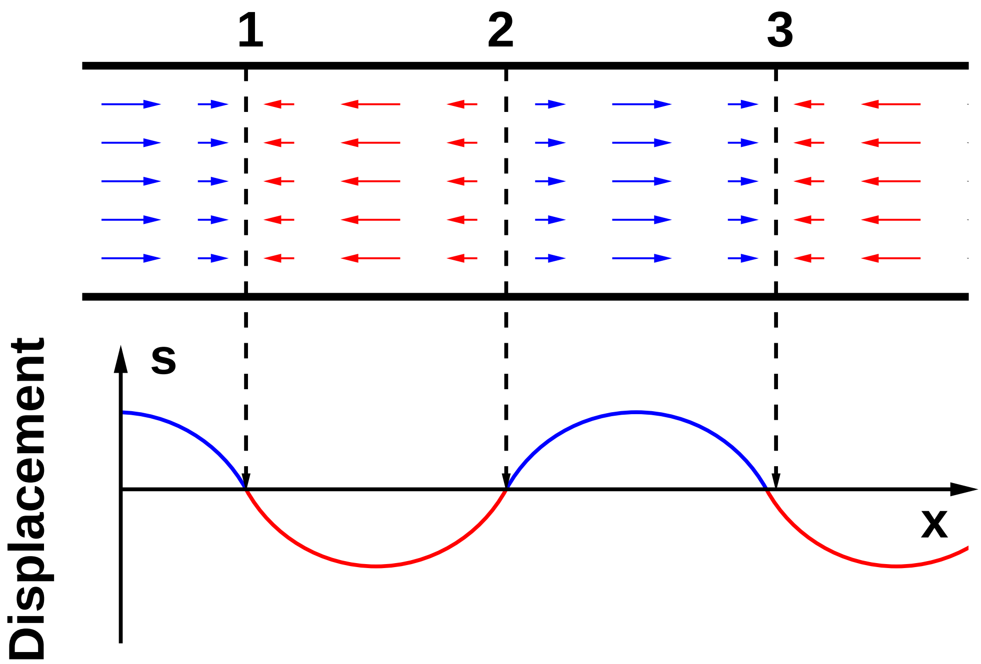
$$u = \sqrt{\frac{B}{\rho}}$$

Examples:

Material	Bulk Modulus	Density	Speed of sound
Air	142 kPa	$1.2\,\mathrm{kg/m^3}$	344 m/s
Helium	179 kPa	$0.18\,\mathrm{kg/m^3}$	1000 m/s
Water	2.05 GPa	$1\,\mathrm{g/cm^3}$	1430 m/s
Steel	165 GPa	$7.8\,\mathrm{g/cm^3}$	4600 m/s

2.3: Sound as longitudinal waves

Let us look closer at the actual motion of the air:



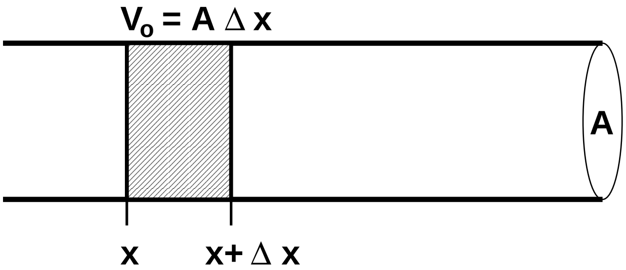
What about the pressure? Intuitively we would say that the pressure is high in regions 1 and 3 since all the air is flowing toward these regions and lowest for region 2.

Assume for the displacement

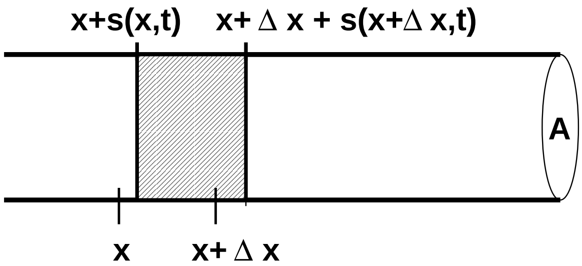
$$s(x,t) = s_0 \cos(kx - \omega t)$$

with $s_0 << \lambda$.

The pressure at each point can be obtained by looking at the change of volume brought on by the wave:



This volume is deformed by the wave:



$$\Delta V = (\Delta x + s_0 \cos(k(x+\Delta x) - \omega t) - s_0 \cos(kx - \omega t))A - V_0$$

Use

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

$$\begin{aligned} \Delta V &= A\Delta x + As_0 \cos(kx - \omega t)\cos(k\Delta x) \\ &\quad - As_0 \sin(kx - \omega t)\sin(k\Delta x) \\ &\quad - As_0 \cos(kx - \omega t) - A\Delta x \end{aligned}$$

We had a very thin slice, therefore $k\Delta x << 1$ and we can approximate $\cos(k\Delta x) \simeq 1$ and $\sin(k\Delta x) \simeq k\Delta x$.

This leaves us with

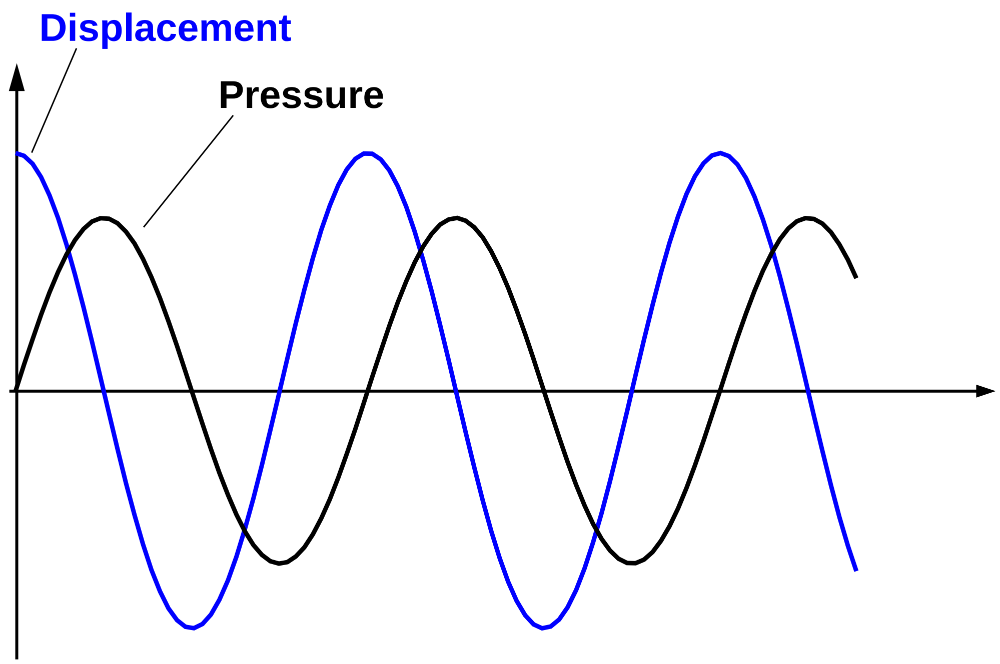
$$\begin{aligned}\Delta V &= -ks_0 A\Delta x \sin(kx-\omega t)\\ &= -ks_0 V_0 \sin(kx-\omega t)\end{aligned}$$

Again use the bulk modulus to relate ΔV and Δp and collect it all ($\Delta p = -B\Delta V/V$):

$$\Delta p = s_0 Bk \sin(kx-\omega t) = p_0 \sin(kx-\omega t)$$

$$\boxed{p_0 = s_0 Bk = s_0 \rho\omega u}$$

The last step uses $u = \omega/k = \sqrt{B/\rho}$.

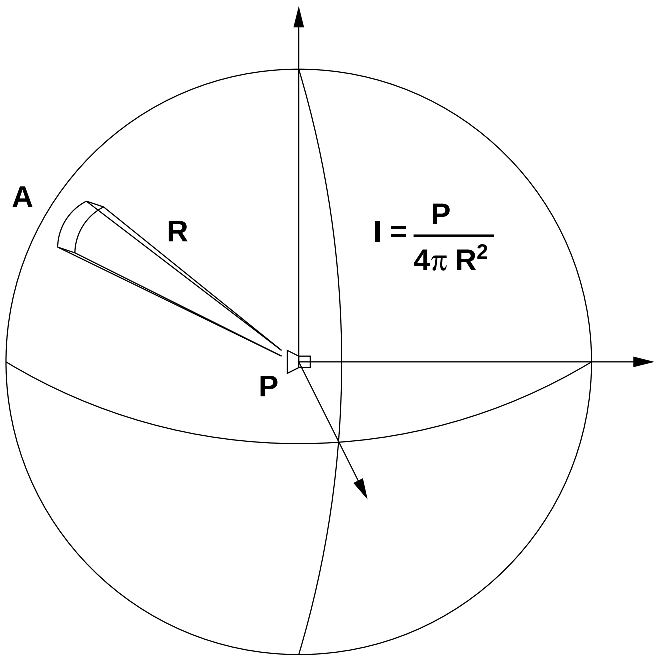


2.4: Sound intensity

What is the power transmitted by a sound wave?

Sound is not restricted to a one dimensional string, it will radiate in 3 dimensions. We define the sound intensity I observed at a distance R from the source as the power received per unit area:

$$I = \frac{P}{4\pi R^2}$$



The displacement at a distance r from the source is then

$$s(r,t) = s_0 \sin(kr - \omega t - \Phi_0)$$

$r = \sqrt{x^2 + y^2 + z^2}$

The kinetic energy in a thin spherical shell around the source is

$$\Delta E_{kin} = \frac{\Delta m}{2}\left(\frac{\mathrm{d}\, s(r,t)}{\mathrm{d}\, t}\right)^2$$

$$\frac{\mathrm{d}}{\mathrm{d}\, t} s(r,t) = s_0 \sin(kr - \omega t - \Phi_0) = -\omega s_0 \cos(kr - \omega t - \Phi_0)$$

$$\Delta E_{kin} = \frac{\Delta m \omega^2 s_0^2}{2} \cos^2(kr - \omega t - \Phi_0)$$

The next steps are completely analogue to the discussion we had for waves on strings:

1) Calculate the average kinetic energy over one period.

$$\langle \Delta E_{kin} \rangle = \frac{\Delta m \omega^2 s_0^2}{4}$$

2) The average potential energy is the same as the average kinetic energy, therefore the average total energy is twice the average kinetic energy:

$$\langle \Delta E_{tot} \rangle = \frac{\Delta m s_0^2 \omega^2}{2}$$

compare wave on a string:

$$\langle \Delta E_{tot} \rangle = \frac{\mu \Delta x y_m^2 \omega^2}{2}$$

For a spherical shell with radius R and thickness Δr we have $\Delta m = 4\pi R^2 \Delta r \rho$

$$\langle \Delta E_{tot} \rangle = \frac{4\pi R^2 \Delta r \rho s_0^2 \omega^2}{2}$$

3) Compute the power (energy transmitted per time)

$$\begin{aligned} P &= \frac{\langle \Delta E_{tot} \rangle}{\Delta t} \\ &= \frac{4\pi R^2 \rho s_0^2 \omega^2}{2} \frac{\Delta r}{\Delta t} \\ &= \frac{4\pi R^2 \rho s_0^2 \omega^2 u}{2} \end{aligned}$$

This gives the intensity as

$$I = \frac{P}{4\pi R^2} = \frac{\rho s_0^2 \omega^2 u}{2}$$

Sound intensity is measured in W/m^2.

2.5: Example

We generate sound with a radio. The power of the radio is $P_s = 0.25\,\mathrm{W}$, and we listen to it at a distance of $R = 1\,\mathrm{m}$. What is the sound intensity at the ear?

$$I = \frac{P_s}{4\pi R^2} = \frac{0.25\mathrm{W}}{12.56(1\mathrm{m})^2} = 19.9\mathrm{mW/m^2}$$

If the average human ear has a size of $A = 50\,\mathrm{cm}^2$, how much power reaches the ear?

$$P = \frac{P_s}{4\pi R^2}A = \frac{0.25\mathrm{W}}{12.56(100\mathrm{cm})^2}50\mathrm{cm}^2 = 0.1\mathrm{mW}$$

Is that a lot?

We cannot really answer that question because we do not yet have a scale against which to measure sound intensities. So we must create one.

The softest sounds that the human ear can hear has a pressure amplitude of $2.8 \times 10^{-5}\,\mathrm{Pa}$ at a frequency of $1000\,\mathrm{Hz}$ (very low whisper).

The loudest sounds we can tolerate (at the pain threshold) has a pressure amplitude of $28\,\mathrm{Pa}$ - that is a starting jumbo jet from a few meters beside the runway!

What are the displacement amplitudes for these sounds?

We use $p_0 = s_0 \rho \omega u$.

$f = 1000\,\mathrm{Hz}$

$u = 340\,\mathrm{m/s}$

$\rho = 1.2\,\mathrm{kg/m^3}$

$p_0 = 2.8 \times 10^{-5}\,\mathrm{Pa}$

This gives for the whisper

$$s_0 = p_0/(\rho \omega u) = 1.1 \times 10^{-11} \frac{\mathrm{Pa\,m^3 s^2}}{\mathrm{kg}}$$

$$s_0 = 1.1 \times 10^{-11}\mathrm{m}$$

and for the jumbo jet

$$s_0 = 1.1 \times 10^{-5}\mathrm{m} = 11\mu\mathrm{m}$$

The ear is sensitive over 6 orders of magnitude of pressure!

2.6: Summary

- Soundwaves are longitudinal pressure waves.
- The speed of sound u is related to the bulk modulus B and the mass density ρ via $u = \sqrt{B/\rho}$.
- The bulk modulus relaes the change in volume due to a change in pressure: $\Delta p = -B\Delta V/V$
- The displacement amplitude s_0 and the pressure amplitude p_0 are related via $p_0 = s_0\rho\omega u$.
- The sound intensity is the power radiated by the source per unit area: $I = P/(4\pi R^2)$.
- The sound intensity is related to the displacement amplitude s_0 via $I = \rho S_0^2\omega^2 u/2$.

2.7: Exercises

Halliday Resnick & Walker:

Reading: HRW p. 398-404

Exercises: p. 420 ff.: Q3, Q10, Q12, 4E, 9E, 17E, 27P

2.7: The decibel scale

The human ear covers an enormous range of sound intensities: The lowest whispers have an intensity $I = 10^{-12}\,\mathrm{W/m^2}$, while the sound of a starting jet has an intensity of $1\,\mathrm{W/m^2}$. That is 12 orders of magnitude.

It is therefore useful to introduce a logarithmic scale: the decibel scale.

It is not possible to take logarithms of physical quantities which have units:

$$1\,\mathrm{kg} = 1000\,\mathrm{g}$$

$$\begin{aligned} \log_{10}(1\,\mathrm{kg}) &= \log_{10}(1000\,\mathrm{g}) \\ 0 &= 3 \qquad\qquad ???? \end{aligned}$$

It is possible to take the logarithm of a **ratio** of physical quantities.

We can define the lowest audible sound intensity $I_0 = 10^{-12}\,\mathrm{W/m^2}$. Then the definition

$$\beta = (10\,\mathrm{dB})\log_{10}(I/I_0)$$

makes sense.

$1\,\mathrm{dB} = 1$ decibel $= 0.1$ bel

The unit **bel** was named after Alexander Graham Bell.

The decibel scale

Sound	Sound Intensity Level [dB]	Intensity [W/m^2]	Pressure amplitude [Pa]	Displacement amplitude [m]
Jet takeoff	130	10	90	3.6×10^{-5}
Construction site	120	1	28	1.1×10^{-5}
Rock concert	110	0.1	9	3.6×10^{-6}
Shout (1.5m)	100	0.01	2.8	1.1×10^{-6}
Heavy Truck (15m)	90	0.001	0.9	3.6×10^{-7}
Urban Street	80	10^{-4}	0.28	1.1×10^{-7}
Automobile interior	70	10^{-5}	0.09	3.6×10^{-8}
Normal Conversation	60	10^{-6}	0.028	1.1×10^{-8}
Large Office	50	10^{-7}	0.009	3.6×10^{-9}
Living Room	40	10^{-8}	0.0028	1.1×10^{-9}
Bedroom at night	30	10^{-9}	0.0009	3.6×10^{-10}
Broadcast studio	20	10^{-10}	0.00028	1.1×10^{-10}
Rustling leaves	10	10^{-11}	9.0×10^{-5}	3.6×10^{-11}
Lowest audible whisper	0	10^{-12}	2.8×10^{-5}	1.1×10^{-11}

How do we work with the decibel scale?

Remember basic logarithms:

$$\log(ab) = \log(a) + \log(b) \qquad \log(a/b) = \log(a) - \log(b)$$

$$\log(a^b) = b\log(a)$$

One loundspeaker plays music at an intensity $I = 10^{-4}\,\mathrm{W/m^2}$. The sound intensity level is

$$\begin{aligned}\beta &= (10\,\mathrm{dB})\log_{10}\left(\frac{10^{-4}\,\mathrm{W/m^2}}{10^{-12}\,\mathrm{W/m^2}}\right)\\ &= (10\,\mathrm{dB})\log_{10}(10^8) = 80\,\mathrm{dB}\end{aligned}$$

The loudspeaker doubles its intensity:

$$\begin{aligned}\beta &= (10\,\mathrm{dB})\log_{10}\left(\frac{2\times 10^{-4}\,\mathrm{W/m^2}}{10^{-12}\,\mathrm{W/m^2}}\right)\\ &= (10\,\mathrm{dB})\log_{10}(2\times 10^8)\\ &= (10\,\mathrm{dB})\log_{10}(2) + (10\,\mathrm{dB})\log_{10}(10^8)\\ &= 3\,\mathrm{dB} + 80\,\mathrm{dB} = 83\,\mathrm{dB}\end{aligned}$$

Doubling the intensity only adds 3 dB.

Two sounds have sound intensity levels $\beta_1 = 50\,\mathrm{dB}$ and $\beta_2 = 30\,\mathrm{dB}$. What is the ratio of the two sound intensities?

$$\begin{aligned}
\beta_1 - \beta_2 &= (10\,\mathrm{dB})\log_{10}\left(\frac{I_1}{I_0}\right) - (10\,\mathrm{dB})\log_{10}\left(\frac{I_2}{I_0}\right) \\
50\,\mathrm{dB} - 30\,\mathrm{dB} &= (10\,\mathrm{dB})\left[\log_{10}\left(\frac{I_1}{I_0}\right)\right] - \log_{10}\left(\frac{I_2}{I_0}\right) \\
20\,\mathrm{dB} &= (10\,\mathrm{dB})\log_{10}\left(\frac{I_1}{I_0}\frac{I_0}{I_2}\right) \\
20\,\mathrm{dB} &= (10\,\mathrm{dB})\log_{10}\left(\frac{I_1}{I_2}\right) \\
2 &= \log_{10}\left(\frac{I_1}{I_2}\right) \\
100 &= \frac{I_1}{I_2}
\end{aligned}$$

The ratio of the two sound intensities is $I_1 : I_2 = 100 : 1$.

What are the sound intensities for these two sounds in $\mathrm{W/m^2}$?

$$\begin{aligned}
50\,dB &= (10\,\mathrm{dB})\log_{10}\left(\frac{I_1}{I_0}\right) \\
5 &= \log_{10}\left(\frac{I_1}{10^{-12}\,\mathrm{W/m^2}}\right) \\
10^5 &= \frac{I_1}{10^{-12}\,\mathrm{W/m^2}} \\
10^{-7}\,\mathrm{W/m^2} &= I_1
\end{aligned}$$

$$\frac{I_1}{I_2} = \frac{100}{1} \Rightarrow I_2 = 10^{-9}\,\mathrm{W/m^2}$$

2.8: The human ear

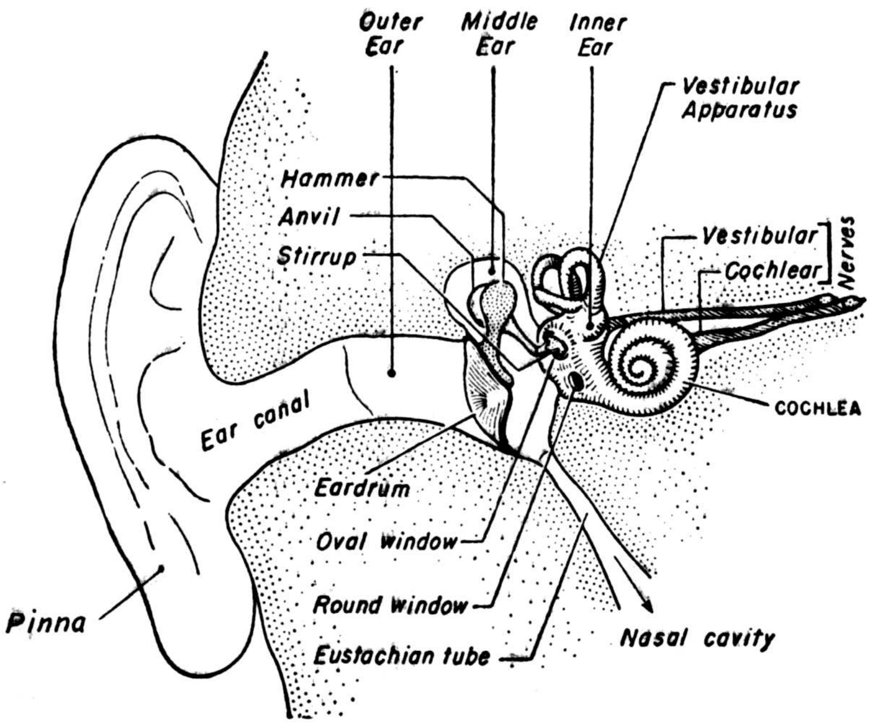


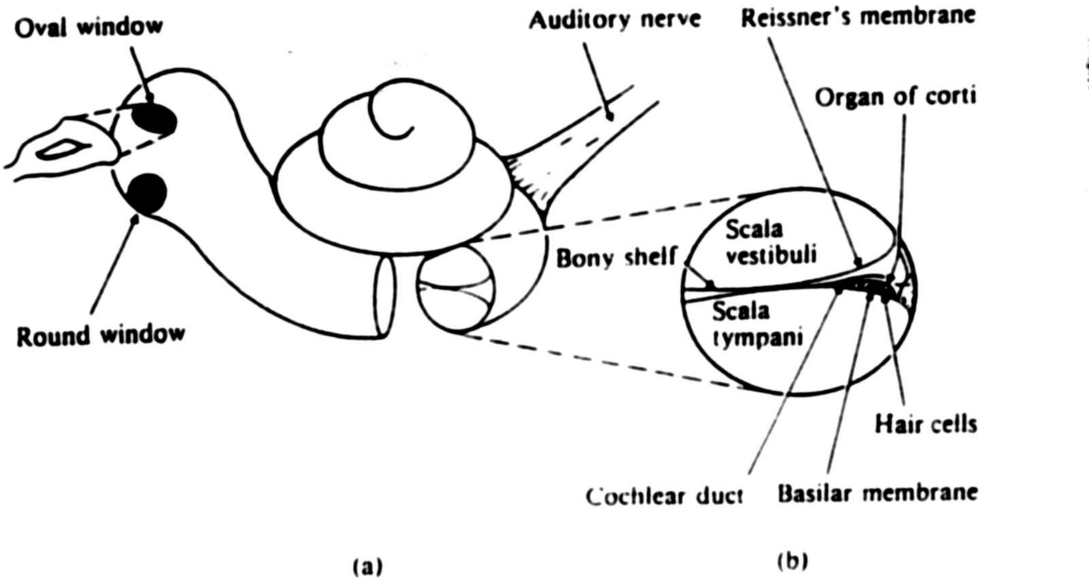
Fig. 8.11. Simplified drawing of the human ear. [After Shortley and Williams, *Elements of Physics*, 2nd ed., Prentice-Hall, Inc., 1955.]

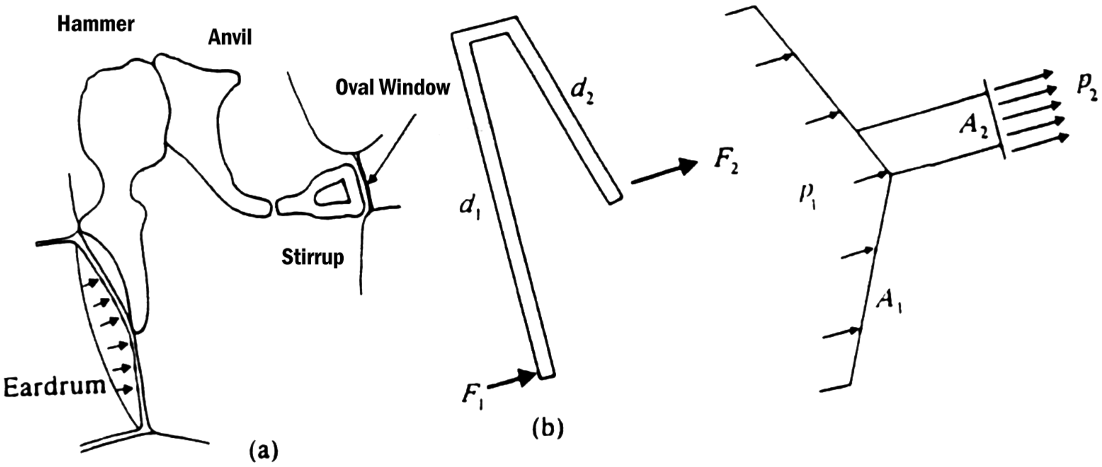
Outer ear: Concentrate the sound onto the eardrum.

Middle ear: Amplify the sound even more, act as safety against too loud noise

Inner ear: Decode the different frequencies, transmit information to the brain.

The pressure on the eardrum is transferred to the oval window on the cochlea by a system of levers (Hammer, Anvil, Stirrup)

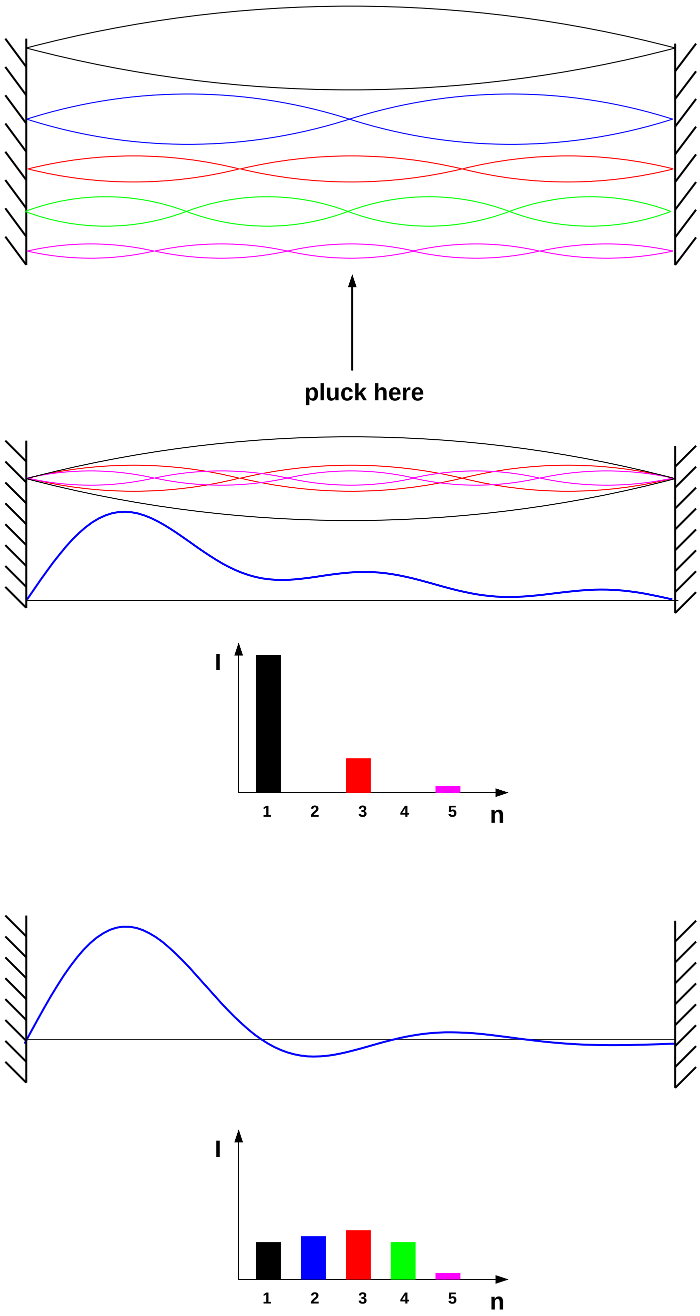


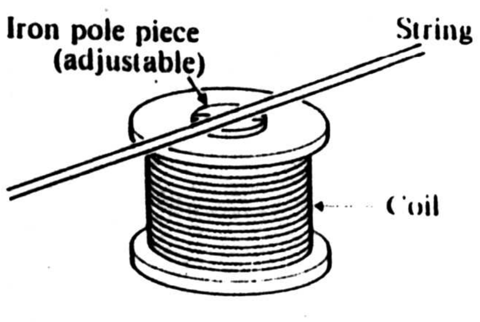


2.9: Physics of musical instruments

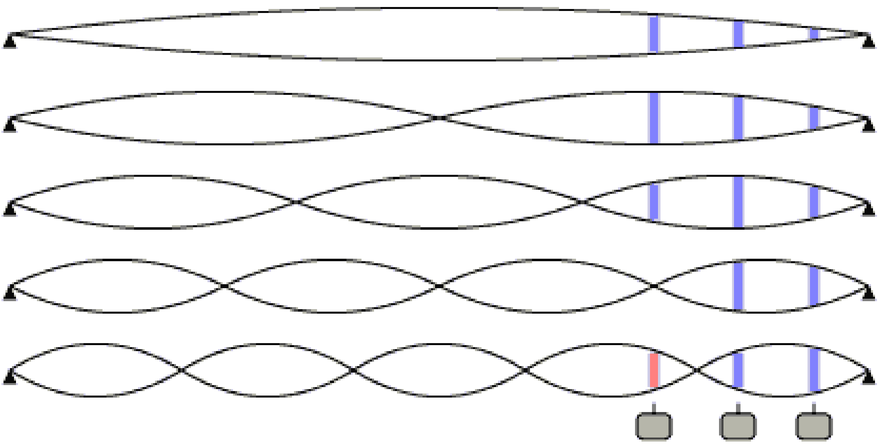
Strings - Pipes - Membranes

Standing waves on a string:

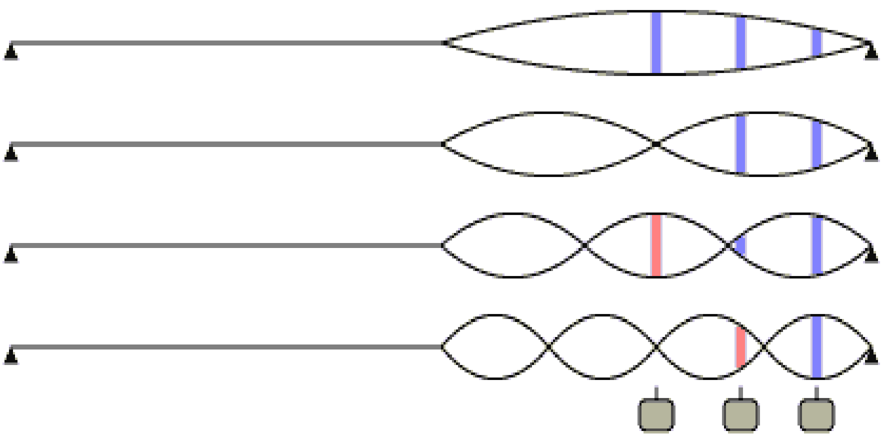




Open String



Fingered String



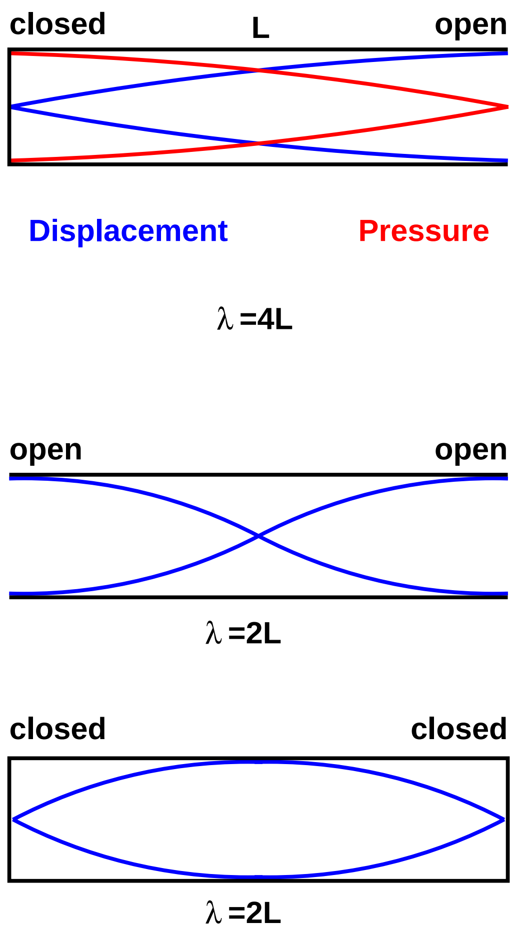
2.10: Pipes

Take a pipe with a closed and an open end and excite a standing soundwave in it.

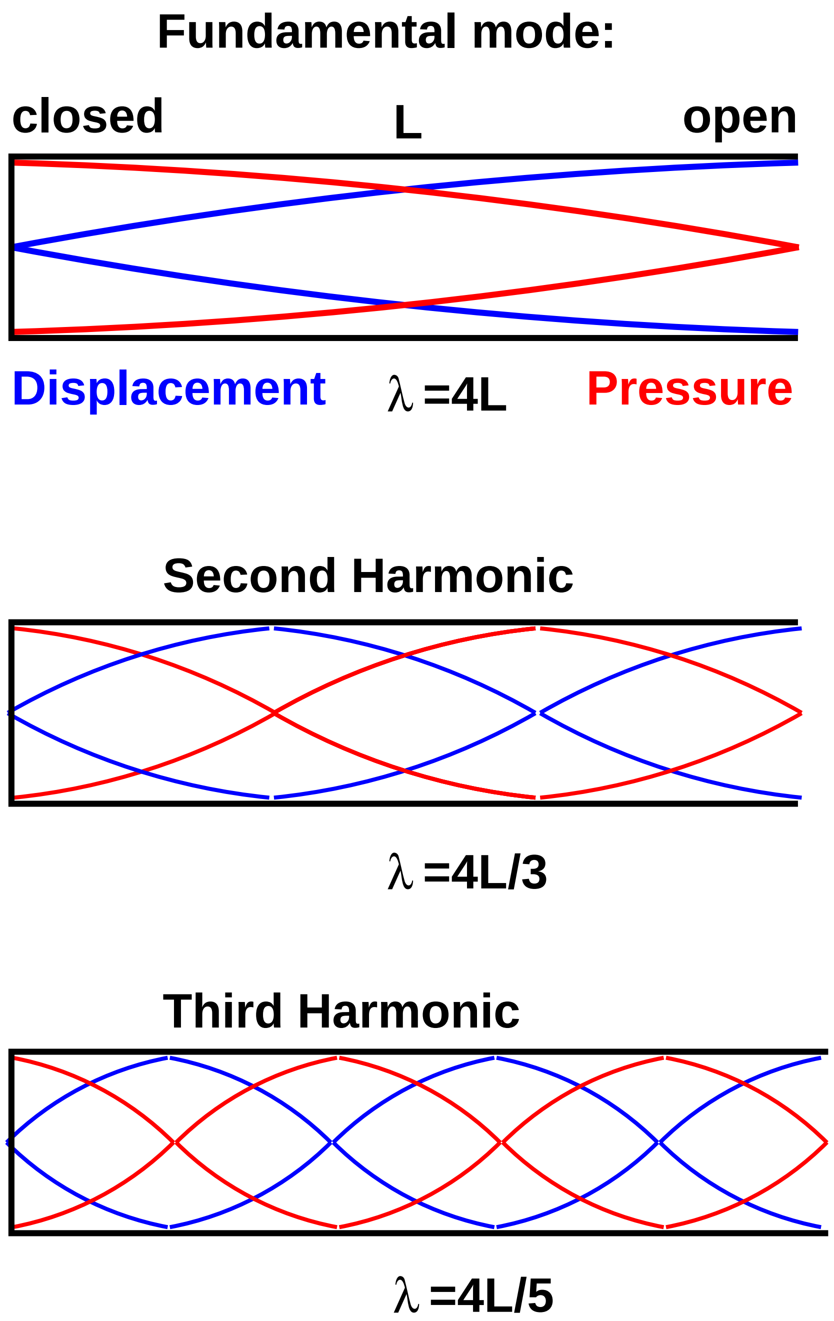
At the closed end the displacement has to vanish.

At the open end, the pressure has to stay the same as in the outside world.

The pressure nodes coincided with the dislacement antinodes:



Of course, all the higher modes are also possible in pipes:

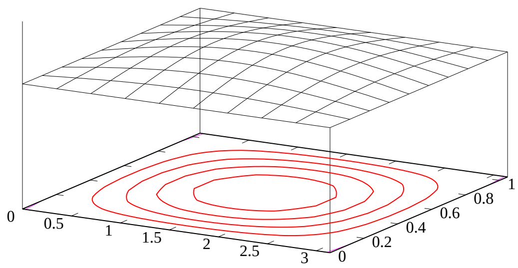


2.11: membranes

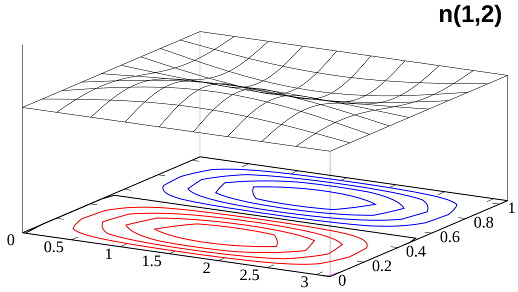
We are not restricted to one-dimensional patterns.

Take a rectangular membrane fixed around the rim:

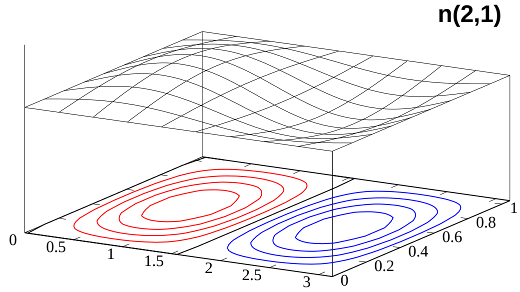
Fundamental mode: n(1,1)



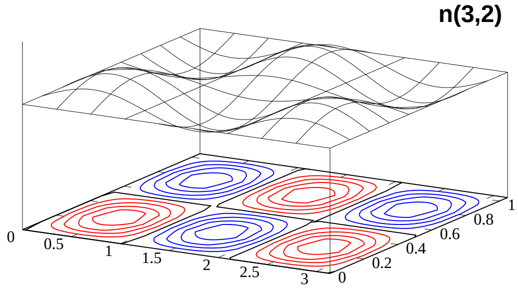
n(1,2)



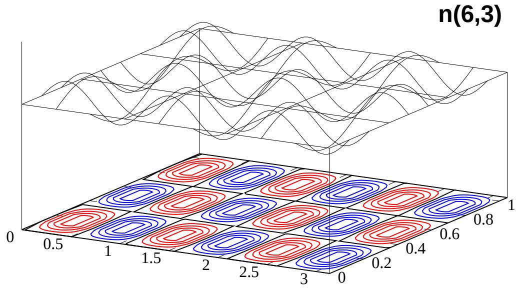
n(2,1)

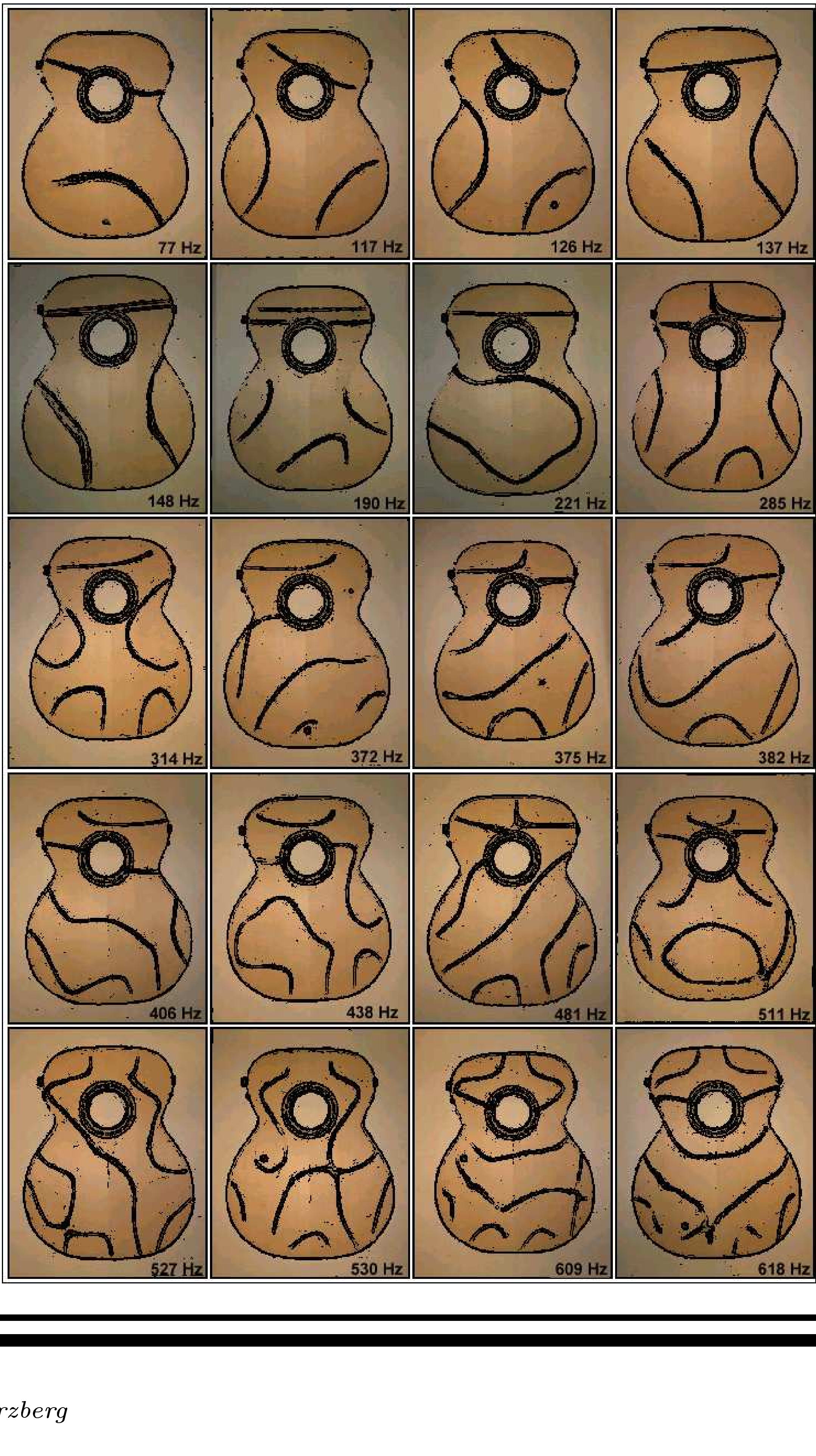


n(3,2)



n(6,3)




77 Hz
117 Hz
126 Hz
137 Hz
148 Hz
190 Hz
221 Hz
285 Hz
314 Hz
372 Hz
375 Hz
382 Hz
406 Hz
438 Hz
481 Hz
511 Hz
527 Hz
530 Hz
609 Hz
618 Hz

2.12: Summary

- The decibel scale is a logarithmic scale for the sound intensity level:

$$\beta = (10\,\mathrm{dB})\log_{10}(I/I_0)$$

- Adding 10 dB to a sound intensity level means the sound intensity has increased by a factor of 10.
- The reference sound intensity is $I_0 = 10^{-12}\,\mathrm{W/m^2}$.
- In a pipe reflection of a soundwave at an open end means the pressure must have a node at the open end.
- Reflection at a closed end means the displacement must have a node there.
- Standing waves on strings, pipes, membranes, are used to build musical instruments.
- The distribution of power into the harmonic modes determines the sound quality.

Exercises

Halliday Resnick & Walker:

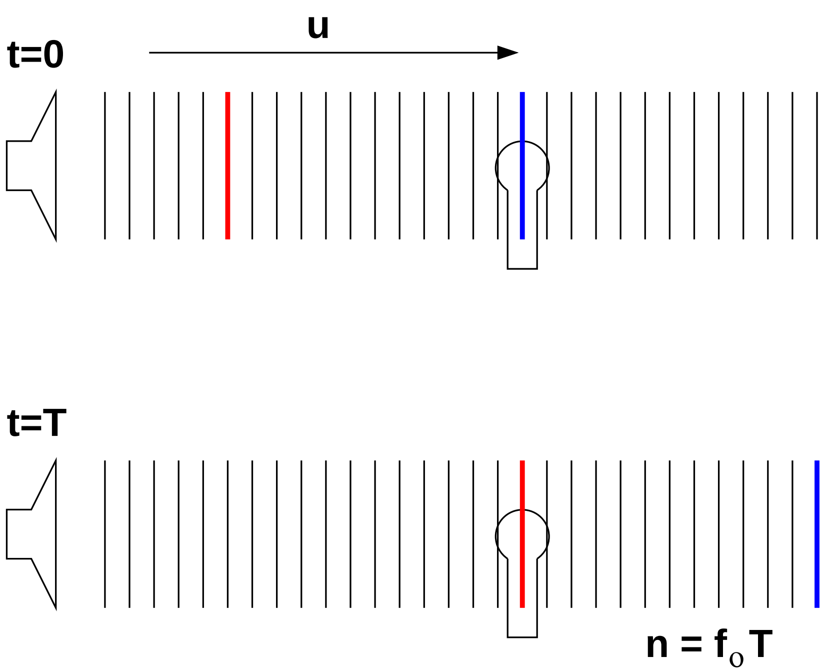
Reading: HRW p. 406-412

Exercises: p. 420 ff.: 17E, 20E, 23E, 26P, 28P, 31E, 33E, 36P

2.13: The Doppler effect

So far we have always assumed that the emitter, the detector and the medium are at rest. That is not generally true. From experience we know that the frequency emitted by a car coming toward us is higher than that of a car moving away from us. That is the Doppler effect.

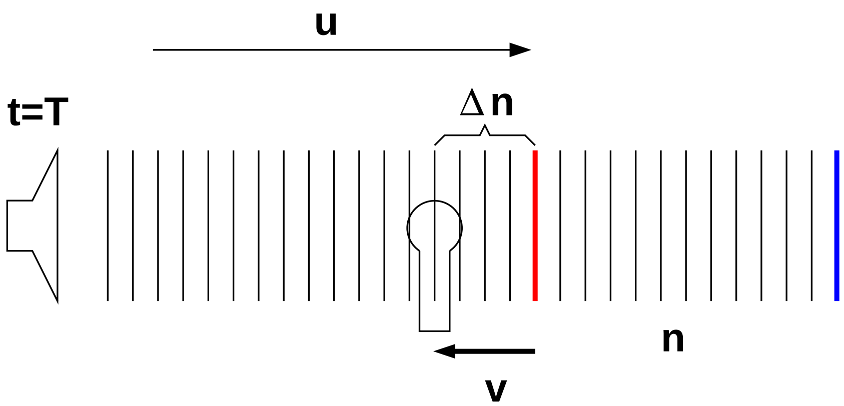
Consider the folluwing setup:



The source emits waves with a frequency $f_0 = 1/T_0$ and a wavelength $\lambda = u/f_0$ with u being the speed of sound.

We assume that the medium (air) is at rest and measure all velocities (of source and detector) relative to the air.

Detector moving, source stationary



The detector moves toward the source at a speed v. After the time T we have encountered $n + \Delta n$ wavelengths!

$$n = \frac{T}{T_0} = f_0 T$$

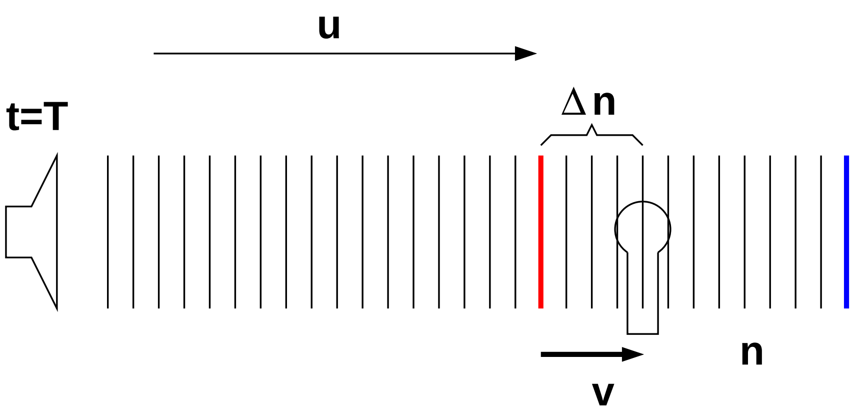
$$\Delta n = \frac{vT}{\lambda} = \frac{vTf_0}{u}$$

So the perceived frequency is

$$f = \frac{n + \Delta n}{T} = \frac{f_0 T + v f_0 T/u}{T} = f_0 \left(1 + \frac{v}{u}\right)$$

higher than the emitted frequency, in accordance with our experience.

If the detector moves away from the source we have a similar picture:



We now count $n - \Delta n$ wavelengths. That gives a frequency

$$f = \frac{n - \Delta n}{T} = \frac{f_0 T - v f_0 T/u}{T} = f_0 \left(1 - \frac{v}{u}\right)$$

lower than the emitted frequency.

In total we have for a moving detector and a stationary source:

$$\boxed{f = f_0 \left(1 \pm \frac{v}{u}\right)}$$

"+": detector moves **toward** the source.
"-": detector moves **away** from the source.

Example

A siren wails with a frequency $f = 2000\,\mathrm{Hz}$. What frequency can you hear if you drive toward the siren at $100\,\mathrm{km/h}$? The speed of sound is u=$344\,\mathrm{m/s}$.

$$v = 100\mathrm{km/h} = \frac{100\,000}{3600}\mathrm{m/s} = 27.8\mathrm{m/s}$$

We travel toward the source so we will perceive a greater frequency:

$$\begin{aligned} f &= f_0\left(1 + \frac{v}{u}\right) = 2500\left(1 + \frac{27.8}{344}\right)\ \mathrm{Hz} \\ &= 2702\,\mathrm{Hz} \end{aligned}$$

We hear the siren at a frequency $f = 2702\,\mathrm{Hz}$.

When driving away from that siren we hear a frequency $f = 2450\,\mathrm{Hz}$. What is our speed?

We travel away from the source so we will perceive a lower frequency:

$$f = f_0 \left(1 - \frac{v}{u}\right)$$

$$2450\,\mathrm{Hz} = 2500 \left(1 - \frac{v}{344}\right)\,\mathrm{Hz}$$

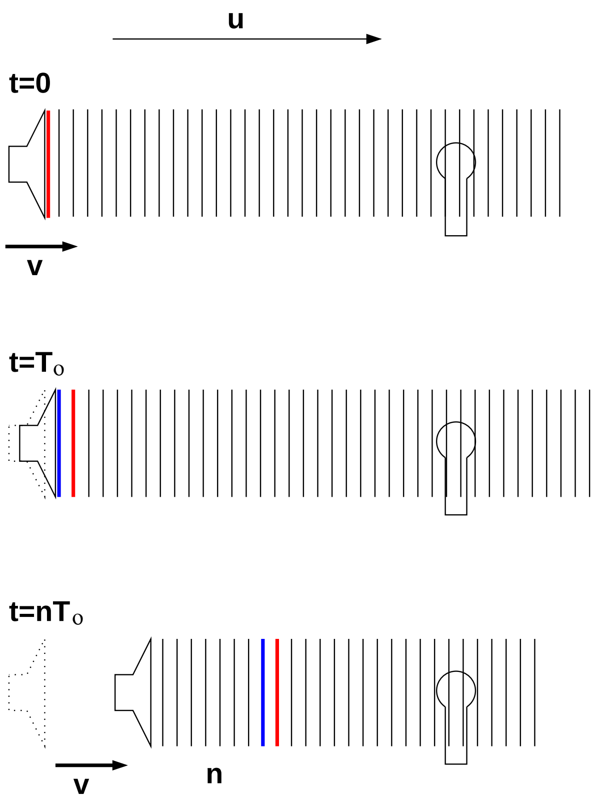
$$\frac{2450}{2500} - 1 = -\frac{v}{344}$$

$$v = 344 \left(1 - \frac{2450}{2500}\right)\mathrm{m/s} = 6.88\mathrm{m/s}$$

$$v = 24.77\mathrm{km/h}$$

We drive at a speed of $v = 24.77\,\mathrm{km/h}$ **away from** the source.

Source moving, detector stationary



The source moves toward the detector at speed v. Once the sound is emitted, it moves with the speed of sound. At the time $t = 0$ source and detector are a distance d_1 apart.

The crest emitted at $t = 0$ reaches the detector after the time $t_1 = d_1/u$. The next crest is emitted at the time $t = T_0$ and only has to travel a distance

$$d_2 = d_1 - vT_0 = d_1 - v/f_0$$

It arrives at a time $t_2 = T_0 + d_2/u = T_0 + (d_1 - vT_0)/u$

We therefore see one wavelength arrive in the time

$$\Delta t = t_2 - t_1 = T_0 + \frac{d_1 - vT_0}{u} - \frac{d_1}{u} = T_0 - T_0 v/u = T_0 \left(1 - \frac{v}{u}\right)$$

The frequency is

$$f = \frac{1}{\Delta t} = \frac{1}{T_0 \left(1 - \frac{v}{u}\right)} = \frac{f_0}{1 - \frac{v}{u}}$$

As the source moves toward the detector we hear a higher frequency!

If the source moves away from the detector the same arguments hold:

The crest emitted at $t = 0$ reaches the detector after the time $t_1 = d_1/u$. The next crest is emitted at the time $t = T_0$ and has to travel a distance

$$d_2 = d_1 + vT_0 = d_1 + v/f_0$$

It arrives at a time $t_2 = T_0 + d_2/u = T_0 + (d_1 + vT_0)/u$

We therefore see one wavelength arrive in the time

$$\Delta t = t_2 - t_1 = T_0 + \frac{d_1 + vT_0}{u} - \frac{d_1}{u} = T_0 + T_0 v/u = T_0\left(1 + \frac{v}{u}\right)$$

The frequency is

$$f = \frac{1}{\Delta t} = \frac{1}{T_0\left(1 + \frac{v}{u}\right)} = \frac{f_0}{1 + \frac{v}{u}}$$

The frequency is lower!

In total we have for a moving source and a stationary detector:

$$\boxed{f = \frac{f_0}{1 \mp \frac{v}{u}}}$$

"-": source moves **toward** the detector.
"+": source moves **away** from the detector.

Example

A train drives past a stationary listener along the track. The train whistles at a frequency of $f = 3500\,\mathrm{Hz}$ and drives at a speed of $v = 50\,\mathrm{m/s}$.

What frequency do we hear before and after the train passed us?

$$f = \frac{f_0}{1 \mp \frac{v}{u}}$$

First the train comes toward us, and we will hear a greater frequency:

$$f = \frac{f_0}{1 - \frac{v}{u}}$$

$$f = \frac{3500\,\mathrm{Hz}}{1 - \frac{50}{344}} = 4095\,\mathrm{Hz}$$

When the train moves away from us we hear a smaller frequency:

$$f = \frac{f_0}{1 + \frac{v}{u}}$$

$$f = \frac{3500\,\mathrm{Hz}}{1 + \frac{50}{344}} = 3056\,\mathrm{Hz}$$

2.14: General Doppler Effect

In general both the source and the detector can be moving:

Let v_s be the speed of the source and v_d be the speed of the detector. The speed of sound is u and the frequency of the emitted sound is f_0. All speeds are measured relative to the air which is assumed to be at rest.

Then we had:

$$\text{(1): moving detector} \qquad f_1 = f_0\left(1 \pm \frac{v_d}{u}\right)$$

$$\text{(2): moving source} \qquad f_2 = \frac{f_0}{1 \mp \frac{v_s}{u}}$$

These can be combined if we replace f_0 in (1) with the frequency associated with the moving source:

$$f = f_2\left(1 \pm \frac{v_d}{u}\right) = f_0 \frac{1}{1 \mp \frac{v_s}{u}}\left(1 \pm \frac{v_d}{u}\right)$$

$$f = f_0\left(\frac{u \pm v_d}{u \mp v_s}\right)$$

The signs can be confusing: If source and detector move **toward** each other, the frequency will be **greater**! Choose your signs accordingly.

2.15: Examples

Speed of sound in air $u = 344\,\mathrm{m/s}$. Frequency $f_0 = 5000\,\mathrm{Hz}$. Calculate the perceived frequency in each of these cases:

a) $v_s = 30\,\mathrm{m/s}$ toward the detector

$$f = \frac{f_0}{1 - \frac{v_s}{u}} = \frac{5000\,\mathrm{Hz}}{1 - 30/344} = 5000/0.9128 = 5478\,\mathrm{Hz}$$

b) $v_d = 30\,\mathrm{m/s}$ away from the source

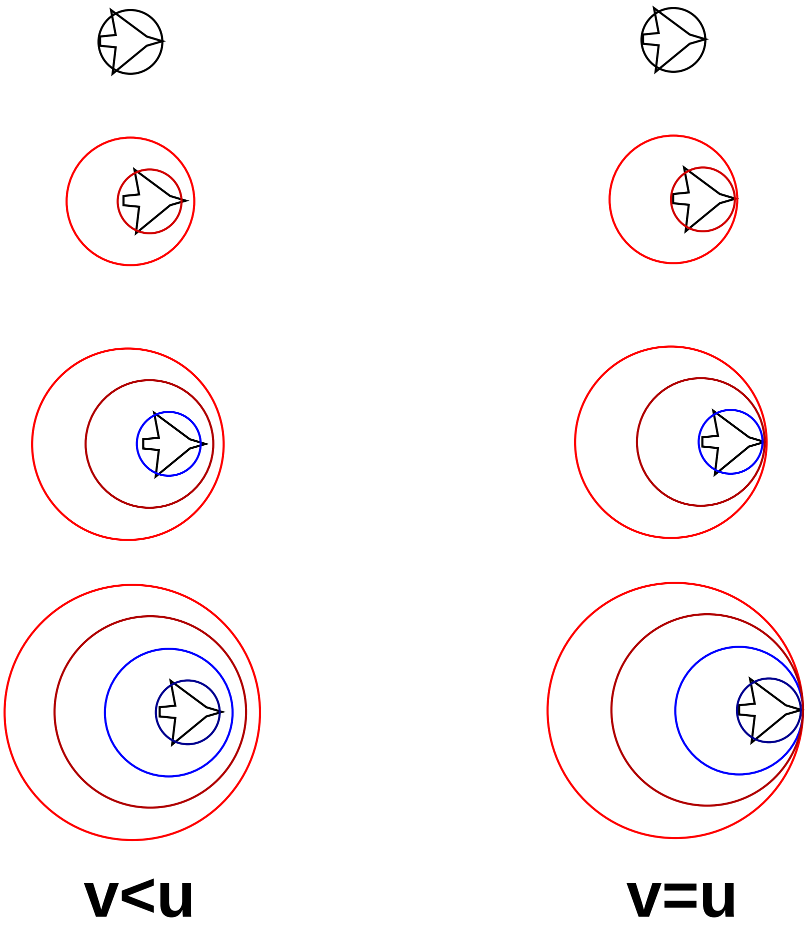
$$f = f_0(1 - \frac{v_d}{u}) = 5000\,\mathrm{Hz}(1 - 30/344) = 4564\,\mathrm{Hz}$$

c) $v_s = 60\,\mathrm{m/s}$ toward the detector, $v_d = 30\,\mathrm{m/s}$ away from the source.

$$f = f_0\left(\frac{u \pm v_d}{u \mp v_s}\right) = 5000\,\mathrm{Hz}\left(\frac{344\,\mathrm{m/s} - 30\,\mathrm{m/s}}{344\,\mathrm{m/s} - 60\,\mathrm{m/s}}\right)$$

$$f = 5000\,\mathrm{Hz}\frac{314}{284} = 5528\,\mathrm{Hz}$$

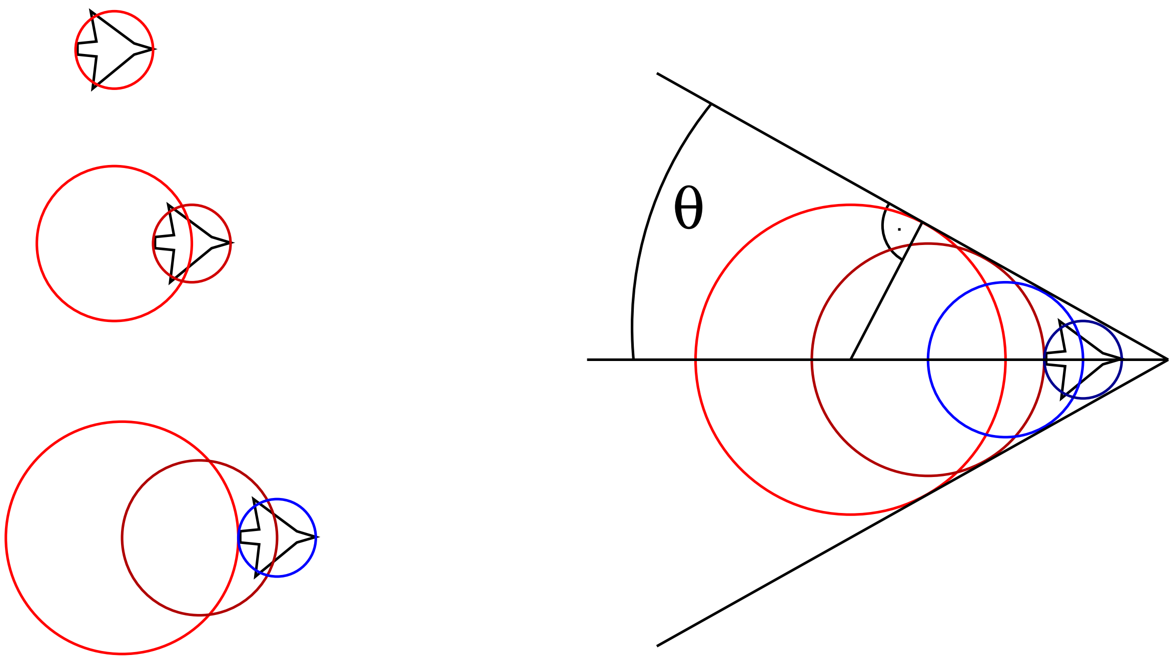
2.16: Supersonic speeds



At $v = u$ each wavecrest is emitted at the position of the previous crest. We have maximum constructive interference: The "sound barrier". It is this constructive superposition of all soundwaves that causes enormous stress on the plane.

2.17: Mach Cone

An aircraft flying at supersonic speeds creates a **Mach Cone**:



The opening angle of the cone is $\phi_{Mach} = 2\theta = 2\sin^{-1}(u/v_s)$ with v_s the speed of the aircraft and u the speed of sound.

2.18: Summary

- The source emits waves with a frequency f_0. We assume that the medium (air) is at rest and measure all velocities (of source and detector) relative to the air. u is the speed of sound.

- moving detector and a stationary source:

$$f = f_0 \left(1 \pm \frac{v_d}{u}\right)$$

  "+": detector moves **toward** the source.
  "-": detector moves **away** from the source.

- moving source and a stationary detector:

$$f = \frac{f_0}{1 \mp \frac{v_s}{u}}$$

  "-": source moves **toward** the detector.
  "+": source moves **away** from the detector.

- general Doppler effect:

$$f = f_0 \left(\frac{u \pm v_d}{u \mp v_s}\right)$$

  If source and detector move **toward** each other, the frequency will be **greater**.

Exercises

Halliday Resnick & Walker:

Reading: HRW p. 414-420

Exercises: p. 420 ff.: 46E, 48E, 51P, 52P, 55P, 59P

2.19: Beats

We have looked at the superposition of waves before. What hapens if the two waves do not have the same frequencies?

Take two sinusoidal sound waves

$$s_1(x,t) = s_0 \cos(k_1 x - \omega_1 t)$$
$$s_2(x,t) = s_0 \cos(k_2 x - \omega_2 t)$$

and choose the origin so that for $x = 0$, $t = 0$ both waves are in phase.

What is the resulting wave?

$$\begin{aligned} s(x,t) &= s_1(x,t) + s_2(x,t) \\ &= s_0 \cos(k_1 x - \omega_1 t) + s_0 \cos(k_2 x - \omega_2 t) \end{aligned}$$

Another useful trigonometric identity is

$$\cos(\alpha) + \cos(\beta) = 2 \cos \tfrac{1}{2}(\alpha + \beta) \cos \tfrac{1}{2}(\alpha - \beta)$$

With $\alpha = k_1 x - \omega_1 t$ and $\beta = k_2 x - \omega_2 t$ we get:

$$\begin{aligned} s(x,t) &= 2s_0 \cos(\tfrac{1}{2}((k_1 - k_2)x - (\omega_1 - \omega_2)t)) \\ &\times \cos(\tfrac{1}{2}((k_1 + k_2)x - (\omega_1 + \omega_2)t)) \end{aligned}$$

If $\omega_1 = \omega_2$ this reduces to the expression for two waves interfering constructively we had before.

In sound waves we are most interested in the frequencies. Since the speed of sound is $u = \omega/k$ we can always find k from ω. To make matters convenient, we look at the displacement as a function of time at the origin ($x = 0$).

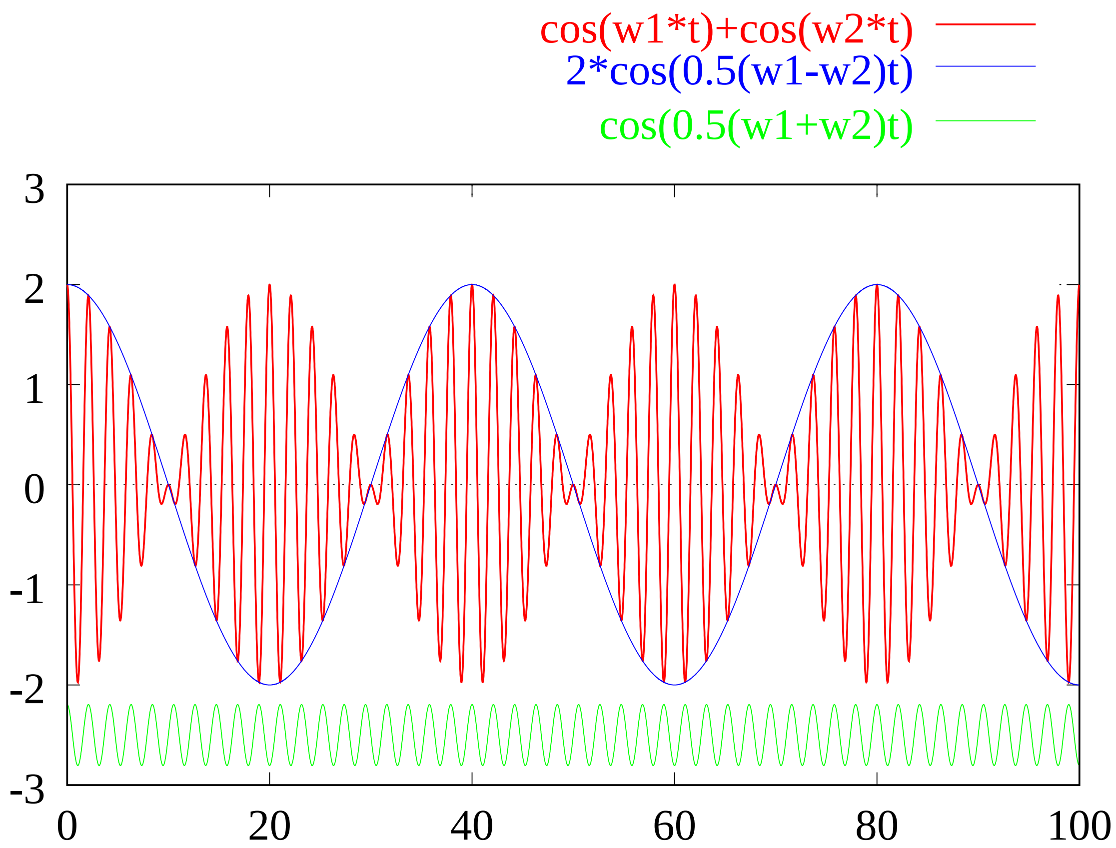
That gives

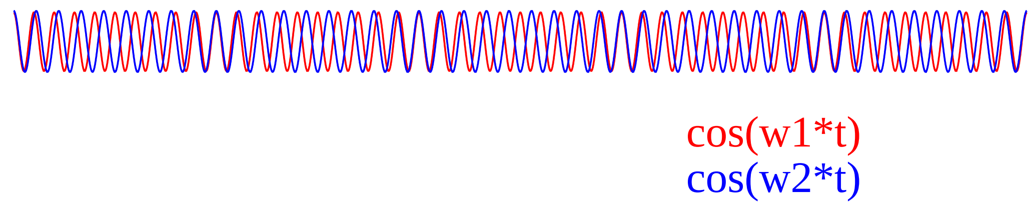
$$s(x,t) = s(0,t) = 2s_0 \cos \tfrac{1}{2}(\omega_1 - \omega_2)t \ \cos \tfrac{1}{2}(\omega_1 + \omega_2)t$$

Or, defining $\omega_a = \frac{1}{2}(\omega_1 + \omega_2)$, and $\omega_d = \frac{1}{2}(\omega_1 - \omega_2)$ we have

$$s(0,t) = 2s_0 \cos(\omega_d t) \cos(\omega_a t)$$

If we observe the displacement amplitude at $x = 0$ then we get this picture:





These big oscillations are called "beats". Since there are **two** maxima during each period, their angular frequency is the difference between the two angular frequencies:

$$\omega_{beat} = |(\omega_1 - \omega_2)|$$

In the picture before I used

$$f_1 = 0.5\text{Hz} \qquad\qquad f_2 = 0.45\text{Hz}$$

with a beat frequency of $f_{beat} = |(f_1 - f_2)| = 0.05\,\text{Hz}$.

Usually the amplitudes are not the same. We still will get beats:

$$\begin{aligned}
s(0,t) &= s_1\cos(\omega_1 t) + s_2\cos(\omega_2 t) \\
&= \tfrac{1}{2}s_1\cos(\omega_1 t) + \tfrac{1}{2}s_1\cos(\omega_1 t) \\
&+ \tfrac{1}{2}s_2\cos(\omega_2 t) + \tfrac{1}{2}s_2\cos(\omega_2 t) + \\
&+ \tfrac{1}{2}s_1\cos(\omega_2 t) - \tfrac{1}{2}s_1\cos(\omega_2 t) \\
&+ \tfrac{1}{2}s_2\cos(\omega_1 t) - \tfrac{1}{2}s_2\cos(\omega_1 t) \\
&= \tfrac{1}{2}(s_1+s_2)(\cos(\omega_1 t) + \cos(\omega_2 t)) \\
&+ \tfrac{1}{2}(s_1-s_2)(\cos(\omega_1 t) - \cos(\omega_2 t))
\end{aligned}$$

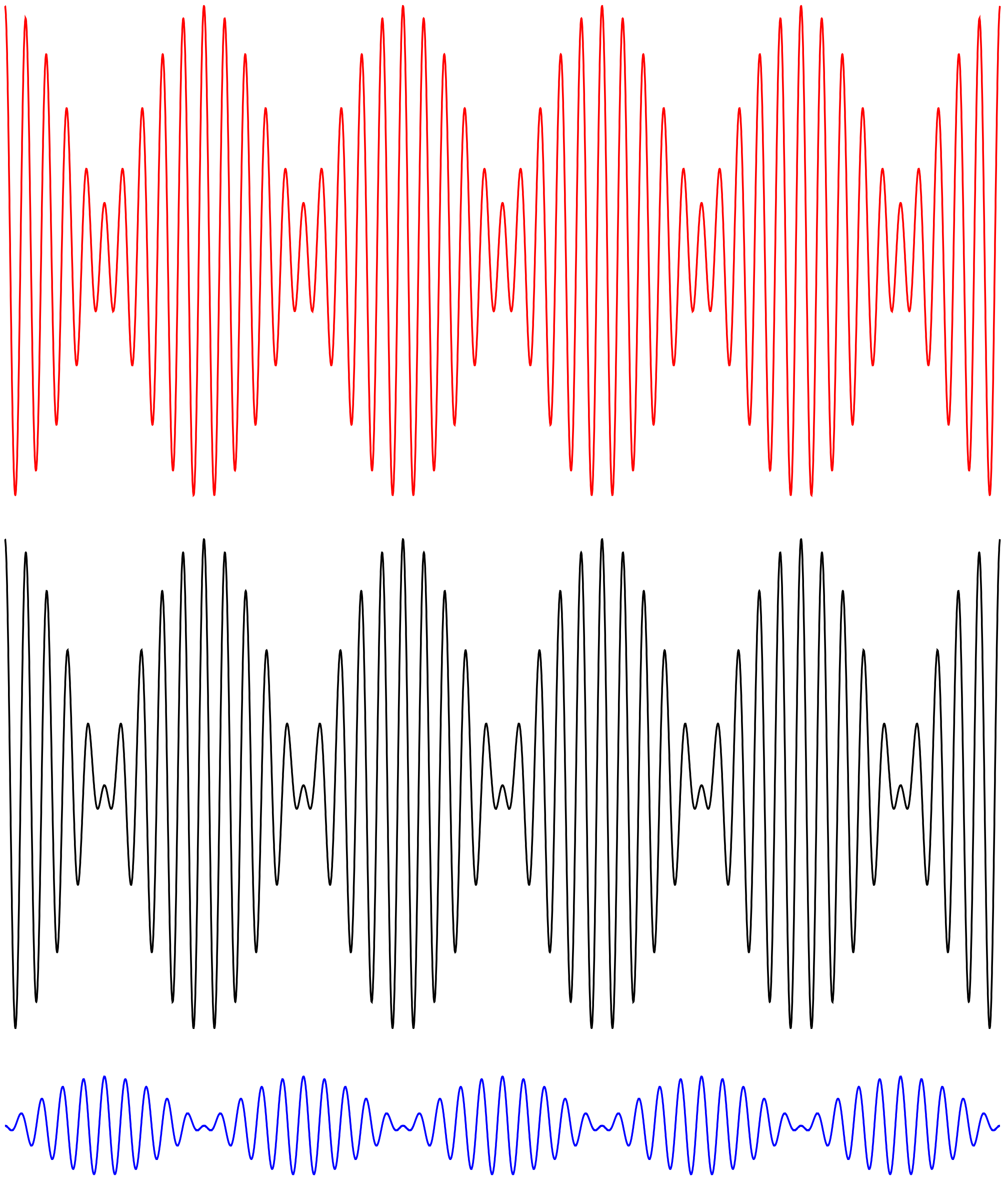
Again use

$$\cos(\alpha) + \cos(\beta) = 2\cos\tfrac{1}{2}(\alpha+\beta)\cos\tfrac{1}{2}(\alpha-\beta)$$

and

$$\cos(\alpha) - \cos(\beta) = -2\sin\tfrac{1}{2}(\alpha+\beta)\sin\tfrac{1}{2}(\alpha-\beta)$$

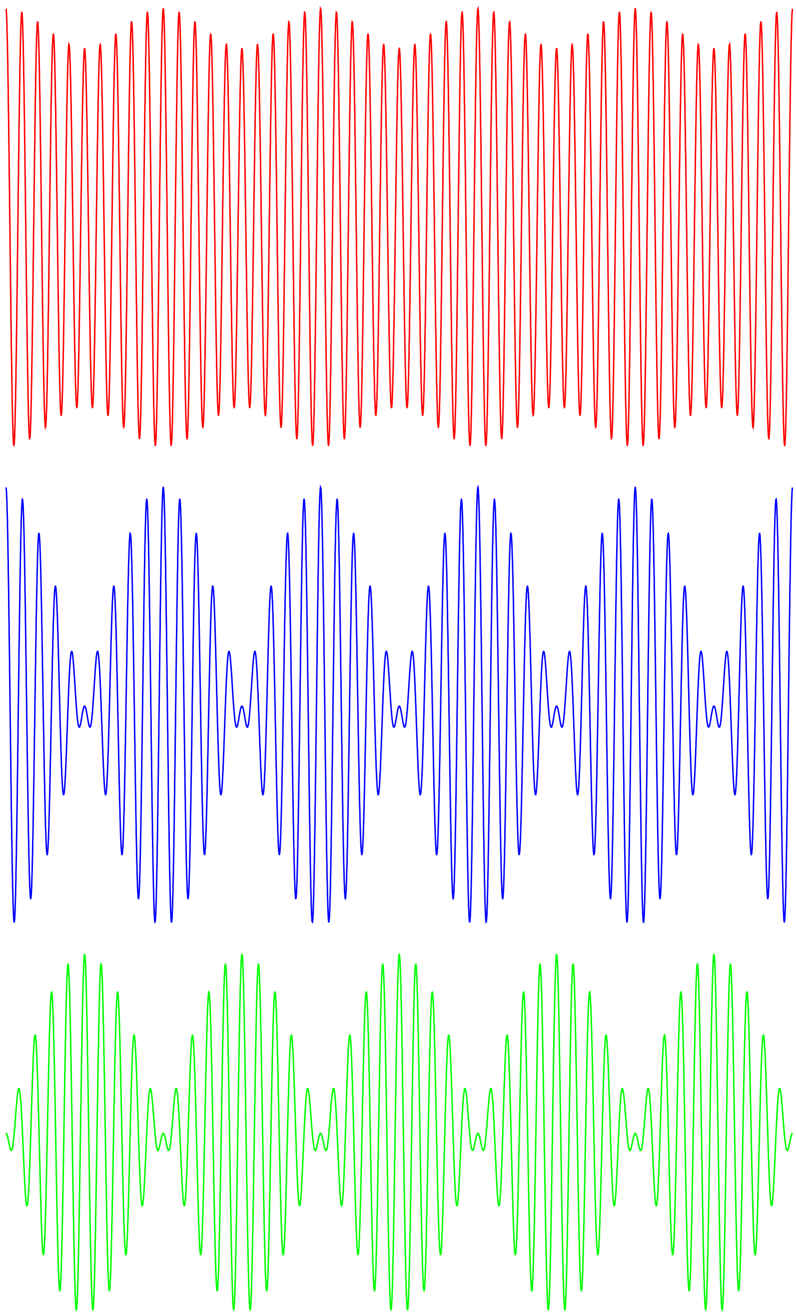
$$s(0,t) = (s_1+s_2)\cos(\omega_a t)\cos(\omega_d t) - (s_1-s_2)\sin(\omega_a t)\sin(\omega_d t)$$



Top: $s(0,t) = s_1 \cos(\omega_1 t) + s_2 \cos(\omega_2 t)$

Middle: $s(0,t) = (s_1 + s_2) \cos(\omega_a t) \cos(\omega_d t)$

Bottom: $s(0,t) = (s_1 - s_2) \sin(\omega_a t) \sin(\omega_d t)$



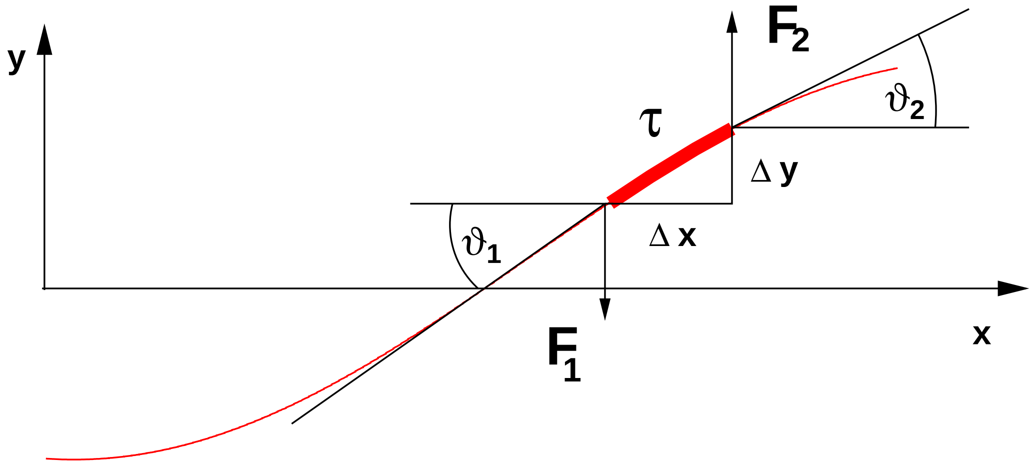
Top: $s(0,t) = s_1 \cos(\omega_1 t) + s_2 \cos(\omega_2 t)$

Middle: $s(0,t) = (s_1 + s_2) \cos(\omega_a t) \cos(\omega_d t)$

Bottom: $s(0,t) = (s_1 - s_2) \sin(\omega_a t) \sin(\omega_d t)$

2.20: The Wave Equation

Go back to the wave on a string.



We have some disturbance in the string $y(x,t)$, but we make no assumptions about the nature of this disturbance except that it is small and "well behaved" (continuous and differentiable).

The total **vertical** force on that linesegment is $F_1 + F_2$.

$$F_1 = -\tau \sin\vartheta_1 \simeq -\tau \tan\vartheta_1 = -\tau \left.\frac{\partial y(x,t)}{\partial x}\right|_{x=x_1}$$

$$F_2 = \tau \sin\vartheta_2 \simeq \tau \tan\vartheta_2 = \tau \left.\frac{\partial y(x,t)}{\partial x}\right|_{x=x_2}$$

$$F_y = F_1 + F_2 = \tau \left(\left.\frac{\partial y(x,t)}{\partial x}\right|_{x=x_2} - \left.\frac{\partial y(x,t)}{\partial x}\right|_{x=x_1} \right)$$

Define

$$u(x,t) = \frac{\partial y(x,t)}{\partial x}$$

$$F_y = \tau(u(x_2,t) - u(x_1,t))$$

For a small difference between x_1 and x_2 we can use the Taylor expansion:

$$u(x_2,t) = u(x_1,t) + (x_2 - x_1)\left.\frac{\partial u(x,t)}{\partial x}\right|_{x=x_1} + \cdots$$

The vertical force is therefore

$$F_y = \tau \Delta x \frac{\partial u(x,t)}{\partial x}$$

or, putting $u(x,t) = \frac{\partial y(x,t)}{\partial x}$:

$$F_y = \tau \Delta x \frac{\partial^2 y(x,t)}{\partial x^2}$$

Newton's law $F = ma$ then becomes

$$\tau \Delta x \frac{\partial^2 y(x,t)}{\partial x^2} = \mu \Delta x \frac{\partial^2 y(x,t)}{\partial t^2}$$

Or:

$$\tau \frac{\partial^2 y(x,t)}{\partial x^2} = \mu \frac{\partial^2 y(x,t)}{\partial t^2}$$

We already know that $v = \sqrt{\tau/\mu}$, so we can write

$$v^2 \, \frac{\partial^2 y}{\partial x^2} = \frac{\partial^2 y}{\partial t^2}$$

This is the **Wave Equation**. It appears whenever waves are present:

In mechanics, acoustics, optics, electrodynamics, quantum mechanics, etc.

It says that at each point in space and time the acceleration of a piece of string is proportional to the curvature at this point.

How do we deal with this differential equation?

First let us see if the sinusoidal waves we've considered so far are actually solutions of the wave equation.

To do so we must verify that
$y(x,t) = y_0 \sin(kx - \omega t - \Phi_0)$ fulfills the wave equation.

We need the second derivatives $\frac{\partial^2 y}{\partial x^2}$ and $\frac{\partial^2 y}{\partial t^2}$:

$$\frac{\partial y(x,t)}{\partial x} = k \cos(kx - \omega t - \Phi_0)$$
$$\frac{\partial^2 y(x,t)}{\partial x^2} = -k^2 \sin(kx - \omega t - \Phi_0)$$

$$\frac{\partial y(x,t)}{\partial t} = -\omega \cos(kx - \omega t - \Phi_0)$$
$$\frac{\partial^2 y(x,t)}{\partial t^2} = -\omega^2 \sin(kx - \omega t - \Phi_0)$$

$$v^2 \, \frac{\partial^2 y}{\partial x^2} = \frac{\partial^2 y}{\partial t^2}$$
$$-k^2 v^2 \sin(kx - \omega t - \Phi_0) = -\omega^2 \sin(kx - \omega t - \Phi_0)$$

If we set $v^2 = \omega^2/k^2$, the wave equation is fulfilled!

Is this the only solution?

No, $y(x,t) = y_0 \sin(kx + \omega t - \Phi_0)$ also fulfills te wave equation. That is good, because in the derivation of the wave equation we made no assumptions about the direction in which the wave is moving.

Are there more solutions?

Take $f(x,t) = f(kx - \omega t) = f(\alpha)$ with $\alpha = kx - \omega t$.
f can be **any** function, as long as it is differentiable at least twice.

$$\frac{\partial f(\alpha)}{\partial x} = \frac{\partial f(\alpha)}{\partial \alpha}\frac{\partial \alpha}{\partial x} = k\frac{\partial f(\alpha)}{\partial \alpha}$$

$$\frac{\partial^2 f(\alpha)}{\partial x^2} = k\frac{\partial}{\partial \alpha}\frac{\partial f(\alpha)}{\partial \alpha}\frac{\partial \alpha}{\partial x} = k^2\frac{\partial^2 f(\alpha)}{\partial \alpha^2}$$

And the time derivatives:

$$\frac{\partial f(\alpha)}{\partial t} = \frac{\partial f(\alpha)}{\partial \alpha}\frac{\partial \alpha}{\partial t} = -\omega\frac{\partial f(\alpha)}{\partial \alpha}$$

$$\frac{\partial^2 f(\alpha)}{\partial t^2} = -\omega\frac{\partial}{\partial \alpha}\frac{\partial f(\alpha)}{\partial \alpha}\frac{\partial \alpha}{\partial t} = \omega^2\frac{\partial^2 f(\alpha)}{\partial \alpha^2}$$

Put this into the wave equation:

$$v^2\frac{\partial^2 y}{\partial x^2} = \frac{\partial^2 y}{\partial t^2} \quad \leftrightarrow \quad v^2 k^2\frac{\partial^2 f(\alpha)}{\partial \alpha^2} = \omega^2\frac{\partial^2 f(\alpha)}{\partial \alpha^2}$$

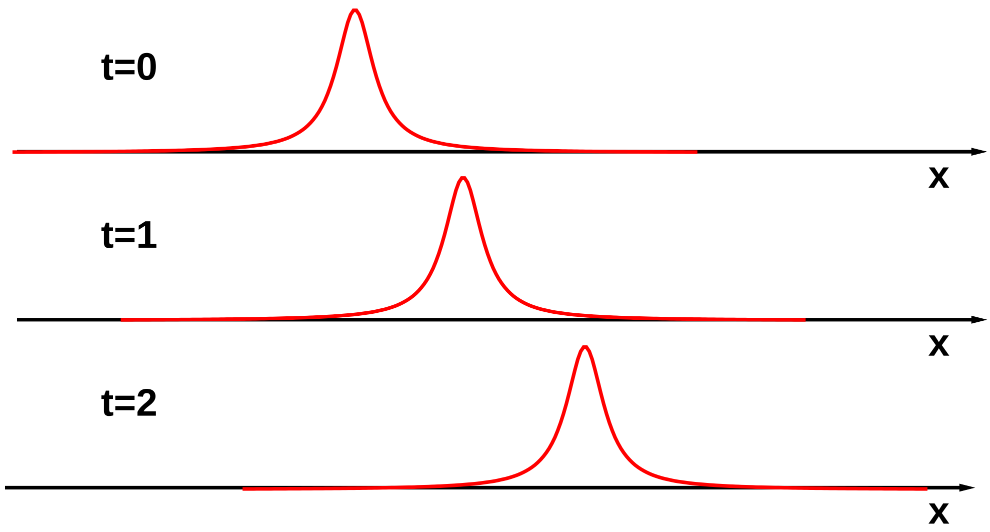
Again, with $v^2 = \omega^2/k^2$ the wave equation is fulfilled!

In just the same way we find that any function $f(x,t) = f(kx + \omega t)$ fulfills the wave equation.

Any function of $(kx \pm \omega t)$ is a wave. That is exactly what the principle of superposition implied.

Example: Take a pulse given by

$$f(x,t) = \frac{1}{1 + (kx - \omega t)^2}$$



$$\frac{\partial}{\partial x}\left(\frac{1}{1 + (kx - \omega t)^2}\right) = \frac{-2k(kx - \omega t)}{(1 + (kx - \omega t)^2)^2}$$

$$\frac{\partial^2}{\partial x^2}\left(\frac{1}{1 + (kx - \omega t)^2}\right) = -2k^2 \frac{1 - 3(kx - \omega t)^2}{(1 + (kx - \omega t)^2)^3}$$

The time derivatives are:

$$\frac{\partial}{\partial t}\left(\frac{1}{1+(kx-\omega t)^2}\right) = \frac{2\omega(kx-\omega t)}{(1+(kx-\omega t)^2)^2}$$

$$\frac{\partial^2}{\partial t^2}\left(\frac{1}{1+(kx-\omega t)^2}\right) = -2\omega^2\frac{1-3(kx-\omega t)^2}{(1+(kx-\omega t)^2)^3}$$

The wave equation was:

$$v^2\ \frac{\partial^2 y}{\partial x^2} = \frac{\partial^2 y}{\partial t^2}$$

$$-2k^2v^2\frac{1-3(kx-\omega t)^2}{(1+(kx-\omega t)^2)^3} = -2\omega^2\frac{1-3(kx-\omega t)^2}{(1+(kx-\omega t)^2)^3}$$

Again the wave equation is fulfilled. This time we did not have any periodic wave, but just a pulse.

2.21: Summary

- Two travelling waves with different frequencies produce **beats**. The beat frequency is given by $f_{beat} = |f_1 - f_2|$

- The wave equation is

$$v^2 \frac{\partial^2 y}{\partial x^2} = \frac{\partial^2 y}{\partial t^2}$$

  with the wave speed v.

- Any function of $(kx \pm \omega t)$ fulfills the wave equation. $(kx - \omega t)$ describes a wave travelling in the positive x–direction,
  $(kx + \omega t)$ describes a wave travelling in the negative x–direction.

2.22: Useful Equations

$$\cos\alpha + \cos\beta = 2\cos(\tfrac{1}{2}(\alpha-\beta))\cos(\tfrac{1}{2}(\alpha+\beta))$$
$$\cos\alpha - \cos\beta = -2\sin(\tfrac{1}{2}(\alpha-\beta))\sin(\tfrac{1}{2}(\alpha+\beta))$$
$$\sin\alpha + \sin\beta = 2\cos(\tfrac{1}{2}(\alpha-\beta))\sin(\tfrac{1}{2}(\alpha+\beta))$$
$$\sin\alpha - \sin\beta = 2\sin(\tfrac{1}{2}(\alpha-\beta))\cos(\tfrac{1}{2}(\alpha+\beta))$$

$$\cos(\alpha\pm\beta) = \cos\alpha\cos\beta \mp \sin\alpha\sin\beta$$
$$\sin(\alpha\pm\beta) = \sin\alpha\cos\beta \pm \cos\alpha\sin\beta$$

$$\cos 2\alpha = \cos^2\alpha - \sin^2\alpha$$
$$\sin 2\alpha = 2\sin\alpha\cos\alpha$$
$$\cos^2\alpha = \tfrac{1}{2}(1+\cos 2\alpha)$$
$$\sin^2\alpha = \tfrac{1}{2}(1-\cos 2\alpha)$$
$$\sin\alpha = \alpha - \alpha^3/6 + \cdots$$
$$\cos\alpha = 1 - \tfrac{1}{2}\alpha^2 + \cdots$$

$$\cos\alpha + \cos(\alpha+\beta) + \cos(\alpha+2\beta) + \cdots + \cos(\alpha+(n-1)\beta)$$
$$= \cos(\alpha + \tfrac{1}{2}(n-1)\beta)\frac{\sin\frac{1}{2}n\beta}{\sin\frac{1}{2}\beta}$$

Exercises

Halliday Resnick & Walker:

Reading: HRW p. 412-414

Exercises: p. 420 ff.: 42E, 45P

Show explicitely that the following functions fulfill the wave equation:

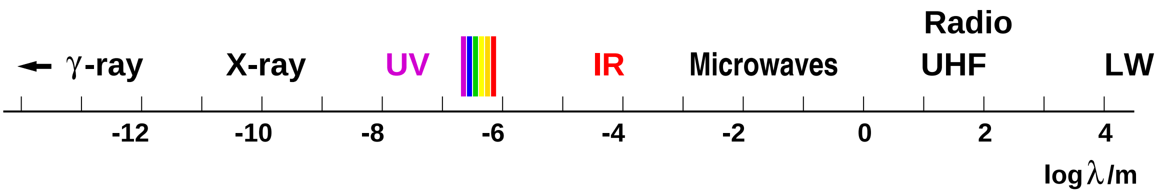
$$y(x,t) = A\cos(kx + \omega t)$$

$$y(x,t) = \frac{1}{1 + (kx + \omega t)^2}$$

3: Geometrical Optics

What is light? Light is electromagnetic radiation. The spectrum ranges from very long radio waves (100 m and longer) to very short γ-rays (10^{-12} m) and shorter.

The visible spectrum covers a range of wavelengths from 400 nm (blue) to 700 nm (red).



This wavelength is small compared to the typical dimensions of optical instruments. We can therefore learn a lot about optics without treating light as waves.

Treating light as rays travelling in straight lines we can now study optical instruments, lenses, mirrors, telescopes, etc.

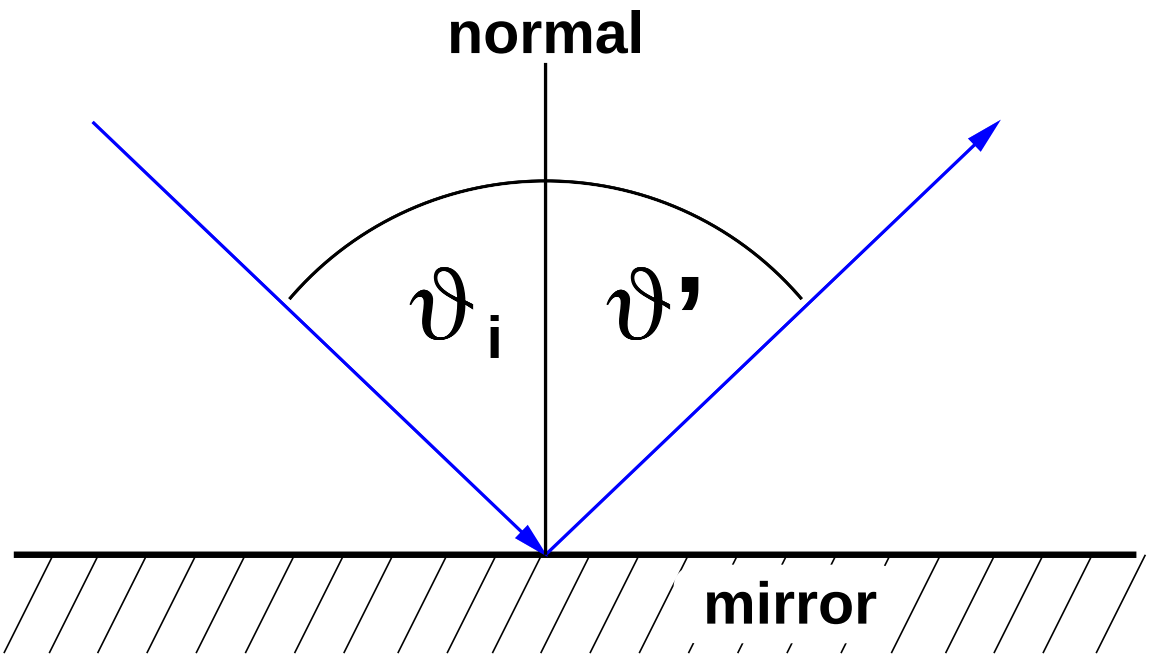
$\Rightarrow$ **Geometrical Optics**

Later we will go back and justify our conclusions by taking the wave nature of light into account.

3.1: Reflection

The first phenomenon is **Reflection**.

"A ray incident on a plane mirror is reflected so that the reflected ray lies in the plane of incidence and the angle of reflection is the same as the angle of incidence."



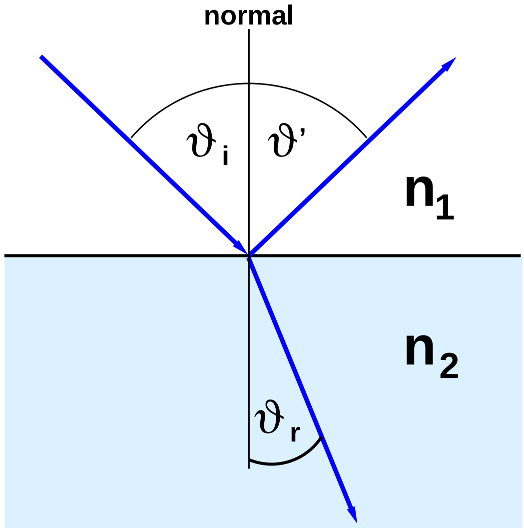
$$\vartheta' = \vartheta_i$$

This law holds regardless of the wavelength of the incident light.

3.2: Refraction

The second phenomenon is **Refraction**. A ray of light incident on a transparent surface (say, a lake) will split. Part of the beam is reflected (you can see reflections on a lake) and part of it will enter the water.

But the ray entering the water does not travel straight, it is refracted.



The diffraction depends on the material and is given by the index of refraction n.

The law of refraction is (Snell's law):

> The refracted ray lies in the plane of incidence and the angle of refraction ϑ_r is connected to the angle of incidence ϑ_i by
> $$n_1 \sin\vartheta_i = n_2 \sin\vartheta_r$$

The index of refraction is connected to the speed of light **in the medium** by

$$n = c/v$$

In water we have $n = 1.33$ and thus the speed of light in water is $v = c/1.33 = 225500\,\mathrm{km/s}$. We shall discuss the reasons for this later.

One consequence is, that the refractive index of vacuum is $n_{vac} = 1$ and that no material can have a refractive index less than 1.

MEDIUM	n
Vacuum	1
Air	1.00029
Water	1.33
Quartz	1.46
Crown glass	1.52
Flint glass	1.65
Sapphire	1.77
Diamond	2.42

3.3: Examples

A ray of light is incident on a glass cube with refractive index $n = 1.5$. The angle of incidence is $\vartheta_i = 20°$.

What is the angle of reflection? $\vartheta' = 20°$.

What is the angle of refraction? $n_{air} = 1$

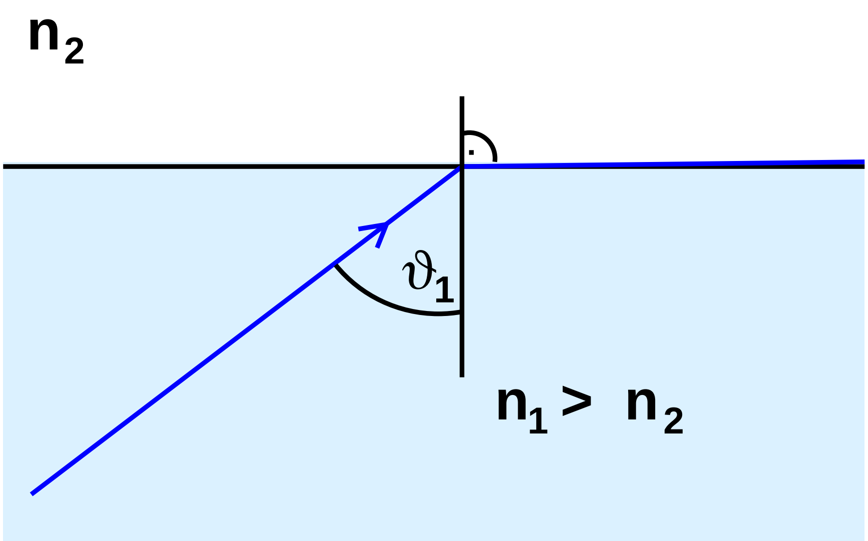
$$n_{air} \sin 20° = n_{glass} \sin \vartheta_r$$
$$\sin \vartheta_r = \frac{1}{1.5} 0.342$$
$$\vartheta_r = 13.2°$$

The glass cube is now under water $n_W = 1.33$. What is the angle of refraction?

$$n_W \sin 20° = n_{glass} \sin \vartheta_r$$
$$\sin \vartheta_r = \frac{1.33}{1.5} 0.342$$
$$\vartheta_r = 17.7°$$

3.4: Total internal reflection

If we shine a ray of light from under the water into the air, what is the largest angle the incident ray can have?



The law of refraction gives

$$n_1 \sin\vartheta_1 = n_2 \sin\vartheta_2$$

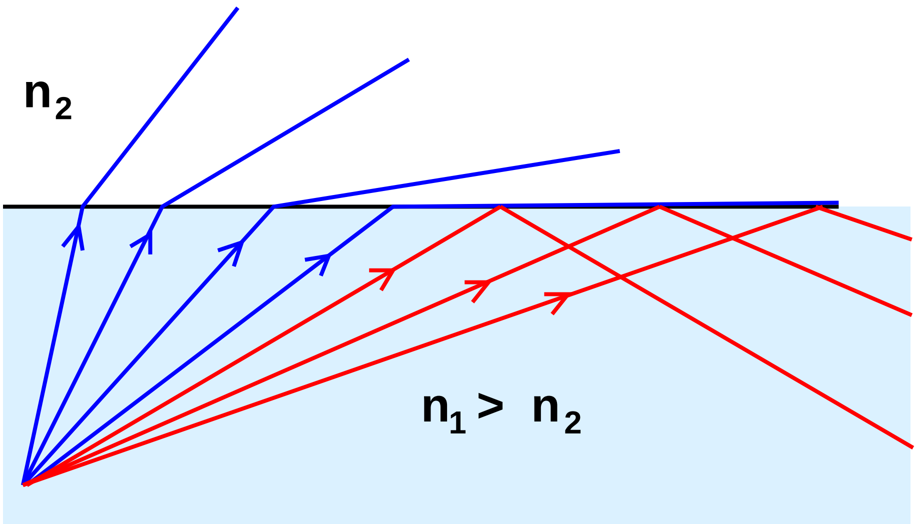
The refracted ray must have an angle smaller than $\vartheta_2 = 90°$

$$\sin\vartheta_1 = \frac{n_2}{n_1}\sin 90° < 1$$

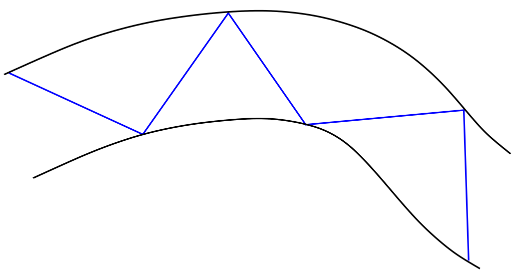
For water and air we get

$$\vartheta_{crit} = \sin^{-1}\frac{1}{1.33} = \sin^{-1} 0.75 = 48.6°$$

What happens if we try a larger incident angle? The law of refraction cannot be fulfilled, we can not get a refracted beam. We are now left only with a reflected beam, the ray never leaves the water.

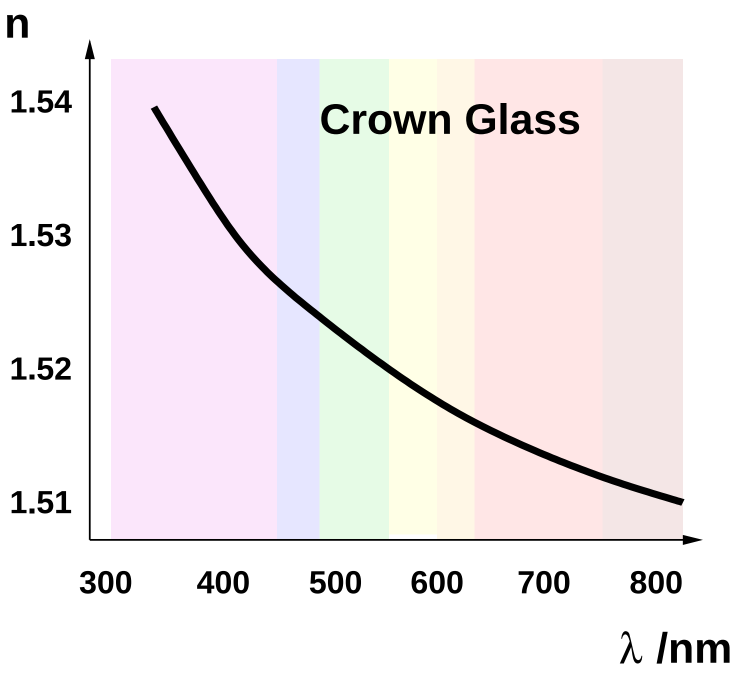


This phenomenon is called **total internal reflection**. Total internal reflection is very useful in optical fibres: It will transmit light with a minimum of loss. A reflective coating (Silver etc.) will have much larger losses!



3.5: Dispersion

When you shine a beam of white light into refracting bodies it is sometimes refracted into its colors. The index of refraction depends on the wavelength of the light!

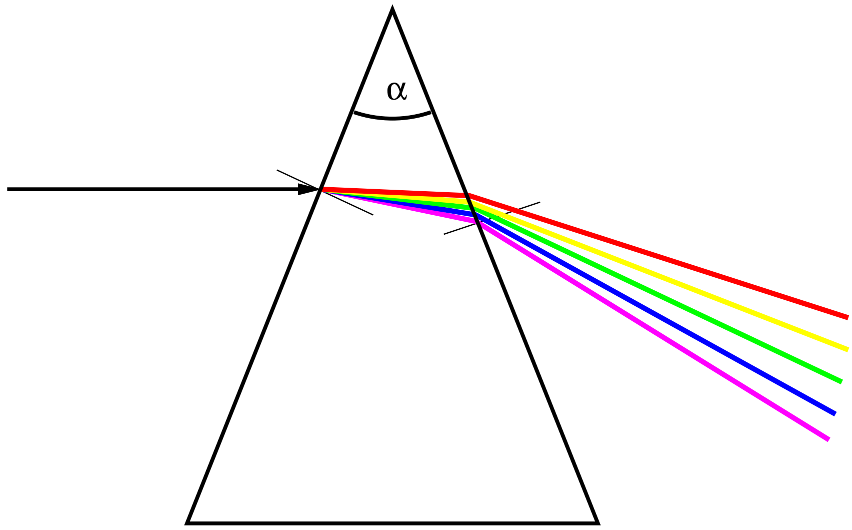


This means the speed of light in the medium depends on the wavelength. In glass, red light is faster than blue light!

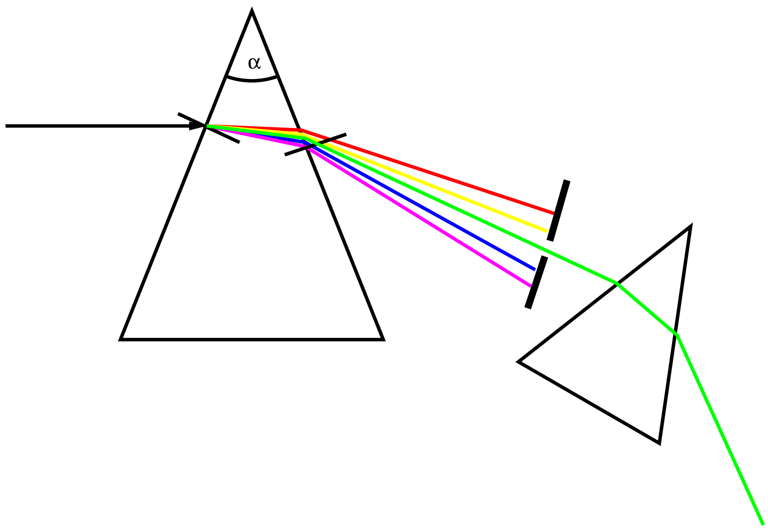
This is the first time that we find the wave speed to depend on the wavelength — in contrast to waves on strings or sound.

3.6: Prisms

We can use this effect in a prism to break white light into its component colors:



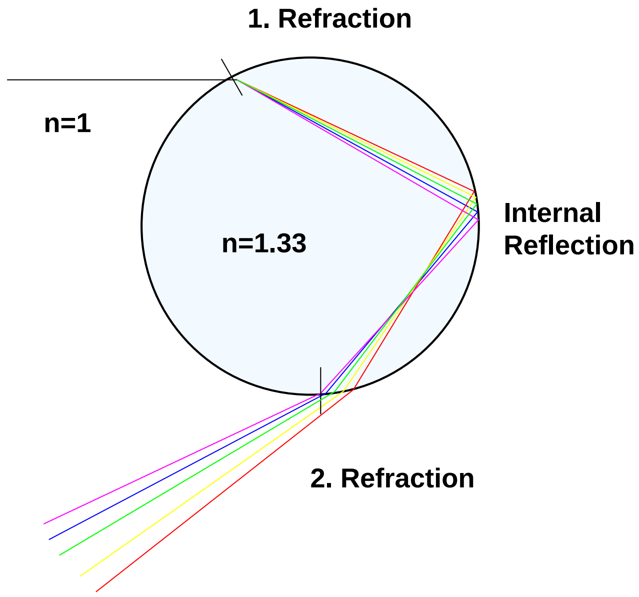
These colors can **not** be further refracted using another prism.



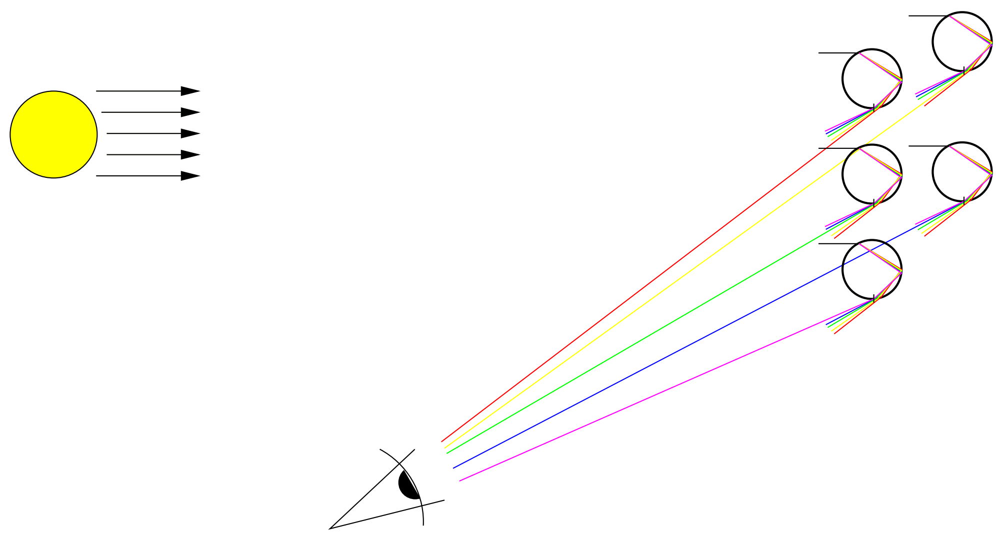
3.7: Rainbows

A rainbow is formed when you look at a cloud consisting of billions of little spherical droplets of water with the sun behind you.

Rene Descartes was the first scientist to explain the rainbow in 1637 by analysing several paths of light through a spherical droplet with two refractions at the surface and one total reflection.



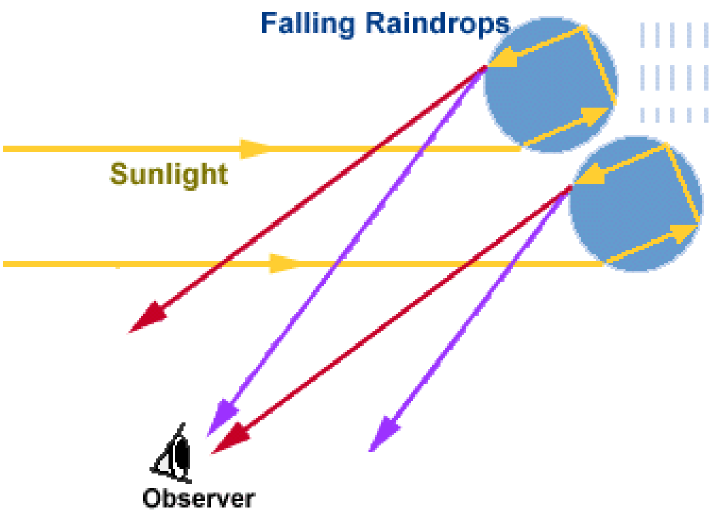
Of all scattering angles, those around 42° for red and 40° for blue light are most intense. They correspond to the shortest optical paths through the droplet.



Note that the sun is always behind you when you see a rainbow.

Sometimes we can see a second, inverted rainbow over the first one:

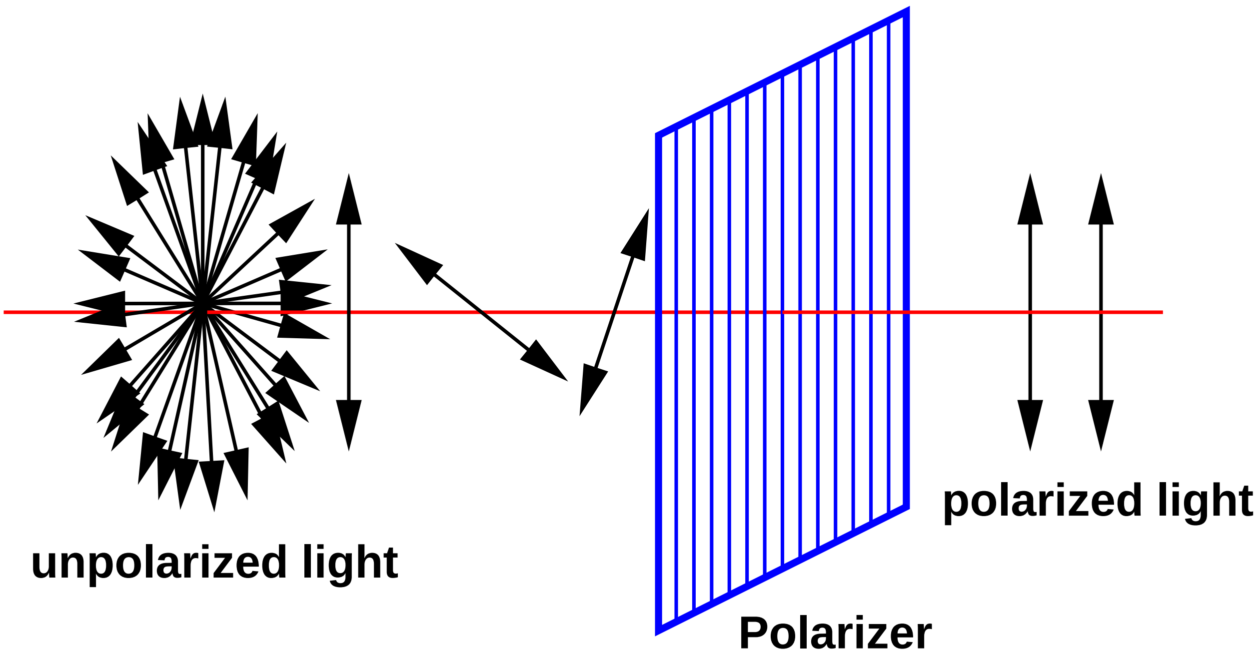
This is formed if the sunlight undergoes **two** internal reflections:



3.8: Polarization

Light is a purely transversal wave. As light waves move in one direction, the electric field vector can oscillate freely in a plane perpendicular to that direction.

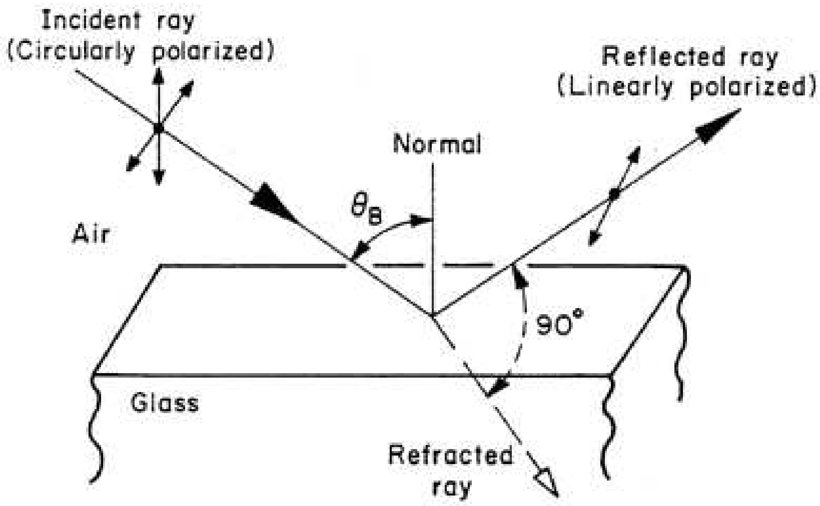
Normal sunlight contains many rays with random polarizations - sunlight is unpolarized.



By applying filters that transmit light with one polarization only, we can generate **polarized** light.

3.9: The Brewster angle

We find experimentally that light reflected off a refracting surface (A normal mirror won't work!) is polarized when viewed under a specific angle, the **Brewster angle**.



The Brewster angle is the angle where the refracted ray and the reflected ray are perpendicular to each other: $\vartheta' + \vartheta_r = 90$.

$$n_1 \sin\vartheta_B = n_2 \sin\vartheta_r = \sin(90 - \vartheta_B) = n_2 \cos\vartheta_B$$

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

If medium 1 is air we have $n_1 = 1$ and we obtain for the Brewster angle:

$$\tan\vartheta_B = n_2$$

3.10: Summary

- Reflection: "A ray incident on a plane mirror is reflected so that the reflected ray lies in the plane of incidence and the angle of reflection is the same as the angle of incidence."
- Refraction (Snell's law): The refracted ray lies in the plane of incidence and the angle of refraction θ_r is connected to the angle of incidence θ_i by $n_1 \sin\theta_i = n_2 \sin\theta_r$
- The index of refraction n is the ratio of the speed of light in vacuum to the speed of light in the medium $n = c/v$.
- A ray of light trying to cross a boundary from a large refractive index n_1 to a small ref. index n_2 can undergo total internal reflection. The critical angle is given by $\sin\vartheta_{crit} = \frac{n_2}{n_1}$
- The refractive index also depends weakly on the wavelength. This allows prisms to disperse white light into its colors.
- Light can be polarized by reflection off a refractive boundary. The angle at which the reflected ray is maximally polarized is called the Brewster angle $\tan\vartheta_B = \frac{n_2}{n_1}$.

Exercises

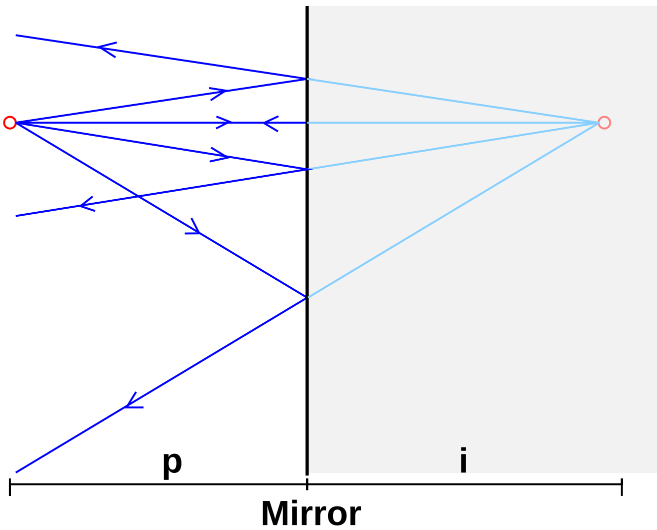
Halliday Resnick & Walker:

Reading: HRW pp. 814-825

Exercises: p. 826 ff.: Q6, Q8, Q9, Q12, 45E, 46P, 51P, 61E

3.11: Mirrors

The most important quality of a mirror is the ability to form an image. If you look at your reflection in the mirror your eyes intercept rays of light that have been reflected off the mirrors surface and the image they originate from seems to be **behind** the mirror.



The distance of the image from the mirror $|i|$ is equal to the distance of the object from the mirror $|p|$ and its size is equal to the size of the object.

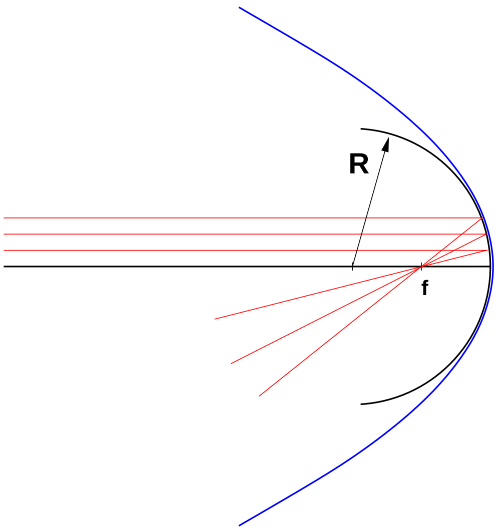
This solves an old puzzle: It is not left and right that are swapped in the mirror, but front and back!

Since the image is behind the mirror we can put a solid block of concrete there without changing the image. The image is **virtual**.

3.12: Spherical Mirrors

The rays from the sun approach us from far away and can be taken as parallel. A curved concave mirror with a radius of curvature R will focus these parallel rays into a single spot if they are close enough to the optical axis.

Taking a spherical mirror is not ideal, a parabola will indeed focus **all** incoming rays into a point. A spherical mirror is a good approximation to the parabola - for small distances from the optical axis.



The focus is positioned halfway between the center of the sphere and the mirror itself.

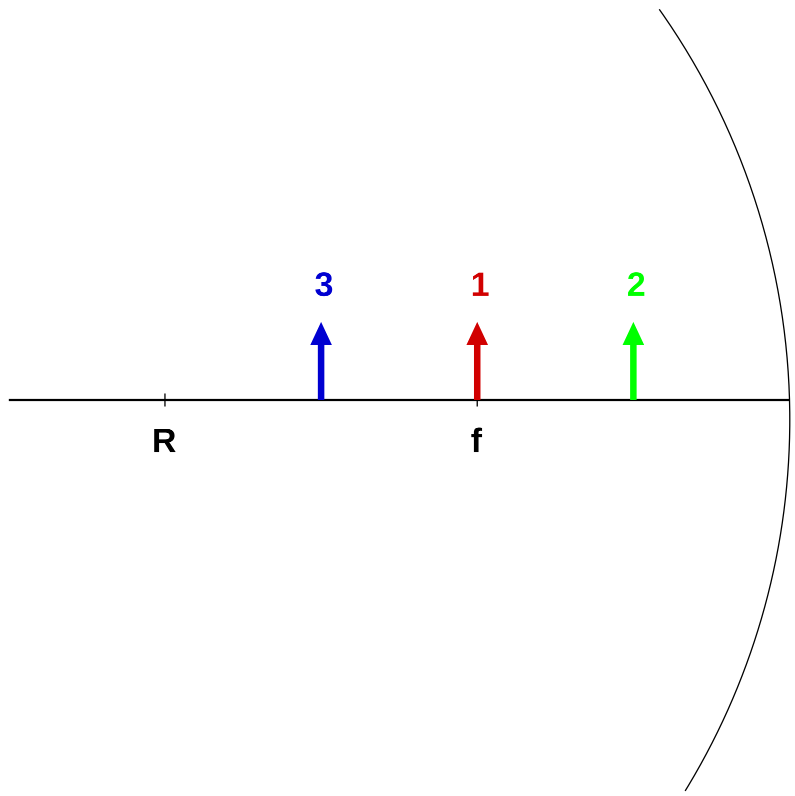
$$|f| = R/2$$

All the rays actually pass through the focus, it is a **real** image of the sun.

What happens if the object is not infinitely far away?

We can distinguish 3 cases:

1. The object is at the focus
2. The object is between focus and mirror
3. The object is beyond the focus

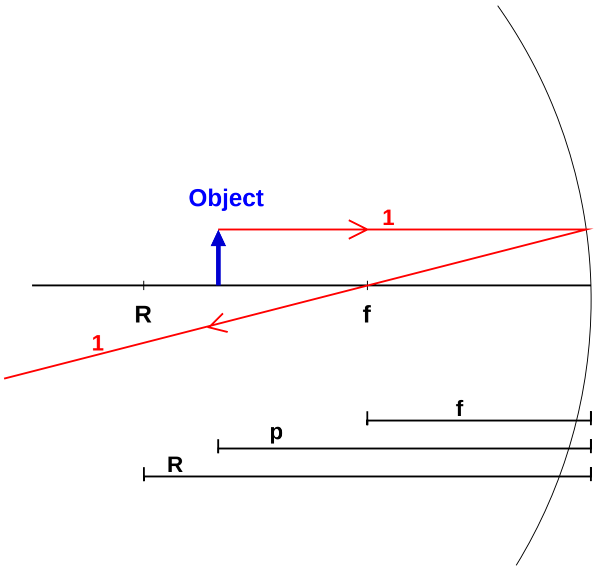


To find out where the image will be in all these cases we must come up with a quick way to construct the image.

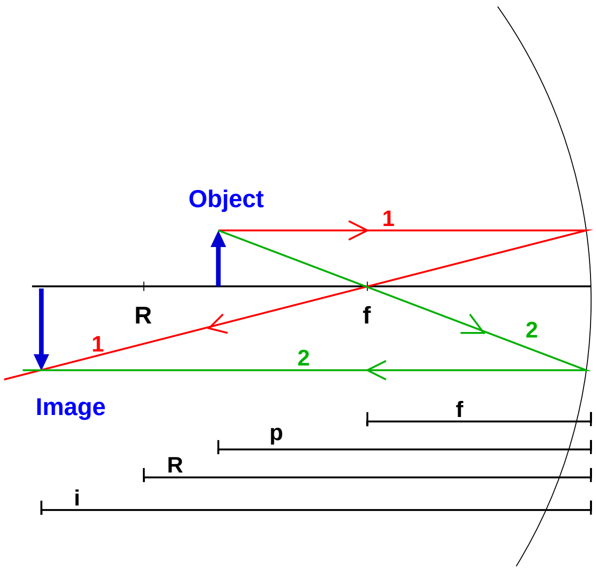
3.13: Constructing Images

We can easily construct the image using only a ruler if we recall the properties of the focal point:

1) Construct a ray from the object parallel to the optical axis. Its reflection will go through the focus.

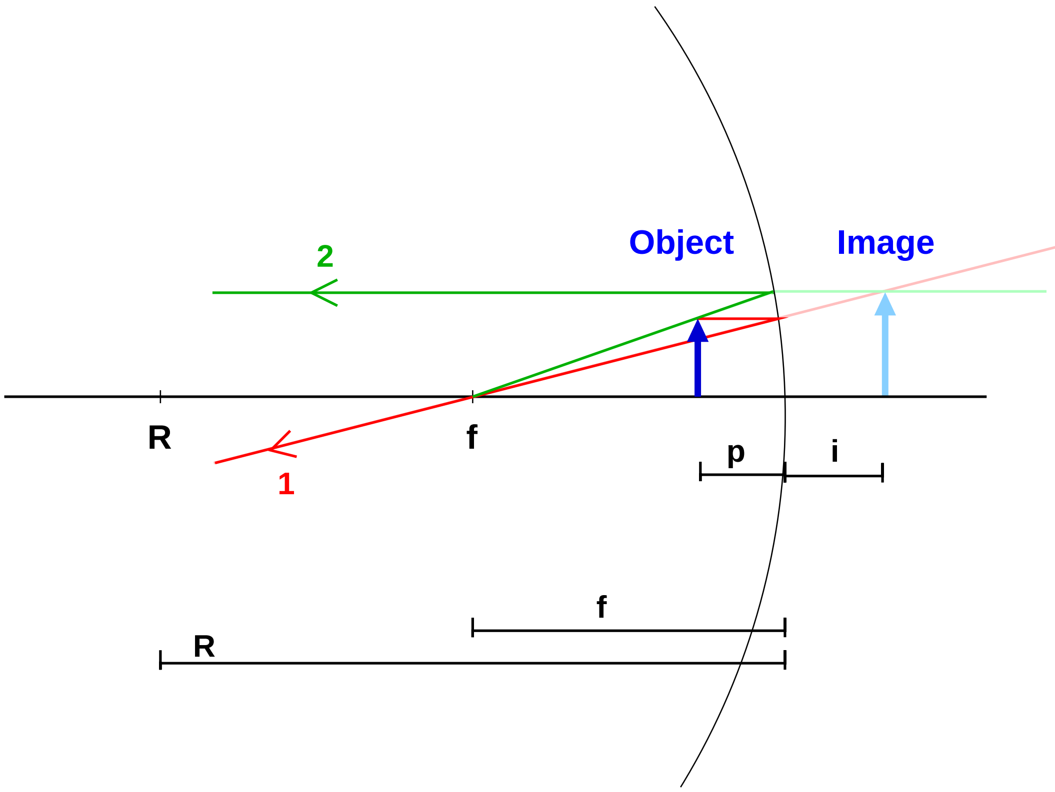


2) Construct a ray from the object through the focus to the mirror. It will be reflected parallel to the optical axis and it will intersect the first ray at the position of the image.



Here we get a real but inverted image!

If necessary, continue to draw the reflected rays on the other side of the mirror.



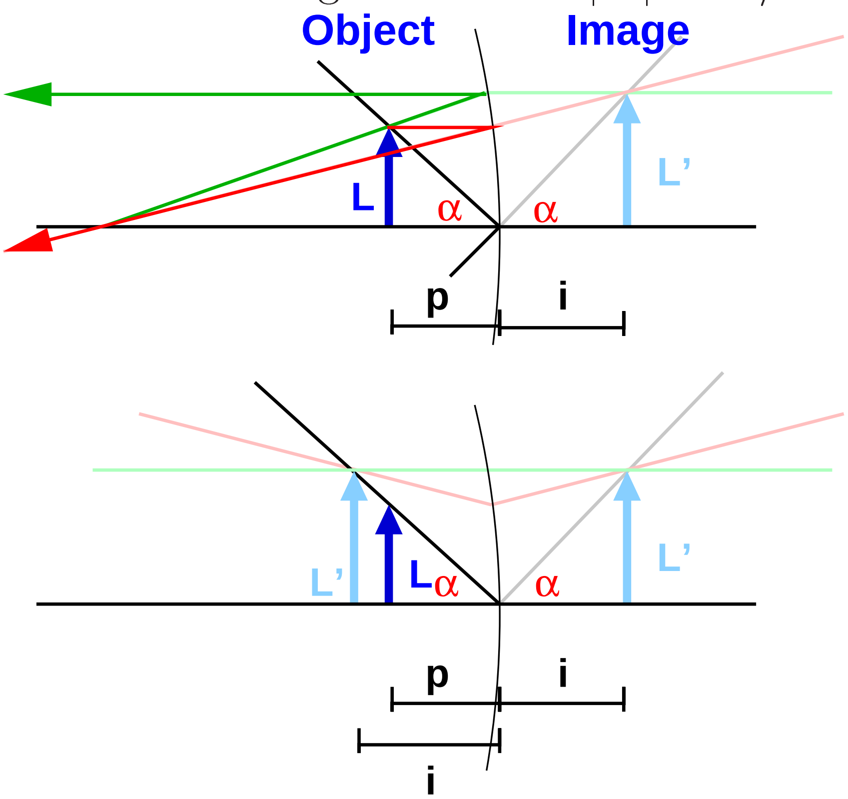
This construction gave us a virtual but upright image.

This method to generate images is universal and can be used for spherical mirrors as well as for lenses. We will use it a lot!

3.14: Magnification

When we look at a plane mirror, the image has the same size as the object. Clearly, with a curved mirror that is not the case. What is the magnification of a concave mirror with a radius R?

Let the height of the arrow be L. The height of the image is L' and the magnification is $|m| = L'/L$.

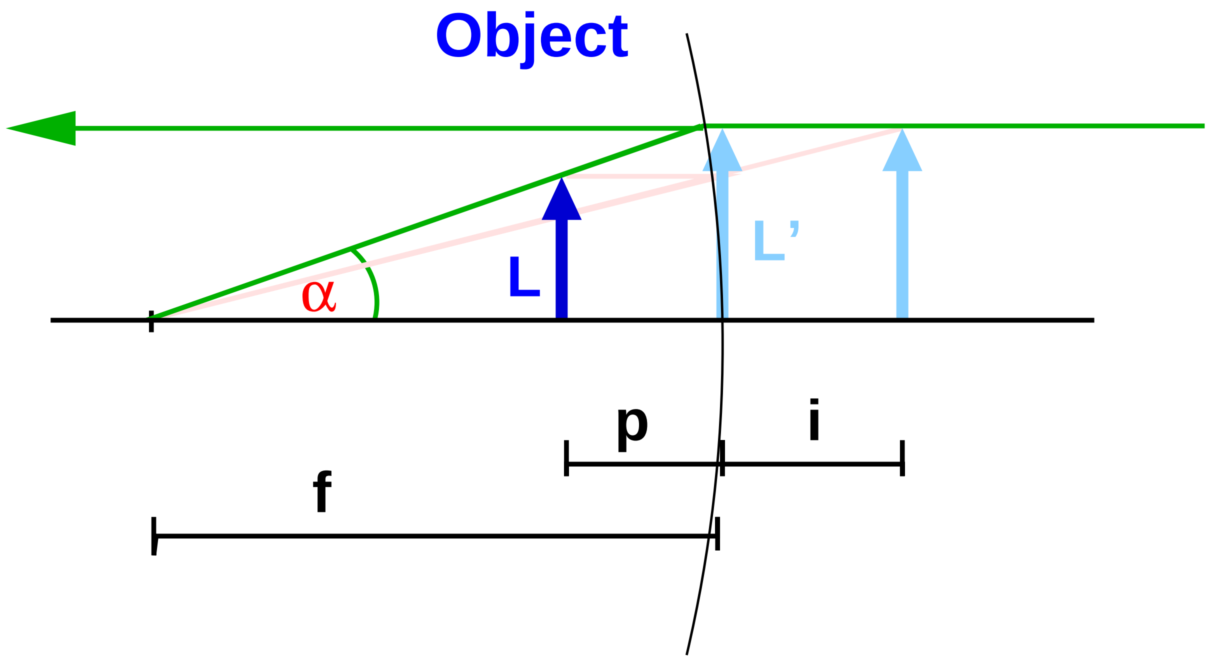


From the similar triangles we find

$$L' = |i| \tan\alpha \qquad L = |p| \tan\alpha \qquad |m| = \frac{L'}{L} = \pm\frac{|i|}{|p|}$$

The magnification has a sign: It is positive if the image is upright, it is negative if the image is inverted.

We can also easily relate the image distance i the object distance p and the focal length f:



$$L' = |f| \tan\alpha \qquad L = (|f| - |p|) \tan\alpha$$

$$\frac{L}{L'} = |m^{-1}| = \frac{|f| - |p|}{|f|} = \frac{|p|}{|i|}$$

$$|f| - |p| = \frac{|f||p|}{|i|}$$

$$|f| \left(1 - \frac{|p|}{|i|}\right) = |p|$$

$$|f| = \frac{|p||i|}{|i| - |p|}$$

$$\frac{1}{|f|} = \frac{1}{|p|} - \frac{1}{|i|}$$

Now we have to take into account the signs.

We define the radius R and the focal length $f = R/2$ to be positive for concave mirrors. We also take the object distance p to be positive.

How do we deal with virtual images? Easy: We define the image distance for a virtual image to be negative! Specifically, when we wrote earlier for a plane mirror $|p| = |i|$ what we meant was that for a plane mirror we have

$$p = -i$$

If we put this sign convention into effect we get

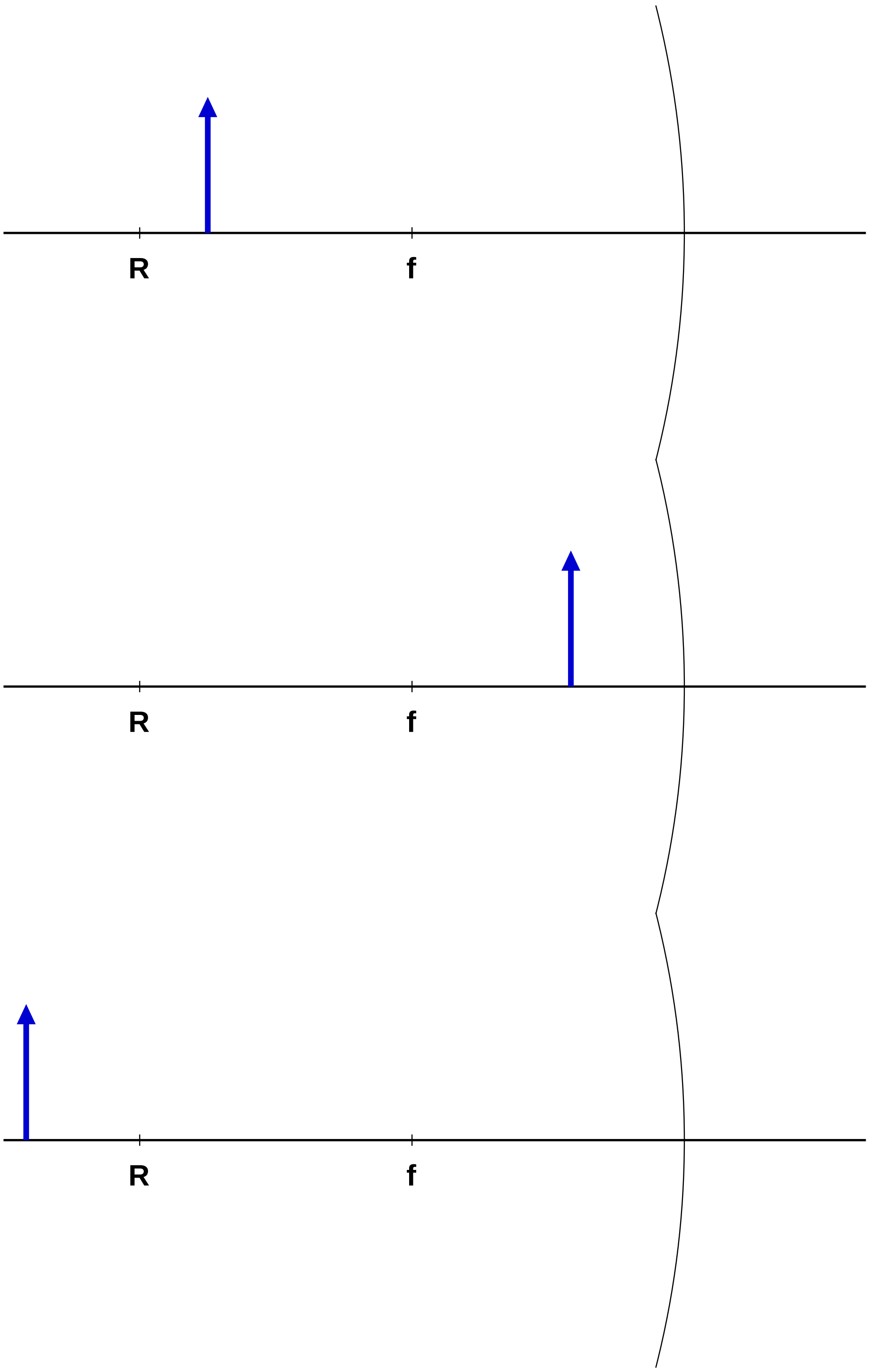
$$\frac{1}{f} = \frac{1}{p} + \frac{1}{i}$$

And the magnification becomes

$$m = -\frac{i}{p}$$

This way we automatically have $m > 0$ for virtual images which means they are upright. The real image in our earlier example was inverted. Since it is real, i and p are both positive, and $m = -i/p$ is negative - everything is consistent!

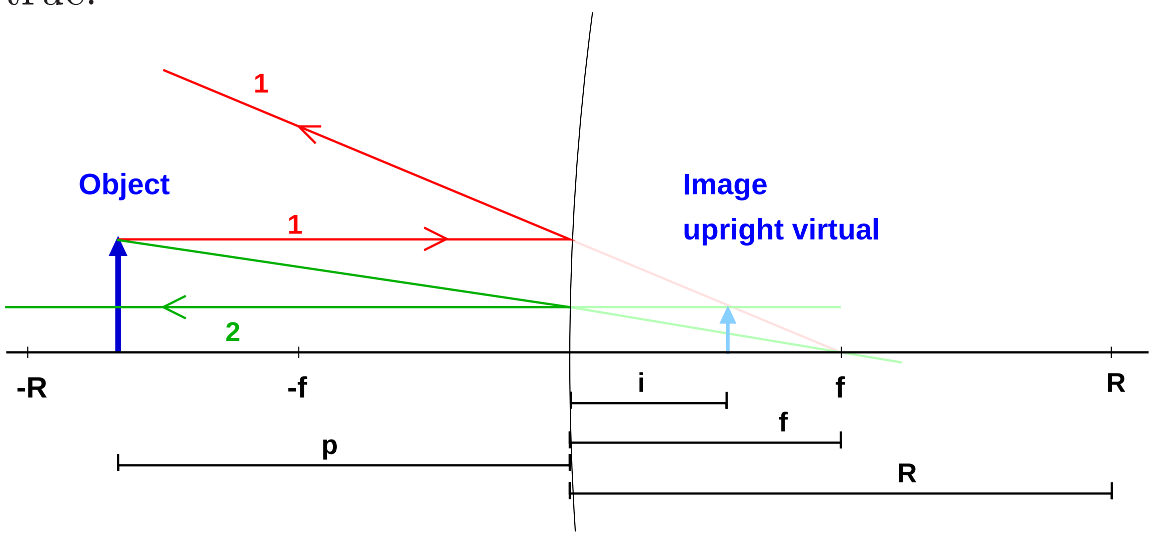
3.15: Examples

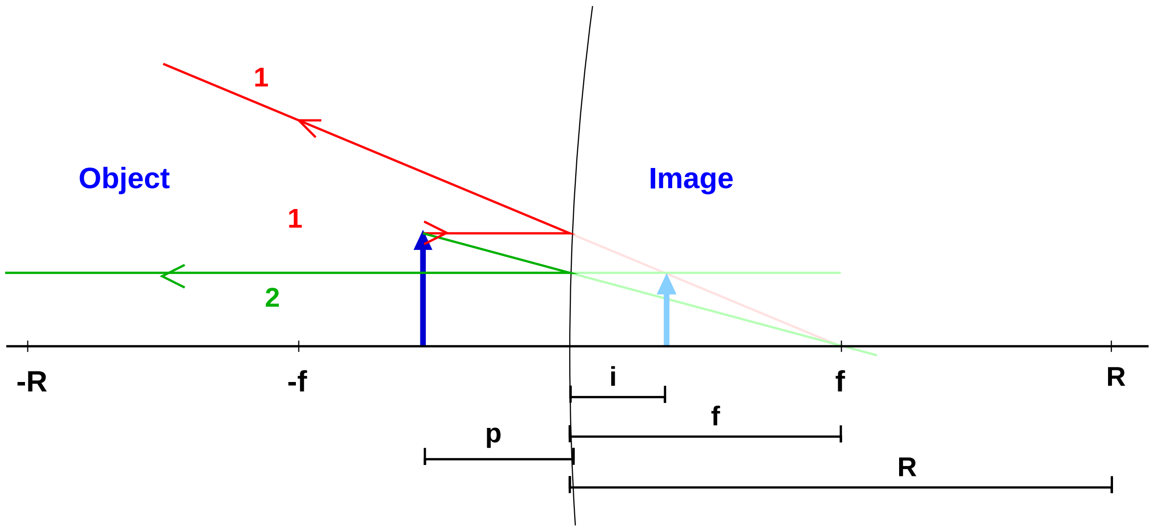


3.16: Convex mirrors

Convex mirrors can be dealt with in exactly the same way as before. The focal point is now behind the mirror, it is a **virtual** focus. All our rules for constructing images are still true, though.

This means that for a convex mirror both the focal length f and the radius R are negative. $f = R/2$ is still true.





Mirror Checklist

Mirror	Object	Image		Sign of			Magnif.
Type	location	location	type	f	i	p	—m—
plane	anywhere	behind m.	upr. virtual	$+\infty$	-	+	1
concave	beyond R	betw. R and f	inv. real	+	+	+	< 1
	betw. R and f	beyond R	inv. real	+	+	+	> 1
	inside f	behind m.	upr. virtual	+	-	+	> 1
convex	anywhere	inside f	upr. virtual	-	-	+	< 1

3.17: Thin lenses

Another very common optical imaging device is the lens. At first we restrict ourselves to the simplest case: symmetrical thin lenses.

The radius of curvature is the same for both sides of the lens, and the distance between the two surfaces is so small, that we can assume both refractions to take place at the same position.

Since the index of refraction can vary from lens to lens, we no longer have a simple relationship between the focal length f and the radius R.

We also now have **two** focal points on either side of the lens.

To alleviate all fears of boredom, the sign conventions for lenses are not the same as for mirrors.

We can have converging and diverging lenses.

3.18: Converging Lenses

The construction of the image is very similar:

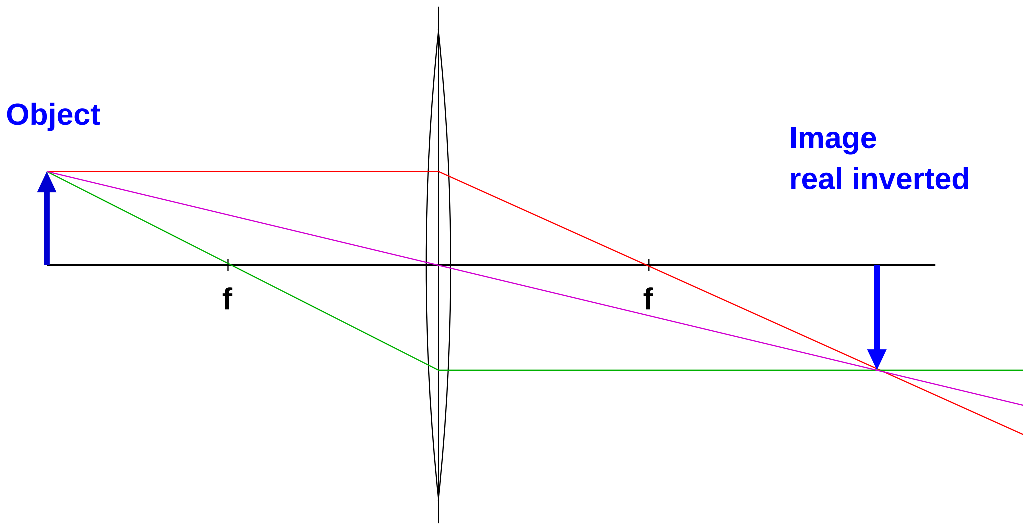
First draw a ray from the object parallel to the optical axis. It will be refracted into the focus on the opposite side of the lens.

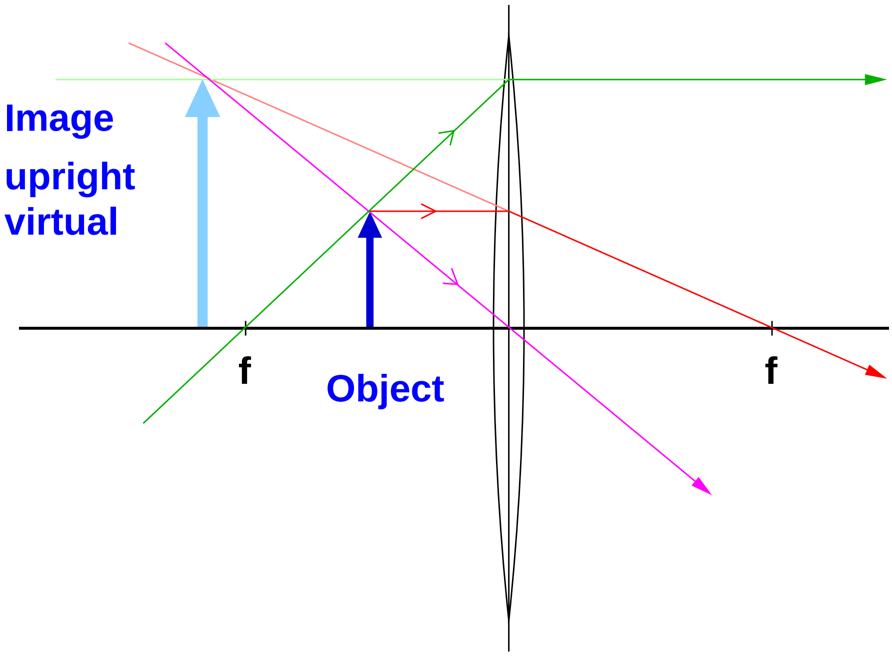
Second draw a ray connecting the object and the focus on the same side of the lens. It will be refracted parallel to the optical axis.

Third draw a ray from the object through the center of the lens. It will pass through unrefracted.

All three rays meet in the image.

If necessary, the rays have to be continued onto the other side of the lens to form a virtual image.





The relationship between the image distance, the object distance and the focal length is still

$$\frac{1}{f} = \frac{1}{p} + \frac{1}{i}$$

For a converging lens f and p are positive. i is positive for a real image and negative for a virtual image.

The magnification is $m = -i/p$.

A converging lens creates inverted real images on the far side of the object and upright virtual images on the same side as the object.

3.19: Summary

- Plane mirrors produce upright virtual images behind the mirror.
- The focal length f is related to the object and image distances p and i via $f^{-1} = p^{-1} + i^{-1}$ The signs of f,p, and i are important!
- The magnification of an image is $m = -i/p$. Again, the signs are important. A negative value of m means an inverted image.
- A concave mirror can produce real and virtual images. Its radius of curvature is $R = 2f > 0$.
- A convex mirror can only produce virtual images. Its radius of curvature is $R = 2f < 0$.
- The images from mirrors can be constructed with a ruler.
- Thin lenses obey the same rules as mirrors: $f^{-1} = p^{-1} + i^{-1}$ but the signs are different: converging lenses (with convex surfaces) have negative radii and local lengths, diverging lenses (with concave surfaces) have positive radii and focal lengths. The relations $m = -i/p$ and $f^{-1} = p^{-1} + i^{-1}$ still hold.

Exercises

Halliday Resnick & Walker:

Reading: HRW pp. 834-848

Exercises: p. 856 ff.: Q4, Q5, Q6, Q8, 10P, 13P, 14P, 24P, 30P

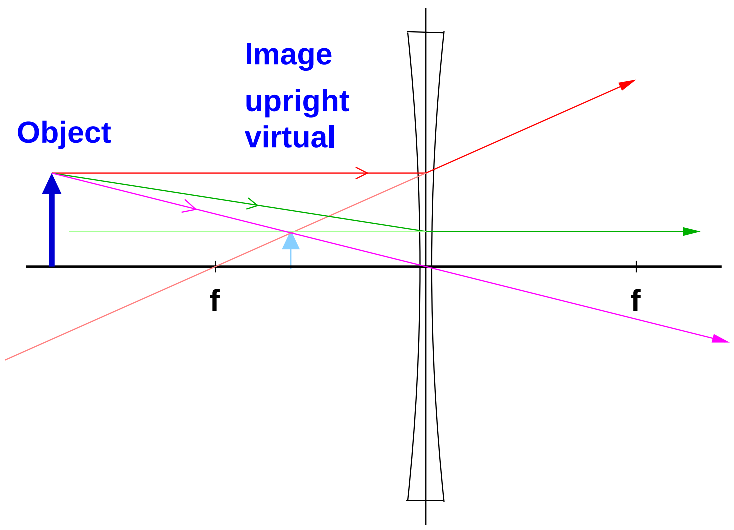
3.20: Diverging Lens

First draw a ray from the object parallel to the optical axis. It will be refracted so that the backward continuation of the refracted ray goes through the focus on the same side of the lens.

Second draw a ray from the object through the lens into the focus on the opposite side of the lens. It will be refracted parallel to the optical axis.

Third draw a ray from the object through the center of the lens. It will pass through unrefracted.

All three rays meet in the image.



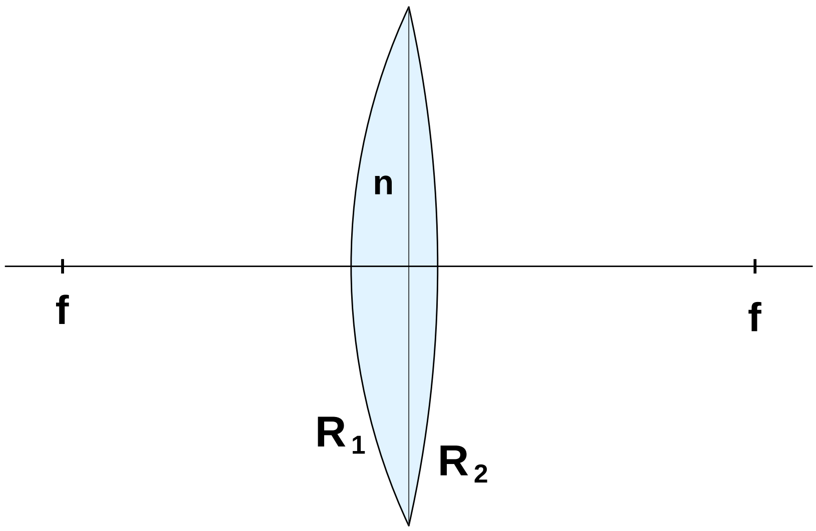
For a diverging lens f is negative and p is positive. Only virtual images can be formed with a diverging lens and i is always negative. The magnification is $m = -i/p > 0$.

3.21: The Lensmakers Equation

The focal length f and radius of curvature R of a mirror were connected by $f = R/2$. In lenses we can have two radii of curvature, but only one focal length. The focal length must also depend on the index of refraction n. So how can we calculate the focal length from the radii of a lens?

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

R_1 is the radius of the first surface.



For a symmetrical ($R_2 = -R_1$) thin lens made of flint glass ($n = 1.5$) we get as a crude approximation:

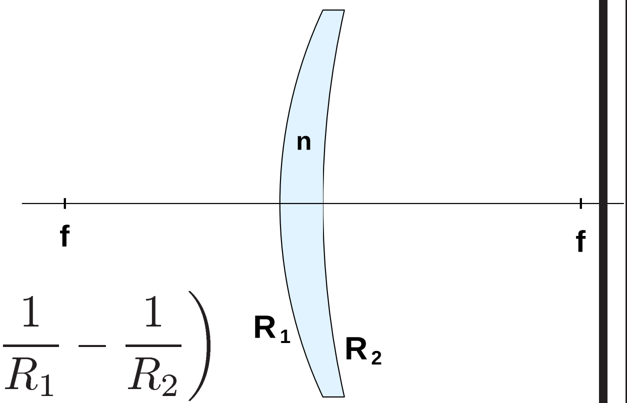
$$f = \frac{1}{1.5-1}\frac{R_1}{2} = R_1$$

3.22: Examples

Glasses have lenses with one concave and one convex surface: $R_1 = +60\,\text{cm}$, $R_2 = +90\,\text{cm}$. $n = 1.60$. What is

the focal length? Is this lens converging or diverging?

If an object is placed 150 cm in front of the lens, what kind of image will be formed where? What is the magnification?



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

$$f = \frac{1}{(1.60-1)\left(\frac{1}{60\,\text{cm}} - \frac{1}{90\,\text{cm}}\right)} = 300\,\text{cm}$$

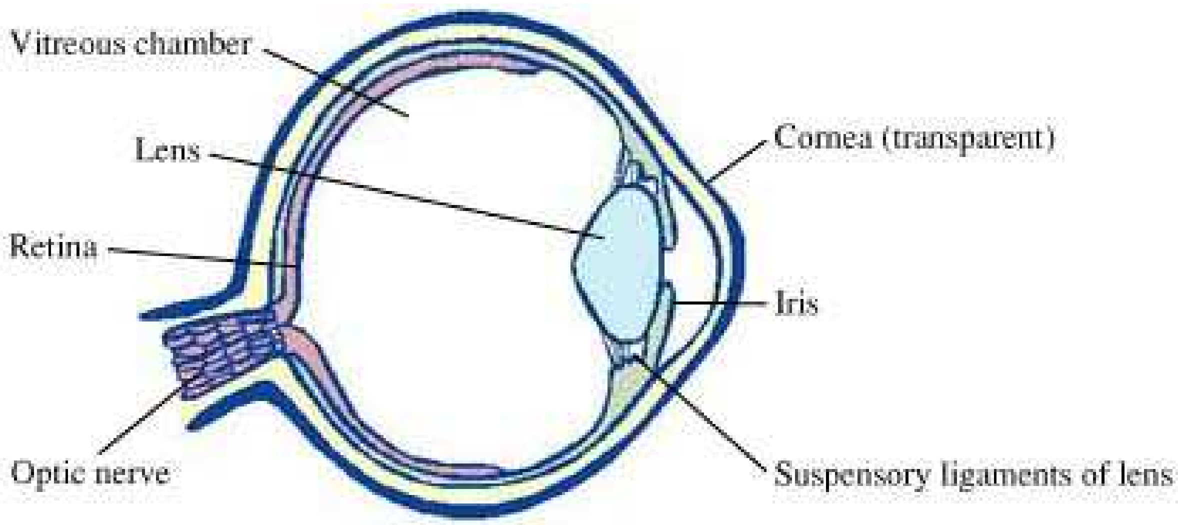
The focal length is positive: it is a converging lens.

The image distance is calculated as

$$\frac{1}{f} = \frac{1}{p} + \frac{1}{i} \Leftrightarrow i = \frac{1}{\frac{1}{300\text{cm}} - \frac{1}{150\text{cm}}} = -300\text{cm}$$

Magnification: $m = -i/p = -(-300)/150 = +2.0$. It is a virtual, upright image twice the size of the object.

3.23: The human eye



The lens creates an inverted real image on the retina where sensors process the image into a nerve pattern which is passed to the brain.

The image on the retina is a real, inverted image, the brain automatically compensates.

The lens can change its focal length to produce sharp images on the retina. The closest distance for a sharp image is typically 25 cm.

The size of the object is determined by the **viewing angle** ϑ.

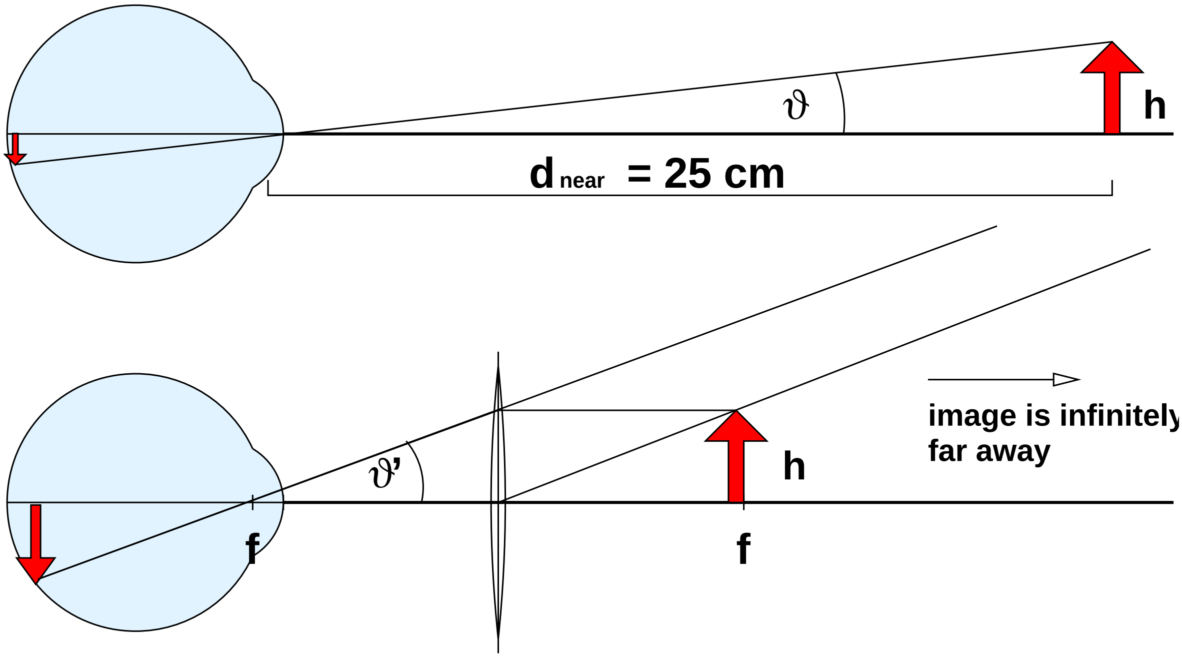
3.24: Optical Instruments

The **angular magnification** (do not confuse that with the *lateral* magnification!) is the ratio of the viewing angle with an optical instrument ϑ' to the maximum viewing angle obtainable with the naked eye ϑ:

$$m_\vartheta = \vartheta' / \vartheta$$

The reference angle ϑ depends on a convention for the viewing distance. In HR&W d_{near} is chosen as the reference distance. Usually the distance of most distinct vision ($d_d = 25\,\mathrm{cm}$) is chosen as the reference distance. In earlier editions of HRW $d_d = 15\,\mathrm{cm}$ is sometimes used.

The easiest optical instrument is the magnifying glass. It consists of a single converging lens. The object is placed at or just inside the focal length of the lens.



The maximum viewing angle without the lens is
$\tan\vartheta \simeq \vartheta \simeq h/d_{near}$

The image at an infinite distance appears under the viewing angle $\tan\vartheta' \simeq \vartheta' \simeq h/f$. The total angular magnification is therefore:

$$m_\vartheta = \frac{d_{near}}{f} = \frac{25\text{cm}}{f}$$

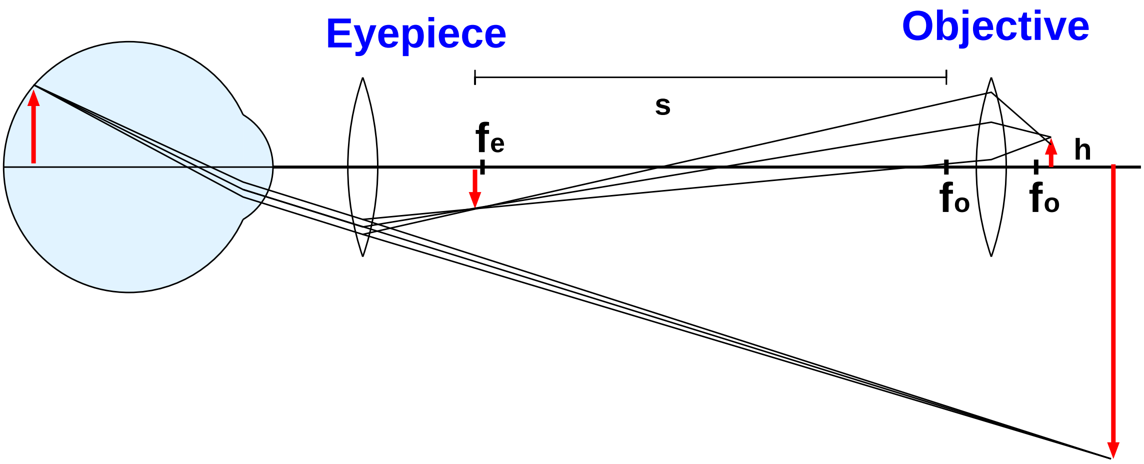
We can obtain magnifications of 10-20 with good magnifiers.

3.25: The Microscope

The microscope probably was invented by the Dutch lensmakers Hans and Zacharias Janssens between 1590 and 1610 in Middleburg, The Netherlands. The modern, two lens compound microscopes, however, have not been widely used until the end of the 18^{th} century, again by Dutch lensmakers Jan and Harmanus van Deyl.

It consists of two converging lenses, the Objective and the Eyepiece.

The object is placed just outside the focus of the objective, creating a huge, inverted real image. This real image is then viewed with the eyepiece acting as a magnifier.



The magnification of the real image is

$$m_o = -\frac{i}{p} = -\frac{f_o + s}{f_o} \simeq -\frac{s}{f_o}$$

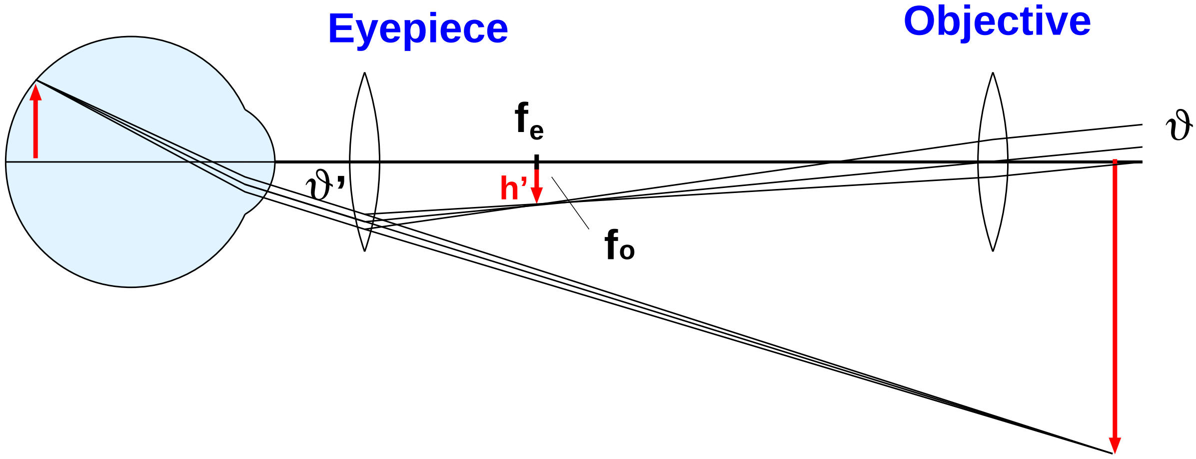
The angular magnification of the eyepiece acting as a magnifier is

$$m_\vartheta = \frac{25\text{cm}}{f_e}$$

The total magnification is the product of the magnifications of objective and eyepiece:

$$M_{tot} = m_\vartheta m_o = -\frac{25\text{cm}}{f_e}\frac{s}{f_o}$$

3.26: The Refracting Telescope



The object is at a distance D very far away, so the image will be produced very close to the focus of the objective. The focus of the eyepiece is chosen to coincide with the focus of the objective and again acts as a magnifier to view the enlarged, real, inverted image of the objective.

What is the magnification obtainable with this setup?

$$\tan\vartheta \simeq \vartheta = h/D$$

The real image inside the telescope has a size

$$h' = -f_o \tan\vartheta \simeq -f_o\vartheta$$

It is viewed under a viewing angle of

$$\tan\vartheta' \simeq \vartheta' = h'/f_e$$

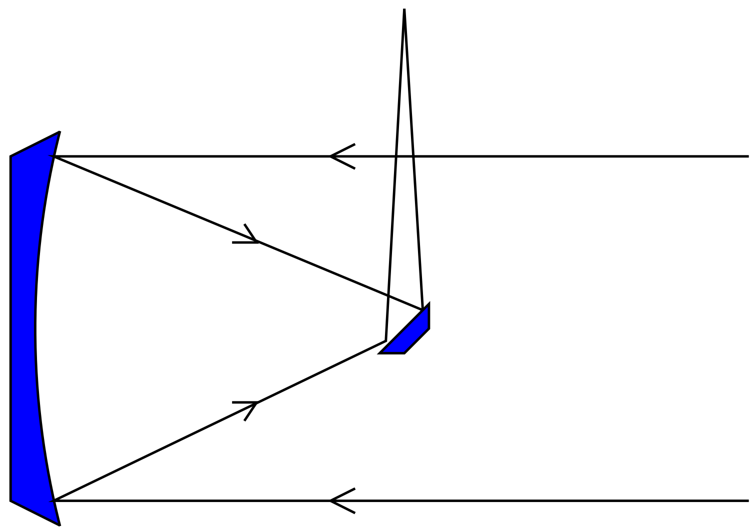
We have for the angular magnification

$$m_\vartheta = \frac{\vartheta'}{\vartheta} = \frac{h'}{f_e}\frac{1}{\vartheta} \simeq \frac{-f_o\vartheta}{f_e\vartheta} = -\frac{f_o}{f_e}$$

3.27: The Astronomic Telescope

Astronomical telescopes consist of a system of parabolic mirrors rather than lenses. The main advantage is that they do not suffer from chromatic aberration.

The most popular design goes back to Sir Isaac Newton:



How do astronomical telescope help us seeing faint objects in the sky?

Point sources: A point source will remain a point source even with very high magnification. But all the light from that source will stay in one point while all the background light will be distributed over a larger area. Thus the contrast between the point source and the background is improved and you can see it.

Extended objects: Extended objects are magnified and can therefore be seen more easily.

3.28: Summary

- Diverging lenses can only produce virtual images.
- The magnifier lens has an angular magnification of

$$m_\vartheta \simeq \frac{25\text{cm}}{f}$$

.

- The microscope has a magnification that is the product of the individual magnifications of objective and eyepiece.

$$m_\vartheta = \simeq -\frac{25\text{cm} \cdot s}{f_o f_e}$$

- The refracting telescope has a magnification

$$m_\vartheta \simeq -\frac{f_o}{f_e}$$

Exercises

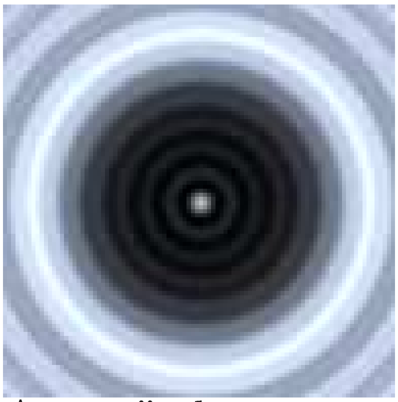
Halliday Resnick & Walker:

Reading: HRW Ch35, pp843-855.

Exercises: pp 858ff: 17E, 18E, 19E, 20E, 23P, 24P, 26P, 32E, 33E, 37P

4: Wave Optics

In 1818, Augustin Fresnel submitted a paper on the theory of diffraction for a competition sponsored by the French Academy. His theory represented light as a wave, as opposed to a bombardment of hard little particles, which was the subject of a debate that lasted since Newton's day. Simeon Poisson, a member of the judging committee for the competition, was very critical of the wave theory of light. Using Fresnel's theory, Poisson deduced the seemingly absurd prediction that a bright spot should appear behind a circular obstruction, a prediction he felt was the last nail in the coffin for Fresnel's theory.

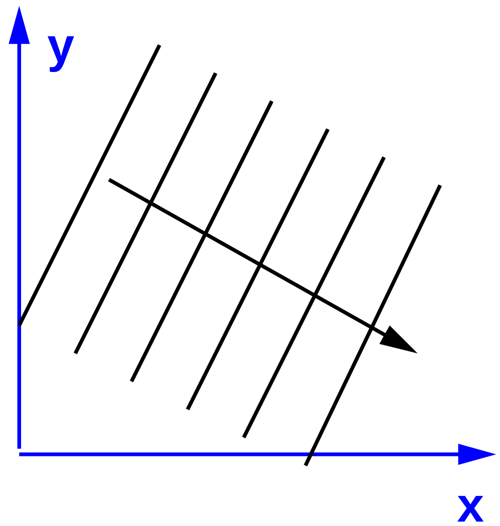


However, Dominique Arago, another member of the judging committee, almost immediately verified the spot experimentally. Fresnel won the competition, and, although it may be more appropriate to call it "the Spot of Arago," the spot goes down in history with the name "Poisson's bright spot" like a curse.

4.1: Plane and spherical waves

We can no longer use a one dimensional description for waves. We distinguish two main cases:

1. Plane waves or wavefronts
2. Spherical waves

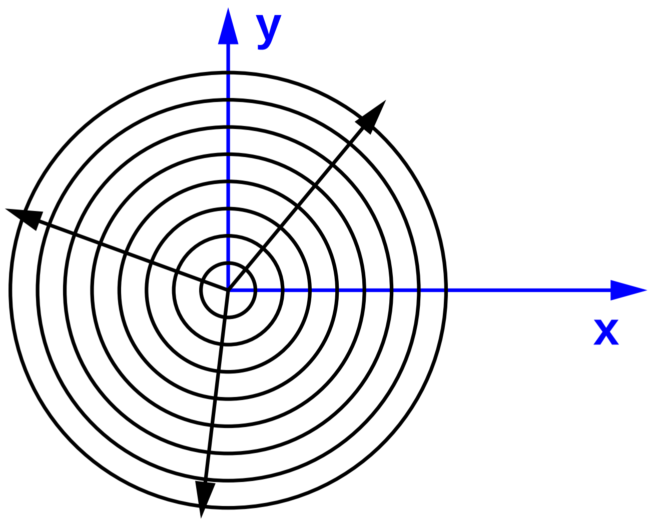


Plane waves are descibed by

$$s(\vec{r}, t) = s_0 \sin(\vec{k}\vec{r} - \omega t - \Phi_0)$$

with the position $\vec{r} = (x, y, z)$ and the wave vector $\vec{k} = (k_x, k_y, k_z)$. If you choose the direction of motion as one of your coordinate axes, say the x-axis, then the wave vector $\vec{k}$ only has components in that direction $\vec{k} = (k_x, 0, 0)$ and you get back the simple description

$$s(\vec{r}, t) = s_0 \sin(k_x x - \omega t - \Phi_0)$$



A spherical wave is described by

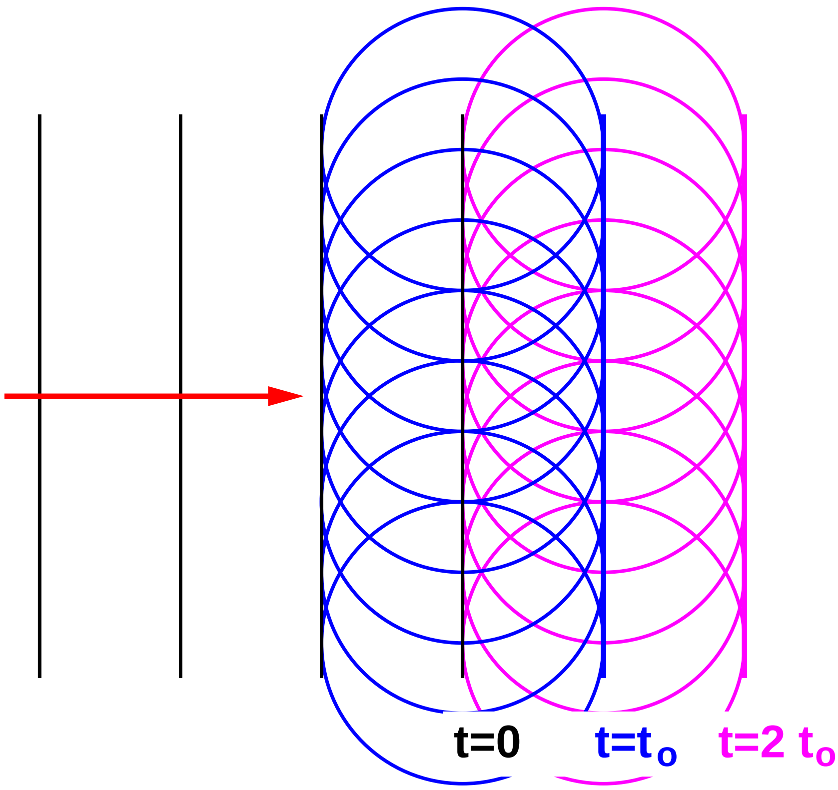
$$s(\vec{r}, t) = s_0 \sin(kr - \omega t - \Phi_0)$$

Note: no vectors. The phase of the wave depends on the distance to the origin, all points $\vec{r}$ on a sphere with radius $r = \sqrt{x^2 + y^2 + z^2}$ have the same displacement as a function of time.

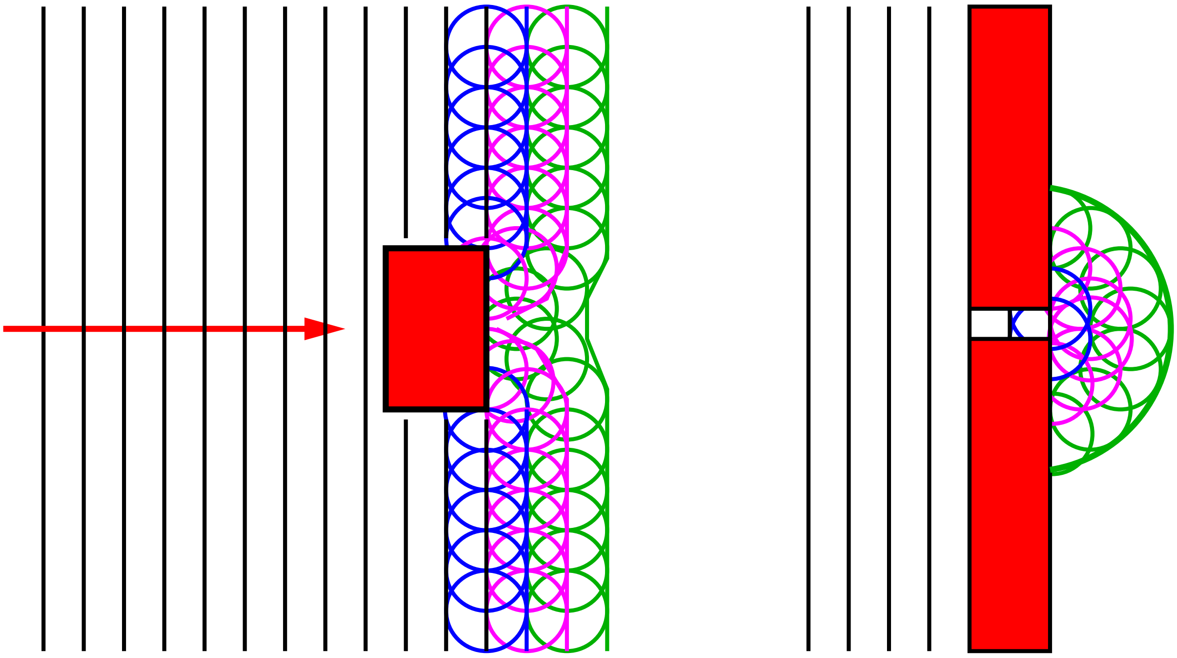
4.2: Huygens' Principle

The mechanism by which **any** wave propagates can be described by Huygen's Principle:

Each point on a wavefront acts as a source of a new spherical wavelet with the same phase. The envelope around these secondary wavelets after a time t is the new wavefront.



If the wavefront encounters an obstacle, this principle determines the motion of the wavefront:

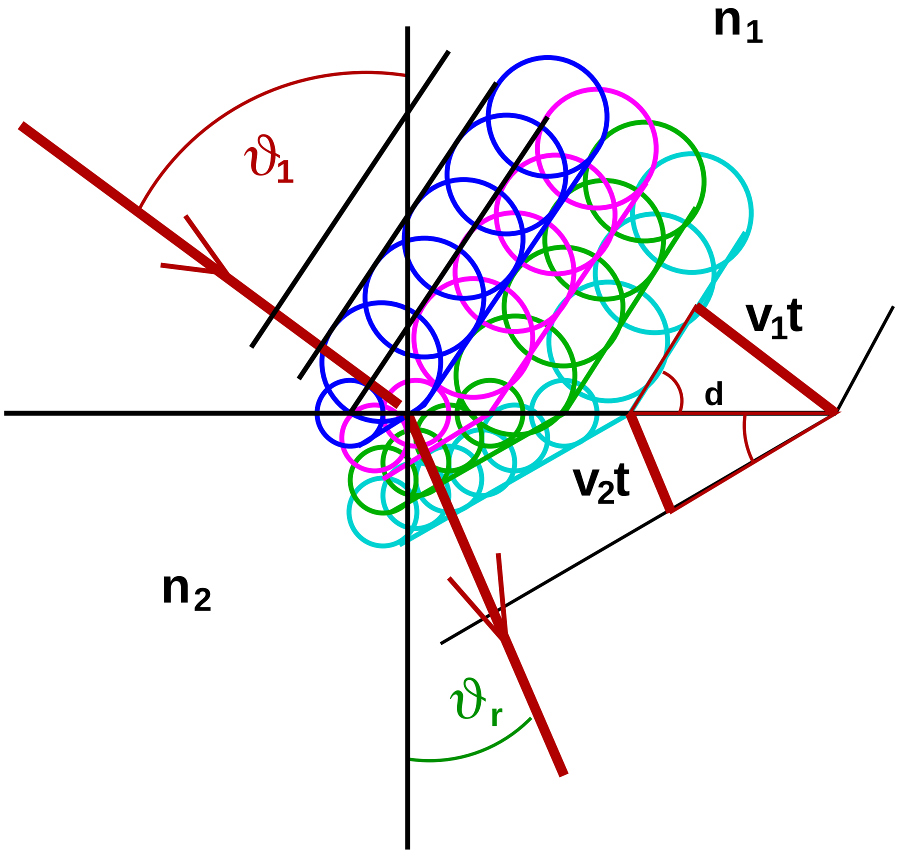


In general waves will diffract around obstacles. We distinguish two cases:

1. Fraunhofer diffraction: Here the waves are plane waves. The source and detector are infinitely far away from the obstacle.
2. Fresnel diffraction: The source and detector are at a finite distance from the obstacle. Here we have to consider spherical waves and the discussion becomes a bit more involved.

4.3: Snell's Law with waves

Take a wavefront incident on a refracting surface at an angle ϑ_1 as before. The speed of the wave in medium 1 is v_1, the speed in the second medium is v_2. The wavefront at the time $t = 0$ has just reached the surface, and according to Huygen's principle we must construct new wavelets from each point and find the envelope:



After a time t the left part of the wavefront will have progressed a distance $d = v_2 t$ while the right side of the wavefront has not quite reached the boundary yet.

We find that the wavefront in the second medium moves at an angle to the incident wave: it is refracted.

We can also determine the angle of refraction: From the construction we read that

$$d = \frac{v_1 t}{\sin\vartheta_1} = \frac{v_2 t}{\sin\vartheta_r}$$

$$\frac{1}{v_1}\sin\vartheta_1 = \frac{1}{v_2}\sin\vartheta_r$$

This is valid for any wave crossing a boundary between two media with a different wave speeds!

For light we use $n = c/v$ and get:

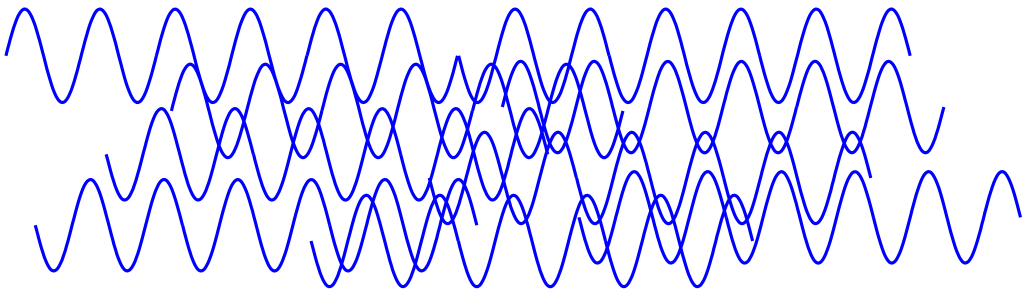
$$n_1 \sin\vartheta_1 = n_2 \sin\vartheta_r$$

Snell's law!

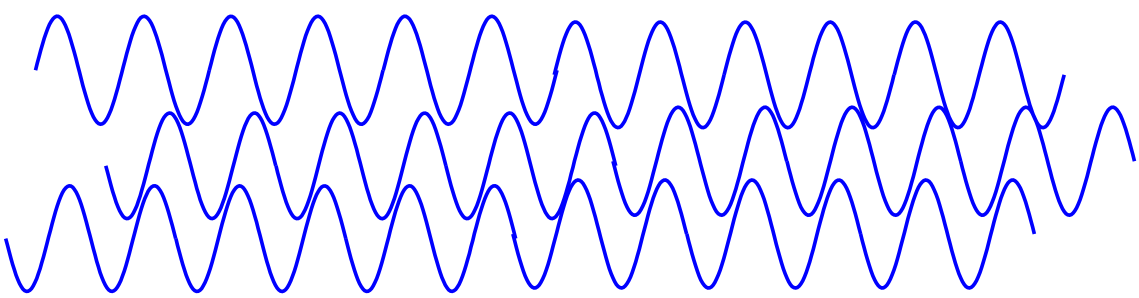
4.4: Interference

For light just like for any other waves the principle of superposition holds and we can observe constructive and destructive interference.

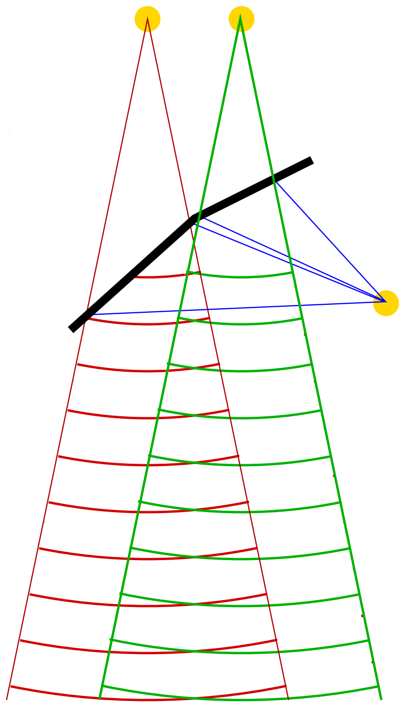
In general light from a lamp consists of many small wavetrains with random phases between them. We see only an average intensity where all constructive and destructive interferences cancel out.



If we want to observe an interference pattern, we have to provide lightwaves with a **fixed** phase between them - we need coherent light.



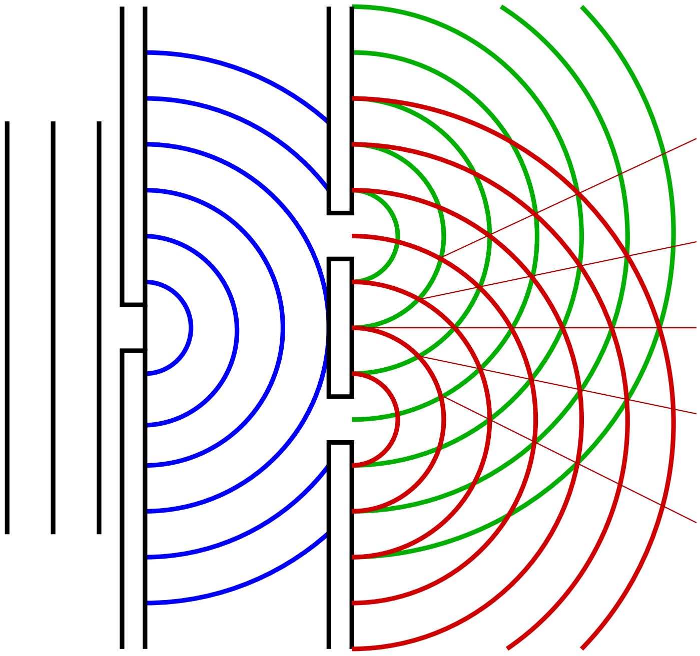
One easy way to create two coherent sources of light is given through Fresnel's double mirrors:



Here the two virtual images of the source act as two coherent sources of light and we can observe an interference pattern.

4.5: Young's Double Slits

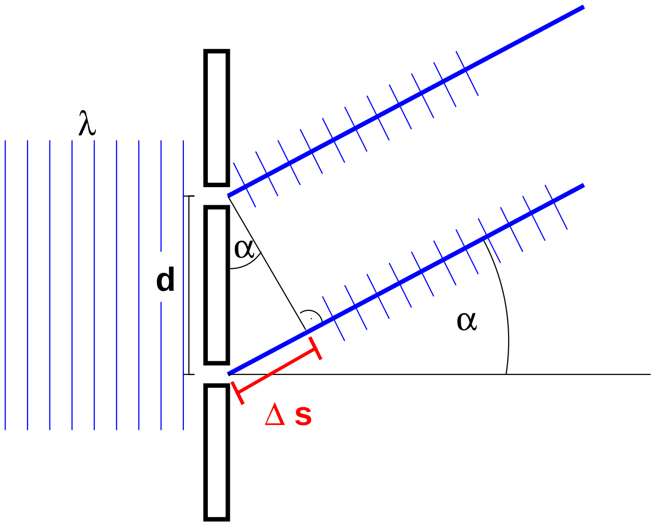
Another way to create coherent light is used in Young's double slit experiment.



A plane wavefront hits a narrow slit and is diffracted. If the slit is small enough, only one Huygens wavelet becomes the source for the entire wave beyond the slit.

This (now spherical) wavefront hits another set of two slits. These two slits now act as two coherent sources of light since they were created from the same wavelet and the only phase difference at the slits is created through the constant different in pathlengths.

We can now calculate the positions of the bright spots on a screen far behind the double slits.



The two wavefronts travelling at an angle α will be in phase at their respective origins, the slits. The phase difference on the screen will be determined by the different pathlengths they have to travel to get there.

To observe a bright spot, we must have constructive interference, i.e. the path difference must be an integer multiple of the wavelength λ:

$$\frac{\Delta s}{\lambda} = n \quad \Leftrightarrow \quad \frac{d \sin \alpha_{max}}{\lambda} = n \quad \Leftrightarrow \quad \sin \alpha_{max} = \frac{n\lambda}{d}$$

Similar arguments lead to the observation of minima when the path difference is just half a wavelength:

$$\Delta s = n\lambda + \tfrac{1}{2}\lambda \quad \Leftrightarrow \quad \sin \alpha_{min} = \frac{2n+1}{2}\frac{\lambda}{d}$$

4.6: Summary

- Waves will diffract around obstacles. The effect is strongest when the dimensions of the object and the wavelength are comparable.
- Huygens principle states that **each point on a wavefront acts as a source of a new spherical wavelet with the same phase. The envelope around these secondary wavelets after a time t is the new wavefront.**
- Fresnel diffraction has to be calculated with spherical waves going in and out.
- Fraunhofer diffraction assumes the wave source and detector to be infinitely far away from the obstacle.
- Snell's law is a dircet consequence of the fifferent wave speeds in two media.
- Light obeys the principle of superposition. In order to see stable interference patterns we must use coherent light sources.
- Young's Double Slit Experiment shows a stable interference pattern. The bright spots are found at an angle $\sin \alpha_m ax = n\lambda/d$, the interference minima are found at angles $\sin \alpha_{min} = (2n + 1)\lambda/(2d)$.

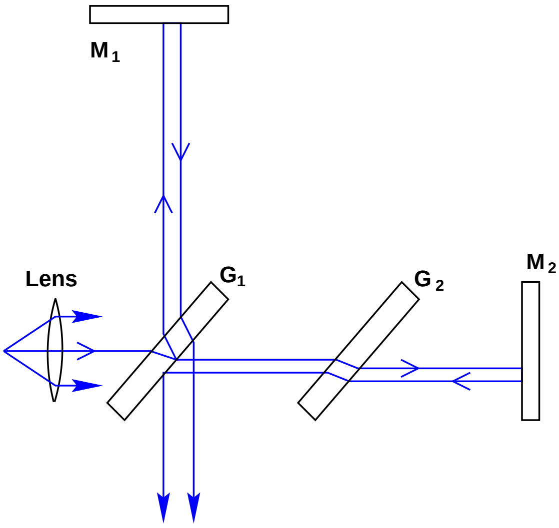
Exercises

Halliday Resnick & Walker:

Reading: HRW CH36, pp. 862-870

Exercises: p. 882 ff.: Q3, Q4, 5P, 11E, 13E, 16E, 21P

4.7: Interferometry



In the previous examples (Young's slits, Fresnel's Mirrors) the coherent light was provided through the division of the wave front. Another way to get coherent light is by division of the amplitude. The most prominent representative of this class of interferometers is the Michelson Interferometer:

The glass plate G_1 is sometimes silvered so that the transmitted and reflected beams have equal intensities.

The glass plate G_2 is inserted so that the optical path lengths in glass are equal in both arms.

In order to observe interference fringes three requirements must be fulfilled:

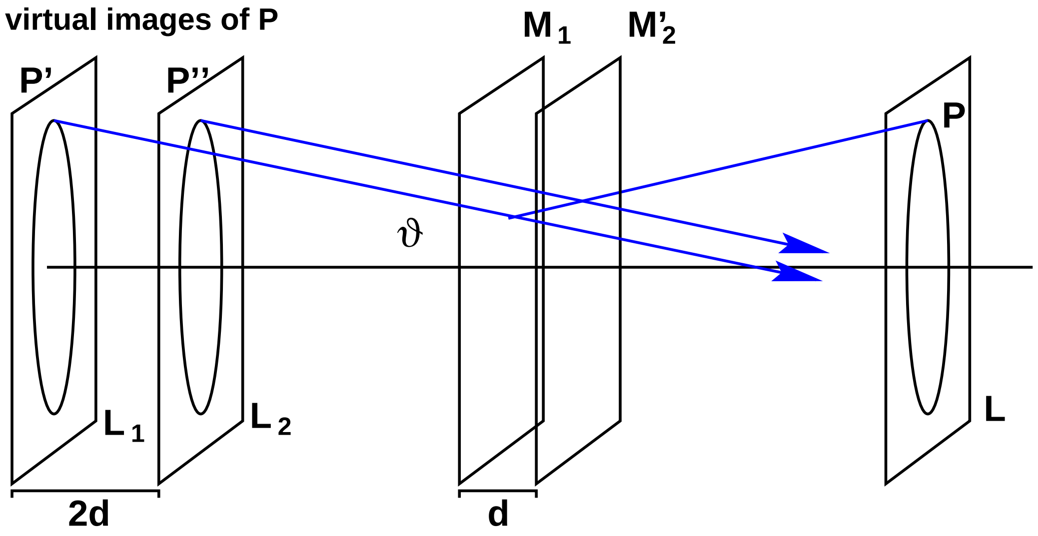
1) The light source must be extended

2) The mirrors must be absolutely perpendicular.

3) The light used must be monochromatic

4.8: Fringe formation

In the Michelson interferometer circular fringes are seen:



The reflection of the same spot of the extended source in both mirrors act as two coherent light sources. If the difference of the armslengths of the interferometer is d, then the two sources will be at a distance $2d$. The path difference for two rays at an angle ϑ to the optical axis is $2d\cos\vartheta$ and we get bright fringes when that pathdifference is a multiple of the wavelength:

$$2d\cos\vartheta = n\lambda$$

4.9: Fringe Intensity

What is the intensity of the fringes?

An electromagnetic wave is given by

$$\vec{E} = \vec{E}_0 \sin(\vec{k}\vec{r} - \omega t)$$

The intensity is the average power delivered by the wave and is given through the Poynting vector:

$$\vec{S} = \frac{1}{\mu_0} \vec{E} \times \vec{B}$$

Its magnitude is

$$S = \frac{EB}{\mu_0} = \frac{E^2}{c\mu_0} = \frac{E_0^2 \sin^2(\vec{k}\vec{r} - \omega t)}{c\mu_0}$$

We get the intensity as the time average of S:

$$I = \frac{1}{Tc\mu_0} \int_0^T E_0^2 \sin^2(\vec{k}\vec{r} - \omega t) \mathrm{d}\, t$$

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

The intensity is proportional to the square of the amplitude!

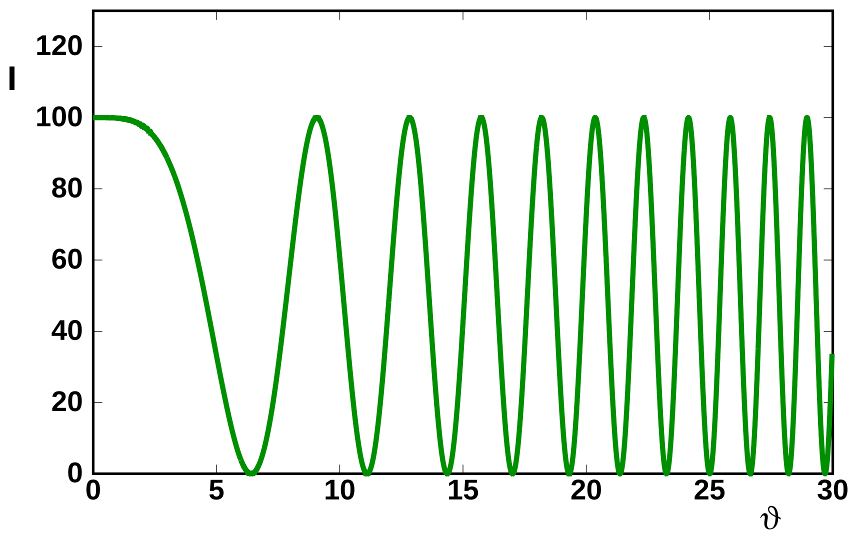
To calculate the intensity of the fringes in the Michelson interferometer, we calculate the superposition of waves at each point on the screen. Then we square the wave and take the time average.

We know the path difference on each point on the screen is $\delta = 2\pi 2d\cos\vartheta/\lambda$ so the total wave on a circle seen at an angle ϑ is:

$$\begin{aligned} y(R,t) &= E_0\sin(kR-\omega t) + E_0\sin(kR-\omega t-\delta) \\ &= 2E_0\cos(\delta)\sin(kR-\omega t-\delta/2) \end{aligned}$$

And the intensity is

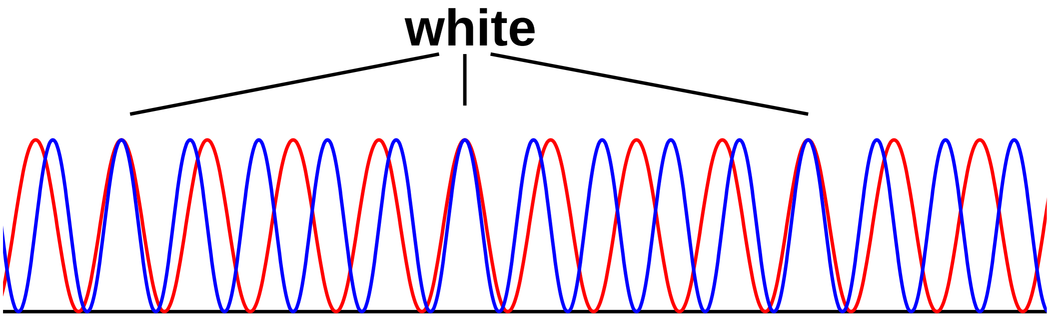
$$I = \frac{1}{2\mu_0 c}(2E_0\cos(\delta))^2 = I_0\cos^2(2\pi 2d\cos\vartheta/\lambda)$$



Plotted for $\lambda = 500\,\mathrm{nm}$, $d = 20\lambda$, $I_0 = 100\,\mathrm{W}$.

4.10: White Light Fringes

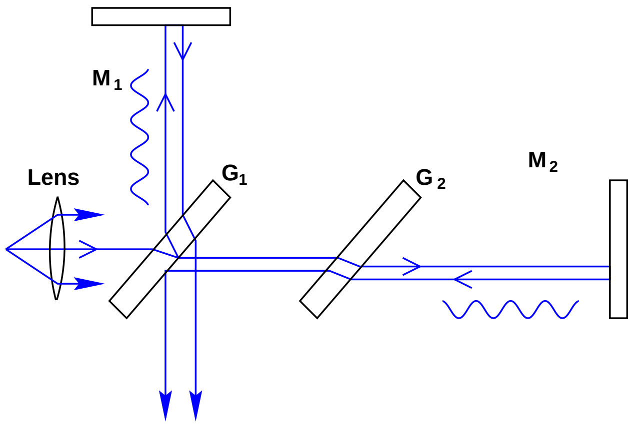
If white light is used the fringes disappear, except when the difference in optical path lengths vanishes: $d = 0$.



The spacing between fringes of different colours is different. In the center the path difference for all colours vanishes and we get a bright white spot surrounded by coloured fringes. Eventually enough coloured fringes meet again to form a white fringe.

4.11: Limits

How long can the path difference be? For white light it can not be more than a few wavelengths. For more monochromatic light the path difference can be found to be a few centimeters.



We can not get truly monochromatic light, even the narrowest spectral lines contain a range of wavelengths. They will eventually get out of step, just like white light fringes.

A different way of looking at the problem is that the light emitted by the source is emitted over a finite time producing a wavetrain of finite length. If the path difference for the wavetrain going through either arm of the interferometer exceeds the length of the wave, it can no longer produce an interference pattern.

The two pictures are actually equivalent and are at the heart of the uncertainty principle.

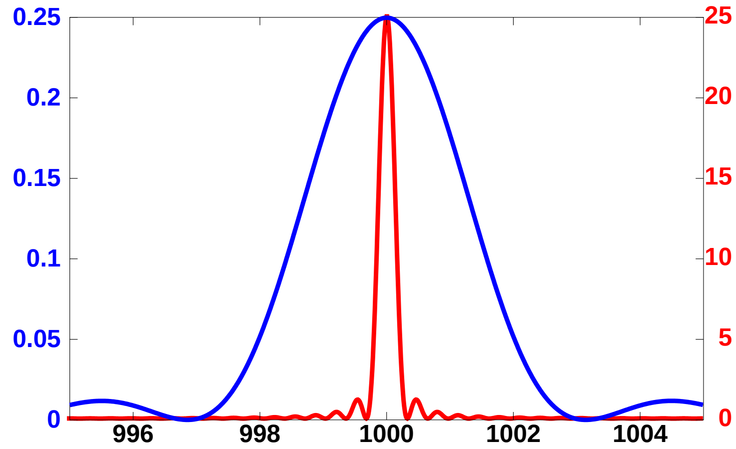
Consider this finite wavetrain:

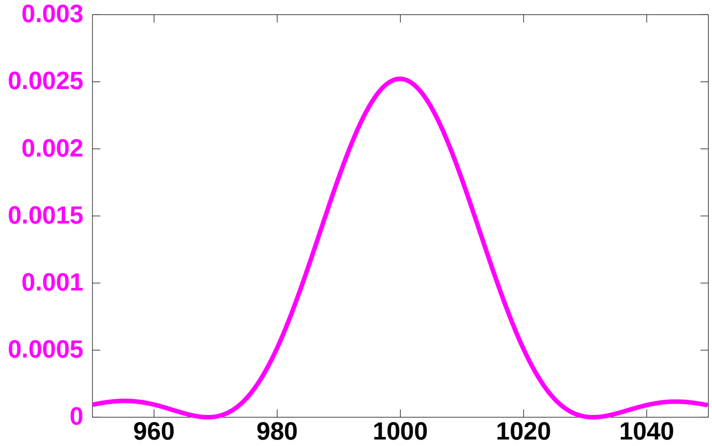
$$y(x,t) = \sin(kx - 1000t) \qquad 0 \leq t \leq T$$

What is its frequency? It seems like $\omega_0 = 1000 rad/s$, but that is only true for an infinite wave. We have to do a Fourier analysis and find that the range of frequencies required to form a finite wave increases as the wave grows shorter:

$$F(\omega) = \sqrt{\frac{2}{\pi}} \int_0^T \sin(\omega_0 t)\sin(\omega t)\,\mathrm{d}\,t$$

$$F(\omega) = \sqrt{\frac{2}{\pi}} \frac{\omega_0 \cos(\omega_0 T)\sin(\omega T) - \omega \sin(\omega_0 T)\cos(\omega T)}{\omega_0^2 - \omega^2}$$





Here $2(F(\omega))^2/\pi$ are plotted for $T = 10\,\mathrm{s}$ (red) $1\,\mathrm{s}$ (blue) $0.1\,\mathrm{s}$ (magenta).

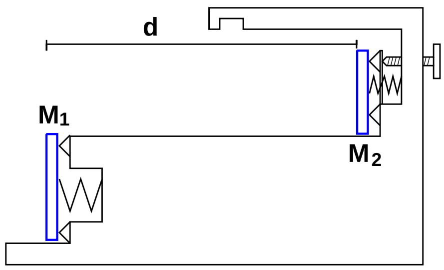
$$\Delta\omega \cdot T = \pi \quad \Leftrightarrow \quad \Delta E \Delta t = \frac{h}{2}$$

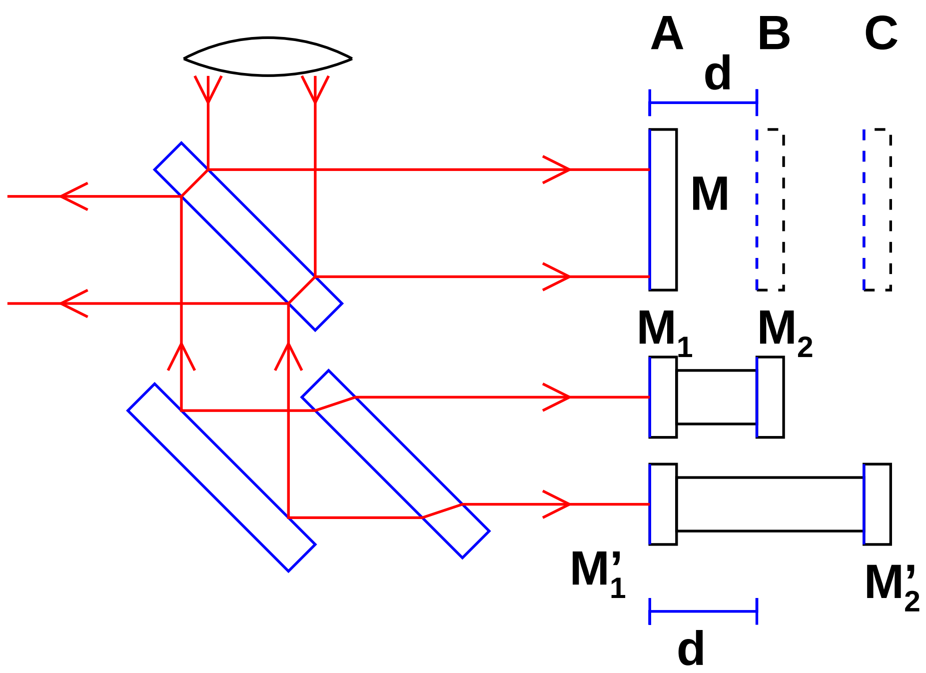
The last step uses $E = hf = h\omega/2\pi$.

4.12: Wavelength Measurements

The Michelson Interferometer was used by Michelson and Benoit to measure the wavelengths of three intense green, red and blue lines of Cadmium against the standard meter in Paris.

They measured in steps using nine etalons, each twice the length of the other. The longest was 10 cm long.





1. Make M_1 and M_1' coplanar using white light fringes
2. Using Cd light, count the number of fringes as M is moved to position B and white light fringes reappear in M_2. This gives the length of the short etalon in wavelengths.
3. Now move the shorter etalon until the white light fringes reappear in M_1.
4. Move M to position C until white light fringes reappear in $M2$
5. Using Cd light, count the number of fringes passing to make M_2 and M_2' coplanar. This gives the length of the longer etalon in terms of wavelengths.
6. Repeat until the longest etalon was moved through its length 10 times. The difference between the mark on the meter and the etalon was measured counting fringes.

This process gives the wavelengths of the Cd lines as

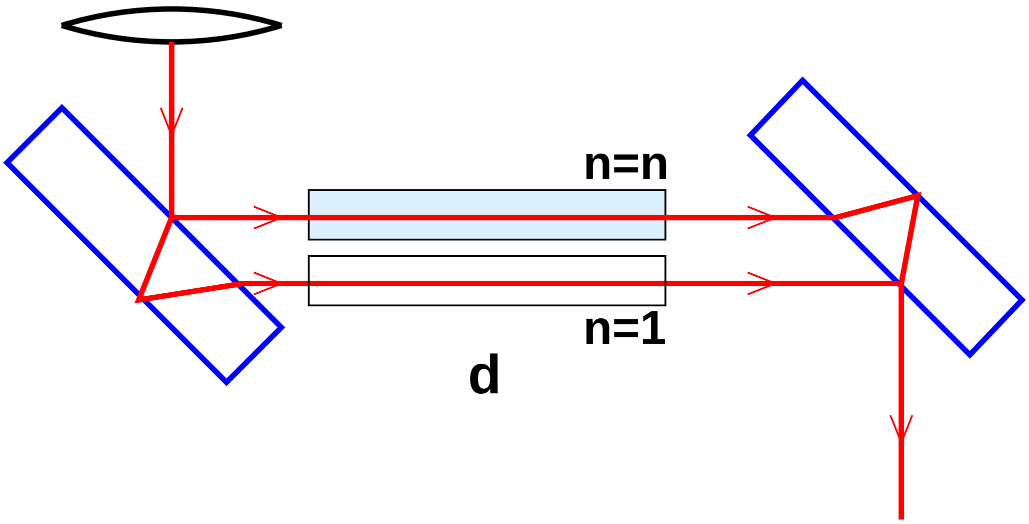
Colour	Wavelength [nm]
Red	643.84722
Green	508.58240
Blue	479.99107

4.13: Jamin Interferometer

Another useful application of an interferometer is the measurement of the refractive index of gases.

Two evacuated tubes of length d are inserted into the arms of the interferometer. One tube is slowly filled with a gas with refractive index n and the passing fringes are counted. If the Number of fringes is m, then the optical path length has changed by $m\lambda$ and we have

$$(n-1)d = m\lambda$$



Exercises

Halliday Resnick & Walker:

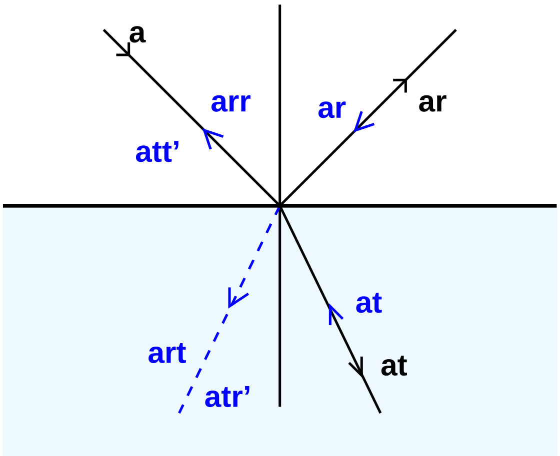
Reading: Ch36 HRW pp. 880-882

Exercises: p. 887 ff.: 54E, 55E, 57P, 58P, 60

4.14: Interference by multiple reflection

A thin film of soap or oil produces coloured stripes. The mechanism responsible is interference.

We have to take a closer look at reflections at a boundary between two media before we can understand the coloured thin film interference patterns:



Let a be the amplitude of the incident wave. The fraction of the amplitude transmitted is t and the fraction of the amplitude reflected is r.

Conservation of energy requires $t^2 + r^2 = 1$.

Now consider the time-reversed process where two waves of amplitude ar and at meet to combine. The first ray coming from above splits into a reflected ray of amplitude arr and a refracted ray art. The second one coming from below splits into a refracted ray att' and a reflected ray atr' where r' and t' denote the reflection and transmission coefficients from below.

Since the time-reversed process must be the same as the original we must have

$$arr + att' = a \qquad art + atr' = 0 \Leftrightarrow r = -r'$$

Either a phase change occurs on reflection from above or from below. Experimentally we find that the phase change occurs for the ray travelling in the medium with the faster wave speed.

This is like the hard and soft reflection of a mechanical wave.

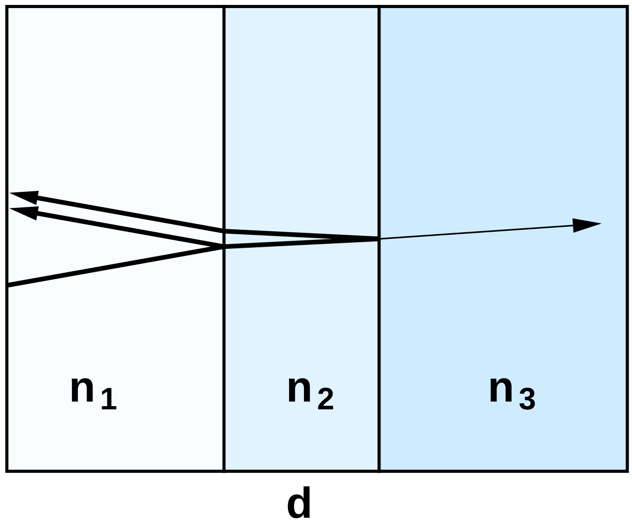
A reflection at a boundary to a medium with higher index of refraction (ray in air reflected off glass) is a hard reflection and a phase change of π occurs.

A reflection at a boundary to a medium with lower index of refraction (ray in glass reflected off air) is a soft reflection and no phase change occurs.

Reflections of silvered surfaces are soft reflections.

4.15: Thin Films

We take a film of material with refractive index n_2 between two materials with refractive indices n_1 and n_3.



A ray incident on the boundary is partially reflected and partially transmitted into the film. Another reflection happens at the back of the film and the two rays can interfere with each other.

The Phase difference due to the extra distance travelled in the film is

$$2\pi\frac{2d}{\lambda_2} = 2\pi\frac{2n_2 d}{\lambda}$$

with the wavelengths in the film λ_2 and in vacuum λ.

An additional phase difference of π can come from either reflection.

If both reflections are hard ($n_1 < n_2 < n_3$) or soft ($n_1 > n_2 > n_3$) there is no additional phase change to consider and the condition for constructive interference becomes

$$2n_2 d = m\lambda$$

If one of the two reflections is hard ($n_1 < n_2$ and $n_1 < n_3$) or ($n_1 > n_2$ and $n_1 > n_3$) then we have an additional phase of π and the condition for constructive interference becomes

$$2n_2 d = (m + \tfrac{1}{2})\lambda$$

The transmitted light must also show interference effects. Energy is conserved, so if a lot of light is reflected, very little can be transmitted and vice versa. If the condition for an interference minimum in the reflected light is fulfilled we get an interference maximum in the transmitted light.

Similarly if the condition for an interference maximum in the reflected light is fulfilled we get an interference minimum in the transmitted light.

Thin Films Checklist

	Transmitted Light		Reflected Light	
	Maximum	Minimum	Maximum	Minimum
	$2n_2d =$	$2n_2d =$	$2n_2d =$	$2n_2d =$
$n_1, n_3 < n_2$ $n_1, n_3 > n_2$	$m\lambda$	$(m + \frac{1}{2})\lambda$	$(m + \frac{1}{2})\lambda$	$m\lambda$
$(n_1 < n_2 < n_3)$ $(n_1 > n_2 > n_3)$	$(m + \frac{1}{2})\lambda$	$m\lambda$	$m\lambda$	$(m + \frac{1}{2})\lambda$

In all cases λ is the wavelength of light in vacuum

4.16: Nonreflective coatings

An important application of thin film interference is a nonreflective coating.

Consider a ray incident on a thin film with refractive index n_2 deposited on a piece of glass with refractive index $n_3 > n_2$. We assume that the ray strikes the surface almost perpendicular. Since now both rays have one hard reflection in them, the condition for an interference minimum in the reflected light is

$$2n_2 d = (m + \tfrac{1}{2})\lambda$$

For $m = 0$ that gives a condition

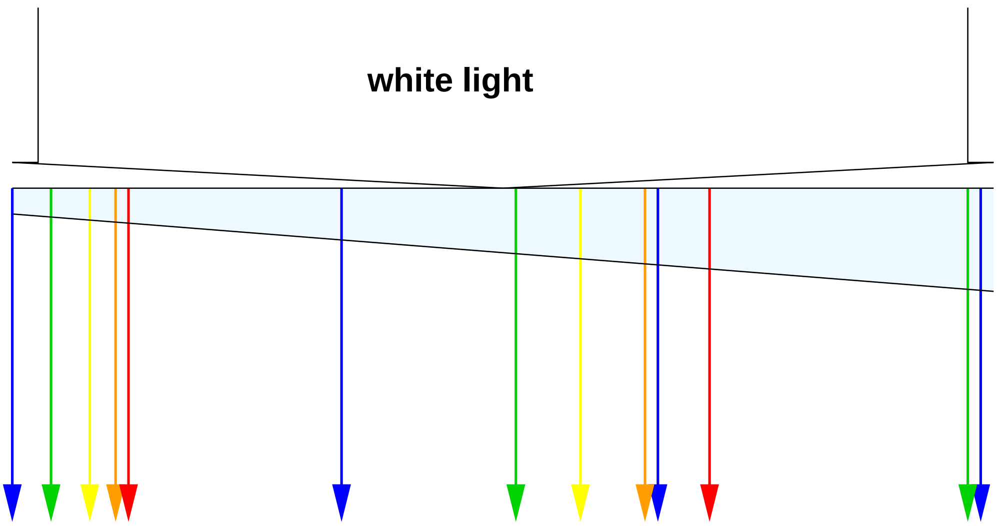
$$d = \frac{\lambda}{4n_2}$$

For orange light ($\lambda = 600\,\mathrm{nm}$) and $n_2 = 1.5$ we get a thickness $d = 100\,\mathrm{nm}$

A coated lens has a purple hue. The condition $d = \frac{\lambda}{4n_2}$ can only be fulfilled exactly for one wavelength, usually chosen near the middle of the spectrum. The outer ends (red and blue) then are reflected more strongly and the coating looks purple.

4.17: Soap films

Now we can understand the colours of an oil film: The thickness of the film is of the order of the wavelength of light and the maximum reflection (or transmission) condition is fulfilled for each wavelength in turn:



The film changes thickness between two green fringes by: $\Delta d = \frac{1}{2n} 550\,\text{nm} = 225/1.33\,\text{nm} = 170\,\text{nm}$.

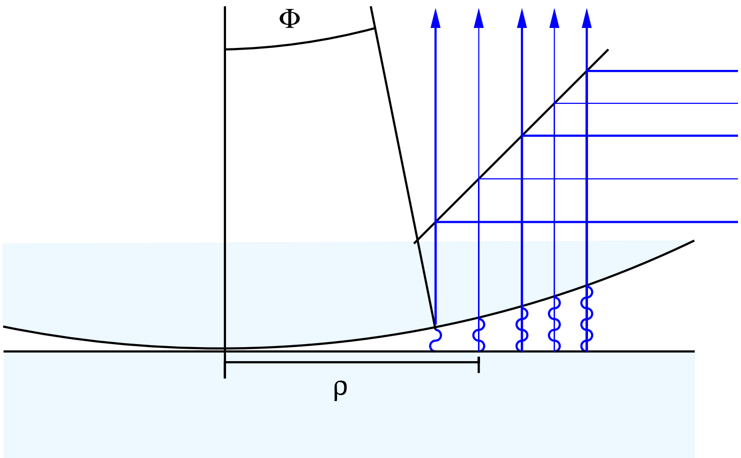
The fringes of the same colour are repeating with different rates. This gives rise not only to pure rainbow colours, but to mixed colours like pink or brown as well.

When the thickness of the film is less than 150 nm no bright fringe can be formed anymore.

4.18: Newton's Rings

If the thickness of the film changes continually, the condition for a bright reflected fringe will be met periodically. This can be used to measure the radius of curvature R for lenses.

Here the film is made of air ($n = 1$) contained between two glass surfaces. The condition for a bright fringe in the reflected light is $2d = (m + \frac{1}{2})\lambda$. The thickness of the film changes by $\lambda/2$ between bright fringes.



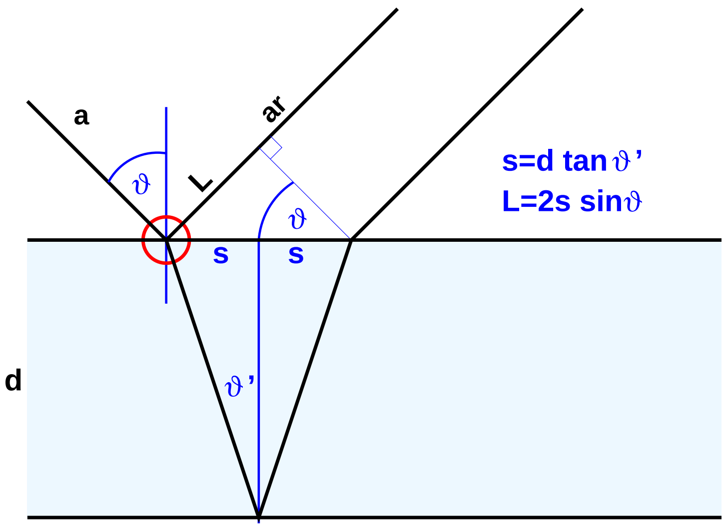
The thickness of the film is $d = R(1 - \cos\phi) \simeq \frac{1}{2}R\phi^2$ and we see a circular bright fringe of radius $\rho = R\tan\phi \simeq R\phi$ if $R\phi^2 = (m + \frac{1}{2})\lambda$

Plotting the square of the radius of the m^{th} fringe against m gives a rather precise measurement of ρ.

$$\rho^2 = (m + \tfrac{1}{2})\lambda \cdot R$$

4.19: Thin film - again

Take a closer look at a thin plane parallel film illuminated from above:



Along the length $L = 2d \tan\vartheta' \sin\vartheta$ ray one changes phase by

$$\Delta\Phi_1 = \pi + 2\pi\frac{L}{\lambda} = \pi + 2\pi\frac{2d\tan\vartheta' \sin\vartheta}{\lambda} = \pi + 2\pi\frac{2nd\sin^2\vartheta'}{\lambda\cos\vartheta'}$$

Ray 2 changes phase by

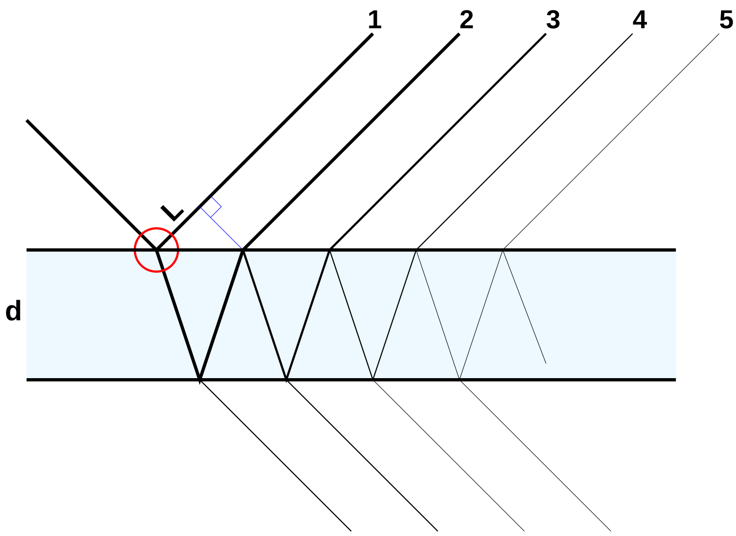
$$\Delta\Phi_2 = 2\pi\frac{2d}{\cos\vartheta'\lambda'} = 2\pi\frac{2nd}{\cos\vartheta'\lambda}$$

The total Phase difference is

$$\Delta\Phi = \pi + 2\pi\frac{2nd}{\lambda}\left(\frac{\sin^2\vartheta' - 1}{\cos\vartheta'}\right) = \pi + 2\pi\frac{2nd}{\lambda}\cos\vartheta'$$

For a nearly vertical angle of incidence ($\cos\vartheta' = 1$) this reduces to

$$\Delta\Phi = \pi + 2\pi\frac{2nd}{\lambda}$$



Only ray 1 contains a phase inverting hard reflection, rays 2,3,... have only internal (soft) reflections.

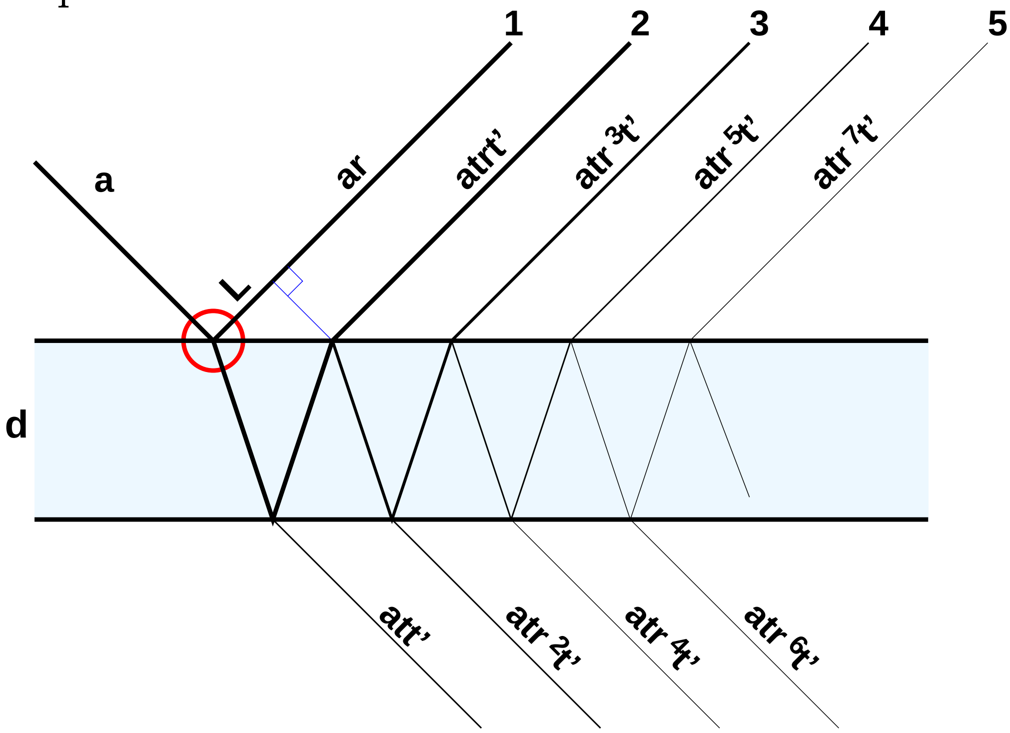
The phase difference between any two neighboring rays except the first one is therefore

$$\Delta\Phi = 2\pi\frac{2nd}{\lambda}\cos\vartheta'$$

Look at an angle where rays 1 and 2 interfere destructively. Then rays 2,3,4... must interfere constructively.

To finally decide if we see a bright or dark reflection we must look at the amplitudes.

If a is the amplitude of the incident ray then ray 1 has an amplitude ar, ray 2 has an amplitude $atrt'$, ray 3 has an amplitude atr^3t' etc.



Summing all amplitudes from ray 2 on gives

$$\begin{aligned} A &= atrt' + atr^3t' + atr^5t' + atr^7t' + \cdots \\ &= atrt'(1 + r^2 + r^4 + r^6 + \cdots) \\ &= atrt'\frac{1}{1-r^2} \\ &= ar \end{aligned}$$

The last step uses $tt' = 1 - r^2$.

So the first ray has an amplitude of ar and ALL other rays have a combined amplitude ar and they interfere destructively: We get an interference minimum in the reflected light if $2nd\cos\vartheta' = m\lambda$ like we had before.

4.20: Summary

- Reflections of light at a refracting boundary are **hard** if the "mirror" has a **larger** index of refraction. These reflections change the phase of the reflected light by π.
- All other reflections including those off silvered surfaces are **soft** reflections without a phase change.
- Interference of light from a thin film will produce bright fringes in the reflected and transmitted light.
- A nonreflective coating is made of a film with a thickness $d = \lambda/4n$ to give a maximum in the transmitted light (=minimum in reflected light).

Exercises

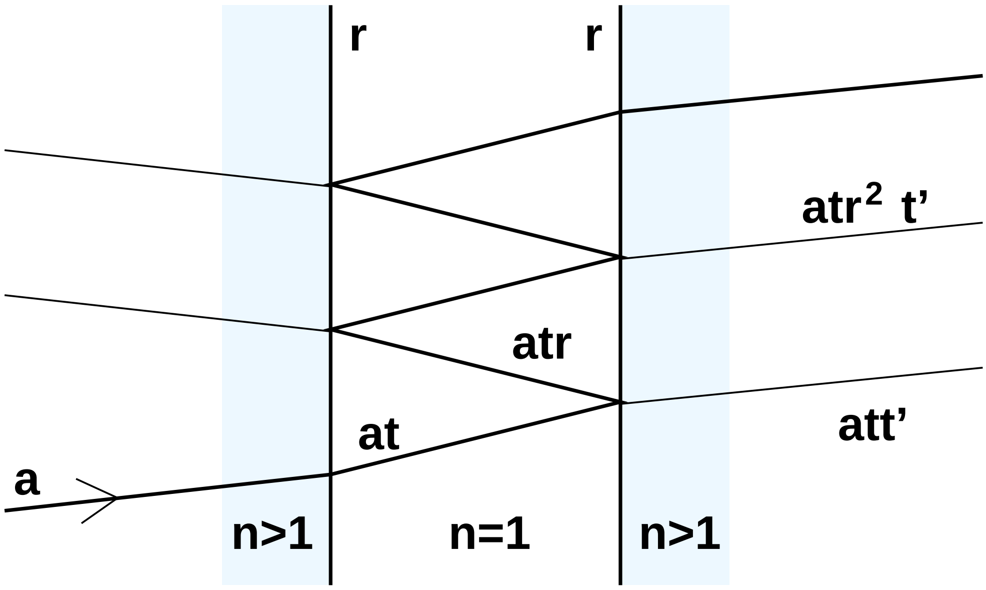
Halliday Resnick & Walker:

Reading: Ch36 HRW pp. 874-880

Exercises: p. 885 ff: 30E, 31E, 32E, 34E, 36E, 38P, 40P, 43P, 49P, 51P.

4.21: Fabry-Perot Interferometer

The sharpness of fringes in transmitted light through a thin film of air can be used to build highly sensitive spectrometer and narrow colour filters.

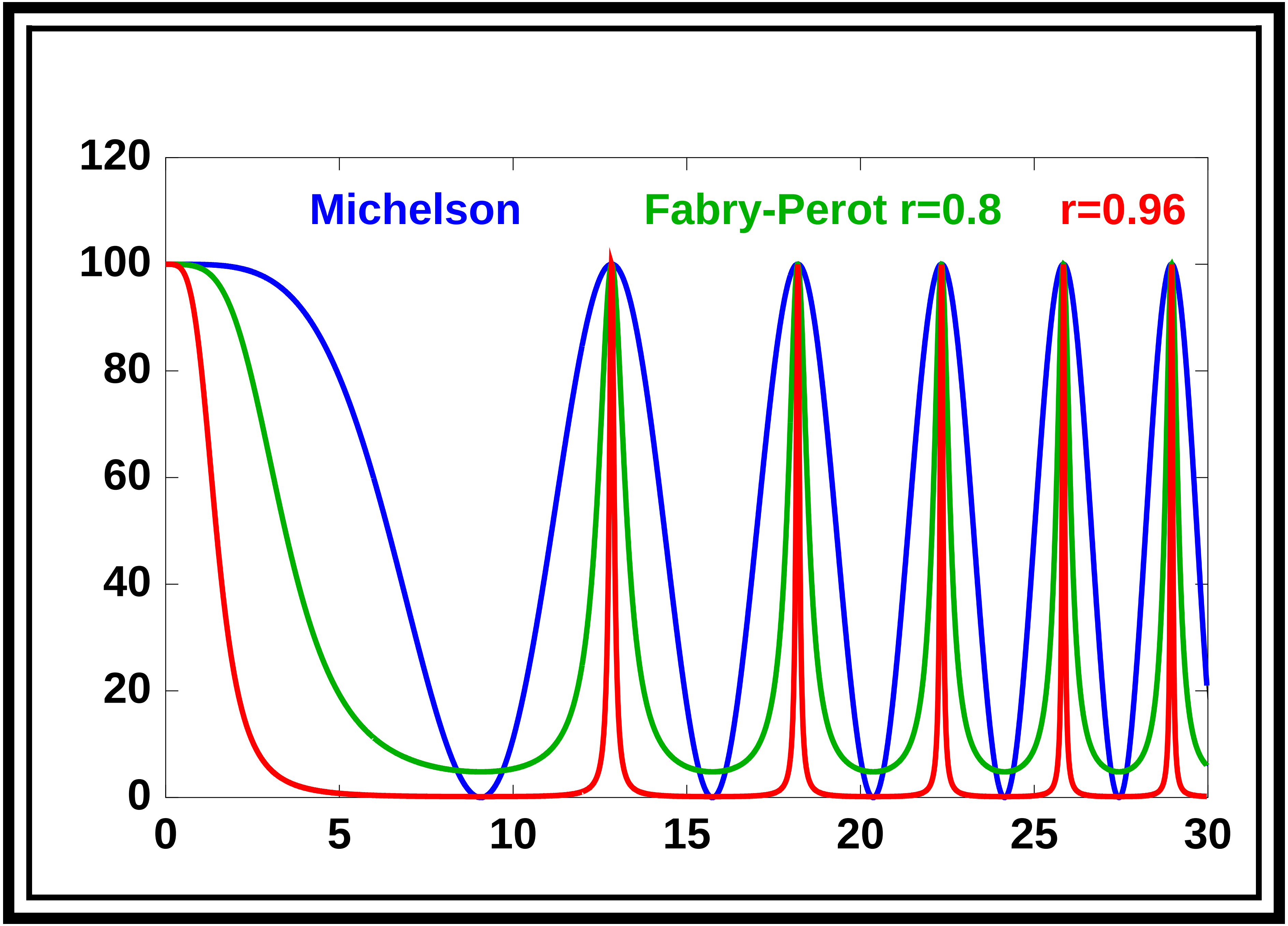


The inner surfaces are silvered, therefore we have only soft reflections and the condition to get a bright fringe in the reflected light becomes

$$2d\cos\vartheta = (m + \tfrac{1}{2})\lambda$$

If the interferometer is illuminated with an extended light source we can observe similar circular fringes as in the Michelson interferometer. If the reflectivity of the surfaces is high enough, the fringes will be very sharp.

PHYS126 Waves and Optics

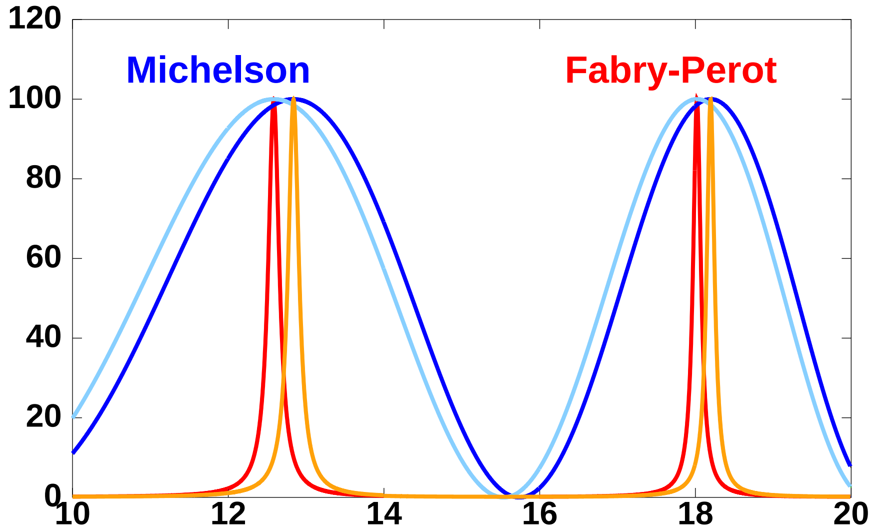


R-D Herzberg

4.22: Resolving power

If we want to analyse the spectrum of light emitted from a source (e.g. an excited atom) we must be able to decompose the light into its wavelength components. A prism will do that, but it is a much too coarse instrument for the fine analysis of a spectrum.

Let an atom emit light of two wavelengths λ_1 and $\lambda_2 = \lambda_1 + \Delta\lambda$.



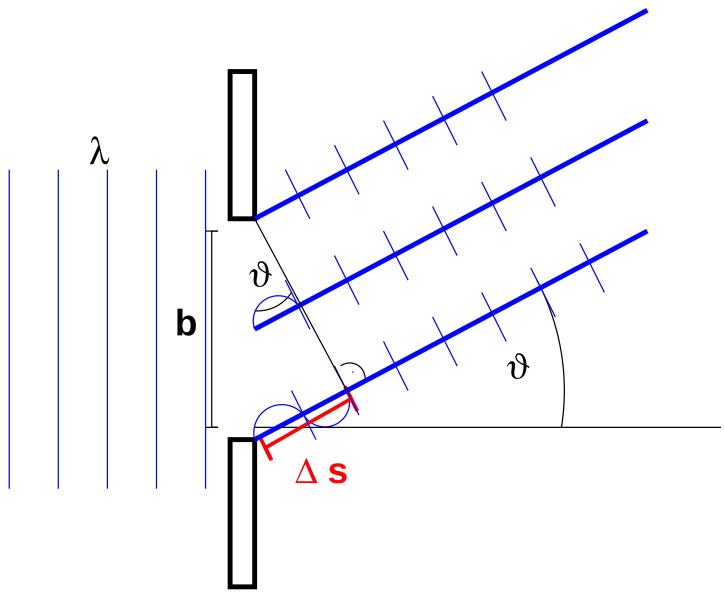
These are for wavelengths 500 nm and 501 nm respectively.

A Fabry-Perot interferometer with a reflectivity of $r = 0.8$ has a tenfold higher resolving power than a Michelson interferometer under the same conditions.

4.23: Single Slit Diffraction

We can now compute the diffraction pattern from a single slit. We will proceed in two stages: First we will obtain the angles at which we observe interference maxima and minima in a purely geometrical way.

Secondly we will compute the complete intensity pattern as a function of the angle of diffraction.



We position the screen sufficiently far away from the slit so we can deem all rays to be parallel (Fraunhofer diffraction).

Let ϑ_1 be the angle for which the top and bottom rays from the edges of the slit have a path difference of exactly λ. Then the top ray and the ray in the center of the slit will have a path difference of $\frac{1}{2}\lambda$ and will cancel.

Each ray in the top half of the slit will find exactly one ray in the bottom half to cancel. The total effect at ϑ_1 is a minimum. We find $b \sin \vartheta_1 = \lambda$

The second minimum is found when the path difference is 2λ: Here you divide the slit into 4 equal parts and each pair of rays in the first and second as well as in the third and fourth part will cancel.

Generally a minimum is found if the path difference between the top and bottom part of the slit is a full wavelength:

$$b \sin \vartheta = m\lambda \qquad \text{Minimum}$$

Straight behind the slit the central (principal) maximum will be located.

To obtain a secondary maximum you can divide the slit into 3 parts such that the path difference between the top and bottom rays is $3/2\lambda$. Then each pair of rays from the first and second partition cancel, leaving those from the third.

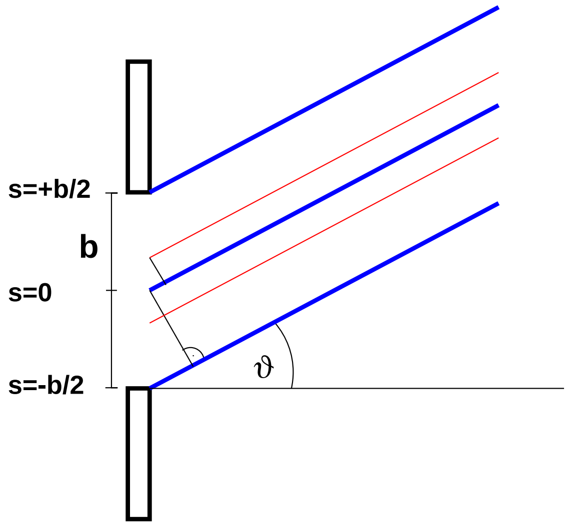
Generally a maximum is found if the path difference between the top and bottom part of the slit is a full wavelength and a half:

$$b \sin \vartheta = \tfrac{1}{2}(2m+1)\lambda \qquad \text{Maximum}$$

4.24: Intensity

The intensity can be computed easily if one divides the slit into narrow parts of width Δs. All rays then have an amplitude $y_0 \Delta s/b$.

The ray emerging from the center of the slit is chosen as the reference ray. The phase difference between a ray emitted at position $s = s$ and a ray emitted at position $s = 0$ is $\delta = 2\pi \frac{s \sin\vartheta}{\lambda} = ks\sin\vartheta$



The sum of two rays emitted from positions $+s$ and $-s$ is

$$dy = dy_{+s} + dy_{-s} = \frac{y_0 \Delta s}{b}(\sin(kx - \omega t - \delta) + \sin(kx - \omega t + \delta))$$

We use $\sin\alpha + \sin\beta = 2\sin(\frac{1}{2}(\alpha+\beta))\cos(\frac{1}{2}(\alpha-\beta))$ and find

$$dy = \frac{2y_0 \Delta s}{b}\sin(kx - \omega t)\cos(\delta)$$

$$dy = \frac{2y_0 \Delta s}{b} \sin(kx - \omega t) \cos(ks \sin \vartheta)$$

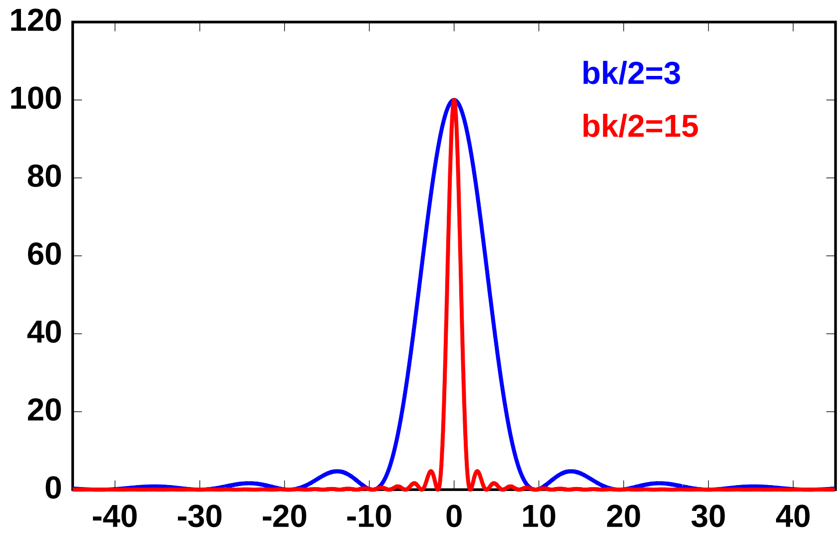
We now must sum up all the contributions from each pair:

$$\begin{aligned} y &= \frac{2y_0 \sin(kx - \omega t)}{b} \int_0^{b/2} \cos(ks \sin \vartheta) \mathrm{d}\, s \\ &= \frac{2y_0}{b} \left[\frac{\sin(ks \sin \vartheta)}{k \sin \vartheta} \right]_0^{b/2} \sin(kx - \omega t) \\ &= y_0 \left(\frac{\sin(\frac{1}{2} kb \sin \vartheta)}{\frac{1}{2} kb \sin \vartheta} \right) \sin(kx - \omega t) \end{aligned}$$

The amplitude can be rewritten as $A_0 \frac{\sin \alpha}{\alpha}$ with $\alpha = \frac{1}{2} kb \sin \vartheta$.

The intensity pattern then is

$$I = I_0 \frac{\sin^2 \alpha}{\alpha^2}$$



4.25: Examples

A slit of width $b = 40\,\mu$m is illuminated with blue light $\lambda = 400\,$nm. What is the distance between the central maximum and the second minimum (the fourth interference maxima on either side) in the interference pattern on a screen 2 m away?

Interference minimum: $b \sin\vartheta = m\lambda$ The first minimum is obtained for $m = 1$, the second minimum is obtained for $m = 2$.

$$\sin\vartheta = 2\lambda/b = 2\frac{400\text{nm}}{40\mu\text{m}} = 0.02$$

The distance on the screen is
$d = 200\text{cm}\tan(\sin^{-1} 0.02) = 4\,\text{cm}$

The condition for a maximum is $b \sin\alpha = \frac{1}{2}(2m+1)\lambda$. The first secondary maximum is obtained for $m = 1$ and the fourth secondary maximum is at $m = 4$.

$$\sin\vartheta = 4.5\lambda/b = 4.5\frac{400\text{nm}}{40\mu\text{m}} = 0.045$$

The distance between the fourth maxima is
$d = 2 \cdot 200\text{cm}\tan(\sin^{-1} 0.045) = 18\,\text{cm}.$

What is the intensity ratio between the third and seventh maximum?

$$I = I_0 \frac{\sin^2 \alpha}{\alpha^2} \qquad \alpha = \tfrac{1}{2} kb \sin\vartheta$$

The condition for a maximum is then

$$\alpha_m = \tfrac{1}{2} kb \sin\vartheta = \tfrac{1}{2} k \tfrac{1}{2}(2m+1)\lambda = \tfrac{1}{2}(2m+1)\pi$$

The third maximum is found at $\alpha_3 = 3.5\pi$, the seventh maxiumum at $\alpha_7 = 7.5\pi$

$$\frac{I_3}{I_7} = \frac{I_0 \frac{\sin^2 3.5\pi}{(3.5\pi)^2}}{I_0 \frac{\sin^2 7.5\pi}{(7.5\pi)^2}} = \frac{7.5^2}{3.5^2} = 4.6$$

4.26: Resolving power

By the resolving power of an optical instrument we mean its ability to produce two separate images of two objects very close together.

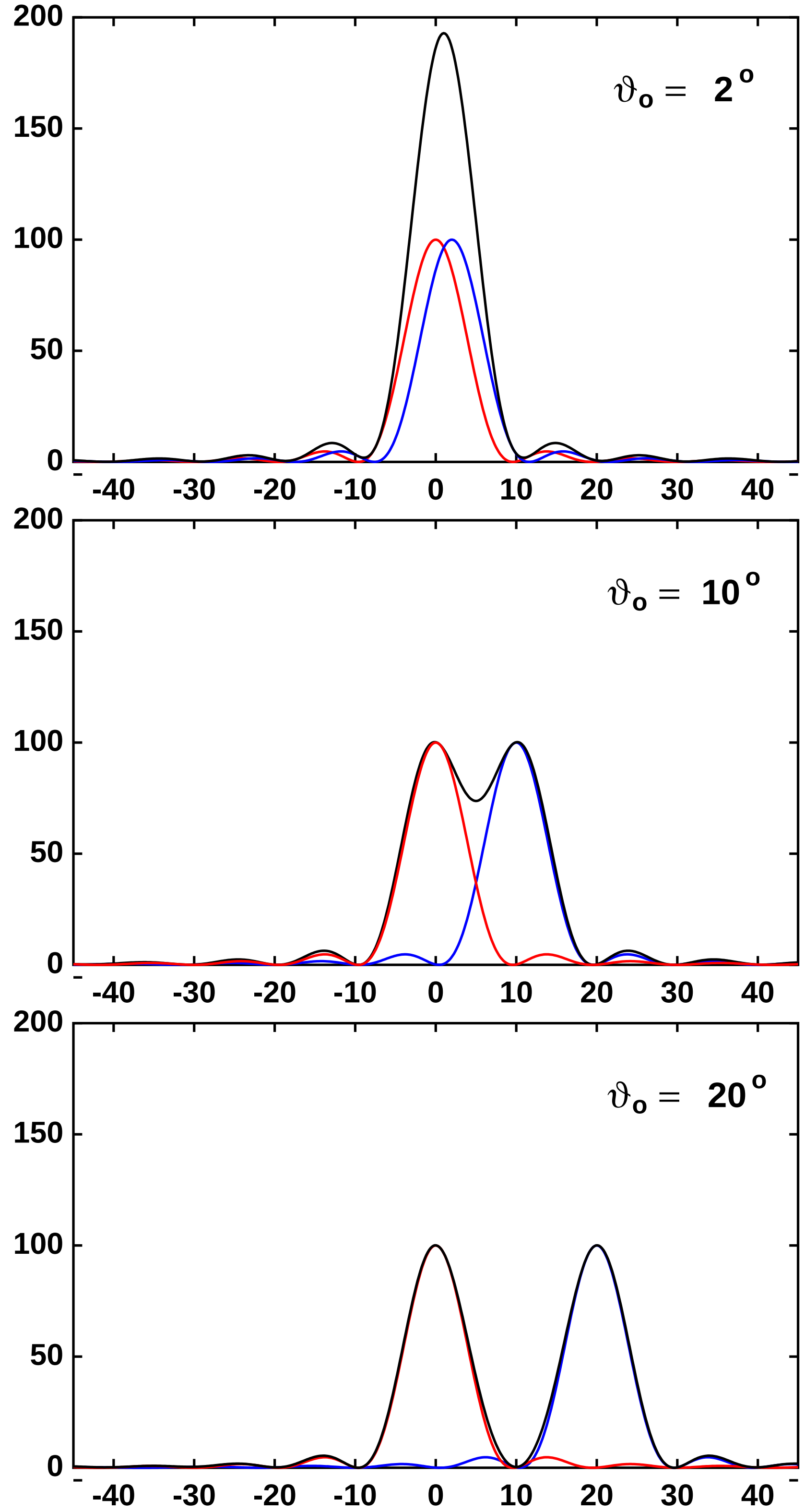
A slit of width b produces a central maximum with a width inversely proportional to b. If the two images are much closer than the width of either central image, they can clearly not be seen as separate images.

If the principal maximum of the second image falls into the first minimum of the diffraction pattern of the first image, we can just barely see them as two separate maxima.

The resolving power is therefore defined as as the minimum angle of resolution ϑ_0. For a slit it is given by

$$\vartheta_0 = \frac{\lambda}{b}$$

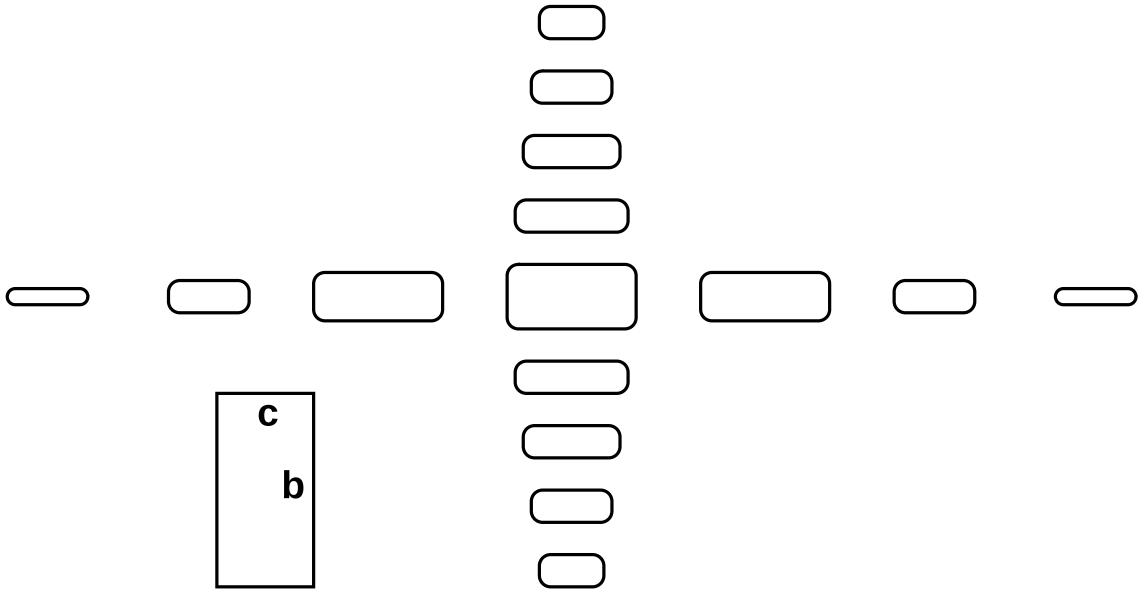
Note: As ϑ_0 increases, the resolving power decreases!


200
150
100
50
0
-40 -30 -20 -10 0 10 20 30 40
$\vartheta_o = 2^o$
$\vartheta_o = 10^o$
$\vartheta_o = 20^o$

4.27: Diffraction from Extended Apertures

In general the diffraction pattern will involve rays from every point on the aperture. Two special cases are worth mentioning: A rectangular aperture and a circular aperture.

The rectangular aperture can be seen as a combination of a vertical slit and a horizontal slit.



The circular aperture is important because almost all lenses are circular, and the rims of the lenses or the tube of a telescope or a microscope is a circular opening that ultimately limits the resolving power of these instruments.

4.28: Rayleigh's criterion

Unfortunately the exact treatment of a circular aperture is rather involved and does not teach us anything new. One finds that the first minimum is found at a slightly different angle:

$$\vartheta_0 = 1.22\frac{\lambda}{b}$$

Compared with a slit the only difference is the factor 1.22.

We can now formulate **Rayleigh's criterion** for the resolution of two images:

> Two images can be resolved if the diffraction maximum of one image coincides with the first diffraction minimum of the second image. For a circular aperture of diameter d this means that the angular separation must be
>
> $$\vartheta_0 = 1.22\frac{\lambda}{d}$$

4.29: Summary

- Rayleigh's Criterion: Two images can be resolved if the diffraction maximum of one image coincides with the first diffraction minimum of the second image. For a circular aperture of diameter d this means that the angular separation of the two images must be at least $\vartheta_0 = 1.22\lambda/d$.
- A single slit of width b will produce an interference pattern of the form
$$I(\vartheta) = I_0 \frac{\sin^2 \alpha}{\alpha^2} \qquad \text{with} \quad \alpha = \tfrac{1}{2} kb \sin\vartheta$$
- The secondary minima are found at angles ϑ such that $b \sin\vartheta = m\lambda$.
- The secondary maxima are found at angles ϑ such that $b \sin\vartheta = (m + \frac{1}{2})\lambda$.
- The resolving power of a Fabry-Perot interferometer is highest for high reflectivities.

Exercises

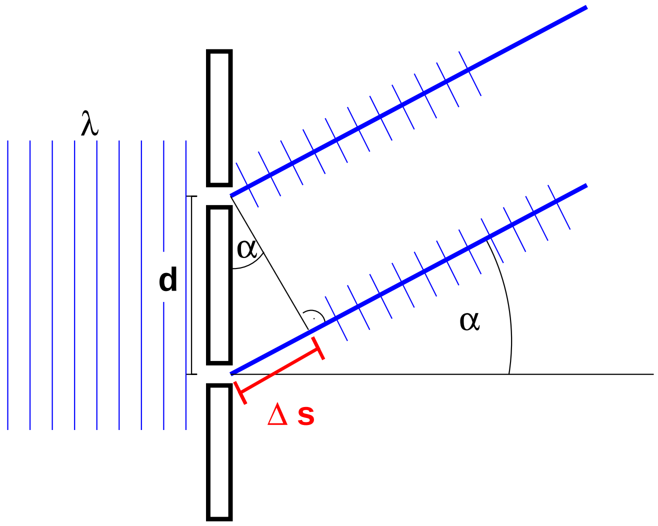
Halliday Resnick & Walker:

Reading: Ch37 HRW pp. 891-900

Exercises: p. 912 ff.: Q1, Q4, Q6, 1E, 2E, 3E, 6P,10E, 15E, 17E, 25P

4.30: Double Slit Diffraction

We have already discussed Young's double slit experiment and found the conditions for interference maxima and minima. Now we can look at the complete intensity pattern produced on a screen behind the slits



At an angle ϑ the path difference is $\Delta s = d \sin\vartheta$ and the phase difference is

$$\Delta\phi = 2\pi \frac{d \sin\vartheta}{\lambda} = kd \sin\vartheta$$

On the screen the two waves will therefore have amplitudes

$$\begin{aligned} y_1(x,t) &= a \sin(kx - \omega t) \\ y_2(x,t) &= a \sin(kx - \omega t - kd \sin\vartheta) \end{aligned}$$

We know how to deal with this sum:

$$\begin{aligned} y(x,t) &= y_1(x,t) + y_2(x,t) \\ &= 2a\cos(\tfrac{1}{2}kd\sin\vartheta)\sin(kx - \omega t - \tfrac{1}{2}kd\sin\vartheta) \end{aligned}$$

The intensity is the square of the amplitude:

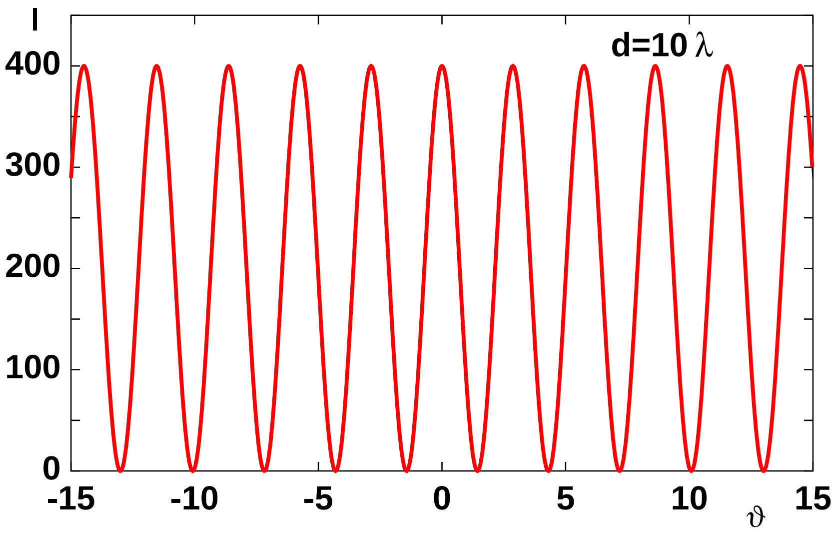
$$I = 4a^2\cos^2(\tfrac{1}{2}kd\sin\vartheta) = 4I_0\cos^2(\tfrac{1}{2}kd\sin\vartheta)$$

Or, if we substitute $\beta = \frac{1}{2}kd\sin\vartheta$, we have

$$I = 4I_0\cos^2(\beta)$$

The conditions for a maxima/minima of order m are:

Maximum:	$d\sin\vartheta = m\lambda$
Minimum:	$d\sin\vartheta = (m + \frac{1}{2})\lambda$



4.31: Examples

Two narrow slits are a distance $d = 40\,\mu$m apart. They are illuminated with light of two different wavelengths: $\lambda_1 = 600\,$nm and $\lambda_2 = 550\,$nm. A screen is 300 cm behind the slits.

What is the distance of the third maximum from the center on the screen in each case?

Condition for a maximum with narrow slits:

$$d \sin\vartheta = m\lambda$$

$$\sin\vartheta_1 = 3\frac{600\text{nm}}{40\mu\text{m}} = 3 \cdot \frac{0.6\mu\text{m}}{40\mu\text{m}} = 0.045$$

$$s_1 = L\tan\vartheta_1 = 300\text{cm} \cdot 0.045 = 13.5\text{cm}$$

$$\sin\vartheta_2 = 3\frac{550\text{nm}}{40\mu\text{m}} = 3 \cdot \frac{0.55\mu\text{m}}{40\mu\text{m}} = 0.041$$

$$s_2 = L\tan\vartheta_2 = 300\text{cm} \cdot 0.041 = 12.4\text{cm}$$

We use $\sin\vartheta \simeq \tan\vartheta \simeq \vartheta$ valid for angles less than 0.1 rad. ($\simeq 6°$)

Red light with a wavelength $\lambda = 638\,\text{nm}$ illuminates a double slit. On a screen 200 cm behind the slits the distance between the central and the second order maximum is found to be 7.8 cm. What is the distance between the two slits?

$$d \sin\vartheta = m\lambda$$

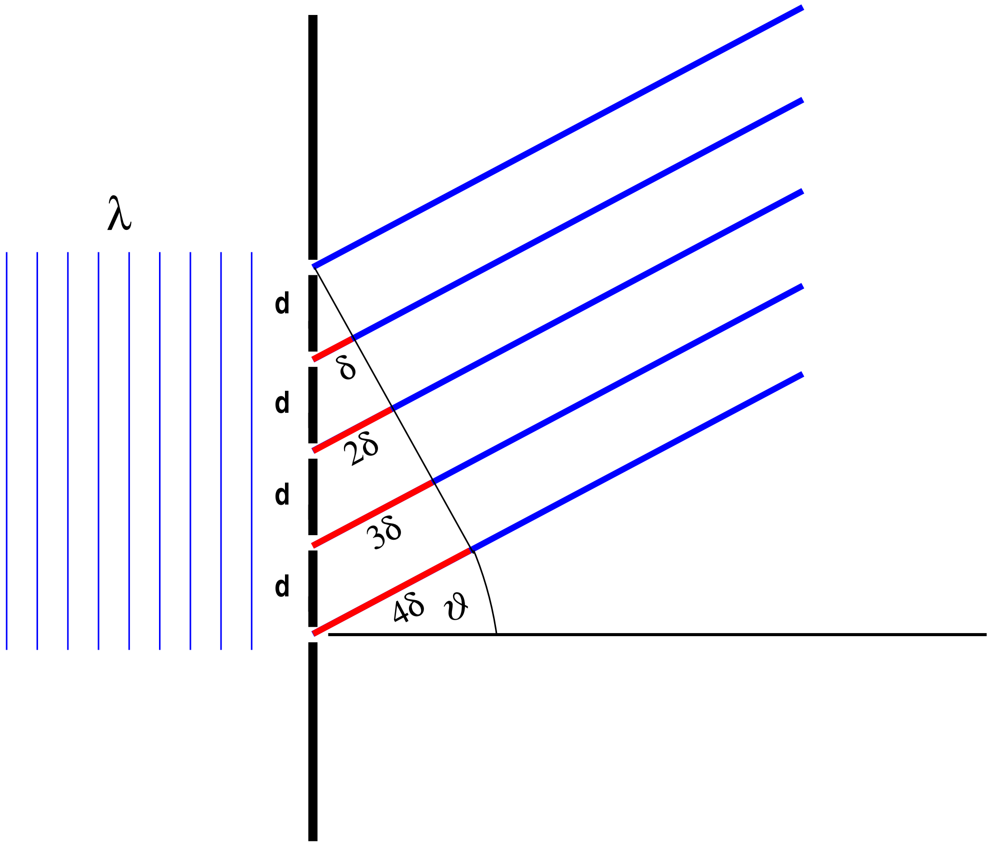
$$\tan\vartheta = \frac{7.8\text{cm}}{200\text{cm}} = 0.039$$

$$\vartheta = 2.23^\circ$$

$$d = \frac{2 \cdot 638\text{nm}}{\sin 2.23^\circ} = 32700\,\text{nm} = 32.7\mu\text{m}$$

4.32: Diffraction Grating

What happens if we take more than two very narrow slits with equal spacings d between each slit?



If we take the ray from the topmost slit as the reference ray then the ray from the N^{th} slit down will have a phase difference of

$$N\delta = 2\pi N d \sin\vartheta/\lambda = kNd\sin\vartheta$$

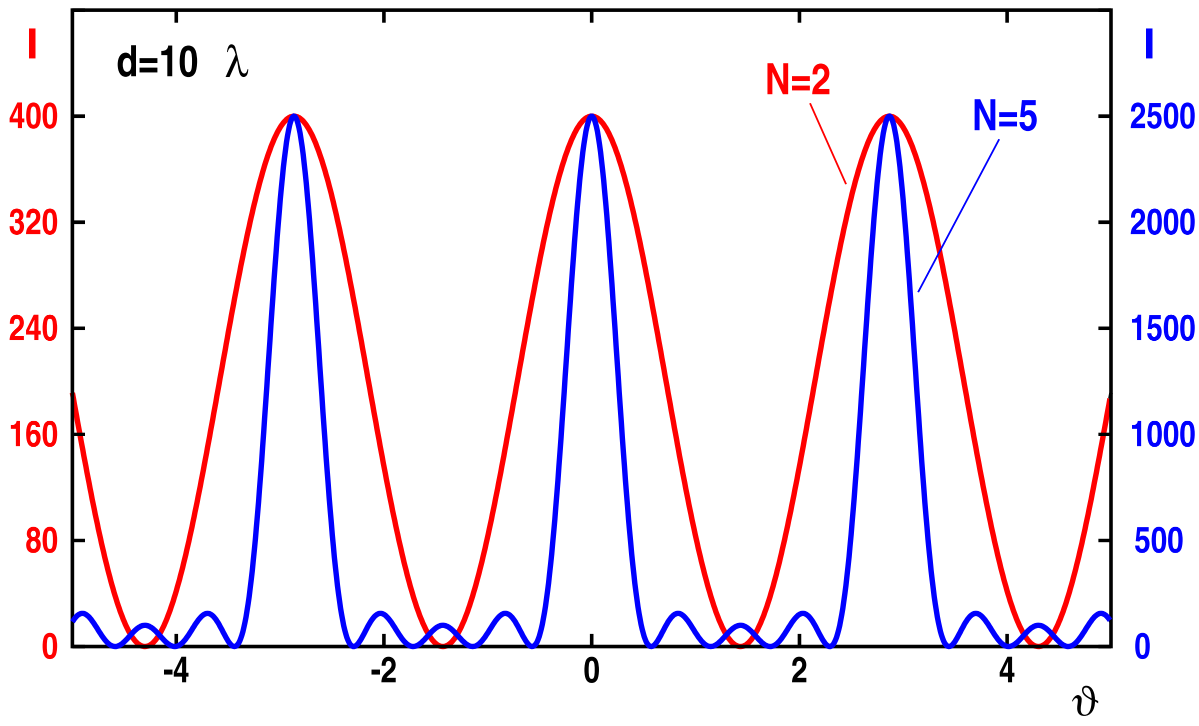
The total wave at that angle then is obtained by summing the waves from all slits:

$$\begin{aligned} A(x,t) &= a[\cos(kx-\omega t)+\cos(kx-\omega t-\delta) \\ &\quad +\cos(kx-\omega t-2\delta)+\cdots \\ &\quad +\cos(kx-\omega t-(N-1)\delta)] \\ &= a\cos(kx-\omega t-\tfrac{1}{2}(N-1)\delta)\frac{\sin\frac{1}{2}N\delta}{\sin\frac{1}{2}\delta} \end{aligned}$$

And its intensity is

$$I = I_0\frac{\sin^2\frac{1}{2}N\delta}{\sin^2\frac{1}{2}\delta} \qquad \delta = kd\sin\vartheta$$

The maximum intensity in each order is $I_{max} = I_0 N^2$!



4.33: Resolving Power

In order to use the grating as a spectrometer we must be able to distinguish two wavelengths λ and $\lambda + \Delta\lambda$.

According to Rayleigh's criterion we can distinguish two lines, if the main maximum of one line falls into the first minimum of the second line. Clearly, if the maximum is high and narrow, we will be able to resolve lines differing by only a small wavelength difference and the **resolving power** is large.

We define the resolving power

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

We can find the position of the minima closest to the main peaks easily by looking at the intensity equation:

$$I = I_0 \frac{\sin^2 \frac{1}{2} N\delta}{\sin^2 \frac{1}{2}\delta} \qquad\qquad \delta = kd \sin\vartheta$$

We denote the angle for the main maximum in m^{th} order with ϑ_m and the angle for the immediately adjacent minimum with ϕ_m.

The main maxima are found at those angles where the denominator vanishes:

$$\tfrac{1}{2}kd\sin\vartheta_m = m\pi \quad \Leftrightarrow \quad d\sin\vartheta_m = m\lambda$$

In the numerator this becomes

$$Nd\sin\vartheta_m = Nm\lambda$$

Next we need to find the angles at which the next minimum is located, that is, where the numerator vanishes next.

$$Nd\sin\phi_m = (Nm+1)\lambda$$

We can now write the criterion for resolution of the wavelengths λ and $\lambda + \Delta\lambda$: The maximum for $(\lambda + \Delta\lambda)$ must coincide with the minimum for λ.

$$Nm(\lambda + \Delta\lambda) = (Nm+1)\lambda$$

$$Nm\lambda + Nm\Delta\lambda = Nm\lambda + \lambda$$

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

A grating with many lines has a higher resolving power than one with few lines **independent** of the spacing of the lines!

The resolving power is higher in the higher orders.

4.34: Dispersion

We also want to be able to measure the angles reasonably comfortably. This means that they must be fairly large. The grating must have a large **dispersion**.

$$D = \frac{\mathrm{d}\,\vartheta}{\mathrm{d}\,\lambda}$$

Again we can easily derive an expression for D. The principal maximum for the m^{th} order is found at an angle:

$$d \sin\vartheta_m = m\lambda$$

$$\frac{\mathrm{d}\,\lambda}{\mathrm{d}\,\vartheta} = \frac{d\cos\vartheta}{m}$$

Or

$$D = \frac{\mathrm{d}\,\vartheta}{\mathrm{d}\,\lambda} = \frac{m}{d\cos\vartheta}$$

The dispersion is proportional to the order and inversely proportional to the spacing but is **independent** of the number of lines in the grating or the wavelength of the light!

The dispersion is inversely proportional to the spacing.

Resolving power R and dispersion D are independent of each other, and a grating is characterized through both.

Two red lasers with wavelengths $\lambda_1 =?$ and $\lambda_2 =?$ shine through a grating with 300 lines/mm onto a screen

$L =$ ______ cm behind the grating. $d = 3333\,\text{nm}$.

The distance between the principal and the first order maximum on the screen is

$s_1 =$ ______ cm

$s_2 =$ ______ cm

What are the wavelengths of the lasers?

$$\tan\vartheta = \frac{s}{L}$$

$\vartheta_1 =$ ______ $\quad \sin\vartheta_1 =$ ______

$\vartheta_2 =$ ______ $\quad \sin\vartheta_1 =$ ______

$\lambda_1 = d\sin\vartheta_1 = 3333\text{nm}\cdot$ ______ $=$ ______ nm

$\lambda_2 = d\sin\vartheta_2 = 3333\text{nm}\cdot$ ______ $=$ ______ nm

4.35: Examples

A grating is 1 mm wide and has 80 lines. What is the dispersion and resolving power for red light ($\lambda = 700\,\mathrm{nm}$) in 3^{nd} order?

$$d = 1\mathrm{mm}/80 = 12500\mathrm{nm}$$

$$D = \frac{m}{d\cos\vartheta}$$

$$d\sin\vartheta = 3\lambda \quad \Rightarrow \vartheta = \sin^{-1}\left(3\frac{700\mathrm{nm}}{12500\mathrm{nm}}\right) = 0.169\mathrm{rad} = 9.7^\circ$$

$$D = \frac{3}{12500\mathrm{nm}\cos(9.7^\circ)} = 0.00024\frac{\mathrm{rad}}{\mathrm{nm}} = 0.014\frac{\mathrm{degrees}}{\mathrm{nm}}$$

$$R = \frac{\lambda}{\Delta\lambda} = Nm = 3\cdot 80 = 240$$

This means we can resolve wavelength differences as small as $\Delta\lambda = \lambda/240$.

The grating is illuminated with yellow sodium light $\lambda_1 = 589.00$, $\lambda_2 = 589.59$. How many lines must the grating have to allow the two lines to be resolved in second order?

$$N = \frac{\lambda}{m\Delta\lambda} = \frac{589\mathrm{nm}}{2\cdot 0.59\mathrm{nm}} = 499$$

A grating with 500 lines will resolve the two sodium lines in second order.

4.36: Summary

- The intensity distribution far behind two narrow slits is

$$I(\vartheta) = I_0 \cos^2 \beta \qquad \beta = \tfrac{1}{2} kd \sin \vartheta$$

- The maxima are found at angles $d \sin \vartheta = m\lambda$.
- The minima are found at angles $d \sin \vartheta = (m + \frac{1}{2})\lambda$.
- A diffraction grating consisting of N very narrow slits with a uniform spacing d between slits produces an intensity pattern

$$I(\vartheta) = I_0 \frac{\sin^2 \frac{1}{2} N\delta}{\sin^2 \frac{1}{2}\delta} \qquad \delta = kd \sin \vartheta$$

- The maximum intensity in the peaks is $N^2 I_0$.
- The resolving power of a grating in m^{th} order is

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

- The dispersion of a grating with spacing d is

$$D = \frac{\mathrm{d}\,\vartheta}{\mathrm{d}\,\lambda} = \frac{m}{d \cos \vartheta}$$

- Dispersion and resolving power are **independent** quantities!

Exercises

Halliday Resnick & Walker:

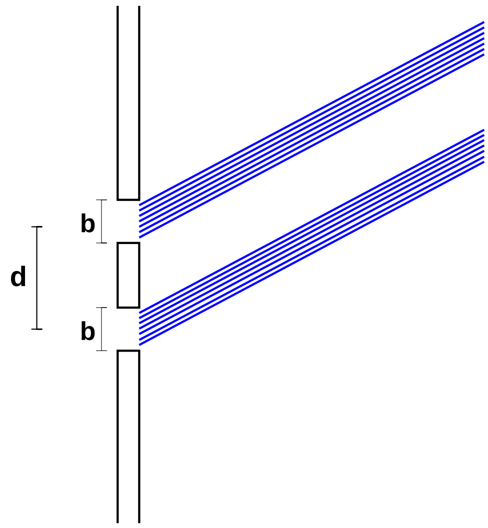
Reading: Ch37 HRW pp. 901-912

Exercises: p. 914 ff.: 27E, 31P, 32P, 33E, 35E, 37P, 41P, 43P, 45P, 47E, 51P

4.37: Double Slit Diffraction

We have already discussed Young's double slit experiment and found the conditions for interference maxima and minima. Then we used very narrow slits.

In general the slits have a finite width, and the intensity pattern on a screen behind the double slits will be a combination of two single slit diffraction patterns and the double slit interference pattern.



The single slit had an intensity pattern

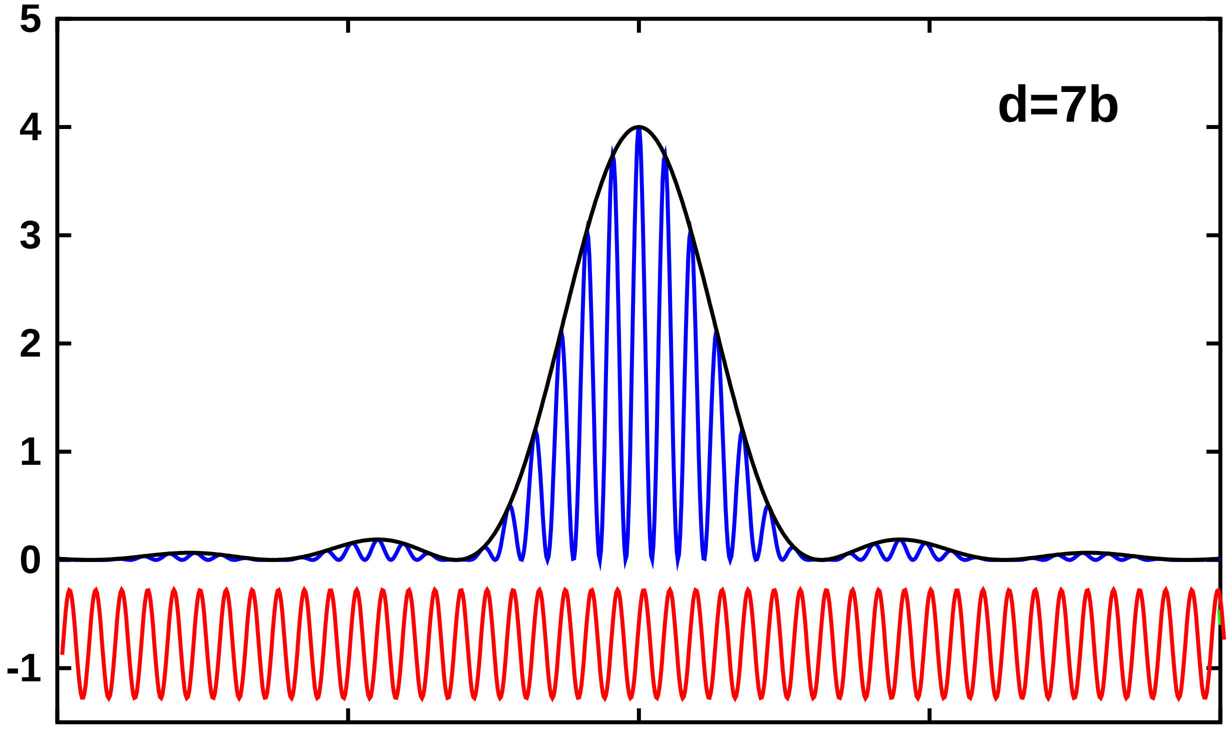
$$I_{ss} = I_0 \frac{\sin^2 \alpha}{\alpha^2} \qquad \alpha = \tfrac{1}{2} kb \sin \vartheta$$

Young's double slits with infinitesimal width at a distance d created an interference pattern

$$I_{Y\,ds} = 4 I_0 \cos^2 \beta \qquad \beta = \tfrac{1}{2} kd \sin \vartheta$$

We can combine the two if we define d to be the distance between the centers of the two slits of width b and look at a screen sufficiently far away so that we can treat all rays as parallel:

$$I = I_{ss} \cdot I_{Yds} = 4I_0 \frac{\sin^2 \alpha}{\alpha^2} \cos^2 \beta$$



If b tends towards 0, the central single slit maximum tends towards an infinite width, and the intensity pattern of Young's double slits is obtained. If $b >> d$ the pattern becomes that of a single slit.

4.38: Realistic Grating

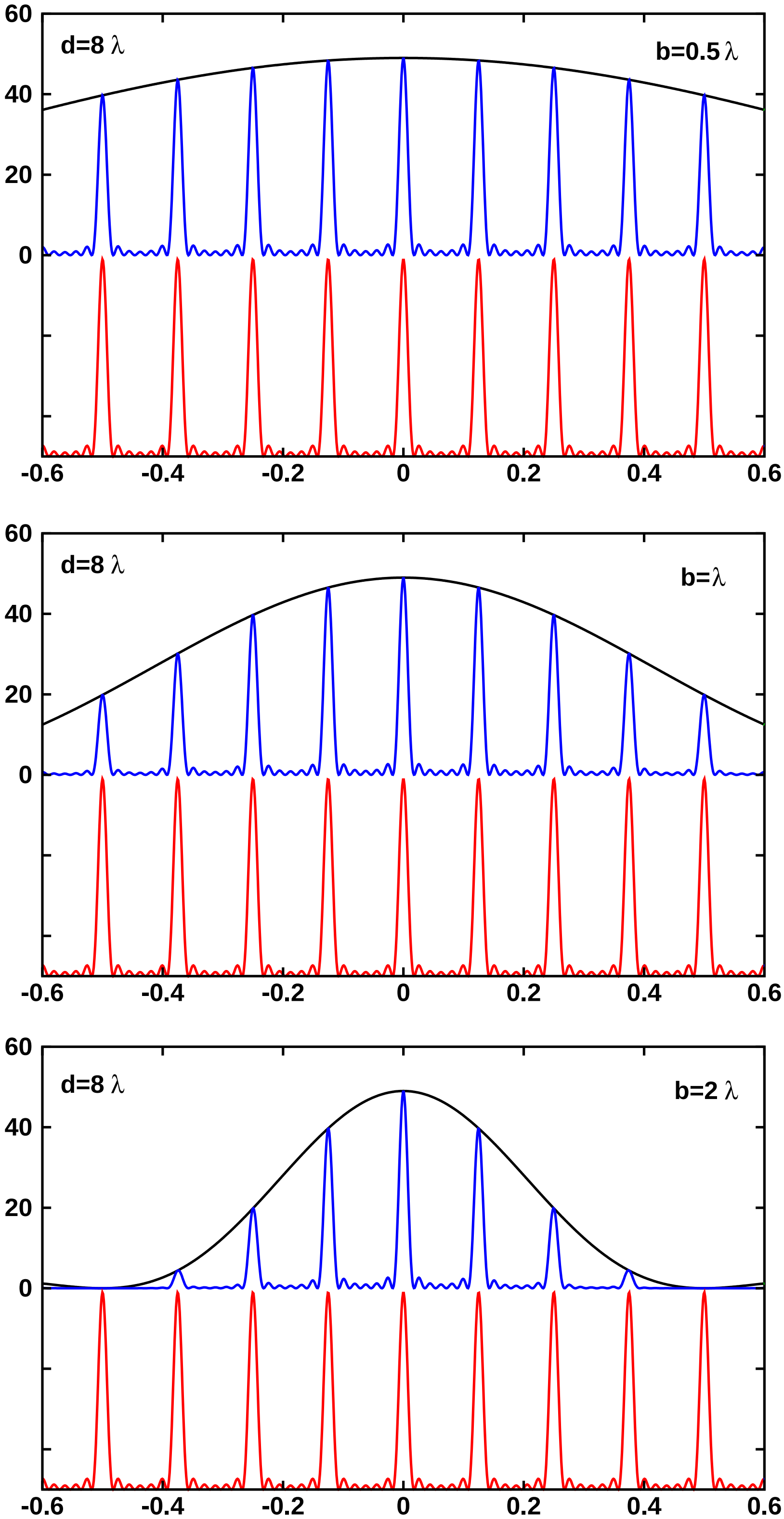
The same is true for a realistic diffraction grating with finite slitwidth.

The total intensity pattern will be a product of the diffraction pattern created by one finite slit and the interference between the contribution from the N different slits.

$$I = I_0 \frac{\sin(\frac{1}{2}N\delta)}{\sin(\frac{1}{2}\delta)} \frac{\sin^2\alpha}{\alpha^2}$$

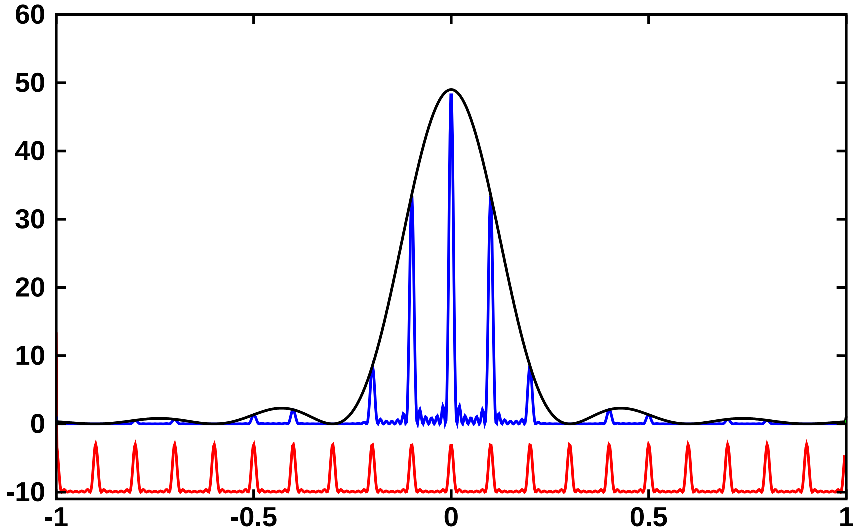
with $\delta = kd\sin\vartheta$ and $\alpha = \frac{1}{2}kb\sin\vartheta$.

One must be careful here. A diffraction pattern is really nothing but a complicated interference pattern created by all the Huygens wavelets originating at the slit. It is therefore a matter of semantics to call one a diffraction pattern and the other an interference pattern.


60
40
20
0
d=8 λ
b=0.5 λ
-0.6
-0.4
-0.2
0
0.2
0.4
0.6
60
40
20
0
d=8 λ
b=λ
-0.6
-0.4
-0.2
0
0.2
0.4
0.6
60
40
20
0
d=8 λ
b=2 λ
-0.6
-0.4
-0.2
0
0.2
0.4
0.6

4.39: Missing orders

Take a grating with slits 2 μm wide and spaced 6 μm apart illuminated with light at a wavelength $\lambda = 600$ nm. What is the intensity distribution?



We seem to lose every other line from the grating pattern because the diffraction pattern created by each slit does not produce any light at those angles!

4.40: Group velocity

If a pulse of light moves through a dispersive medium, all the different wavelengths making up the pulse move at different speeds. What then is the speed of the signal progressing through the medium?

Go back to our discussion of beats. If two waves of slightly different wavelengths λ and λ' and slightly different velocities v and v' travel together, we can find the resulting wave easily:

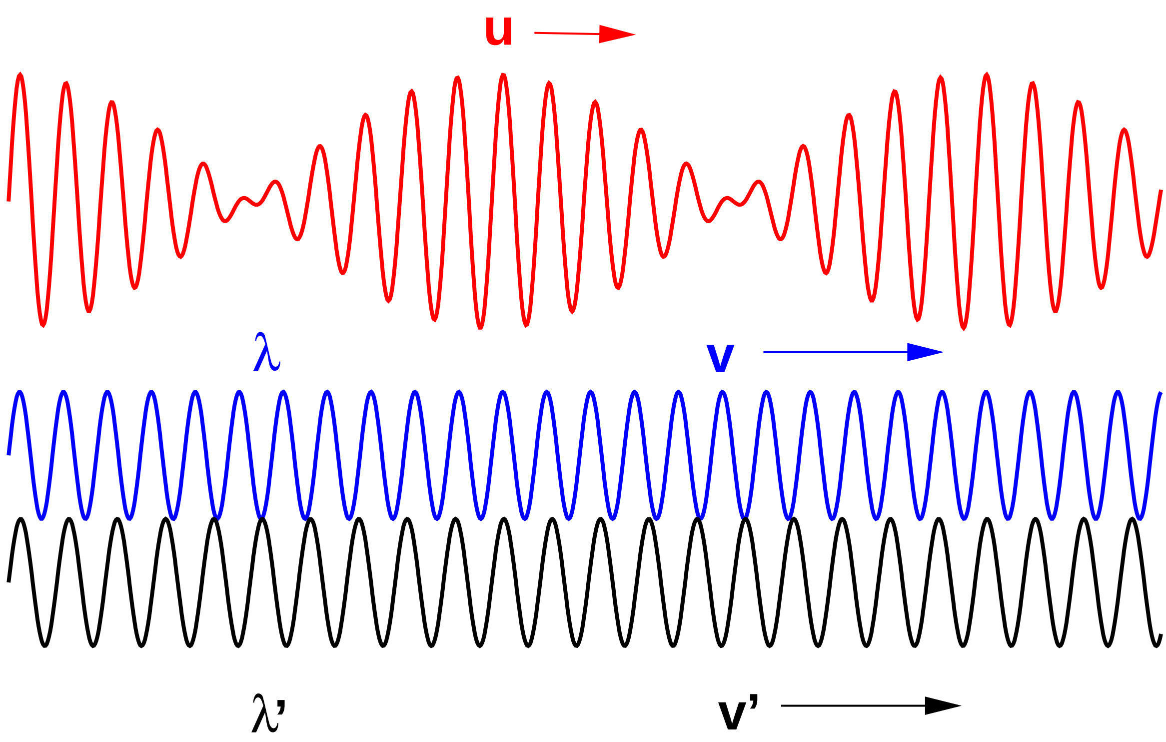
$$y_1(x,t) = y_0 \sin(kx - \omega t) = y_0 \sin(k(x - vt))$$

$$y_2(x,t) = y_0 \sin(k'x - \omega' t) = y_0 \sin(k'(x - v't))$$

with $\omega' = \omega + \mathrm{d}\,\omega$ and $k' = k + \mathrm{d}\,k$.

The resultant wave will be

$$\begin{aligned} y(x,t) &= y_1(x,t) + y_2(x,t) \\ &= y_0 \sin(kx - \omega t) + y_0 \sin(k'x - \omega' t) \\ &= 2 \sin(\tfrac{1}{2}(k+k')x + \tfrac{1}{2}(\omega + \omega')t) \\ &\quad \times \cos(\tfrac{1}{2}(k-k')x + \tfrac{1}{2}(\omega - \omega')t) \end{aligned}$$



The first part has a velocity

$$v = \frac{\frac{1}{2}(\omega + \omega')}{\frac{1}{2}(k + k')} \simeq \frac{\omega}{k}$$

and represents the small oscillations.

The second part is the speed of the envelope of the beats:

$$u = \frac{\frac{1}{2}(\omega - \omega')}{\frac{1}{2}(k - k')} \simeq \frac{\mathrm{d}\,\omega}{\mathrm{d}\,k}$$

It is this speed that we need to look at if we try to transmit a signal!

We call v the phase velocity and u the group velocity. To get a relationship between u and v we take

$$u = \frac{\mathrm{d}\,\omega}{\mathrm{d}\,k} = \frac{\mathrm{d}\,v(k)k}{\mathrm{d}\,k} = v(k)\frac{\mathrm{d}\,k}{\mathrm{d}\,k} + k\frac{\mathrm{d}\,v(k)}{\mathrm{d}\,k} = v + k\frac{\mathrm{d}\,v(k)}{\mathrm{d}\,k}$$

We can rewrite this in terms of the wavelength $\lambda = 2\pi/k$:

$$\frac{\mathrm{d}\,v(k)}{\mathrm{d}\,k} = \frac{\mathrm{d}\,v(k)}{\mathrm{d}\,\lambda}\frac{\mathrm{d}\,\lambda}{\mathrm{d}\,k} = \frac{\mathrm{d}\,v(k)}{\mathrm{d}\,\lambda}\frac{-2\pi}{k^2}$$

$$u = v + k\frac{\mathrm{d}\,v(k)}{\mathrm{d}\,k} = v + k\frac{\mathrm{d}\,v(k)}{\mathrm{d}\,\lambda}\frac{-2\pi}{k^2} = v - \lambda\frac{\mathrm{d}\,v(k)}{\mathrm{d}\,\lambda}$$

This form is useful because it shows that the group velocity is always smaller than the phase velocity, provided the wave speed is low for small wavelengths and increases monotonously.

This is the normal case in all dispersive media.

4.41: Summary

- The intensity pattern produced by realistic gratings and double slits is the product of the intensity pattern for a single slit and the intensity pattern for an appropriate number of very narrow slits:

  Double slit:

  $$I = I_{ss} \cdot I_{Y\,ds} = 4I_0 \frac{\sin^2\alpha}{\alpha^2}\cos^2\beta$$

  Grating:

  $$I = I_0 \frac{\sin(\frac{1}{2}N\delta)}{\sin(\frac{1}{2}\delta)} \frac{\sin^2\alpha}{\alpha^2}$$

  with $\alpha = \frac{1}{2}kb\sin\vartheta$, $\beta = \frac{1}{2}kd\sin\vartheta$ and $\delta = kd\sin\vartheta$.

- The interplay of minima from both patterns can lead to missing orders.

- A signal can only be transmitted with a pulse. In a dispersive medium this pulse travels with the group velocity

  $$u = \frac{\mathrm{d}\,\omega}{\mathrm{d}\,k}$$

- The group velocity u and the phase velocity v are connected via

  $$u = v - \lambda\frac{\mathrm{d}\,v(k)}{\mathrm{d}\,\lambda}$$

Exercises

Halliday Resnick & Walker:

Reading: Ch37 HRW pp. 901-906

Exercises: p. 914 ff.: 27E, 31P, 32P, 33E, 35E, 37P, 41P, 43P, 45P, 47E, 51P