2.12. Wave optics: Diffraction

Why do the interference patterns diminish in intensity with higher order?

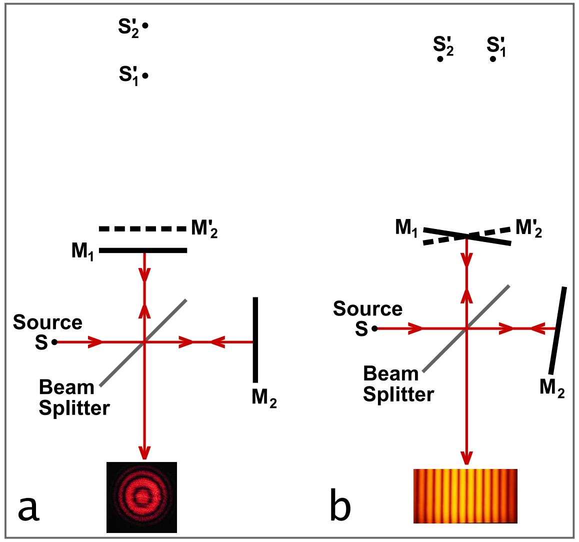
`os26`

$\Rightarrow$ **Diffraction**

Intermezzo: Michelson interferometer

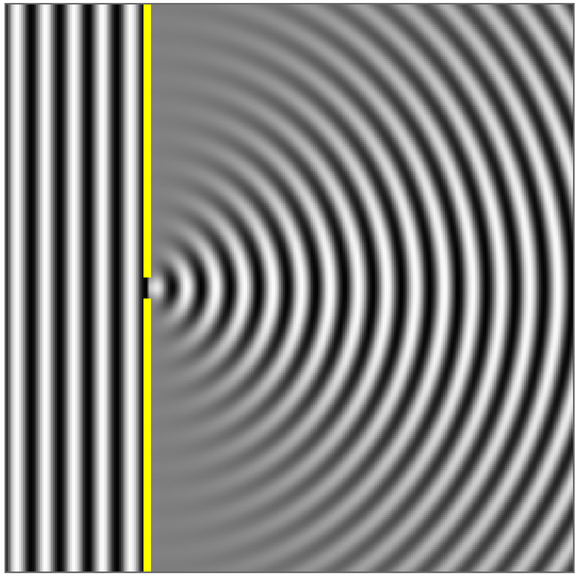
`os26`

- splits a light beam into two paths using a beam splitter
- each path reflects off a mirror and recombines at the detector
- interference depends on **optical path difference (OPD)** between arms
- if one mirror moves by distance Δx, the OPD changes by $2\Delta x$
- used in metrology, gravitational wave detection (LIGO), and coherence experiments



Introduction to diffraction

- **diffraction**: bending of waves as they pass through an aperture or around an obstacle
- direct consequence of light's wave nature
- noticeable when aperture/obstacle size is comparable to wavelength



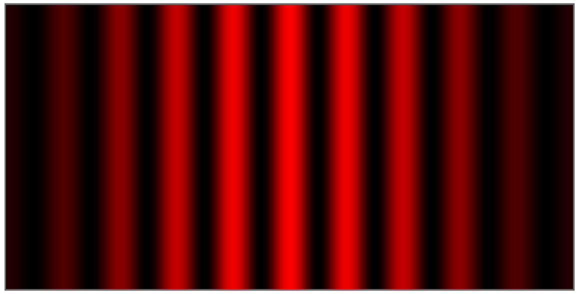
Revisiting the double-slit experiment: What we know

`os05 - double slit w\ red and green`

- previously, double-slit: interference of two discrete waves
- waves **bend** (diffract) at interfaces comparable to wavelength
- interference pattern: bright (constructive) and dark (destructive) fringes

Revisiting the double-slit experiment: The catch

- simple interference predicts equally bright fringes
- **reality**: intensity of fringes is modulated
  - central ones brightest
  - intensity decreases away from center
- each slit acts as a source of waves (Huygens' principle) that interfere in a more complex way
- this phenomenon, responsible for intensity variations, is **diffraction**



from ***wikipedia***, ***Attribution-Share Alike 3.0 Unported***

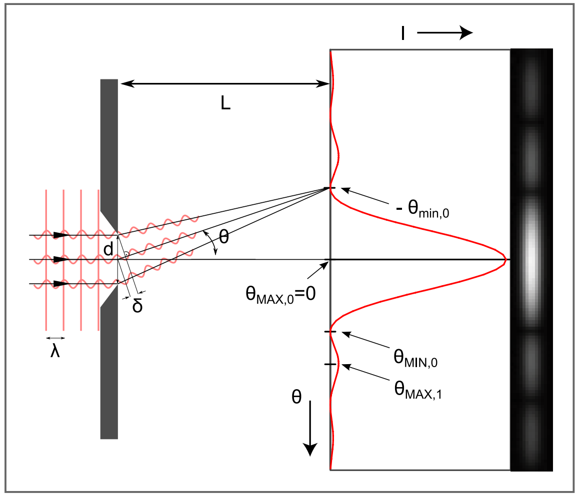
Interference vs. diffraction

- **fundamentally the same phenomenon**: superposition of coherent waves
- **distinction often lies in conceptualization/source arrangement**:
  - **interference**: superposition from a few discrete sources (e.g., two rays)
  - **diffraction**: superposition from a continuous distribution of sources or many closely spaced sources

Diffraction at a single-slit

`os05 - single slit`

- monochromatic light from coherent source (λ, phase)
- single narrow slit of width d ($d \approx \lambda$)
- results in a diffraction pattern on a distant screen:
    - central bright maximum
    - flanked by minima and weaker secondary maxima



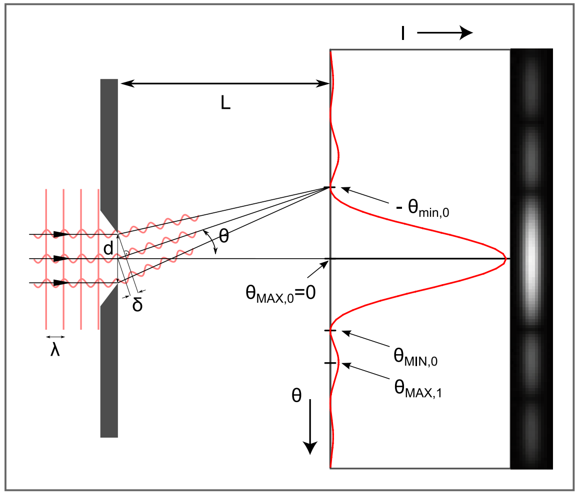
Single-slit pattern formation

`sim - single slit`

- path difference is $\Delta = d\sin\theta$, thus relation to wavelength λ:

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

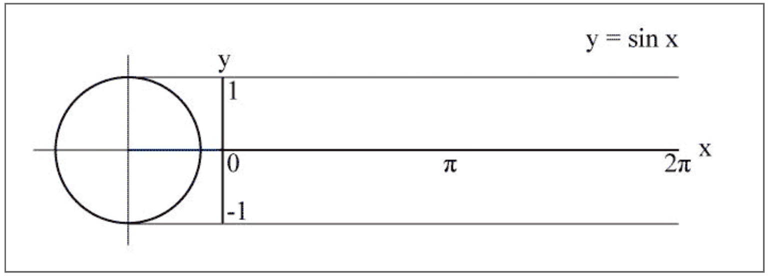
- m is the order *(note: m is a rational number, not integer)*
- **minima**: $d\sin\theta = m\lambda, \quad m = \pm 1, \pm 2, \ldots$
- **central maxima** at $m = 0$
- **higher-order maxima**:$d\sin\theta \approx (m + \frac{1}{2})\lambda, \quad m \approx \pm\frac{3}{2}, \pm\frac{5}{2}, \ldots$



Intensity in single-slit diffraction pattern

`sim - single slit`

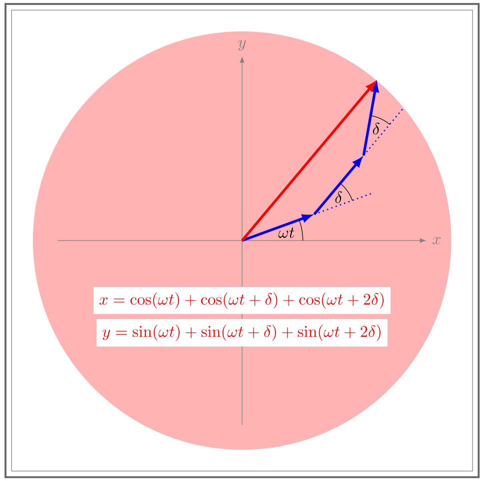
- split slit into N thin strips, each thickness Δy
- each strip emits coherent wavelets (Huygens' principle), i.e. can be considered an independent, coherent light "source"
- consider parallel rays emitted by these "sources" at angle θ
- path difference: $\Delta = \Delta y \sin\theta$
- phase difference: $\Delta\beta = \frac{2\pi}{\lambda}\Delta = \frac{2\pi}{\lambda}\Delta y \sin\theta$
- phase difference gives:
  - minima for: $\Delta\beta = \pm 2\pi, \pm 4\pi, \ldots$
  - maxima for: $\Delta\beta = \pm 3\pi, \pm 5\pi, \ldots$



Phasor technique to understand diffraction

- each strip/"source" has its own **electric field** with amplitude ΔE_0
- **phase differs between strips** $\Rightarrow$ electric field is a **vector** (magnitude & phase)
- intensity on screen: **vector sum of all strips**
- **total phase difference** β across all slits (phase per strip $\Delta\beta = \frac{2\pi}{\lambda}\Delta y \sin\theta$ & width $D = N\Delta y$):

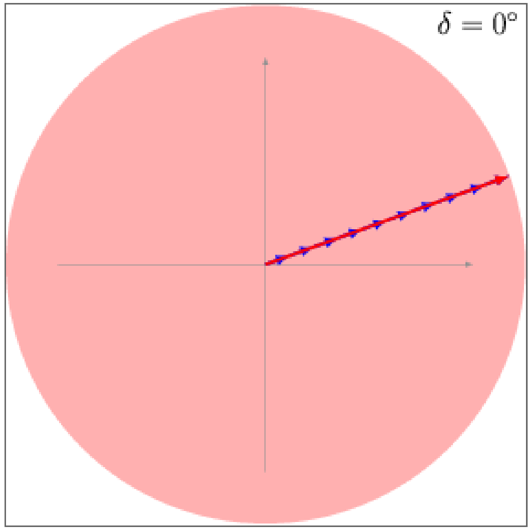
$$\beta = N\Delta\beta = \frac{2\pi}{\lambda}N\Delta y \sin\theta = \frac{2\pi}{\lambda}D\sin\theta$$



Note: δ in plot is our $\Delta\beta$, i.e. the angle of each individual phasor

Phasor summation

- if $\beta = 2\pi$, all vectors cancel (first minimum)
- minima for $\beta = \pm 2\pi, \pm 4\pi, \ldots$
- higher-order maxima for $\beta = \pm 3\pi, \pm 5\pi, \ldots$
  - portion of vectors cancel, reducing intensity
- from first minimum, $\beta = 2\pi = \frac{2\pi}{\lambda} D \sin\theta \Rightarrow \lambda = D \sin\theta \Rightarrow \sin\theta = \frac{\lambda}{D}$

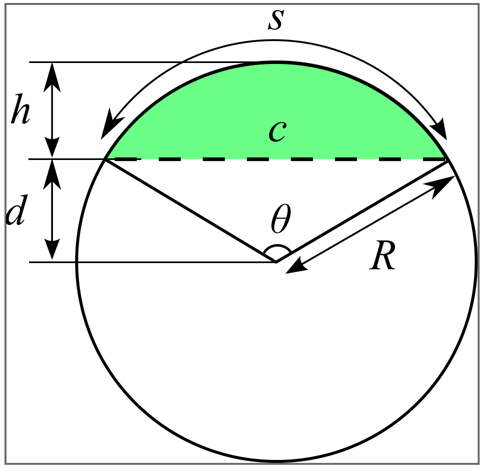


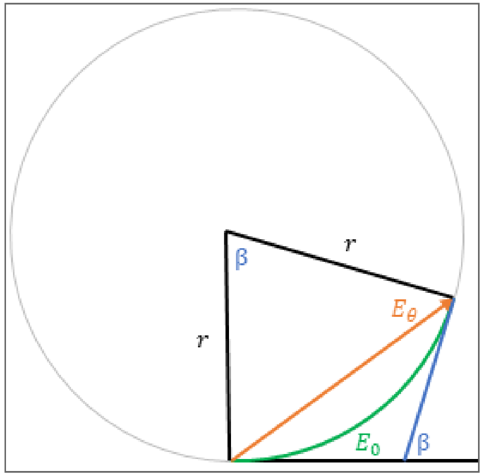
from ***wikipedia***, ***Attribution-Share Alike 4.0 International***

Note: δ in plot is our $\Delta\beta$, i.e. the angle of each individual phasor

Concept to deriving the intensity formula

- **phasors form an arc**
  - **arc length** s is E_0 (our input)
  - **chord** c of the arc, i.e. arrow connecting start and end of arc, is E_θ (our output)
- **for central maximum** ($\theta = 0$): phasors in phase, resultant $E_\theta = E_0 = N\Delta E_0$
- in general:
  - **arc length:** $s = \theta R \Leftrightarrow E_0 = \beta r$
  - **chord length:** $c = 2R\sin(\frac{\theta}{2}) = \Leftrightarrow E_\theta = 2r\sin(\frac{\beta}{2})$





*[left]from **wikipedia**, public domain*

Single-slit intensity formula

- rewrite equations:
  - $E_0 = \beta r \Rightarrow \frac{E_0}{2} = r\frac{\beta}{2}$
  - $E_\theta = 2r\sin(\frac{\beta}{2}) \Rightarrow \frac{E_\theta}{2} = r\sin(\frac{\beta}{2})$
- divide them: $\frac{E_\theta}{E_0} = \frac{\sin(\beta/2)}{\beta/2}$
- intensity is proportional to square of electric field amplitude ($I \propto E^2$):

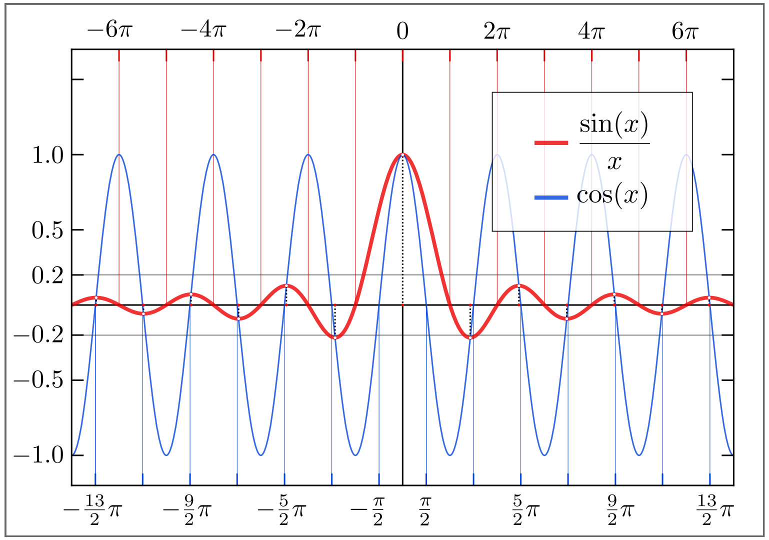
$$\frac{I_\theta}{I_0} = \left(\frac{E_\theta}{E_0}\right)^2 = \left(\frac{\sin(\beta/2)}{\beta/2}\right)^2$$

- substituting $\beta = \frac{2\pi}{\lambda} D \sin\theta$:

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

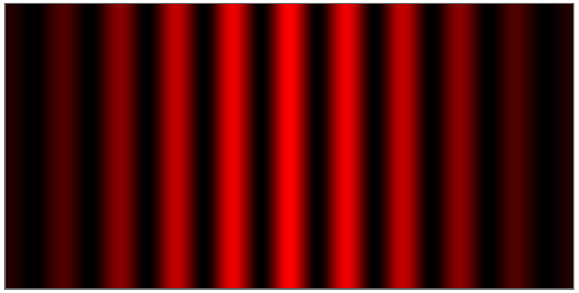
Interpretation single-slit intensity

- $I_\theta = I_0 \left(\frac{\sin\left(\frac{\pi D \sin\theta}{\lambda}\right)}{\frac{\pi D \sin\theta}{\lambda}} \right)^2$
- use sinc function $\mathrm{sinc}(x) = \frac{\sin x}{x}$:
- $I_\theta = I_0 \cdot \mathrm{sinc}^2\left(\frac{\pi D \sin\theta}{\lambda}\right)$
- $\mathrm{sinc}^2(x) = (\frac{\sin(x)}{x})^2$ is sinc squared function:
  - a central maximum at $x = 0$
  - amplitude decreases for larger values of x
  - zeros occurring at multiples of π.
- $\Rightarrow$ **mathematically explains the central bright maximum and the decreasing intensity of the secondary maxima**


-6π
-4π
-2π
0
2π
4π
6π
$\frac{\sin(x)}{x}$
$\cos(x)$
1.0
0.5
0.2
-0.2
-0.5
-1.0
$-\frac{13}{2}\pi$
$-\frac{9}{2}\pi$
$-\frac{5}{2}\pi$
$-\frac{\pi}{2}$
$\frac{\pi}{2}$
$\frac{5}{2}\pi$
$\frac{9}{2}\pi$
$\frac{13}{2}\pi$

Revisiting diffraction at a double-slit

- **previously**: interference analysis determines maxima/minima positions
- **reality**: finite number of peaks, brightest at center, lower intensity surrounding
- this is due to **diffraction** from each slit



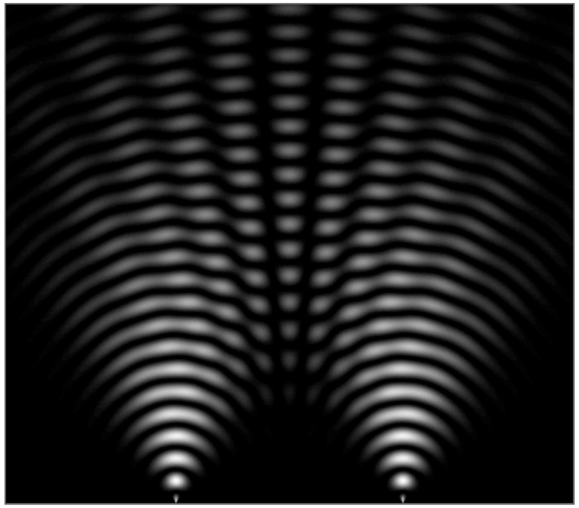
from ***wikipedia**, **Attribution-Share Alike 3.0 Unported***

Double-slit as two single slits to combine diffraction & interference

- double-slit: each slit has width D, separation d
- each slit contributes an electric field modulated by its own diffraction, i.e. **reuse single-slit results**:

$$E_{\text{single}} = E_{0,\text{single}} \frac{\sin(\beta/2)}{\beta/2}$$

- with $\frac{\beta}{2} = \frac{\pi D \sin\theta}{\lambda}$



*from **wikipedia**, gemeinfrei*

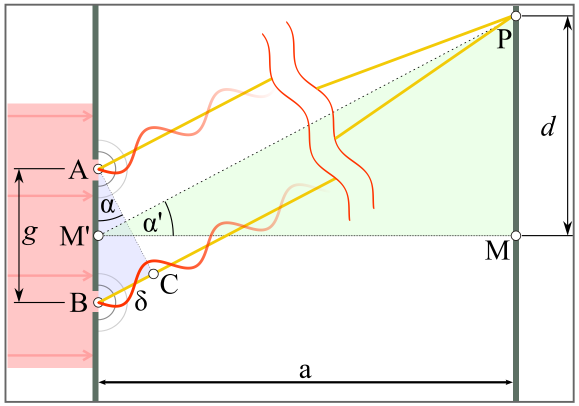
Phase difference between slits

- path difference between light from two slits: $\Delta = d \sin\theta$
- phase difference δ:

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

- electric fields from the two slits:

$$E_1 = E_{single} e^{i\delta/2}, \quad E_2 = E_{single} e^{-i\delta/2}$$



Total electric field for double slit

- superposition:

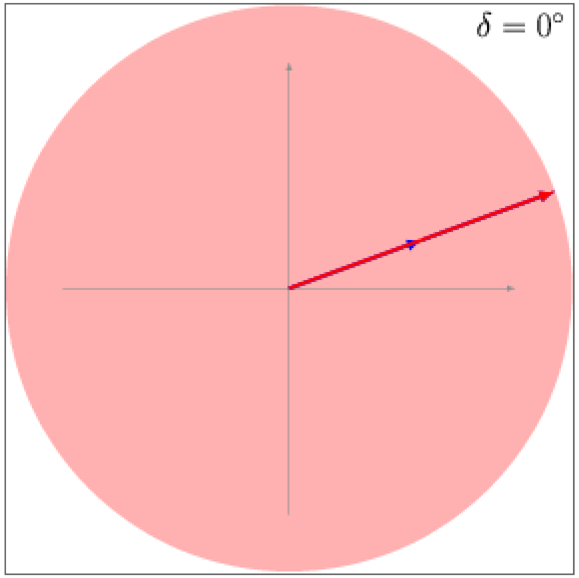
$$E_{total} = E_1 + E_2 = E_{single}(e^{i\delta/2} + e^{-i\delta/2})$$

- using $e^{ix} + e^{-ix} = 2\cos x$:

$$E_{total} = 2E_{single}\cos\left(\frac{\delta}{2}\right)$$

- substituting E_{single}:

$$E_{total} = 2E_{0,single}\frac{\sin(\beta/2)}{\beta/2}\cos\left(\frac{\delta}{2}\right)$$


$\delta = 0°$

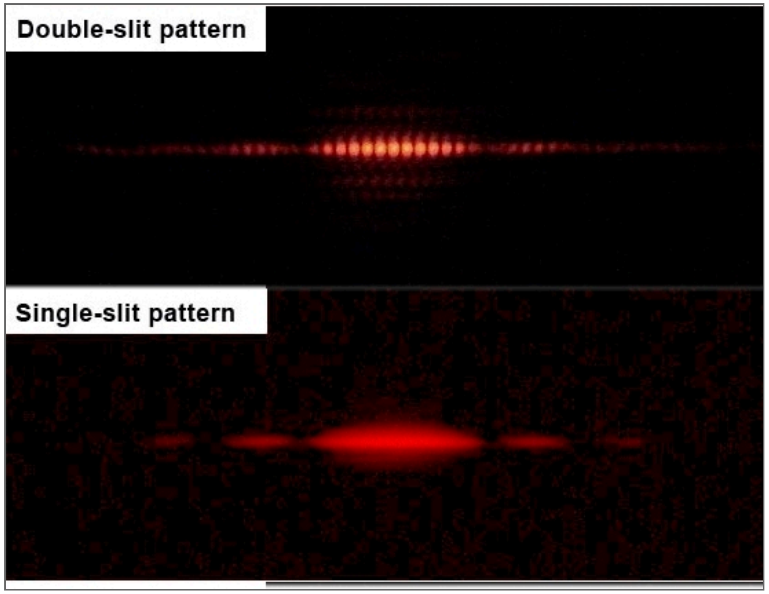
Double-slit intensity formula & interpretation

`sim - intensity`

- intensity $I_\theta \propto E_{total}^2$
- let I_0 be the intensity of the central maximum:

$$I_\theta = I_0 \left(\frac{\sin(\beta/2)}{\beta/2} \right)^2 \cos^2\left(\frac{\delta}{2}\right)$$

- **diffraction factor (envelope)**: $\left(\frac{\sin(\beta/2)}{\beta/2} \right)^2$ from each slit (D)
- **interference factor**: $\cos^2\left(\frac{\delta}{2}\right)$ from path difference between slits (d)
- diffraction envelope modulates finer interference fringes
- zeros of diffraction pattern cause disappearance of interference fringes



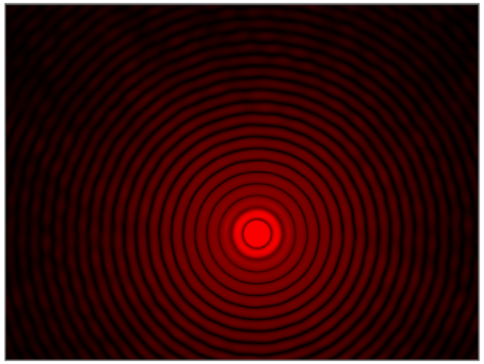
from ***wikipedia**, **Attribution-Share Alike 3.0 Unported***

Diffraction at circular apertures & resolution limit

`ow05 - Lochblende`

- lenses (circular apertures of diameter D) cannot image a point perfectly due to **diffraction & aberration**
- light from a point source forms an **Airy disk**
- angular half-width θ of Airy disk:

$$\theta \approx 1.22 \frac{\lambda}{D}$$



Resolution limit of a pinhole camera

`os07 - resolution limit`

- the resolution is limited mainly by two effects: **diffraction and geometric blur**
- **diffraction** occurs because light waves spread out when passing through the pinhole, causing image blur
- **geometric blur** happens if the pinhole is too large, letting rays from one point spread on the image plane
- there is an optimal pinhole diameter **balancing diffraction and geometric blur** for the sharpest image
- the **optimal pinhole diameter** to minimize blur is approximately

$$d_{\mathrm{opt}} \approx 1.9\sqrt{\lambda f}$$

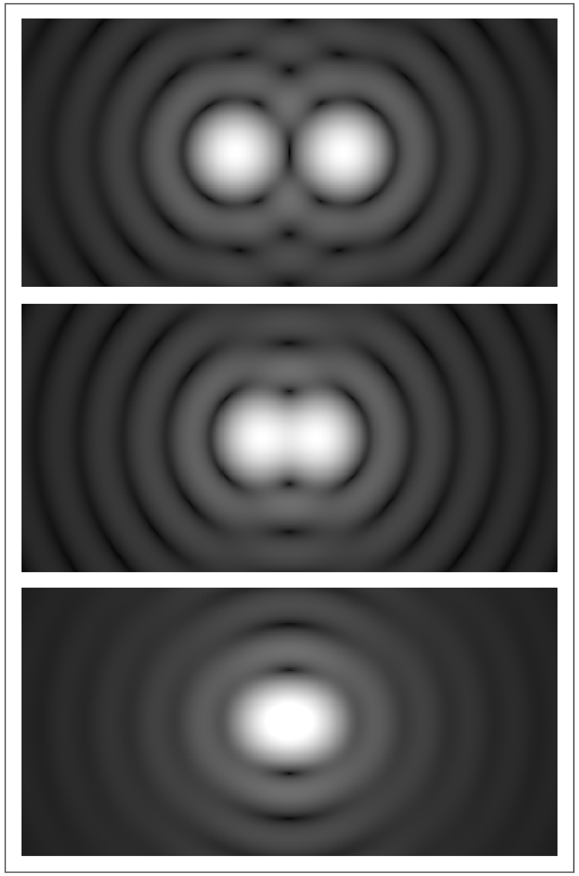
where f is the distance from the pinhole to the image plane (focal length)

Rayleigh criterion

- resolution limit: ability to distinguish two closely spaced objects
- **Rayleigh criterion**: just resolvable when one center overlaps other's first minimum
- minimum angular separation:

$$\theta_{min} = 1.22\frac{\lambda}{D}$$

- smaller $\theta_{min} \rightarrow$ better resolution
- applies to telescopes and mirrors (D = objective diameter)

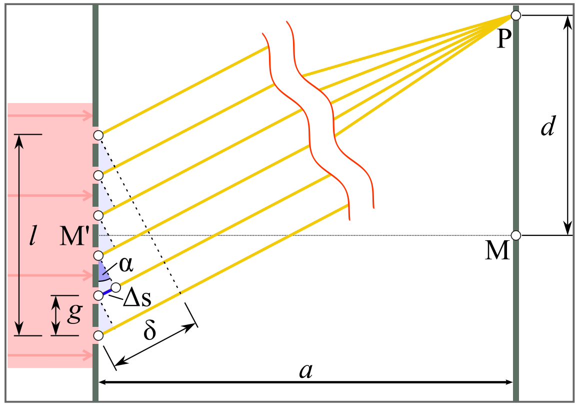


from ***wikipedia****, public domain*

Diffraction grating with monochromatic light

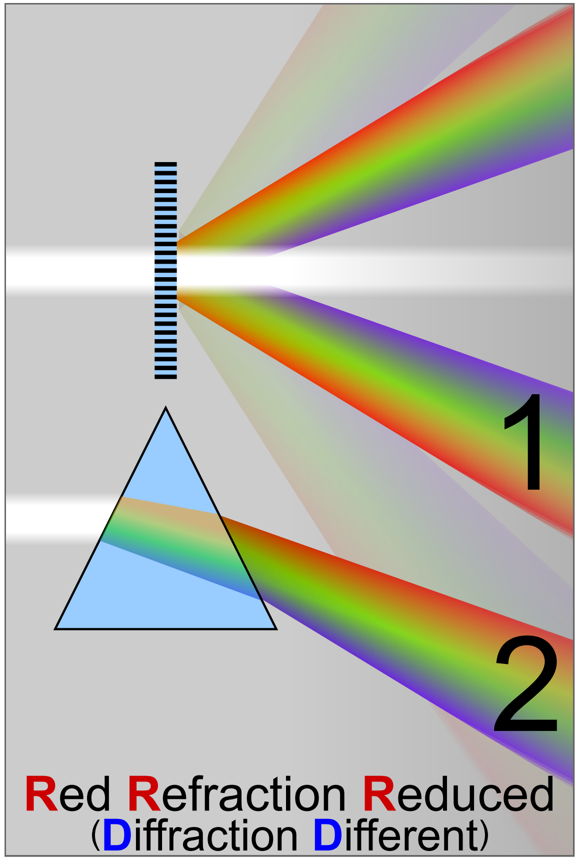
`ow05 - Gitter`

- **diffraction grating**: many equally spaced slits (spacing g)
- thousands of lines per cm/mm
- maxima occur at angles: $\sin\theta = \frac{m\lambda}{g}, \quad m = 0, \pm 1, \pm 2, \ldots$



Diffraction grating & spectroscopy

- **transmission grating**: light passes through
- **reflection grating**: lines ruled on mirror $\rightarrow$ light reflected
- **maxima sharper than for double-slit** as even slight angle change cause destructive interference across many slits
- used for precise wavelength measurements and **basis for spectroscopy**



from ***wikipedia***, ***Attribution-Share Alike 3.0 Unported***

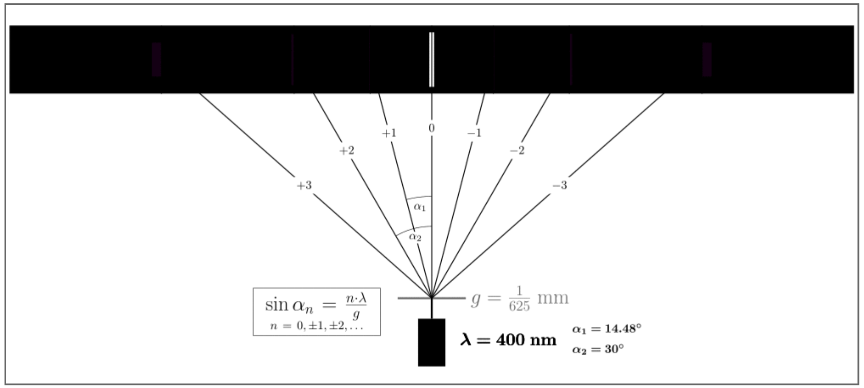
White light with diffraction grating

```
ow13 + ow10 - spect grating
```

- white light instead of monochromatic
- **central white peak** ($m = 0$): all wavelengths overlap constructively
- for $m \neq 0$: different λ diffract at different angles:

$$\sin\theta = \frac{m\lambda}{d}$$

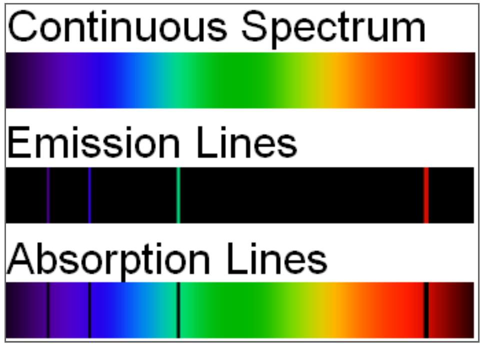
- **result**: spectrum of colors per order



Spectrometer / spectroscope

`ow15 + ow30 - spect lamps`

- **continuous spectrum**: emitted by hot solids, liquids, or dense gases; shows an unbroken range of colors (wavelengths)
- **emission lines**: produced by excited atoms in a low-density gas; specific wavelengths corresponding to electron transitions
- **absorption lines**: occur when light passes through a cooler gas; atoms absorb specific wavelengths, creating dark lines in the spectrum
- **dependency on temperature**: higher temperatures excite more electrons, increasing the intensity and number of emission lines; cooler gases lead to stronger absorption features



*from **wikipedia**, gemeinfrei*