Optical astronomical spectroscopy at the VLT

(Part 1)

F. Pepe

Observatoire de Genève

Course outline

- PART 1 - Principles of spectroscopy
  - Fundamental parameters
  - Overview of spectrometry methods
- PART 2 – 'Modern' spectrographs
  - 'Simple' spectroimager
    - FORS
  - Echelle spectrographs
    - UVES
    - CRIRES
- PART 3 - Spectroscopy on the VLTs
  - Multi-Object spectrographs (MOS) and Intergral-Field Units (IFU) and spectro-imagers
    - Giraffe
    - VIMOS
    - Sinfoni
  - Future instruments
    - X-shooter
    - MUSE
    - ESPRESSO

The observables

Propagation of light wave from stable source at infinity:

$$\vec{E}(\vec{x},t) = \vec{A} \cdot e^{-i(\vec{k}\vec{x}-\omega \cdot t)}$$

which is a solution of wave equations if:

$$c_n = \frac{\omega}{k}, \text{ where } k = \left|\vec{k}\right|$$

c_n is the speed of light in a medium with refractive index n = n(k)

Independent observables are:

$\vec{A}$ = Amplitude of electric field

(k_x, k_y) = Direction vector projected on sky

ω = Frequency or k = Wave vector

The distance between two spatial maxima of the light wave in a given medium and at fixed *t* is called wavelength and results to be:

$$\lambda_n = \frac{2\pi}{n(k) \cdot k}$$

The observables

Astronomical spectroscopy aims at measuring:

$$\vec{A}(\nu,(k_x,k_y)) \quad \text{or} \quad \vec{A}(\lambda,(k_x,k_y))$$

At optical wavelength $\nu=\omega/2\pi$ is 10^{15} Hz, thus too fast to be resolved by detectors. The observable becomes the (surface) brightness or specific intensity I_ν or I_λ:

$$I_\nu(k_x,k_y) = \overline{\left|\vec{E}_{\nu,\lambda}(t,\vec{x}_{obs})\right|^2}^{\,t} = \frac{1}{2}\left|\vec{A}(\nu,(k_x,k_y))\right|^2 \qquad [\mathrm{W\ m^{-2}\ sterad^{-1}\ Hz^{-1}}]$$

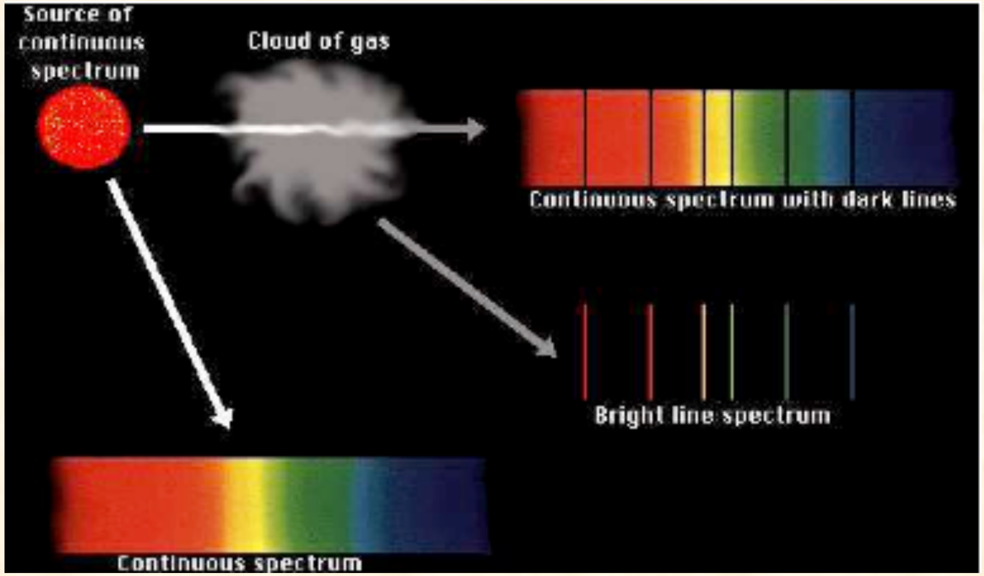
$$\text{or} \qquad I_\lambda(k_x,k_y) = \frac{1}{2}\left|\vec{A}(\lambda,(k_x,k_y))\right|^2 \qquad [\mathrm{W\ m^{-2}\ sterad^{-1}\ \mu m^{-1}}]$$

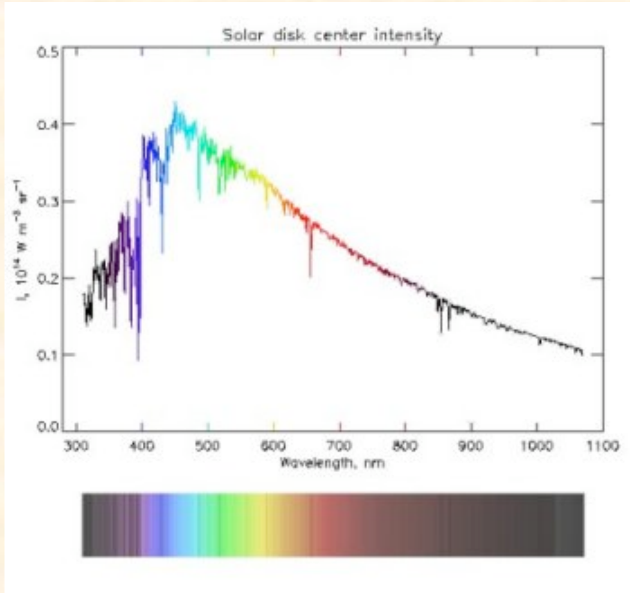
The observables

If we integrate the surface brightness over a given source or sky aperture, we get the spectral flux density F_ν or F_λ at a given light frequency or wavelength:

$$F_\nu = F(\nu) = \int S(\nu,(k_x,k_y)) \cdot \cos\Theta \cdot d\Omega \cong \int S(\nu,(k_x,k_y)) \cdot d\Omega$$

$$F_\lambda = F(\lambda) = \int S(\lambda,(k_x,k_y)) \cdot \cos\Theta \cdot d\Omega \cong \int S(\lambda,(k_x,k_y)) \cdot d\Omega$$





Wavelength or frequency?

Frequency (ν, ω), associated to time-evolution of E-field

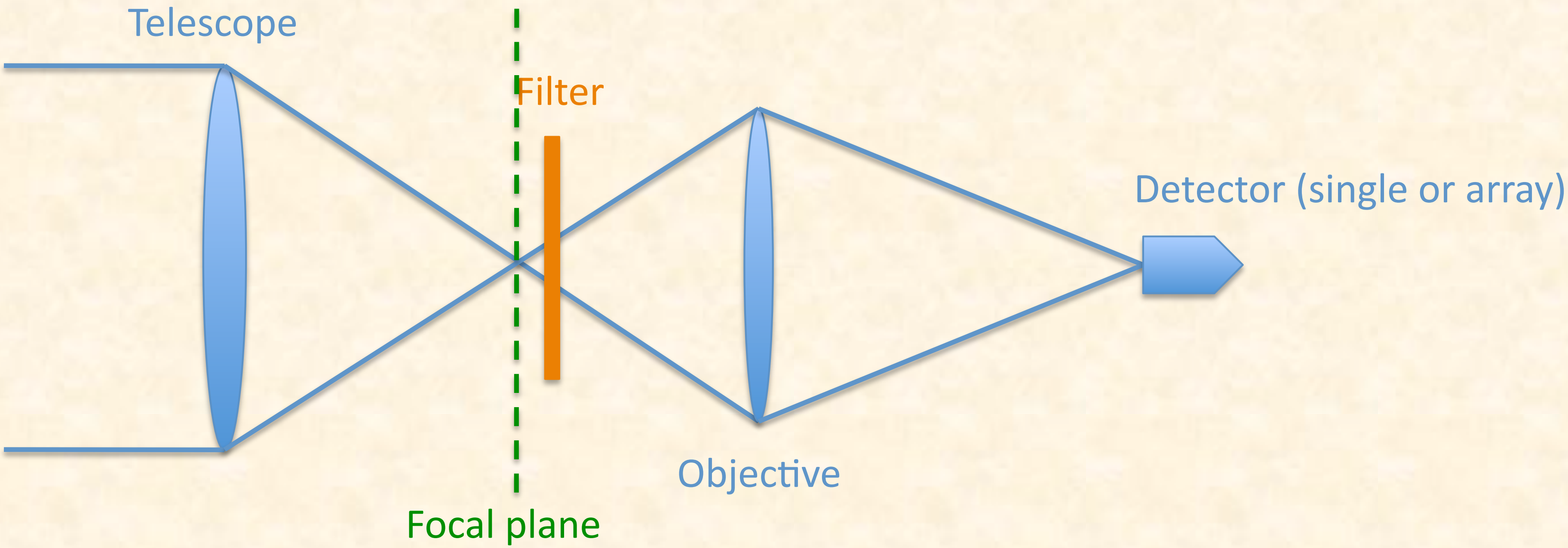
- Fixed location (detect oscillation amplitude as a function of time)
- Need for precise clock (time metric)
- Need for time resolution (or other tricks, e.g. heterodyne detection)

Wavelength (λ, k), represents light propagation in space

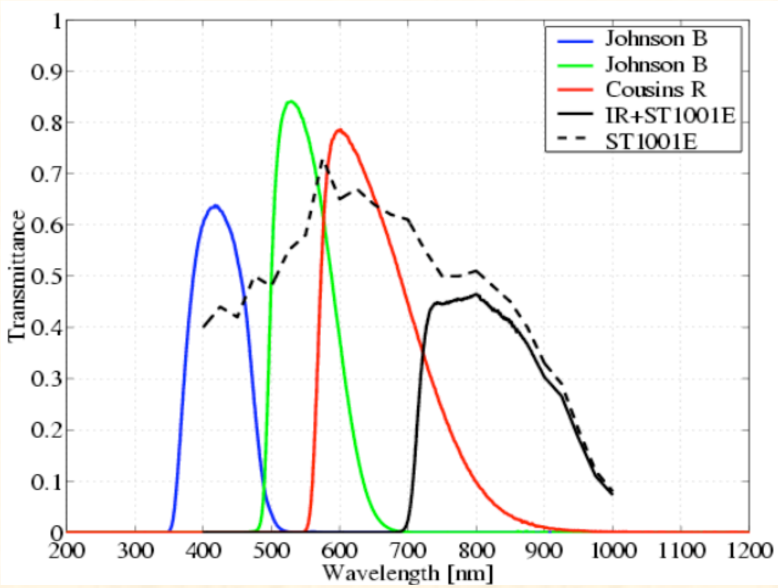
- Fixed time (project wave into physical space or spatially separate various spectral components)
- Depends on medium ->
- Need for precise metrology (space metric) or calibration

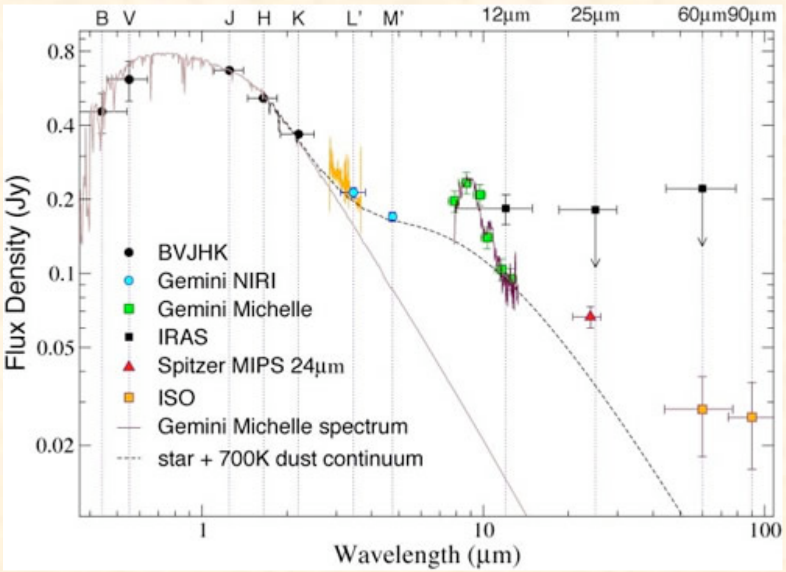
$$\vec{E}(\vec{x},t) = \vec{A} \cdot e^{-i(\vec{k}\vec{x} - \omega \cdot t)}$$

Filter spectrometer



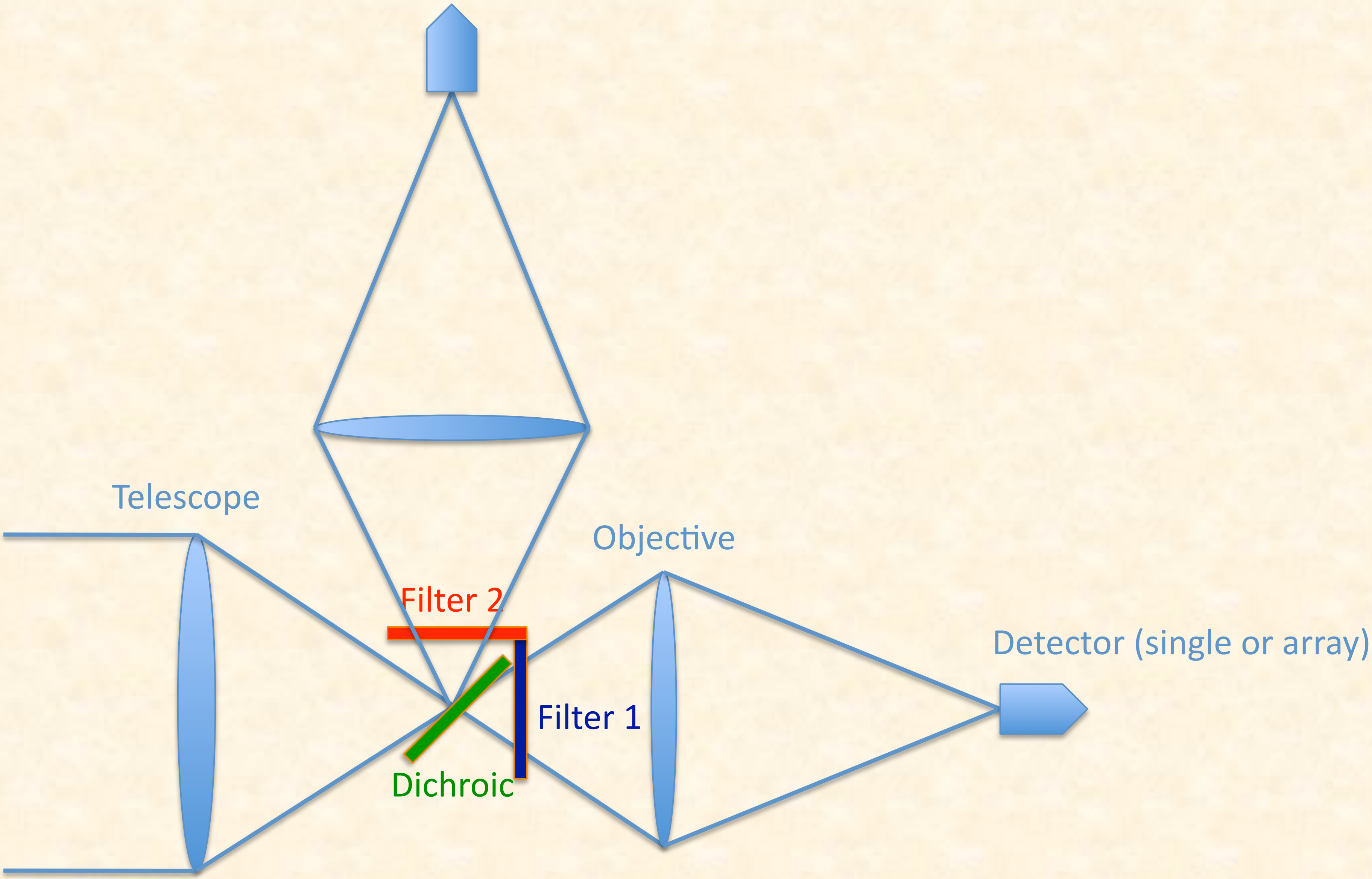
Filter spectrometer



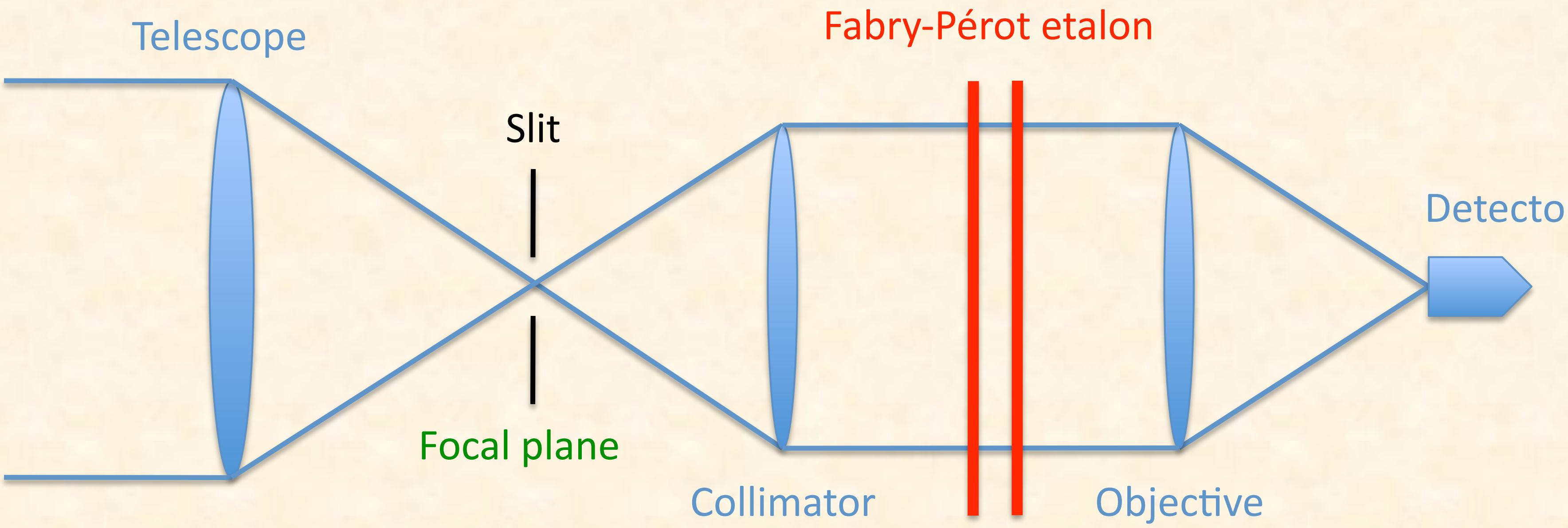


- Detector records I_λ for a given filter with transmittance t_c, central wavelength λ_c, and band width $\Delta\lambda$
- t_c, $\Delta\lambda$ and λ_c need to be calibrated on standard sources
- Appropriate for broad-band spectra
- Only one channel per measurement (unless dichroics are used)

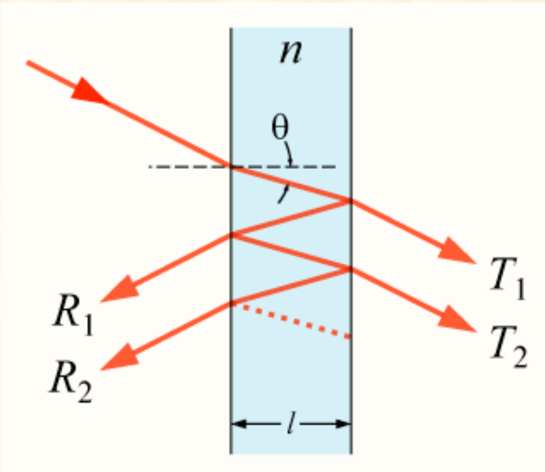
Filter spectrometer

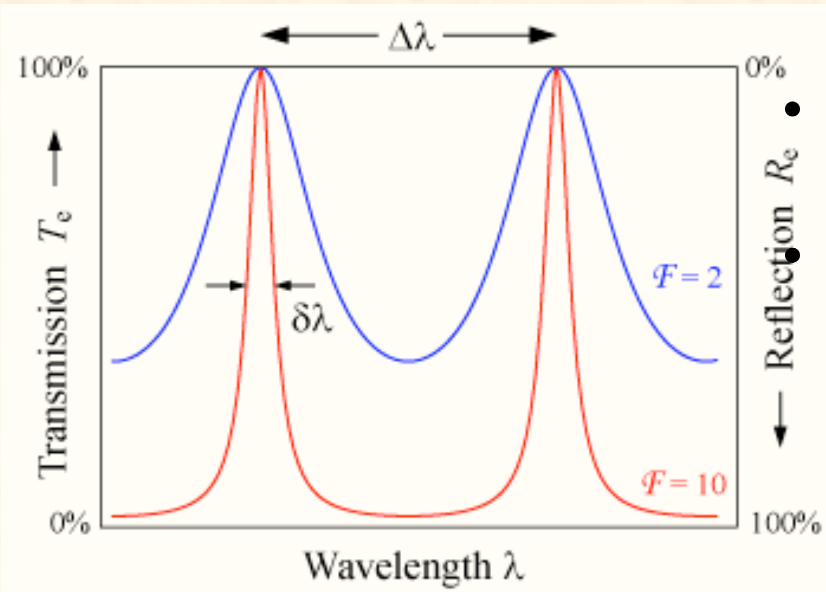


Fabry-Pérot spectrometer



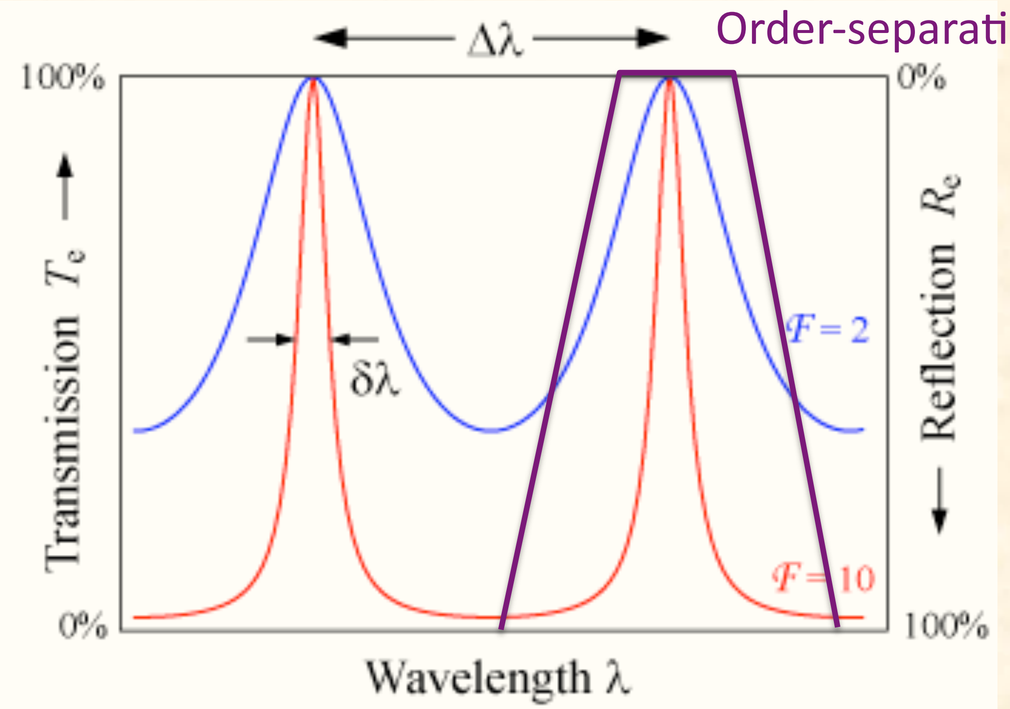
Fabry-Pérot spectrometer





- Similar to filter spectrometer, but spacing can be made tunable
- Detector records $I(\lambda)$ as a function of the transmitted wavelength $m\lambda=2l$, where m is an integer and enumerates the transmitted order.
- Only one spectral channel per measurement
- Transmittance and wavelength must be calibrated.
- Allows high spectral resolution, if the finesse F or the order m is high. In the latter case, a pre-filtering is required to select only one wavelength (order-selection).

Fabry-Pérot spectrometer



Order-separating filter (e.g. with second FP in 1st order)

$$T(\lambda) = \frac{1}{1 + (2F/\pi)^2 \cdot \sin^2(\delta(\lambda)/2)},$$

where

$$\delta(\lambda) = \frac{2\pi}{\lambda} 2nl\cos\Theta \ , \ F = \frac{\pi\sqrt{r}}{1-r}$$

(r = mirror reflectance)

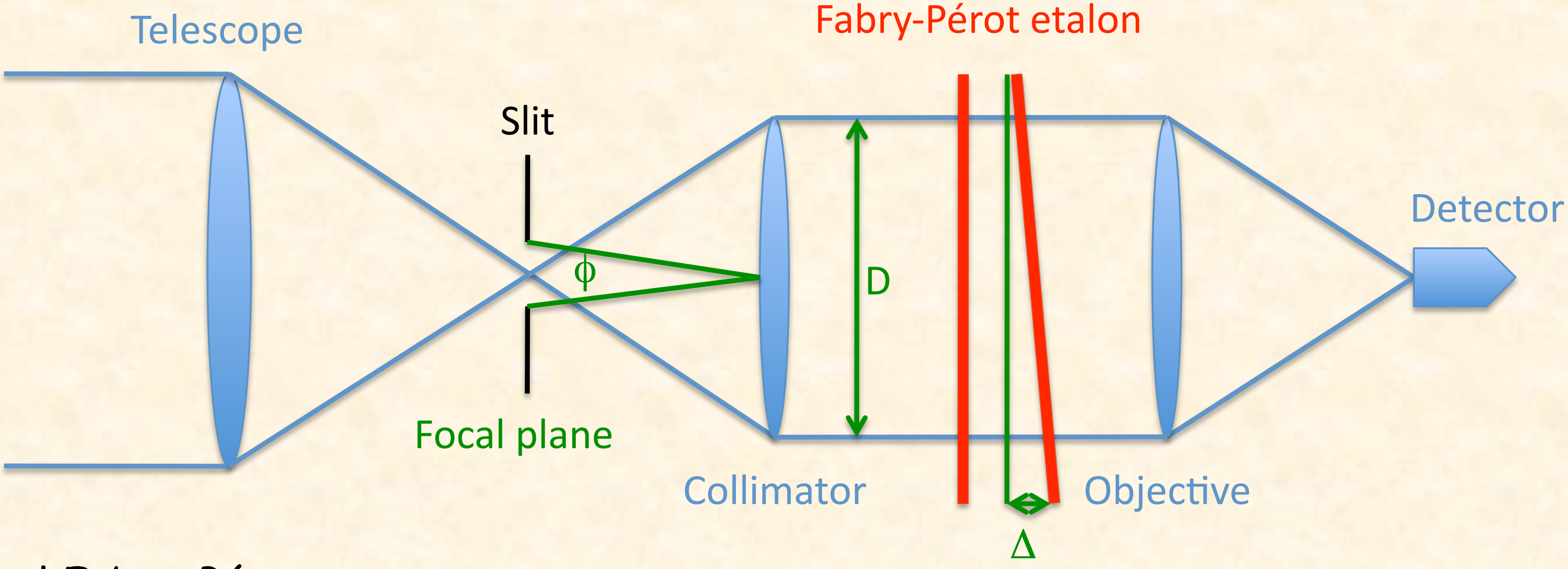
Transmitted wavelength: $\lambda_m = 2nl\cos\Theta / m$

Order separation: $\Delta\lambda = \lambda_{m-1} - \lambda_m \approx 2nl\cos\Theta / m^2$

Finesse: $F = \Delta\lambda / \delta\lambda$

Spectral resolution: $R := \dfrac{\lambda}{\delta\lambda} = m \cdot F$

Fabry-Pérot spectrometer



Real Fabry-Pérot:

$$T(\lambda) = \frac{1}{1 + (2F_E/\pi)^2 \cdot \sin^2(\delta(\lambda)/2)},$$

where $\dfrac{1}{F_E^2} = \dfrac{1}{F_R^2} + \dfrac{1}{F_D^2} + \dfrac{1}{F_P^2} + \dfrac{1}{F_\phi^2}$

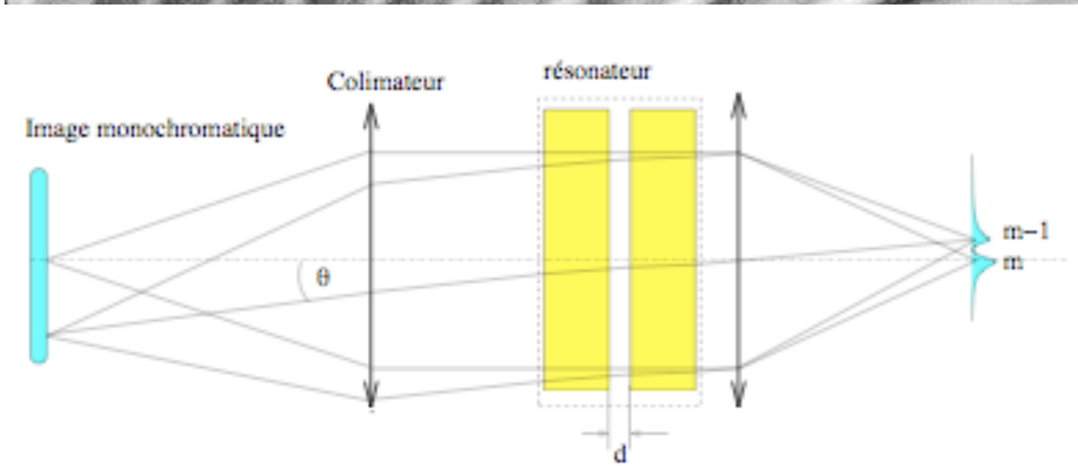
$F_R = \dfrac{\pi \cdot r}{1 - r},$ reflectance finesse

$F_D = \lambda/\delta/\sqrt{2},$ defect finesse ($\delta =$ defect rms)

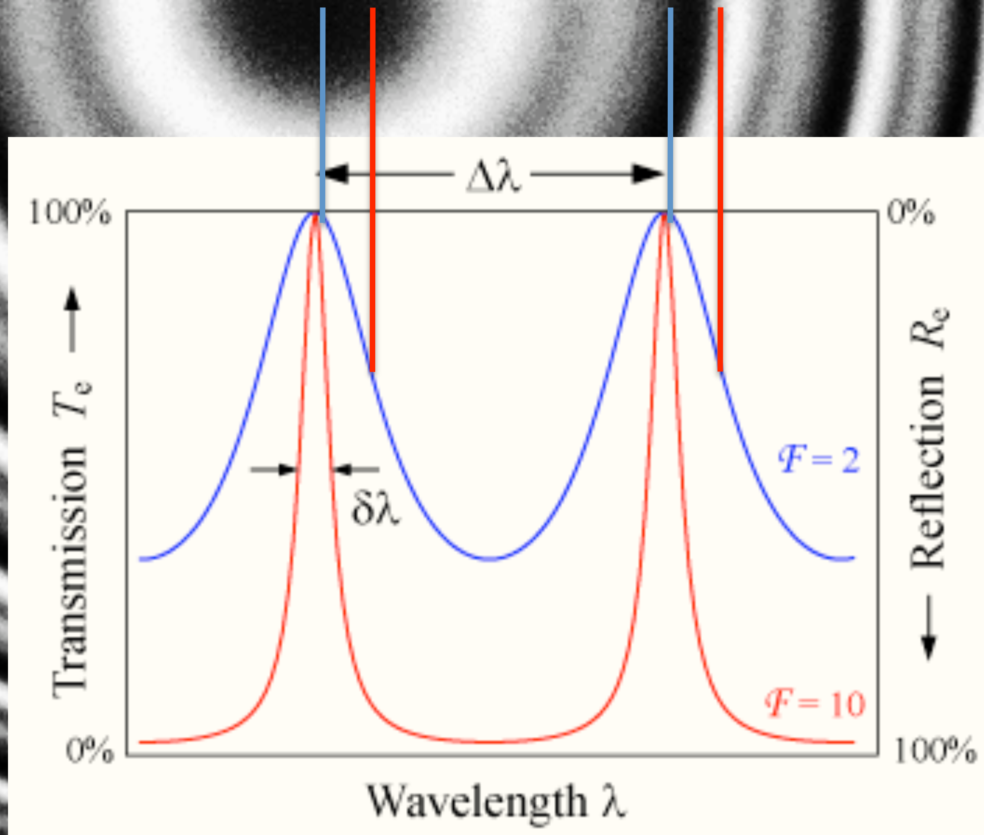
$F_P = \lambda/\Delta/2,$ parallism finesse

$F_\phi = \dfrac{4\lambda}{\phi^2 l},$ aperture finesse

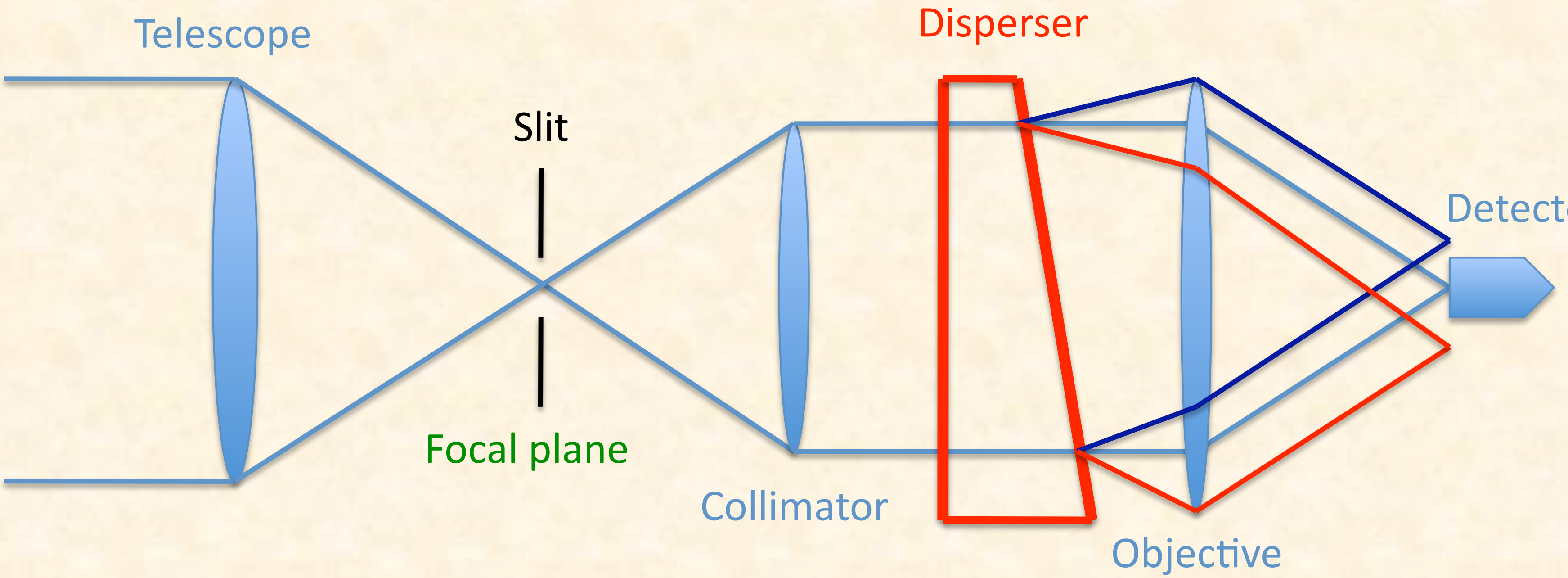
Example of apearture finesse



Transmitted wavelength is unique only for given angle
T
If slit is too wide, the aperture (angle cone) is enlarged and the range of transmitted wavelength increased.
The 'contrast is the reduced, thus the finesse and the spectral resolution



General spectrograph layout



Single detector -> monochromator (may be used with movable part to scan over wavelengths

Array detector -> spectrograph with N wavelength channels (N = number of detectors or pixels)

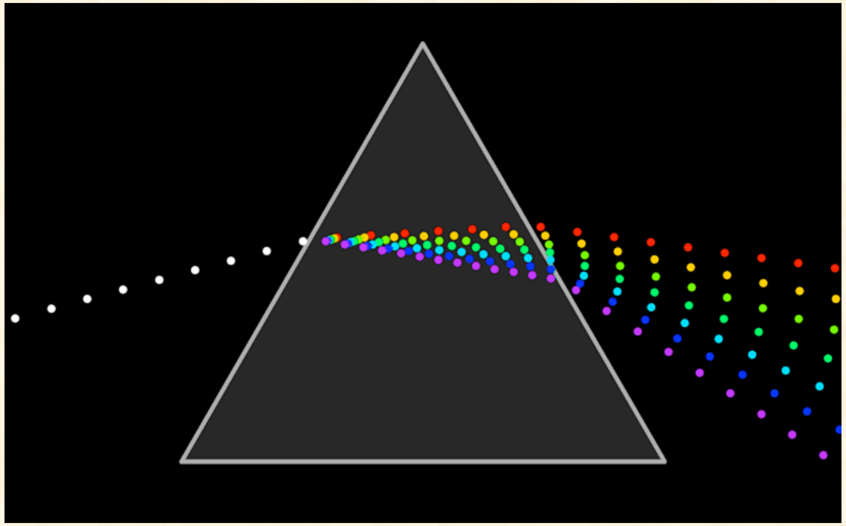
Dispersers

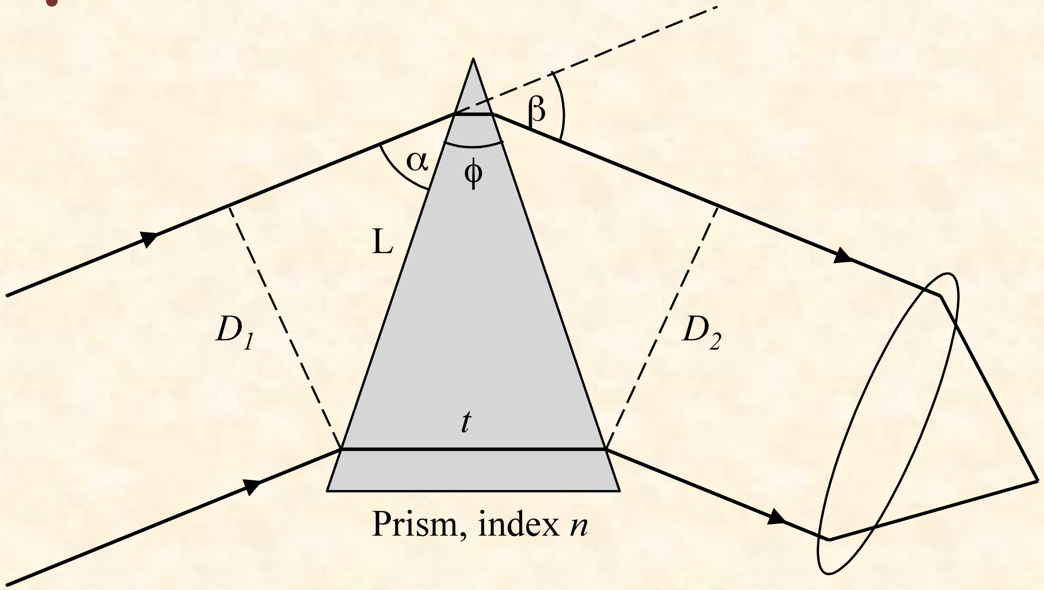
The disperser separates the wavelengths in angular direction. To avoid angular mixing, the beam is collimated. The disperser is characterized by its angular dispersion:

$$D = \frac{\partial \beta}{\partial \lambda}$$

where β is the deviation angle from the un-dispersed direction

The prism





$$\text{Minimum deviation condition: } \beta = \pi - \phi - 2\alpha$$

$$\text{Fermat priciple: } n \cdot t = 2L\cos\alpha$$

$$\Rightarrow \frac{dn}{d\beta} = -\frac{1}{2}\frac{dn}{d\alpha} = \frac{L\sin\alpha}{t} = \frac{D_1}{t}$$

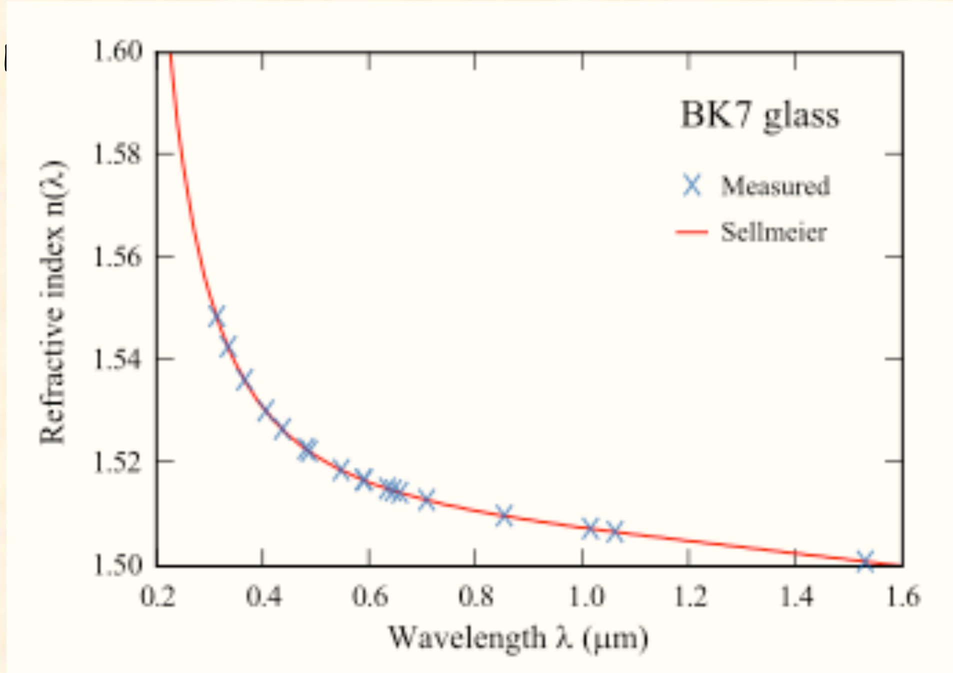
$$\Rightarrow \frac{1}{D_{prism}} = \frac{d\lambda}{d\beta} = \frac{d\lambda}{dn}\cdot\frac{dn}{d\beta} = \frac{D_1}{t}\cdot\frac{d\lambda}{dn} \quad \text{(inverse dispersion)}$$

Prism characteristics

-> High transmittance

-> When used at minimum deviation,
compression or enlargement)

-> Produces 'low' dispersion



Prism example: BK7 (normal glass), t = 50 mm, D = 100 mm

$$D_{prism} = \frac{d\beta}{d\lambda} = \frac{t}{D_1} \cdot \frac{dn}{d\lambda} \cong 0.03 \text{ rad/}\mu\text{m @ } 550 \text{ nm}$$

Prism characteristics

-> Depends mainly on glass material (internal transmittance)

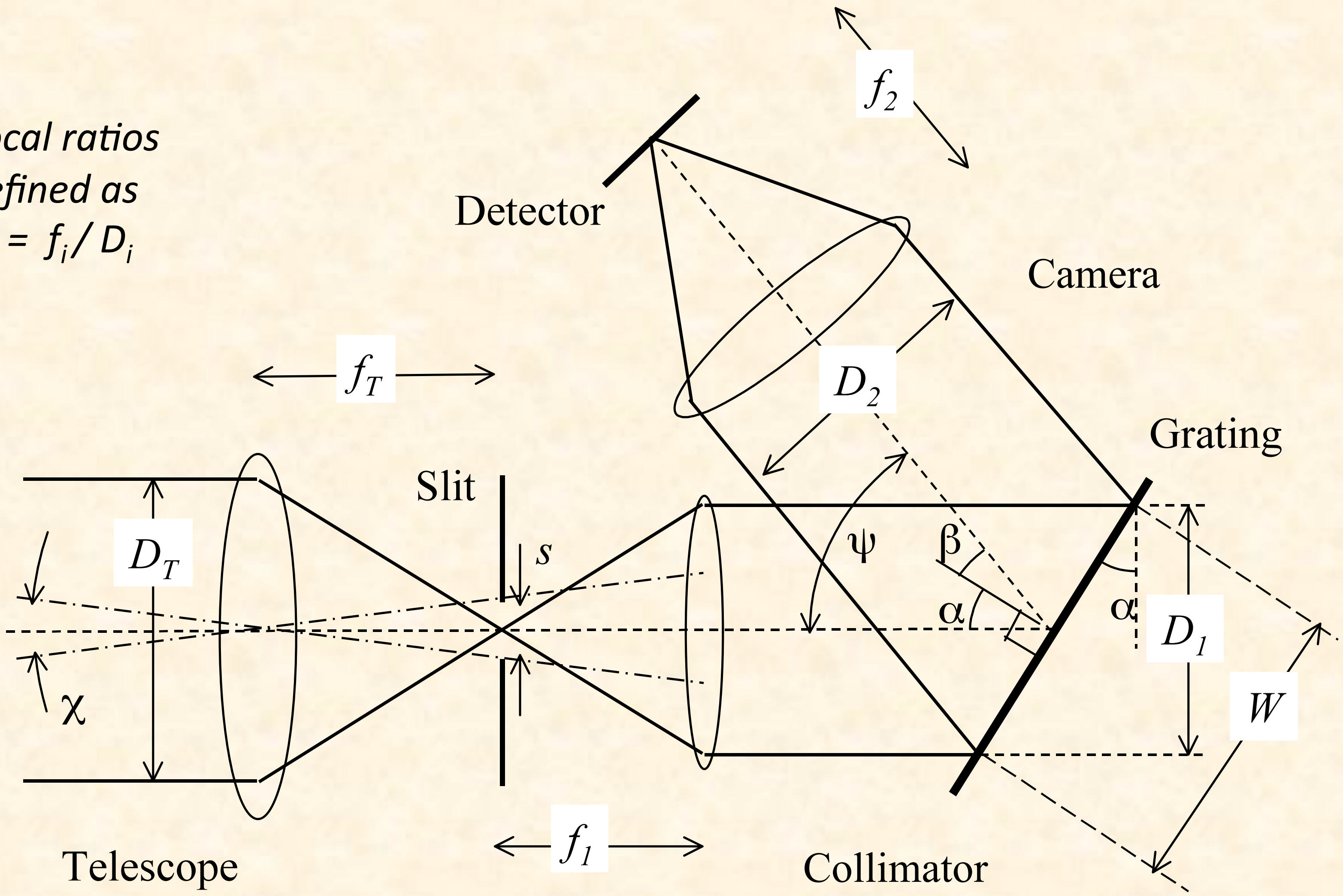
-> Anti-reflection coatings are needed to avoid reflection losses, especially for large apex angles (and large α). The coating must be optimized for the glass and the used angles.

-> Efficiency can be as high as 99%

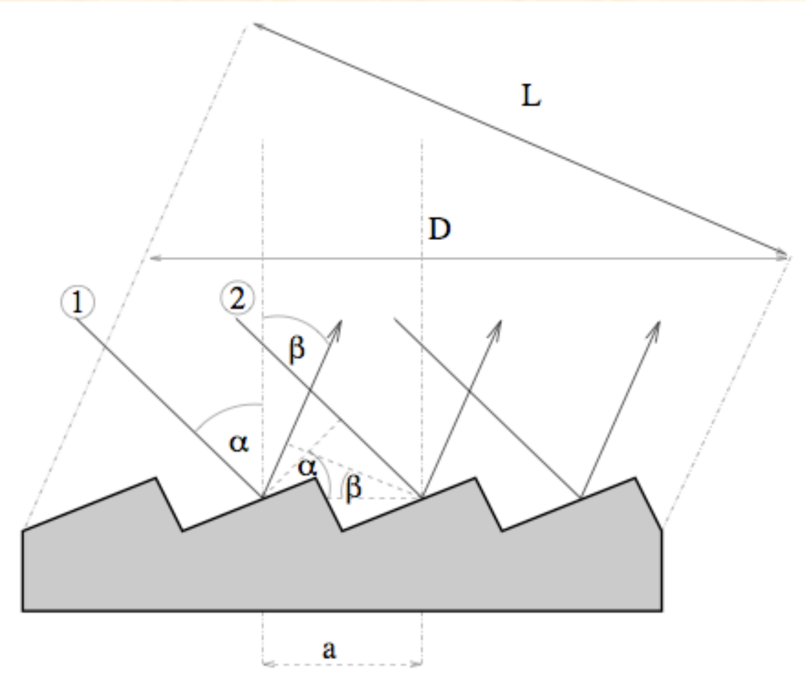
-> The dispersion increases towards the blue wavelengths. For Crown glasses (contain Potassium) the ratio of the dispersion between blue and red is lower than for Flint glasses (contain lead, titanium dioxide or zirconium dioxide) .

Grating spectrograph

Focal ratios defined as
$F_i = f_i / D_i$



The diffraction grating



Generic grating equation from the condition of positive interference between various 'grooves':

$$\boxed{\begin{array}{l} m\rho\lambda = n_1 \sin\alpha + n_2 \sin\beta \quad \text{where} \quad \rho = \dfrac{1}{a} \\ m\rho\lambda = n(\sin\alpha + \sin\beta) \quad \text{reflection grating} \end{array}}$$

$$\text{Angular dispersion:} \quad \frac{d\beta}{d\lambda} = \frac{m\rho}{\cos\beta}$$

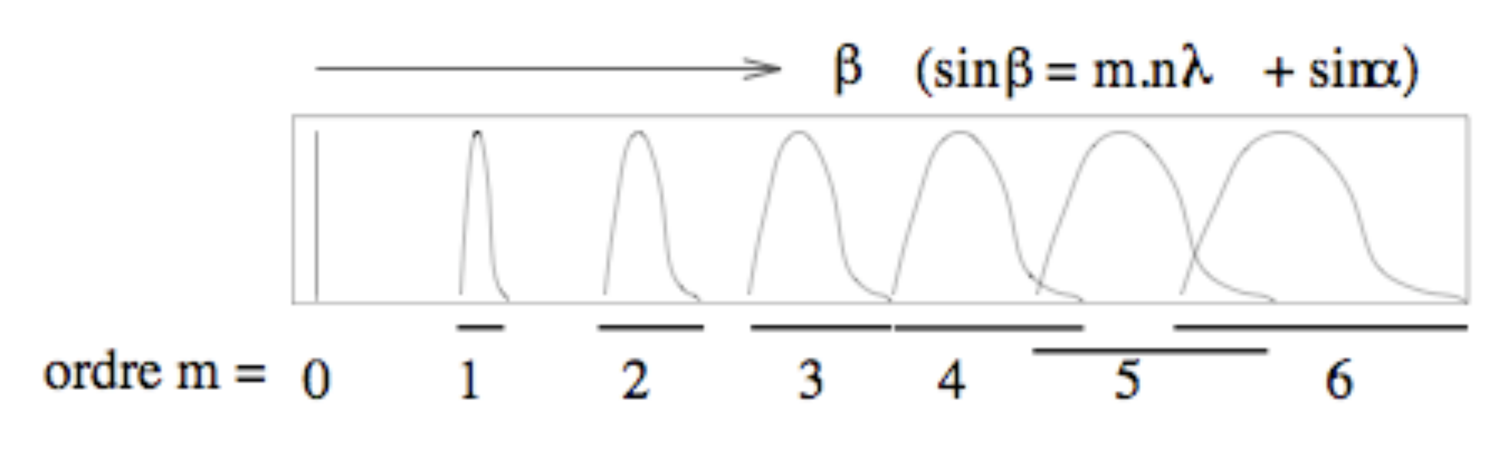
$$\text{Linear dispersion:} \quad \frac{dx}{d\lambda} = \frac{dx}{d\beta}\frac{d\beta}{d\lambda} = f_2 \frac{m\rho}{\cos\beta}$$

Grating characteristics

-> Several orders result for a given wavelength

-> m = 0 for a grating which acts like a mirror -> no dispersion!

-> Orders overlap spatially -> must be filtered or use at m=1

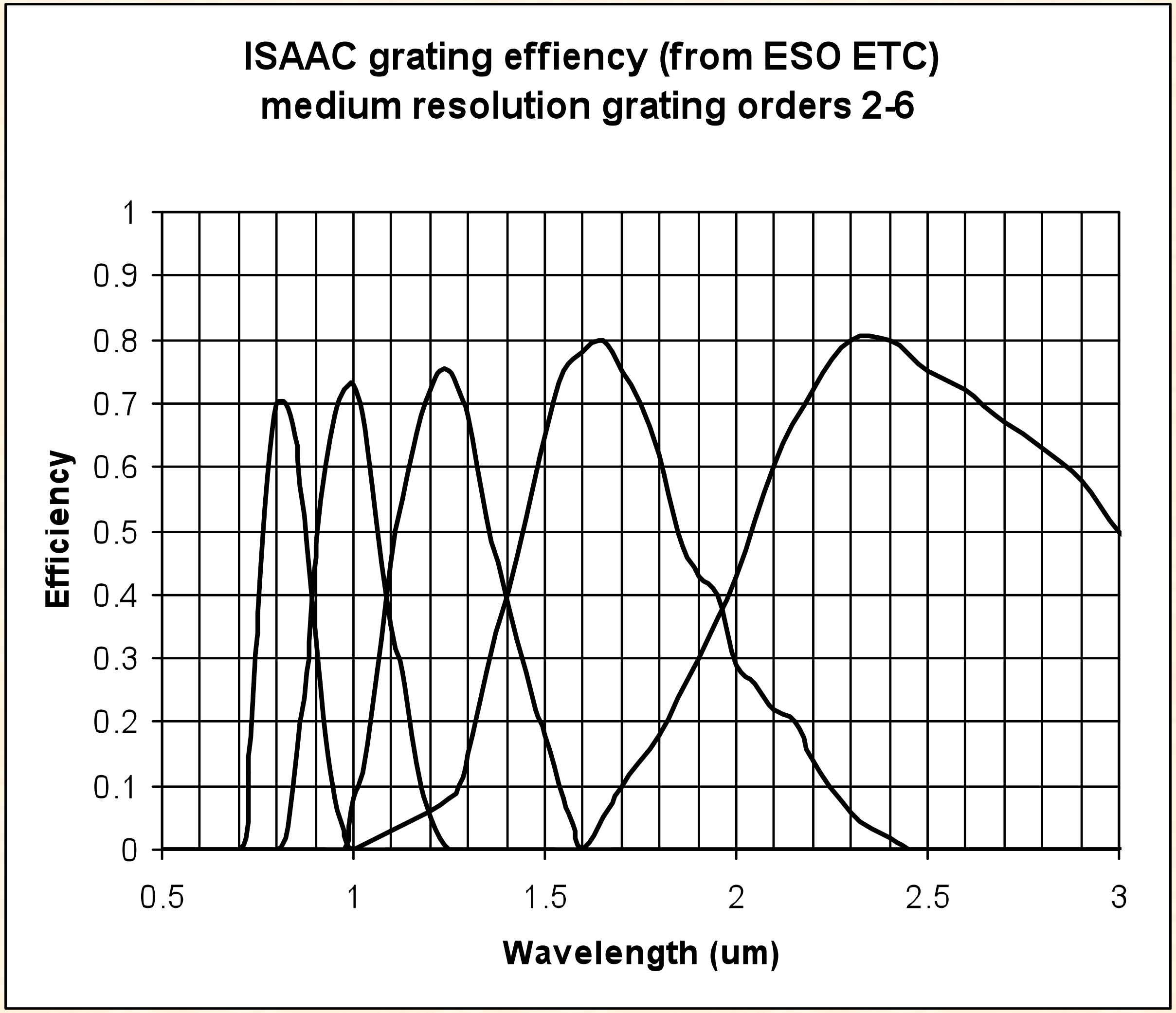


Typical grating example: m = 1, ρ = 1000 gr/mm, sinα + sinβ = 1, cosβ=1/2

$$\Rightarrow \frac{d\beta}{d\lambda} = \frac{m\rho}{\cos\beta} = 2 \text{ rad}/\mu\text{m}$$

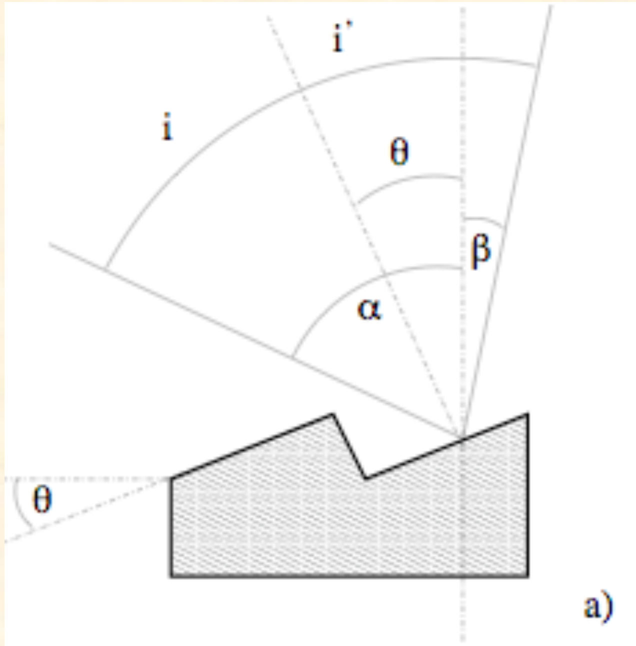
Dispersion typically much higher than for prisms!

Grism efficiency

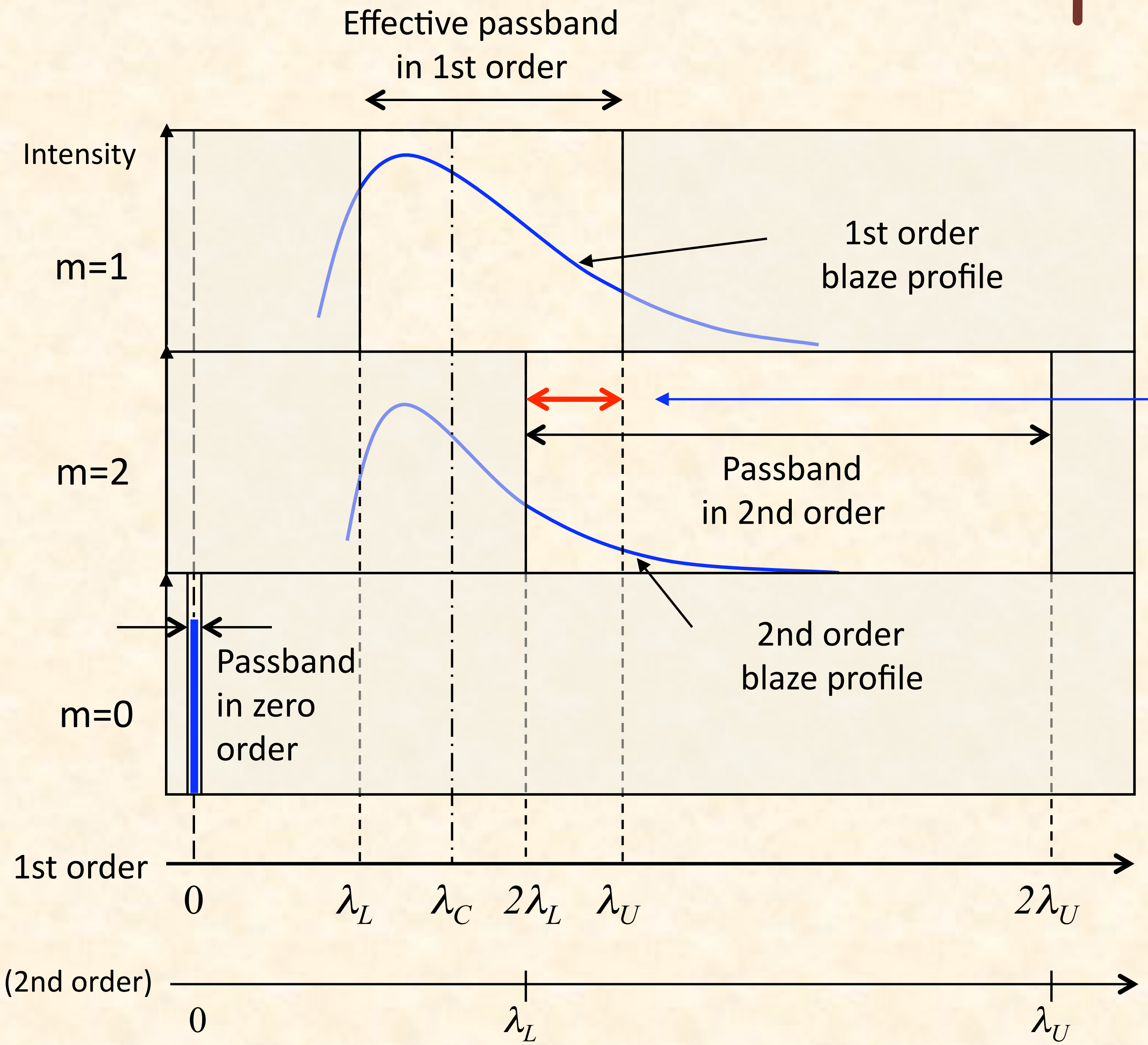


Maximum efficiency obtained when specular groove reflection is matches (Blaze condition):

$$\alpha + \beta = 2\Theta$$



Order overlaps



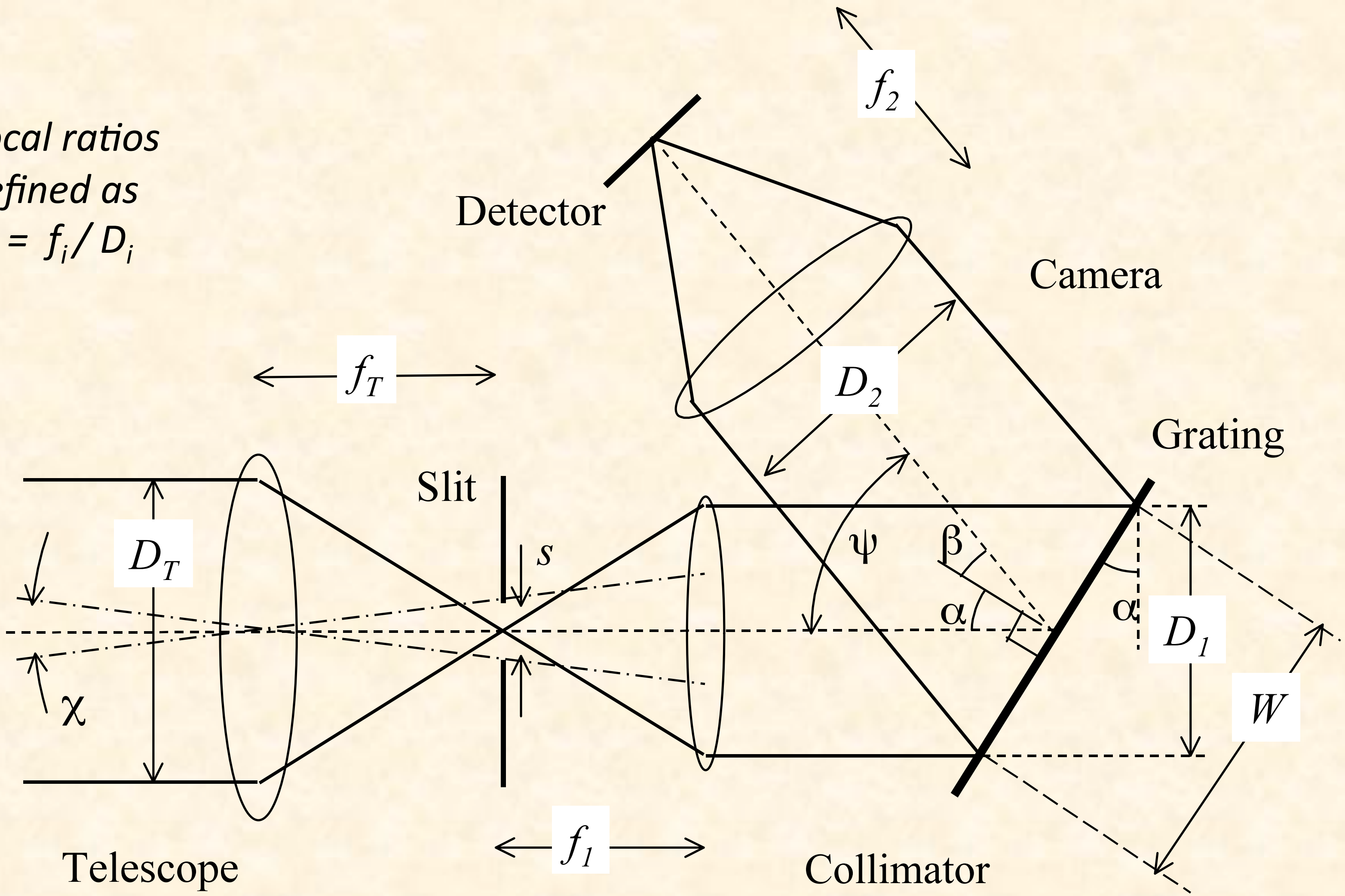
Don't forget higher orders!

Zero order matters for MOS

Wavelength in first order marking position on detector in dispersion direction (if dispersion ~linear)

Resolving power and resolution

Focal ratios defined as
$F_i = f_i / D_i$



Resolving power and resolution

The objective translates angles into positions on the detector. Each position (pixel) of the detector 'sees' a given angle of the parallel (collimated) beam

The collimated beam is never perfectly parallel, because either of the limited diameter of the beam, which produces diffraction $\delta\phi = 1.22\ \lambda/D_1$, or because of the finite slit, which produces and angular divergence $\delta\Theta = s/f_1$.

The angular divergence is translated into a distance $\delta\lambda = f_2\ \delta\Theta$ or $\delta\lambda = f_2\ \delta\phi$ on the CCD. This means that over this distances the wavelengths are mixed (cannot be separated angularly.

Resolving power and resolution

Resolving power is the maximum spectral resolution which can be reached if the slit s = 0 and the angular divergence is limited by diffraction arising from the limited beam diameter. For a given Dispersion D we get the resolving power:

$$RP := \frac{\lambda}{\delta\lambda} = \frac{\lambda}{\delta\Phi / D} = \frac{\lambda}{\delta\Phi \cdot \dfrac{d\lambda}{d\beta}} = \frac{\lambda}{\delta\Phi} \cdot \frac{d\beta}{d\lambda}$$

Spectral resolution is the effective spectral resolution which is finally reached when assuming a finite slit s. For a given Dispersion D we get the spectral resolution:

$$R := \frac{\lambda}{\delta\lambda} = \frac{\lambda}{\delta\Theta / D} = \frac{\lambda}{\delta\Theta \cdot \dfrac{d\lambda}{d\beta}} = \frac{\lambda}{\delta\Theta} \cdot \frac{d\beta}{d\lambda}$$

Conservation of the ‘étendue’

The étendue is defined as E=A x O, where A is the area of the beam at a given optical surface and O is the solid angle under which the beam passes through the surface.

When following the optical path of the beam through an optical system, E is constant, in particular, it cannot be reduced

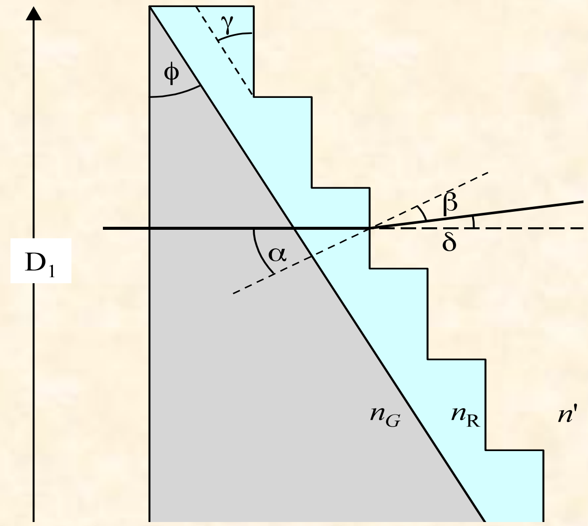
For a telescope, E is the product of the primary mirror surface and the two-dimensional field (in sterad) transmitted by the optical system. Normally, the transmitted field is defines a slit width. When entering spectrograph, the slit x beam aperture at the slit is equal to the etendue E of the telescope. This implies that at fixed spectral resolution, the slit width and the beam diameter cannot be chosen independently, since dΘ depends on both.

Other dispersers

- Grisms
- VPHG
- Echelle grating

Grisms

- Transmission grating attached to prism
- Allows in-line optical train:
  - simpler to engineer
  - quasi-Littrow configuration - no variable anamorphism
- Inefficient for ρ > 600/mm due to groove shadowing and other effects



Grism equations

- Modified grating equation: $$m\rho\lambda = n\sin\alpha + n'\sin\beta$$
- Undeviated condition: $$m\rho\lambda_U = (n-1)\sin\phi$$

  $n' = 1,\ \beta = -\alpha = \phi$

  θ = phase difference from centre of one ruling to its edge
- Blaze condition: $$\theta = 0 \Rightarrow \lambda_B = \lambda_U$$
- Resolving power *(same procedure as for grating)*

$$R = \frac{m\rho\lambda W}{\chi D_T}$$

$$W = D_1 / \cos\phi$$

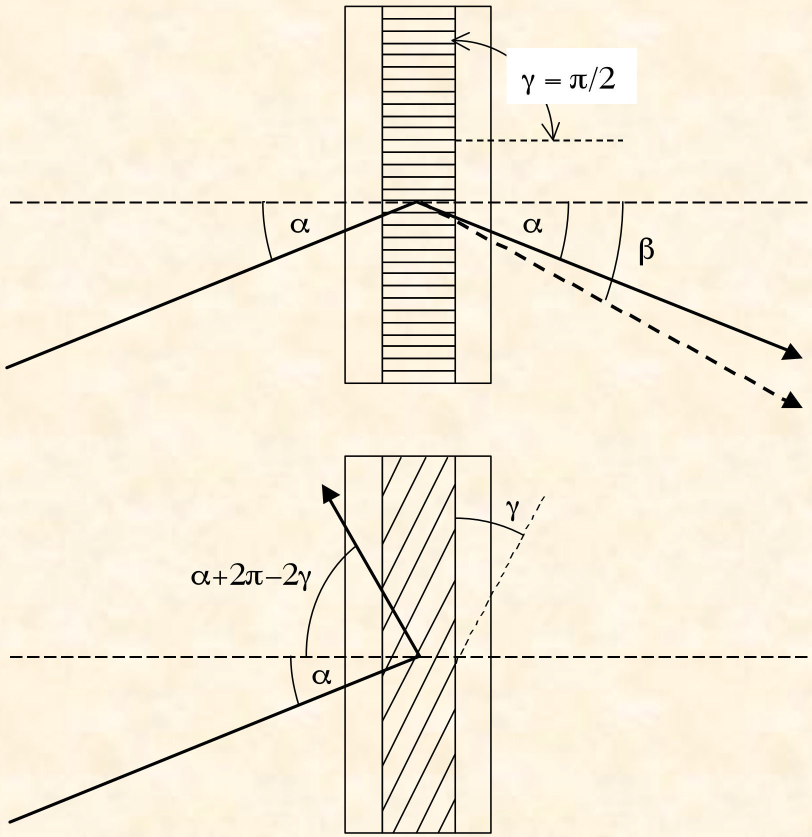
$$R = \frac{(n-1)\tan\phi D_1}{\chi D_T}$$

Volume Phase Holographic gratings

- So far we have considered *surface relief* gratings
- An alternative is *VPH* in which refractive index varies harmonically throughout the body of the grating:
- Don't confuse with '*holographic*' gratings (SR)
- Advantages: $n_g(x,z) = n_g + \Delta n_g \cos\left\lfloor 2\pi\rho_g\left(x\sin\gamma + z\cos\gamma\right)\right\rfloor$
  - Higher peak efficiency than SR
  - Possibility of very large size with high ρ
  - Blaze condition can be altered *(tuned)*
  - Encapsulation in flat glass makes more robust
- Disadvantages
  - Tuning of blaze requires *bendable spectrograph!*
  - Issues of wavefront errors and cryogenic use

VPH configurations

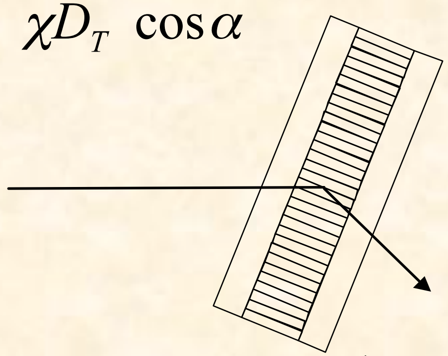
- *Fringes* = planes of constant *n*
- Body of grating made from *Dichromated Gelatine* (DCG) which permanently adopts fringe pattern generated holographically
- Fringe orientation allows operation in transmission or reflection

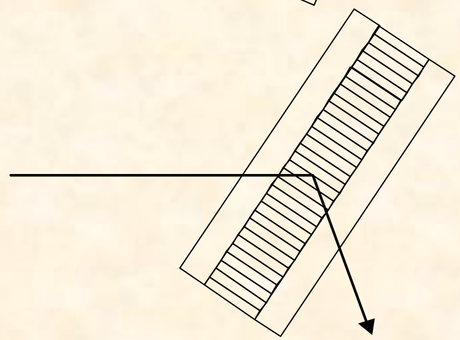


VPH equations

- Modified grating equation: $m\rho\lambda = \sin\alpha + \sin\beta$
- Blaze condition: $m\rho\lambda_B = 2n_g \sin\alpha_g = 2\sin\alpha$

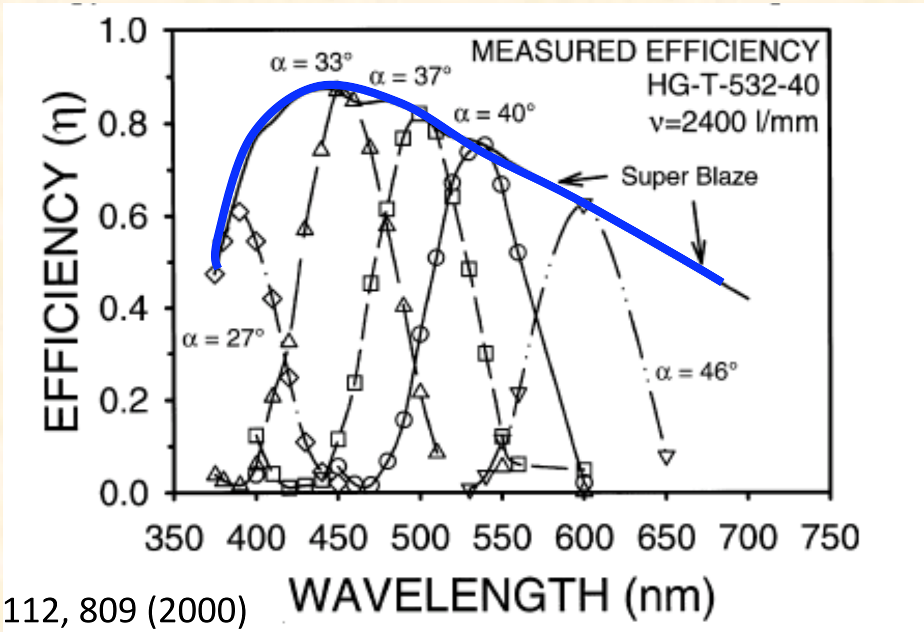
  = *Bragg* diffraction $n_g \sin\alpha_g = \sin\alpha$
- Resolving power: $R = \dfrac{m\rho\lambda W}{\chi D_T} = \dfrac{m\rho\lambda}{\chi D_T}\dfrac{D_1}{\cos\alpha}$
- Tune blaze condition by tilting grating (α)
- Collimator-camera angle must also change by 2α $\Rightarrow$ mechanical complexity





VPH efficiency

- Kogelnik's analysis when: $\dfrac{2\pi\lambda d\rho_g^2}{n_g}. > 10$
- Bragg condition when: $\Delta n_g d. \approx \dfrac{\lambda}{2}$
- Bragg envelopes (efficiency FWHM):
  - in wavelength: $$\Delta\lambda \propto \left(\frac{1}{\rho_g \tan\alpha_g}\right)\Delta n_g = \left(\frac{1}{\rho_g \tan\alpha_g}\right)\frac{\lambda}{d}$$
  - in angle: $$\Delta\alpha \propto \frac{1}{\rho_g d}$$
- Broad blaze requires
  - thin DCG
  - large index amplitude
- *Superblaze*



Barden et al. PASP 112, 809 (2000)

Other dispersers

- Grisms
- VPHG
- Echelle grating …