Data extraction from high-resolution spectra

Marina Lafarga Magro (she/her)

marina.lafarga-magro@warwick.ac.uk

WARWICK
THE UNIVERSITY OF WARWICK

Multiwavelength Approach to Exoplanetary Systems
ICE Summer School 2024

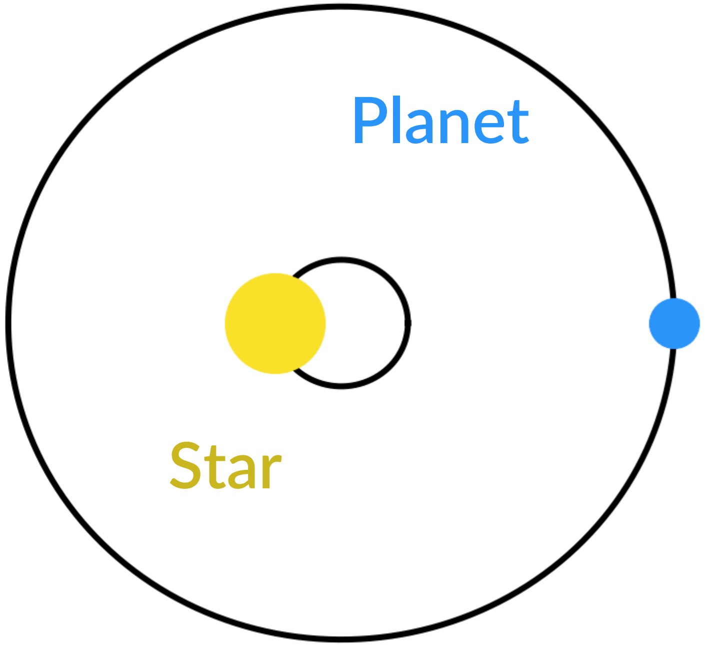
Doppler spectroscopy

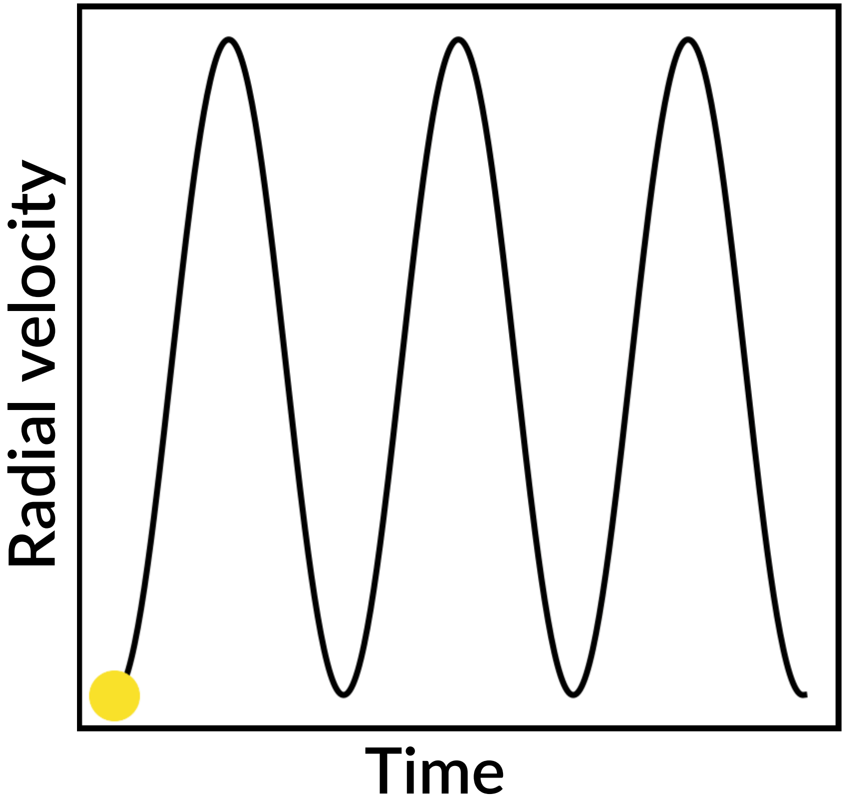
Doppler shift

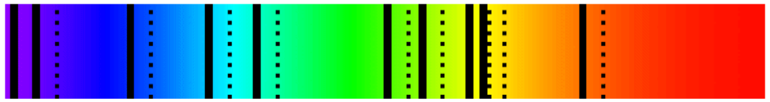
$$\mathrm{RV} = \mathrm{c}\left(1 - \frac{\lambda'}{\lambda}\right)$$

Stellar motion semi-amplitude

$$\mathrm{K} = \left(\frac{2\pi \mathrm{G}}{\mathrm{P}}\right)^{\frac{1}{3}} \frac{M_p \sin i}{\left(M_s + M_p\right)^{\frac{2}{3}}} \left(1 - e^2\right)^{-\frac{1}{2}}$$

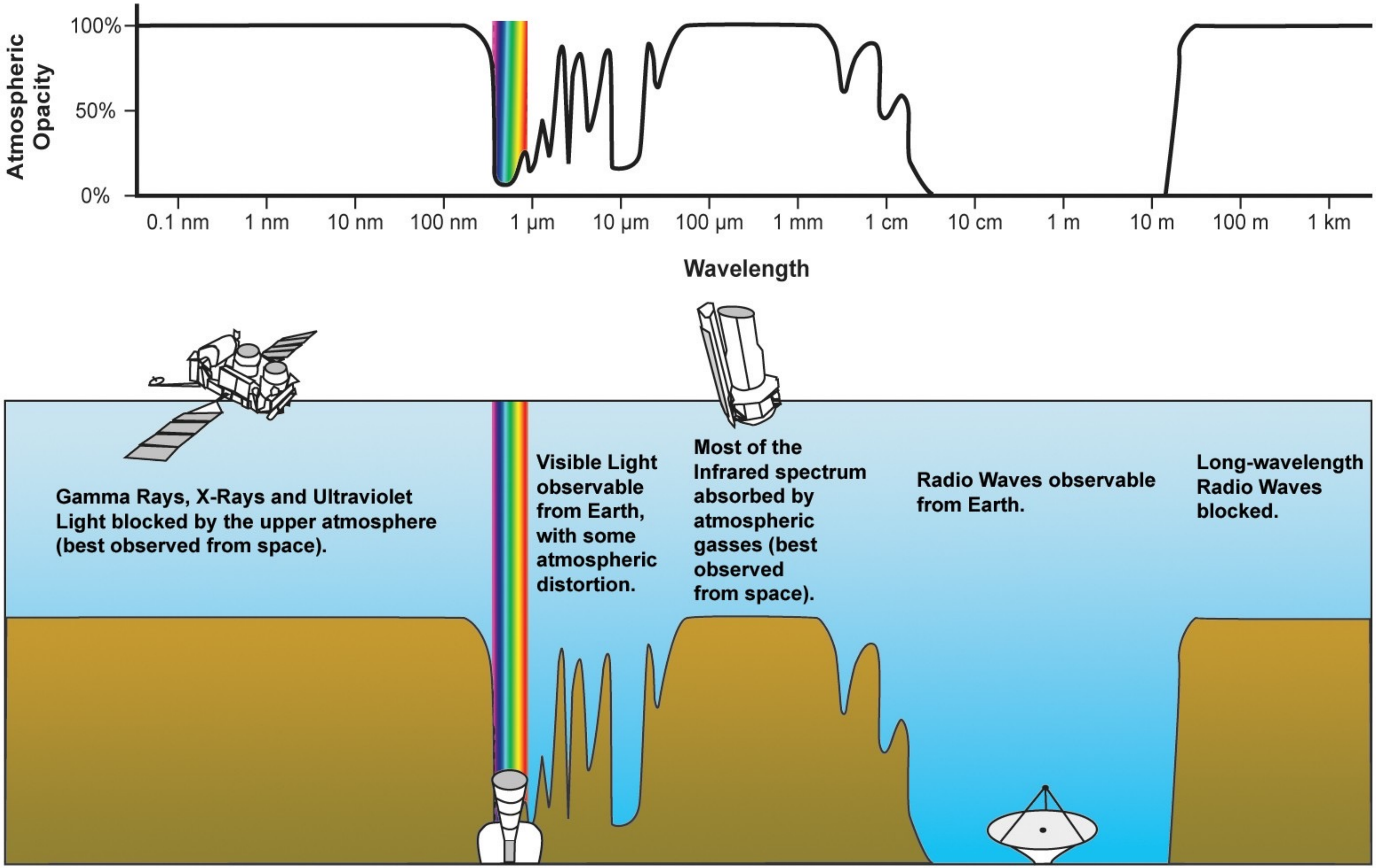






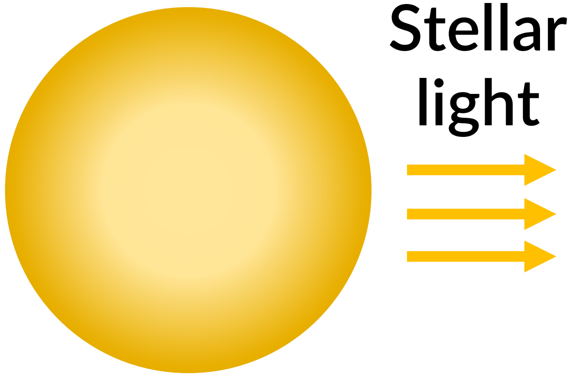
Observed stellar spectra

Image credit: Alysa Obertas

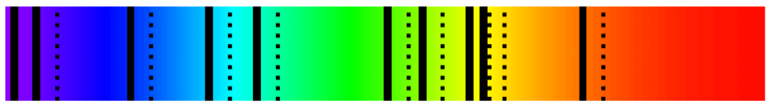

Atmospheric Opacity
100%
50%
0%
0.1 nm
1 nm
10 nm
100 nm
1 µm
10 µm
100 µm
1 mm
1 cm
10 cm
1 m
10 m
100 m
1 km
Wavelength
Gamma Rays, X-Rays and Ultraviolet Light blocked by the upper atmosphere (best observed from space).
Visible Light observable from Earth, with some atmospheric distortion.
Most of the Infrared spectrum absorbed by atmospheric gasses (best observed from space).
Radio Waves observable from Earth.
Long-wavelength Radio Waves blocked.

NASA

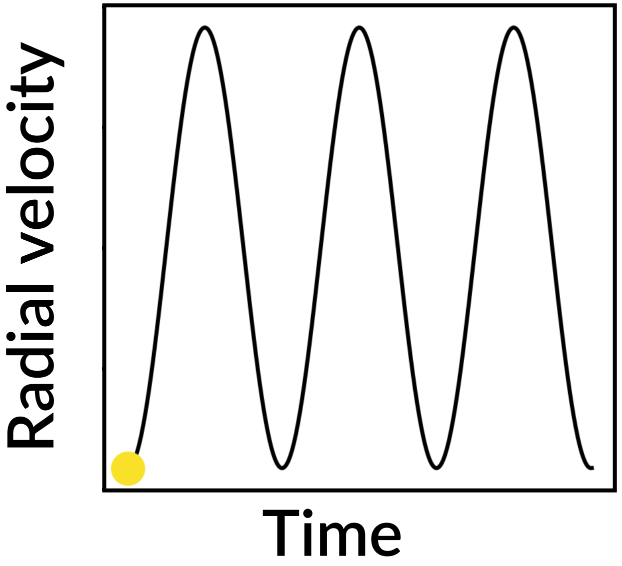
Outline



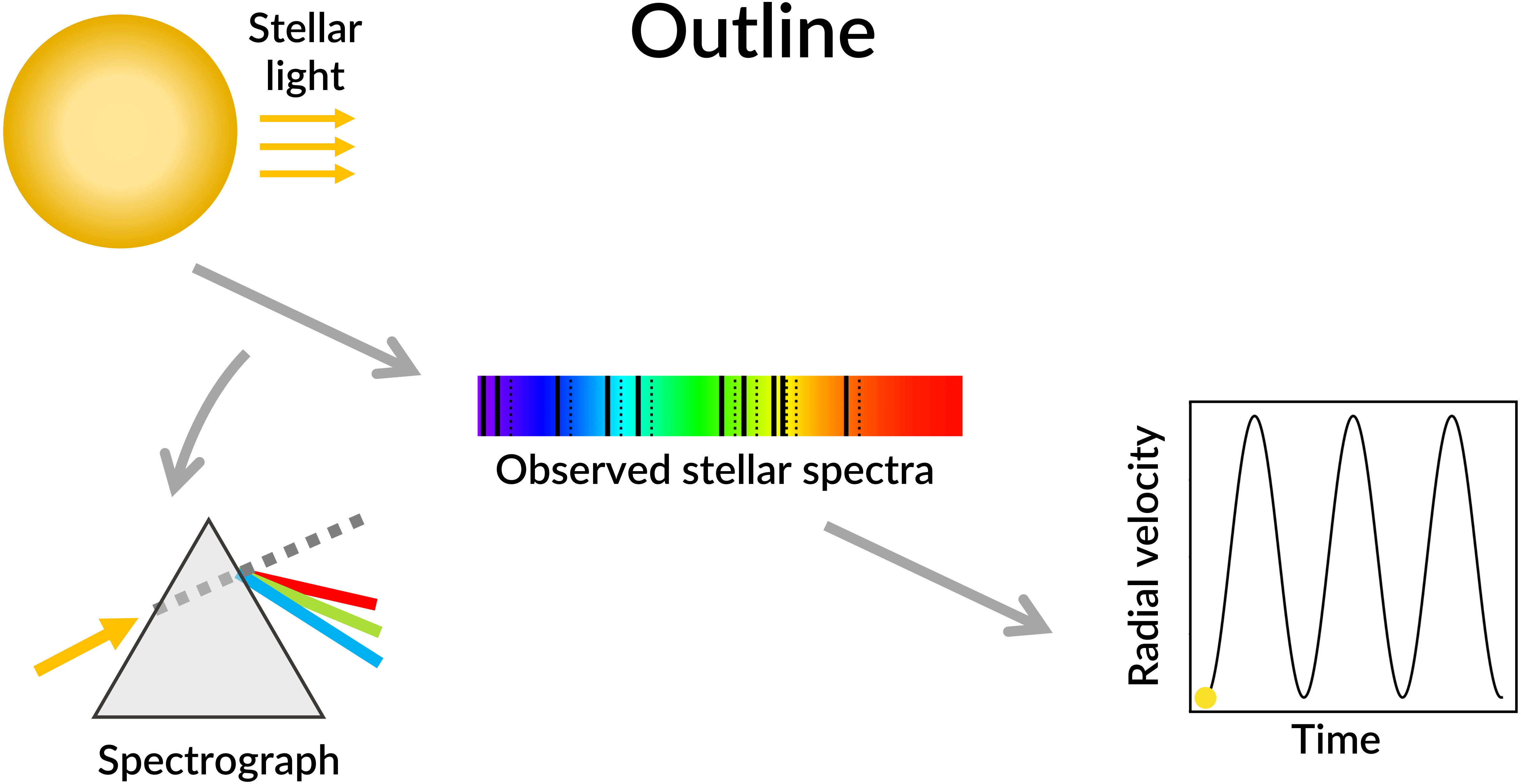
Stellar light



Observed stellar spectra



Outline



Stellar light

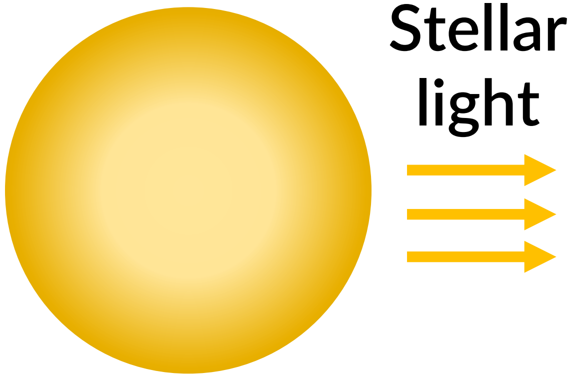
Spectrograph

Observed stellar spectra

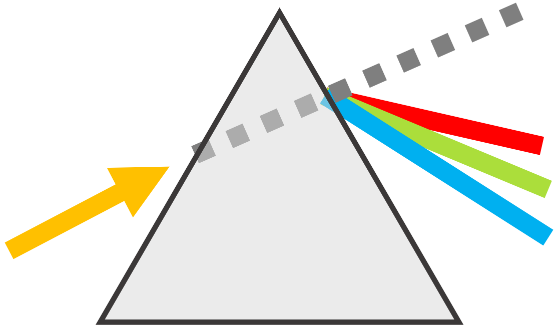
Radial velocity

Time

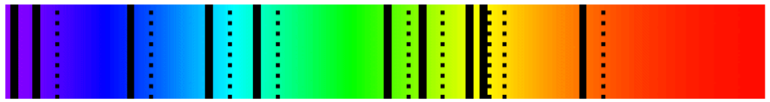
Outline



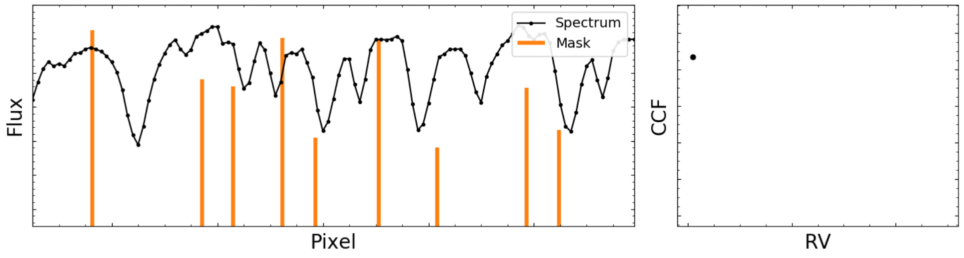
Stellar light



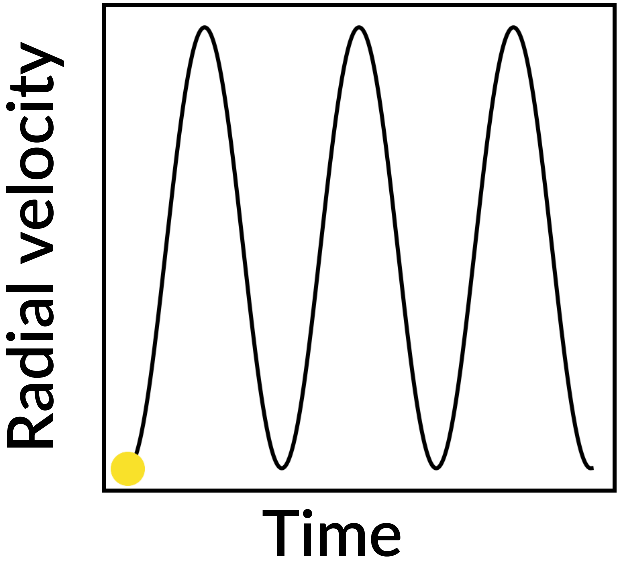
Spectrograph

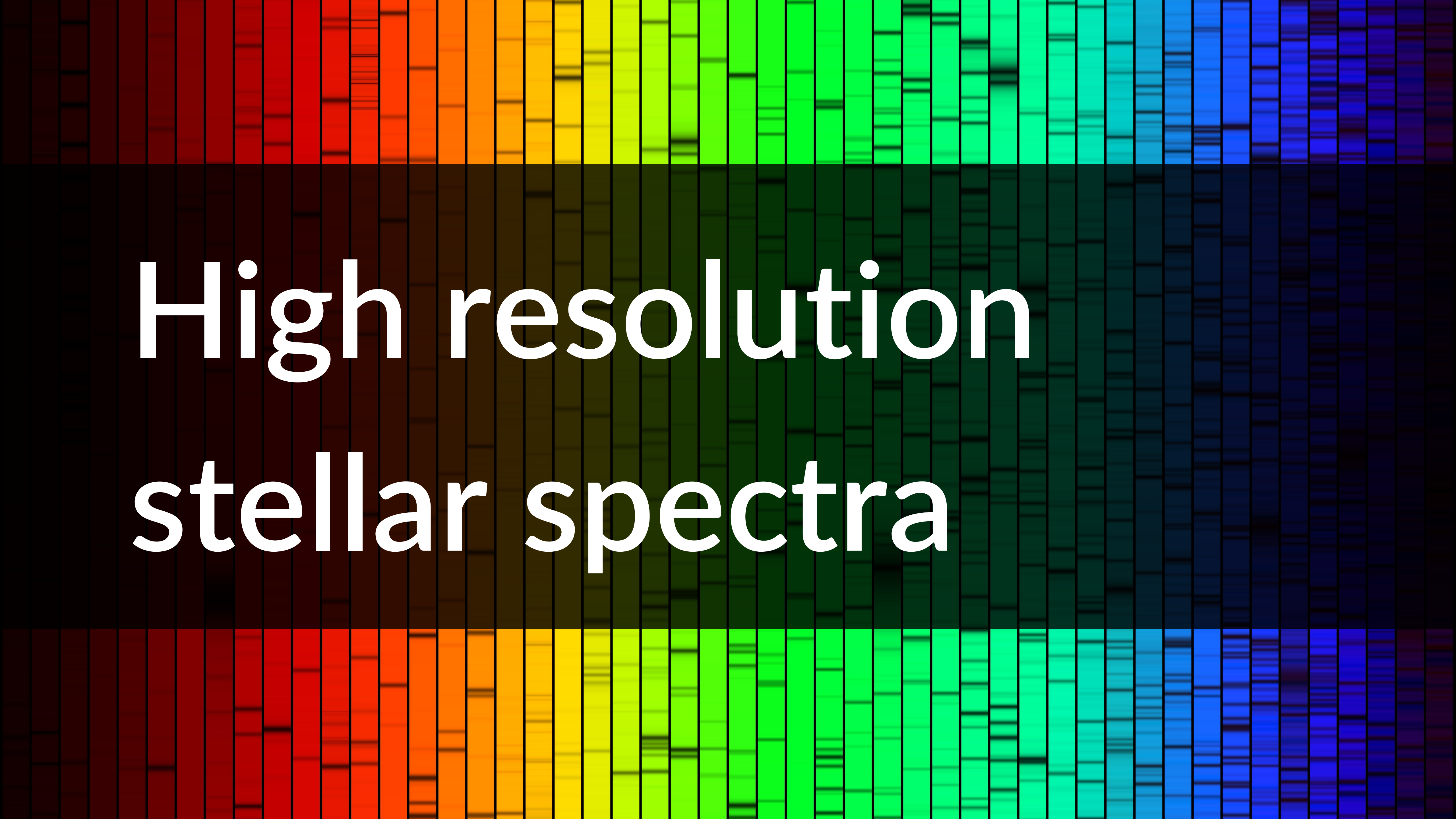


Observed stellar spectra



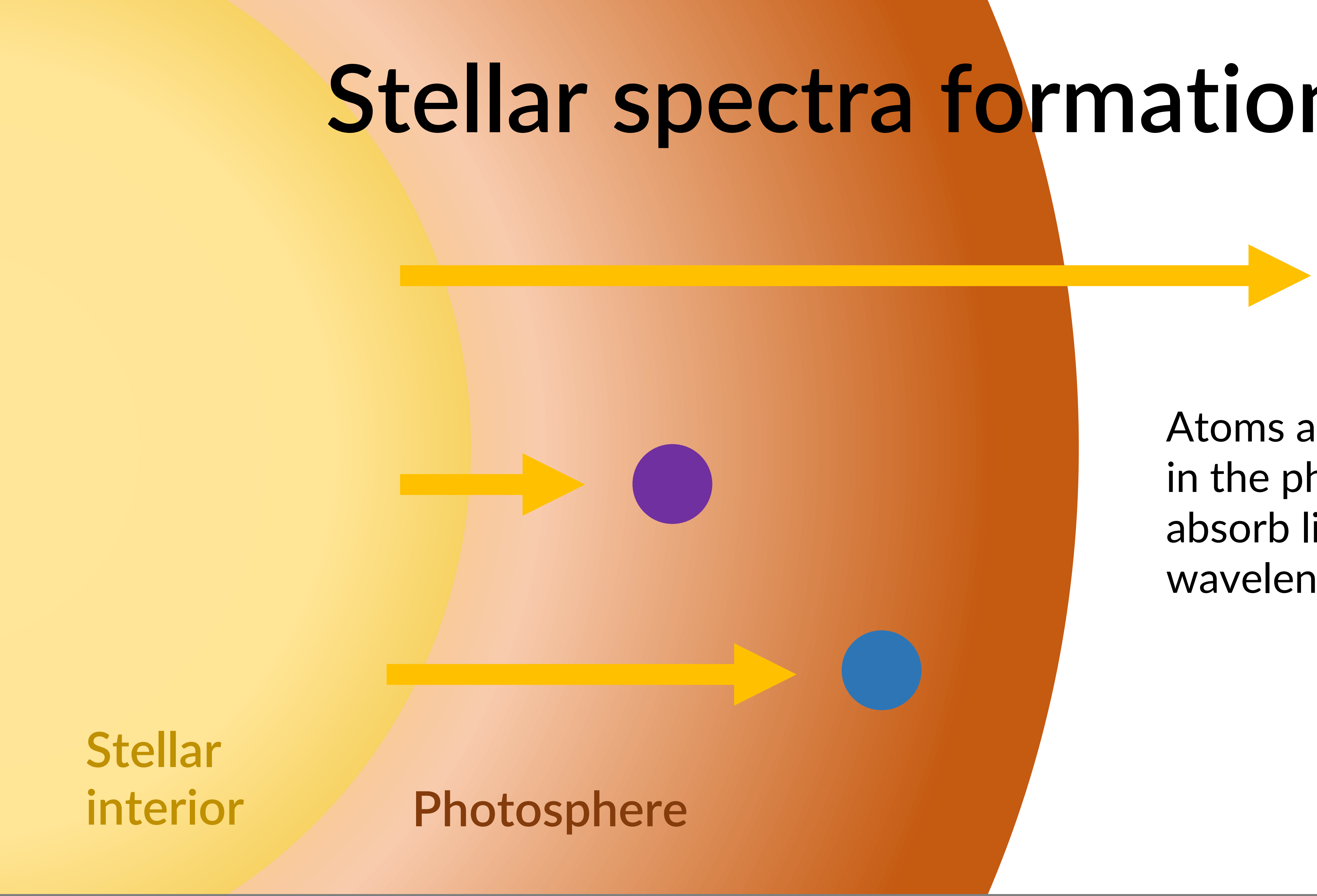
Radial velocity extraction methods





High resolution stellar spectra

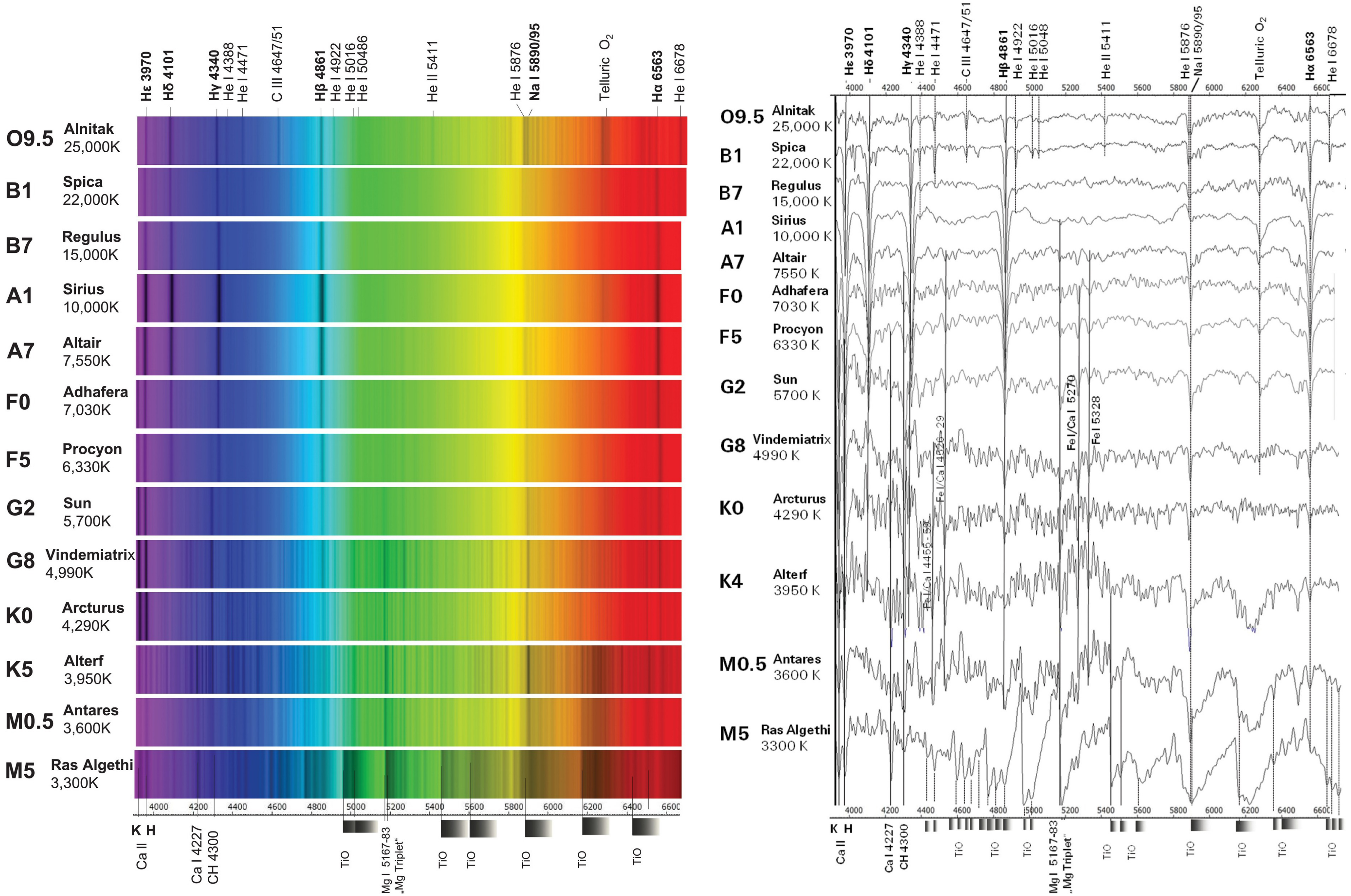
Stellar spectra formation



Atoms and molecules in the photosphere absorb light at specific wavelengths

Stellar spectra

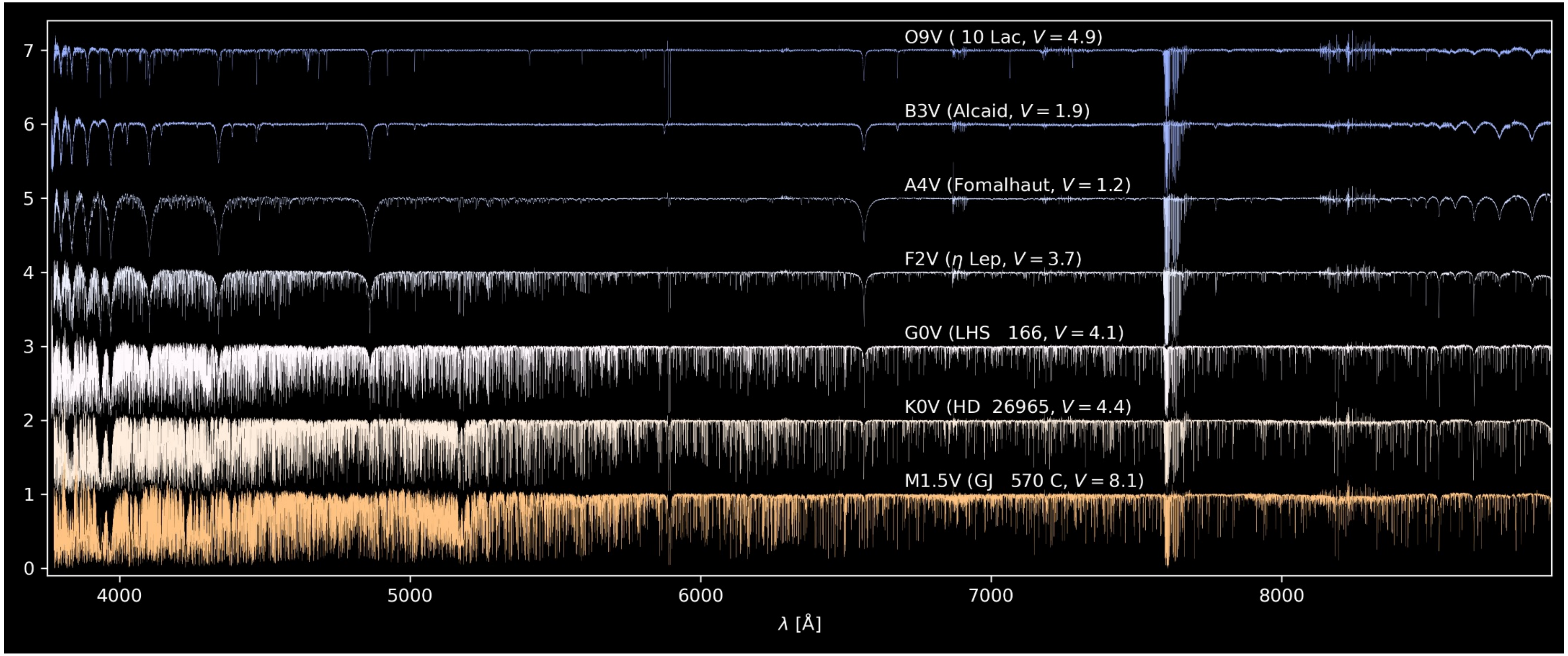
Different absorption lines for different stellar types

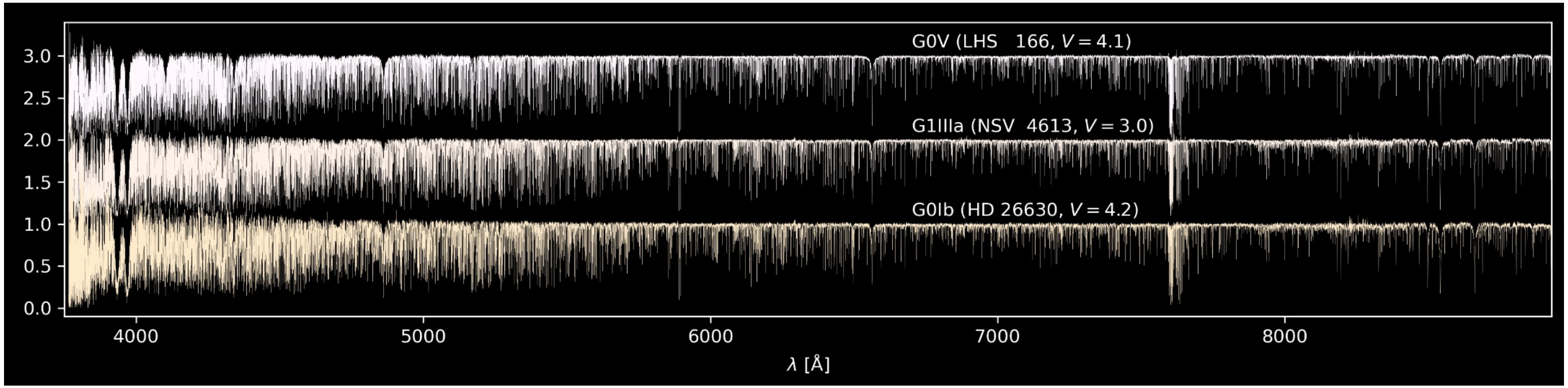


Walker 2017

Stellar spectra

Different absorption lines for different stellar types

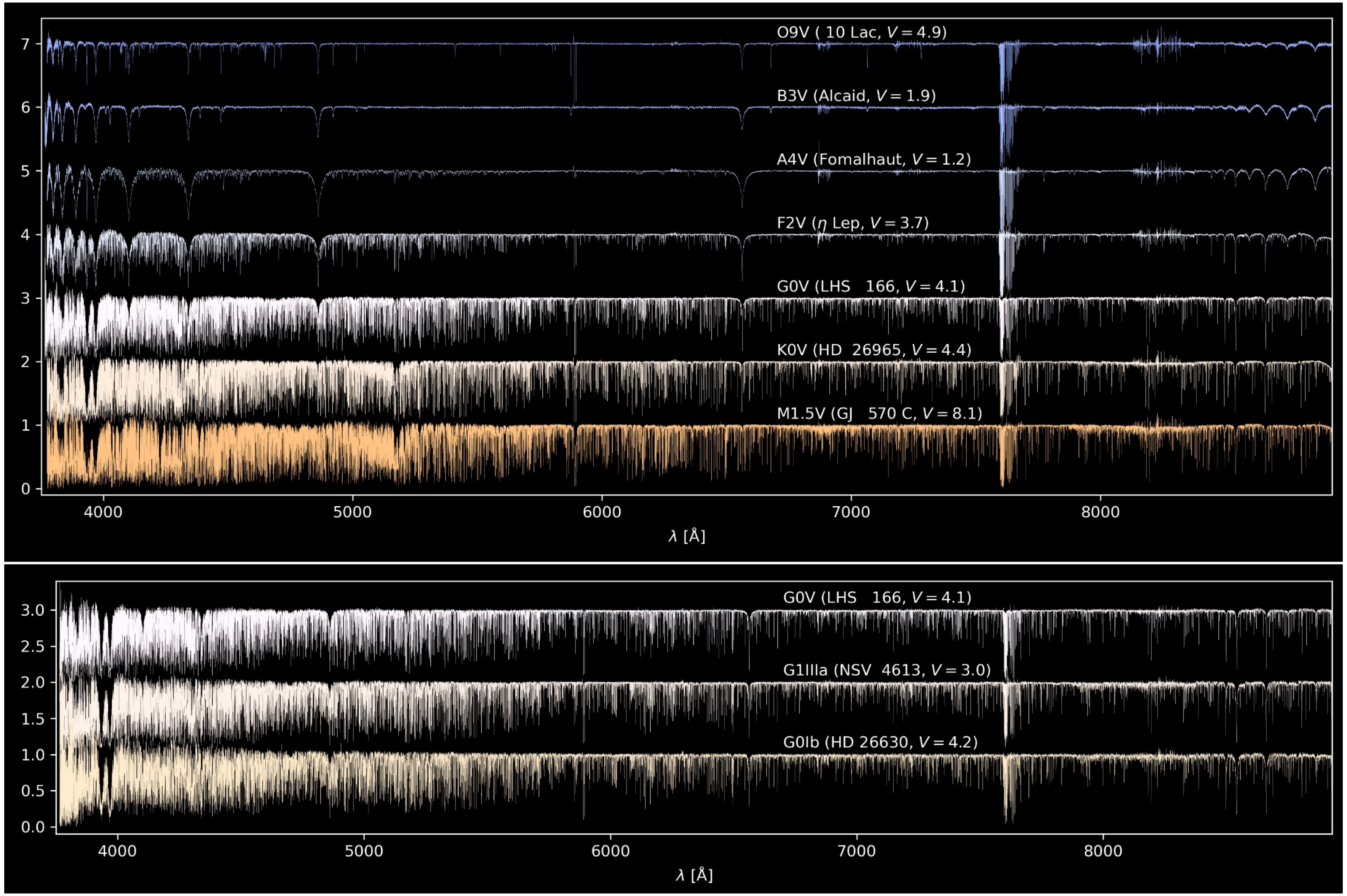




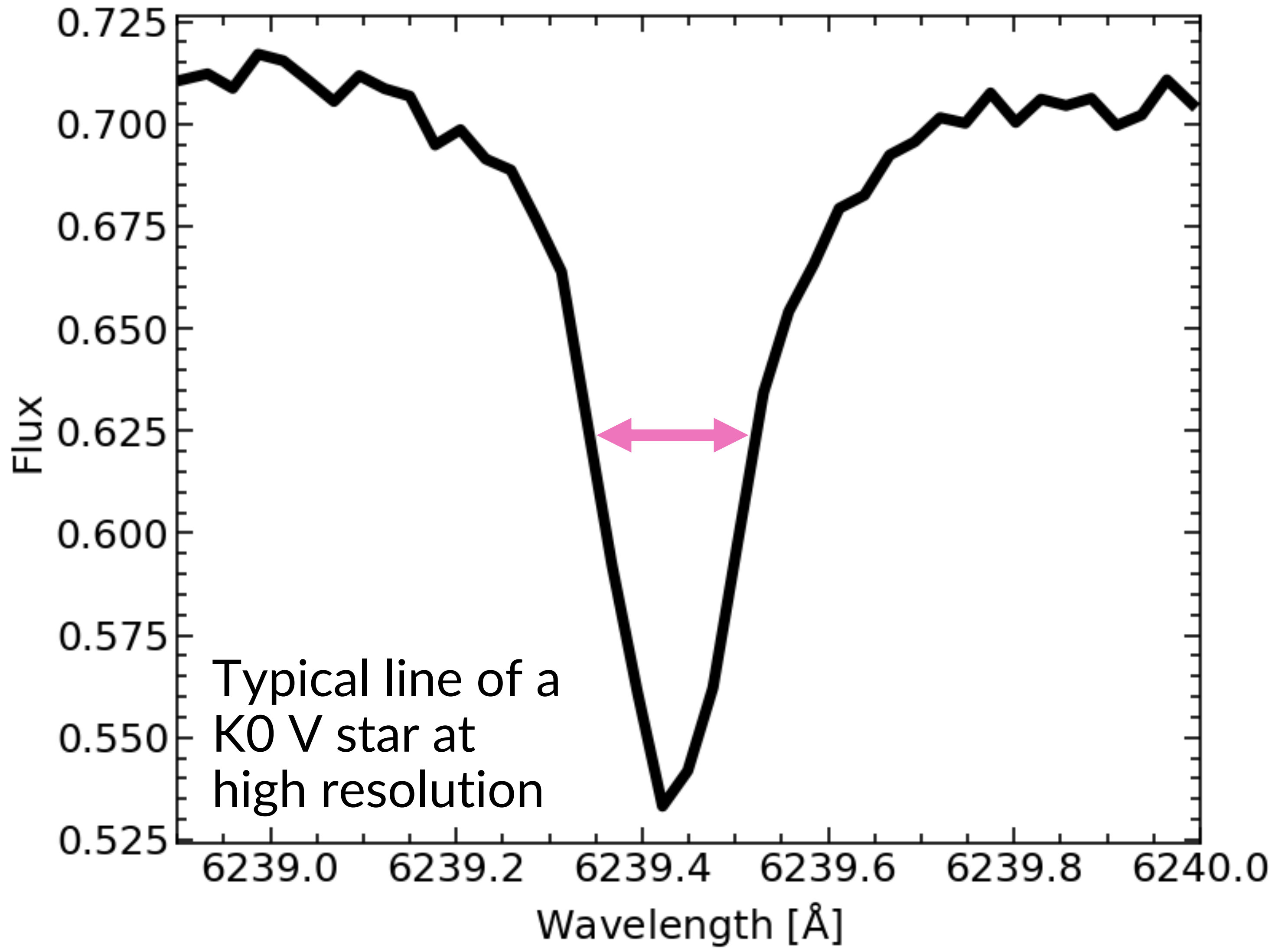
Royer et al. 2024

There is a wealth of information on a stellar spectrum

- Effective temperature
- Surface gravity
- Metallicity, chemical abundances
- Magnetic field
- Radial velocity
- Rotational velocity
- Stellar variability

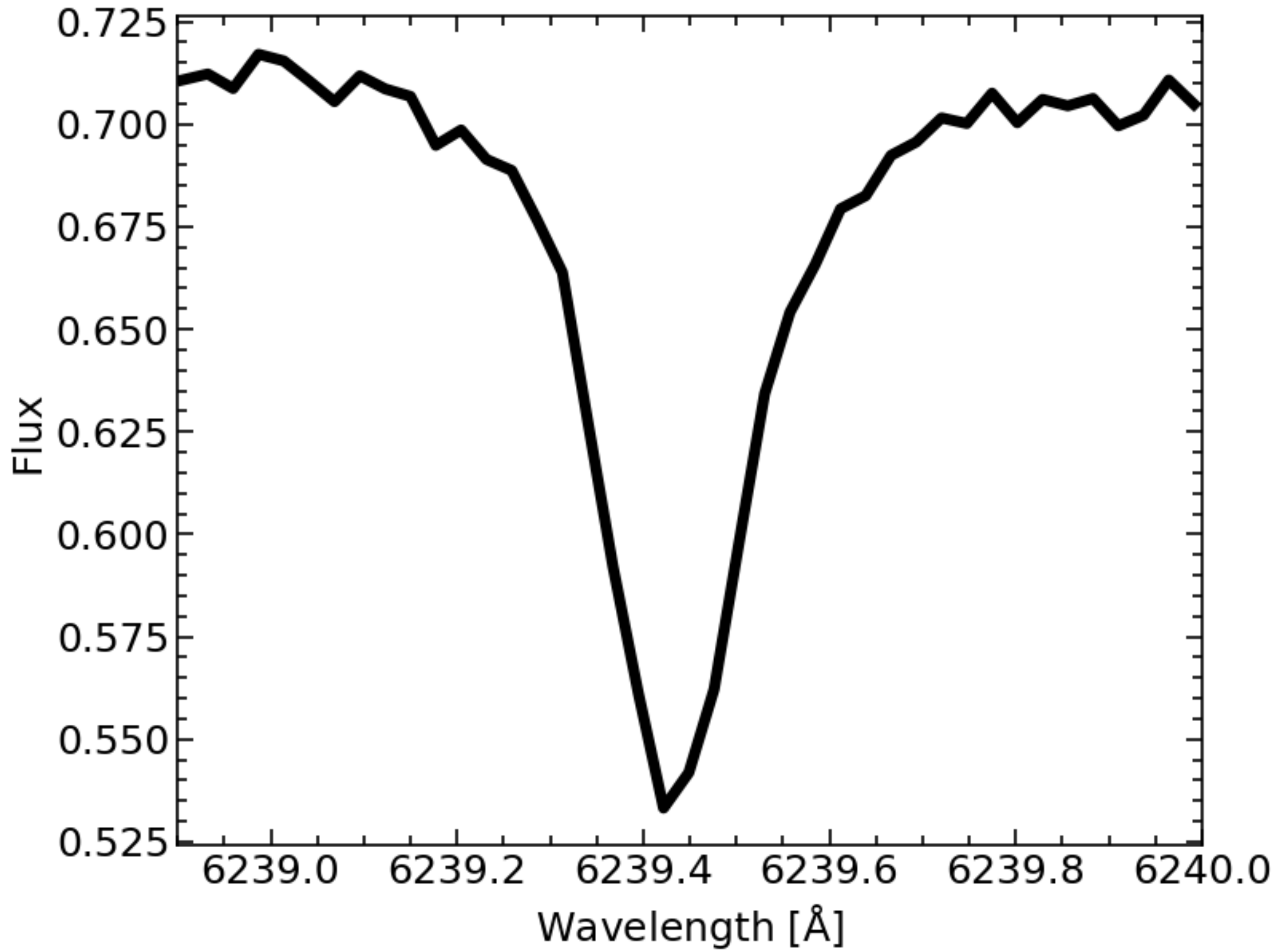


Radial velocity

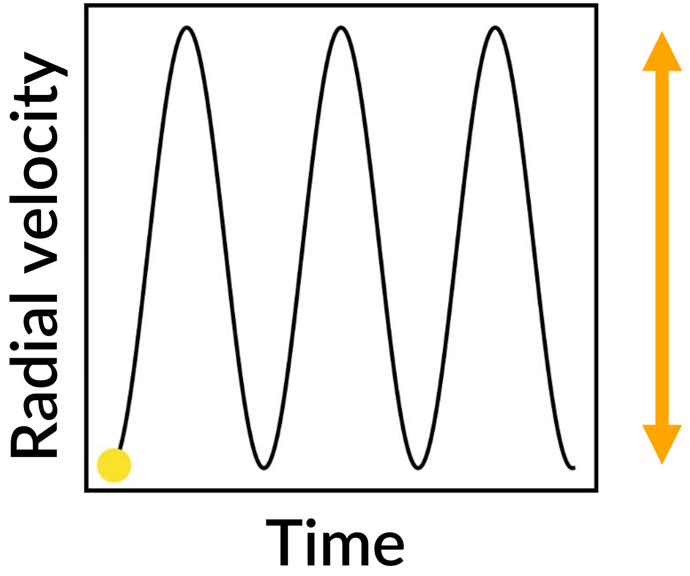


- Typical line width: 10 km/s

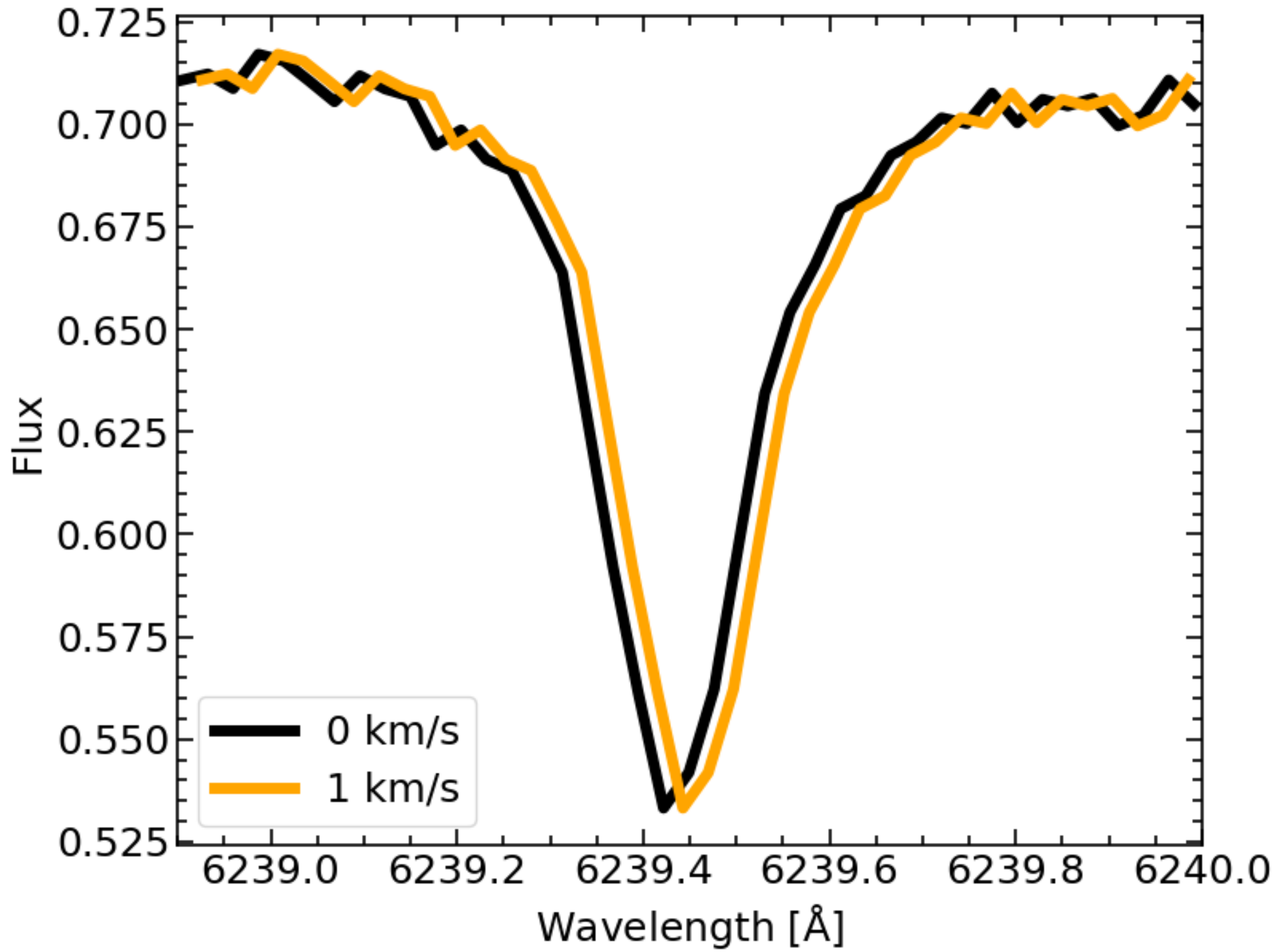
Radial velocity



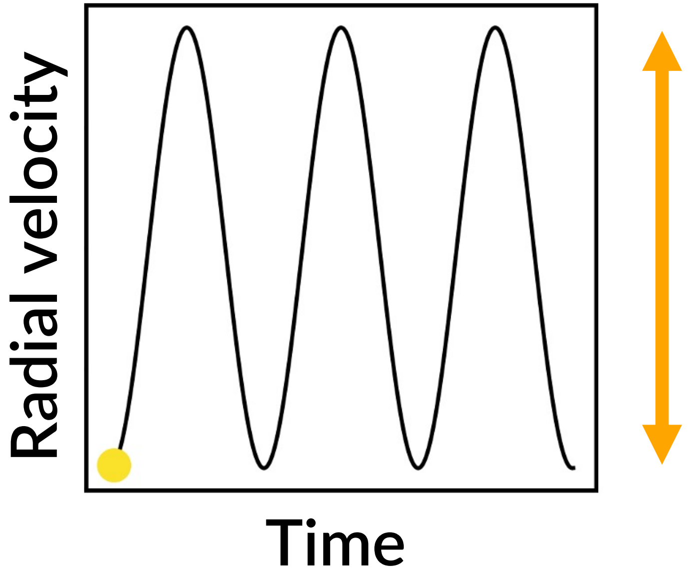
- Typical line width: 10 km/s
- RV shift
  - Binary star: ~ few km/s



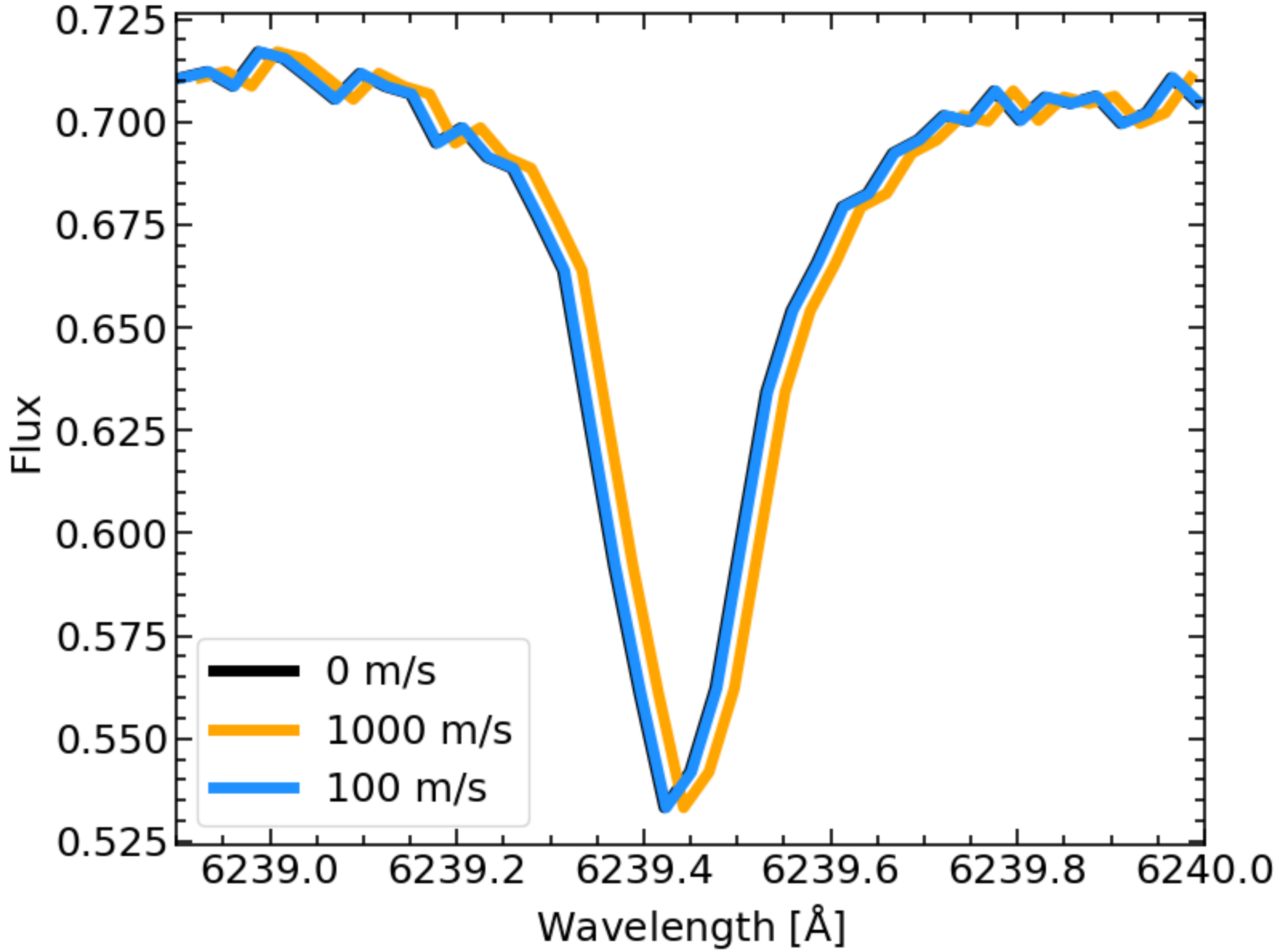
Radial velocity



- Typical line width: 10 km/s
- RV shift
  - Binary star: ~ few km/s

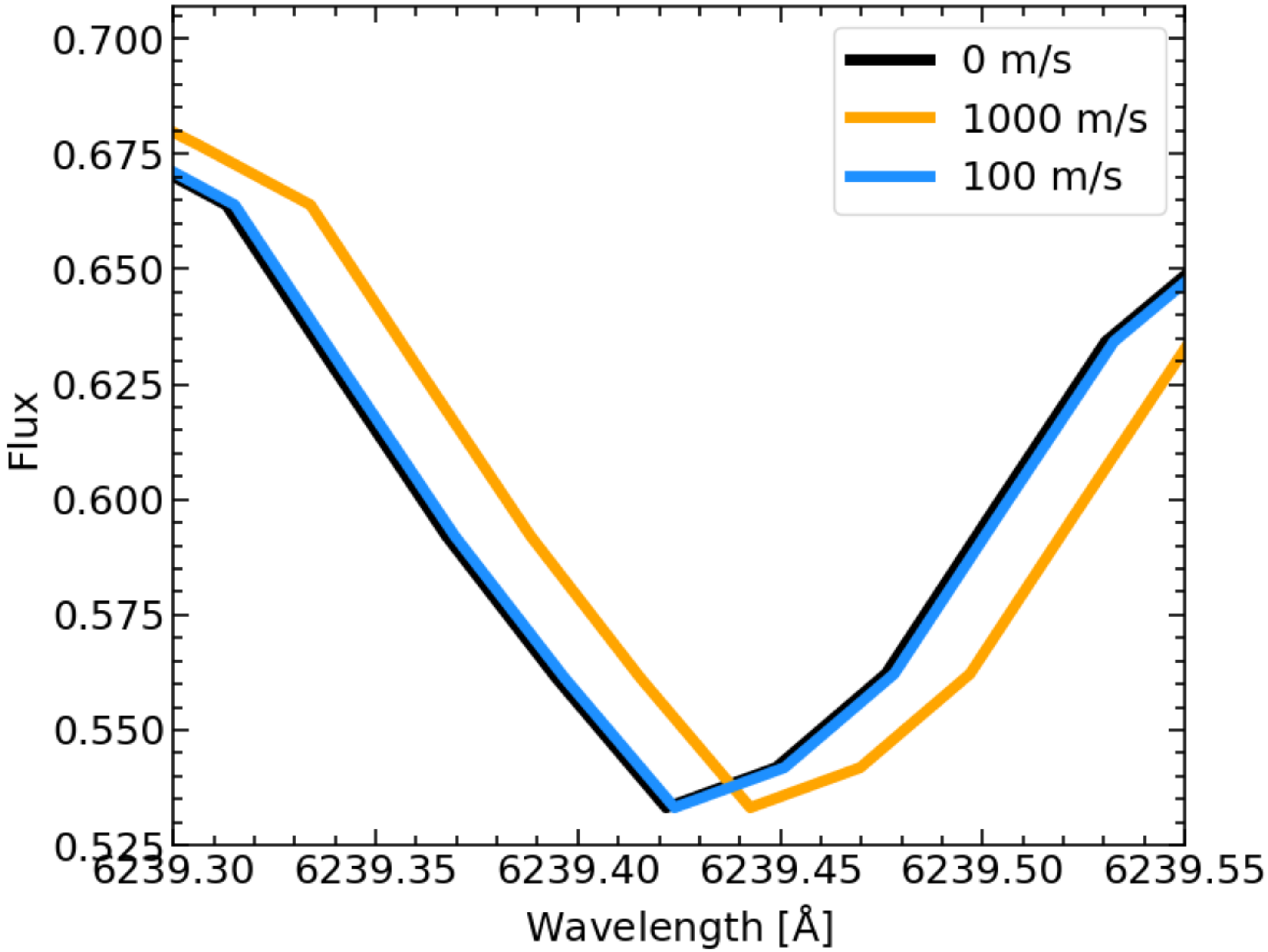


Radial velocity



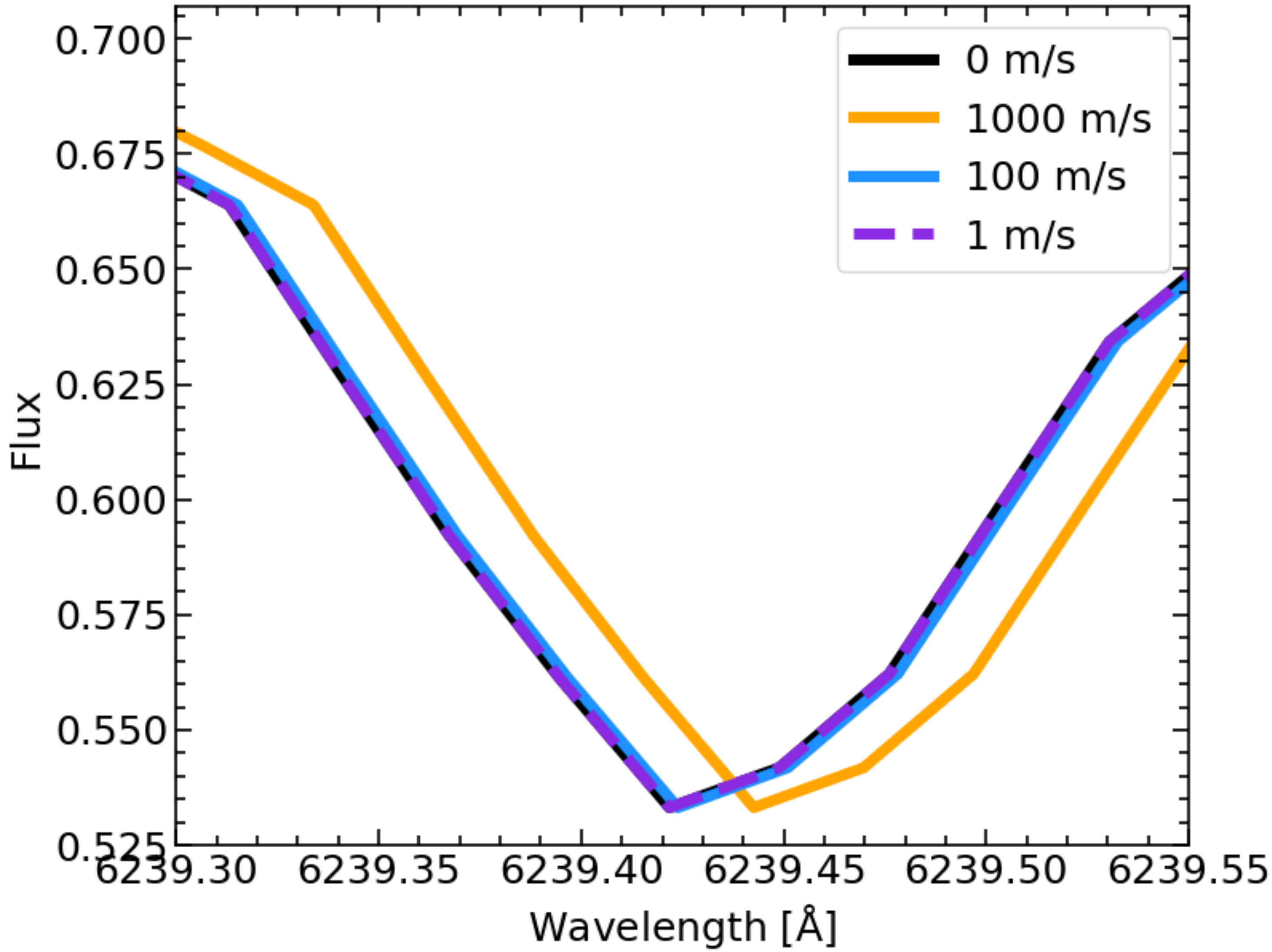
- Typical line width: 10 km/s
- RV shift
  - Binary star ~ km/s
  - Hot Jupiter ~100 m/s

Radial velocity



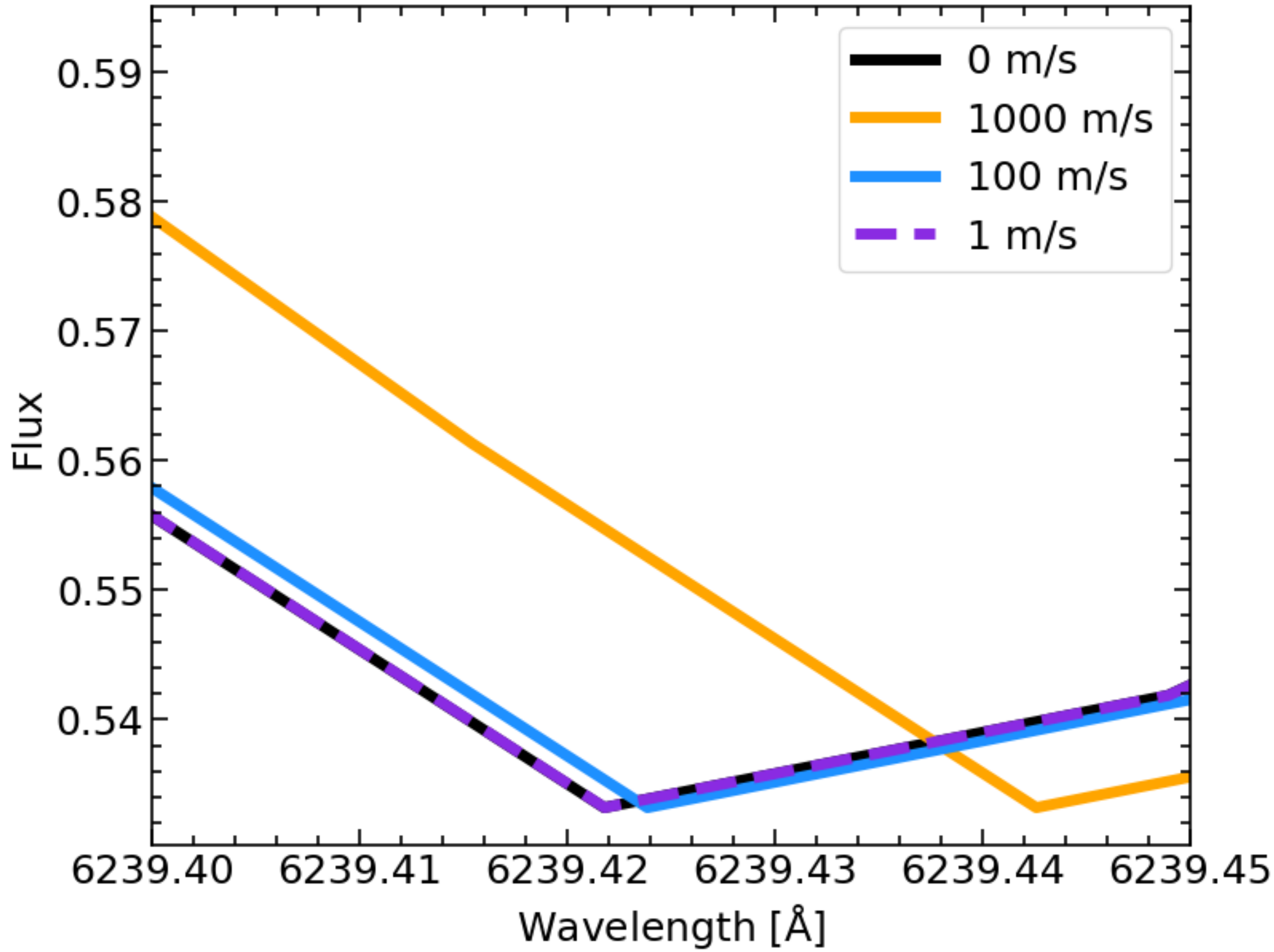
- Typical line width: 10 km/s
- RV shift
  - Binary star ~ km/s
  - Hot Jupiter ~100 m/s

Radial velocity



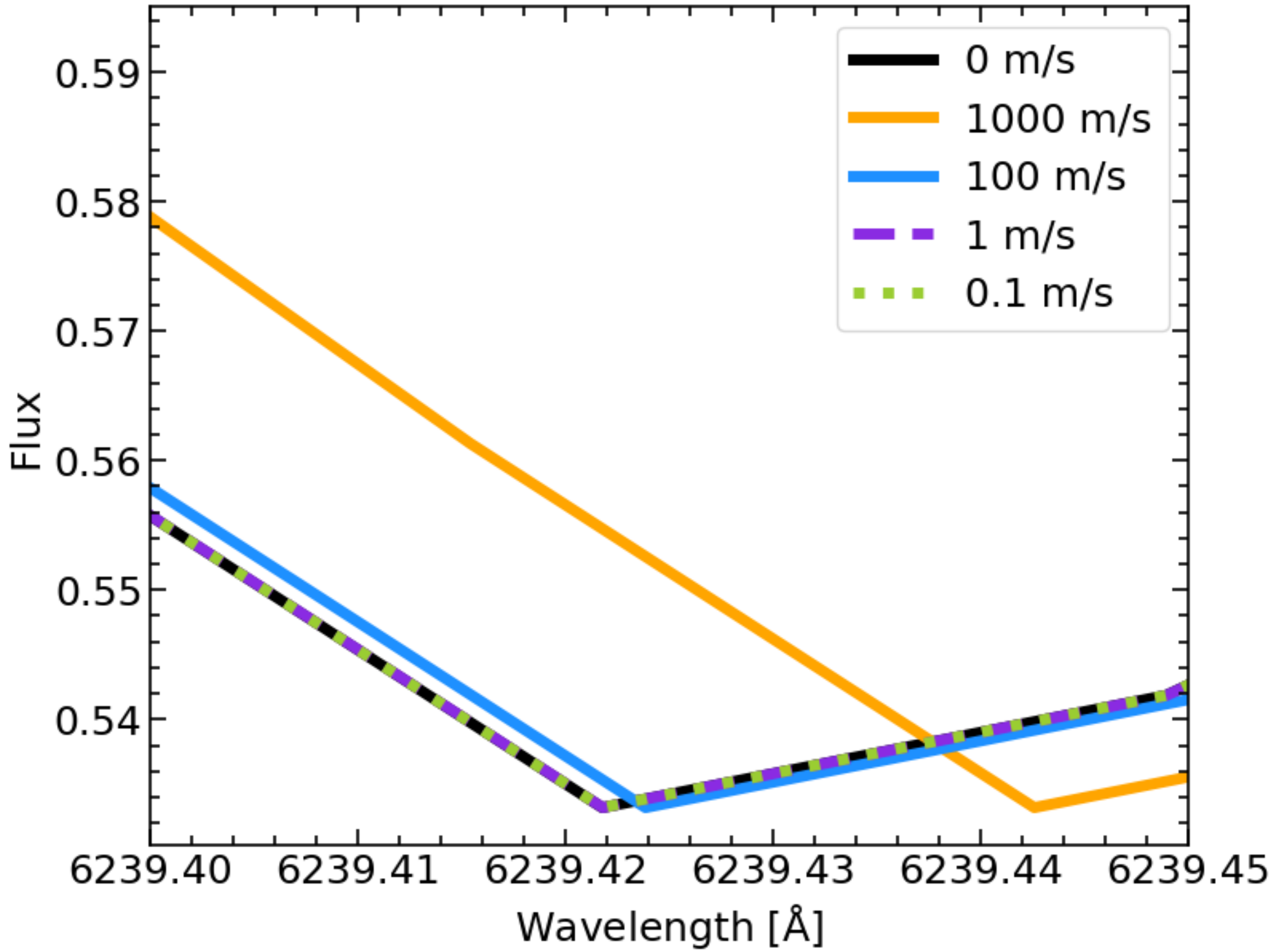
- Typical line width: 10 km/s
- RV shift
  - Binary star ~ km/s
  - Hot Jupiter ~100 m/s
  - RV precision ~1m/s

Radial velocity



- Typical line width: 10 km/s
- RV shift
  - Binary star ~ km/s
  - Hot Jupiter ~100 m/s
  - RV precision ~1m/s

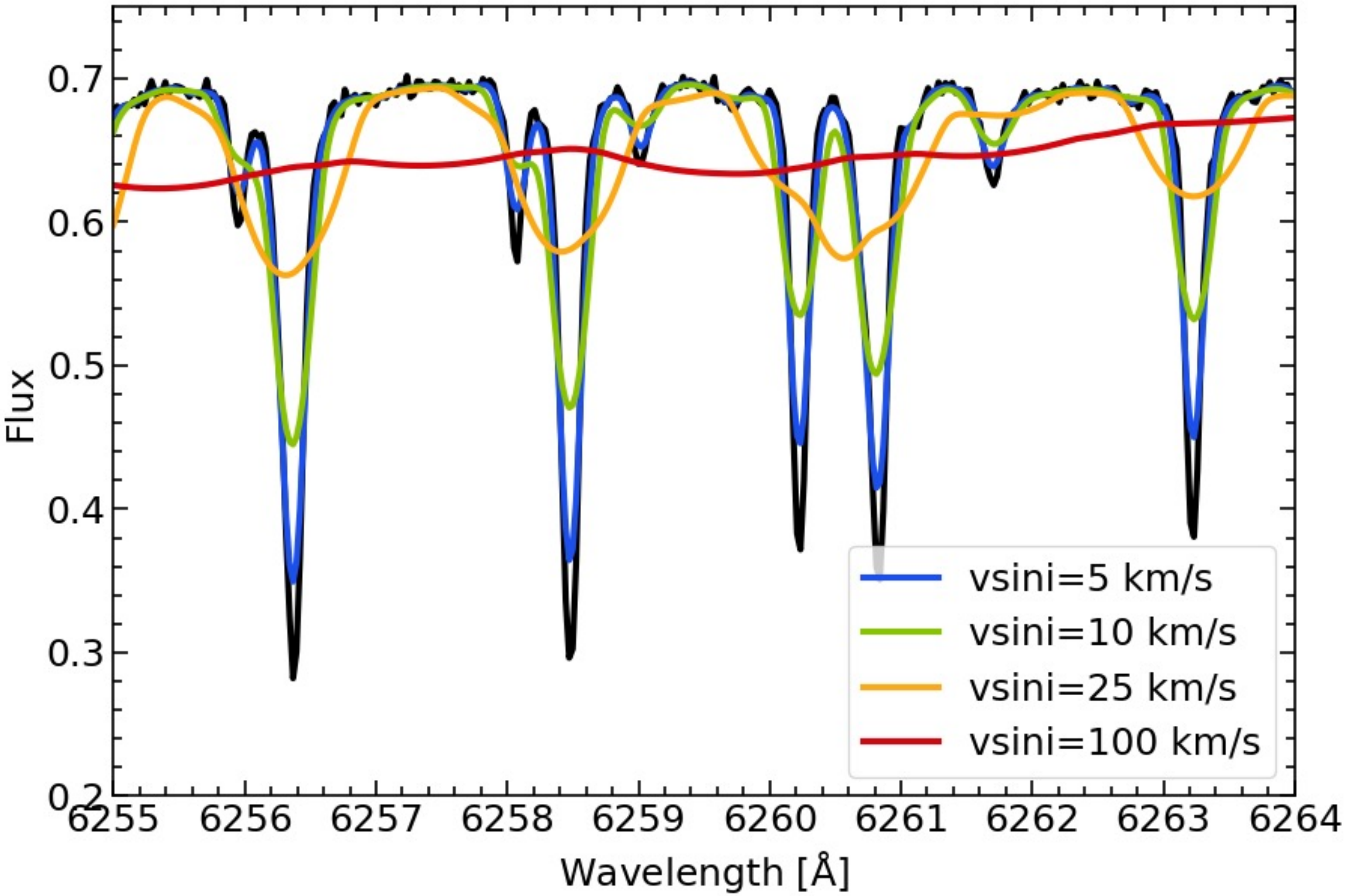
Radial velocity

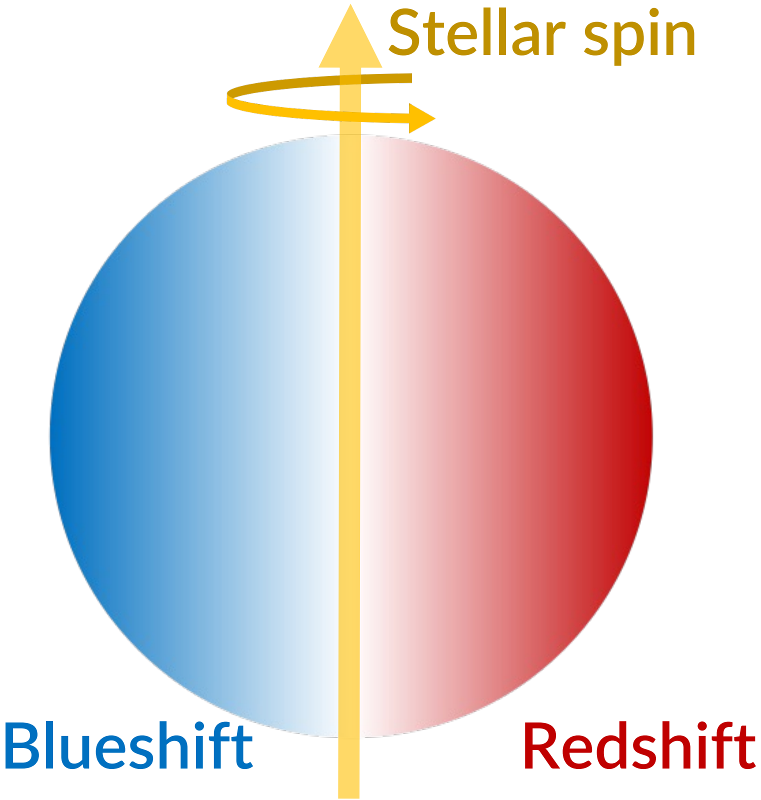


- Typical line width: 10 km/s
- RV shift
  - Binary star ~ km/s
  - Hot Jupiter ~100 m/s
  - RV precision ~1m/s
  - Earth-Sun ~0.1 m/s

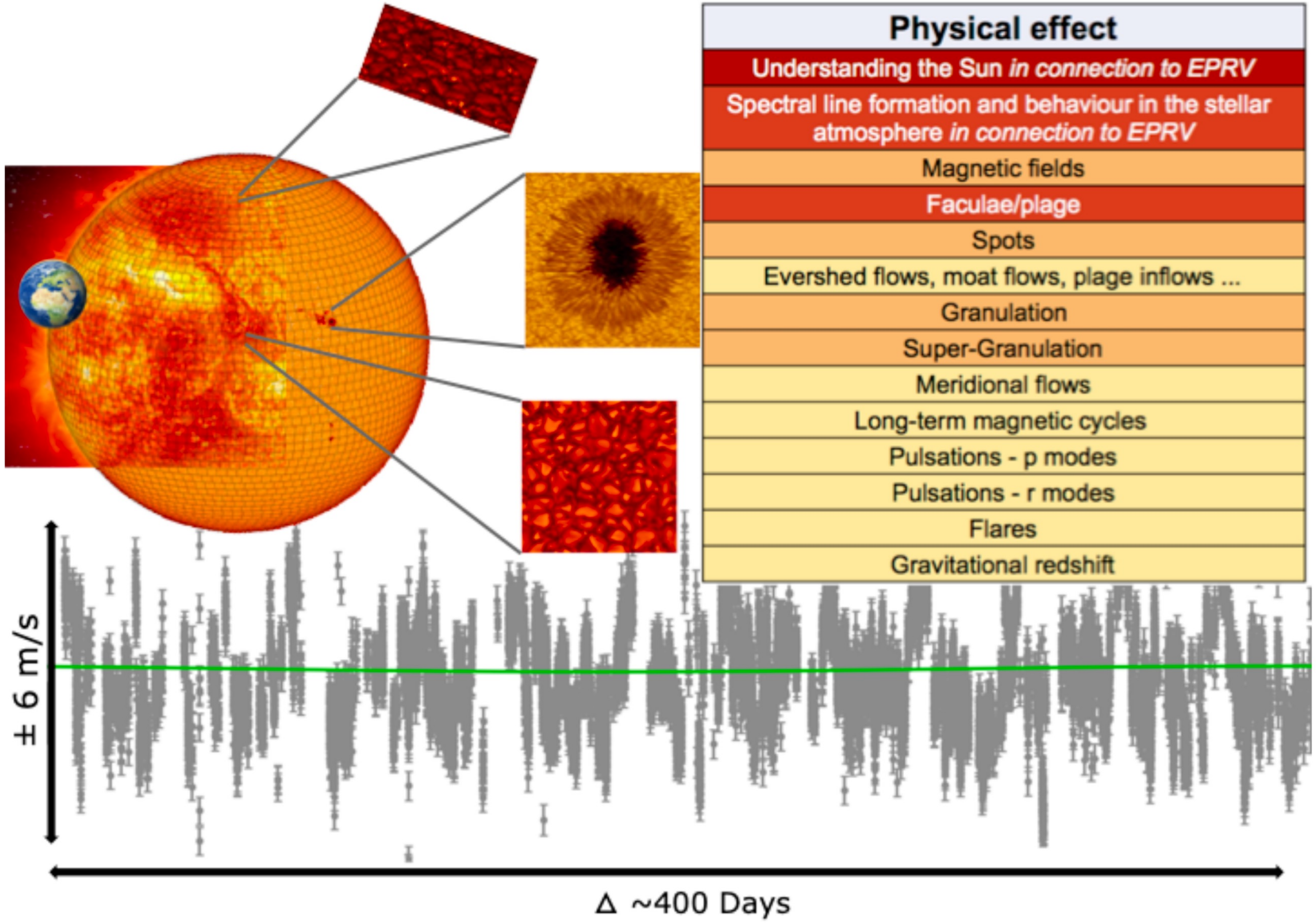
We are measuring
very small RV shifts!

Rotational broadening

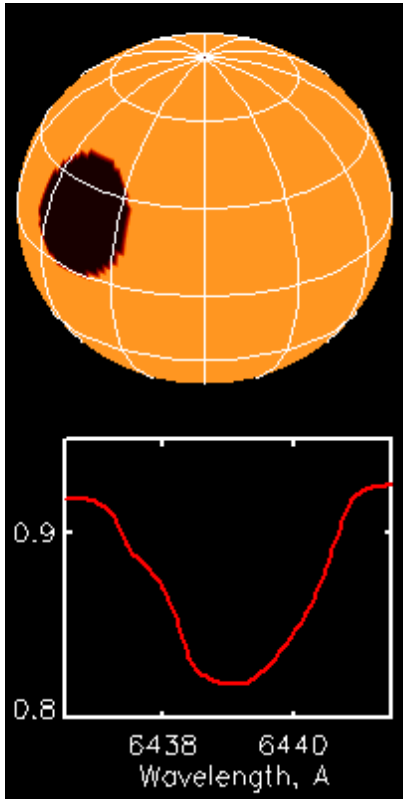




Stellar variability

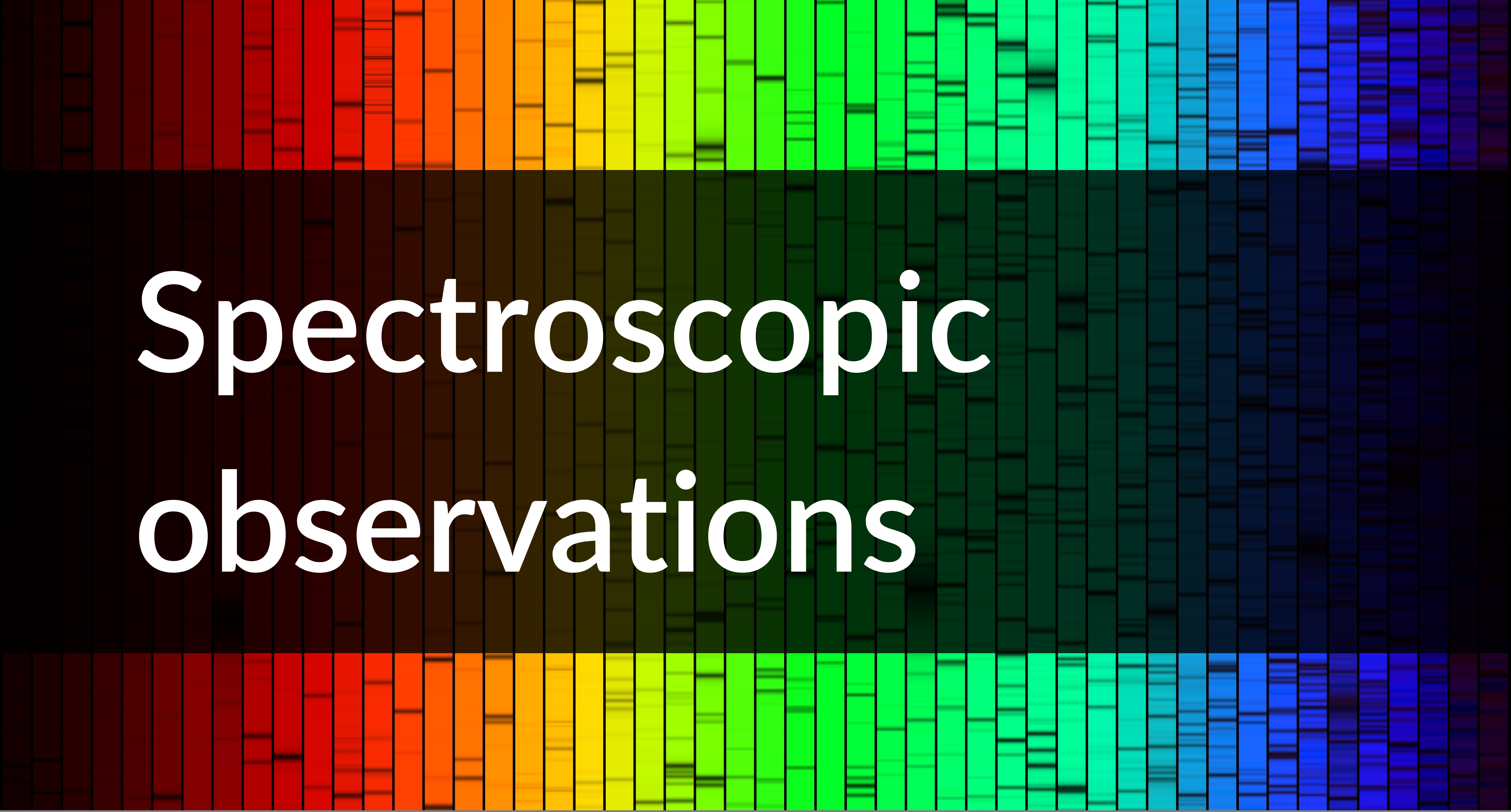


Physical effect
Understanding the Sun *in connection to EPRV*
Spectral line formation and behaviour in the stellar atmosphere *in connection to EPRV*
Magnetic fields
Faculae/plage
Spots
Evershed flows, moat flows, plage inflows ...
Granulation
Super-Granulation
Meridional flows
Long-term magnetic cycles
Pulsations - p modes
Pulsations - r modes
Flares
Gravitational redshift



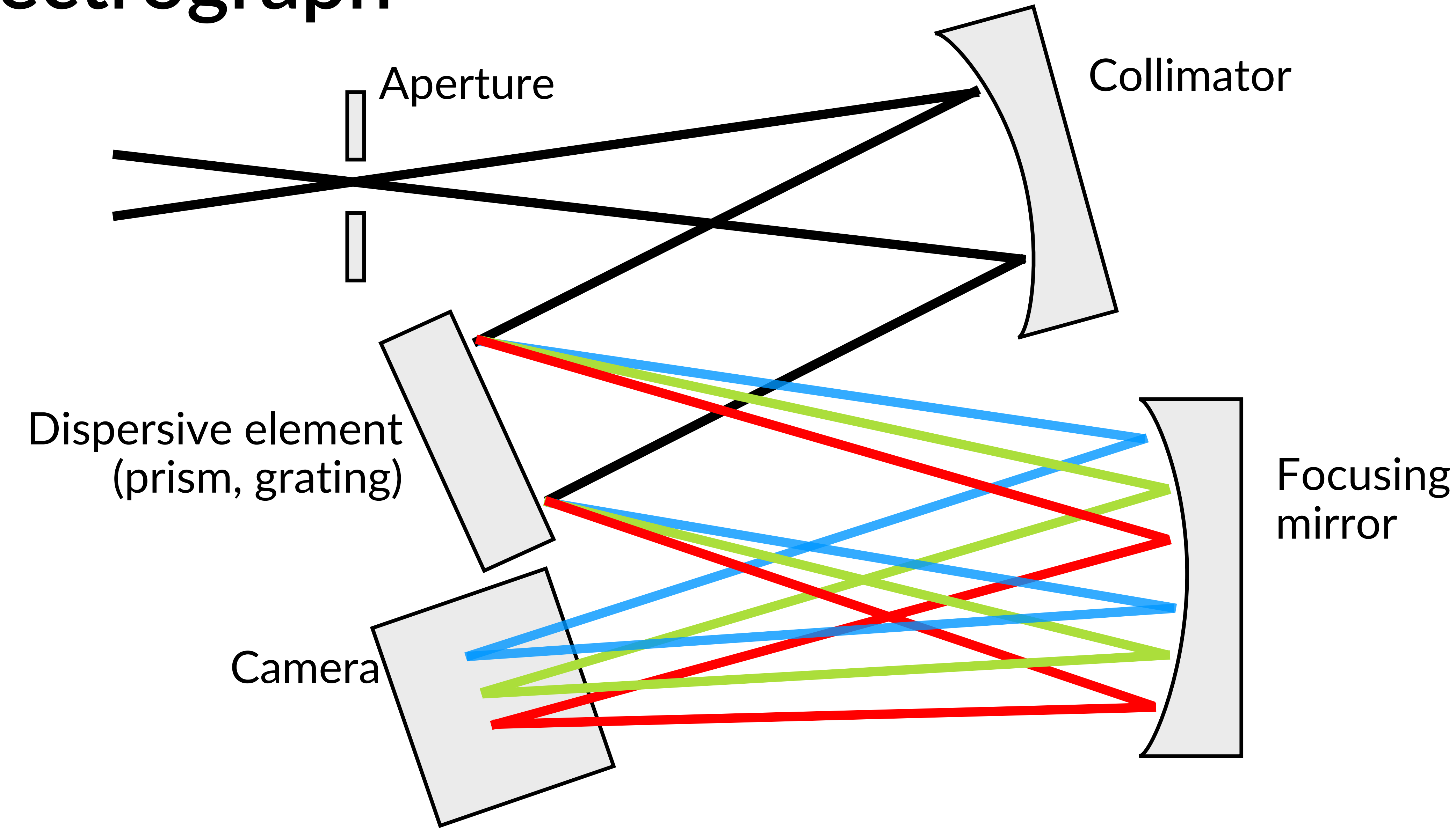
Svetlana Berdyugina

Crass et al. 2021 Extreme Precision Radial Velocity Working Group Final Report, figure NASA, ESA, SDO/HMI, MURaM, Big Bear Solar Observatory, solar RV observations from HARPS-N, Cegla/Haywood/Watson

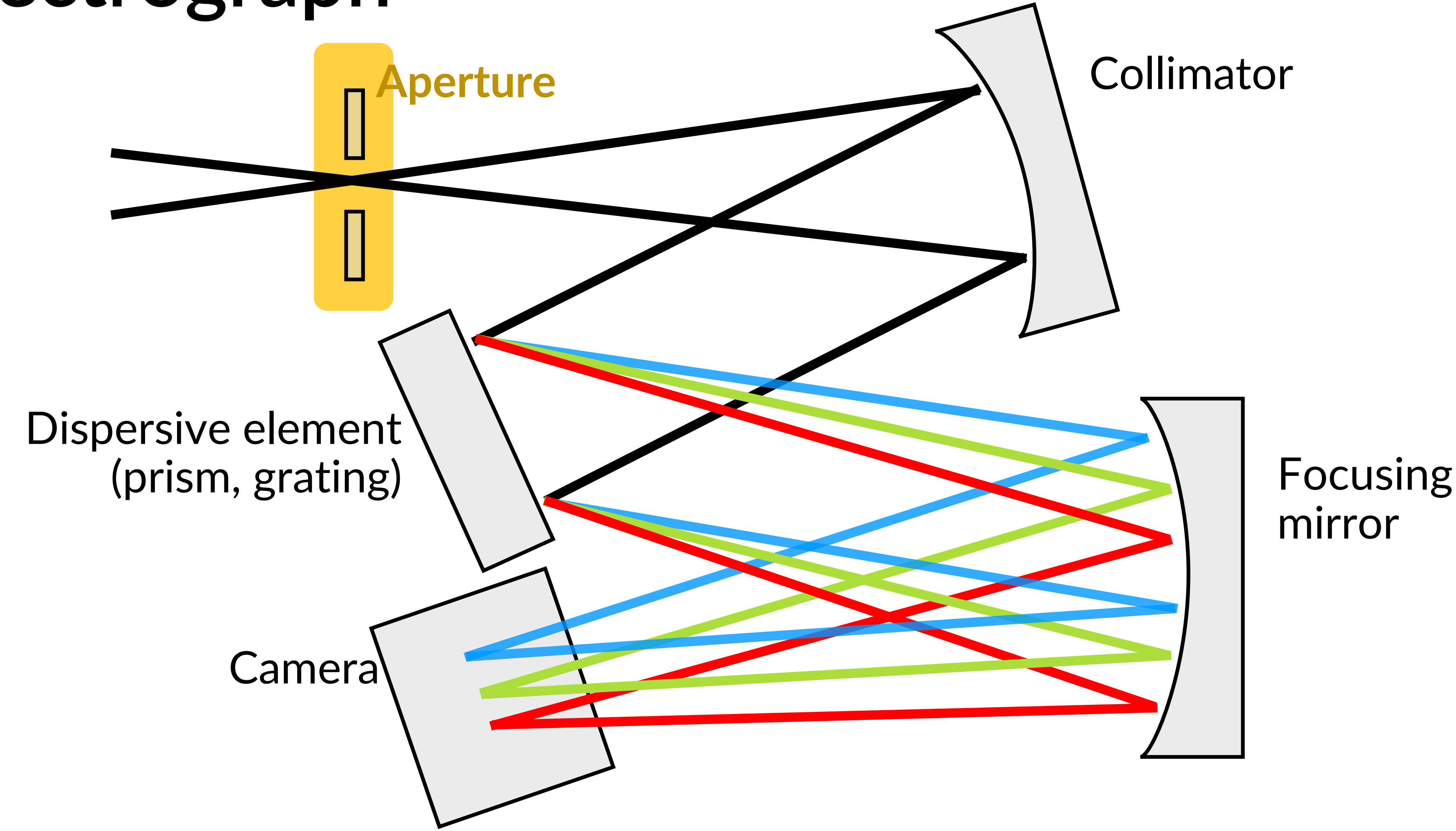


Spectroscopic observations

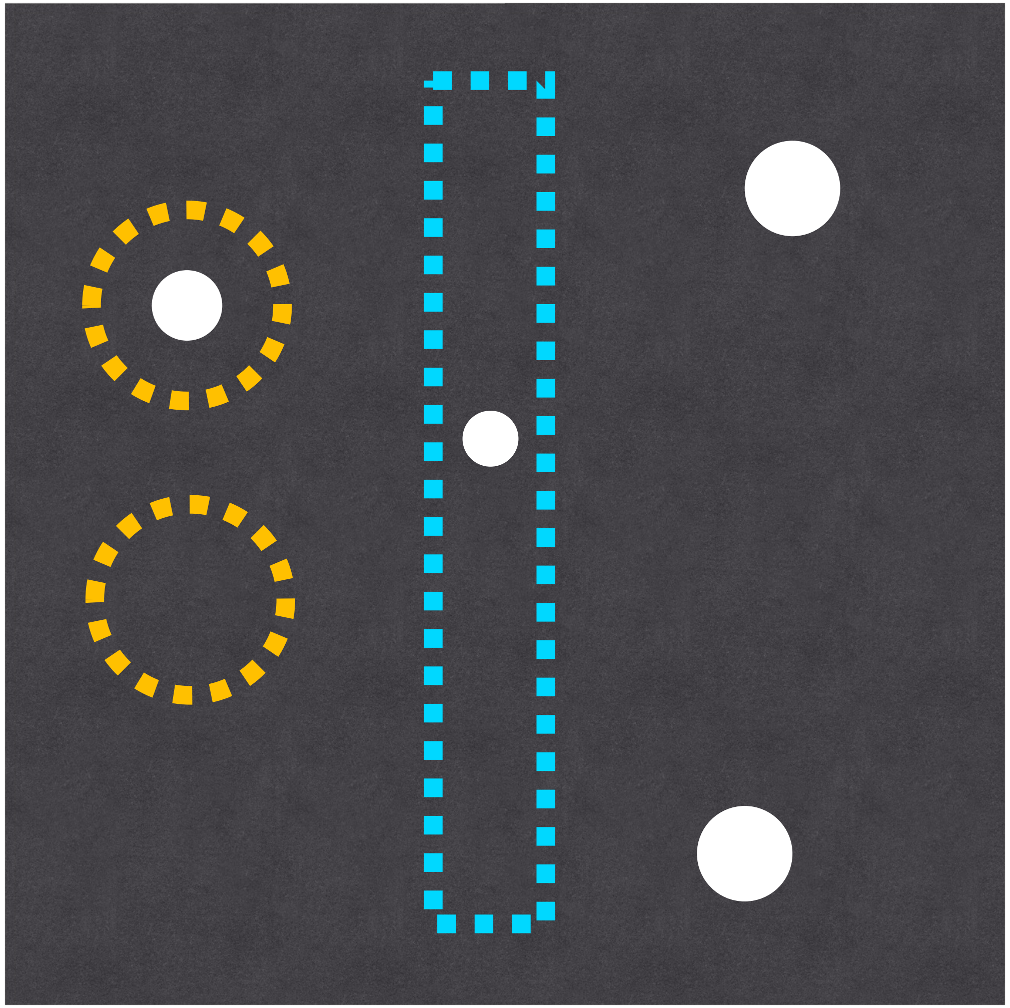
Spectrograph



Spectrograph



Collecting light



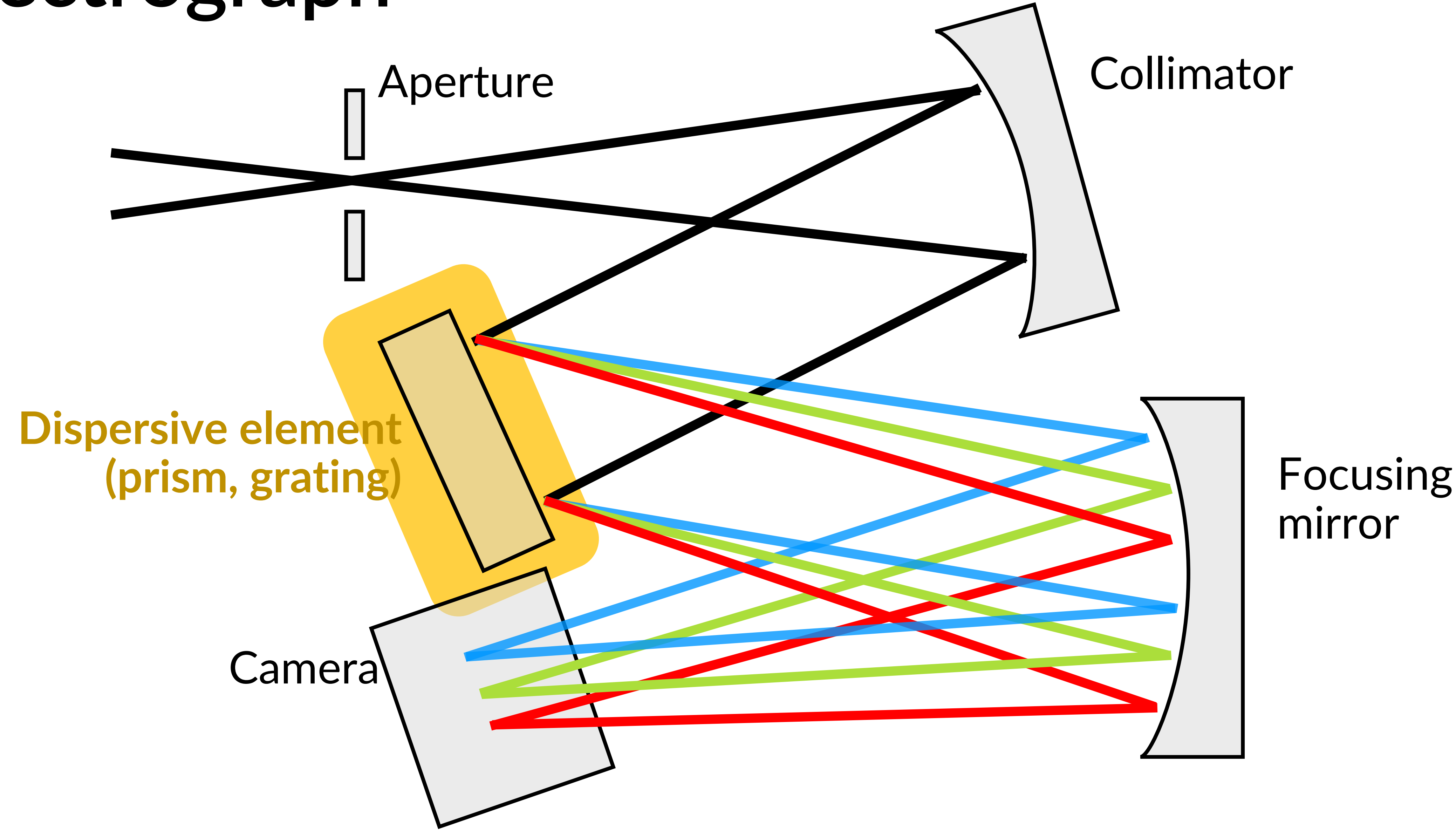
Slit: Mechanical aperture with 2 parallel jaws

- Width can be easily changed
- One spatial direction (along the slit) preserved
- Simultaneous spectrum of the sky
- Typical slit widths: 0.2 – 2.0″

Fibre: Optical guide transmitting light through multiple reflections

- Very constant output (↑ stability)
- **Instrument can be moved off-telescope** (↑ stability)
- Additional fibre(s) for **sky or calibration source**
- Typical fibre diameters: 1.0 – 1.5″ (match seeing)

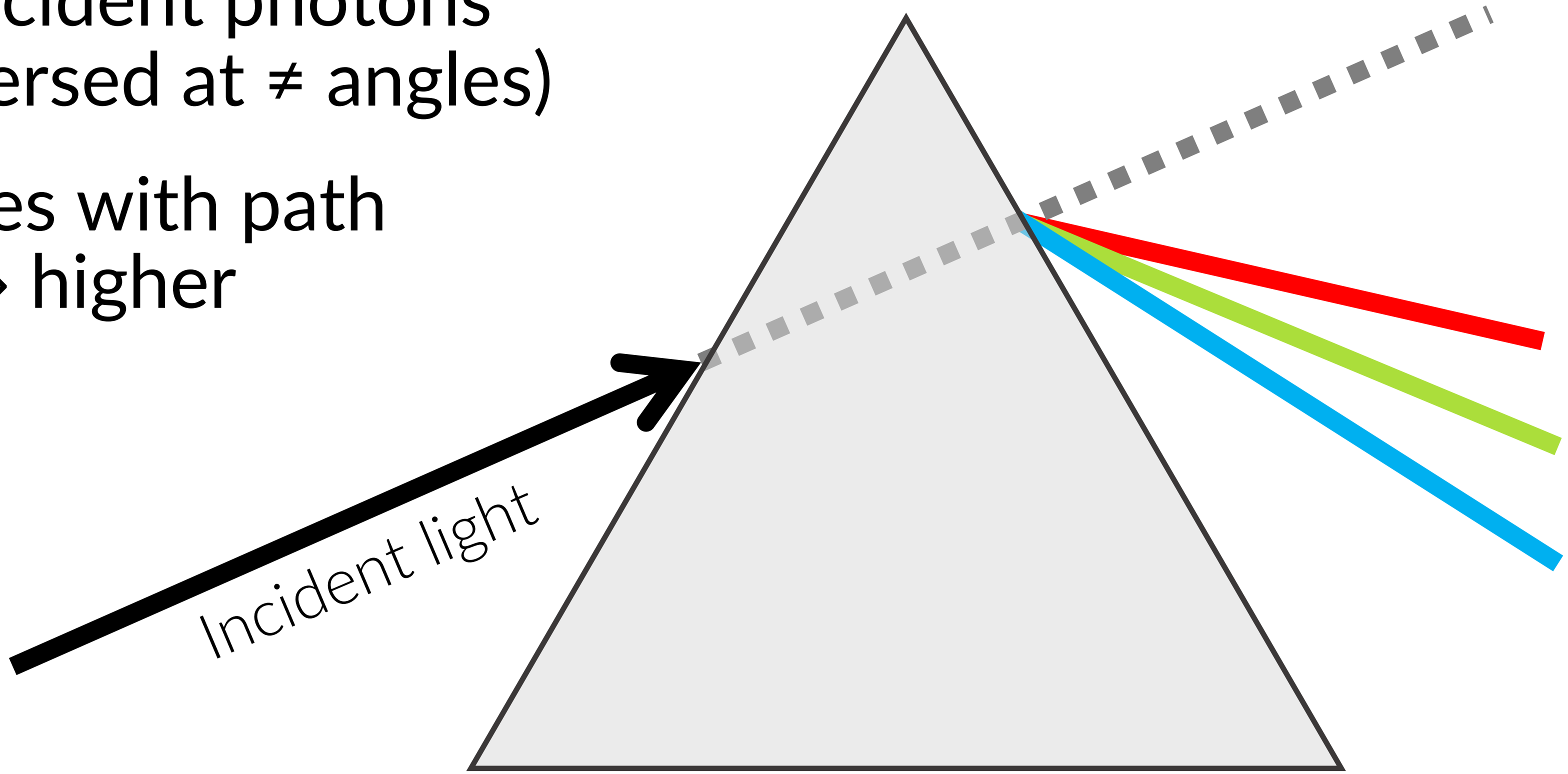
Spectrograph



Dispersing light: prisms, gratings

Prism

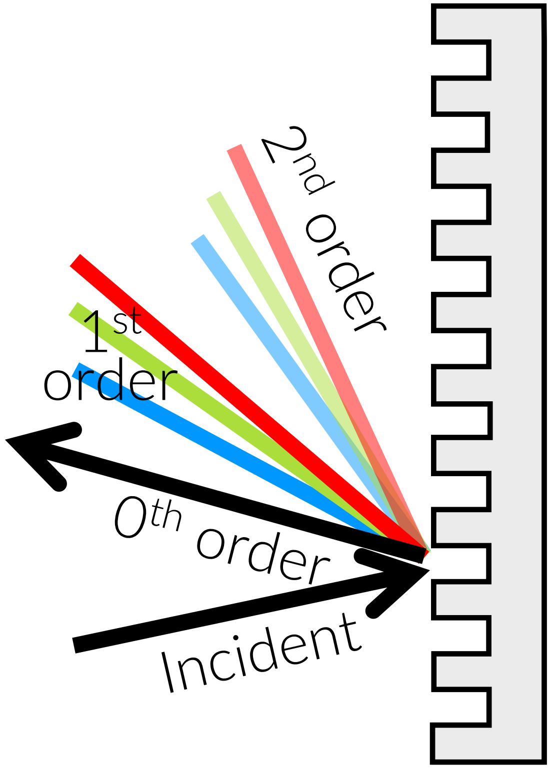
- Uses **variable index of refraction** $\boldsymbol{n(\lambda)}$ to separate incident photons (≠ colours are dispersed at ≠ angles)
- Dispersion increases with path (i.e. larger prism $\Longrightarrow$ higher resolution)



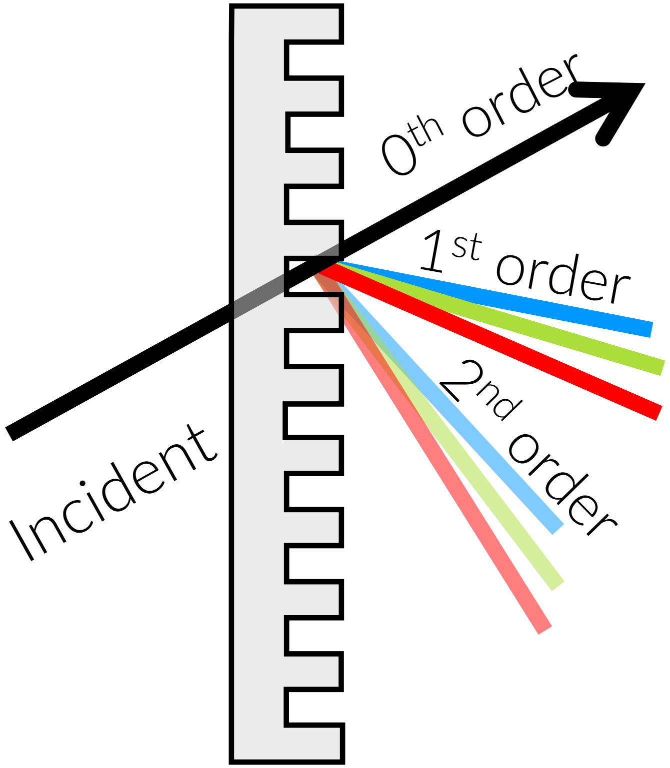
Dispersing light: prisms, gratings

Diffraction grating

- Uses **diffraction + interference** to separate incident photons
- Periodic carving in material with spacing ~ λ of light
- Diffraction orders at interference maxima



Reflection grating
(reflective material)

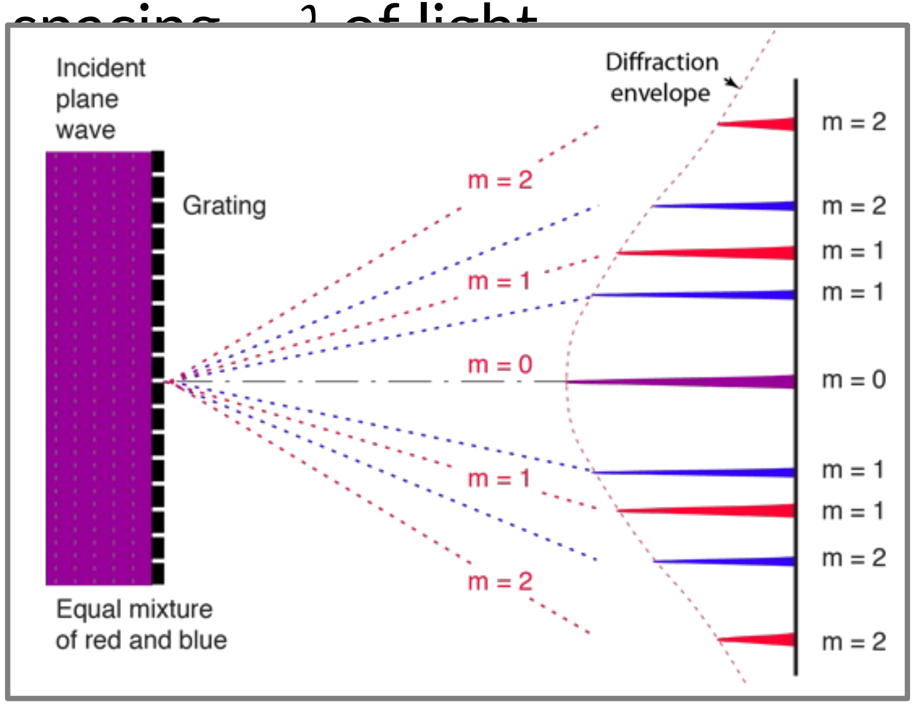


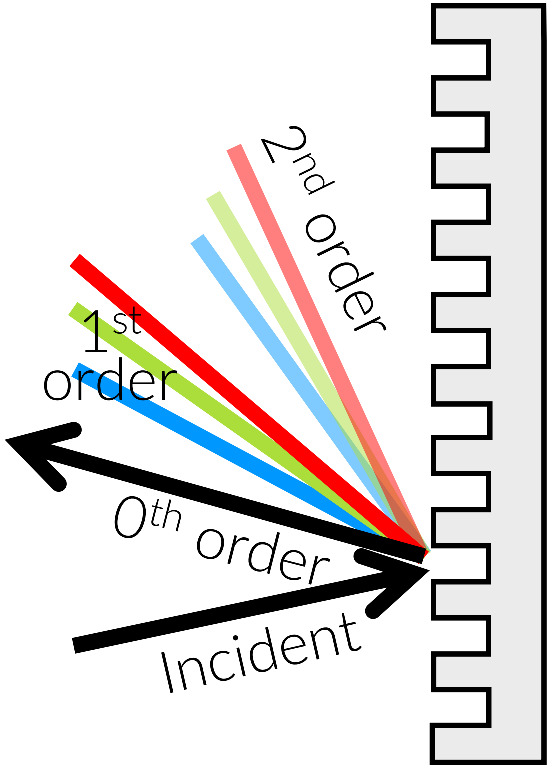
Transmission grating
(transmissive material)

Dispersing light: prisms, gratings

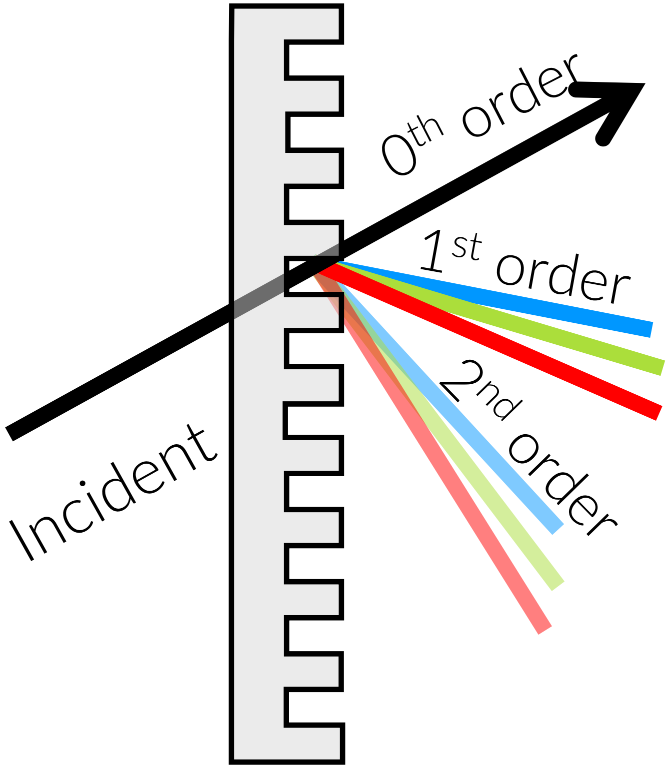
Diffraction grating

- Uses **diffraction + interference** to separate incident photons
- Periodic carving in material with spacing λ of light
- e





Reflection grating
(reflective material)

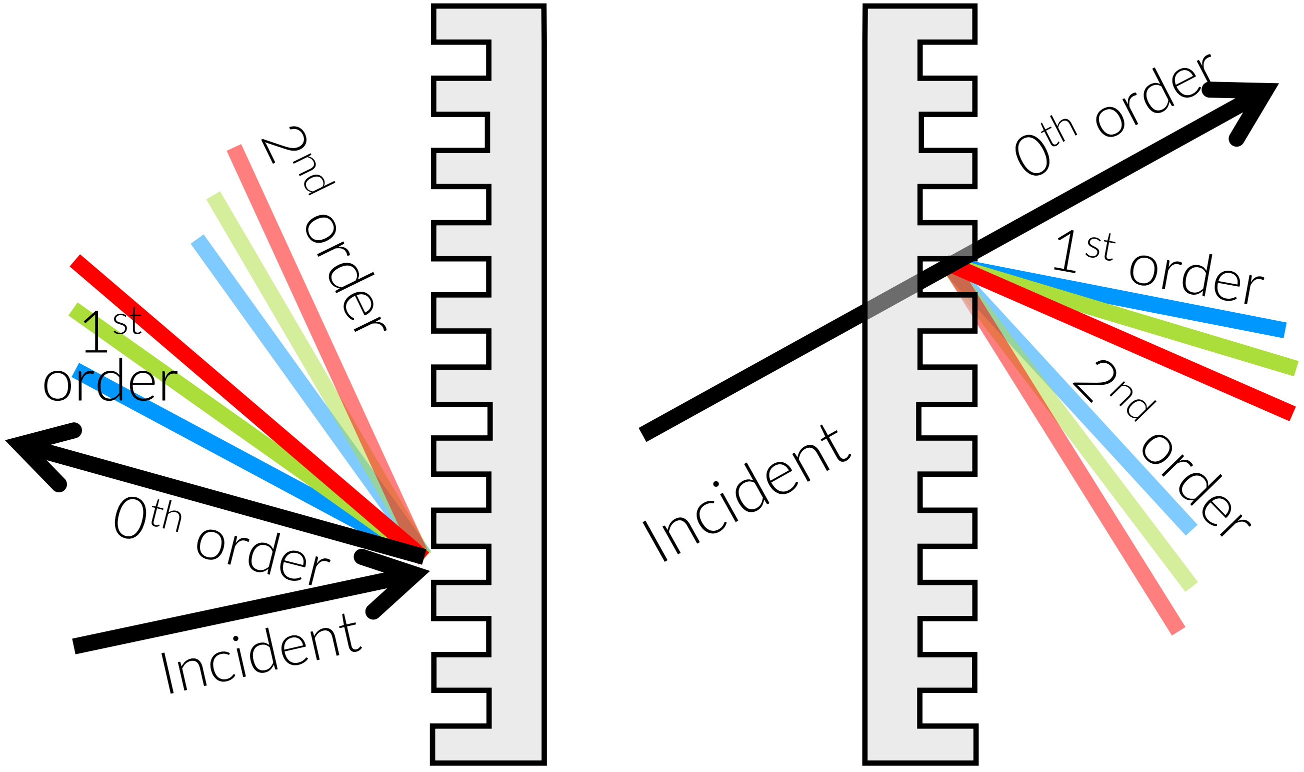


Transmission grating
(transmissive material)

Dispersing light: prisms, gratings

Diffraction grating

- Uses **diffraction + interference** to separate incident photons
- Periodic carving in material with spacing ~ λ of light
- Diffraction orders at interference maxima
- Resolution ↑ with line density
- Resolution ↑ for higher orders (but intensity ↓)
- Most of the light goes to the first orders

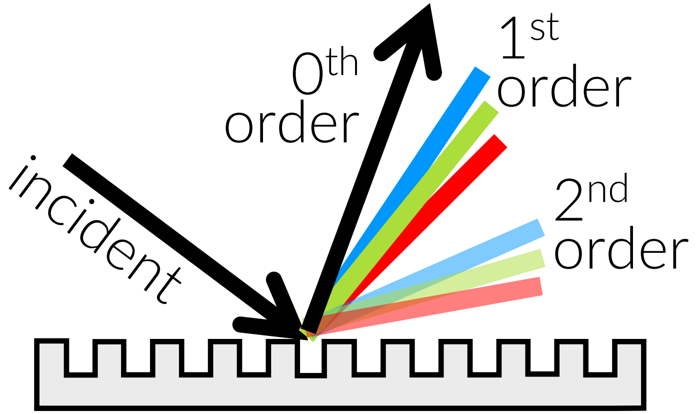


Reflection grating (reflective material)

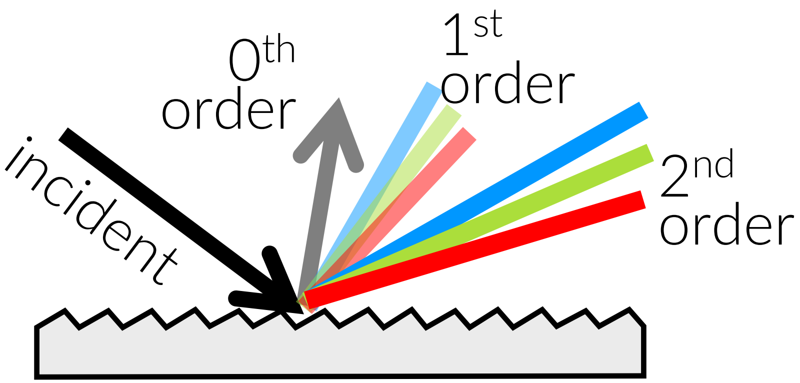
Transmission grating (transmissive material)

Dispersing light: blazed gratings

- Gratings can be ***blazed*** to **concentrate light** away from 0th order and **towards higher orders**
- Reflecting surfaces oriented at a specific *blaze angle* with respect to the surface of the grating
- Called echelette if used for low orders or **echelle** if used for high orders (large blaze angle, > 45°)
- **Order overlap** becomes a problem



Regular (unblazed) grating

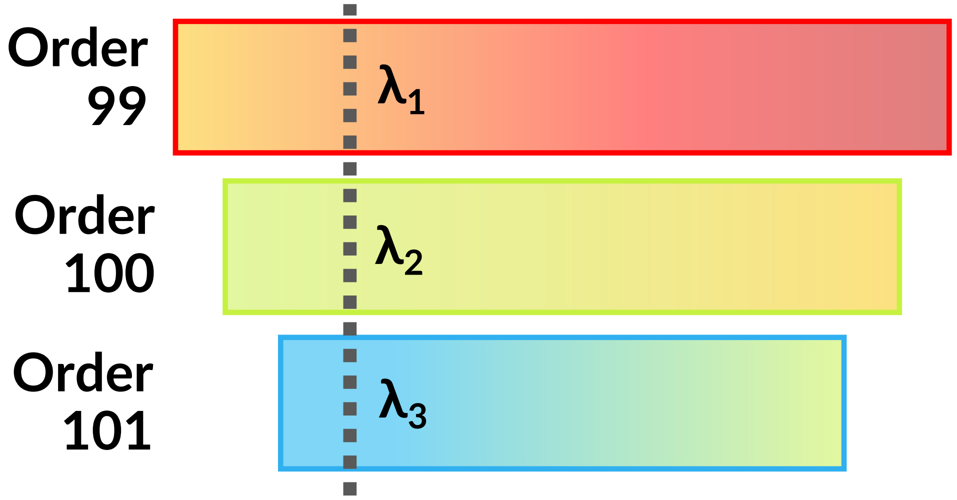


Blazed grating

Dispersing light: cross-dispersion

- The wavelength range of high orders strongly overlaps
- Want to measure λ_1 in order 99, but λ_2 in order 100 and λ_3 in order 101 contaminate your spectra

Orders separated vertically for clarity

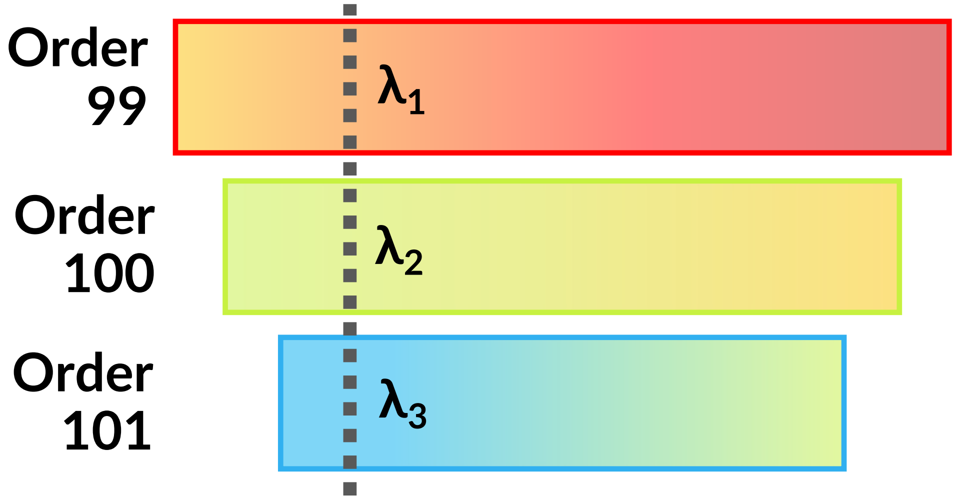


In reality, orders overlap

Dispersing light: cross-dispersion

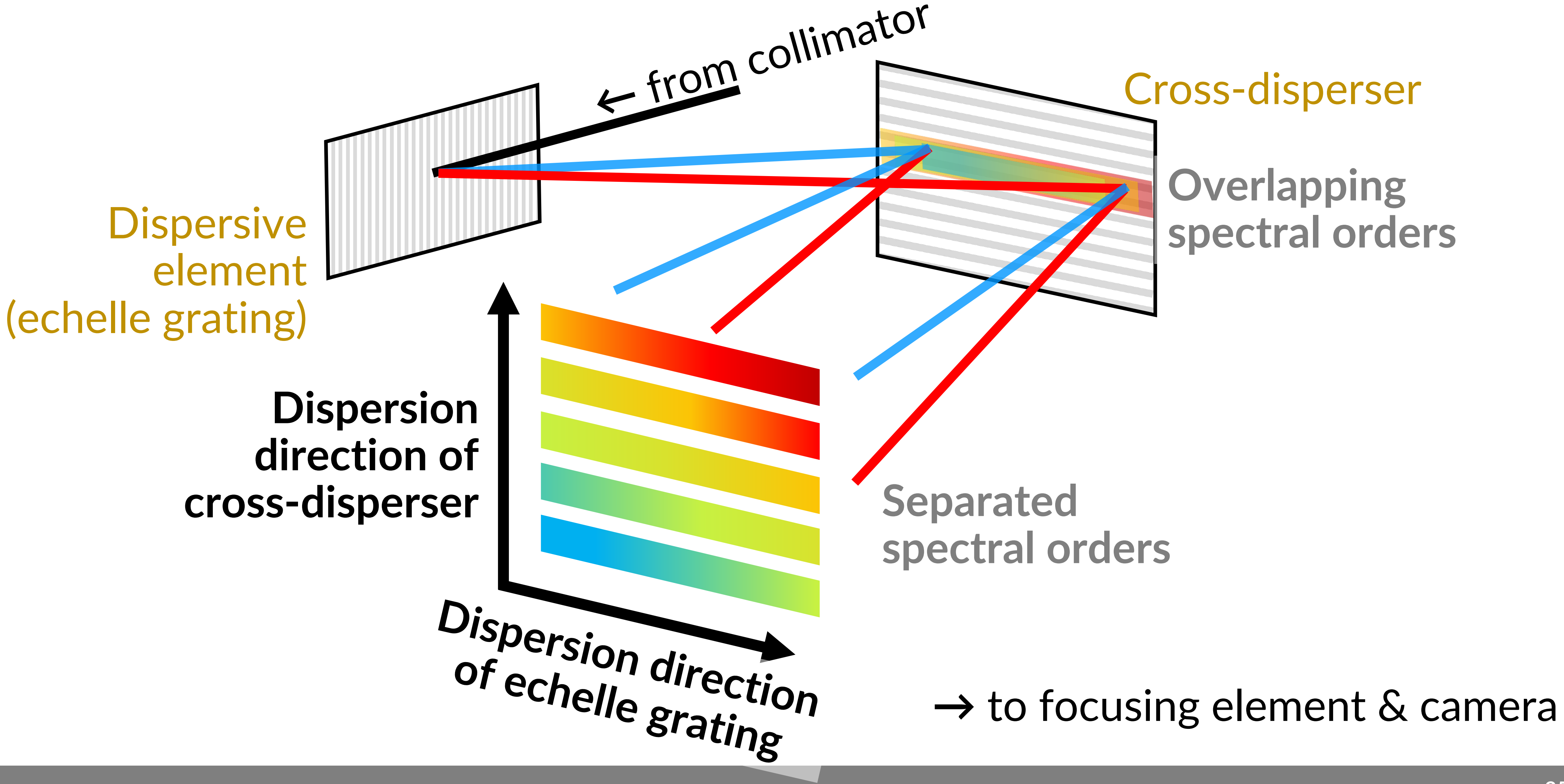
- The wavelength range of high orders strongly overlaps
- Want to measure λ_1 in order 99, but λ_2 in order 100 and λ_3 in order 101 contaminate your spectra
- Solutions:
    - Bandpass filters to selected desired λ range, but lose light
    - Cross-dispersion perpendicular to the initial spectral dispersion to separate the orders

Orders separated vertically for clarity



In reality, orders overlap

Dispersing light: cross-dispersion



Spectral resolution: separating spectral lines

What is the minimum distance between lines ($\Delta\lambda$) to be considered spectrally resolved?

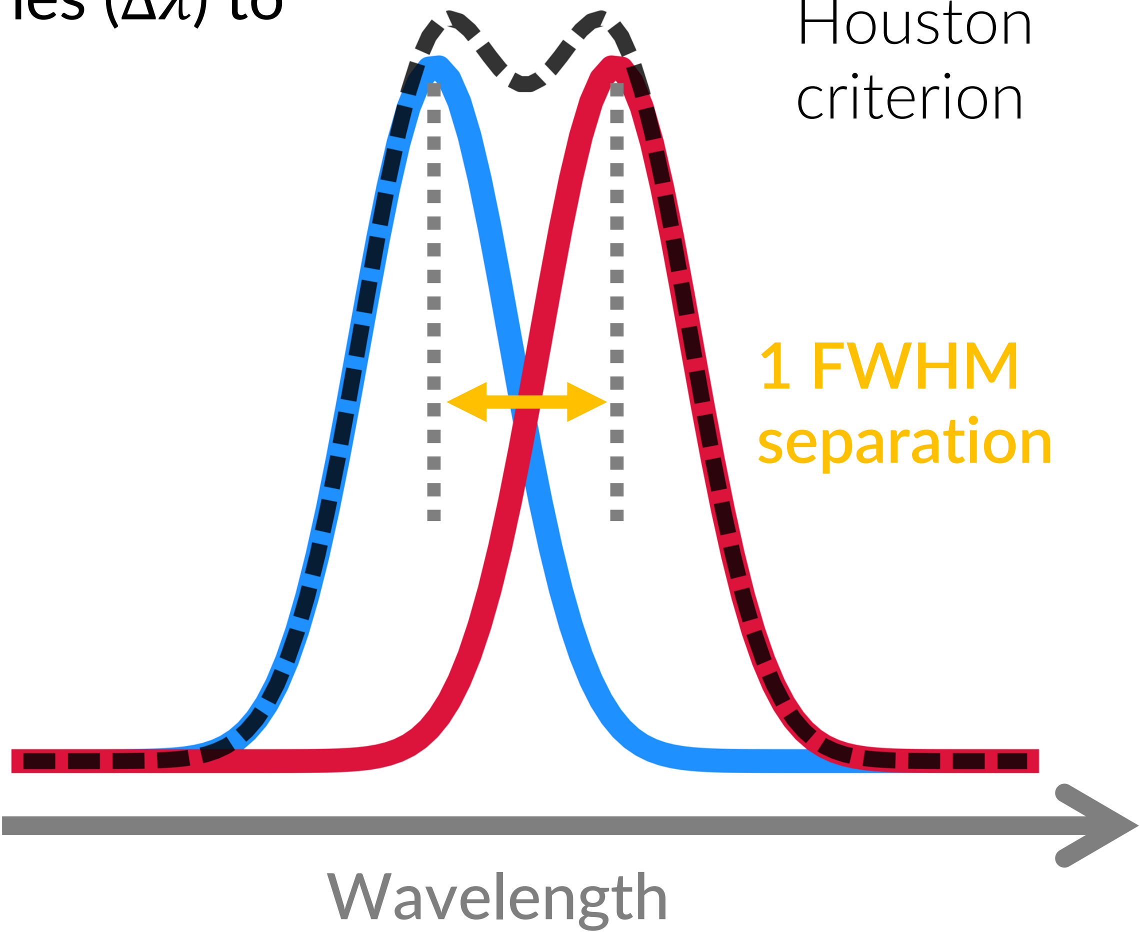
- **Spectral resolution**: $\Delta\lambda$
- **Resolving power**: $\mathrm{R} = \lambda/\Delta\lambda$

Related to **Doppler shift:** $\Delta v \sim c/\mathrm{R}$

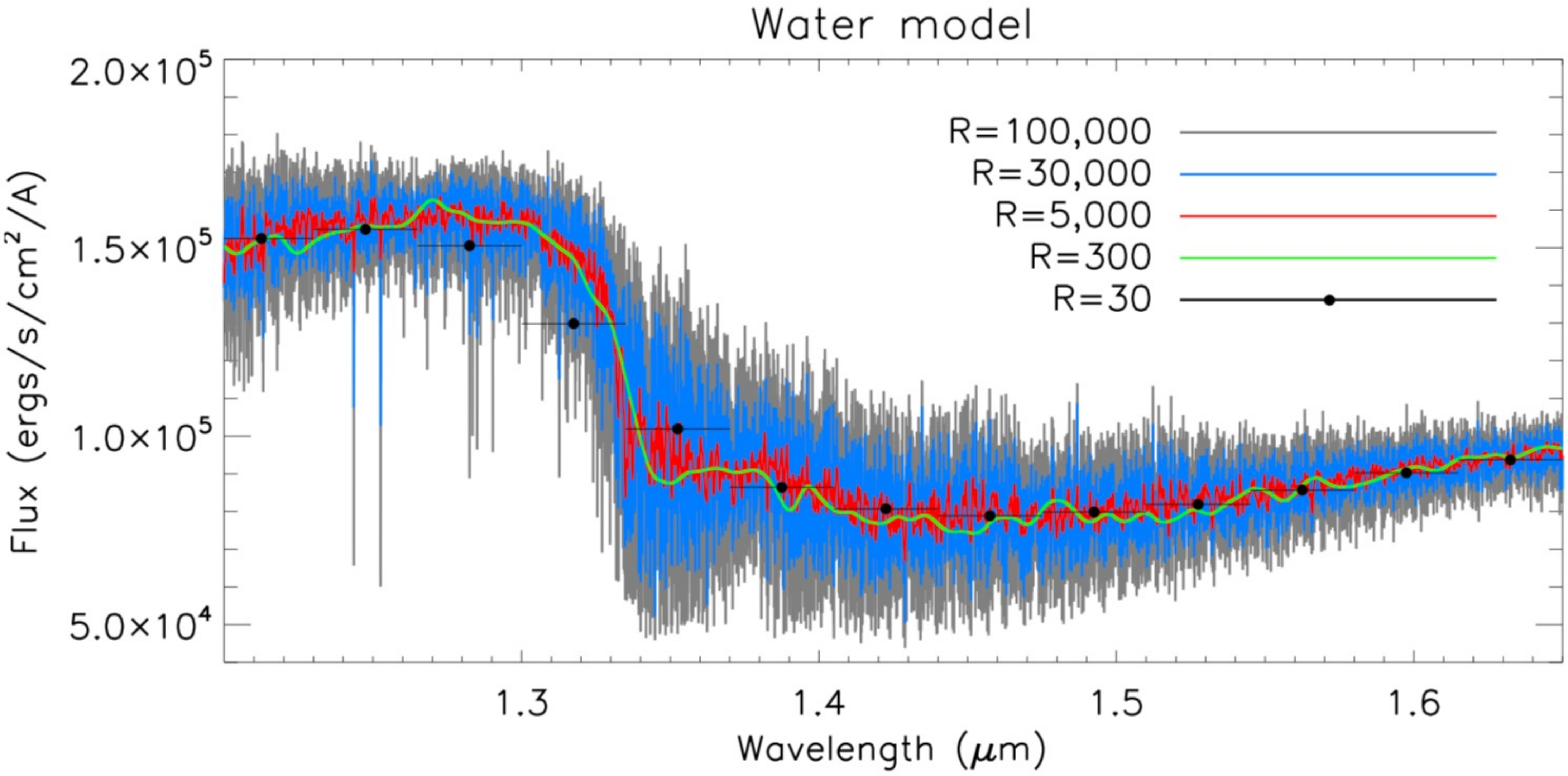
E.g. R = 100 000 ⟹
Resolve 3 km/s in velocity or
0.005 nm in wavelength at 500 nm

Resolution increases with **density of lines in grating** & **order number**

Trade-off between resolution & amount of photons!

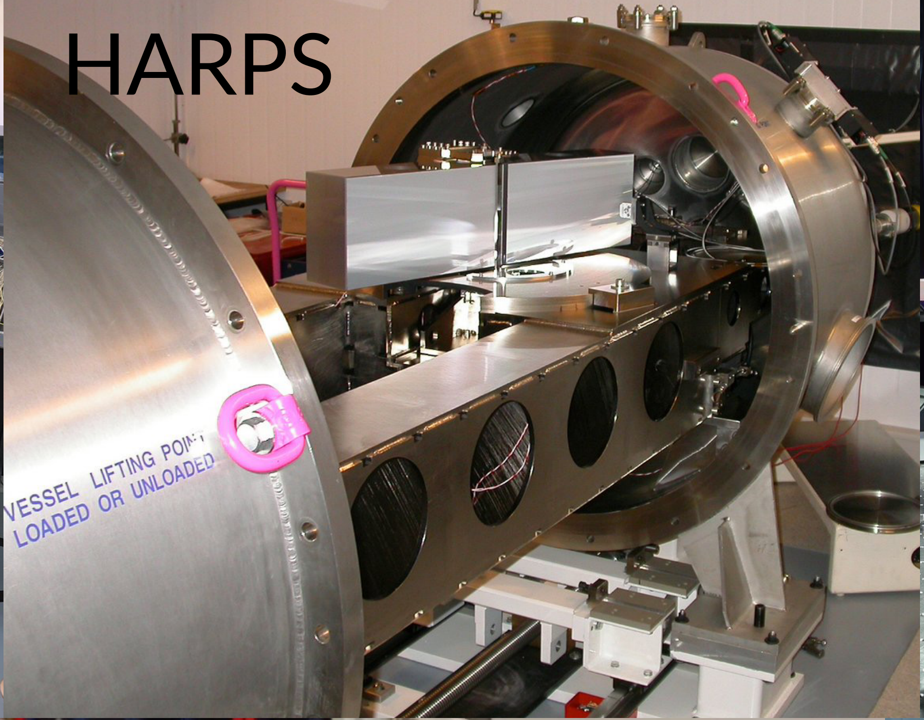


Spectral resolution



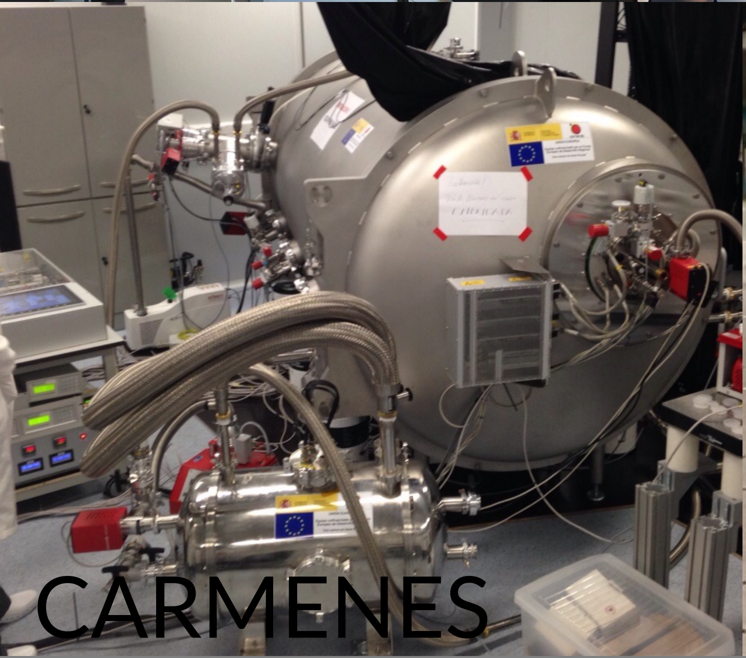
Birkby 2018

ESPRESSO

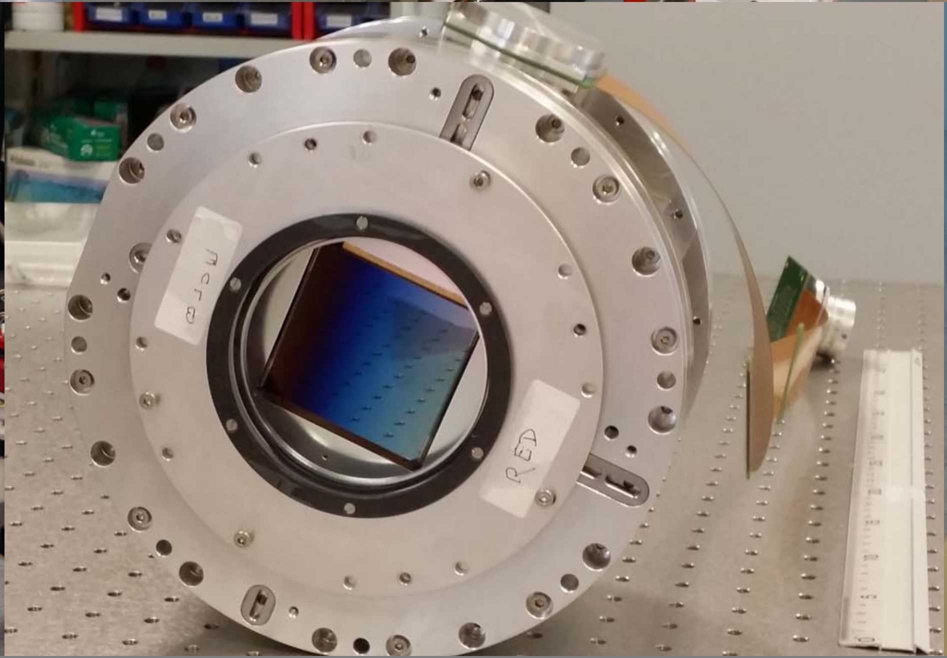


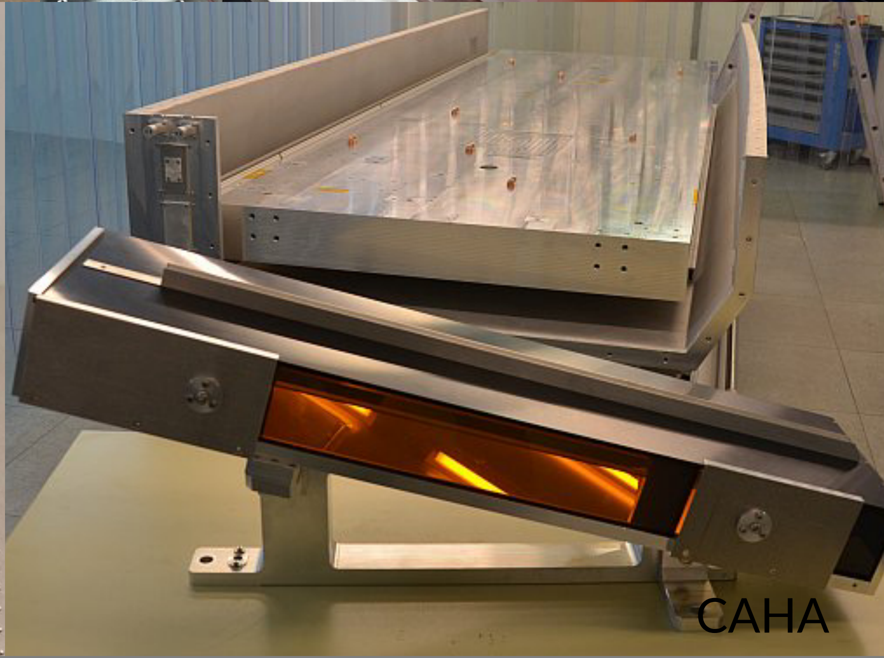
HARPS

ESO



CARMENES

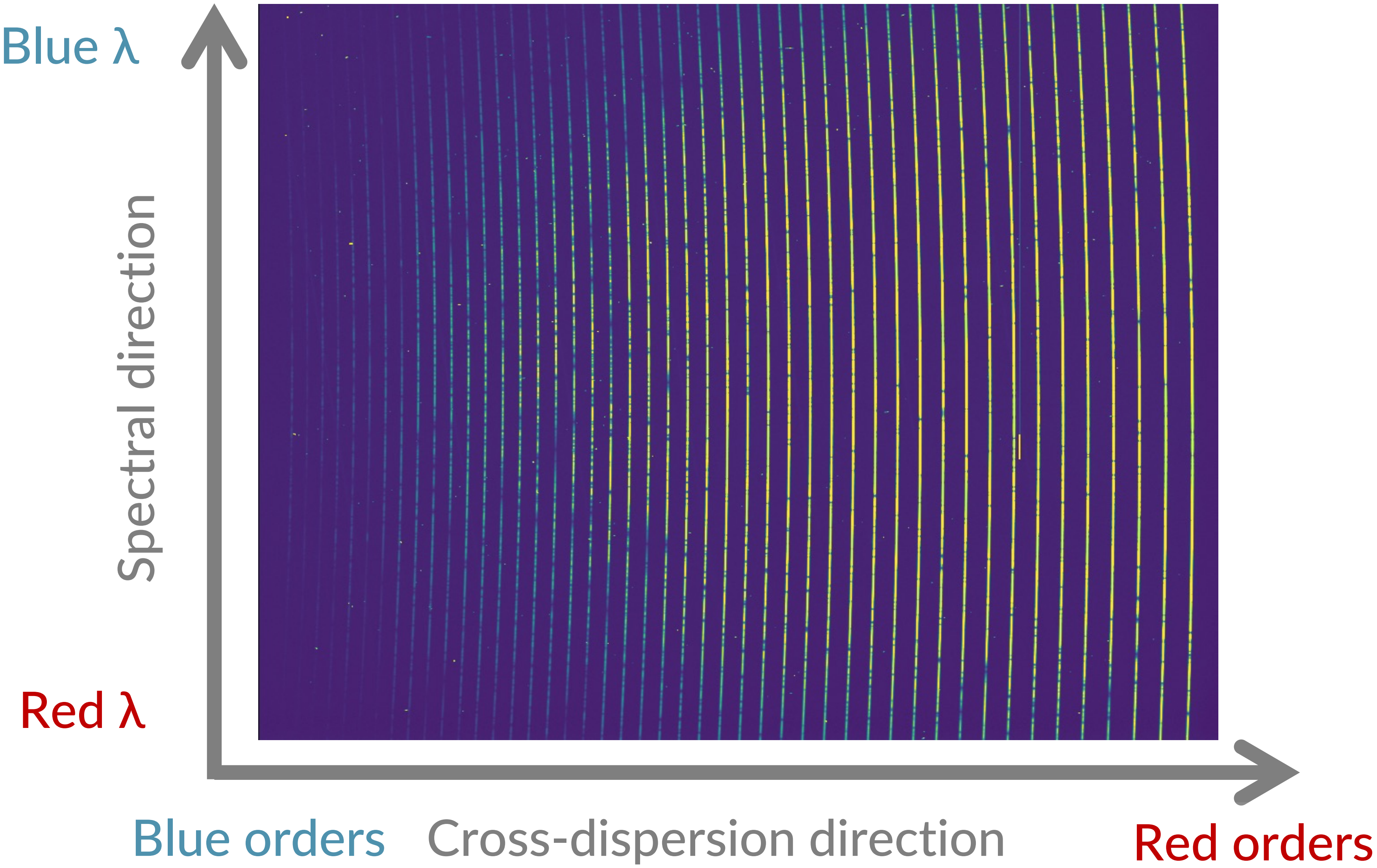




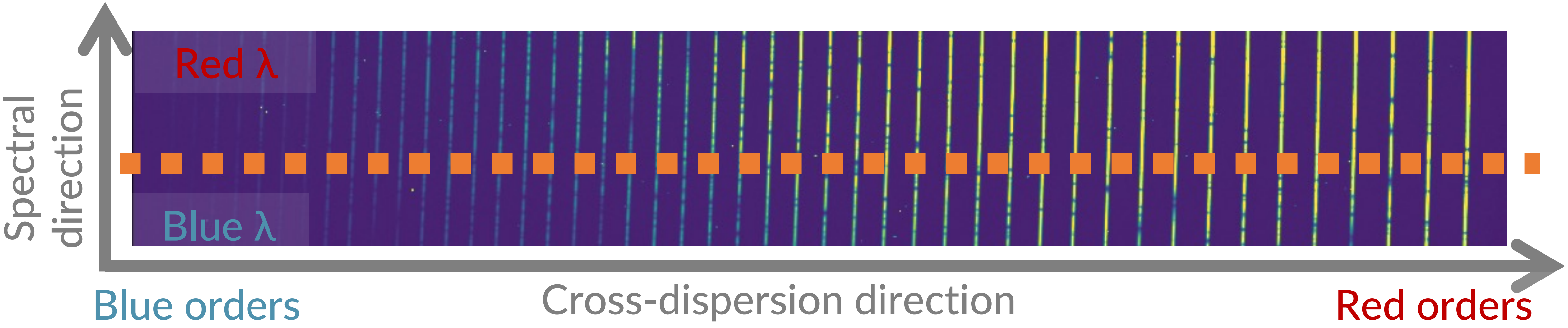
CAHA

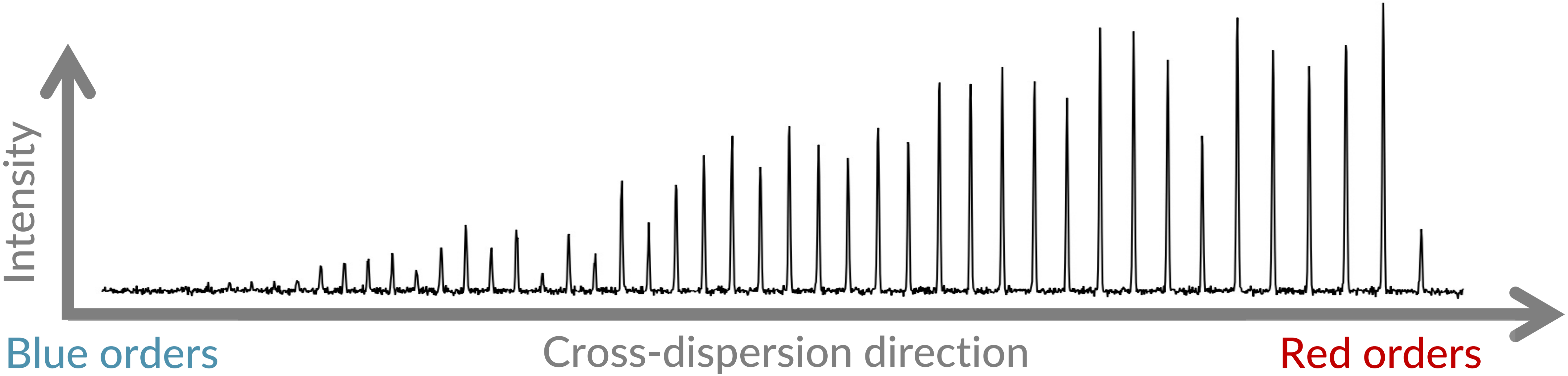
Extracting the spectrum

Example of a cross-dispersed spectrum

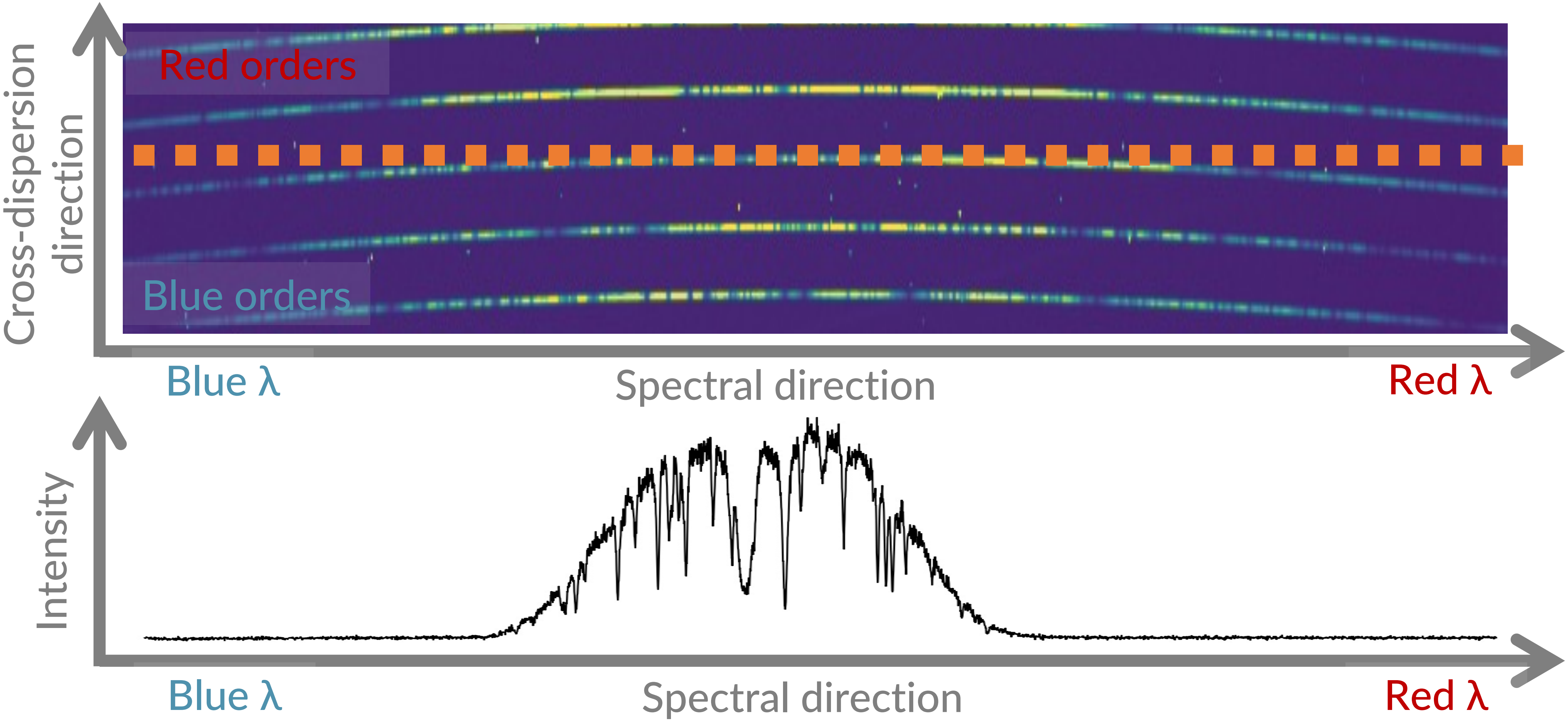


Example of a cross-dispersed spectrum



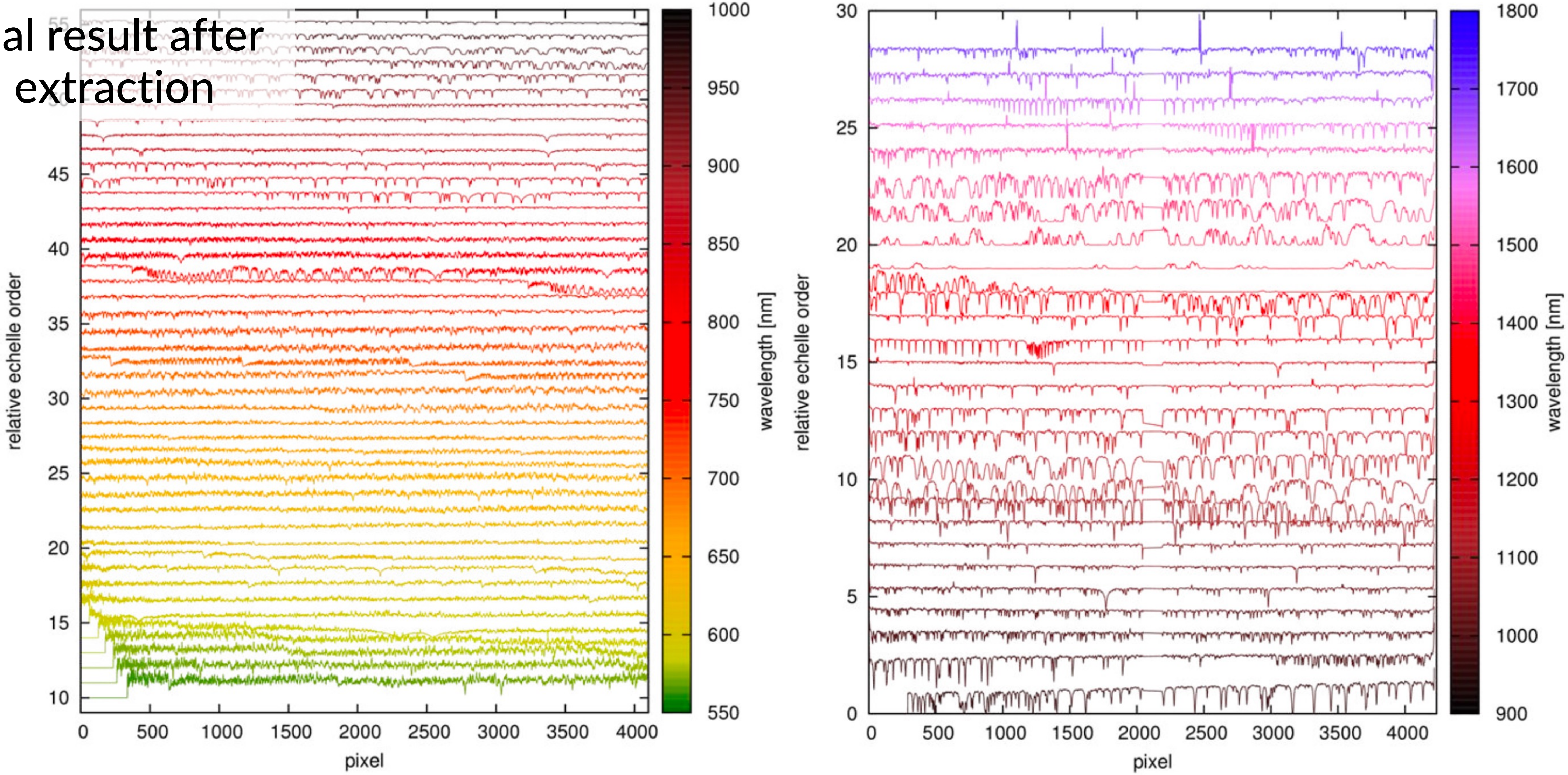


Example of a cross-dispersed spectrum



Example of a cross-dispersed spectrum

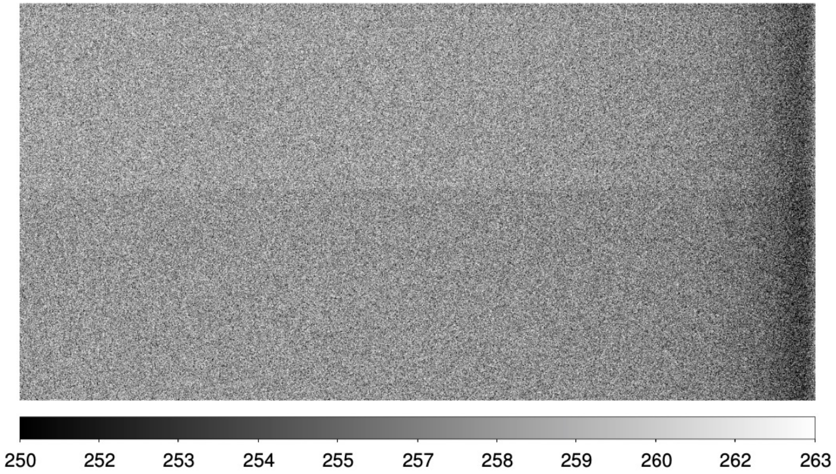
Final result after
full extraction



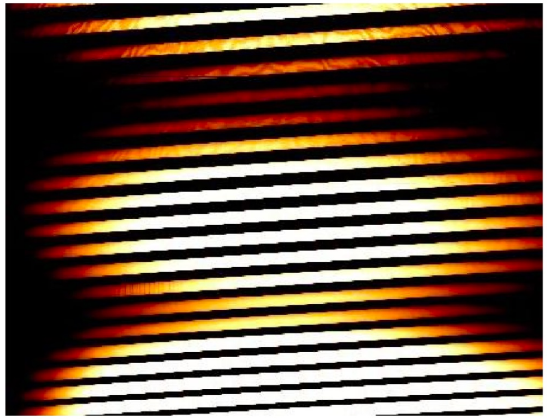
Quirrenbach et al. 2016

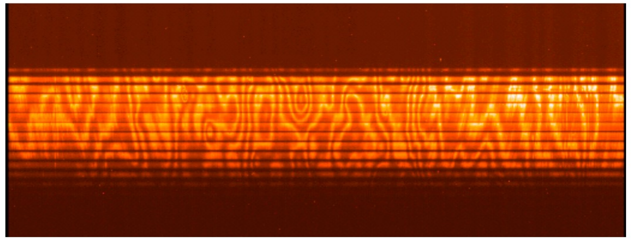
Calibrating the detector

- **Bias** correction
  - 0 s exposure time
  - Pre-/over-scan region in science observations



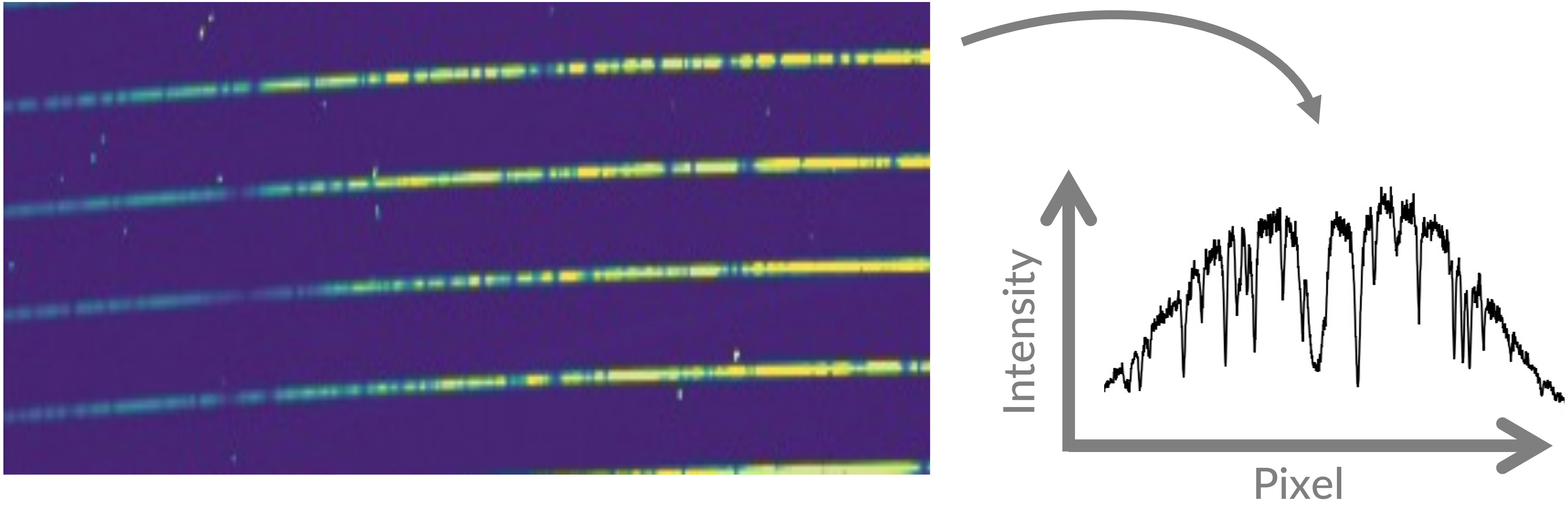
- **Flat fielding**
  - Pixel-to-pixel variations
  - Fringing (interference pattern)
  - Blaze function from the echelle grating
  - λ-dependent efficiency of the instrument
  - Dome flat, sky flat at twilight





Order definition & extraction

- Identify location and width of the spectrum
- Optimal extraction: weight by a smoothed 2D profile (Horne, 1986), as opposed to linear extraction

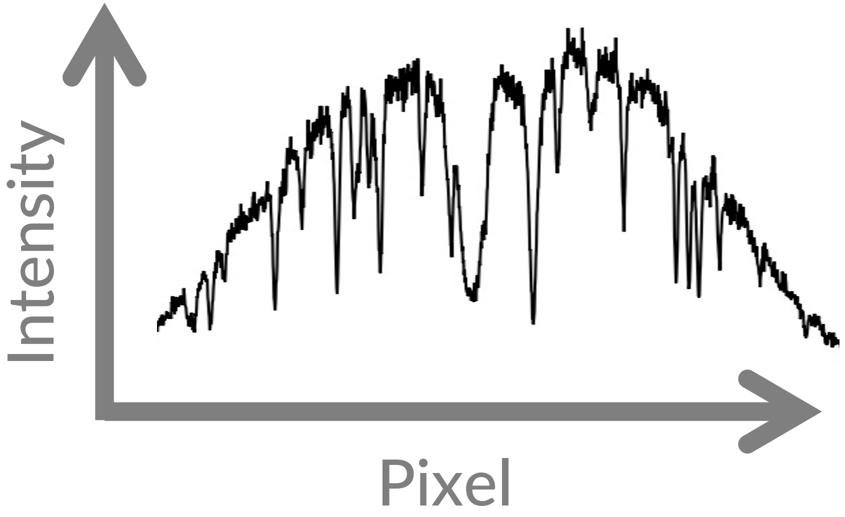


Wavelength calibration: from pixel to λ

- Associate a wavelength to each of the pixels along the spectral direction
- Requires a **reference spectrum** with known wavelengths

Sky absorption/emission lines
- Simultaneous, observer reference frame
- Not accurate
- Few lines in optical, more in near-infrared

Stellar spectral lines
- Simultaneous, stellar frame
- Need well characterised star



Wavelength

Wavelength calibration: from pixel to λ

- Associate a wavelength to each of the pixels along the spectral direction
- Requires a **reference spectrum** with known wavelengths

Sky absorption/emission lines

- Simultaneous, observer reference frame
- Not accurate
- Few lines in optical, more in near-infrared

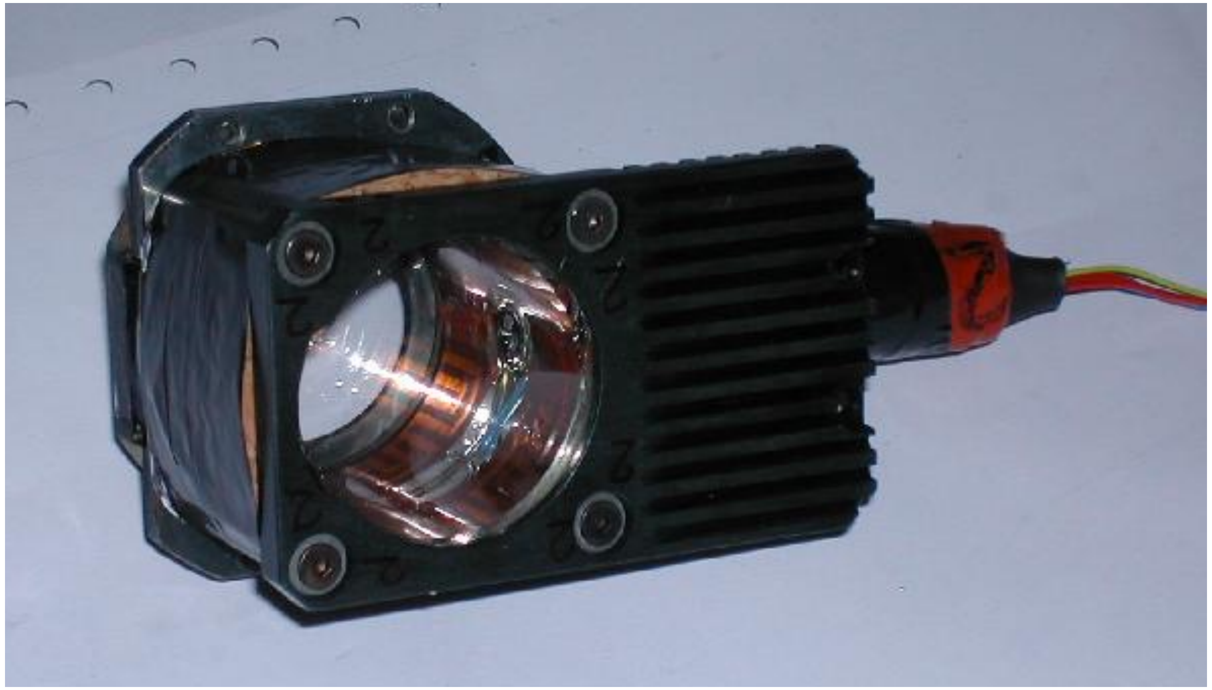
Stellar spectral lines

- Simultaneous, stellar frame
- Need well characterised star

Gas absorption cell in optical path

- Simultaneous, observer reference frame
- Reduces amount of photons from source
- Very accurate (m/s)

HARPS Iodine cell



https://www.eso.org/sci/facilities/lasilla/instruments/harps/inst/description.html

Wavelength calibration: from pixel to λ

- Associate a wavelength to each of the pixels along the spectral direction
- Requires a **reference spectrum** with known wavelengths

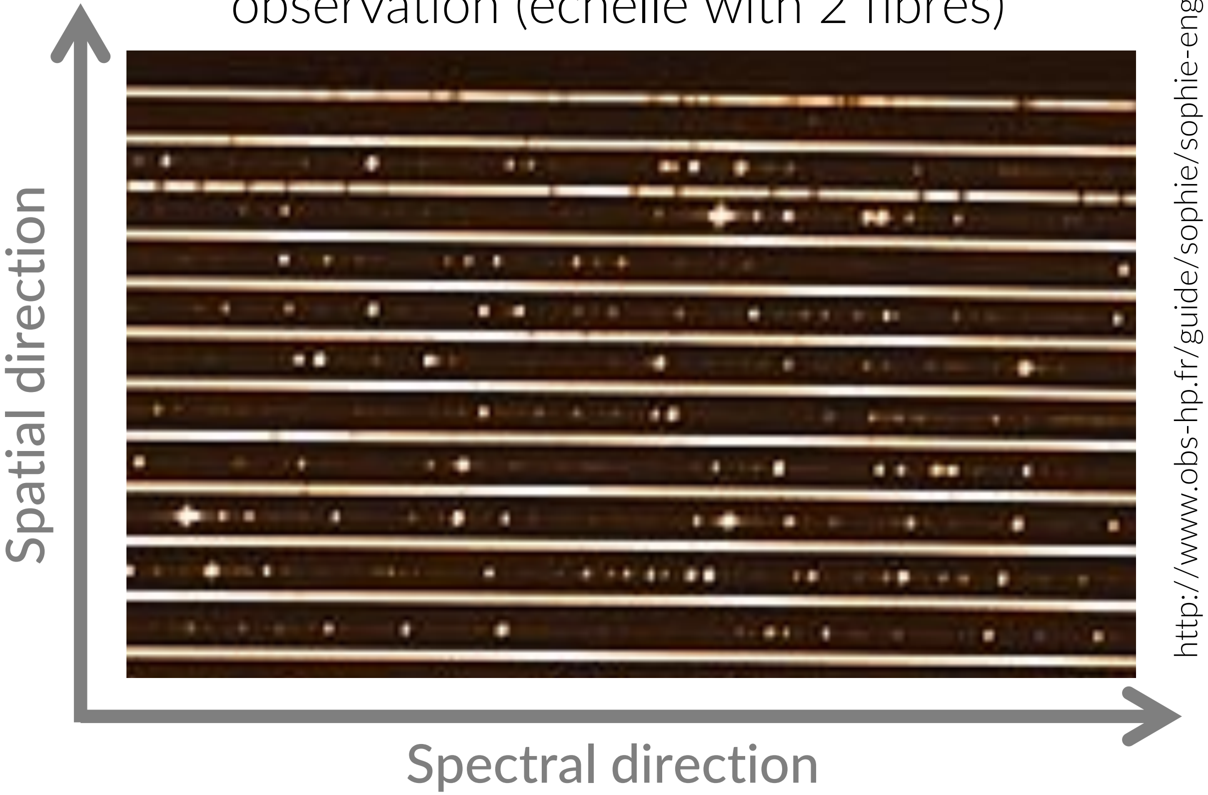
Sky absorption/emission lines
- Simultaneous, observer reference frame
- Not accurate
- Few lines in optical, more in near-infrared

Stellar spectral lines
- Simultaneous, stellar frame
- Need well characterised star

Gas absorption cell in optical path
- Simultaneous, observer reference frame
- Reduces amount of photons from source
- Very accurate (m/s)

Emission line lamps (arcs): Ar, Th, He, Ne, Cu…
- Not always simultaneous, observer reference frame
- Stable source, very accurate

Simultaneous target and emission lamp observation (echelle with 2 fibres)



http://www.obs-hp.fr/guide/sophie/sophie-eng.shtml

Wavelength calibration: from pixel to λ

- Associate a wavelength to each of the pixels along the spectral direction
- Requires a **reference spectrum** with known wavelengths

Sky absorption/emission lines
- Simultaneous, observer reference frame
- Not accurate
- Few lines in optical, more in near-infrared

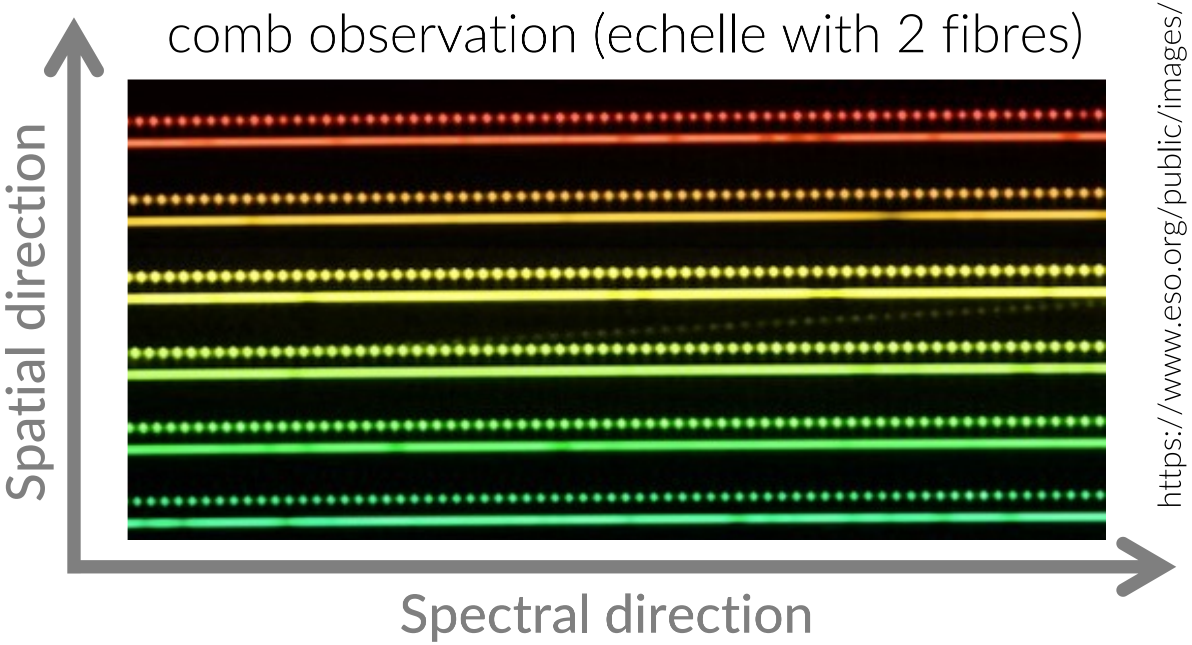
Stellar spectral lines
- Simultaneous, stellar frame
- Need well characterised star

Gas absorption cell in optical path
- Simultaneous, observer reference frame
- Reduces amount of photons from source
- Very accurate (m/s)

Emission line lamps (arcs): Ar, Th, He, Ne, Cu…
- Not always simultaneous, observer reference frame
- Stable source, very accurate

Fabry-Perot & Laser frequency comb
- Simultaneous, observer reference frame
- Very stable, very accurate

Simultaneous target and laser frequency comb observation (echelle with 2 fibres)



https://www.eso.org/public/images/ann12037a/

Wavelength calibration: from pixel to λ

- Associate a wavelength to each of the pixels along the spectral direction
- Requires a **reference spectrum** with known wavelengths

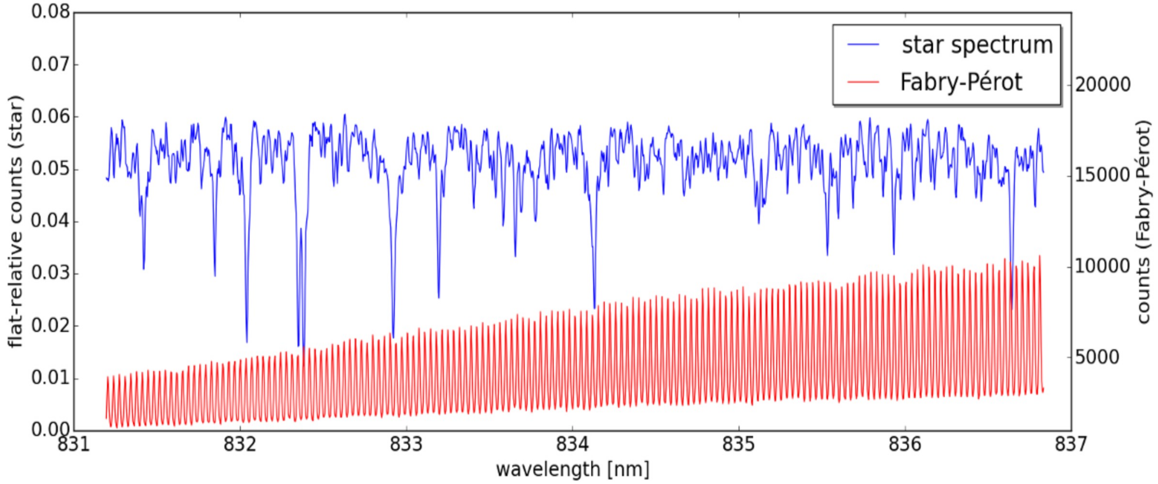
Sky absorption/emission lines
- Simultaneous, observer reference frame
- Not accurate
- Few lines in optical, more in near-infrared

Stellar spectral lines
- Simultaneous, stellar frame
- Need well characterised star

Gas absorption cell in optical path
- Simultaneous, observer reference frame
- Reduces amount of photons from source
- Very accurate (m/s)

Emission line lamps (arcs): Ar, Th, He, Ne, Cu…
- Not always simultaneous, observer reference frame
- Stable source, very accurate

Fabry-Perot & Laser frequency comb
- Simultaneous, observer reference frame
- Very stable, very accurate



Wavelength calibration: from pixel to λ

- Associate a wavelength to each of the pixels along the spectral direction
- Requires a **reference spectrum** with known wavelengths

Sky absorption/emission lines
- Simultaneous, observer reference frame
- Not accurate
- Few lines in optical, more in near-infrared

Stellar spectral lines
- Simultaneous, stellar frame
- Need well characterised star

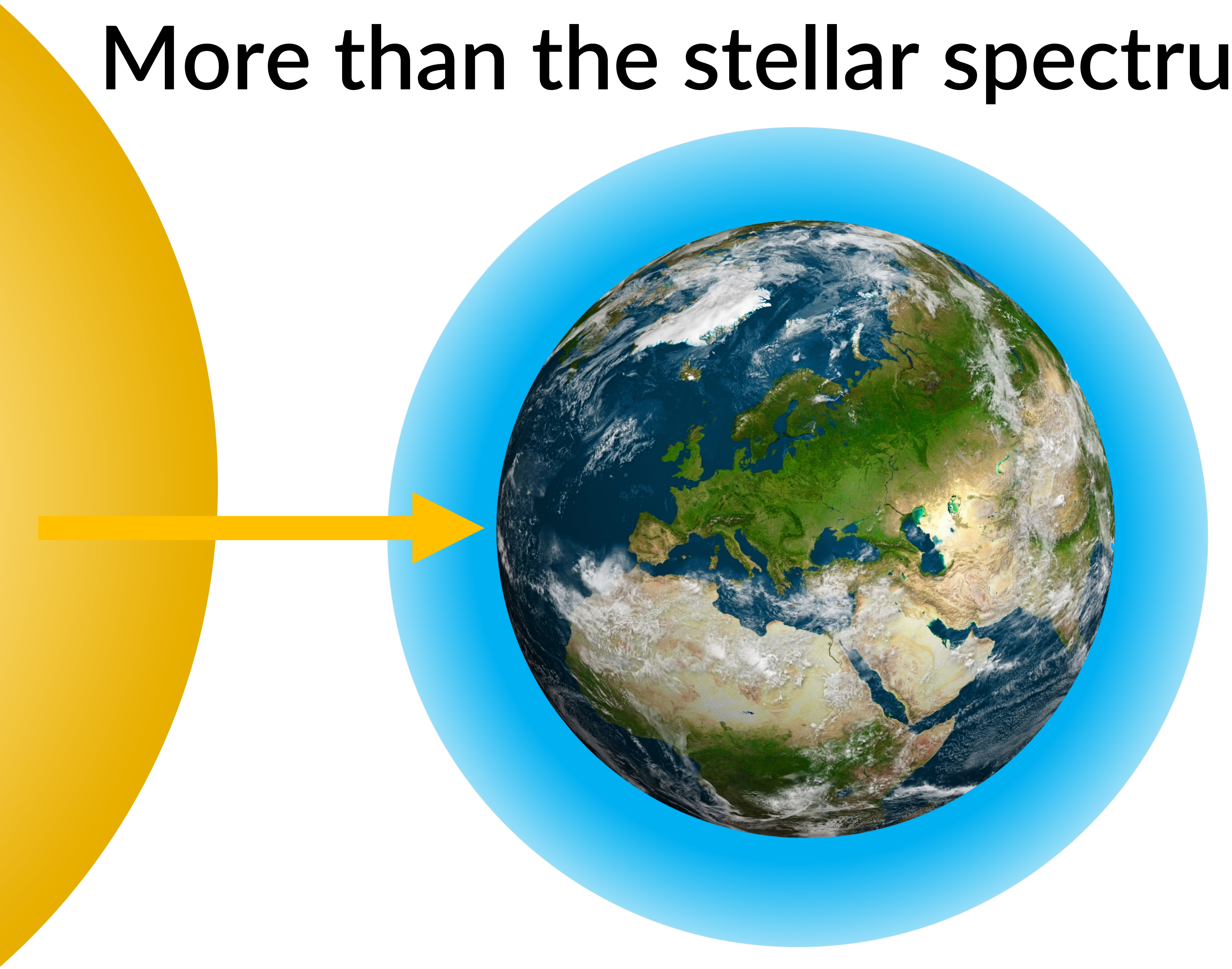
Gas absorption cell in optical path
- Simultaneous, observer reference frame
- Reduces amount of photons from source
- Very accurate (m/s)

Emission line lamps (arcs): Ar, Th, He, Ne, Cu…
- Not always simultaneous, observer reference frame
- Stable source, very accurate

Fabry-Perot & Laser frequency comb
- Simultaneous, observer reference frame
- Very stable, very accurate

Instrument flexures, seeing/pointing variations, temperature and pressure changes… all influence the wavelength solution

More than the stellar spectrum…

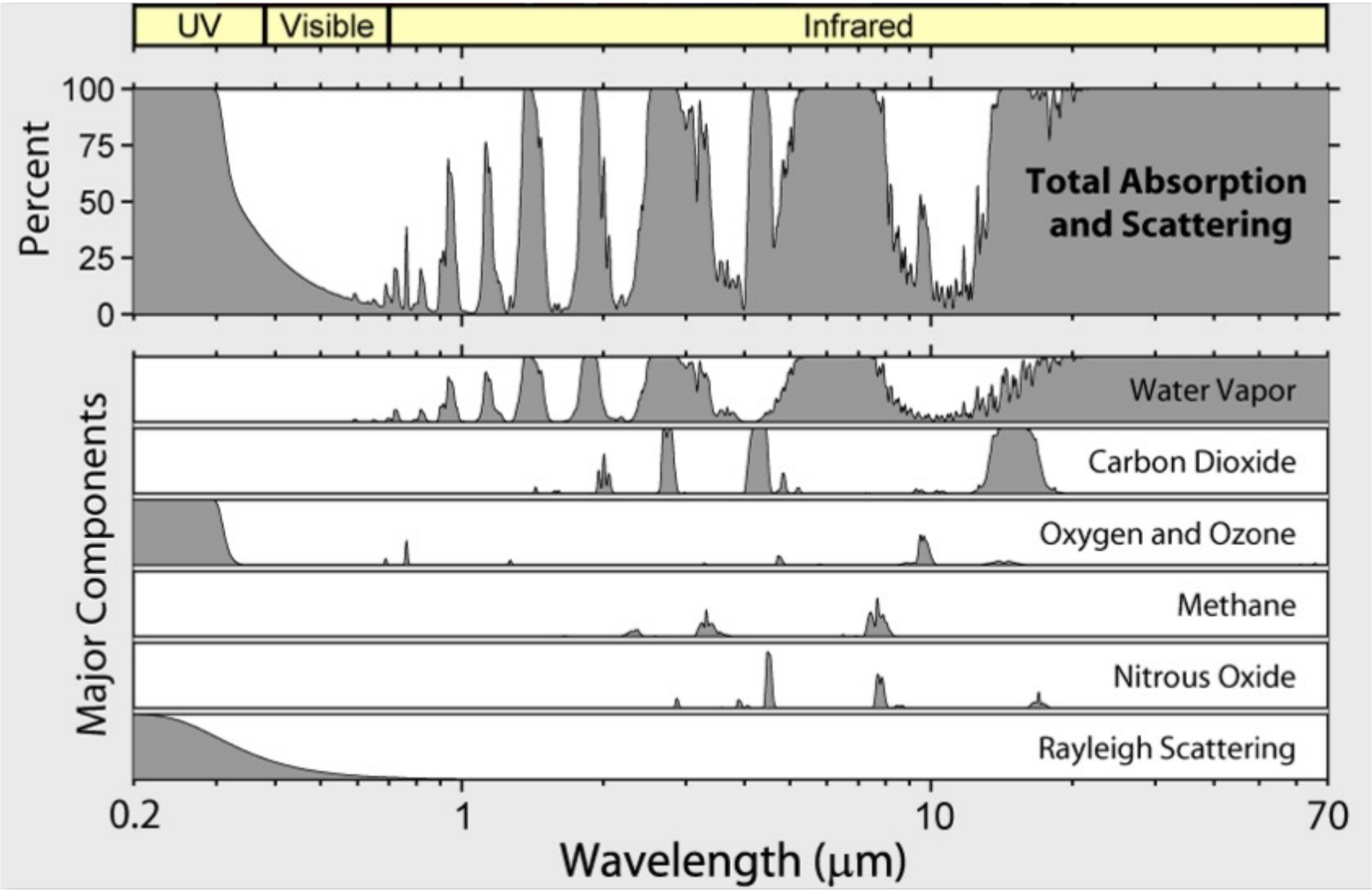


- High-resolution spectrographs are on the ground
- We observe stellar light through the Earth’s atmosphere

NASA Johnson Space Center

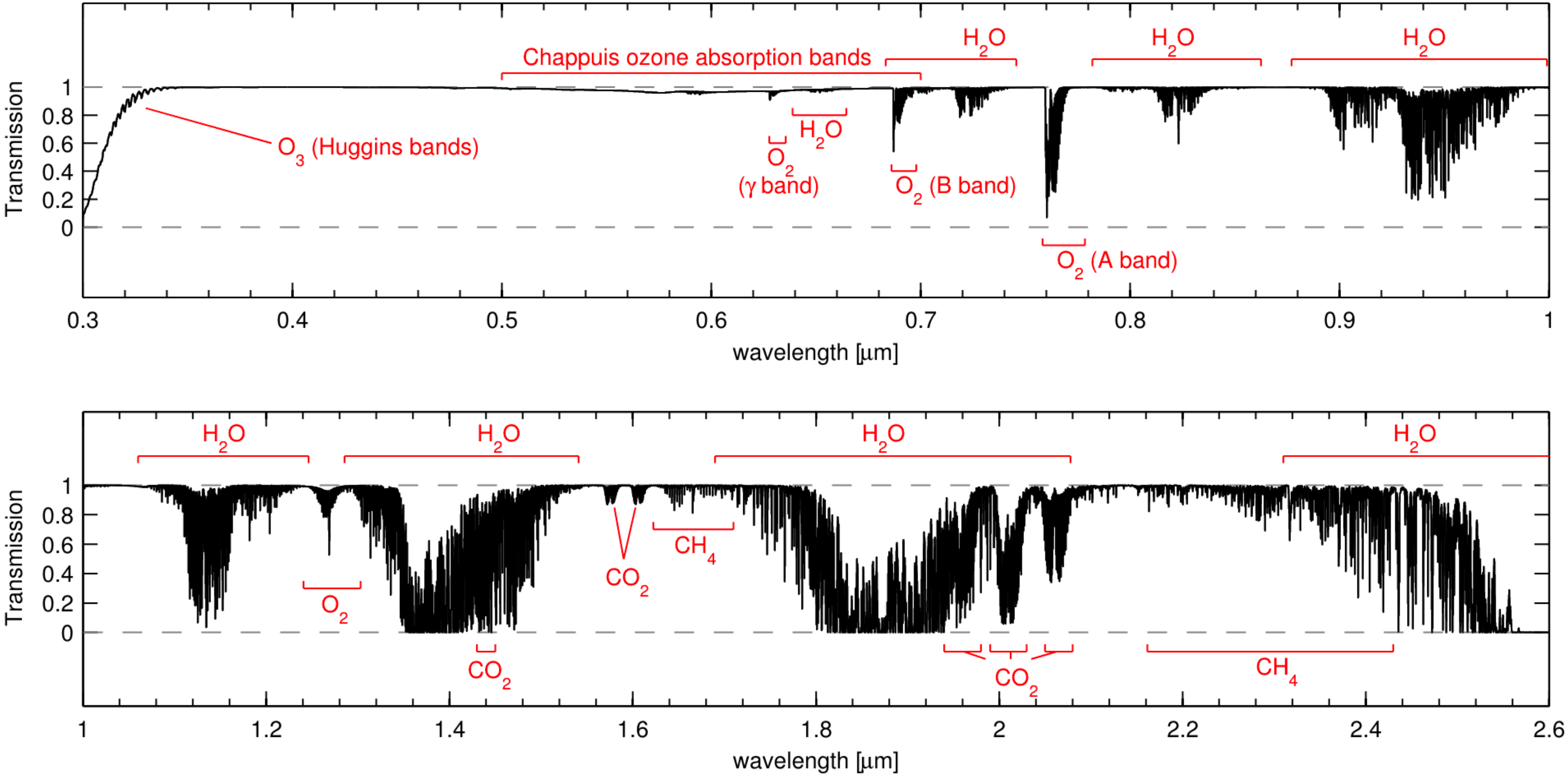
Atmospheric transmission

- Atmospheric transmission strongly depends on λ
- Source spectrum will be imprinted by Earth's transmission spectrum
- At visible wavelengths Earth atmosphere almost transparent

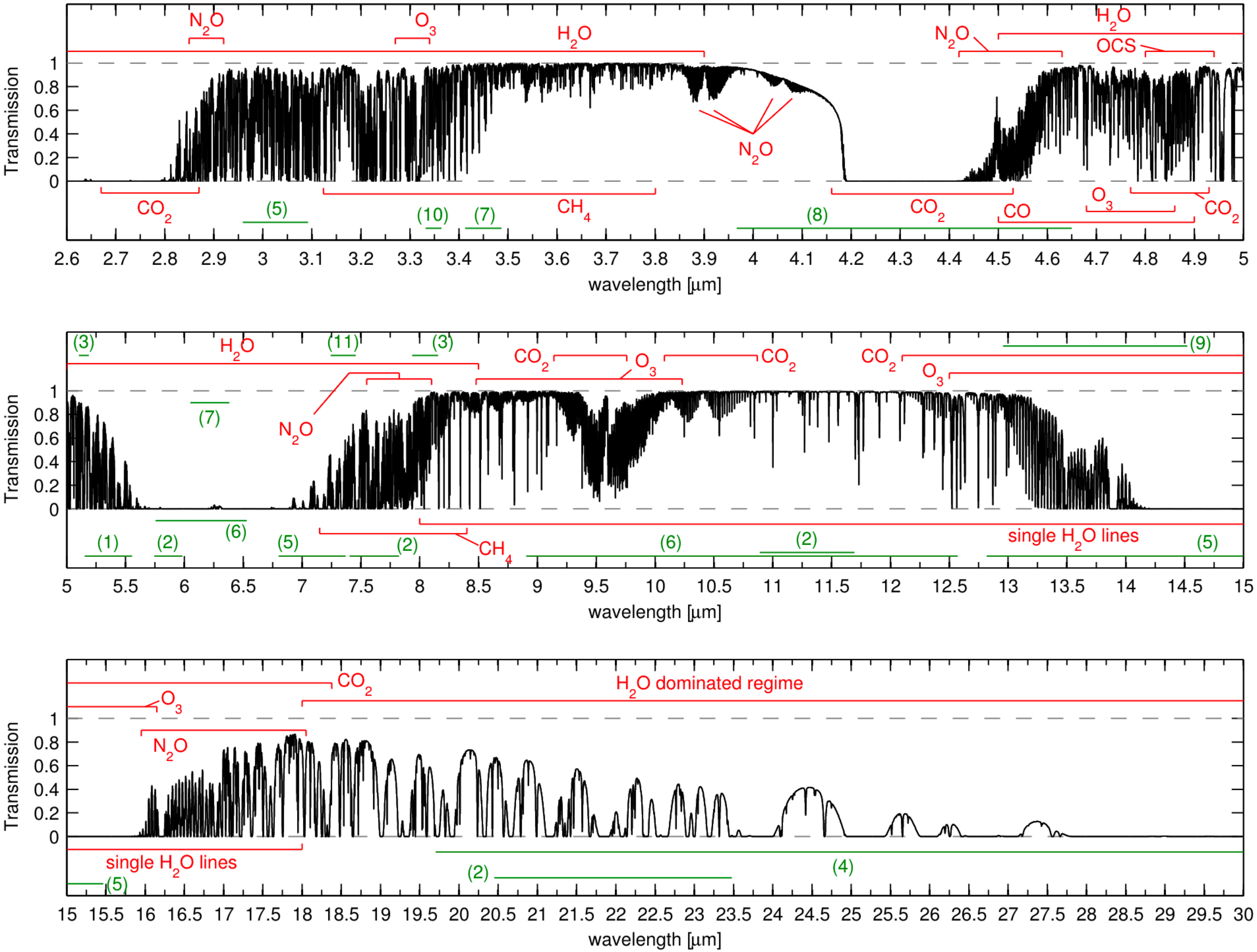


Earth’s telluric absorption spectrum

Smette et al. 2015

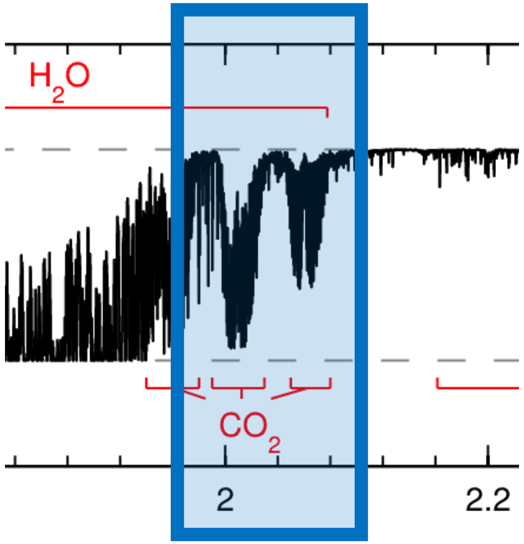


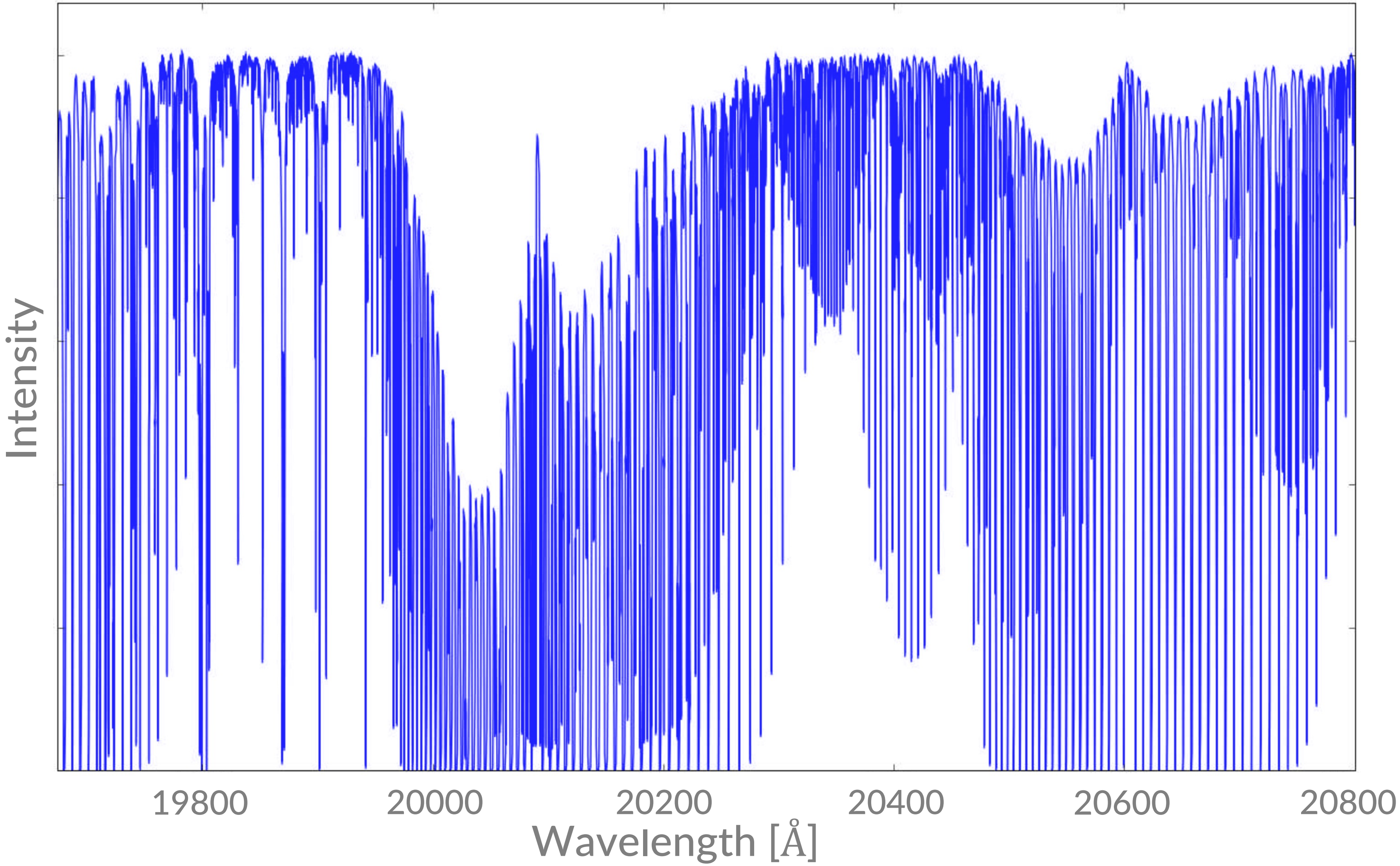
Earth's telluric absorption spectrum



Smette et al. 2015

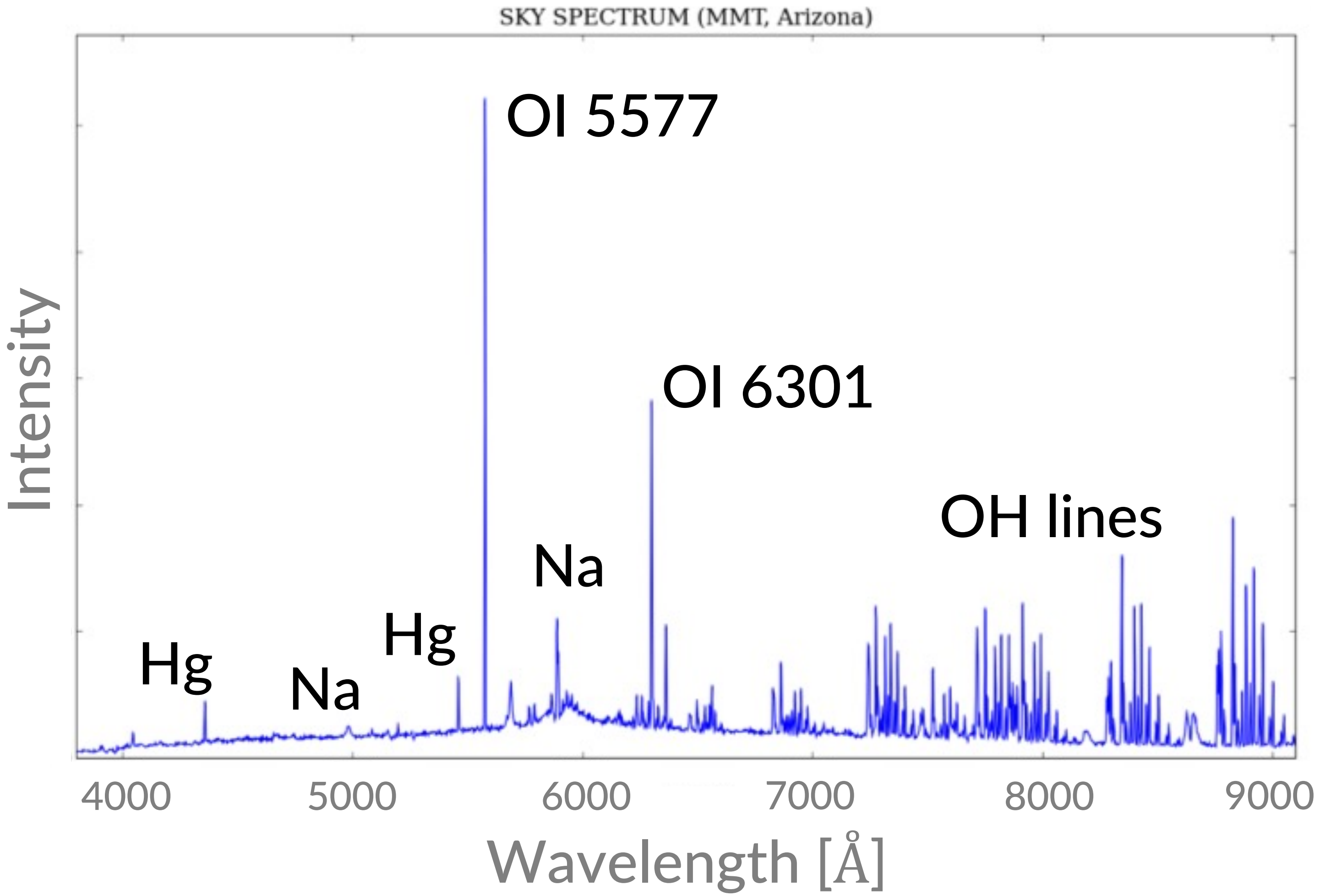
Earth’s telluric absorption spectrum





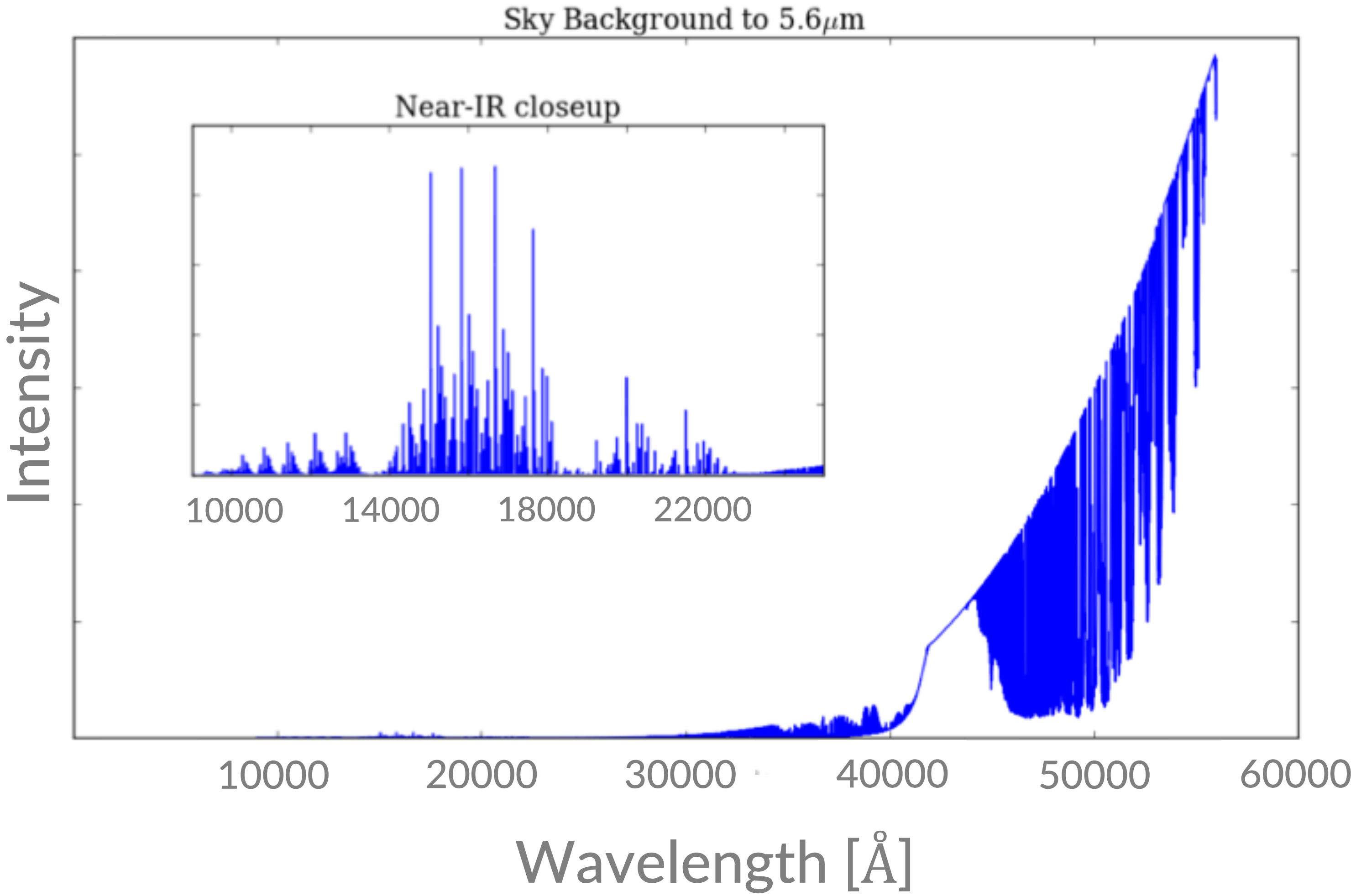
Sky background: Optical

- Background has contributions from many sources
  - Air glow: Strong, discrete emission lines (fluorescence of atmospheric OH, O, Na, & city lights Hg)
  - Zodiacal light
  - Sun/Moonlight
  - Auroare
  - Light pollution
  - Thermal emission from sky, telescope and buildings
  - Non-resolved astronomical background



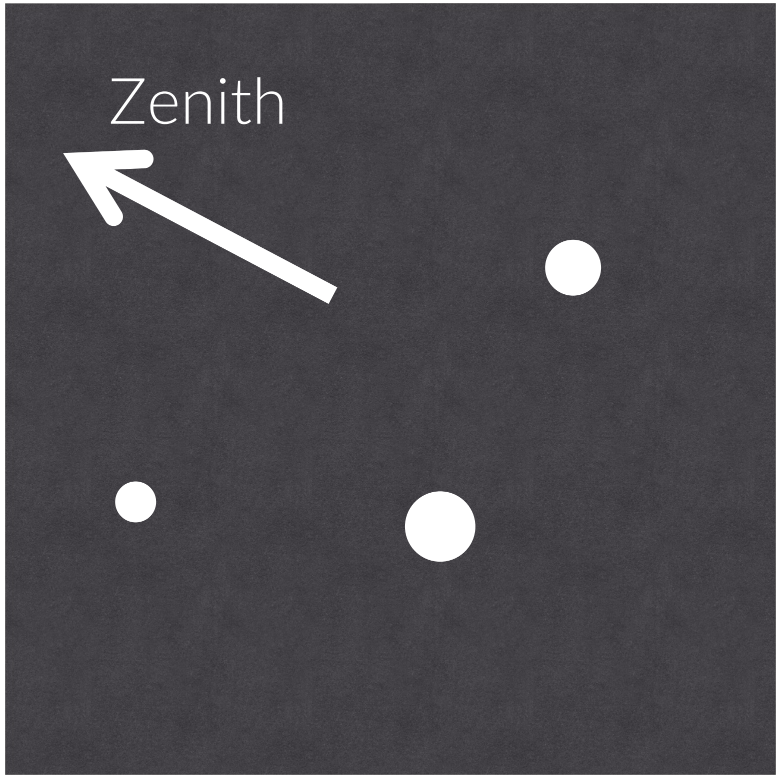
Sky background: Infrared

- Thermal emission from the sky, ground and telescope dominates
- Observations become very challenging for. $\lambda > 5\ \mu m$



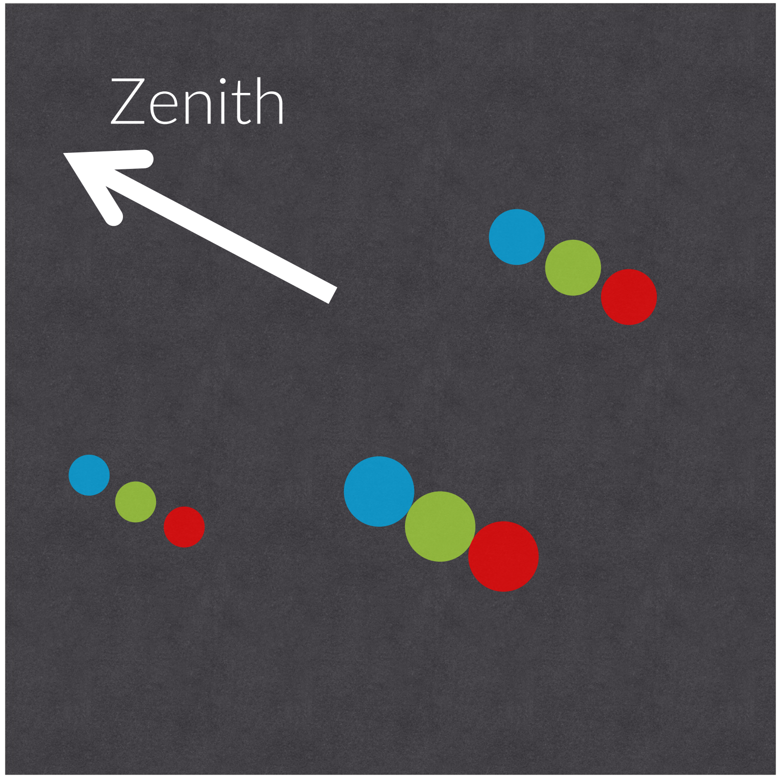
Atmospheric dispersion

- Earth’s atmosphere refracts source light ⟹ Sky position of the source is λ-dependent!



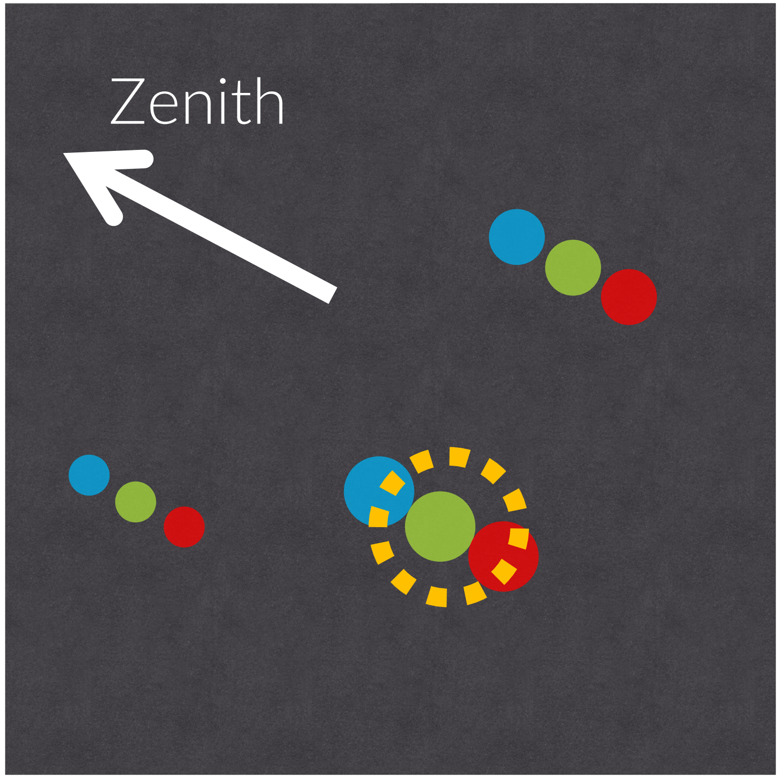
Atmospheric dispersion

- Earth's atmosphere refracts source light ⟹ Sky position of the source is λ-dependent!
  - Index of refraction depends on wavelength, temperature, pressure, water vapour
  - Dispersion happens along the horizon-zenith direction (airmass)
  - Dispersion larger for shorter wavelengths
  - Dispersion direction changes with time
- Affects acquisition
- Atmospheric Dispersion Compensator (ADC)



Atmospheric dispersion

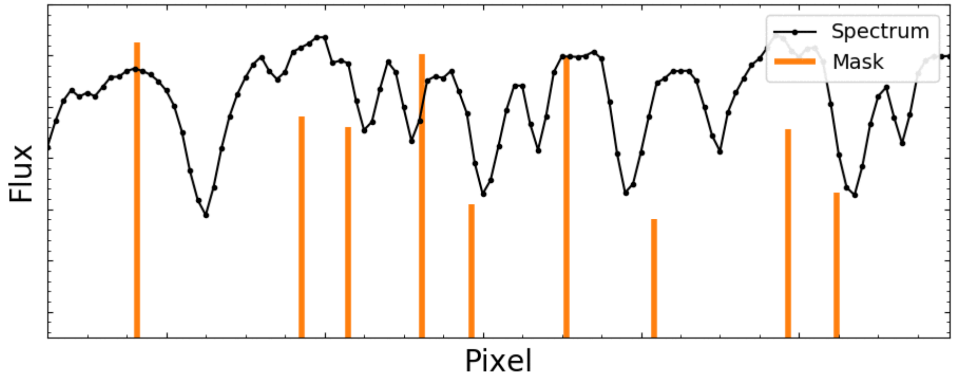
- Earth's atmosphere refracts source light ⟹ Sky position of the source is λ-dependent!
    - Index of refraction depends on wavelength, temperature, pressure, water vapour
    - Dispersion happens along the horizon-zenith direction (airmass)
    - Dispersion larger for shorter wavelengths
    - Dispersion direction changes with time
- Affects acquisition
- Atmospheric Dispersion Compensator (ADC)

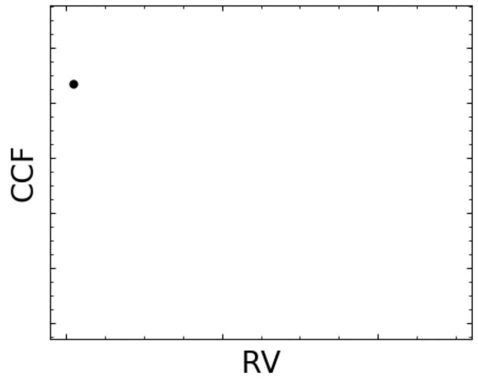


Extracting time-series data from the reduced spectra

Measuring RVs

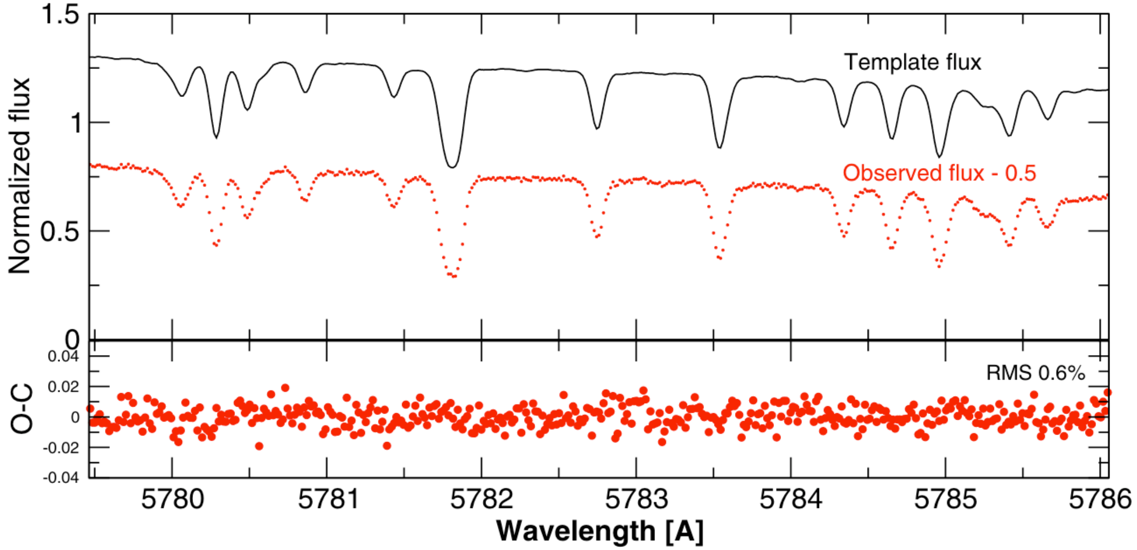
Cross-correlation function





e.g.
Queloz 1995
Baranne et al. 1996
Pepe et al. 2002

Template matching



Anglada-Escudé & Butler 2012

e.g.
Anglada-Escudé & Butler 2012
Astudillo-Defru et al. 2015
Zechmeister et al. 2018

Spectrum RV content

How does the RV precision depend on the spectral line shape?

- Continuum signal-to-noise
- Depth
- Width

→ Better RV precision for deep, narrow lines with high S/N

→ ↑ *vsini* or ↓ resolution will reduce the RV precision

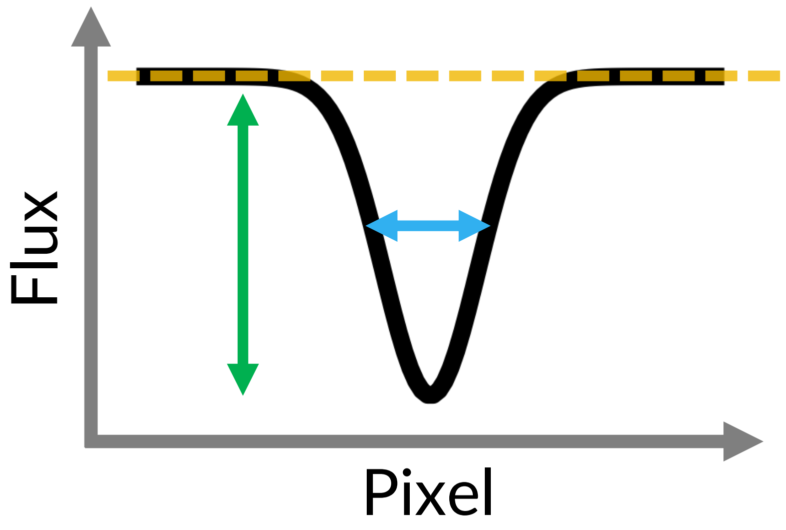
e.g. Connes 1985, Bouchy et al. 2001, Lovis & Fischer 2010

Line width

$$\sigma_{\mathrm{RV}} \sim \frac{\sqrt{\mathrm{FWHM}}}{C \cdot \mathrm{S/N}}$$

Line depth
(contrast = depth / continuum)

Continuum
signal-to-noise
ratio



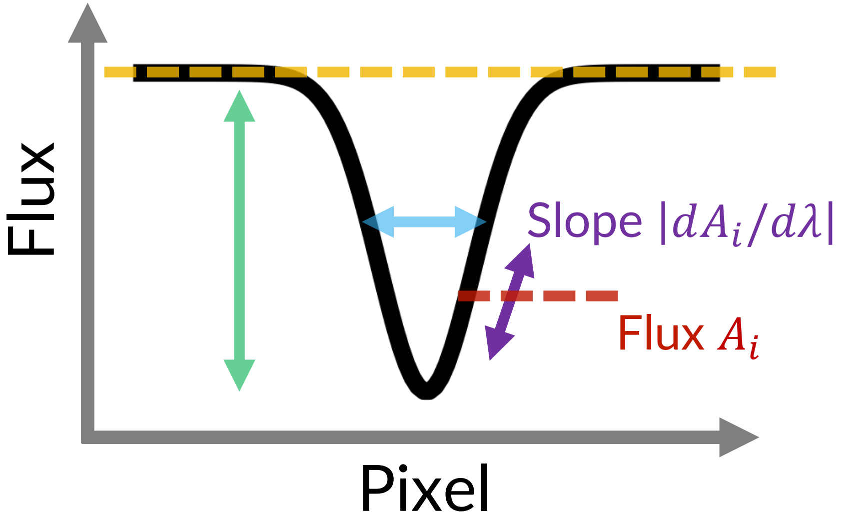
Spectrum RV content

RV precision for pixel i :

$$\sigma_{\mathrm{RV},\,i} = c\,\frac{\sqrt{A_i + \sigma_D^2}}{\lambda_i \cdot |dA_i/d\lambda|}$$

→ Photon noise $\sqrt{A_i}$ + Detector noise σ_D

→ Spectrum slope



RV precision for full spectrum (all pixels):

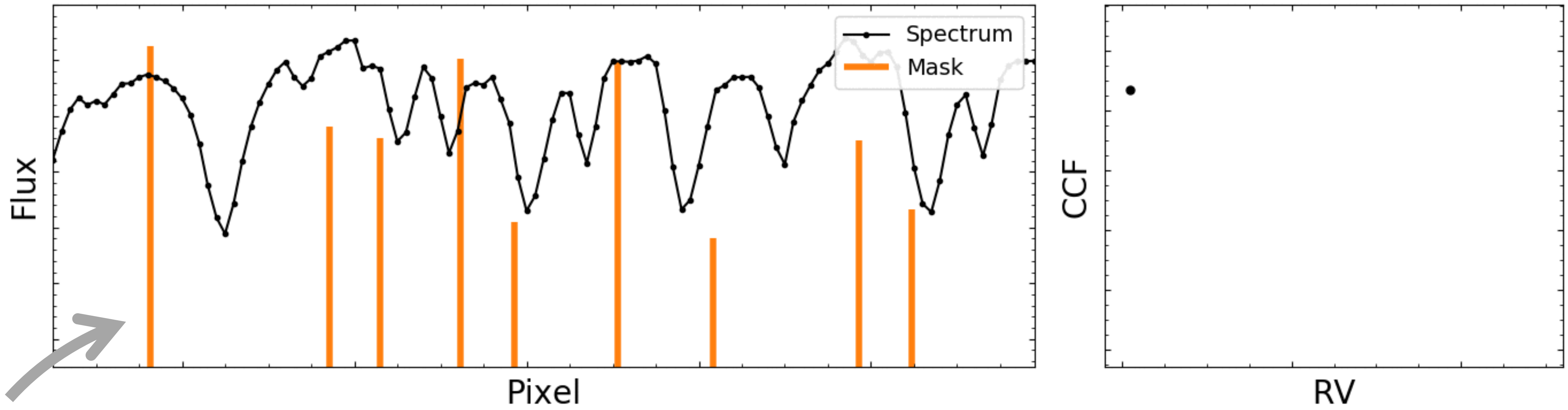
$$\sigma_{\mathrm{RV}} = c\left(\sum_i \frac{\lambda_i^2 \cdot |dA_i/d\lambda|^2}{A_i + \sigma_D^2}\right)^{-1/2}$$

A_i : Flux in pixel i
λ_i: Wavelength in pixel i
$dA_i/d\lambda$: Spectrum shape, slope
σ_D : Detector noise

→ The steeper the spectrum, the higher the RV content

Connes 1985, Bouchy et al. 2001

Cross-correlation function (CCF)



Spectrum ⊗ Mask

Cross-correlation function

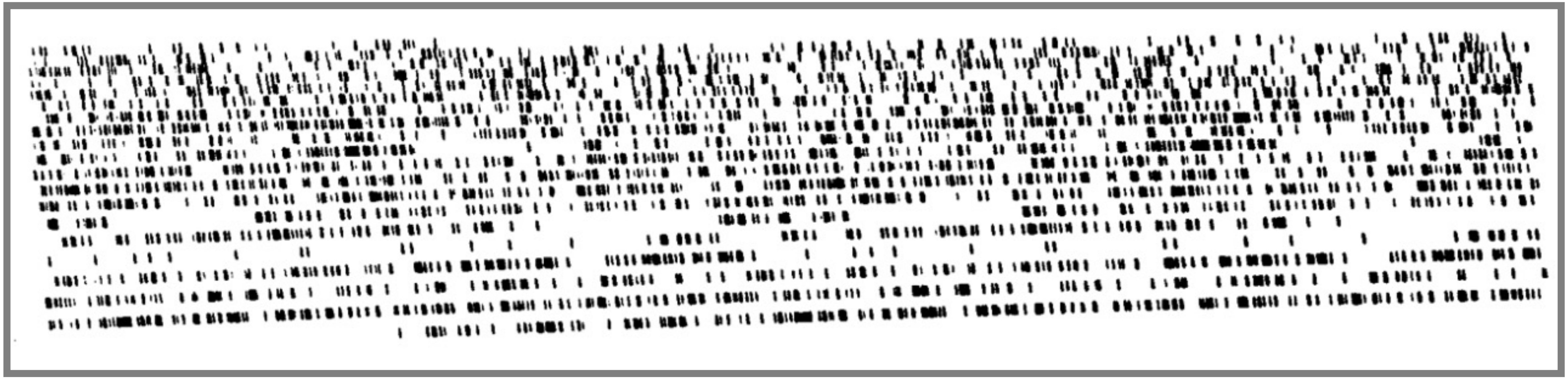
Weighted binary mask

$$\mathrm{CCF}(v) = \sum_{l=1}^{m}\sum_{x=1}^{n} w_l \cdot f_x \cdot \Delta_{xl}(v)$$

w_l : Line weight
f_x : Pixel flux
Δ_{xl} : Overlap between line & pixel
v : Doppler shift of the mask

m : Number of mask lines
n: Number of spectrum pixels

CCF binary masks

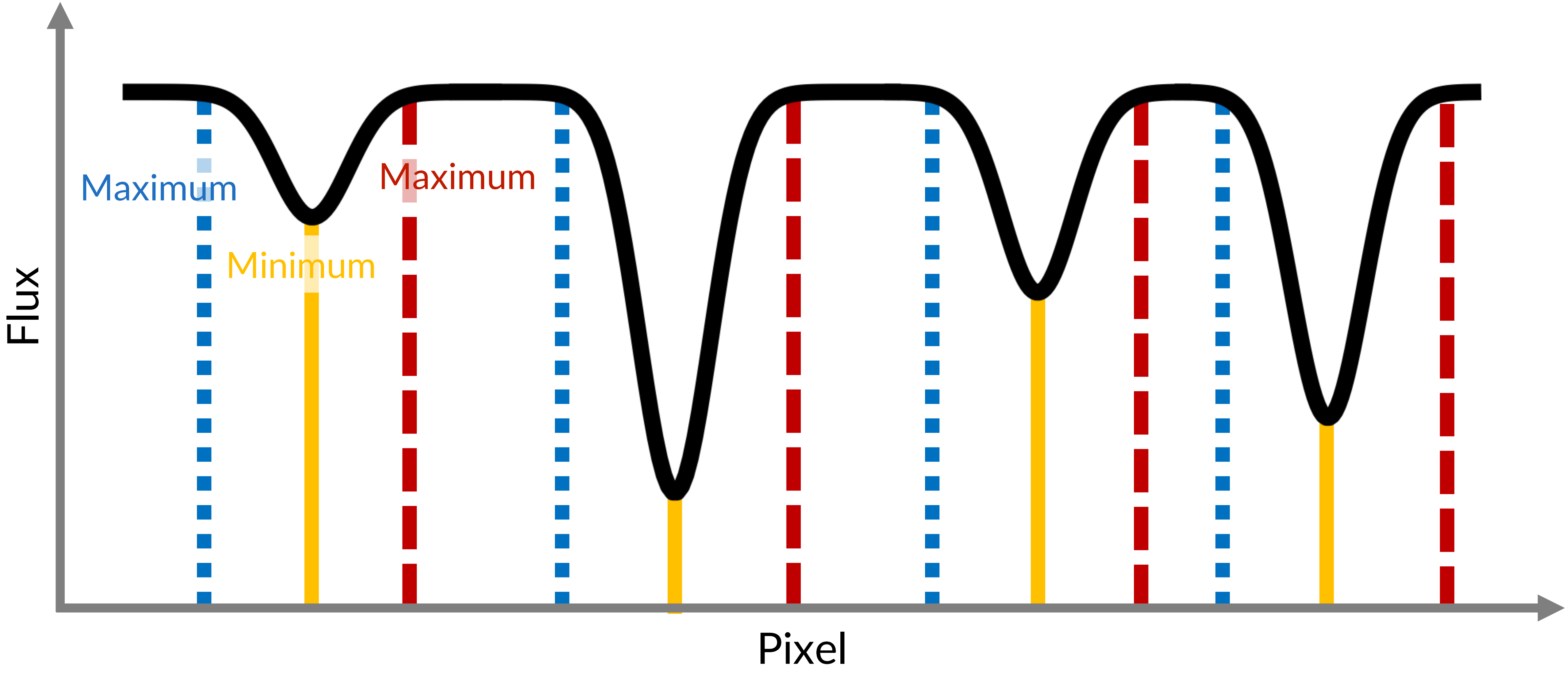


Baranne et al. 1979

Spatial filter or mask (negative) used by CORAVEL, derived from Arcturus (~3000 lines, size ~70 mm x 14 mm)

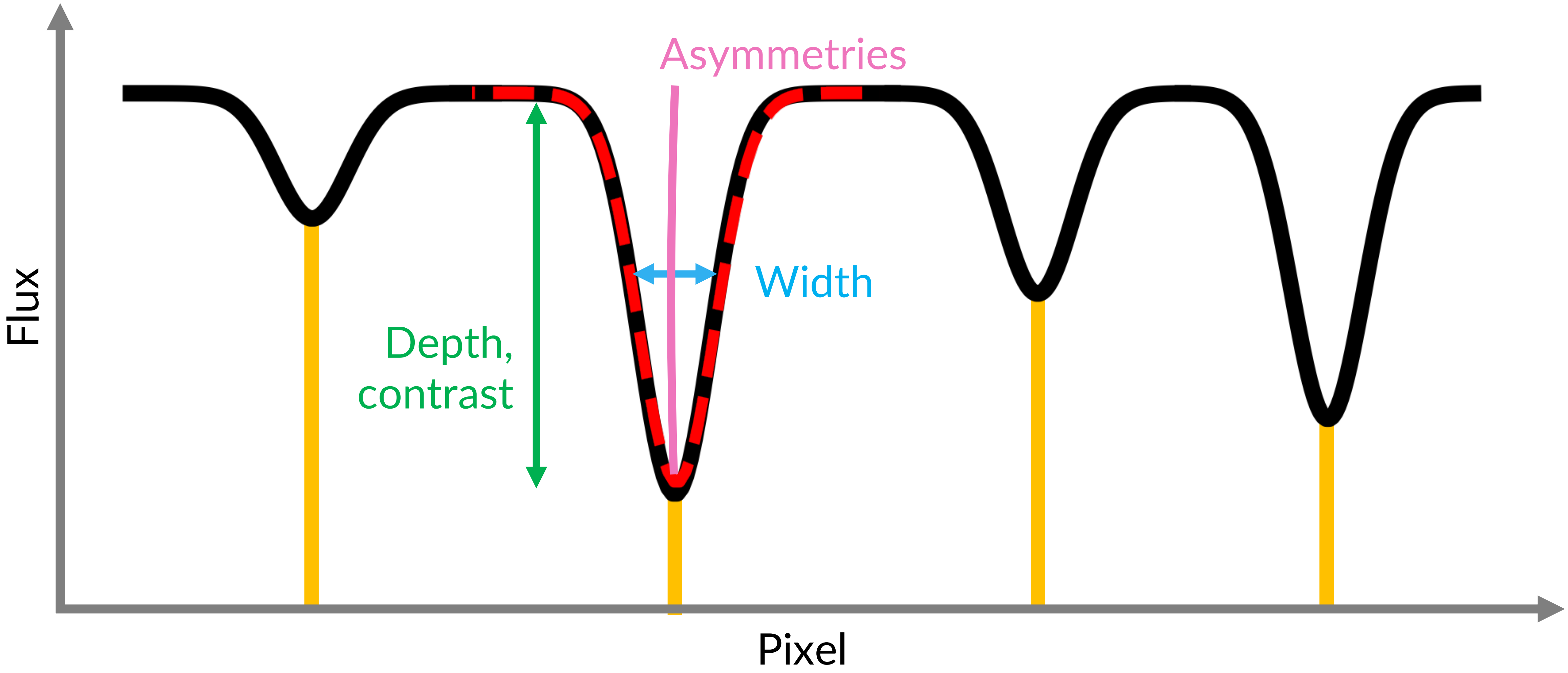
Creating weighted binary masks

On a high S/N stellar template, identify lines by finding minima



Creating weighted binary masks

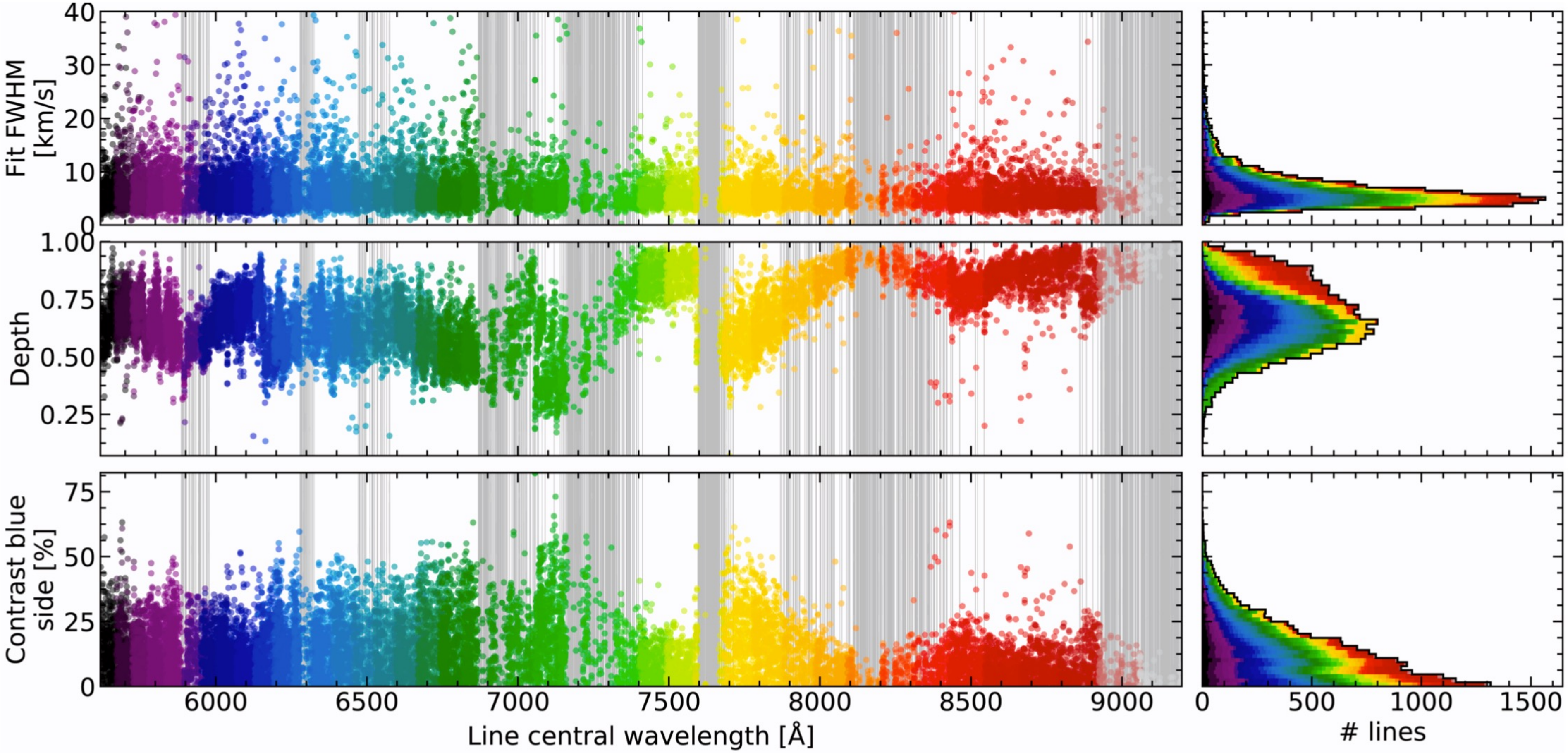
Characterise line shape



Creating weighted binary masks

Select “good” lines based on their shape properties (deep, narrow, symmetric)

Line weight ~ depth (and ~ 1/width)



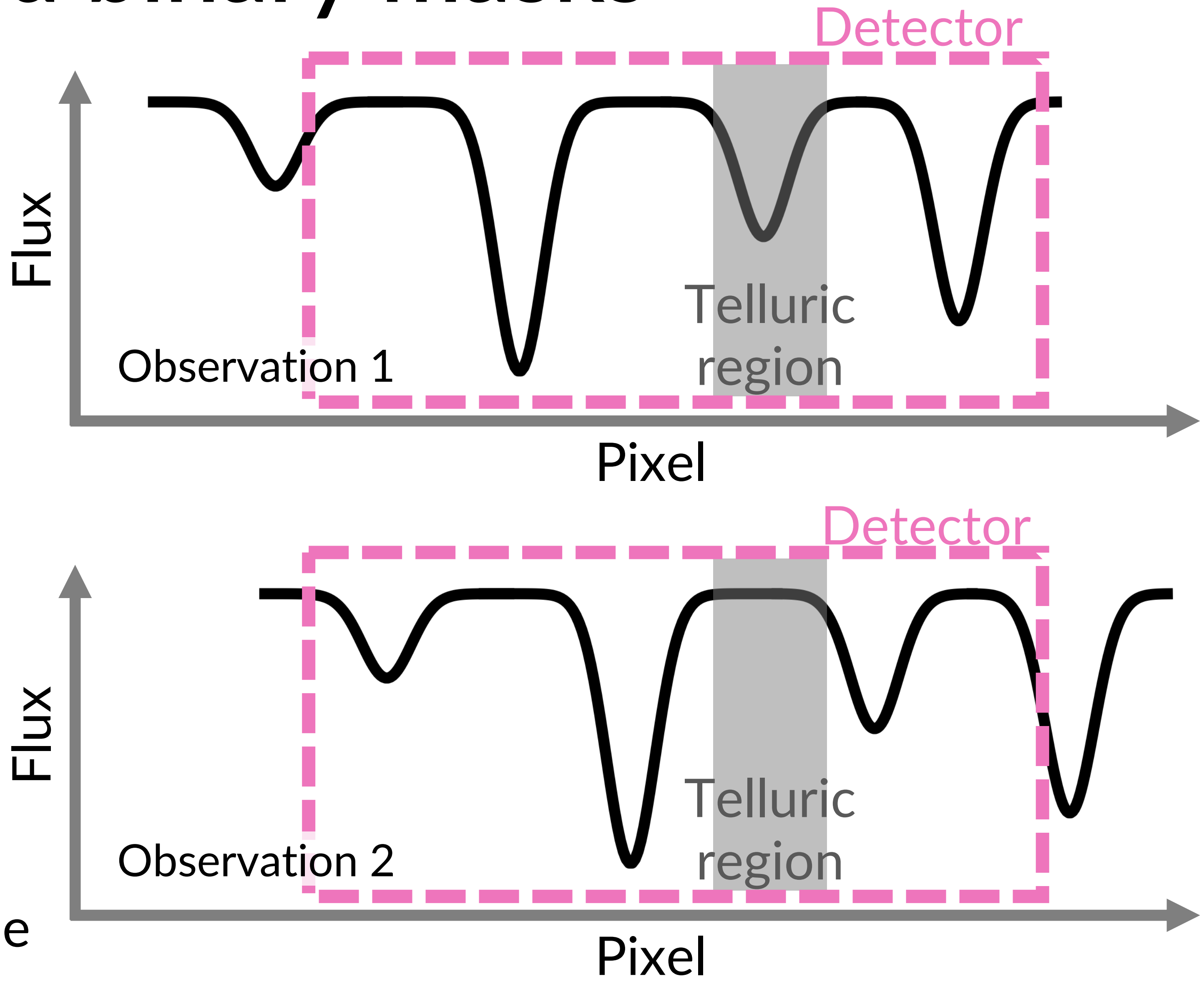
Lafarga et al. 2020, line selection for an M dwarf mask

Creating weighted binary masks

Different lines "visible" for different observations of the same star due to different shifts

- Earth barycentric movement
- Instrumental shifts
- Different line overlap with telluric regions
- Lines "out" of the detector

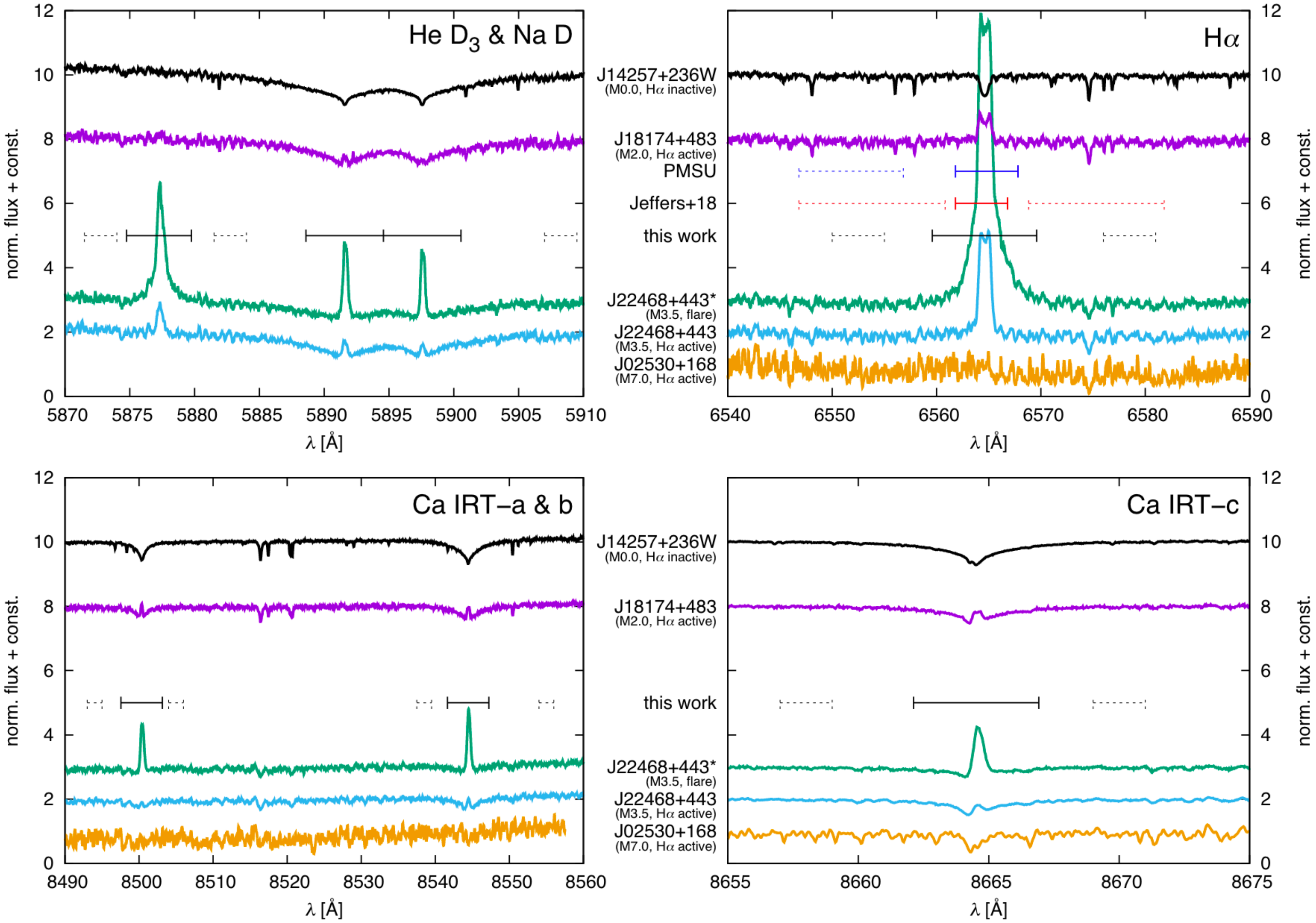
→ Consider line "visibility" throughout the observing times



In observer reference frame

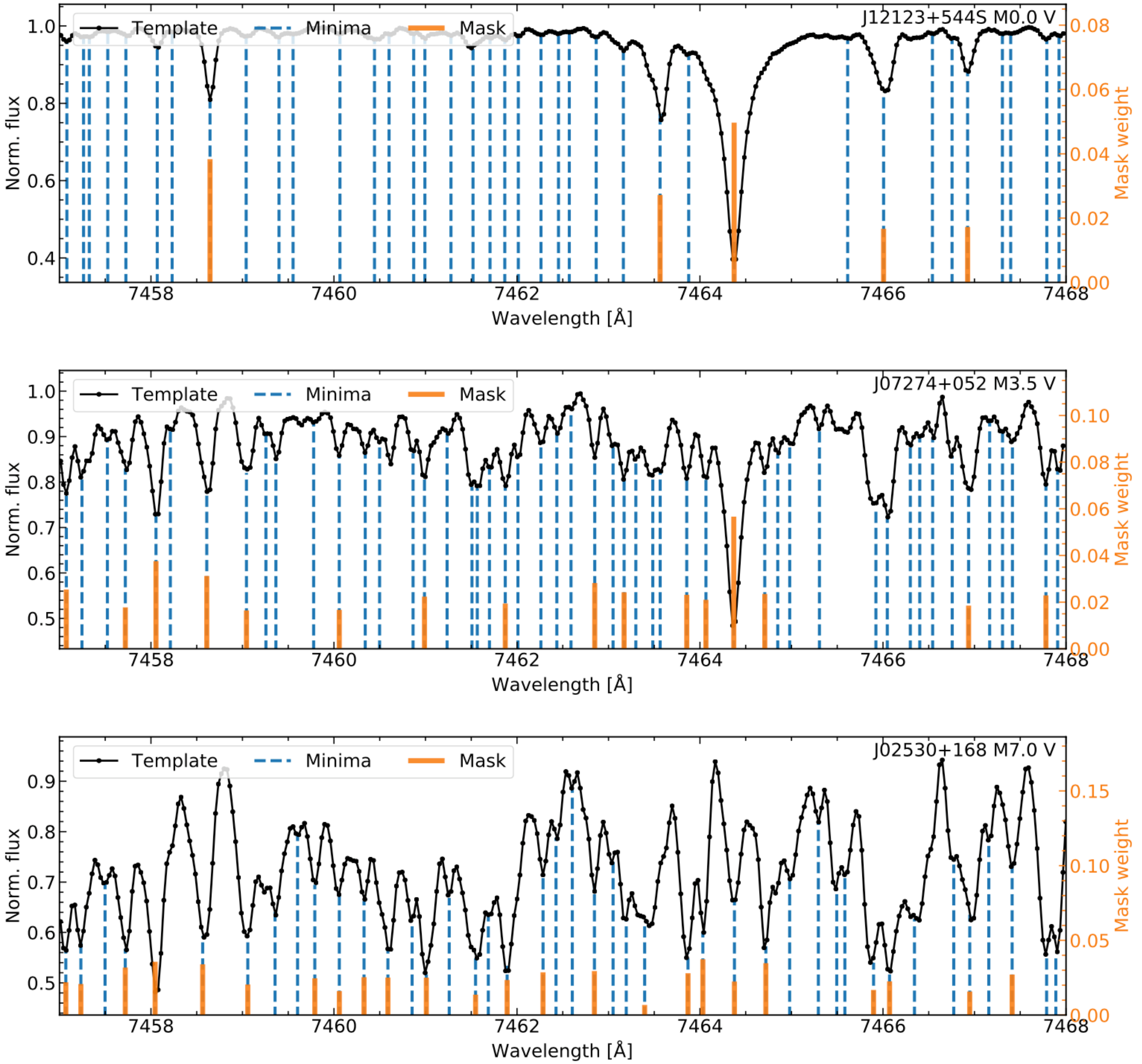
Creating weighted binary masks

Chromospheric lines
(emission)

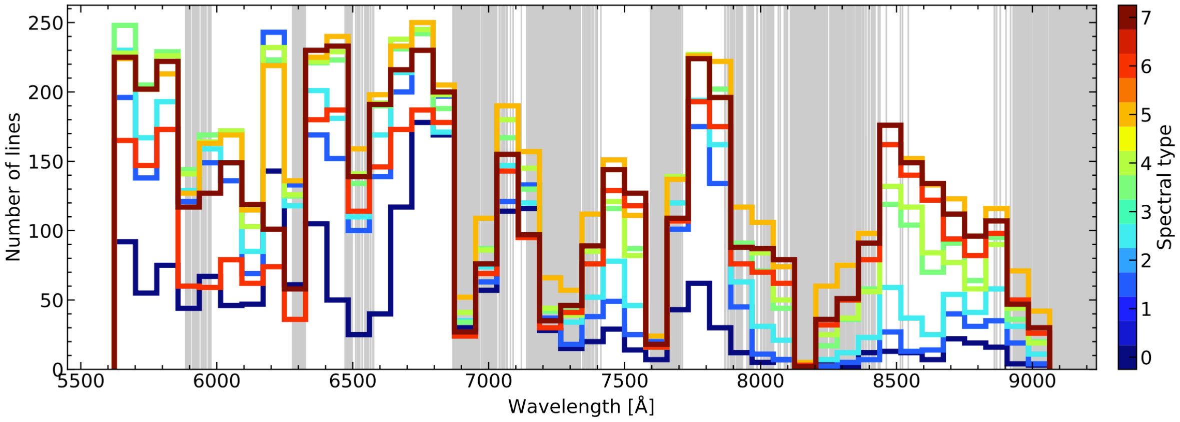


Schöfer et al. 2019

Creating weighted binary masks



Lafarga et al. 2020, number of lines for M dwarf masks of different spectral sub-type



Different stars show different spectral lines

Computing the CCF

$$\mathrm{CCF}(v) = \sum_{l=1}^{m}\sum_{x=1}^{n} w_l \cdot f_x \cdot \Delta_{xl}(v)$$

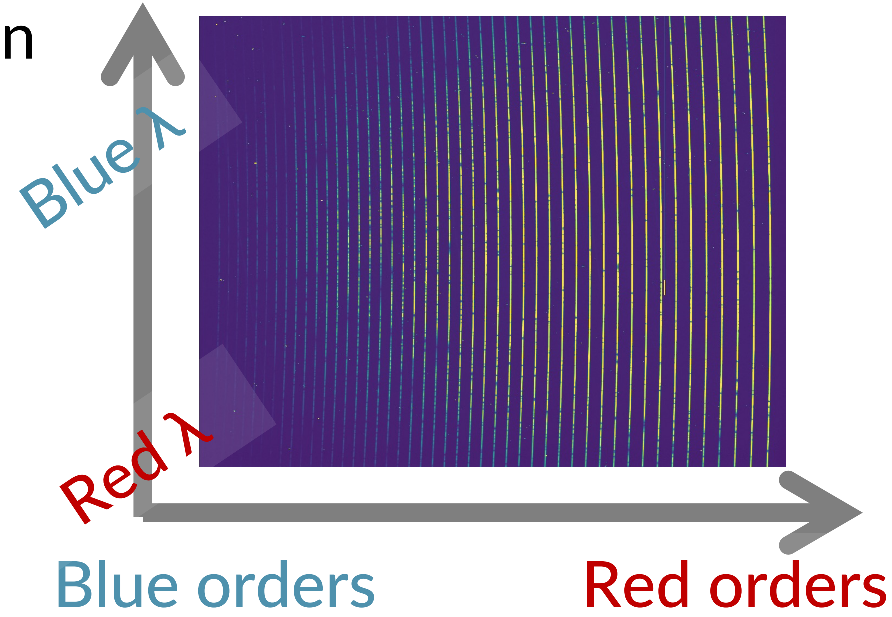
Some things to consider

- Mask line width ~ 1 pixel
- RV step when shifting mask (CCF RV grid) ~ average pixel size in velocity units Δv
  - e.g. HARPS s=3.2 pix/SE, R=115 000, Δv=820 m/s
  - Smaller Δv "counts" the same photons more than once, underestimate RV uncertainties
- CCF computed order-by-order, coadd them to obtain a "final" CCF per observation
- Blue orders tend to have lower S/N
- Order edges tend to have lower S/N
- Some orders are heavily affected by tellurics

$$\Delta v = \frac{c}{R \cdot s}$$

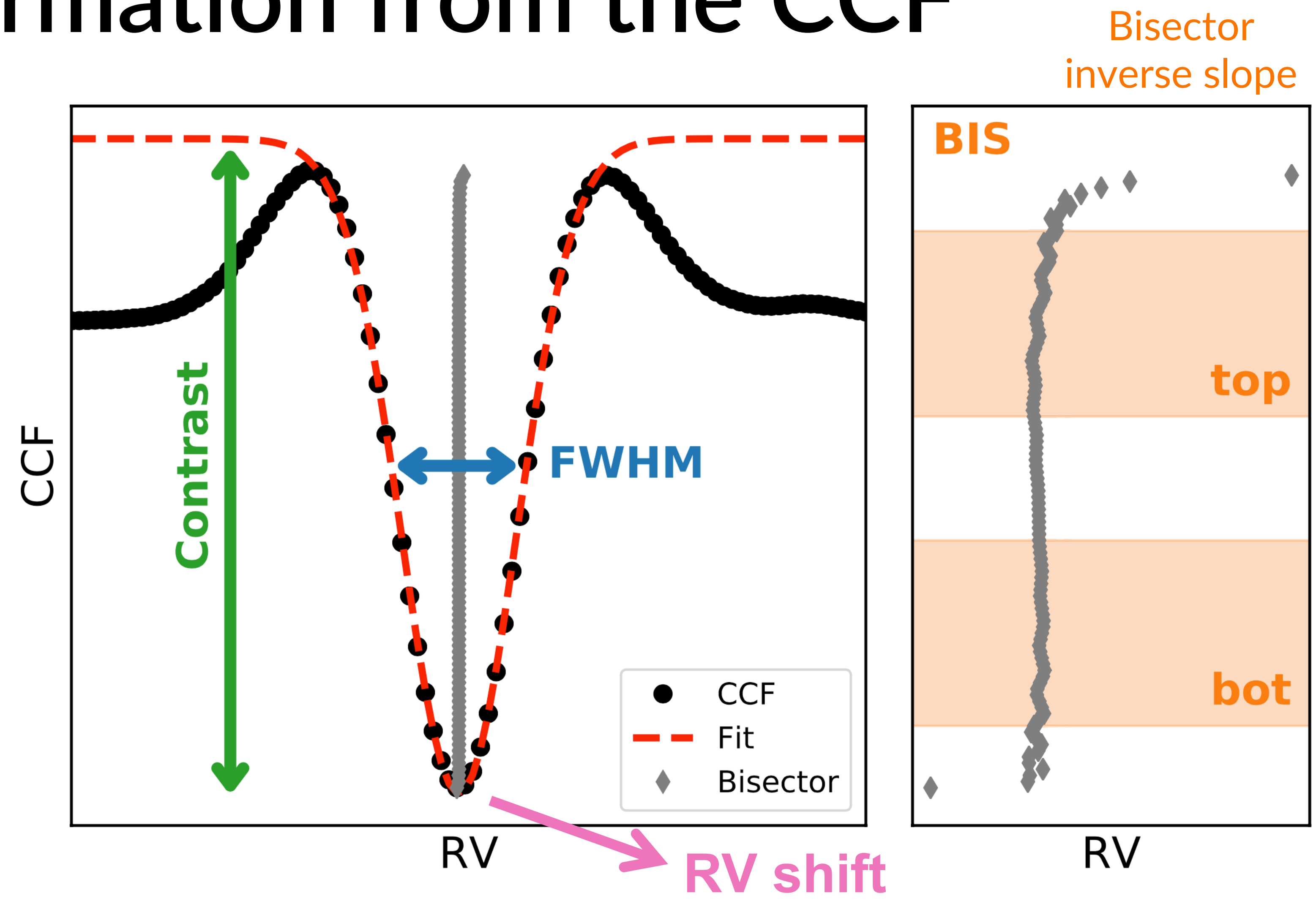
Sampling (pixels / spectral element)

Resolving power



Extracting information from the CCF

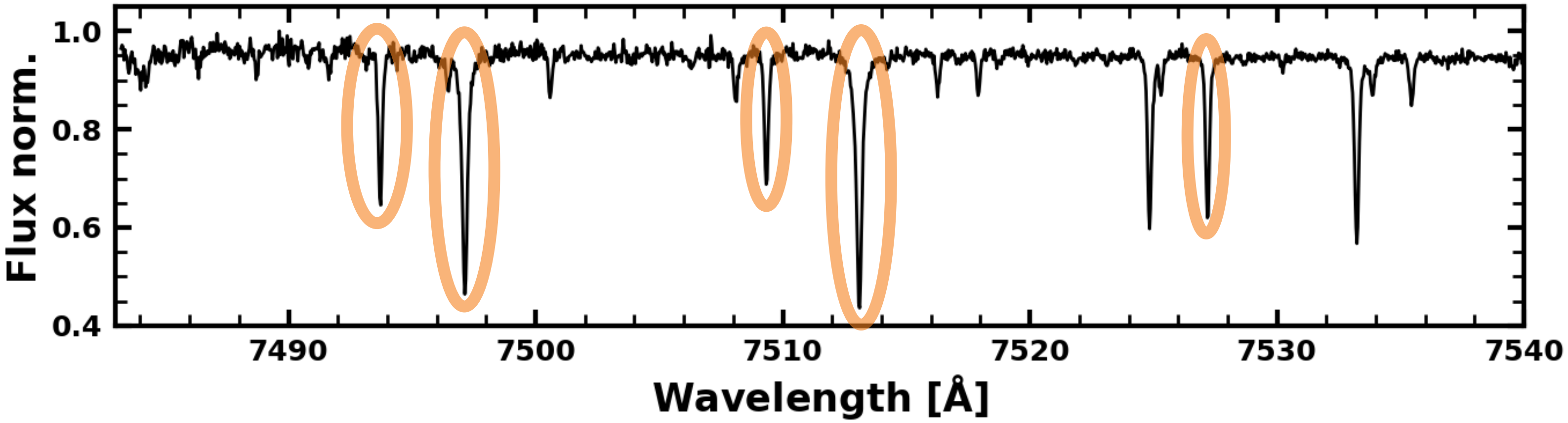
- RV shift = CCF centroid
- Uncertainty from coadded photon noise & CCF slope
- Fit e.g. Gaussian to measure RV and other profile properties
- CCF ~ average spectral line, also contains stellar variability information
- Typical stellar variability proxies: FWHM, contrast, bisector



Queloz et al. 2001, Nardetto et al. 2006, Boisse et al. 2011, Figueira et al. 2013, Lanza et al. 2018, Simola et al. 2019

The CCF method assumes that stellar lines are well isolated and unblended

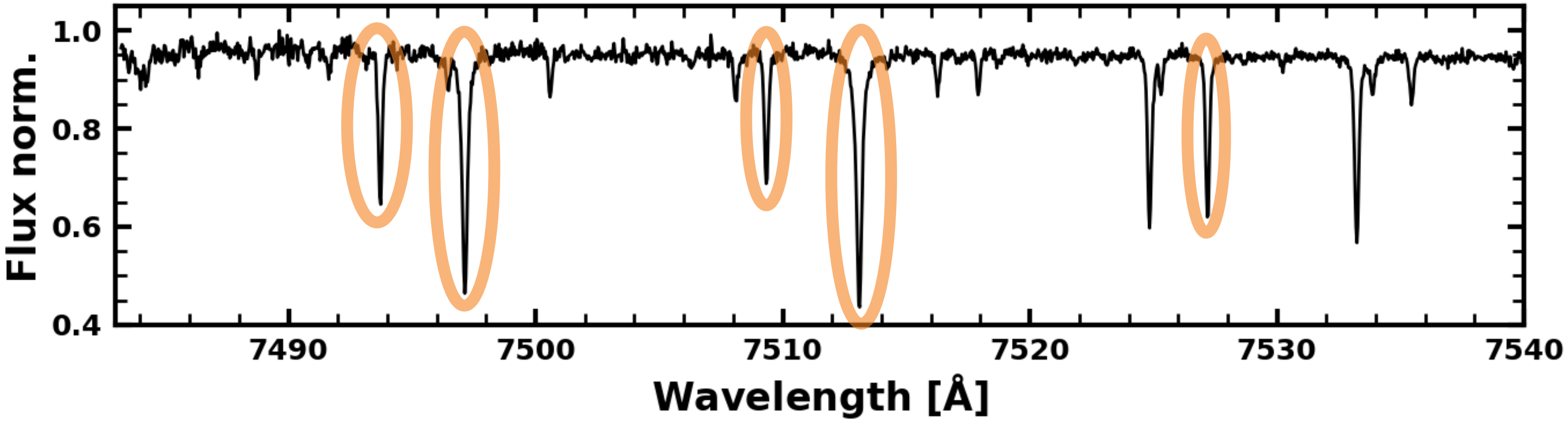
G3
dwarf



CARMENES
observation
S/N~140

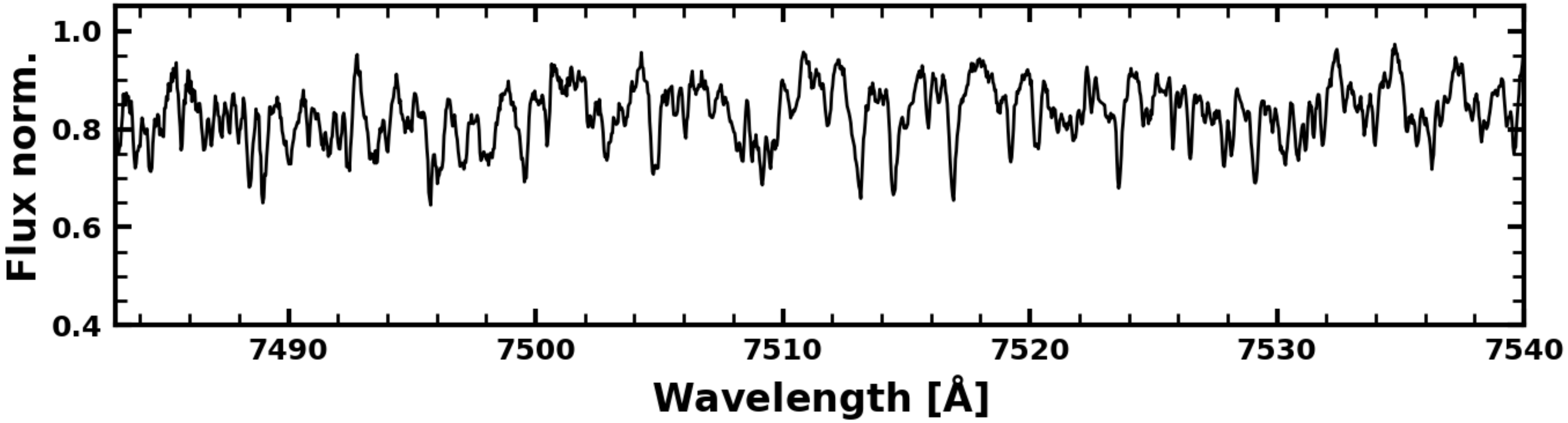
Lines are not as “well-defined” in cooler stars
By selecting good lines we lose a lot of the stellar information

G3
dwarf



CARMENES
observation
S/N~140

M4.5
dwarf

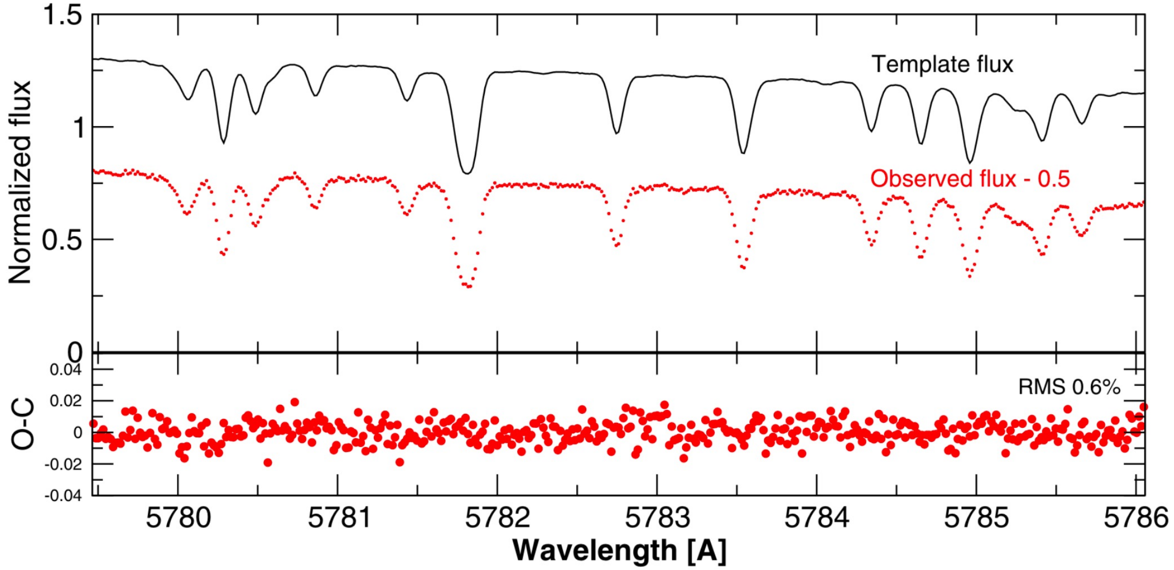


CARMENES
observation
S/N~140

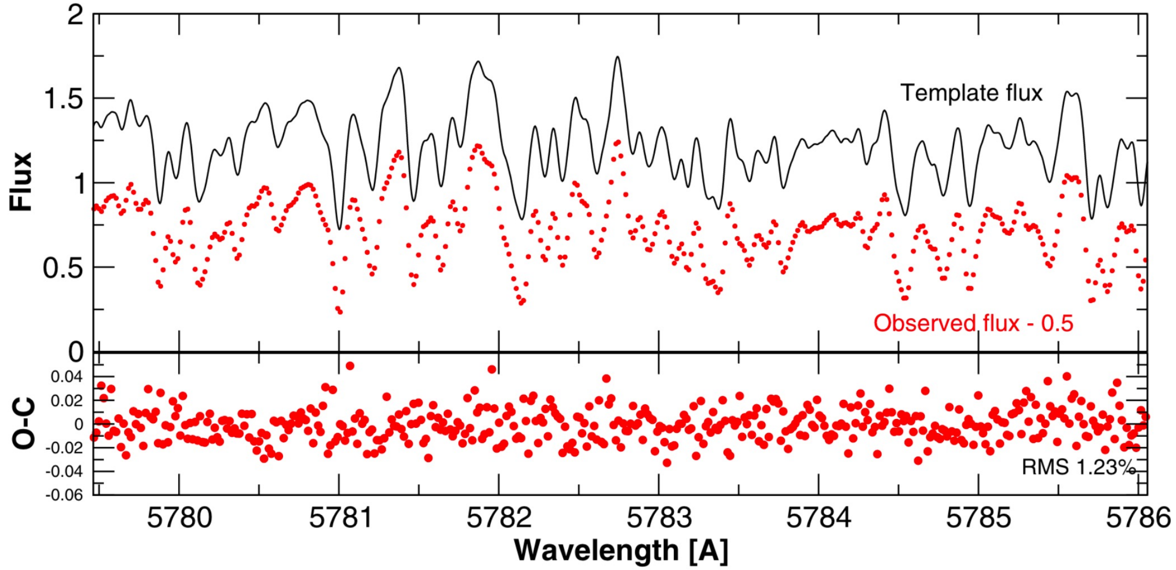
Template matching

Anglada-Escudé & Butler 2012, Astudillo-Defru et al. 2015, Zechmeister et al. 2018

G8.5 V



M4.0 V



- Least-squares matching of the observed spectrum with a high S/N template (minimise the difference between the observation and the template)
- More precise RVs for cool stars than CCF approach

Observed flux

Model flux

$$\chi^2 = \sum_{x=1}^{n} \frac{[f_x - g(\lambda_i)]^2}{{\sigma_i}^2}$$

Flux uncertainty

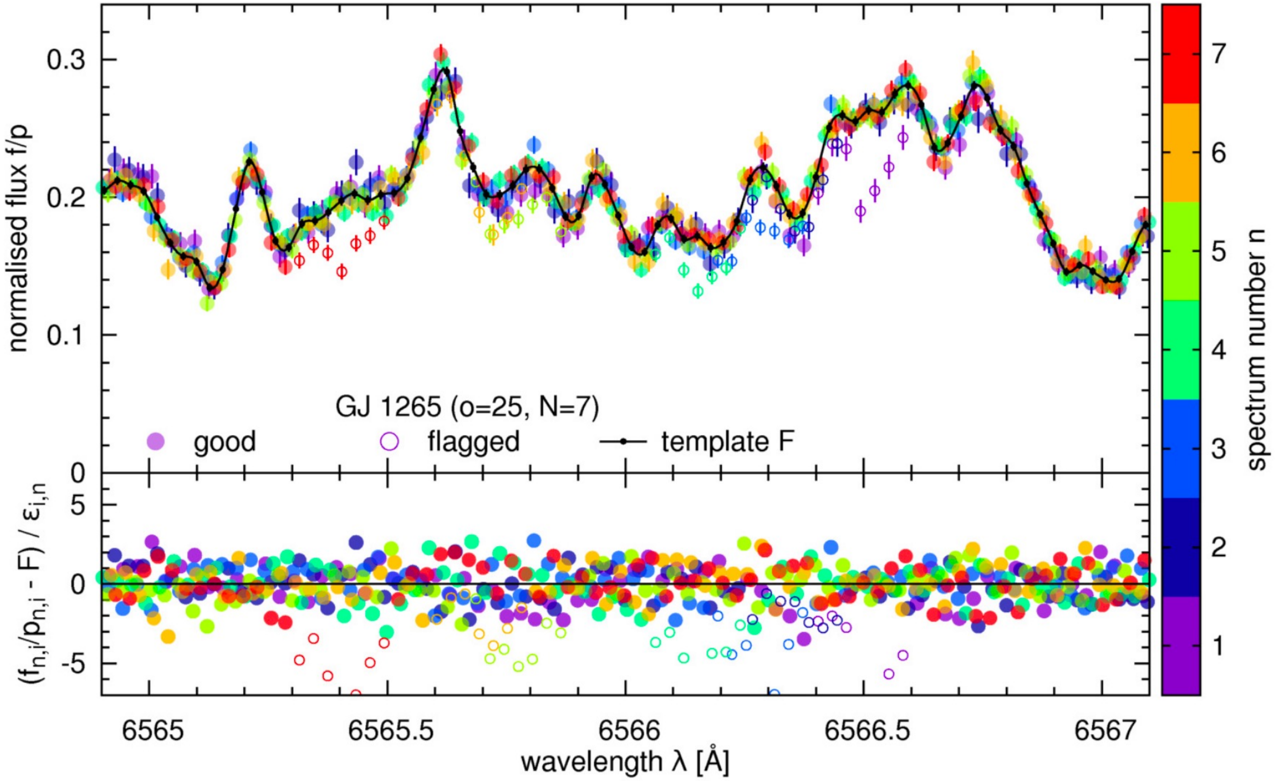
$$g(\lambda_i) = p(\lambda) \cdot \mathrm{G}(\lambda, v)$$

Polynomial to account for flux variations

Spectrum model

Building the template

- Synthetic template with similar properties to observed star (e.g. PHOENIX models, Husser et al. 2013)
- Observation with highest S/N
- Coadded all observations into a high S/N
  - Compute preliminary RVs using observation with highest S/N as template
  - Shift observations by preliminary RVs and coadd them into high S/N template
  - Re-compute RVs with new template
  - Iterative process (usually 1 iteration is sufficient)

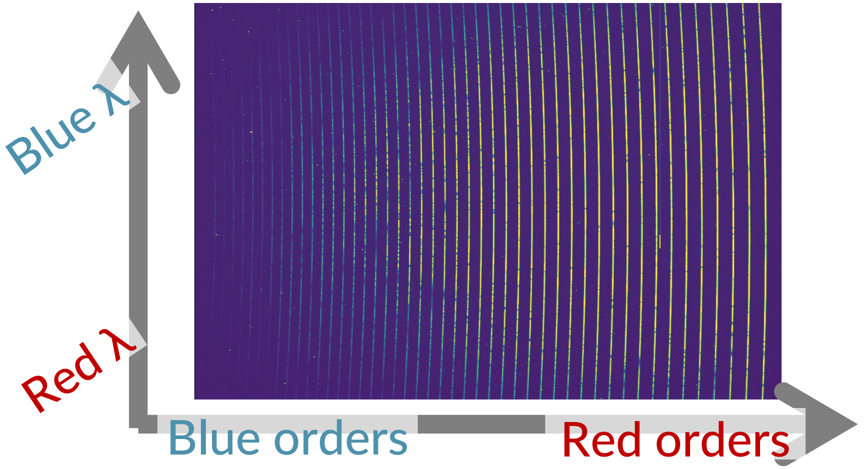


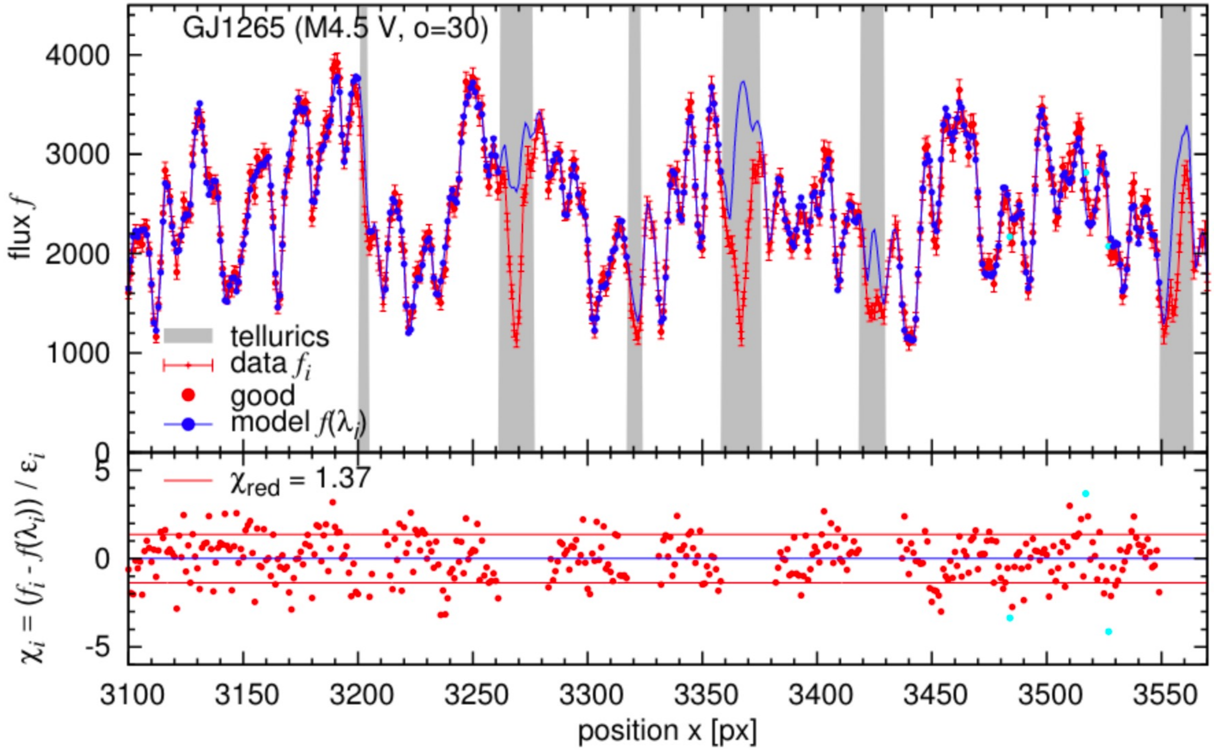
Zechmeister et al. 2018, constructing a template from a B-spline fit to 7 observations

RV computation

Some things to consider

- Minimisation computed order-by-order
  → 1 RV measurement per order
  → Weighted average of the order RVs to obtain "final" RV per observation
- Blue orders tend to have lower S/N
- Order edges tend to have lower S/N
- Some orders are affected by tellurics
  - Discard heavily-affected orders
  - Down-weight/exclude telluric affected pixels





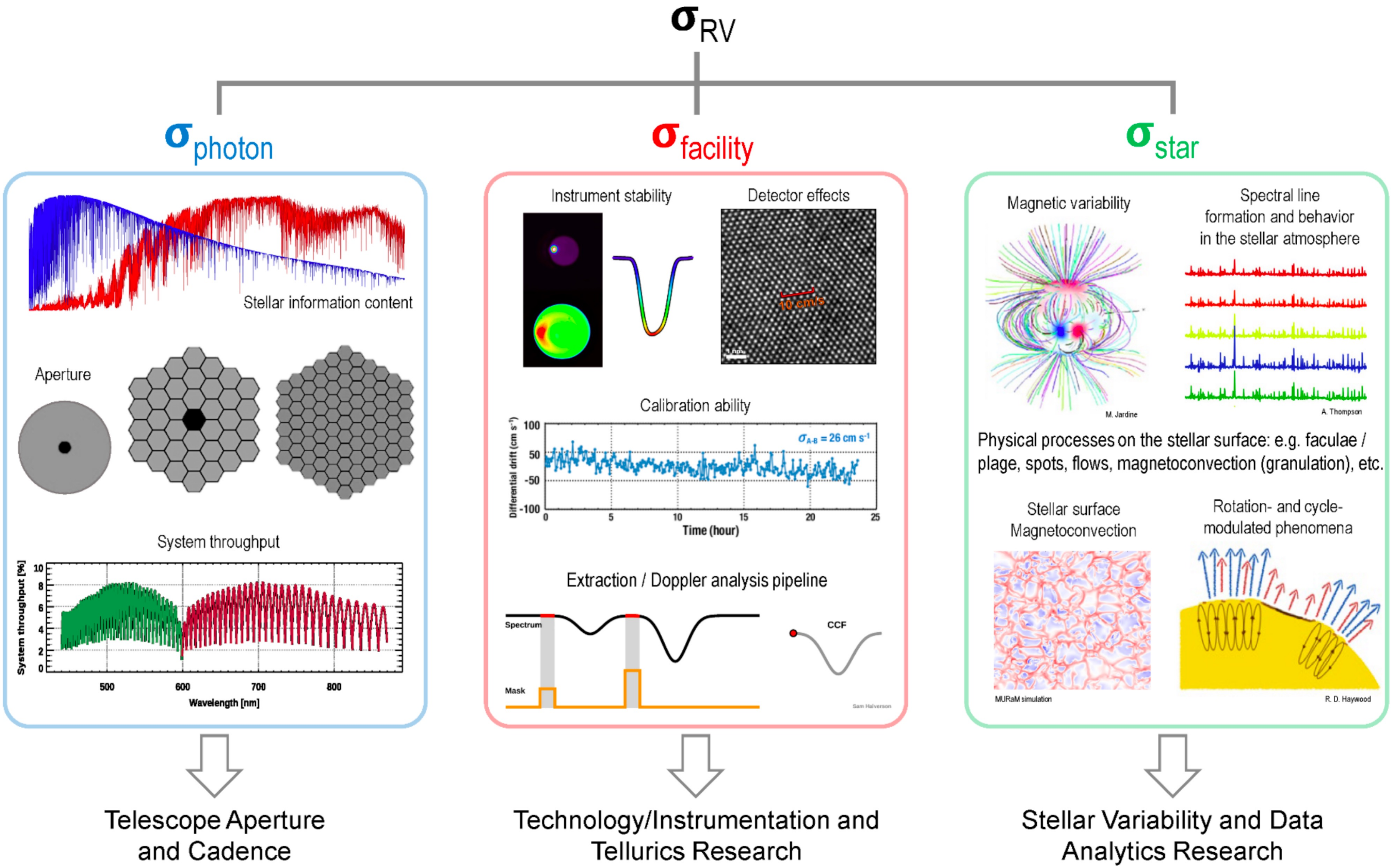
Zechmeister et al. 2018

More ways to measure RVs

Many different approaches that include instrumental, telluric and/or stellar variability effects at the spectral level

- Forward modelling approaches (e.g. Butler et al. 1996, Hirano et al. 2020, Bedell et al. 2019, Gilbertson et al. 2020, Jones et al. 2020)
- Line-by-line approaches (e.g. Dumusque 2018, Cretignier et al. 2020, Artigau et al. 2022, Siegel et al. 2022, Lafarga et al. 2023)
- Least squares deconvolution (e.g. Belloti et al. 2022, Lienhard et al. 2022)
- Fourier domain (e.g. Zhao & Tinney 2020)
- Machine learning (e.g. Czekala et al. 2017, Rajpaul et al. 2020, Colwell et al. 2023)

Almost done!



σ_{RV}

σ_{photon}

Stellar information content

Aperture

System throughput

Telescope Aperture and Cadence

$\sigma_{\mathrm{facility}}$

Instrument stability

Detector effects

Calibration ability

Extraction / Doppler analysis pipeline

Technology/Instrumentation and Tellurics Research

σ_{star}

Magnetic variability

Spectral line formation and behavior in the stellar atmosphere

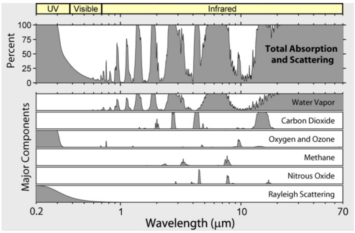
Physical processes on the stellar surface: e.g. faculae / plage, spots, flows, magnetoconvection (granulation), etc.

Stellar surface Magnetoconvection

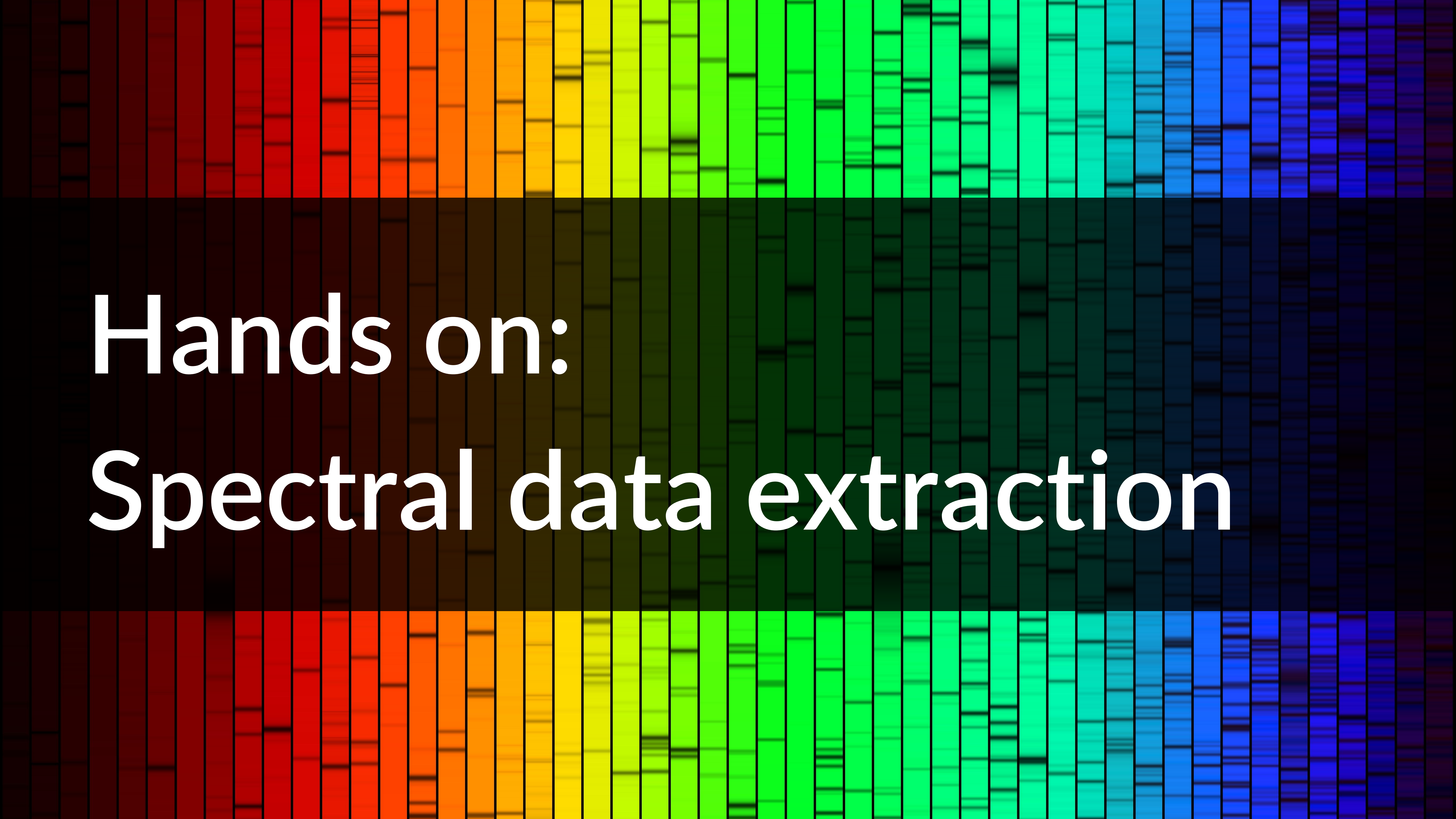
Rotation- and cycle-modulated phenomena

Stellar Variability and Data Analytics Research

& Tellurics



Crass et al. 2021 Extreme Precision Radial Velocity Working Group Final Report, figure by Sam Halverson



Hands on:
Spectral data extraction

Hands-on: Spectral data extraction

Download data and code from:

https://livewarwickac-my.sharepoint.com/:f:/g/personal/u2070295_live_warwick_ac_uk/EgFkwFdfvCpArE6_22uHXl8BNJv66AkrgHBO-Mef_36xqg?e=kKRk92