PHY 517 / AST 443:
Observational Techniques in Astronomy

Lecture 4:
Working with astronomical images
Spectroscopy /
Statistics

So you took all those images, now what?

Images

FITS: Flexible Image Transport System

- open standard for astronomical images
- two parts:
    - image (binary format, integer or float)
    - ASCII *header*

Header Keywords

Mandatory Structure:

Table 5.1: Mandatory keywords for primary header.

1	SIMPLE
2	BITPIX
3	NAXIS
4	NAXISn, n = 1, ..., NAXIS
	(other keywords)
last	END

Conforms to standard:
T(rue) / F(alse)

Bits per pixel:
16: integer
32: float

Number of axes:
2d image -> 2

Image dimensions

End of header

Example

Images from our CCD camera:

```
====> file m13.00000077.FIT (main) <====
SIMPLE  =                    T/CCDSOFT-SOFTWARE BISQUE 3
BITPIX  =                   16
NAXIS   =                    2
NAXIS1  =                 1024
NAXIS2  =                 1024
BSCALE  = +1.00000000000E+000
BZERO   = +3.27680000000E+004
BIAS    =                  100
FOCALLEN= +3.55600000000E+003
OPTAREA = +0.00000000000E+000
```

How big is the image (pixels by pixels)?

Example

Images from our CCD camera:

```
APTAREA = +0.000000000000E+000
APTDIA  = +3.560000000000E+002
TELESCOP= 'Meade LX200             '
OBSERVER= 'T. Cohen, B. Schultz, X. Liu, B. Baserdem'
DATE-OBS= '2016-08-30T03:11:28.477'
TIME-OBS= '03:11:28.477            '
SWCREATE= 'CCDSoft Version 5.00.210'
SET-TEMP= -5.000000000000E+000
COLORCCD=                       0
DISPCOLR=                       1
IMAGETYP= 'Light Frame             '
CCDSEPT =                       1
```

```
OBSTDVER= 'SBFITSEXT Version 1.0'
FILTER  = 'Visual                  '
EXPTIME = +1.000000000000E+001
EXPOSURE= +1.000000000000E+001
ICK     =                     342
TE      =                     447
```

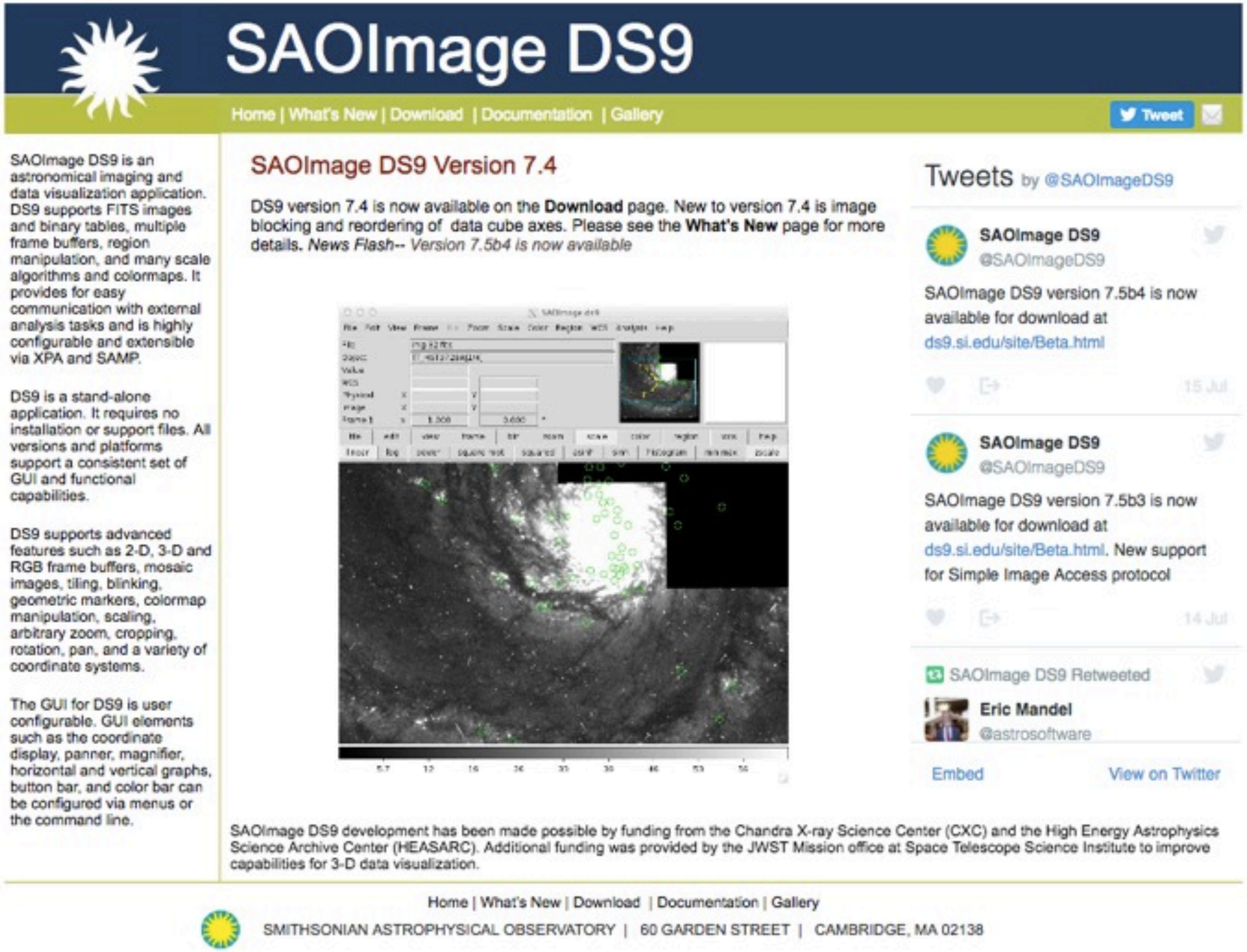
```
EXPSTATE=                     255
CCD-TEMP= -5.232156845990E+000
TEMPERAT= -5.232156845990E+000
INSTRUME= 'SBIG STL-1001 3 CCD Camera'
EGAIN   = +2.060000000000E+000
E-GAIN  = +2.060000000000E+000
```

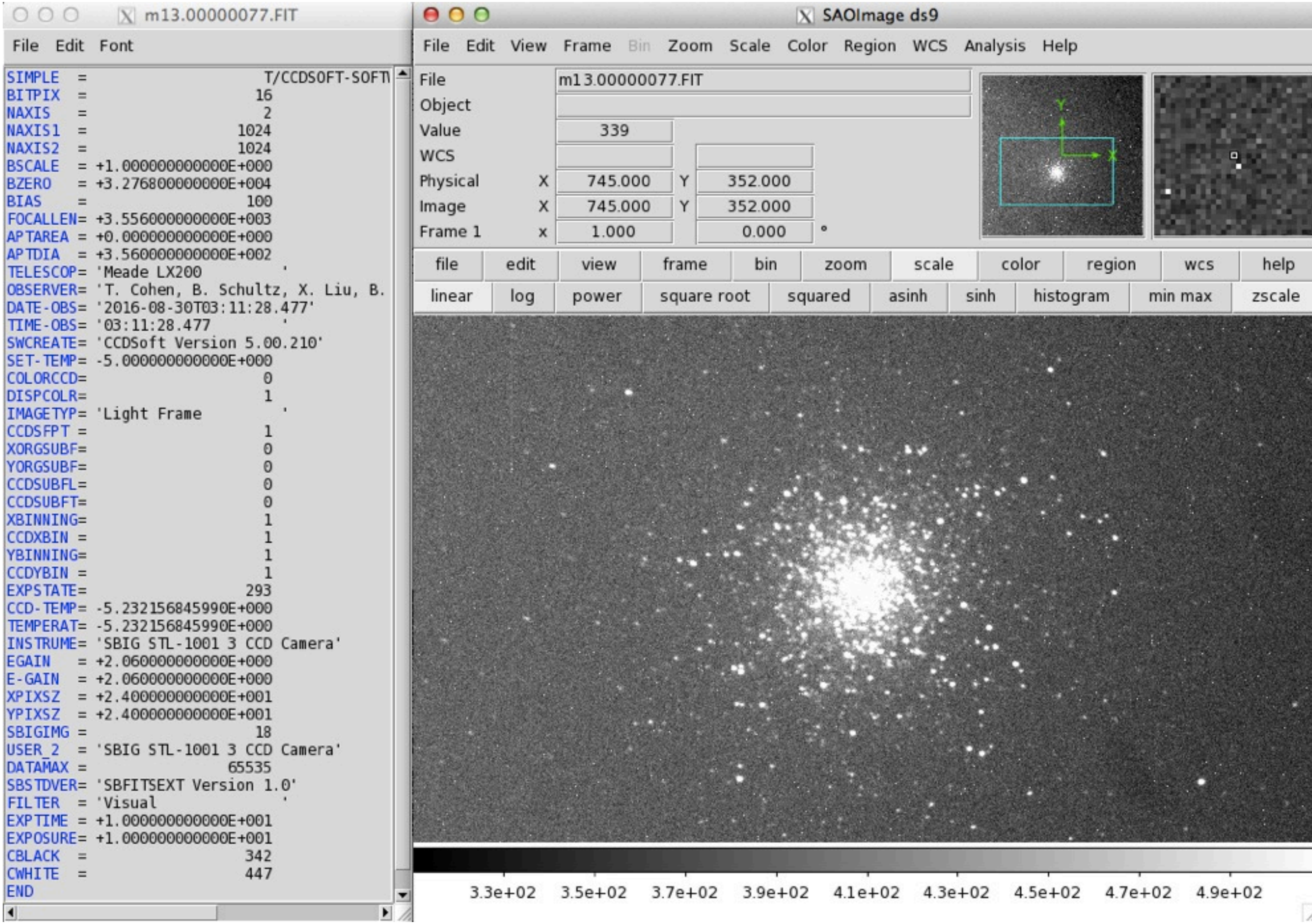
Viewing FITS images

best done with specialized software

e.g. ds9 (by Smithsonian Observatory)

http://ds9.si.edu




m13.00000077.FIT
File Edit Font
SIMPLE = T/CCDSOFT-SOFT
BITPIX = 16
NAXIS = 2
NAXIS1 = 1024
NAXIS2 = 1024
BSCALE = +1.000000000000E+000
BZERO = +3.276800000000E+004
BIAS = 100
FOCALLEN= +3.556000000000E+003
APTAREA = +0.000000000000E+000
APTDIA = +3.560000000000E+002
TELESCOP= 'Meade LX200 '
OBSERVER= 'T. Cohen, B. Schultz, X. Liu, B.
DATE-OBS= '2016-08-30T03:11:28.477'
TIME-OBS= '03:11:28.477 '
SWCREATE= 'CCDSoft Version 5.00.210'
SET-TEMP= -5.000000000000E+000
COLORCCD= 0
DISPCOLR= 1
IMAGETYP= 'Light Frame '
CCDSFPT = 1
XORGSUBF= 0
YORGSUBF= 0
CCDSUBFL= 0
CCDSUBFT= 0
XBINNING= 1
CCDXBIN = 1
YBINNING= 1
CCDYBIN = 1
EXPSTATE= 293
CCD-TEMP= -5.232156845990E+000
TEMPERAT= -5.232156845990E+000
INSTRUME= 'SBIG STL-1001 3 CCD Camera'
EGAIN = +2.060000000000E+000
E-GAIN = +2.060000000000E+000
XPIXSZ = +2.400000000000E+001
YPIXSZ = +2.400000000000E+001
SBIGIMG = 18
USER_2 = 'SBIG STL-1001 3 CCD Camera'
DATAMAX = 65535
SBSTDVER= 'SBFITSEXT Version 1.0'
FILTER = 'Visual '
EXPTIME = +1.000000000000E+001
EXPOSURE= +1.000000000000E+001
CBLACK = 342
CWHITE = 447
END
SAOImage ds9
File Edit View Frame Bin Zoom Scale Color Region WCS Analysis Help
File m13.00000077.FIT
Object
Value 339
WCS
Physical X 745.000 Y 352.000
Image X 745.000 Y 352.000
Frame 1 x 1.000 0.000 °
file edit view frame bin zoom scale color region wcs help
linear log power square root squared asinh sinh histogram min max zscale
3.3e+02 3.5e+02 3.7e+02 3.9e+02 4.1e+02 4.3e+02 4.5e+02 4.7e+02 4.9e+02

Specifying coordinates

The astrometric information in FITS images (also referred to as the WCS) is stored in the header using a standard set of keywords. The reference location is defined by the following keywords:

- CRVAL1: defines the right (α) ascension of the reference pixel
- CRVAL2: defines the declination (δ) of the reference pixel
- CRPIX1: the x location of the reference pixel
- CRPIX2: the y location of the reference pixel

The plate scale and rotation of the image is contained in the CD MATRIX (CD?_? keywords).

- CD1_1 is the partial of first axis coordinate w.r.t. x
- CD1_2 is the partial of first axis coordinate w.r.t. y
- CD2_1 is the partial of second axis coordinate w.r.t. x
- CD2_2 is the partial of second axis coordinate w.r.t. y

$$\begin{pmatrix} CD1\_1 & CD1\_2 \\ CD2\_1 & CD2\_2 \end{pmatrix} = scale * \begin{pmatrix} cos(\theta) & sin(\theta) \\ -sin(\theta) & cos(\theta) \end{pmatrix}$$

Specifying coordinates

```
CTYPE1  = 'RA---TAN-SIP' / TAN (gnomic) projection + SIP distortions
CTYPE2  = 'DEC--TAN-SIP' / TAN (gnomic) projection + SIP distortions
EQUINOX =               2000.0 / Equatorial coordinates definition (yr)
LONPOLE =                180.0 / no comment
LATPOLE =                  0.0 / no comment
CRVAL1  =        250.418630769 / RA  of reference point
CRVAL2  =        36.5118440685 / DEC of reference point
CRPIX1  =        351.470682144 / X reference pixel
CRPIX2  =        386.277894974 / Y reference pixel
CUNIT1  = 'deg     ' / X pixel scale units
CUNIT2  = 'deg     ' / Y pixel scale units
CD1_1   =   -3.33986320359E-05 / Transformation matrix
CD1_2   =    0.000411933007076 / no comment
CD2_1   =   -0.000411849476697 / no comment
CD2_2   =   -3.33825508905E-05 / no comment
```

Image Processing

class github page has starter python code for the image processing (dark and flat field correction)

the code is incomplete, you have to read and modify it!

the code is distributed through the classroom github assignment on exoplanets

get a github account, set up your group repositories

watch some github tutorials

collaborate on the code; check in new versions

your github submissions are part of how we determine that everybody is contributing to the lab work

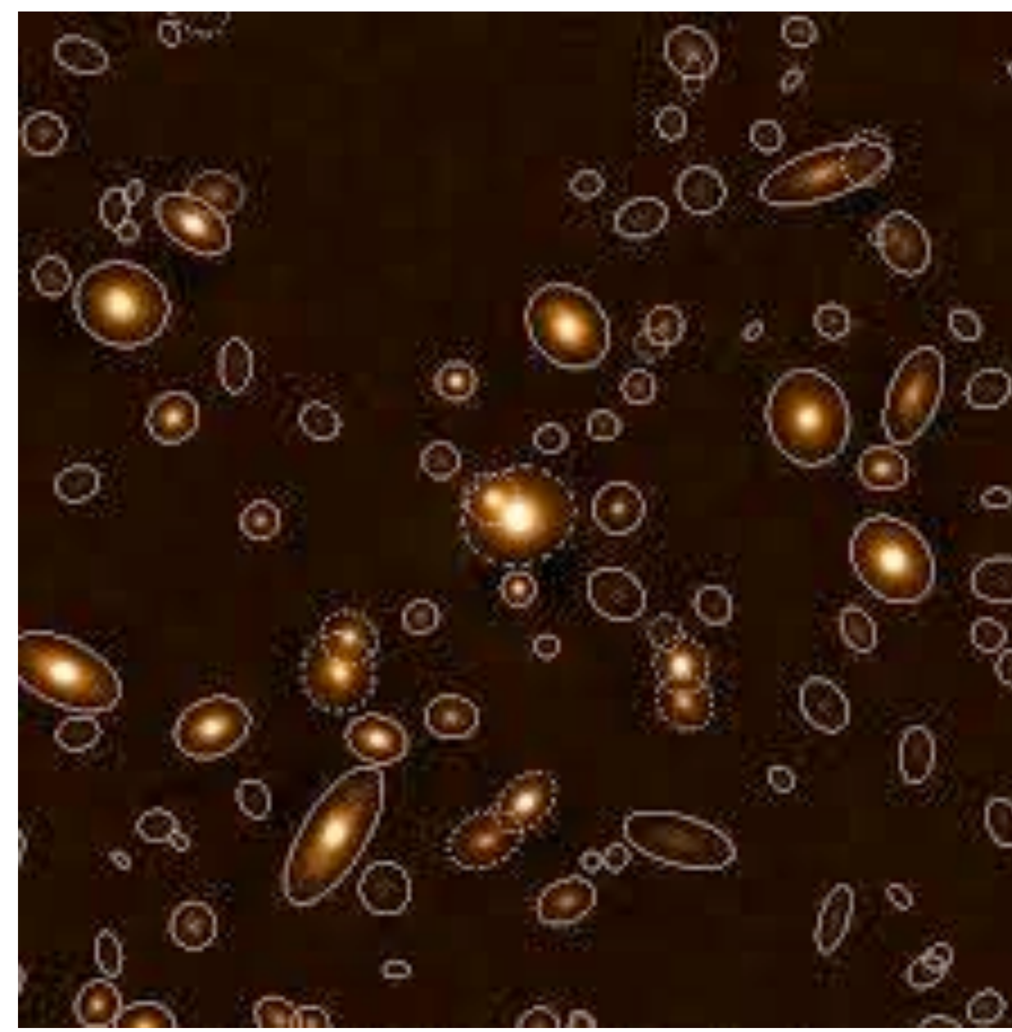
From images to catalogs

the (updated) exoplanet lab description gives guidance on the full analysis steps: from raw images to flux measurements

an important step is identify objects in an image, and measure their properties

standard tool: Source Extractor

check guide on class github page



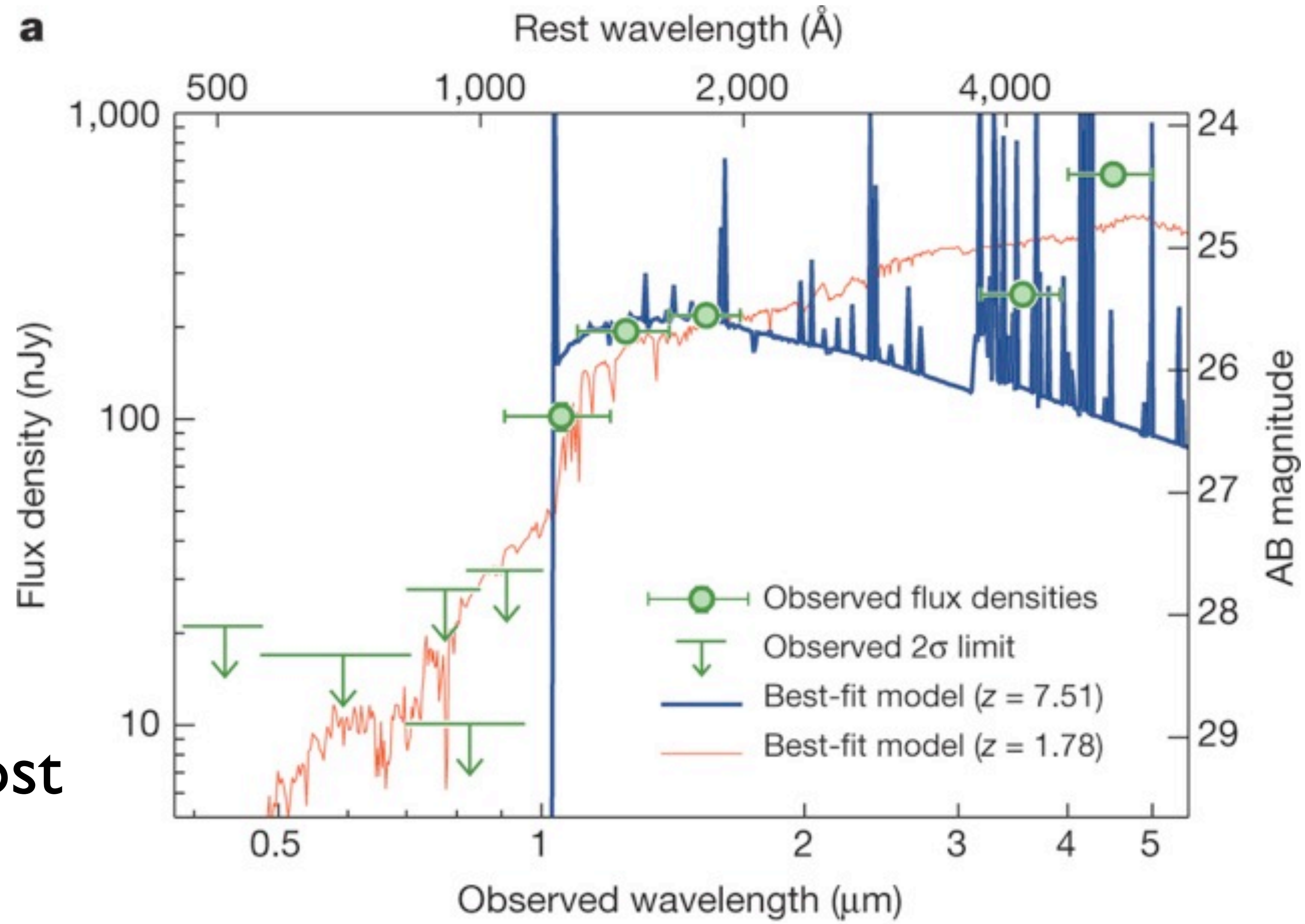
Spectroscopy

Motivation

photometry (measuring flux from images) only measures integrated flux

gives some information about the object properties, but (usually) not enough

e.g.: finding the most distant galaxies

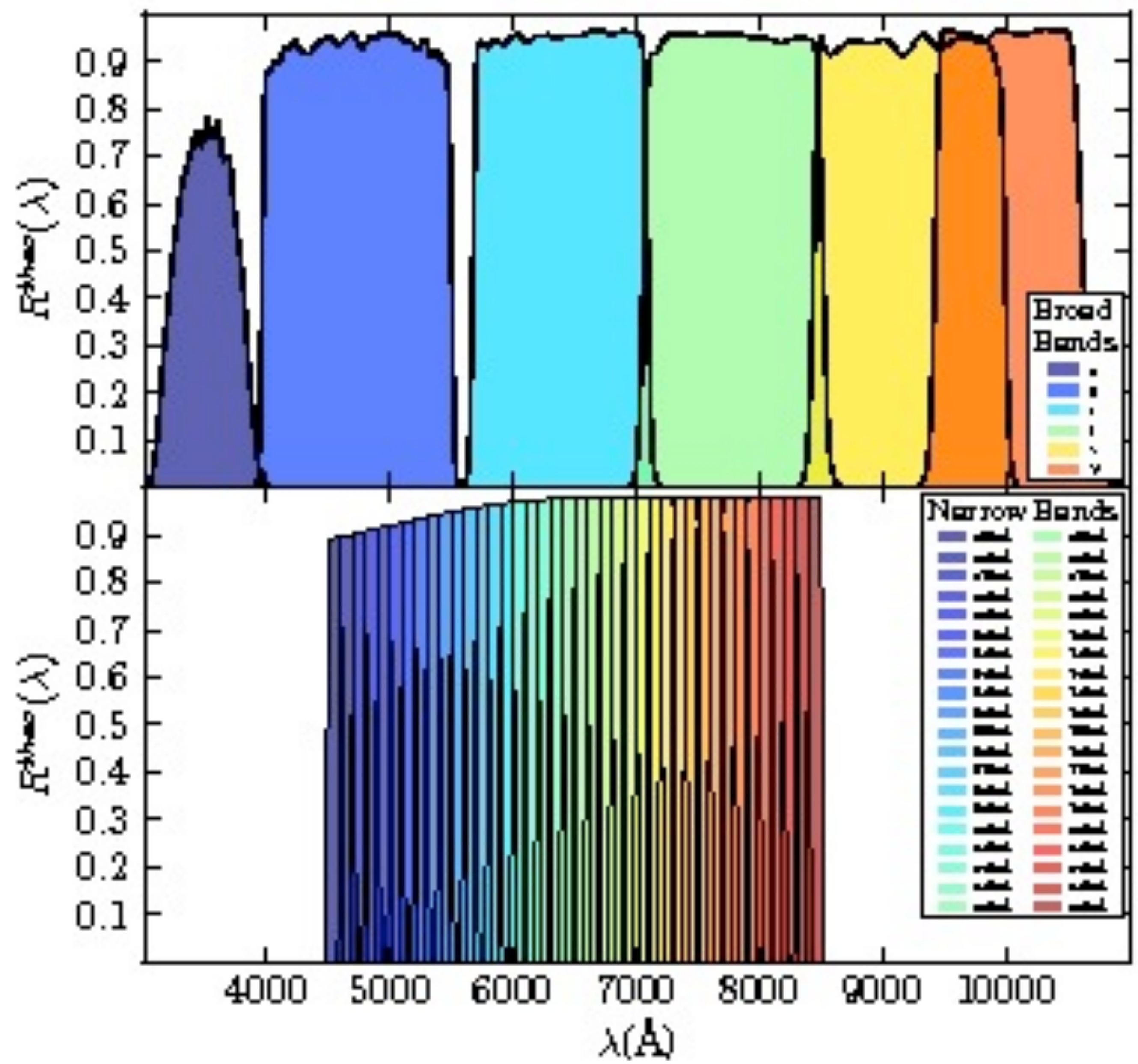


Narrow-band imaging

can determine spectrum of object with images in many narrow-band filters

advantage: can determine spectra of all objects in the same FOV

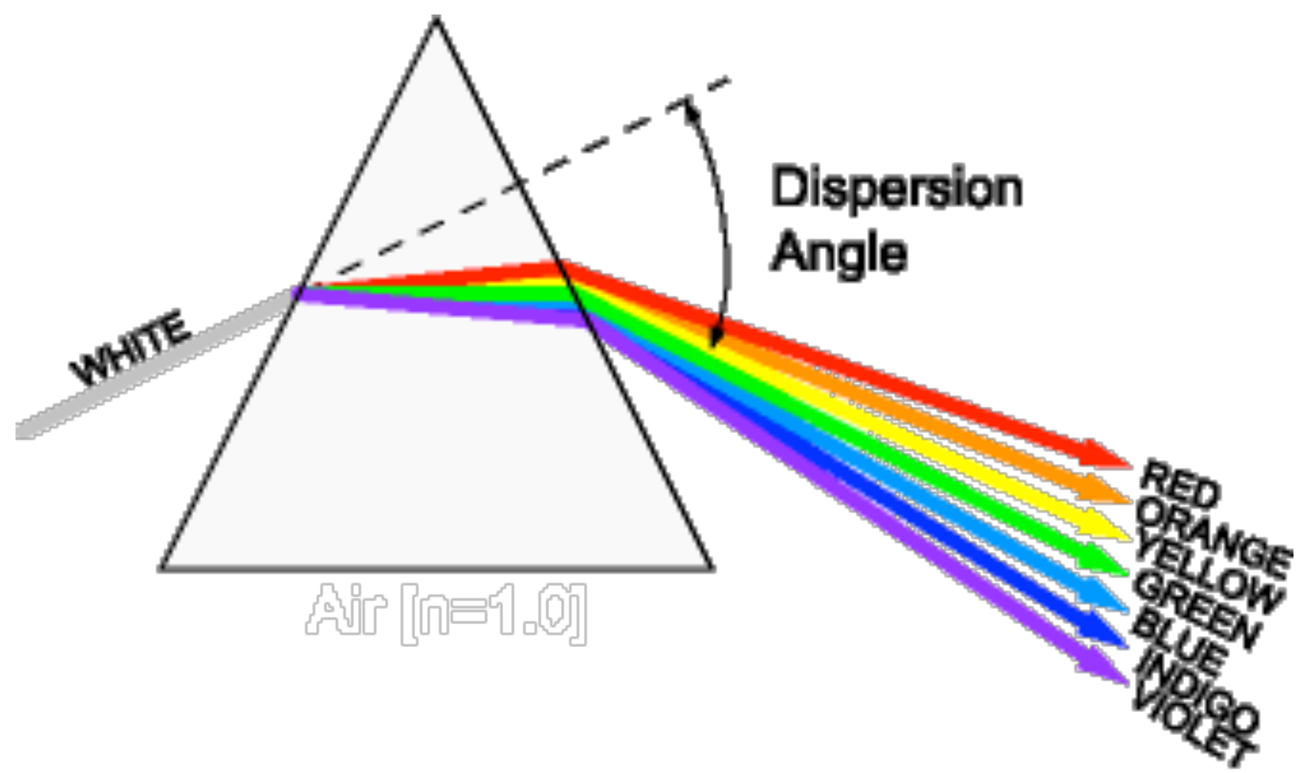
disadvantage: have to take a lot of images!



Spectroscopy

add a dispersing element to split up the light from an object: measure the spectrum directly

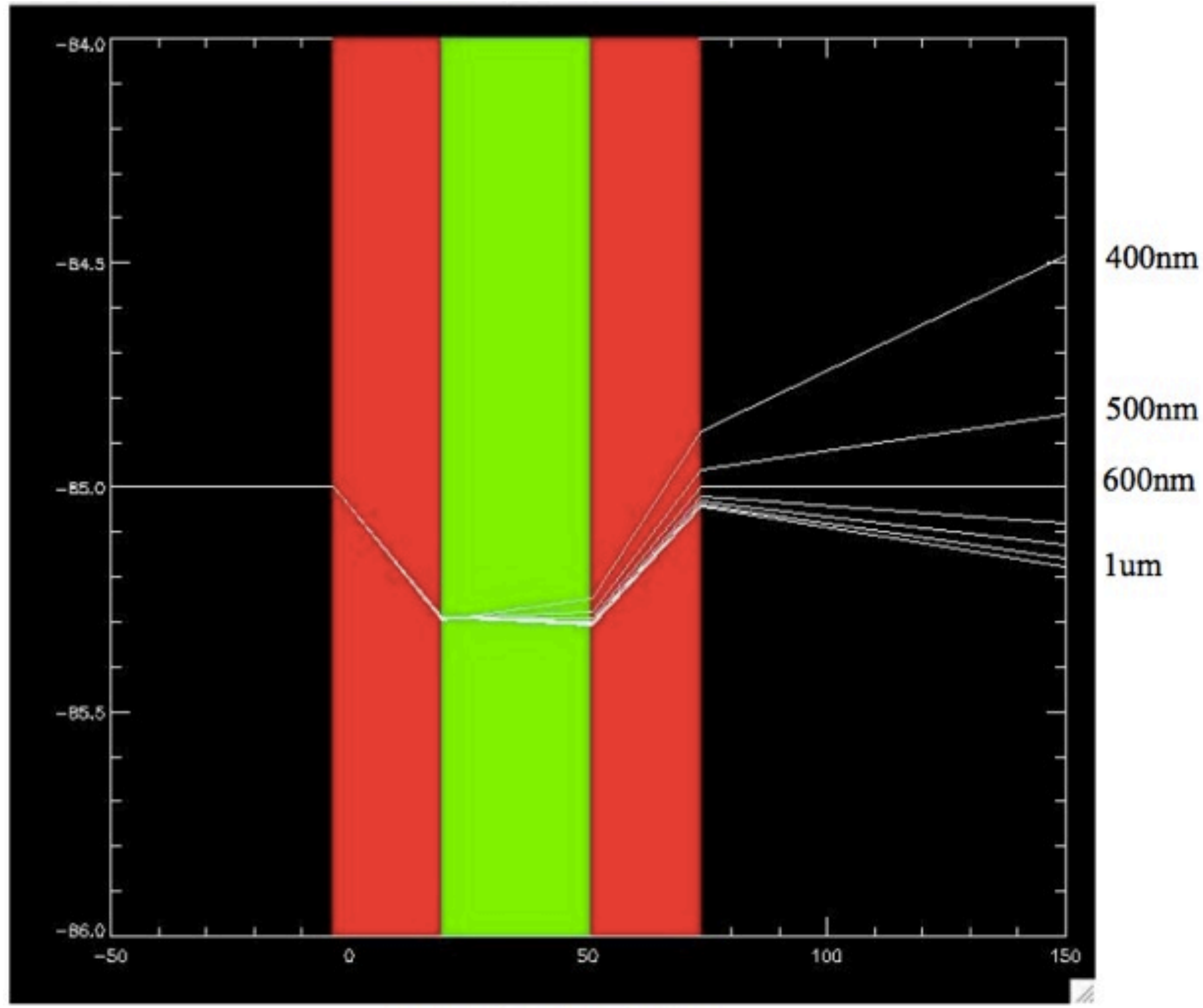
e.g. a prism:



Prism Spectroscopy

only few astronomical spectrographs use prisms

- low dispersion (resolution)
- dispersion varies with wavelength



LDP Prism description

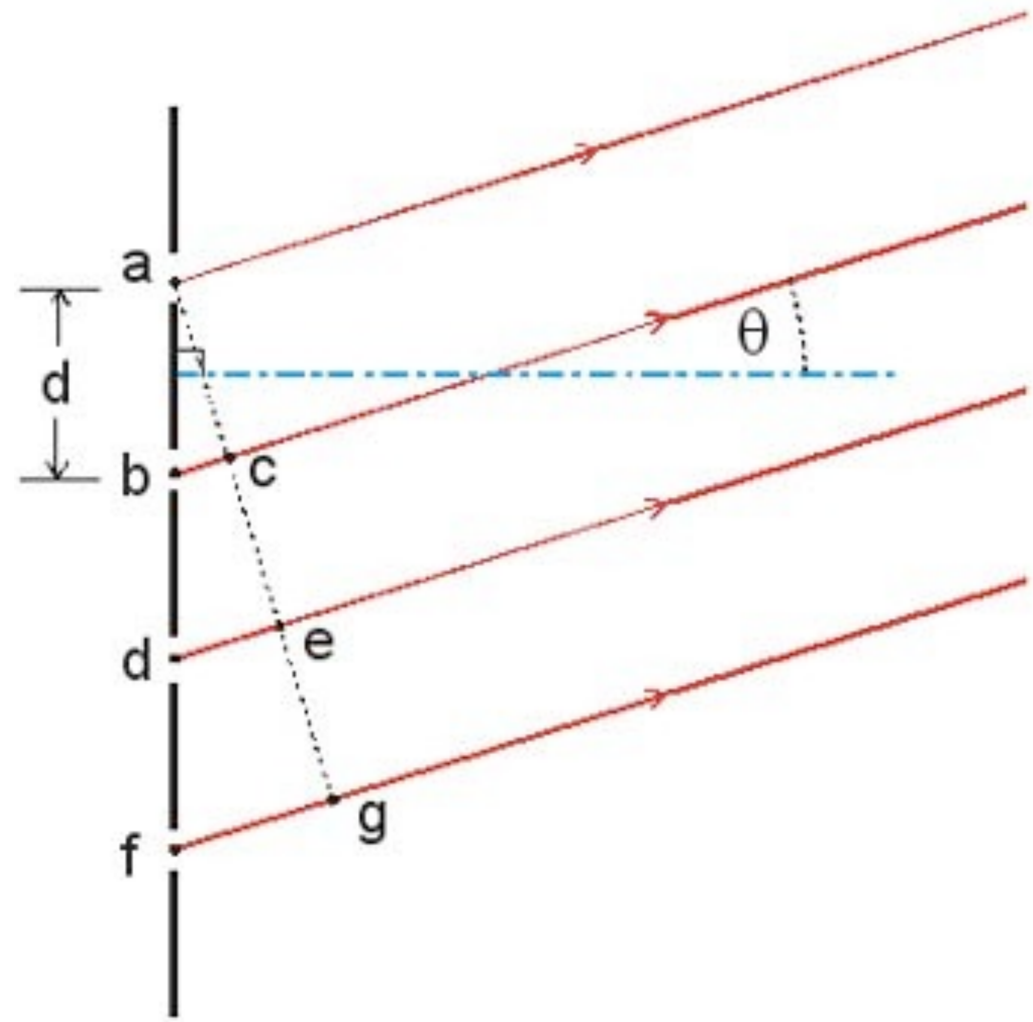
"low dispersion prism" for IMACS spectrograph on Magellan 6-m telescope; uses 3 prisms

Diffraction gratings

make use of wave properties of light: interference

grating: many parallel lines (~500/mm)

similar to single-slit and double-slit experiments



if b-c = λ: maximum at θ and d-e = 2λ, etc.

position of *n*th order:

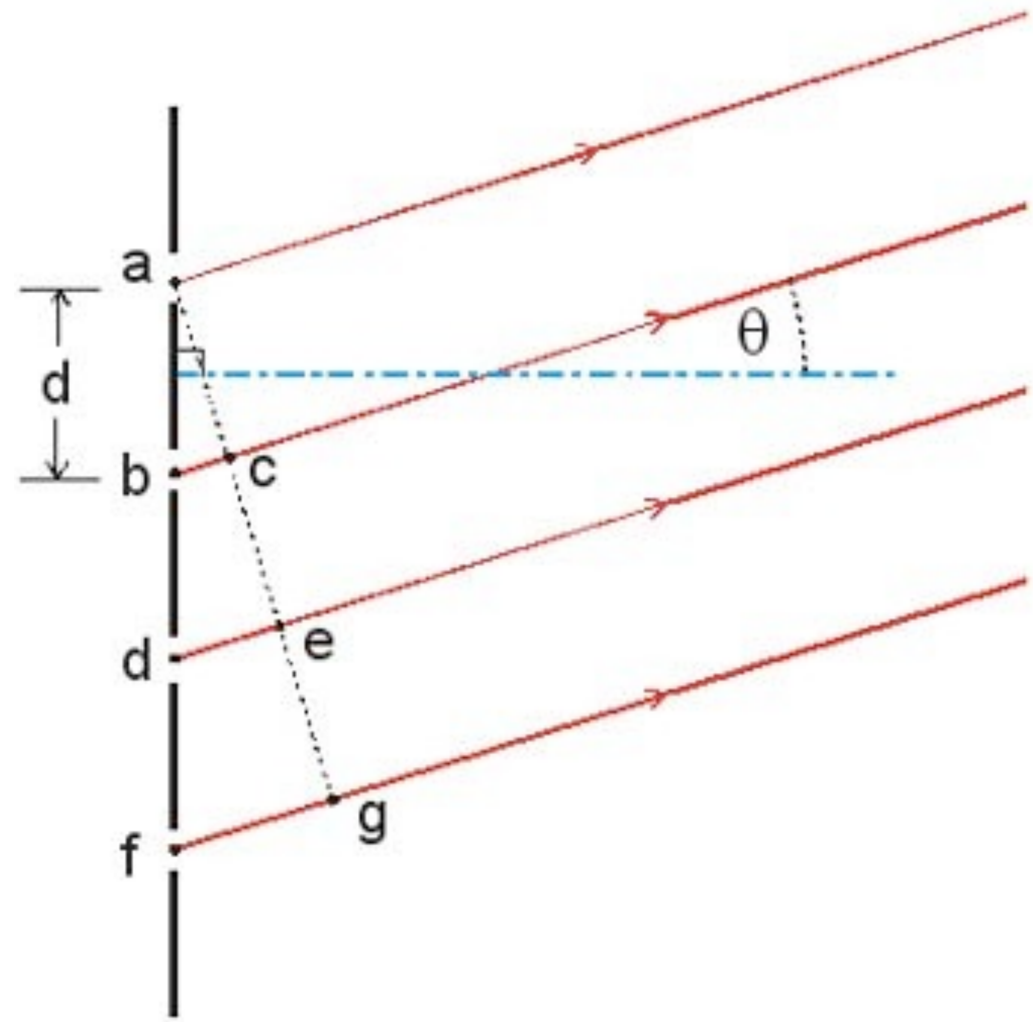
$$n\lambda = d \sin\theta$$

Diffraction gratings

make use of wave properties of light: interference

grating: many parallel lines (~500/mm)

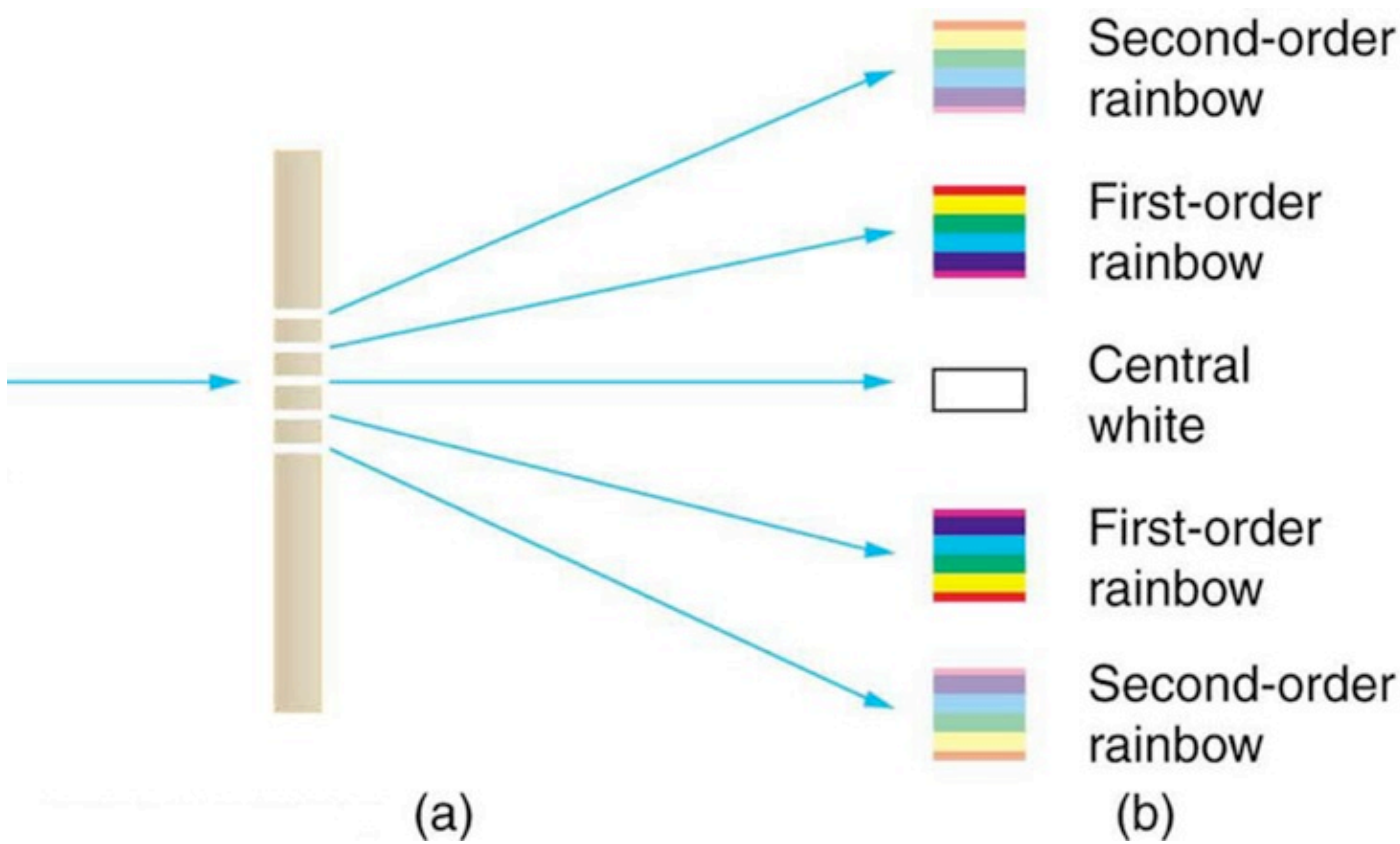
similar to single-slit and double-slit experiments



if b-c = λ: maximum at θ and d-e = 2λ, etc.

position of *n*th order: $$n\lambda = d \sin\theta$$

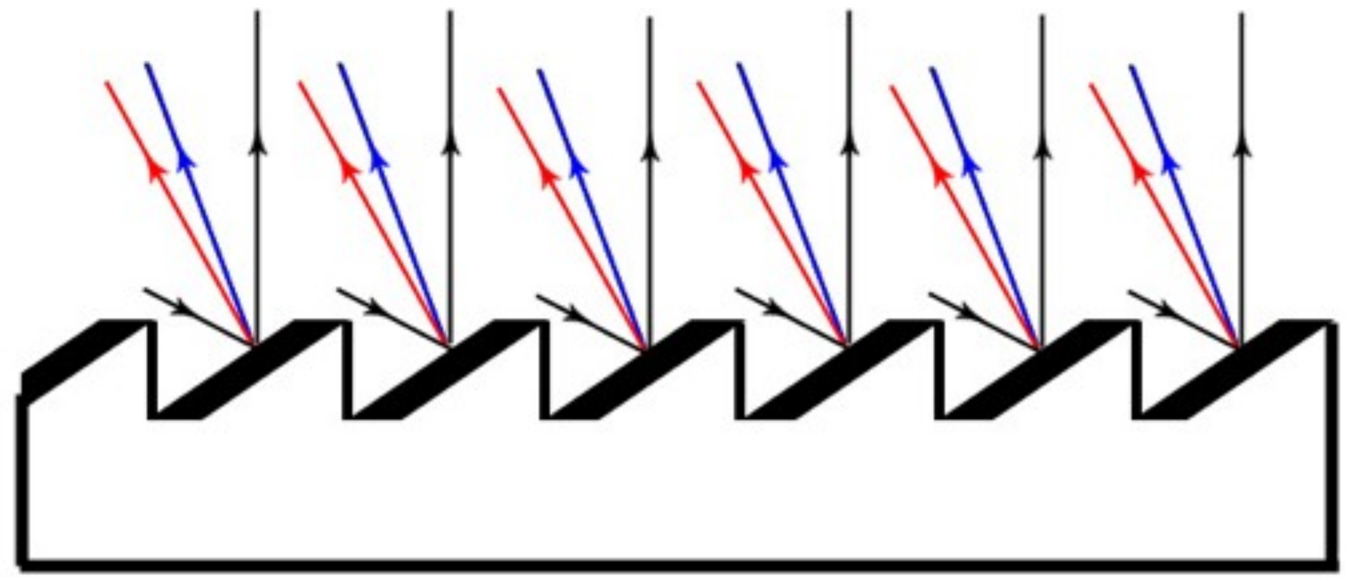
Diffraction gratings



Diffraction gratings

can be transmission gratings
or reflection gratings

most astronomical
spectrographs use reflection
gratings

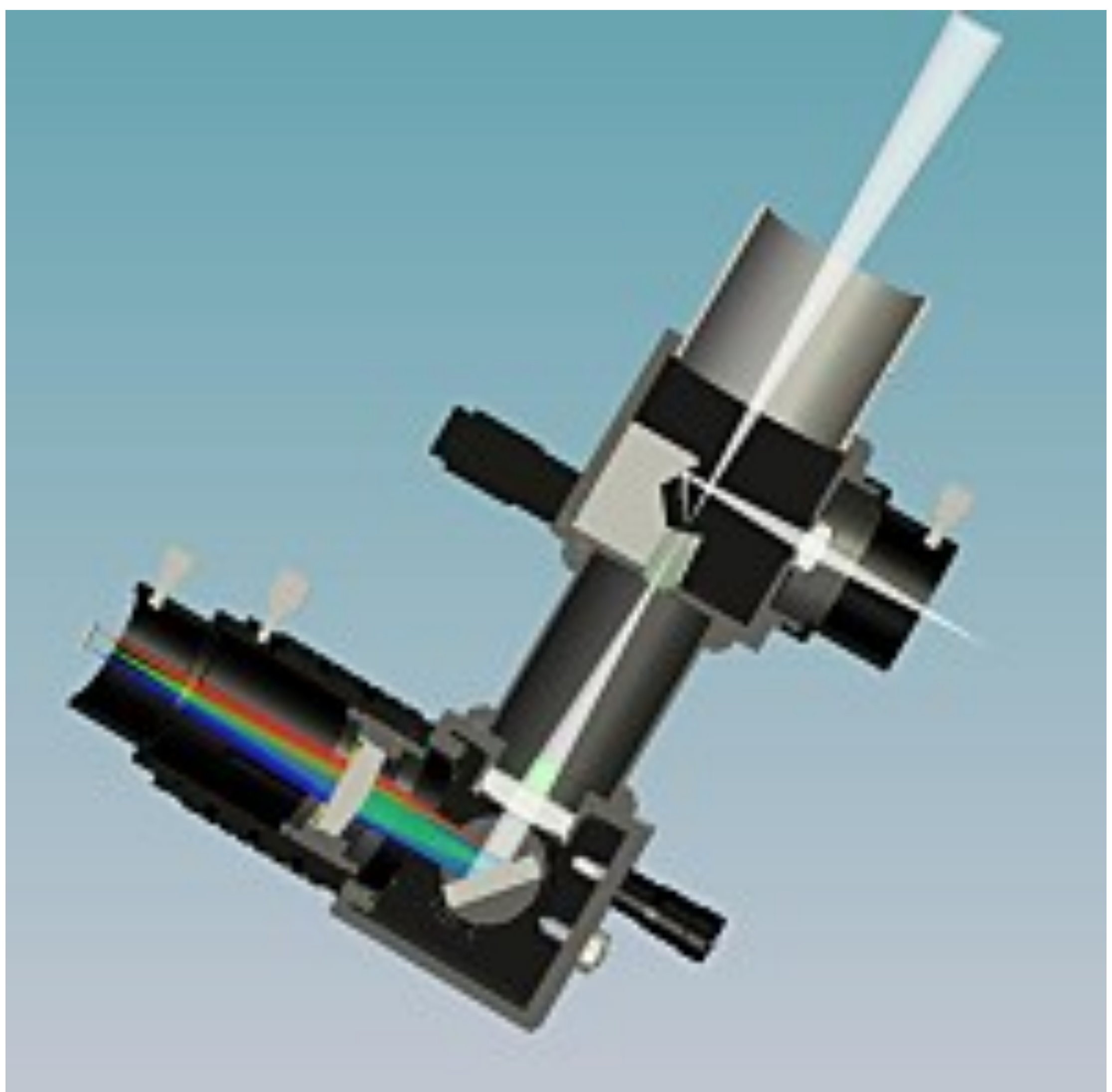


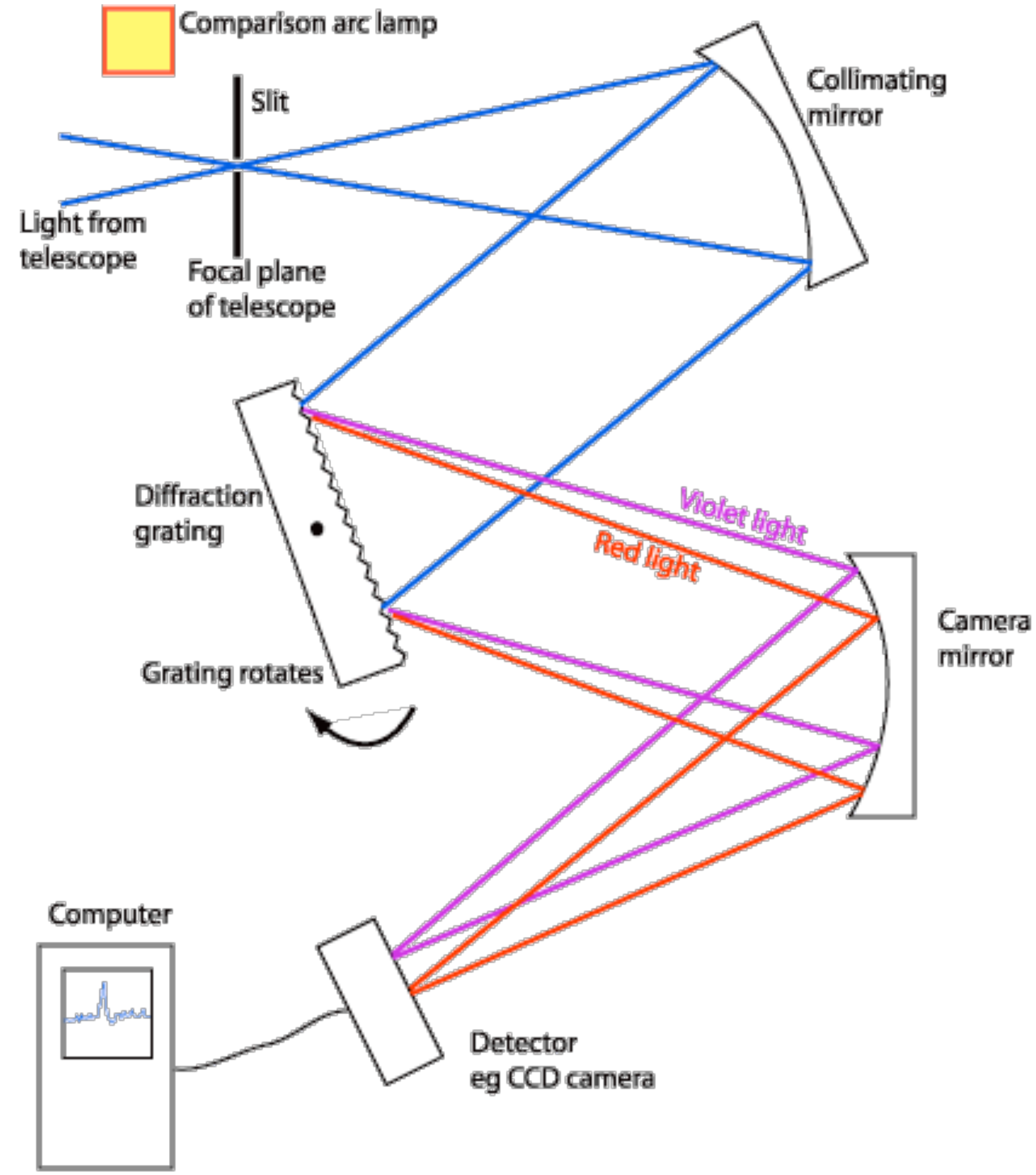
blaze wavelength: wavelength for direction of reflection coincides with desired spectral order
→ maximal efficiency

Typical spectrograph

entrance: usually a slit, similar to seeing size

collimator: converts a diverging beam to a parallel beam





A Schematic Diagram of a Slit Spectrograph

Spectral Resolution

defined by smallest wavelength difference $\Delta\lambda$ that can be distinguished at wavelength λ

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

determined by:

- grating (line density)
- width of entrance slit
- size of CCD pixels

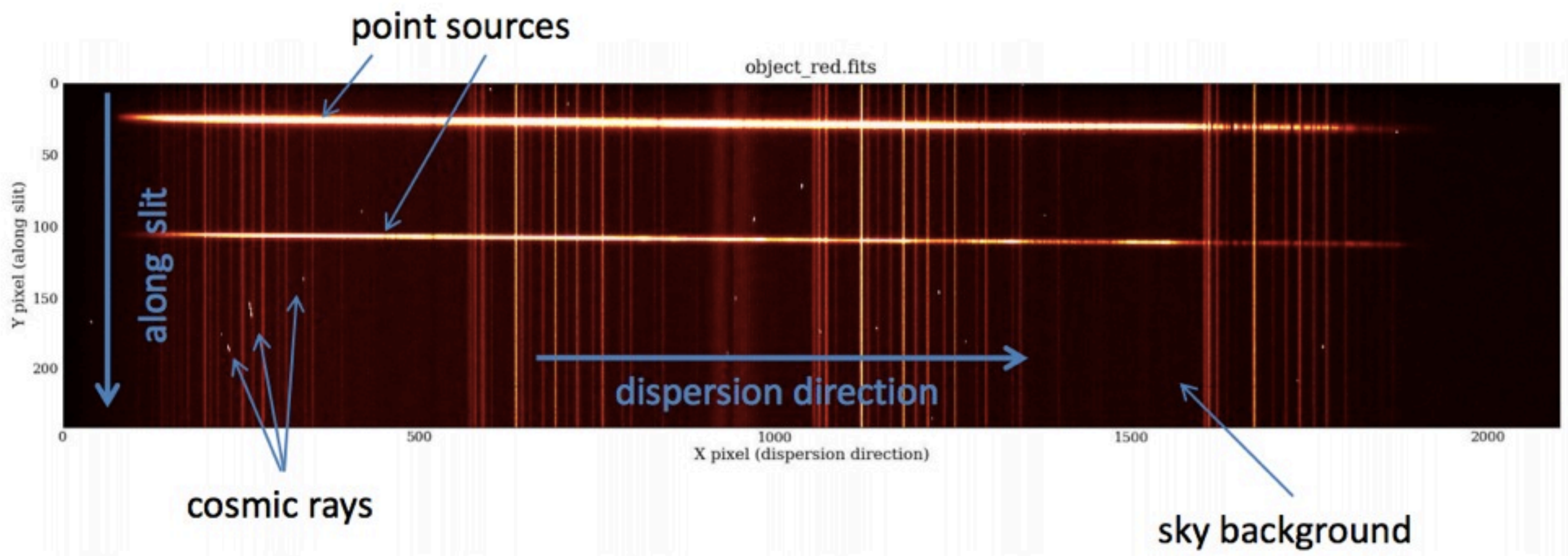
dispersion: length $\Delta\lambda'$ of spectrum over single pixel, [Å/px]

resolution: R or $\Delta\lambda$

to properly sample the spectrum:

$$\Delta\lambda \sim 2 - 3\,\Delta\lambda'$$

A long-slit observation



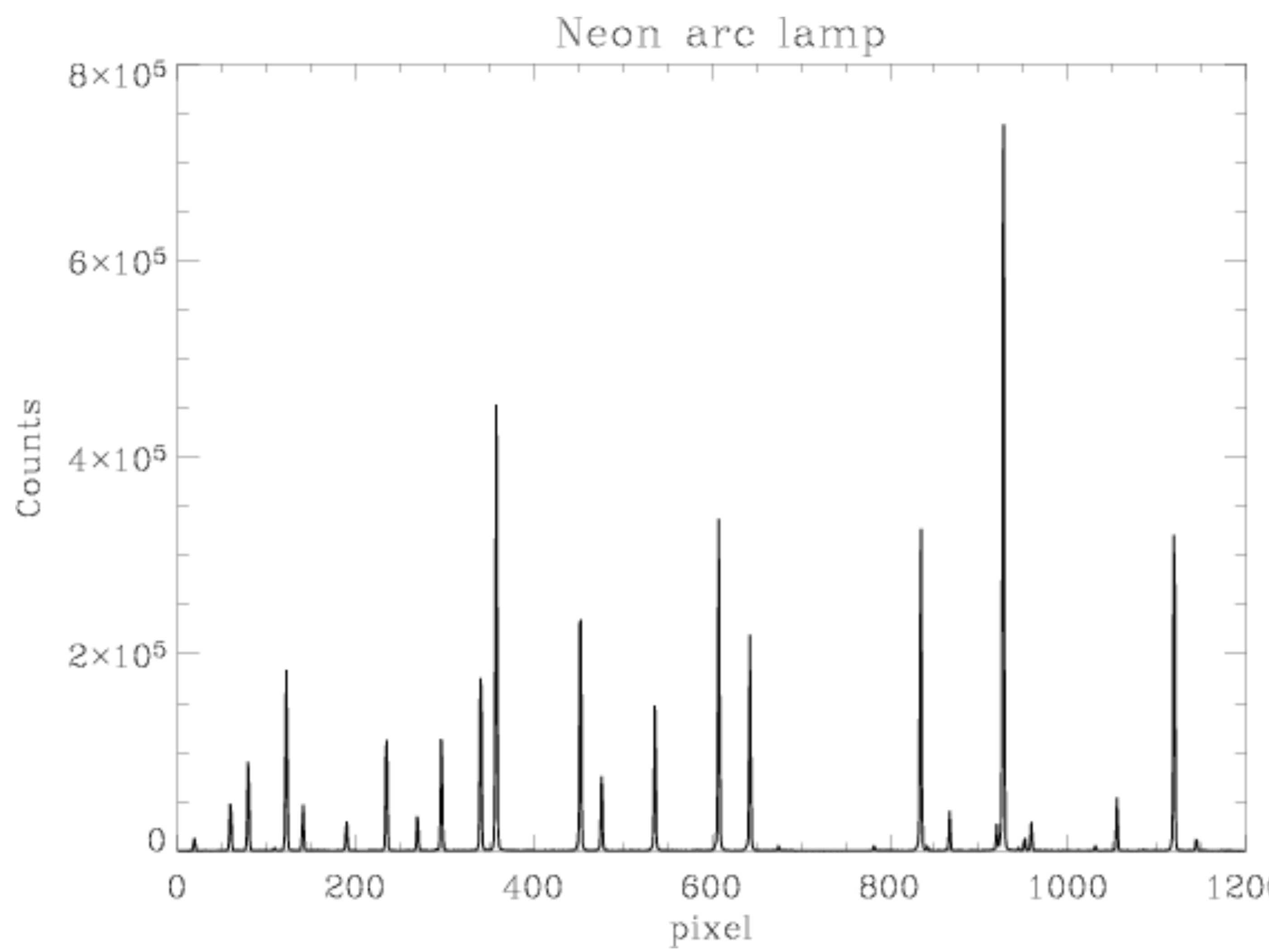
- long axis of CCD used to sample spectrum
- spatial information along slit still available: two objects, lots of sky
- sky background has a lot of emission lines!

Spectroscopic Calibration

- dark frames!
- flat field: use bright continuum source
  - small-scale pixel sensitivity variation
  - variations in slit width
- wavelength calibration: which position on the CCD corresponds to which wavelength?
  - use "arc" lamps with discrete emission lines
  - can also use sky emission lines
- flux calibration:
  - "spectrophotometric" standard stars: stars with known spectral shapes, smooth continua

Spectroscopic Calibration

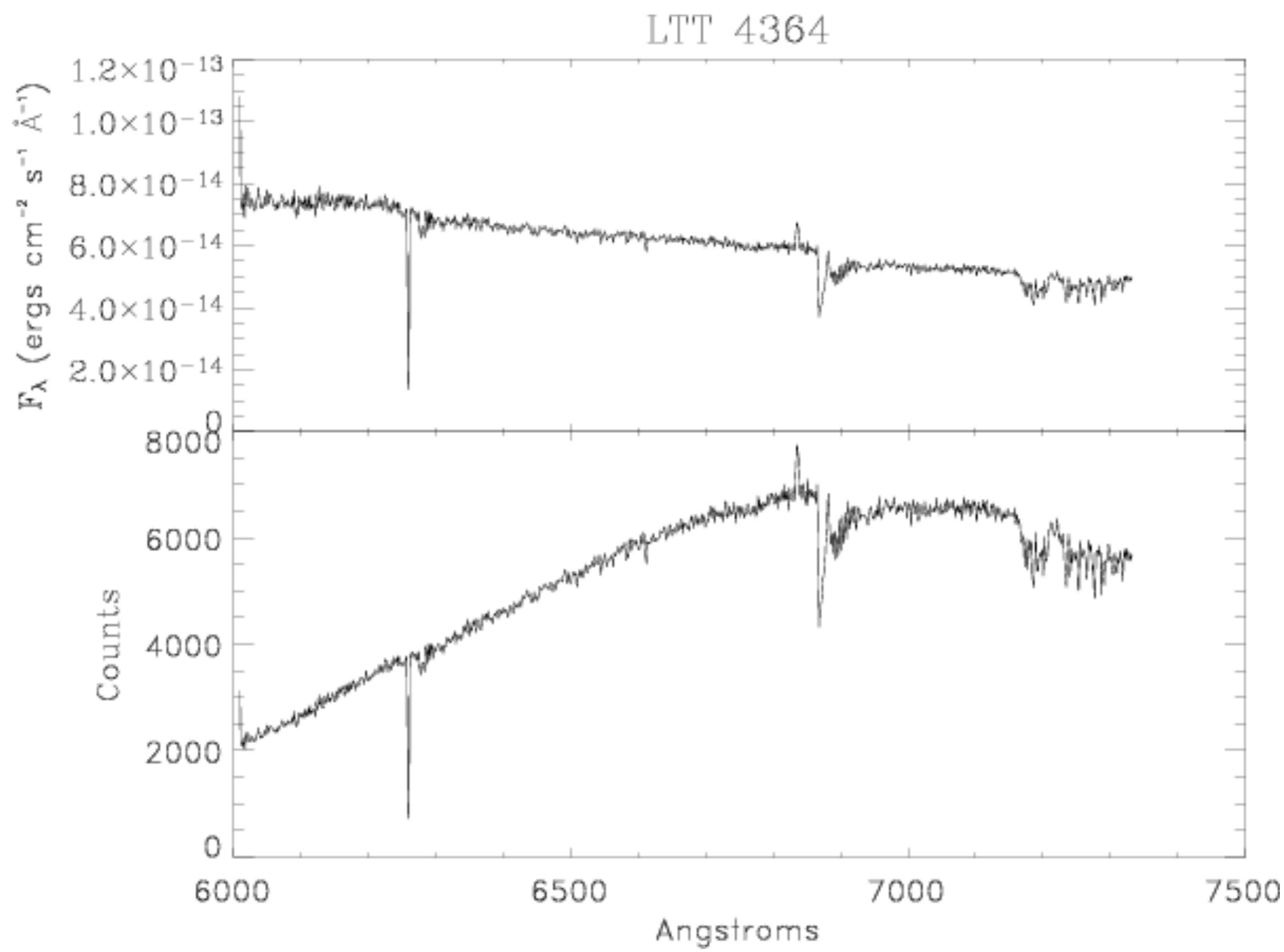
wavelength calibration: map pixel position to emission lines



Spectroscopic Calibration

flux calibration:
observe spectrophotometric standard star

compare observed spectrum (counts) to known spectrum



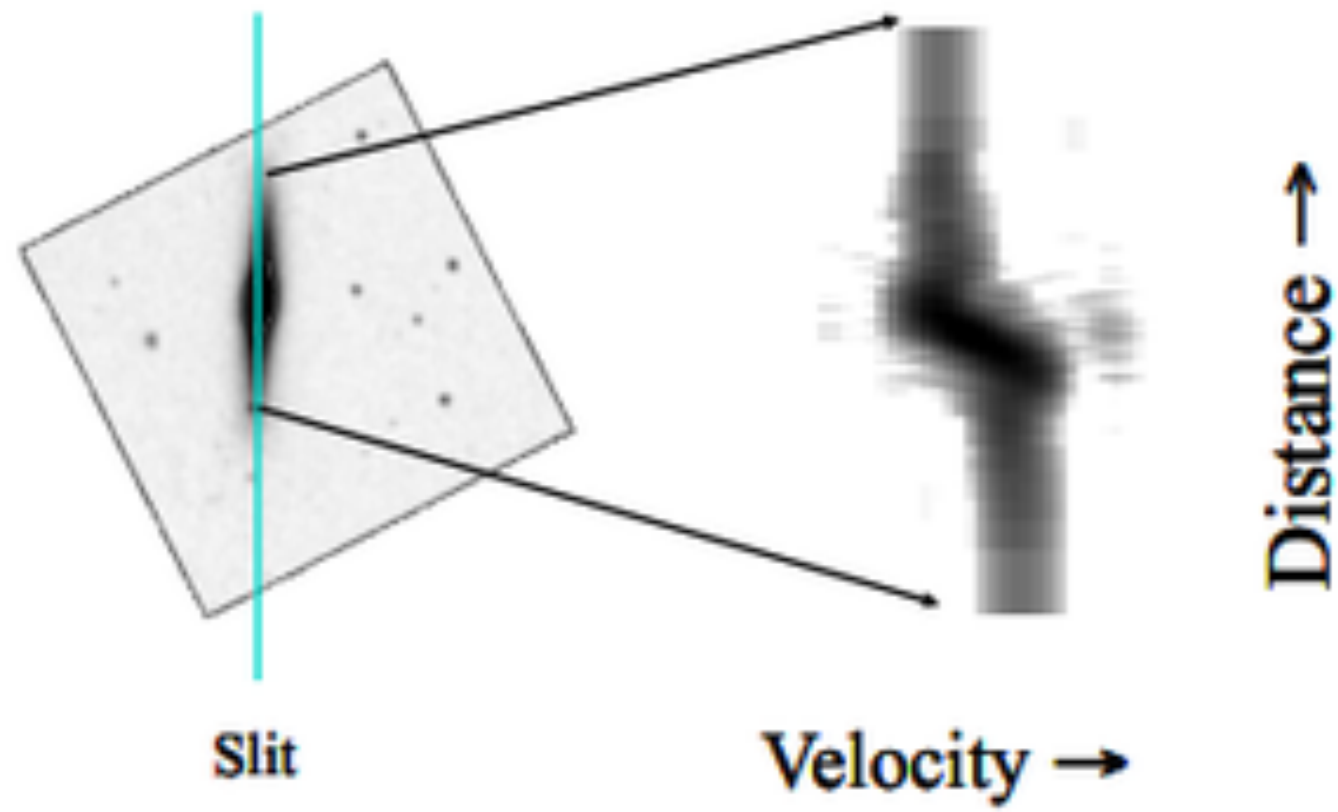
Long-slit spectrographs

most common spectrograph

can only target one (or a few) objects

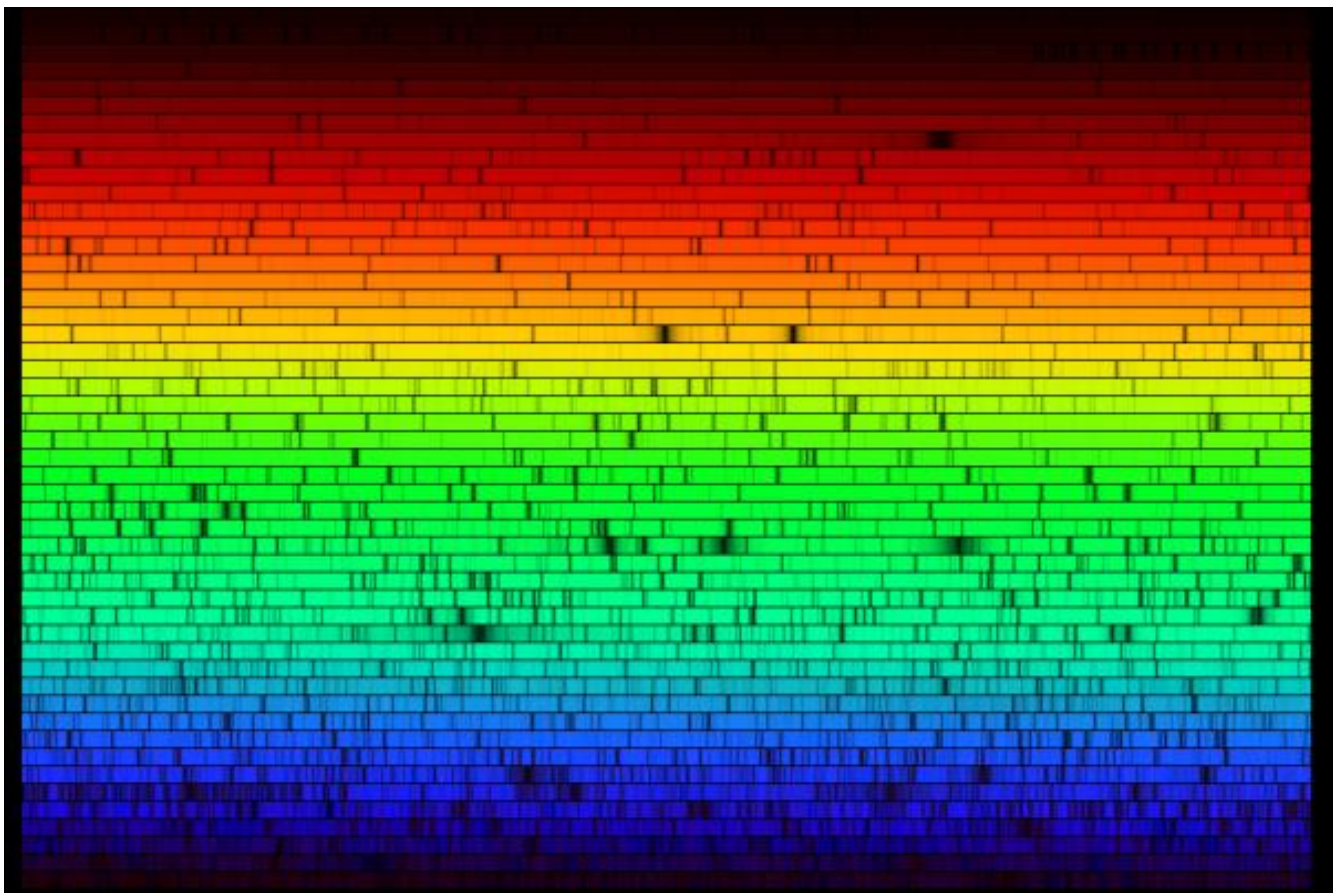
gives spatial variation

very good estimate of sky background



Echelle spectrographs

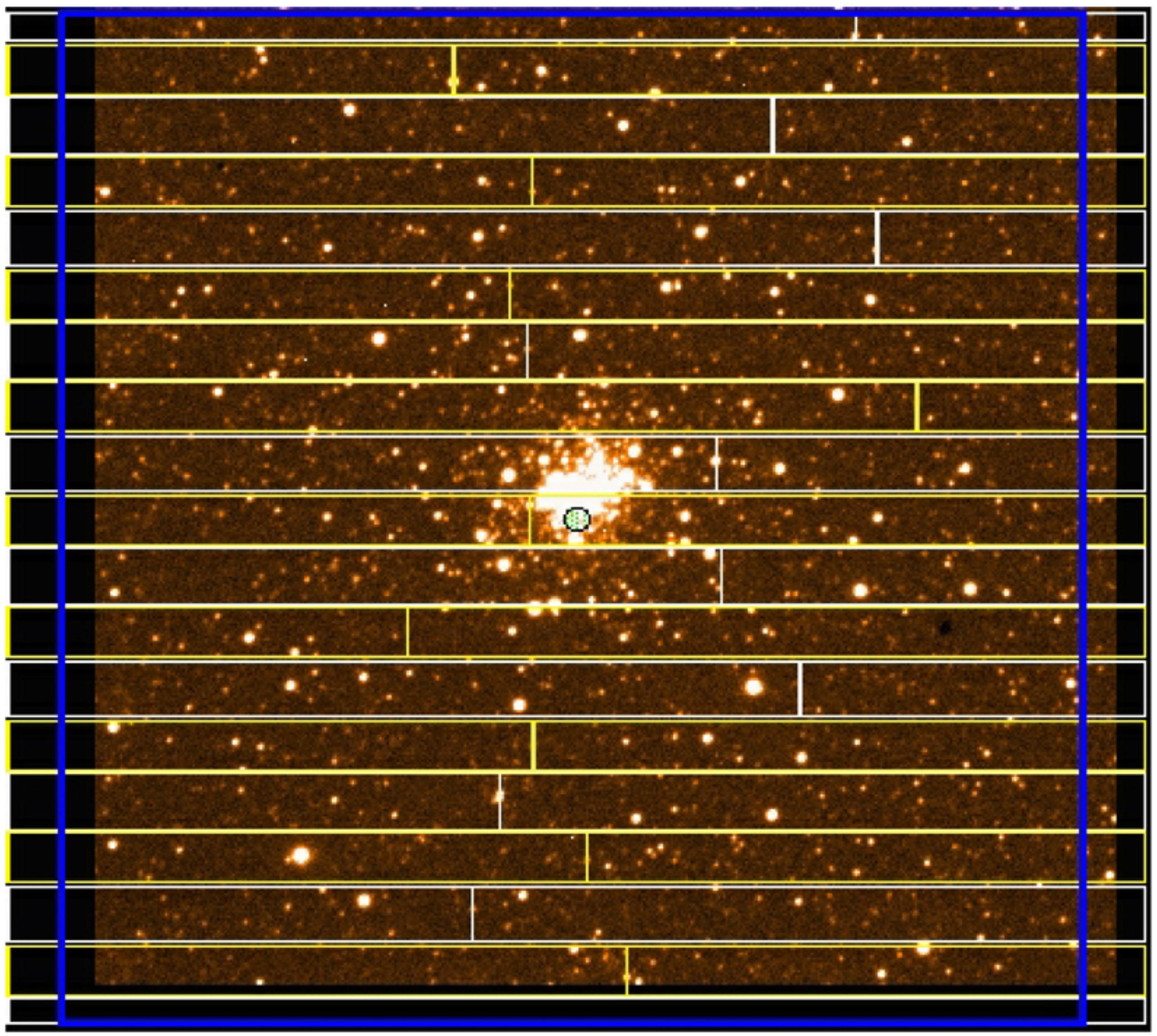
- very high resolution long-slit spectrographs
- have additional elements to fit entire spectrum onto CCD
- only for bright objects



Echelle spectrum of the Sun, 4000-7000Å

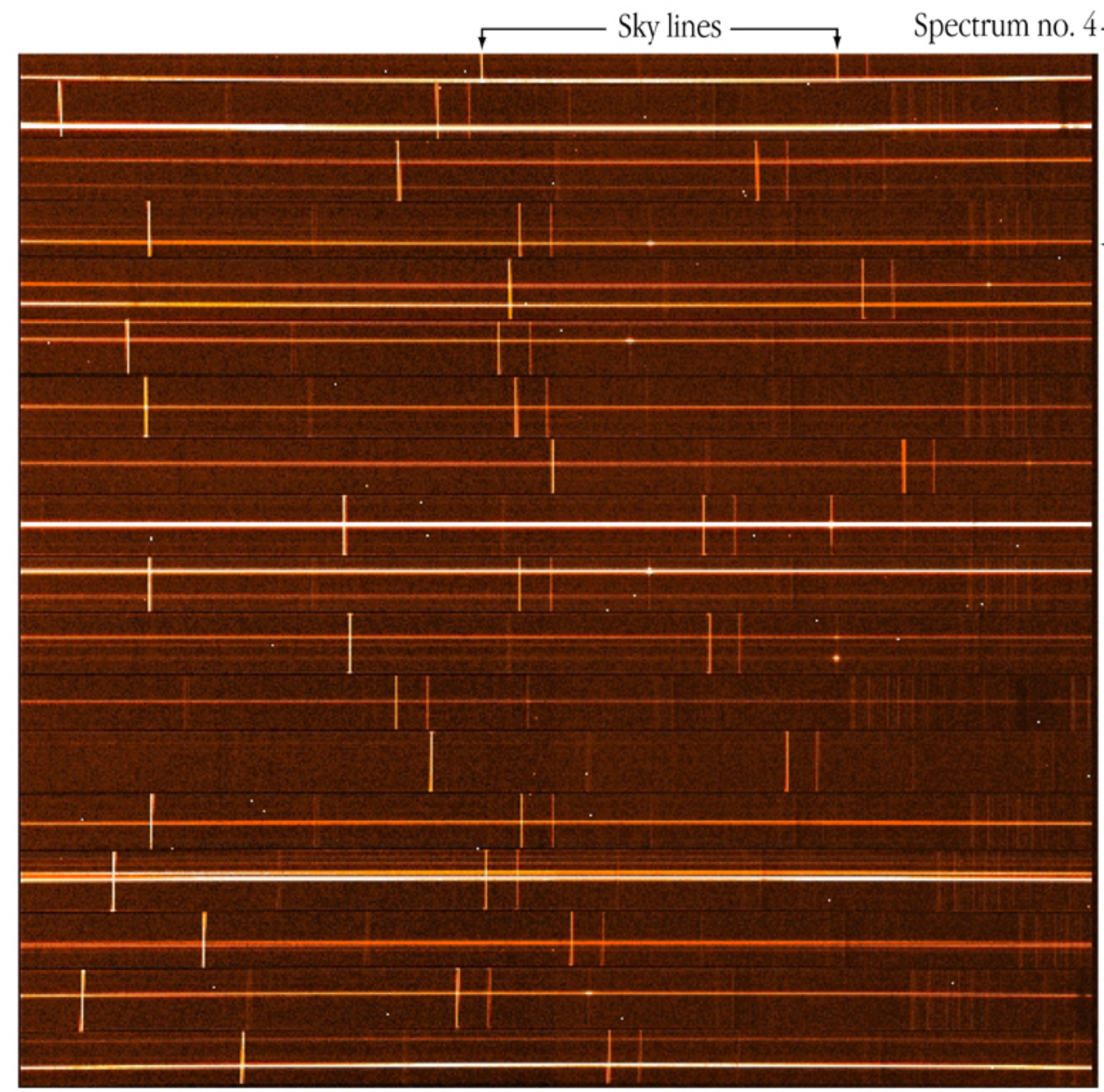
Multi-object spectrographs

make a mask with multiple slits, one per target



Open Cluster NGC 330 in SMC - VLT UT1 + FORS1 (MOS-mode)

ESO PR Photo 38c/98 (7 October 1998)

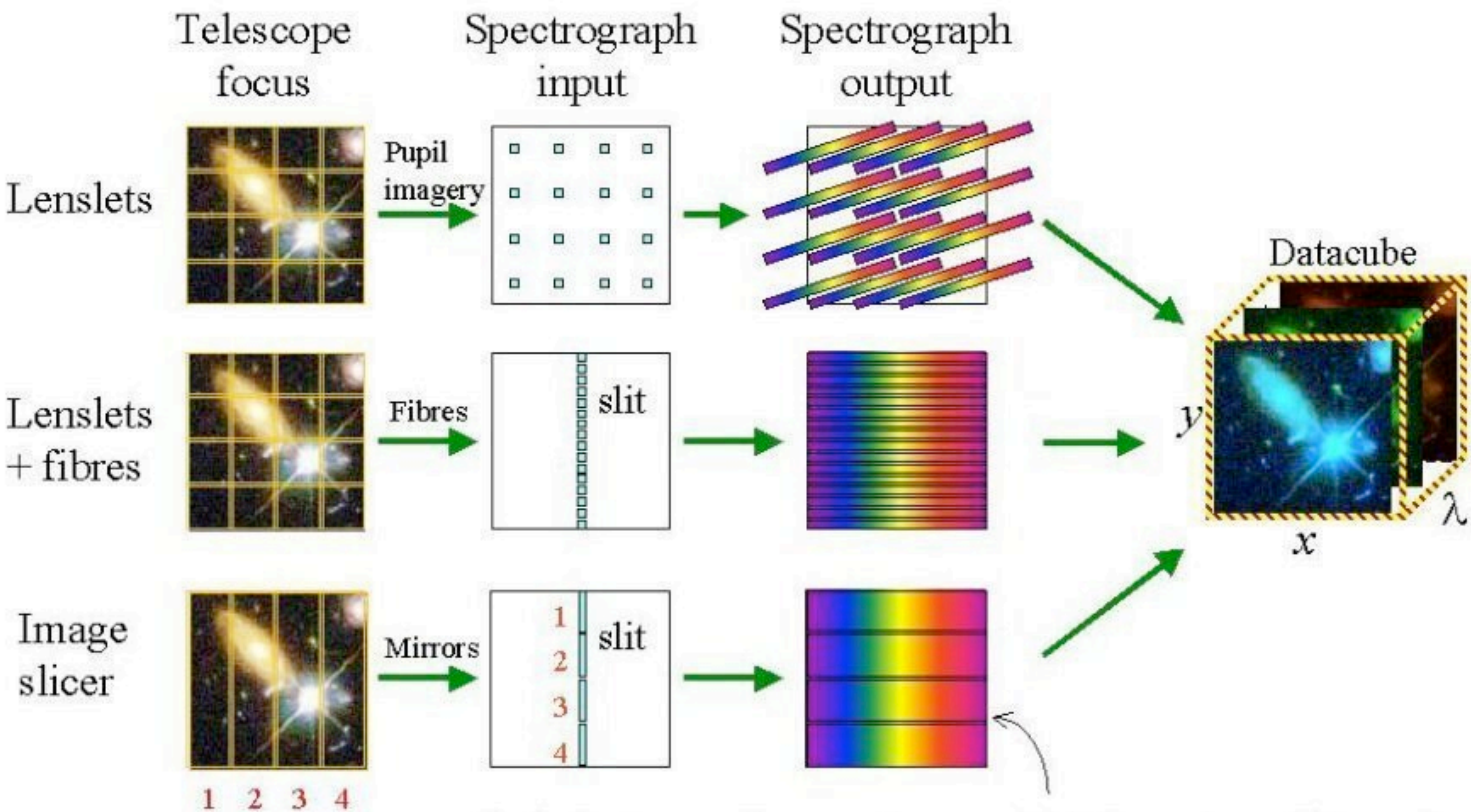


Spectra of Stars in Open Cluster NGC 330 in SMC - VLT UT1 + FORS1 (MOS-mode)

ESO PR Photo 38d/98 (7 October 1998)

Integral-Field Units

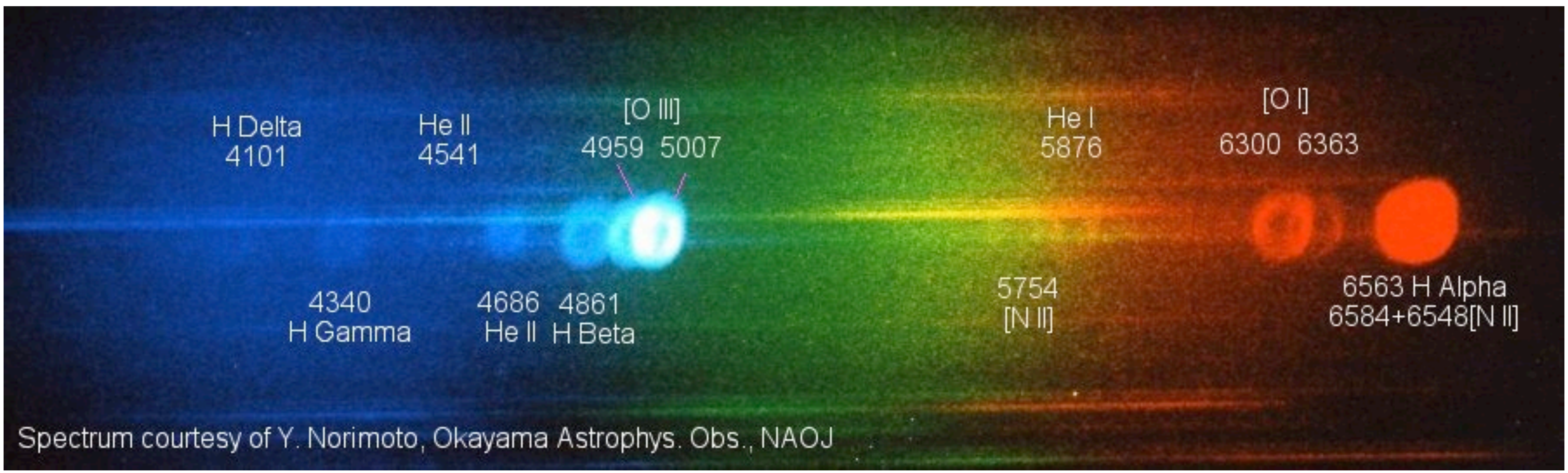
divide image into "spaxels" (not quite pixels)



Spectroscopy Lab

measure strengths of several emission lines of a bright planetary nebula or star-forming nebula

use line ratios to determine gas temperate



slitless(!) spectrum of the Ring Nebula

Spectroscopy Lab - Preparation

- http://www.astro.sunysb.edu/anja/PHY517_AST443/emission_nebulae_lab_draft.pdf
- elevated sky levels are detrimental to spectroscopy
- ‣ need to choose dark nights: Moon close to New Moon, no football games??
- select a suitable target: bright emission line nebula (star-forming region, such as the Orion nebula, or bright planetary nebula, such as the Ring nebula) *google is your friend...*
- ensure that it has good visibility for several hours per night
- pick three dates (two for back-up due to weather / technical issues) between today and Nov 4th; e-mail me your observing request

(A brief intro to)

Statistics

Statistics in Astronomy

- we are almost always working in the low signal-to-noise regime
- have to be very careful to make correct inferences from our data!
- robust (and advanced) statistical techniques play a very important role in astronomy

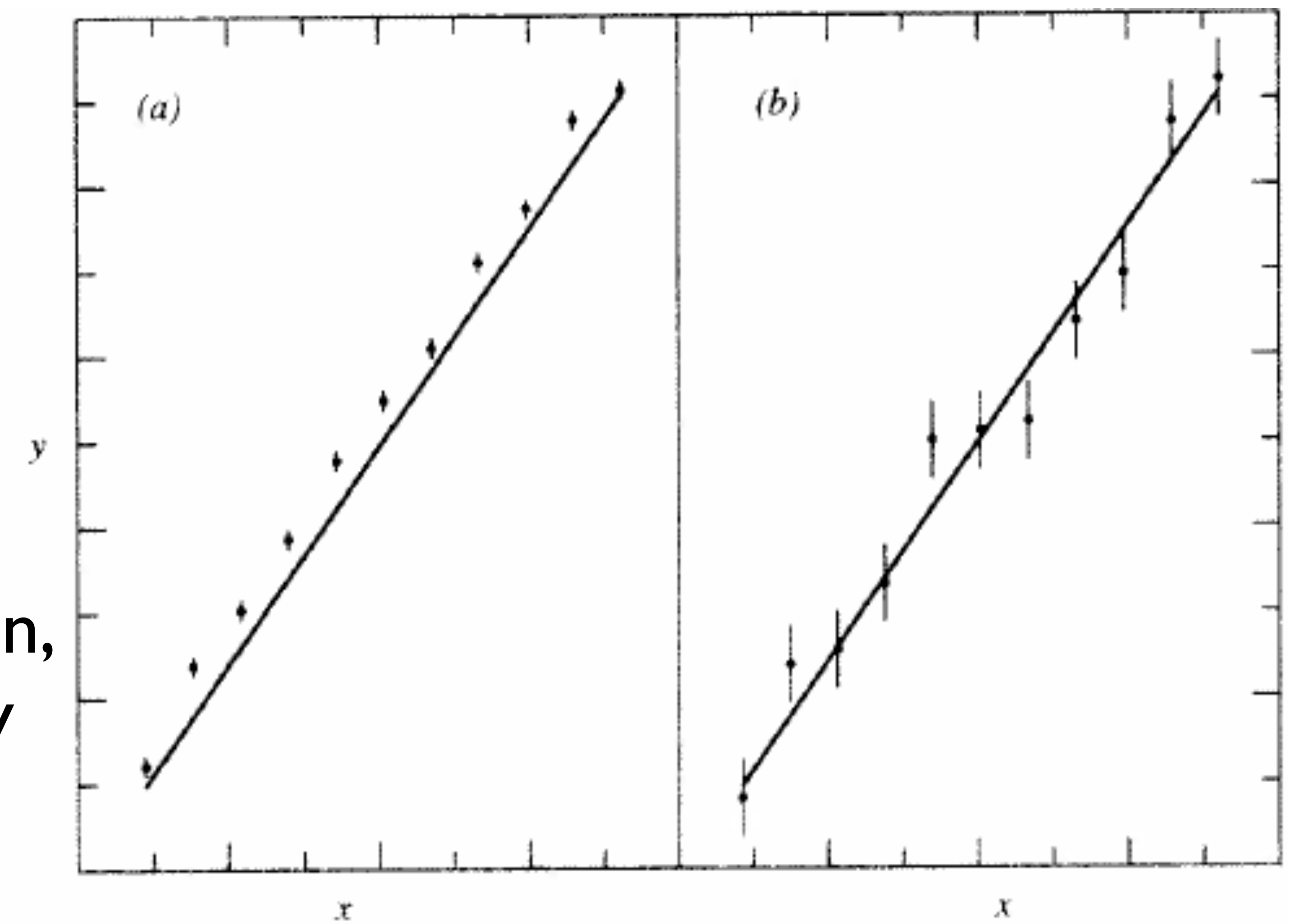
“Error”

- error: difference between *measured* and *true* value
- can be due to:
  - random fluctuations (statistical error)
  - instrumental / algorithmic limitations (systematic error)
  - mistakes (illegitimate error)
- measurements are meaningless if not accompanied by an estimate of the error
- but truth is unknown, have to estimate error indirectly

Accuracy vs. Precision

- accuracy: how close a measurement is to the truth
- precision: size of (statistical) error bars

high precision,
low accuracy



low precision,
high accuracy

FIGURE 1.1
Illustration of the difference between precision and accuracy. (*a*) Precise but inaccurate data. (*b*) Accurate but imprecise data. True values are represented by the straight lines.

Sample vs. Parent Distribution

- measurement x_i of a quantity x:
  - approximates x
  - not necessarily equal to x because of statistical errors
- many measurements x_i:
  - expected to be distributed about true value
  - sample distribution
- parent distribution:
  - probability of particular result from single measurement
  - idealized outcome of infinite number of measurements

the sample distribution *samples* the parent distribution

Mean / average

- (unweighted) mean of the sample distribution:

$$\bar{x} = \frac{1}{N} \sum_i x_i$$

- IF there are no systematic errors, the mean of the parent population is:

$$\mu = \lim_{N \to \infty} \frac{1}{N} \sum_i x_i$$

Deviation / variance / std. deviation

- deviation of one measurement: $d_i = x_i - \mu$

- sample variance: average of the squares of the deviations

$$\sigma^2 = \frac{1}{N}\sum_i (x_i - \mu)^2$$

- when computing from sample population:

$$s^2 = \frac{1}{N-1}\sum_i (x_i - \bar{x})^2$$

- standard deviation: $\sigma = \sqrt{\text{variance}}$

indicates how much the measurements deviate from the mean

Weighted mean

- previously, all measurements had equal weight
- some measurements are more precise than others; can assign weight w_i to each measurement x_i
- weighted mean:

$$\bar{x} = \frac{\sum_i w_i \, x_i}{w_i}$$

- for reasonable (Gaussian) distributions, optimal weight is the inverse of the variance of each measurement:

$$\bar{x} = \frac{\sum_i x_i/\sigma_i^2}{1/\sigma_i^2}$$

Uncertainty on the mean

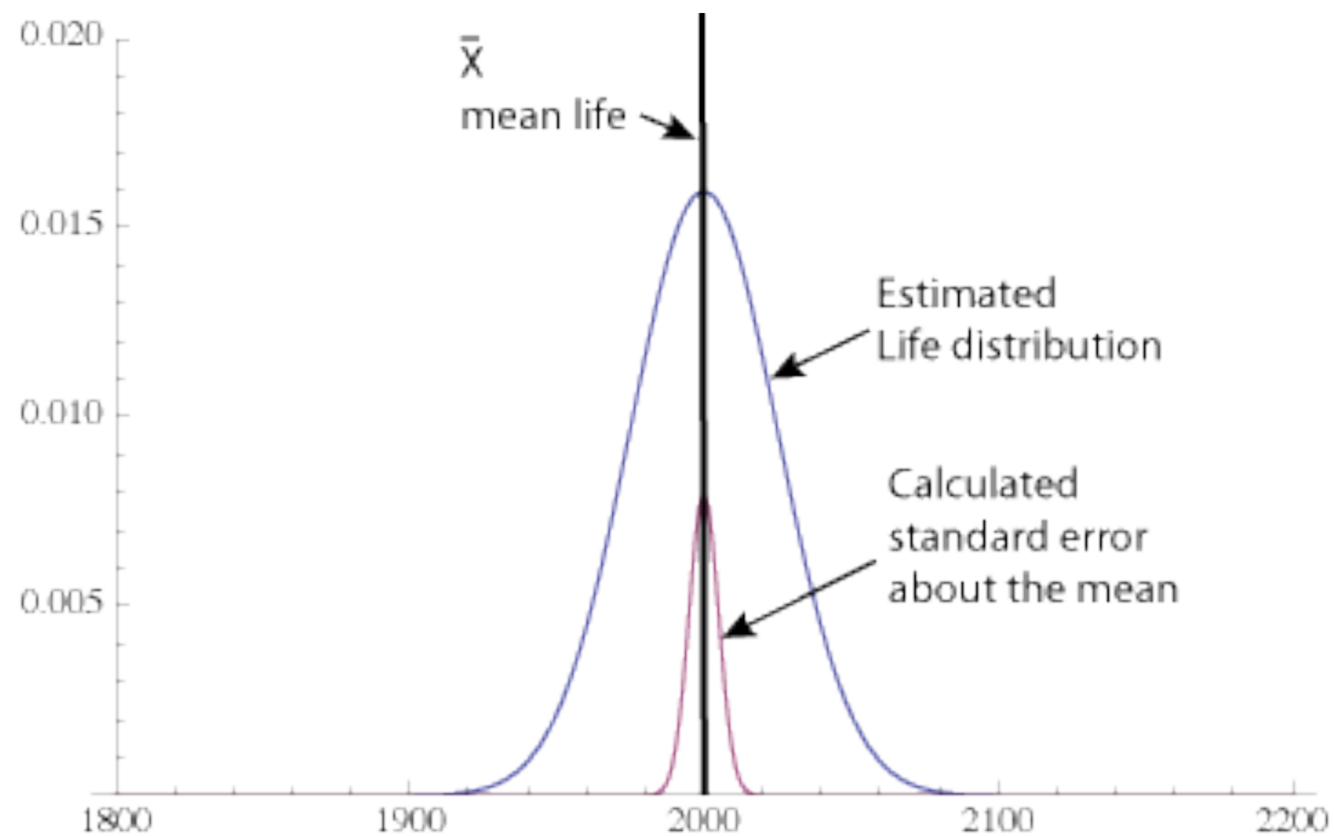
- variance and std. deviation are measures of the *width* of a distribution

- the location of the mean can be determined more precisely:

- uncertainty on the mean:

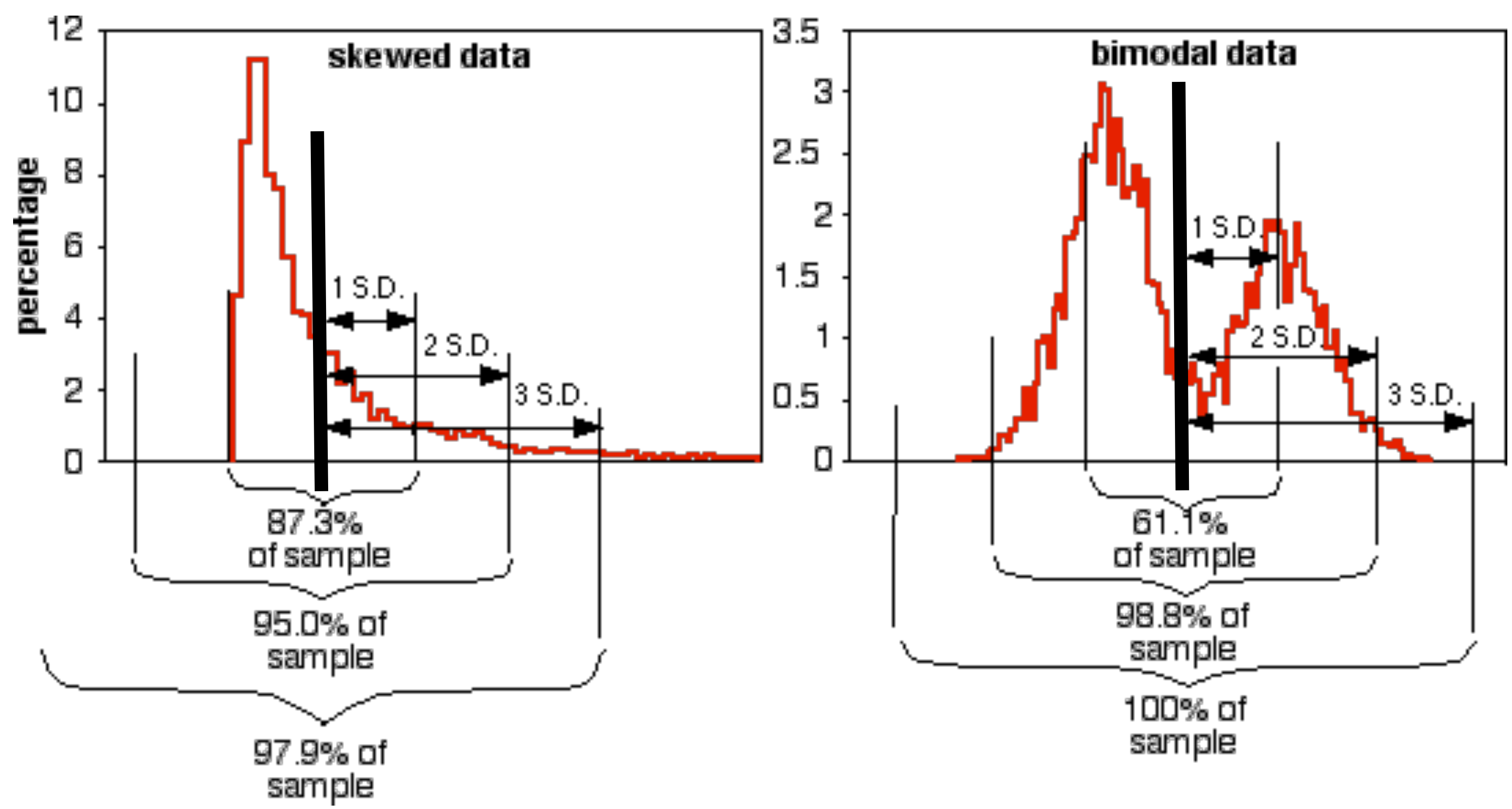
$$\sigma_{\bar{x}}^2 = \frac{\sigma^2}{N}$$

$$\sigma_{\bar{x}}^2 = \frac{1}{\sum_i 1/\sigma_i^2}$$



But what does it all mean?

- can calculate mean, variance, etc. for any set of data points
- *that does not guarantee that they are useful descriptions of the distribution !*

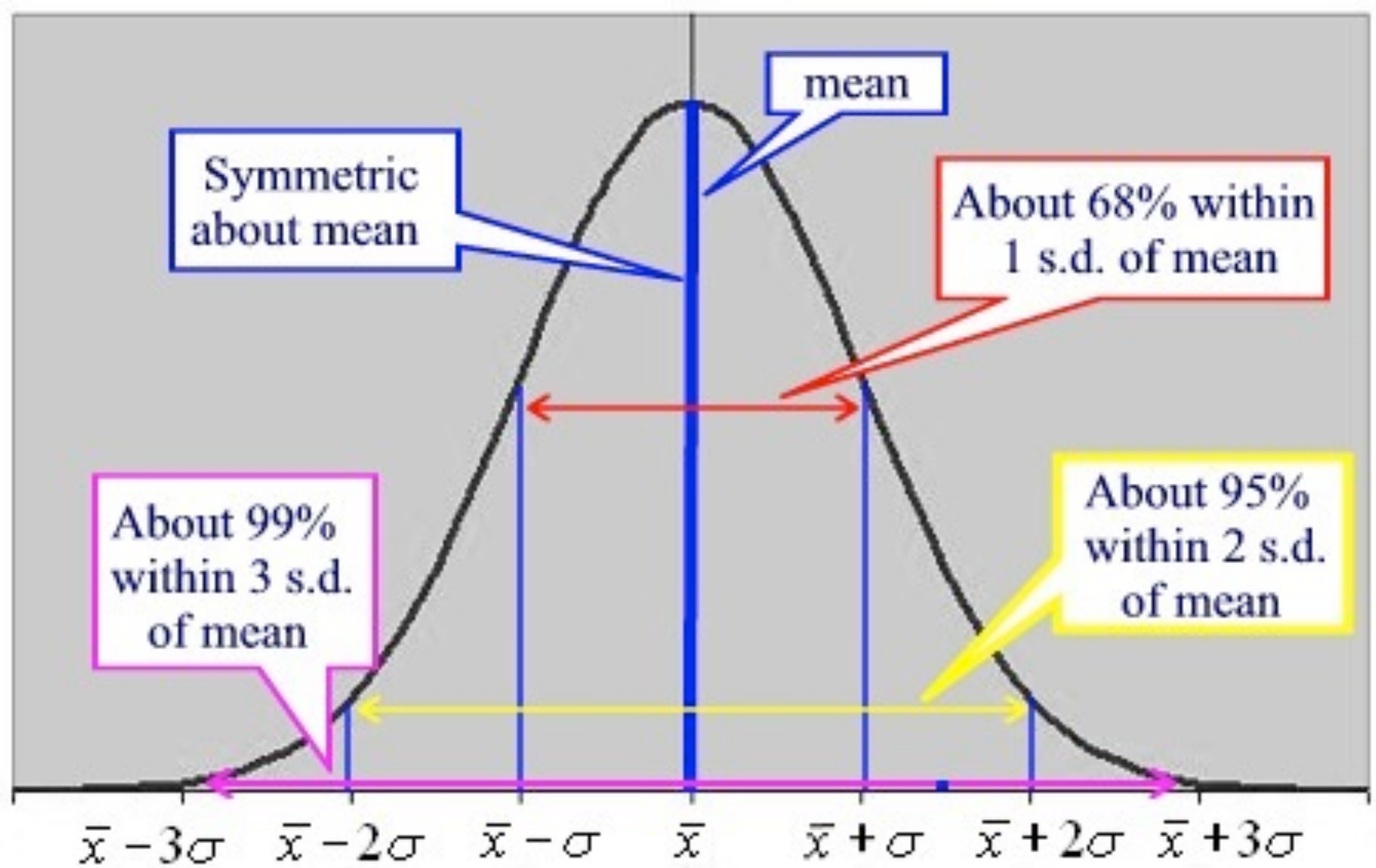


Gaussian / Normal Distribution

- probability for single measurement to yield *x* :

$$p(x;\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]$$

- entirely described by mean μ and std. dev. σ
- ideal case scenario!



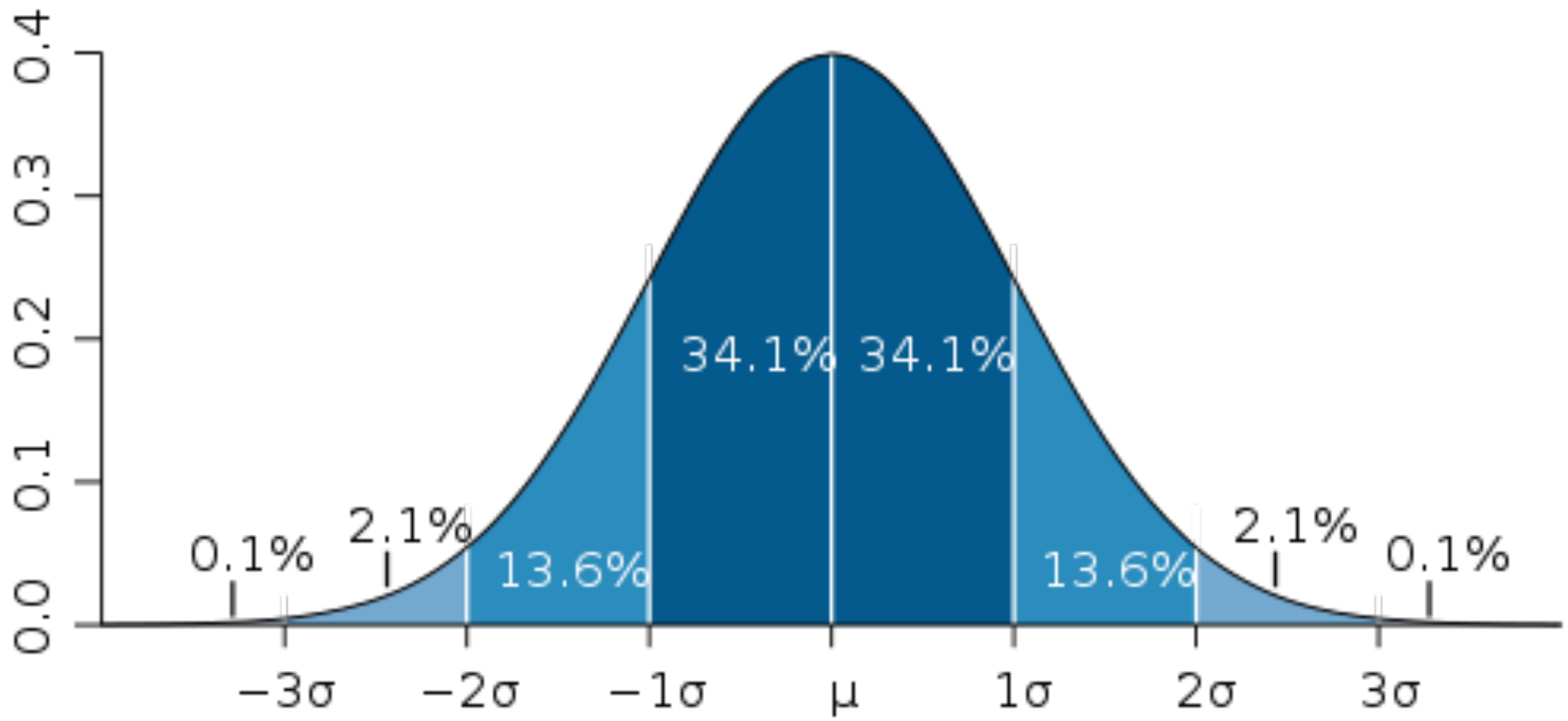
Gaussian / Normal Distribution

- many processes in Nature described by a normal (or log-normal) distribution

- central limit theorem: the distribution of a *large number* of *random, independent draws* will tend to a normal distribution

Gaussian / Normal Distribution

well-defined relation between the probability of occurrence and number of standard deviations away from mean



measurements should fall:

- within 1σ of the mean 68.3% of the time
- within 2σ of the mean 95.4% of the time
- within 3σ of the mean 99.73% of the time

Significant detections?

- the significance of a detection is often quoted in “sigmas” to indicate the probability that the signal is (in)consistent with a random fluctuations

- only a valid measure of probability if the background distribution is Gaussian!

- in particle physics: need $>5\sigma$ to claim detection

- in astronomy: detections are claimed at $>3\sigma$

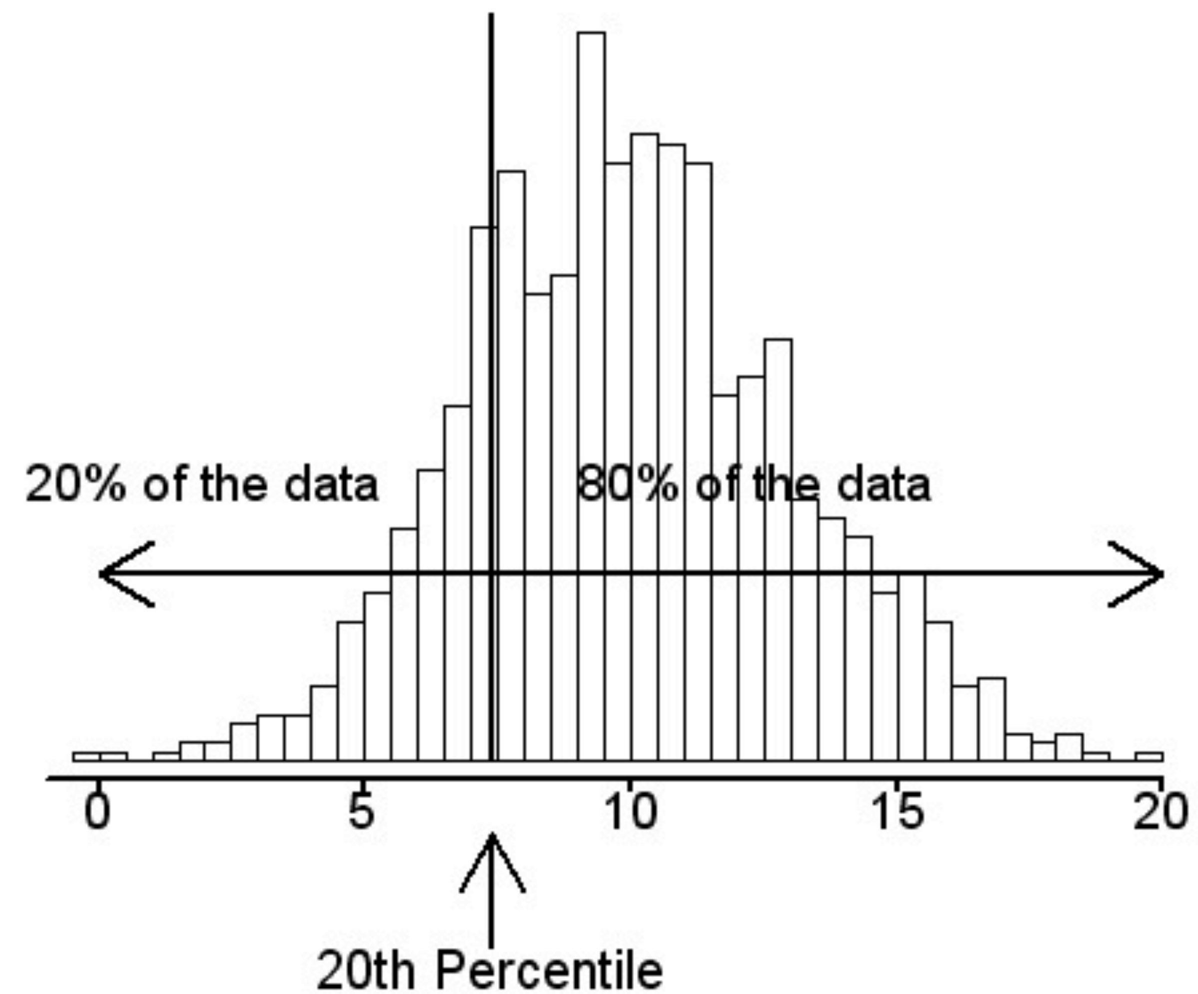
- don’t trust claims below 3σ

Non-Gaussian distributions

- what if your distribution is non-Gaussian?
- have to decide on case-by-case basis

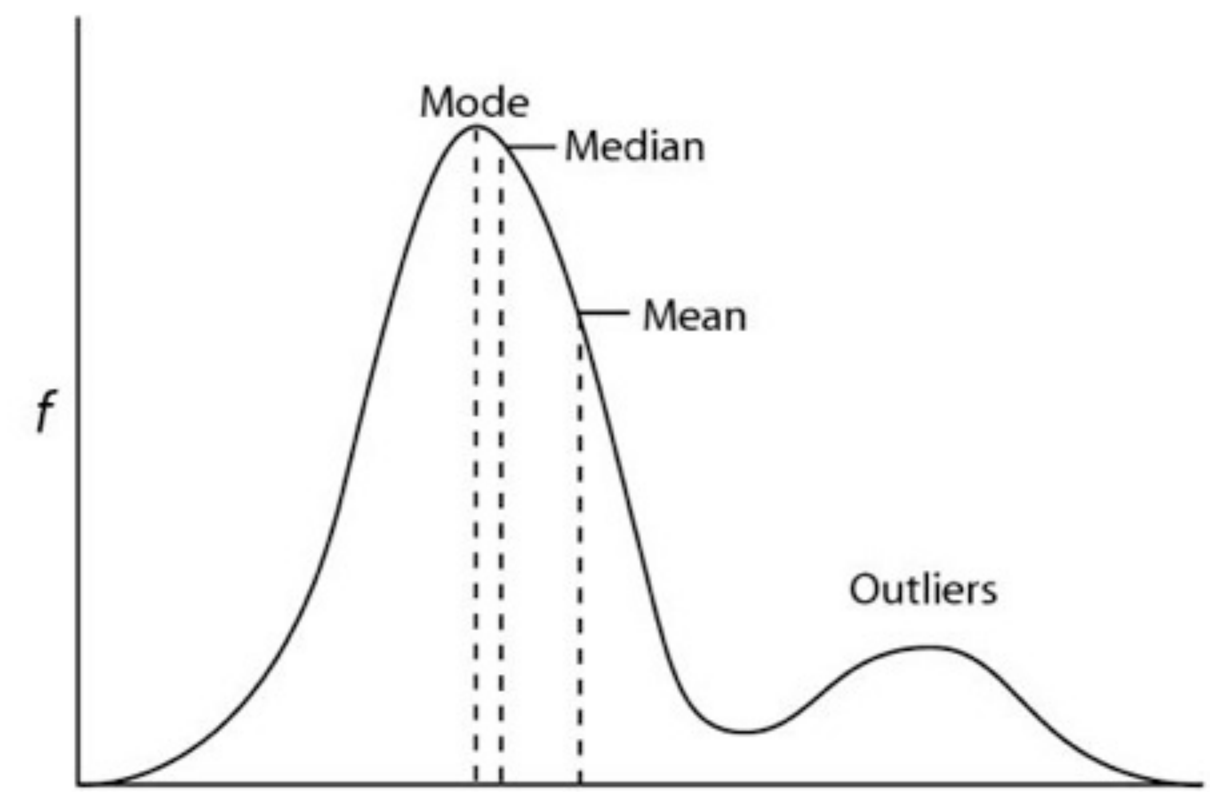
- percentiles (quartiles): can always sort your data, quote values that are above certain percentage of population

- median: 50th percentile; half the data above, half below



Outliers

- for normal distribution, median = mean
- what if distribution is "almost" normal, but has a few outliers? *e.g. cosmic rays in dark frame*
- mean: significantly affected by outliers
- median: robust against (small number of) outliers
- sometimes, it's ok to remove gross outliers ("sigma-clipping"), but need to make sure not to bias your results!



Error propagation

- often, want to determine dependent variable *x* that is a function of one or more measurements

e.g. $x = f(u, v)$ *u* and *v* have (measured variances):

$$\sigma_u^2 = \frac{1}{N-1}\sum_i (u_i - \bar{u})^2 \qquad \sigma_v^2 = \frac{1}{N-1}\sum_i (v_i - \bar{v})^2$$

covariance between *u* and *v*:

$$\sigma_{uv}^2 = \frac{1}{N-1}\sum_i (u_i - \bar{u})(v_i - \bar{v})$$

note: if *u* and *v* are independent, covariance vanishes for large N

Error propagation

- Gaussian case: variance in *x* can be expressed in terms of variance in *u* and *v*, and the covariance between them:

$$\sigma_x^2 = \sigma_u^2 \left(\frac{\partial x}{\partial u}\right)^2 + \sigma_v^2 \left(\frac{\partial x}{\partial v}\right)^2 + 2\sigma_{uv}^2 \left(\frac{\partial x}{\partial u}\right)\left(\frac{\partial x}{\partial v}\right)$$

- if u and v independent:

$$\sigma_x^2 = \sigma_u^2 \left(\frac{\partial x}{\partial u}\right)^2 + \sigma_v^2 \left(\frac{\partial x}{\partial v}\right)^2$$

e.g. $x = a\,u\,v$

with a = constant: $\dfrac{\sigma_x^2}{x^2} = \dfrac{\sigma_u^2}{u^2} + \dfrac{\sigma_v^2}{v^2}$

Resampling

- what about errors on dependent quantities in non-Gaussian case?

- can use resampling methods (bootstrap, jackknife, ...)

- e.g. bootstrap: resampling with replacement

- *N* measurements
- draw from your measurements *N* times (can draw same measurement more than once)
- determine derived quantity
- repeat *n* times
- quantify the distribution up from *n* iterations (e.g. percentiles)

Model fitting

- to fit a model to a dataset, need to quantify how good the fit fits the data

- if errors are Gaussian, optimal statistic is χ^2 ("chi-squared")

$$\chi^2 = \sum_i \frac{(D[x_i] - M[x_i])^2}{\sigma_i^2}$$

$D[x_i]$ are the data values; $M[x_i]$ are the values of the model evaluated at positions x_i

(note similarity to normal probability distribution!)

Model fitting

- the best-fitting model is the one that minimizes the χ^2 value

$$\chi^2_{\min} = \sum_i \frac{(D[x_i] - M_{\mathrm{best}}[x_i])^2}{\sigma_i^2}$$

how to find the best-fit model:

- brute force: make a grid of parameter values, calculate χ^2 for each
- use a minimization algorithm

Model fitting

- you have found the "best-fit" parameters of the model that minimize the χ^2, but is that model actually a good fit?

$$\chi^2_\nu = \frac{\chi^2_{\min}}{\nu}$$

- reduced chi-square: scale best-fit chi-square by ν, the number of free parameters, number of data points minus the number of free model parameters

- example: fitting a line: two model parameters (slope and intercept)

$$\nu = \text{number of data points} - 2$$

Model fitting

- given a random realization of an experiment with ν degrees of freedom, the probability to obtain $\chi^2_{\rm min}/\nu$ or larger is described by the chi-squared distribution
- comparing the measured reduced chi-squared to the expectation is an indication whether the model is an acceptable fit to the data
- for an acceptable model, the remaining deviations should be well described by a random (Gaussian) process

$\chi^2_{\rm min}/\nu \approx 1$	model is a good fit
$\chi^2_{\rm min}/\nu \gg 1$	model is a bad fit
$\chi^2_{\rm min}/\nu \ll 1$	model is a overfitting the data