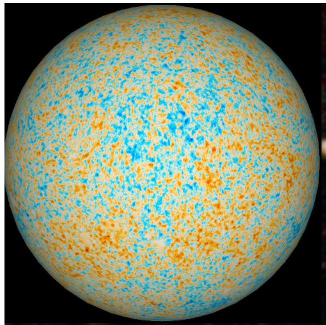
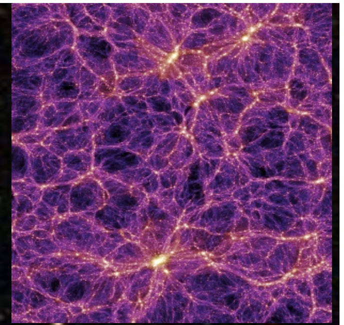
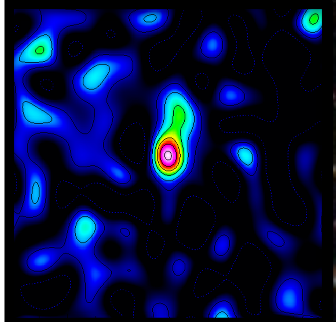


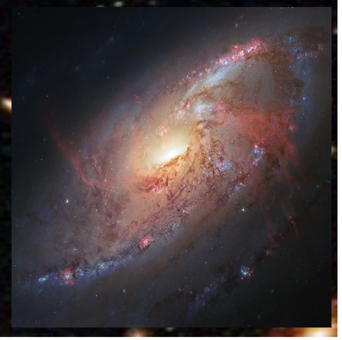


DATA ANALYSIS IN ASTROPHYSICS

Simona Gallerani

SCUOLA NORMALE SUPERIORE





School on Scientific Data Analysis

25th-27th November

Program of the day

- <u>Stefano Carniani</u>: Exploration of 3D spectroscopy data

  HST and ALMA data of distant galaxies:

  1. Why are early galaxies important?
  2. What is redshift?
  3. What is a data cube?
  4. How do we get data?

- <u>Graziano Ucci</u>: Machine learning on Galaxy spectra

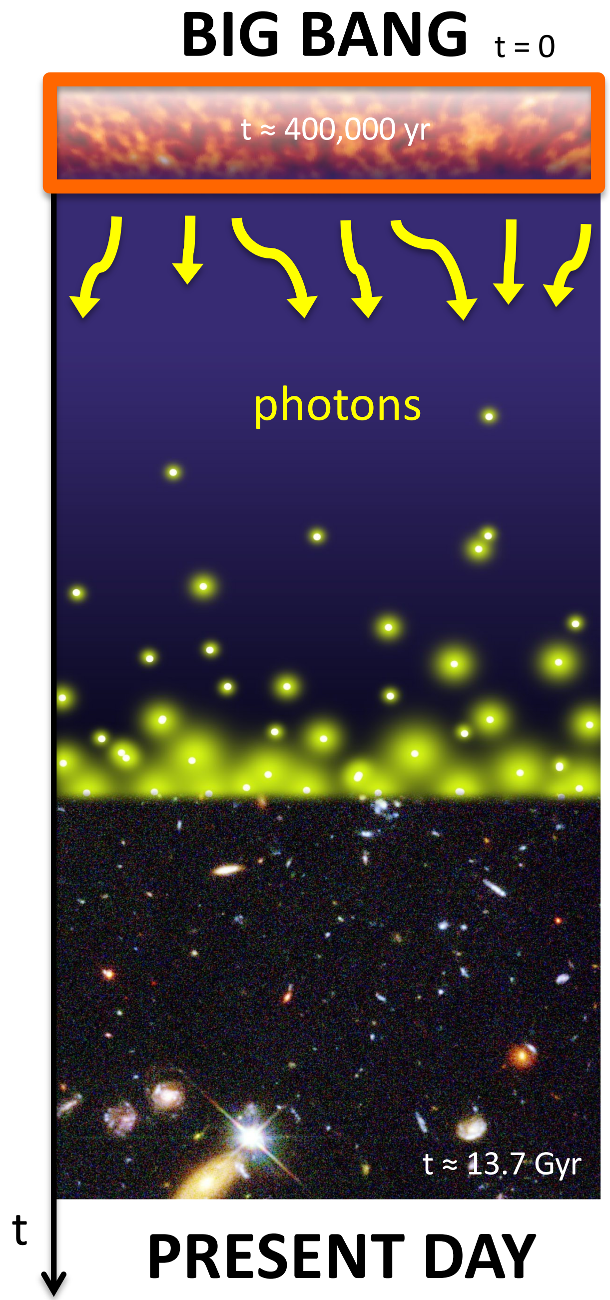
  Emission lines arising from the interstellar medium

  1. What is the interstellar medium?
  2. What is the radiative transfer equation?
  3. What is CLOUDY?

- <u>Andrea Pallottini</u>: Synthetic spectra of galaxies in cosmological simulations

- <u>Viviana Acquaviva</u>: The Science of Data Science (After-dinner speech)

From the Big Bang to the first picture of the Universe



According to the standard cosmological model, the Big Bang represents the origin of our Universe. In this primordial stage it can be described as a hot and dense gas composed of matter and radiation.

As a consequence of the expansion of the Universe this primordial plasma cools down rapidly. Radiation and matter decouple and start evolving separately.

What happens to radiation?

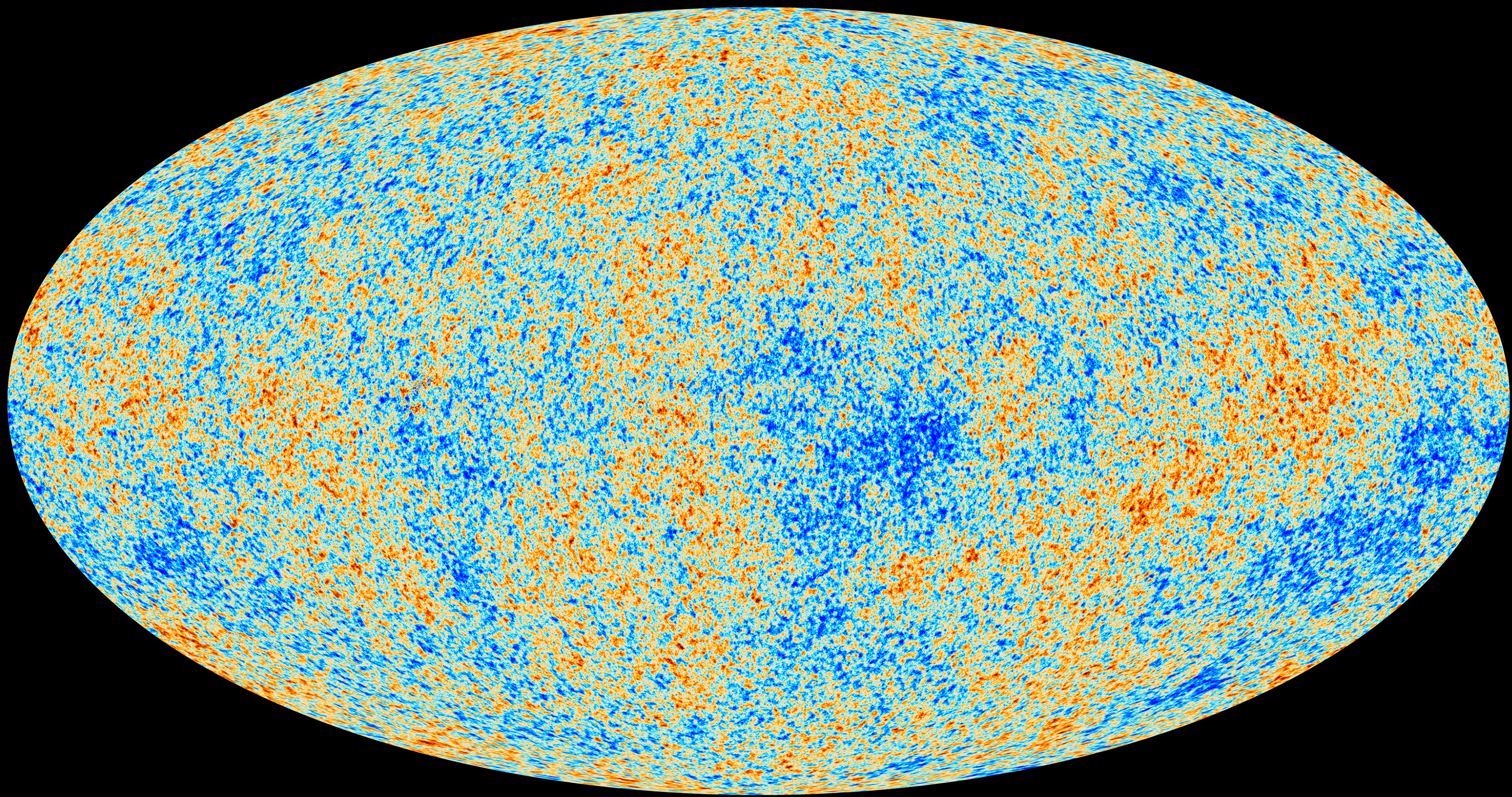
Photons can freely travel towards the Earth, providing us the first picture of the Universe:

THE COSMIC MICROWAVE BACKGROUND

The Planck Space Telescope

Launched into orbit in the 2009

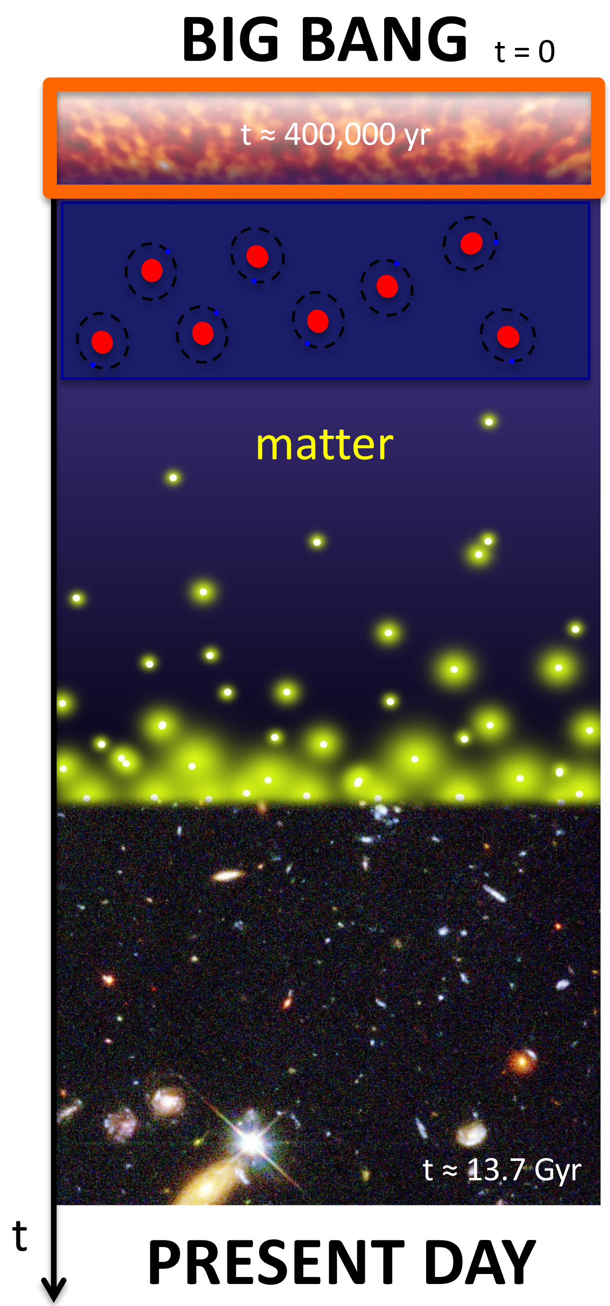
The Cosmic Microwave background



Almost 400 thousands years after the Big Bang, the Universe was a fluid almost 3000 K hot, highly omogeneous but with small fluctuations:
hot (blue) spots represent regions that are hotter (colder) than the average.

TEMPERATURE ⟷ DENSITY

From the Big Bang to the Dark Ages of the Universe



According to the standard cosmological model, the Big Bang represents the origin of our Universe. In this primordial stage it can be described as a hot and dense gas composed of matter and radiation.

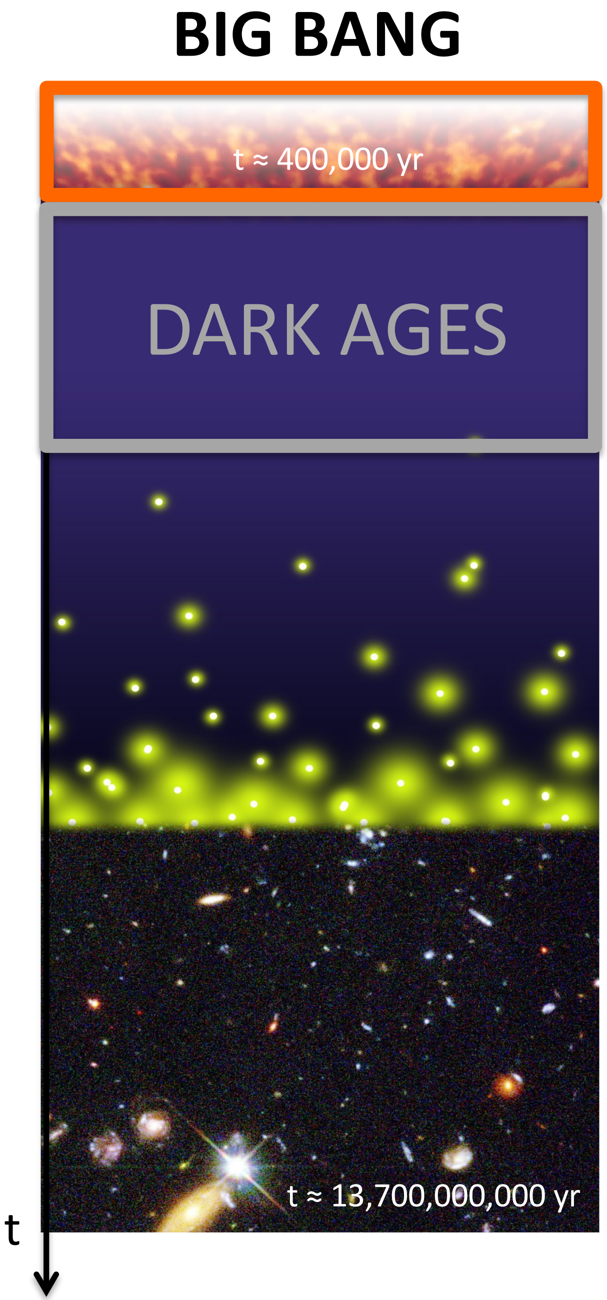
As a consequence of the expansion of the Universe this primordial plasma cools down rapidly. Radiation and matter decouple and start evolving separately.

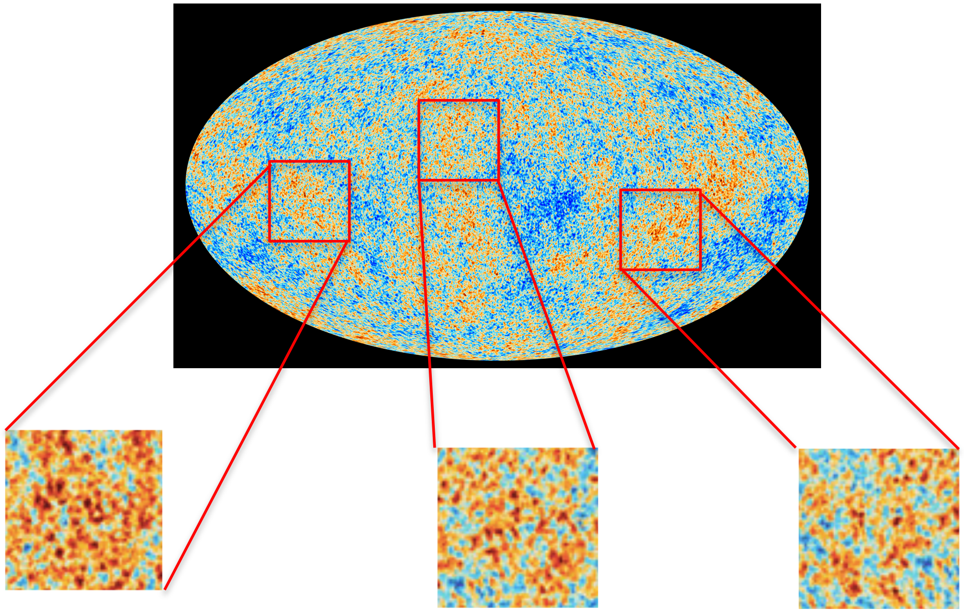
What happens to matter?

The first atoms of hydrogen start to form.
No sources of light are present.

The Universe enters in the so-called **Dark Ages**.

First stars and galaxies





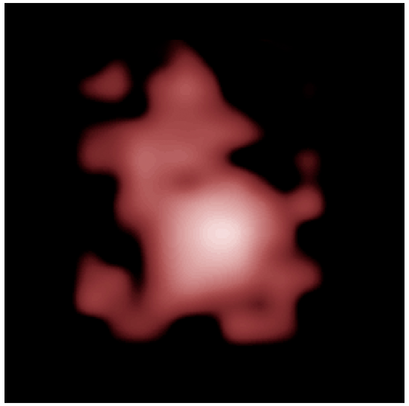
Starting from the denser regions of the Universe
at the time of decoupling, the first sources of light will form.

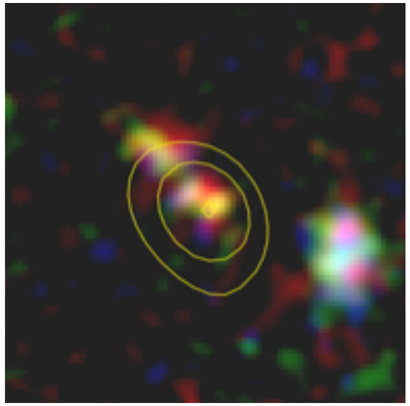
The hotter regions (red spots) revealed by the CMB
represent overdense regions of the Universe.
These grow thanks to gravity and give rise to
the first structures of the Universe.

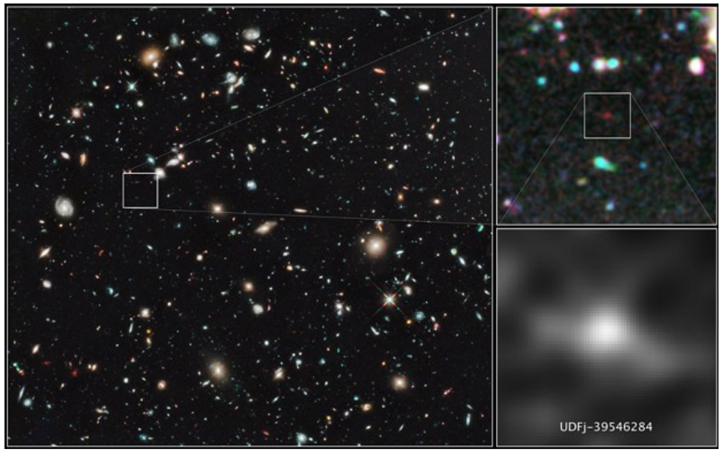
Theoretical models predict that
the first stars and galaxies formed in the Universe
appeared almost 200 million years after the Big Bang

The Cosmic Renaissance: Galaxies

Thanks to powerful ground and space-based telescopes
It is possible to detect light from very distant sources







The Universe is still very young! Less than 1 billion years old:

A child at its first year of primary school

The Hubble Space Telescope

Launched into orbit in the 1990

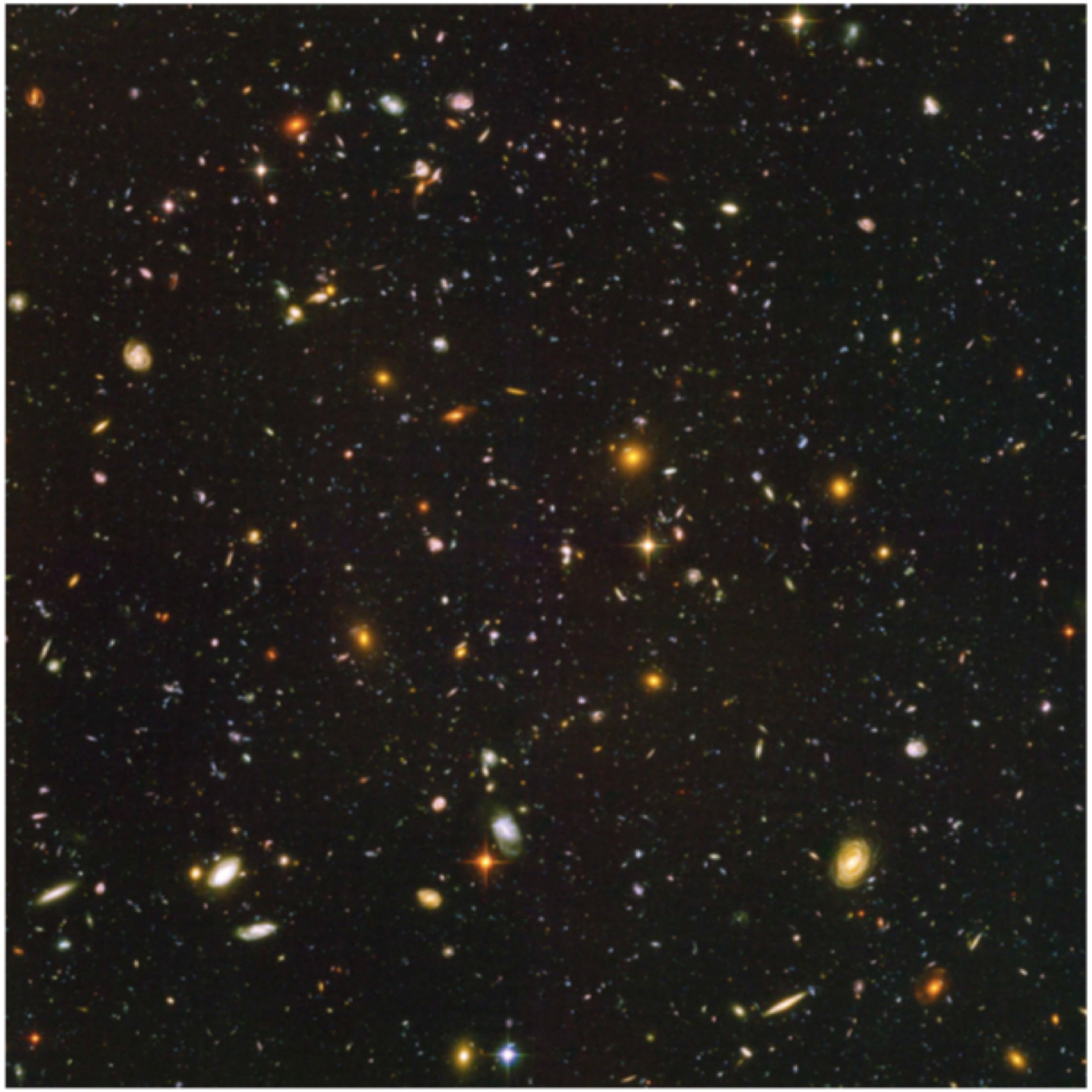
Thousands of new galaxies discovered.

Several of them are extremely far, formed more than 10 Gyr ago!

The Hubble Deep Field

Travelling in the Hubble Deep Field is like travelling in time.
Observing very distant galaxies means exploring very early epochs, close to the Dark Ages of the Universe.

The Hubble Ultra Deep Field



Beckwith et al. (2006)

10^6 second exposure of an 11'x11' region

The exposure time was divided among 4 filters, centered at different wavelengths:

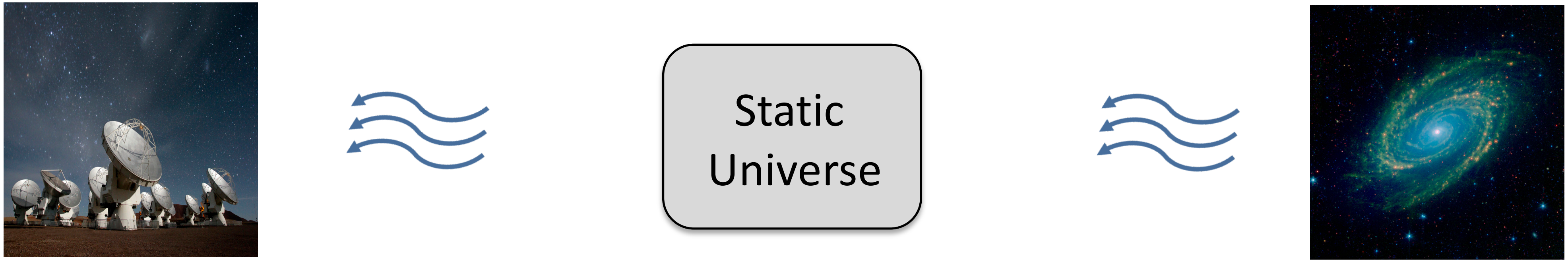
F435W @ 4347 Å
F606W @ 6033 Å
F775W @ 7730 Å
F850LP @ 9082 Å

The image contains > 10^4 objects

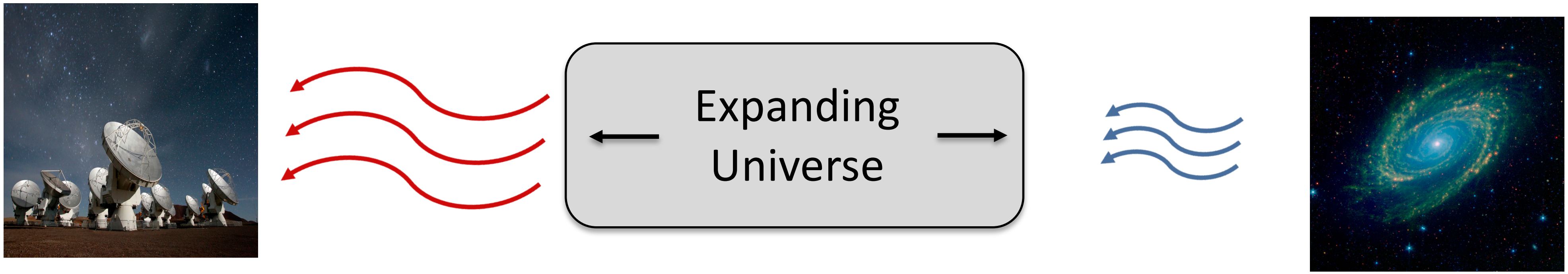
Color rendition of the final HUDF image cropped to display an area of uniform exposure. The color mapping used to produce this rendition is:

blue = combination of F435 & F606W,
green = combination of F606W & F775W,
red = combination of F775W & F850LP.

The expansion of the Universe



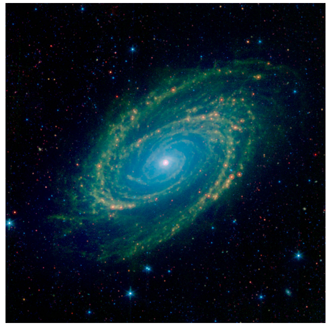
the wavelength λ_e of the radiation emitted by a source at a given distance d from on observer travels into a static space-time. It arrives to the observer with the same wavelength $\lambda_o = \lambda_e$.



the wavelength λ_e travels into an expanding space-time. This results into a stretching of the radiation wavelength. It arrives to the observer with a wavelength $\lambda_o > \lambda_e$.

The cosmological redshift

Given the *finite speed of light*,
the farther is the source, the longer is the travel,
and the stronger is the radiation stretching



The *cosmological redshift* measures this radiation stretching and it is defined as:

$$\boxed{1 + z = \frac{\lambda_o}{\lambda_e}}$$

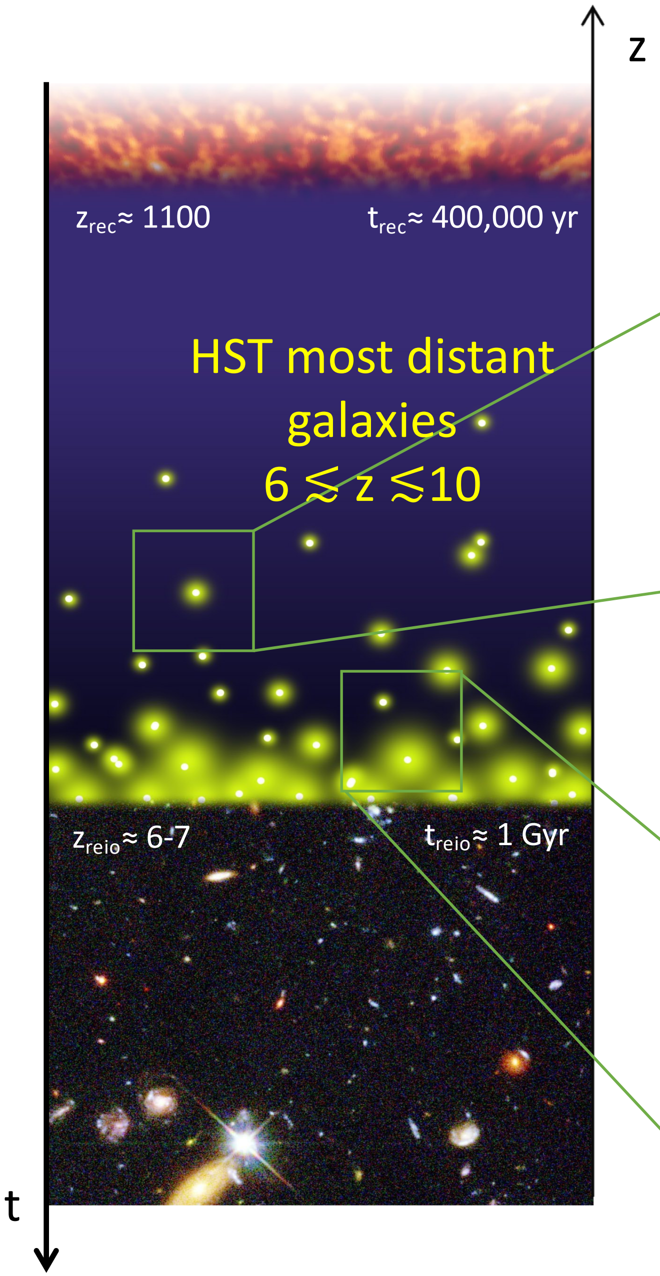
It provides information on the *past epochs* at which photons were originarily emitted

Long wavelengths come from *high-redshift sources*, namely sources *far away from the Earth*.

HST is sensitive to infrared photons.

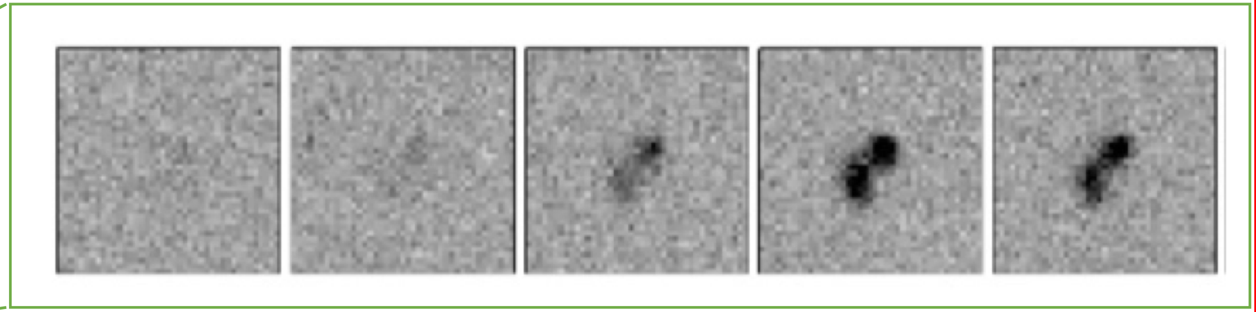
It can detect galaxies up to *z≈10* (age of the Universe: ≈ 500 million years)

How do we detect and characterize high-z galaxies?



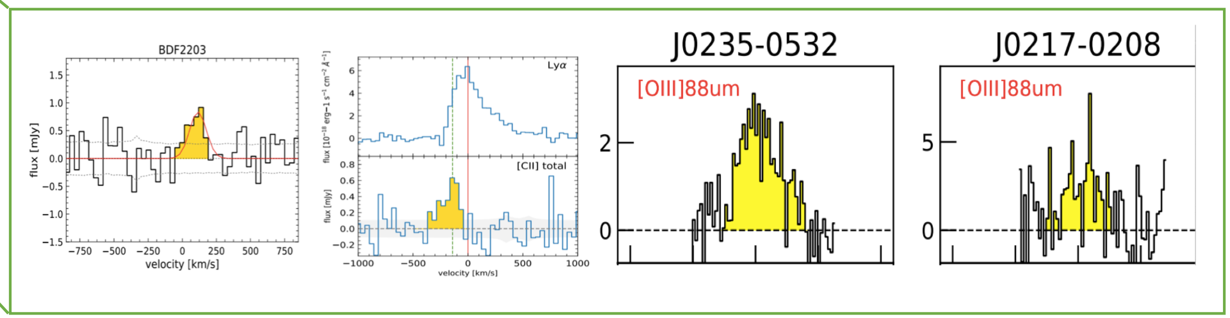
IMAGING

Dropout technique



SPECTROSCOPY

Far infrared emission lines



What is the type of data coming out from a telescope?

It depends on the instruments on-board

(e.g.) HST INSTRUMENTS

- **ACS: Advanced Camera for Surveys**
- **WFC3: Wide Field Camera 3**

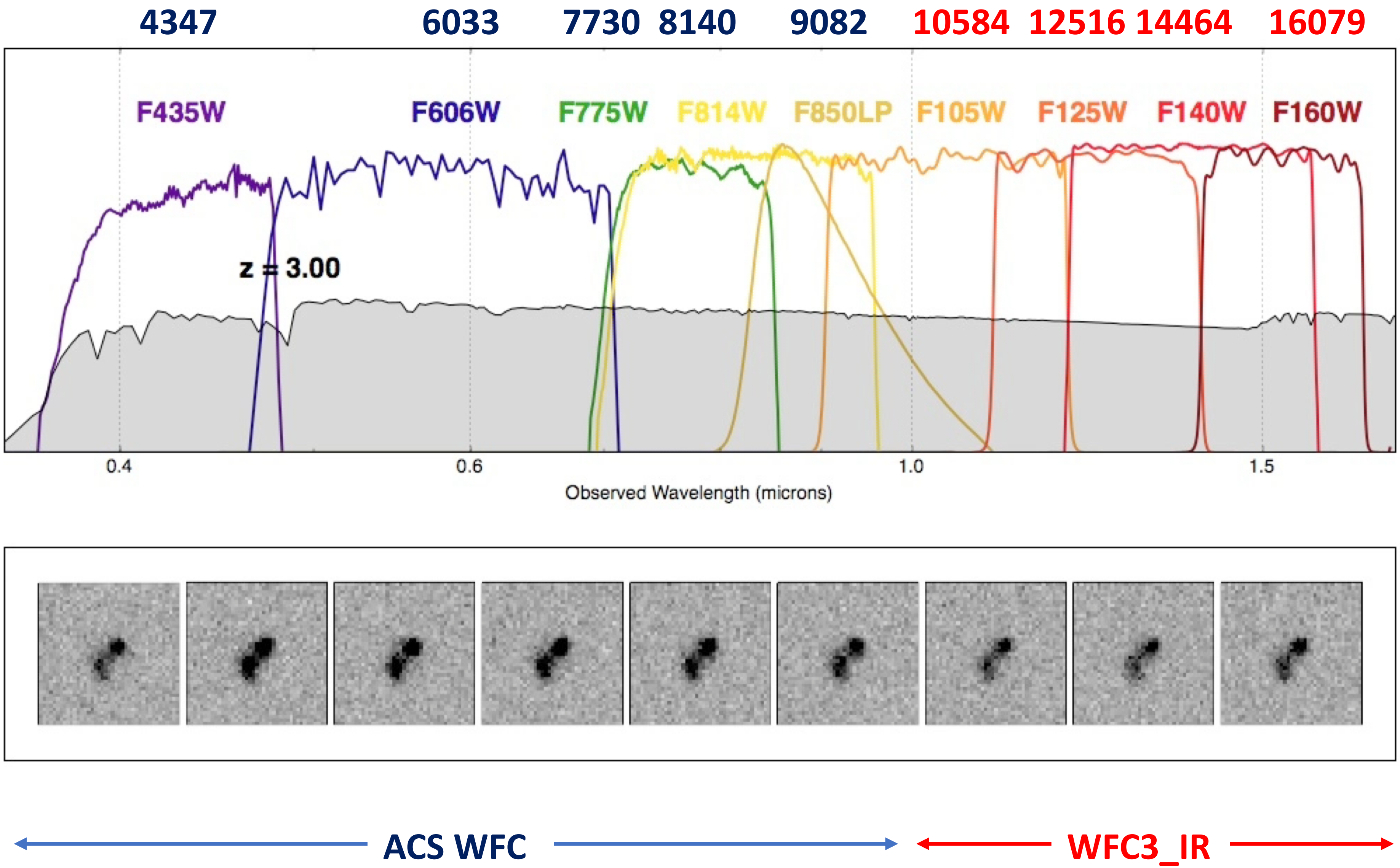
} **IMAGES**

- NICMOS: NIR Camera and Multi Object Spectrograph
- COS: Cosmic Origins Spectrograph
- STIS: Space Telescope Imaging Spectrograph

} SPECTRA

Instrument	Wavelength coverage (micron)	Pixel size (arcsec)	Field of View (arcsec)
ACS WFC	0.4-1.1	0.05 0.3 kpc @z=6	202 x 202 ~1.1 Mpc @z~6
WFC3 IR	0.8-1.7	0.13 0.8 kpc @z=6	136 x 123 ~0.7 Mpc @z~6

Dropout technique

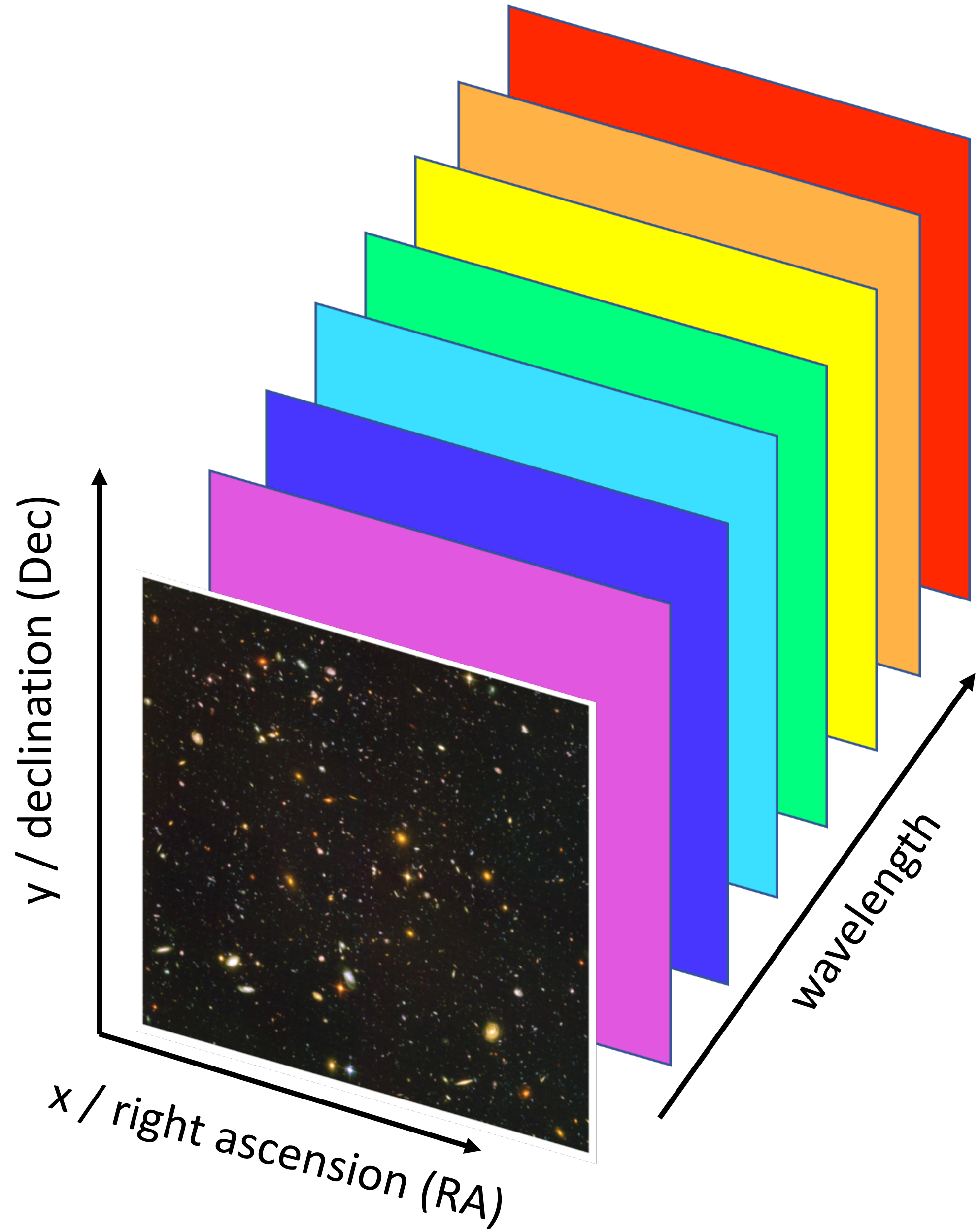


Hands on by Stefano Carniani

Exploration of an HST data cube:

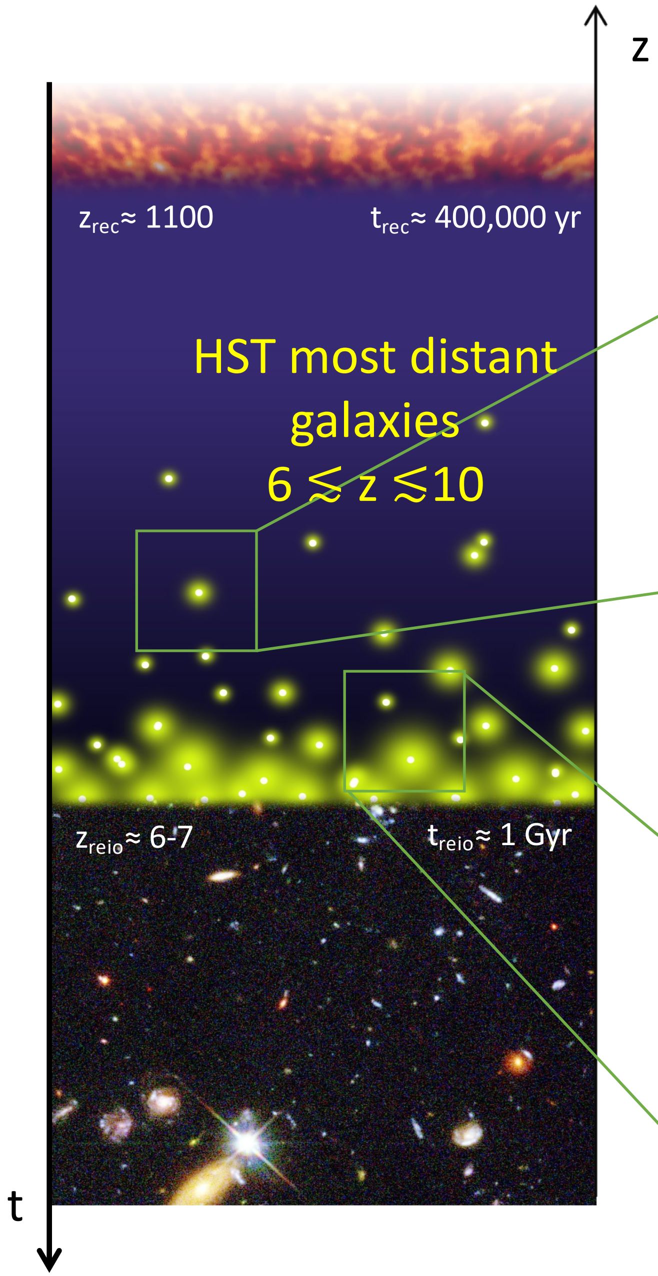
A data cube contains:
two spatial dimensions
(x and y or RA and Dec),
and *one spectral dimension*
(the wavelength).

You will learn how to determine
the redshift of a source
by using the dropout technique



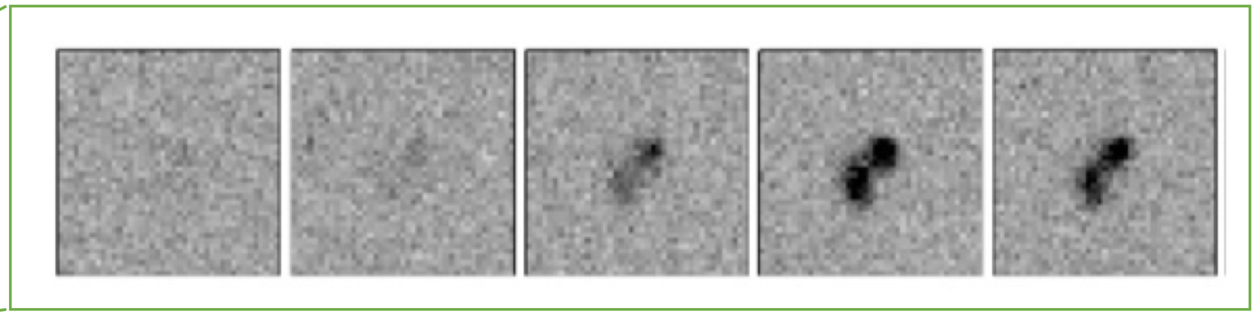
Right Ascension (RA) and Declination (Dec) are
object's celestial coordinates.
RA corresponds to longitude and Dec to latitude.

How do we detect and characterize high-z galaxies?



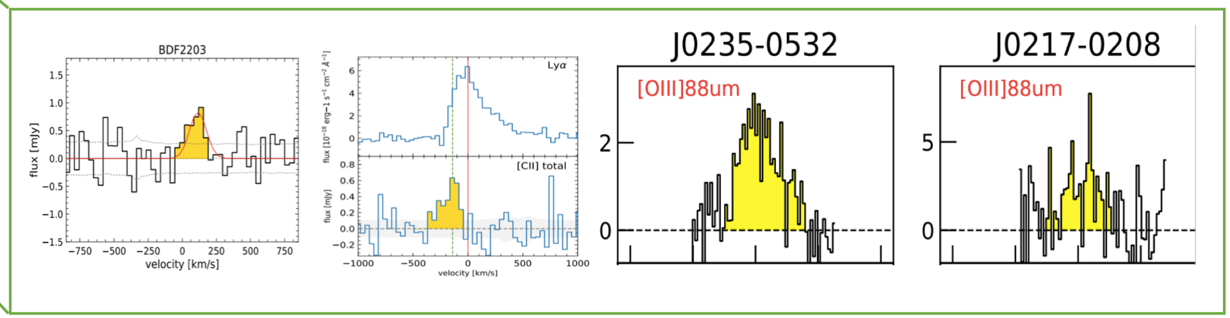
IMAGING

Dropout technique

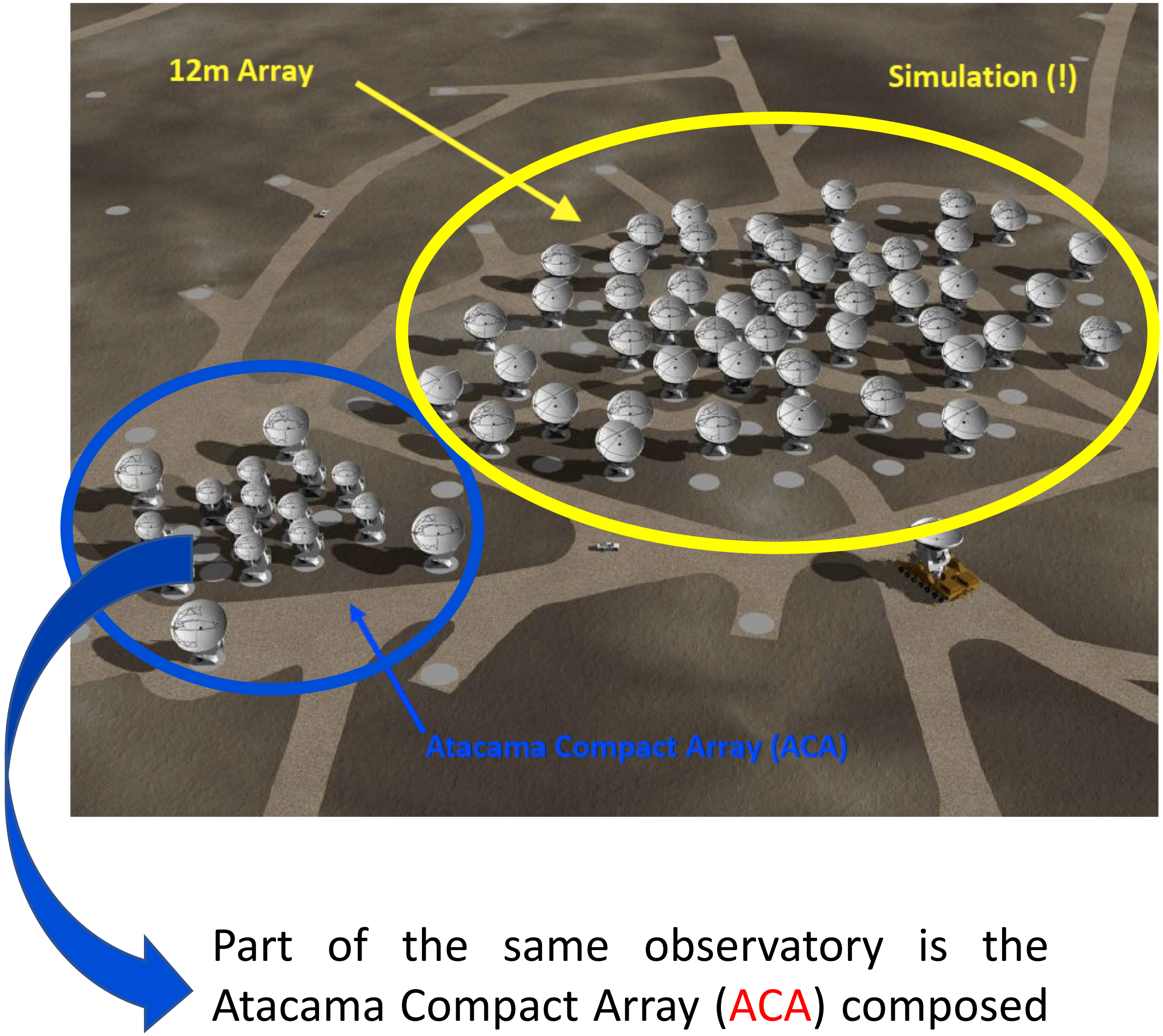


SPECTROSCOPY

Far infrared emission lines



ALMA: Atacama Large Millimeter Array



ALMA is an interferometer composed by an array of 50 antennas of 12 m (diameter) equipped with receivers sensitive to (sub-)mm wavelengths, located in the Atacama desert in Chile at 5000 m above the sea level.

Antennas can be moved over distances from 150 m to 16 km.

Part of the same observatory is the Atacama Compact Array (ACA) composed by 12 antennas, 4 (8) are 12 (7) m large.

Why do we need an array of antennas?

Why do we need to go to the desert at 5000 m a.s.l.?

Why do we need an array of antennas?

Telescope (spatial/angular) resolution (R_θ)

Capability to distinguish two point sources into separate images

The minimum scale of a source that a telescope is able to distinguish

$$R_\theta = 1.22\frac{\lambda}{D}$$

Wavelength of the observations

Diameter of the telescope

The **Rayleigh criterion** (in units od radians)

Comparing telescope resolutions:

HST resolution ($\boldsymbol{\lambda}$=1 **µ**m; D=2.4 m) → R_θ=0.1''

versus

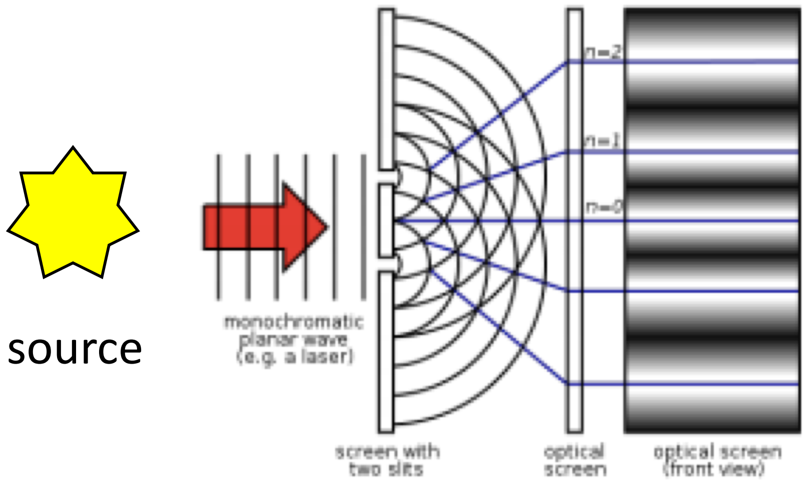
ALMA resolution ($\boldsymbol{\lambda}$=1 mm; D=12 m) → R_θ=21''
(single antenna)

If we want to have a telescope
in the mm range
with resolution similar to HST
we need to build an antenna 1 km large.
This is technically impossible.

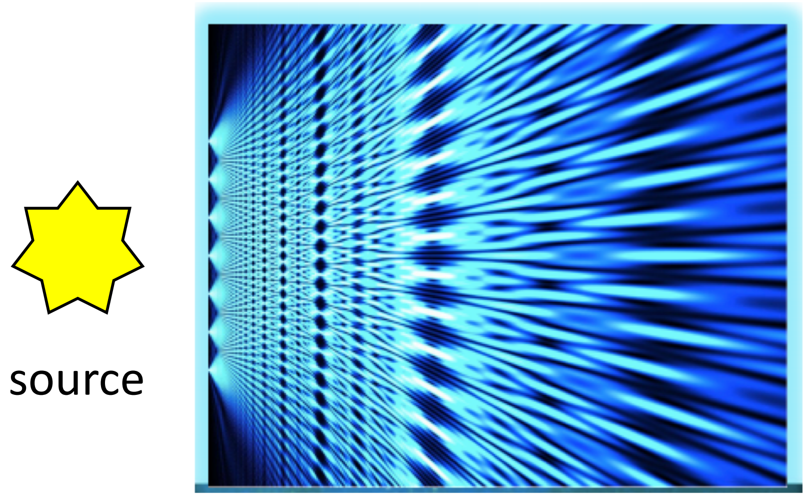
INTERFEROMETRY

The technique of interferometry

Two-slits experiment



Multi-slits experiment

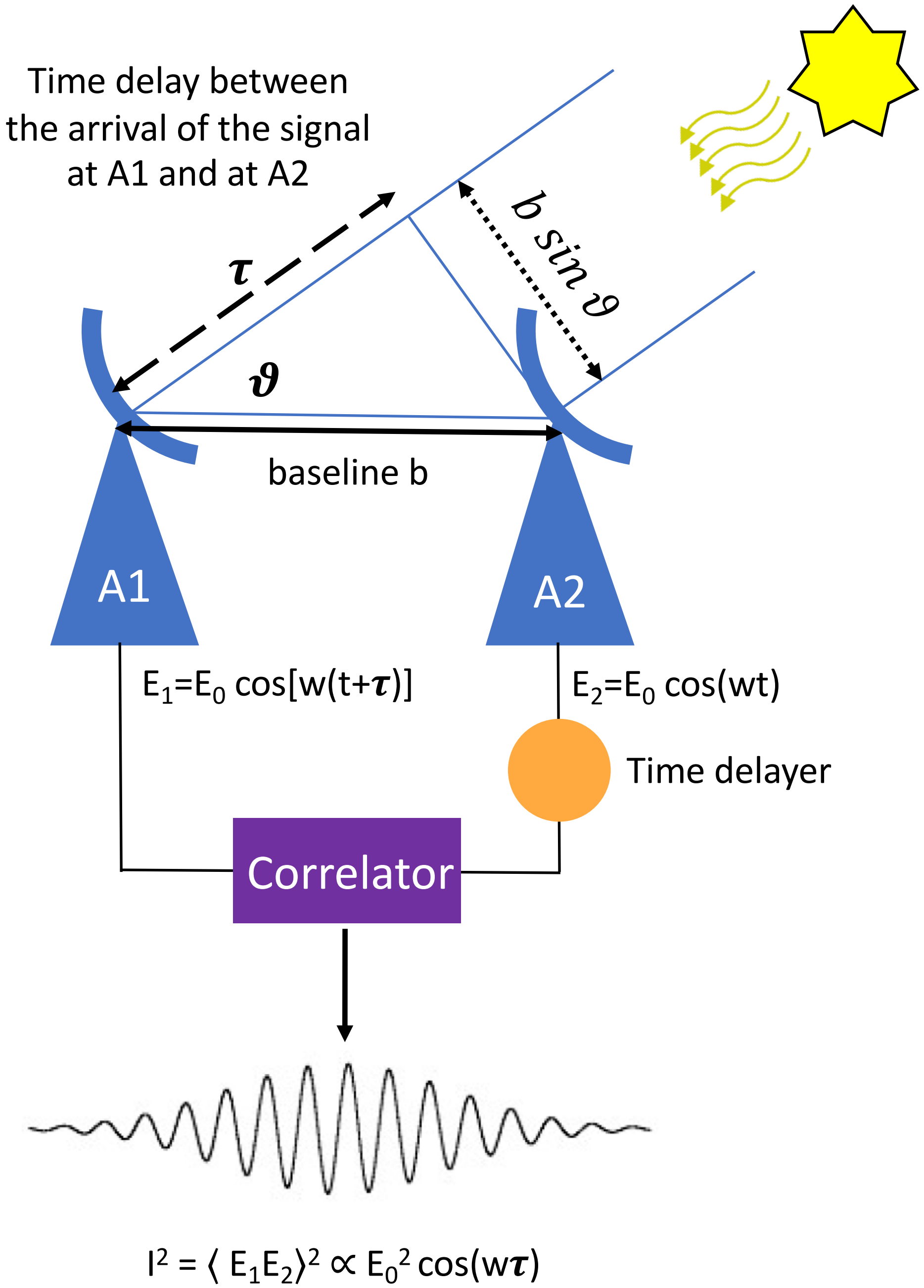


Differently from what happens with «conventional» telescopes, an interferometer does not produce an image of the source directly. It measures instead the interference pattern produced by the source radiation impinging on multiple apertures.

In the case of the Young experiment, the interference pattern is produced by 2 slits (or apertures). In the case of (sub-)mm and radio intereferometer the interference pattern is obtained by correlating the signal arriving on a larger number of antennas (50 in the ALMA case).

As we will see, the image of the source can be reconstructed from the interference pattern. The larger is the number of slits, the more detailed is the measurement of the interference pattern, the higher is the quality of the final image of the source we can get.

The technique of interferometry



At each antenna, the signal arrives at slightly different times, depending on its location in the array.

Dedicated softwares compensate this *geometrical time delay*: $\tau = \frac{b}{c} \cos \vartheta$, where b (called *baseline*) measures the distance between the two antennas and c is the speed of light.

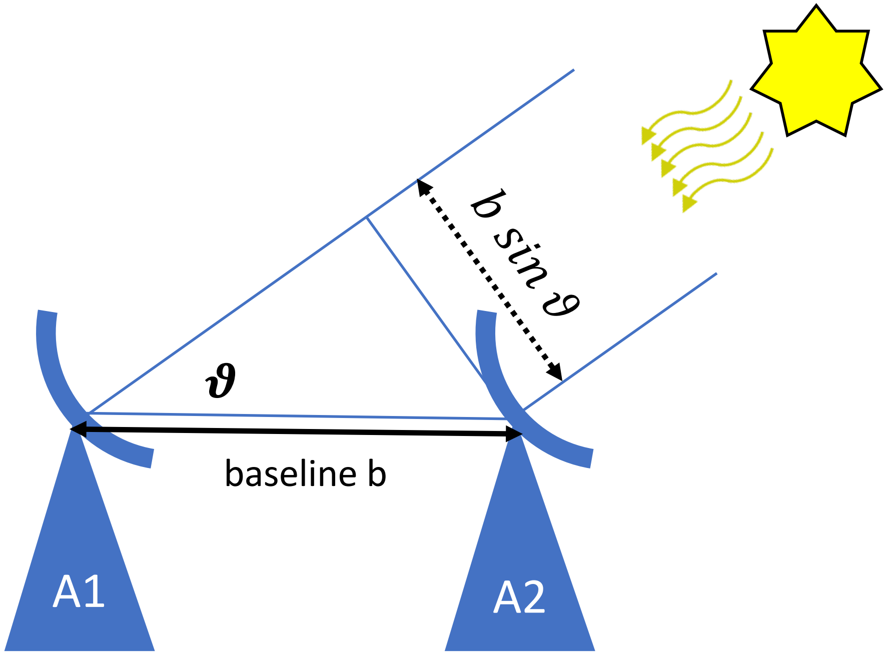
The signal from antenna A1 is then combined with that from antenna A2 in the *correlator*.

To each couple of antennas a *complex visibility** (that depends on the amplitude E_0 and phase $w\tau$) remains associated.

the visibility is defined as:

$$V = \frac{I_{max} - I_{min}}{I_{max} + I_{min}}$$

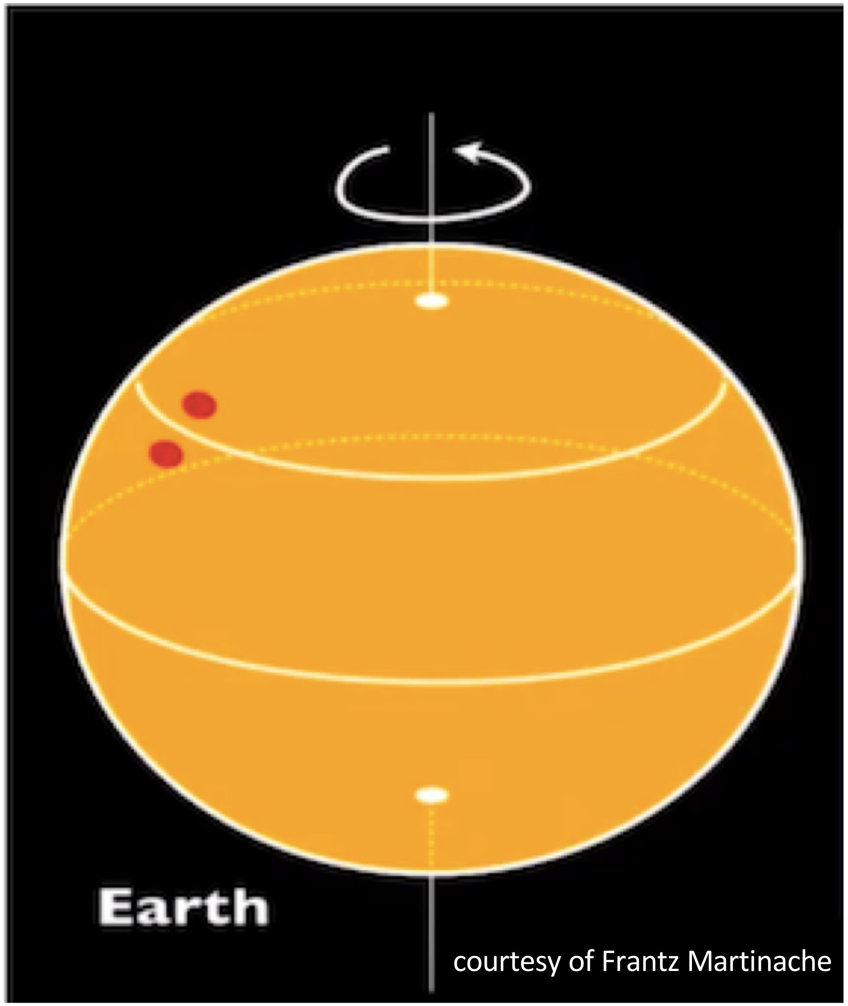
The uv plane



To make the final step through which we can get a source image by means of interferometry, it is necessary to introduce the *uv plane*. To understand the uv plane we have to *move on the source* and imagine how it perceives the geometry of the interferometer.

Let us still consider the simple case of only two antennas. From the point of view of the source the *lenght of the baseline* $(b\ sin\ \vartheta)$ changes as the Earth rotates. It is *maximum* when the source is *at the zenith* $(\vartheta = 90°)$ and minimum as it is setting $(\vartheta = 0°)$.

The uv plane



North-South orientation

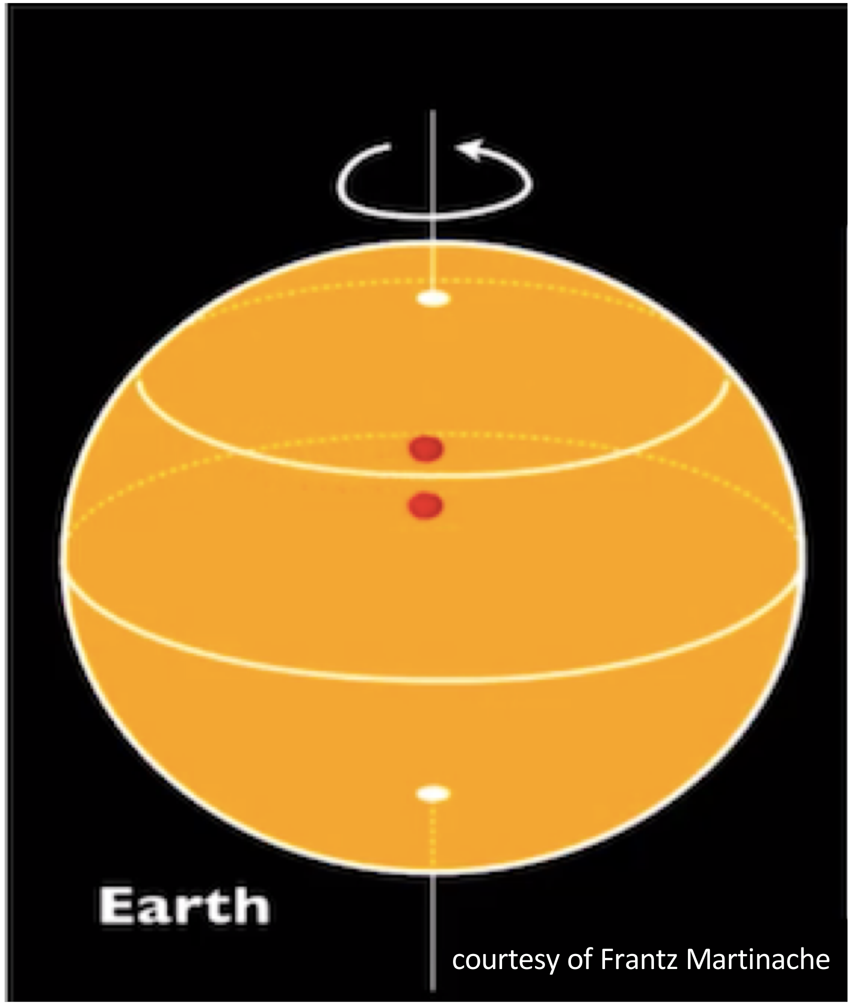
The baseline rotates
during the transit of the source

To make the final step through which we can get a source image by means of interferometry, it is necessary to introduce the *uv plane*. To understand the uv plane we have to *move on the source* and imagine how it perceives the geometry of the interferometer.

Let us still consider the simple case of only two antennas. From the point of view of the source the *lenght of the baseline* $(b\ sin\ \vartheta)$ changes as the Earth rotates. It is *maximum* when the source is *at the zenith* $(\vartheta = 90°)$ and minimum as it is setting $(\vartheta = 0°)$.

The same lenght of the baseline will also depend on the *orientation of the baseline* with respect to the Earth's rotation axis.

The uv plane



North-South orientation

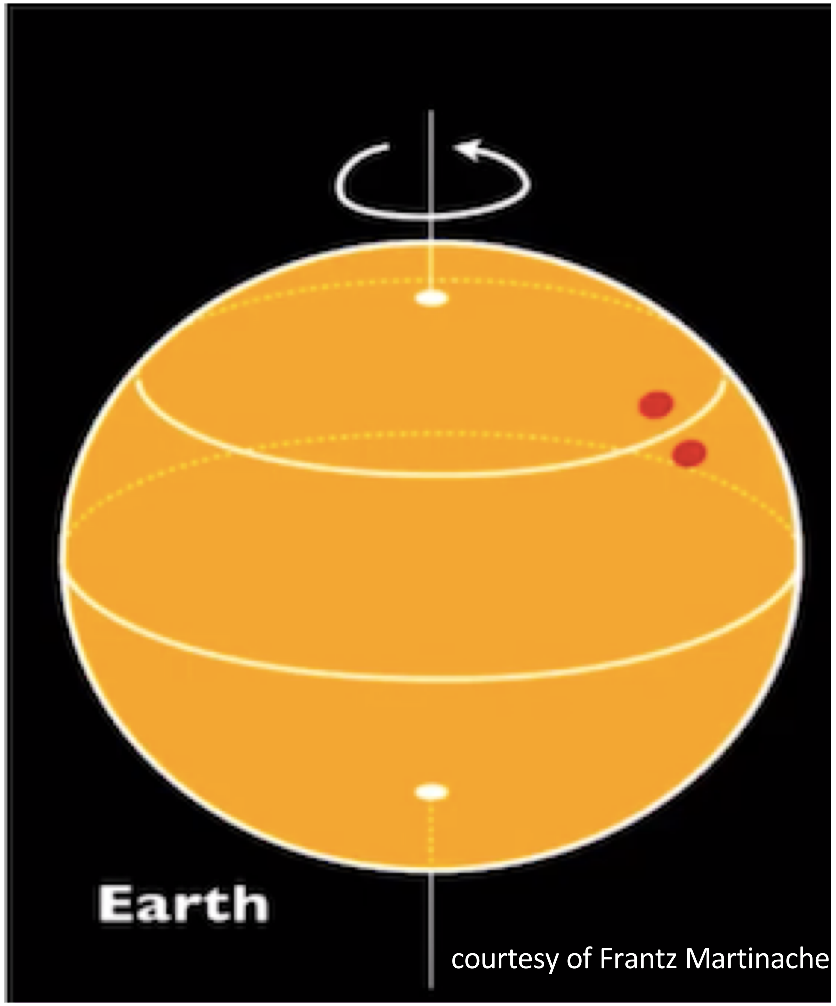
The baseline rotates
during the transit of the source

To make the final step through which we can get a source image by means of interferometry, it is necessary to introduce the *uv plane*. To understand the uv plane we have to *move on the source* and imagine how it perceives the geometry of the interferometer.

Let us still consider the simple case of only two antennas. From the point of view of the source the *lenght of the baseline* $(b \sin \vartheta)$ changes as the Earth rotates. It is *maximum* when the source is *at the zenith* $(\vartheta = 90°)$ and minimum as it is setting $(\vartheta = 0°)$.

The same lenght of the baseline will also depend on the *orientation of the baseline* with respect to the Earth's rotation axis.

The uv plane



North-South orientation

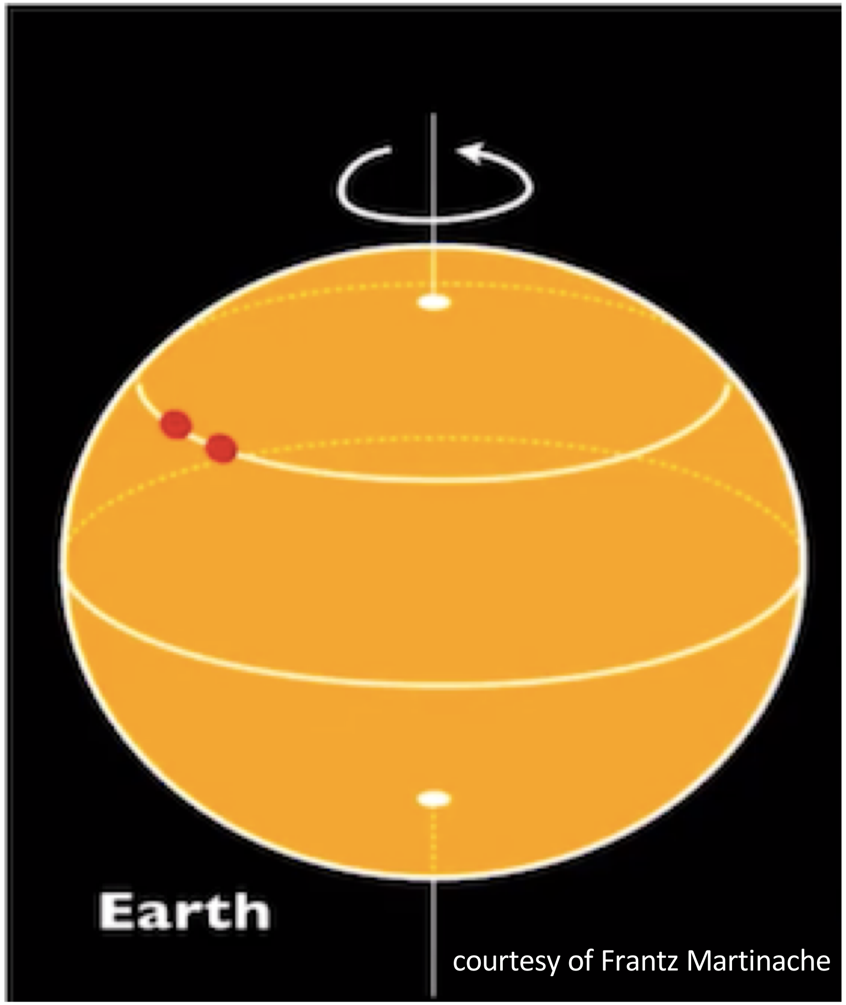
The baseline rotates
during the transit of the source

To make the final step through which we can get a source image by means of interferometry, it is necessary to introduce the *uv plane*. To understand the uv plane we have to *move on the source* and imagine how it perceives the geometry of the interferometer.

Let us still consider the simple case of only two antennas. From the point of view of the source the *lenght of the baseline* $(b\ sin\ \vartheta)$ changes as the Earth rotates. It is *maximum* when the source is *at the zenith* $(\vartheta = 90°)$ and minimum as it is setting $(\vartheta = 0°)$.

The same lenght of the baseline will also depend on the *orientation of the baseline* with respect to the Earth's rotation axis.

The uv plane



East-West orientation

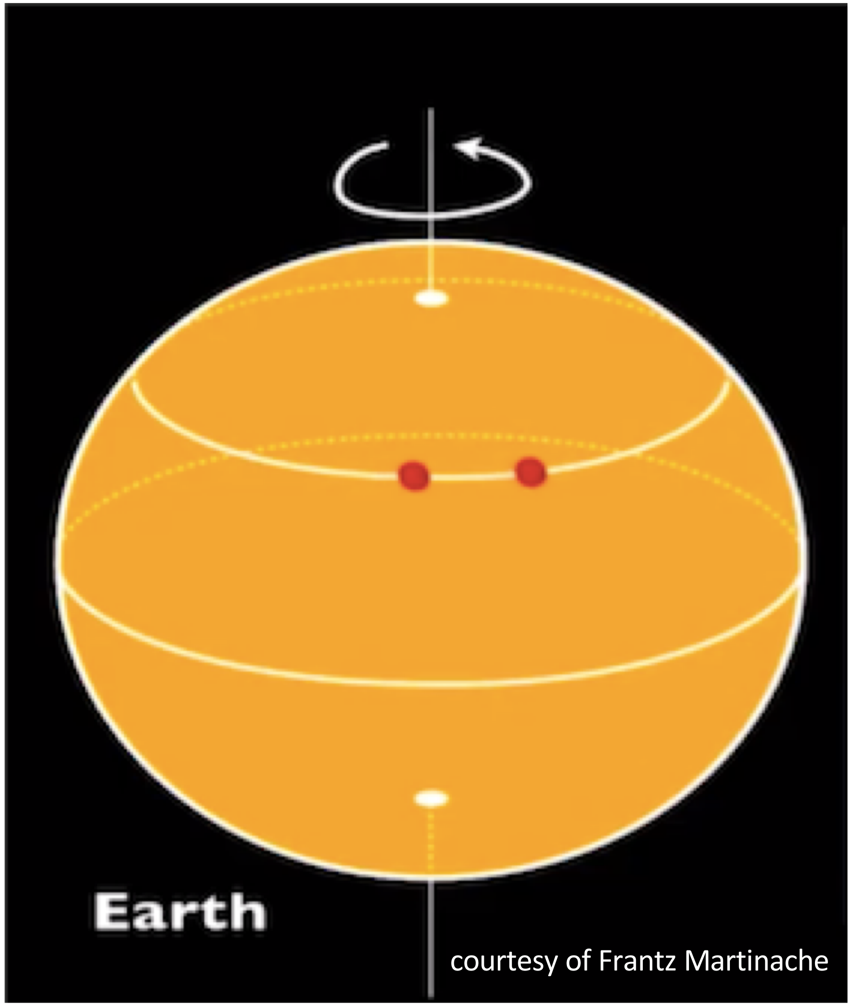
The baseline both rotates AND scales up and down during the transit of the source

To make the final step through which we can get a source image by means of interferometry, it is necessary to introduce the *uv plane*. To understand the uv plane we have to *move on the source* and imagine how it perceives the geometry of the interferometer.

Let us still consider the simple case of only two antennas. From the point of view of the source the *lenght of the baseline* $(b \sin \vartheta)$ changes as the Earth rotates. It is *maximum* when the source is *at the zenith* $(\vartheta = 90°)$ and minimum as it is setting $(\vartheta = 0°)$.

The same lenght of the baseline will also depend on the *orientation of the baseline* with respect to the Earth's rotation axis.

The uv plane



East-West orientation

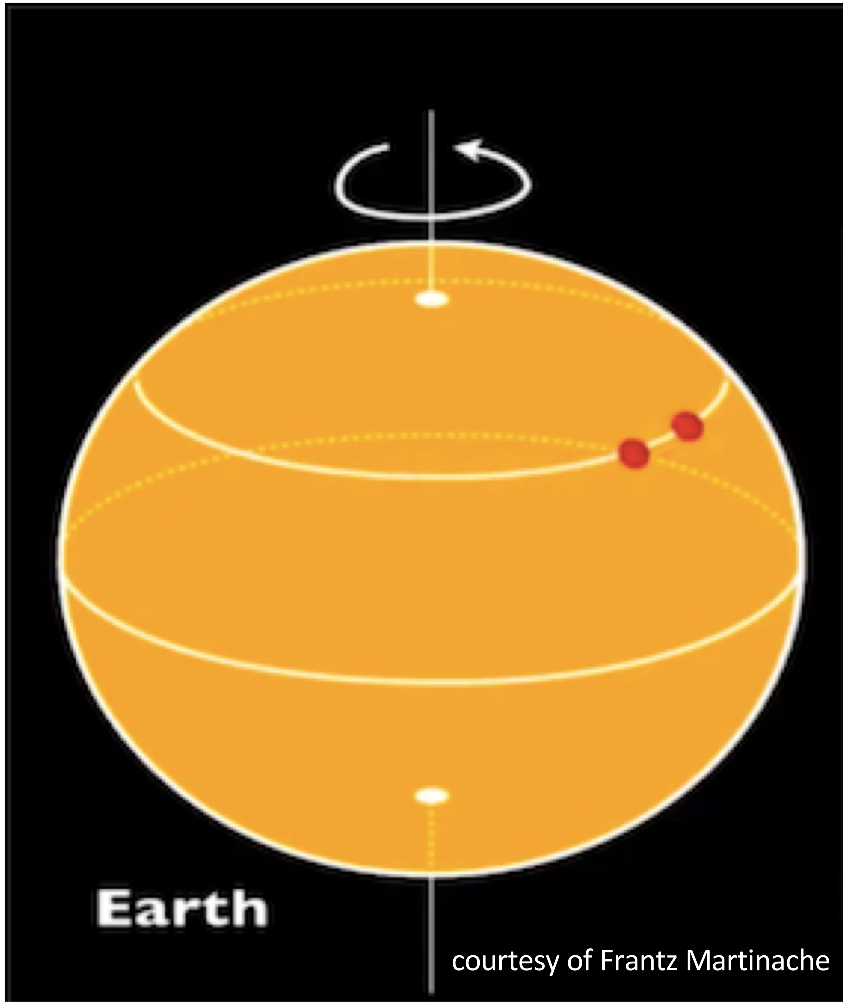
The baseline both rotates AND
scales up and down
during the transit of the source

To make the final step through which we can get a source image by means of interferometry, it is necessary to introduce the *uv plane*. To understand the uv plane we have to *move on the source* and imagine how it perceives the geometry of the interferometer.

Let us still consider the simple case of only two antennas. From the point of view of the source the *lenght of the baseline* $(b \sin \vartheta)$ changes as the Earth rotates. It is *maximum* when the source is *at the zenith* $(\vartheta = 90°)$ and minimum as it is setting $(\vartheta = 0°)$.

The same lenght of the baseline will also depend on the *orientation of the baseline* with respect to the Earth's rotation axis.

The uv plane



East-West orientation

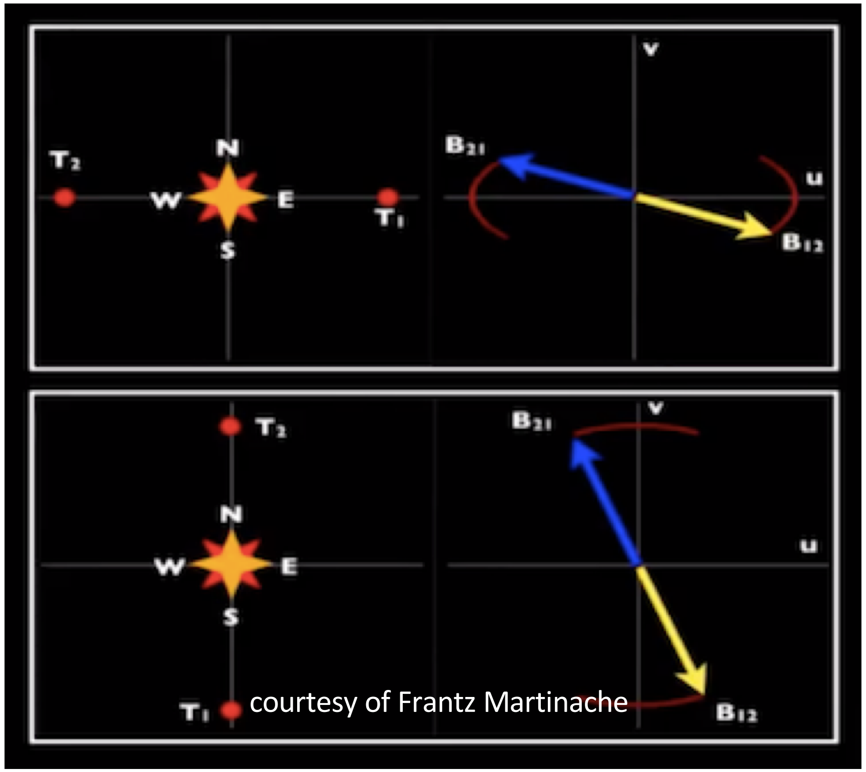
The baseline both rotates AND scales up and down during the transit of the source

To make the final step through which we can get a source image by means of interferometry, it is necessary to introduce the *uv plane*. To understand the uv plane we have to *move on the source* and imagine how it perceives the geometry of the interferometer.

Let us still consider the simple case of only two antennas. From the point of view of the source the *lenght of the baseline* $(b \, sin \, \vartheta)$ changes as the Earth rotates. It is *maximum* when the source is *at the zenith* $(\vartheta = 90°)$ and minimum as it is setting $(\vartheta = 0°)$.

The same lenght of the baseline will also depend on the *orientation of the baseline* with respect to the Earth's rotation axis.

The uv plane



The uv plane is a system of coordinates

$$\frac{\vec{b}}{\lambda} \equiv (\vec{u}, \vec{v})$$

that represents the lenght of a baseline in units of wavelength,

where $\vec{u}$ is the East-West component of the baseline
and $\vec{v}$ is the North-South component of the baseline.

The *interference pattern* is finally related to the *source brightness* through the *«van Cittert-Zernike» theorem*

$$V(u, v) = \int\int dx\, dy\, B(x, y) e^{2\pi i(ux+vy)}$$

the *complex visibility* V(u,v) in the *uv plane* is the 2D Fourier transform of the *source brightness* B(x,y) in the *sky plane*.

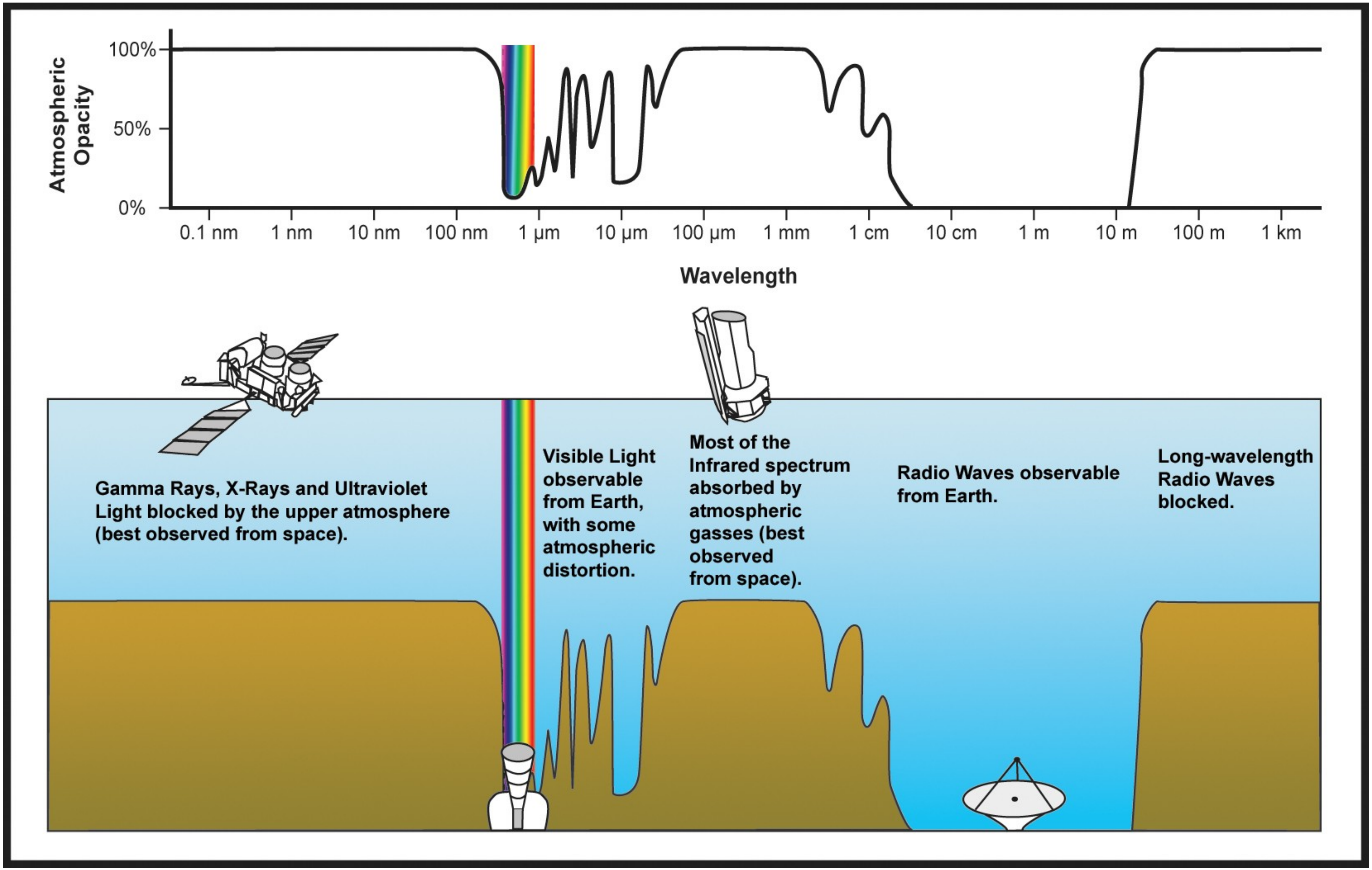
Why do we need to go to the desert?

If we want to observe at long wavelengths

Why do we need to go to space?

If we want to observe at short wavelengths

Atmospheric windows and absorption bands



Main gas absorbers in the atmosphere

- <u>Oxygen and Ozone</u> (O_2,O_3)

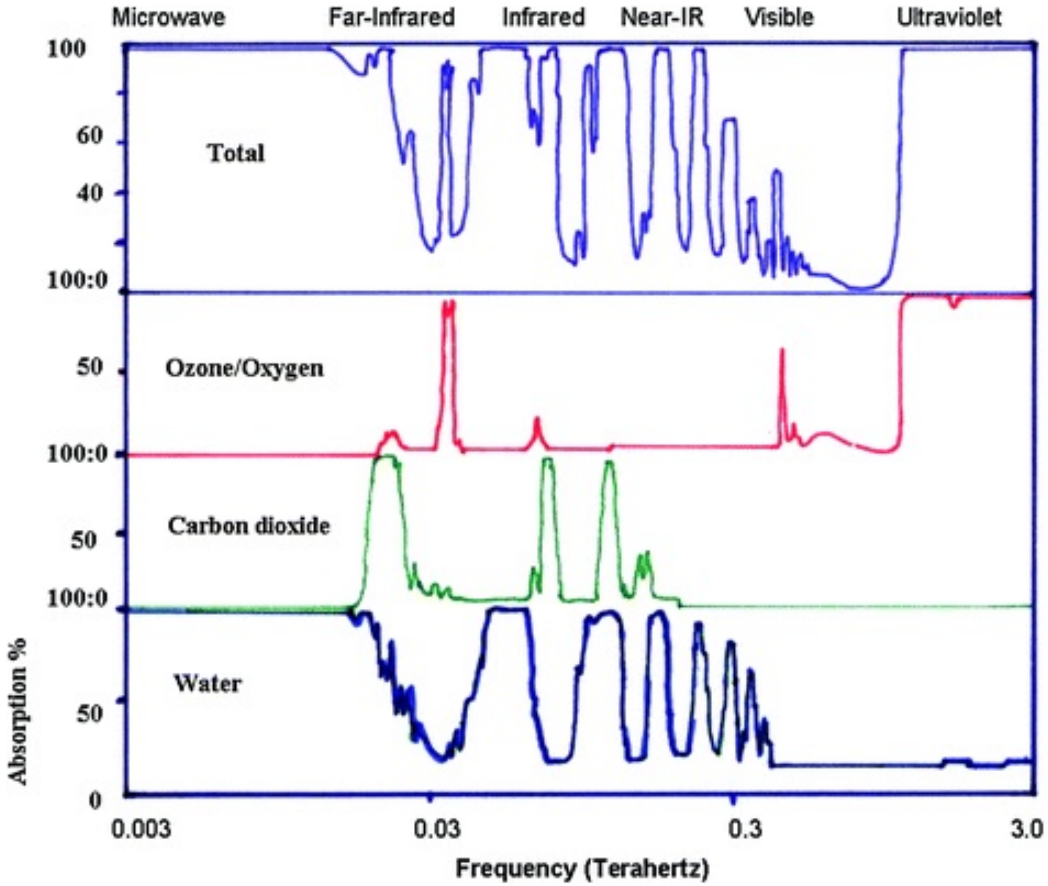
  from UV to higher frequencies

- <u>Carbon dioxide</u> (CO_2)

  visible and Far IR

- <u>Water vapor</u> (H_2O)

  from (sub-)mm to lower frequencies



Goody & Robinson (1951)

Main gas absorbers in the atmosphere

- Oxygen and Ozone (O_2,O_3)

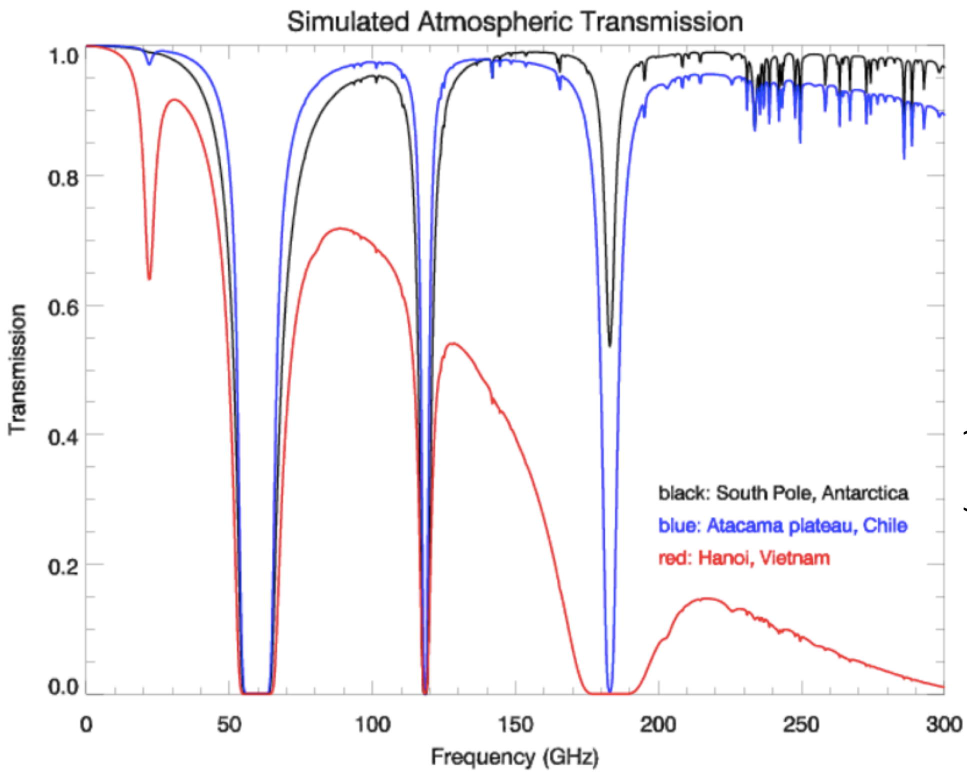
  from UV to higher frequencies

- Carbon dioxide (CO_2)

  visible and Far IR

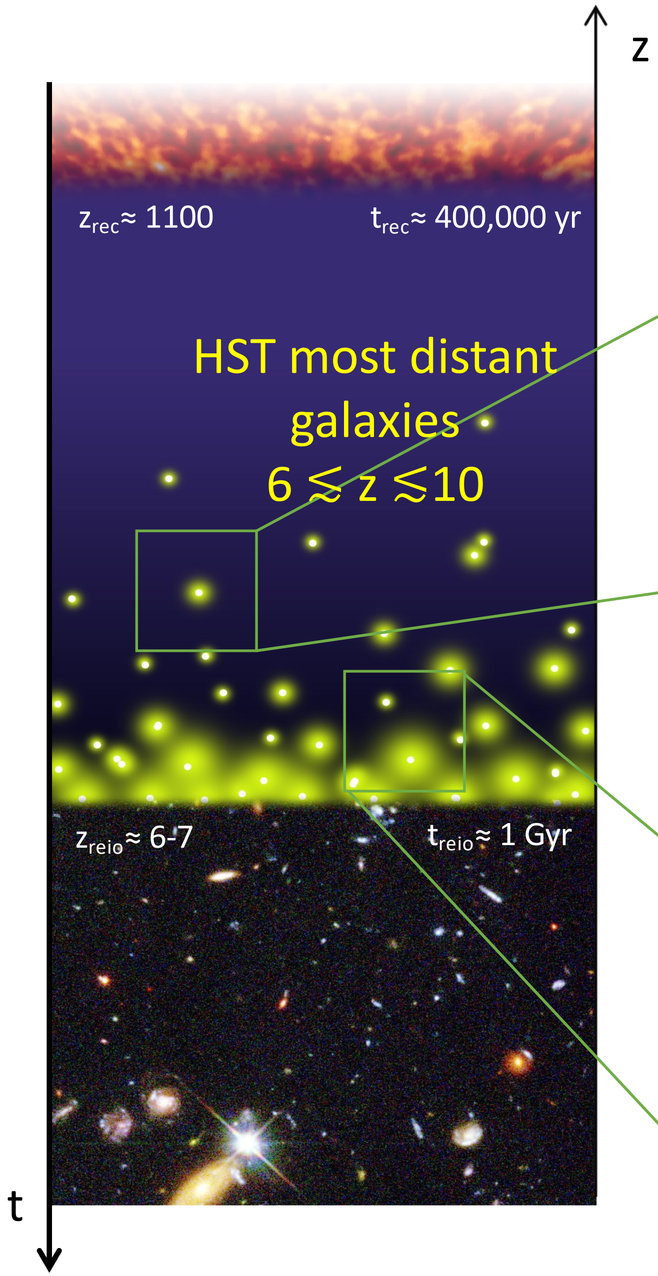
- Water vapor (H_2O)

  from (sub-)mm to lower frequencies



Kovac & Barkatz (2007)

How do we detect and characterize high-z galaxies?

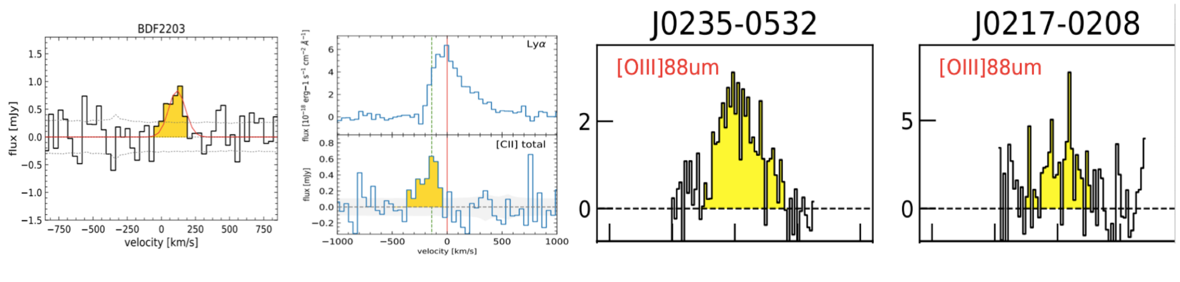


IMAGING

Dropout technique

SPECTROSCOPY

Far infrared emission lines

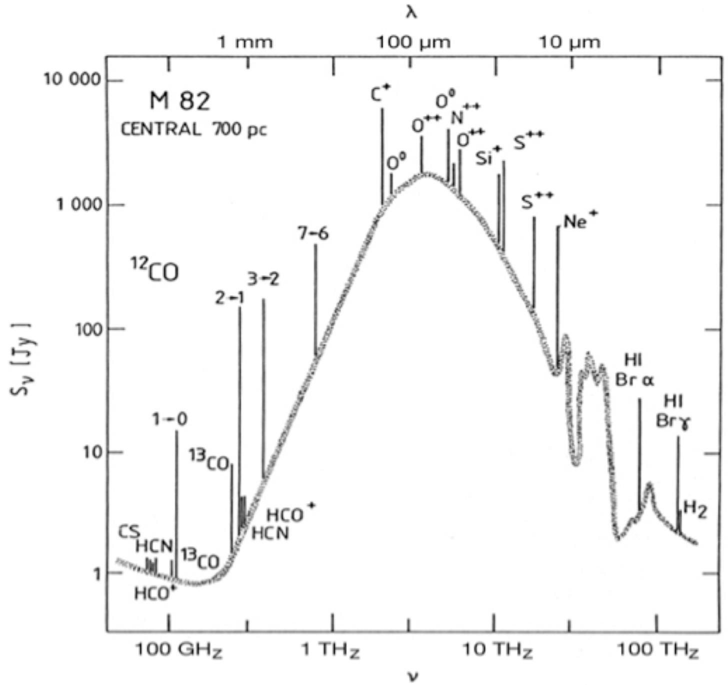


Properties of galaxies from their spectra

A galaxy spectrum contains a wealth of information on the fundamental physical processes occurring within the galaxy.

A single spectrum alone can tell us about:

- Galaxy metallicity (Z/Z_{sun}),
- Amount of dust,
- Gas density (n_{gas}),
- Age of the stellar population,
- Rate of star formation (SFR)
- Presence of AGN (active galactic nuclei)

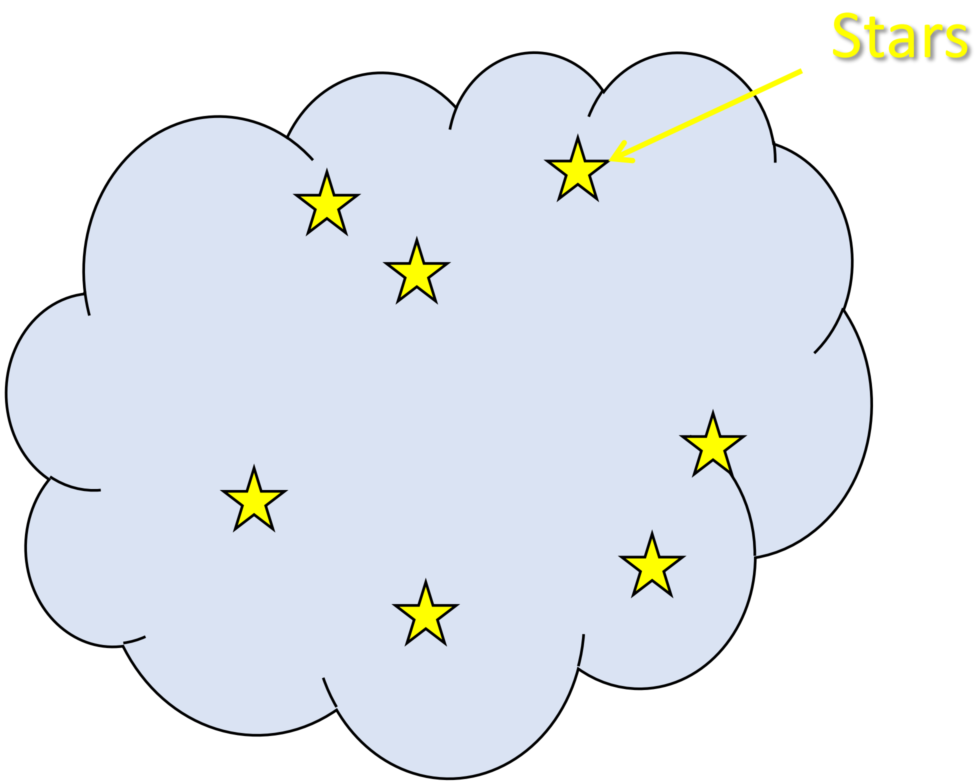


Emission lines in galaxy spectra
are extremely powerful tools for inferring this information.
Thery arise from the
interstellar medium of galaxies

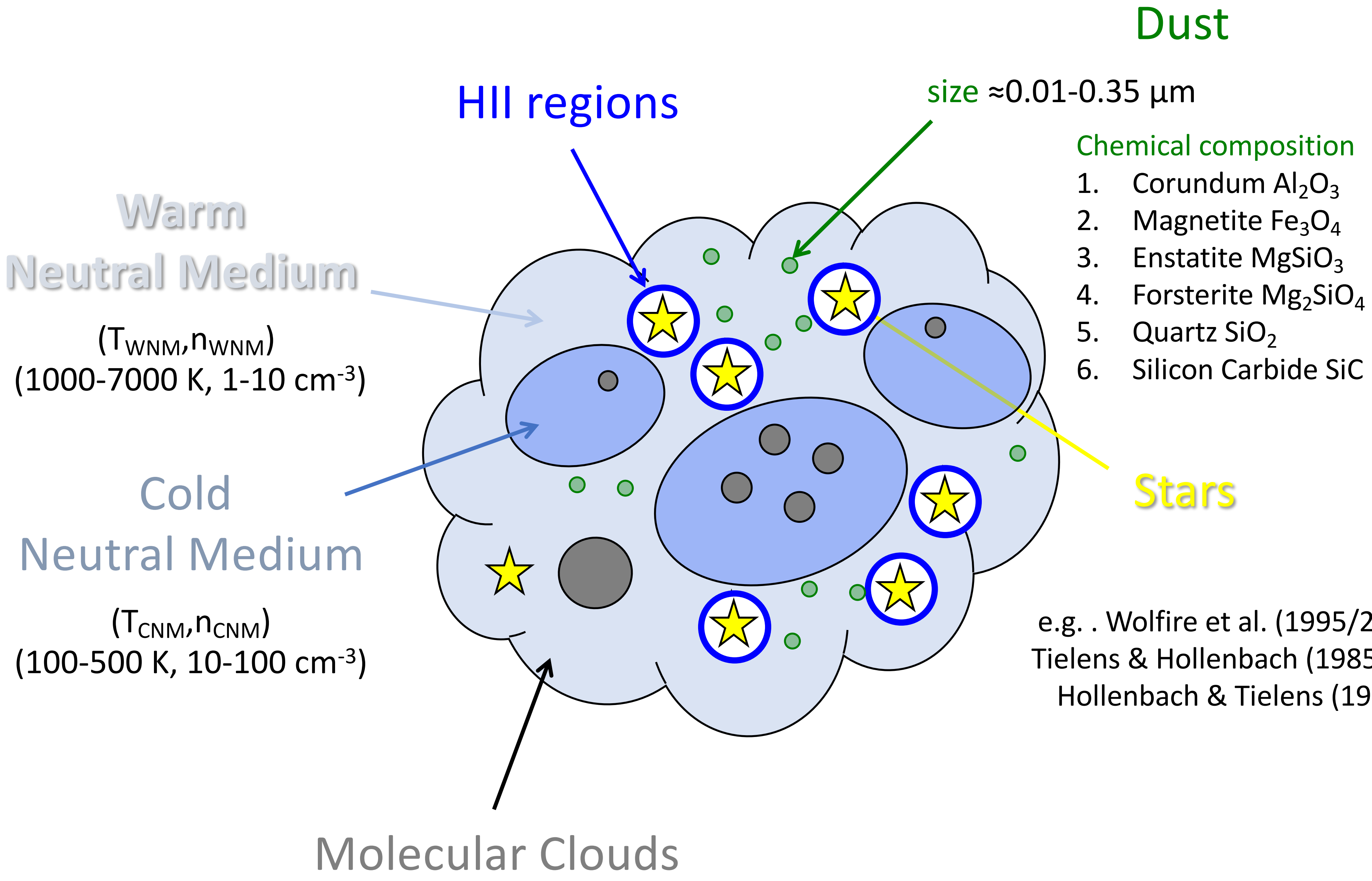
The interstellar medium (ISM)

With the term
«**interstellar medium**»
(as it is clear from the word itself)
we *simply* refer
to the space in galaxies
that is among stars.

However the ISM
is *everything but simple*!

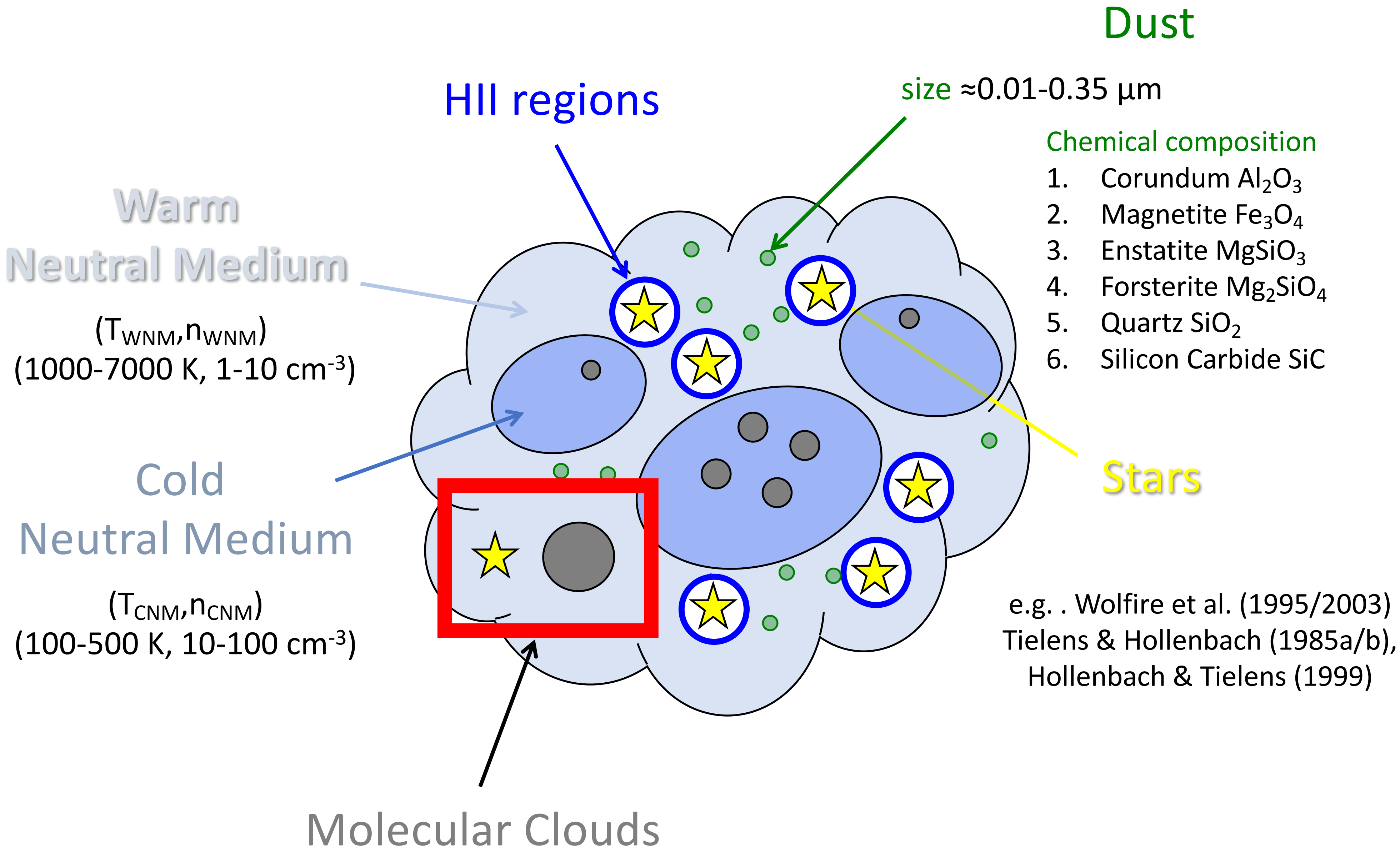


The interstellar medium (ISM): a very complex environment



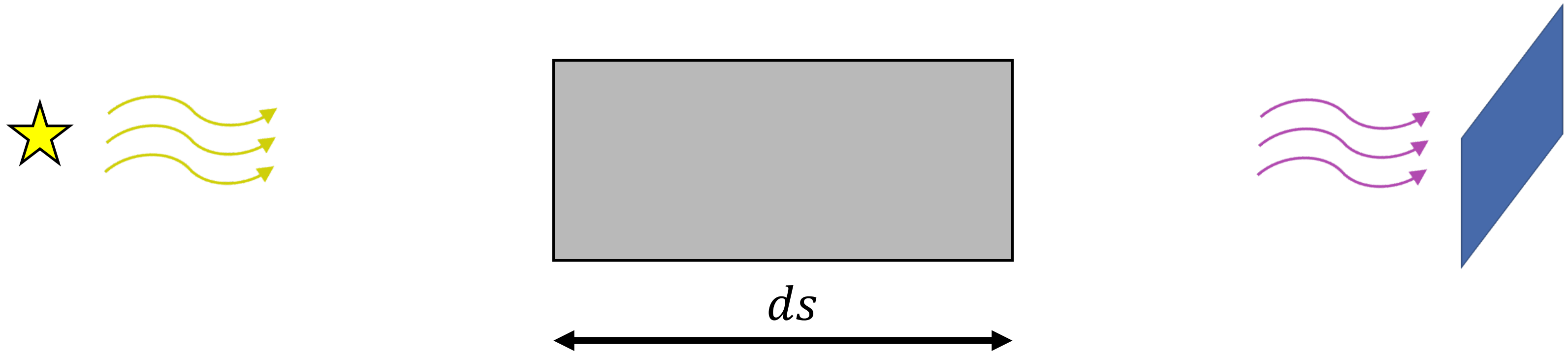
e.g. . Wolfire et al. (1995/2003)
Tielens & Hollenbach (1985a/b),
Hollenbach & Tielens (1999)

The interstellar medium (ISM): a very complex environment



The radiative transfer equation

1° case: ABSORPTION ONLY



I_v^0

Intrinsic
specific intensity

$$dp = k_v ds = d\tau$$

Absorption probability
(k_v = absorption coefficient,
τ= optical depth)

$$I_v = I_v^0 \exp(-\tau)$$

Observed
specific intensity

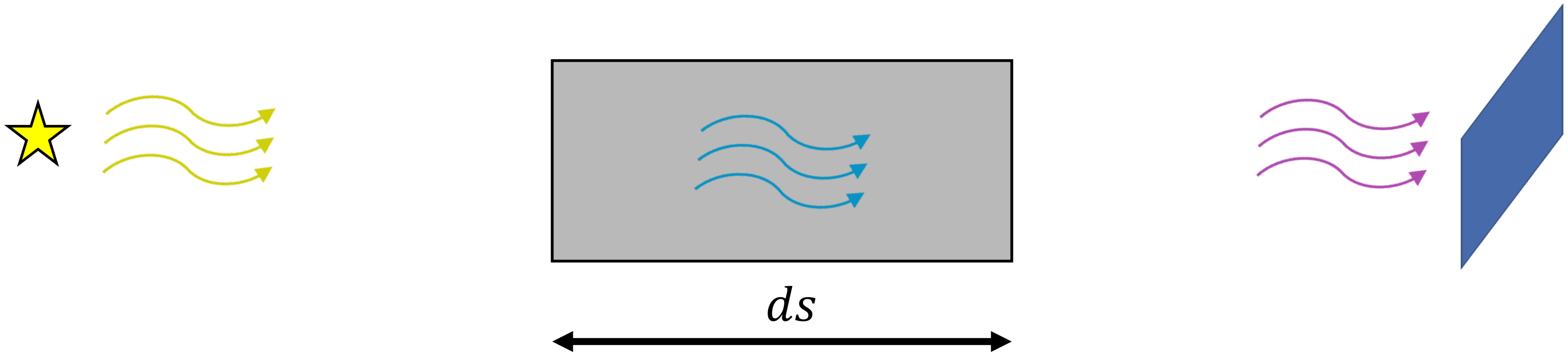
RT
solution

$$\frac{dI_v}{I_v} = -d\tau$$

Intensity fraction that will be lost
because of absorption

The radiative transfer equation

2° case: ABSORPTION+EMISSION



I_υ^0

$$I_\upsilon = [1 - \exp(-\tau)]S_\upsilon \; + \; I_\upsilon^0 \exp(-\tau)$$

$$dI_\upsilon = \epsilon_\upsilon ds = d\tau$$

$$S_\upsilon = \frac{\epsilon_\nu}{k_\upsilon}$$

Absorption only

$$\frac{dI_\upsilon}{ds} = -\, I_\upsilon\, k_\upsilon$$

specific intensity
emitted by the cloud
(ϵ_υ = emission coefficient)

RT
solution

Emission only

$$\frac{dI_\upsilon}{ds} = \epsilon_\upsilon$$

ABSORPTION + EMISSION

$$\frac{dI_\upsilon}{ds} = -\, I_\upsilon\, k_\upsilon + \epsilon_\upsilon$$

CLOUDY

simulations of astronomical clouds and their spectra

Cloudy is a code largely used in Astrophysics that *solves the radiative transfer equation* and *calculates the ionization structure of the cloud*.

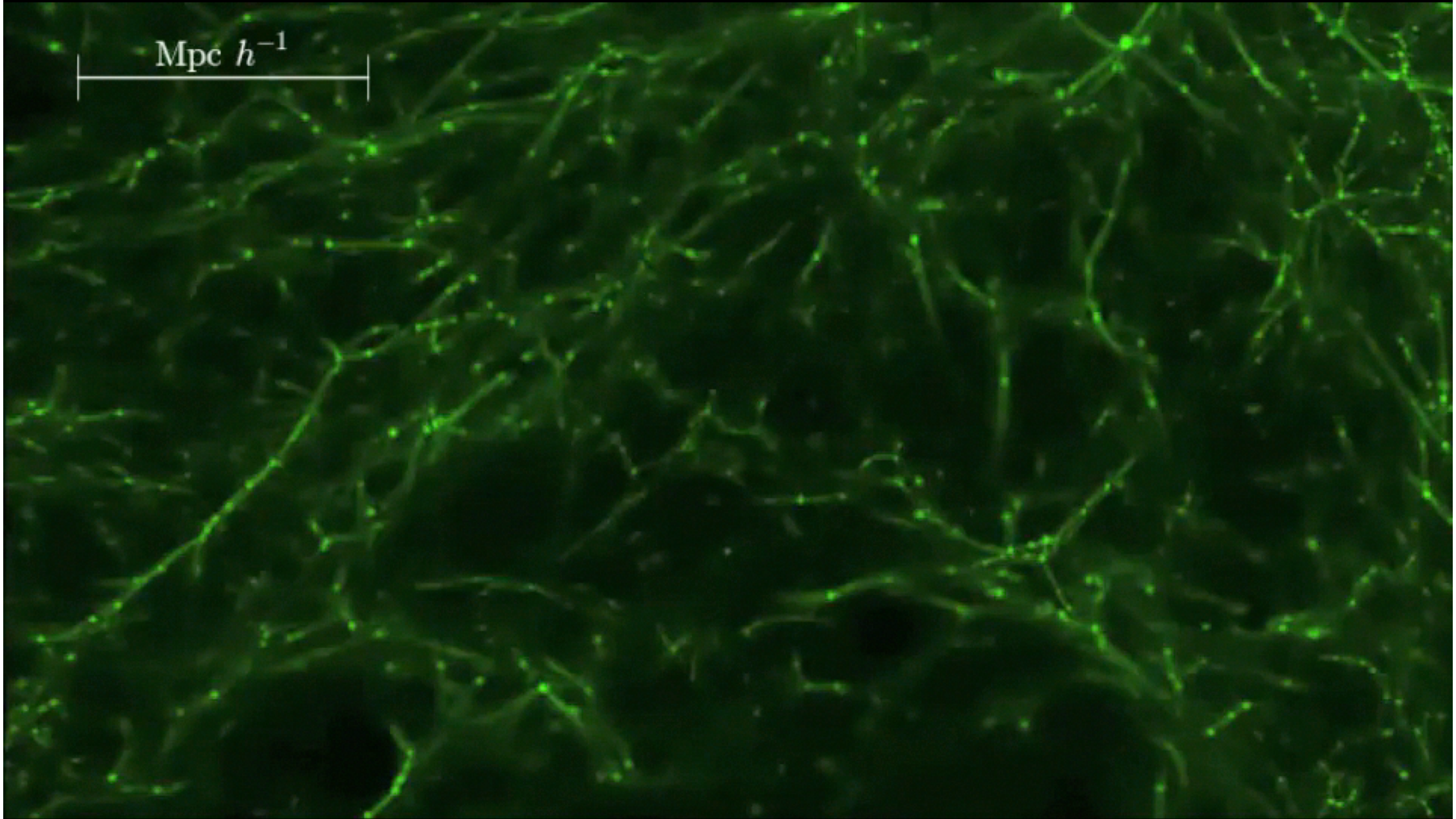
The output is a prediction of the *gas physical conditions* and the *resulting observed spectrum*.

These predictions depend on the *cloud properties*

(gas number density n_{gas} [cm^{-3}], column density $N_{gas} = n_{gas} \times ds$ [cm^{-2}], metallicity Z/Z_{sun})

and on the *intensity of the radiation field.*

NUMERICAL SIMULATIONS OF THE HIGH REDSHIFT UNIVERSE



Courtesy of Andrea Pallottini

SUMMARY

- Stefano Carniani: Exploration of 3D spectroscopy data

1. HST data cube → Dropout technique
2. ALMA data cube → emission lines arising from the ISM

- Graziano Ucci: Machine learning on Galaxy spectra

1. emission lines arising from the ISM
2. library of CLOUDY models

- Andrea Pallottini: Synthetic spectra of galaxies in cosmological simulations

1. Cosmological simulation
2. Mock emission lines from simulated galaxies