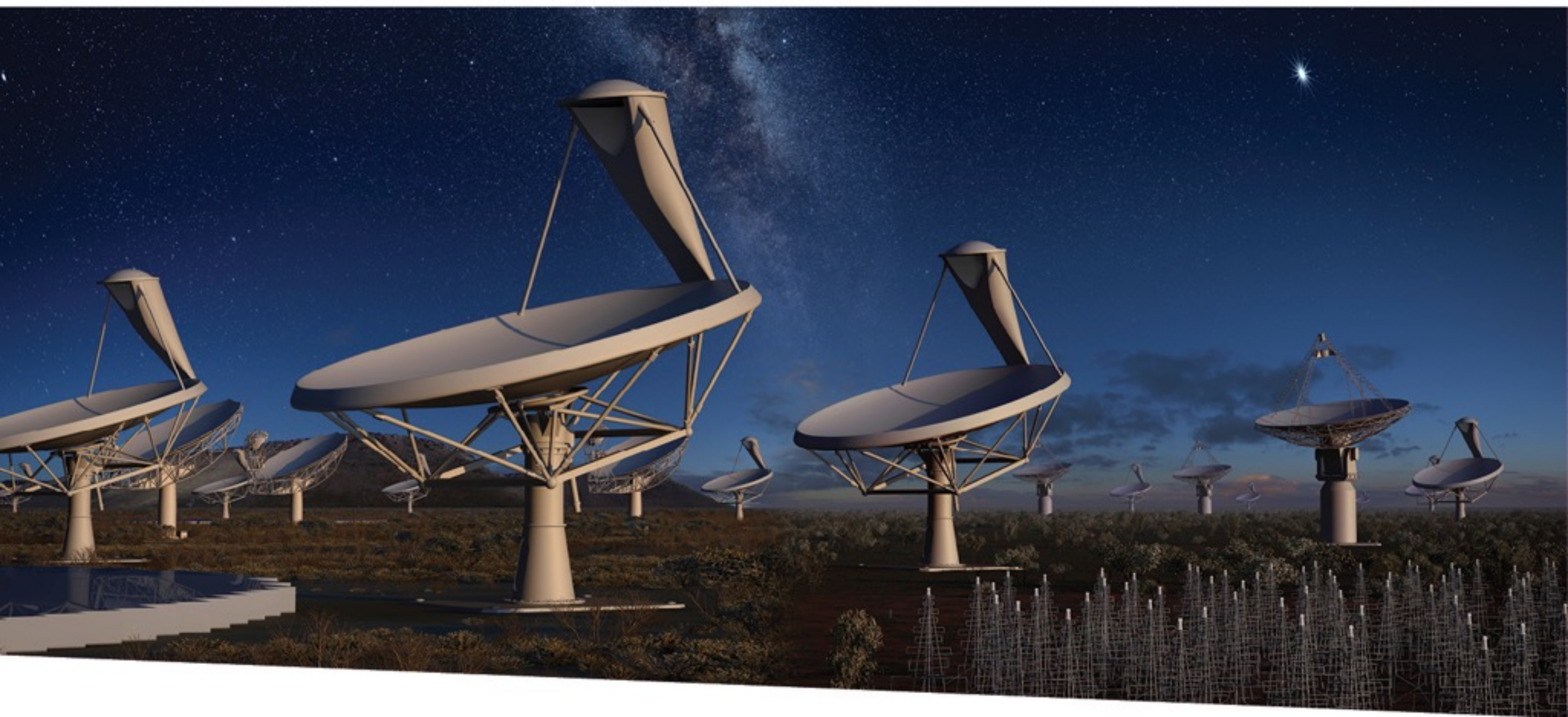
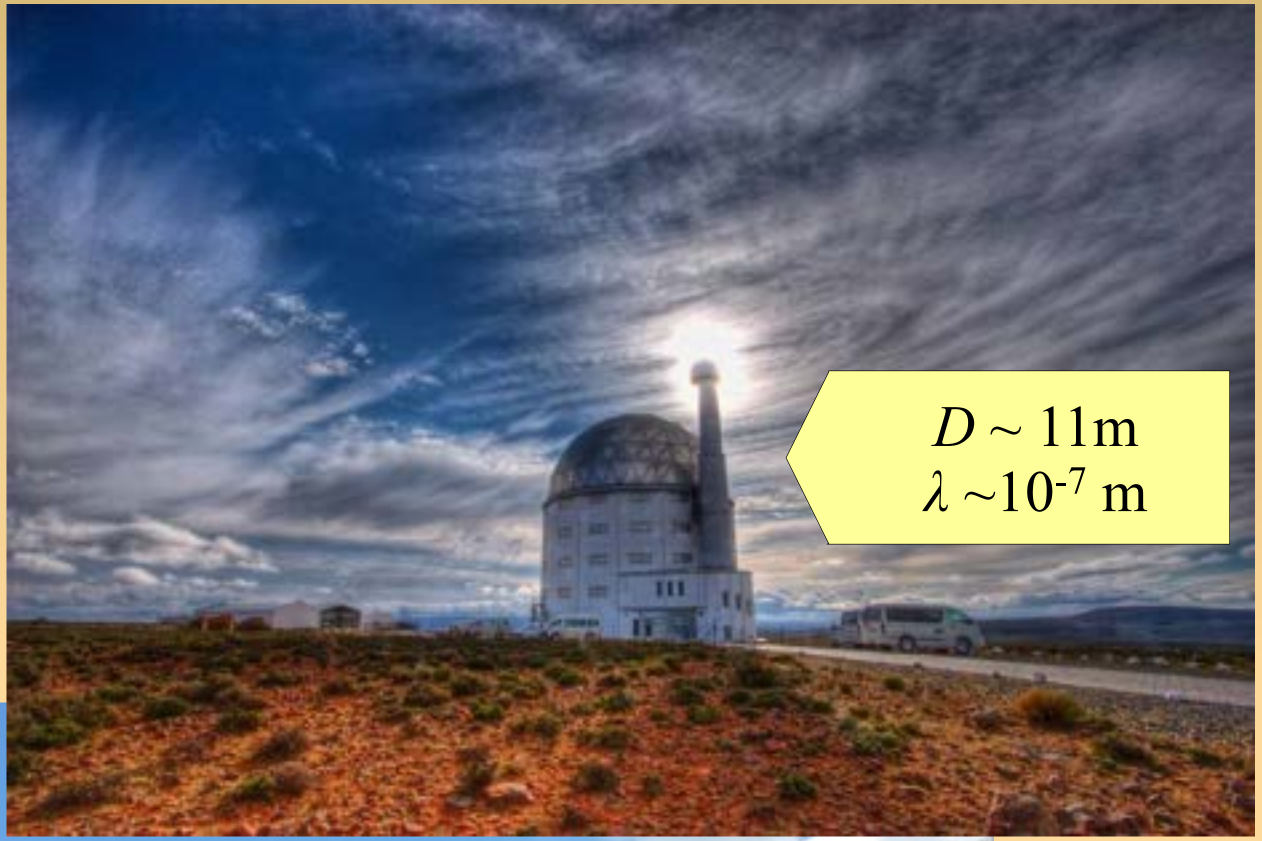
Facing current and future calibration/imaging challenges of the Square Kilometre Array (SKA)

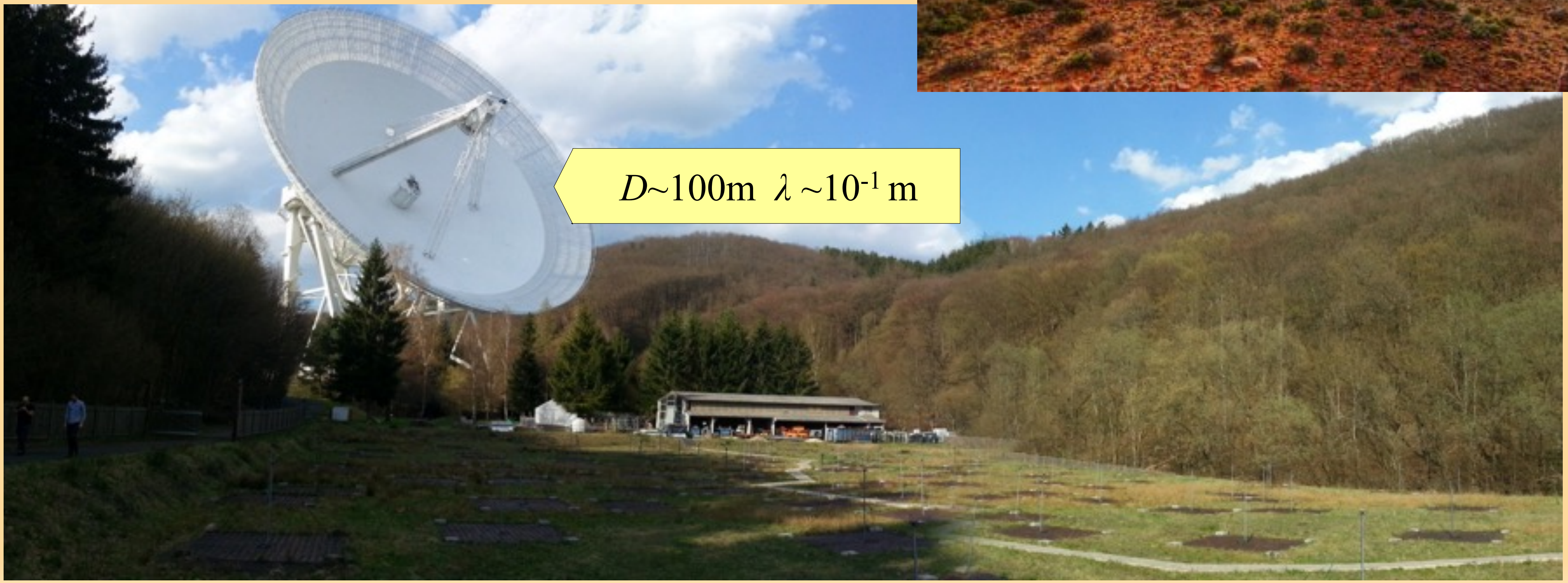
Most slides are stolen from O. Smirnov with his approval



Why Interferometry?

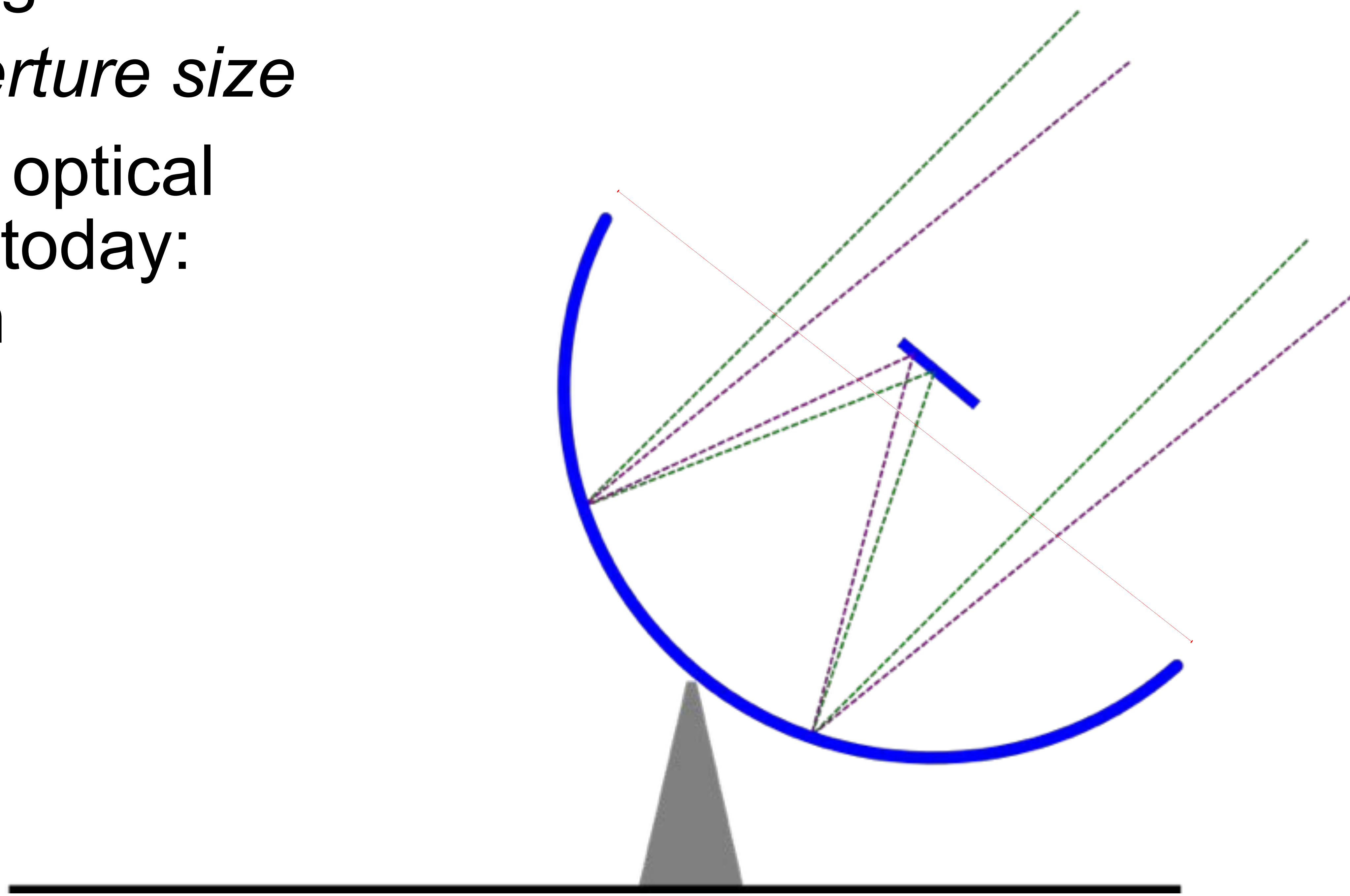
- Aperture size ***D*** is crucial:
  - Resolution $\sim \lambda / D$
  - Sensitivity $\sim D^2$
- Very poor resolution in radio due to long wavelengths
- SALT resolution: ~1000 km radio dish?





A Typical Telescope

- A digital camera with delusions of grandeur
- Lens = mirror (metal reflects radio), focuses incoming radiation on a detector
- *D = aperture size*
- Biggest optical scopes today: *D*~10 m

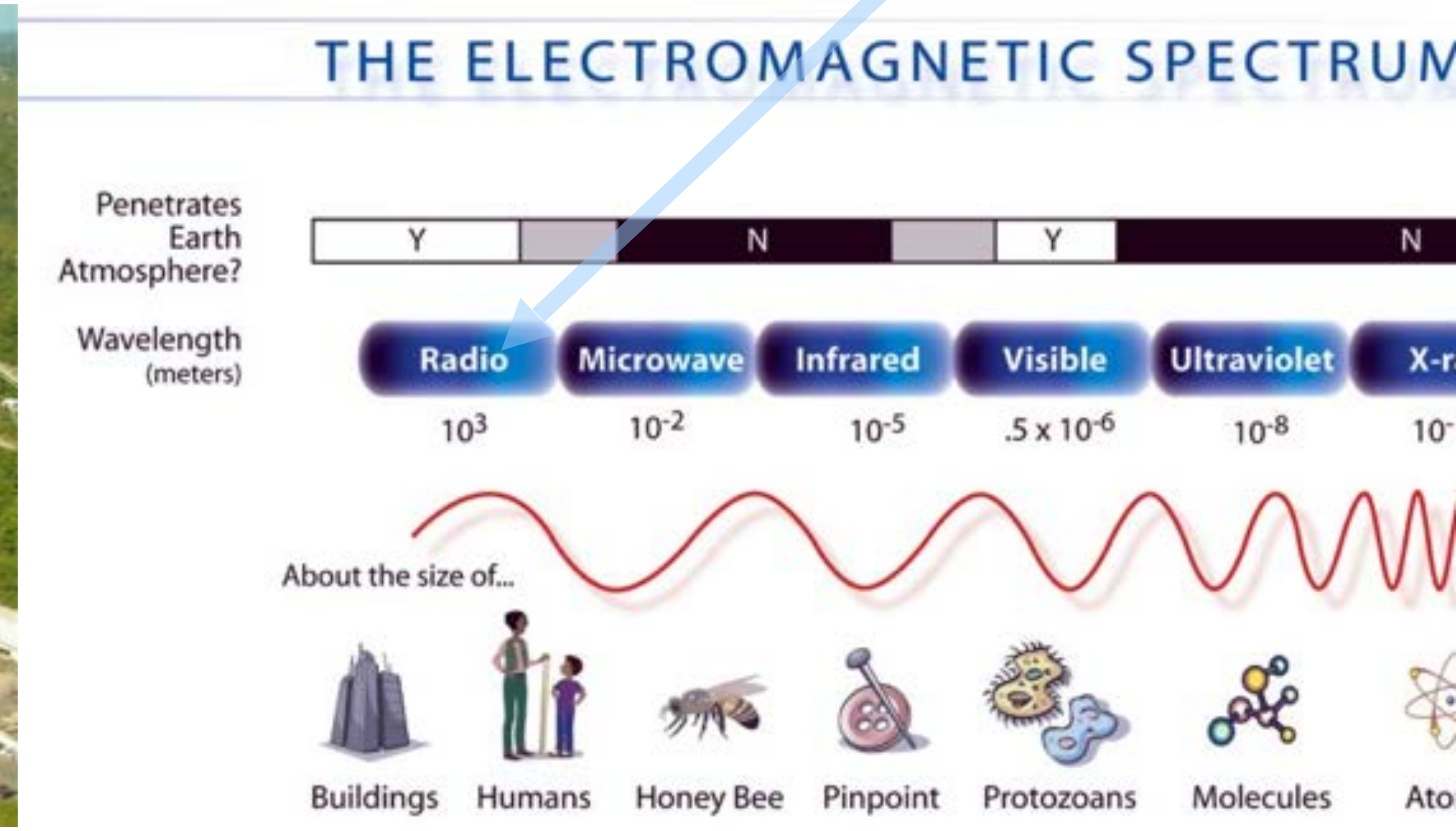


Wavelength Matters

- Longer wavelength
  = poorer resolution
- At 21cm, a radio telescope would need to be miles across to match just a human eye's resolution...
- Biggest we have is 300m (Arecibo), and even that took some “cheating”...

=

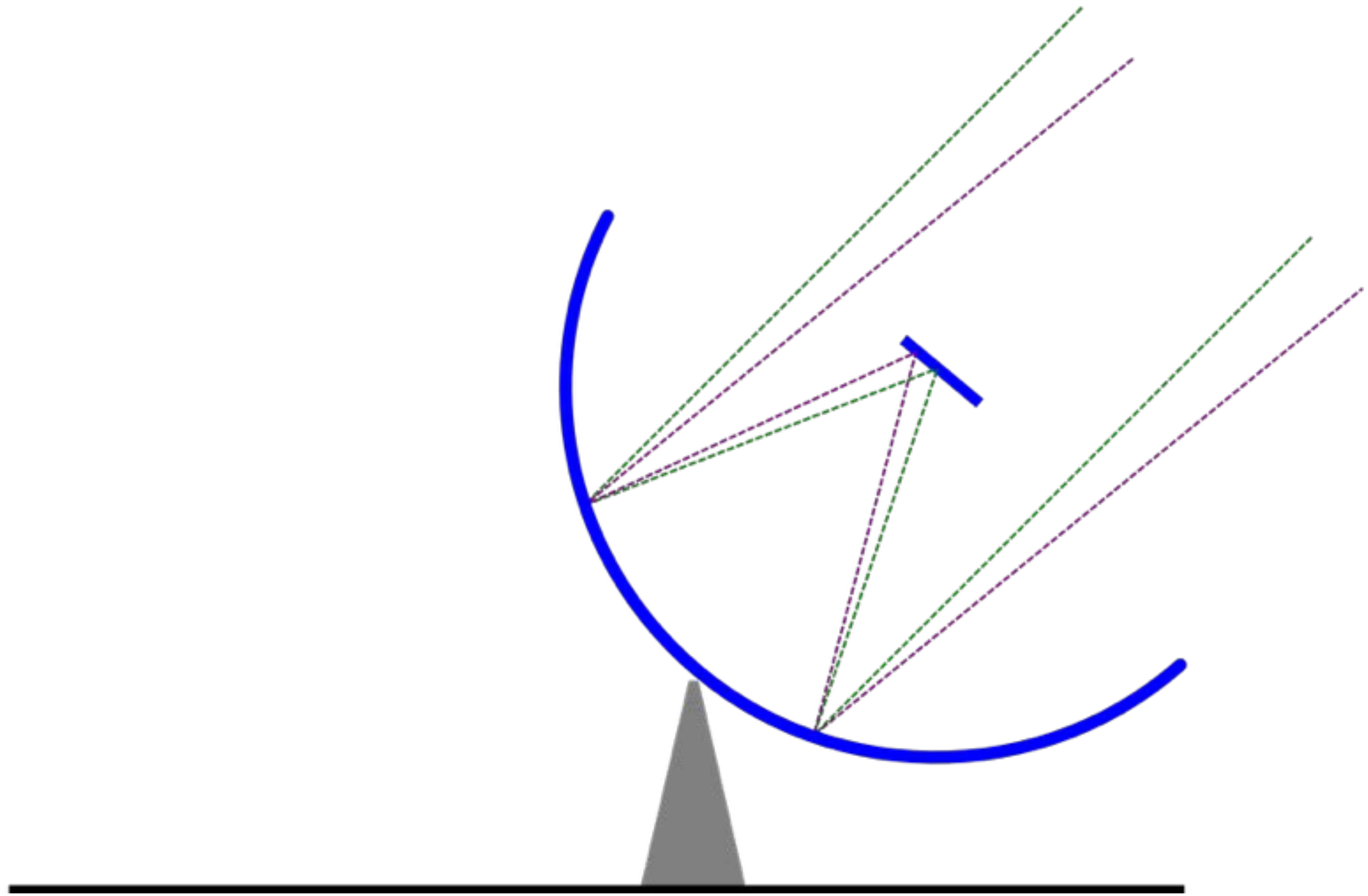
$$\text{resolution} \sim \lambda/D$$



Aperture Synthesis To The Rescue

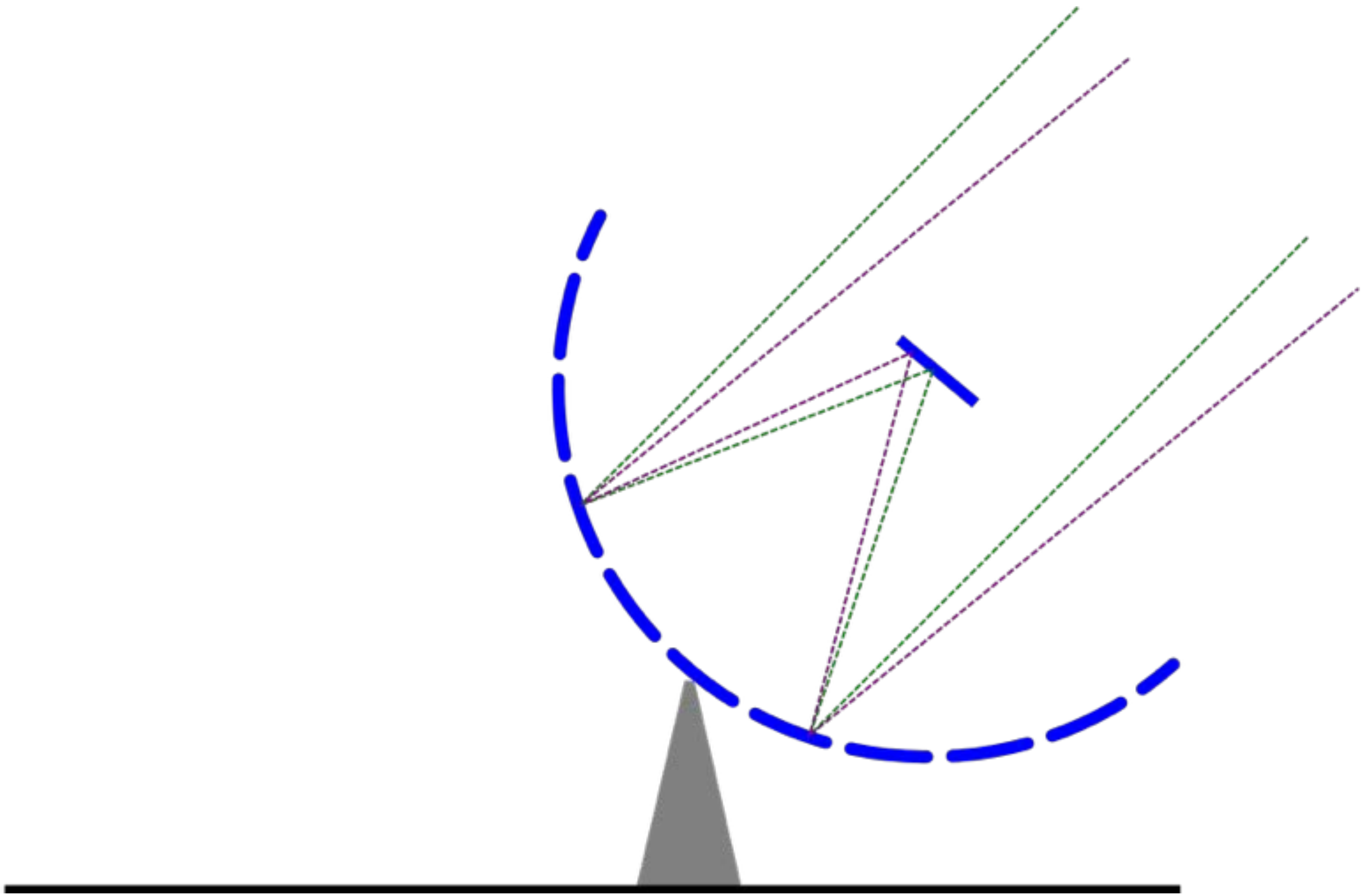
- Radio wavelengths are very large compared to visible light
  - >1,000,000 times longer
- Therefore, a single radio telescope has very poor angular resolution
  - 25m dish ~ size of the full Moon
  - Single “pixel” feeds (the detectors need to be correspondingly large!)
- **1950-60s: Sir Martin Ryle**'s group at Cambridge developed the technique of *aperture synthesis interferometry,* which worked around this problem.

How To Make An Interferometer 1



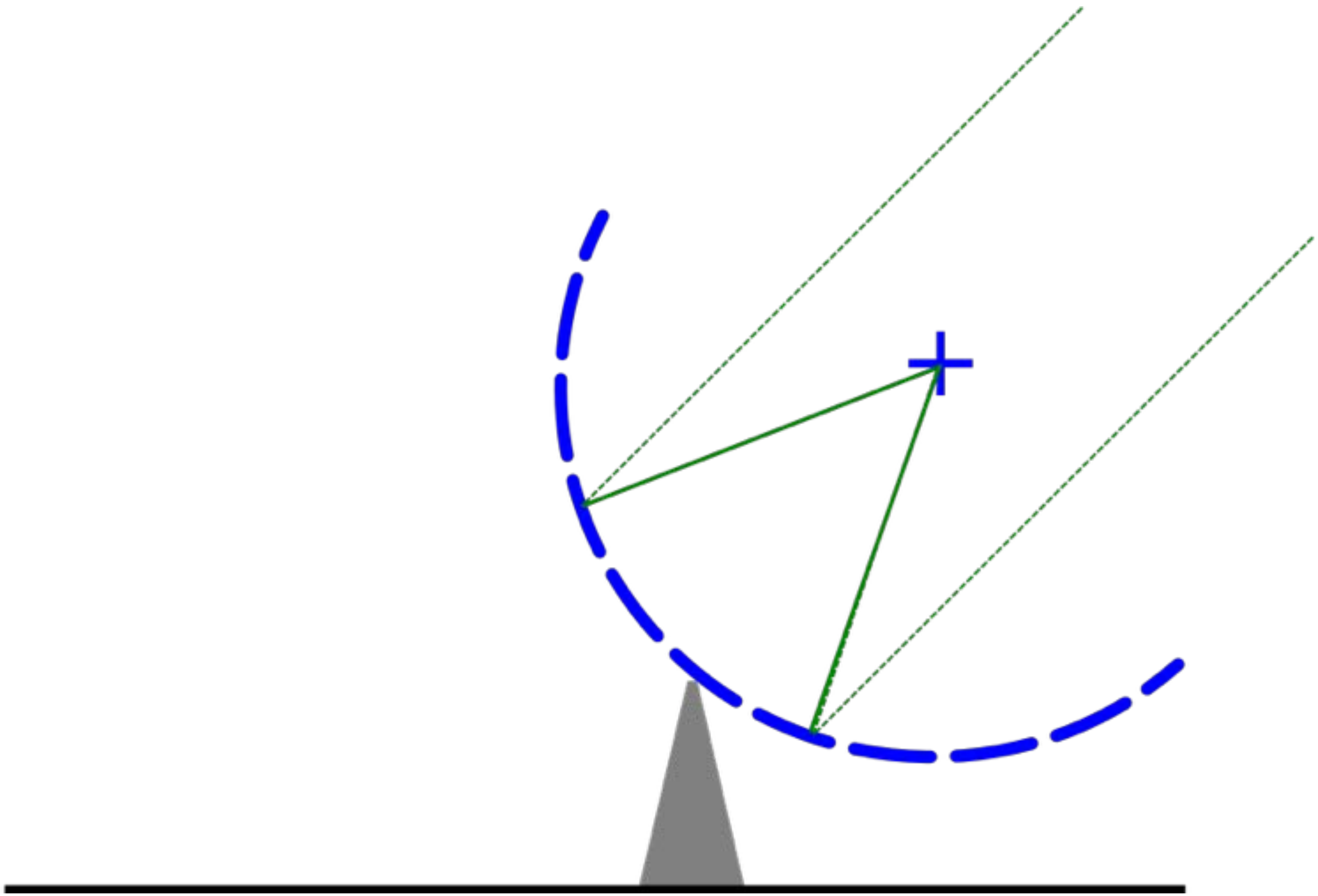
- Start with a normal reflector telescope....

How To Make An Interferometer 2



- Then break it up into sections...

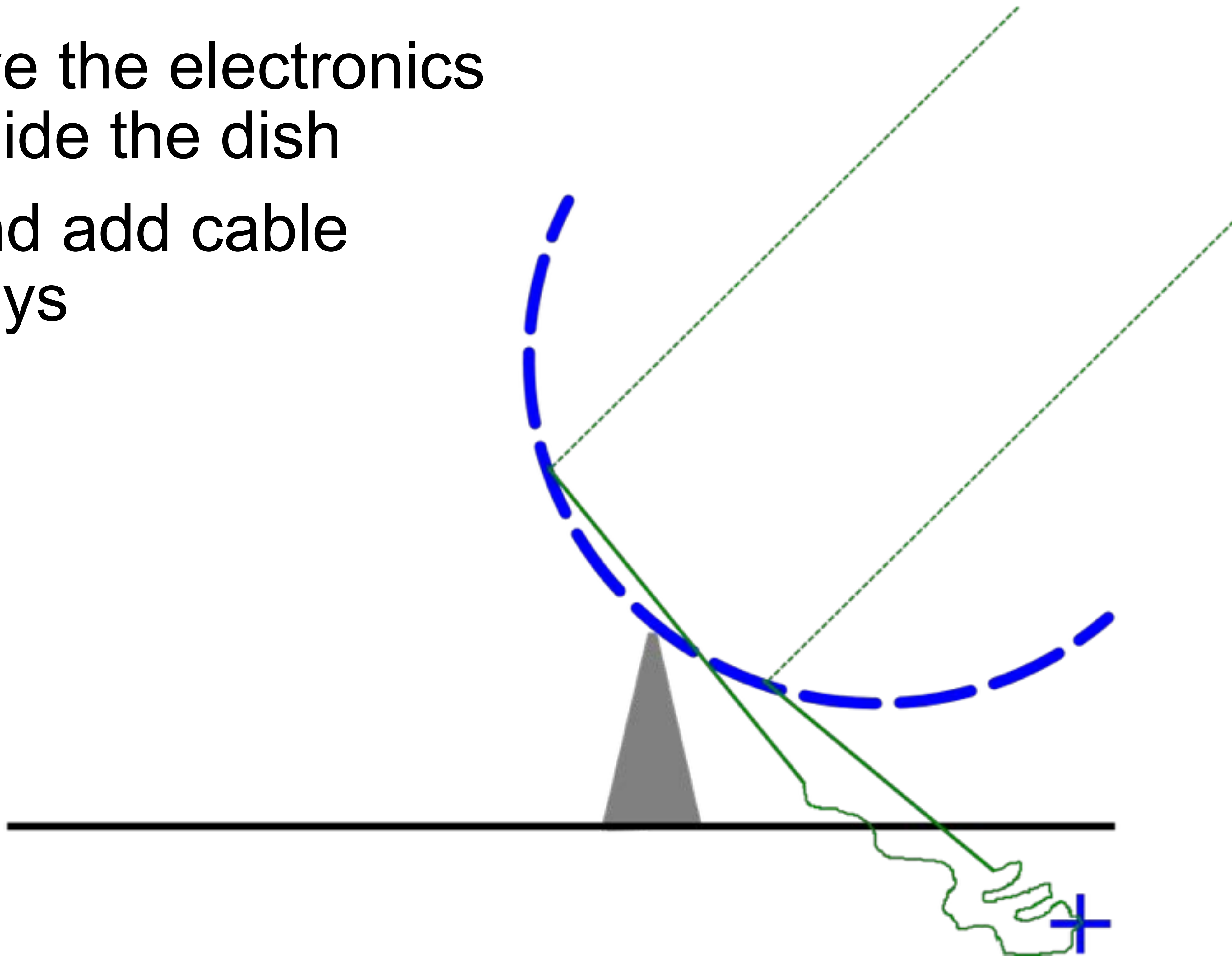
How To Make An Interferometer 3



- Replace the optical path with electronics

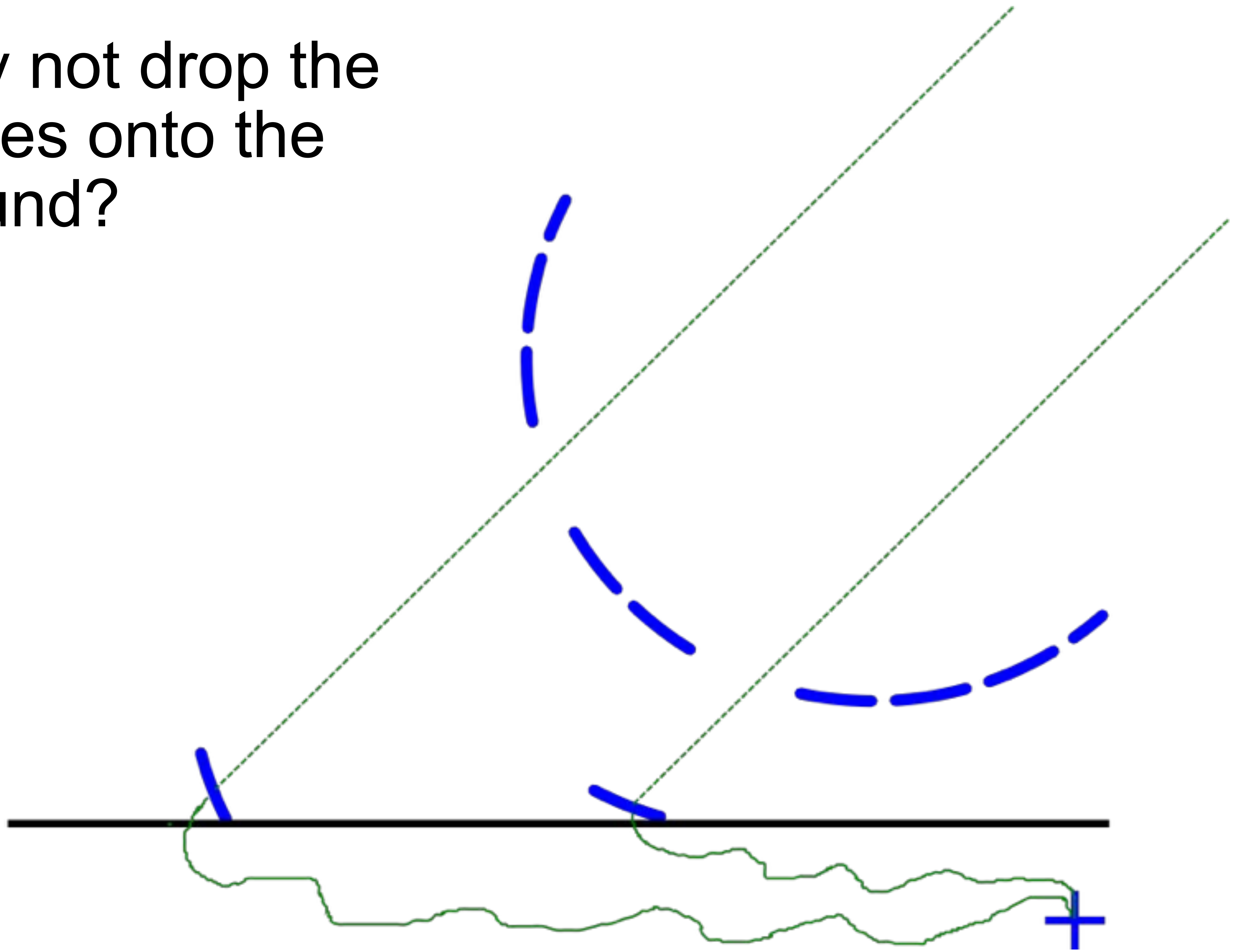
How To Make An Interferometer 4

- Move the electronics outside the dish
- ...and add cable delays



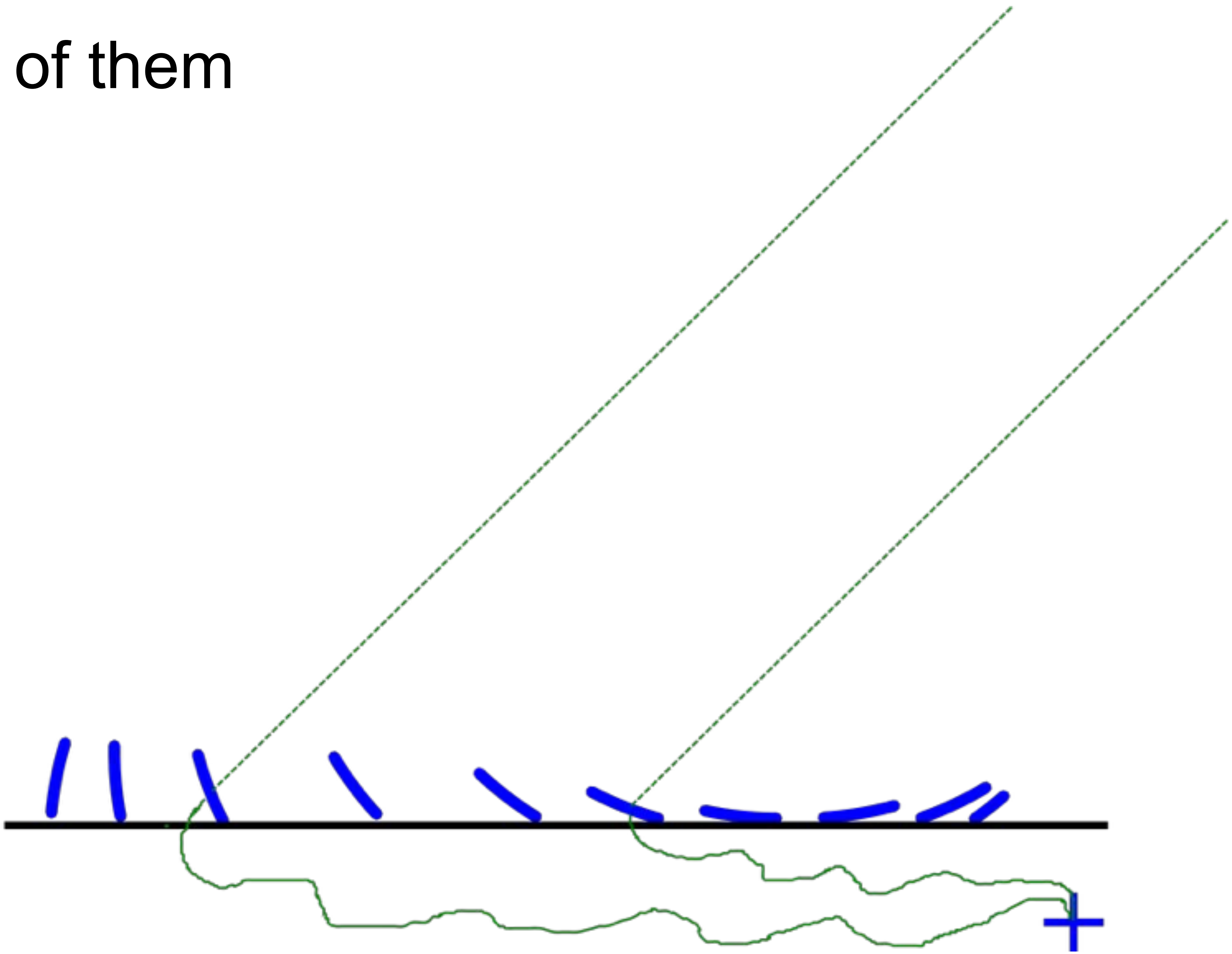
How To Make An Interferometer 5

- Why not drop the pieces onto the ground?



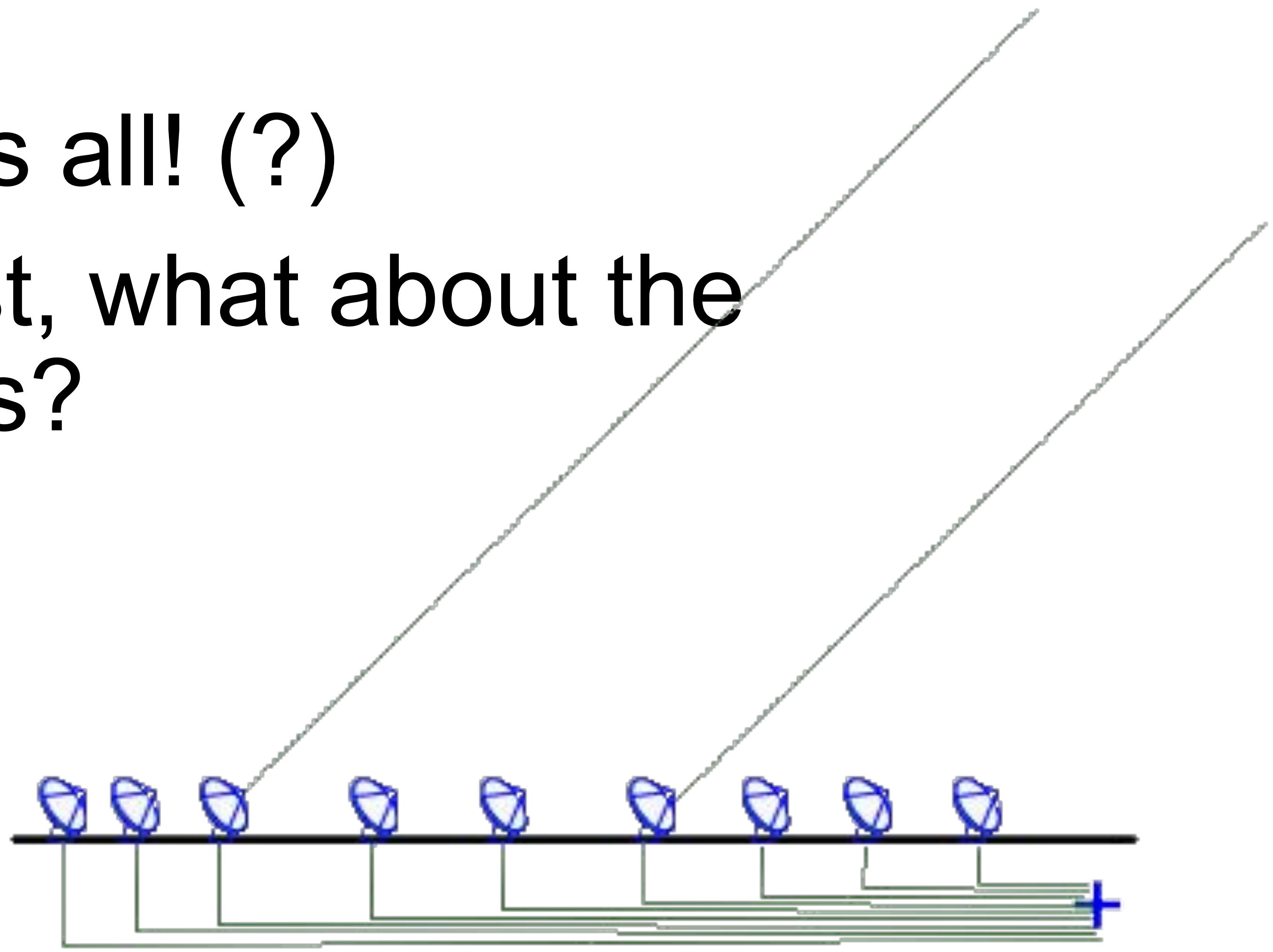
How To Make An Interferometer 6

- ...all of them

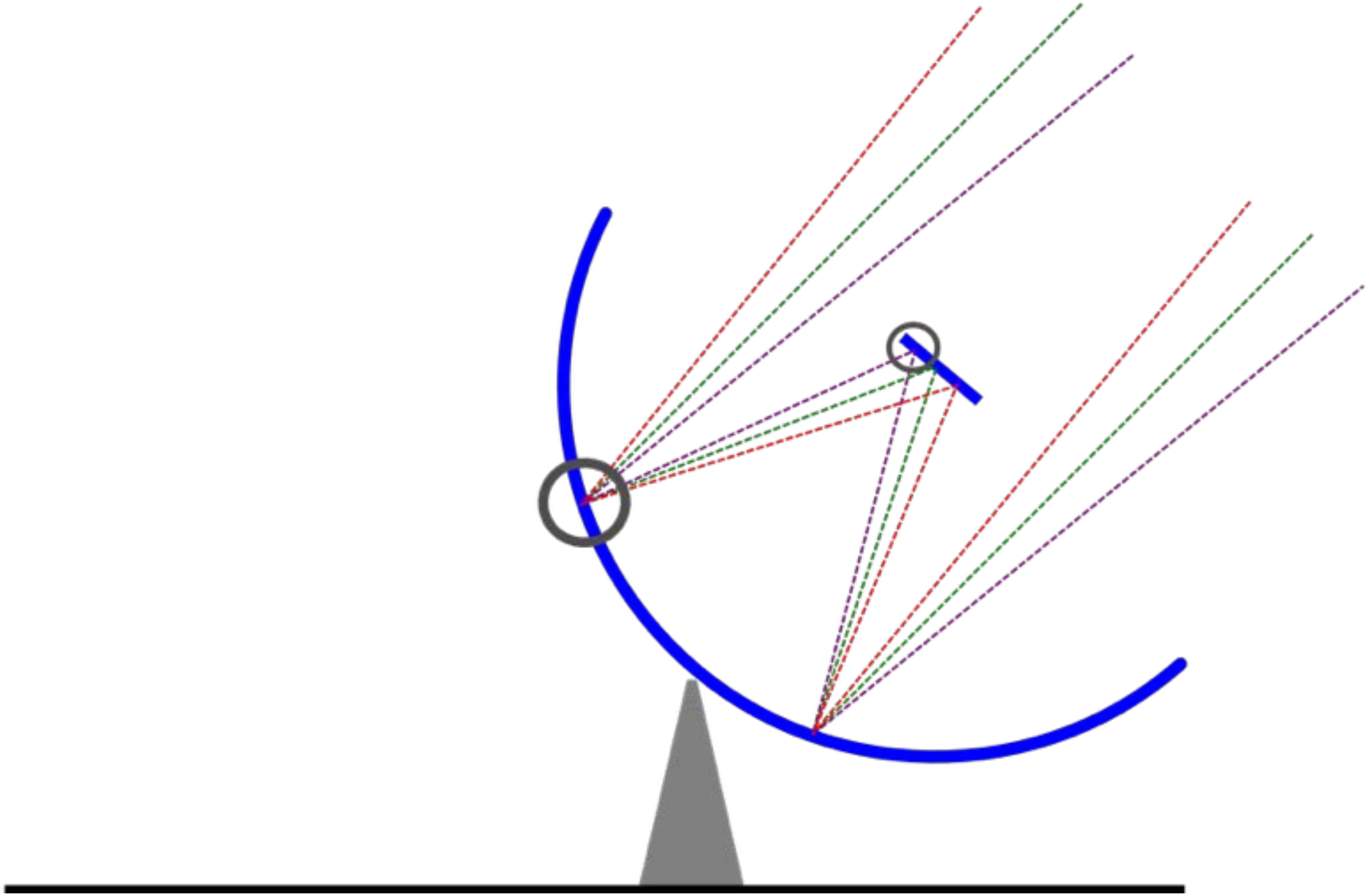


How To Make An Interferometer 7

- And now replace them with an ***array*** of proper radio dishes.
- ...and that's all! (?)
- Well almost, what about the other pixels?

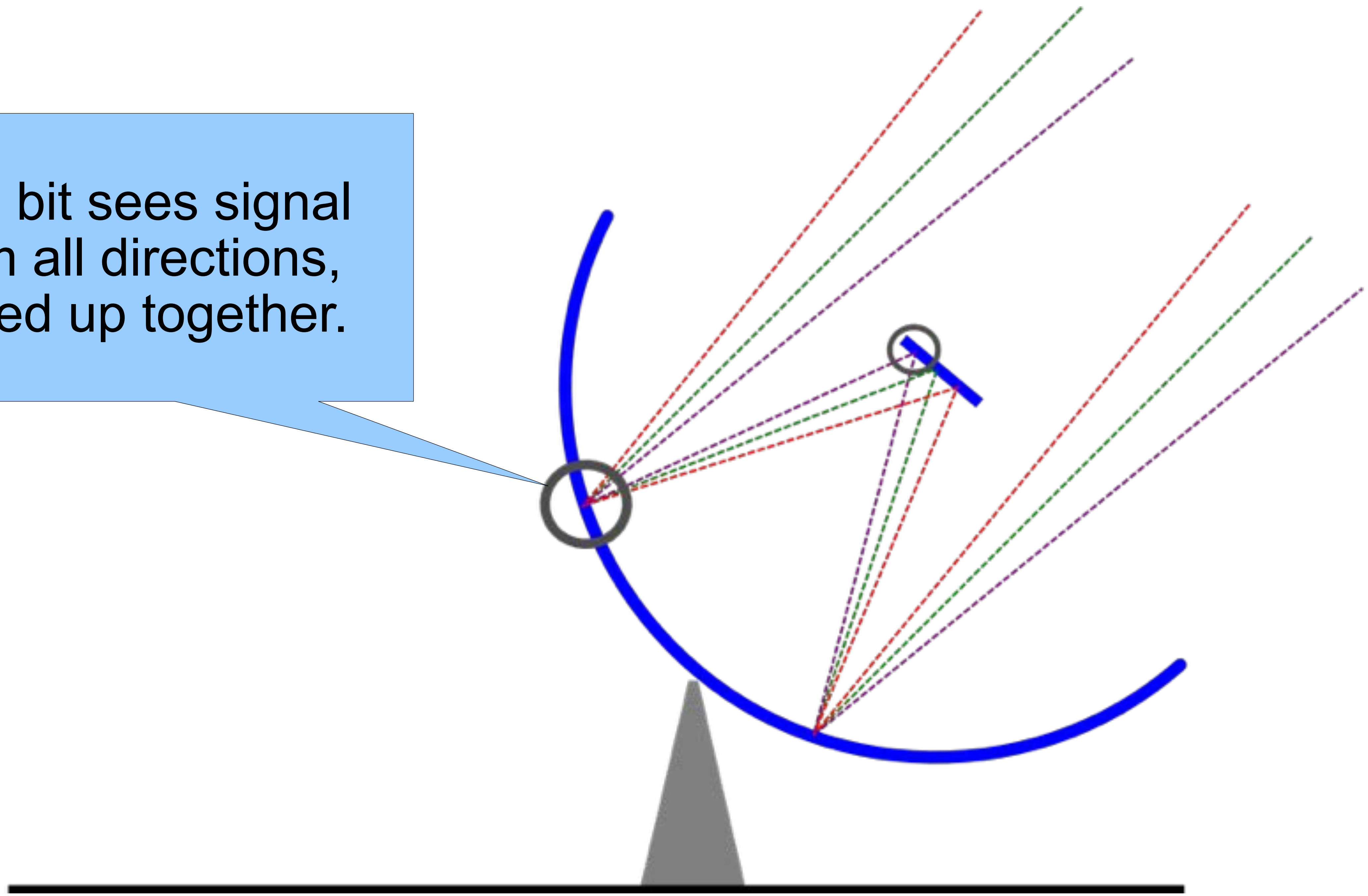


How Does Optical Imaging Do It?

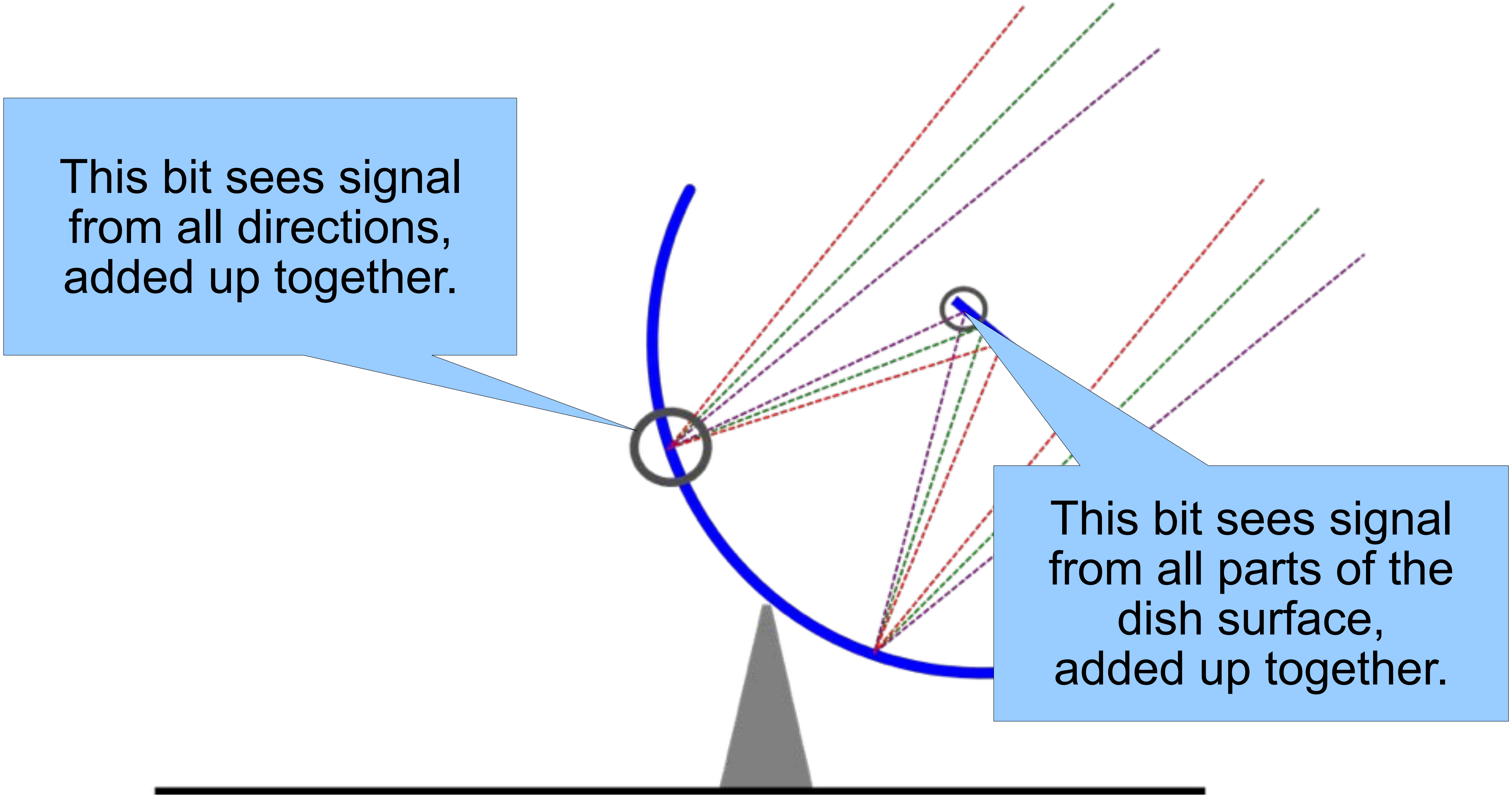


How Does Optical Imaging Do It?

This bit sees signal
from all directions,
added up together.



How Does Optical Imaging Do It?

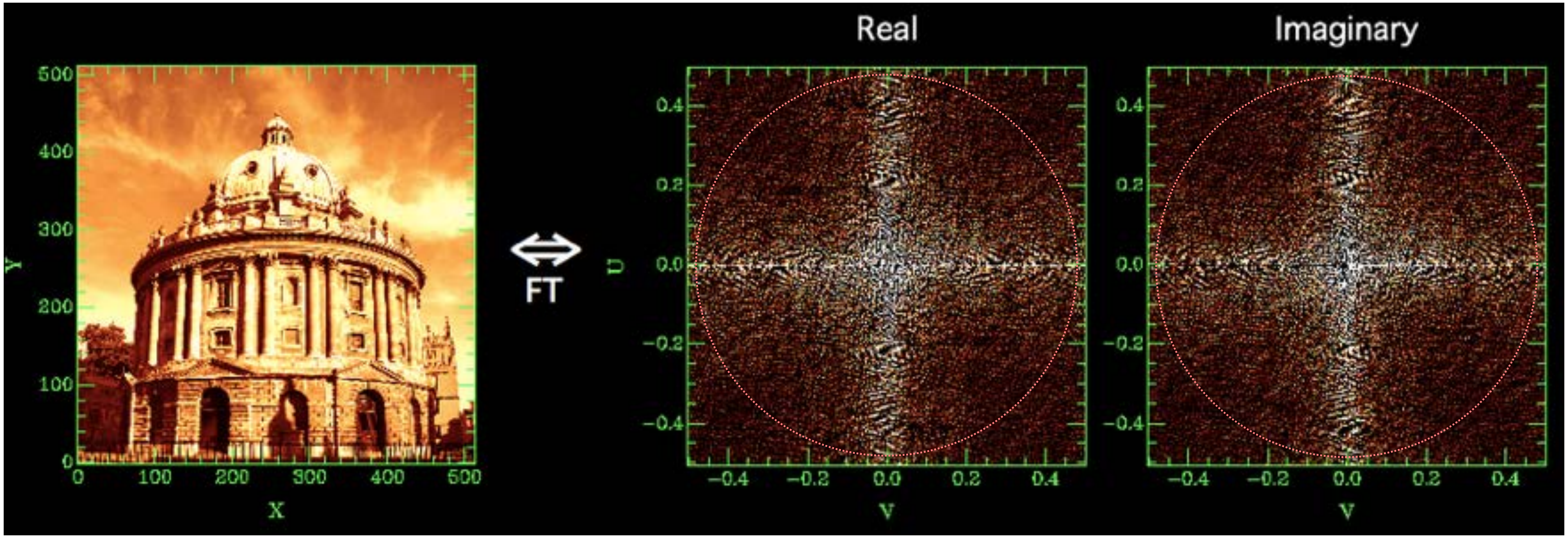


Fourier Transforms

 = The world's fastest computer for doing Fourier transforms

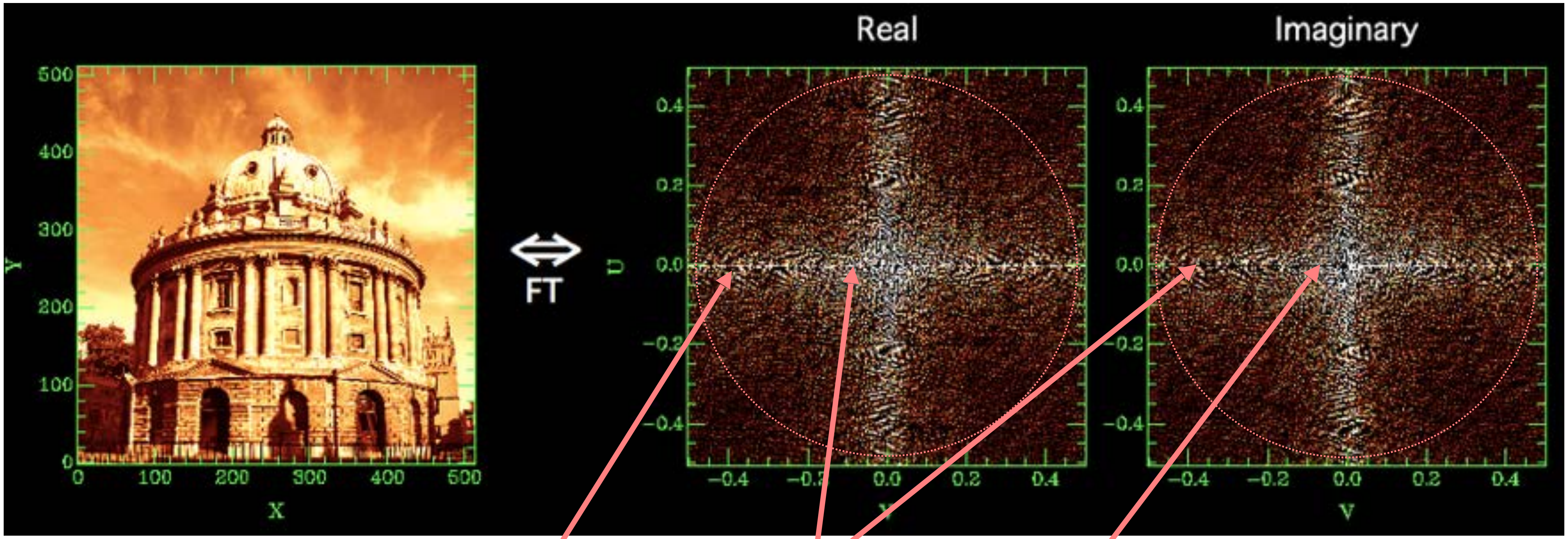
- An optical imaging system implicitly performs two Fourier transforms:
  1. Signal over the aperture = FT of the sky
  2. Signal over focal plane = inverse FT of the aperture
- A radio interferometer array measures (1)
  - Then we do the second FT in software
  - Hence, “aperture synthesis”

The Fourier (aka *uv*-) Plane

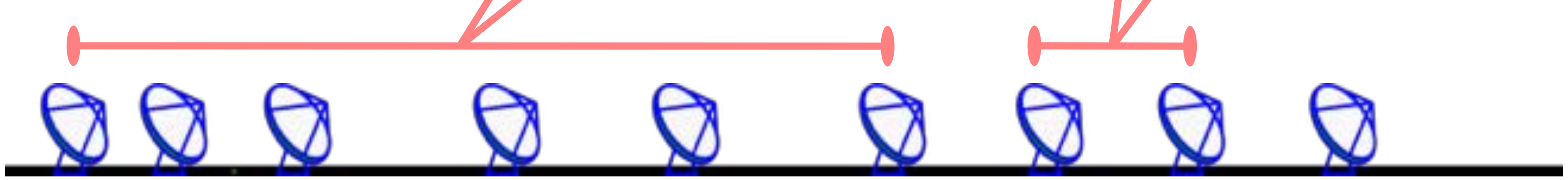


- The Fourier plane (***uv*-plane**) is a mathematically equivalent representation of the **image** plane, we can recover one from the other via Fourier transforms
- One pair of antennas (**baseline**) samples one point in the *uv*-plane (one **complex visibility**)
- An optical system (e.g. your eye) samples the entire plane at once

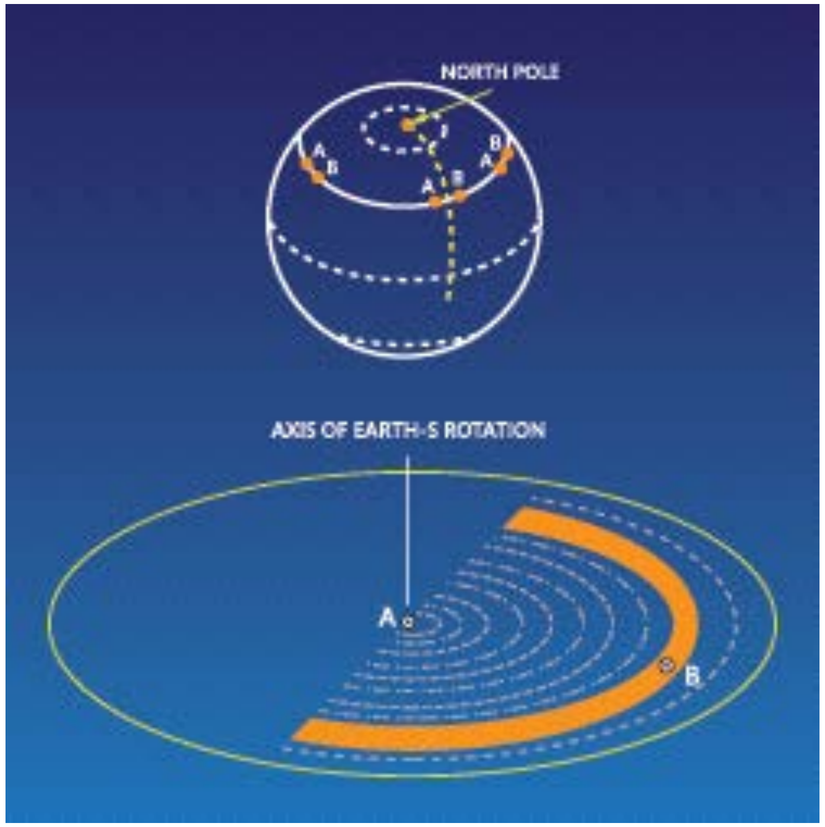
The Fourier (aka *uv*-) Plane



- The Fourier plane (***uv*-plane**) is a mathematically equivalent representation of the **image** plane, we can recover one from the other via Fourier transforms
- One pair of antennas (**baseline**) samples one point in the *uv*-plane (one **complex visibility**)
- An optical system (e.g. your eye) samples the entire plane at once



Earth Rotation Aperture Synthesis



- Each baseline samples one *uv*-point
- An array of *N* antennas will sample *N*(*N*-1)/2 points
- This is usually not enough to recover a good image, but...
- As the Earth rotates, each baseline's *uv*-point is rotating as well
  - ...while the Universe is (mostly) still
- Over several hours, we can sample most of the *uv*-plane with even a modest number of antennas
- Resolution determined by longest baseline:

$$\sim \lambda/B$$

Some Real-Life Arrays: WSRT

- WSRT (Westerbork Synthesis Radio Telescope), The Netherlands
- 14x25m dishes on an East-West line
- Max baseline 2.7km
- Completed 1970, upgraded since
- World record dynamic range, still

Some Real-Life Arrays: JVLA

- JVLA (Karl G. Jansky Very Large Array), New Mexico, USA
- 27x25m dishes
- Reconfigurable (via rail tracks) to a longest baseline of 36km
- Completed in 1980 (as simply the “VLA”), recently upgraded and renamed
- The most successful radio telescope in the world, in terms of science produced

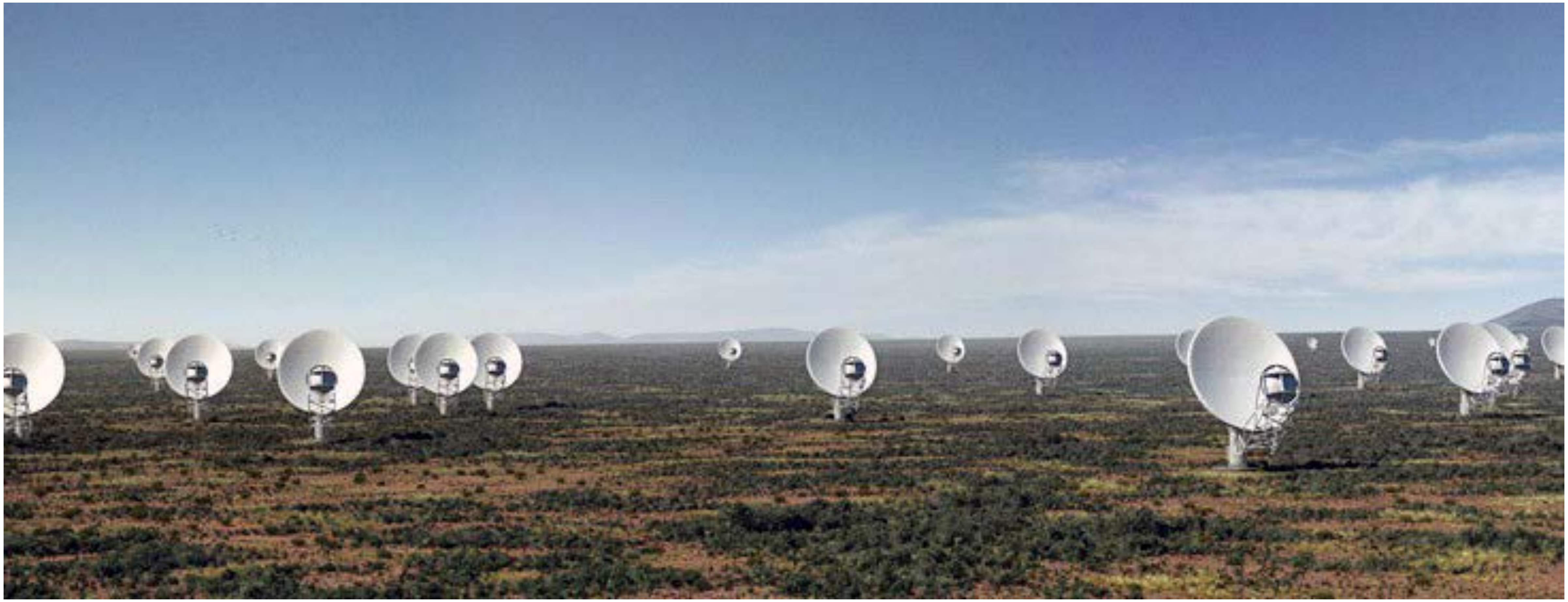
Some Real-Life Arrays: LOFAR

- Low Frequency Array (LOFAR)
- 36 stations (not dishes!) across The Netherlands
- 8 (and counting) international stations
- Inaugurated 2010

Some Real-Life Arrays: EVN & VLBA

- [not only] European VLBI (Very Long Baseline Interferometry) Network
- Connects radio telescopes around the world into an *ad hoc* interferometer
- Baselines of thousands of kilometres
- US analogue: VLBA (Very Long Baseline Array)
- African VLBI project will fill in the North-South gaps

Future Arrays: MeerKAT (2016)



- 64x13.5m dishes, 8km longest baseline
- Now under construction in the Karoo

SKA (2024)

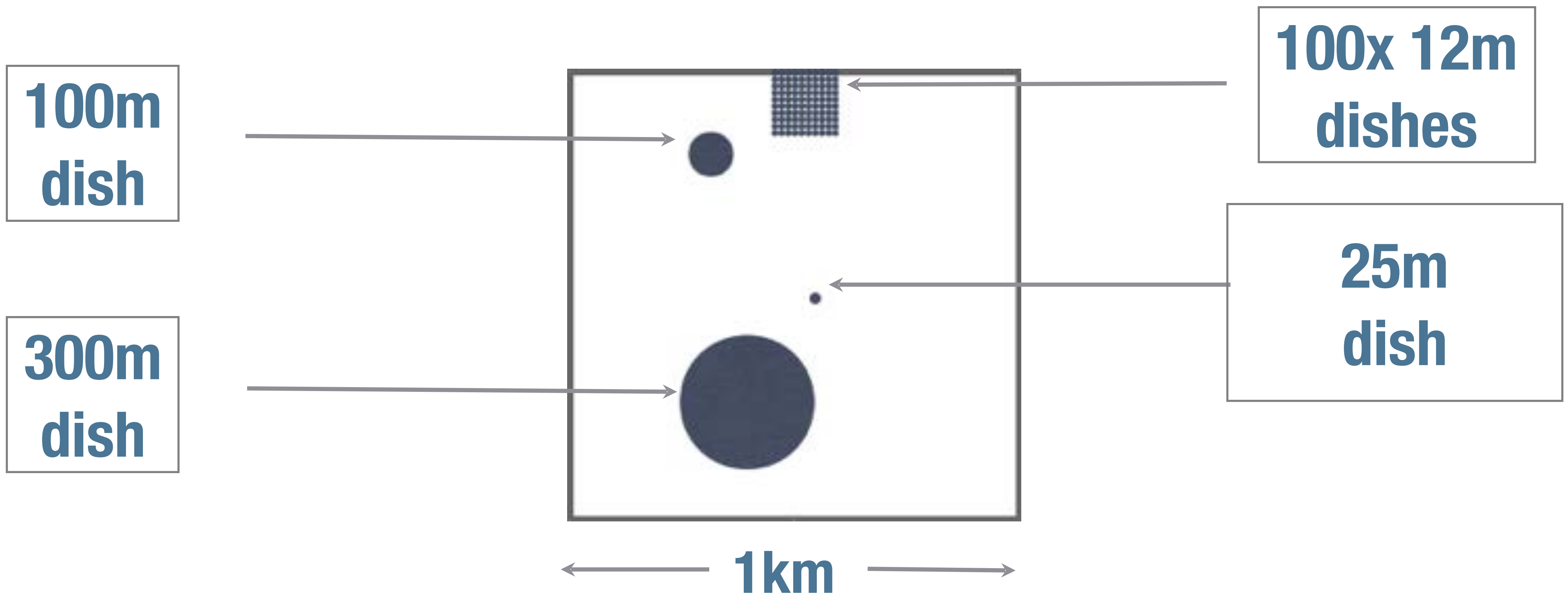
- Dish component: ~3000x15m dishes (= 1 km^2)
  - Half of them within a 5km "core" in the Karoo
- Longest baselines to 3000 km
- 2020: Phase One (SKA1) with 250 dishes (incorporates MeerKAT)

What Is The SKA?

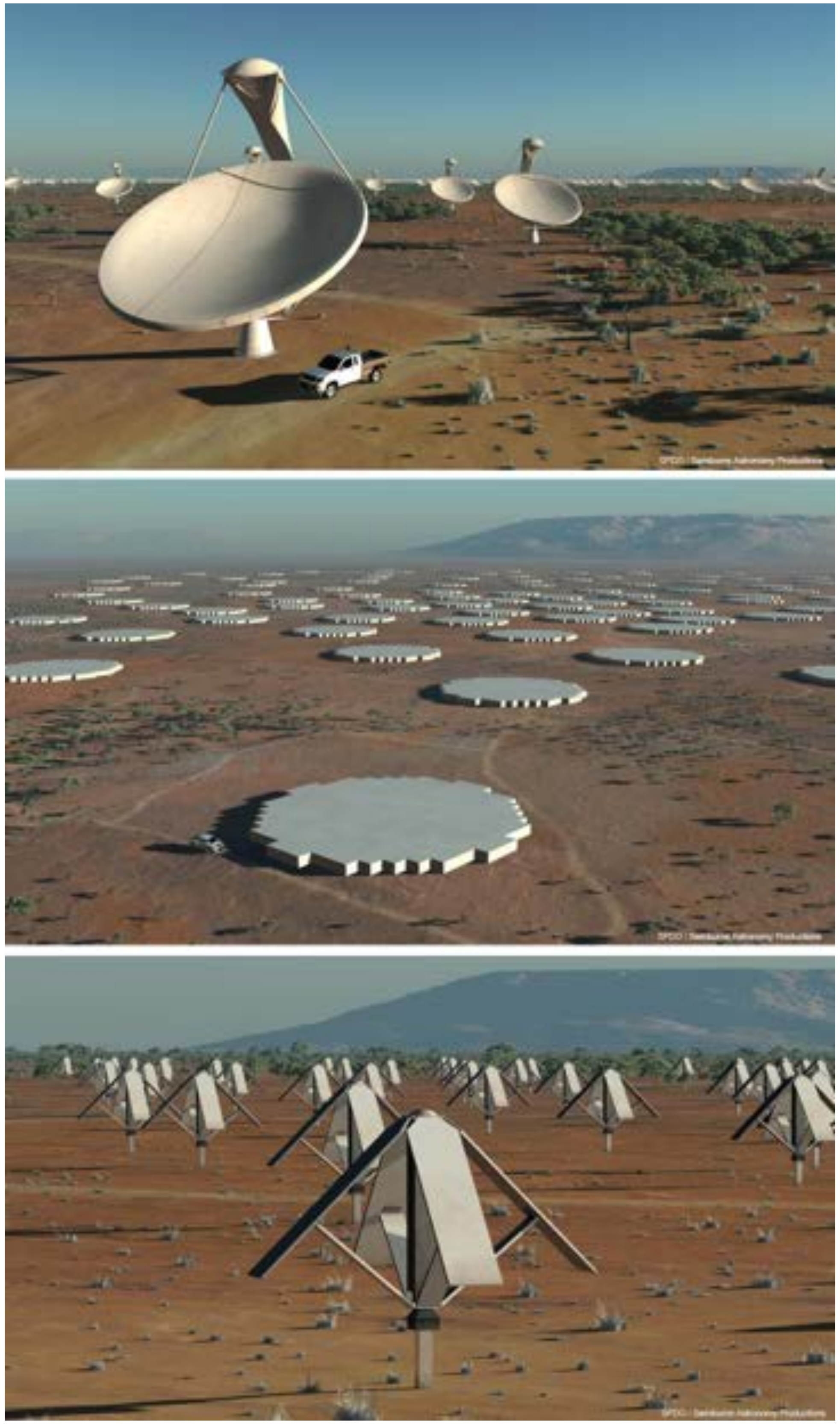
- The Square Kilometre Array will be the biggest radio telescope in the world, and one of the biggest and most challenging projects in science
  - The SKA Organization: Australia, Canada, China, Germany, Italy, Netherlands, New Zealand, South Africa, Sweden, UK
- Cost R20+ billion
- On 25 May 2012, the SKA Organisation announced that the SKA would be jointly sited in South Africa and Australia, with ~70% coming to South Africa

Why SKA?

- Aperture synthesis has nailed the *resolution* problem
  - VLBI (and Space VLBI) routinely achieves higher res than optical astronomy
- Sensitivity still limited by collecting area D^2



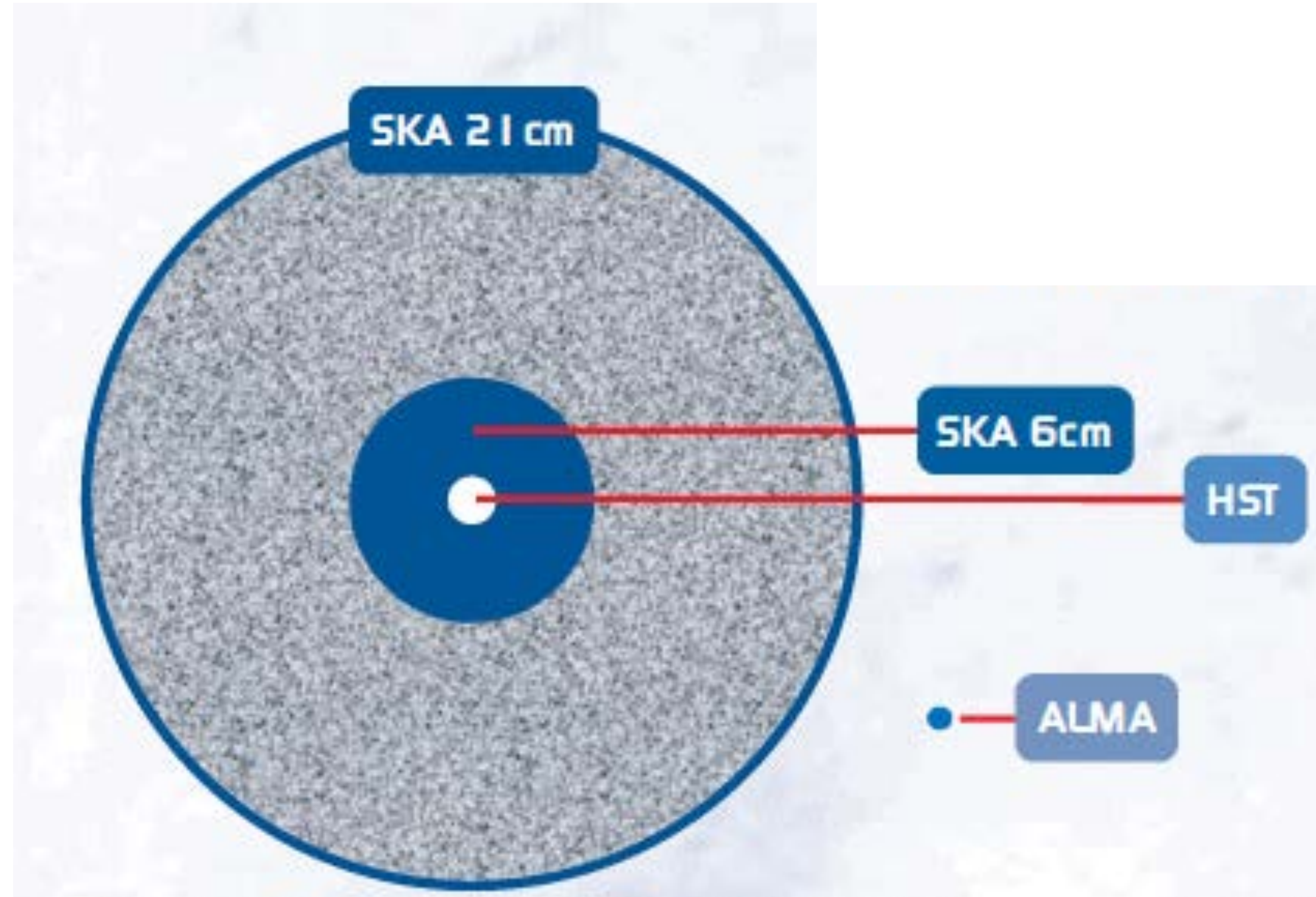
SKA Components

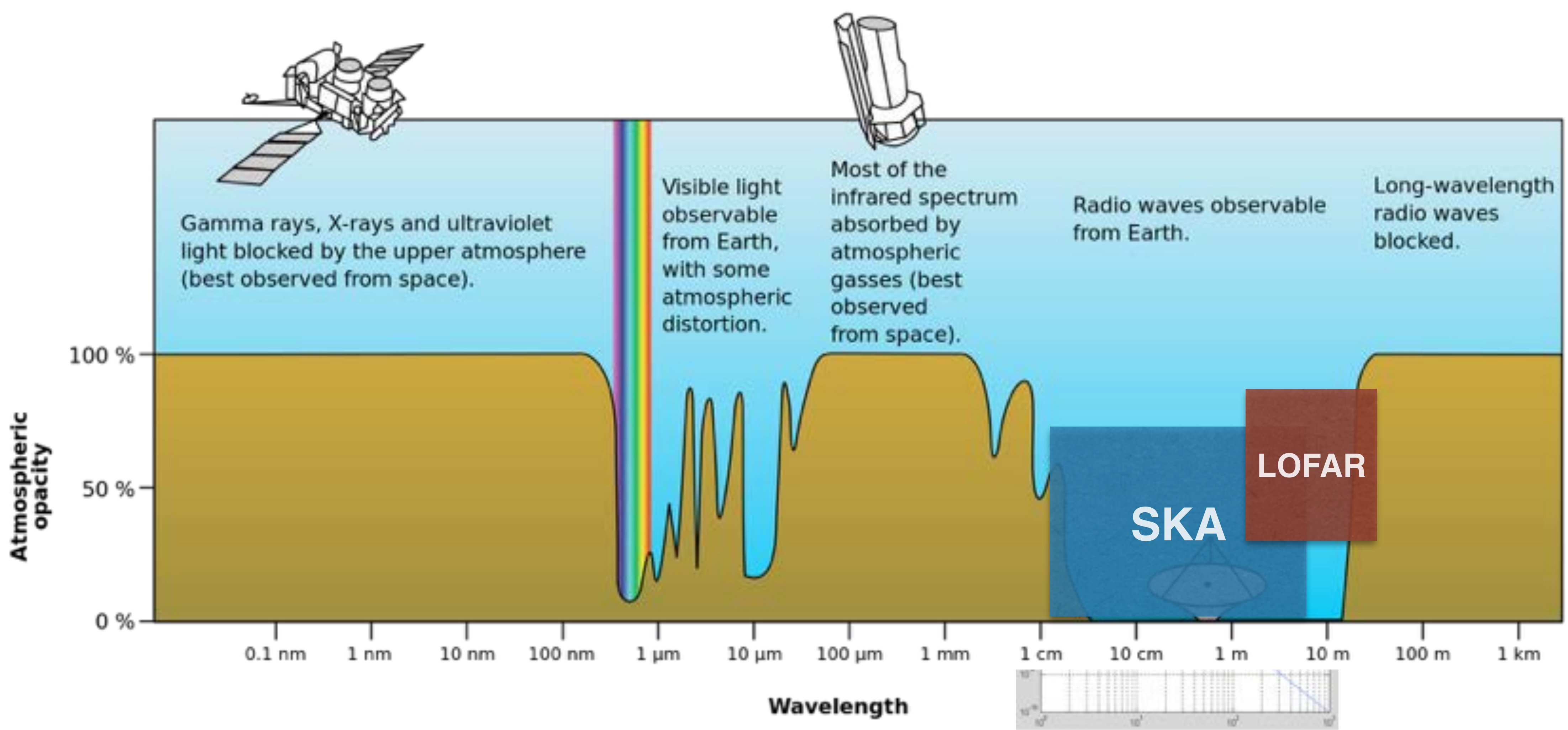


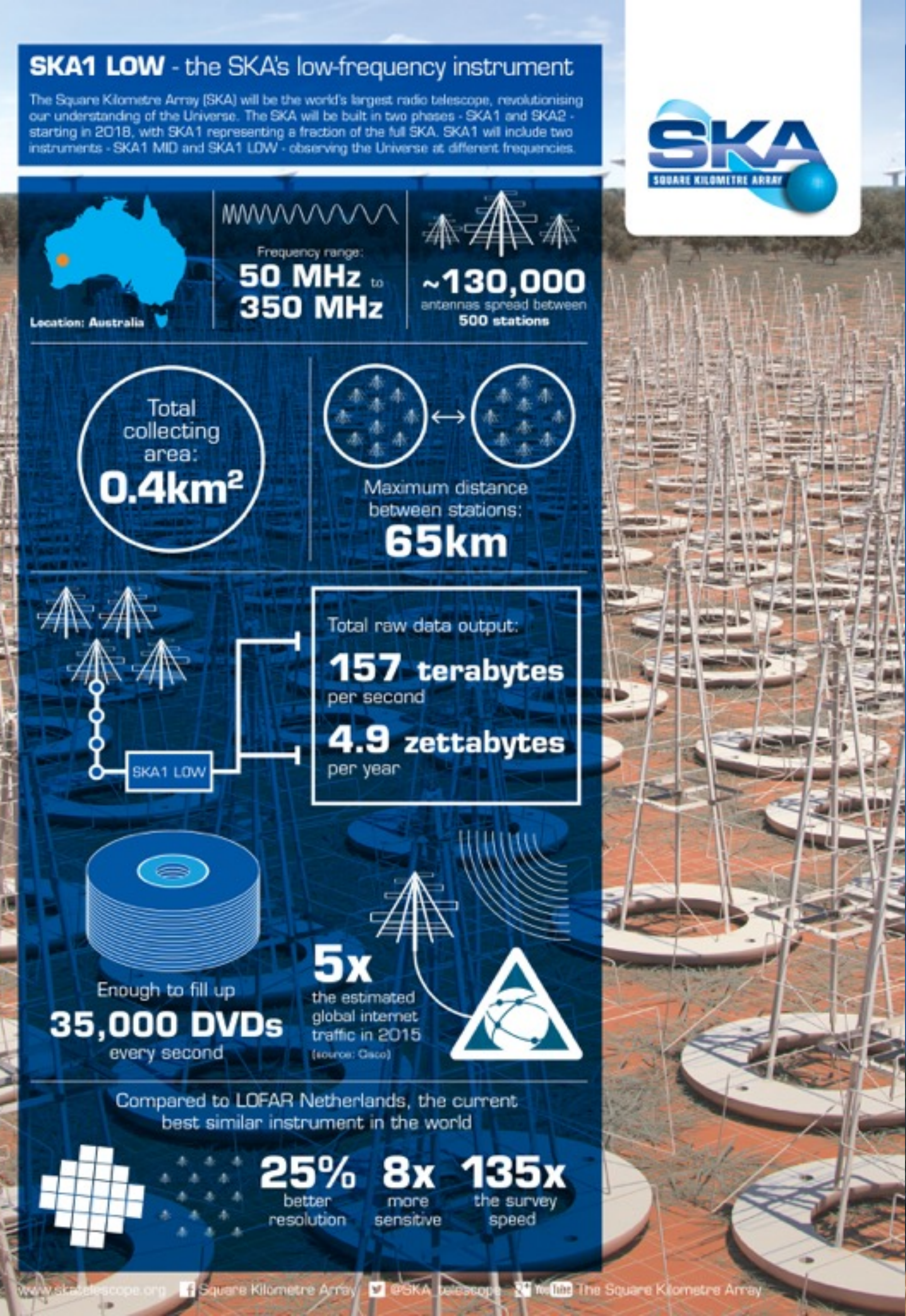
- SKA is actually three instruments in one, for three frequency ranges
  - SKA-Low: sparse aperture arrays, 70-500 MHz, Australia
  - SKA-Mid: dense aperture arrays, 500-800 MHz, South Africa
  - SKA-Dishes: 500 MHz – 10 GHz, South Africa
- SKA-Low is about 30% of the instrument

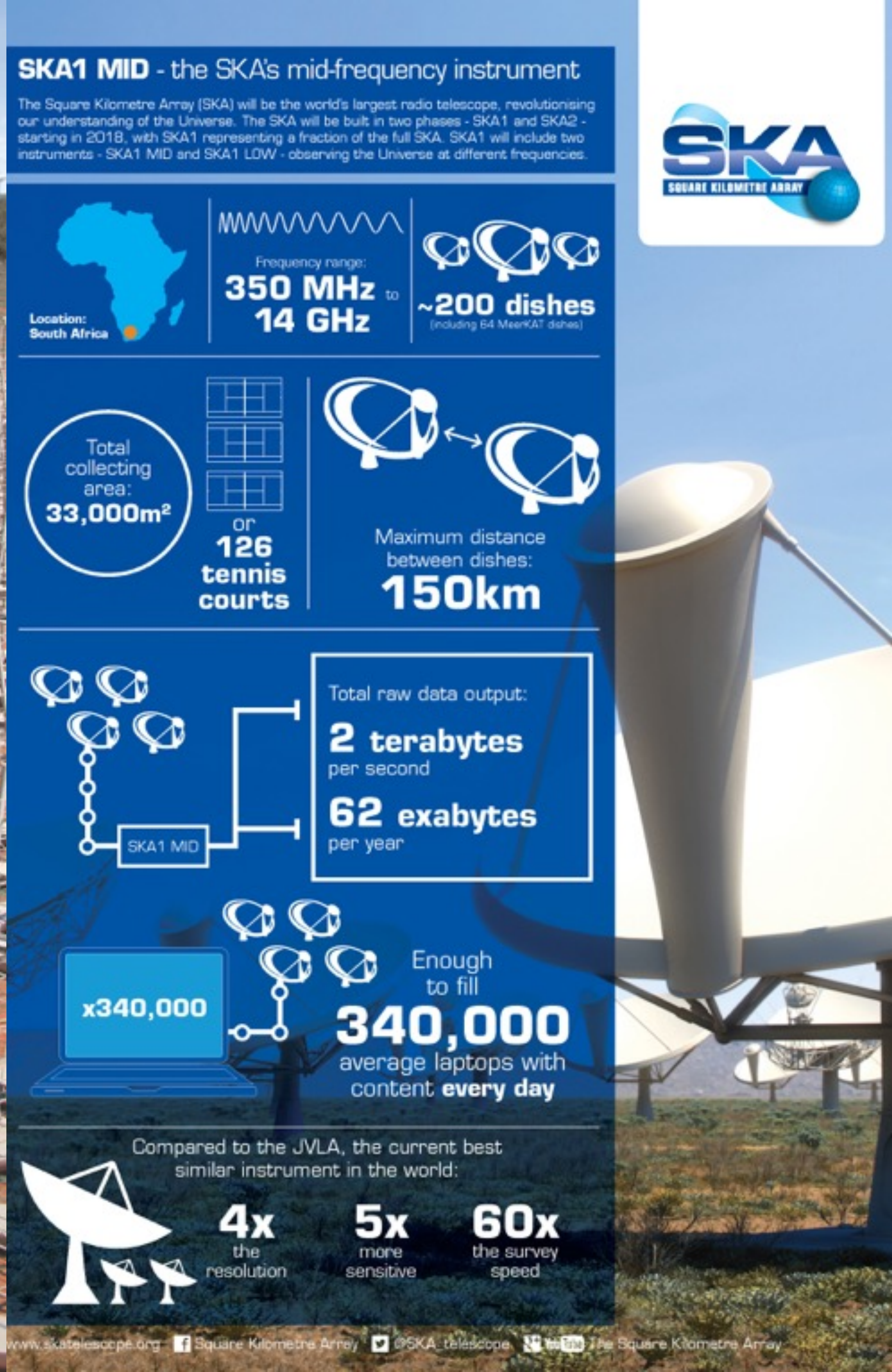
SKA numbers

FoV	1°x1° @ 21 cm
Ang. Res.	<0.1" @ 21 cm
Data flow	~1 TB / min
Ang. Res.	<0.1" @ 21 cm
Freq	50 MHz to 14 GHz
λ	6 m to 2.1 cm





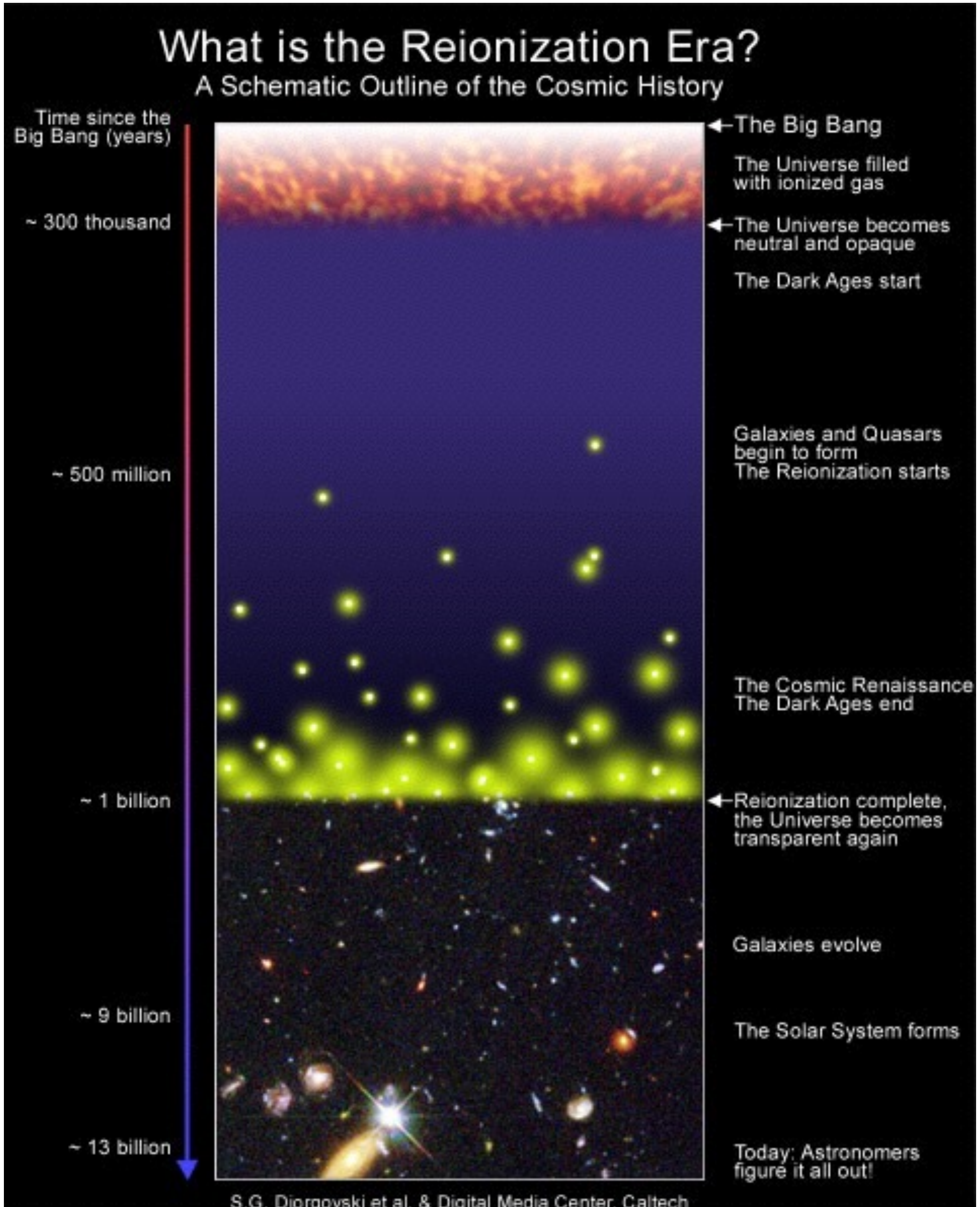

SKA1 LOW - the SKA's low-frequency instrument
The Square Kilometre Array (SKA) will be the world's largest radio telescope, revolutionising our understanding of the Universe. The SKA will be built in two phases - SKA1 and SKA2 - starting in 2018, with SKA1 representing a fraction of the full SKA. SKA1 will include two instruments - SKA1 MID and SKA1 LOW - observing the Universe at different frequencies.
SKA
SQUARE KILOMETRE ARRAY
Location: Australia
Frequency range:
50 MHz to 350 MHz
~130,000
antennas spread between
500 stations
Total collecting area:
0.4km²
Maximum distance between stations:
65km
Total raw data output:
157 terabytes
per second
4.9 zettabytes
per year
SKA1 LOW
Enough to fill up
35,000 DVDs
every second
5x
the estimated global internet traffic in 2015
(source: Cisco)
Compared to LOFAR Netherlands, the current best similar instrument in the world
25%
better resolution
8x
more sensitive
135x
the survey speed
www.skatelescope.org Square Kilometre Array @SKA_telescope The Square Kilometre Array


SKA1 MID - the SKA's mid-frequency instrument
The Square Kilometre Array (SKA) will be the world's largest radio telescope, revolutionising our understanding of the Universe. The SKA will be built in two phases - SKA1 and SKA2 - starting in 2018, with SKA1 representing a fraction of the full SKA. SKA1 will include two instruments - SKA1 MID and SKA1 LOW - observing the Universe at different frequencies.
SKA
SQUARE KILOMETRE ARRAY
Location: South Africa
Frequency range:
350 MHz to 14 GHz
~200 dishes
(including 64 MeerKAT dishes)
Total collecting area:
33,000m²
or
126 tennis courts
Maximum distance between dishes:
150km
Total raw data output:
2 terabytes
per second
62 exabytes
per year
SKA1 MID
x340,000
Enough to fill
340,000
average laptops with content every day
Compared to the JVLA, the current best similar instrument in the world:
4x
the resolution
5x
more sensitive
60x
the survey speed
www.skatelescope.org Square Kilometre Array @SKA_telescope The Square Kilometre Array

SKA Headline Science

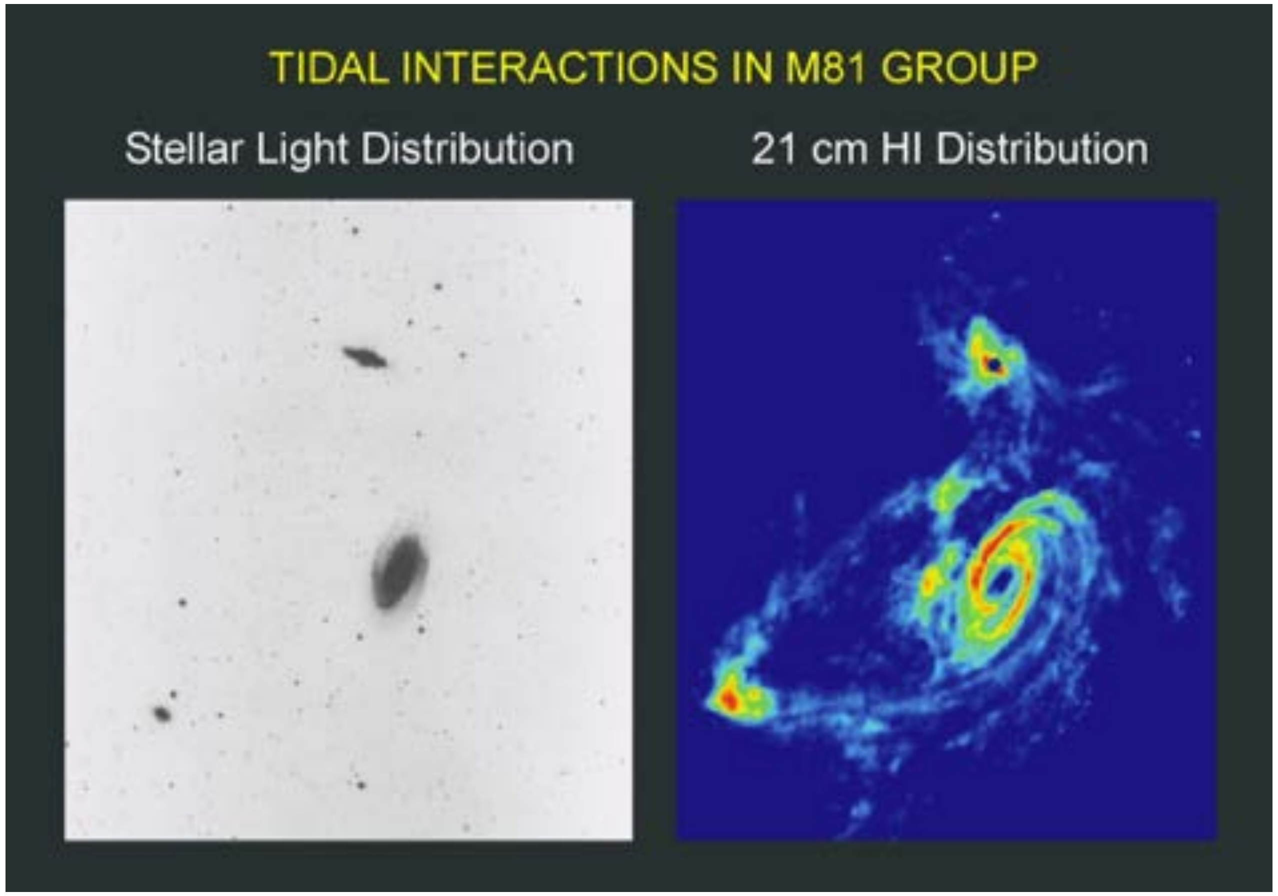
- Probing the Dark Ages
- Galaxy evolution, cosmology and dark energy
- Cosmic magnetism
- Strong field tests of gravity (“Was Einstein Right?”)
- The Cradle Of Life

Probing The Dark Ages



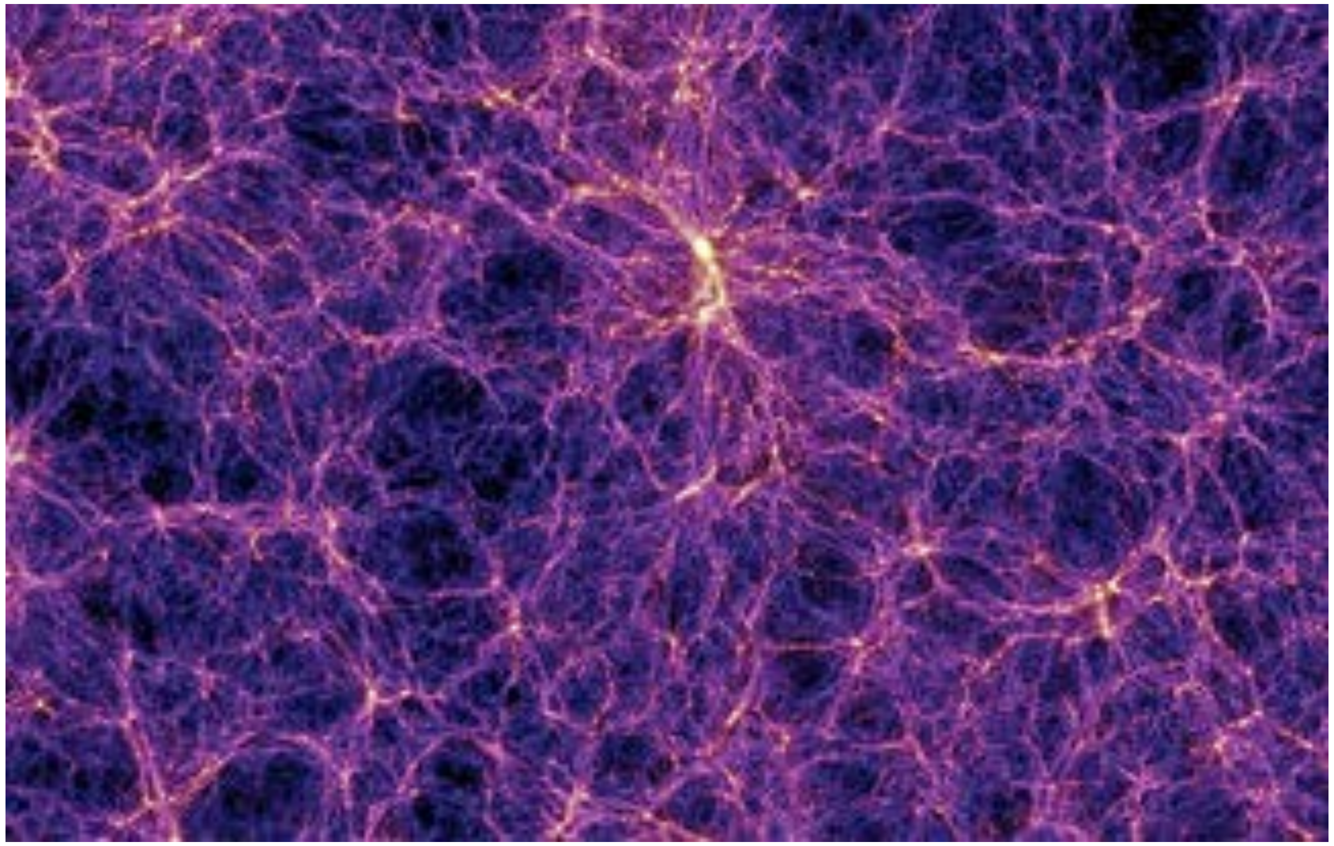
- Direct imaging of HI at extremely high redshift (i.e. at <200 MHz – down from 1420!) will open a window on the Epoch of Reionization (EoR)

Galaxy Evolution, Cosmology, Dark Energy



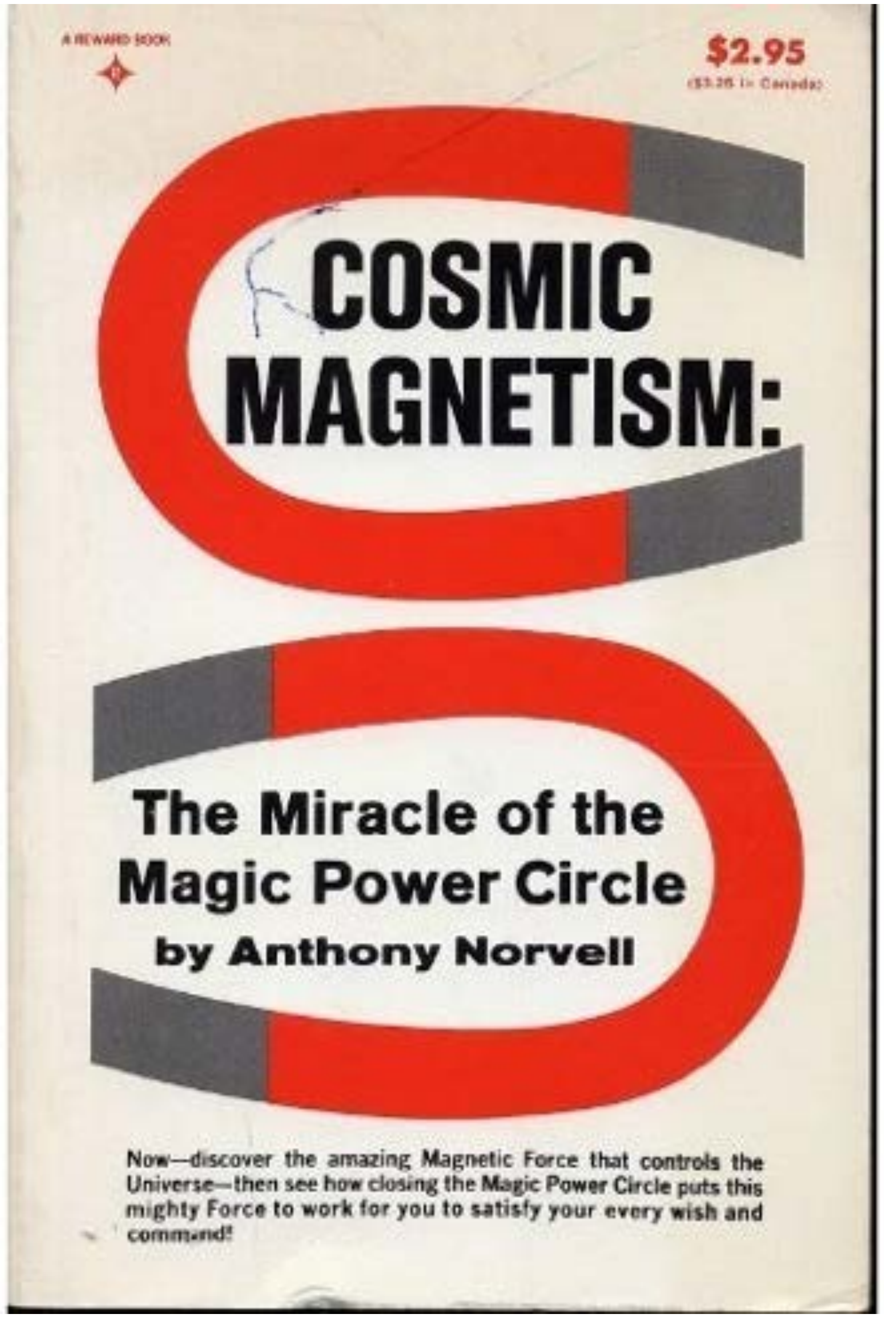
HI is the main ingredient of galaxies, and can be observed unobscured. “Cosmic census”: how do galaxies form and evolve?

On very large scales, galaxies tend to form in sheets and filaments of a “cosmic web”. Can be observed in HI, with redshifts giving a 3D view!

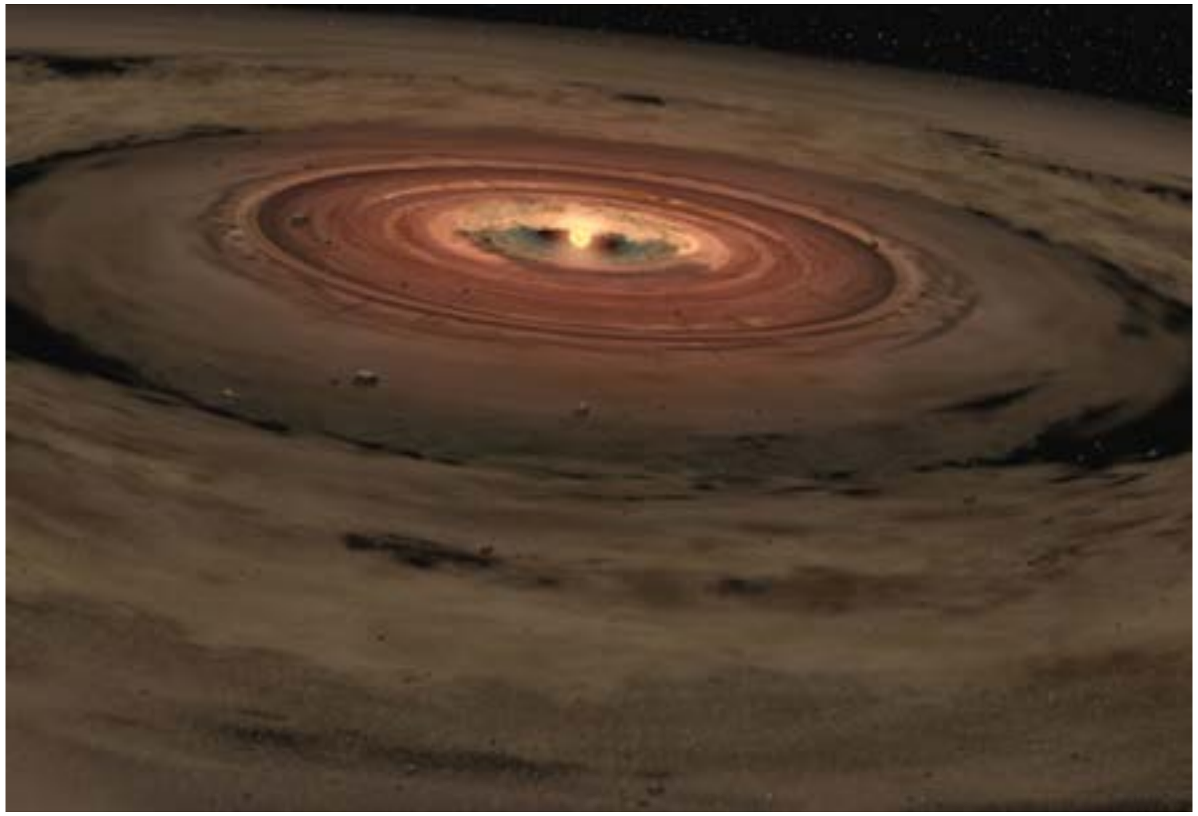


Cosmic Magnetism

- Magnetic fields are vital ingredients of many astrophysical phenomena
- … yet the origin, evolution and structure of cosmic magnetism is still unclear
- At SKA sensitivities, can be probed with very accurate radio polarization measurements



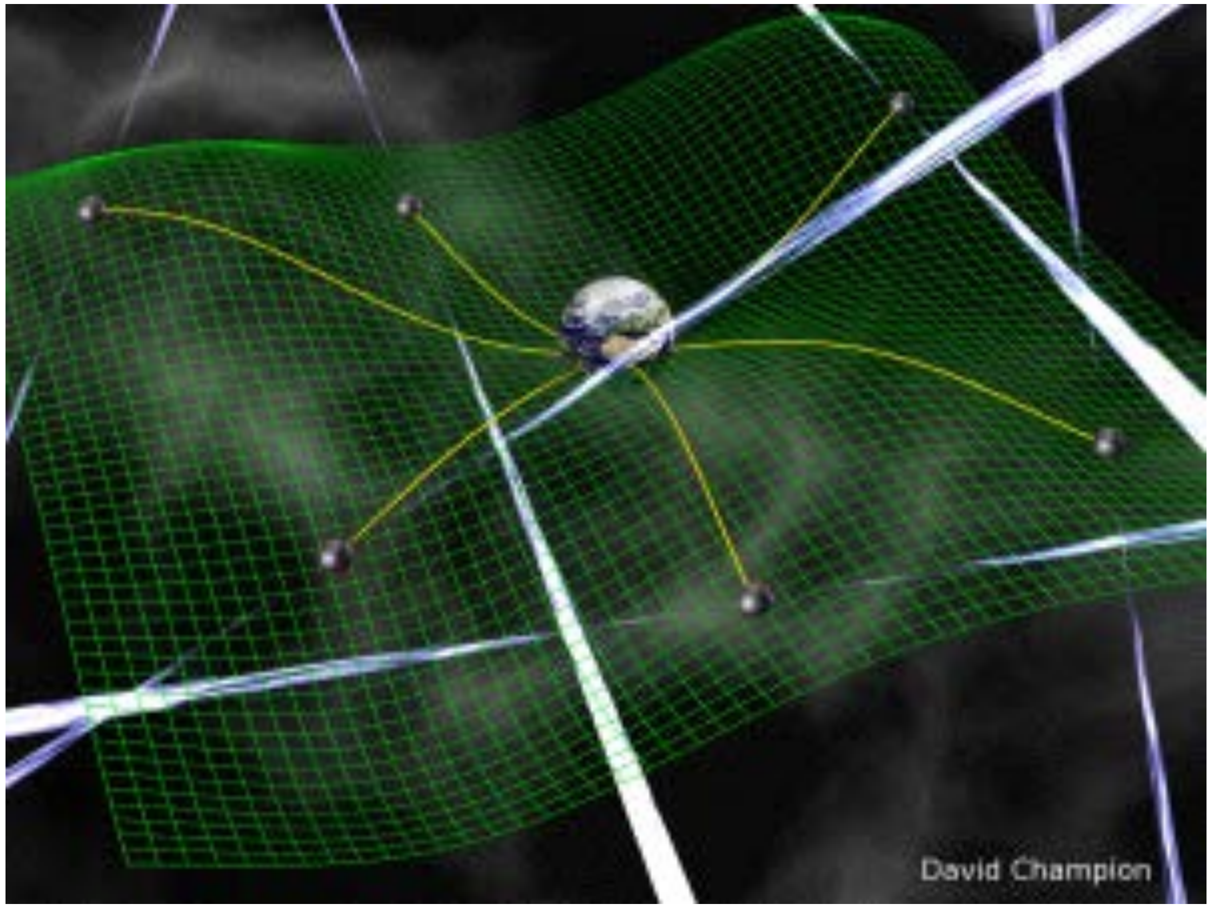
The Cradle Of Life



- The SKA will be able to image dust-obscured protoplanetary disks, thus directly observing the process of planet formation
- Astrobiology: detection of complex molecules, study of prebiotic chemical evolution in interstellar clouds

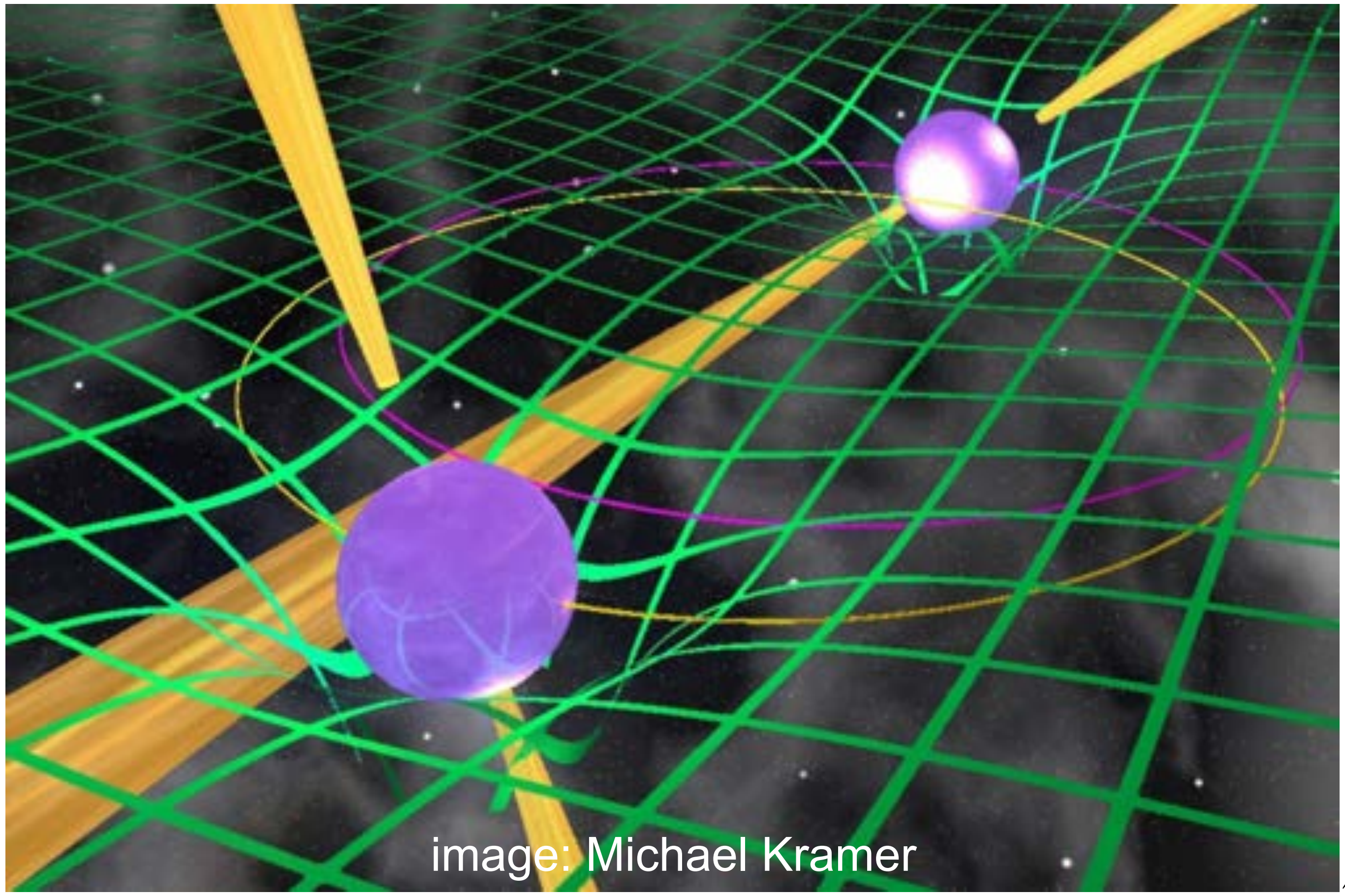
  - Sensitive enough to detect airport radar up to ~50 light years away

Gravitational Waves & Strong Field Tests



- A "timing array" of pulsars should allow for direct observation of gravitational waves

- Double pulsars and pulsars near black holes probe extreme gravitational regimes – if Einstein was wrong, this is where we find out.



Challenges Of The SKA

- Engineering: dishes, arrays, optical fibre, electric power...
- Computational
  - Data rate: 10-100 times today's <u>worldwide</u> internet traffic
  - Will require a future-world-class supercomputer
  - Only feasible if both Moore's Law holds out, **and** we make regular algorithmic advances...
- Mathematical
  - ...those very advances. Need better methods and better math!

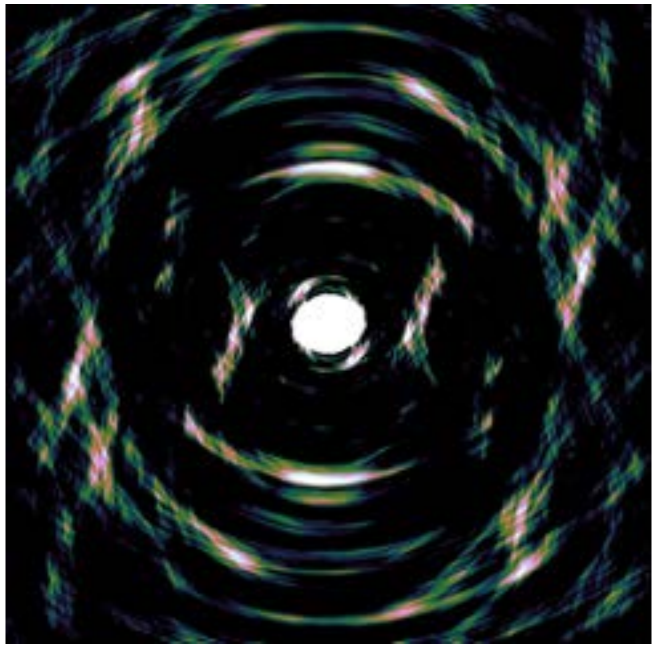
Why Is It Difficult?

- Gaps in the *uv*-plane:
  - At the end of the day, we're still trying to fill a huge "virtual aperture" with small dishes
  - Gaps in the *uv*-plane sampling are unavoidable
  - Fundamentally missing information
  - Can never recover the image fully
- Every measurement is distorted by the instrument response, needs to be calibrated.
  - And some measurements affected by man-made radio interference (RFI)
  - The "Fourier scramble": **one bad point** in the Fourier plane affects **all points** in the image plane

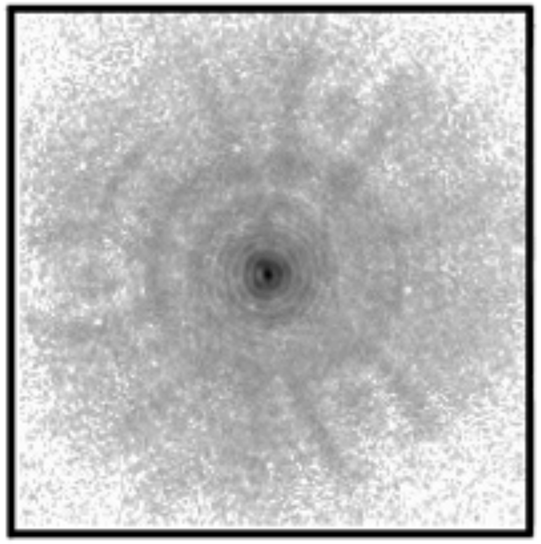
Gaps ⇒ Point Spread Function

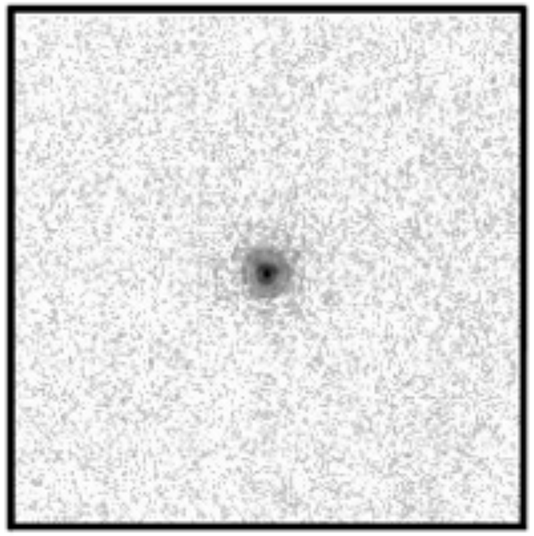
PSF of the WSRT.
The regular rings are due to the regular spacing of its antennas in the East-West direction.

- Response to a point source: Point Spread Function (PSF)
- Observed "dirty image" is *convolved* with the PSF
- Structure in the PSF = uncertainty in the flux distribution (corresponding to missing data in the *uv*-plane)



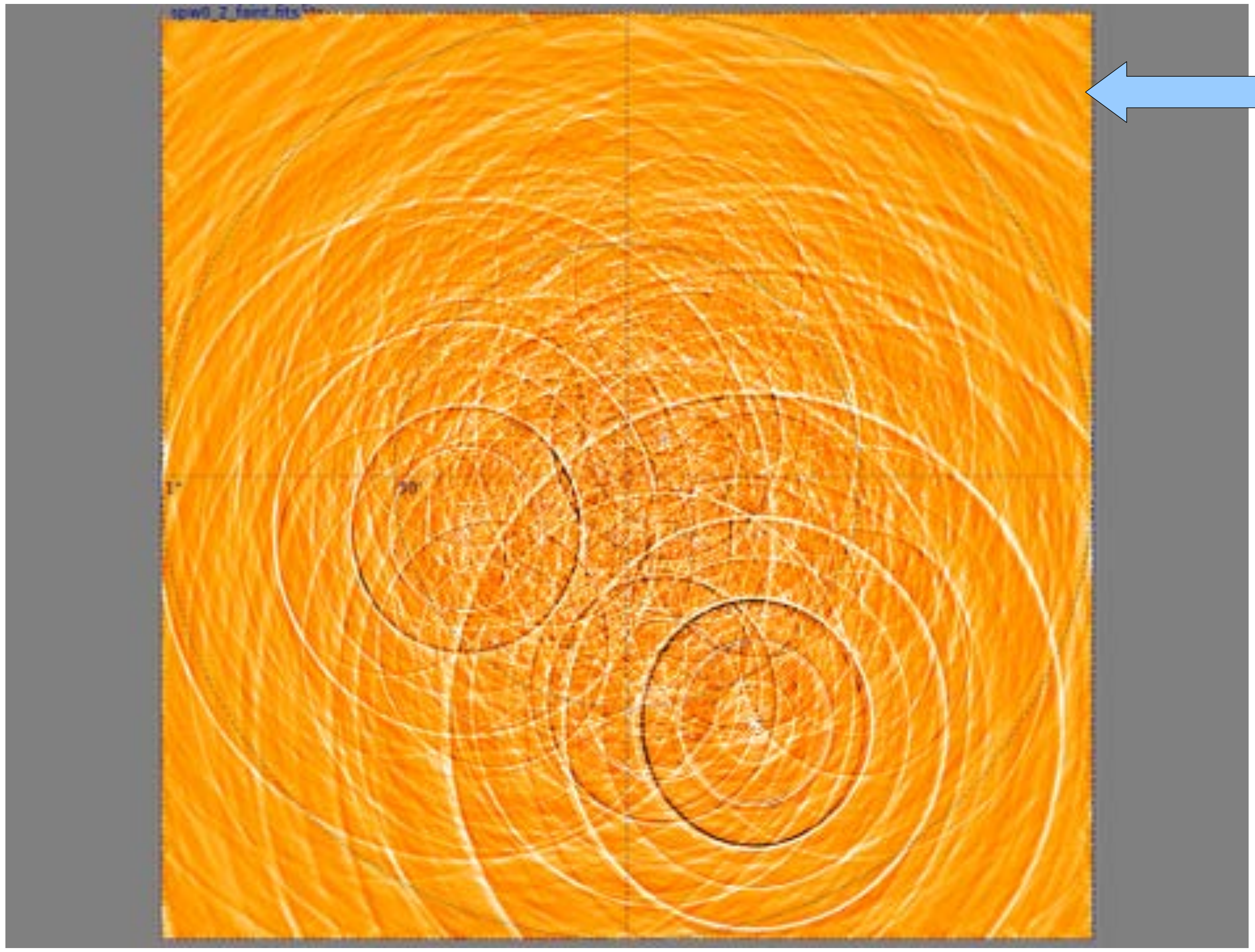
PSF of MeerKAT





PSF of the Hubble Space Telescope before & after servicing mission

Deconvolution: from dirty to clean images



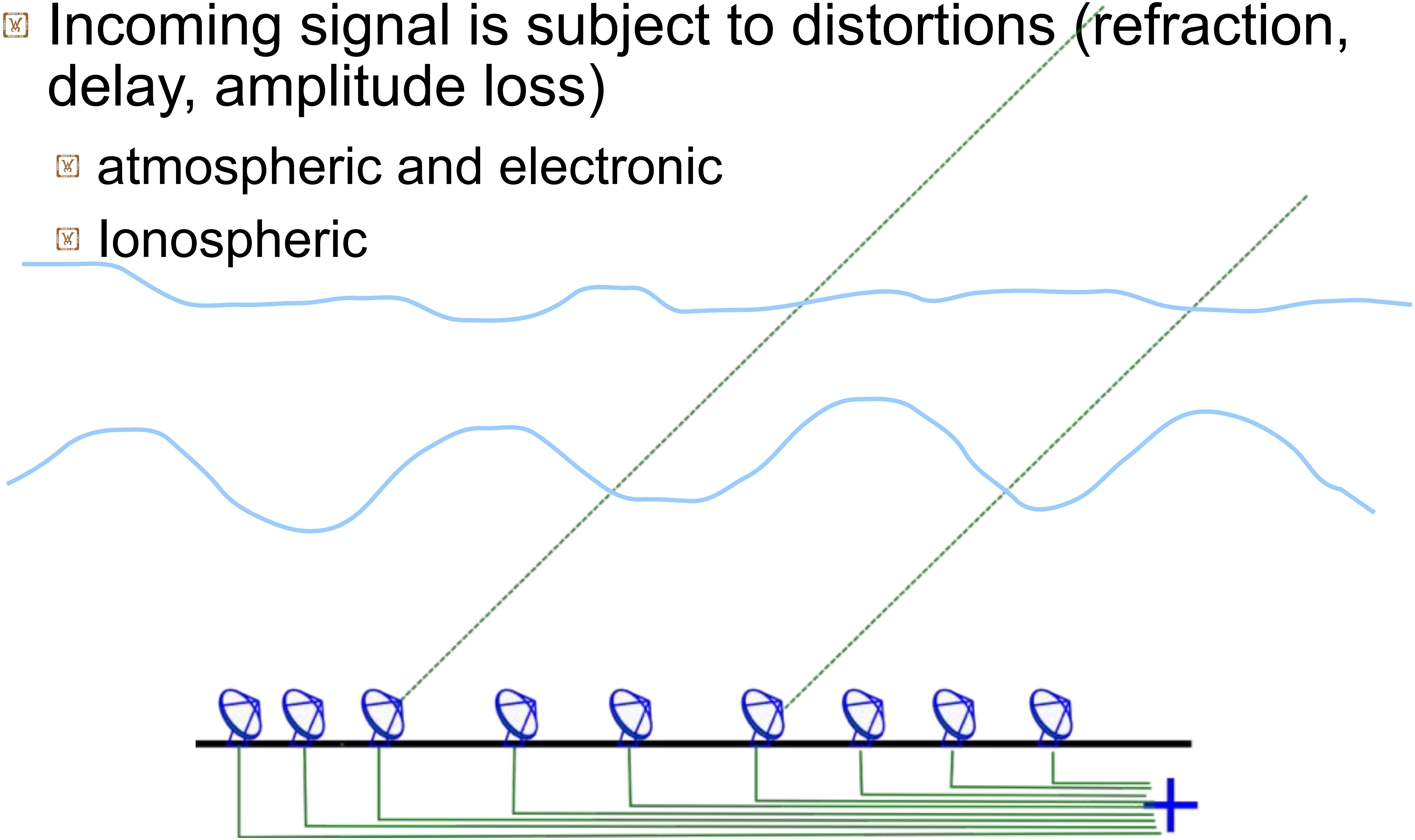
Real-life WSRT "dirty" image

- Dirty image dominated by PSF *sidelobes* from the strongest sources
- *Deconvolution* required to get at the faint stuff underneath.

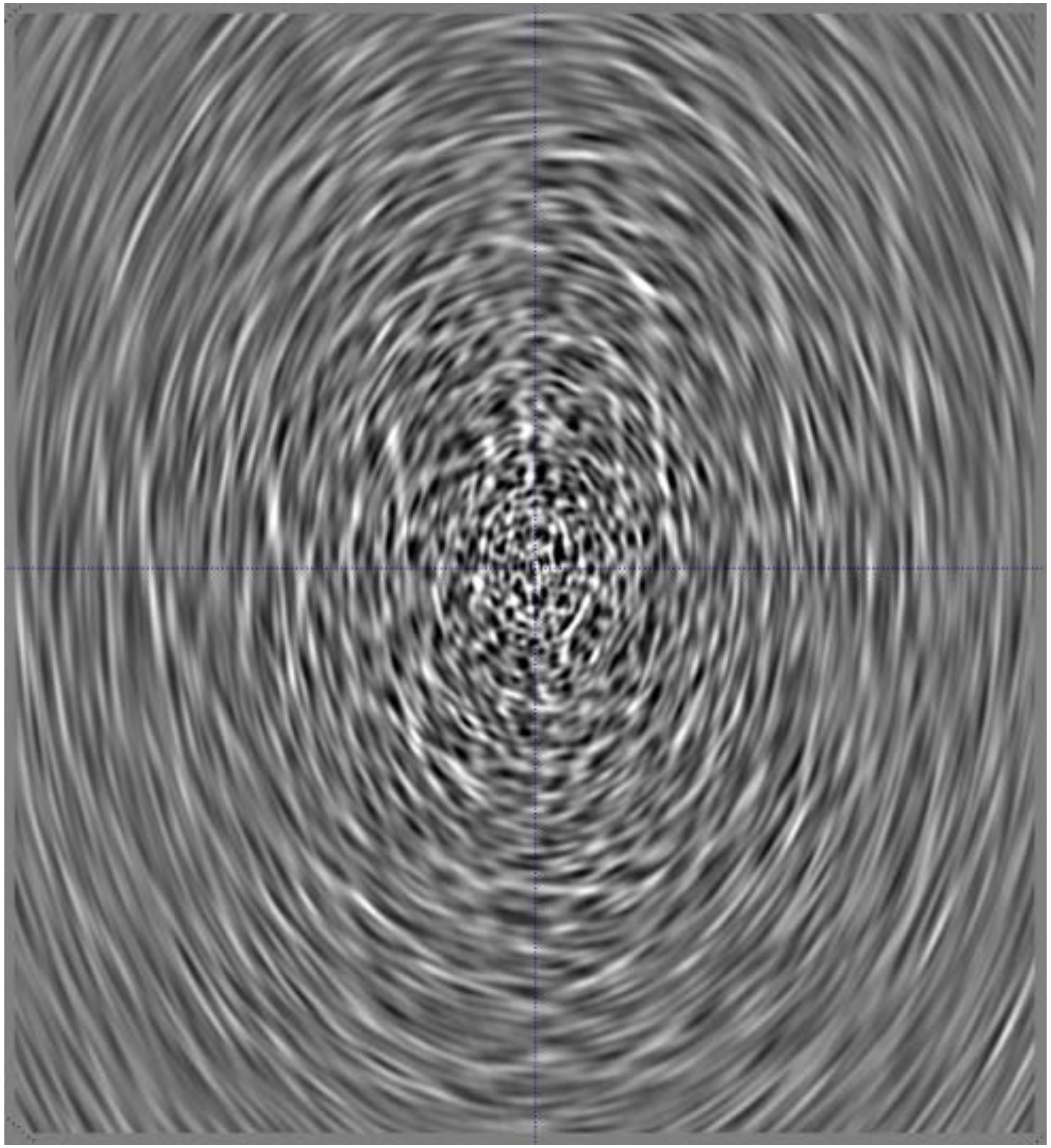
- A whole continuum of skies fits the dirty image (pick any value for the missing *uv*-components)
- Deconvolution picks one = interpolates the missing info from extra assumptions (e.g.: "sources are point-like").

Distorted Measurements

- Incoming signal is subject to distortions (refraction, delay, amplitude loss)
  - atmospheric and electronic
  - Ionospheric



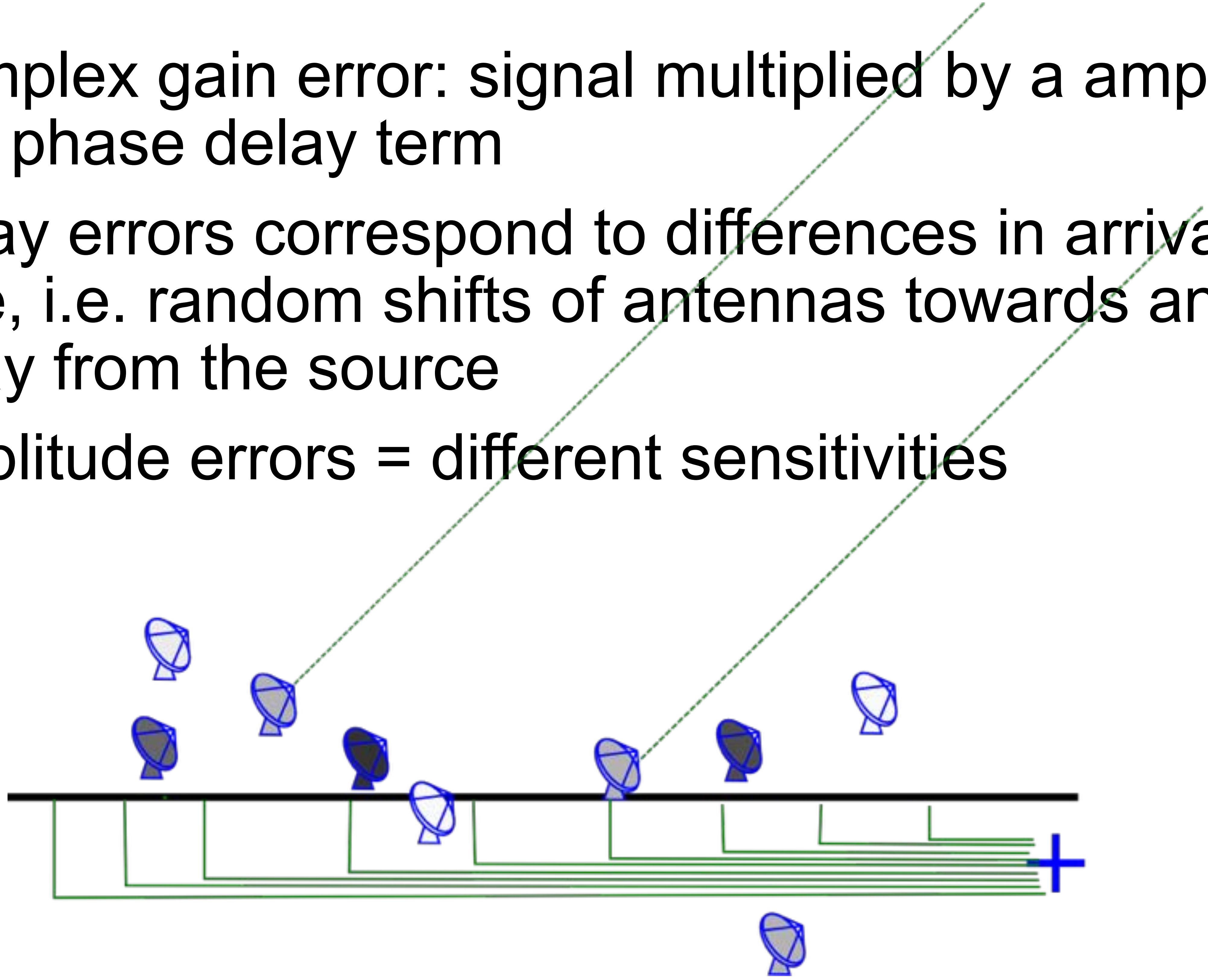
Distorted Measurements



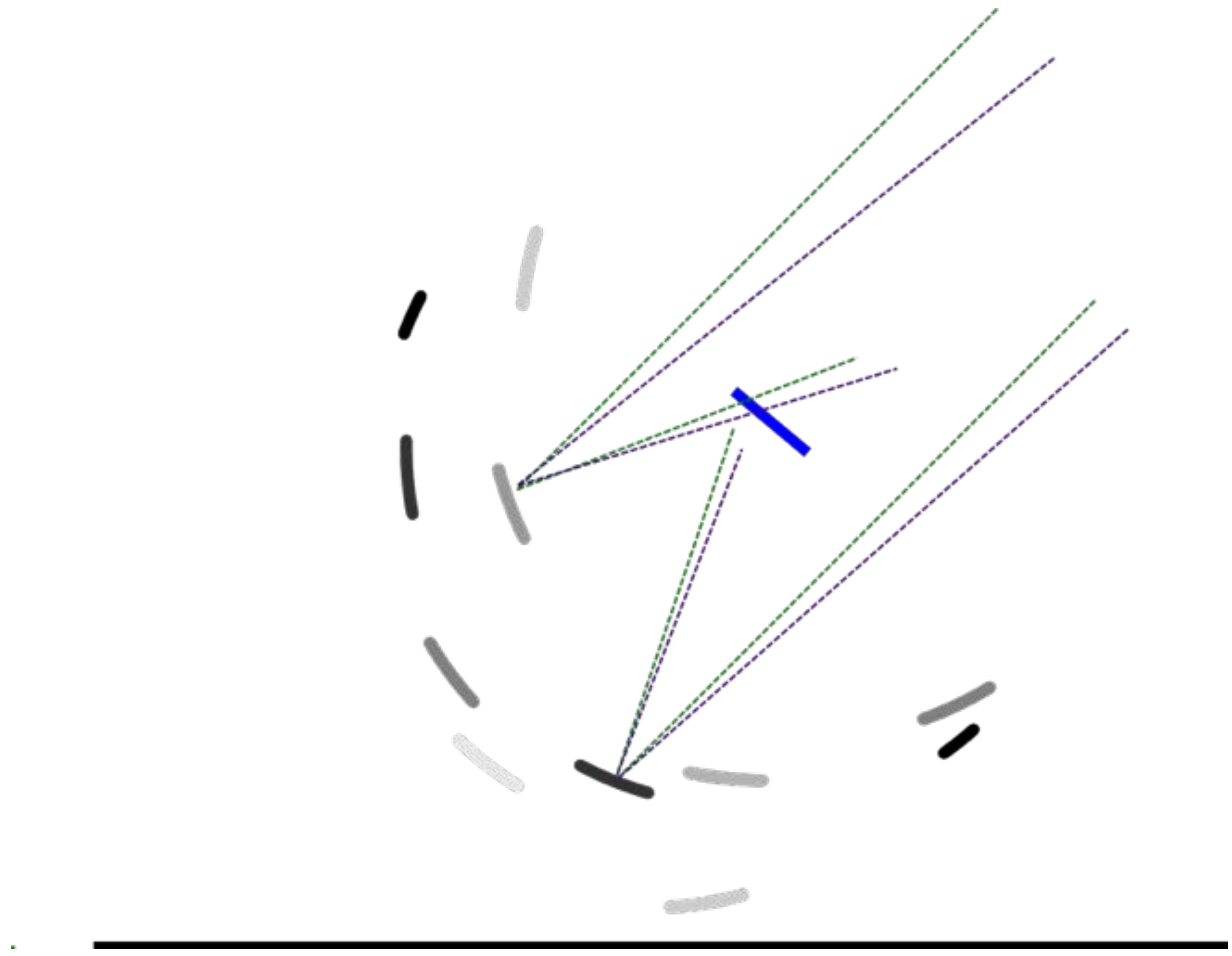
- One point-like source, but observed with complex phase errors
  - Ionosphere, troposphere, electronics
- In the *uv*-plane, phase encodes information about <u>location</u>
- Phase errors tends to spread the flux around
- Calibration of complex gains required before we can see anything at all!

Distorted Measurements

- Complex gain error: signal multiplied by a amplitude and phase delay term
- Delay errors correspond to differences in arrival time, i.e. random shifts of antennas towards and away from the source
- Amplitude errors = different sensitivities



Optical equivalent



Stone-Age Calibration
(First-Generation, or 1GC)

- Calibrate gains using a known calibrator source
- Move antennas to target, cross your fingers, and hope that everything stays stable enough to get an image

$$V_{pq} = g_{pq} M_{pq}$$

Observed data

Gain of baseline (i.e. antenna pair) *p-q*

Known calibrator model

- Dynamic range: ~100:1

1980: The Selfcal Revolution (2GC)

- A very simple realization: per-baseline gains are actually products of per-antenna complex gains!

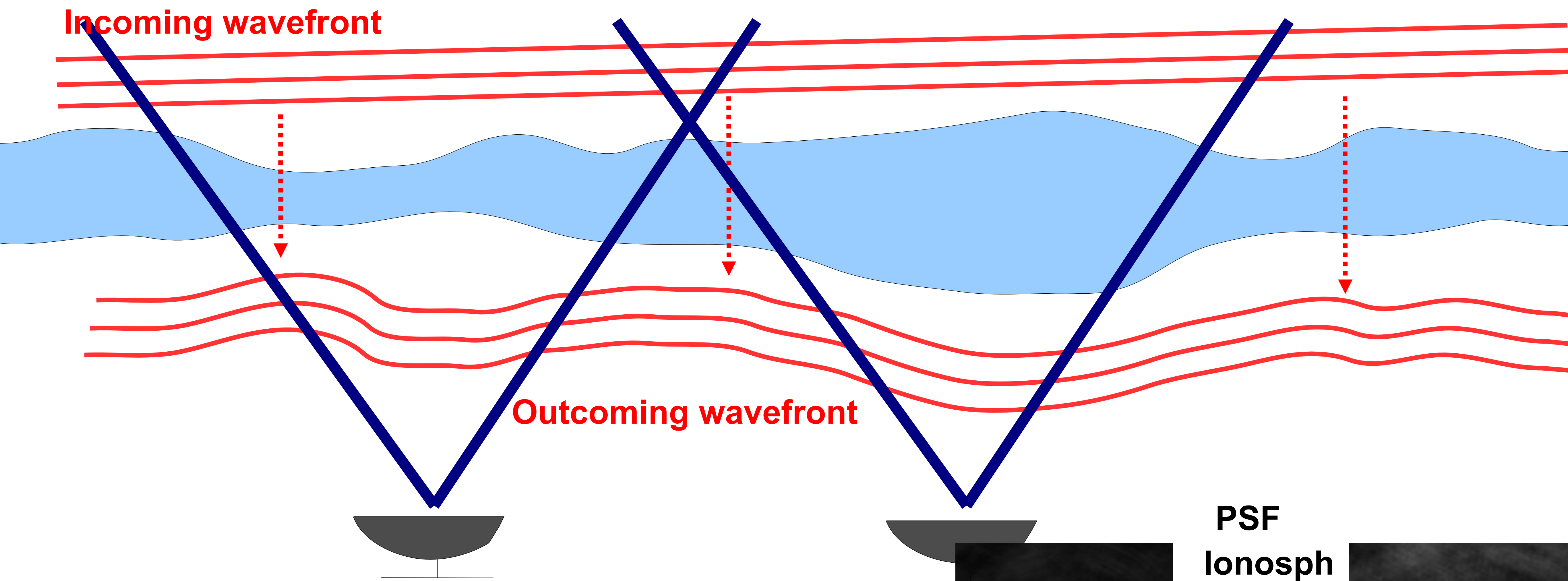
$$V_{pq} = g_p \bar{g}_q M_{pq}$$

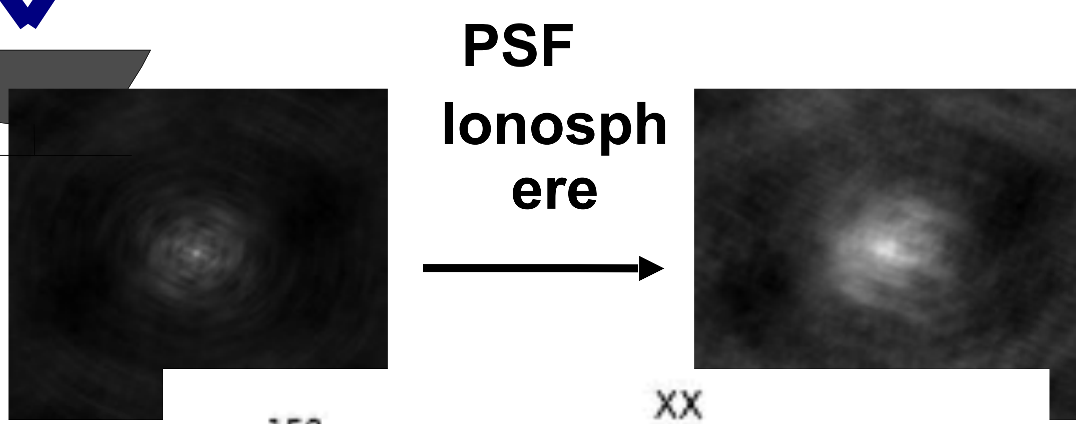
- $N(N-1)/2$ visibilities >> N gains
  - Start with simple M
  - Solve for g's
  - Improve M, rinse & repeat

dynamic range > 10^6:1

Huge body of experience suggests that this works rather well, **BUT** there's no formal proof (!!!) Current practice is a collection of *ad hoc* methods, dark art and lore passed down the generations in what is virtually an oral tradition.

… When Direction Dependent Effects (DDE) become a problem : Ionosphere

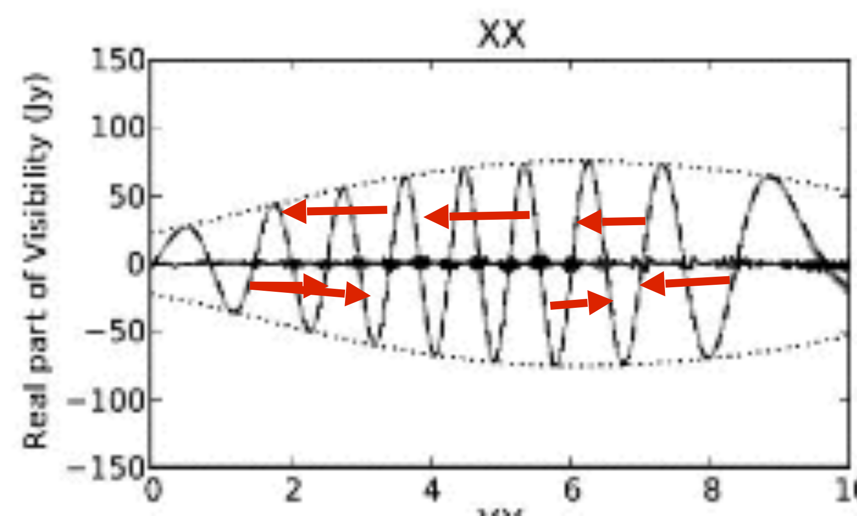




Big field of view : station, direction, time and frequency dependent

Other direction dependent effects :

- Frequency dependent phase shift
- Amplitude variation (sometimes)
- Faraday rotation
- **Effect on the polarisation**

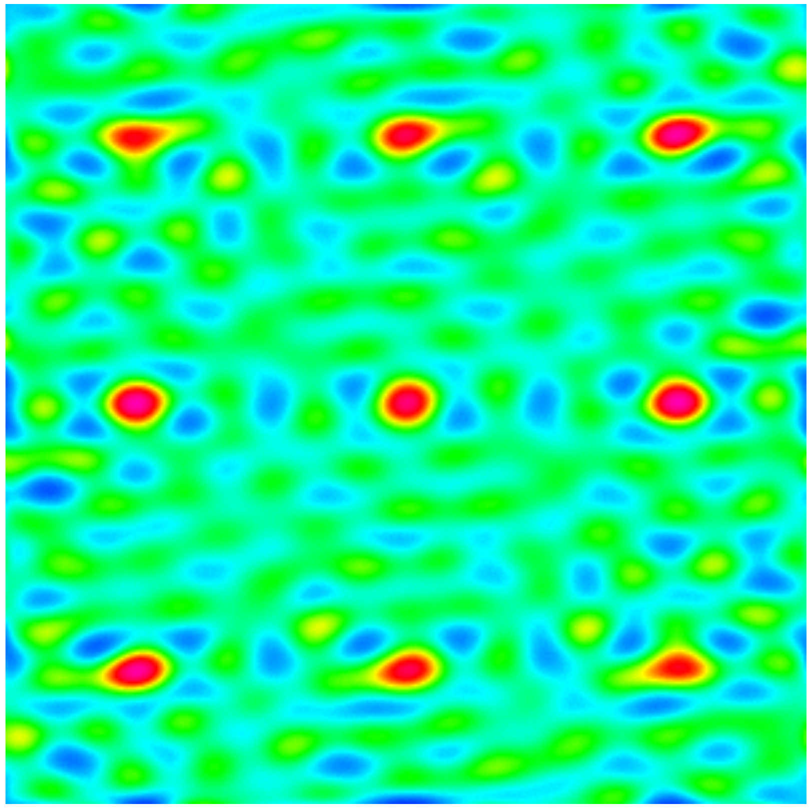


The Fundamental Limit of 2GC...

- The calibration equation assumes the unknown gains are direction-independent

$$V_{pq} = g_p g_q^* M_{pq}$$

  - ...when they are at best only approximately so
  - But incorporating directional dependence is devilishly tricky due to the nature of the Fourier transform
- Example: ionosphere
  - Introduces phase errors which appear to “shift” the sources
  - 2GC can calibrate the overall shift, but not the differential shifts

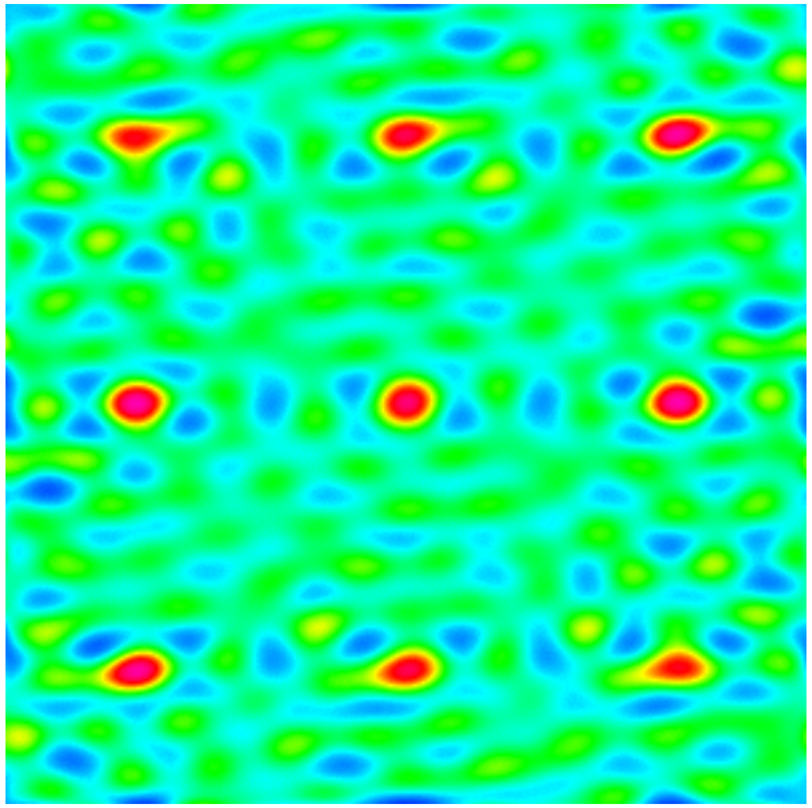


The Fundamental Limit of 2GC...

- The calibration equation assumes the unknown gains are direction-independent

$$V_{pq} = g_p g_q^* M_{pq}$$

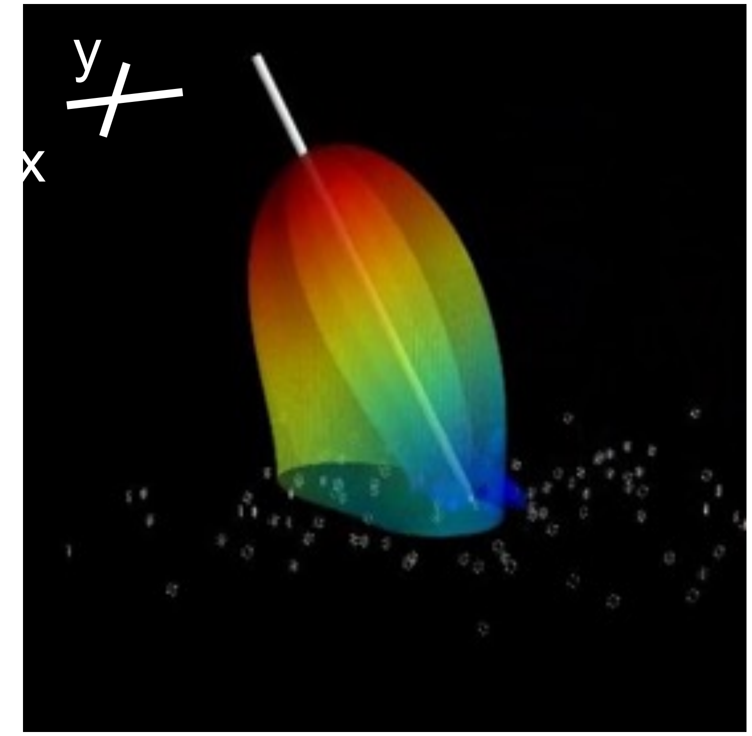
  - ...when they are at best only approximately so
  - But incorporating directional dependence is devilishly tricky due to the nature of the Fourier transform
- Example: ionosphere
  - Introduces phase errors which appear to “shift” the sources
  - 2GC can calibrate the overall shift, but not the differential shifts

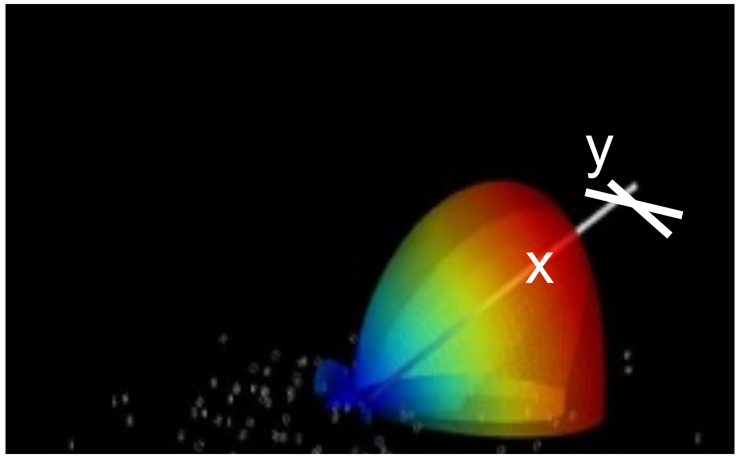


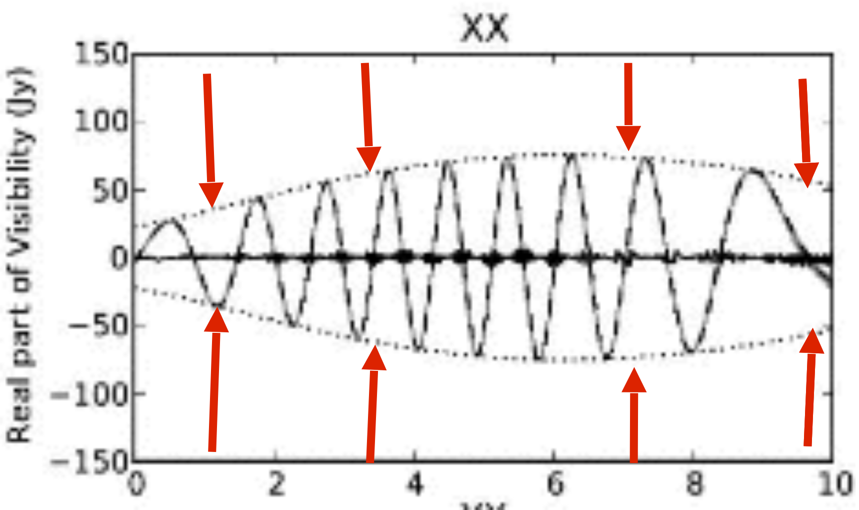
Why Is This a Problem Now?

- Older telescopes are also prone to direction-dependent effects
  - but if your sensitivity is not high enough to even see the distortions, then who cares?
- The SKA will have record sensitivity
  - Which means it will also be sensitive to far more subtle errors
- Also, it will be built from “cheap junk”....

… When Direction Dependent Effects (DDE) become a problem : Beam







LOFAR stations are phased arrays

- Beam is variable in frequency and time
- Projection of the dipoles in the sky is non trivial
- Beam can be station-dependent
- Individual clock effects

--> Strong effects on polarisation

The Measurement Equation

Hamaker et al. 1994

Mesurement (10^6-10^9 points)

$$\boxed{V_{pq}} = \underbrace{\boxed{G_p \left(\sum_{i=1}^{N} \mathrm{B}_{pi} \mathrm{K}_{pi} \mathrm{I}_{pi}}}_{\text{Calibration}} \underbrace{\boxed{F_i \, . \, F_i^{+}}}_{\text{Unknown Sky}} \underbrace{\boxed{\mathrm{I}_{qi}^{+} \mathrm{K}_{qi}^{+} \mathrm{B}_{qi}^{+} \right) G_q^{+}}}_{\text{Calibration}}$$

Calibration **Unknown Sky** **Calibration**

The Measurement Equation

Hamaker et al. 1994

Mesurement (10^6-10^9 points)

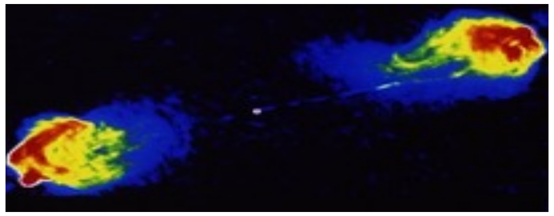
$$\boxed{V_{pq}} = G_p \left(\sum_{i=1}^{N} \mathrm{B}_{pi} \mathrm{K}_{pi} \mathrm{I}_{pi} F_i \,.\, F_i^{+} \mathrm{I}_{qi}^{+} \mathrm{K}_{qi}^{+} \mathrm{B}_{qi}^{+} \right) G_q^{+}$$

Calibration

Unknown Sky

Calibration

Electric field



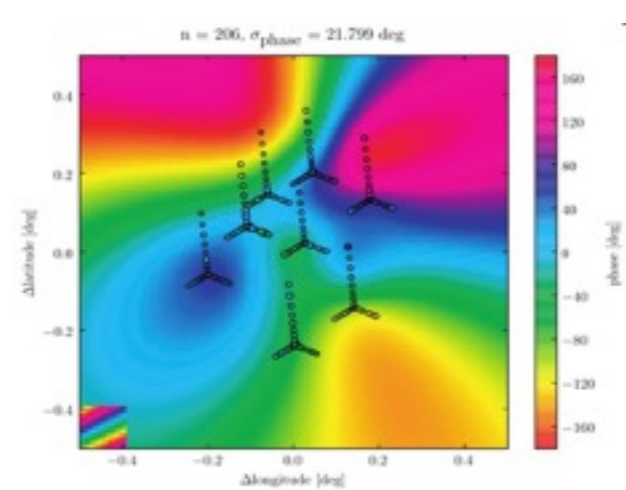
The Measurement Equation

Hamaker et al. 1994

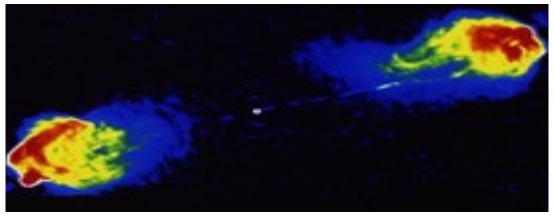
Mesurement (10^6-10^9 points)

$$\boxed{V_{pq}} = \underbrace{G_p \left(\sum_{i=1}^{N} \mathrm{B}_{pi} \mathrm{K}_{pi} \mathrm{I}_p \right.}_{\text{Calibration}} \underbrace{F_i \, . \, F_i^+}_{\text{Unknown Sky}} \underbrace{\left. \mathrm{I}_{qi}^+ \mathrm{K}_{qi}^+ \mathrm{B}_{qi}^+ \right) G_q^+}_{\text{Calibration}}$$

Ionosphere



Electric field



The Measurement Equation

Hamaker et al. 1994

Mesurement (10^6-10^9 points)

$$\boxed{V_{pq}} = \boxed{G_p \left(\sum_{i=1}^{N} \mathrm{B}_{pi} \mathrm{K}_{pi} \mathrm{I}_{p}} \boxed{F_i \,.\, F_i^{+}} \boxed{\mathrm{I}_{qi}^{+} \mathrm{K}_{qi}^{+} \mathrm{B}_{qi}^{+} \right) G_q^{+}}$$

Calibration

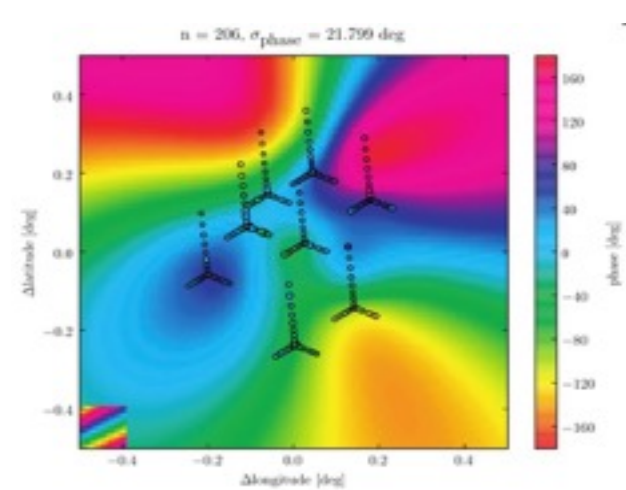
Unknown Sky

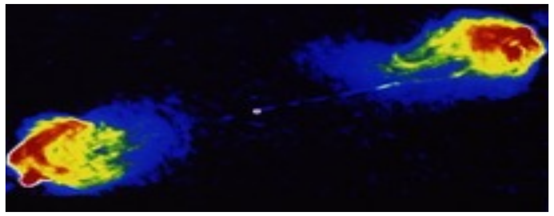
Calibration

Geometrical delay +Correlator

Ionosphere

Electric field





$$\mathrm{K}_p \mathrm{K}_q^{+} = \exp(-2i\pi\phi_{pq}) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$\phi_{pq} = u_{pq} l + v_{pq} m + w_{pq}(\sqrt{1 - l^2 - m^2} - 1)$$

The Measurement Equation

Hamaker et al. 1994

Mesurement (10^6-10^9 points)

$$V_{pq} = G_p \left(\sum_{i=1}^{N} \mathrm{B}_{pi} \mathrm{K}_{pi} \mathrm{I}_p F_i \;.\; F_i^{+} \mathrm{I}_{qi}^{+} \mathrm{K}_{qi}^{+} \mathrm{B}_{qi}^{+} \right) G_q^{+}$$

Calibration

Unknown Sky

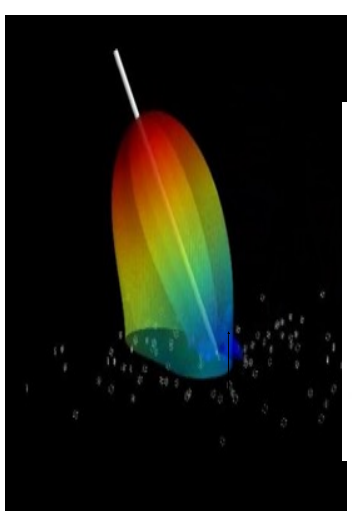
Calibration

Beam

Geometrical delay +Correlator

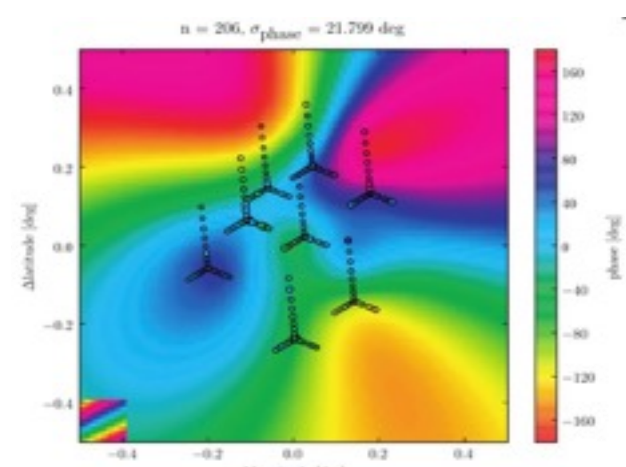
Ionosphere

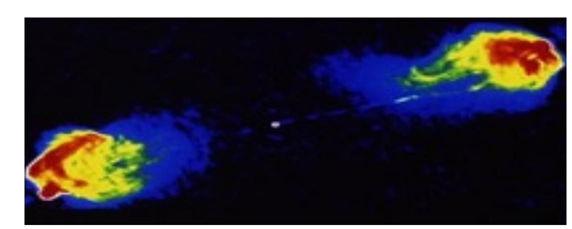
Electric field



$$\mathrm{K}_p \mathrm{K}_q^{+} = \exp(-2i\pi\phi_{pq}) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$\phi_{pq} = u_{pq} l + v_{pq} m + w_{pq}\left(\sqrt{1 - l^2 - m^2} - 1\right)$$





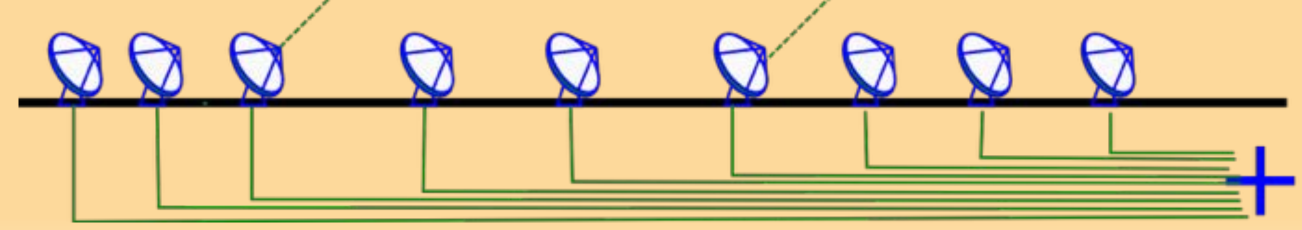
Radio Interferometry 101

- One *station* (i.e. dish or aperture array) measures two complex voltages

$$\vec{v} = \begin{pmatrix} v_x \\ v_y \end{pmatrix}$$

- A correlator computes the complex visibility for each pair of stations *p,q*:

$$V_{pq} = \langle \vec{v}_p \vec{v}_q^H \rangle = \langle \begin{pmatrix} v_{px}v_{qx}^* & v_{px}v_{qy}^* \\ v_{py}v_{qx}^* & v_{py}v_{qy}^* \end{pmatrix} \rangle$$

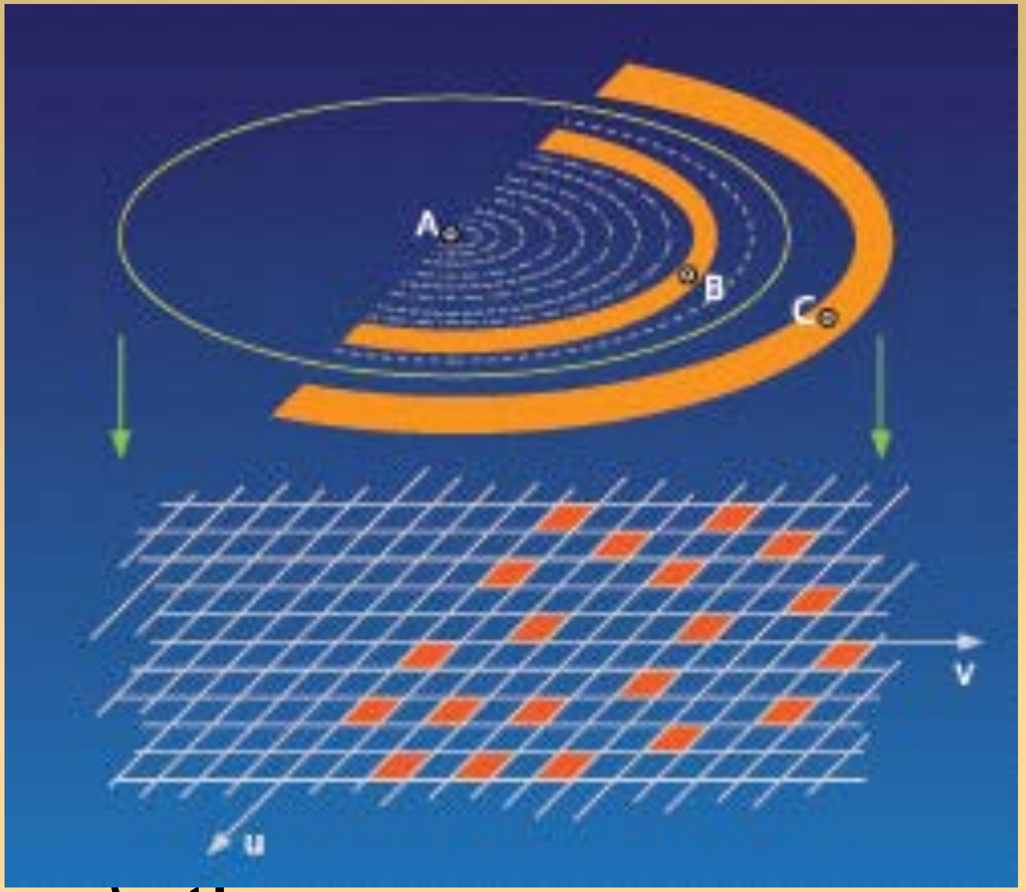


Radio Interferometry 102

- Complex visibility = Fourier component
- Thus, each pair of stations *p,q* measures one point in the Fourier plane (the *uv*-plane), corresponding to the separation between the stations (baseline vector) $\vec{u}_{pq}$
- With *N* stations, *N(N-1)/2* baselines can be measured instantaneously
  - WSRT, 14 stations: 91 baselines
  - LOFAR, 44 stations: 946 baselines
  - SKA, 2500 stations: >3 million baselines

Convolutional Gridding & Imaging

- *uv*-points are measured along elliptical tracks
- Traditional imaging performs a gridding step to resample these onto a regular grid (for the FFT):
  $V_{grid} = V \circ G$



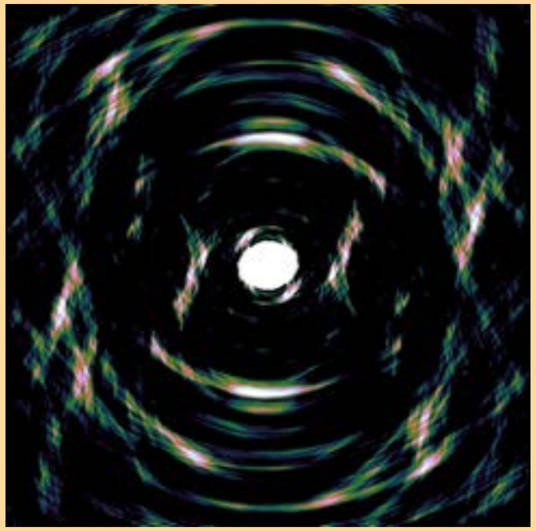
- If we know the DDE distribution $E_p(l,m)$, then we can correct for it via some extra convolutions:

$$V_{grid} = F_p \circ V \circ F_q^H \circ G \qquad (F_p \xleftrightarrow{FT} E_p^{-1})$$

Point Spread Functions

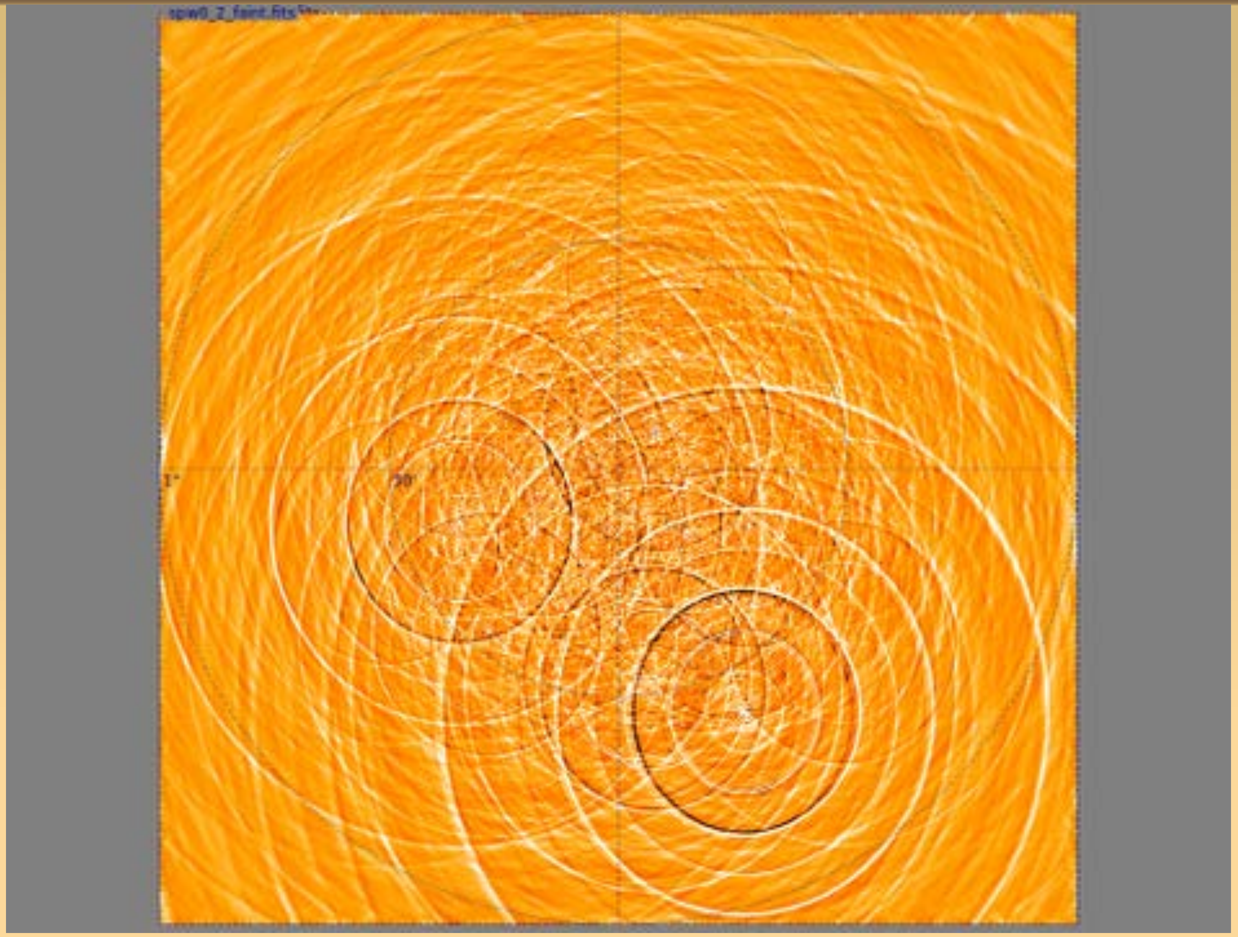
PSF of the WSRT.
The regular rings are due to the regular spacing of its antennas in the East-West direction.

- Response to a point source: Point Spread Function (PSF)
- Observed "dirty image" is *convolved* with the PSF
- Structure in the PSF = uncertainty in the flux distribution (corresponding to missing data in the *uv*-plane)
- Strong near-in PSF sidelobes; far sidelobes that do not go to 0 with distance



PSF of MeerKAT

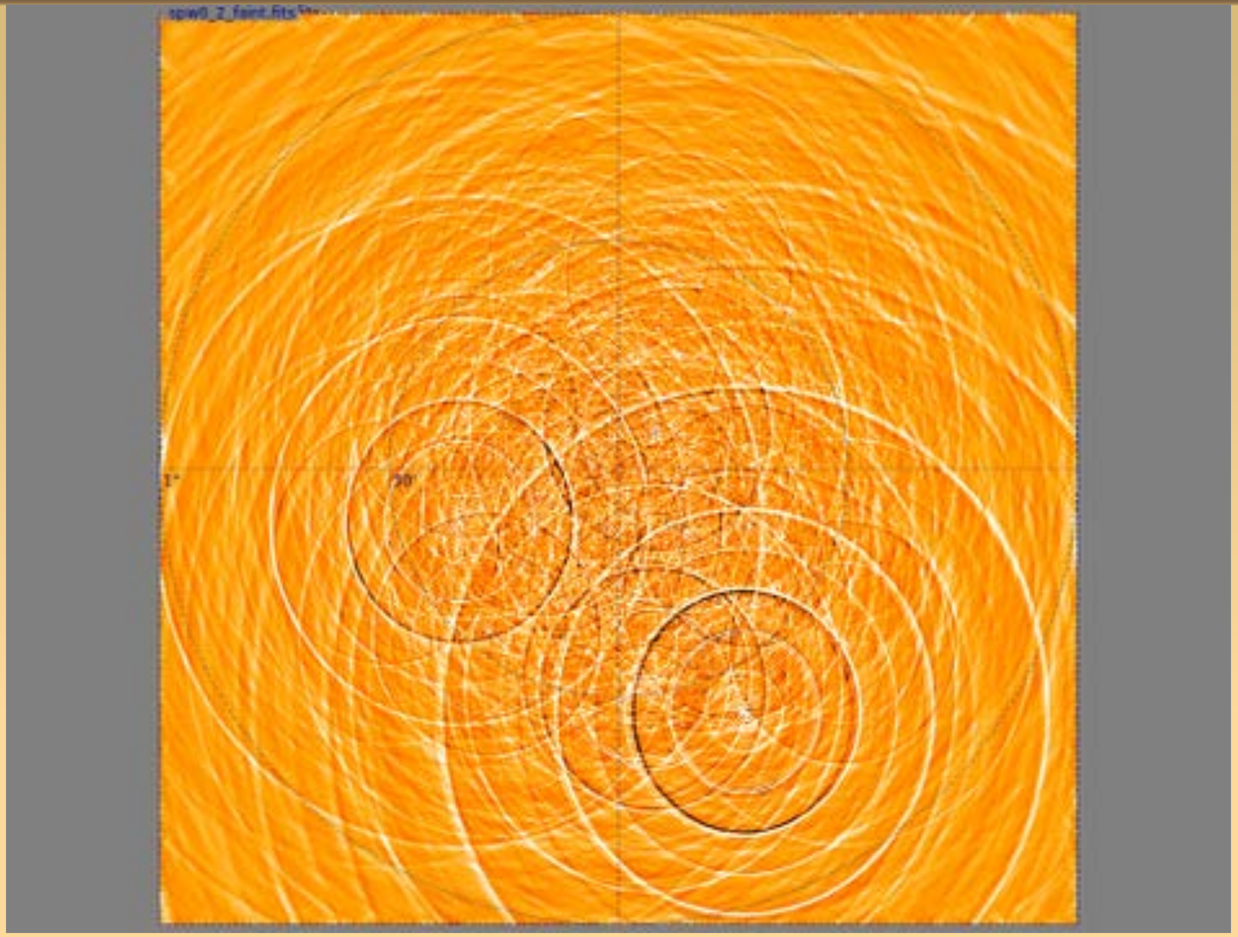
The Problem With Sidelobes



- "Dirty image" dominated by PSF sidelobes of the brightest sources
- Deconvolution is fundamentally ill-posed
  - But sky is sparse
- CLEAN, MEM, etc.
- Promising new methods are being proposed

The Problem With Sidelobes

- "Dirty image" dominated by PSF sidelobes of the brightest sources
- Deconvolution is fundamentally ill-posed
  - But sky is sparse
- CLEAN, MEM, etc.



- Promising new methods are being proposed

  Bayesianally sparse (BS) techniques

Apparent Skies

- “Classic” interferometry:
  - sky $B(l,m) \xleftrightarrow{FT}$ visibilities $V(u,v)$
- Interferometry in the presence of DDEs:
  - sky $B(l,m)$ + per-station effect $E_p(l,m)$
  - each station pair sees an “apparent sky”

$$B_{pq}(l,m) = E_p(l,m)\, B(l,m)\, E_q^H(l,m) \xleftrightarrow{FT} V_{pq}(u,v)$$

  - each visibility point sampled from a different “apparent *uv*-plane”

Apparent Skies

- "Classic" interferometry:
  - sky $B(l,m) \xleftrightarrow{FT}$ visibilities $V(u,v)$
- Interferometry in the presence of DDEs:
  - sky $B(l,m)$ + per-station effect $E_p(l,m)$
  - each station pair sees an "apparent sky"

(The Devil In) *E*-Jones

$$B_{pq}(l,m) = E_p(l,m)\, B(l,m)\, E_q^H(l,m) \;\xleftrightarrow{FT}\; V_{pq}(u,v)$$

  - each visibility point sampled from a different "apparent *uv*-plane"

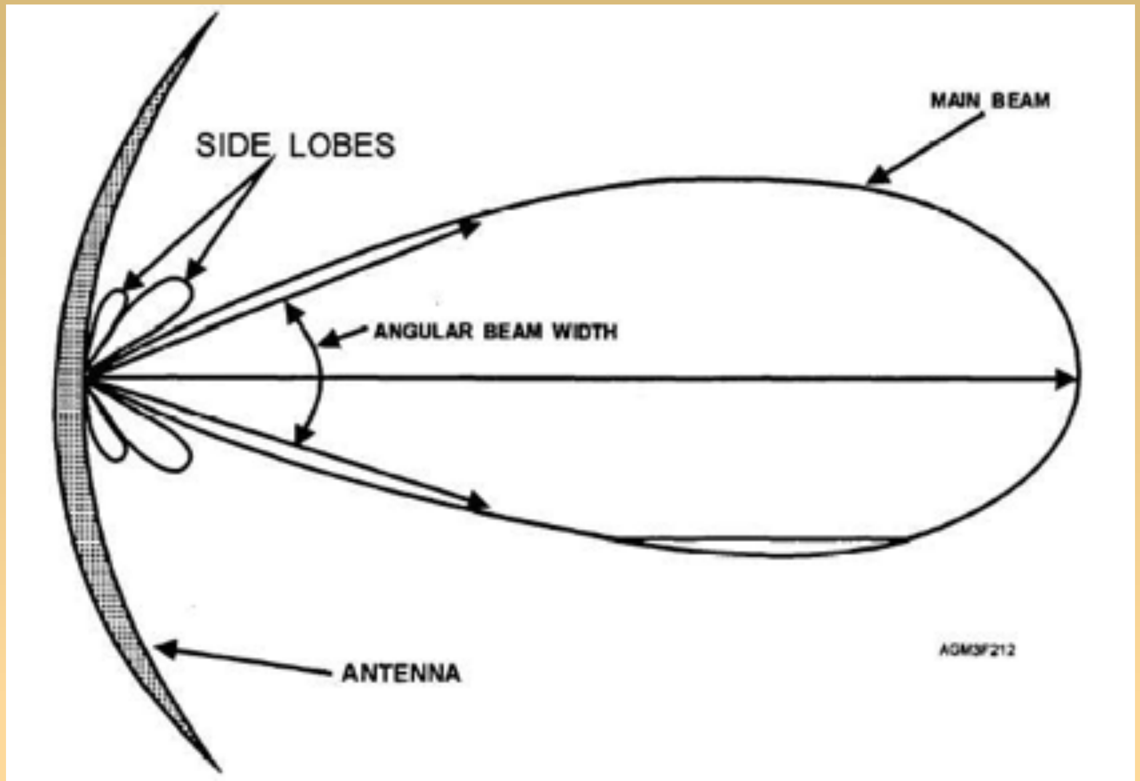
AW-Projection

- Application of DDEs via convolutional gridding forms the basis of the AW-projection algorithm
  - (Initially formulated as *w*-projection for correcting the *w*-term)
  - Current state-of-the-art
- Problems:
  - Need to know the DDE (calibration!)
  - If support of F_p is extended, convolution is expensive

Primary Beam

Primary Beams

- Field of view of an interferometer is determined by the main lobe of the primary beam (PB) pattern
- PB sidelobes are unavoidable, pick up unwanted signal
- Interferometer measures the sky multiplied by the square of the PB
  - ...assuming stable and identical beams!

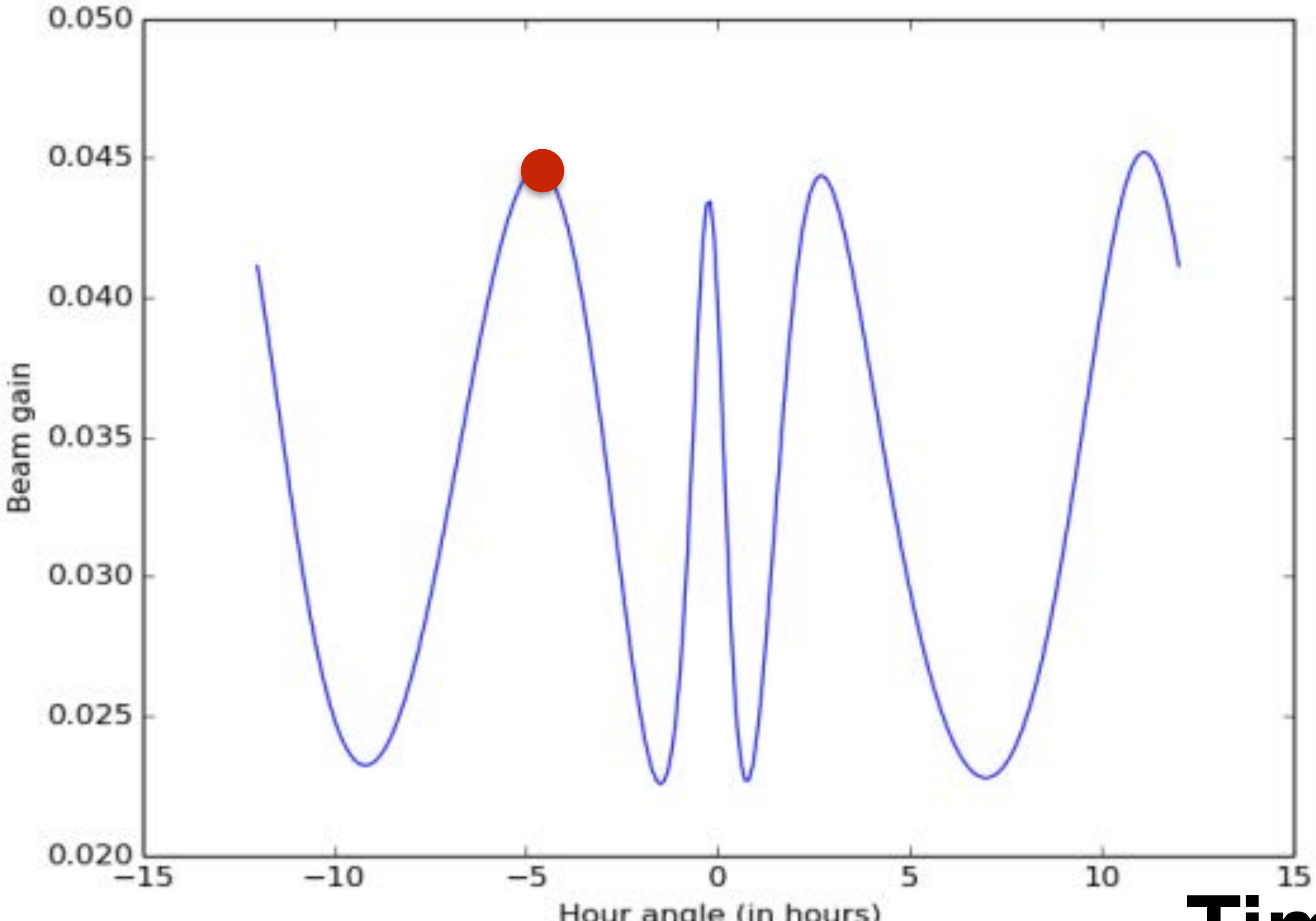


Primary Beam Problems

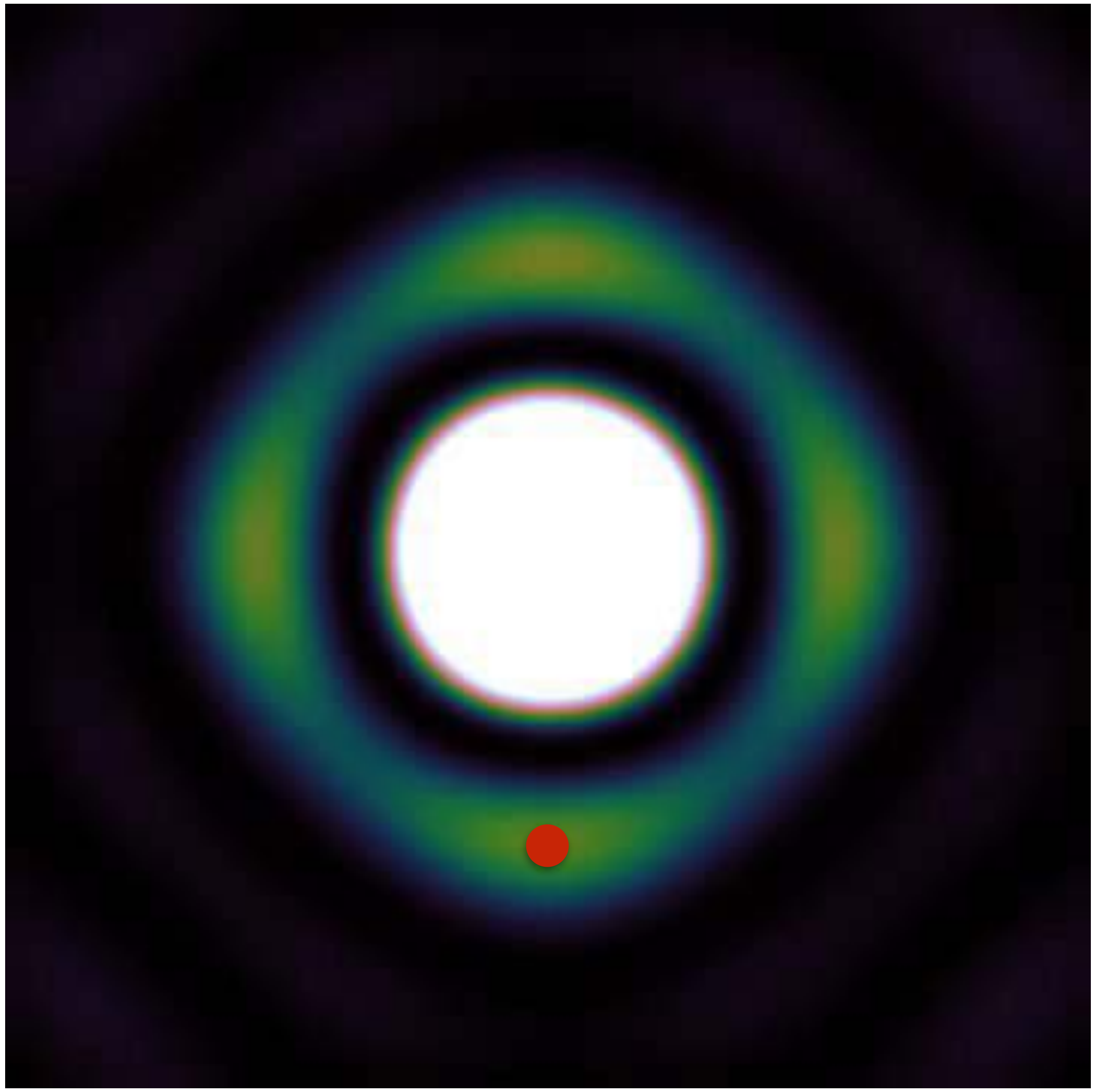
- PB's stable and identical only to first order
  - Pointing errors
  - Mechanical deformations
  - Beamformer gain drifts (for AAs/PAFs)
- PB variations = change in apparent source phase and amplitude, per *uv*-point
- FT translates this into spatial artefacts
  - Can be attenuated via calibration
- Plus, PB sidelobes pick up unwanted junk

Primary beam rotation with time

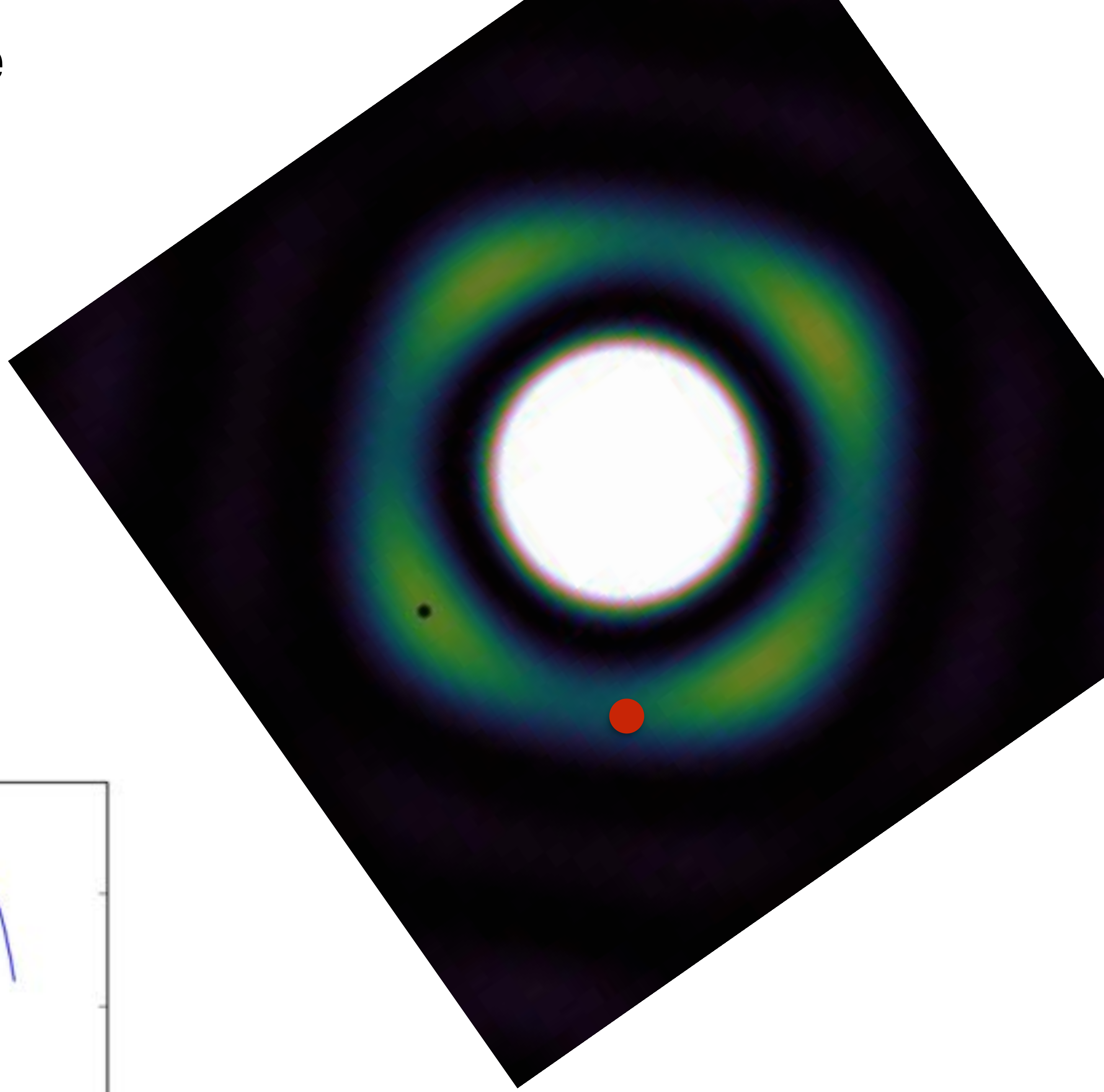
Gain



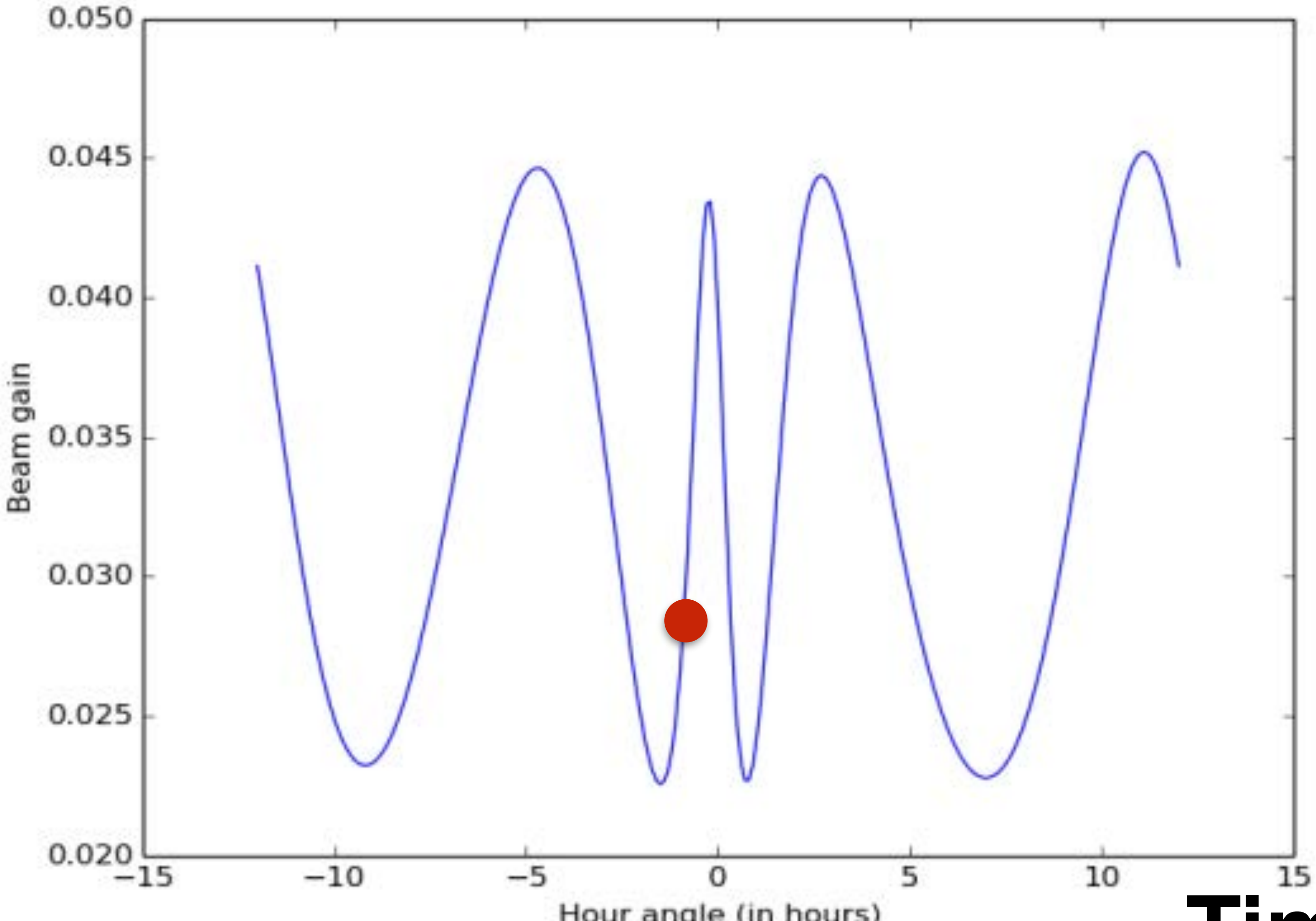
Time



Primary beam rotation with time



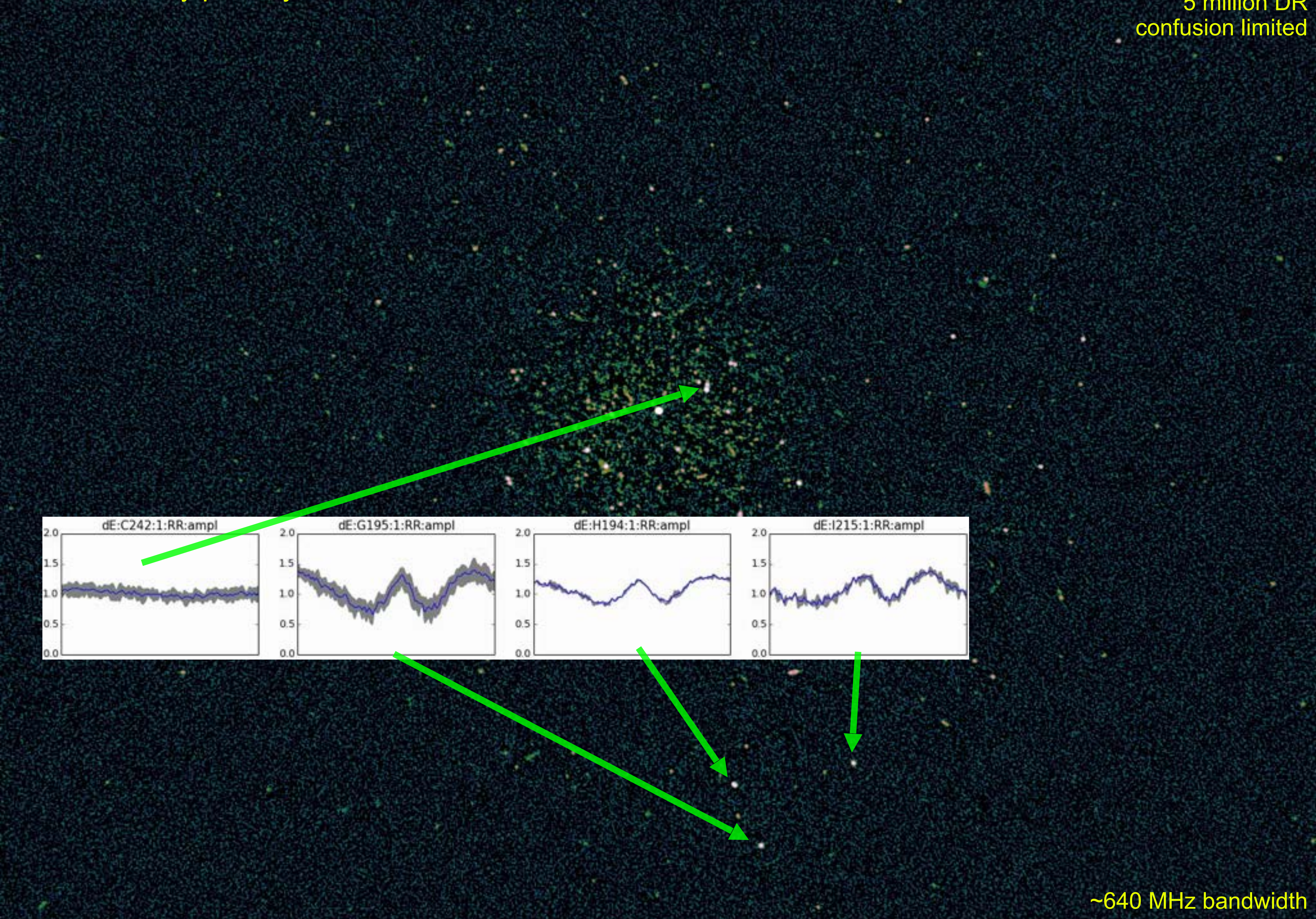
Gain



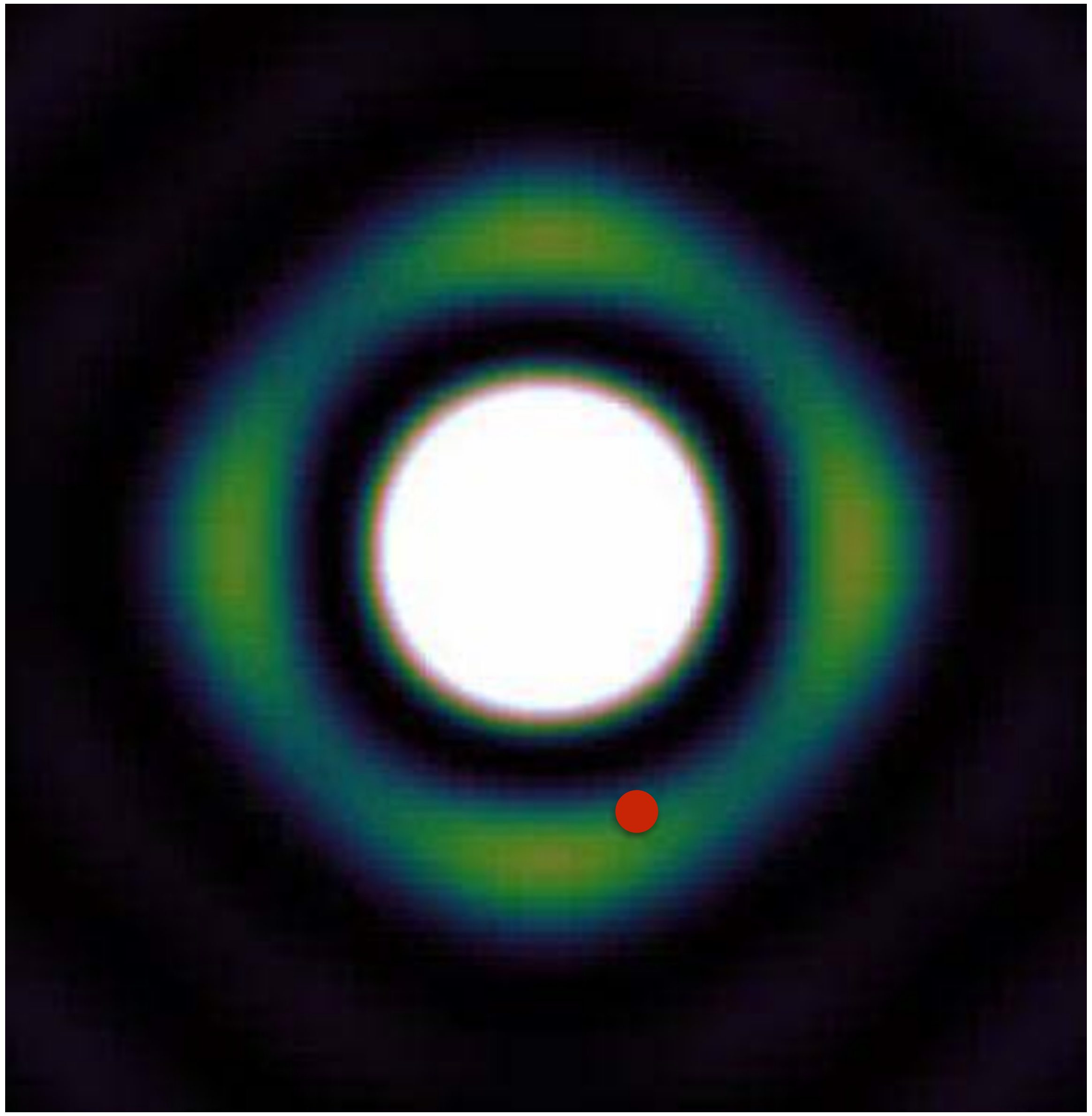
Time

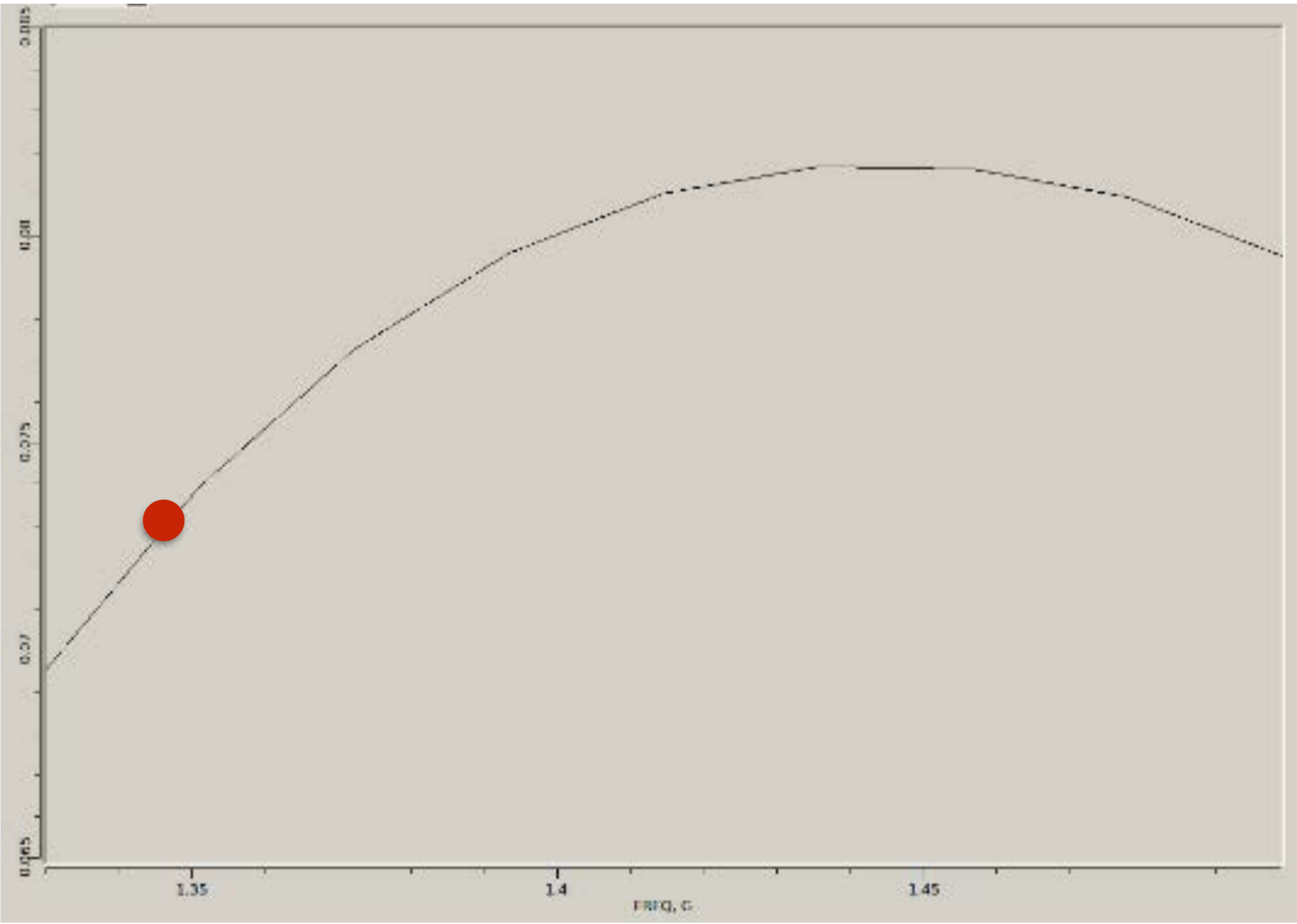
anim-beam.avi

Example differential gain solutions for JVLA (2014 image)
Dominated by primary beam rotation



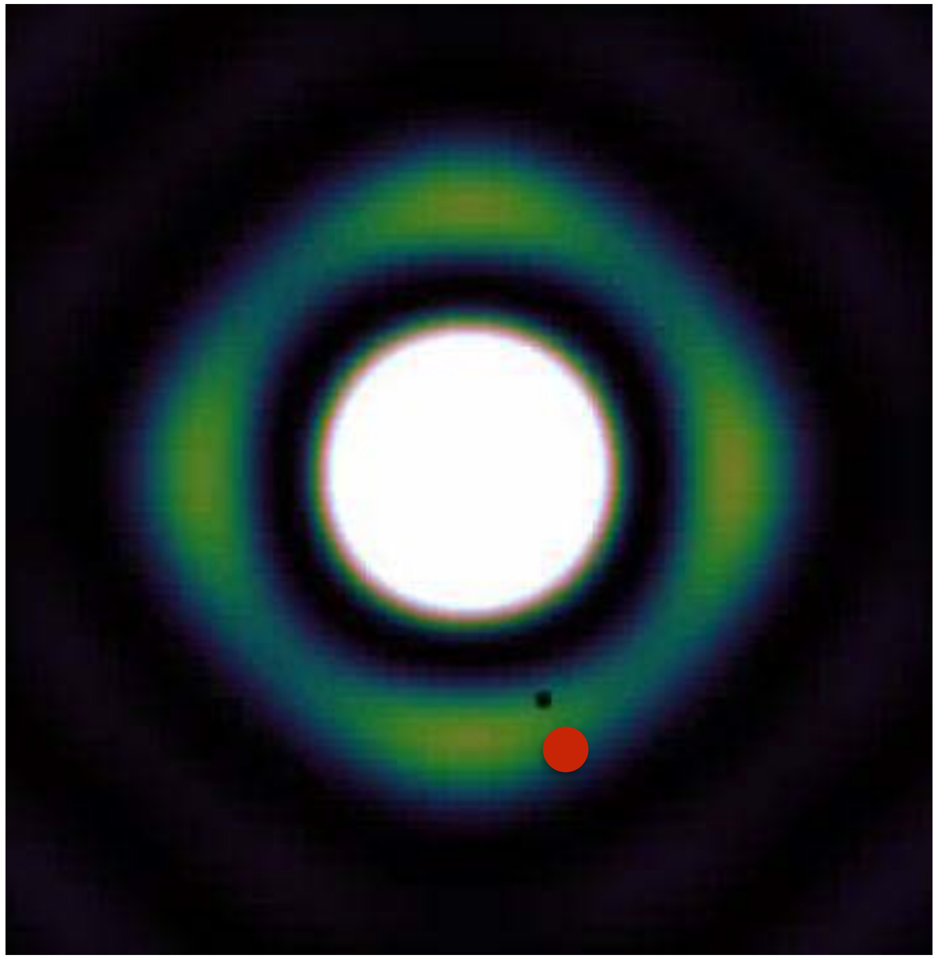
Primary beam scaling with frequency

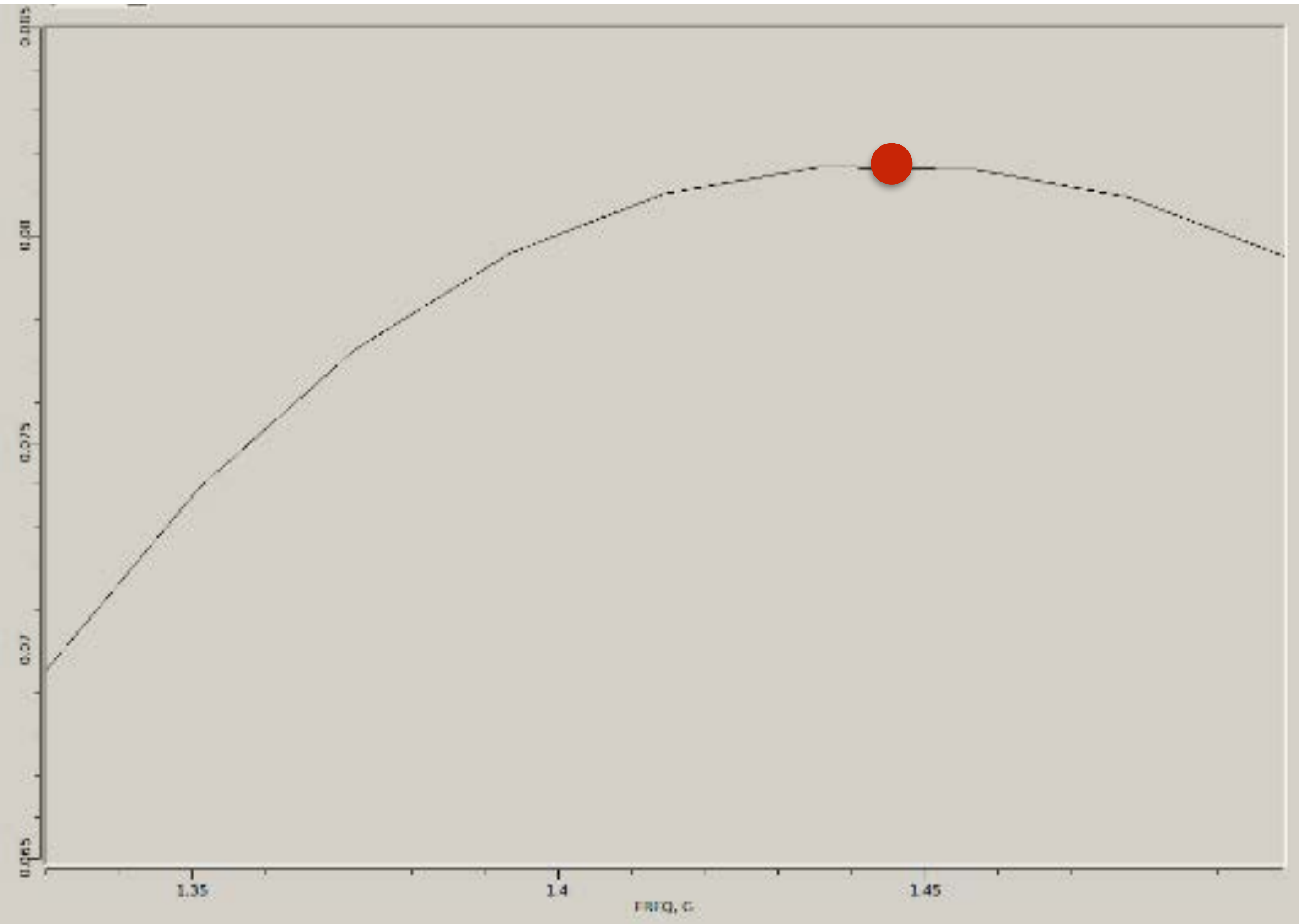




Freq /^

Primary beam scaling with frequency

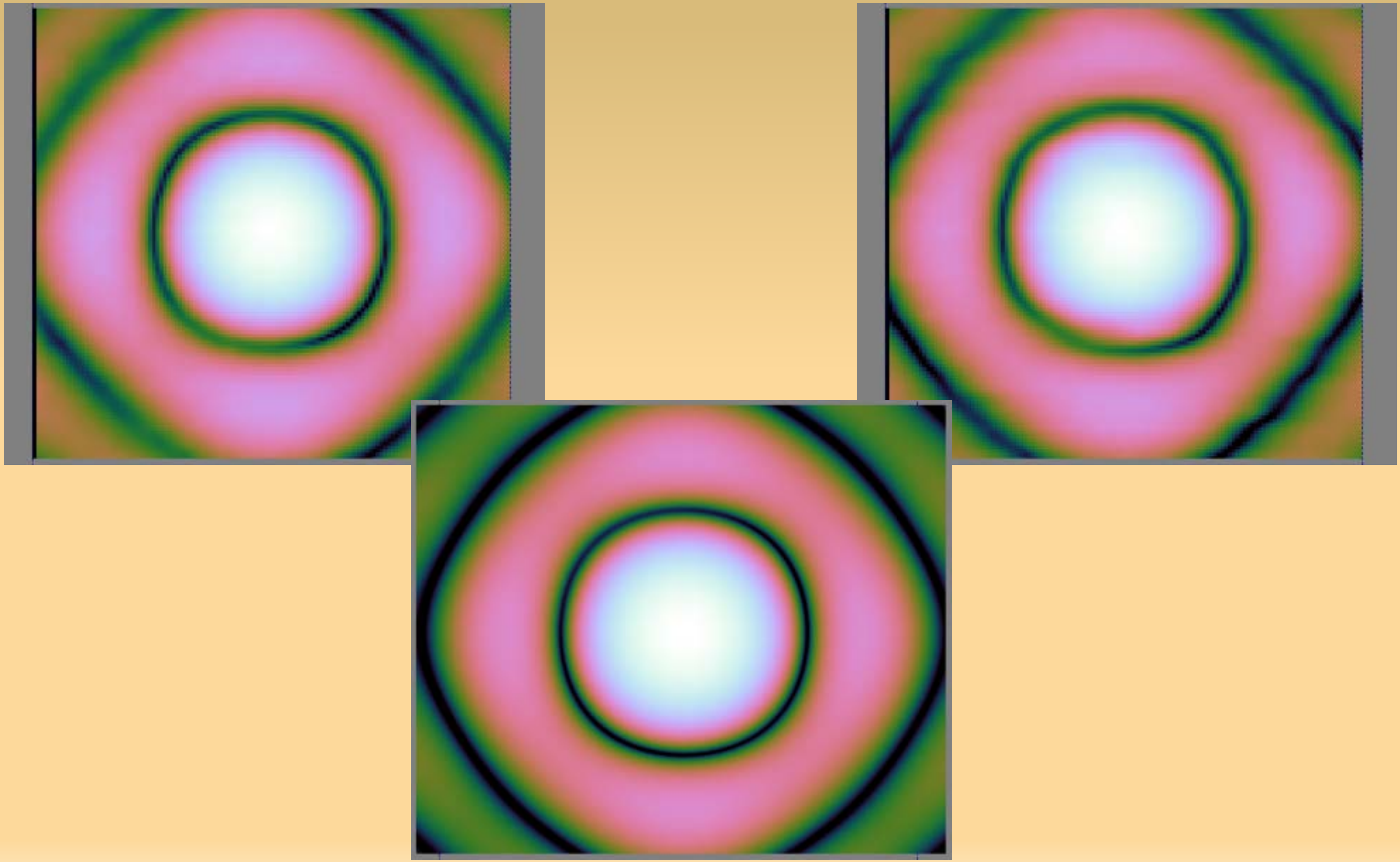




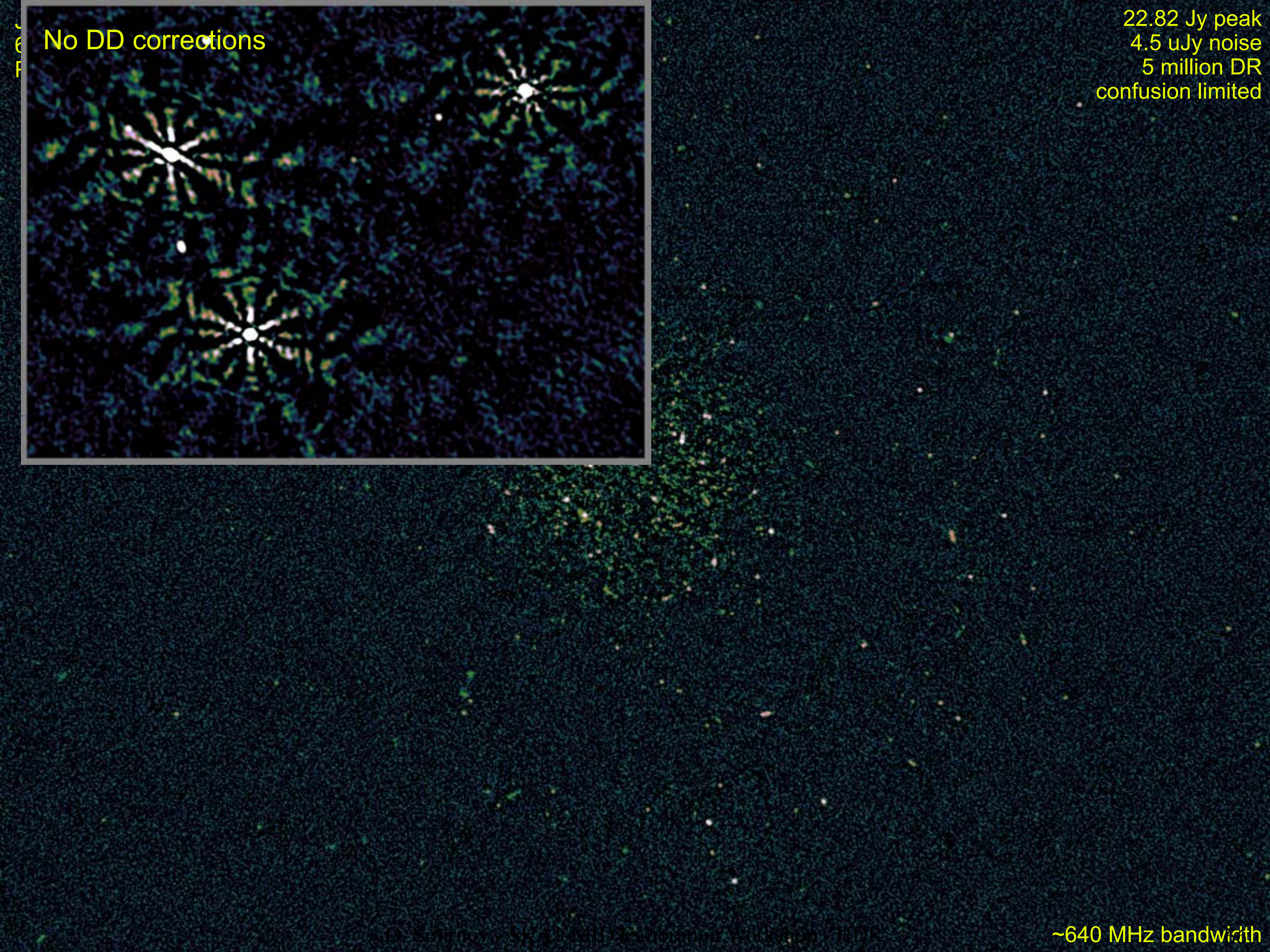
Freq /^

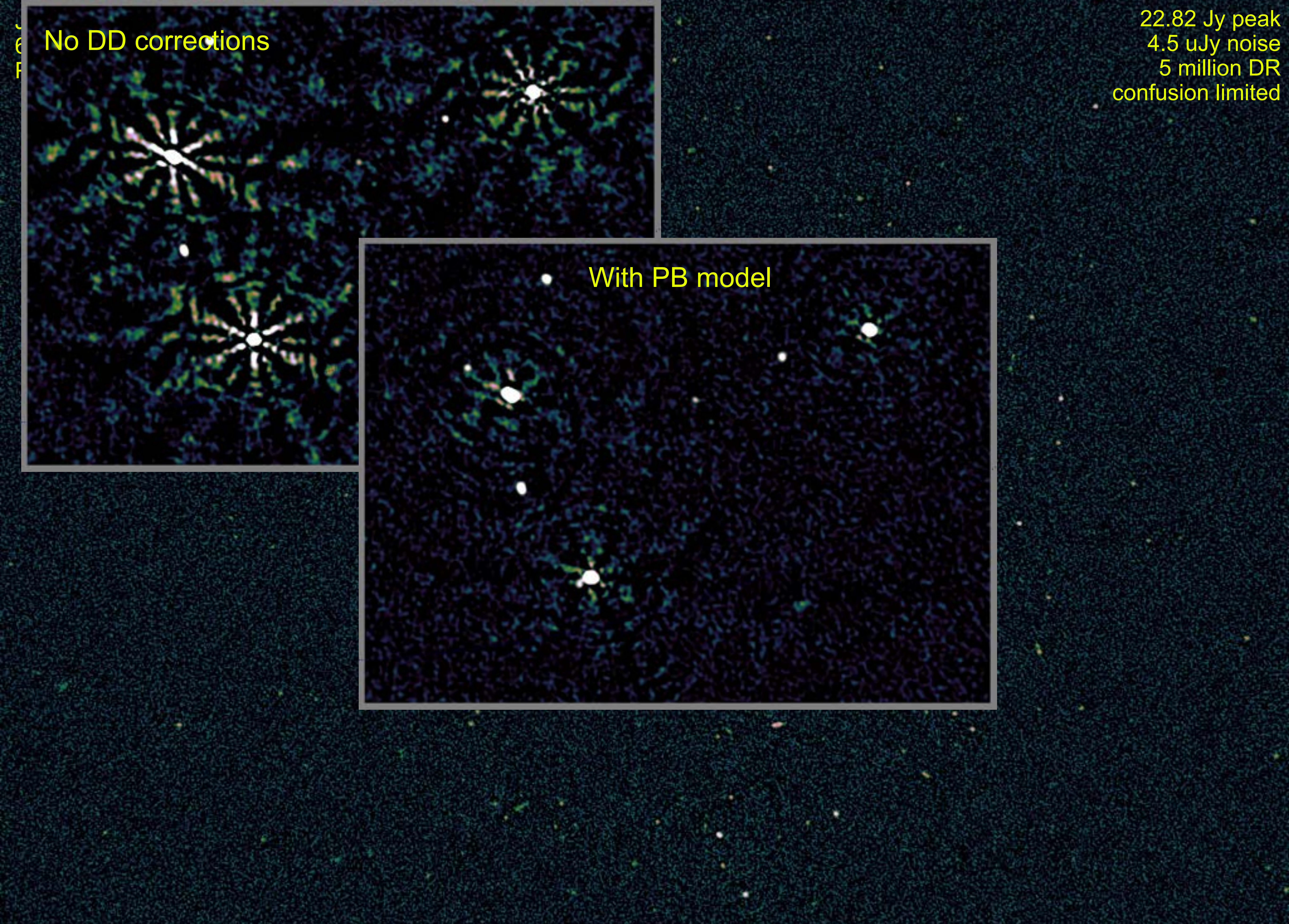
anim-5-28-LL.avi

JVLA beam holography

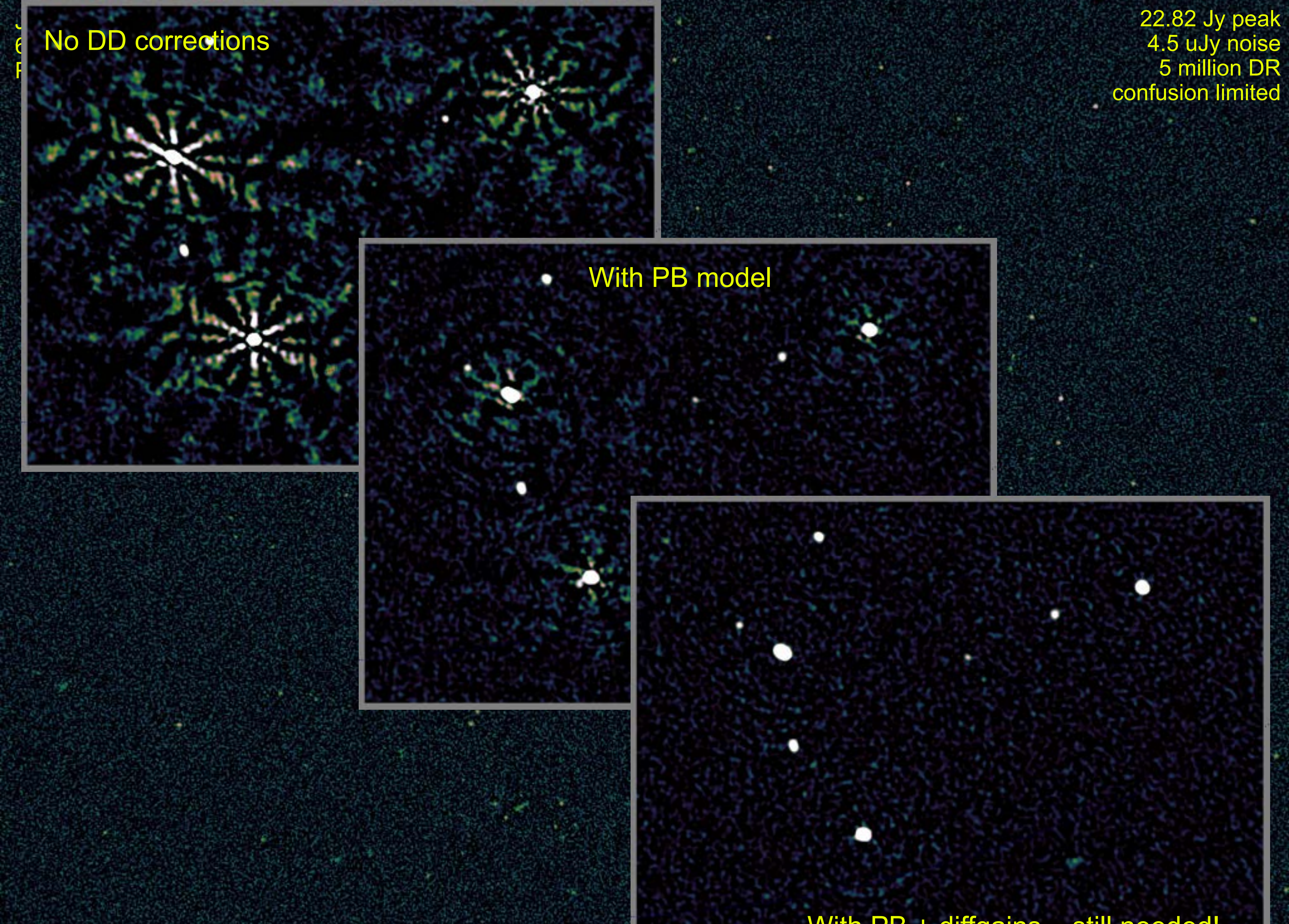


MuellerMatrixAnimation.avi


No DD corrections
22.82 Jy peak
4.5 uJy noise
5 million DR
confusion limited
~640 MHz bandwidth



22.82 Jy peak
4.5 uJy noise
5 million DR
confusion limited
No DD corrections
With PB model
~640 MHz bandwidth


22.82 Jy peak
4.5 uJy noise
5 million DR
confusion limited
No DD corrections
With PB model
With PB + diffgains – still needed!

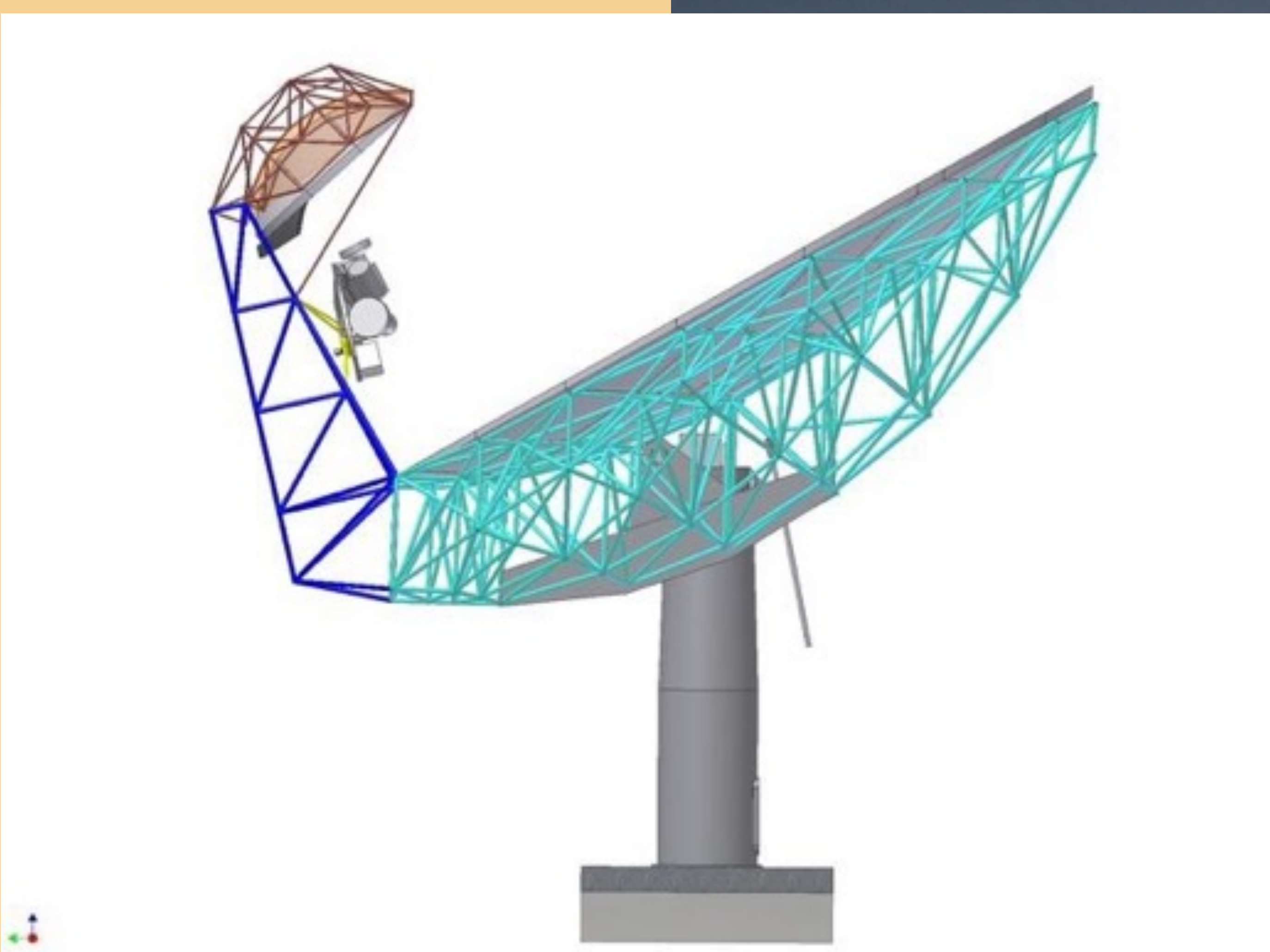
Algorithmic Advances

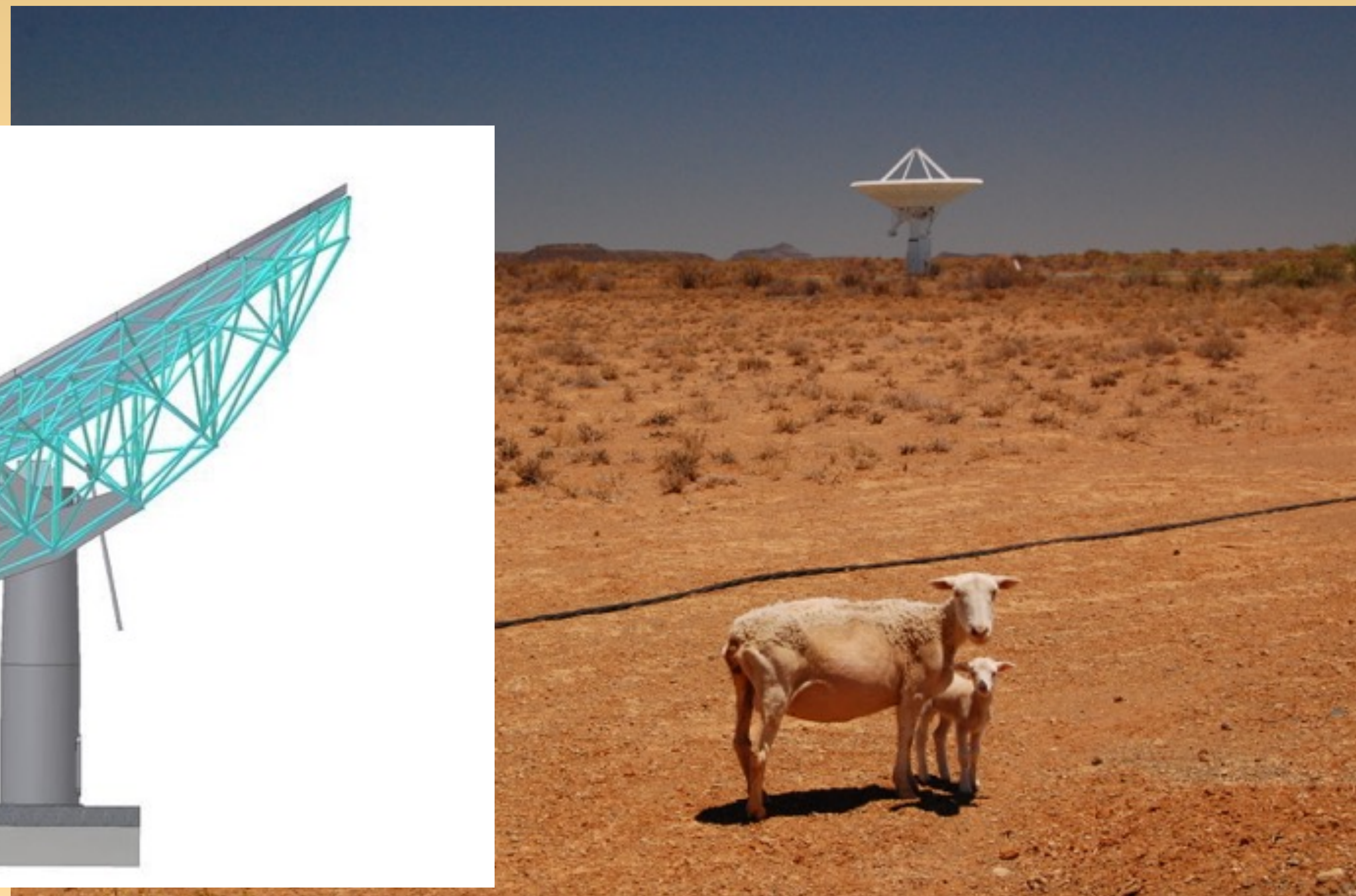
- Better deconvolution (CS and its ilk)
- Artefact discrimination
- COHJONES (aka DD-Stefcal) – much faster DD solutions
  - Extending this to beam parameters
- KAFCA (C. Tasse) – even faster DD solutions via a Kalman-type filter
  - Can contemplate **G-** and **dE-**calibration in an online regime, then average the data down
- DD faceting (C. Tasse) – faceted imager with on-the-fly BDA – x10-x100 gridding savings

=> Convex optimization + sparsity for calibration & imaging

Case Study: KAT-7 vs. MeerKAT

- MeerKAT: offset Gregorian dishes
- KAT-7 (MeerKAT precursor): prime focus dishes





KAT-7 vs. MeerKAT Beams
Prime focus vs. offset Gregorian