Gaia

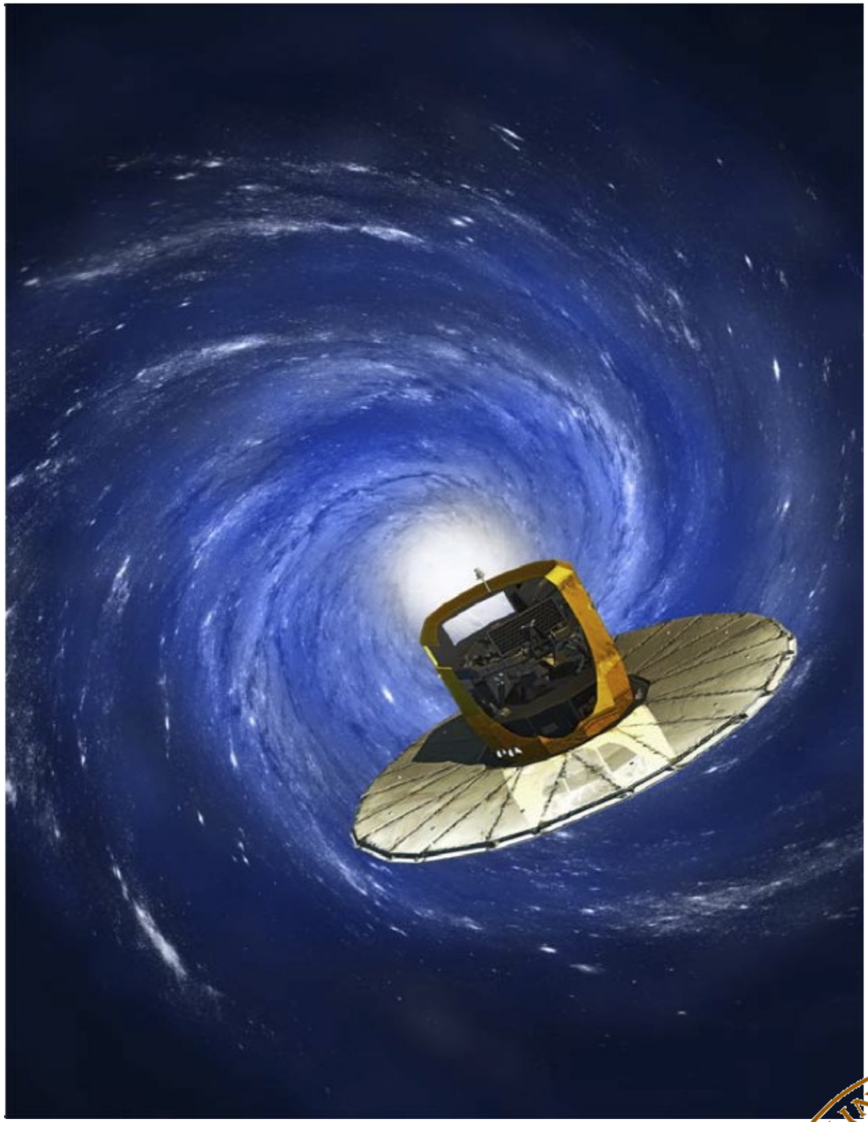
Solving equations with a billion unknowns

Lennart Lindegren

Lund Observatory
Lund University, Sweden

Outline of talk

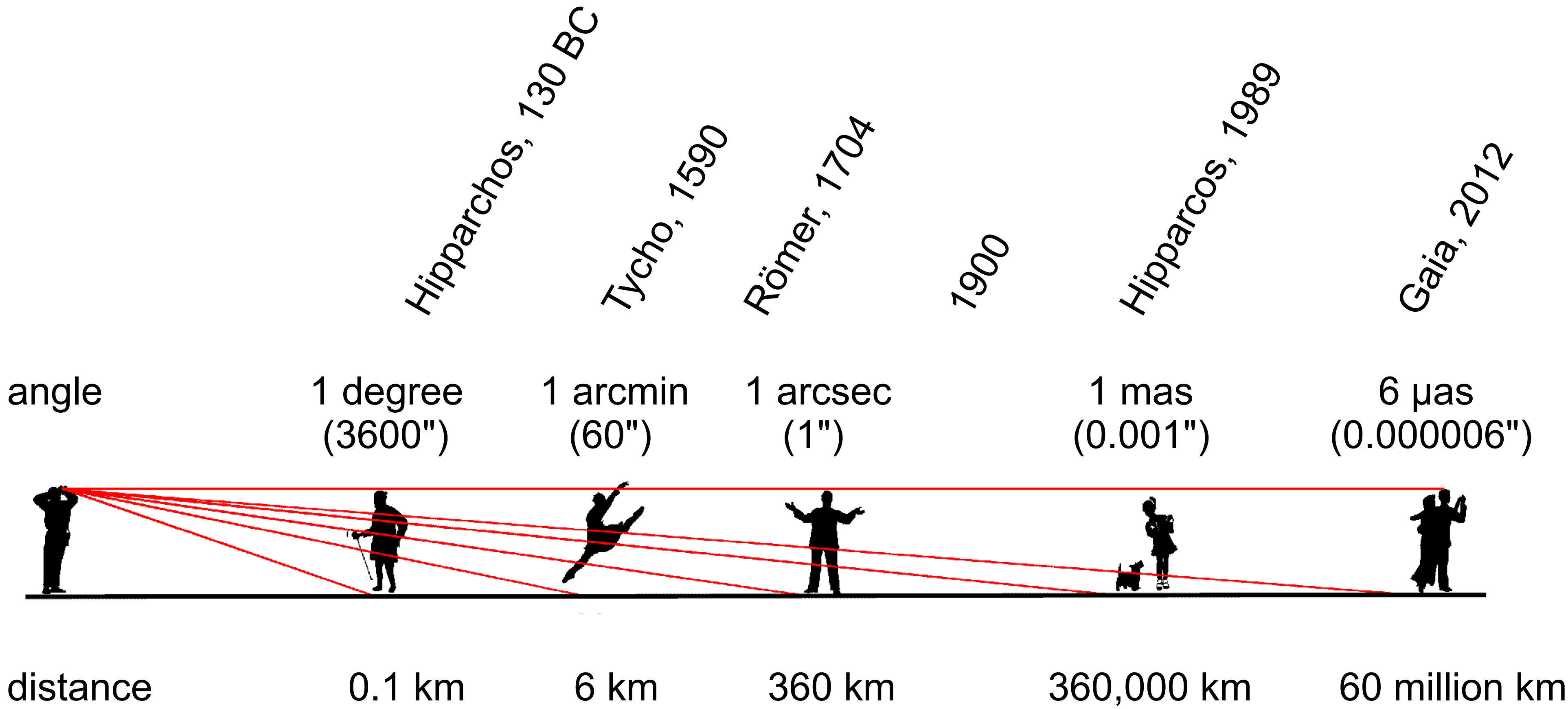
- Gaia in a nutshell
- Observation principle
- Hardware
- Data processing challenge
- The solution



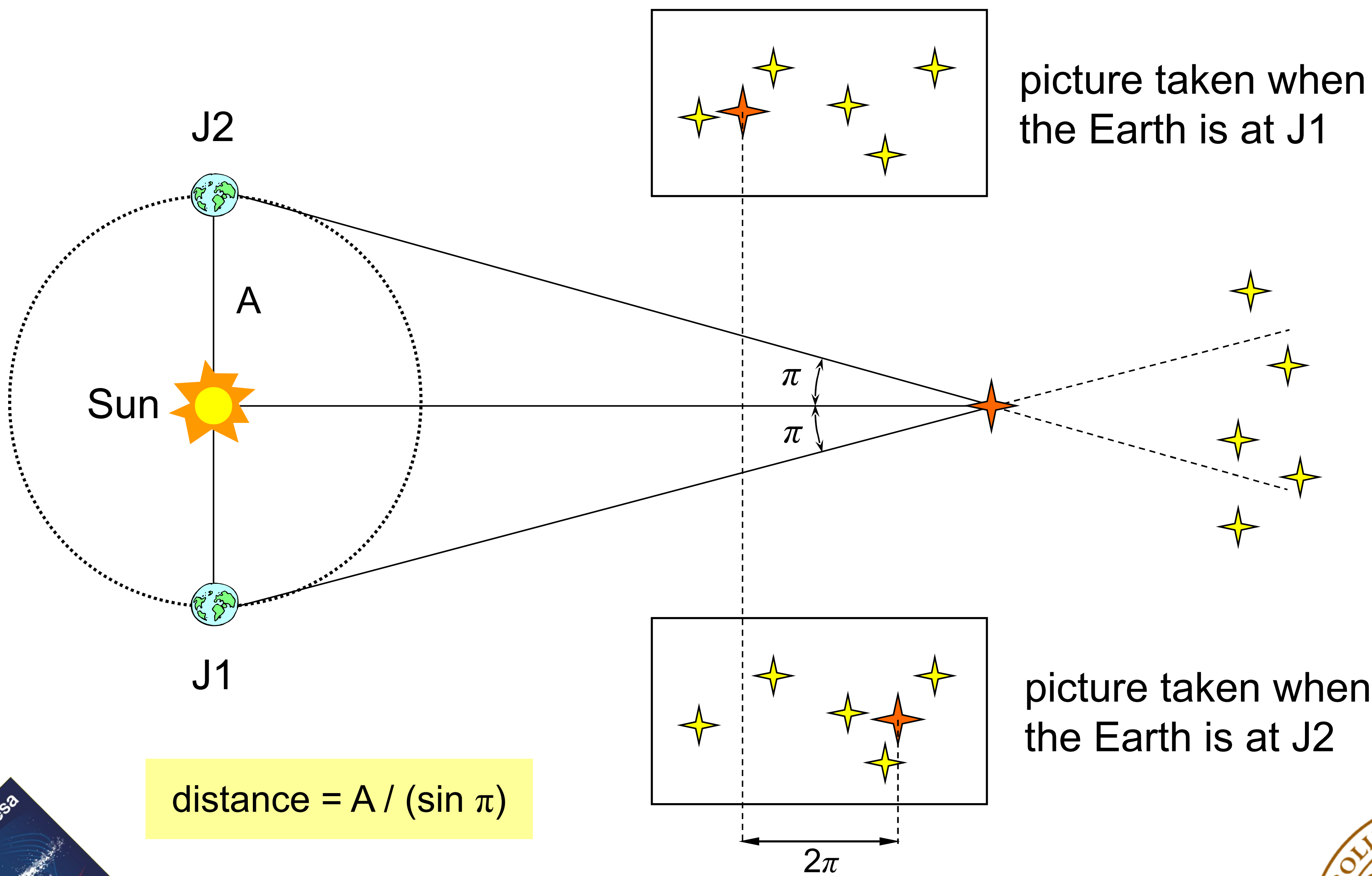
Gaia in a nutshell

- **All-sky astrometric survey carried out 2012 – 2017**
  - final results around 2020
  - positions, parallaxes (distances), proper motions (transverse vel.)
  - flux measurements at low spectral resolution
- **All point objects in the detection range (magnitude 6 to 20)**
  - stars, asteroids, quasars, extragalactic supernovae, etc
  - about 10^9 objects
- **Using Hipparcos principle (scanning, two fields of view)**
  - positional accuracy from 6 μas (bright stars) to 200 μas (faint)
  - tied to a non-rotating frame via ~500,000 quasars
- **Spectroscopic radial velocities (V < 16, few km/s)**

How small is 6 micro-arcseconds?



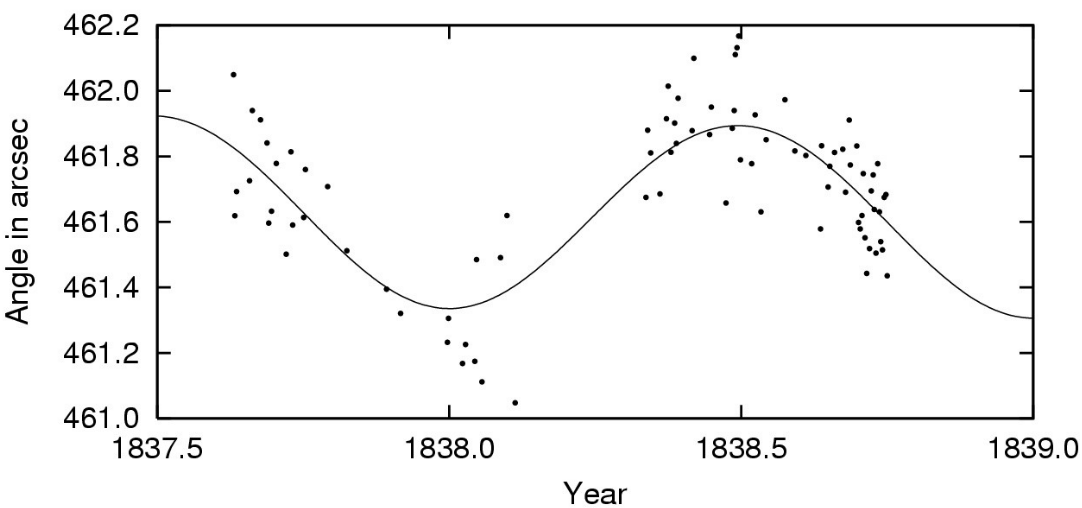
Measurement of parallax (π)

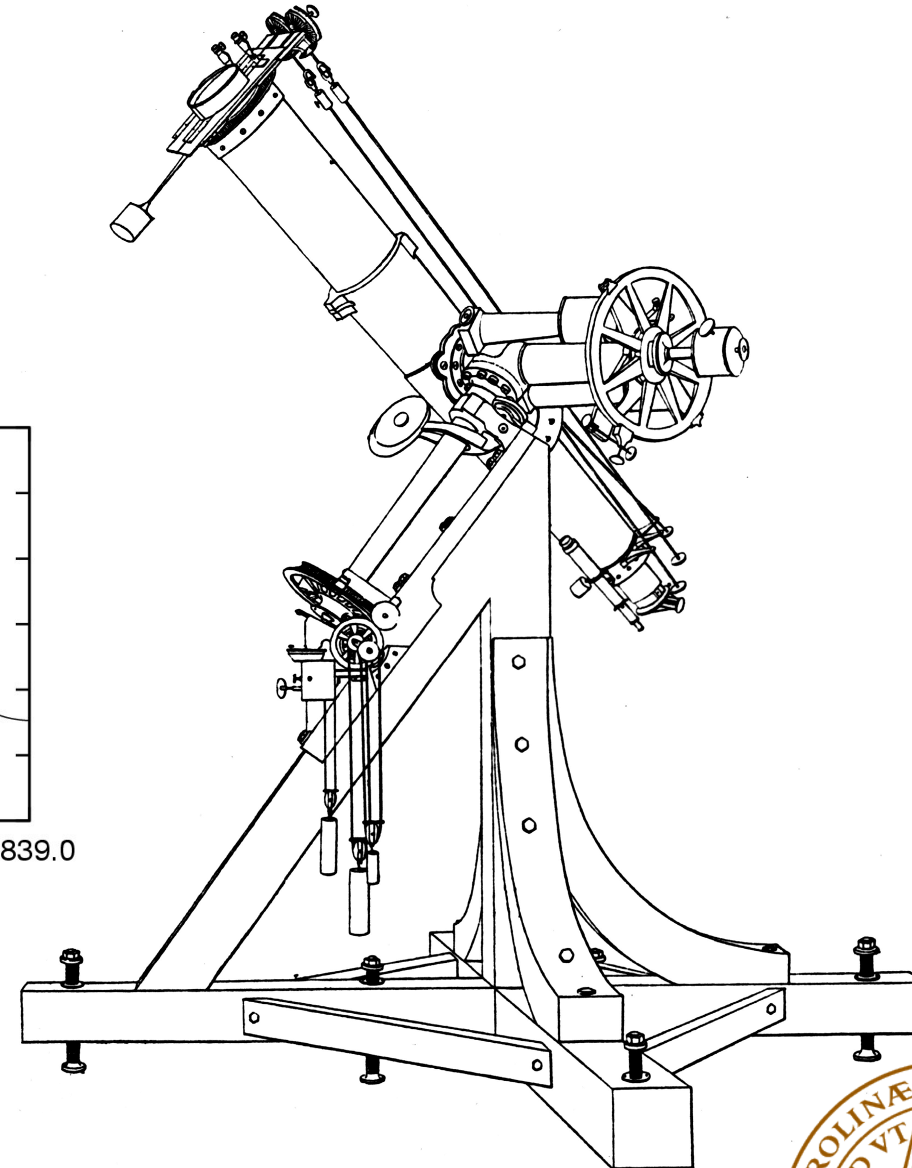


distance = A / (sin π)

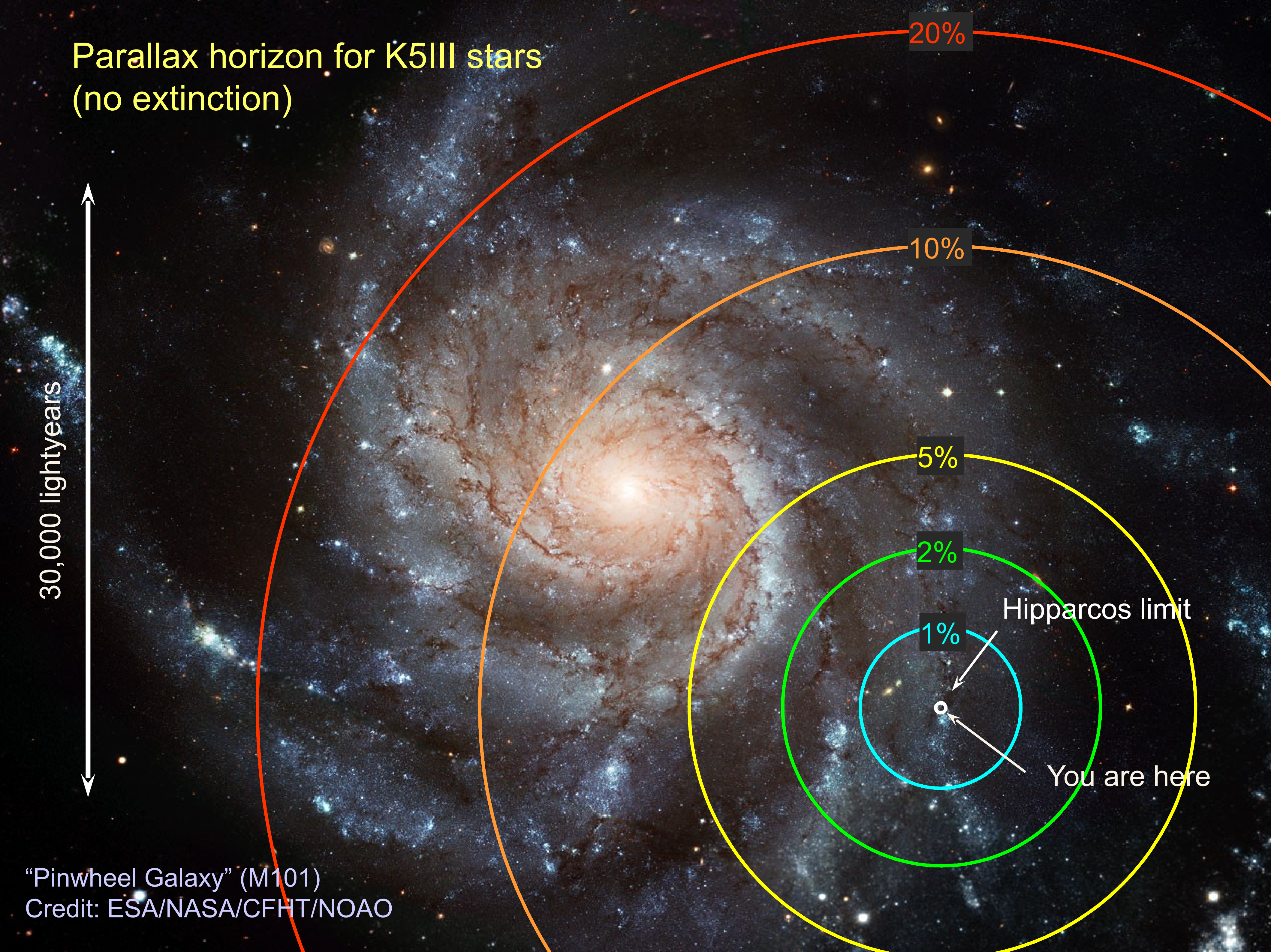
Principle of parallax measurement - from the ground

Friedrich Wilhelm Bessel's measurement of the parallax of 61 Cyg (1838) - relative to two background stars - gave = 0.294” (error was about 0.01”)



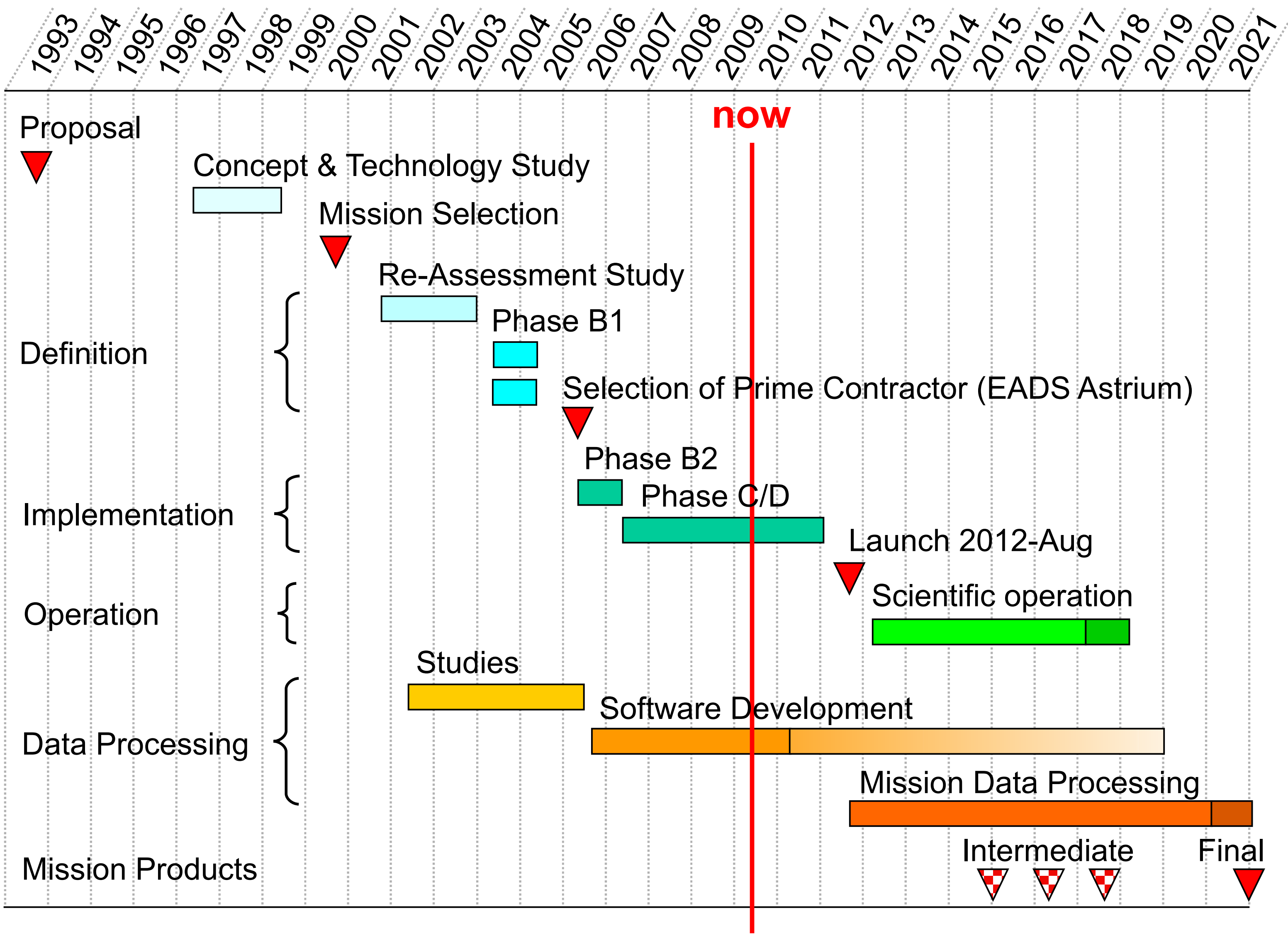


Fraunhofer's 70 1/8 f-Heliometer, 1826



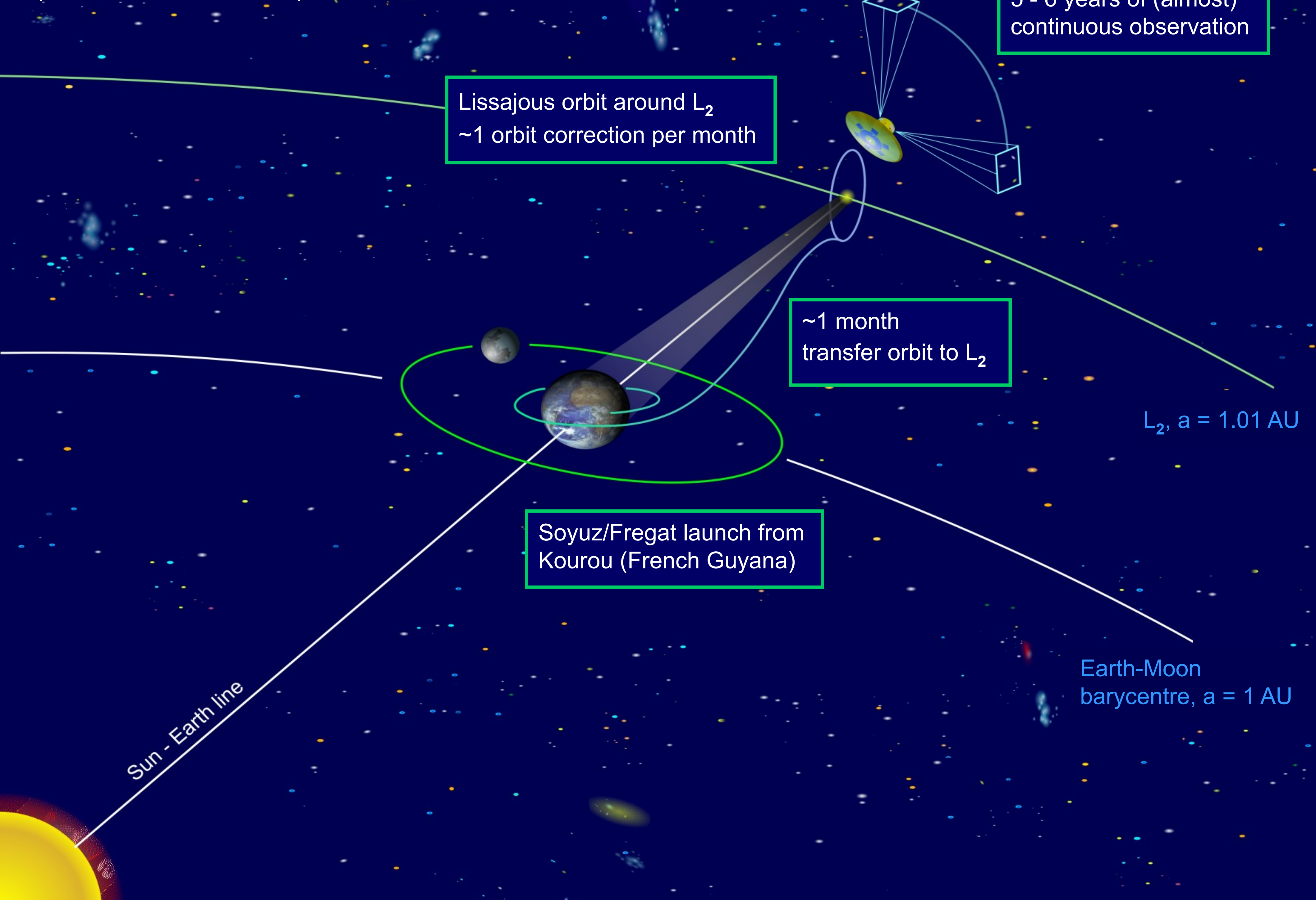
Parallax horizon for K5III stars
(no extinction)
30,000 lightyears
20%
10%
5%
2%
1%
Hipparcos limit
You are here
“Pinwheel Galaxy” (M101)
Credit: ESA/NASA/CFHT/NOAO

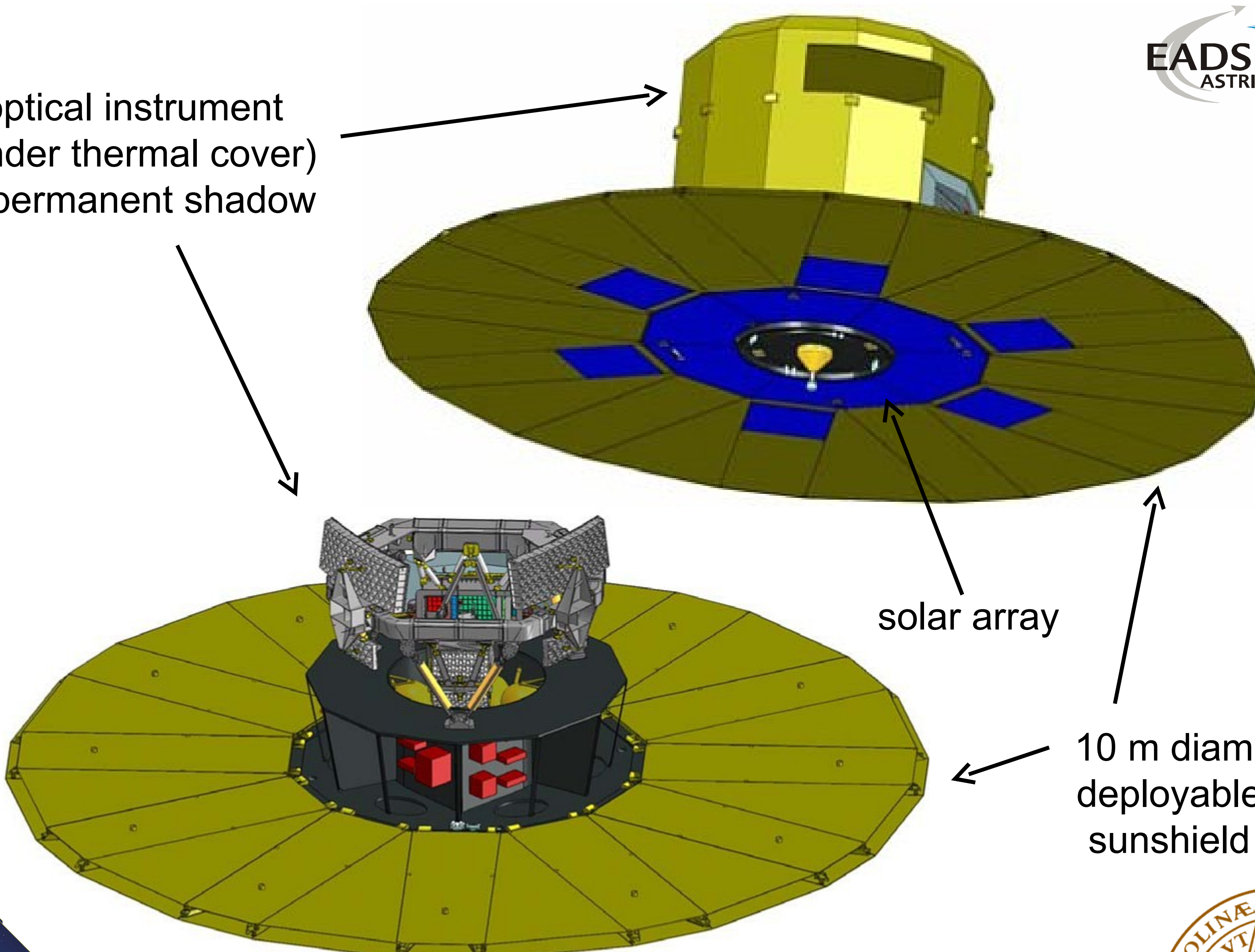
Gaia - schedule



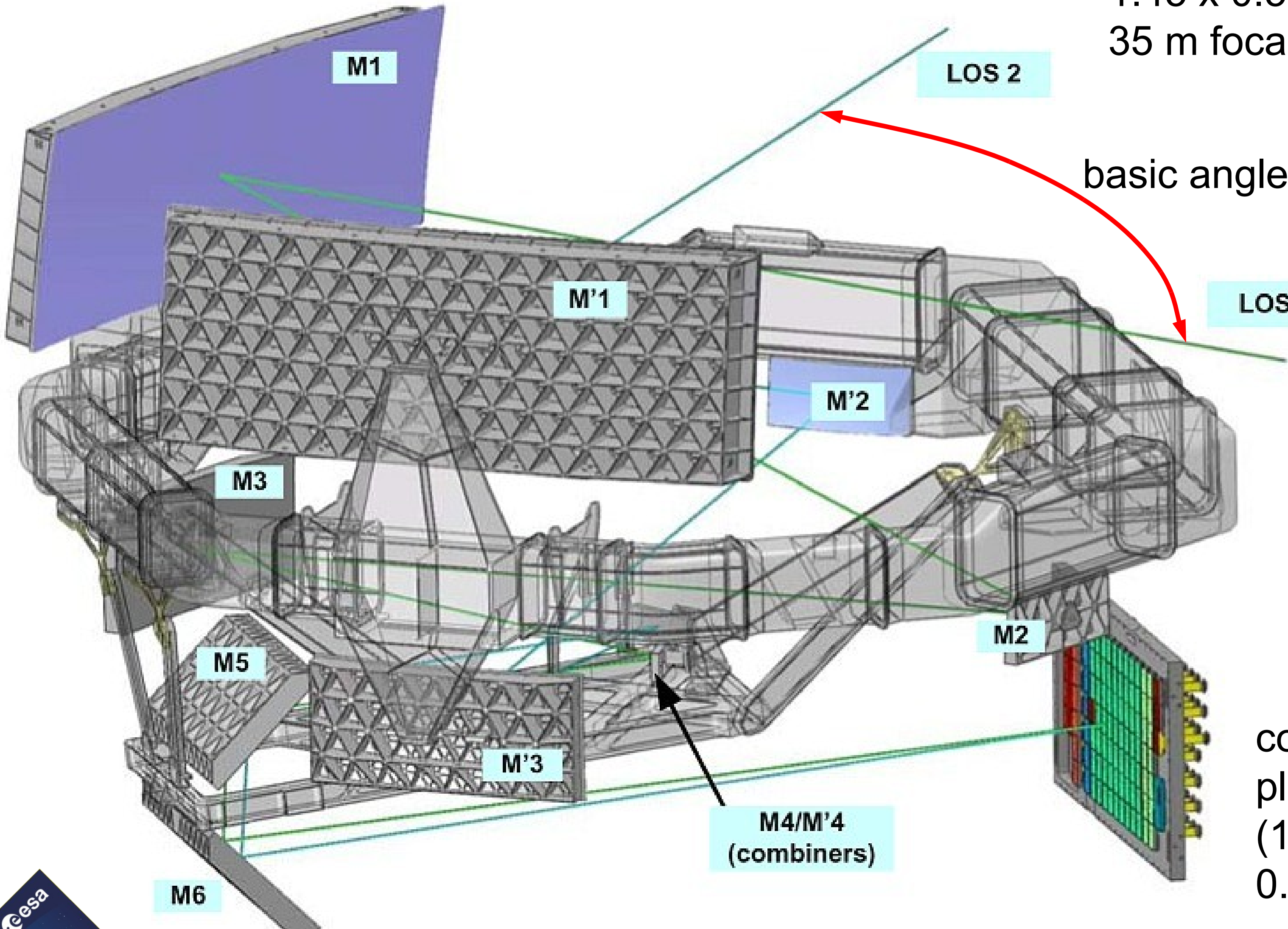
Gaia launch and orbit

(Credit: EADS Astrium)



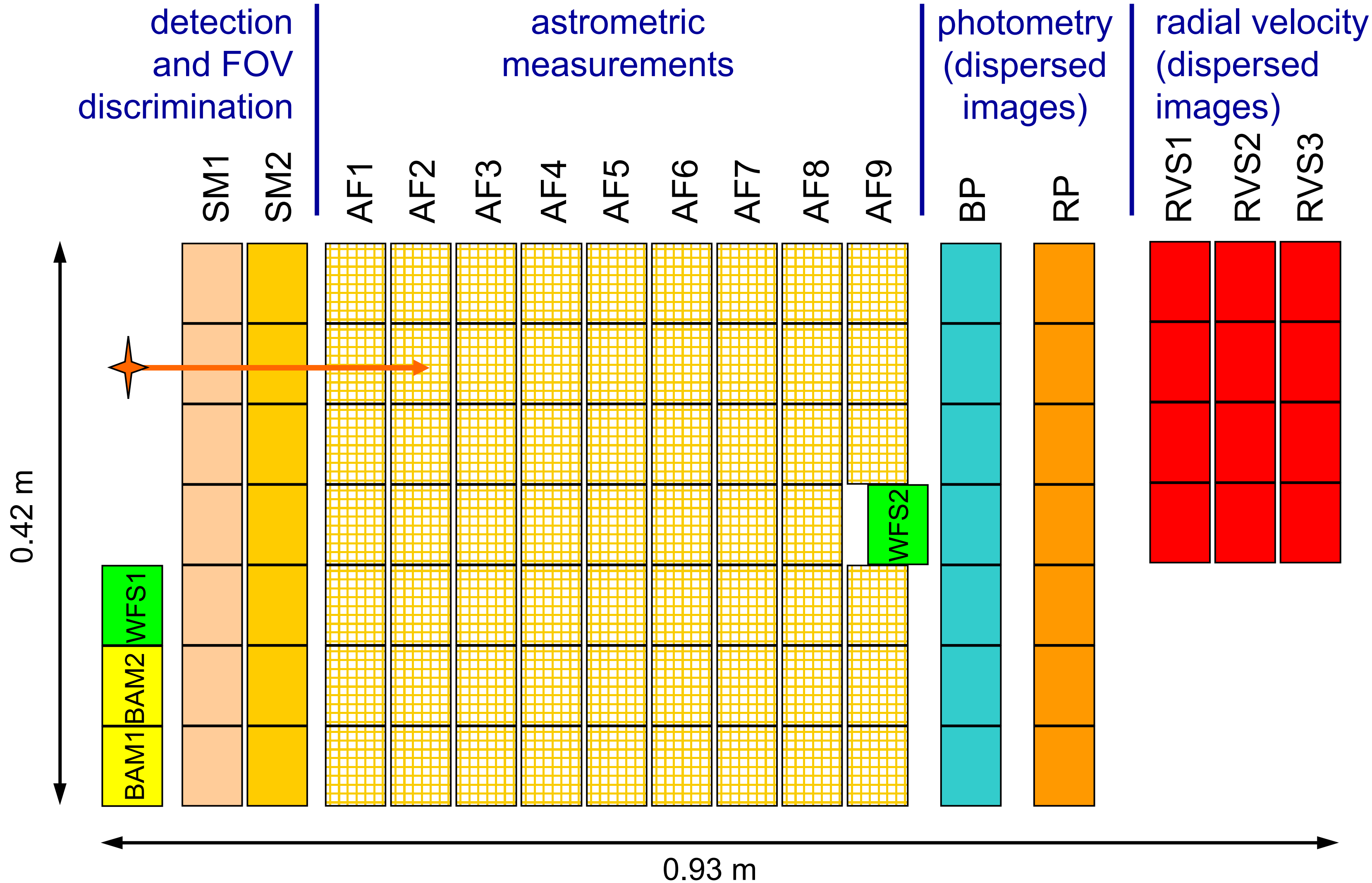

optical instrument
(under thermal cover)
in permanent shadow
solar array
10 m diam.
deployable
sunshield

2 off-axis telescopes
1.45 x 0.5 m² aperture
35 m focal length



common focal plane, 106 CCDs (1 Gigapixel)
0.93 x 0.42 m²

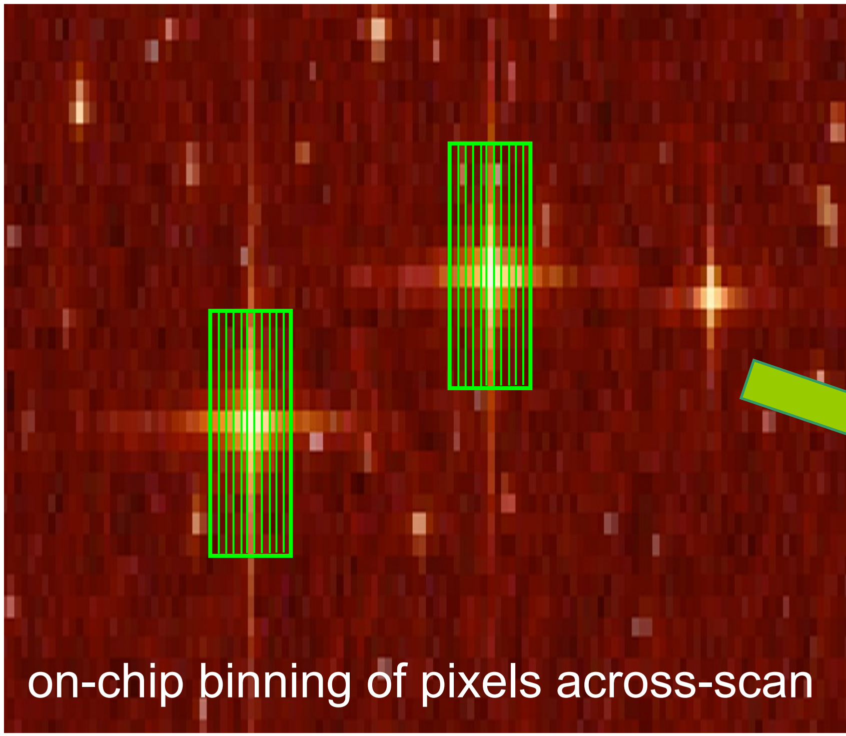
Gaia focal plane (106 CCDs)



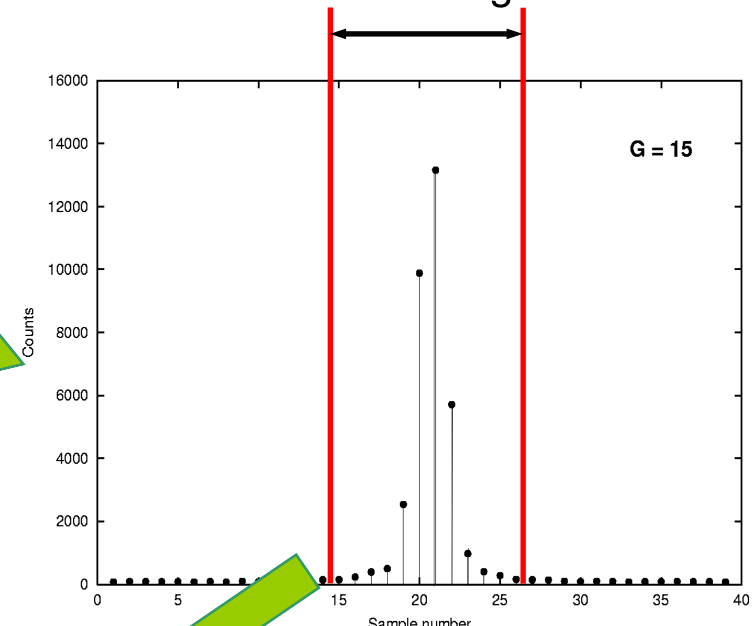
BAM = basic angle monitor, WFS = wavefront sensor

CCD data collection in the astrometric field

scan



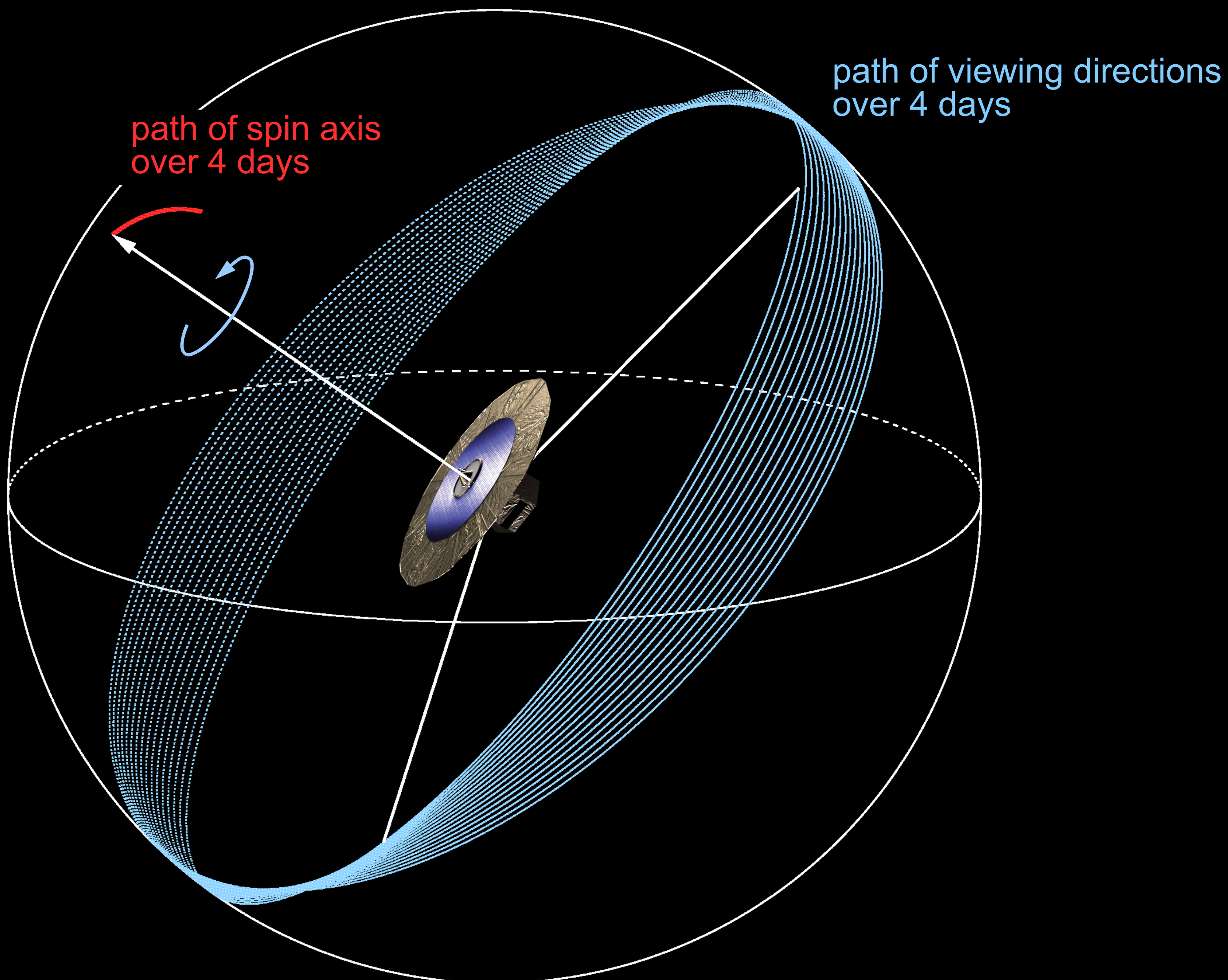
window of 6-18 samples
transmitted to ground



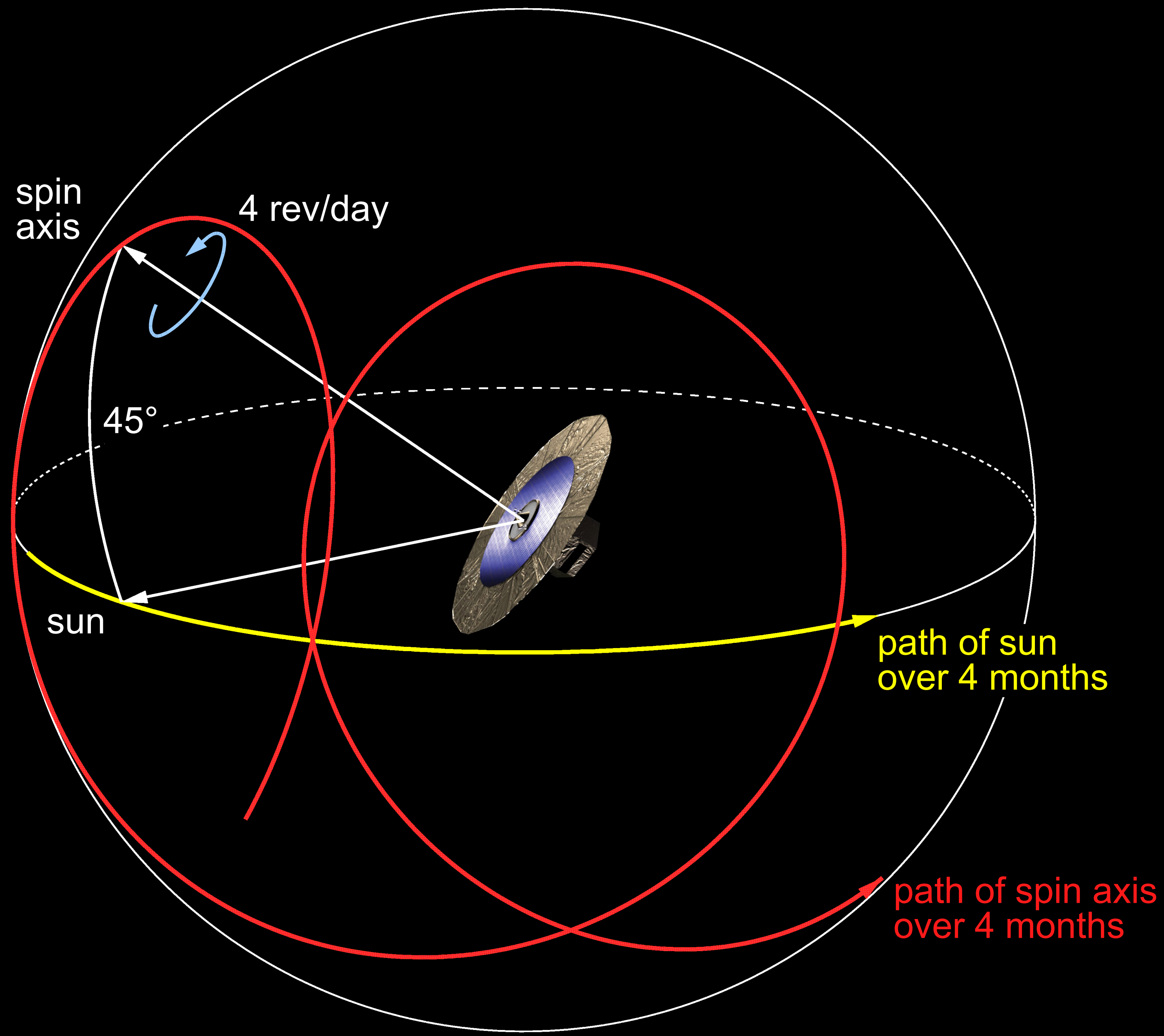
"Time of observation" for image centre relative to CCD determined to ~200 μas precision (magnitude 15)

Some 700 such measurements per object in 5 years

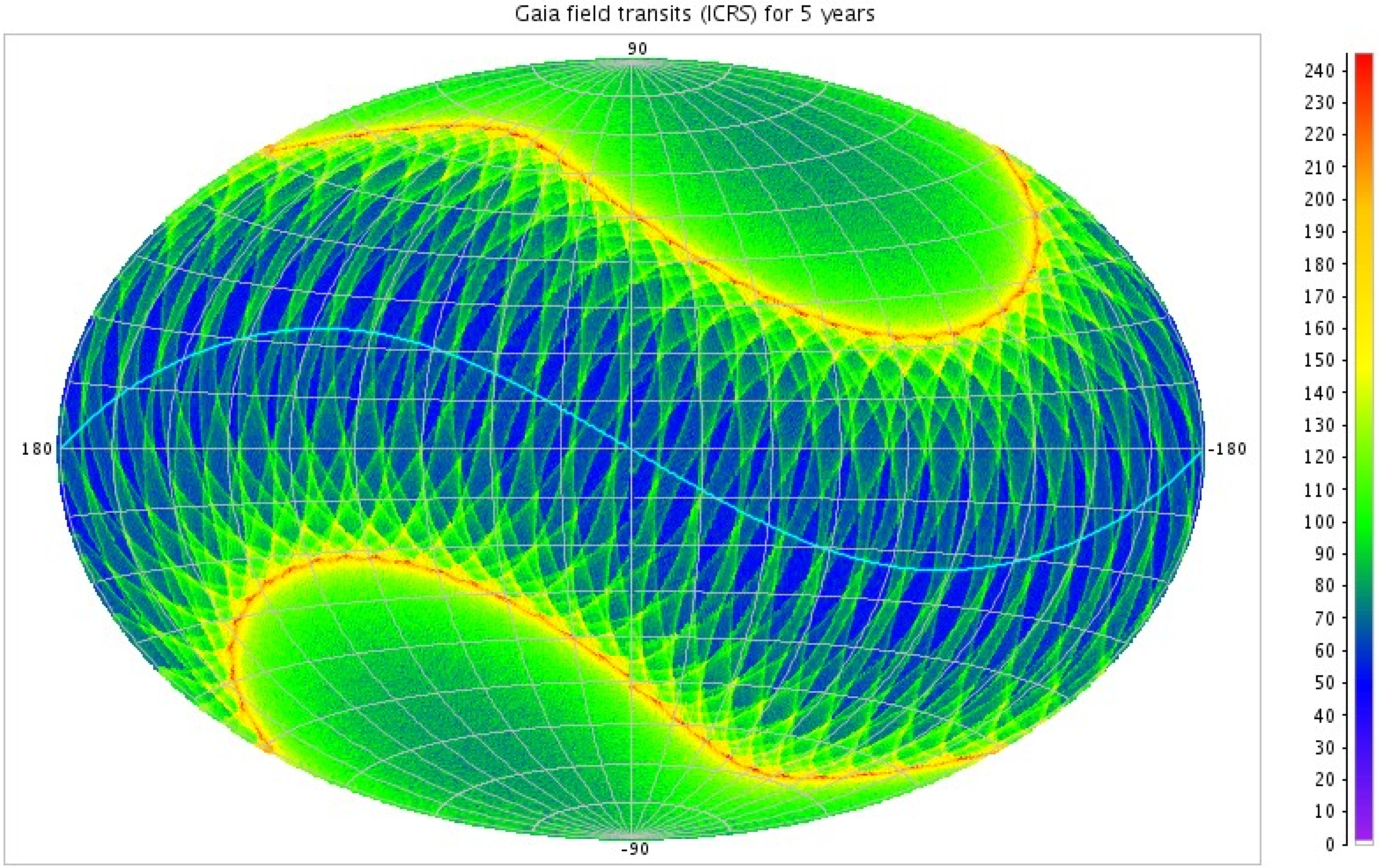
Gaia scanning: Motion of viewing directions over 4 days



Gaia scanning: Motion of the spin axis over 4 months



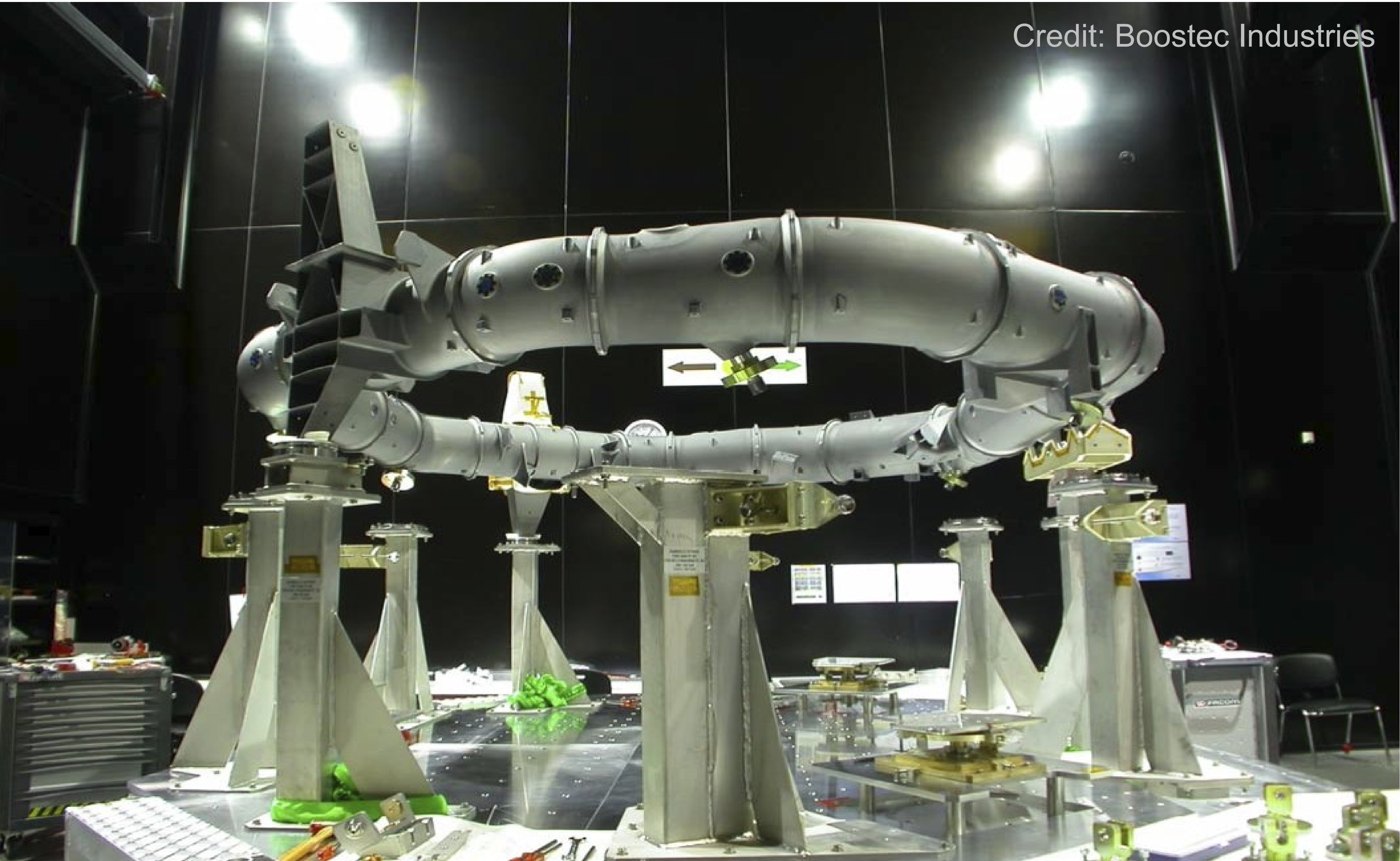
Number of FoV crossings per star (5 years)



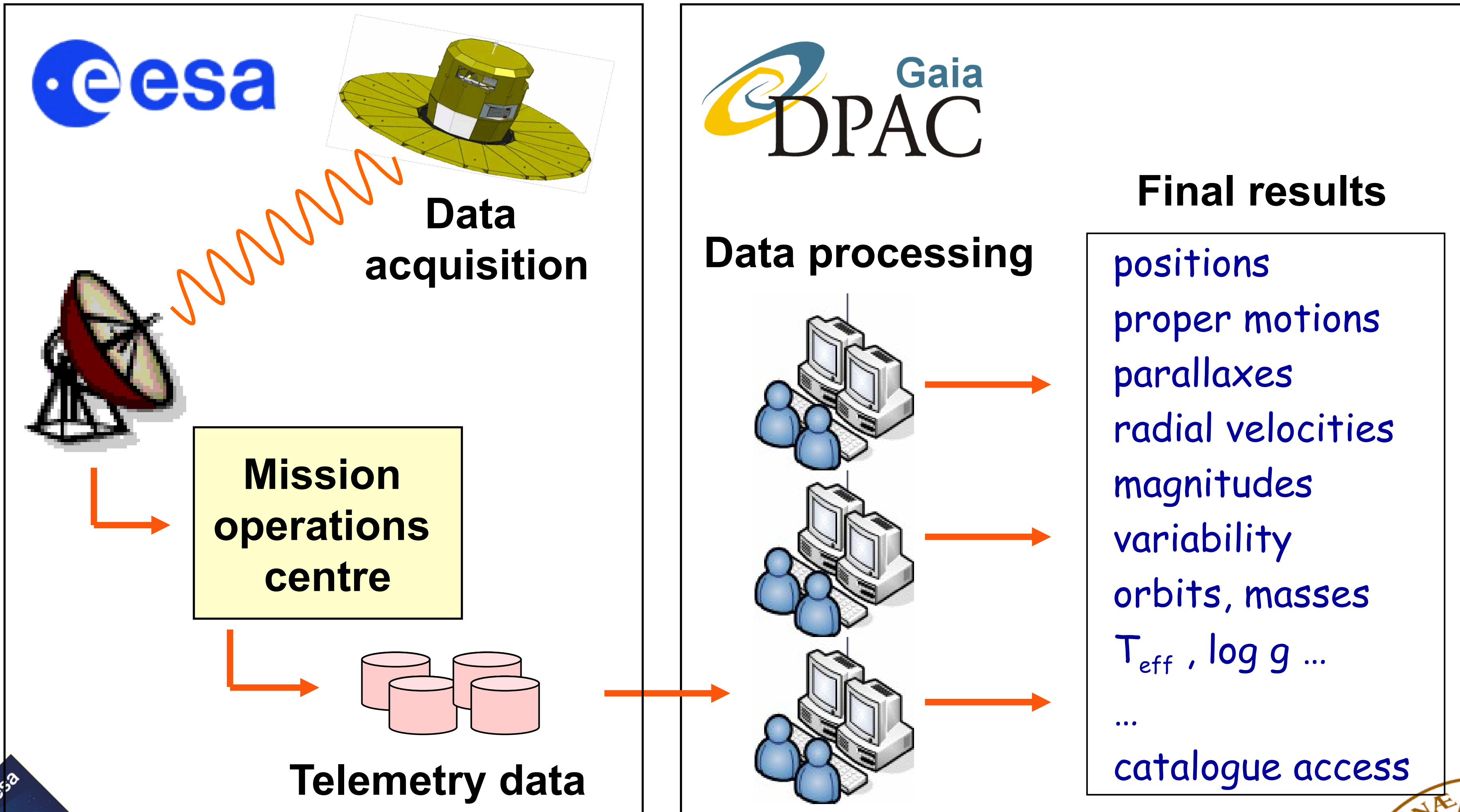
sky average = 80

Equatorial projection

SiC torus (optical bench) finished (Jan 2010)

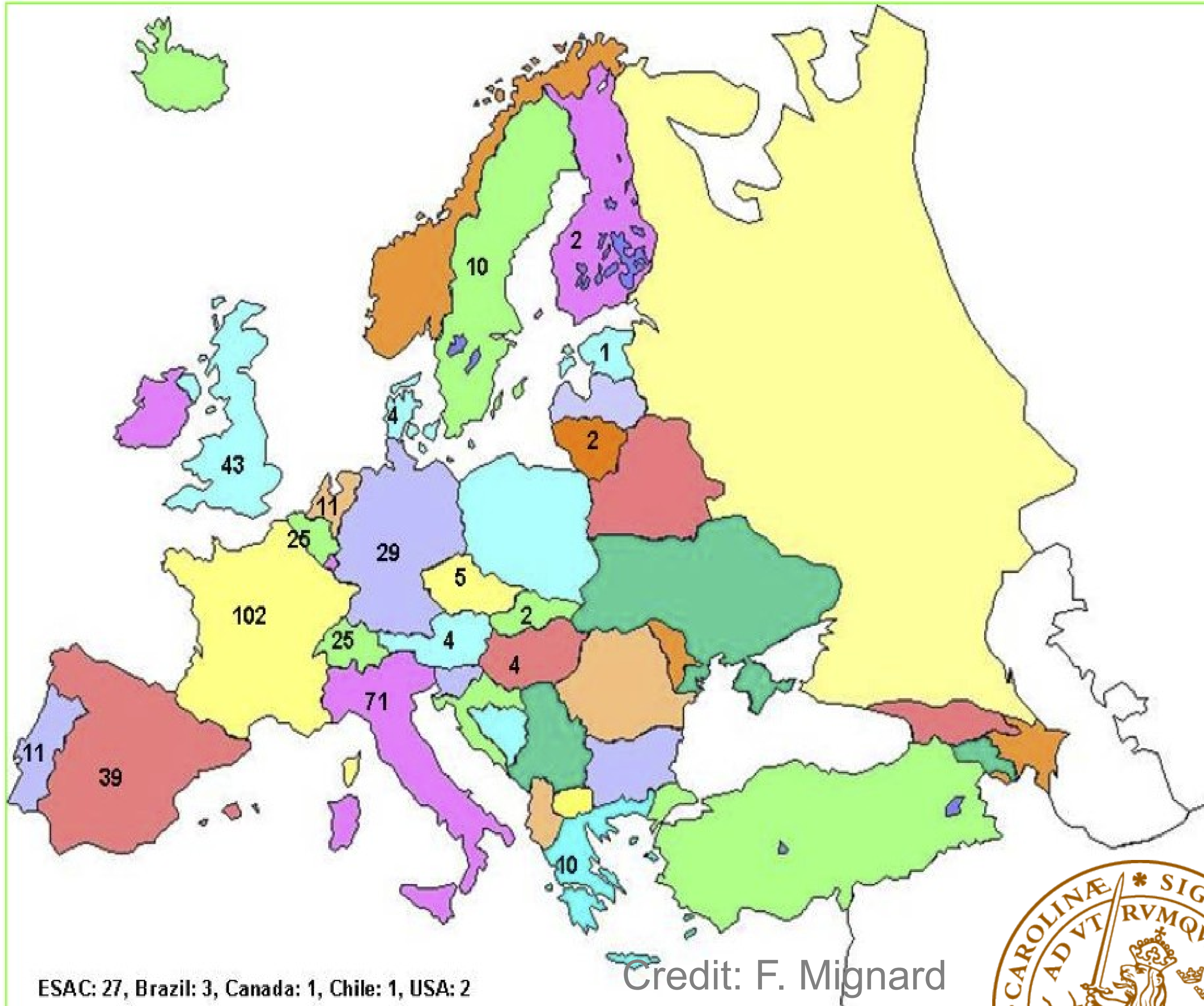


Data collection and reduction: ESA and DPAC responsibilities



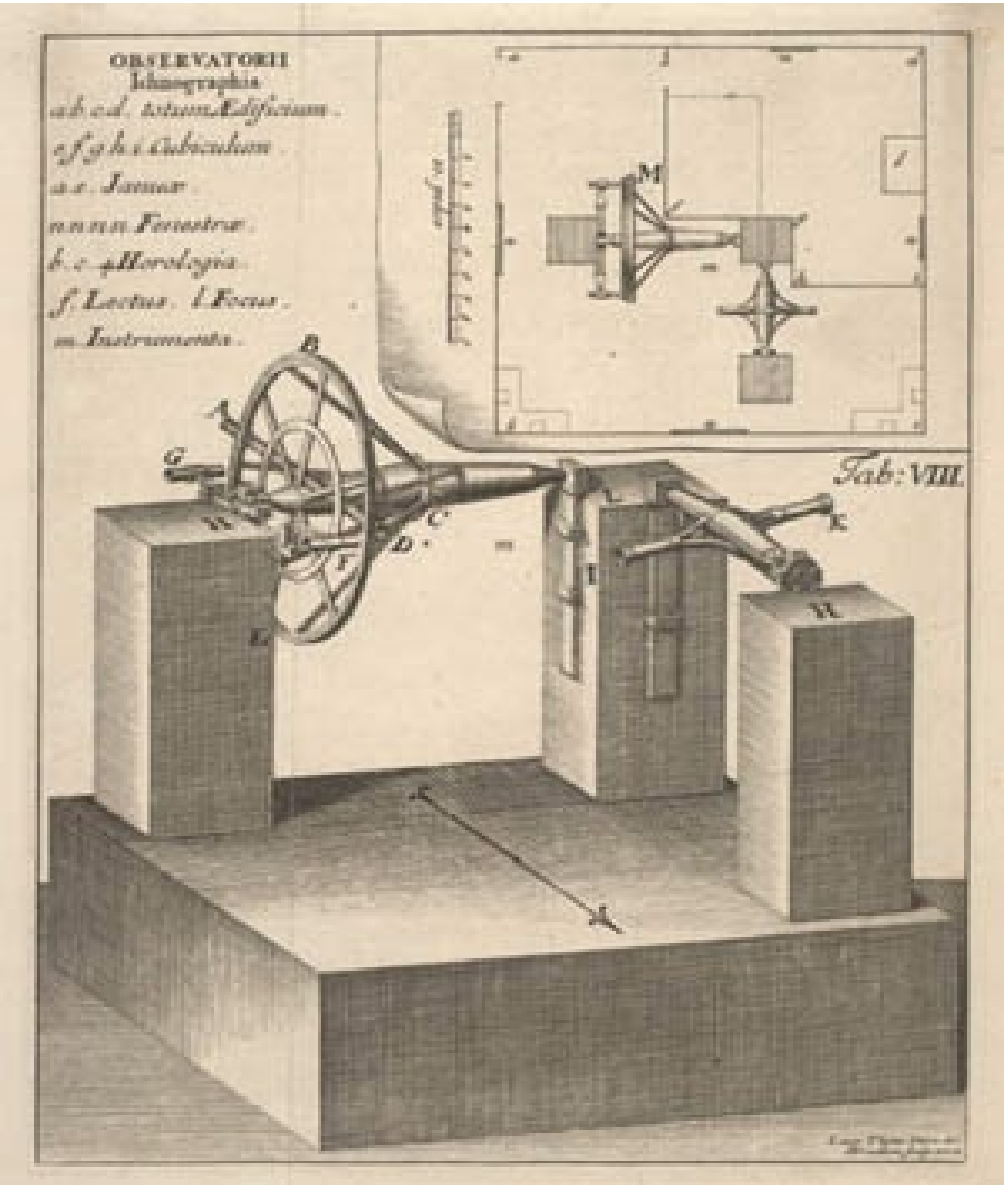
Gaia Data Processing and Analysis Consortium (DPAC)

- Individual persons organized in an ad hoc structure
- About 430 members (January 2010)
- 24 funding agencies
- 6 data processing centres



Credit: F. Mignard

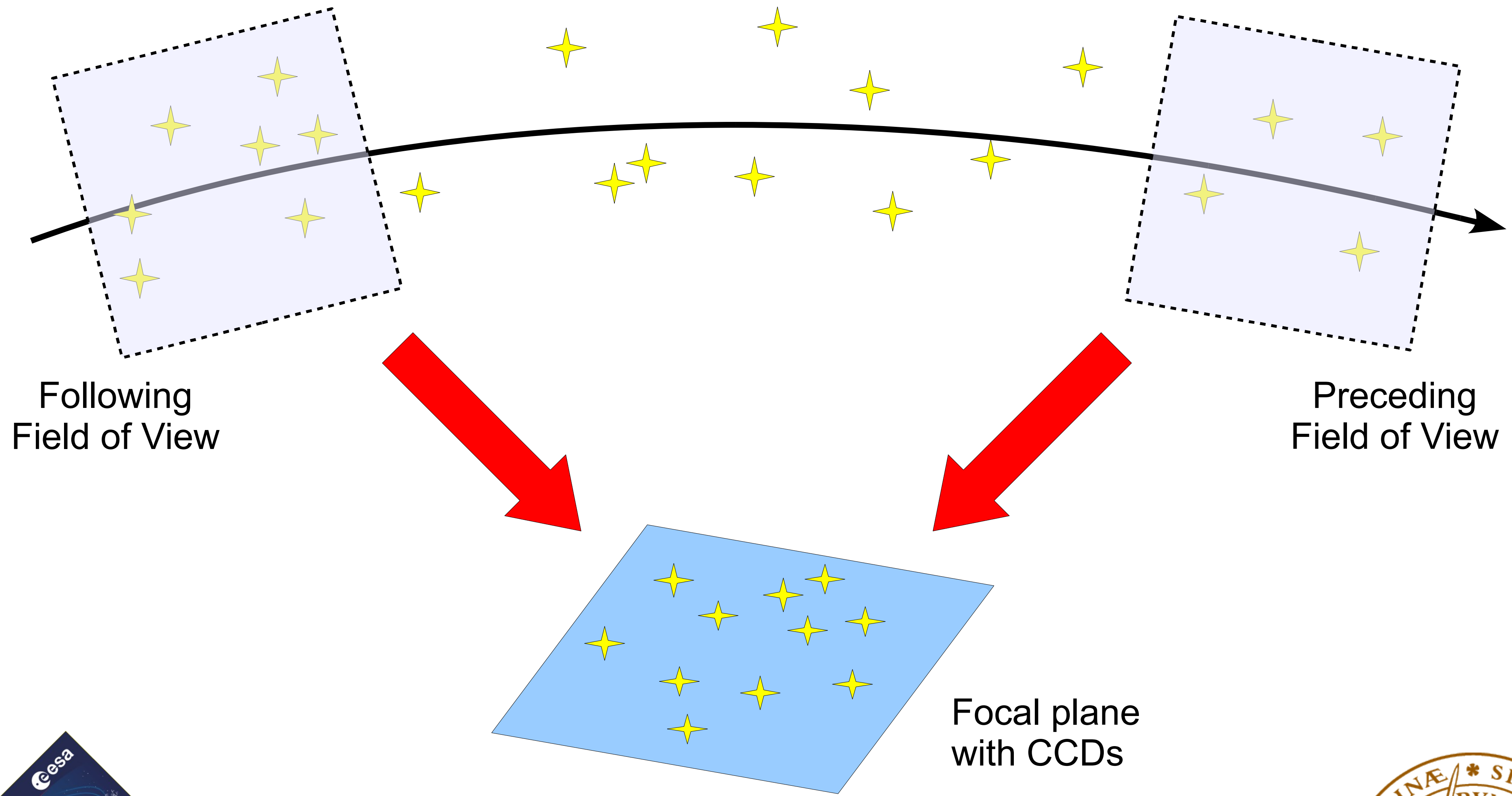
How does the Gaia data processing differ from other (ground-based) astrometric projects?



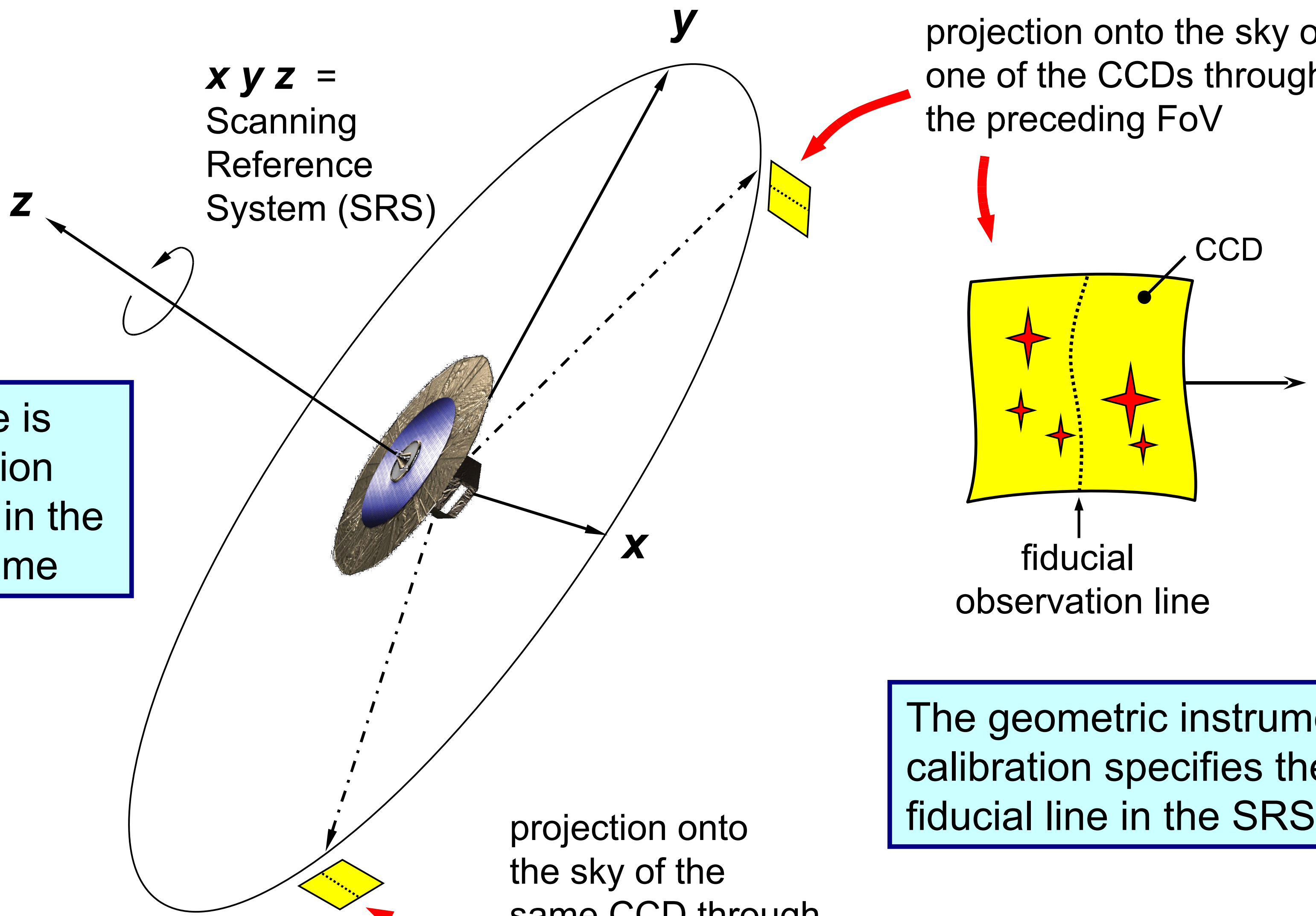
Ole Rømer's meridian circle (ca 1704)

The meridian circle at Lund Observatory (ca 1920)

Linking stars by focal plane measurements



The observations depend on many things besides the star positions...

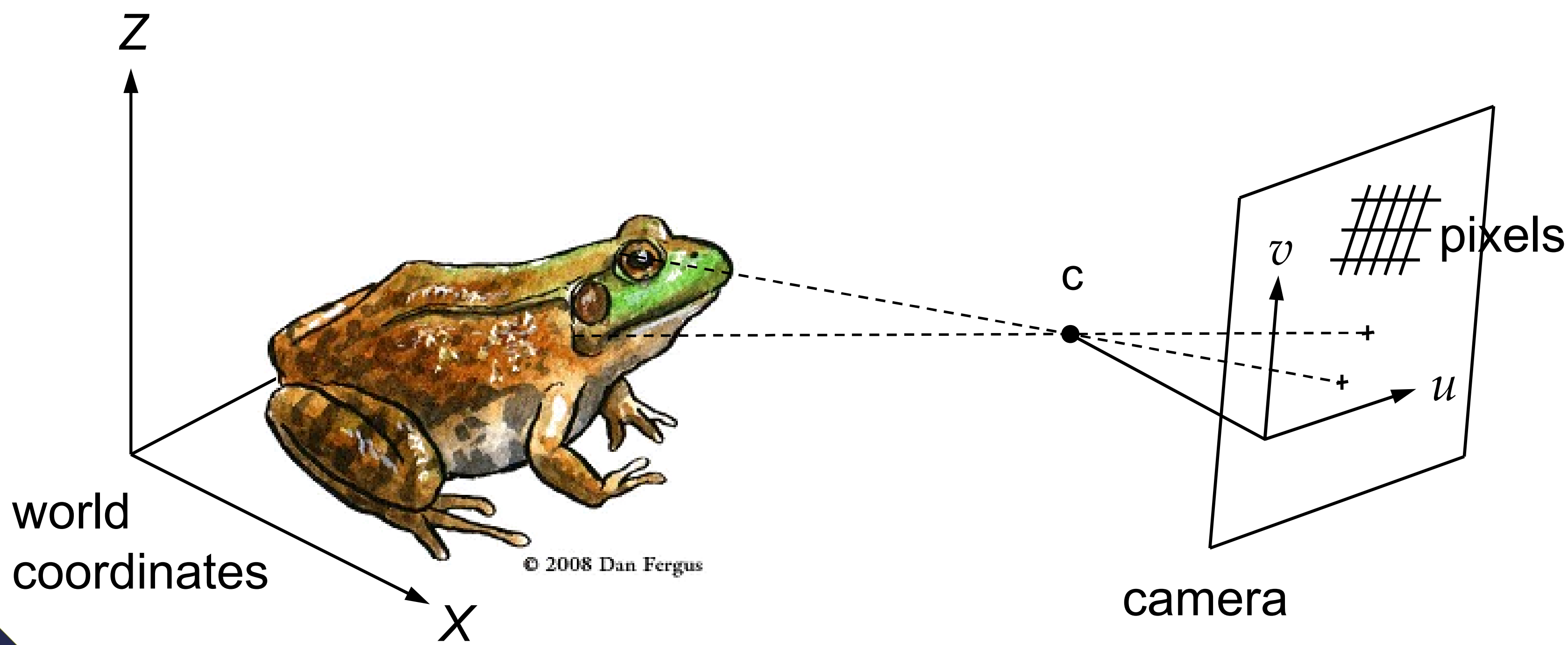


The attitude is the orientation of the SRS in the celestial frame

The geometric instrument calibration specifies the fiducial line in the SRS

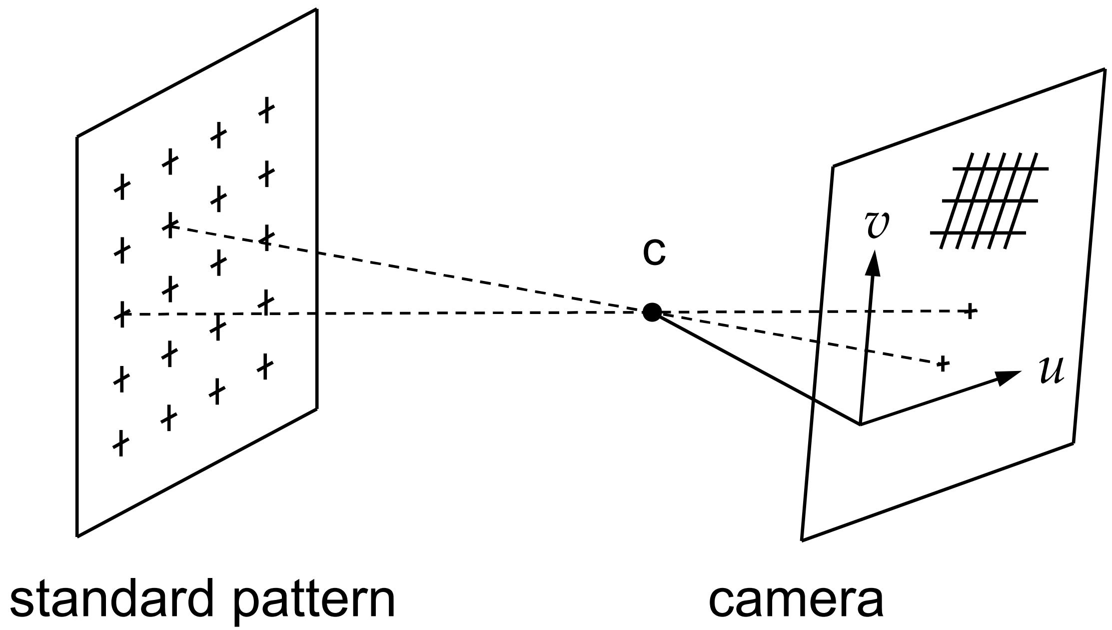
Principle of self-calibration: An example from computer vision

- Extrinstic camera parameters: (X, Y, Z) → (u, v) [rotation, translation]
- Intrinsic camera parameters: (u, v) → pixels [origin, scales, shear]



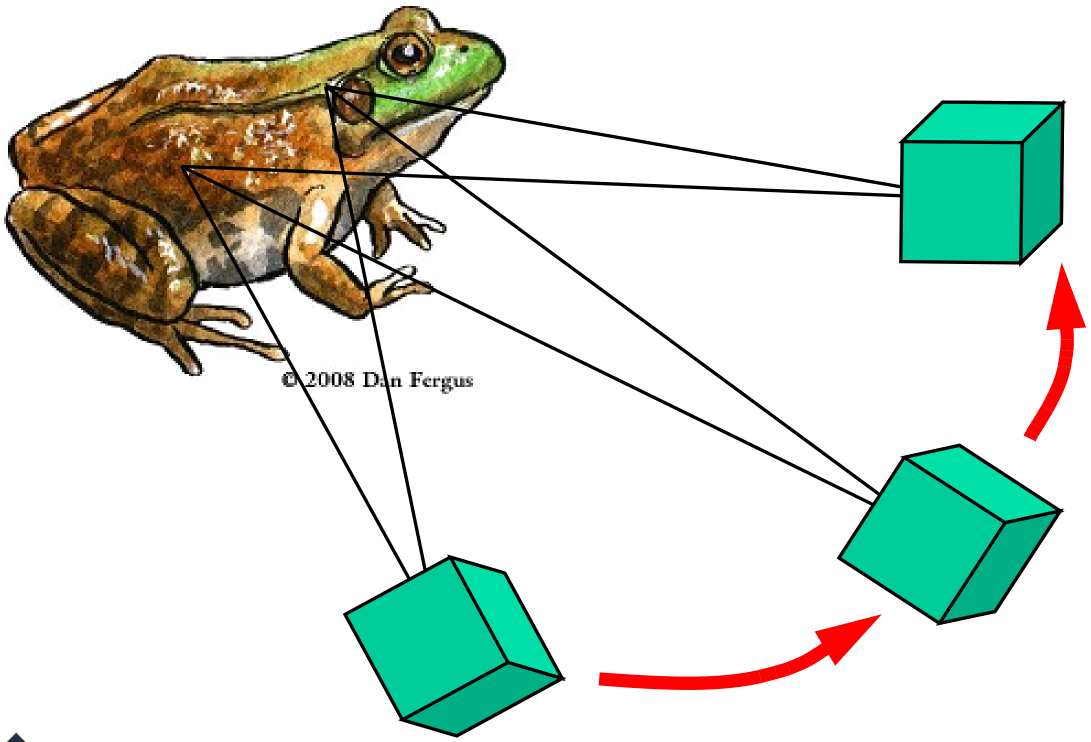
Camera calibration in computer vision (1/2)

- Classical (standard) calibration method:



Camera calibration in computer vision (2/2)

- Self-calibration method using a moving camera



Maybank & Faugeras (1992) showed that with ≥ 7 point correspondences and ≥ 3 camera positions, the extrinsic and intrinsic camera parameters can be recovered up to a scale factor.

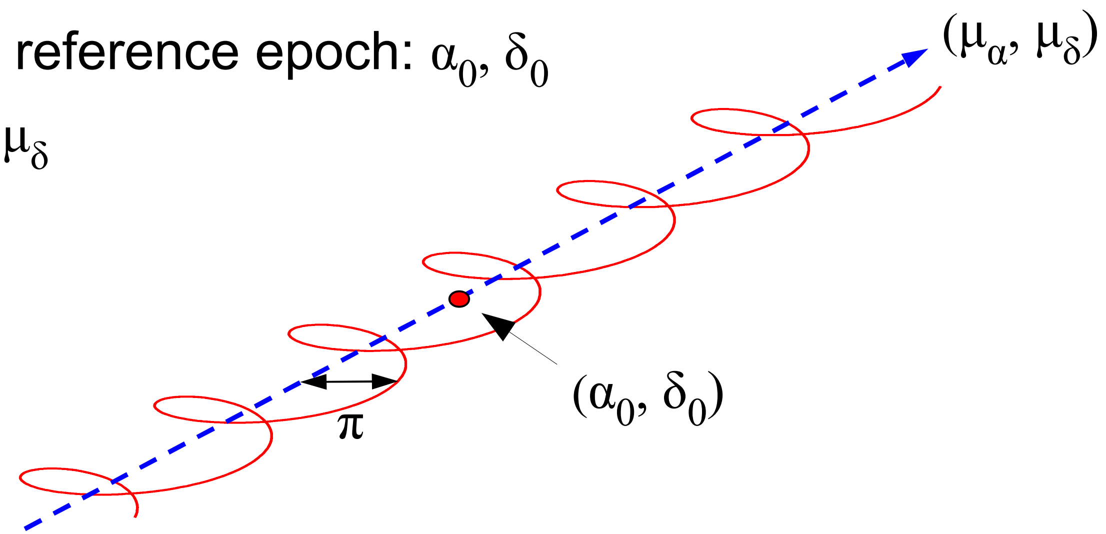
Assumptions:
- intrinsic param constant
- object constant in (*X*,*Y*,*Z*)

Gaia is a self-calibrating instrument

- Pre-launch laboratory calibration is not feasible to required accuracy
- Gaia cannot not rely on any previous astrometric measurements for its calibration
- No special "calibration observations" are made
- Most of the observations contribute to the calibration
- Instrument stability on short time scales is essential
- **The data reduction becomes very complicated because the observations are all coupled together in a big system of equations**

The global astrometric solution

- **The inertial motion of a single star on the celestial sphere is described by 5 astrometric parameters (unknowns):**
  - two for the position at a fixed reference epoch: α_0, δ_0
  - two for the angular rates: μ_α, μ_δ
  - one for the parallax: π



- **About 200 million stars can be used as ”primary sources” for the global astrometric solution**
  - this gives ~ 1000 million astrometric unknowns
  - they are linked through simultaneous measurements on the CCDs

Structure of the observation equations

$$\frac{\partial t}{\partial s_i}\Delta s_i + \frac{\partial t}{\partial a_j}\Delta a_j + \frac{\partial t}{\partial c_k}\Delta c_k \simeq t_{\mathrm{obs}} - t_{\mathrm{calc}}(s_i, a_j, c_k)$$

- Every observation t_{obs} depends on:
  - the astrometric parameters s_i of exactly one star (i)
  - the attitude parameters a_j of exactly one time interval (j)
  - the calibration parameters c_k of exactly one calibration unit (k)

Method of Least Squares (C.F. Gauss, 1795)

- The observation equations form an overdetermined system of equations

$$\mathbf{A}\mathbf{x} \simeq \mathbf{b}$$

where A is an extremely sparse matrix of size ~ $10^{11} \cdot \; 10^{9}$

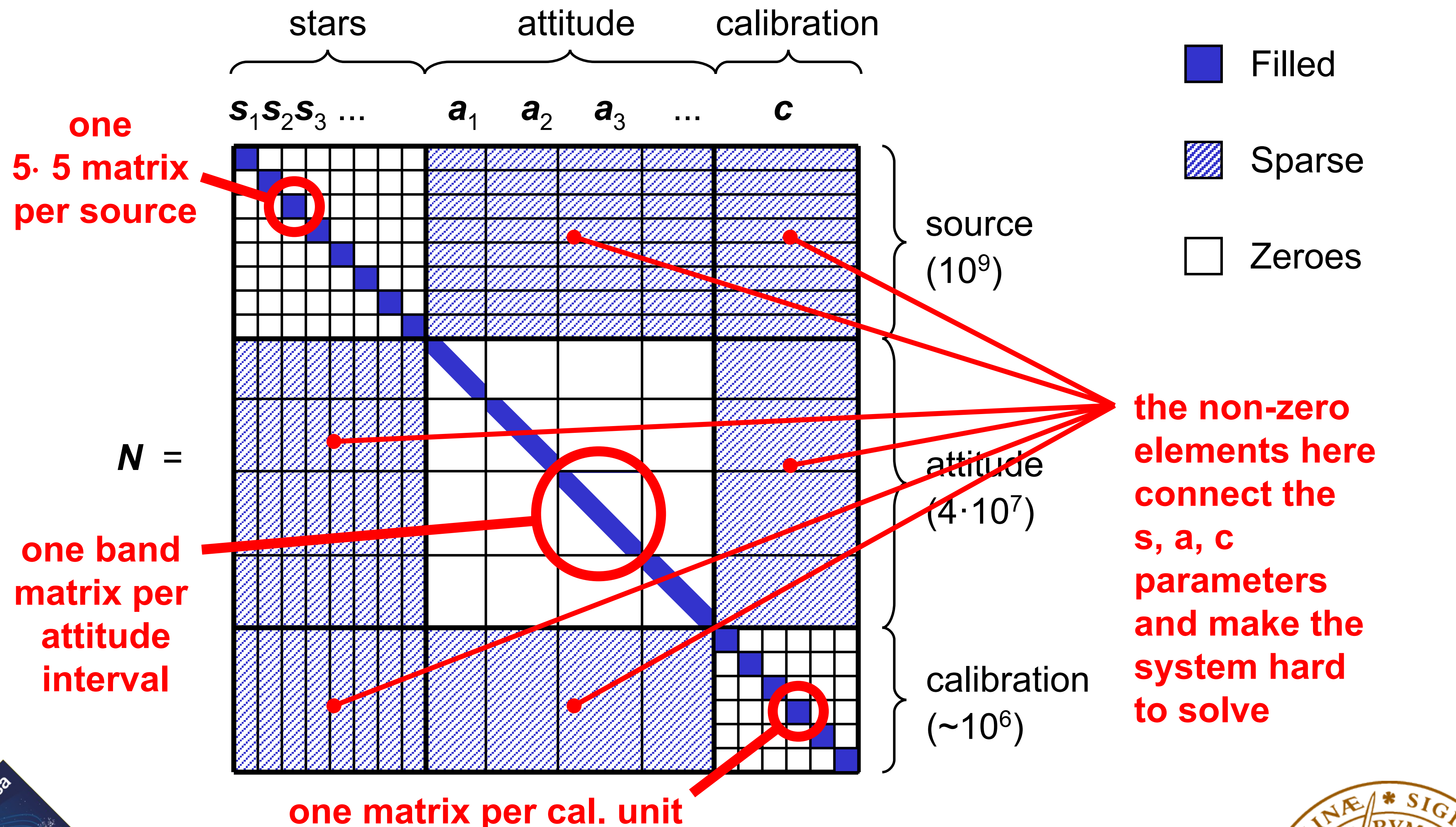
- The "best" solution minimizes the sum of the squares of the residuals:

$$\min_{\mathbf{x}} \quad |\mathbf{b} - \mathbf{A}\mathbf{x}|^2$$

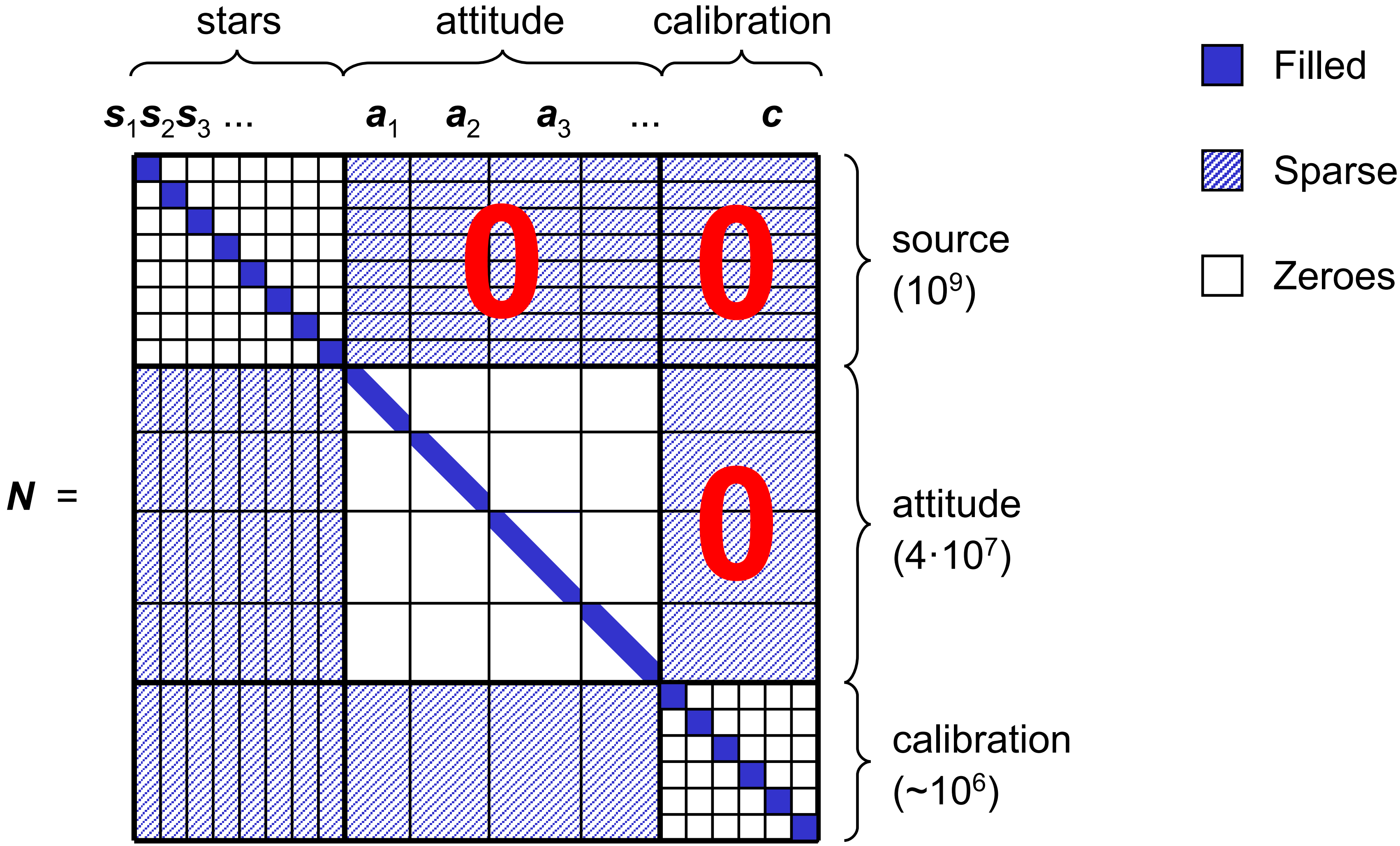
- It can be found by solving the normal equations

$$\mathbf{A}^{\mathrm{T}}\mathbf{A}\,\mathbf{x} = \mathbf{A}^{\mathrm{T}}\mathbf{b} \qquad \Leftrightarrow \qquad \mathbf{N}\mathbf{x} = \mathbf{h}$$

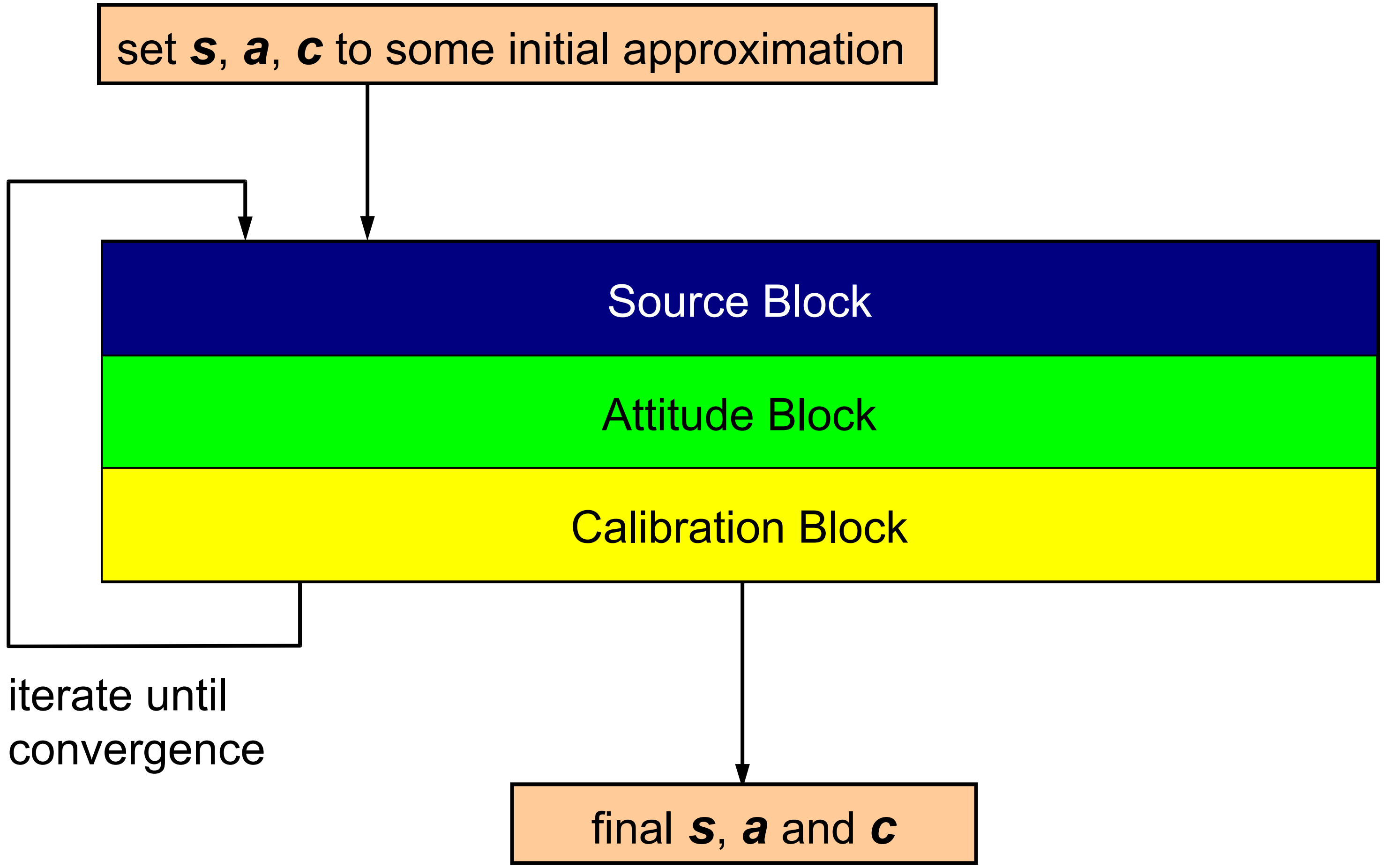
Structure of the normal equations matrix



Iterative solution method



Astrometric Global Iterative Solution (AGIS)

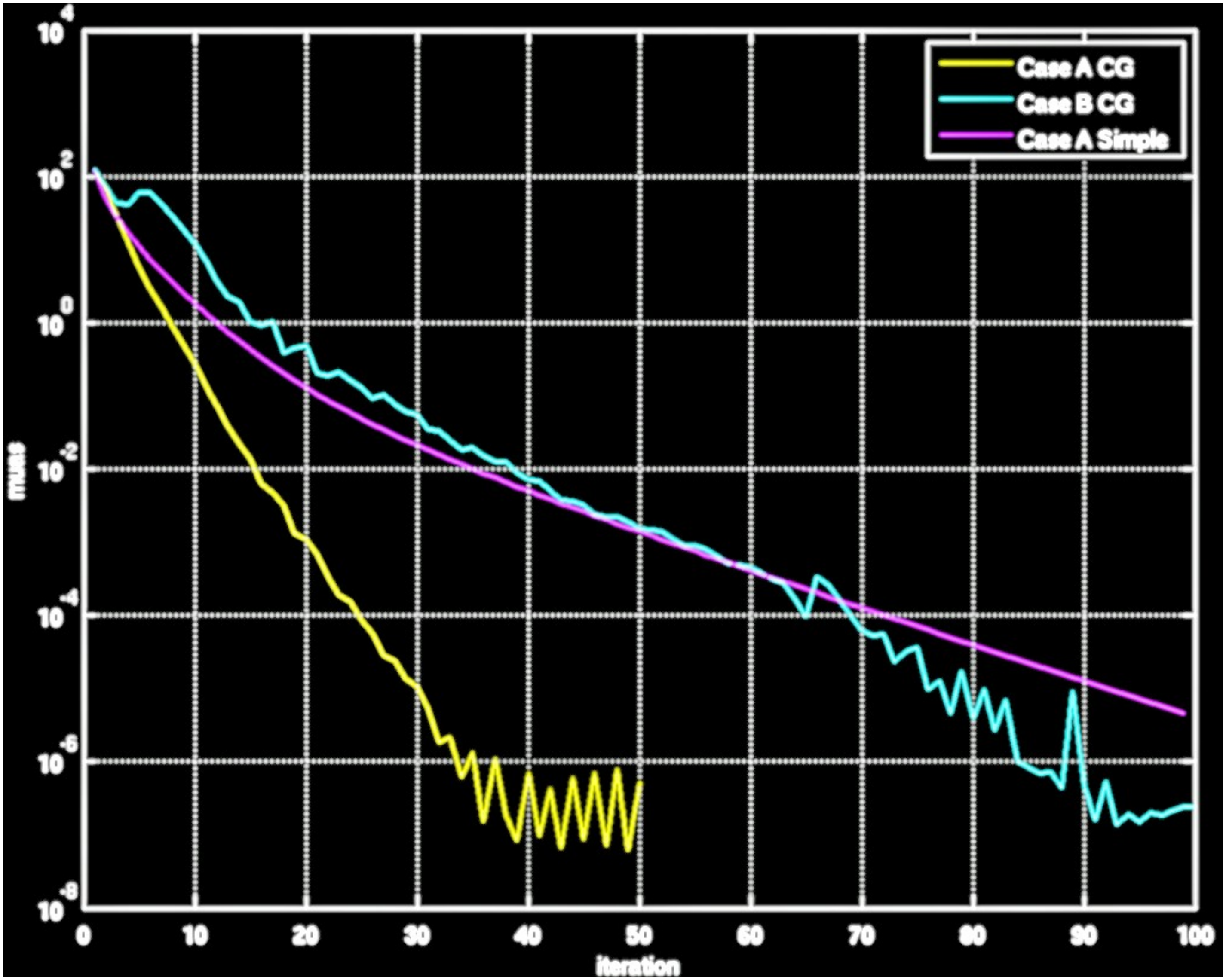


40 - 100 iterations needed for full convergence

initial errors

x 10^{-9}

double precison rounding errors

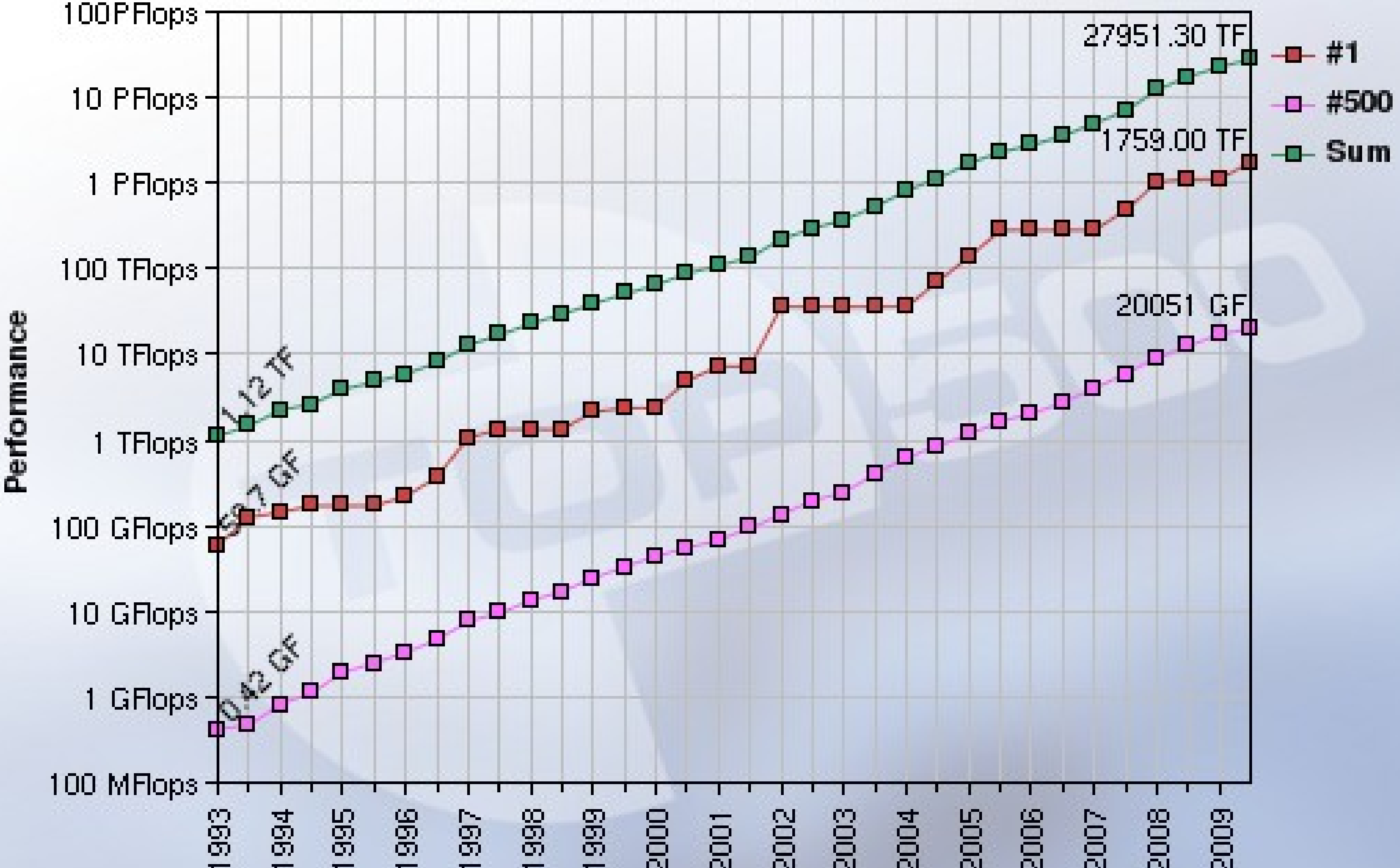


Number of floating-point operations (flop)

- The astrometric solution is currently run on a cluster ~1 Tflop/s
- Can handle about 10 million stars (1/20 of final size)
- A solution takes 10 days for convergence
- Final system (2012) ~10 Tflop/s
- Full solution ~20 days
- About 10 such solutions needed (~$2 \cdot 10^{20}$ flop)
- Total data reduction effort estimated to 10^{21} flop

We are helped by Moore's law....

Performance Development

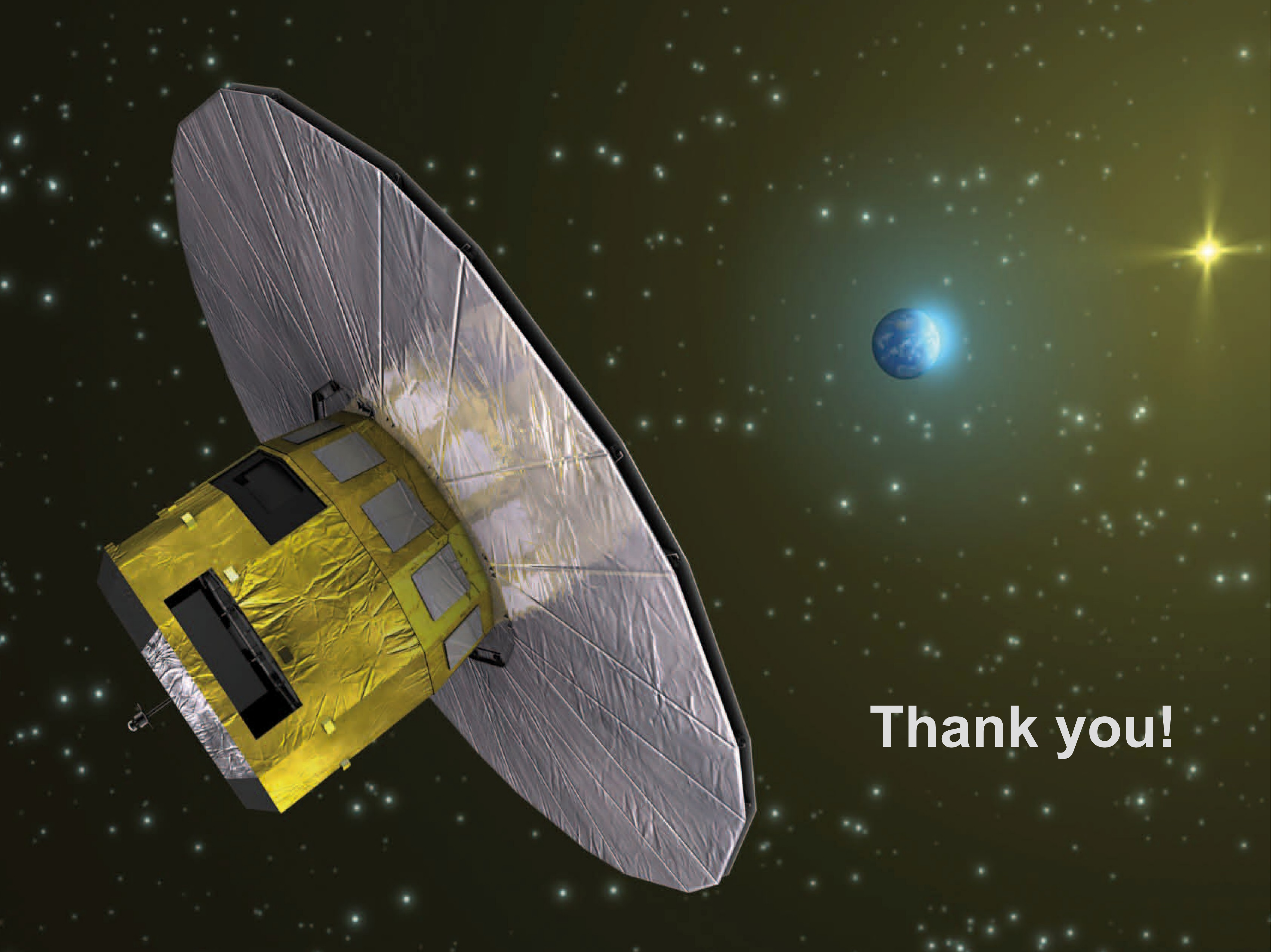


Number of floating-point operations (flop)

- The astrometric solution is currently run on a cluster ~1 Tflop/s
- Can handle about 10 million stars (1/20 of final size)
- A solution takes 10 days for convergence
- Final system (2012) ~10 Tflop/s
- Full solution ~20 days
- About 10 such solutions needed (~$2 \cdot 10^{20}$ flop)
- Total data reduction effort estimated to 10^{21} flop

This is one of the largest computational challenges in observational astronomy, but the rewards will be immense:

A 3D map of our Galaxy!


Thank you!