Tools for data analyses in Cosmology

- Aula 1 -

Camila Novaes

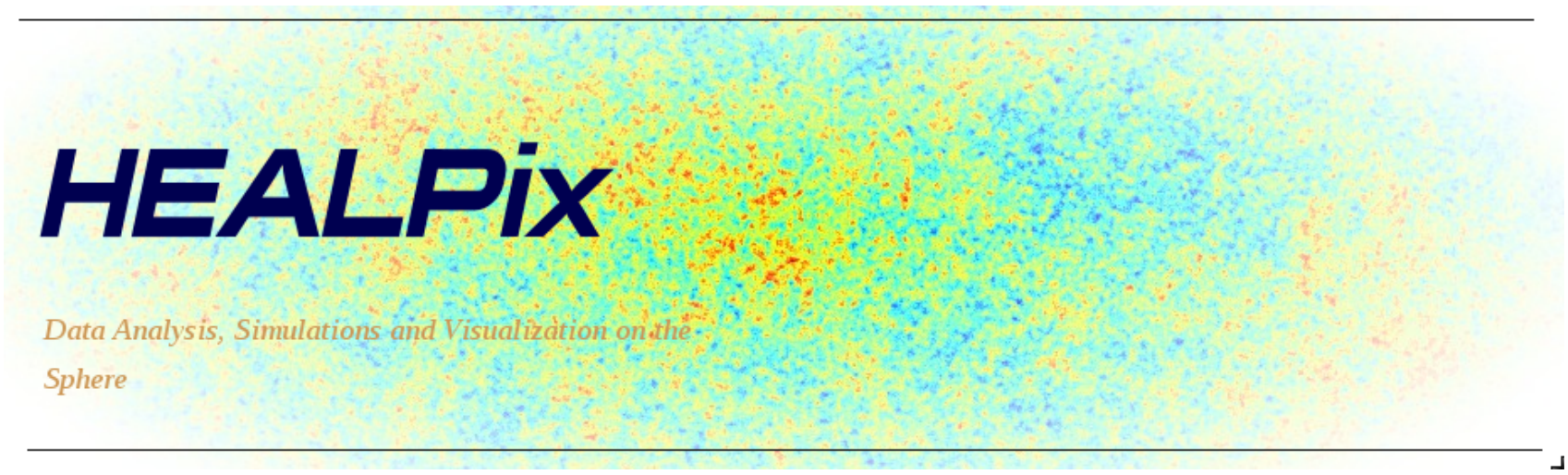
Observatório Nacional

May 9, 2017

Overview

- *Hierarchical Equal Area isoLatitude Pixelation* (HEALPix)[1].***
- *Tool for OPerations on Catalogues And Tables* (TOPCAT)[2].
- *Aladin Sky Atlas*[3]
- + some good "tips" ...

[1] http://healpix.sourceforge.net/documentation.php
[2] http://www.star.bris.ac.uk/ mbt/topcat/#docs
[3] http://aladin.u-strasbg.fr/


HEALPix
Data Analysis, Simulations and Visualization on the Sphere

The Astrophysical Journal, 622:759–771, 2005 April 1

HEALPix: A FRAMEWORK FOR HIGH-RESOLUTION DISCRETIZATION AND FAST ANALYSIS OF DATA DISTRIBUTED ON THE SPHERE

K. M. Górski,[1,2] E. Hivon,[3] A. J. Banday,[4] B. D. Wandelt,[5,6] F. K. Hansen,[7]
M. Reinecke,[4] and M. Bartelmann[8]

Received 2004 September 21; accepted 2004 December 10

ABSTRACT

HEALPix—the Hierarchical Equal Area isoLatitude Pixelization—is a versatile structure for the pixelization of data on the sphere. An associated library of computational algorithms and visualization software supports fast scientific applications executable directly on discretized spherical maps generated from very large volumes of astronomical data. Originally developed to address the data processing and analysis needs of the present generation of cosmic microwave background experiments (e.g., BOOMERANG, *WMAP*), HEALPix can be expanded to meet many of the profound challenges that will arise in confrontation with the observational output of future missions and experiments, including, e.g., *Planck*, *Herschel*, *SAFIR*, and the Beyond Einstein inflation probe. In this paper we consider the requirements and implementation constraints on a framework that simultaneously enables an efficient discretization with associated hierarchical indexation and fast analysis/synthesis of functions defined on the sphere. We demonstrate how these are explicitly satisfied by HEALPix.

Subject headings: cosmic microwave background — cosmology: observations — methods: statistical

Motivation

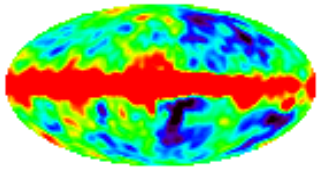
1. INTRODUCTION

Advanced detectors in modern astronomy generate data at huge rates over many wavelengths. Of particular interest to us are those data sets that accumulate measurements distributed on the entire sky, or a considerable fraction thereof. Typical examples include radio, cosmic microwave background (CMB), submillimeter, infrared, X-ray, and gamma-ray sky maps of diffuse emission, and full-sky or wide-area surveys of extragalactic objects. Together with this wealth of gathered information comes an inevitable increase in complexity for data reduction and science extraction. In this paper we are focused on those issues related to the distinctive nature of the spherical spatial domain over which the data reside. Our original motivations arose from work related to the measurement and interpretation of the CMB anisotropy. The growing complexity of the associated science extraction problem can be illustrated by the transition between

1. global analysis problems: harmonic decomposition, estimation of the power spectrum, and higher order measures of spatial correlations;
2. real space morphological analyses, object detection, identification, and characterization;
3. the simulation of models of the primary and foreground sky signals to study instrument performance and calibrate foreground separation and statistical inference methods; and
4. spatial and/or spectral cross-correlation with external data sets.

These tasks, and many others, necessitate a careful definition of the data models and proper construction of the mathematical framework for data analysis such that the algorithmic and computing time requirements can be satisfied in order to achieve the successful and timely scientific interpretation of the observations. A particular method of addressing some of these issues is

Motivation



SKY

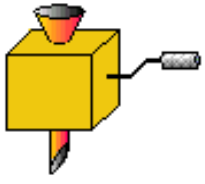
MEASUREMENT

RAW DATA

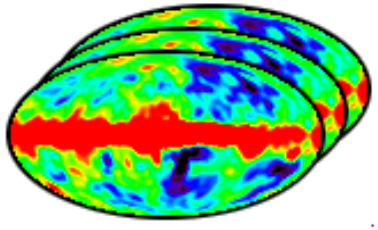
CLEANING

Pixel 1	Pixel 2	ΔT
6422347	6443428	-454.841
3141592	2718281	141.421
8454543	9345593	654.766
1004356	8345388	-305.567
...	...	...

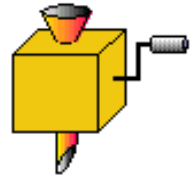
TIME-ORDERED DATA



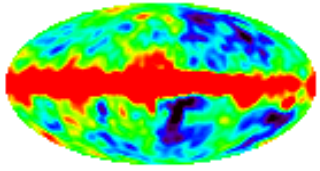
MAPMAKING



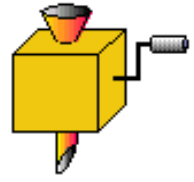
MULTI-FREQUENCY MAPS



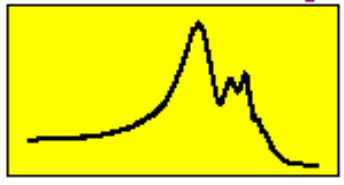
FOREGROUND REMOVAL



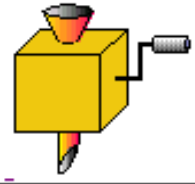
SKY MAP



POWER ESTIMATION



POWER SPECTRUM

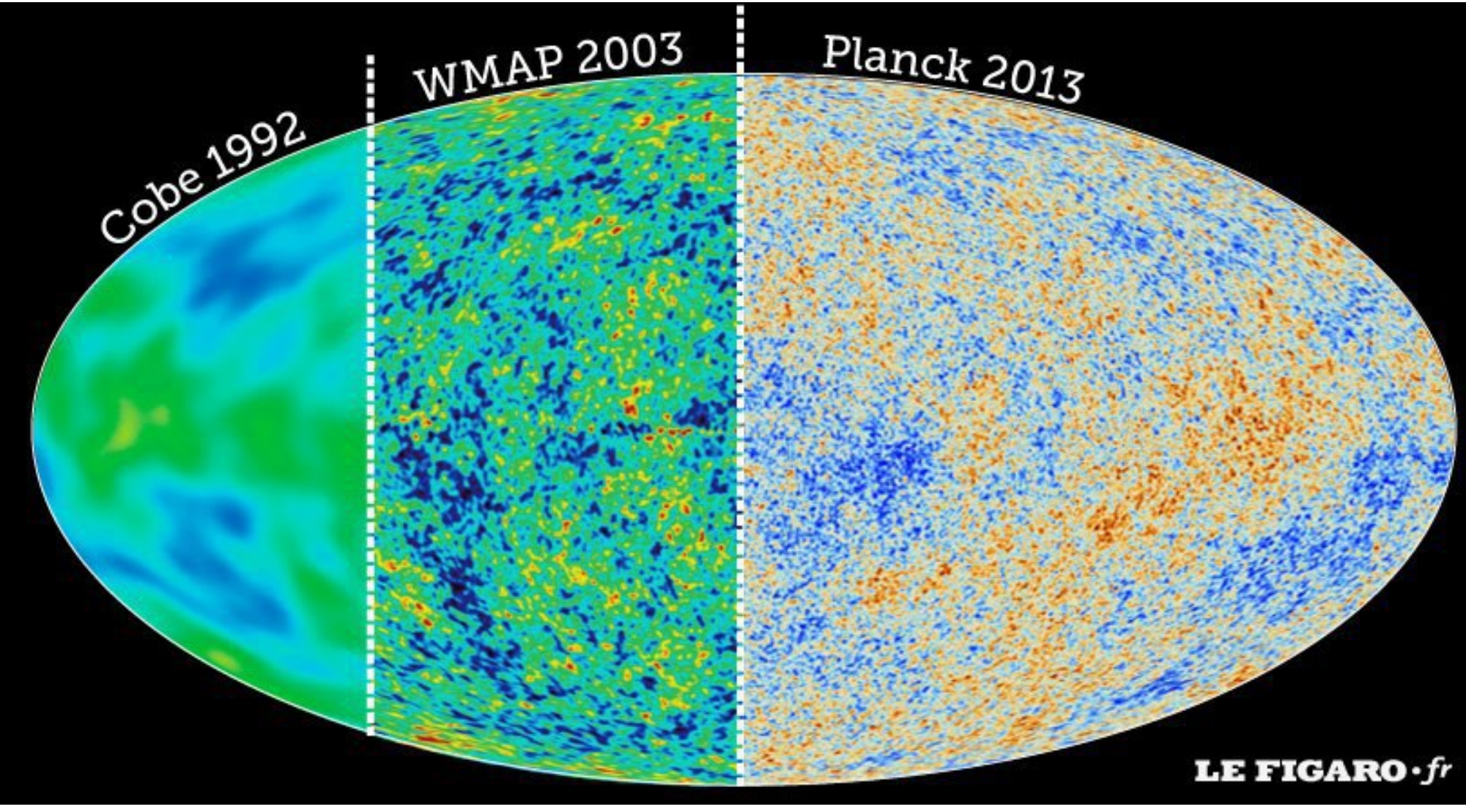


MODEL TESTING

Ω, Ω_b, Λ, τ, h
n, n_T, Q, T/S

PARAMETER ESTIMATES

Motivation

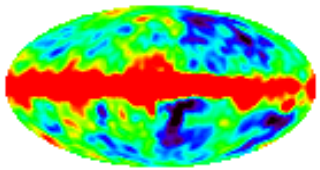


Motivation

To illustrate:

- COBE (early 1990s): 3 freqs., 7° FWHM, $\rightarrow \sim$ 6000 pixels,
- BOOMERANG (late 1990s): 4 freqs., 12' FWHM, $\rightarrow \sim 2 \times 10^5$ pixels ... PARTIAL SKY,
- WMAP (early 2000s): 5 freq., 14' FWHM, $\rightarrow \sim 3 \times 10^6$ pixels,
- Planck (2010s): 9 freqs., 5' FWHM, $\rightarrow \sim 50 \times 10^6$ pixels.

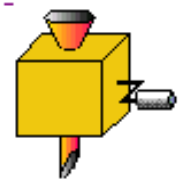
Motivation



SKY

MEASUREMENT

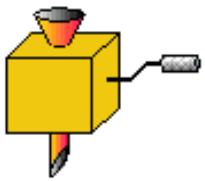
RAW DATA



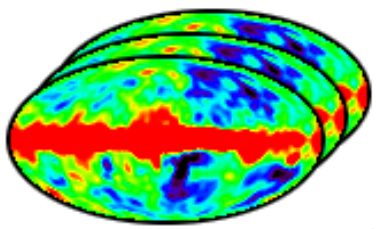
CLEANING

Pixel 1	Pixel 2	ΔT
6422347	6443428	-454.841
3141592	2718281	141.421
8454543	9345593	654.766
1004356	8345388	-305.567
...	...	...

TIME-ORDERED DATA

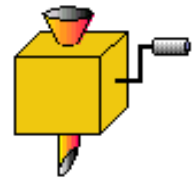


MAPMAKING

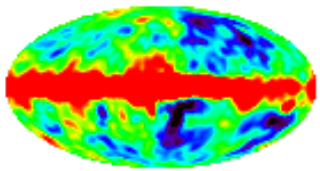


MULTI-FREQUENCY MAPS

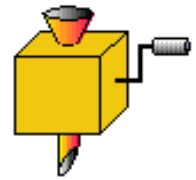
HEALPix



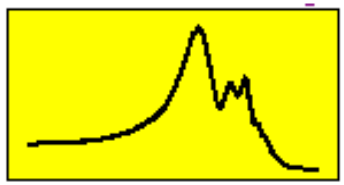
FOREGROUND REMOVAL



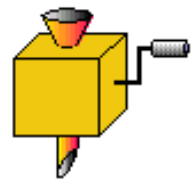
SKY MAP



POWER ESTIMATION



POWER SPECTRUM



MODEL TESTING

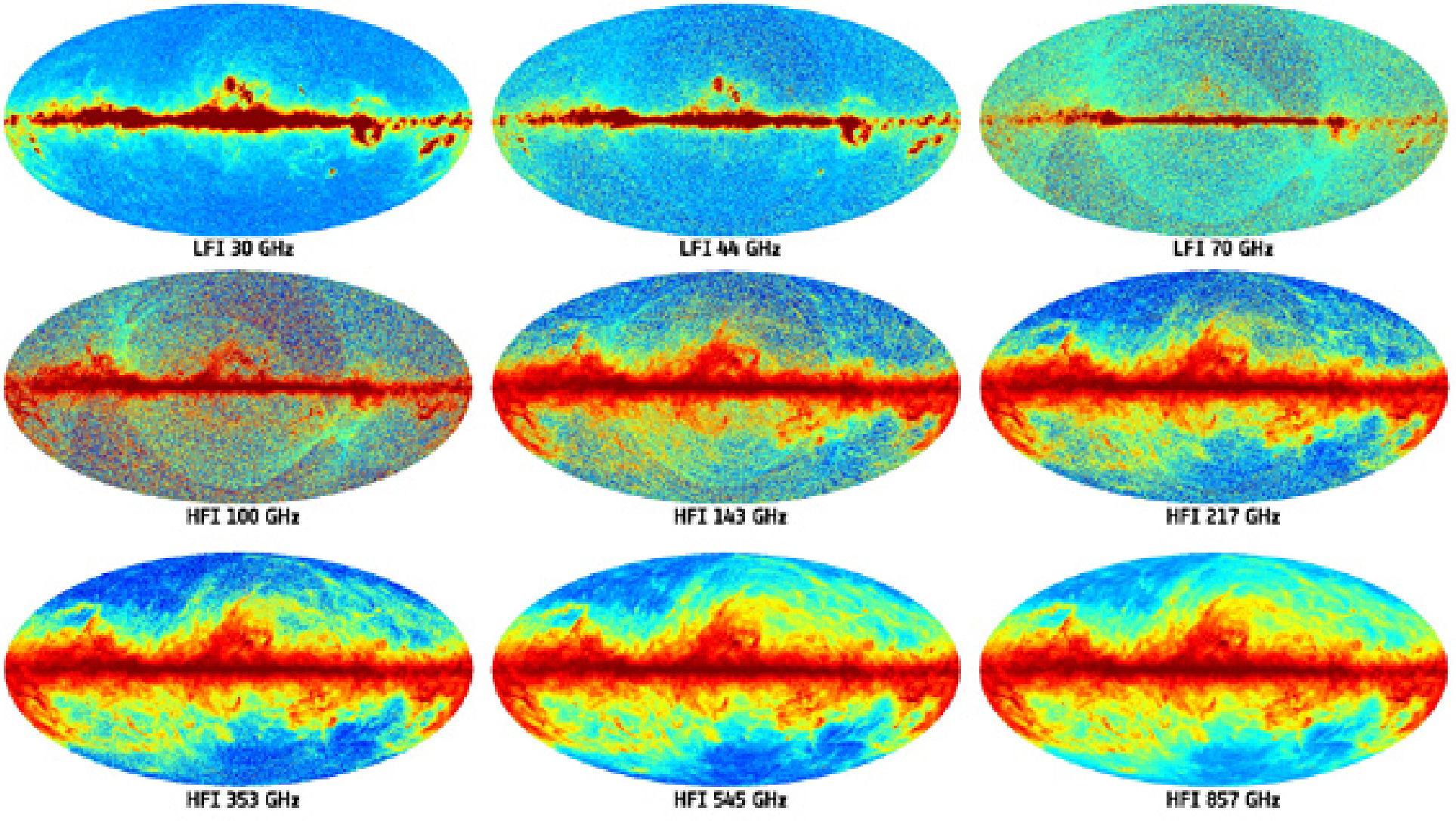
Ω, Ω_b, Λ, τ, h
n, n_T, Q, T/S

PARAMETER ESTIMATES

HEALPix

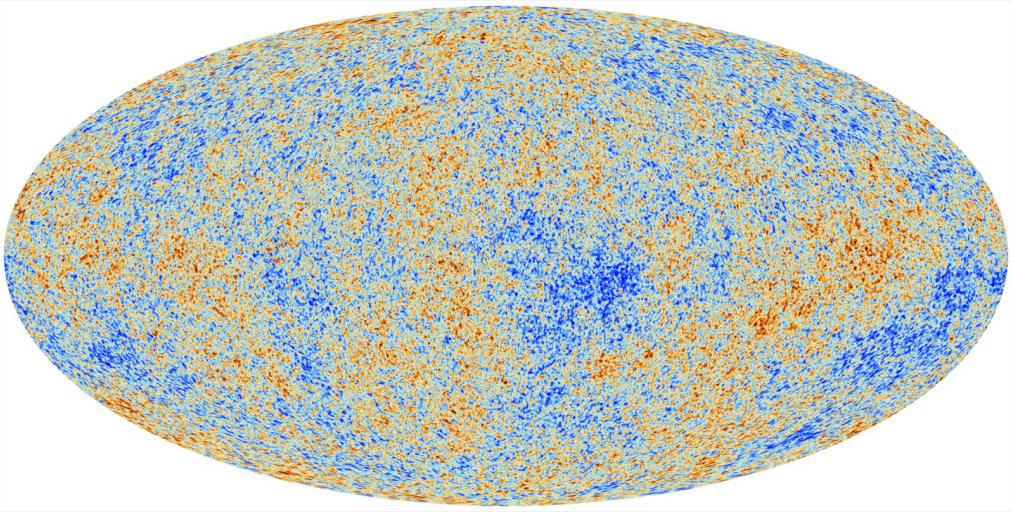
Motivation

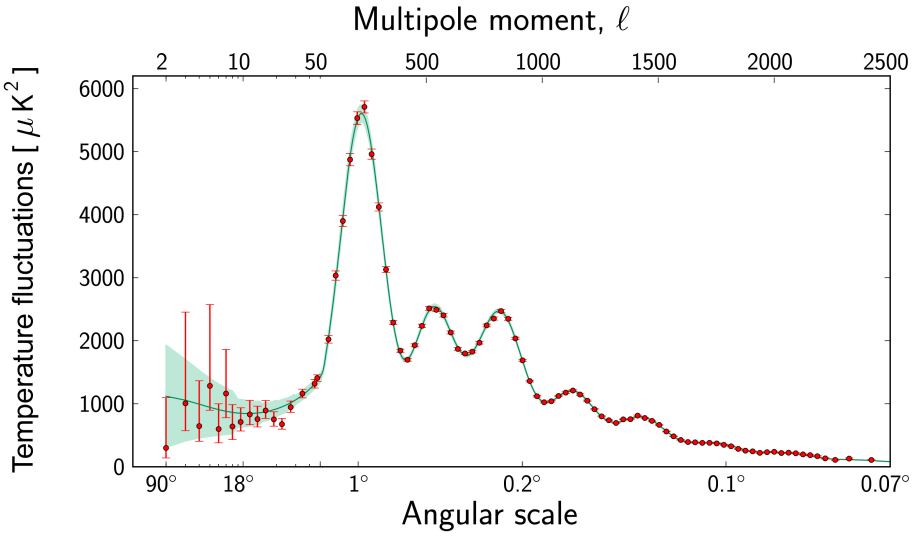
Planck data



Motivation

Planck data





Motivation

1. INTRODUCTION

Advanced detectors in modern astronomy generate data at huge rates over many wavelengths. Of particular interest to us are those data sets that accumulate measurements distributed on the entire sky, or a considerable fraction thereof. Typical examples include radio, cosmic microwave background (CMB), submillimeter, infrared, X-ray, and gamma-ray sky maps of diffuse emission, and full-sky or wide-area surveys of extragalactic objects. Together with this wealth of gathered information comes an inevitable increase in complexity for data reduction and science extraction. In this paper we are focused on those issues related to the distinctive nature of the spherical spatial domain over which the data reside. Our original motivations arose from work related to the measurement and interpretation of the CMB anisotropy. The growing complexity of the associated science extraction problem can be illustrated by the transition between

1. global analysis problems: harmonic decomposition, estimation of the power spectrum, and higher order measures of spatial correlations;
2. real space morphological analyses, object detection, identification, and characterization;
3. the simulation of models of the primary and foreground sky signals to study instrument performance and calibrate foreground separation and statistical inference methods; and
4. spatial and/or spectral cross-correlation with external data sets.

These tasks, and many others, necessitate a careful definition of the data models and proper construction of the mathematical framework for data analysis such that the algorithmic and computing time requirements can be satisfied in order to achieve the successful and timely scientific interpretation of the observations. A particular method of addressing some of these issues is

→ **CMB** ←

Motivation

1. INTRODUCTION

Advanced detectors in modern astronomy generate data at huge rates over many wavelengths. Of particular interest to us are those data sets that accumulate measurements distributed on the entire sky, or a considerable fraction thereof. Typical examples include radio, cosmic microwave background (CMB), submillimeter, infrared, X-ray, and gamma-ray sky maps of diffuse emission, and full-sky or wide-area surveys of extragalactic objects. Together with this wealth of gathered information comes an inevitable increase in complexity for data reduction and science extraction. In this paper we are focused on those issues related to the distinctive nature of the spherical spatial domain over which the data reside. Our original motivations arose from work related to the measurement and interpretation of the CMB anisotropy. The growing complexity of the associated science extraction problem can be illustrated by the transition between

1. global analysis problems: harmonic decomposition, estimation of the power spectrum, and higher order measures of spatial correlations;
2. real space morphological analyses, object detection, identification, and characterization;
3. the simulation of models of the primary and foreground sky signals to study instrument performance and calibrate foreground separation and statistical inference methods; and
4. spatial and/or spectral cross-correlation with external data sets.

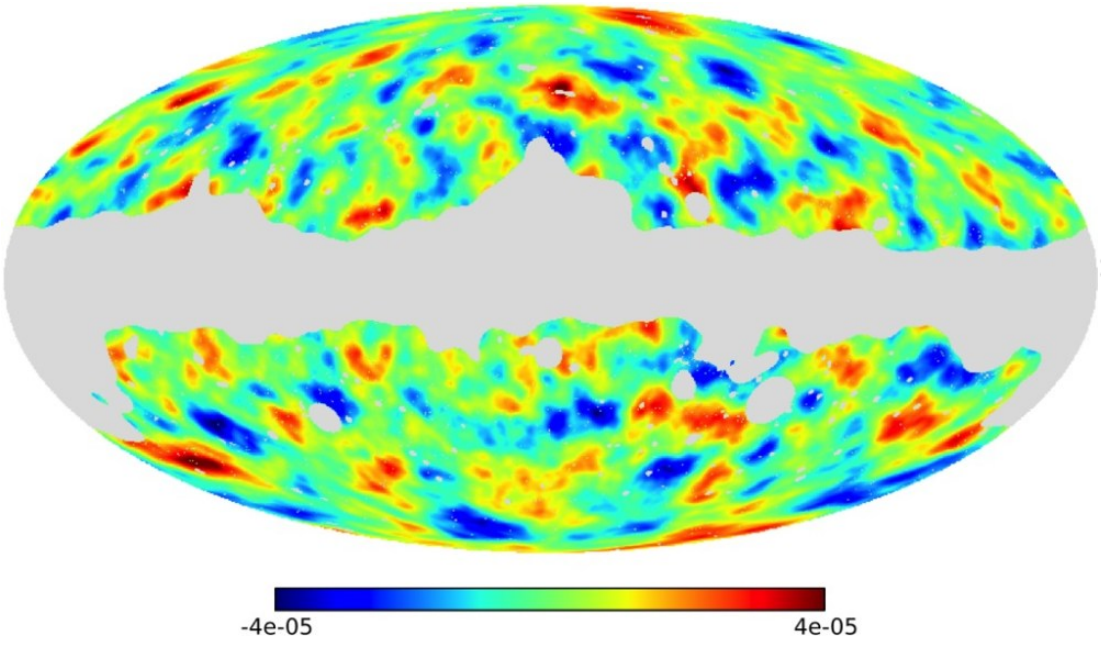
These tasks, and many others, necessitate a careful definition of the data models and proper construction of the mathematical framework for data analysis such that the algorithmic and computing time requirements can be satisfied in order to achieve the successful and timely scientific interpretation of the observations. A particular method of addressing some of these issues is

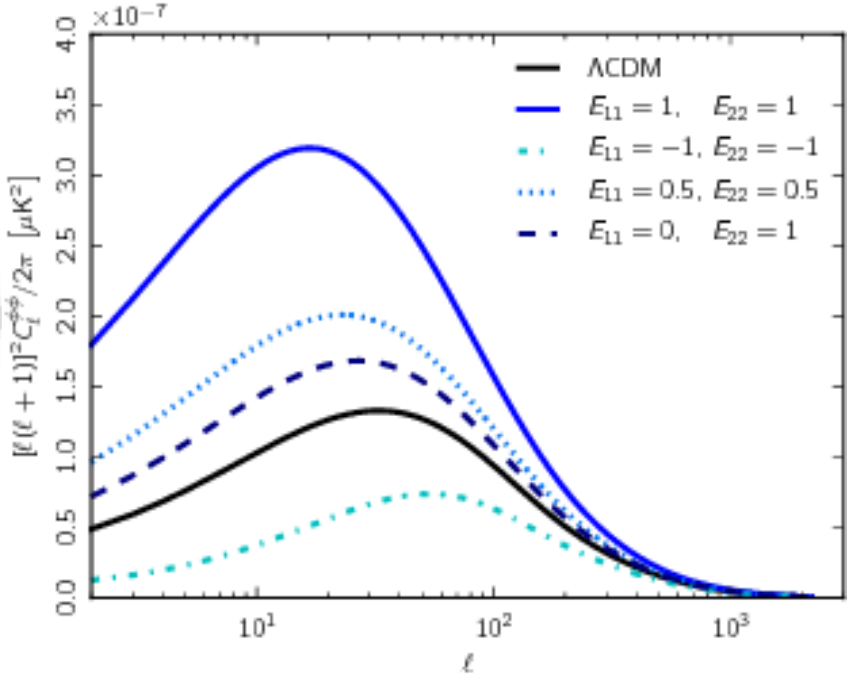
→ CMB ←

What am I doing here?

Motivation

Planck lensing potential map

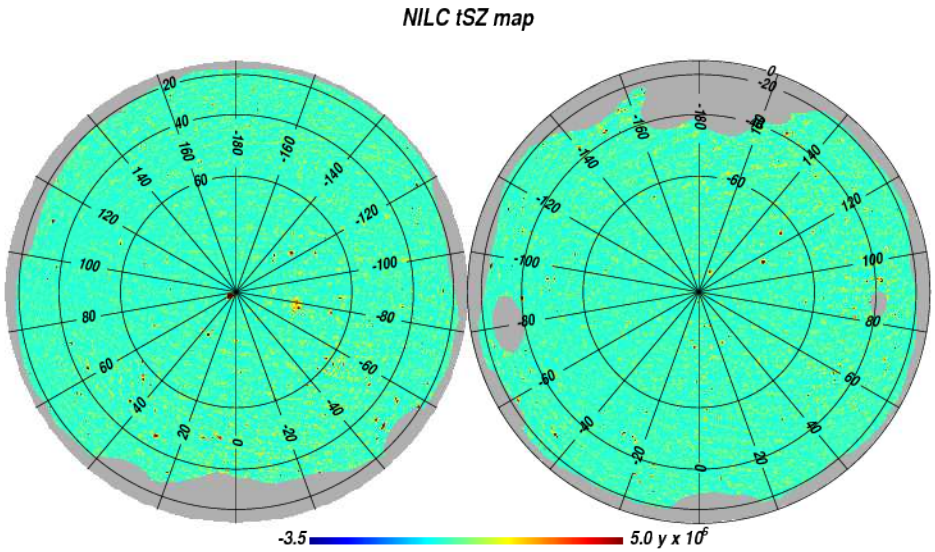




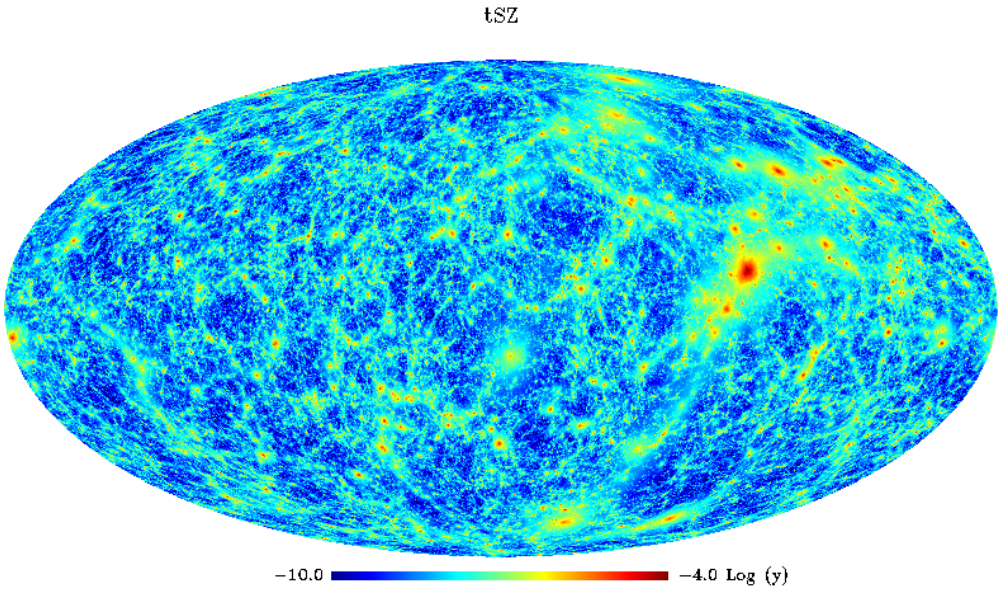
Motivation

Sunyaev-Zeldovich effect

Planck data

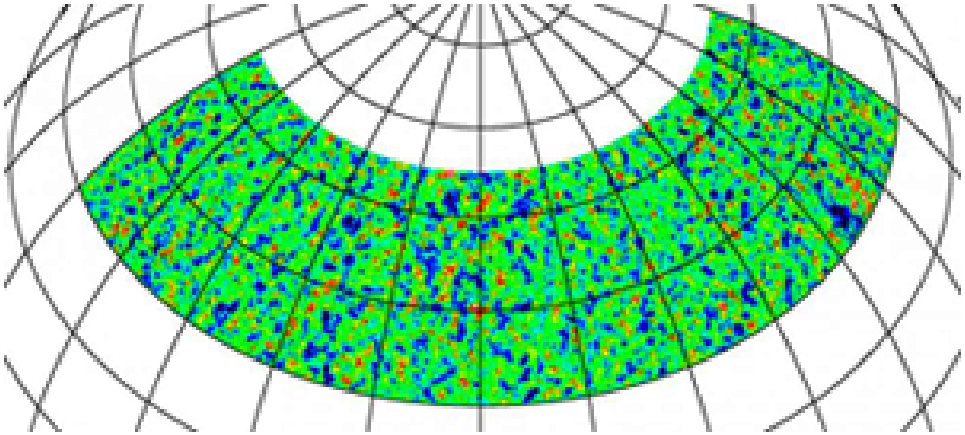


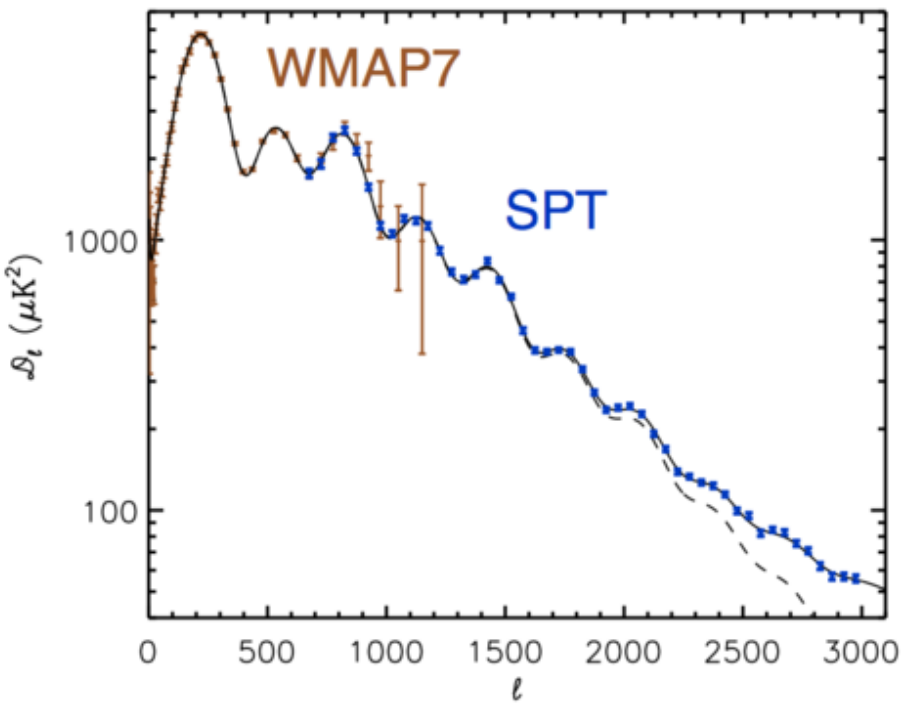
Simulation



Motivation

South Pole Telescope

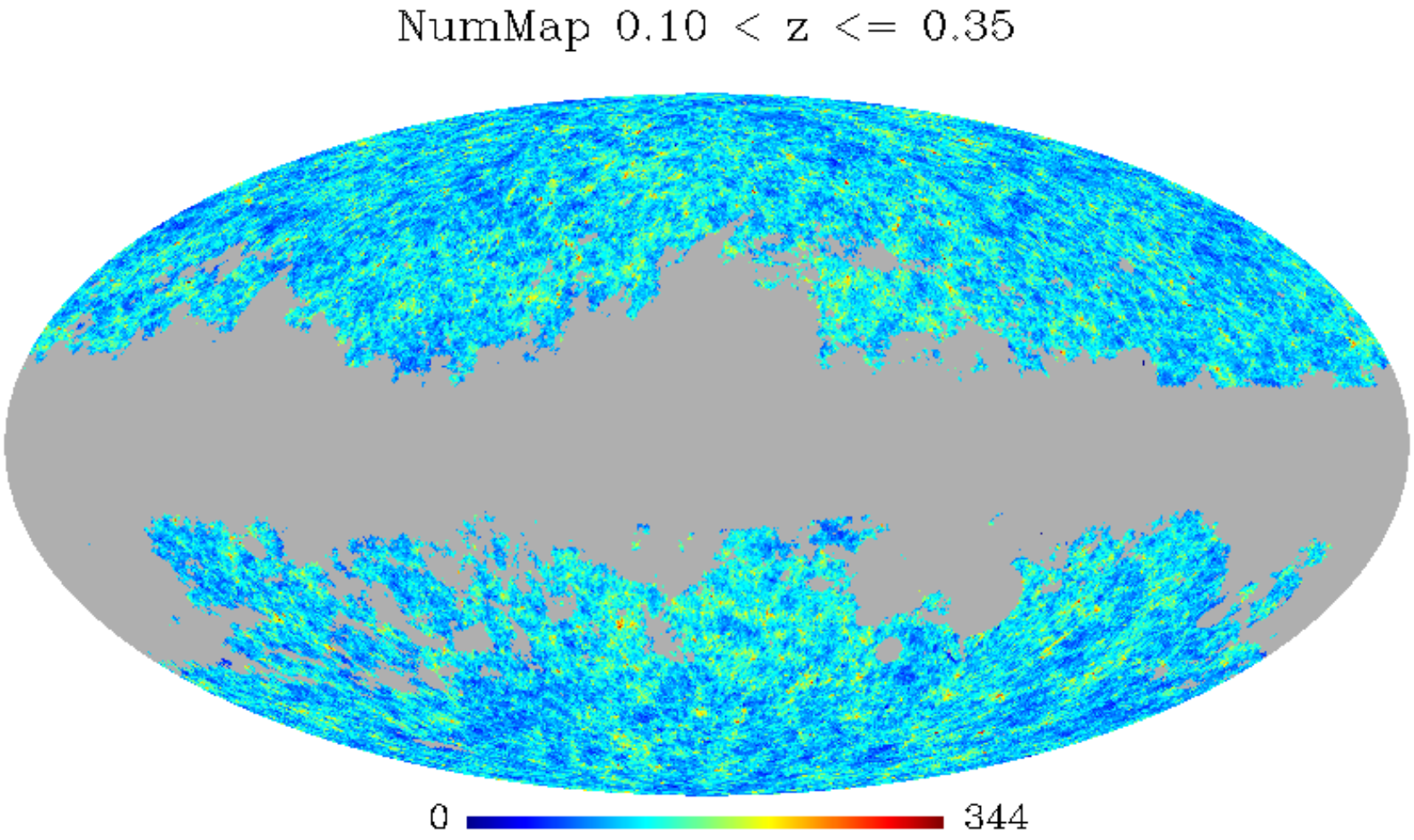




Resolution: 1.58 – 0.44 arcmin.

Motivation

WISE x SuperCOSMOS photometric redshift catalog



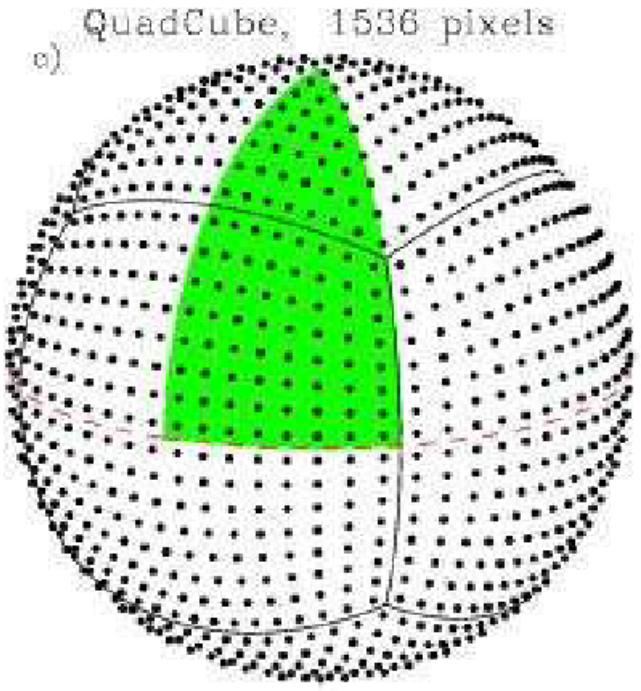
Fundamentals

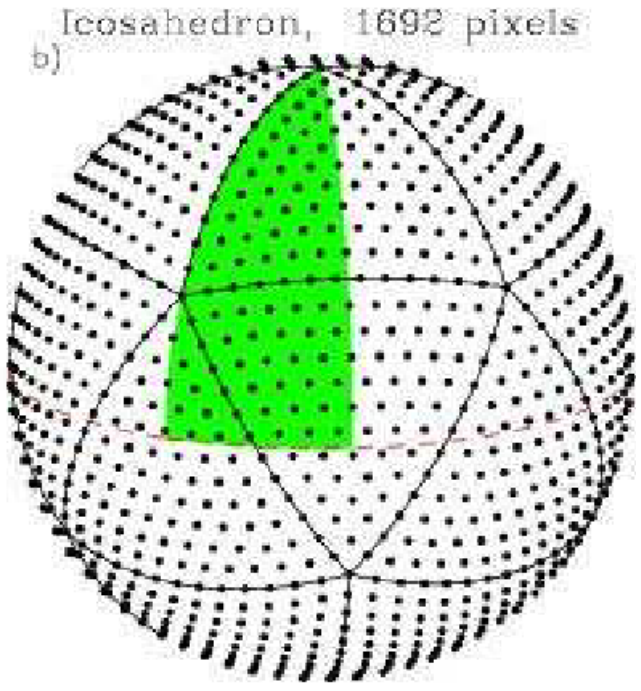
HEALPix: **H**ierarchical **E**qual **A**rea iso**L**atitude **Pix**elation

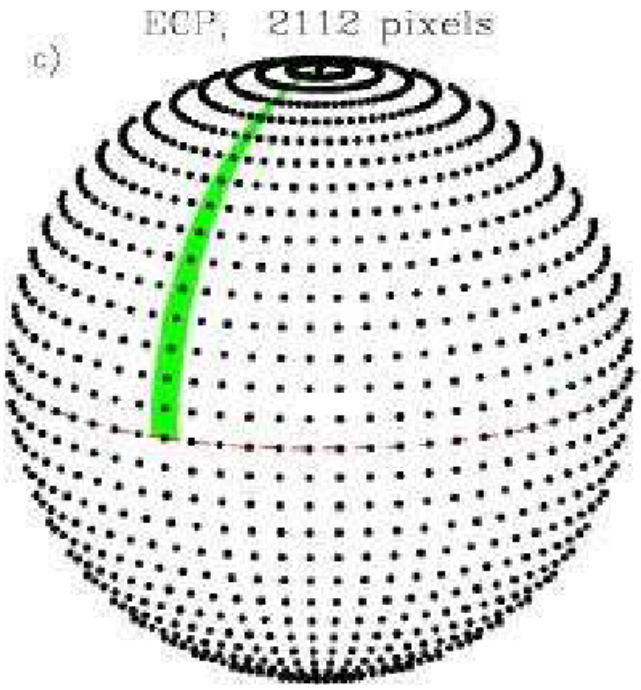
Essential properties:

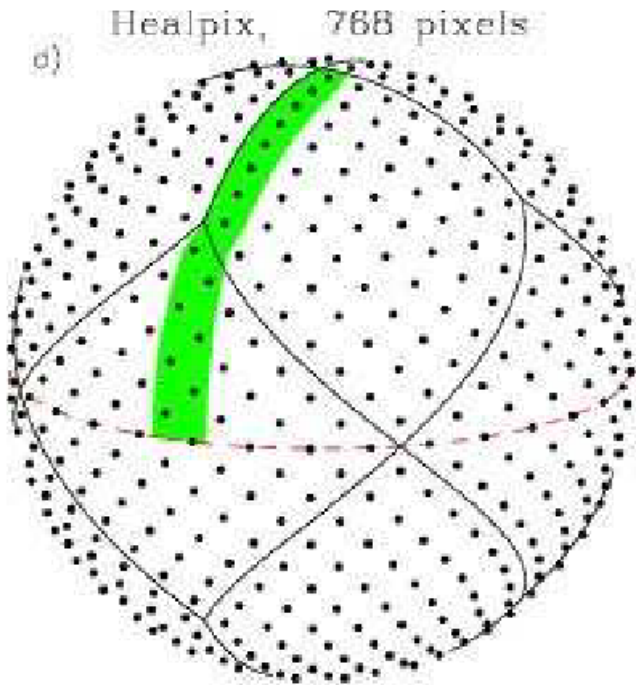
1. The sphere is **hierarchically** tessellated into curvilinear quadrilaterals (12 base pixels): to support efficient data handling.
   - e.g.: region definitions, data selection, map resolution changes, etc.
2. **Areas** of all pixels at a given resolution are identical.
3. Pixels are distributed on lines of **constant latitude**: to support effective harmonic analysis.

Other possible tessellations: attempts



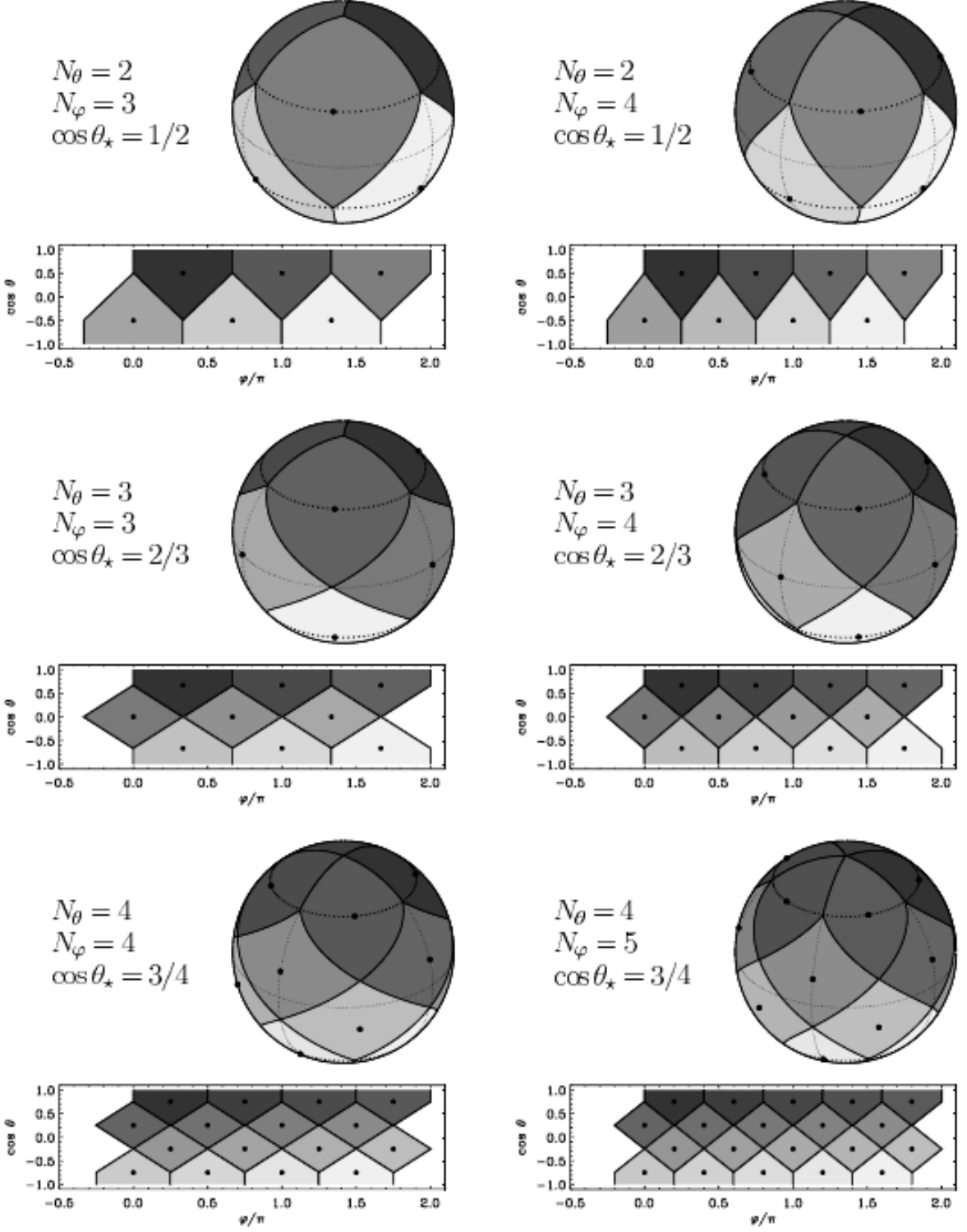






Fundamentals

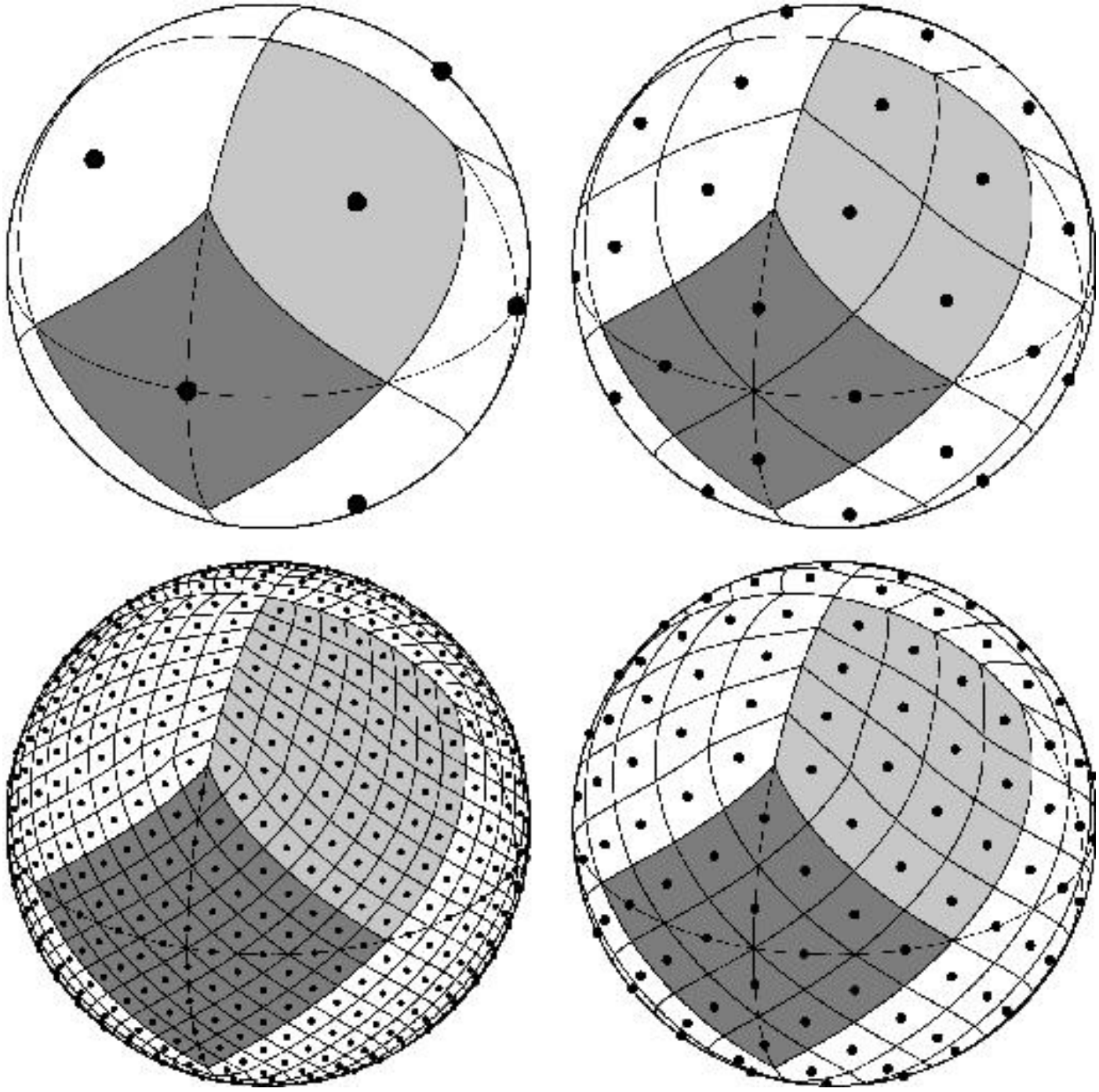
Several possible equal-area isolatitude tessellations of the sphere, which can support a hierarchical tree for the further subdivision of each large base-resolution pixel.



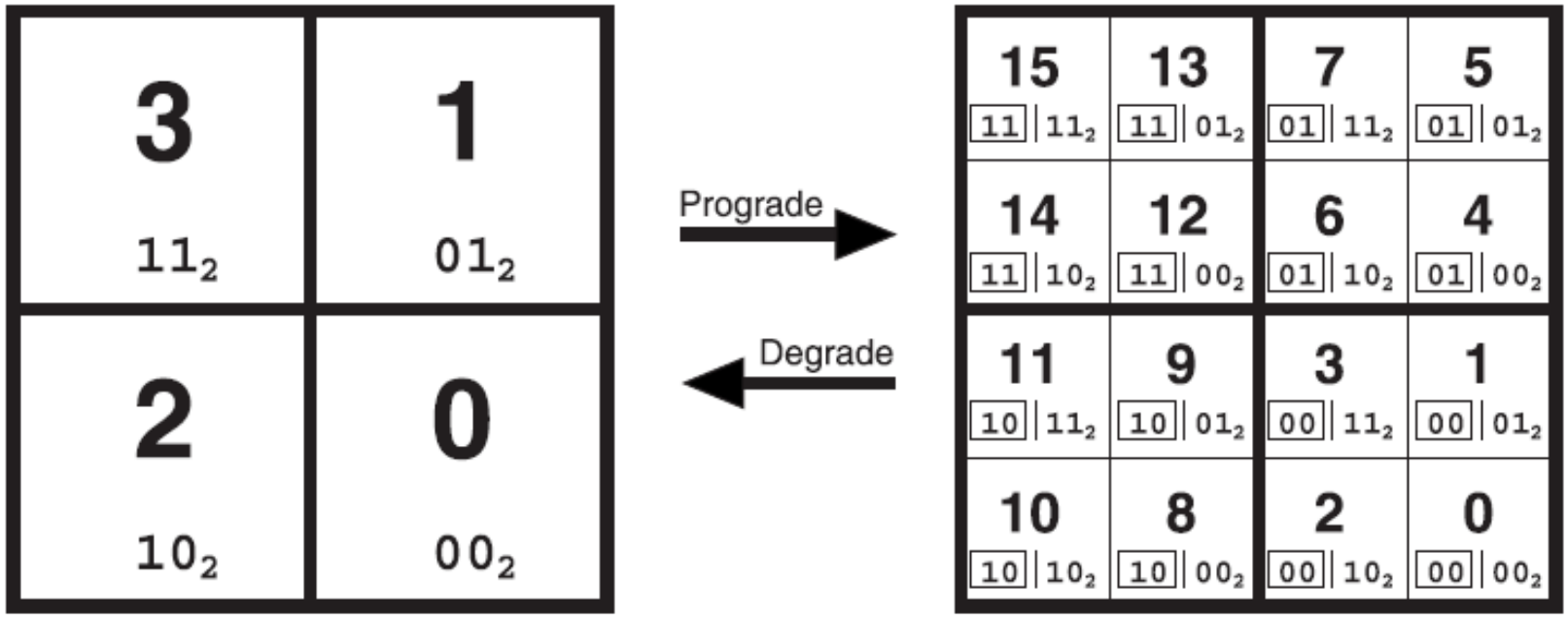
Fundamentals

of pixels $\rightarrow$

12, 48,
768, 192



Fundamentals



The math of the pixelisation

The resolution of the grid is expressed by the parameter $N_{\rm side}$ (power of 2: 2k).

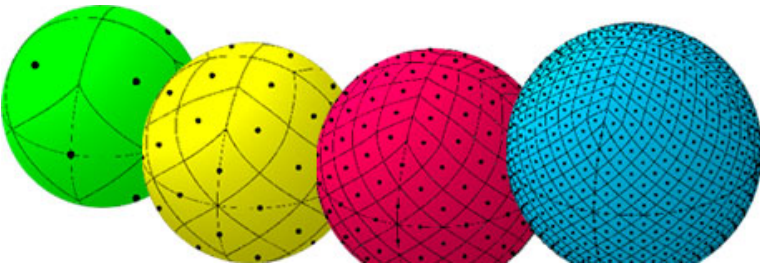
↓

Total number of pixels:

$$N_{\rm pix} = 12 \times N_{\rm side}^2$$

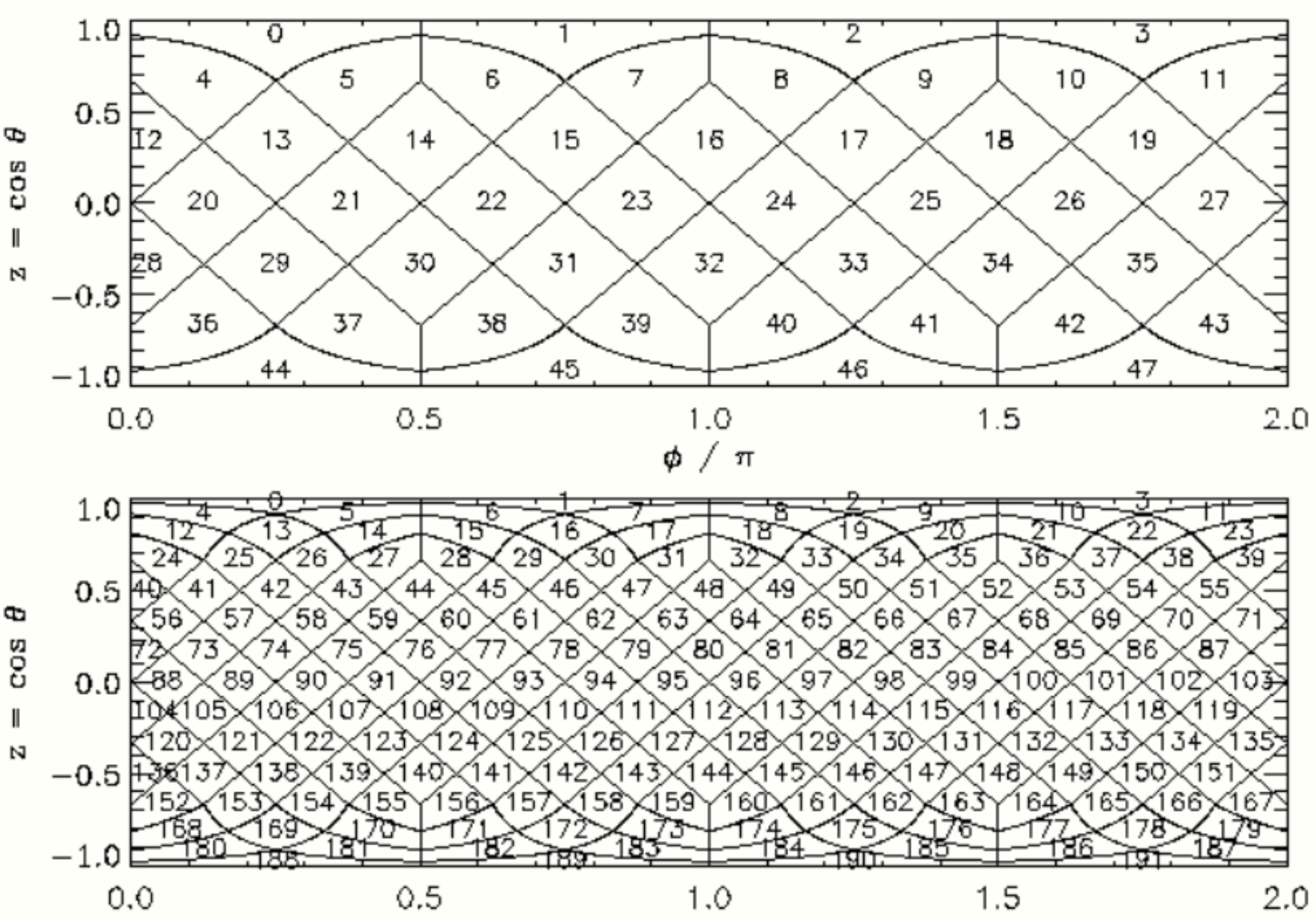
Area: $\Omega = 4\pi / N_{\rm pix}$.

All pixel centers are located on $N_{\rm ring} = 4 \times N_{\rm side} - 1$ rings of constant latitude.

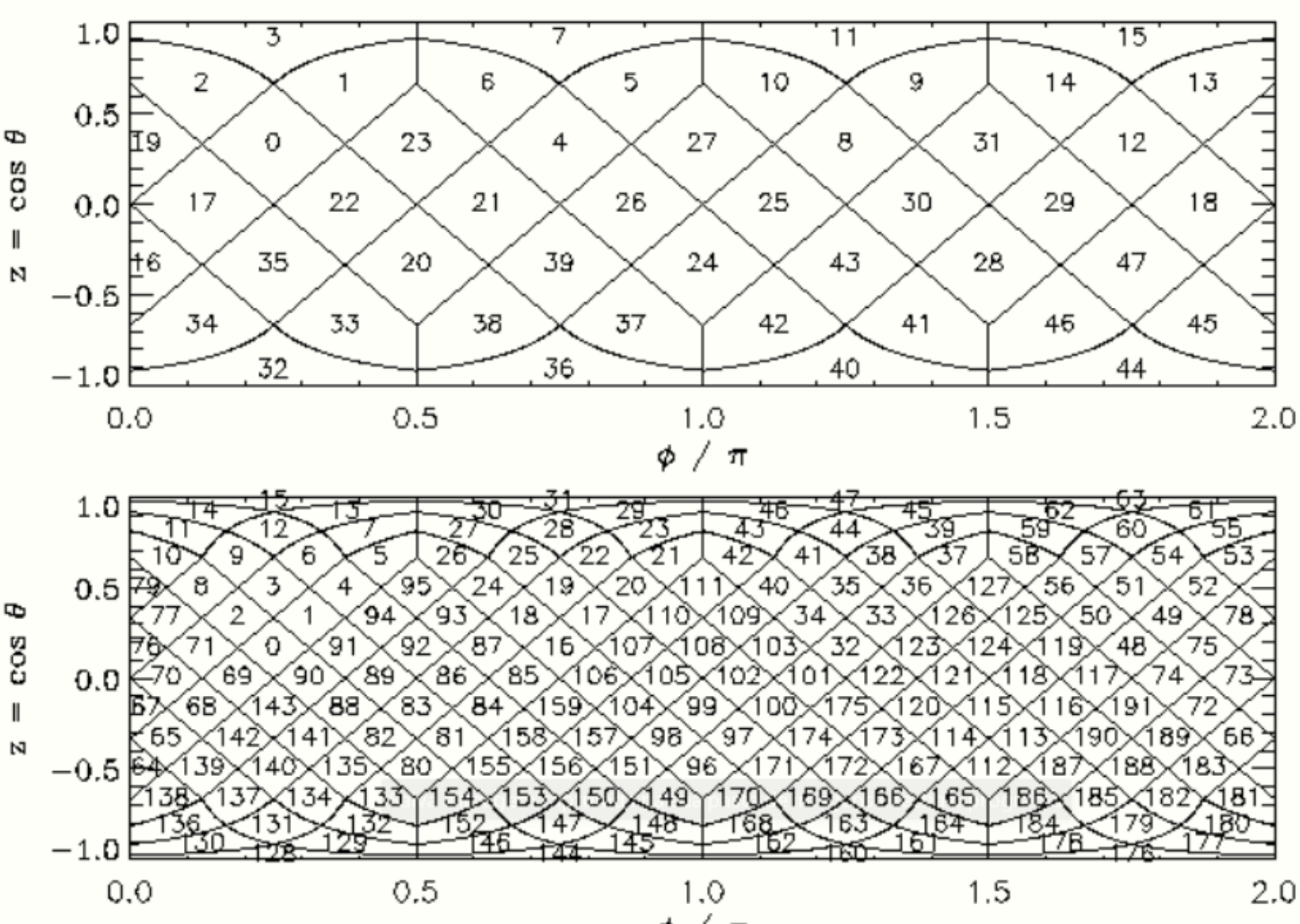


k	$N_{\rm side} = 2^k$	$N_{\rm pix} = 12 N_{\rm side}^2$	$\theta_{\rm pix} = \Omega_{\rm pix}^{1/2}$
0	1	12	58.6^o
1	2	48	29.3^o
2	4	192	14.7^o
3	8	768	7.33^o
4	16	3072	3.66^o
5	32	12288	1.83^o
6	64	49152	$55.0'$
7	128	196608	$27.5'$
8	256	786432	$13.7'$
9	512	3145728	$6.87'$
10	1024	12582912	$3.44'$
11	2048	50331648	$1.72'$
12	4096	201326592	$51.5''$
13	8192	805306368	$25.8''$
14	2^{14}	3.22×10^9	$12.9''$
15	2^{15}	1.29×10^{10}	$6.44''$
16	2^{16}	5.15×10^{10}	$3.22''$
17	2^{17}	2.06×10^{11}	$1.61''$

Numbering schemes: **Ring**



Numbering schemes: Nest



Installation

HEALPix **vs** Healpy

HEALPix

Data Analysis, Simulations and Visualization on the Sphere

Brought to you by: hivon, kmgorski, reinecke

Summary | Files | Reviews | Support | Wiki | Mailing Lists | Bugs | News | Discussion | Code

★ 0.0 Stars (6)

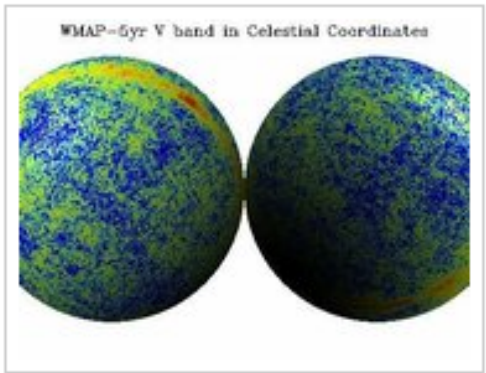
93 Downloads (This Week)

Last Update: 4 hours ago

Tweet G+1 1 Curtir 1

Download Healpix_3.31_2016Aug26.tar.gz

Browse All Files



Description

Software for pixelization, hierarchical indexation, synthesis, analysis, and visualization of data on the sphere. Please acknowledge HEALPix by quoting the web page http://healpix.sourceforge.net and publication: K.M. Gorski et al., 2005, Ap.J., 622, p.759

(https://sourceforge.net/projects/healpix/)

Directory structure

```
camila@cosmo:~$ cd Instalacoes/Healpix_3.31/
camila@cosmo:~/Instalacoes/Healpix_3.31$ ls -l
total 148
-rwxr--r-- 1 camila camila  2653 Ago 26  2016 configure
-rw-r--r-- 1 camila camila 18011 Nov 24  2010 COPYING
drwxr-xr-x 2 camila camila  4096 Ago 26  2016 data
drwxr-xr-x 4 camila camila  4096 Ago 26  2016 doc
-rwxr-xr-x 1 camila camila  6627 Dez  4  2014 healpix_doc
-rw-r--r-- 1 camila camila 69887 Ago 16  2016 hpxconfig_functions.sh
-rw-r--r-- 1 camila camila  3309 Ago 26  2016 INSTALL
-rw-r--r-- 1 camila camila  7582 Ago 26  2016 Makefile.in
-rw-r--r-- 1 camila camila  5753 Ago 26  2016 READ_Copyrights_Licenses.txt
-rw-r--r-- 1 camila camila   224 Ago 26  2016 README
drwxr-xr-x 8 camila camila  4096 Ago 26  2016 src
drwxr-xr-x 2 camila camila  4096 Ago 26  2016 test
-rw-r--r-- 1 camila camila    33 Ago 26  2016 Version
```

Directory structure ... Let's see!