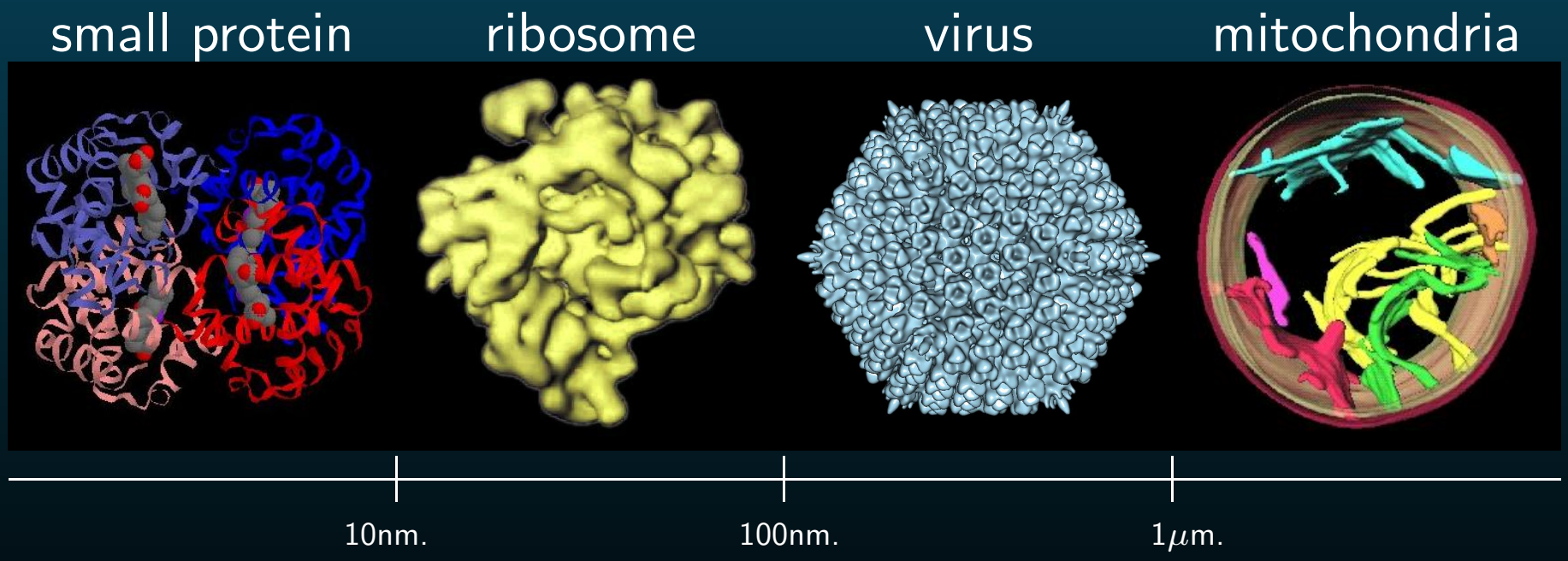
3D Reconstruction of Biological Macromolecules using Crystalline Samples and Electron Microscopy

R Marabini

April 2001

Reconstruction of biological macromolecules from projections obtained with an Electron Microscope.



Transmission Electron Microscopy in Biology

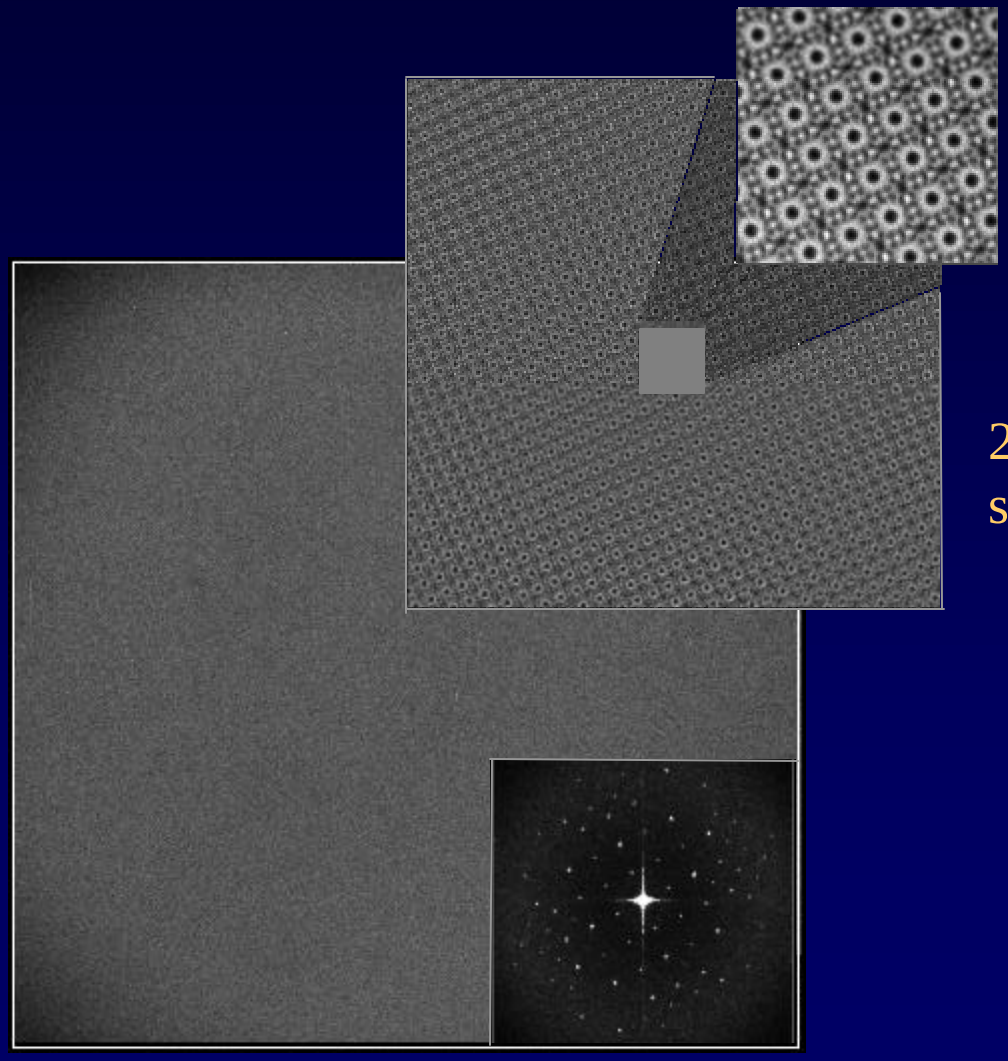
- Understanding the structure of biological macro-molecules is central to the interpretation of their function in the cell.
- TEM makes possible direct imaging of biological structures at molecular level.
- Critical aspects to reach high resolution:
  - Specimen preparation
  - Imaging Process
- Image processing techniques aim to alleviate many of the problems in TEM images.

Specimen Preparation: 2D Biological Crystals

- Biological macromolecules are sensitive to electron radiation.
- To preserve high resolution details and minimize the radiation damage, specimen images are recorded at low dose (1-10 e^-/Å).
- As a consequence, the SNR is very poor and images are extremely noisy.
- To improve the SNR, 2D crystalline specimens are used.

2D Biological Crystals

Cryomicroscopy TEM image:
- Low e^- dose
- Low SNR



2D Crystalline specimens

Image Processing of 2D Biological Crystals in TEM

- The General Problem:

  To obtain the 3D reconstruction at high resolution of biological macromolecules from projection images.

- The Key:

  The Fourier Central Section Theorem.

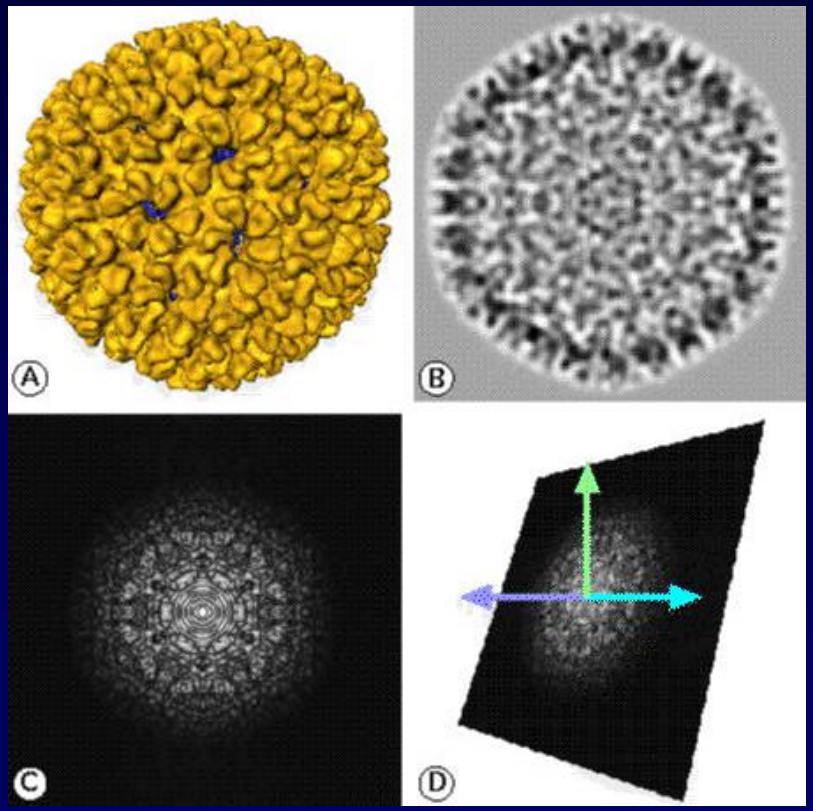
- The Method to Be Used:

  Direct Fourier Method of reconstruction

- Limitations of the Method:
  - *Interpolation in reciprocal space.*
  - *Lack of information in certain areas of the space.*

    (the range of tilt angles is limited: ±60^{o})

Fourier Central Section Theorem



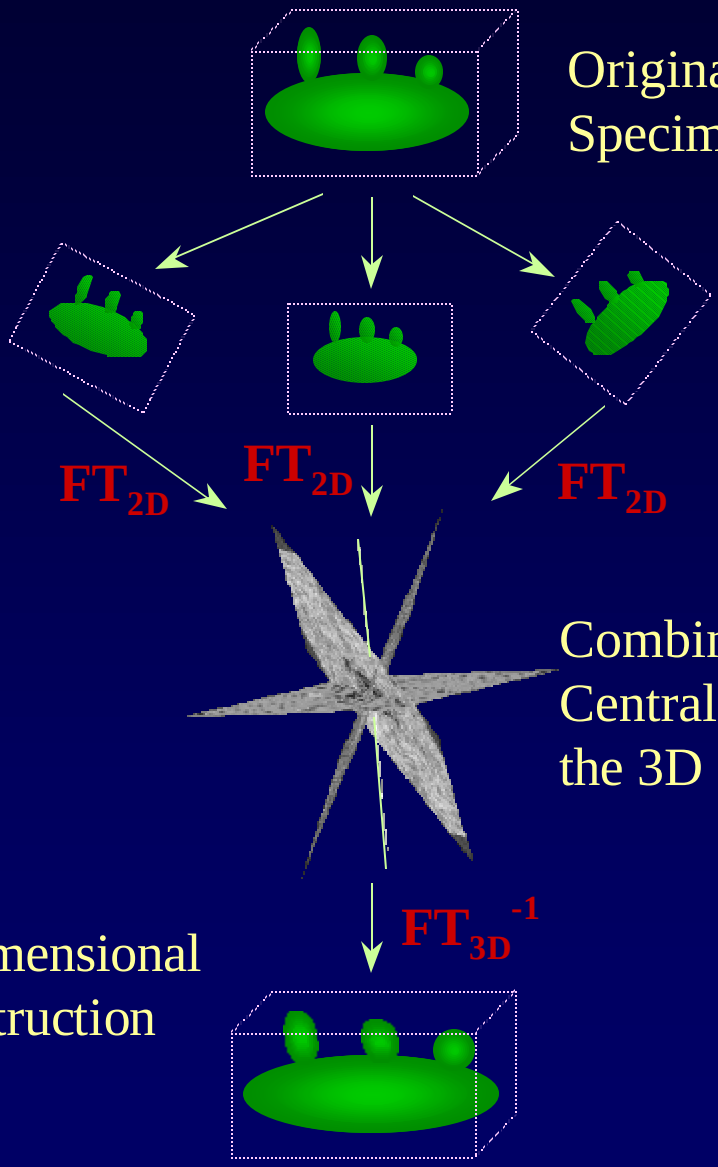
A) 3D Object.

B) Projection from (A).

C) Fourier transform from (B).

D) The Fourier transform of a 2D projection is a plane in the 3D Fourier space which crosses the origin: *Central Section.*

Direct Fourier Method



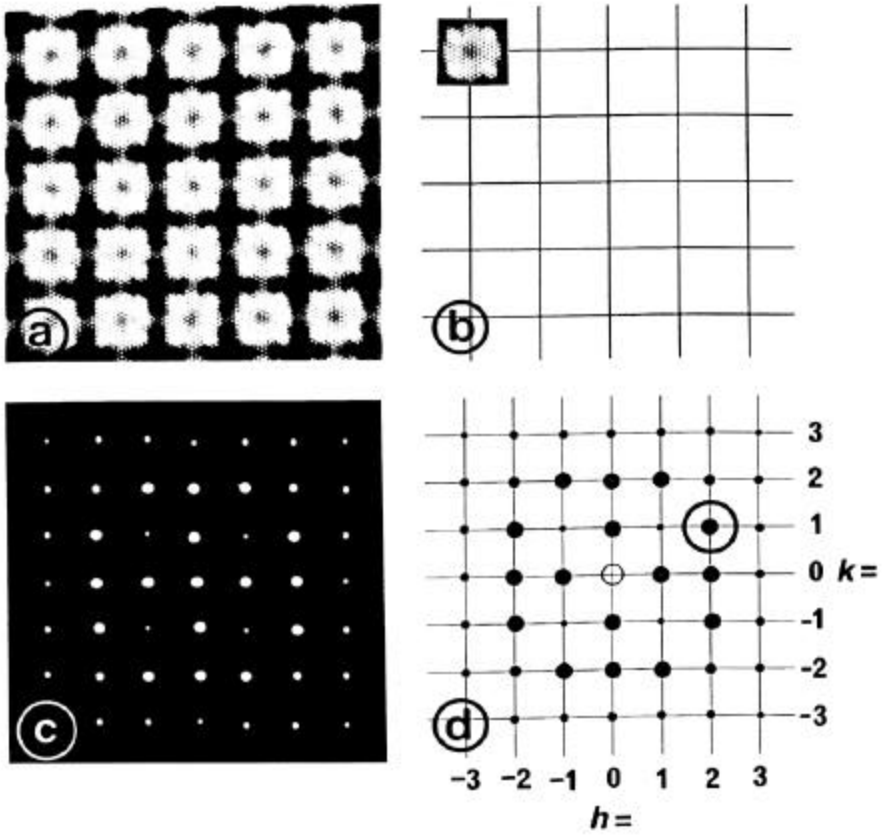
3D Fourier Transform of a 2D Crystal

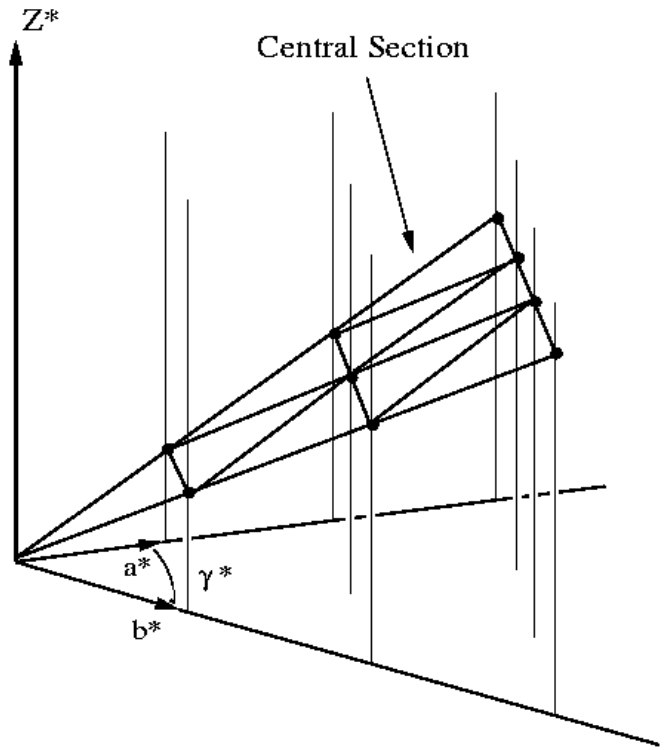
FT of a projection image from a 2D crystal:

Structural information arranged in a lattice in Fourier space.

Combination of central sections:

Lattice lines along Z^* axis.





Structure Determination for 2D Crystals

1. Collection and digitalization of sufficient good electron micrographs at suitable distributed viewing angles.
2. Analysis of each image into Fourier components.
3. Combination of the amplitudes and phases from each image to provide samples of the continuous Fourier transform at regular intervals along lattice lines.
4. Calculation of the structure via inverse 3D Fourier transform.

Analysis of the Image into Fourier Components

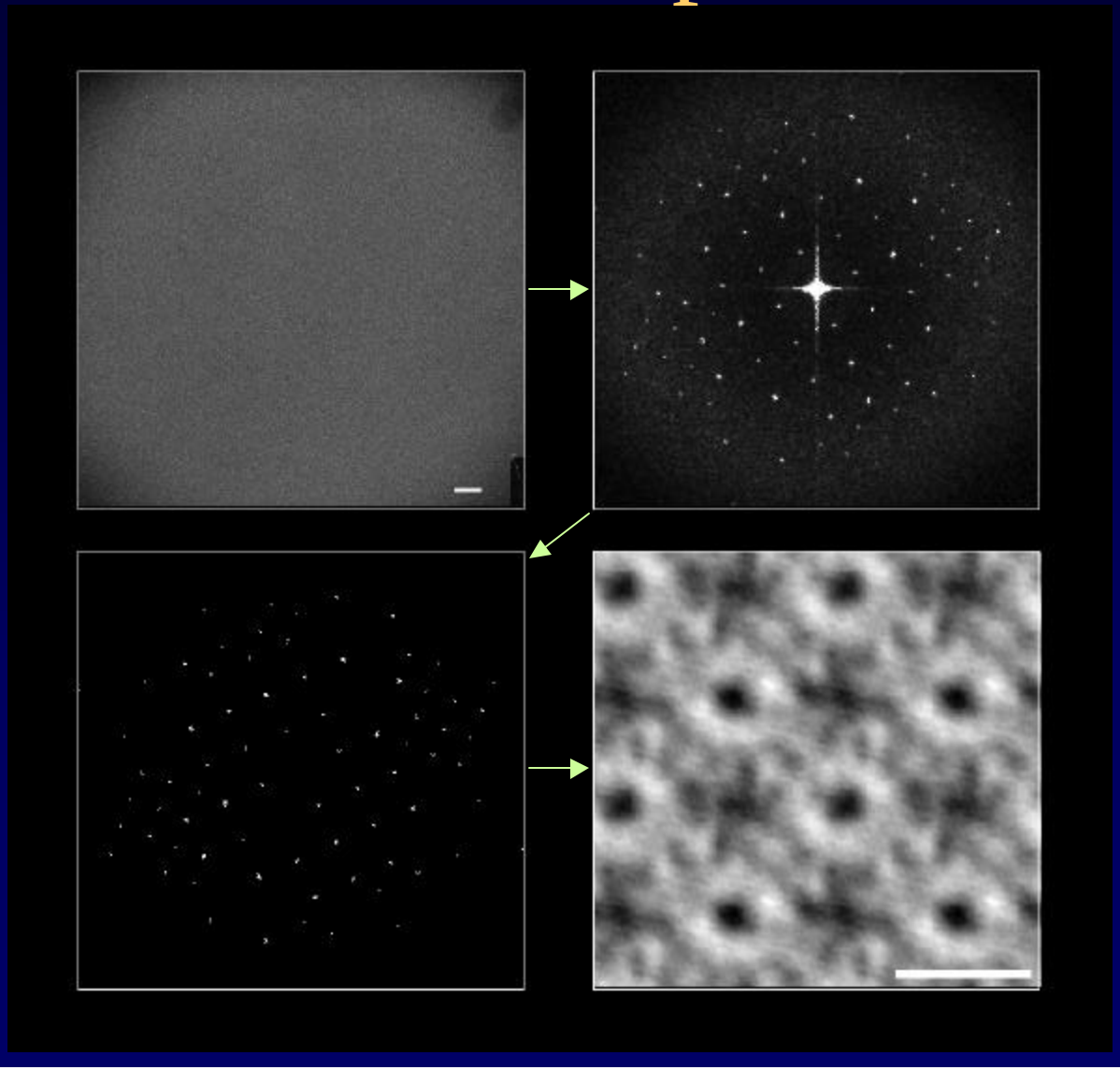
- Aim:

  To obtain reliable measures of the amplitude and phase of the frequency components.

- Problems:
  - Disorder in 2D biological crystals.
  - Structural variability present in crystals
  - Effect of the TEM image formation system.

Analysis of the Image into Fourier Components

Original Image



Fourier transform

Extraction of Fourier components

Fourier synthesis

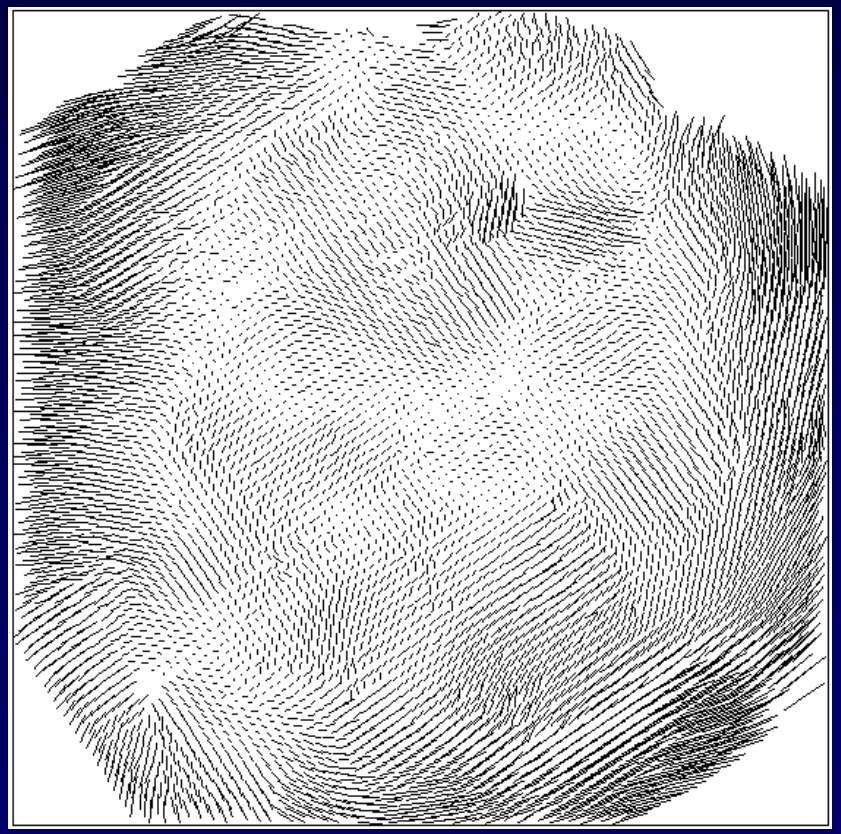
Disorder in 2D Biological Crystals

- What does *disorder* mean?

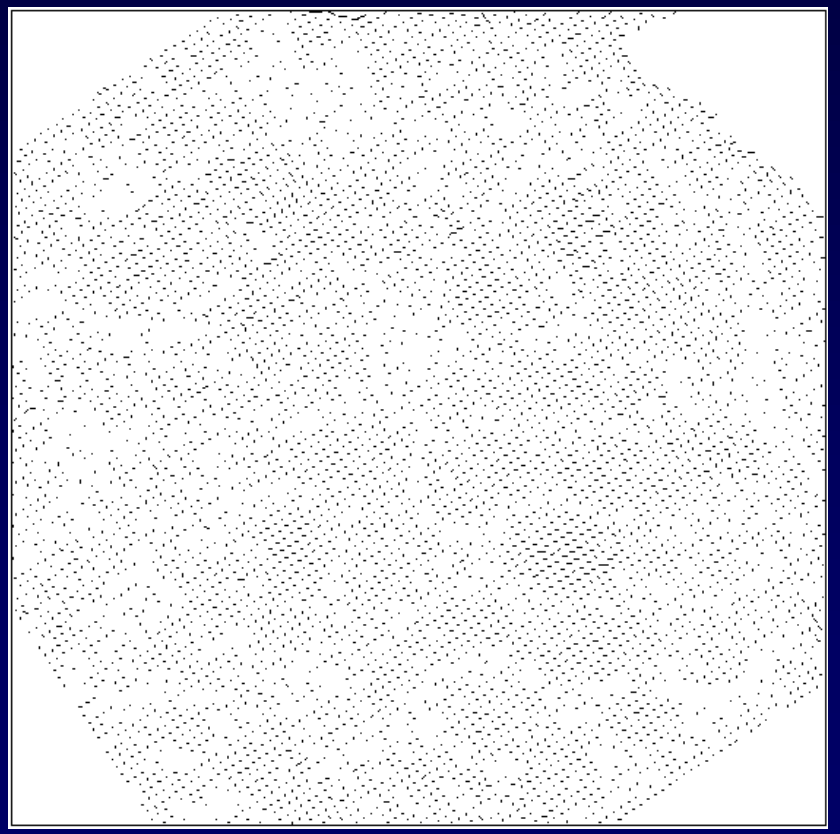
  Unit cells slightly displaced from their ideal lattice positions.
- Effects:
  - Broaden the reflections in the FT.
  - Attenuate higher spatial frequencies.
  - Loss of high resolution information.
- Correction:
  - Unbend the real space lattice by re-interpolating the image.
- Result:
  - Reflections in the FT become sharper.
  - SNR is improved
  - High resolution information is recovered.

Effect of the Unbending Process. Distortion Map

Before unbending

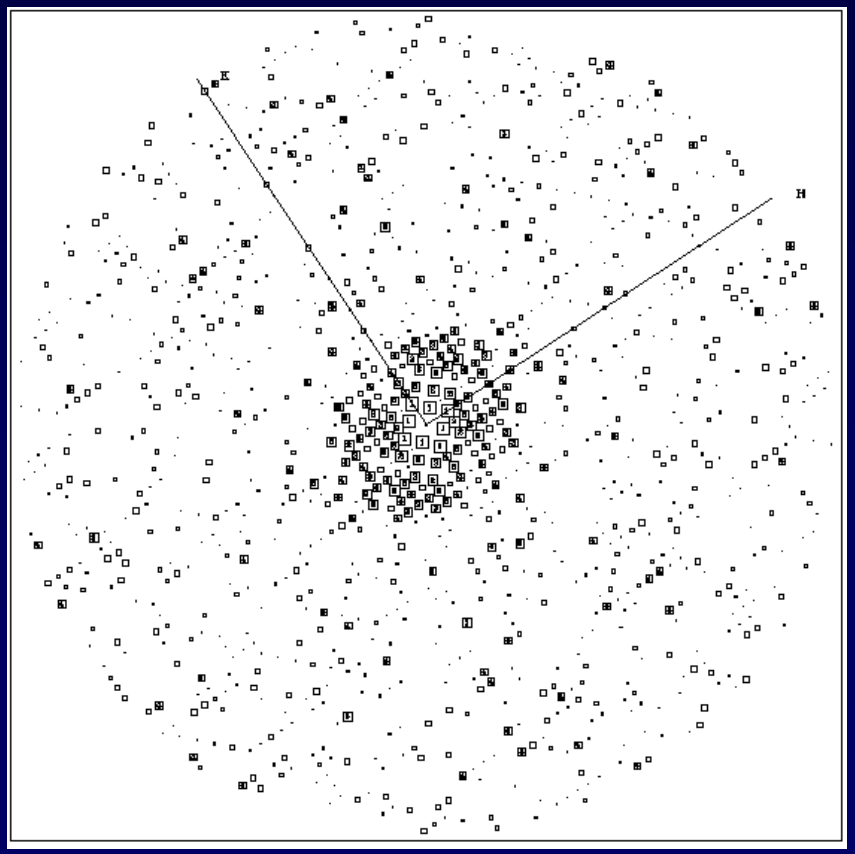


After unbending

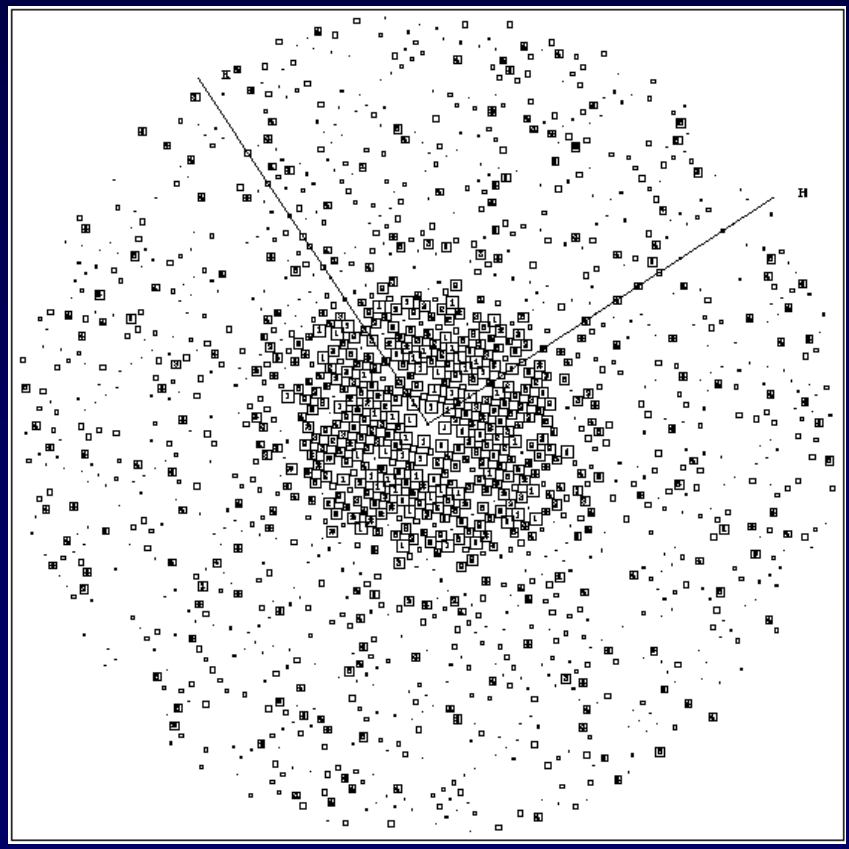


Effect of the Unbending Process. Fourier Spectrum

Before unbending

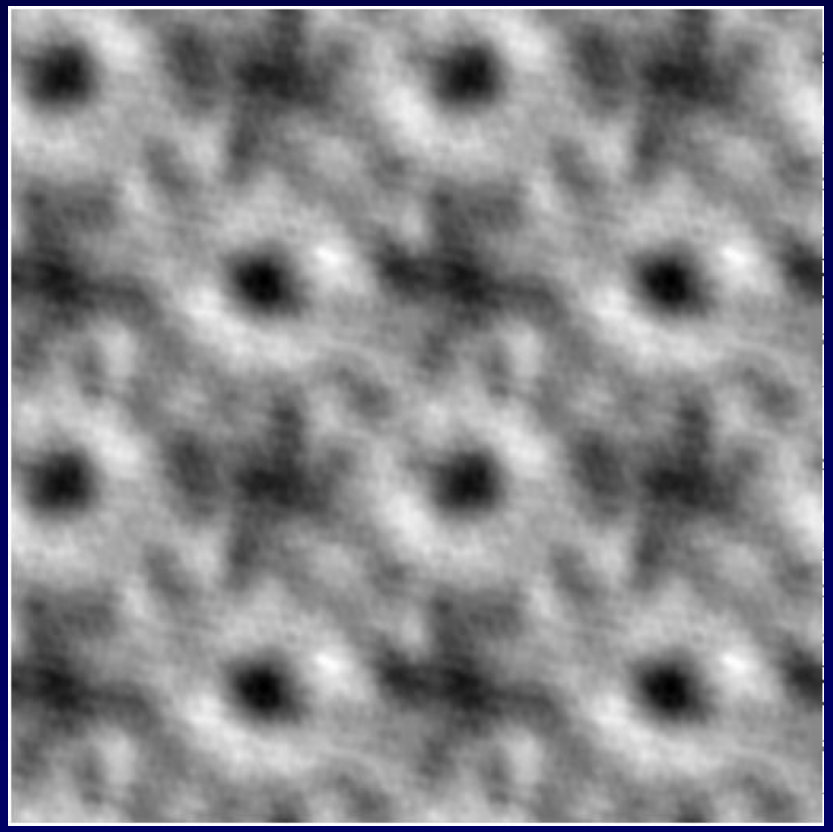


After unbending

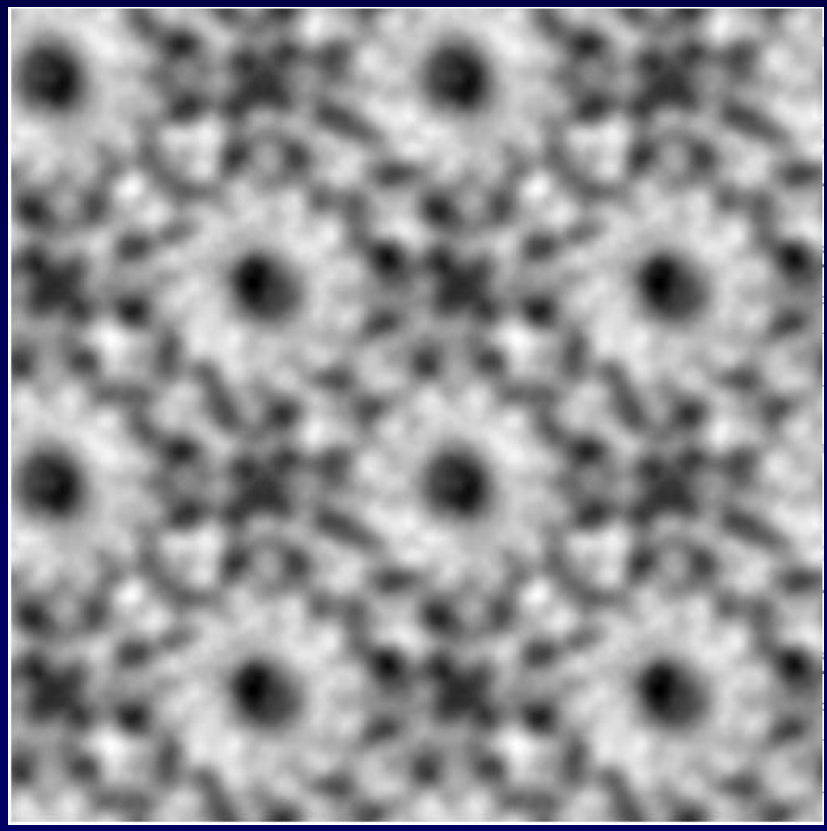


Effect of the Unbending Process. Fourier Synthesis

Before unbending



After unbending

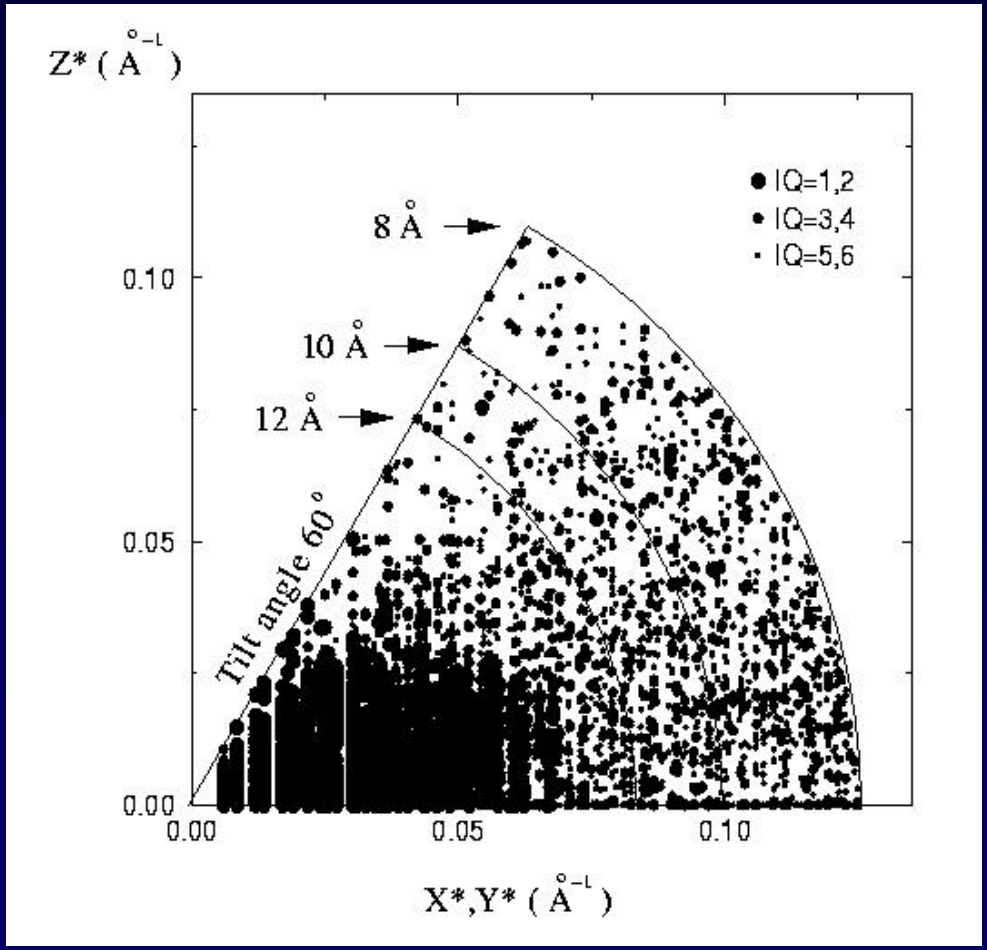


TEM Contrast Transfer Function

- Represents the change made by the TEM imaging system at each spatial frequency.
- CTF Detection: Spectral estimation approaches help find the CTF affecting the image.
- CTF Correction: Reversed phases are easily restored. Amplitudes are restored by dividing by the CTF level (beware of noise).

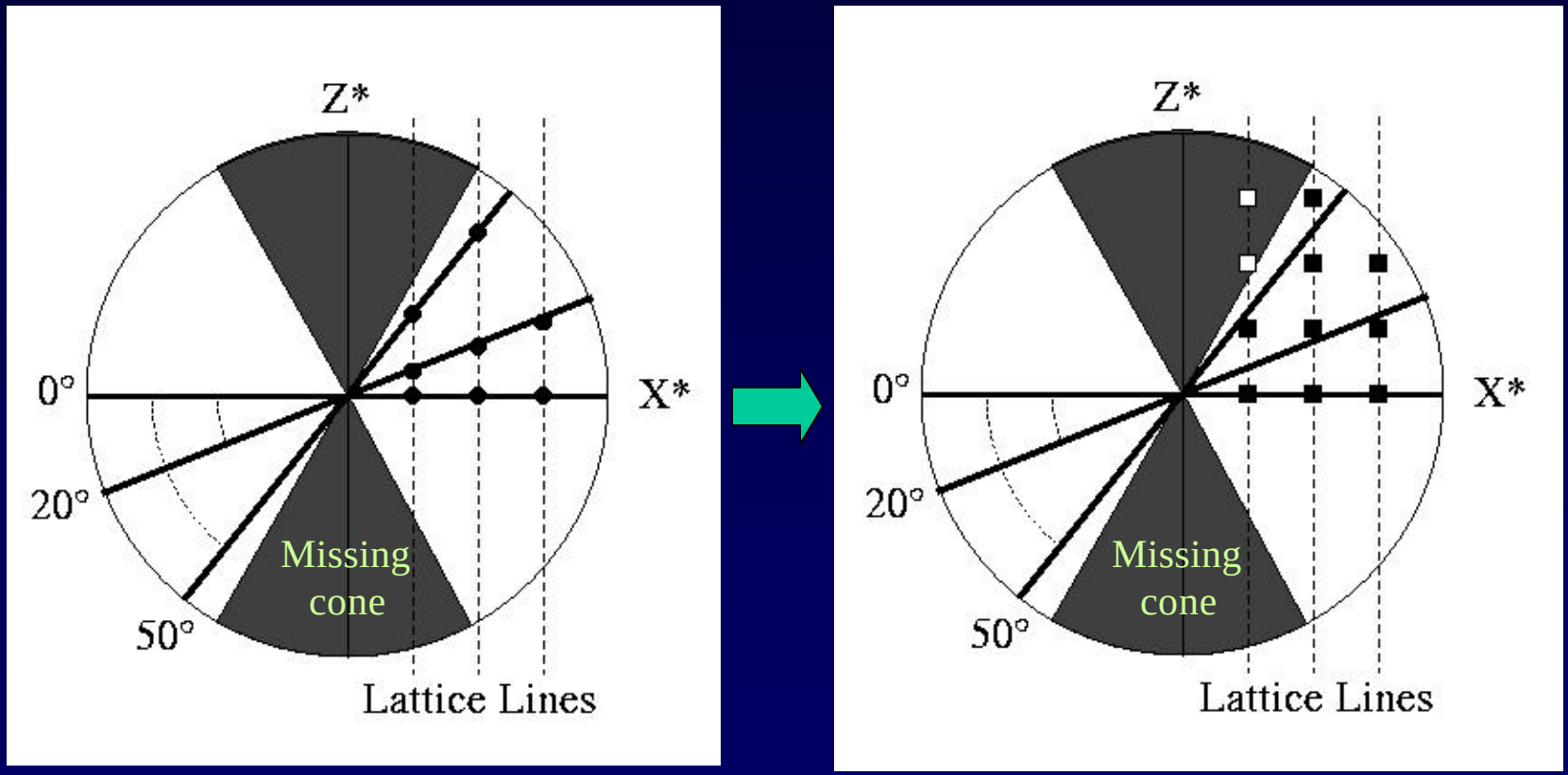
Sampling of Fourier Space

after combination of all central sections.



Interpolation along Lattice Lines

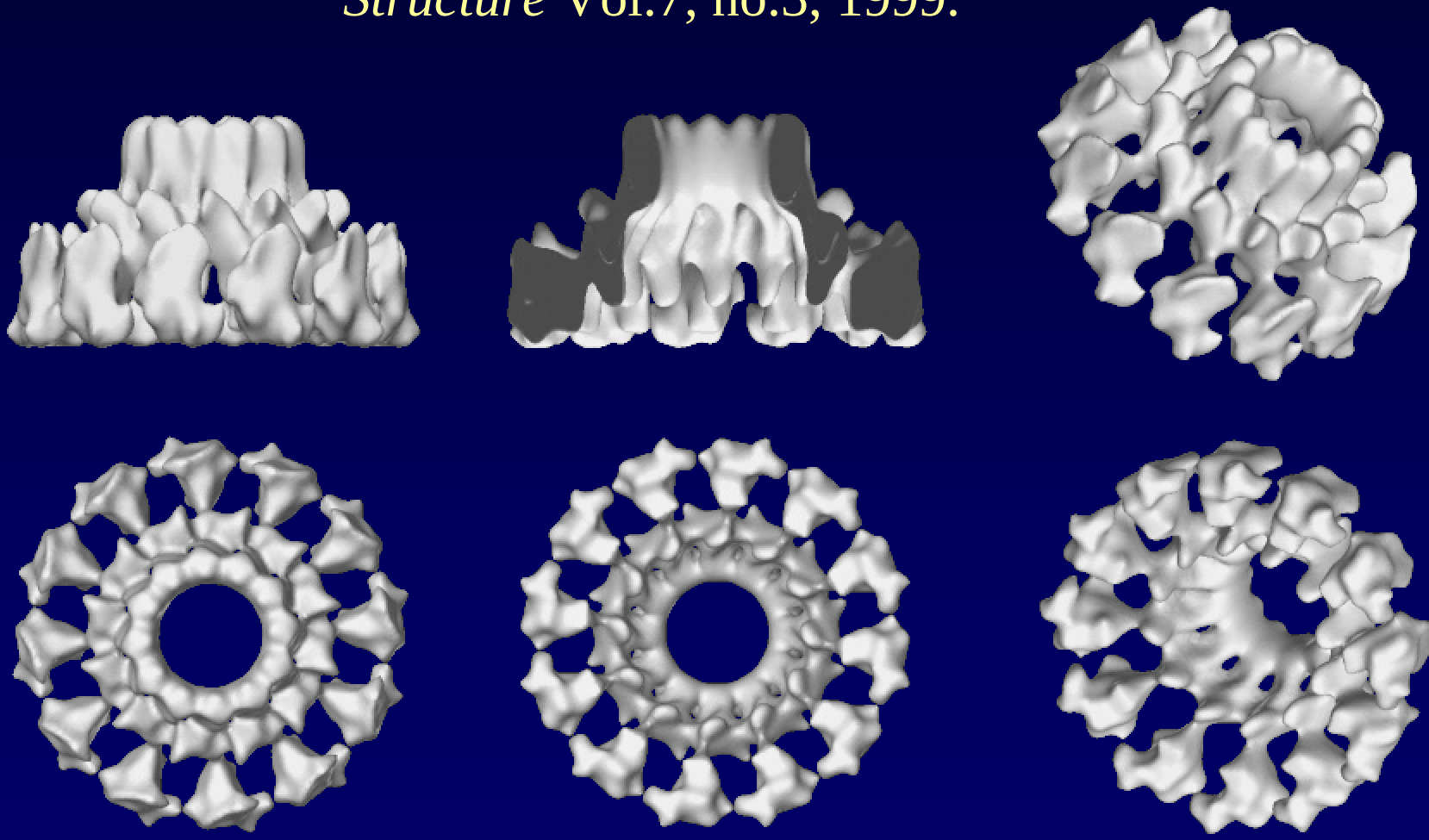
To apply the 3D FT^{-1}, lattice lines have to be uniformly sampled.



Experimental application

- 3D Structure of the *f29* bacteriophage connector.

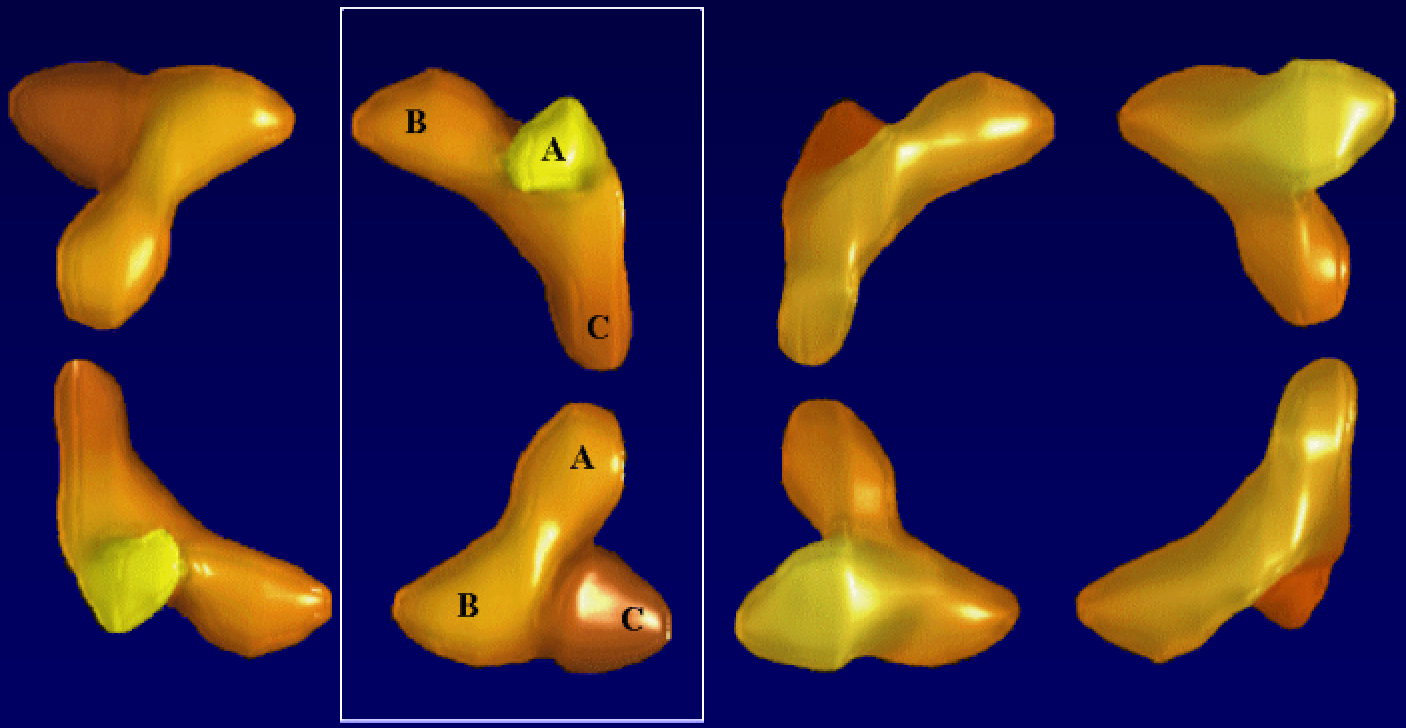
Structure Vol.7, no.3, 1999.



Experimental application

- 3D structure of the *StnII* protein.

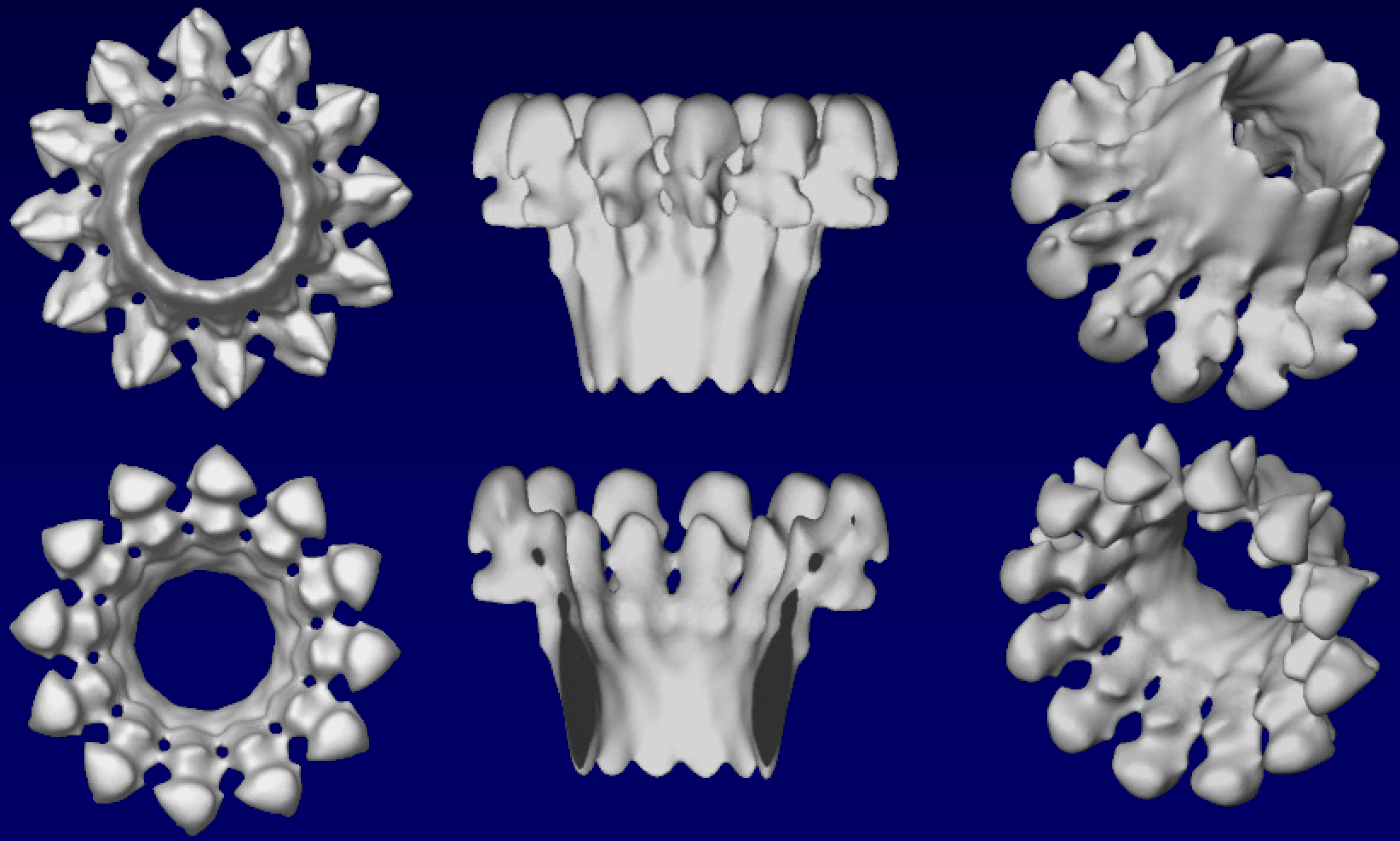
BioPhysical J., Vol 78, no.6, 2000.



Experimental application

- 3D structure of the *T3* bacteriophage connector.

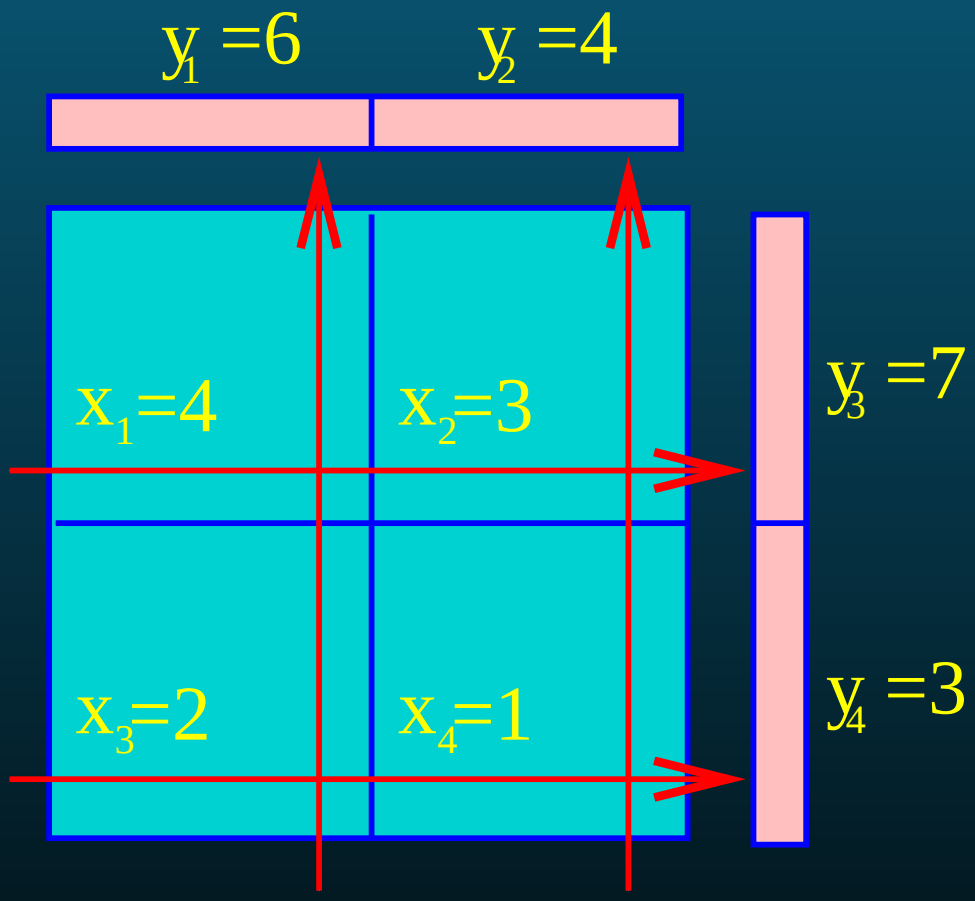
J. Structural Biology, Vol., pp., 2000.



- The very general aim: **Reconstruction of proteins (biological macromolecules) from images (projections) obtained with an Electron Microscope**.
- 3D Reconstruction using **ART**. (What is ART and why should we try it?)
- Different samples require different reconstruction techniques. (Or, How did we adapt ART for single particles, viruses, crystals...)
- Implementing Constraints (extra information independent of the projections or symmetry)

Reconstruction as a Set of Linear Equations

(Or what is ART)

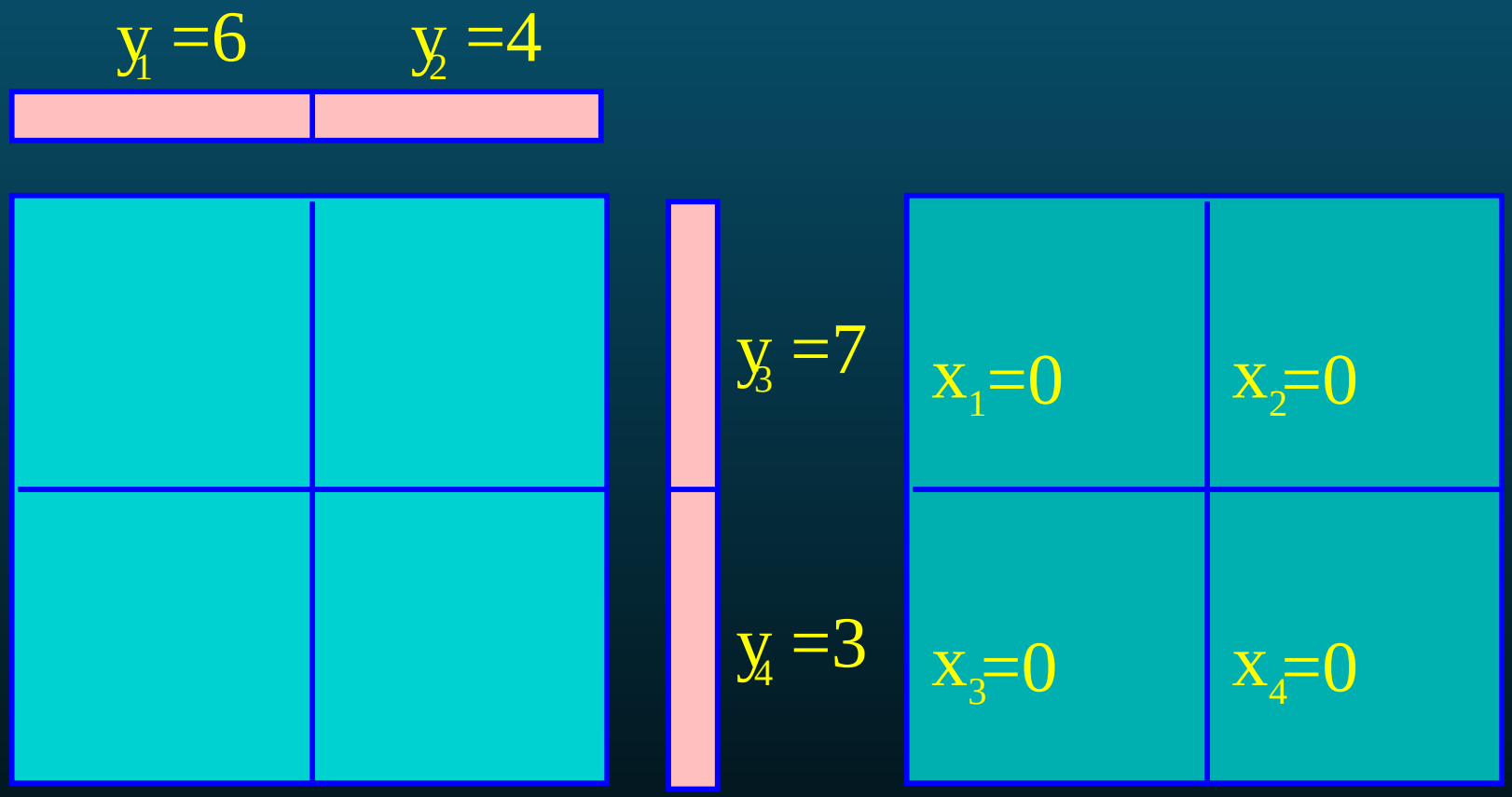


$$
\begin{aligned}
x_1 + x_2 &= 7 \\
x_3 + x_4 &= 3 \\
x_1 + x_3 &= 6 \\
x_2 + x_4 &= 4
\end{aligned}
$$

$$64 \times 64 \times 64 = 262,144; 64 \times 64 \times 1000 = 4,096,000$$

$$315 \times 315 \times 315 = 31,255,875; 315 \times 315 \times 60,000 = 5,953,500,000$$

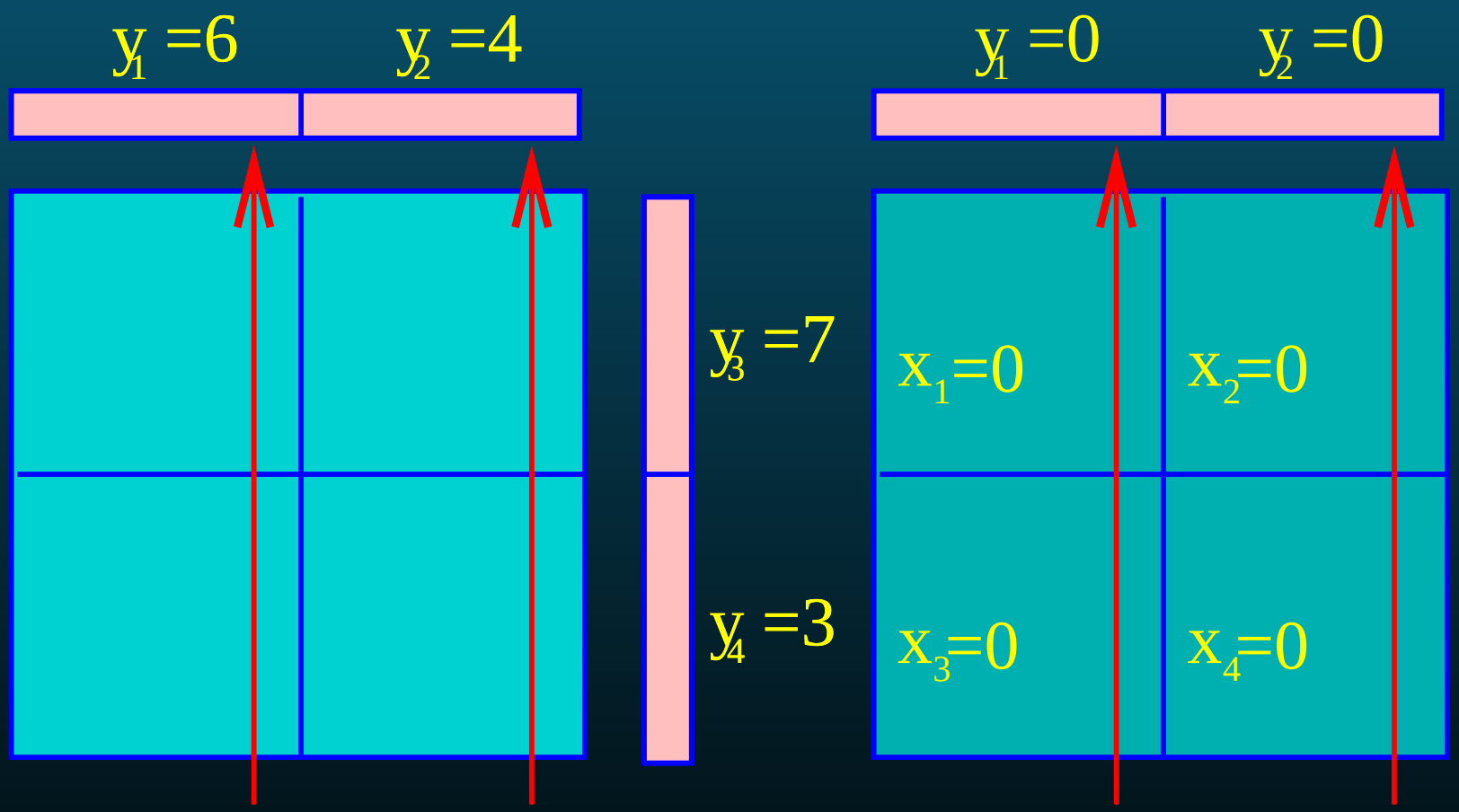
Reconstruction as a Set of Linear Equations



Microscope

Computer

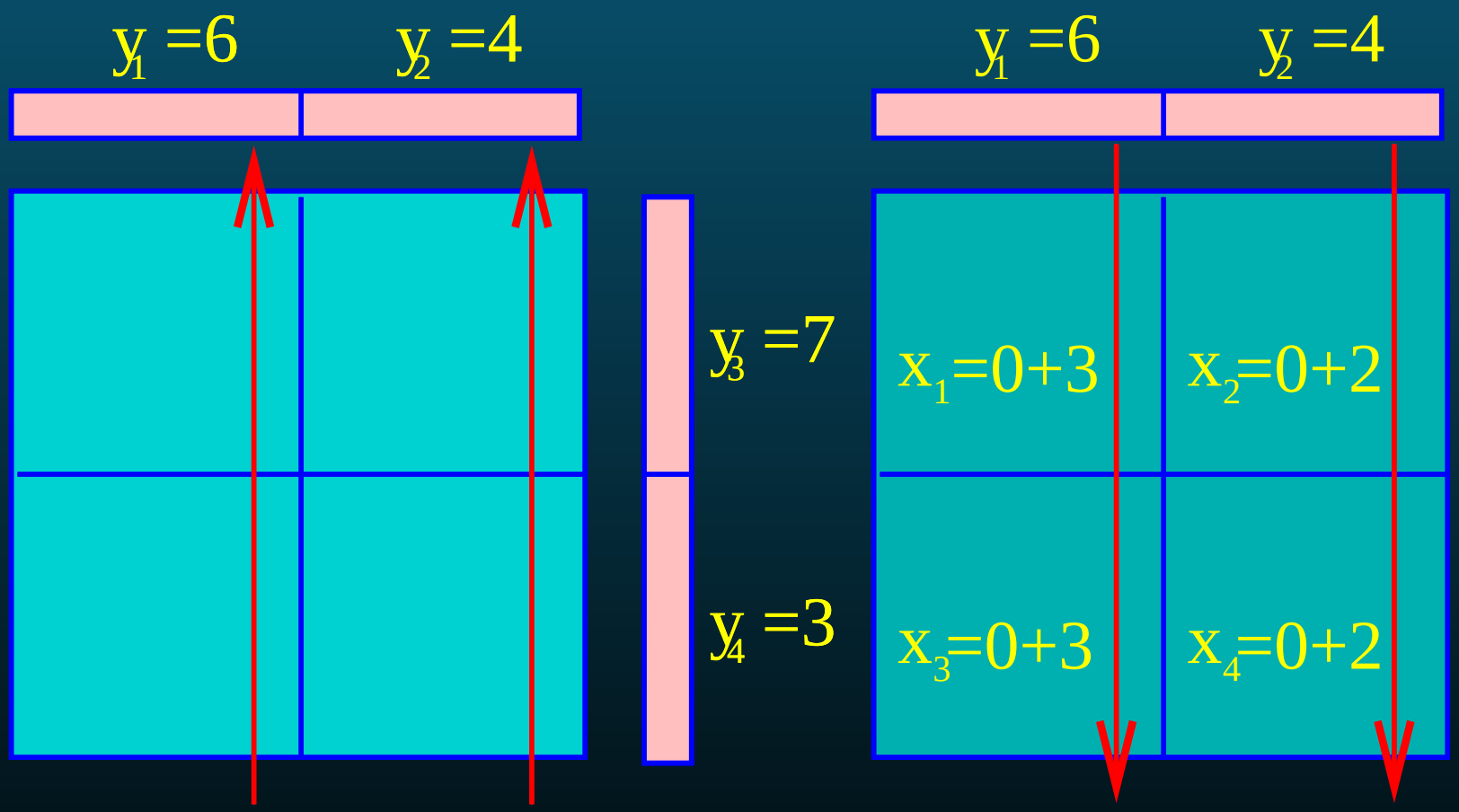
Reconstruction as a Set of Linear Equations



Microscope

Computer

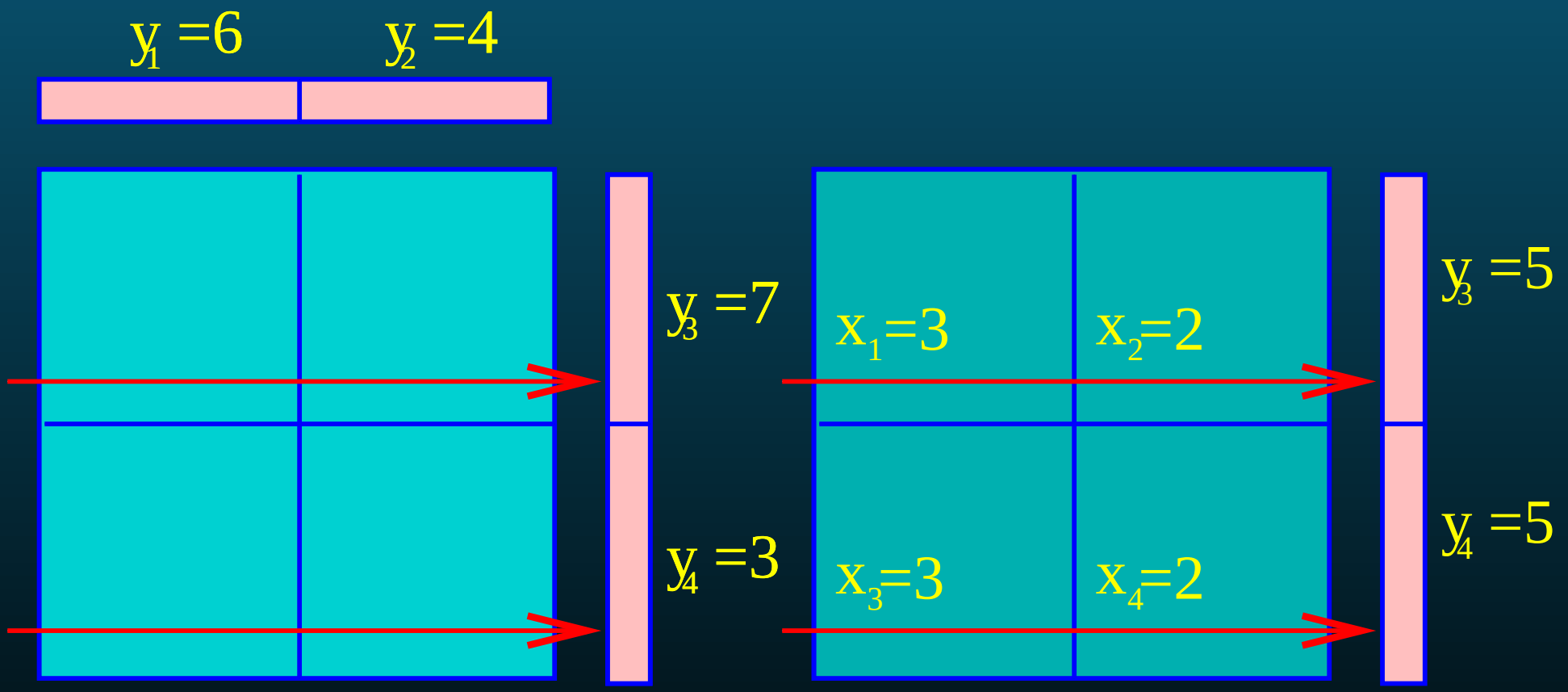
Reconstruction as a Set of Linear Equations



Microscope

Computer

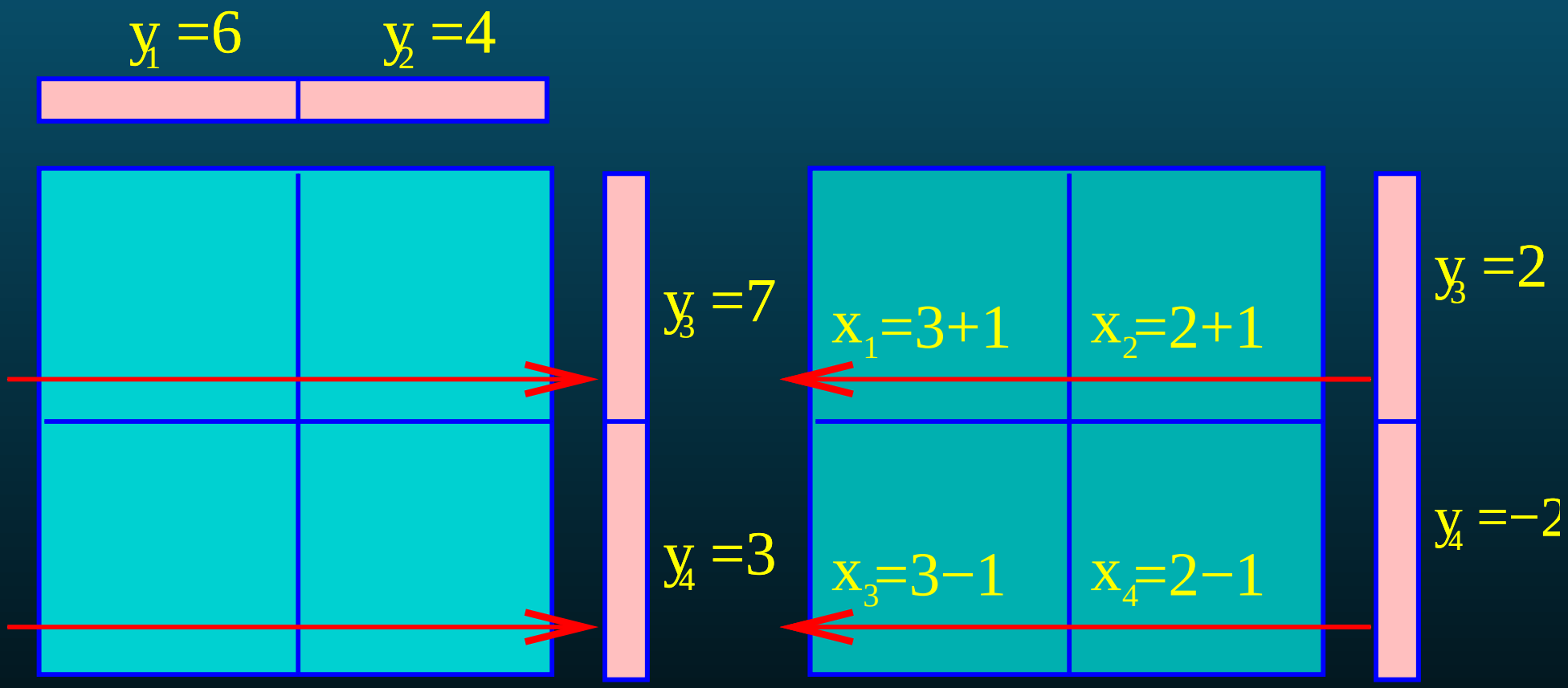
Reconstruction as a Set of Linear Equations



Microscope

Computer

Reconstruction as a Set of Linear Equations



This is already a solution!!!

Microscope

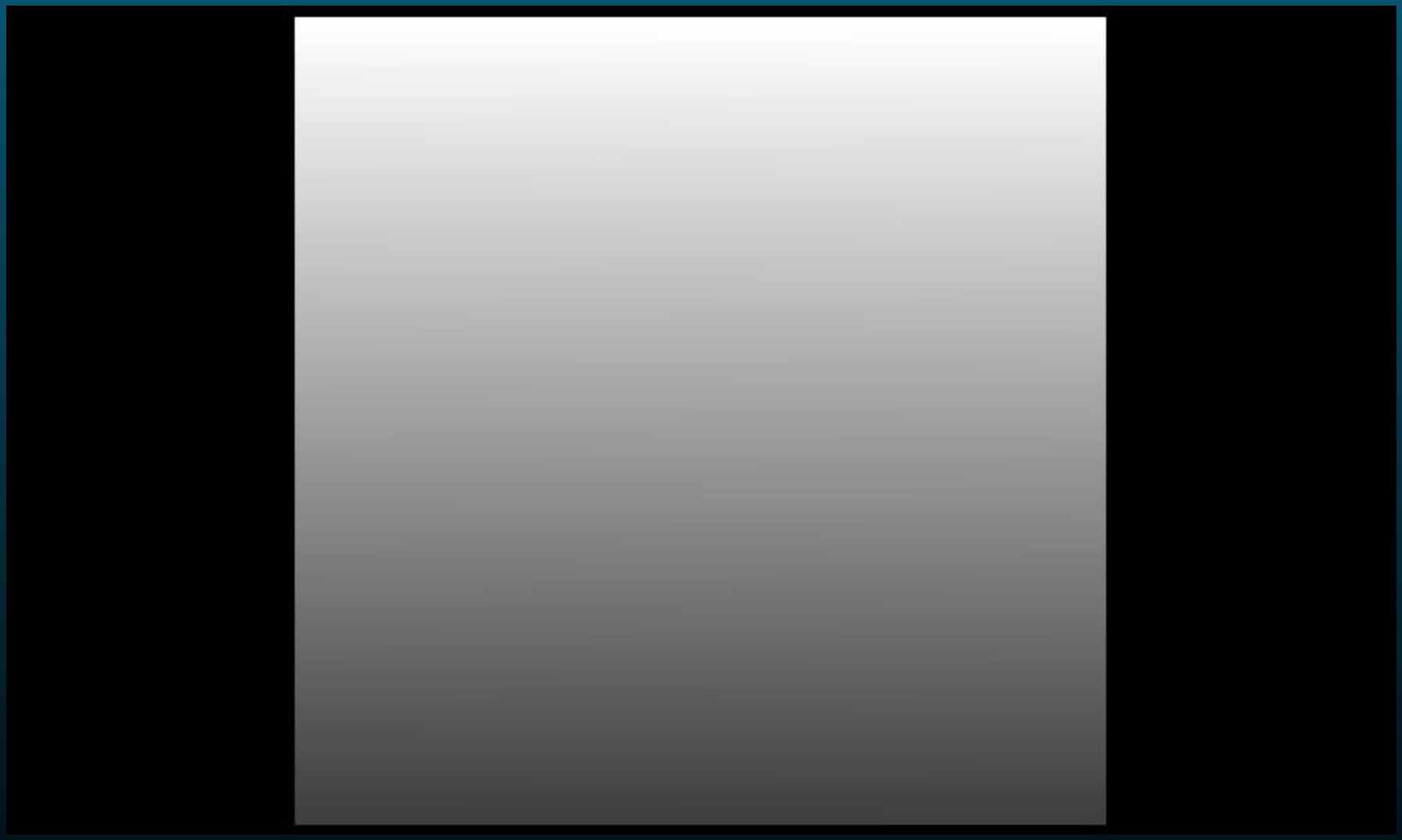
Computer

Procesamiento digital de volúmenes sin vóxeles, o cómo describir la naturaleza sin cubicarla

R. Marabini

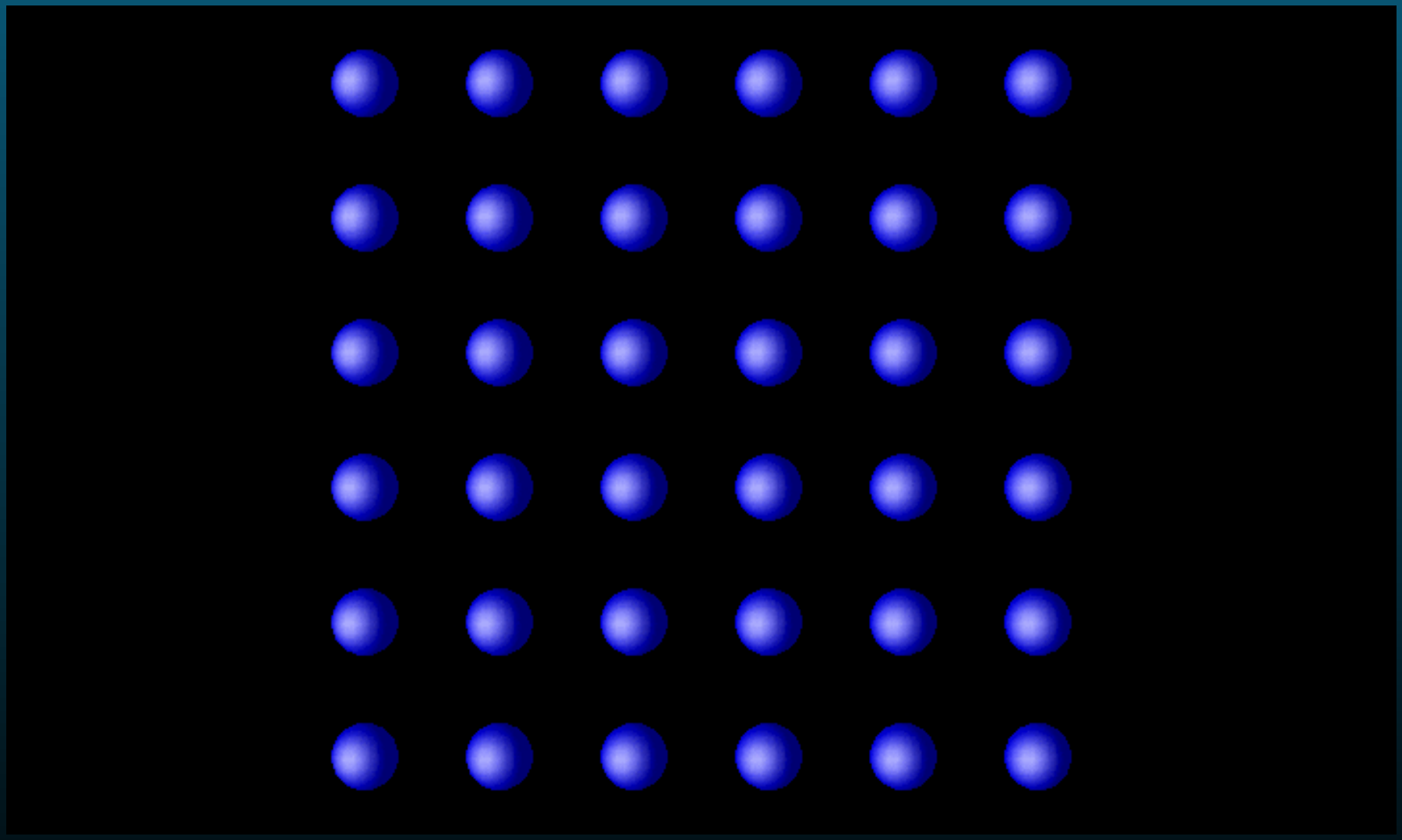
Septiembre 2001

Theoretical Background: Discrete Images



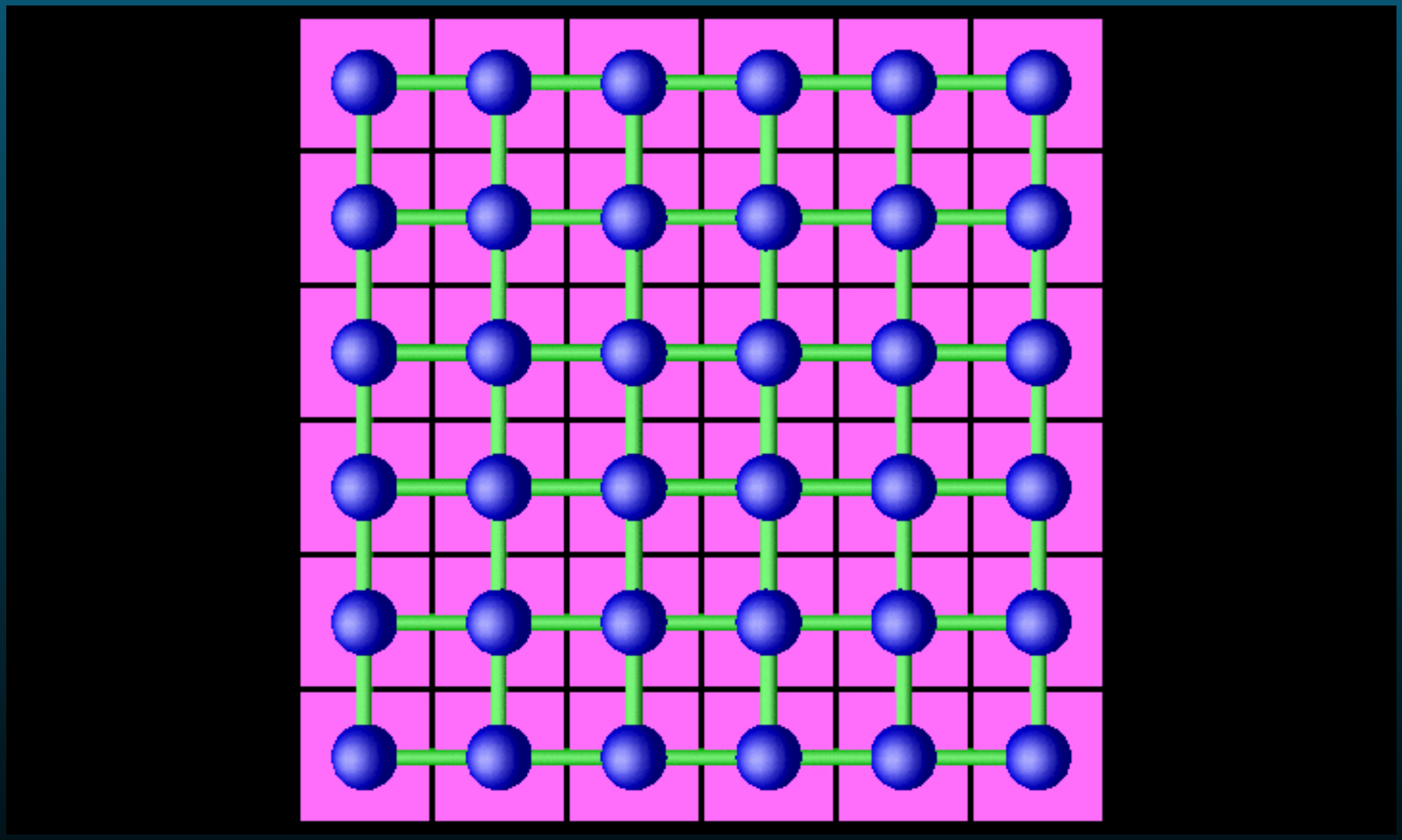
Analog Image

Theoretical Background: Discrete Images



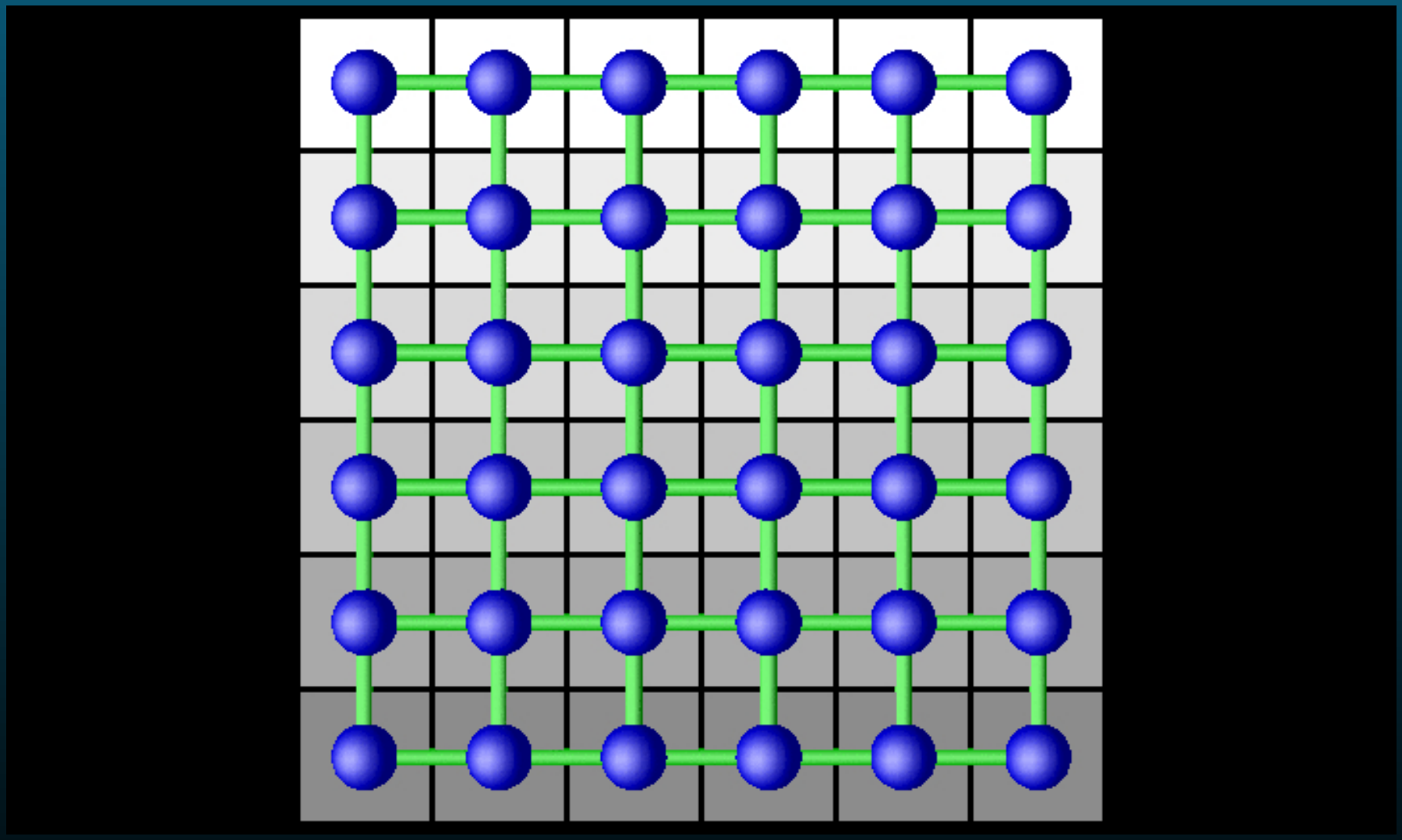
Grid (in N-D): set of points $\in \mathbb{Z}^N$

Theoretical Background: Discrete Images



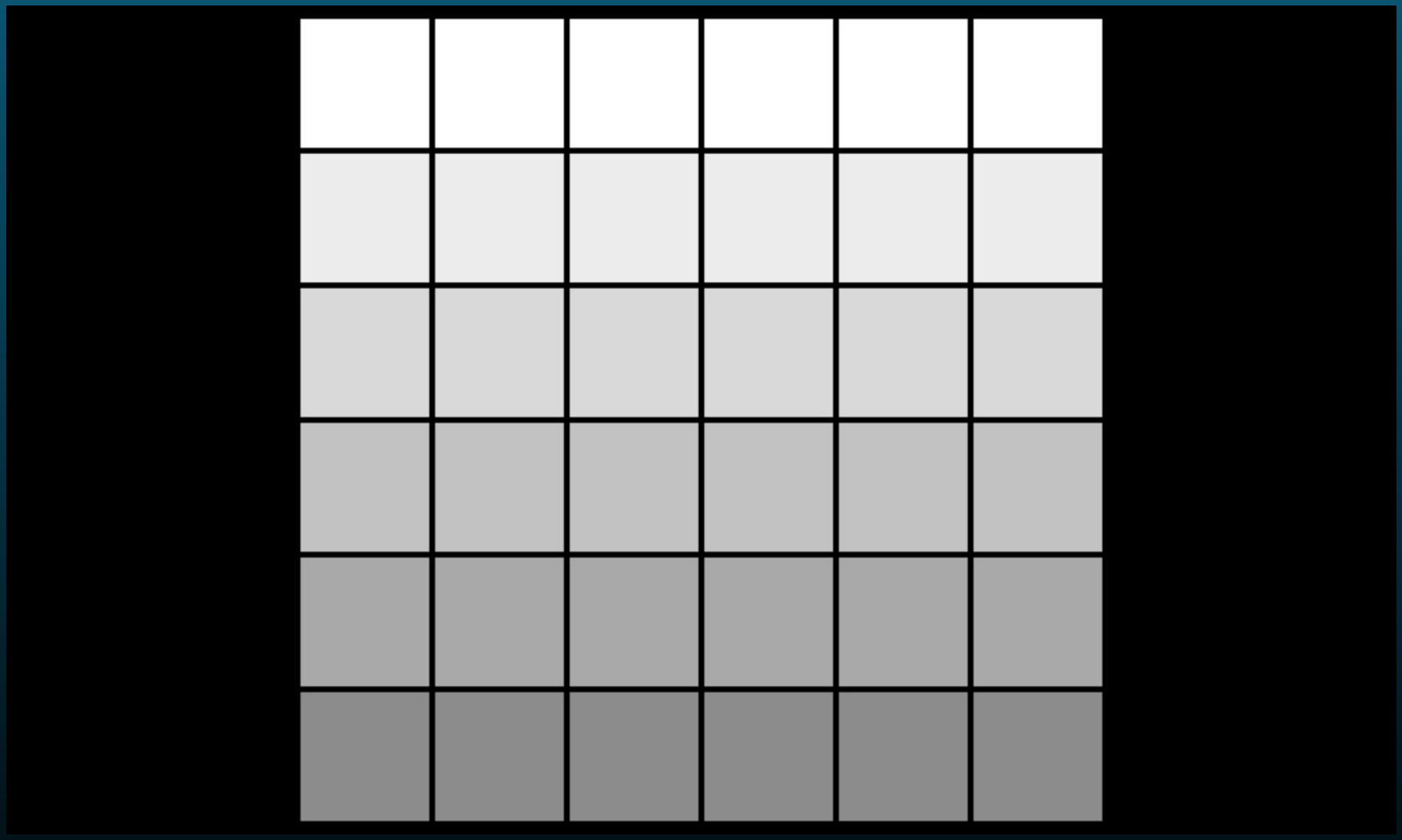
a basis function

Theoretical Background: Discrete Images

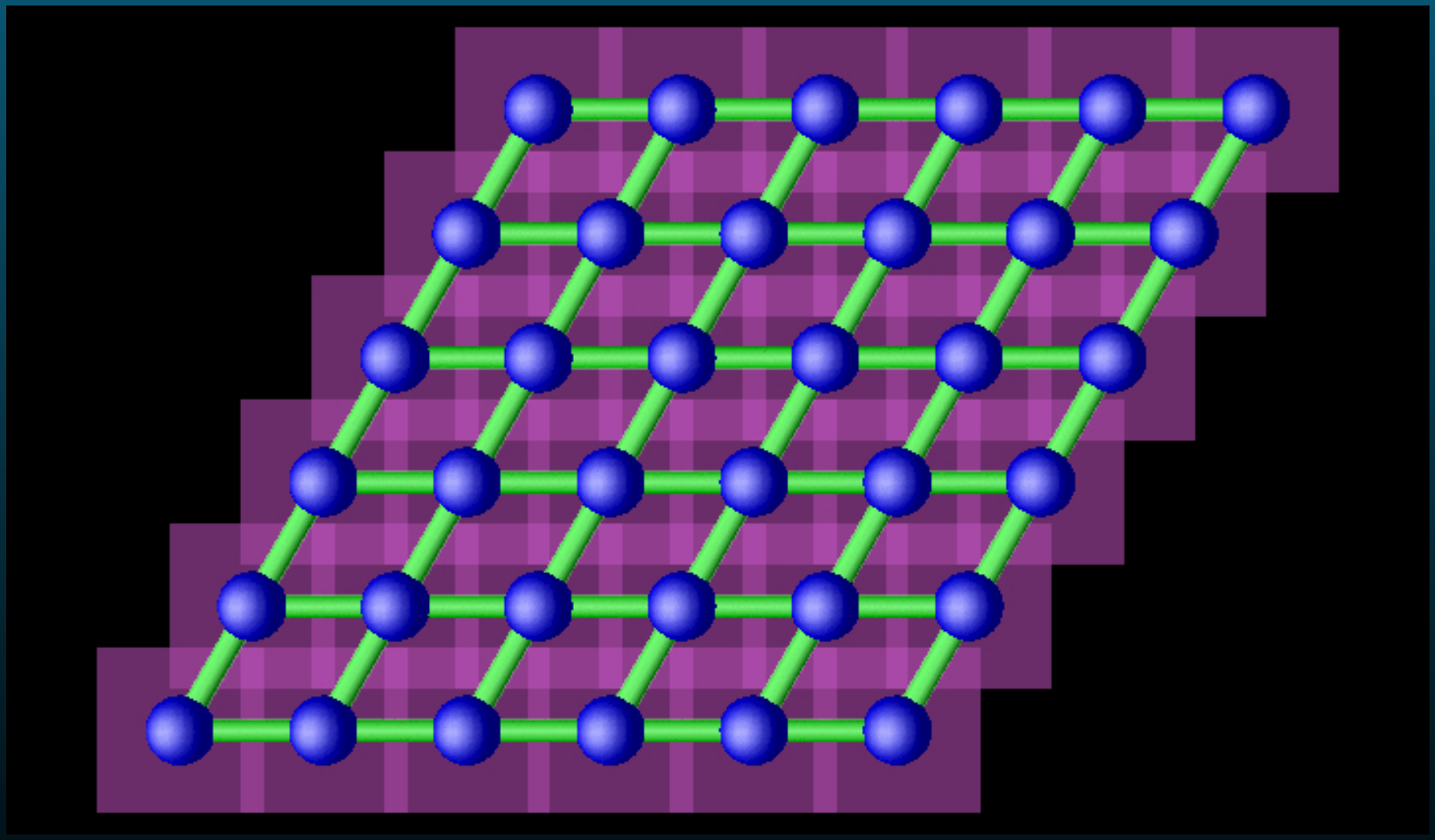


multiplied by a coeficient

Theoretical Background: Discrete Images

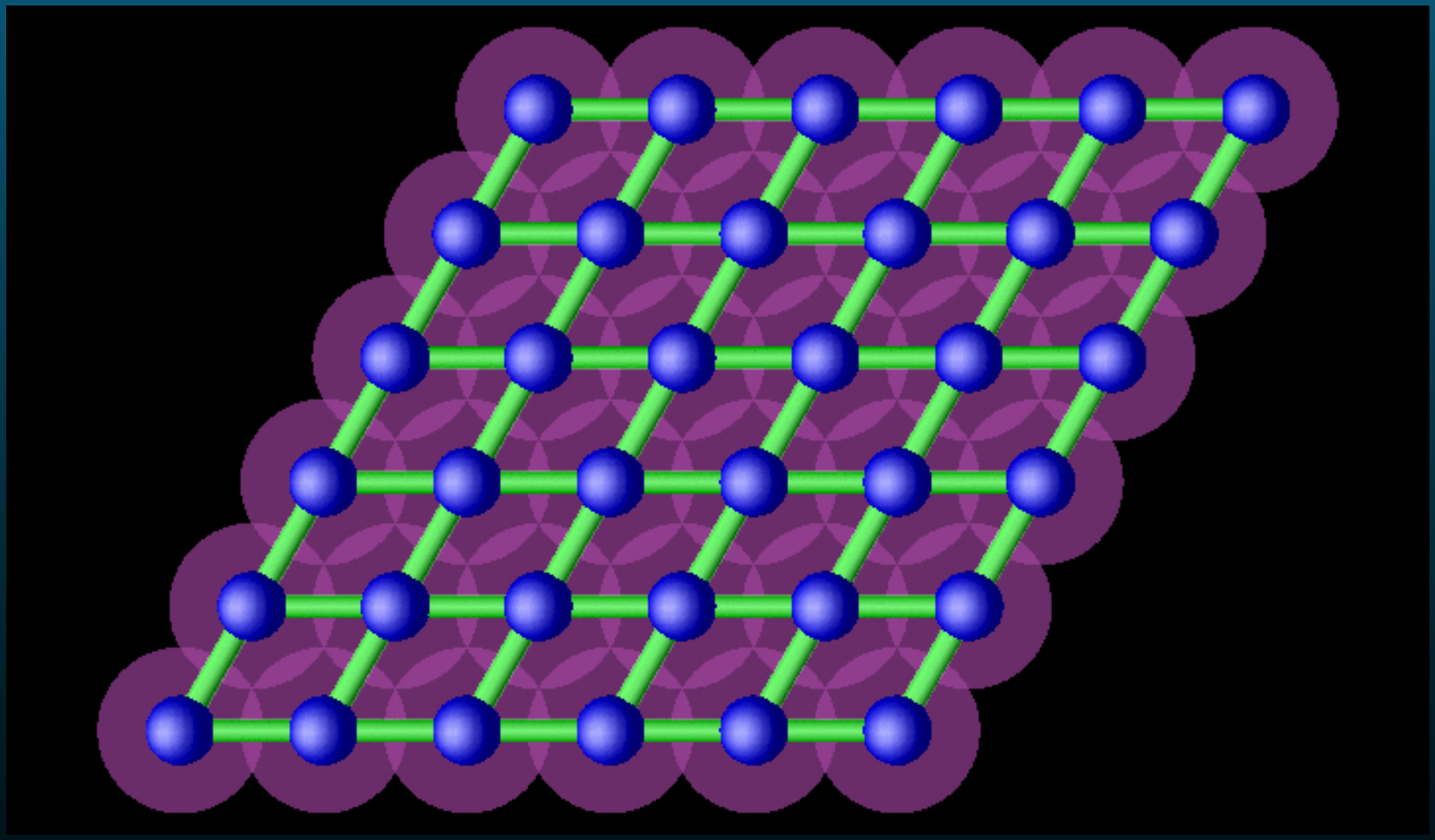


Theoretical Background: Discrete Images



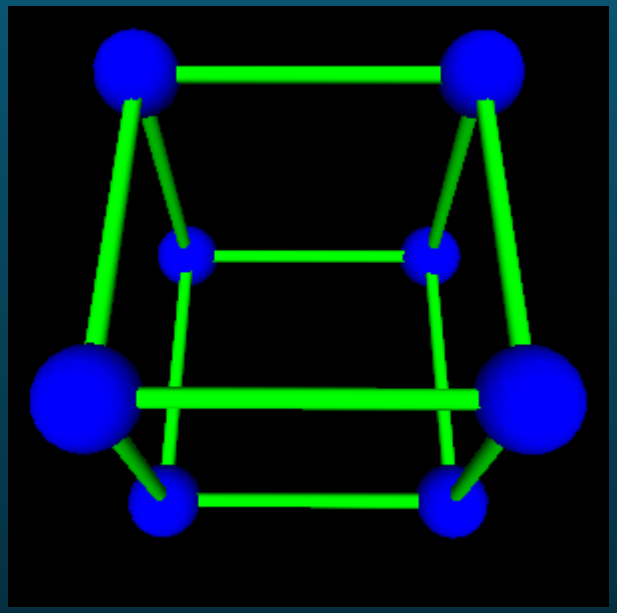
Voxels as basis functions

Theoretical Background: Discrete Images

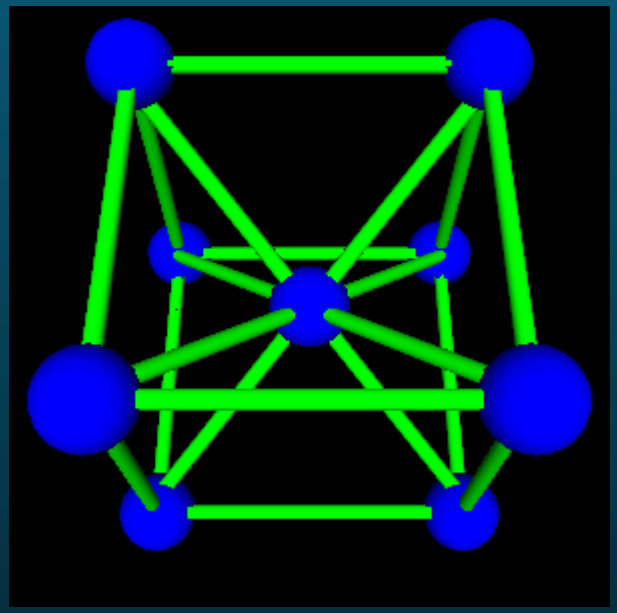


A circular symmetric set of basis functions

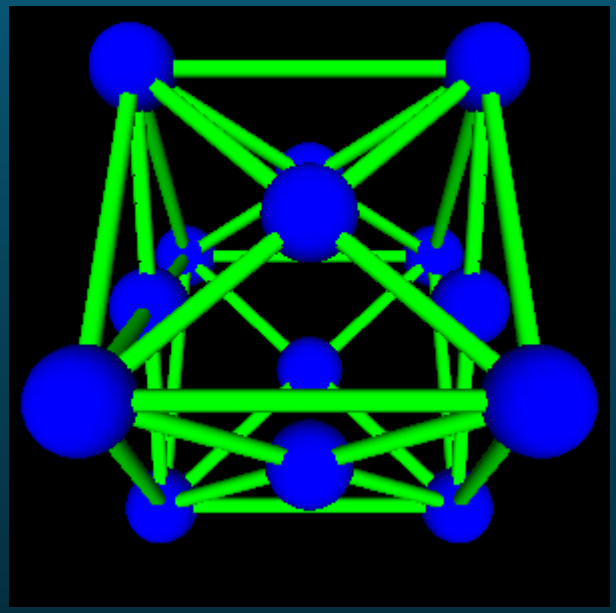
Cubic Grids



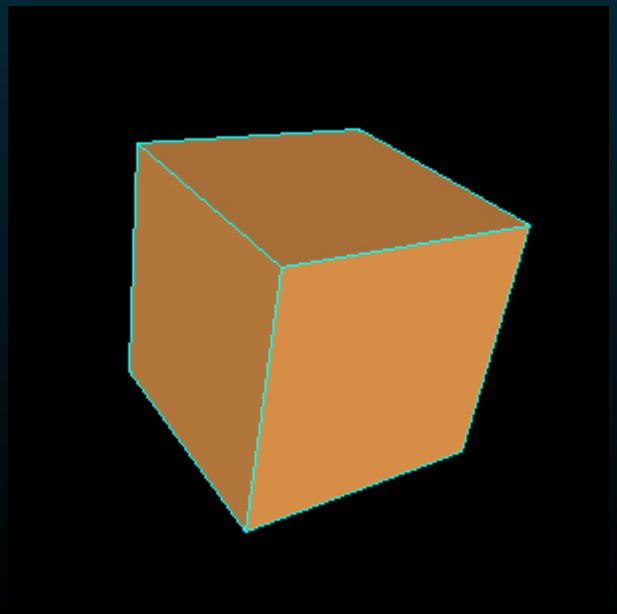
SC

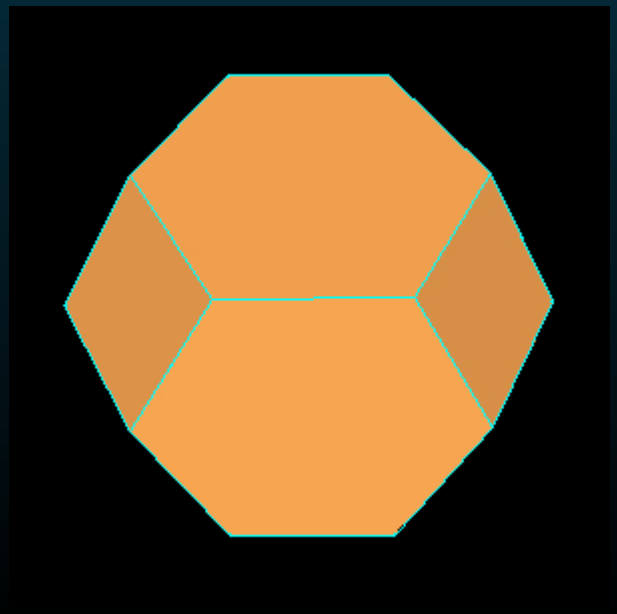


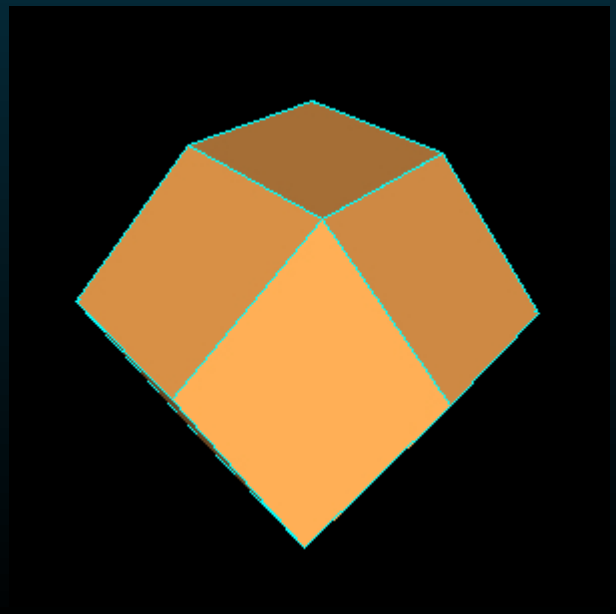
BCC



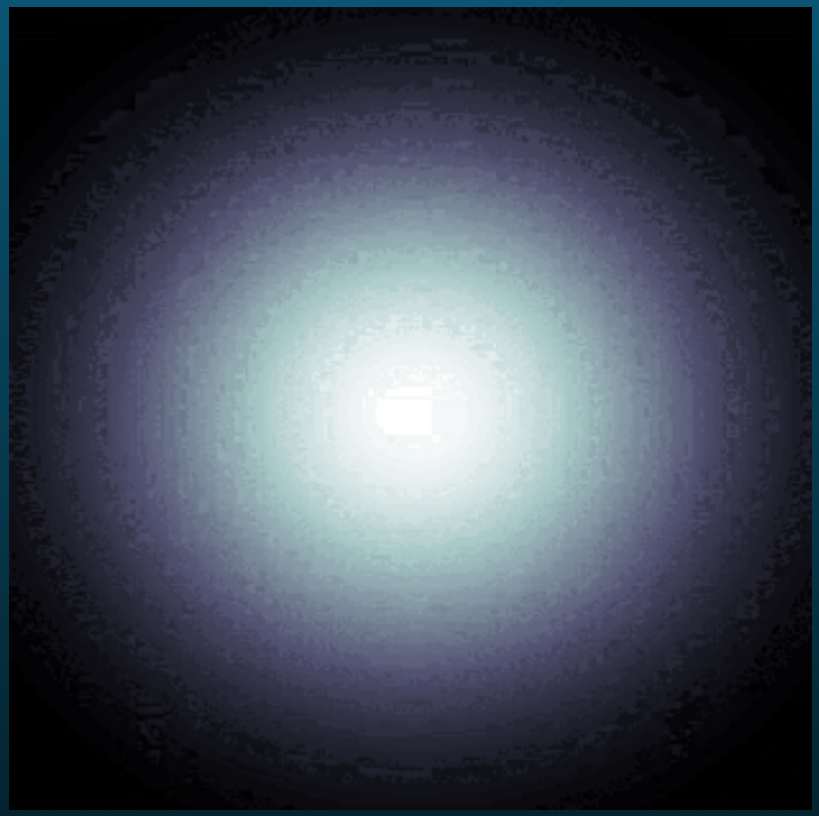
FCC

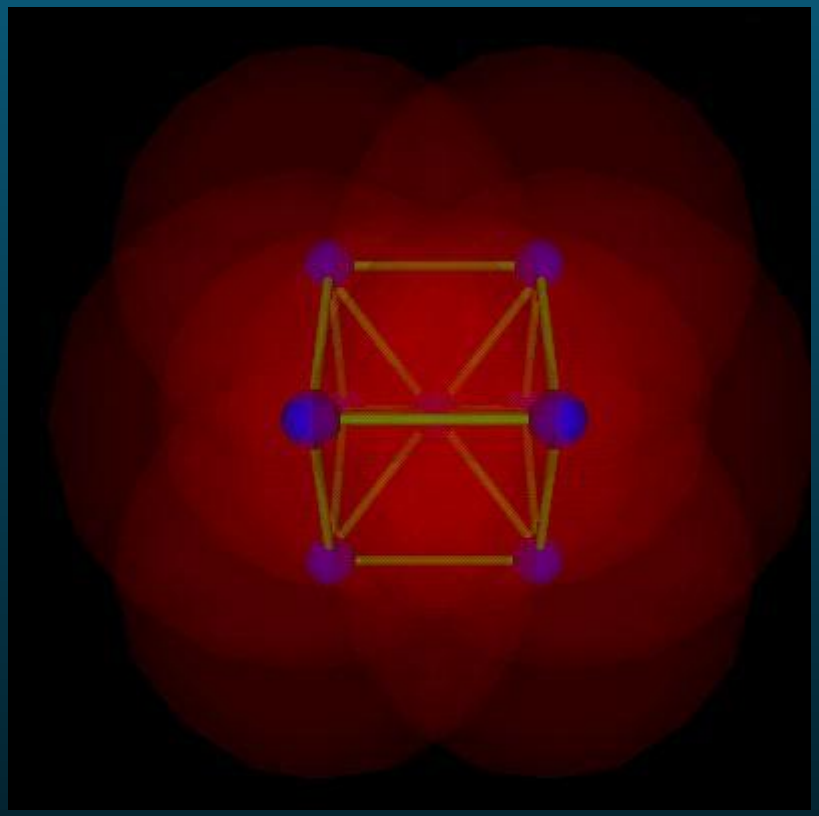






Blobs

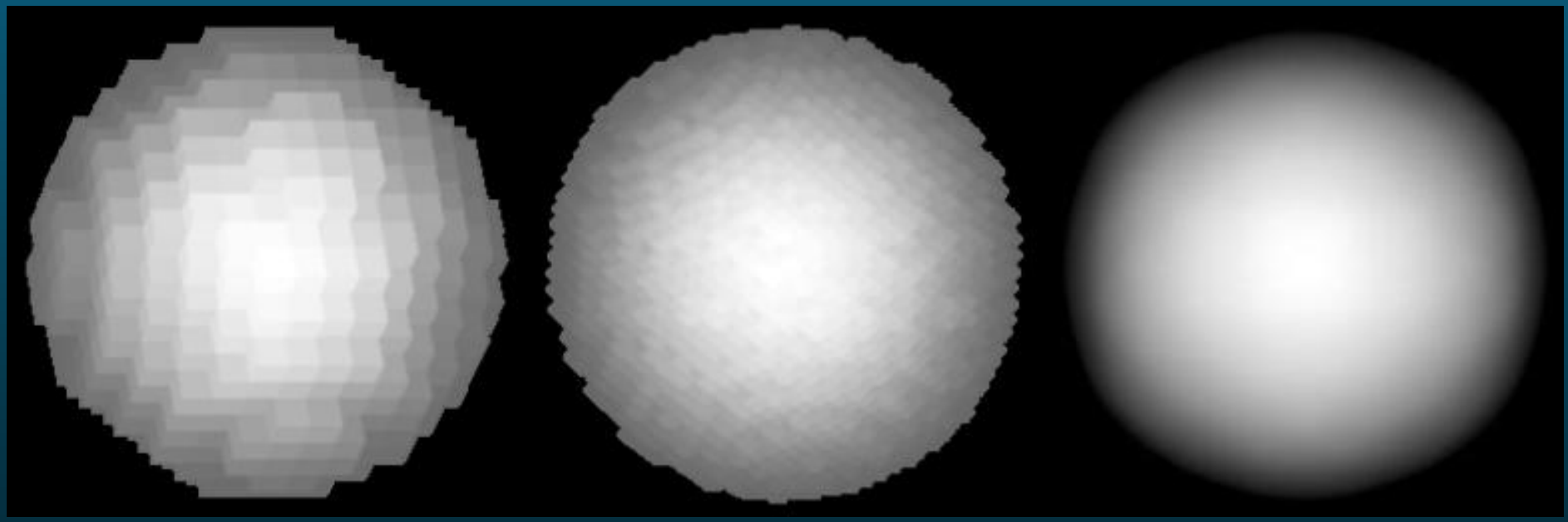




Kaiser-Bessel window functions: rotationally symmetric basis functions, smooth, localized in space, effectively localized in Fourier space, continuous derivatives, analytic calculation of: projection, gradient and Laplacian

Lewitt, JOSA, 7:1834-1846

Displaying with Blobs



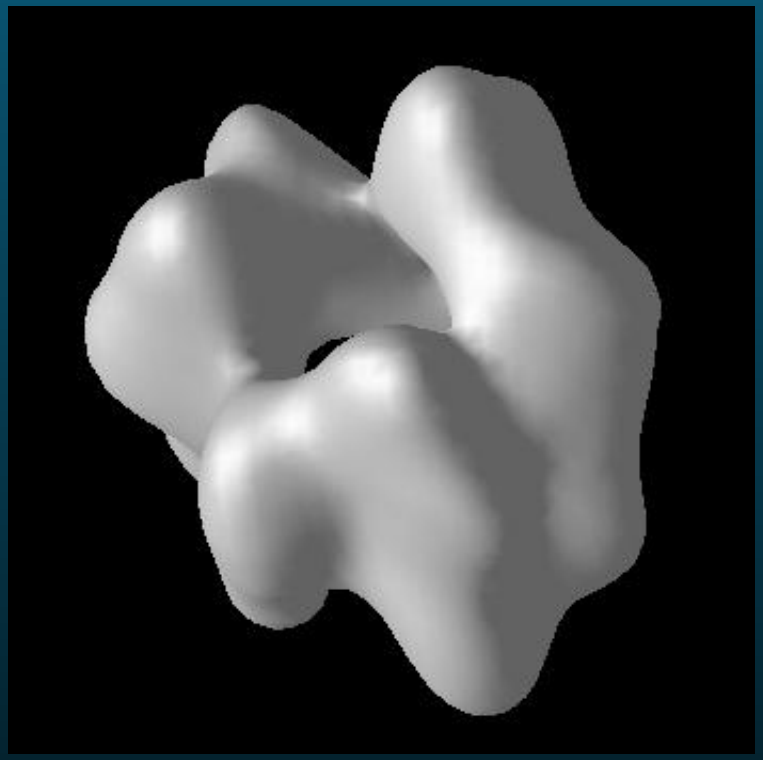
SC FCC+Vox FCC+Blobs

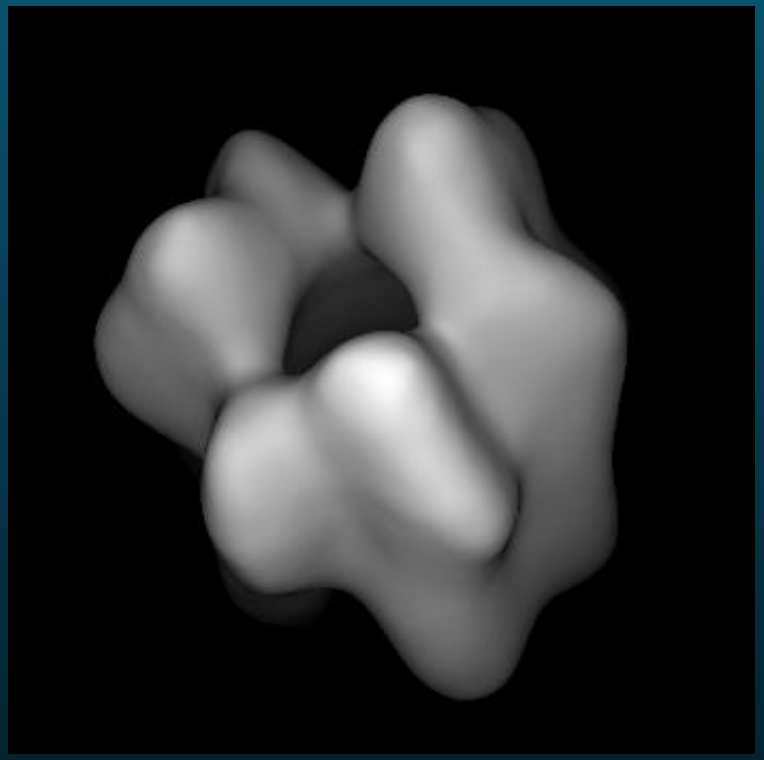
Computer graphic display of a sphere based in different grids and basis functions. Left: the cubic grid $2\mathbb{Z}^3$. Middle: The FCC grid $F_{4^{\frac{3}{2}}}$ (i.e., the same volume $voxel$ than in the cubic grid $2\mathbb{Z}^3$). Right: same FCC grid using blobs. In all cases the display has been produced in the square grid $\frac{1}{8}\mathbb{Z}^3$

Gabor T. Herman, *Geometry of Digital Spaces*, Birkhäuser, Berlin

Garduño *et al.*, IWCIA 2001 proceedings 425-441

Displaying with Blobs





Surface rendering of the implicit surface for the reconstruction of the macromolecule complex DnaB-DnaC using $OpenDX$ and a methodology directly based on Blobs

Barcena $et\ al.$, EMBO J. 2001 Mar 15;20(6):1462-8.

Garduño $et\ al.$, IWCIA 2001 proceedings 425-441

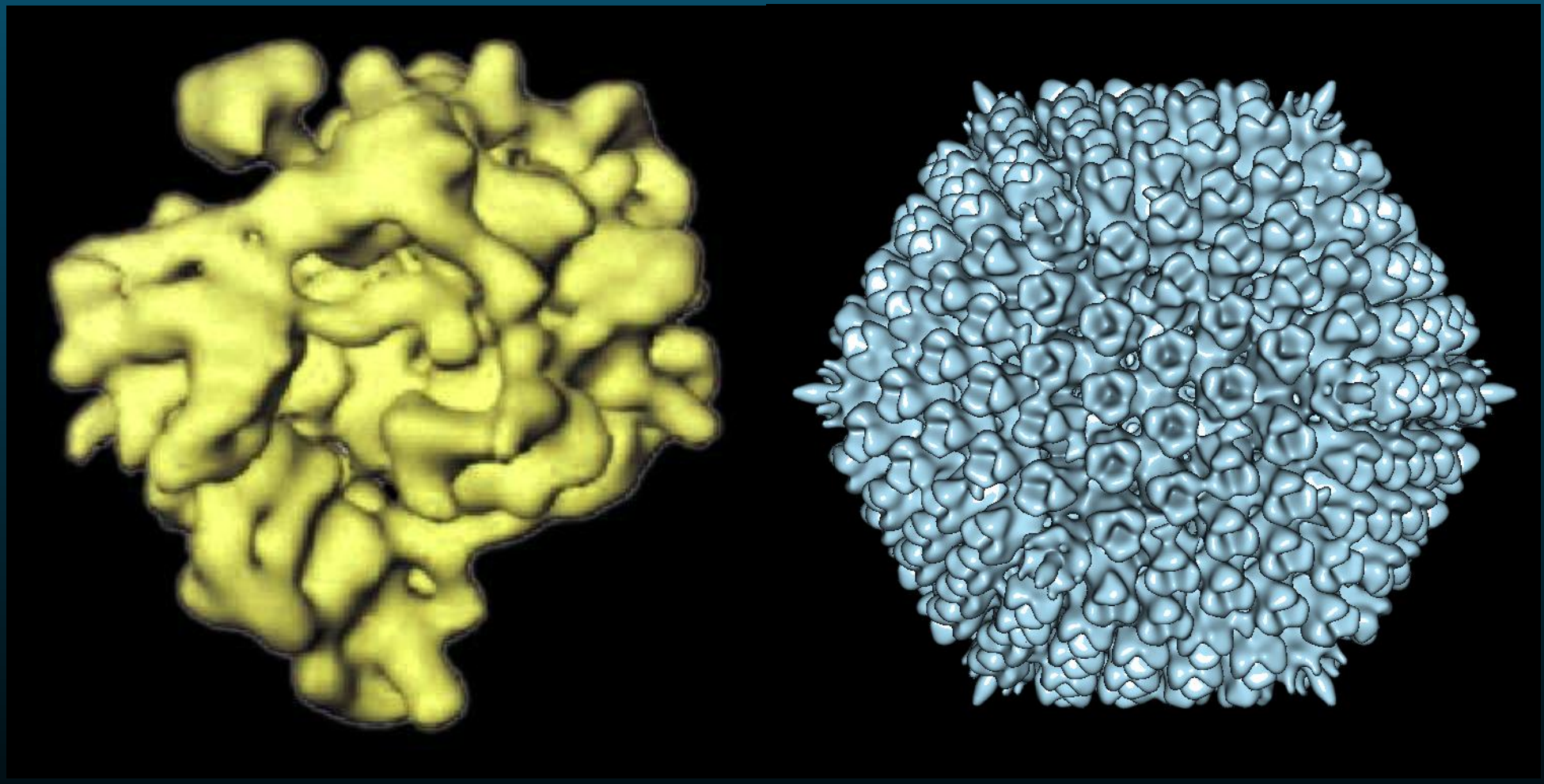
Why do we use ART?

- Matej *et al.* Phys. Med. Biol. (1994) 355-367
- Kinahan *et al.* IEEE Trans. Nucl. Sci. 42 (1995) 2281
- Matej and Lewitt, IEEE Trans. Nucl Sci. 42 (1995) 1361
- Marabini *et al.* J. Struc. Biol. 120 (1997) 363
- Marabini *et al.* Ultramic. 72 (1998) 53
- Sorzano *et al.* J. Struc. Biol. (to appear)
- Better framework for incorporation of *a priori* knowledge

Summary

- The very general aim: **Reconstruction of proteins (biological macromolecules) from images (projections) obtained with an Electron Microscope**.
- 3D Reconstruction using **ART**. (What is ART and why should we try it?)
- Different samples require different reconstruction techniques. (Or, How did we adapt ART for single particles, viruses, crystals...)
- Implementing Constraints (extra information independend of the projections or symmetry)

Sample Geometry (symmetry) and 3D Reconstruction-I

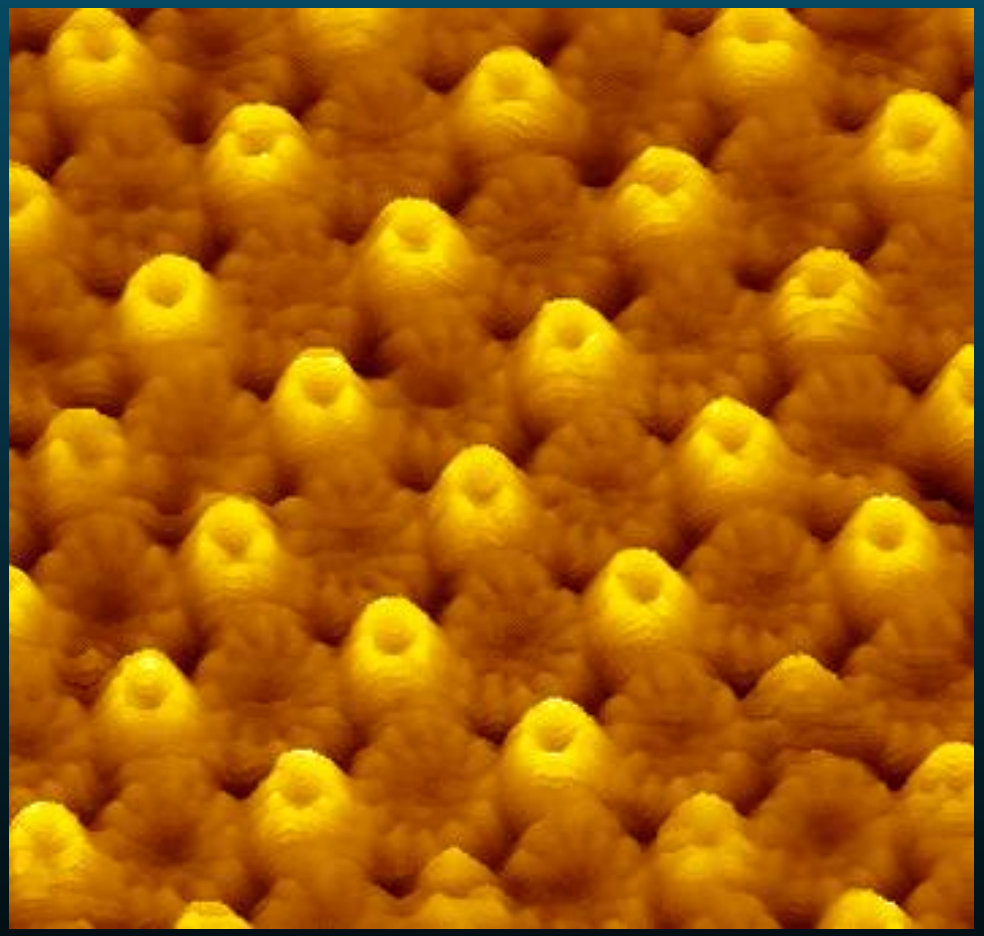


Ribosome (Low)

Icosahedral Virus (Medium)

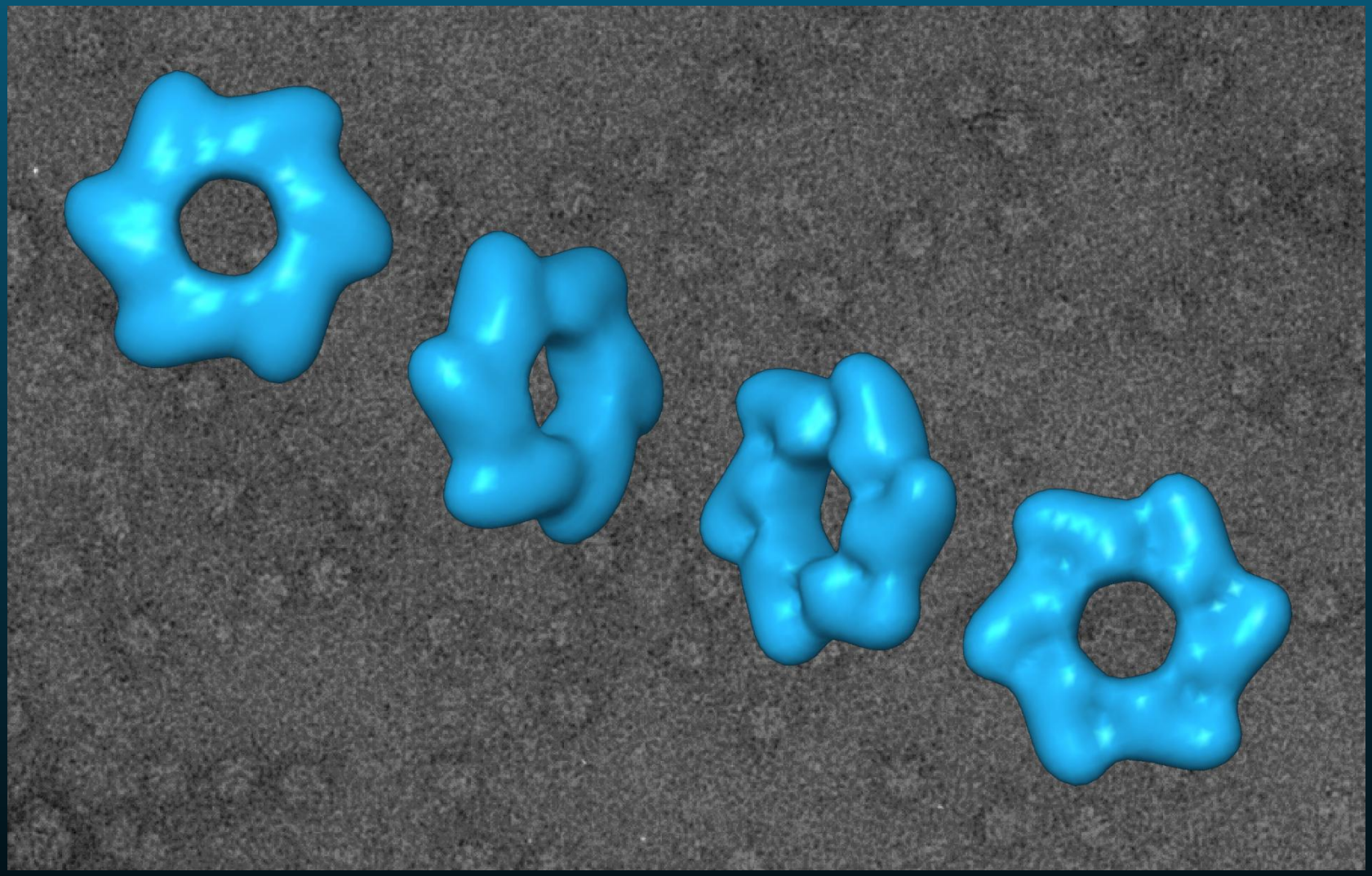
Sample Geometry (symmetry) and 3D Reconstruction-II

Helix (High)



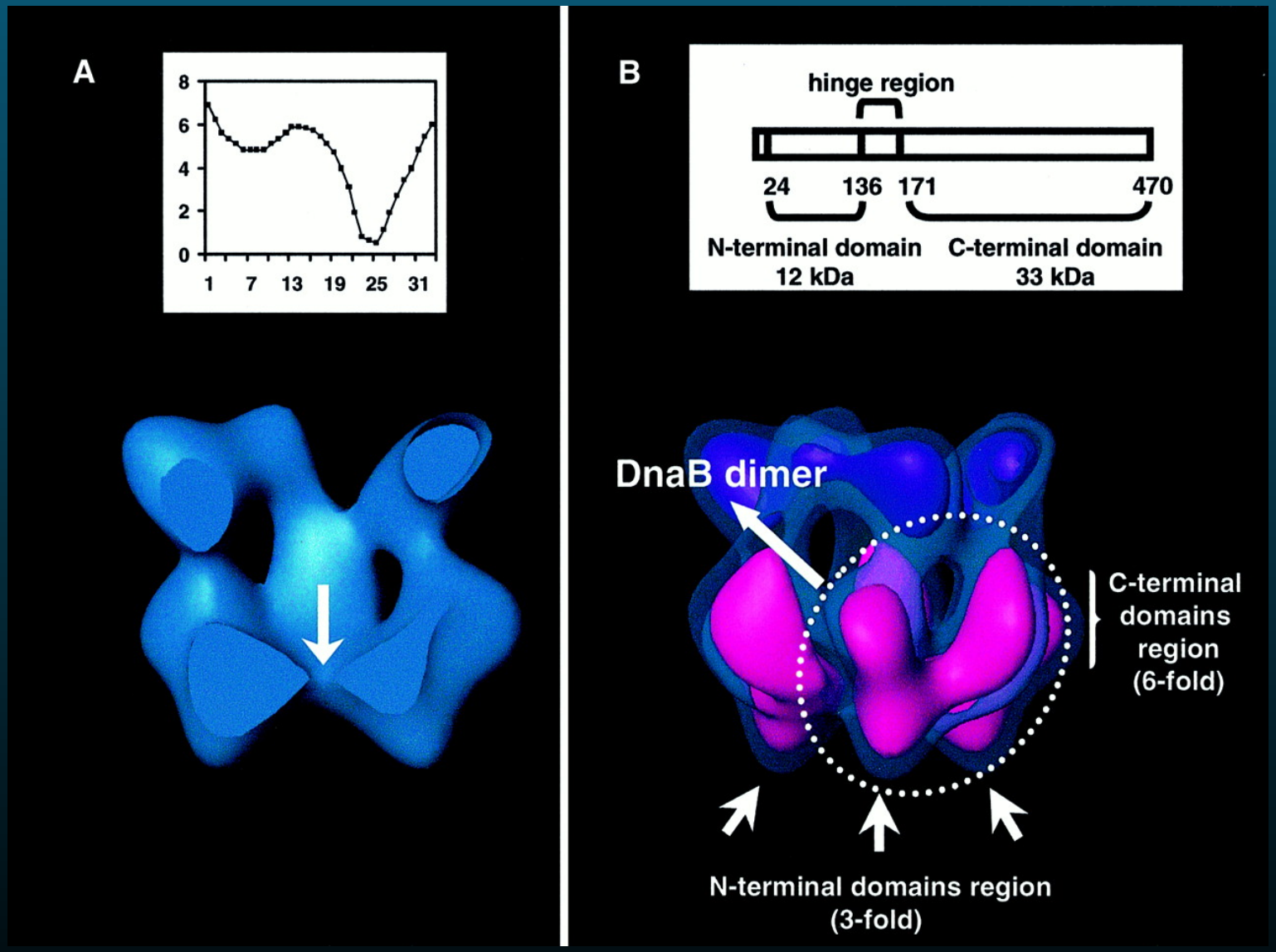
Crystal (High, but different)

ART for Single Particles-I



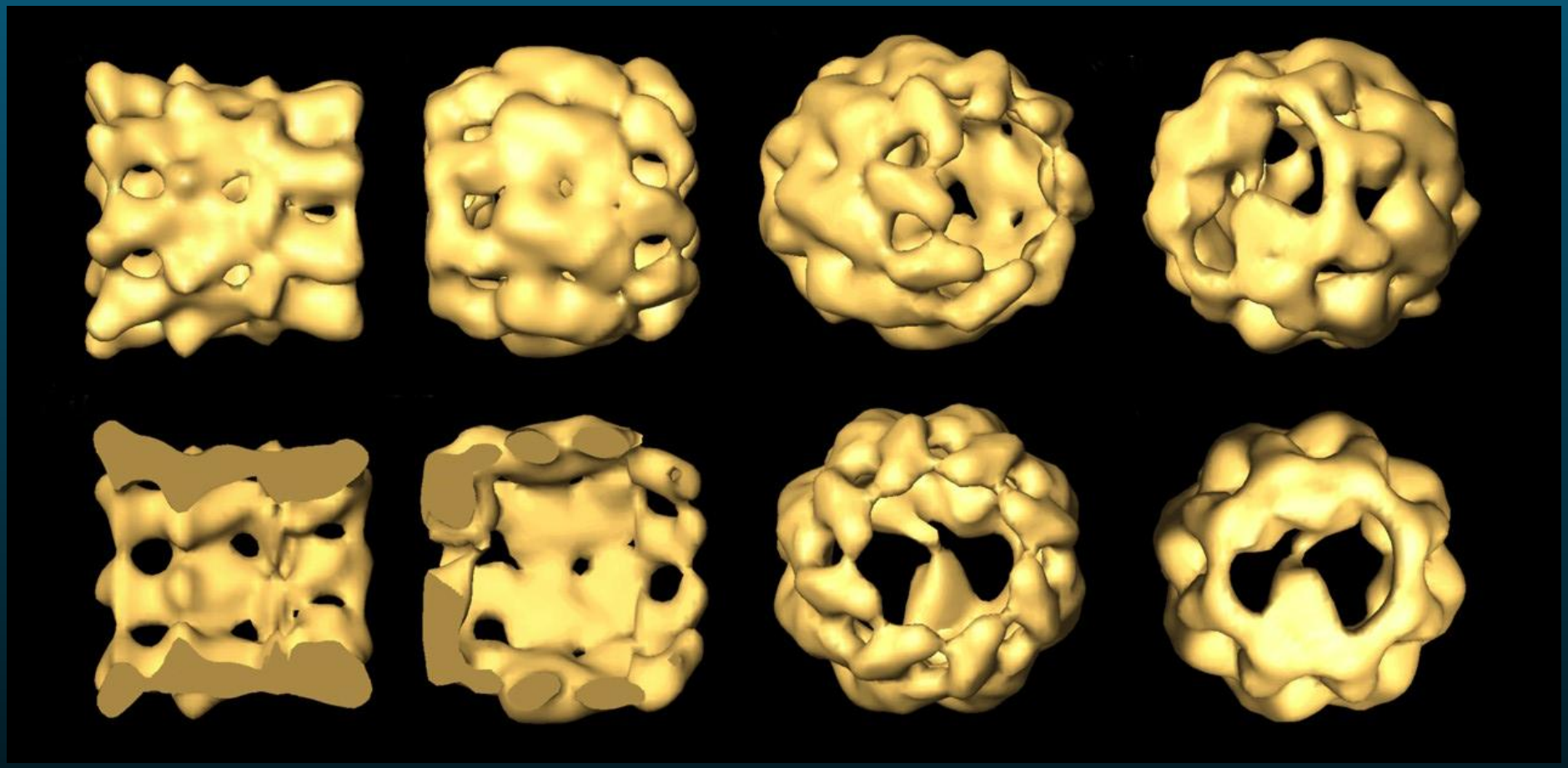
SanMartin *et al.* J. Mol. Biol., 268, 15-20 (1997)

ART for Single Particles-II



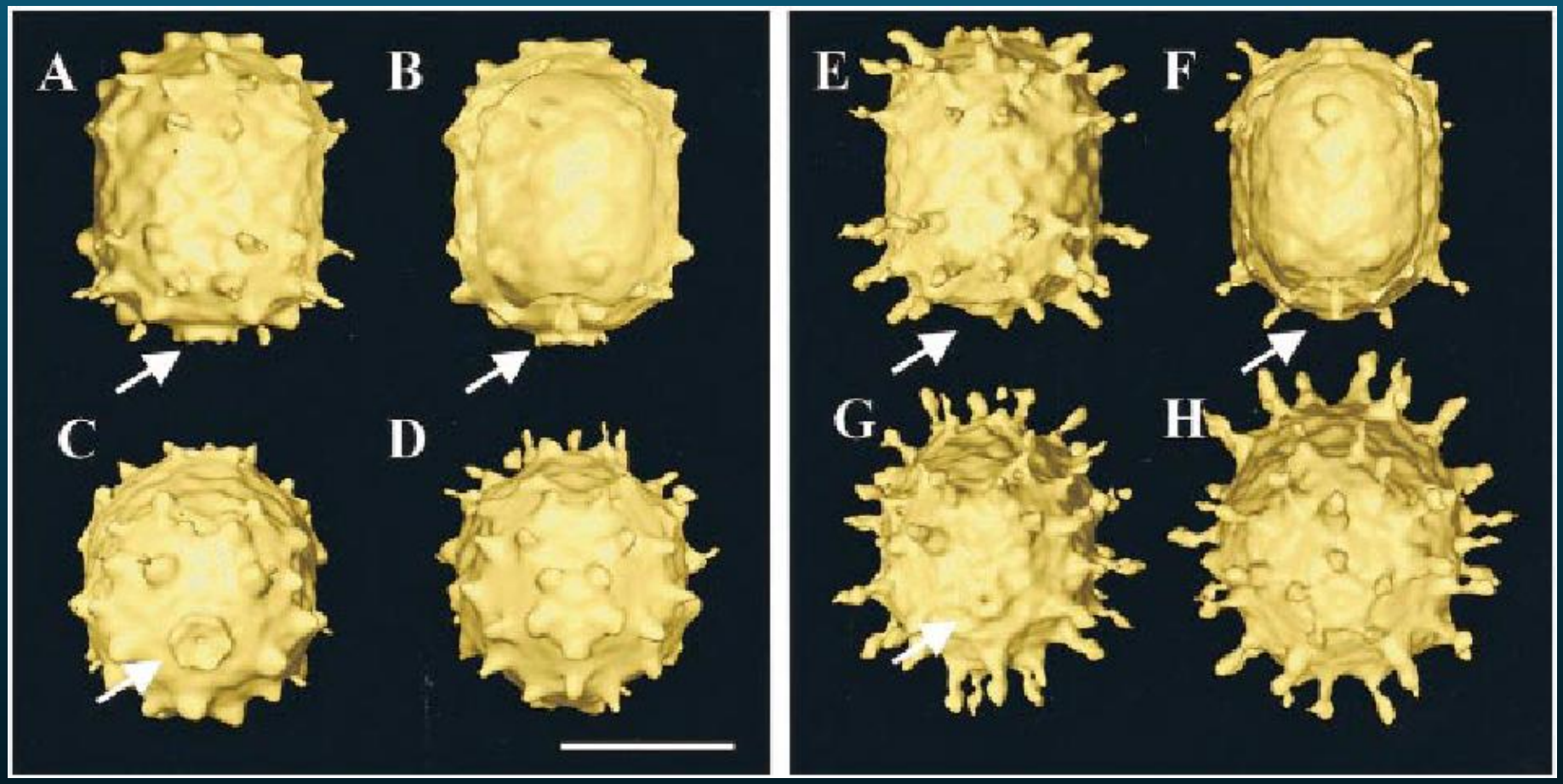
Barcena *et al.* EMBO J 2001 Mar 15;20(6):1462-8

ART for Single Particles-III



Llorca *et al.* EMBO J 2000 Nov 15;19(22):5971-9

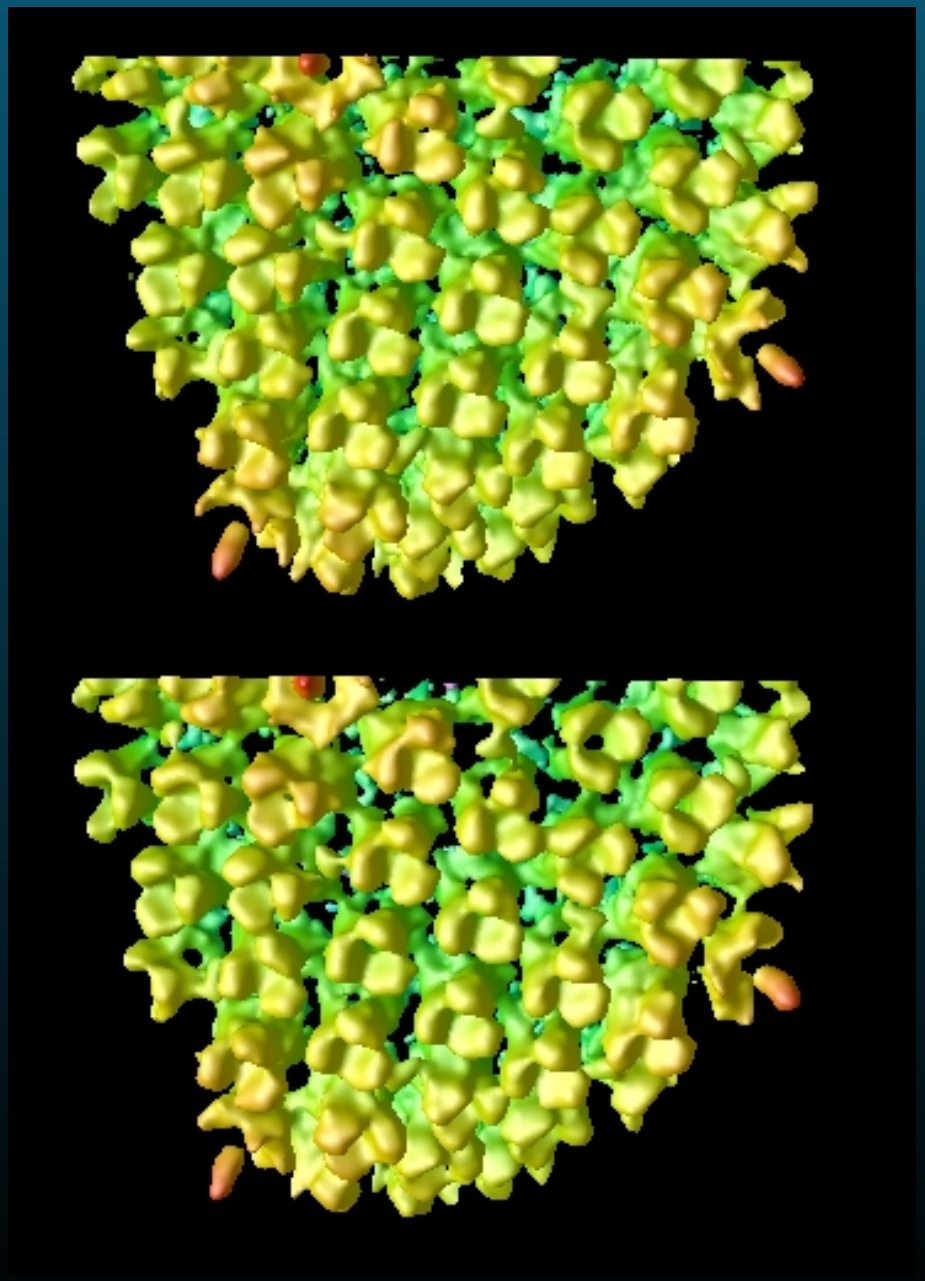
ART for Single Particles-IV



Ibarra *et al.* J Mol Biol 2000 May 19;298(5):807-15

Connector phi-29

Virus

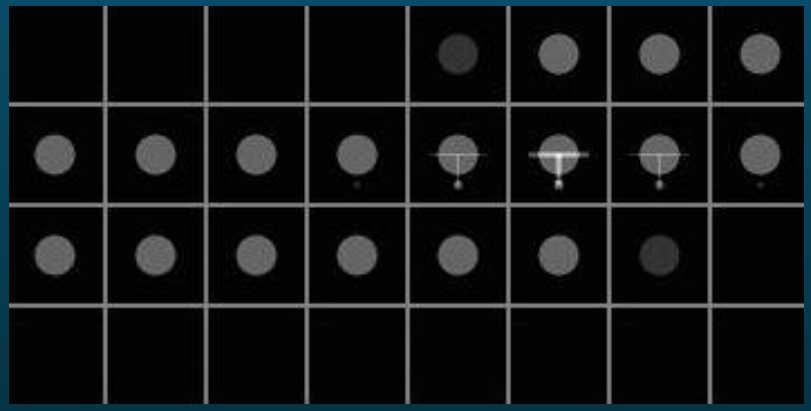


- Symmetry (6,5 10,3 15,2)
- Smooth
- CPU time: 4-days
  - 315^3
  - $200 * 60$
- Speed up
  - Parallelization?
  - Data Reorganization?
  - Cache Memory?

Summary

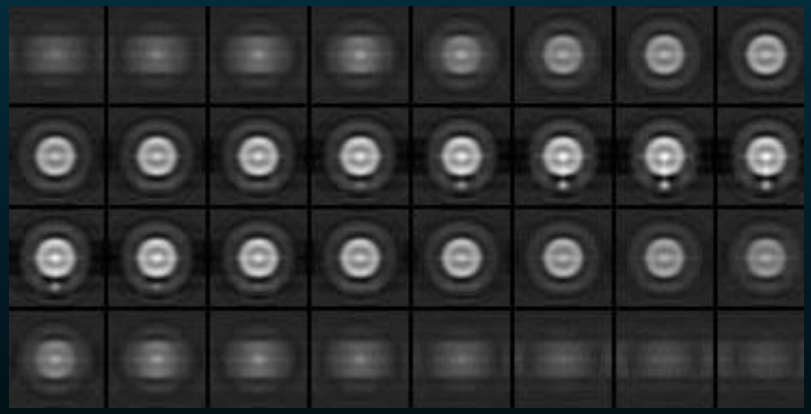
- The very general aim: **Reconstruction of proteins (biological macromolecules) from images (projections) obtained with an Electron Microscope**.
- 3D Reconstruction using **ART**.
- Different samples require different reconstruction techniques.
- Implementing Constraints (extra information independend of the projections or symmetry)
  - ⋆ Modeling the microscope aberrations (CTF)
  - ⋆ Spatial Constraints

Modeling the microscope aberrations (CTF)

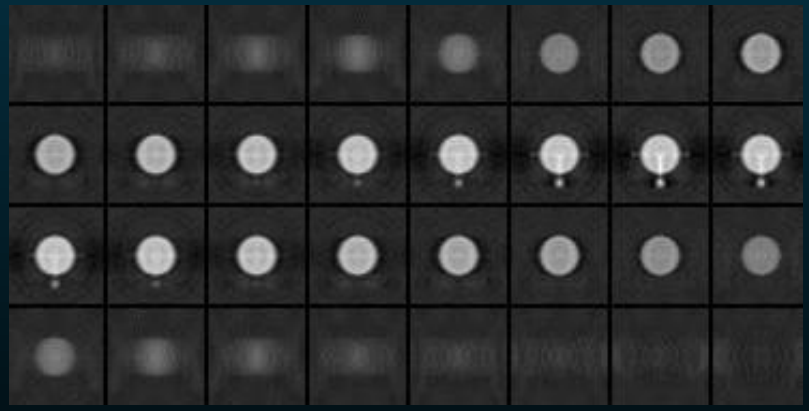


Phantom

Phantom$\otimes FT^{-1}$(CTF)



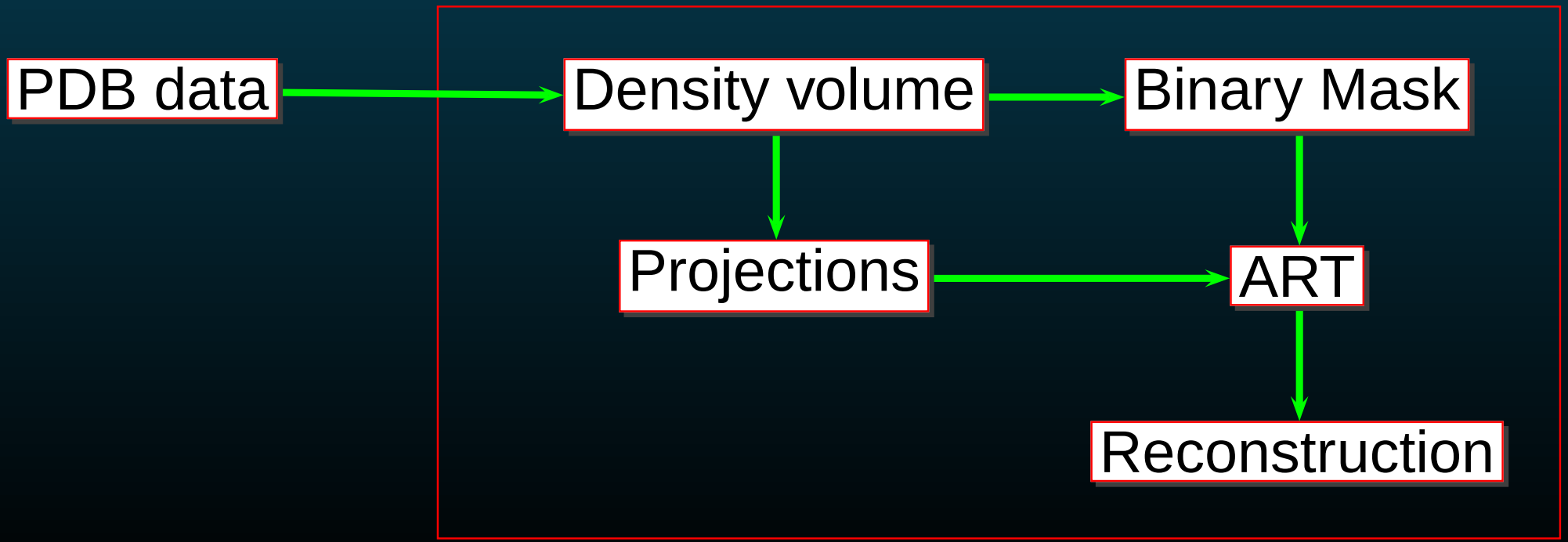
Phantom$\otimes FT^{-1}$(CTF*SGN(CTF))



IDR

Introducing Spatial Constrain in the Reconstruction Process

There are some techniques (as FM or MS) that give very accurate information about the surface of the proteins but not the interior. This constrains can be incorporate easily in ART (or other iterative and real space methods).



Summary

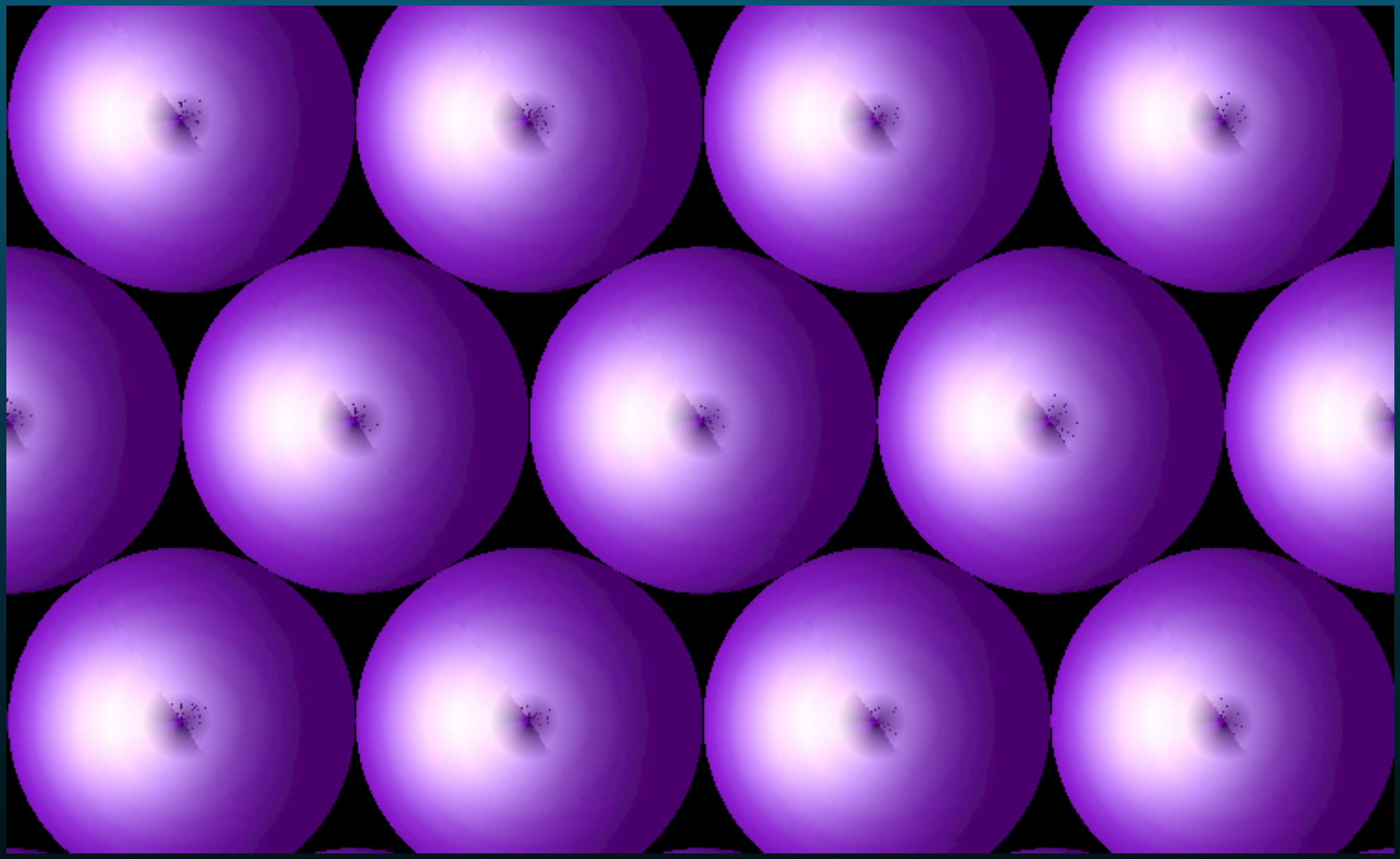
- The very general aim: **Reconstruction of proteins (biological macromolecules) from images (projections) obtained with an Electron Microscope**.
- 3D Reconstruction using **ART**. (What is ART?)
- Different samples require different reconstruction techniques.
- Implementing Constraints
- Description of our particular implementation of ART for crystals.

ART for Crystals-0

The analysis by transmission electron microscopy of biological material is inconvenienced by its sensitivity to electron radiation. In order to reduce the damaged caused by the radiation, the electron dose is kept low and consequently the signal to noise ratio is poor.

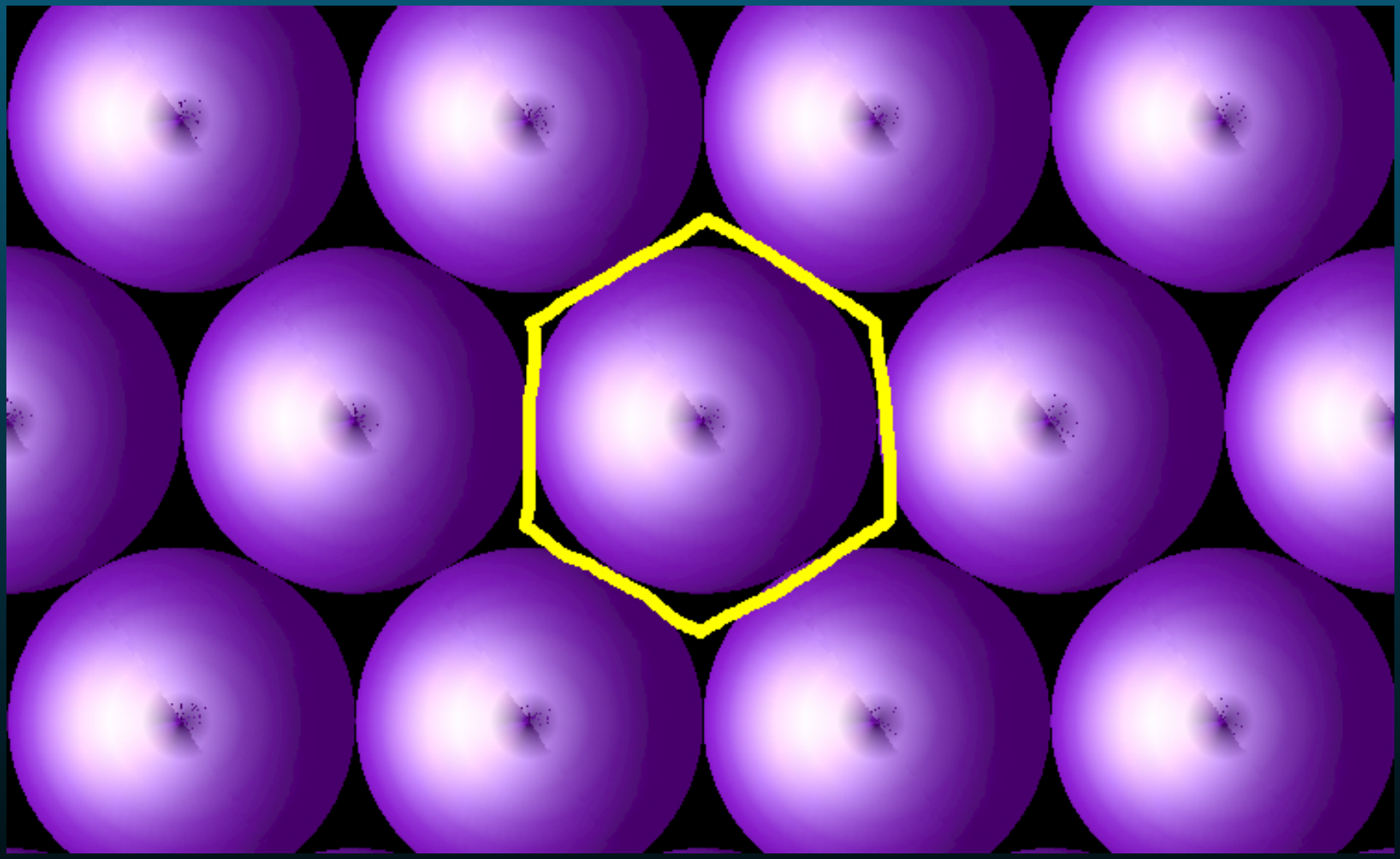
Image processing methods have been developed to counter this problem. Although ideally, these methods can be applied to any image, they are usually most powerful for objects in which subunits are arranged in a regular manner, such a two dimensional crystals

ART for Crystals-I



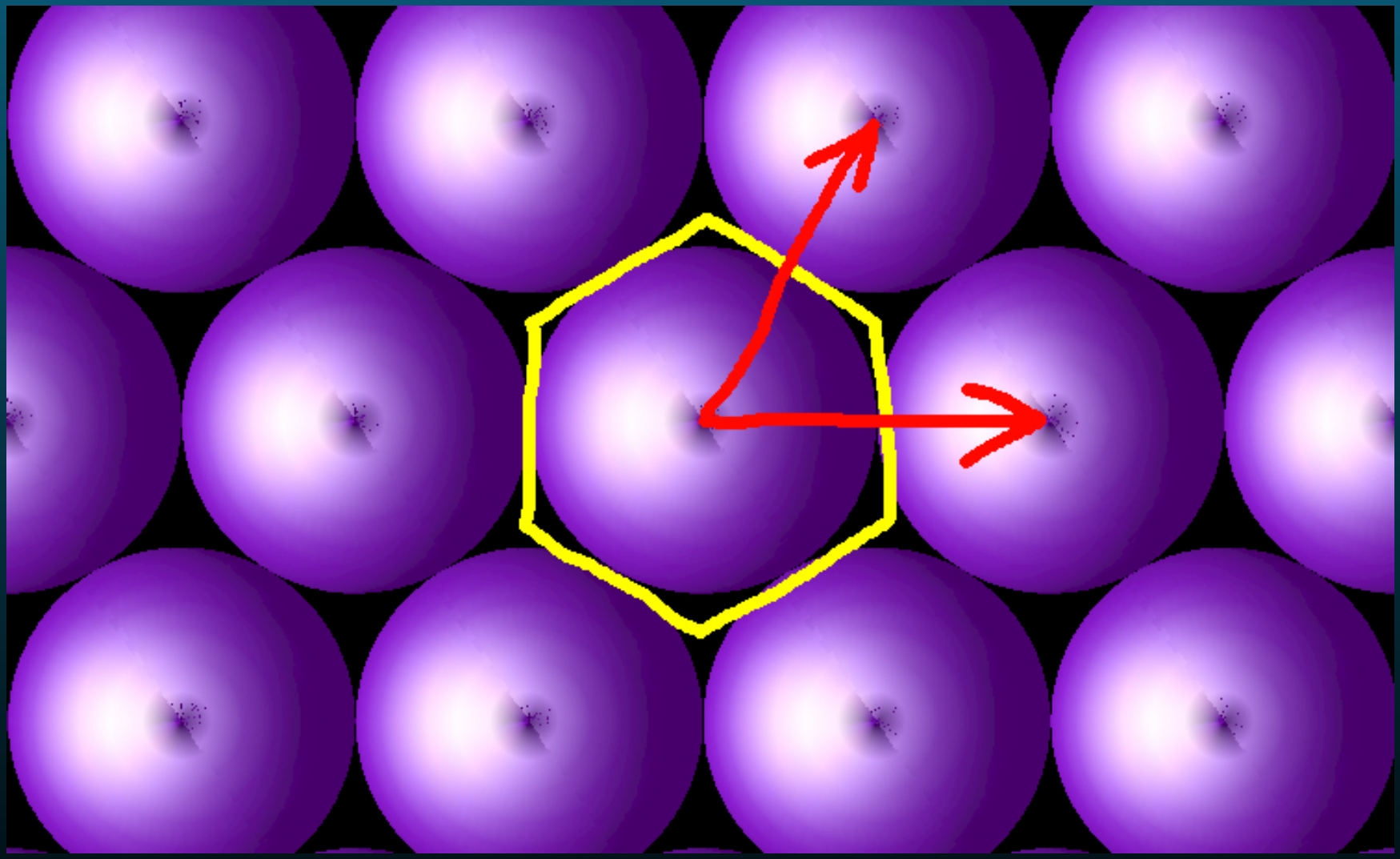
2D-Crystal (motif repets in plane XY)

ART for Crystals-II



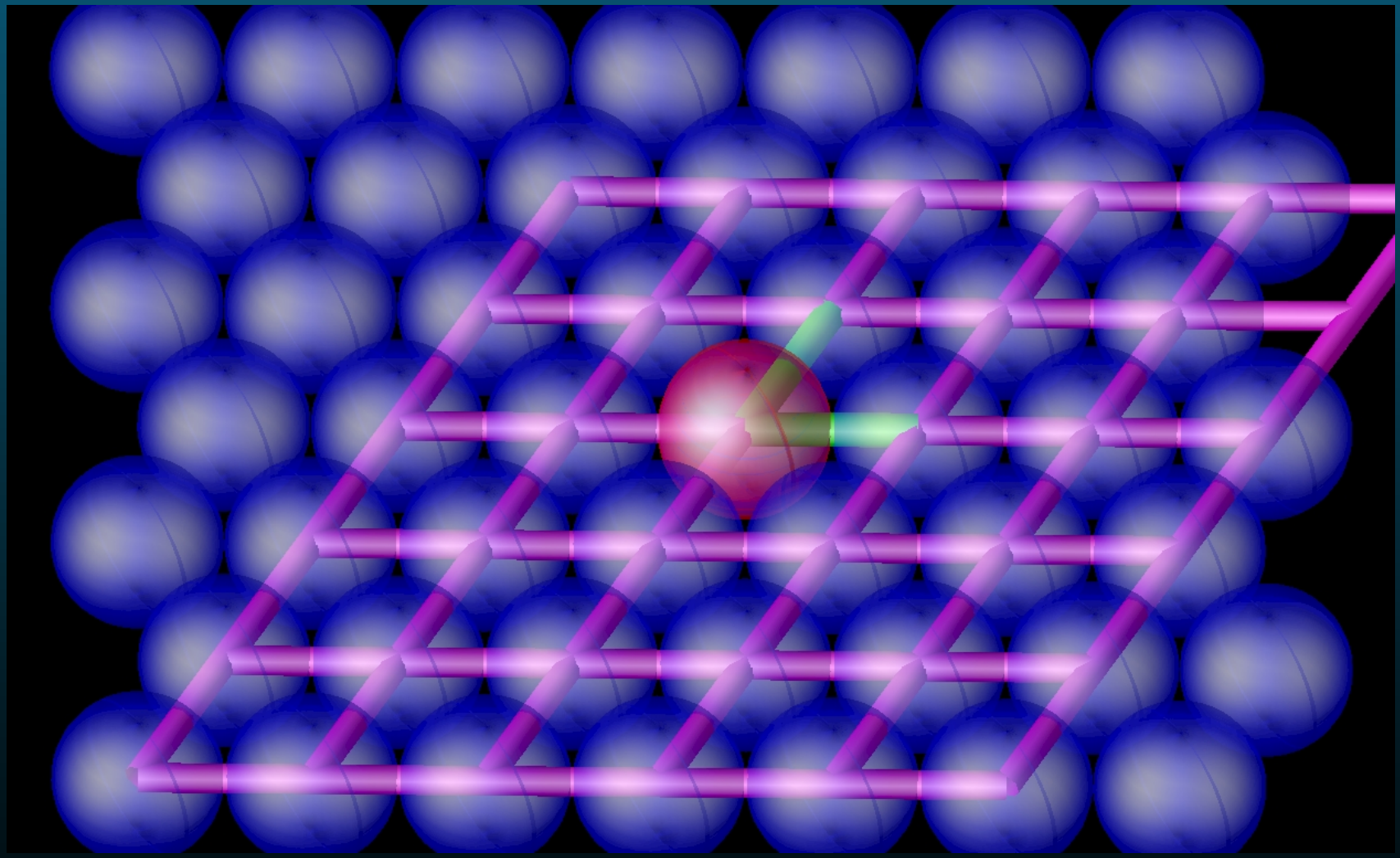
Motif

ART for Crystals-III

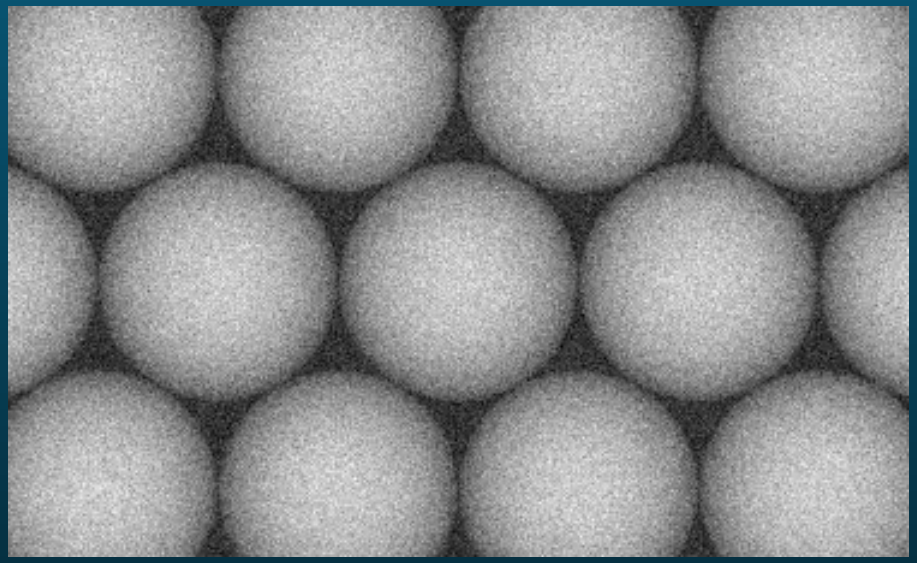
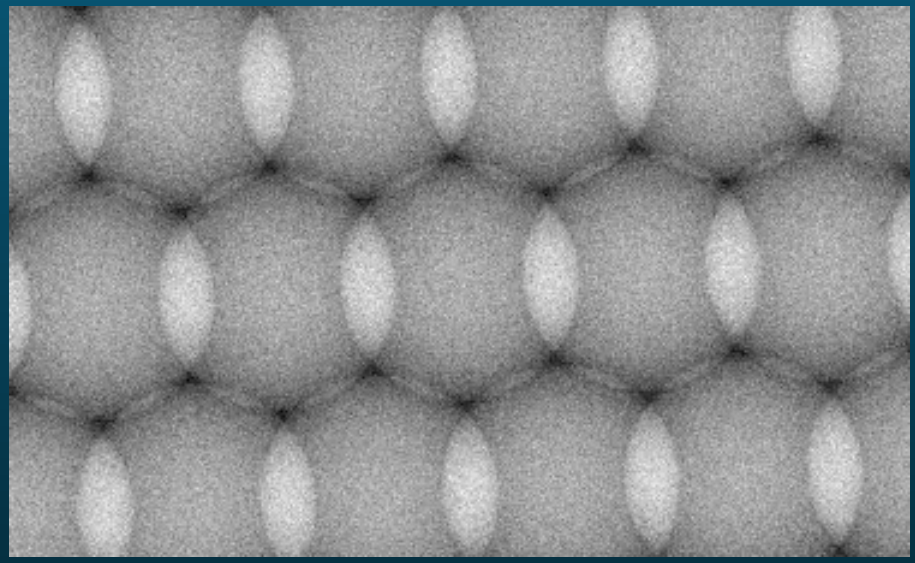
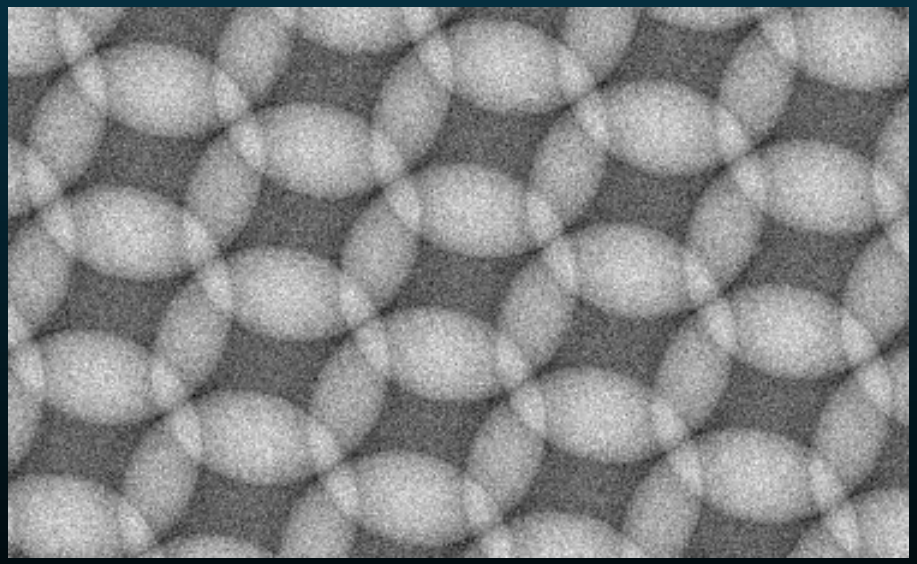
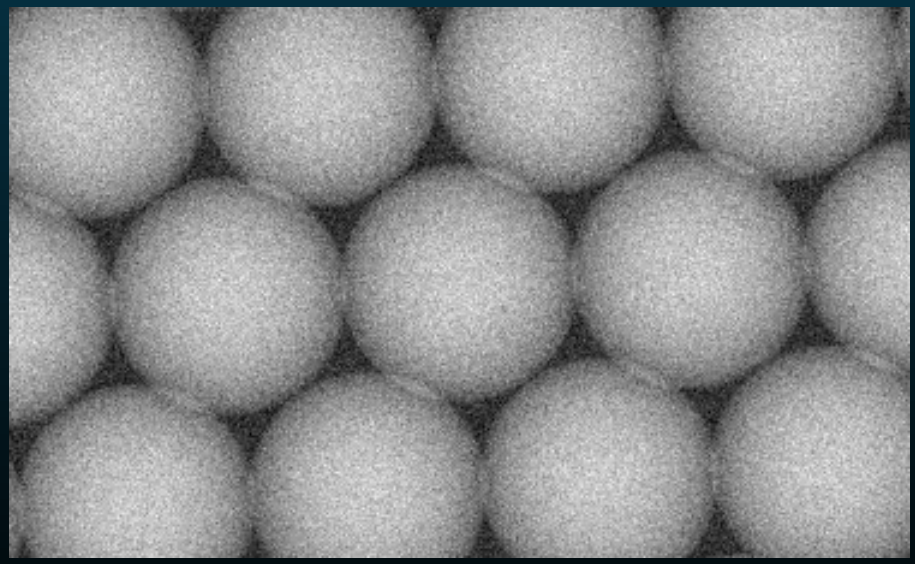


Motif and fundamental vectors

What is a crystal-I?

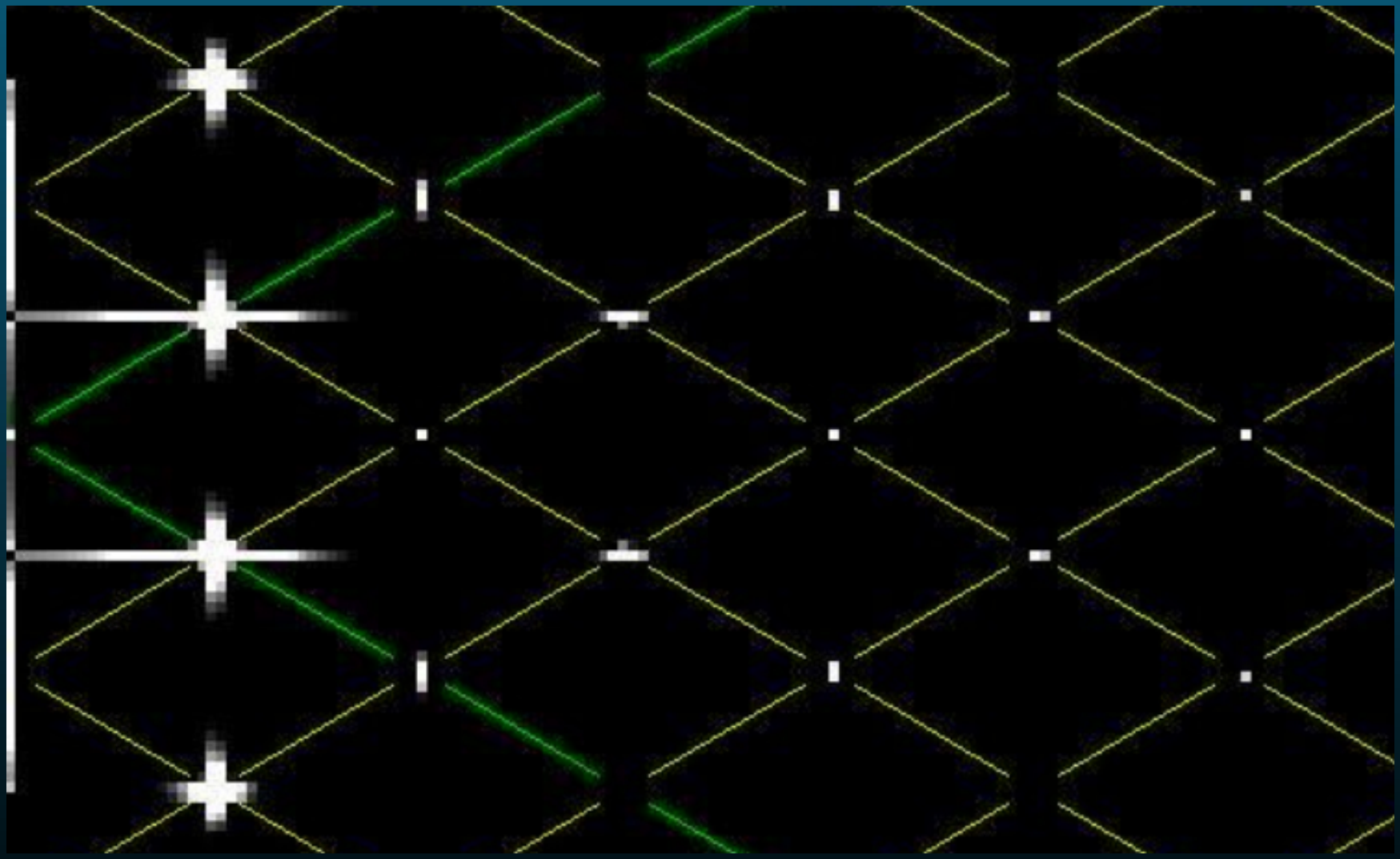


ART for Crystals-I

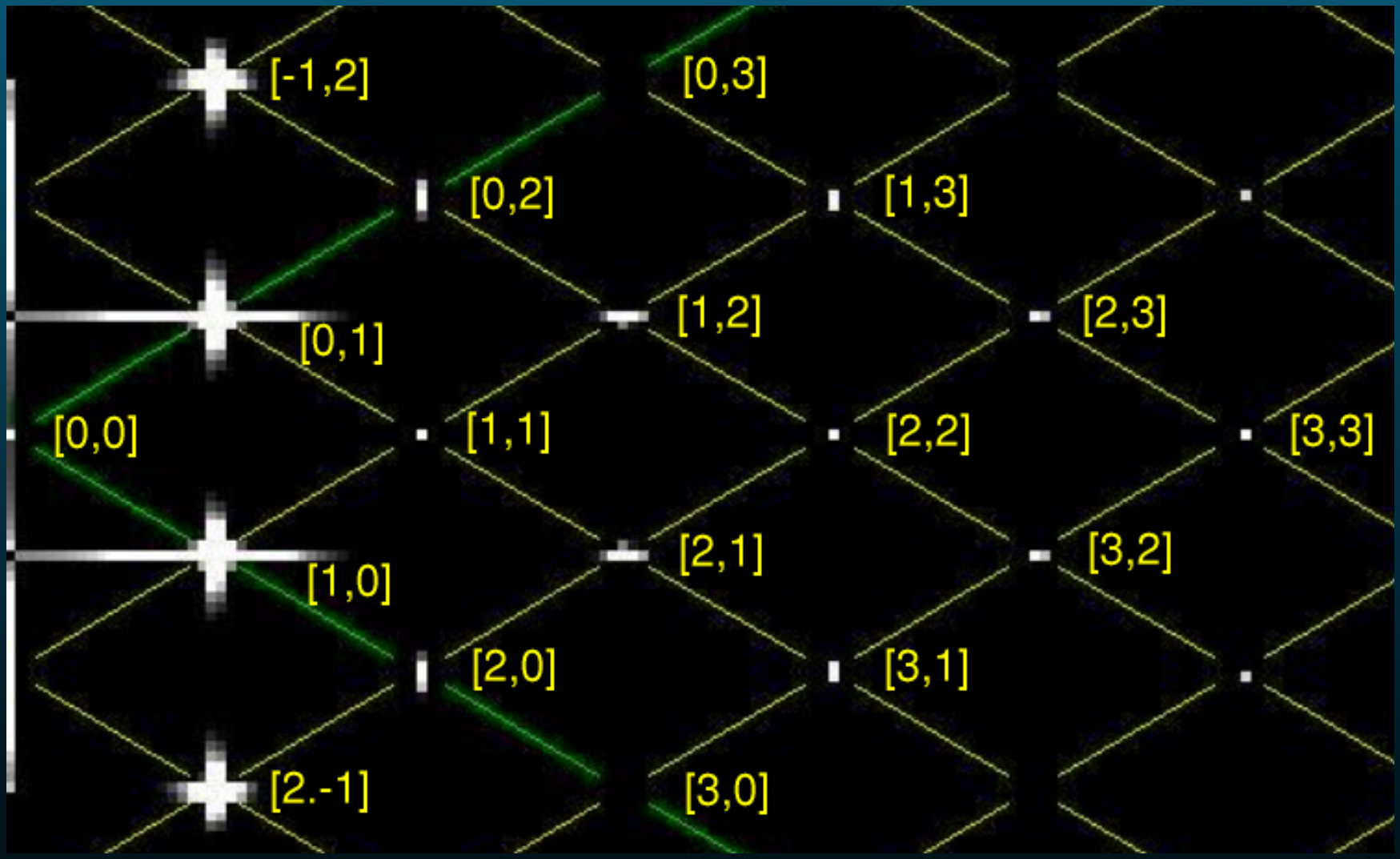
Projections (2D data recorded by the microscope)

ART for Crystals-I



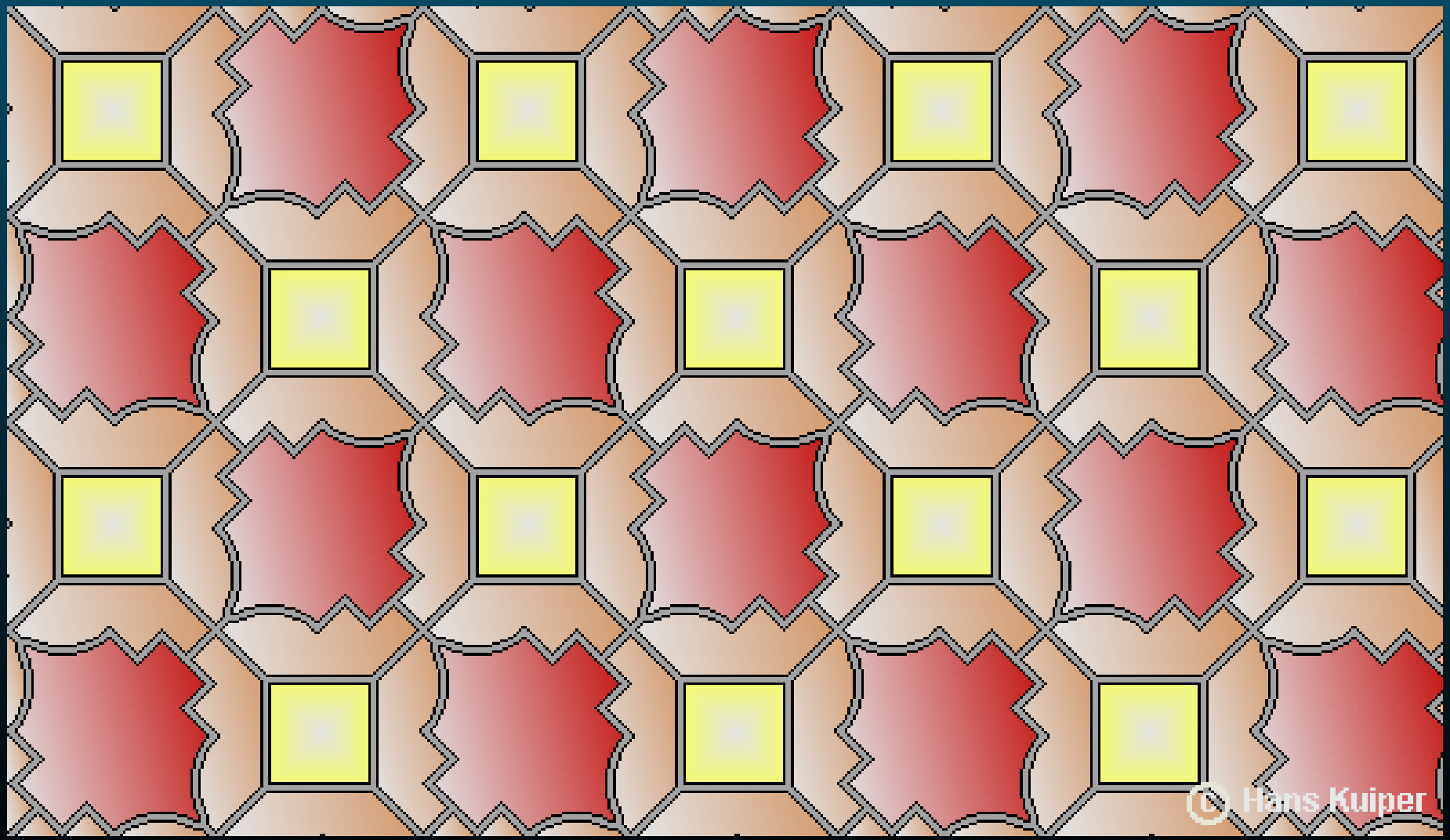
Projection Fourier transform

ART for Crystals-I

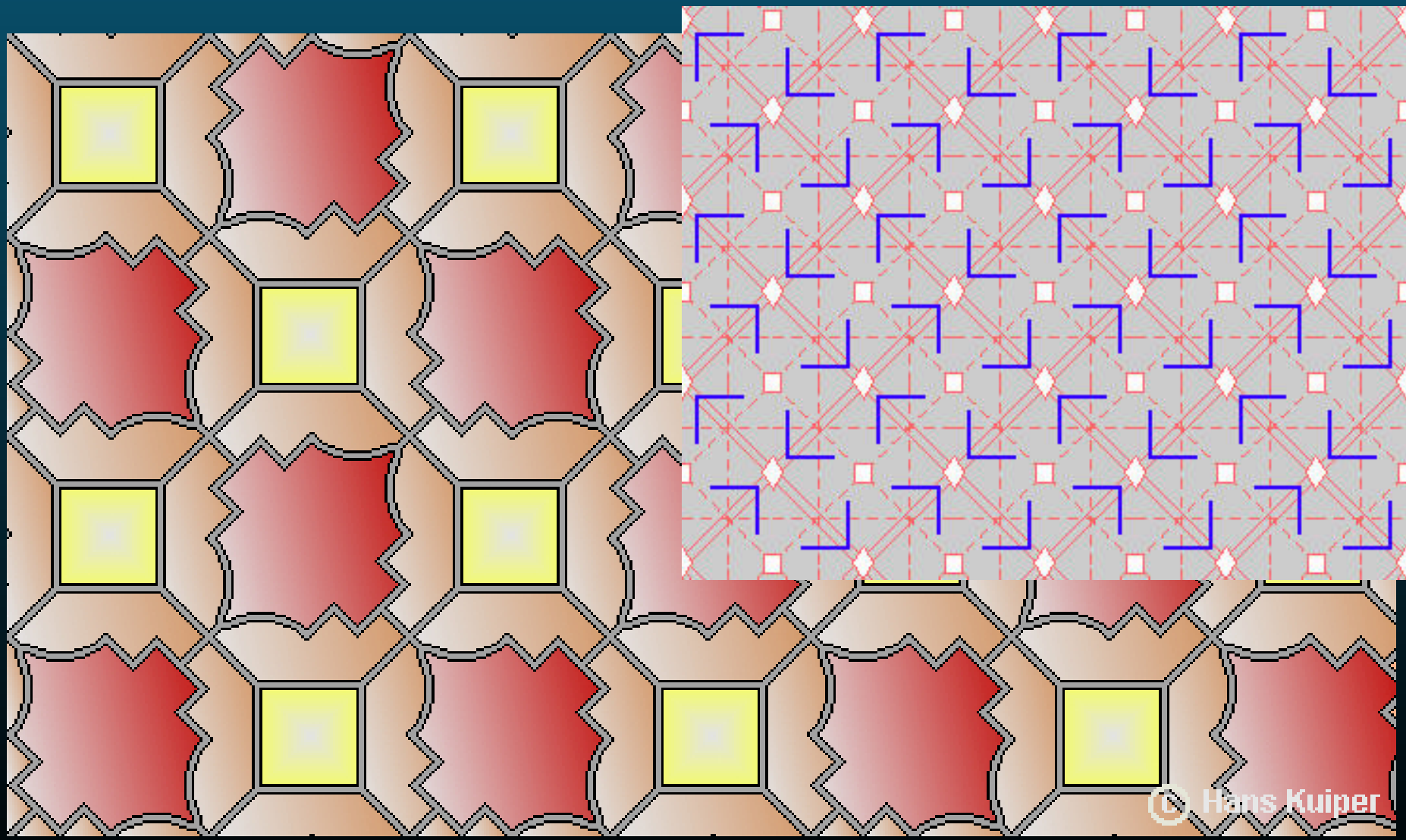


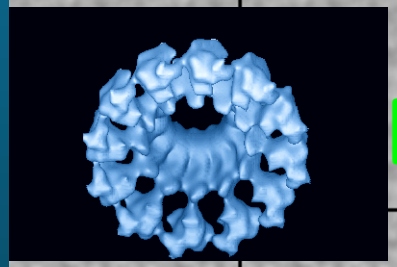
Indexed projection

Symmetry

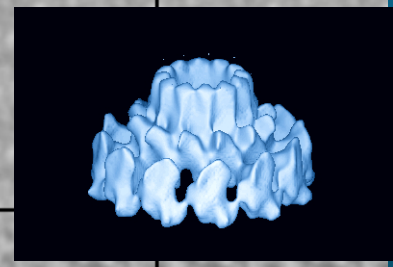


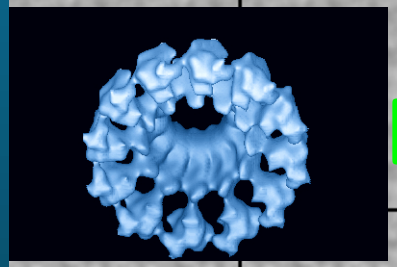
Symmetry



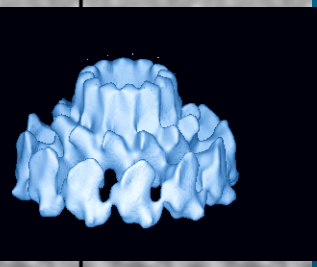


Reconstructions with real-data

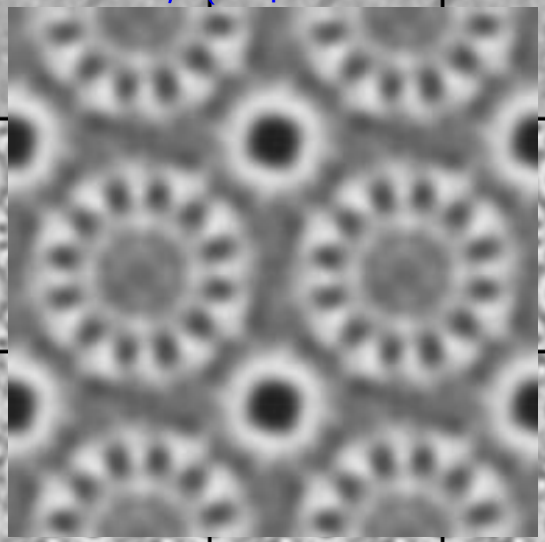


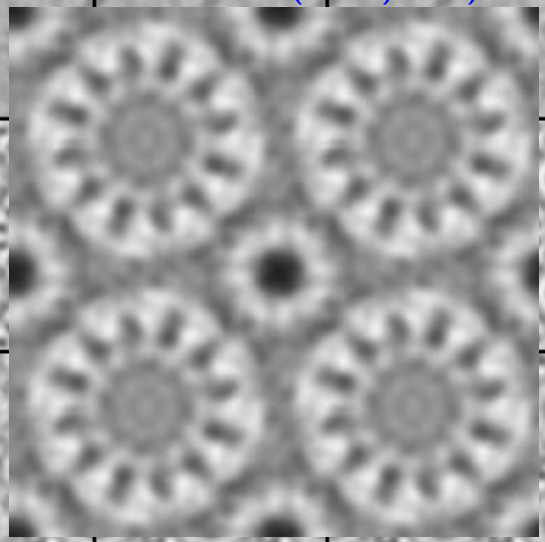


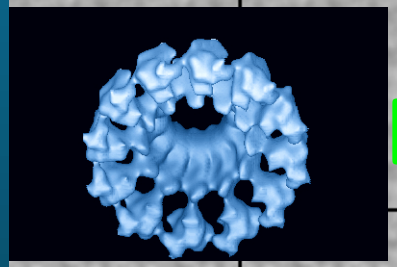
Reconstructions with real-data



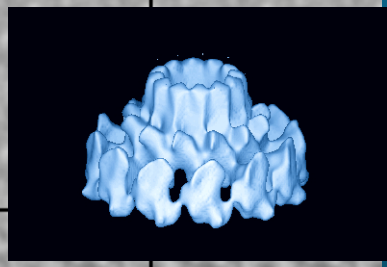
Phi-29, (Valpuesta *et al.* Struct. Fold. & Des. 7 (1999) 289)



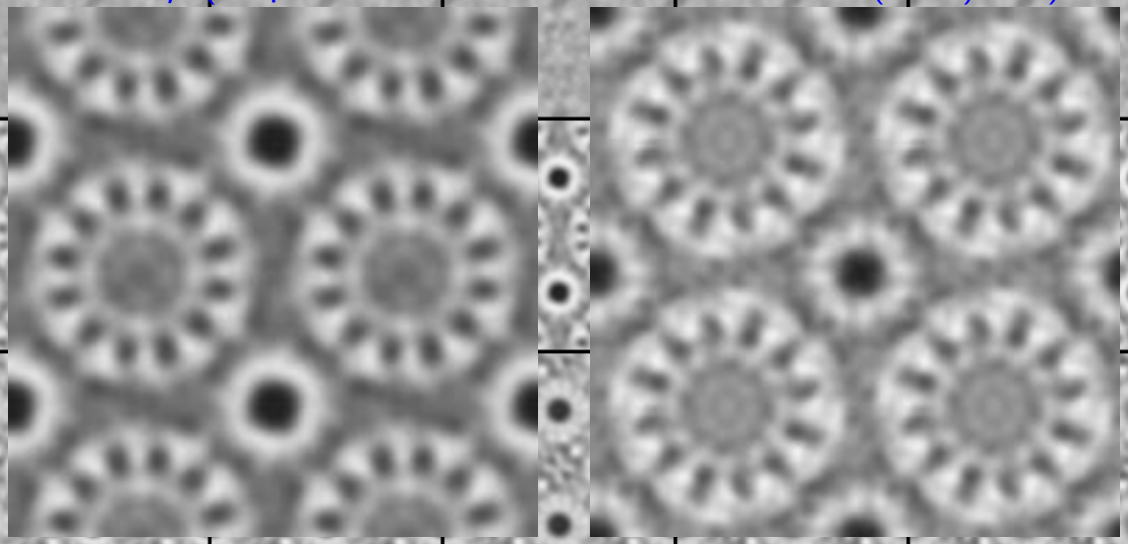




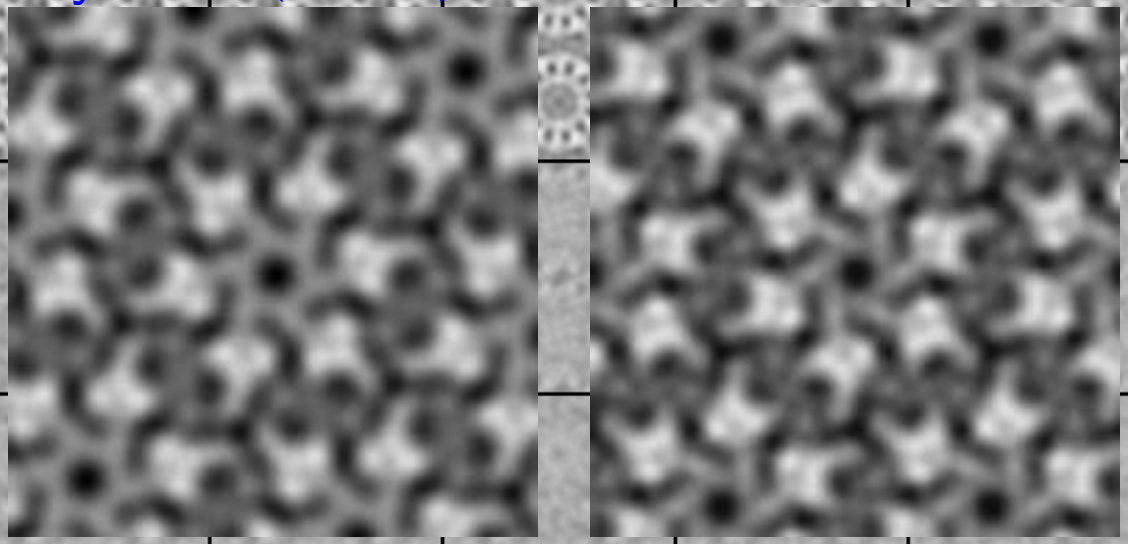
Reconstructions with real-data



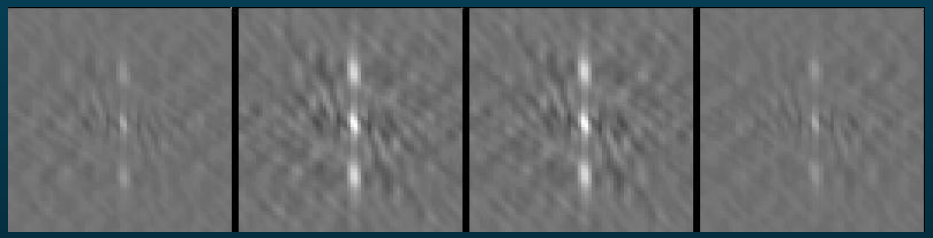
Phi-29, (Valpuesta *et al.* Struct. Fold. & Des. 7 (1999) 289)

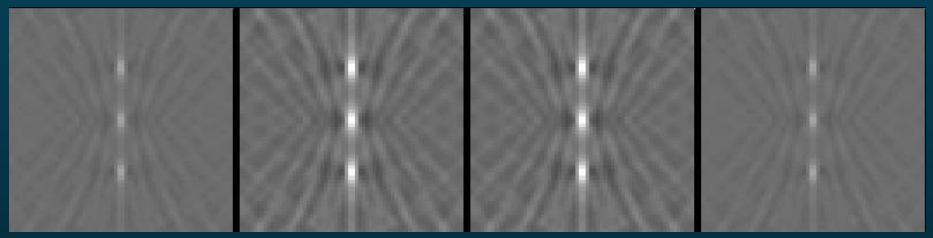


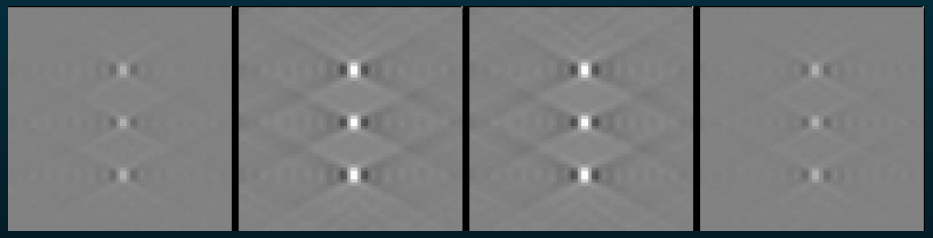
Polyhead, (unpublish)



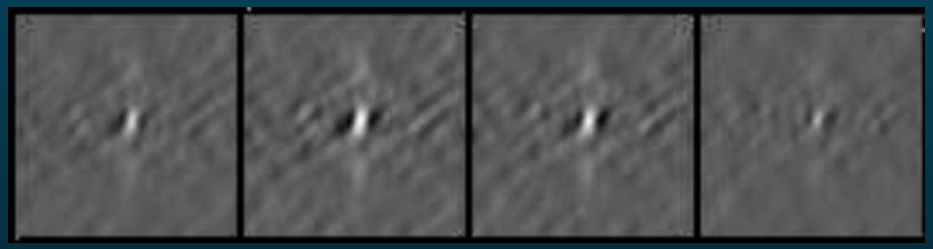
Reconstruction with Phantoms

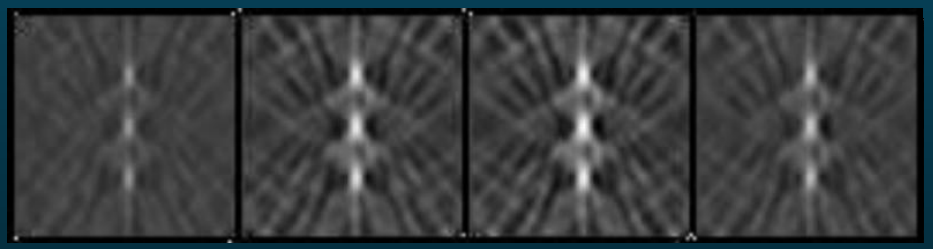


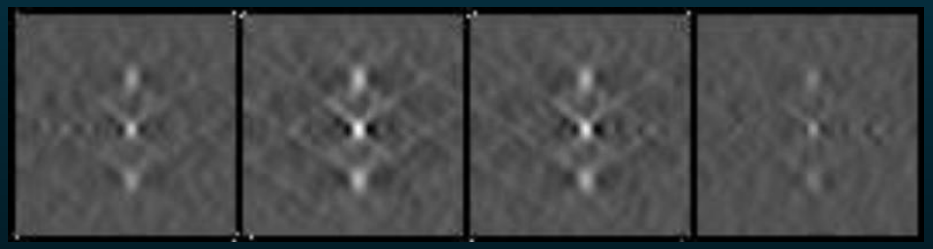


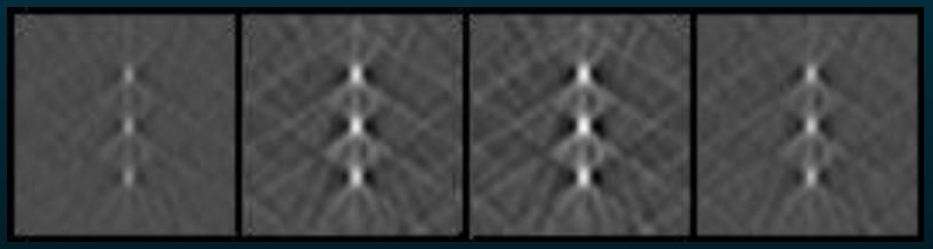


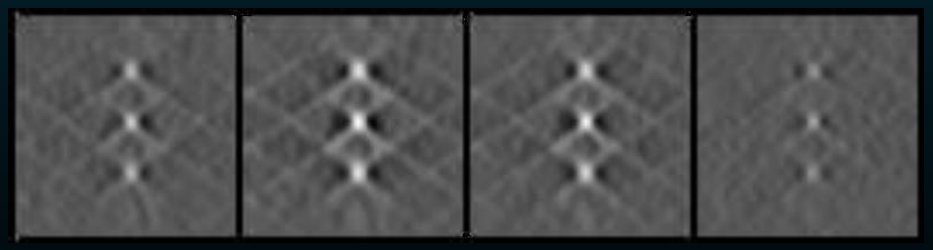
Reconstruction with Phantoms

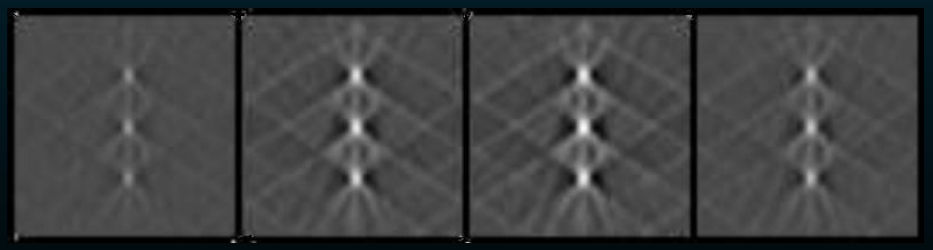




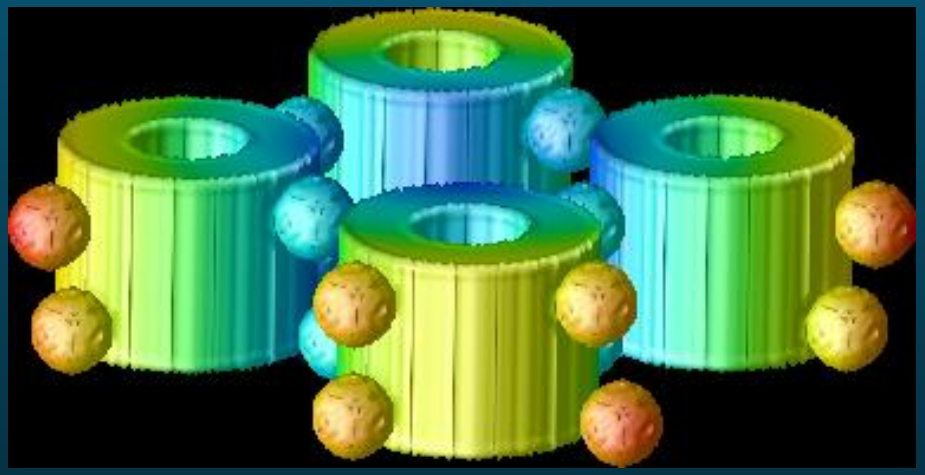


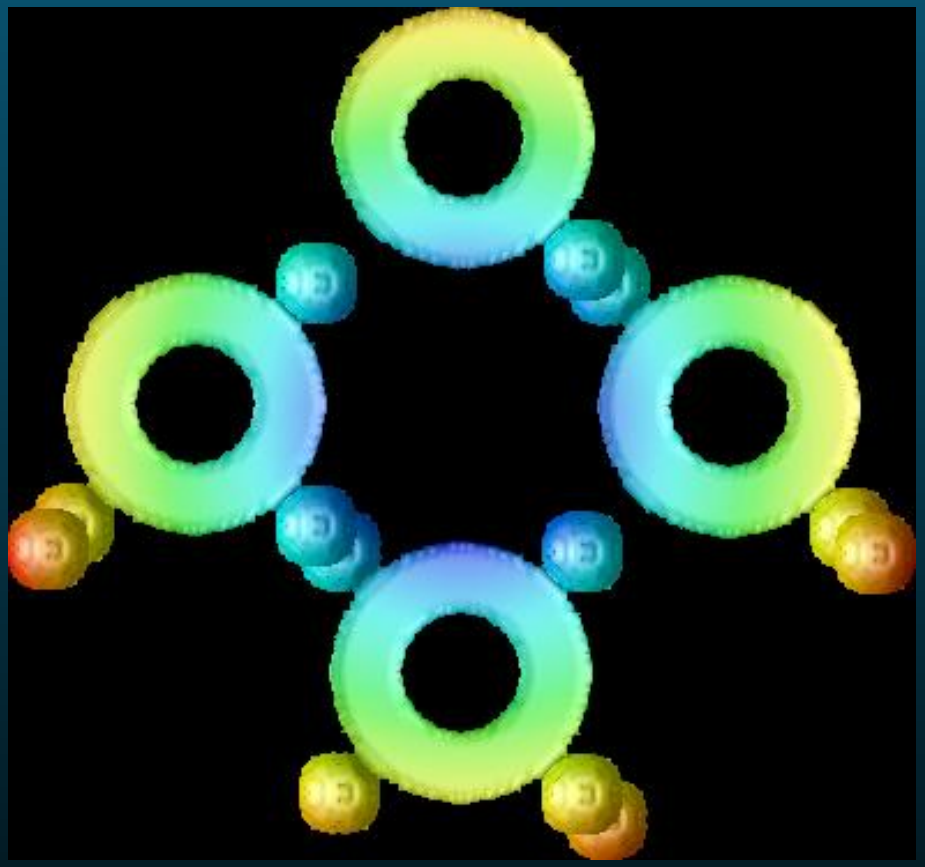






Phantoms and FOMs





Phantom: 2^{16} different configurations.

FOM Task: Detect the particular member by computing the average density in the 16 spheric zones in which the spheres can be localizated.

Acknowledgments

- Algorithms and software
  - ⋆ Gabor T. Herman (Temple U, Phil.)
  - ⋆ Robert Lewitt (UPENN)
  - ⋆ Samuel Matej (UPENN)
  - ⋆ Jose Maria carazo (CNB-Madrid)
  - ⋆ Carlos Oscar Sorzano (CNB-Madrid)
  - ⋆ Tomas Valbueno (CNB-Madrid)
- Standard Crystal Reconstruction
  - ⋆ Jose Jesus Fernandez (U. of Almeria)
- PHI 29 data
  - ⋆ Andreas Engel (Muller Inst., Basel)
  - ⋆ Jose Maria Valpuesta (CNB, Madrid)
- Polyhead Data
  - ⋆ Heinz Gross (Zurich)
  - ⋆ Eva Dimmeler (Zurich)
- Adenovirus Data
  - ⋆ C. SanMartin (Wistar I., Phil.)