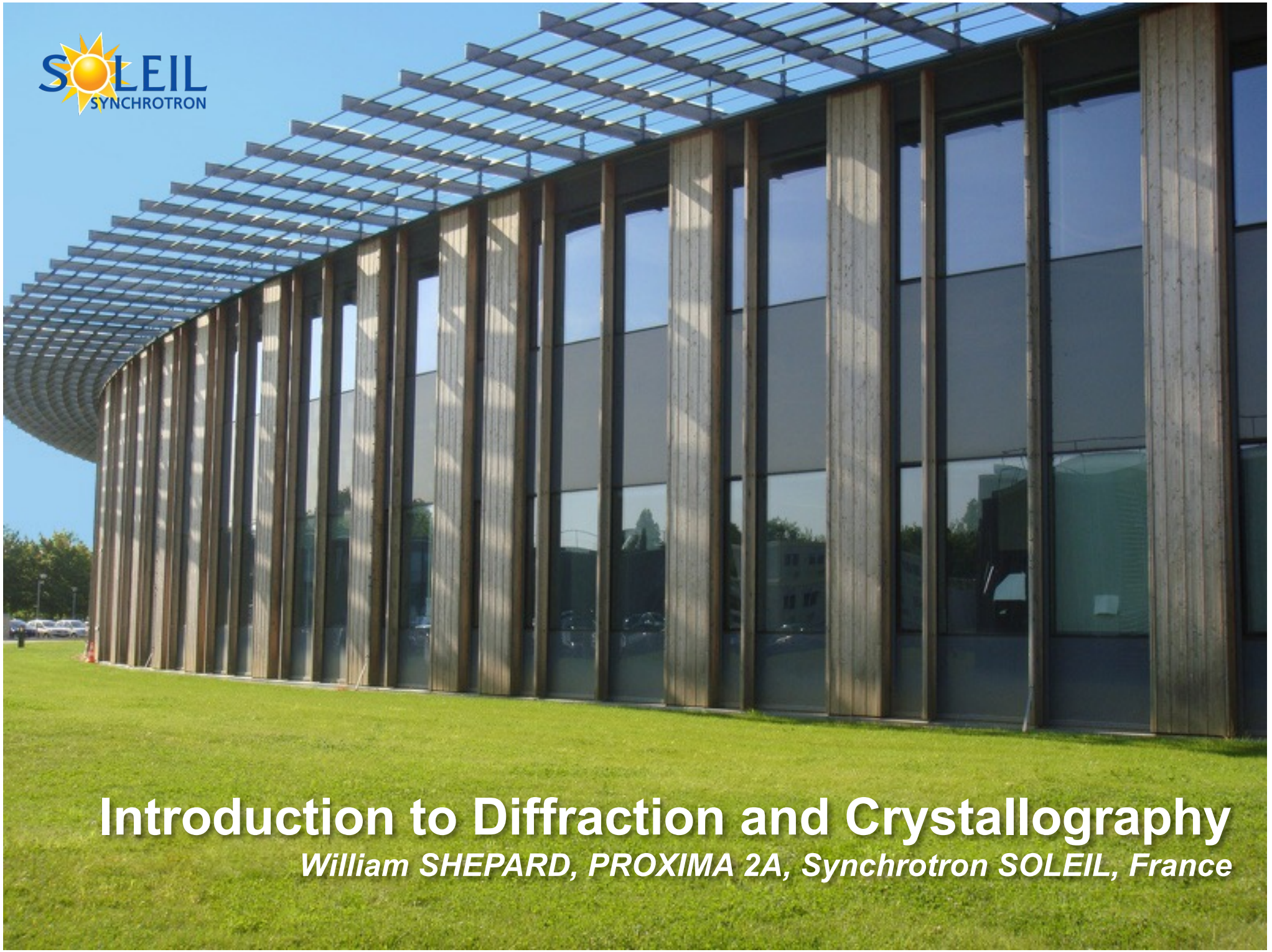


Introduction to Diffraction and Crystallography

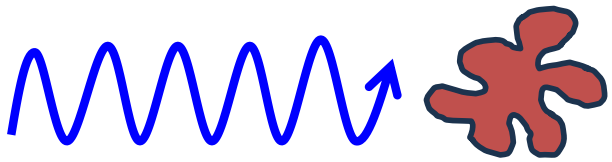
William SHEPARD, PROXIMA 2A, Synchrotron SOLEIL, France

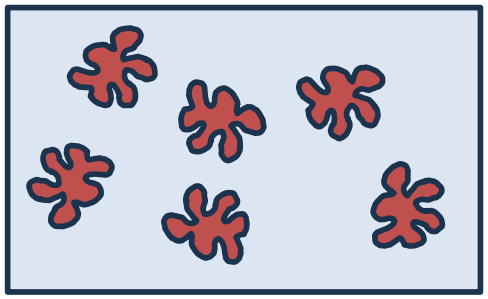
What is the Question?

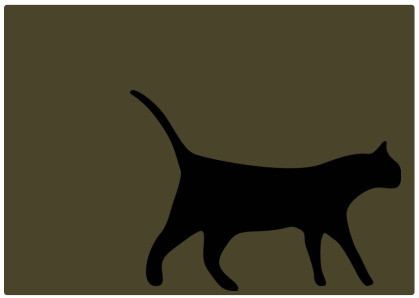
- I am studying an atom, molecule or particle…
  - Metal complex, organic, protein, DNA, RNA, complex, etc…
    - What is its shape? Domains? Oligomer?
    - Which parts are important for its function?
    - Where or what is its active site?
    - Which residues are critical for binding substrates?
    - What are the factors of specificity?
    - What are the structural changes between different states?
    - What are the dynamics?
    - etc…

- **What is its 3D structure at atomic resolution?**

Problems, problems, problems…



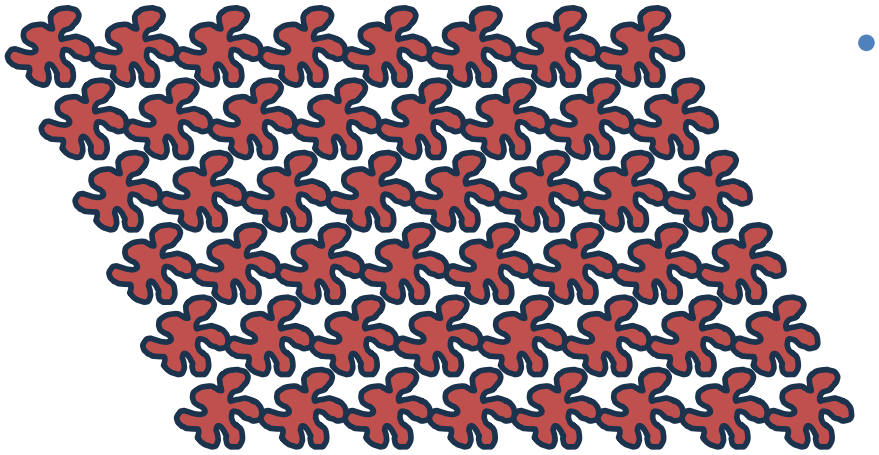




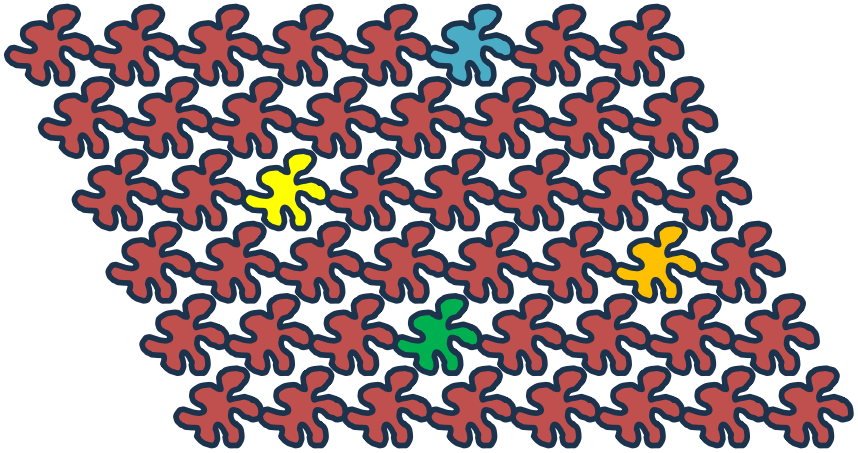
Where is the cat?

- Some hurdles
  - Particle size
    - 10 – 100 Å
    - Probing photon wavelength < particle size
      - λ < 10 Å
  - Orientation
    - Random in solution, amorphous solid & gas
  - Fragility
    - C – C bond 5.76e-19 J
    - 12.4 keV photon 1.99e-15 J
    - 3454 times more !!!
  - Weak signal
    - to see a particle, the probe must interact with it…

One solution…



2-dimensional crystal

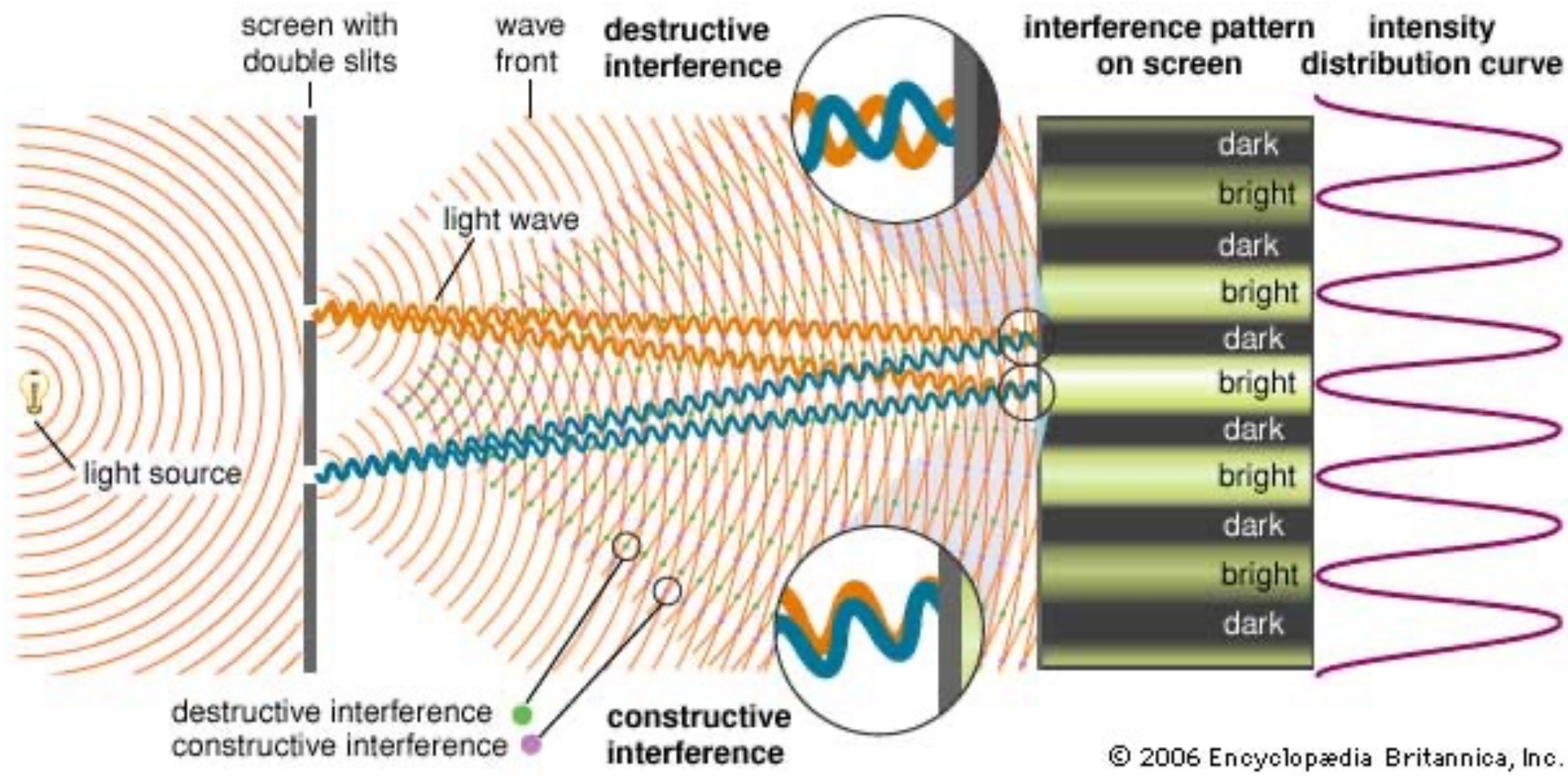


2-dimensional crystal with damaged particles

- Crystallise it!
  - Orient the particles the same way
    - Crystals have top, bottom, front, back, left & right
  - Repetitive units
    - Combines weak signals
    - Multiple copies resists radiation damage
  - Employ phenomena of **diffraction**
    - *Interaction without destruction!*

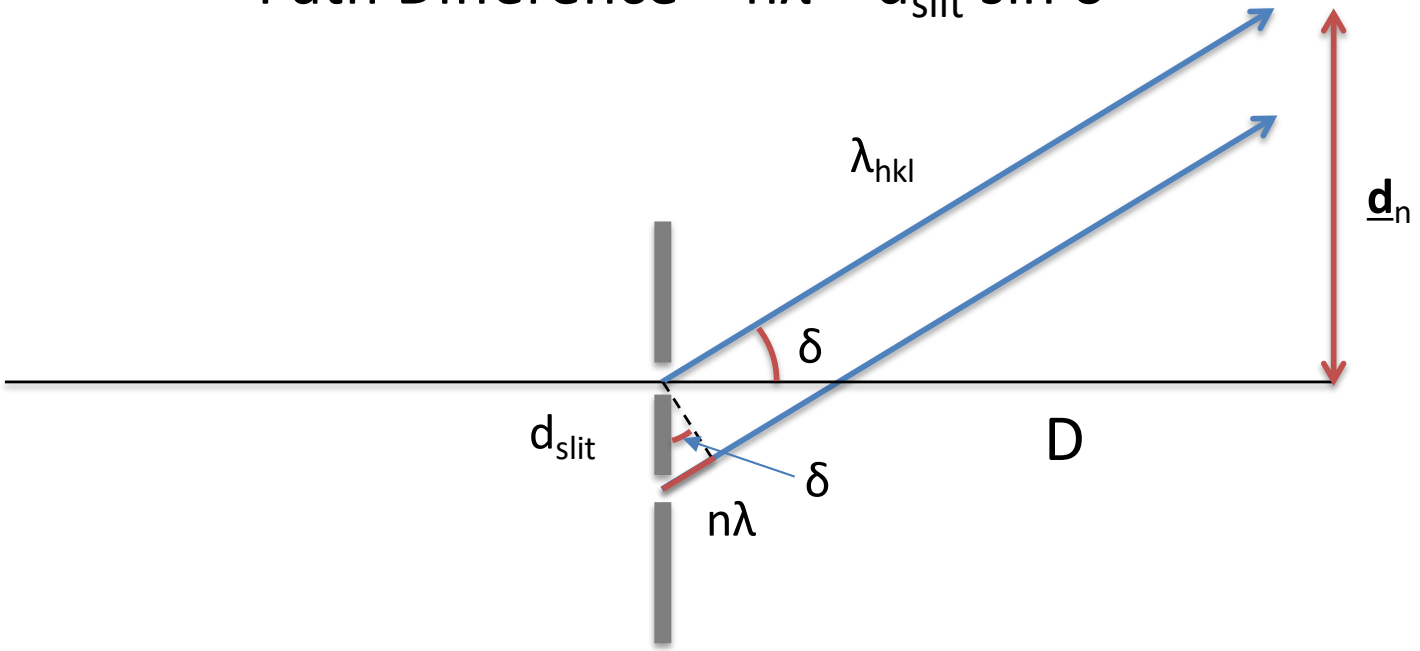
Diffraction & Reciprocal Space

Young's Experiment (1803) Revisited



Diffraction Geometry from a Slit

Path Difference = $n\lambda = d_{slit} \sin \delta$

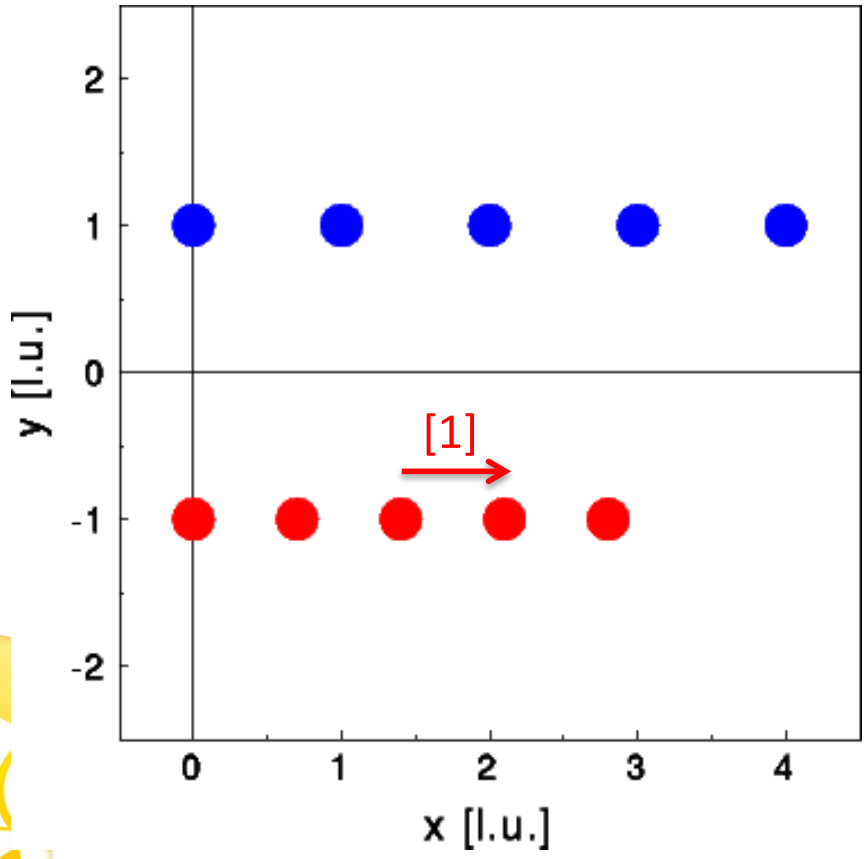


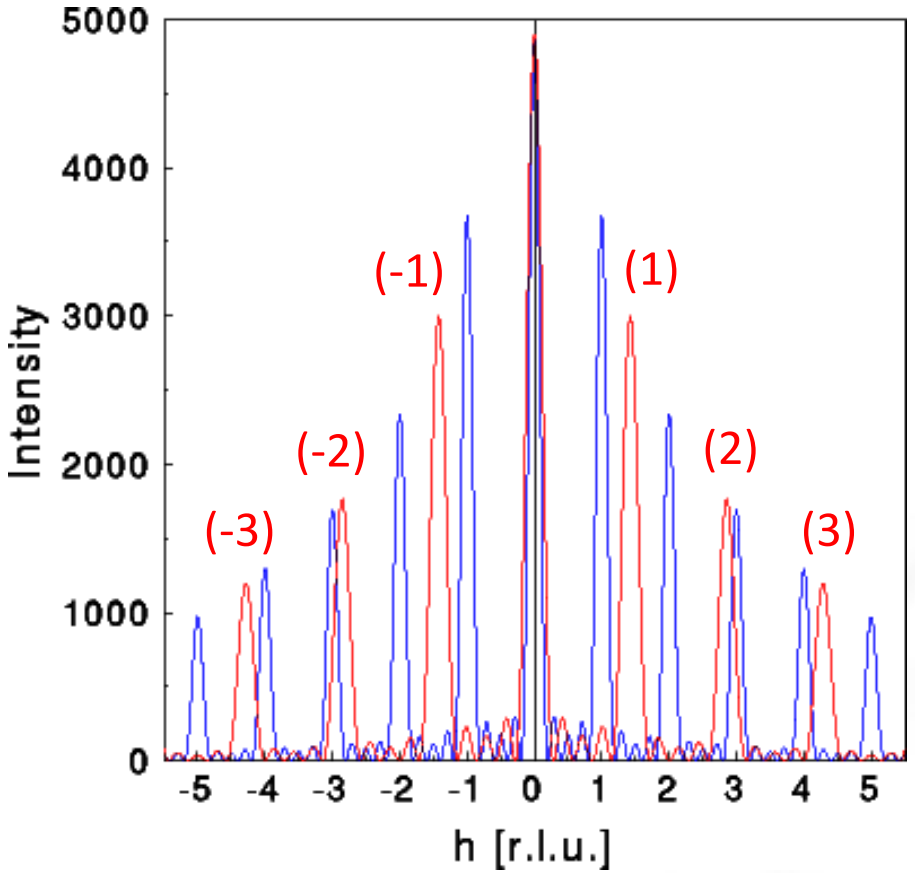
$| \mathbf{d}_n | = D \tan \delta$

$| \mathbf{d}_n | = D \ (n\lambda / d_{slit}) / \cos \delta$

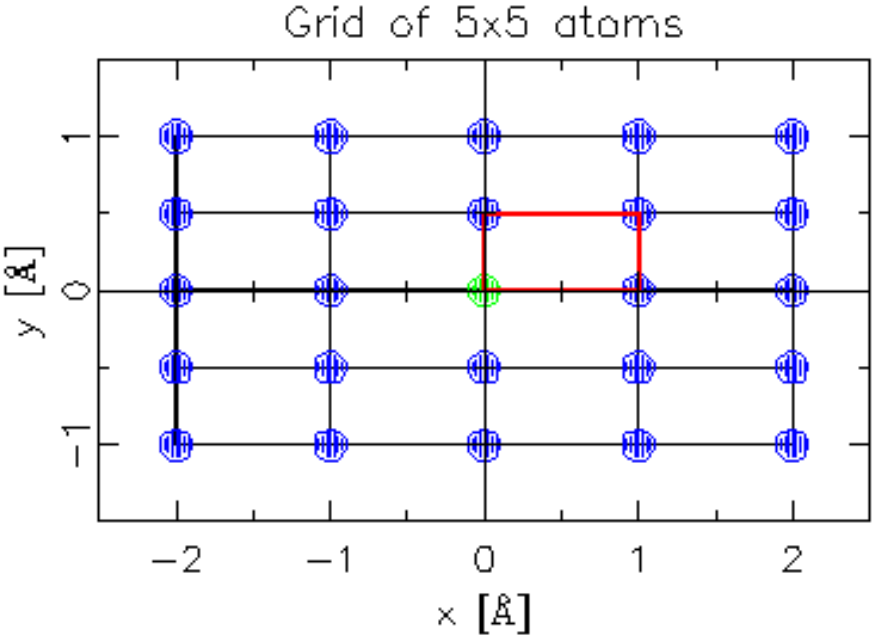
1D Diffraction (Slits) & Reciprocity

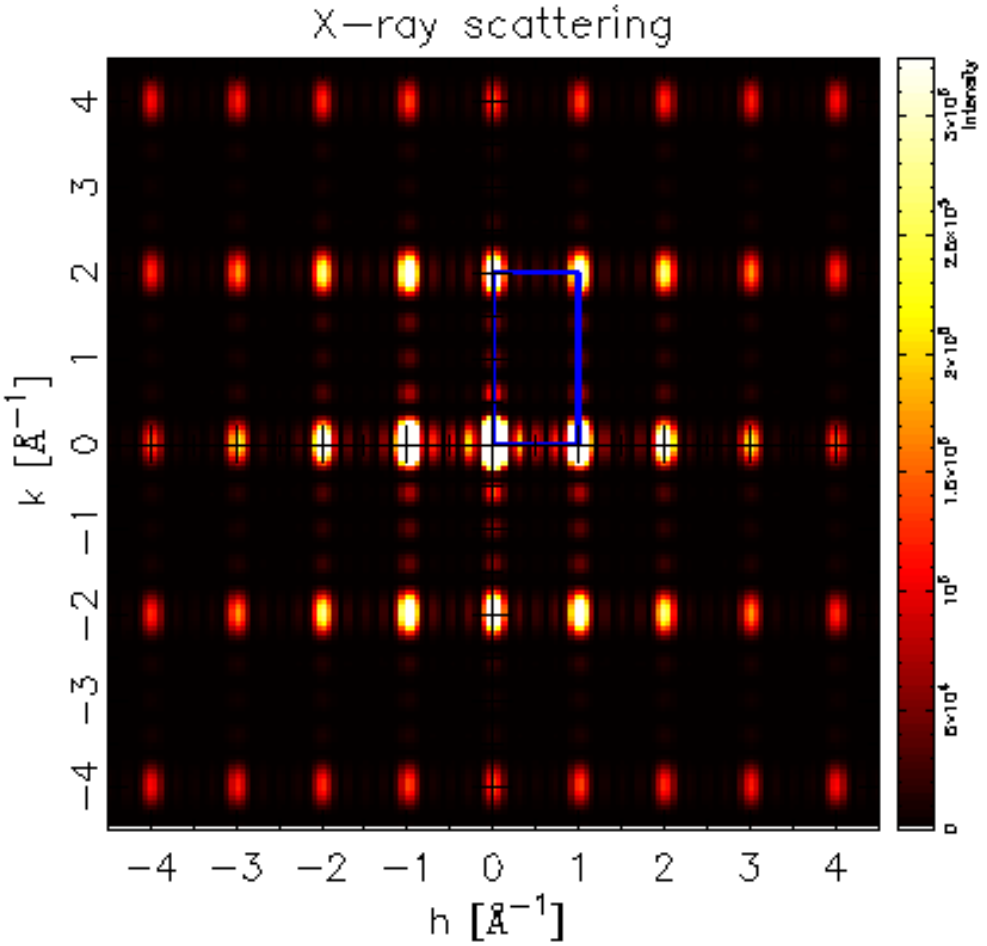
- Consider the multiple atom experiment
- A smaller spacing (x) between slits or atoms in the lattice yields….
  - a larger separation between diffraction spots (1/x)
    - Each diffraction spot has its own Miller index, (h)
    - Each Miller index represents a vector in real space, [h]



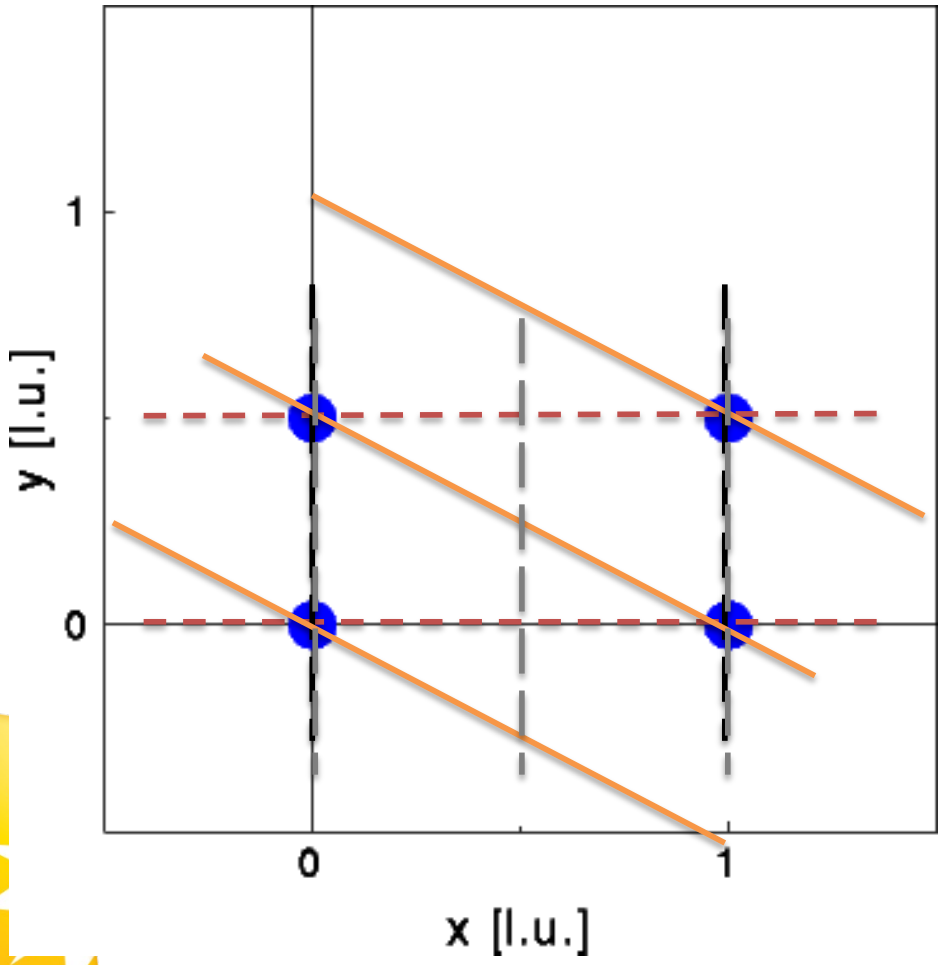


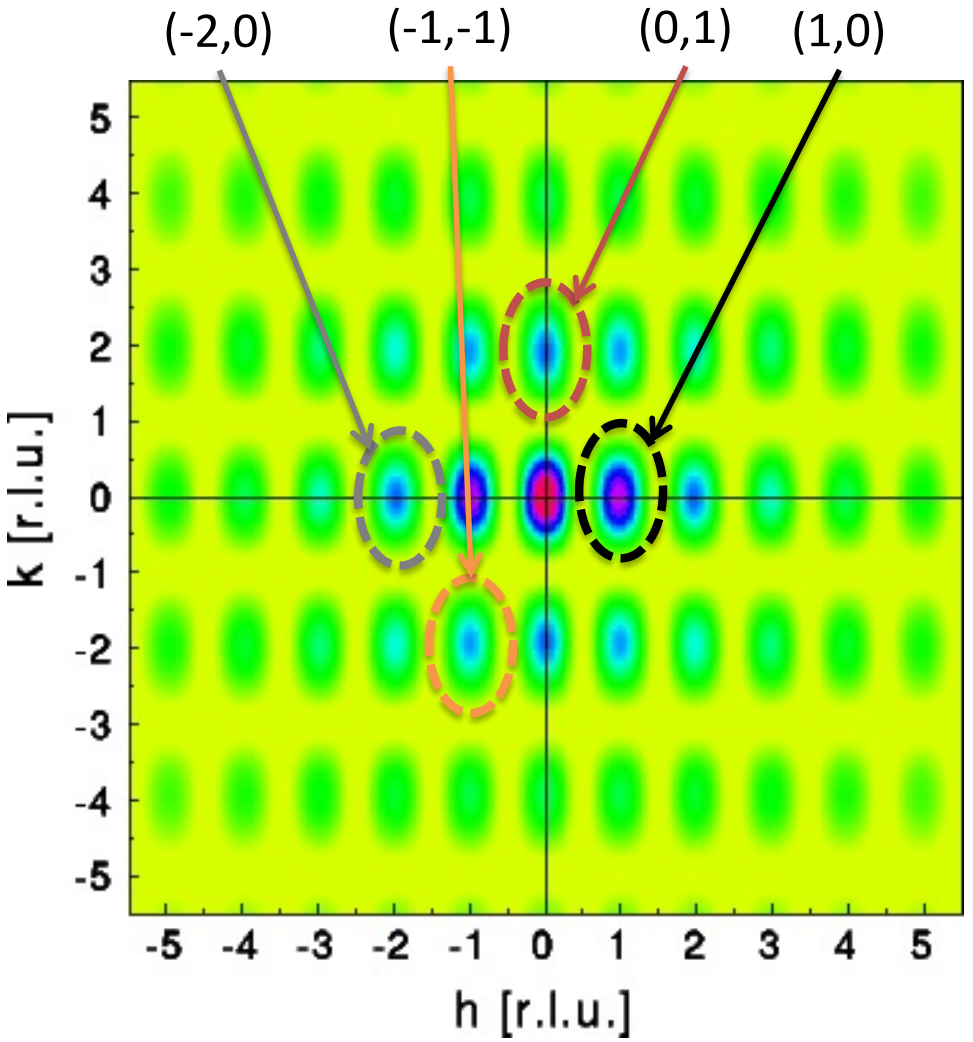
Diffraction from a Grid (2D) & its Reciprocal Space



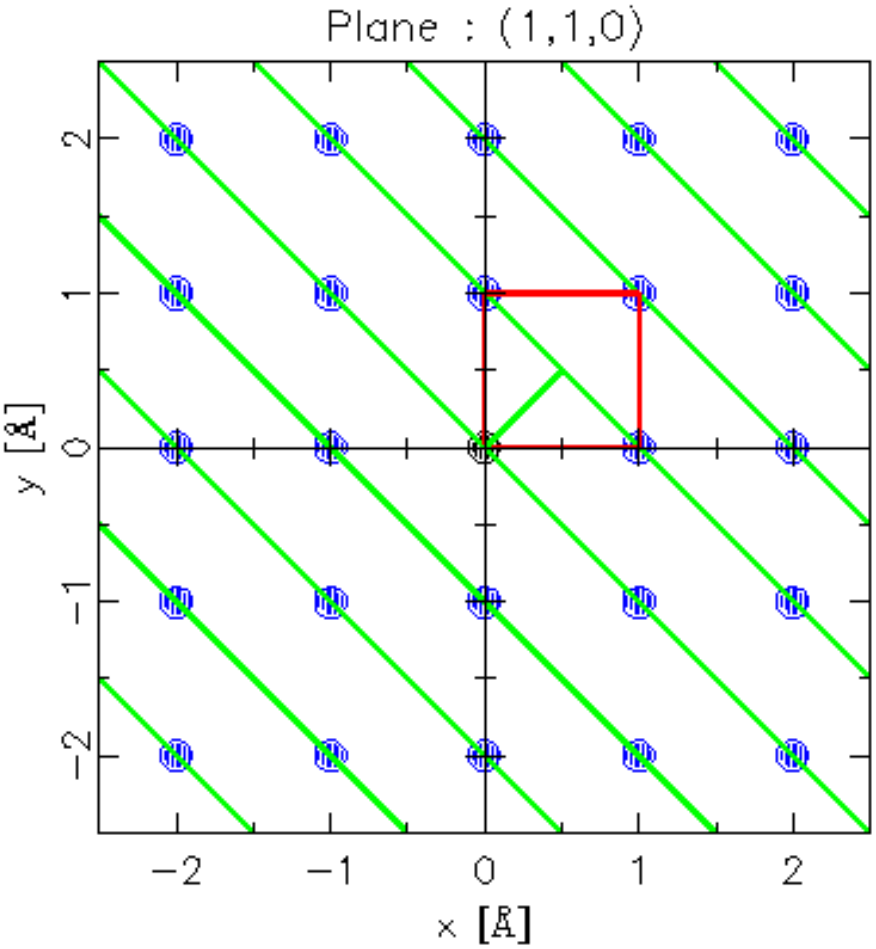


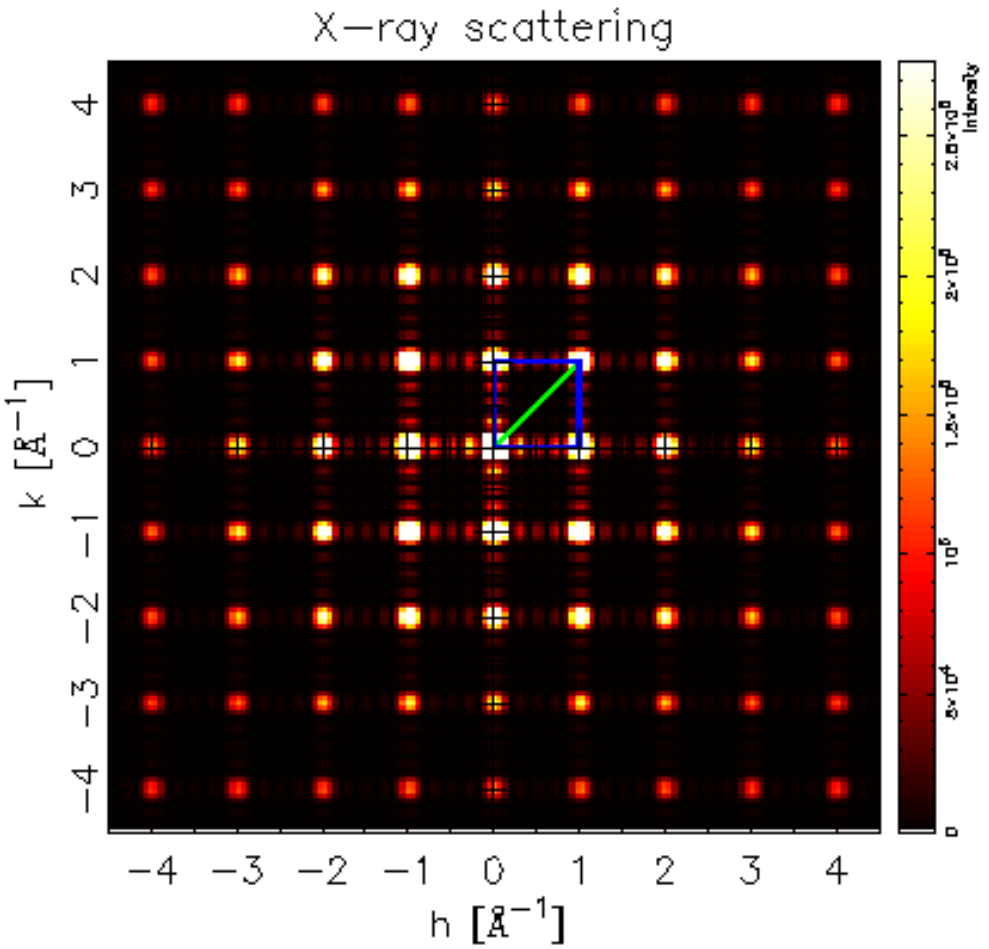
2D Reciprocal Space Miller indices & Lattice lines



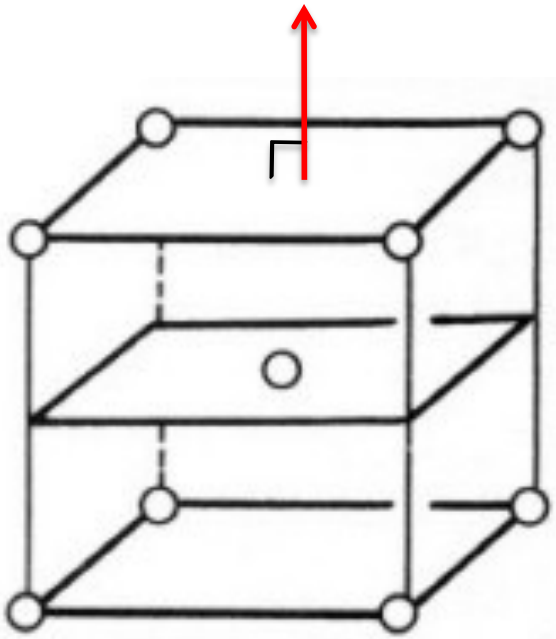


Lattice Lines, Vectors & Indices

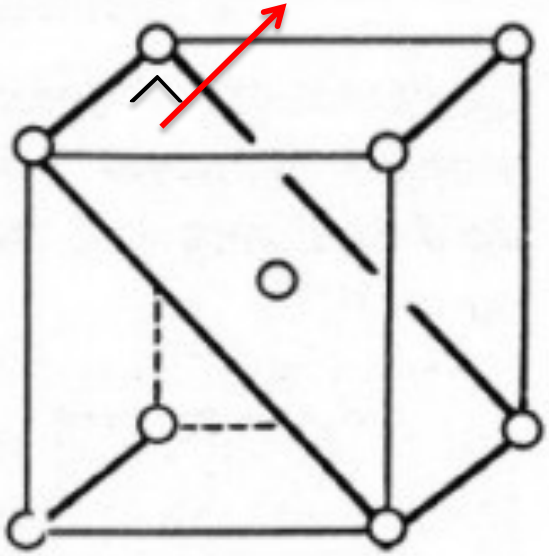




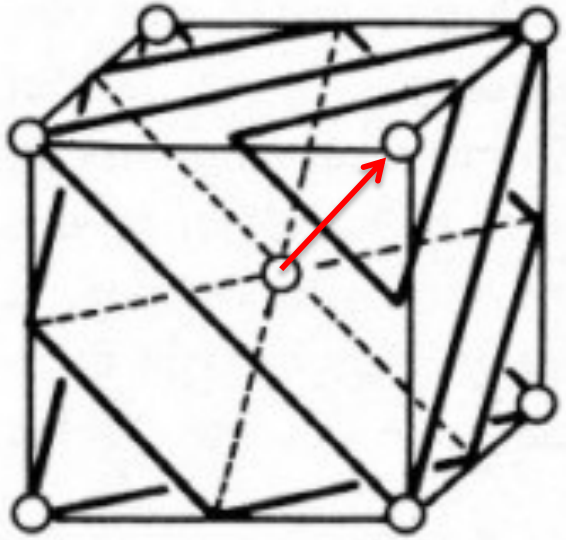
Crystals: Lattice Planes in 3D Space



(100)



(110)

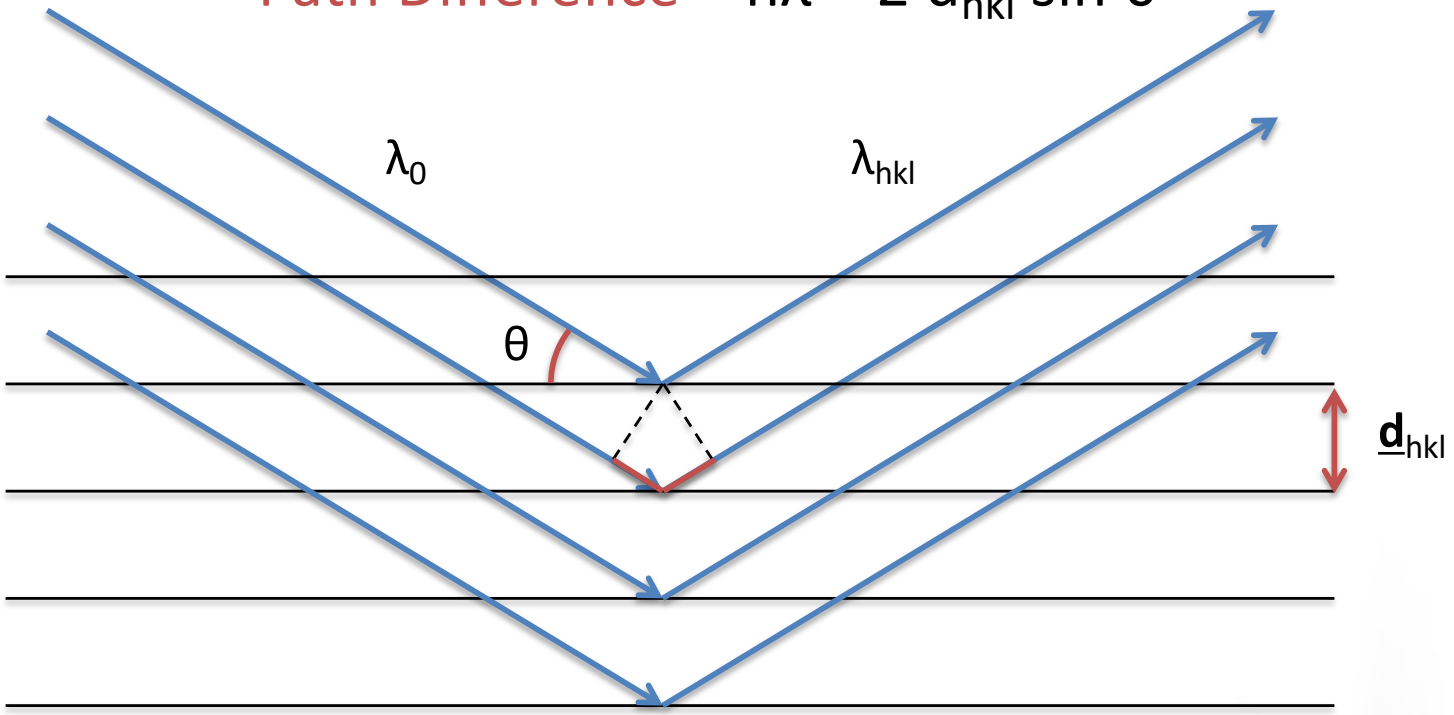


(111)

- Crystals contain sets of lattice planes denoted by Miller indices (hkl)
  - The orientation of a set of lattice planes can be represented by a vector
  - This vector is perpendicular (normal) to the set of planes

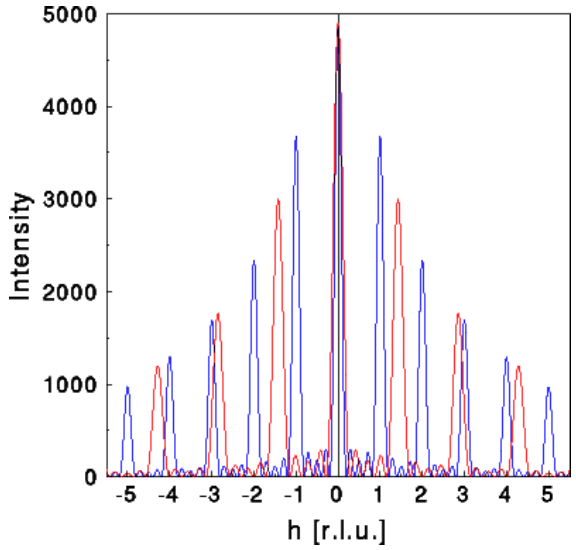
Bragg's Law

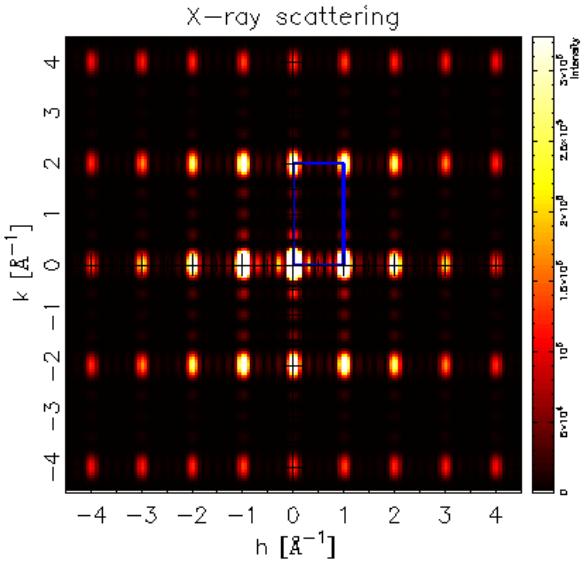
Path Difference = $n\lambda = 2\, d_{hkl} \sin\theta$



- X-ray energy conserved
  - $\lambda_0 = \lambda_{hkl}$
- When the crystal is rotated
  - A different set of planes diffract

1D, 2D & 3D Diffraction



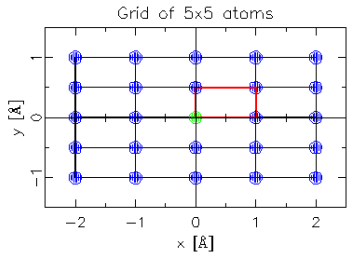


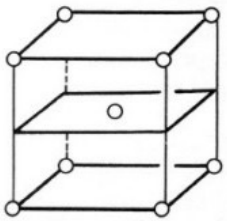
?

1D
Lattice points
[h]

2D
Lattice lines
[hk]

3D
Lattice planes
[hkl]





Unit cell parameters

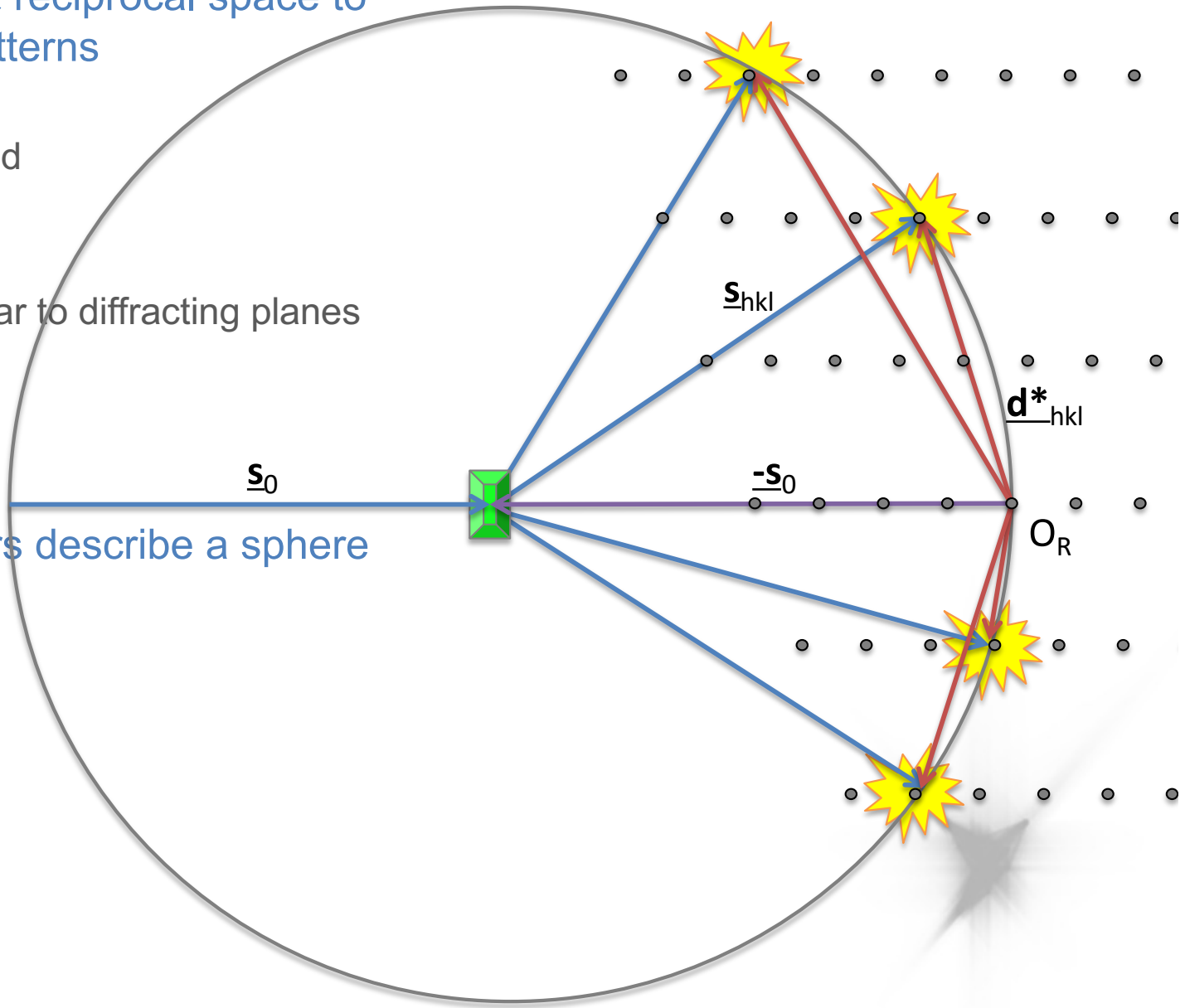
a

a, b, α

a, b, c, α, β, γ

Ewald Construction Derived

- Uses wave vectors & reciprocal space to explain diffraction patterns
  - $|\underline{s}_0| = 1 / \lambda$
  - Energy is conserved
    - $|\underline{s}_0| = |\underline{s}_{hkl}|$
  - $\underline{s}_{hkl} - \underline{s}_0 = \underline{d^*}_{hkl}$
  - $\underline{d^*}_{hkl}$ is perpendicular to diffracting planes d_{hkl}
  - $|\underline{d^*}_{hkl}| = 1/d_{hkl}$
- The scattering vectors describe a sphere
  - Radius = 1 / λ
- Reciprocal space
  - Lattice of points
  - a*, b*, c*



Stills: No Rotation

Figures courtesy of Z. Dauter (1999)

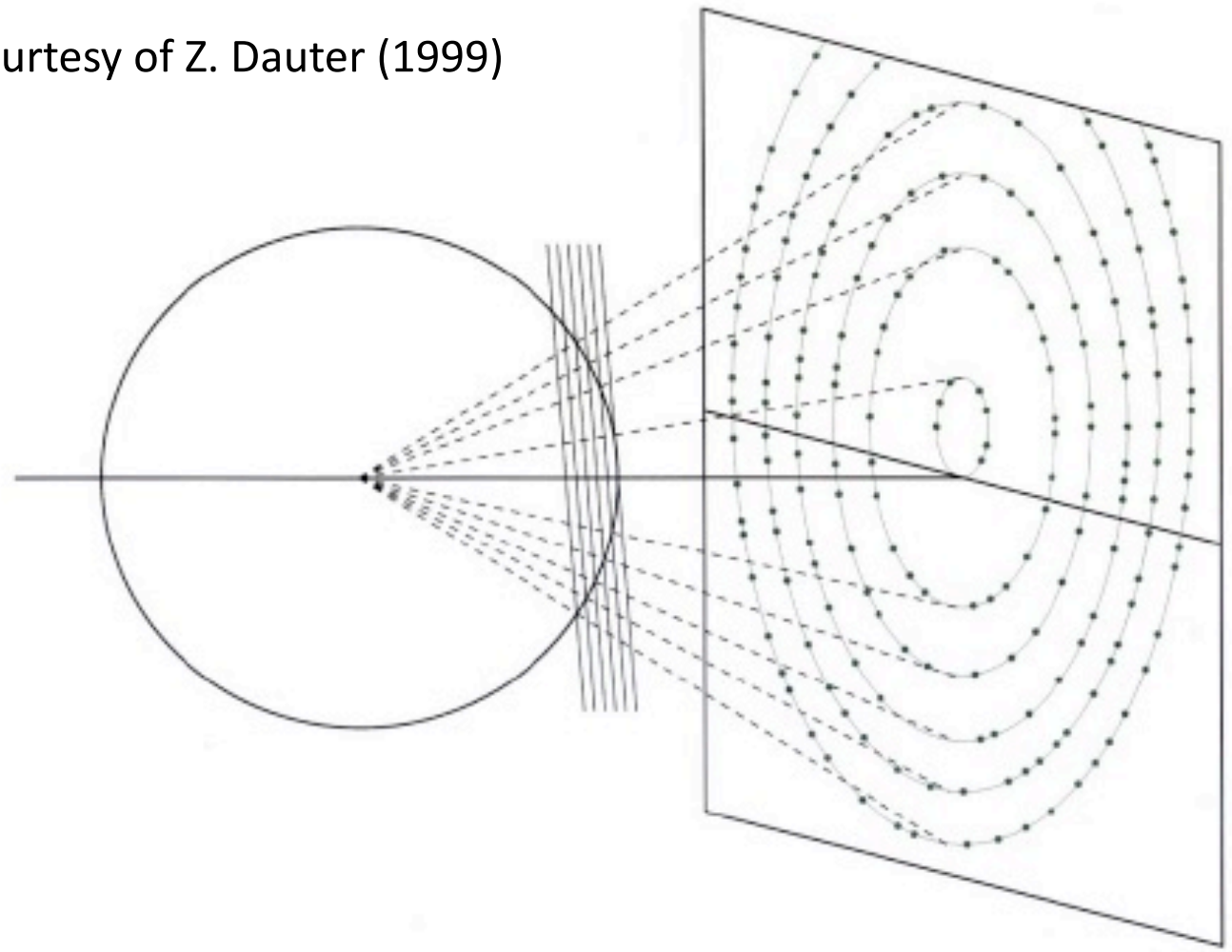


Figure 3
A still exposure with a stationary crystal contains only a small number of reflections arranged in a set of narrow ellipses.

Rotation Method

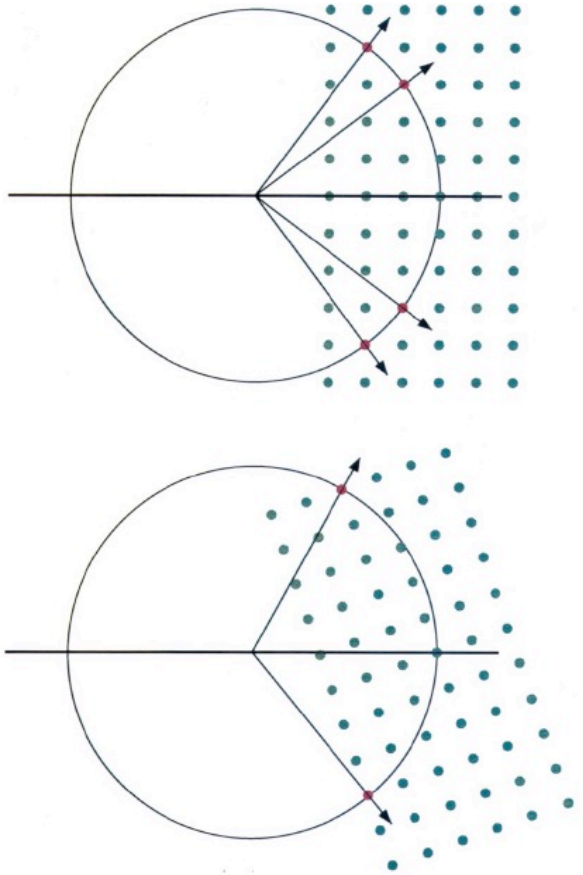


Figure 2
To bring more reflection into diffraction, the crystal represented by the reciprocal lattice has to rotate.

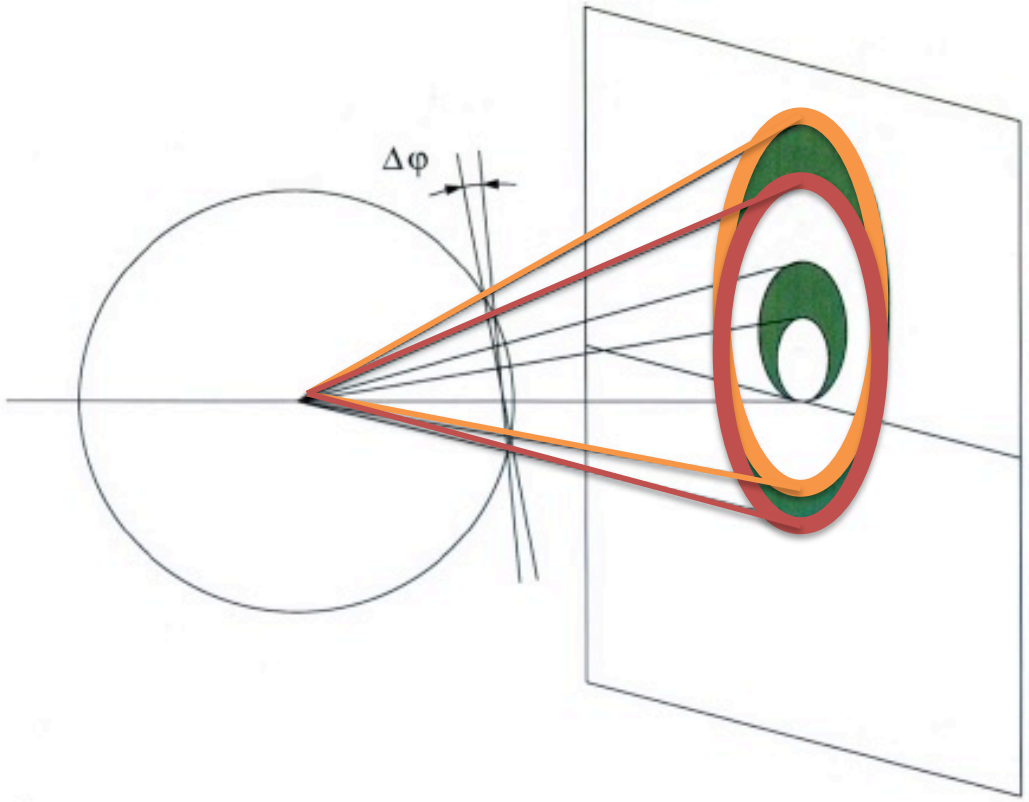
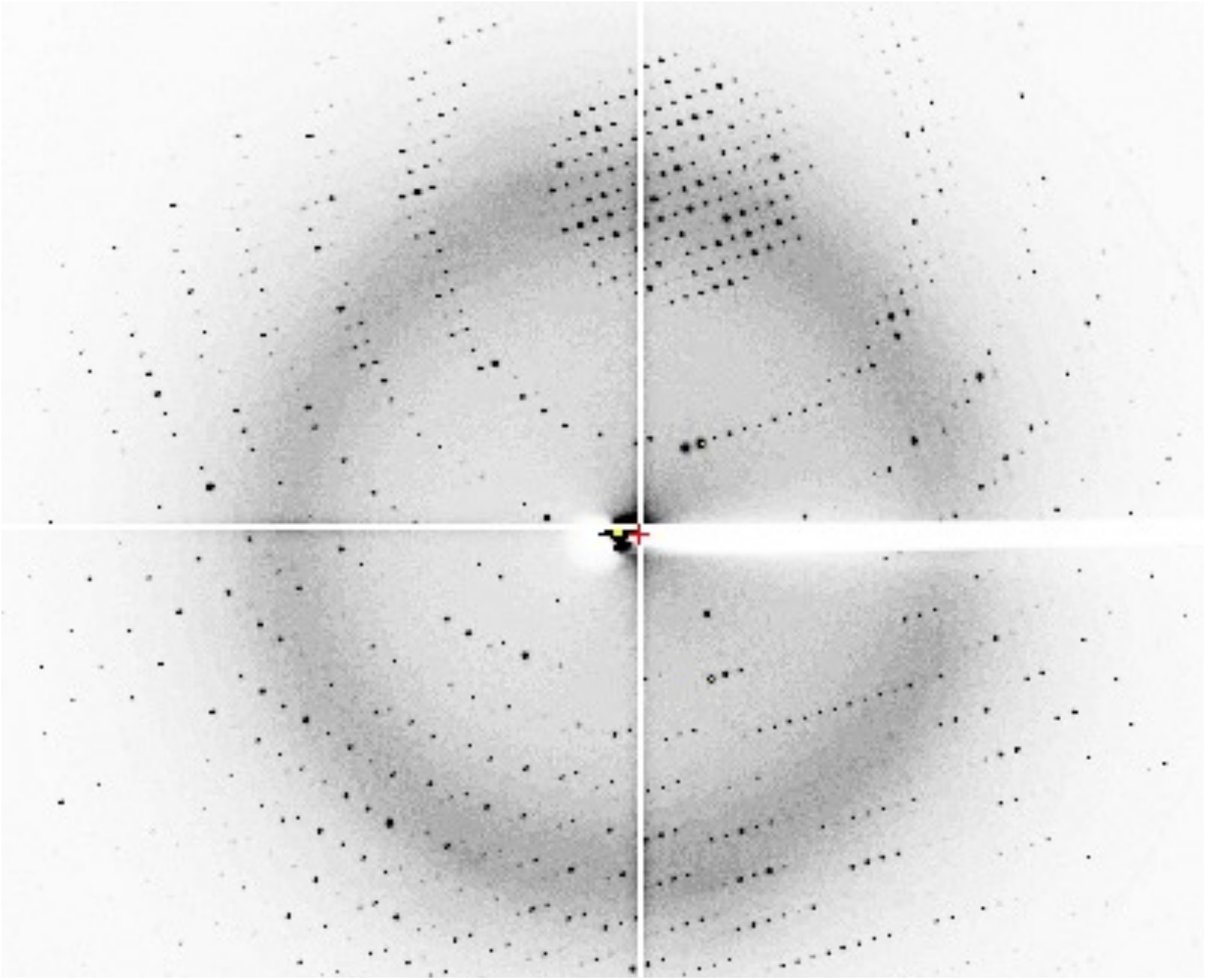


Figure 4
When the crystal is rotated, reflections from the same plane in the reciprocal lattice form a lune, limited by two ellipses corresponding to the start and end positions.

An X-ray diffraction pattern from a crystal…

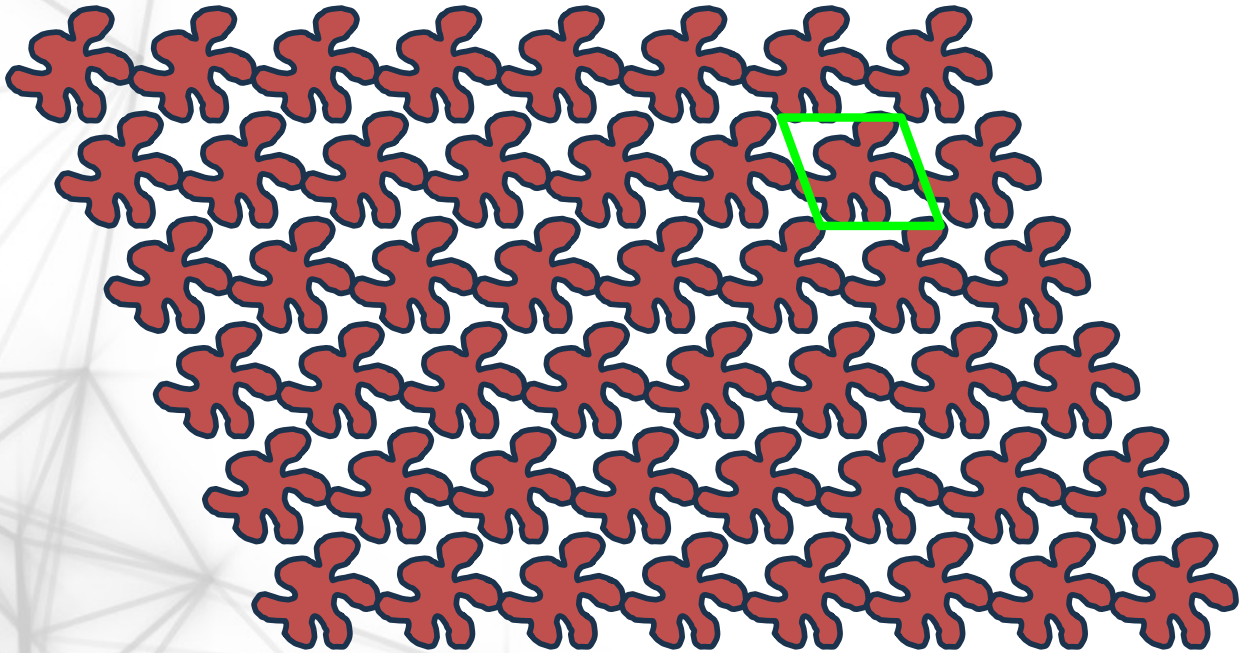


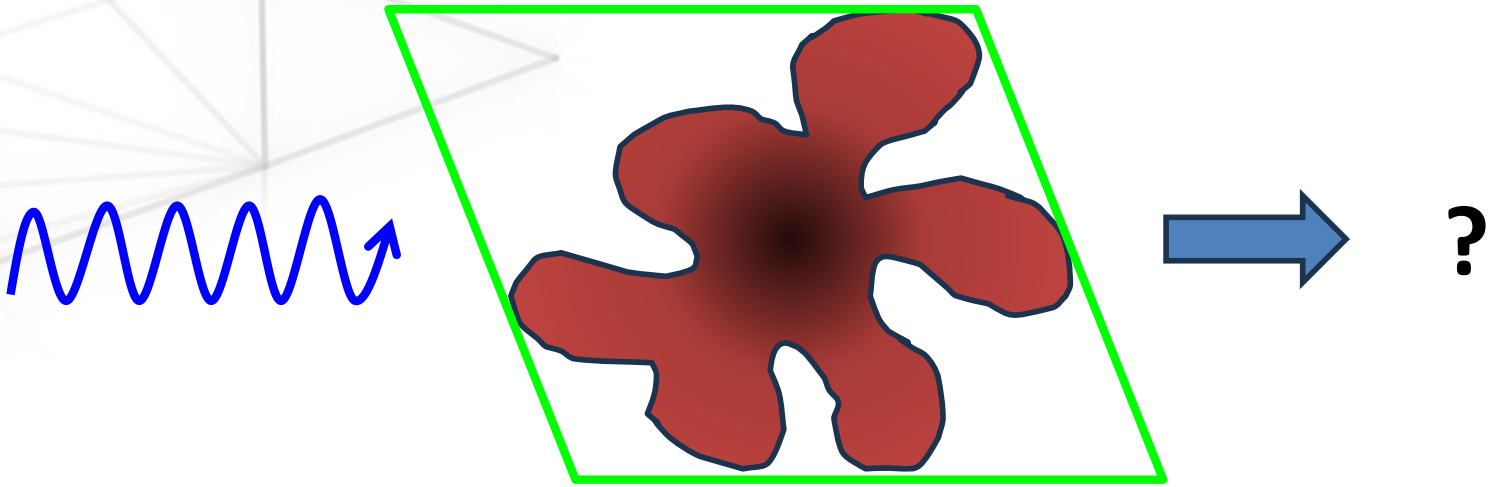
Summary of Reciprocal Space

- Reciprocal space is NOT discrete, but rather continuous
  - but Bragg **reflections are discrete points in Reciprocal space (h,k,l)**
- For larger / smaller unit cell lengths (a, b, c)
  - Diffraction spots are closer together / further apart (a*, b*, c*)
- The crystal lies at the center of the Ewald sphere
  - Scattering vectors have the same length because energy is conserved
  - The ensemble of all of the possible scattering vectors $\mathbf{s}_{hkl}$ describe a sphere
- The radius of the Ewald sphere is 1/λ
  - Long X-ray wavelengths produce a small Ewald sphere
  - Short X-ray wavelengths produce a large Ewald sphere
- The origin of Reciprocal space:
  - lies at the intersection of $\mathbf{s}_0$ and the Ewald sphere
  - All of the diffraction vectors **d*** originate from the origin of reciprocal space
- Not all points in Reciprocal Space will diffract
  - Only those points that lie on the Ewald sphere will diffract
- The origin of Reciprocal Space is fixed
  - but **Reciprocal Space rotates exactly as the crystal rotates**
- **a***, **b*** & **c*** are not necessarily orthogonal to **a**, **b**, **c**

From Scattering to Structure Factors

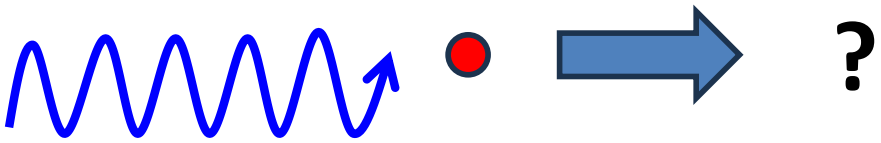
Diffraction from the Unit Cell Contents?





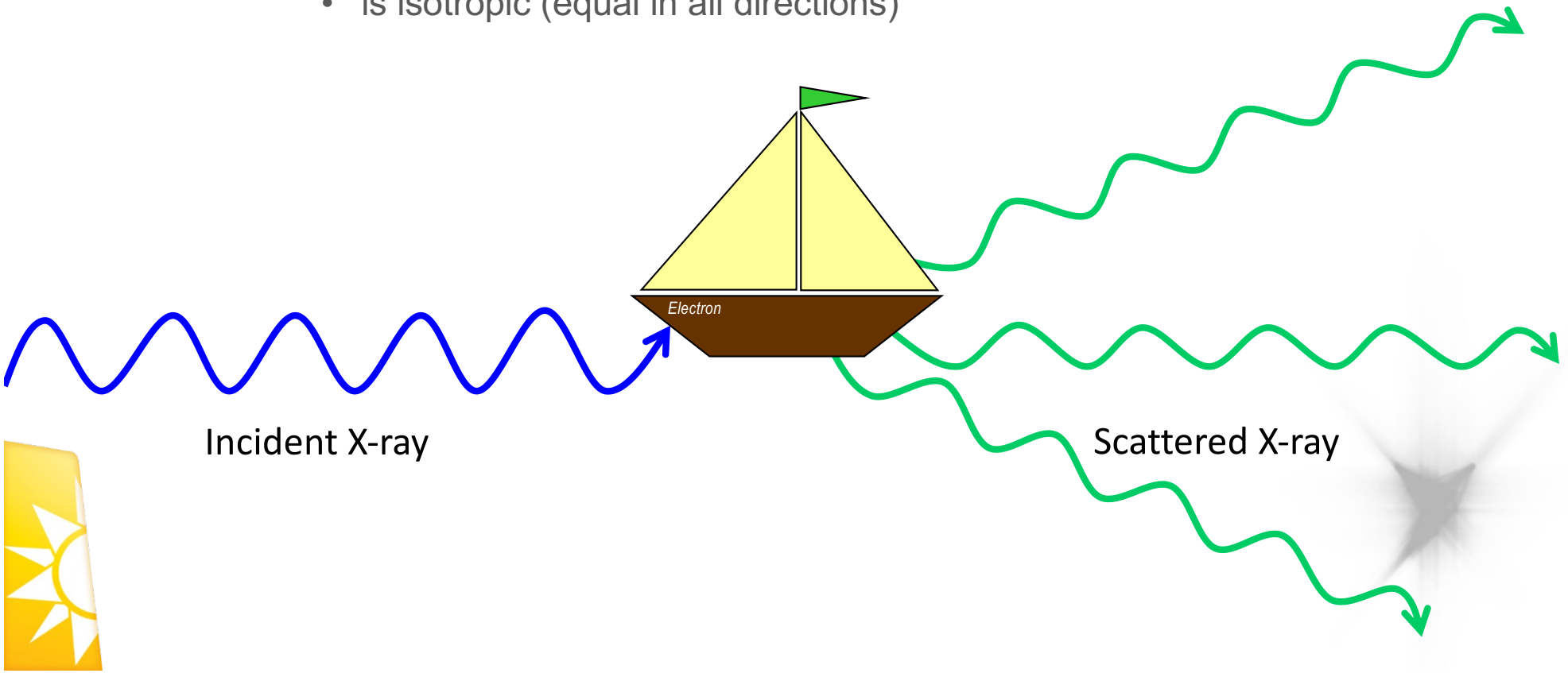
Types of Scattering

- X-rays
  - Electrons (density)
    - Elastic (diffraction)
    - Inelastic (absorption)
- Neutrons
  - nuclei (H/D)
- Electrons
  - Electrostatic charges or fields

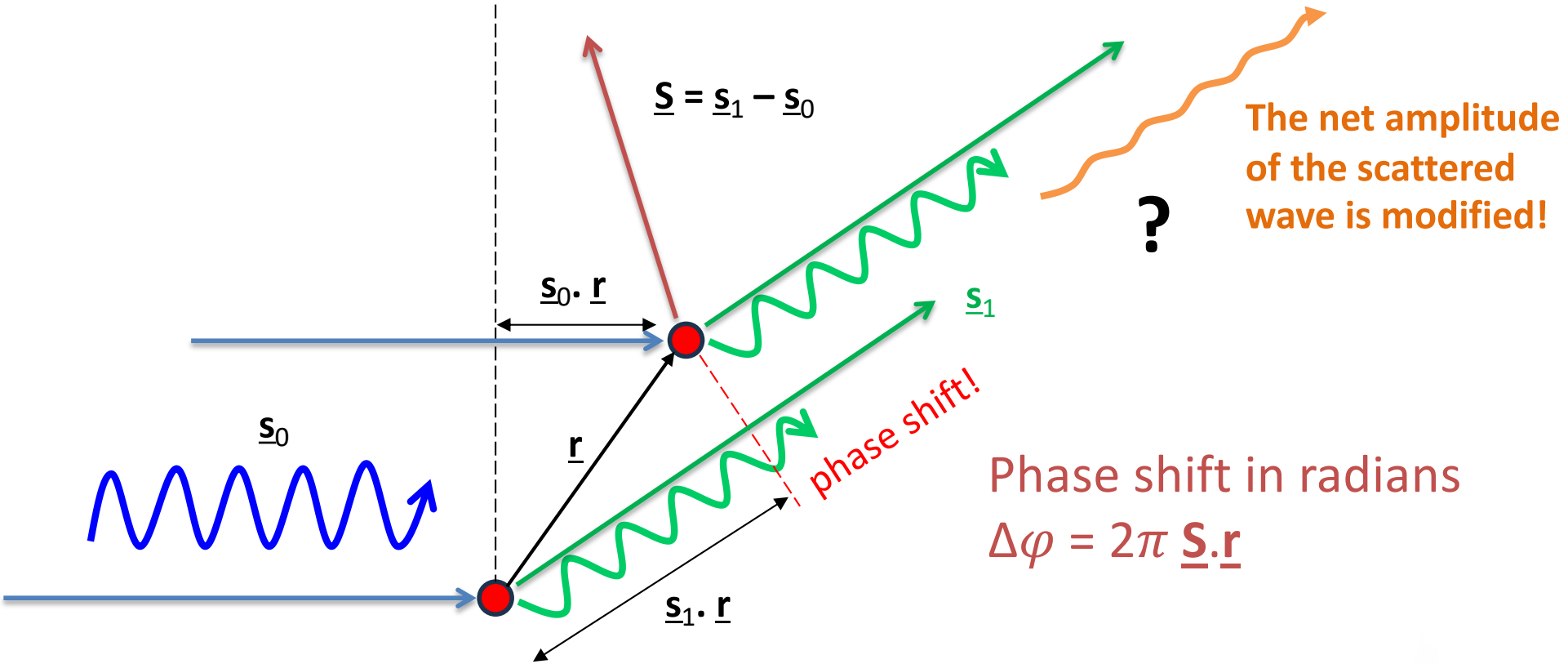


X-ray Scattering Analogy

- X-ray scattering is the interaction between the electric vector, **E**, of the incident X-ray and an electron.
  - The scattered X-ray:
    - has the same wavelength (elastic scattering)
    - is isotropic (equal in all directions)



Scattering from 2 Electrons



Phase shift in radians

$\Delta\varphi = 2\pi\ \underline{\mathbf{S}}.\underline{\mathbf{r}}$

Path difference

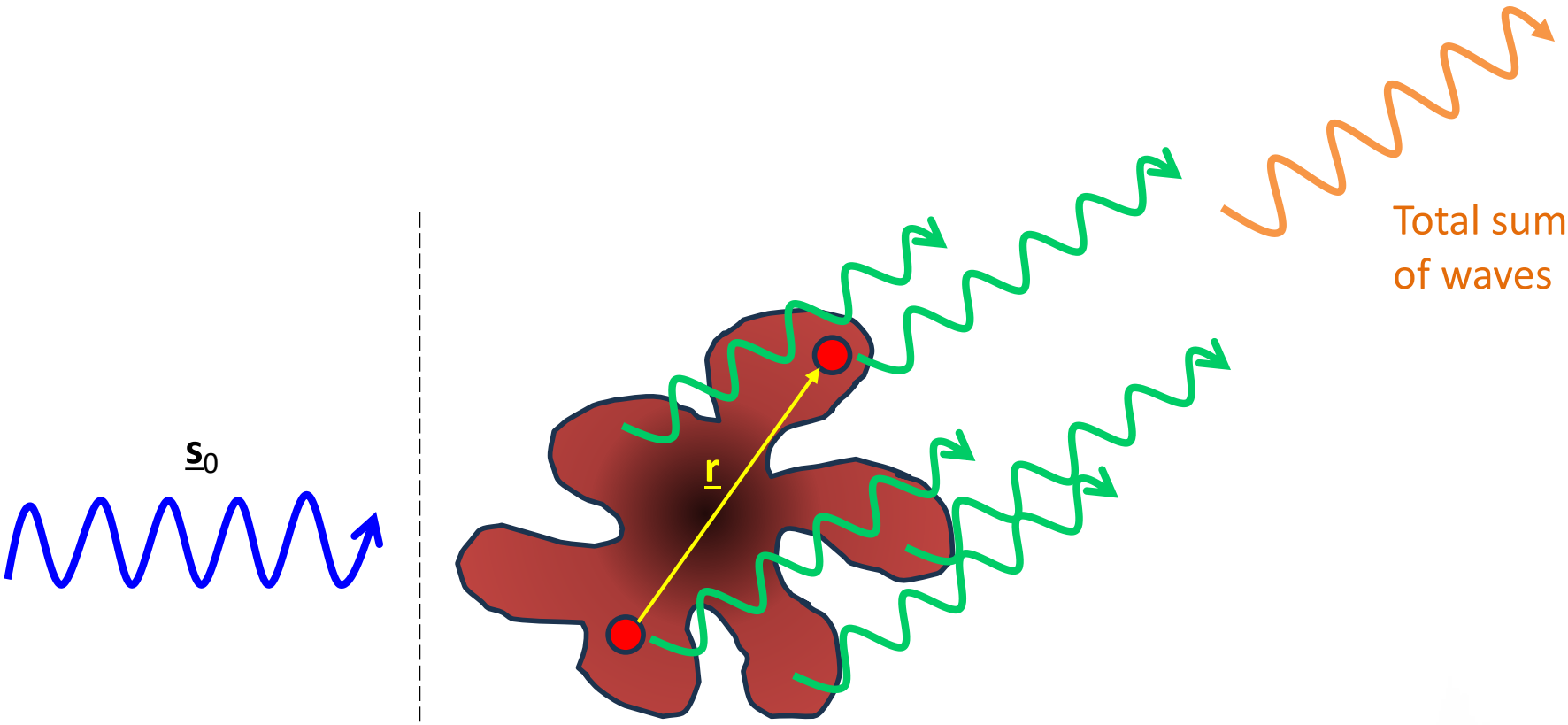
$= (\underline{\mathbf{s}}_1.\ \underline{\mathbf{r}}) - (\underline{\mathbf{s}}_0.\ \underline{\mathbf{r}})$

$= (\underline{\mathbf{s}}_1 - \underline{\mathbf{s}}_0)\ .\ \underline{\mathbf{r}}$

$= \underline{\mathbf{S}}.\underline{\mathbf{r}}$

The *phase* shift is derived from the *positions* of the scatterers!

Scattering from Many Voxels

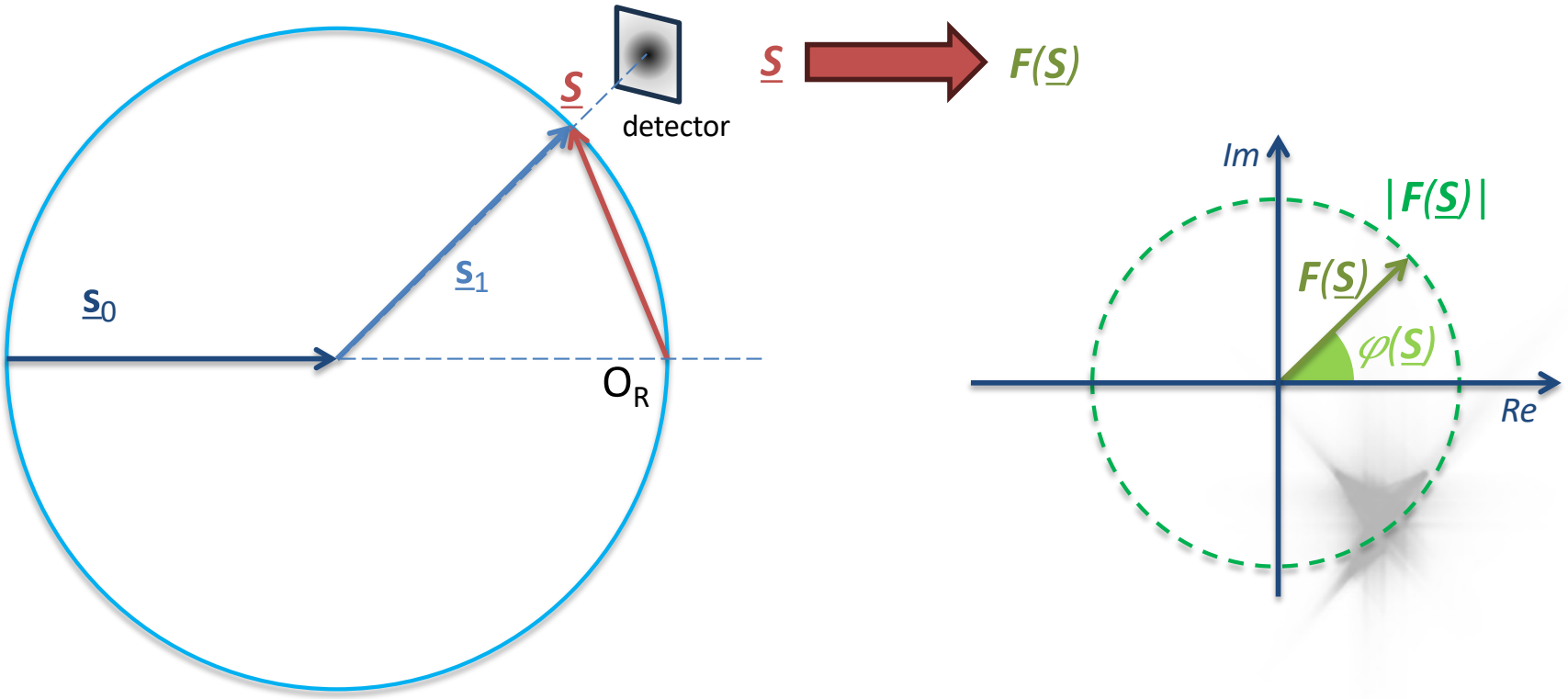


Total phase shifts in radians

$$F(\underline{S}) = \int_{-\infty}^{+\infty} \rho(\underline{r})\, e^{i2\pi \underline{S}.\underline{r}}\, dr$$

Structure Factors

- Scattering is a *wave function* described by a *structure factor*
  - Any given point, $\underline{\boldsymbol{S}}$, in reciprocal space is associated with a *complex number* called a *structure factor*, $\boldsymbol{F(\underline{S})}$.
  - A structure factor, $\boldsymbol{F(\underline{S})}$,
    - is the summation of all scattered waves along $\underline{\boldsymbol{S}}$,
    - has amplitude, $|\boldsymbol{F(\underline{S})}|$, and phase, $\varphi(\underline{\boldsymbol{S}})$.



Structure Factor Equations

For **continuous** reciprocal space (e.g. WAXS) employ ***S*** and ***r***

$$\boldsymbol{F}(\underline{\boldsymbol{S}}) = \int_{-\infty}^{+\infty} \rho(\underline{\boldsymbol{r}})\, e^{2\pi i \underline{\boldsymbol{S}}.\underline{\boldsymbol{r}}}\; dr$$

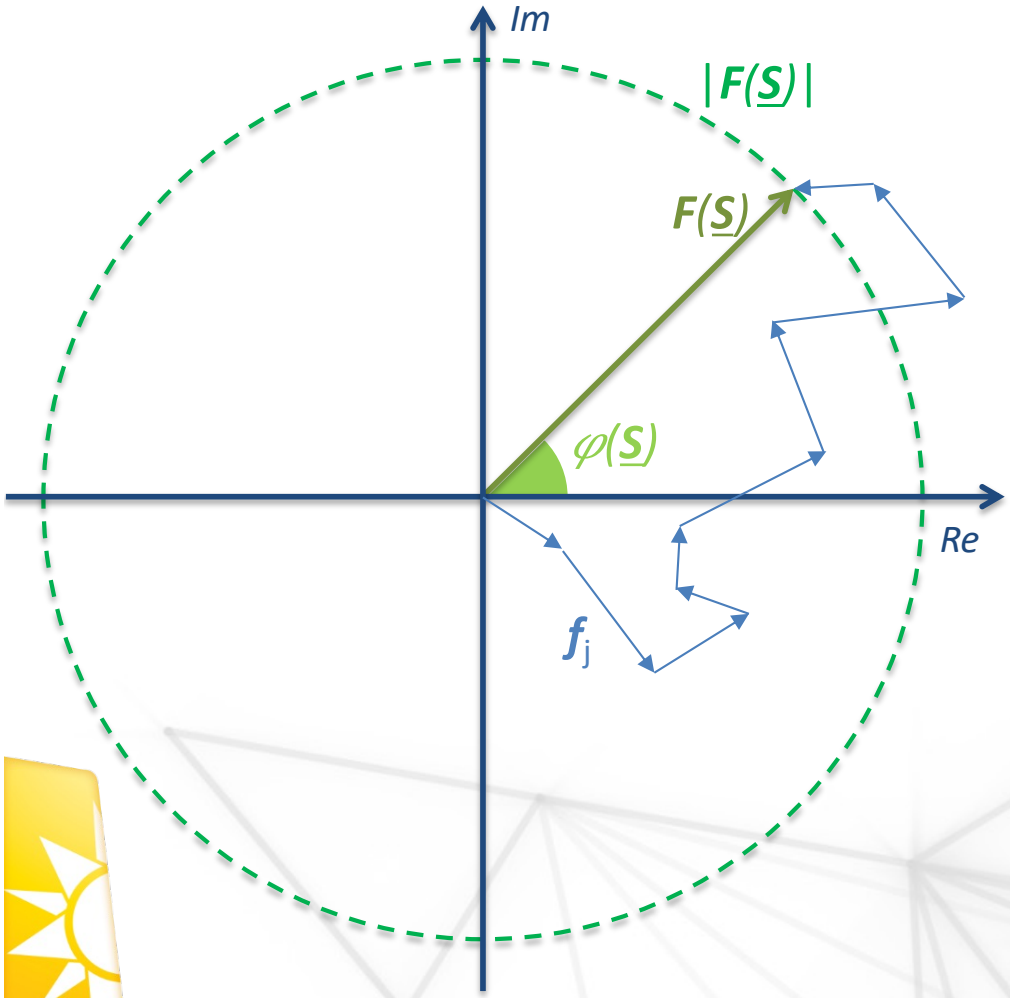
For **discrete** reciprocal space (e.g. MX) employ ***h*** and ***x***

- Miller indices, ***h***
- fractional coordinates, ***x***

$$\boldsymbol{F}(\underline{\boldsymbol{h}}) = \int_{0}^{1} \rho(\underline{\boldsymbol{x}})\, e^{2\pi i \underline{\boldsymbol{h}}.\underline{\boldsymbol{x}}}\; dx$$

Form Factor Equation

$$\boldsymbol{F}(\underline{\boldsymbol{h}}) = \int_0^1 \rho(\underline{\boldsymbol{x}})\, e^{2\pi i \underline{\boldsymbol{h}}.\underline{\boldsymbol{x}}}\, dx$$

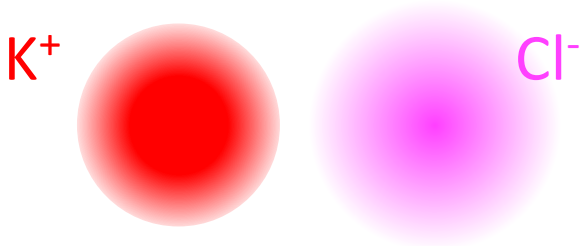


$$\boldsymbol{F}(\underline{\boldsymbol{h}}) = \sum_{j=1}^{N} |f_j|\, e^{2\pi i\, \underline{\boldsymbol{h}}.\underline{\boldsymbol{x}}_j} = \sum_{j=1}^{N} \boldsymbol{f}_j$$

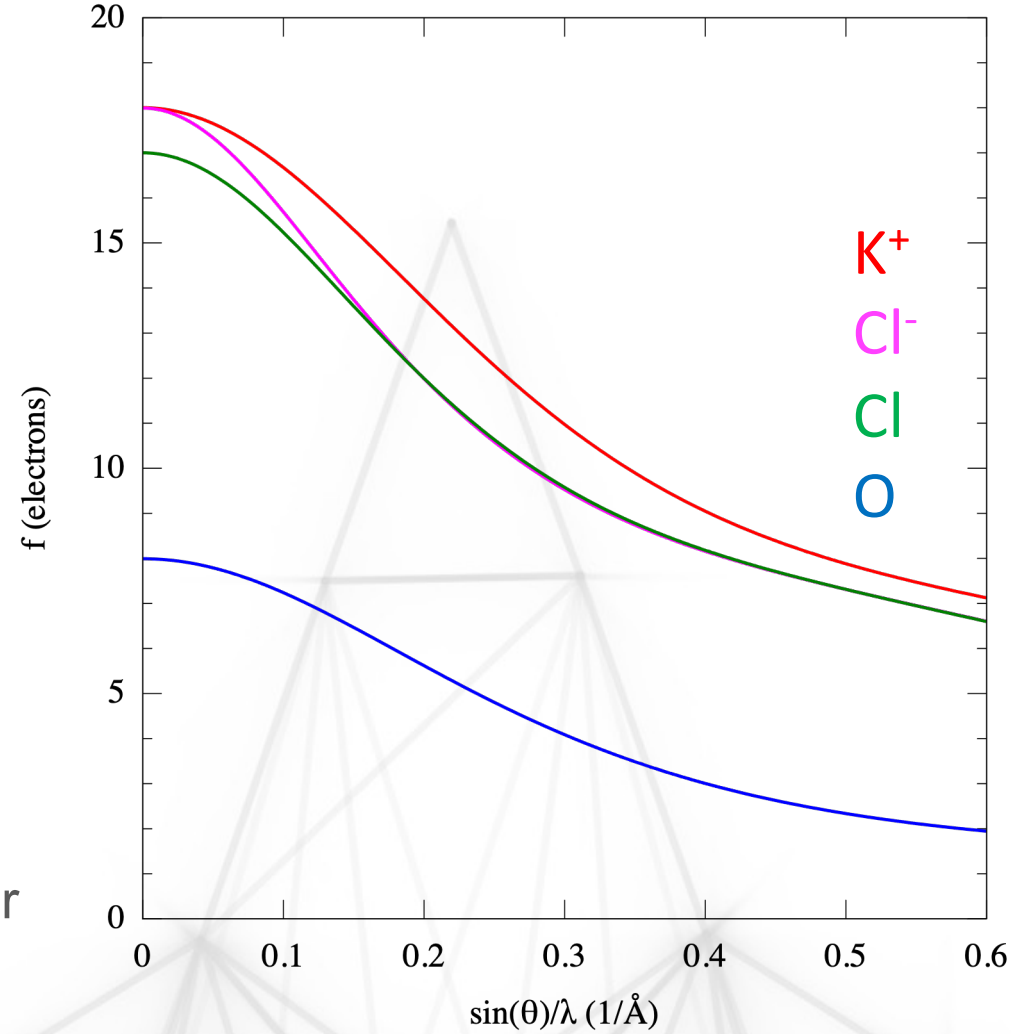
- for *N* particles at positions $\underline{\boldsymbol{x}}_j$
- form factor of jth particle = $\boldsymbol{f}_j$
  - Amplitude = $|\boldsymbol{f}_j|$
  - Phase = $\varphi_j = 2\pi\, \underline{\boldsymbol{h}}.\underline{\boldsymbol{x}}_j$

Atomic Form (Scattering) Factors

- Atoms & ions are diffuse
  - Like clouds of electrons



- Atomic Form Factors
  - Units = electrons (X-rays)
  - For θ=0, f = number electrons
    - Depends upon oxidation state
  - Decays with sin(θ) / λ
  - More diffuse ρ(r) leads to faster decay in f(θ)



Intensity Equation

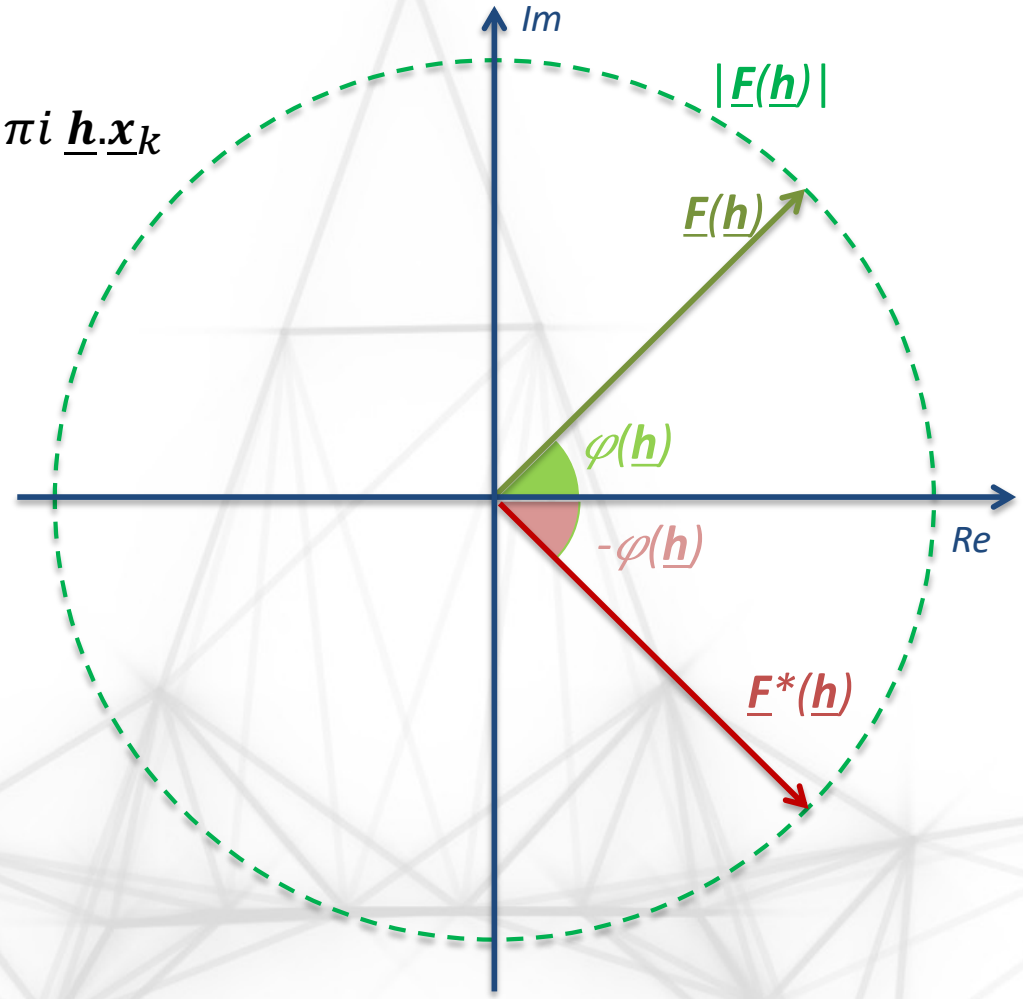
- The intensity of the scattered wave, $I(\underline{\boldsymbol{h}})$, is determined by multiplying the structure factor, $\underline{\boldsymbol{F}}(\underline{\boldsymbol{h}})$, by its complex conjugate

$$I(\underline{\boldsymbol{h}}) = \boldsymbol{F}(\underline{\boldsymbol{h}}) \times \boldsymbol{F}^*(\underline{\boldsymbol{h}})$$

$$= \sum_{j=1}^{N} \boldsymbol{f}_j \, e^{2\pi i \, \underline{\boldsymbol{h}}.\underline{\boldsymbol{x}}_j} \times \sum_{k=1}^{N} \boldsymbol{f}_k \, e^{-2\pi i \, \underline{\boldsymbol{h}}.\underline{\boldsymbol{x}}_k}$$

$$= \sum_{j=1}^{N} \sum_{k=1}^{N} \boldsymbol{f}_j \, \boldsymbol{f}_k \, e^{2\pi i \, \underline{\boldsymbol{h}}.(\underline{\boldsymbol{x}}_j - \underline{\boldsymbol{x}}_k)}$$

$$= \left|\boldsymbol{F}(\underline{\boldsymbol{h}})\right|^2$$



Reconstucting Density, $\rho(\underline{x})$

Fourier Theorem & Series

Jean-Baptisite Joseph FOURIER (1768-1830)

$$\boldsymbol{F}(\underline{\boldsymbol{S}}) = \int \rho(\underline{\boldsymbol{r}})\, e^{2\pi i \underline{\boldsymbol{S}}.\underline{\boldsymbol{r}}}\, dr$$

$$\rho(\underline{\boldsymbol{r}}) = \int F(\underline{\boldsymbol{S}})\, e^{-2\pi i \underline{\boldsymbol{S}}.\underline{\boldsymbol{r}}}\, dS$$

For repeating systems, such as a crystal (***x***), the integral develops into a summation of the coefficients at discrete points in reciprocal space (***h***):

$$\rho(\underline{\boldsymbol{x}}) = \sum_{\underline{\boldsymbol{h}}} \boldsymbol{F}(\underline{\boldsymbol{h}})\, e^{-2\pi i \underline{\boldsymbol{h}}.\underline{\boldsymbol{x}}}$$

What else did Joseph Fourier discover?

Fourier Series

- For repeating systems, the integral develops into a sum
  - crystal with a repeating density, $\rho(\underline{\boldsymbol{x}})$,
  - sum the coefficients at discrete points in reciprocal space, $F(\underline{\boldsymbol{h}})$

$$\rho(\underline{\boldsymbol{x}}) = \sum_{\underline{\boldsymbol{h}}} \boldsymbol{F}(\underline{\boldsymbol{h}})\, e^{-2\pi i \underline{\boldsymbol{h}}.\underline{\boldsymbol{x}}}$$

$$\rho(\underline{\boldsymbol{x}}) = \sum_{\underline{\boldsymbol{h}}} \left|\boldsymbol{F}(\underline{\boldsymbol{h}})\right| e^{i\varphi(\underline{\boldsymbol{h}})} e^{-2\pi i \underline{\boldsymbol{h}}.\underline{\boldsymbol{x}}}$$

Amplitude Phase Sinusoid

Example of a Fourier Series

- From an *Outline of Crystallography for Biologists* Blow (2002)

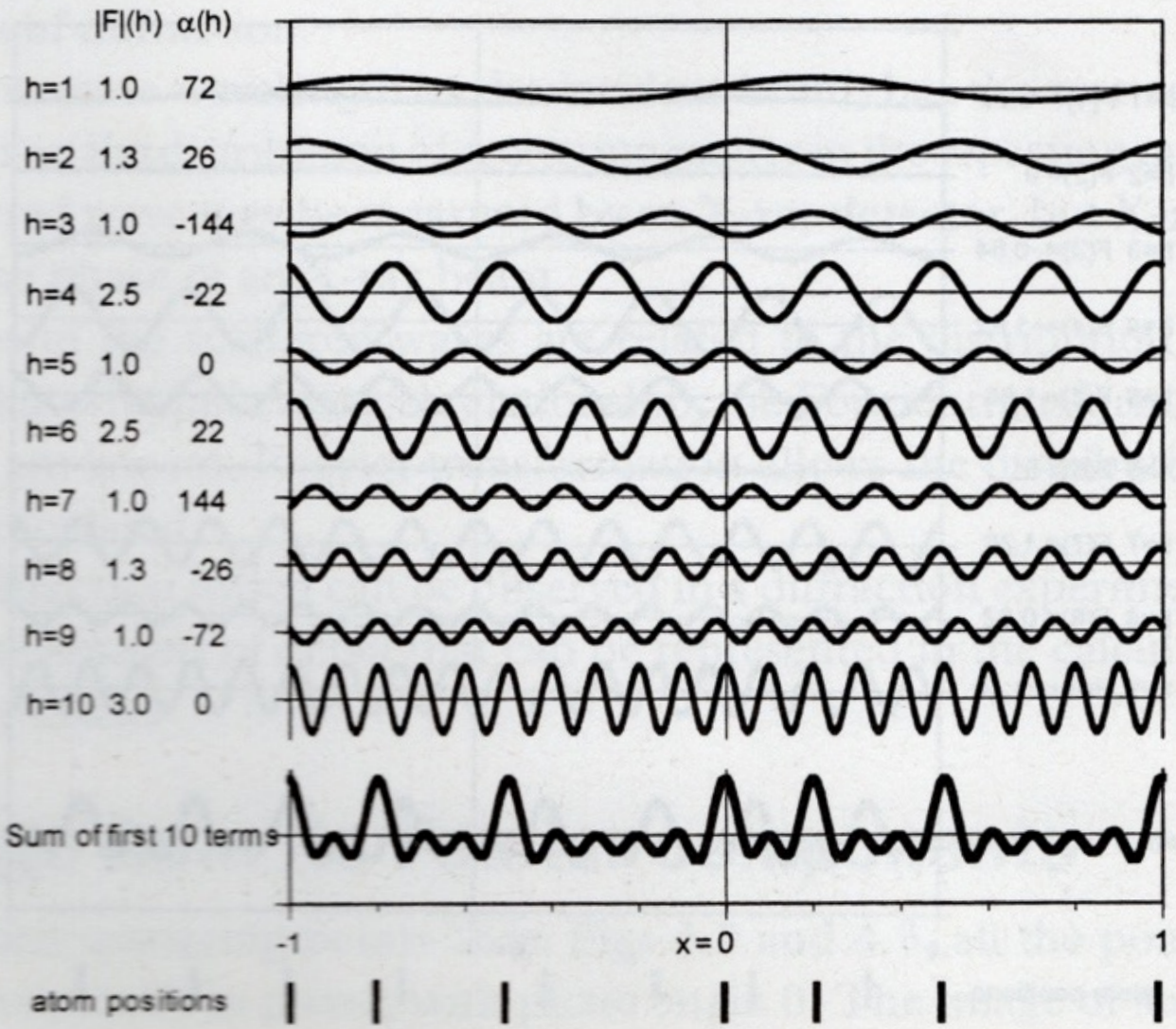
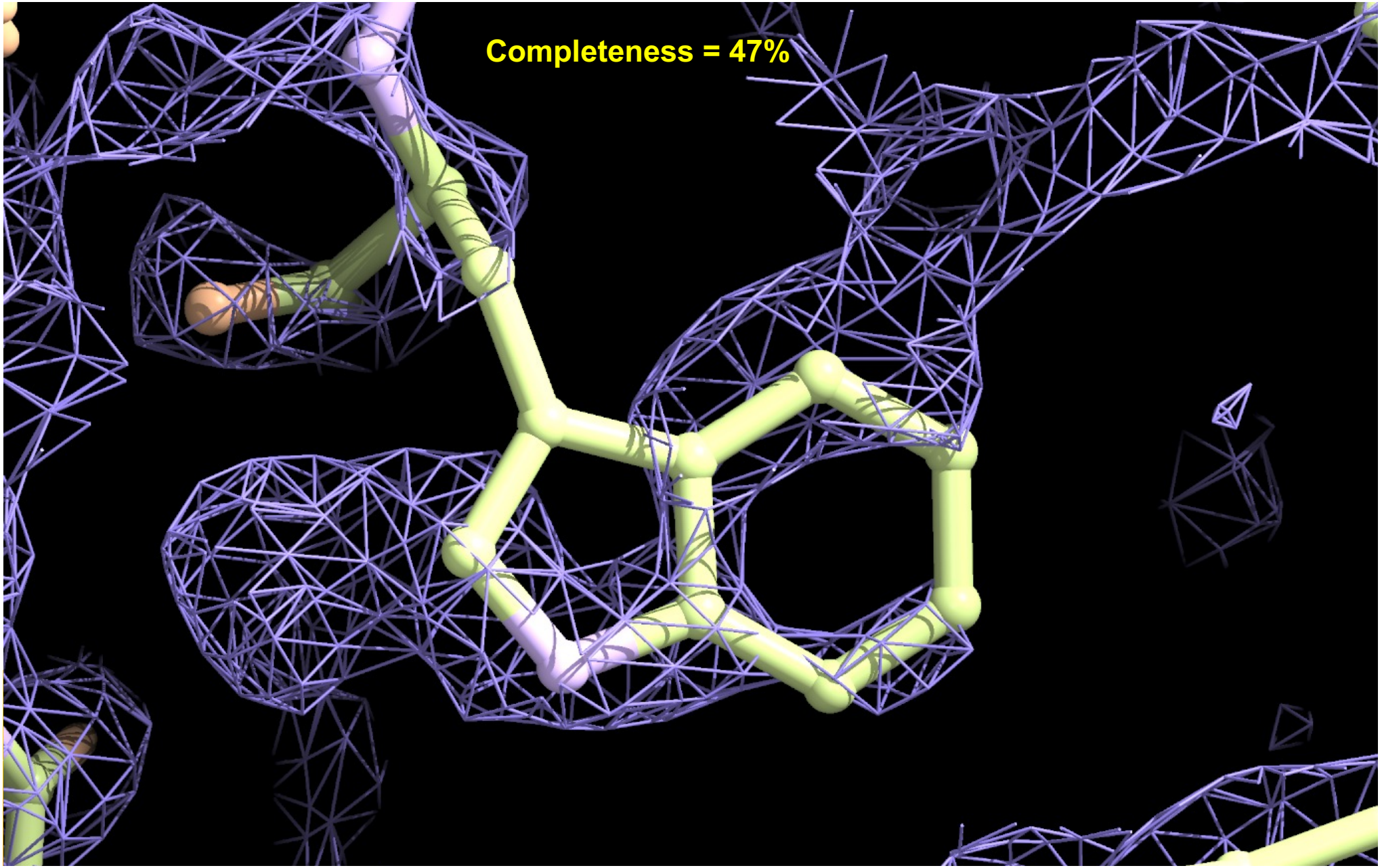


Fig. 4.12 Fourier summation based on a non-centrosymmetric structure with atoms at $x=0$, 0.2, 0.5. The Fourier terms have various phases, $\alpha(h)$.

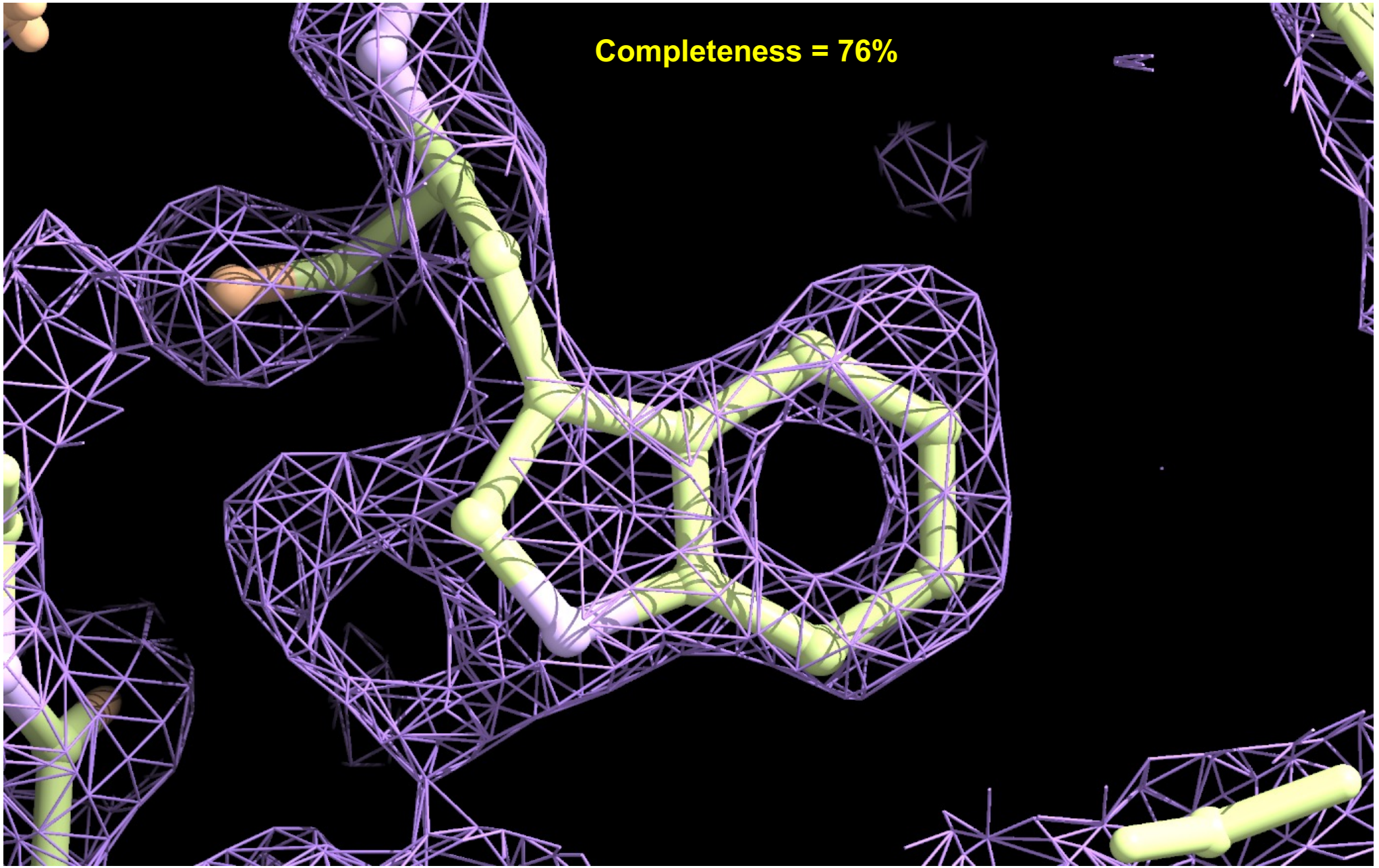
Properties of Fourier Series

- Conception of “resolution”
  - Higher order harmonics (***h***, Miller indices) provide more detail!
- Completion
  - All coefficients |***F***(***h***)|contribute to the reconstruction of the map
    - Stronger amplitudes contribute more
    - Weaker amplitudes contribute less

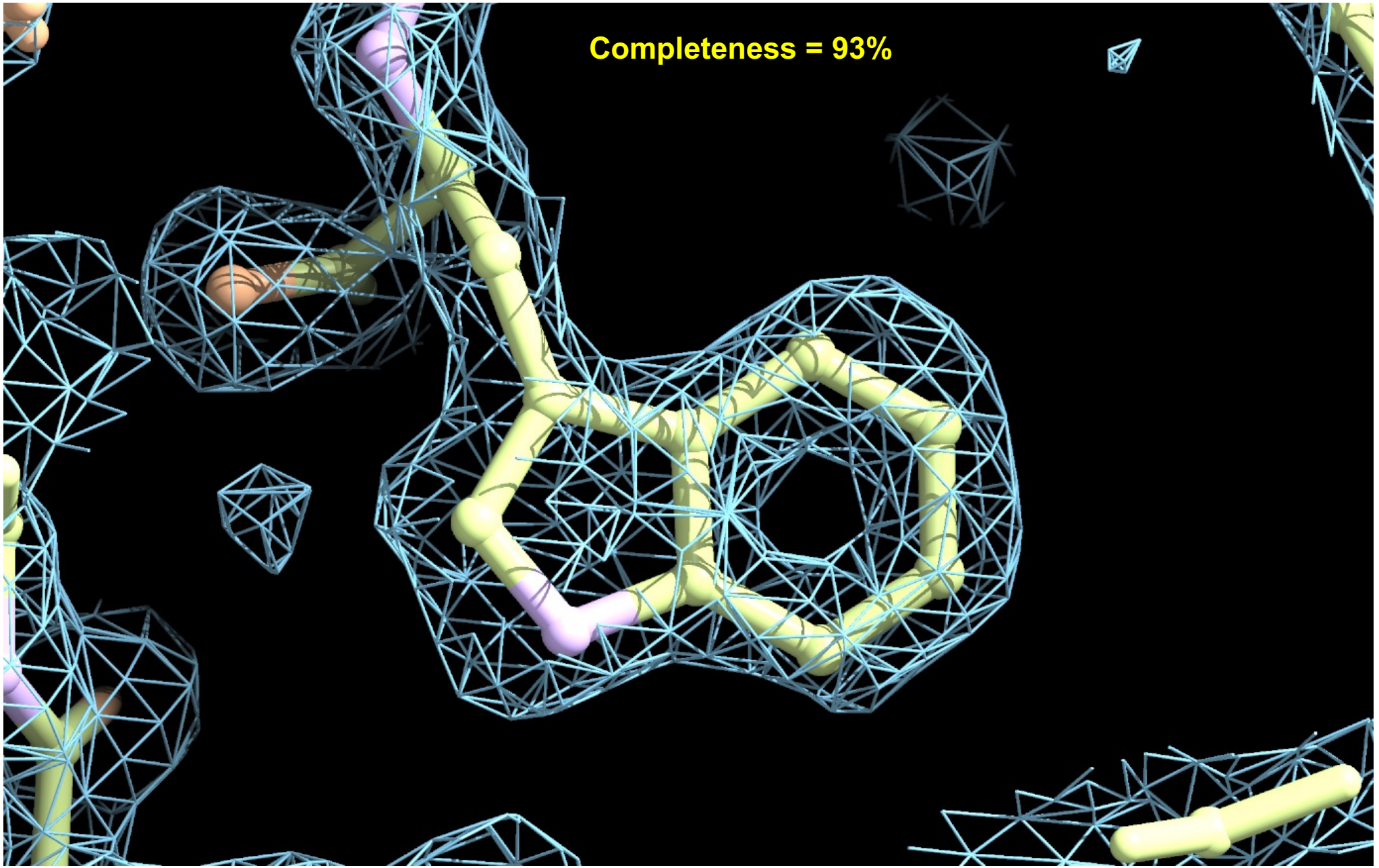
Effects of Completion



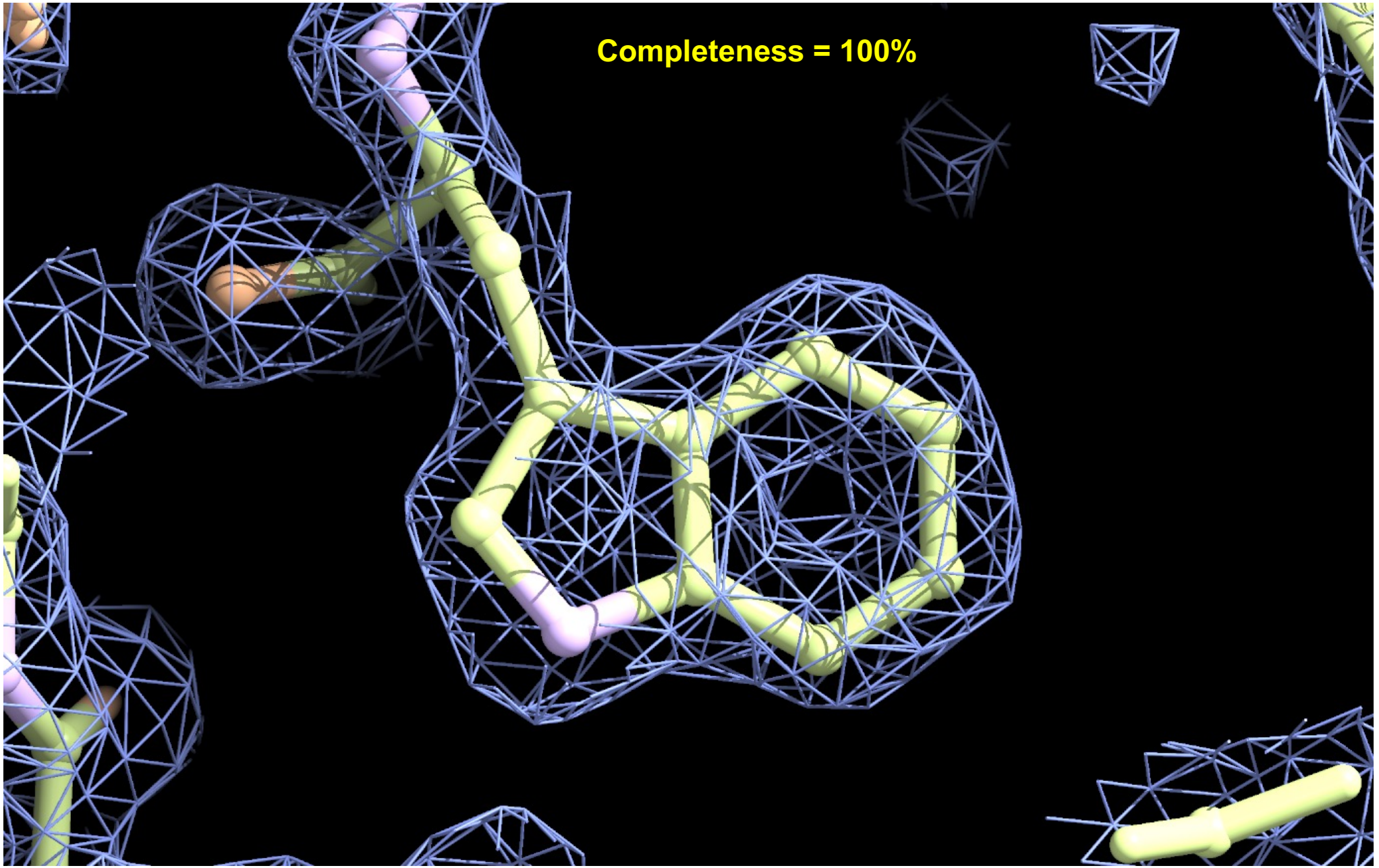
Effects of Completion



Effects of Completion



Effects of Completion



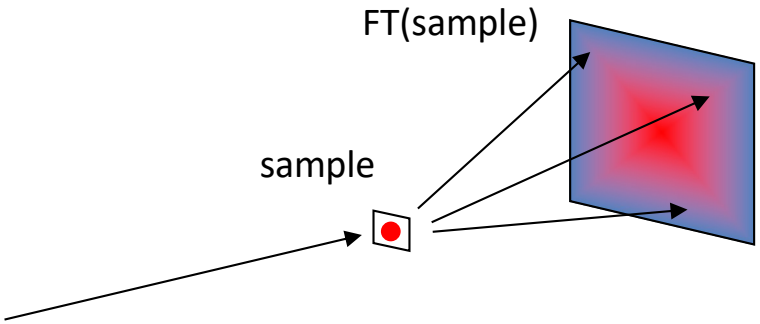
Properties of Fourier Series

- Conception of “resolution”
  - Higher order harmonics (***h***, Miller indices) provide more detail!
- Completion
  - All coefficients |***F***(***h***)|contribute to the reconstruction of the map
    - Stronger amplitudes contribute more
    - Weaker amplitudes contribute less
- Importance of phases
  - The phases, $\varphi(\mathbf{h})$, are very, very important…
  - This is known as the ***Phase Problem***

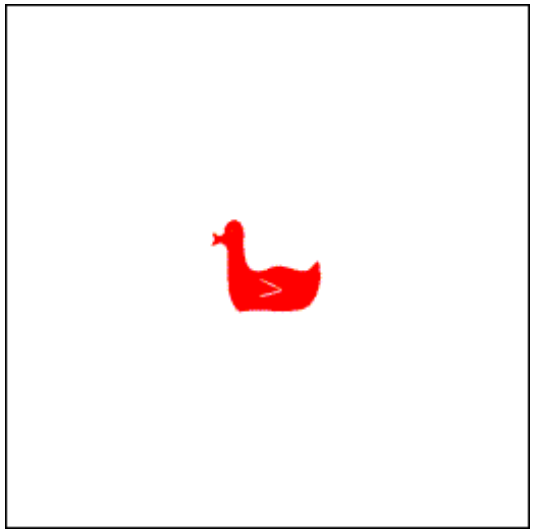
The Phase Problem

Duck & Cat Fourier Transforms

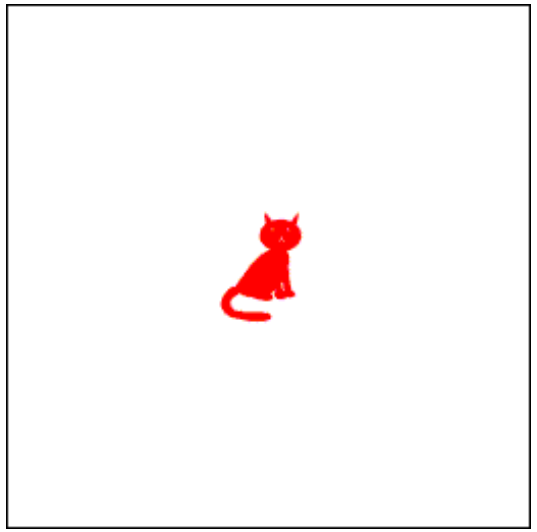
Courtesy of Kevin Cowtan http://www.yorvic.york.ac.uk/~cowtan/fourier/fourier.html



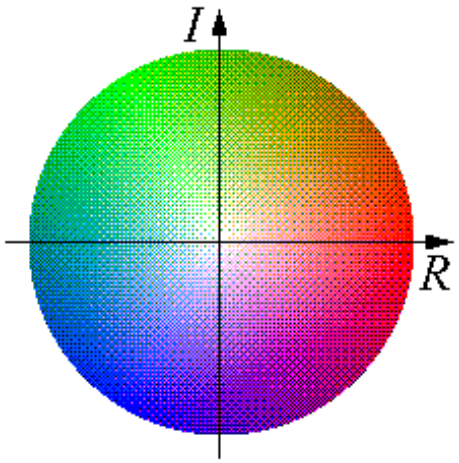
0-dimensional diffraction experiment



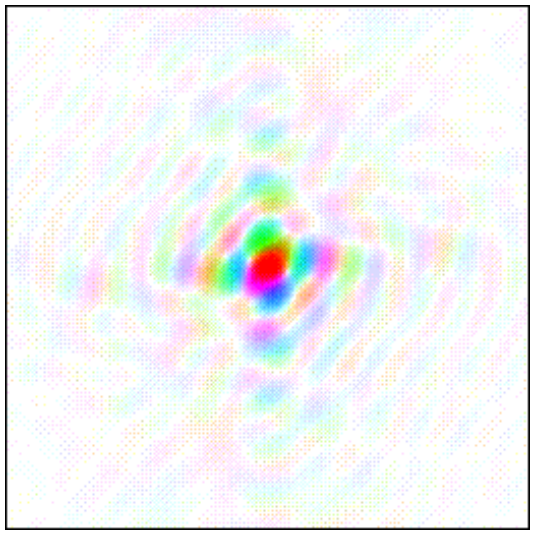
duck



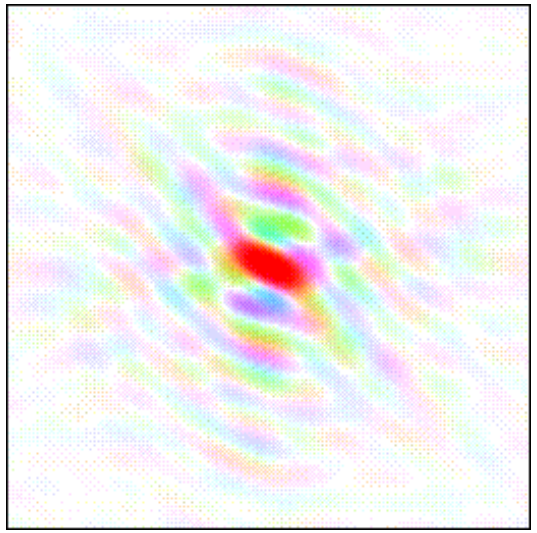
cat



Amplitude = intensity
Phase = colour

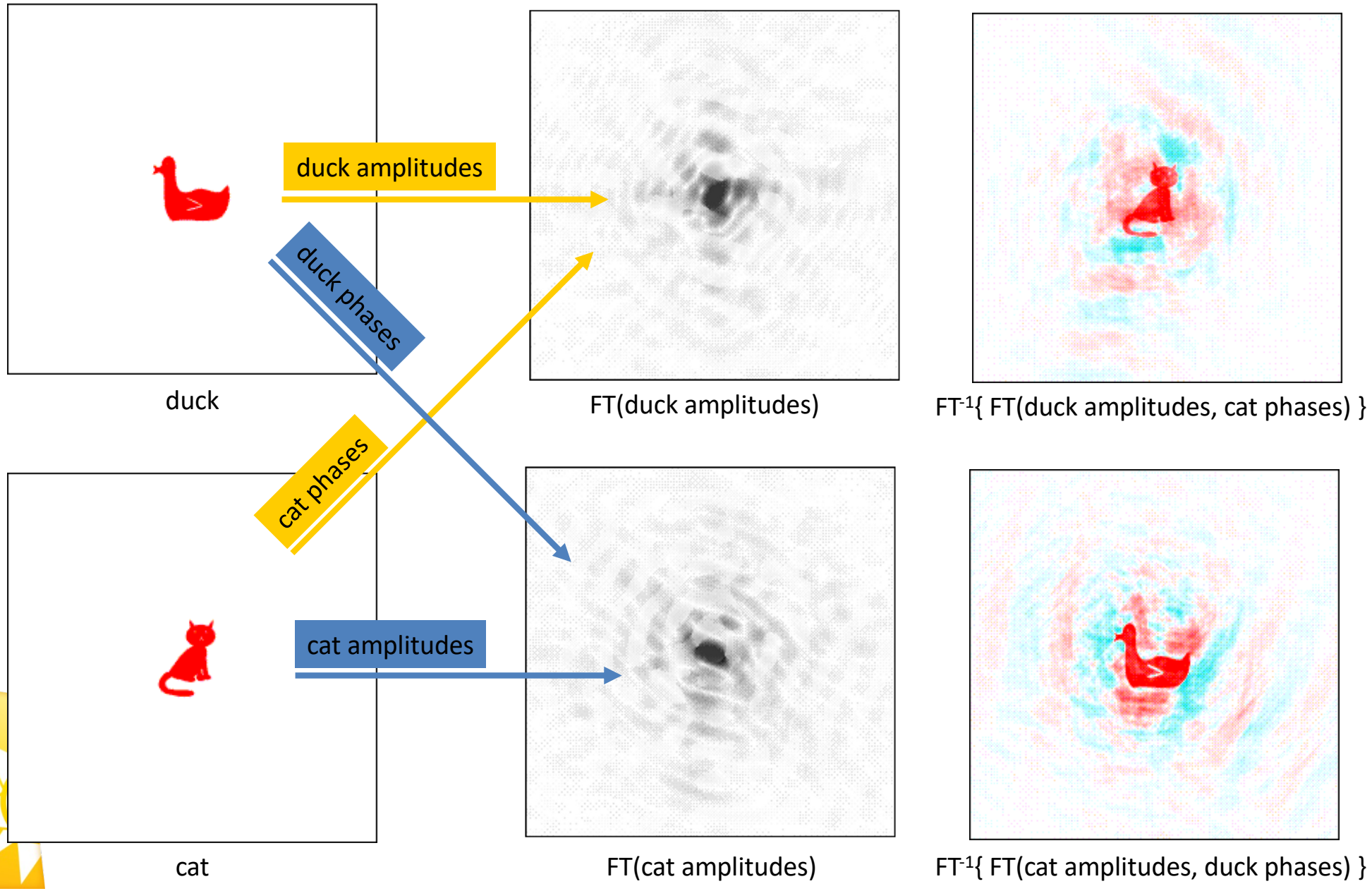


FT(duck)



FT(cat)

Duck & Cat Cross Phasing



The Phase Problem

If we wish to determine a structure with the Fourier series…

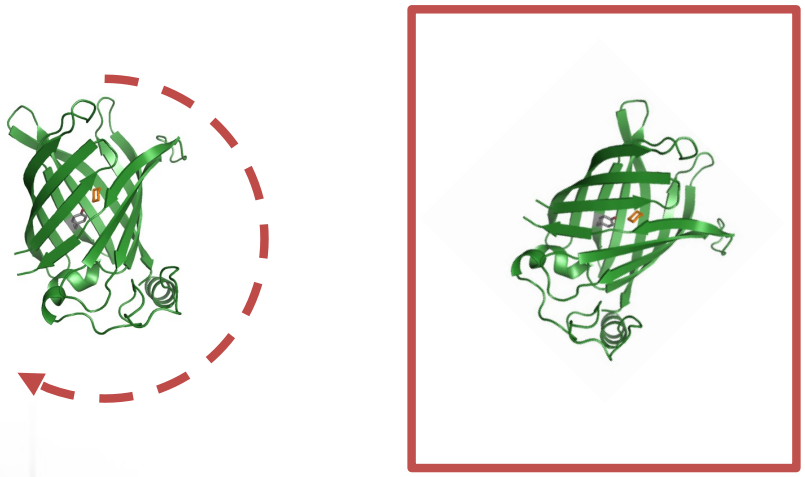
- Q: Can we *directly measure* phases?
  - No, although phases can be determined *indirectly*…
- Q: Can we *apply random* phases?
  - No, we will just get noise…
- Q: Can we set all phases to a *constant*?
  - No, this is akin to an auto-correlation (Patterson) function…
- Q: Can we *ignore* the phases?
  - No, because the ***phases are derived from the positions of the scatterers!***
  - Thus, the inverse is also true, ***the positions of the scatterers are derived from the phases***!

Solving the Phase Problem

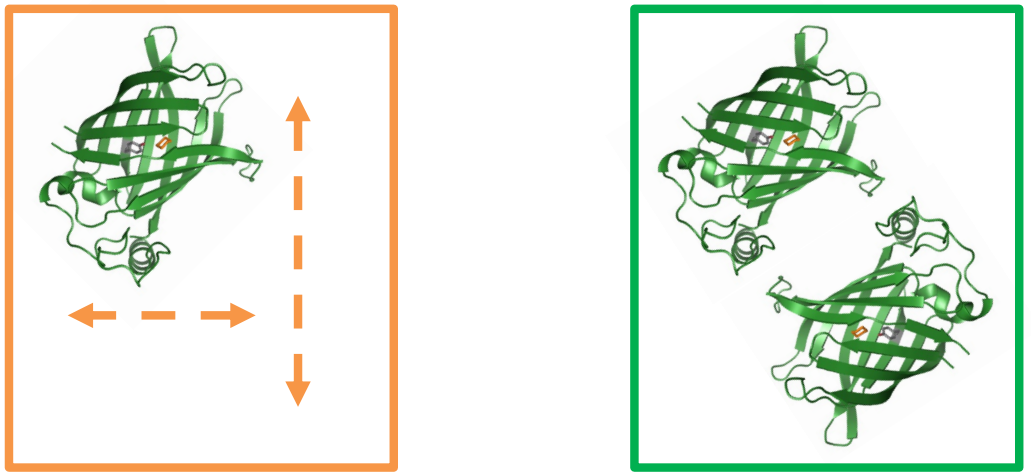
- Small molecules (< 200 non-H atoms)
  - Heavy Atom Method
    - A small number of relatively heavier atoms with dominate the phases…
  - Direct Methods
    - $\varphi(\underline{\boldsymbol{H}}) + \varphi(\underline{\boldsymbol{H}} - \underline{\boldsymbol{K}}) + \varphi(\underline{\boldsymbol{H}} - \underline{\boldsymbol{L}}) = 0$
      - if $\underline{\boldsymbol{H}} - \underline{\boldsymbol{K}} - \underline{\boldsymbol{L}} = \mathbf{0}$, and if $|E(\underline{\boldsymbol{H}})|, |E(\underline{\boldsymbol{K}})|, |E(\underline{\boldsymbol{L}})| > 1.5$
      - Very high resolution data
        - beyond 1.2Å
- Macromolecules (> 200 non-H atoms)
  - Molecular Replacement
    - Six-dimensional search with a suitable starting model
      - > 30% sequence identity
      - Very big impact by AlphaFold2 et al.
  - Single & Multiple Isomorphous Replacement (SIR & MIR)
    - Add a heavy atom to the crystal structure
      - Hg, Pt, etc…
  - Single & Multiple-wavelength Anomalous Diffraction (SAD & MAD)
    - Collect at X-ray wavelengths near and/or above an absorption edge
      - Intrinsic heavy atom (P, S, Mn, Fe, Cu, Zn, Se, Br, etc…)
      - Incorporated heavy atom (Hg, Pt, I, lanthanides, polyoxometalate clusters, etc…)

Molecular Replacement

- 3-dimensional Rotation Search

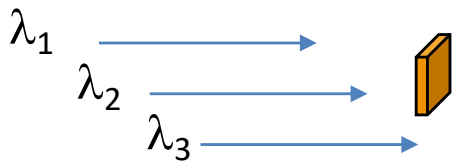


- 3-dimensional Translation Search

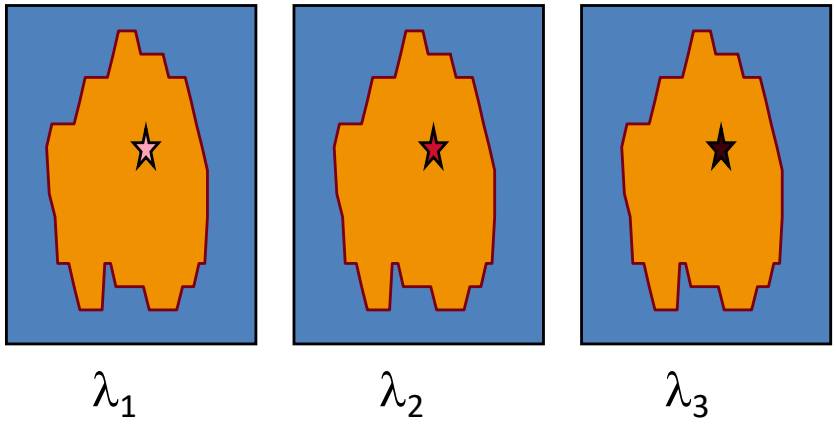


Comparison of MAD & MIR

MAD

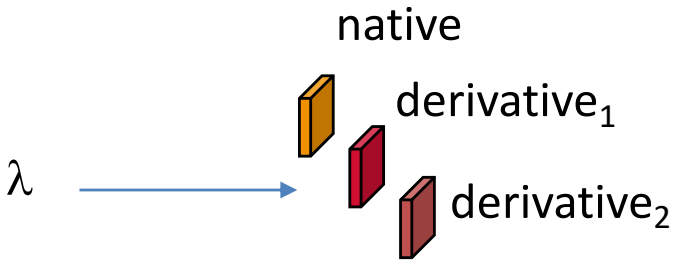


Strict isomorphism

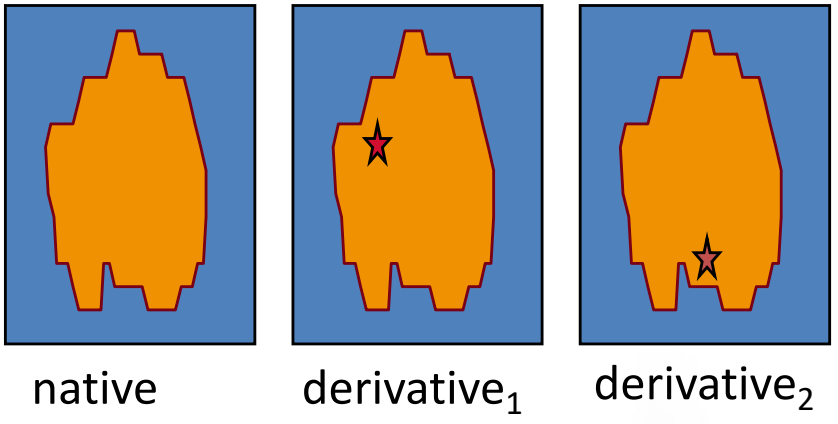


Weak signal < 5%

MIR



Issues of non-isomorphism



Strong signal > 10%

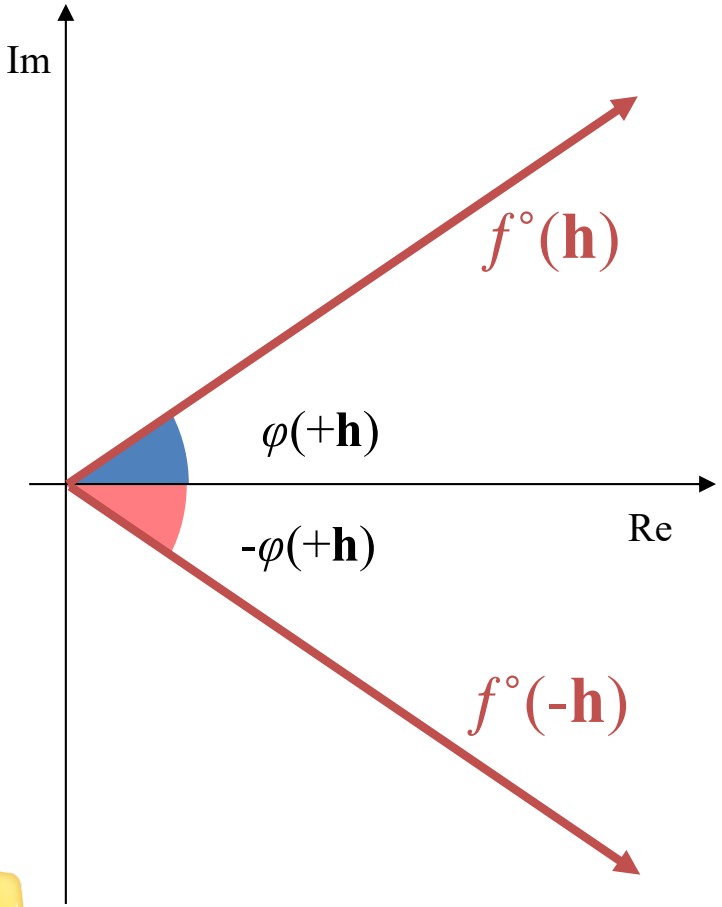
- Both methods use the heavy atom sites as **references** to phase for the crystal structure

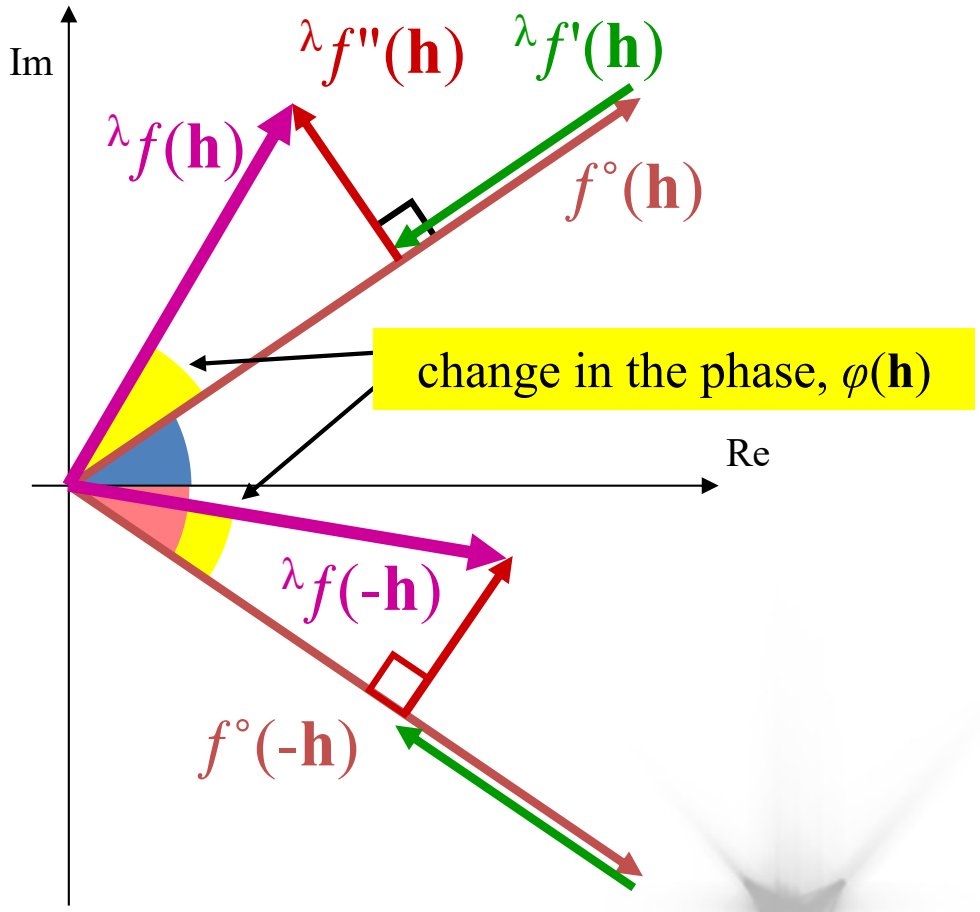
The Effect of Anomalous Scattering

normal scattering

anomalous scattering

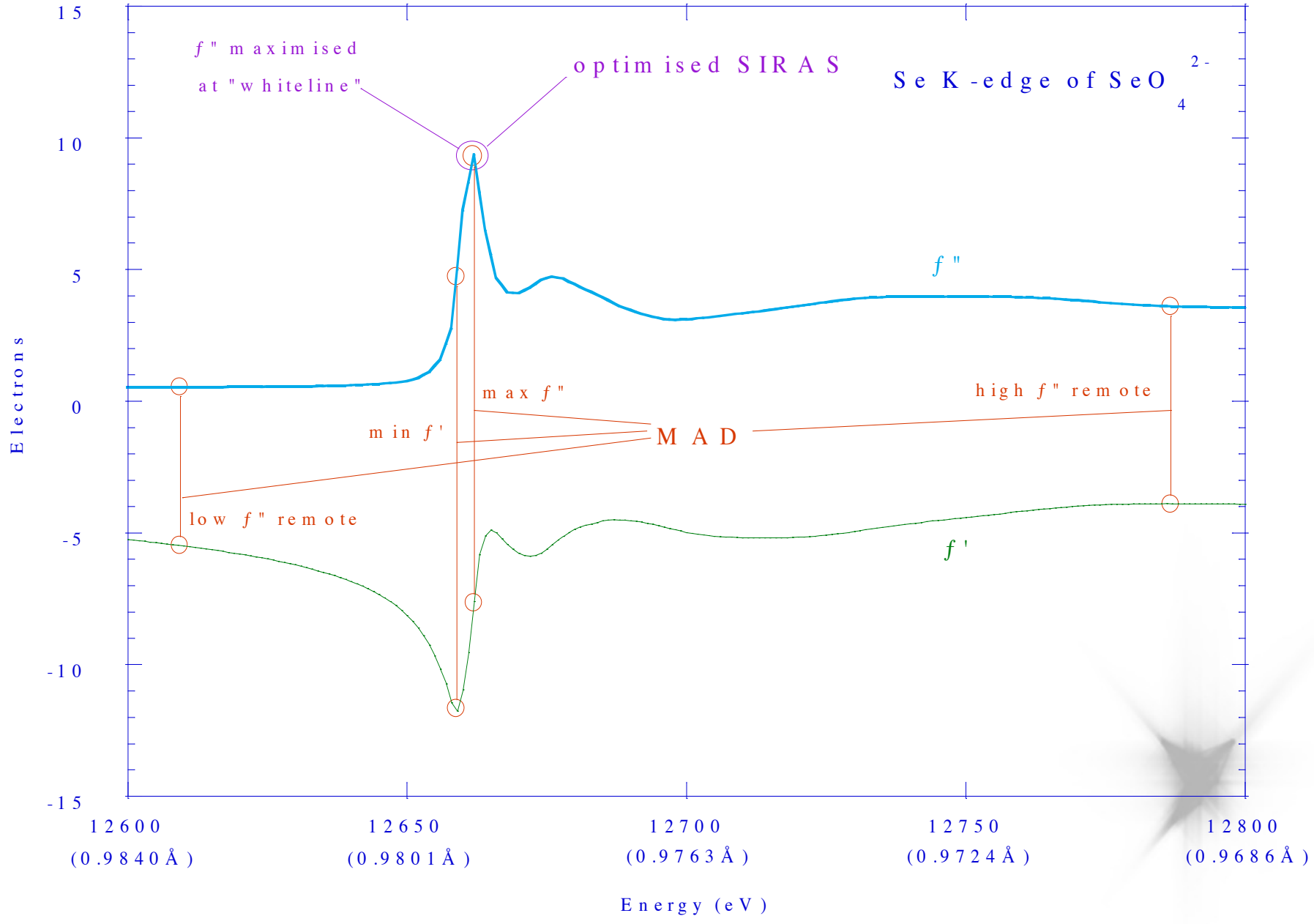
$^{\lambda}f(\mathbf{h}) = f^{\circ}(\mathbf{h}) + {}^{\lambda}f'(\mathbf{h}) + i\,{}^{\lambda}f''(\mathbf{h})$





$^{\lambda}f''(\mathbf{h})$ always lags 90° (+i) behind $f^{\circ}(\mathbf{h})$!

Variation of Anomalous Scattering Factors *f'* & *f"*

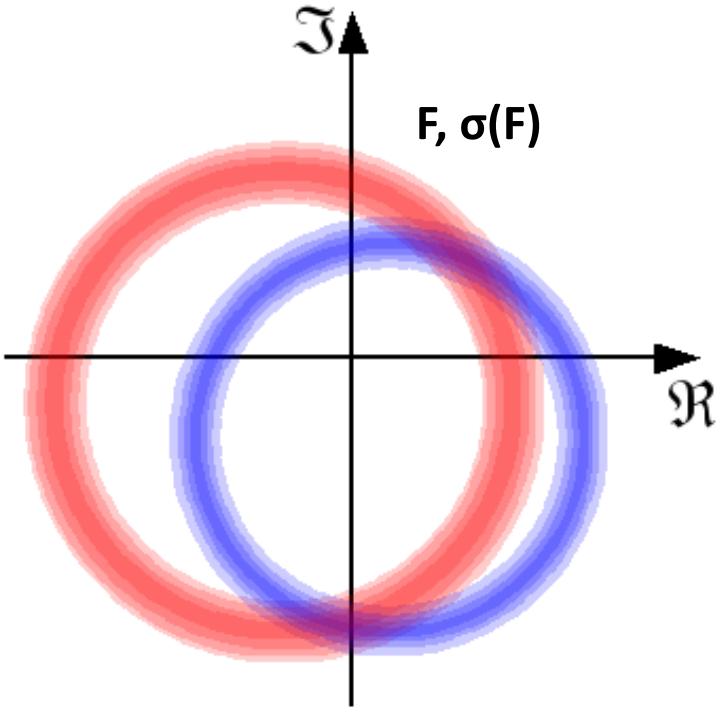


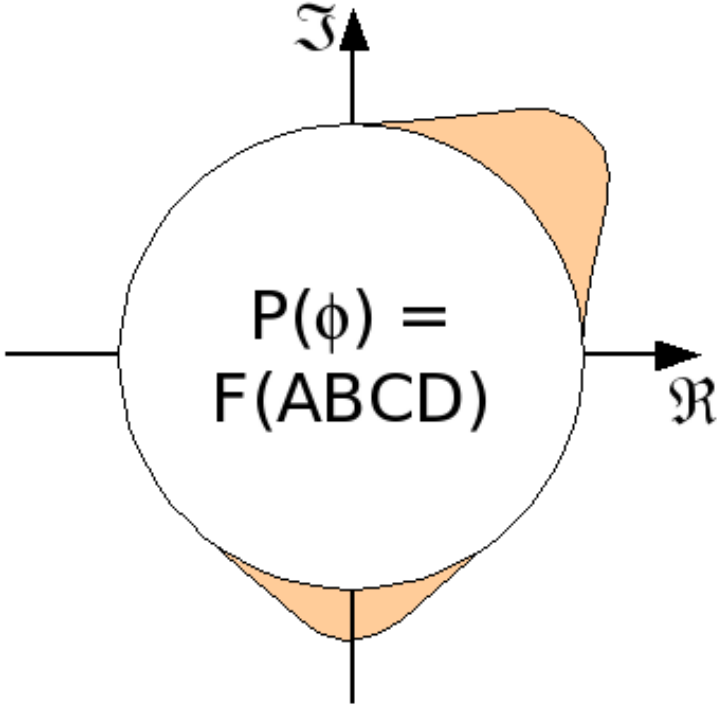
Finding Heavy (Reference) Atoms

- Patterson Methods
  - $|\underline{F}_{PH}(h)| - |\underline{F}_{P}(h)|$ or $|\underline{F}_{PH}(+h)| - |\underline{F}_{PH}(-h)|$
  - Good for few heavy atoms
    - $N < 5$
  - Equivalent to heavy atom method
    - Small molecule metal complexes
- Direct Methods
  - Good for several heavy and equal atoms
    - $5 < N < 150$
  - More complicated & statistical method
    - Originally used for organic molecules (C, H, N, O, etc…)

Hendrickson-Lattman Coefficients

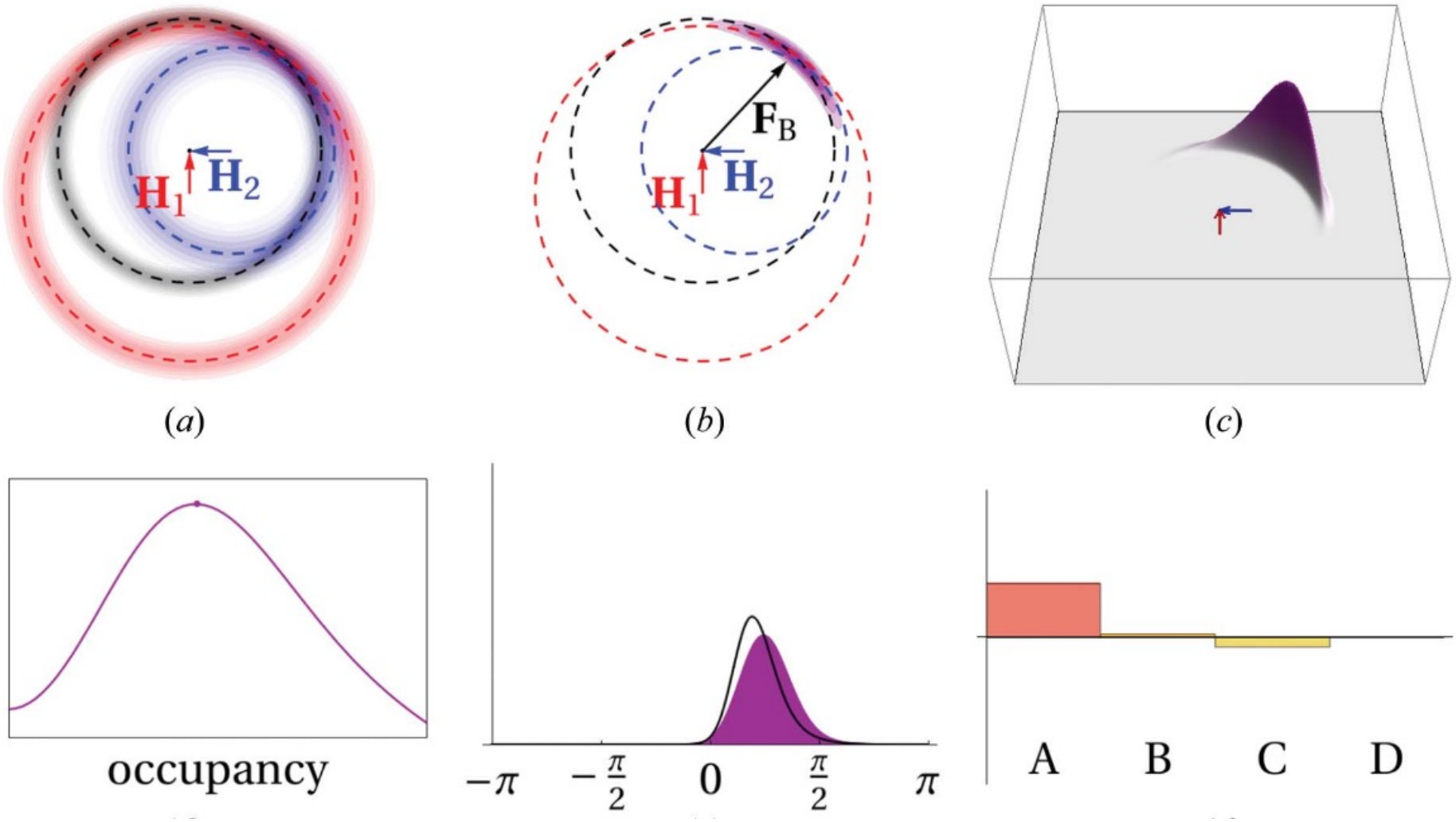
$$P(\varphi) = N \exp(A \cos(\varphi) + B \sin(\varphi) + C \cos(2\varphi) + D \sin(2\varphi))$$





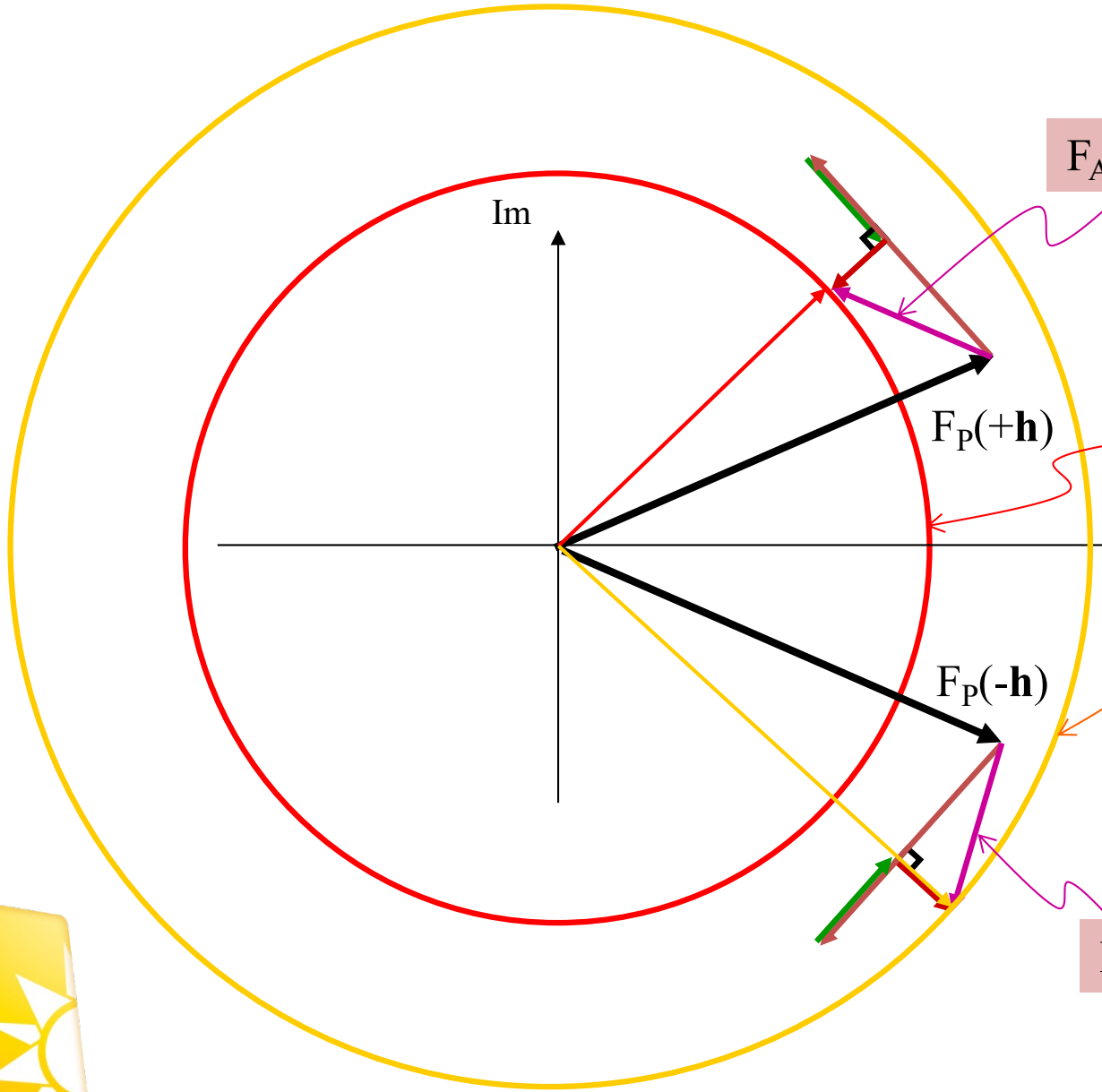
Courtesy http://ccp4wiki.org/~ccp4wiki/wiki/index.php?title=Hendrickson_Lattman_coefficients

MIR or MAD Phasing



Courtesy McCoy & Read, 2010

Anomalous Scattering & The Breakdown of Friedel's Law



$F_A(+\mathbf{h})$

$|F_T(+\mathbf{h})|$ circle

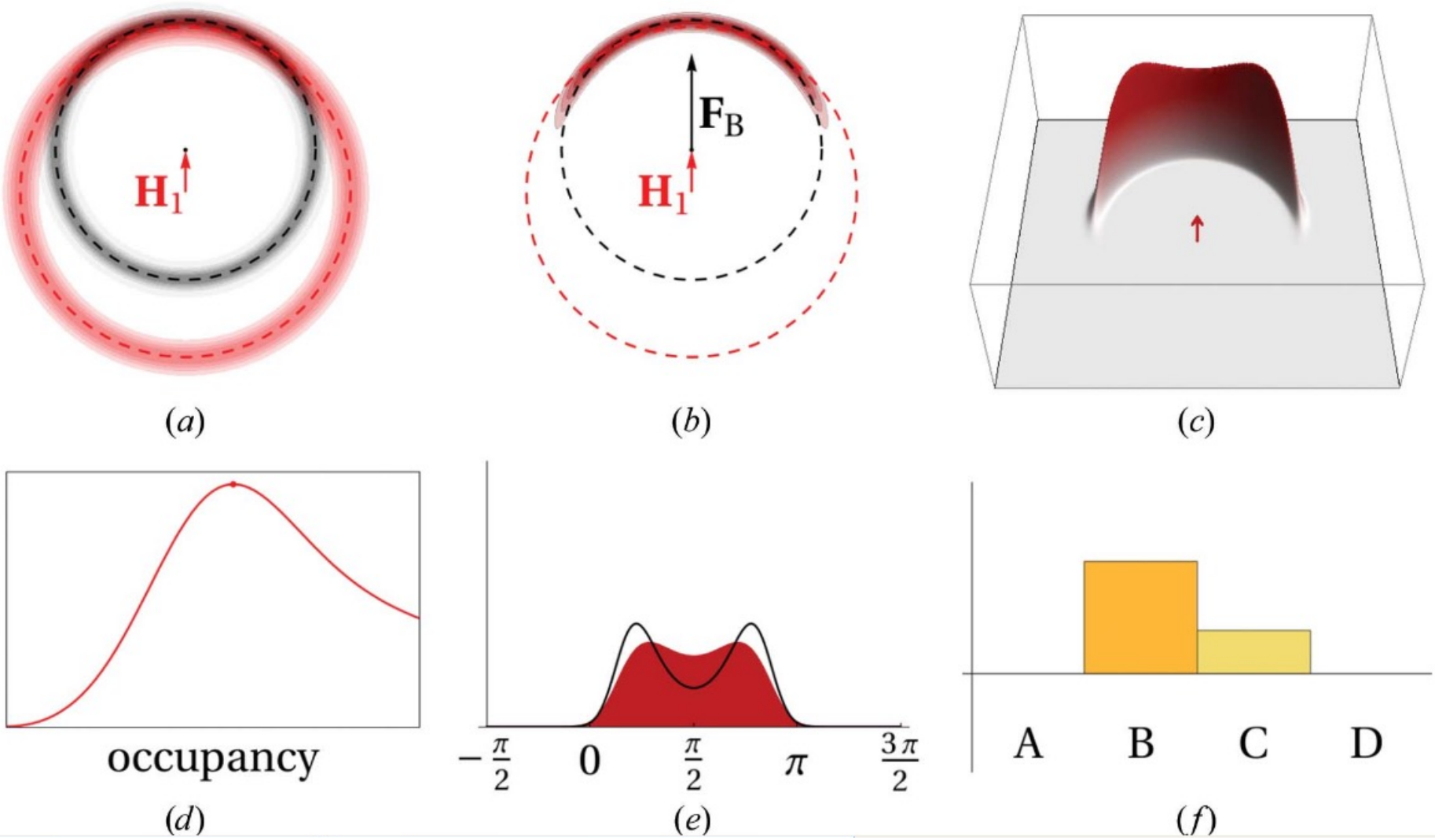
$|F_T(-\mathbf{h})|$ circle

$F_A(-\mathbf{h})$

Friedel's law
$|F(+\mathbf{h})| = |F(-\mathbf{h})|$

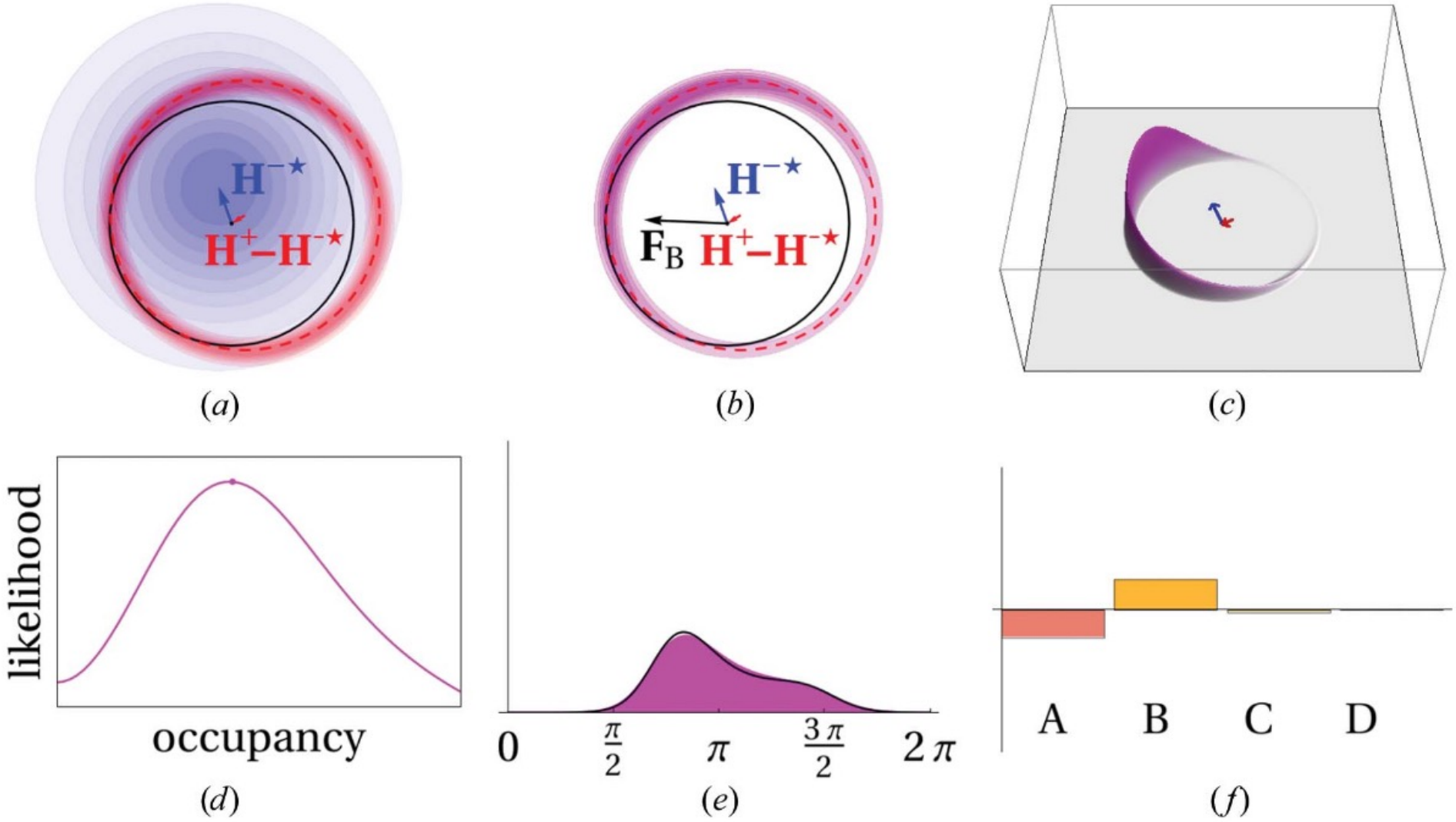
With anomalous scattering ($\lambda f''$)
$|F(+\mathbf{h})| \neq |F(-\mathbf{h})|$

SIR Phasing



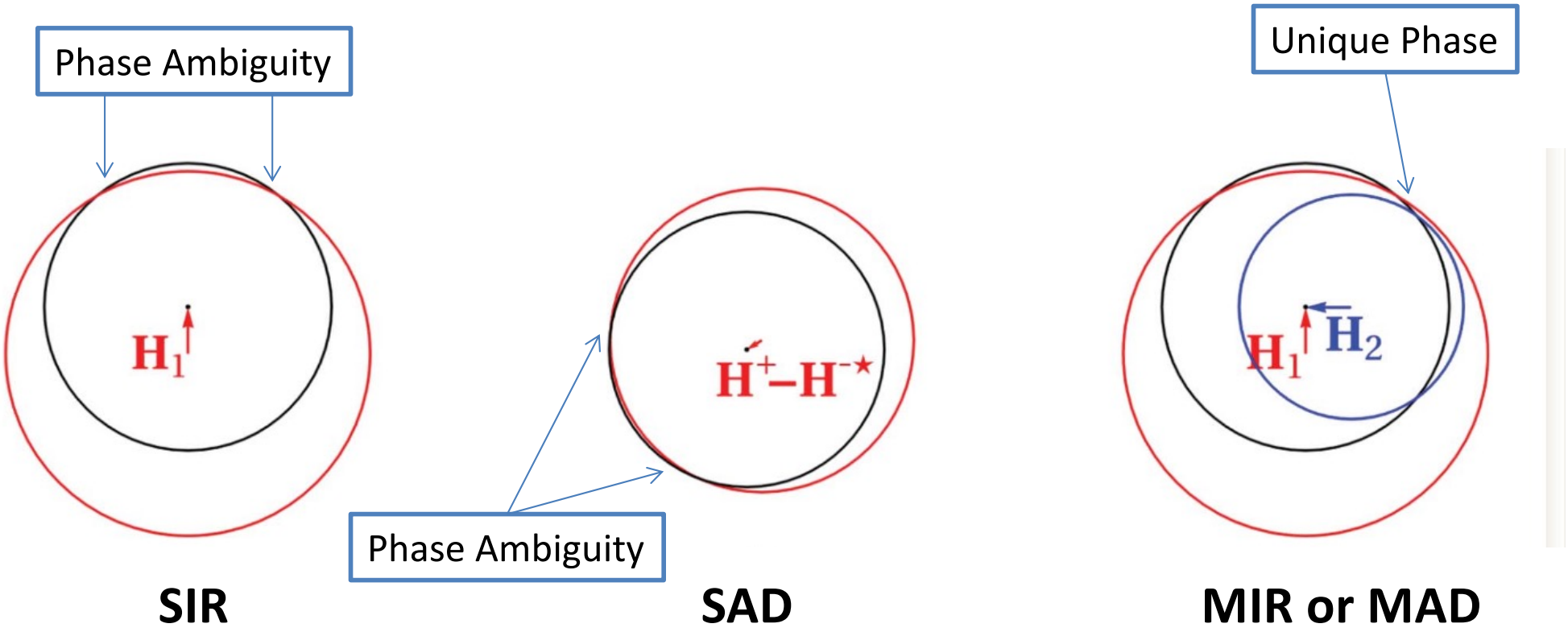
Courtesy McCoy & Read, 2010

SAD Phasing



Courtesy McCoy & Read, 2010

Different Phasing Methods

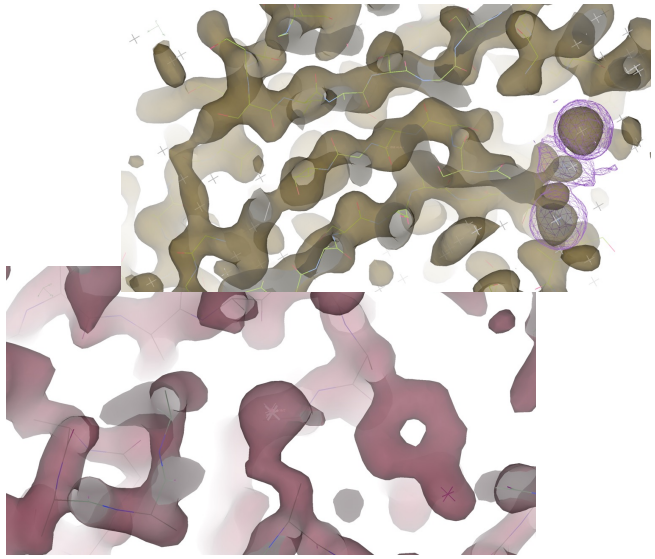


$$P(\varphi) = N \exp(A \cos(\varphi) + B \sin(\varphi) + C \cos(2\varphi) + D \sin(2\varphi))$$

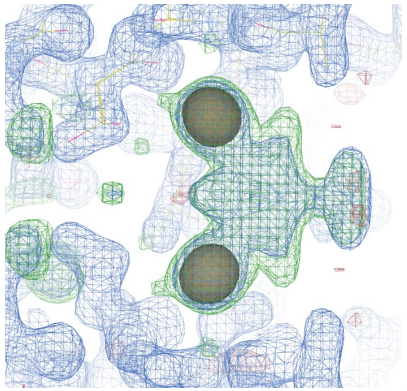
Courtesy McCoy & Read, 2010

Interpretation, Refinement & Validation

- Interpretation
  - Tracing polypeptide mainchain
  - Fitting of sidechains
  - Solvent molecules
- Refinement
  - Constrained & Restrained
    - Low observable:parameter ratio
  - Least Squares & Maximum Likelihood
  - R-factors
    - Rwork, Rfree
- Validation
  - Checks on geometry (stereochemistry)
    - Bond lengths & angles, planarity, etc…
  - Clashes
    - Intermolecular contacts
  - Spurious density
    - $2F_{obs} - F_{calc}$ & $F_{obs} - F_{calc}$



$$R = \frac{\sum_{\underline{h}} \left| |F_{obs}(\underline{h})| - |F_{calc}(\underline{h})| \right|}{\sum_{\underline{h}} |F_{obs}(\underline{h})|}$$

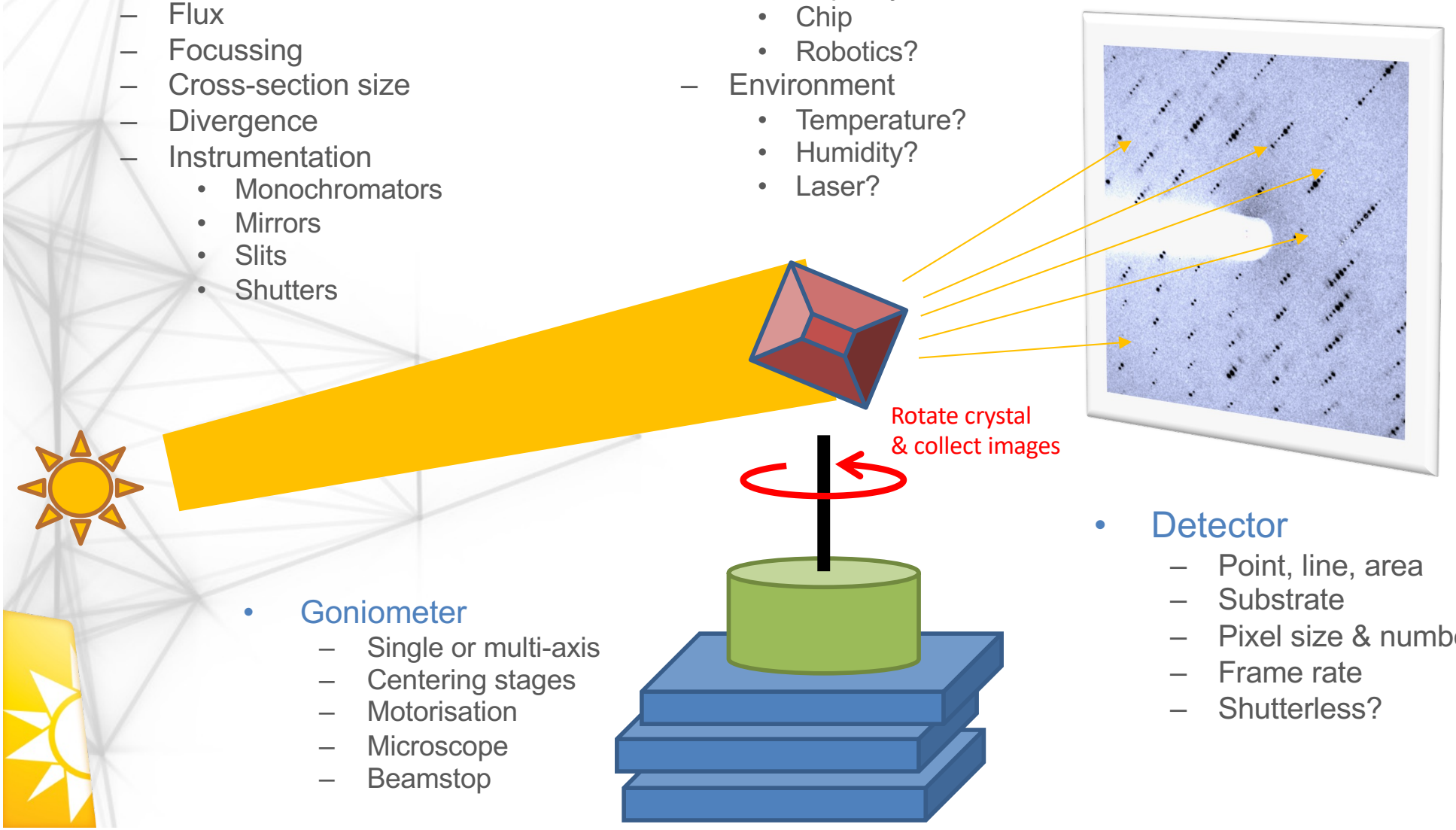


The Experimental Set Up

The Basic Experimental Set Up

- Source
  - X-rays, electrons, neutrons
  - Energy
  - Bandpass
  - Flux
  - Focussing
  - Cross-section size
  - Divergence
  - Instrumentation
    - Monochromators
    - Mirrors
    - Slits
    - Shutters

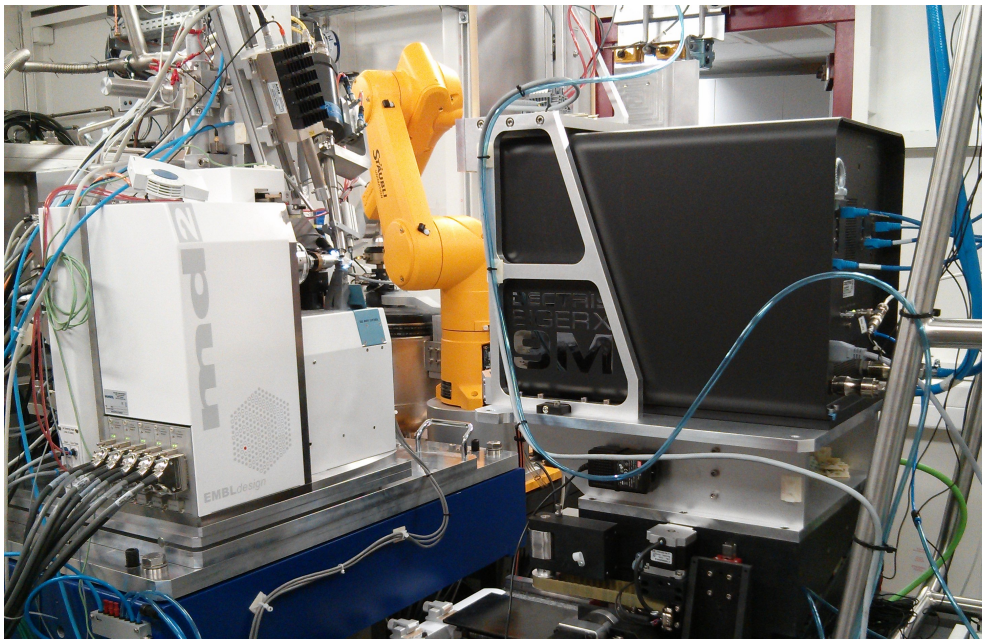
- Sample
  - Mounting
    - Loop
    - Capillary
    - Chip
    - Robotics?
  - Environment
    - Temperature?
    - Humidity?
    - Laser?

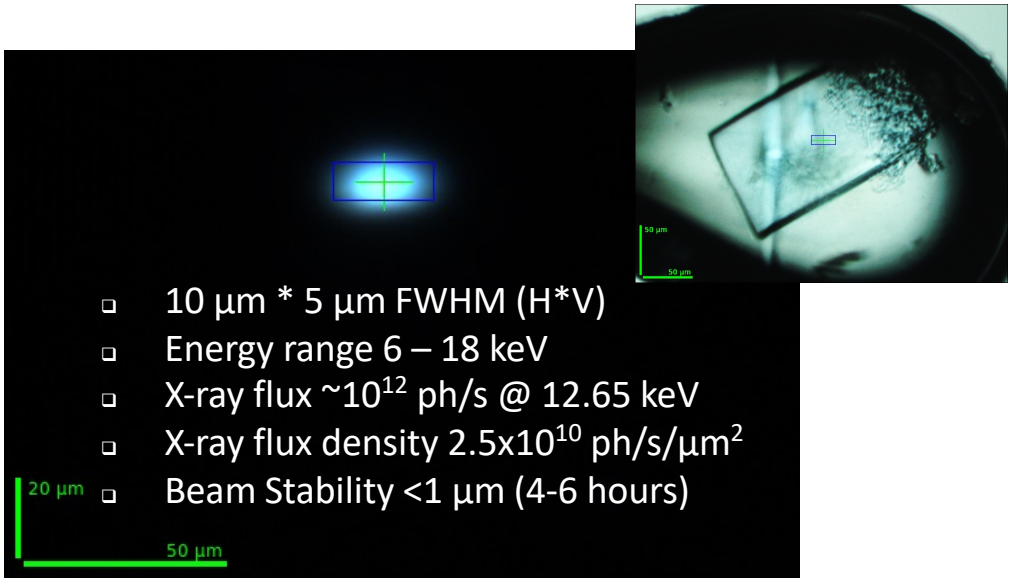


- Goniometer
  - Single or multi-axis
  - Centering stages
  - Motorisation
  - Microscope
  - Beamstop

- Detector
  - Point, line, area
  - Substrate
  - Pixel size & number
  - Frame rate
  - Shutterless?

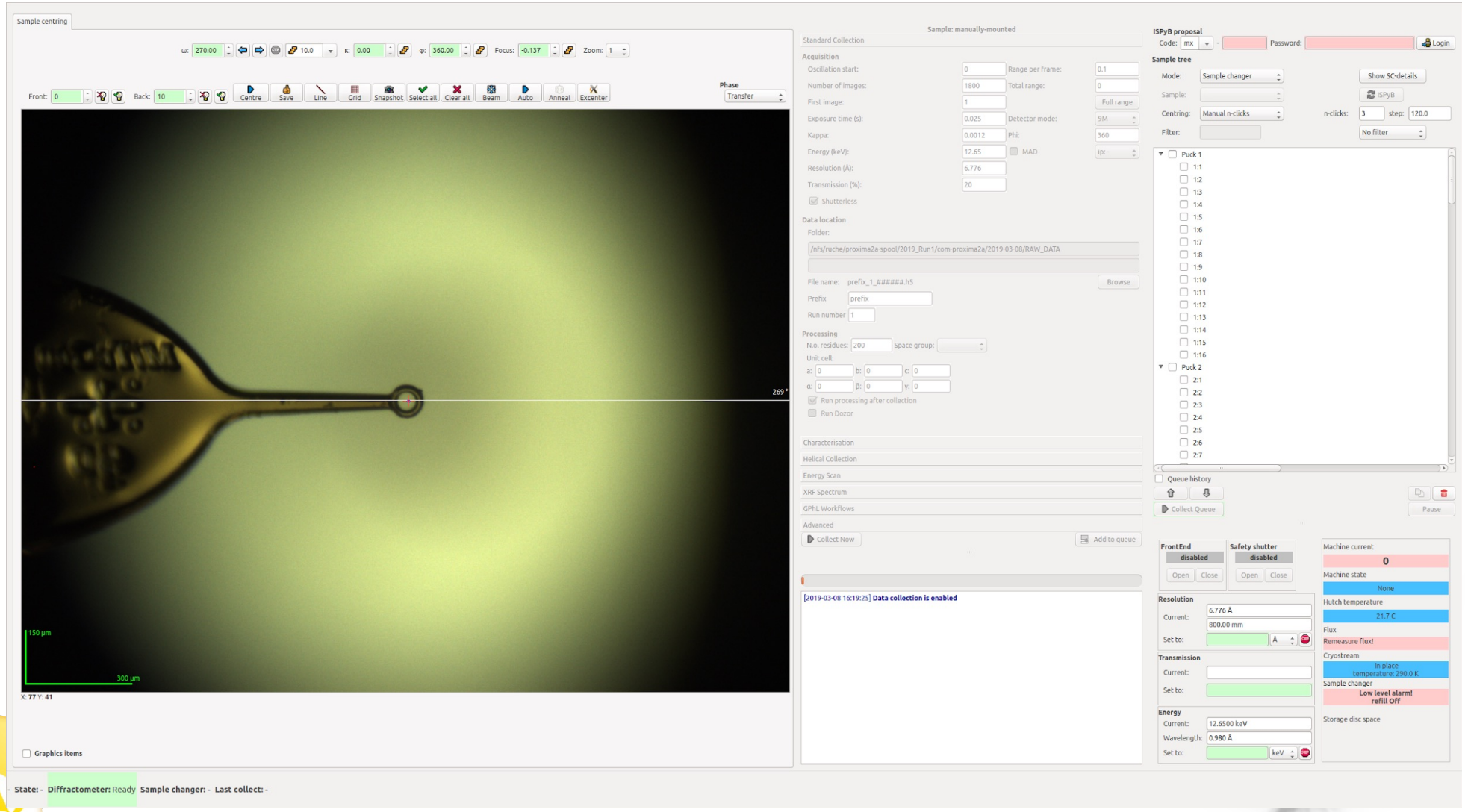
Synchrotron Beamline: PROXIMA 2A



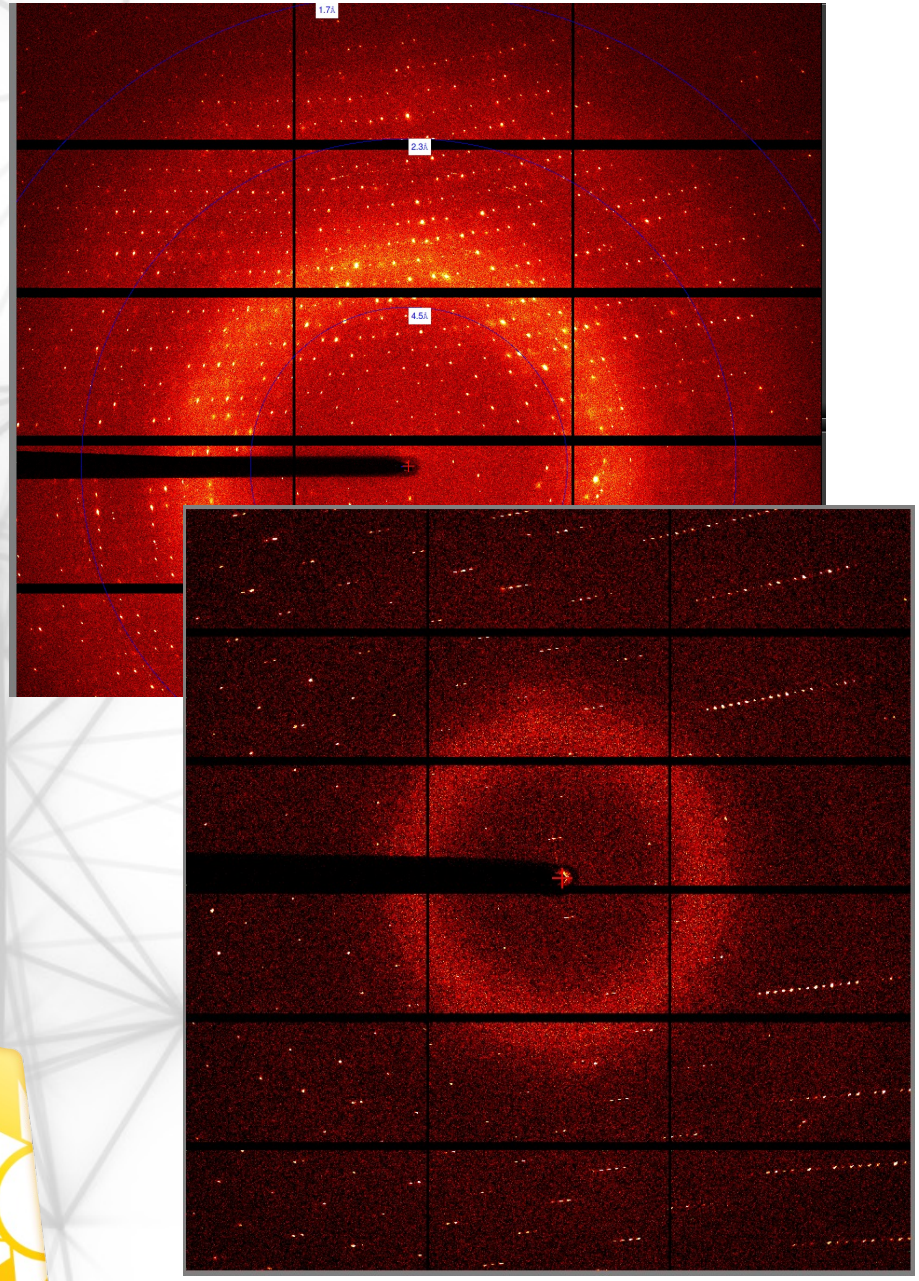


- 10 µm * 5 µm FWHM (H*V)
- Energy range 6 – 18 keV
- X-ray flux ~10^{12} ph/s @ 12.65 keV
- X-ray flux density 2.5x10^{10} ph/s/µm^2
- Beam Stability <1 µm (4-6 hours)

MXCuBE


Sample centring
Centre
Save
Line
Grid
Snapshot
Select all
Clear all
Beam
Auto
Anneal
Excenter
Phase
Transfer
269 °
150 µm
300 µm
X: 77 Y: 41
Graphics items
Sample: manually-mounted
Standard Collection
Acquisition
Data location
Processing
Characterisation
Helical Collection
Energy Scan
XRF Spectrum
GPhL Workflows
Advanced
Collect Now
Add to queue
[2019-03-08 16:19:25] Data collection is enabled
ISPyB proposal
Login
Sample tree
Show SC-details
Puck 1
Puck 2
Queue history
Collect Queue
Pause
FrontEnd
disabled
Safety shutter
disabled
Resolution
Transmission
Energy
Machine current
Machine state
None
Hutch temperature
Flux
Remeasure flux!
Cryostream
Sample changer
Storage disc space
State: - Diffractometer: Ready Sample changer: - Last collect: -

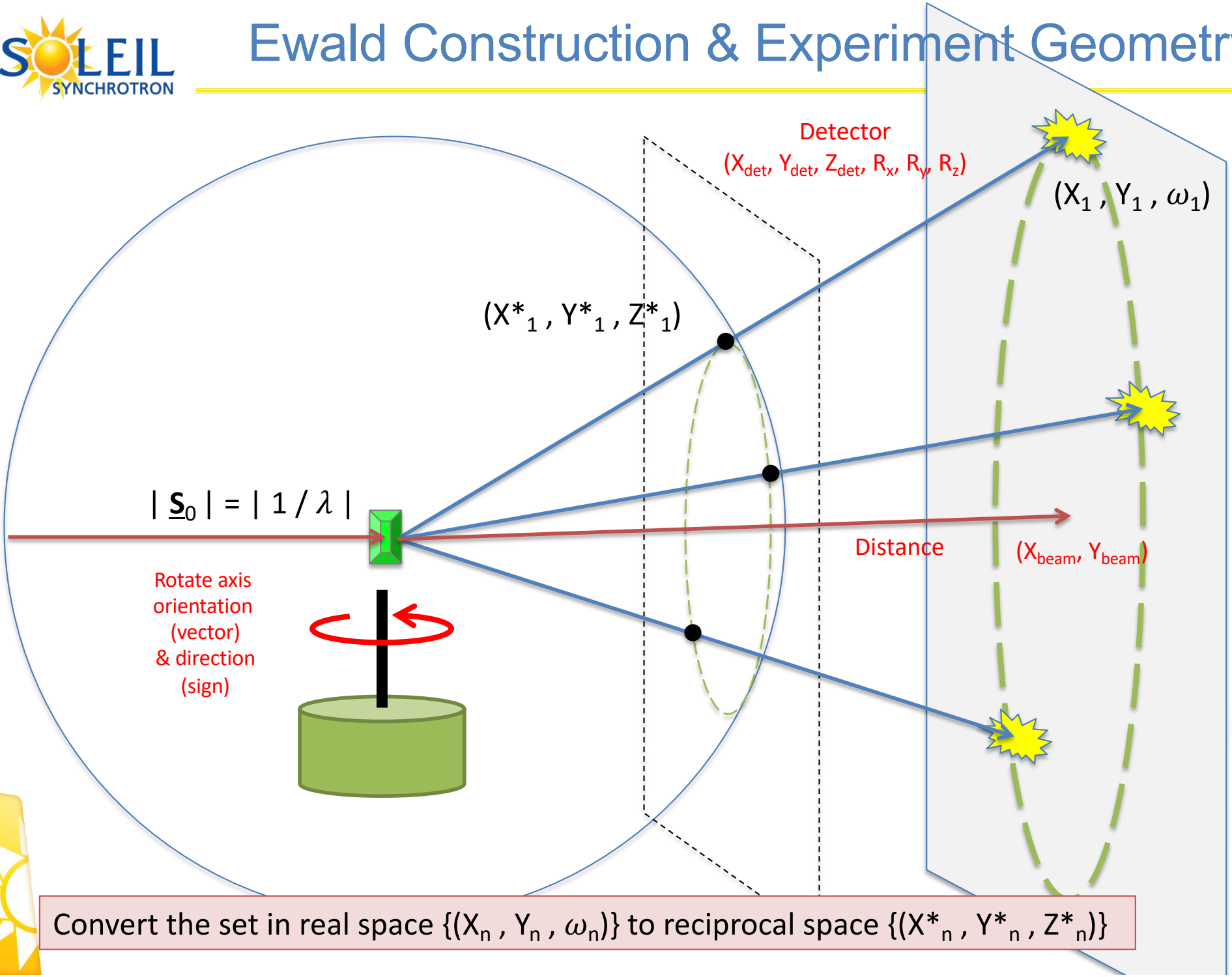
Steps in a Diffraction Experiment



- Align the beam & centre the crystal
  - Align the beam to the rotation axis
  - Center the crystal visually
    - Center with X-rays, if necessary
- Characterise
  - Small wedges of images
    - e.g. 10 × 0.1° at ω = 0°, 90°
  - Check centering
  - Check for diffraction
    - Quality and resolution
    - Inspect summed images ($\Delta\omega$ > 1°)
  - Check auto-indexing results
    - If implemented
- Collect
  - Adjust data collection parameters
    - Resolution or distance
    - Attenuation factor
    - Rotation step size ($\Delta\omega$)
    - Set starting angles (ω_{start})

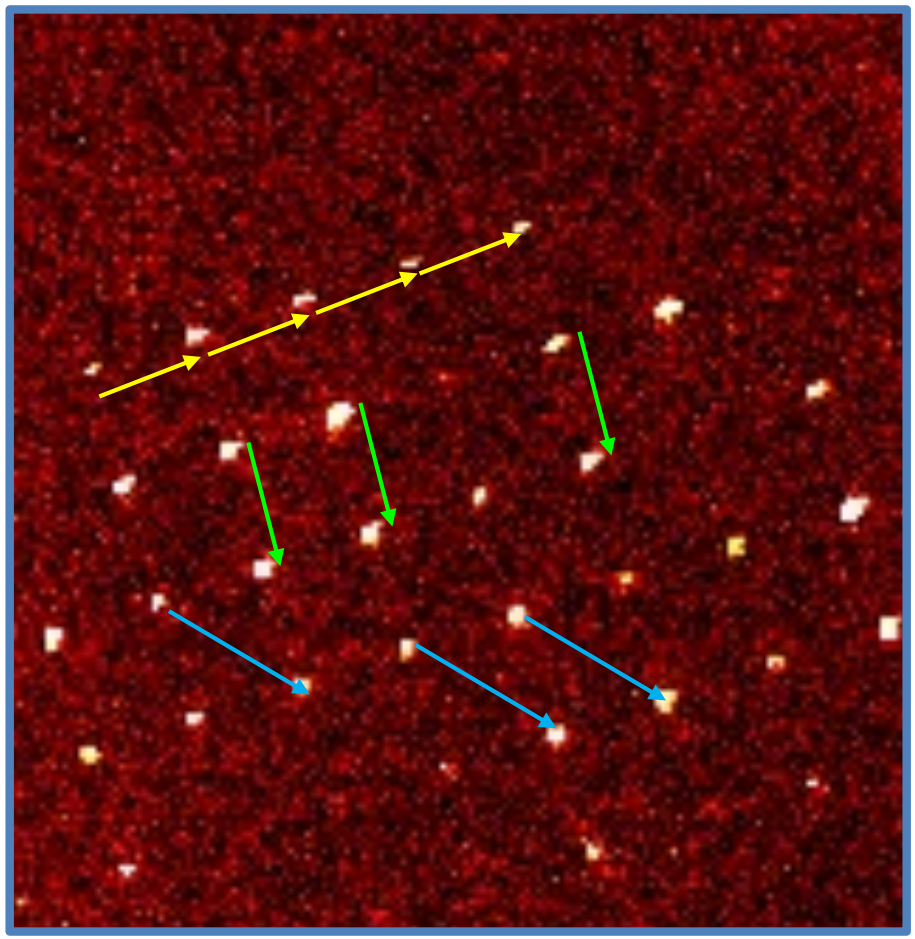
Data Processing

Ewald Construction & Experiment Geometry



Convert the set in real space $\{(X_n, Y_n, \omega_n)\}$ to reciprocal space $\{(X^*_n, Y^*_n, Z^*_n)\}$

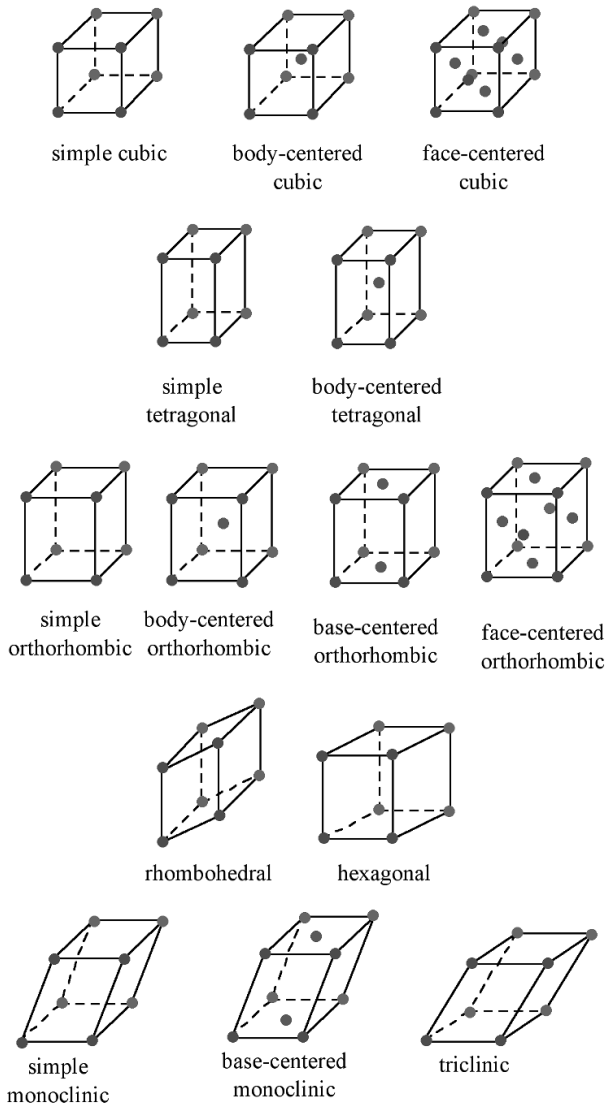
Reciprocal Space Vectors & Indexing



- Indexing
  - Generate all vectors between spots in "Reciprocal Space"
  - Find the "common vectors"
  - Determine a "Reduced Cell"
    - a*, b*, c*, α^*, β^*, γ^*
  - Assign Miller indices (h,k,l) to each spot
  - Transform to all lattices
    - Calculate a penalty

Unit Cells & Bravais Lattices

- 14 Bravais Lattices
- 7 Lattice Systems
  - Trinclinic (a = anorthic)
    - $a \neq b \neq c$, $\alpha \neq \beta \neq \gamma \neq 90°$
  - Monoclinic (m)
    - $a \neq b \neq c$, $\alpha = \gamma = 90°$, $\beta \neq 90°$
  - Orthorhombic (o)
    - $a \neq b \neq c$, $\alpha = \beta = \gamma = 90°$
  - Tetragonal (t)
    - $a = b \neq c$, $\alpha = \beta = \gamma = 90°$
  - Rhombohedral (r)
    - $a = b = c$, $\alpha = \beta = \gamma \neq 90°$
  - Hexagonal (h)
    - $a = b \neq c$, $\alpha = \beta = 90°$, $\gamma = 120°$
  - Cubic (c)
    - $a = b = c$, $\alpha = \beta = \gamma = 90°$
- Centering
  - Primitive (P)
  - Axis centered (A,B,C)
  - Body-centered (I)
  - Face-centered (F)
  - Rhombohedral (R)



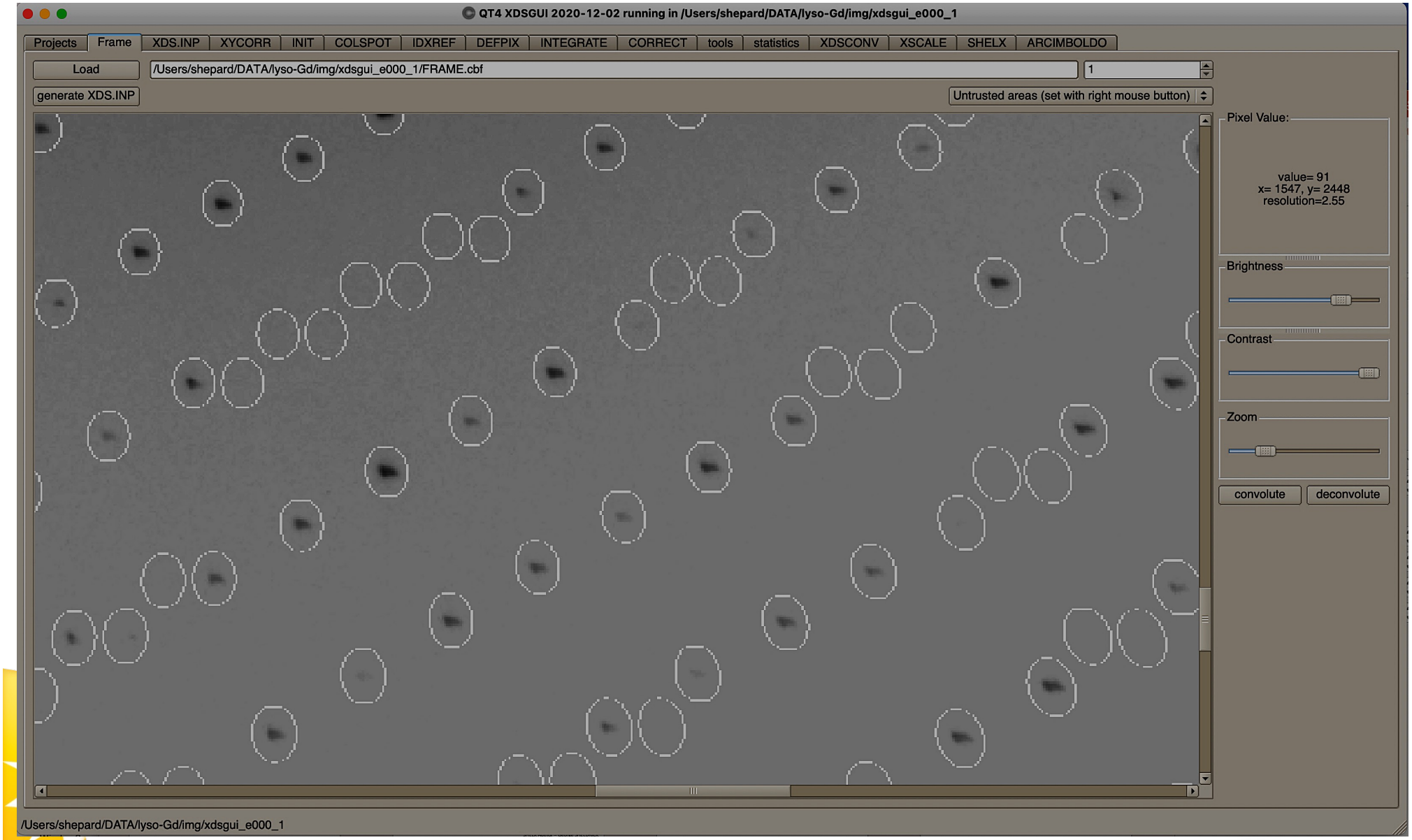
Indexing & Space Groups

LATTICE-CHARACTER	BRAVAIS-LATTICE	QUALITY OF FIT	UNIT CELL CONSTANTS (ANGSTROEM & DEGREES) a	b	c	alpha	beta	gamma
* 44	aP	0.0	98.5	103.5	106.3	90.1	90.0	90.0
* 31	aP	0.0	98.5	103.5	106.3	89.9	90.0	90.0
* 35	mP	0.2	103.5	98.5	106.3	90.0	90.1	90.0
* 34	mP	0.6	98.5	106.3	103.5	90.1	90.0	90.0
* 33	mP	0.7	98.5	103.5	106.3	90.1	90.0	90.0
* 32	oP	0.8	98.5	103.5	106.3	90.1	90.0	90.0
* 25	mC	30.7	148.3	148.4	98.5	90.0	90.0	88.5
* 23	oC	30.8	148.3	148.4				
* 20	mC	30.8	148.4	148.3				
* 21	tP	31.4	103.5	106.3				
* 14	mC	52.1	142.9	142.9				
* 13	oC	52.2	142.9	142.9				
* 10	mC	52.2	142.9	142.9				
* 11	tP	52.3	98.5	103.5	106.3	90.1	90.0	90.0
4	hR	82.8	142.9	145.0	178.0	93.5	87.8	117.9
2	hR	83.0	142.9	145.0	178.2	93.6	87.7	118.0
3	cP	83.4	98.5	103.5	106.3	90.1	90.0	90.0
5	cI	248.9	144.9	142.9	148.3	59.7	58.3	62.1
39	mC	249.9	229.2	98.5	106.3	90.0	90.1	64.6
37	mC	250.1	234.3	98.5	103.5	90.0	90.1	65.1
38	oC	250.5	98.5	229.2	106.3	89.9		
29	mC	250.5	98.5	229.2	106.3	89.9		
28	mC	250.6	98.5	234.4	103.5	90.0		
36	oC	250.7	98.5	234.3	103.5	89.9		
41	mC	275.4	236.4	103.5	98.5	90.0		
30	mC	275.4	103.5	236.4	98.5	90.0	90.0	64.1
40	oC	275.4	103.5	236.4	98.5	90.0	90.0	115.9

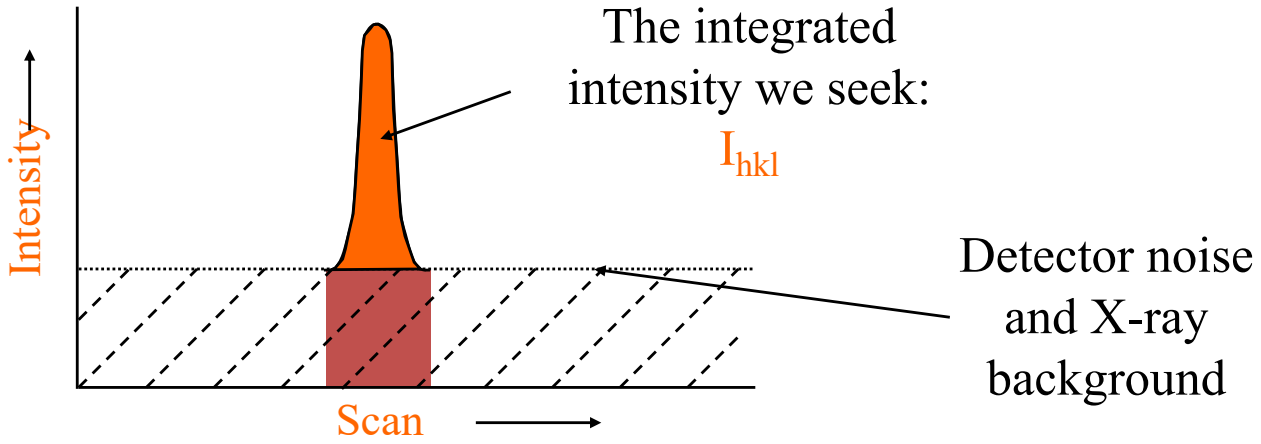
Possible solutions
- All equally likely
- Usually we select the highest symmetry, but BEWARE the correct cell could have lower symmetry

Possible polymorphism?
- Proteins can crystallise in more than one lattice and space group

Display Predictions



Integration



85	90	73	65	74	78	76	69	69	50	57
76	78	71	66	75	83	75	59	78	71	61
75	88	81	69	86	123	130	88	84	76	74
70	81	81	85	188	546	694	147	79	73	80
95	94	90	86	159	408	508	143	91	69	63
82	81	80	83	107	153	163	102	88	73	62
77	77	75	79	83	84	85	78	76	79	83
80	78	80	77	74	74	82	76	76	71	83
79	77	73	74	69	67	80	79	69	64	87

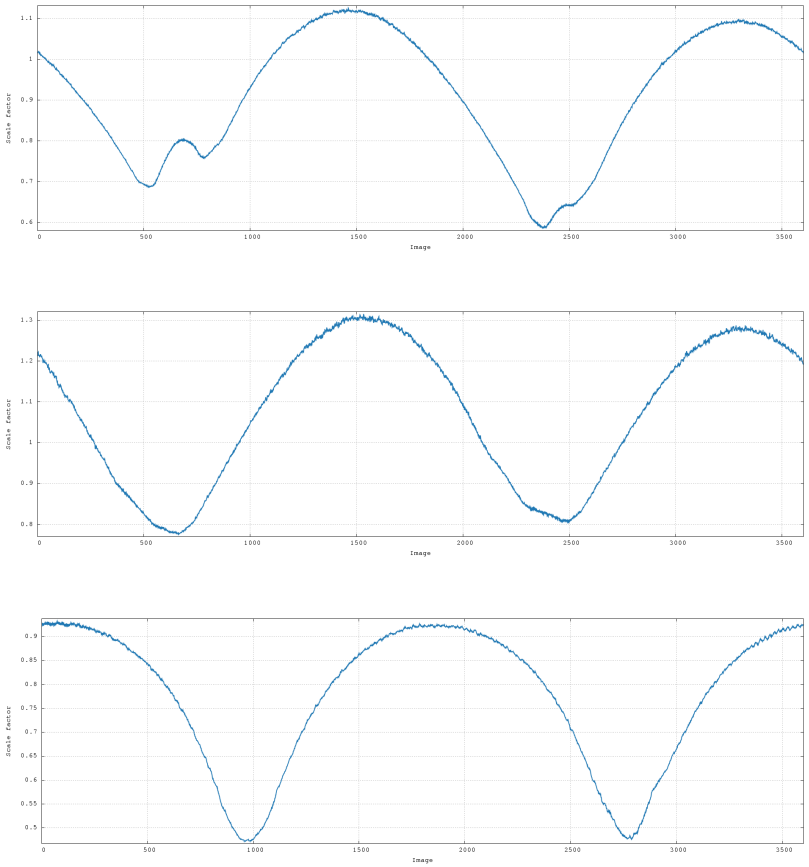
Data Quality Table

- **Resolution limits**
  - Shells of resolution
  - Determine high limit cut-off
- **Completeness**
  - Should be close 100%
  - Typically less in low and high resolution shells
- **R-factors**
  - Residual factor
- **CC(1/2)**
  - Pearson's Correlation Coefficient
- **I / sigma**
  - Signal-to-noise ratio
- **Anomalous**
  - Differrences between `I(hkl)` and `I(-h,-k,-l)`
  - Pearson's Correlation Coefficient
  - Signal-to-noise ratio
  - R-factor (R-anom, but not shown)
- **Multiplicity (not shown)**
  - Number observed / Number unique
  - Very important

SUBSET OF INTENSITY DATA WITH SIGNAL/NOISE >= -3.0 AS FUNCTION OF RESOLUTION

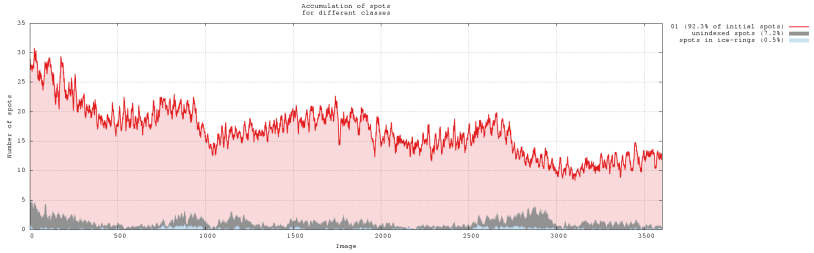
RESOLUTION LIMIT	NUMBER OF REFLECTIONS OBSERVED	UNIQUE	POSSIBLE	COMPLETENESS OF DATA	R-FACTOR observed	R-FACTOR expected	COMPARED	I/SIGMA	R-meas	CC(1/2)	Anomal Corr	SigAno	Nano
5.17	10113	848	890	95.3%	3.3%	4.5%	10104	54.84	3.5%	99.9*	100*	19.960	309
3.67	19342	1587	1588	99.9%	3.7%	4.6%	19340	55.66	3.9%	99.9*	99*	15.770	677
3.00	24467	2016	2016	100.0%	4.2%	4.6%	24467	53.71	4.4%	99.9*	99*	14.968	891
2.60	28113	2383	2383	100.0%	4.7%	4.8%	28113	48.64	5.0%	99.9*	99*	15.249	1077
2.33	31467	2727	2727	100.0%	5.2%	5.0%	31467	43.92	5.4%	99.9*	99*	13.790	1245
2.13	34446	2999	2999	100.0%	5.4%	5.4%	34446	40.18	5.7%	99.9*	99*	12.101	1381
1.97	37481	3257	3257	100.0%	6.2%	6.3%	37481	33.94	6.5%	99.8*	99*	10.461	1514
1.84	36454	3478	3489	99.7%	8.3%	8.8%	36453	23.78	8.8%	99.8*	98*	7.638	1626
1.74	18105	3400	3771	90.2%	12.0%	14.2%	17860	10.35	13.2%	98.6*	94*	3.698	1417
total	239988	22695	23120	98.2%	4.5%	5.0%	239731	36.47	4.7%	99.9*	99*	11.244	10137

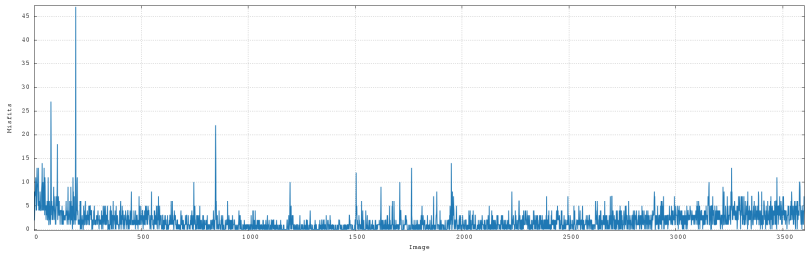
Scaling Diffraction Data

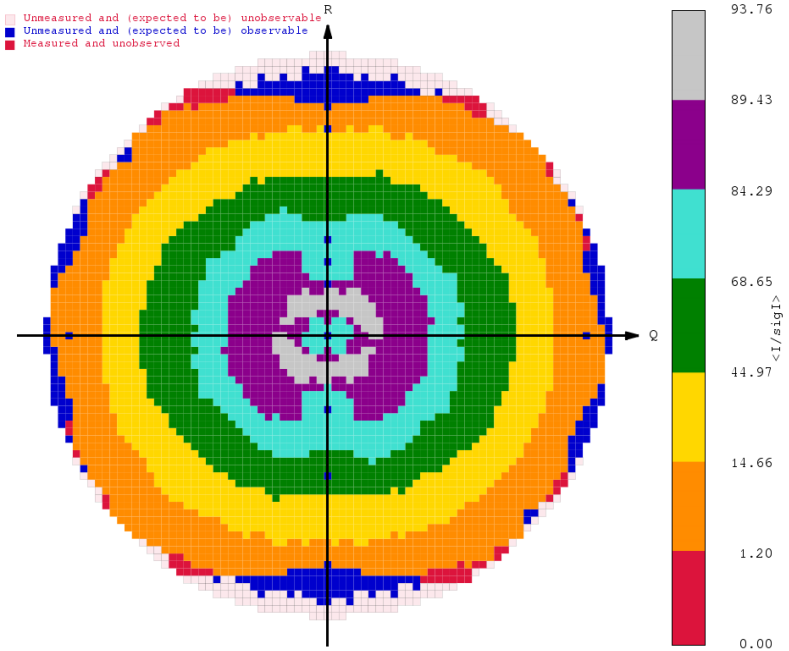


- Scale equivalent reflections
  - `I(h,k,l) = I(-h,-k,-l) =` etc…
  - Depends upon crystal's "point group"
- Corrections for:
  - Incident beam variations
  - Changes in illuminated volumes
  - Detector inhomogeneities
  - Effects of absorption
  - Lorentz and polarisation effects
- Over-scaling is possible….
  - Squashing of anomalous signal
    - e.g. radiation damage

Automated Data Processing Software







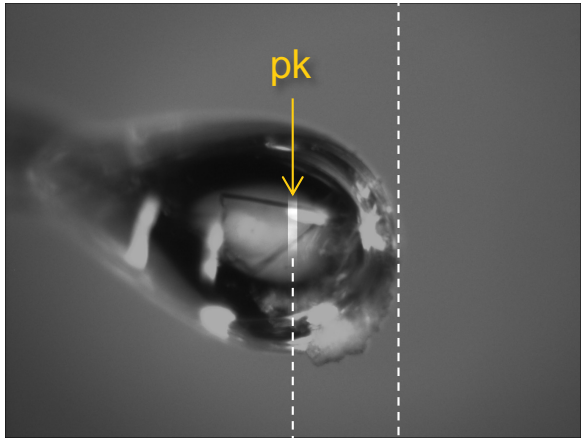
- Common processing programs
  - XDS = X-ray Detector Systems
    - Wolfgang Kabsch & Kay Dietrichs
  - MOSFLM
    - Andrew Leslie & Harry Powel
  - DENZO-HKL2000
    - Z. Otwinowski & W. Minor
  - DIALS
    - DIAMOND Light Source
- Pipelines
  - AUTOPROC (Global Phasing Ltd)
    - XDS
  - XDSME (SOLEIL)
    - XDS Made Easy
    - XDS & CCP4
  - XIA2 (DLS)
    - DIALS or XDS
  - XDSGUI
    - XDS

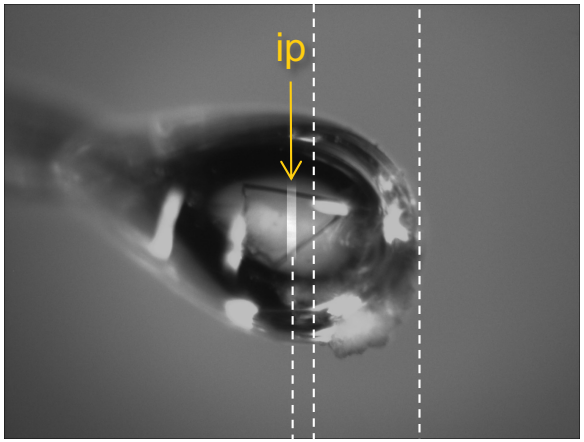
Data File Formats

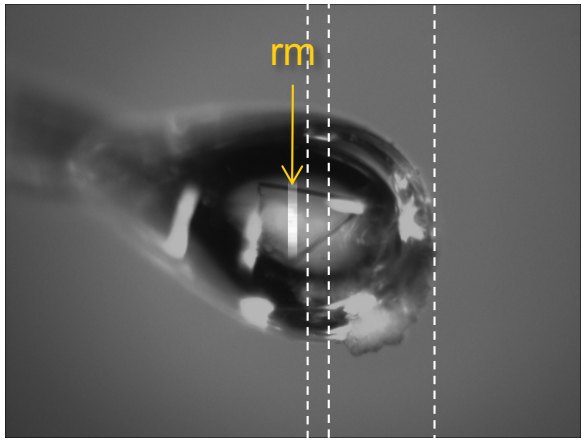
- Data formats
  - XDS (ascii)
    - XDS_ASCII.HKL
    - Header and colums
    - H,K,L,I,IOBS,XDET,YDET,PHI,…
  - CCP4 MTZ (binary with columns)
    - Project.mtz
    - Columns = H,K,L,I/SYMI.FOBS,SIGFOBS,…
  - DENZO SCA (ascii)
    - Project.sca
  - DIALS REFL
    - MessagePack with columns
    - Integrate.refl
  - SHELX (ascii)
    - H,K,L,F

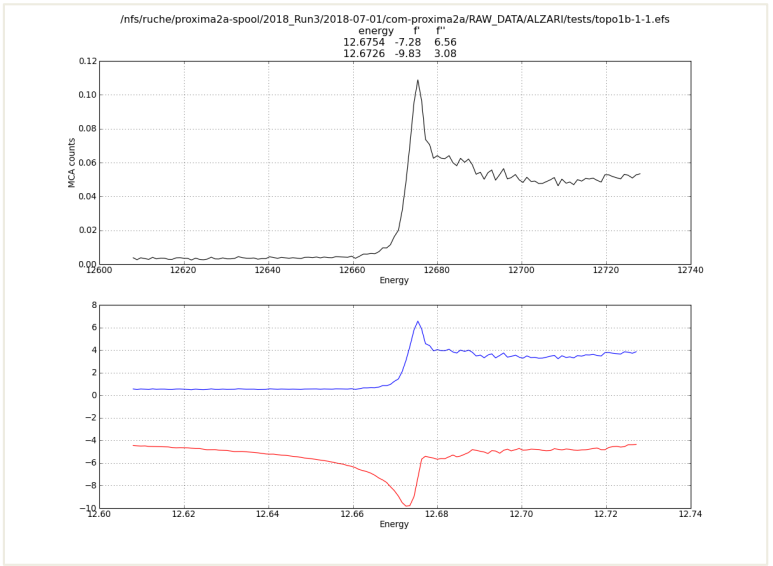
3D Maps & Examples

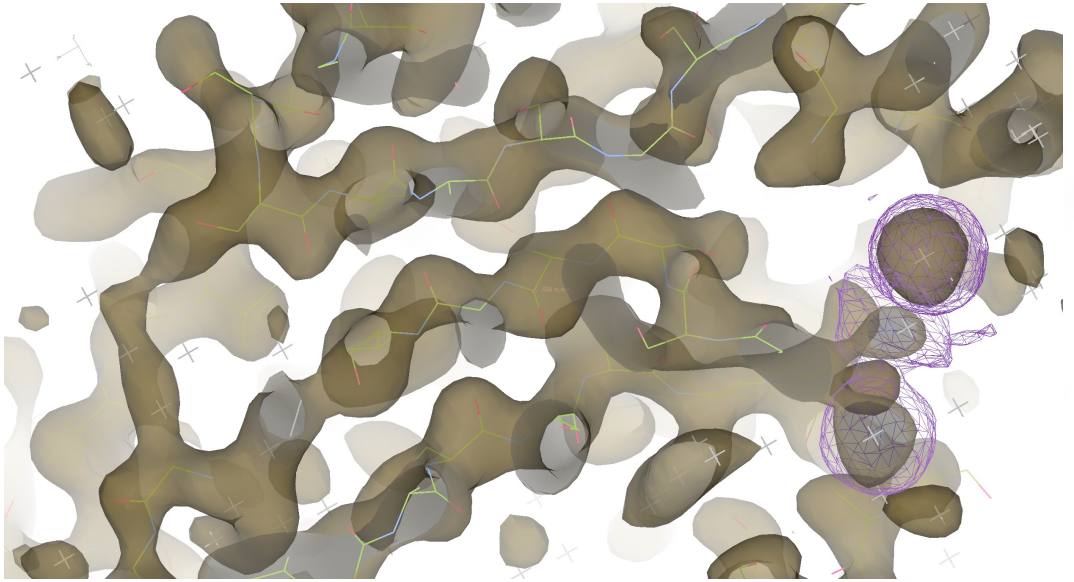
Multi-point MAD



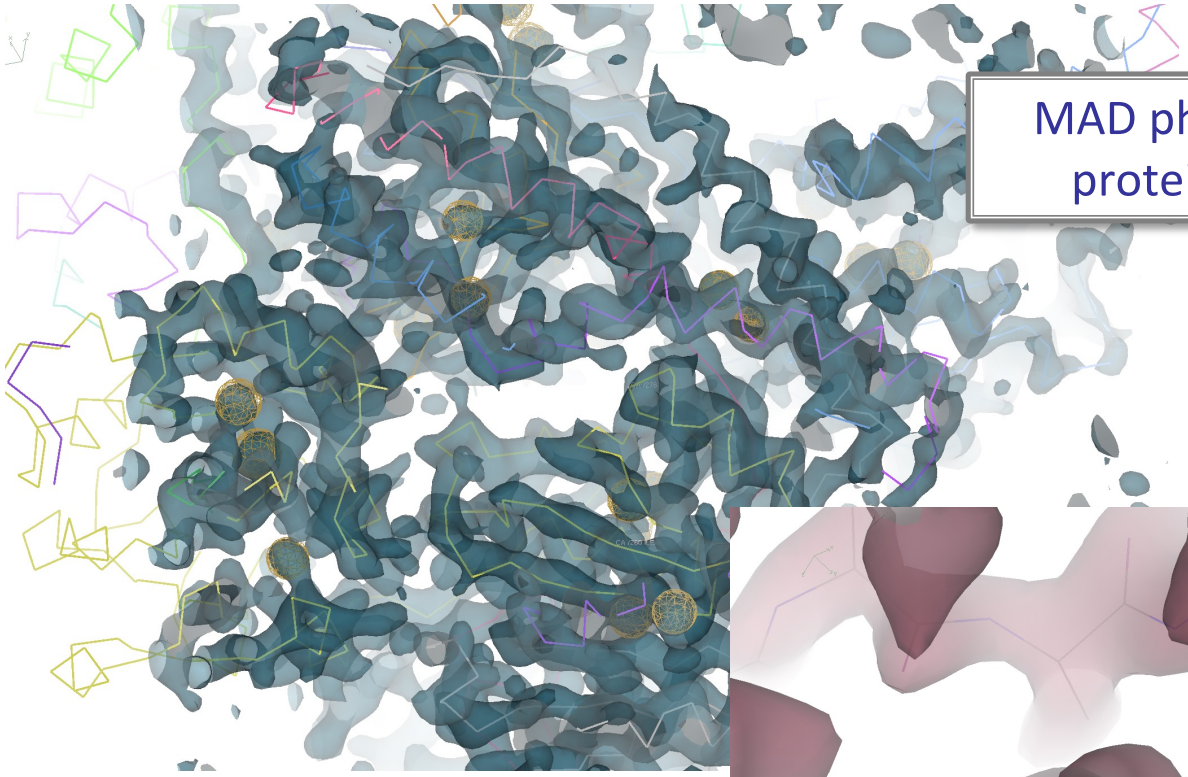




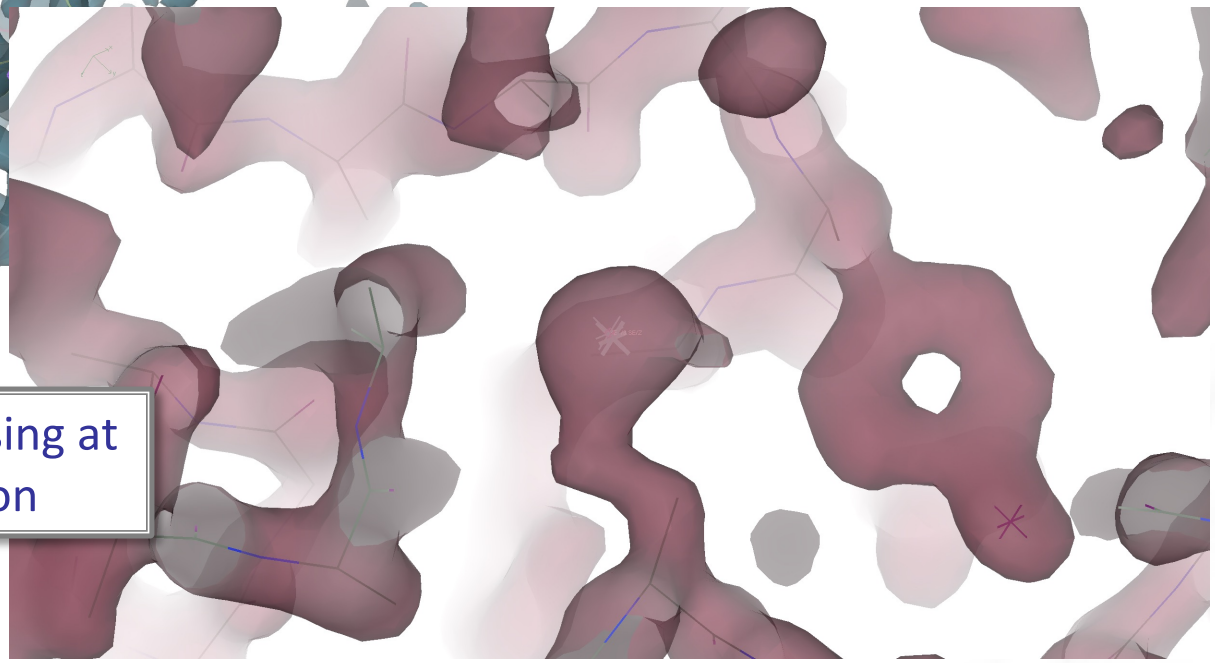




Other multi-point MAD structures

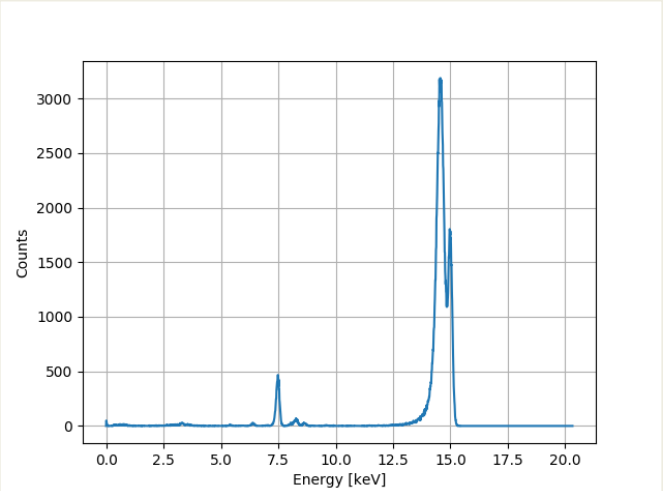


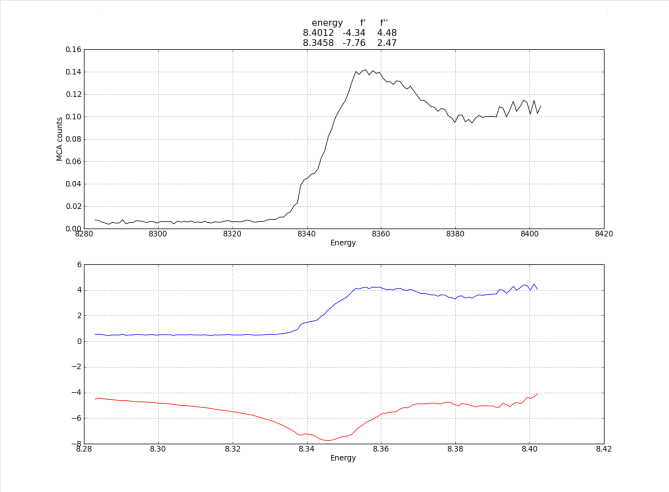
MAD phasing of a 1400 aa protein with 48 SeMet

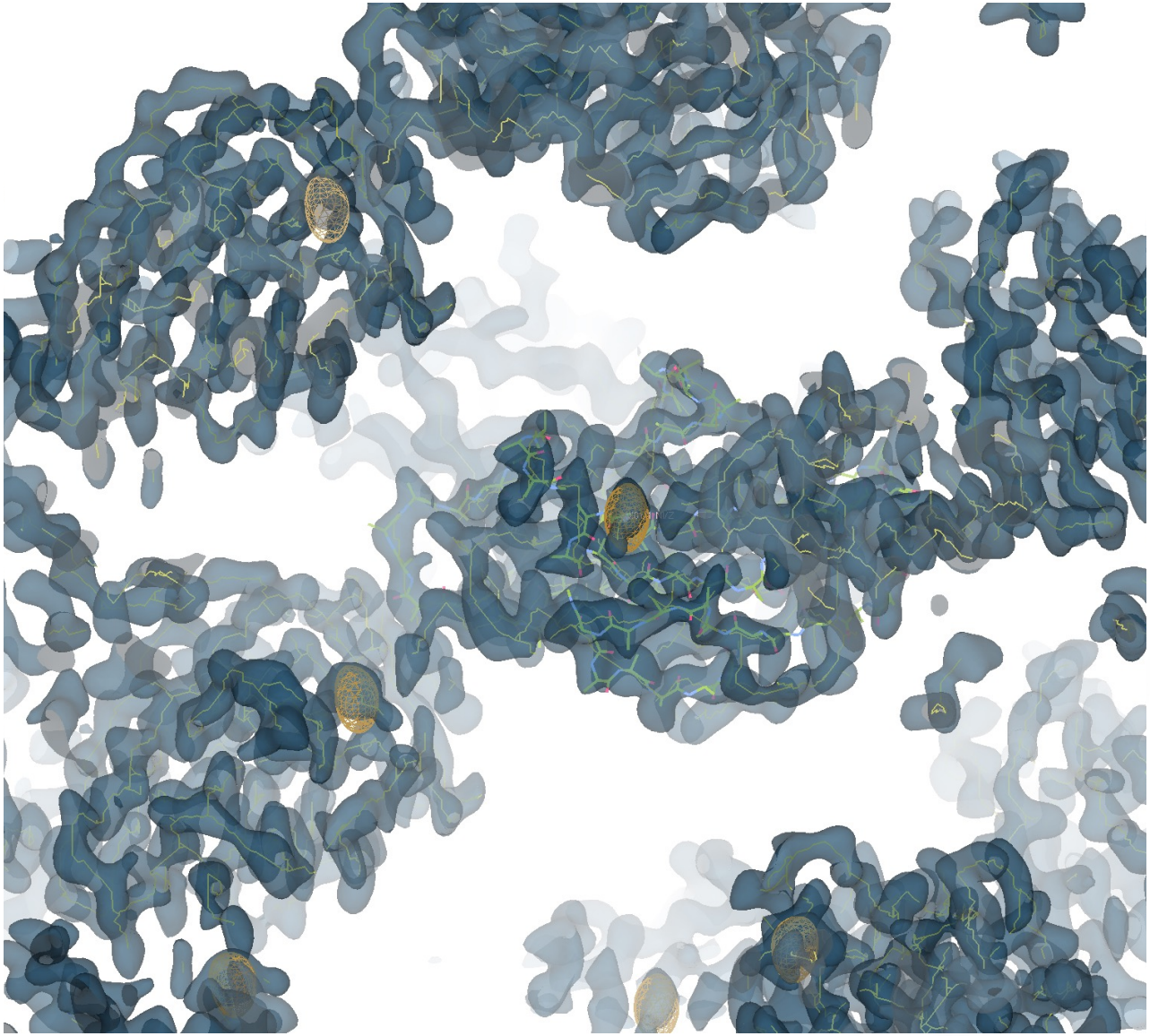


SeMet MAD phasing at 1.5Å resolution

Unknown Heavy Atoms: Finding a Nickel

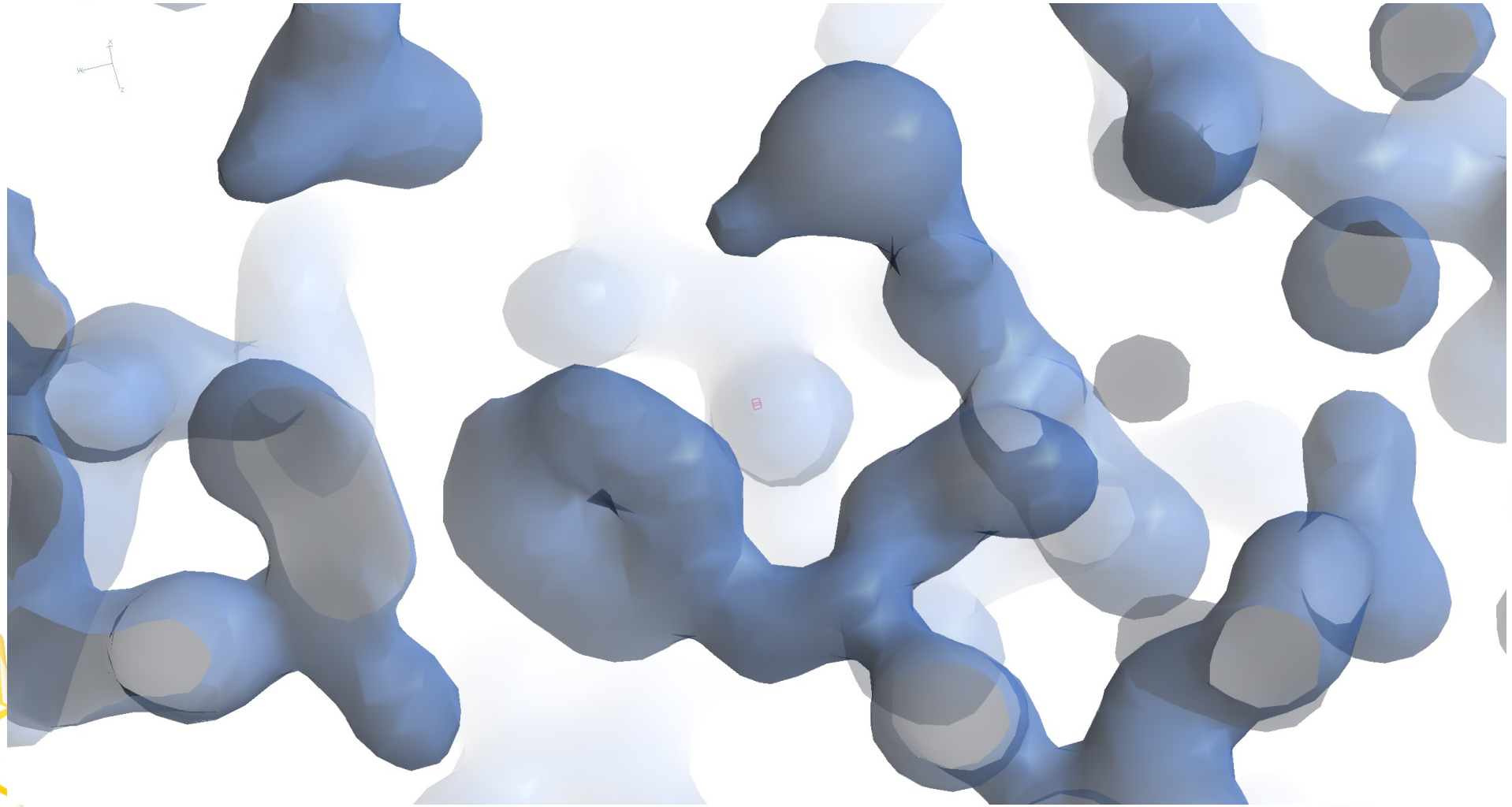




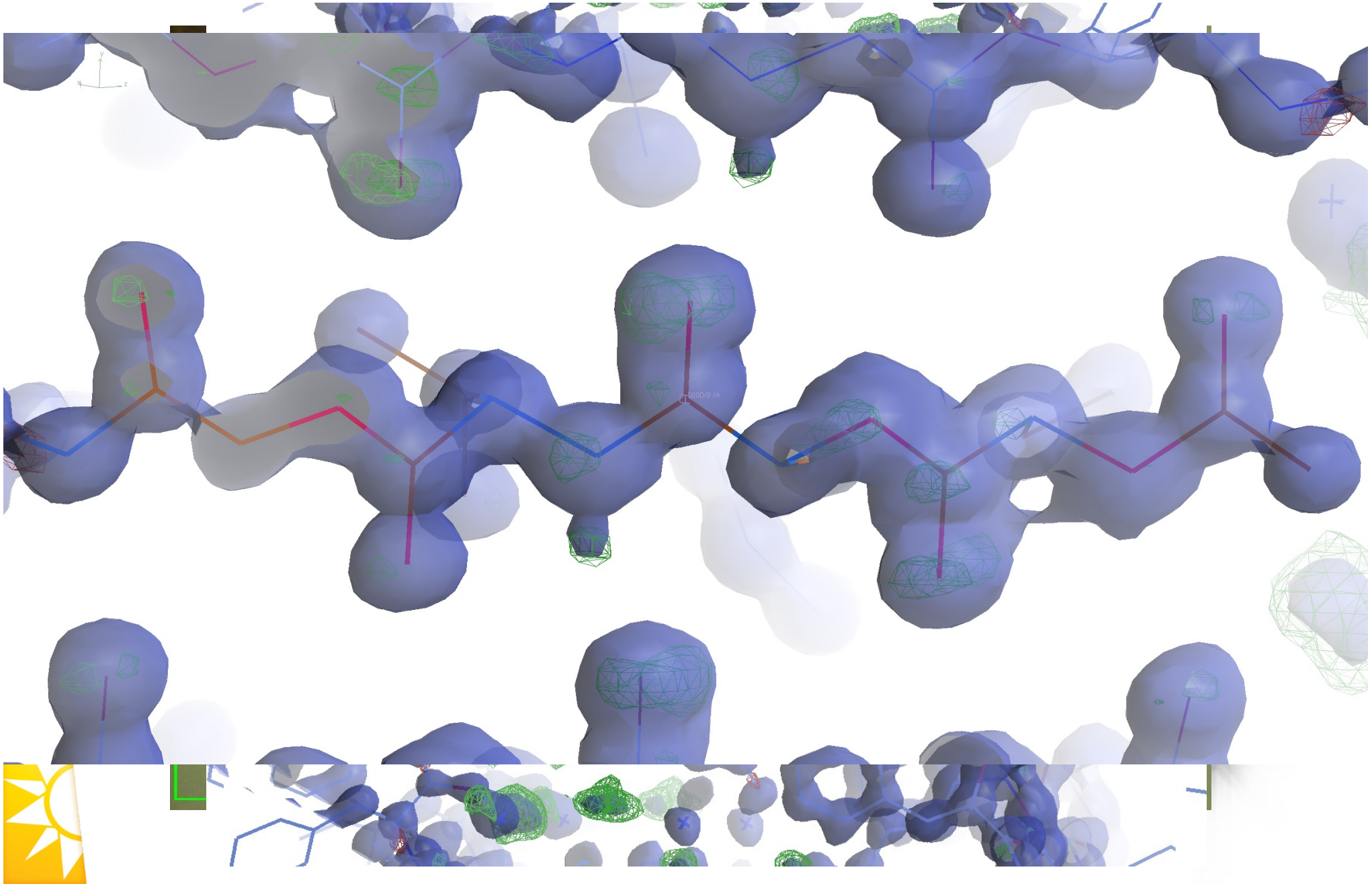


Helical Sulphur SAD

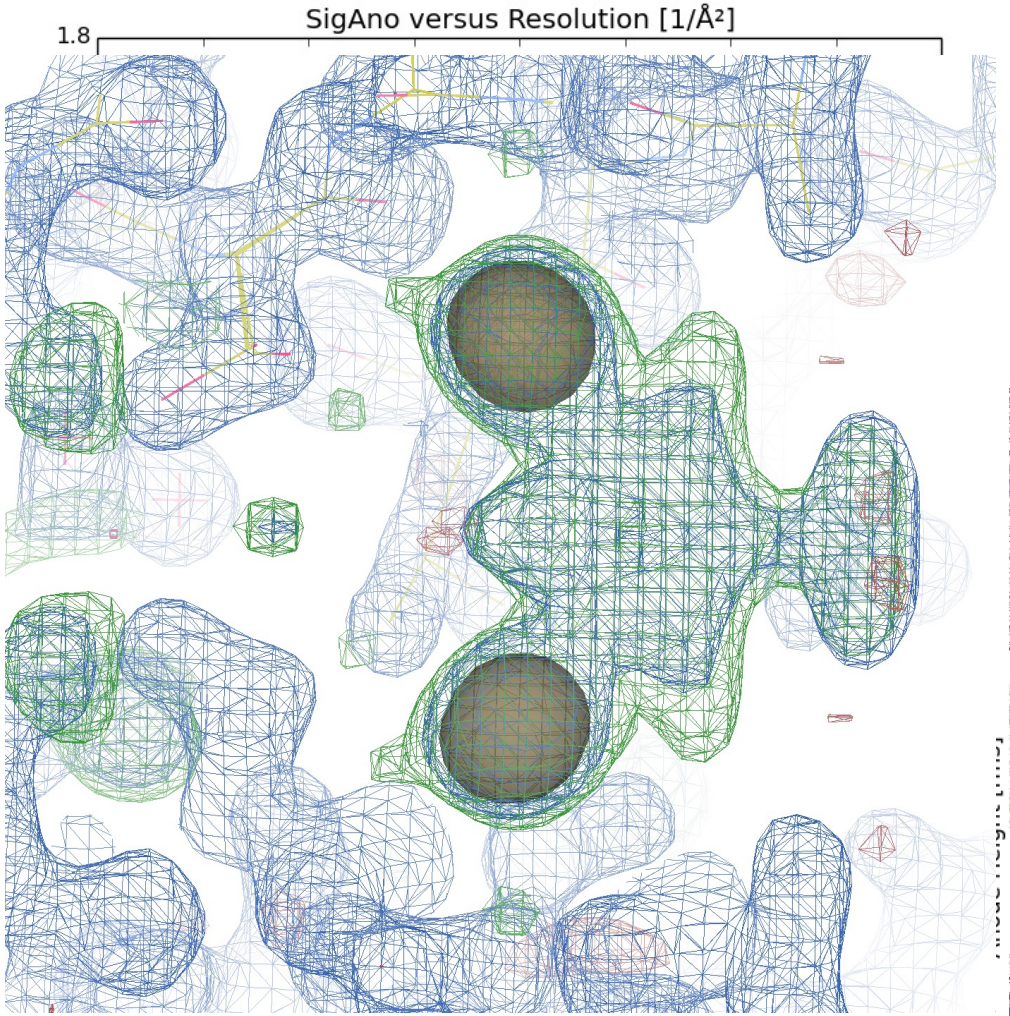
- Unknown protein crystal
  - Diffracts to 1.4 Å resolution

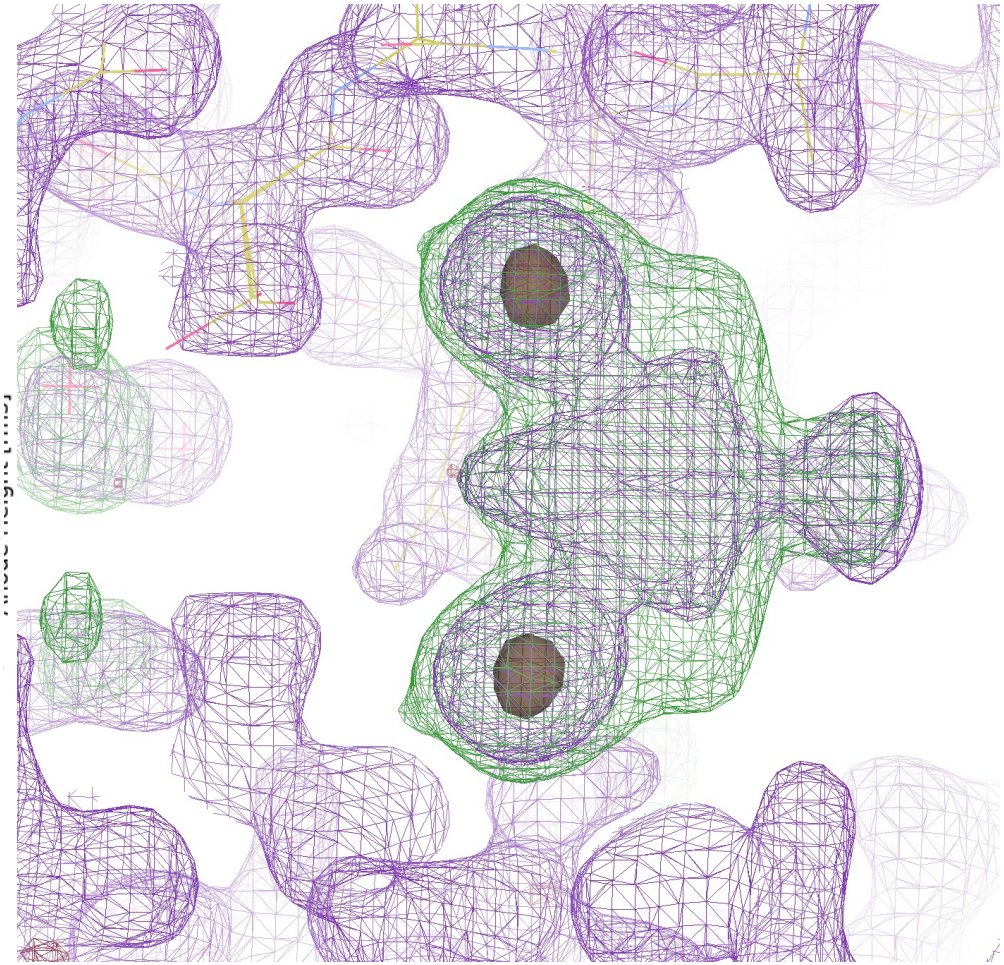


The Ultimate Helical Scan



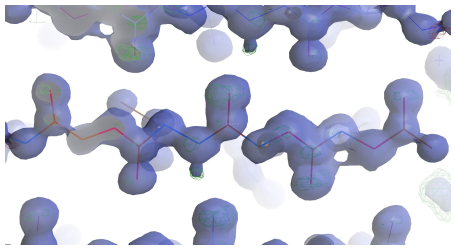
Radiation Damage & Heavy Atoms

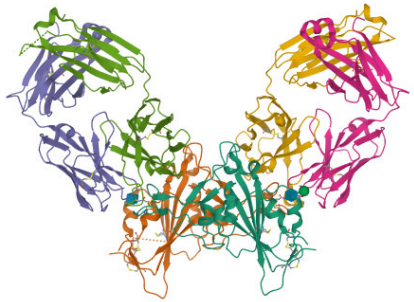


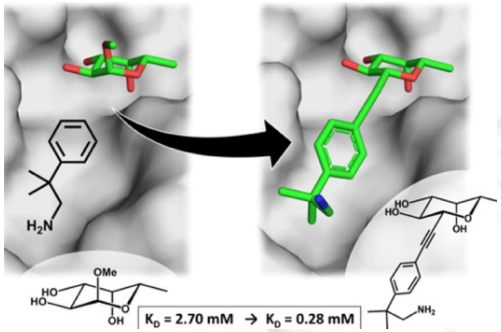


Conclusions

Crystallography after 100 years







- Structural Chemistry
  - CCDC crystal structures
    - > 1,300,000 entries total
    - > 18,000 entries added annually
- Structural Biology
  - PDB X-ray crystal structures
    - >180,000 entries total
    - >9k entries added annually
- 48 Nobel Laureates
  - William Henry Bragg & William Lawrence Bragg (1915)
    - For their services in the analysis of crystal structure by means of X-rays
  - David Julius & Ardem Patapoutian (2021)
    - For the discoveries of receptors for temperature and touch
- Industrial applications
  - Pharmaceutical
  - Drug Discovery

Acknowledgements

Resources & Acknowledgements

- Websites
  - http://www.lks.physik.uni-erlangen.de/diffraction/teaching.html
  - http://www.yorvic.york.ac.uk/~cowtan/fourier/fourier.html
  - http://ccp4wiki.org/~ccp4wiki/wiki/index.php?title=Hendrickson_Lattman_coefficients
- Dauter (1999) Acta Cryst. **D55**, 1703 -17
- Blundell & Johnson (1976)
- Drenth (1999)
- Als-Nielsen & McMorrow (2001)
- Blow (2002)
- Rhodes (2006)
- Rupp (2009)
- Read & McCoy (2010)
- Sherwood & Cooper (2010)
- International IUCr Volumes
- Numerous colleagues!
  - Especially: Richard Kahn, Roger Fourme,…

Questions

Email: william.shepard@synchrotron-soleil.fr