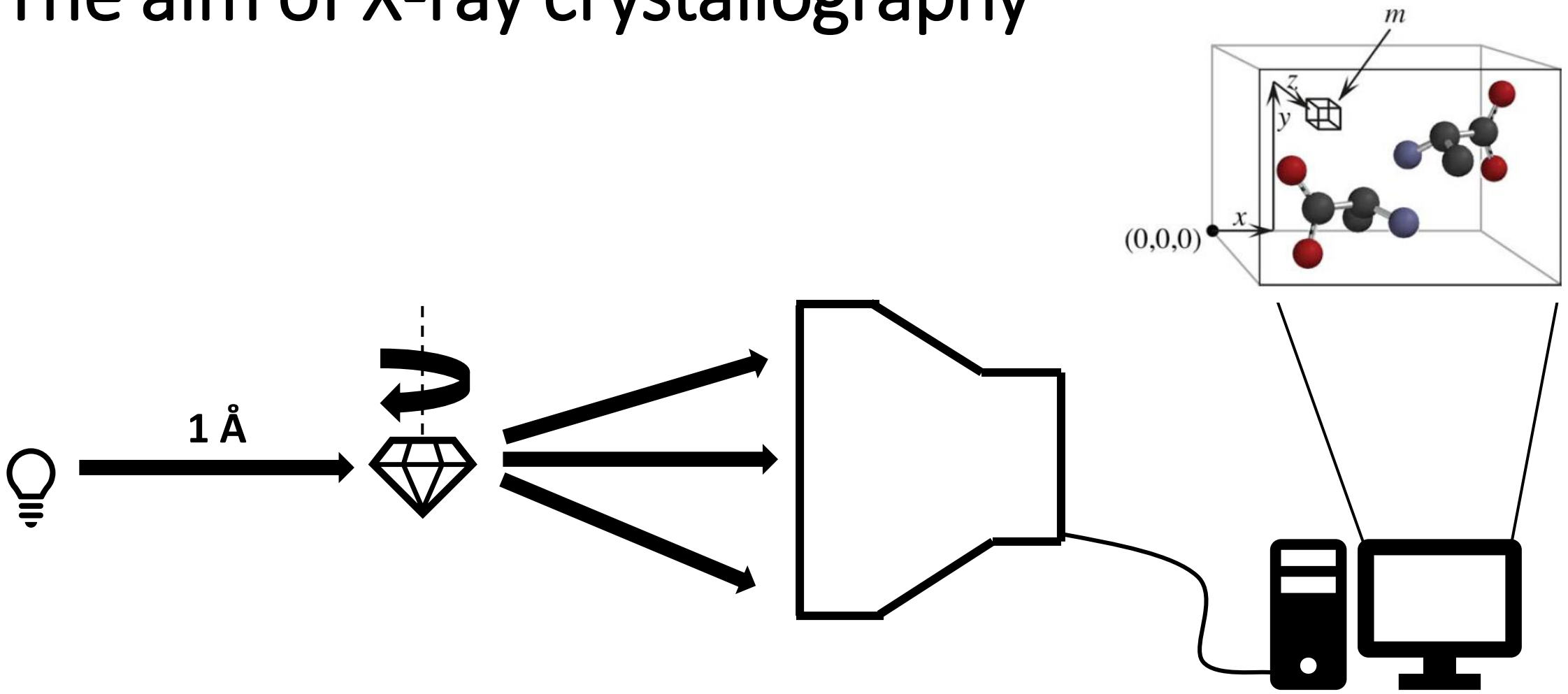
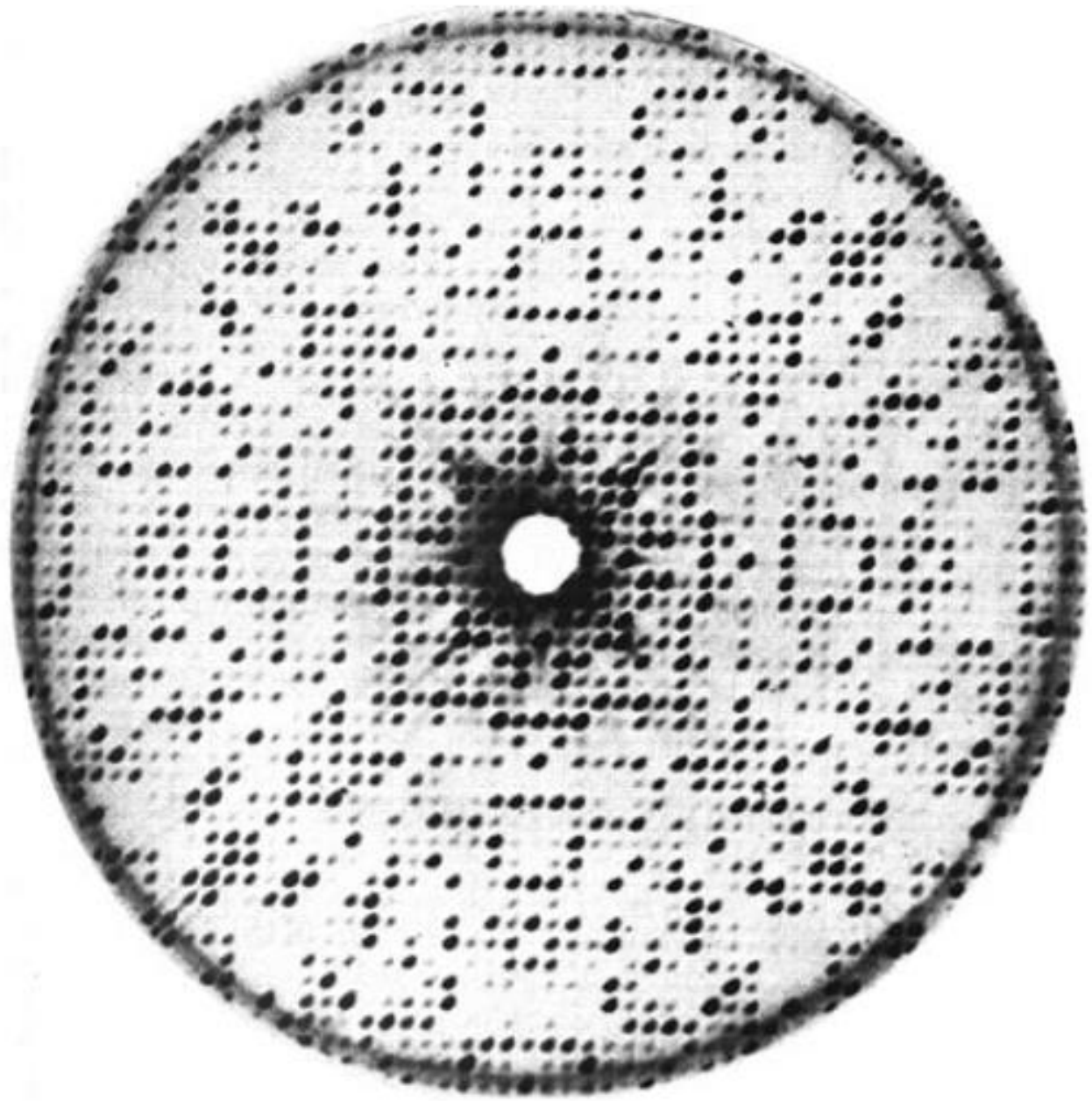
From crystal to structure

Filip Melicher

S1004 Methods for structural characterization of biomolecules

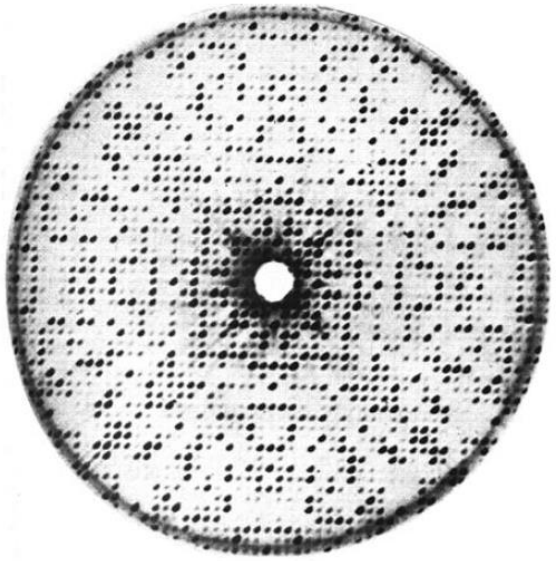
The aim of X-ray crystallography





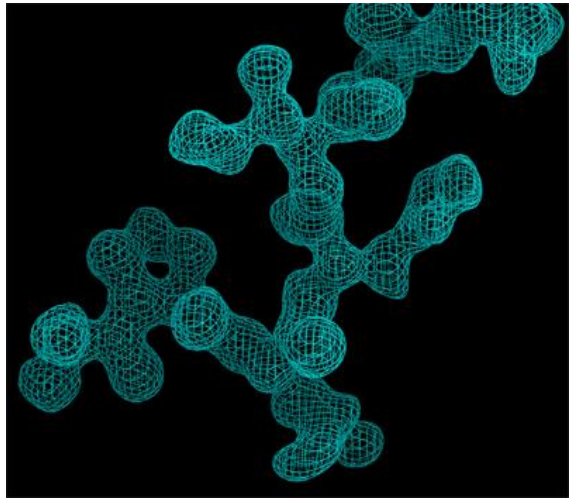
Workflow

Step 1



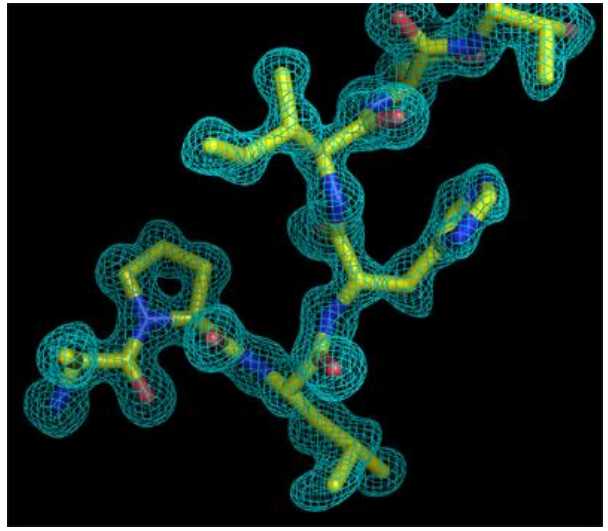
Diffraction data

Step 2

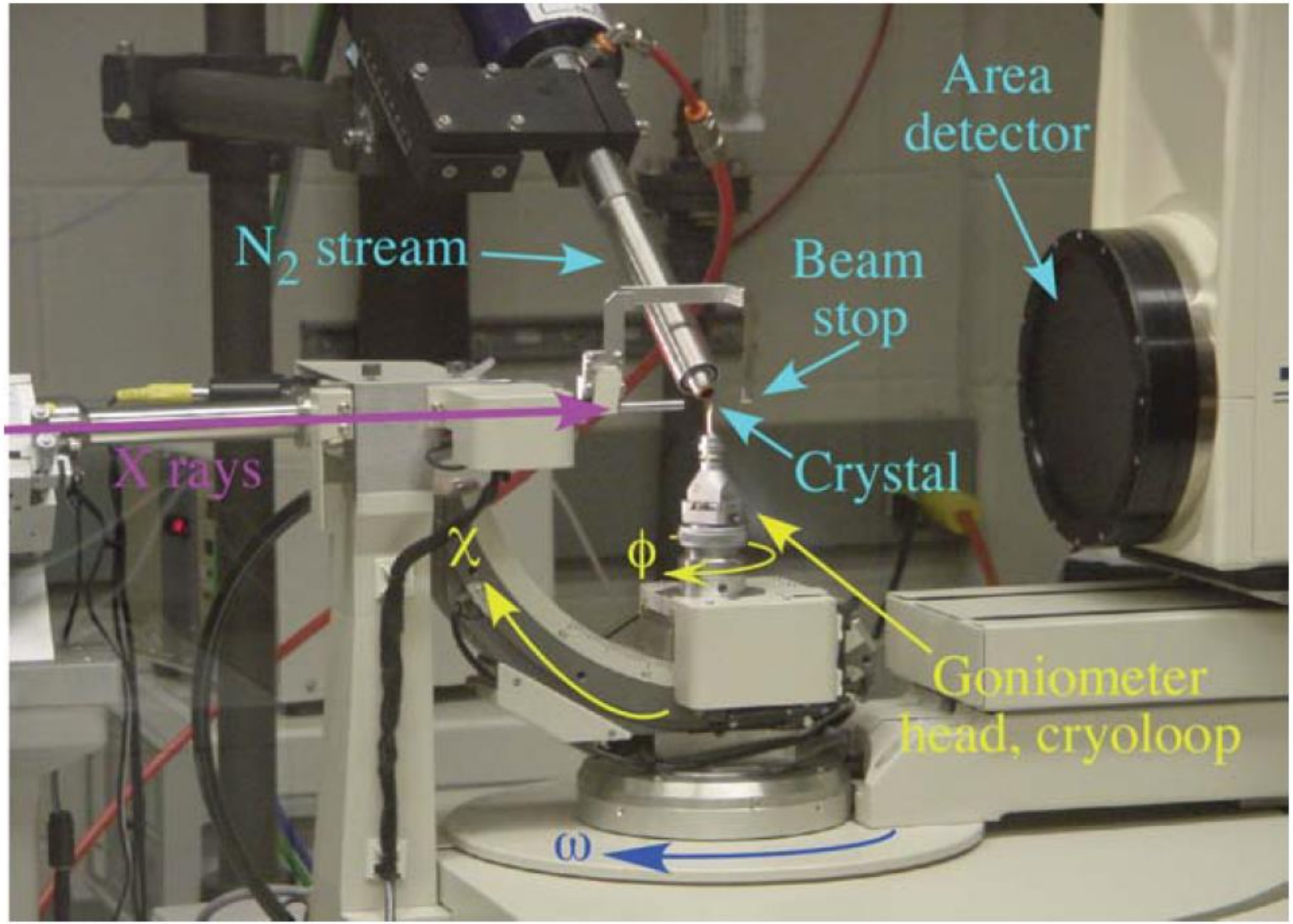


Electron density calculation

Step 3



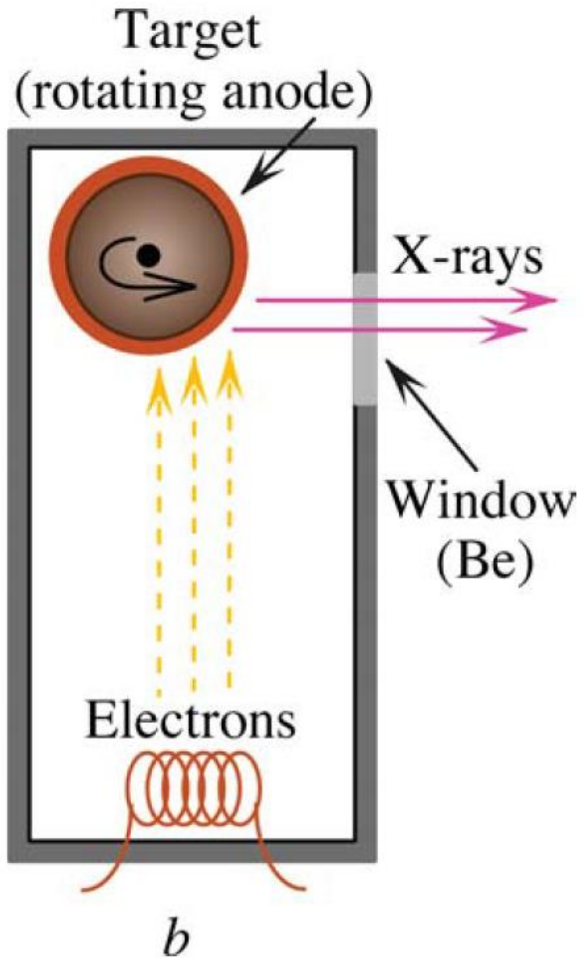
Electron density interpretation


Area
detector
N_2 stream
Beam
stop
X rays
Crystal
χ
ϕ
Goniometer
head, cryoloop
ω

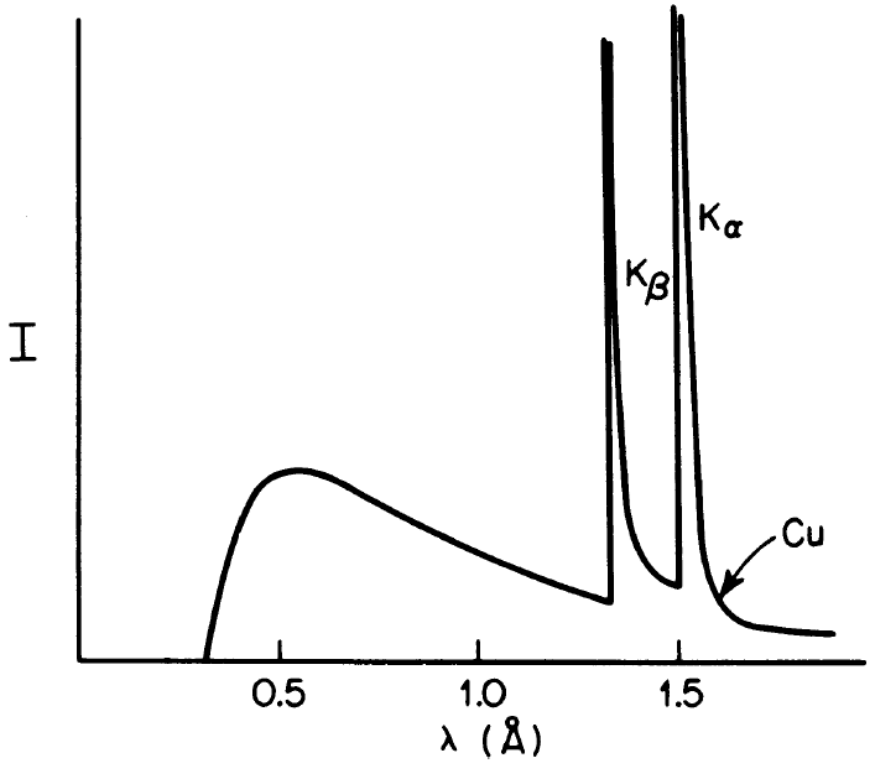
In-house

Anode: Cu, Mo, Cr, W

Transition of electrons from L, M levels to K level



0.1 nm = 1 Å



Synchrotrons

Monochromatic coherent waves

Multiple wavelengths

High intensity

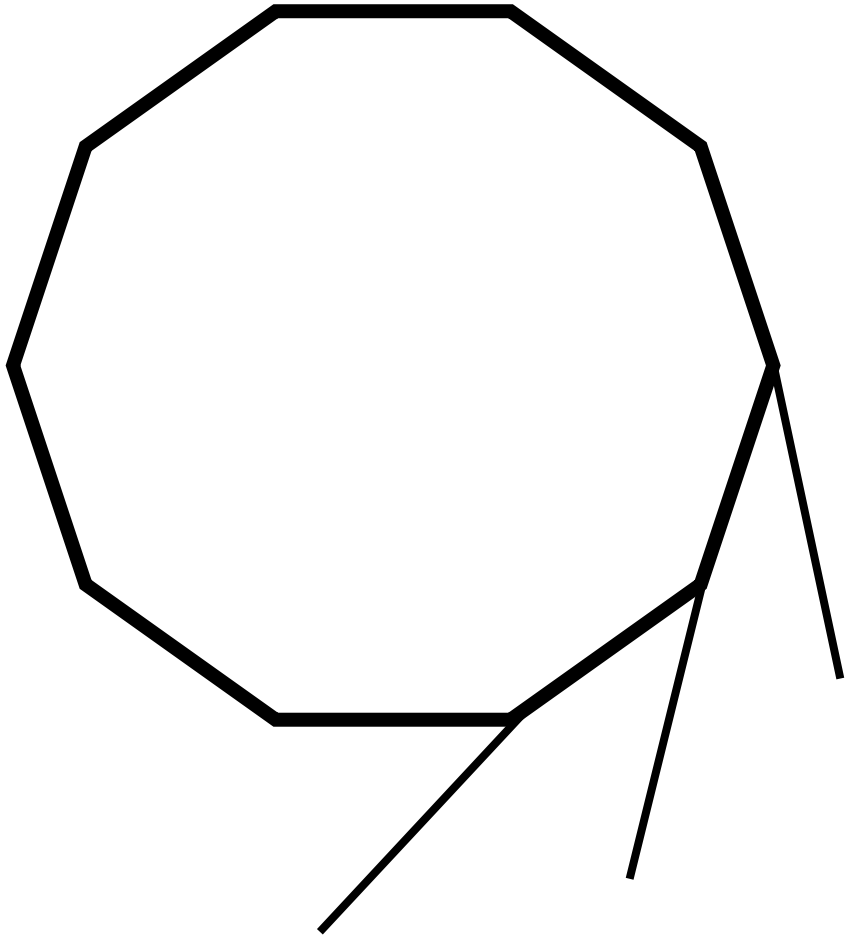
http://www.proconsystems.net/sincrotron-alba-el-acelerador-de-electrones-espanol-n-21-es

Synchrotrons

Monochromatic coherent waves

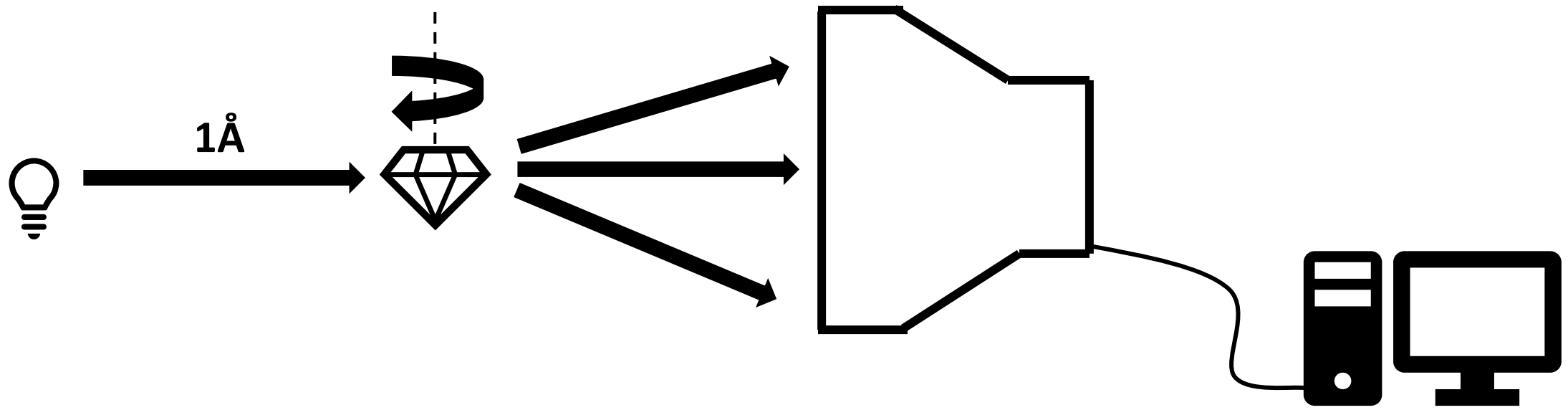
Multiple wavelengths (tuneable)

High intensity

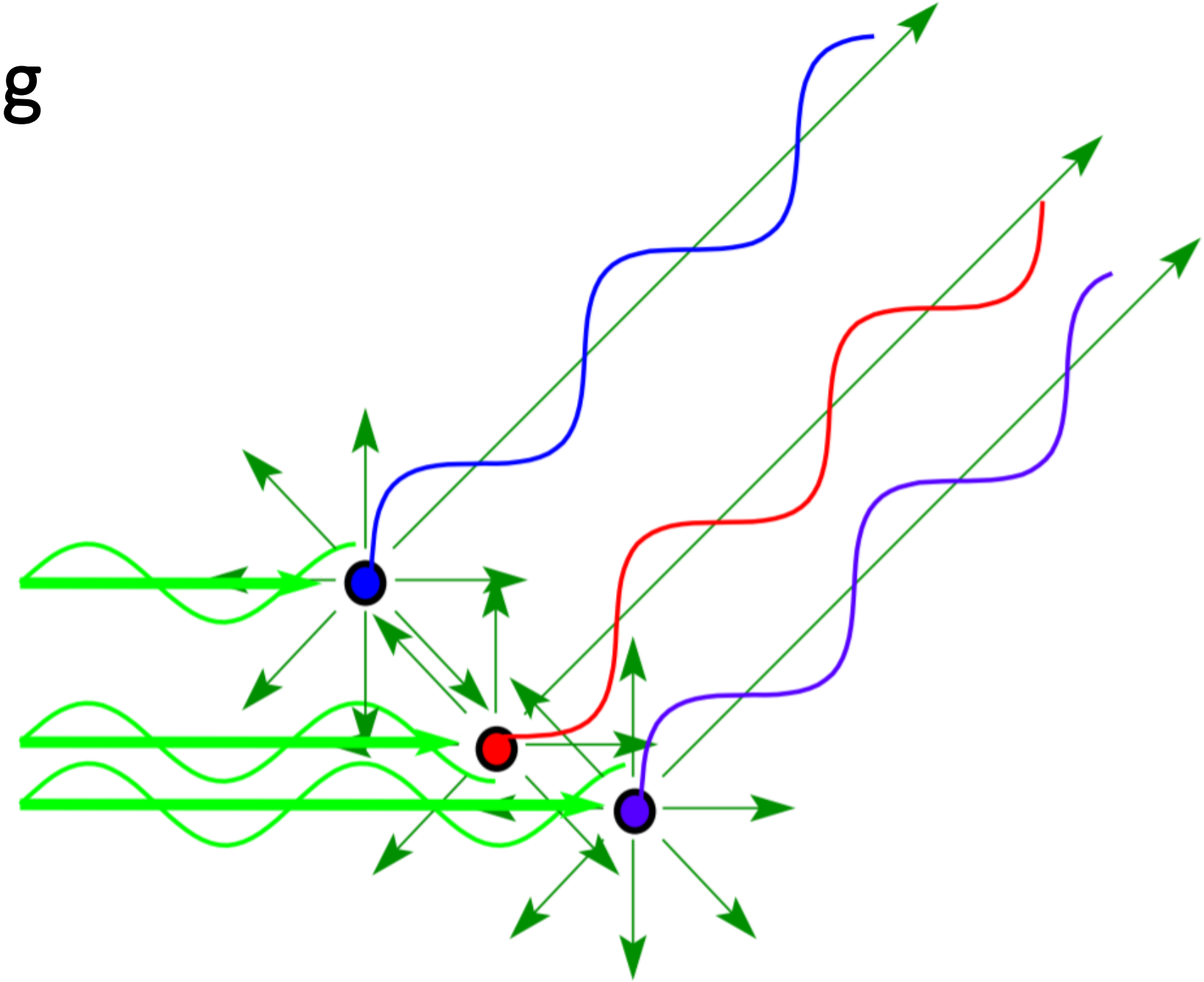


http://www.proconsystems.net/sincrotron-alba-el-acelerador-de-electrones-espanol-n-21-es

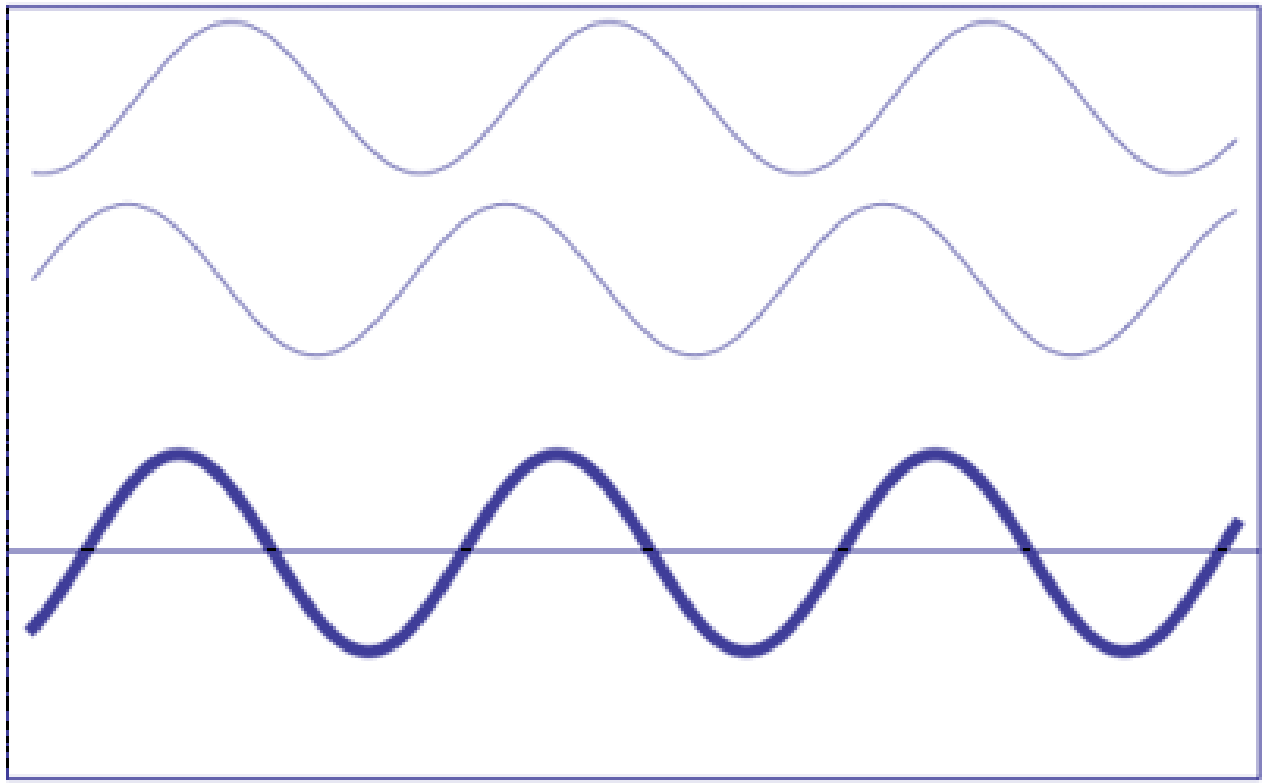
The aim of X-ray crystallography



Elastic scattering

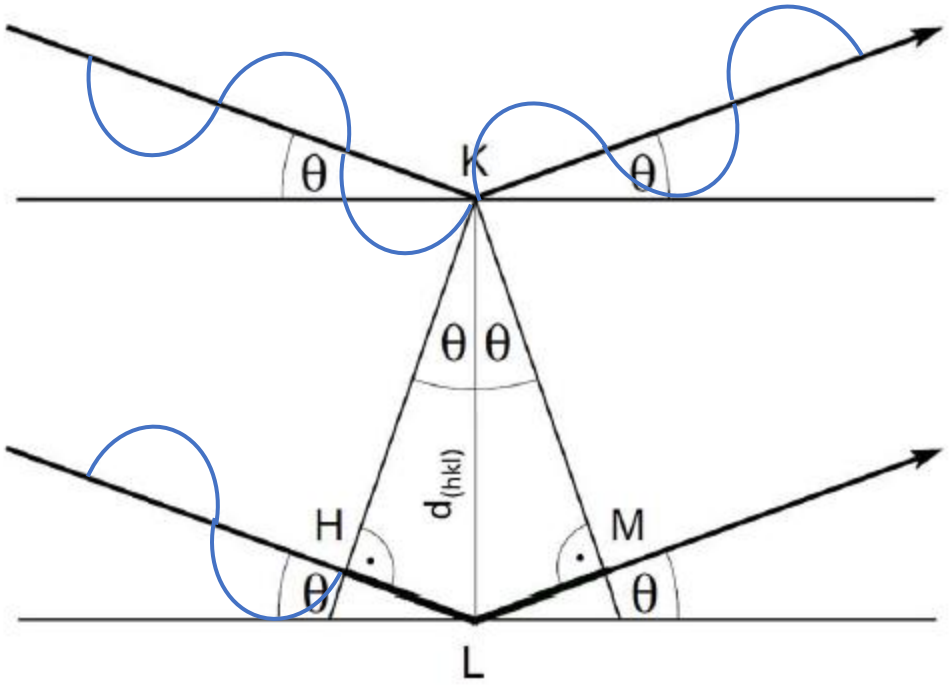


Reminder



Bragg´s law

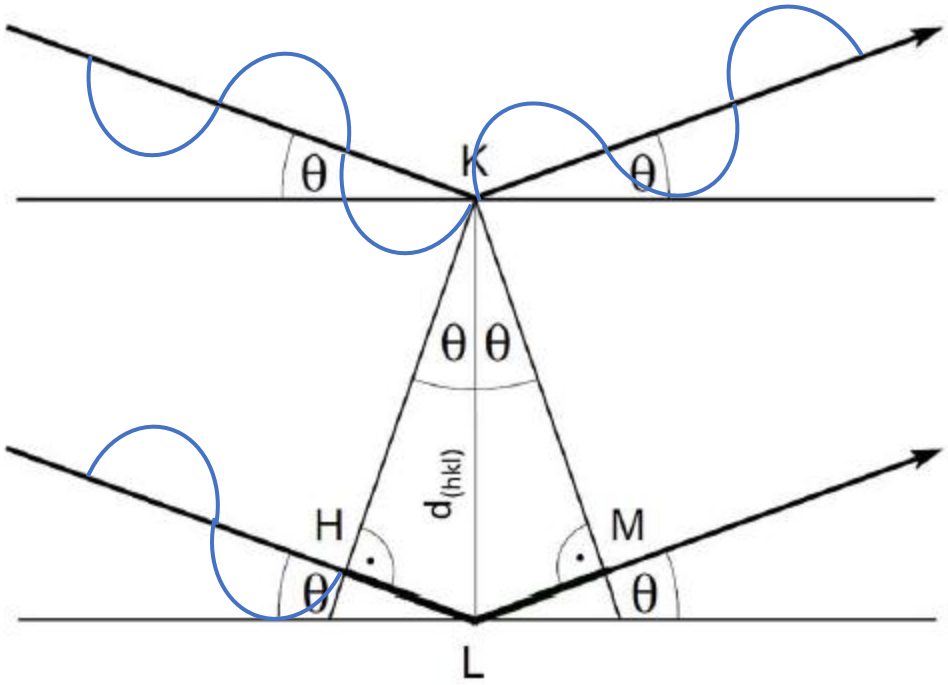
$\sin\theta = \mathrm{HL/d}$



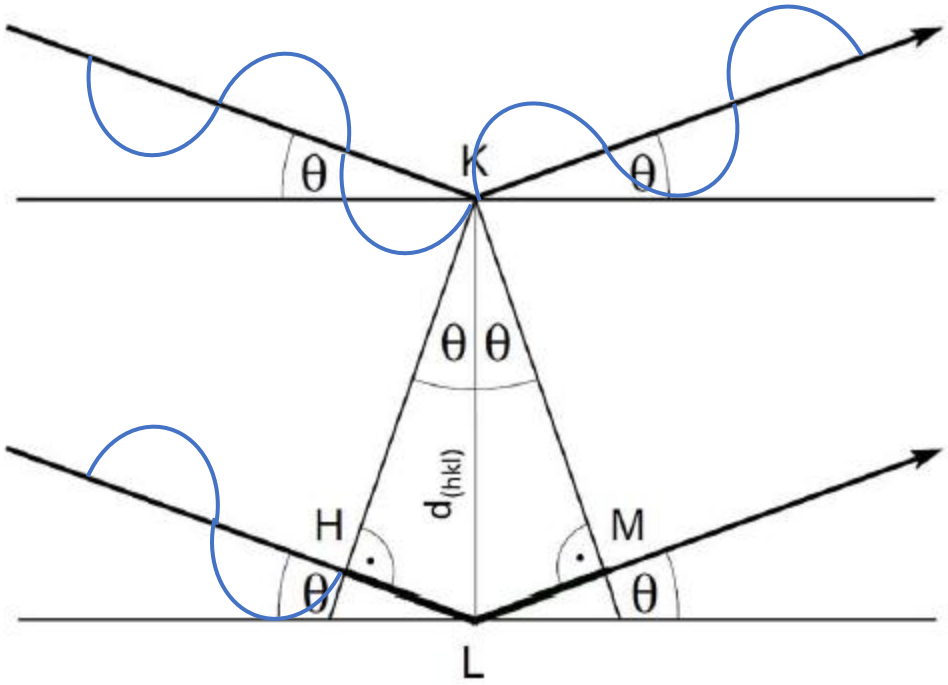
Bragg´s law

$\sin\theta = \mathrm{HL/d}$

HL=LM=w



Bragg´s law

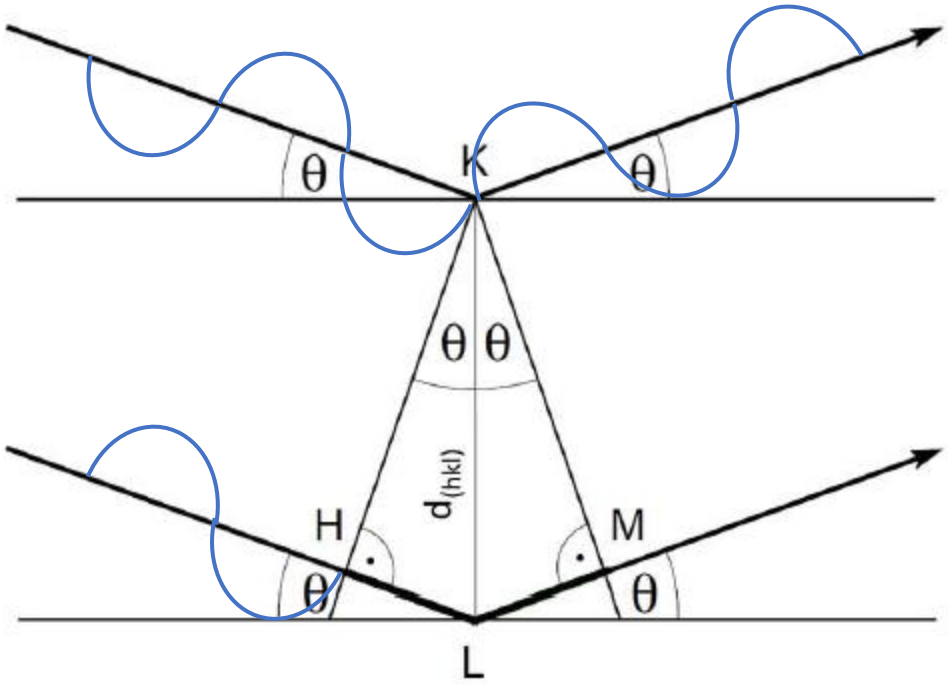


$\sin\theta = \mathrm{HL/d}$

HL=LM=w

2w= $n\lambda$=>w= $n\lambda/2$

Bragg´s law



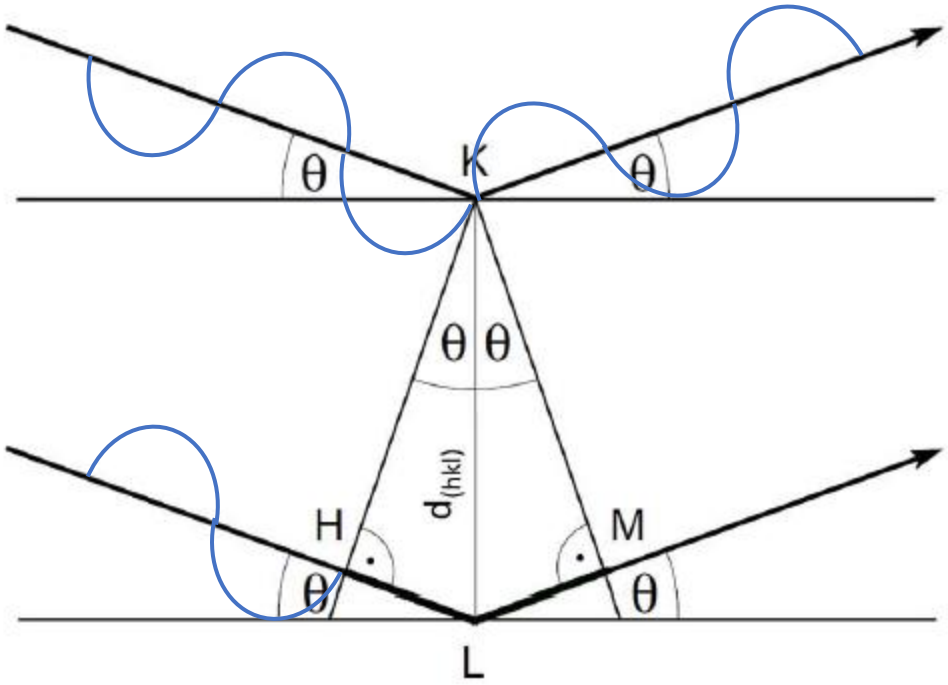
$\sin\theta = \text{HL/d}$

HL=LM=w

2w= $n\lambda$=>w= $n\lambda/2$

$\sin\theta = \text{w/d}$

Bragg´s law



$\sin\theta = \text{HL/d}$

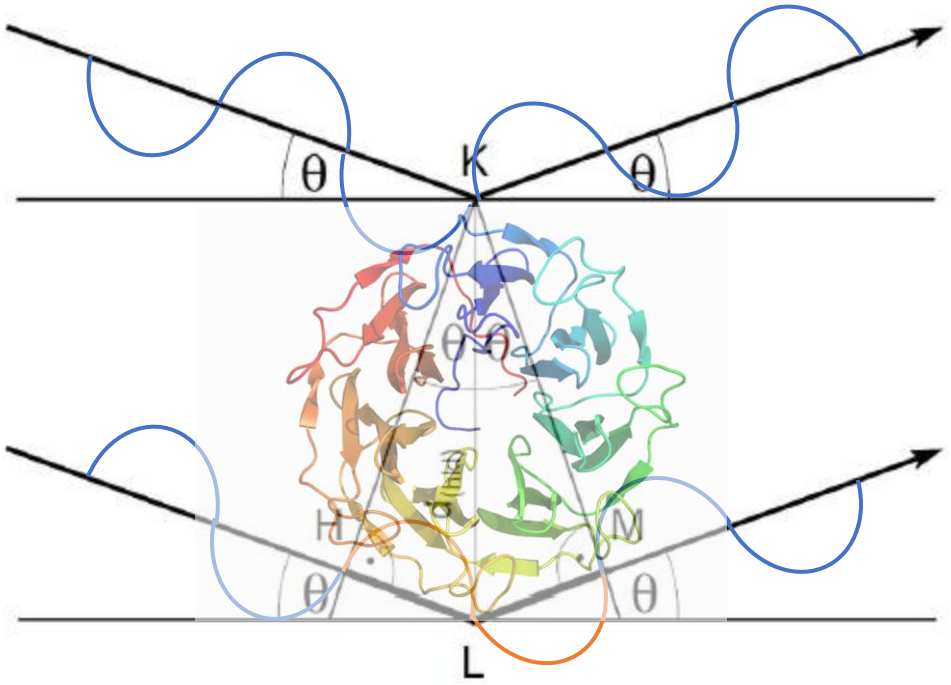
HL=LM=w

$2\text{w} = n\lambda \Rightarrow \text{w} = n\lambda/2$

$\sin\theta = \text{w/d}$

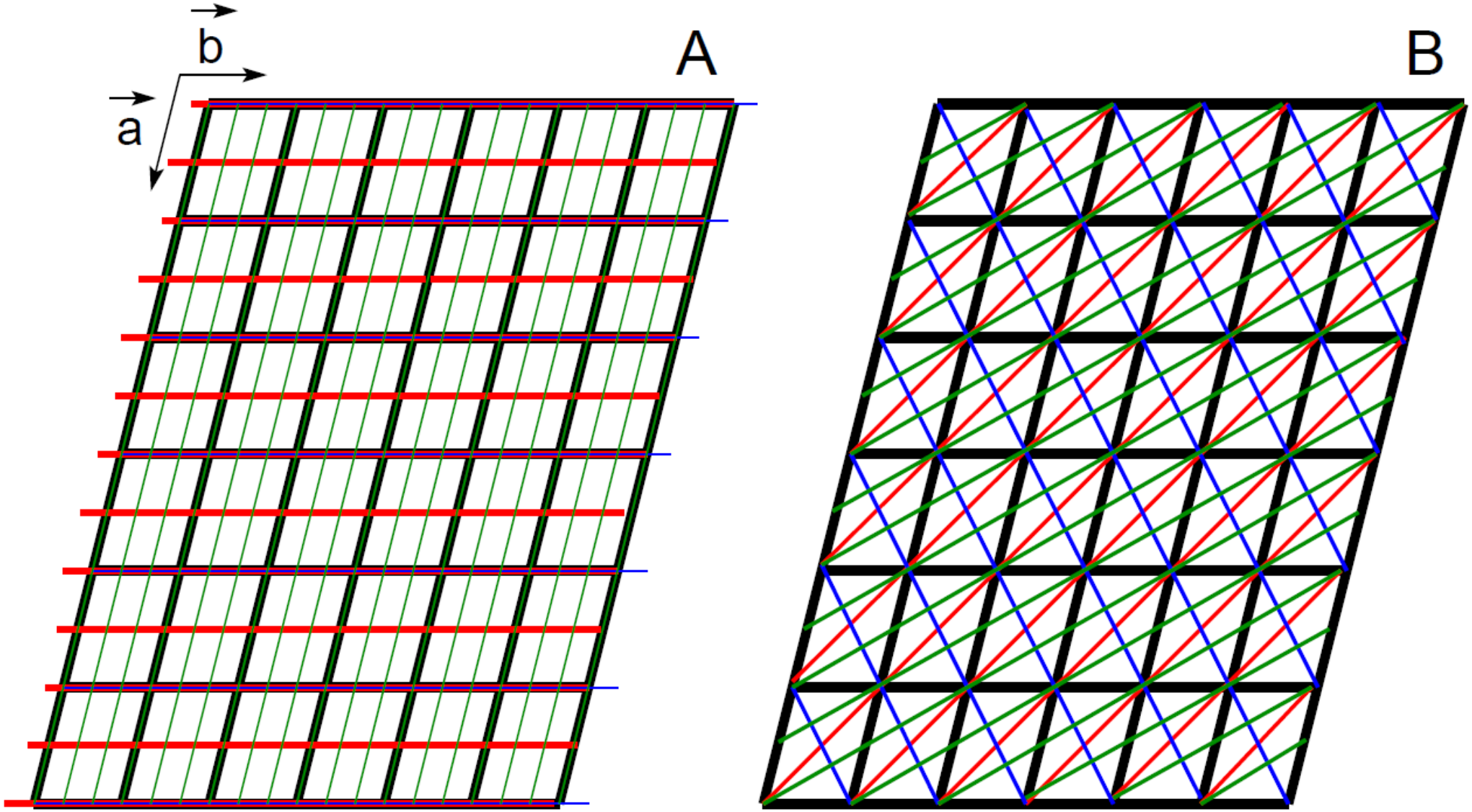
$\sin\theta = n\lambda/2\text{d} \Rightarrow$
$2\text{d}\sin\theta = n\lambda$

Bragg´s law

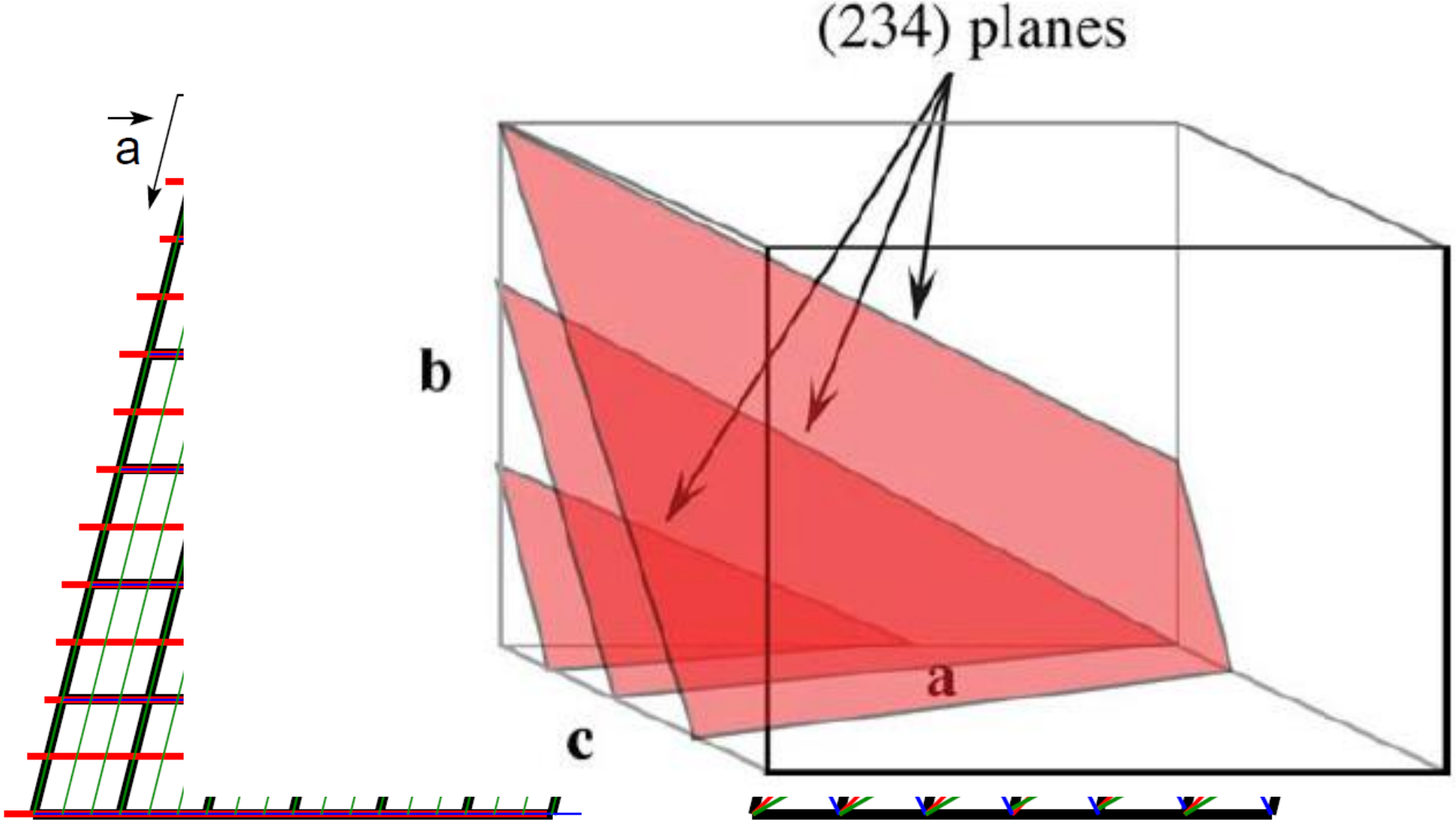


$$2d_{(hkl)} \sin\theta = n\lambda$$

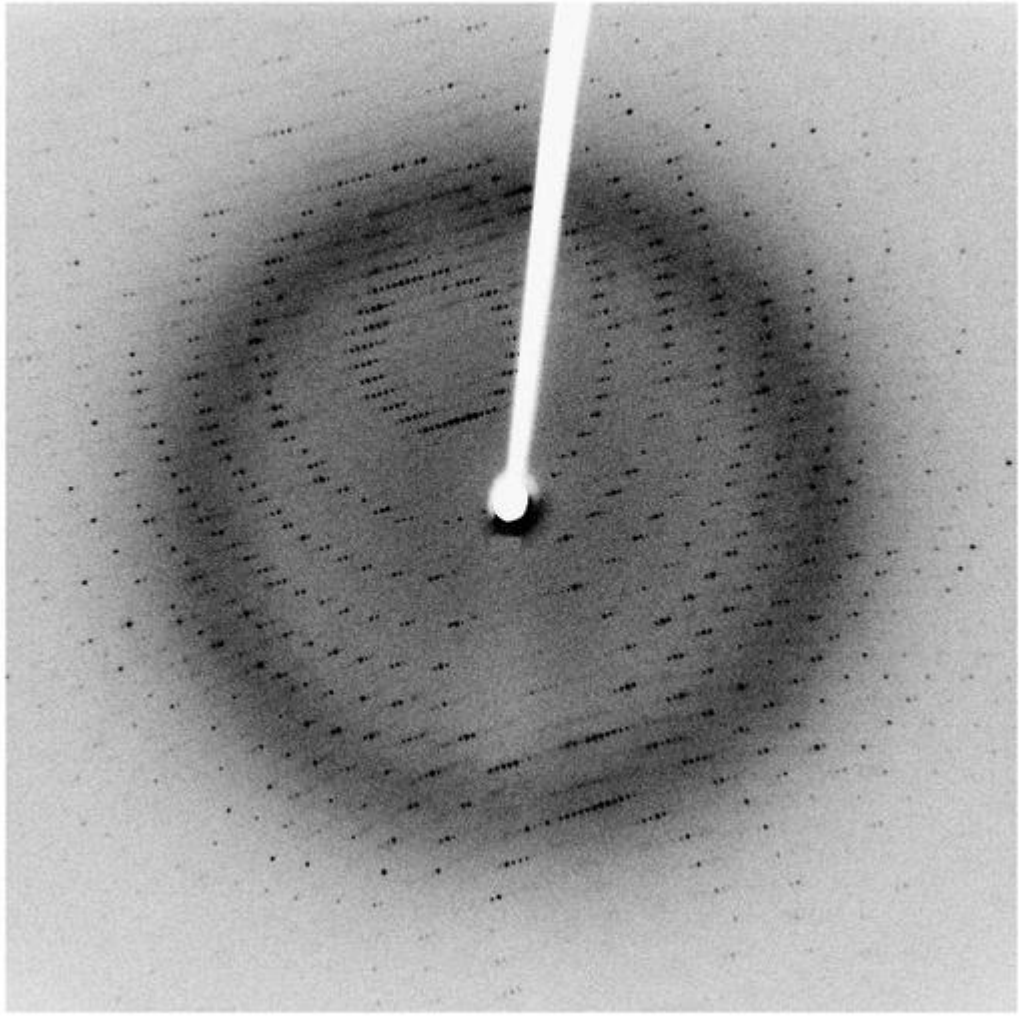
Bragg´s planes [0,0,1]; [0,0,2]; [1,0,0]; [3,0,1]...



Bragg´s planes [0.0.1]; [0.0.2]; [1.0.0]; [3.0.1]...

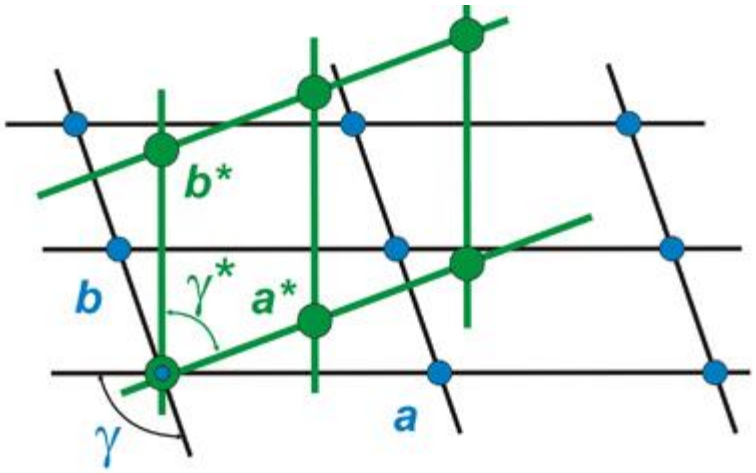


Now reciprocal space…



https://phys.libretexts.org/Courses/University_of_California_Davis/UCD%3A_Biophysics_200A_-_Current_Techniques_in_Biophysics/X-ray_Protein_Crystallography

Now reciprocal space…



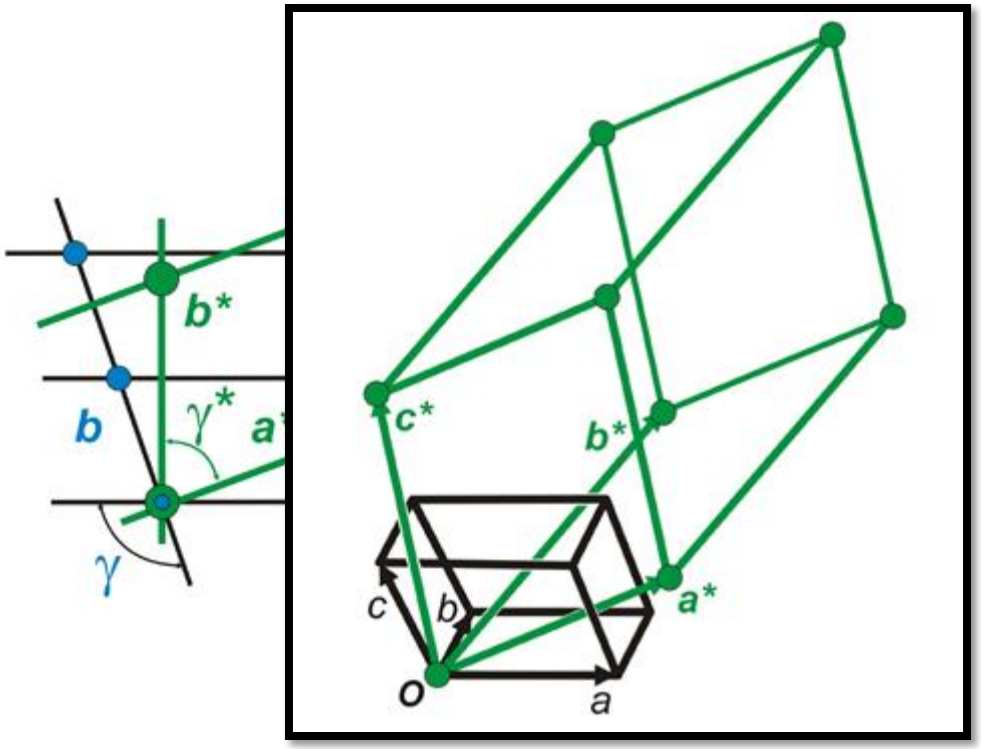
$$a^* = \frac{b \times c}{V}; \quad \cos\alpha^* = \frac{\cos\beta \,.\, \cos\gamma - \cos\alpha}{\sin\beta \,.\, \sin\gamma} \quad ; \quad a^* \text{ perpendicular to: } b \,,\, c$$

$$b^* = \frac{c \times a}{V}; \quad \cos\beta^* = \frac{\cos\alpha \,.\, \cos\gamma - \cos\beta}{\sin\alpha \,.\, \sin\gamma} \quad ; \quad b^* \text{ perpendicular to: } c \,,\, a$$

$$c^* = \frac{a \times b}{V}; \quad \cos\gamma^* = \frac{\cos\alpha \,.\, \cos\beta - \cos\gamma}{\sin\alpha \,.\, \sin\beta} \quad ; \quad c^* \text{ perpendicular to: } a \,,\, b$$

https://www.xtal.iqfr.csic.es/Cristalografia/parte_04-en.html

Now reciprocal space…

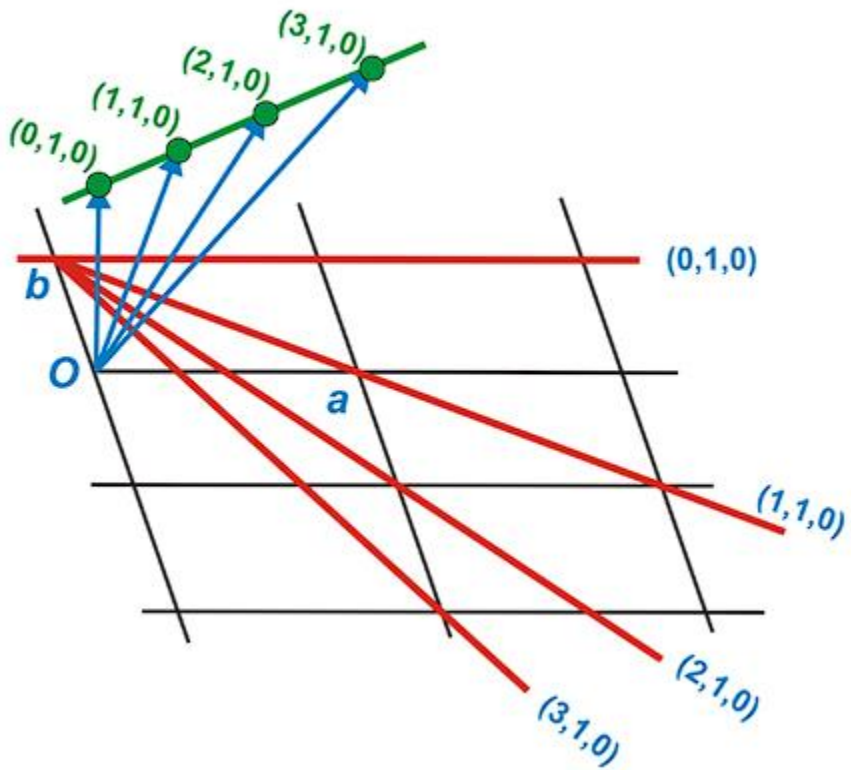


$$a^* = \frac{b \times c}{V}; \quad \cos\alpha^* = \frac{\cos\beta \,.\, \cos\gamma - \cos\alpha}{\sin\beta \,.\, \sin\gamma} \quad ; \quad a^* \text{ perpendicular to: } b \,,\, c$$

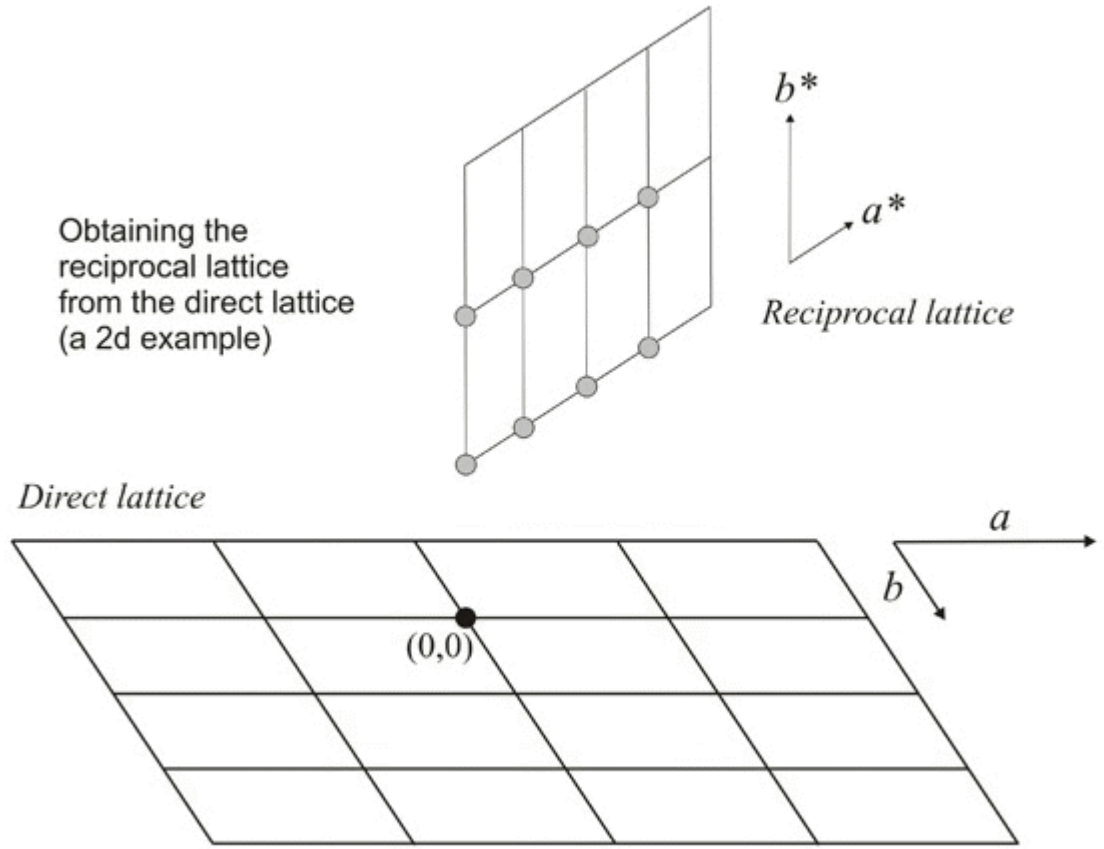
$$b^* = \frac{c \times a}{V}; \quad \cos\beta^* = \frac{\cos\alpha \,.\, \cos\gamma - \cos\beta}{\sin\alpha \,.\, \sin\gamma} \quad ; \quad b^* \text{ perpendicular to: } c \,,\, a$$

$$c^* = \frac{a \times b}{V}; \quad \cos\gamma^* = \frac{\cos\alpha \,.\, \cos\beta - \cos\gamma}{\sin\alpha \,.\, \sin\beta} \quad ; \quad c^* \text{ perpendicular to: } a \,,\, b$$

https://www.xtal.iqfr.csic.es/Cristalografia/parte_04-en.html



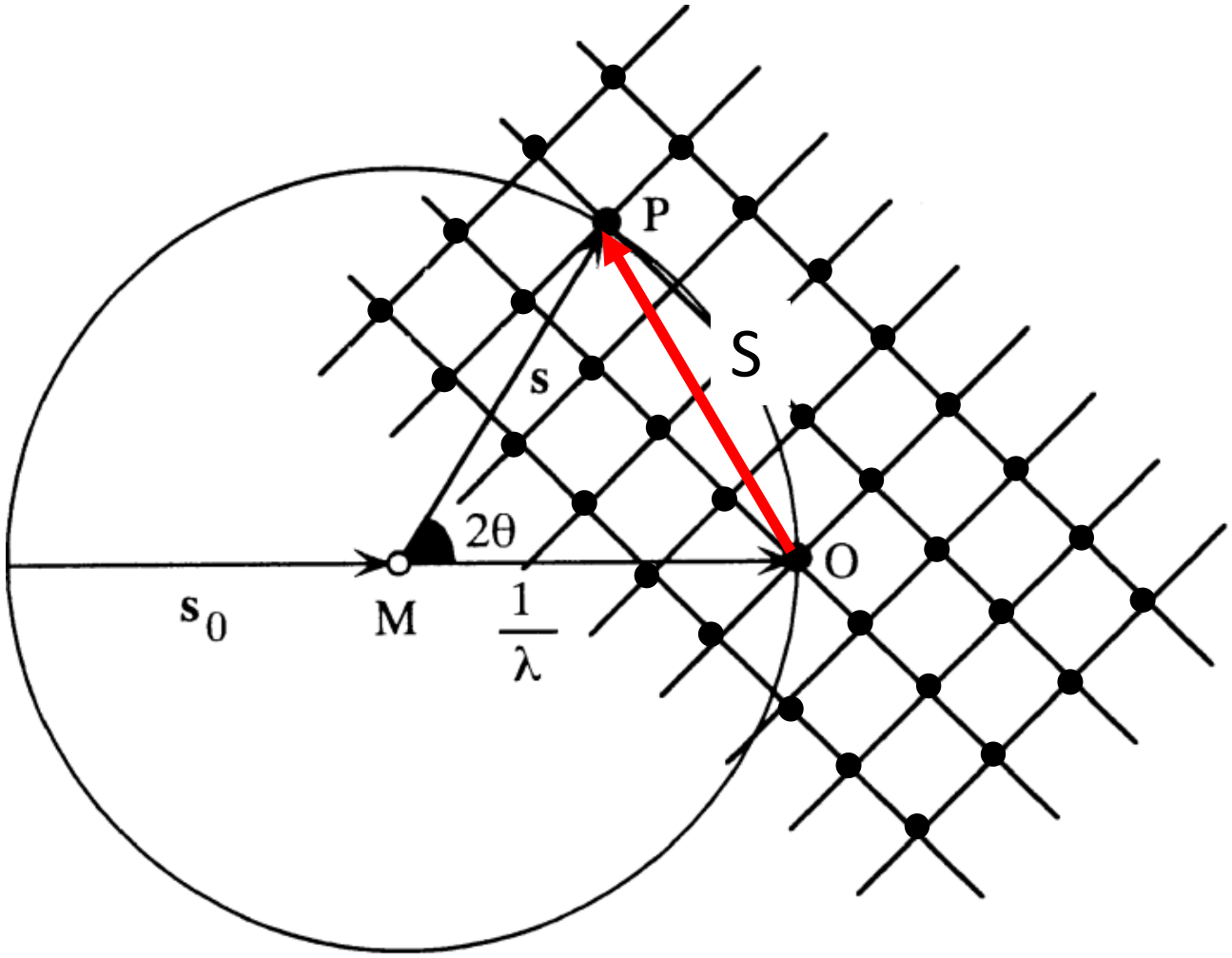
Obtaining the reciprocal lattice from the direct lattice (a 2d example)



https://www.xtal.iqfr.csic.es/Cristalografia/parte_04-en.html

Ewald´s sphere

Geometrical description of the Bragg´s law in the reciprocal space.



Ewald´s sphere

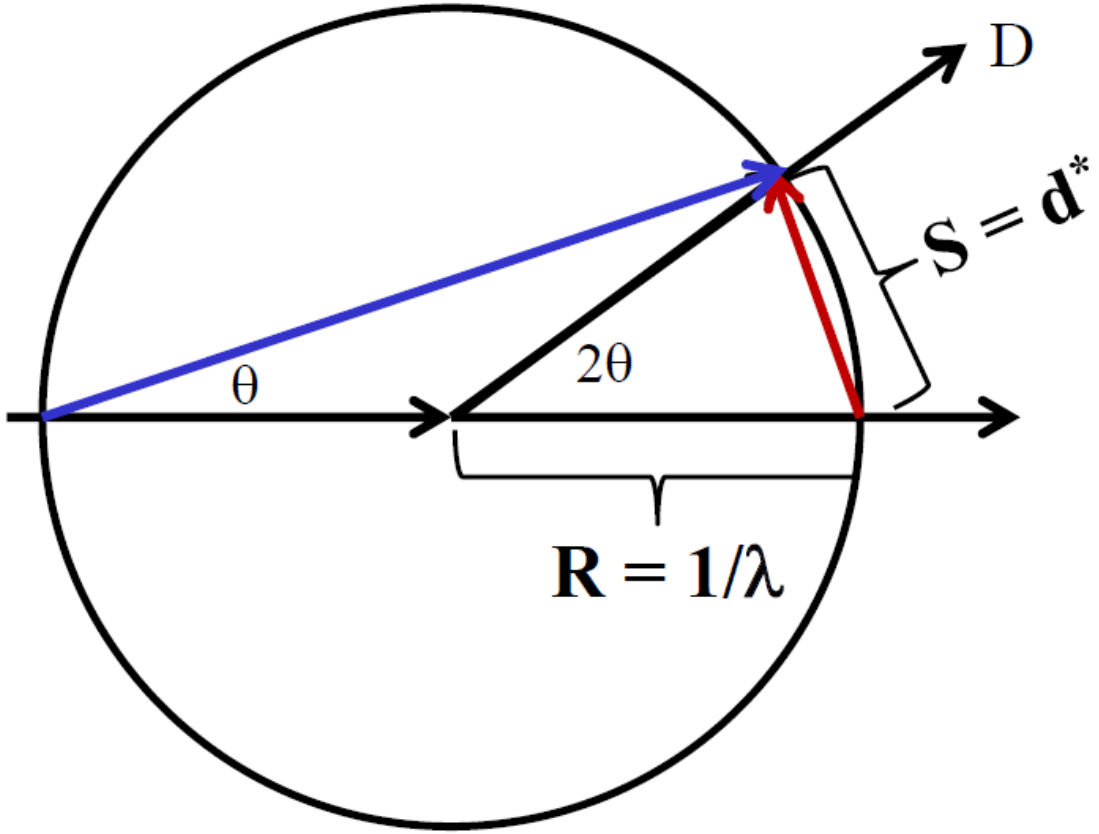
$\mathrm{Sin}\theta = S / 2R$

$S = d^*_{hkl} = 1/d_{hkl}$

$R = (1/\lambda)$

$\Rightarrow \mathrm{Sin}\theta = (1/d_{hkl}) / 2(1/\lambda)$

$\mathrm{Or\ Sin}\theta = 1\ \lambda / 2d_{hkl}$



Andrew McCarthy

Ewald´s sphere

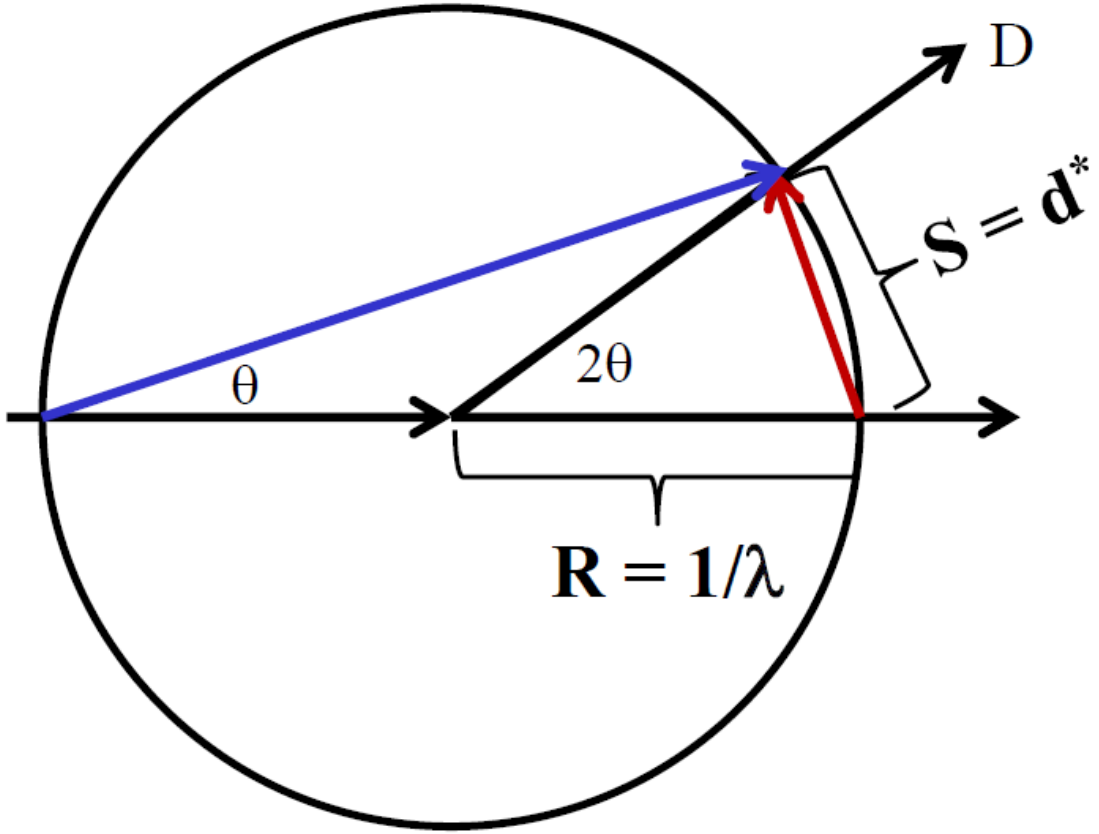
$\mathrm{Sin}\theta = S / 2R$

$S = d^*_{hkl} = 1/d_{hkl}$

$R = (1/\lambda)$

$\Rightarrow \mathrm{Sin}\theta = (1/d_{hkl}) / 2(1/\lambda)$

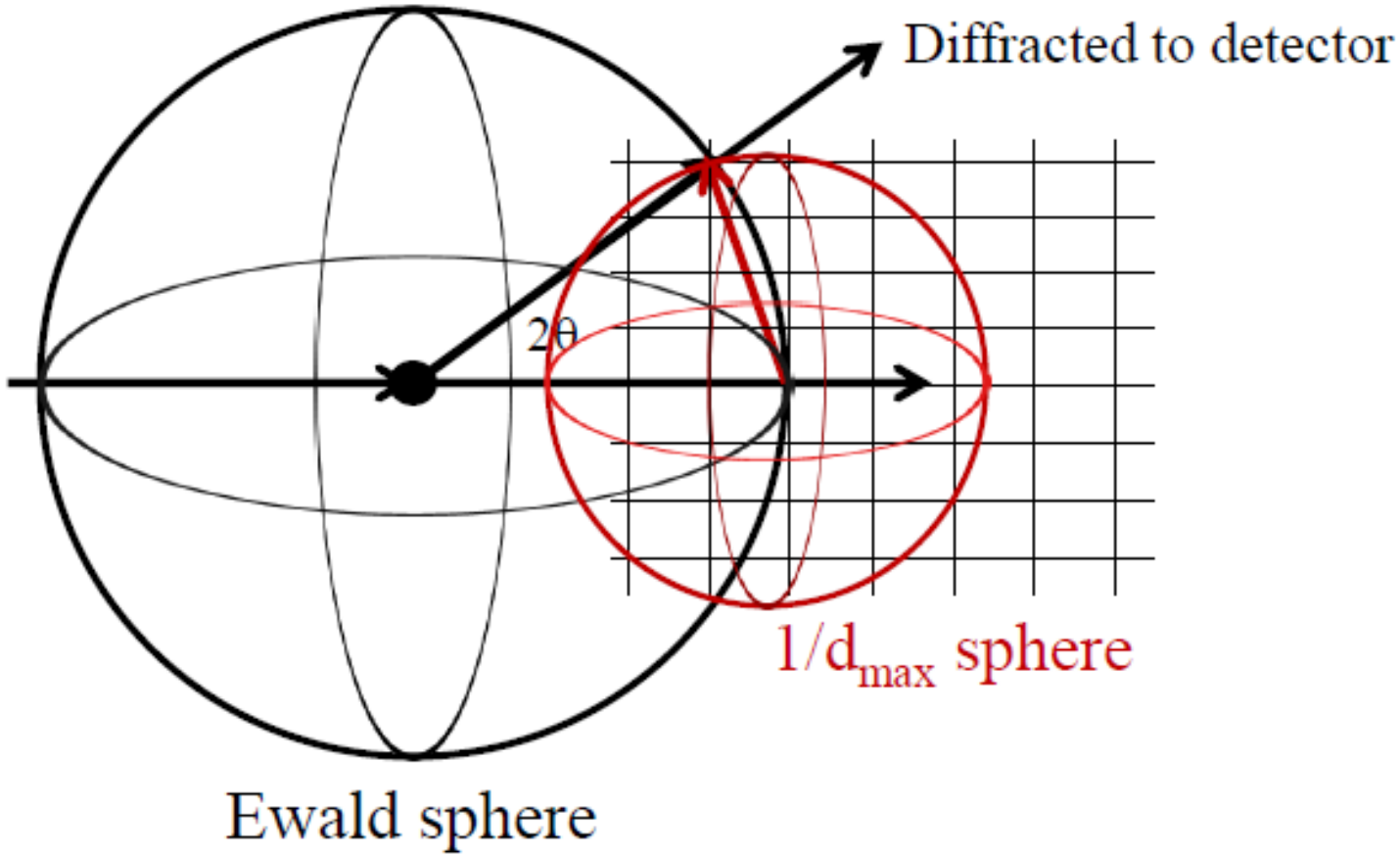
$\mathrm{Or\ Sin}\theta = 1\ \lambda / 2d_{hkl}$



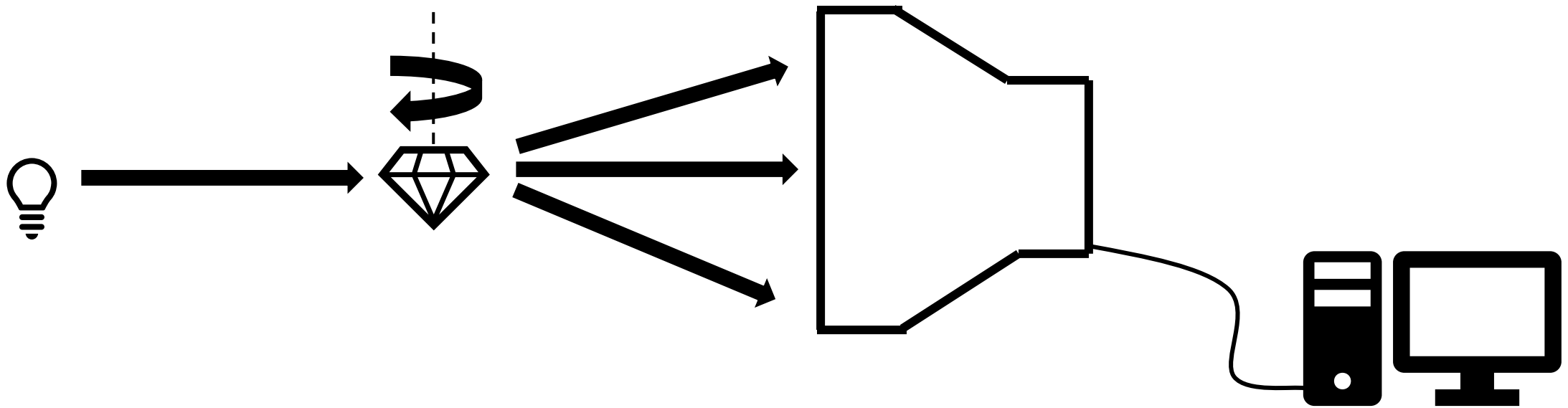
$$2d_{(hkl)}\,\mathrm{Sin}\,\theta = n\lambda$$

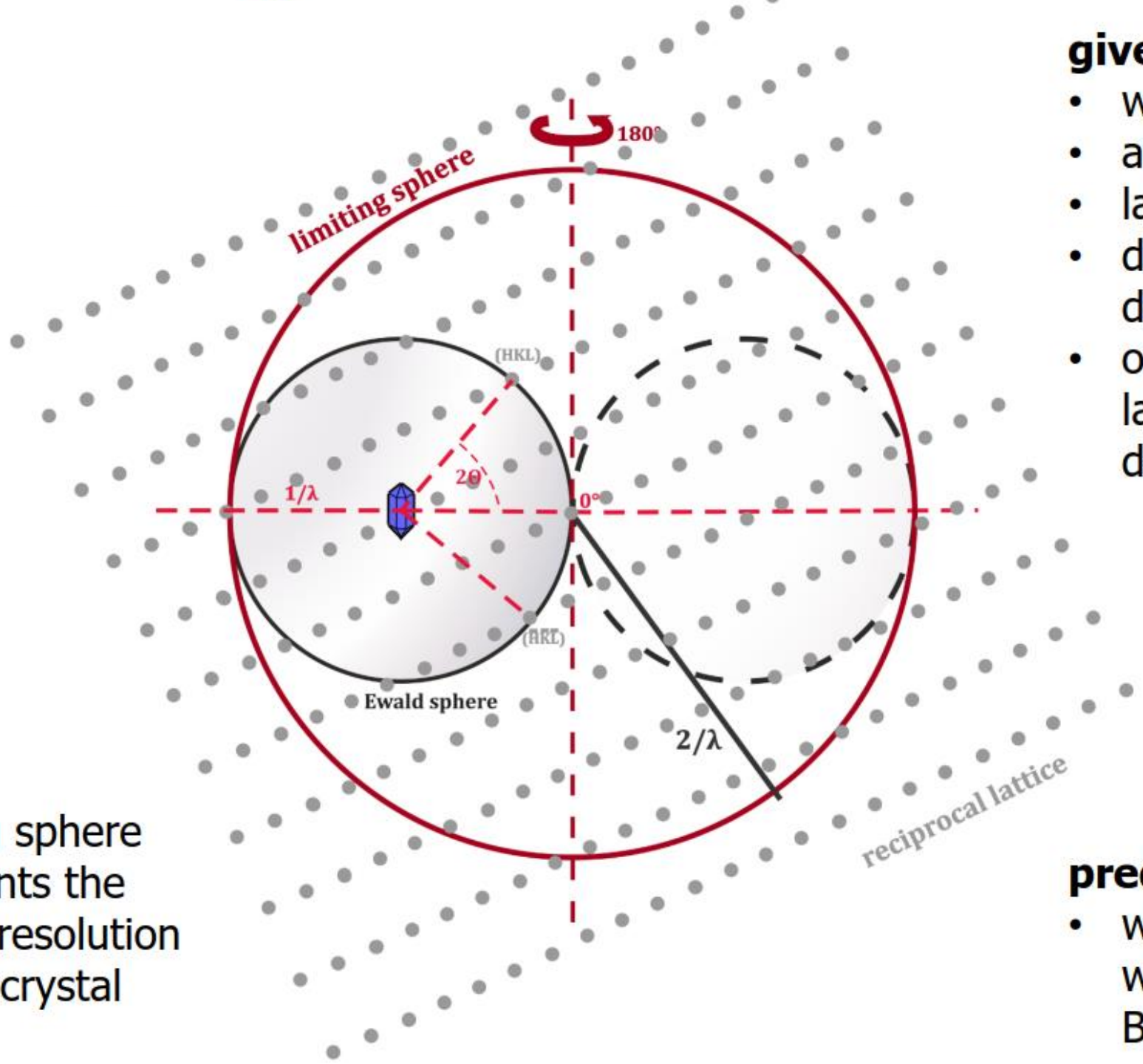
Andrew McCarthy

Ewald´s sphere



Andrew McCarthy





given:

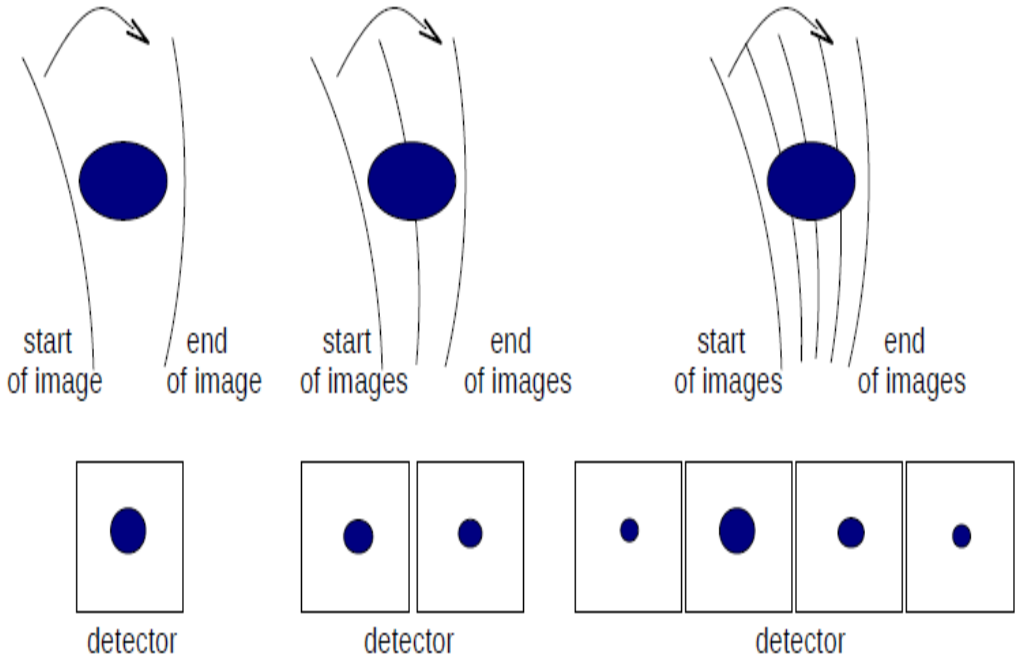
- wavelength
- angle
- lattice
- distance from detector
- orientation of lattice relative to detector

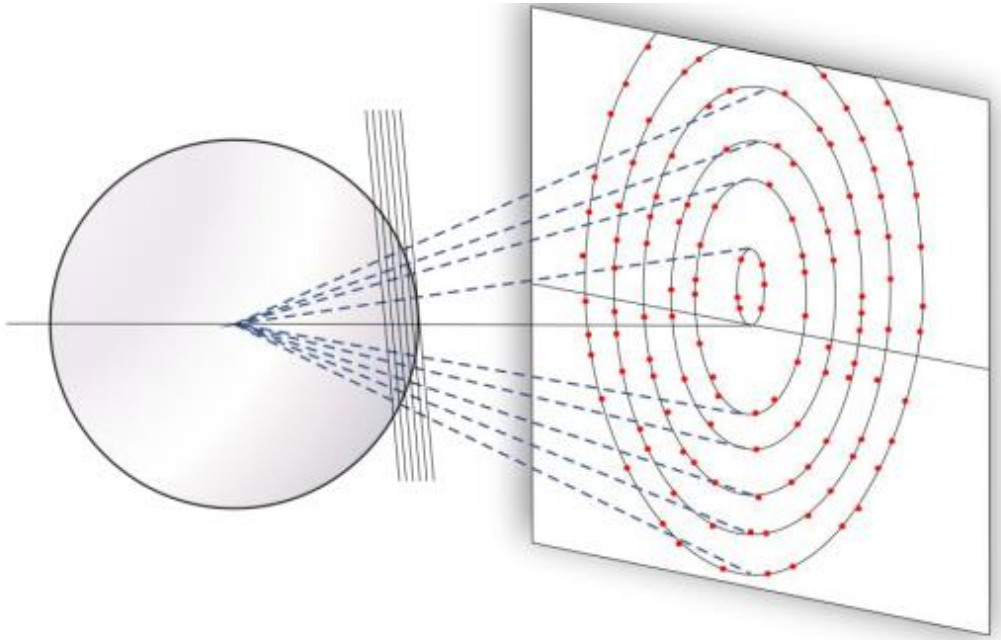
Limiting sphere represents the limit of resolution of your crystal

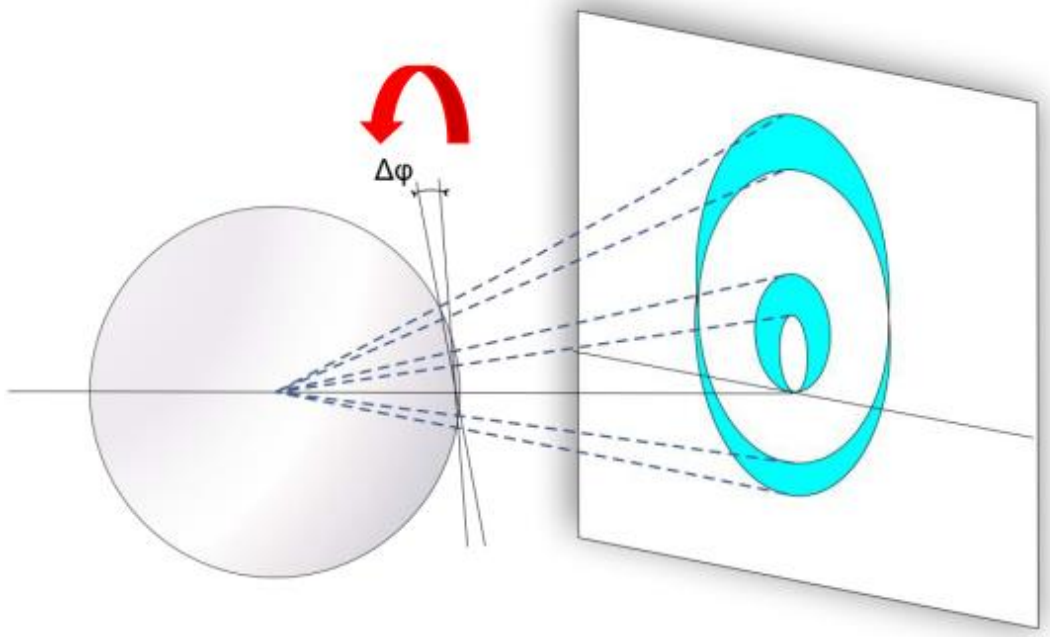
predicts:

- which diffracted waves satisfy Bragg's law

Tilt angle


start
of image
end
of image
start
of images
end
of images
start
of images
end
of images
detector
detector
detector



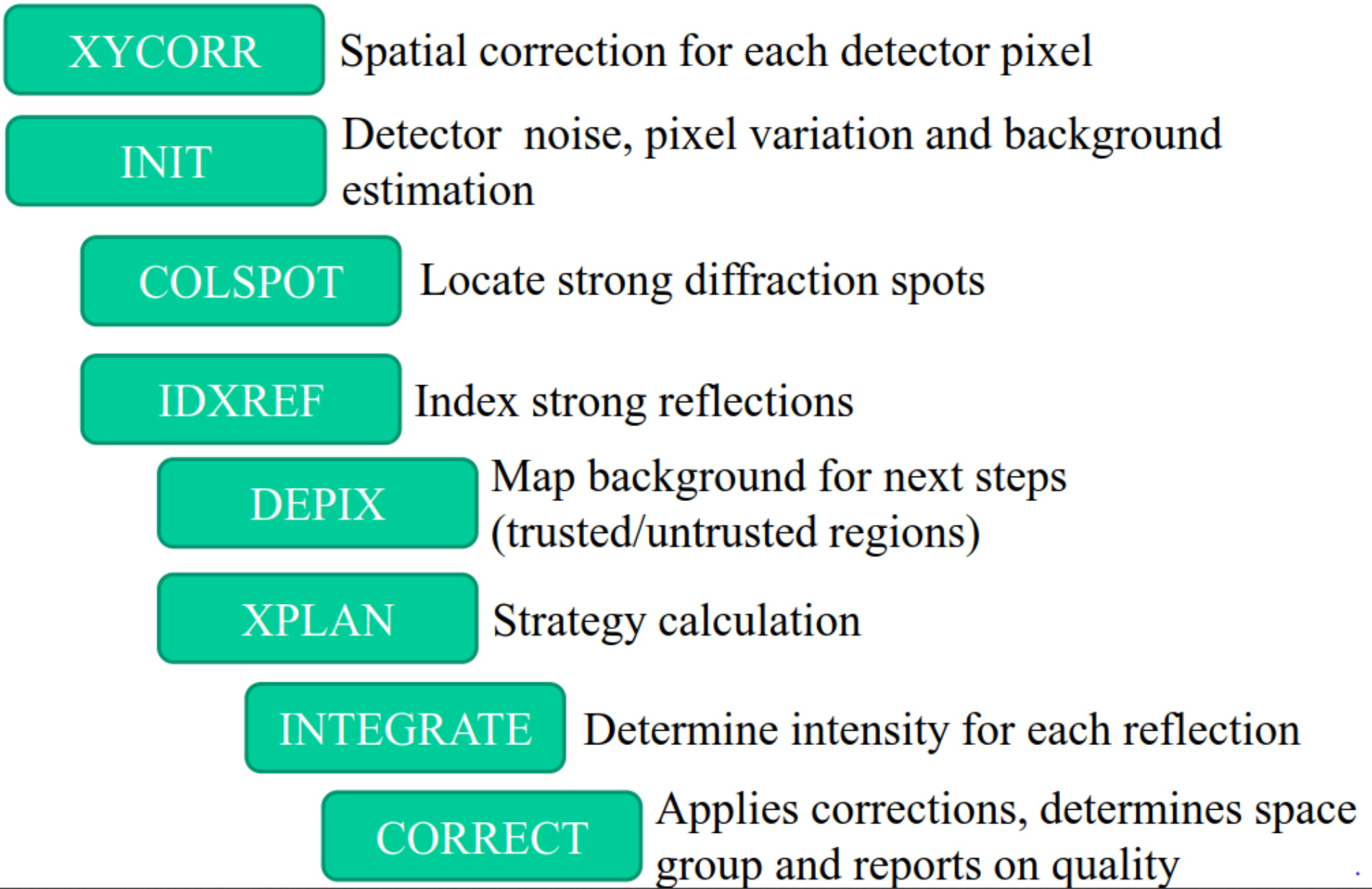

Δφ

Data collection strategy

- Trial 2 pictures of the crystal for quality check in 90°

- Wavelength
- Tilt angle
- Rotation range
- Starting angle of data collection
- Intensity of the beam
- Resolution?

XDS


XYCORR
Spatial correction for each detector pixel
INIT
Detector noise, pixel variation and background estimation
COLSPOT
Locate strong diffraction spots
IDXREF
Index strong reflections
DEPIX
Map background for next steps (trusted/untrusted regions)
XPLAN
Strategy calculation
INTEGRATE
Determine intensity for each reflection
CORRECT
Applies corrections, determines space group and reports on quality

Data processing

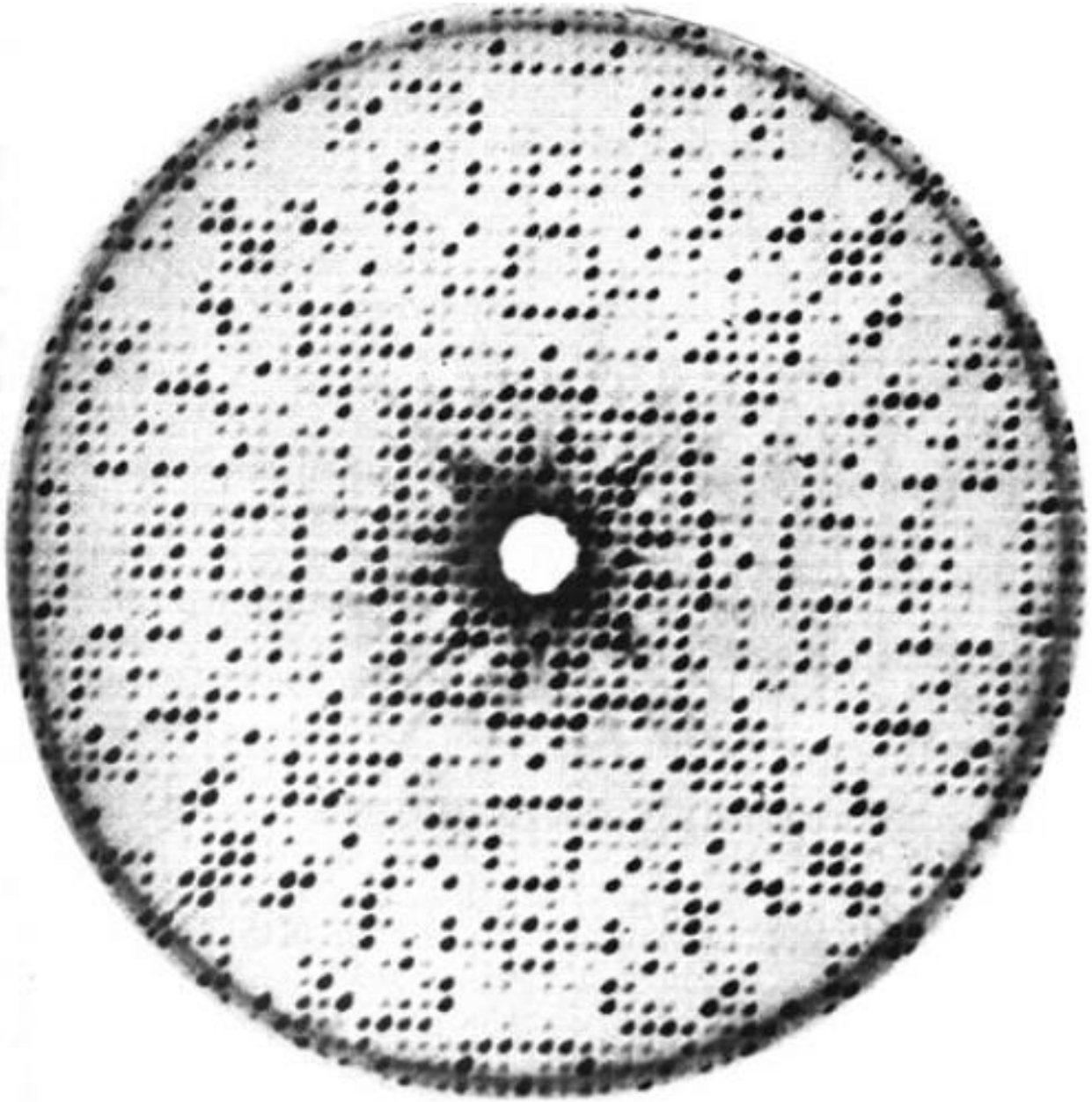
- Spot finding and autoindexing
- Parameter refinement
- Integration (the actual measurement of spot intensities)
- Scaling and merging

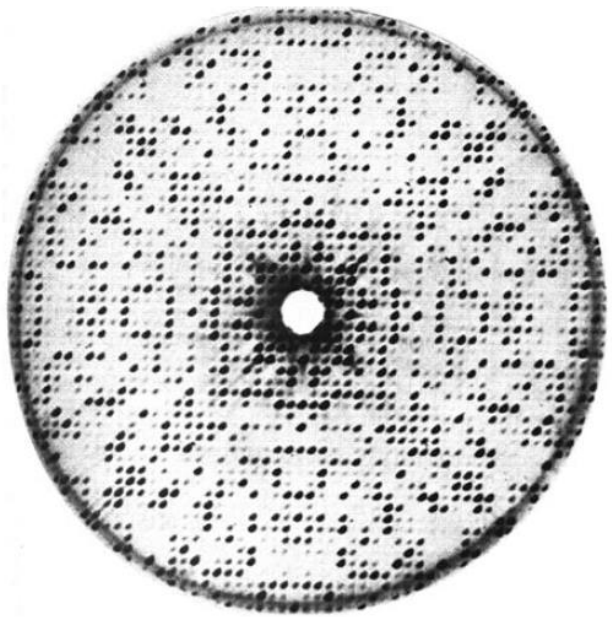
Scale and Merge

- Scale:
  - Illuminated volume of the crystal
  - Change in incident radiation intensity
  - Non-uniformity of detector response
  - Radiation damage
  - …
- Merge:
  - Merging same or symetry equivalent reflections

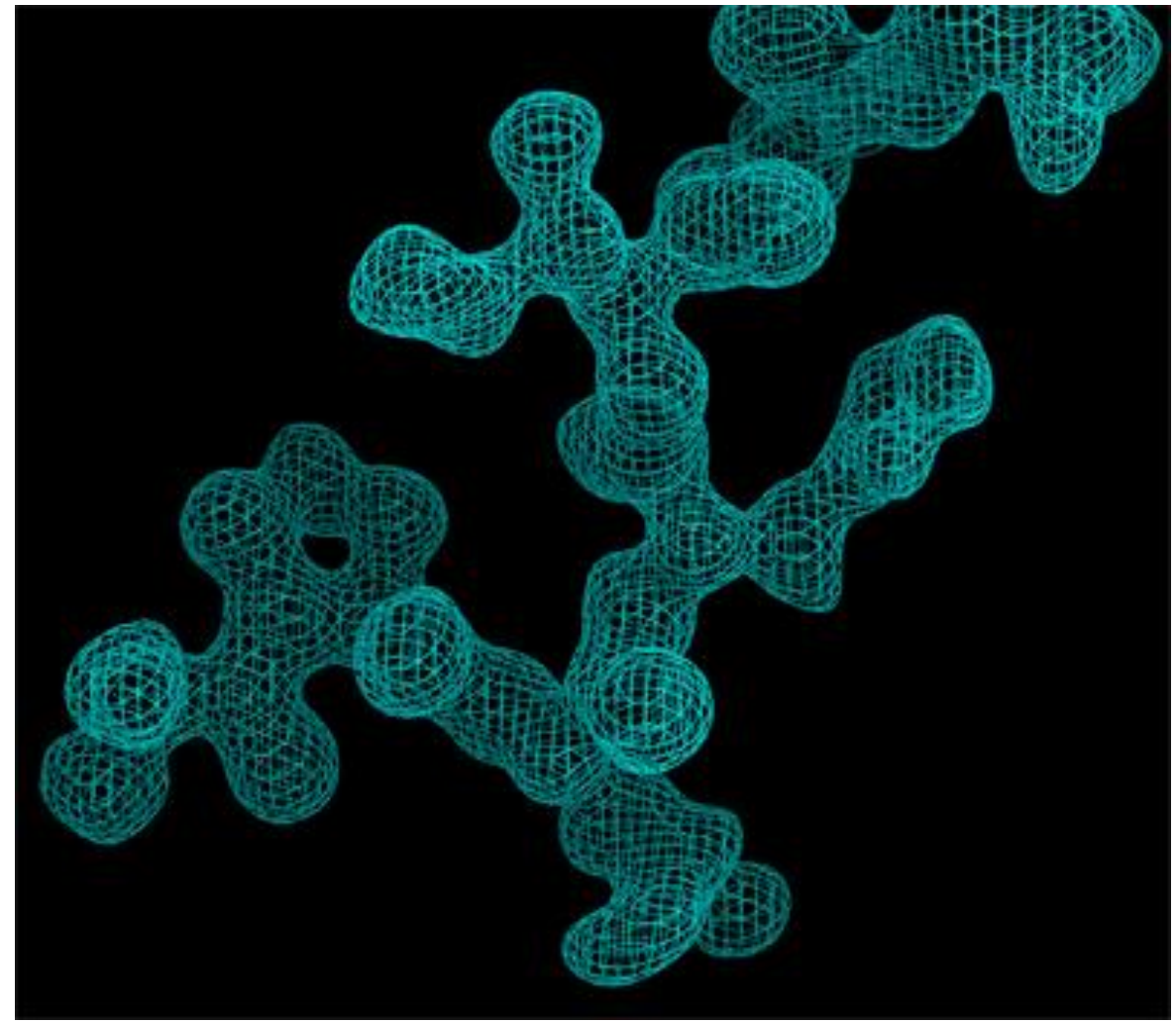
Scale an

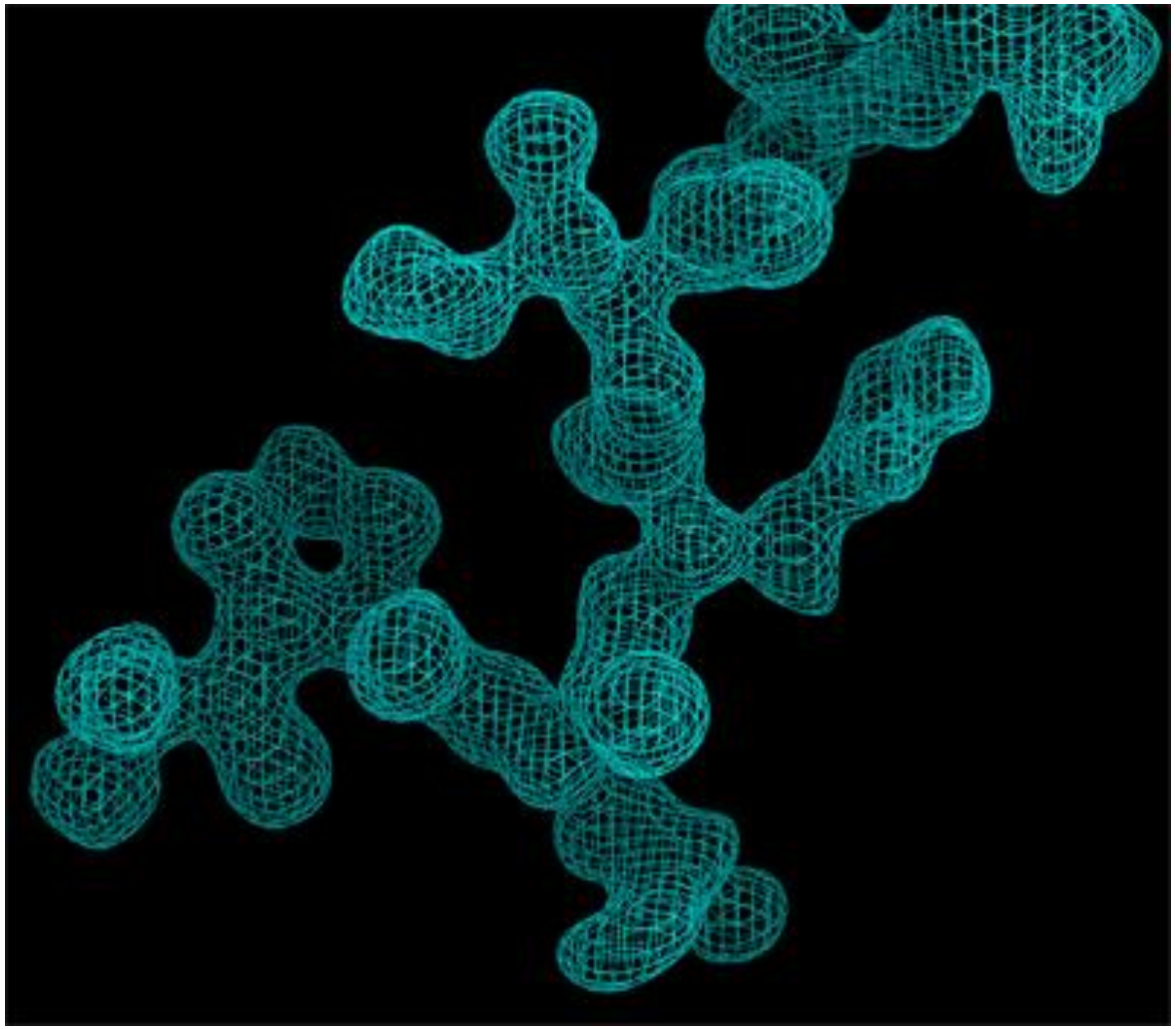
- Scale:
  - Illumin
  - Change
  - Non-ur
  - …
- Merge:
  - Mergir

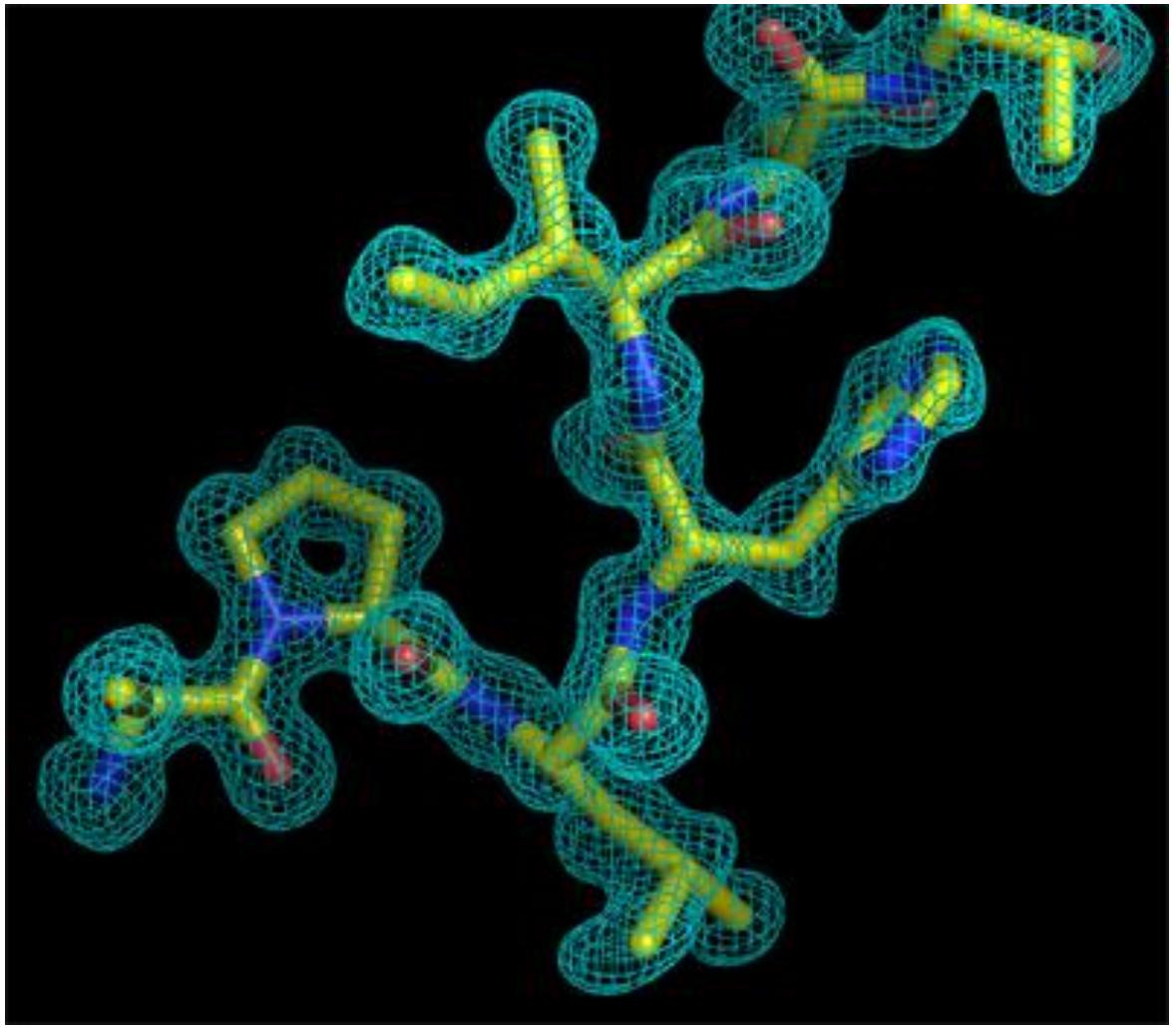




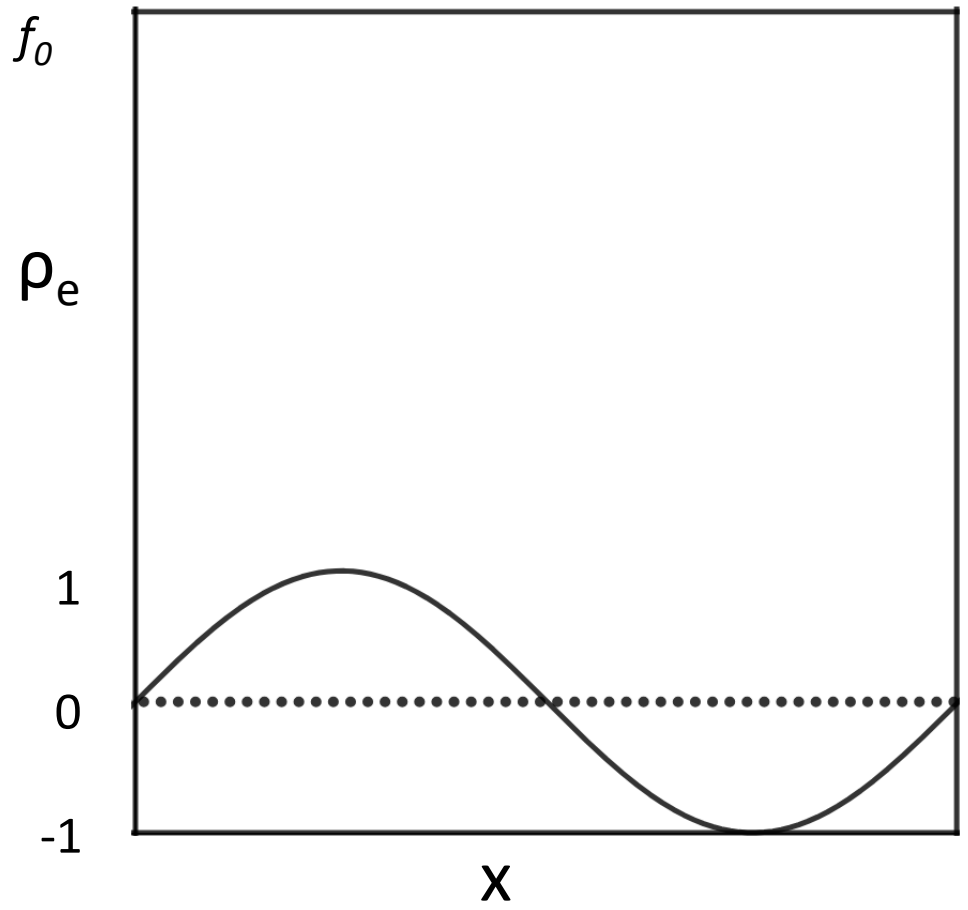
FT



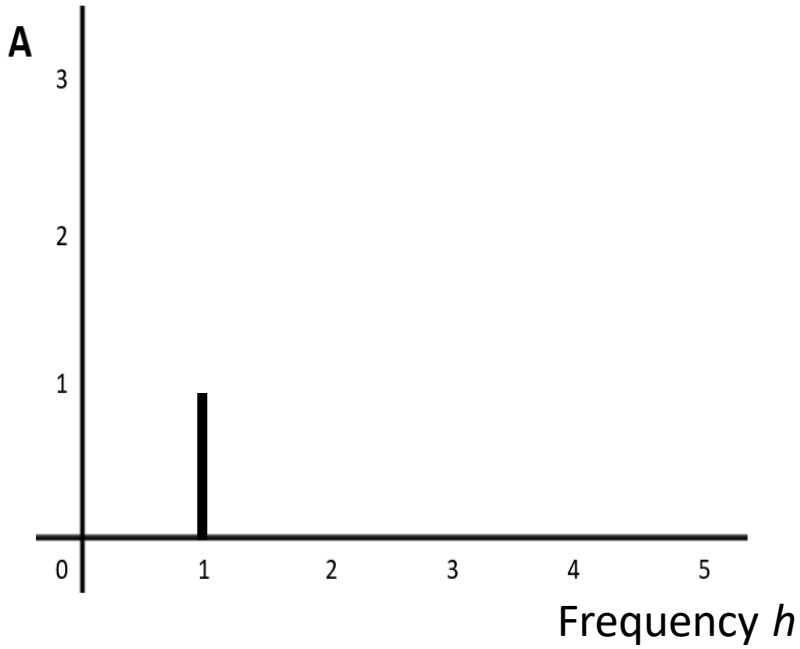




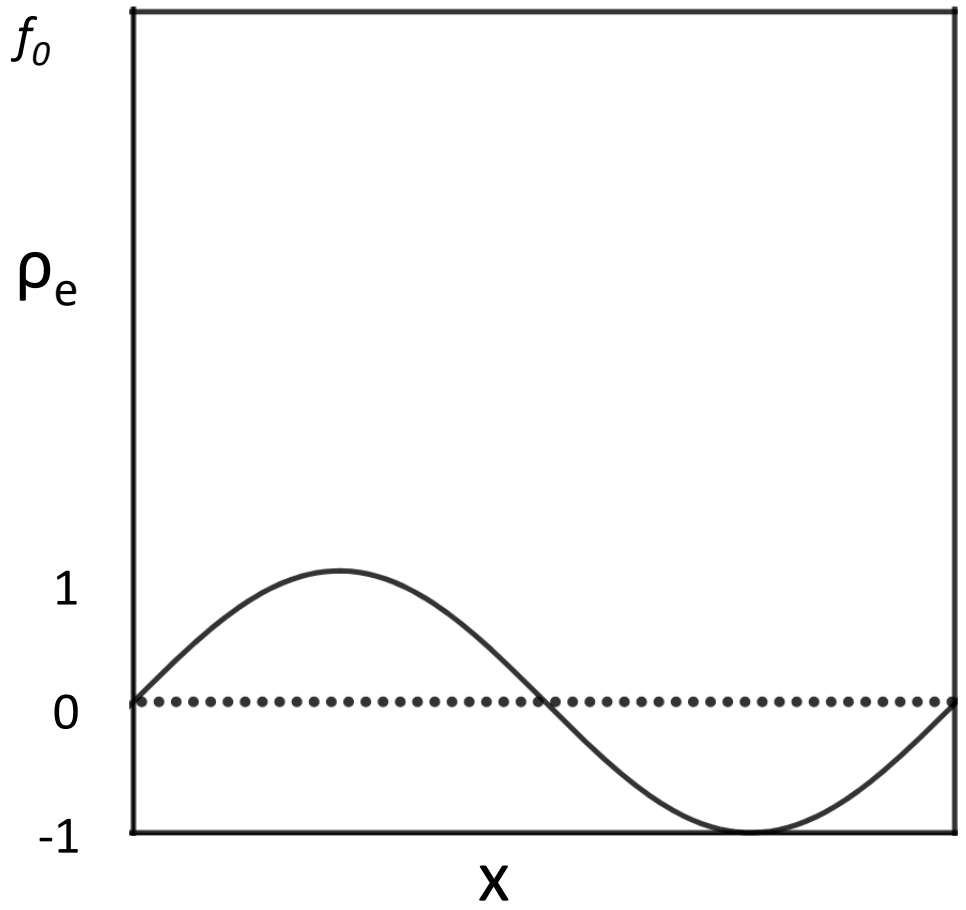
1D experiment



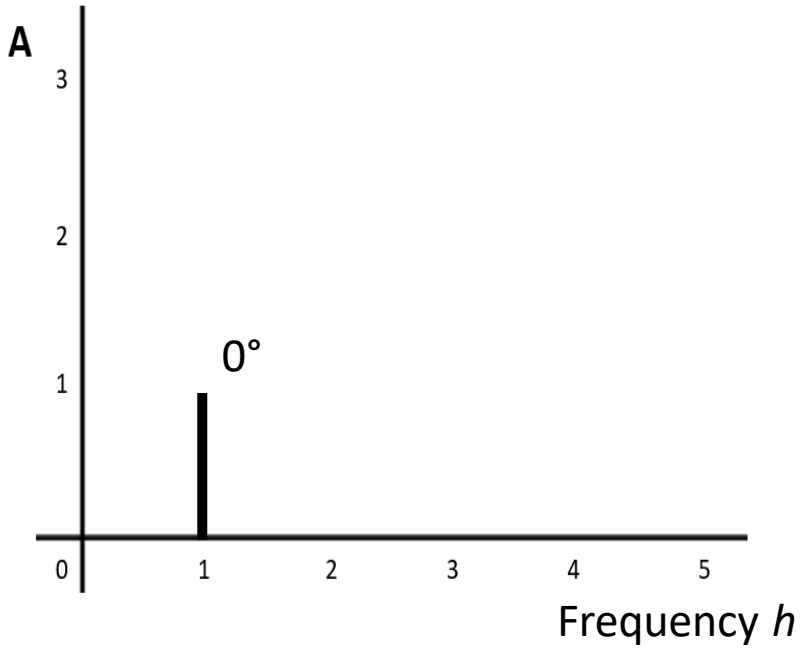
FT



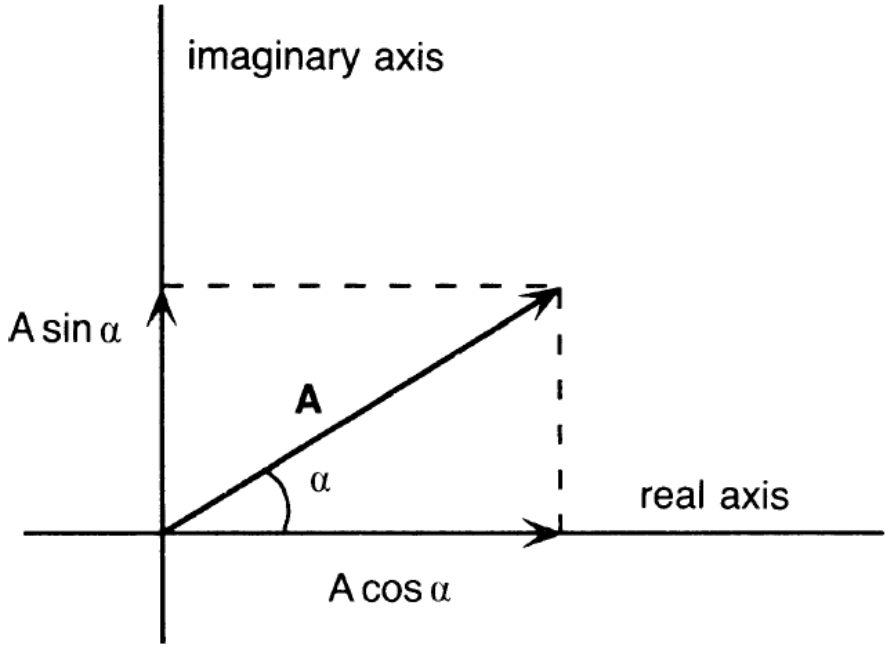
1D experiment



FT



Argand diagram

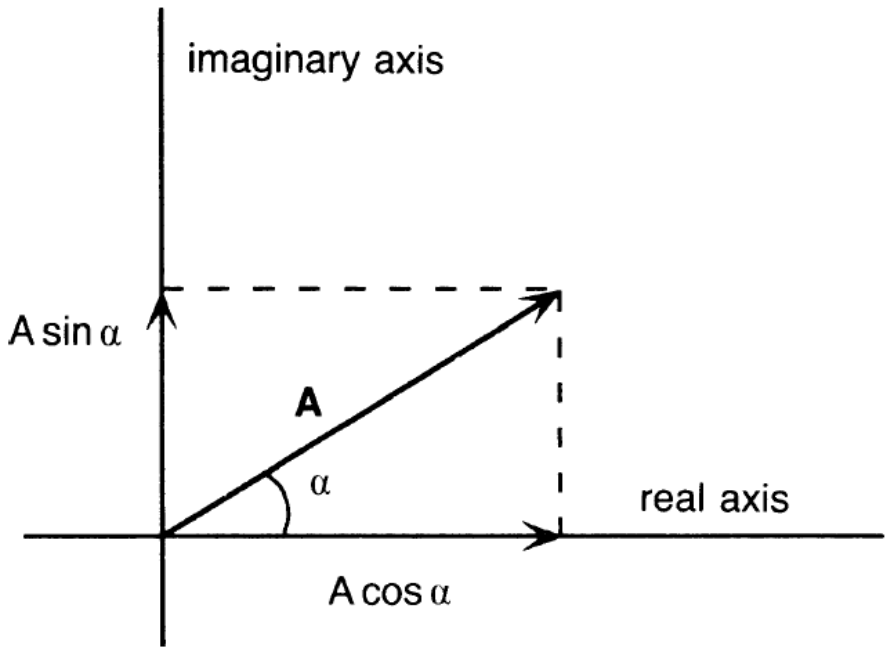


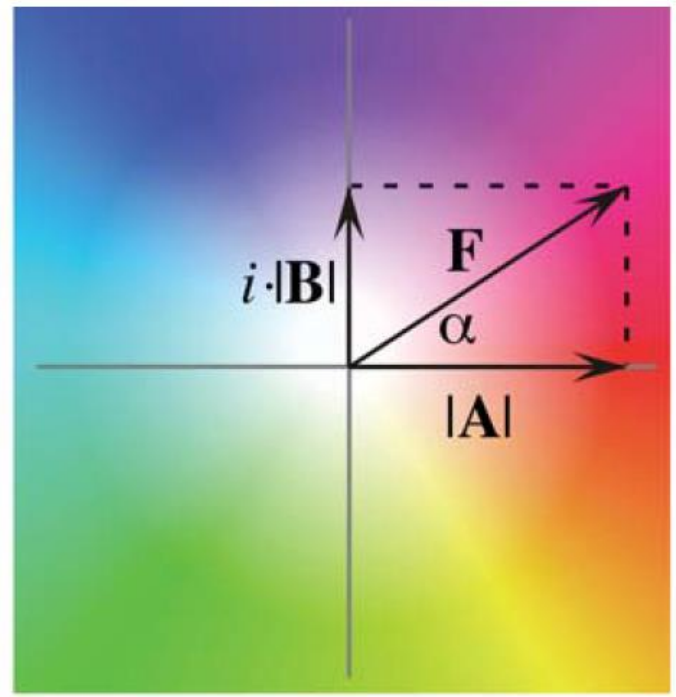
$$A = |A|(\cos\ \alpha + \ i\ sin\ \alpha)$$

$$A = |A| \exp i\alpha$$

$$F_{(hkl)} = \left|F_{(hkl)}\right| (cos\ \alpha_{(hkl)} + isin\ \alpha_{(hkl)})$$

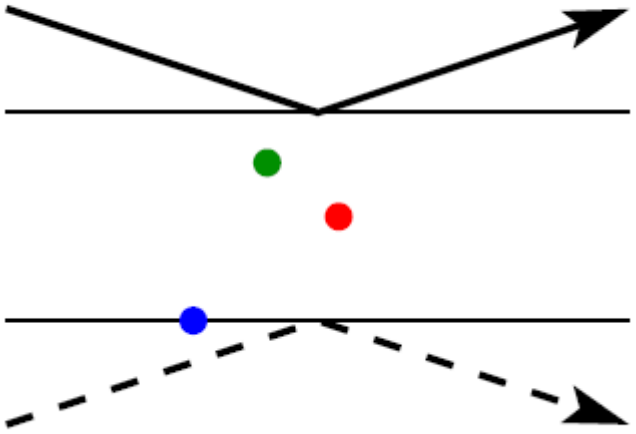
$$F_{(hkl)} = \left|F_{(hkl)}\right| exp\ i\alpha_{(hkl)}$$

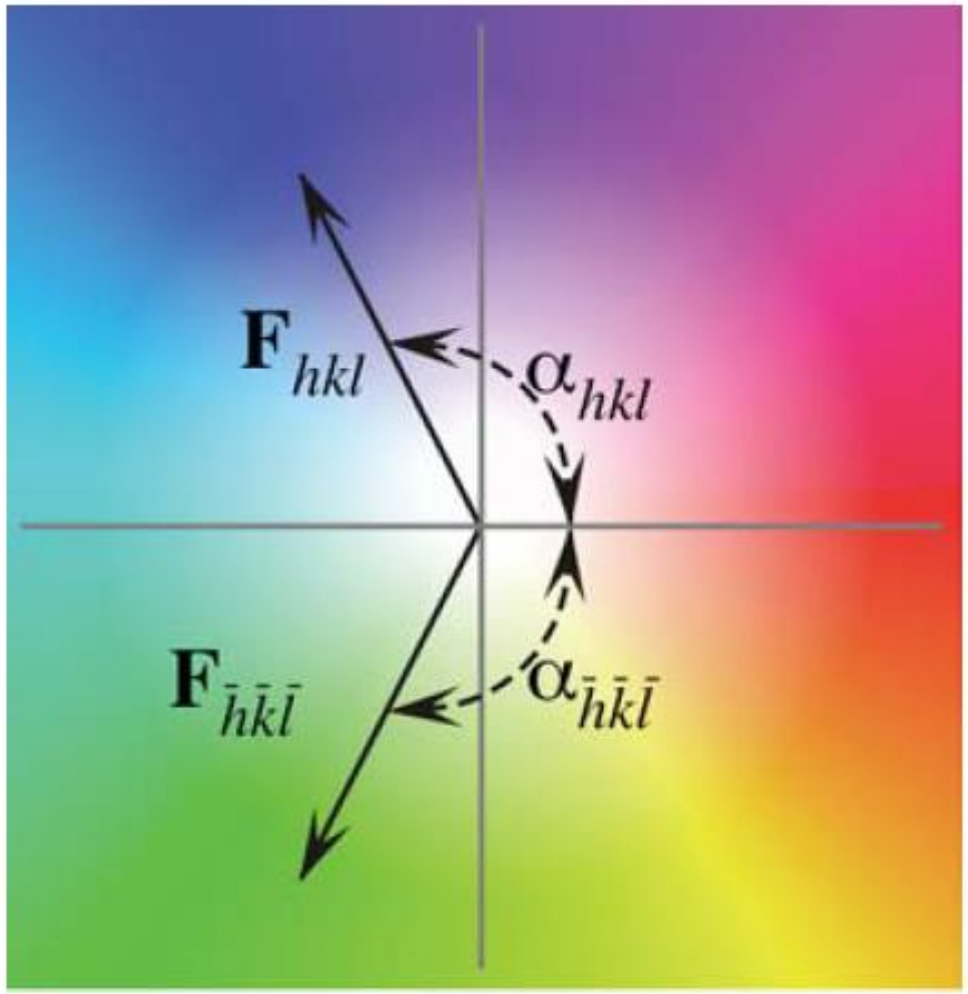

imaginary axis
A sin α
A
α
real axis
A cos α


i·|B|
F
α
|A|

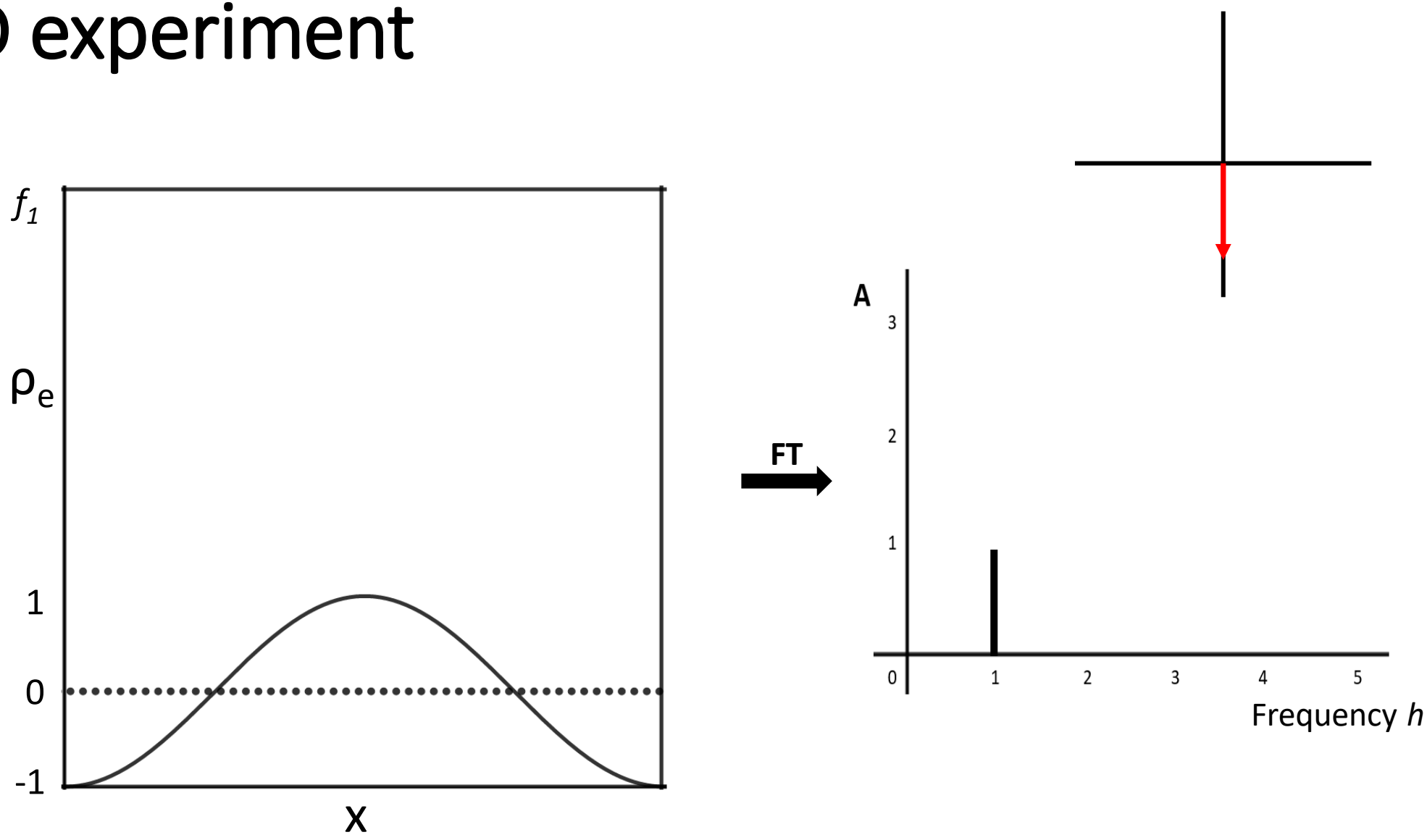
Friedel´s law – Friedel´s pairs

- $\alpha_{\bar{h}\bar{k}\bar{l}} = 360 - \alpha_{hkl}$
- $|F_{\bar{h}\bar{k}\bar{l}}| = |F_{hkl}|$

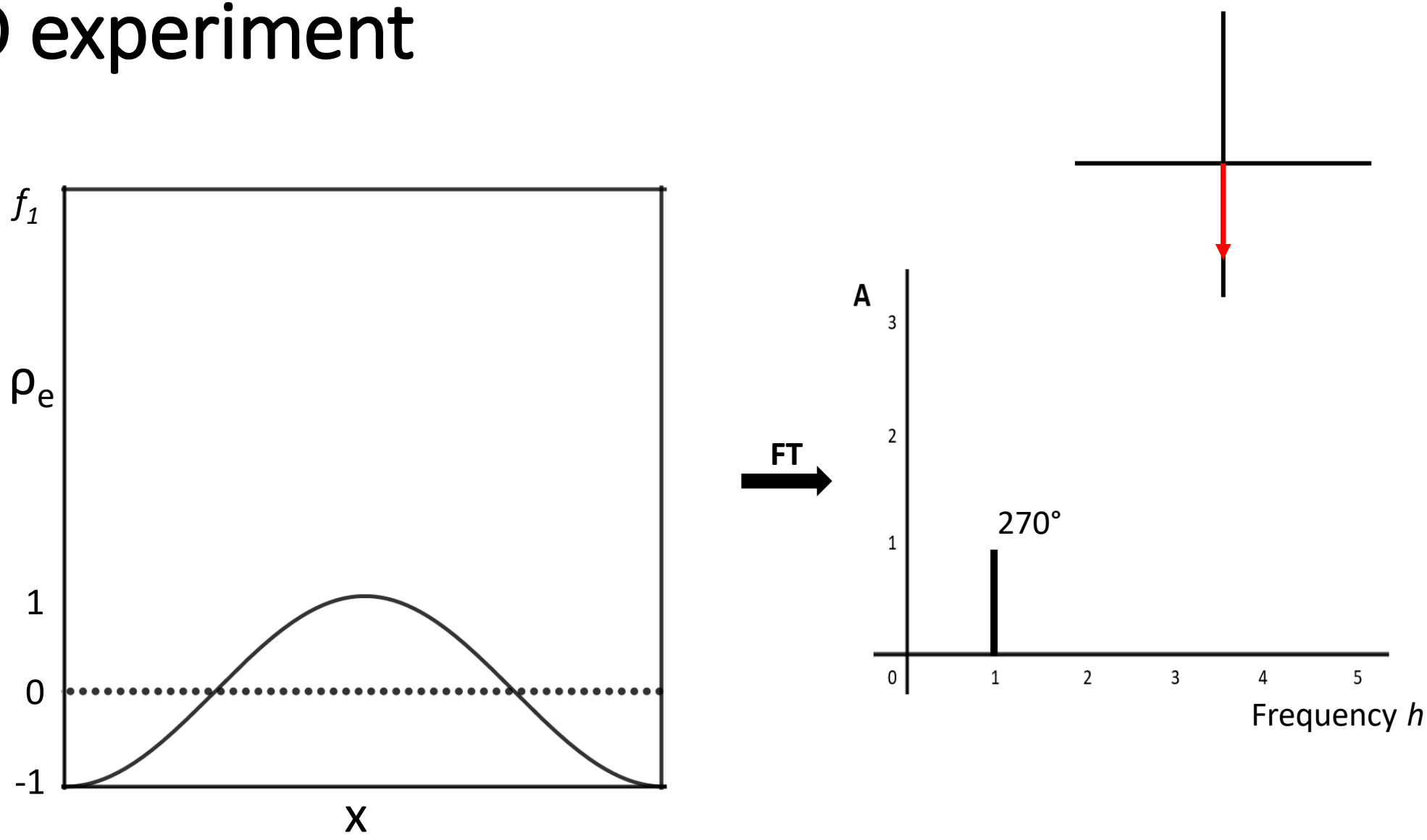




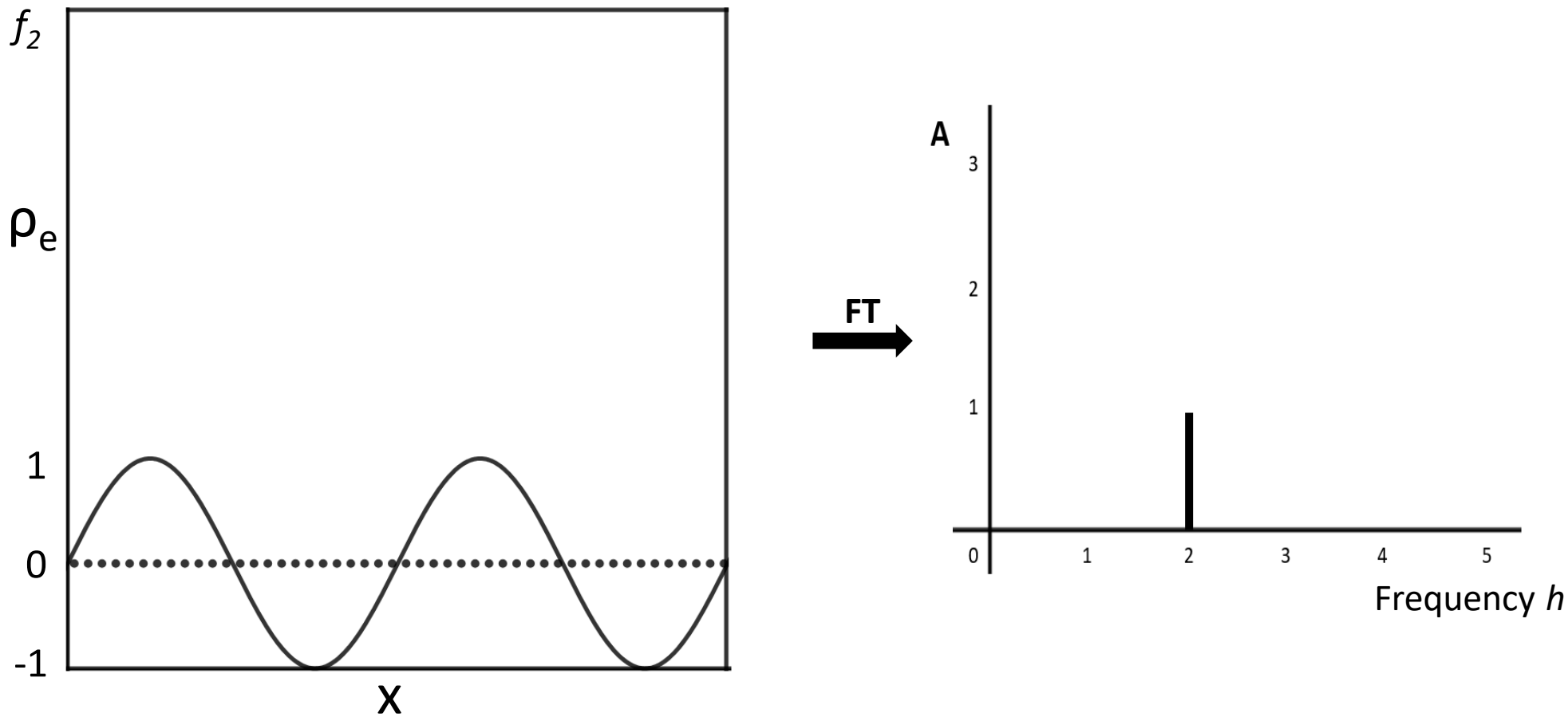
1D experiment



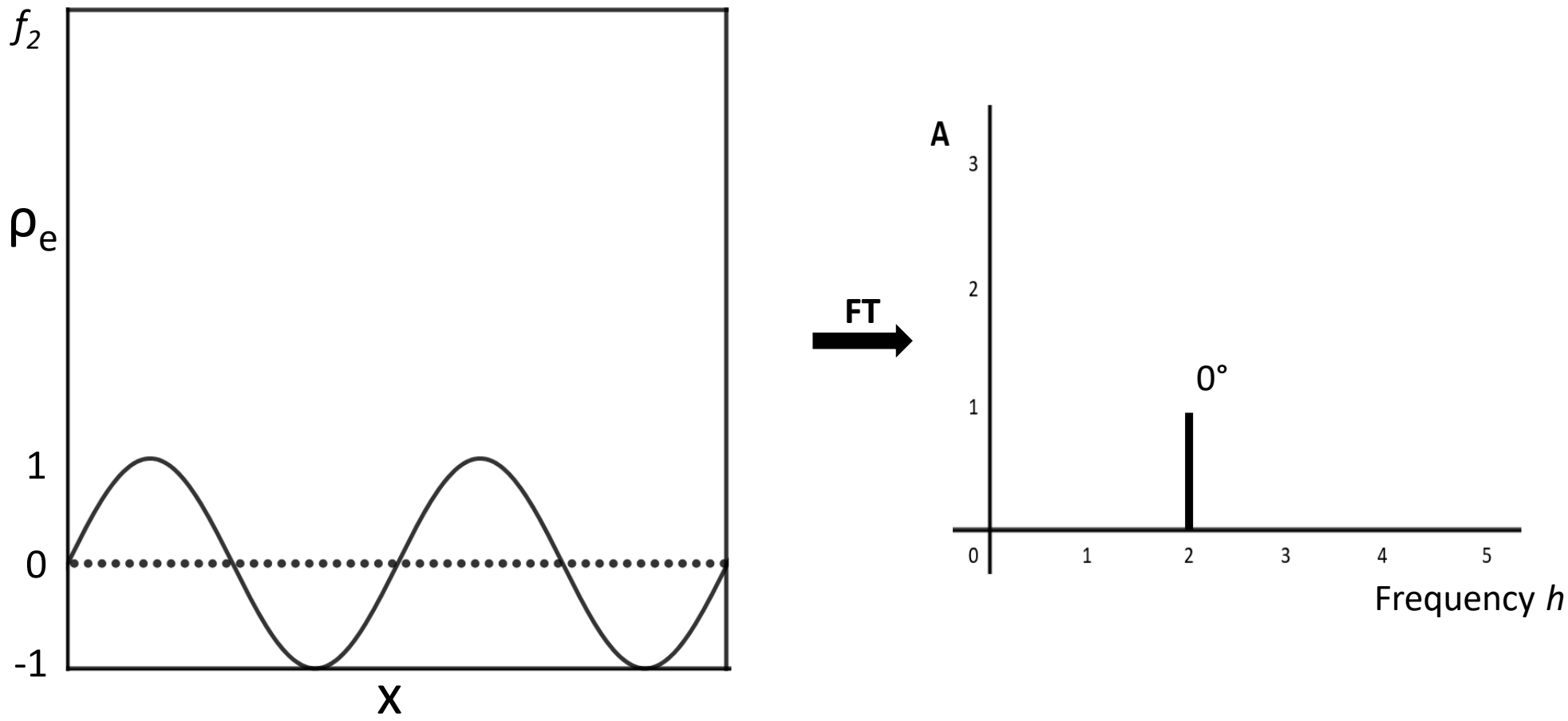
1D experiment



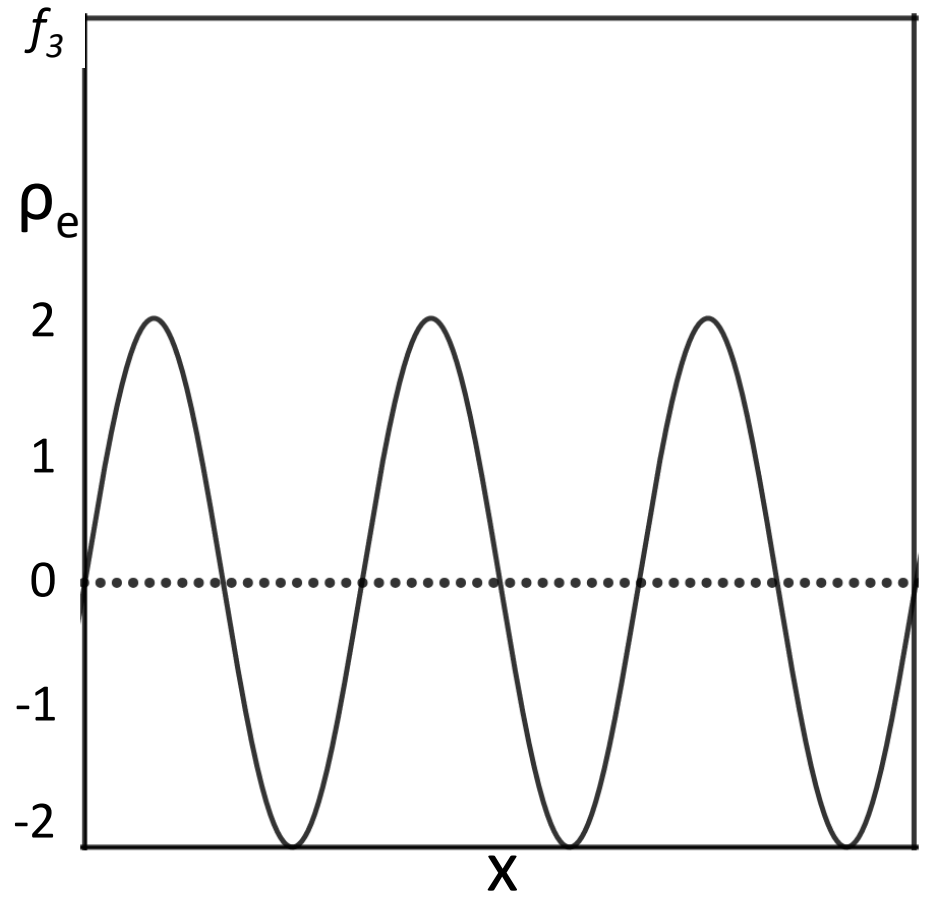
1D experiment



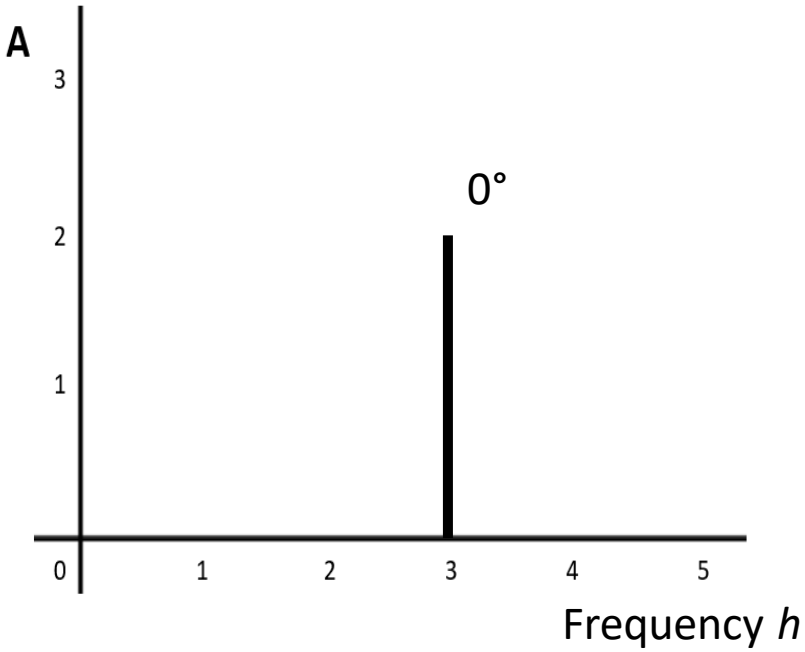
1D experiment



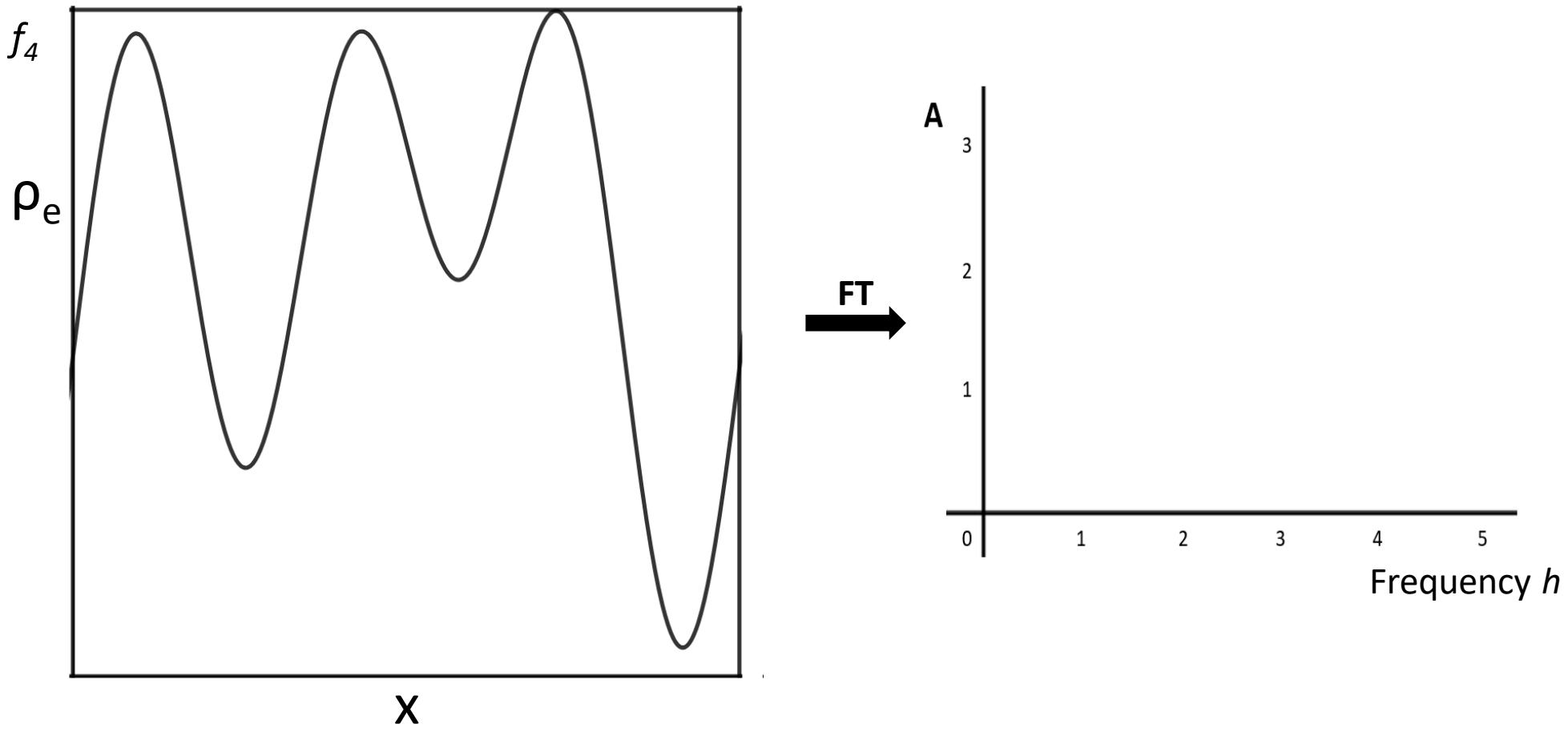
1D experiment

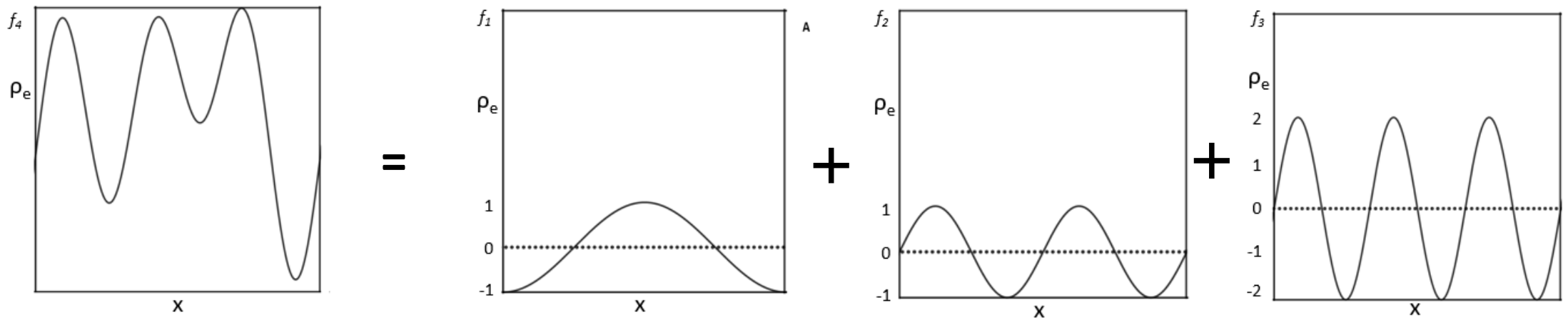


FT

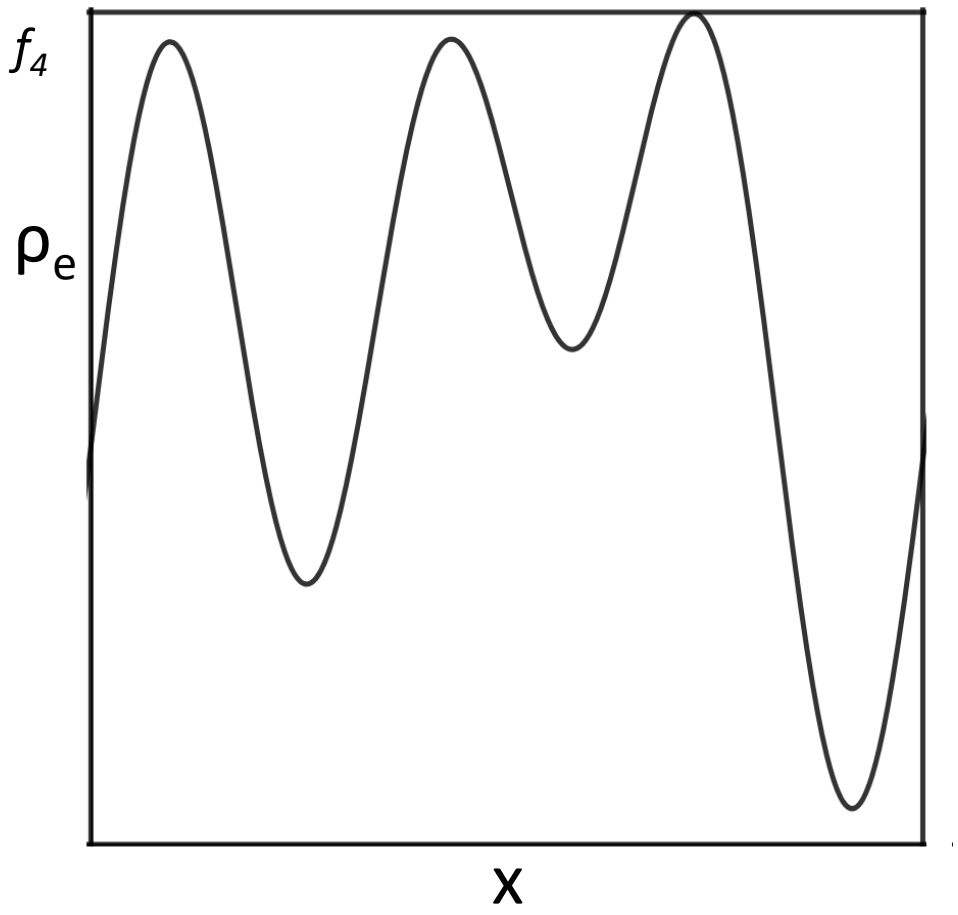


1D experiment

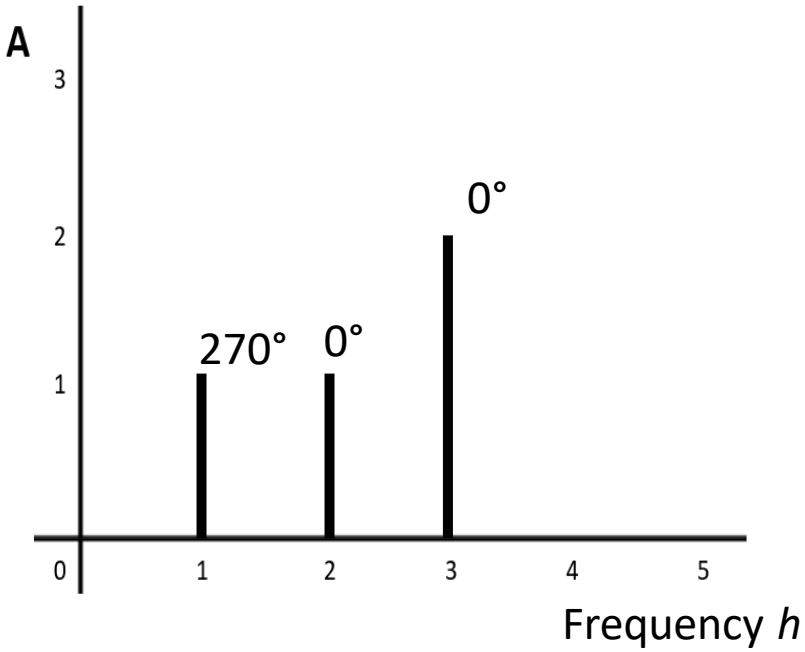



f4
ρe
x
=
f1
ρe
1
0
-1
x
A
+
f2
ρe
1
0
-1
x
+
f3
ρe
2
1
0
-1
-2
x

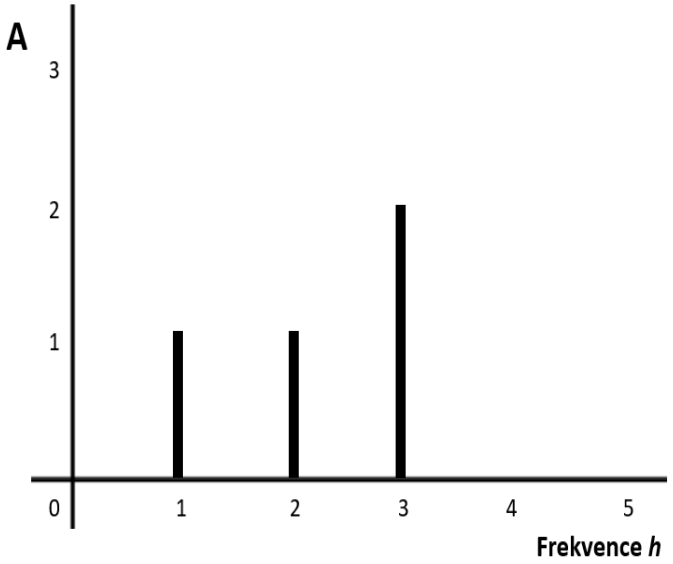
1D experiment

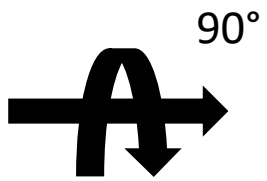


FT

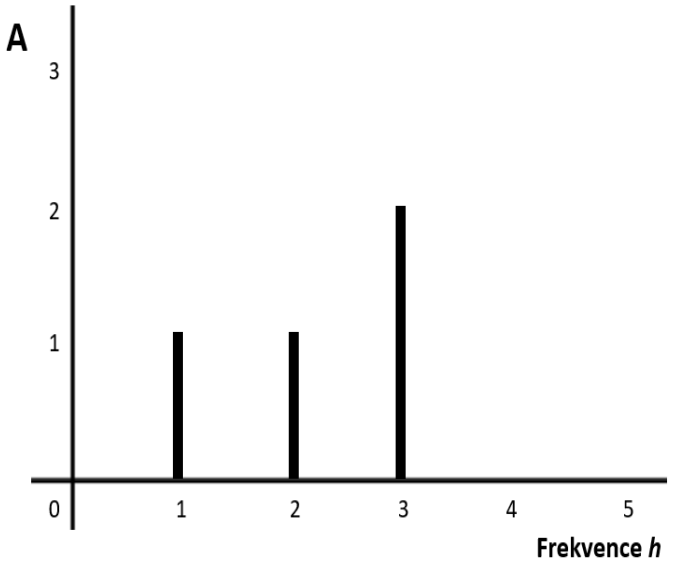


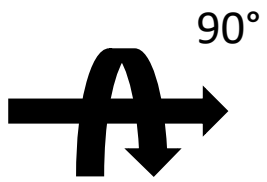
1D experiment to 2D experiment

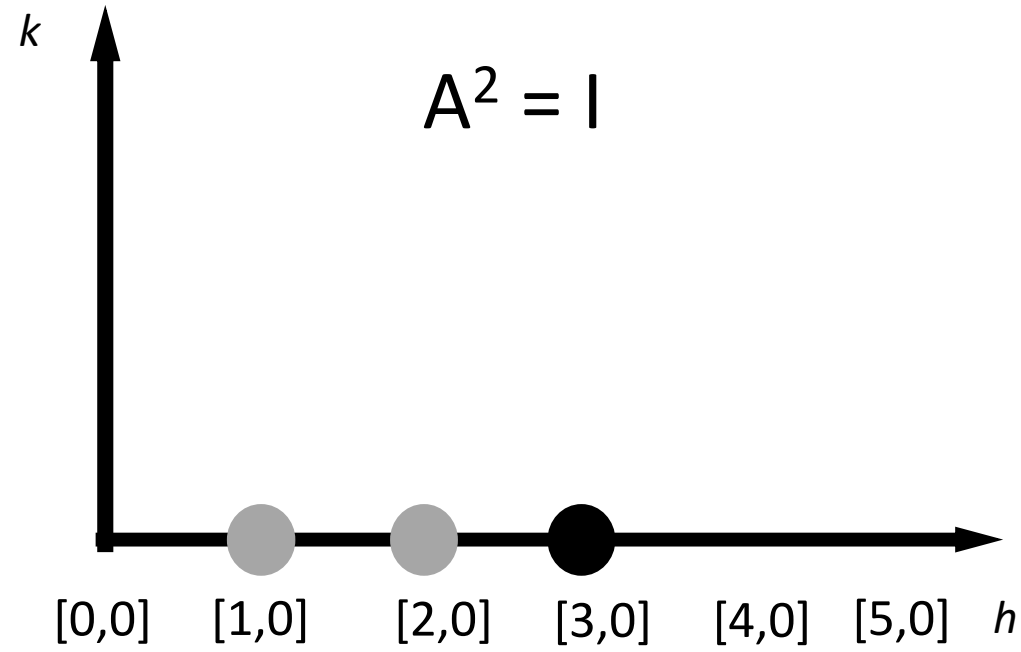


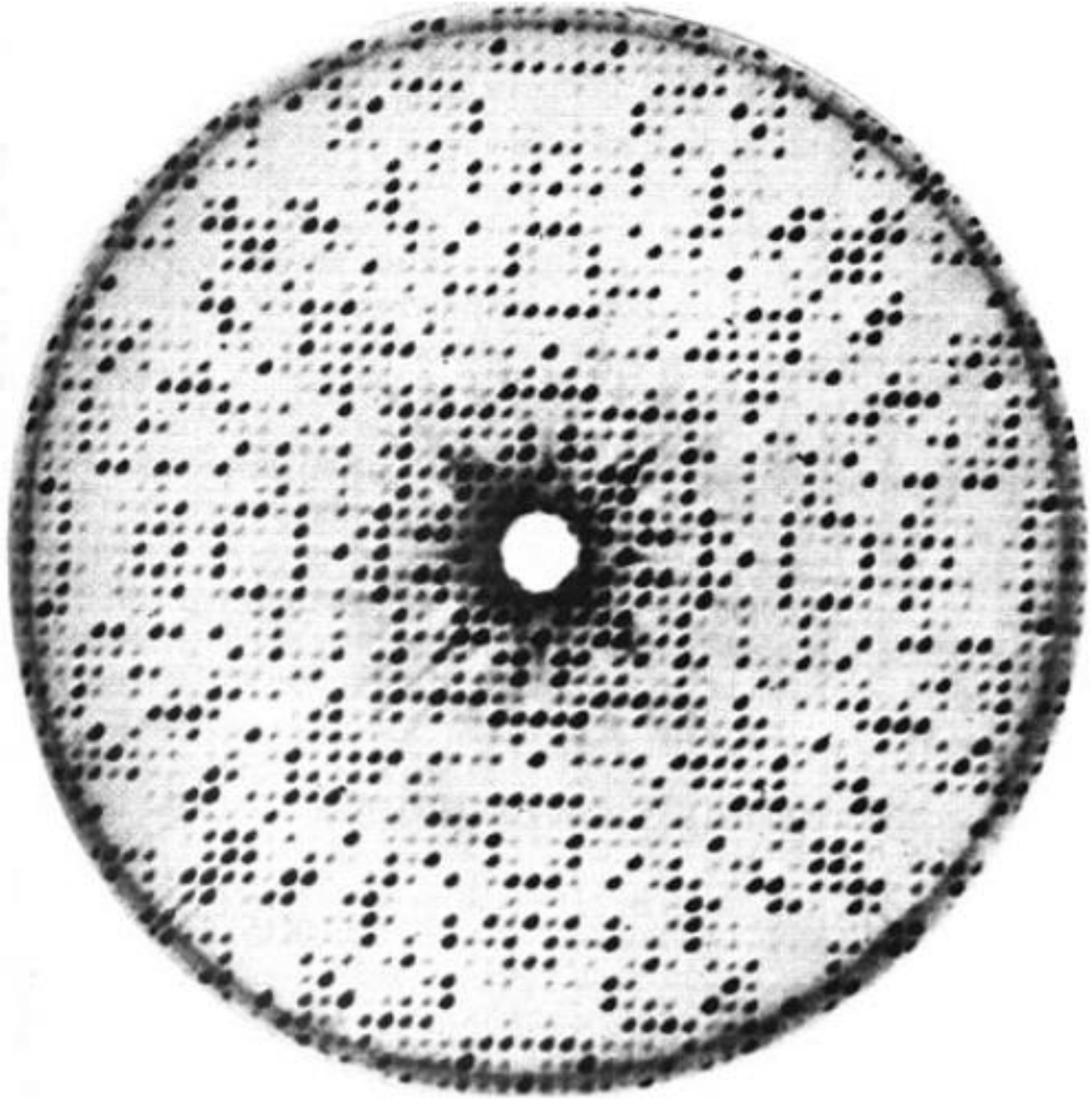


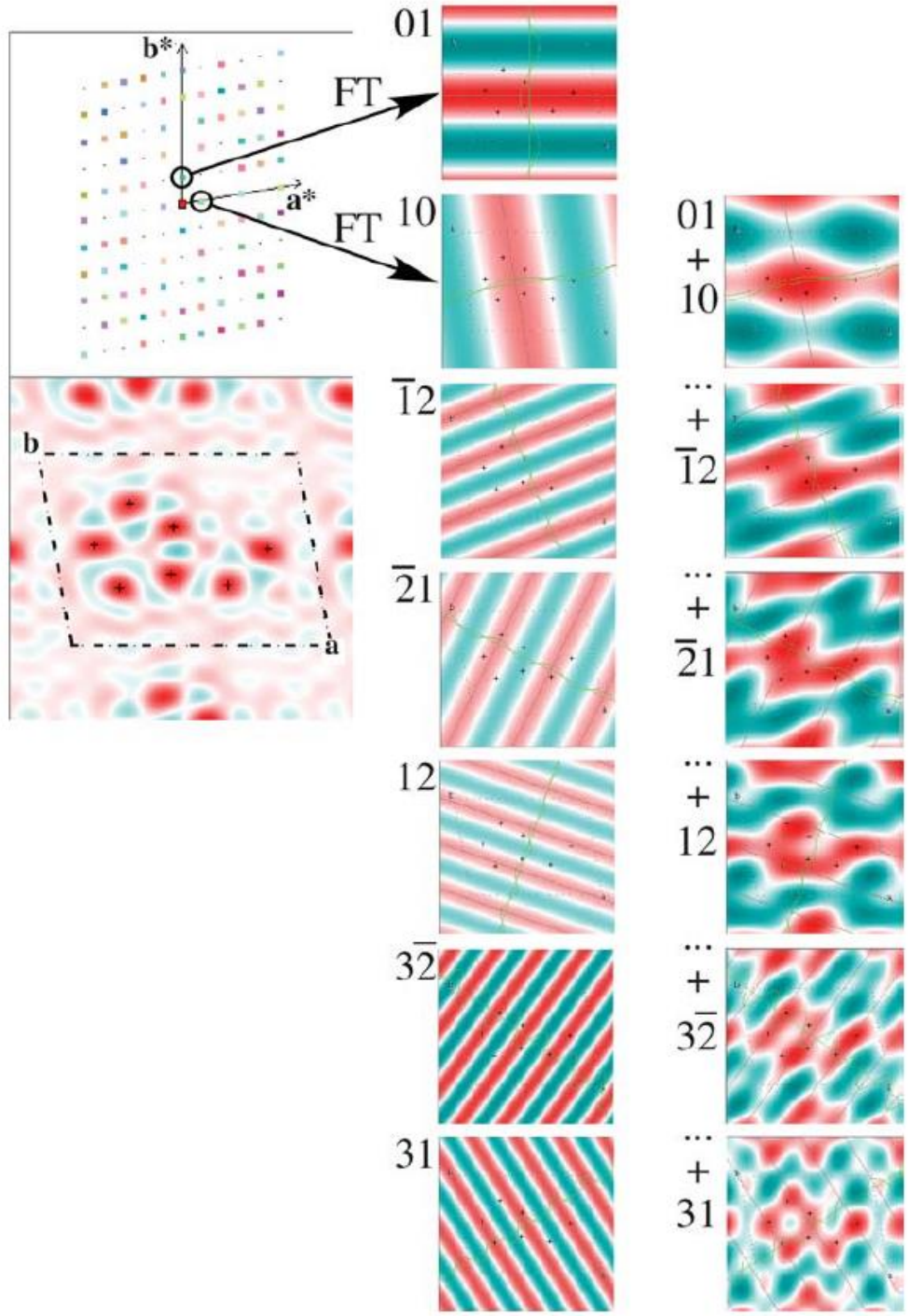
1D experiment to 2D experiment










b*
a*
FT
FT
01
10
b
a
$\bar{1}2$
$\bar{2}1$
12
$3\bar{2}$
31
01
+
10
…
+
$\bar{1}2$
…
+
$\bar{2}1$
…
+
12
…
+
$3\bar{2}$
…
+
31

Mathematic description

$$F_{(hkl)} = V_C \int_{x=0}^{1} \int_{y=0}^{1} \int_{z=0}^{1} \rho_{(xyz)}\, exp\,[2\pi i(hx + ky + lz)]dxdydz$$

FT ⇅ FT^{-1}

$$\rho_{(xyz)} = \frac{1}{V_C} \sum_{h} \sum_{k} \sum_{l} \left|F_{(hkl)}\right| exp\left[-2\pi i(hx + ky + lz) + i\alpha_{(hkl)}\right]$$

Mathematic description

$$F_{(hkl)} = V_C \int_{x=0}^{1} \int_{y=0}^{1} \int_{z=0}^{1} \rho_{(xyz)}\, exp\,[2\pi i(hx + ky + lz)]dxdydz$$

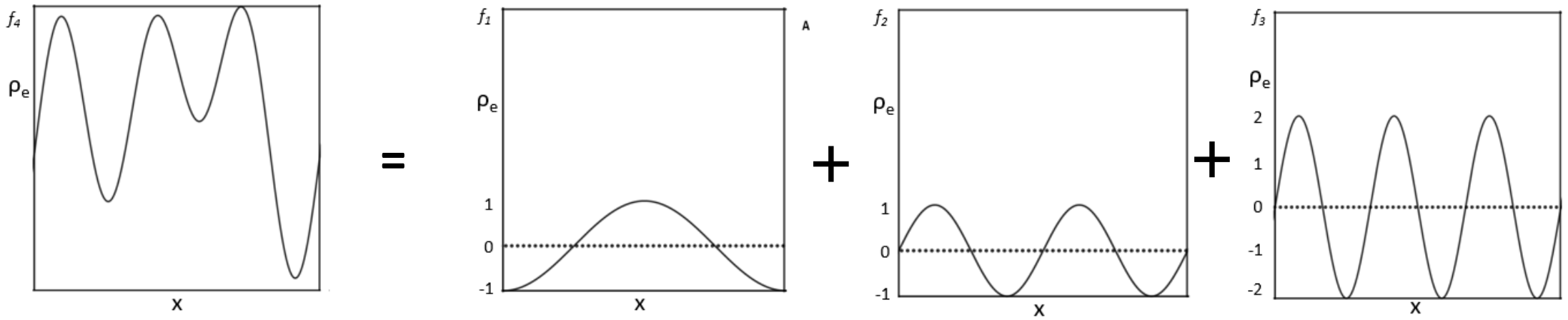
$$F_{(hkl)} = \left|F_{(hkl)}\right| exp\, i\alpha_{(hkl)}$$

FT ⇅ FT^{-1}

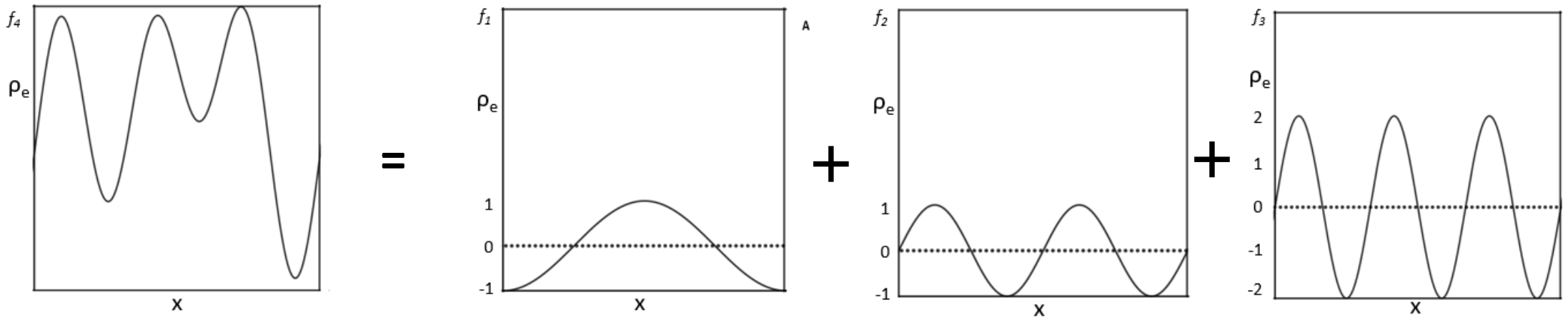
$$\rho_{(xyz)} = \frac{1}{V_C} \sum_{h} \sum_{k} \sum_{l} \left|F_{(hkl)}\right| exp\left[-2\pi i(hx + ky + lz) + i\alpha_{(hkl)}\right]$$

Mathematic description: 1D

$$\rho_{(x)} = \frac{1}{l}\sum_{n=1}^{N} |F_h| \ \exp\left[-2\pi i \ hx + \ i\alpha_{(hkl)}\right]$$

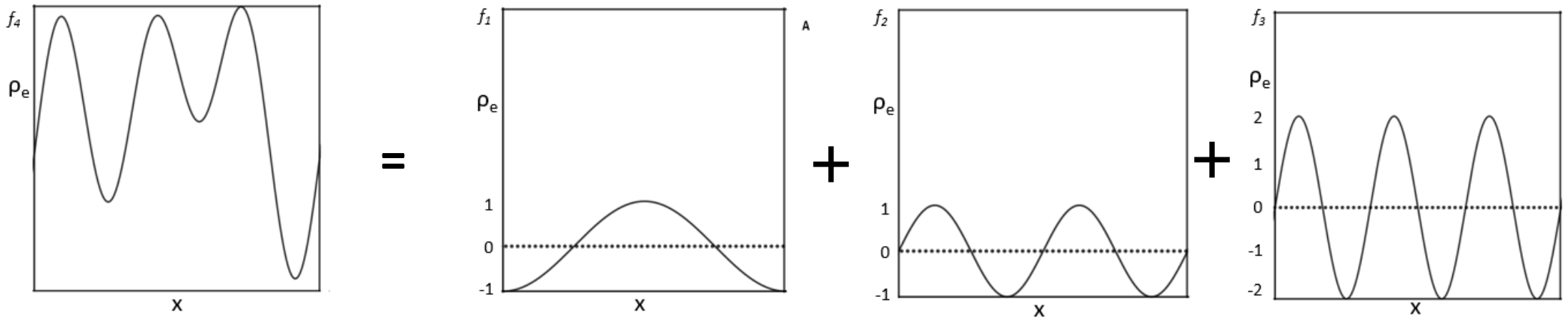


Electron density ⟵ $$\rho_{(x)} = \frac{1}{l}\sum_{n=1}^{N} |F_h| \left[-2\pi i\, hx + i\alpha_{(hkl)}\right]$$



Sum of all structural factors

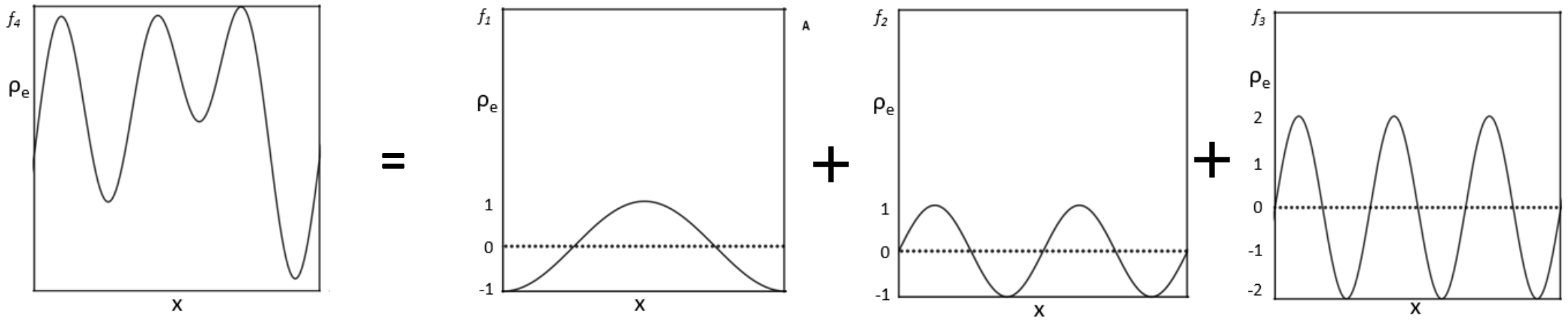
Electron density ⟵ $\rho_{(x)} = \frac{1}{l} \sum_{n=1}^{N} |F_h| \left[-2\pi i\, hx + i\alpha_{(hkl)}\right]$



Sum of all structural factors

Electron density ⟵ $\rho_{(x)} = \frac{1}{l} \sum_{n=1}^{N} |F_h| \left[-2\pi i\, hx + i\alpha_{(hkl)}\right]$

Amplitude



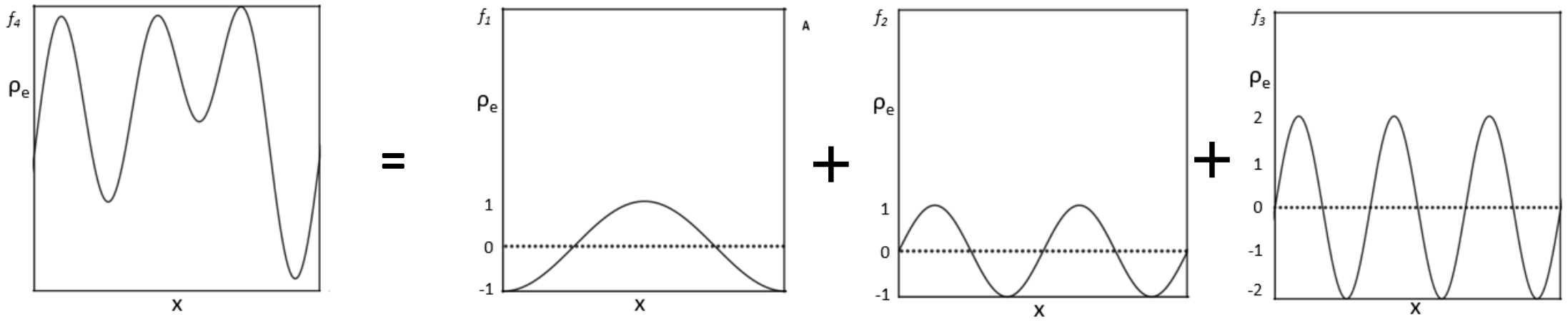
Sum of all structural factors

Frequency

Electron density ⟵

$$\rho_{(x)} = \frac{1}{l}\sum_{n=1}^{N} |F_h| \left[-2\pi i\, hx + i\alpha_{(hkl)}\right]$$

Amplitude



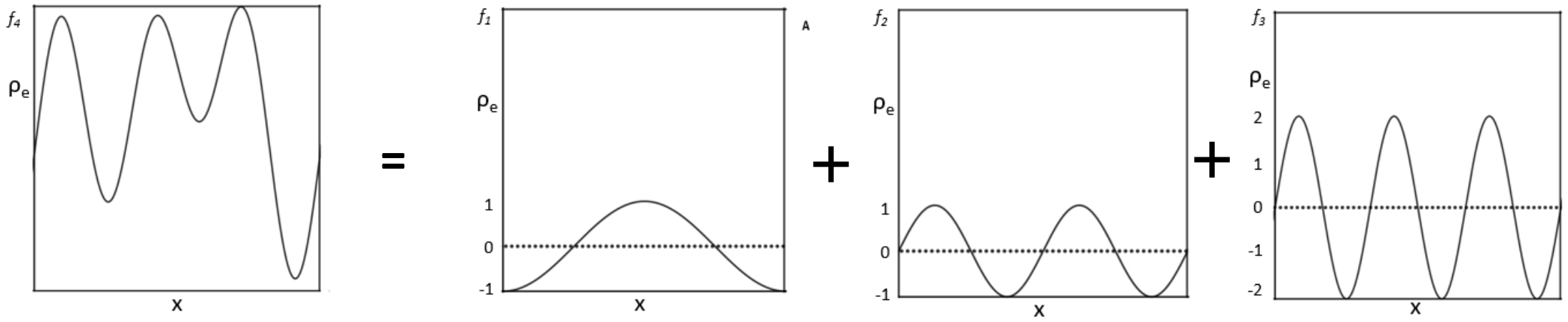
Sum of all structural factors

Frequency

Electron density ⟵ $\rho_{(x)} = \frac{1}{l} \sum_{n=1}^{N} |F_h| \left[-2\pi i\, hx + i\alpha_{(hkl)}\right]$

Amplitude

Phase



Sum of all structural factors

Frequency

Electron density ⟵

$$\rho_{(x)} = \frac{1}{l} \sum_{n=1}^{N} |F_h| \left[-2\pi i\, hx + i\alpha_{(hkl)} \right]$$

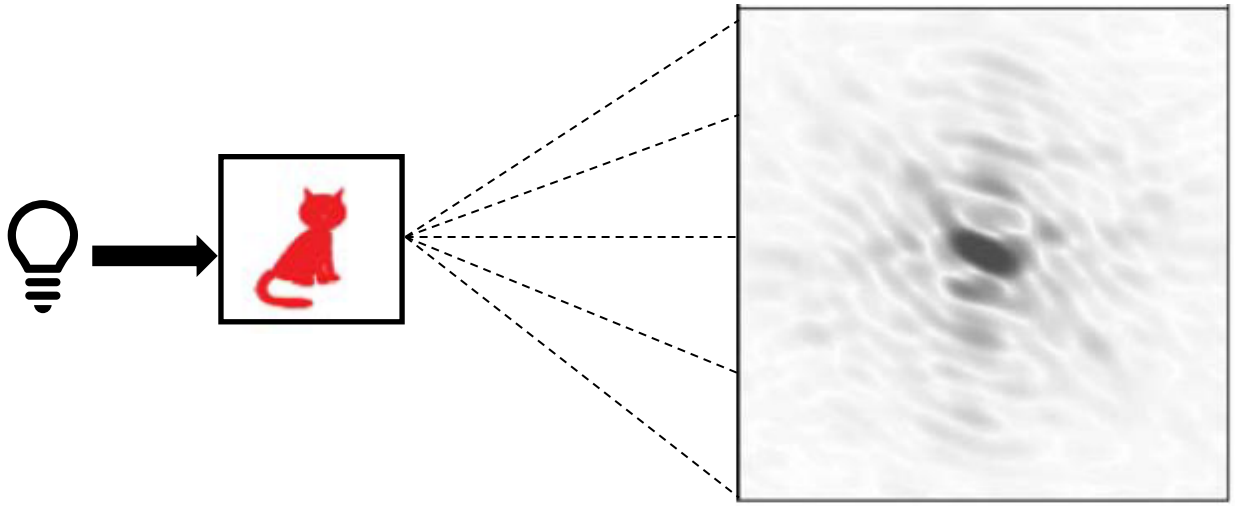
Length of unit cell **Amplitude** **Phase**

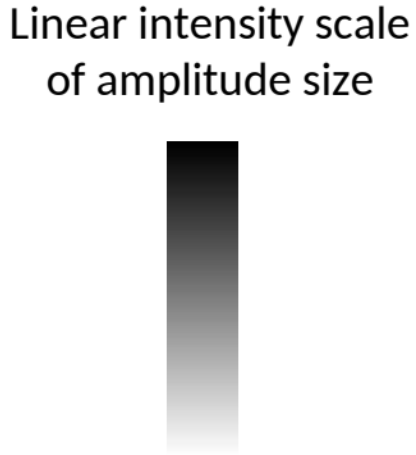
Methods for phasing:

- Molecular replacement (70 %)
- Isomorphous replacement
- Anomalous scattering

Molecular replacement

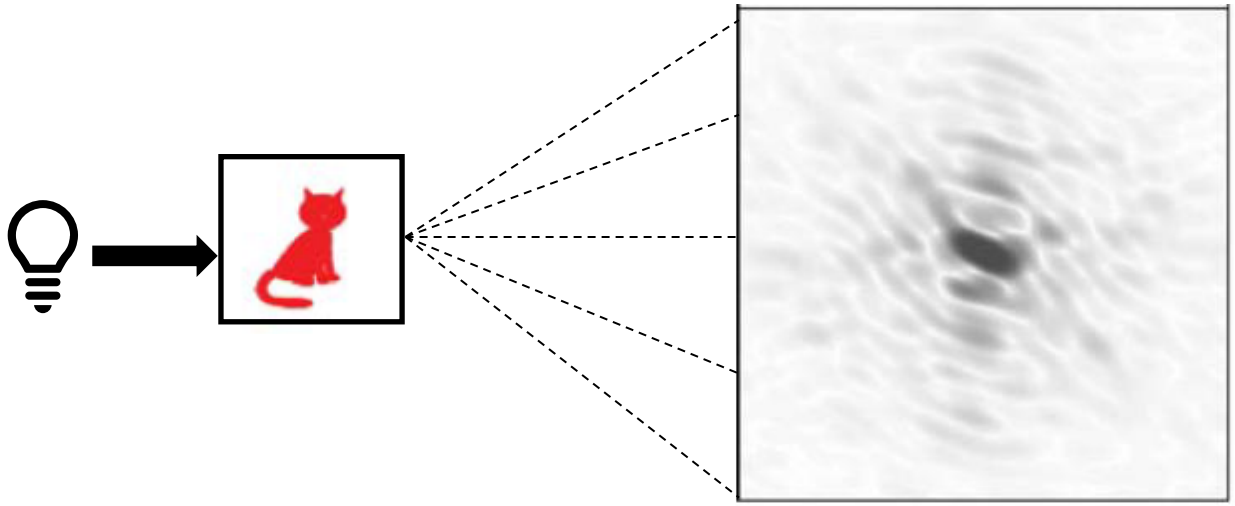
Molecular replacement - basics



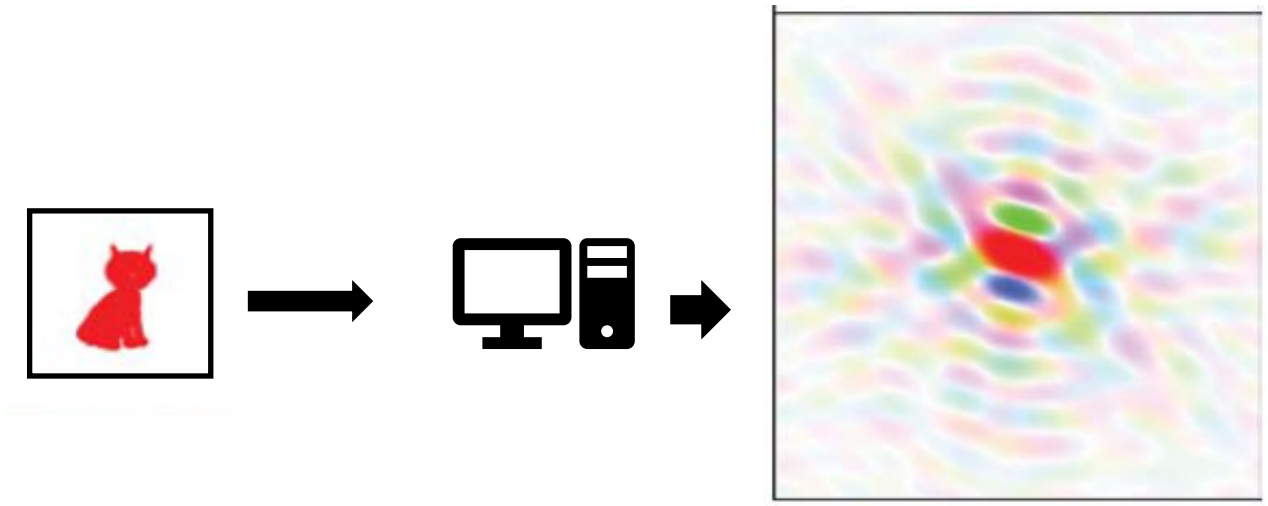


http://www.ysbl.york.ac.uk/~cowtan/fourier/magic.html

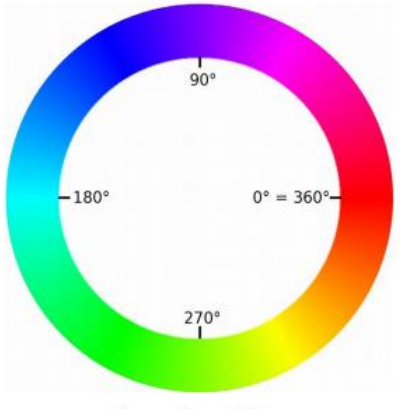
Molecular replacement - basics



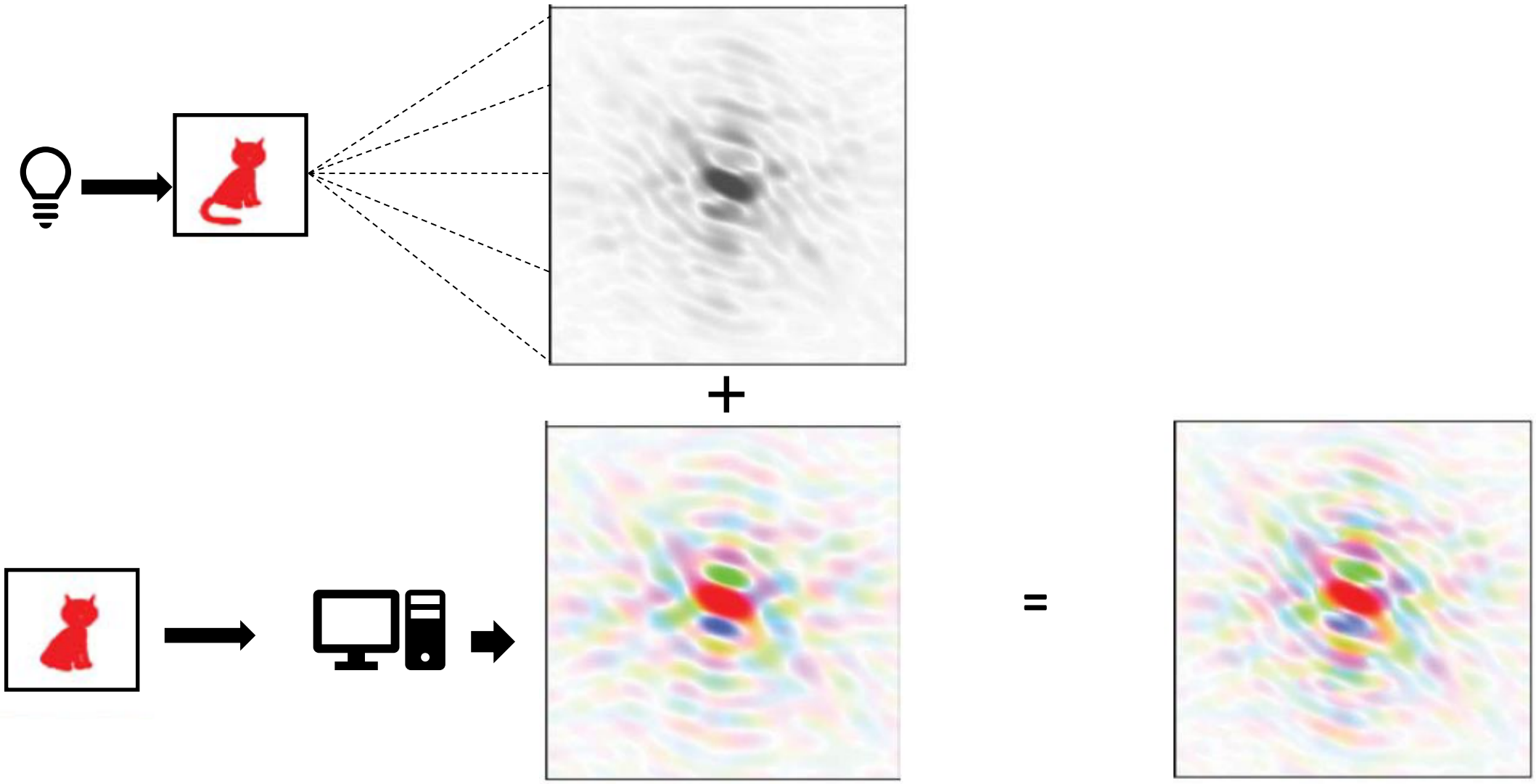
Linear intensity scale of amplitude size



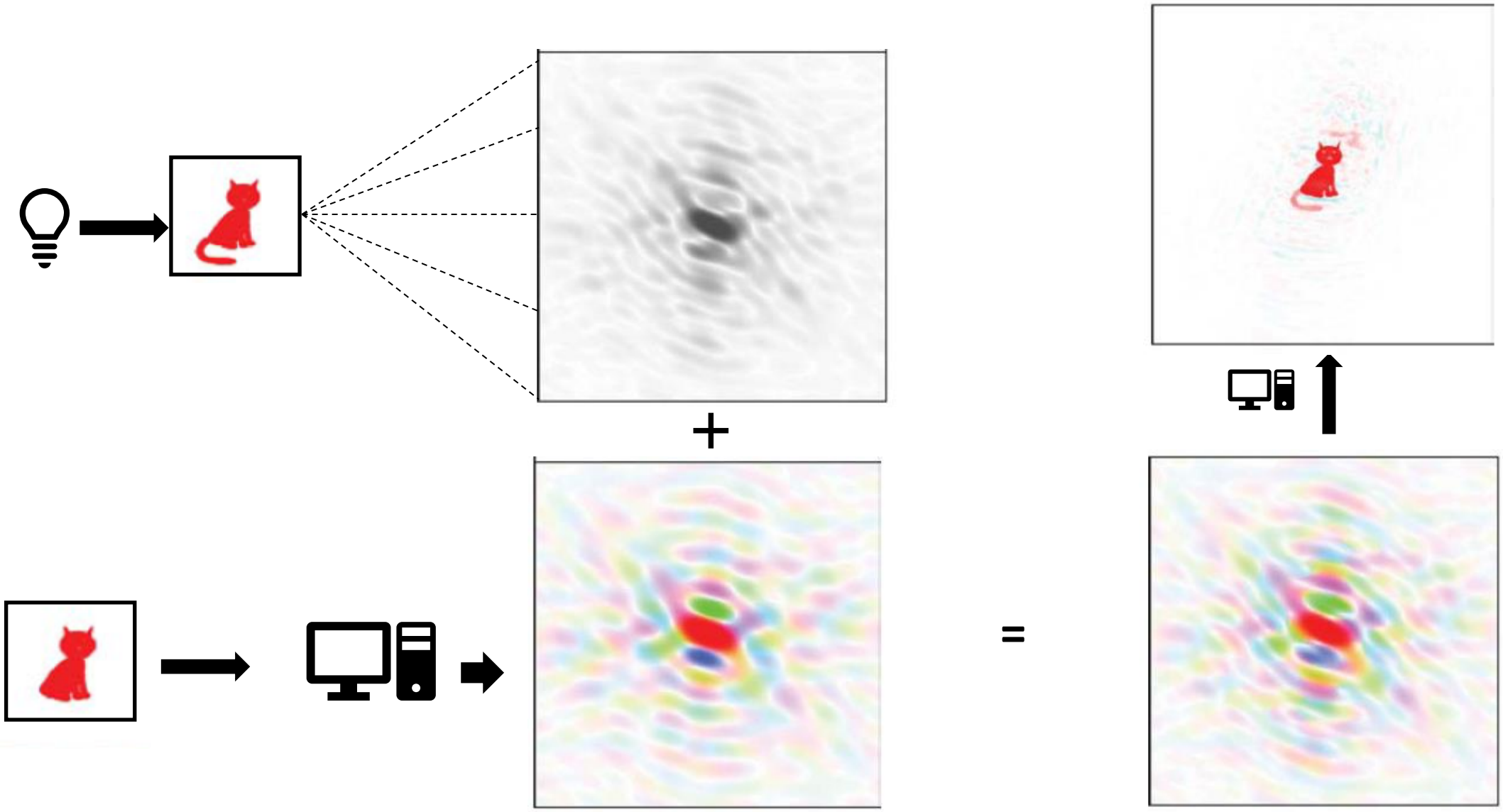
Circular rainbow scale of phases



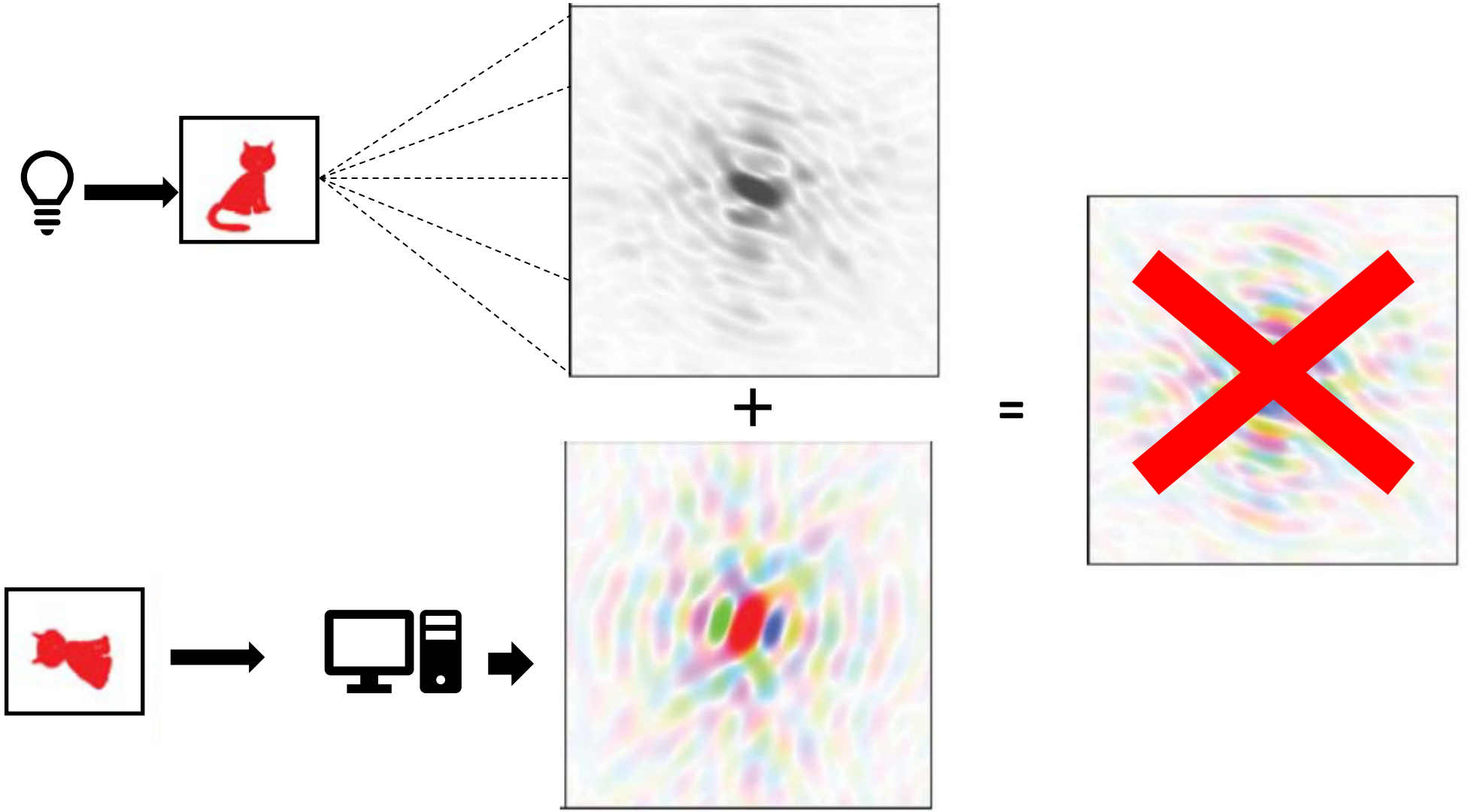
Molecular replacement - basics



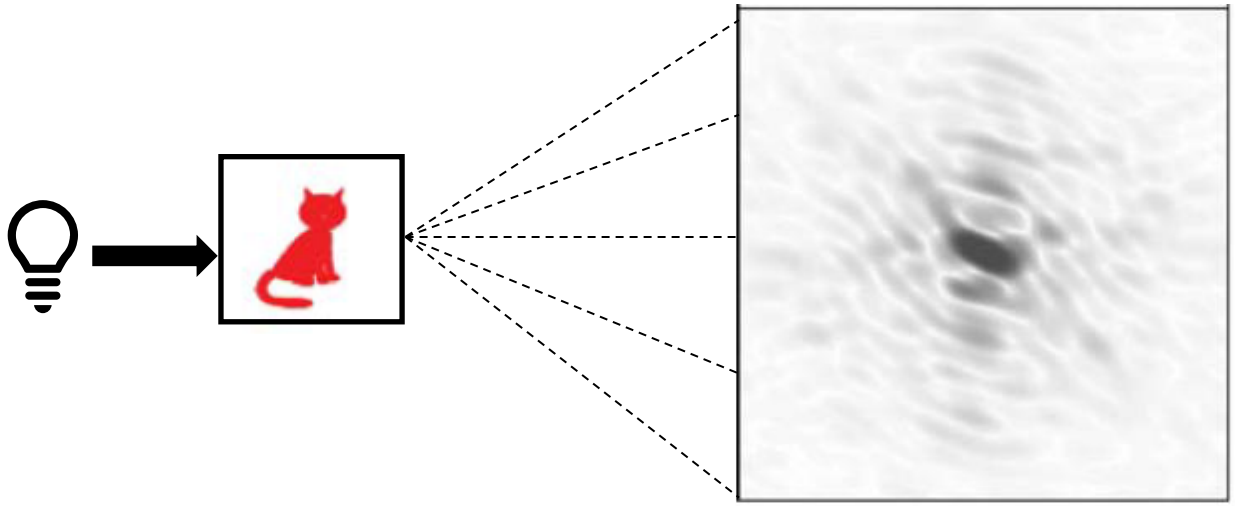
Molecular replacement - basics



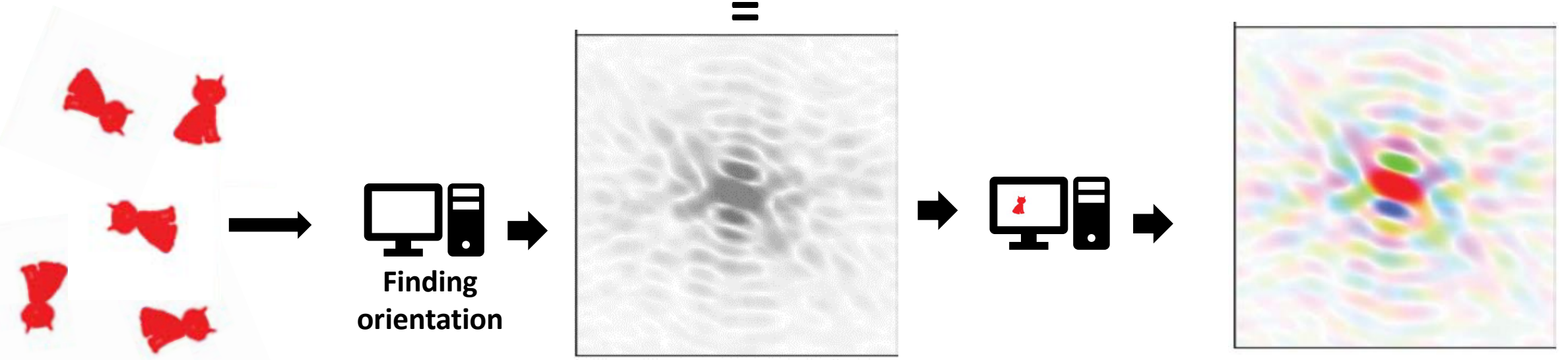
Orientation problem



Finding orientation



=



Patterson map

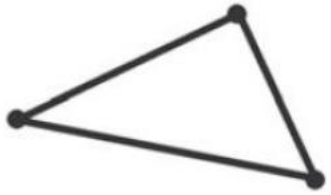
$$P_{(uvw)} = \frac{1}{V_C} \sum_h \sum_k \sum_l \left|F_{(hkl)}\right|^2 exp\,[-2\pi i(h\mathrm{u} + k\mathrm{v} + lw)\,]$$

All phases α_{hkl} are set to 0°

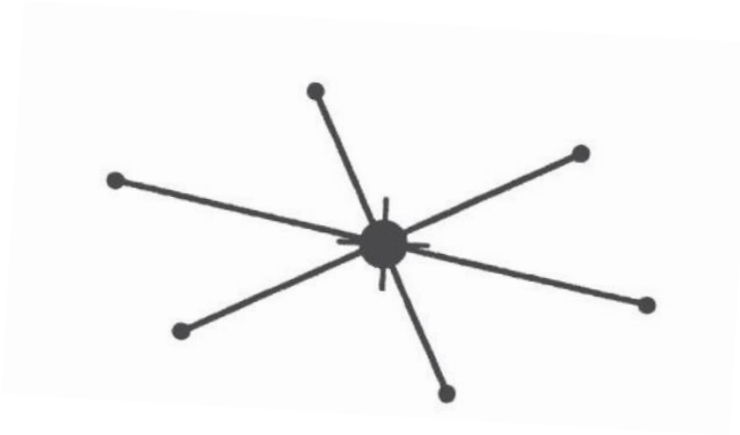
Peaks coresponds to VECTORS BETWEEN ATOMS

Patterson map

$$P_{(uvw)} = \frac{1}{V_C} \sum_h \sum_k \sum_l \left|F_{(hkl)}\right|^2 exp\,[-2\pi i(h\mathrm{u} + k\mathrm{v} + lw)\,]$$

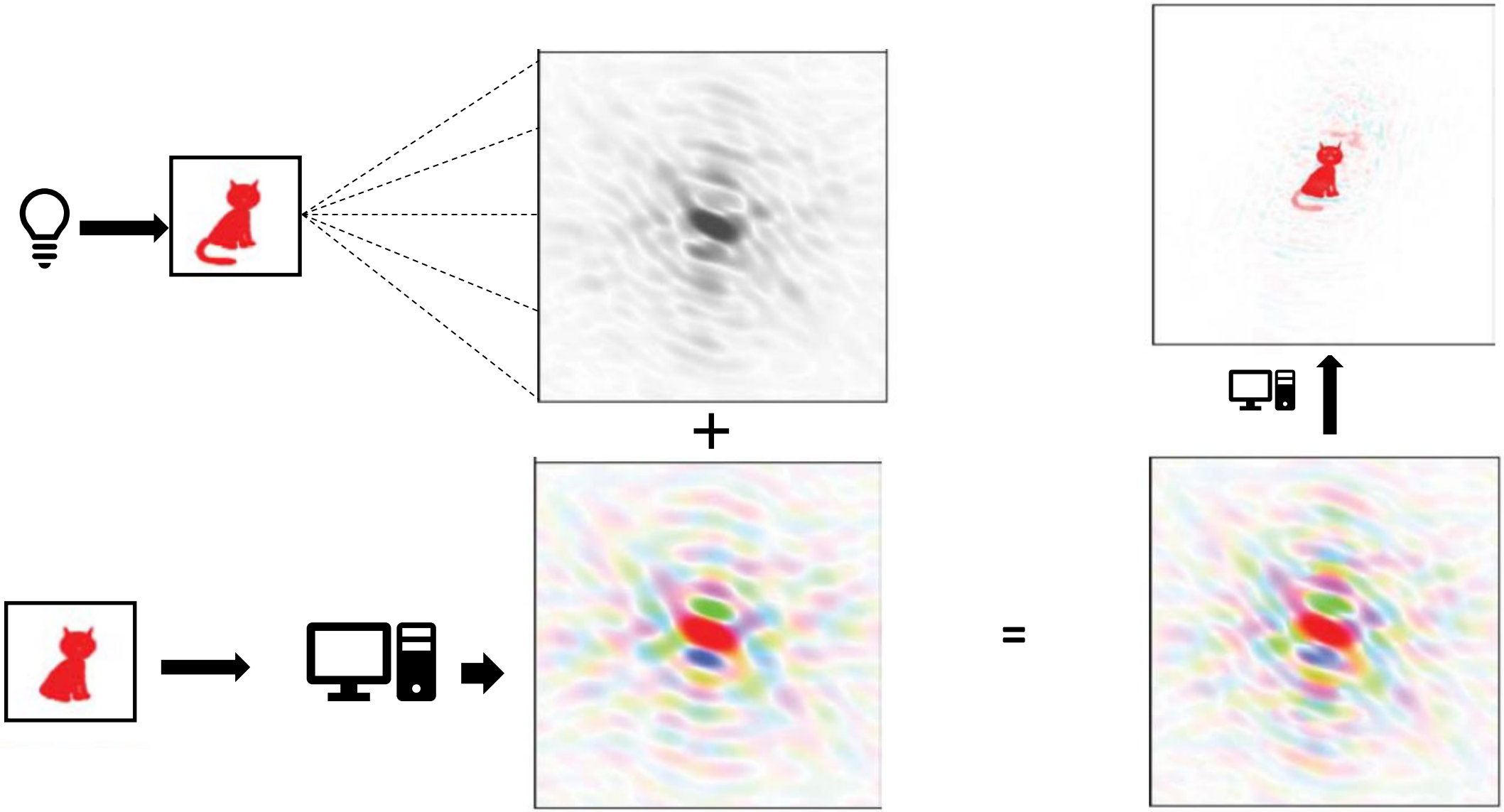


molecule

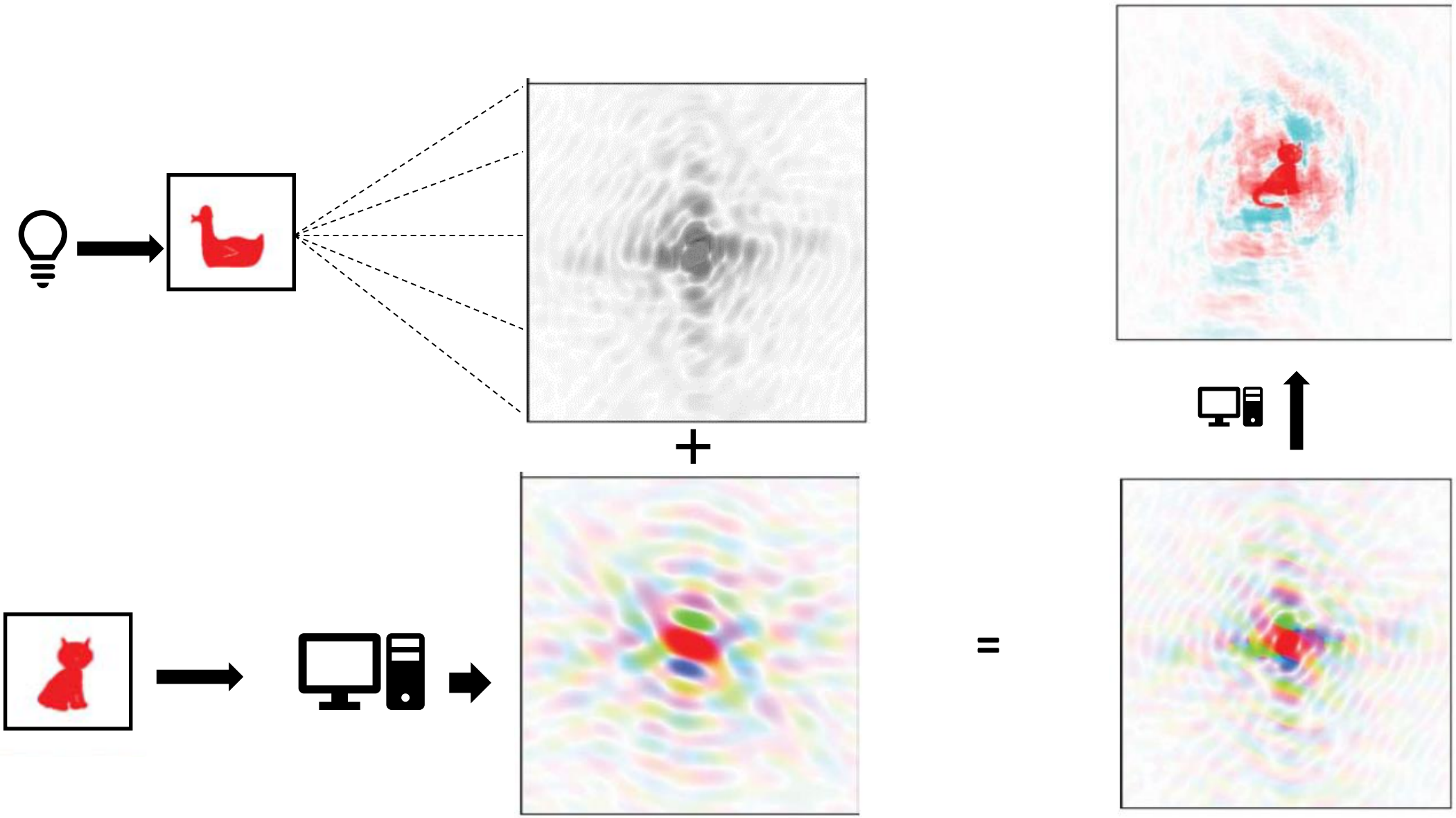


Patterson map

Molecular replacement



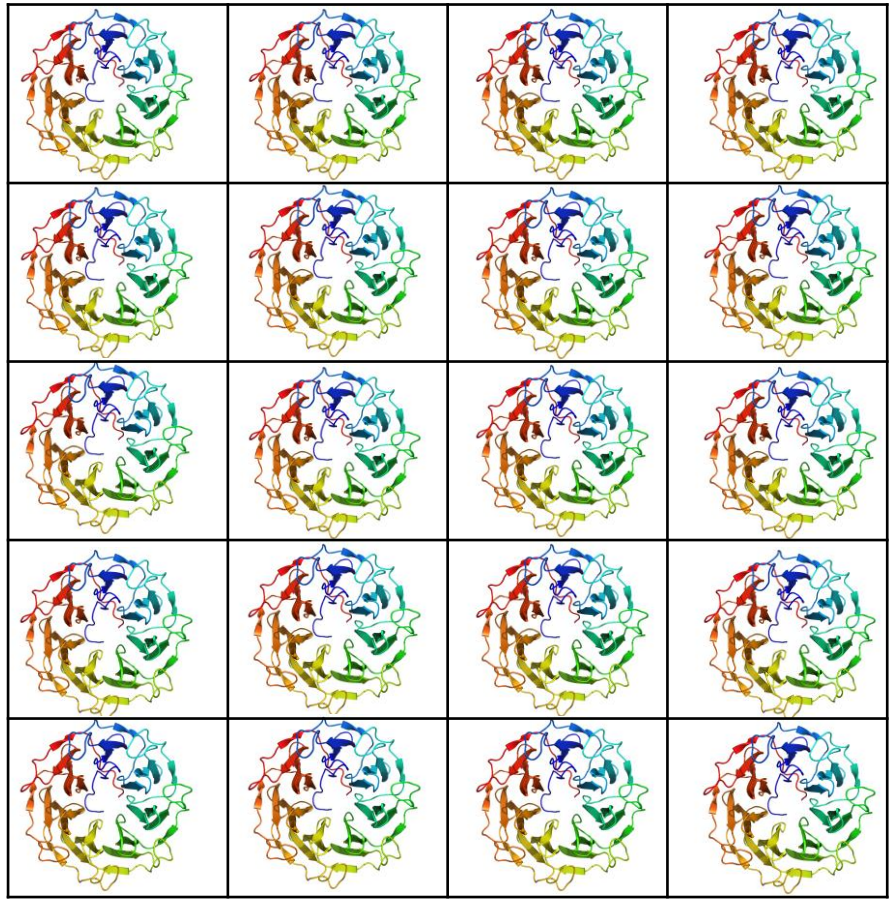
MR problems



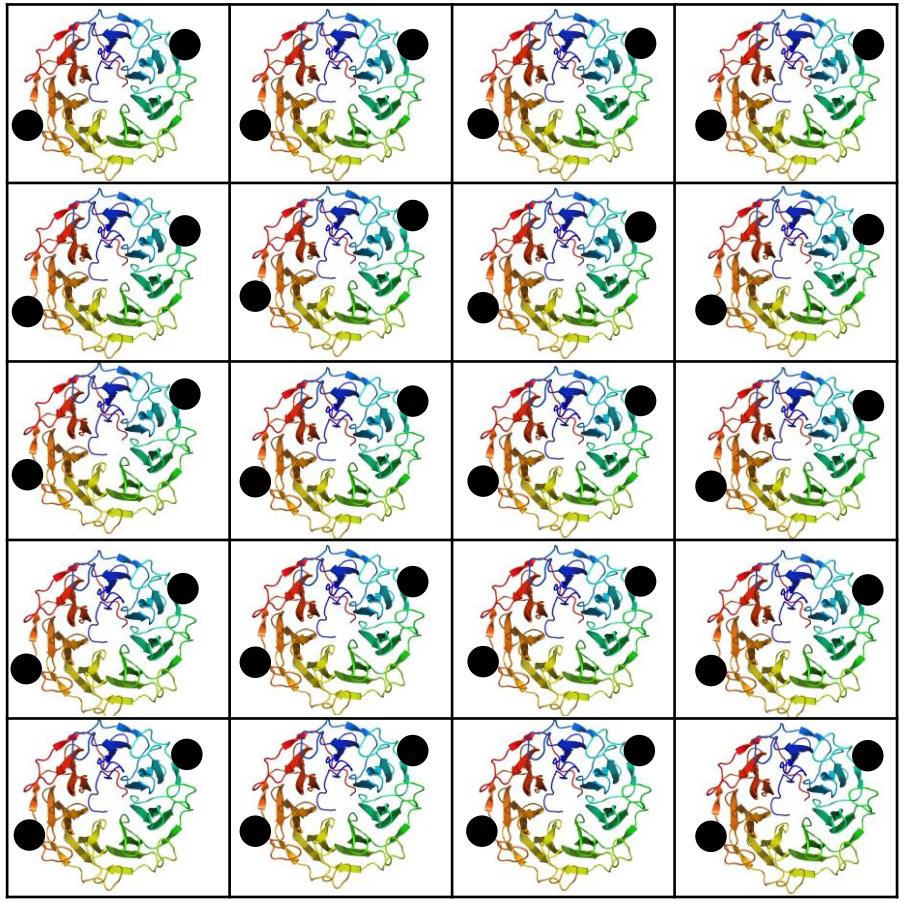
Isomorphous replacement

Uniform binding of heavy atoms

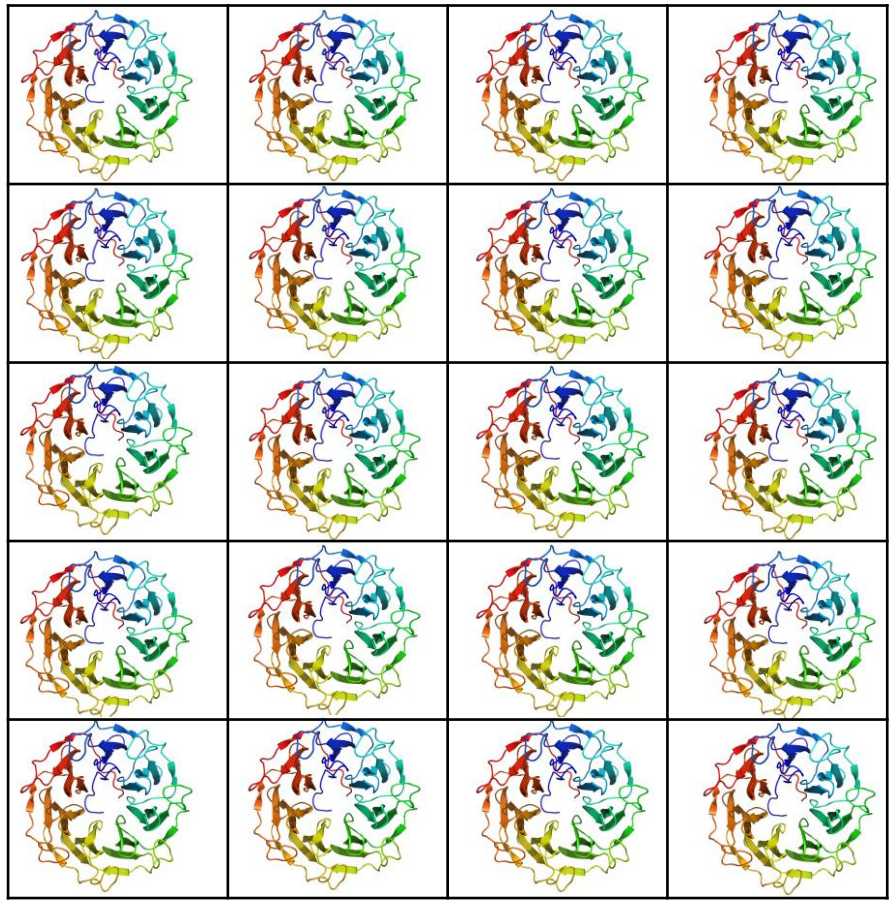
Native crystal



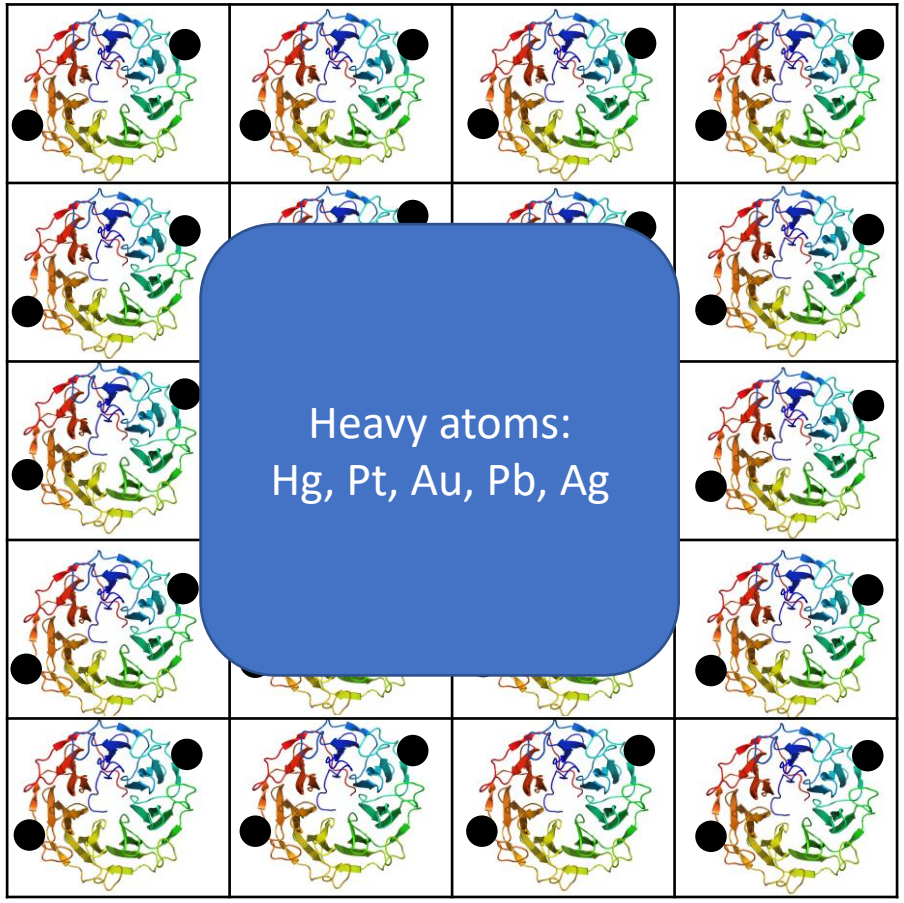
Heavy atom derivatives



Native crystal

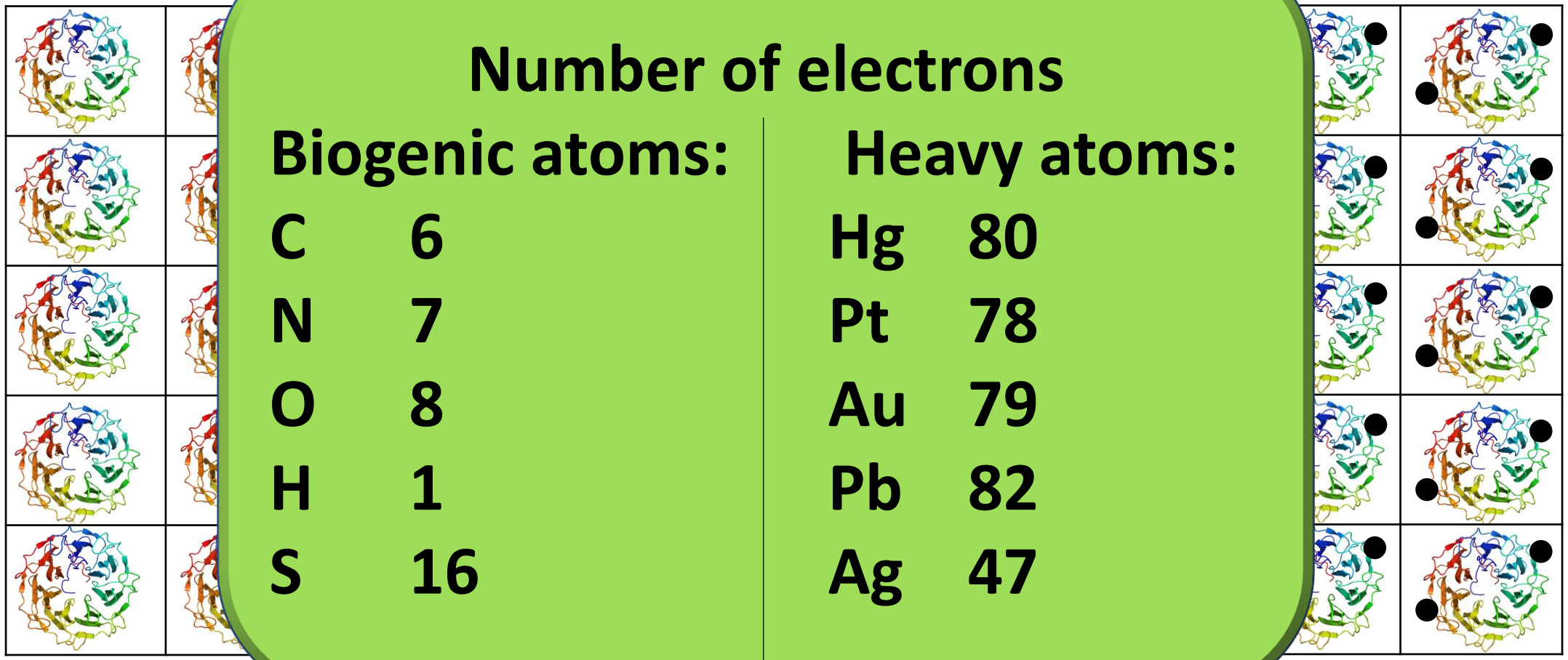


Heavy atom derivatives



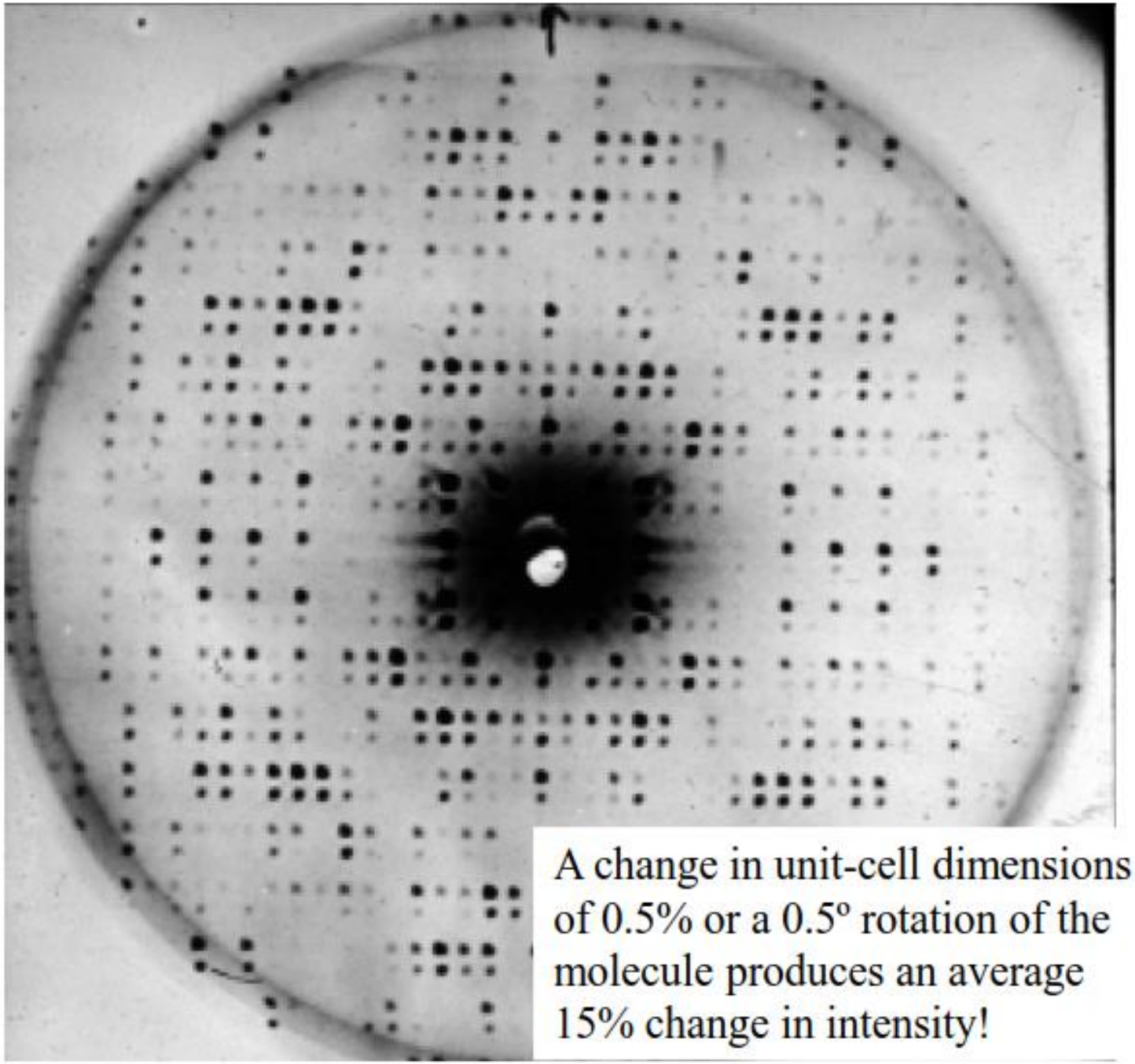
Native

derivatives



Number of electrons

Biogenic atoms:		Heavy atoms:	
C	6	Hg	80
N	7	Pt	78
O	8	Au	79
H	1	Pb	82
S	16	Ag	47

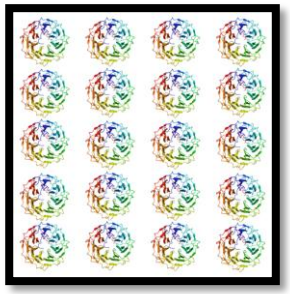


A change in unit-cell dimensions of 0.5% or a 0.5º rotation of the molecule produces an average 15% change in intensity!

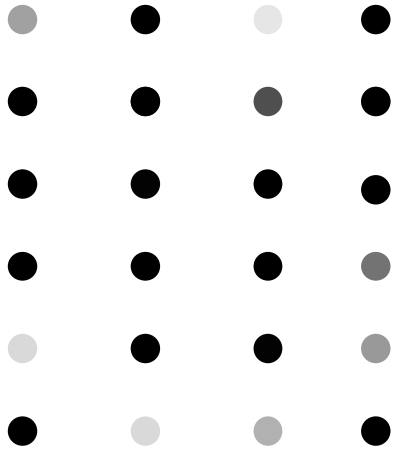
Addition of 1 Hg atom to a protein of 1000 atoms will produce an average fractional change of intensity of ~25% so differences should be easy to measure and data doesn't have to be that accurate

Crick, F. H. C. & Magdoff, B. S. (1956). Acta Cryst. 9, 901-908.

Native crystal

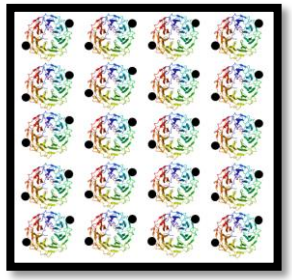


Crystal

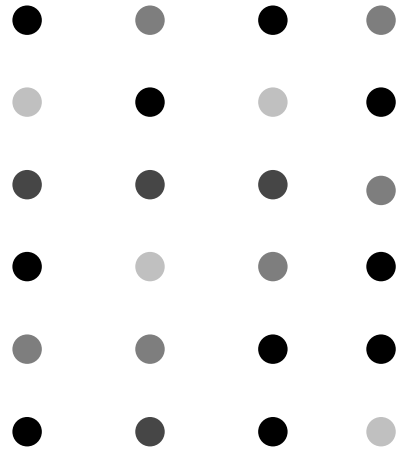


Diffraction pattern

Heavy atom derivatives



Crystal

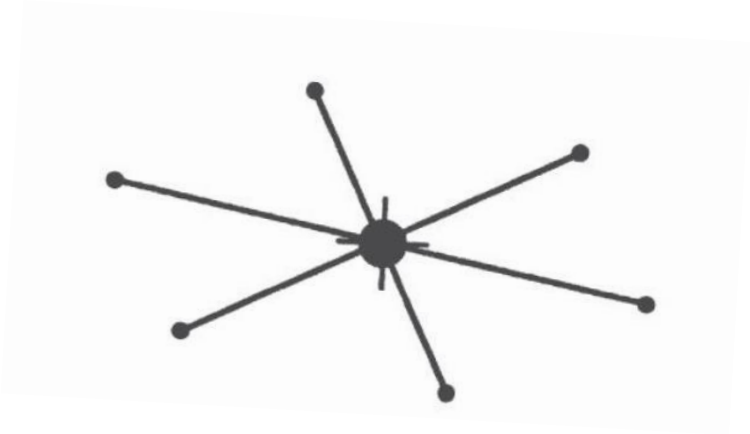


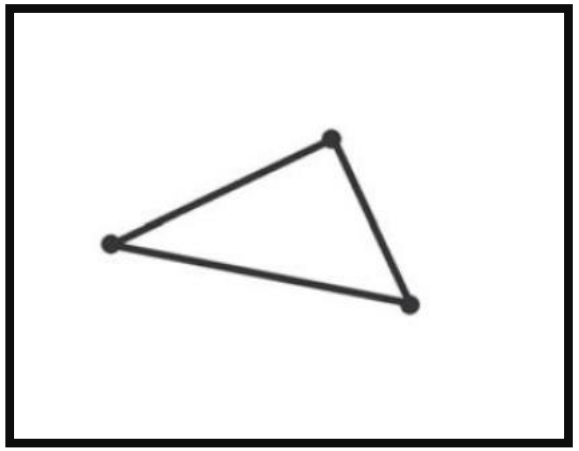
Diffraction pattern

$$F_{PH} - F_P = F_H$$

Patterson map of Heavy atoms

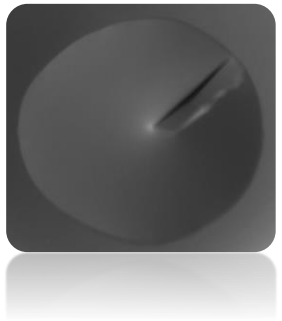
$$P_{(uvw)} = \frac{1}{V_C} \sum_h \sum_k \sum_l |F_{PH} - F_P|^2 \, exp\,[-2\pi i(h\mathrm{u} + k\mathrm{v} + lw)\,]$$



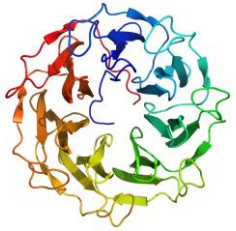


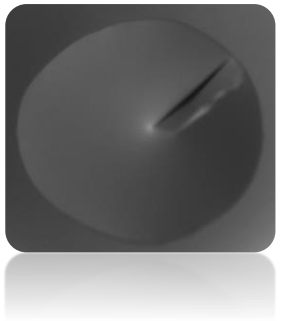
Heavy atom derivatives

+ Heavy atom



Co-crystallization



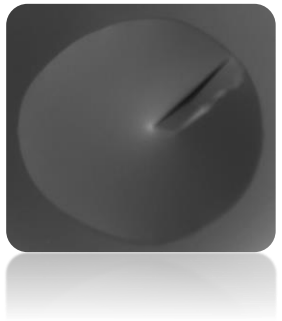


+ Heavy atom

Soaking

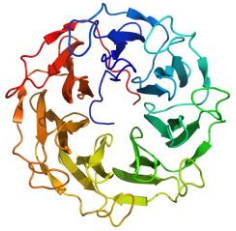
Heavy atom derivatives

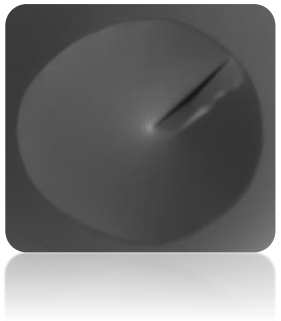
+ Heavy atom



Co-crystallization

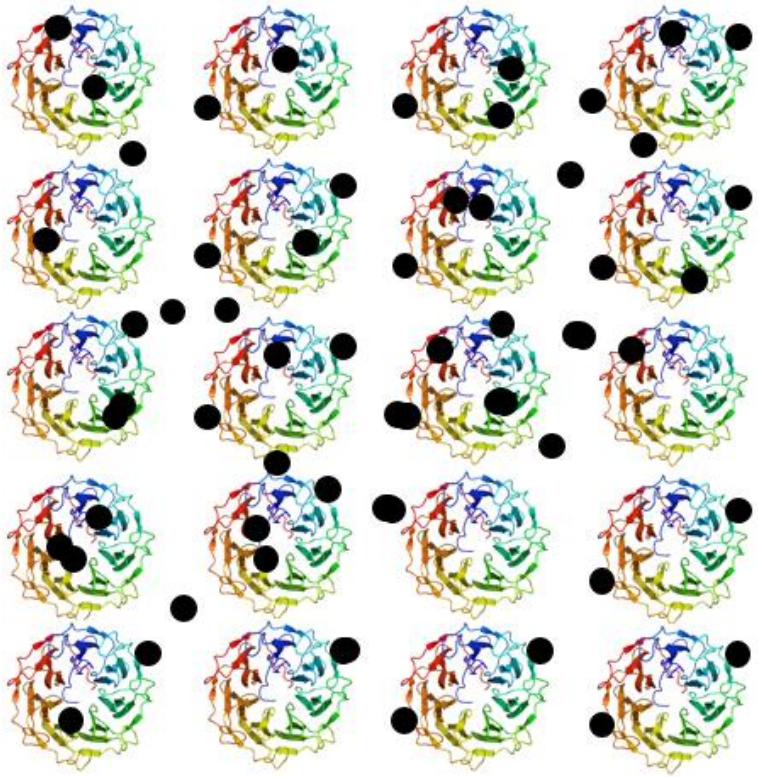
Optimization: pH, substance, time of soaking, concentration ...



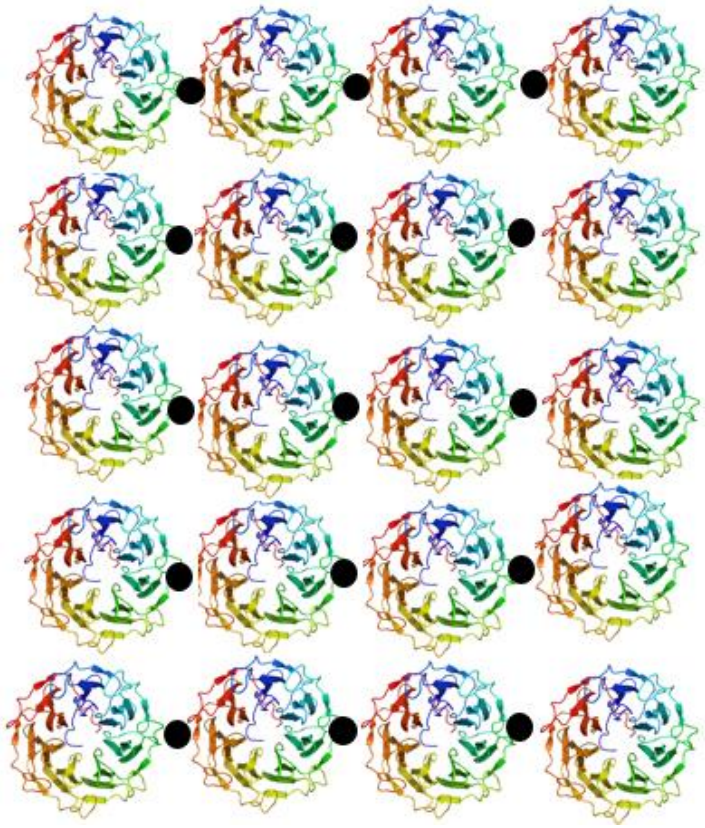


+ Heavy atom

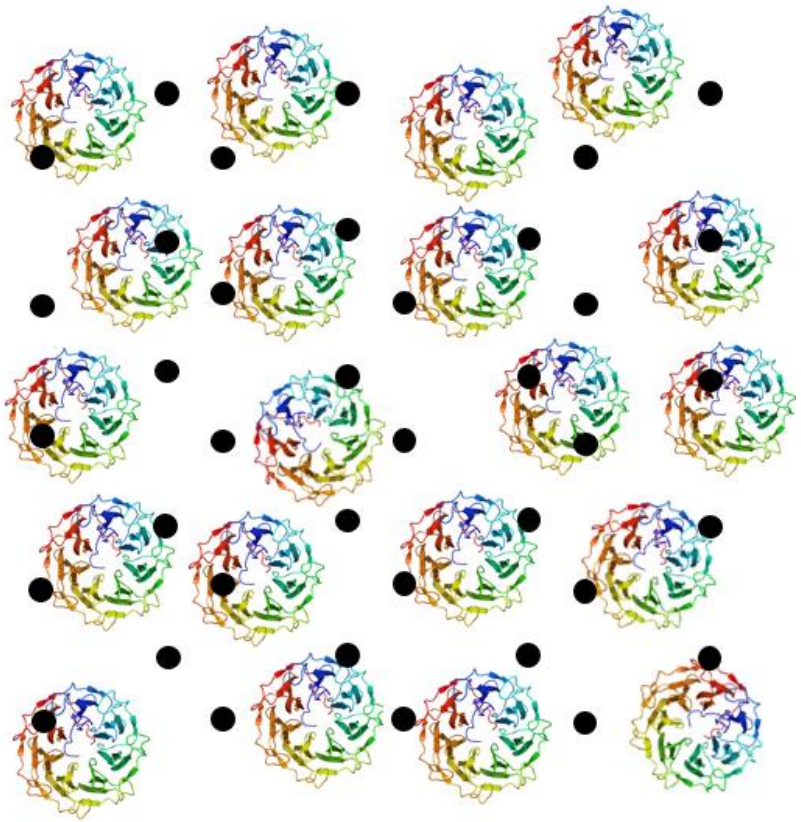
Soaking



1. Non-specific binding



2. Change of the unit cell parameters



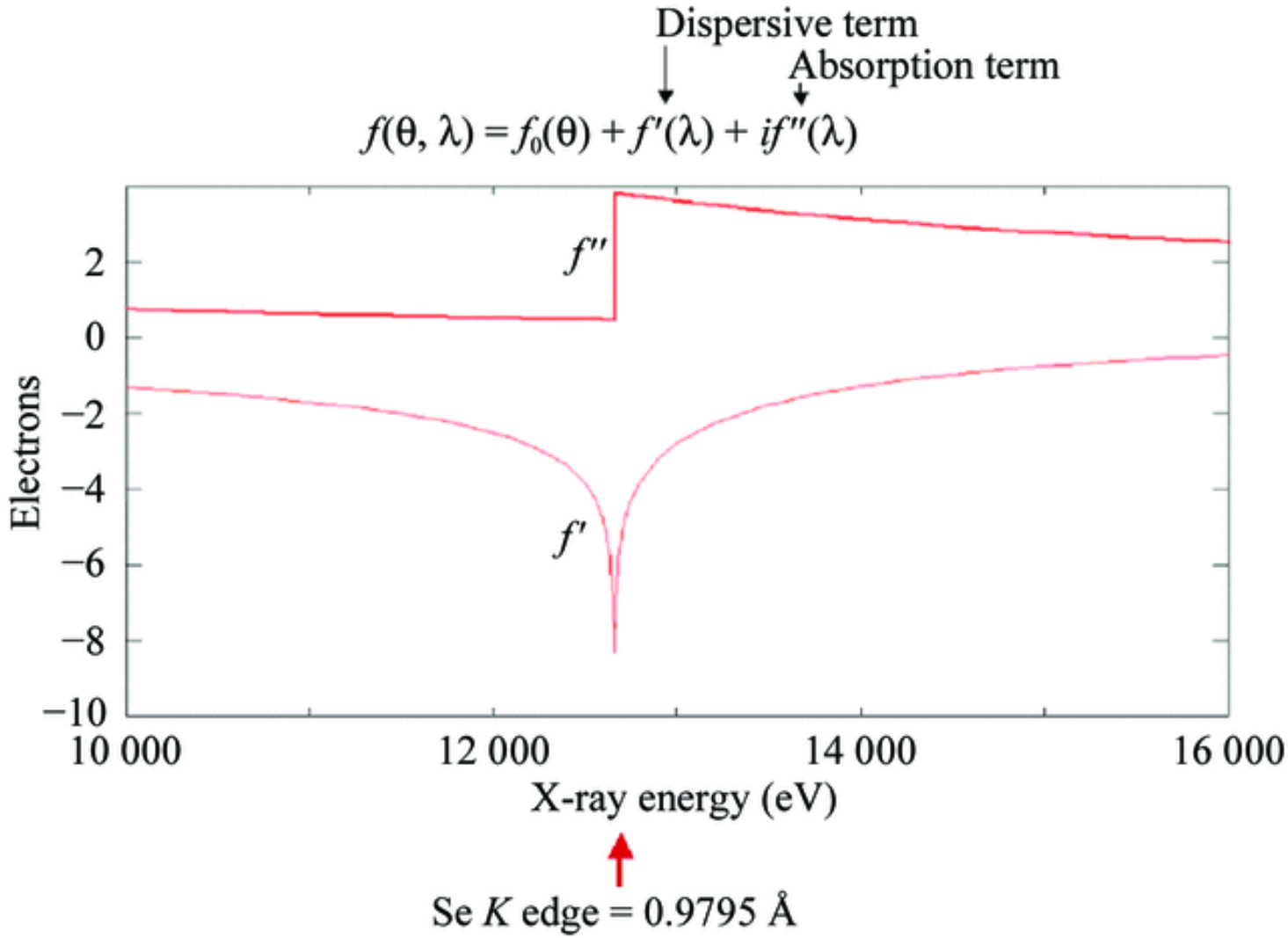
3. Disruption of the crystal

Problems with binding of HA

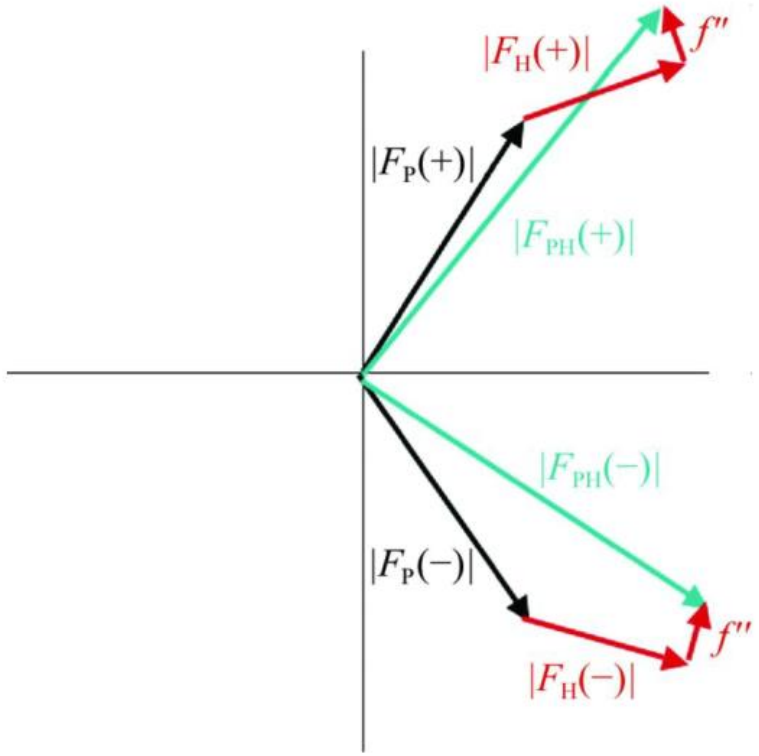
Anomalous scattering

Anomalous scattering

- Absorbtion of X-ray
- Reemiting alternated phases
- Heavy atoms
- Fe, Mn, Se (SeMet)



Anomalous scattering

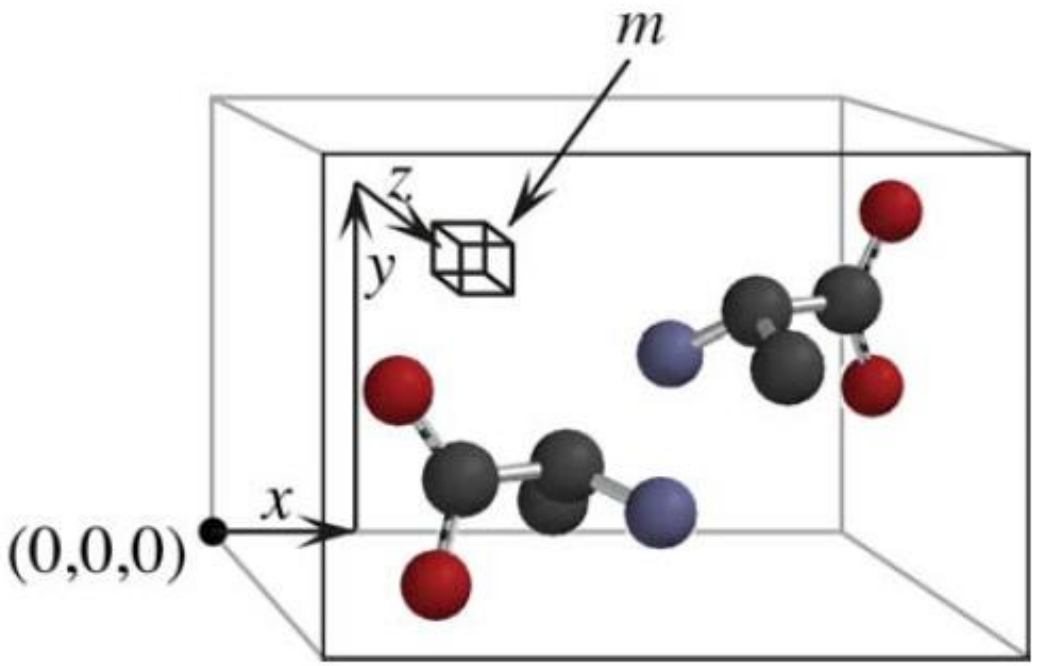


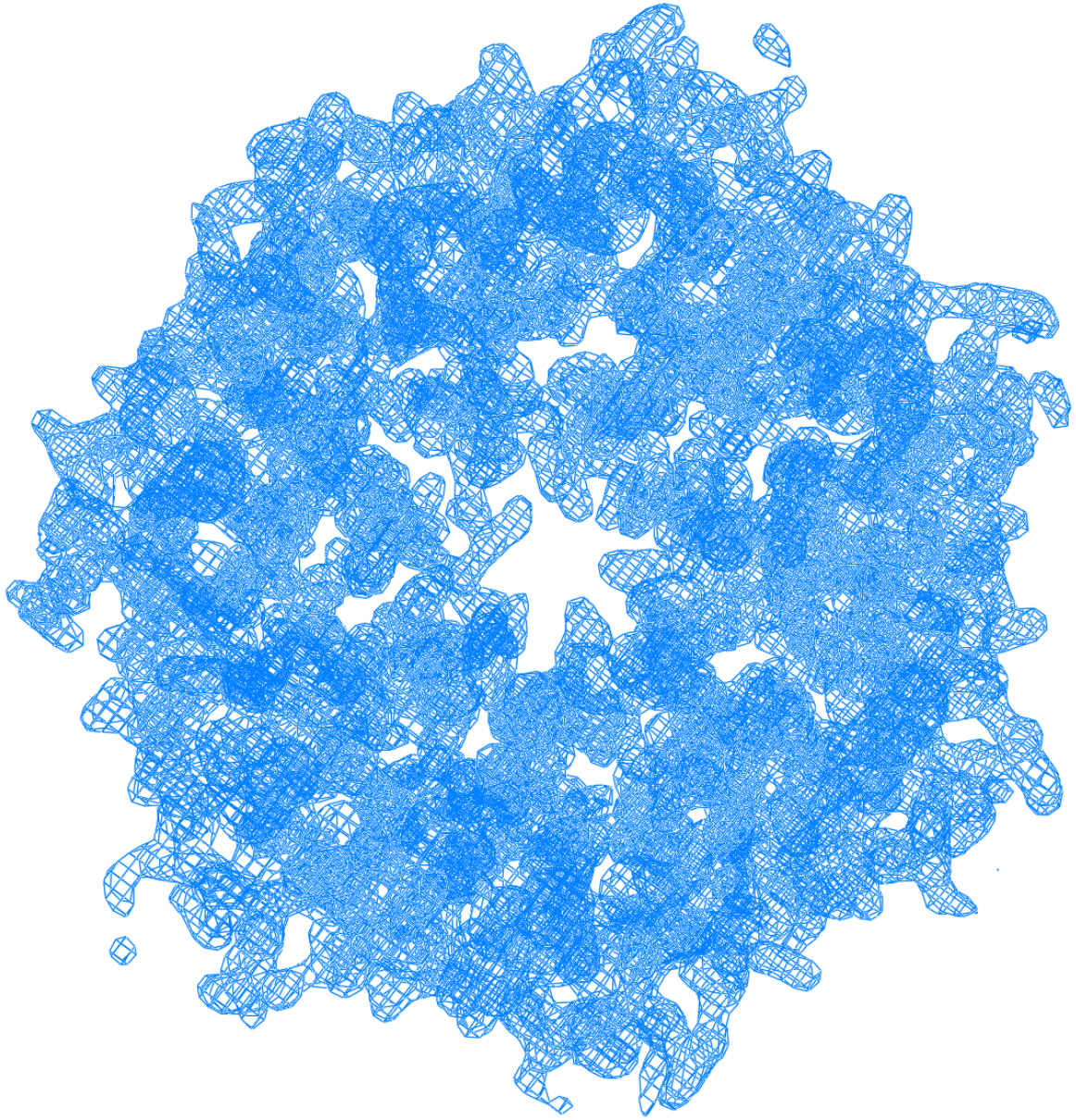
- Friedels law false

$$|F_{\bar{h}\bar{k}\bar{l}}| \neq |F_{hkl}|$$

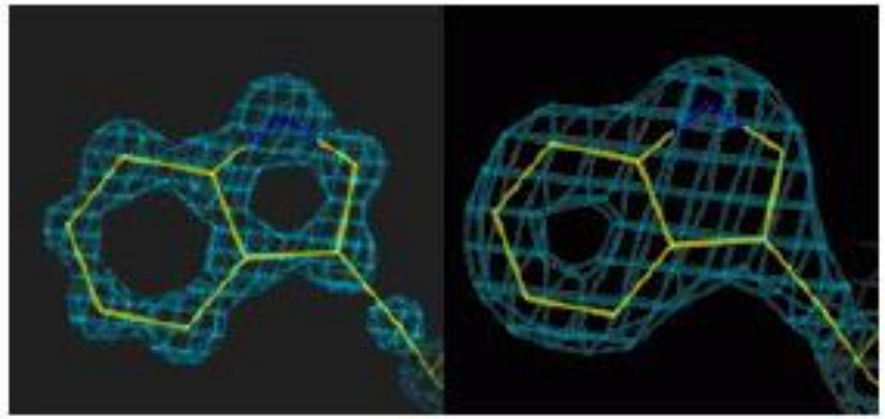
- Structure factors are alternated
- The most common: SeCys
- **Finding Heavy atoms by Patterson map!**

Electron density maps

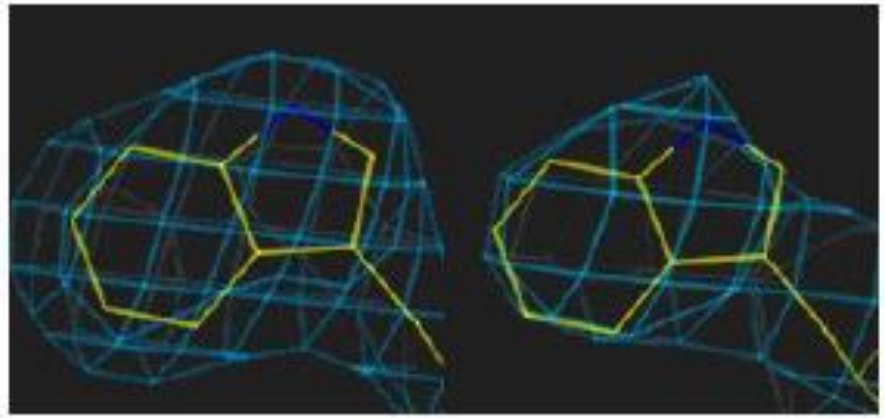




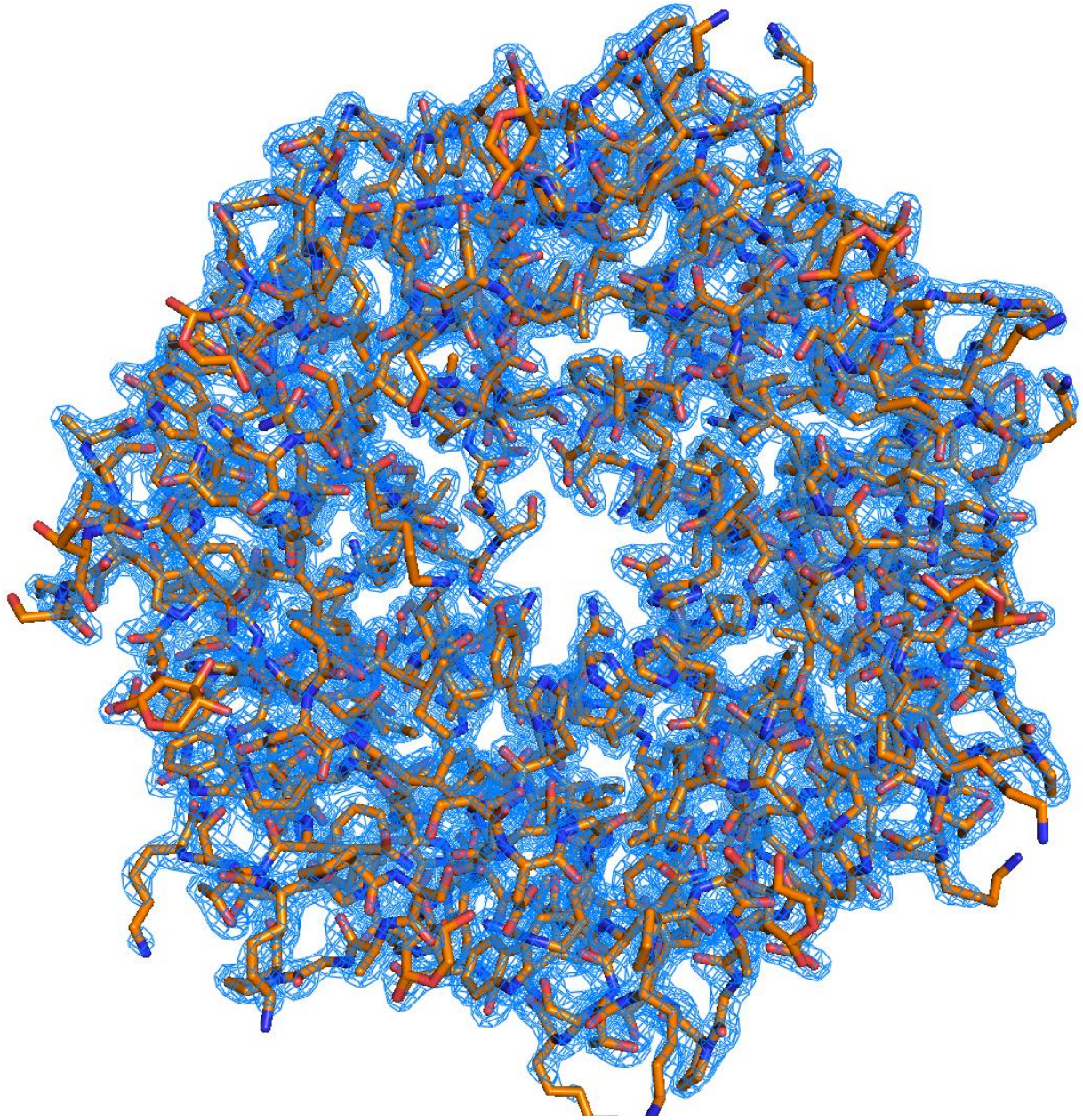
Model building & resolution

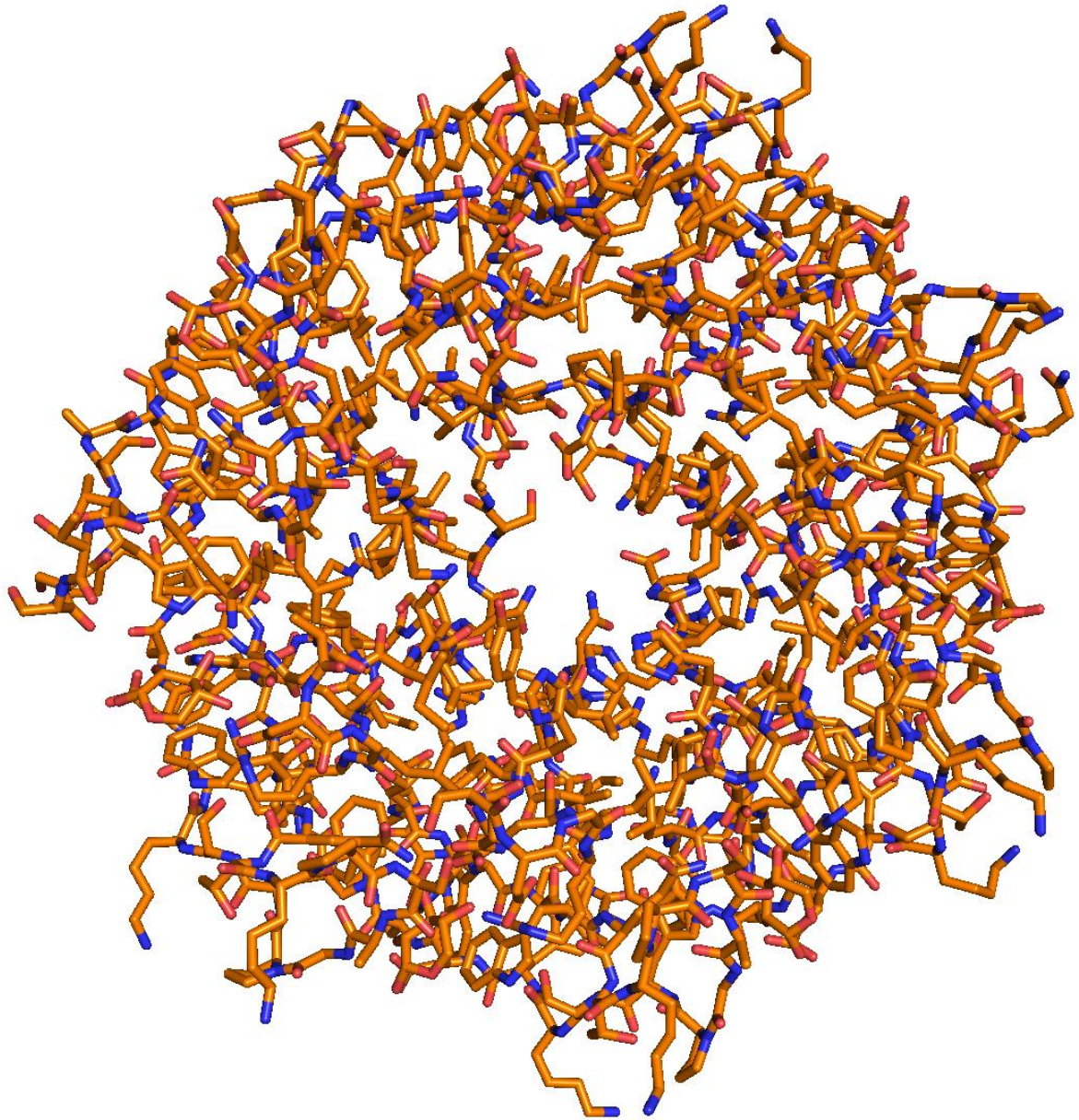


1.0Å 2.5Å



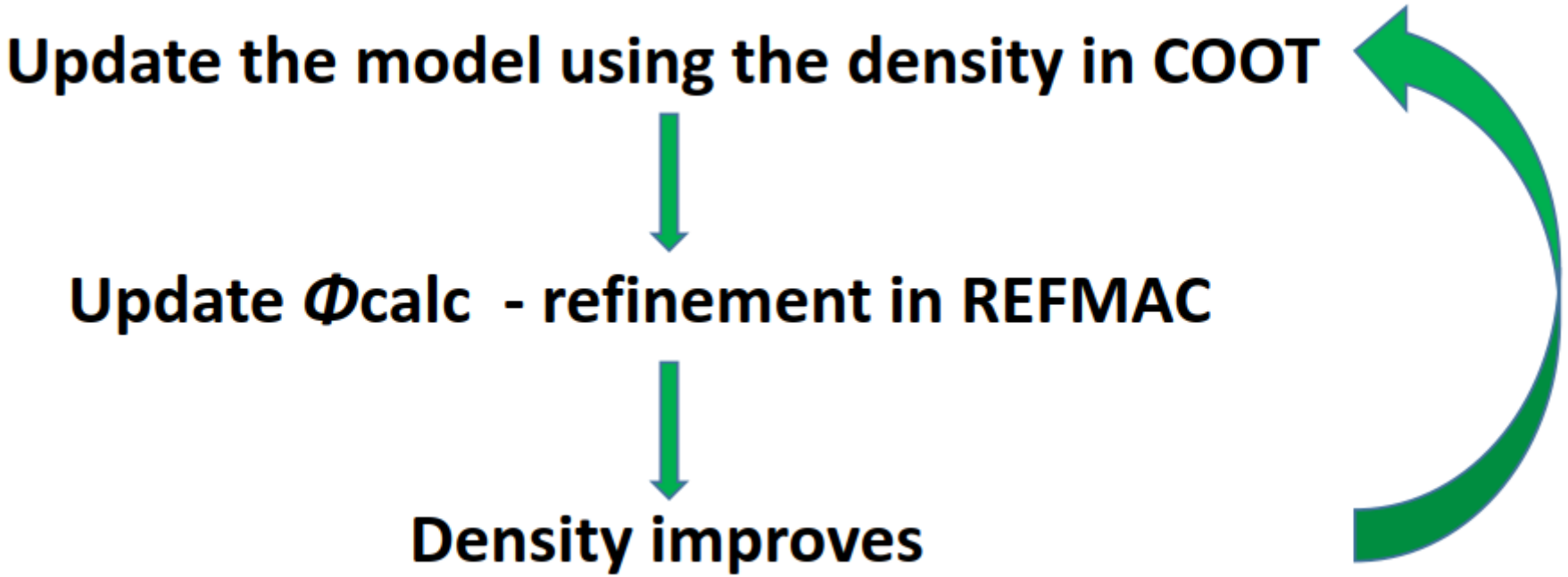
3.0Å 4.0Å





Refinement of the model

Main objective of the refinement process – build a model that explains diffraction data as good as possible and without violation of chemistry



Ramachandran diagram

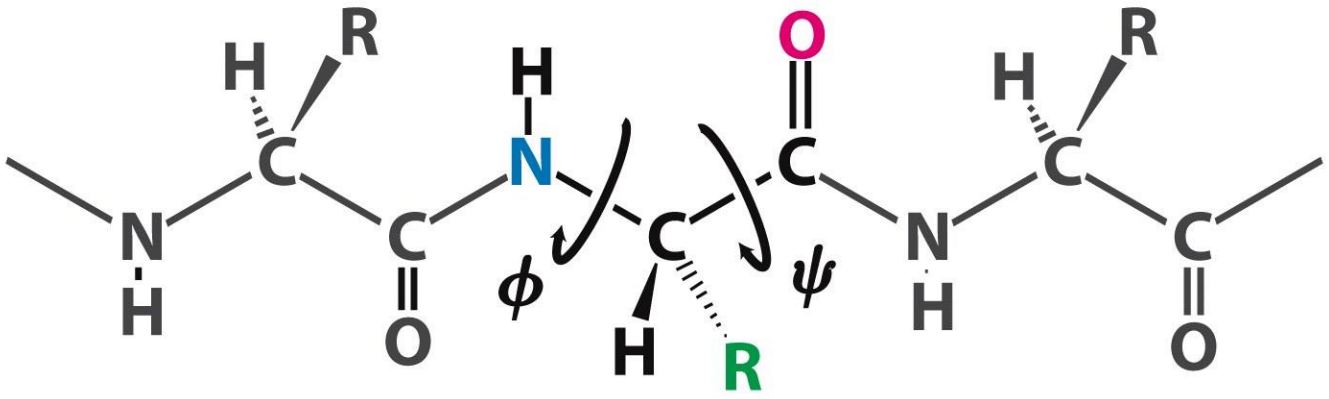
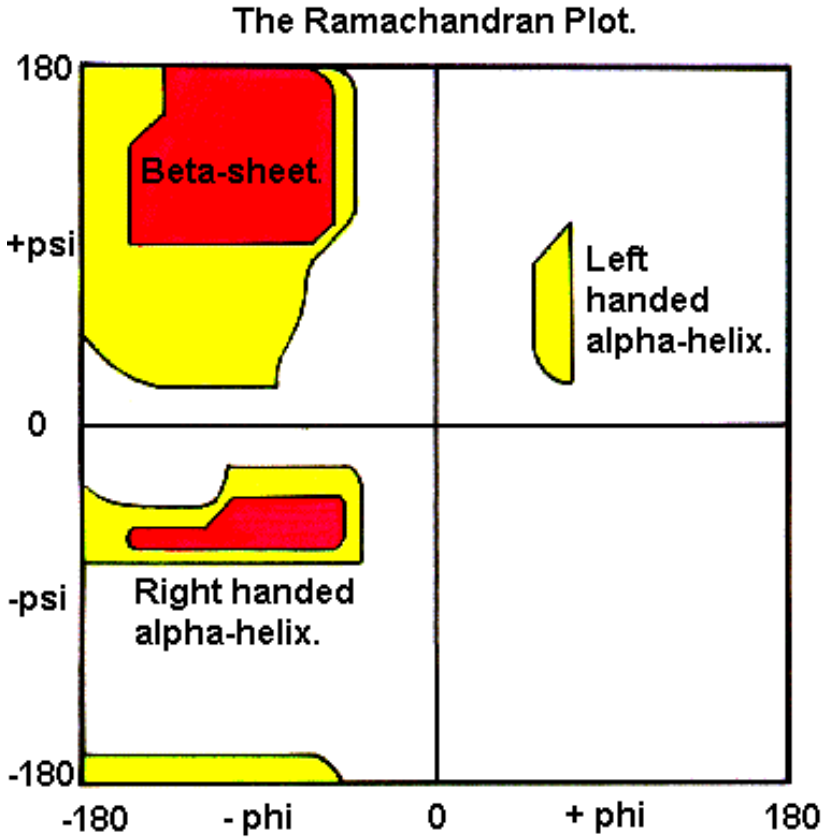
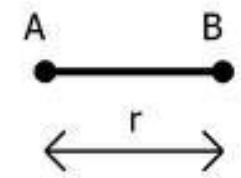


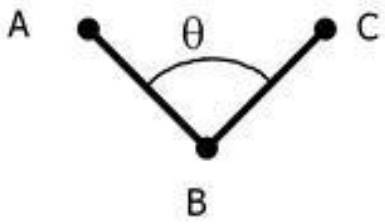
Figure 2.22a
***Biochemistry*, Seventh Edition**



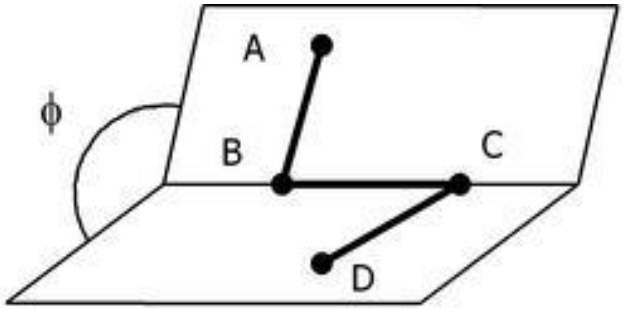
Geometrical parameters



Bond Length r



Bond Angle θ



Dihedral Angle ϕ

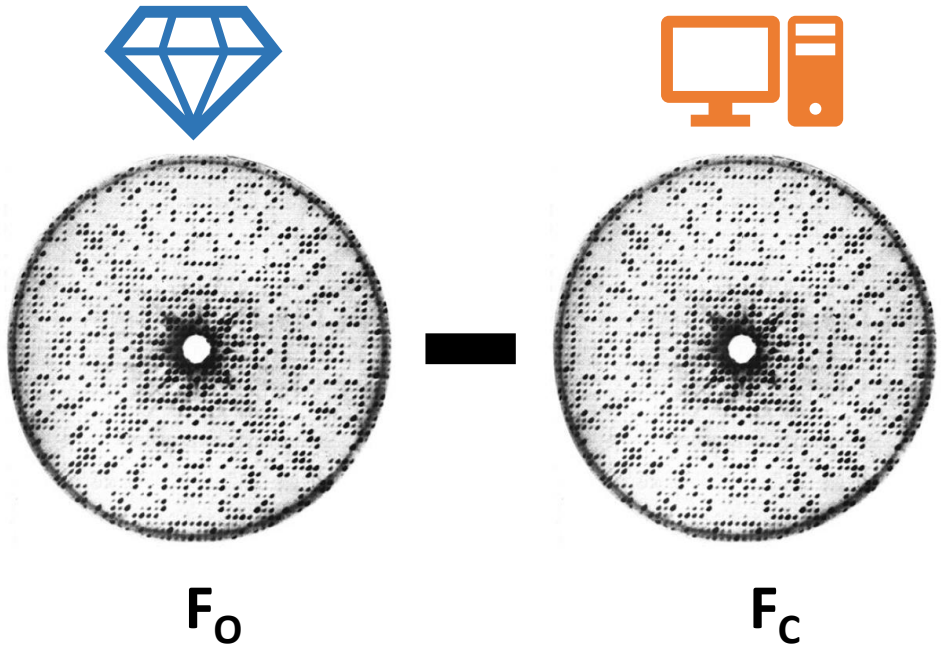
$$\mathrm{RMSD}=\sqrt{\frac{\sum_{n=1}^{N}(y_n-\hat{y}_n)^2}{N}}$$

https://www.webmo.net/link/help/AdjustTool.html

Validation (Reciprocal space)

$$R = \frac{\sum_{hkl}\left||F_O| - k|F_C|\right|}{\sum_{hkl}|F_O|}$$

- 95 % of the reflections
- Model bias
- Overfitting



Random structure:

R = 0.65

Validation (Reciprocal space)

$$R = \frac{\sum_{hkl}\left||F_O| - k|F_C|\right|}{\sum_{hkl}|F_O|}$$

- 95 % of the reflections
- Model bias
- Overfitting

$$R_{free} = \frac{\sum_{hkl\in T}\left||F_O| - k|F_C|\right|}{\sum_{hkl\in T}|F_O|}$$

- 5 % of the reflections
- **No model bias!!**

Validation (Reciprocal space)

$$R = \frac{\sum_{hkl}\left||F_O| - k|F_C|\right|}{\sum_{hkl}|F_O|}$$

- 95 % of the reflections
- Model bias
- Overfitting

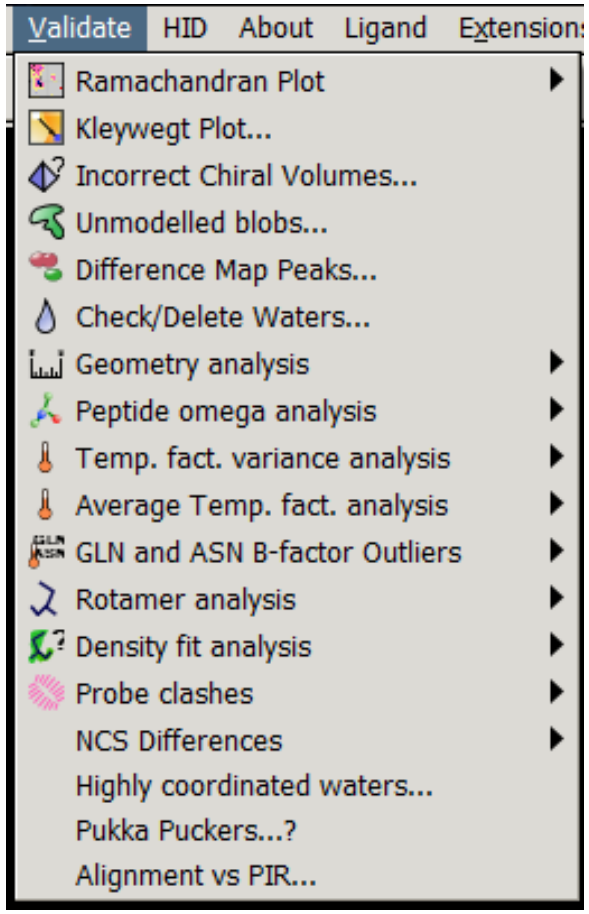
$$R_{free} = \frac{\sum_{hkl\in T}\left||F_O| - k|F_C|\right|}{\sum_{hkl\in T}|F_O|}$$

- 5 % of the reflections
- No model bias

Good structure:

- Resolution * 10
- R- R_{free}= 4 – 6 %

Validation



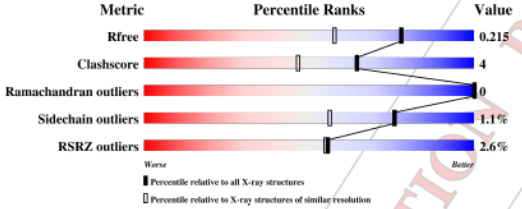
WinCoot

1 Overall quality at a glance

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 1.65 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.

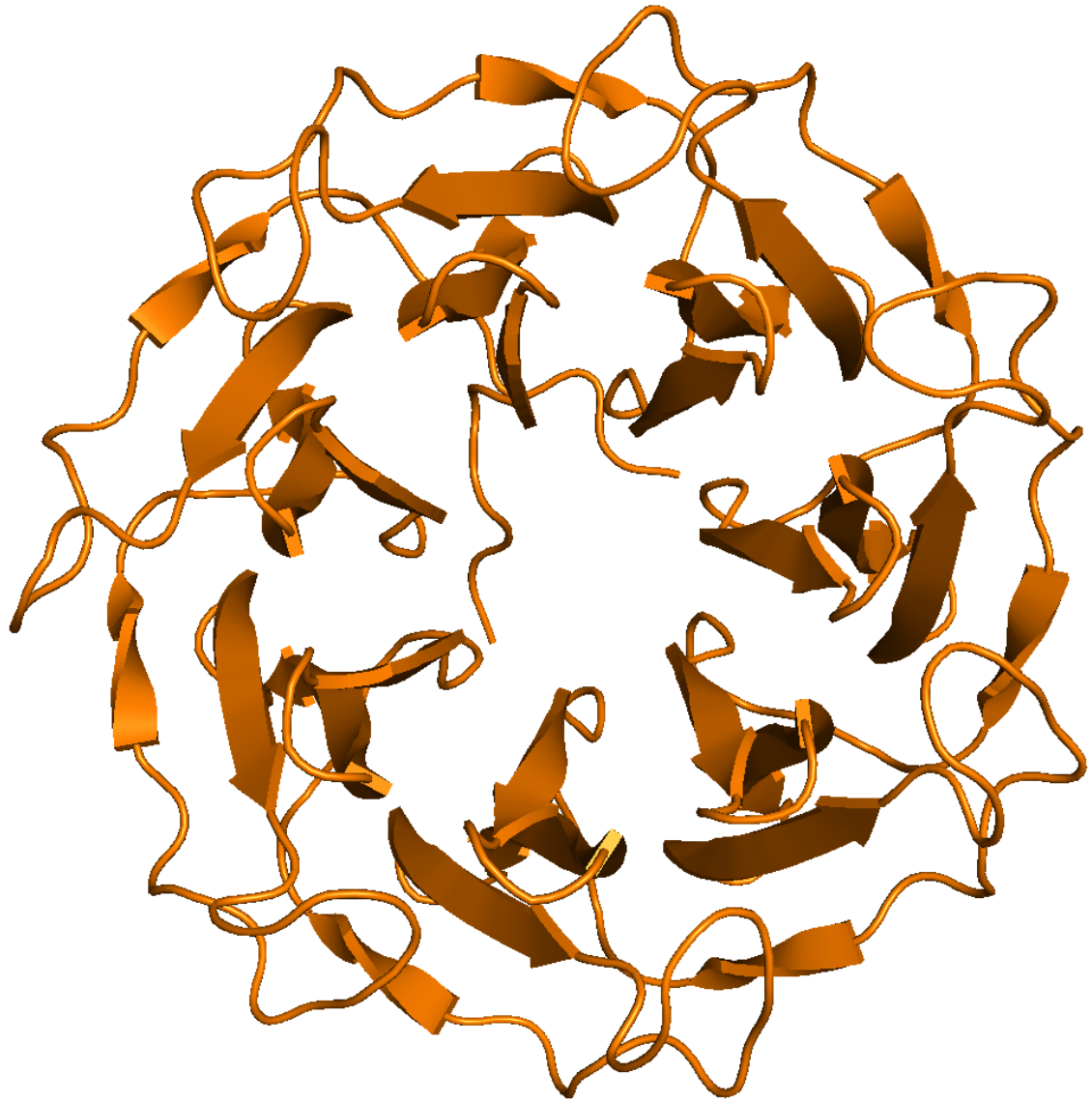


Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	111664	2539 (1.66-1.62)
Clashscore	122126	2672 (1.66-1.62)
Ramachandran outliers	120053	2629 (1.66-1.62)
Sidechain outliers	120020	2629 (1.66-1.62)
RSRZ outliers	108989	2498 (1.66-1.62)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	AAA	348	3% 91% 8%

PDB validation system



Structural database

Biological Assembly 1

3D View: Structure | Electron Density | Validation Report | Ligand Interaction

Global Symmetry: Cyclic - C3 (3D View)
Global Stoichiometry: Homo 3-mer - A3

Display Files | Download Files

2BT9

Lectin from Ralstonia solanacearum complexed with Me-fucoside

DOI: 10.2210/pdb2BT9/pdb

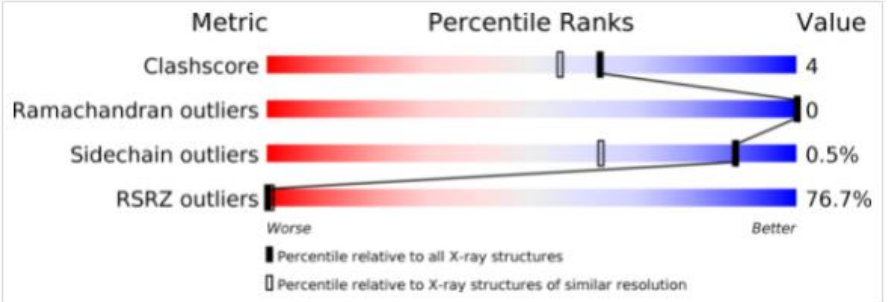
Classification: LECTIN
Organism(s): Ralstonia solanacearum
Expression System: Escherichia coli BL21(DE3)
Mutation(s): No

Deposited: 2005-05-27 **Released:** 2005-06-02
Deposition Author(s): Mitchell, E.P., Kostlanova, N., Wimmerova, M., Imberty, A.

Experimental Data Snapshot

Method: X-RAY DIFFRACTION
Resolution: 0.94 Å
R-Value Free: 0.120

wwPDB Validation 3D Report | Full Report



Ligand Structure Quality Assessment