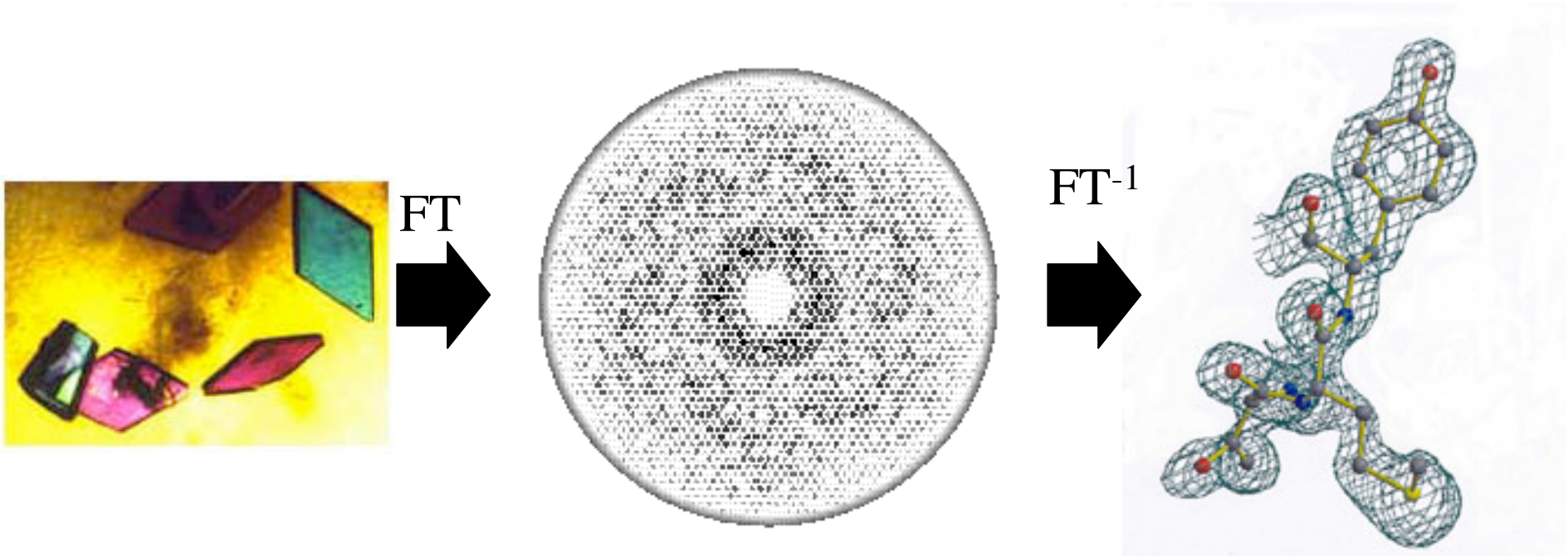
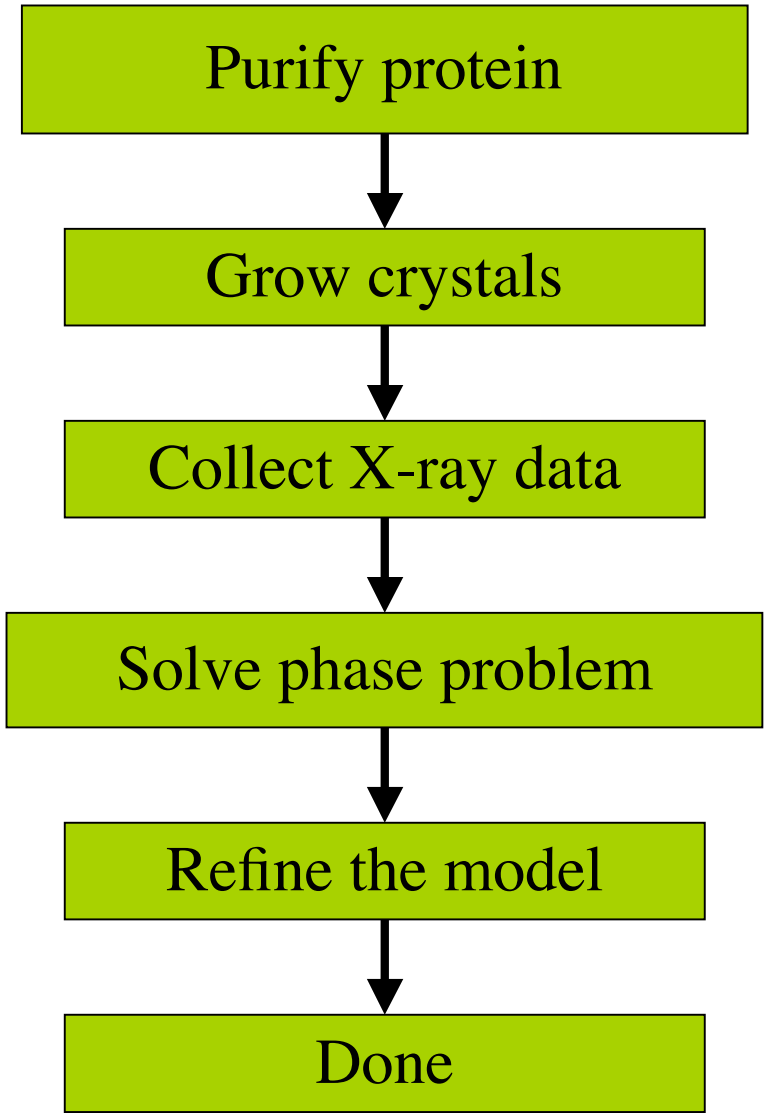
Protein Crystallography



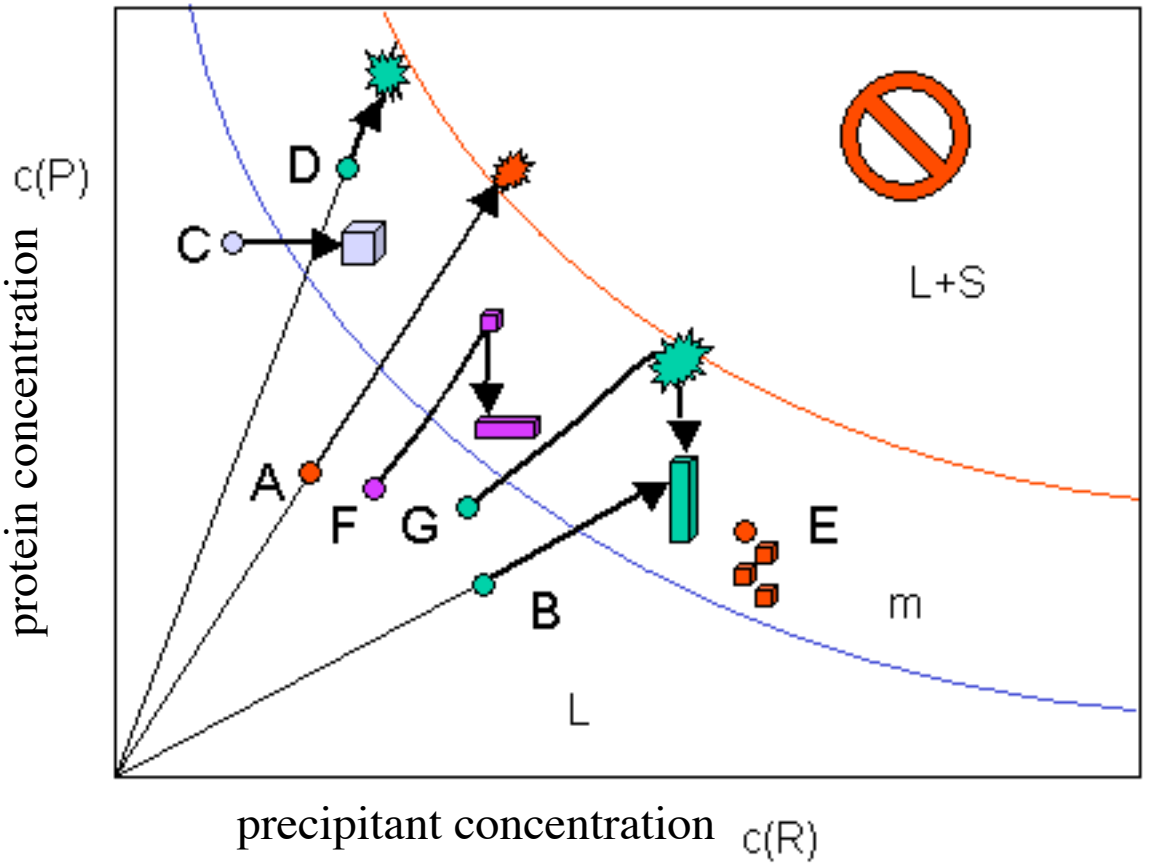
Crystals *X-rays* *Atoms*

From protein to protein structure



Crystals and X-rays

Protein crystal growth



Arrows indicate different diffusion experients.

A,B,D,F,G. Vapor diffusion.

E. Bulk

C. Microdialysis

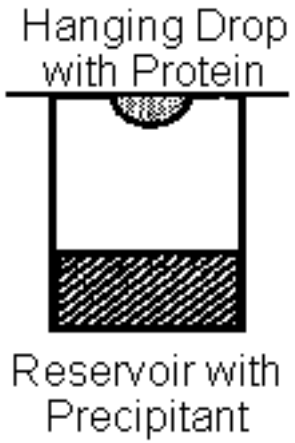
L=liquid
S=solid
m=metastable state (supersaturated)

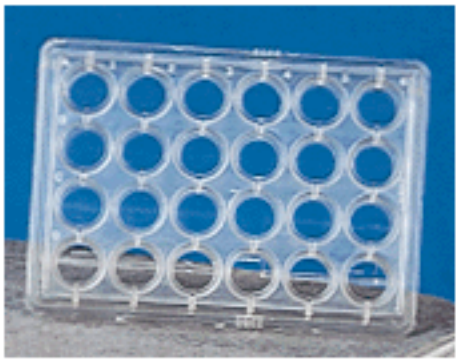
blue line = saturation of protein

red line = supersaturation limit

Crystal growth occurs between these two limits. Above the supersaturation limit, proteins form only disordered precipitate.

vapor diffusion setup





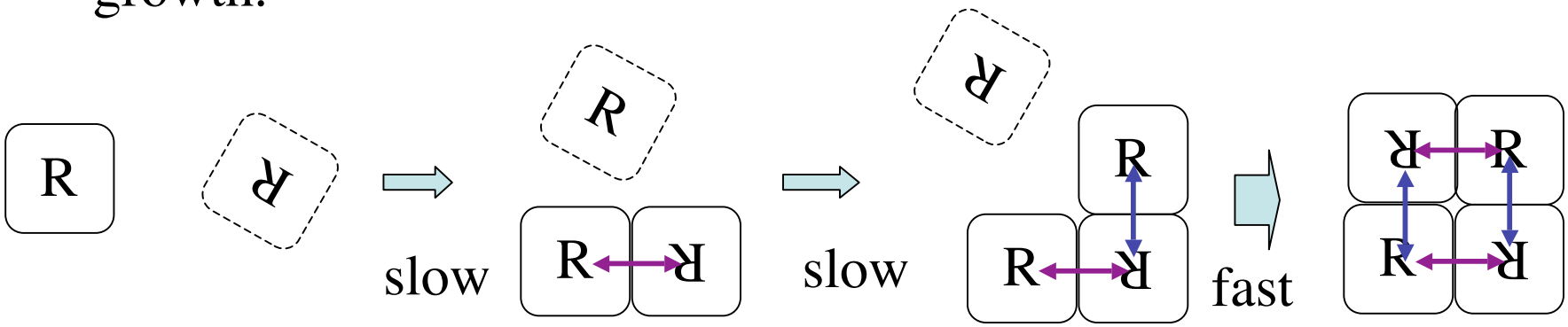
a Linbro plate

Volatiles (i.e. water) evaporate from one surface and condence on the other.

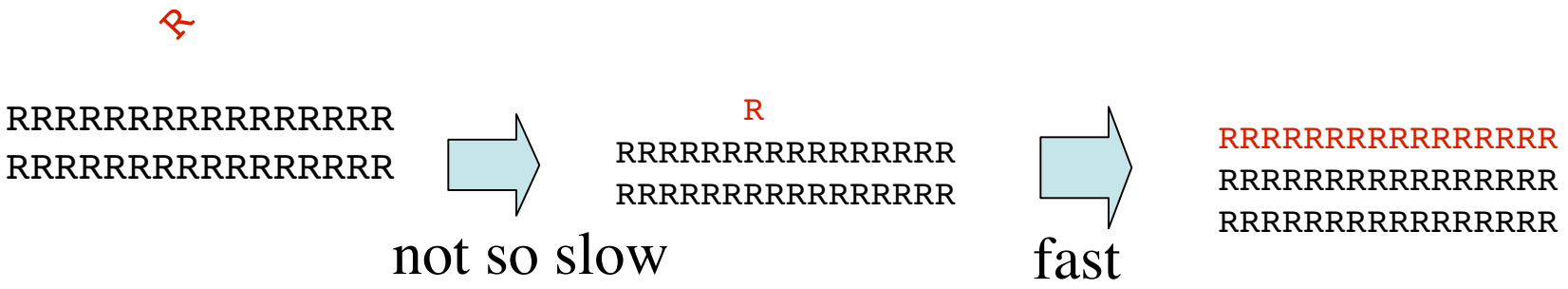
Drop has *higher* water concentration than reservoir, so drop slowly evaporates.

Crystallization theory

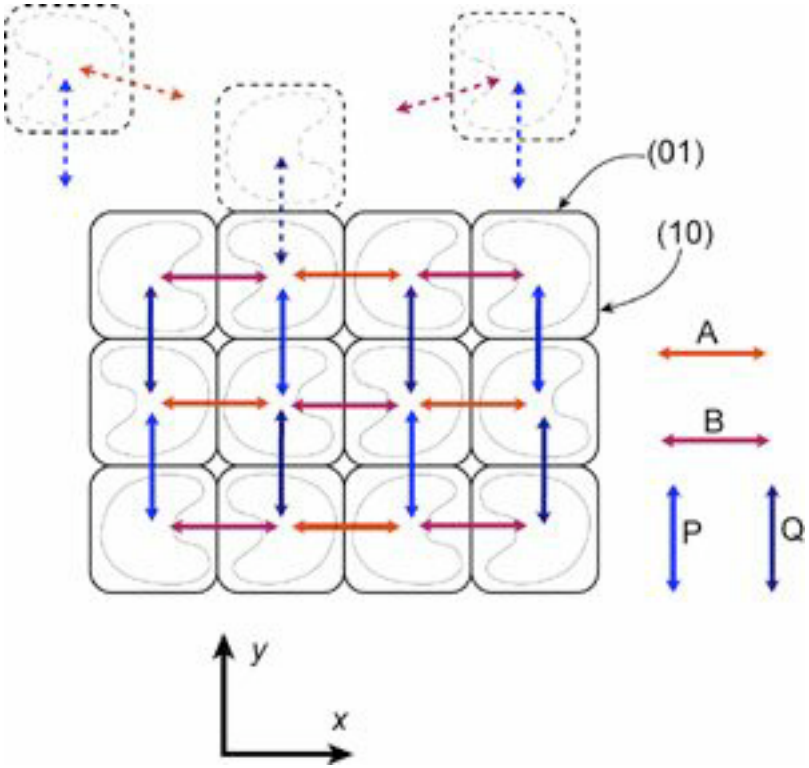
Nucleation takes higher concentration than crystal growth.



After nucleation, the large size of a face makes the weak bond more likely.



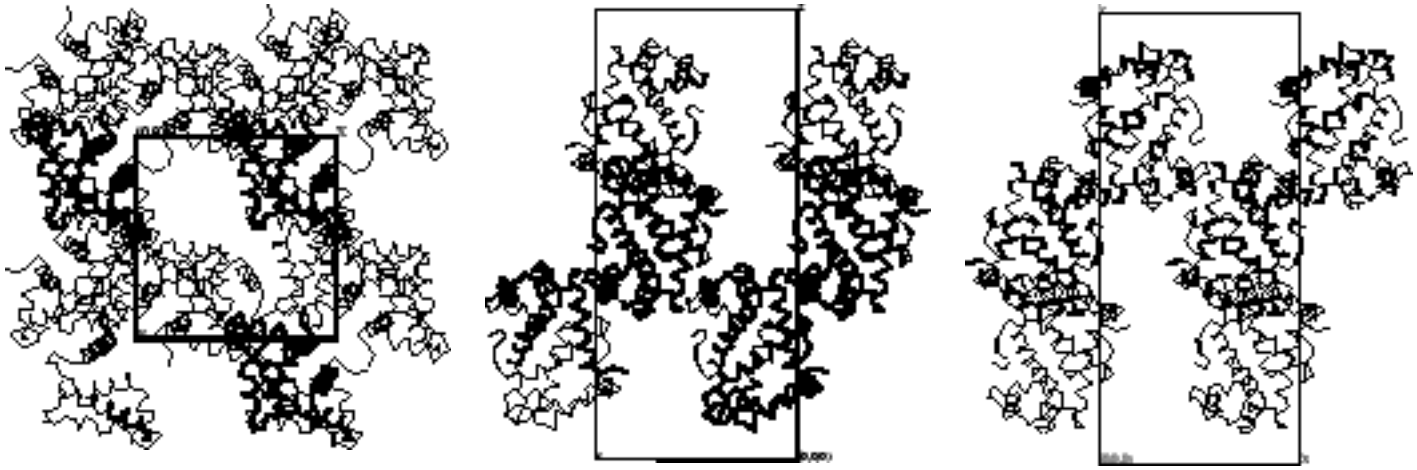
Crystallization theory



Bonds A,B are stronger than P,Q.

More on Periodic Bond Chain theory: http://www.che.utoledo.edu/nadarajah/webpages/PBC.htm

Crystal packing



Protein crystal packing interactions are salt-bridges and H-bonds mostly. These are much weaker than the hydrophobic interactions that hold proteins together. This means that (1) *protein crystals are fragile*, and (2) proteins in crystals are probably *not* significantly distorted from their native conformations.

Crystallization robot

High-throughput crystallography labs use pipeting robots to explore thousnds of “conditions”. Each condition is a formulation of the **crystal drop** and the **reservoir solution**.

<u>Conditions can have different</u>:

•protein concentration

•pH

•precipitant, precipitant concentration

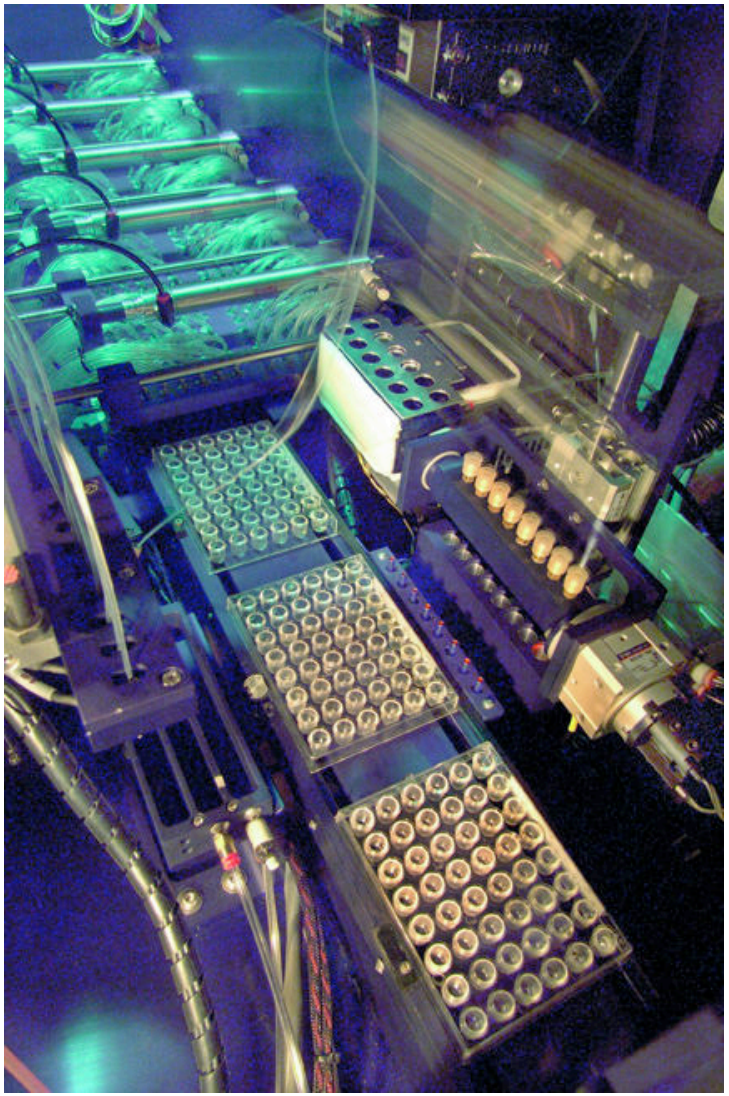
•detergents

•organic co-solvents

•metal ions

•ligands

•concentration gradient



protein crystals

~1mm

subtilisin

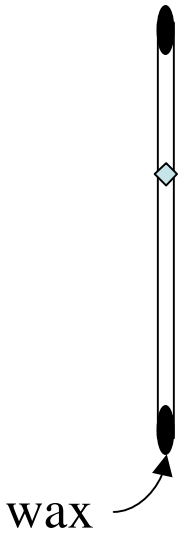
cellulase

The color you see is "**birefringence**", the wavelength-dependent rotation of polarized light.

Crystal mounting

If not freezing

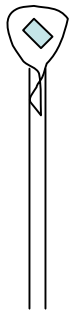
Xtal is mounted in a thin-walled glass capillary tube



Crystal must be kept at proper humidity and temperature!! Very fragile!

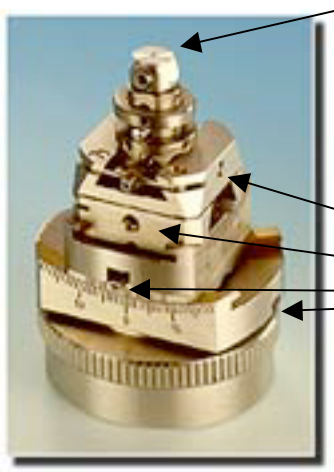
If freezing (preferred)

Xtal is mounted on a thin film of water in a wire loop. The loop is fixed to a metal or glass rod.



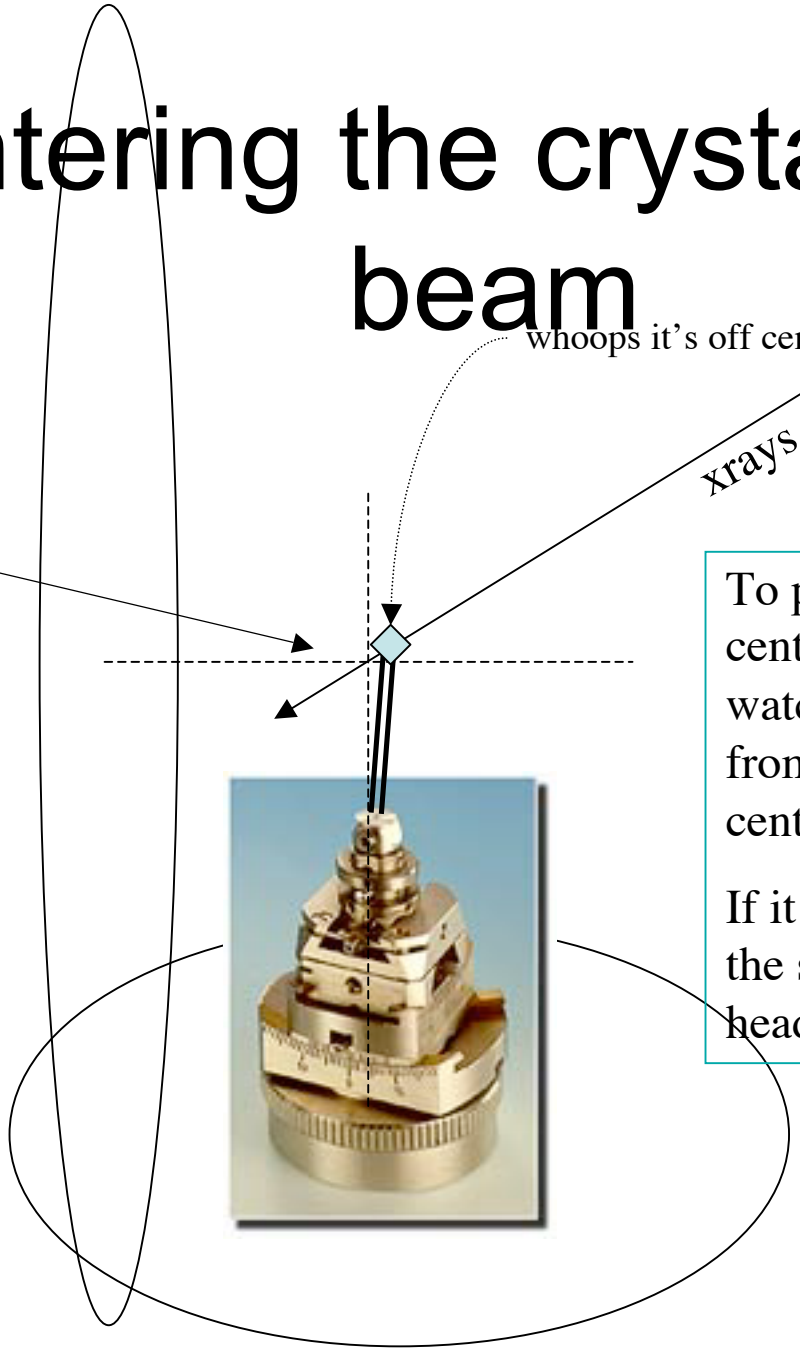
Must freeze immediately or film will dry out.!

Mounted xtal is attached to a goniometer head for precise adjustment.



eucentric goniometer head (made by Nonius)

Centering the crystal in the beam

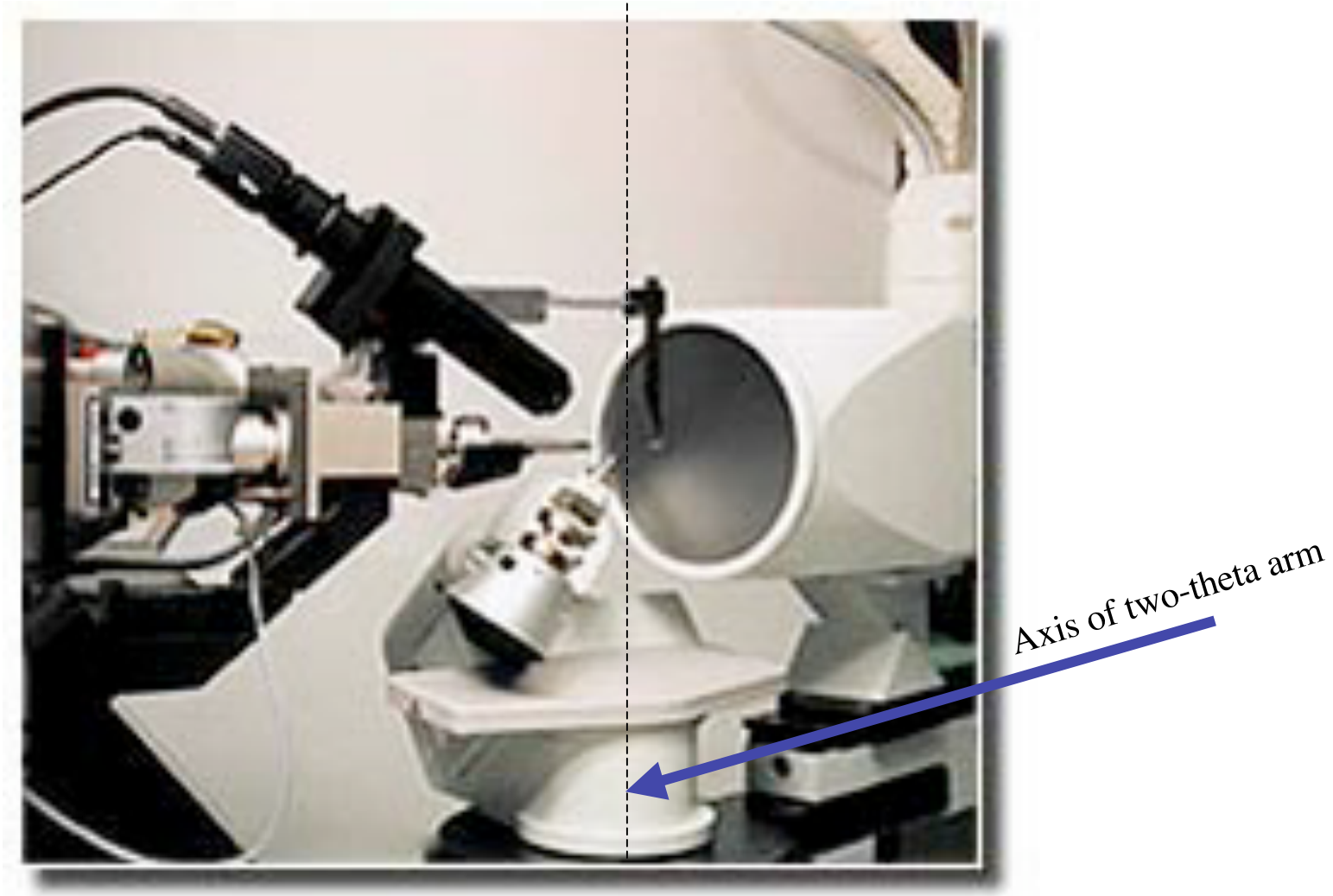


"**machine center**" is the intersection of the beam and the two goniostat rotation axes. Must be set by manufacturer!

To place crystal at machine center, rotate ω and κ and watch the crystal. If it moves from side to side, it is off center.

If it is off-center, we adjust the screws on the goniometer head.

X-ray diffractometer with area detector



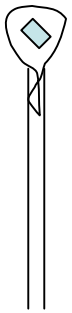
The detector (or film) sits on a “Two-theta arm” that can swing out, away from the beam to collect high-resolution data.

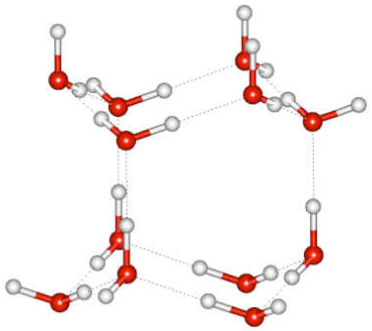
Synchrotrons provide tunable, monochromatic X-rays

Crystals must be flash frozen

Water must be frozen to < –70°C very fast to prevent the formation of **hexagonal ice crystals**. Water glass forms.

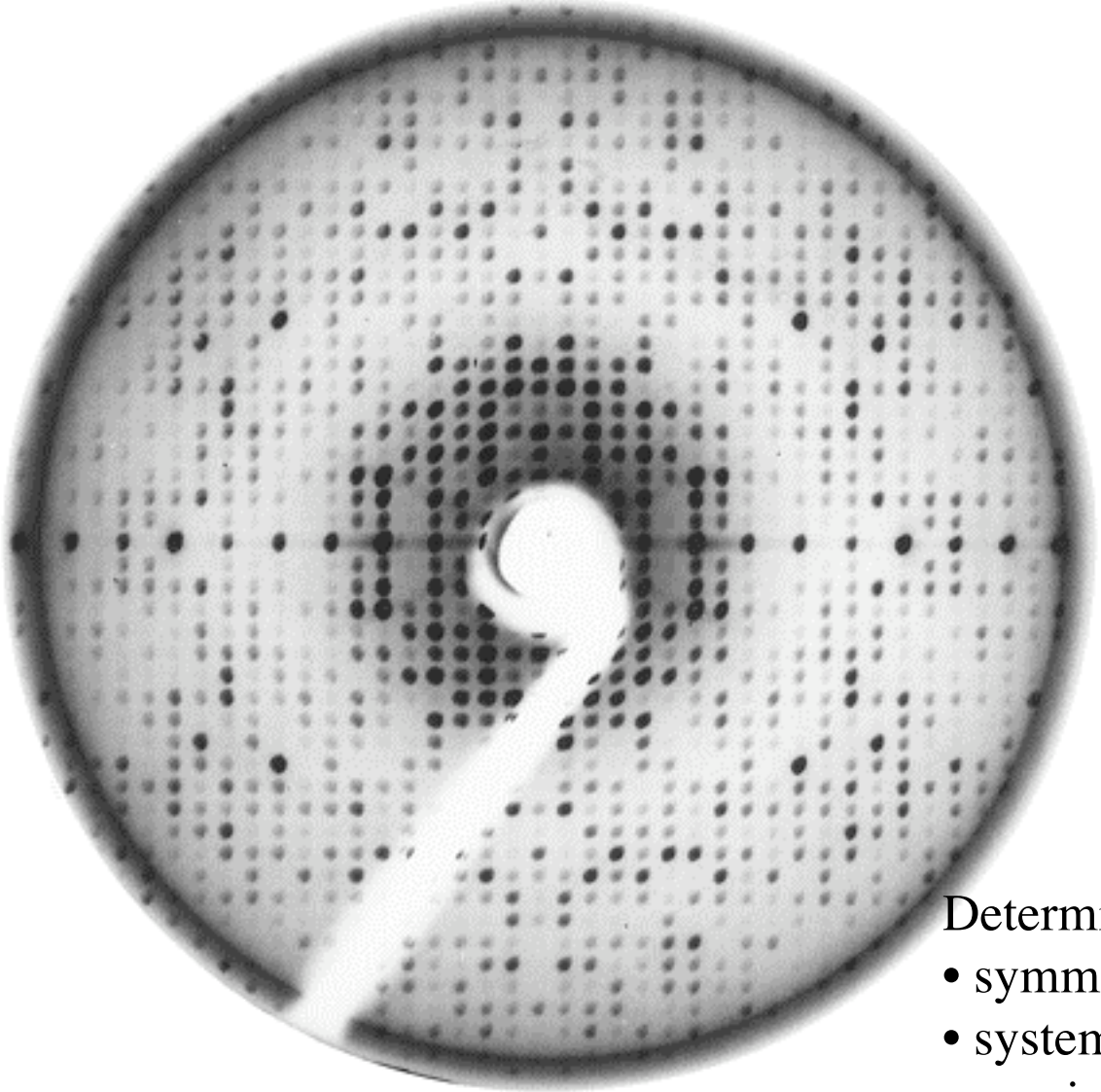
How? Crystals, mounted on loops, are flash frozen by dipping in liquid propane or freon at –70°, or by instant exposure to N_2 gas at –70°C.





hexagonal ice

Precession photograph



Determine Space Group

- symmetry
- systematic absences
- spacing of spots
- angle between axes.

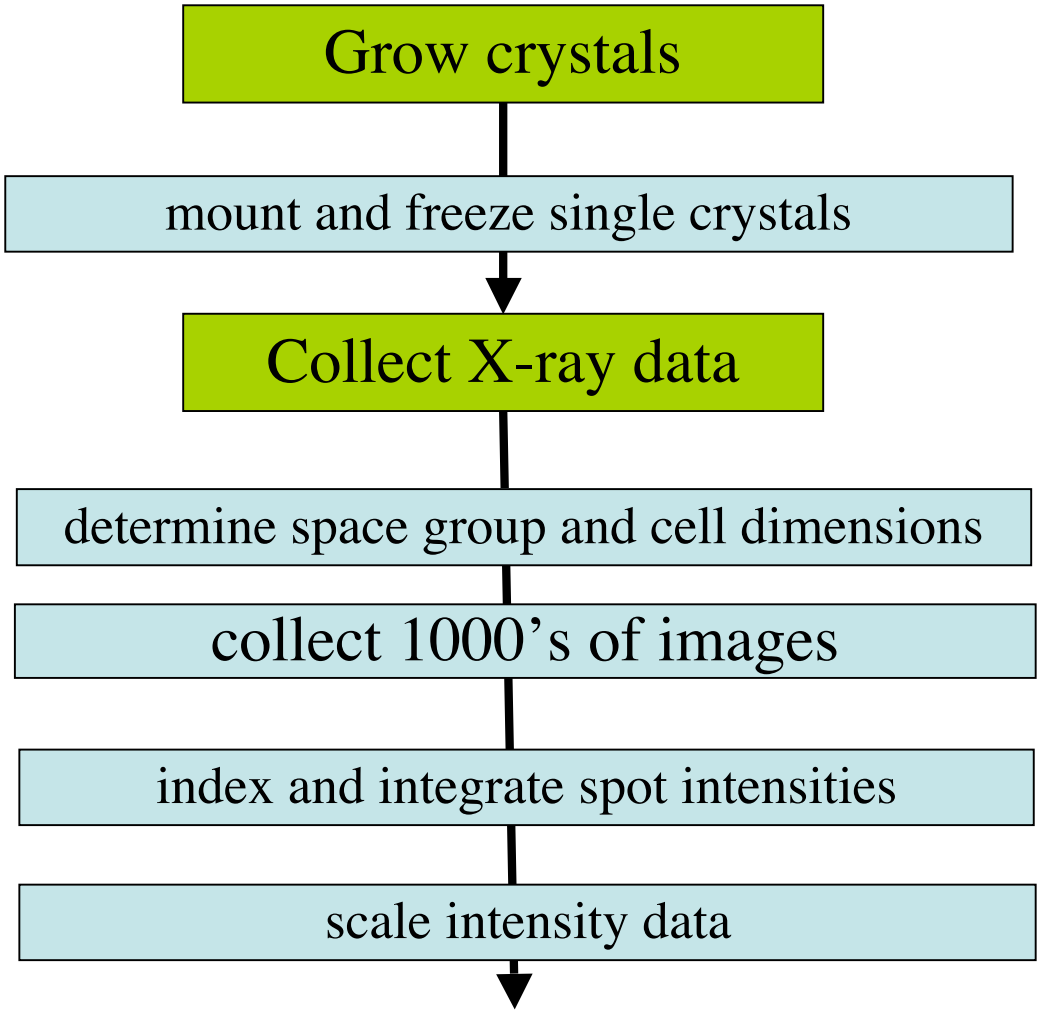
Data collection

Measure the intensity (amplitude squared) of each reflection.

Output of data collection, thousands of ***reflections***, each with 5 parameters:

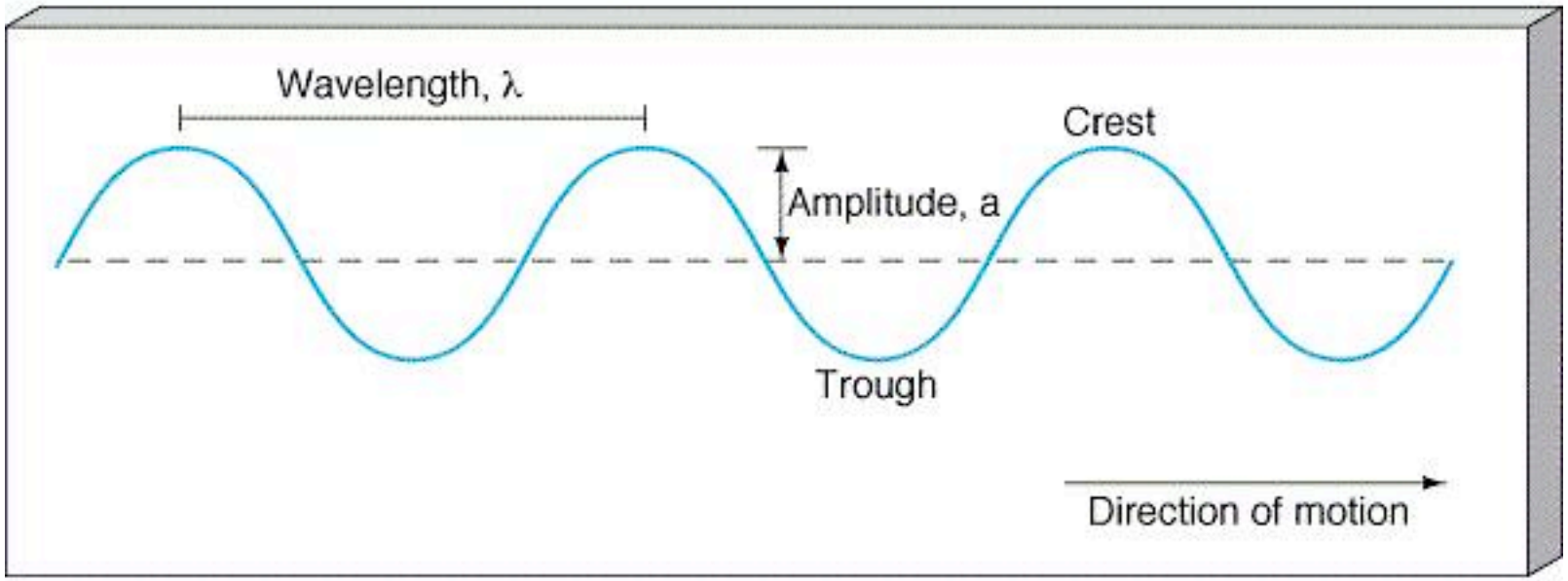
`h k l F sigma`

From protein to data



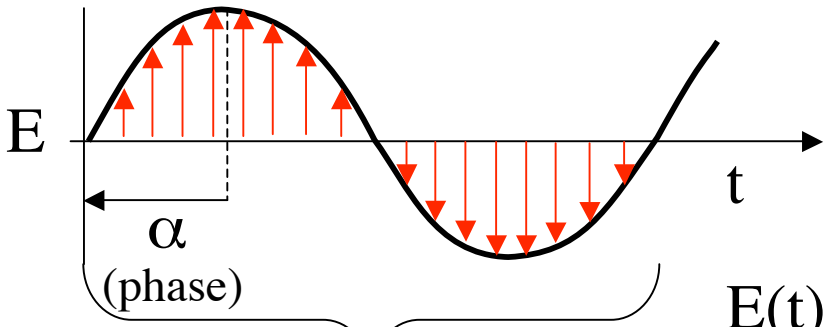
X-ray scattering & the Fourier tranform

light



The general equation for wave

Photons are oscillating electric fields*.



Amplitude

$E(t) = A \cos(\omega t + \alpha)$

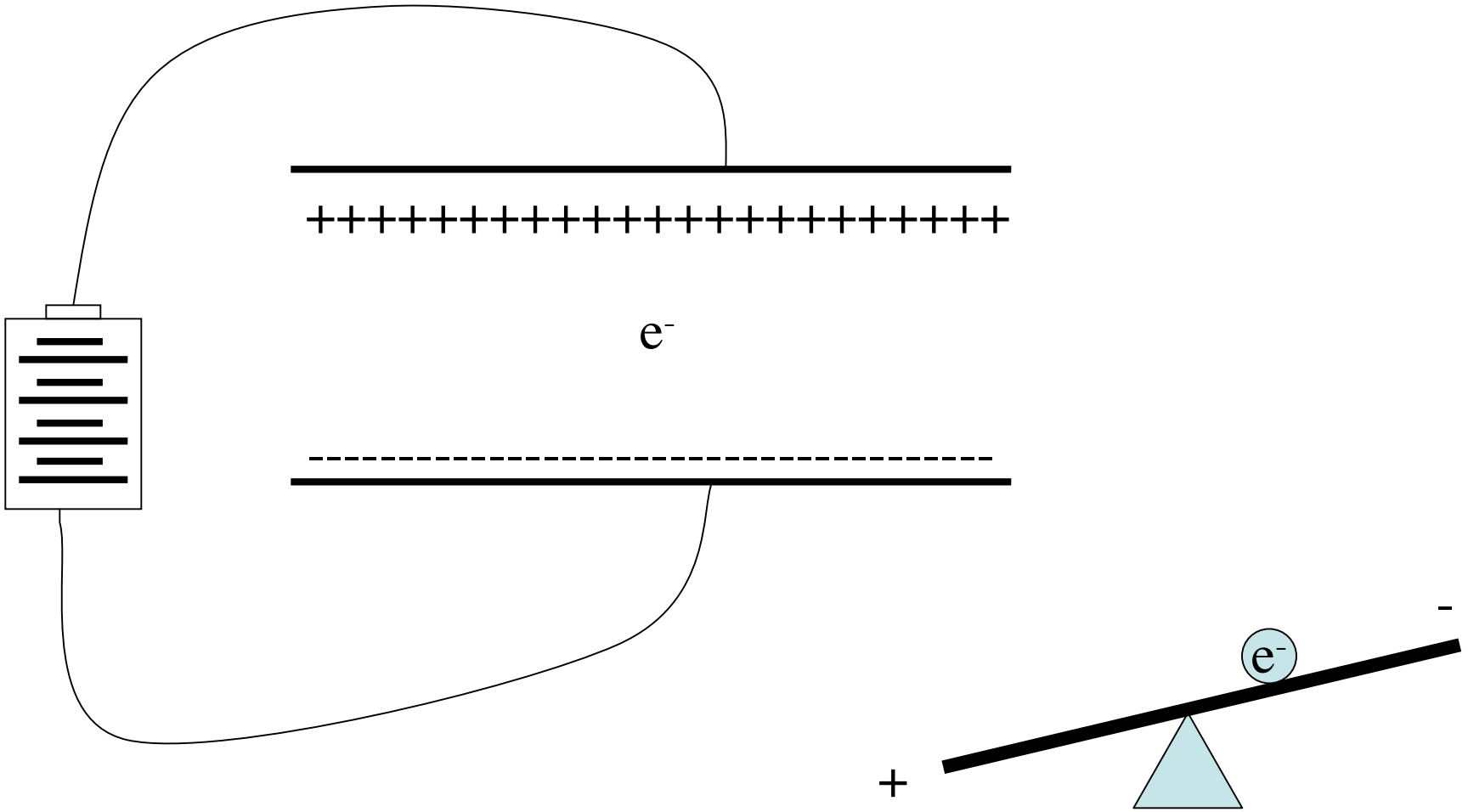
$\omega = 2\pi c/\lambda$

The instantaneous electric field at time **t**

oscillation rate in cycles/second

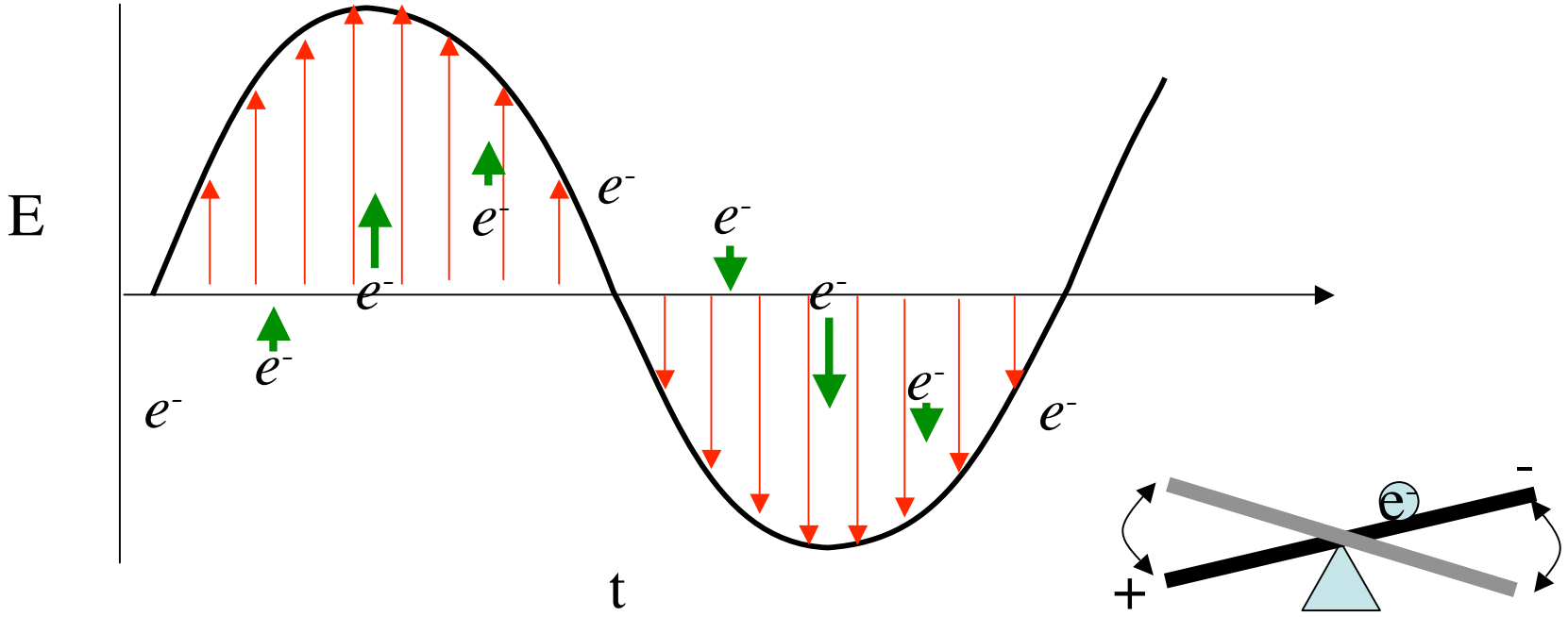
*also an oscillating magnetic field of the same frequency, 90 degrees out of phase.

An electric field accelerates charged particles



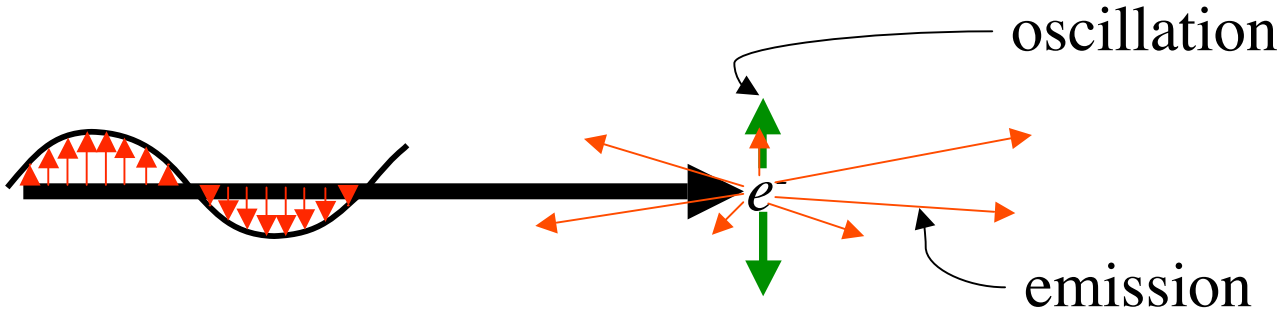
e^- oscillates in an electric field...

- *e*- oscillation is the same frequency as the X-rays
- *e*- oscillation is much faster that orbiting motion of *e*- around nucleus (no significant Doppler effect).
- The amplitude of the *e*- oscillation is large because the mass of an e^- is small. [Atomic nuclei don't oscillate much!]



oscillating e^- create photons

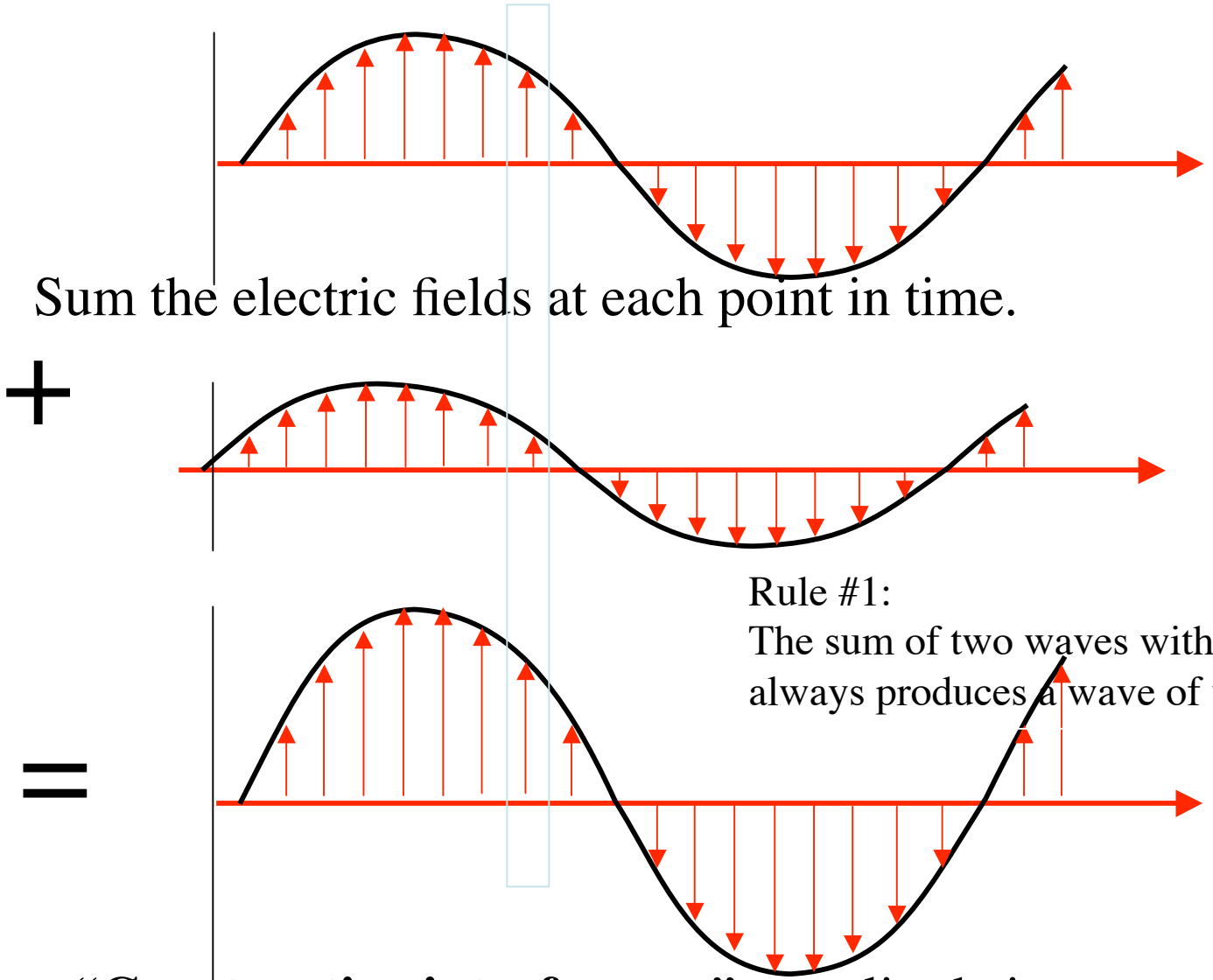
…in all directions ⊥ to the oscillation of incoming.



That's scattering.

X-ray sources may be partially polarized.

Wave addition

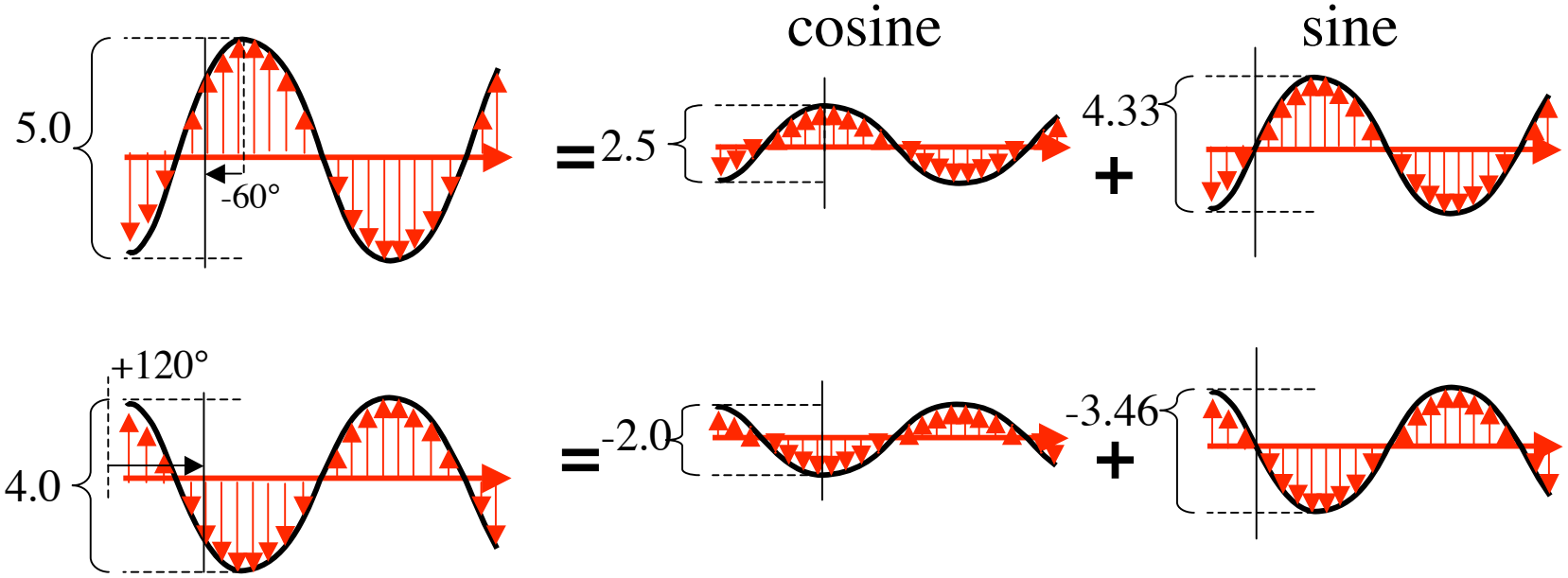


Sum the electric fields at each point in time.

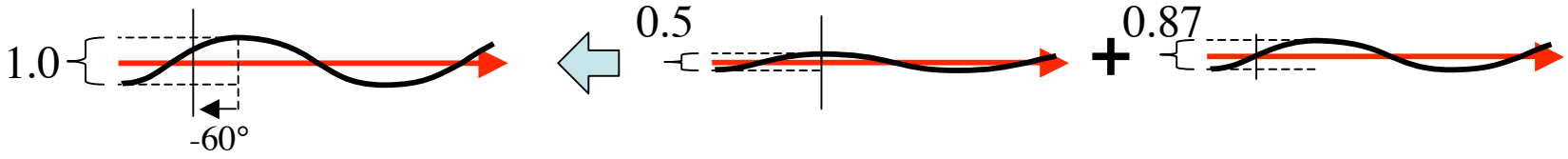
Rule #1:
The sum of two waves with wavelength λ always produces a wave of wavelength λ.

"**Constructive interference**": amplitude increases.

Adding two waves by parts



Add amplitudes of cosine and sine parts, then recombine them.



Cosine parts and Sine parts can be summed independently.
(like “orthogonal coordinates”)

The sum of angles rule

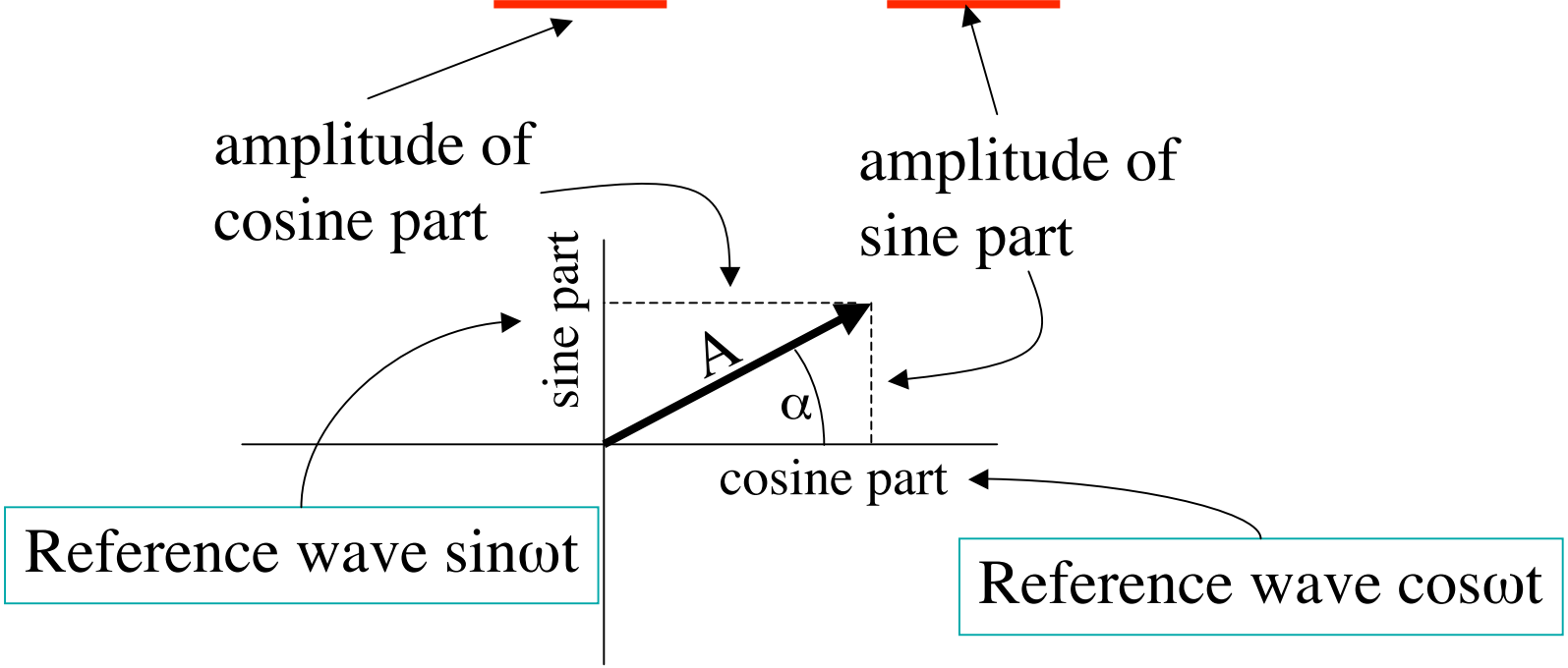
$$\cos(\alpha + \beta) = \cos \alpha \cos \beta - \sin \alpha \sin \beta$$

Applying the sum of angles rule to the wave equation decomposes it into sine and cosine parts

$E(t) = A \cos(\omega t + \alpha)$

Using the sum of angles rule:

$A \cos(\omega t + \alpha) = A \cos\alpha \cos\omega t - A \sin\alpha \sin\omega t$



A waves can be represented as a complex exponential

Euler's Theorem: $e^{i\alpha} = \cos\alpha + i\sin\alpha$

Proof: write the expansions and sum them

$$e^{\alpha} = 1 + \alpha + \alpha^2/2! + \alpha^3/3! + \alpha^4/4! + \alpha^5/5! - \cdots$$

$$e^{i\alpha} = 1 + i\alpha - \alpha^2/2! - i\alpha^3/3! + \alpha^4/4! + i\alpha^5/5! - \cdots$$

$$\cos\alpha = 1 - \alpha^2/2! + \alpha^4/4! - \alpha^6/6! - \cdots$$

$$i\sin\alpha = i\alpha - i\alpha^3/3! + i\alpha^5/5! - i\alpha^7/7! + \cdots$$

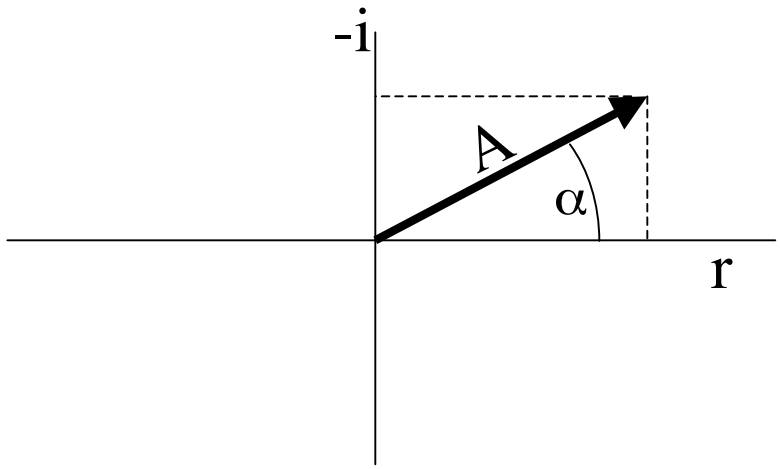
A wave as a complex exponential

$$E(t) = A\cos(\omega t + \alpha)$$

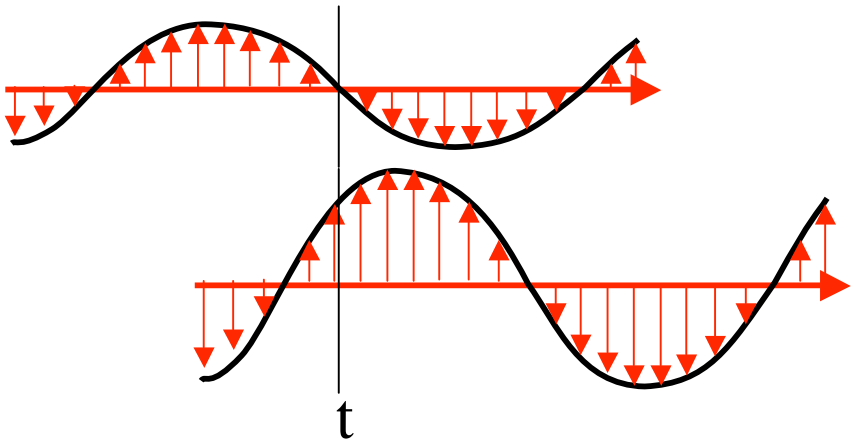
$$= A\cos\alpha\cos\omega t - A\sin\alpha\sin\omega t$$

($\cos\omega t$ → **1**, $\sin\omega t$ → ***-i***)

$$= \mathbf{A\, e^{i\alpha}}$$

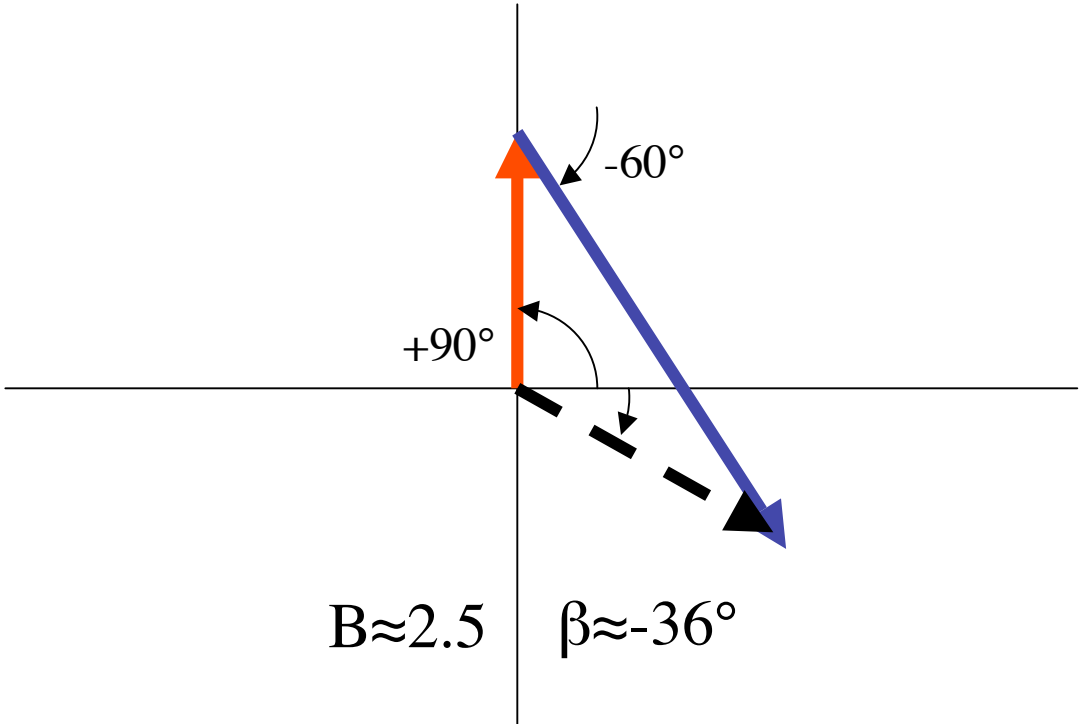


Adding waves using vector addition



$A_1=2.0 \quad \alpha_1= +90°$

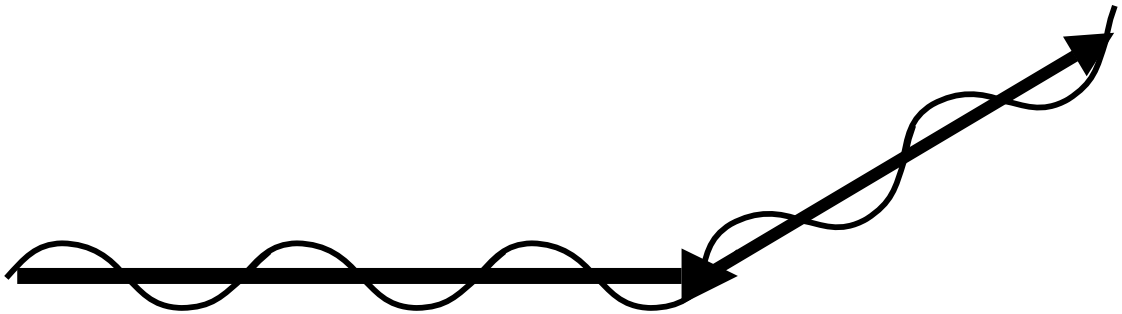
$A_2=4.0 \quad \alpha_2= -60°$



$B \approx 2.5 \quad \beta \approx -36°$

Phase depends on the distance traveled

Phase = D/ λ – nearest integer(D/λ)

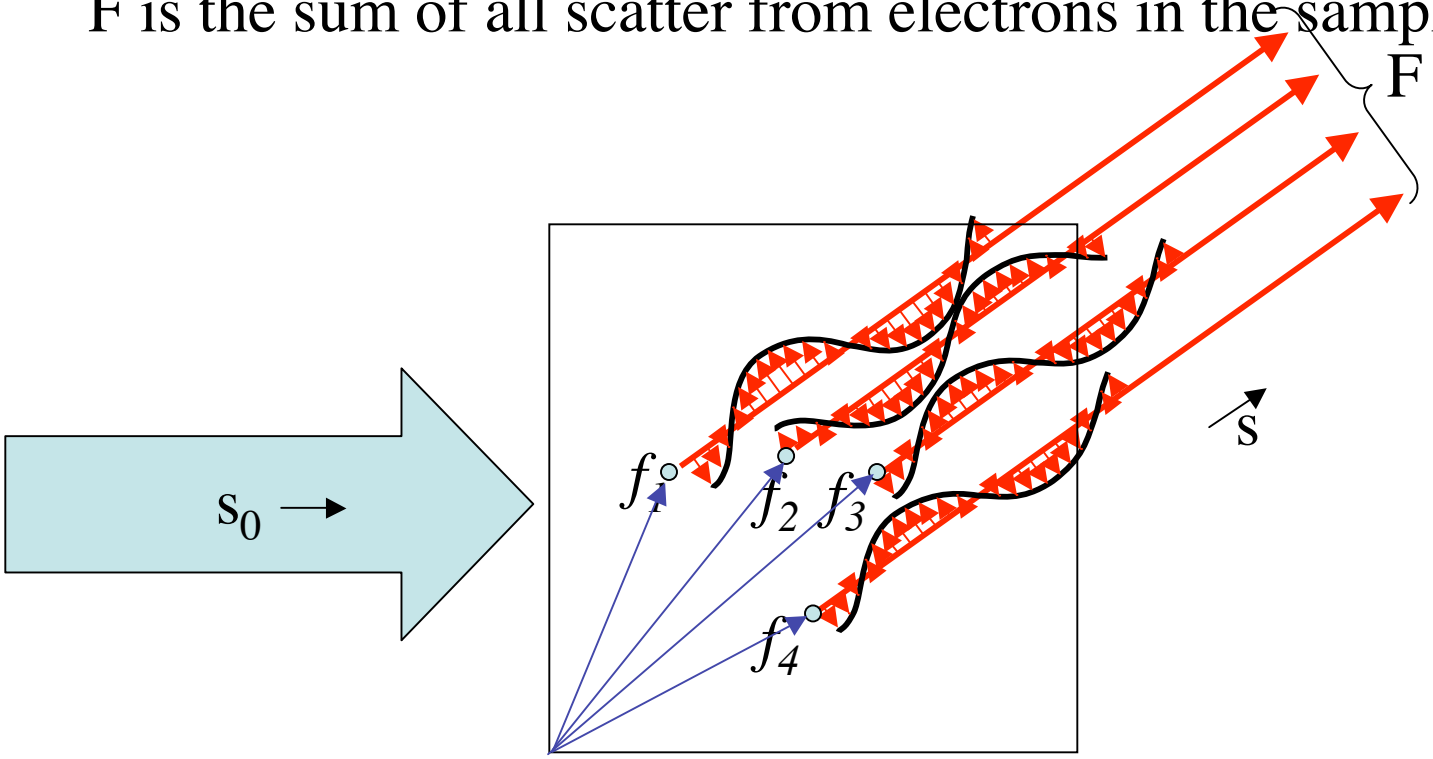


Same for scattered path

definition: Structure factor (F)

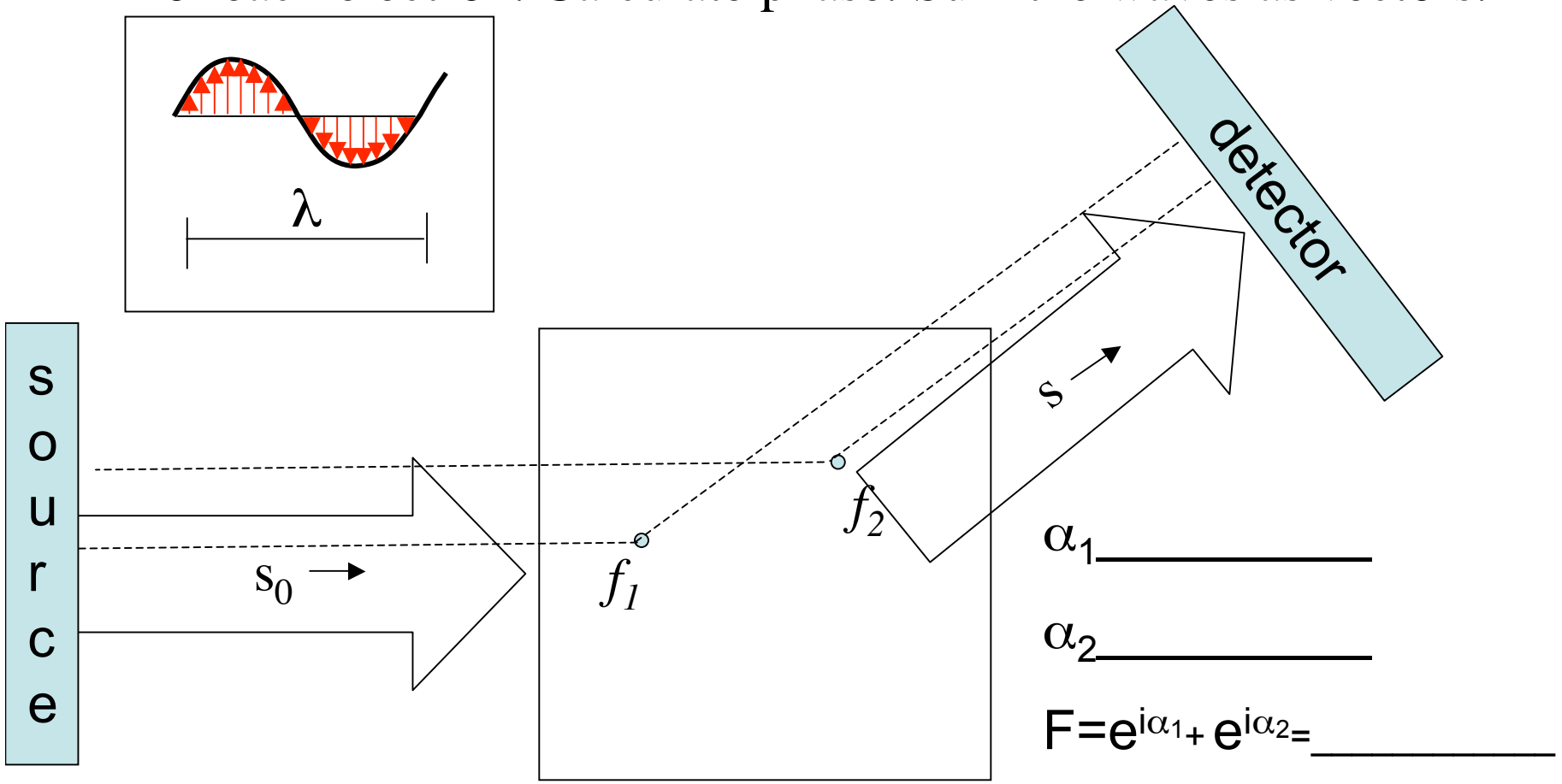
...is a single scattered wave (photon) *F*, having an *amplitude*, *phase* and *direction* relative to the crystallographic reference frame.

F is the sum of all scatter from electrons in the sample.

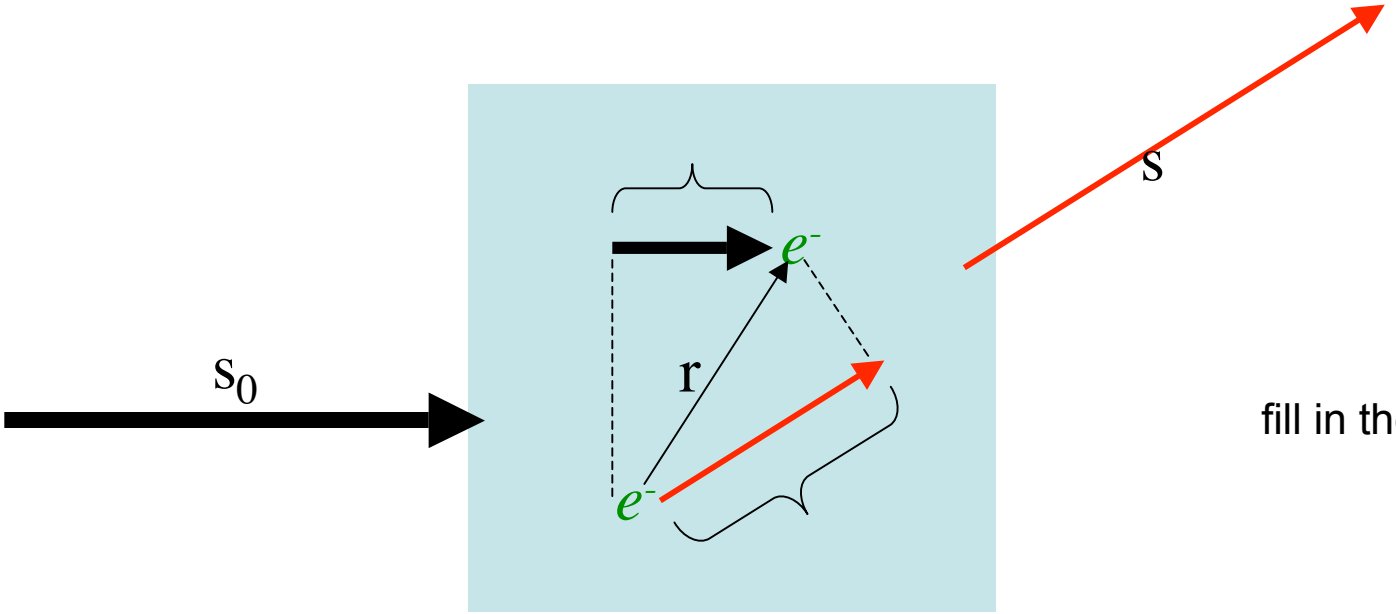


In class exercise: sum the scattered waves.

Measure the distance traveled from the source to the detector for each electron. Calculate phase. Sum the waves as vectors.



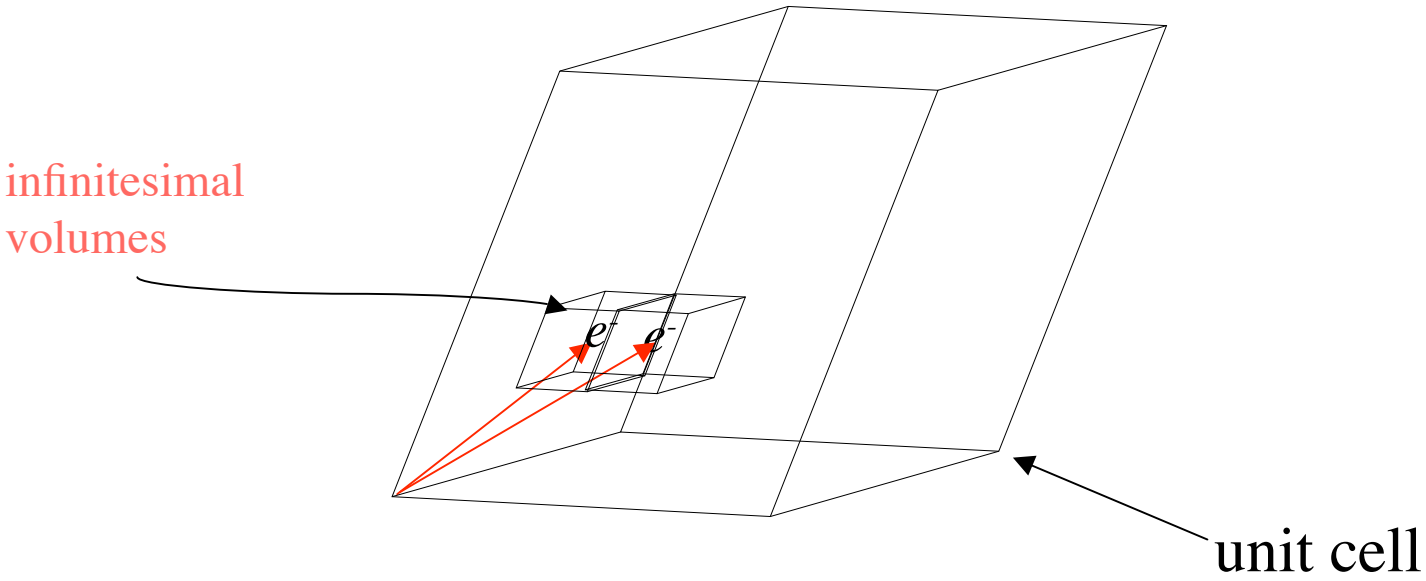
Path difference to get phase



fill in the dot products

Phase difference = length difference divided by the wavelength, multiplied by 2π.

Fourier Transform, an integral of waves

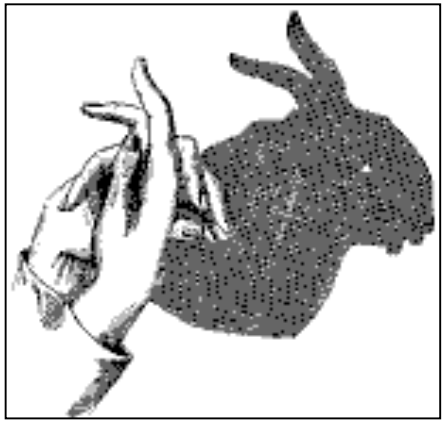


$$F(S) = \int \rho(r) e^{i2\pi S \bullet r} dr$$

ρ=electron density

Inverse Fourier Transform

There exists an inverse Fourier Transform which when substituted into the FT produces an identity.



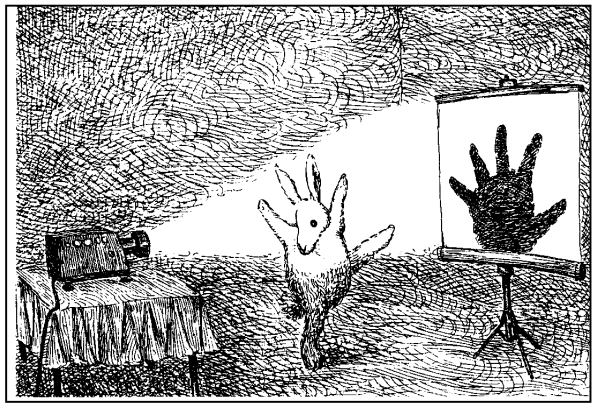
forward transform

$$F(S) = \int \rho(r) e^{i 2\pi S \bullet r} dr$$

reverse transform

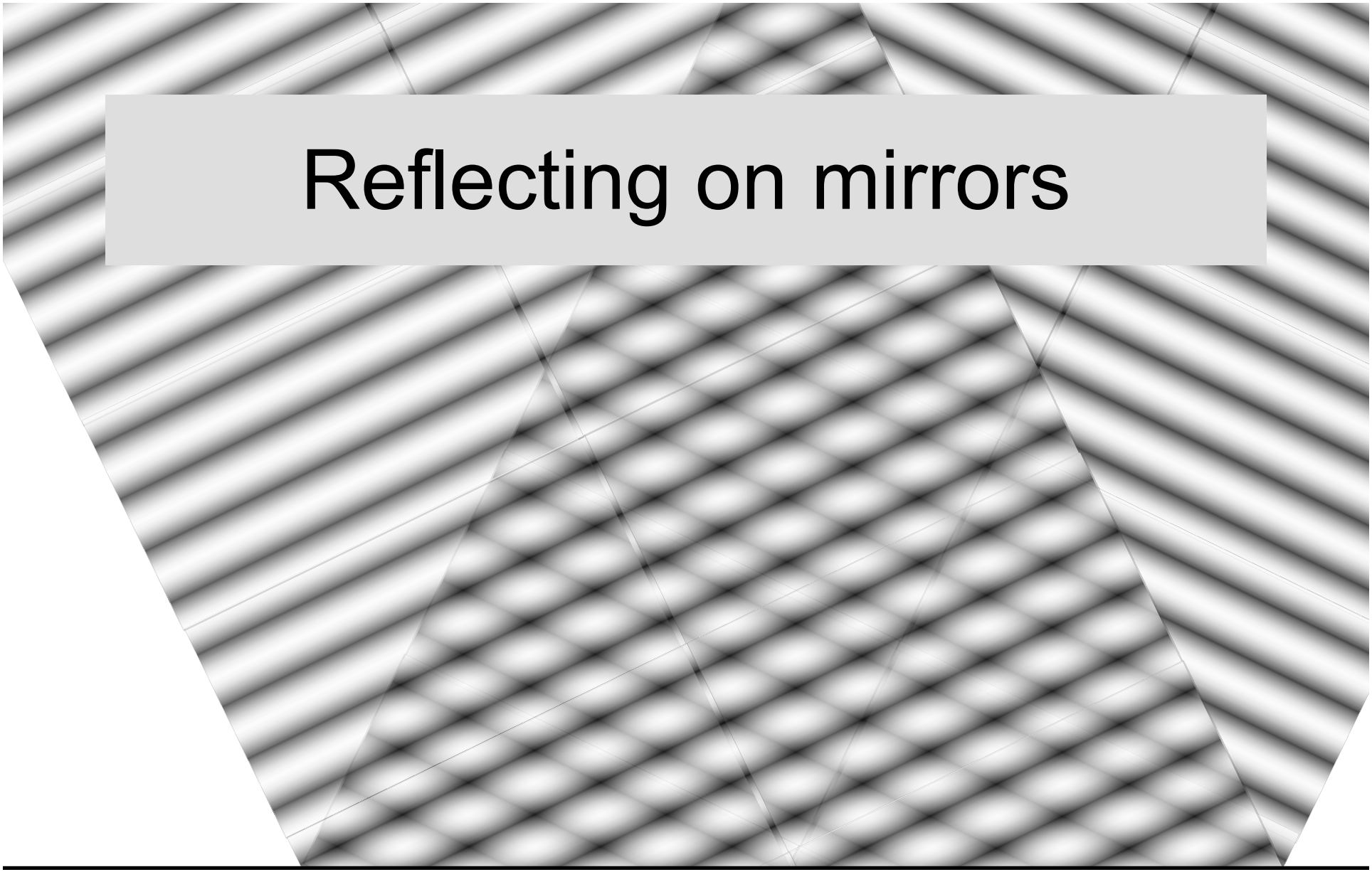
$$\rho(r) = \int F(S) e^{-i 2\pi S \bullet r} d(S)$$

note the minus sign



Bragg's Law and Diffraction

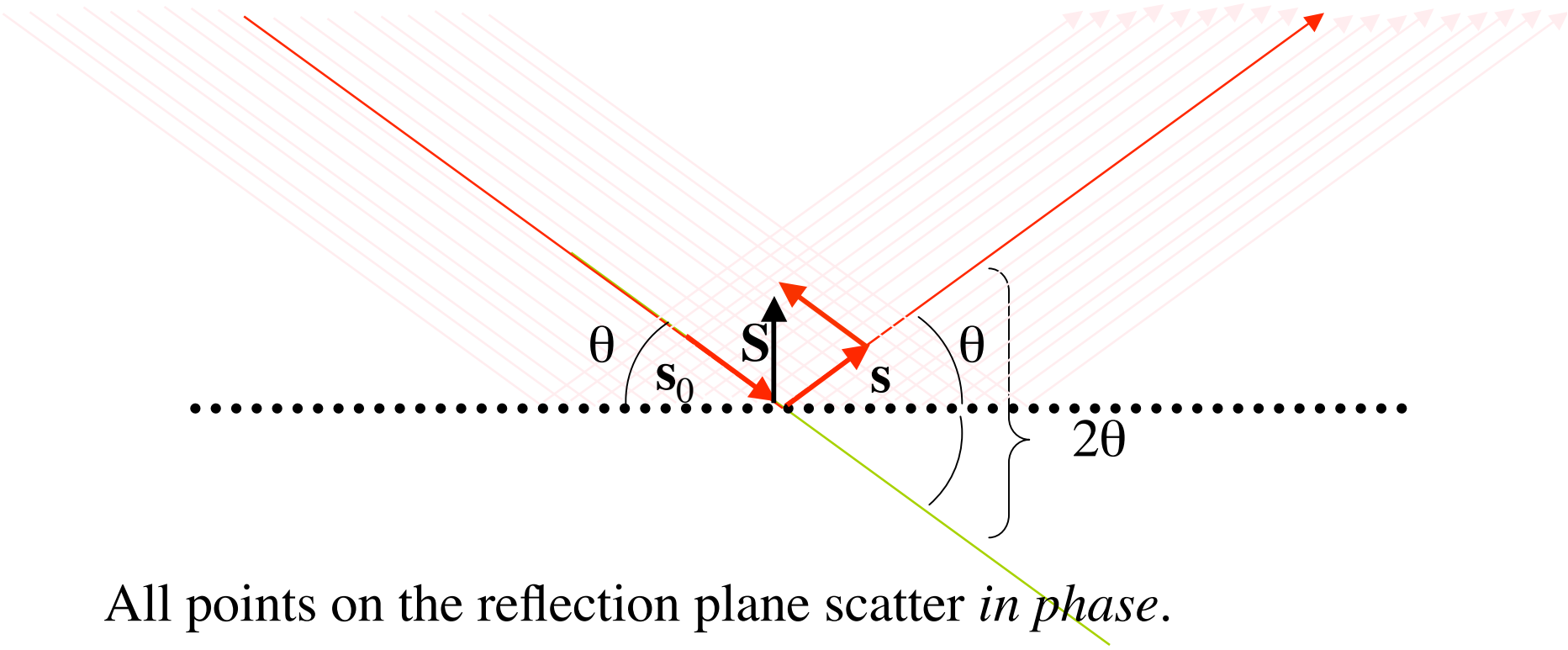
Reflecting on mirrors



Although sometimes drawn as a vector, X-rays are plane waves.

Reflection plane

same angle, θ, with beam and scattered Xray



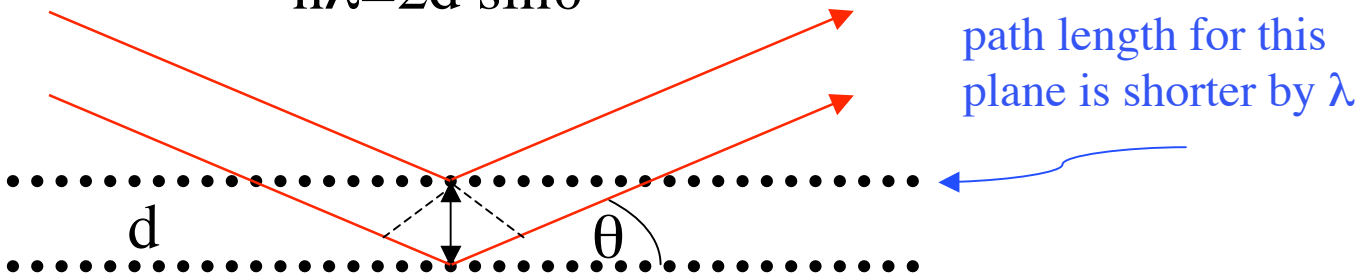
All points on the reflection plane scatter *in phase*.

That's why a single direction of scatter is called a ***reflection***.

Sir Lawrence Bragg, winner of 1914 Nobel Prize

Bragg's law

$n\lambda = 2d \sin\theta$

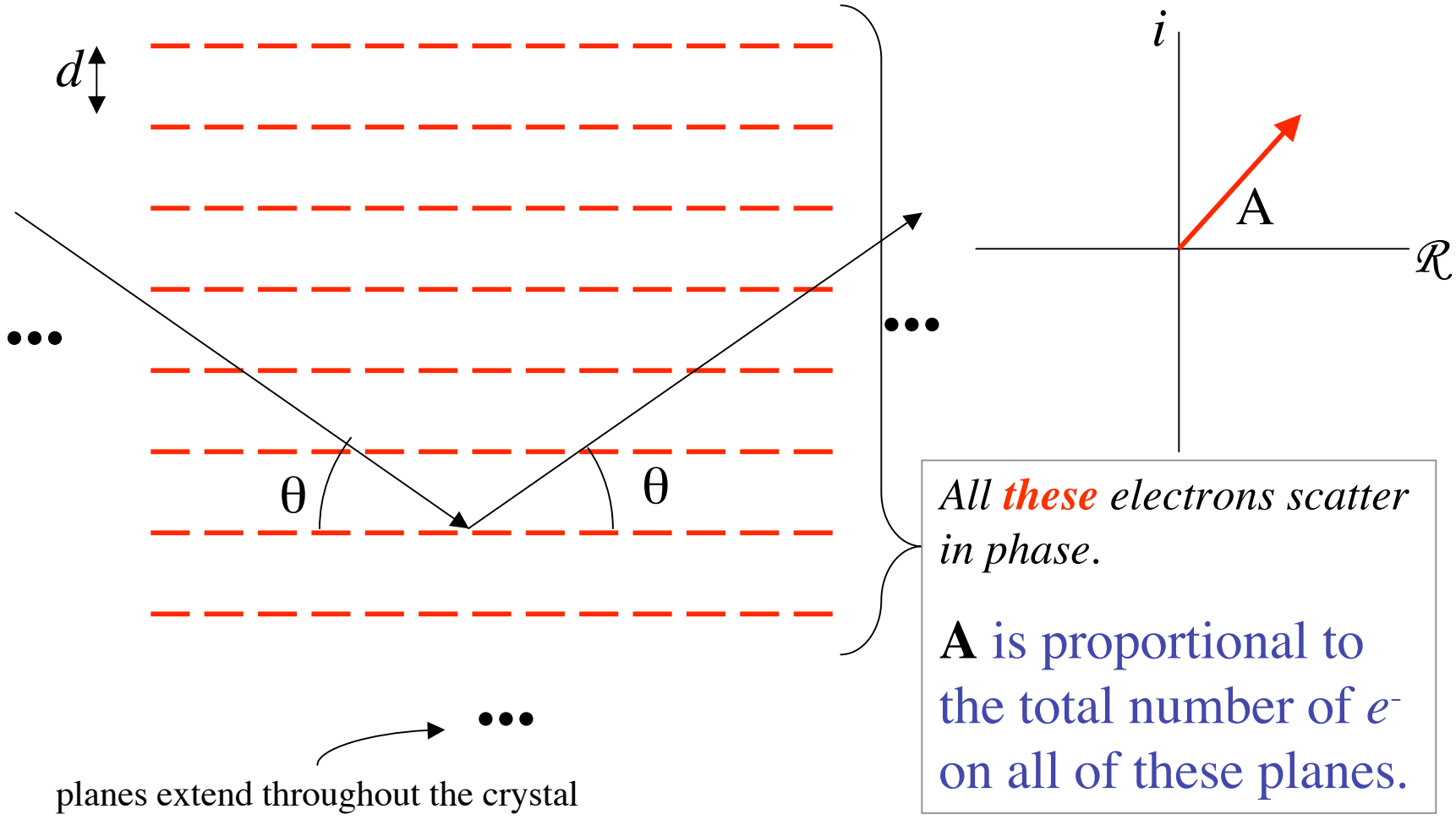


Reflection planes separated by *d* scatter in phase.

If θ is larger, *d* is smaller.

$d = \lambda / 2\sin\theta$

Scattering by Bragg planes

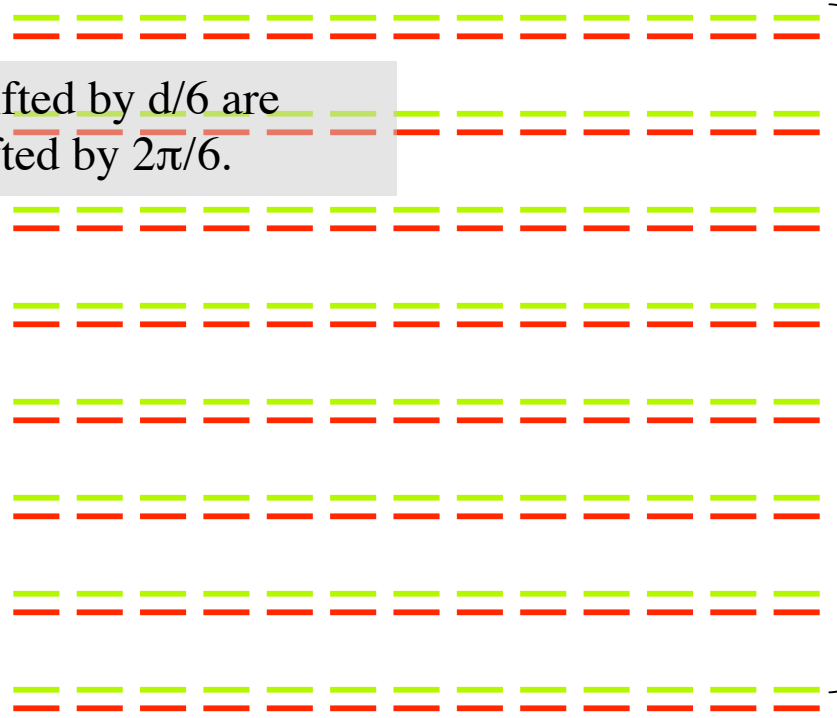


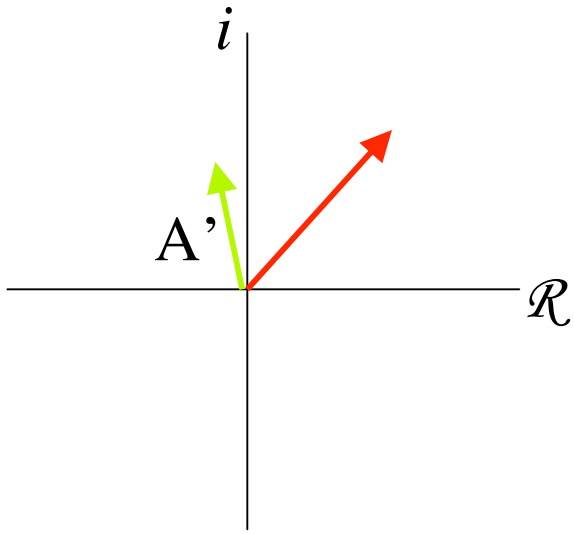
*All **these** electrons scatter in phase.*

A is proportional to the total number of e^- on all of these planes.

Integrating planes separated by *d*

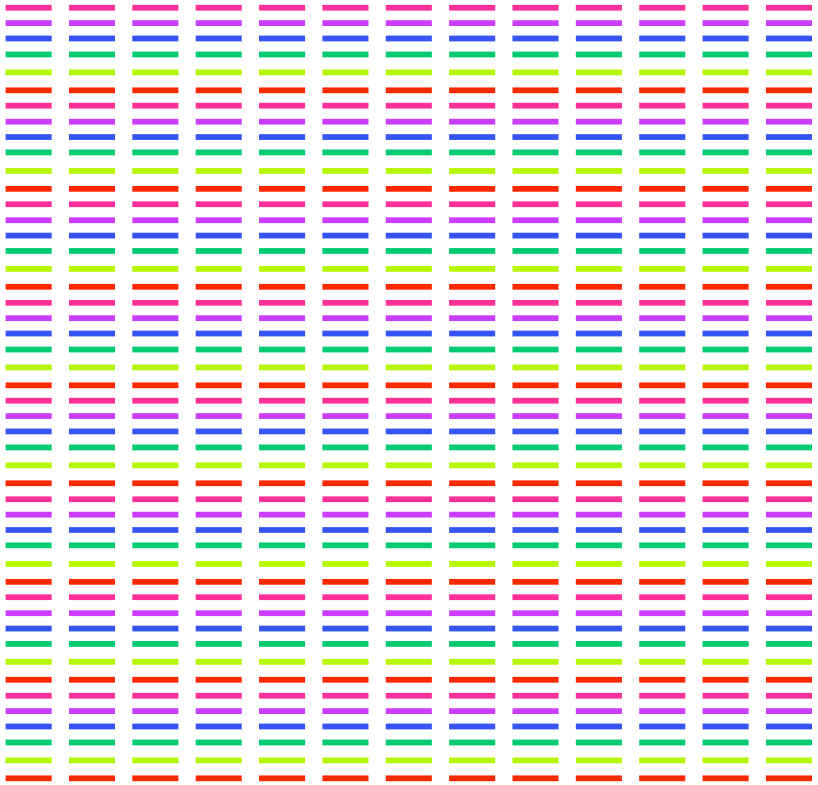
Planes shifted by d/6 are phase shifted by 2π/6.





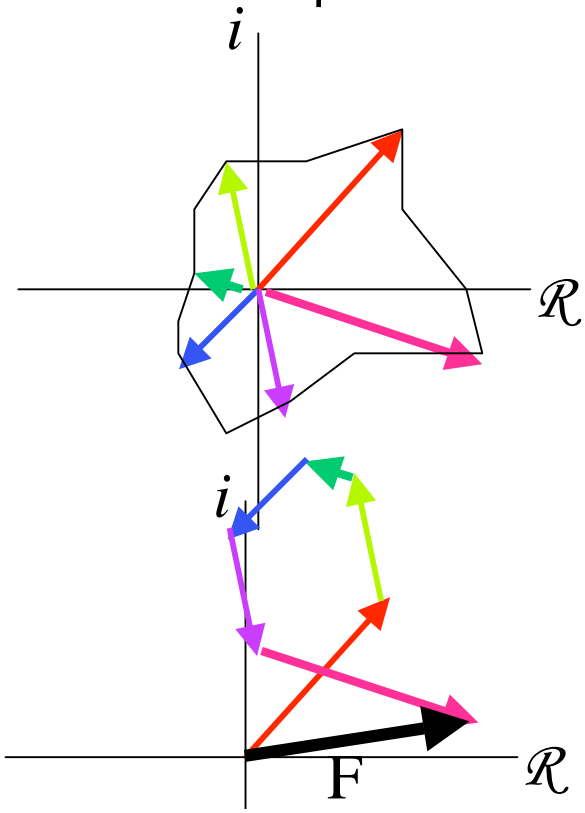
*All **these** electrons scatter in phase, and the phase is shifted by 2π/6 or 60°.*

Integrating planes separated by *d*



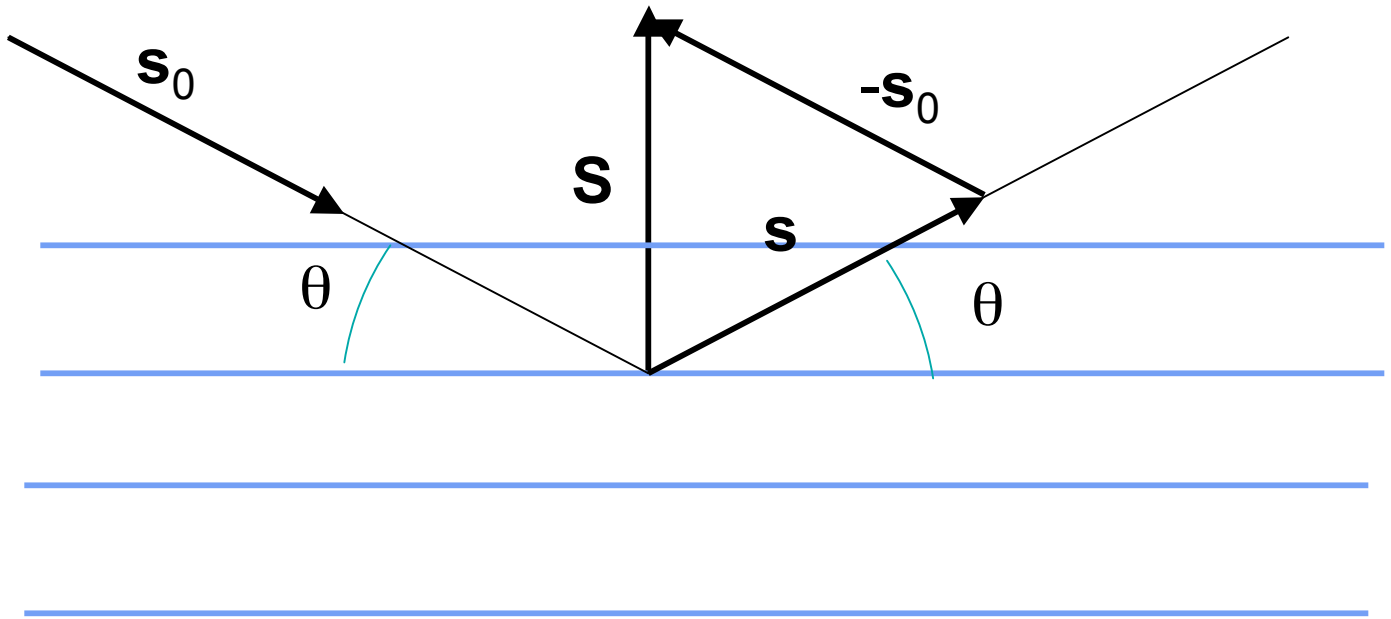
The total F is the wave sum over all Bragg planes.

amplitude as a function of phase



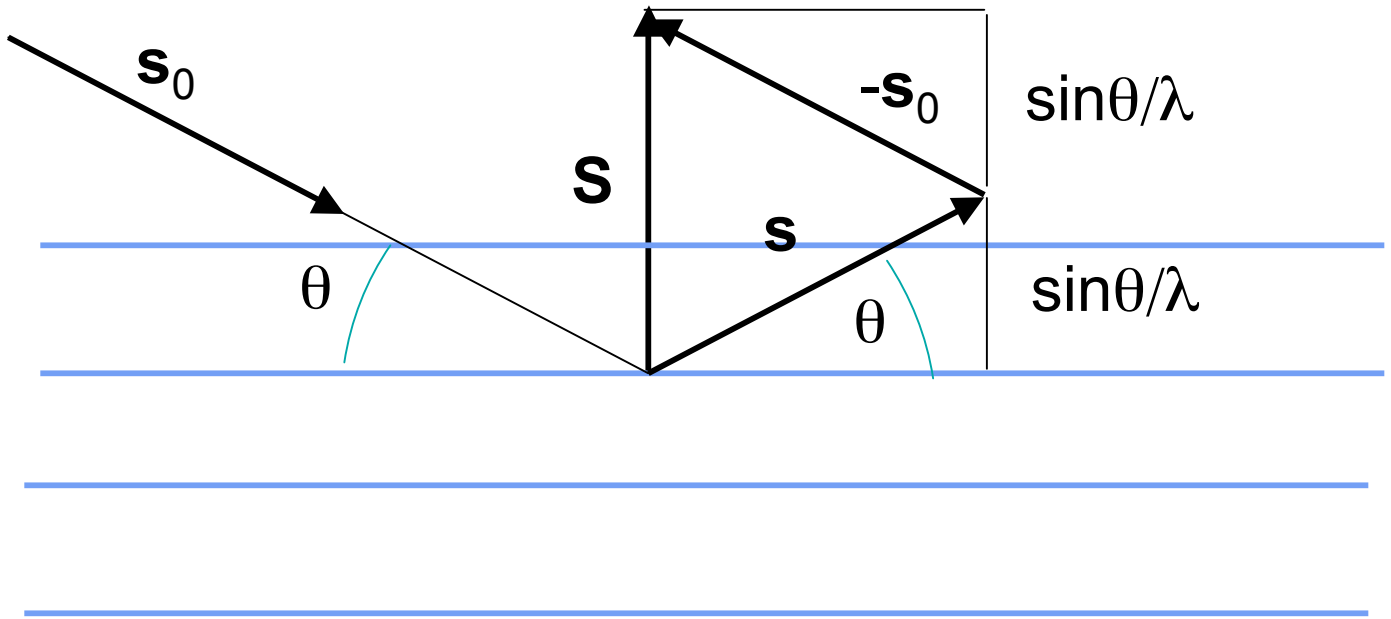
integrating (complex!) over phase gives **F**.

Bragg planes are always perpendicular to **S**



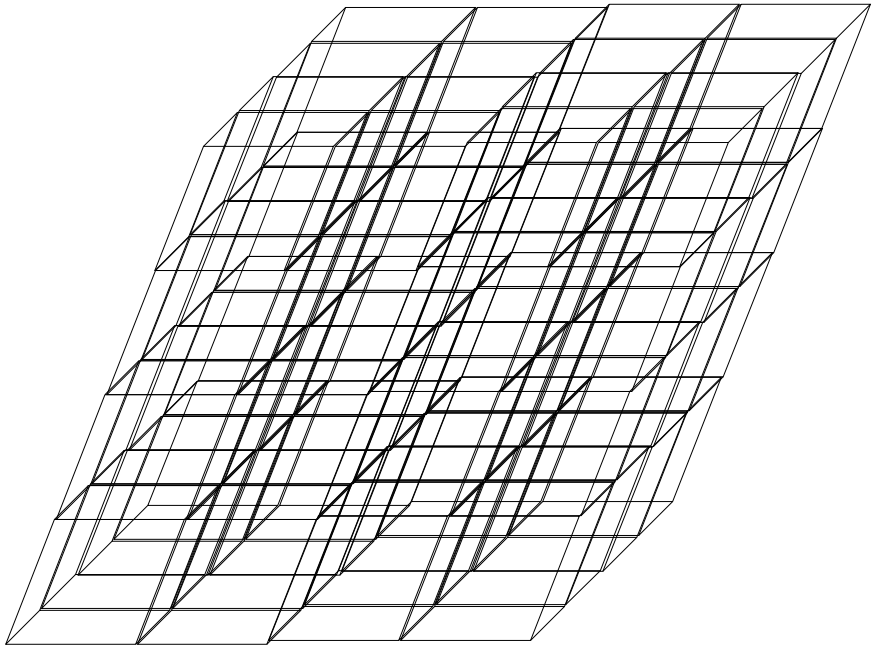
Since $\mathbf{s}_0$ and **s** are the same length and have the same angle to the reflection plane, $\mathbf{S} = (\mathbf{s}-\mathbf{s}_0)/\lambda$ is normal to the plane.

The length of **S** is 1/d



The length of **S** is 2sinθ times the lengths of **s** and $\mathbf{s}_0$, which is 1/λ. So |**S**| = 2sinθ/λ = 1/d

Crystal = 3D lattice



A vector expression that defines a crystal lattice:

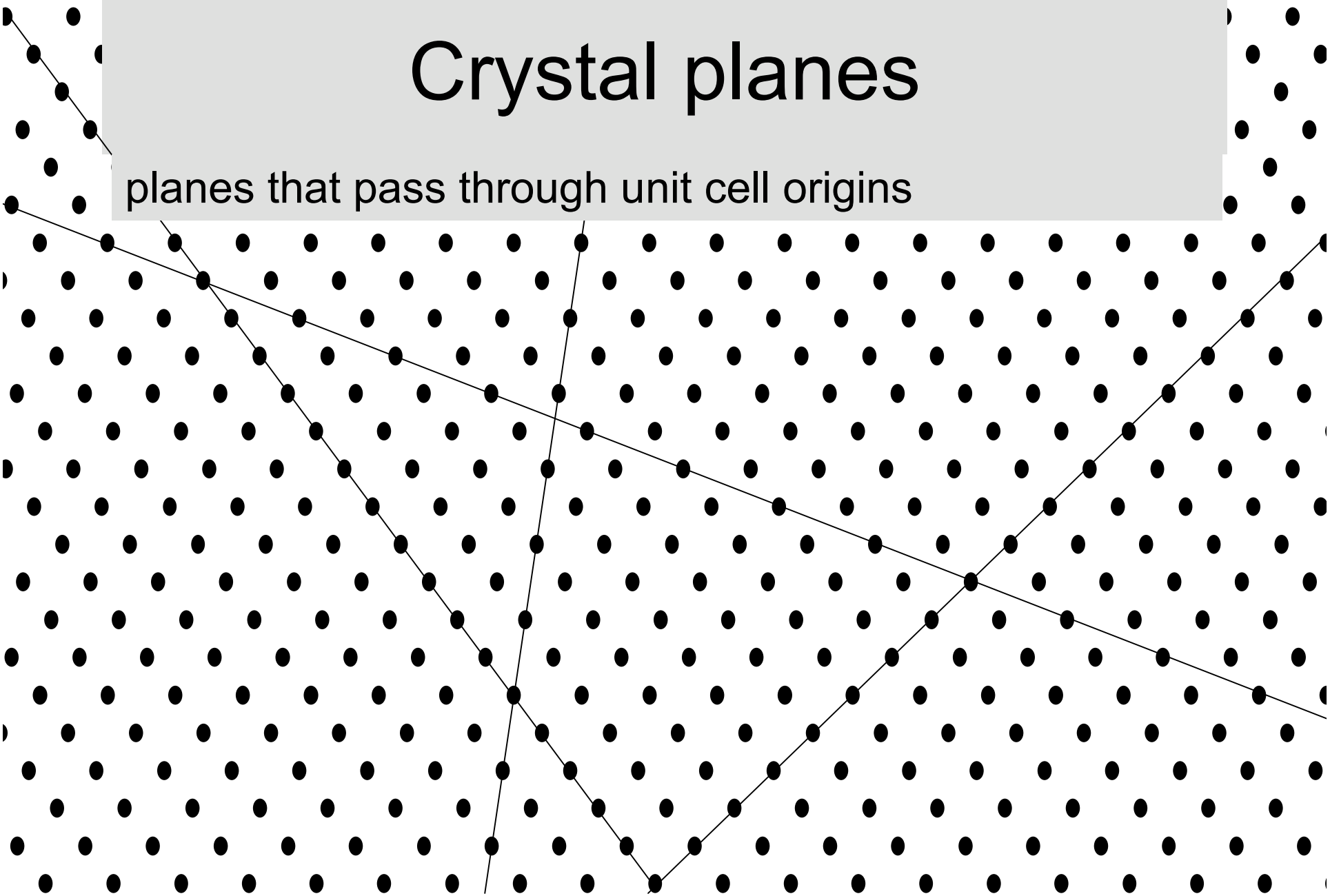
$$\rho(r) = \rho(\boldsymbol{r} + t\mathbf{a} + u\mathbf{b} + v\mathbf{c})$$

where a,b,and c are the unit cell axes and t,u, and v are integers.

ρ=electron density

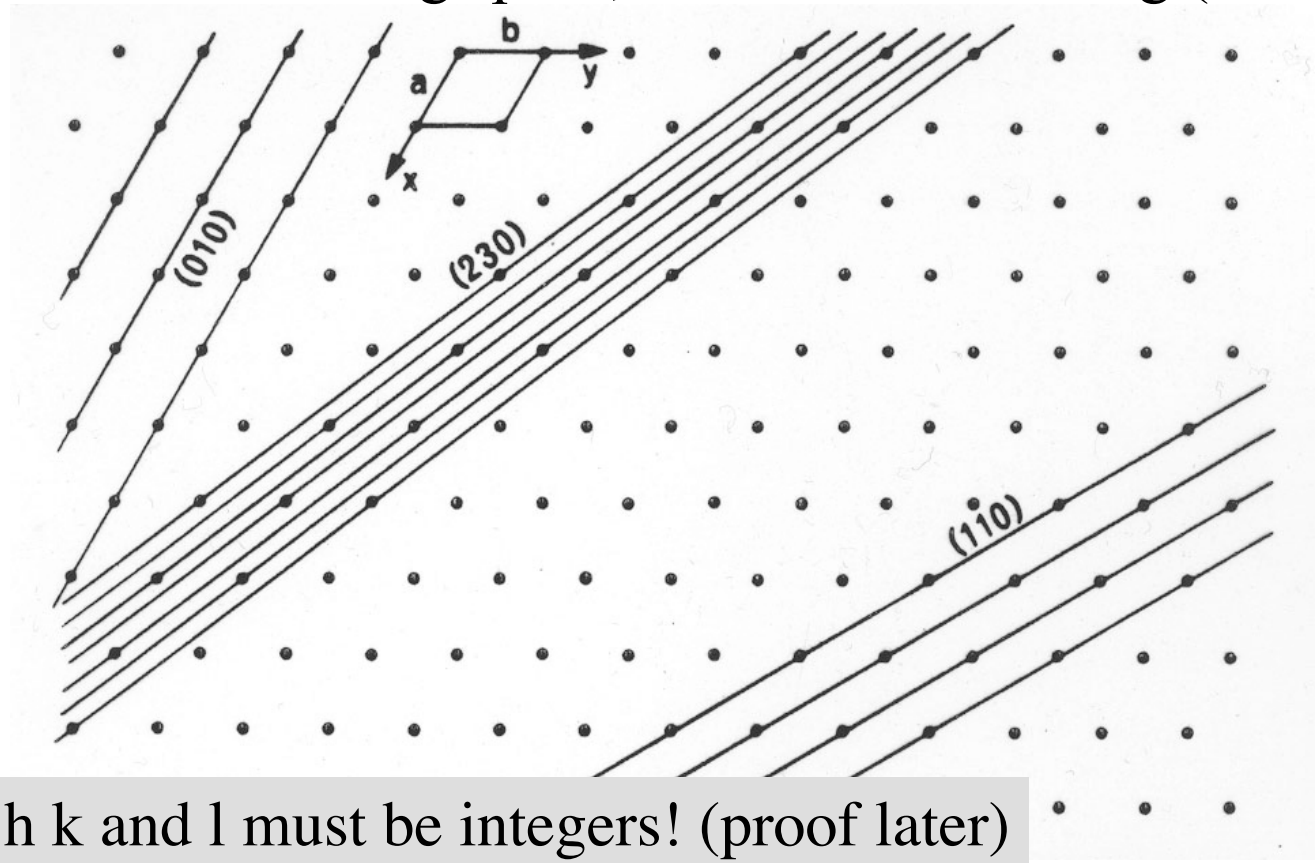
Crystal planes

planes that pass through unit cell origins



Crystal plane numbers: Miller indeces

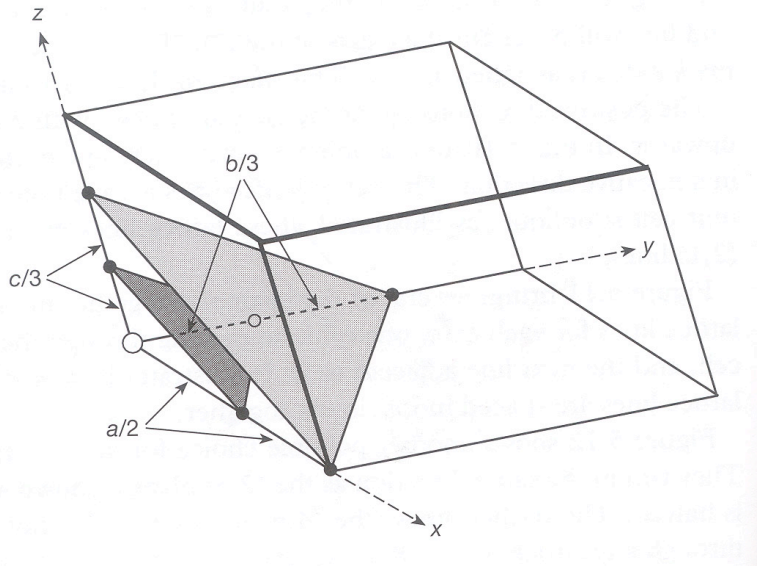
Starting from the Origin and moving to the next plane, if it intersects
the **a** axis at 1/h,
the **b** axis at 1/k and
the **c** axis at 1/l,
...then the planes are called (h k l). Each set of crystal planes defines a "reflection" in scattering space, also numbered using (h k l).



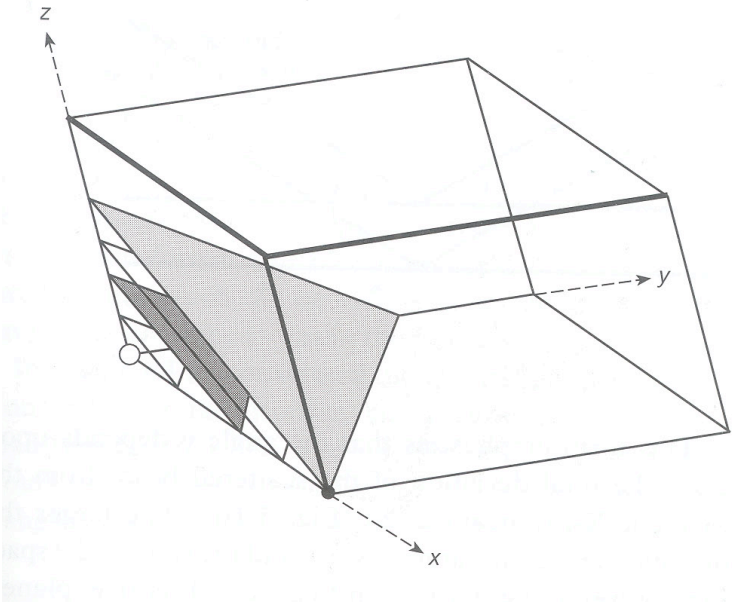
NOTE: h k and l must be integers! (proof later)

Examples of 3D crystal planes

(2 3 3) crystal planes



(4 6 6) crystal planes



planes intersect the cell axes at fractional coordinates

(1/h,0,0), (0,1/k,0),(0,0,1/l)

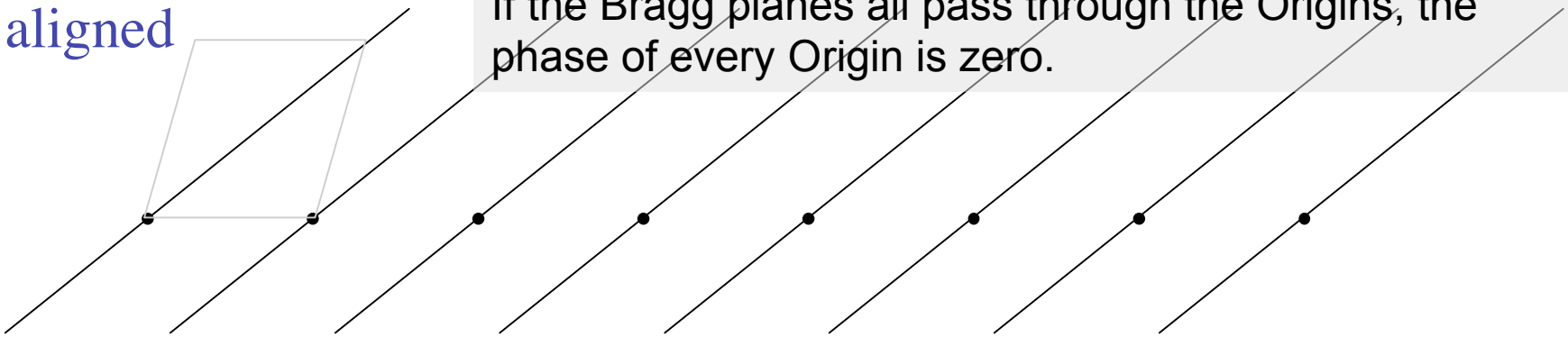
Proof: The only Bragg planes that diffract X-rays are those that match crystal planes

In other words, if we have seen a reflection on the film, that reflection corresponds to a set of crystal planes. Since all crystal planes pass through the unit cell origins, and since the phase of the Origin can be set to zero, *all (observable) Bragg planes of phase zero pass through the Origins*

Proof: All Bragg planes of phase zero pass through the Origins

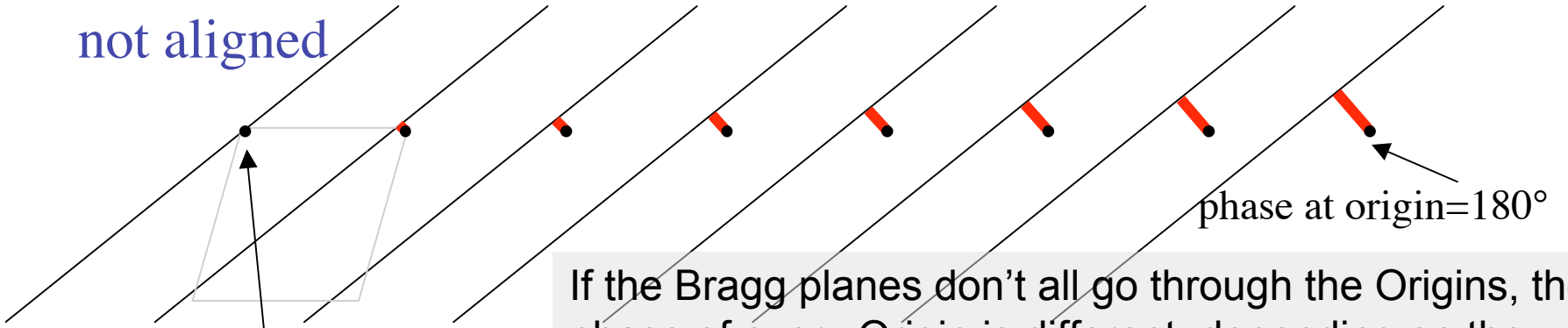
1. Bragg planes are either aligned with the Unit Cell Origins, or they are not.

aligned



If the Bragg planes all pass through the Origins, the phase of every Origin is zero.

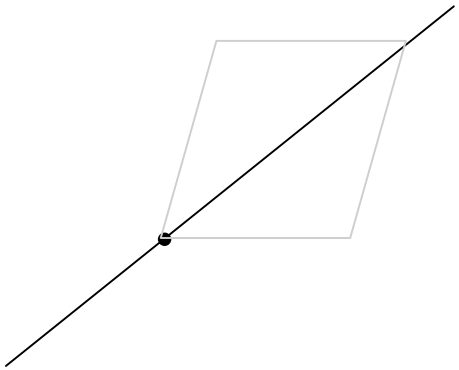
not aligned



If the Bragg planes don't all go through the Origins, then phase of every Origin is different, depending on the distance to the Bragg plane.

Proof: All Bragg planes of phase zero pass through the Origins

2. All planes that pass through the Origins have the same number of electrons



The angle and intercept with the Unit Cell determine with atoms are on the plane.

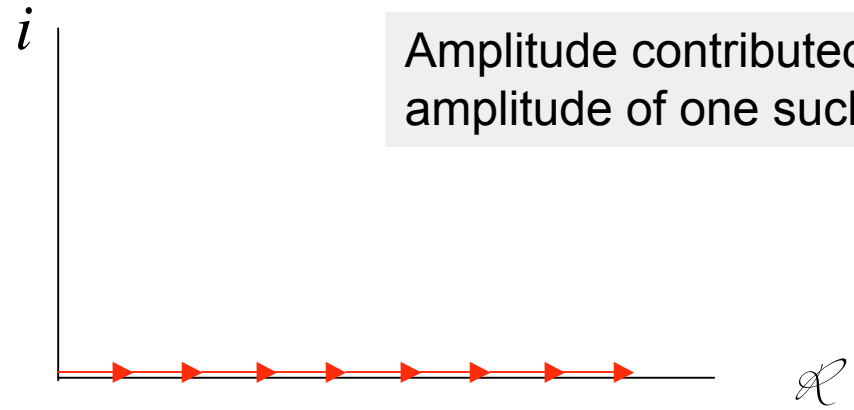
3. All planes that pass through the Origins contribute the same amplitude.

...because amplitude is proportional to number of electrons, and (statement 2).

→

Proof: All Bragg planes of phase zero pass through the Origins

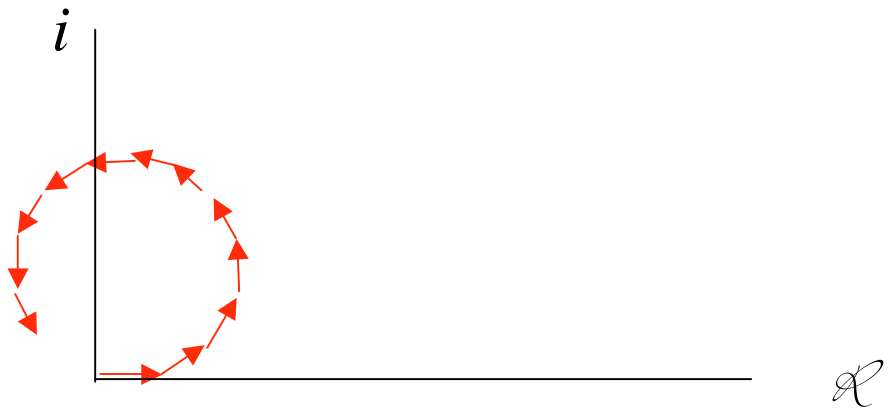
4. Total amplitude is the sum of the amplitudes of the planes if the planes have the same phase.



Amplitude contributed by origin planes is 10K times the amplitude of one such plane, if there are 10K unit cells.

5. Total amplitude is approximately zero if the planes have different phases.

Phase shifts by a constant for each unit cell. Vectors sum in a circle. Summed over 10K unit cells, vector length is small.



Proof: All Bragg planes of phase zero pass through the Origins

6. Any point in the Unit Cell can be the Origin.

7. All equivalent positions by lattice symmetry have the same phase.

Because of (statement 6), statements 1-5 apply to any point in the Unit cell.

8. If the Bragg planes do not pass through all Origins, the diffraction amplitude is zero.

Because the total diffraction amplitude is the wave sum over all points in the Unit Cell.

Conclusion: Bragg planes that pass through all of the Origins diffract X-rays. Bragg planes that do not pass through all of the origins do not diffract X-rays.

Definition of the Reciprocal Lattice

Let's define the reciprocal lattice as the subset of points in S for which the Laue conditions hold: $S = (h\mathbf{a}^*, k\mathbf{b}^*, l\mathbf{c}^*)$ then...

$$\mathbf{a}^* \bullet \mathbf{a} = 1$$

$$\mathbf{b}^* \bullet \mathbf{b} = 1$$

$$\mathbf{c}^* \bullet \mathbf{c} = 1$$

reciprocal lattice axes

See Drenth Ch 4, p 86, Table 4.1 for how to calculate the reciprocal lattice vectors a*b*c*

Complete Laue conditions

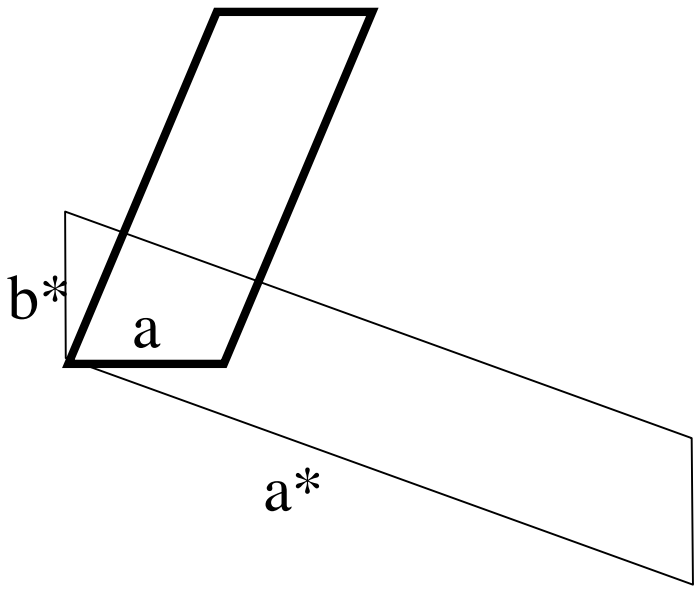
$$\mathbf{a}^* \bullet \mathbf{a} = 1 \qquad \mathbf{b}^* \bullet \mathbf{a} = 0 \qquad \mathbf{c}^* \bullet \mathbf{a} = 0$$

$$\mathbf{a}^* \bullet \mathbf{b} = 0 \qquad \mathbf{b}^* \bullet \mathbf{b} = 1 \qquad \mathbf{c}^* \bullet \mathbf{b} = 0$$

$$\mathbf{a}^* \bullet \mathbf{c} = 0 \qquad \mathbf{b}^* \bullet \mathbf{c} = 0 \qquad \mathbf{c}^* \bullet \mathbf{c} = 1$$

Real cell relationship to reciprocal cell

2D for simplicity



If $a < b$, the $a^* > b^*$

$a^* \perp b \quad a^* \perp c$

$b^* \perp a \quad b^* \perp c$

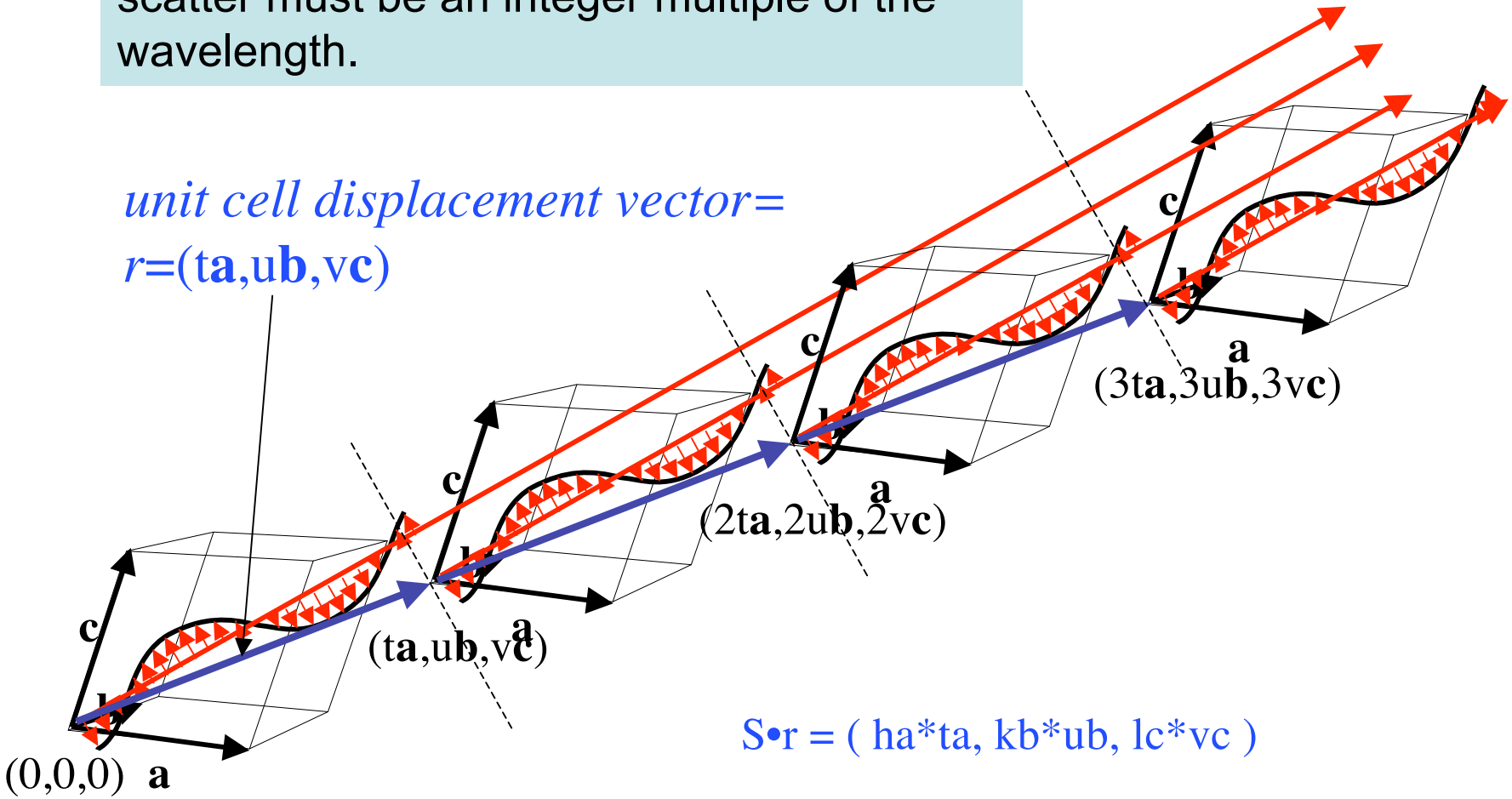
$c^* \perp a \quad c^* \perp b$

Adding unit cells

For unit cell scatter to add constructively, the dot product of the displacement and scatter must be an integer multiple of the wavelength.

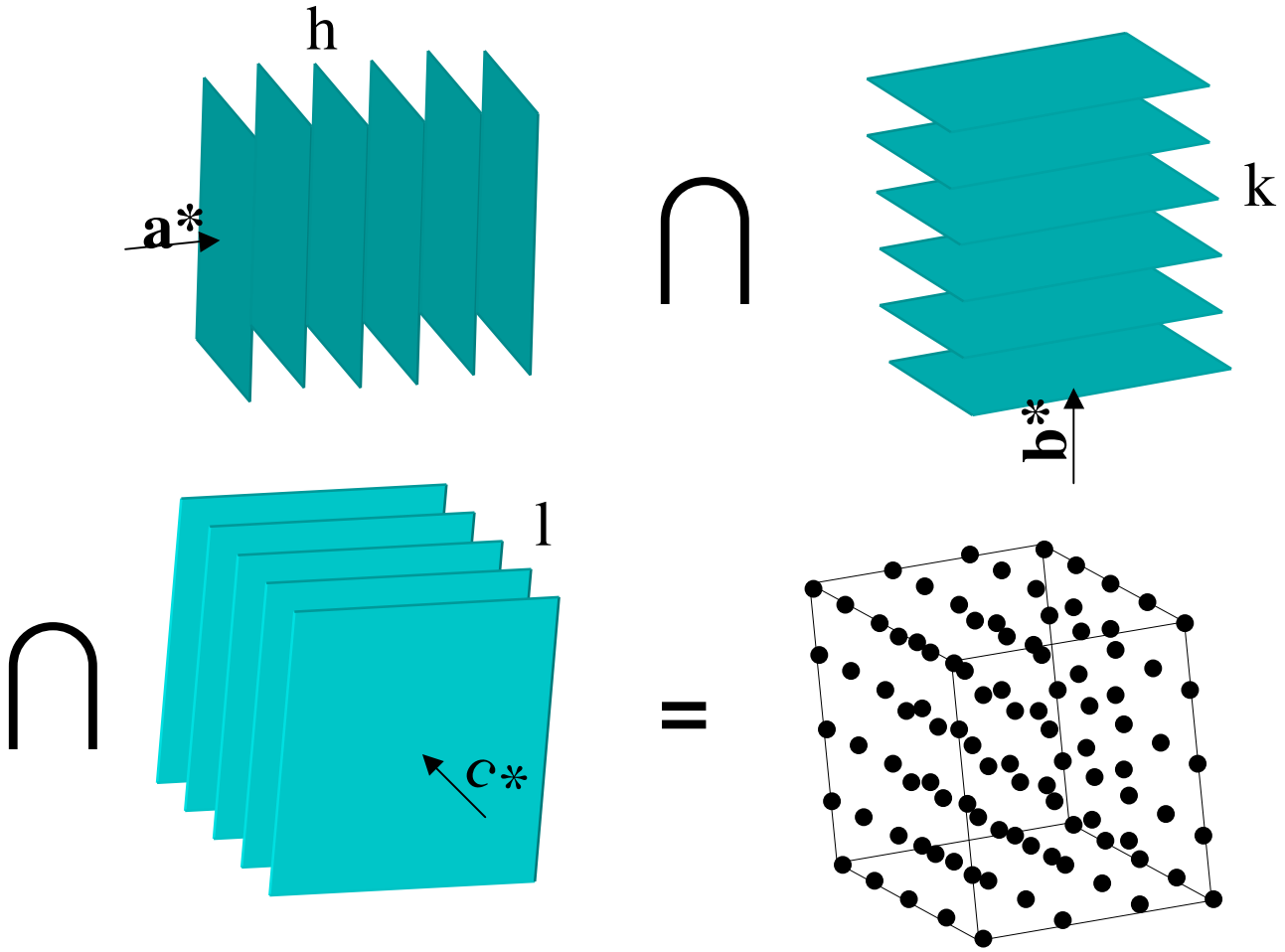
scattering vector=
S=(h**a***,k**b***,l**c***)

unit cell displacement vector=
r=(t**a**,u**b**,v**c**)



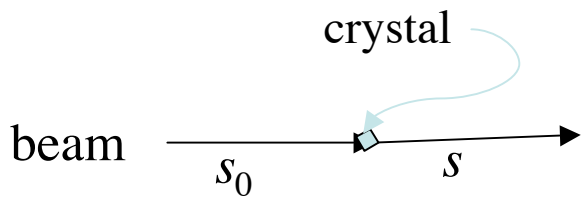
S•r = (ha*ta, kb*ub, lc*vc)

Reciprocal lattice: A periodic delta function in three dimensions

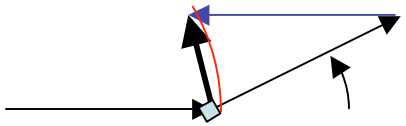


…results in points of diffraction, in **S** space.

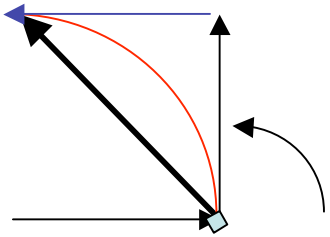
Ewald Sphere
Visible reciprocal space with the beam fixed



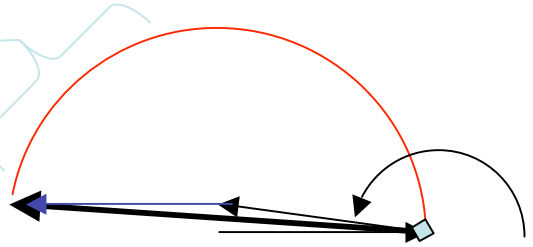
$2\theta = 0.$



$2\theta = 30°$



$2\theta = 90°$



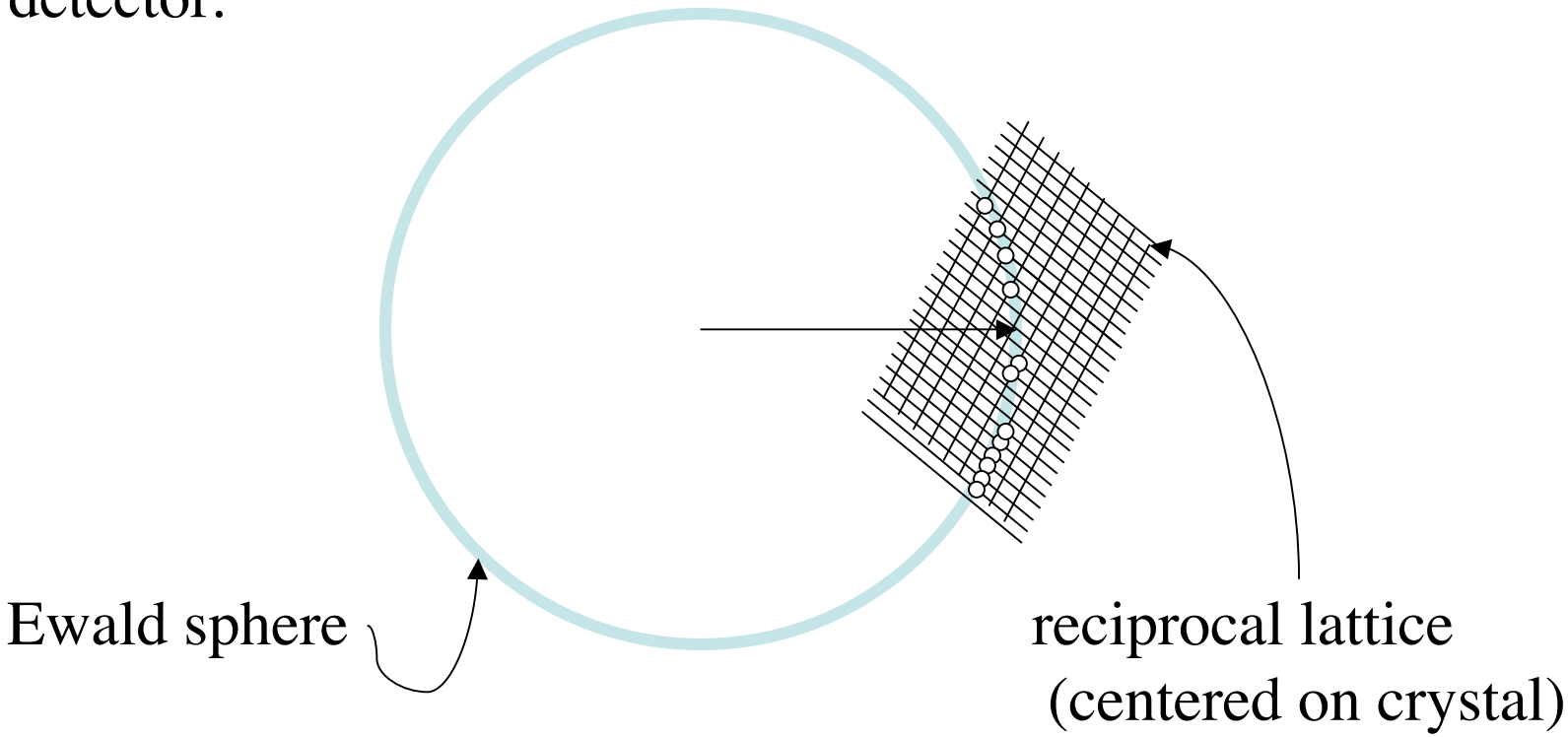
$2\theta = 170°$

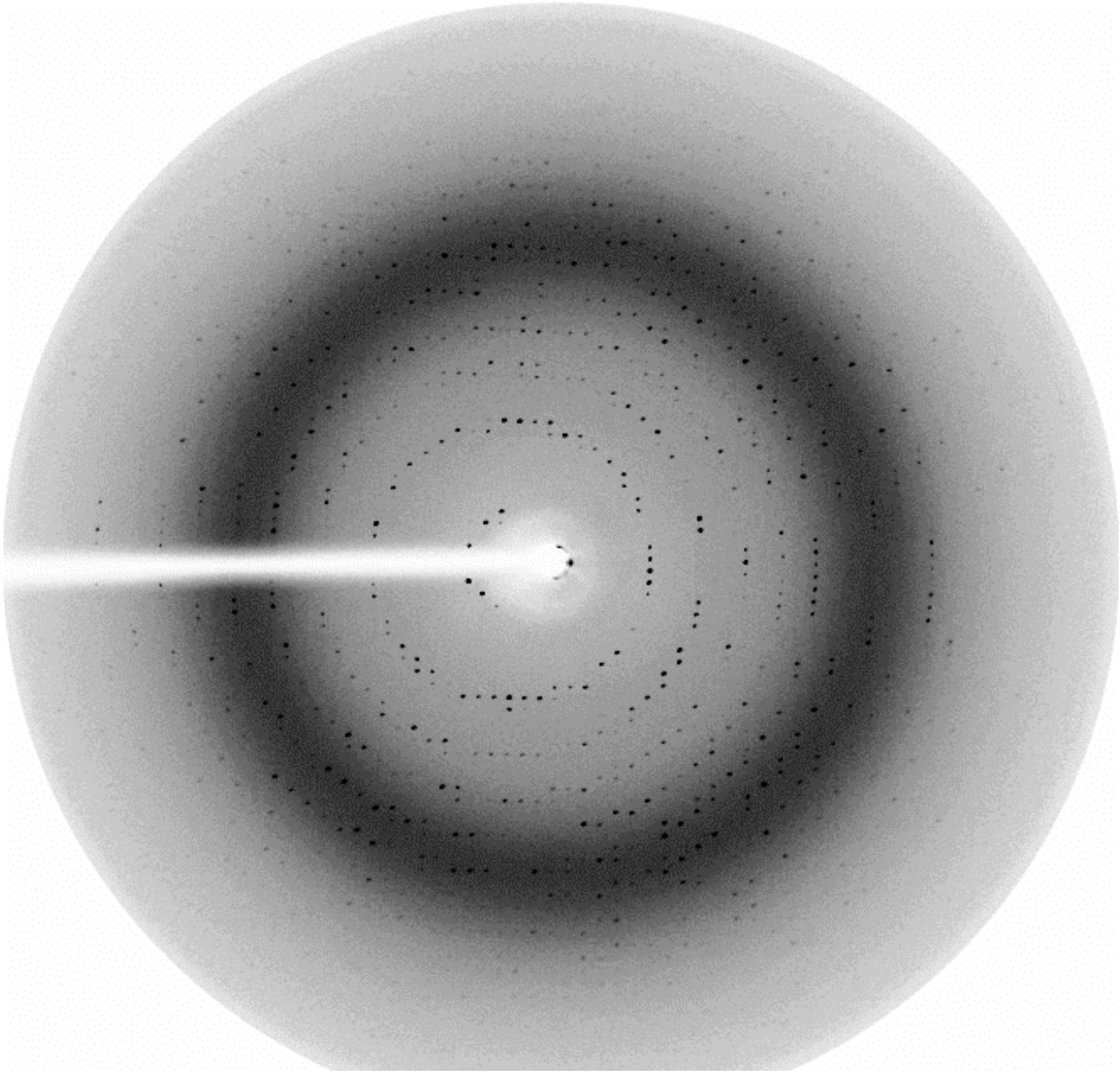
In 3D, a spherical shell of space is visible.

Radius of the sphere is $1/\lambda$.

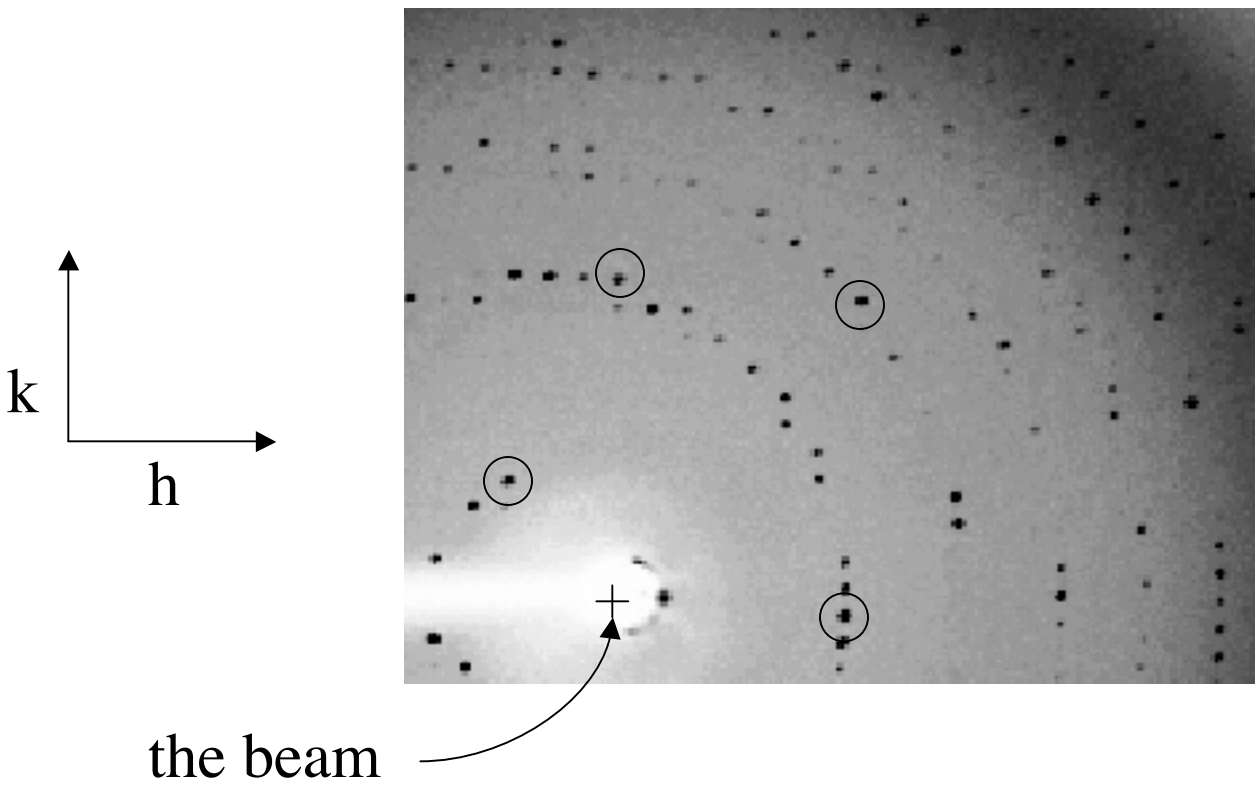
Ewald sphere intersecting the reciprocal lattice

For a given orientation of the crystal with respect to the X-rays, only those value of S that are on the Ewald sphere are visible to the detector.

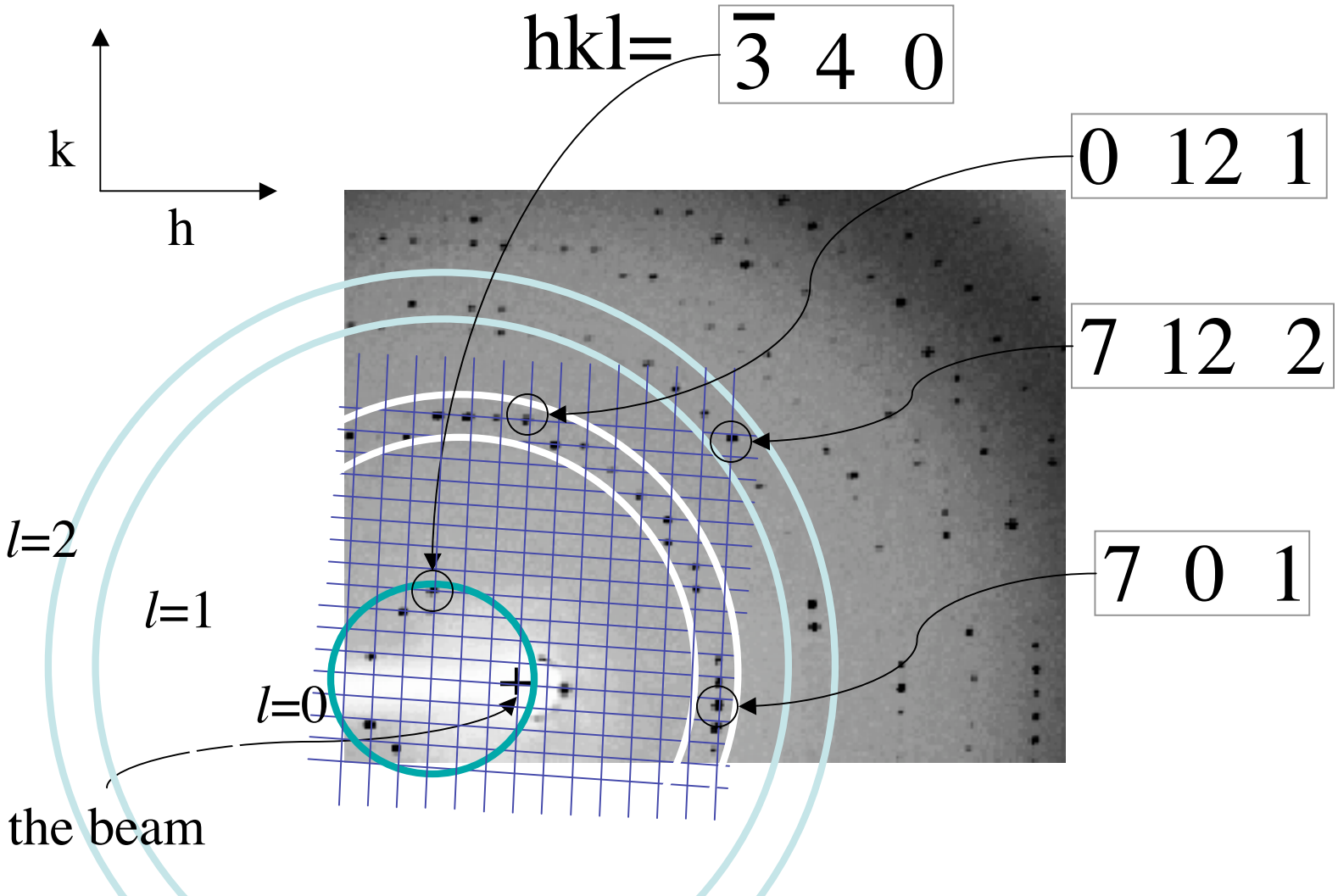




In class exercise: index these spots



answer



Phasing, maps

To get ρ, invert the FT

forward transform -- electron density --> reflections

$$F(hkl) = \sum_{r} \rho(r) e^{i2\pi(hx+ky+lz)}$$

reverse transform -- reflections --> electron density

$$\rho(r) = \sum_{hkl} F(hkl) e^{-i2\pi(hx+ky+lz)}$$

The Phase Problem

Oh no. We can't measure phases!

X-ray detectors (film, photomultiplier tubes, CCDs, etc) **can measure only the intensity of the X-rays**

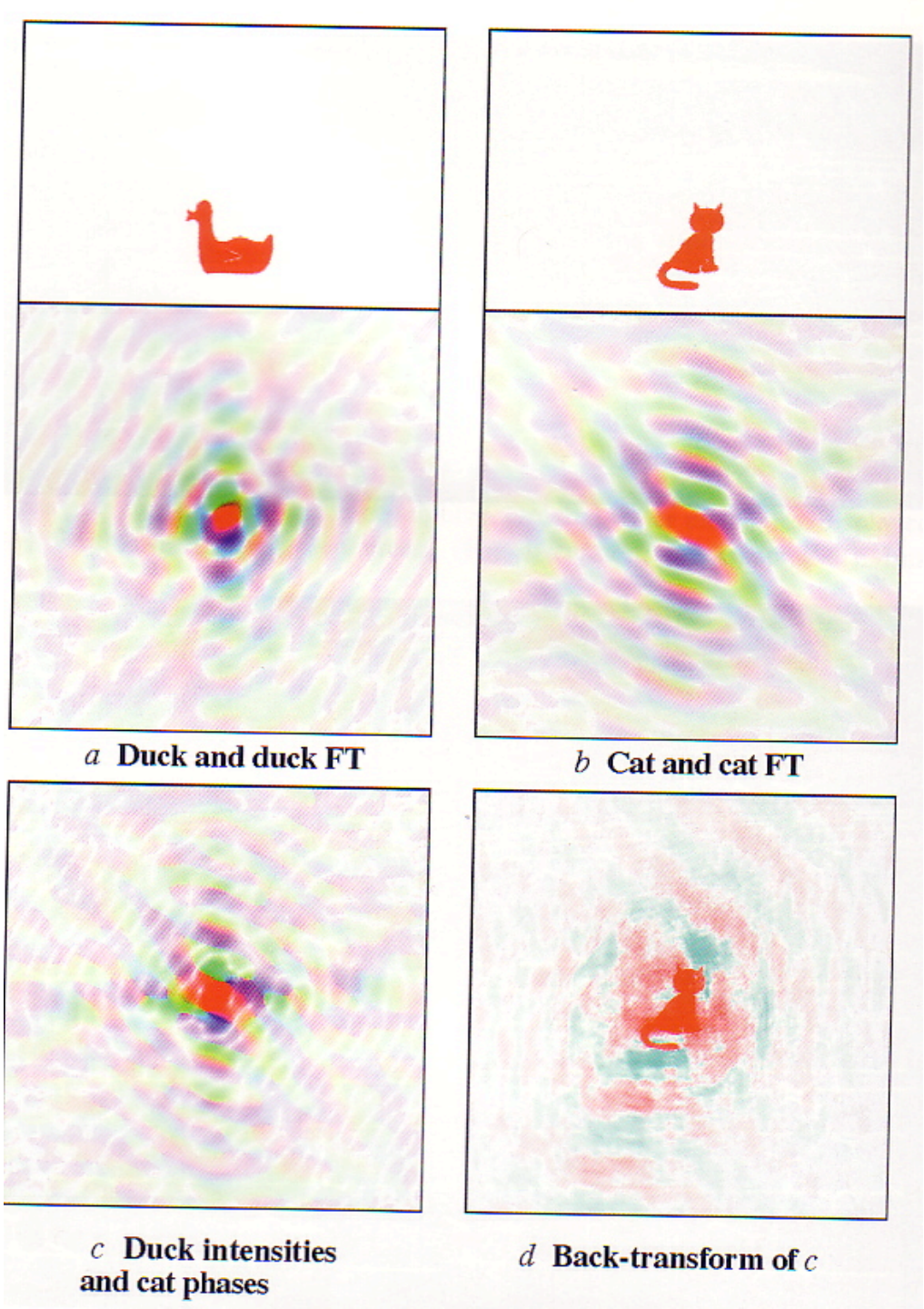
$$F(hkl) = |F(hkl)|e^{i\alpha(hkl)}$$

We can only measure this part

Phase is more important than amplitude

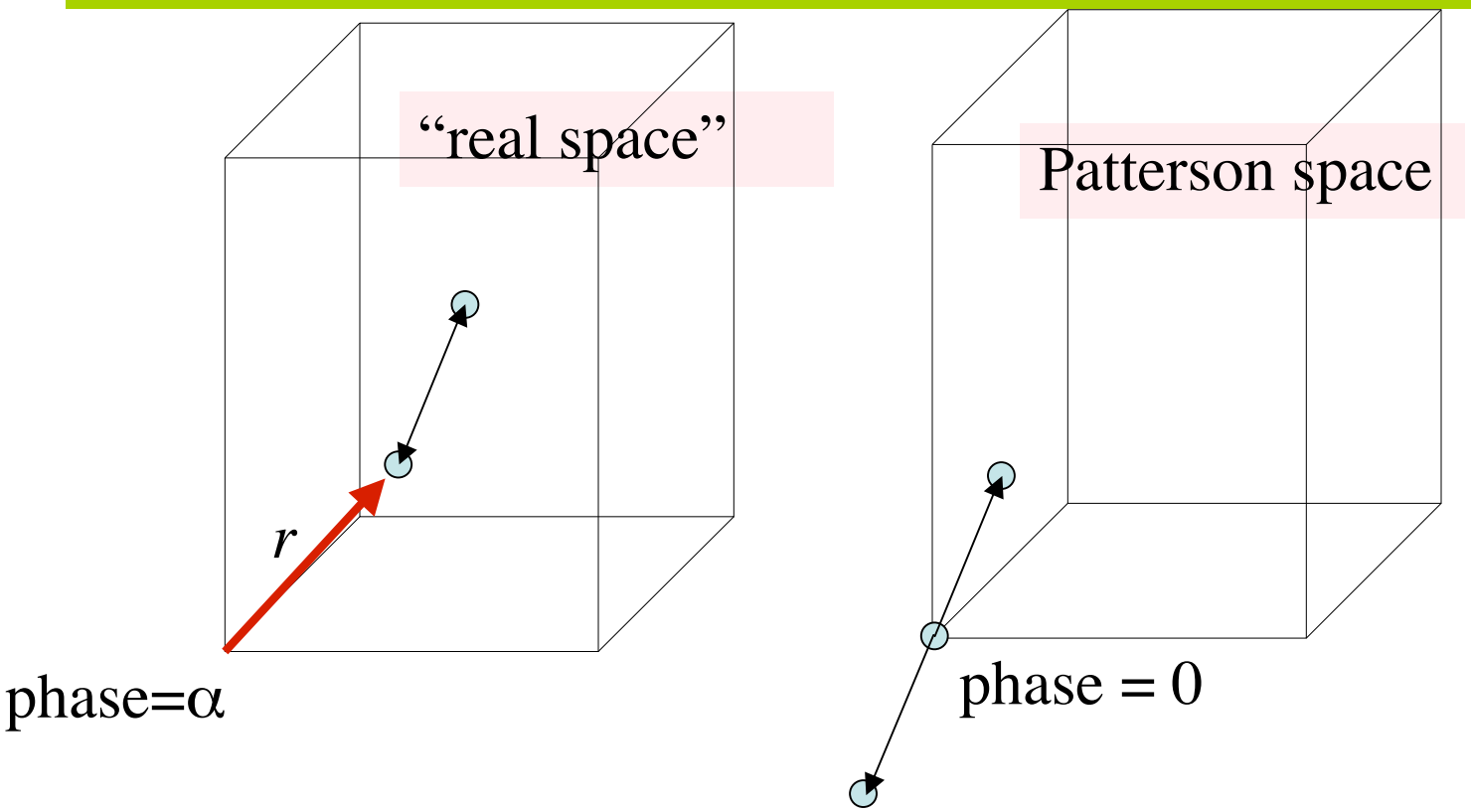
hue=phase angle

brightness=amplitude



What if we just invert the FT with all phases = zero ?

It would be like moving every atom to the origin,
every pair of atoms to a triplet centered at the origin.

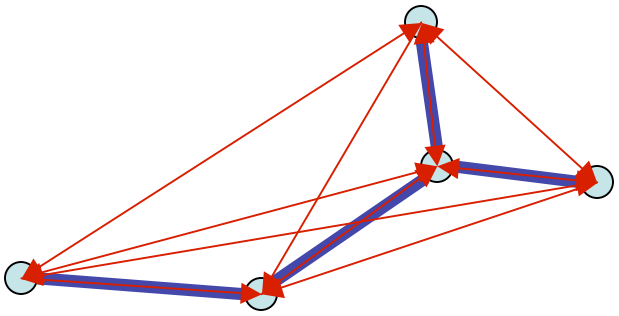


Setting all phases to 0 creates a Patterson Map of the molecule.

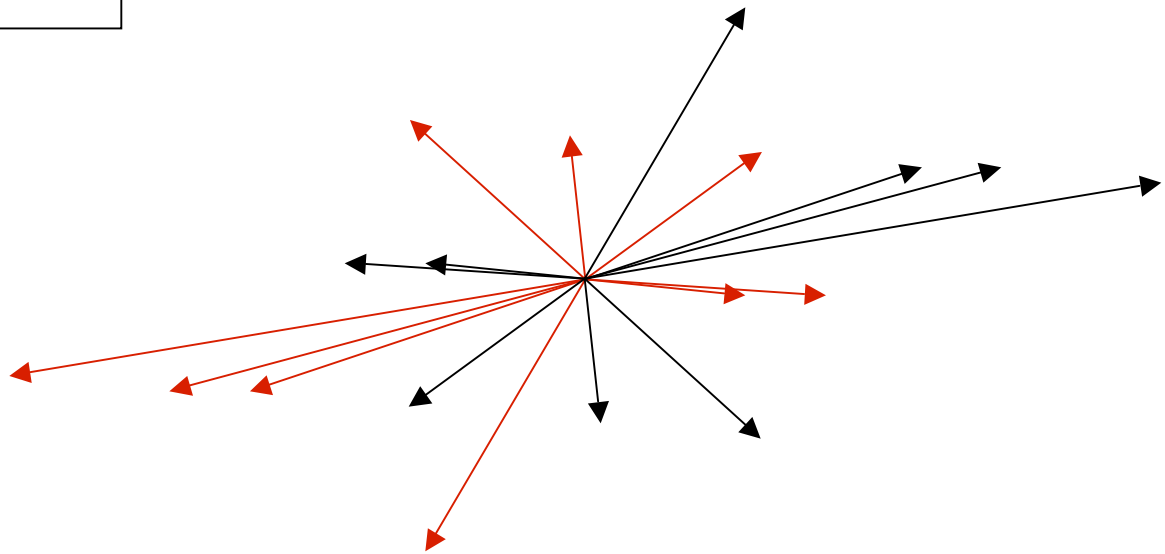
Patterson map represents all inter-atomic vectors

To generate a fake Patterson map in 2D, draw all inter-atomic vectors, then move the tails to the origin. The heads are where peaks would be.

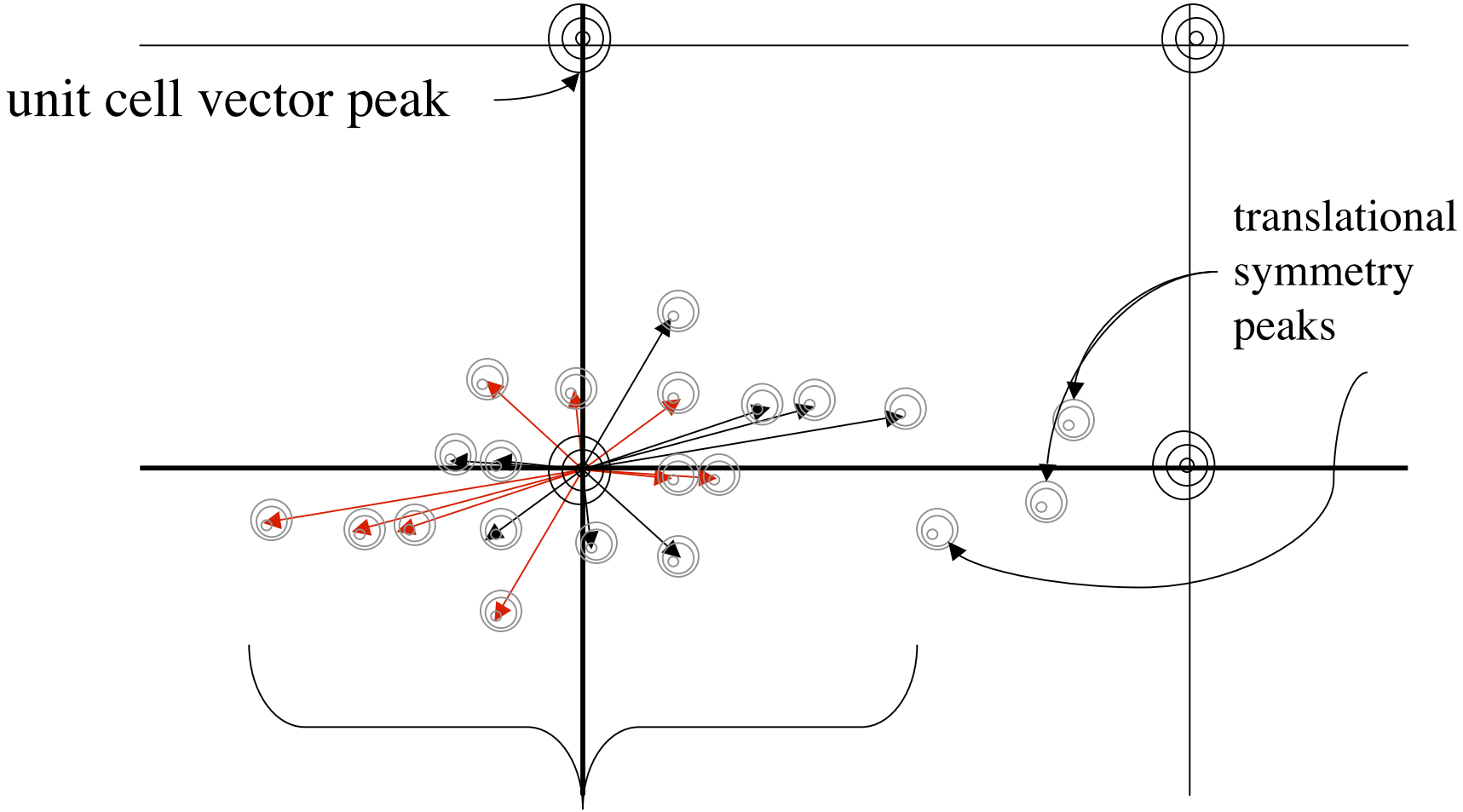
For example, take glycine, 5 atoms (not counting H's)



Move each vector to the origin

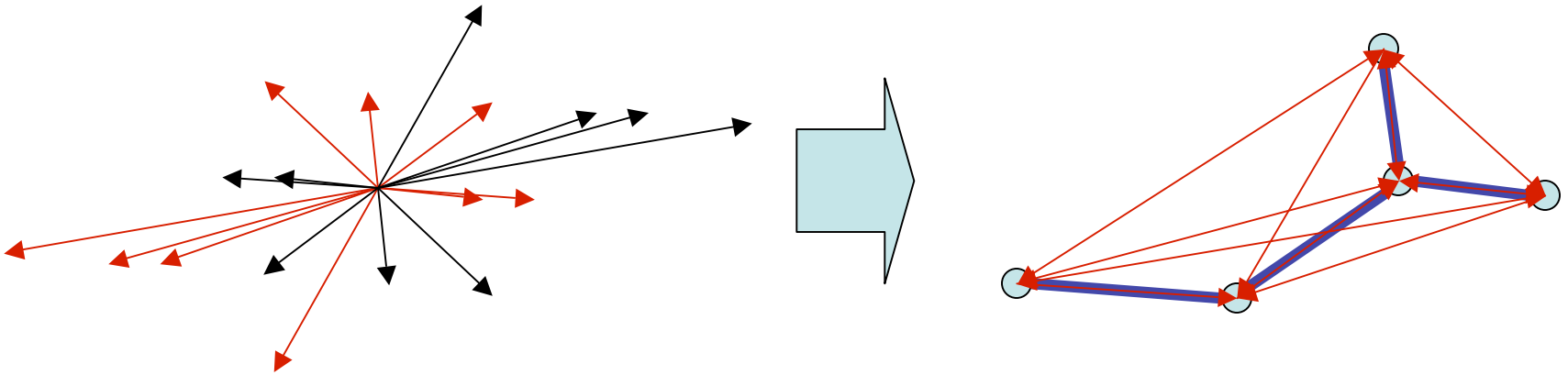


Patteron map for Gly



Can you reassemble glycine from this?

For small molecules, vector/geometry problem can be solved



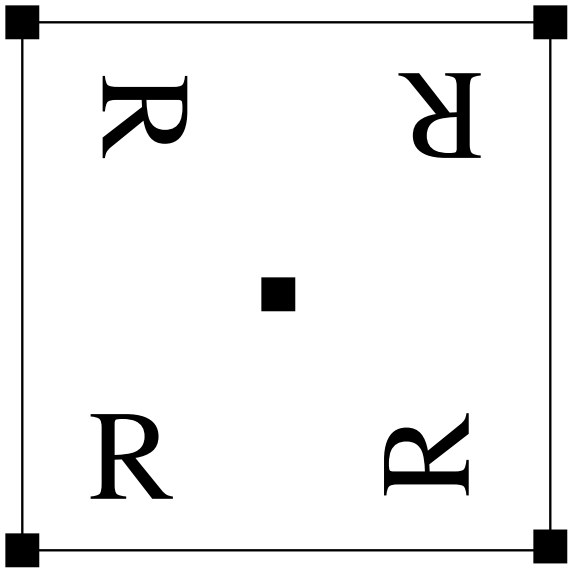
...if you know the geometry (bond lengths, angles) of the molecule

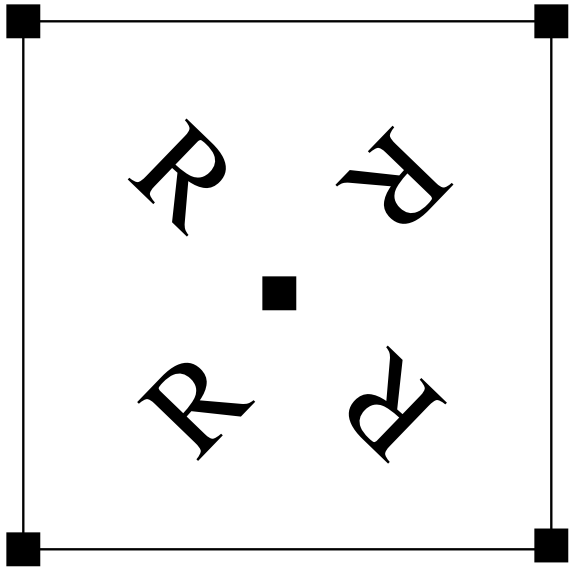
http://www.cryst.bbk.ac.uk/xtal/mir/patt2.htm

molecular replacement

If the structure of the molecule is known approximately, then the phases can be calculated.

BUT. We need to know how the molecule is oriented.





The diffraction patterns of these two crystals are not the same.

6-dimensional search space

Every possible rigidbody transformation of a molecule can be described using 6 parameters. 3 angles of rotation (defining a matrix of 9 coefficients), and a vector of translation (3 values). i.e.

$$x' = c_{11}x + c_{21}y + c_{31}z + v_x$$

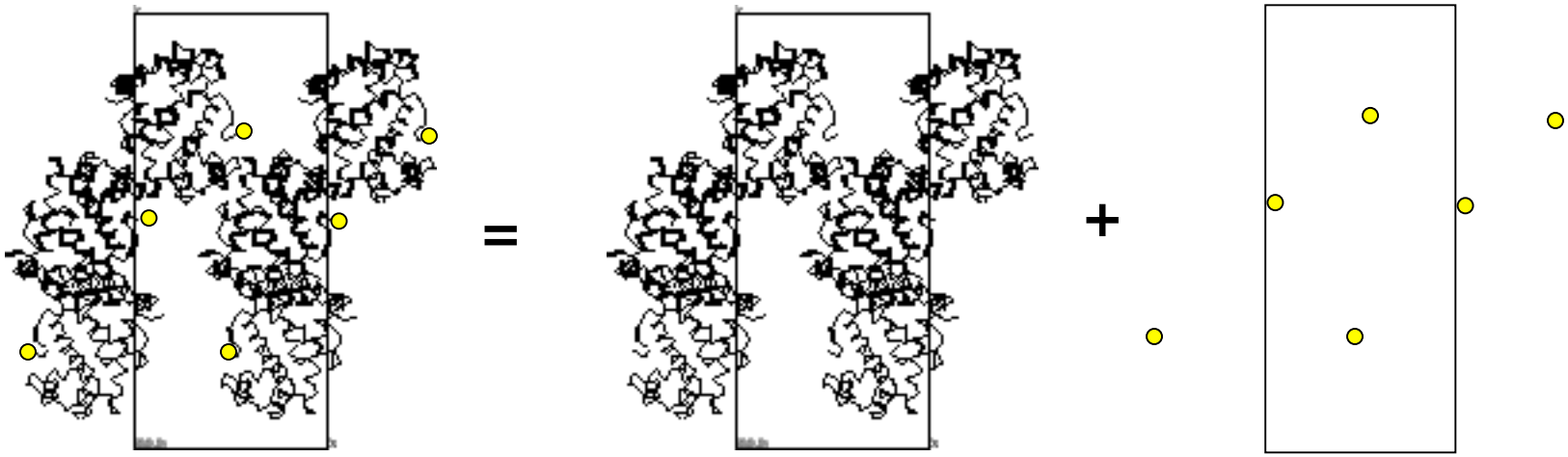
$$y' = c_{12}x + c_{22}y + c_{32}z + v_y$$

$$z' = c_{13}x + c_{23}y + c_{33}z + v_z$$

Therefore, the position of our molecule in the crystal unit cell must be a 6D transformation of its current position. Molecular replacement is the method for finding the angles and vector that define the transformation.

Multiple isomorphous replacement

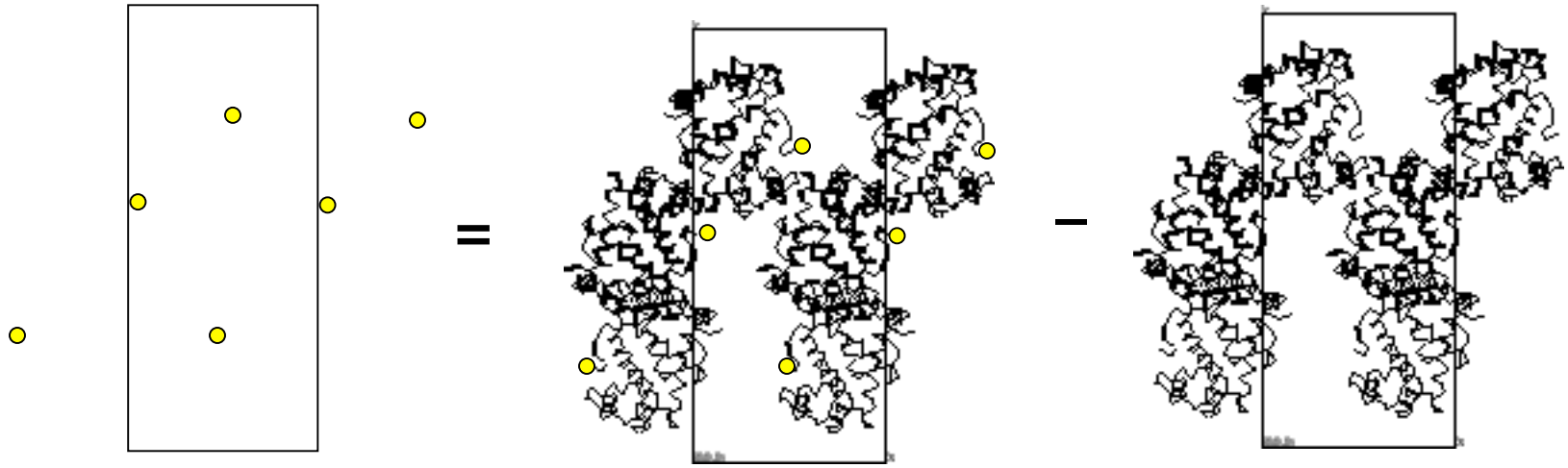
=Turning proteins into small molecules by soaking in heavy atoms



The Fourier transform (i.e. diffraction pattern) of a heavy atom derivitive is the vector sum of the transforms of the protein and the heavy atoms.

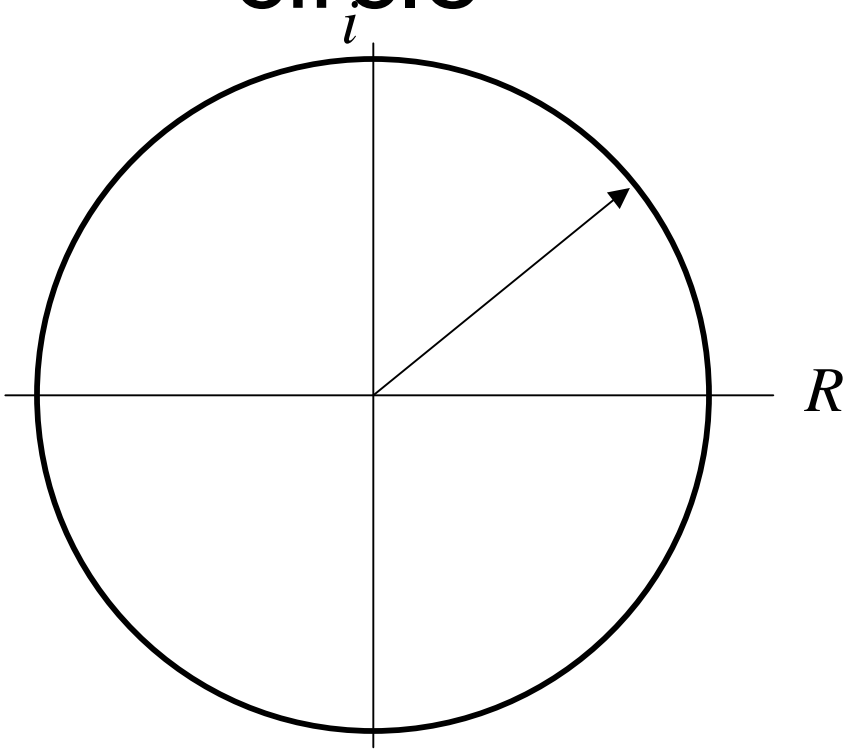
NOTE: protein and protein-heavyatom crystals must be *isomorphous*.

Subtracting Fourier transforms



The Fourier transform of the heavy atoms is the vector difference of F_{PH} and F_H.

We can represent a structure factor of *unknown phase* as a circle



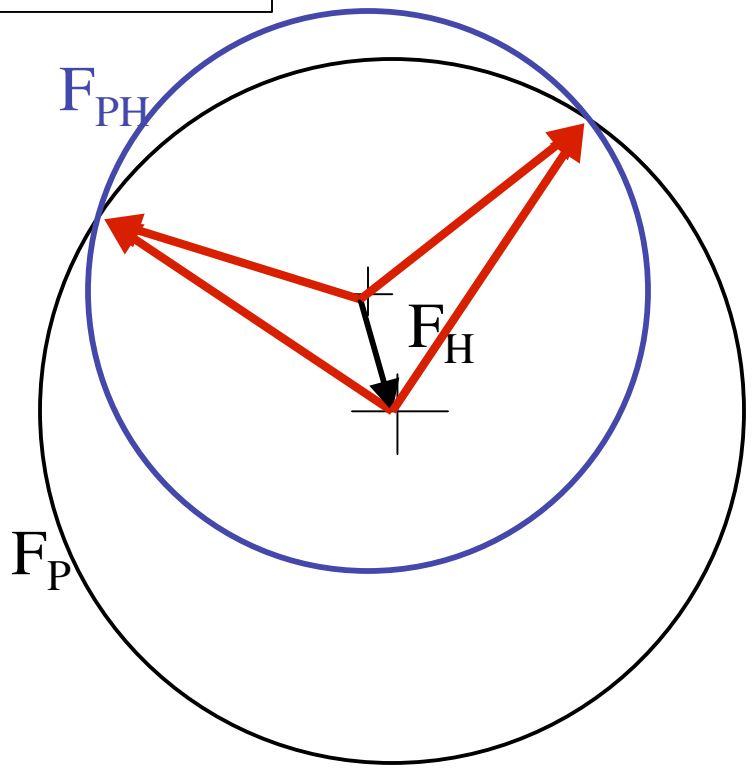
Radius of the circle is the amplitude. The true F lies somewhere on the circle.

Summing structure factors geometrically

We know only amplitude

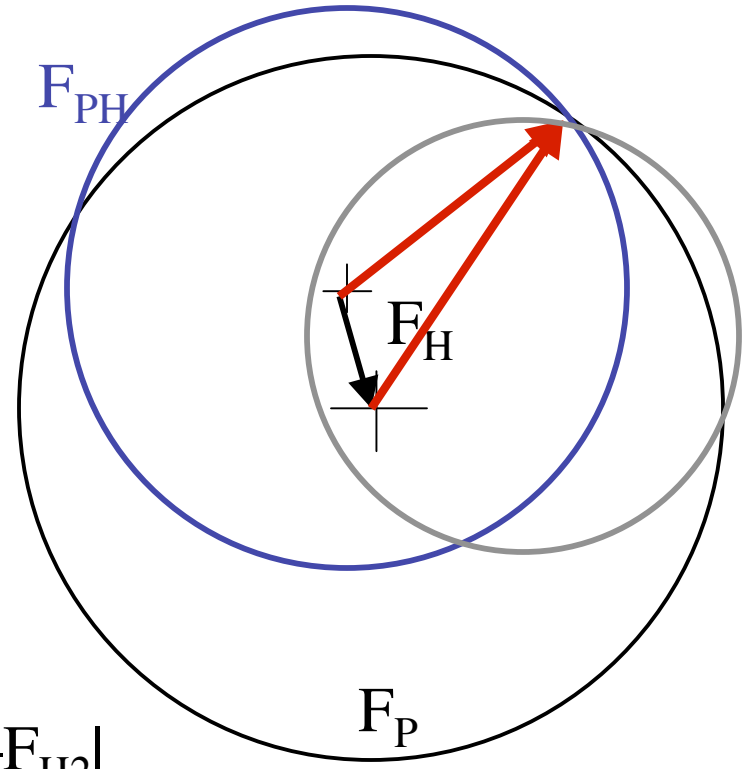
We know amplitude and phase

$$|F_P + F_H| = |F_{PH}|$$



There are two ways to make the vector sums add up.

Two heavy atom derivatives are enough, but more is better



$|F_P + F_{H1}| = |F_{PH1}|$

$|F_P + F_{H2}| = |F_{PH2}|$

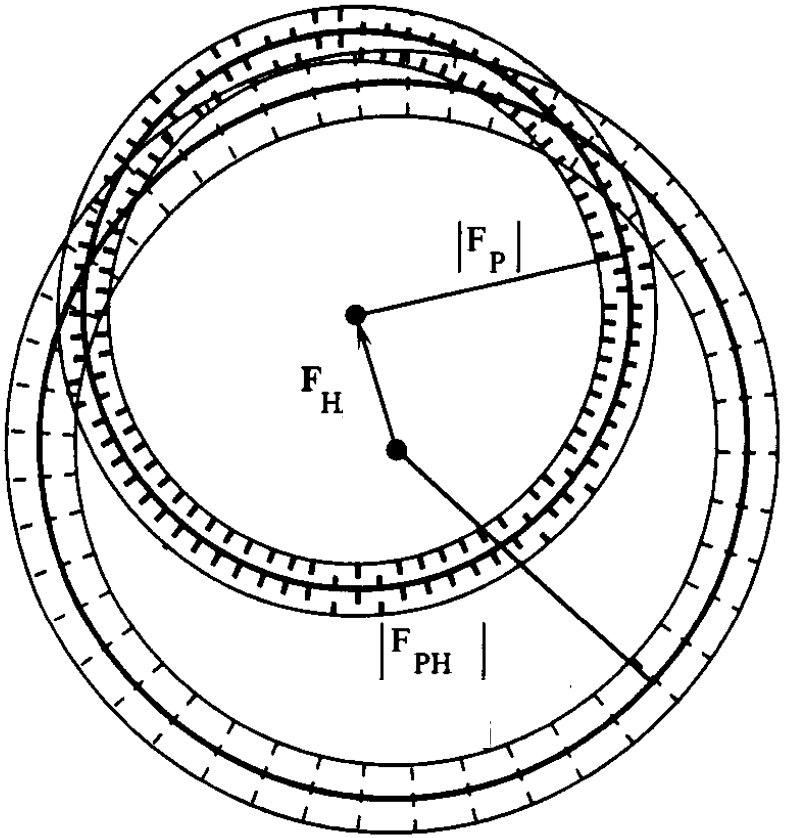
Or

$|F_P| = |F_{PH1} - F_{H1}| = |F_{PH2} - F_{H2}|$

Initial phases

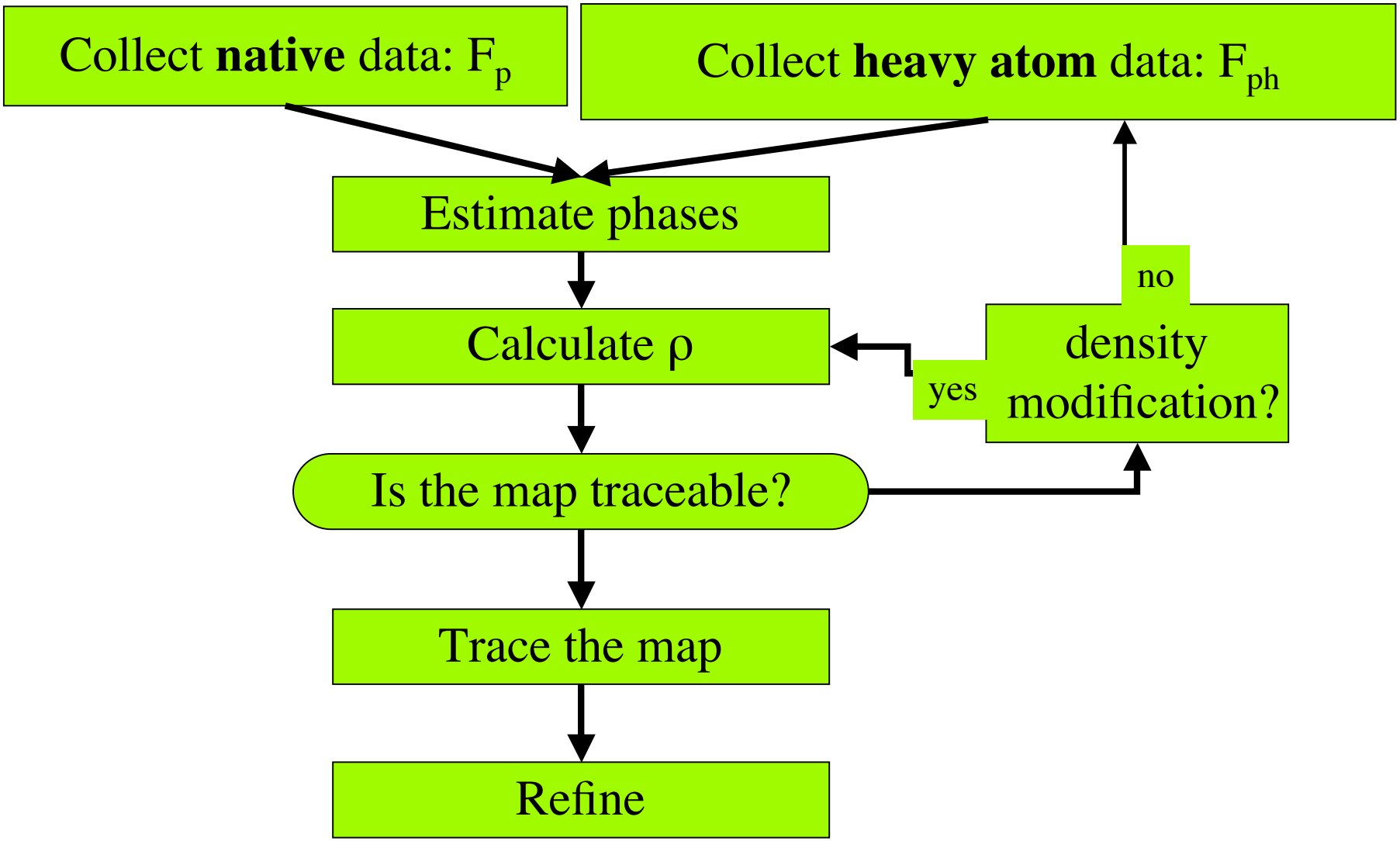
Phases are not measured exactly because amplitudes are not measured exactly.

Error bars on F_P and F_{PH} create a distribution of possible phase values α.

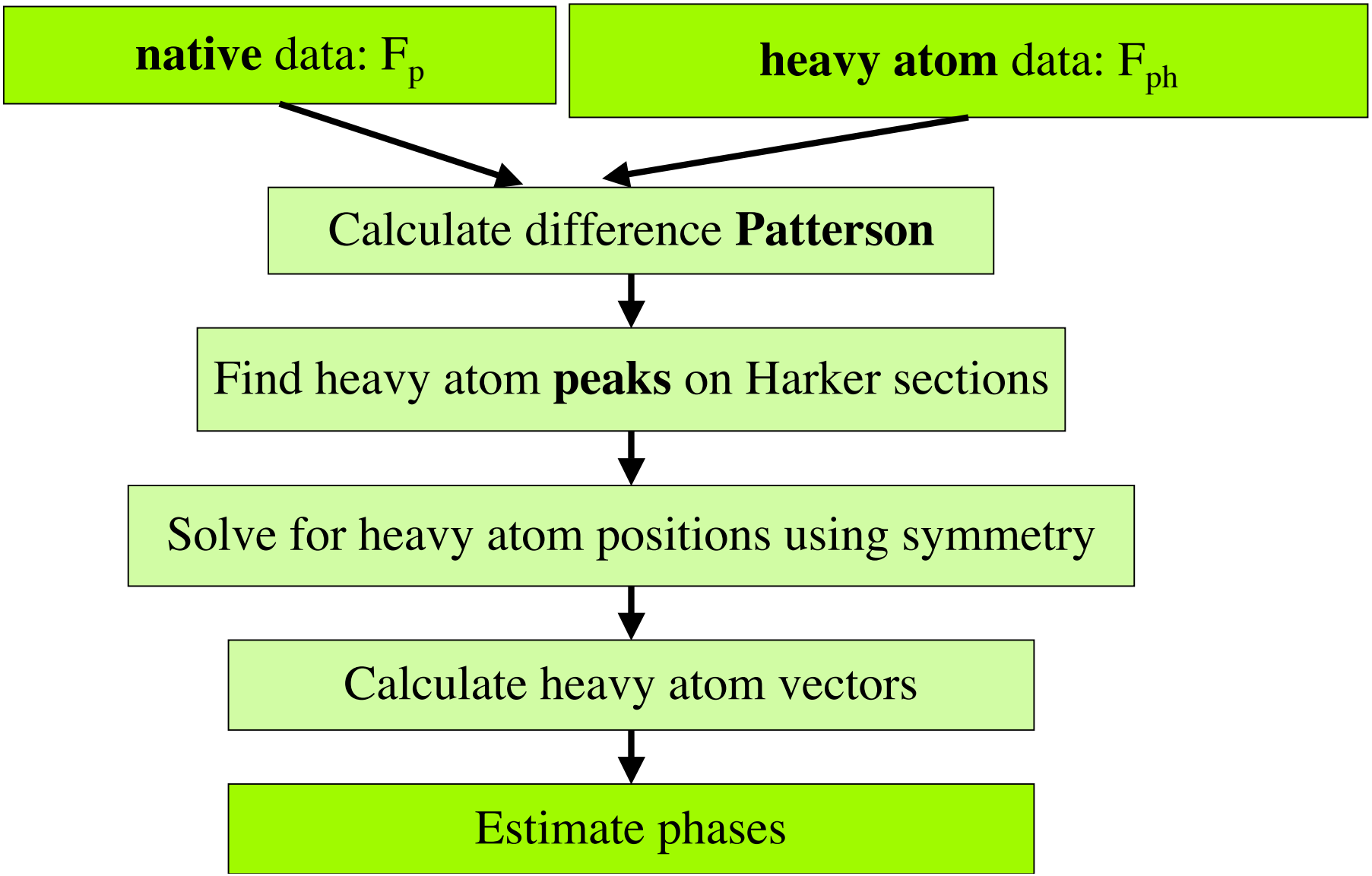


width of circle is 1σ deviation, derived from data collection statistics.

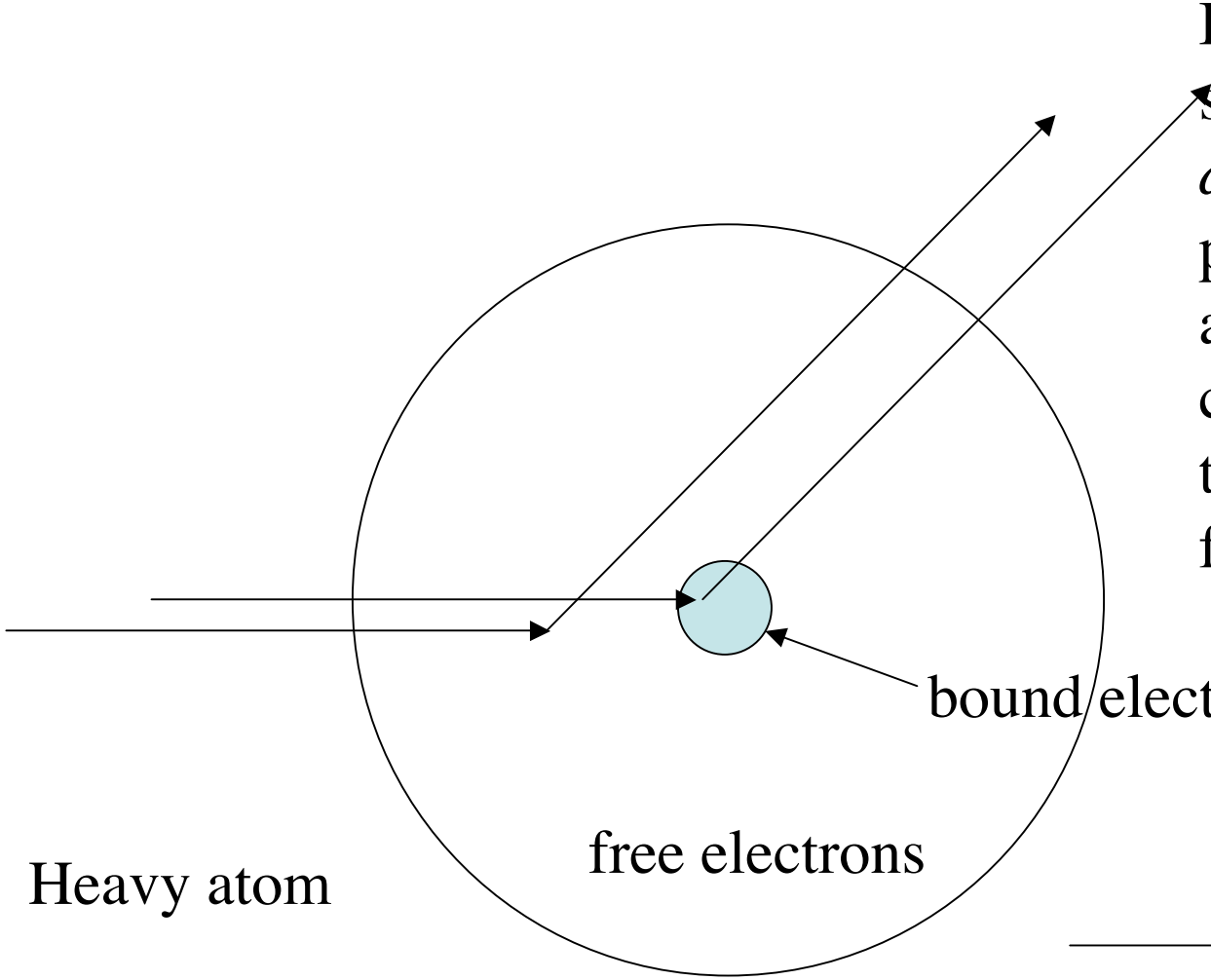
From data to model



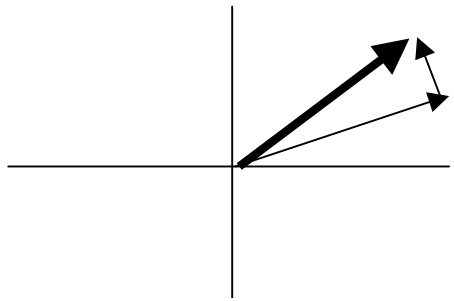
From data to phases



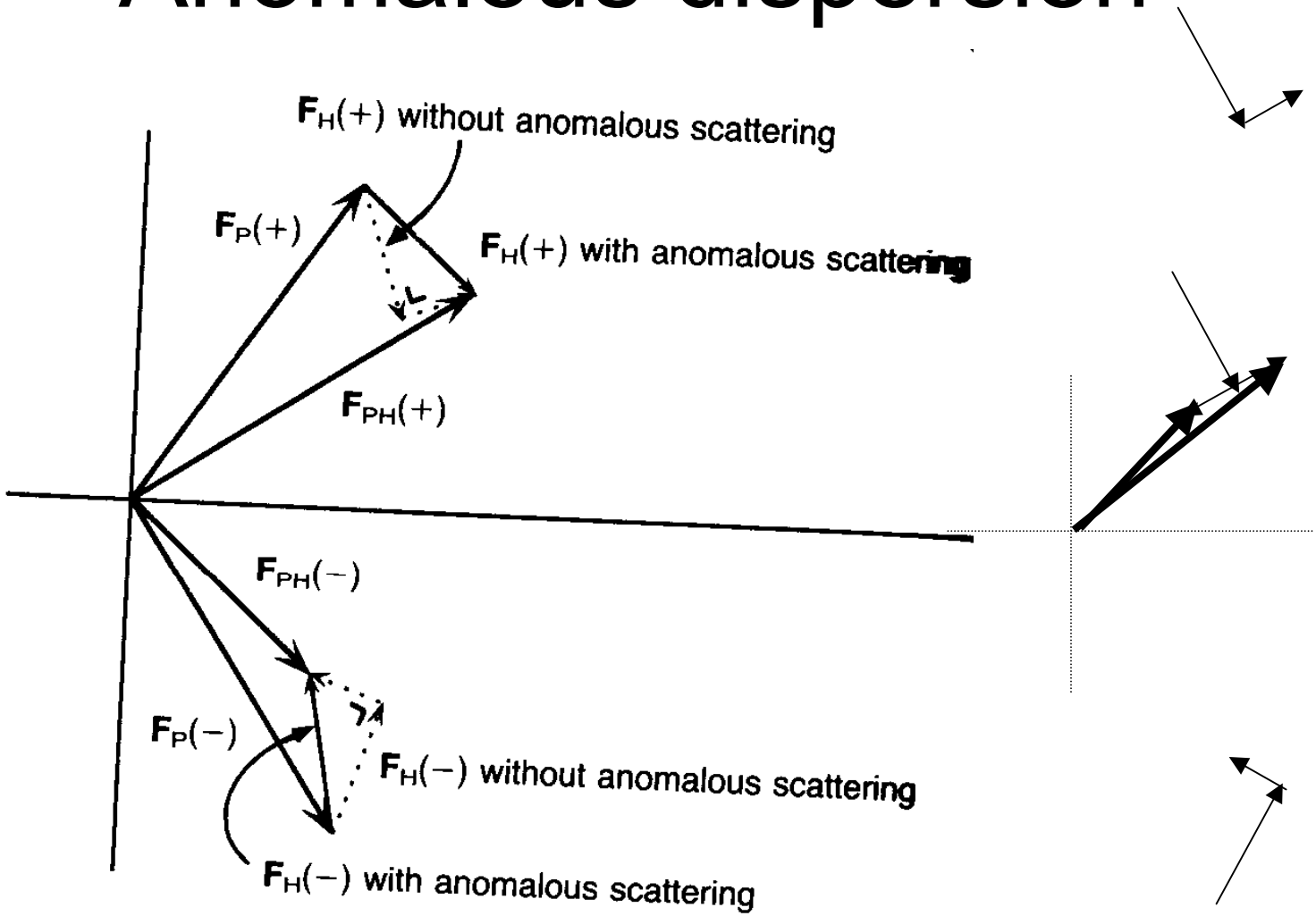
Anomalous dispersion



Inner electrons scatter with a *time delay*. This is a phase shift that is always counter-clockwise relative to the phase of the free electrons.



Anomalous dispersion



Resolution, R-factor

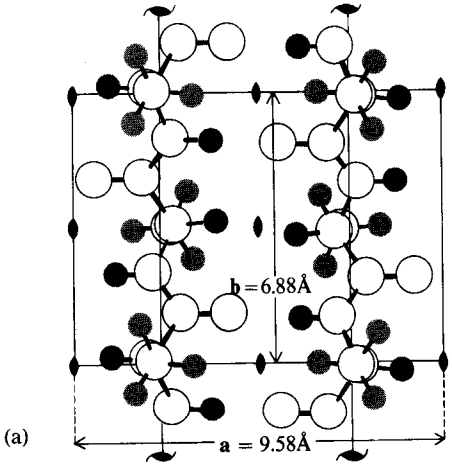
Resolution

low

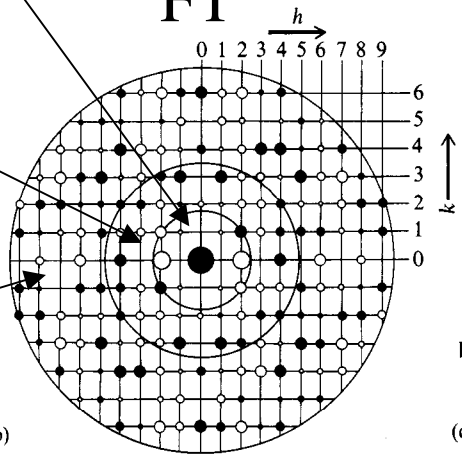
medium

high

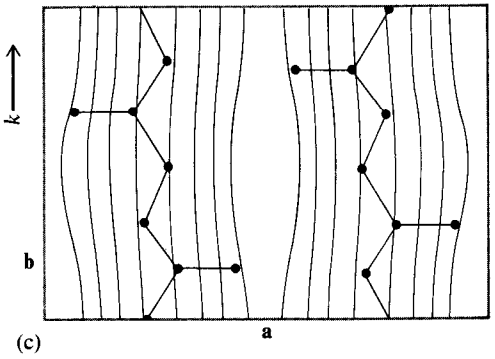
image: 2 beta strands

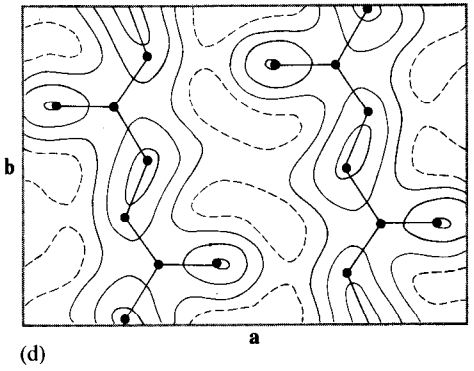


FT

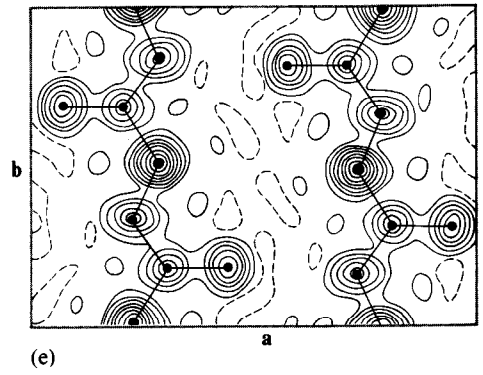


back transform, low



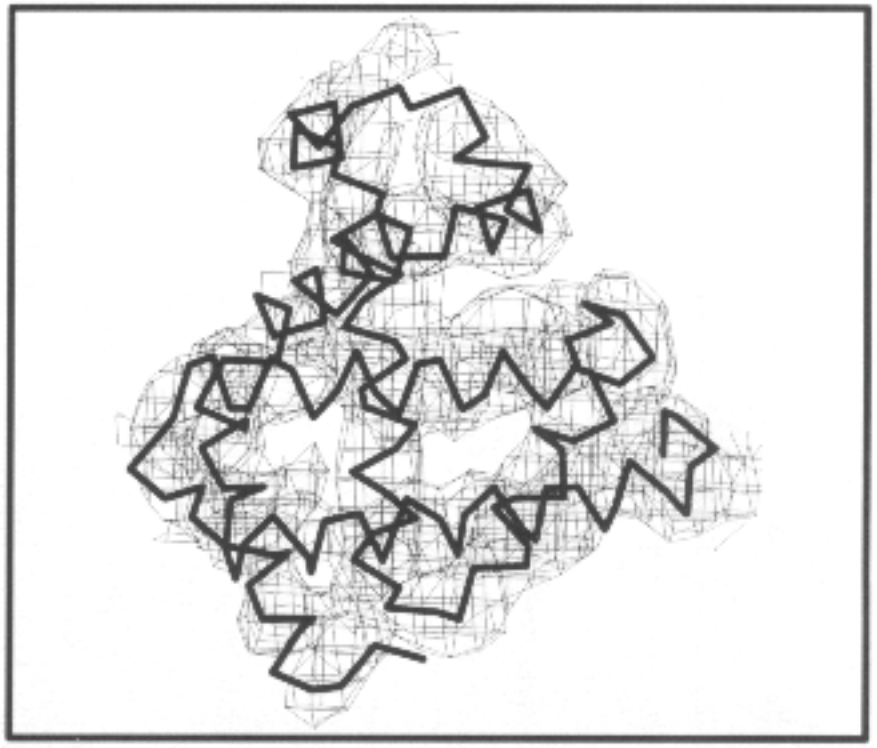


back transform, medium



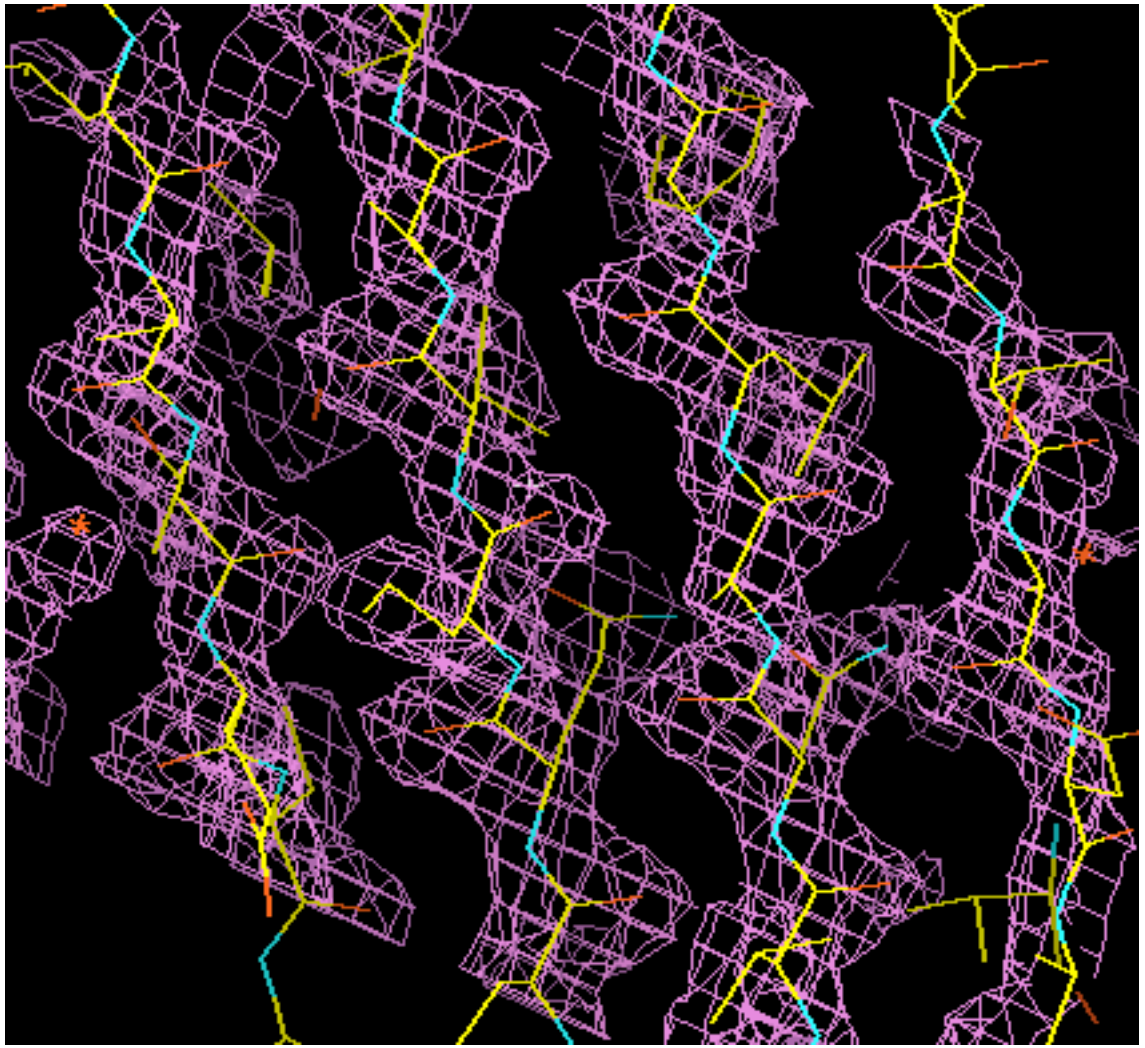
back transform, high

Low-resolution

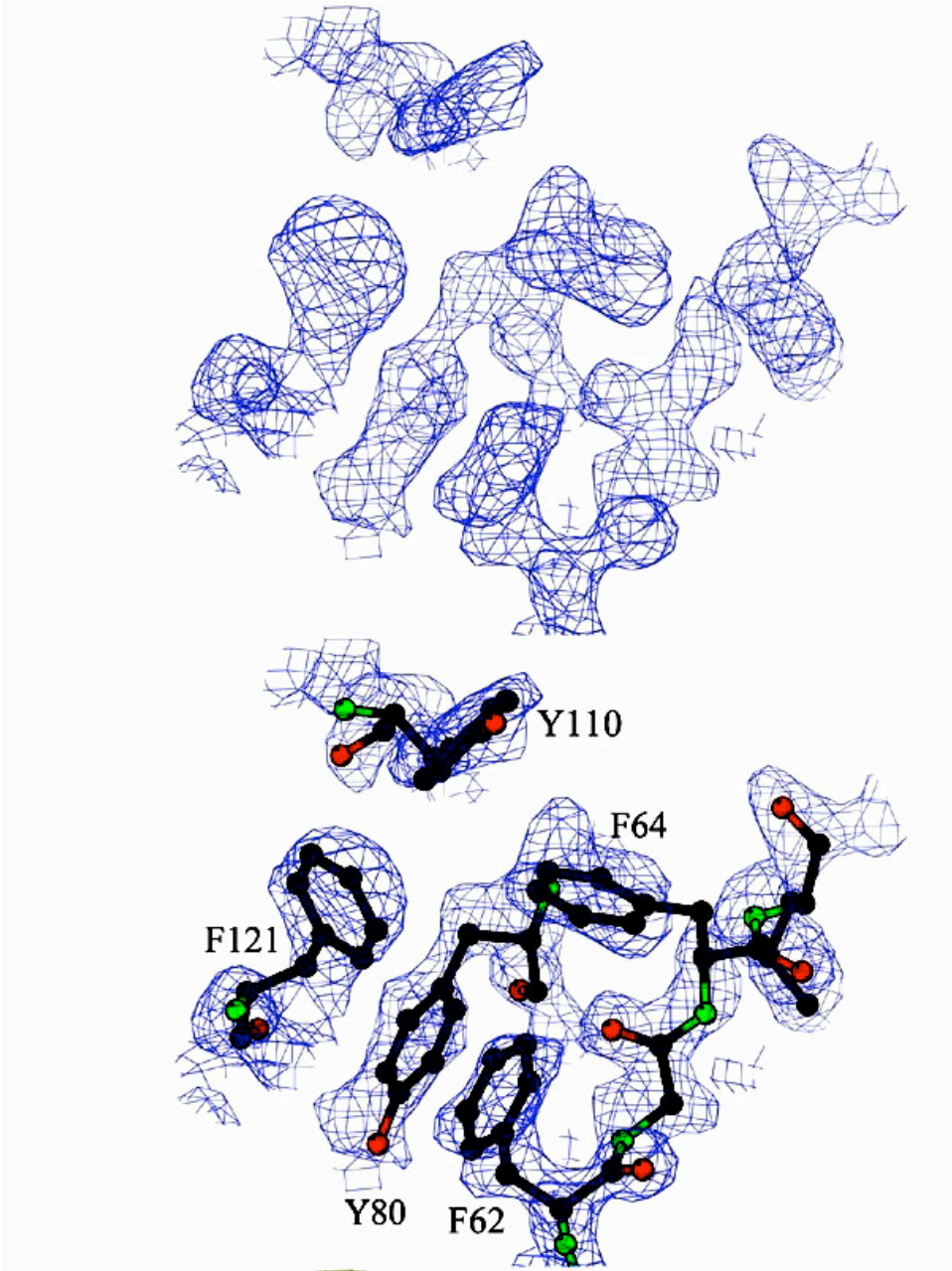


At 4-6Å resolution, alpha helices look like sausages.

Medium resolution

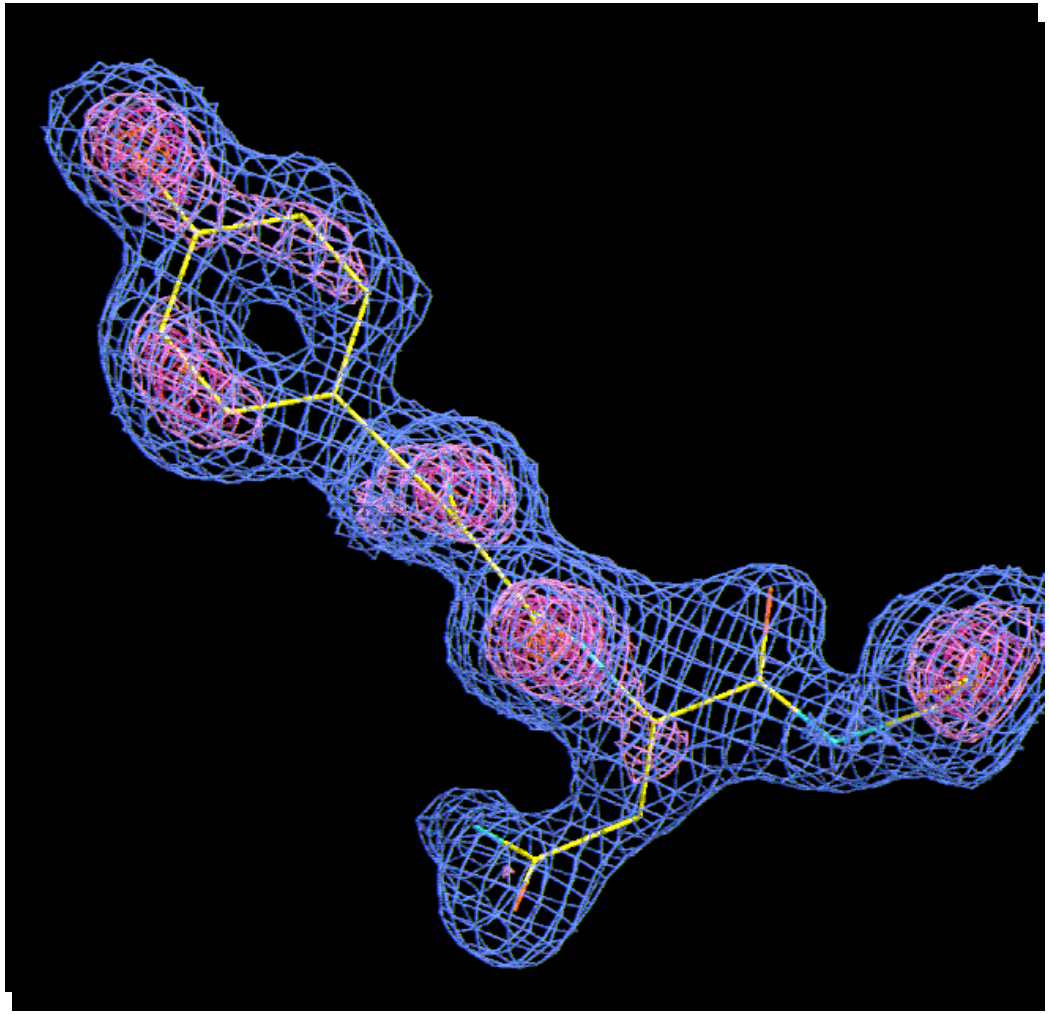


~3Å data is good enough to se the backbone with space inbetween.



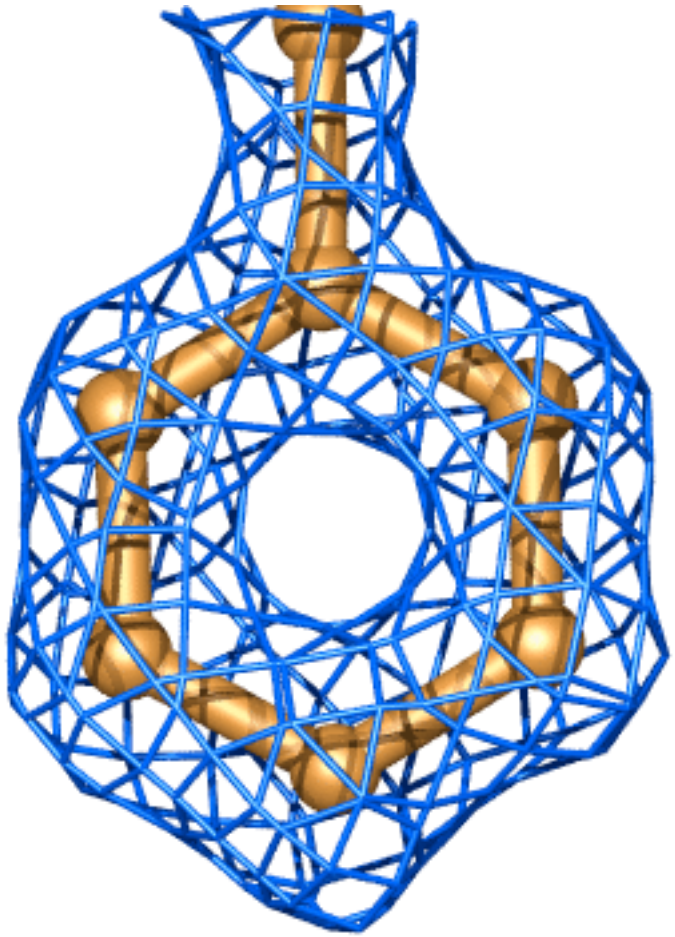
Errors in the phases make some connections ambiguous.

Holes in rings are a good thing

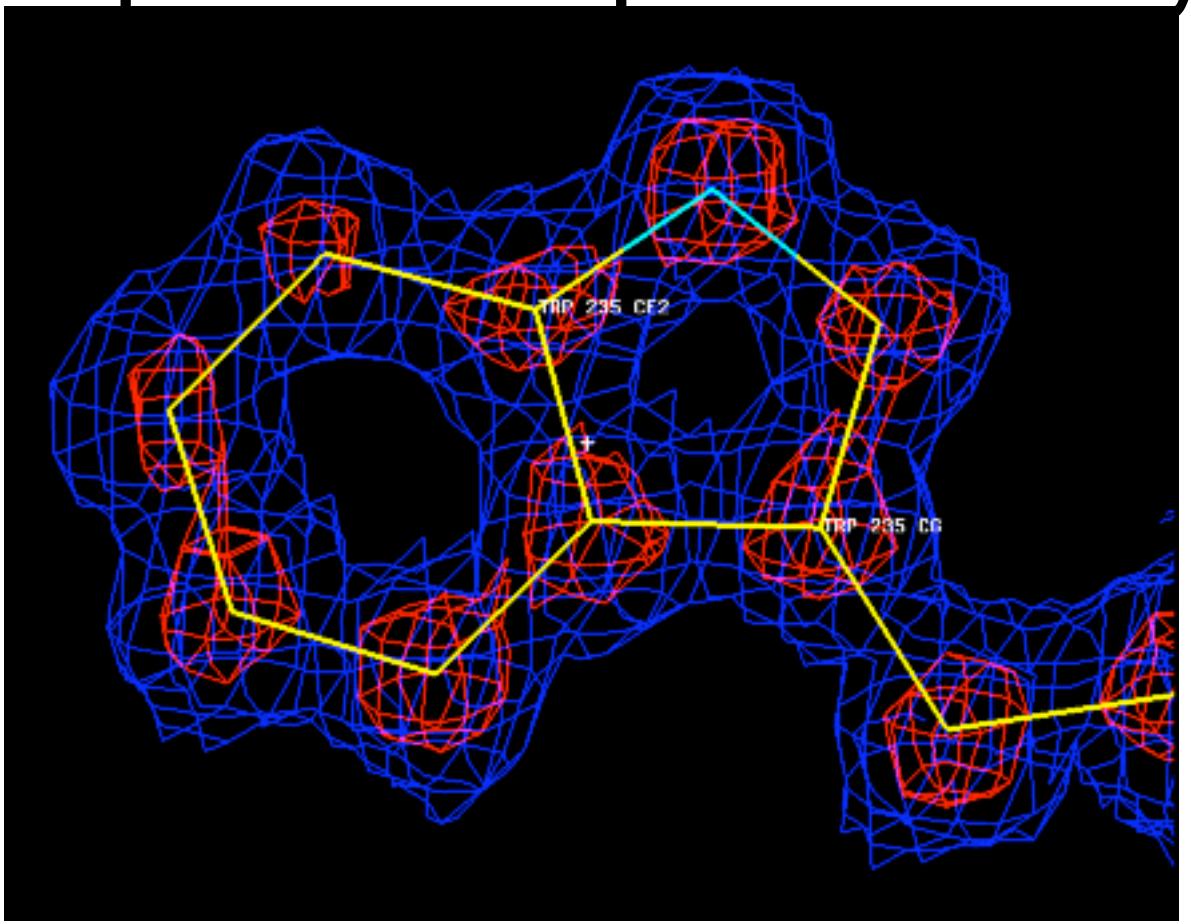


Seeing a hole in a tyrosine or phenylalanine ring is universally accepted as proof of good phases. You need at least 2Å data.

Great map: holes in rings

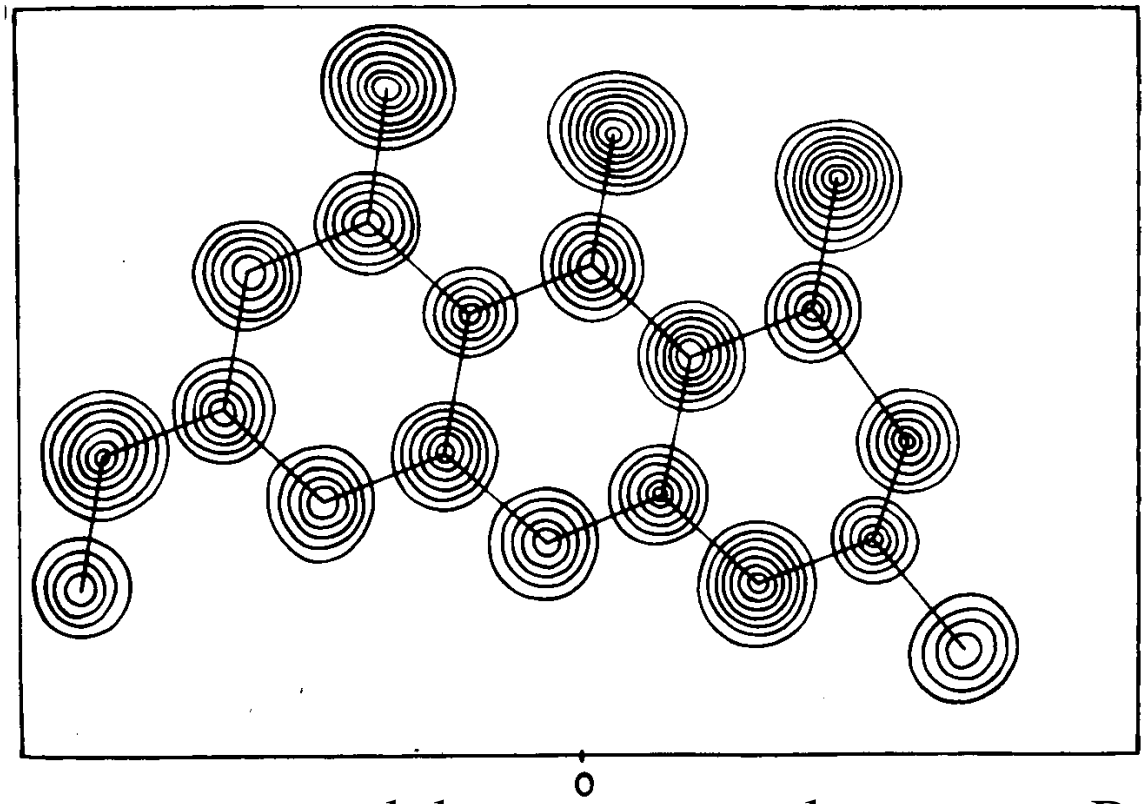


Superior map: Atomicity



Rarely is the data this good. 2 holes in Trp. All atoms separated.

Only small molecule structures are this good



Atoms are separated down to several contours. Proteins are never this well-ordered. But this is what the density really looks like.

The “R-factor”

The R-factor is the residual between measured and calculate amplitudes.

The “free R-factor” is calculated the same as the R-factor, but on a reserved test set that was not used to fot the model.

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

Note: *k* is a scaling factor.

Refinement

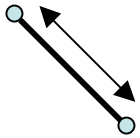
- The *gradient* of the R-factor with respect to each atomic position may be calculated.
- Each atom is moved down-hill along the gradient.
- "Restraints" may be imposed.

$$\frac{dR_{factor}}{d\vec{v}_i}$$

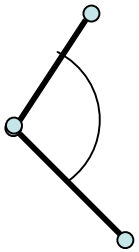
Restraints enforce good stereochemistry

A restraint is a function of the coordinates that is lowest when the coordinates are “ideal”, and which increases as the coordinates become less ideal..

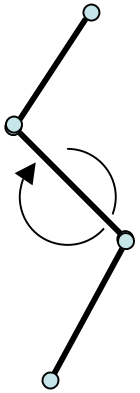
Stereochemical restraints



bond lengths



bond angles



torsion angles

also...

planar groups

temprature factors

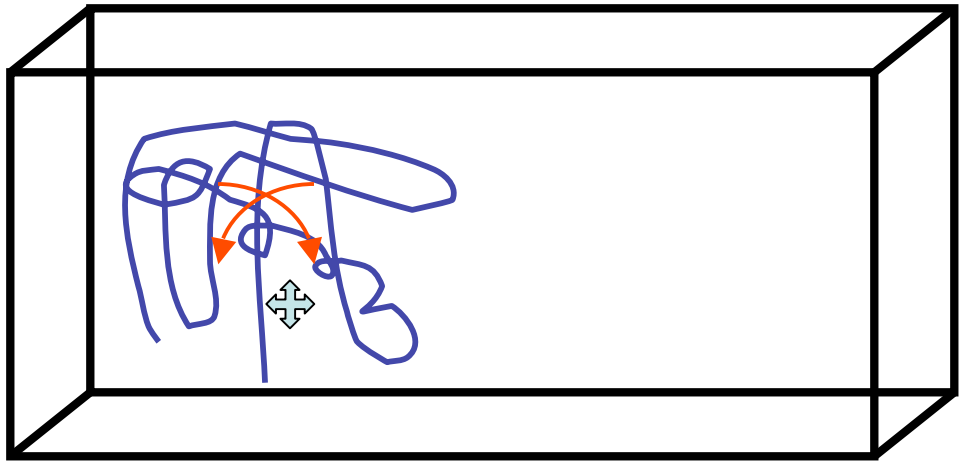
Rigid body refinement

(1) Rigid body refinement.

After molecular replacement only. Whole molecule treated as a rigid group. After it is moved, new phases are calculated, α_{calc}.

$$\rho(\vec{r}) = \sum_{h} \left| F_{obs}\left(\vec{h}\right) \right| e^{-i\left(2\pi\left(\vec{h}\bullet\vec{r}\right)+\alpha_{calc}\left(\vec{h}\right)\right)}$$

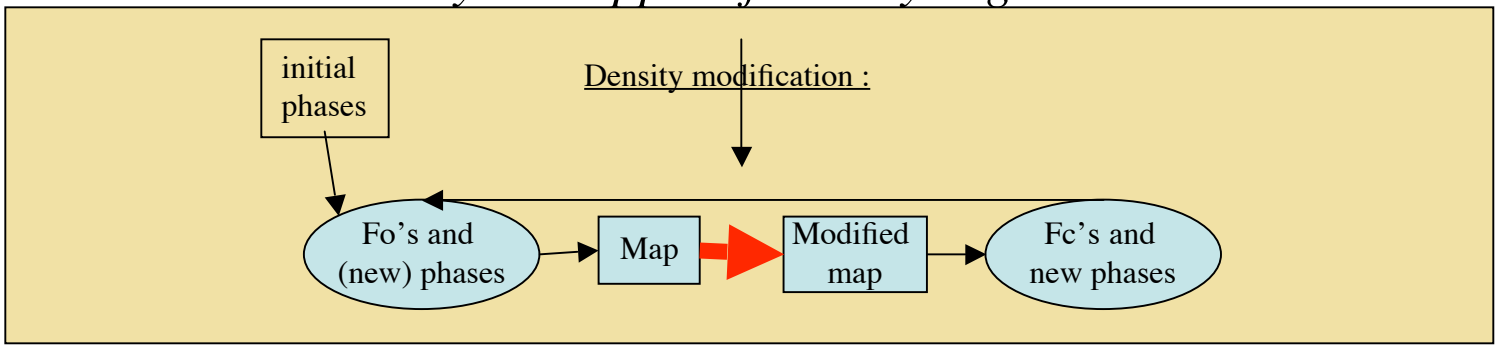
.



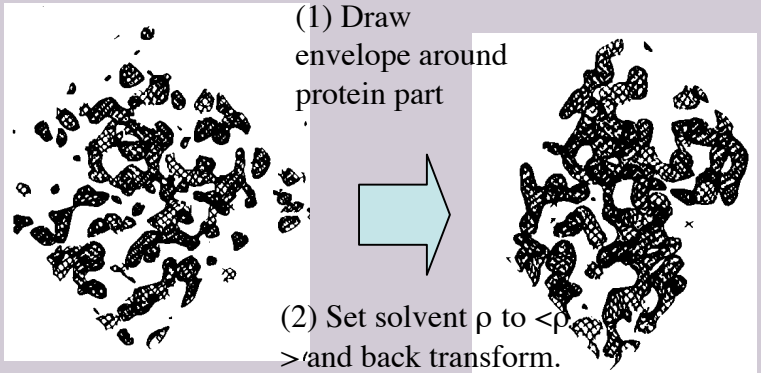
Density modification.

(2) Density modification.

After molecular replacement or heavy-atom isomorphous replacement. Coordinate-free refinement. The map is modified directly, then new phases are calculated. *May be skipped if density is good.*



Solvent Flattening: Make the water part of the map flat.

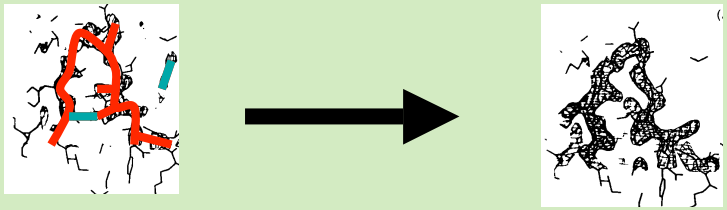


(1) Calculate map.

(2) Skeletonize the map

(3) Make the skeleton "protein-like"

(4) Back transform the skeleton.



Protein-like means: (a) no cycles, (b) no islands

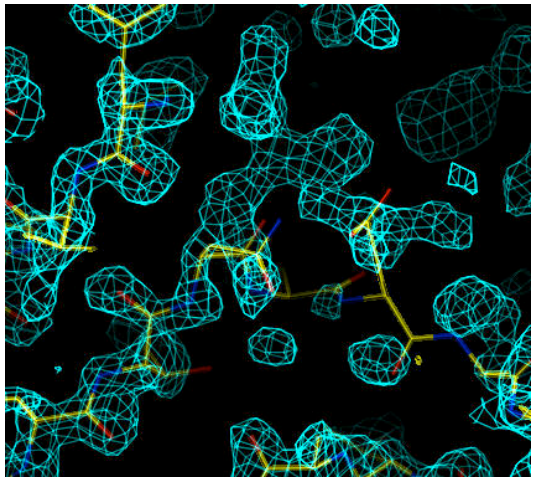
Difference maps

(3) Difference maps are used to find the missing atoms. *e.g.* solvent, substrate, ligand, ions, etc.

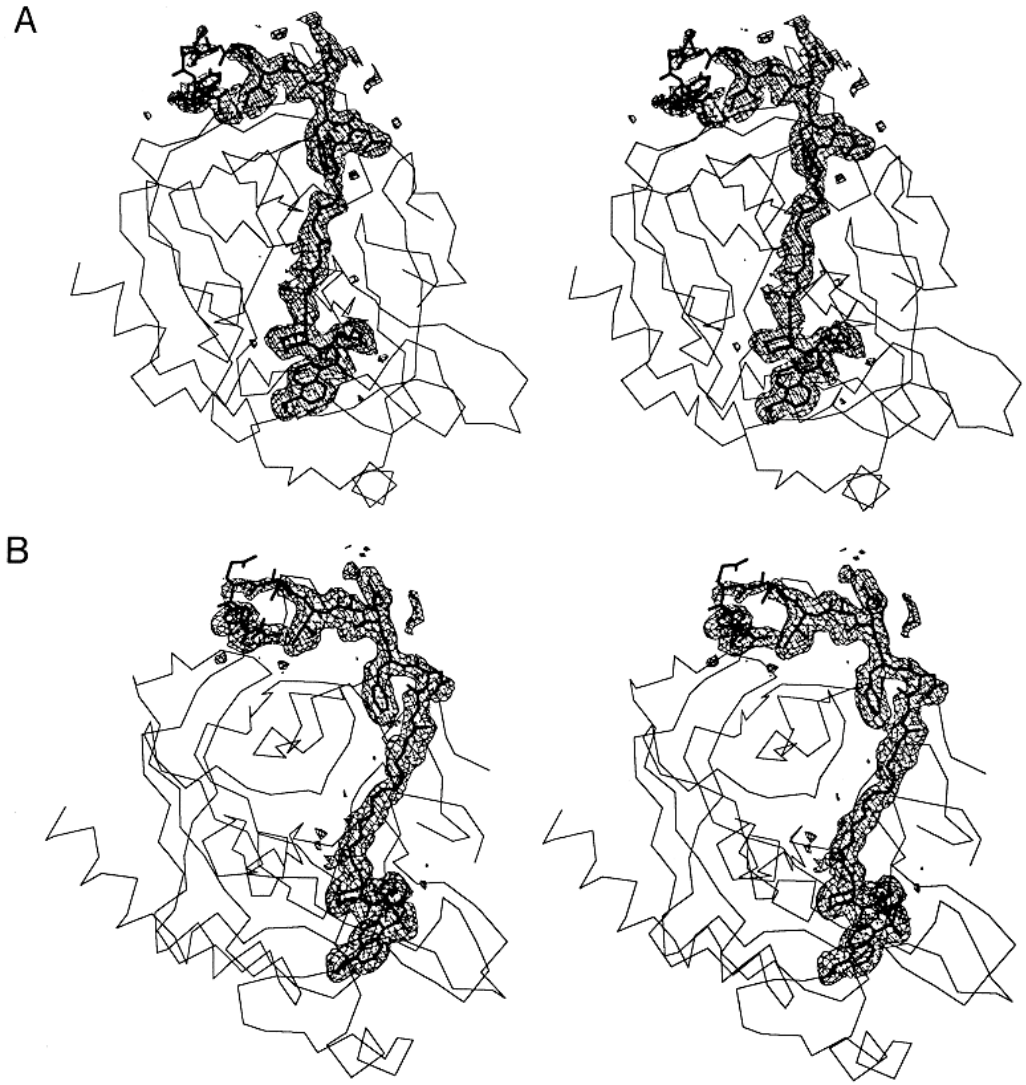
$\rho(F_o\text{-}F_c)$ = Difference map. Fc is calculate from the coordinates. This map shows missing or wrongly placed atoms.

$\rho(2F_o\text{-}F_c)$ = This is a "native" map (F_o) *plus* a difference map (F_o-F_c). This map should look like the corrected model.

Omit map = Difference map after removing *suspicious* coordinates. Removes "phase bias" density that results from least-squares refinement using wrong coordinates.



NOTE: " ρ(X) " means "maps calculated using amplitudes X and calculated phases"



FÉTHIÈRE et al, Protein Science (1996), 5: 1174- 1183.

Two inhibitor peptides in two different crystals of the protease *thrombin*.

The inhibitor coordinates were omitted from the model before calculating F_c.

Then maps were made using F_o-F_c amplitudes and Fc phases.

(stereo images)

Least-squares refinement

(4) Least squares, protein coordinates + overall B-factor.

- The *partial derivative* of the R-factor with respect to <u>each atomic position</u> can be calculated, because we know the change in amplitudes with change in coordinates.
- A 3D derivative is a "gradient". Each atom is moved down-hill along the gradient.
- "Restraints" may be imposed to maintain good stereochemistry.

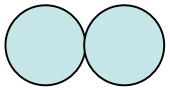
$$\frac{dR_{factor}}{d\vec{v}_i}$$

bond lengths

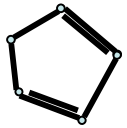
bond angles

torsion angles

Restraint types:



van der Waals



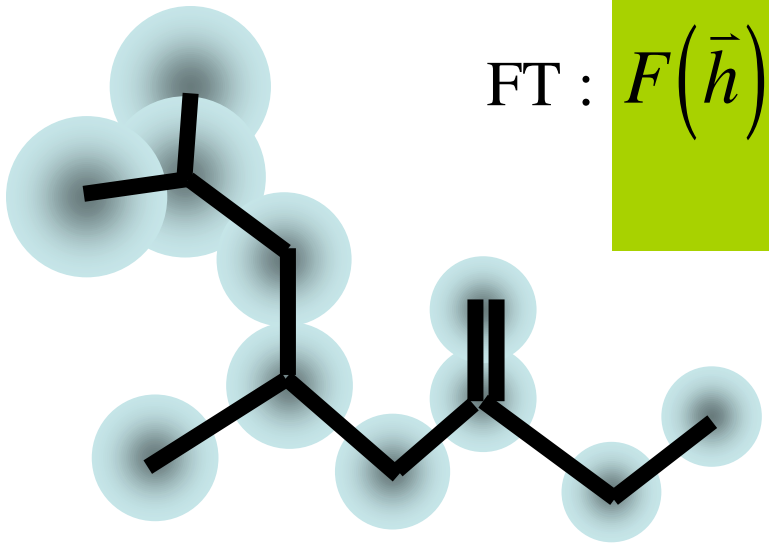
planar groups

Atomic B-factor refinement

(5) Least squares, protein coordinates + atomic B-factors.

B = "temperature factor" = Gaussian d^{-2}-dependent scale factor

Gaussian equation : $e^{-\left(B\frac{\sin^2\theta}{\lambda^2}\right)} = e^{-\left(\frac{B}{4d^2}\right)}$



FT : $$F(\vec{h}) = \sum_{g=all\ \ atoms} f(g) e^{2\pi i\left(\vec{h}\bullet\vec{r}_g\right)} e^{-\left(B_g\frac{\sin^2\theta}{\lambda}\right)}$$

Restraint: Atoms that are *bonded to each other* should not have large differences in B.

The derivative of the R-factor with respect to B can be calculated, since B-effects the amplitudes.

Because the high resolution amplitudes depend on B more than low-resolution amplitudes, high resolution (2.5Å or better) is required to refine atomic B-factors.

Refinement

Initial model to final model

Steps after initial modeling:

(1) Rigid body refinement.

(2) Density modification.

(3) Difference maps. Add waters, ions, ligands.

(4) Least squares, protein coordinates + overall B-factor.

(5) Least squares, protein coordinates + atomic B-factors.

(7) Least squares, multiple occupancy and anisotropic B-factors.

(8) Validation.