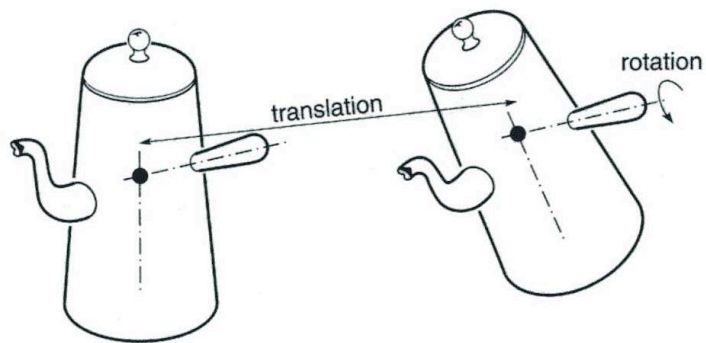
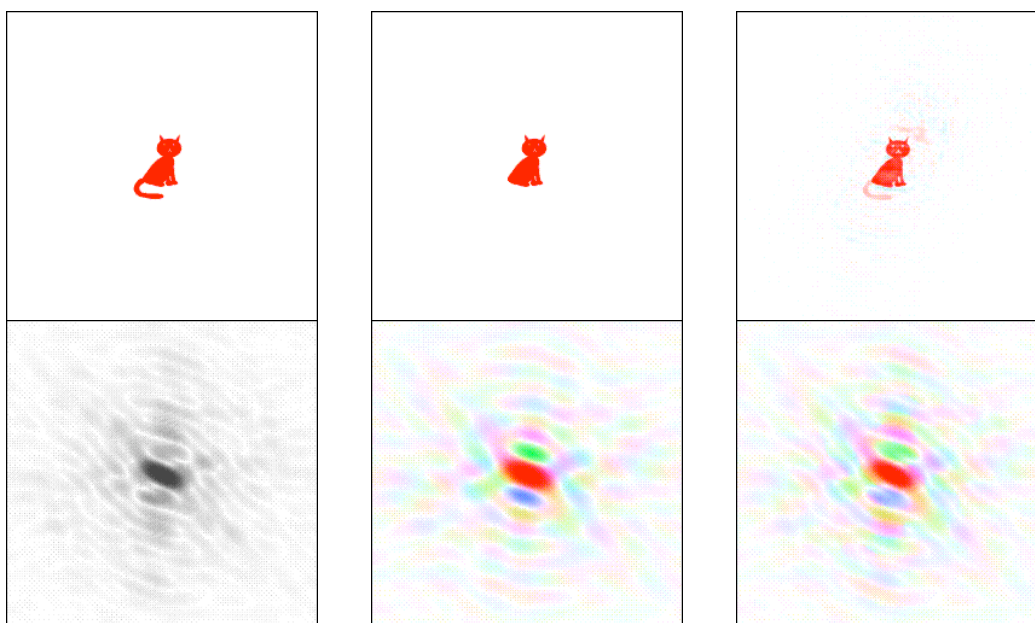


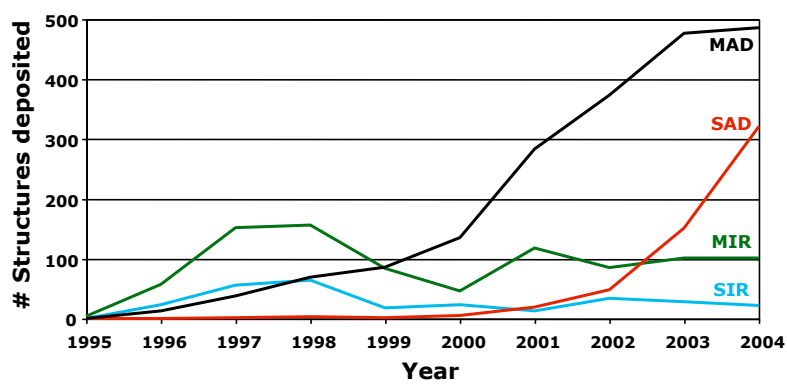
Molecular Replacement:
Rotation + Translation of the transform of a known structure





MAD is the dominant method for determining new crystal structures of proteins.

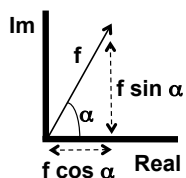
Change in Methods of Structure Determination



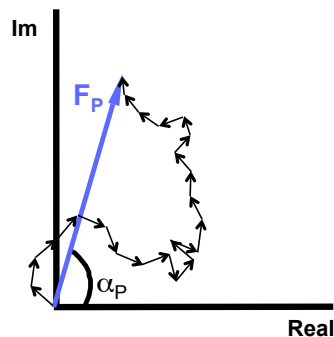
Argand Diagrams for Primary Phasing

Structure factor contribution from one atom:

$$\begin{aligned}
 &= f \exp(2\pi i(hx+ky+lz)) \\
 &= f (\cos 2\pi(hx+ky+lz) + i \sin 2\pi(hx+ky+lz)) \\
 &= f (\cos \alpha + i \sin \alpha)
 \end{aligned}$$



From all protein atoms

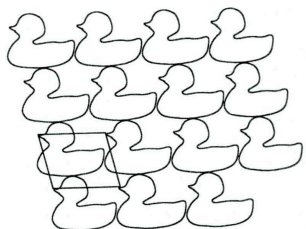


Atomic scattering (f) can be diminished by

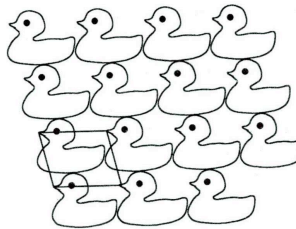
- thermal vibration (B = temperature factor)
 - partial occupancy (q = occupancy fraction)
- $$q f \exp(-B \sin^2 \theta / \lambda^2)$$

$$F(hkl) = \sum_{j=1}^{\text{\# atoms}} f_j \exp(2\pi i(hx_j+ky_j+lz_j))$$

Isomorphous Replacement



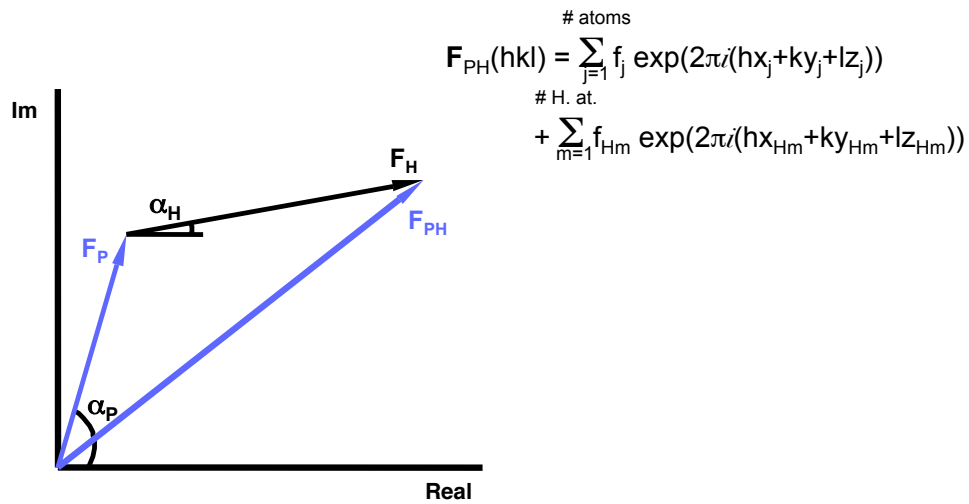
“Native” crystal



“Derivative” crystal

“Isomorphous” means that the ducks in the native crystal are identical to the ducks in the derivative crystal, except for the addition of the strongly scattering label, or “heavy atom”.

Single Isomorphous Replacement (SIR)



$$|F_{PH}|^2 = |F_P|^2 + |F_H|^2 + 2|F_P| |F_H| \cos(\alpha_P - \alpha_H)$$

How much does a heavy atom change diffracted intensities?

Because phase shifts for individual atoms are uncorrelated:

$$\sqrt{\text{avg}(|F|^2)} = \text{rms}|F| = \sqrt{\sum f^2} = \sqrt{N} f, \text{ for like atoms}$$

$$\frac{\text{rms}|F_{PH} - F_P|}{\text{rms}|F_P|} = \sqrt{\frac{N}{2}} \frac{f_H}{\text{rms}|F_P|}$$

This ratio is comparable to the “ R_{iso} ” ratio:
 $\frac{\sum (I_{PH} - I_P)}{\sum I_P}$
 because $I \propto |F|^2$

Protein of MW 50,000: $\text{rms}|F_P| \approx 400$

1 Hg atom, $q = 1$: $qf_H = 80$

“signal” = $1/\sqrt{2} (80/400) \approx 0.14$

1 Hg atom, $q = 0.5$: $qf_H = 40$

“signal” = $1/\sqrt{2} (40/400) \approx 0.07$

Estimates of Scattering Strength for Biological Macromolecules

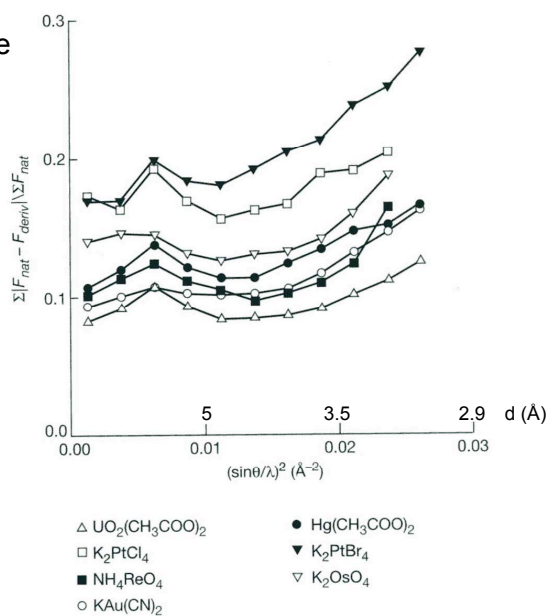
Expressions for $\langle |F| \rangle$ (rms $|F| = \sqrt{n} \times f$) for n atoms with scattering factor f :

Protein	$6.70 \times (\# \text{ atoms})^{1/2}$	$(346 \times \# \text{ residues})^{1/2}$	$(3.14 \times \text{MW})^{1/2}$
DNA	$7.20 \times (\# \text{ atoms})^{1/2}$	$(1128 \times \# \text{ residues})^{1/2}$	$(3.87 \times \text{MW})^{1/2}$
RNA	$7.26 \times (\# \text{ atoms})^{1/2}$	$(1183 \times \# \text{ residues})^{1/2}$	$(3.89 \times \text{MW})^{1/2}$

Characteristics of an average residue in protein, DNA or RNA :

amino acid	8 non-H atoms of $6.7 e^-$	MW 110
deoxyribonucleotide	21 non-H atoms of $7.2 e^-$	MW 292
ribonucleotide	21 non-H atoms of $7.26 e^-$	MW 305

Example of isomorphous signal (ΔF_{iso}):
7 derivatives of glycerate dehydrogenase



The Patterson function is used to find heavy atom positions in a derivative.

$$P(u,v,w) = 1/V \sum (|\mathbf{F}_{PH}(hkl)| - |\mathbf{F}_P(hkl)|)^2 \cos 2\pi(hu + kv + lw)$$

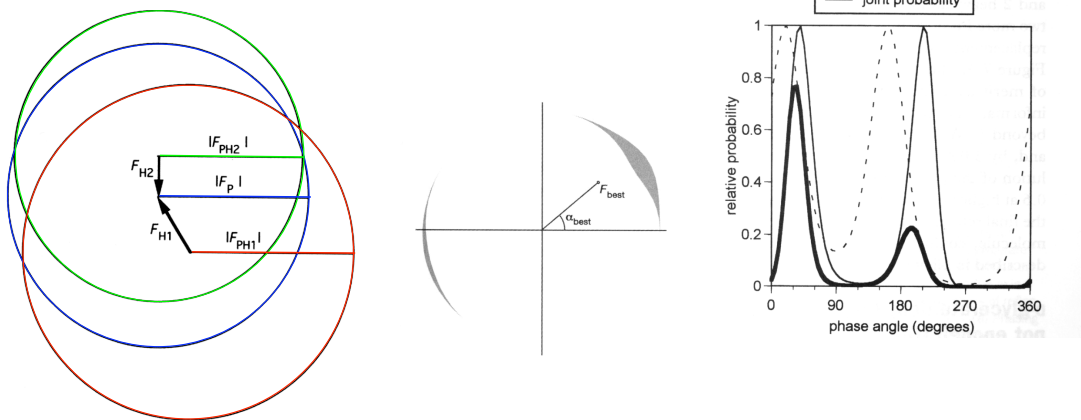
Where $(|\mathbf{F}_{PH}(hkl)| - |\mathbf{F}_P(hkl)|)$ approximates $|\mathbf{F}_H(hkl)|$.

The approximation is most accurate for reflections with
 large $|\mathbf{F}_{PH}(hkl)| - |\mathbf{F}_P(hkl)|$
and
 large $|\mathbf{F}_H(hkl)|$

$$\text{On average: } \langle ||\mathbf{F}_{PH}(hkl)| - |\mathbf{F}_P(hkl)|| \rangle = \frac{1}{\sqrt{2}} \langle |\mathbf{F}_H(hkl)| \rangle$$

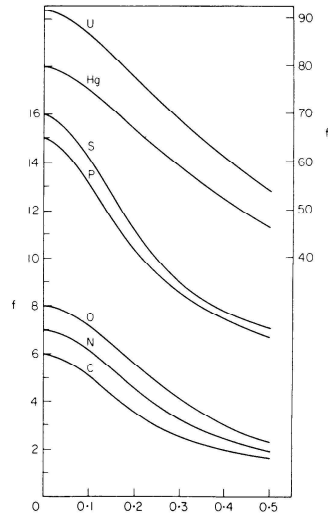
“Most probable” and “best” phase estimates

“Most probable” is the phase with highest probability.
 “Best” is the phase at the centroid of the distribution.



“Anomalous scattering” is expressed in the scattering factor, f .

“Normal” scattering, f or f_0 .



“Anomalous” scattering is complex:

$$f = f_0 + f' + i f''$$

f' = real component of anomalous scattering

f'' = imaginary component of anomalous scattering

f' and f'' are much smaller than f_0 .

f' and f'' are due to inner-shell electronic effects, so their values don't “fall off” with resolution as does f_0 .

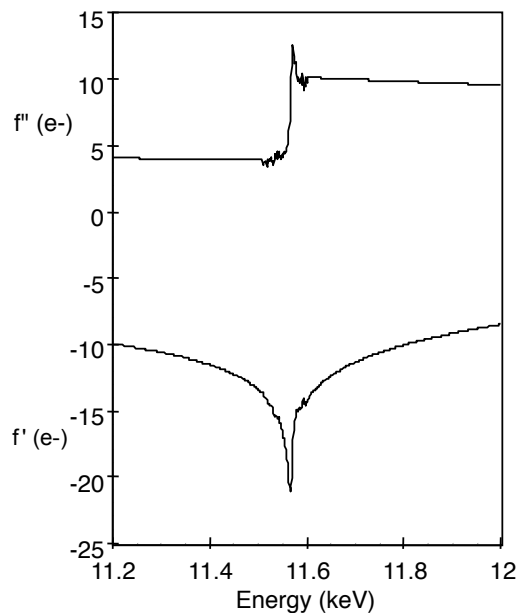
atoms

$$F(hkl) = \sum_{j=1}^{\text{# atoms}} (f_{0j} + f'_j + i f''_j) \exp(2\pi i(hx_j + ky_j + lz_j))$$

$$F''(hkl) = \sum (i f''_j) \exp(2\pi i(hx_j + ky_j + lz_j))$$

$$F'(hkl) = \sum f'_j \exp(2\pi i(hx_j + ky_j + lz_j))$$

Anomalous Scattering Factors for Pt L_{III} edge



f'' is proportional to the atomic absorption coefficient (μ_a):

$$f''(E) = \frac{mc}{4\pi e^2 \hbar} E \mu_a(E)$$

f' is calculated from f'' by Kramers-Kronig transformation:

$$f'(E) = \frac{2\delta}{\pi} \sum_{i=0}^{\infty} \frac{E_i f''_i}{E^2 - E_i^2}$$

(δ = energy sampling increment)

Positions of X-ray absorption edges

Periodic table form:

http://www.bmsc.washington.edu/scatter/AS_periodic
(including calculation of theoretical f'' and f' spectra)

Tabular form:

http://xdb.lbl.gov/Section1/Table_1-1.pdf

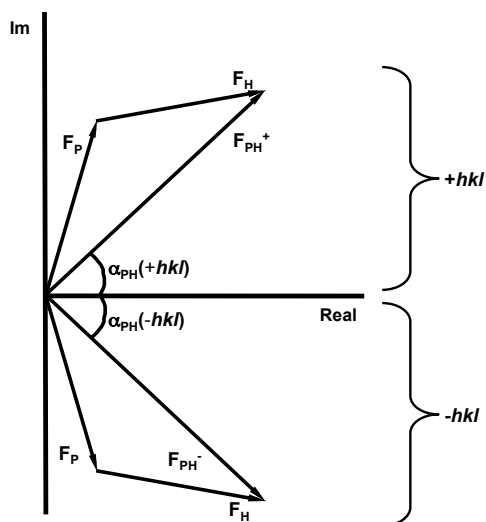
Edges of interest to biological crystallography:

Se - K	12.658 keV
Br - K	13.474 keV
Fe - K	7.112 keV
Zn - K	9.659 keV
Hg - L _{III}	12.284 keV
Pt - L _{III}	11.564 keV

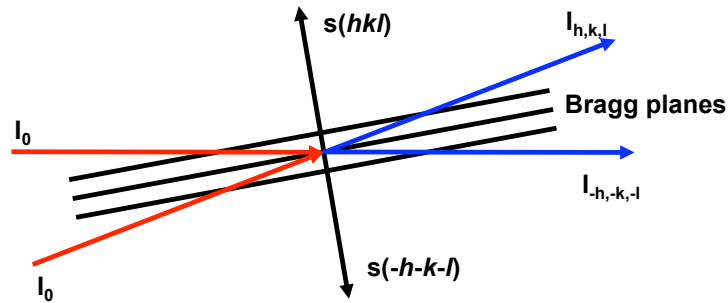
Accessible at most beamlines: 7-15 keV

“Sweet” zone: 11-14 keV

Relationship of $F(h,k,l)$ and $F(-h,-k,-l)$ without anomalous scattering



Friedel pairs: reflections h,k,l and $-h,-k,-l$



Bijvoet pairs: any symmetry-equivalent reflections related by a mirror or inversion
Example, in an orthorhombic space group, such as $P2_12_12_1$:

A	B
h,k,l	$-h,-k,-l$
$h,-k,-l$	$-h,k,l$
$-h,k,-l$	$h,-k,l$
$-h,-k,l$	$h,k,-l$

Any reflection in group A and any reflection in group B are Bijvoet mates.
Each reflection in group A and the reflection on the same line in group B are Friedel mates.

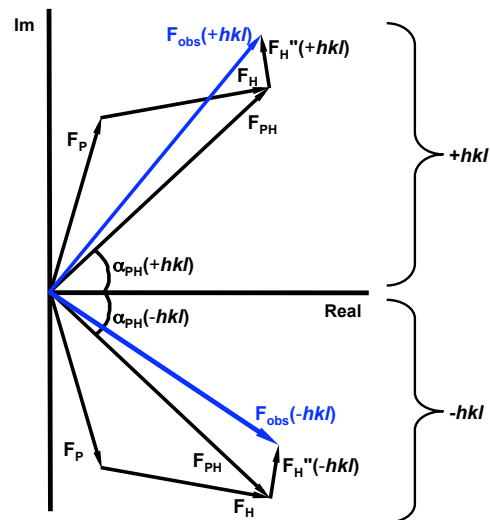
Relationship of $F(h,k,l)$ and $F(-h,-k,-l)$ with anomalous scattering

$$f = f_0 + f' + i f''$$

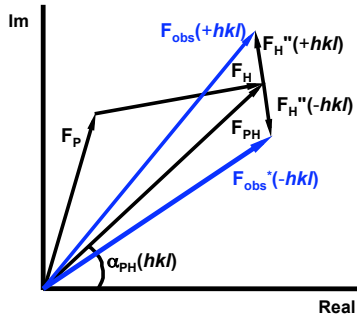
$$F(hkl) = \sum_{j=1}^{\text{\# atoms}} (f_{0j} + f'_j + i f''_j) \exp(2\pi i(hx_j + ky_j + lz_j))$$

$$F''(hkl) = \sum (i f''_j) \exp(2\pi i(hx_j + ky_j + lz_j))$$

$$F'(hkl) = \sum f'_j \exp(2\pi i(hx_j + ky_j + lz_j))$$



How does anomalous scattering provide phase information?



$$|F_{PH}^+|^2 = |F_P|^2 + |F_H|^2 + |F_H''|^2 + 2|F_P| |F_H| \cos(\alpha_P - \alpha_H) + 2|F_P| |F_H''| \cos(\alpha_P - \alpha_H + \omega) + 2|F_H| |F_H''| \cos(\omega)$$

$$|F_{PH}^-|^2 = |F_P|^2 + |F_H|^2 + |F_H''|^2 + 2|F_P| |F_H| \cos(\alpha_P - \alpha_H) - 2|F_P| |F_H''| \cos(\alpha_P - \alpha_H + \omega) - 2|F_H| |F_H''| \cos(\omega)$$

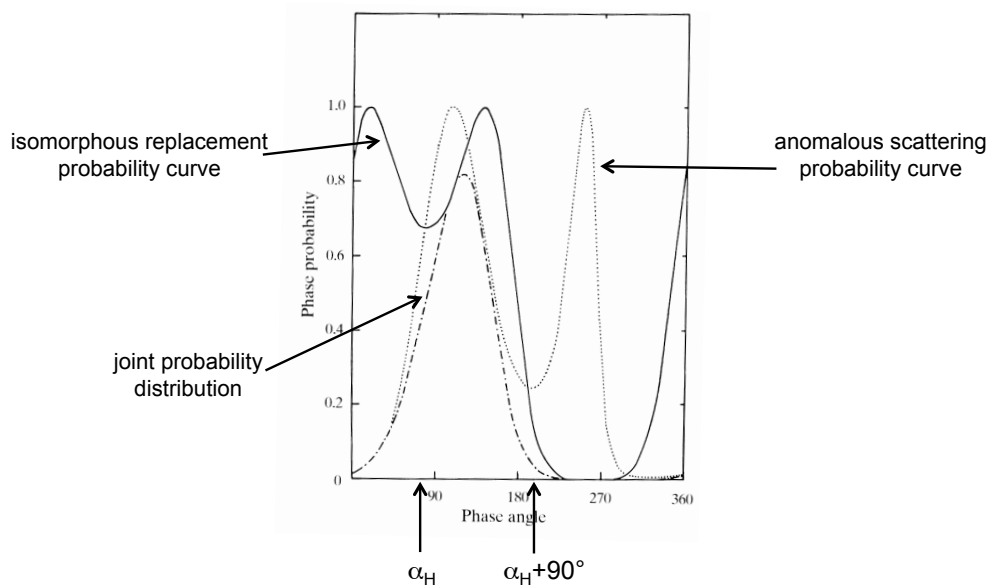
$$|F_{PH}^+|^2 - |F_{PH}^-|^2 = 4|F_P| |F_H''| \cos(\alpha_P - \alpha_H + \omega) + 4|F_H| |F_H''| \cos(\omega)$$

$$F''(hkl) = \sum_{j=1}^{\text{\# atoms}} (f_j'') \exp(2\pi i(hx_j + ky_j + lz_j))$$

Good news: $\omega = \pi/2$ if only one atomic type has detectable anomalous scattering.

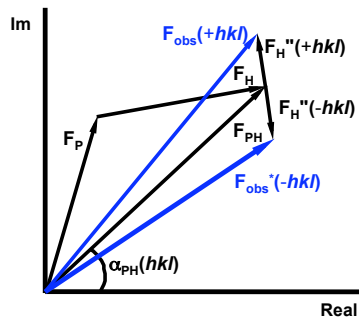
$$|F_{PH}^+|^2 - |F_{PH}^-|^2 = 4|F_P| |F_H''| \sin(\alpha_P - \alpha_H)$$

Phase probability distributions for anomalous scattering with isomorphous replacement



Source: Matthews, 2001

Single-wavelength Anomalous Diffraction, a.k.a. **SAD**



$$|F_{PH}^+|^2 = |F_{PH}|^2 + |F_H''|^2 + 2|F_{PH}| |F_H''| \cos(\alpha_{PH} - \alpha_H + \omega)$$

$$|F_{PH}^-|^2 = |F_{PH}|^2 + |F_H''|^2 - 2|F_{PH}| |F_H''| \cos(\alpha_{PH} - \alpha_H + \omega)$$

$$|F_{PH}^+|^2 - |F_{PH}^-|^2 = 4|F_{PH}| |F_H''| \cos(\alpha_{PH} - \alpha_H + \omega)$$

$$\omega = \pi/2$$

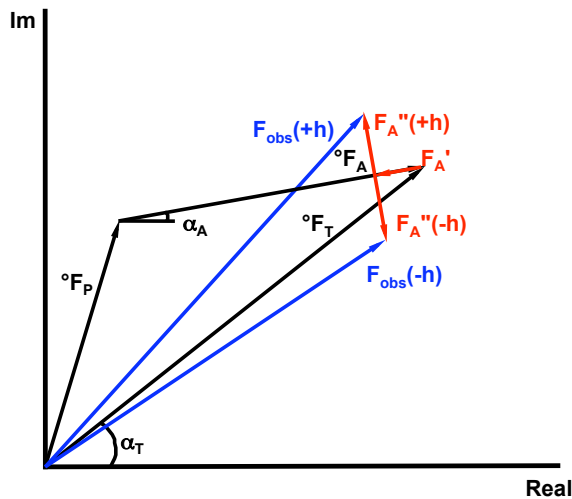
$$|F_H''| = f''/f_0 |F_H|$$

$$|F_{PH}^+|^2 - |F_{PH}^-|^2 = 4|F_{PH}| |F_H''| \sin(\alpha_{PH} - \alpha_H)$$

$$P(\alpha) = N \exp\left\{\frac{-\varepsilon^2}{2E^2}\right\}$$

$$\varepsilon = F^+ - F^- + 2 \frac{f_A''}{f_A} F_A \sin(\alpha_T - \alpha_A)$$

Structure factor definitions for anomalous scattering (without isomorphous replacement)



$^{\circ}F_T$ = Total normal scattering

$^{\circ}F_A$ = Normal scattering from anomalously scattering atoms

F_A'' = Imaginary component of anomalous scattering

F_A' = Real component of anomalous scattering

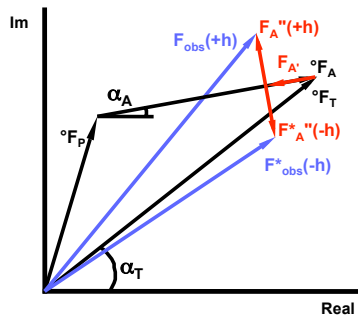
Color code

Due to "normal" scattering only

Due to anomalous scattering only

From experiment

Multiwavelength Anomalous Diffraction, a.k.a. MAD



$$|F_{\text{obs}}(\lambda, \pm h)|^2 = |^0F_T|^2 + |F'_A(\lambda)|^2 + |F''_A(\lambda)|^2 \\ + |^0F_T| |F'_A(\lambda)| \cos(\alpha_T - \alpha_A) \\ \pm |^0F_T| |F''_A(\lambda)| \sin(\alpha_T - \alpha_A)$$

$$|F_{\text{obs}}(\lambda, \pm h)|^2 = |^0F_T|^2 + a(\lambda) |^0F_A|^2 \\ + b(\lambda) |^0F_T| |^0F_A| \cos(\alpha_T - \alpha_A) \\ \pm c(\lambda) |^0F_T| |^0F_A| \sin(\alpha_T - \alpha_A)$$

$$a(\lambda) = (f'_{\lambda}{}^2 + f''_{\lambda}{}^2) / ^0f_A{}^2$$

$$b(\lambda) = 2(f'_{\lambda}) / ^0f_A \quad F'_A(\lambda) = (f'_{\lambda} / ^0f_A) |^0F_A|$$

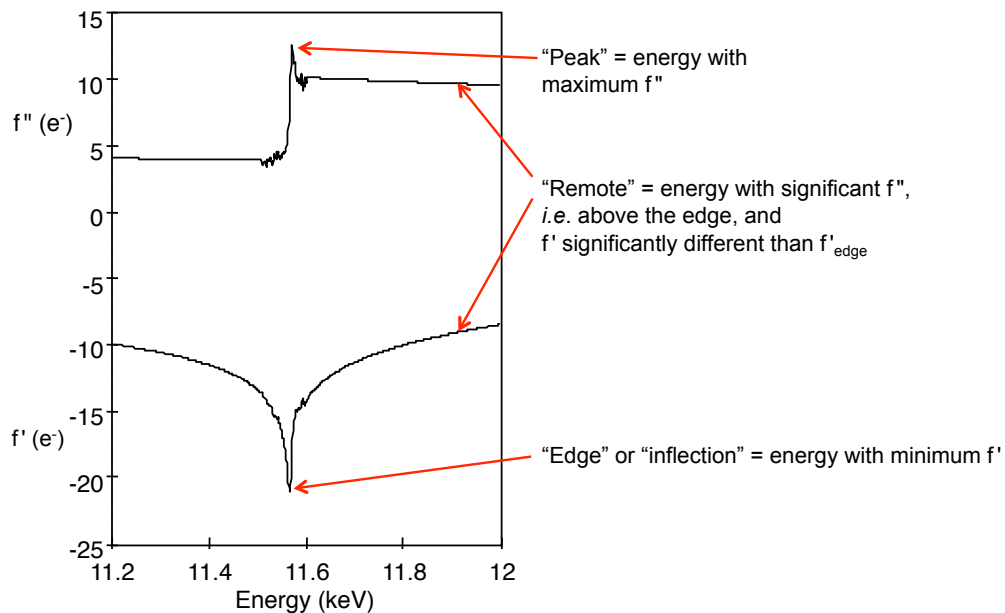
$$c(\lambda) = 2(f''_{\lambda}) / ^0f_A \quad F''_A(\lambda) = (f''_{\lambda} / ^0f_A) |^0F_A|$$

from experiment

wavelength dependent

wavelength independent

Energy selection for a MAD experiment



Phasing signal in MAD

For like atoms: $\sqrt{\text{avg}(|F|^2)} = \text{rms } |F| = \sqrt{\sum f^2} = \sqrt{N} f$

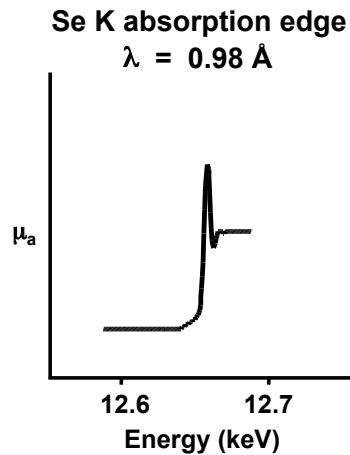
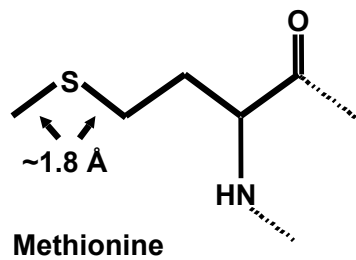
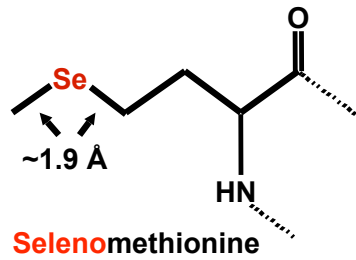
Anomalous signal: $\frac{\text{rms } |F^+ - F^-|}{\text{rms } |F_T|} = \sqrt{\frac{N}{2}} \frac{2f_A''}{\text{rms } |F_T|} = \sqrt{2N} \frac{f_A''}{\text{rms } |F_T|}$

Dispersive signal: $\frac{\text{rms } |F_{\lambda 1} - F_{\lambda 2}|}{\text{rms } |F_T|} = \sqrt{\frac{N}{2}} \frac{f_A'}{\text{rms } |F_T|}$

Isomorphous signal: $\frac{\text{rms } |F_{PH} - F_P|}{\text{rms } |F_P|} = \sqrt{\frac{N}{2}} \frac{f_H}{\text{rms } |F_P|}$

Phasing signal in MAD for example edges

Atom	f_0 (e ⁻)	f'_{edge} (e ⁻)	f''_{peak} (e ⁻)	f'_{remote} (e ⁻)	Edge	Energy (keV)
Se	34	-10	6	-3	K	12.658
Br	35	-7	4	-3	K	13.446
Fe	26	-8	4	-3	K	7.115
Zn	30	-9	4	-3	K	9.666
Hg	80	-20	10	-8	L _{III}	12.283
Pt	78	-21	13	-8	L _{III}	11.565
Yb	70	-33	35	-7	L _{III}	8.947



The average protein has 1 Met per 60 residues
(excluding Met1, which is usually disordered).

Patterson Function

$$P(u,v,w) = 1/V \sum |F(hkl)|^2 \cos 2\pi(hu + kv + lw)$$

Equivalent to a density map from intensities (I) with all phases = 0

Also equivalent to:

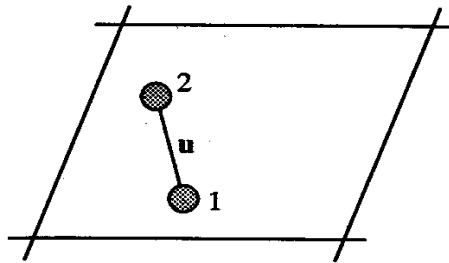
$$P(\mathbf{u}) = \int \rho(\mathbf{r}) \times \rho(\mathbf{r} + \mathbf{u}) d\mathbf{v} \quad \mathbf{r} = x,y,z \quad \mathbf{u} = u,v,w$$

This produces a map of all interatomic vectors in the structure.
 peaks at the end points of vectors from all atom i to atom j
 also atom j to atom i

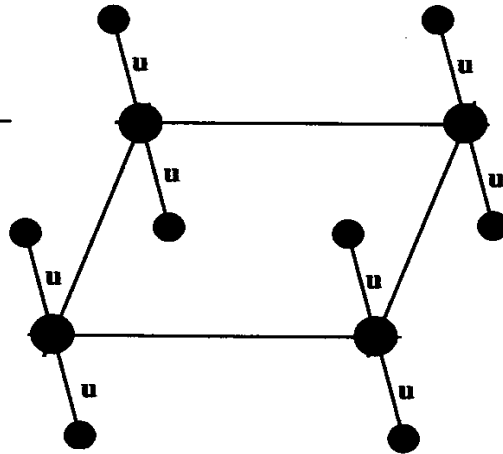
Patterson cell is the same size & shape as the real unit cell.
 Has all symmetry in the real unit cell + inversion symmetry.

<p>Reminder: electron density equation</p> $\rho(x,y,z) = 1/V \sum F(hkl) \exp(i\alpha(hkl) - 2\pi i(hx + ky + lz))$ $= 1/V \sum F(hkl) (\cos(\alpha - 2\pi(hx + ky + lz)) + i \sin(\alpha - 2\pi(hx + ky + lz)))$

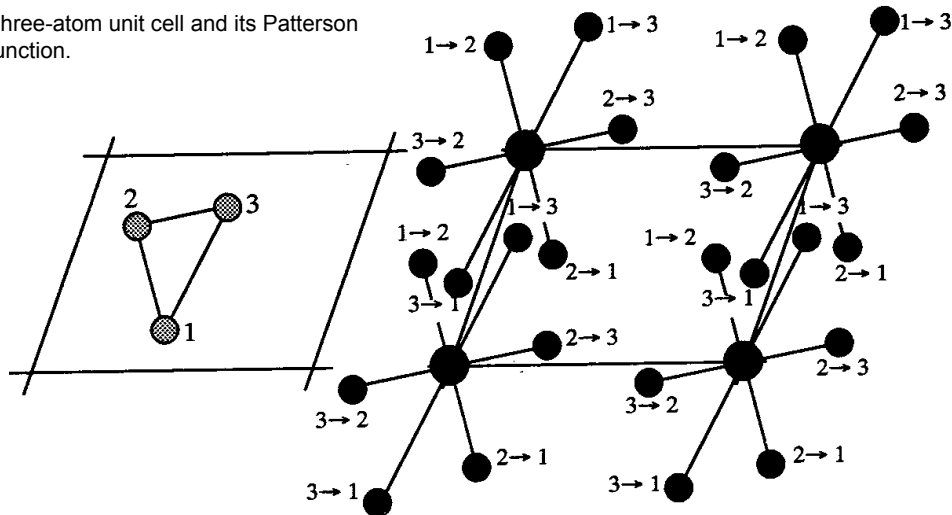
Real unit cell (symmetry $P1$) contains 2 atoms ("1" and "2"), separated by vector \mathbf{u} .



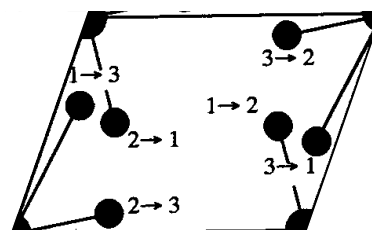
Patterson cell contains all interatomic vectors between atoms "1" and "2". Added symmetry: inversion center.



Three-atom unit cell and its Patterson function.



The Patterson function for an n -atom structure contains n^2 interatomic vectors. The origin peak includes self-vectors (each atom to itself), so the number of off-origin peaks is $n^2 - n$. In this example, there are 6 vectors ($3^2 - 3$).



Harker vectors: special Patterson between symmetry-equivalent atoms

Example: space group $P2$

Equivalent positions:

For every heavy atom at x, y, z
 Another is found at $-x, y, -z$

Vector between these two positions is called a “Harker vector”:

$$u, v, w = (x - (-x), (y - y), (z - (-z)) = 2x, 0, 2z$$

In space group $P2$, we examine the $v = 0$ section first.

Solving heavy-atom Pattersons for the heavy-atom positions usually starts with considering all possible Harker and “cross” vectors.

Patterson vectors between two heavy atoms in space group $P2$

	x_1, y_1, z_1	$-x_1, y_1, -z_1$	x_2, y_2, z_2	$-x_2, y_2, -z_2$
x_1, y_1, z_1	0, 0, 0	$-2x_1, 0, -2z_1$	$x_2 - x_1, y_2 - y_1, z_2 - z_1$	$-x_1 - x_2, y_2 - y_1, -z_1 - z_2$
$-x_1, y_1, -z_1$	$2x_1, 0, 2z_1$	0, 0, 0	$x_1 + x_2, y_2 - y_1, z_1 + z_2$	$x_1 - x_2, y_2 - y_1, z_1 - z_2$
x_2, y_2, z_2	$x_1 - x_2, y_1 - y_2, z_1 - z_2$	$-x_1 - x_2, y_1 - y_2, -z_1 - z_2$	0, 0, 0	$-2x_2, 0, -2z_2$
$-x_2, y_2, -z_2$	$x_1 + x_2, y_1 - y_2, z_1 + z_2$	$x_2 - x_1, y_1 - y_2, z_2 - z_1$	$2x_2, 0, 2z_2$	0, 0, 0

Notice the relationships of these vectors:

Inversion symmetry: $2x_1, 0, 2z_1$ vs. $-2x_1, 0, -2z_1$

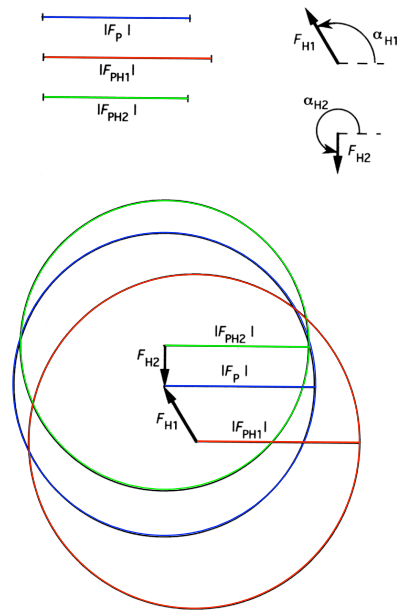
$x_1 - x_2, y_1 - y_2, z_1 - z_2$ vs. $x_2 - x_1, y_2 - y_1, z_2 - z_1$

Mirror symmetry: $x_1 + x_2, y_1 - y_2, z_1 + z_2$ vs. $x_1 + x_2, y_2 - y_1, z_1 + z_2$

Same v section: $x_1 - x_2, y_1 - y_2, z_1 - z_2$ vs. $x_1 + x_2, y_1 - y_2, z_1 + z_2$

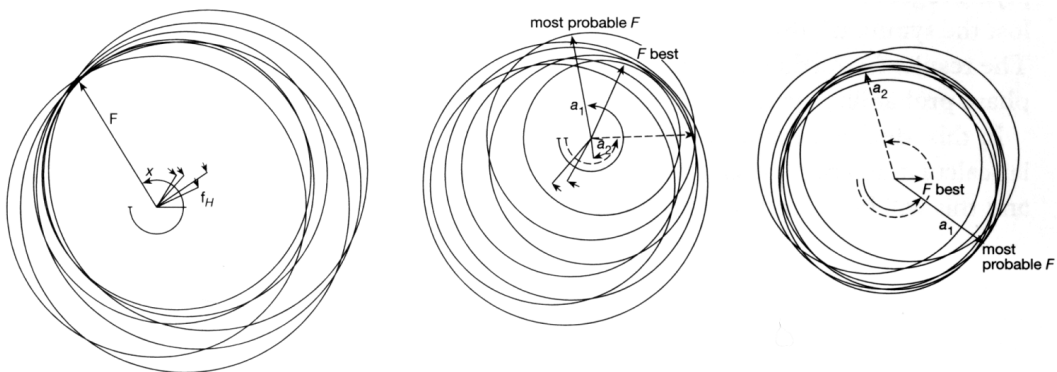
Typical example of a reflection
and two derivatives.

Note the “lack of closure”
where the circles are closest
to intersecting.



Source: Blow, chap. 7

Three example reflections from a six-derivative MIR problem.



6 derivatives agree:
strong phase indication

some agreement:
moderate phase indication

poor agreement:
no phase indication

Source: Blow, chap. 7