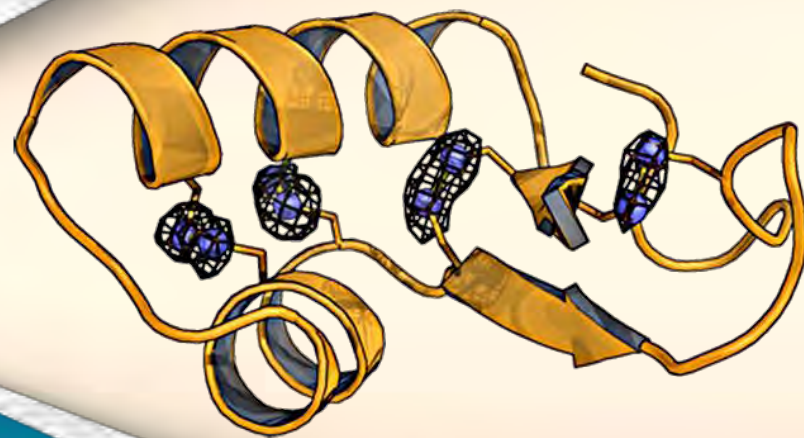


Experimental phasing, Pattersons and SHELX

Andrea Thorn



What is experimental phasing?

**Experimental phasing is what you do
if MR doesn't work.**



What is experimental phasing?

Experimental phasing methods depend on intensity differences.

These differences are caused by a marker substructure of certain elements.

MAD and **SAD** exploit the anomalous signal from one or more data sets from the same crystal.

SIR (special case: **RIP**) and **MIR** utilizes several heavy-atom soaked derivative crystals. They have to be isomorphous to be utilized.



Methods

- Single wavelength anomalous diffraction (**SAD**)
 - Native sulfur-based SAD (**S-SAD**)
- Multiple wavelength anomalous diffraction (**MAD**)
- Single isomorphous replacement (**SIR**)
 - Radiation-induced phasing (**RIP**)
- Single isomorphous replacement with anomalous scattering (**SIRAS**)
- Multiple isomorphous replacement with anomalous scattering (**MIRAS**)



Theory

STRUCTURE FACTORS

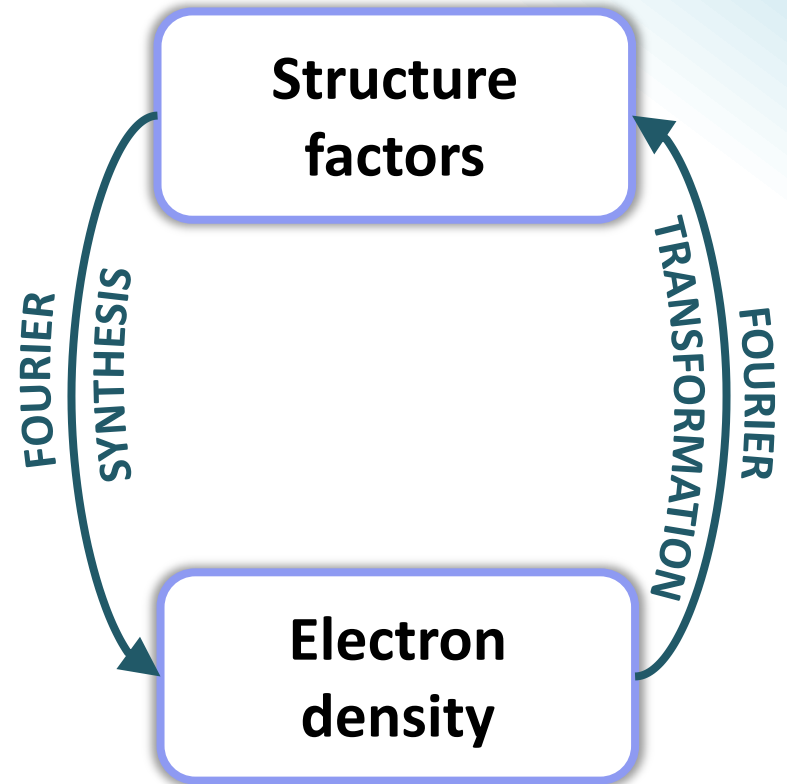
A decorative graphic at the bottom of the slide consisting of multiple thin, parallel lines that form a wavy, ribbon-like shape. The lines are light blue and white, and they curve across the bottom of the page, creating a sense of motion and depth.

Structure factors

For each reflection, there is a

structure factor F_{hkl}

If we know the structure factors including their phases for all reflections, we can easily calculate the electron density map, and hence get the structure.

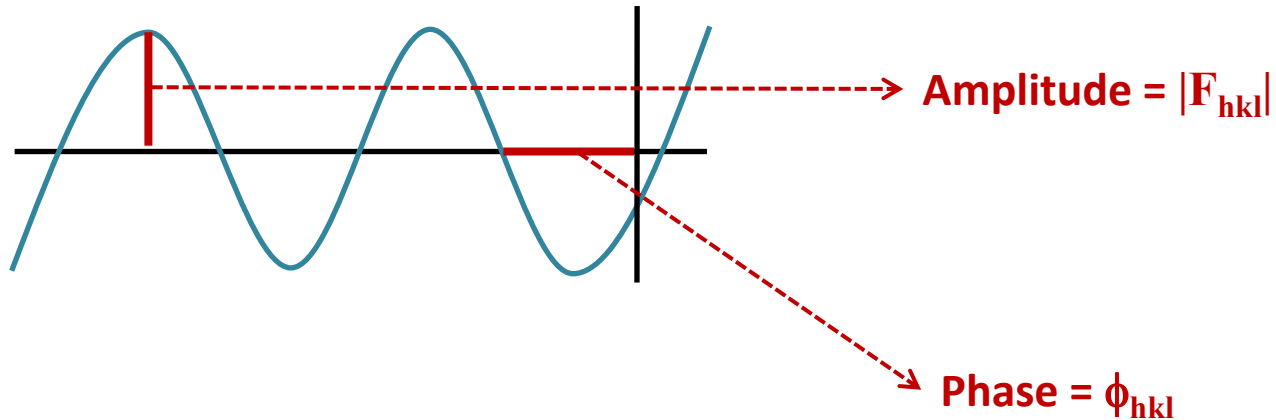


Structure factors

For each reflection, there is a

structure factor F_{hkl}

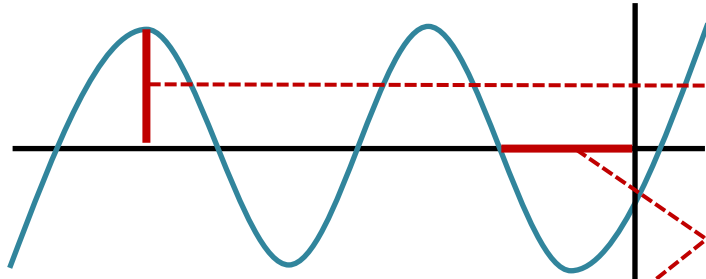
= a wave



Structure factors

structure factor F_{hkl}

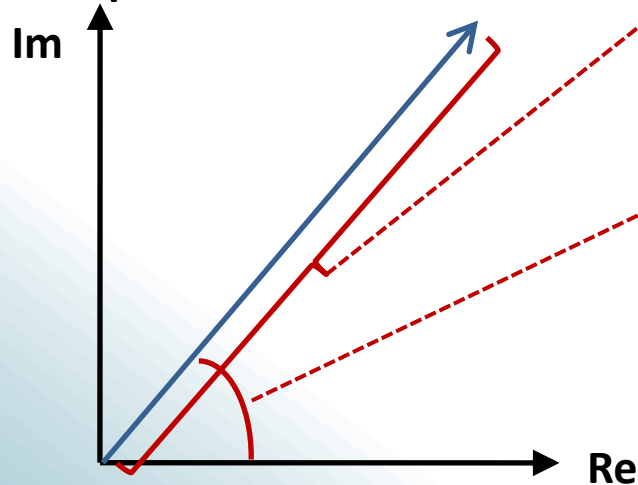
= a wave



Amplitude = $|F_{hkl}|$

$|F_{hkl}|^2 \sim I_{hkl}$ Intensity

= a complex number



Phase = ϕ_{hkl}

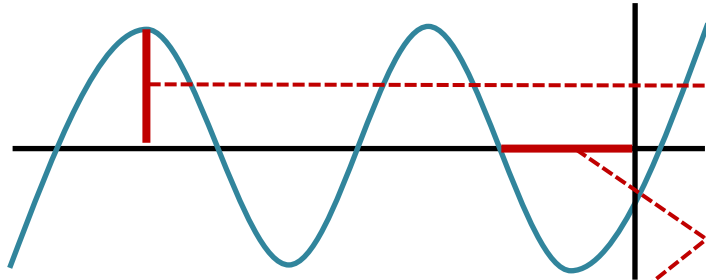
cannot be measured... :-)



Structure factors

structure factor F_{hkl}

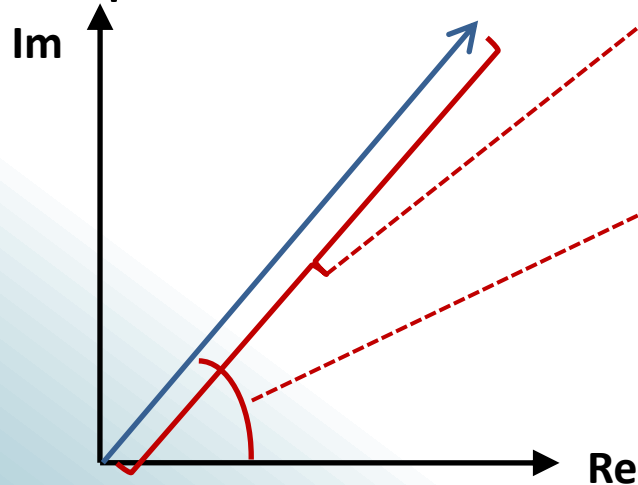
= a wave



Amplitude = $|F_{hkl}|$

$|F_{hkl}|^2 \sim I_{hkl}$ Intensity

= a complex number



Phase = ϕ_{hkl}

cannot be measured... :-)



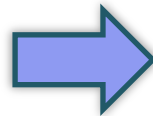
Structure factors

$$\text{Amplitude} = |F_{hkl}|$$

$$|F_{hkl}|^2 \sim I_{hkl} \text{ Intensity } \checkmark$$

$$\text{Phase} = \phi_{hkl}$$

cannot be measured... :-)



PHASE PROBLEM

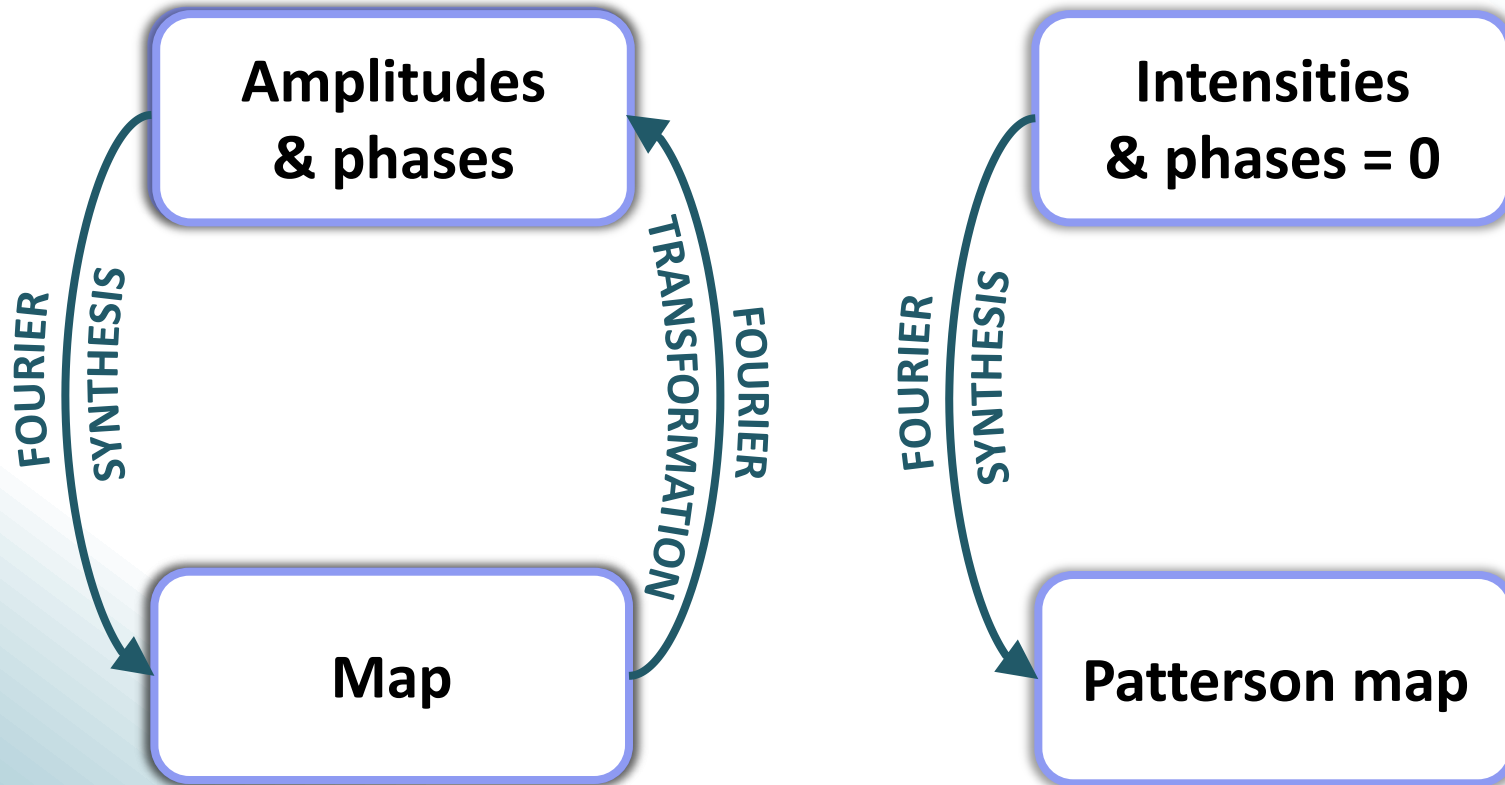
The central problem
of crystallography



Theory

PATTERSON MAPS

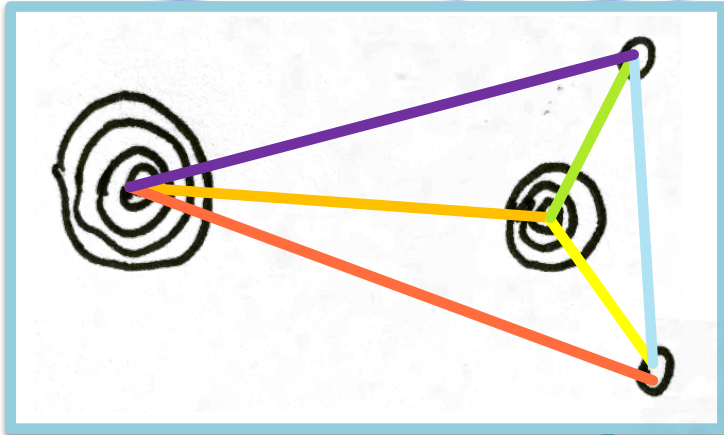
Calculating a map - Patterson



What does it look like?

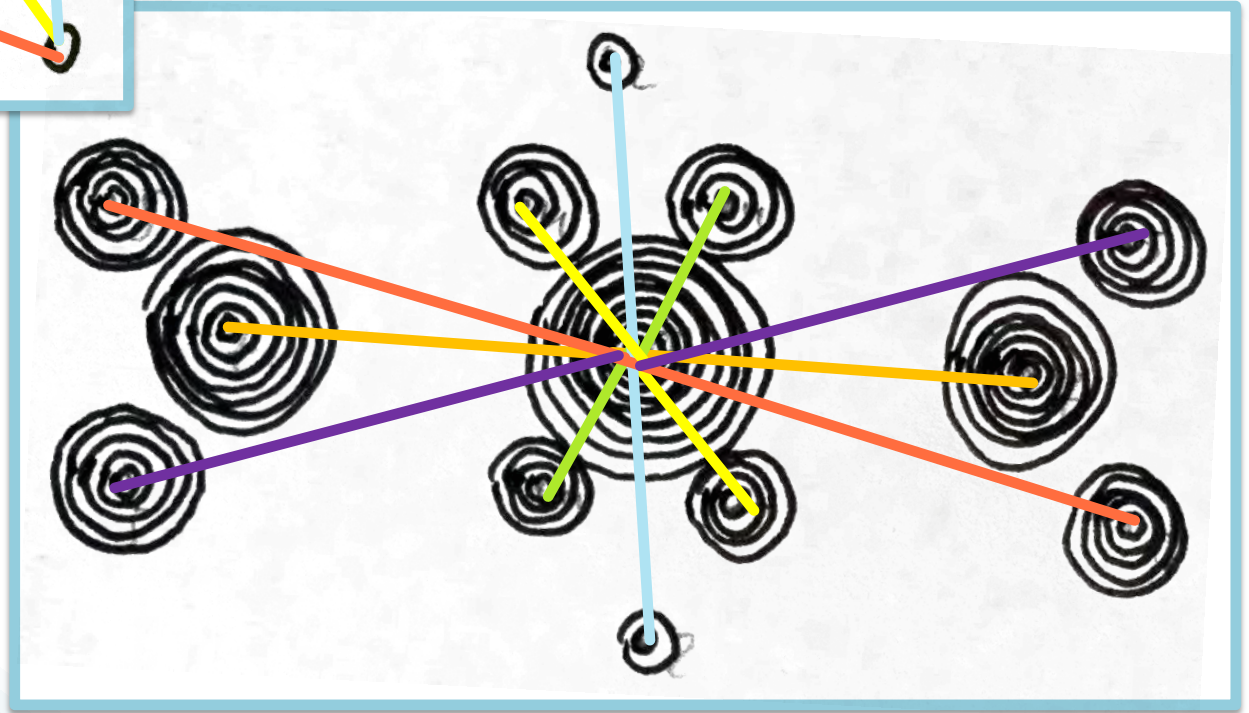


Calculating a map - Patterson

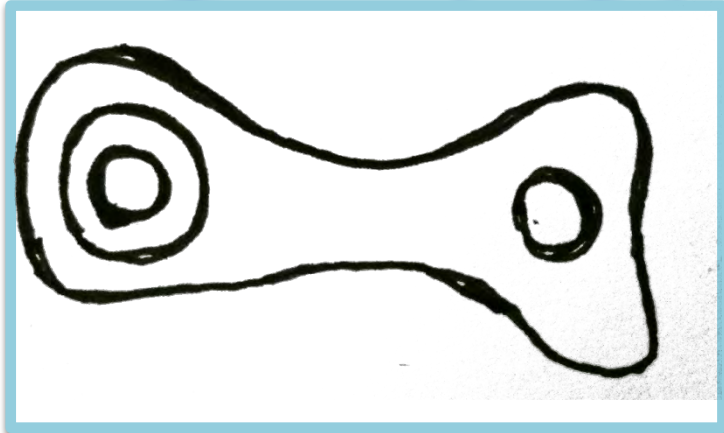


- Interatomic vectors
- No relative positions
- Handedness is not resolved.

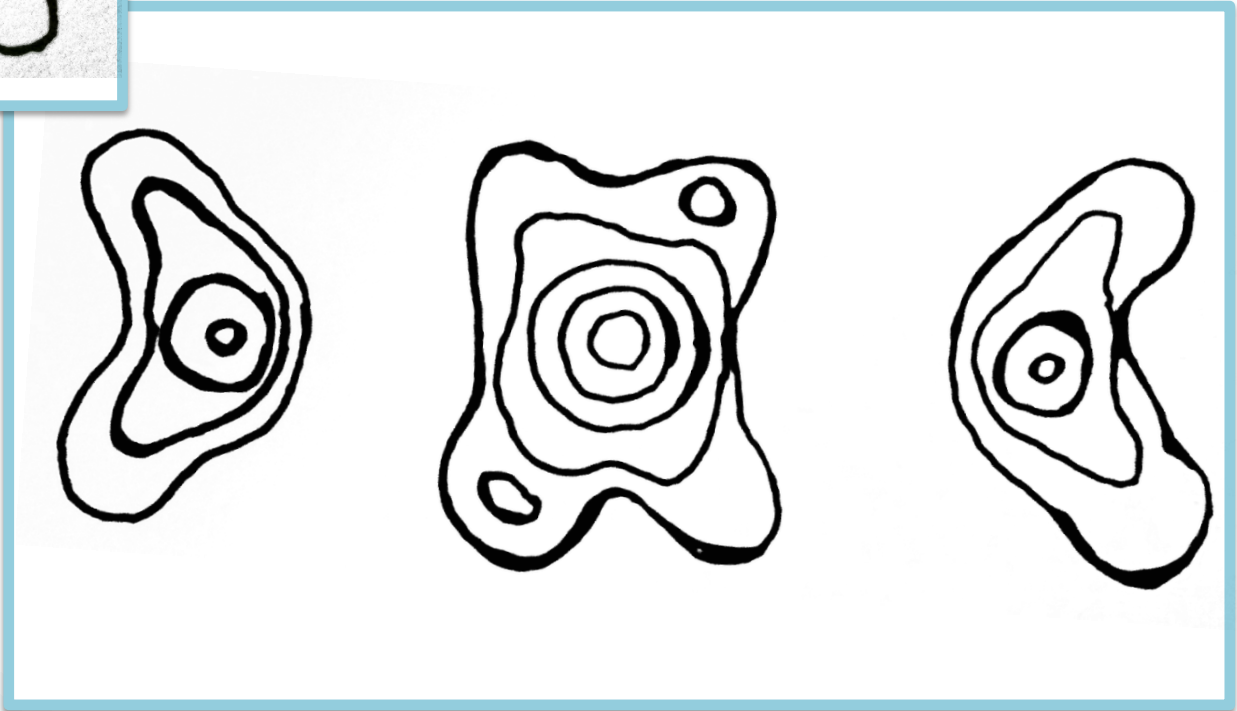
$n^2 - n$
peaks in a
Patterson map



Calculating a map - Patterson



Problem: Resolution
(number of vectors)

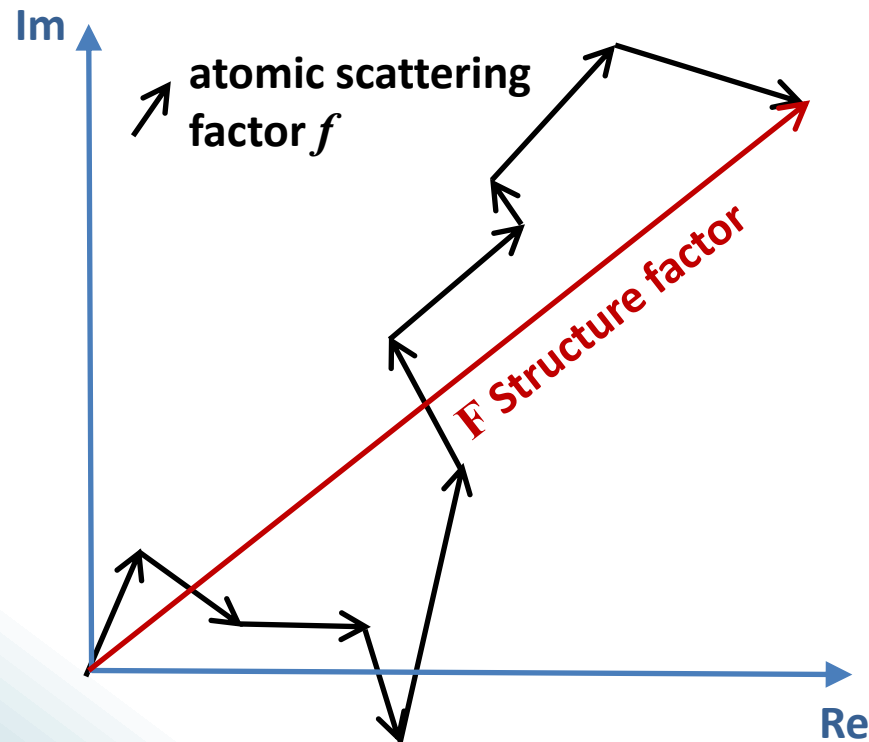


Theory

ANOMALOUS SCATTERING

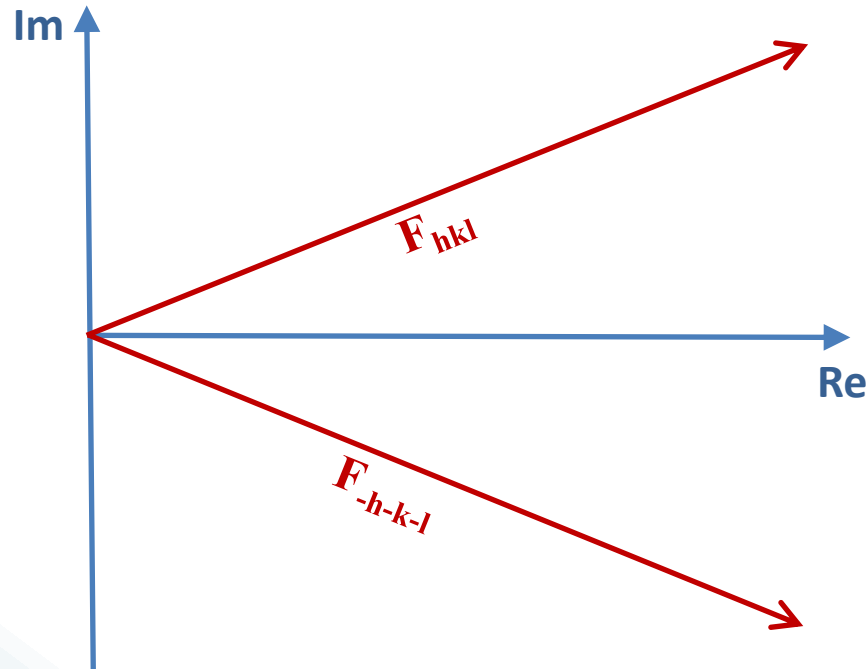
The anomalous signal

Each structure factor is composed of contributions f from each atom:



The anomalous signal

Friedel's law: $|F_{hkl}| = |F_{-h-k-l}|$ $\phi_{hkl} = -\phi_{-h-k-l}$



The anomalous signal

But in reality, there is **anomalous scattering** due to resonance with electronic transitions in the atom:

$$f = f_0 + f' + if''$$

depends solely on resolution

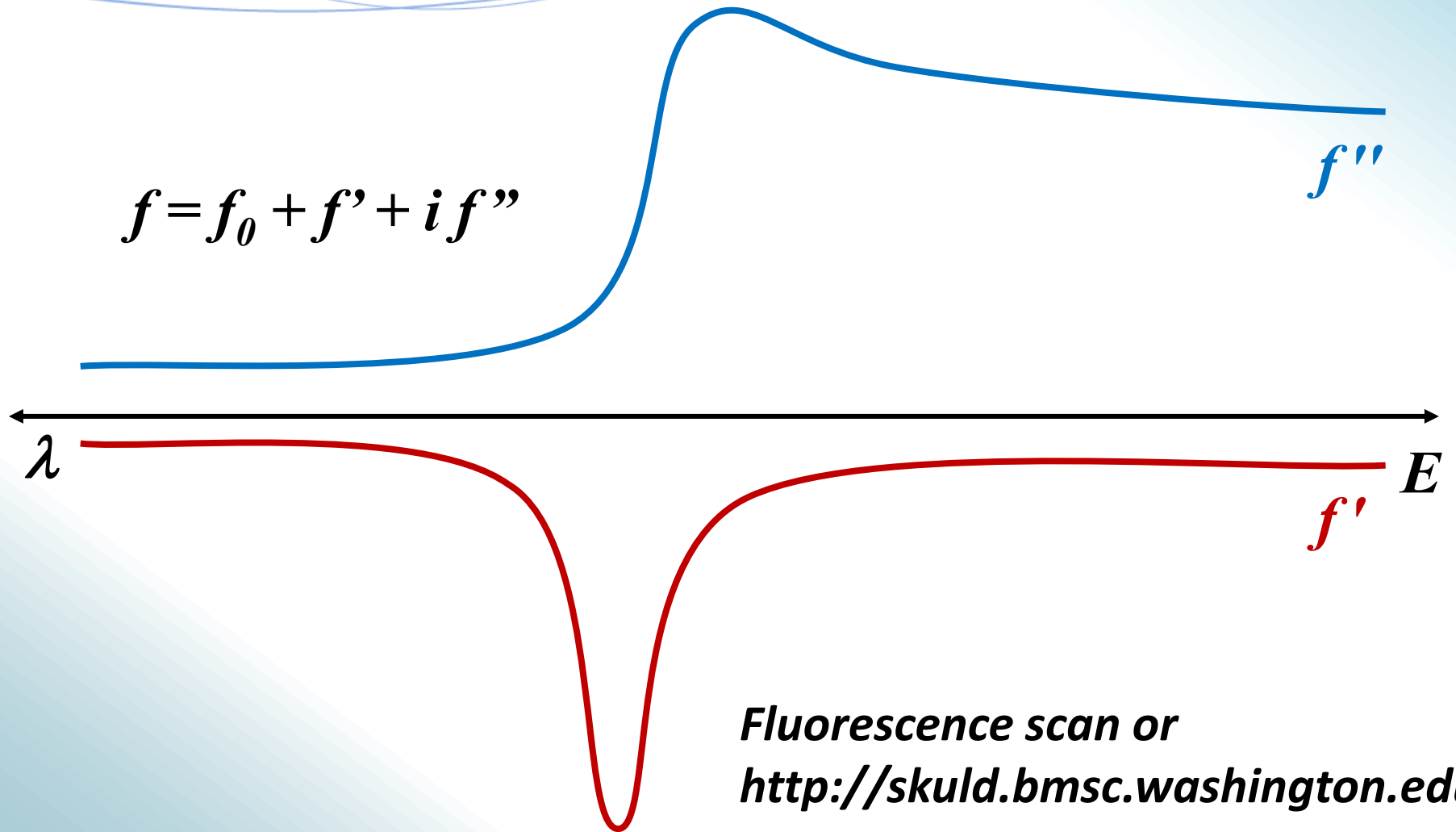
real component

imaginary component

f' and f'' are observed near absorption edges of the atom's element, and are λ -dependent



The anomalous signal



*Fluorescence scan or
<http://skuld.bmsc.washington.edu>*

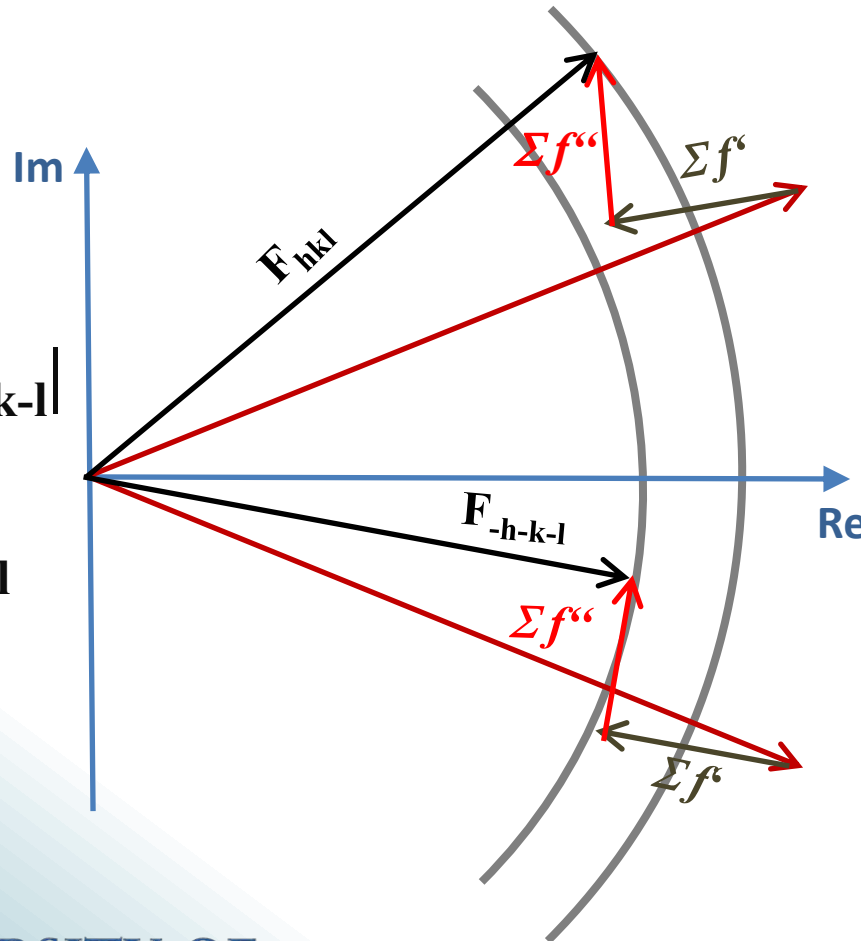


The anomalous signal

f'' breaks Friedel's law:

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

$$\phi_{hkl} \neq -\phi_{-h-k-l}$$

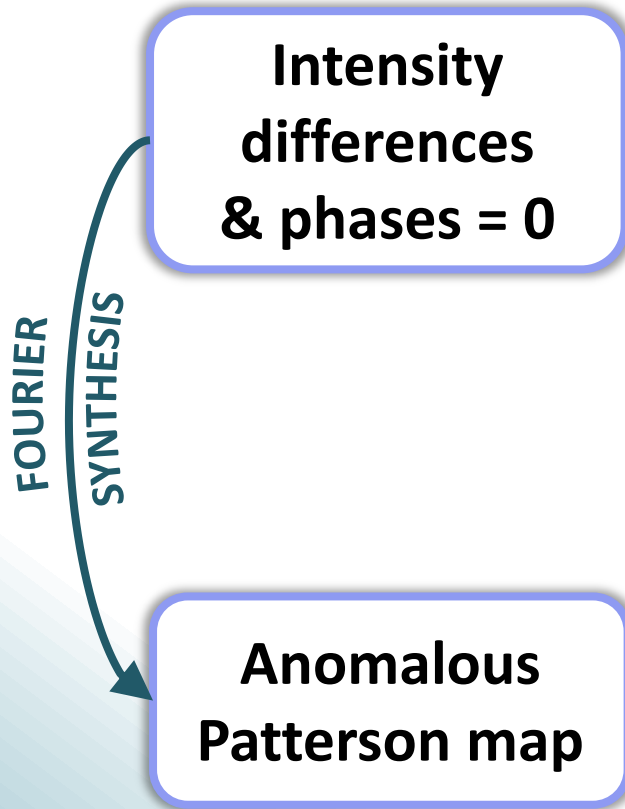


The intensities of Friedel pairs no longer have the same intensity!

This can be used for the absolute structure determination and for experimental phasing!



Anomalous Patterson map



A Patterson map calculated from the anomalous differences only relates to vectors between anomalously scattering atoms:

Anomalous Patterson map

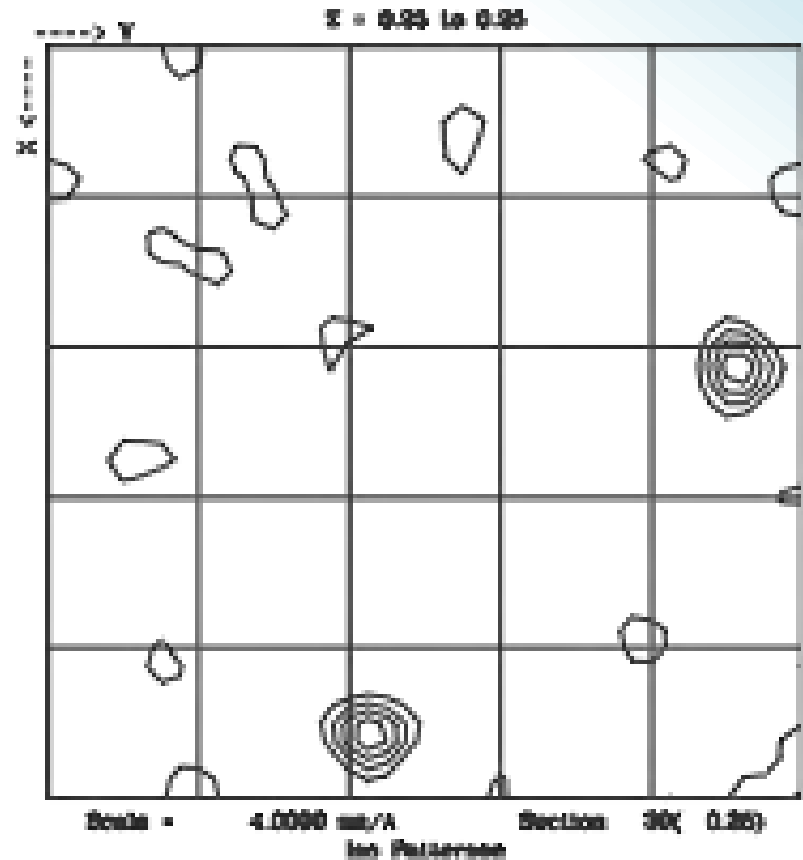
Even at low resolution, atoms can now be differentiated.



Patterson maps



ANOMALOUS



ISOMORPHOUS

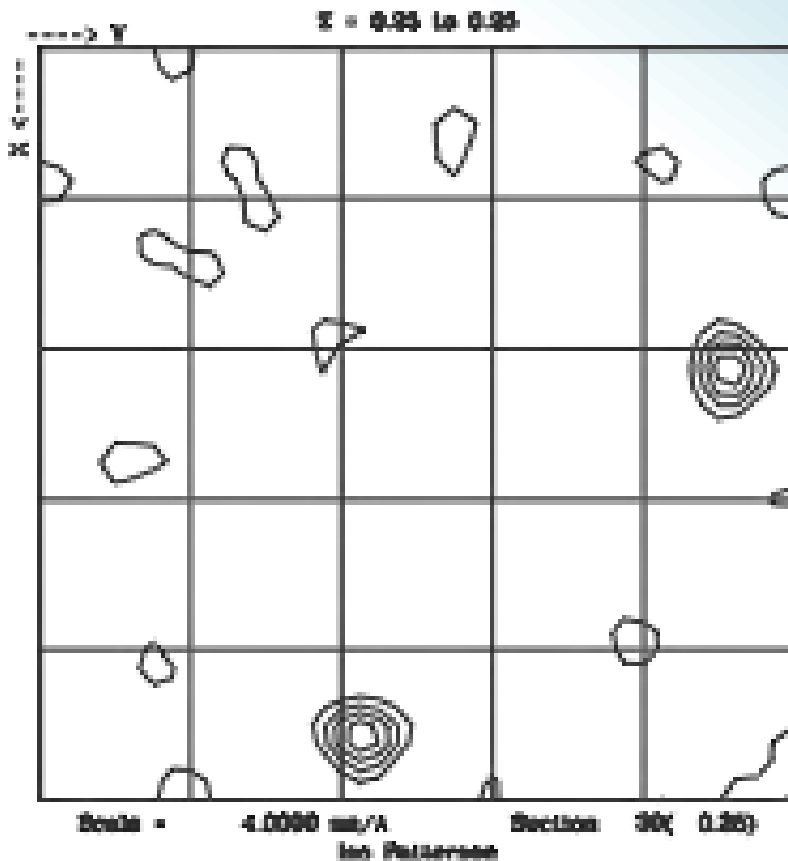


Isomorphous replacement

SIR, MIR and RIP:
Intensities differ by atoms
missing/added.

These differences can be
used for an 'isomorphous'
Patterson

Most heavy atoms do also
scatter anomalously at
common wavelengths.



How to...

SUBSTRUCTURE SEARCH IN SHELXD

Direct methods

- Phases of strong reflections are related (as a result of the non-random distribution of atoms.)
 - Triplett equations
 - Sayre equation
- Relations are relatively easy to resolve for few atoms.
- Usage of normalized structure factors (E values):

$$|E_{hkl}|^2 = \frac{|F_{hkl}|^2 / \epsilon}{\langle |F_{hkl}|^2 / \epsilon \rangle}$$

ϵ scale factor for proper treatment of special position reflections

$$\langle |F_{hkl}|^2 / \epsilon \rangle$$

mean per resolution shell



Substructure search

Finding the substructure of marker atoms

- Direct methods
 - Patterson methods
- Borrowed from small molecule crystallography
- These methods require separate atomic electron densities to locate atoms.
 - They work here because the marker atoms have large interatomic distances.
 - Disulfides become ,supersulfurs‘.



Substructure search

- **Patterson seeding** means starting the search with atoms consistent with the anomalous/isomorphous Patterson maps.
- **Dual space direct methods** recycle and modify trial substructures by peak search in the electron density and refining phases in reciprocal space. Convergence is faster than in reciprocal space alone.



Substructure search

An **overdetermined** problem with **noisy** data...

Critical factors in substructure search:

- Resolution range highly affects the outcome
- Good data quality
- Intensity outliers are problematic
- Scaling (also anisotropic scaling) is needed

BEWARE: Handedness is not resolved at this stage!
(Density modification differentiates later.)

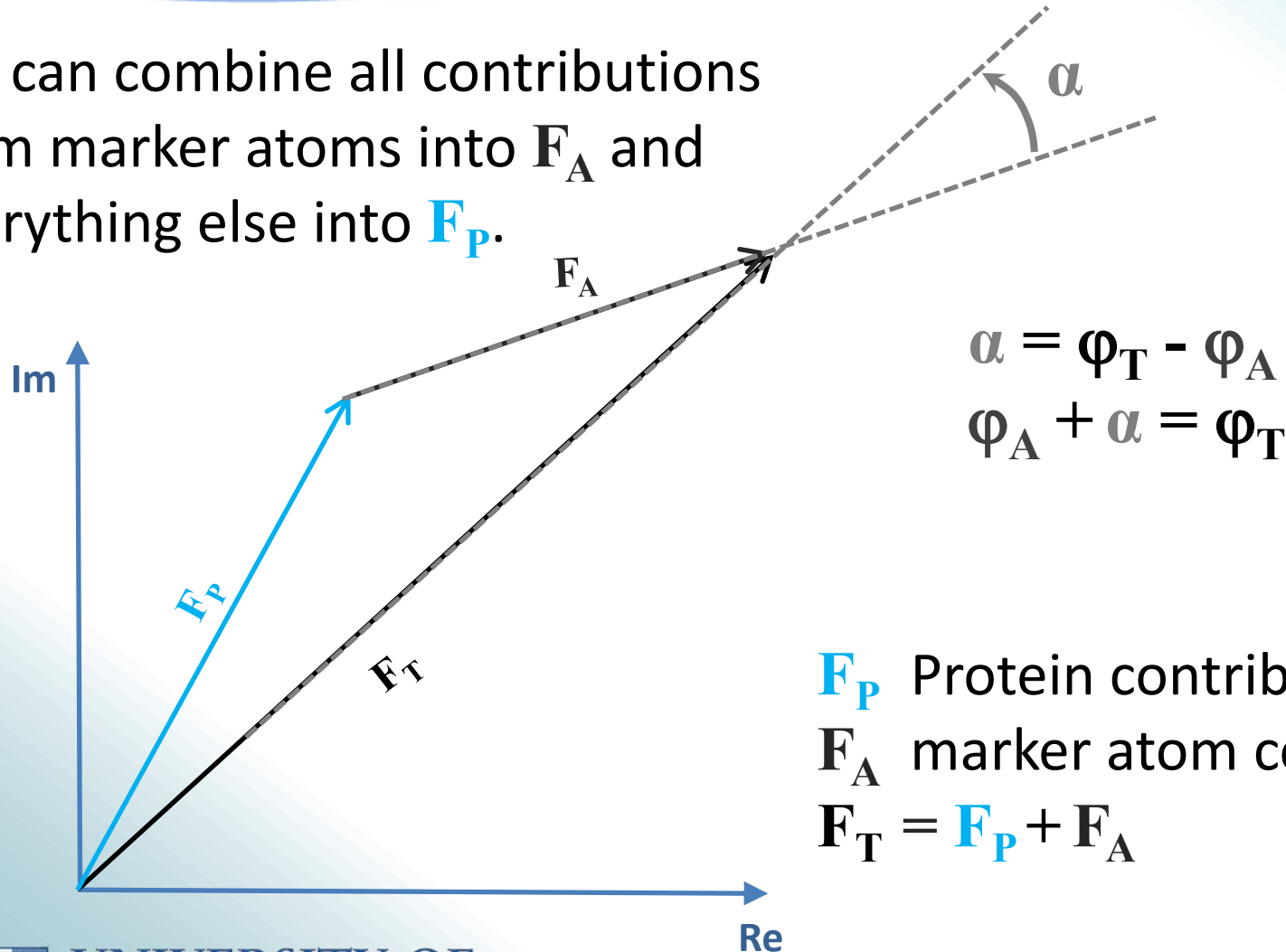


How to...

PHASING THE REST (SHELXC)

From substructure to structure

We can combine all contributions from marker atoms into \mathbf{F}_A and everything else into \mathbf{F}_P .



\mathbf{F}_P Protein contribution
 \mathbf{F}_A marker atom contribution
 $\mathbf{F}_T = \mathbf{F}_P + \mathbf{F}_A$



From substructure to structure

So, if we would know the anomalous scatterer positions (or heavy atom positions), we could calculate F_A :

$$\alpha = \varphi_T - \varphi_A$$

$$\varphi_A + \alpha = \varphi_T$$

If we could then get α , we could calculate φ_T and **solve the phase problem!**



From substructure to structure

Phasing equations

If we would have no errors...

$$|F_{hkl}|^2 = |F_T|^2 + a |F_A|^2 + b |F_T| |F_A| \cos\alpha + c |F_T| |F_A| \sin\alpha$$

$$|F_{-h-k-l}|^2 = |F_T|^2 + a |F_A|^2 + b |F_T| |F_A| \cos\alpha - c |F_T| |F_A| \sin\alpha$$

$$a = \frac{f''^2 + f'^2}{f_0^2} \quad b = \frac{2f'}{f_0} \quad c = \frac{2f''}{f_0}$$

F_T **Total structure factor**

F_A **Marker substructure structure factor**

$$\alpha = \phi_T - \phi_A$$



From substructure to structure

Phasing equations

$$|F_{hkl}|^2 = |F_T|^2 + a |F_A|^2 + b |F_T| |F_A| \cos\alpha + c |F_T| |F_A| \sin\alpha$$

$$|F_{-h-k-l}|^2 = |F_T|^2 + a |F_A|^2 + b |F_T| |F_A| \cos\alpha - c |F_T| |F_A| \sin\alpha$$

For each wavelength, we have different **a**, **b**, **c** and two observations. $|F_A|$, $|F_T|$ and α are unknown. So given good data from at least two wavelengths, the equation can be solved. This would be **MAD** then, and works best if the f' differences and the sum of f'' values would be large!



From substructure to structure

Phasing equations

$$|F_{hkl}|^2 = |F_T|^2 + a|F_A|^2 + b|F_T||F_A|\cos\alpha + c|F_T||F_A|\sin\alpha$$

$$|F_{-h-k-l}|^2 = |F_T|^2 + a|F_A|^2 + b|F_T||F_A|\cos\alpha - c|F_T||F_A|\sin\alpha$$

In a **SAD** experiment, we have only two observables, as we measured only one wavelength. So we assume

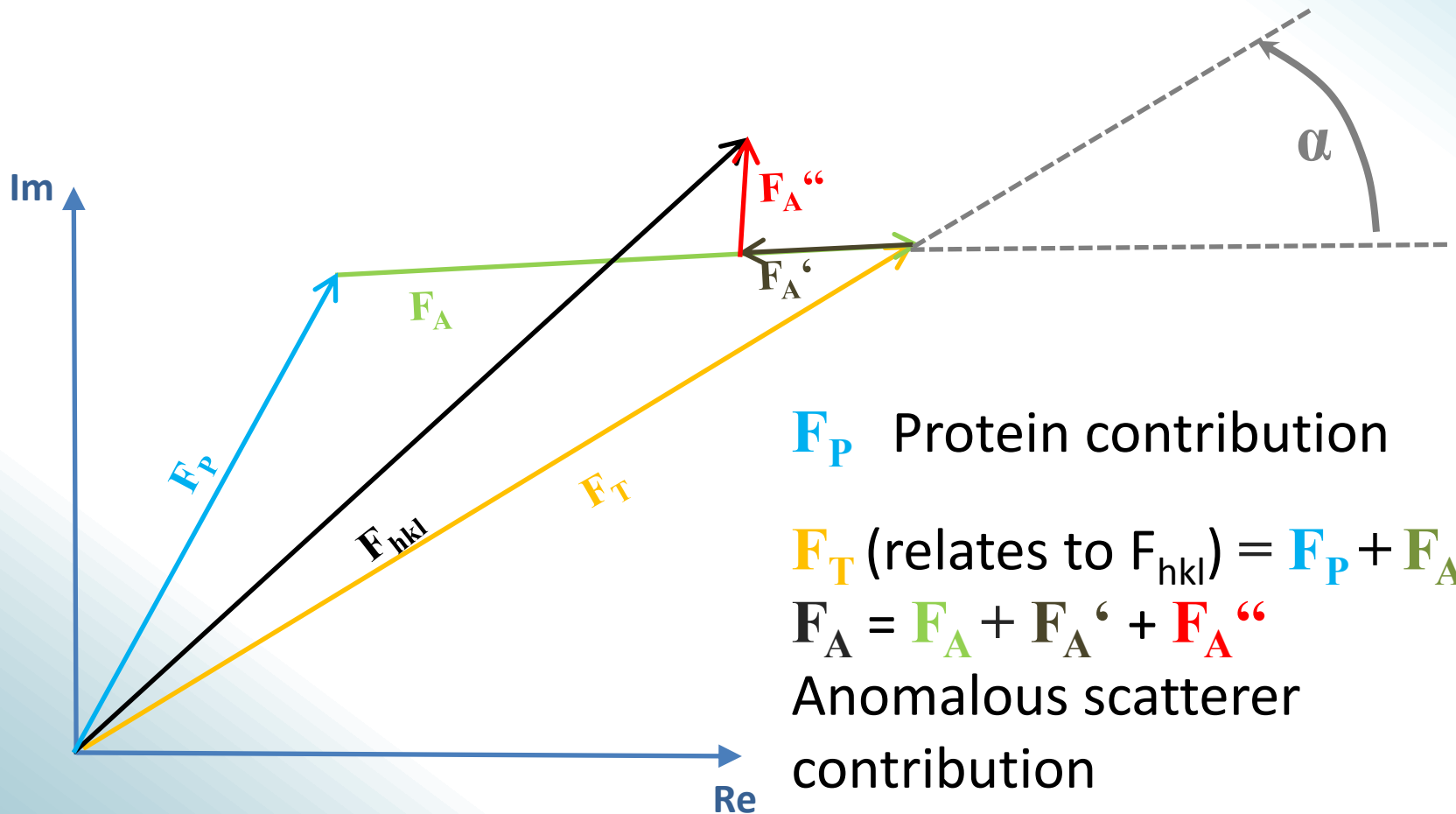
$$|F_T| = 0.5 (|F_{hkl}| + |F_{-h-k-l}|) \text{ and get}$$

$$|F_{hkl}| - |F_{-h-k-l}| = c|F_A|\sin\alpha$$

This is sufficient for the substructure and estimation of φ_T !



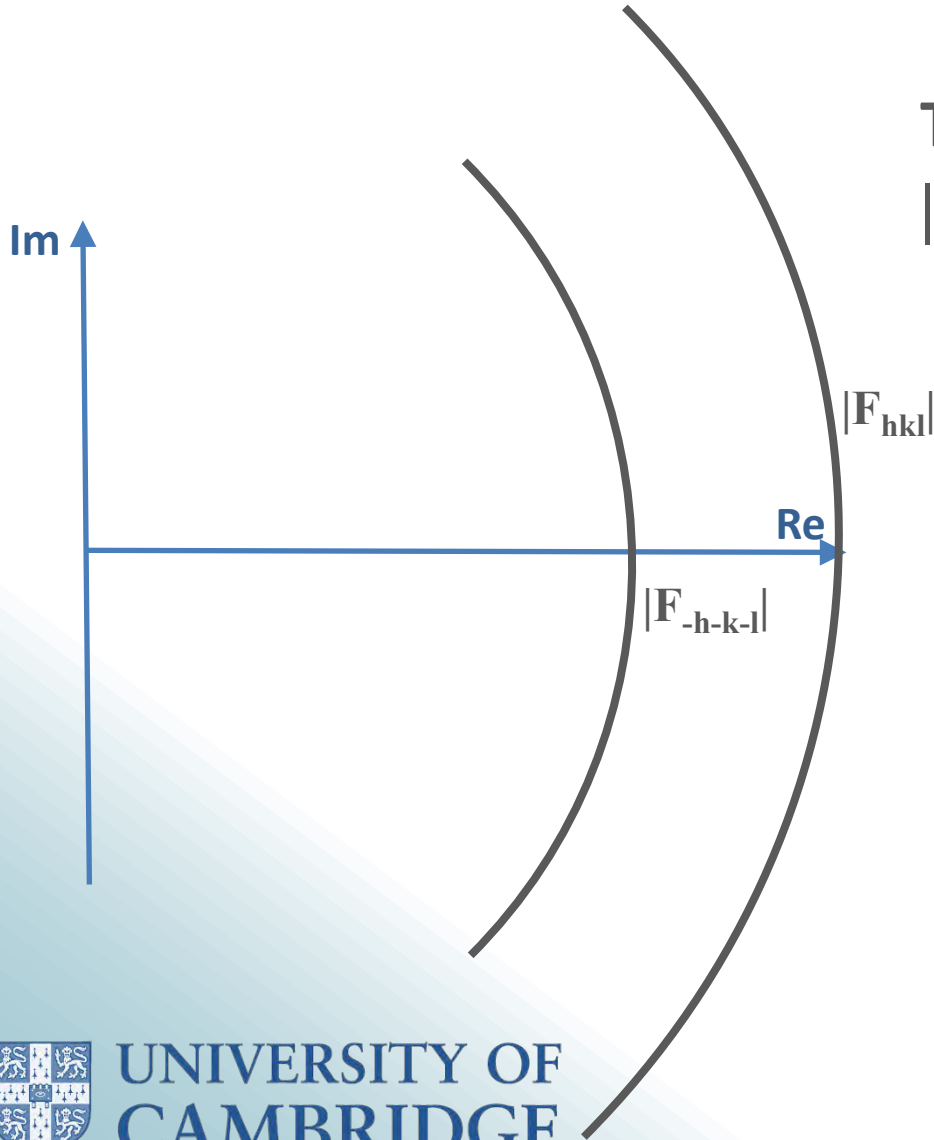
From substructure to structure



From substructure to structure

This is what we know:

$|F_{hkl}|$ and $|F_{-h-k-l}|$

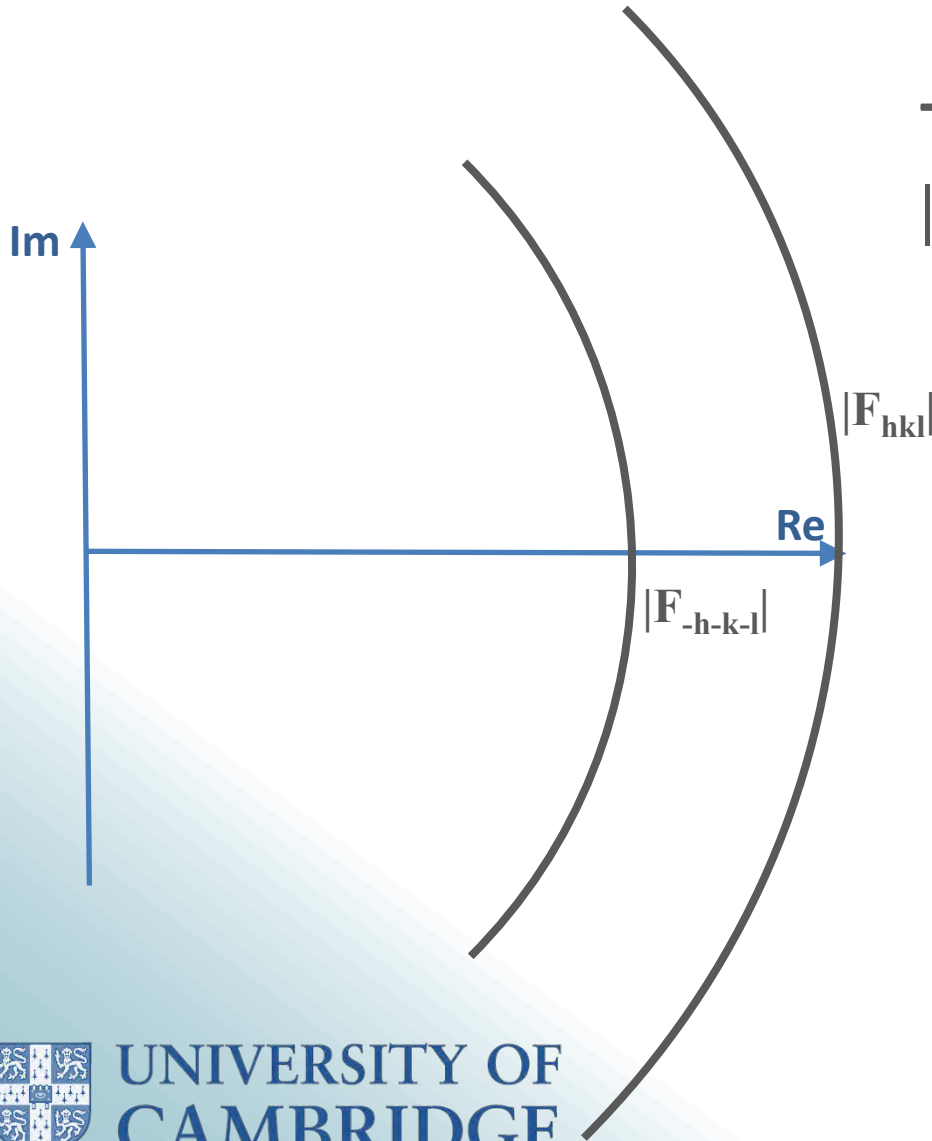


From substructure to structure

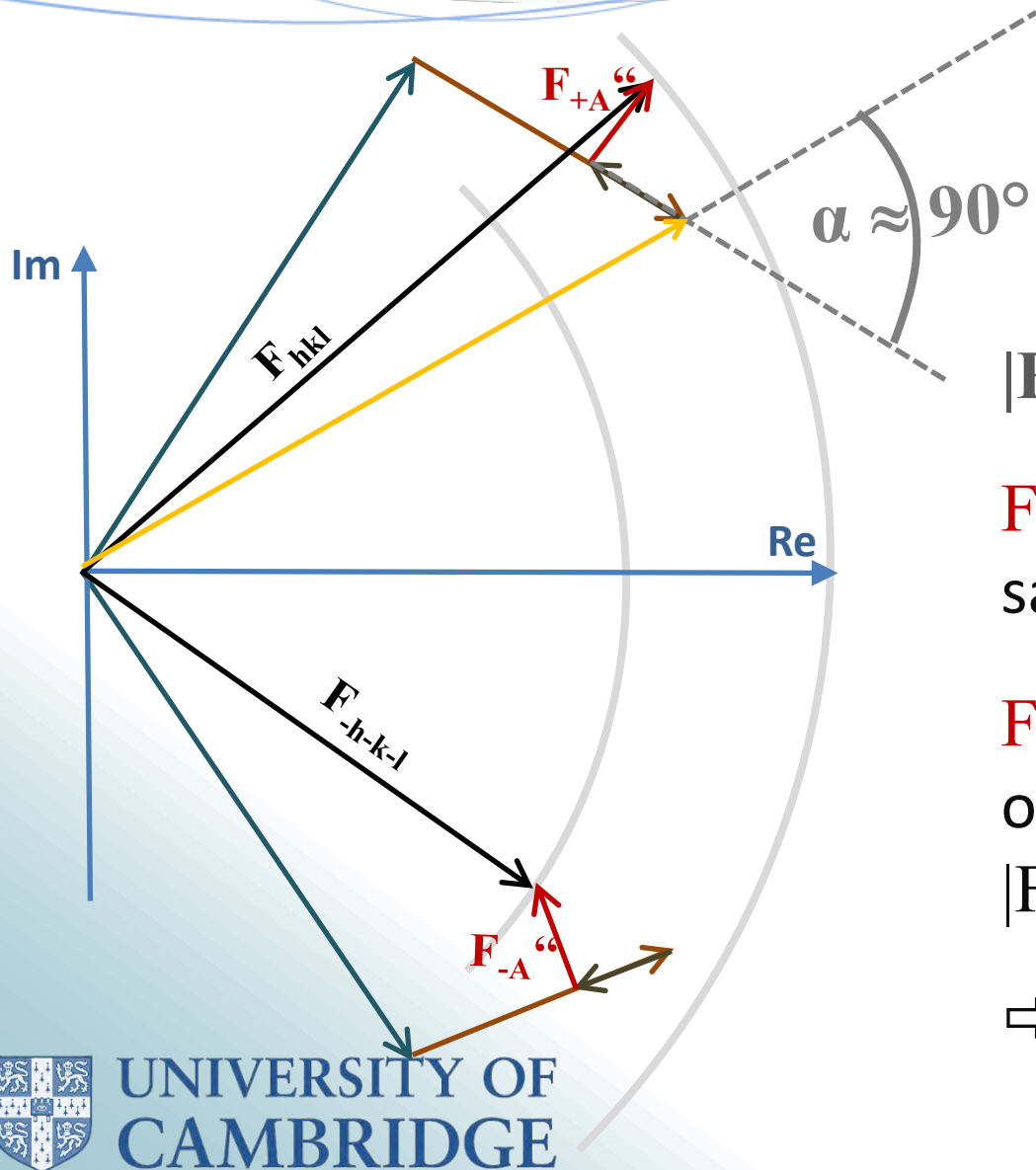
This is what we know:

$$|F_{hkl}| \text{ and } |F_{-h-k-l}|$$

$$|F_{hkl}| \gg |F_{-h-k-l}|$$



From substructure to structure



$$|\mathbf{F}_{hkl}| \gg |\mathbf{F}_{-h-k-l}|$$

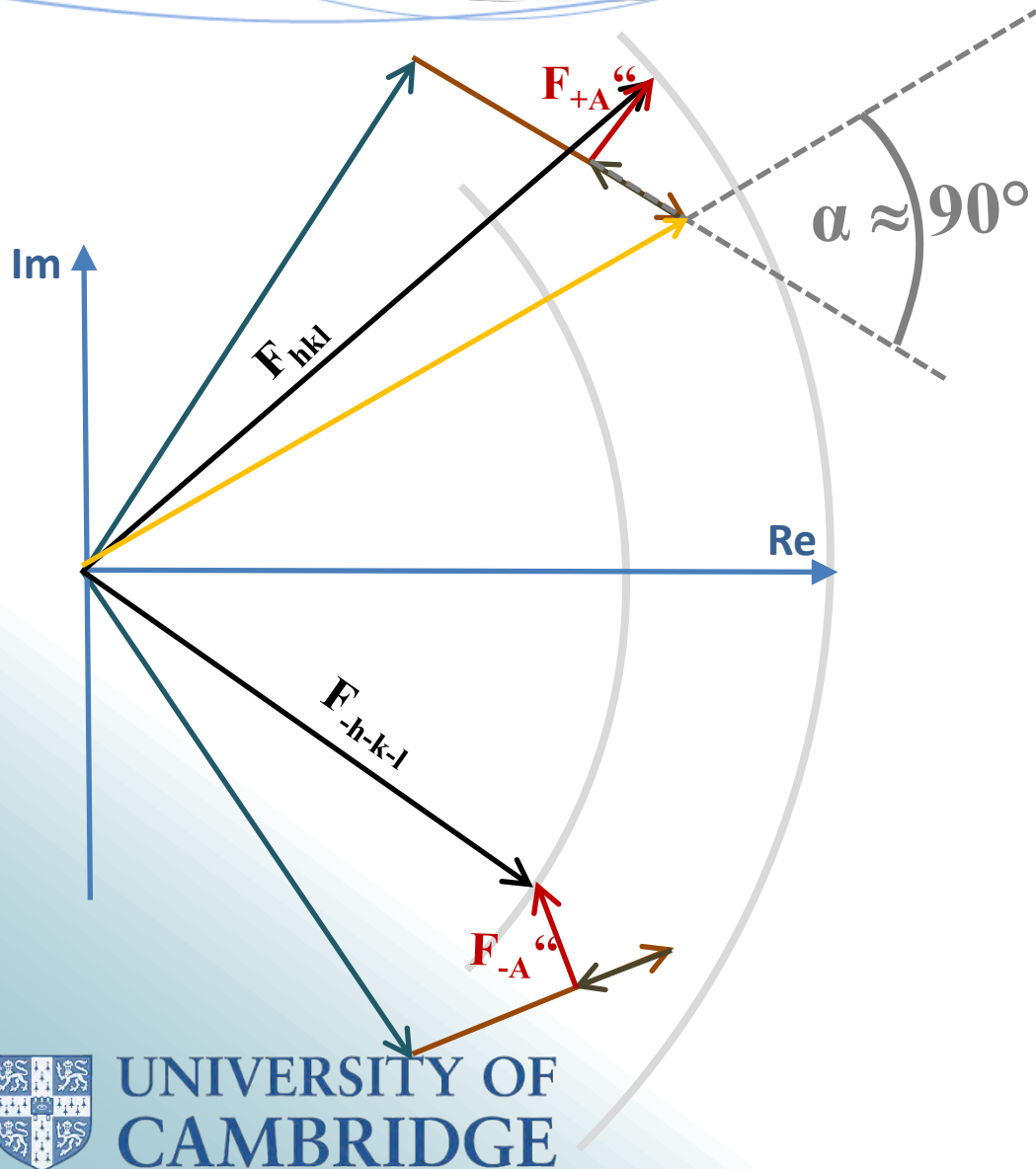
\mathbf{F}_{+A} “ has to point in the same direction as $|\mathbf{F}_{hkl}|$

\mathbf{F}_{-A} “ has to point in the opposite direction as $|\mathbf{F}_{-h-k-l}|$

$\Rightarrow \alpha$ must be close to 90° !



From substructure to structure



$$\text{If: } |F_{hkl}| \ll |F_{-h-k-l}|$$

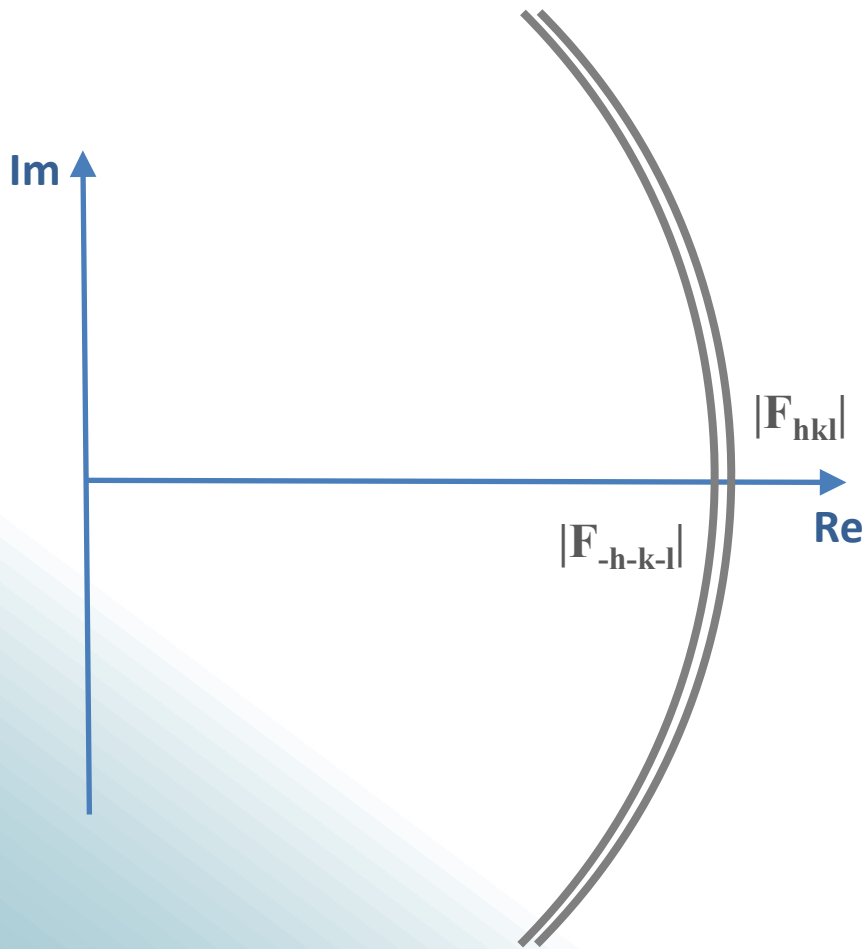
$\Rightarrow \alpha$ must be close to 270° !

Reflections with the largest anomalous differences must be closest to $\alpha = 90^\circ$ or $\alpha = 270^\circ$.

As you can easily see, estimation is rough.



From substructure to structure



$$|F_{hkl}| \approx |F_{-h-k-l}|$$

F_{+A} “ and F_{-A} “ must be very small or almost perpendicular to F_{hkl} or F_{-h-k-l} , respectively.

$\Rightarrow \alpha$ must be close to 0° or 180°



Density modification

- φ_T can now be computed from the phasing equations!

$$\varphi_A + \alpha = \varphi_T$$

Via Fourier synthesis, an initial map is gained.

- By σ_A coefficients and Sim weights the map is improved.
- But most important: **Density modification** is applied.



How to...

DENSITY MODIFICATION IN SHELXE

Density modification

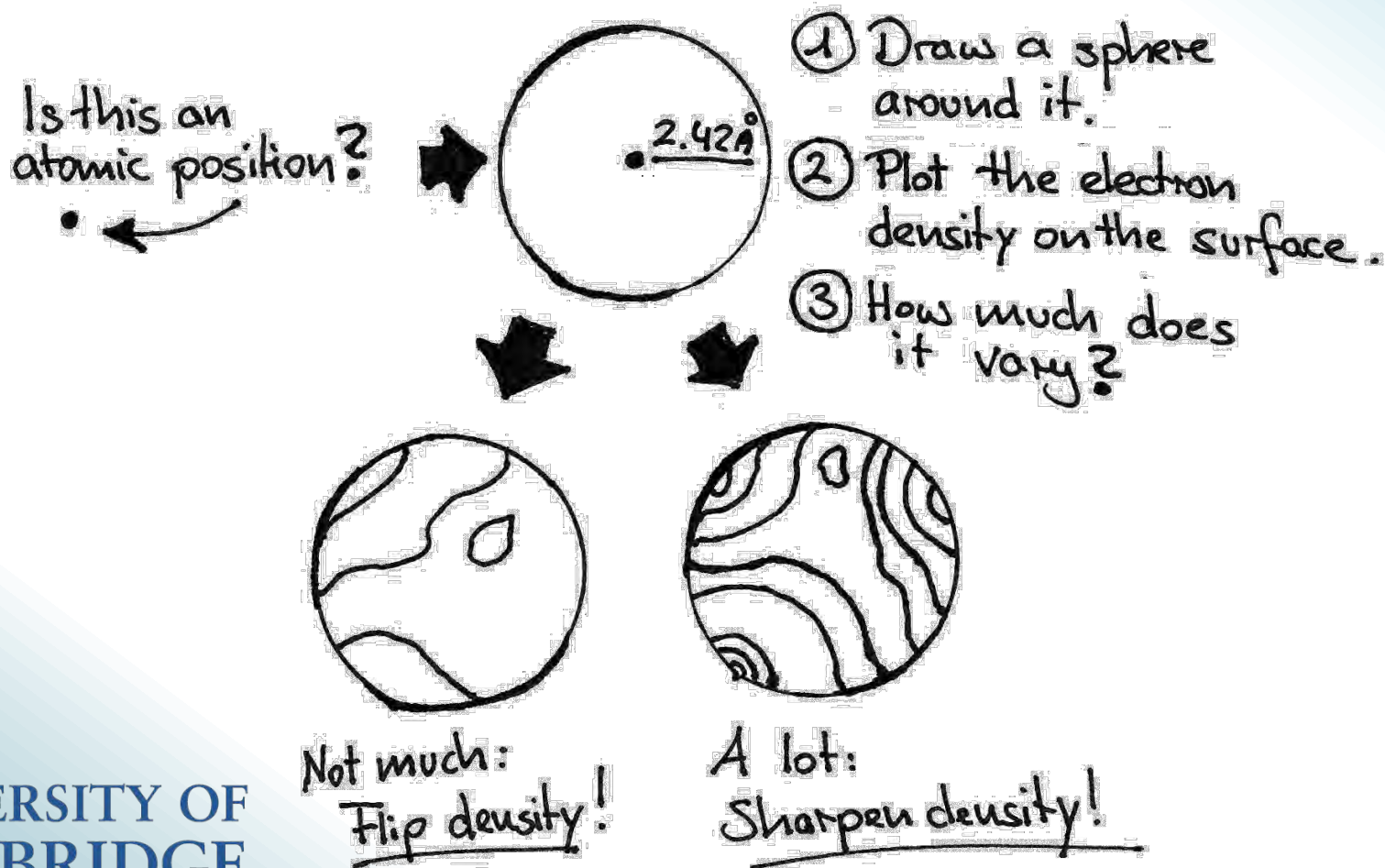
Especially SAD phases are still ambiguous as well as inaccurate. **Density modification** dramatically improves initial phases, electron density and **resolves handedness!**

- Based on areas filled by disordered solvent
- Solvent area is flattened or flipped
- NCS averaging can improve map quality
- High solvent content gives often better improvement



DENSITY MODIFICATION

Most programs use a mask. SHELXE uses the sphere-of-influence method for density modification:



Density modification

After several cycles, one of the two maps (one for each substructure enantiomer) looks ,like protein‘.

The other has less connectivity and looks ,ragged‘.

Example: Elastase

After density modification, the structure is solved!
Experimental phasing has led to initial phases.



Experimental phasing, for real

PRACTICALITIES

Heavy atoms

The main constituents of organic matter

H																			He
Li	Be											B	C	N	O	F			Ne
Na	Mg											Al	Si	P	S	Cl			Ar
K	Ca	Sc	Ti	V	Cr	Mn	Fe	Co	Ni	Cu	Zn	Ga	Ge	As	Se	Br			Kr
Rb	Sr	Y	Zr	Nb	Mo	Tc	Ru	Rh	Pd	Ag	Cd	In	Sn	Sb	Te	I			Xe
Cs	Ba	La	Hf	Ta	W	Re	Os	Ir	Pt	Au	Hg	Tl	Pb	Bi	Po	At			Rn
Fr	Ra	Ac																	
		Th	Pa	U															

Classic heavy-atoms

Gaseous inert heavy-atoms



(Significant) anomalous scatterers

The main constituents of organic matter are not anomalous scatterers

Weak anomalous scatterers at long wavelength

Useful anomalous scatterers @ K absorption edges

Selenomethionine

H																			He
Li	Be											B	C	N	O	F			Ne
Na	Mg											Al	Si	P	S	Cl			Ar
K	Ca	Sc	Ti	V	Cr	Mn	Fe	Co	Ni	Cu	Zn	Ga	Ge	As	Se	Br			Kr
Rb	Sr	Y	Zr	Nb	Mo	Tc	Ru	Rh	Pd	Ag	Cd	In	Sn	Sb	Te	I			Xe
Cs	Ba	La	Hf	Ta	W	Re	Os	Ir	Pt	Au	Hg	Tl	Pb	Bi	Po	At			Rn
Fr	Ra	Ac																	
		Th	Pa	U															

Useful anomalous scatterers at long wavelength

Gaseous inert heavy-atoms

Classic heavy-atoms – isomorphous signal & useful anomalous scattering @ L absorption edges



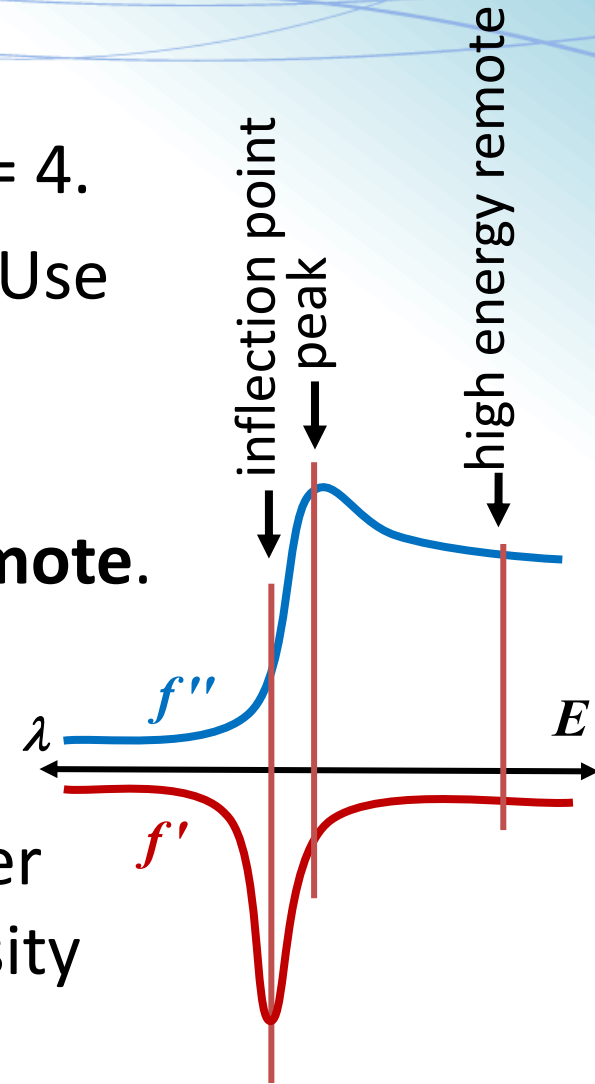
Data collection

- High multiplicity is good.
- Radiation damage is often bad.
- Precise intensity measurements are good.
- Near to the absorption edge, the crystal absorbs most energy, therefore radiation damage is high.
- A **fluorescence scan** can prove the presence of anomalous scatterers in the crystal.



Data collection: MAD

- Collect **peak** with at least multiplicity = 4.
- Radiation damage? Stop and try SAD! Use a second crystal to collect high energy remote.
- No damage? Measure **high energy remote**.
- Last data set should be **inflection** – so f' is maximized.
- A **higher resolution data set** with lower redundancy may prove useful for density modification and for refinement.



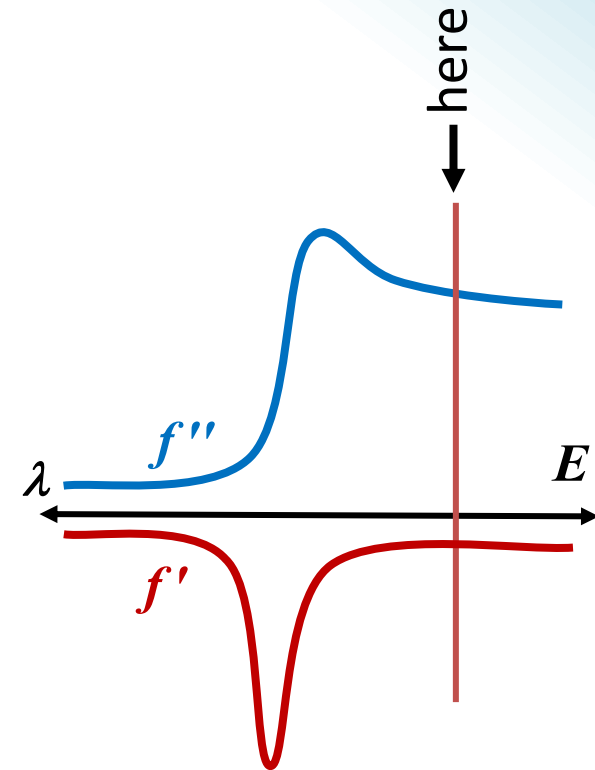
$$f = f_0 + f' + i f''$$



Data collection: SAD

- Best wavelength right and not too close to peak
- Beware not to hit the „white line“ near the peak
- A bit away from peak, radiation damage will be less

SAD data measured at peak are often the result of a MAD experiment attempt!



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



Data evaluation

- The general data quality should be good – multiplicity, completeness, R_{PIM} etc.
- If scaling was applied, check statistics.
- Check the mask; inner shell completeness?
- Data set files well distinguishable?
- If you have made a fluorescence scan, keep it.
- Is there an anomalous signal in the collected data?
 - Anomalous correlation within a data set: $CC_{anom(1/2)}$
 - $\langle d''/\sigma \rangle$ and/or $\langle d'/\sigma \rangle$
 - Anomalous correlation of data sets: CC_{anom}



Things you want to have an idea about

- Space group? (Twinning?)
- How many marker atoms do you expect?
- Substructure: Which elements/molecules?
- What could be the best resolution cut-off?
(SHELXC assumes data resolution + 0.5Å)
- Could any marker atoms ,fuse' into bigger blobs of density because of resolution cut-off? Disulfides?
- Merging of data from different crystals/runs?
- Expected solvent content and residue numbers?



If you use SHELX...

SHELXC: α calculation, data analysis,
file preparation

SHELXD: Substructure search

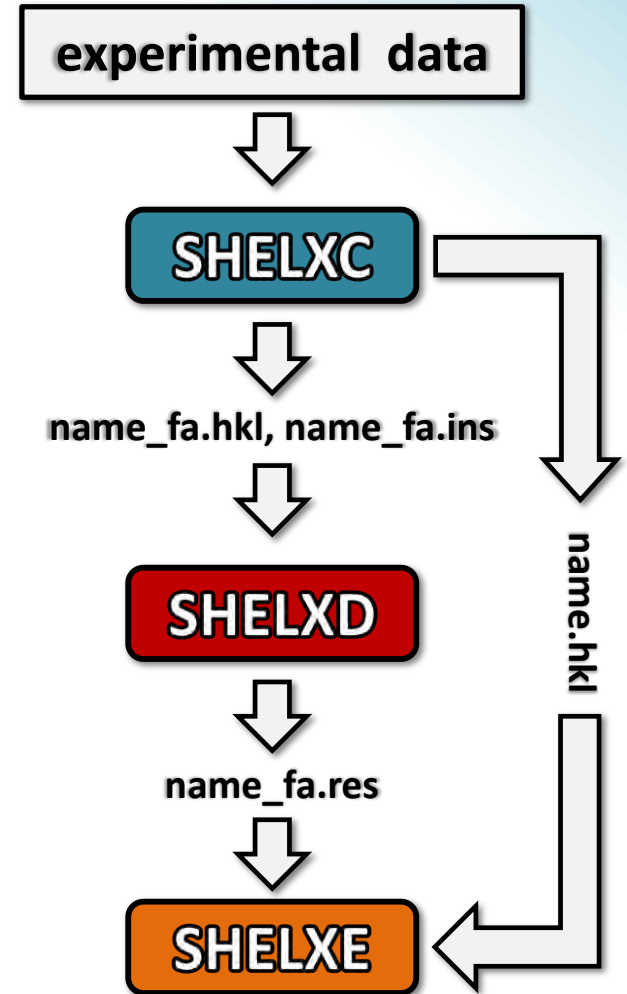
SHELXE: Density modification, tracing*

* *A traced structure is solved; CC (trace against native data) > 25% (for data < 2.5 Å)*

[**ANODE:** Validation]

Pipeline?

Other experimental phasing programs should be considered, in particular for ease of use or problem cases**.



Wait a moment...

ANOMALOUS MAPS

Wait a moment...

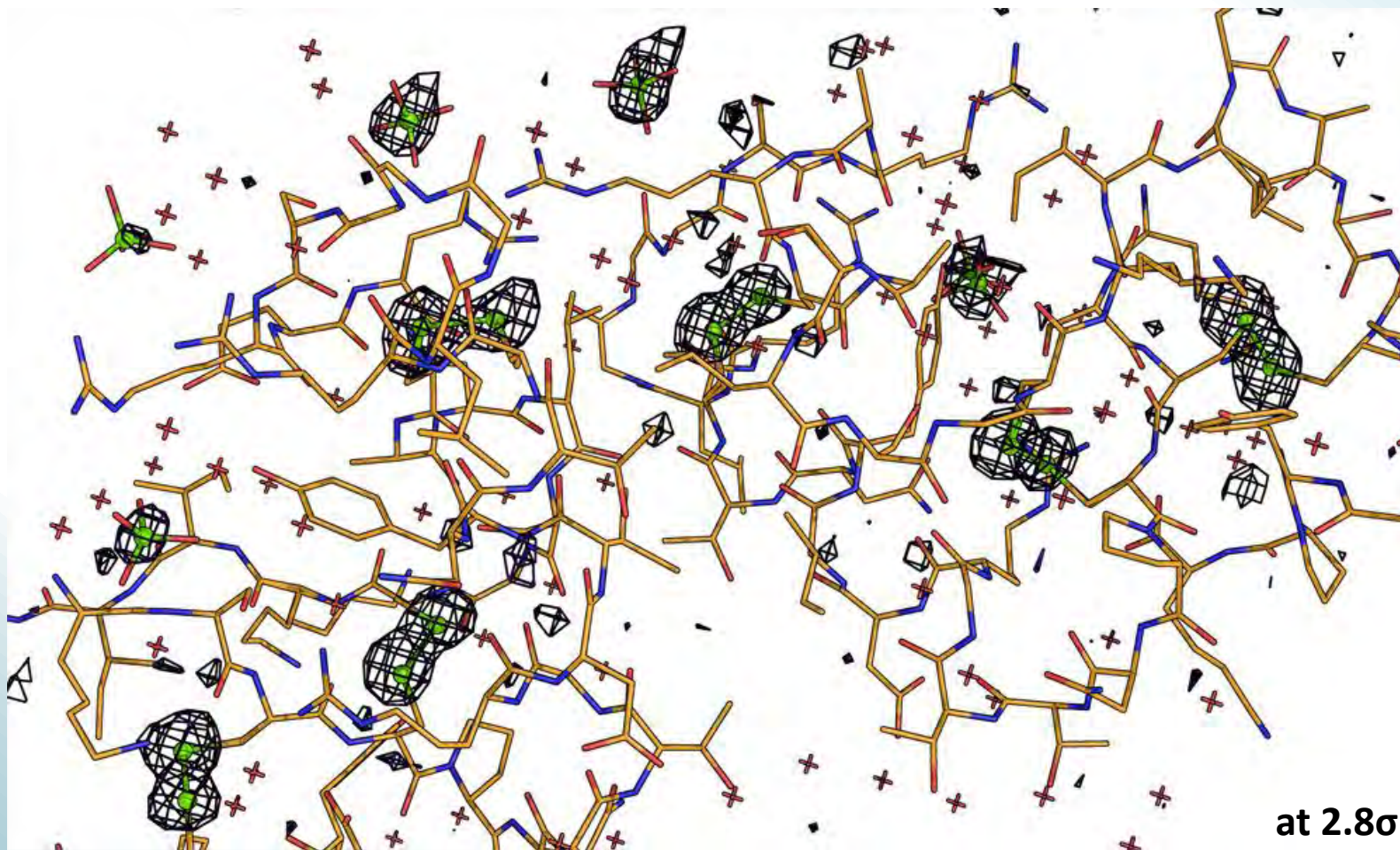
If I can use the anomalous signal for a Patterson, can I also calculate a map from the anomalous signal?

Yes! And it can be used after phasing for atom type identification, radiation damage assessment *et cetera*.

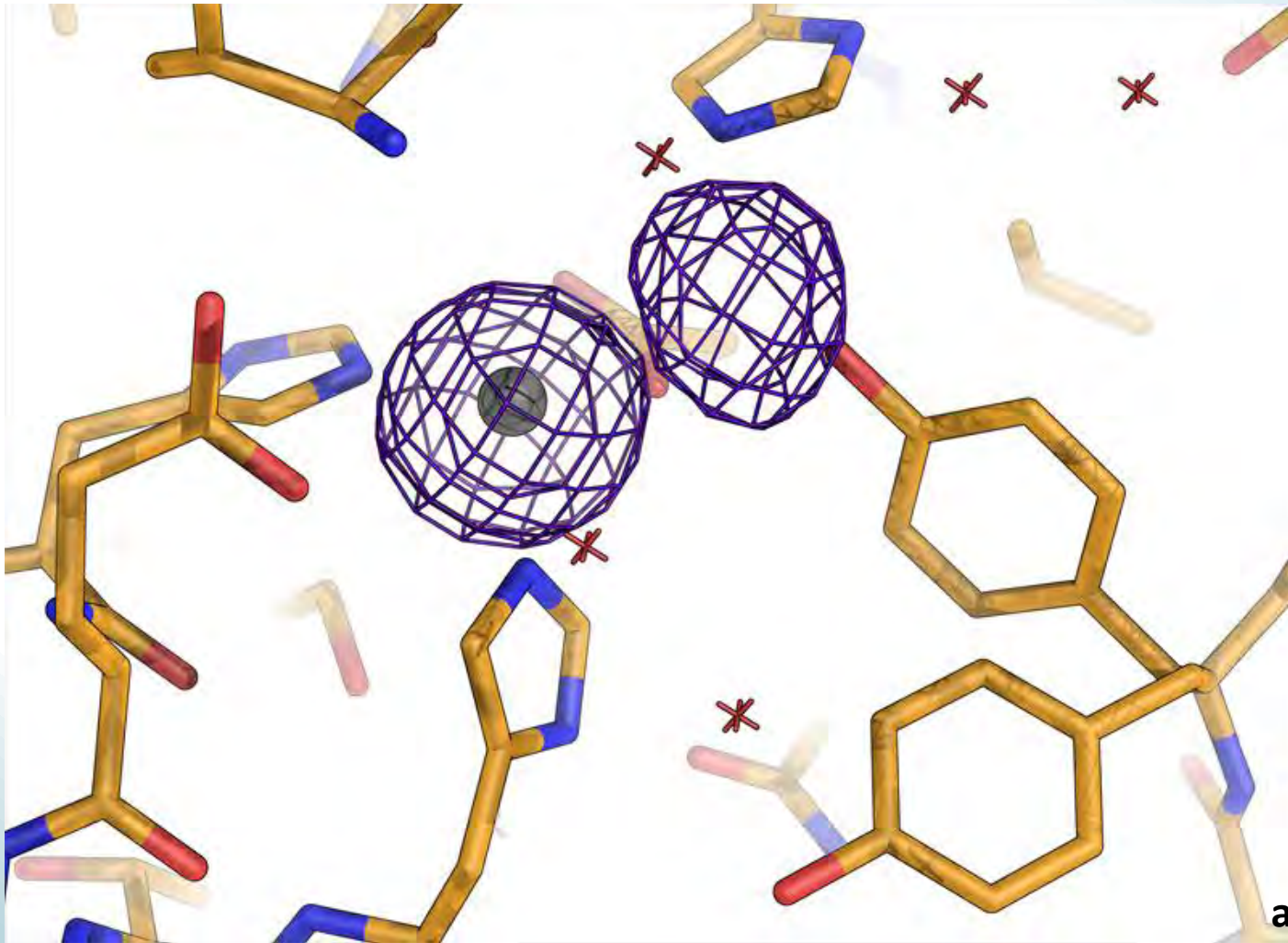
One possibility to get such a map quickly from SHELX type input is the program **ANODE**.



Anomalous density maps



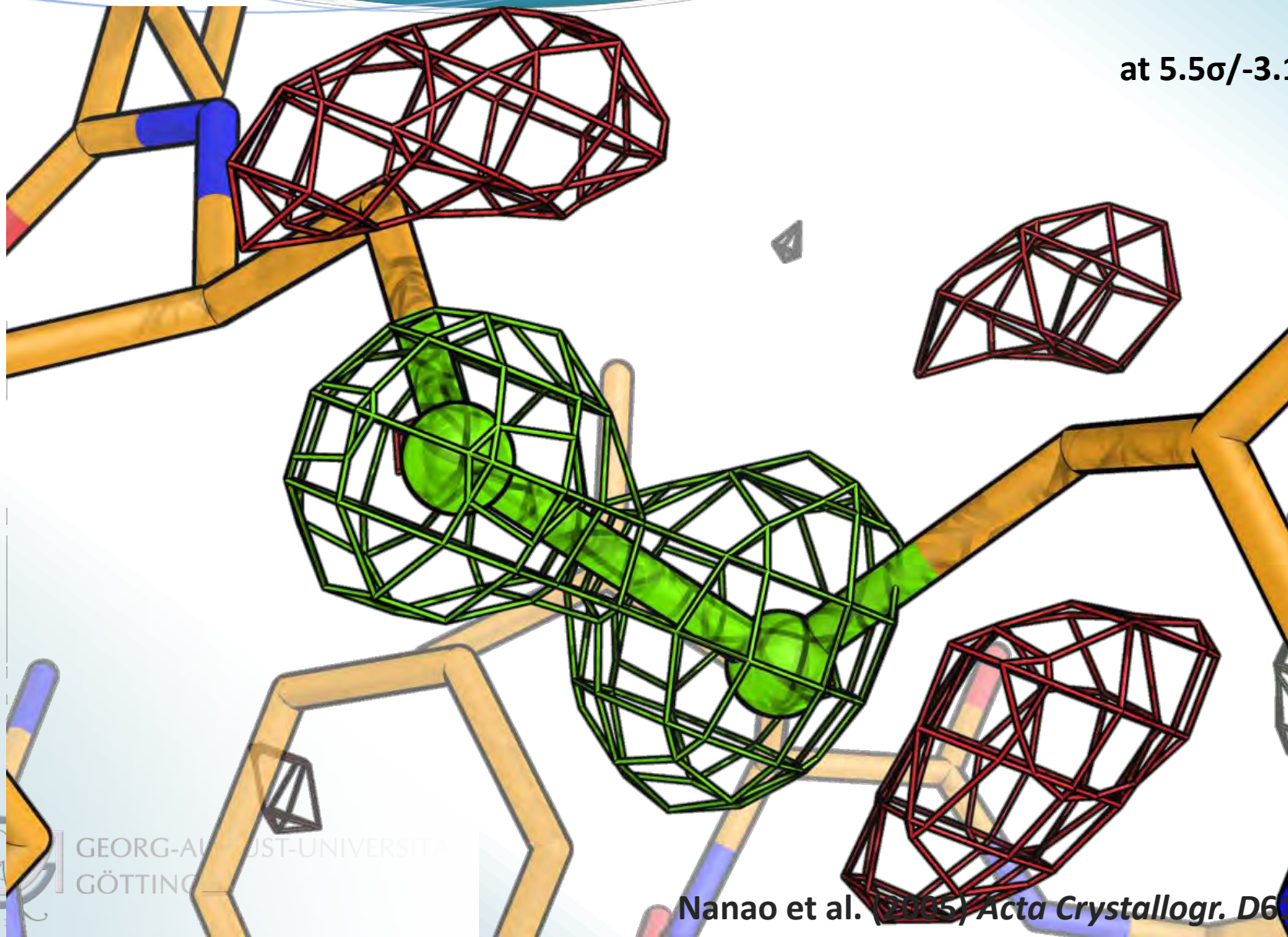
Anomalous density maps



at 3.5 σ

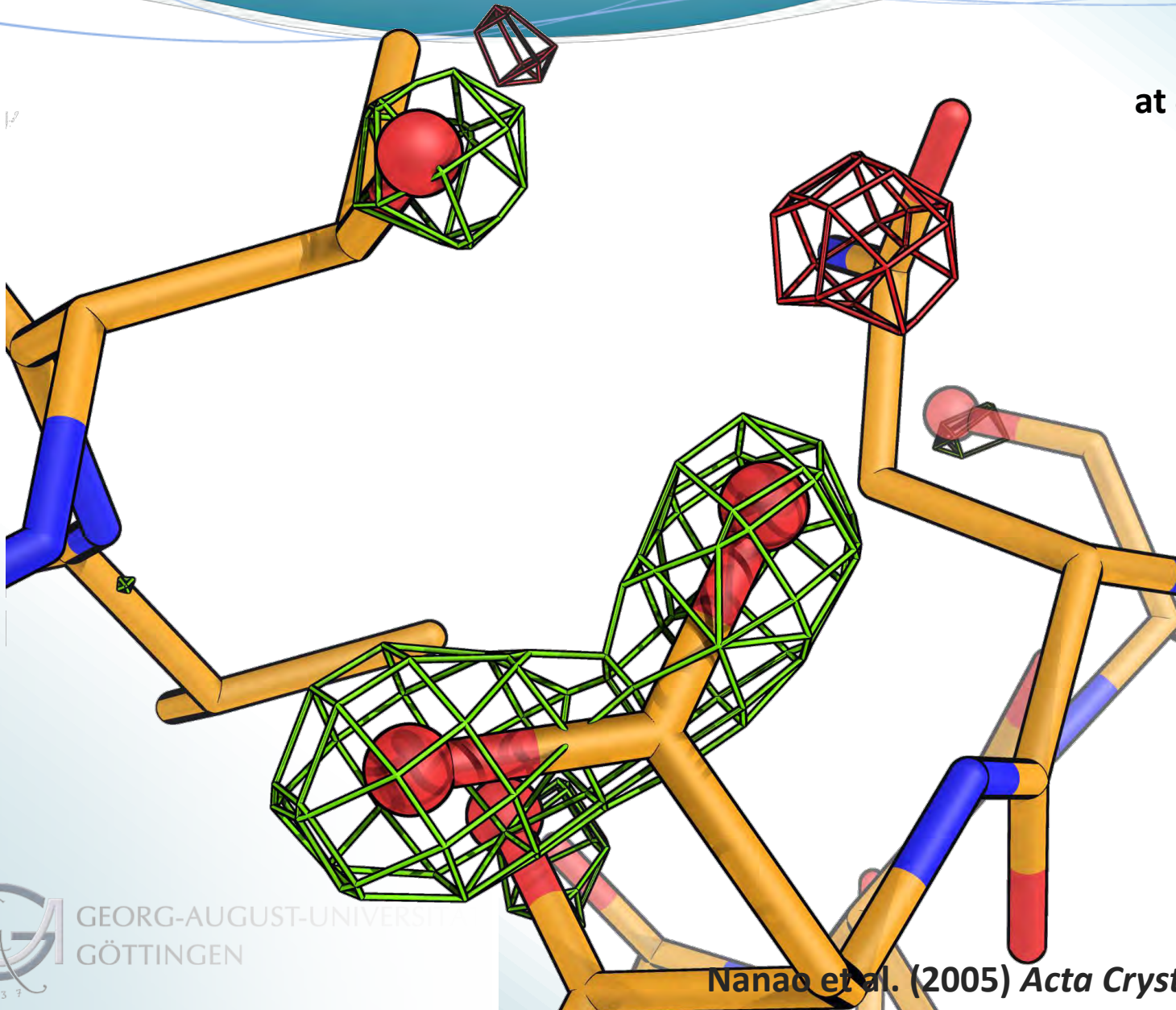
RIP density maps

at $5.5\sigma/-3.1\sigma$



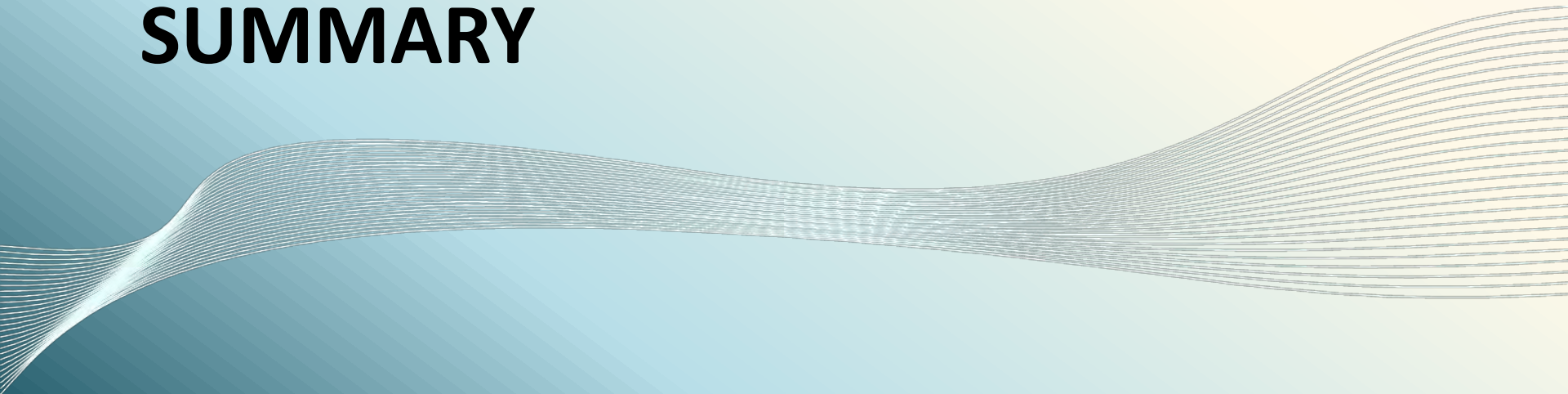
RIP density maps

at 4.8σ / -3.1σ



Final

SUMMARY



SUMMARY

- Experimental phasing methods use marker substructures of certain elements to solve the phase problem via the phasing equations. Patterson maps can help.
- **MAD** and **SAD** exploit the anomalous signal from one or more data sets from the same crystal.
- **SIR** and **MIR** utilizes several heavy-atom soaked derivative crystals. They have to be isomorphous to be utilized.
- Experimental phase solutions **do not define the enantiomorph**; after solution, the map that looks like protein has to be chosen!



ACKNOWLEDGEMENTS

I am grateful to **George Sheldrick, Phil Evans, Randy Read and Airlie McCoy** who freely shared with me their knowledge, ideas and material. I also want to thank the Read/Deane lab for their support!

GET TO KNOW:
YOUR ADVISOR'S
NEGATION FIELD

A Professor's Negation Field is the unexplained phenomenon whereby mere spatial proximity to an experimental set-up causes all working demonstrations to fail, despite the apparent laws of Physics or how many times it worked right before he/she walked into the room.

BEWARE ALSO:

The Sphere of Death.
Allowing your experiment within arm's reach of your Advisor risks the possibility of immediate destruction.



WWW.PHDCOMICS.COM

JORGE CHAM © 2007



UNIVERSITY OF
CAMBRIDGE

LITERATURE

- Bernhard Rupp, **Biomolecular Crystallography: Principles, Practice, and Application to Structural Biology**, 2004
- Kai Diederichs, P. Andrew Karplus, **Improved R-factors for diffraction data analysis in macromolecular crystallography**. Nat Struct Biol. (1997). 4, 269-75.
- Manfred S. Weiss, **Global indicators of X-ray data quality**, J. Appl. Cryst. (2001). 34, 130-135



LITERATURE

- George M. Sheldrick, **A short history of SHELX**, Acta Cryst. (2008). A64, 112-122
- George M. Sheldrick (2002). **Macromolecular phasing with SHELXE**, *Z. Kristallogr.* 217:644-650.
- George M. Sheldrick, **Experimental phasing with SHELXC/D/E: combining chain tracing with density modification**, Acta Cryst. (2010). D66, 479-485
- A. Thorn & G.M. Sheldrick: **“ANODE: ANOmalous and heavy-atom DEnsity calculation”** *J. Appl. Cryst.* 44 (2011), 1285-1287

<http://shelx.uni-ac.gwdg.de/SHELX/>

