

Structure Determination from Powder Diffraction Data



Michael Evans

Bruker AXS, Karlsruhe, Germany



Structure Determination

The Phase Problem



- Simply put, the electron density of a crystal is represented in reciprocal space by a complete set of complex **structure factors, $F(\mathbf{h})$**

$$F(\mathbf{h}) = |F(\mathbf{h})| \exp[i\phi(\mathbf{h})]$$

- If we could measure these structure factors directly, structure solution would simply be an inverse Fourier transform...

$$\rho(x,y,z) = \sum_{hkl} \mathbf{F}_{hkl} \exp \{ -2\pi i (hx + ky + lz) \} / V$$

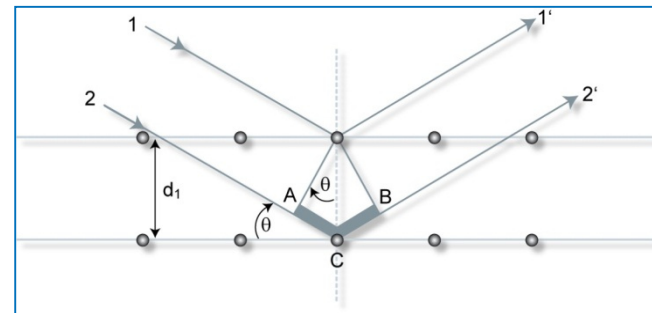
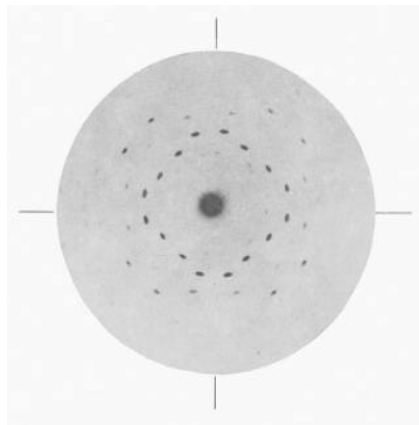
- The intensities from a diffraction experiment can give us $|F(\mathbf{h})|$, *but the phase information, $\phi(\mathbf{h})$, cannot be experimentally determined.*
- Structure solution is essentially the solution of the phase problem...

Structure Determination

Historical Background



- 1912: Friedrich, Knipping and von Laue's scattering experiment show that X-rays are scattered by crystals
- 1913: W.H. and W.L. Bragg introduce simple representation of specular reflection from lattice planes in a crystal



- In the years that followed, 'structure solution' was a trial-and-error technique based on crystal habit and symmetry, priory chemical knowledge, and intuition.

Structure Determination

Historical Background



- 1912: Friedrich, Knipping and von Laue's scattering experiment show that X-rays are scattered by crystals
- 1913: W.H. and W.L. Bragg introduce simple representation of specular reflection from lattice planes in a crystal
- 1950's: Harker and Kasper are among the first to show that combining *a priori* structural knowledge with diffracted intensities can, in fact, solve the phase problem
 - Non-negativity: the electron density function $\rho(\mathbf{r})$ in a crystal must be non-negative everywhere
 - Atomicity: molecules and compounds consist of atoms, thus $\rho(\mathbf{r})$ reaches maximum positive values at atom positions and drops to small values between atoms
- These restrictive conditions became the foundation of **direct methods**

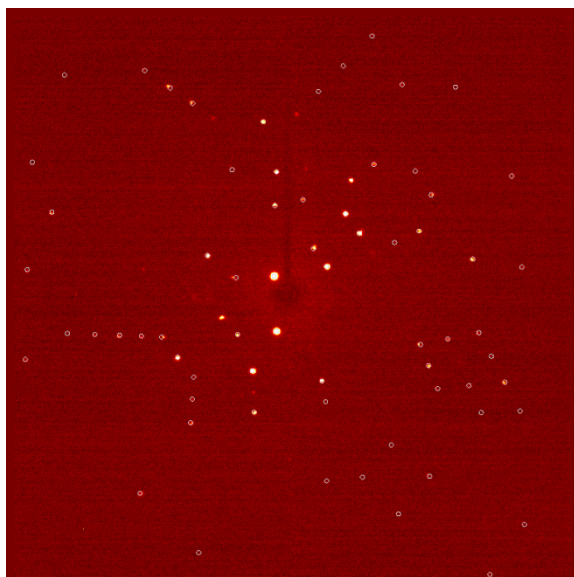
Structure Determination

Direct Methods

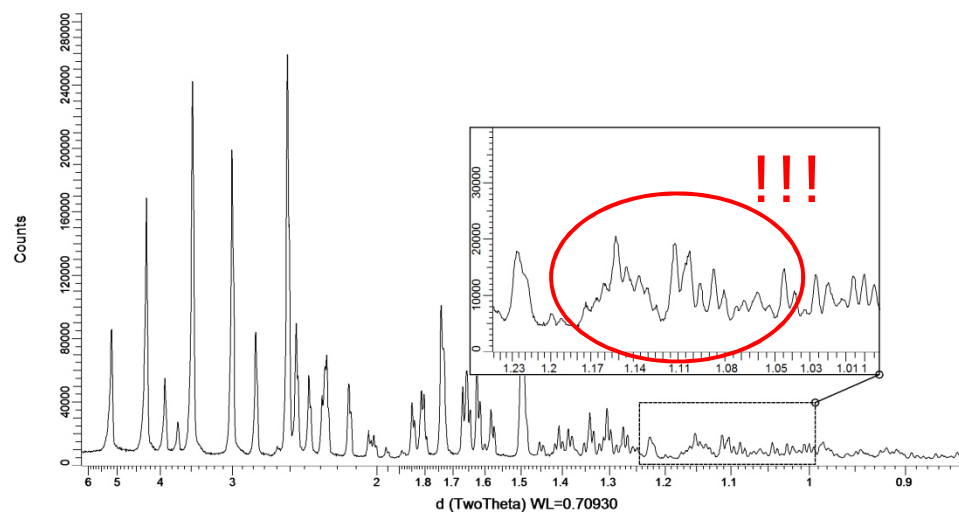


- Direct Methods are *ab initio* crystal structure techniques able to estimate phases $\varphi(\mathbf{h})$ directly from structure factor magnitudes $|F(\mathbf{h})|$
 - Requires data collected to atomic resolution ($\sim 1 \text{ \AA}$ or better)
 - Intensities for each individual peak are extracted

Single crystal diffraction
Well-separated peaks



Powder diffraction
Severe peak overlap



Structure Determination Powder Diffraction Challenges



- Why is structure determination more difficult with powders than single crystals?
 - Collapse of 3D space into a 1D diffraction pattern
 - Possible presence of preferred orientation effects
 - Peak broadening (anisotropic?) due to lattice defects
 - Peak overlap ('sharing' of intensities)
- Many modern materials (metal hydrides and battery materials, alloys, intermetallic compounds, pharmaceuticals) are available only as polycrystalline materials
- Direct methods work well for single crystals, but are not well suited for powders...

- Structure solution methods can be divided into two broad categories: **Direct methods** and **Real-space methods**

Direct Methods

- Works with structure factors (F^2) extracted from powder data
- Requires data measured to atomic resolution ($\sim 0.9 \text{ \AA}$ min)
- Peak overlap!
- Initially developed for single crystal and modified for powder

Examples include: Patterson methods, Fourier methods

Direct-Space Methods (aka "Global Optimization" or "Real-Space")

- Works with extracted structure factors (F^2) or with $y(\text{obs})$
- More suitable for molecules, where this structure is generally known beforehand
- An algorithm is used to search through parameter space

Examples include: Monte Carlo, Simulated Annealing, Genetic Algorithm

SDPD

Methods



- Or, the powder diffractionist's toolbox for SDPD – all available in TOPAS
- Simulated Annealing
 - A direct space approach where all parameters lie in direct space (as opposed to direct methods, which works in reciprocal space)
 - Ideal approach for molecular structures
- 3D Fourier Analysis
 - An ideal tool for structure completion, when other methods give only a partial or incomplete structural model
- Charge Flipping
 - An iterative algorithm which uses both direct and reciprocal space
 - Does not require a trial structure or chemical input -> FAST

SDPD

The Process



- Structure determination through powder diffraction is a sequential process with clearly defined stages:
 1. Indexing
 2. Space Group Determination
 3. Profile Fitting/Intensity Extraction
 4. Structure Determination
 5. Structure Refinement
- Essentially, it is not possible to continue until the previous step has been completed.
- Going backwards or repeating a step is possible (almost guaranteed!)

Live Example #1

LiFePO₄ – Simulated Annealing



- Solving the structure of LiFePO₄ using Simulated Annealing and Fourier Methods

Measurement conditions:

D8 Advance Diffractometer

- Mo tube
- Focusing Goebel mirror
- Capillary geometry
- LynxEye linear detector
- 12-hour measurement

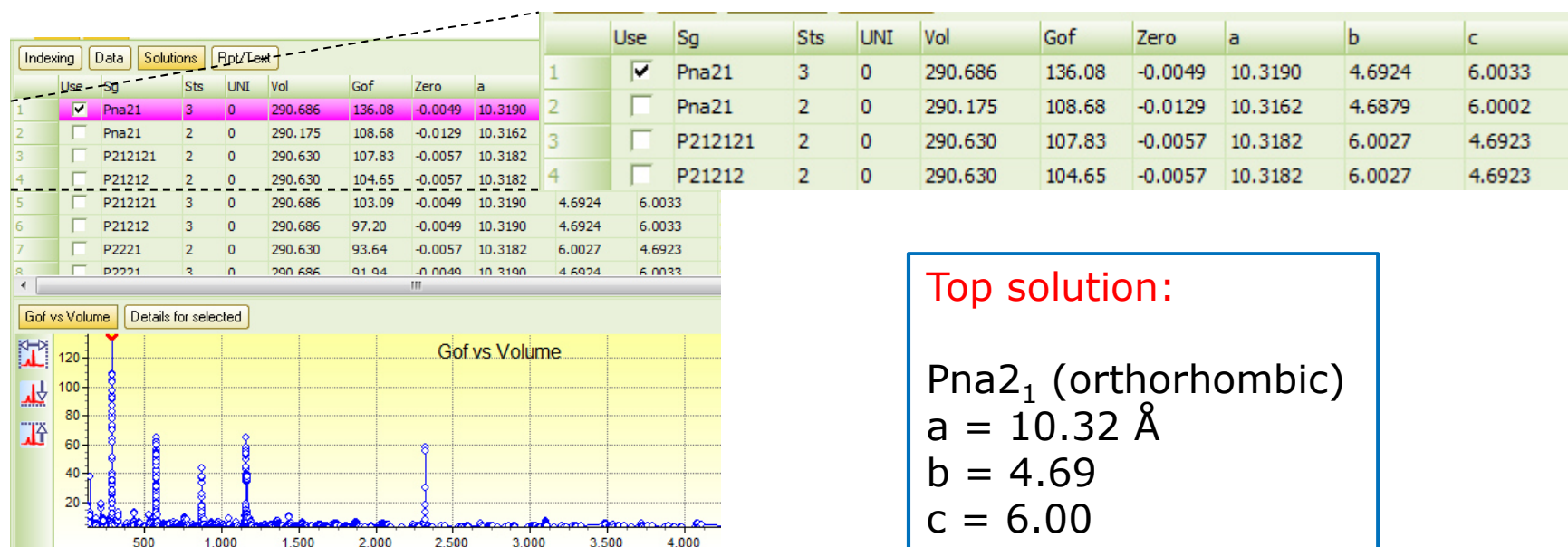


Live Example #1

LiFePO₄ – Simulated Annealing



- Solving the structure of LiFePO₄ using Simulated Annealing and Fourier Methods
- Step 1: Indexing
 - Indexing gives us a possible solution (or solutions) that we can check one by one



Live Example #1

LiFePO₄ – Simulated Annealing



- Solving the structure of LiFePO₄ using Simulated Annealing and Fourier Methods
- Step 2: Space Group Determination
 - Better said, the results from indexing gives us possible crystal systems and *extinction symbols*, not the space group directly
 - Several space groups may share an extinction symbol:

Table 7.2: Unique space group hkl's in powder diffraction.

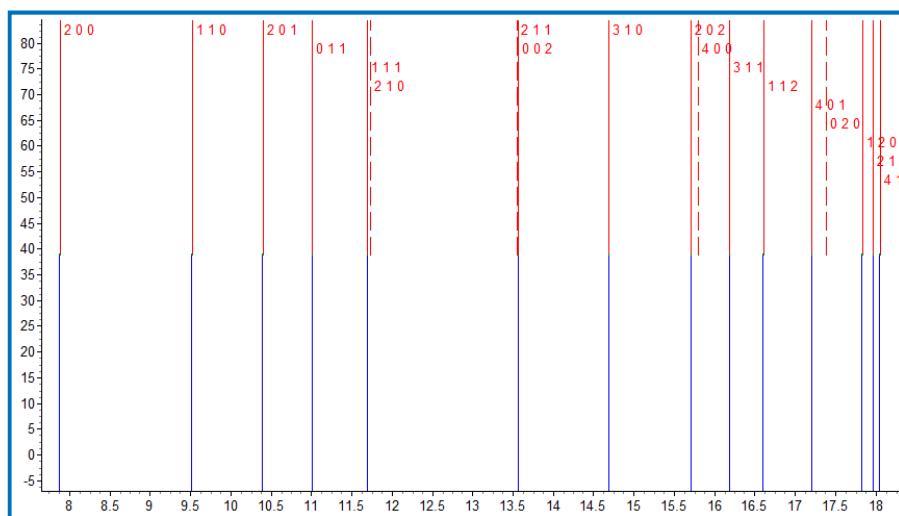
Space group numbers with identical hkl's	Space group symbols with identical hkl's		
Triclinic		54	Pcca
1 2	P1 P-1	50	Pban
Monoclinic		33 62	Pna21 Pnma
9 15	Cc C2/c	34 58	Pnn2 Pnnm
5 8 12	C2 Cm C2/m	32 55	Pba2 Pbam
14	P21/c	30 53	Pnc2 Pmna
7 13	Pc P2/c	29 57	Pca21 Pbcm
4 11	P21 P21/m		
3 6 10	P2 Pm P2/m		

Live Example #1

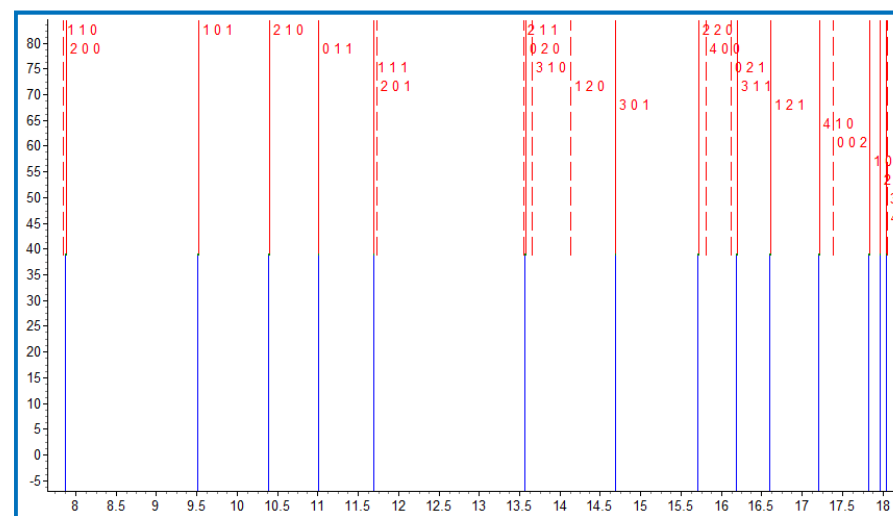
LiFePO₄ – Simulated Annealing



- Step 2: Space Group Determination
 - Because of peak overlap and difficult-to-see weak reflections in PD, unambiguous determination of the correct *extinction symbol* and *space group* is generally not possible...



Calculated reflections for Pna2₁



Calculated reflections for P2₁2₁2₁

Live Example #1

LiFePO₄ – Simulated Annealing



- Step 2: Space Group Determination
 - When more than one choice presents itself, the correct space group may be determined by:
 - molecular properties/presence of polyhedra
 - consideration of space group frequencies
 - trial and error (attempting the complete structure solution process for each probable space group and choosing the most reliable)
 - In this example, let's start with Pna2₁, as we have a higher GOF in the indexing and fewer 'missing' calculated reflections

Table 7.2: Unique space group hkl's in powder diffraction.

Space group numbers with identical hkl's	Space group symbols with identical hkl's
33 62	Pna21 Pnma

Live Example #1

LiFePO₄ – Simulated Annealing



- Step 3: 'Profiling' (Pawley fit)
 - A Pawley fit is similar to a Rietveld refinement, and many of the parameters are the same. The main difference is that the intensity of each reflection, I_{hkl} , is calculated and refined
 - The main purpose of this step is to get 'correct' values for all other refineable parameters so that they may be fixed during the Structure Determination step. These parameters include:
 - Background
 - Zero shift
 - Lattice parameters
 - Profile parameters

Live Example #1

LiFePO₄ – Simulated Annealing



- Step 4: Structure Determination (trial structure)
 - To make a trial structure, first throw the required atoms into the box
 - $V = 290 \text{ \AA}^3$
 - 1 non-H (non-Li) atom $\sim 15 \text{ \AA}^3$
 - 1 formula unit = $6 * 15 = 90 \text{ \AA}^3$
 - $Z = 290 / 90 = 3.22 \rightarrow 4$
 - In space groups $Pna2_1$ and $Pnma$, the highest symmetry site has a multiplicity of 4, so $Z = 4$ makes sense...

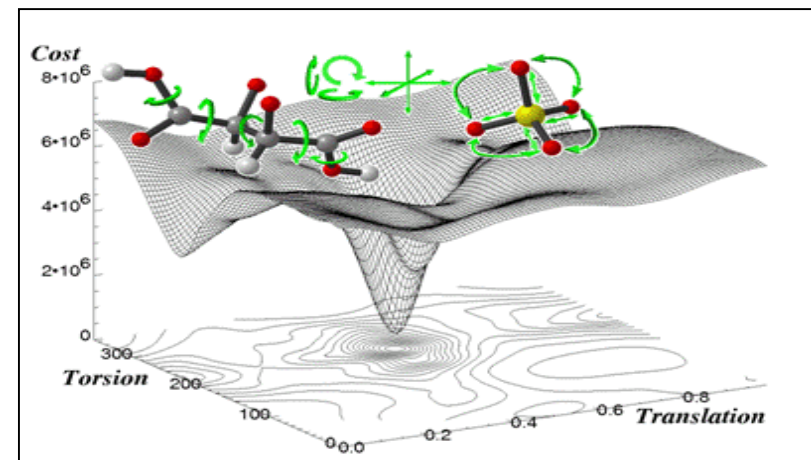
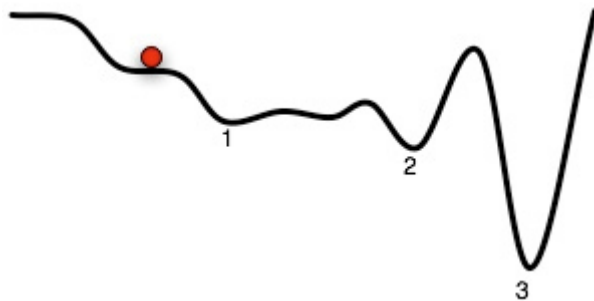
$Pna2_1$	Positions			Coordinates			
	Multiplicity,	Wyckoff letter,	Site symmetry				
No. 33	4	a	1	(1) x, y, z	(2) $\bar{x}, \bar{y}, z + \frac{1}{2}$	(3) $x + \frac{1}{2}, \bar{y} + \frac{1}{2}, z$	(4) $\bar{x} + \frac{1}{2}, y + \frac{1}{2}, z + \frac{1}{2}$

Simulated Annealing

What is it?



- Analogous to the annealing process applied to metals:
 - There is a temperature variable to simulate the heating process
 - The algorithm is set to high temperature which is slowly allowed to 'cool'
 - At higher temperatures, bad moves (solutions) are accepted with higher frequency, which allows the algorithm to jump out of local minima it may find itself in
 - The temperature is then reduced, and so is the chance of accepting worse solutions, so it focuses on an area in search space where an optimum solution may be found



From: Radovan ČERNÝ, Chem. Met. Alloys
1(2008) 120-127,

SDPD

Simulated Annealing



- Simulated annealing is a direct space approach where adjustable parameters lie in direct rather than reciprocal space
- Procedure:
 1. A trial crystal structure is constructed by randomly positioning and orienting individual atoms, molecular fragments or complete molecules taking into account (known or guessed) space group information
 2. After calculating diffraction data and comparing it against the measured diffraction data, the variable parameters of the model are adjusted in order to maximize the level of agreement between the observed and calculated data (i.e., minimize χ^2).

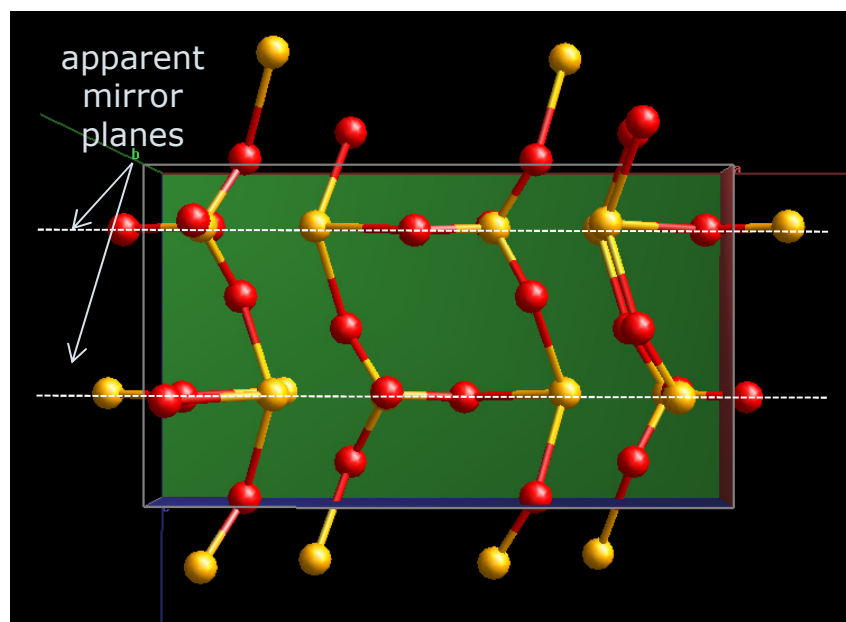
Live Example #1

LiFePO_4 – Simulated Annealing

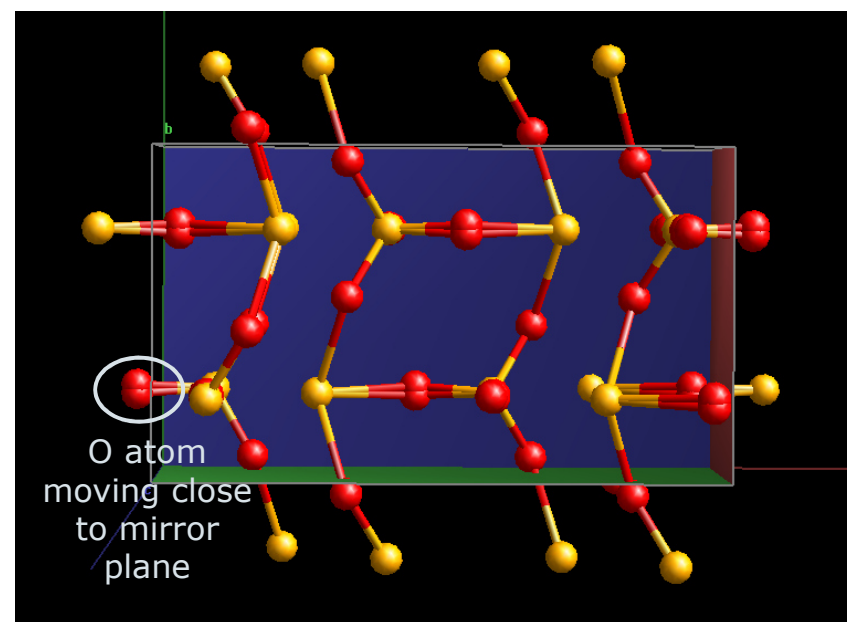


- Step 4: Structure Determination (trial structure)

Trial structure in $\text{Pna}2_1$



Trial structure in Pnma



Live Example #1

LiFePO₄ – Simulated Annealing



- Step 4: Structure Determination (find the missing Li atoms)
- If we have F_{hkl} , we can obtain the electron density through a reverse Fourier transform:

$$\rho(x,y,z) = \sum_{hkl} F_{hkl} \exp \{ -2\pi i (hx + ky + lz) \} / V$$

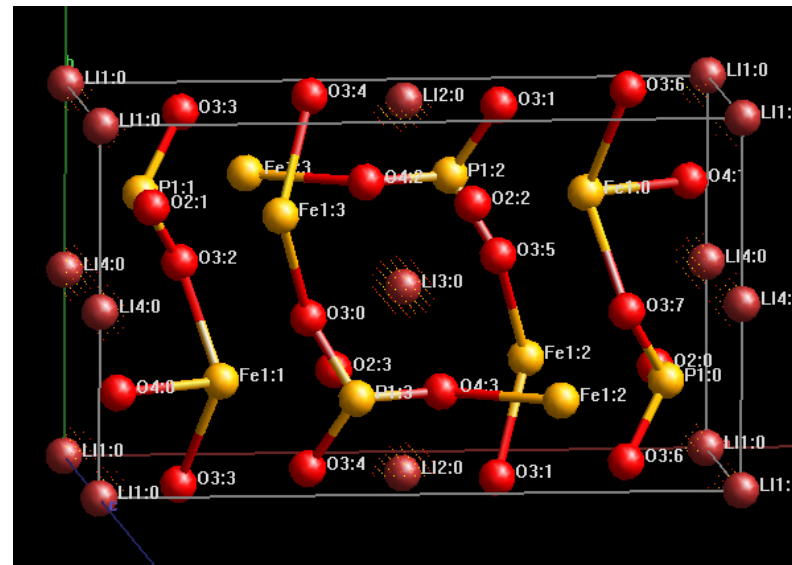
- A **difference Fourier map** will then show us the residual electron density where missing atoms may be located

Live Example #1

LiFePO₄ – Simulated Annealing



- Step 4: Structure Determination (difference Fourier map)
 - fourier_map 1
 - fourier_map_formula = Fobs - Fcalc;
 - load f_atom_type f_atom_quantity
 - {
 - Li = 4;
 - }



Live Example #1

LiFePO₄ – Simulated Annealing



- Step 5: Structure Refinement

Live Example #2

Cimetidine – Charge Flipping



- Structure determination through powder diffraction is a sequential process with clearly defined stages:

1. Indexing
2. Space Group Determination
- 3. Profile Fitting/Intensity Extraction**
- 4. Structure Determination**
5. Structure Refinement



1-cyano-2-methyl-3-(2-{[4-methyl-1*H*-imidazol-5-yl)methyl]sulfanyl}ethyl)guanidine

Charge Flipping

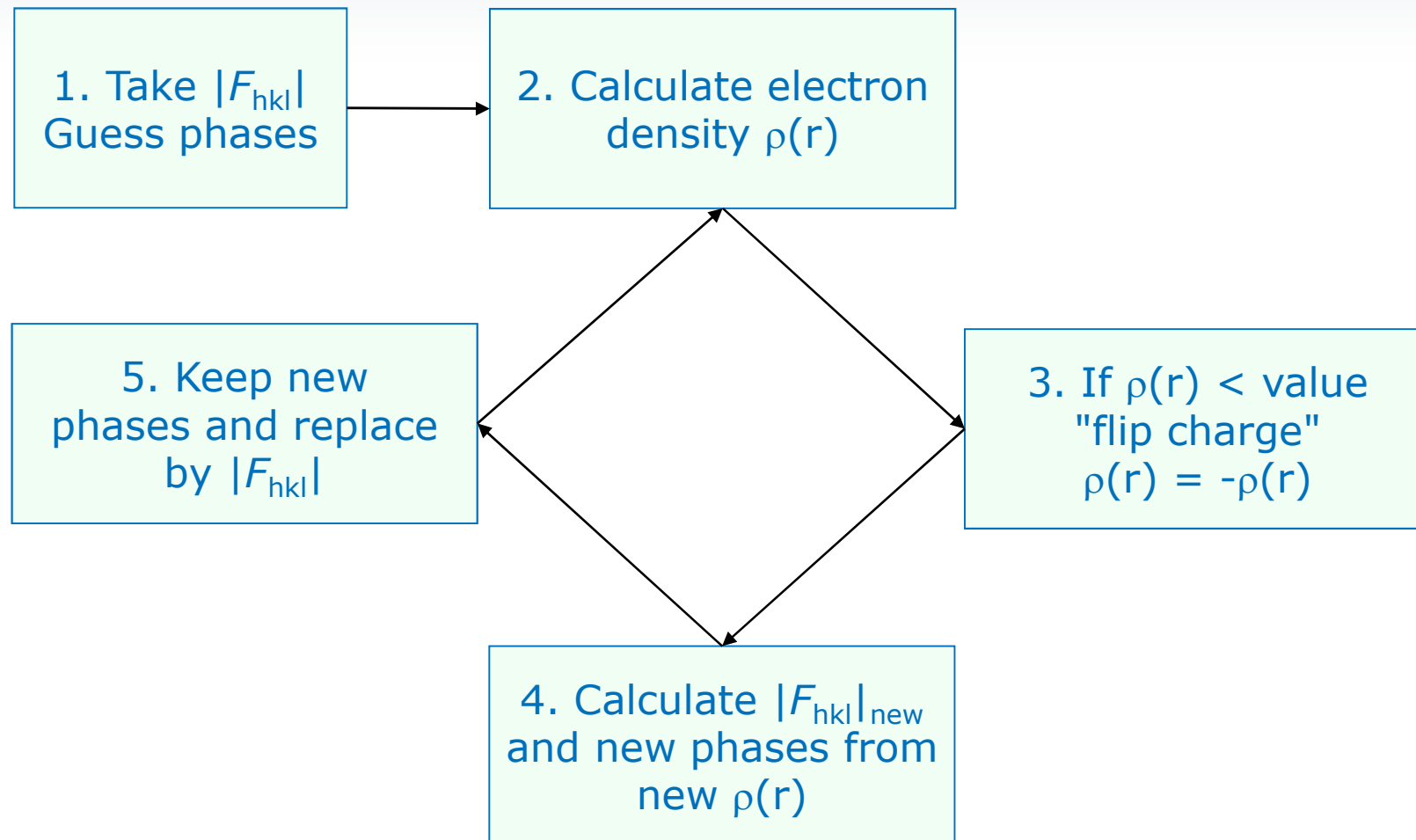
What is it?



- First described by Oszlányi and Sütő in 2004
- Iterative dual-space algorithm
- Required input are lattice parameters and reflection intensities (F_{obs})
- No use of chemistry or trial structure models
- The output is an approximate scattering density (i.e. structure) sampled on a discrete grid
- Charge flipping is very fast
 - Structures can be (partially) solved in a few seconds up to a few minutes (generally faster than it takes to set up a trial structure model for simulated annealing)

Charge Flipping

What is it?



Live Example #2

Cimetidine – Charge Flipping



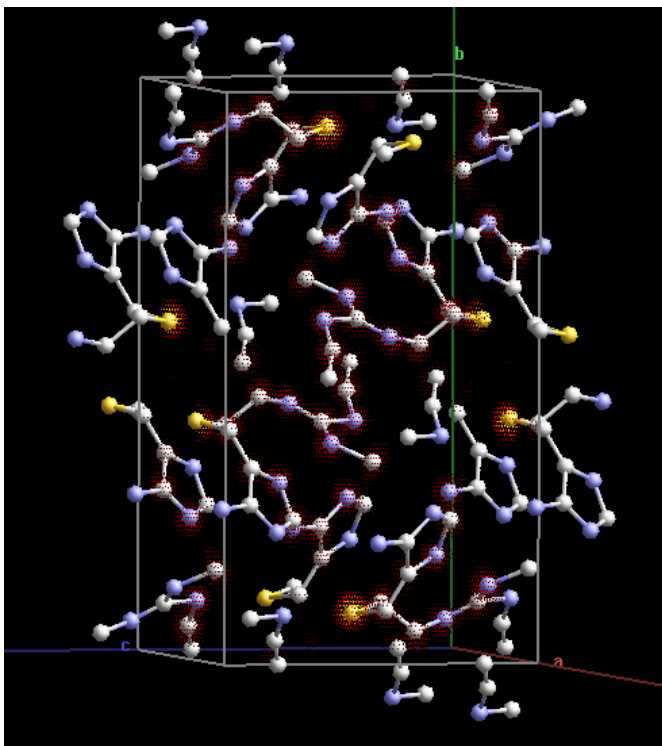
- Step 3: Intensity Extraction, $|F_{\text{obs}}|$
- Step 4: Structure Determination

Live Example #2

Cimetidine – Charge Flipping



- Step 3: Intensity Extraction, $|F_{\text{obs}}|$
- Step 4: Structure Determination



```
Temporary output
----- Picked atoms -----
' Maximum number of atoms to be picked 68
str
a 6.8256
b 18.82
c 10.395
al 90
be 106.4
ga 90
site S1  x 0.78473 y 0.08098 z 0.98495 occ S 0.786 beq 1
' site S1  x 0.21527 y 0.58098 z 0.51505 occ S 0.786 beq 1
' site S1  x 0.78473 y 0.41902 z 0.48495 occ S 0.786 beq 1
' site S1  x 0.21527 y 0.91902 z 0.01505 occ S 0.786 beq 1
site N2  x 0.40321 y 0.23437 z 0.63823 occ N 1.035 beq 1
' site N2  x 0.59679 y 0.73437 z 0.86177 occ N 1.035 beq 1
' site N2  x 0.40321 y 0.26563 z 0.13823 occ N 1.035 beq 1
' site N2  x 0.59679 y 0.76563 z 0.36177 occ N 1.035 beq 1
site N3  x 0.90477 y 0.12494 z 0.45126 occ N 1.043 beq 1
' site N3  x 0.09523 y 0.62494 z 0.04874 occ N 1.043 beq 1
' site N3  x 0.90477 y 0.37506 z 0.95126 occ N 1.043 beq 1
' site N3  x 0.09523 y 0.87506 z 0.54874 occ N 1.043 beq 1
site N4  x 0.78189 y 0.59326 z 0.89211 occ N 1.018 beq 1
' site N4  x 0.21811 y 0.09326 z 0.60789 occ N 1.018 beq 1
' site N4  x 0.78189 y 0.90674 z 0.39211 occ N 1.018 beq 1
' site N4  x 0.21811 y 0.40674 z 0.10789 occ N 1.018 beq 1
site N5  x 0.65487 y 0.18611 z 0.21708 occ N 1.013 beq 1
' site N5  x 0.34513 y 0.68611 z 0.28292 occ N 1.013 beq 1
' site N5  x 0.65487 y 0.31389 z 0.71708 occ N 1.013 beq 1
' site N5  x 0.34513 y 0.81389 z 0.78292 occ N 1.013 beq 1
site N6  x 0.87560 y 0.56121 z 0.12540 occ N 1.021 beq 1
' site N6  x 0.12440 y 0.06121 z 0.37460 occ N 1.021 beq 1
' site N6  x 0.87560 y 0.93879 z 0.62540 occ N 1.021 beq 1
' site N6  x 0.12440 y 0.43879 z 0.87460 occ N 1.021 beq 1
```

Simulated Annealing vs. Charge Flipping

Conclusions

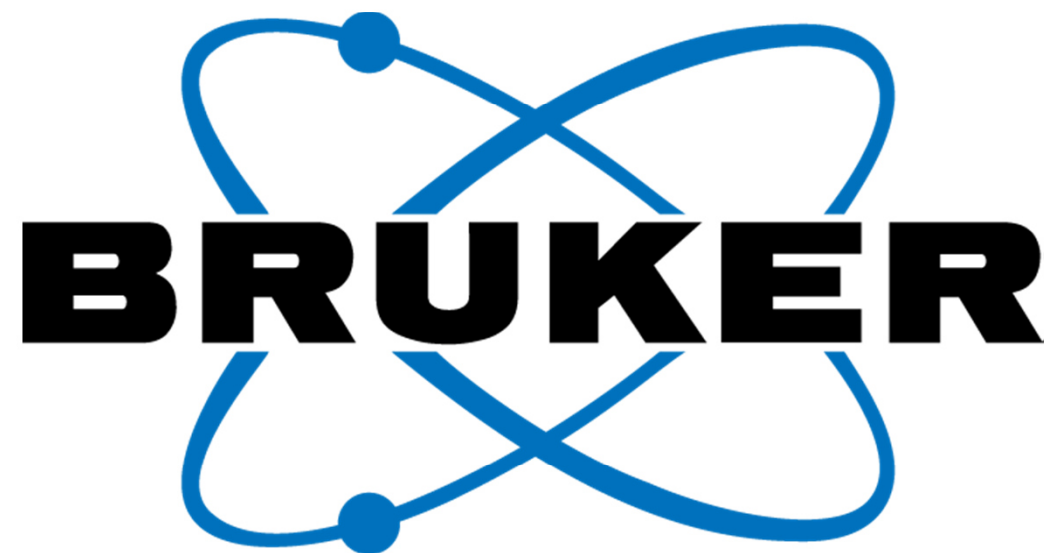


Simulated Annealing:

- Requires a trial structure model, which can be partial or random
- Performs better on poor quality data. **Important advantage!**
- Comparatively slow

Charge Flipping:

- No use of chemistry / trial structure models. **Important advantage!**
- Requires high quality data
- Even if the structure does not solve completely, heavy atoms and / or molecular fragments can often be found very quickly, which greatly assists subsequent simulated annealing structure determination
- Very fast; structures can be (partially) solved in seconds up to a few minutes, i.e. faster than one typically can create a start model / rigid body for simulated annealing



Innovation with Integrity