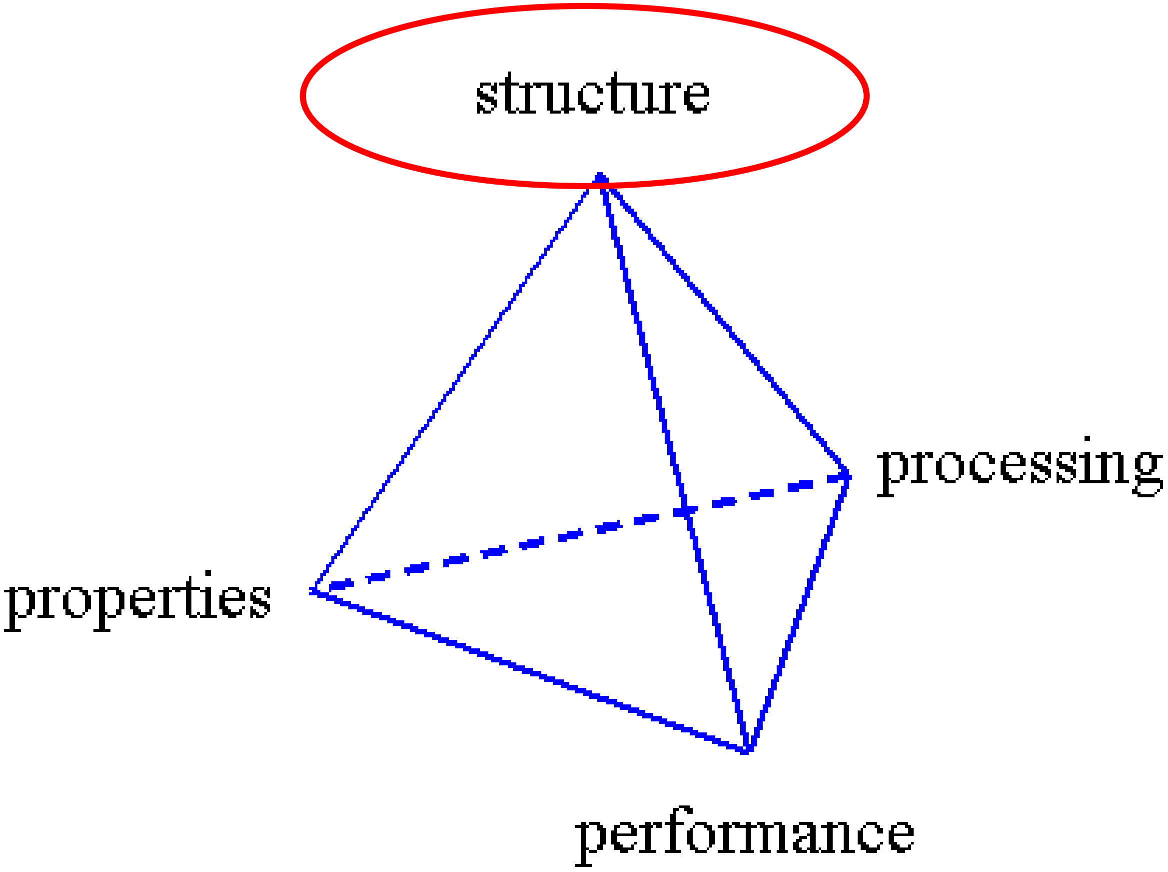
Feasible Sets

L. D. Marks
Northwestern University

Electron Crystallography is the branch of science that uses electron scattering and imaging to study the structure of matter

What is the science?



Why determine the structure?

- To finish my PhD
- To get/keep my job
- Because structure *coupled* with other science really matters – but only when coupled
- Follow the science, not the electron

How to solve a structure?

- Guess, then refine
  - Will always give something, but if the guess is wrong GIGO
- Use Patterson function
  - Difficult for complicated structures (more to come)
- Use DFT[1]
  - If the original guess is wrong, GIGO
  - Functionals are inaccurate for most oxides (energies wrong)
- Get an image
  - STM is hard to interpret
  - HREM, can be ambiguous (more to come)
- Get a Diffraction Pattern
  - Incomplete information (more to come)

[1] For later today

Four basic elements are required to solve a recovery problem

1. A data formation model

   Imaging/Diffraction/Measurement

2. A priori information

   The presence of atoms or similar

3. A recovery criterion:

   A numerical test of Goodness-of-Fit

4. A solution method.

   Mathematical details

Patrick Combettes, (1996). Adv. Imag. Elec. Phys. 95, 155

Four basic elements are required to solve a recovery problem

1. A data formation model

    Imaging/Diffraction/Measurement

    Kinematical Theory/Linear Imaging

        Single Weak Scattering + Ewald Sphere

        Qualitatively correct; Quantitatively inaccurate

    Bragg's Law

        Single Scattering + Zero Excitation Error

        Worse than Kinematical Theory (it is different)

    Dynamical Theory/Non-Linear Imaging

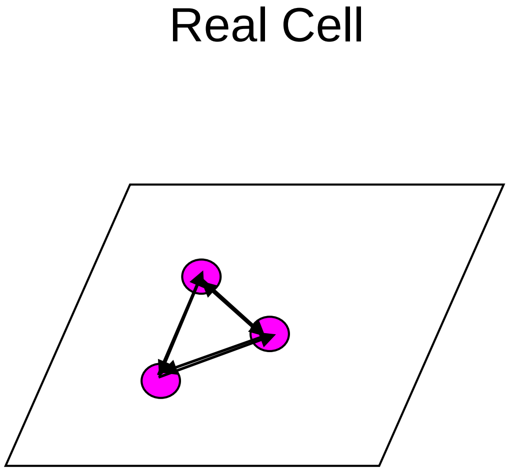
        Quantitatively correct, to the accuracy of the electrostatic potential (exact in principle)

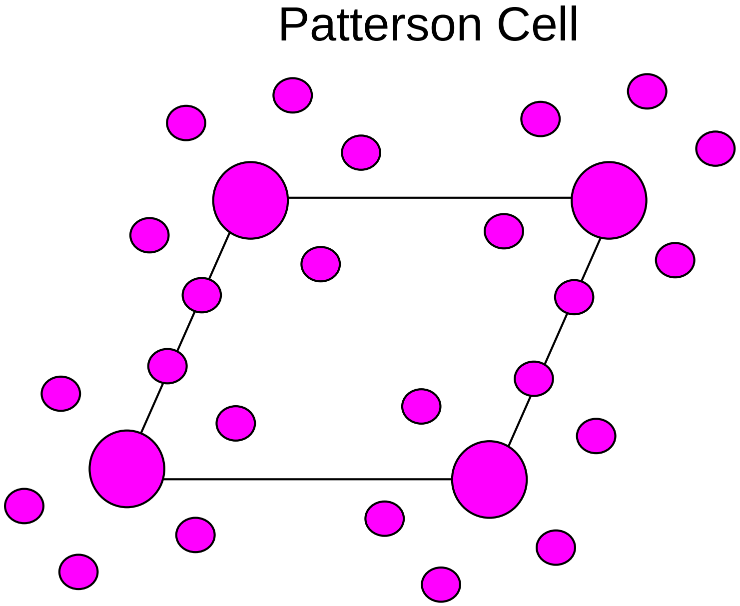
Warning: Errors in the model introduce systematic errors in the recovery which of course can lead to GIGO

$$P(uvw) = \frac{1}{V}\sum_{hkl} I(hkl)e^{-2\pi i(hu+lv+kw)}$$

Patterson Function

$$P(\vec{u}) = \int \rho(\vec{r})\rho(\vec{r}+\vec{u})d^3\vec{r}$$



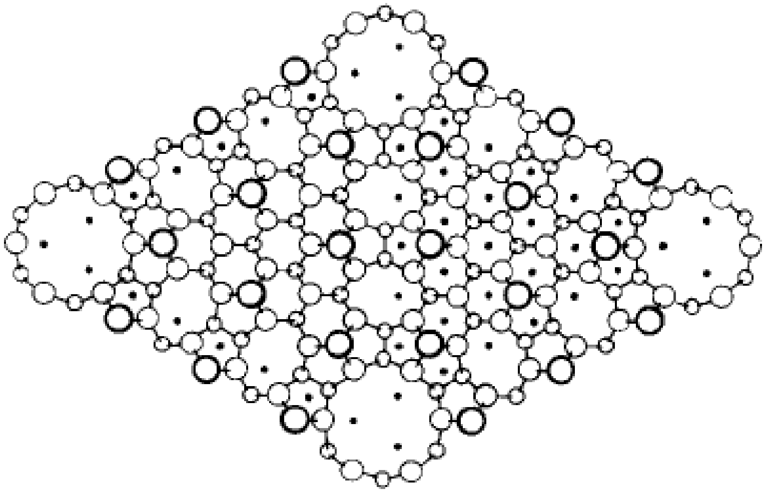


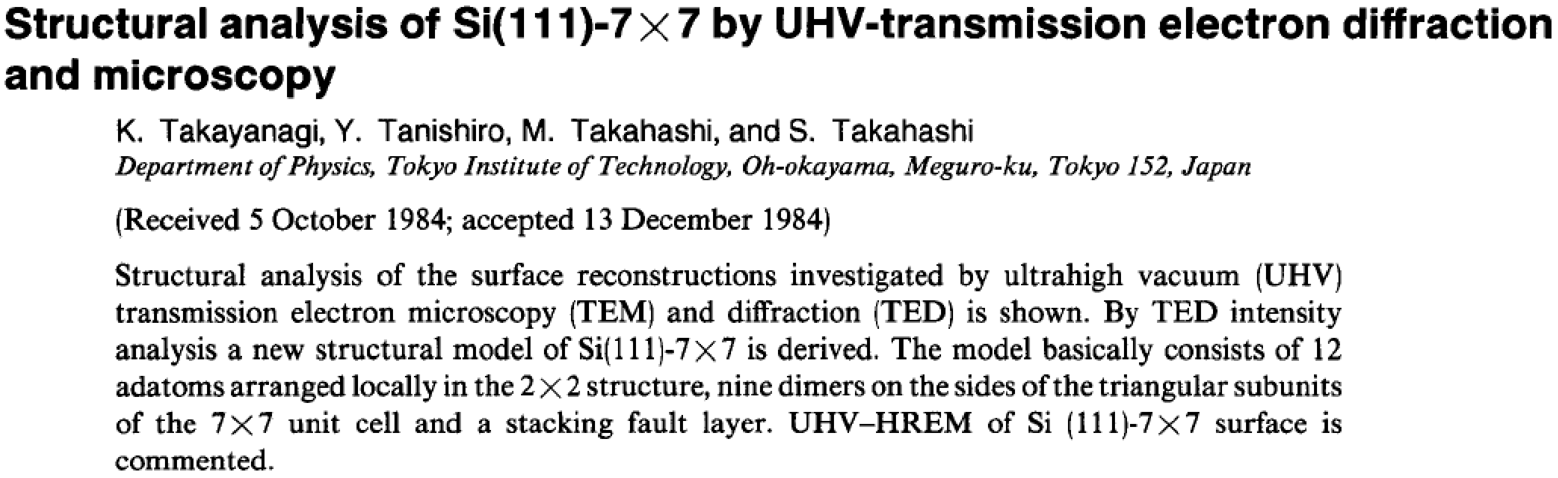
1) Patterson is symmetric about origin (centrosymmetry)

2) Can see pattern of real cell in Patterson cell repeated N times

3) Contains N(N-1) peaks (not counting origin) → gets complicated!

Patterson Function





Structural analysis of Si(111)-7×7 by UHV-transmission electron diffraction and microscopy

K. Takayanagi, Y. Tanishiro, M. Takahashi, and S. Takahashi
Department of Physics, Tokyo Institute of Technology, Oh-okayama, Meguro-ku, Tokyo 152, Japan

(Received 5 October 1984; accepted 13 December 1984)

Structural analysis of the surface reconstructions investigated by ultrahigh vacuum (UHV) transmission electron microscopy (TEM) and diffraction (TED) is shown. By TED intensity analysis a new structural model of Si(111)-7×7 is derived. The model basically consists of 12 adatoms arranged locally in the 2×2 structure, nine dimers on the sides of the triangular subunits of the 7×7 unit cell and a stacking fault layer. UHV–HREM of Si (111)-7×7 surface is commented.

J Vac Sci Technol A3, 1502 (1986) > 1800 Citations

The Phase Problem

- We have an exit wave from the sample
  - $\psi(r)$ wave in real space $= a(r)\exp(-i\phi(r))$
  - $\Psi(u) = \int \exp(-2\pi i u.r)\psi(r)dr = A(u)\exp(-i\phi(u))$
- Observables
  - $I(r) = <|\psi(r)|^2> = <a(r)^2>$ Real Space Image
  - $I(u) = <|\Psi(u)|^2> = <A(u)^2>$ Diffraction Pattern
- Note: "<>" is average over incoherent aberrations and other statistical terms

Phase: Apples & Oranges

FT → $A_a \exp(-i\ \phi_a)$

FT → $A_o \exp(-i\ \phi_o)$

+

$A_o \exp(-i\ \phi_a)$ → IFT → $\left\{ \begin{array}{l} \text{Oranle ?} \\ \text{Appge ?} \end{array} \right.$

Phase of Apple + Amplitude of Orange = ?

Phase of Apple = Apple

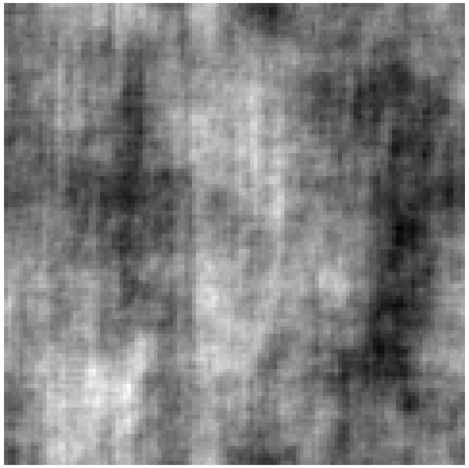
$$FT^{-1}\{A_o\,\underline{\exp(-i\,\phi_a)}\,\} \Rightarrow \text{Apple}$$

Phase is more important than amplitude

The importance of phase information

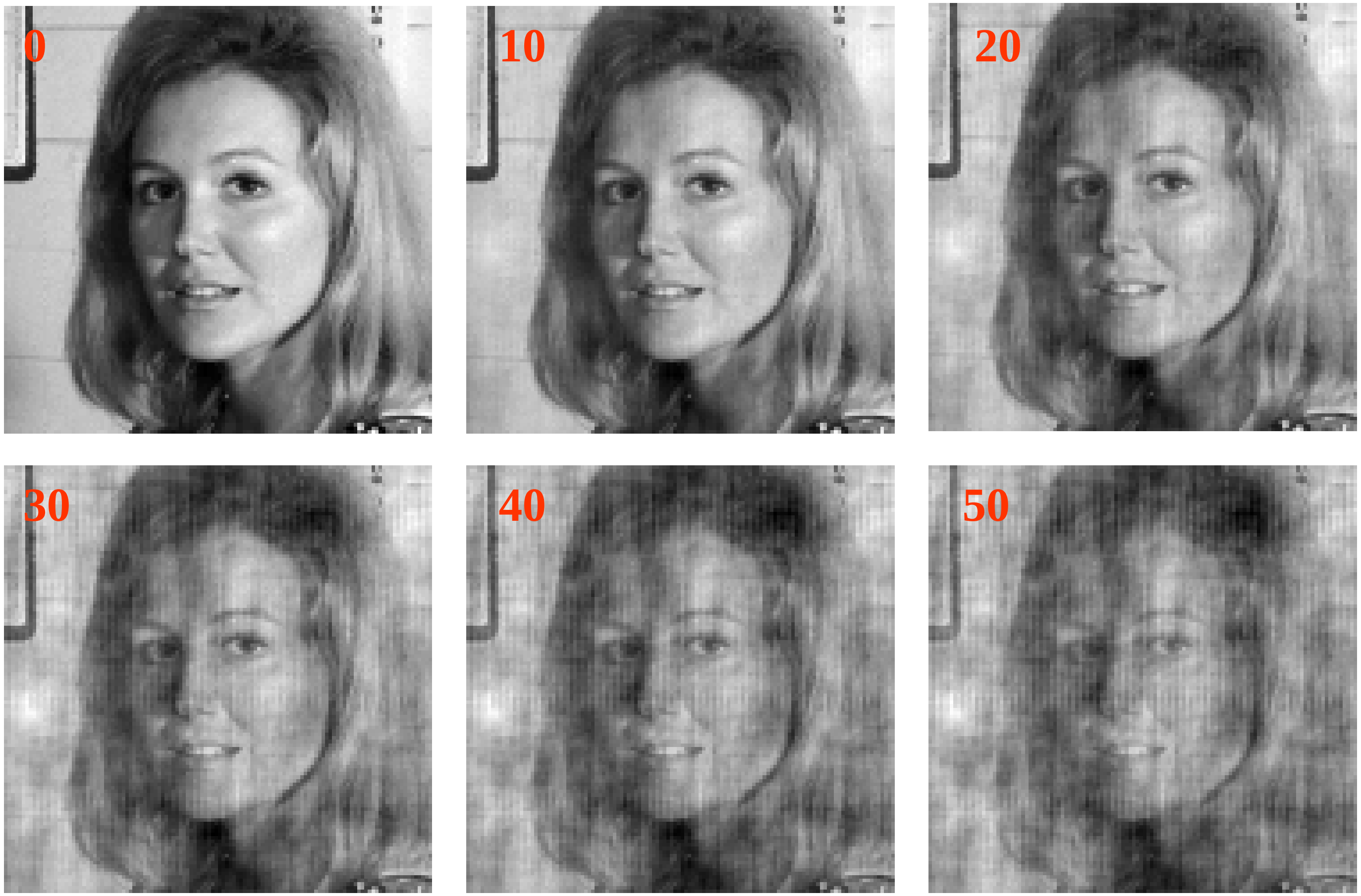
Suzy

Correct Modulus
Random Phases



Correct Phase Random
Modulus

Role of error in phases (degrees)



We would like to find the phases exactly, but we don’t have to

Phase and Modulus Errors

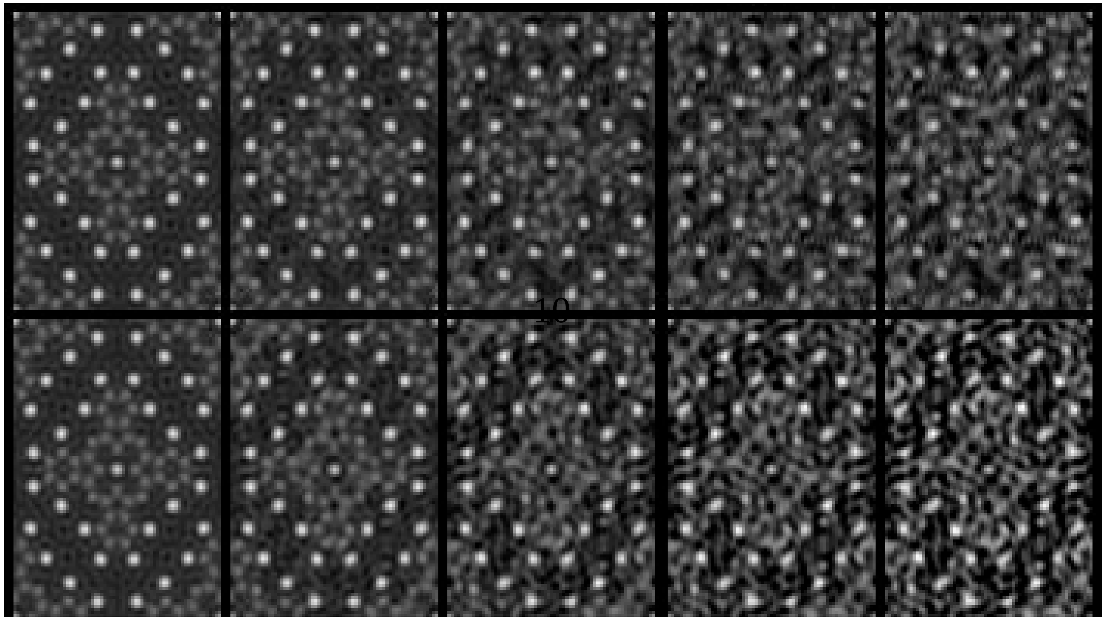
0 10° 20° 30° 40°

Phase Error

Modulus Correct

Modulus Error

Phase Correct



R 0 26% 52% 78% 104%

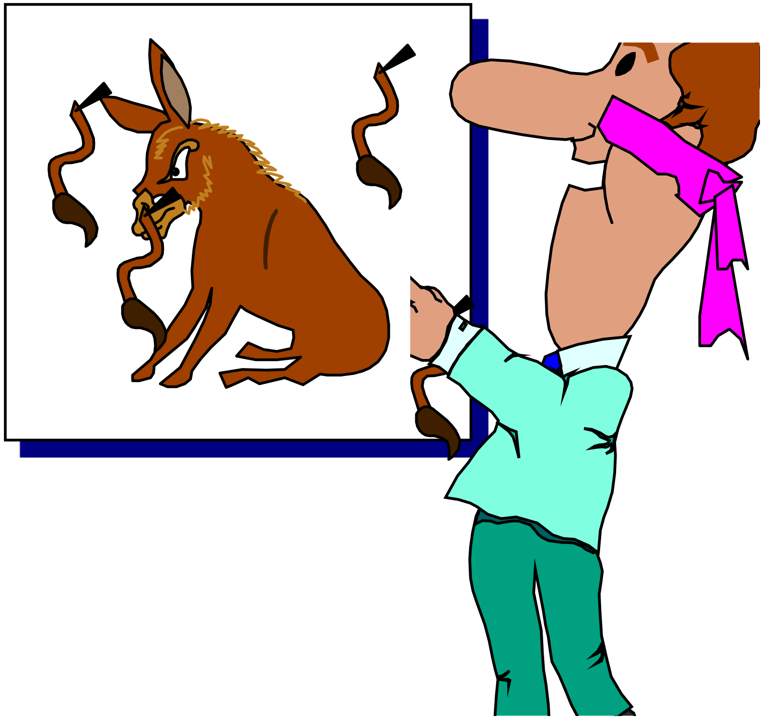
We only need approximately correct phases

We can tolerate modulus errors

Direct Methods vs. Indirect Methods

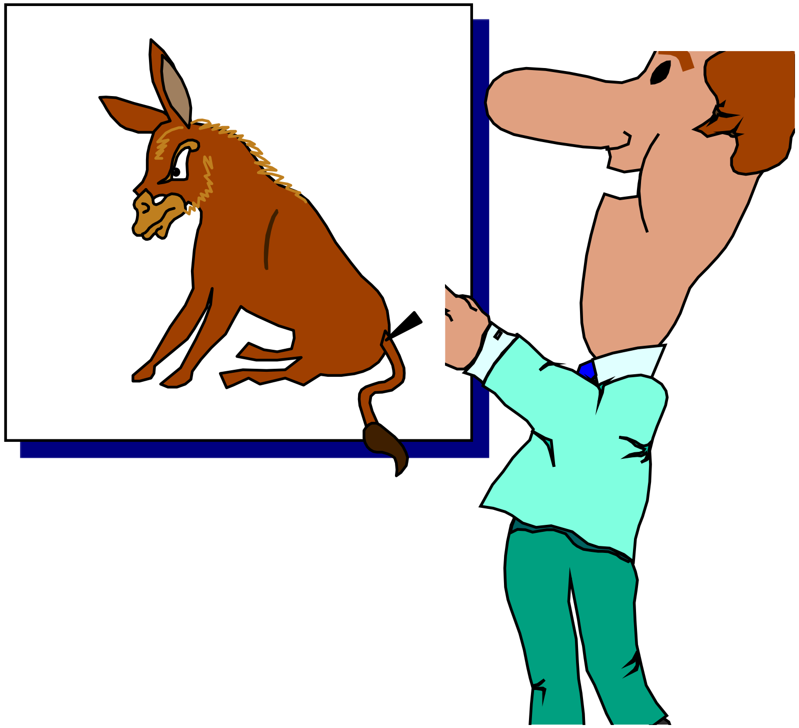
Indirect Methods:

"Trial and Error"

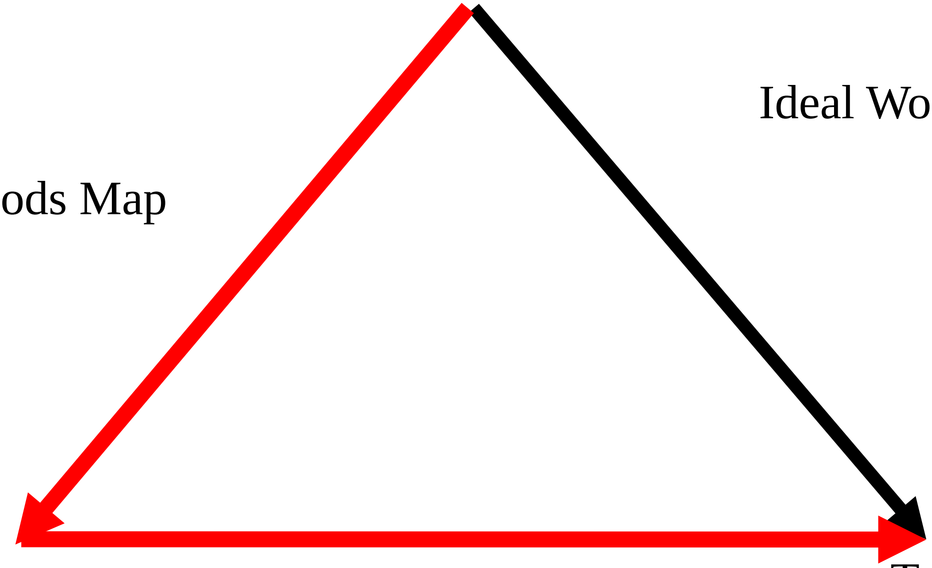


Direct Methods:

Using available information to find solutions



Crystallographic Direct Methods Structure Triangle



Data

Ideal World

Direct Methods Map
or Image

Trial
Structure

Structure Completion
(non-trivial)

True
Structure

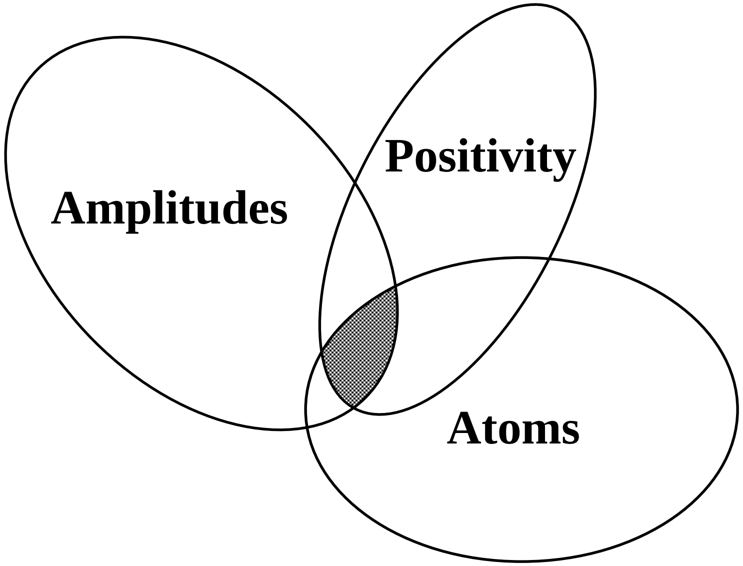
More: 1970’s Mathematics

- C -- Some constraints (e.g. atomicity, probabilities of triplets)
- F -- Some function (e.g. a FOM)
- Minimize, e.g. Lagrangian

$$L = F + \lambda C$$

1990's Mathematics

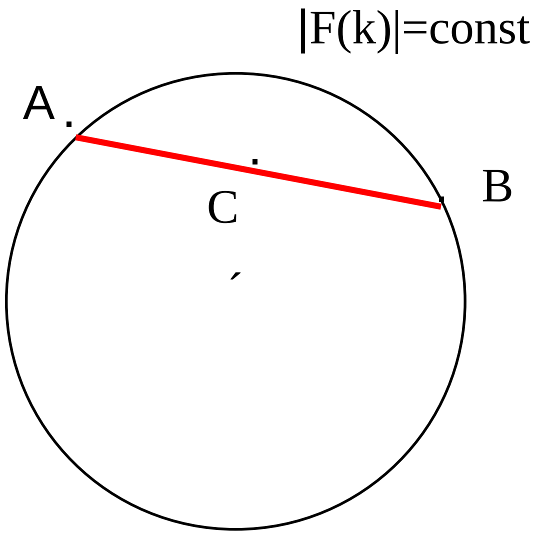
- We have constraints (e.g. atomicity, amplitudes)
  - Treat as sets
- We are looking for the solution as intersection of several constraint sets



Acta Cryst A55, 601 (1999)

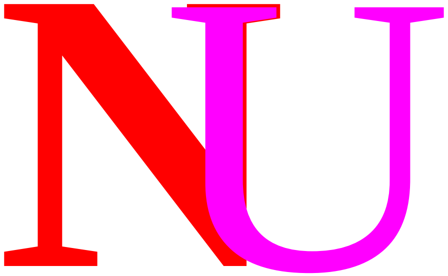
The $64,000 question

- A set is convex if any point between two members is also a member
  - If all the sets are convex, problem has one solution
  - If they are not, there may be more than one local minimum
- Amplitude measurements do not form a convex set
- But…there still may only be one solution.
- Unsolved mathematical problem

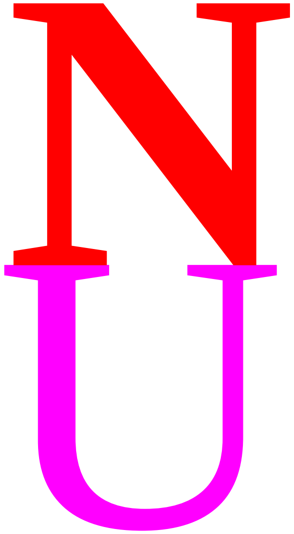


Multiple non-convex constraints

Consider the two sets “N” and “U”

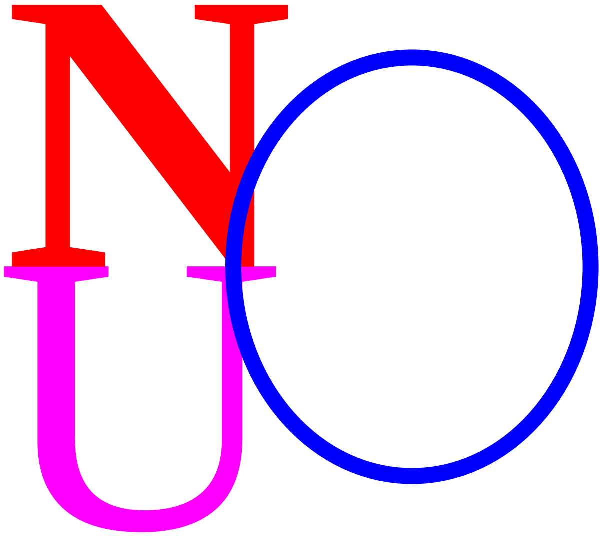


Overall Convex



Overall Non-Convex

Crystallographic methodology



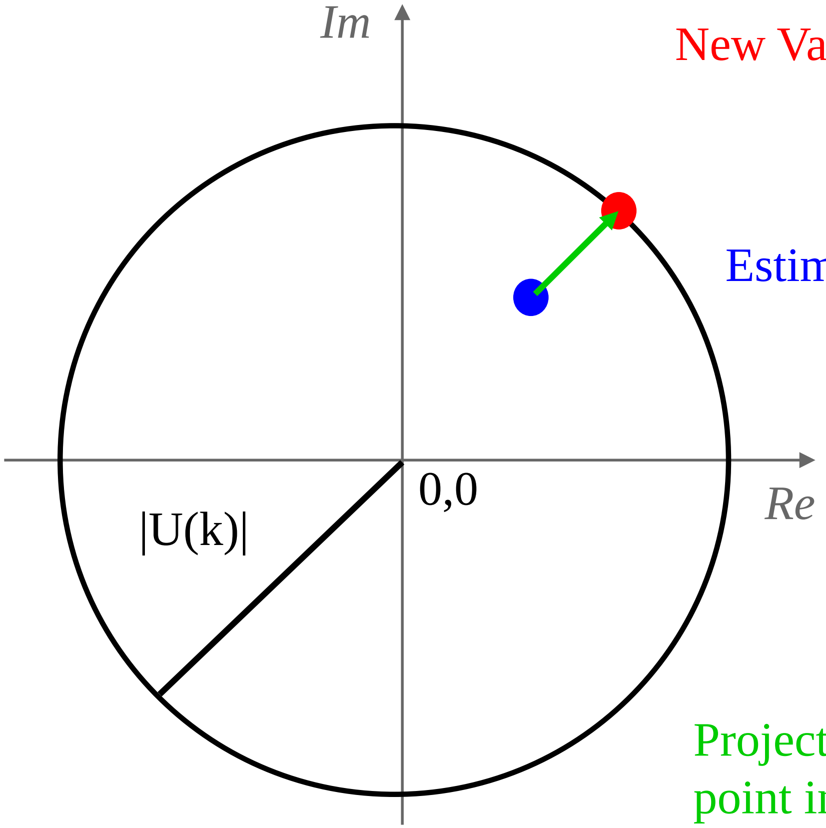
Overall Non-Convex

Overall Unique

Addition of additional convex constraints tends to give a unique solution

Structure Completion: add additional constraints as the phases become known

Orthogonal Projections

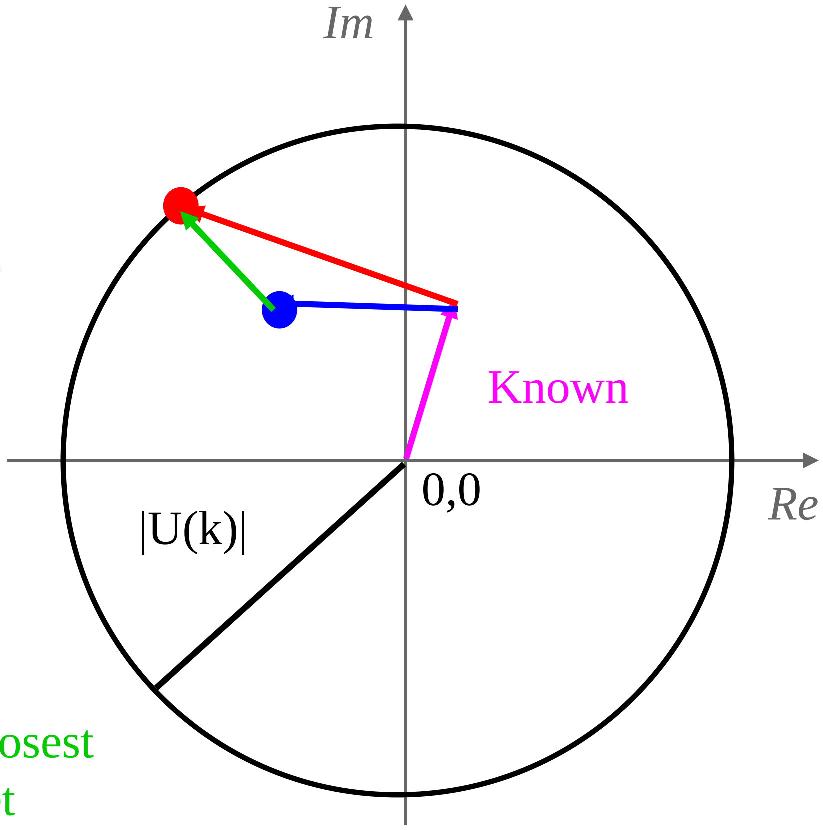


New Value

Estimate

Project: closest point in set

Modulus Only

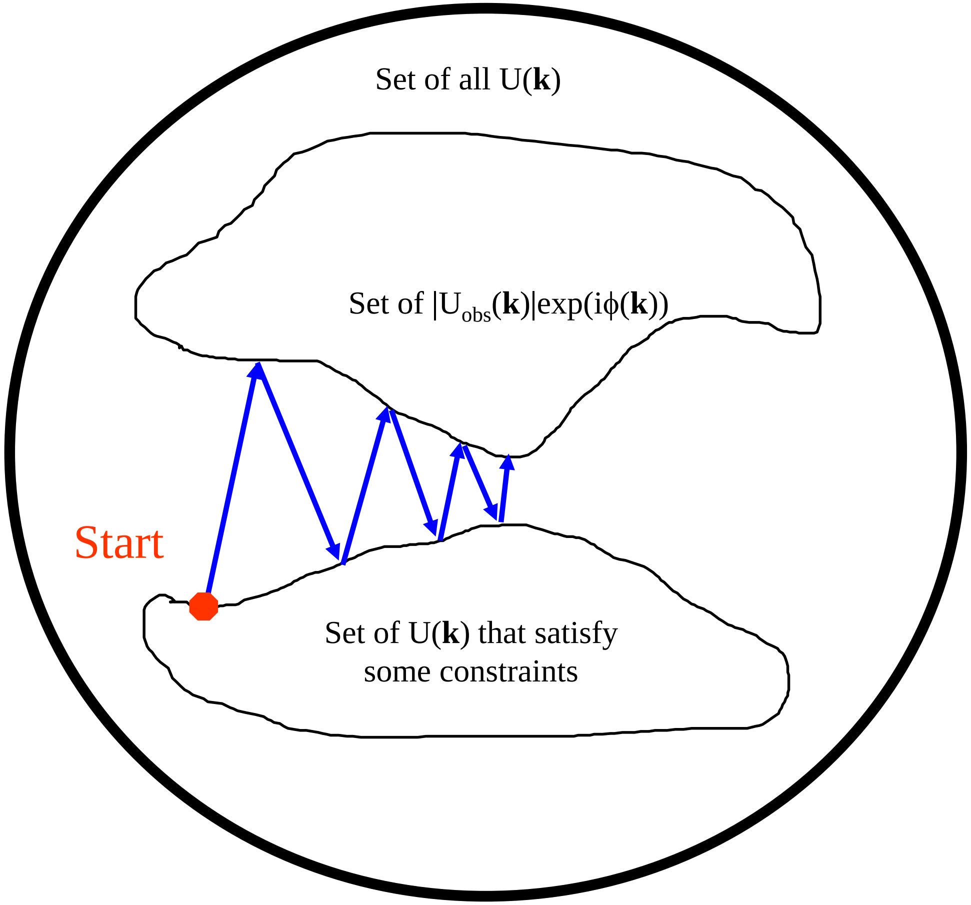


Known

Part of U(k) known

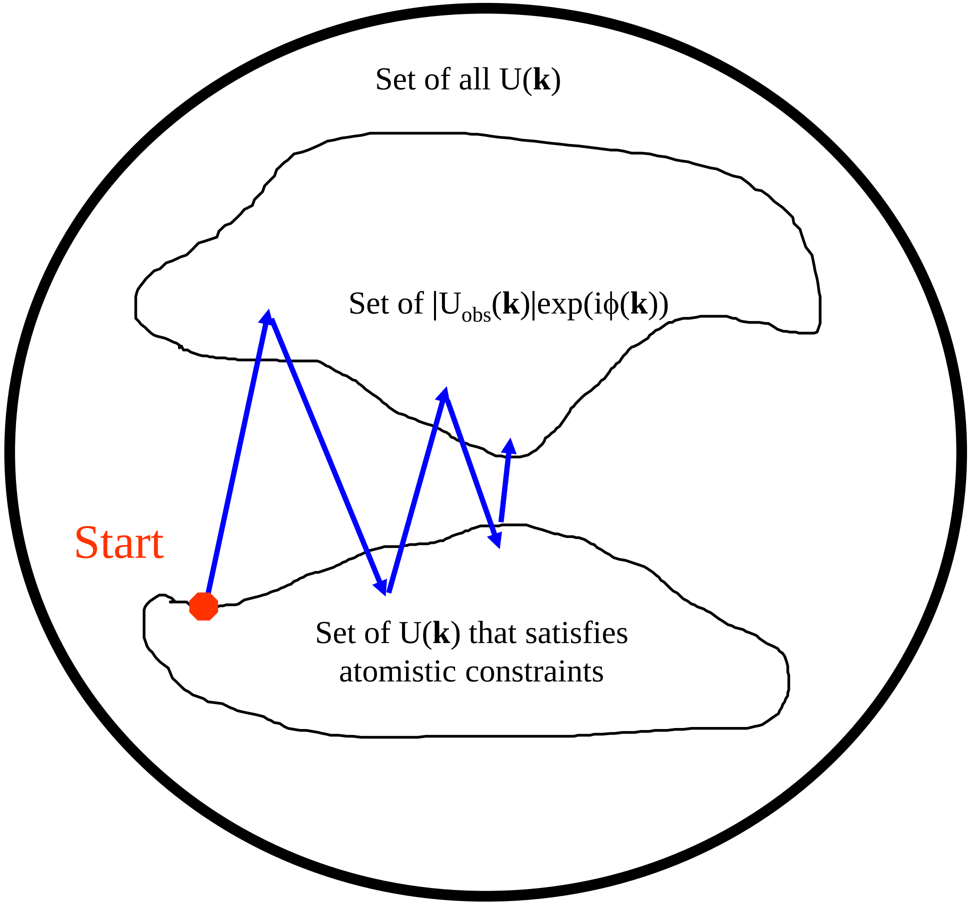
Successive Projections

- Iterate between projections
- Other variants possible (see Combettes, *Advances in Imaging and Electron Physics* **95,** 155-270, 1996)

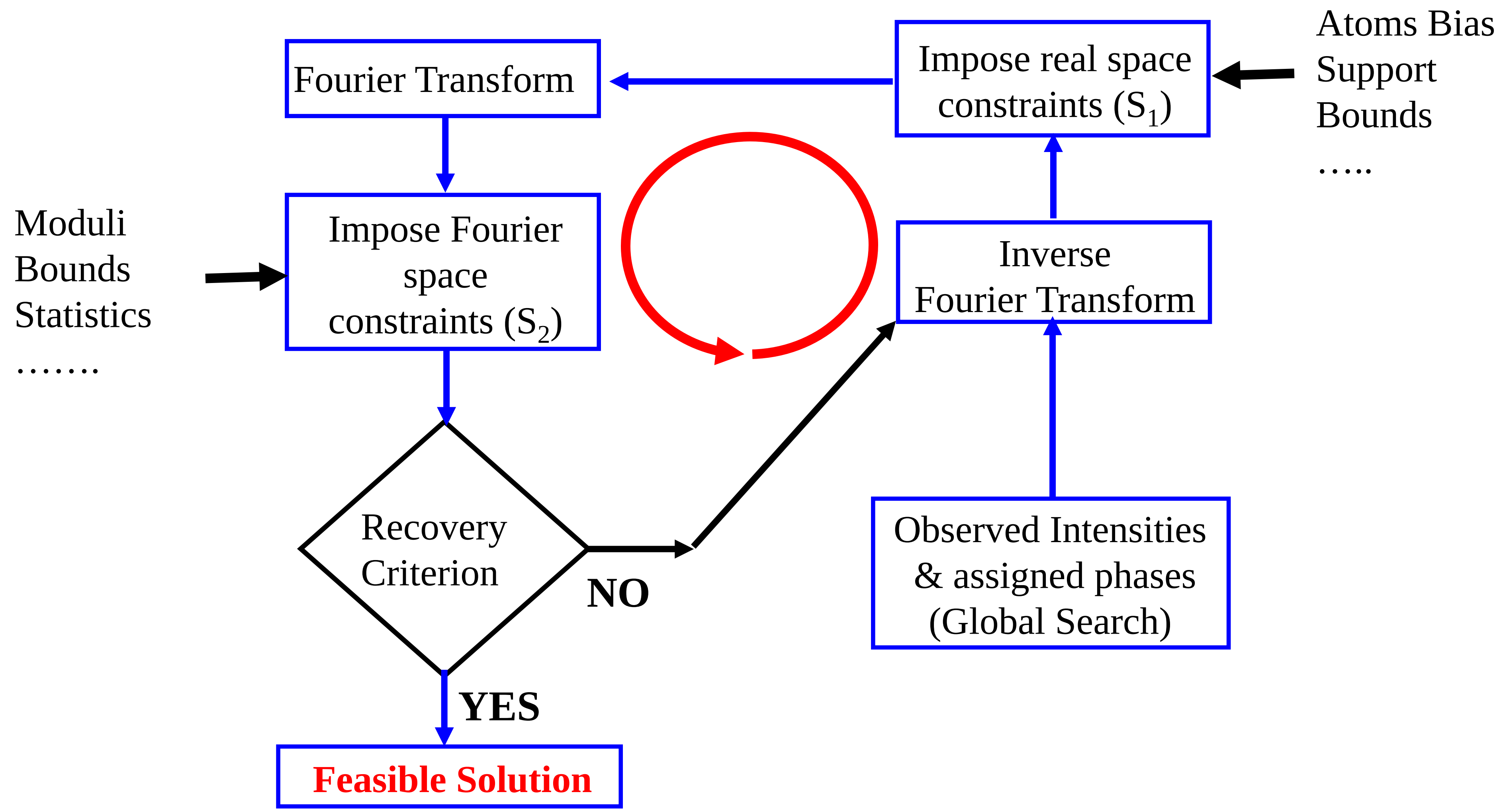


Over-relaxed Projections

- Iterate between projections
- Overshoot (deliberately)
- Converges faster
- Sometimes better solutions



Algorithm Overview (Gerschberg-Saxton)



Atoms Bias
Support
Bounds
…..

Fourier Transform

Impose real space constraints (S_1)

Moduli
Bounds
Statistics
…….

Impose Fourier space constraints (S_2)

Inverse Fourier Transform

Recovery Criterion

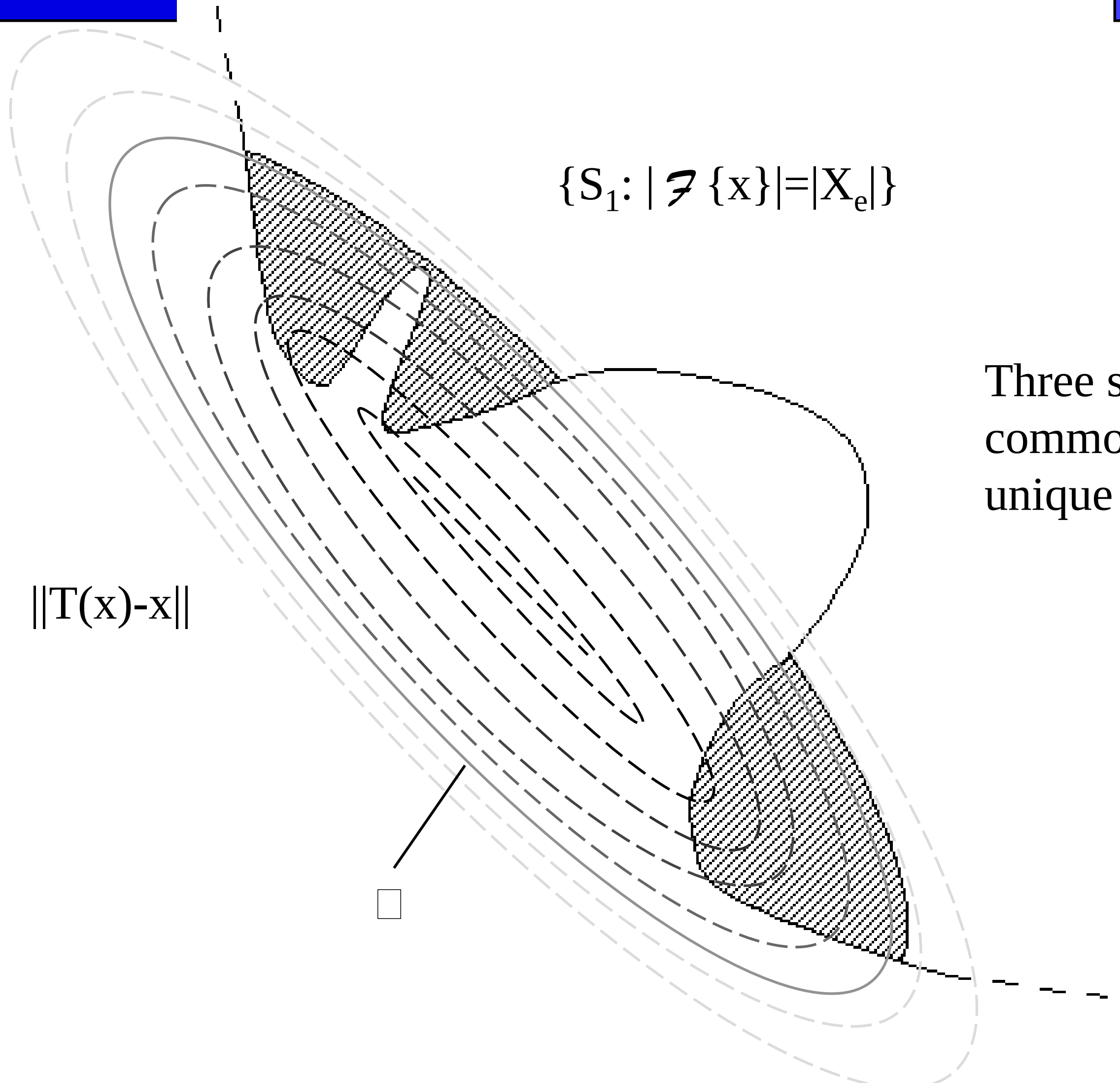
NO

YES

Observed Intensities & assigned phases (Global Search)

Feasible Solution

Multiply-Connected Feasible Set

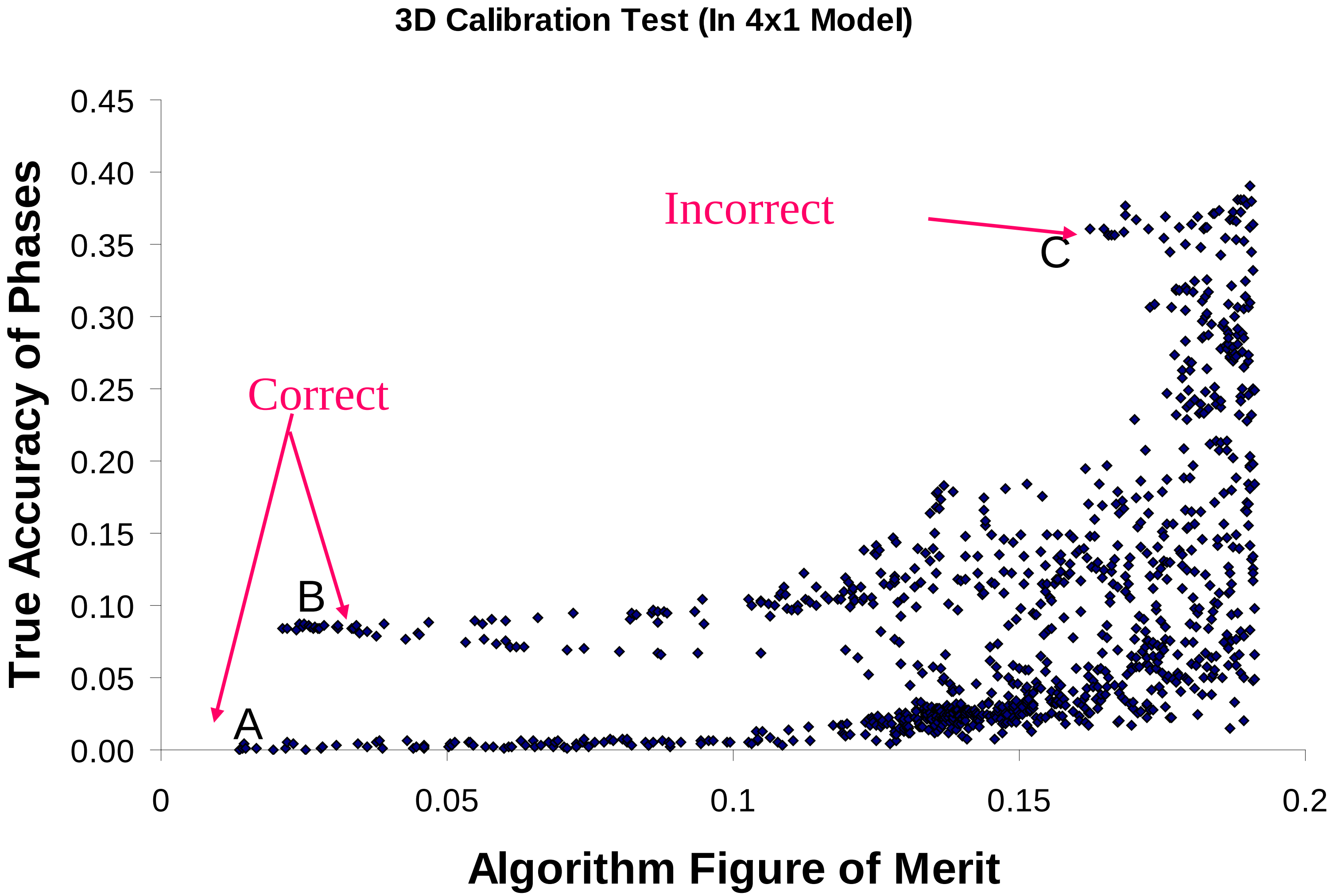


Three shaded regions common to both sets, 3 unique solutions

Typical results

Df = phase error

$$\frac{S|U(k)|\{1-\cos(Df)\}}{S|U(k)|}$$


3D Calibration Test (In 4x1 Model)

Incorrect

Correct

Where do constraints come from

- Physical nature of experiment
  - Limited beam or object size
- Physical nature of scattering
  - Atomic scattering
- Statistics & Probability
  - Minimum Information/Bias = Maximum Entropy

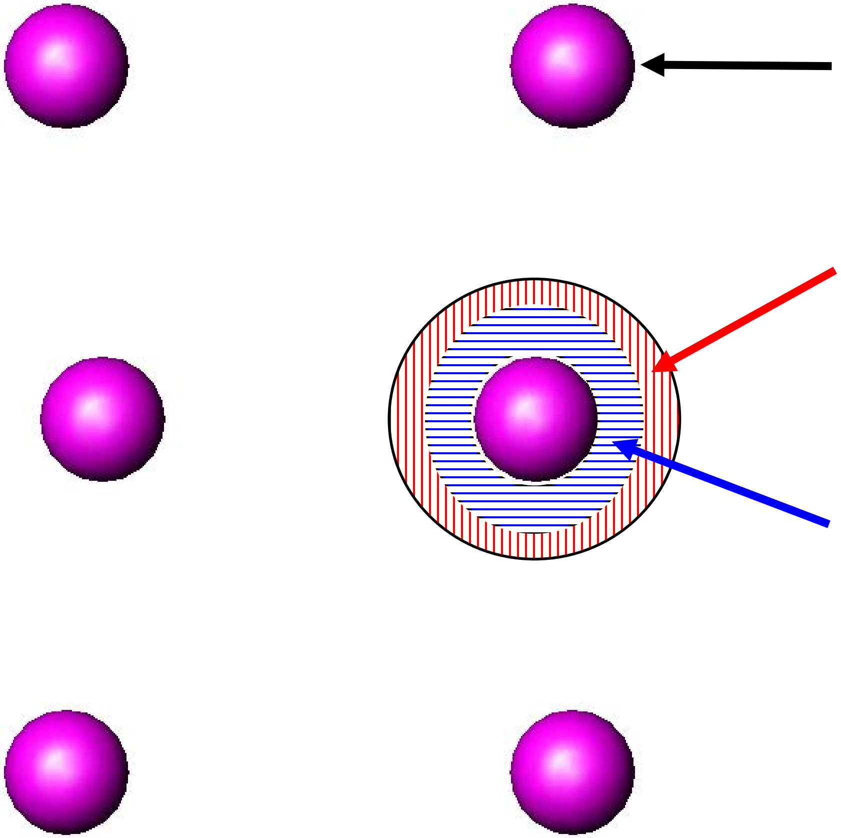
Types of Constraints

- Convex – highly convergent
  - Multiple convex constraints are unique
- Non-convex – weakly convergent
  - Multiple non-convex constraints may not be unique

Other Constraints

Convex	Non-Convex
Positivity (weak)	Presence of Atoms
Atoms at given positions	Bond Lengths
Least bias (MaxEnt)	Interference $A(k)=\|B(k)+Known(k)\|^2$
Intensities & errors $\equiv \chi^2$	Anti-bumping
Statistics (e.g. Σ_2)	Bond angles
Support for gradient	
Symmetry	

Atomistic Constraints



r(r) known (convex if position is known)

Bonding –
another atom

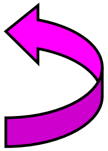
Bumping
r(r)=0

Why we don’t need all the data

- The constraints, e.g. support & atomistic, generate both amplitude & phase estimates.
- The amplitudes and phases of the unmeasured points must also be consistent with the constraints.
- Hence it is often (not always) possible to recover to a good approximation the “missing cone” values

Solvent Flattening & oversampled data

- Constraint: part of real-space x is zero
- Convex constraint
- Iteration
    - x = 0, part of map
    - $|X| = |X_{\text{observed}}|$



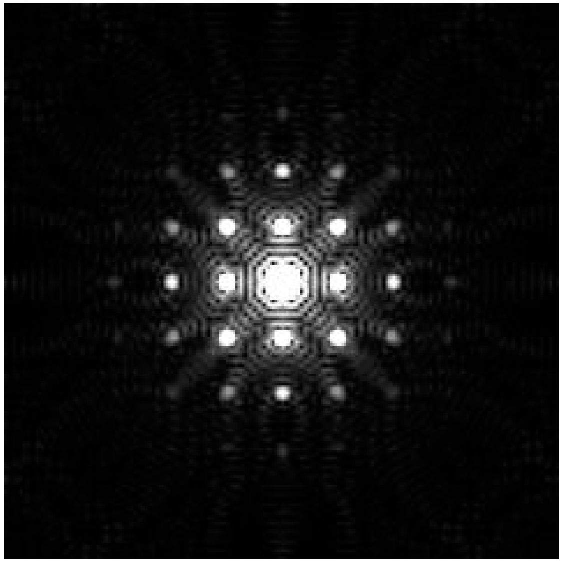
- A version of classic Gerschberg-Saxton algorithm, and “clean” algorithm

Exact Cases

- Suppose we have N pixels, and N/2 are known to be zero (compact support)
- Wave is described by N/2 moduli, N/2 phases (for a real wave) in reciprocal space
- Unkowns – N ; measurements N/2 ; contraints N/2
- Problem is in principle fully solveable

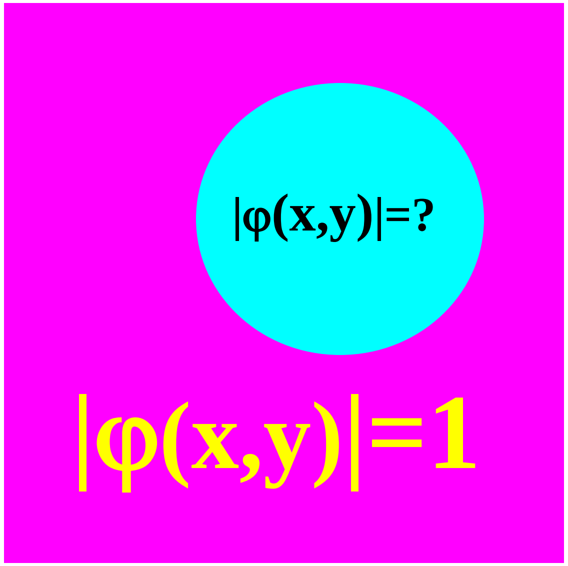
(It can be shown to be unique in 2 or more dimensions, based upon the fundamental theorem of algebra)

Phase Recovery for a Small Particle



True diffraction pattern
for small particle model
(Non-Convex Constraint)

+



$|\varphi(x,y)|=?$

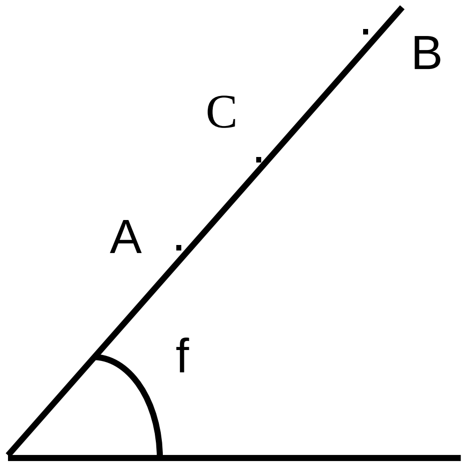
$|\varphi(x,y)|=1$

Convex Support
Constraint

= ?

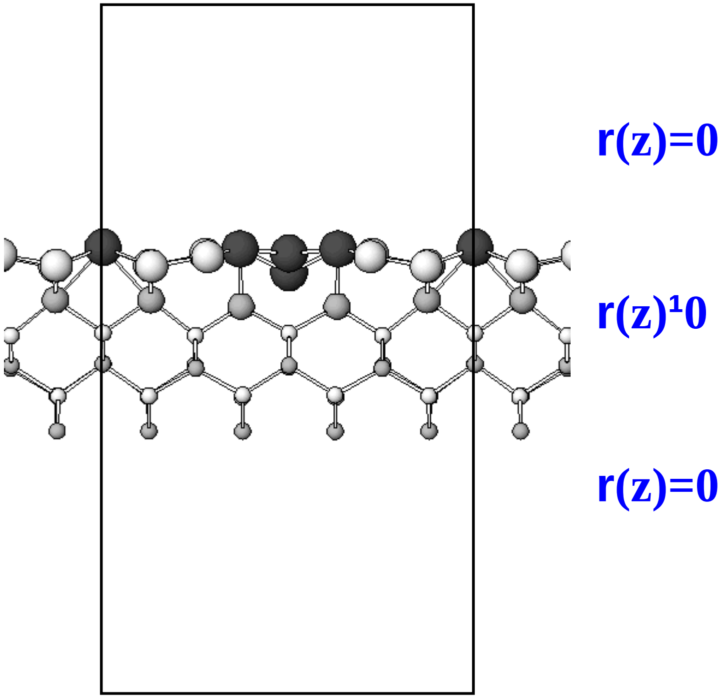
Convex Set for unmeasured |U(h,k,l)|

- Phase of U(h,k,l) can be estimated from other reflections
- Set of U(h,k,l) with a given phase is convex
- Hence |U(h,k,l)| is well specified and can be (approximately) recovered
- Remember, phase is more important than amplitude



3D-Support Constraint

- Displacements decay as $(\alpha+z)\exp(-qz)$ into bulk[1]
- Consider only non-bulk spots
- Real space constraint
  - ❑ $\rho(z)=0$ away from surface
- Convex constraint

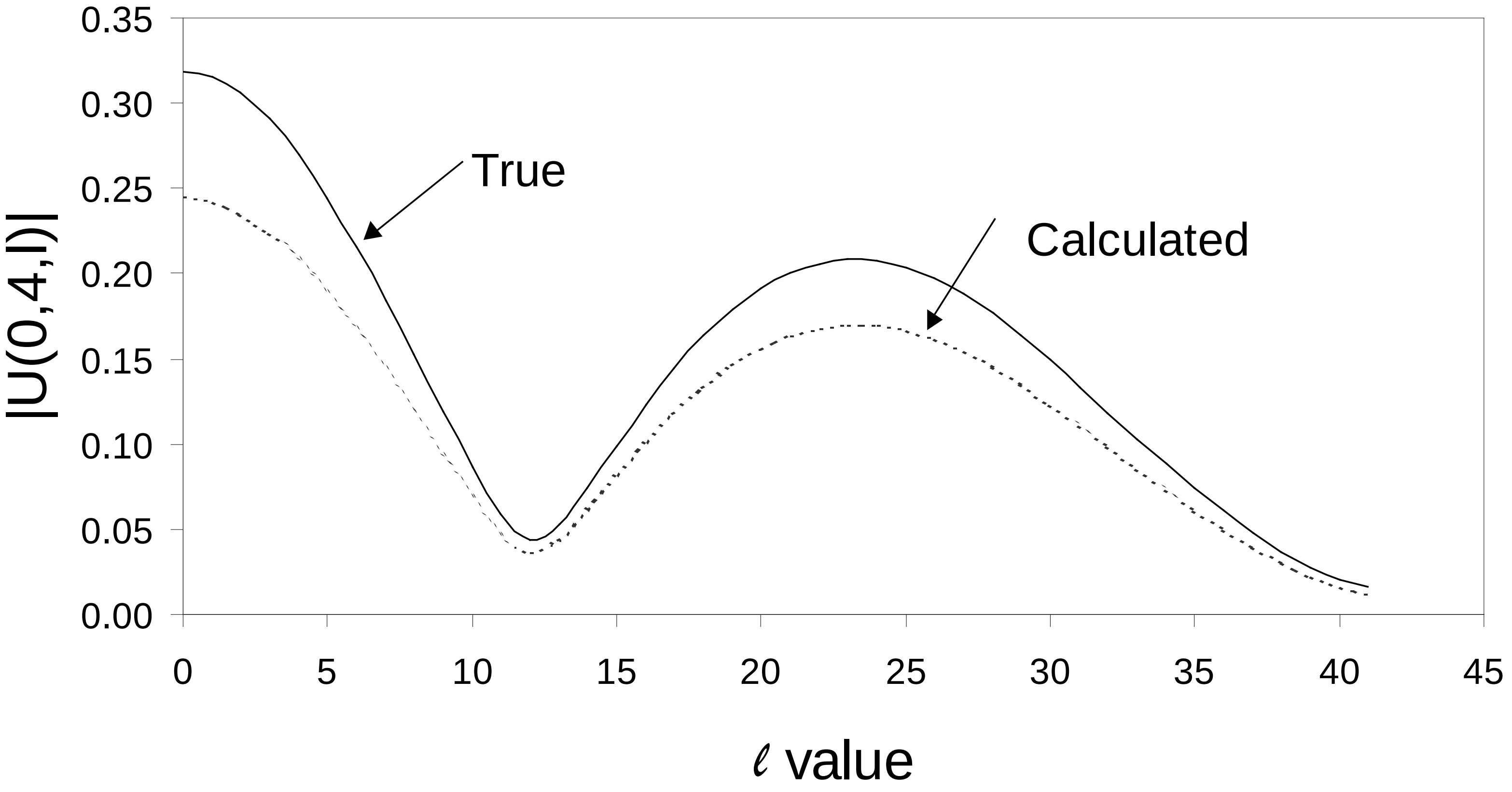


---

[1] Biharmonic expansion of strain field, Surface Science 294, 324 (1993)

Unmeasured Reflections

Recovery of Unmeasured Reflections



Classical Approach

- For periodic object (not finite one)
- Based upon probabilities
- Complex (empirical) FOM
- Uses central-limit theorem
    - Can use other probability theorems
    - Not all work!

Basic Idea

- There are certain relationships which range from exact to probably correct.
- Simple case, Unitary Sayre Equation, 1 type

$$F(k) = \sum_l f(k)\exp(2\pi i k.r_l)$$

- Divide by N, #atoms & f(k), atomic scattering factors

$$U(k) = 1/N \sum_l \exp(2\pi i k.r_l); \; u(r) = 1/N \sum_l \delta(r - r_l)$$

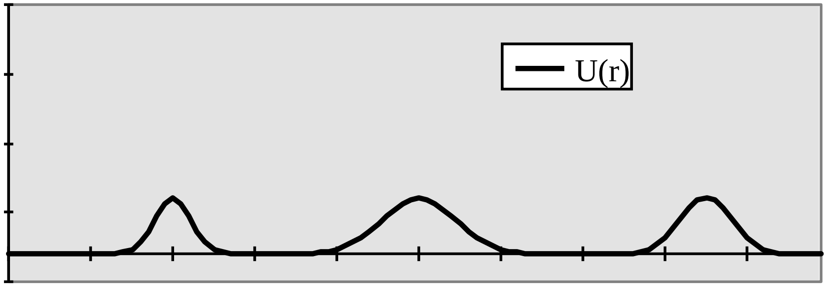
$$u(r) = Nu(r)^2$$

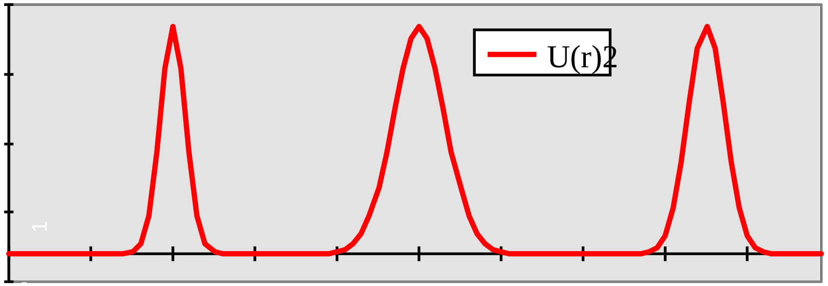
Constraint

Real/Reciprocal Space

$$U(h) \approx \sum_{k} U(k)U(h-k)$$

$$U(r) \approx U(r)2$$





Reinforces strong (atom-like) features

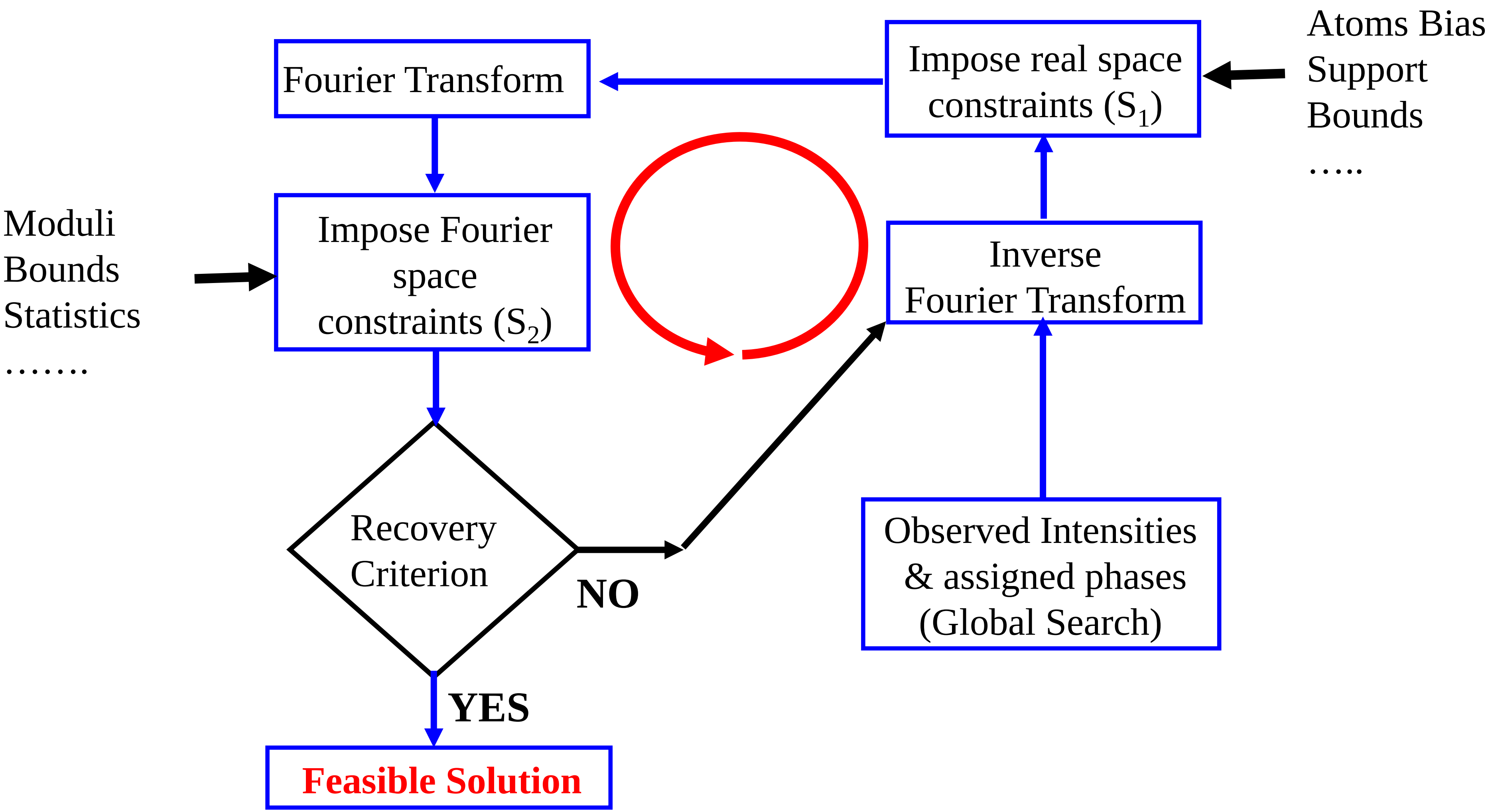
Tangent Formula

- If $U(r) = U(r)^2 = U'(r)$
- Important part is the phase
- $U(u) = |U(u)|\exp(i\theta)$; we know $|U(u)|$ but not θ
- $\exp(i\theta) = \exp(i\theta')$; $\mathrm{Tan}(\theta) = \mathrm{Tan}(\theta')$
- Replace old θ by new one

Tangent Formula

1. Initial $\rho(r)$
2. Project onto “Real Space Constraint” $\rho^2(r)$
3. FFT
4. Project amplitudes onto Observed
5. FFT

Algorithm Overview (Gerschberg-Saxton)



Fourier Transform

Impose real space constraints (S_1)

Atoms Bias
Support
Bounds
…..

Moduli
Bounds
Statistics
…….

Impose Fourier space constraints (S_2)

Inverse Fourier Transform

Recovery Criterion

NO

Observed Intensities & assigned phases (Global Search)

YES

Feasible Solution

Cochran Distribution (Σ_2): I

- Definition: $U(k) = (1/N)\sum_m \exp(2\pi i k.r_m)$
- Consider the product

$$NU(k-h)U(h) = (1/N)\sum_m \exp(2\pi i k.r_m)\sum_l \exp(2\pi i h.(r_m - r_l))$$

- If the atoms are randomly distributed,

$$\left\langle \sum_l \exp(2\pi i h.(r_m - r_l)) \right\rangle = 1$$

(exponential terms average to zero if $m \neq l$)

$$N\langle U(k-h)U(h)\rangle = (1/N)\sum_m \exp(2\pi i k.r_m) = U(k)$$

Cochran Distribution: II

- Consider next

Average is zero

$$\left|NU(k-h)U(h)-U(k)\right|^2$$

Known

$$=|U(k)|^2+N^2|U(k\text{-}h)U(h)|^2$$

$$\text{-}2N|U(k)U(k\text{-}h)U(h)|\cos\left(\phi(k)-\phi(k-h)-\phi(h)\right)$$

Known

Average must be 2nπ

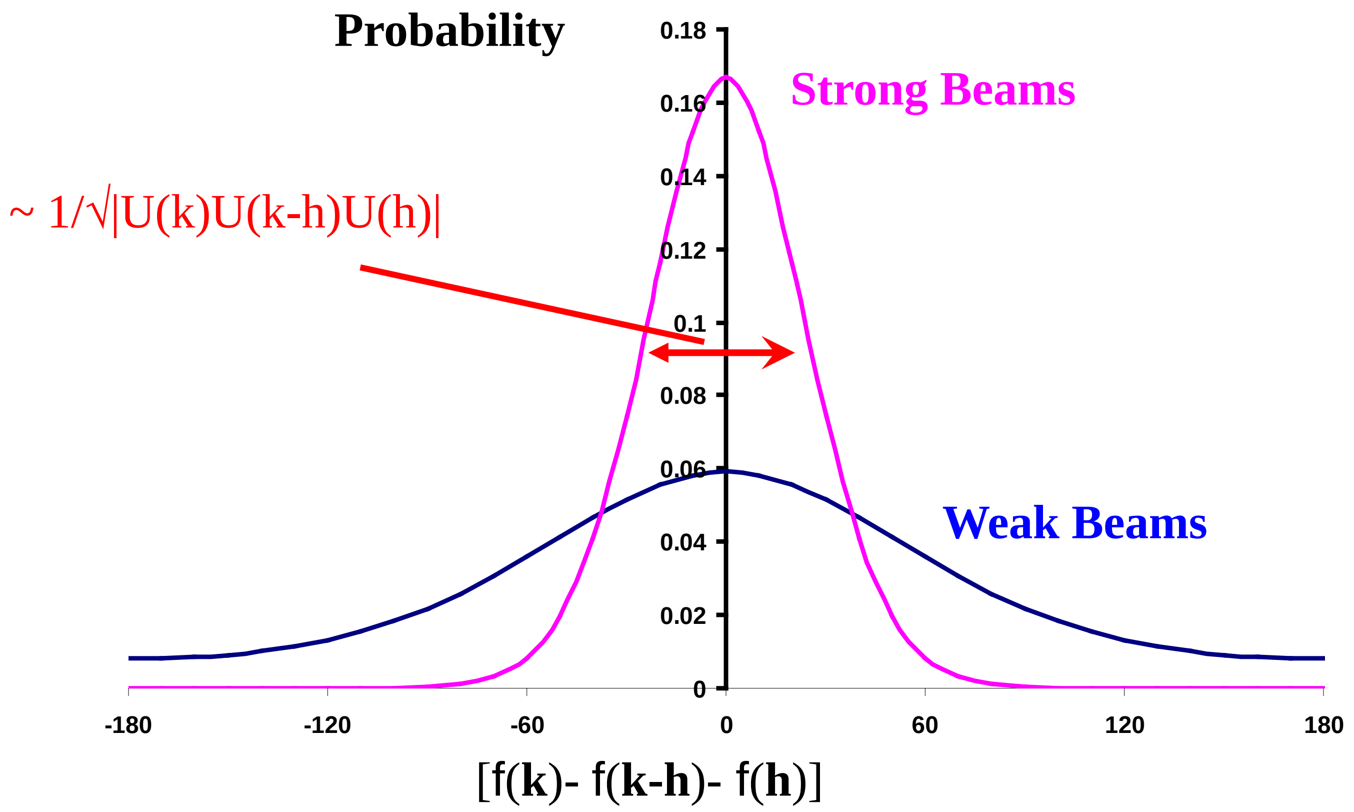
Cochran Distribution: III

- We have a distribution of values. The Central Limit theorem: all distributions tend towards Gaussian. Hence a probability:
- $P(U(\mathbf{k}) - NU(\mathbf{k}-\mathbf{h})U(\mathbf{h}))$

$\sim C\exp(-|U(\mathbf{k}) - NU(\mathbf{k}-\mathbf{h})U(\mathbf{h})|^2)$

$\sim C\exp(2|U(\mathbf{k})U(\mathbf{k}-\mathbf{h})U(\mathbf{h})|\cos[\phi(\mathbf{k}) - \phi(\mathbf{k}-\mathbf{h}) - \phi(\mathbf{h})])$

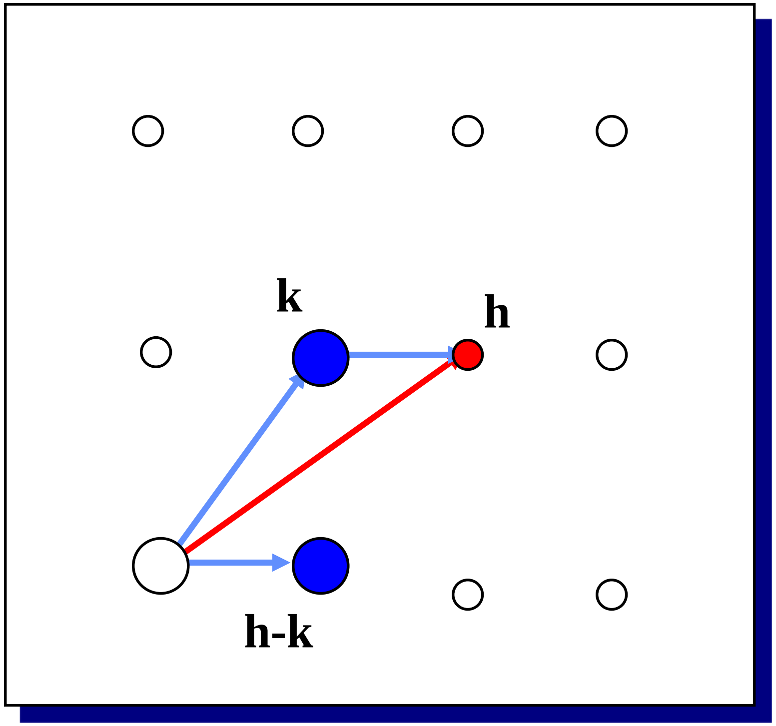
Form of Distribution



~ 1/√|U(k)U(k-h)U(h)|

Note: this is more statistics than the presence of atoms

Σ_2 Triplet



For reflections **h-k**, **k** and **h**:

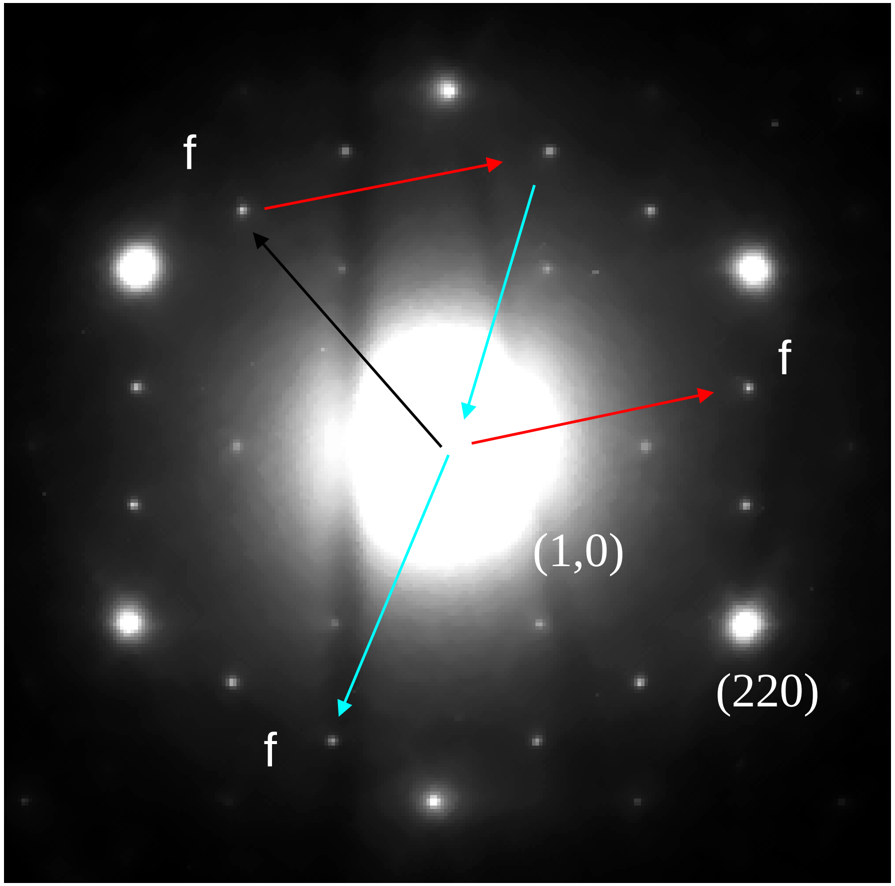
$$\phi(\mathbf{h}) \approx \phi(\mathbf{k}) + \phi(\mathbf{h\text{-}k})$$

W. Cochran (1955). Acta. Cryst. **8** 473-8.

● = known structure amplitude and phase

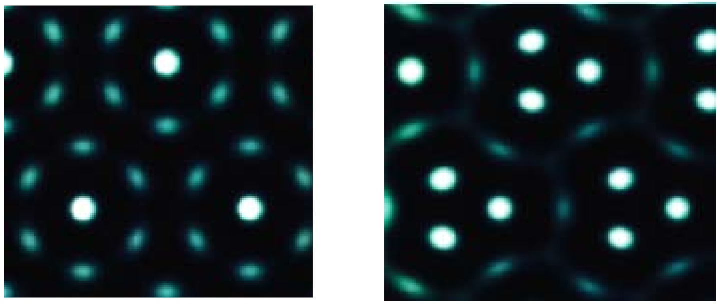
● = known structure amplitude and unknown phase

Example: Si(111) √3x√3 Au

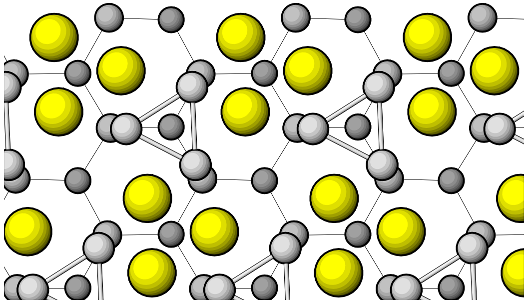


Only one strong reflection

- $3\phi \sim 360n$ degrees
- ϕ=0,120 or 240
- ϕ=0 has only 1 atom
- 120 or 240 have 3



Other information
3 Au



TED data

- Why can DM sometimes (not always) work with transmission electron diffraction data?
- Answer: statistical projection of dynamical intensities onto kinematical is valid in many cases

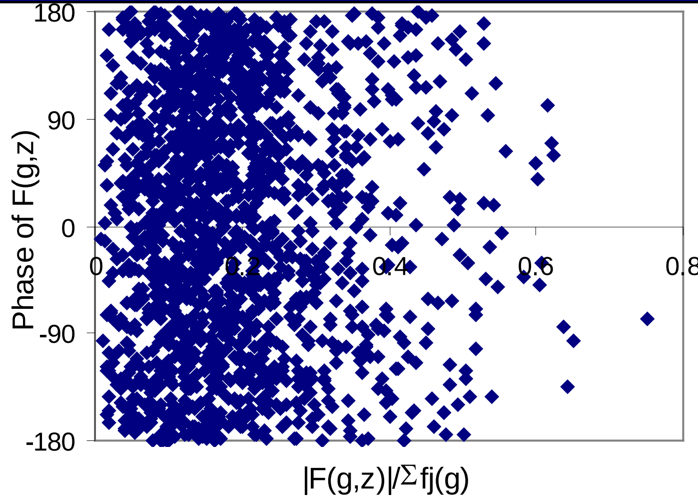
Acta Cryst A54, 591 (1998); A56, 458 (2000) ; A57, 231 (2001)

Statistics in a 1s model

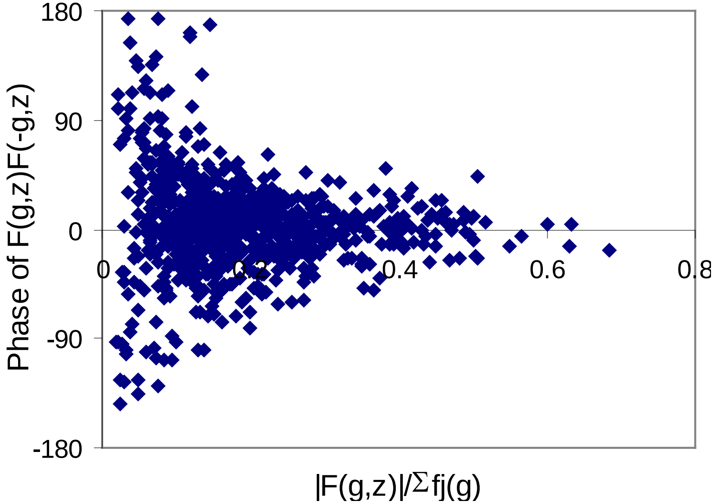
Kinematical	Dynamical
Σ_o:	Σ_o:
$\phi(g)+\phi(-g) = 0$	$\phi(g)+\phi(-g) \sim \omega$
$\lvert F(g)\rvert = \lvert F(-g)\rvert$	$\lvert F(g)\rvert \sim \lvert F(-g)\rvert$
Σ_2:	Σ_2:
$\phi(g)+\phi(h-g)+\phi(-h) \sim 2n\pi$	$\phi(g)+\phi(h-g)+\phi(-h) \sim \zeta$

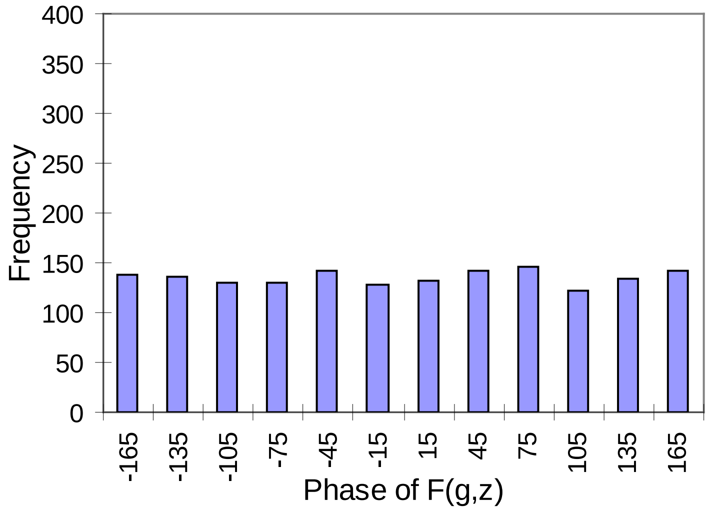
Kinematical Theory for electrons is “wrong”, but statistically “right” in many cases, ***BUT NOT ALL*** (see C252)

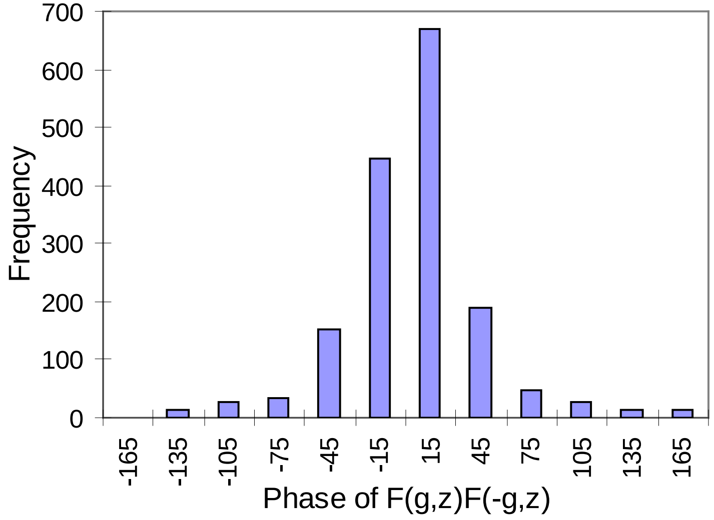
Σ_0 dynamical distribution



non-symm
5.264nm

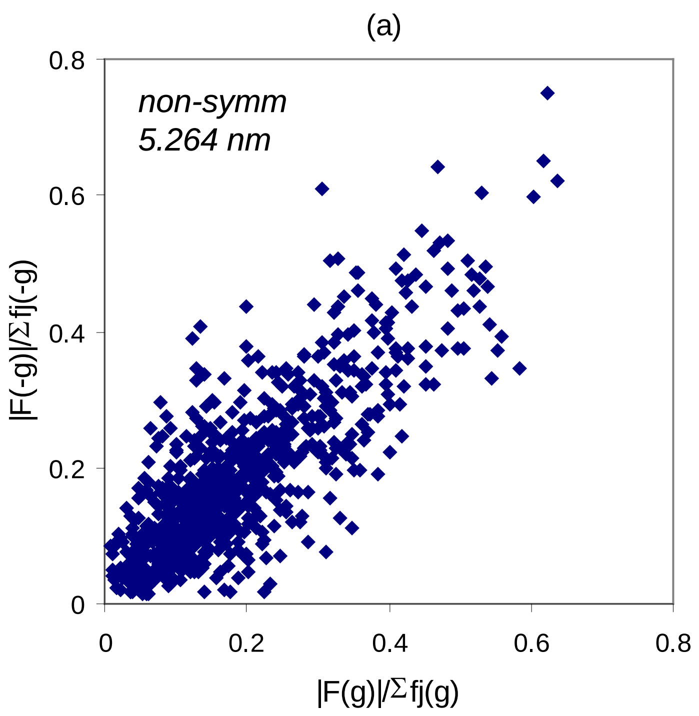


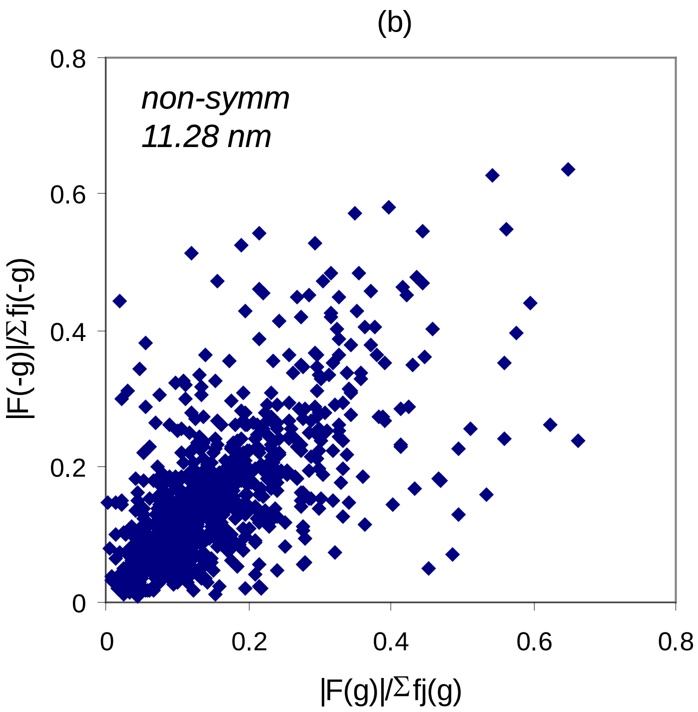




Σ_0 dynamical distribution

Freidel symmetry is statistical





Inequalities

$|\text{Sum } a_i b_i|^2 < \text{Sum } |a_i|^2 \text{ Sum } |b_i|^2$

$a_i = 1/\text{sqrt}(N)\cos(2\pi k r_i)$; $b_i = 1/\text{sqrt}(N)$

$\text{Sum } a_i b_i = U(k)$

Sum 1/N = 1 for N atoms

$\text{Sum } |a_i|^2 = 1/N \text{ Sum } \cos(2\pi k r_i)^2$

$= 1/2N \text{ Sum } (1+\cos(2\pi[2k]r_i))$

$= ½ + U(2k)$

Hence $U^2(k) < ½ + U(2k)/2$

If U(k) is large – can set U(2k)

Quartets

- Phase relationships involving 4 terms for weak reflections
  - Positive and Negative
  - Rarely useful with TEM

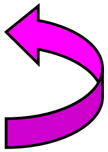
Maximum Entropy

- Constraint on amplitudes – non-convex
  - $\Sigma \ \{I_{obs}(k)-|F_e(k)|^2\}^2/\sigma(k)^2 < \text{tolerance}$
- Bregman projections[1] are generalizations of projections for any convex function
- Positivity constraint via a Bregman projection ***IS***[1] maximization of entropy
- Overall problem is non-convex, requires a multisolution approach

[1] Y. Censor and S. A. Zenios, Parallel Optimization Theory, Algorithms and Applications, Oxford University Press, 1997

Shake & Bake

- Set of potentials due to non-overlapping atoms
  - ❑ $V(r) = \Sigma_j \, v(r-r_j)$
- Not convex
  - ❑ $U(r) = \lambda\Sigma_j \, v(r-r_j) + (1-\lambda) \, \Sigma_j \, v(r-r_j')$
- Iteration
  - ❑ Project (replace) x by atoms
  - ❑ $|X| = |X_{observed}|$

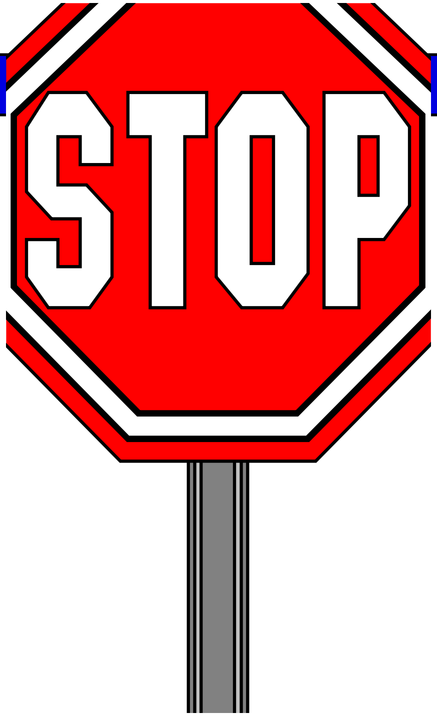


Iterate

Charge-Flipping

- Potential is positive, and always “peaky” or zero
- Small noise (+ve & -ve) is “wrong”
- Flip the sign of anything < tol
  - Comparable to an overprojection
  - Also relates to compressive sensing

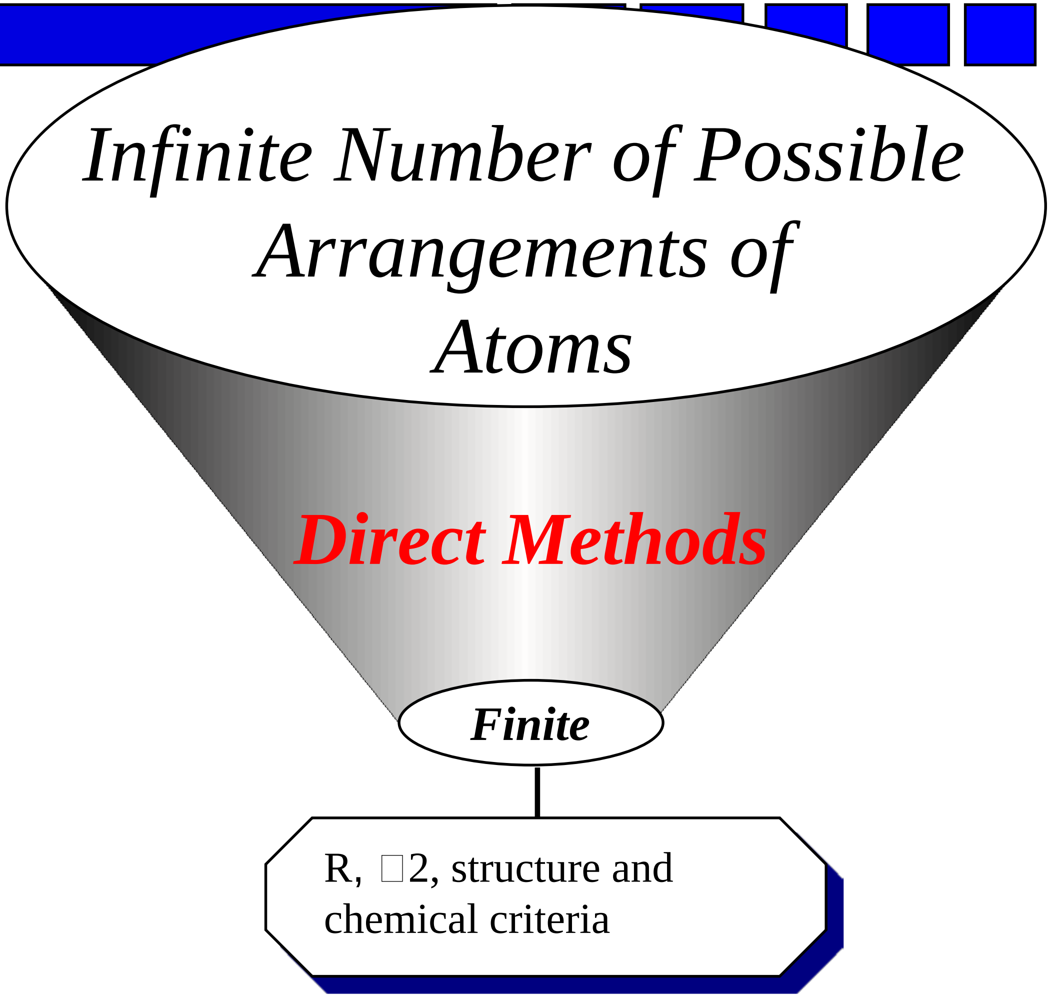
Caveat: Not Physics



This is probability, not an exact "answer"

All one can say is that the "correct" answer will be among those that are found

How is it implemented?



Infinite Number of Possible Arrangements of Atoms

Direct Methods

Finite

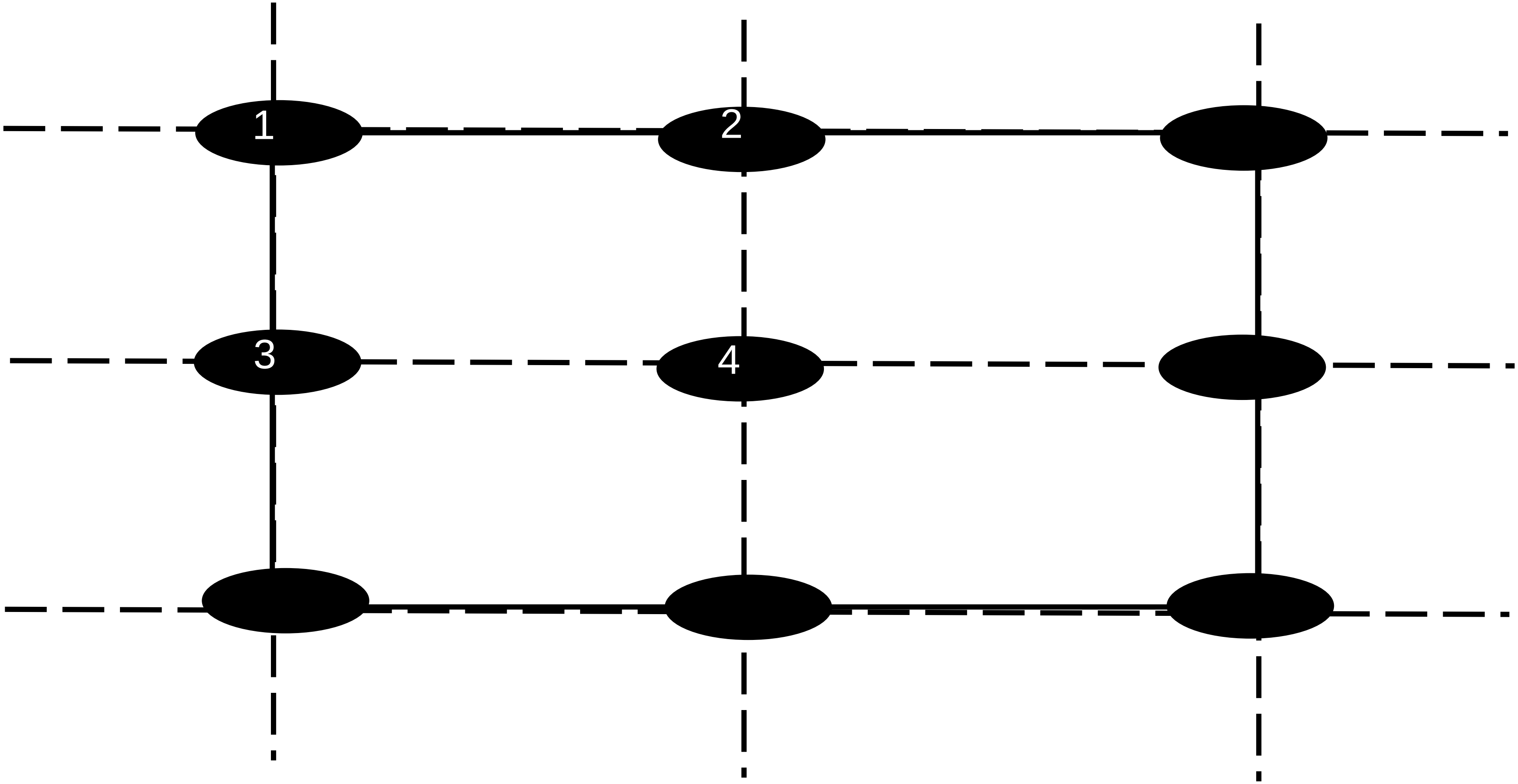
R, □2, structure and chemical criteria

Implementation

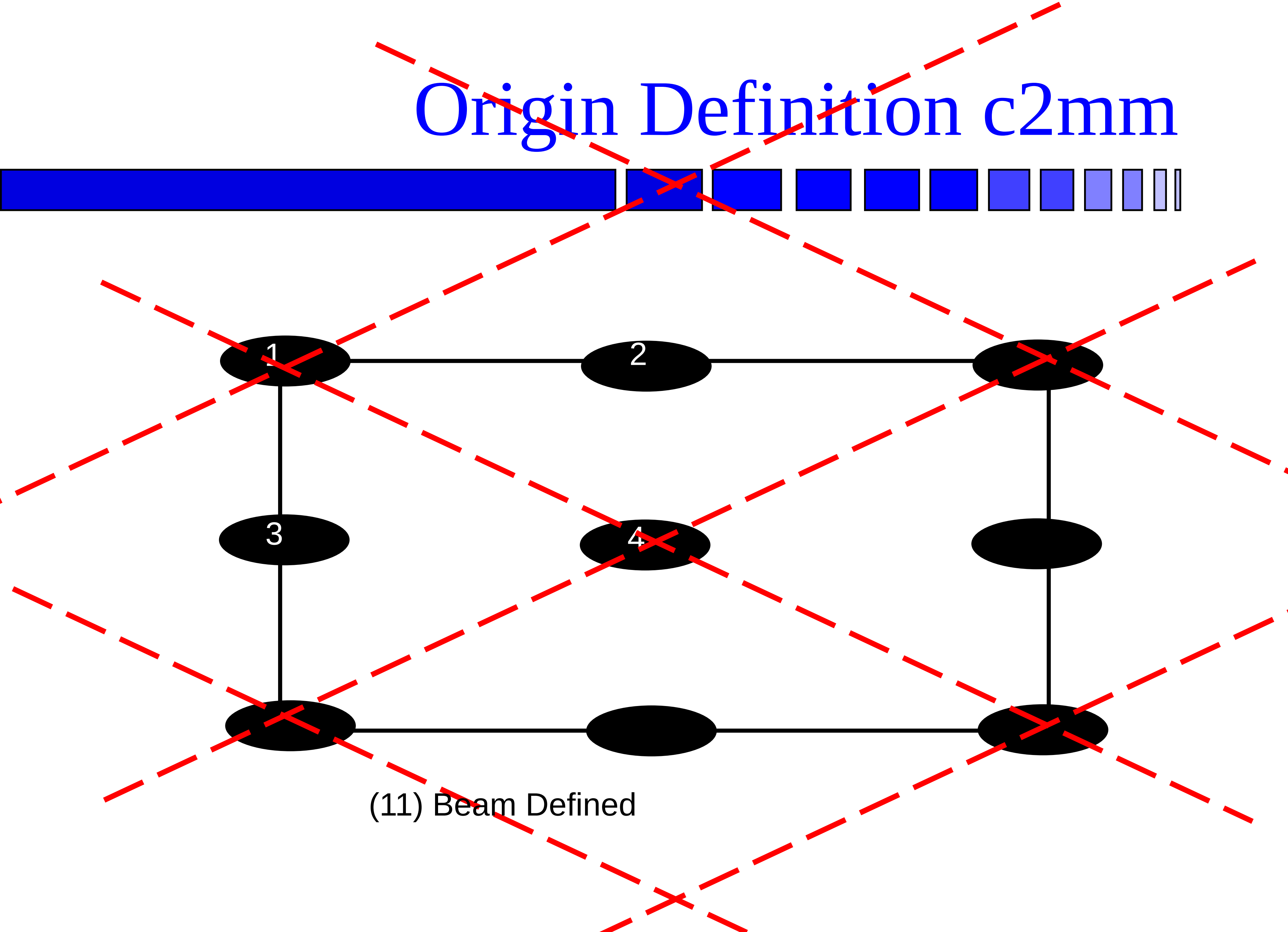
1. Guess phases for some reflections
2. Generate from these phases for others and improved phases for initial set
3. Test consistency of predicted amplitudes and phases
4. Iterate, so long as consistency is improving

Note: permuting phases has lower dimensions than permuting atom positions

Origin Definition c2mm

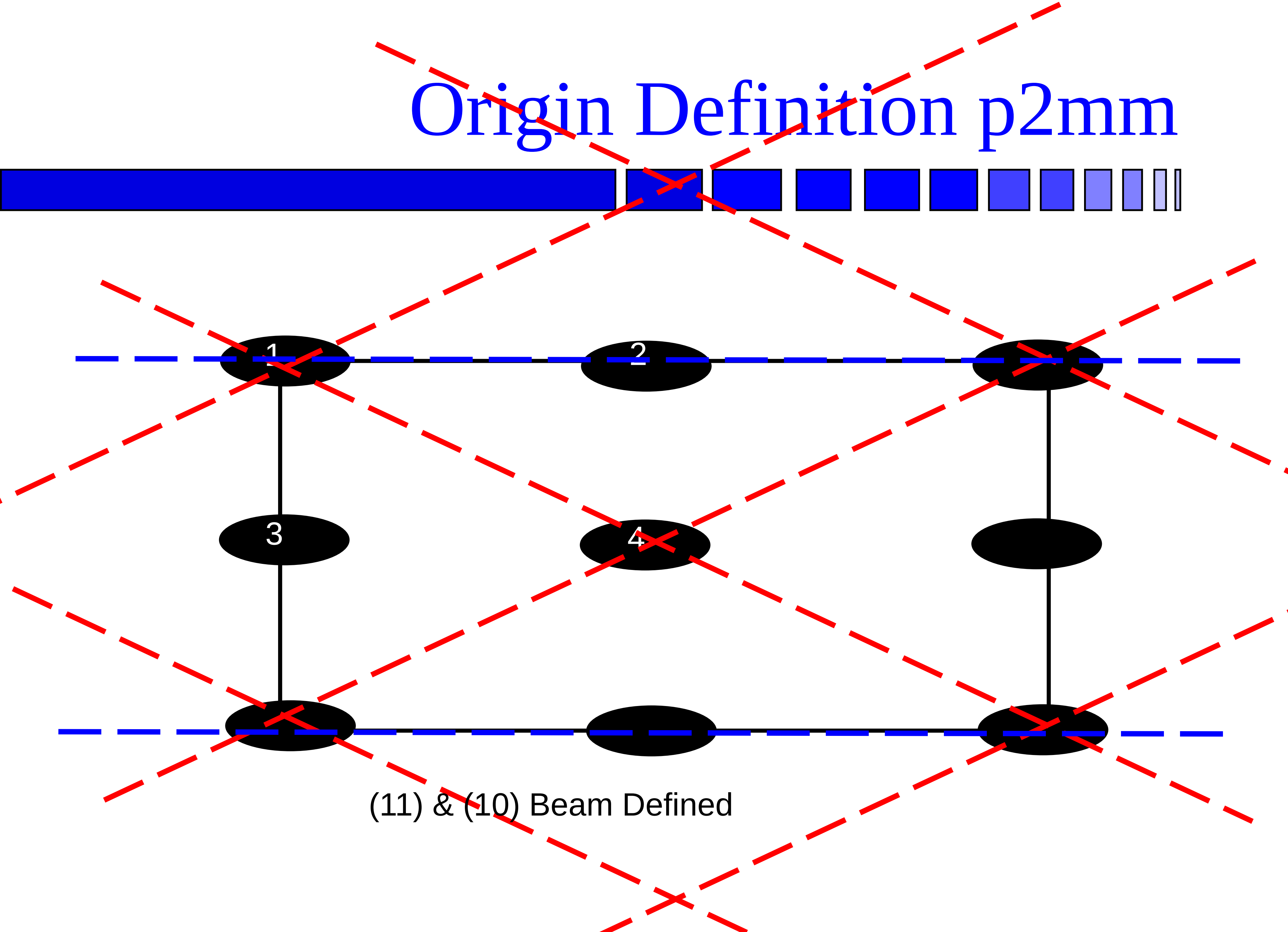


Origin Definition c2mm



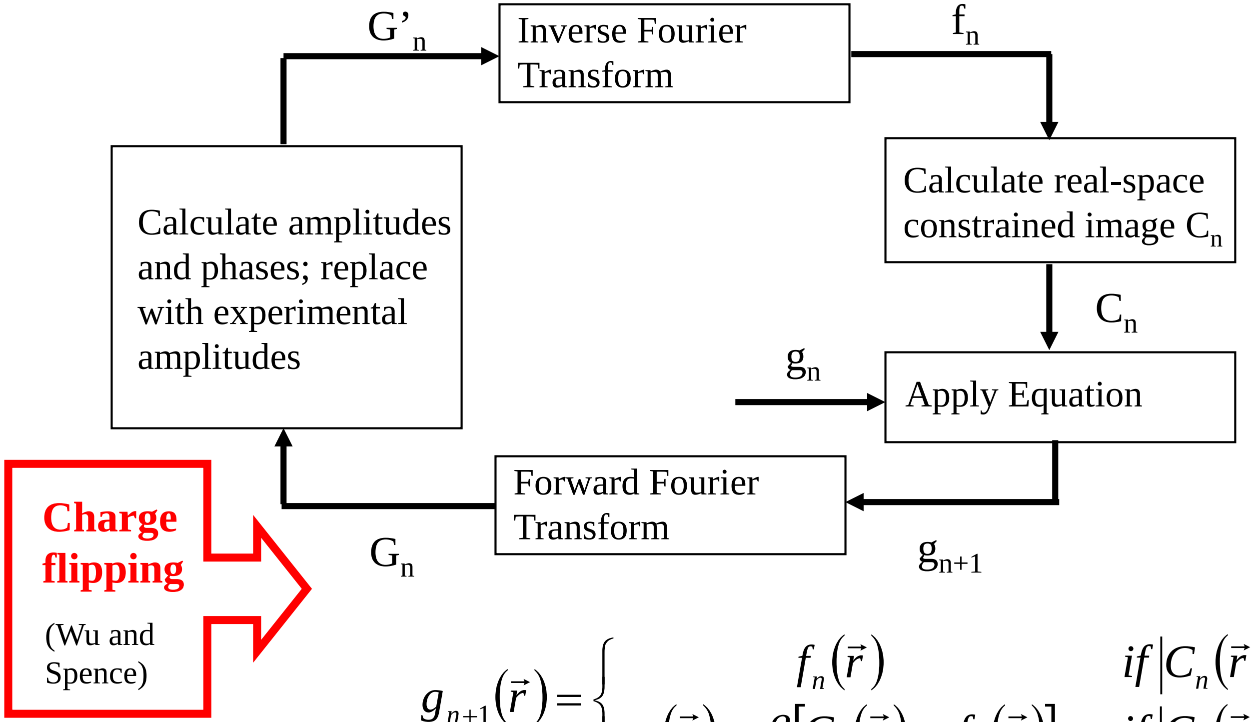
(11) Beam Defined

Origin Definition p2mm



(11) & (10) Beam Defined

The Algorithm



G'_n

Inverse Fourier Transform

f_n

Calculate real-space constrained image C_n

C_n

g_n

Apply Equation

g_{n+1}

Forward Fourier Transform

G_n

Calculate amplitudes and phases; replace with experimental amplitudes

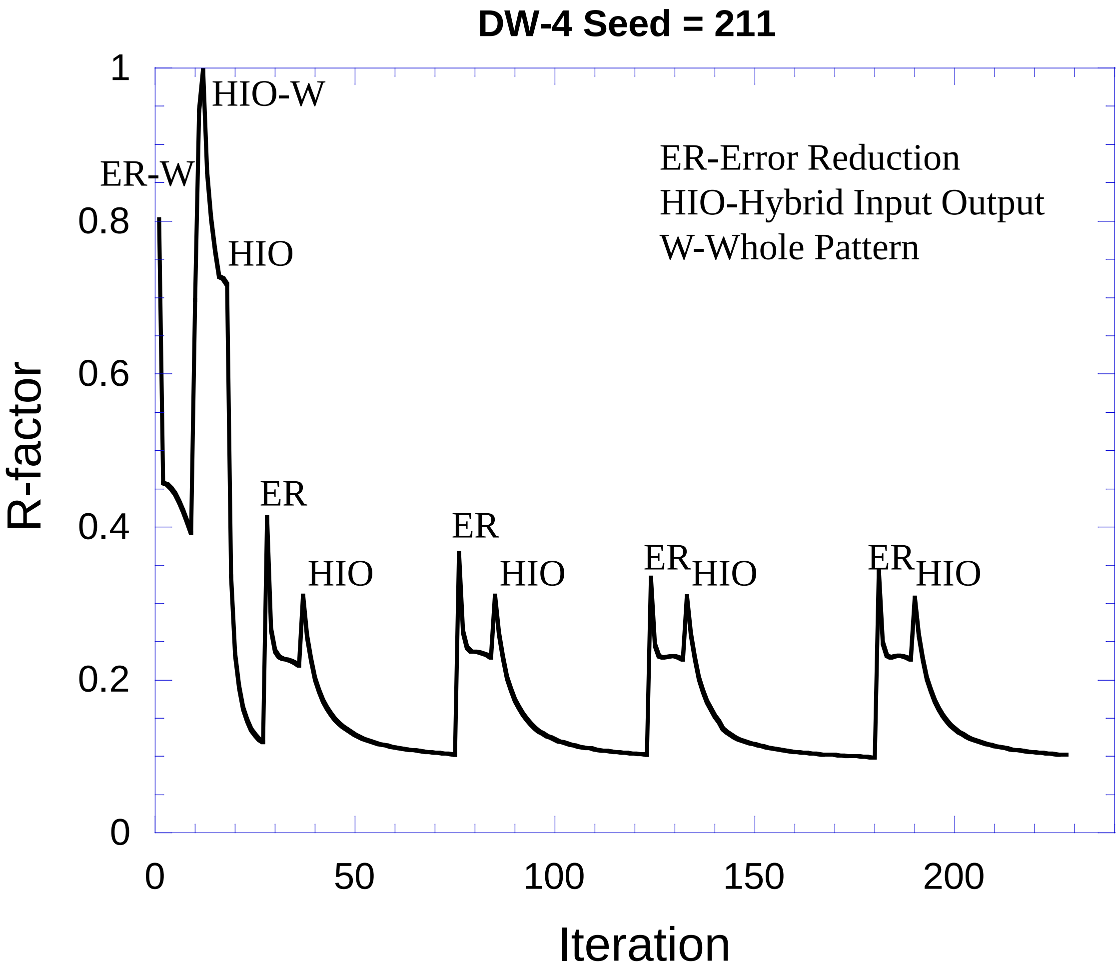
Charge flipping

(Wu and Spence)

$$g_{n+1}(\vec{r}) = \begin{cases} f_n(\vec{r}) & if\,|C_n(\vec{r}) - f_n(\vec{r})| < \varepsilon \\ g_n(\vec{r}) + \beta[C_n(\vec{r}) - f_n(\vec{r})] & if\,|C_n(\vec{r}) - f_n(\vec{r})| > \varepsilon \end{cases}$$

The flow chart of hybrid input and output algorithm for iterative phase retrieval (after Millane and Stroud, 1997).

Convergence and the Missing Central Beam

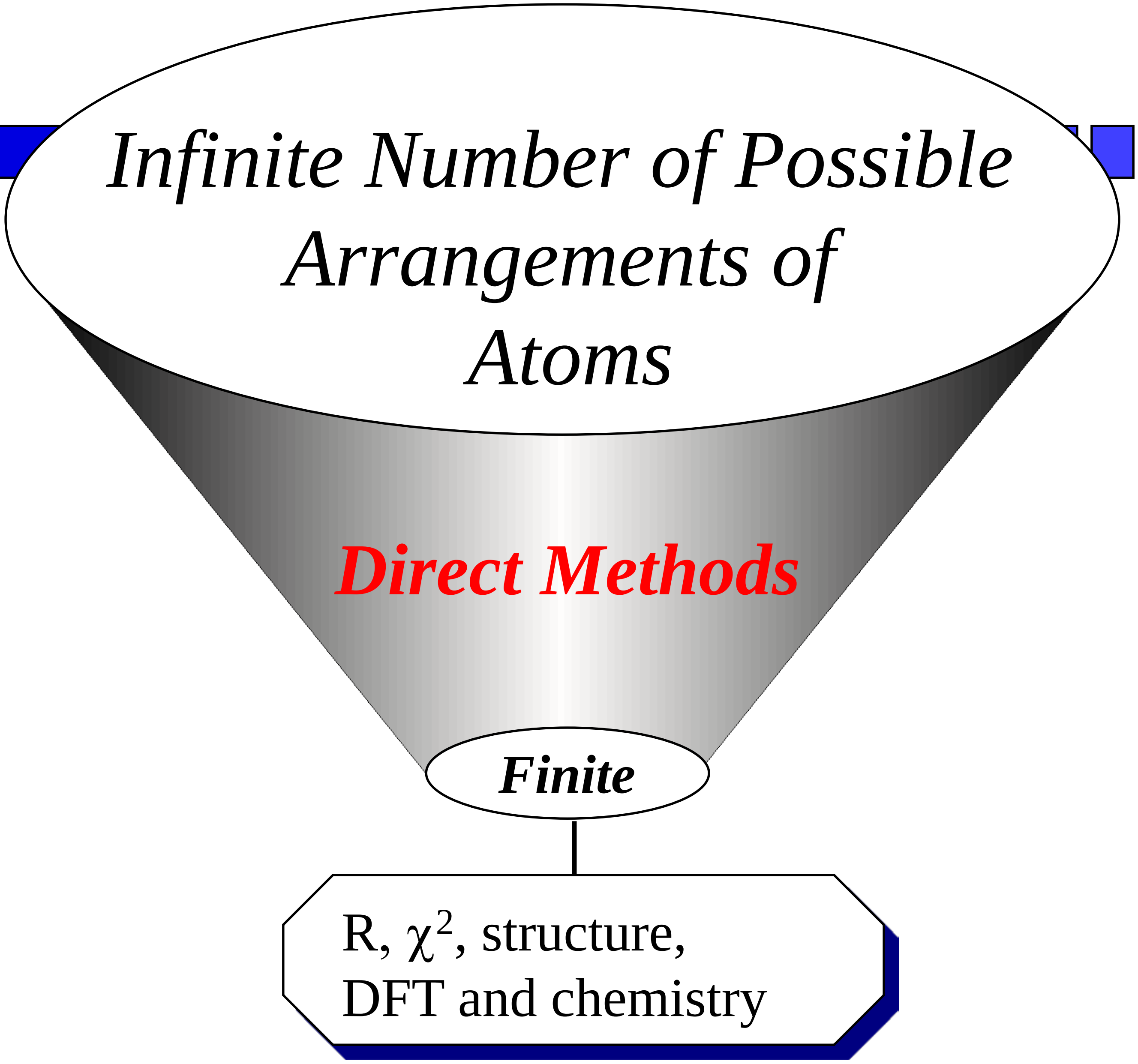


ER-Error Reduction
HIO-Hybrid Input Output
W-Whole Pattern

$$R = \frac{\sum \left\| F^{Exp} \right| - \left| F^{R} \right\|}{\sum \left| F^{Exp} \right|} 100\%$$

- Missing central beam from IP saturation
- Use low mag. TEM image
- Reconstruction start with the whole pattern
- Finish with as recorded diffraction pattern

Implementation

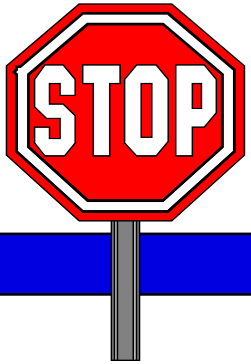


Infinite Number of Possible Arrangements of Atoms

Direct Methods

Finite

R, χ^2, structure,
DFT and chemistry



What do D.M. give us

- With the moon in the right quarter -- real space potential/charge density
- In other cases:
  - Atom positions may be wrong (0.1-0.2 Å)
  - Peak Heights may be wrong
  - Too many (or too few) atoms visible
- But... this is often *(not always)* enough to complete the structure

Chris Gimore

Sufficient Conditions

1. Kinematical Scattering
2. Intensity mapping:I(k)>I(k’) iff |F(k)|>|F(k’)
   - Small deviations from kinematical
   - Precession diffraction
3. Scattering dominated by one type of atom

Four basic elements are required to solve a recovery problem

1. A data formation model

   Imaging/Diffraction/Measurement

2. A priori information

   The presence of atoms or similar

3. A recovery criterion:

   A numerical test of Goodness-of-Fit

4. A solution method.

   Mathematical details

Patrick Combettes, (1996). Adv. Imag. Elec. Phys. 95, 155