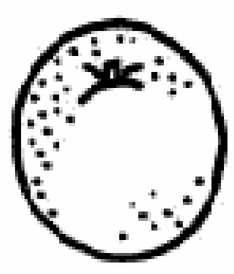
The classic phase problem

- We measure $|F(k)|$, the modulus
- $\rho(r)=\int \exp(2\pi i k.r)|F(k)|\exp(i\phi(k))dk$
- Phase information, $\phi(k)$ is lost
- Does this matter?

Phase: Apples & Oranges

FT → $A_a \exp(-i\ \phi_a)$



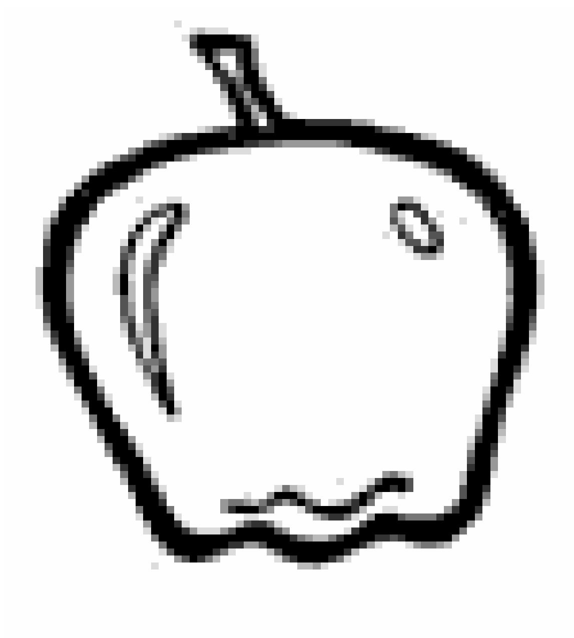
FT → $A_o \exp(-i\ \phi_o)$

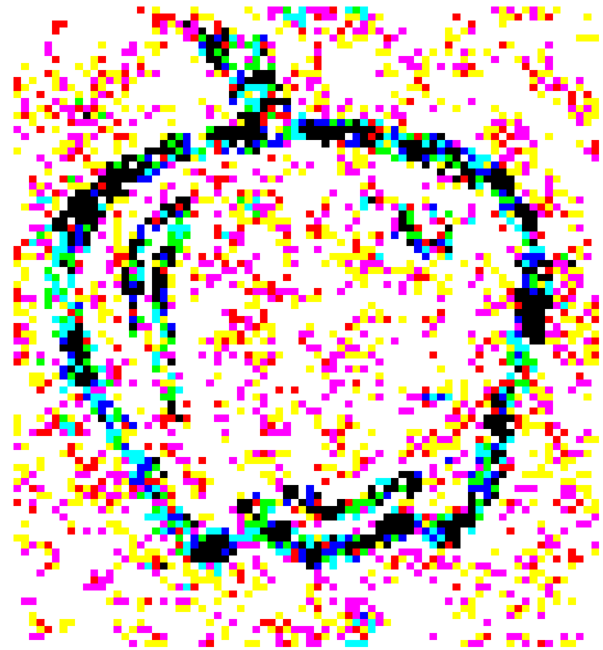
+

$A_o \exp(-i\ \phi_a)$ → IFT → { Oranle ?
Appge ?

Phase of Apple + Amplitude of Orange = ?

Phase of Apple = Apple





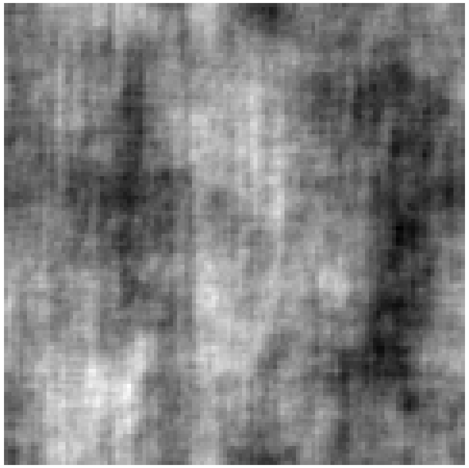
$FT^{-1}\{A_o \underline{\exp(-i\,\phi_a)}\} \Rightarrow$ Apple

Phase is <u>more</u> important than amplitude

The importance of phase information

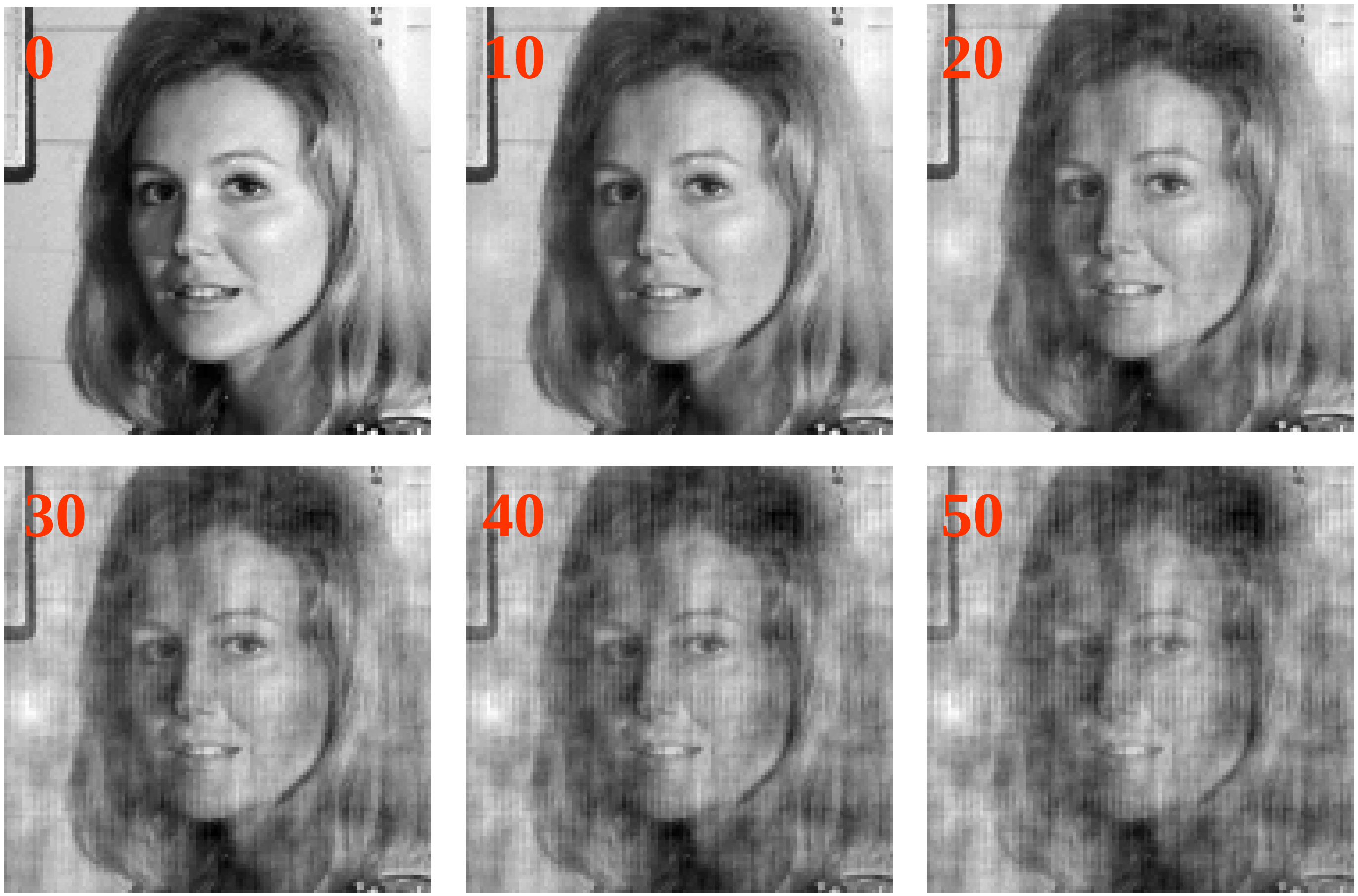
Suzy

Correct Modulus
Random Phases



Correct Phase
Random Modulus

Role of error in phases (degrees)



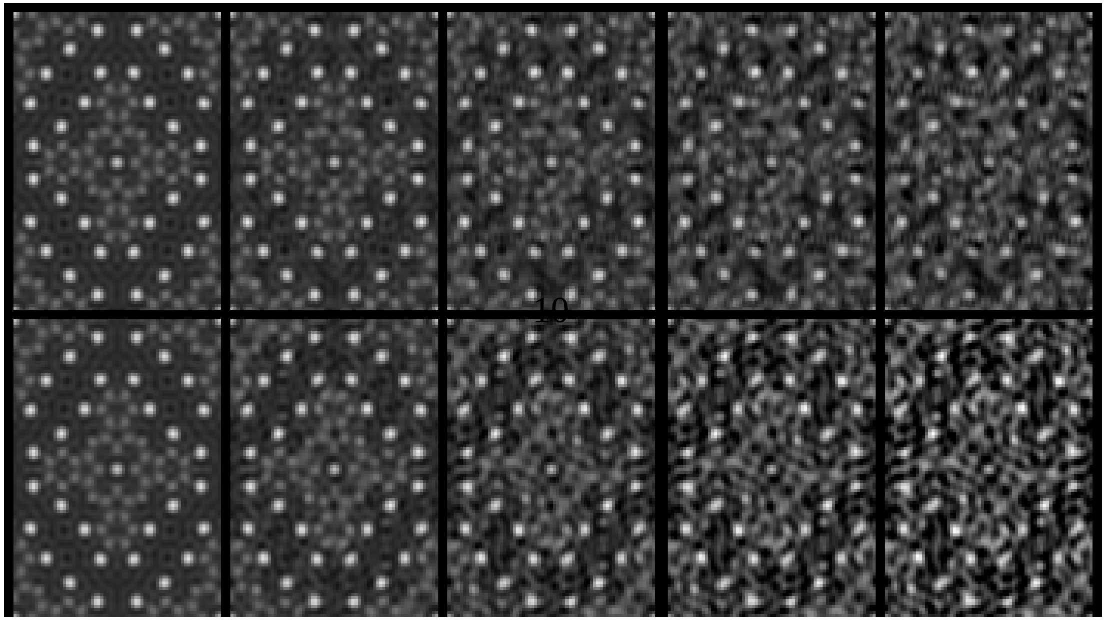
We would like to find the phases exactly, but we don’t have to

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

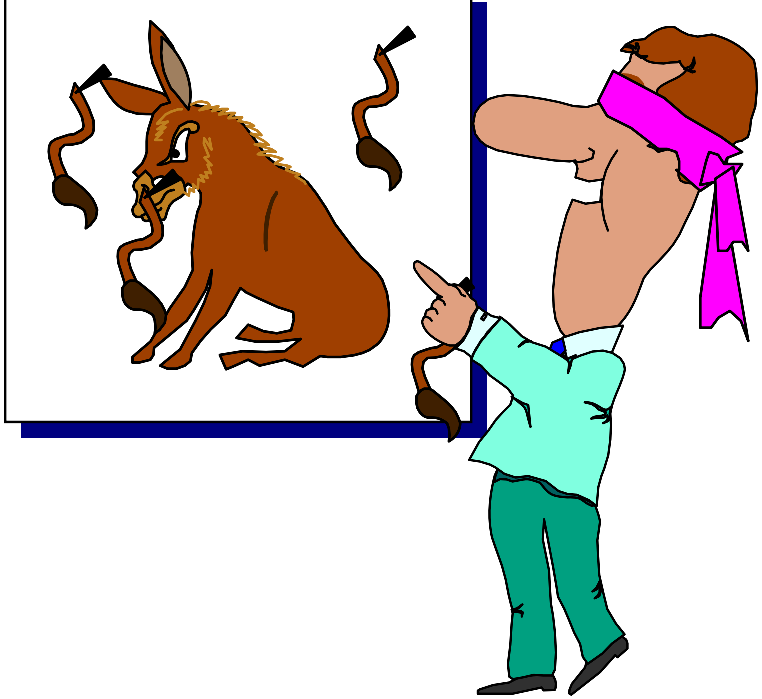
Direct Methods vs. Indirect Methods

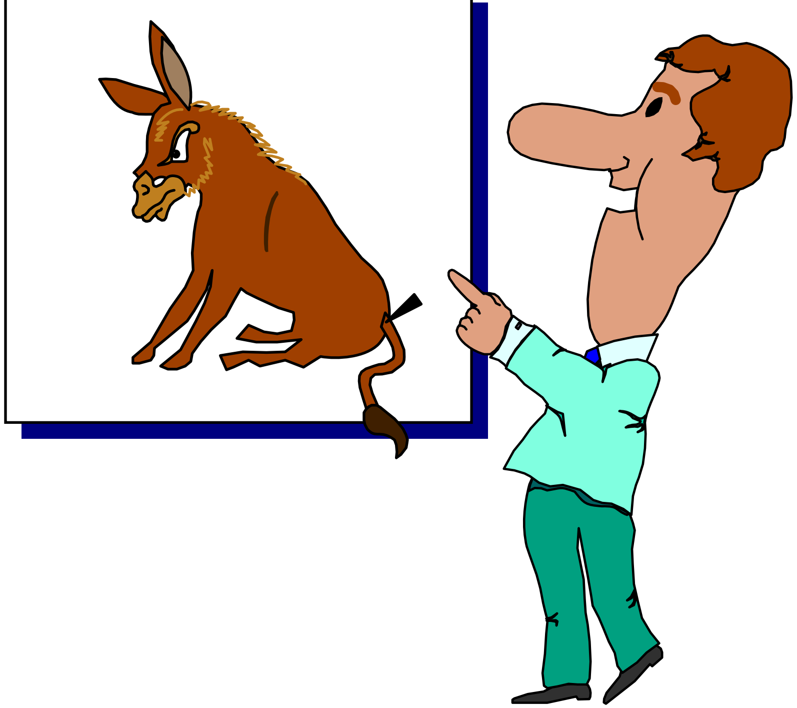
Indirect Methods:

"Trial and Error"

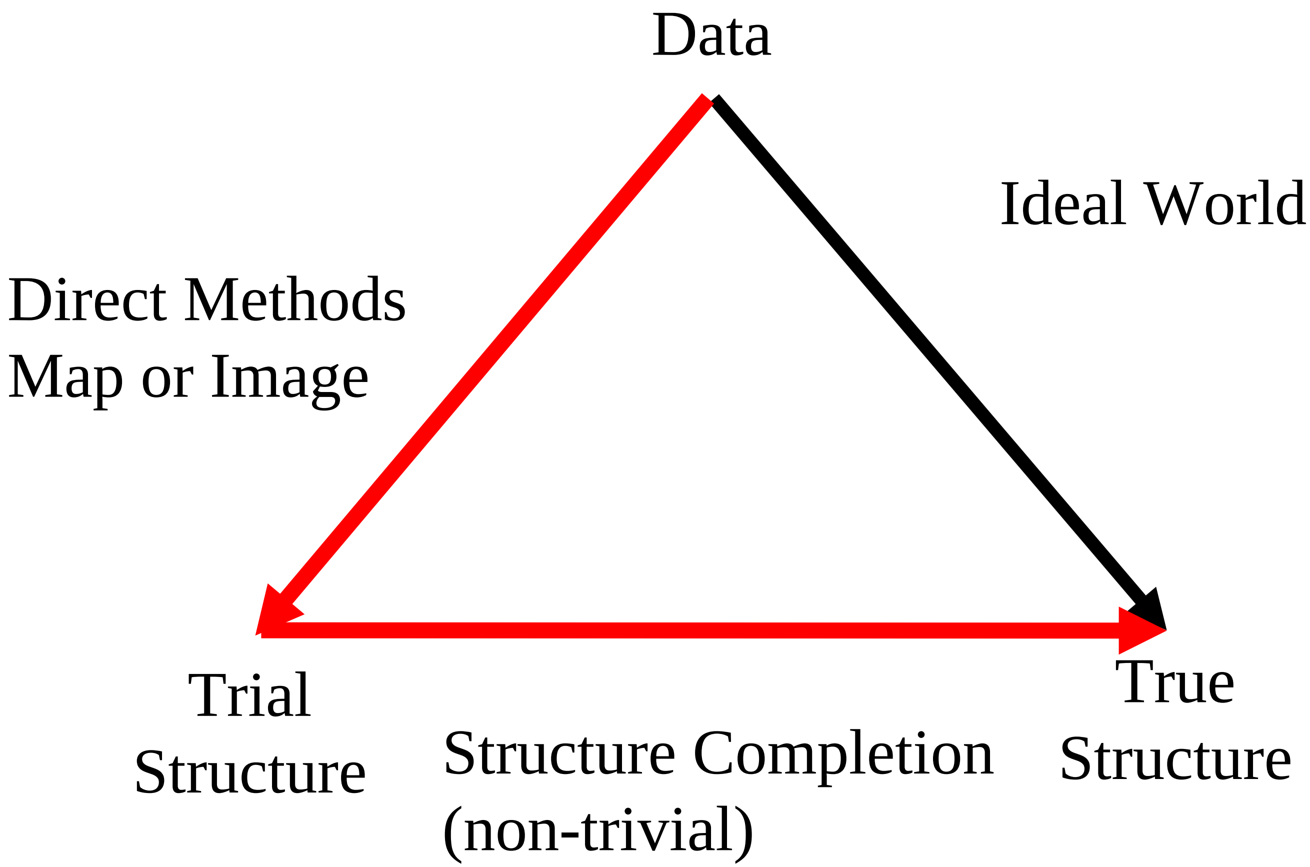
Direct Methods:

Using available information to find solutions





Crystallographic Direct Methods Structure Triangle



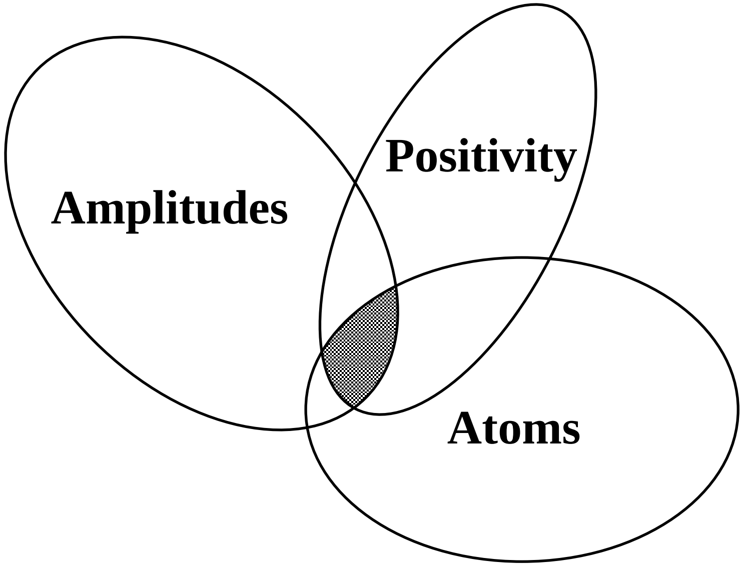
More: 1970's Mathematics

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

$$I = F + \lambda C$$

1990's Mathematics

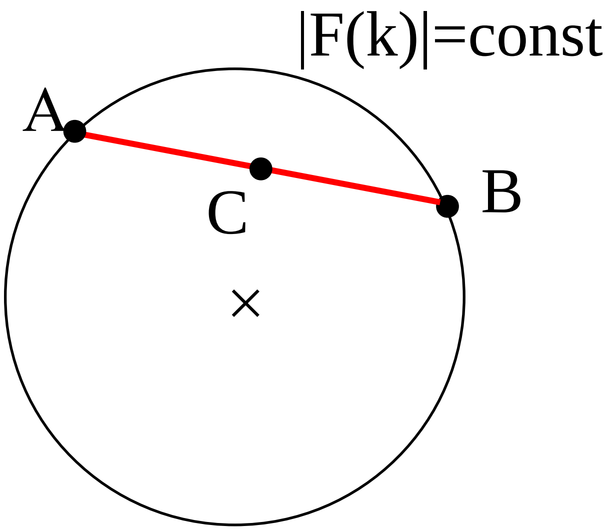
- We have constraints (e.g. atomicity, amplitudes)
  - Treat as sets
- We are looking for the solution as intersection of several constraint sets



Acta Cryst <u>A55</u>, 601 (1999)

The $64,000 question

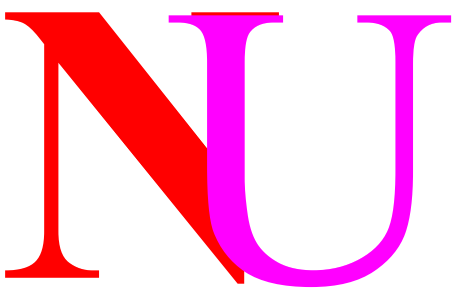
- A set is convex if any point between two members is also a member
  - If all the sets are convex, problem has one solution
  - If they are not, there may be more than one local minimum



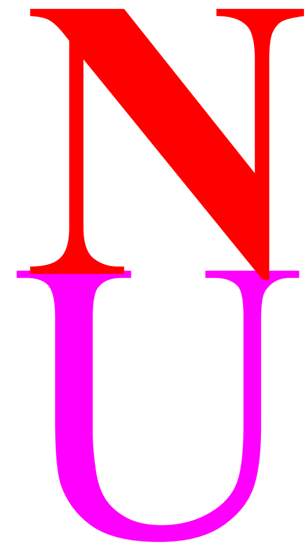
- Amplitude measurements do not form a convex set
- But…there still may only be one solution.
- Unsolved mathematical problem

Multiple non-convex constraints

Consider the two sets “N” and “U”

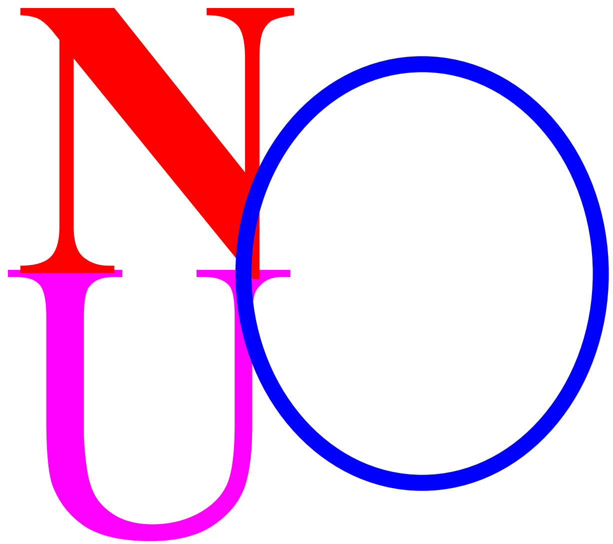


Overall Convex



Overall Non-Convex

Crystallographic methodology



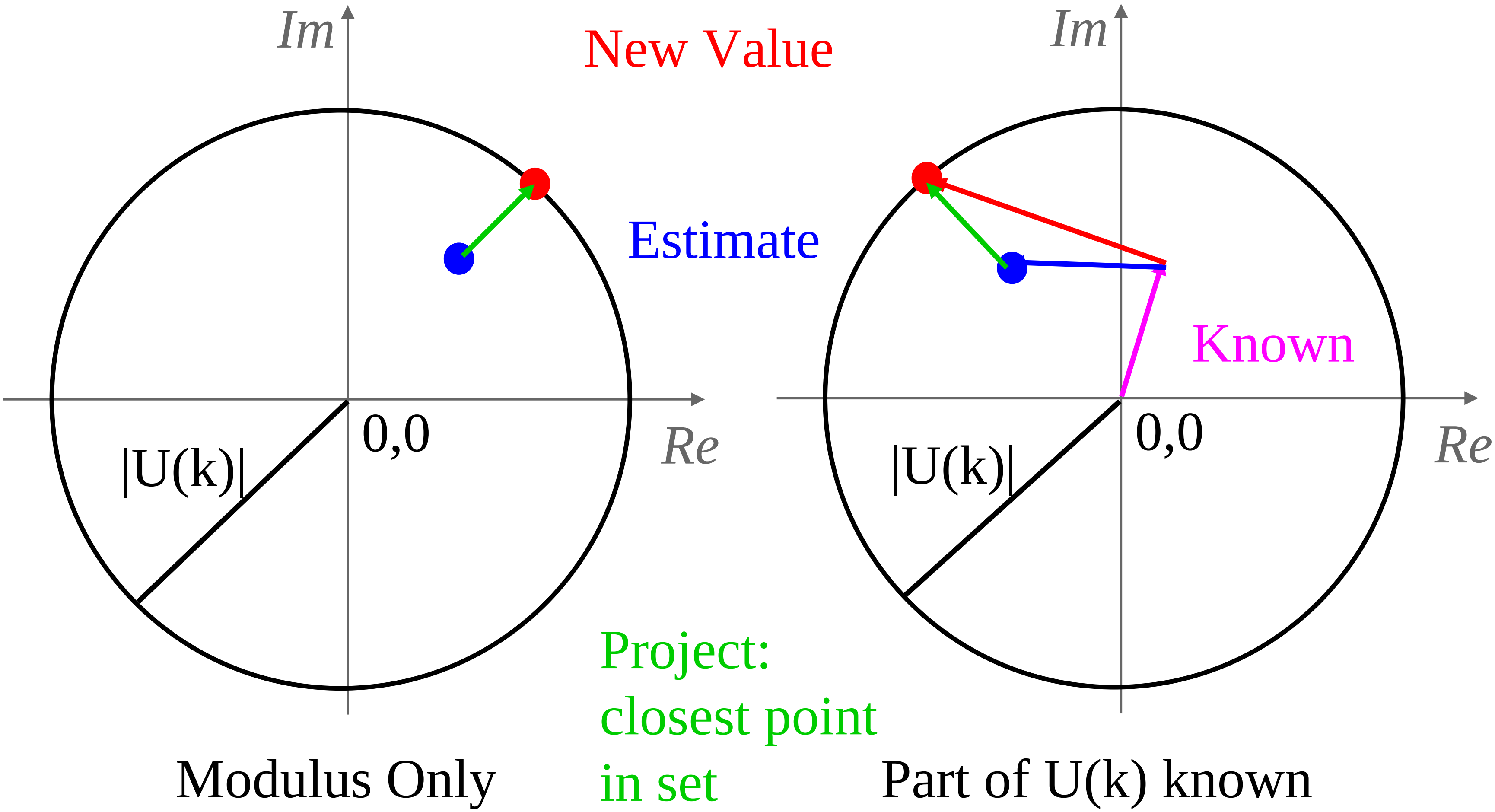
Overall Non-Convex

Overall Unique

Addition of additional convex constraints tends to give a unique solution

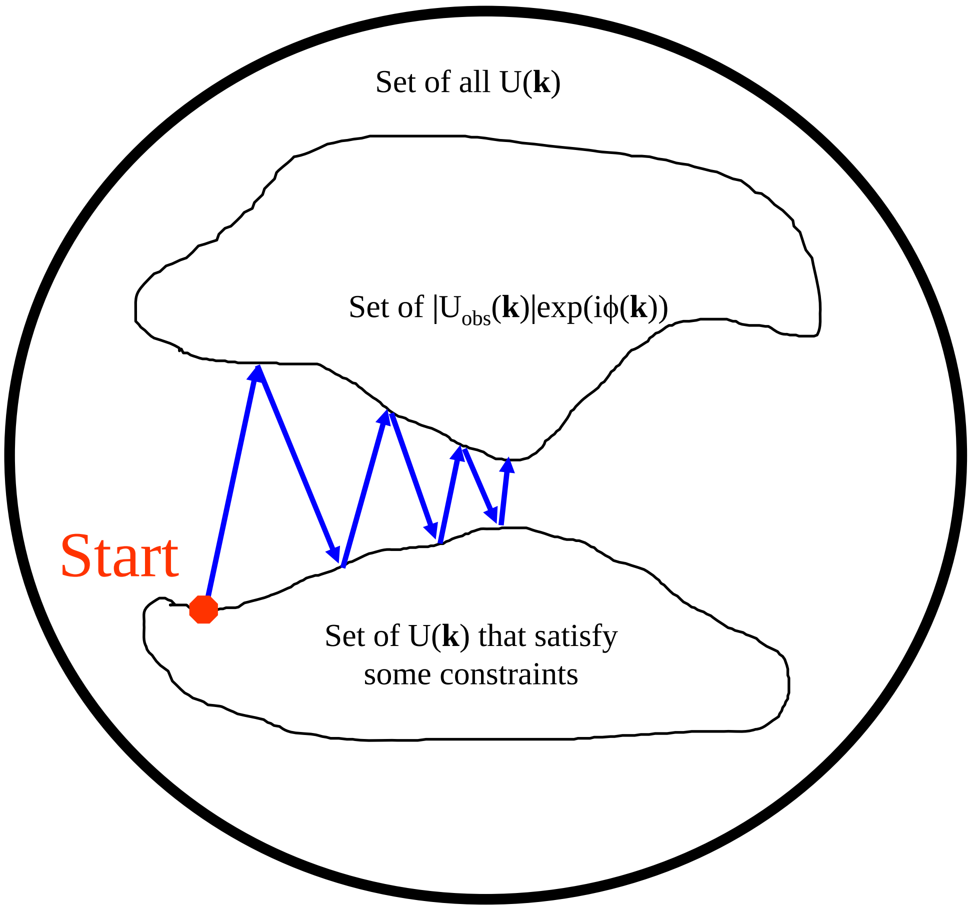
Structure Completion: add additional constraints as the phases become known

Orthogonal Projections



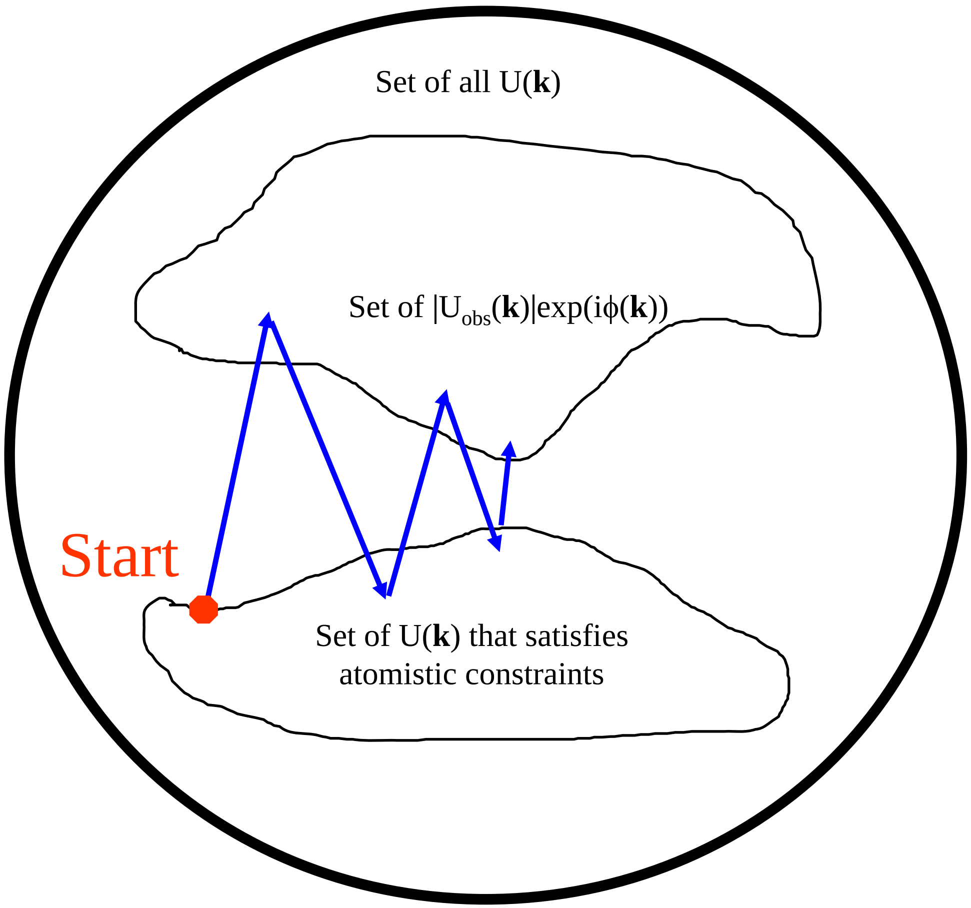
Successive Projections

- 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



Where do constraints come from

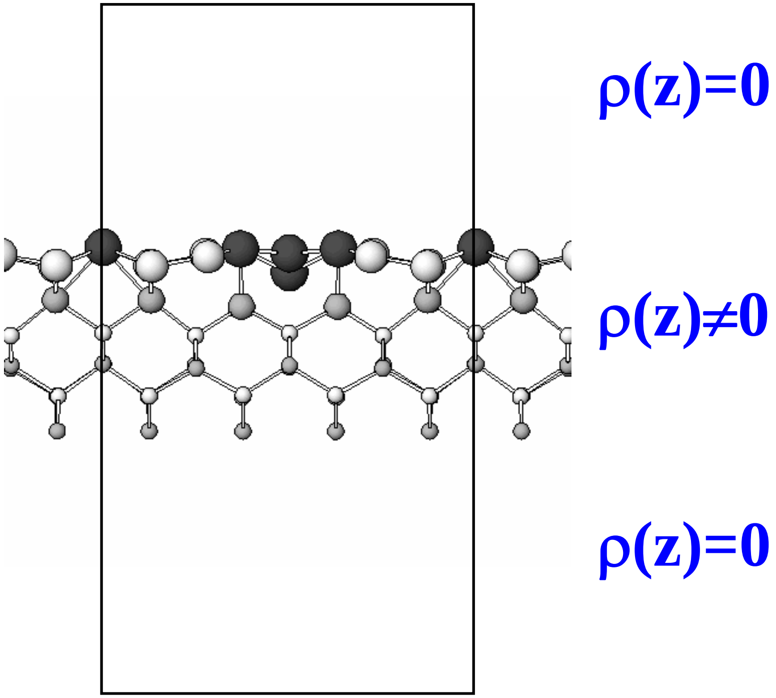
- Physical nature of experiment
  - Limited beam or object size
- Physical nature of scattering
  - Atomic scattering
- Statistics & Probability
  - Minimum Information/Bias = Maximum Entropy

Types of Constraints

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

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)

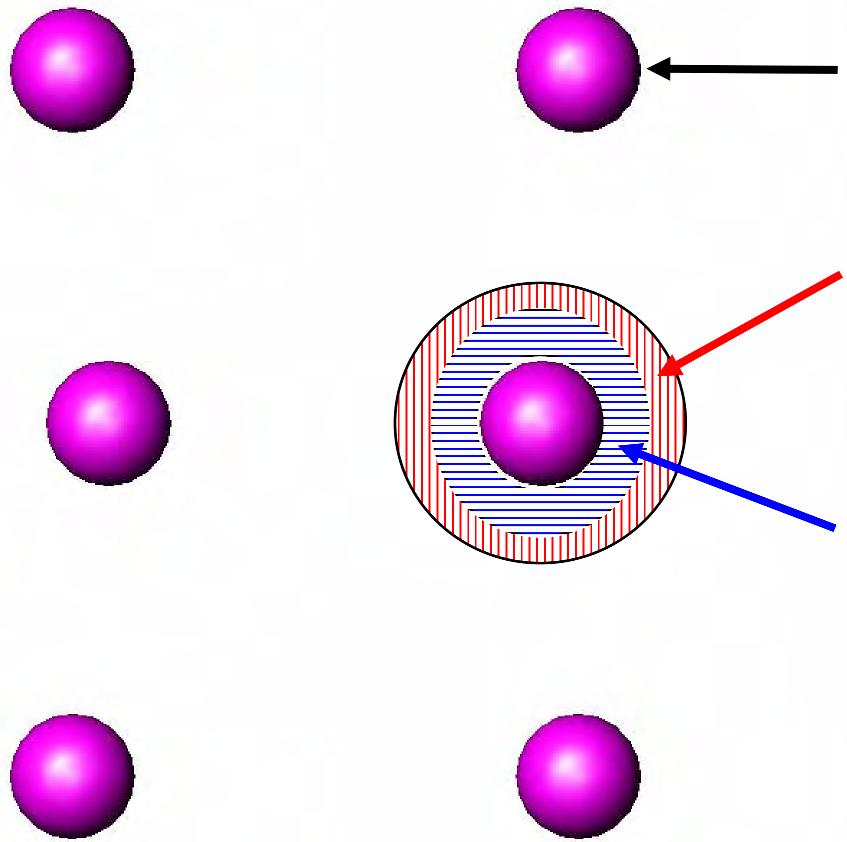
Why we don’t need all the data

- 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

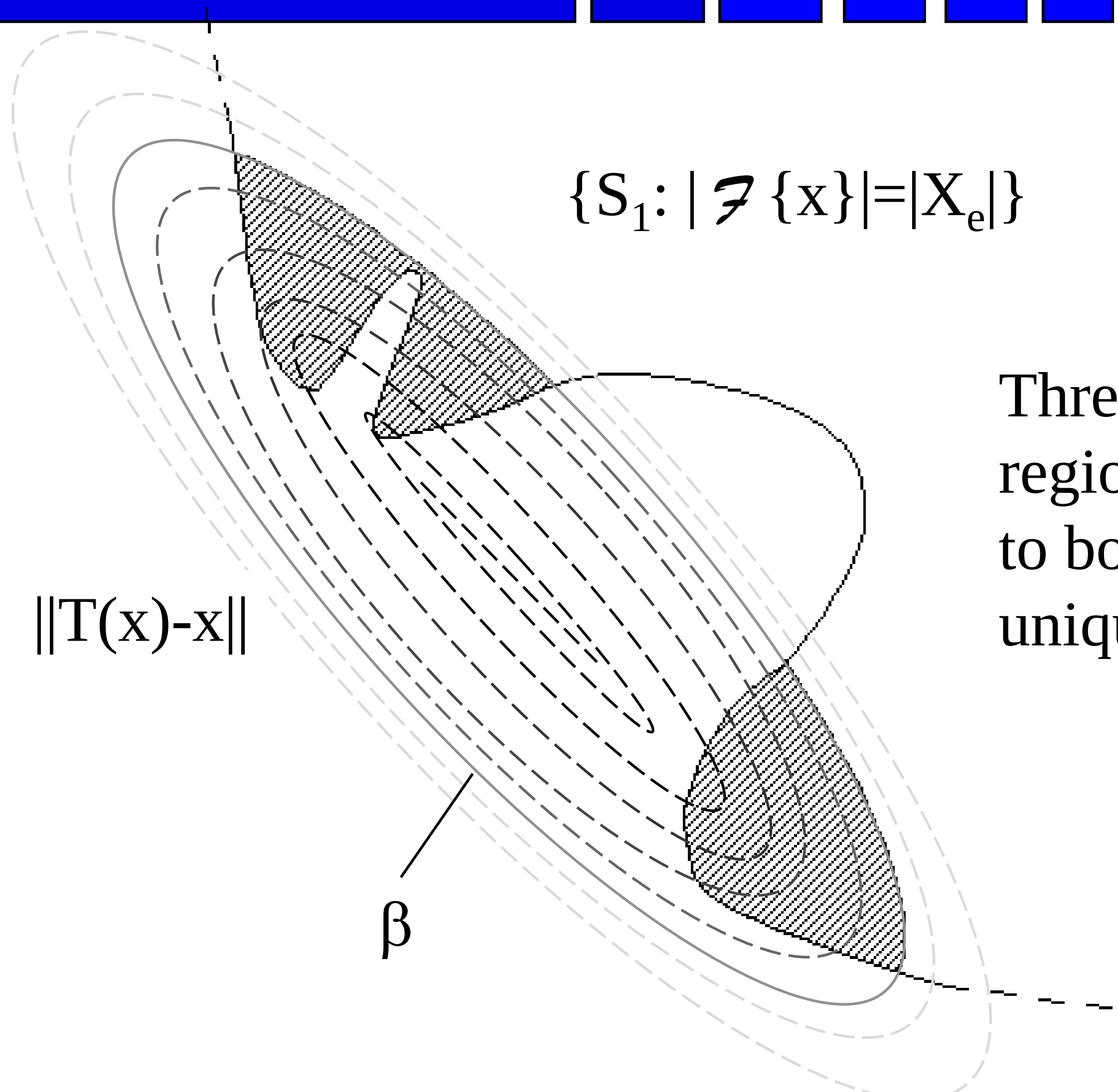
Other Constraints

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



Multiply-Connected Feasible Set

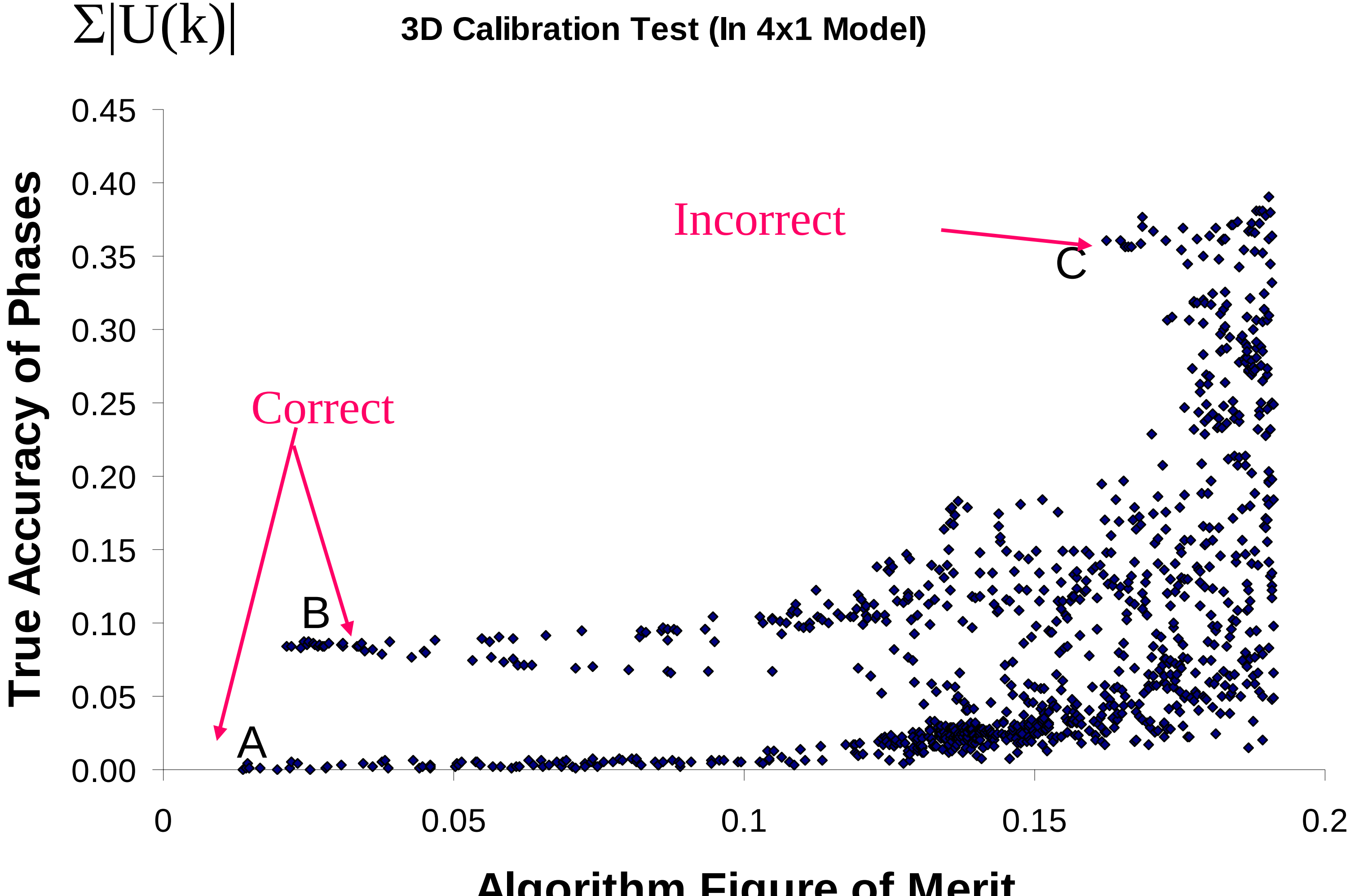


Three shaded regions common to both sets, 3 unique solutions

Typical results

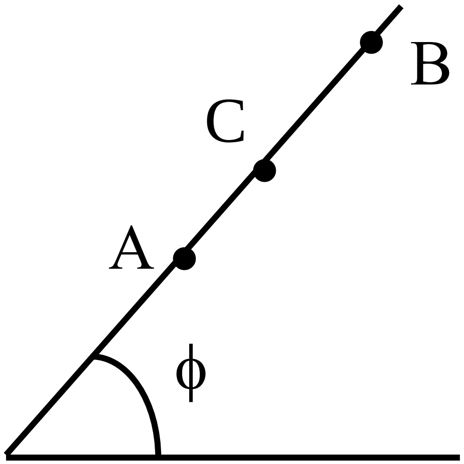
$\Delta\phi$ = phase error

$$\frac{\Sigma|U(k)|\{1-\cos(\Delta\phi)\}}{\Sigma|U(k)|}$$



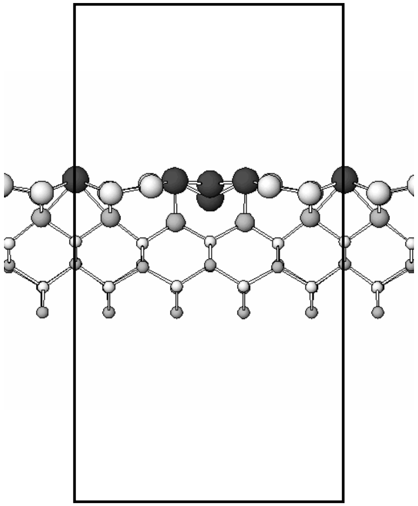
Convex Set for unmeasured |U(h,k,l)|

- 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



Support Constraint

- Displacements decay as $(\alpha+z)\exp(-qz)$ into bulk[1]
- Real space constraint
  - $\rho(z)=\rho(z)w(z)$ $w(z)=1, -L<z<L$
    $=0$, otherwise
- Convex constraint
- Has well documented properties



$\rho(z)=0$

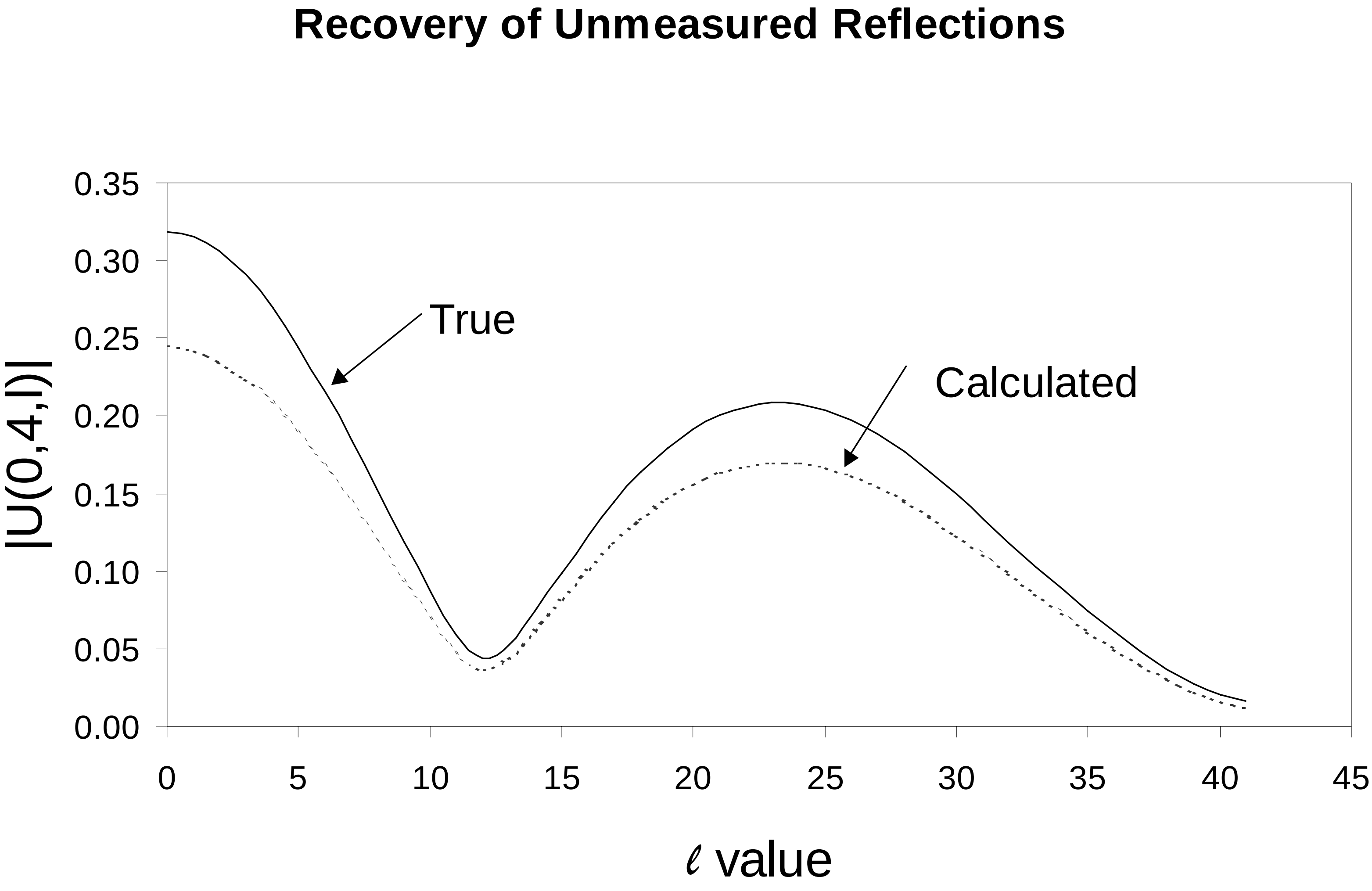
$\rho(z)\neq0$

$\rho(z)=0$

PRB 60, 2771 (1999)

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

Unmeasured Reflections



Addition Information

- Physical nature of scattering
  - Atomic scattering
- Statistics & Probability
  - Minimum Information/Bias = Maximum Entropy
- These can be converted to mathematical constraints

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

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

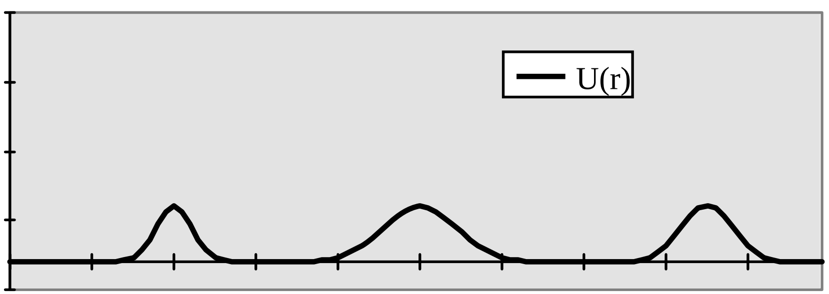
$$u(r) = Nu(r)^2$$

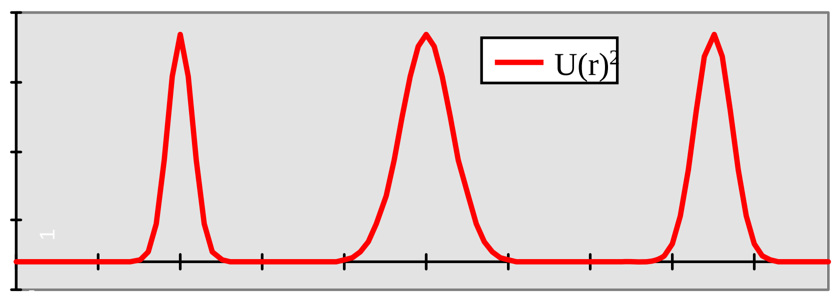
Constraint

Real/Reciprocal Space

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

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





Reinforces strong (atom-like) features

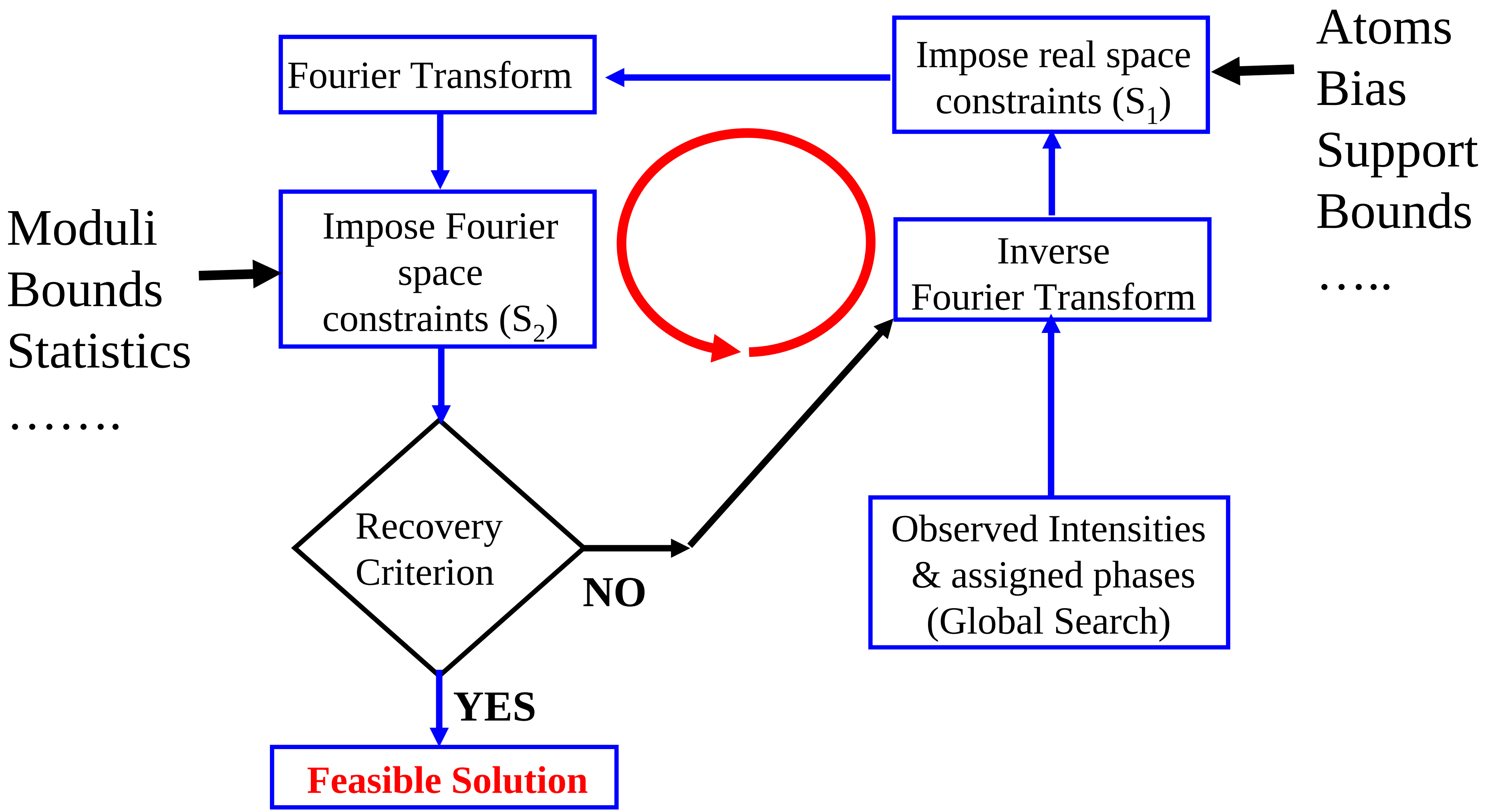
Tangent Formula

- 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)



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

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

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

Known

Average must be 2nπ

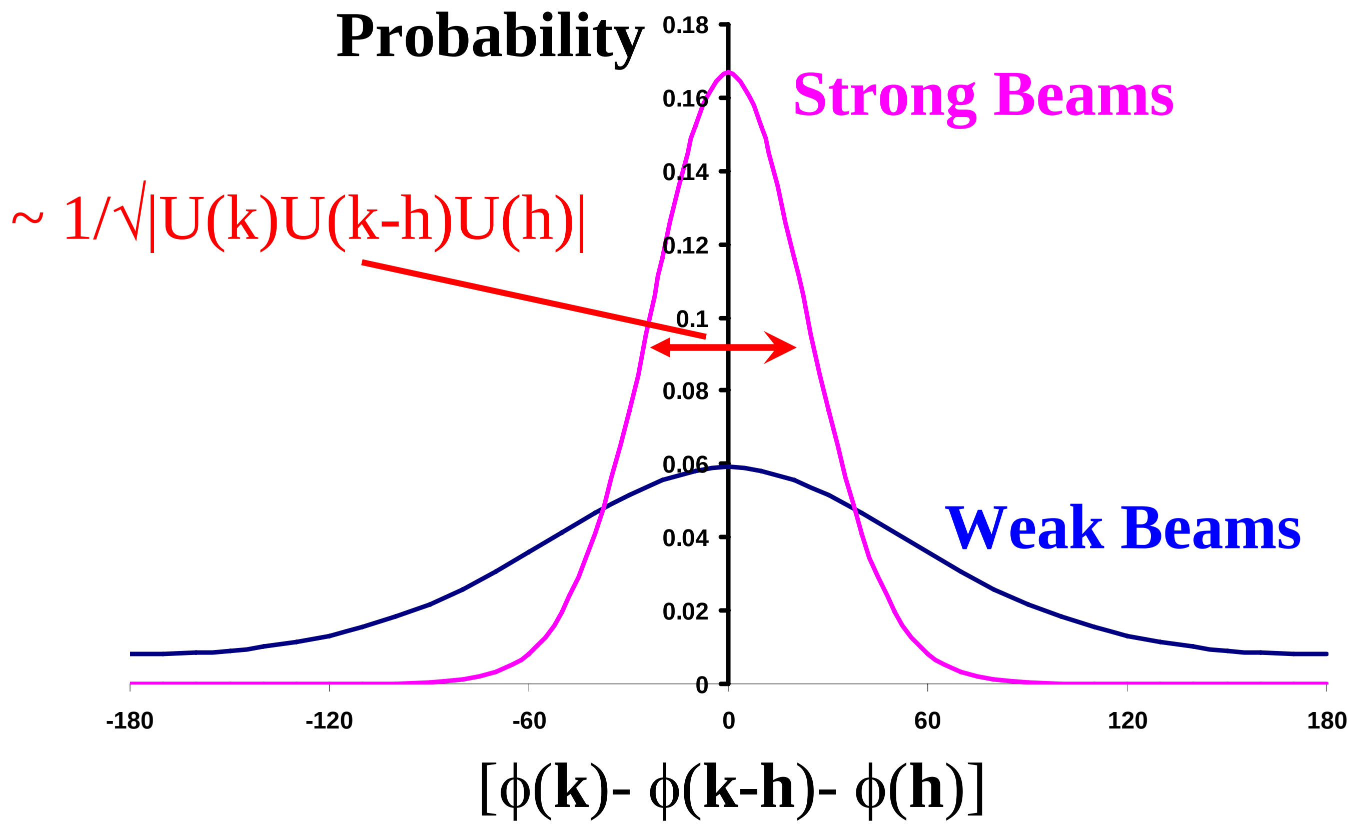
Cochran Distribution: III

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

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

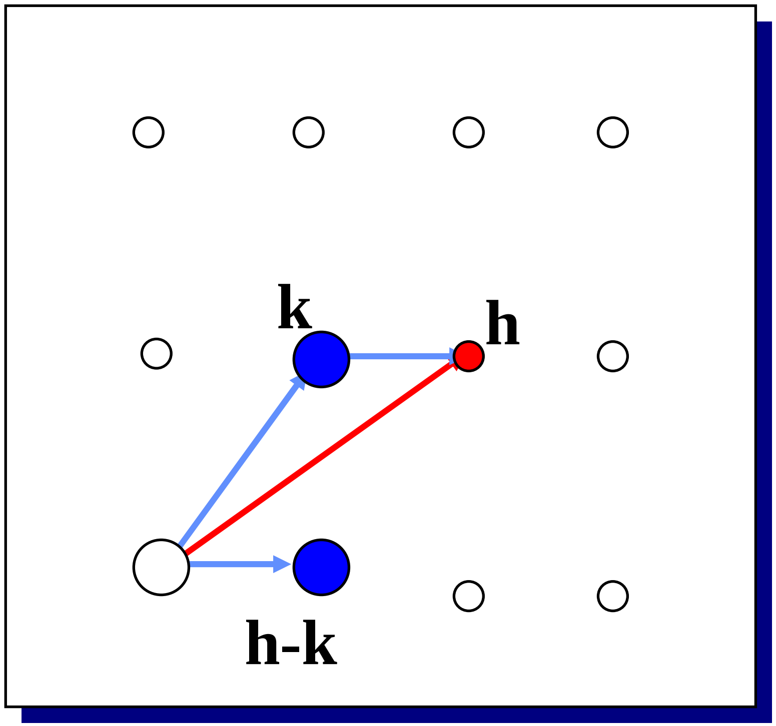
$\sim C\exp(2|U(\mathbf{k})U(\mathbf{k-h})U(\mathbf{h})|\cos[\phi(\mathbf{k}) - \phi(\mathbf{k-h}) - \phi(\mathbf{h})])$

Form of Distribution



Note: this is more statistics than the presence of atoms

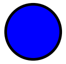
Σ_2 Triplet



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

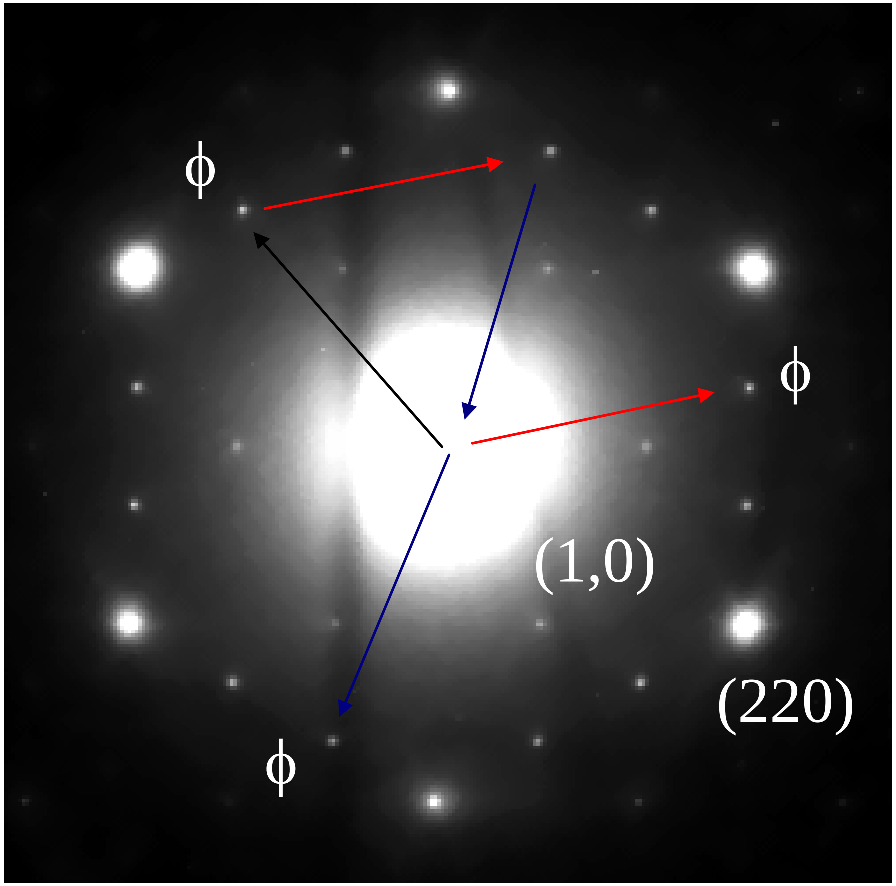
$$\phi(\mathbf{h}) \approx \phi(\mathbf{k}) + \phi(\mathbf{h\text{-}k})$$

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

 = known structure amplitude and phase

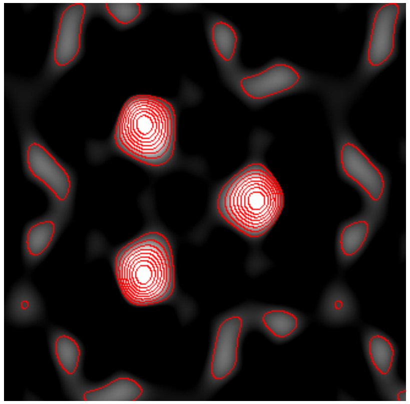
 = known structure amplitude and <u>unknown</u> phase

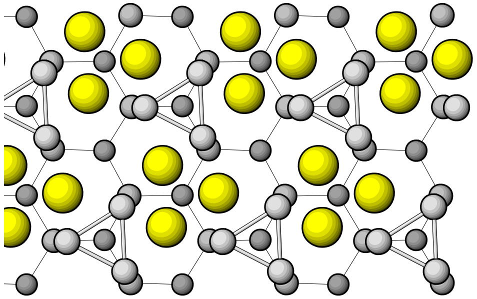
Example: Si(111) √3x√3 Au



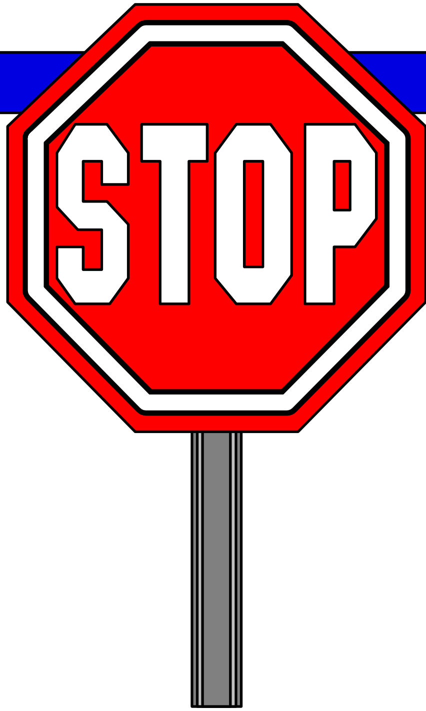
Only one strong reflection

- 3ϕ ~ 360n degrees
- ϕ=0,120 or 240
- ϕ=0 has only 1 atom
- 120 or 240





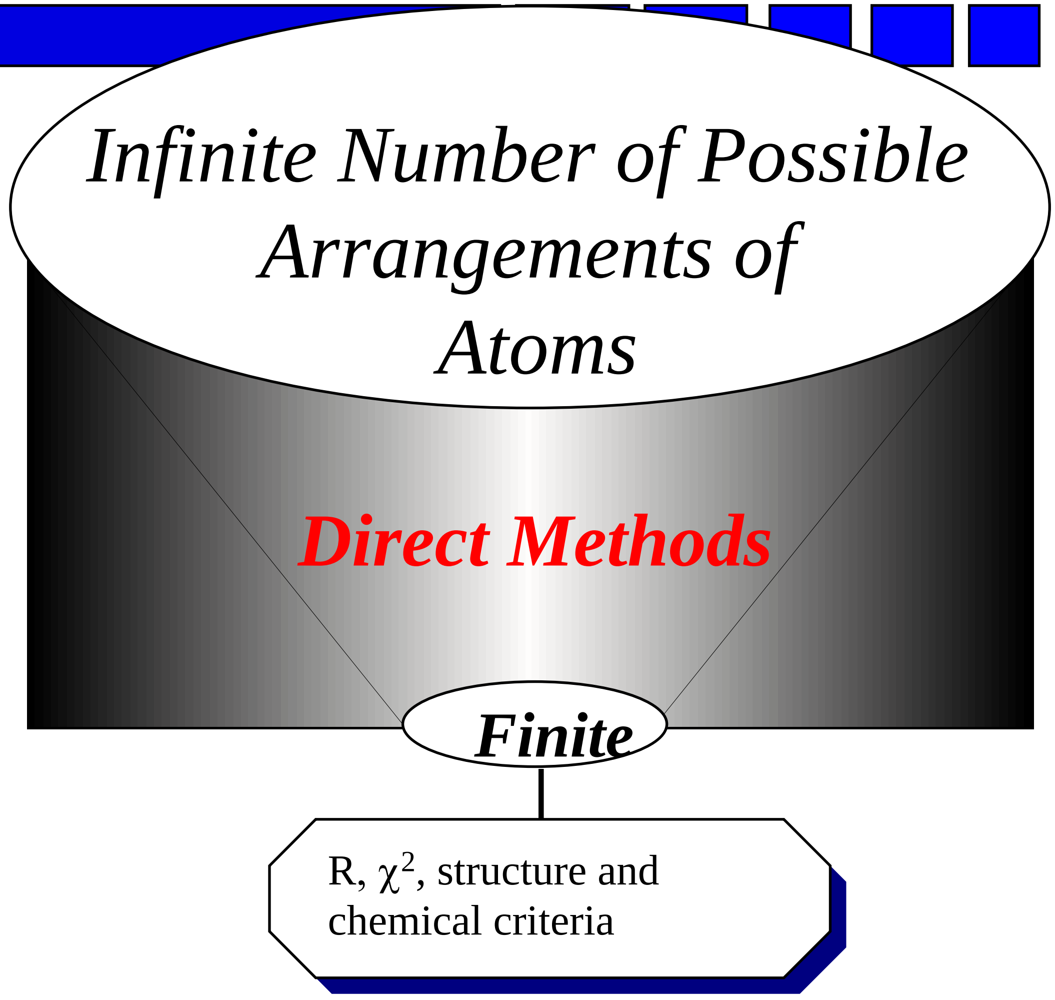
Caveat: Not Physics



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?

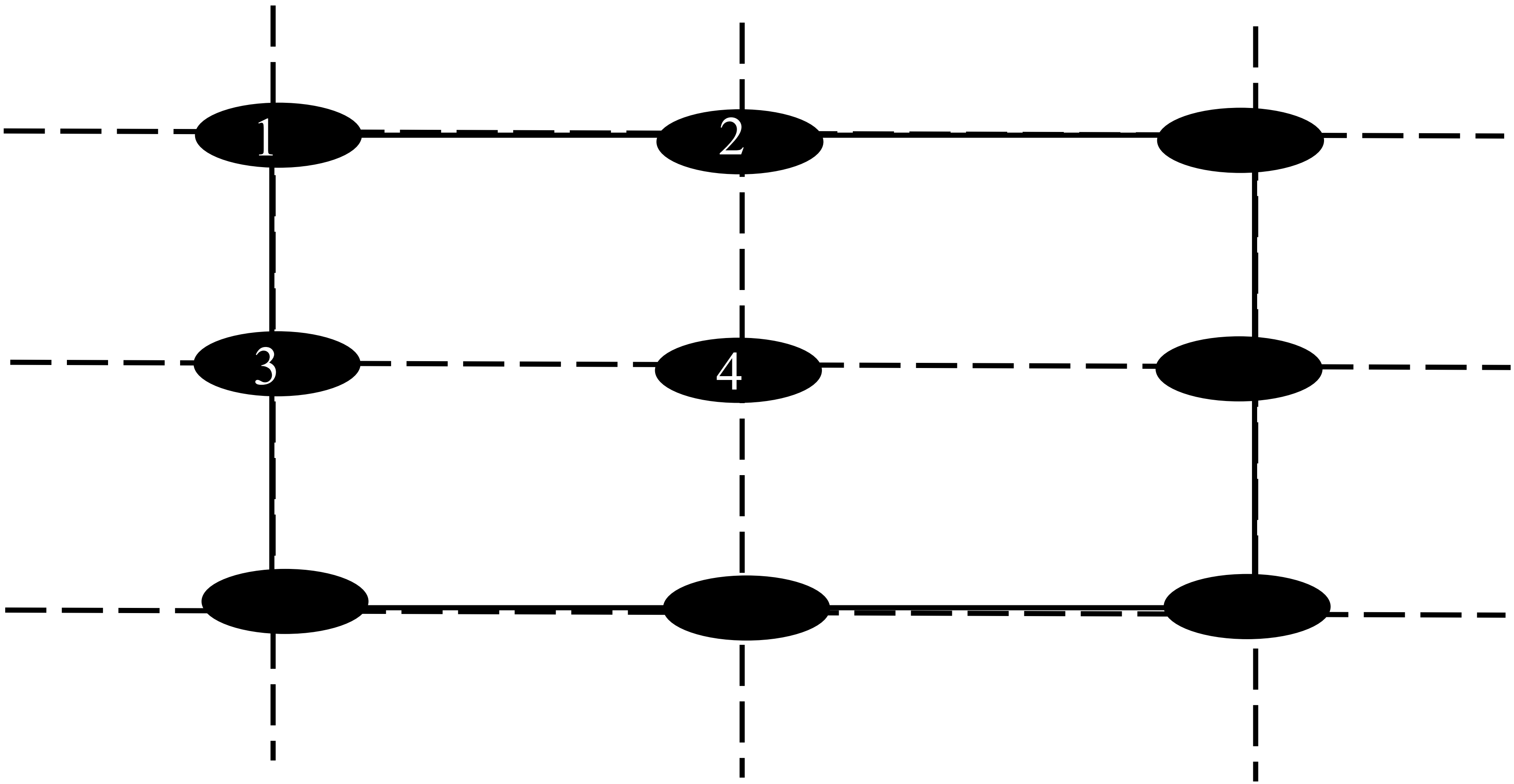


Implementation

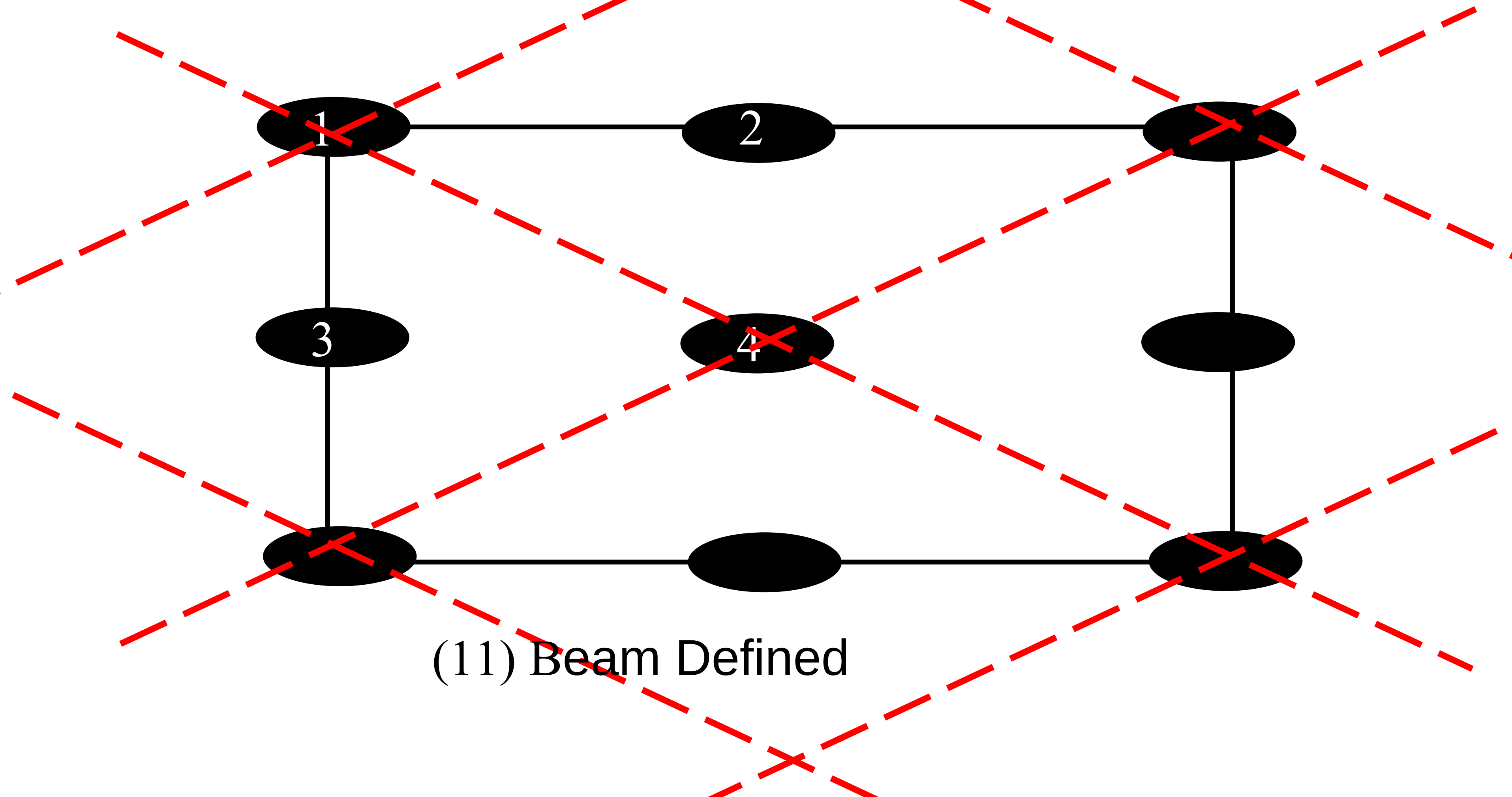
1. Guess phases for some reflections
2. Generate from these phases for others and improved phases for initial set
3. Test consistency of predicted amplitudes and phases
4. Iterate, so long as consistency is improving

Note: permuting phases has lower dimensions than permuting atom positions

Origin Definition c2mm

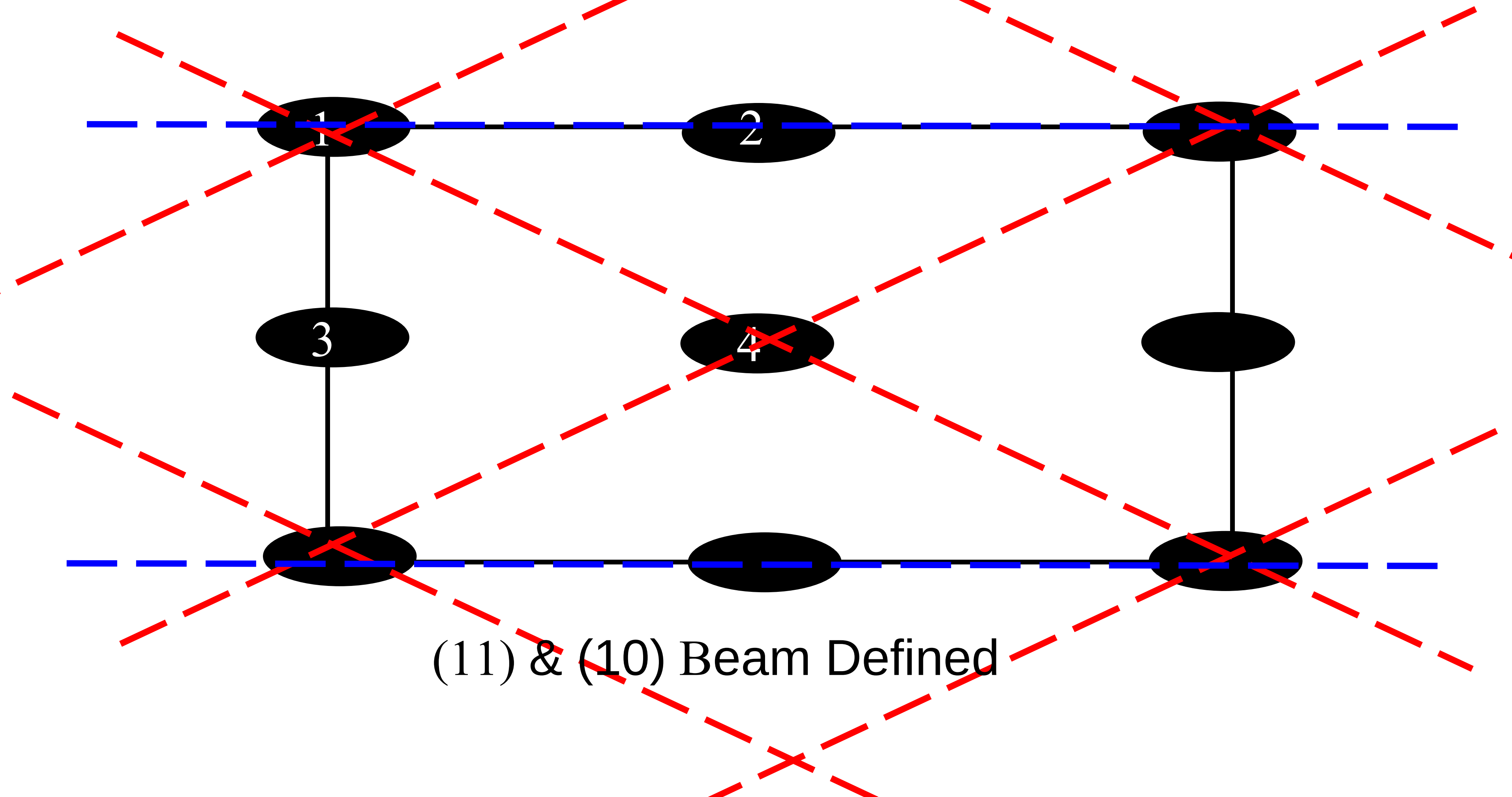


Origin Definition c2mm



(11) Beam Defined

Origin Definition p2mm



(11) & (10) Beam Defined

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

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

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

$= ½ + U(2k)$

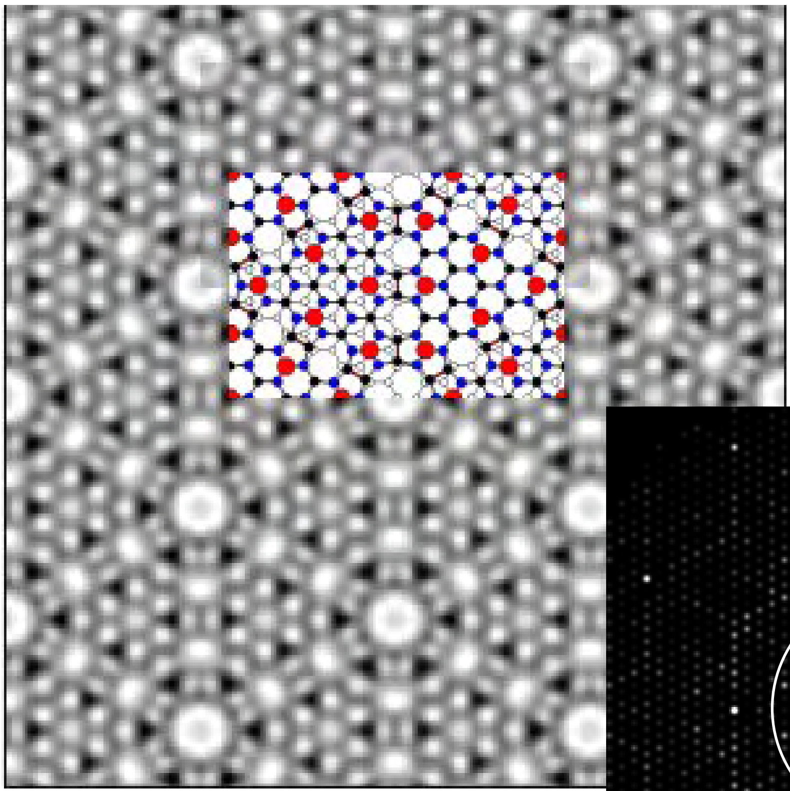
Hence $U^2(k) < ½ + U(2k)/2$

If U(k) is large – can set U(2k)

Quartets

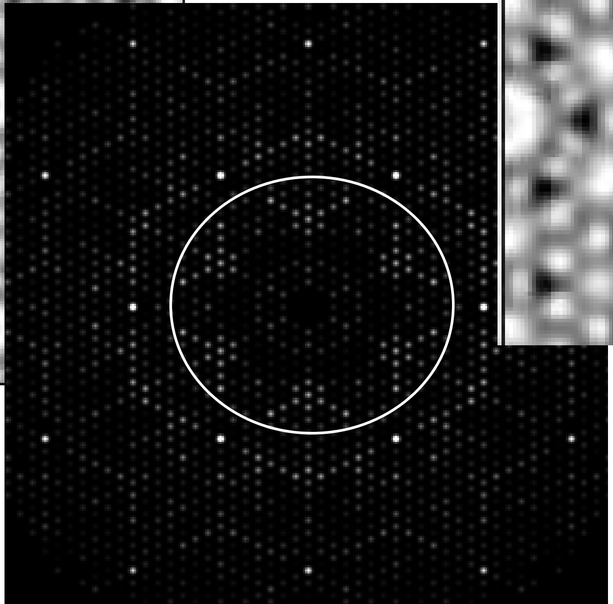
- Phase relationships involving 4 terms for weak reflections
  - Positive and Negative
  - Rarely useful with TEM

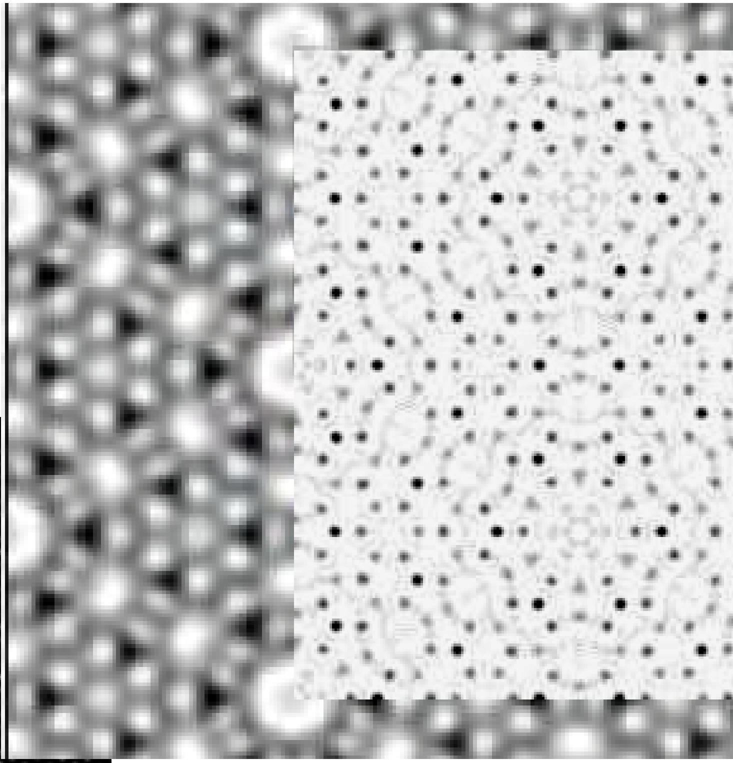
Restoration and Extension



0.3nm Image

+DP

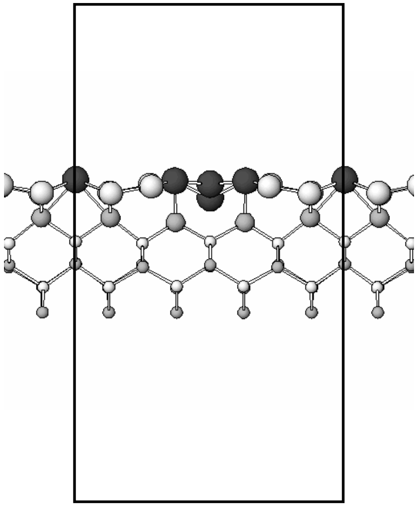




0.05nm Image

Support Constraint

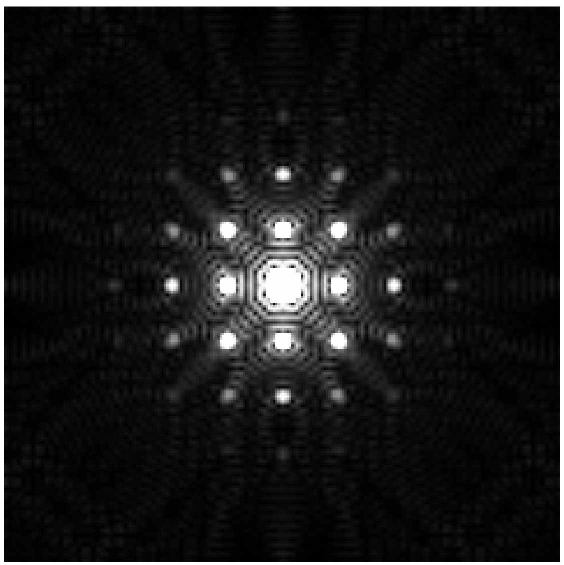
- Displacements decay as $(\alpha+z)\exp(-qz)$ into bulk[1]
- Real space constraint
  - $\rho(z)=\rho(z)w(z)$ $w(z)=1, -L<z<L$
    $=0$, otherwise
- Convex constraint
- Has well documented properties



PRB 60, 2771 (1999)

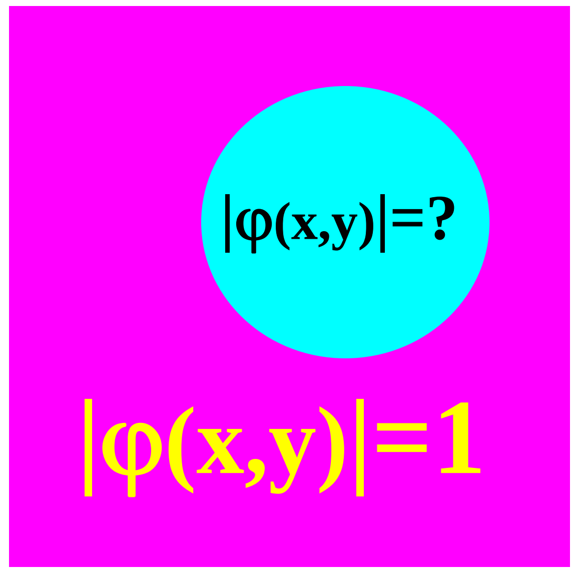
[1]Biharmonic expansion of strain field, SS 294, 324 (1993)

Phase Recovery for a Small Particle



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

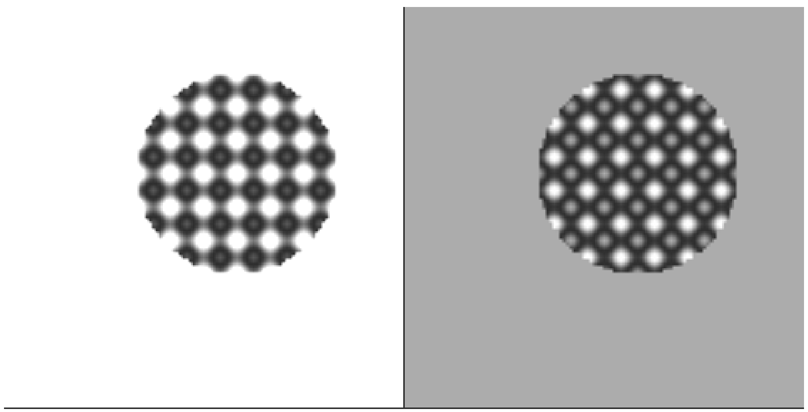
+



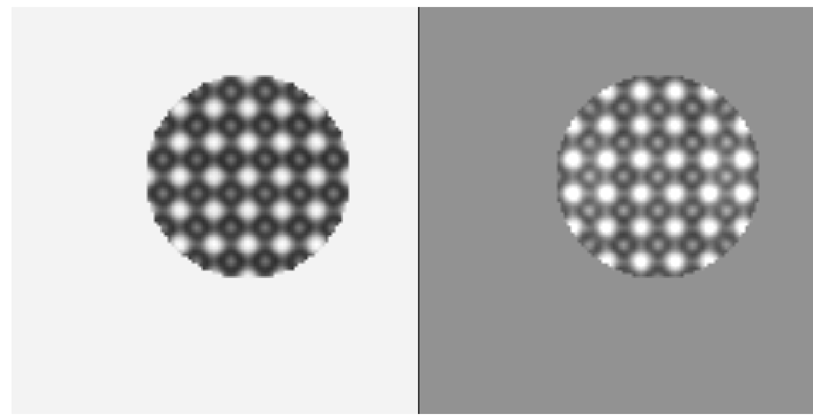
Convex Support
Constraint

= ?

Phase Recovery for a Small Particle

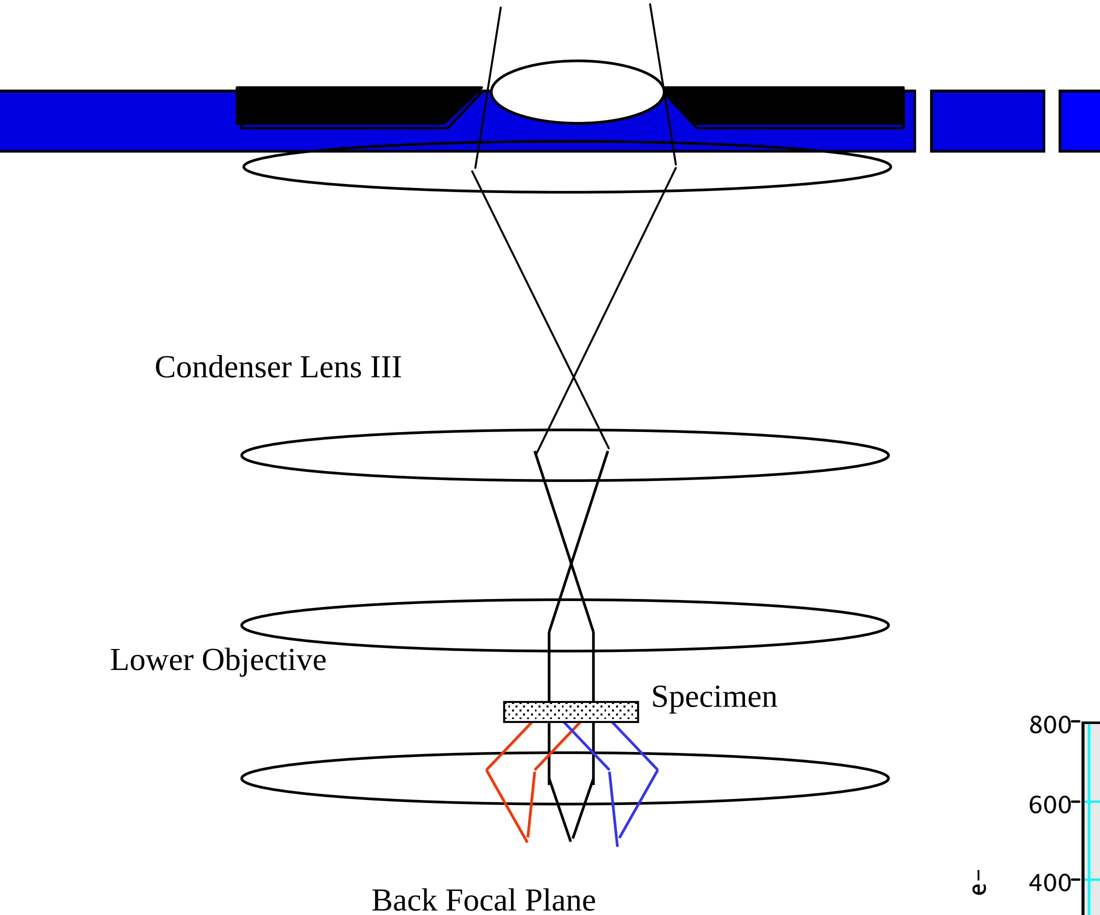


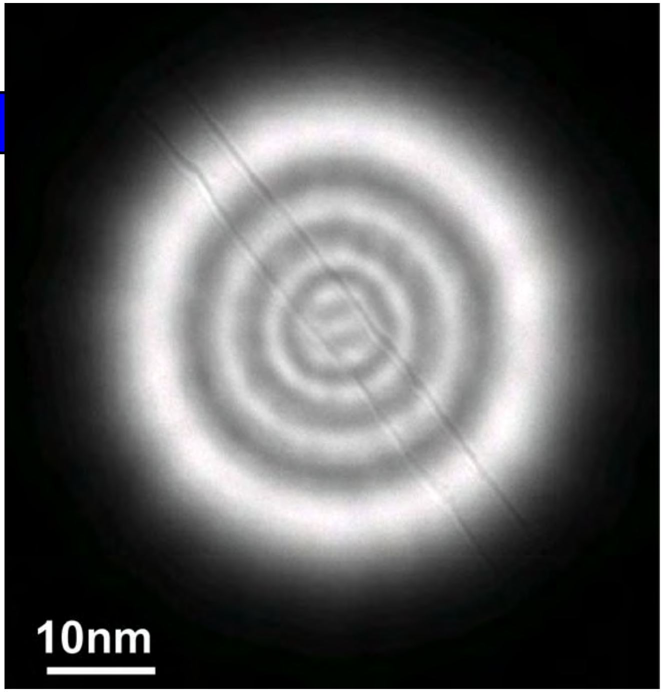
True real space exit wave for small particle model

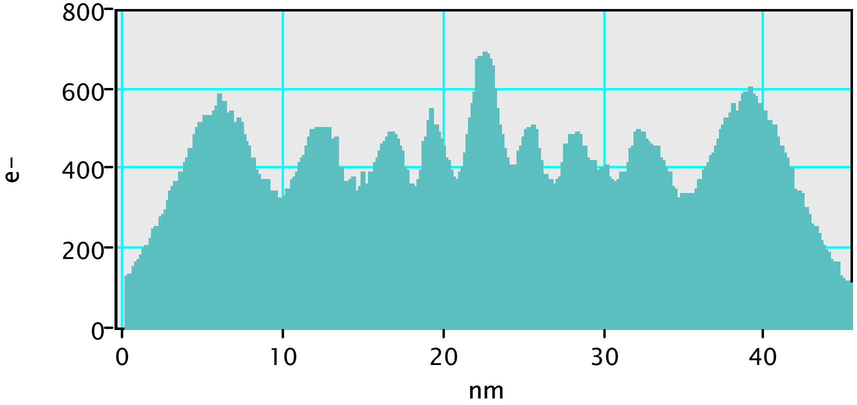


Reconstructed exit wave after 3000 iterations

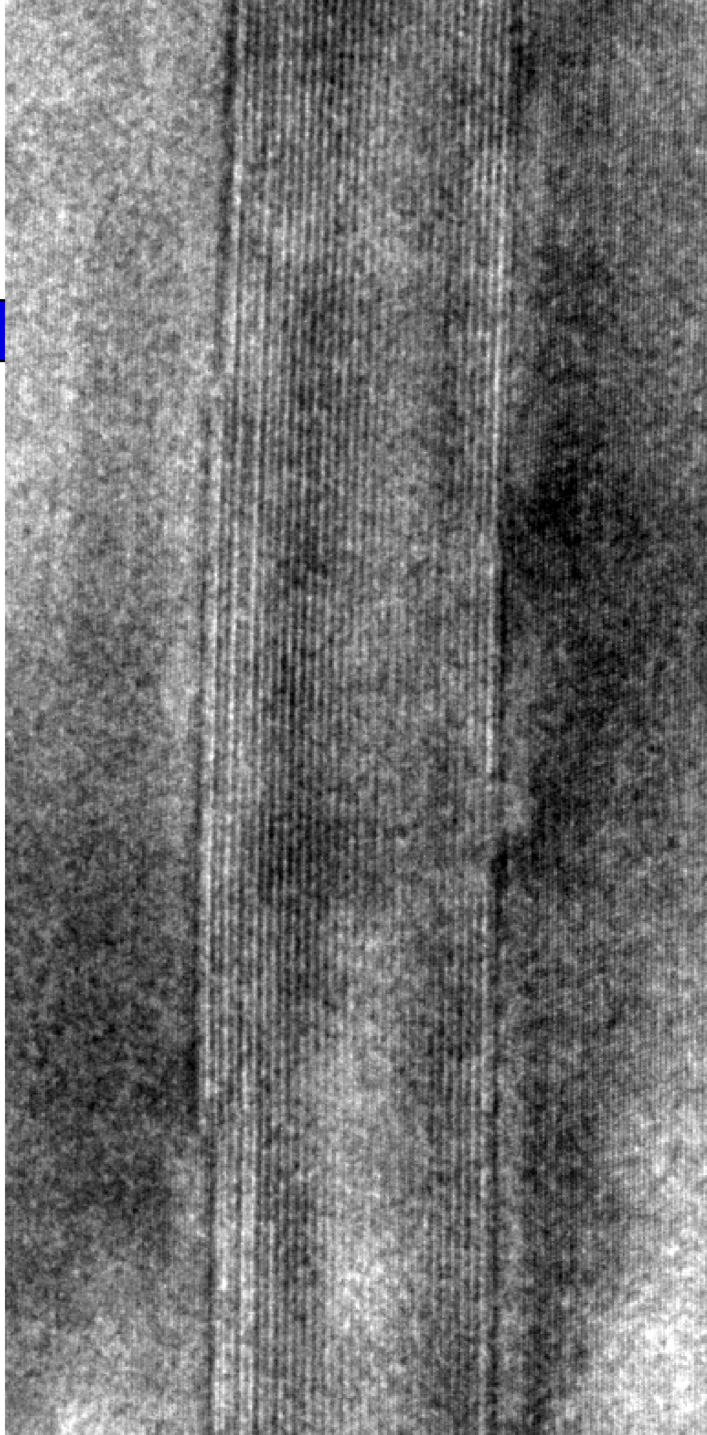
Electron Nanoprobe formation

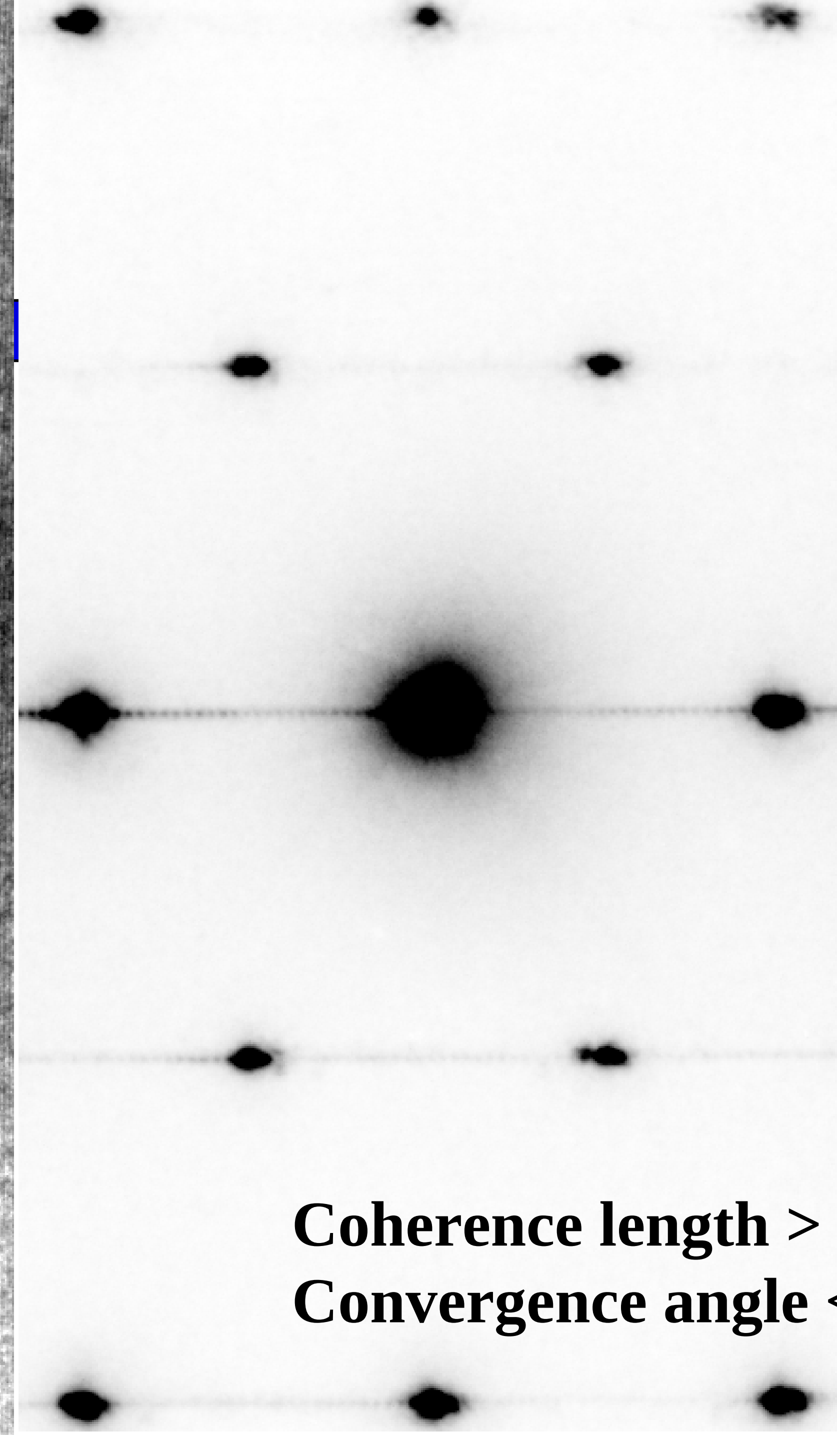






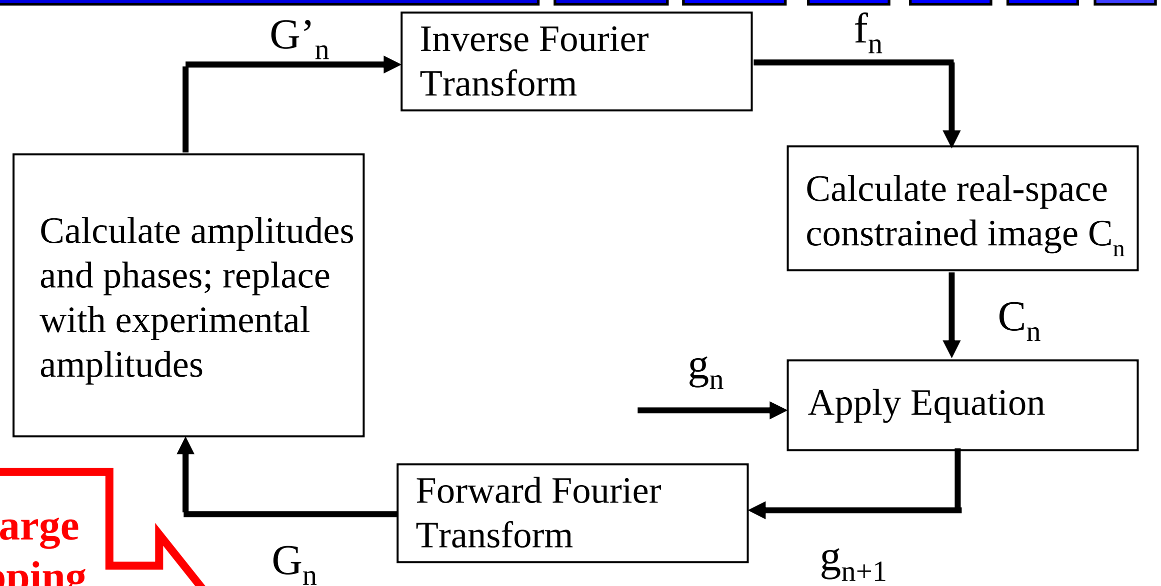
10 µm aperture -> 50 nm beam
M = 1/200

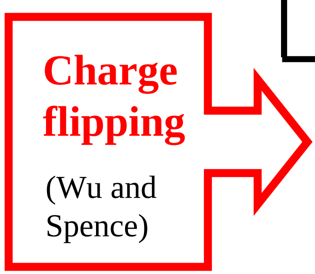




Coherence length > 15 nm
Convergence angle <0.2 mrad

The Algorithm

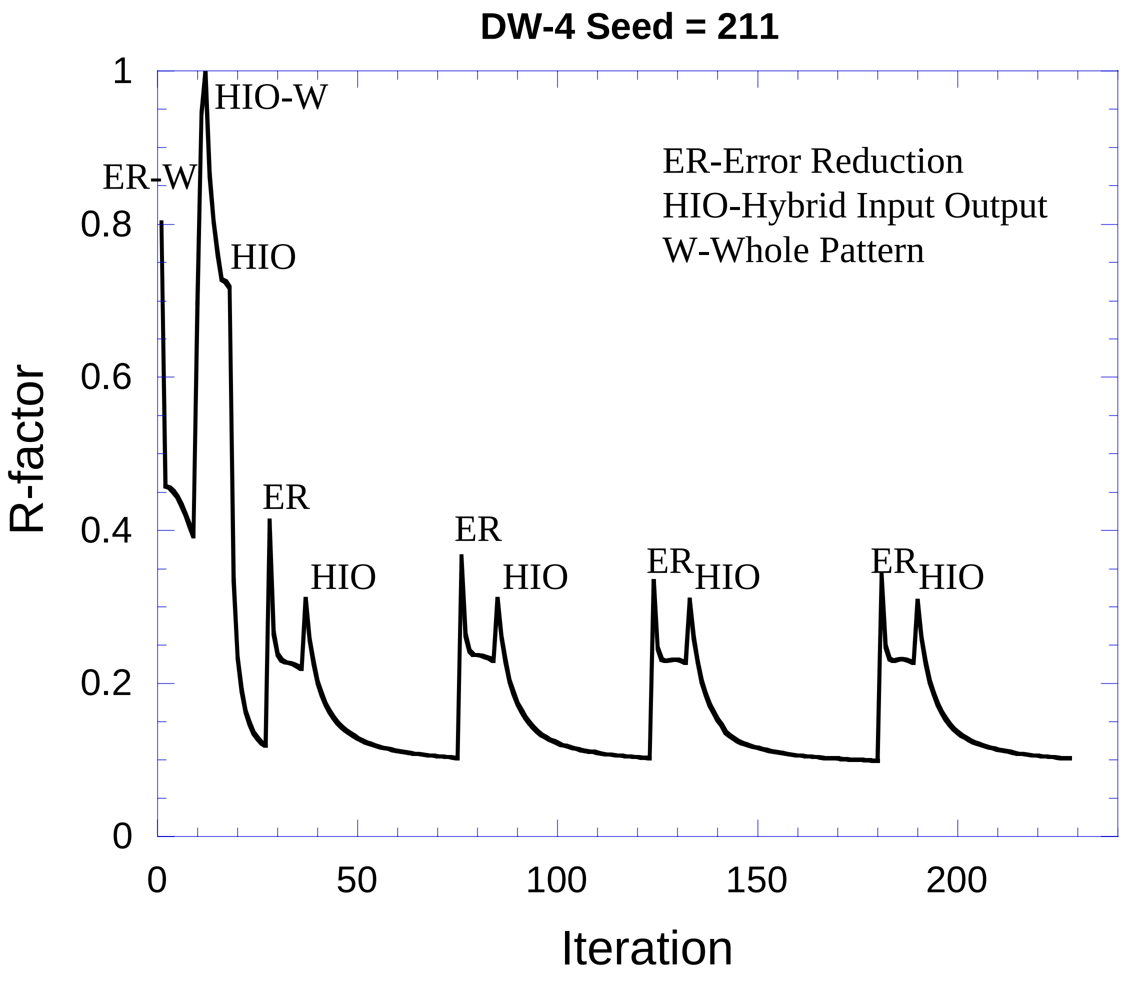




$$g_{n+1}(\vec{r}) = \begin{cases} f_n(\vec{r}) & if\ |C_n(\vec{r}) - f_n(\vec{r})| < \varepsilon \\ g_n(\vec{r}) + \beta[C_n(\vec{r}) - f_n(\vec{r})] & if\ |C_n(\vec{r}) - f_n(\vec{r})| > \varepsilon \end{cases}$$

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

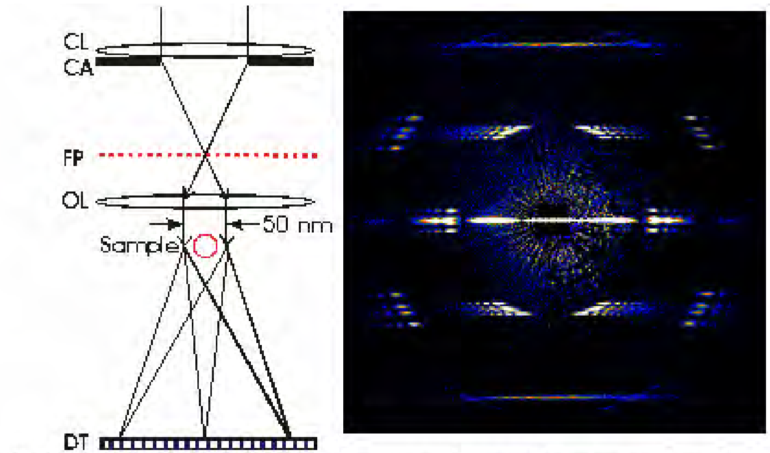
Convergence and the Missing Central Beam



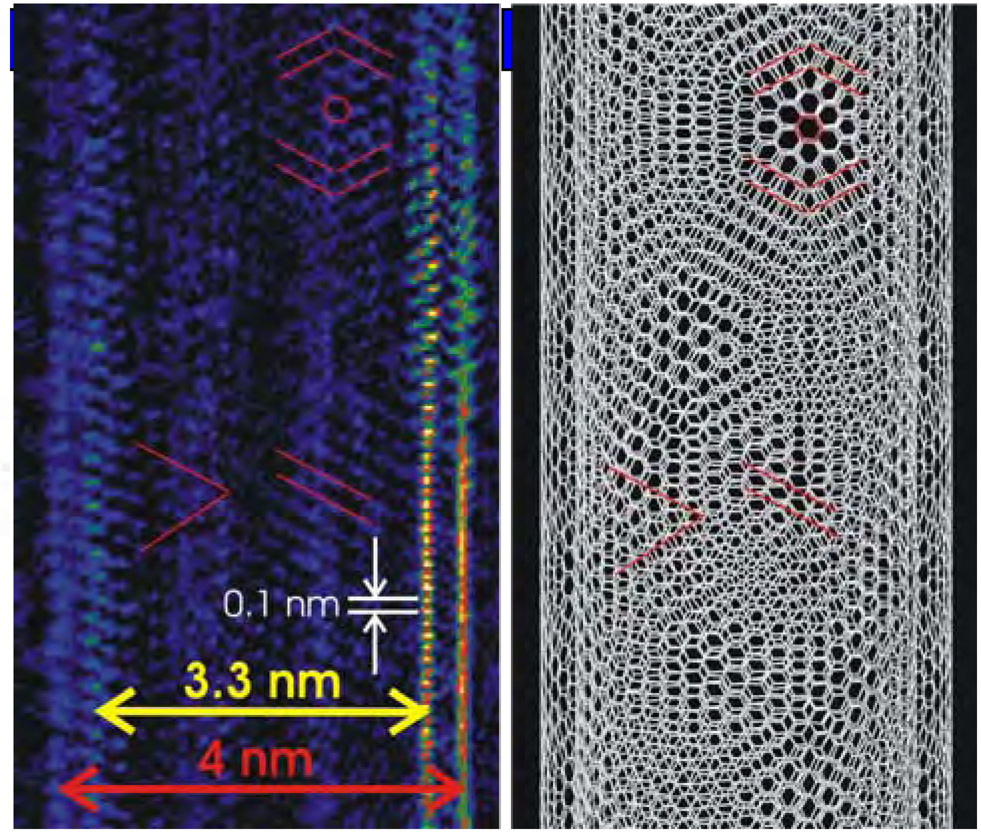
$$R = \frac{\sum \left\| F^{Exp} \right| - \left| F^{R} \right\|}{\sum \left| F^{Exp} \right|} 100\%$$

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

Diffractive Imaging and Phase Retrieval

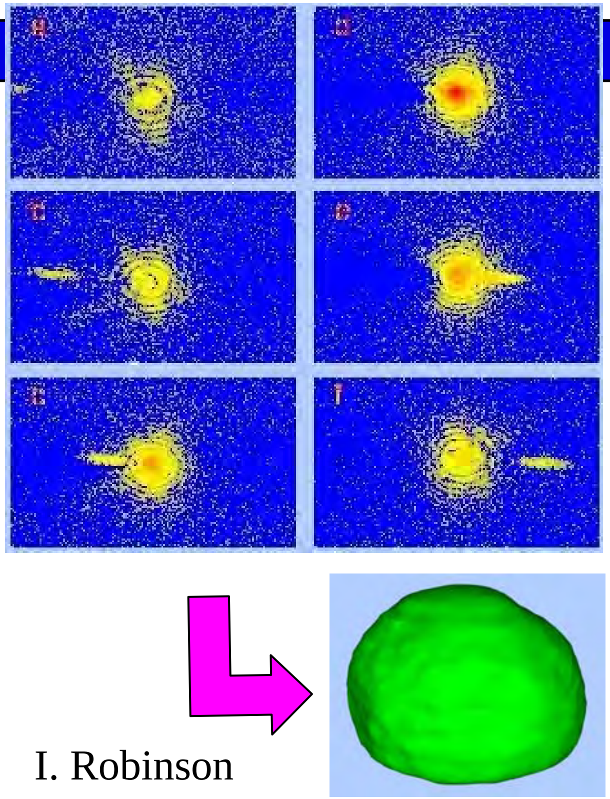


(left) A single double wall nanotube is illuminated with a narrow beam of electrons. (right) The diffraction pattern of the tube

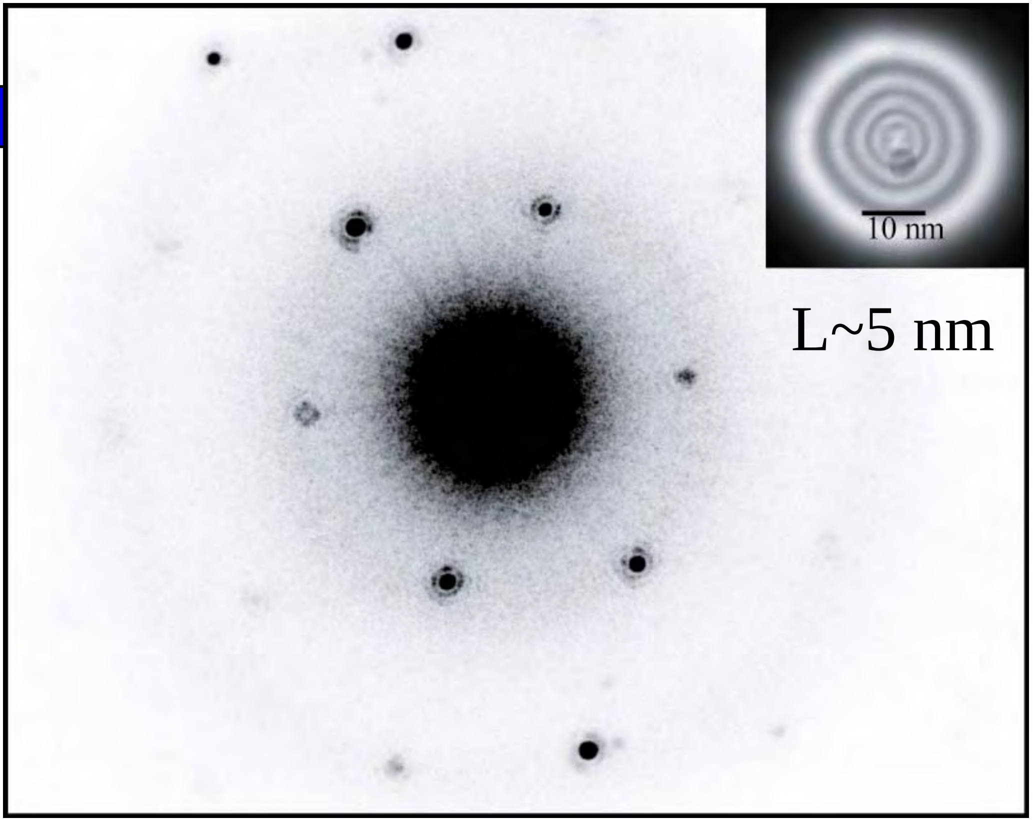


J.M. Zuo, I. Vartanyants, M. Gao, R. Zhang and L.A. Nagahara, *Science, 300,* 1419 *(2003)*

Single Particle Diffraction



I. Robinson
Synchrotron



- Atomic resolution
- Strong interaction of electrons

J. Tao, See Zuo et al, Microscopy Research Techniques, 2004