



In 1935, Patterson showed that the unknown phase information in the equation for electron density:

$$\rho(xyz) = 1/V \sum_h \sum_k \sum_l |\mathbf{F}(hkl)| \exp[i\alpha(hkl)] \exp[-2\pi i(hx + ky + lz)]$$

can be removed if we do the summation on $F^2(hkl)$ and assume that all of the phases are 0. This provides the Patterson function:

$$P(uvw) = 1/V \sum_h \sum_k \sum_l |\mathbf{F}(hkl)|^2 \exp[-2\pi i(hu + kv + lw)]$$

Patterson showed that the position of a “peak” at (u, v, w) provided by this function correspond to the vectors between two atoms (at (x_1, y_1, z_1) and (x_2, y_2, z_2)) in the crystal lattice such that:

$$u = x_1 - x_2 \quad v = y_1 - y_2 \quad w = z_1 - z_2$$

And that there will be a corresponding peak at:

$$u = x_2 - x_1 \quad v = y_2 - y_1 \quad w = z_2 - z_1$$

Note that a “map” in the context of crystallography refers to a 3-dimensional distribution of maxima, or “peaks” within a unit cell or a series of unit cells. For a unit cell that contains N atoms, the corresponding Patterson map will exhibit N^2 peaks, which can complicate things significantly.

There are N peaks at the origin, representing the vector between each atom and itself, and $N(N-1)$ peaks within the unit cell. This can result in a map that is very congested and difficult to interpret. Note that the Patterson map is centrosymmetric, which further congests cells for non-centrosymmetric lattices.

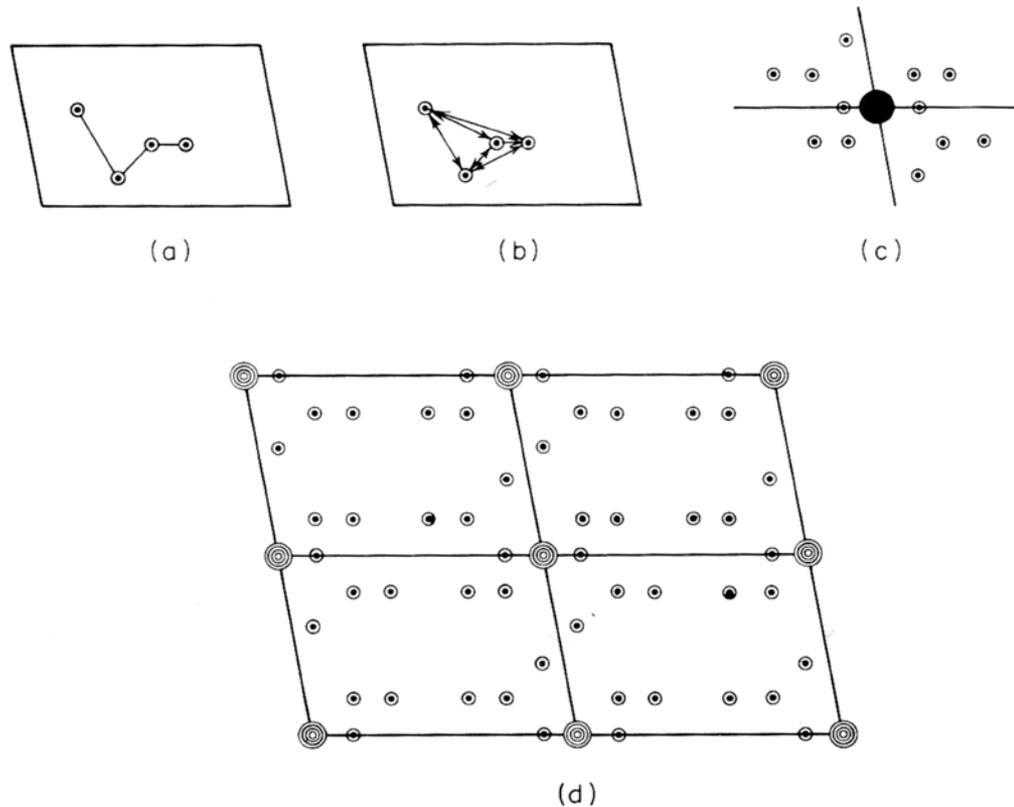


Figure 11.1. (a) Set of points. (b) Interatomic vectors. (c) Patterson peaks about the origin from one set; ●, origin peak. (d) Patterson peaks in four unit cells.

The intensity of each peak is proportional to the product of the atomic numbers of the two atoms to which the vector refers. This makes Patterson functions particularly effective for the identification of the position of heavy atoms (and subsequently the other atoms) when most of the other atoms are light.

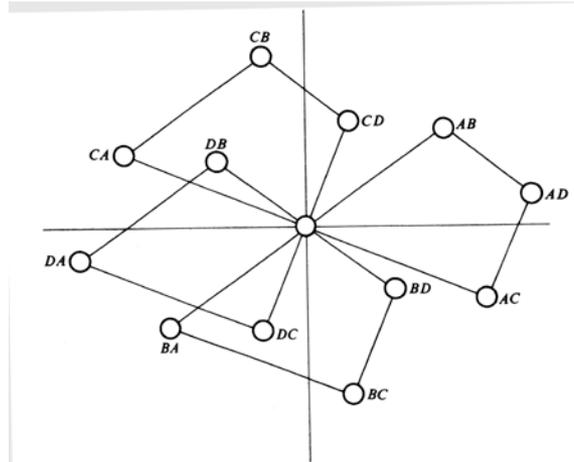
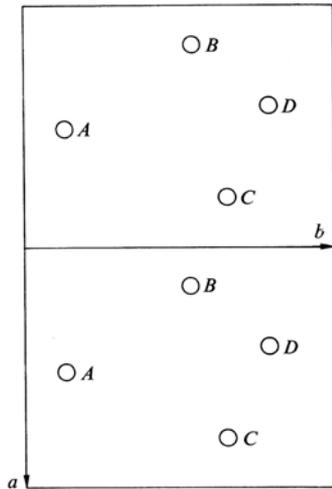


FIG. 6-3 Arrangement of N atoms repeated N times generates the Patterson map. (See Fig. 6-1.)

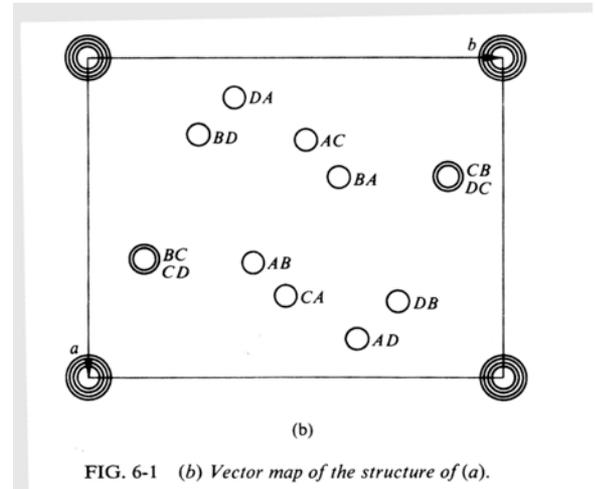


FIG. 6-1 (b) Vector map of the structure of (a).

A Patterson map can be constructed for a unit cell by translating each atom in the cell to the origin and putting a peak of appropriate intensity at the new positions of the other atoms in the cell. Note that the intensity of the origin peak is proportional to: $\sum Z^2$ for all atoms in the unit cell and is usually scaled to e.g. 1000.

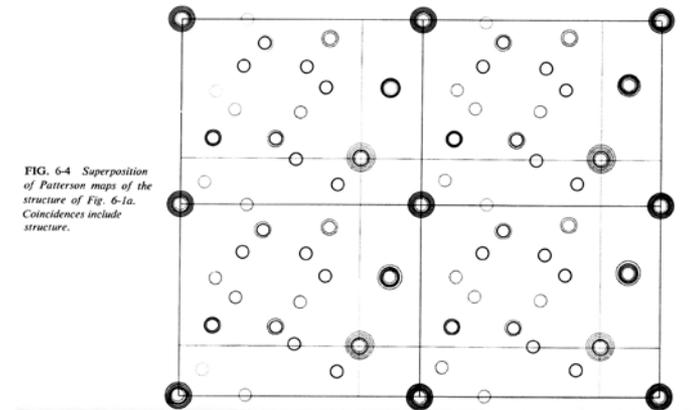
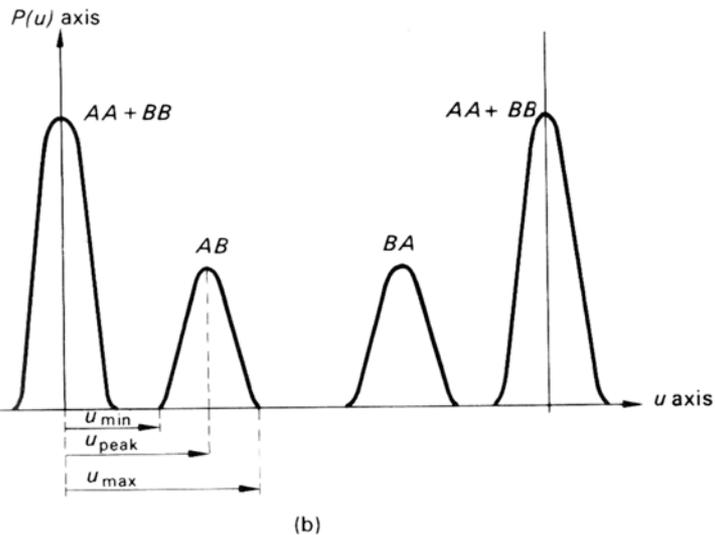
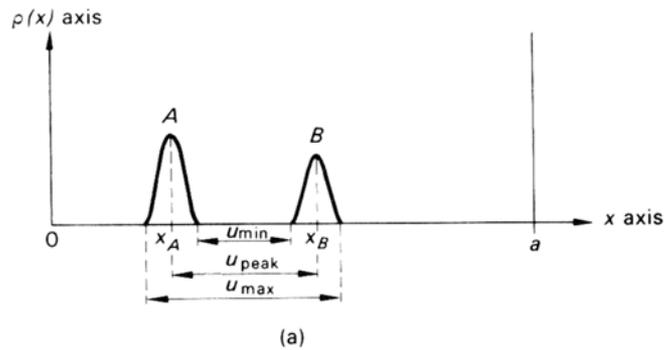


FIG. 6-4 Superposition of Patterson maps of the structure of Fig. 6-1a. Coincidences include structure.

The volume (width) of the peaks in a Patterson map are larger than the peaks that one would find in an electron density map, which can further complicate identification of individual peaks because of excessive overlap. This situation is usually improved in practice by applying a sharpening function that accentuates the peaks.



$$|F_{\text{mod}}(hkl)|^2 = \frac{|F_o(hkl)|^2}{\exp[-2B(\sin^2\theta)/\lambda^2] \{\sum_{j=1}^N f_j\}^2} \quad (6.61)$$

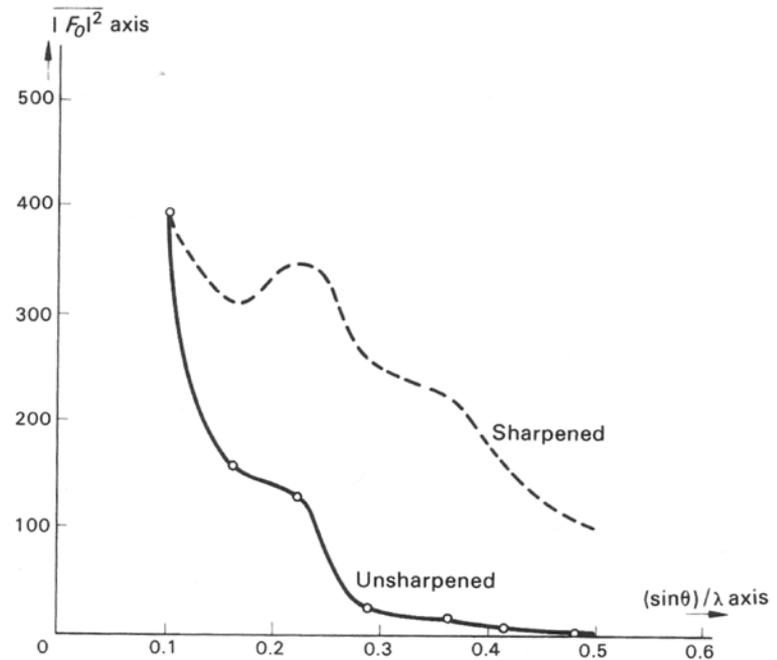


FIG. 6.15. Effect of sharpening on the radial decrease of the local average intensity $|F_o|^2$.

One particularly useful feature of Patterson functions is that the presence of peaks at certain positions provide information about the symmetry elements that are present (**including mirror planes and rotation axes**) in the unit cell and the coordinates of atoms that are related by these elements. The special positions are known as *Harker lines* and *Harker planes* (or sections).

Table 11.1 Some Harker Planes and Lines

2-fold axis $\parallel a, b, c$	$0vw; u0w; uv0$
2-fold screw $\parallel a, b, c$	$1/2vw; u1/2w; uv1/2$
m plane $\perp a, b, c$	$u00; 0v0; 00w$
a glide $\perp b, c$	$1/2v0; 1/20w$
b glide $\perp a, c$	$u1/20; 01/2w$
c glide $\perp a, b$	$u01/2; 0v1/2$

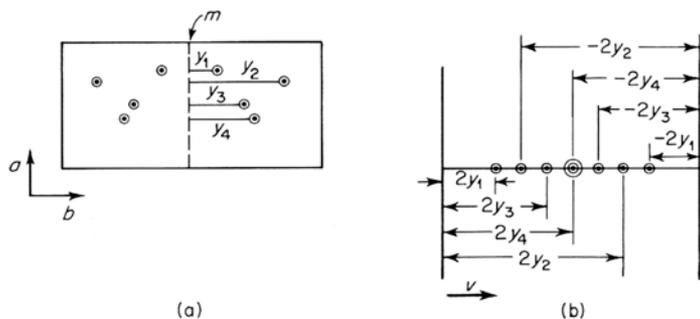


Figure 11.5. (a) Projection of a structure in space group Pm on the ab face. (b) Harker peaks produced along the v axis by the atoms in (a).

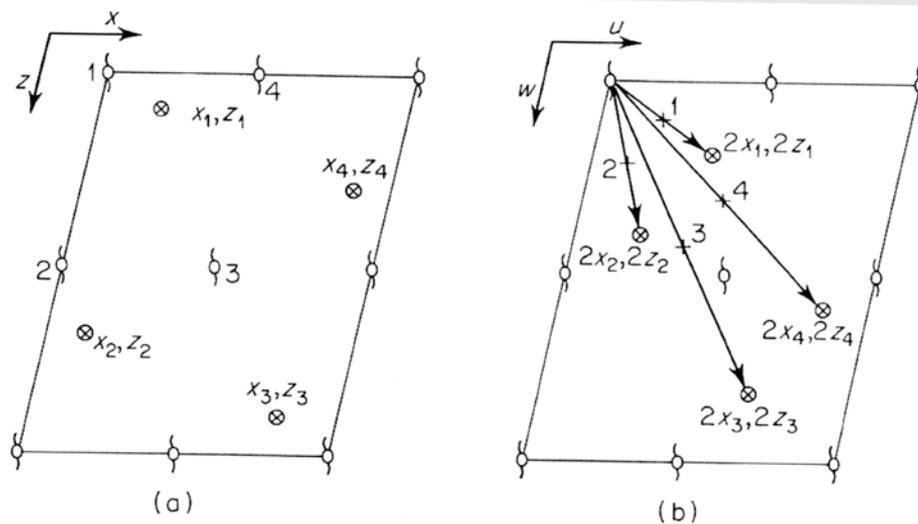


Figure 11.8. (a) Two heavy atoms at x_1, z_1 and x_2, z_2 and their screw-related atoms at x_3, z_3 and x_4, z_4 . In projection. (b) Harker section $P(u, 1/2, w)$ showing the possible atomic locations 1, 2, 3, 4 from peaks $2x_1, 2z_1$; $2x_2, 2z_2$; $2x_3, 2z_3$; $2x_4, 2z_4$. Each location corresponds to choosing the origin on a different screw axis; 1, 2, 3, or 4 in (a).

Patterson methods are often the first choice for the solution of structures containing a few heavy atoms. The initial information provided by the identification of the correct atomic position of the heavy atom and some of the other atoms allows one to begin to calculate structure factors. This initial model can then be refined to determine the rest of the atomic positions (electron density) and thus solve the crystal structure.

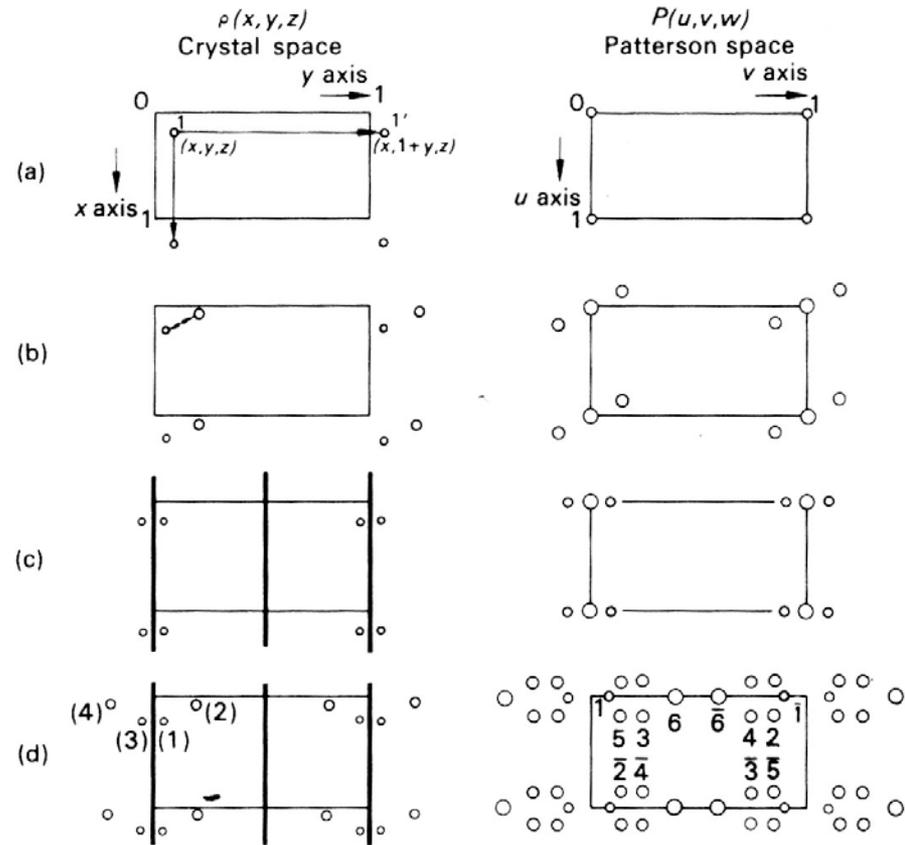


FIGURE 6.14. Effects of symmetry-related and symmetry-independent atoms on the Patterson function. The weights of the peaks are approximately proportional to the diameters of the circles: (a) $P1$ ($N = 1$); (b) $P1$ ($N = 2$)—two atoms per unit cell produce $(2^2 - 2)$ nonorigin peaks; (c) Pm ($N = 2$)—two nonorigin peaks, but with coordinates $\pm\{0, 2y, 0\}$; (d) Pm ($N = 4$)—12 nonorigin peaks per unit cell; for clarity the origin peak has not been drawn. The Patterson space group is $P\bar{1}$ in (a) and (b) and $P2/m$ in (c) and (d). Figure 6.14d is discussed again on page 286.



The calculation of a Patterson map is a tedious process if done by hand (as it was in the past). For a centro-symmetric space group, the equation:

$$P(uvw) = 1/V \sum_h \sum_k \sum_l |F(hkl)|^2 \exp[-2\pi i(h u + k v + l w)]$$

can be replaced with:

$$P(uvw) = 1/V \sum_h \sum_k \sum_l |F(hkl)|^2 \cos [2\pi(h u + k v + l w)]$$

The map is then calculated by determining $P(uvw)$ for increments of each unit cell parameter.

E.g. $(u, w, v) = (0.01, 0, 0), (0.02, 0, 0), (0.03, 0, 0) \dots (1, 1, 0.98), (1, 1, 0.99), (1, 1, 1)$.

The calculated values of $P(uvw)$ at each point are then used to generate the 2-D contour maps or 3-D intensity maps that you can use to determine the atomic positions. The process can be simplified by taking symmetry considerations into account but it is still tedious without a computer.



An example from the Stout and Jensen notes that I gave you in class (see pages 281-284). The a molecule containing a single Br atom crystallizes in the space group $P2_12_12_1$. The Patterson peaks for this space group are given in the following table (from “Patterson Peaks” by D.L. Ward of MSU). Please note that you can replace the “+” signs with “ \pm ” signs in the peak positions.

Orthorhombic	222	D_2^4	$P2_12_12_1$
Patterson symmetry	$P m m m$	$P 2_1 2_1 2_1$	No. 19

Origin at midpoint of three non-intersecting pairs of parallel 2_1 axes

Positions

Multiplicity,
Wyckoff letter,
Site symmetry,
Coordinates

Patterson peaks (U, V, W (Multiplicity))

4	a	1	x,y,z; 1/2-x,-y,1/2+z; etc.			
			0, 0, 0 (4)	1/2+2x, 2y, 1/2 (1)	2x, 1/2, 1/2+2z (1)	1/2, 1/2+2y, 2z (1)

Vectors between two sets of unique atoms

Wyckoff letters

Wyckoff letters

a,	a	$x1-x2, y1-y2, z1-z2$ (1)
		$1/2+x1+x2, y1+y2, 1/2+z1-z2$ (1)
		$x1+x2, 1/2+y1-y2, 1/2+z1+z2$ (1)
		$1/2+x1-x2, 1/2+y1+y2, z1+z2$ (1)

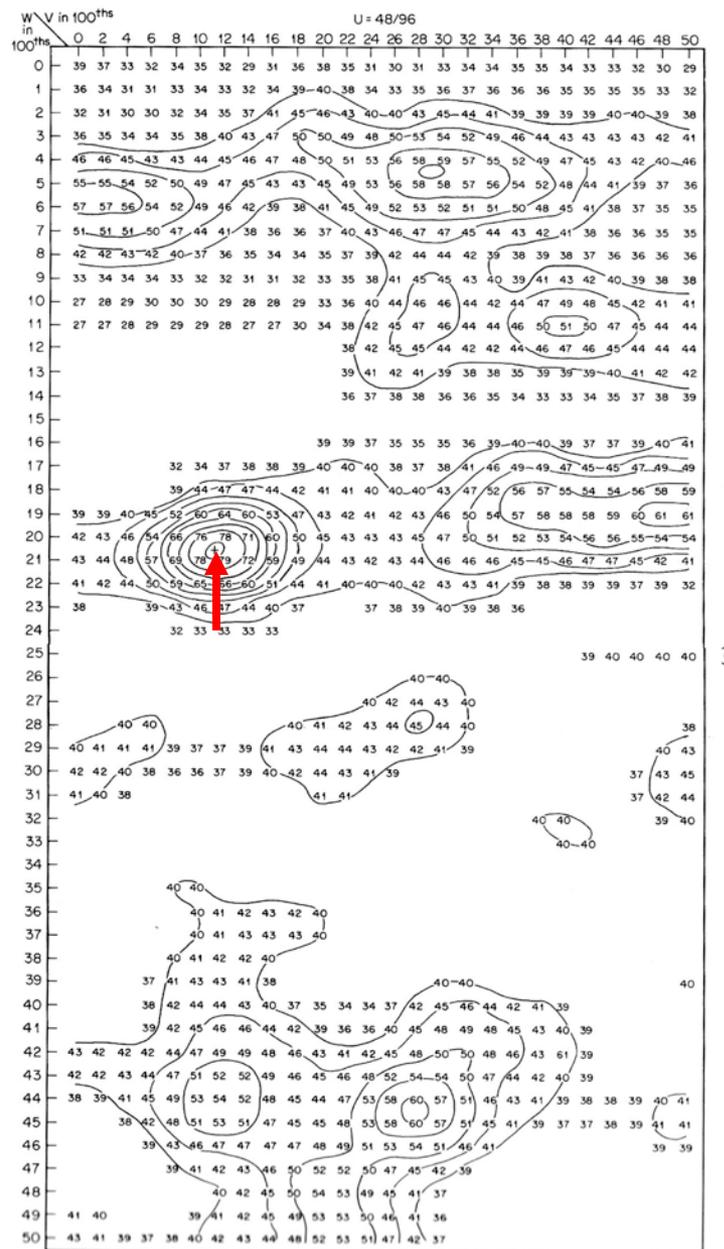


Remember that the most intense peaks in the Patterson maps will be those corresponding to the Br-Br vectors. In the Harker section ($1/2, v, w$) for the 2_1 axis in the a direction, we see a peak at $(48/96, 11/100, 21/100)$. From the table on the last page, this peak must correspond to the real coordinates $(1/2, 1/2-2y, 2z)$, thus:

$$y = (0.11 - 0.5) / (-2) = 0.195$$

$$z = 0.21 / 2 = 0.105$$

Thus we have discovered two of the coordinates and we must find the third from one of the other Harker sections.



In the Harker section ($u, 1/2, w$) for the 2_1 axis in the b direction, we see a peak at $(34/96, 50/100, 29/100)$. From the table on the last page, this peak must correspond to the real coordinates $(2x, 1/2, 1/2-2z)$, thus:

$$x = 17/96 = 0.177$$

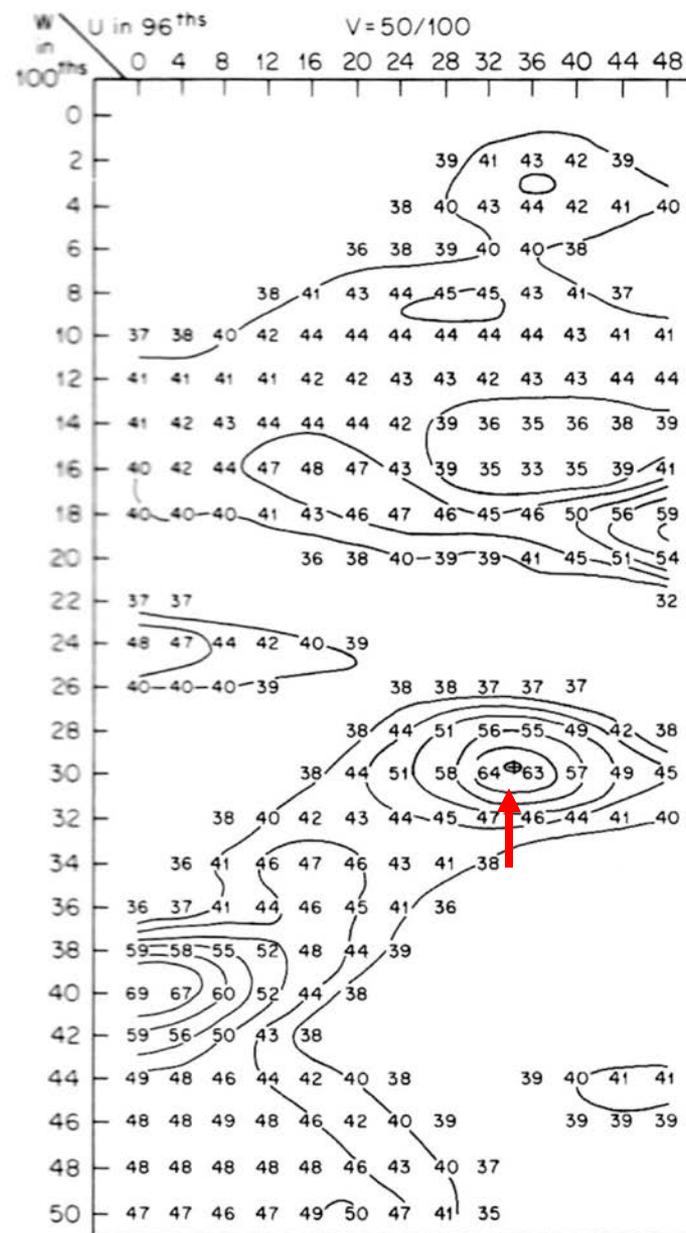
although it is not needed, as a check:

$$\text{does } 0.5 - 2(0.105) = 0.29 ?$$

$$0.5 - 0.21 = 0.29, \text{ so everything is OK}$$

The coordinates of one of the Br atoms in the cell are thus: $(0.177, 0.195, 0.105)$

The other coordinates can be generated by the equivalent position relationships.

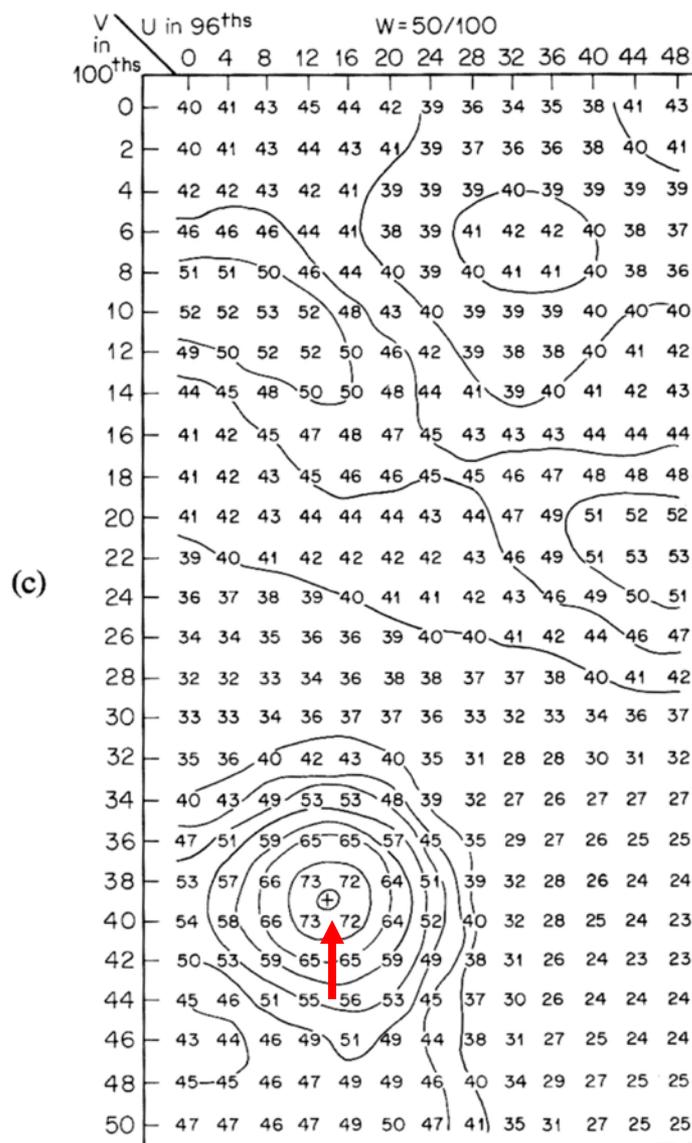


As a final check: In the Harker section (u,v,1/2) for the 2_1 axis in the c direction, we see a peak at (14/96, 39/100, 50/100). From the table on the last page, this peak must correspond to the real coordinates (1/2-2x, 2y, 1/2), thus:

$$x = (0.146 - 0.5) / (-2) = 0.177 \text{ (good)}$$

$$y = 0.39 / 2 = 0.195 \text{ (good)}$$

So everything works out. Note that there are other acceptable solutions available using the various combinations of “+” and “-” signs. These will produce either the a symmetry related set of coordinates or coordinates with the origin shifted (which is not a problem for this non-centrosymmetric space group).





Overall, the application of classical Patterson methods for structure solution can be summarized by the following steps:

1. calculate Patterson function: $P(uvw)$
2. solve for possible Heavy atom position(s)
3. use this position and the electron density to calculate an approximate phase angle (α_{hkl}) for each reflection
4. use these approximate α_{hkl} values to calculate $\rho(xyz)$ and plot it
5. interpret the electron density map in a chemically sensible way

“Direct Methods” are statistical approaches to the phase problem that rely on the (obviously reasonable) assumption that the electron density in the unit cell can never be less than 0, and the (less reasonable) assumption that atoms are discrete entities. The implication of the first requirement is that the phases that we use for the FT of the structure factors must produce a result that has the least amount of “negative” electron density.

Direct methods employs phase relationships that are probably true to generate an initial set of phases/electron density map. Some of the important ideas/relationships used in Direct Methods are presented in the following slides. The mathematics underlying this approach are outlined in the “An Introduction to Direct Methods” handout for those of you who are interested. It should be noted that Direct Methods was the approach used to solve the phase problem for the majority of small molecules for the last few decades.

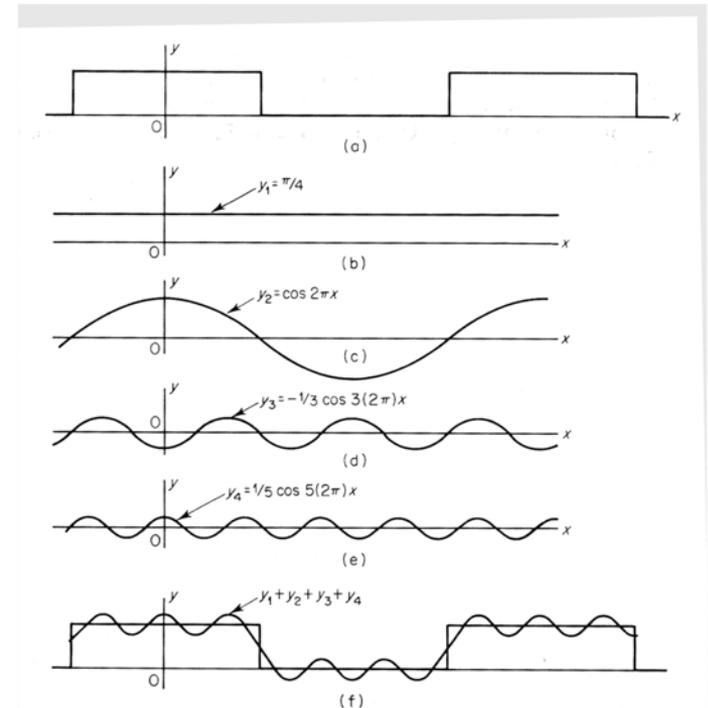


Figure 8.9. (a) A periodic step function. (b to e) Graphs of the first four terms of the Fourier series representing (a). (f) Sum of terms represented by (b to e) as an approximation to the function.



Structure factors with large amplitudes (“strong” reflections) correspond to regions having large amounts of electron density and are thus most important for the determination of phase information. Note that Direct Methods uses “normalized” structure factors $E(hkl)$ instead of $F(hkl)$ to allow the statistical methods to work.

$$E(hkl)^2 = F(hkl)^2 / (\varepsilon \sum f_j^2)$$

Where ε is an integer depending on (hkl) that corrects for the angular dependence of scattering intensity for the given reflection (i.e. the amplitude of $E(hkl)$ does not depend on θ).

Some important relationships used by Direct methods: (note that α_h means $\alpha(h_1k_1l_1)$, α_k means $\alpha(h_2k_2l_2)$ and α_{h-k} means $\alpha((h_1-h_2) (k_1-k_2) (l_1-l_2))$ etc.)

1. For non-centrosymmetric structures: $\alpha_{-h} + \alpha_k + \alpha_{h-k} = 0$
e.g. if $\alpha(-1-10) = -100^\circ$ and $\alpha(2-10) = 30^\circ$ then $\alpha(-300) = -130^\circ$

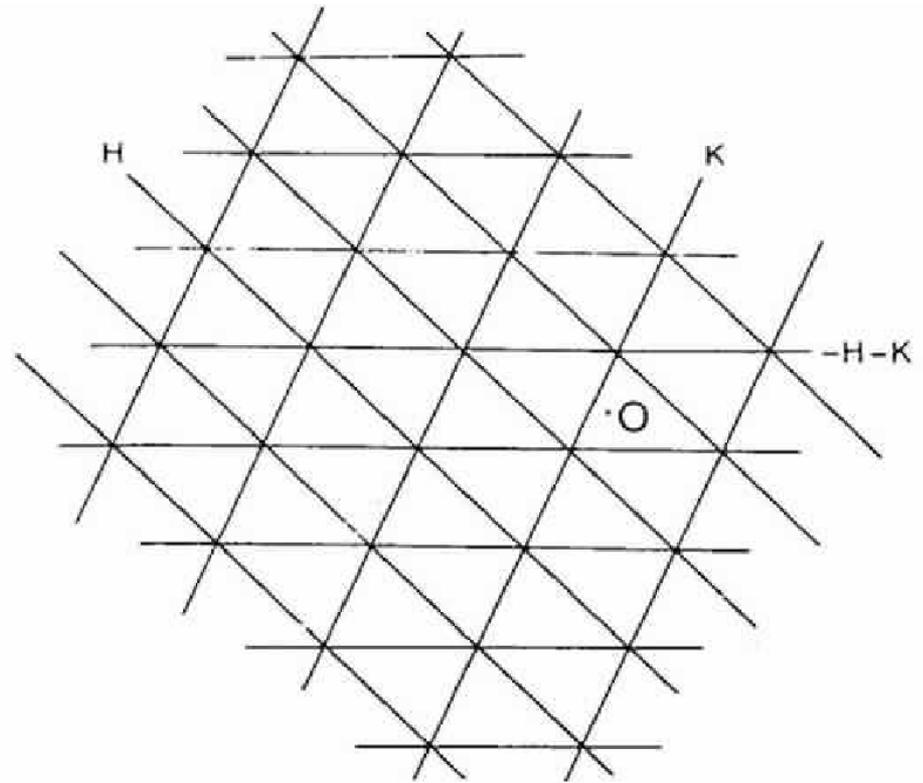
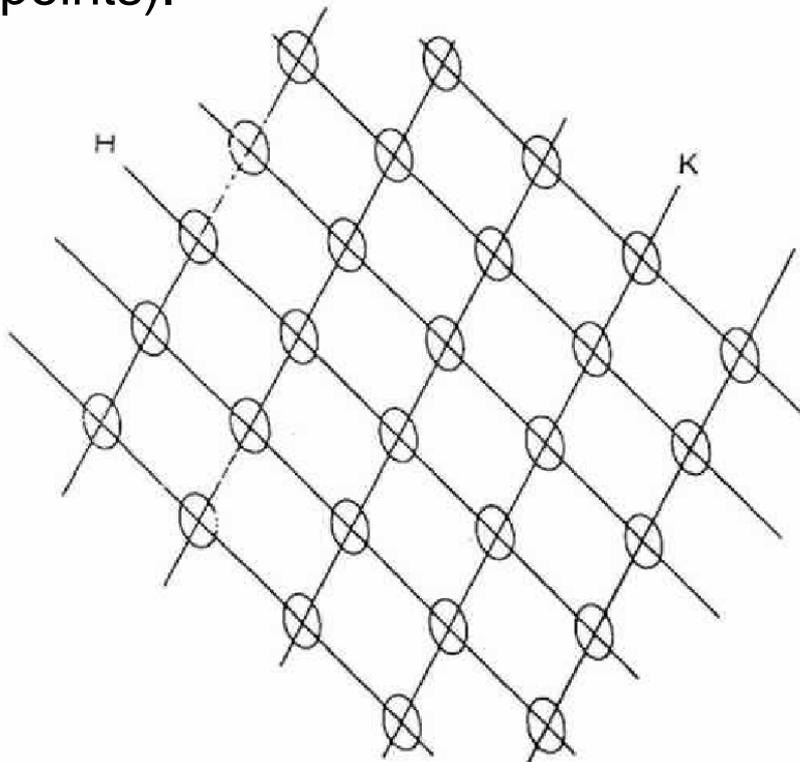
2. For centrosymmetric structures: $S_{-h} \cdot S_k \cdot S_{h-k} = 1$

Where S indicates the sign of the reflection

e.g. if $s(-1-10) = -$ and $s(2-10) = +$ then $s(-300) = -$

3. There are numerous other relationships such as these that are used.
see for example, Structure Invariants, quartet relationships etc. in the handout

A physical and geometrical interpretation of one of these triplet relationships is shown below. Because strong reflections (denoted H and K) correspond to high concentrations of electron density, the intersection of the maxima of the waves (which are indicated by the planes) of two strong reflections is **LIKELY** to be the region of high electron density. In this light, the maxima for another strong reflection are likely to intersect with this point (and thus the phases of these reflections are related since the maxima occur at the same points).



An example of Direct methods for a “1-D” molecule.

For the molecule illustrated below, the $(00l)$ structure factors observed are listed and depicted in the diagram to the right. Using Direct Methods we guess the phases of the most intense reflections and then use phase relationships to build up electron density and avoid negative electron density in the cell. The “guesses” are used to generate an electron density map which we examine to see if it makes sense chemically.

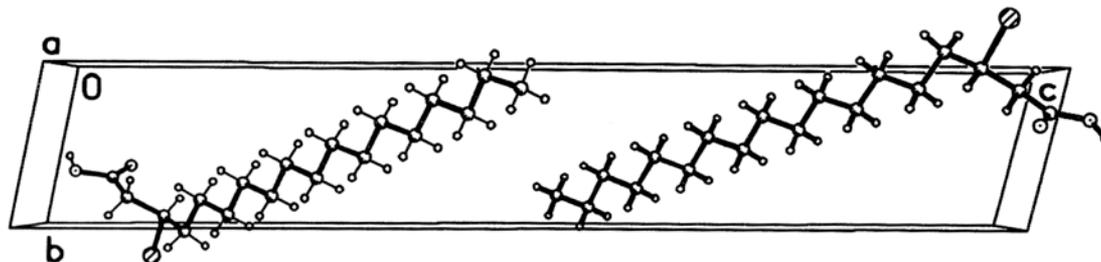


Fig. 2.15 Two molecules in the elongated unit cell of 3-bromo-octadecanoic acid.

Note that the lower angle reflections provide “coarse” estimates and the higher angle data provides finer detail. In this case, the best initial solution is obtained by setting the phases of the (004) , (005) reflections to be negative. This build up amplitude near the first “trough” of each wave they cancel each other in the middle of the cell.

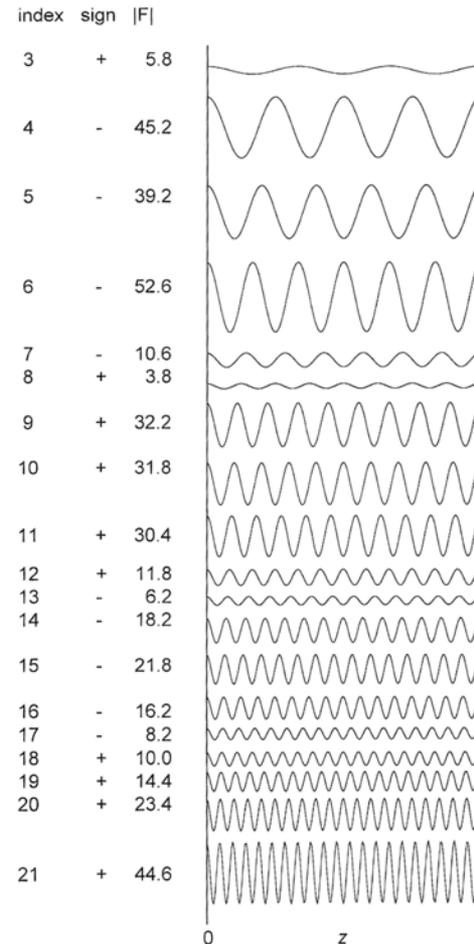


Fig. 2.16 The contributions of the 19 $00l$ reflections to the one-dimensional Fourier summation of equation 2.3, with all their phases set at zero. The correct phases, as signs, are shown with the indices and amplitudes in the left-hand columns. Reflections with larger indices are observed at higher Bragg angles and provide greater resolution of the electron density image, just as light scattered at greater angles by an object on an optical microscope provides better resolution than low-angle scattering. The curves shown here for the different reflections must not be confused with X-ray wavelengths and frequencies (the wavelength is constant for all reflections); these are not the waves themselves, but the contributions they make to the electron density calculation at various points in the unit cell via the one-dimensional Fourier transformation.

Table 2.1 Observed amplitudes and correct phases for 00 l reflections of 3-bromo-octadecanoic acid

l index	Measured $ F(00l) $	Correct sign
3	5.8	+
4	45.2	-
5	39.2	-
6	52.6	-
7	10.6	-
8	3.8	+
9	32.2	+
10	31.8	+
11	30.4	+
12	11.8	+
13	6.2	-
14	18.2	-
15	21.8	-
16	16.2	-
17	8.2	-
18	10.0	+
19	14.4	+
20	23.4	+
21	44.6	+

Since $4+5 = 9$, the sign of the (009) reflection will be positive since: $(-)(-) = (+)$. Using similar phase relationships for the other relatively intense $|F(00l)|$, illustrated in the table below, a reasonable electron density map can be constructed. This is shown in the middle of the figure to the right. Notice the large peaks for the Br atoms and the smaller peaks that correspond to the series of C atoms in the molecule.

Table 2.2 Phase relationships for the strongest 00 l reflections

	<u>4</u>	<u>5</u>	<u>6</u>	9	10	11	<u>14</u>	<u>15</u>	<u>16</u>	20	21
<u>4</u>											
<u>5</u>											
<u>6</u>											
9											
10											
11											
<u>14</u>											
<u>15</u>											
<u>16</u>											
20											
21											

For each relationship the l indices of the three reflections are given by one entry in the table body together with the corresponding number at the head of the column and the number at the left-hand end of the row; an underlined index represents a negative reflection amplitude (for example, reflection 6 has a negative amplitude, but reflection 9 has a positive amplitude).

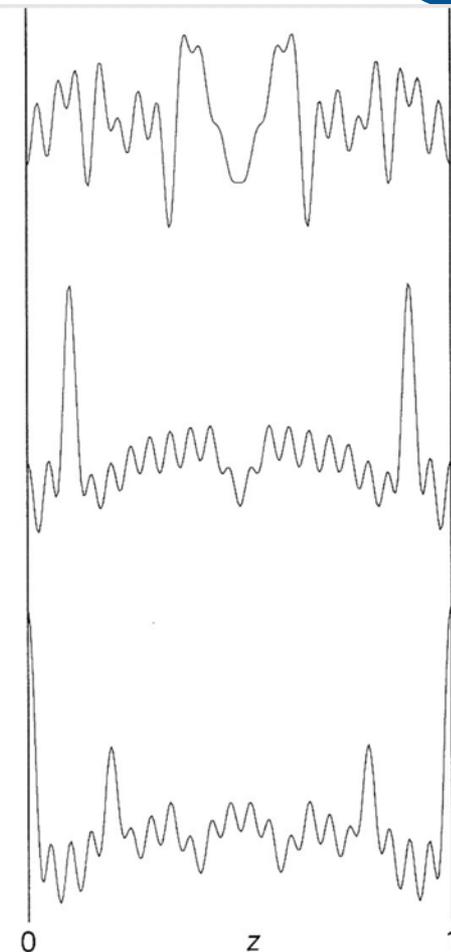


Fig. 2.17 Combinations of the 19 contributions of Fig. 2.16 with different sets of phases: top, randomly chosen phases giving an unrecognizable result; middle, correct phases clearly showing the bromine atoms; bottom, all phases positive, resembling a Patterson synthesis.

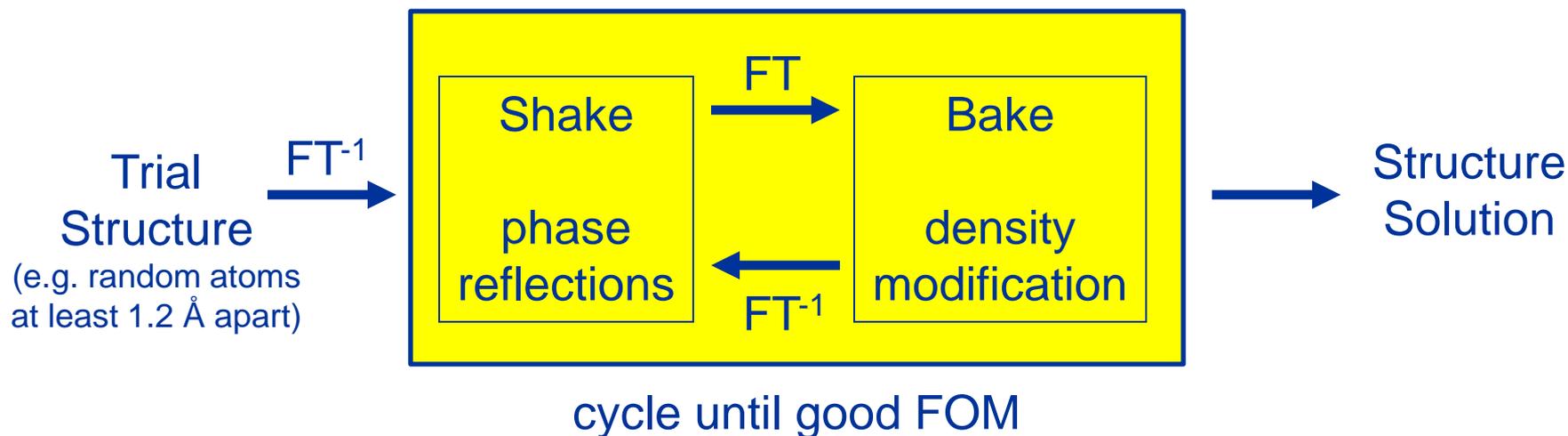


Overall, the application of Direct Methods for structure solution can be summarized by the following steps:

1. Normalize structure factors and identify the strongest reflections
2. Determine the appropriate phase relationships and assign starting phases
3. Use these assignments (guesses) to calculate an approximate phase angle (α_{hkl}) for many other strong reflections
4. Assess the statistical reasonability of the phase set using a figure of merit (FOM)
5. Use these approximate α_{hkl} values to calculate $\rho(xyz)$ and plot it
6. Interpret the electron density map in a chemically sensible way (if it is no good, go to step 2 and repeat with new phases)



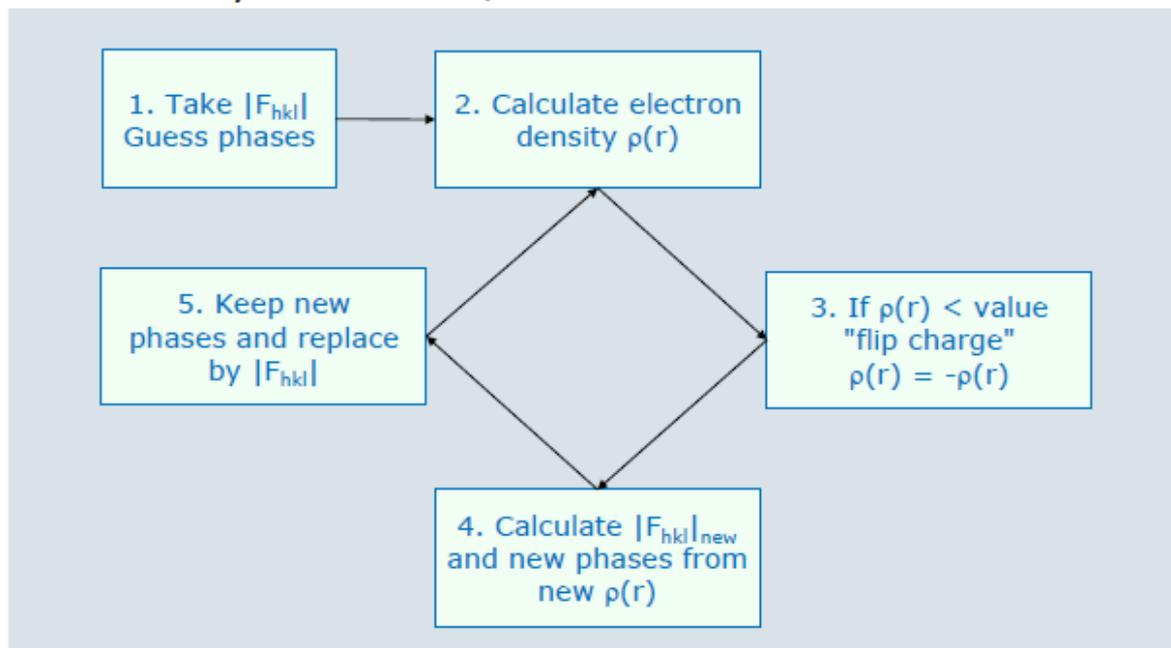
Traditional direct methods approaches work in reciprocal space (i.e. on the structure factors only) – although we examine the real space (electron density) result of the FT^{-1} afterward in order to assess if the solution is chemically reasonable, no information from the electron density map was used to generate the solution. Some more modern direct methods approaches actually use information from trial electron density maps (real space) to improve the phasing assignment (reciprocal space). These methods include “Shake and Bake”, “Half-bake” (SHELX-D), and the SIR200N implementations (see C. Giacovazzo, Phasing in Crystallography; IUCr, 2014).





“Dual Space Methods” (*Charge Flipping*: SuperFlip, SHELXT, OLEX, TOPAS, and others; VLD: SIR2011+) are approaches to the phase problem that rely on the (obviously reasonable) assumption that the electron density in the unit cell can never be less than 0 (similar to Direct Methods) but that proceed iteratively to localize electron density concentration. These methods do not require knowledge of the composition of the structure (e.g. heaviest atom present) or of the space group symmetry of the unit cell.

Charge Flipping Oszlányi and Sütő, 2004





There are many more atoms in macromolecules so solving the phase problem becomes much more difficult. Methods to overcome this problem include:

1. MIR: Multiple Isomorphous Replacement.

In this technique, we replace lighter atoms with heavier analogues and trust that the overall structure does not change (e.g. replace S with Se, or Cl with Br or I). This allows for Patterson methods to be used for initial phasing.

2. MAD: Multi-wavelength Anomalous Dispersion

In this technique, multiple-wavelengths of radiation are used on the same (native) crystal in order to solve the phase problem (see http://skuld.bmsc.washington.edu/scatter/AS_index.html); this technique requires a synchrotron to do.