INTERNATIONAL ATOMIC ENERGY AGENCY
UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION
INTERNATIONAL CENTRE FOR THEORETICAL PHYSICS
I.C.T.P., P.O. BOX 586, 34100 TRIESTE, ITALY, CABLE: CENTRATOM TRIESTE

H4.SMR/642 - 7

College on Methods and Experimental Techniques in Biophysics

28 September - 23 October 1992

Protein Crystallography

C.E. BUGG

University of Alabama, U.S.A.

These are preliminary lecture notes, intended only for distribution to participants.

PROTEIN CRYSTALLOGRAPHY

Charles E. Bugg
University of Alabama at Birmingham

References:

1. T.L. Blundell and L.N. Johnson, "Protein Crystallography," Academic Press: New York (1976)

2. A. McPherson, "Preparation and Analysis of Protein Crystals," John Wiley & Sons: New York (1982)

3. A. McPherson *in*: "Methods in Enzymology," Vol. 114, Diffraction Methods, pp. 77-112. Eds. C.H.W. Hirs, S.N. Timasheff and H.W. Wyckoff, Academic Press: New York (1985)

PROTEIN CRYSTALLOGRAPHY

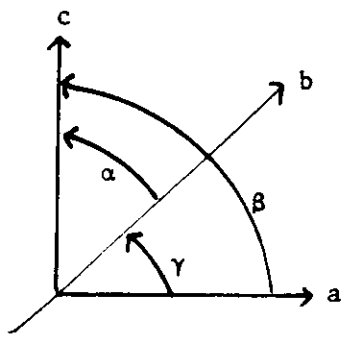
Crystals represent atomic or molecular aggregates which repeat themselves indefinitely in space.

Repetition means an operation by which the system is brought into a state indistinguishable from the initial state.

A **crystal lattice** is defined by a combination of three vectors, which must not be coplanar, multiplied by all of the positive and negative integers.

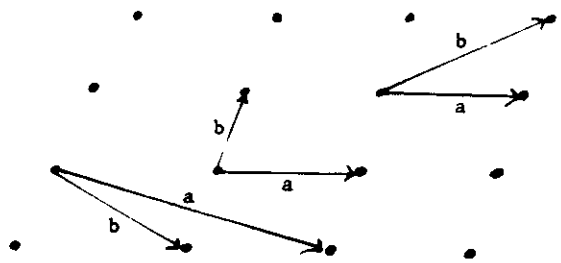
A **lattice** has six defining parameters:

three distances a, b and c
three angles α, β, and γ



The lattice is strictly a geometrical concept, and should not be associated with atoms or molecules. Only in very special cases, such as in the structures of metals, will there be any direct relationship between the lattice and the atomic arrangement.

In general, there are an infinite number of ways in which the vectors a, b and c can be chosen to represent any given lattice: for example -

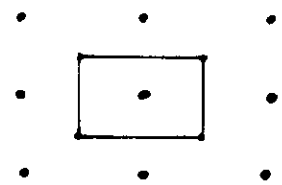


As a rule, the vectors a, b and c, referred to as the unit cell translations, are chosen so that they are the shortest three axes not in the same plane.

If it is possible to choose the axes such that one or more of the interaxial angles are equal to 90°, then the choice is made on this basis.

A cell with only one lattice point is called a primitive cell (denoted by P).

In addition to the lattices which have only one lattice point per unit cell, it is sometimes necessary to establish a lattice which has more than one lattice point per cell in order to end up with interaxial angles of 90°.



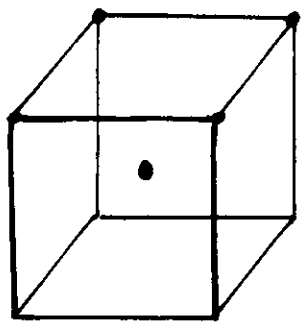
(Point in center belongs completely to the cell; points at corners are shared with other cells and only 1/4 of corner points belong to the cell.)

In three dimensions, centering can result from either an extra lattice point in the very center of the cell (body-centering denoted by the symbol I) or be centered on the faces of the cell.

Centering on faces can be of two basic types:

1. Centering on all of the faces (denoted by symbol F);
2. Centering on only the A, B or C faces (where the A face is a plane perpendicular to A, etc. If there is no lattice plane perpendicular to a given axis, then this type of centering is impossible). Face centering of this type is denoted by A, B or C depending on which face is centered.

e.g.



I Centering

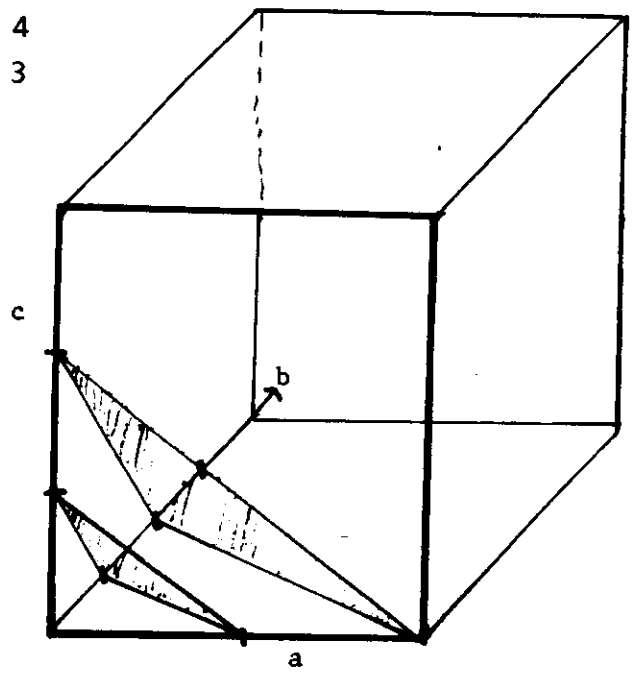
It is useful to have a nomenclature describing all the possible sets of planes within a lattice which can form crystal faces (and which, as we will see later, can also reflect X-rays).

The planes are described by a set of three integers h, k l, referred to as the Miller indices.

The plane corresponding to set of Miller indices h, k, l divides the a axis into h equal parts, the b axis into k equal parts, and the c axis into l equal parts. Thus, the plane intersects the a axis at a/h
b axis at b/k
c axis at c/l

e.g., take plane for which h = 2
k = 4
l = 3

written as 2 4 3.



Plane intersects a = 1/2 a
b = 1/4 b
c = 1/3 c

If this procedure is continued to include the other sets of division points, a family of parallel planes is obtained.

The perpendicular distance between these crystal planes (denoted as h k l) is an important property of the lattice which can be measured by X-ray diffraction techniques.

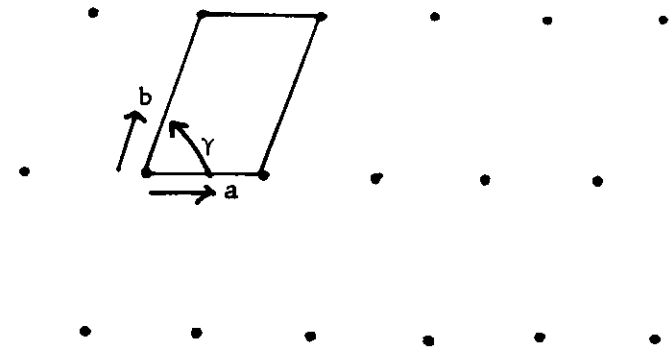
In real terms, the planes, like the lattice, have only geometrical meaning.

Most of what we know about protein structure has been learned by X-ray crystallography. Crystallography can be used to determine the structure of any material that can be crystallized.

The advantage of a crystal is that it contains many sets of molecules all arranged in identical orientation. Thus, we are able to get diffraction effects and the diffracted signals are amplified by the zillions of identical repeating units.

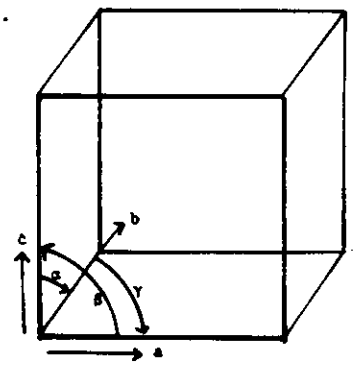
IN SUMMARY,
The scheme that establishes the repeating units in these crystals is the lattice, which is a geometrical construction that describes the repeat vectors. There are 14 different types of lattice systems.

In two-dimensions, it looks like:



Each vector movement $\vec{a}$ or $\vec{b}$ takes us from a position to an identical position elsewhere in the repeating array. We can define this lattice by the lengths of a and b, plus the angle γ between these two vectors.

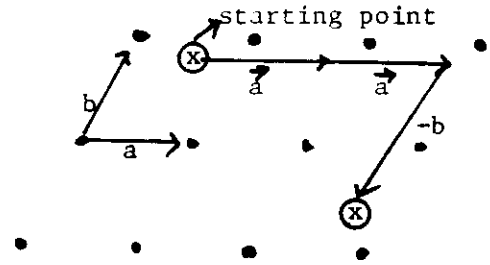
For a three-dimension lattice, we would have an identical vector $\vec{c}$. These three vectors $\vec{a}$, $\vec{b}$, and $\vec{c}$ define a "box" which is called the **unit cell.** This unit cell is the fundamental unit that is repeated throughout the crystal.



This box can also be defined by lengths of $\vec{a}$, $\vec{b}$ and $\vec{c}$ and the intervector angles α, β and γ.

The complete crystal is then made up of the boxes or unit cells repeated continuously by the three vectors $\vec{a}$, $\vec{b}$ and $\vec{c}$.

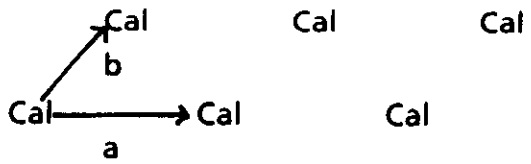
No matter where you start in this crystal, whether on the edge of an oxygen atom, or the center of aromatic ring, an identical position will be reached by any combination of $\vec{a}$, $\vec{b}$ and $\vec{c}$ movements.



A new position is reached by taking ⓧ and moving it $2\vec{a} - \vec{b}$.

What is found within these unit cells that are repeated continuously through a protein crystal?

The unit cell can contain a single protein molecule. For example, in crystals of calmodulin, which is a calcium-binding protein that we are currently investigating, there is one single molecule in the cell, and the structure would be

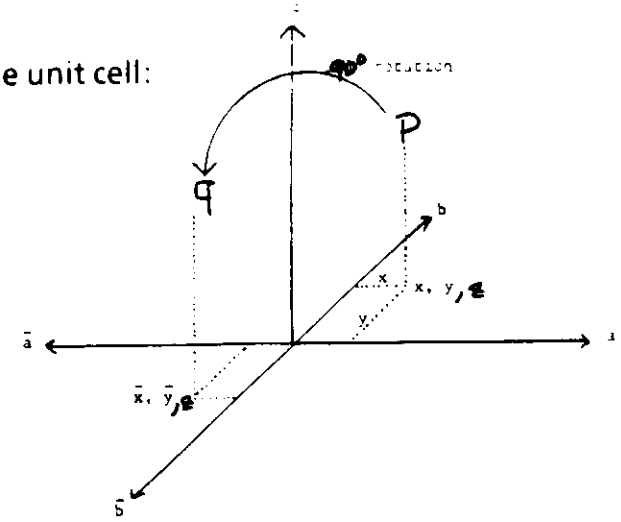


However, it is much more common to find an aggregate of protein molecules in the cell. Each unit cell would have an identical set of molecules, which are then repeated to define the crystal.

The individual molecules within this aggregate are generally related by symmetry, or if the protein contains subunits, the individual subunits may be related by symmetry.

For example, a common symmetry element in protein crystals is the rotation axis. These rotation axes can be 2-fold (very common), 3-fold, 4-fold, etc.

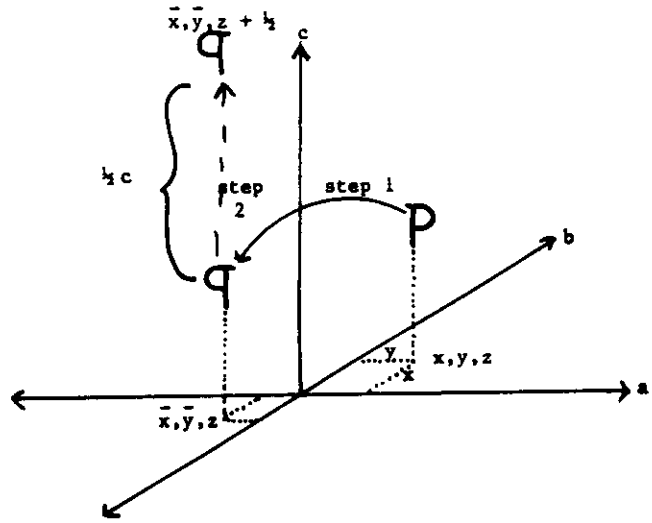
Consider 2-fold rotation around the z-axis of the unit cell:



Every point in the unit cell at x y z will give rise to an identical point at -x, -y, z.

Another common type of symmetry element that relates protein molecules is the screw axis, which combines rotation with translation.

For example, a z-fold screw axis parallel to the C direction would combine a 180° rotation around C with a translation in the C direction of 1/2 C.



step 1: $x, y, z \longrightarrow \bar{x}, \bar{y}, z$

step 2: $\bar{x}, \bar{y}, z \longrightarrow \bar{x}, \bar{y}, z + 1/2\ C$

Most unit cells contain combinations of symmetry elements, so that there are several different molecules contained in the cell, all of which are related by symmetry elements.

Only certain combinations of lattice types and symmetry elements are allowed. The allowed combinations are called **space groups** and there are a total of 230 different possible space groups.

The space groups are designated symbolically, like

$$P\,2_1\,2_1\,2_1$$

primitive lattice 2-fold screw axes parallel to a, b and c

Generally, the space group can be readily idenfied by examining the symmetry present in diffraction patterns from the crystal.

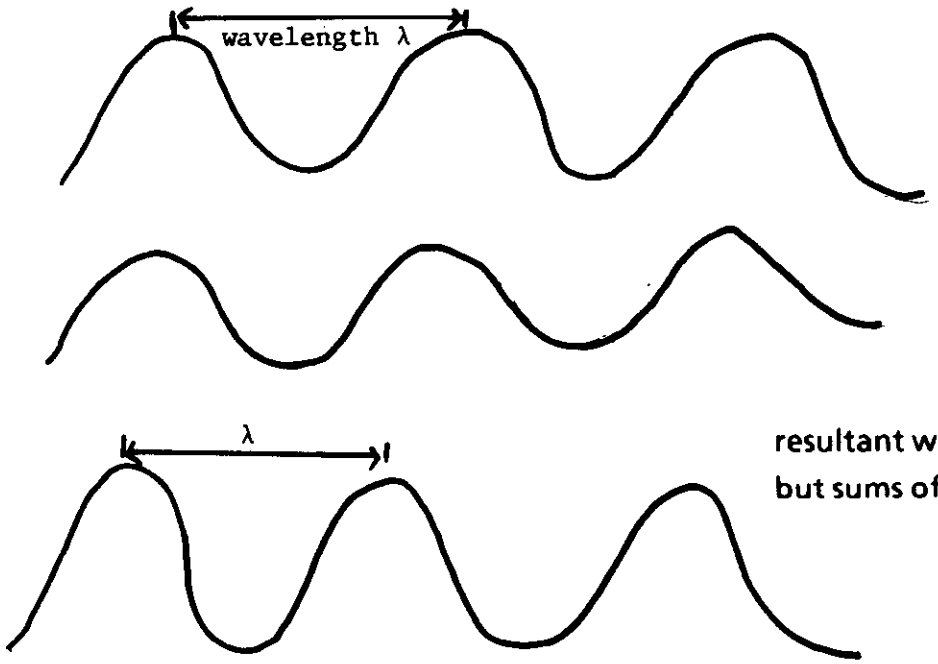
If the space group is known, then it is only necessary to locate one protein molecule and the others can be generated via the symmetry elements.

Why do we observe diffraction when a protein crystal is placed in an X-ray beam?

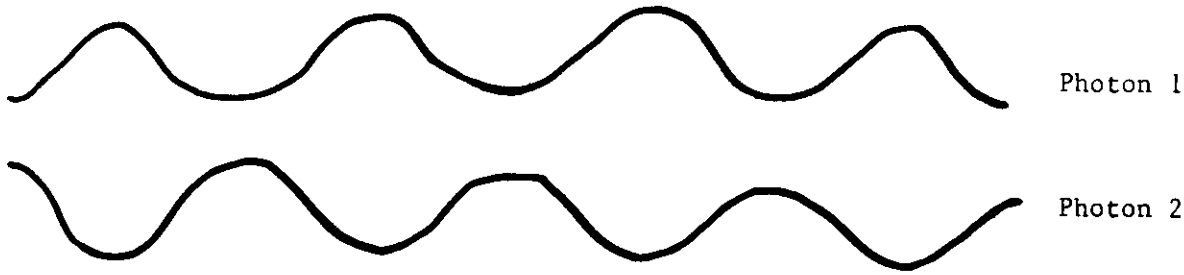
X-rays are photons, and have wave properties. They can interact with electrons; the electrons then scatter the X-ray photons in all directions.

Photons interact with each other in a manner that can be explained in terms of their wave character. Two extreme modes of photon interaction are constructive and destructive interference.

1. Constructive interference: waves are in phase

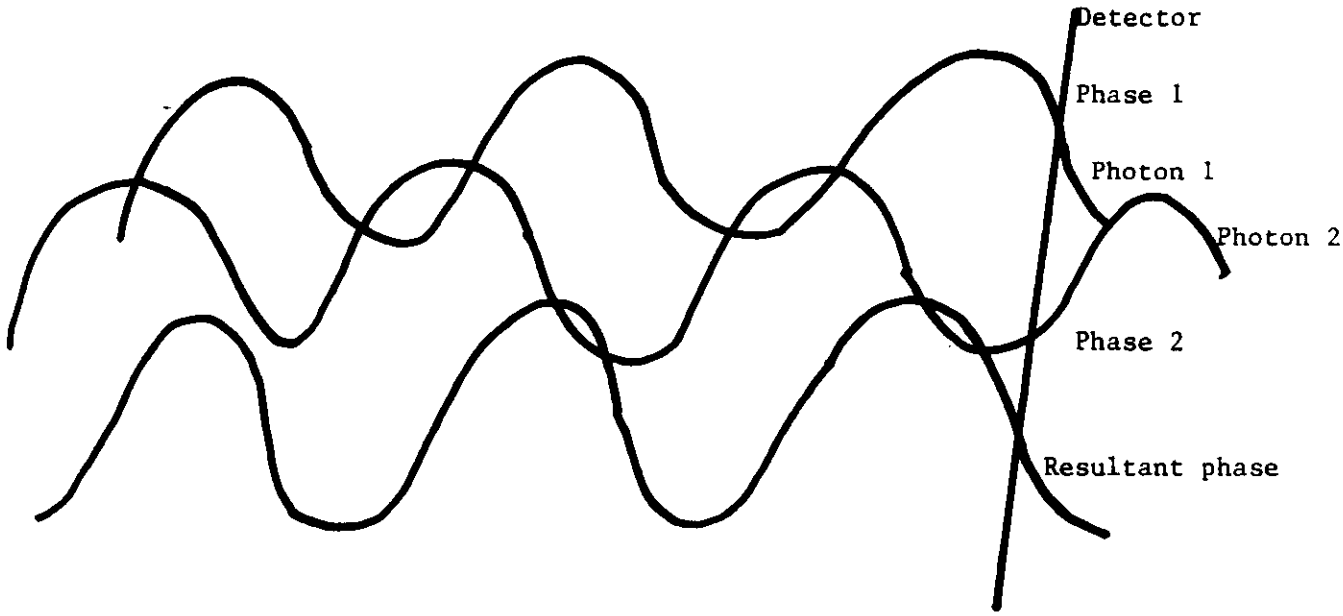


2. Destructive interference: waves are 180° out of phase



The waves cancel each other and resultant amplitude is zero.

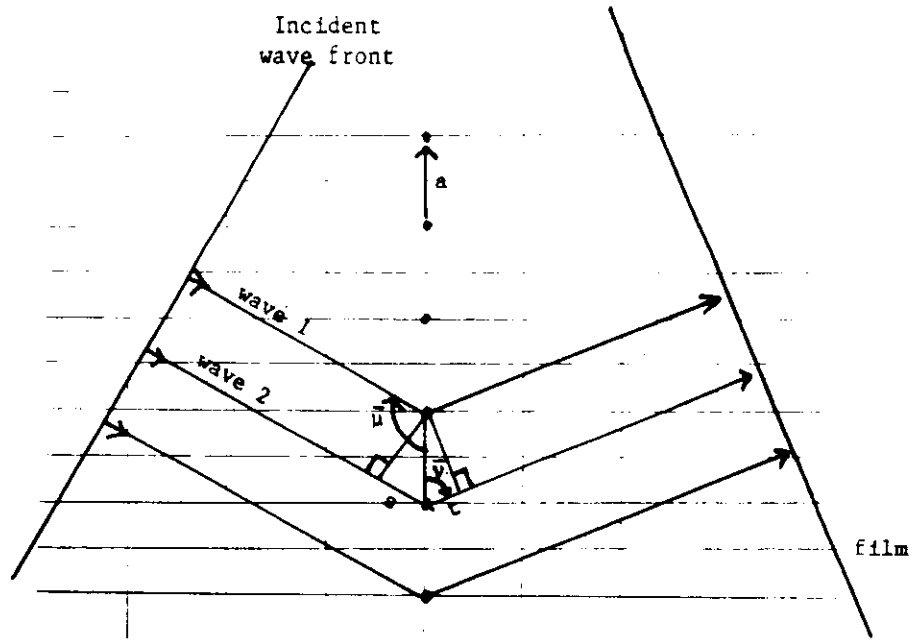
If the waves are partially out of phase, they will result in a wave that has an amplitude and a phase angle (*e.g.*, relative position of peaks and troughs) that are intermediate between constructive and destructive interference.



The actual amplitude of the resultant wave can be measured as the amount of darkening on a photographic film or the signal in a scintillation counter.

Unfortunately, there is no way at present to measure the relative phase angles of the waves - this is known as the **phase problem**. We will return to it later.

How does the presence of a repeating array, such as a lattice, affect the interactions between X-ray photons? First consider a one-dimensional lattice:



Complete reinforcement will occur when the distance traveled by Wave 1 differs by an integral number of wavelengths from the distance traveled by Wave 2.

i.e.,

$$S + t = h\lambda \qquad h = \text{integer}$$

$$\text{Sin}(\vec{\mu} - 90°) = S / a$$

$$-\text{Cos}\ \vec{\mu} = S / a$$

$$S = -\text{Cos}\ \vec{\mu} \cdot a$$

$$\text{Cos}\ \vec{\nu} = t / a$$

$$t = \text{Cos}\ \vec{\nu} \cdot a$$

$$S + t = -\text{Cos}\ \vec{\mu} \cdot a + \text{Cos}\ \vec{\nu} \cdot a = h\lambda$$

$$\text{Cos}\ \vec{\nu} - \text{Cos}\ \vec{\mu} = h\lambda / a$$

When this geometrical relationship is satisfied the photons coming off of all adjacent points will be in phase and a signal will be observed on the film. If the adjacent photons are even slightly out of phase, then they will eventually cancel each other out over the trillions of repeats that are found in a crystal. Thus, there will be a sharp "peak" observed when their geometry is followed.

If we have a 3-D lattice, the other lattice translations b and c will generate corresponding equations.

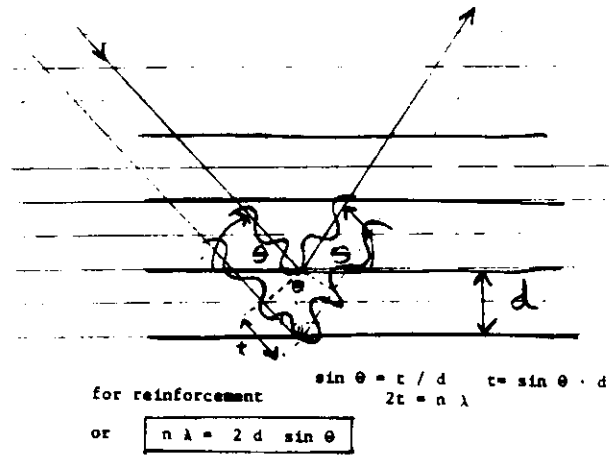
The complete set will then be

$$\cos \bar{\nu}a - \cos \bar{\mu}a = h\lambda / a$$
$$\cos \bar{\nu}b - \cos \bar{\mu}b = k\lambda / b$$
$$\cos \bar{\nu}c - \cos \bar{\mu}c = l\lambda / c$$

These are known as Laue's equation.

When all three equations are satisfied, only a vector spot will appear on the film.

The integers h, k, l are known as Miller indices and each individual spot can be assigned to a set of these indices.

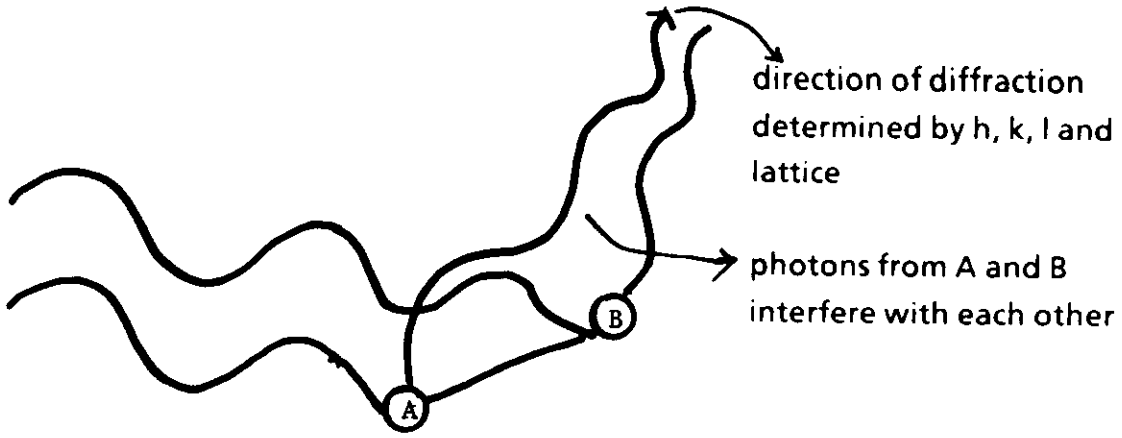


So all of the photon sets that come off of the unit cells are in phase with each other - i.e., unit cell 1 is in phase with unit cell 2, etc. And the directions in which diffraction can occur will be determined by the lattice repeats.

What happens within the unit cells?

If we have a cell full of atoms, we have a large number of electrons, all of which will scatter the X-rays. There will be a variety of orientations, depending on the actual molecular structure, and orientation of the molecules within the cell.

We will get interference between the photons coming off of the individual atoms, and the extent of interference will contain information about the atomic arrangements.



This same event happens within each cell in exactly the same way.

As a result, the intensities of the different diffracted beams will be different and will contain detailed information about the arrangement of atoms.

This information is all contained in the structure factor

$$|F_{hkl}| = \sqrt{I_{hkl}}$$

$$|F_{hkl}|^2 = \left(\sum_{j=1}^{\#\text{ atoms in cell}} f_j \cos 2\pi (h x_j + h y_j + l z_j) \right)^2 + \left(\sum_{j=1}^{\#\text{ atoms in cell}} f_j \sin 2\pi (h x_j + h y_j + l z_j) \right)^2$$

(f_j: scattering power of each atom $\propto$ # of electrons)

Although we generally have many more structure factors than parameters, we are not able to solve these non-linear equations directly.

A more direct representation of the relationship between the structure factors and the molecular arrangement is the electron density map. This is a map that shows the distribution of electrons in the unit cell. At atomic positions, the electron density is very high, so such a map can be used to construct a model of the protein structure, if the necessary information is available.

$$\rho(x,y,z) = \sum_{\substack{\text{all h,k,l} \\ \text{values}}} |F\,h\,k\,l| \cos[2\pi(h\,x + k\,y + l\,z) - \alpha\,h\,k\,l]$$

This gives the value of the electron density at the point x y z in the unit cell. If we divide the cell into a fine grid of x y z values, and calculate the electron density at each of these grid points, we would then have a "map" of the structure.

How do we get $\rho(x,y,z)$?

We can measure $|F\,h\,k\,l|$ since it is the value of $\sqrt{I\,h\,k\,l}$

Unfortunately, we cannot measure $\alpha\,h\,k\,l$, the phase angle for the h k l reflection.

There are various special methods for obtaining $\alpha\,h\,k\,l$, but the one that is commonly used for proteins is the **isomorphous replacement method**.

By this technique, one complexes metals or other heavy-atom multi-electron complexes to the protein. By measuring the change in $|F\,h\,k\,l|$ when the metal is bound, it is possible to establish where in the unit cell the metal is going and to use this information to establish the $\alpha\,h\,k\,l$ values. The heavy atoms are generally diffused into the protein crystals, through large aqueous channels that permeate these crystals.

Normally, there isn't much that can be done with the native data alone. However, there are several exceptions:

1. If the molecule possesses non-crystallographic symmetry, Patterson search methods can be done to locate the symmetry elements; *e.g.*, if the molecule is a subunit enzyme which possesses a rotation axis, the axis may or may not correspond to a crystallographic symmetry element. If not, the Patterson may be searched using a **rotation function** to determine the orientation of the non-crystallographic element. Then a **translation function** can be used to determine where in the cell the symmetry element is placed.

2. **Molecular replacement**. If one has a new crystal form of a molecule where the structure has already been determined, it is possible to search the Patterson for the orientation and position of the molecule. Given this, and assuming that the molecular structure is approximately the same in the 2 crystal environments, one can then take the new coordinates, calculate phase angles and use difference Fourier maps to refine the structure.

However, usually it will be necessary to prepare heavy-atom derivatives of the protein in order to determine its structure. We will study this in great detail, but for right now, we will examine the principal steps:

1. One or two crystals are transferred to 1ml of a solution in which they are not soluble. Usually, this is a solution of normal mother liquor, with a slightly elevated concentration of precipitating agent, *e.g.*, if it crystallizes from 60% A.S., then the solution might be 65% A.S.

2. A stock solution of the heavy-atom compound, in the mother liquor, is prepared, *e.g.*, 1ml of .1M of the heavy atom.

3. The heavy-atom solution is added to the crystal-containing solution (usually gradually), until the desired concentration of heavy-atom is attained. Usually, 0.001M is a good starting point in a search.

   If the crystal cracks, then lower concentrations would be tested. If the heavy-atom doesn't appear to bind (as evidenced by diffraction pattern), then higher concentrations can be tried. Usually, 0.01M is the upper limit.

The idea is to find the concentration where the strong, specific sites will be occupied, but where the multiple non-specific weak sites will be empty.

4. Soak time is often an important variable. The crystals should be observed frequently to see if cracking or glazing is appearing. Usually, a photo will be tried after 24 hours. Longer soak time may be required.

5. After soaking, the crystal is mounted as usual and a precession photo is taken. Normally, one photo is sufficient, and one tries to select a centric zone if possible.

If changes are observed in the intensities (or seen visually), then the derivative is possible. Then it would be a matter of finding the concentration and time limits for producing consistent changes in the pattern.

If no changes are observed, then the soak concentration might be increased and the soak time lengthened.

Once the conditions are established, and the changes are convincing, data are collected.

It is important to follow exactly the same data collection procedure as used for native (*e.g.*, reflections measured in same order, same standards used, crystal mounted in same orientation, *etc.*)

The native set and the derivative set are then placed on a common scale.

Now we are ready to locate the heavy atoms and begin the phasing procedure. Before going into the next stages, we need to digress somewhat and look at the relationship between the structure factors and phases in a slightly different way.

Nomenclature:

(a) F_P = structure factor for protein. Unless written as the absolute value, we are talking about a vector of magnitude $|F_P|$ and phase angle α.
$F_P = F_P(hkl)$ - I'll leave off h k l.

(b) F_{PH} = structure factor for the heavy-atom derivative of the protein.

(c) F_H structure factor contribution of the heavy-atoms alone.

(d) α or $\alpha(h\,k\,l)$ = the phase angle for the h k l reflection.

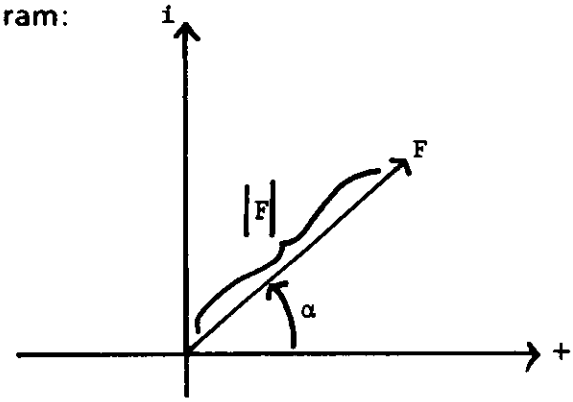
The structure factor (any of the above) can best be represented as a complex number:

$$F(h\,k\,l) = |F(h\,k\,l)| \exp[i\,\alpha(h\,k\,l)]$$

where $i = \sqrt{-1}$

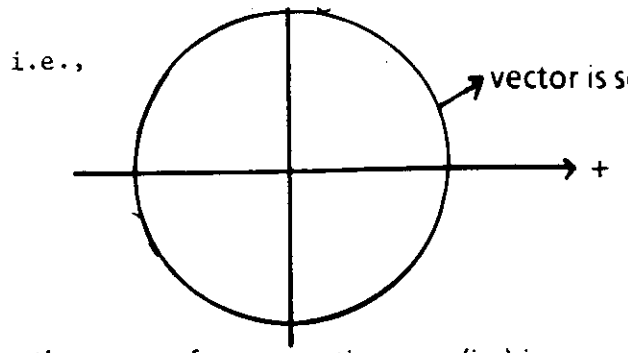
In the future, I will write this simply as $F = |F| \exp.(i\alpha) = |F| e^{i\alpha}$

A complex number of this type can in turn be represented as a vector in the complex plane or Argand diagram:



where the phase angle α can range from 0 to 2π (0 to 360°).

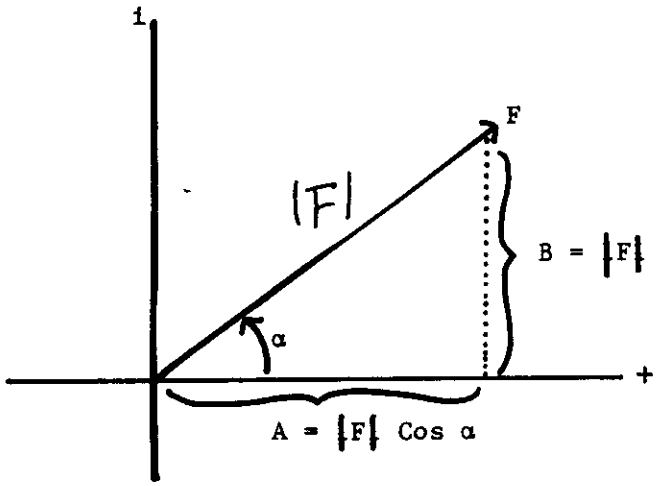
If one knew the magnitude, but not the phase angle, then one would only know the length of the vector in the complex plane, but not its phase angle α.



Another way of representing exp (i α) is

$$e^{i\alpha} = \cos\alpha + i \sin\alpha$$
$$= A + i B$$

This is equivalent to the following for a structure factor:



α is the angle that vector makes with real axis (horizontal axis)

A final structure factor is composed of the vector sum from all the component parts (atoms or groups of atoms)

$$F = \sum_j \vec{f}_j \qquad \text{where } \vec{f}_j \text{ is a vector}$$

$$= \sum_j f_j \exp(i \phi_j)$$

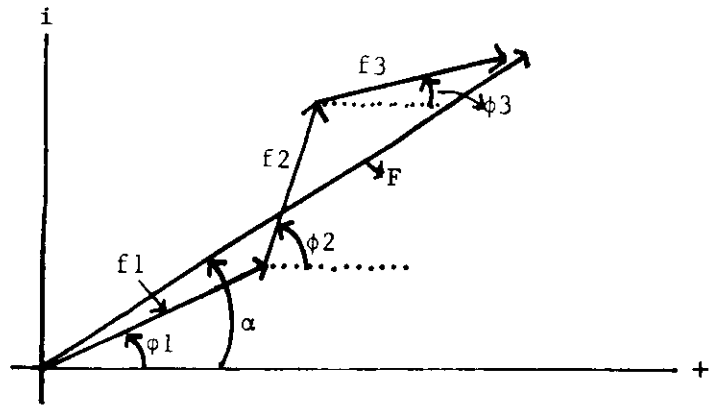
where the ϕ's are the phase angles for each of the individual components.

If we have a series of atoms, then the phase angle component for reflection h k l is given by

$$\phi_j = 2\pi(h x_j + k y_j + l z_j)$$

and $$F = F(h k l) = \sum_j f_j \exp 2\pi i (h x_j + k y_j + l z_j)$$

So that if we had three atoms, the vector sum would look like so:



F can also be broken down into individual real and imaginary components:

$$F = \sum_j f_j \cos 2\pi (h x_j + k y_j + l z_j) + i \sum_j f_j \sin 2\pi (h x_j + k y_j + l z_j)$$

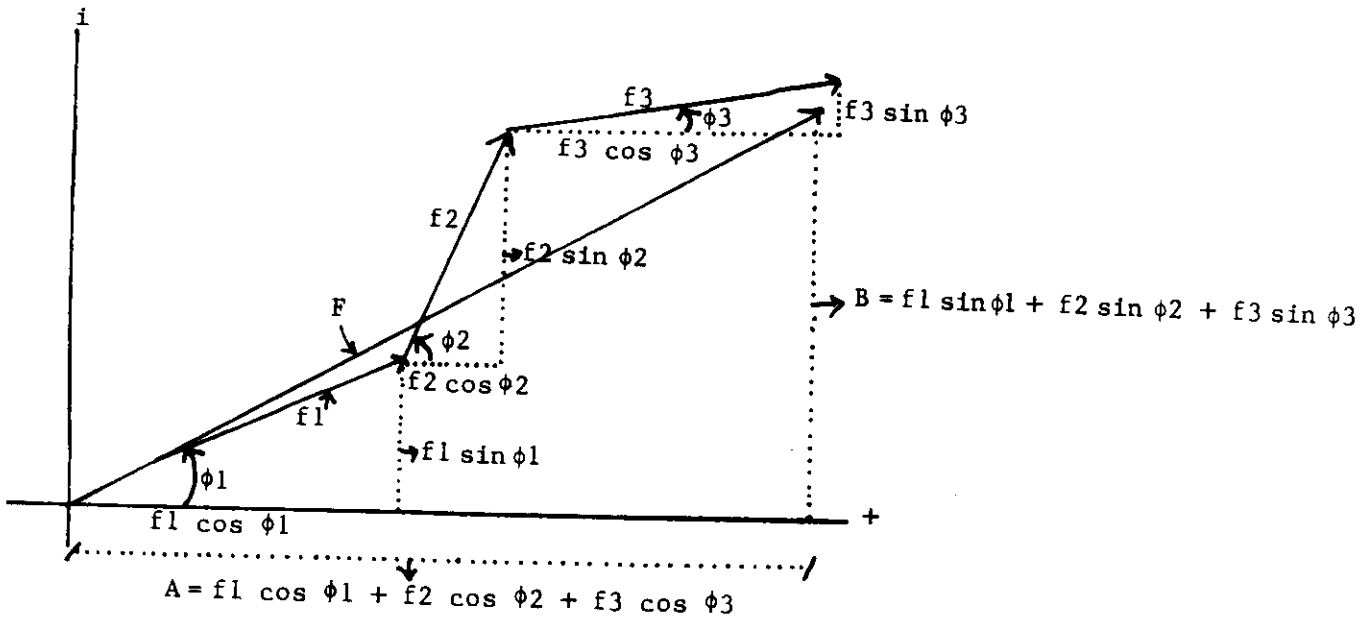
A

B

$j = 1, \#\text{of atoms}$

$F \equiv A + iB$

Again, this can be appreciated in vector notation:



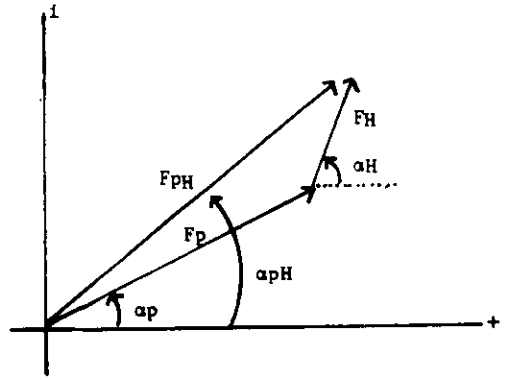
Usually, all of the individual components that go to make Fp and F_{pH} are assumed to be contained in a single vector

$$F_H = \sum_{\text{heavy atoms}} f_H \exp(2\pi i (h x_H + k y_H + k z_H))$$

Sum is over the heavy atom sites

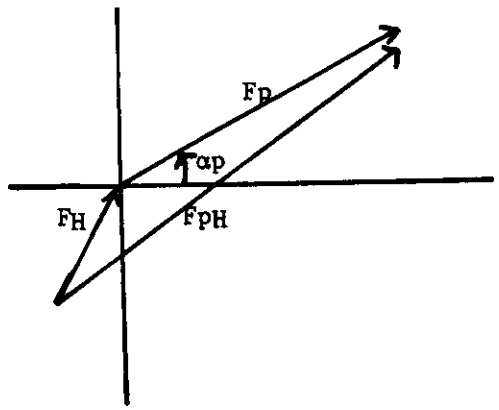
and the relationship would be expressed as

$$\overrightarrow{F_{pH}} = \overrightarrow{F_p} + \overrightarrow{F_H}$$



In fact, usually F_H is very small relative to Fp and F_{pH} and $\alpha_p \approx \alpha_{pH}$

The origin of the complex plane has no meaning (α is really the angle that the vector makes with respect to the real axis); the above situation can also be depicted as follows:



If we knew the phase angle α for the vector Fp, then we could calculate an electron density map using the equation

$$\rho(x\,y\,z) = \frac{1}{V_c} \sum_h \sum_k \sum_l |Fp| e^{i\alpha} \exp(-2\pi i (h x + k y + l z))$$

which can be reduced to

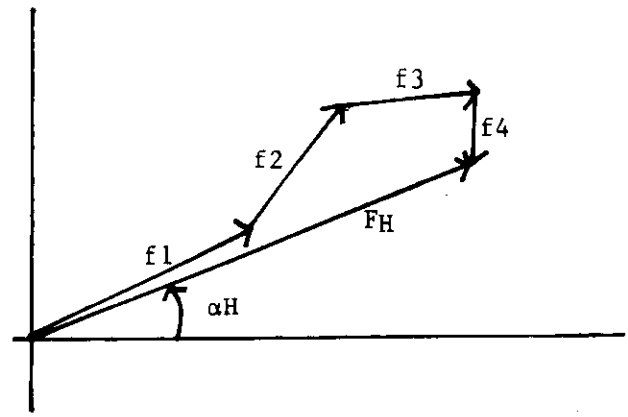
$$\rho(x\,y\,z) = \frac{1}{V_c} \sum_{\text{all reflections}} |Fp| \cos[2\pi (h x + k y + l z) - \alpha]$$

WHERE V_c IS THE VOLUME OF THE UNIT CELL

In order to determine the phase angle α we first have to locate the heavy atoms in one or more derivatives, so that we can calculate the vector F_H (both the magnitude and the phase angle for F_H). How do we locate the heavy atoms?

Obviously, we could find the heavy atoms by classical methods (Patterson, direct methods, *etc.*), if we knew the modulus of F_H, $|F_H|$, since the F_H's represent the Fourier transform of the heavy-atom electron density.

Obviously, FH is composed of individual contributions from all of the heavy atoms in a given derivative, *i.e.*,



If we knew the position of the heavy atoms in the cell, we could simply calculate the modulus and phase angle for F_H via

$$F_H(h\,k\,l) = \sum_{j=1}^{\#\text{ of heavy atoms}} f_i \cos 2\pi (h x_j + k y_j + l z_j) + i \sum f_i \sin 2\pi (h x_j + k y_j + l z_j)$$

$$F_H = A + iB$$

$$F_H^2 = A^2 + B^2$$

$$|F_H| = \sqrt{A^2 + B^2}$$

$$\tan \alpha_H = B / A$$

Patterson Techniques

If one carries out a Fourier summation using the absolute values of the measured intensities in place of the structure factors, the result is called a Patterson synthesis. The value of the Patterson function at a position u, v, w of the unit cell is then given by

$$P(u, v, w) = \frac{1}{V} \sum_h \sum_k \sum_l |F_{hkl}|^2 \cos 2\pi (hu + kv + lw)$$

NOTE: α is now missing from this equation!

If P (u, v, w) is calculated for a number of points in the unit cell, the result is a map which contains information about the possible vectors between atoms in the unit cell.

A large value for P will result when

$u = Xj - Xi$

$v = Yj - Yi$

$w = Zj - Zi$

where j and i correspond to any two atoms.

P will have a large value wherever these conditions are satisfied. Therefore, if there are n atoms in the unit cell, there will be n^2 positions at which P will have large values.

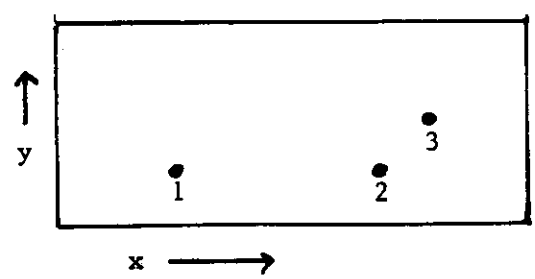
i.e., $j = 1, n$

$i = 1, n$

number of possible values for $Xj-Xj = n^2$

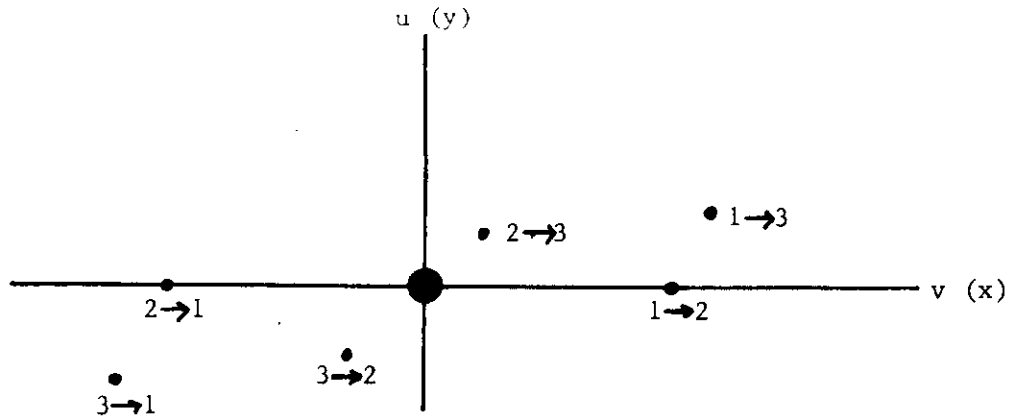
The actual magnitude of P will be proportional to the product of the atomic numbers (*i.e.*, number of electrons) of the atoms involved in the vectors. Thus, P vectors between "heavy atoms" will be especially pronounced in a Patterson map and may be easily distinguished from the lighter peaks.

As an example of the relationship between a Patterson map, take a simple atomic arrangement and the corresponding Patterson:



ATOMIC POSITIONS

The Patterson of this system would consist of nine peaks, three of which (vectors between an atom and itself) would be superimposed at the origin:



For a simple structure, it is sometimes possible to work back from the Patterson map to the original structure.

A more usual application is the use of a Patterson to determine information about the position of a heavy atom in the structure; *e.g.*, if atoms 1 and 3 were heavy vectors between these atoms would be especially pronounced and information about the relative positions of the heavy atoms in the cell could be easily obtained.

[redacted]

The x, y, z coordinates of the heavy atoms are determined using Patterson maps.

If we simply calculated a Patterson map using $|F_{PH}|^2$ as coefficients, the heavy atom vectors would be swamped by the vectors from the many protein atoms.

Consequently, we calculate "difference Patterson" maps in which the coefficients are $|F_{PH} - F_P|^2$.

The resulting coefficient is <u>mainly</u> influenced by the $\vec{F_H}$ component, since we have separated out the Fp term.

What we are actually doing is assuming that

$$|F_H| \approx |F_{PH} - F_P|$$

which is crude, but it usually works!

THE PATTERSON MAP USED FOR LOCATING HEAVY ATOMS IS:

$$P(u,v,w) = \frac{1}{k}\sum_{k}\sum_{l}\sum_{l} |F_{PH} - F_P|^2 \cos 2\pi (hu + kv + lw)$$

<u>Calculating an electron density map for determining the protein structure:</u>

Once we know the x, y, z coordinates of the heavy-atom sites in a derivative, we can begin to calculate the possible phase angles for the protein structure factors.

Given the heavy-atom coordinates from a difference Fourier map, we can actually calculate the magnitude of the $\vec{F_H}$ vector in an Argand diagram:

$$|F_H(hkl)|^2 = \left[\sum_{j=1}^{\#\text{ of heavy atoms}} f_j \cos 2\pi (hx_j + ky_j + lz_j)\right]^2$$

$$+ \left[\sum_{j=1}^{\#\text{ of heavy atoms}} f_j \sin 2\pi (hx_j + ky_j + lz_j)\right]^2$$

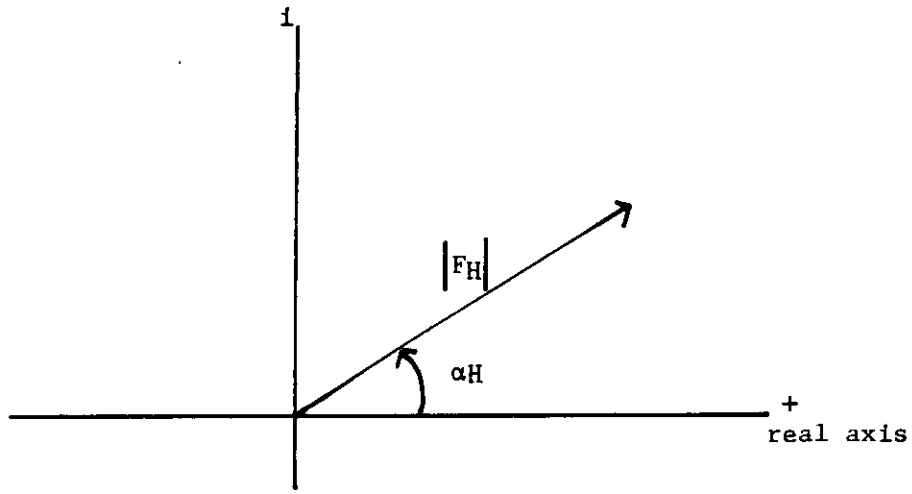
$$|F_H(hkl)| = \sqrt{|F_H(hkl)|^2}$$

and $\alpha H = \tan^{-1}\left(\frac{B}{A}\right)$

where $B = \Sigma f_j \sin 2\pi (h x_j + k y_j + l z_j)$

$A = \Sigma f_j \cos 2\pi (h x_j + k y_j + l z_j)$

then we know



We also know the length of the $\vec{Fp}$ vector and the length of the $\vec{FpH}$ vector, since

$|Fp(h,k,l)| = (Ip(h,k,l))^{1/2}$

where Ip (h, k, l) is the measured diffraction intensity for the h, k, l reflection (after corrections are applied and the Ip value is properly scaled).

Similarly, we know the magnitude of the $\vec{FpH}$ vector from the experimental measurement of the intensity of the h k l reflection from the heavy-atom derivatives.

So, we begin the phase determination knowing

$|FH|$

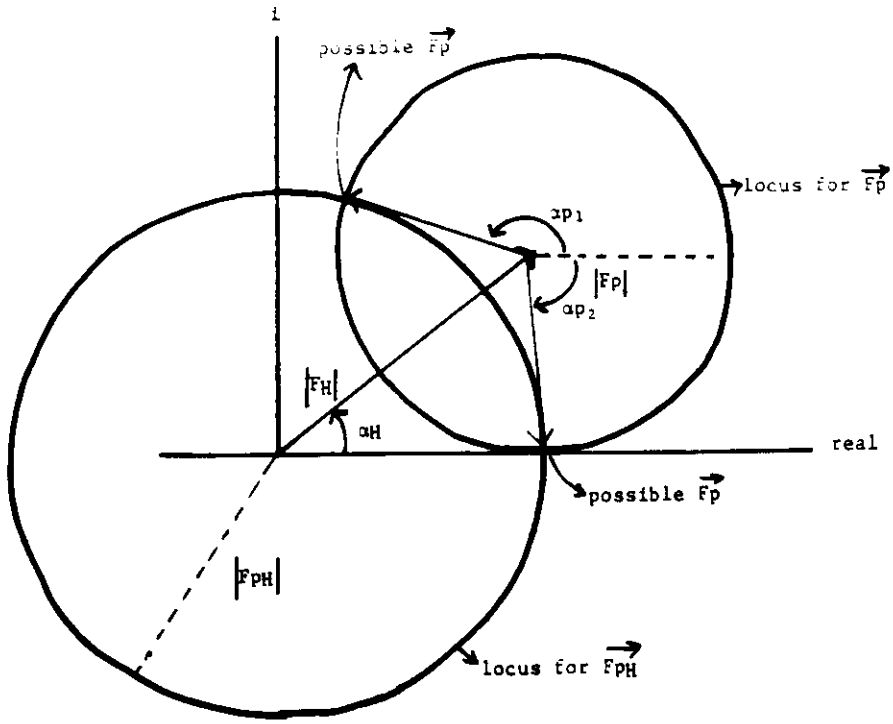
αH

$|Fp|$

$|FpH|$

To determine the phases that are consistent with these four parameters:

1. Place the $\vec{F_H}$ in the Argand diagram:



2. Draw a circle of radius $|F_{PH}|$ around the origin of the Argand diagram.
   (i.e. AROUND THE ORIGIN OF THE $\vec{F_H}$ VECTOR.
3. Draw a circle of radius $|F_P|$ around the end of the $\vec{F_H}$ vector.
4. The Fp and FpH circles will intersect at two points. Vectors from the center of the $|F_P|$ circle (*i.e.*, the head of the $\vec{F_H}$ vector), to the two points of intersection of the Fp and FpH circles define two α_P values that are consistent with the $|F_H|$, α_H, $|F_P|$, and $|F_{PH}|$ values.

To tell which one of these is actually correct, we would prepare a second heavy-atom derivative, and repeat the process. The second derivative would then give new values for $|F_H|$, α_H, and $|F_{PH}|$.

The diagram for the second derivative would also give two possible phase solutions for α_P; only one of these would agree with one of the possible solutions based on the first derivative - this would be the correct value for α_P !

Once α p is determined for all reflections (all h k l values), it is then possible to calculate an electron density map that can be used to determine the structure of the protein. This is accomplished by computing the Fourier summation:

$$\rho(x, y, z) = \frac{1}{Vc} \sum_{h} \sum_{k} \sum_{l} |F h k l| \cos[2\pi(h x + k y + l z) - \alpha(h k l)]$$

where

$\rho(x, y, z)$ is the electron density at position x, y, z in the unit cell

Vc is the volume of the crystallographic unit cell.

$|F h k l|$ is the magnitude of $\vec{Fp}$ for the h, k, l reflection

α (h k l) is the α p value for the h k l reflection

and the summation includes the entire set of h k l reflections.