Lecture – 50
Direct Methods Part - 01

Welcome back to the course of Chemical Crystallography. In the previous lecture, we started understanding the direct methods, different mathematical roots to determine the phases.
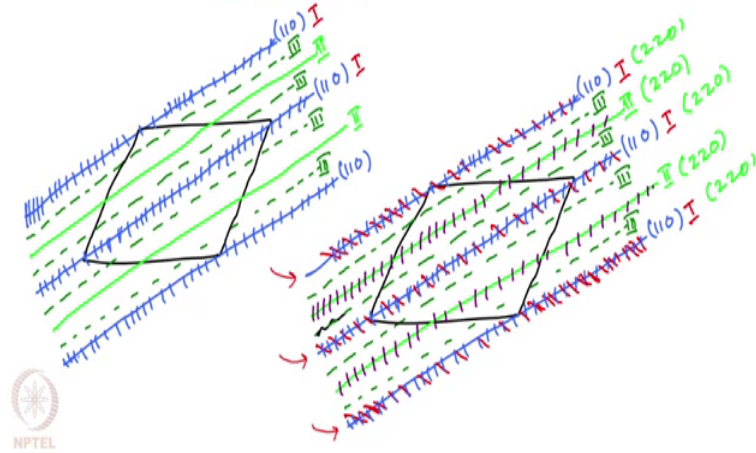
(Refer Slide Time: 00:29)

So, in during that discussion, we just discussed about the expressions for normalized structure factor which is you should remember, we discussed it as we wrote it as, E H equal to 1 by sum over j equal to 1 to N f j square root of that and sum over j equal to 1 to N f j exponential 2 pi i h x j plus k y j plus l z j.

So and then we went on to define the corresponding unitary structure factor, U H you remember that when we write U H it actually means we are talking about a reflection of miller indices h, k and l. So, we wrote that U H is equal to the mod of F H square divided by the sum of j equal to 1 to N f j. And then we understood that the value of U H varies from 0 to 1 and further we could write that E H square is equal to N U H square. So, we could write an expression between the normalized structure factor and the unitary structure factor. So, from that point we went on to explain two different methods.

The sigma 1 relationship which we then used to draw the Bragg-Lipson plots and try to identify the electron density regions, where the electron density is maximum for a given reflection and it is a family that is 1 0 0 and 2 0 0 like for example, if it is h k l, then 2 h 2 k 2 l type of reflection, where should the more intensity region or more electron dense region should appear. So, what we saw that the regions where we have the overlap of maxima regions of the most probability of finding the atoms and electrons close to those regions in the unit cell. And this could lead to a reasonable structure solution. When we do it in 3-dimension and do it for a large number of sets of planes which are parallel and intersecting set of planes.

(Refer Slide Time: 03:39)

Direct Methods

➢ The triplet relationship for large $|E_H|$, $|E_K|$ and $|E_{\bar{H}\bar{K}}|$
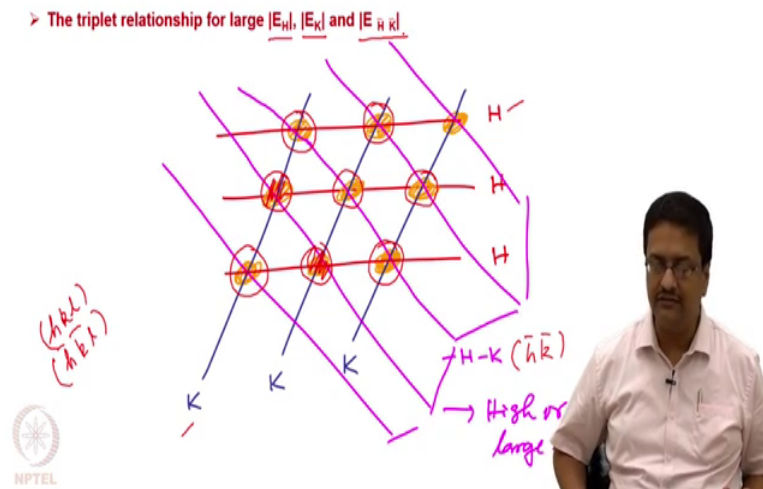
So, then we all went on to talk about the triplet relationship where if you know the that the E H and E K that is the normalized structure factor for two different reflection, two different planes are both high. And we can identified those planes as H and K and they intersect at a given point. So, we understand that the point of intersection of these two planes, which give large values of E would have the concentration of electron density at those regions where these two planes in depth.

So, then the plane which contains these large density points, which joins those large electron density points are actually the plane which is nothing but h bar, k bar in two dimension or in case of h k l; it could be h bar k bar l that plane which then contains the large electron density region. So, from this relationship one could conclude that E H bar K bar will also have the large intensity. So, all these understandings give us an indication that how the large values of E can lead to the under gaze of the actual location of atoms, but it only talks about the closeness or a proximity of those atoms to these planes.

(Refer Slide Time: 05:25)

Direct Methods

➤ Most widely used method for direct determination of phases of reflections through various mathematical reasoning and understanding

➤ Between 1930 and 1965, many attempts were made for direct phasing like the Bragg-Lipson plot

➤ Harker and Kasper[1] introduced Harker-Kasper inequality to relate the phases of individual reflections.

➤ Using the general structure factor expression, one can write

① $-\ F_{hkl}^2 \leq F_{000}^2$, where $|F_{000}|$ is the amplitude of str factor at the R.L. origin. This is nothing but equal to the total no. of e's present in the unit cell.

➤ For centrosymmetric structures, one can further write

② $F_{hkl}^2 \leq F_{000}\left[\frac{1}{2} F_{000} + \frac{1}{2} F_{2h2k2l}\right] \longrightarrow$ ②

[1. D. Harkar & J. S. Kasper, Acta. Cryst., 1, 70, 1948]

The direct methods are a list of methods, a large number of different methods which we will discuss a few of them have been used for a long time. So, this is most widely used method for direct determination of phases of reflections through various mathematical reasoning and understanding. Between 1930 and 1965, many attempts were made for direct phase determination like the Bragg-Lipson and plot and so on. The most important introduction or discovery in this region is by Harker and Kasper, who introduced the Harker-Kasper inequality to relate the phases of individual reflections.

So, the reference here is given for Harker-Kasper inequality. So, in for using Harker-Kasper inequality one can write that F h k l square is less than equal to F 0 0 0 square, where F 0 0 0 is the amplitude of structure factor at the reciprocal lattice origin. This is nothing but equal to the total number of electrons, present in the unit cell. So, we write this equation as equation number 1. So, this expression can be then modified for centrosymmetric structures as the following, which we can write as the equation number 2; F h k l square is less than equal to F triple naught into half of F triple naught plus half of F 2 h 2 k 2 l and this we write as equation number 2.

(Refer Slide Time: 08:25)

Direct Methods

➢ Assuming $u_{hkl} = \dfrac{F_{hkl}}{F_{000}}$ ——③

➢ Dividing equation (2) by F_{000}^2 we can write

$$\dfrac{F_{hkl}^2}{F_{000}^2} \leq \dfrac{1}{2} + \dfrac{F_{2h2k2l} \cdot F_{000}}{F_{000}^2}, \text{ or } u_{hkl}^2 \leq \dfrac{1}{2} + \dfrac{1}{2} u_{2h2k2l} \text{ ——④}$$

➢ If both, magnitude and sign (+ve) of U_{hkl}^2 is known, then the phase of U_{2h2k2l} remains to be determined only

➢ Thus,

$$|U_{hkl}|^2 \leq \dfrac{1}{2} \pm \dfrac{1}{2} |U_{2h2k2l}| \text{ ——⑤}$$

➢ If both $|U_{hkl}|$ and $|U_{2h2k2l}|$ are large then +ve sign on RHS has to be selected to make the inequality valid.

So, now if we assume that small u h k l is equal to F h k l by F is 0 0 0. And then we divide the equation number 2 by F square 0 0 0, one can write that F h k l square by F 0 0 0 square is less than equal to half plus F 2 h 2 k 2 l by F 0 0 0 square into F 0 0 0. So, we can write it as u h k l square is less than equal to half plus half of u 2 h 2 k 2 l and we write it as equation number 4. If both the magnitude and sign the positive sign of U h k l square is unknown is known, then the phase of U 2 h 2 k 2 l remains to be the only determined and only parameter that has to be determined.

So, thus you can write mod of U h k l square is less than equal to half plus minus half u 2 h 2 k 2 l, as equation number 5. So, if both of U h k l and u 2 h 2 k 2 l that is this and that are large, then positive sign on the right hand side is valid to make the inequality work.

Direct Methods

Example of phase determination by inequality

$|u_{hkl}|^2$	$|u_{2h2k2l}|^2$	Phase (+)	Phase (-)	Comment
0.60	0.20	0.60	0.40	u_{2h2k2l} must be +ve
0.50	0.10	0.55	0.45	u_{2h2k2l} must be +ve
0.40	0.10	0.55	0.45	u_{2h2k2l} could be either +ve or -ve
0.40	0.30	0.65	0.35	u_{2h2k2l} must be +ve
0.25	0.50	0.75	0.25	u_{2h2k2l} must be +ve
0.25	0.30	0.65	0.35	u_{2h2k2l} could be either +ve or -ve

➢ To obtain definite result, the reflections must have amplitude that are large fractions of F

$$|U_{hkl}|^2 \leq \frac{1}{2} \pm \frac{1}{2}|U_{2h2k2l}|$$

NPTEL

So, here is one example of this set of values for phase determination by inequality method. Here what we can see is that if I have the values for u h k l square listed in the first column and the calculated u 2 h or observed u 2 h 2 k 2 l square values are listed in second column. Then the phase is if it is the is positive, then you get the value of inequality on the right hand side is 0.6 and the other hand if it is minus, you get it as 0.4 what we should keep in mind that we are talking about the relationship, where u h k l square should be less than equal to half plus minus half of u 2 h 2 k 2 l.

So, we have various values of u h k l and u 2 h 2 k 2 l from the data. And we are trying to find out what should be the phase of u 2 h 2 k 2 l reflection compared to u h k l reflection. So, if these are the values which are listed here for u h k l and if these are the values for the reflection u 2 h 2 k 2 l; then by selecting plus or minus we get different sets of values for the right hand side, these are the values for the right hand side.

So, now we can see that if this is the condition, these are the values then u 2 h 2 k 2 l must be positive, must have a positive phase to have the inequality valid. If it is second one it also must be positive, but in case of third one the phase can be both positive and negative. Similarly, for the next one and that following this should be positive and u 2 h 2 k 2 l could again be a positive or negative for a smaller value of u h k l like this.

So, this indicates that there are indications of possibilities of phases of different reflections related to one given reflection u h k l. So, to obtain a definite result, the

reflections must have amplitudes that are large fractions of F. So, the larger the value of u h k l, it is it becomes easier to assign the correct phase of those reflections.

(Refer Slide Time: 14:01)

Direct Methods

> Unfortunately, reflections with large $|U_{hkl}|$ is less in number and hence this method becomes less useful for complex structures.

> Further $|U_{hkl}|$ is dependent on $\sin\theta$ and hence at high angle, $|U_{hkl}|$ is generally small and cannot be used for phase determination.

> As the value of $|F_{hkl}|$ or $|U_{hkl}|$ reduces with $\sin\theta$, $|F_{hkl}|$ /$|F_{000}|$ falls so low that phase determination becomes impossible.

> To overcome this, we assume that all the scattering power of an atom is concentrated at a point (point atom) and represent F for such assumption as F_{atom}

> This assumption leads to the fact that the scattering power of an atom (f) is independent of $\sin\theta$ and is constant (=Z) for all values of $\sin\theta/\lambda$

NPTEL

But unfortunately, the reflections with large huge U h k l is less in number in general for all data sets. And hence this method becomes less useful for complex structures with large number of atoms. Further, U h k l which is dependent on F h k l is dependent on sin theta and hence at high angle that is at sin theta U h k l is generally small and cannot be used for phase determination purposes; that means, the data that we collect is up to say 56, 50 or 60 degree to theta at U h k l values fall off rapidly at higher angles and then those reflections will not have a significant value of U h k l and we will would not be able to determine the phases correctly

So, as the value of F h k l or U h k l reduces with sin theta, F h k l by F triple naught falls so low that the phase determination becomes impossible. So, to overcome this, what we do is, we assume that all the scattering power of an atom is concentrated at a point, so we consider it as a point atom. And represent the F for such assumption as F atom. And then this assumption leads to the fact that the scattering power of F is independent of sin theta and is a constant for all values of sin theta by lambda.

Direct Methods

➤ Characteristic of point atom model is that the scattering power of an atom (f) is does not fall with sinθ/λ rather is constant (=Z) for all values of sinθ/λ.
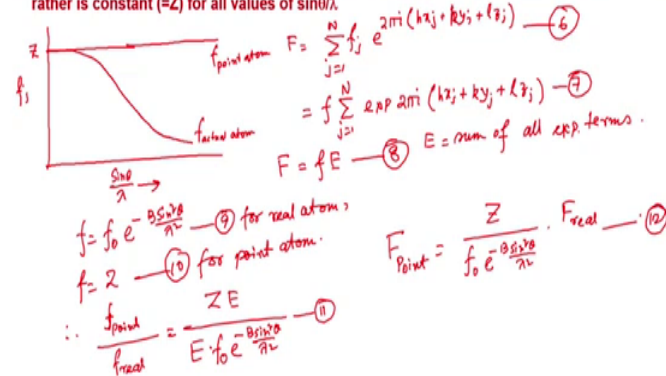
So, with this understanding if we try to plot, the scattering factor f j with sin theta by lambda. In general case f j should fall like this with sin theta by lambda that is for f for actual atom, but by considering the point atom the value of f should remain constant that is f point atom should be equal to z for all values of theta. So, in this case we can write F equal to sum over j equal to 1 to N f j e to the power 2 pi i h x j plus k y j plus l z j as equation number 6.

And then if we assume that the structure is consisting of only one type of atom, then the scattering factor f can be taken out of the summation and we can rewrite it as j equal to 1 to N e to the power 2 pi i h k j plus k y j plus l z j as equation number 7. So, one can then write that F equal to small f multiplied by E, where E is the sum of all exponential term.

So, now as we know that the scattering factor f is equal to f 0 e to the power minus B sin square theta by lambda square, for real atom. We should write f equal to Z for point atom. So, we can then write that f point by f real is nothing but Z E divided by E into f 0 e to the power minus will be sin square theta by lambda square; so this is my equation number 11. So, from here we can write that F point is equal to Z by f 0 e to the power minus B sin square theta by lambda square into F real that is my equation number 12.

Direct Methods

> As most of the structures contain more than one type of atoms,

$$F_{point} = \frac{\sum_{j=1}^{} Z_j}{e^{-\frac{B\sin^2\theta}{\lambda^2}} \sum_{j=1}^{N} f_j} \cdot F_{real} \quad \text{——} \quad (13)$$

> Now, we define unitary structure factor, U_{hkl} as

$$U_{hkl} = \frac{F_{hkl, point}}{F_{000}} \quad \text{——} \quad (14)$$

$$\text{or} \quad U_{hkl} = \frac{F_{hkl, point}}{\sum_{j=1}^{N} f_j} \quad \text{——} \quad (16)$$

$\longrightarrow$ includes the thermal motion (B)

So using eq^n (13) and $\sum_{j=1}^{N} Z_j = f_{000}$, we can write

$$U_{hkl} = \frac{F_{hkl, point}}{\left(e^{-\frac{B\sin^2\theta}{\lambda^2}}\right) \sum_{j=1}^{N} f_{0j}} \quad \text{——} \quad (15)$$

> U ranges from 0 to 1 and U_{hkl} represents a structure factor which has same phase as that of F_{hkl}

As most of the structures contain more than one type of atoms then the same expression has to be changed in this manner, F point is equal to sum over j equal to 1 to N Z j divided by e to the power minus B sin square theta by lambda square into sum over j equal to 1 to N f j f o j into F real. So, we write it as equation number 13.

So, now we will define again the unitary structure factor U h k l as F h k l point divided by F 0 0 0, which is equation number 14. So, using equation number 13 and that j equal to n Z j is equal to F triple naught, we can write U h k l is equal to F h k l point divided by e to the power minus B sin square theta by lambda square into sum over j equal to 1 to N f o j or we can further simplify and write U h k l is equal to F h k l point divided by sum over j equal to 1 to N f i f j, where this f j includes the thermal motion that is B; so we write this equation as 16.

So, as this U actually represents a structure factor h k l. So, this must have the same phase and the value of U should range from 0 to 1, so that is the restriction on values of 0 to 1 and the U h k l would represent the structure factor with structure factor F h k l with the same phase as that of U.

Direct Methods

➤ This indicates that complete in-phase scattering can occur only when the atoms are located at the maxima of Bragg-Lipson chart, in 3D.

➤ Although this will not locate the position of the atoms but will reduce the number of possible sites for the atoms.

➤ Generally |U| close to 1 are less in number but large values of |U| allows us to use more constrains on atom positions.

➤ We know that the information contained in a reflection is determined by the intensity of the reflection relative to the average intensity of the neighboring reflections.

➤ So, it is observed that |U|s are large (close to 1) for reflections with higher value of $\sin\theta$ compared to the |F| for many low angle reflections (low $\sin\theta$)

So, this indicates that the complete in-phase scattering can occur only when the atoms are located at the maxima of Bragg-Lipson chart, in three dimension as well. Although this will not actually locate at the positions that is the coordinates of the atoms, but we will reduce the number of possible positions or sides of those atoms and we will be able to so identify the more electron density regions in the unit cell.

Generally, U close to 1 are very less in number, which we have already indicated in the previous few in slides, but large values of U allows us to use more constraints on atom positions. We know that the information contained in a reflection is determined by the intensity of the reflection, relative to the average intensity of neighboring reflections. So, it is obvious that U s are large close to 1 for reflections with higher value of sin theta compared to F for many low angle reflections with low sin theta. So, this is the possibility.

Direct Methods

➤ To address the question that which reflections are most important we need to arrange the reflections with decreasing value of |U|.

➤ This can be done by summation of U over all atoms for a centrosymmetric structure as

$$U_{hkl} = 2 \sum_{j=1}^{N} n_j \cos 2\pi (hx_j + ky_j + lz_i) \longrightarrow \text{(17)} \quad \text{Valid for centrosymmetric str.}$$

$$\text{Where } n_j = \frac{f_j}{\sum_{j=1}^{N} f_j} \quad \text{(18)}$$

$$\longrightarrow \text{scattering power of the } j\text{th atom.}$$

So, to address this question that which reflections are most important we need to arrange the reflections with decreasing value of module that is the absolute value of U. This can be done by summation of U over all the atoms. So, one can do this sum as U h k l is equal to 2 times j equal to 1 to N n j cos 2 pi h x j plus k y j plus l z j, this is equation number 17; this is valid for centrosymmetric structures. You can have only one cos term that is valid for centrosymmetric structures only. Where we can also write, n j is equal to f j divided by sum over j equal to 1 to N f j. So, this is the scattering power of the jth atom.

Direct Methods

➤ As n_j is a function of $\sin\theta$, if we assume all atoms are same (which is not inappropriate for light atoms structure having C, N, O and H only), we can write

$$n_j = \frac{1}{N} \quad \text{(19)}$$

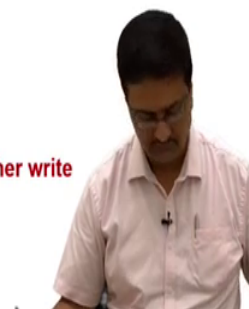$$\text{Similarly, } \langle F^2 \rangle = \sum_{j=1}^{N} f_j^2 \quad \text{(20)}$$

➤ The expectation (average) value of U² is

$$\therefore \langle U^2 \rangle = \sum_{j=1}^{N} n_j^2 \quad \text{(21)}$$

$$\text{The RMS value of } U, \ U_{RMS} = \sqrt{\sum_{j=1}^{N} n_j^2} \quad \text{(22)}$$

➤ For light atom structures with similar Z, we can further write

$$U_{rms} \approx \frac{1}{\sqrt{N}} \quad \text{(24)}$$

As n j is a function of sin theta, we assume all atoms are same which is not inappropriate for light atoms structures having only carbon, nitrogen, oxygen and hydrogen's in the structure. So, we can further write that n j is equal to 1 by N. Similarly, the expectation value of F square is equal to sum over j equal to 1 to N f j square. So, the expectation value of U square is given as sum over j equal to 1 to N n j square. So, the RMS value of U that is U RMS is nothing but equal to sum over j equal to 1 to N n j square. So, for light atom structures with similar Z, we can further write that U RMS is nearly equal to 1 by square root of N.

(Refer Slide Time: 28:55)

Direct Methods

> As the average value of |U| drops well below the value needed for using the inequality appropriately for structure determination with large N

> Therefore, we need better methods phase determination based on certainties than probabilities

> Karle & Hauptman (1956) introduced the concept of normalized structure factor, E_{hkl} as

$$E_{hkl}^2 = \frac{U_{hkl}^2}{\langle U^2 \rangle} \quad — \quad (24)$$

> The E values allow he normalization of all classes of reflections to a common basis

So, as the average value of U drops well below the value needed for using the inequality to work for structure determination with large N. Therefore, we need better methods to determine the phases based on certainties and then probabilities. To overcome that Karle and Hauptman in 1956 introduced the concept of normalized structure factored where E h k l was represented as, E h k l square is equal to U h k l square by expectation value of U square; this equation is number 24. So, the E values are now allow the normalization of all classes of reflections to a common basis.

Direct Methods

> Combining the equation 24 and 21, we can write

$$E_{hkl}^2 = \frac{U_{hkl}}{\varepsilon \sum\limits_{j=1}^{N} n_j^2} \quad —— (25)$$

$$\text{and} \quad E_{hkl}^2 = \frac{|F_{hkl}|^2}{\varepsilon \sum\limits_{j=1}^{N} f_j^2} \quad —— (26)$$

ε = an integer, generally 1 but may be different for diff. space groups.

$P2_1/c$ space group. $\varepsilon = 2$ for $(h0l)$ & $(0k0)$
$\varepsilon = 1$ for all other (hkl).

So, when we combine these two equations 24 and 21, we can write that E h k l square is equal to U h k l square divided by epsilon sum over j equal to 1 to N n j square. And E h k l square as equal to mod of F h k l square divided by epsilon sum over j equal to 1 to N f j square, where this epsilon is equal to an integer which is generally 1, but may be different for different space groups. So, for example, in case of P 2 1 by C space group, epsilon is equal to 2 for the h o l set of reflections and 0 k 0 set of reflections and epsilon is equal to 1 for all other h k l.