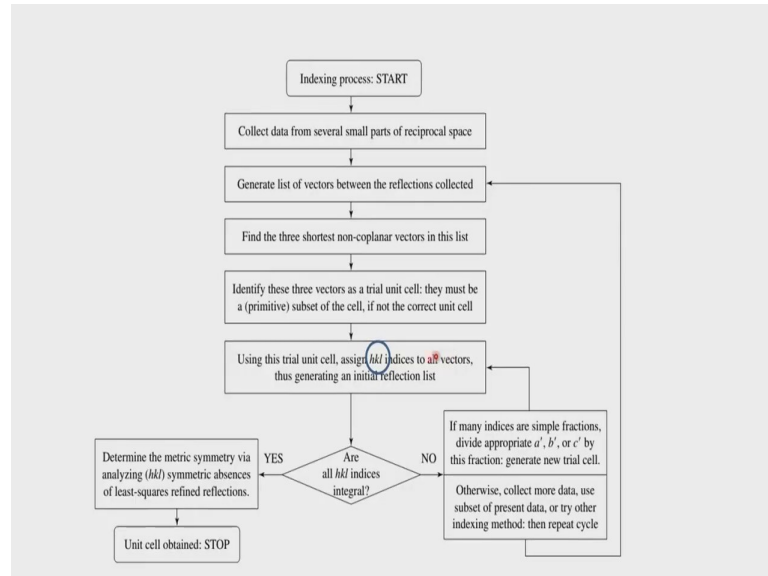


Symmetry and Structure in the Solid State
Prof. T. N. Guru Row
Department of Chemistry and Biochemistry
Indian Institute of Science, Bangalore

Lecture – 49
Solid State and Structural Chemistry Unit

(Refer Slide Time: 00:31)



So, after looking at this indexing process and identifying the unit cell dimensions, the next job will be to calculate and improve all the a, b, c, and α , β , γ and so on.

(Refer Slide Time: 00:47)

| Crystal System | $1/d_{hkl}^2$ |
|------------------------|--|
| Triclinic | $(1 - \cos^2 \alpha - \cos^2 \beta - \cos^2 \gamma + 2 \cos \alpha \cos \beta \cos \gamma)^{-1} \left(\frac{h^2}{a^2} \sin^2 \alpha + \frac{k^2}{b^2} \sin^2 \beta + \frac{l^2}{c^2} \sin^2 \gamma + \frac{2hk}{bc} (\cos \beta \cos \gamma - \cos \alpha) + \frac{2hl}{ca} (\cos \gamma \cos \alpha - \cos \beta) + \frac{2hk}{ab} (\cos \alpha \cos \beta - \cos \gamma) \right)$ |
| Monoclinic | $\frac{h^2}{a^2 \sin^2 \beta} + \frac{k^2}{b^2} + \frac{l^2}{c^2 \sin^2 \beta} - \frac{2hl \cos \beta}{ac \sin^2 \beta}$ |
| Trigonal (R) | $\frac{1}{a^2} \left(\frac{(h^2 + k^2 + l^2) \sin^2 \alpha + 2(hk + hl + kl)(\cos^2 \alpha - \cos \alpha)}{1 + 2 \cos^3 \alpha - 3 \cos^2 \alpha} \right)$ |
| Hexagonal/Trigonal (P) | $\frac{4}{3a^2} (h^2 + k^2 + hk) + \frac{l^2}{c^2}$ |
| Orthorhombic | $\frac{h^2}{a^2} + \frac{k^2}{b^2} + \frac{l^2}{c^2}$ |
| Tetragonal | $\frac{h^2 + k^2}{a^2} + \frac{l^2}{c^2}$ |
| Cubic | $(h^2 + k^2 + l^2)/a^2$ |

$a=b=c$
 $\alpha=\beta=\gamma=90^\circ$

So, we can do that very comfortably by looking at this particular expression. So for example, you see that the 7 crystal systems are listed here. The value of $1/d_{hkl}^2$ as we saw in the earlier classes also depends upon the nature of the crystal. Suppose it is a cubic crystal, the $1/d_{hkl}^2$ will be $(h^2+k^2+l^2)/a^2$ because in a cubic system what we have is $a=b=c$, α, β, γ is 90 degrees. This is in a cubic system.

So, in a cubic system since we have these values the way in which the d_{hkl}^2 comes up, in this system is $1/d_{hkl}^2$ is $(h^2+k^2+l^2)/a^2$. So, we have done this identification, we have calculated the hkl values and the $h^2 k^2 l^2$ square can be calculated and then, we know where the d values are coming with respect to the reciprocal lattice. So therefore, we can calculate the cell dimension here. The same way in a tetragonal system, we can calculate the cell dimension a and c, these are straightforward equations and orthorhombic we can calculate a b and c.

So, because the $1/d_{hkl}^2$ defines this equal to that; in the case of an orthorhombic system. So, as you see here trigonal and then rhombohedral setting of the trigonal system, the monoclinic system where the β angle becomes important. We have discussed this earlier on when we discussed the reciprocal lattice and it's relationship with direct space. So, we can use these expressions therefore; to kind identify the d values.

So, in principle we have the d values of every spot in the data collected identified with respect to the cell dimensions a, b, c the three non co planar vectors which we were talking about. And so, what we do in principle, is what the program does in principle is to see first take those reflections which you have indexed as hkl. Take the three non-co planar vectors which you have identified and then what it does is, it starts from fixing the cubic lattice to the system and then see whether hkl remains integrals. If the hkl values are not integral as we discussed in the previous viewgraph or the previous flowchart, then it will see you all it is not cubic let me try tetragonal.

So, like that it will try the systems to bring down the symmetry and you see triclinic one is fairly complicated because α, β, γ are not equal to 90 degrees. So all these expressions it will fit in and see which is the most probable fail and then feed it back to the flow chart we saw in the previous case. So, we can keep on doing the refinement associated with the value of d. So, since we have a very large number of reflections identified the $1/d_{hkl}^2$

is very very large in number with respect to the measurements which we have done and so, these values of a, b, c will become extremely accurate.

The accuracy associated with the cell dimensions could be at the third decimal place or sometimes even the fourth decimal place in a single crystal. In fact, its accuracy becomes much more when we do powder diffraction, mainly because powder diffraction has a large number of very small what we call as crystallites each one is a small crystal. Since we have a random collection of all these crystals, there is an averaging out effect on determining the cell dimensions. So, we can even go to the fifth decimal place.

So normally, for example in case of a cubic system if the a value is 6.123 let us say, we can also find out from this over fitting of all the d values we can find out how far is the deviation of this determination. So, it can be like 6.123 which are standard deviation of 2; that means, the standard deviation is 0.002 so; that means, within the error bar of 0.002 the determined cell dimension is valid.

So we can therefore, get very accurate cell dimensions and powder diffraction gives more accurate cell dimensions than single crystal data. This is something which is to be noted.

(Refer Slide Time: 05:17)

Determination of the Space Group

- Cell Dimensions, Laue Symmetry, Intensity Statistics (Centro/Non-Centro), **Systematic Absences**
- **Point symmetry** of unit cell determines **symmetry of diffraction pattern**
- **Translational symmetry** elements result in **systematically absent Bragg spots**. [For example **Centering, screws, glides**]

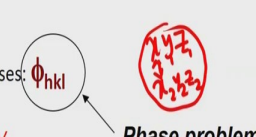
Structure Determination

- Experiment $\rightarrow I_{hkl} \rightarrow |F_{hkl}| = \text{Sqrt}(I_{hkl})$
- Needed for 3D structure (approximate) Phases: ϕ_{hkl}

$|F_{hkl}| + \phi_{hkl} = F_{hkl} \rightarrow$ 3D-Fourier Synthesis

$\rho(x,y,z) = [\sum_{hkl} F_{hkl} \exp\{-2\pi i(hx + ky + lz)\}] / V$

x,y,z are fractional coordinates (range 0 \rightarrow 1)



Phase problem

So, having got this up to this particular point we now go to the description of how we determine the structure? How do we use the symmetry information which we already

have learnt so far and how the structure determination protocol per say goes. I will take you through this whole series of determination of the structure in a few slides rather rapidly. However, we will be visiting each one of them separately with respect to the later date the next few lectures.

So, the first step is one once you have the data, we have determined the a , b , c , α , β , γ you know the crystal system, you now go and look at the systematic absences. The systematic absences will tell us the space group information. So we have the Cell Dimensions and therefore, we also have the information on the Laue Symmetry and then of course, we now have the intensities collected so we can do the Intensity Statistics. In fact, the intensity statistics will tell us whether the crystal system belongs to a centro symmetric system or a non-centro symmetric system. Please make a note of this at this moment, we will discuss it later on probably in the next class or something.

So, the intensity statistics will be deciding the Centro Symmetric or Non-Centro symmetric nature of the given material. Systematic absences of course, will determine the space group, but in case the space group does not have any systematic absences like $P1$ and $P-1$. How do we distinguish between $P1$ and $P-1$? We distinguish between $P1$ and $P-1$ by looking at what we call as the intensity statistics. How the intensities are distributed with respect to $(\sin\theta)/\lambda$.

So, we have an average intensity over the experiment. So, we calculate I/I_{average} and see how the $(\sin\theta)/\lambda$ varies with respect to I/I_{average} and that will give us a distribution which can be plotted in the form of a curve and if that particular curves are two types. One is the exponential type the other is the error function type. The error function type tells us it is the centro symmetric system if it is an exponential type it tells us it is a non-centric system. We will see that in more detail when we go for the discussion of scaling and the intensity statistics which we will evaluate in the next few minutes.

So, the point group of the unit cell determines the symmetry of diffraction pattern. So, that is, the Laue symmetry that will determine the symmetry of the diffraction pattern. Then we translational symmetry elements result in systematically absent Bragg spots. For example, this is centering, screws and glides which we discussed in the previous class. So, we have all these information that is available. So, once you have collected the data, once you have taken that particular data onto your system determine the a , b , c , α ,

β , γ by means of which every spot which was identified with respect to four angles: 2θ , ϕ , ψ and ω is converted to hkl .

So we have therefore, the information about the indices and when once we have the indices we have the information on the crystal system. We have the information on the nature of space based on the systematic absences we also have information on the space group. So, this is the next step which we covered up to the previous class. So, we have therefore, the uniquely indexed pattern at our disposal from the experiment. We have determined the space group to the extent reasonably reliable and in most cases systematic absences as we saw will give us a unique space group information having gotten all that we now go and go for the so called structure determination.

Here is a some issues which we have to discuss, but if before we go into the detail of it I will just tell you the protocol that is followed since we are now doing, we are in an experimental mode, so to say. Because we have talked about how what kind of crystals we choose, how we mount and how we collect the data on the diffractometer, how do we look at all those data then determine a , b , c and d of hkl and then of course, systematic absence is giving us the space group.

So, I thought we will finish off with the protocol of a structure determination. Now any structure determination is done in two steps. The reason is that we have to determine the phases and this is the so called phase problem in crystallography. So, experiment. we got the intensities, the square root of the intensities gave us the modulus of F_{hkl} which is coming from the square root of I_{hkl} . So, if we want to calculate now the 3 dimensional structure by using the Fourier Synthesis, which we will discuss in the coming class. The modulus of F_{hkl} has to be combined with the phase information. To get the F_{hkl} and then the Fourier transformation takes it from the f of hkl information which is up here to the ρ of x , y , z which is the electron density distribution.

So, we plot now the 3 dimensional x , y , z coordinate system. The ρ value, the electron density value which is calculated from these expression. You are quite familiar with this expression F_{hkl} exponential $\{2\pi i (hx + ky + lz)\}$. It now depends upon n atoms. So, all every atom will have this formulation divided by V the volume of the unit cell. So, here xyz of course, are the fractional coordinates it can range between 0 and 1. The idea of this part of the slide which is shown here is to bring in the importance of the

determination of the phase problem, a solution to the phase problem. How do we get to the phases because we cannot get them from experiment, we cannot calculate them directly and obviously, we have to make estimates of the phases.

So, the estimates of the phases is done by several methods. One is the direct space method which is also referred to as the direct methods and the other is to go through the vector space which we discussed earlier on. That is, take the intensity and directly do a Fourier transform and that Fourier transform in turn will give us the information on the inter atomic vectors that is to say let me draw it a little bit.

Suppose you have an atom x, y, z here, you have an atom at x_2, y_2, z_2 , you have an atom at x_1, y_1, z_1 . Then what you will get in a Patterson map is the vector between these two. The problem with the Patterson map is the origin will be the starting point for these vector. It will also give the vector between x_2, y_2, z_2 and x_1, y_1, z_1 .

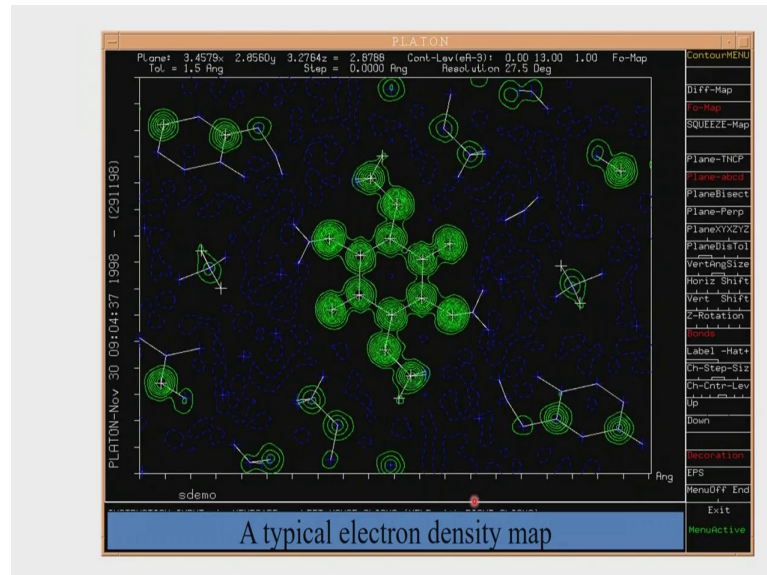
So, both these vectors appear with respect to the origin. Both these atoms will also appear at the origin because the self vectors will also be depicted in the Patterson. So however, in principle we should be able to find where the atoms are and that is the so called Patterson method. We will be discussing detailed analysis of the Patterson approach particularly with respect to this symmetry and structure. You see the whole idea of this course is symmetry in structure. So, we will bring in the importance of symmetry in the vector space which will allow us to determine the structure.

So, we will discuss Patterson synthesis in that context. Of course, the direct phase structural determinate or structure determination using the direct phase solution problem is something which is based on probabilities and we will develop it very little but not in full detail, but give us the basics of the physical properties that are involved in direct methods approach. So, the physics of it we will try and understand. Again, in terms of the planes and in terms of the electron density contained in the planes and so on.

So, that way the phase problem solution we will not spend too much time; however, we will find the methods to solve the phase problem. Once we have the phase we will try to get an understanding of what is this Fourier synthesis is all about and how do we get the Fourier transform done from F_{hkl} to ρ_{xyz} . We already have seen that if we have the information on the ρ_{xyz} which means the we know the positions of the atoms, we can go and calculate F_{hkl} .

But what we get from the experiment now is the modulus of F_{hkl} phase solved we have the F_{hkl} values, this is the expression for the Fourier transformation. So, we go from F_{hkl} to ρ_{xyz} . From F_{hkl} to ρ_{xyz} using the Fourier transform.

(Refer Slide Time: 14:31)



Let us go further and see how this all happens. This is a picture which is a typical electron density map which comes out after all these process is done; that means, when you have done the Fourier transformation you get a picture like this. This is a sort of very nicely featured picture you may not always get this kind of a picture, but for discussions sake, I have taken a picture like this.

Now, this is the electron density map. So, what we have done is, we have plotted the electron density map some of the details are given here. This is an output from the program. So, we will see that we will get electron densities drawn in this particular fashion. Now what it is what is it that is represented in this is the position of position x y z of each one of these atoms up here, and then there is a contour which is drawn at a certain specific intervals and those contours which are drawn at specific intervals have some meaning; that means, they carry that amount of electron density.

So; that means, when you calculate the ρ_{xyz} you have a map you calculate the value of 0, 0.1, 0.1, 0.2, 0.2, 0.2, 0.3 and so on, covering the entire space. We normally cover the symmetric unit of the unit cell and in this particular diagram of course, the isometric unit is combined with the rest of the other symmetric unit in the unit cell so that we are

This diagram is drawn from the structure of this particular compound whatever that is, we have a ring system here, we have the substitutions, this corresponds to the diagram here and these are the values which will appear and we see that based on the strength of the electron density at these positions, we get the positions of the oxygen atoms. The rest of the atoms are carbon atoms.

(Refer Slide Time: 18:03)

Structure Completion

- Extract the 3D Coordinates (x, y, z) of the atoms.
- Assign Atom Types (Scattering type C, O etc.)
- Assign Additional Parameters to Model the *Thermal Motion* (T) of the Atoms.
- Other Parameters: Extinction, Twinning, Flack x
- Model: $F_{hkl} = \sum_{j=1, n} f_j T \exp\{2\pi i(hx + ky + lz)\}$
- Non-linear Least-squares Parameter Refinement until Convergence.
- Minimize: $\sum_{hkl} w [(F_{hkl}^{obs})^2 - (F_{hkl}^{calc})^2]^2$
- Agreement Factor: $R = \sum |F^{obs} - F^{calc}| / \sum |F^{obs}|$

$F(hkl) \leftarrow \sum_{j=1}^n f_j T$

This is the beginning of the structure determination.

That is why I said the structure is determined in two steps. The first step covers basically the solving the phases. So, the phases which we have determined or approximate and therefore, the structure is also going to be approximate. In other words if we go back here, the distance between these two may not correspond to the exact distance between carbon and carbon. Some of them may be elongated, some of them may be short, some of them may be long because there are errors that are associated with the phase.

So, the structure therefore, is first the initial estimate of the phases. You will get what this is known as a starting model. So, we get a starting model and that starting model has to be refined. In order to refine the starting model, we need to have a large amount of data. The fact that we have a large number of hkl values and less number of parameters to determine will allow us to do the so called least squares refinement.

I do not know whether it is within the purview of this course, but I will give some basic details of the least squares refinement. So that we know how the structure which we determine initially which is called which is giving us the trial structure. We will develop into a final structure of this kind where the atoms positions are known, the distance between atoms are accurate, the angle between atoms are accurate and this is a refined structural model of that same compound.

So the way the way in which we have to do is refer to as the structure completion. So, this is the second part of the structure determination. First part is to get this trial structure by solving the phase problem and then using the Fourier transform on the F_{hkl} to get the initial ρ_{xyz} . You can call it the initial electron density map. Now after that we have to do this protocol of structure completion. So, this lists out the protocol of structure completion. So, extract the 3D coordinates of the atoms as we saw just now and assign atom types, scattering type C, O etc. Assign additional parameters to model the thermal motion of the atoms.

Now, here again a little digression because the as we always know that there is these atoms in a structure or not at rest, they are vibrating about their mean positions. The reason for vibration is purely thermodynamical and the reason is also accounted by the fact that these atom may not be associated with that x, y, z which you have determined. So, there is always a, there is always a area around the x, y, z which is possible for this particular atom to occupy and this error in this $\Delta x, \Delta y, \Delta z$ could be in all three directions.

And therefore, these thermal motion is expressed in terms of what we call as a thermal ellipsoid. We will be analyzing this in a short while from now by one once we do the Wilson plot which is essentially the plot which puts the relative measurements we got from the experimental measurement onto a final scale value which is equal to the total number of electrons inside the scattering material we are looking for. And this one once we have this the thermal motion of the atoms will get additional parameters along with the x, y, z .

So, if there is a carbon atom it is associated with the coordinates x, y, z . It is also associated with the thermal portion parameters. Since the value of a, b and c are different from each other in many crystal systems and also the α, β, γ , depending upon the symmetry information the thermal ellipsoids, the ellipsoids that is associated with the

position x , y , z can have different shapes and the values of these major and minor axis may be different. The orientation of the ellipsoid may be different.

So, all these now come as additional parameters. So, we have we will have six additional parameters which are referred to as an isotropic thermal parameters. Then we have other parameters which will come along with the data. It is due to the nature of the crystal and also due to the nature of the material. The extinction and the twinning is essentially associated with the crystal. When we have twinning associated with the crystal, the crystal diffraction will be affected by twinning and there is a way to treat it which we are not going into the detail of it.

The flack parameter which is called the x parameter will allow us to determine the absolute configuration particularly in case when we have the anomalous scattering effects. So, every one of the atoms in the structure will be anomalously affected, but if you have a very heavy atom like bromine or something in an organic molecule, the affect associated on the variation of I_{hkl} and I_{-h-k-l} which we discussed in the last class. The difference between them, the Friedel's law will be violated. The Bijvoet pairs will be created and those Bijvoet pairs will allow us to determine this parameter. The parameter which will now, tell us about the absolute configuration of the system. We will go into the detail with an example in the later classes.

So essentially, we build up a model. The model is based on F_{hkl} it depends upon the scale factor the f_j value, the thermal parameters and so on, this should be under the summation j equals 1 to n . S is to replace the summation. This is the Bill Gates effect on my computer because on my computer this comes as sum here it comes as S . So, I should have checked it before, but anyway this is summation over j equals 1 to n , f_j is the thermal parameter exponential $2\pi (hx + ky + lz)$. Surprisingly, the 2π has been retained somehow.

The non-linear least squares parameter refinement until convergence, this is the next step where what we do is you see we have a trial structure. The trial structure will give us a ρ_{xyz} so let me write it down. So, we have now arrived at a trial structure and this trial structure which we have is based upon solution to the phase problem. So, we solved the phase problem we got a trial structure, this is the starting model.

Now, the starting model values we can use to calculate rho of x, y z. So, we have any rho of x, y, z which is the electron density. Now using the electron density you go back and calculate. This is now the calculated F_{hkl} . So, the calculated F_{hkl} will have errors because we do not have the accurate determination of the phases. So, the F_{hkl}^{calc} is now put back here in this expression, you have to see hkl which is coming from the trial structure is put here which is, F_{hkl} calculator and we already have the F_{hkl} observed, which is the experimental F_{hkl} .

So, F_{hkl} observed minus the F_{hkl} calculated whole square. This is a quantity which is which can be called as ΔF , that is the difference between the actual experimental observation and the calculated value which has come from a trial structure calculation. So, this difference therefore, has to be minimized. So, we have to minimize the summation $\sum_{hkl} w$ where w is a weighting scheme we will not discuss it further. There is some kind of a weight it has to be attached for each one of these reflections because we do not measure every reflection under the same condition.

So, that will attach a factor called w , we will not go into the detail of it; but sum over hkl you have F_{hkl} observed minus F_{hkl} calculated. This value has to be minimized from the least square sense. So, this is the non-linear least squares refinement protocol we will discuss it a little more. This is just to finish the protocol we have to follow for one once we have put a crystal collected the data and so on.

So now, you have an agreement factor. Now this agreement factor is referred to as the R factor. Now this R factor is equal to the sum over this should be $\sum_{obs} |F_{obs} - F_{calc}|$ divided by $\sum_{obs} F_{obs}$. I will correct it in the possibly not now, but I will correct it later because I don't have the necessary software here. So, sum over F_{obs} minus F_{calc} that is the delta F divided by sum over F_{obs} . So this gives us agreement factor or what is called a reliability index.

Now, this reliability index is expressed with respect to 100. So, if the reliability value is somewhere around 10 or something, you say that your structure is fully determined. There are other issues which will come up with the least squares refinement which we will be discussing when we go in detail with the non-linear least squares refinement protocol. So at this moment, what we have to do is one once we have the solved the what we call as the phase problem we get a trial structure. The trial structure is now associated with a thermal motion for each and every atom and other parameters which are

extinction, twinning, flack etcetera are covered along with it, Then we go and develop a model and that particular model will give us the trial structure with an electron density ρ_{xyz} .

Now, based on this ρ_{xyz} we calculate the structure factor again and subtract the structure factor from the observed structure factor because that will tell us what is the difference between the two which for which we have to account for and that particular difference has to be minimized; and that minimization process is using what we call as a non-linear least squares refinement protocol and we will discuss that a little more detail in the coming classes right.

(Refer Slide Time: 28:44)

Refinement of the Structural Model

Refinement Steps (Programs SHELXL, Crystals, XTAL etc)

1. Refine positional parameters + isotropic U
2. Refine positional + anisotropic parameters
3. Introduce H-atoms
4. Refine H-atoms with x,y,z,U(iso) or riding on their carrier atoms
5. Refine weighting scheme
6. ORTEP presentation →

So, at this stage how do we go about this refinement there are automatic packages. Since we are going through the protocol of structure determination, I thought I will take you through this particular slide as well. So, what you do is, you use programs like SHELXL, Crystals, XTAL. So many programs are available. So, use all these programs and what you do is you improve on the positional parameters by doing the F_{obs} minus F_{calc} and trying to minimize the value of F_{obs} minus F_{calc} along with the isotropic thermal and later it could be converted to anisotropic thermal parameters as well.

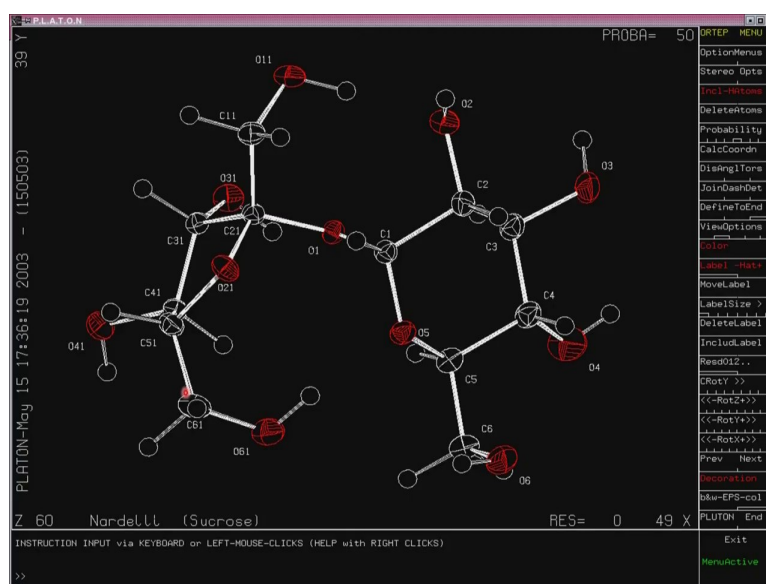
Now having done this, when once we reach a reliability index of a decent amount and then that does not change much with changing the F_{obs} minus F_{calc} values, you now introduce the hydrogen atoms. Now the issue of hydrogen atom is very interesting.

Hydrogen atoms do not have a core. Since hydrogen atoms do not have a core, we have an issue. The issue is whether the $\sin F$ will varying a $(\sin\theta)/\lambda$ can be existing at all for the hydrogen atom. Even if it exists, it will be very very small.

So, when we actually determine the structure in the first instance using the phase refined a protocol and then eventually trying to get to the starting model, the starting model will cannot give up hydrogen atoms. So, we have to now introduce the hydrogen atoms there are various ways in which we can introduce the hydrogen atom. The detail of which I will not probably go into in full swing, except to tell that hydrogen atoms can be introduced because we can fix the hydrogen atom stereo chemically on the associated atom or we can refine the hydrogen atoms with the x,y,z , U_{iso} or riding on their carrier atoms and so on.

So, this particular detail I will not tell you right now, but when we actually discuss the refinement and the interaction of hydrogen atoms, we will bring in these details later. After all these, we do what is known as an ORTEP presentation. This is the OakRidge Thermal Ellipsoid Plot. This will tell us how the atoms are oriented with respect to each other in the crystal structure along with their thermal features which are now the features that have to be discussed.

(Refer Slide Time: 31:14)



Let me show the picture which will be clear. So here is a structure when in which we have shown the thermal ellipsoid. So, each and every atom now is associated with a

thermal ellipsoid. Hydrogens are shown as circles because they are associated with a isotropic thermal ellipsoid in this particular structure which you have taken as an example. So, these therefore, tells us that the oxygen atom location O5 is with respect to some x, y, z here and then the uncertainty in locating that particular atom because the atom is vibrating at any given time. When we do the x-ray diffraction experiment, it may not be at that point, but it may be distributed in any place, any space 3 Dimensional space within that particular ellipsoid.

So, this tells us the limit to which we can rely on the determination of the position of the atom and therefore, this is the thermal ellipsoid. We can calculate the thermal ellipsoid for each and every atom and this therefore, now represents the structure of the molecule as such, as it at that particular temperature. So if you reduce the temperature, the size of the thermal ellipsoids will reduce. If you increase the temperature the size of the thermal ellipsoids will increase. So the atomic vibration therefore, goes hand in hand with temperature. So it is therefore, suggested that all the structures have to be determined at the temperatures which are lower than the ambient temperature.