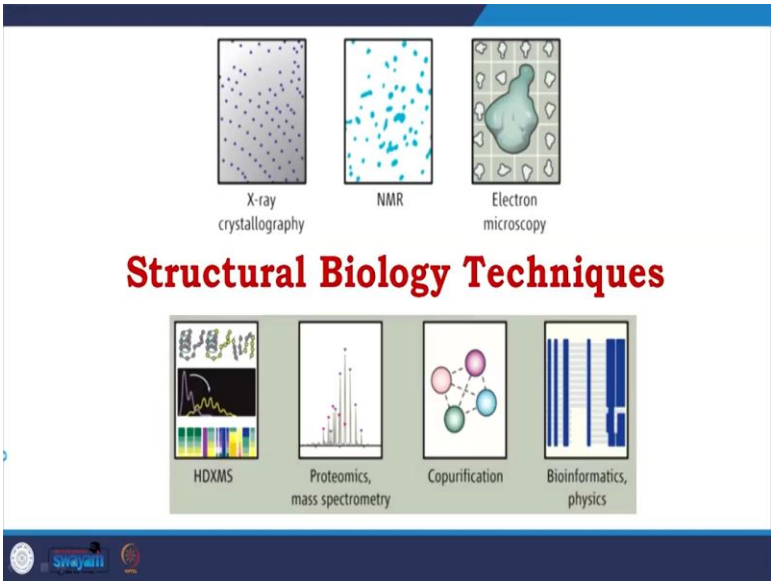
Structural Biology
Prof. Saugata Hazara
Department of Biotechnology
Indian Institute of Technology - Roorkee

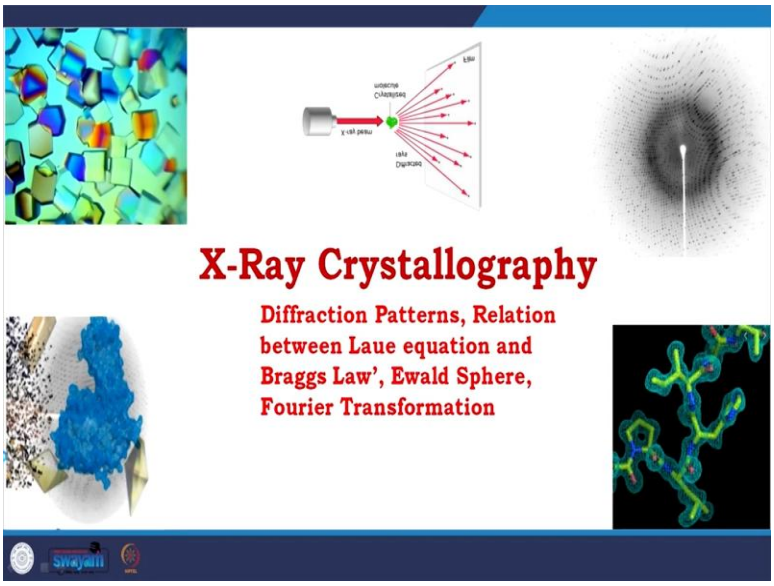
Lecture – 21
X-Ray Crystallography: Data Analysis, Part-I

Hi everyone, welcome back to the structural biology course. We are discussing structural biology techniques, and today we are starting on the 5th module, where we will continue with X-ray crystallography.

(Refer Slide Time: 00:44)

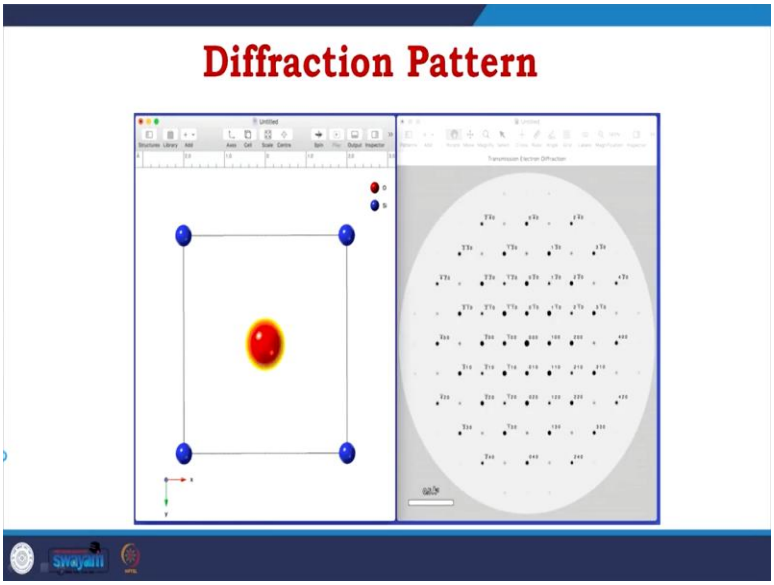


(Refer Slide Time: 00:46)



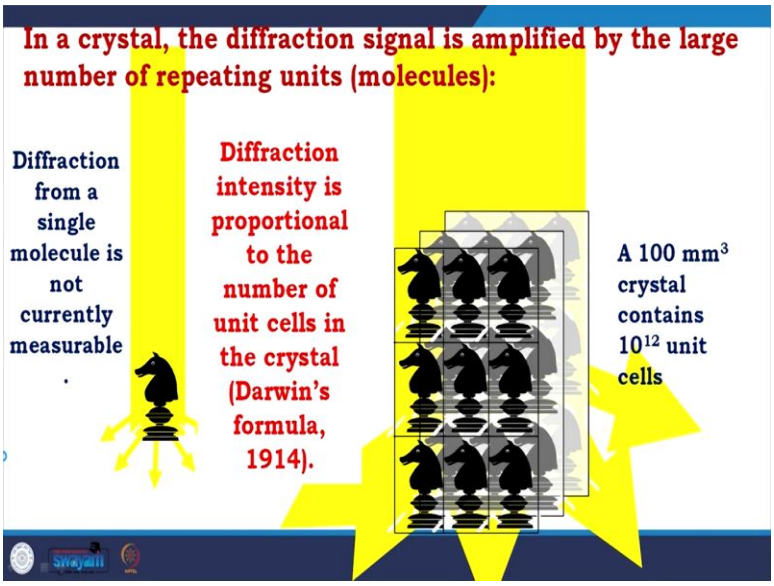
Today, we will discuss its diffraction patterns, the relation between the Laue equation and Braggs law. I talked about Braggs law, but now we will look at instead of an atom, if it is a lattice, how the Braggs law is different, and then we will see the relation between Bragg and Laue. And then, we will look at how Ewald Sphere and Fourier transformation is also working.

(Refer Slide Time: 01:22)



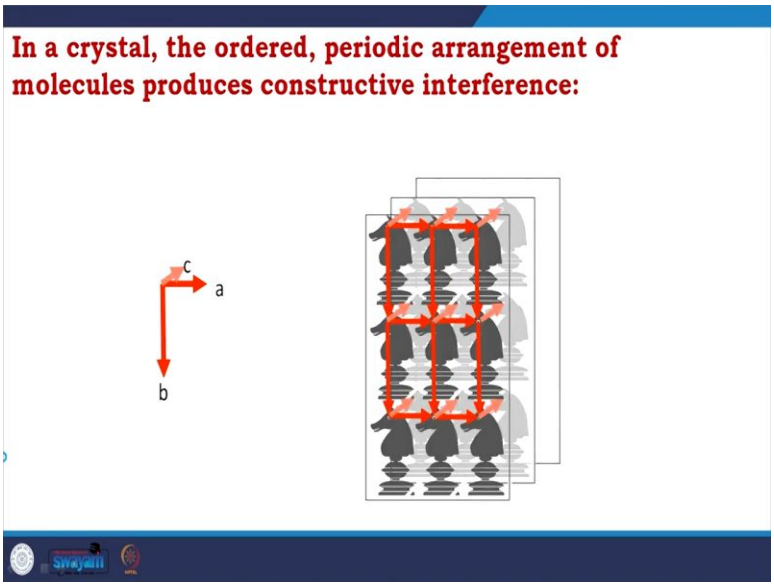
So, diffraction pattern I talked about data collection, but the core of data collection is always the diffraction pattern, as you see in this picture where the atom is heated, and a pattern is developed

(Refer Slide Time: 01:40)



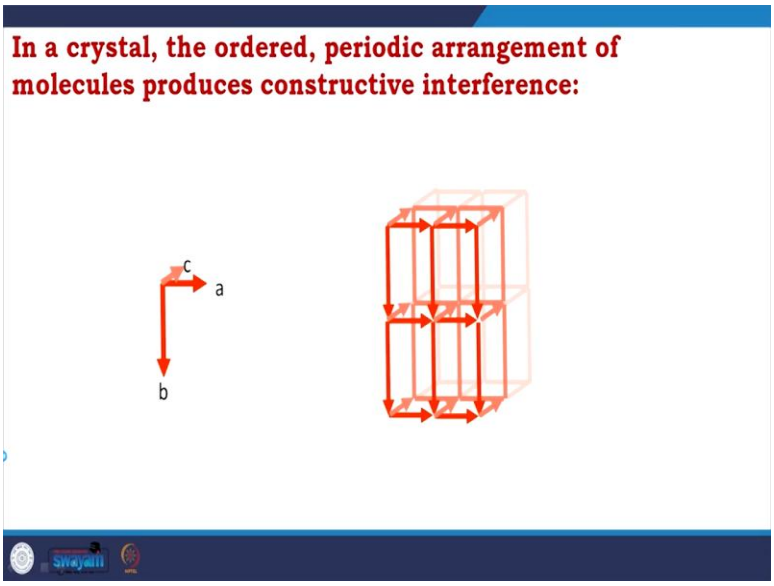
Do you remember in the introductory classes where I talked about why we need a crystal instead of working on a single protein? I told in a crystal the large number of repeating units amplifies the diffraction signal. So, when it is a single molecule, the diffraction from a single molecule is not certainly miserable. It is a very weak signal, whereas when you see the proper arrangement, diffraction intensity is proportional to the number of unit cells. So, instead of 1, there is 10^{12}, so the intensity is much higher. So, you could easily measure that.

(Refer Slide Time: 02:38)



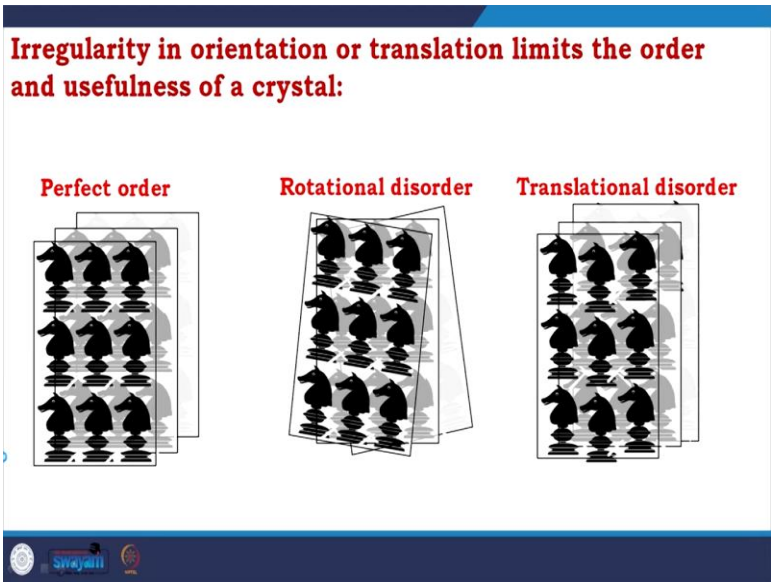
In a crystal, the ordered periodic arrangement of the molecule produces constructive interference

(Refer Slide Time: 03:07)



So, Crystal organized a very beautiful compact 3D arrangement, but if this 3D arrangement did not perform properly, there would be a problem. And this problem would be identified by looking at the diffraction pattern. That is why distraction pattern is very important in protein crystallography. You got a crystal you are happy that you will get information. But when you diffract it, the diffraction pattern will immediately tell you. You got a good crystal or not when I say good crystal now you understand properly arranged.

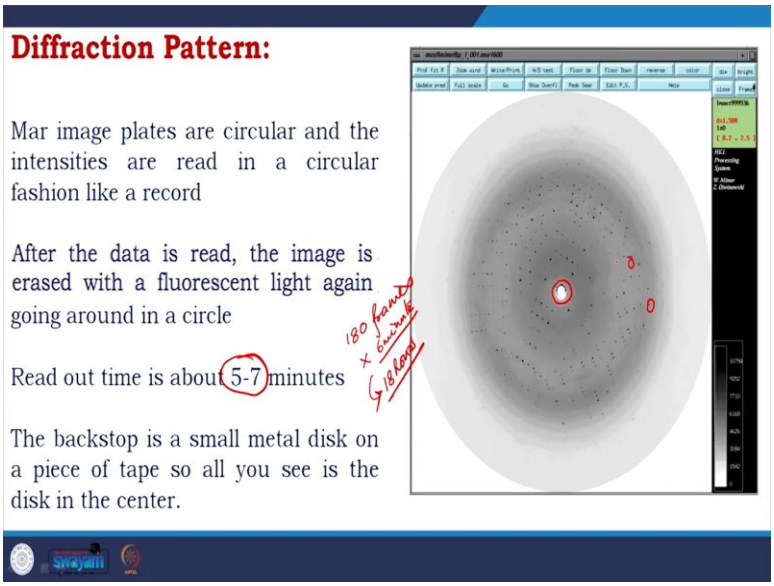
(Refer Slide Time: 04:01)



So, what could be bad? It could be bad in many ways, but in general, if you look at this crystal, it is perfect, then there might be rotational disorder as you could see here or translational disorder as you could see here. Now, if you see the irregularity in orientation translation limit the order

and usefulness of the crystal, I agree that crystallization is the most important thing. Still, once you get a crystal, that does not mean that you are done. You have to have the crystal in perfect order, and you have no role in it. This is the way the process of crystal growth the crystal order would be developed.

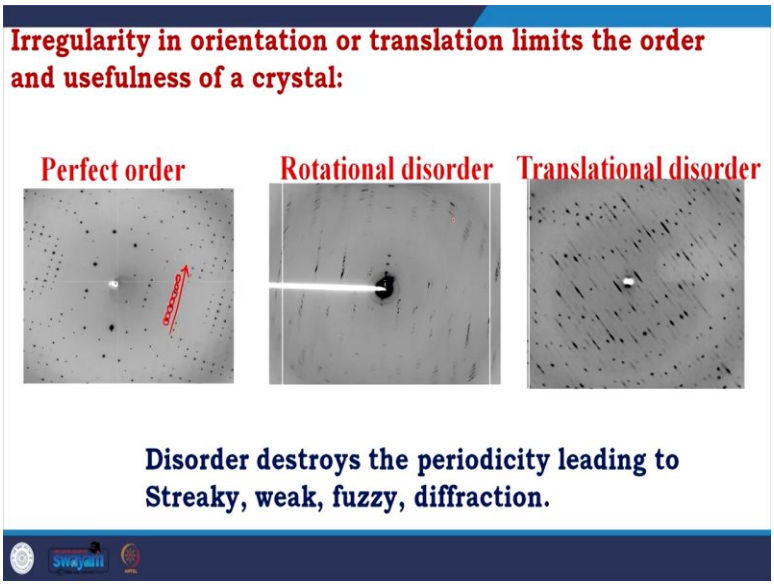
(Refer Slide Time: 04:56)



So, I talked about the diffraction pattern you see here: a mass films screen. So, this is Mar image plates the circular as you see. And the intensities are read circularly like a record. After the data is read, the image is erased, with the fluorescent light goes around in a circle again. So, it takes time it takes about 5 to 7 minutes. I am talking about these make data collection at home hugely time-consuming. You see, this is the effect of the backstop. You do not see any diffraction here because it is a small metal disc on a piece of tape. So, all you see is the disc in the center. So, when you have the beam, you do not want the direct ray to hit. So, you put a beam stop. The beam stops are different types. I will show you another type of beam stop once we are going through the journey of our diffraction pattern.

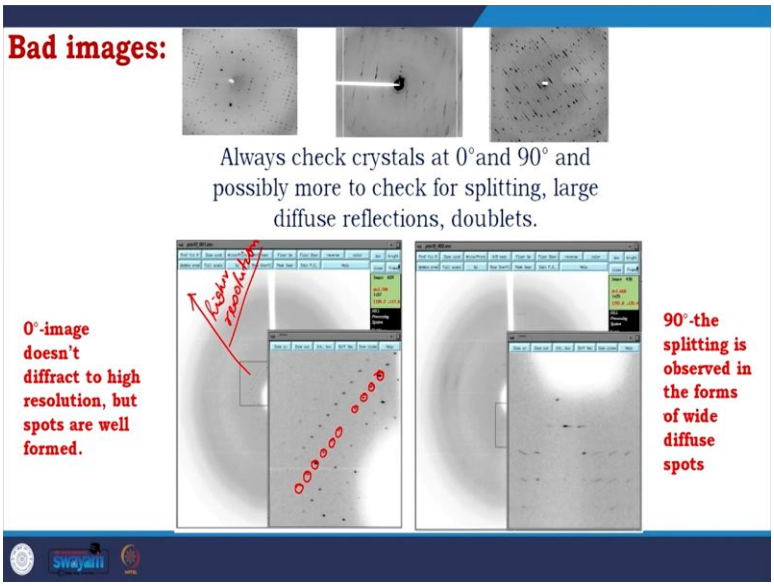
So, we look at the diffraction pattern, understand the read-out time, and understand the back stopper beam stop. If you see that spot, you will see that some sports are very dark. Let us say this one in comparison to the other one. Let us say this one this is where that difference in the data collection makes because of the angles where it is hitting and diffracting.

(Refer Slide Time: 08:12)



So, I have shown you these slides, but to show the irregularity in orientation on translation limits the order and usefulness of the crystal.l Perfect order. You see, you get all of them in 1 line. So, you get a pattern. That is the practice and pattern, and here you see rotational disorder, and none of the spots are individually clear. Here you see translational disorder, and the spots are merged between each other. So, this order destroys the periodicity leading to streaky, weak, fuzzy diffraction.

(Refer Slide Time: 09:41)



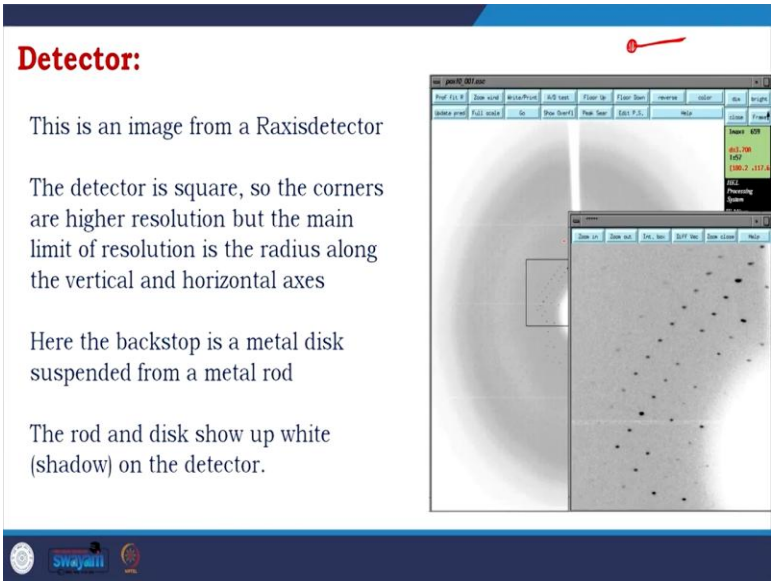
So that is what I was talking about bad images. And another very important thing when you are going for data collection before you should always start like 0 degrees. So, if you look at the 0-degree image does not diffract to high resolution here, but the spots are very well-formed. So, a

clear pattern is developed now how I know that they are not in high resolution looking at those layers more these spots going on the corner higher the resolution.

But what you should do you should immediately go to 90 degrees, 180 degrees, 270 degrees because you cannot check all the images right then, but you should do 0, 90, 180, 360, in that way you get the same crystal and come to 90 degrees the splitting is observed in the form of diffuse white spots you could see it. So, now you know that this crystal is not good.

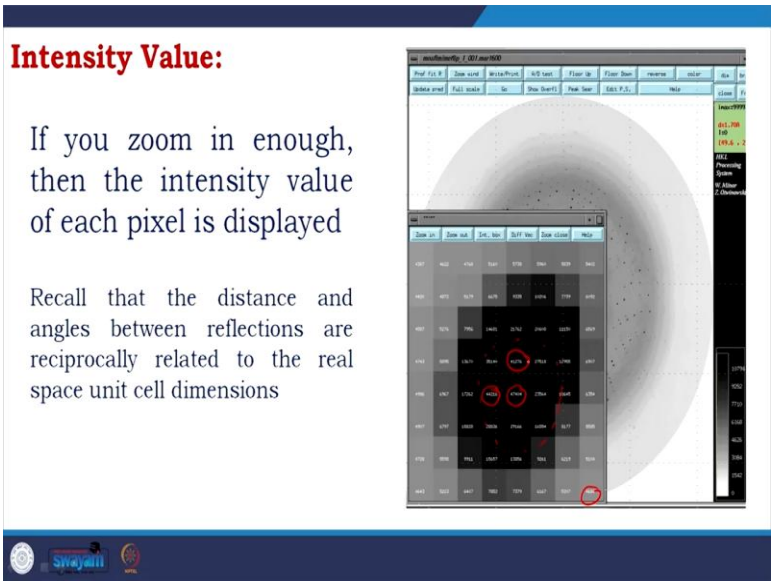
So, always check crystal at 0 degrees and 90 degrees and possibly more to check for splitting large diffuse reflections and dublets, as I am repeatedly telling. So, before data collection starts, you should look at the diffraction patterns very carefully.

(Refer Slide Time: 12:03)



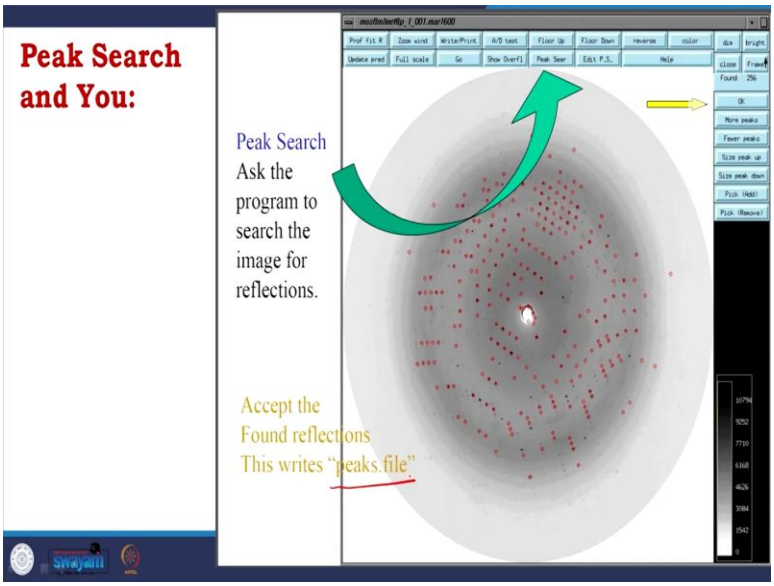
Detectors I talked about, I show you the Marwan. This is an image from a Raxisdetector, which is from Rigaku. The detector is square, so the corners are higher resolution, but the main resolution limit is the radius along the vertical and horizontal axes. Here very interesting if you see the backstop or beam stop is a metal disc followed by a rod. So, here the backstop is a metal disc suspended from a metal rod. The rod and disc show up white on the detector.

(Refer Slide Time: 12:58)



You want to look at the intensity values. So, if you zoom in enough, then each pixel's intensity value would be displayed. So, if you see here, you see that each pixel has its numbers. So, how do you understand which one is the dark one? You see the core ones. If I take this as that core, you see that it is 47000, 44000, 41000, whereas here you get 4000, so 10 times difference in the intensity. And from here, you can understand that the distance and angle between reflections are reciprocally related to the real space unit cell dimensions.

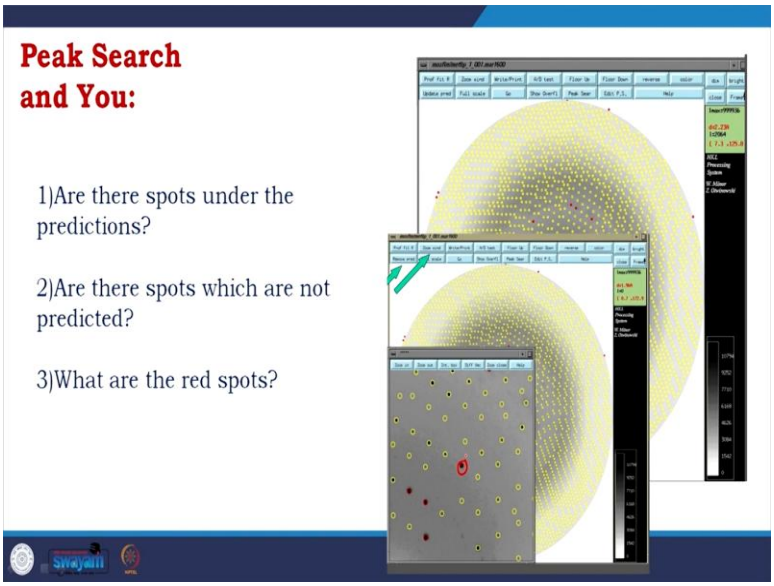
(Refer Slide Time: 13:55)



Very important is how to do the Peak search? So, you could ask the program to search the image for reflection, and when the program is picking up spots if you are happy, you could accept the found reflection, and when you do that, this will write a file called peaks file. So, all the accepted

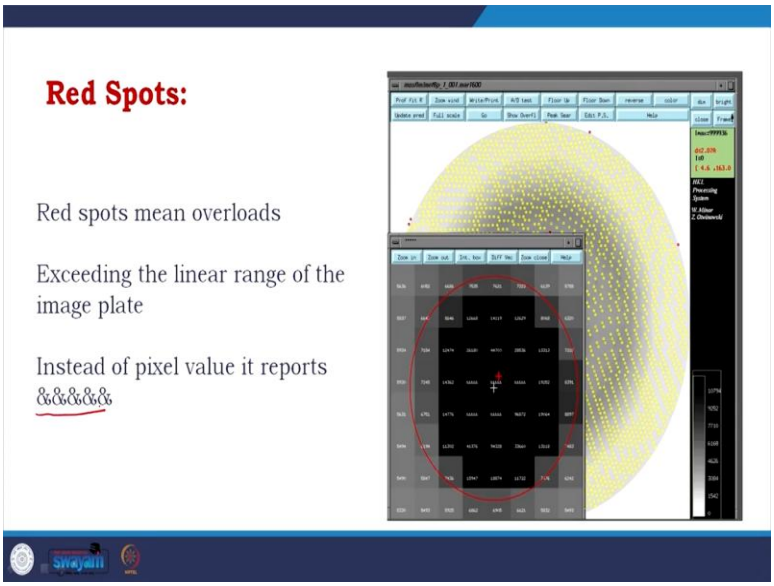
peaks would now be recorded into a separate file. But here actually you could do good work. You could think about putting your mind active towards the questions.

(Refer Slide Time: 14:44)



Are there spots under the prediction? Like all the spots are predicted, or some are somehow not predicted by the program, are the spots not predicted? And what are the red spots? If you see the red spot and I zoom you will understand these are the red spots.

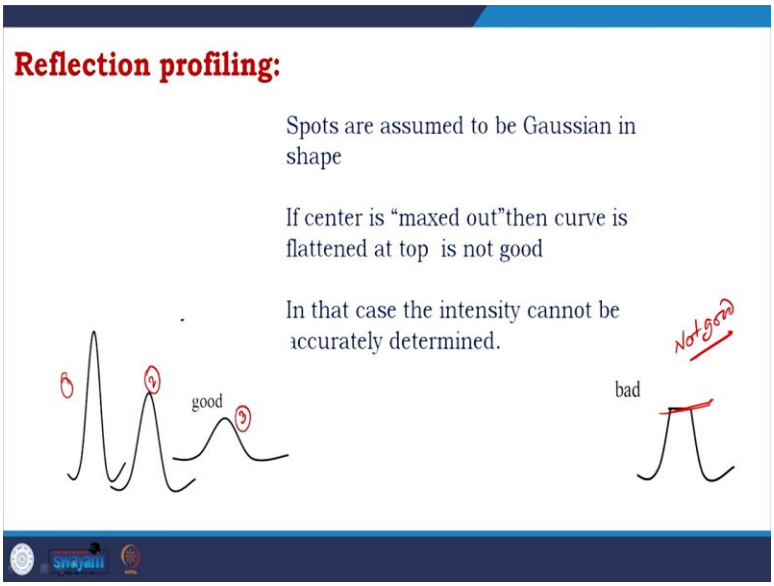
(Refer Slide Time: 15:10)



So, the red spots mean overloads because of the exceeding intensity in the linear range of the image plate, so the image plate has a capacity. If the intensity is more than that, then you get red

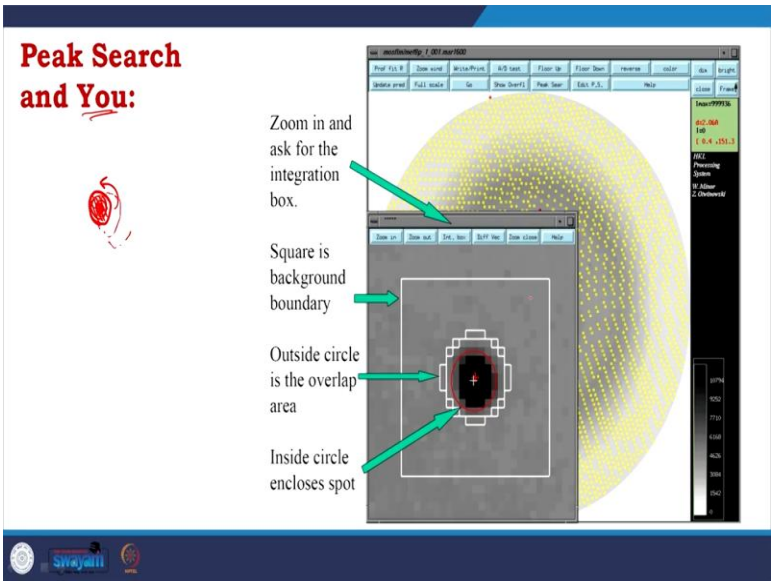
spots. Here instead of pixel value, it reports with this sign. So, now you are looking at the spots and modifying them.

(Refer Slide Time: 16:06)



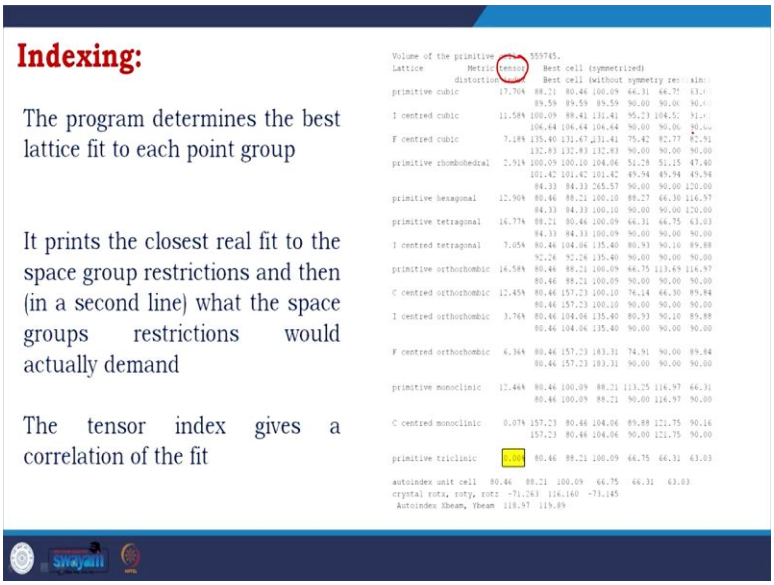
So that is called reflection profiling; spots are assumed to be Gaussian in shape. You get these once all are good when you see a builds up, but if the center is maxed out, somehow it gets the maximum value, then the curve is flattened at the top like this, no cut at here, and this is not good, in that case, the intensity cannot be accurately determined. So, you have to understand those basics.

(Refer Slide Time: 16:48)



Now how could you work on it? You zoom in and ask for the integration box. What is an integration box? The box; where the spots are. Then, this square is the background boundary. The square is the noise, the outside circle is the overlap area, and the inside circle encloses the spot. So, now you open it, and you could modify it to a little bit, you know, centering or something, and in that way, if you do that you could find some spots which are not taken automatically by the program. You could do that manually; what is the benefit of that? By doing that, you could make some false spots out, and you could make some real spots in, which would help your indexing and further data processing with higher quality.

(Refer Slide Time: 18:44)

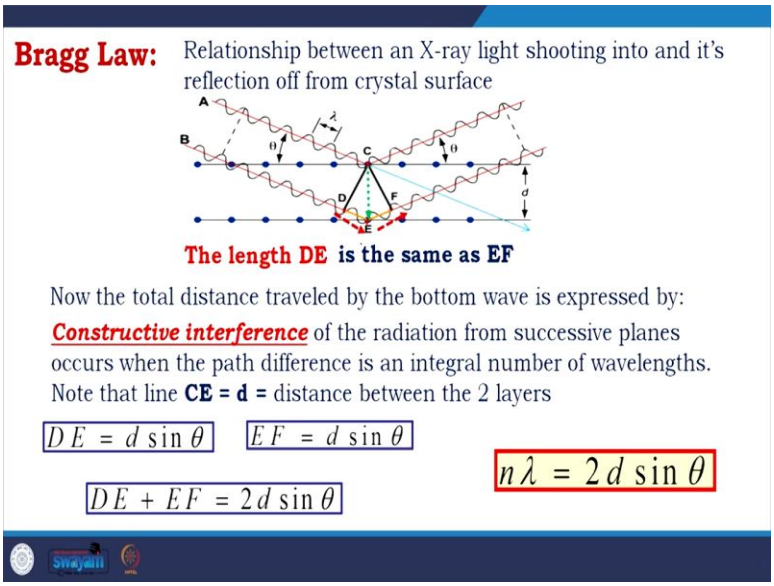


So, indexing we talked about, the program determines the best lattice fit to each point group. It brings the closest real fit to the space group restrictions and what it would demand. So, you got the first data. You look at the quality of the data by looking at the diffraction pattern. And now you go to the next step, indexing the tensor index is the column you will see which gives a correlation of the fit.

So, if you look at what you learn? The diffraction pattern you learn is the first critical thing you have to understand while going for crystallographic data collection. If you see that the spots are not in a pattern if they are diffuse if they are like merging and all these things, you know that the crystal quality the ordering of crystal there is some mistake, but some time you do the auto collection of spots to get some spot, now by working manually, you could improve that and once you do that you come to the indexing.

So, this is a file as I talked about, this is the tensor index. In the tensor index, if you are getting values for different space groups primitive cubic, I centered cubic, F center cubic, primitive rhombohedral, primitive hexagonal, primitive tetragonal, I centered tetragonal, primitive orthorhombic, C center orthorhombic, I center orthorhombic, F center orthorhombic, primitive monoclinic, C center monoclinic, primitive triclinic and then you get this tensor index and you know which cells are best with the symmetry distance and without the symmetry distance. So, here you see, this has a 0% tensor index.

(Refer Slide Time: 21:13)



We have discussed Bragg law. Bragg's law talked about the relationship between an X-ray light shooting into and reflecting from the crystal surface, but we will look at it from a different perspective. So, what we know here is the length when there are two reflected lights the length DE which is same as the EF. So, this D to F, now the total distance traveled by the bottom web could be expressed by its constructive interference of the radiation from successive planes occurs when the path difference is an integral number of wavelengths.

Now the distance CE, which the green line equals d the distance between the two layers and

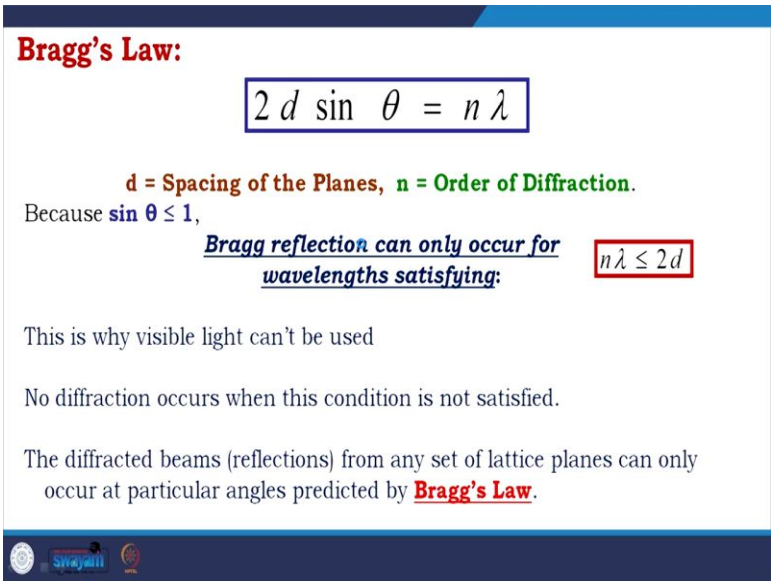
$DE = d \sin\theta$

$EF = d \sin\theta$.

So, $DE + EF = 2d \sin\theta$

and $n\lambda = 2d \sin\theta$. So that is what Bragg's law we have discussed before.

(Refer Slide Time: 22:43)



As I told d is the spacing of the plane, n = order of diffraction

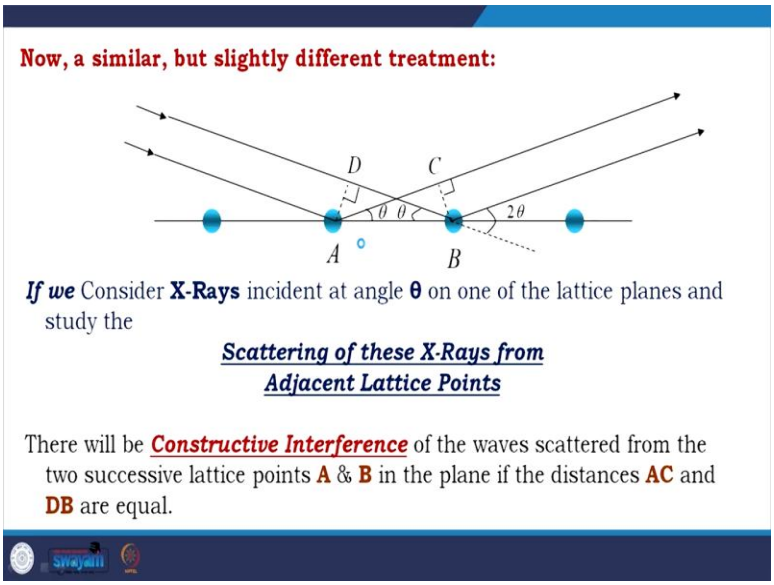
$2d \sin\theta = n\lambda$.

Now because $\sin\theta \leq 1$,

Bragg reflection can only occur for the wavelengths satisfying

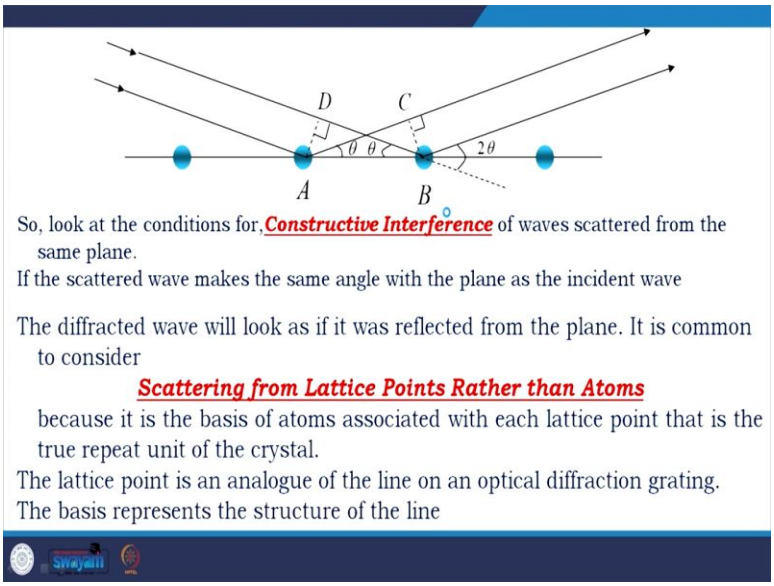
$n\lambda \leq 2d$. And that is the reason visible light cannot be used. No diffraction occurs when this condition would not be satisfied. The diffracted beams from any set of lattice planes can only occur at particular angles predicted by Bragg's law.

(Refer Slide Time: 23:29)



So, this is what we discussed. And I have to bring them again because I want to discuss a similar but slightly different treatment here instead of the atoms we are discussing lattice point. So, let's consider X-rays incident at angle theta on one of the lattice planes and study the scattering of these X rays from adjacent lattice points. There will be constructive interference of the waves scattered from the two successive lattice points A and B in the plane if the distance is AC and DB are equal.

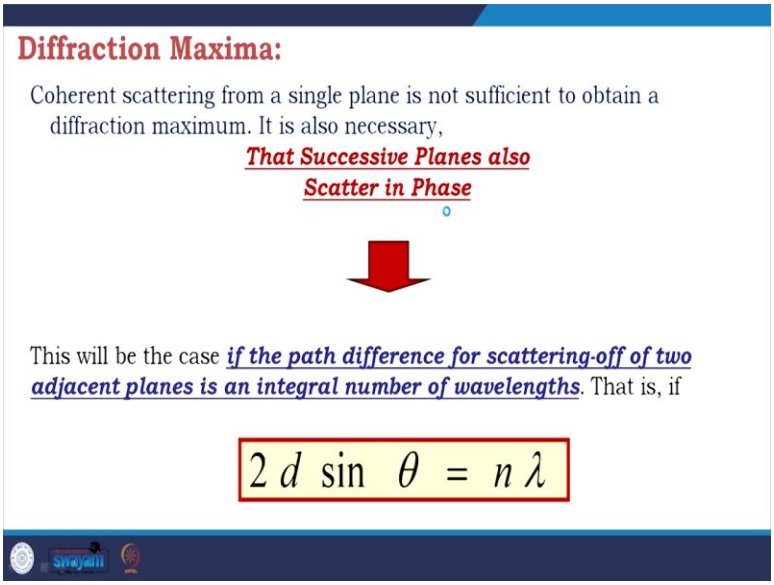
(Refer Slide Time: 24:23)



So, look at the condition for constructive interference of waves scattered from the same plane. If the scattered wave makes the same angle with the plane as the incident wave, the diffracted wave will look as if it was reflected from the plane. So, instead of diffraction, it will look as it is

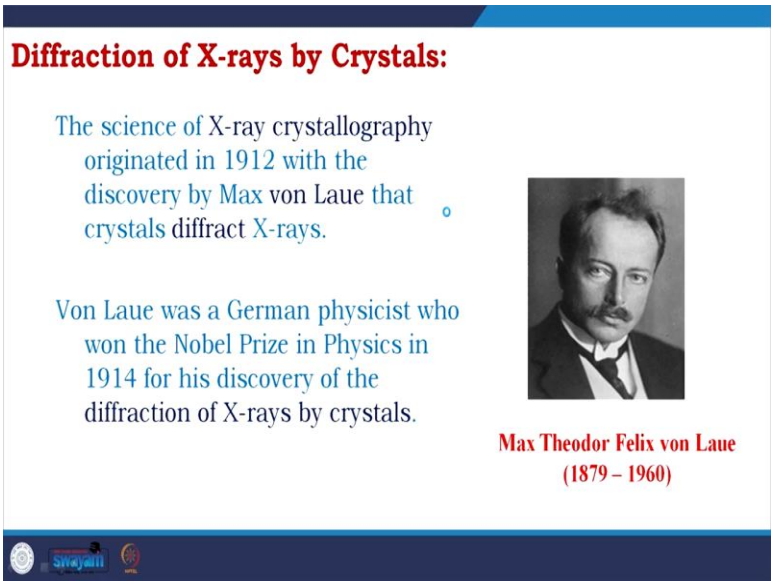
reflected. So, it is common to consider the scattering from lattice points rather than atoms because atoms are there if you have followed the discussion and understood. But the orientation would be in the lattice. So, the lattice points are important because it is the basis of atoms associated with each lattice point is. The true repeat unit of the crystal atoms is there, but the unit cell is the lattice. The lattice point is an analog of the line on an optical diffraction grating the basis, representing the line's structure.

(Refer Slide Time: 25:41)



Coming to diffraction maxima, coherent scattering from a single plane is not sufficient to obtain diffraction maxima. It is also necessary that successive planes also scatter in phase. So, this will be the case if the path difference for the scattering-off of 2 adjacent planes is an integral number of wavelengths, which happens if $2d \sin\theta = n\lambda$. So, rather than the atom, it comes more important when it is the lattice points we are now sipping our journey from Bragg to Laue.

(Refer Slide Time: 26:35)

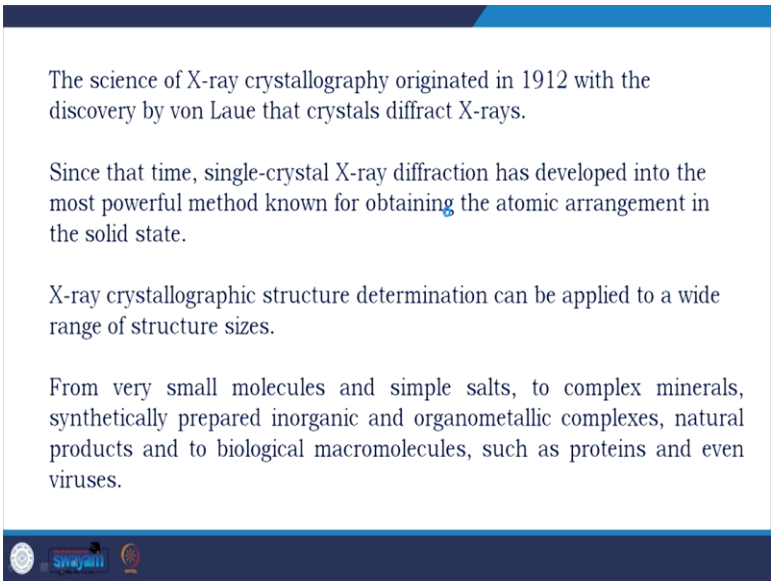


So, diffraction of X rays by crystal: The science of X-ray crystallography originated in 1912 with the discovery of Max Von Laue that crystal diffracts X rays. Von Laue was a German physicist who won the Nobel Prize in Physics in 1914 to discover the diffraction of X rays by crystal. So, he is considered the father of X-ray diffraction.

When Von Laue was working as a lecturer in LMU under Summer fields, Evald was writing his doctoral thesis under the same guide. Their interaction was critical and hugely helped each other. They interacted with Bragg and influenced each other interestingly. When Bragg came to Cambridge, they made Watson and Crick, and now you know the history.

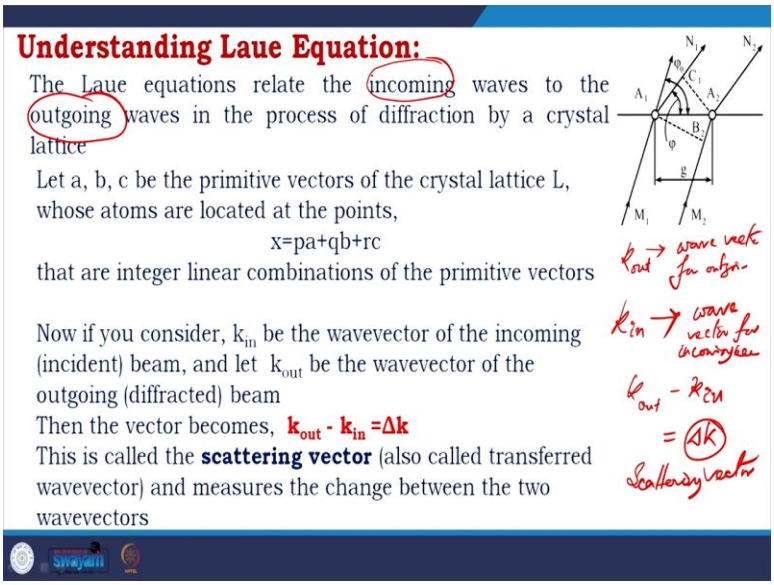
So, you know how the technique of X-ray crystallography helped Watson and Crick to get the first structure of the biological macromolecules DNA which probably happens one of the most significant discoveries in the history of scientific discovery to understand DNA to understand the pattern to understand its structure and further the birth of molecular level biology based on visualization.

(Refer Slide Time: 28:49)



The science of X-ray crystallography originated in 1912 with the discovery of von Laue that crystal diffracts X-rays. Since then, single-crystal X-ray diffraction has developed into the most powerful method for obtaining the atomic arrangement in the solid-state. X-ray crystallographic structure determination can be applied to a wide range of structure sizes that we have discussed and will discuss. From very small molecule and simple salts to complex minerals, synthetically prepared inorganic and organometallic complexes, natural products, and at the base to the biological macromolecules, such as protein and now even the big viruses. So, you could easily understand Von Laue, and his discovery helped change this spectrum or the research of X-ray diffraction.

(Refer Slide Time: 30:06)



The Laue equation relates the incoming waves to the outgoing waves in diffraction by a crystal lattice. So, if you look at this picture and consider a, b, c to be the primitive vectors of the crystal lattice L. So, this is the crystal lattice we are considering where we are considering a, b, c be the primitive vector the atoms would be located at the point

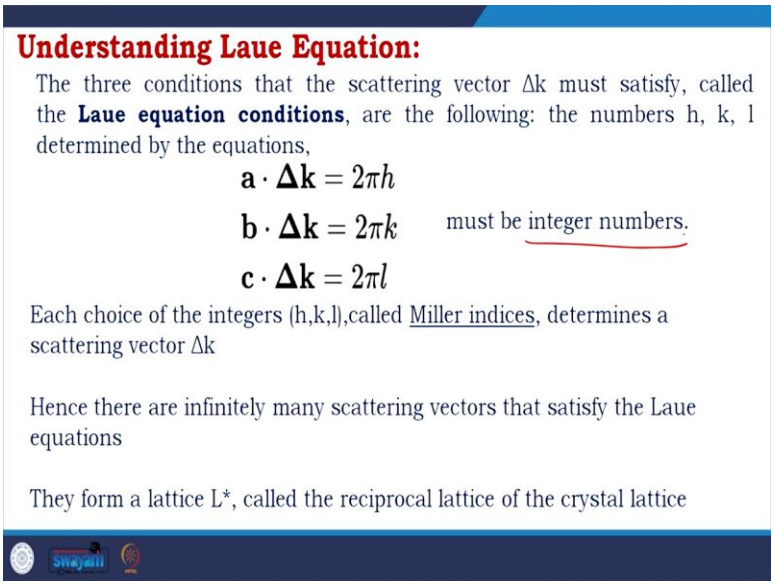
x (any point) = pa + qb + rc

these are the integer linear combination of primitive vectors.

Now, if we consider k_{in} be the wavevector of the incoming beam, and let k_{out} be the wavevector of the outgoing beam, the difference $k_{out} - k_{in} = \Delta k$

Δk is called the scattering vector, it is also called transferred wavevector, and it measures the change between the two wavevectors. So, what happened as I told it is the comparison between the incoming waves and the outgoing waves and considering k_{in} as the incoming beam wavevector for incoming beam, k_{out} again wavevector for outgoing beam, their difference $k_{out} - k_{in}$ is Δk, this Δk is also vector which is scattering vector.

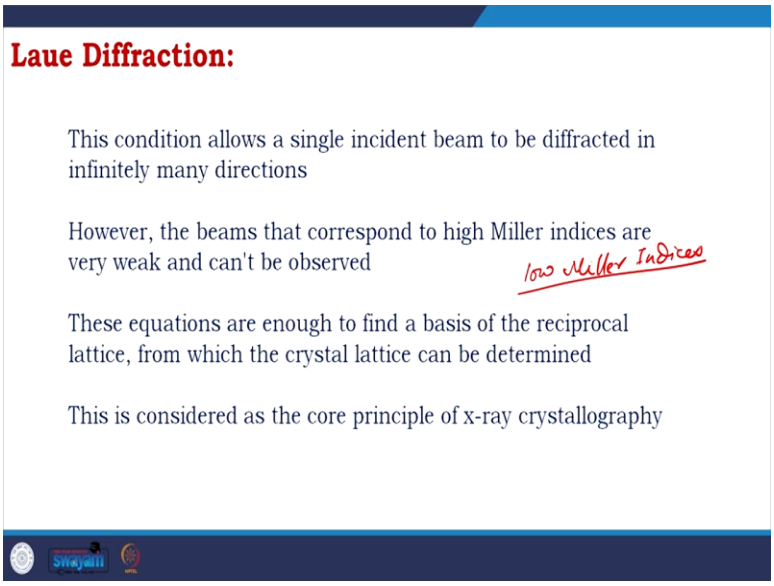
(Refer Slide Time: 32:41)



The three conditions that the scattering vector Δk must satisfy, called the Laue equation conditions are the following the numbers h, k, l determined by the equation a dot del k = 2pi h, b dot del k = 2pi k and c dot del k = 2pi l and they should be integer number. So, this is a conversion from a to h, b to k and c to l which is the relation between a real lattice and a reciprocal lattice each choice of the integers h k l called Miller indices determines a scattering vector del k.

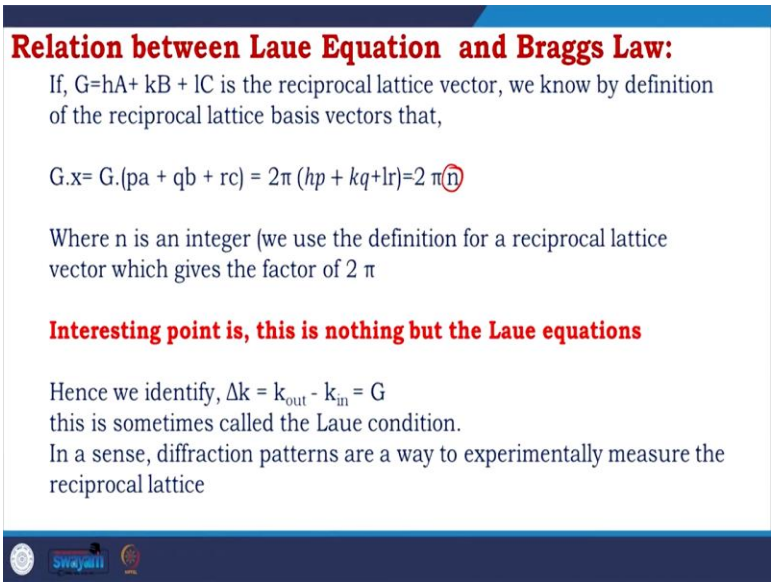
So, now you understand the relation because you already know about the Miller indices you already know about the reciprocal space and you know how amazingly the Miller system is developed the Miller indices system so that we could get integers when it is parallel. So, many cases could be solved by using Miller indices. Hence there are infinitely many scattering vectors that satisfy the Laue equation. So, they who satisfy they form a lattice L star called the reciprocal lattice of the crystal lattice.

(Refer Slide Time: 34:16)



So, now you understand the correlation this condition allows a single incident beam to be diffracted in infinitely many directions. However the beams that correspond to high Miller indices are very weak and cannot be observed. So, the beams with low Miller indices are important or it could be measured. These equations are enough to find a basis of the reciprocal lattice from which the crystal lattice can be determined. So, you could get each other this is considered as the core principle of X ray crystallography. So, you know the condition.

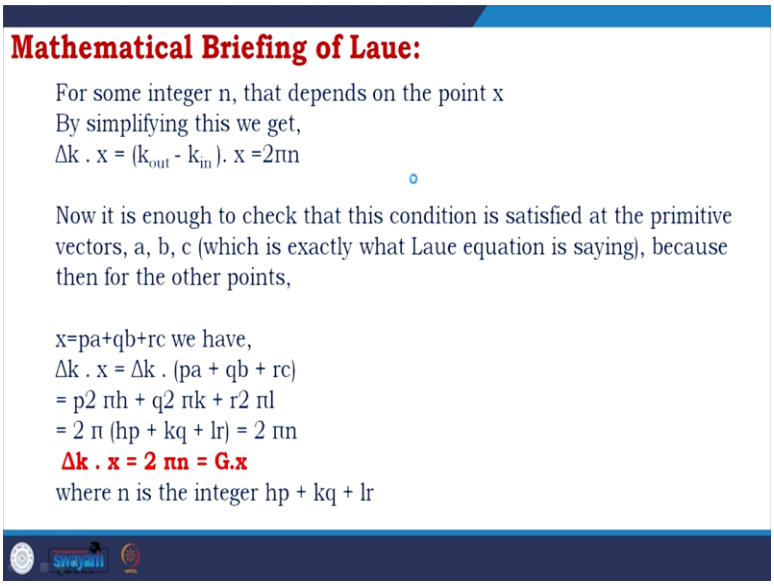
(Refer Slide Time: 35:10)



Now what we will do we will find the relation between the Laue equation and Bragg's law. So, if G = hA + kB + IC is the reciprocal lattice vector we know by definition of the reciprocal lattice basis vectors that G dot x= G dot pa + qb + rc. Now that means 2pi hp + kq + lr = 2pi n where n

is an integer we use the definition for a reciprocal lattice vector which gives the factor of 2pi. Now the interesting point is that this relation is nothing but a Laue equation. You do not get it yes because I did not go into the mathematical part of the Laue equation but let us go there.

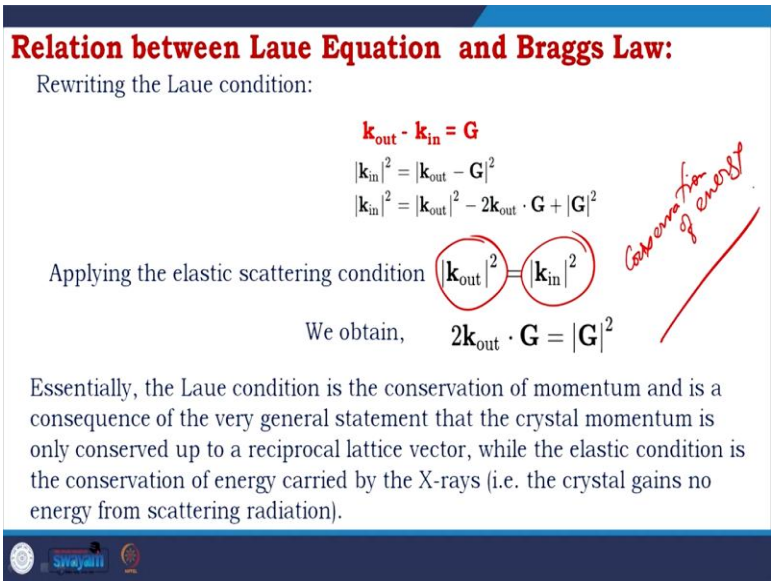
(Refer Slide Time: 36:26)



So, mathematical briefing up Laue for some integer n that depends on the point x by simplifying these we could get delta k dot x = k out – k in dot x = 2pi n. Now this relation is enough to check that this condition is satisfied at the primitive vectors a, b, c which is exactly what Laue equation is saying because then for the other points if you see x = pa + qb + rc we have delta k dot x = delta k dot pa + qb + rc because x is replaced by equal to p2 pi h + q2 pi k + r2 pi l that means 2pi hp + kq + lr which means now you note 2pi n.

So, now if you see delta k dot x = 2pi n which is equal to G dot x when n is the integer hp + kq + lr. So, now you understand that the interesting point is this is nothing but the Laue equation. Hence we identify del k = k out - k in which you know equal to G. This is called Laue condition and in the sense diffraction patterns are a way to experimentally measure the reciprocal lattice.

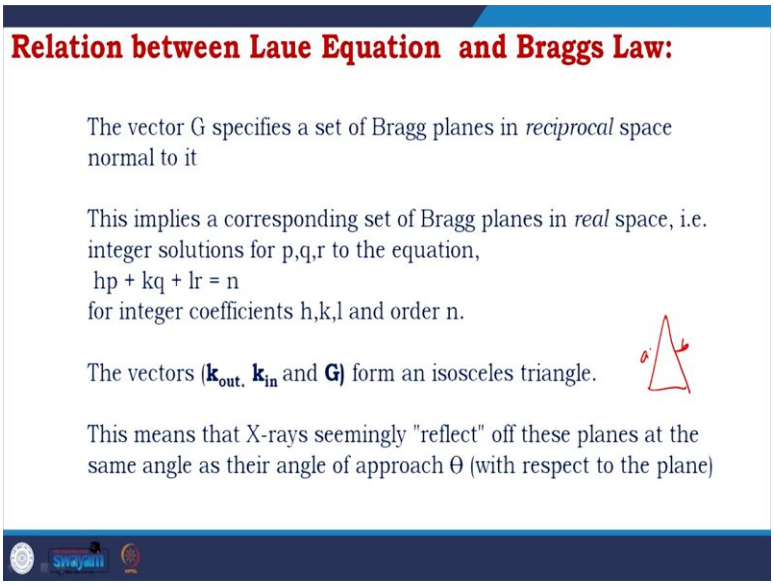
(Refer Slide Time: 38:25)



Now if we rewrite the Laue condition k out - k in = G. So, absolute value of k in whole square equal to capsulate value of k out minus G whole square. So, k in absolute value square equal to k out absolute value square - 2k out dot G + G absolute value square. Now we apply elastic scattering condition for these which means that wavevector of incident beam is equal to the wavevector of the outgoing beam.

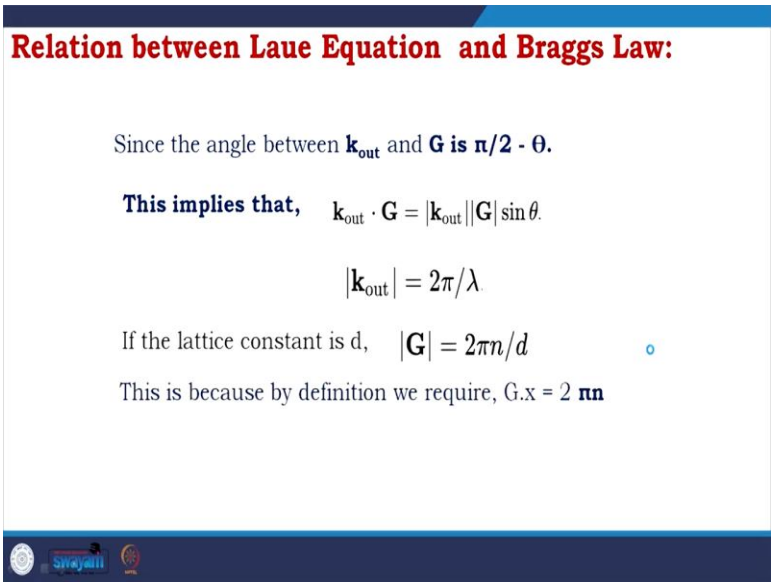
So, 2k out. G equal to absolute G is whole square essentially the Laue condition is the conservation of momentum and is a consequence of a very general statement that the crystal momentum is only concept up to a reciprocal lattice vector while the elastic condition is conservation of energy carried by the X rays as you see here. This equal to this so you could say conservation of energy.

(Refer Slide Time: 40:02)



Now the victor G specifies a set of Bragg planes in the reciprocal space normal to it. These implies a corresponding set of Bragg planes in real space that is integer solution for p, q, r to the equation which is hp + kq + lr = n for integer coefficients h, k, l and order n. Now the victors k out and k in and G will form an isosceles triangle and isosceles triangle means 2 sides are equal this means that x rays seemingly reflect up these planes at the same angle as their angle up approach theta with respect to the plane.

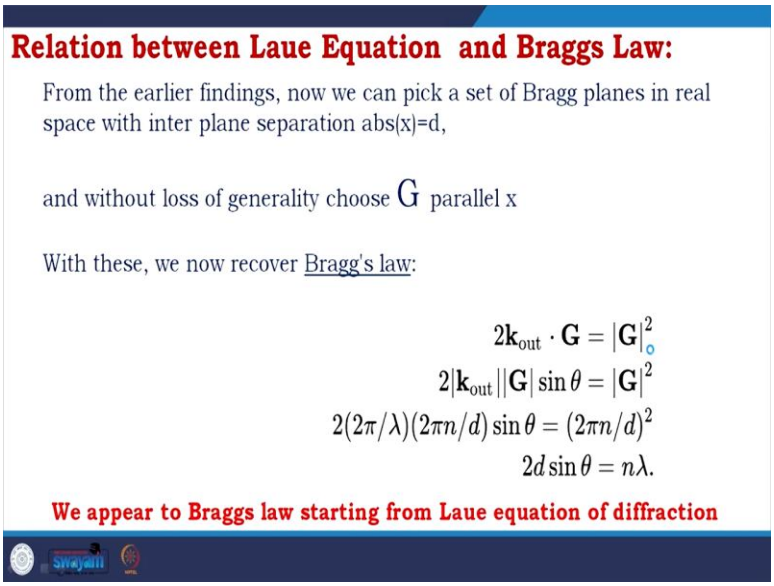
(Refer Slide Time: 41:07)



Since the angle between k out and G is pi by 2 - theta this implies that k out dot G equal to absolute value of k out into absolute value of G sin theta. So, absolute value of k out equal 2pi /

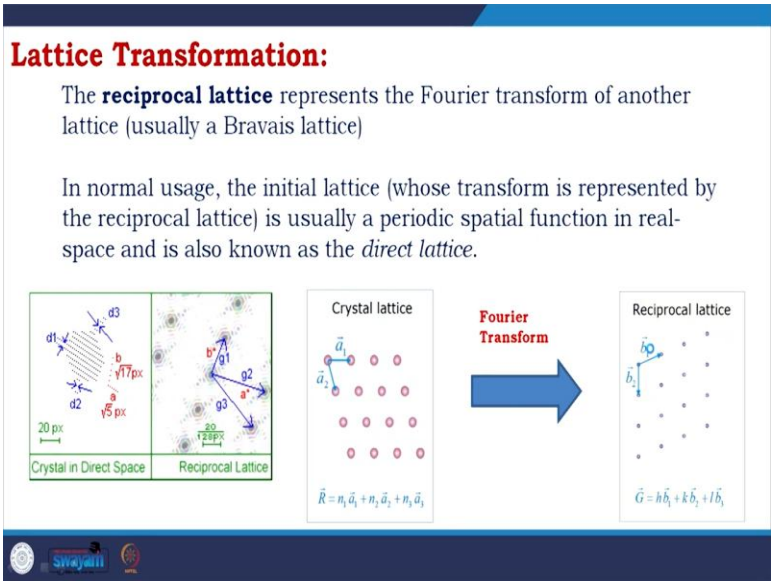
lambda if the lattice constant is d the spacing between the 2 layers G absolute equal to 2pi n / d this is because by definition we require G dot x = 2pi n.

(Refer Slide Time: 41:50)



So, from the earlier findings now we can pick a set of Bragg planes in real space with inter plane separation absolute x = d and without loss of generality choose G parallel x with these we could now recover Bragg's law you will see 2k out. G equal to absolute G squared 2k out G sin theta is G squared 2. So, we replaced with the values we got 2 2pi / lambda 2pi n / d sin theta = 2pi n / d square which is 2d sin theta equal to n lambda. So, in that way we appear to Bragg's laws equations starting from Laue equation of fraction

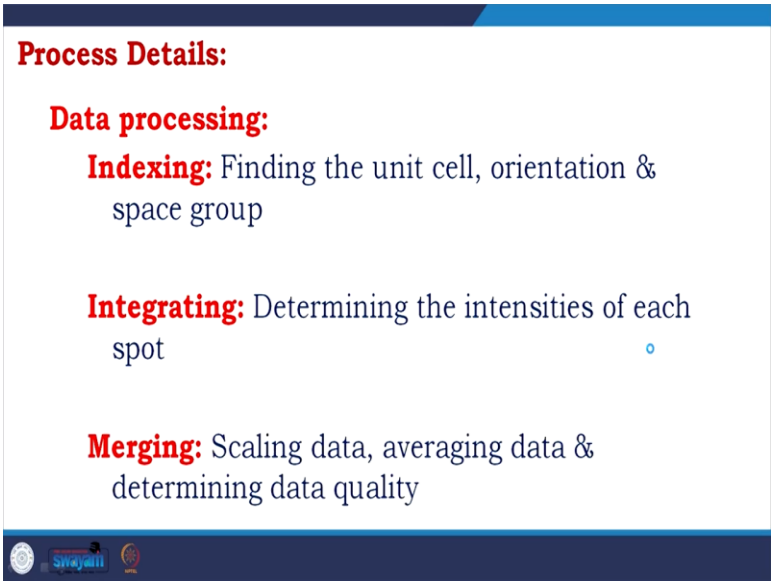
(Refer Slide Time: 42:51)



If we look for lattice transformation the reciprocal lattice represents the Fourier transform up another lattice usually a Bravais lattice in the normal usage the initial lattice whose transform is represented by the reciprocal lattice is usually a periodic spatial function in real space and is also known as the direct lattice. So, you see the crystal lattice that is the direct lattice and you see the reciprocal lattice.

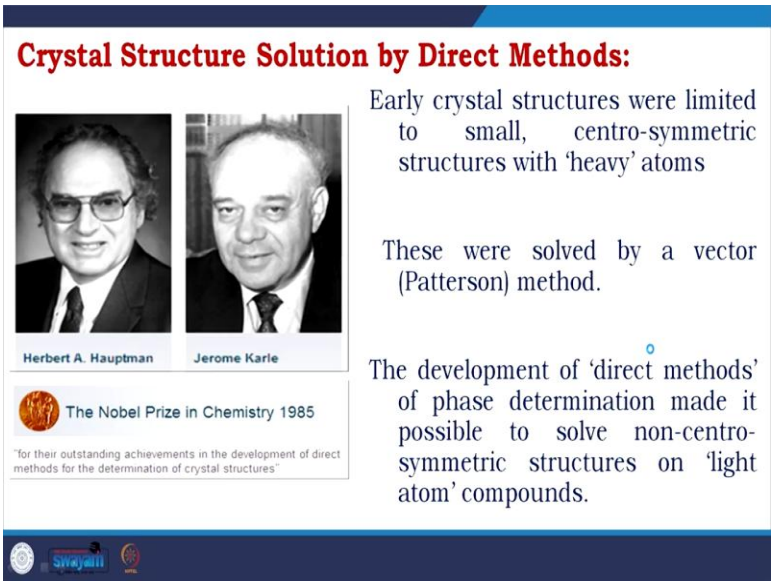
So, there inter connections is very important and how you could to do that you could do that with a mathematical operation which known as Fourier transform and if you see you will see the picture, pictorial representation for crystal in a direct space and crystal in a reciprocal space. So that difference is clear here and our data actually connect as who as shown in different aspects we are getting data from direct and some from reciprocal lattice measurement.

(Refer Slide Time: 44:25)



So, the process details in data analysis the first step indexing as I have discussed finding the unit cell, orientation and space group than we have to integrate the determining the intensities of each spot and then we have to do merging which is scaling, averaging and determining the data quality.

(Refer Slide Time: 44:54)



The crystal structure solution was first happen by direct method Herbert A. Hauptman and Jerome Karle they awarded it noble prize chemistry in 1995 for the outstanding achievements in the developments of direct method for the determinations of the crystal structure. So, as we know early crystal structure as limited to small centro- symmetric structures with heavy atoms it was easy this was solve by vector method which is Patterson method.

The development of direct method of phase that determinations made it possible to solve non-centro-symmetric structures on light atoms compounds and then we will discuss about what we call the problems and all you phase in solving the protein crystals in the next class we will look at about in direct method, about the famous phase problem and then about structure factor, intensity and all regarding that thank you very much.