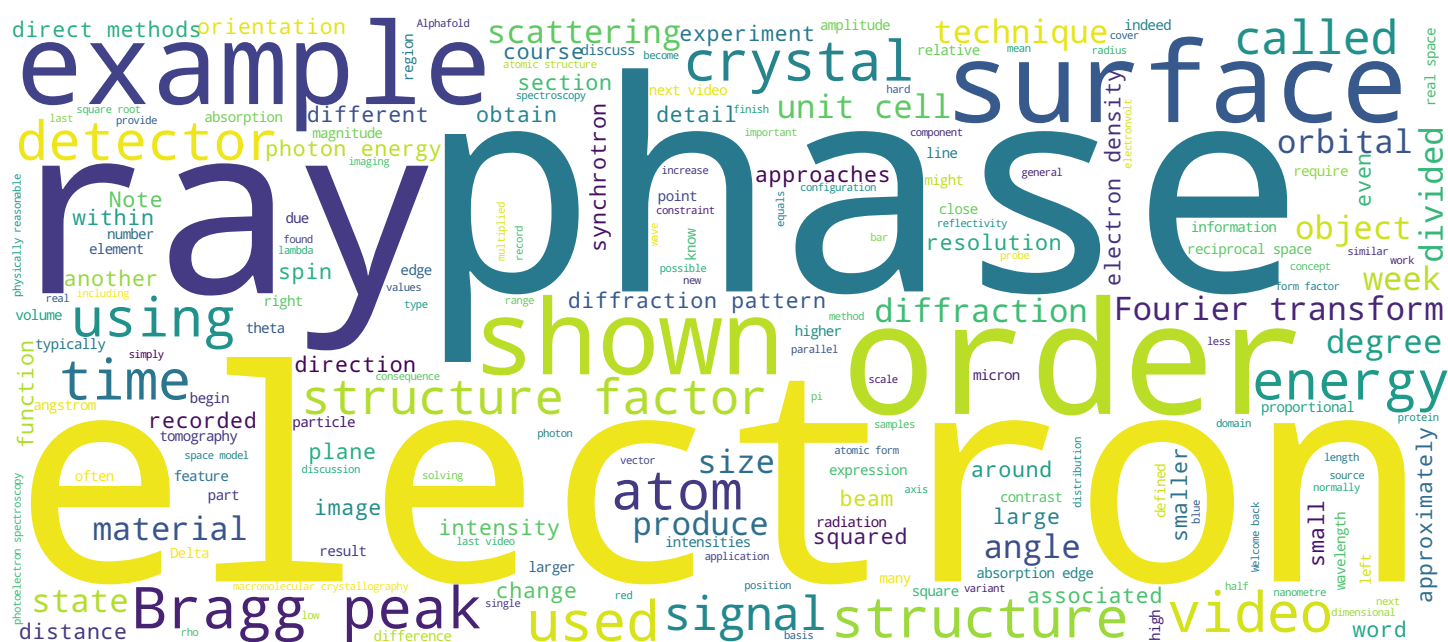


Synchrotrons and x-ray free-electron lasers

Techniques and applications

Prof. Philip Willmott



Search MOOC



Video



Contents and objectives of this video



- What's the problem?
- Solving the problem outside MX
 - Direct methods
 - Iterative methods

Hello again. In this video, I'll outline in more detail what the phase problem is and some approaches to resolving it, particularly outside the domain of macromolecular crystallography, which I defer to videos four, five, and six of this section.

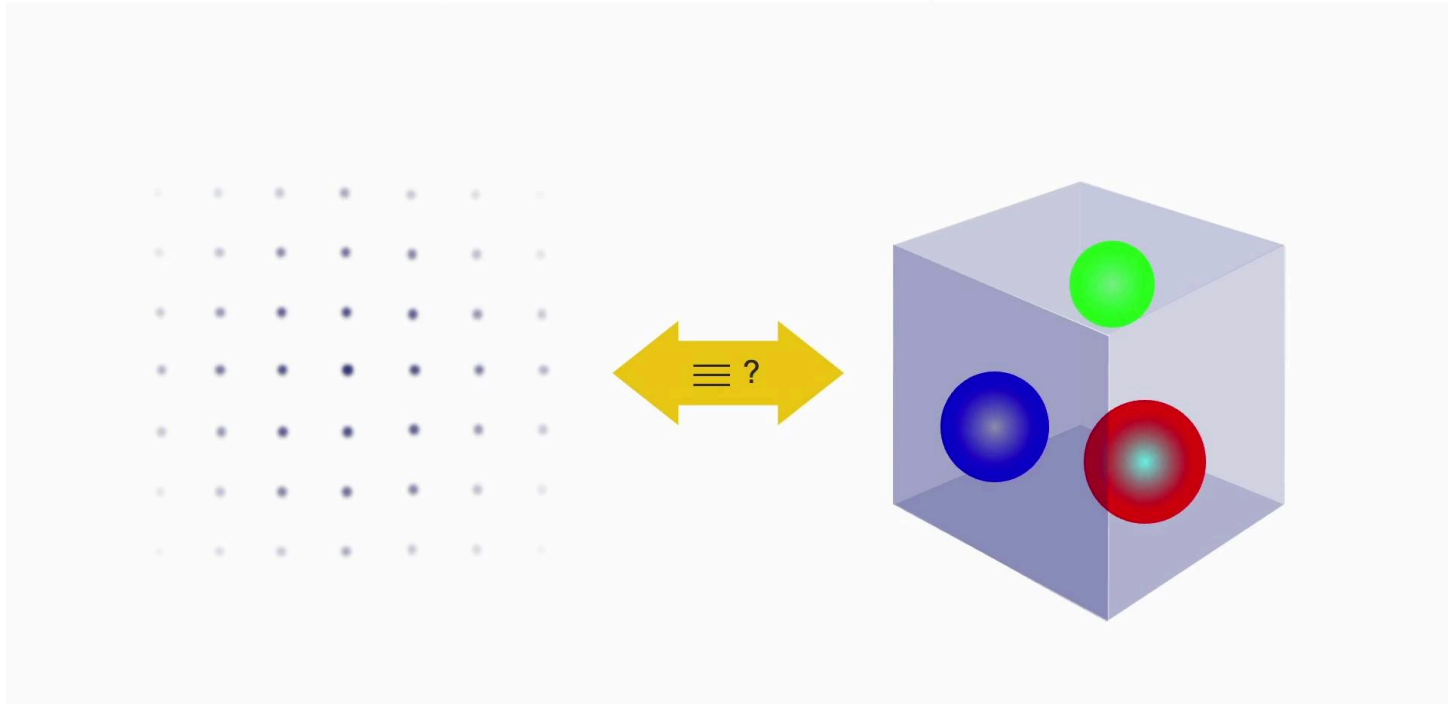
Notes

Summary



0m 05s

Constructing the structure factor, F



Now, if we already know the atomic positions, we can easily determine the phases as we've just demonstrated. However, what we normally have is a set of intensities from which we would like to work out the atomic structure. But we don't know a priori the phases because we also don't know where the atoms sit within the unit cell.

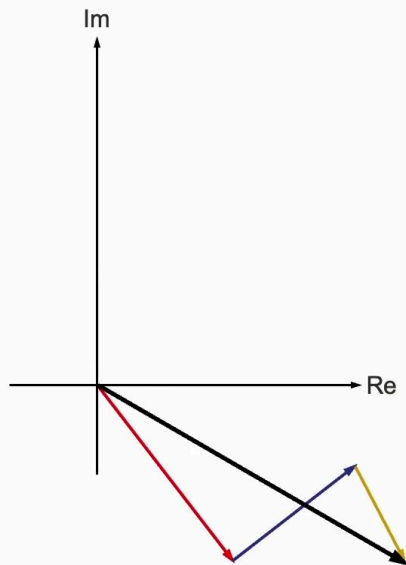
Notes

Summary



0m 27s

Constructing the structure factor, F



- F = vector sum of atomic form factors

$$F_{hkl} = \sum_j f_j \exp \left[-i2\pi \underbrace{(hx_j + ky_j + lz_j)}_{\phi_j} \right]$$

- Intensity of Bragg peak @ (hkl) is proportional to $|F_{hkl}|^2$
- All phase information ϕ_j is lost!
- All we know is $|F_{hkl}|$ and f_j

To recap the statement we briefly made about the phase problem in the last video, we can work out the magnitudes of the structure factors F_{hkl} from the Bragg peak intensities as the latter are proportional to the square of the former. What we don't know, however, is the orientation of F_{hkl} , or indeed, the orientation of the atomic form factors that go together to produce the structure factor. An obvious but strong constraint in solving the phase problem is that a given atomic configuration that produces the right magnitude for the structure factor from the set of atomic form factors for one Bragg peak, must also be consistent with the intensities of all the other measured Bragg peaks as well.

Notes

Summary

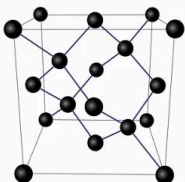


0m 54s

Early successes to solving the phase problem



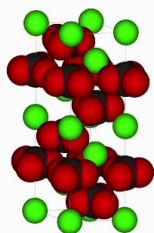
Rock salt



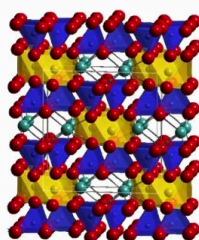
Diamond

Trial-and-error
structural models
W.L. Bragg

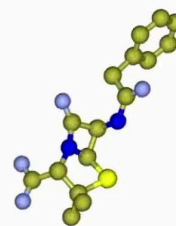
Symmetry and
physicochemical arguments,
few parameters



Calcite

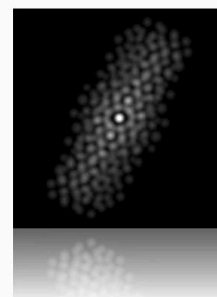


Diopside



Fourier analysis
Patterson maps

Penicillin-G



[M.F. Perutz *Acta Crystallogr. A* 46 633 \(1990\)](#)

Early successes in determining the configurations of relatively simple structures were achieved through simple physicochemical arguments and symmetry considerations, notably in the pioneering work of Lawrence Bragg and his research group. This is beautifully summarized in the article by Max Perutz, one of Bragg's most luminary students, given here in the link. By using so-called Patterson functions and Fourier analysis, more and more complex systems could be solved. These approaches were trailblazed by Dorothy Hodgkin in her seminal studies of biomolecular structures such as penicillin, insulin, and vitamin B12.

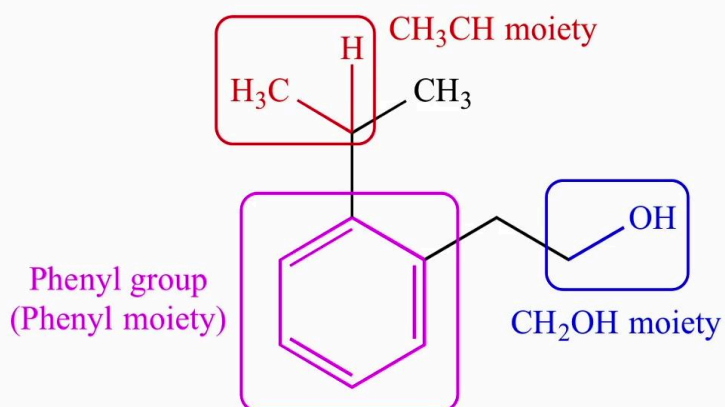
Notes

Summary



1m 57s

Approaches to solving the phase problem



Chemistry

- Atomic/covalent/ionic radii
- Bond lengths
- Chemical affinity
- Electronegativity/electrostatics
- Known substructures
 - "moieties"
 - Functional groups

...

Courtesy University of California at Los Angeles: <http://www.chem.ucla.edu/~harding/IGOC/M/moiety.html>

Simple chemical arguments can be used to set constraints on atomic structures, such as the known radii of atoms involved in a given bond type and the bond length with chemical partners. Other chemical properties, such as chemical affinity and electronegativity can also be applied, while in some cases, known substructures or moieties can be assumed to remain rigid within a unit cell, such as benzene rings, for example.

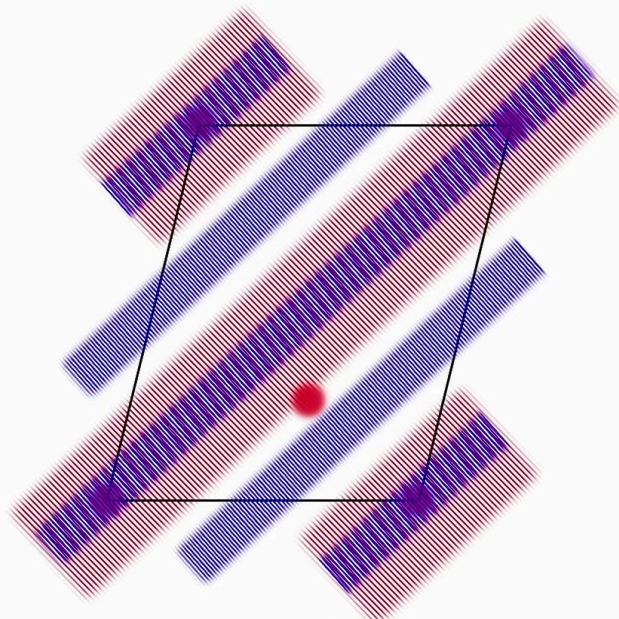
Notes

Summary



2m 54s

Direct methods



https://www.iucr.org/_data/assets/pdf_file/0018/14544/17.pdf

- Consider phases of Fourier components
 ϕ_{hkl}
- Exploit constraints
 - Atomicity – e-density piles up in small volumes associated with atomic nuclei
 - Electron density always ≥ 0
- Statistical correlations, based on probabilities
- Suitable for u.c.s with \leq few 100 atoms

A highly successful approach to solving the phase problem is that of so-called direct methods. The electron scattering density resulting in the diffraction maxima is lumpy due to the existence of atoms. Moreover, there is no such thing as a negative electron density. The consequence is that the phases can't be completely random, but must have well-defined relationships that lead to the set of structure factors, in other words, the diffraction pattern. Direct methods exploit these constraints by attempting to guess the phases directly using the observed amplitudes as a guide, hence their name. Separation between lumps is typically between one and two angstroms, hence structure factors associated with electron waves equal to or smaller than the minimum physically reasonable atomic separation must be recorded in order to obtain the necessary information. In the example shown here for the electron waves associated with the Bragg peak hkl in a given direction shown in red, and that for the Bragg peak $2h2k2l$ shown here in blue, if both of these peaks are strong, their relative phase is likely to be zero. It's beyond the scope of this course to go into the details of direct methods, though the interested reader should read through the didactic review article given here.

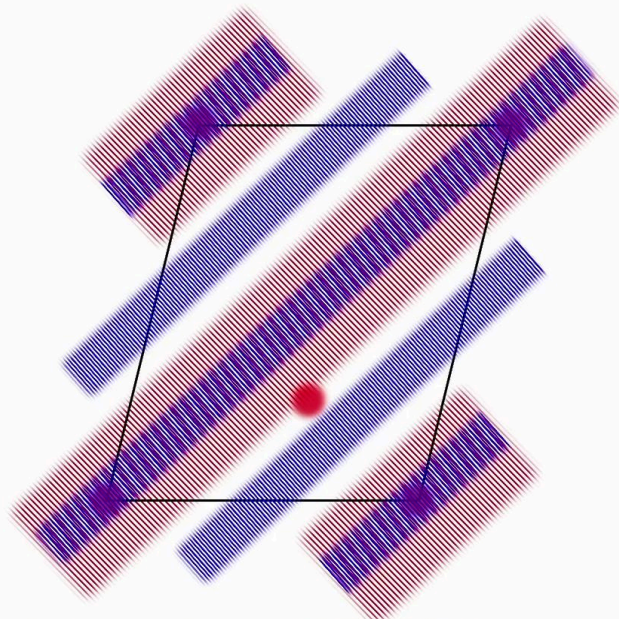
Notes

Summary



3m 31s

Direct methods



https://www.iucr.org/_data/assets/pdf_file/0018/14544/17.pdf

- Consider phases of Fourier components
 ϕ_{hkl}
- Exploit constraints
 - Atomicity – e-density piles up in small volumes associated with atomic nuclei
 - Electron density always ≥ 0
- Statistical correlations, based on probabilities
- Suitable for u.c.s with \leq few 100 atoms

Direct methods are the most common approach to solving the structures of unit cells with volumes of approximately 1,000 cubic angstroms or smaller. For larger structures typically found in macromolecular crystals, other approaches are required. We cover these in the fourth and fifth videos of this section.

Notes

Summary



5m 19s

Iterative methods

$$\{|F_{hkl}|\}_{\text{meas}}$$

- Flipflop between real- and reciprocal space via FT and IFT
- Physical constraints in real space
 - Positivity
 - Atomicity
 - u.c. size
 - ...
- Use phases generated by FT when go from real space to reciprocal space $\{\phi_{hkl}\}$
- Use magnitudes of measured structure factors in reciprocal space $\{I_{hkl}^{1/2}\}$
- Iterate until self-consistent

Iterative methods are becoming more popular as their algorithms become more sophisticated and modern computer power means that solutions can swiftly be found. These techniques flip-flop between real and reciprocal space, and in each of these, changes are made according to physical constraints. In general terms, iterative techniques proceed as follows. The measured set of structure factor amplitudes provided by the square roots of the measured Bragg peak intensities are assigned an entirely random set of phases. Note that this is exactly what we try not to do with direct methods, but let's forget this and continue anyway.

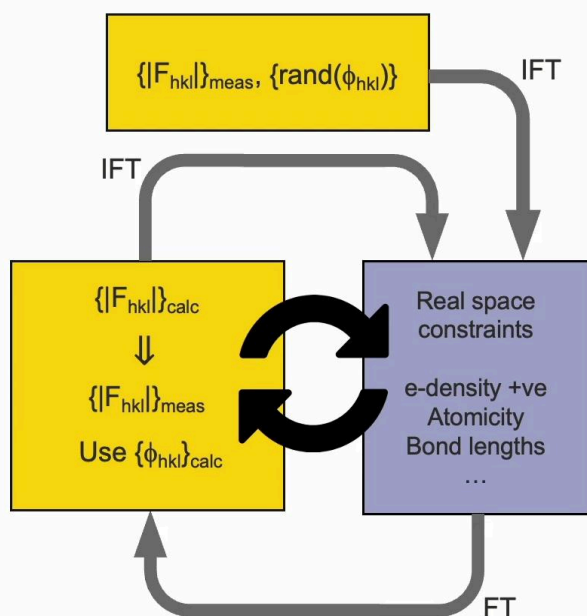
Notes

Summary



5m 42s

Iterative methods



- Flipflop between real- and reciprocal space via FT and IFT
- Physical constraints in real space
 - Positivity
 - Atomicity
 - u.c. size
 - ...
- Use phases generated by FT when go from real space to reciprocal space $\{\phi_{hkl}\}$
- Use magnitudes of measured structure factors in reciprocal space $\{I_{hkl}^{1/2}\}$
- Iterate until self-consistent

Using these amplitude and phase sets, an inverse Fourier transform is performed to produce a real-space structure. Now, because we assigned a random set of phases, it's extremely unlikely that we obtained a physically reasonable real-space model. We can modify this first guess by imposing some physically reasonable constraints, such as, for example, the need to have only positive electron density; atomicity, which requires the lumpiness, which we already talked about with reasonable separated lump maxima; the size of the unit cell; and so on. After this massaging of the real space model, we perform a Fourier transform, which generates a new set of structure factor amplitudes and phases. Now here, in reciprocal space, we impose the constraint that the amplitudes need to equal the measured amplitudes, or in other words, the square roots of the measured intensities. Hence, we replace the set generated by the Fourier transform with the measured values. We then perform once more an inverse Fourier transform and continue in this loop, iterating each time the real space model and the reciprocal space amplitudes until we achieve a self-consistent model in which the real and reciprocal space data agree with one another.

Notes

Summary



In the next video...



In the next video, we take a short time-out in order to discuss anomalous diffraction, that is, the change in the structure factor as a consequence of the photon energy being close to an absorption edge. This has been the basis of many phasing techniques in macromolecular crystallography, in which the size of the structures and the large number of atoms involved precludes the use of direct methods or iterative techniques. Note that I just said, "This has been the basis," as since the end of 2020, a revolutionary new approach based on artificial intelligence has more or less solved the phase problem for macromolecular structures. Hence, in the two videos after the next, we take a historical trip about phasing of macromolecules. Although these approaches have, to a large extent, now become redundant, they do remain important in order to understand in other areas of crystallography and indeed, the newest approach of using artificial intelligence could only have succeeded thanks to the huge database accumulated over the last half-century using these older phasing techniques. In the last video, the present front-runner in AI of macromolecular structure determination, AlphaFold 2, will be briefly summarized.

Notes

Summary



8m 03s