Institut de Biologie Structurale

Principles of X-ray scattering by macromolécules

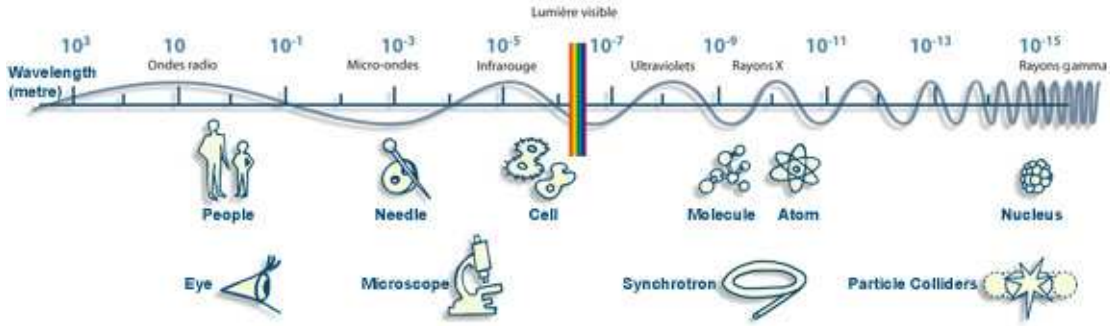
Ecole RéNaFoBiS, Oléron, juin 2014

Parcours

- Physicien (DEA en physique fondamentale en 1987)
- Thèse sur la dynamique moléculaire par simulation sur une protéine, l'utéroglobine (1990)
- Post-doc (Service Militaire) au National Cancer Institute (Maryland): cristallographie sur le BPTI et une asparaginase utilisée en cancérothérapie
- Chercheur CEA à l'Institut de Biologie structurale
  - ➢ Cristallographie sur différentes protéines
    - •Toxine de scorpion et de serpent => résolution atomique
    - •Deshydrogénase
    - •Ferredoxine
    - •GTPases
    - •Récepteur pour l'antigène du lymphocyte T
  - ➢ Immunologie structurale
    - •Interaction, reconnaissance, lien entre interaction à l'echelle moléculaire et activation au niveau cellulaire

How can we “see” macromolecules?

- Based on interaction between radiations and matter
- The wavelength of the radiation used should be suited to the size of the smallest details we want to observe:
  - ➢ Resolution limit in diffraction experiments is λ/2
- Proteins and macromolecules assemblies are too small (1 to 50 nm, 10 to 500 Å) to be observed with visible light (300 nm < □ < 800 nm)
- To see atoms, a wavelength of about 1 Å is required



Three types of “radiation” commonly used

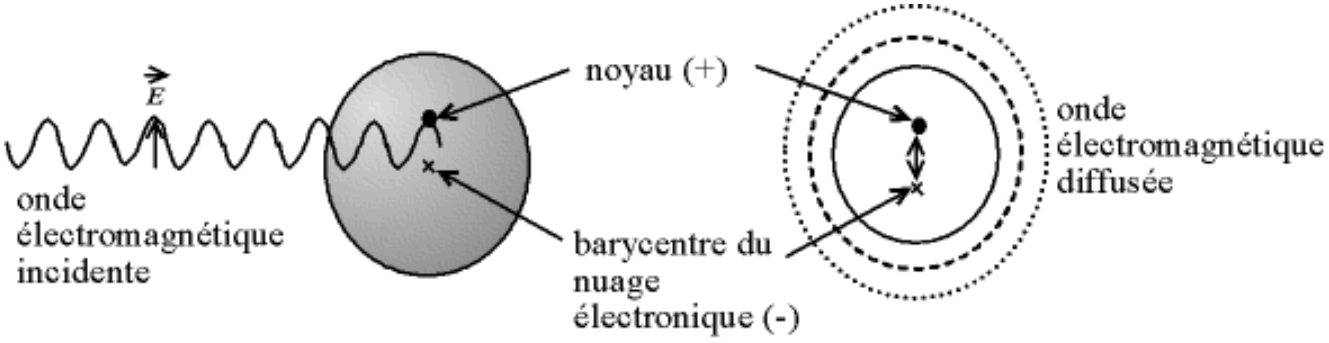
- X-rays
  - ➢ Electromagnetic waves / photons
    - •Photon energy w = hc/□ = h□ with h = 6.6257 10^{-34} J.s (Plank constant)
    - •7 keV < w < 17 keV or 1.7 Å > λ > 0.7 Å
- Neutrons
  - ➢ Neutral particle (m_n = 1.6749 10^{-27} kg)
    - •Energy w = 1/2$m_n v^2$, wavelength λ = h / m_nv
    - •λ = 1.5 Å for v = 2600 m/s and w = 3.6 10^{-2} eV
- Electrons
  - ➢ Negatively charged particle (q = 1.6 10^{-19} C, m_e = 9.1091 10^{-31} kg)
    - •Energy w = 1/2$m_e v^2$, wavelength λ = h / m_ev
    - •λ = 1.2 Å for v = 6000 km/s and w = 100 eV
- Here, we will focus only on X-rays!

X-rays interact with electrons in atoms

- Elastic scattering (no loss of energy, wavelength is conserved)
  - ➢Rayleigh scattering: bound electron (photon energy << electron binding energy)
  - ➢Thomson scattering: free electrons (photon energy >> electron binding energy)
    - •Carbon atom E(1s) = -1013 eV, E(2s,2p) = -36 eV to be compared to 7 to 15 keV for X-ray photons
    - •Photon energy should differ from element absoption edges
- Inelastic scattering (loss of energy, scattered photon have a larger wavelength)
  - ➢Compton scattering (a bound electron is ejected in the continuum)
  - ➢Raman effect (a bound electron jump to a free state of the atom or molecule)
- Ionisation

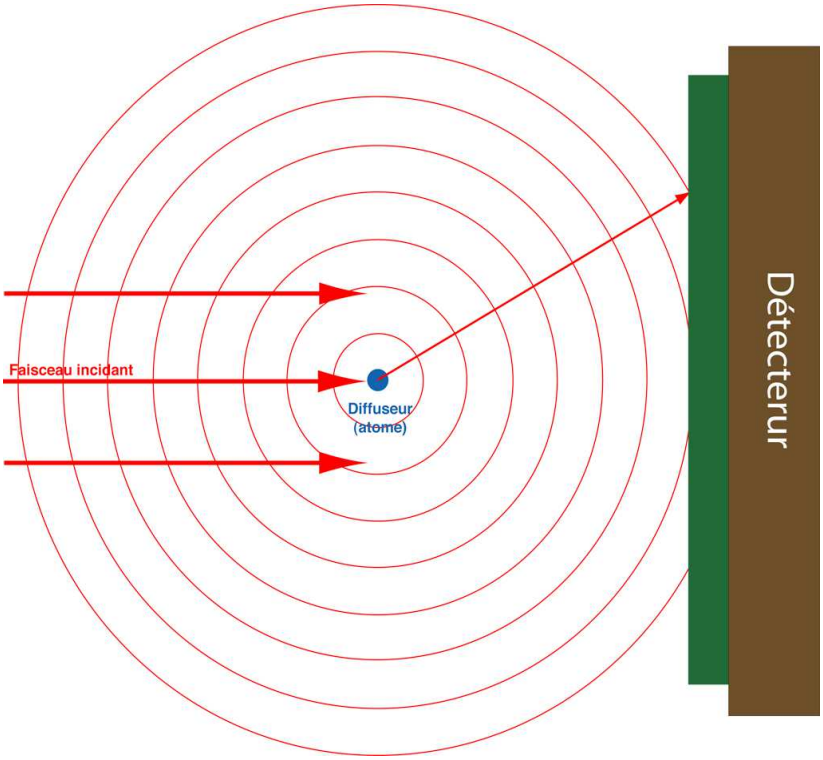
How a photon is scattered by an atom ?

- Classical model
  - ➢The photon is scattered by the electronic cloud of the atom
  - ➢The incident electromagnetic wave induce the oscillation of the electronic cloud
  - ➢The negatively charged electronic cloud et the positively charged nucleus become an oscillating dipole, thus emitting a spherical electromagnetic wave of same wavelength and a phase difference of π.



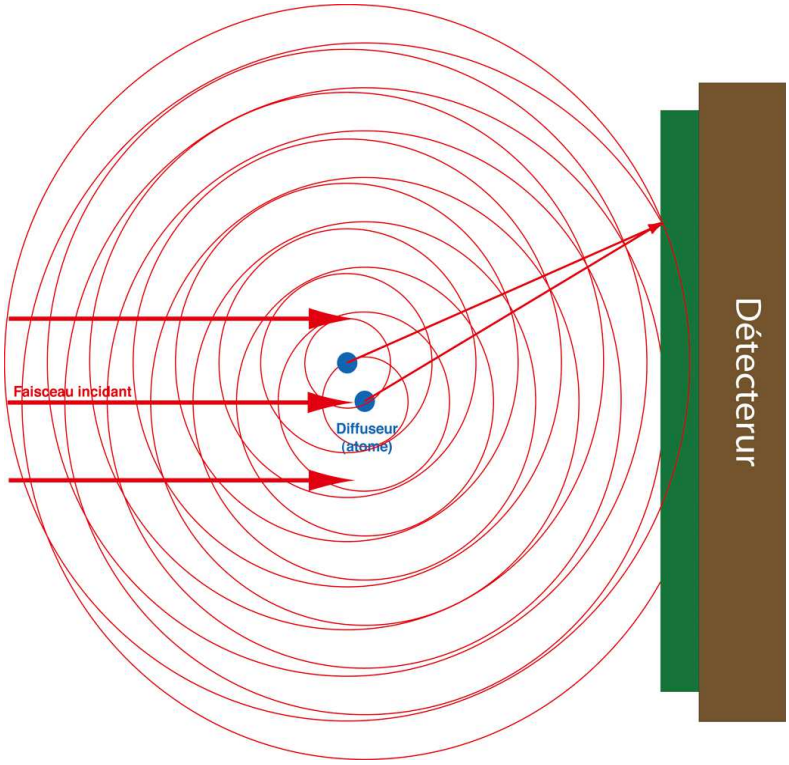
Scattering by one atom

- The planar incident wave induces the emission of a spherical wave of same wavelength and with a π phase shift



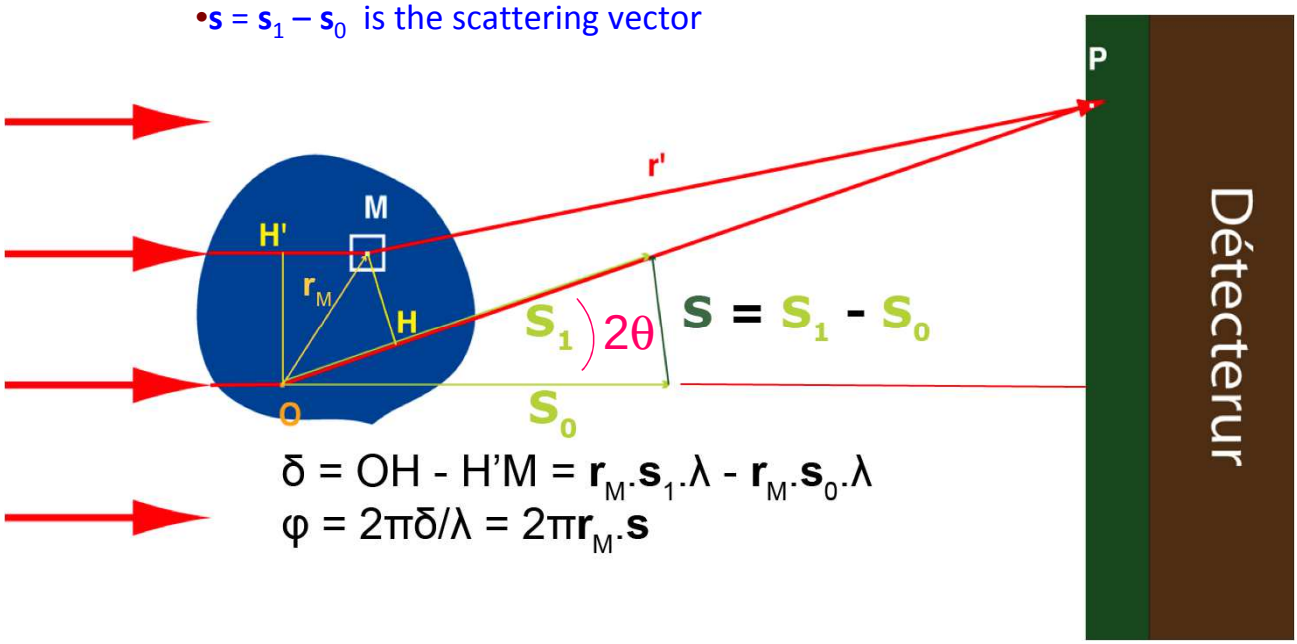
Scattering by two atoms

- The two atoms are emitting a spherical electromagnetic wave. If the two atoms are not mobile and close by, the two waves interfere



Scattering by a solid sample

- What do we measure on the detector ?
  - •Incident wave described by its wave vector $\mathbf{q}_0$ ($|\mathbf{q}_0| = 2\pi/\lambda$) or $\mathbf{s}_0 = \mathbf{q}_0 / 2\pi$ ($|\mathbf{s}_0| = 1/\lambda$) : $E_0 \cos[\, \omega t - \mathbf{q}_0.\mathbf{r}\,]$
  - •Scattered wave described by its wave vector $\mathbf{q}_1$ ($|\mathbf{q}_1| = 2\pi/\lambda$) or $\mathbf{s}_1 = \mathbf{q}_1 / 2\pi$ ($|\mathbf{s}_1| = 1/\lambda$)
  - •$\mathbf{s} = \mathbf{s}_1 - \mathbf{s}_0$ is the scattering vector



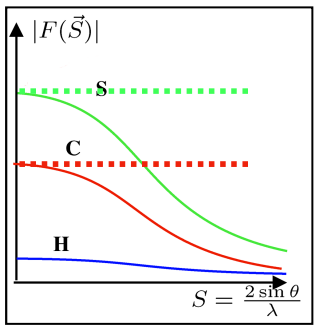
Relation between scattered wave and solid sample?

- Incident wave : $E_0 \exp[\, i(\omega t - \mathbf{q}_0.\mathbf{r})\,]$
- $f(\mathbf{r}_M)$ is the scattering power at point M defined by $\mathbf{r}_M$
  - ➢The wave scattered by the small volume ∂v around M, as observed in P is:
    - •$\partial E(\mathbf{s}) = E_0/r'\, f(\mathbf{r}_M) \exp[\, i(\omega t - \mathbf{q}_0.\mathbf{r}_M)\,].\exp(-i\mathbf{q}_1.(\mathbf{r}-\mathbf{r}_M)).\exp(i\pi)\,\partial\mathbf{r}_M$
    - • $= E_0/r'\, f(\mathbf{r}_M) \exp(i(\omega t - \mathbf{q}_1.\mathbf{r})).\exp(2i\pi\,\mathbf{r}_M.\mathbf{s}).\exp(i\pi)\,\partial\mathbf{r}_M$
- To calculate the wave scattered by the entire sample, we have to sum on the entire volume of the sample:
  - ➢$E(\mathbf{s}) = \int_v E_0/r'\, f(\mathbf{r}_M) \exp(i(\omega t - \mathbf{q}_1.\mathbf{r})).\exp(2i\pi\,\mathbf{r}_M.\mathbf{s}).\exp(i\pi)\,\partial\mathbf{r}_M$
  - ➢ $= E_0/r' \exp(i(\omega t - \mathbf{q}_1.\mathbf{r})) \exp(i\pi) \int_v f(\mathbf{r}_M) \exp(2i\pi\,\mathbf{r}_M.\mathbf{s})\,\partial\mathbf{r}_M$
  - ➢E(**s**) is proportional to the Fourier transform of the scattering power of the sample
- For X-rays, this scattering power is proportional to the electron density
  - ➢E(**s**) is proportional to the Fourier transform of the electron density

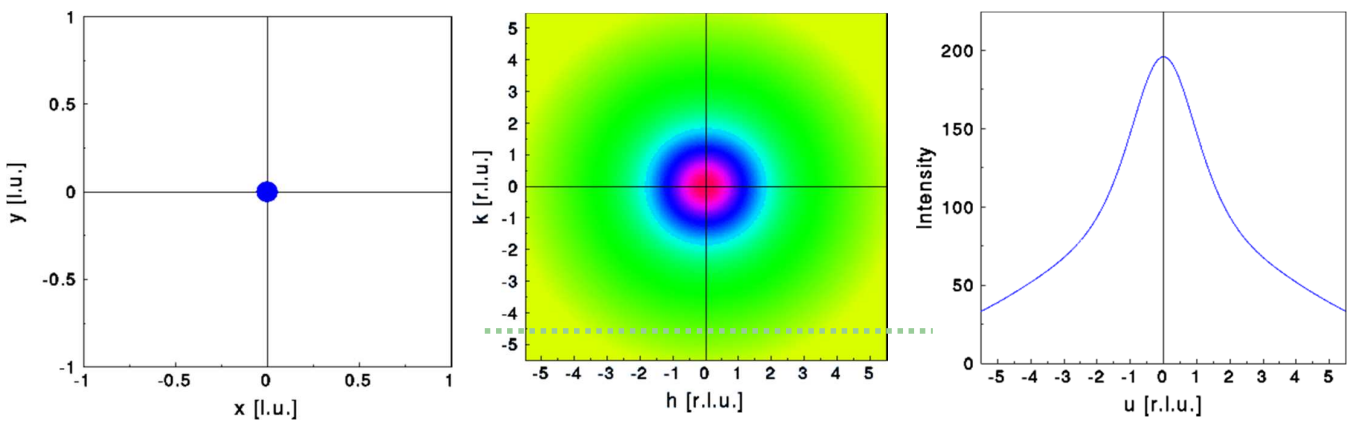
Structure factor

- $E(\mathbf{s}) = E_0/r' \sigma \int_v \rho(\mathbf{r}) \exp(2i\pi\, \mathbf{r}.\mathbf{s})\, \partial\mathbf{r}$
- Structure factor:
  - Direct Fourier transform of the electron density
  - $F(\mathbf{s}) = \int_v \rho(\mathbf{r}) \exp(2i\pi\, \mathbf{r}.\mathbf{s})\, \partial\mathbf{r}$
  - Complex number (amplitude and phase)
- Electron density
  - Indirect Fourier transform of the structure factors
  - $\rho(\mathbf{r}) = \int_v F(\mathbf{s}) \exp(-2i\pi\, \mathbf{r}.\mathbf{s})\, \partial\mathbf{s}$
- The detector measures the intensity of the scattered wave
- The intensity of the scattered wave is proportional to the square of the structure factor modulus
  - $I(\mathbf{s}) \alpha F(\mathbf{s}).F^*(\mathbf{s}) = |F(\mathbf{s})|^2$

A few exemples …

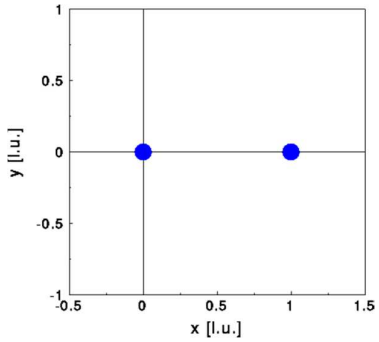


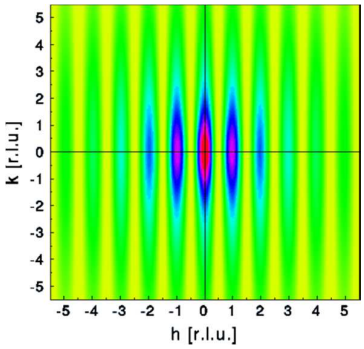
- Diffusion spectra of one atom
  - Spherical electron density
  - $f_{at}(s) = 2 \int_{r=0,+\infty} \rho_{at}(r)\, r\, \sin(2\pi\, sr) / s\, \partial r$
  - $f_{at}(0)$ = number of electrons in atom
  - $f_{at}(s)$ is constant when atom is ponctual (neutron diffraction)
  - $f_{at}(s)$ is the atomic scattering factor

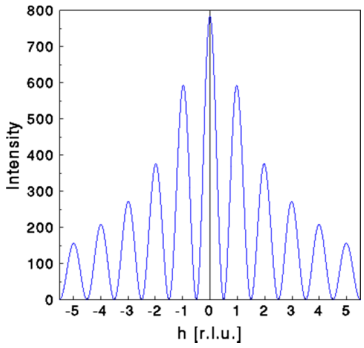


More atoms …

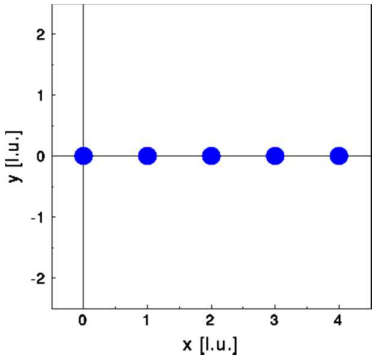
- Diffusion by two atoms

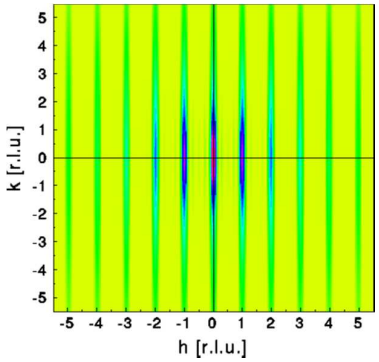


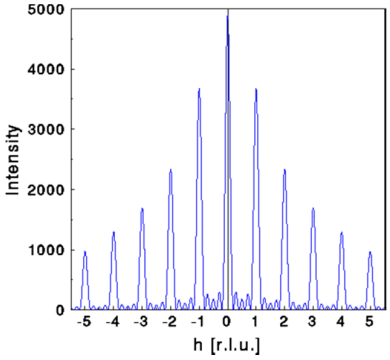




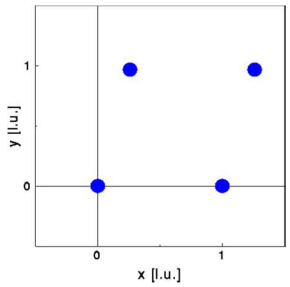
- Diffusion by five atoms

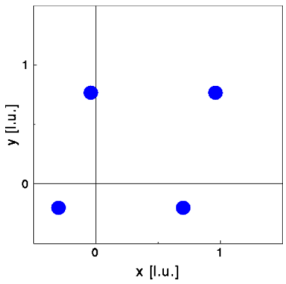


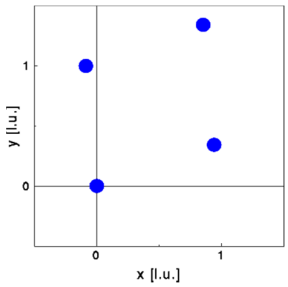




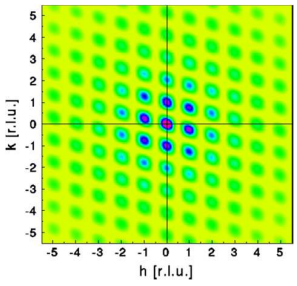
- Four atoms: effect of translation and rotation of the set of atoms

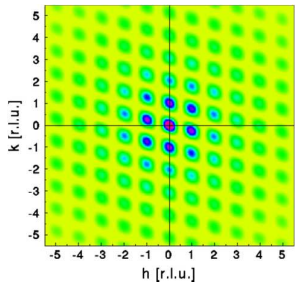


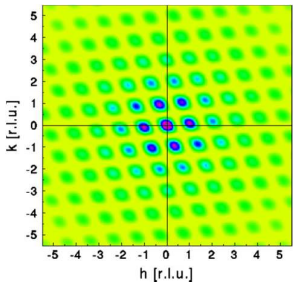




- No change upon translation
- Diffusion spectra is rotated when the set of atoms are rotated







Structure factor of an assembly of atoms

- $f_j(s)$ is the scattering factor of the atom j

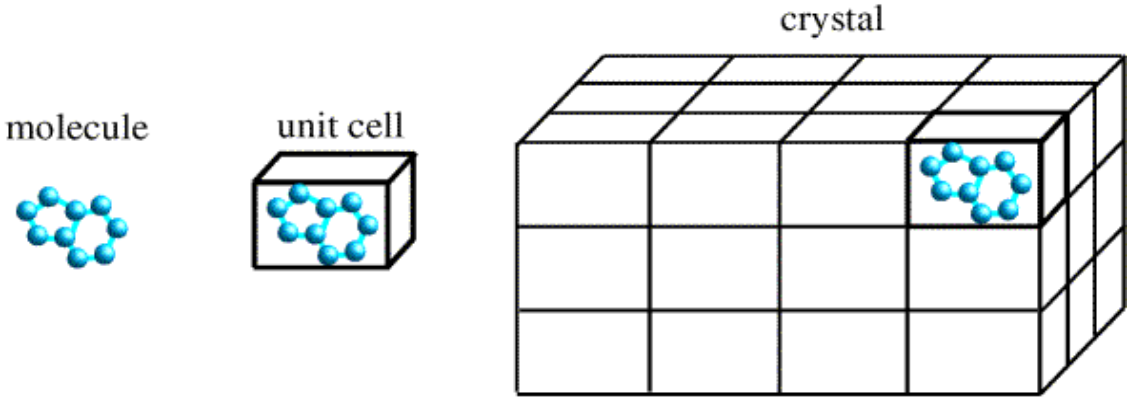
- The phase shift of the wave scattered by an atom in position $\mathbf{r}_j$, in the direction defined by scattering vector **s** is (relatively to an atom at the origin):
  - $2\pi\mathbf{r}_j\mathbf{s}$

- The structure for an assembly a N atoms is thus:
  - $F(\mathbf{s}) = \sum_{j=1,N} f_j(s).\exp(2i\pi\mathbf{r}_j\mathbf{s})$

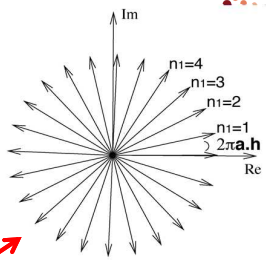
- $f_j(s)$ are known and tabulated for each element
  - Calculating the structure factor of an assembly of atoms is straightforward

A specific case of solid sample: the crystal

- The crystal is described by a lattice which characterize the three-dimensional periodicity
- The lattice is define by three vector, **a**, **b**, **c**, that define the unit cell
- The content of each unit cell is identical



Structure factor of a crystal



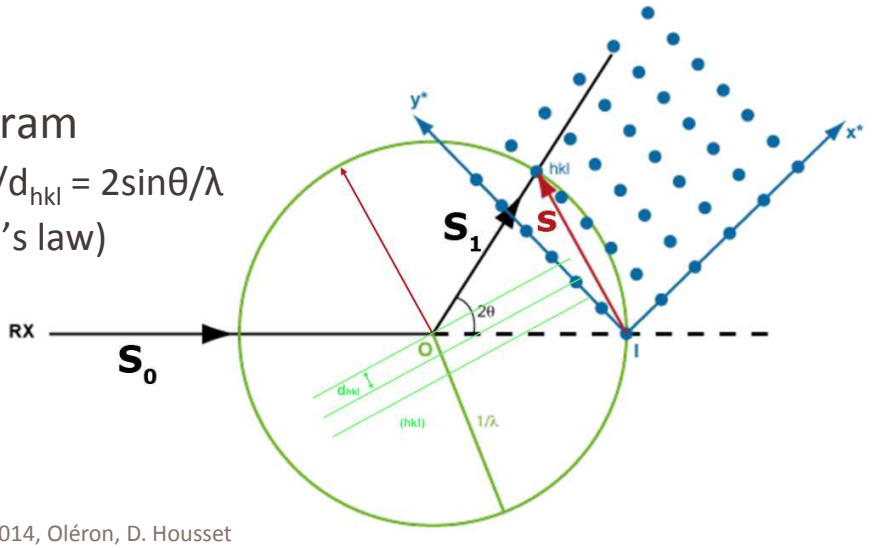
- The structure factor is:
  - $F(\mathbf{s}) = \int_{vech.} \rho(\mathbf{r}) \exp(i\, 2\pi \mathbf{s}.\mathbf{r})\, \partial\mathbf{r}$
  - $F(\mathbf{s}) = \sum_n \int_{vmaille} \rho(\mathbf{r}+\mathbf{r}_n) \exp(i\, 2\pi\, \mathbf{s}.(\mathbf{r}+\mathbf{r}_n))\, \partial\mathbf{r}$
  - $F(\mathbf{s}) = \sum_n \exp(i\, 2\pi\, \mathbf{s}.\mathbf{r}_n) \int_{vmaille} \rho(\mathbf{r}) \exp(i\, 2\pi\, \mathbf{s}.\mathbf{r})\, \partial\mathbf{r}$
  - $F(\mathbf{s}) = \sum_n \exp(2i\pi\, n_1 \mathbf{a}.\mathbf{s}) \exp(2i\pi\, n_2 \mathbf{b}.\mathbf{s}) \exp(2i\pi\, n_3 \mathbf{c}.\mathbf{s}) \int_{vmaille} \rho(\mathbf{r}) \exp(i\, 2\pi\, \mathbf{s}.\mathbf{r})\, \partial\mathbf{r}$

- This structure factor close to zero, except for specific values of **s** :
  - **a**.**s** = h, **b**.**s** = k, **c**.**s** = l where h, k, l are integers (Laue equations)
  - Diffraction spots instead of diffusion spectra
  - Discrete structure factors

- In directions that satisfy Laue equation, the structure factor is the one of the unit cell, multiplied by the number of cells in the crystal.
- The crystal is a signal amplifier!

Reciprocal lattice

- If **s** satisfy the Laue equation:
  - **a**.**s** = h, **b**.**s** = k, **c**.**s** = l where h, k, l are integers
- Then **s** is a vector of a lattice named reciprocal lattice
  - **s** = h.**a*** + k.**b*** + l.**c*** where
    - $\mathbf{a}^* = \mathbf{b} \wedge \mathbf{c} / V_{maille}$, $\mathbf{b}^* = \mathbf{c} \wedge \mathbf{a} / V_{maille}$, $\mathbf{c}^* = \mathbf{a} \wedge \mathbf{b} / V_{maille}$
    - $V_{maille} = (\mathbf{a} \wedge \mathbf{b}).\mathbf{c}$
    - Larger unit cell => smaller reciprocal cell => closer spots on diffraction spectra
- Ewald diagram
  - $|\mathbf{s}| = 1/d_{hkl} = 2\sin\theta/\lambda$
  - (Braggs's law)

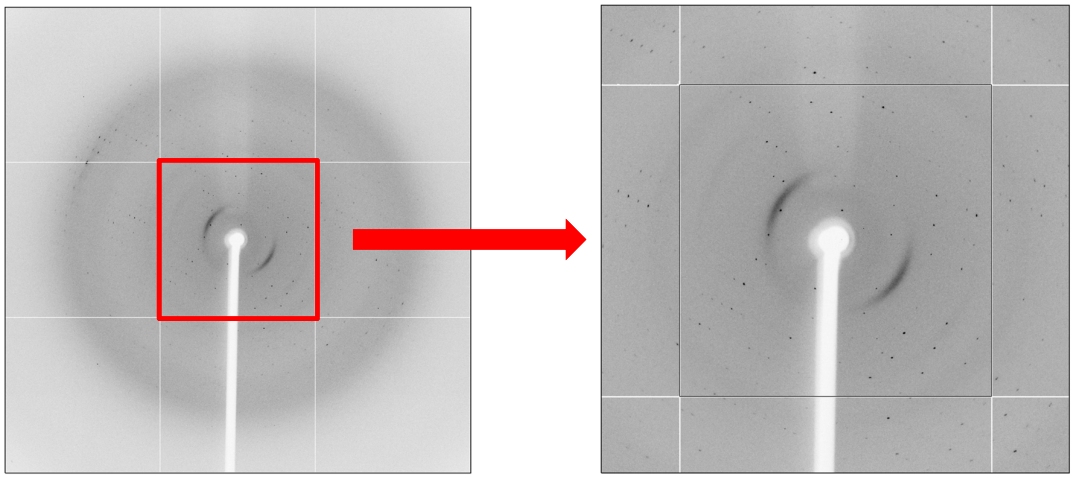


What does represent the structure factor

- A parallel with sound
  - in crystals, ρ(**r**) is a periodic function in the three dimensional space
  - The sound signal is periodic function of time (one dimension)
- F(**s**) are the coefficient of the development of ρ(**r**) in a Fourier serie :
  - $F(\mathbf{s}) = \int_{vmaille} \rho(\mathbf{r}) \exp(i\, 2\pi\, \mathbf{s}.\mathbf{r})\, \partial\mathbf{r}$
  - $\rho(\mathbf{r}) = 1/V_{maille} \sum_{\mathbf{s}} F(\mathbf{s}) \exp(-i\, 2\pi \mathbf{s}.\mathbf{r})$
- The Fourier transform of the sound signal provide the frequencies contained in the sound
  - Include the fundamental frequency and the harmonics (2 times, 3 times, … the fundamental frequency)
- The largest **s** or frecency is, the more detailed is the drescription of ρ(**r**) or of the sound signal

What does the diffraction tell us ?

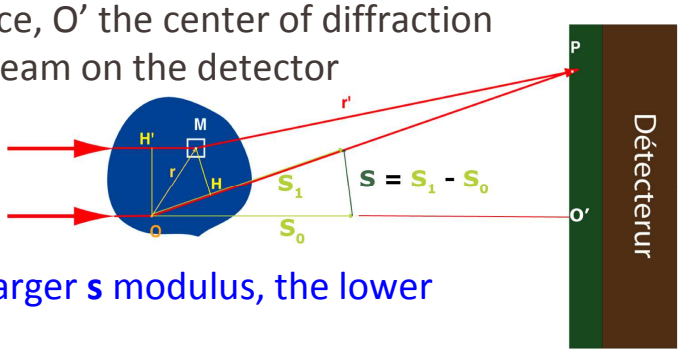
- The position of spots on diffraction images will allow the determination of the unit cell dimension and the orientation of the crystal
- The intensity of each spot contains information on the electron density of the entire cell
  - $I(\mathbf{s}) = F(\mathbf{s}).F^*(\mathbf{s})$
  - $= (\int_{vmaille} \rho(\mathbf{r}) \exp(i\, 2\pi\, \mathbf{s}.\mathbf{r})\, \partial\mathbf{r}) (\int_{vmaille} \rho(\mathbf{r}) \exp(-i\, 2\pi\, \mathbf{s}.\mathbf{r})\, \partial\mathbf{r})$



Rap1 diffraction image Courtesy of Dr. Marie-Hélène Le Du

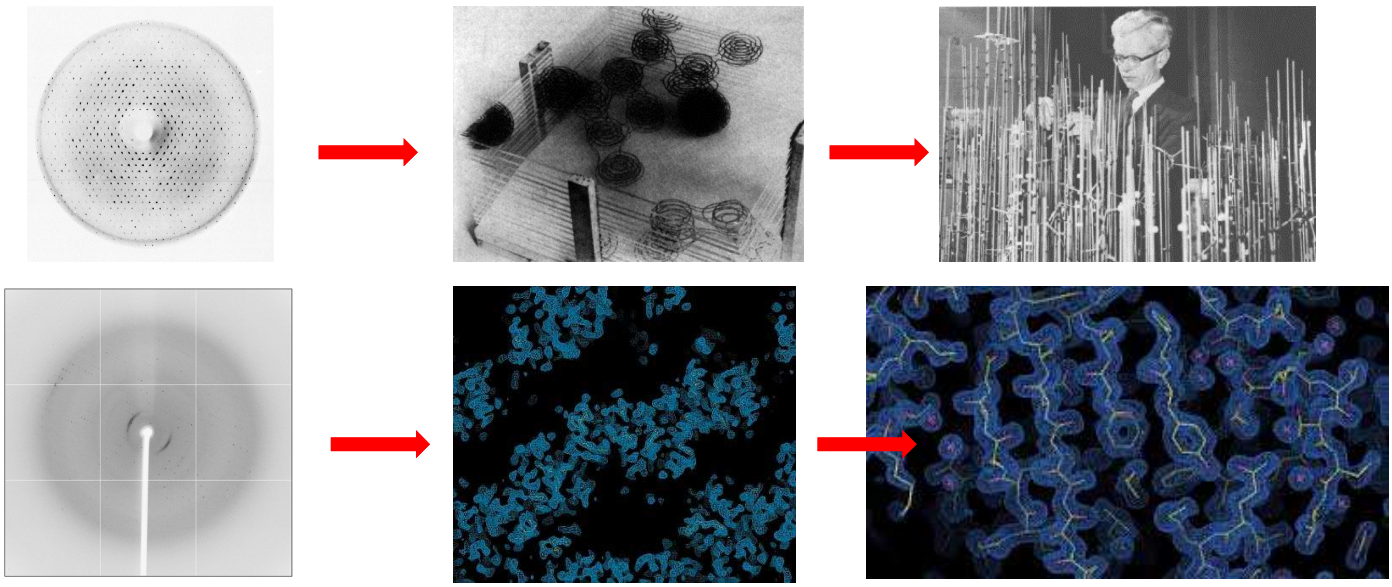
Resolution

- In crystallography
  - d = 1 / |**s**| in Å
- Directly related with the distance from the center (direct beam position) of the diffraction image:
  - If D is the crystal - detector distance, O' the center of diffraction image, P the position of diffracted beam on the detector
  - |O'P| = D sin2θ
  - 1 / d = |**s**| = 2sinθ / λ



  - The further from the center, the larger **s** modulus, the lower value of d (higher resolution) in Å
- Maximum resolution in diffraction experiments
  - 2θ = 180° => θ=90° => d = λ / 2
- Details that can be seen (separated)
  - d/√2 in crystallography. What about SAXS ?

Calculating the electron density ?



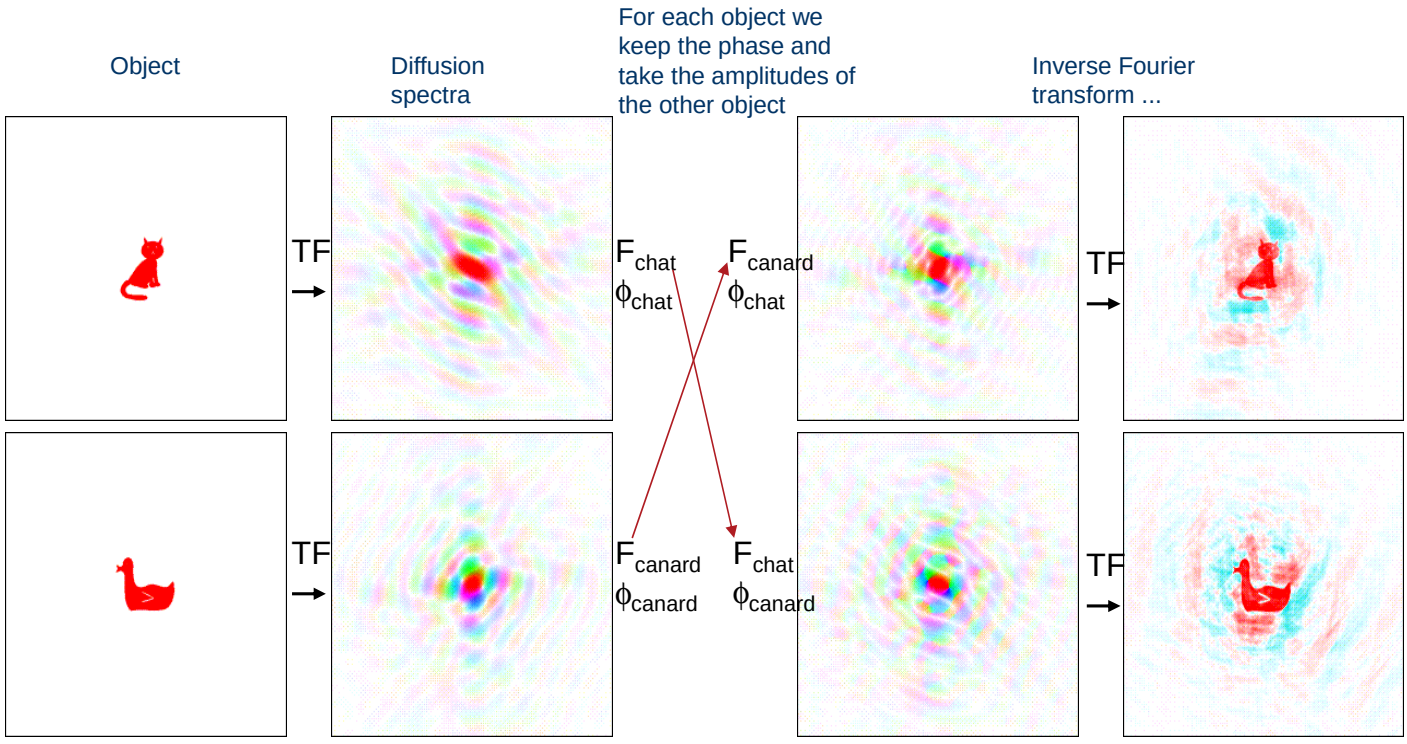
- Simple in principle
  - The electron density within the unit cell can be derived from the structure factors :
    - $\rho(\mathbf{r}) = 1/V_{maille} \sum_{\mathbf{s}} F(\mathbf{s}) \exp(-i\, 2\pi \mathbf{s}.\mathbf{r})$
    - But …

Calculating the electron density

- We do not measure F(**s**) but only an intensity I(**s**) which is proportional to the square of F(**s**) modulus: the phase of the scattered wave is lost
  - ➢ $F(\mathbf{s}) = |F(\mathbf{s})|\exp(i\phi(\mathbf{s}))$
  - ➢ Intensity measurements do not allow the calculation of the electron density for maromolecules

- Methods have been developed to provided a reasonable estimate of the phases $\phi(\mathbf{s})$ and allow the calculation of □(**r**) => Jean-Luc's Talk

Most of the structural information is in the phases



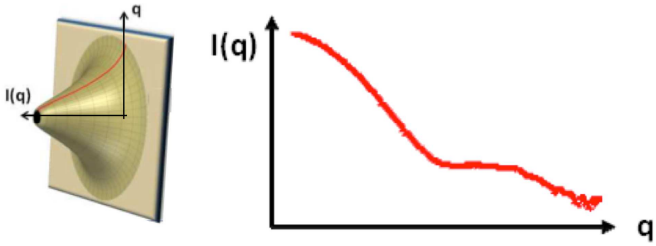
Patterson function

- The Patterson function can be calculated without phase information
  - $P(\mathbf{u}) = 1/V \sum_s F(\mathbf{s})\, F^*(\mathbf{s}) \exp(-2i\pi\, \mathbf{u}.\mathbf{s})$
  - $= 1/V \sum_s |F(\mathbf{s})|^2 \exp(-2i\pi\, \mathbf{u}.\mathbf{s})$
  - It is the Fourier tranform of the squared amplitude of structure factors
- It represents the autocorrelation function of the electron density:
  - $P(\mathbf{u}) = \int_V \rho(\mathbf{r}).\, \rho(\mathbf{r}+\mathbf{u})d\mathbf{r}$
  - P(u) has a local maximum when u is an interatomic vector
- Can be used to solved structures when the number of atoms is small (less than 50) : not suitable for proteins
- Very useful for phasing (MR, Heavy atom derivatives, anomalous signal)

What about macromolecules in solution ?

- Macromolecules move due to thermal motion
  - Each macromolecules adopts all orientations during the measurements
    - Spherical averaging
- The distance between scatterers is not fixed
  - The distance between two equivalent atoms in two distant molecules vary during the measurement
    - => incoherent scattering
    - => no interference (≠ diffraction by a crystal)
    - => Intensities scattered by each molecules adds up
- Scattering of the solvent is substracted
  - $I_{macromolecule}(\mathbf{q}) = I_{sample}(\mathbf{q}) - I_{buffer}(\mathbf{q})$
- Provide information on electron density difference between the sample and the buffer, $\Delta\rho(\mathbf{r})$

- The scattering spectra has a radial symmetry
- One dimensional data due to the spherical averaging



- $I(q) = \langle TF(\Delta\rho(\mathbf{r})) \,.\, TF(\Delta\rho(-\mathbf{r}))\rangle$
- $I(q) = \langle TF(\int_{Vparticle}\Delta\rho(\mathbf{r+u}).\,\Delta\rho(\mathbf{u})d\mathbf{u})\rangle$
- $I(q) = 4\pi \int_{0,inf} p(r).\sin(qr)/qr\, dr$
  - p(r) is the distribution of distances between all pairs of points within the particle weighted by the respective electron densities
- $p(r) = r/2\pi \int I(q).q.\sin(qr)\, dq$
- Some structural characteristics of the macromolecule can be derived from p(r).

Comparison of diffraction and SAXS spectra