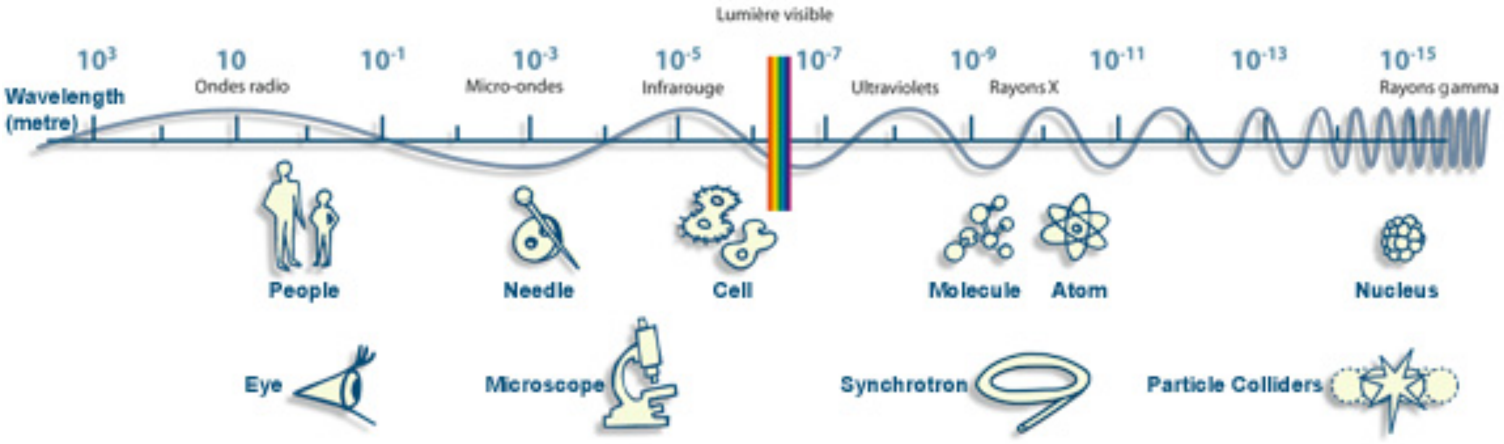
Institut de Biologie Structurale

Principles of X-ray scattering by macromolécules

Ecole RéNaFoBiS, Oléron, juin 2015

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 $\lambda/2$
- Proteins and macromolecules assemblies are too small (1 to 50 nm, 10 to 500 Å) to be observed with visible light (300 nm < λ < 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 $w = hc/\lambda = h\nu$ with $h = 6.6257\ 10^{-34}$ J.s (Plank constant)
    - •$7\ keV < w < 17\ keV$ or $1.7\ Å > \lambda > 0.7\ Å$
- Neutrons
  - ➢Neutral particle ($m_n = 1.6749\ 10^{-27}$ kg)
    - •Energy $w = 1/2m_n v^2$, wavelength $\lambda = h / m_n v$
    - •$\lambda = 1.5\ Å$ 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 $w = 1/2m_e v^2$, wavelength $\lambda = h / m_e v$
    - •$\lambda = 1.2\ Å$ for v = 6000 km/s and $w = 100$ eV

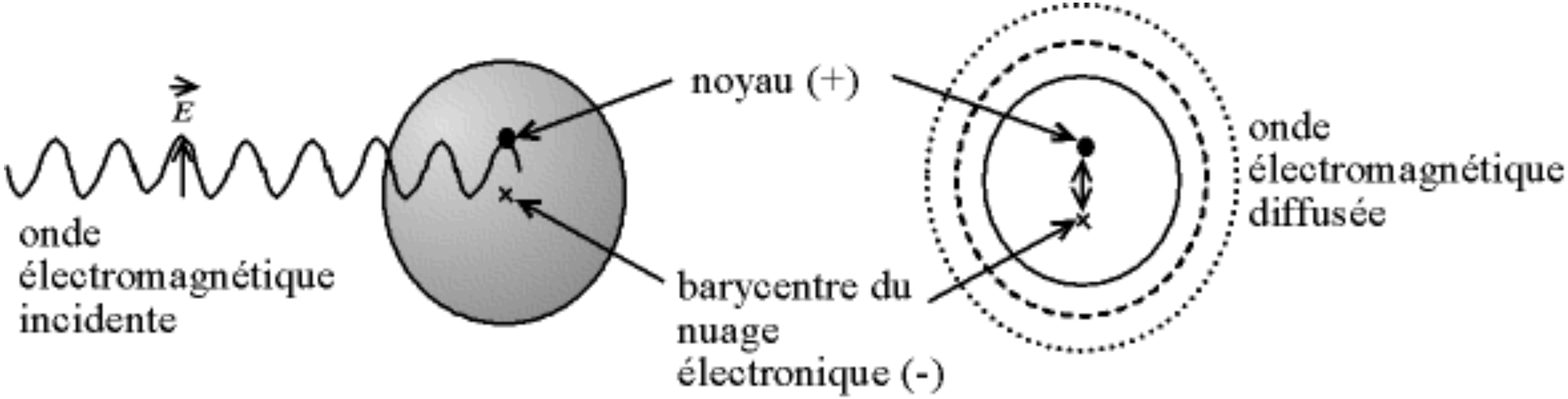
- Here, we will focus only on X-rays!

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 ?

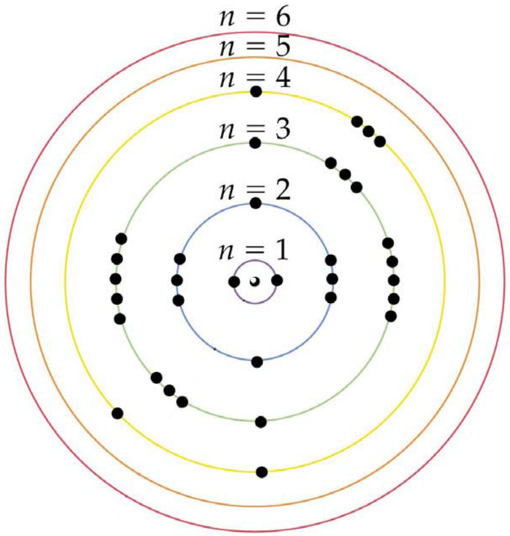
- Classical model
  - ➢The photon is scattered by the electronic cloud of the atom
  - ➢The incident electromagnetic wave induce the oscillation of the electronic cloud
  - ➢The negatively charged electronic cloud et the positively charged nucleus become an oscillating dipole, thus emitting a spherical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

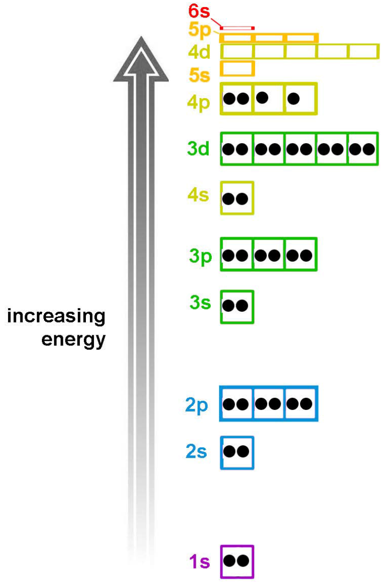


Anomalous scattering

- Hypothesis of 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

- For some elements the energy required to exite a core electron (1S or 2S orbitals) is in the X-ray energy range





http://ocw.mit.edu/courses

	absorption edges (keV/ Å)				
Element	**Edges K**	**Edge L1**	**Edgel L2**	**Edge L3**	**Edgel M**
C	0.28 / 44				
O	0.54 / 23				
N	0.40 / 31				
S	2.47 / 5	0.193	0.164	0.162	0.017
Se	12.66 / 0.98	1.653	1.477	1.436	0.232
Pt	78.39	13.89 / 0.89	13.27 / 0.93	11.56 / 1.07	3.29 / 3.77
W	69.52	12.09 / 1.02	11.54 / 1.07	10.20 / 1.21	2.82 / 4.89
Re	71.67	12.52 / 0.99	11.96 / 1.03	10.53 / 1.17	2.93 / 4.23
Tc	21.04	3.04 / 4.07	2.79 / 4.44	2.68 / 4.62	0.54
U	115.6	21.75	20.95	17.17 / 0.72	5.55 / 2.23

Atomic form factor at the edge

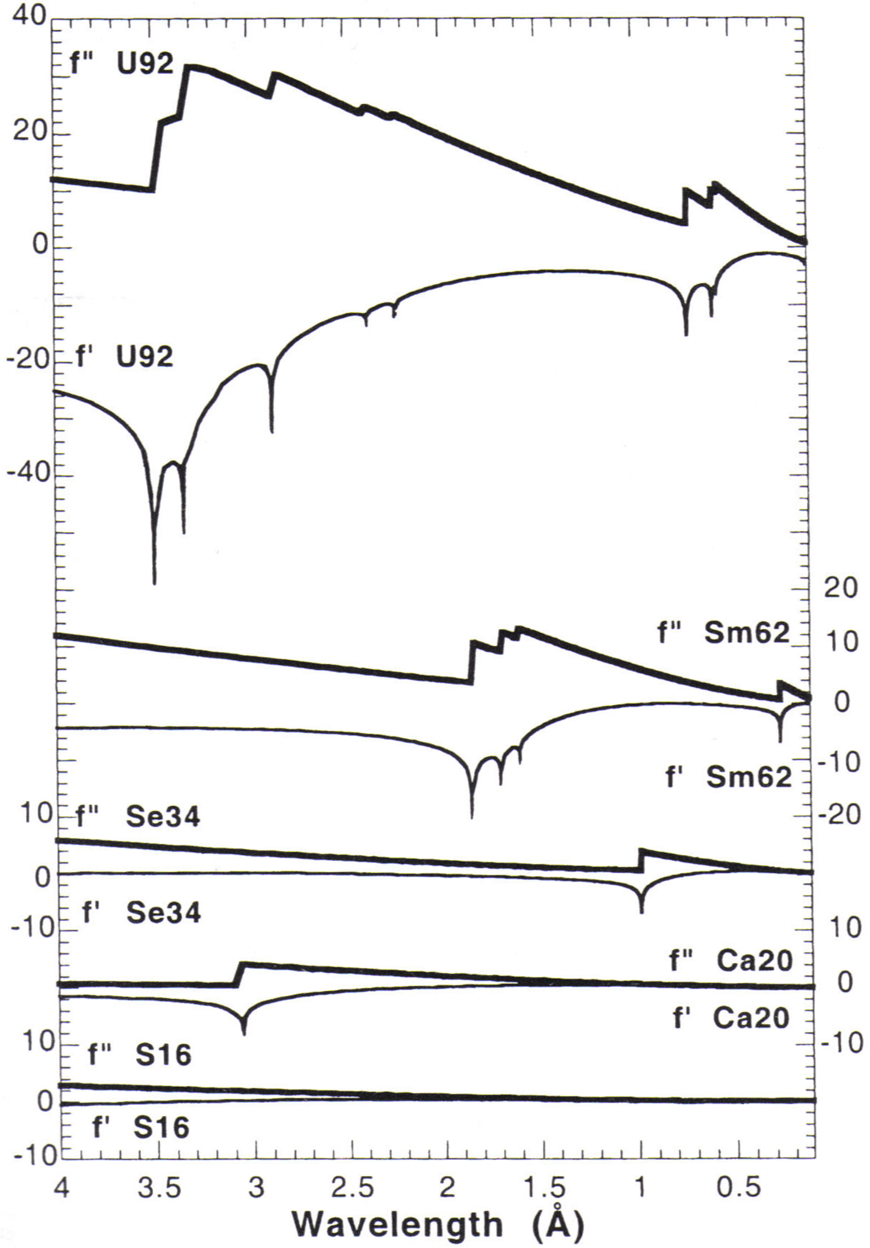
- In the vicinity of the edge
  - ➢Resonance phenomena between the incident wave and the electrons of the atom
  - ➢The phase difference between the incident and the scattered wave is no longer π: change in the phase.
  - ➢The atomic form factor is then:
    - $f_j(\lambda) = {}^0f_j + f_j'(\lambda) + i.f_j''(\lambda)$
    - The $f_j''(\lambda)$ term introduce the change in phase: component with phase shift of $\pi/2$
    - Calculed from the absoption or fluorescence spectra
    - $f_j'(\lambda)$ is calculated by the Kramers-Kroning formula:

$$f'(E) = \frac{2}{\pi}\int_0^\infty \frac{E'.f''(E').dE'}{E^2 - E'^2}$$

- E is the energy

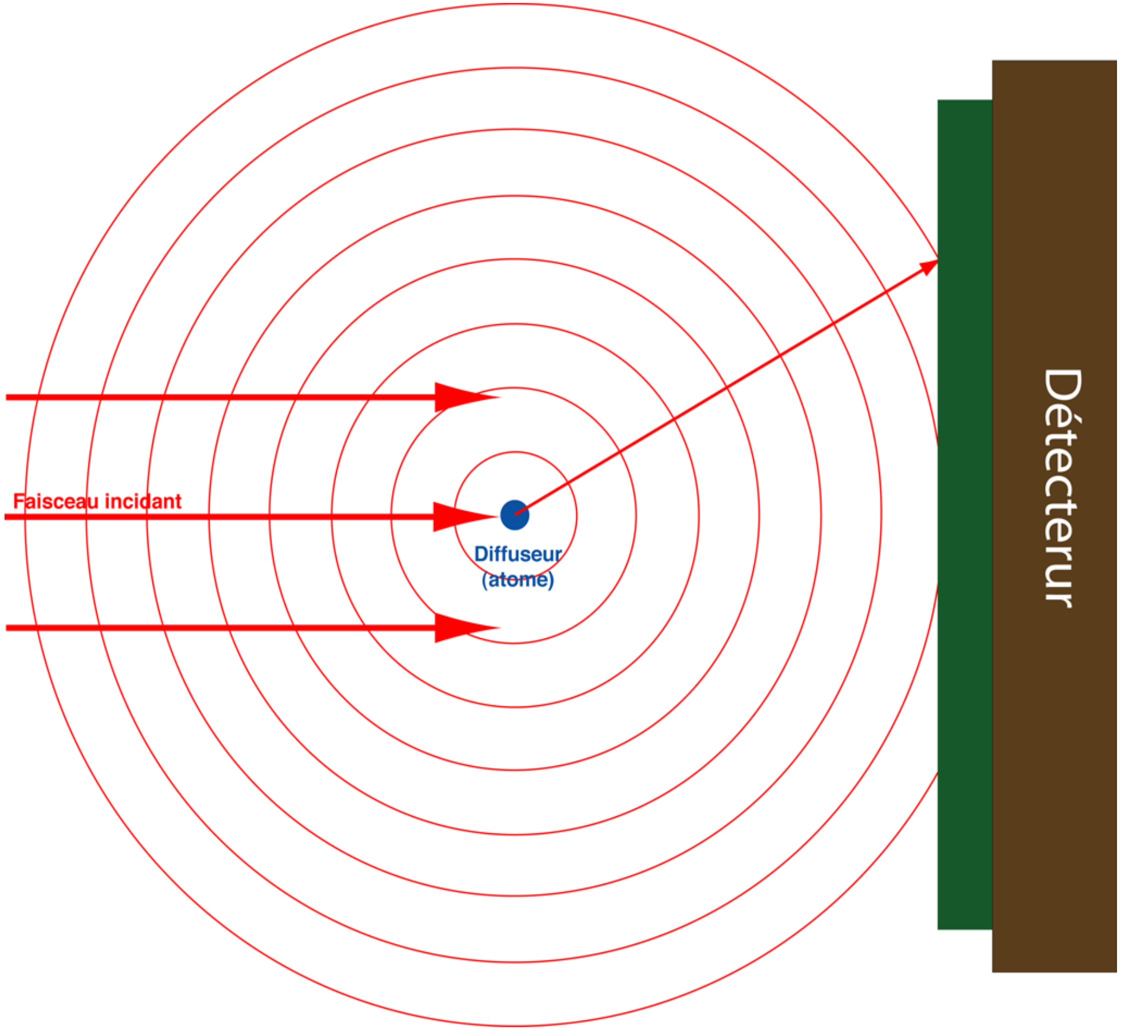
Exemple of absorption spectra

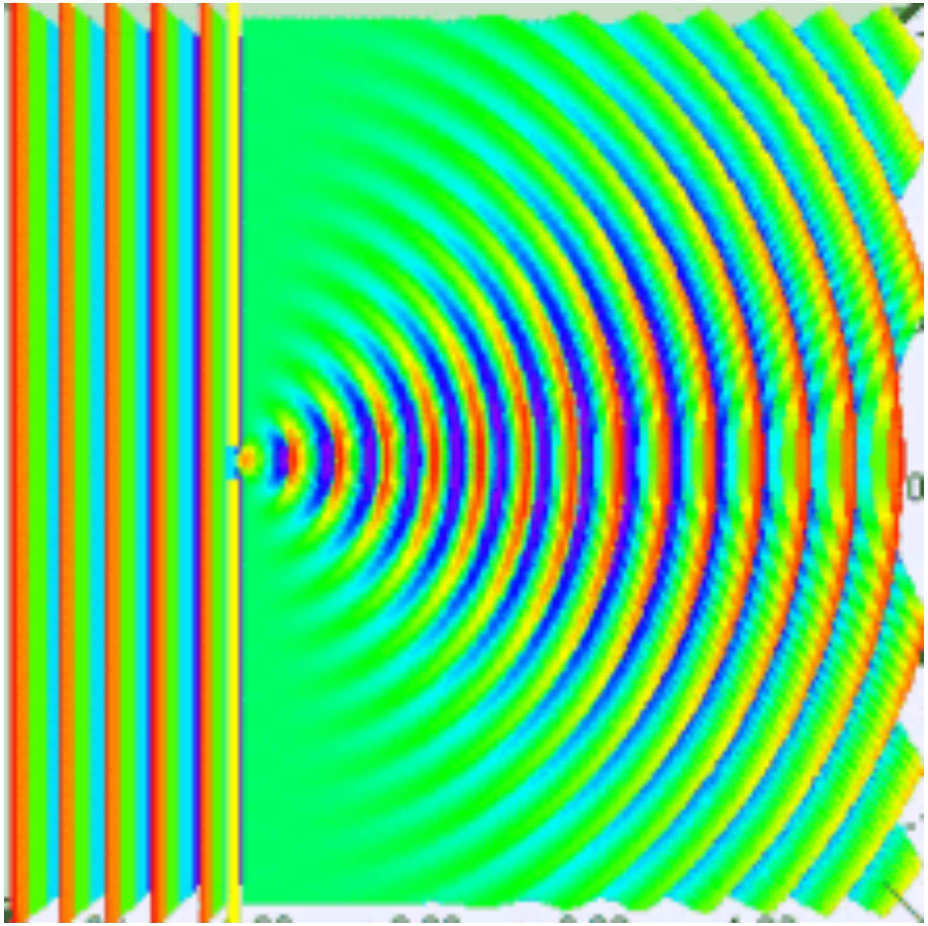
- Different height of the edge

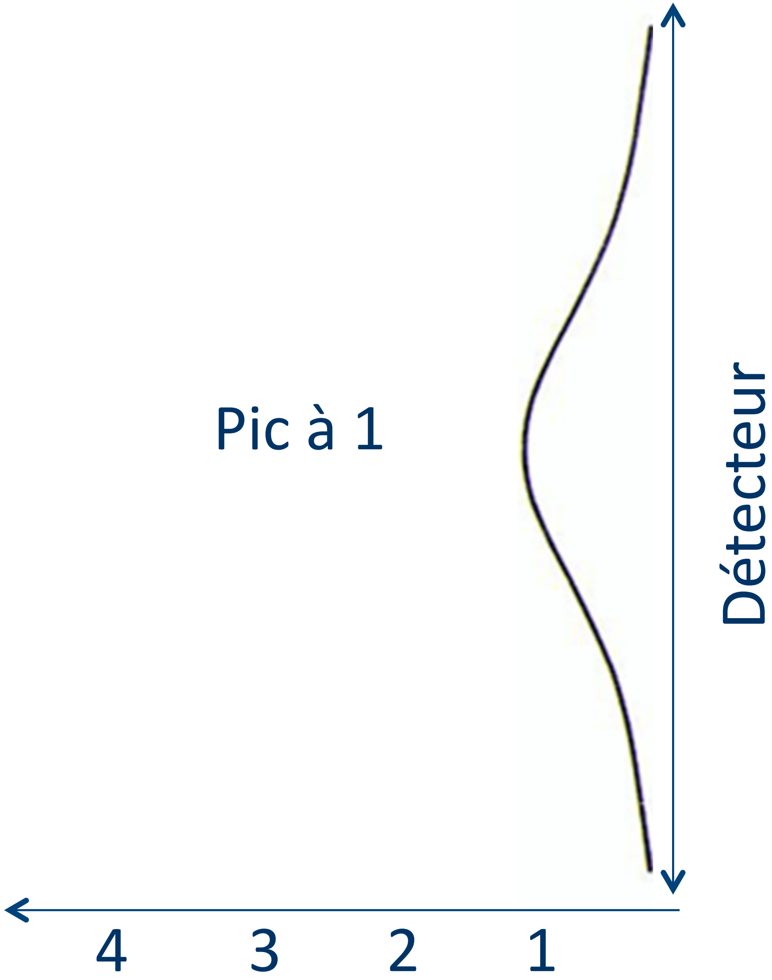


Scattering by one atom

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



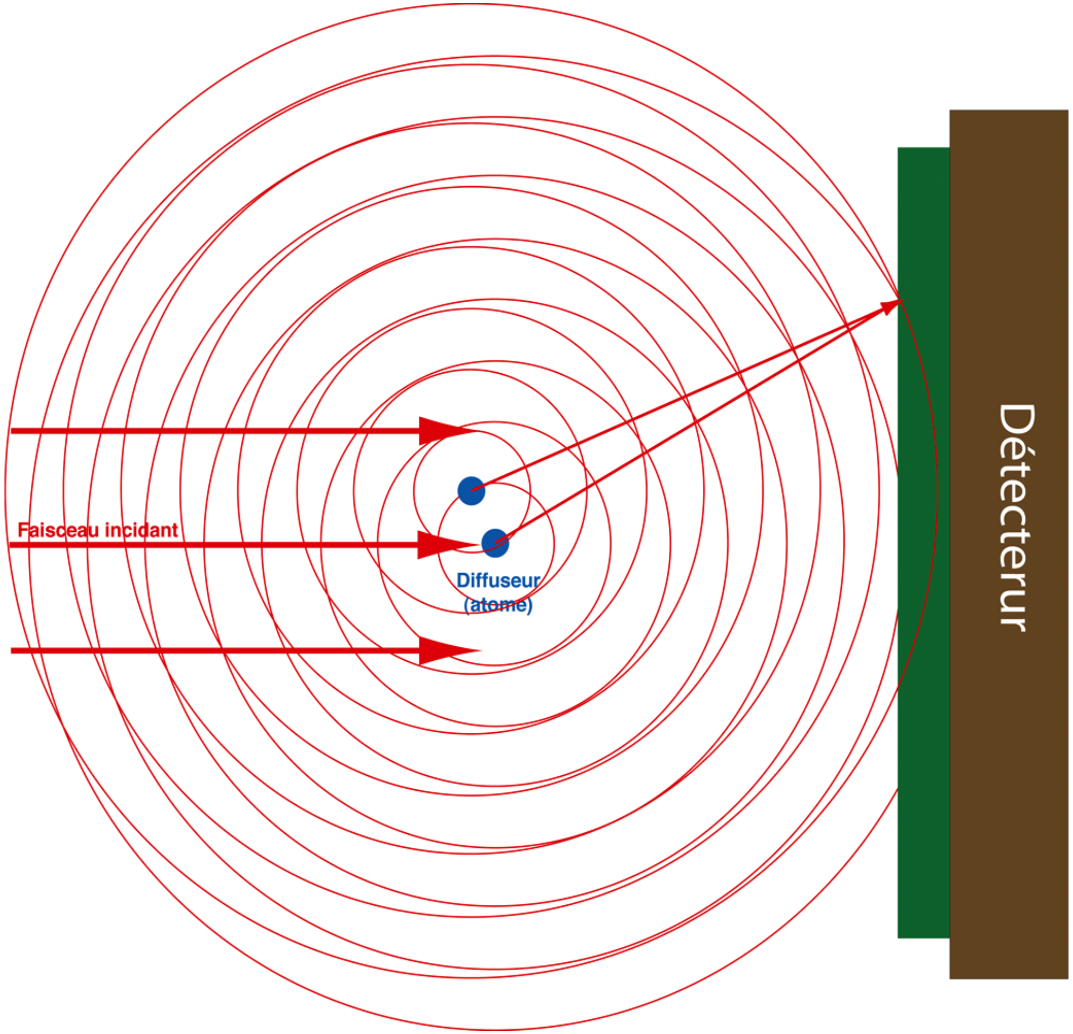


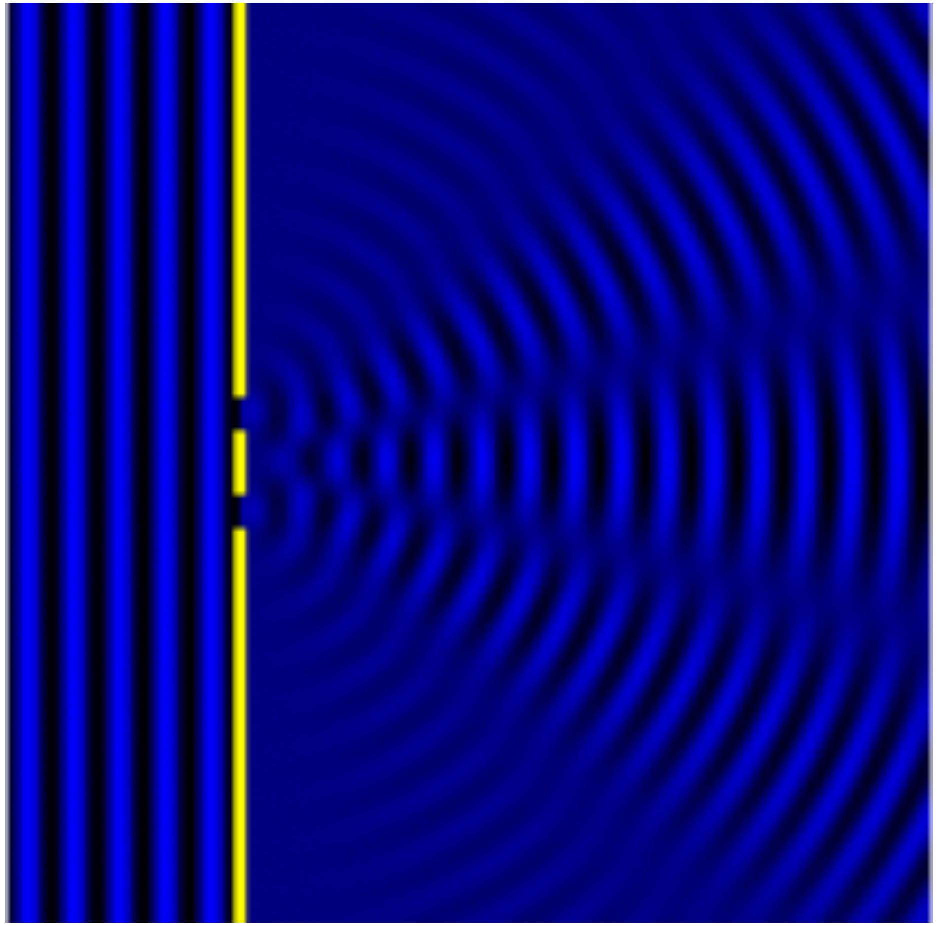


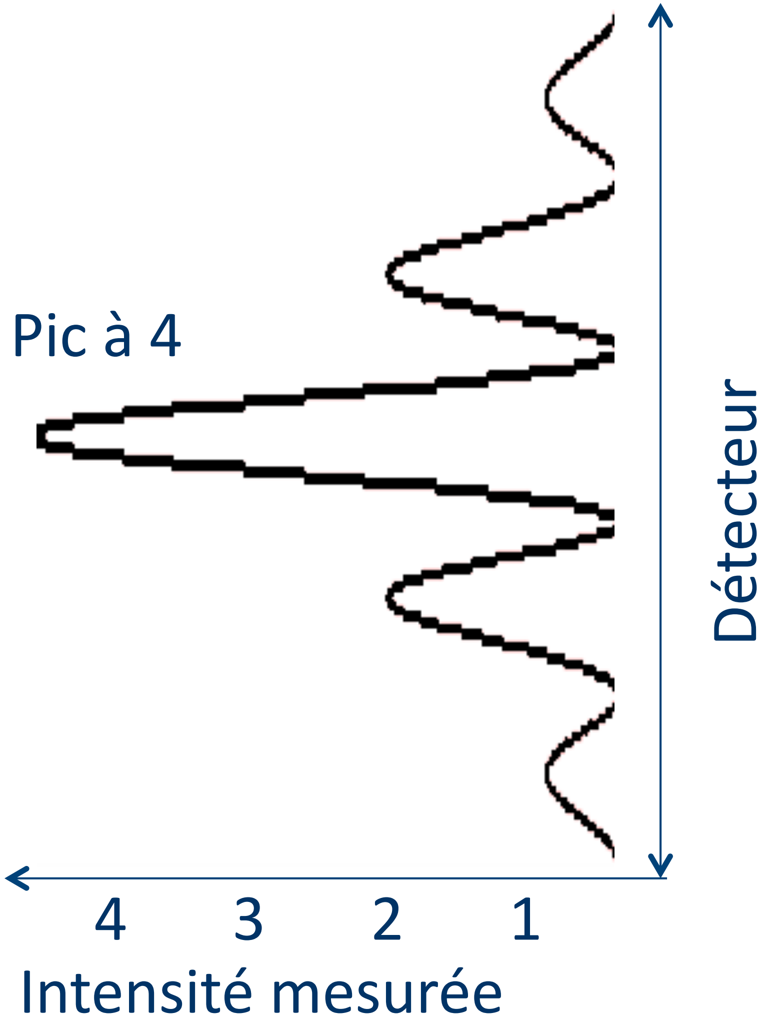
Pic à 1

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

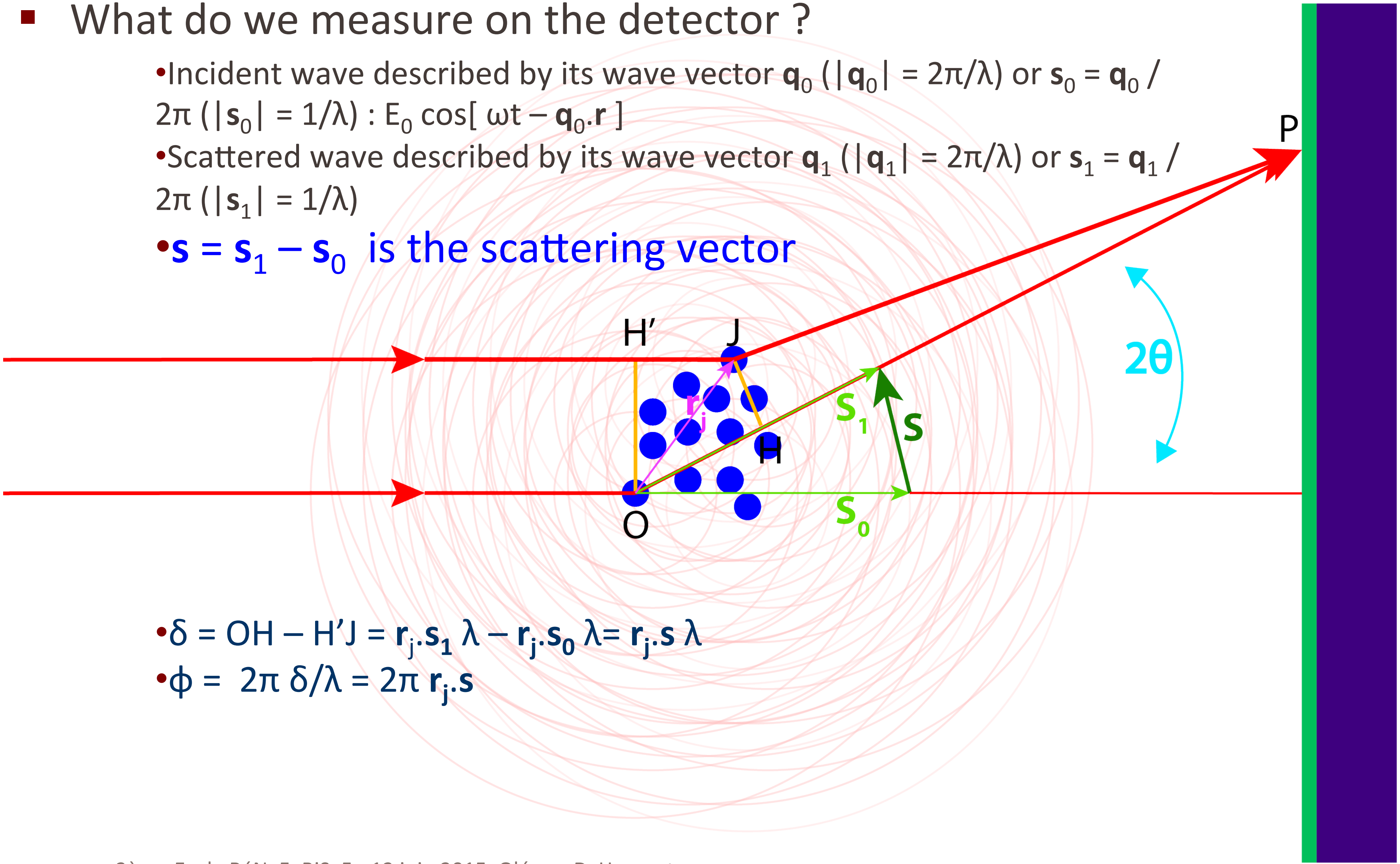





Pic à 4
Détecteur
4
3
2
1
Intensité mesurée

Scattering by numerous atoms

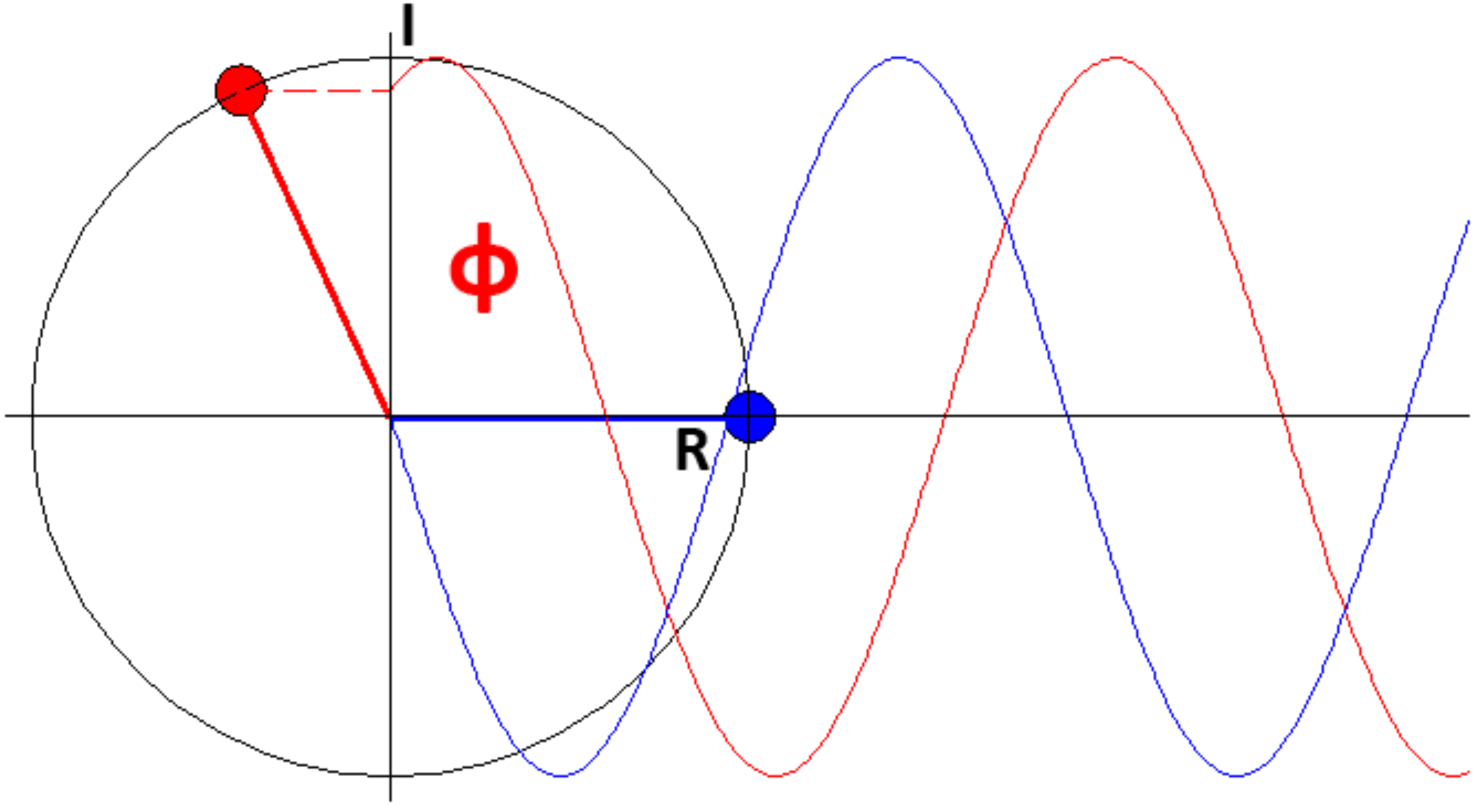
- What do we measure on the detector ?
  - Incident wave described by its wave vector $\mathbf{q}_0$ ($|\mathbf{q}_0| = 2\pi/\lambda$) or $\mathbf{s}_0 = \mathbf{q}_0 / 2\pi$ ($|\mathbf{s}_0| = 1/\lambda$) : $E_0 \cos[\,\omega t - \mathbf{q}_0.\mathbf{r}\,]$
  - Scattered wave described by its wave vector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



  - $\delta = OH - H'J = \mathbf{r}_j.\mathbf{s}_1\,\lambda - \mathbf{r}_j.\mathbf{s}_0\,\lambda = \mathbf{r}_j.\mathbf{s}\,\lambda$
  - $\phi = 2\pi\,\delta/\lambda = 2\pi\,\mathbf{r}_j.\mathbf{s}$

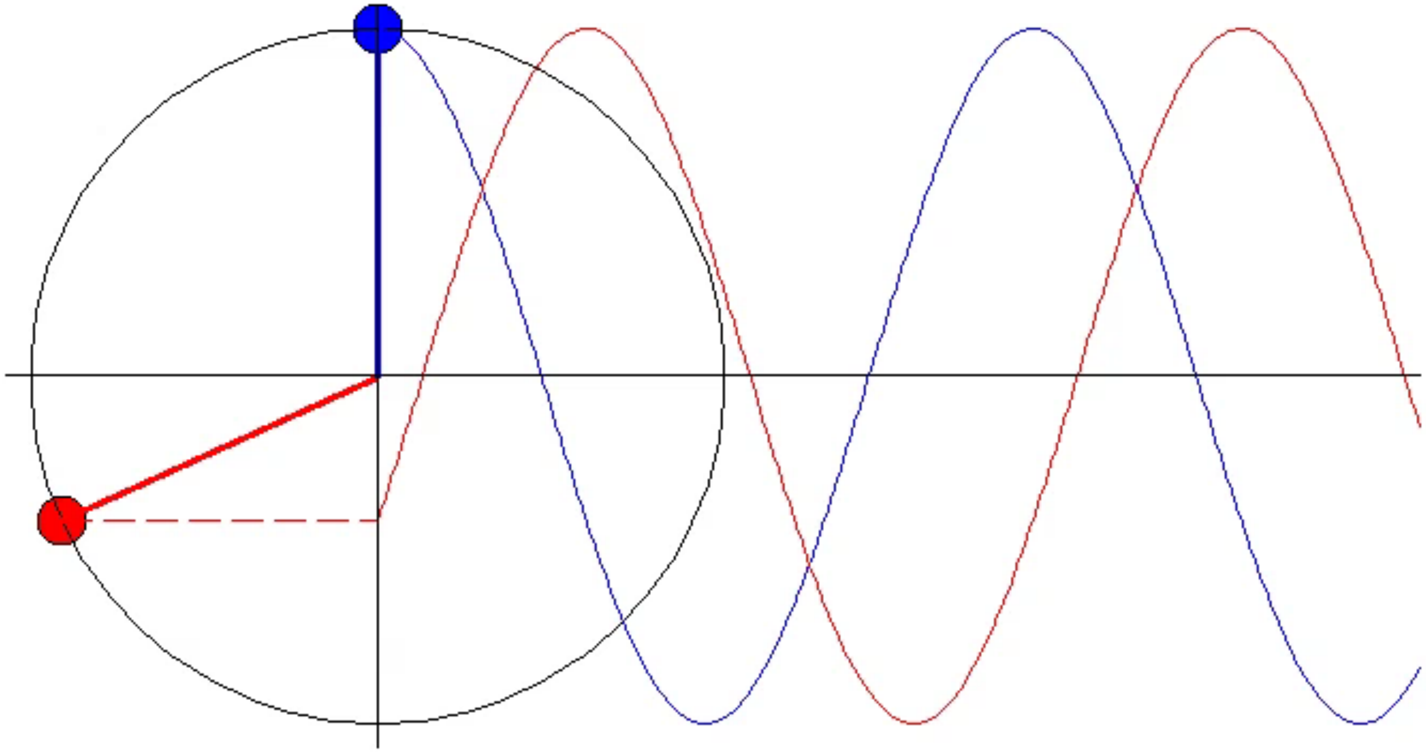
Représentation de Fresnel des ondes

- $E_0 \cos[\,\omega t - \mathbf{q}_0.\mathbf{r}\,] \Leftrightarrow E_0 \exp[i(\omega t - \mathbf{q}_0.\mathbf{r})]$
- $E_0 \cos[\,\omega t - \mathbf{q}_0.\mathbf{r} + \phi\,] \Leftrightarrow E_0 \exp[i(\omega t - \mathbf{q}_0.\mathbf{r} + \phi)]$
- $\Leftrightarrow E_0 \exp[i(\omega t - \mathbf{q}_0.\mathbf{r})].\exp(i\phi)$



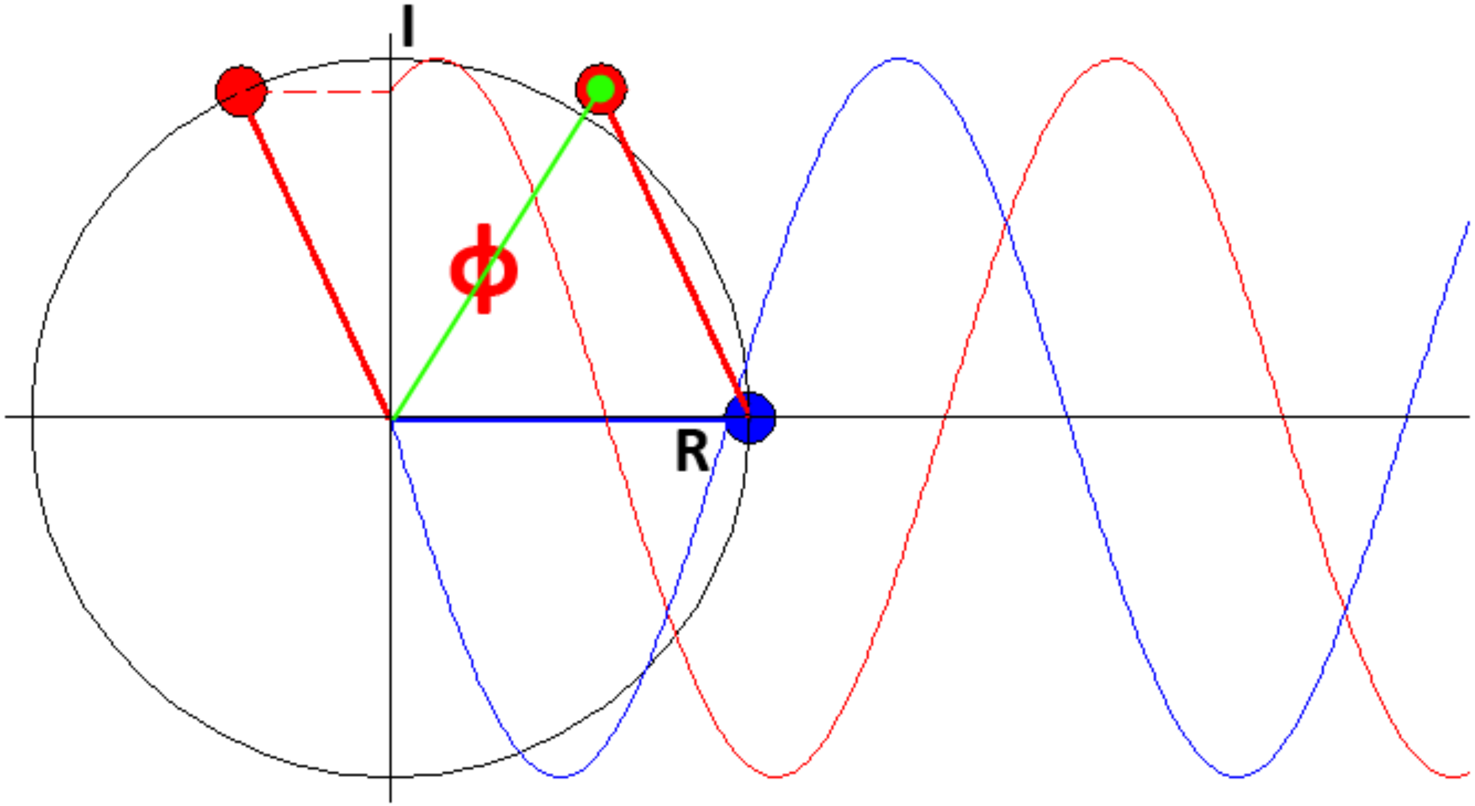
Représentation de Fresnel des ondes

- $E_0 \cos[\,\omega t - \mathbf{q}_0.\mathbf{r}\,] \Leftrightarrow E_0 \exp[i(\omega t - \mathbf{q}_0.\mathbf{r})]$
- $E_0 \cos[\,\omega t - \mathbf{q}_0.\mathbf{r} + \phi\,] \Leftrightarrow E_0 \exp[i(\omega t - \mathbf{q}_0.\mathbf{r} + \phi)]$
- $\Leftrightarrow E_0 \exp[i(\omega t - \mathbf{q}_0.\mathbf{r})].\exp(i\phi)$



Représentation de Fresnel des ondes

- $E_0 \cos[\ \omega t - \mathbf{q}_0.\mathbf{r}\ ] \Leftrightarrow E_0 \exp[i(\omega t - \mathbf{q}_0.\mathbf{r})]$
- $E_0 \cos[\ \omega t - \mathbf{q}_0.\mathbf{r} + \phi\ ] \Leftrightarrow E_0 \exp[i(\omega t - \mathbf{q}_0.\mathbf{r} + \phi)]$
- $\Leftrightarrow E_0 \exp[i(\omega t - \mathbf{q}_0.\mathbf{r})].\exp(i\phi)$



Scattered wave by an ensemble of atoms?

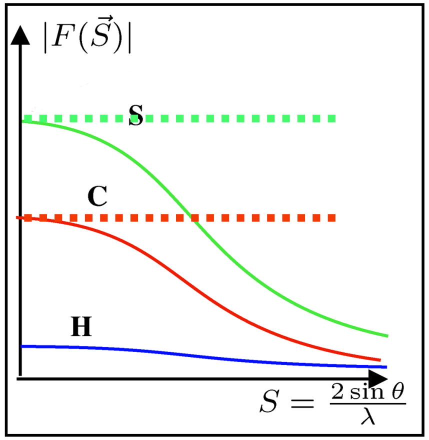
- Incident wave : $E_0 \exp[\, i(\omega t - \mathbf{q}_0.\mathbf{r})\,]$
- f_i is the scattering factor for the atom j at point J (defined by $\mathbf{r}_j$)
  - The wave scattered by atom j at J, as observed in P (defined by **s**) is:
    - $E_j(\mathbf{s}) = E_0/r'\ f_j \exp[\, i(\omega t - \mathbf{q}_0.\mathbf{r}_j)\,].\exp(-i\mathbf{q}_1.(\mathbf{r}-\mathbf{r}_j)).\ \exp(i\pi)$
    - $\quad = E_0/r'\ f_j \exp(i(\omega t - \mathbf{q}_1.\mathbf{r})).\ \exp(2i\pi\ \mathbf{r}_j.\mathbf{s}).\exp(i\pi)$
- To calculate the wave scattered by the entire sample, we have to sum on all the atoms of the sample:
  - $E(\mathbf{s}) = \Sigma_j\ E_0/r'\ f_j \exp(i(\omega t - \mathbf{q}_1.\mathbf{r})).\ \exp(2i\pi\ \mathbf{r}_j.\mathbf{s}).\exp(i\pi)$
  - $\quad = E_0/r' \exp(i(\omega t - \mathbf{q}_1.\mathbf{r}))\ \exp(i\pi)\ \Sigma_j\ f_j \exp(2i\pi\ \mathbf{r}_j.\mathbf{s})$
  - E(**s**) is proportional to the Fourier transform of the scattering factors of the atoms present in the sample
- For X-rays, this scattering factor is proportional to the electron density (number of electrons) of the atom
  - E(**s**) is proportional to the Fourier transform of the electron density

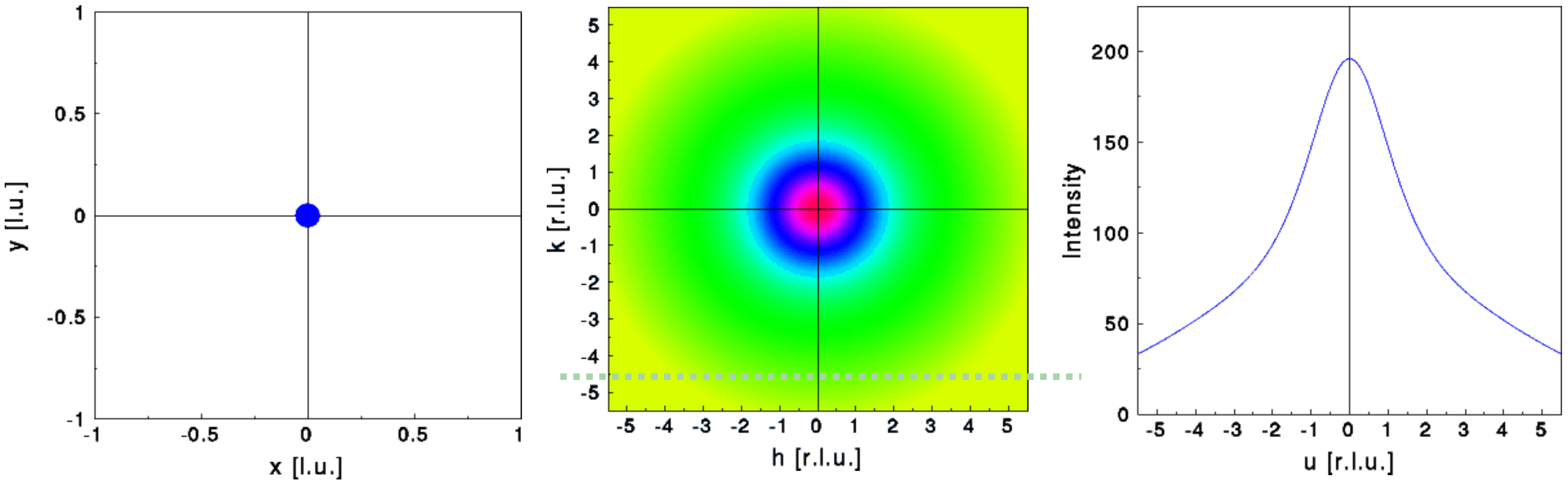
Structure factor

- $E(\mathbf{s}) = E_0/r' \, \sigma \int_v \rho(\mathbf{r}) \exp(2i\pi \, \mathbf{r}.\mathbf{s}) \, \partial\mathbf{r}$
- Structure factor:
  - Direct Fourier transform of the electron density
  - $F(\mathbf{s}) = \int_v \rho(\mathbf{r}) \exp(2i\pi \, \mathbf{r}.\mathbf{s}) \, \partial\mathbf{r} = \Sigma_j \; f_j \exp(2i\pi \, \mathbf{r}_j.\mathbf{s})$
  - 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 …

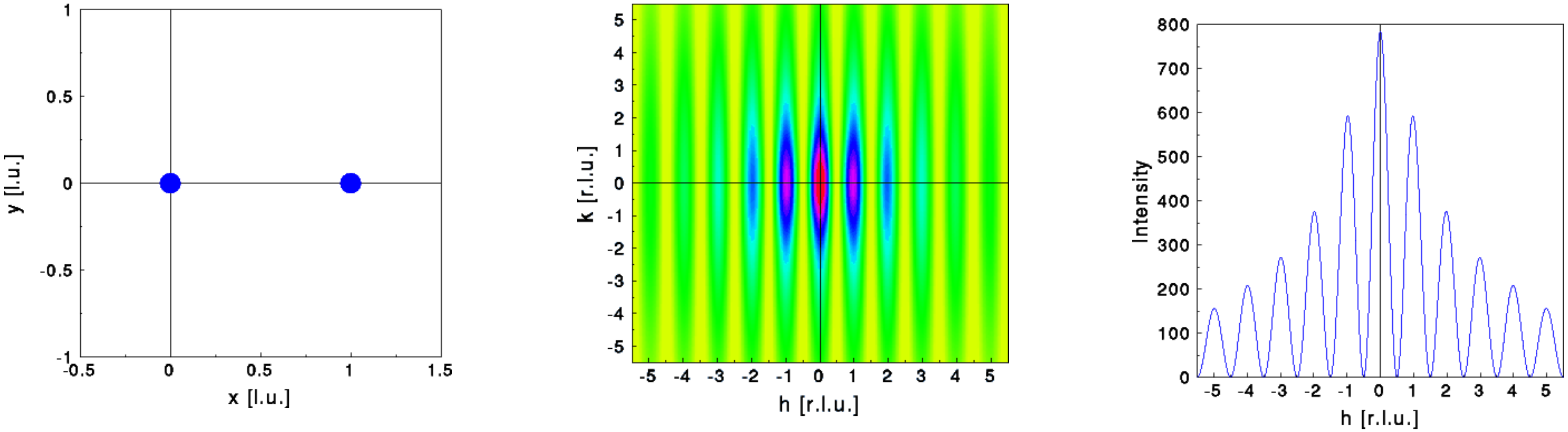
- Diffusion spectra of one atom
  - ➢Spherical electron density
  - ➢$f_{at}(s) = 2 \int_{r=0,+\infty} \rho_{at}(r)\, r\ \sin(2\pi\, sr) / s\ \partial r$
  - ➢$f_{at}(0)$ = number of electrons in atom
  - ➢$f_{at}(s)$ is constant when atom is ponctual (neutron diffraction)
  - ➢$f_{at}(s)$ is the atomic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



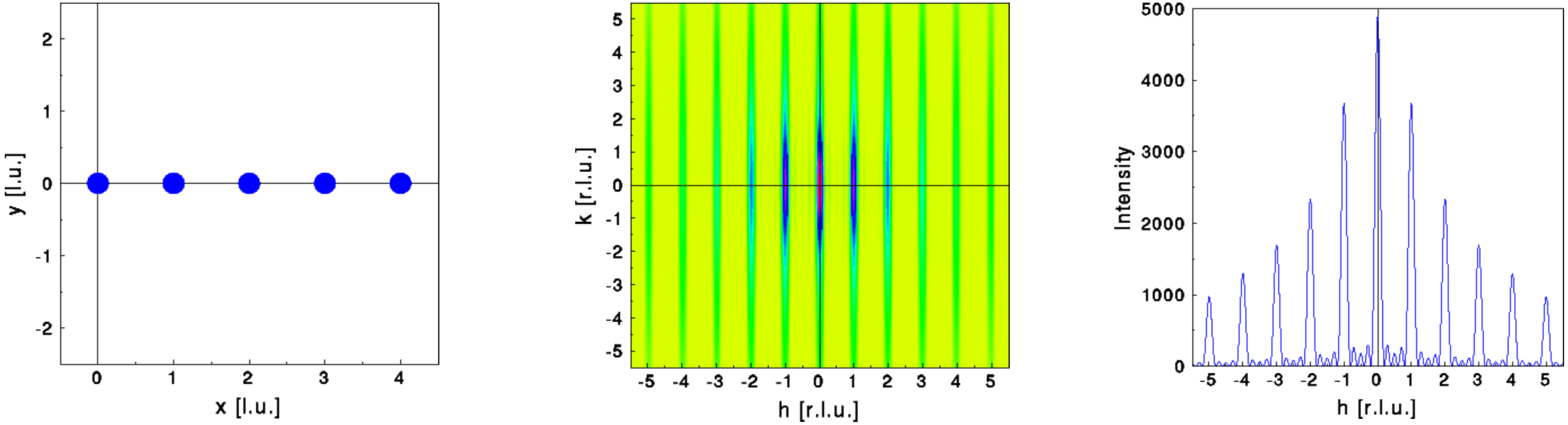


More atoms …

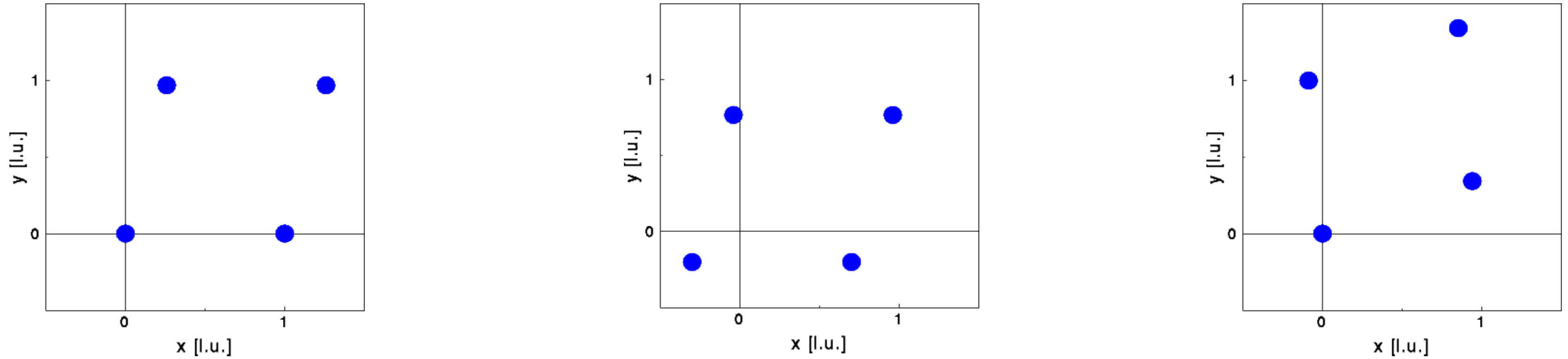
- Diffusion by two atoms



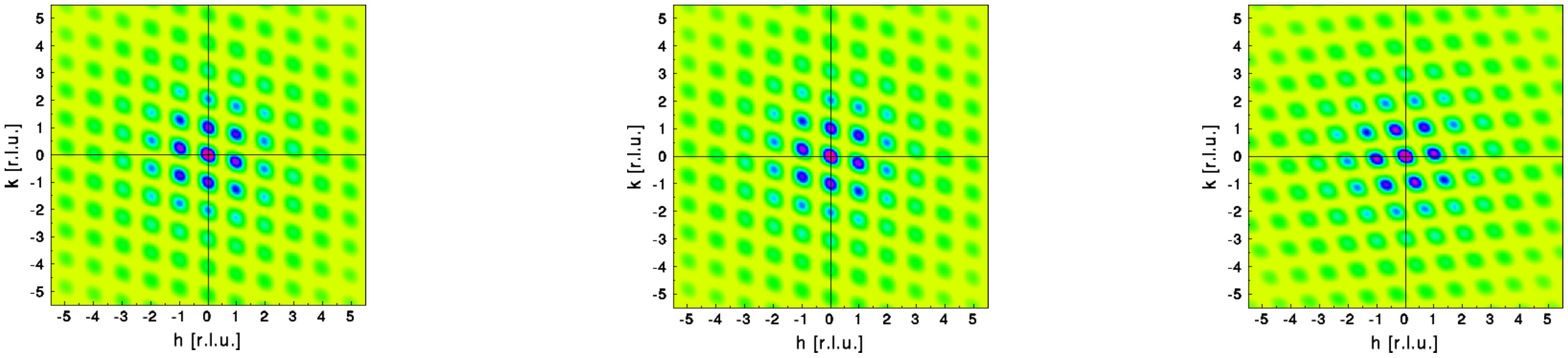
- Diffusion by five atoms



- 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

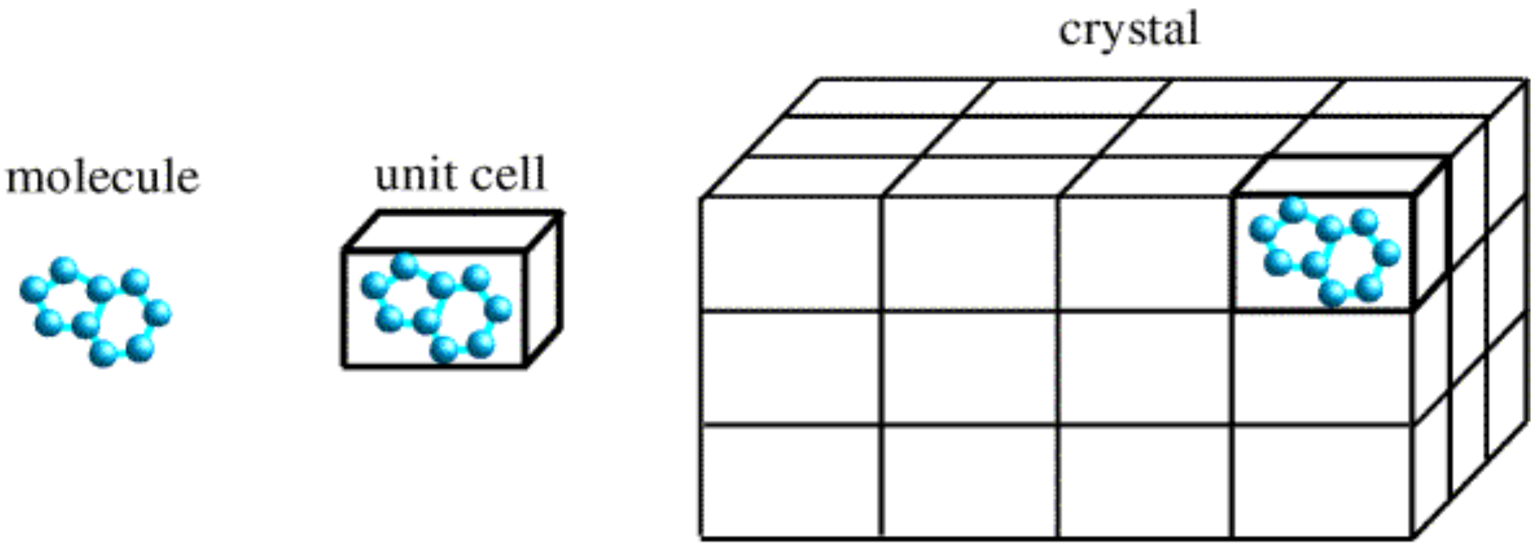


Structure factor of an assembly of atoms

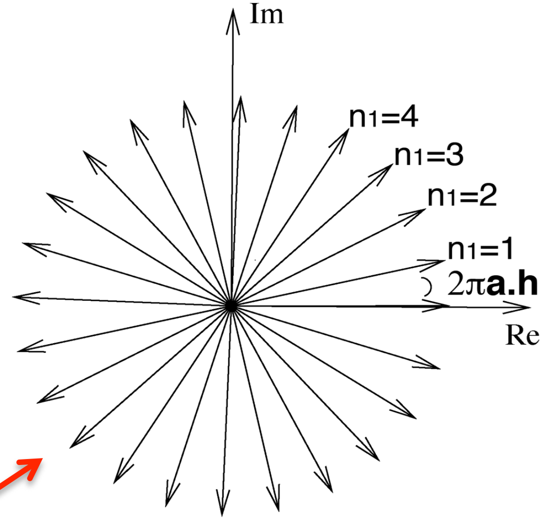
- $f_j(s)$ is the scattering factor of the atom j
- The phase shift of the wave scattered by an atom in position $\mathbf{r}_j$, in the direction defined by scattering vector **s** is (relatively to an atom at the origin):
  - ➢ $2\pi\mathbf{r}_j\mathbf{s}$
- The structure for an assembly a N atoms is 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:
  - $F(\mathbf{s}) = \int_{vech.} \rho(\mathbf{r}) \exp(i\, 2\pi \mathbf{s}.\mathbf{r})\, \partial\mathbf{r}$
  - $F(\mathbf{s}) = \sum_{n=1,Ncell} \int_{vcell} \rho(\mathbf{r}+\mathbf{r}_n) \exp(i\, 2\pi\, \mathbf{s}.(\mathbf{r}+\mathbf{r}_n))\, \partial\mathbf{r}$
  - $F(\mathbf{s}) = \sum_{n=1,Ncell} \exp(i\, 2\pi\, \mathbf{s}.\mathbf{r}_n) \int_{vcell} \rho(\mathbf{r}) \exp(i\, 2\pi\, \mathbf{s}.\mathbf{r})\, \partial\mathbf{r}$
  - $F(\mathbf{s}) = \sum_{n=1,Ncell} \exp(2i\pi\, n_1\mathbf{a}.\mathbf{s})\exp(2i\pi\, n_2\mathbf{b}.\mathbf{s})\exp(2i\pi\, n_3\mathbf{c}.\mathbf{s}) \int_{vcell} \rho(\mathbf{r}) \exp(i\, 2\pi\, \mathbf{s}.\mathbf{r})\, \partial\mathbf{r}$

- This structure factor close to zero, except for specific values of **s** :
  - $\sum_{n=1,Ncell} \exp(2i\pi\, n_1\mathbf{a}.\mathbf{s})\exp(2i\pi\, n_2\mathbf{b}.\mathbf{s})\exp(2i\pi\, n_3\mathbf{c}.\mathbf{s}) \sim 0$ except when
  - **a.s** = h, **b.s** = k, **c.s** = l where h, k, l are integers (Laue equations)
  - Diffraction spots instead of diffusion spectra
  - $\sum_{n=1,Ncell} \exp(2i\pi\, n_1\mathbf{a}.\mathbf{s})\exp(2i\pi\, n_2\mathbf{b}.\mathbf{s})\exp(2i\pi\, n_3\mathbf{c}.\mathbf{s}) = Ncell$ (number of cells in the crystal)

Structure factor of a crystal

- 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.
  - $F(\mathbf{s}) = N_{cell} \int_{vcell} \rho(\mathbf{r}) \exp(i\, 2\pi\, \mathbf{s}.\mathbf{r})\, \partial\mathbf{r}$
- The intensity of each spots is proportional to the square of the number of cells.
  - $I(\mathbf{s})\ \alpha\ N_{cell}^{2}\ (\int_{vcell} \rho(\mathbf{r}) \exp(i\, 2\pi\, \mathbf{s}.\mathbf{r})\, \partial\mathbf{r})(\int_{vcell} \rho(\mathbf{r}) \exp(-i\, 2\pi\, \mathbf{s}.\mathbf{r})\, \partial\mathbf{r})$
- The bigger your crystal is, the higher is your signal
- 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

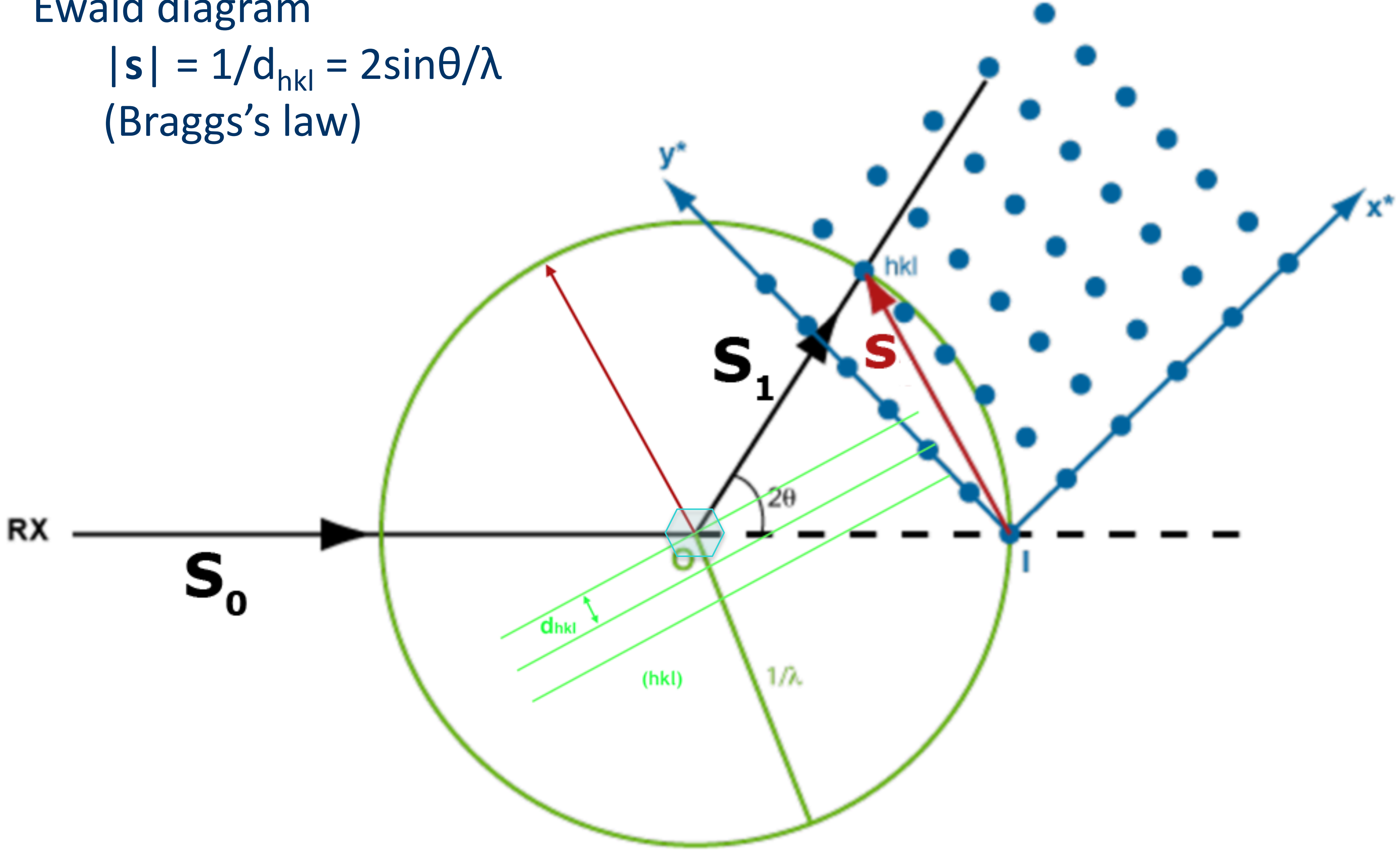
- Then **s** is a vector of a lattice named reciprocal lattice
  - **s** = h.**a*** + k.**b*** + l.**c*** where
    - **a*** = **b** ∧ **c** / V_{maille}, **b*** = **c** ∧ **a** / V_{maille}, **c*** = **a** ∧ **b** / V_{maille}
    - V_{maille} = (**a** ∧ **b**).**c**
    - Larger unit cell => smaller reciprocal cell => closer spots on diffraction spectra
- A reciprocal lattice is associated to each real space lattice (defined by the unit cell vectors)
- Each diffraction spots correspond to vectors in the reciprocal space
  - **s** = h.**a*** + k.**b*** + l.**c*** where h, k, l are integers

Ewald sphere and reticular plans

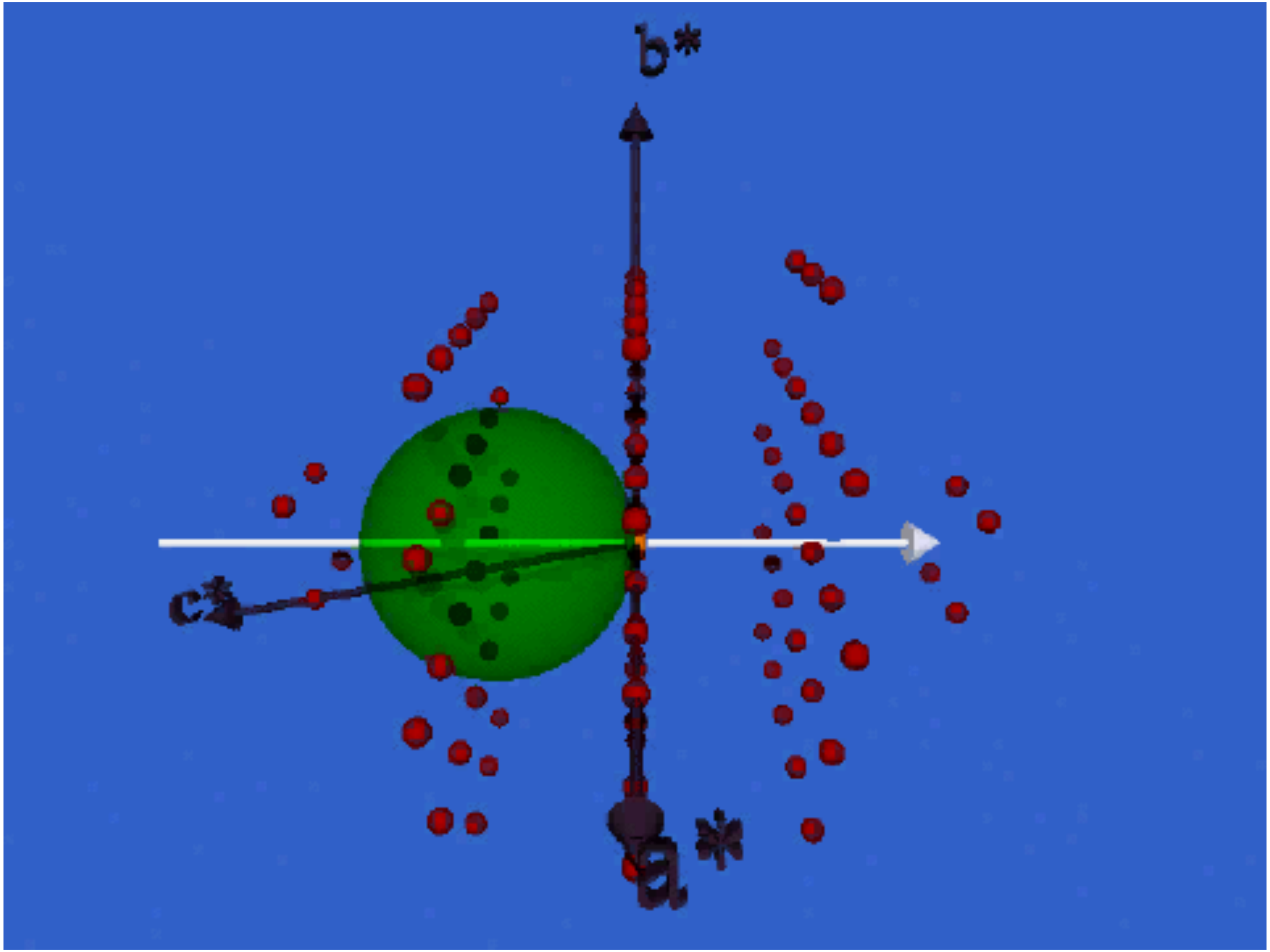
Ewald diagram

$|\mathbf{s}| = 1/d_{hkl} = 2\sin\theta/\lambda$

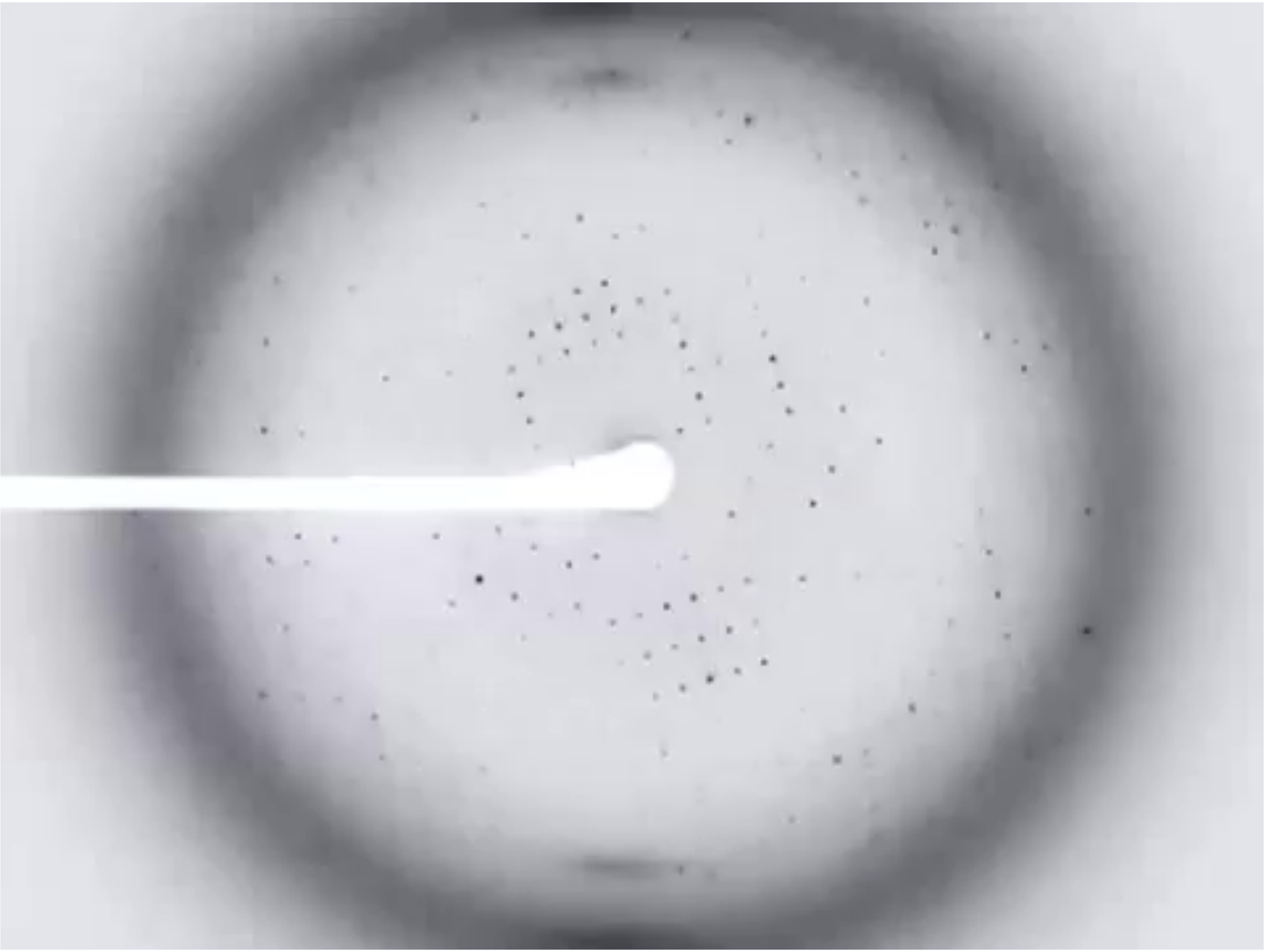
(Braggs's law)



Ewald sphere and data collection

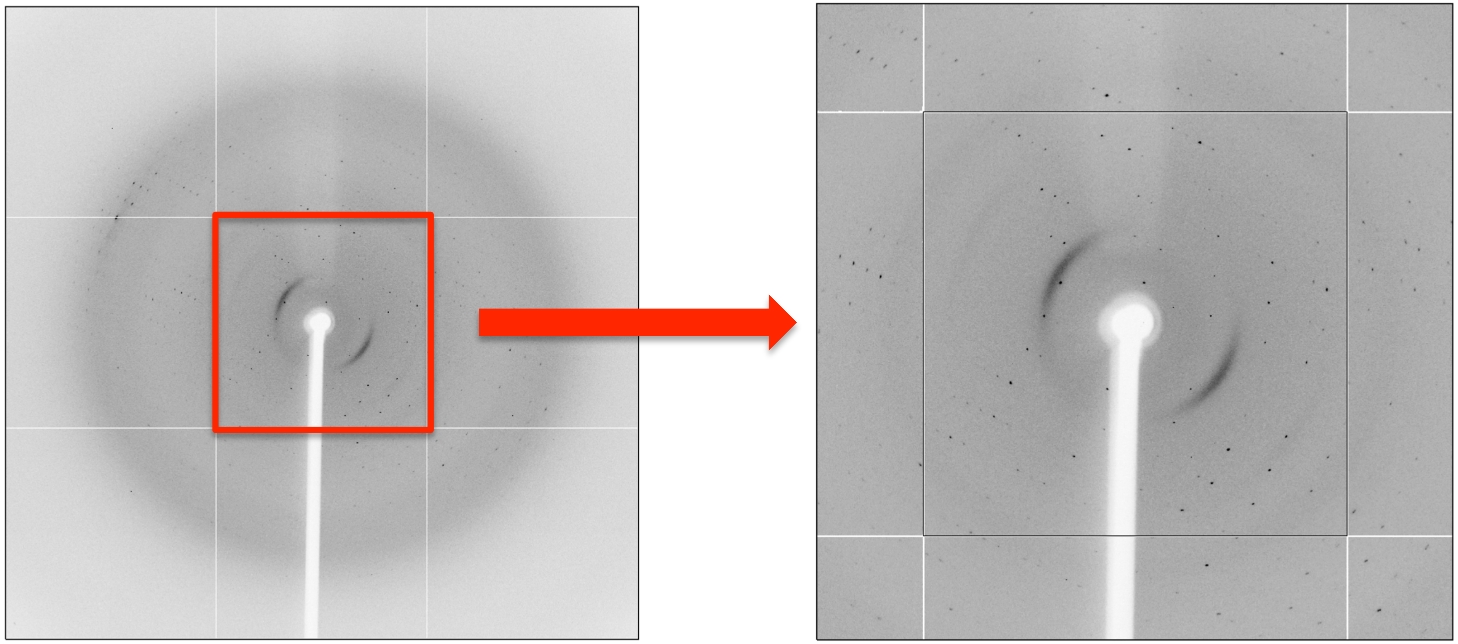


A real data collection (medium resolution)



What does the diffraction tell us ?

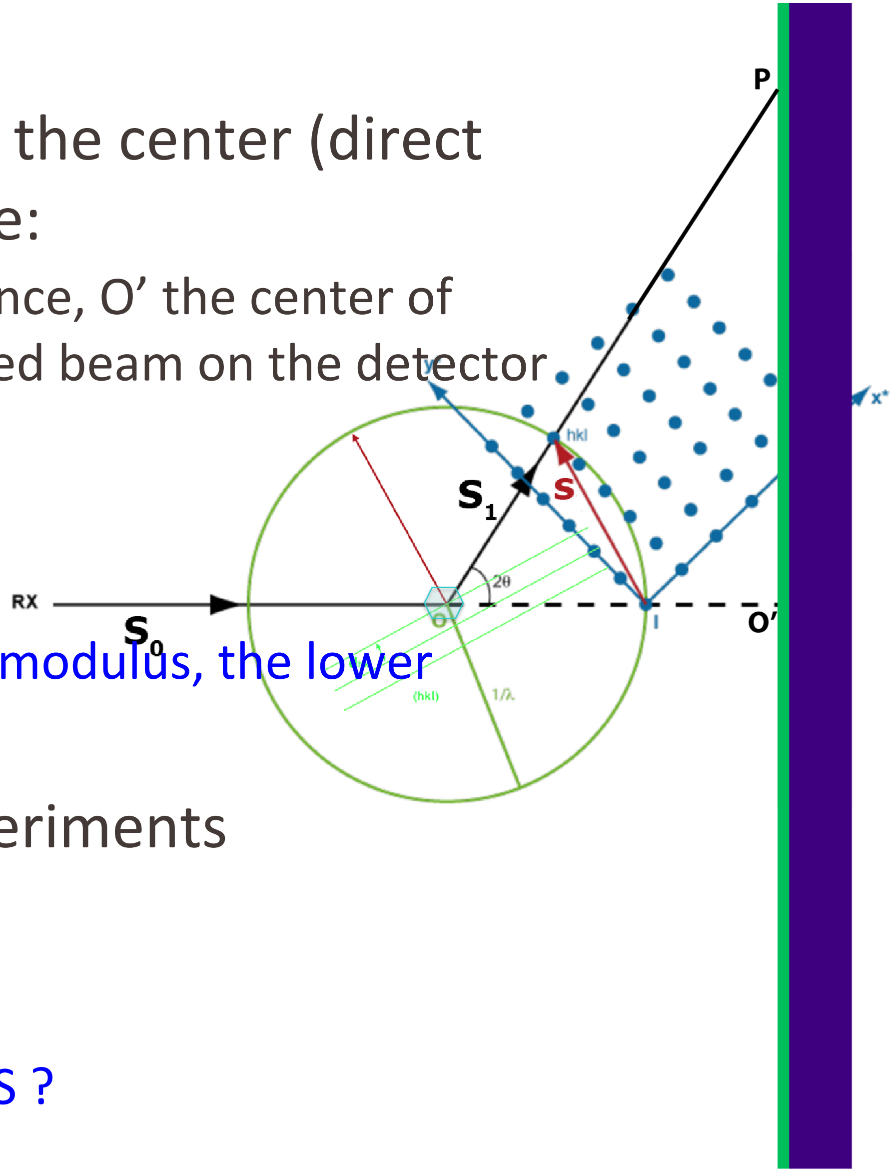
- The position of spots on diffraction images will allow the determination of the unit cell dimension and the orientation of the crystal
- The intensity of each spot contains information on the electron density of the entire cell
  - ➢ $I(\mathbf{s}) = F(\mathbf{s}).F^*(\mathbf{s})$
  - ➢ $= (\int_{vcell}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



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

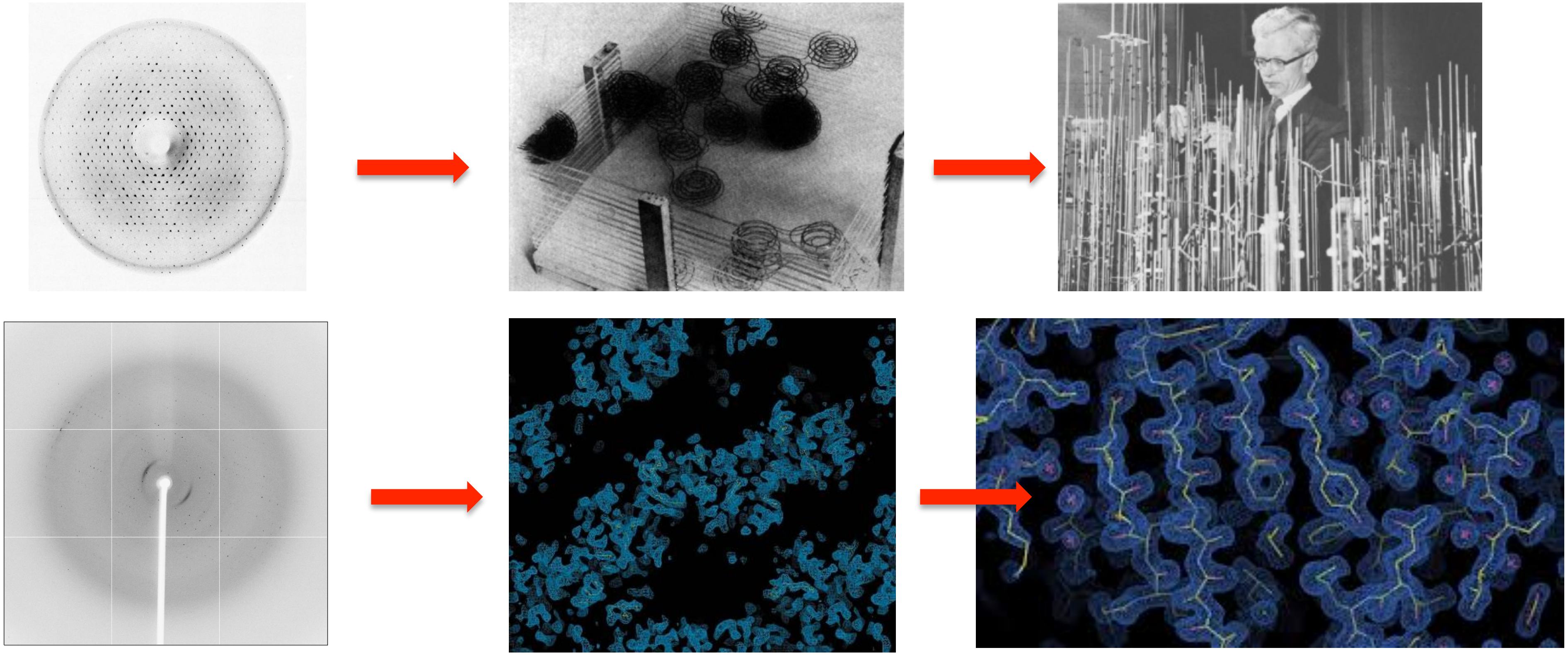
Resolution

- In crystallography
  - ➢d = 1 / |**s**| in Å
- Directly related with the distance from the center (direct beam position) of the diffraction image:
  - ➢If D (|OO'|) 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 ?



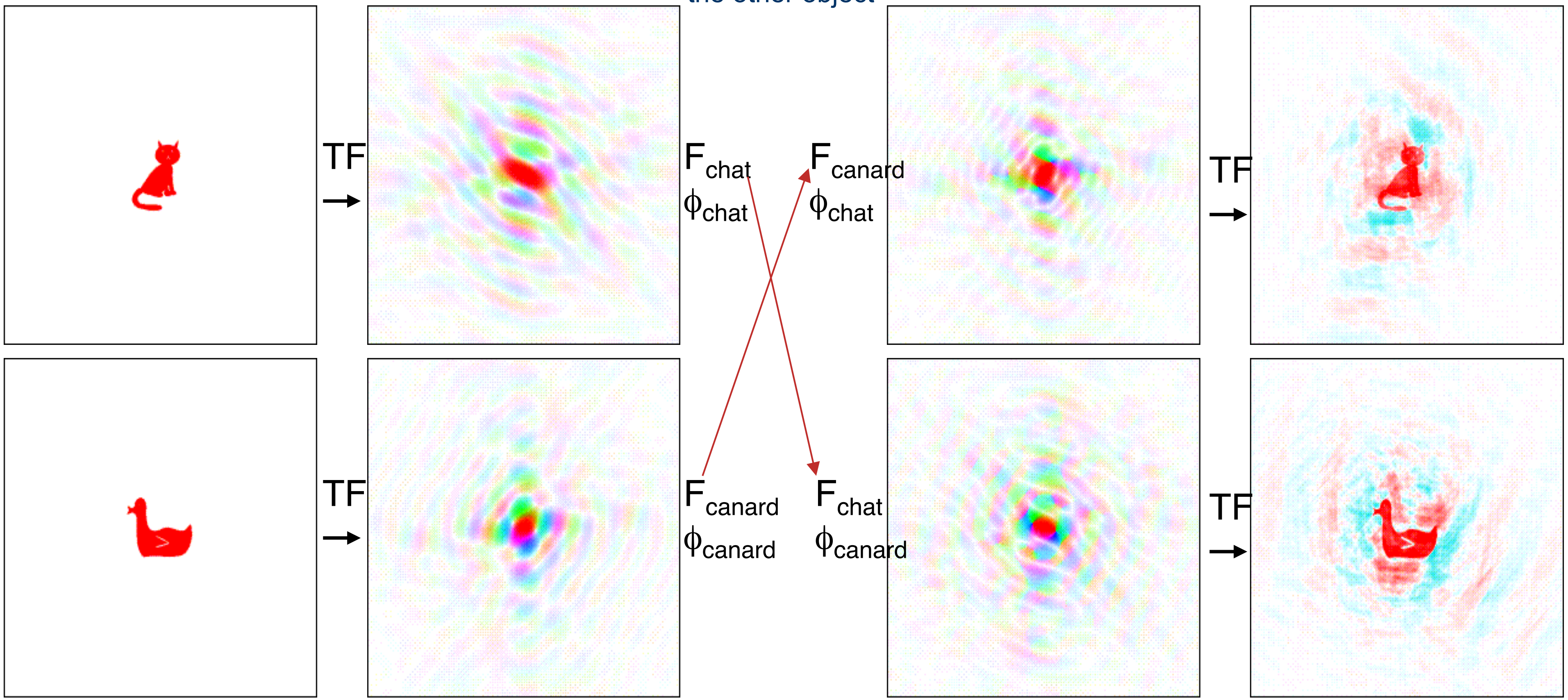
- Simple in principle
  - ➢The electron density within the unit cell can be derived from the structure factors :
    - •$\rho(\mathbf{r}) = 1/V_{\text{maille}}\, \Sigma_{\mathbf{s}}\, F(\mathbf{s}) \exp(-i\, 2\pi \mathbf{s}.\mathbf{r})$
    - •But …

Calculating the electron density

- We do not measure F($\mathbf{s}$) but only an intensity I($\mathbf{s}$) which is proportional to the square of F($\mathbf{s}$) modulus: the phase of the scattered wave is lost
    - ➢ $F(\mathbf{s}) = |F(\mathbf{s})| \exp(i\phi(\mathbf{s}))$
    - ➢ $I(\mathbf{s}) = |F(\mathbf{s})|^2$
    - ➢ 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 $\rho(\mathbf{r})$ => Patrice's Talk

Most of the structural information is in the phases



Object

Diffusion spectra

For each object we keep the phase and take the amplitudes of the other object

Inverse Fourier transform ...

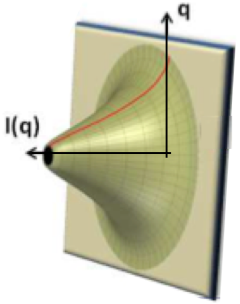
Patterson function

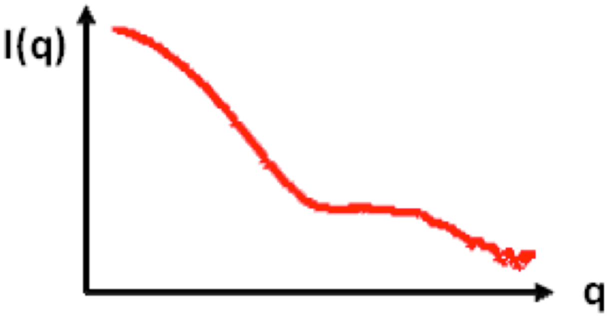
- The Patterson function can be calculated without phase information
  - $P(\mathbf{u}) = 1/V \sum_s F(\mathbf{s})\, F^*(\mathbf{s}) \exp(-2i\pi\, \mathbf{u} \cdot \mathbf{s})$
  - $= 1/V \sum_s |F(\mathbf{s})|^2 \exp(-2i\pi\, \mathbf{u} \cdot \mathbf{s})$
  - It is the reverse Fourier tranform of the 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
  - ➢$I_{\text{macromolecule}}(\mathbf{q}) = I_{\text{sample}}(\mathbf{q}) - I_{\text{buffer}}(\mathbf{q})$
- 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





- $I(q) = \langle TF(\Delta\rho(\mathbf{r}))\ .\ TF(\Delta\rho(-\mathbf{r}))\rangle$
- $I(q) = \langle TF(\int_{V\text{particle}} \Delta\rho(\mathbf{r+u}).\ \Delta\rho(\mathbf{u})d\mathbf{u})\rangle$
- $I(q) = 4\pi \int_{0,\inf} p(r).\sin(qr)/qr\ dr$
  - 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