

Diffraction Methods in Chemistry

University of Cambridge Part II Natural Sciences Tripos

Yue Wu

*Yusuf Hamied Department of Chemistry
Lensfield Road,
Cambridge, CB2 1EW*

yw628@cam.ac.uk

Acknowledgements

Nothing in these lecture notes is original. They are largely based on the notes by Dr. Andrew Bond, who lectured this course in 2025. They are nowhere near accurate representations of what was actually lectured, and in particular, all errors are almost surely mine.

Contents

1	Diffraction	4
1.1	Addition of Electromagnetic Waves	4
1.2	Diffraction of an Object	7
1.3	Diffraction of X-rays from Electrons	8
2	The Phase Problem and the Patterson Function	11
2.1	The Phase Problem	11
2.2	Patterson Function	12
2.3	Illustrating the Patterson Function	14
2.4	Patterson Function of a Molecule	17
2.5	The Patterson Function of Ideal Gas	20

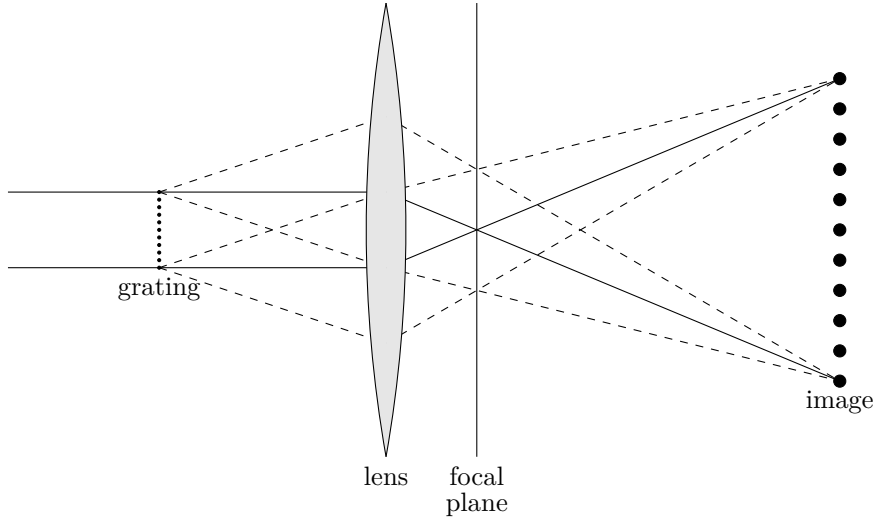


Figure 1.1: A magnified image of the diffraction grating is formed using a convex lens to focus the diffraction pattern.

1 Diffraction

It might be familiar from physics that if you shine a beam of light on a diffraction grating, then a diffraction pattern will emerge. If you put a lens after the diffracted beam, they may be brought to a focus, and a magnified image of the diffraction grating will be seen, as shown in the figure.

Atoms in a crystal are, in some sense, very fine gratings. Can we use a similar method to image the atoms in a crystal directly? However, to form a diffraction pattern, we need the wavelength of the light to be in roughly the same length scale as the pattern we are trying to image. The atoms are in the length scales of Angstroms 10^{-10} m, therefore, we need to use X-ray for diffractions at atomic length scales. However, there is no lens for X rays (because lenses are made of atoms too), so the imaging process does not work. What we have to do is to measure the intensity of the diffracted beams at different directions, and infer the atomic structure mathematically.

This is, however, a hard thing to do. We will spend the vast majority of our lectures on how to translate the diffracted beam intensities to atomic structure.

1.1 Addition of Electromagnetic Waves

X-rays are electromagnetic waves, so they are described by Maxwell's equations, from which we can infer that the electromagnetic fields are essentially just oscillating electric and magnetic fields, travelling in some direction \mathbf{k} (the magnitude of \mathbf{k} also encodes the wavelength via $|\mathbf{k}| = 2\pi/\lambda$) with the speed of light c .¹ We will label the distance as x along the direction of propagation of the electromagnetic wave \mathbf{k} , and consider

¹In free space, $\rho = 0$ and $\mathbf{J} = 0$, so by Maxwell's equations

$$\nabla \times (\nabla \times \mathbf{E}) = -\nabla \times \frac{\partial \mathbf{B}}{\partial t} = -\frac{\partial}{\partial t} (\nabla \times \mathbf{B}) = -\frac{\partial}{\partial t} \left(\mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t} \right) = -\mu_0 \epsilon_0 \frac{\partial^2 \mathbf{E}}{\partial t^2}. \quad (1.1)$$

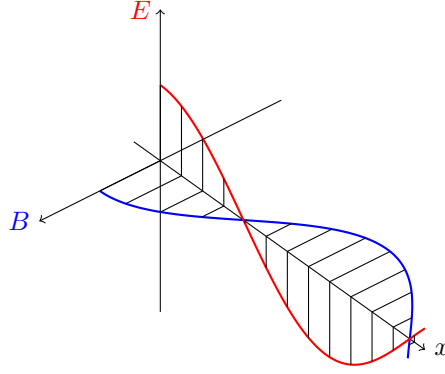


Figure 1.2: Electromagnetic wave propagating in space.

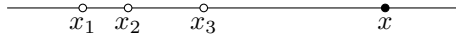
the wave at some fixed time t . Then the electric (or equivalently the magnetic) field strength is given by

$$\psi(x) = A \sin\left(\frac{2\pi x}{\lambda} + \phi\right), \quad (1.7)$$

where the real number A is the amplitude and ϕ is the phase offset. However, it turns out that when considering oscillations, it is always easier to use complex numbers, so we will alternatively see the electric field as the real part of the complex wave

$$\psi(x) = A \exp\left(\frac{2\pi i x}{\lambda}\right), \quad (1.8)$$

where $A = |A| e^{i\phi}$ is now complex and includes both the amplitude $|A|$ and the initial phase offset ϕ .



Now what happens if we have multiple sources, each sending off magnetic waves of the same frequencies but with different amplitudes and phases? Let's consider the simple 1D case, where n sources are located at x_j . We will let all three sources to have

By vector calculus identities, we also have

$$\nabla \times (\nabla \times \mathbf{E}) = \nabla(\nabla \cdot \mathbf{E}) - \nabla^2 \mathbf{E} = -\nabla^2 \mathbf{E}, \quad (1.2)$$

and hence

$$\nabla^2 \mathbf{E} = \mu_0 \epsilon_0 \frac{\partial^2 \mathbf{E}}{\partial t^2}. \quad (1.3)$$

This is the wave equation, with wave speed $c = 1/\sqrt{\mu_0 \epsilon_0}$, with general complex solution

$$\mathbf{E} = \mathbf{E}_0 e^{i(\mathbf{k} \cdot \mathbf{r} - \omega t)} \quad (1.4)$$

for $\mathbf{E} \in \mathbb{C}$ and $\mathbf{k}^2 c^2 = \omega^2$. By Maxwell's equations, we also have

$$\frac{\partial \mathbf{B}}{\partial t} = -\nabla \times \mathbf{E} = -i\mathbf{k} \times \mathbf{E}, \quad (1.5)$$

and so

$$\mathbf{B} = \frac{\mathbf{k}}{\omega} \times \mathbf{E}. \quad (1.6)$$

Therefore, \mathbf{E} and \mathbf{B} are both transverse wave that are in the same phase but perpendicular to each other.

a zero initial phase², so that the resulting wave at point x due to source j is given by

$$\psi_j(x) = A_j \exp\left(\frac{2\pi i(x - x_j)}{\lambda}\right), \quad (1.9)$$

where A_j is real because we have a zero initial phase. We will denote the phase of the wave due to source j at x as

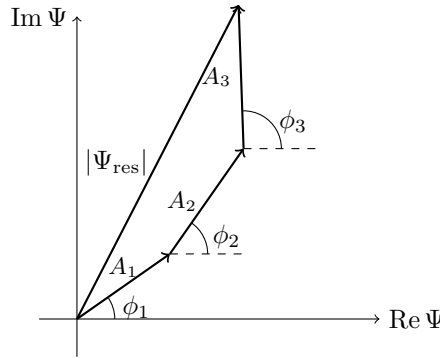
$$\phi_j = \frac{2\pi(x - x_j)}{\lambda}. \quad (1.10)$$

Then to work out the total wave, we only need to sum up the contributions from each sources to get

$$\Psi_{\text{res}}(x) = \sum_j A_j \exp(i\phi_n). \quad (1.11)$$

It is easy to see that the phase angle of the resultant amplitude is

$$\Phi_{\text{res}} = \tan^{-1} \left[\frac{\sum_j A_j \sin \phi_j}{\sum_j A_j \cos \phi_j} \right]. \quad (1.12)$$



Now what if we move the measuring point x by an amount of δx ? This will increase the phases of the wave from all the sources by

$$\delta\phi = \frac{2\pi\delta x}{\lambda}. \quad (1.13)$$

The net result is that the resultant magnitude of the combined wave does not change, but the phase changes by $\delta\phi$, as

$$\Psi'_{\text{res}}(x) = \sum_j A_j \exp(i(\phi_n + \delta\phi)) \quad (1.14)$$

$$= e^{i\delta\phi} \sum_j A_j \exp(i\phi_n). \quad (1.15)$$

What if we change the position of one of the sources x_k ? This will in general lead to a change in the phase ϕ_k . Now both the phase and the amplitude of the resulting wave is different.

²It is a trivial generalisation for the sources to have different non-zero initial phases. We make such assumption because this is the only situation we will need for later discussion, and it makes my diagram cleaner. It is actually also trivial to generalise it to 3D by replacing $x - x_j$ with $\|\mathbf{x} - \mathbf{x}_j\|$ in all the expressions below.

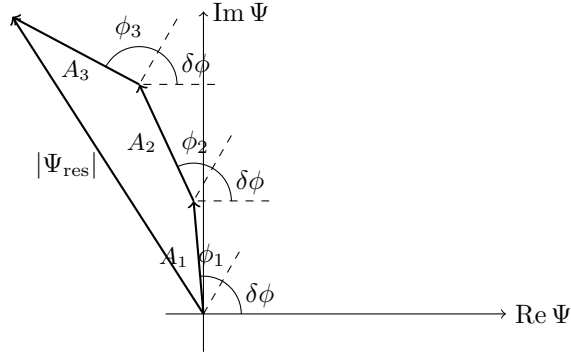


Figure 1.3: Movement of the measuring point results in a change of measured phase only.

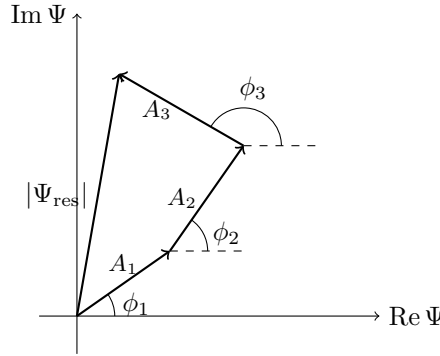


Figure 1.4: A change in the position of one of the sources leads to a change in both the magnitude and the phase measured.

1.2 Diffraction of an Object

We would now like to consider the diffraction of an extended object. Suppose we have parallel incident beams from the direction $\hat{\mathbf{s}}_0$ illuminating the whole object, and we are measuring the diffracted beam at some distant point Q . The dimension of this object is small compared to its distance to the measuring point Q so we can sensibly say that Q is at direction $\hat{\mathbf{s}}$ for all points in the object.³ To produce the diffraction pattern, we need to sum up the radiation scattered from all the points in the object. To do this, we first pick an arbitrary origin O in the object, and we need to figure out the intensity and phase difference between the beams diffracted in direction $\hat{\mathbf{s}}$ from O and from any other point P with position vector \mathbf{r} .

From figure 1.5, it is easy to see that the beam diffracted from P and O has a path difference

$$\Delta x = \mathbf{r} \cdot \hat{\mathbf{s}}_0 - \mathbf{r} \cdot \hat{\mathbf{s}}. \quad (1.16)$$

In particular, a positive path difference means that the beam from P falls behind the beam from O , leading to a negative phase difference, so

$$\Delta\phi = -\frac{2\pi\Delta x}{\lambda} = \frac{2\pi\mathbf{r} \cdot (\hat{\mathbf{s}} - \hat{\mathbf{s}}_0)}{\lambda}. \quad (1.17)$$

³This type of diffraction is known as *Fraunhofer diffraction*. If you are observing near the diffracting object, then the diffraction is *Fresnel diffraction*.

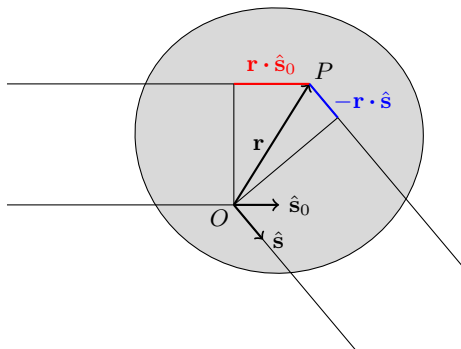


Figure 1.5: Path difference for diffractions in a general object.

The amplitude of the beam diffracted from O and P may also be different, since the object may be inhomogeneous. Let $A(\mathbf{r})$ be a real function representing the diffraction amplitude, then by integrating over all points, the wave measured at point Q is

$$\Psi(\hat{\mathbf{s}}) = \int d^3\mathbf{r} A(\mathbf{r}) \exp\left(\frac{2\pi i \mathbf{r} \cdot (\mathbf{s} - \mathbf{s}_0)}{\lambda}\right). \quad (1.18)$$

We can define the *scattering vector* \mathbf{S} by

$$\mathbf{S} := \frac{\hat{\mathbf{s}} - \hat{\mathbf{s}}_0}{\lambda}. \quad (1.19)$$

It is a vector bisecting the incident and diffracted beam, as shown in the figure below. If we define the angle of deflection to be 2θ , then the length of the scattering vector is

$$\|\mathbf{S}\| = \frac{2 \sin \theta}{\lambda}. \quad (1.20)$$

The scattering vector has dimension $[\text{L}]^{-1}$, so it exists in the reciprocal space, which might be familiar from Part IB Chemistry A.⁴ This allows us to simplify our expression

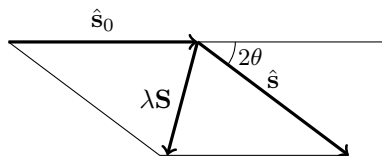


Figure 1.6: Construction of the scattering vector.

of the measured wave as

$$\Psi(\mathbf{S}) = \int d^3\mathbf{r} A(\mathbf{r}) \exp(2\pi i \mathbf{r} \cdot \mathbf{S}). \quad (1.22)$$

1.3 Diffraction of X-rays from Electrons

We will now consider how X-rays interacts with matter. The interaction generally fall into two categories.

⁴In some derivation, one uses *momentum transfer vector* \mathbf{Q} defined by

$$\mathbf{Q} = \frac{2\pi(\hat{\mathbf{s}} - \hat{\mathbf{s}}_0)}{\lambda}. \quad (1.21)$$

instead of \mathbf{S} .

- *Elastic scattering.* This is our main focus. In an *elastic scattering*, or *Thomson scattering*, there is no net transfer of energy from the photon to the material. It can be think of as some kind of resonance of the atomic electrons. When X-rays pass through the material, the oscillating electric field causes the atomic electrons to oscillate, which in turn emit X-rays in all directions. The incoming and outgoing rays retain a non-random phase relationship and are said to be *coherent*.
- *Inelastic processes.* *Inelastic processes* involve energy transfer from the incoming X-rays to electrons in atoms. The simplest example is photoelectron emission, where an X-ray photon providing energy for an atom to eject a core-shell electron. The photons are then re-emitted with lower energy in an incoherent fashion. This incoherence means that they cannot interfere with each other to produce diffraction patters. Other processes of this type may result from interaction with valence shell electrons or vibrations, all being incoherent. We are therefore not interested in inelastic processes at this stage — we will return to this later.

The quantitative description of Thomson scattering strictly only applies to free electrons, but the results are found to be generally applicable to crystals under normal scattering conditions. Crucially, the amplitude of X-rays diffracted from some point is proportional to the local electron density $\rho(\mathbf{r})$. Therefore,

$$\Psi(\mathbf{S}) = \int d^3\mathbf{r} \sigma \rho(\mathbf{r}) \exp(2\pi i \mathbf{r} \cdot \mathbf{S}), \quad (1.23)$$

where σ is the scattering power of a single electron. We can divide through this constant of proportionality to obtain the *structure factor* $F(\mathbf{S})$ given by

$$F(\mathbf{S}) = \int d^3\mathbf{r} \rho(\mathbf{r}) \exp(2\pi i \mathbf{r} \cdot \mathbf{S}). \quad (1.24)$$

Mathematically, $F(\mathbf{S})$ is the *Fourier transform* of the electron density $\rho(\mathbf{r})$.

Definition 1.1. Let $f : \mathbb{R}^n \rightarrow \mathbb{C}$ be a function. The *Fourier transform* of f is $\mathcal{F}[f] \equiv \tilde{f} : \mathbb{R}^n \rightarrow \mathbb{C}$ given by⁵

$$\tilde{f}(\mathbf{k}) := \int d^n \mathbf{x} f(\mathbf{x}) e^{2\pi i \mathbf{k} \cdot \mathbf{x}}. \quad (1.26)$$

Theorem 1.2. Let $\tilde{f}(\mathbf{k})$ be the Fourier transform of $f(\mathbf{x})$. The *inverse Fourier transform* that converts \tilde{f} back to f is given by

$$f(\mathbf{x}) = \mathcal{I}[\tilde{f}(\mathbf{k})] := \int d^3\mathbf{k} \tilde{f}(\mathbf{k}) e^{-2\pi i \mathbf{k} \cdot \mathbf{x}}. \quad (1.27)$$

⁵You might be more familiar with the Fourier transform defined as

$$\tilde{f} = \frac{1}{\sqrt{2\pi}} \int dx f(x) e^{-ikx}, \quad (1.25)$$

where the normalisation factor of $1/\sqrt{2\pi}$ may or may not be there. It is just a matter of convention. However, this difference in convention is annoying since it results in slight tweaks in a lot of formulae.

Proof.

$$\begin{aligned}
\mathcal{F}[\tilde{f}](\mathbf{x}) &= \int d^n \mathbf{k} e^{-2\pi i \mathbf{k} \cdot \mathbf{x}} \int d^3 \mathbf{s} e^{2\pi i \mathbf{k} \cdot \mathbf{s}} f(\mathbf{s}) \\
&= \int d^n \mathbf{s} f(\mathbf{s}) \int d^n \mathbf{k} e^{2\pi i \mathbf{k} \cdot (\mathbf{s} - \mathbf{x})} \\
&= \int d^n \mathbf{s} f(\mathbf{s}) \delta(\mathbf{s} - \mathbf{x}) \\
&= f(\mathbf{x}).
\end{aligned} \tag{1.28}$$

□

Hence to convert the $F(\mathbf{S})$ back to $\rho(\mathbf{r})$, one simply needs to perform the inverse Fourier transform

$$\rho(\mathbf{r}) = \int d^3 \mathbf{S} F(\mathbf{S}) \exp(-2\pi i \mathbf{r} \cdot \mathbf{S}). \tag{1.29}$$

This means that if we can determine the $F(\mathbf{S})$, then we can easily reconstruct the electron density in a material, no matter how complex the structure is. (1) Measure the phases and the amplitudes of the diffracted beams. (2) Divide by single-electron scattering power σ to get $F(\mathbf{S})$. (3) Do the inverse Fourier transform to obtain $\rho(\mathbf{S})$. Simple as that! Three steps to solve any chemical structure in the world!

This should be the happy ending of our story on X-ray diffraction.

2 The Phase Problem and the Patterson Function

2.1 The Phase Problem

If the problem is really that simple, then this would be a one-lecture course instead of a 12-lecture one. There is a huge issue. We cannot measure the phase of an X-ray.

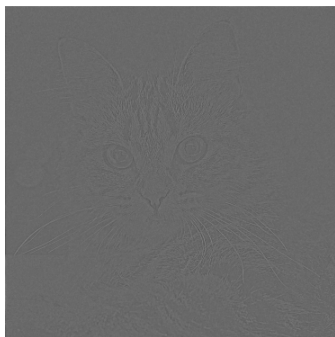
To measure the magnitude and the phase of an electromagnetic wave, we would use an antenna. However, this requires the antenna to have its dimensions comparable to the radiation wavelength. The X-ray we used for diffraction is at atomic length scale — that's why we can see the diffraction, but it is impossible to make an antenna at atomic length scale since antennas should be made from atoms! All we can measure in a diffraction experiment is the number of X-ray photons encountered in a given time, which is the diffracted *intensity*. This is proportional to $I = |F(\mathbf{S})|^2$. The phase information is completely lost. This is referred to as the *phase problem* in diffraction.

Now the question is: is getting the correct phase important? If using a wrong phase of the diffracted beam has little impact on the electron density calculated, then this would not be a huge problem. We can calculate the modulus of $F(\mathbf{S})$ by taking the square root of the intensity, and just plug in some random phases to calculate $\rho(\mathbf{r})$. However, it turns out that getting the correct phase is extremely important — it even plays a more important role than the magnitude to some extent. This can be illustrated by the following little experiment.

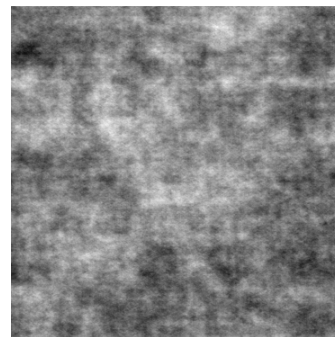
Instead of doing Fourier transform of a electron density function, we do a Fourier transform on a picture. A black and white picture is also a function of the pixels, where the values of the function is the gray scale. We can transform it into the frequency domain, and then transform it back, and we get the original image. However, if we use some randomly generated phase values to replace the correct phase values in the Fourier transform of the image, and we transform it back, we see a complete mess. This is what will happen if we use random phase value for the structure factor — we will get no valuable information. If fact, if we use some random value for the magnitudes of the Fourier transform, with the correct phase values, we can still vaguely see the shape of a cat.



(a) Correct magnitude + correct phase



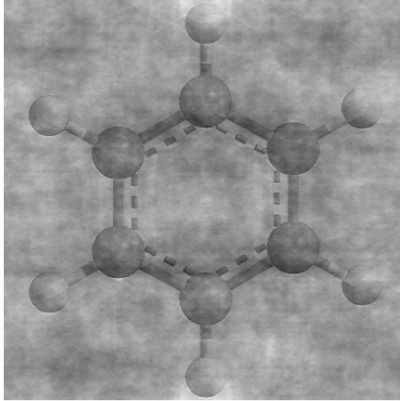
(b) Random magnitude + correct phase



(c) Correct magnitude + random phase

Figure 2.1: The effect of using the wrong phases and magnitudes in an inverse Fourier transform.

Just for fun, we can extend this experiment further. We have two images: one is a benzene and one is a cat. What will happen if we use the magnitude value of one figure with the phase value of the other?



(a) Magnitude of cat + phase of benzene



(b) Magnitude of benzene + phase of cat

Figure 2.2: Phases govern what we see in an inverse Fourier transform. If you use the phases of a benzene, you see a benzene. If you use the phases of a cat, you see a cat.

Moral: phase is important, but we can't measure it. The main task of structure determination using X-ray diffraction is to deduce the phases of the structure factor. This is what we will spend the remaining 11-lectures worths of time on.

2.2 Patterson Function

Since all what we can get from measuring a diffraction pattern is the intensity $I(\mathbf{S}) = |F(\mathbf{S})|^2$, let's see what its inverse Fourier transform gives us anyway.

Before doing that, let's first introduce convolution and correlation.

Definition 2.1. For two functions $f, g : \mathbb{R}^n \rightarrow \mathbb{C}$, their *convolution* is

$$(f * g)(\mathbf{x}) := \int d^n \mathbf{s} f(\mathbf{s}) g(\mathbf{x} - \mathbf{s}), \quad (2.1)$$

and their *correlation* is

$$(f \otimes g)(\mathbf{x}) := \int d^n \mathbf{s} f(\mathbf{s})^* g(\mathbf{x} + \mathbf{s}). \quad (2.2)$$

Theorem 2.2 (Convolution theorem). The Fourier transform of a convolution is the product of Fourier transform.

$$\mathcal{F}[f * g] = \mathcal{F}[f] \mathcal{F}[g]. \quad (2.3)$$

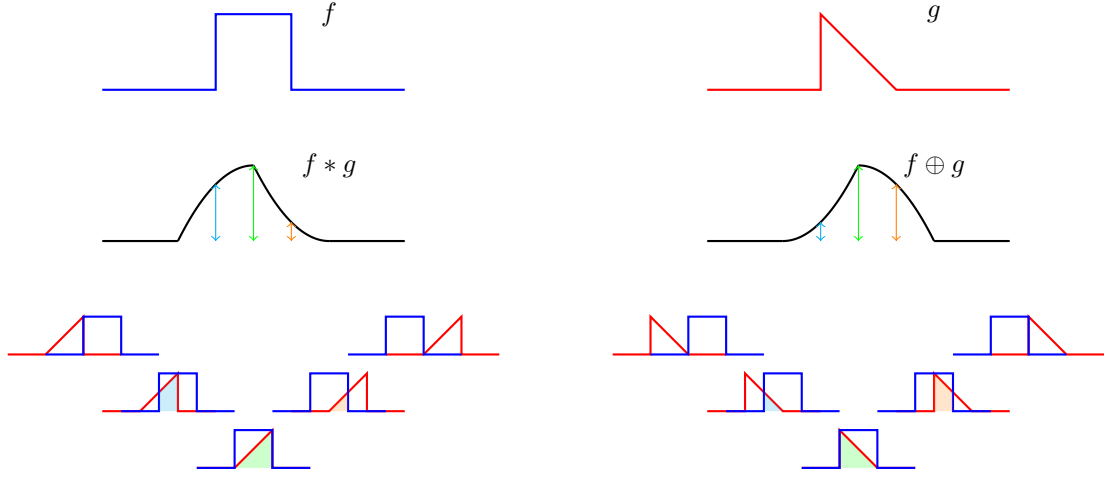


Figure 2.3: Graphic explanation of convolution and correlation.

Proof.

$$\begin{aligned}
\mathcal{F}[f * g](\mathbf{k}) &= \int d^n \mathbf{x} e^{2\pi i \mathbf{k} \cdot \mathbf{x}} \int d^n \mathbf{s} f(\mathbf{s}) g(\mathbf{x} - \mathbf{s}) \\
&= \int d^n \mathbf{s} f(\mathbf{s}) \int d^n \mathbf{x} e^{2\pi i \mathbf{k} \cdot \mathbf{x}} g(\mathbf{x} - \mathbf{s}) \\
&= \int d^n \mathbf{s} f(\mathbf{s}) \int d^n \mathbf{u} e^{2\pi i \mathbf{k} \cdot (\mathbf{u} + \mathbf{s})} g(\mathbf{u}) \quad \text{define } \mathbf{u} = \mathbf{x} - \mathbf{s} \\
&= \int d^n \mathbf{s} e^{2\pi i \mathbf{k} \cdot \mathbf{s}} f(\mathbf{s}) \int d^n \mathbf{u} e^{2\pi i \mathbf{k} \cdot \mathbf{u}} g(\mathbf{u}) \\
&= \tilde{f}(\mathbf{k}) \tilde{g}(\mathbf{k}) .
\end{aligned} \tag{2.4}$$

□

Theorem 2.3. Conversely, the Fourier transform of a product fg is the convolution of their Fourier transforms

$$\mathcal{F}[fg] = \mathcal{F}[f] * \mathcal{F}[g] . \tag{2.5}$$

Proof.

$$\begin{aligned}
\mathcal{F}[fg](\mathbf{k}) &= \int d^n \mathbf{x} e^{2\pi i \mathbf{k} \cdot \mathbf{x}} f(\mathbf{x}) g(\mathbf{x}) \\
&= \int d^n \mathbf{x} e^{2\pi i \mathbf{k} \cdot \mathbf{x}} f(\mathbf{x}) \int d^n \ell e^{-2\pi i \ell \cdot \mathbf{x}} \tilde{g}(\ell) \\
&= \int d^n \ell \tilde{g}(\ell) \int d^n \mathbf{x} e^{2\pi i \mathbf{x} \cdot (\mathbf{k} - \ell)} f(\mathbf{x}) \\
&= \int d^n \ell \tilde{g}(\ell) \tilde{f}(\mathbf{k} - \ell) \\
&= \tilde{f} * \tilde{g}(\mathbf{k}) .
\end{aligned} \tag{2.6}$$

□

We see a beautiful duality here. Product in the real domain is the convolution in the Fourier domain, and *vice versa*.

Theorem 2.4 (Wiener–Khinchin theorem). The Fourier transform of the autocorrelation of a function is its *power spectral intensity*.

$$\mathcal{F}[f \otimes f] = \left| \tilde{f} \right|^2. \quad (2.7)$$

Proof.

$$\begin{aligned} \mathcal{F}[f \otimes f](\mathbf{k}) &= \int d^n \mathbf{x} e^{2\pi i \mathbf{k} \cdot \mathbf{x}} \int d^n \mathbf{s} f^*(\mathbf{s}) f(\mathbf{x} + \mathbf{s}) \\ &= \int d^n \mathbf{s} f^*(\mathbf{s}) \int d^n \mathbf{x} e^{2\pi i \mathbf{k} \cdot \mathbf{x}} f(\mathbf{x} + \mathbf{s}) \\ &= \int d^n \mathbf{s} f^*(\mathbf{s}) \int d^n \mathbf{u} e^{2\pi i \mathbf{k} \cdot (\mathbf{u} - \mathbf{s})} f(\mathbf{u}) \quad \text{define } \mathbf{u} = \mathbf{x} + \mathbf{s} \\ &= \int d^n \mathbf{s} e^{-2\pi i \mathbf{k} \cdot \mathbf{s}} f^*(\mathbf{s}) \int d^n \mathbf{u} e^{2\pi i \mathbf{k} \cdot \mathbf{u}} f(\mathbf{u}) \\ &= \tilde{f}^*(\mathbf{k}) \tilde{f}(\mathbf{k}) = \left| \tilde{f}(\mathbf{k}) \right|^2. \end{aligned} \quad (2.8)$$

□

Finally, if we apply both sides on the Wiener–Khinchin theorem, we get

$$\mathcal{I} \left[\left| \tilde{f} \right|^2 \right] = f \otimes f. \quad (2.9)$$

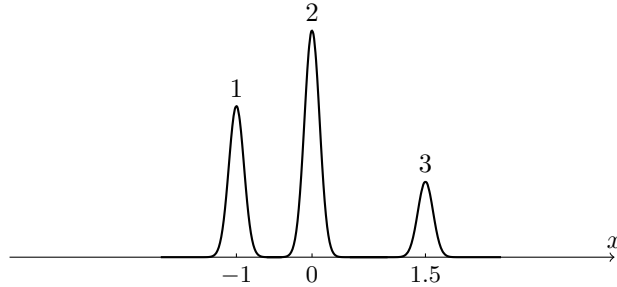
We can apply this to our diffraction intensity! We previously identified the structure factor to be the Fourier transform of the electron density, $F(\mathbf{S}) = \mathcal{F}[\rho(\mathbf{x})]$, and the diffraction intensity is exactly $I(\mathbf{S}) = |F(\mathbf{S})|^2$. Therefore, the inverse Fourier transform of the diffraction intensity is exactly the autocorrelation of the electron density!

$$\mathcal{I}[I(\mathbf{S})] = \rho \otimes \rho(\mathbf{r}). \quad (2.10)$$

We often call this the *Patterson function*, denoted $P(\mathbf{r})$.

2.3 Illustrating the Patterson Function

The above derivation is a bit abstract. Let's do a simple 1D example of an electron density function with three peaks, modelling three atoms in a line, and see how its Patterson function look like.

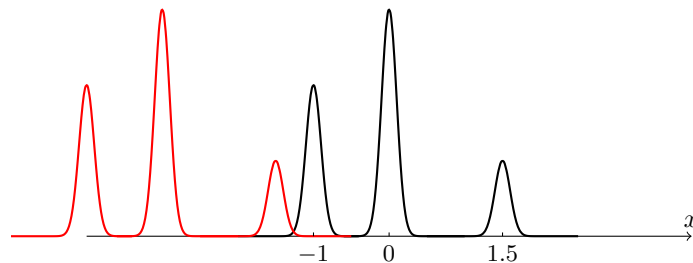


We put the three peaks at positions $x = -1$, $x = 0$ and $x = 1.5$ with weights 2, 3 and 1 respectively. From the definition of the Patterson function, which is the correlation of the electron density with itself

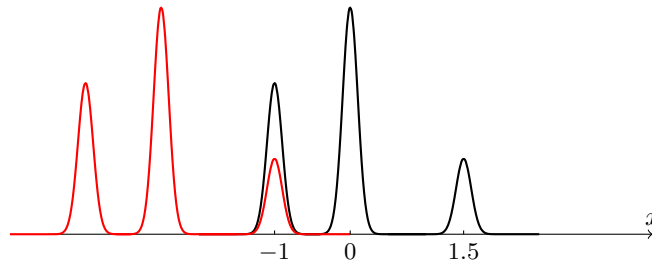
$$P(u) = \int dx \rho(x)\rho(x+u), \quad (2.11)$$

we need to take another copy of electron density function, shift it by some distance u , and evaluate the integral of their product, which is roughly governed by how well do the peaks of the two functions overlap. Let's set u to a range of different values and see what happens.

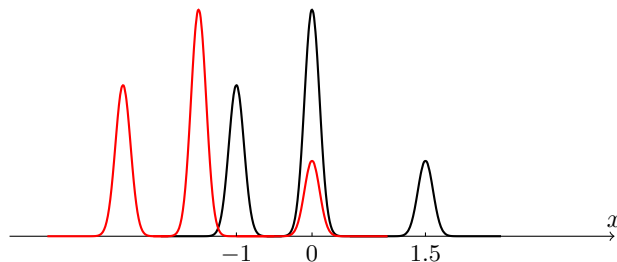
- $u = -3$: First, when u is really small, the two function has almost no overlap, so at this value of u , $F(u) \approx 0$.

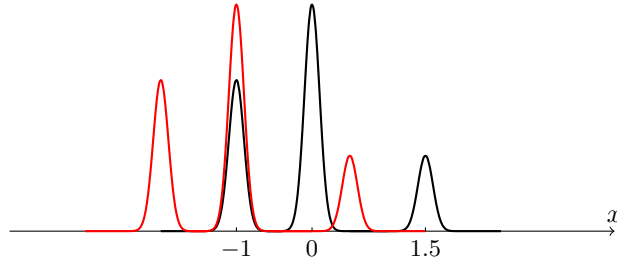


- $u = -2.5$: At this value of u , peak 1 and peak 3 just overlaps. Without the actual functional form of the peak, we can't evaluate this integral of their product, but since peak 1 has weight 2 and peak 3 has weight 1, we can roughly assign a relative weight of $2 \times 1 = 2$ to the Patterson function here.

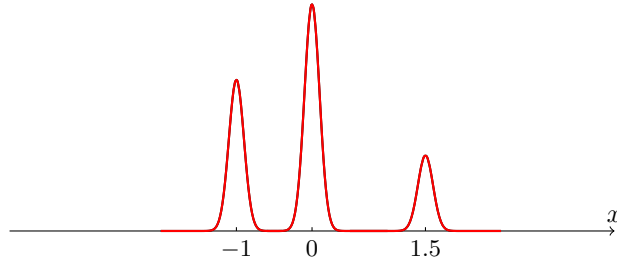


- $u = -1.5$: Now at this value of u , peak 3 and peak 2 overlaps, giving a peak in the Patterson function with relative weight $3 \times 1 = 3$.





- $u = -1$: Peak 1 and peak 2 overlaps, giving a peak in the Patterson function with relative weight $2 \times 3 = 6$.
- $u = 0$: At this special value of u , the second function is not moved at all, so all the three peaks are perfectly aligned. This leads to a very large peak in the Patterson function with weight $2^2 + 3^2 + 1^2 = 14$.



You can go further to check other values of u , but we should have done enough to discover a pattern. The Patterson function will have peaks at the values when the are two peaks in the electron density function align, which are exactly the “inter-peak” distances in the electron density. The relative weight of the peak in the Patterson function is the product of the weight of the peaks in the electron density.

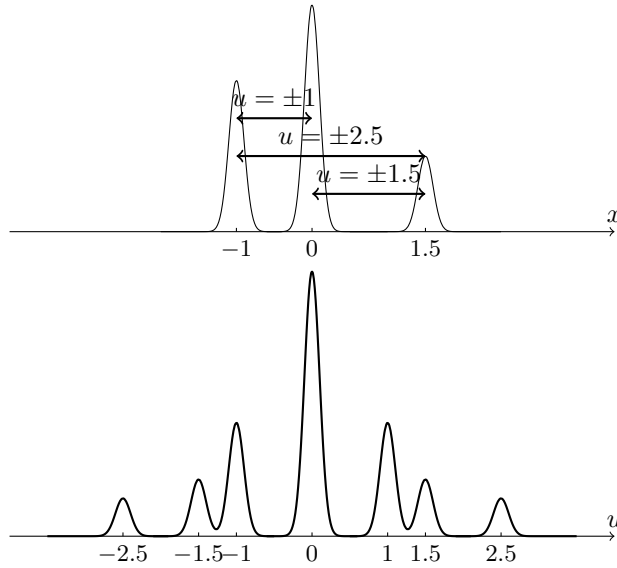


Figure 2.4: Peaks in the Patterson function corresponds to the inter-peak vectors in the electron density function.

To conclude, let's summarise some features of the Patterson function:

1. The Patterson function is defined in the same space as $\rho(\mathbf{r})$ (as opposed to the structure factor defined in the reciprocal space), and it spans a range twice that spanned by $\rho(\mathbf{r})$.
2. The Patterson function is always centrosymmetric, even if $\rho(\mathbf{r})$ is not.
3. The maximum of the Patterson function occurs at $\mathbf{u} = \mathbf{0}$, and its value is

$$P(\mathbf{0}) = \int d^3\mathbf{r} \rho^2(\mathbf{r}). \quad (2.12)$$

4. Local maxima in $P(\mathbf{u})$ occurs when peaks in $\rho(\mathbf{r})$ overlaps with those in $\rho(\mathbf{u} + \mathbf{r})$. This happens if the offset vector \mathbf{u} is some peak \leftrightarrow peak vector in $\rho(\mathbf{r})$. If we take an atomic picture of the electron density, in which electron densities are concentrated on the atomic sites (we will formalise this latter), then peaks in $P(\mathbf{u})$ occurs when \mathbf{u} is an *interatomic vector*.
5. If there are N peaks in $\rho(\mathbf{r})$, there are N^2 interatomic vectors. Of these, N will occur at $\mathbf{u} = \mathbf{0}$, corresponding to vectors of an atom pointing itself. Consequently, there will in general be $N^2 - N$ non-origin peaks in $P(\mathbf{u})$ (some of them may overlap as well).
6. For a function $\rho(\mathbf{r})$ consists of resolved peaks, the total weight (area) of the peak produced in $P(\mathbf{u})$ is proportional to the product of the weights of the peaks in $\rho(\mathbf{r})$ and $\rho(\mathbf{r} + \mathbf{u})$ producing it. If atoms in $\rho(\mathbf{r})$ and $\rho(\mathbf{r} + \mathbf{u})$ have atomic numbers Z_1 and Z_2 , then the weight of the peaks in $P(\mathbf{u})$ will be proportional to $Z_1 Z_2$. Consequently, Patterson peaks involving heavy atoms (where Z_1 and/or Z_2 is large) will show up most clearly.
7. Peaks in $P(\mathbf{u})$ will be broadened than in $\rho(\mathbf{r})$, and peaks corresponding to vectors from light atoms may merge into a poorly-defined backgrounds.

Hence, as we will see later, the Patterson function will be helpful in locating heavy atoms in a structure.

2.4 Patterson Function of a Molecule

Now consider a SO_3 molecule. For simplicity, we will make this problem 2 dimensional. The atomic number of S and O are 16 and 8 respectively, so the electron density function has one peak of weight 16, which we set at the origin, and three peaks of weight 8 at the same distance from the origin, separated by 120° .

To evaluate the Patterson function, we need to find all the interatomic vectors. Let's first consider vectors from the central S atom. There will be a self-vector ($\text{S} \rightarrow \text{S}$) at the origin, with weight $16 \times 16 = 256$, and there will be three $\text{S} \rightarrow \text{O}$ vectors arranged symmetrically around origin, with weight $16 \times 8 = 128$ each. Their distance to the origin is the same as the S – O bond length, which we will denote as r_1 . Note that this looks like the molecule itself, with the S atom sitting at the origin.

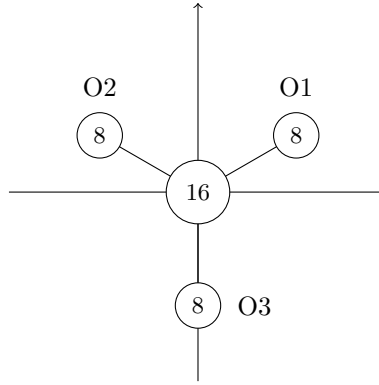


Figure 2.5: The electron density of a SO_3 molecule.

Then consider the vectors from the atom O1. There will be a self vector ($\text{O1} \rightarrow \text{O1}$) at the origin, with weight $8 \times 8 = 64$. There are two other $\text{O} \rightarrow \text{O}$ vectors, $\text{O1} \rightarrow \text{O2}$ and $\text{O1} \rightarrow \text{O3}$, at 60° to each other, with weight 64 and length $\sqrt{3}r_1$. Bisecting these two is a $\text{S} \rightarrow \text{O1}$ vector with weight 128 and length r_1 . Note this again look like the molecule itself, but now with O1 sitting at the origin. Similarly, we can consider vectors from atoms O2 and O3 as well. They are shown in the figure below.

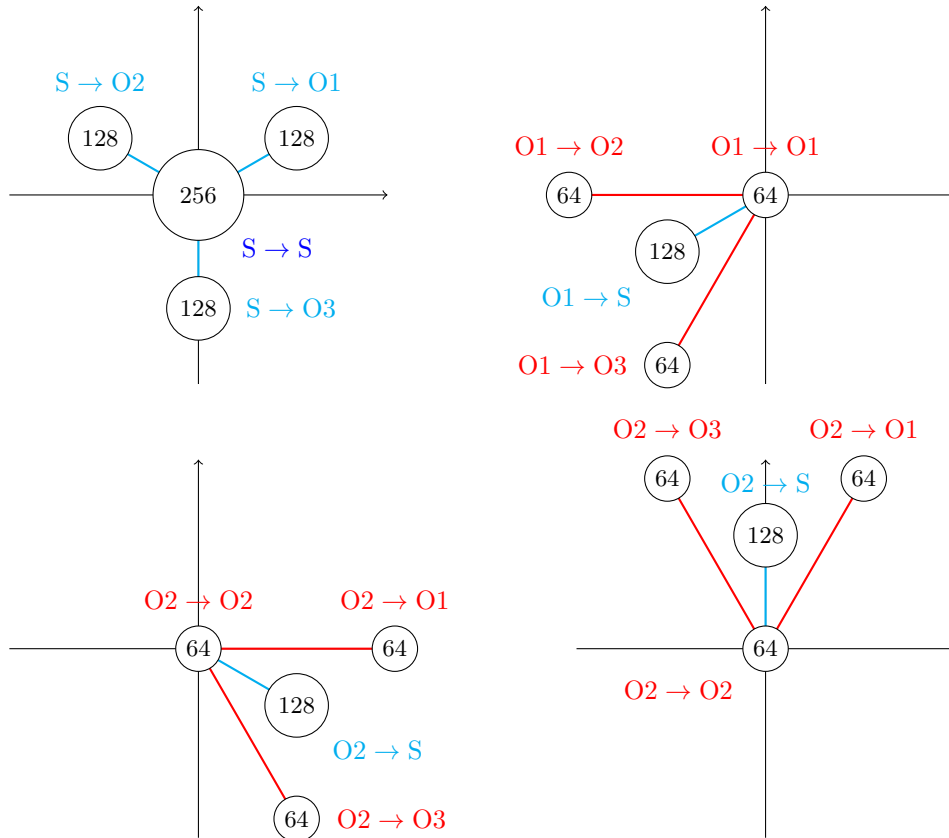


Figure 2.6: Interatomic vectors of SO_3 .

Then to construct the Patterson function, we only need to add these contributions up. The result is shown in the figure below.

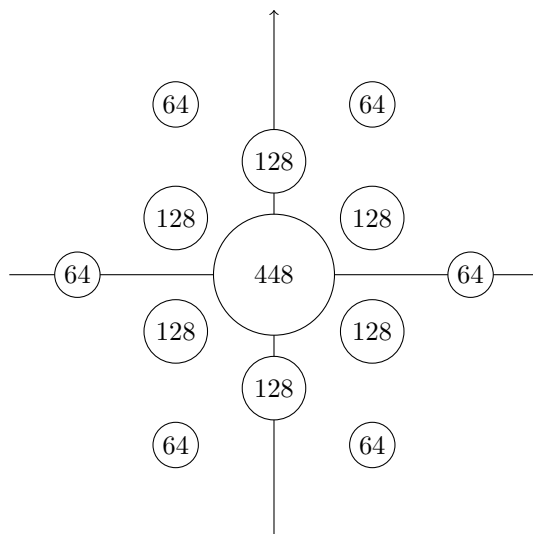


Figure 2.7: The Patterson function of a SO_3 molecule.

Some remarks:

- The weight of the origin is the sum of the squares the atomic numbers of all atoms.
- The Patterson function includes vectors between all atoms, not just between the chemically bonded ones.
- $P(\mathbf{u})$ must have a centre of symmetry at the origin, even though SO_3 does not have one. The point group of the SO_3 molecule is D_{3h} , and the centre of symmetry have upgraded the point group of the Patterson function to D_{6h} .

One can use the same principle to construct the Patterson of a 3D molecule. Below an example of the Patterson function of SO_4^{2-} ion, with T_d symmetry. The Patterson function has an extra inversion centre, so its point group is O_h .

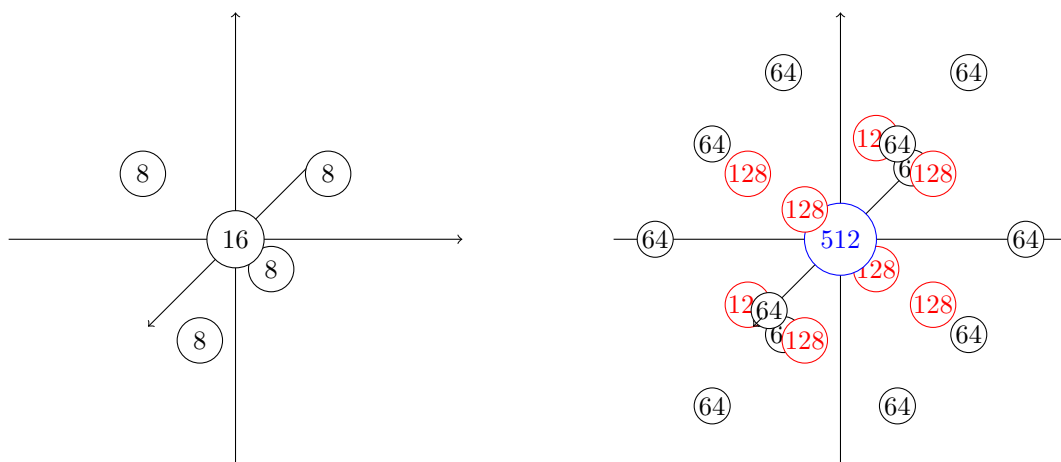


Figure 2.8: The electron density and the Patterson function of SO_4^{2-} .

2.5 The Patterson Function of Ideal Gas

We now move on to a more realistic case where there is a whole collection of molecules, and we will consider two extremes — an ideal gas and a crystalline solid.

First, let's consider an ideal gas. There are no intermolecular forces, negligible molecular volumes, and we will assume that all molecules are perfectly rigid and geometrically identical. The Patterson peaks will have contribution from all molecules: both intermolecular vectors and intramolecular vectors. For the vectors pointing within a molecule, the interatomic vectors should be identical to that for a single molecule, multiplied by the total number of molecules that are present. But as we are in the gas phase, we have another complexity. Due to the random molecular orientation, the Patterson function of an ideal gas appears as a spherically averaged version of the 3D Patterson that we constructed for one molecule. Thus, only the length of the interatomic vector is important, and we will plot the resulting Patterson function as a radial vector map. For example, for the SO_4^{2-} ideal gas (which apparently will not be ideal in reality), the radial Patterson function is shown in the figure below.⁶

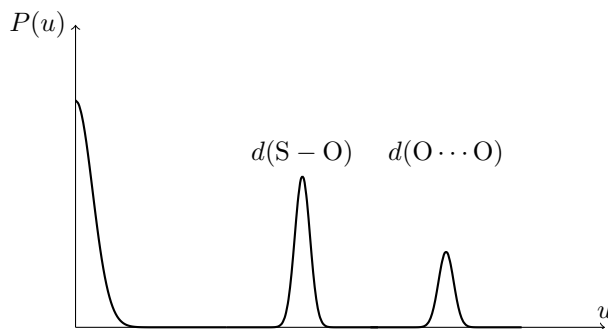


Figure 2.9: Intramolecular contribution to the radial Patterson function of SO_4^{2-} .

We then need to consider the intermolecular vectors. If the gas is indeed ideal with no intermolecular forces, then the positions and orientations of all molecules should be completely random at any time. Hence, all distances between molecules are equally possible, and so the intermolecular vectors will sum over time to form a constant background. If we are going to inject a little bit of reality into this crude model, we can say that the molecules cannot be too close from each other, so this background is zero at small u .

As a result, in an ideal gas, the radial Patterson function has discernible peaks corresponding solely to the spherical averaged Patterson function of one molecule, superimposed on a featureless background. Non-ideal gases, liquid and non-crystalline solid will show similar radial Patterson function due to their disorderedness. However, they may show local coordination peaks.

⁶As for the atomic wavefunction, you can choose to plot a radial Patterson function or a radial distribution function — we are being rather loose here since we are doing everything schematically. Anyway, the true weight of a peak is the integral of the 3D Patterson function.

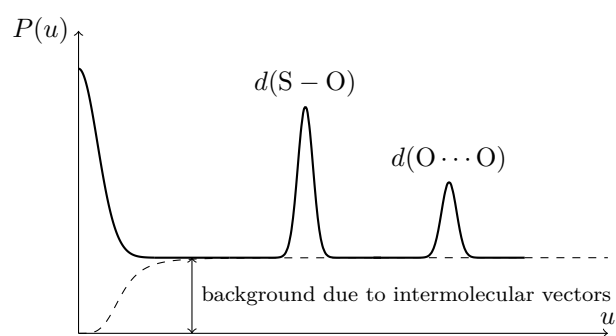


Figure 2.10: Radial Patterson function of SO_4^{2-} with both intermolecular and intramolecular contributions.