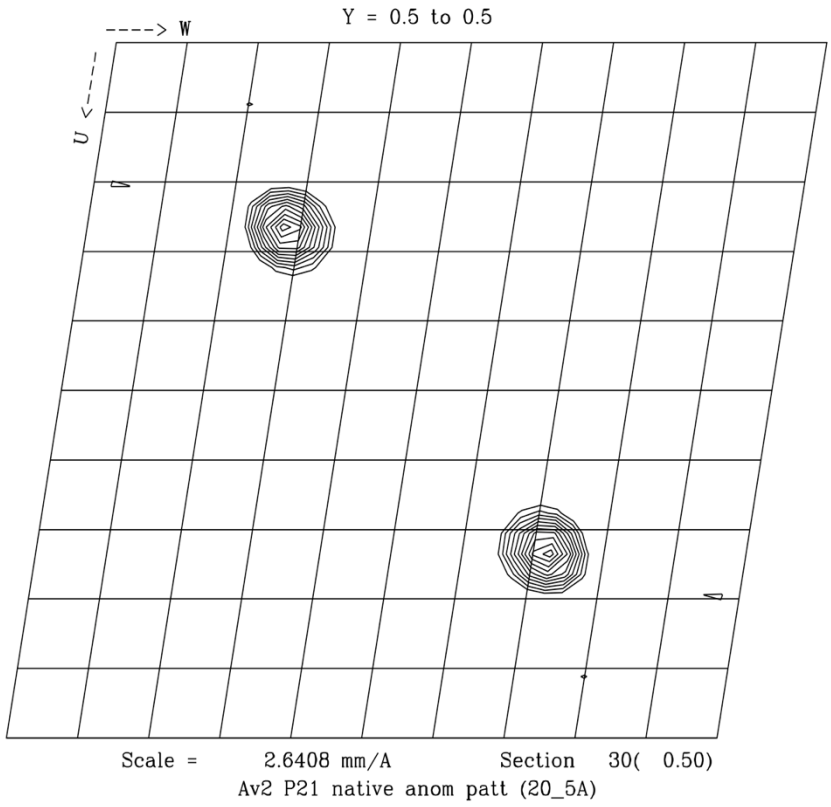
Notes on Macromolecular Crystallography from the Turn of the Century



Douglas C. Rees
California Institute of Technology
dcrees@caltech.edu

Table of Contents

Introduction and Overview 3
Chapter 1: Introduction to Lattices and Scattering 4
Chapter 2: Patterson Methods 32
Chapter 3: Phasing Methods in Macromolecular Crystallography 41
Chapter 4: Non-crystallographic Symmetry Averaging and Molecular Replacement 53
Chapter 5: Structure Refinement 64

Introduction and Overview

I started these notes in the mid-1980s when I taught an informal course on the foundations of macromolecular crystallography for interested students and postdocs at UCLA. They were not intended as a comprehensive and systematic overview of crystallography, but rather as an introduction to some of the more mysterious or challenging aspects encountered in these studies that were of particular interest to me at that time – specifically, the relationship between X-ray scattering and structure, how to find heavy atoms and calculate phases, how to identify any non-crystallography symmetry operators, and how to refine an initial set of parameters. Phasing was generally the rate determining step in that era and so these activities naturally occupied a significant amount of the crystallographer's time (along with preparing suitable crystals in the first place). The foundational material discussed in these notes was motivated by problems that arose in our research and the examples used to illustrate various topics were often based on these experiences. I am the first to acknowledge that many important topics, including data collection, data processing, model building and model refinement, are not discussed in the depth they merit, primarily because I was unable to improve upon the available programs and so I focused my efforts on phasing the projects under investigation in the group.

These notes reflect the state of macromolecular crystallography in the early 2000s. Beginning at that time, the availability of MAD phasing methods, increasing numbers of models for molecular replacement, superb synchrotron beamlines and remarkable software packages, revolutionized the practice of macromolecular crystallography. Consequently, the ability to solve heavy atom derivatives from Patterson maps was no longer a critical survival skill. The subsequent revolutions in cryo-electron microscopy and computation (i.e., AlphaFold) have completely transformed structural biology. Not coincidentally, the driving force in structural biology has largely changed from an emphasis on solving the "first" in a family of structures, to an emphasis on the "function" part of the structure-function paradigm exemplified from the earliest days of the field by Max Perutz. As a result of these developments, there is no longer any compelling need to update the material in these notes, particularly the references, and so they reflect the state of the macromolecular crystallography at the turn of the century.

I will take this opportunity to recognize the remarkable teachers of macromolecular structure that I had as a graduate student. Steve Harrison and Don Wiley taught the Biochemistry 112 course that opened the world of structural biology to me; the scattering treatment has been heavily influenced by their approach, including my appreciation for the Atlas of Optical Transforms. My graduate advisor William Lipscomb ("The Colonel") had seen it all and could illuminate any aspect of crystallography, particularly space groups and interpreting Patterson maps. Mitch Lewis has been an incredible friend and resource, having studied crystallography at the source and bringing his expertise to the New World. The subsequent opportunities to teach crystallography with my colleagues David Eisenberg, Pamela Bjorkman and Bil Clemons provided essential "real world" experience trying to convey this material, with a highlight being Pamela's lecture on "how to read a structure paper" that condensed all the lessons and take-home messages into one compelling presentation (and yet another important topic not discussed in these notes).

I would especially like to acknowledge the incredible work and passion of the graduate students, postdoctoral fellows and staff in my research group who have provided the driving force for these studies and taken us in directions I would never have dreamed possible when I started these notes. Thank you.

Chapter 1: Introduction to Lattices and Scattering

General references

A. Guinier (1963) **X-ray Diffraction in Crystals, Imperfect Crystals and Amorphous Bodies**, Freeman. Excellent presentation of the general theories for X-ray scattering applied to any systems.

R.W. James (1948) **The Optical Principles of the Diffraction of X-rays**, Ox Bow Press (reprinted 1982). Classic treatment of X-ray diffraction.

Steve Harrison recommended both Guinier and James for background reading on X-ray diffraction when I started my rotation and they are still invaluable resources after nearly 50 years.

C. Giacovazzo, ed. (1992) **Fundamentals of Crystallography**, Oxford. Great overview of the foundations of crystallography.

J. Drenth (1994) **Principles of Protein X-ray Crystallography**, Springer. Clear and concise treatment of macromolecular crystallography by one of the early practitioners.

B. Rupp (2010) **Biomolecular Crystallography**, Garland. Complete and accessible treatment of the theory and practice of macromolecular crystallography.

International Tables of Crystallography, Vols A, B, C, F; (plus vols I, II, III, IV of the previous edition). The standard reference works for crystallography.

Introduction

When electromagnetic radiation interacts with matter, two things can happen – absorption and scattering. Many spectroscopic methods, including NMR and EXAFS, are based on absorption effects that can yield useful structural information. The most powerful methods for determining structures (X-ray, neutron and electron diffraction), however, involve scattering measurements (although we will also briefly touch on the consequences of absorption in these experiments, too). Not surprisingly, the wavelength of the incident radiation specifies the "resolution" or degree of spatial detail that can be obtained. Since we are often (although not exclusively) interested in atomic level structure, we will need radiation with wavelengths in the Å range. (This statement is true for X-rays and neutrons, but is not correct for electron crystallography). The basic experimental design is to shine X-rays (or electrons or neutrons) on the sample, measure the scattering by means of a suitable detector, and then infer (by methods to be described) the structure of the material that did the scattering. If the sample has a collection of molecules in random orientations, then only an orientationally averaged structure can be obtained directly from the experiment. But, if the sample is ordered (as in a crystal), then an unaveraged molecular structure can be determined - this is the heart of the most powerful applications of scattering methods in structural biology.

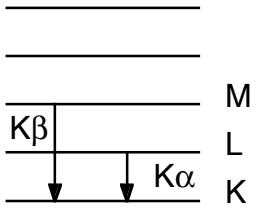
X-rays

SAFETY CONSIDERATIONS

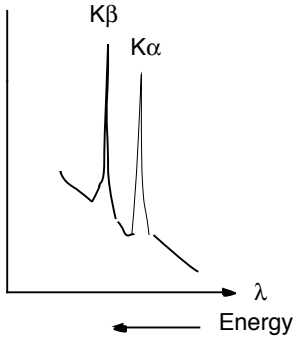
Extreme care MUST be used when working around X-ray generating equipment. Always check the status of the shutter before manipulating crystals, adjusting beam stops, etc. Even when the shutter is closed, assume it is open and work accordingly. Constant vigilance must be maintained to prevent irradiation – the highly collimated nature of the beam and the relatively low energy of the X-rays we use can produce severe burns. NEVER leave disassembled equipment such that the next person can be inadvertently exposed to X-rays. Interlocks, while usually reliable, can fail, leading to irradiation of unsuspecting people.

X-ray sources

Where do X-rays come from? There are several ways to answer this question. In the lab, hospital, etc., X-rays are typically generated by accelerating electrons against a metal target, resulting in the ejection of electrons from the metal. If inner shell electrons are ejected, then electrons from outer shells can drop down, with the consequent emission of a photon (X-rays).



If the electrons are accelerated with a voltage V_{acc}, then the highest energy X-ray (shortest λ) is given by $\lambda = 12{,}398\text{Å}/V_{acc}$. Typically, $V_{acc} \sim 50\text{kV}$ or $\lambda_{min} \sim 0.25$ Å. If a K shell electron is ejected, then electrons dropping from L $\rightarrow$ K and M $\rightarrow$ K give rise to Kα and Kβ X-rays, respectively. The emission spectra for a typical metal looks something like this:



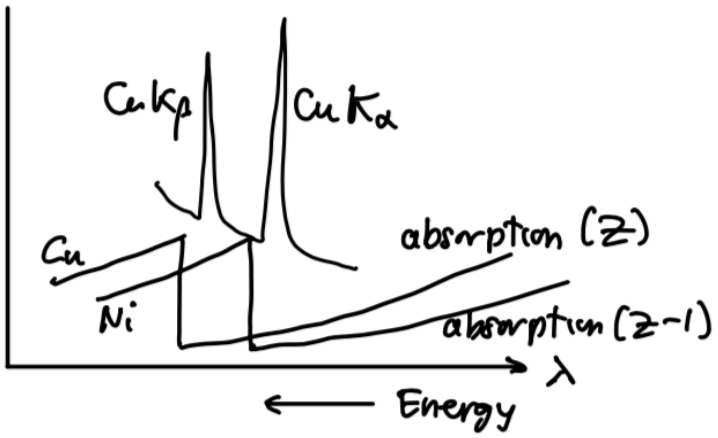
For Cu, λ for Kα, Kβ = 1.54Å, 1.38Å, respectively, which correspond to minimum V_{acc}'s of 8.05 kV and 8.98 kV, respectively. These lines are actually doublets due to the presence of two electronic configurations that differ slightly in energy. The positions of the emission lines are element specific; for example, for Mo, λ Kα = 0.71Å. Consequently, some control over λ is possible by suitable choice of target.

Often (although not always) it is desirable to have a monochromatic source of X-rays, and the emission spectra from the metal is not such a source. Some general approaches to cleaning up conventional X-ray sources include:

monochromator: put a crystal, such as graphite, in the beam, and use diffraction in a particular 2θ direction to get the desired λ. The spacing between graphite layers is 3.35Å, so to get $\lambda = 1.54$Å (CuKα radiation), $2\theta = 2\sin^{-1}(\lambda/2d) = 2\sin^{-1}(1.54/2*3.35) = 26.6^\circ$ (this equation will be derived later).

filter: not only do metals emit X-rays, but they absorb them, too. The absorption spectra is shifted to just slightly higher energy (lower λ) than the emission spectra, since the X-rays need to be sufficiently energetic to knock electrons out of an orbital. The absorption spectra for element Z-1 is shifted to lower energy (longer λ) from element Z, since the electrons are less tightly bound. It

so happens that element Z-1 can be used to filter Kβ radiation from Kα for element Z. Example: use of a Ni filter with Cu radiation, as illustrated in the following schematic.



mirrors: Although X-ray lenses do not exist (since the refractive index of most materials to X-rays is ~1), it is possible to manipulate X-rays by total internal (grazing incidence) reflection from appropriate mirrors. The angle of grazing incidence depends on the wavelength, so that it is possible to separate Kα and Kβ lines, for example (see James, pp. 171-176).

Synchrotron radiation

High-power, tunable X-ray sources have been developed at **synchrotron** radiation facilities, as a consequence of the acceleration of charged particles (typically electrons or positrons) maintained in an approximately circular orbit through the dipole magnets of a storage ring. First generation synchrotron sources were based on bending magnets that produce radiation with a spectral distribution over a wide range of energy. The primary purpose of these bending magnets is to maintain the circulating electron or positron bunches, by bending the particles into a closed orbit. Second and third generation sources contain magnetic devices called "insertion devices" which are inserted into straight sections in a storage ring. These devices, either wigglers (second generation) or undulators (third generation), are much more efficient radiation sources than bending magnets, and do not produce a net displacement or deflection of the stored beam. Monochromators and mirrors are used to manipulate the beam, although the high power loads greatly complicate design of these devices (in 3rd generation sources, liquid nitrogen is needed to cool the mirrors, etc.). (see BESAC report for more details)

The intensity of these sources has made it possible to collect high quality data from small crystals, and the tunable wavelength nature permits the application of powerful phasing methods based on anomalous dispersion.

Back to the original question – where do X-rays come from? Despite the preceding discussion, in the US, X-rays come from the Department of Energy, and in other countries from the DOE's functional counterparts. Simply stated, X-rays are generated by the organizations that pay the power bills, etc. for synchrotrons and the staff at these facilities deserve our gratitude and support.

Crystals and Crystallization

Crystals are generated by translational repeats of a basic building block (the unit cell), and so contain many repeats of the same unit in a specific orientation. Although biological X-ray diffraction studies generally use three-dimensionally ordered crystals, two-dimensional crystals can be studied by electron microscopy/diffraction, and one-dimensional crystals have been used to study the electron density profiles of membrane multilayers.

The basic idea behind macromolecular crystallization is to bring a macromolecular solution to slight supersaturation, and hope that crystals appear rather than amorphous precipitant. The main peculiarities of macromolecular crystallizations, as opposed to salt, etc., are related to sample stability - have to use T, pH, precipitants, etc. that don't denature the sample.

Three basic approaches are used for macromolecular crystallizations: vapor diffusion, dialysis and batch methods.

vapor diffusion: The sample with protein is physically separated from a reservoir solution - only water or volatile components can go between solutions until equilibrium is reached. Two general experimental designs: sitting drops (10-20 μl samples) and hanging drop (~2 μl samples). With robotic systems, 10-100x smaller volumes can be used in crystallization trials.

dialysis: protein solution is contained in a semipermeable membrane surrounded by a solution. Water and low molecular weight compounds (salts) are permeable, but not protein. Sample volumes ~20 μl and up. This technique is not used so much these days, since it is a pain to set up, but there are a lot of advantages to dialysis methods and they should be used more often. The Colonel would remind us that bigger volumes can give bigger crystals.

batch: protein and precipitant solution are placed together in a closed container or covered with a layer of oil - either mixed, or layered in small tubes or capillaries. Useful for anaerobic work.

Typical crystallization variables include protein concentration, precipitant(s), pH (buffer), ionic strength (salt), temperature. Protein concentration ~10 mg/ml; precipitants such as $(NH_4)_2SO_4$, polyethylene glycols (PEG), methylpentanediol (MPD), molybdate; pH ~ 5-10, depending on stability; ionic strength - no salt to highly concentrated; temperature, typically 4˚ to 20˚; and miscellaneous additives (heavy atoms, polyamines, ligands). Of course, these days, Hampton kits are used for screening crystallization conditions. For membrane proteins, also need detergents.

It can be sometimes difficult to distinguish macromolecular crystals from salt (phosphate, calcium, etc.) crystals (this is an advantage of using colored metalloproteins!). Typical tests involve diffraction (size of unit cell), birefringence (usually proteins are and salts aren't, since proteins are chiral), the "crush" test (like the Salem witch test), the use of dyes such as “Izit”, and, more recently, a UV fluorescence microscope to detect tryptophan fluorescence.

Soluble protein crystals typically contain ~50% solvent corresponding to a Matthews coefficient V_m (= unit cell volume ($Å^3$) divided by the total molecular weight of protein in the unit cell) of ~2.4 $Å^3$/Dalton - "exact" content can be found by measuring the crystal density (although this is rarely done). Membrane proteins typically have a higher solvent content and V_m ~ 4.0 $Å^3$/Dalton.

Some useful relationships between V_m, solvent content , the number of molecules in the unit cell and the crystal density may be derived as follows (B.W. Matthews JMB 82, 513-526 (1974)), with ρ_c, ρ_p, ρ_s denoting the density (gm/cm^3) of the crystal, protein and solvent; V_c, V_p, V_s corresponding to the volumes (cm^3) per unit cell; n = number of molecules per unit cell and M = molecular weight (gm/mole). The crystal density can be calculated from the wet and dry weight of crystals or from the equilibrium position in a calibrated density gradient of organic solvents.

$$V_m = \frac{V_C \times 10^{24}}{nM} = \frac{V_C N_A}{nM}\frac{10^{24}}{N_A} = \frac{V_C N_A}{nM}\frac{1}{0.6023}$$

$\rho_c V_c, \rho_p V_p, \rho_s V_s =$ weight of crystal, protein (neglect bound water), solvent

$\rho_c V_c = \rho_p V_p + \rho_s V_s$ conservation of mass

$V_c = V_p + V_s$ conservation of volume

$$V_p = V_c \frac{\rho_c - \rho_s}{\rho_p - \rho_s}$$

$$nM = N_A V_P \rho_p = N_A V_c \frac{\rho_c - \rho_s}{1 - \bar{v}_P \rho_s} \quad \text{with } \bar{v}_P = 1/\rho_p \sim 0.74 \text{ cm}^3/\text{gm}$$

$$n = \frac{N_A V_c}{M}\frac{\rho_c - \rho_s}{1 - \bar{v}_P \rho_s}$$ determination of n from crystal density

$$nM = N_A V_P \rho_p = N_A V_C (1 - X_S)\rho_p \quad \text{with } X_S = \text{frac. solvent}$$

$$\frac{nM}{N_A V_c} = (1 - X_S)\rho_p = \frac{1}{0.6023 V_M}$$

$$X_S = 1 - \frac{1}{0.6023 \rho_p V_M} \approx 1 - \frac{1.23}{V_M} \quad \text{with } \rho_p = 1.35 \text{ gm/cm}^3$$

So, typical water soluble and membrane protein crystals (V_m ~ 2.4 and 4.0 Å^3/dalton) are ~49% and 69% solvent, respectively.

A nice size crystal is ~ a tenth of a millimeter on each side, say 0.1 x 0.1 x 0.1 mm^3. A typical protein density is 1.3 gm/cm^3 = 1.3 mg/mm^3. So, a protein of molecular weight M has a specific volume of [(M gm)/6.02x10^{23}] [cm^3/1.35 gm] [10^{24}Å^3/cm^3] / M = 1.23 Å^3/dalton, or about half the Matthew's coefficient (ie, crystals are ~50% protein). Now, this typical crystal has a volume of 0.001 mm^3, which contains ~0.001 mm^3 x 1.35 mg/mm^3 x 1/2 ~ 0.7 μg protein. For M = 70,000, this is about .01 nmole protein or ~10^{13} molecules (equivalent to a concentration of ~10 mM protein). 1 μg protein ~ 0.1 μl of a 10mg/ml protein solution, so that usable crystals can be grown from 100 nl hanging drops.

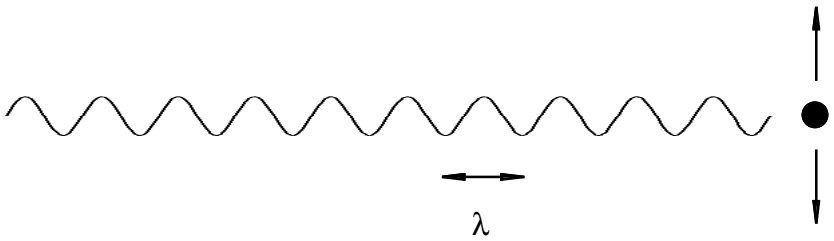
Odds of success? I would estimate these as ~50% if you have reasonable amounts of pure, water soluble proteins. Typical ways of improving odds - screening for different crystallization conditions and changing macromolecular specimens. Approaches to improve odds include (a) factorial screens (exorbitantly overpriced kits); (b) screening multiple homologs, (c) adding ligands (inhibitors, Fvs, to form complexes etc.); (d) adding other additives; (e) mutating residues at potential contact sites (Glu/Lys to Ala) (f) proteolysis or truncated forms; (g) crystallization crosslinkers (polyamines, polytungstates); and (h) incorporate crystallization sites (fusion proteins) = desperation time.

Two last bits of advice:
(1) even though crystallization is a purification technique, in general, **the purer the protein, the better the crystals** (with some notable exceptions, particularly involving membrane proteins and phospholipids)
2) it is essential to **check that you've crystallized the molecule you think you have**! Gels and mass spectrometry can be useful analytical tools in this regard.

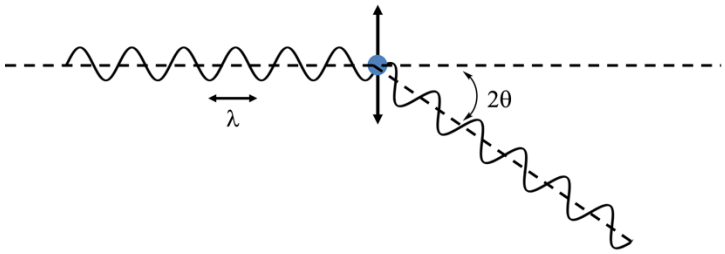
Diffraction Theory

(the section on "waves and complex variables" in the Mathematical overview may be helpful).

Consider electromagnetic radiation that is incident on an object much smaller than λ.



This interaction induces an electric dipole in the object. Neglecting absorption, re-radiation or scattering of the radiation occurs at the same λ:



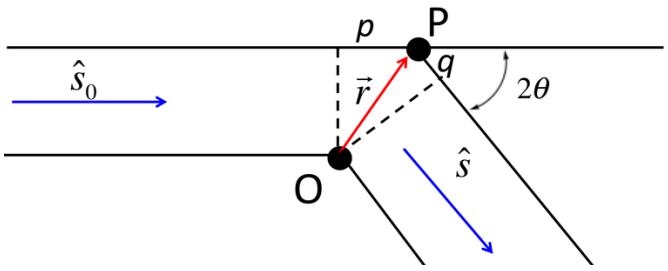
The fraction of unpolarized radiation scattered at an angle 2θ relative to the incident beam is given by the classical Rayleigh expression for the elastic scattering from bound electrons:

$$\frac{I}{I_0} = \frac{8\pi^4\alpha^2}{r^2\lambda^4}\left(1+\cos^2 2\vartheta\right)$$

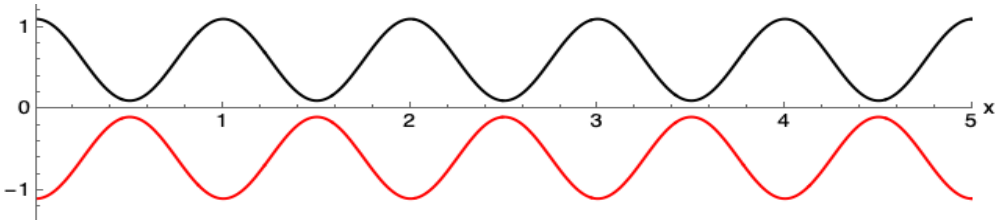
where α is the polarizability of the scattering object.

Unfortunately, this doesn't help us too much, since to get useful atomic information, we need to consider the case where the object is much larger than λ (for example, proteins are typically ~100 Å, while λ ~1.5 Å). So, our treatment has to be expanded to include large molecules.

Consider scattering of radiation from a point at the origin, O, and from a point, P, at a position $\vec{r}$ from the origin. The scattering from each point individually is given by the Rayleigh scattering expression - we need to see how to combine these individual contributions. Let $\hat{s}_0$ be the direction of the incident radiation (assumed coherent over a sufficiently large spatial region), and $\hat{s}$ be the direction of the scattered radiation, which forms an angle 2θ relative to the incident beam.



Recall 2 waves of same λ



These waves are identical, except for a phase shift. In this particular case, the waves are 180° out of phase. This phase factor can be written as $e^{i\phi}$ where $i = \sqrt{-1}$. When $\phi = 180° = \pi$ radians, $e^{i\phi} = -1$. Going back to the original problem, let's assume that all the incoming radiation ($\hat{s}_0$) is in phase. How could $\hat{s}$ be out of phase? Because different path lengths are traveled! The difference in path length is equal to the excess in path length 2 - the excess in path length 1. This can be used to calculate the phase difference, which is **minus** number of wavelengths the radiation through P travels relative to O, -(p-q)/λ, multiplied by 2π (the change in sign is because an increase in pathlength corresponds to a decrease in phase shift, since the wave on the longer path will be behind the waves on the shorter path). Now

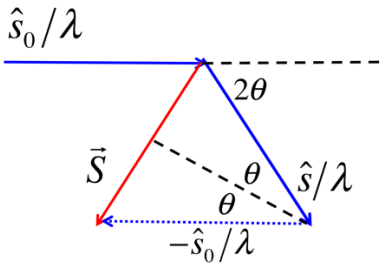
$$p = \vec{r} \cdot \hat{s}_0 \text{ and } q = \vec{r} \cdot \hat{s}$$

and the phase shift-term becomes

$$e^{-2\pi i \frac{(p+q)}{\lambda}} = e^{-2\pi i \frac{(\vec{r}\cdot\hat{s}_o - \vec{r}\cdot\hat{s})}{\lambda}} = e^{2\pi i \frac{\vec{r}\cdot(\hat{s}-\hat{s}_o)}{\lambda}} \equiv e^{2\pi i \vec{r}\cdot\vec{S}}$$

where the diffraction vector $\vec{S} = \frac{(\hat{s} - \hat{s}_0)}{\lambda}$ and the phase shift-term becomes $e^{2\pi i \vec{r}\cdot\vec{S}}$.

Geometrically, the diffraction vector bisects the incident and outgoing radiation directions:



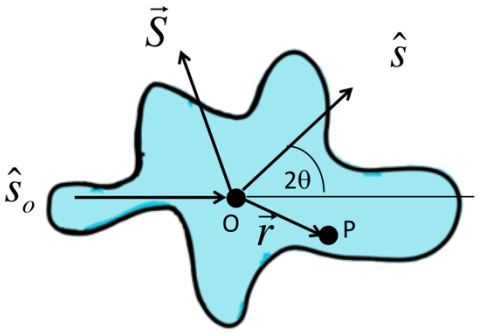
$$\vec{S} \equiv \frac{\hat{s} - \hat{s}_o}{\lambda}$$

$$\left|\vec{S}\right| = \frac{2\sin\vartheta}{\lambda} = \frac{1}{d}$$

If the scattering from point O in the direction 2θ is g, then the scattering in the same direction from the point P is $ge^{2\pi i\vec{r}\cdot\vec{S}}$, and the total scattering from the two points is

$$F\left(\vec{S}\right)=g+ge^{2\pi i\vec{r}\cdot\vec{S}}$$

Consider a scattering object



Then $F\left(\vec{S}\right)=\int ge^{2\pi i\vec{r}\cdot\vec{S}}d\vec{r}=\int\rho\left(\vec{r}\right)e^{2\pi i\vec{r}\cdot\vec{S}}d\vec{r}$

$$F\left(\vec{S}\right)=\left|F\left(\vec{S}\right)\right|e^{i\varphi_{\vec{S}}}$$
$$=\left|F\left(\vec{S}\right)\right|\left(\cos\varphi_{\vec{S}}+i\sin\varphi_{\vec{S}}\right)$$
$$=A_{\vec{S}}+iB_{\vec{S}}$$

these are variants of the fundamental diffraction equation. If ρ(r) is known, the diffraction pattern can be calculated via this relationship, which is a Fourier transform. For example, if the electron density distribution of an atom is known from a quantum mechanical calculation, the atomic form factor describing the scattering of X-rays from this atom can be calculated by (numerical or analytical) evaluation of the Fourier transform integral.

Now

F($\vec{S}$), $\rho\left(\vec{r}\right)$ are related by a **Fourier** transform

usually, we are interested in the **inverse** transform

$$\rho\left(\vec{r}\right)=\frac{1}{V}\int \mathrm{F}(\vec{S})e^{-2\pi i\vec{r}\cdot\vec{S}}d\vec{S}$$

given the diffraction pattern, this equation lets us find the electron density of an object.

Comment: Even though $\rho\left(\vec{r}\right)$ is real, the form of the Fourier transform shows that, in general $F\left(\vec{S}\right)$ is a complex number, or $F\left(\vec{S}\right)=\left|F\left(\vec{S}\right)\right|e^{i\varphi_{\vec{S}}}$, which is an amplitude times a phase term. Unfortunately, when recording the diffraction data, only the intensity $\left|F\left(\vec{S}\right)\right|^2=F\left(\vec{S}\right)F^*\left(\vec{S}\right)$ is measured - **all** the phase information is lost. As a result, it is easy to calculate F(S) given $\rho\left(\vec{r}\right)$, but can't usually calculate $\rho\left(\vec{r}\right)$ given $\left|F\left(\vec{S}\right)\right|$. This is the only part of F that is easily available

experimentally. The challenge to macromolecular crystallography is to find ϕ, so that $\rho(\vec{r})$ can be calculated. This is the origin of the so-called "phase problem".

There are two major techniques based on these scattering equations:

1. solution scattering. A collimated X-ray beam is passed through a macromolecular solution, and the scattering pattern is recorded. This is (conceptually) a simple experiment, but, since the molecules are randomly oriented, the |F| that is measured is not for a single molecule, but rather for a spherically averaged molecule. This means that only a radially averaged ρ can be determined, which can provide useful information about the size and shape of the molecule.

For problems with spherical symmetry, the Fourier transform relationship for solution scattering reduces to

$$F(\vec{S}) = \int \rho(\vec{x}) e^{2\pi i \vec{S}\cdot\vec{x}} d\vec{x}$$

$$F(S) = \int \rho(\vec{r}) e^{2\pi i \vec{S}\cdot\vec{r}} r^2 \sin\theta d\phi d\theta dr$$

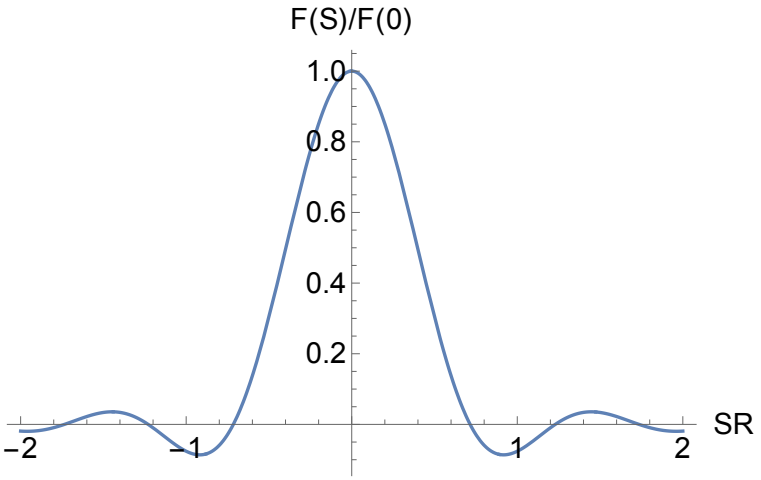
$$= \int_0^{2\pi} d\phi \int_0^{\infty} r^2 \rho(r) dr \int_0^{\pi} \sin\theta e^{2\pi i S r \cos\theta} d\theta$$

$$= \frac{2}{S} \int_0^{\infty} r\rho(r) \sin(2\pi S r) dr$$

For a uniform sphere with $\rho(r) = 1, r \le R$ and $\rho(r) = 0, r > \infty$

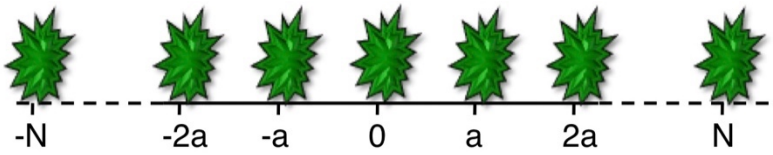
$$F(S) = \left(\frac{4\pi}{3} R^3\right) \left[\frac{3(\sin(2\pi SR) - 2\pi SR\cos(2\pi SR))}{(2\pi SR)^3}\right]$$

This function is illustrated below; the zeroes occur at SR = 0.72, 1.23, 1.73, …



A great example is provided by the small angle X-ray scattering of from bacteriophage P22 heads (Earnshaw, Casjens and Harrison, JMB 104, 387 (1976))

2. scattering from oriented sample. If there was some way to hold all the molecules in a fixed orientation relative to the beam, then the orientationally unaveraged structure could be determined. This can be achieved by crystallization. The translational symmetry of crystals assures that the orientation of one molecule with respect to another is defined throughout the entire sample. But, we can also be sure that the presence of >1 molecule in a sample will also alter the scattering pattern. Let's see how.

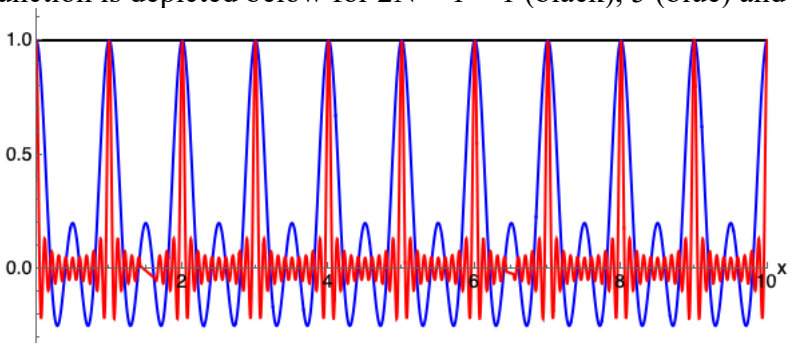


Imagine that there are 2N+1 objects equally spaced in a one-dimensional lattice. If G(S) is the transform of the objects at x = 0, then the others will be identical, except for a phase shift. For the object, at n•a, the scattering contribution relative to the original equals $G(\vec{S})e^{2\pi in\vec{a}\cdot\vec{S}}$. The total transform from all 2N+1 is

$$F(\vec{S}) = \sum_{j=-N}^{N} G(\vec{S})e^{2\pi ij\vec{a}\cdot\vec{S}} = G(\vec{S})\sum_{j=-N}^{N} e^{2\pi ij\vec{a}\cdot\vec{S}}$$

The transform is modified by the summation term, called the interference function, which does not depend on the structure, but only on the spacing and the number of objects in the crystal. The interference function is plotted below as a function of $\vec{a}\cdot\vec{S}$ and N. The maxima get very sharp as $N \to \infty$, and $\vec{a}\cdot\vec{S}$ = integer $\equiv h$. So, diffraction only occurs in certain specific directions, which depend only on the lattice, and the direction of the incident beam with respect to the lattice.

The interference function is depicted below for 2N + 1 = 1 (black), 5 (blue) and 21 (red).



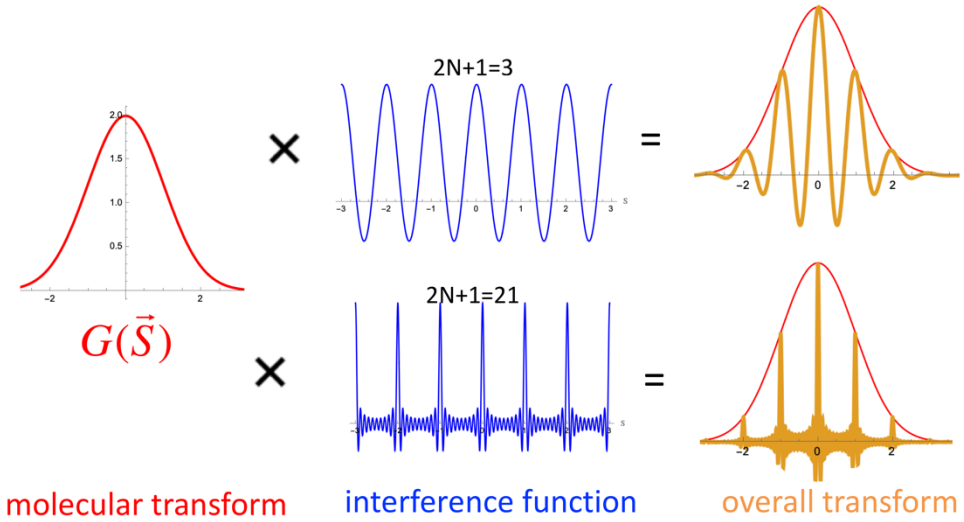
Some important properties:

1. The larger a is, the closer the spacing between allowed values of $\vec{S}$ will be. The lattice spacing a characterizes the structure in "real" space. The spacings in $\vec{S}$ are in "reciprocal" space. Large separations in one space correspond to small separations in the other.

2. The sharpness of the transform depends on the number of diffracting objects in the lattice.

3. The observed transform is the product of the diffraction pattern of a single object, times the diffraction pattern of the lattice. In a crystal, this means that the Fourier transform of a single object is "sampled" at points in reciprocal space where the interference function is non-zero.

This can be illustrated in one-dimension for a Gaussian peak, placed in a lattice with either 3 or 21 copies. The interference function samples the underlying molecular transform (in this case, a Gaussian). The sharpness of the "peaks" reflects the number of points in the lattice.



Optical transforms are useful for demonstrating these effects. The **Atlas of Optical Transforms** by G. Harburn, C.A. Taylor, and T.R. Welberry, Cornell University Press (1975) is a fascinating resource

The reciprocal lattice: the Fourier transform of a lattice from N objects starting at the origin is given by

$$F(\vec{S}) = \sum_{n=1}^{N} G(\vec{S}) e^{2\pi i(n-1)\vec{a}\cdot\vec{S}}$$

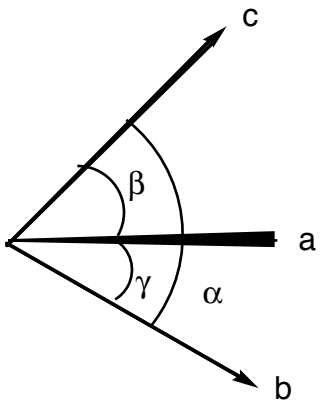
$$G(\vec{S}) = \int_{\text{repeat unit}} \rho(\vec{X}) e^{2\pi i \vec{S}\cdot\vec{X}} d\vec{X}$$

let $\vec{X} = x\vec{a}$, where x is called the fractional coordinate, with $0 \le x \le 1$ in the unit cell

$$\vec{S}\cdot\vec{X} = \vec{S}\cdot x\vec{a}$$
$$= (\vec{S}\cdot\vec{a})x$$
$$= hx$$

$$F(h) = F(\vec{S}) = \int_0^1 \rho(x) e^{2\pi i h x} a dx$$

In three-dimensional space, crystals are generated by infinite translations along the repeat vectors a, b, c which are linearly independent, but not necessarily orthogonal. These three vectors define the unit cell, which is the basic unit of a crystal:



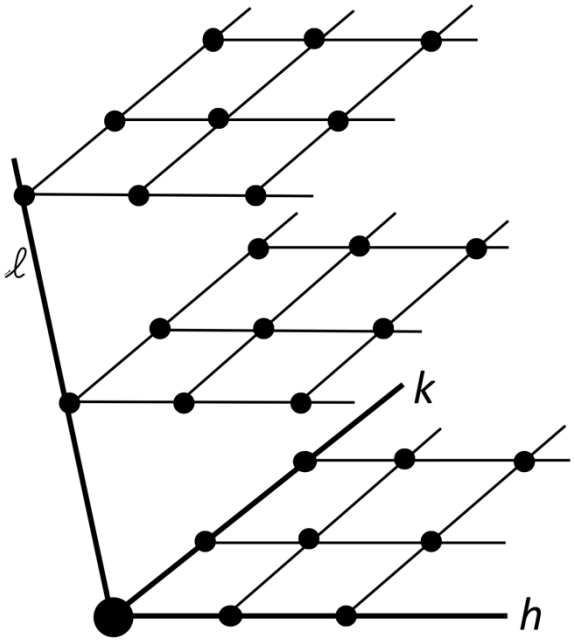
The diffraction condition is now given by:

$$\vec{S}\cdot\vec{a} = h$$
$$\vec{S}\cdot\vec{b} = k$$
$$\vec{S}\cdot\vec{c} = l$$

where h, k, l are integers. The Fourier transform equation becomes:

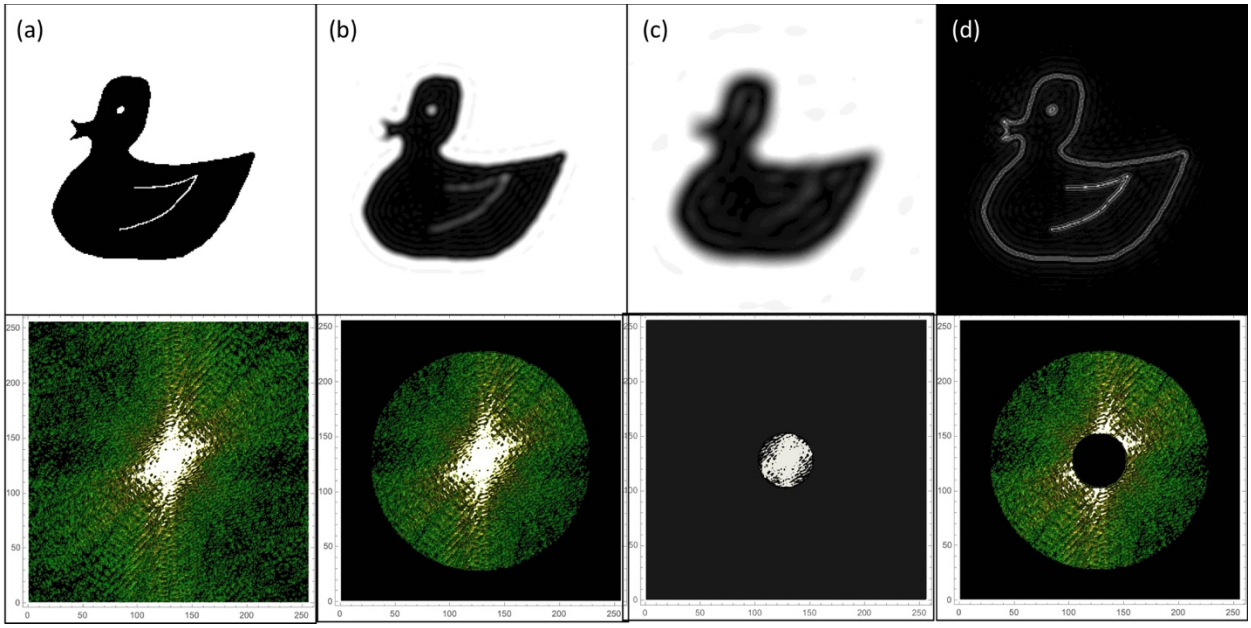
$$F(hkl) = F(\vec{S}) = \iiint_{\text{unit cell}} \rho(x,y,z)e^{2\pi i(hx+ky+lz)}dV$$
$$dV = Vdxdydz$$
$$V = \vec{a}\cdot\vec{b}\times\vec{c} = \text{volume of the unit cell}$$

The integers h, k, l are known as Miller indices, and they specify the coordinates of the allowed diffraction vector in reciprocal space. They are determined only by the unit cell dimensions, and they define a set of points known as the "reciprocal" lattice.



We mentioned that real and diffraction space are reciprocal. What does that mean? Points further away from the origin in reciprocal space correspond to details which are closer together in real space - this brings us to the concept of resolution, which is a measure of the detail in an object. For example, let's try to paint an apple (life size). With a 2" brush, can get the general shape; 1/2" brush, get a better defined outline; 1/8" brush, get the stem, etc. The smaller the brush, the higher the resolution (degree of detail). In diffraction space, higher resolution means we collect scattered radiation out to a further distance from the origin, which corresponds to information about closely spaced points in real space.

The diffracting duck analysis was to my knowledge introduced by Taylor and Cochran (**Optical Transforms**, Bell (1964)) to illustrate the effect of resolution on Fourier transformations. A black and white duck was used as the test object, and Fourier transforms were implemented using an optical bench. The following example, with a hand-sketched duck, was generated in Mathematica, guided by the exposition in J.W. Goodman's **Fourier Transforms Using Mathematica**, SPIE (2020). In the figure below, there are pairs of images with the upper panel corresponding to the real image and the lower panel the corresponding Fourier transform. (a) gives the original object with the Fourier transform below; in (b) and (c) the inverse Fourier transform is calculated after imposing masks of decreasing resolution. (d) is an example of imposing both high and low resolution limits to yield a so-called dark-field imagine that accentuates edges (regions of greatest contrast). [Note: the low resolution duck in (c) resembles the Shmoo from the ancient L'il Abner cartoon strip.]



How is resolution defined? Here is a one-dimensional analysis - look at the smallest separation where scattered radiation from adjacent points differs by one wavelength λ.

$\vec{a} \cdot \vec{S} = h$ (integer)

$\vec{d} \cdot \vec{S} = n$

the smallest spacing, $\vec{d}$, occurs when $\vec{d} \cdot \vec{S}$=d$\left|\vec{S}\right| = 1$

$$\vec{S} = \frac{\hat{s} - \hat{s}_o}{\lambda}$$

$$\left|\vec{S}\right| = \frac{2\sin\theta}{\lambda} = \frac{1}{d}$$

If we look at the scattering from two points separated by d, where d is parallel to S, then

$$\vec{d} \cdot \vec{S} = \left|\vec{d}\right|\left|\vec{S}\right| = n$$

or

$$\left|\vec{S}\right| = \frac{n}{d} = \frac{2\sin\theta}{\lambda}$$

which can be rearranged to Bragg's law:

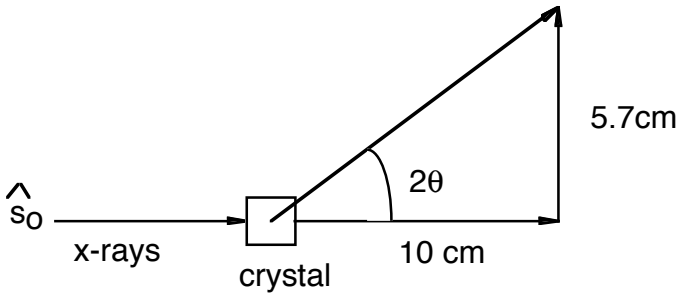
$$2d\sin\theta = n\lambda$$

By convention, the resolution in macromolecular analyses can be derived from the maximum scattering angle for the observed diffraction data

$$\text{resolution} = d = \frac{\lambda}{2\sin\theta_{max}}$$

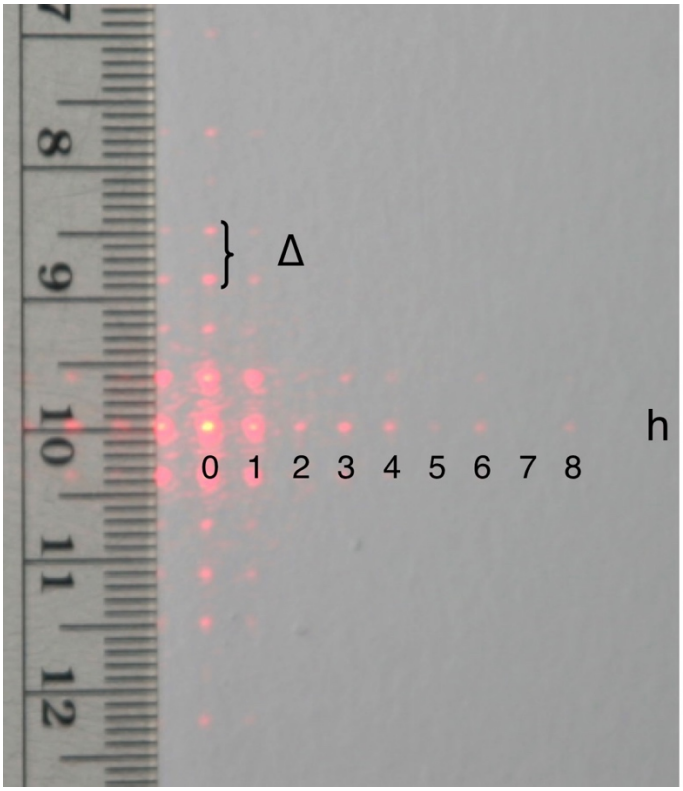
although this is rather subjective.

As an example, for a 3 Å resolution data set, with Cu Kα radiation (λ = 1.54Å), θ_{max} = 14.9˚. This corresponds to scattering at 5.7 cm from the incident beam at 10 cm from the crystal (= 10 cm * tan(2θ)).



Example: Optical diffraction of a 60 mesh wire sieve

The diffraction pattern of a 60 mesh wire sieve was obtained using a red laser ($\lambda \sim 6.5 \times 10^{-4}$ mm) projected against the wall at d = 2400 mm.



The lattice spacing "*a*" along one dimension may be calculated from the scattering relationships that have been derived:

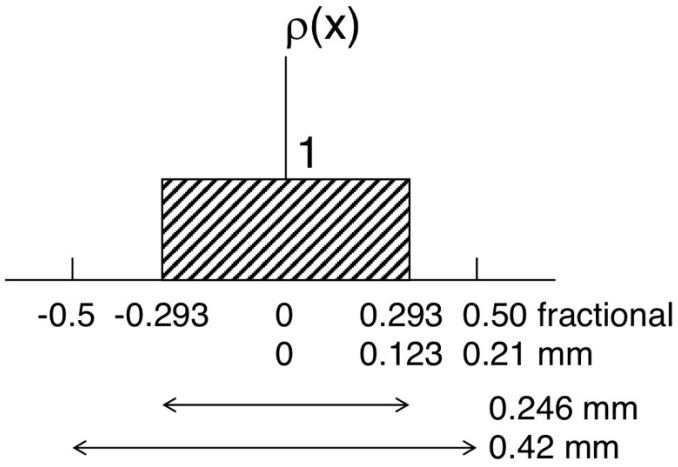
$a = \frac{\lambda}{2\sin\vartheta} \sim \frac{\lambda}{2\vartheta} = \frac{\lambda d}{\Delta}$ with d = 2400 mm and $\lambda = 6.5 \times 10^{-4} mm$

where Δ is the spacing between adjacent points in the reciprocal space lattice of the diffraction pattern.

in the vertical direction, the h = 6 reflection is 23 mm from the origin or Δ = 3.8 mm
in the horizontal direction, the k = 6 reflection is 21 mm from the origin or Δ = 3.5 mm

From Δ, d and λ, *a* is calculated to be 0.41 mm in the vertical direction and 0.45 mm in the horizontal. For reference, a 60 mesh screen has a spacing of 0.42 mm.

While the spacing of the wires determines the location of the diffraction maxima, the intensity reflects the size of the opening, or equivalently, the thickness of the wire. We will examine this along the horizontal axis, using a one-dimensional electron density profile. At the outset, we do not know the size of the opening, but let's "guess" it is 0.246 mm.

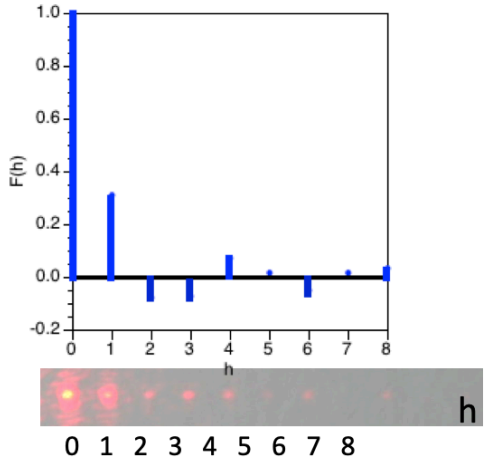


An opening of 0.246 mm corresponds to +/- 0.123 mm, or in fractional coordinates, +/- 0.293 (using 0.42 Å for the lattice spacing).

The diffraction pattern is given by the Fourier transform of this "electron density" profile:

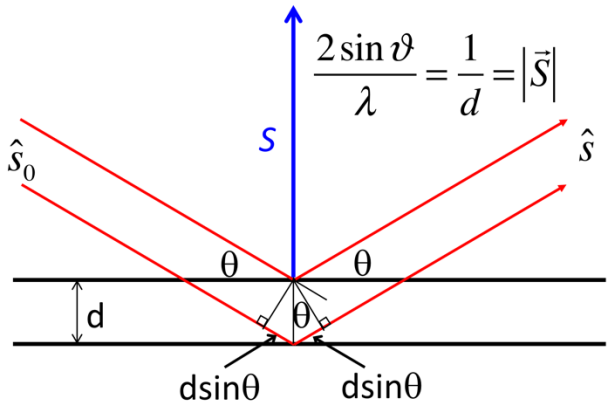
$$F(S) = F(h) = \int_{-0.293}^{0.293} \cos 2\pi hx \, dx = \frac{\sin 2\pi h(0.293)}{\pi h}$$

Qualitatively, this appears to give a reasonable fit to the observed diffraction pattern (a detailed fit would require estimating the intensities from the image of the diffraction pattern, and then taking the square root to get the corresponding amplitude values).

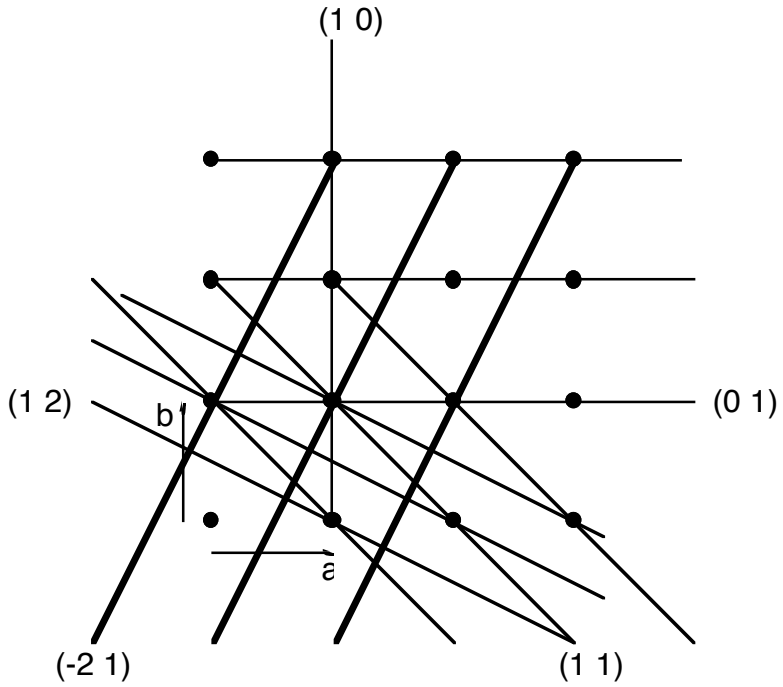


Crystal Lattice Spacings and Reciprocal Space

X-rays are diffracted from crystal planes when they are oriented normal to the diffraction vector, S, and the spacing between the planes is equal to 1/|S|:



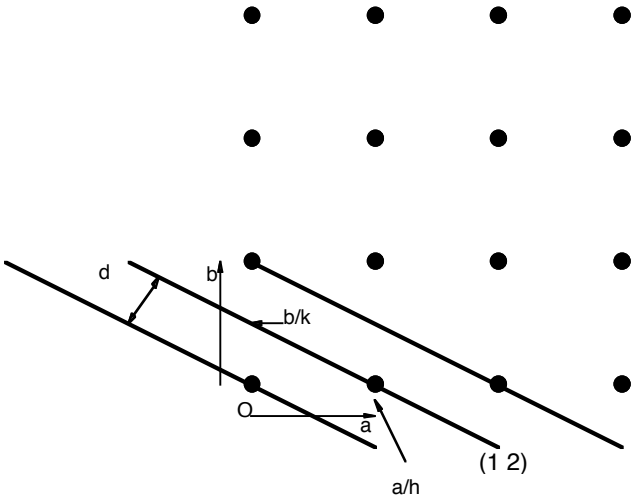
Now, crystals consist of atoms and not planes, but because of the periodic nature of the crystal lattice, there are sets of parallel planes known as lattice planes that have equivalent electron density distributions. These planes are described by their Miller indices, conventionally designated *h, k, l* that are the reciprocals of the intercepts, in units of the cell edges, that a plane makes with the unit cell axes. For simplicity, this is illustrated below for a two-dimensional lattice.



When a plane is parallel to a given axis, its intercept along that axis is at infinity, so that the Miller index is zero.

When a given hkl plane is perpendicular to the diffraction vector, that plane is in the diffraction condition. The spacing of between the adjacent hkl lattice planes gives the reciprocal of the

diffraction vector amplitude. The perpendicular spacing between adjacent hk planes in an orthorhombic, two-dimensional lattice may be calculated as follows:



A line, such as one of these lattice "planes", is defined by the equation y = mx + p, where m and p are the slope and y intercept, respectively. The slope of the *hk* line is given by the change in y divided by the change in x within a unit cell, or m = (-(b/k)/(a/h) = -(bh)/(ak)), and p = b/k. Hence, the equation of the *hk* line is:

$$y = \frac{-bh}{ak}x + \frac{b}{k}$$

d, the spacing between adjacent lattice planes, is given by the perpendicular distance of the *hk* line from the origin, which is equivalent to finding where y^2+x^2 is a minimum on this line. From the equation of the line, this is equal to minimizing $(mx+p)^2 + x^2$ with respect to x. This gives:

$$\frac{d}{dx}\left(\left(mx+p\right)^2 + x^2\right) = 0$$

$$0 = 2m(mx+p) + 2x$$

$$\left(m^2+1\right)x = -bm$$

$$x = \frac{-bm}{\left(m^2+1\right)}$$

d will be the distance from the origin to the hk plane at this value of x, giving

$$d^2 = \left(\frac{m\left(-bm\right)}{m^2+1} + b\right)^2 + \left(\frac{bm}{m^2+1}\right)^2$$

$$= \left(\frac{b}{m^2+1}\right)^2\left(m^2+1\right)$$

$$= \frac{b^2}{m^2+1}$$

In terms of hk, this then equals

$$d^2 = \left(\frac{(b/k)^2}{(hb/ka)^2 + 1} \right)^2$$
$$= \frac{b^2 a^2}{h^2 b^2 + k^2 a^2}$$

The diffraction vector amplitude equals 1/d; squaring this gives:

$$\left|\vec{S}\right|^2 = \frac{1}{d^2} = \left(\frac{h}{a}\right)^2 + \left(\frac{k}{b}\right)^2$$
$$= h^2 a^{*2} + k^2 b^{*2}$$
$$\equiv d^{*2} = \left|ha^* + kb^*\right|^2$$

where the variables with the "*" represent reciprocal lattice quantities. This analysis for an orthogonal, two dimensional lattice shows that the amplitude of the diffraction vector equals the distance from the origin to the diffraction spot *hk*; ie |S| = d*, and this value is the reciprocal of the spacing between the appropriate *hk* planes in real space.

A complete analysis would show that this result is also true for a three-dimensional, non-orthogonal lattice, but that the reciprocal lattice vectors in the general case need not be equal to the reciprocal of the real space vectors; ie, in general a* ≠ 1/a, etc. As suggested by the preceding analysis, a* is perpendicular to the bc plane (that is, it is the perpendicular distance between adjacent 100 planes); b* is perpendicular to the ca plane, and c* is perpendicular to the ab plane. Using vector notation, real and reciprocal space axes are related as follows:

$$a^* = \frac{b \times c}{a \bullet b \times c}, \text{ etc. and } a = \frac{b^* \times c^*}{a^* \bullet b^* \times c^*}$$

By definition, the cross product ("×") of two vectors is perpendicular to both of them. The unit cell volumes in real and reciprocal space are given by the triple product in the denominator of these expressions: V = a•b×c and V* = a*•b*×c*. Equivalent expressions can be obtained by circularly permuting the axes (to maintain a right-handed coordinate system), which has the following pattern:

$$a \to b \to c \to a, \text{ etc.}$$
$$\alpha \to \beta \to \gamma \to \alpha, \text{ etc.}$$

Working through these expressions explicitly yields expressions for the cell lengths and angles:

$$a^* = \frac{bc \sin\alpha}{V} \qquad \cos\alpha^* = \frac{\cos\beta\cos\gamma - \cos\alpha}{\sin\beta\sin\gamma}$$
$$V = abc \sin\alpha^* \sin\beta \sin\gamma$$
$$= abc\sqrt{1 - \cos^2\alpha - \cos^2\beta - \cos^2\gamma + 2\cos\alpha\cos\beta\cos\gamma}$$
$$V^* = \frac{1}{V}$$

Again, expressions for b*, β*, etc. can be obtained by circular permutation of the variables.

In general, crystallographic lattices in real and reciprocal space are non-orthogonal, and this must be remembered when computing lengths, etc. To calculate the length of a vector in a non-orthogonal coordinate system, the metric tensor G or G*, for real and reciprocal space, is used. For example, the length squared of the hkl vector in reciprocal space is given by:

$$
\begin{aligned}
d^{*2} &= \left(hkl\right)^T G^* \left(hkl\right) \\
&= \begin{pmatrix} h & k & l \end{pmatrix} \begin{pmatrix} a^* \bullet a^* & a^* \bullet b^* & a^* \bullet c^* \\ b^* \bullet a^* & b^* \bullet b^* & b^* \bullet c^* \\ c^* \bullet a^* & c^* \bullet b^* & c^* \bullet c^* \end{pmatrix} \begin{pmatrix} h \\ k \\ l \end{pmatrix} \\
&= h^2 a^{*2} + k^2 b^{*2} + l^2 c^{*2} + 2hk\cos\gamma^* + 2hl\cos\beta^* + 2kl\cos\alpha^* \\
&= \left|\vec{S}\right|^2
\end{aligned}
$$

Crystal Symmetry and the Diffraction Pattern

A crystal is defined by the translational symmetry operation on the unit cell - this results in the diffraction pattern of the contents of a single unit cell being sampled by the reciprocal lattice. Additional symmetry operations can also be incorporated into the crystal lattice, such that implementation of these symmetry operations leaves the crystal lattice apparently unchanged.

There are exactly 230 different combinations of symmetry elements that can be accommodated in an infinitely repeating crystal lattice. These different combinations of symmetry elements are known as space groups.

The symmetry properties of the crystal lattice are reflected in both the amplitude and phase of the diffraction pattern. Example, in a centrosymmetric structure, there are pairs of atoms related by an inversion center. This gives rise to pairs of equivalent positions, with coordinates *x,y,z* and *-x,-y,-z* $\equiv \bar{x},\bar{y},\bar{z}$. The consequences of this for the diffraction pattern are as follows:

$$F(hkl) = \sum_{j=1}^{N} f_j e^{2\pi i(hx_j+ky_j+lz_j)} = \sum_{j=1}^{N} f_j e^{2\pi i(\vec{h}\cdot\vec{x}_j)}$$

$$= \sum_{\text{pairs } j'=1}^{N/2} f_{j'}\left[e^{2\pi i(\vec{h}\cdot\vec{x}_{j'})} + e^{-2\pi i(\vec{h}\cdot\vec{x}_{j'})} \right]$$

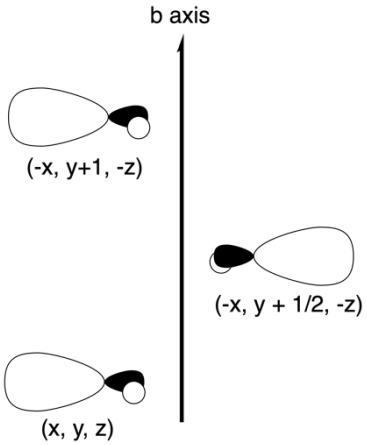
$$\text{recall } \cos\phi = (e^{i\phi} + e^{-i\phi})/2$$

$$= 2\sum_{\text{pairs } j'=1}^{N/2} f_{j'}\left[\cos\left(2\pi(\vec{h}\cdot\vec{x}_{j'})\right) \right]$$

In this case, F(hkl) is **real**, not complex, and the phase is given by the sign: F(hkl) > 0, α_{hkl}=0˚; F(hkl) < 0, α_{hkl}=180˚. Now, in general, macromolecules only crystallize in one of 65 space groups that do not contain mirror planes or inversion centers. Certain "zones" (planes) of reflections can be centrosymmetric, however. An example - projecting (looking) down a twofold axis along y gives pairs of points related by (x,0,z), (-x,0,-z).

Not only do symmetry operators influence phases, but they can also influence amplitudes. Following the outline above, it can be shown that $F(hkl)=F(\bar{h}k\bar{l})$, when a two fold axis is present along the y axis (equivalent positions $(x,y,z),(\bar{x},y,\bar{z})$). In many cases, the space group symmetry can be deduced from the symmetry properties of the diffraction pattern. A more detailed discussion of the symmetry imposed relationships on the amplitudes and phases is provided below.

Space group examples. A common space group for biological macromolecules is $P2_1$, which has a 2_1 screw axis along the b axis in the "standard" setting. Screw axes are helical symmetry operations. $P2_1$ is an example of a monoclinic space group, with a,b , c in general different, $\alpha = \gamma = 90^\circ$, and β has no restrictions. There are two equivalent positions in $P2_1$, x,y,z and $\bar{x},y+\frac{1}{2},\bar{z}$. This space group has two **asymmetric units** that are related by the screw axis, and not just by the translational operators of the lattice.

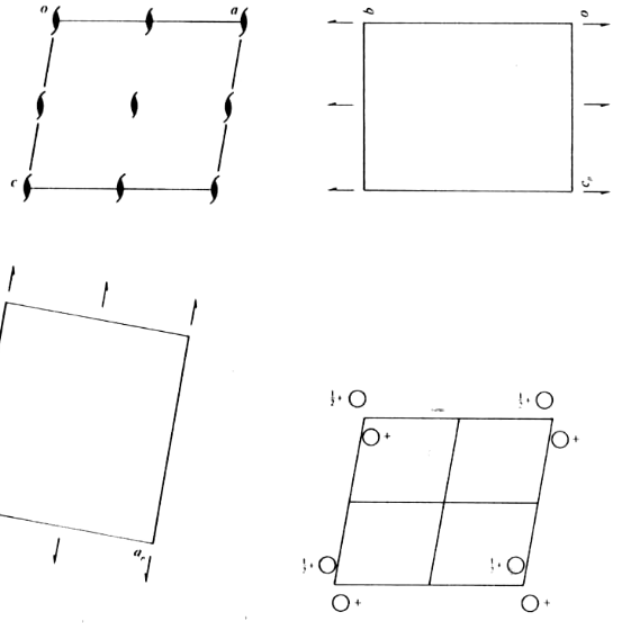


The standard representations for space groups are provided by the International Tables for X-ray Crystallography, with the page for space group $P2_1$ reproduced below:

$P\,2_1$ C_2^2 2 Monoclinic

No. 4 $P\,1\,2_1\,1$ Patterson symmetry $P\,1\,2/m\,1$

UNIQUE AXIS b



Origin on 2_1

Asymmetric unit $0 \le x \le 1;\quad 0 \le y \le 1;\quad 0 \le z \le \frac{1}{2}$

Symmetry operations

(1) 1 (2) $2(0,\frac{1}{4},0)$ $0,y,0$

Generators selected (1); $t(1,0,0)$; $t(0,1,0)$; $t(0,0,1)$; (2)

Positions

Multiplicity, Wyckoff letter, Site symmetry	Coordinates	Reflection conditions
		General:
2 a 1	(1) x,y,z (2) $\bar{x},y+\frac{1}{2},\bar{z}$	$0k0$: $k=2n$

Symmetry of special projections

Along [001] $p\,1g\,1$	Along [100] $p\,1\,1g$	Along [010] $p\,2$
$\boldsymbol{a}'=\boldsymbol{a}_p$ $\boldsymbol{b}'=\boldsymbol{b}$	$\boldsymbol{a}'=\boldsymbol{b}$ $\boldsymbol{b}'=\boldsymbol{c}_p$	$\boldsymbol{a}'=\boldsymbol{c}$ $\boldsymbol{b}'=\boldsymbol{a}$
Origin at $0,0,z$	Origin at $x,0,0$	Origin at $0,y,0$

Maximal non-isomorphic subgroups

I [2]$P\,1$ 1

IIa none

IIb none

Maximal isomorphic subgroups of lowest index

IIc [3]$P\,1\,2_1\,1$ $(\boldsymbol{b}'=3\boldsymbol{b})(P\,2_1)$; [2]$P\,1\,2_1\,1$ $(\boldsymbol{c}'=2\boldsymbol{c}$ or $\boldsymbol{a}'=2\boldsymbol{a}$ or $\boldsymbol{a}'=\boldsymbol{a}+\boldsymbol{c},\boldsymbol{c}'=-\boldsymbol{a}+\boldsymbol{c})(P\,2_1)$

Minimal non-isomorphic supergroups

I [2]$P\,2_1/m$; [2]$P\,2_1/c$; [2]$P\,2\,2\,2_1$; [2]$P\,2_1\,2_1\,2$; [2]$P\,2_1\,2_1\,2_1$; [2]$C\,2\,2\,2_1$; [2]$P\,m\,c\,2_1$; [2]$P\,c\,a\,2_1$; [2]$P\,m\,n\,2_1$; [2]$P\,n\,a\,2_1$; [2]$C\,m\,c\,2_1$; [2]$P\,4_1$; [2]$P\,4_3$; [3]$P\,6_1$; [3]$P\,6_5$; [3]$P\,6_3$

II [2]$C\,1\,2\,1(C\,2)$; [2]$A\,1\,2\,1(C\,2)$; [2]$I\,1\,2\,1(C\,2)$; [2]$P\,1\,2\,1$ $(2\boldsymbol{b}'=\boldsymbol{b})(P\,2)$

In some cases, space groups have "special positions". In P2, with a pure twofold rotation along b, the equivalent positions are x,y,z and $\overline{x},y,\overline{z}$. A point on the b axis (0,y,0) is rotated into itself at this position - also true for (0.5,y,0), (0.5,y,0.5), (0,y,0.5). Although there are two asymmetric units in space group P2, you can have only one molecule in the unit cell, if it is a dimer, with a molecular twofold axis that coincides with the crystallographic twofold. Sometimes this is useful in defining the molecular symmetry just from the space group.

A useful and important property of structure factors can be deduced from the fact that the electron density of a structure is real, which means that $\rho(\vec{x}) = \rho^*(\vec{x})$ for all $\vec{x}$, where * represents the complex conjugate.

$$\rho(\vec{x}) = \frac{1}{V}\sum_{\vec{h}} F\left(\vec{h}\right)e^{-2\pi i\vec{h}\cdot\vec{x}}$$

$$\rho^*(\vec{x}) = \frac{1}{V}\sum_{\vec{h}} F^*\left(\vec{h}\right)e^{2\pi i\vec{h}\cdot\vec{x}}$$

replace $\vec{h}$ with-$\vec{h}$ (OK, since the sum is over all reflections)

$$\rho^*(\vec{x}) = \frac{1}{V}\sum_{\vec{h}} F^*\left(-\vec{h}\right)e^{-2\pi i\vec{h}\cdot\vec{x}}$$

since $\rho(\vec{x}) = \rho^*(\vec{x})$ for all $\vec{x}$, then

$$F^*(-\vec{h}) = F(\vec{h})$$

$$\left|F(-\vec{h})\right|e^{-i\alpha_{-\vec{h}}} = \left|F(\vec{h})\right|e^{i\alpha_{\vec{h}}}$$

Therefore, reflections h and -h have the same amplitudes, and the phase of one is minus that of the other. As a consequence of the realness of the electron density, the diffraction pattern is centrosymmetric. This property is also known as Friedel's law, and it is usually adequately obeyed, although we'll later examine some cases where the breakdown has interesting consequences.

How are |F| and α determined? The first part (measuring |F|) is (at least conceptually) relatively easy and involves data collection; the second (determining α), requires solution of the “phase problem” and can be more challenging. Methods for doing this will be discussed later.

Unit Cell Transformations

On occasion, it may be necessary to transform coordinates and reflection lists between different choices of unit cell. This may arise if the unit cell chosen by auto-indexing during data collection is not the cell that you would like, or if there is some relationship between different crystal forms that one wishes to emphasize. These transformations are relatively easy to implement once one go through the process. More details can be found on pages 70-72 of Volume A of the International Tables.

Let **P** be the matrix that transforms the unit cell axes ($\mathbf{a_1}$) of crystal form 1 into the unit cell axes of crystal form 2 ($\mathbf{a_2}$):

$$\mathbf{a_2}^{\mathbf{T}} = \mathbf{a_1}^{\mathbf{T}}\ \mathbf{P}$$

where $\mathbf{a_1}^{\mathbf{T}}$ is the row vector (a b c), etc. The determinant of **P** gives the unit cell volume of crystal 2 relative to crystal 1 (and will be positive if right-handed coordinate systems are used).

P also transforms the reflection indices from crystal 1 ($\mathbf{h_1}$) to the indices of crystal 2 ($\mathbf{h_2}$):

$$\mathbf{h_2}^{\mathbf{T}} = \mathbf{h_1}^{\mathbf{T}}\ \mathbf{P}$$

where $\mathbf{h_1}^{\mathbf{T}}$ is the row vector (h k l). The inverse transform from crystal 2 to crystal 1 is given by the matrix $\mathbf{Q} = \mathbf{P}^{-1}$ (and usually, $\mathbf{P}^{-1}$ is not the same as $\mathbf{P}^{\mathbf{T}}$). Q transforms the basis vectors:

$$\mathbf{a_2}^* = \mathbf{Q}\ \mathbf{a_1}^*$$

$$\mathbf{x_2} = \mathbf{Q}\ \mathbf{x_1}$$

where $\mathbf{a_1}^*$ is the column vector of the reciprocal space vectors, $\mathbf{x_1}$ is the column vector of the coordinates of a point in real space, etc. The eigenvectors of **Q** with unit eigenvalues correspond to directions (**x** vectors) that are unchanged by this transformation.

If the real space lattice is translated by a vector **p**, then the inverse shift is given by $\mathbf{q} = -\mathbf{Q}\,\mathbf{p}$.

The real space metric tensor, $G_{ij} = \mathbf{a}_i \cdot \mathbf{a}_j$, transforms as

$$\mathbf{G_2} = \mathbf{P}^{\mathbf{T}}\mathbf{G_1}\mathbf{P}$$

and the reciprocal space metric tensor transforms as

$$\mathbf{G_2}^* = \mathbf{Q}\ \mathbf{G_1}^*\ \mathbf{Q}^{\mathbf{T}}$$

Equivalent Reflections and Phase Relationships

Let the j^{th} symmetry operation have a rotation matrix C_j and a translation vector t_j. Then

$$\mathbf{h}_j^T = \mathbf{h}^T \mathbf{C}_j$$

$$\alpha(\mathbf{h}_j) = \alpha(\mathbf{h}) - 2\pi (\mathbf{h}^T \cdot \mathbf{t}_j)$$

For centric reflections, $\mathbf{h}^T \mathbf{C}_j = -\mathbf{h}^T$ (ie, the rotated reflection is the Friedel mate of the original reflection)

$$\alpha(\mathbf{h}_j) = \pi (\mathbf{h}^T \cdot \mathbf{t}_j)$$

Proof:

$$F(h) = \sum_j f e^{2\pi i h(C_j x + t_j)}$$

$$F(h_l) = F(hC_l) = \sum_j f e^{2\pi i h C_l (C_j x + t_j)}$$

$$= e^{-2\pi i h t_l} \left\{ \sum_j f e^{2\pi i h C_l (C_j x + t_j) + t_l} \right\}$$

If C_j, t_j and C_l, t_l are crystallographic symmetry operators, then by definition, transformation of a point x by any C_j,t_j generates an equivalent position. Hence, if $(C_j x + t_j) = x_j$, then $C_l x_j + t_l$ is another equivalent position. Therefore, the term in {} = F(h), and

$$F(hC_l) = e^{-2\pi i h t_l} F(h)$$

If $hC_l = -h$, then F(h) and F(hC_l) are both Friedel mates and centric. Since $\alpha(-h) = -\alpha(h)$ by Friedel's law, then for centric reflections:

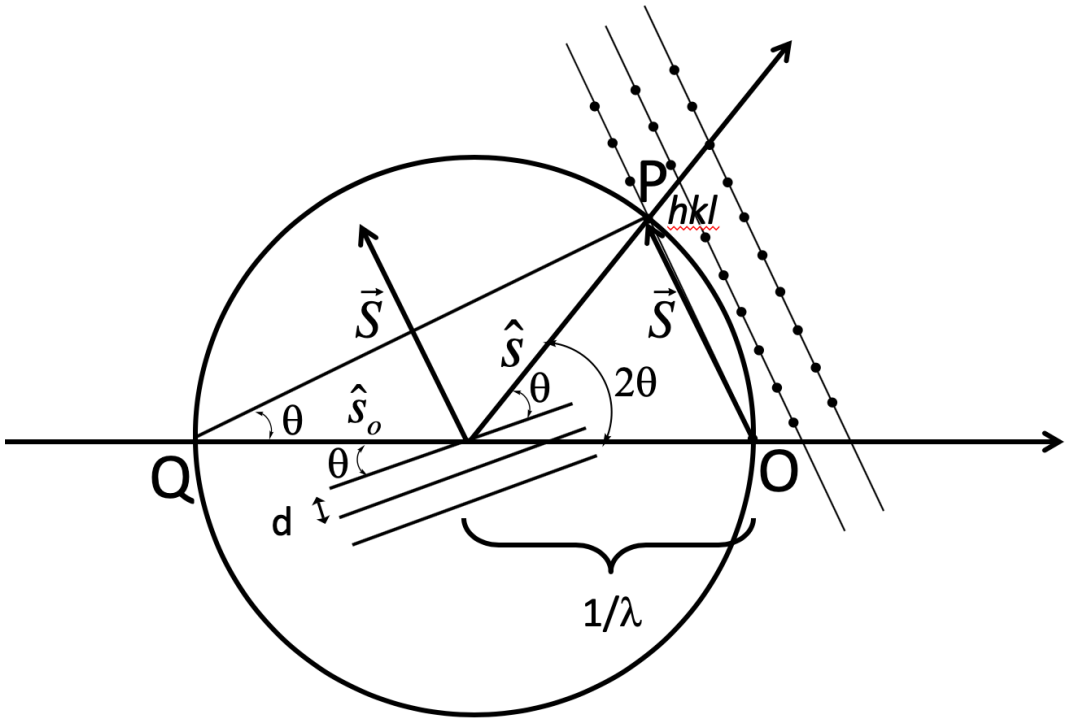
$$\alpha(\bar{h}) = \alpha(h) - 2\pi h t_l = -\alpha(h), \text{ or}$$

$$\alpha(h) = \pi h t_l \quad (\text{modulo } \pi)$$

Data collection

The sphere of reflection (Ewald sphere):

As we have seen, associated with the real space lattice of a crystal is a reciprocal space lattice, with reflections occurring at lattice points possessing integer indices. In the reflecting condition, the diffraction vector is normal to the real space lattice planes that are scattering, and has the length 1/d, where d is the spacing between these planes. A geometrical construct that summarizes the diffraction condition is the Ewald or reflection sphere. This is generated by a sphere of radius $1/\lambda$, with the primary beam, $\hat{s}_o$, passing along a diameter QO and exiting through the origin, O, on the surface of the sphere. The scattered beam exits the sphere in the direction $\hat{s}$ from the center of the sphere to the surface at point P. If a lattice point is positioned at P, then the diffraction vector $\vec{S}$ or OP goes from the origin to this point. The corresponding direct lattice planes will be parallel to QP, and perpendicular to $\vec{S}$. The length of OP, by definition, is the distance of the point P from the reciprocal space origin. Under these conditions, $OP = QO \sin\theta = 2\sin\theta/\lambda$ which is Braggs law. Hence, any lattice point which is positioned on the Ewald sphere will be in the diffracting position. For example, the X-ray diffraction pattern of a stationary crystal (which is known as a still photograph) will consist of spots on a series of concentric rings (called lunes), since the intersection of a set of parallel planes and a sphere generate these rings. Data collection strategies are designed such that all the necessary lattice points pass through the Ewald sphere, so that they are in the diffracting position and the intensities of these spots can be measured experimentally.



(Inspired by Figure 17-10 of Eisenberg and Crothers **Physical Chemistry with Applications to the Life Sciences** (1979))

X-ray sensitive detectors come in several flavors with image plates and CCDs the current workhorses in structural biology. Several types of data collection strategies can be employed:

Rotation geometry/monochromatic radiation: The crystal is rotated through a defined angular range in a collimated, monochromatic beam while reflections are recorded on the detector. Critical parameters for data collection include the wavelength, crystal to detector distance, oscillation axis, oscillation range per photograph, starting and ending oscillation range, and exposure time. These parameters will define the resolution, completeness and quality of the data by ensuring a complete data set is collected to the desired resolution and that spots are sufficiently resolved from each other. Whenever possible, highly redundant data sets should be obtained for accurate intensity estimates.

Laue/white radiation: This approach is not so widely employed, but can be useful in certain time resolved studies. Here, a fixed crystal is placed in a collimated, polychromatic (white) radiation source - perhaps $\lambda = 0.5$ - 2Å, so that many reflections can be recorded simultaneously.

Mosaicity and Crystal Perfection

If crystals were perfect, a given reflection would only diffract in a precisely defined position, and any rotation away from this position would move the reflection from the diffraction position. As quickly learned, however, reflections diffract over a finite range, due to imperfections in the crystal and optical system. Crystals are typically composed of small blocks of perfect crystals estimated as $\sim 10^3$ Å in length (Glusker). These blocks are, however, slightly misoriented with respect to each other and are consequently termed mosaic blocks. The orientational distribution of these blocks contributes to a finite width for reflections, along with the wavelength spread, collimation, etc. of the incident X-ray beam.

Curiously, perfect crystals diffract less intensely than imperfect, mosaic crystals, since the perfect alignment of all the unit cells means that multiple scattering events need to be taken into consideration, and a detailed analysis shows that the scattered intensity is proportional to $|F|$, rather than $|F^2|$ for mosaic crystals. These multiple scattering events give rise to extinction effects that modulate the observed intensities. As a result, small molecule crystallographers will occasionally immerse their crystals in liquid nitrogen to generate small cracks that reduce crystal perfection.

Multiple scattering effects are particularly important for electron diffraction studies, since electrons are scattered much more effectively than X-rays. They are also present in X-ray studies; there was concern in the early days of crystallography that macromolecular structures may be greatly complicated by multiple scattering events, because so many reflections are simultaneously in the diffraction position at any given time (a condition for multiple scattering). As events have subsequently shown, this is not a serious issue, although perhaps for very accurate, high resolution structural studies may need to take this into account. Certain multiple scattering events ("three-beam" studies) have also been shown to provide experimental phase estimates, by measuring the change in intensity of a given reflection as other reflections pass through the Ewald sphere.

Chapter 2: Patterson Methods

After data collection and processing (which we haven't discussed, but should have), we have the amplitudes, but not the phases, for the diffraction pattern of a crystal. The fun part now starts to get the phases/structure (these are effectively interchangeable goals for macromolecular structures (at least as of the time these notes were originally prepared)).

Historically, the first approach used to solve structures was the method of trial and error, where the structure was guessed and the diffraction pattern calculated to see if there was agreement with the observed diffraction pattern. Pauling called this the stochastic approach, and developed it to its highest art form. This can be useful for small structures, or ones with high symmetry, but as the number of atoms increases, it becomes less and less practical.

The next important advance was the introduction of the Patterson function by Patterson in 1934. Patterson found that the Fourier transform of the intensities (which are experimentally available) gave a useful function that he modestly called the F^2 synthesis, and which we now call the Patterson function. As Pauling described in the 1989 Pauling Lecture, the introduction of this function marked a revolution in crystal structure determination.

The Patterson function, P, as the Fourier transform of the intensities, turns out to be equivalent to an autocorrelation function of the structure, and has a large value at positions corresponding to vectors between pairs of atoms:

$$P(\vec{u}) = \frac{1}{V}\sum_{\vec{h}} I(\vec{h}) e^{-2\pi i \vec{h}\cdot\vec{u}}$$

(the usual convention is that $\vec{u} = (u,v,w)$=Patterson space vector

and $\vec{x} = (x,y,z)$= real space vector

$$P(\vec{u}) = V\int_0^1 \rho(\vec{x})\rho(\vec{x}+\vec{u})d\vec{x}$$

Proof:

$$P(\vec{u}) = V\int_0^1 \rho(\vec{x})\rho(\vec{x}+\vec{u})d\vec{x} = V\int_0^1 \left[\frac{1}{V}\sum_{h'} F(h')e^{-2\pi i h' x}\right]\left[\frac{1}{V}\sum_{h} F(h)e^{-2\pi i h(x+u)}\right]dx$$

$$= \frac{1}{V}\sum_h\sum_{h'} F(h)F(h')e^{-2\pi i h u}\int_0^1 e^{-2\pi i x(h+h')}dx \quad \text{(vector notation implicit in what follows)}$$

Note: $\int_0^1 e^{-2\pi i x(h+h')}dx = 1; \quad h+h' = 0$

$$=0; \ h+h' \neq 0$$

$$= \frac{1}{V}\sum_h F(h)F(\bar{h})e^{-2\pi i h u}$$

$$= \frac{1}{V}\sum_h F(h)F^*(h)e^{-2\pi i h u}$$

$$P(u) = \frac{1}{V}\sum_h I(h)e^{-2\pi i h u}$$

This can be further simplified, since the diffraction pattern has inversion symmetry

$$P(u) = \frac{1}{V}\sum_{\text{all } h} I(h)e^{-2\pi ihu}$$

since $I(h) = I(\bar{h})$

$$P(u) = \frac{1}{V}\sum_{h>0} I(h)\left[e^{-2\pi ihu} + e^{2\pi ihu}\right]$$

$$P(u) = \frac{2}{V}\sum_{h>0} I(h)\cos(2\pi hu)$$

An equivalent derivation can be performed starting from the reciprocal space definition of P:

$$\begin{aligned} P(u) &= \frac{1}{V}\sum_{\text{all } h} I(h)e^{-2\pi ihu} \\ &= \frac{1}{V}\sum_{h} F(h)F^*(h)e^{-2\pi ihu} \\ &= \frac{1}{V}\sum_{h}\left[\int_0^1 \rho(x')e^{2\pi ihx'}Vdx'\right]\left[\int_0^1 \rho(x)e^{-2\pi ihx}Vdx\right]e^{-2\pi ihu} \\ &= V\int_0^1 \rho(x)\left[\int_0^1 \rho(x')\left(\sum_h e^{2\pi i(x'-x-u)}\right)dx'\right]dx \end{aligned}$$

Note: the Dirac delta function ($\delta(x)$) is defined as

$$\begin{aligned} \sum_{h=-\infty}^{h=+\infty} e^{2\pi ihx} &\equiv \delta(x) \\ &= \infty, \quad x = 0 \\ &= 0, \quad x \neq 0, \text{ and} \\ \sum_h e^{2\pi ih(x'-x-u)} &= \delta(x'-x-u), \text{ so} \end{aligned}$$

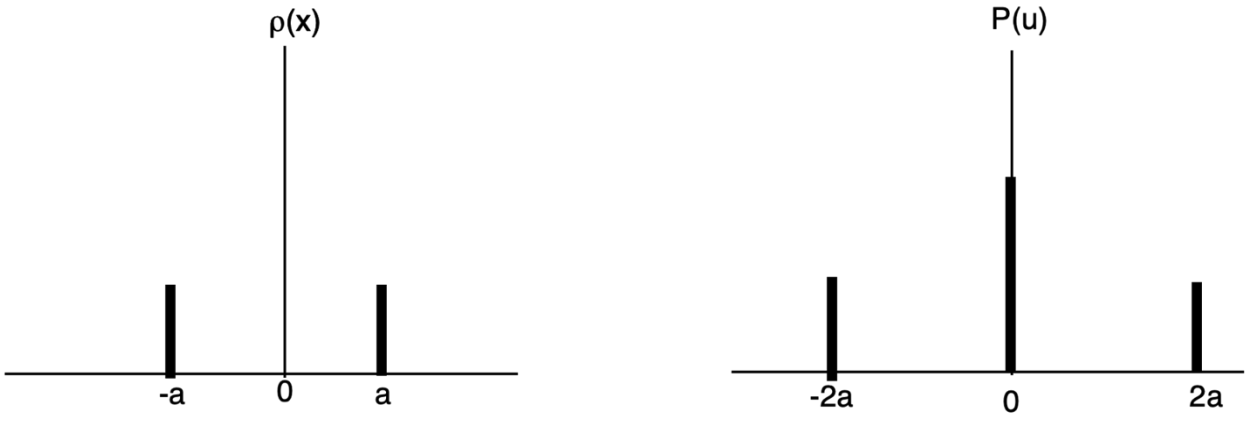
$$\int_0^1 \rho(x')\delta(x'-x-u)dx' = \rho(x+u), \text{ and}$$

$$P(u) = V\int_0^1 \rho(x)\rho(x+u)dx$$

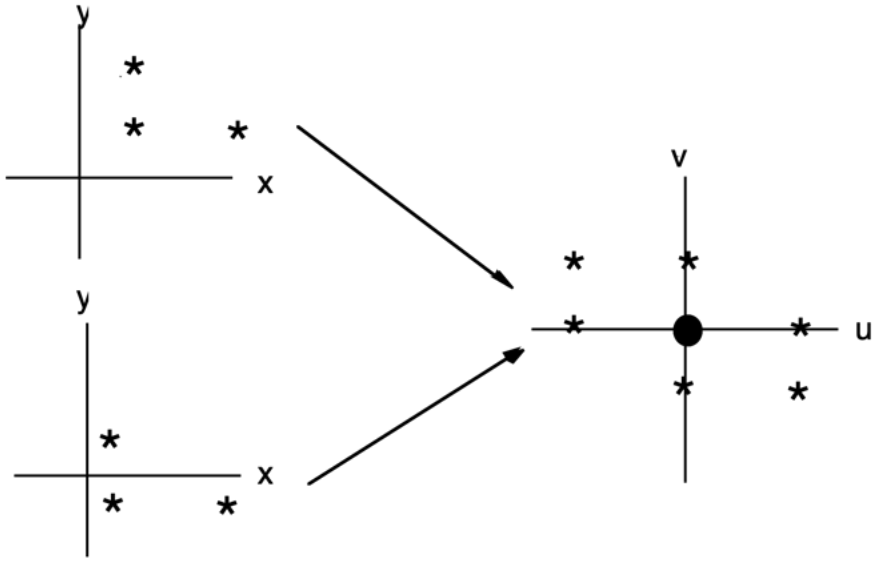
Properties of the Patterson function

The Patterson function consists of the vectors between atoms. The relationship between $\rho(x)$ and $P(u)$ can be illustrated by some simple examples.

(a) one-dimensional example, with two point scatterers at $x = \pm a$.



(b) two-dimensional example. Note, translation of entire object has no effect on Patterson function - recall the demonstration of the optical transform of the screen, where a translation of the screen has no effect on the intensity of the diffraction pattern (which is the Fourier transform of the Patterson function). The Patterson function can be seen to be generated by sequentially moving all atoms in the structure to the origin (method of superimposed images).



The relationship between the atomic positions in crystals and the corresponding Patterson function can be determined by looking at the vectors between all atoms (plus symmetry related sites (equivalent positions)). For example, in space group $P2_1$, the equivalent positions are x, y, z and -x, y+1/2,-z. Consequently, Patterson peaks will be present that correspond to the interatomic vectors between atoms related by crystallographic symmetry. Patterson peaks between equivalent sites are known as **self-vectors**, and in space group $P2_1$ these have the components:

(u,v,w) = (x,y,z) - (-x, y+1/2,-z) = (2x, -1/2, 2z) = (2x, 1/2, 2z)

(u',v',w') = (-x, y+1/2, -z) - (x,y,z) = (-2x, 1/2, -2z)

The Patterson function for a $P2_1$ structure does not have a screw axis, but it does have a center of inversion (since vectors from atoms i to j, and j to i, are both found). All Patterson functions are centrosymmetric (neglecting anomalous dispersion), have the same lattice type (P, C, etc.) as the space group, and have all translational symmetry elements (screw axes and glide planes) replaced by their non-translational symmetry components (rotation axes and mirror planes). Because of the inversion center, the asymmetric unit of the Patterson function is 1/2 that of the real cell (see the new International Tables). For example, space group $P2_1$ has Patterson symmetry P2/m, as does space group P2. Space group $P2_12_12_1$ has Patterson symmetry Pmmm.

The significance of the Patterson function for structure determinations is that atomic positions can be derived from the positions of Patterson function peaks. Surprisingly, remarkably little has been published about how to do this, even though crystallographers traditionally have spent lots of time on this problem. One good introduction is Chapter 7 of Lipson and Cochran.

It can be useful to start with interpreting the self-vectors when solving Patterson maps manually (but see the upcoming example). In space group $P2_1$, the self-vectors appear in the v = 1/2 plane. David Harker, a 1936 CIT Chemistry PhD, noted (*J. Chem. Phys.* **4**, 381 (1936)) that certain lines or planes in Patterson space have a high concentration of self-vectors, which are now called Harker sections. For example, space group $P2_1$ has a v = 1/2 Harker section, while space group P2 has a v = 0 Harker section. Space group $P2_12_12_1$ has three Harker sections at u = 1/2, v = 1/2, w = 1/2.

To find an atomic position in space group $P2_1$, look at the Harker section (v = 1/2) and get the u, w coordinates of a point and divide these by 2 to get x ,z. So, if (u, v, w) = (0.26, 1/2, 0.28), then (x, y, z) = (0.13, ?, 0.14). In this space group, the y value for any one point is arbitrary, and often selected to be 0 - but this is only true for the first position found. In space group P1, the coordinates of the first position found are all arbitrary, and can be taken as (0, 0, 0). This situation is known as the origin ambiguity.

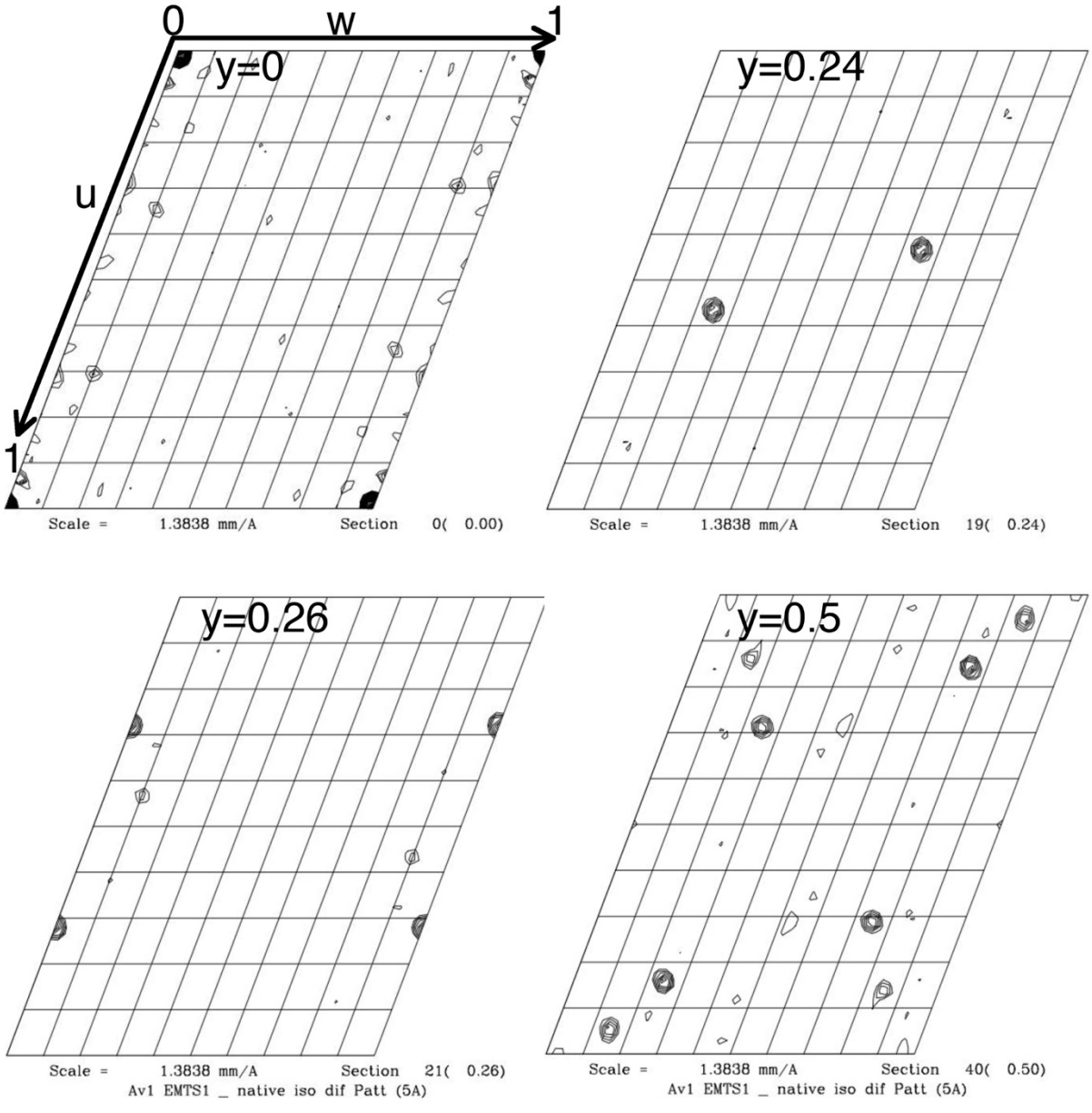
Further consideration of the possible solutions to a given Patterson map will reveal other types of origin ambiguities. In space group $P2_1$, the following single sites would all give the same sets of vectors in the Patterson function: (0.13, y, 0.14), (0.63, y, 0.14), (0.13, y, 0.64), and (0.63, y, 0.64) – i.e., in addition to the arbitrary y coordinate, it is possible to add 1/2 to either or both x and z. Because the Patterson peaks involve quantities like 2x or 2z, they are insensitive to the addition (or not) of 1/2 to these coordinates. Examination of the arrangement of symmetry operators in space group $P2_1$ reveals that screw axes are spaced by 1/2 along a and c, and which one is selected to be the origin is arbitrary. In other space groups, the specifics of the origin ambiguity will differ, but the same general considerations are valid.

In addition to the origin ambiguity, there is also an enantiomeric ambiguity. Since the Patterson function is centrosymmetric, a given structure and its enantiomer will have the same Patterson function (so, for example, in space group $P2_1$, atoms at (x, y ,z) and (-x ,-y ,-z) will have the same Patterson peaks). When the position of the first atom is established, there is a 50:50 chance it will be consistent with a particular enantiomer. Once the coordinates of the first atom are assigned, then that defines the choice of enantiomer for all the remaining sites.

To summarize this discussion, in going from the Patterson function to the crystal structure, there are ambiguities with respect to the choice of origin and enantiomer. For the first site that is located, the choice of origin and enantiomer are arbitrary. However, the origin and enantiomer are no longer arbitrary for all subsequent sites, and they need to be assigned consistently with the first site. This is done through examination of the **cross-vectors** that appear between sites that are not related by crystallographic symmetry, such as between the first and subsequent sites. The following example demonstrates how this can be done.

A real example

The 3-dimensional peak listing and the Harker section for a heavy atom difference Patterson map (the EMTS1 derivative of the *A. vinelandii* MoFe-protein), calculated at 5 Å resolution in space group $P2_1$, are illustrated below. Interpret this Patterson map in terms of two major sites, and two minor sites. The two minor sites have the approximately the same y coordinates as the two major sites. (These figures were generated in CCP4 with reflection file av1_emts1_sca.mtz)



peaks with I/σ > 3 as interpolated from this difference Patterson map by PEAKMAX

pk	u	v	w	I/σ
1	0.000	0.000	0.000	75.4
9	0.061	0.000	0.984	6.2
13	0.290	0.000	0.000	4.4
14’	0.295	0.000	0.897	4.3
17	0.346	0.000	0.096	3.7
17'	0.654	0.000	0.904	3.7
14	0.705	0.000	0.103	4.3
10	0.930	0.000	1.000	5.4
8	0.939	0.000	0.016	6.2
21	0.363	0.031	0.062	3.4
20	0.637	0.031	0.938	3.4
27	0.056	0.035	0.029	3.2
26	0.945	0.035	0.971	3.2
11	0.000	0.038	0.062	4.6

pk	u	v	w	I/σ
11'	0.000	0.038	0.938	4.6
18	0.358	0.166	0.219	3.4
19	0.642	0.166	0.781	3.4
24	0.127	0.225	0.847	3.2
25	0.873	0.225	0.153	3.2
4	0.438	0.237	0.755	7.8
5	0.563	0.237	0.245	7.8
2	0.280	0.262	0.000	9.1
2'	0.720	0.262	0.000	9.1
15	0.432	0.267	0.099	3.7
16	0.568	0.267	0.901	3.7
23	0.380	0.273	0.902	3.2
22	0.620	0.273	0.098	3.2

pk	u	v	w	I/σ
29	0.026	0.425	0.224	3.1
31	0.087	0.425	0.224	3.1
30	0.913	0.425	0.777	3.1
28	0.974	0.425	0.776	3.1
7'	0.056	0.500	0.856	6.3
12	0.137	0.500	0.146	4.6
3	0.160	0.500	0.758	8.3
6	0.291	0.500	0.254	6.8
6'	0.709	0.500	0.746	6.8
3'	0.840	0.500	0.242	8.3
12'	0.863	0.500	0.854	4.6
7	0.944	0.500	0.144	6.3

The solution is at the end of this handout – if you want to work this first, don’t peek (peak?)!

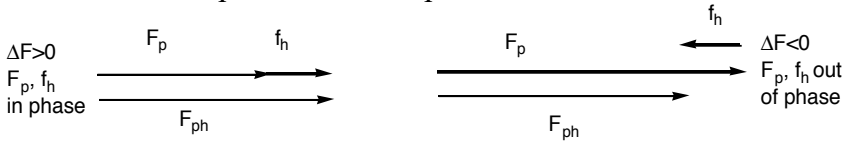
Why can't every structure be solved from the Patterson function? Because, then life would be too easy! Actually, there is an even simpler answer. Consider a structure with N atoms. Each atom has N-1 vectors to other atoms, and 1 vector to itself. So, the Patterson function has N(N-1) vectors to other atoms, and N origin vectors, for a total of N^2 peaks in the same Patterson volume that only contains N peaks in the real structure. Hence, as the size of the structure increases, there are more and more peaks overlapping, and they can no longer be resolved and identified with vectors between a particular pair of atoms.

What happens when the structure cannot be solved from the Patterson function (the typical case in macromolecular crystallography)? Then, more indirect methods need to be used. The primary method used to solve new structures today (i.e. in the golden era of macromolecular X-ray crystallography) is the method of multiple isomorphous replacement, which was developed by Perutz in his pioneering studies of the hemoglobin structure. The basic idea is to bind a heavy atom to a crystallized structure, so that nothing is changed except for the addition of the heavy atom (i.e. isomorphous replacement). The heavy atom will also contribute to the scattering (here, heavy is defined operationally in the sense that measurable changes (~10-20%) in the intensities must occur). From the change in diffraction intensities, ideally (1) the heavy atom positions can be determined (from Patterson functions), and (2) the protein phases can be established. We'll see how to do this shortly.

But first, how are heavy atom derivatives prepared? The traditional approach is through trial and error experiments, either by soaking existing crystals or by co-crystallizations. Common heavy atom reagents include mercurials (mercuric chloride, ethylmercurithiosalicylate, ethyl mercury phosphate, ethyl mercury chloride), platinates (K_2PtCl_4), trimethyl lead acetate, K_3IrCl_6, gold sodium thiomalate, etc. More recently, useful MAD derivatives are prepared by producing the protein in a cell line that can incorporate a heavy atom (selenomethionine), by soaking crystals in solutions (with cryosolvents) that contain I^- or Br^-, or by binding a derivatized ligand (inhibitor) to the protein.

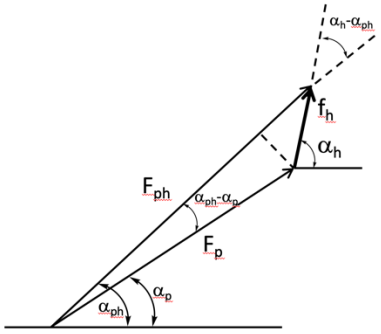
How are heavy atom positions determined? If the heavy atom scattering factor amplitude, $|f_h|$, was available, then the Patterson function of this would give the heavy atom - heavy atom (h-h) vectors, which could be analyzed to give the heavy atom position. But, this quantity isn't available. Instead, "all" we have (after data collection) are $|F_p|$ and $|F_{ph}|$, the amplitudes of the protein and derivative structure factors. The Patterson function of $|F_p|$ would have the protein atom - protein atom vectors (p-p), while the Patterson function of $|F_{ph}|$ will have p-p, p-h, and h-h vectors. While the individual heavy atom vectors (h-h) will be larger than the individual p-p vectors, there are so many more p-p vectors that the h-h vectors cannot be identified. One possibility would be to do a type of "difference" Patterson map, by taking the Fourier transform of $|F_p|^2 - |F_{ph}|^2$, which will cancel out the p-p vectors, leaving just the p-h and h-h vectors. This works better, but still isn't very useful because the h-h vectors are hard to see against the background of p-h vectors. (This is actually something that is well worth checking out with real data.)

The current solution is to use another type of "difference" Patterson calculation, the so-called $(\Delta F)^2$ synthesis, first introduced by Rossmann (*Acta Cryst*. **13**, 221 (1960)). The logic of this approach is most clear for centrosymmetric structures, where the scattering from heavy atoms is either in phase, or 180˚ out of phase, from the protein:



In either case, $|\Delta F| = | |F_{ph}| - |F_p| | = |f_h|$ (as long as both $|F|$'s are greater than $|f_h|$) so that the Fourier transform of $|\Delta F|$ gives the h-h vectors only (ideally!).

For non-centrosymmetric structures, the situation is situation is slightly more complex, since the heavy atoms and protein atoms no longer scatter exactly in or out of phase (in general). For this case, the vector diagram becomes:



If $|F_{ph}|, |F_p| >> |f_h|$, then $\Delta F \sim |f_h| \cos(\alpha_h - \alpha_{ph})$: Consequently, $|\Delta F|$ is always less than or equal to $|f_h|$. Although Bragg disliked this, Rossmann's approximation turned out to be close enough to permit location of heavy atoms in many cases. We'll later see that if anomalous data is available, even better approximations can be determined.

$$\Delta F = \left|f_h\right| \cos\left(\alpha_h - \alpha_{ph}\right) = \frac{1}{\left|F_{ph}\right|}\left(Aa + Bb\right)$$

with $F_{ph} = A + iB$ and $f_h = a + ib$

Solution to Patterson problem

Space group $P2_1$, equivalent positions x, y, z and -x, y+1/2, -z, has Harker vectors $\pm$ (2x, 1/2, 2z)

peak heights in a Patterson map are ~ $f_i f_j$ for the vector between atoms i and j. So, the largest peaks will involve the two major sites, the next largest will be between the major and minor sites, and the weakest will involve only the minor sites.

(NOTE: for historical reasons, all v's with $0.225 \leq v \leq 0.237$ are assigned v = 0.24, and with $0.262 \leq v \leq 0.272$ are assigned v= 0.26)

Structure solution strategy:

start with self-vectors? Hard, due to complications of hand and origin ambiguities.
start with cross-vectors? Yes!

site 1 1a: x1, y1, z1 1b: -x1, y1+1/2, -z1
site 2 2a: x2, y2, z2 2b: -x2, y2+1/2, -z2

crosspeaks: 2a-1a: x2-x1, y2-y1, z2-z1
1a-2b: x2+x1, 1/2-(y2-y1), z2+z1
these peaks are located in v sections symmetric about v=1/4

so, if 2a-1a is assigned to the peak .28, .26, .0 (peak 2)
and 1a-2b is assigned to the peak .57, .24, .25 (peak 5)

then, these can be solved to give:
x2, y2, z2 .43, .26, .12
x1, y1, z1 .15, .0, .12

checking the self-peaks for these two sites shows that they are the two largest peaks (3' and 6) on the Harker section.

These sites have the same z coordinate, which reflects the ncs twofold in this structure, which is perpendicular to the crystallographic a and b axes; ie, it is along c*.

Finding the minor sites - in a way, this is done by trial and error. In this particular case, we'll use the fact that pairs of sites have the same y coordinates. Hence, one minor site has y = 0 and the other has y = .26. The crosspeak from (x1, y1 ,z1) = .15,. 0, .12 to the minor site with y = .26 will be found in the v = .26 section of the Patterson map. There are 4 possibilities for this crosspeak (once crosspeaks 2 and 5 and symmetry equivalents are excluded)

.43, .26 ,.10 (peak 15); .57, .26 ,.90 (peak 16);
and .62, .26, .10 (peak 22) ; .38, .26, .90 (peak 23)

(**note**: these peaks are not always evident when Patterson maps are calculated and contoured, since their precise height varies with different resolutions, $|\Delta F|$ magnitude, etc. cutoffs. <u>This emphasizes the need to experiment with these different parameters when trying to locate heavy atom positions</u>).

By adding the coordinates of this potential crosspeak to (x1, y1 ,z1) =.15, .0, .12, the possibilities for the coordinates x3, y3, z3 for site 3 are:

.58, .26, .22; .72, .26, .02; .77, .26, .22; .53, .26, .02

Now, these sites can be checked with crosspeaks to the major site x2 ,y2, z2 = .43, .26, .12:

.58, .26, .22; and symmetry equivalent .42, -.24, .78 will have crosspeaks to .43, .26, .12 of: .15, .0, .10 and .01, .5, .34 which are not present. This rules out this possibility for x3.

.72, .26, .02 and symmetry equivalent .28, -.24, .98 will have crosspeaks to .43, .26, .12 of: .29, .0, .90 (peak 14') and .15, .5 ,.14 (peak 12), which are present. Thus, this site is probably x3, which is confirmed by testing out the other two possibilities.

The coordinates for the 4th site can be found using the strongest unaccounted for peak on the Harker section (.05, 1/2, .85; peak 7) and (.95, 1/2, .15; peak 7'). These should represent a potential crosspeak involving the minor site with y = 0, with the major site at .15, .0, .12 (because this pair will have a crosspeak at y = 1/2). Hence the possibilities for x4 are:

.20, .5, .97 = .80, .0, .03 and .10, .5, .27 = .90, .0, .73.

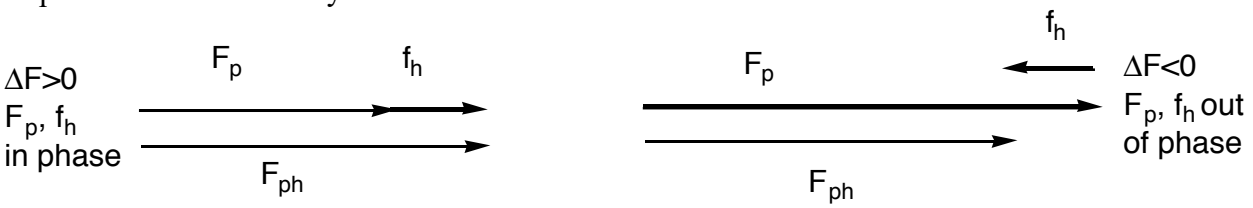
The first possibility would have crosspeaks to the site .43, .26, .12 of .63, .26, .09 (peak 22), while the second would have crosspeaks at .53, .26, .39 (unobserved). Hence the final solutions are

major sites:	.43, .26, .12	.15, .0, .12
minor sites	.72, .26, .02	.80, .0, .03

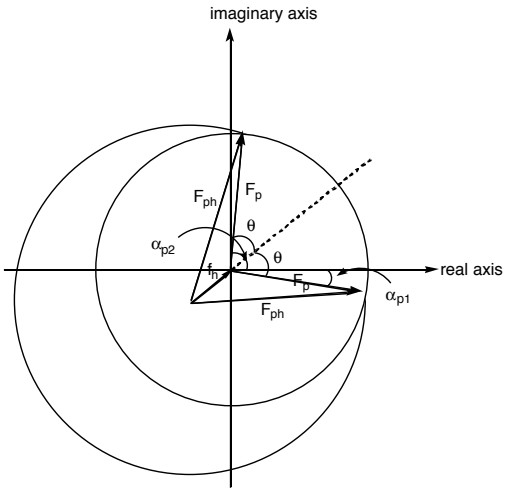
While this is a real problem, it is unlikely the solution would actually be found by manual interpretation of the Patterson function. Instead, one of a variety of programs would be used that work in either real space (Patterson search programs) and reciprocal space (direct methods like Shelx and SnB). The Patterson search programs can use various strategies, but the one I like (if there are not too many sites) is to systematically step through the unit cell (or appropriate region taking into account origin and hand ambiguities) and find single sites consistent with the Harker sections and other self-vectors. Subsequent sites are then identified by systematically stepping through the asymmetric unit looking for sites with cross-vectors to input single sites; as more sites are established, these can all be used in the search for additional sites. The scoring function used to identify potential sites is critical, and can include sum function (Patterson density is summed up over all predicted peak positions) or minimum function (lowest Patterson density over all predicted peak positions is used). The sum function is sensitive to very large peaks (like the origin), and usually isn't too useful. The minimum function is most sensitive, but it can be mislead by ripples, etc. It is possible to use the minN function, and take the Nth lowest peak as the score. Implementation of this also requires consideration of multiplicity effects (proper weighting of peaks on symmetry axes, planes, etc.)

Chapter 3: Phasing Methods in Macromolecular Crystallography

The most powerful experimental phasing methods are based on perturbations introduced into the diffraction pattern of macromolecular crystals by the incorporation of suitable heavy atoms. From the location of the heavy atoms (determined by Patterson methods discussed in part 2), the complex scattering factor for the heavy atoms, $\vec{f}_h$, can be calculated from the basic scattering (Fourier transform) equation. The relationship between the native and derivative structures is given by $\vec{F}_p + \vec{f}_h = \vec{F}_{ph}$. After the heavy atoms have been located, then $\vec{f}_h$, $|\vec{F}_p|$ and $|\vec{F}_{ph}|$ are known. Of course, it is something else, namely α_p, that is desired. For centric reflections, all of the vectors are either in phase or exactly out of phase. So, if $\Delta F = |F_{ph}| - |F_p| > 0$, then the h and p atoms are scattering in phase, and $\alpha_p = \alpha_h$. Conversely, if $\Delta F < 0$, then the h and p atoms are scattering out of phase, and $\alpha_p = \alpha_h + 180^\circ$. So, the protein phases can be determined "by inspection" in the centrosymmetric case:



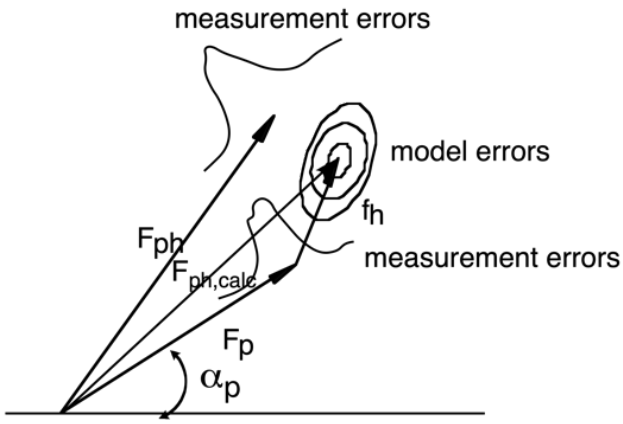
Unfortunately, in macromolecular work, few reflections are centrosymmetric (with the exception of certain zones that are perpendicular to evenfold rotation axes, this can only occur if racemic mixtures of L- and D- proteins are crystallized, using peptide synthesis to make the D-amino acid protein (which has been done by Jeremy Berg)). Once again, the noncentrosymmetric case is more complex. The same vector equation ($\vec{F}_p + \vec{f}_h = \vec{F}_{ph}$) holds, but now there are no phase restrictions. For the correct phases, a closed triangle is ideally formed by these three vectors (known as a Harker diagram, after the CIT Chemistry PhD). Again, we want to find α_p when $\vec{f}_h$, $|\vec{F}_p|$ and $|\vec{F}_{ph}|$ are known, but now, two different solutions are possible that are symmetrical about α_h: $\alpha_{p1,2} = \alpha_h \pm \theta$, where $\theta = (\alpha_p - \alpha_h)$:



$$F_{ph}^2 = f_h^2 + F_p^2 - 2f_h F_p \cos(\pi - \vartheta)$$
$$= f_h^2 + F_p^2 + 2f_h F_p \cos(\vartheta)$$
$$\vartheta = \cos^{-1}\left(\frac{F_{ph}^2 - f_h^2 - F_p^2}{2f_h F_p}\right) \cong \cos^{-1}\left(\frac{\Delta F}{f_h}\right)$$

With one heavy atom derivative (single isomorphous replacement, or SIR), in general there are two phase choices – and it is not possible to tell which one is correct in the absence of supplemental phase information, such as solvent flattening or non-crystallographic symmetry averaging. It is possible to use both phase choices - the correct phase will give the real structure, but against a background of noise from the incorrect phase choice. This can be useful, but the best thing to do is to prepare a second (or even more), different derivative, and ideally, only one phase choice will be common between the two derivatives. This is the basis of the method of multiple isomorphous replacement (MIR). (Note: there has been recent interest in the use of single wavelength anomalous diffraction (SAD) phasing, where the phase ambiguity is resolved through solvent flattening (see LM Rice, TN Earnest, AT Brunger, *Acta crystallogr*. **D56**, 1413-1420 (2000)); presumably the same approach would work with single isomorphous replacement (SIR), except that the isomorphism isn't as good as with anomalous data).

In reality, even with multiple derivatives, a unique phase solution is often not obtained, due to errors in data, errors in the heavy atom model, and non-isomorphism. These can be schematically represented as shown below:



Due to these sources of error, the "phasing triangle" may not be closed even for the correct phase. A rigorous treatment of these errors is complicated, but a useful approximation was developed by Blow and Crick (*Acta Cryst*. **12**, 794 (1959)). For a particular protein phase, α, the derivative $|F_{ph,calc}|$ can be calculated:

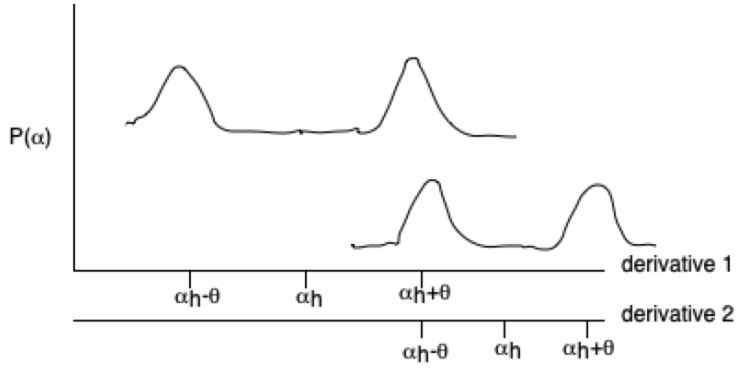
$$\left|F_{ph,calc}(\alpha)\right| = \left|\left|F_p\right| e^{i\alpha} + \vec{f}_h\right|$$

the "lack of closure" for this derivative may be defined

$$\varepsilon(\alpha) = \left|F_{ph,obs}\right| - \left|F_{ph,calc}(\alpha)\right|$$

To a reasonable approximation, the probability that α is the correct phase is given by the Gaussian expression: $P(\alpha) \propto \exp\left[-\frac{\varepsilon^2(\alpha)}{2E^2}\right]$, where E is the rms lack of closure for that derivative.

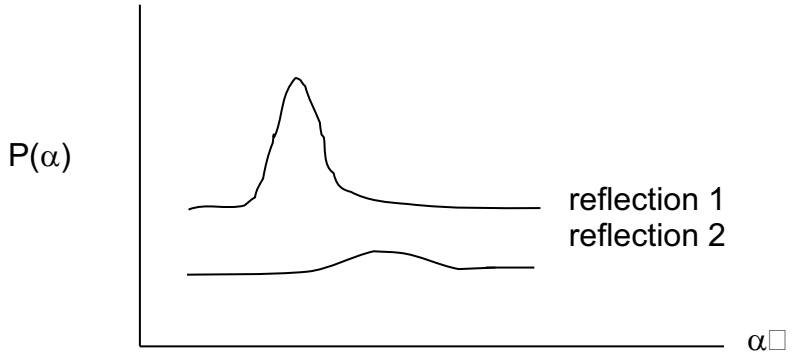
From this expression, P(α) can be calculated for each derivative:



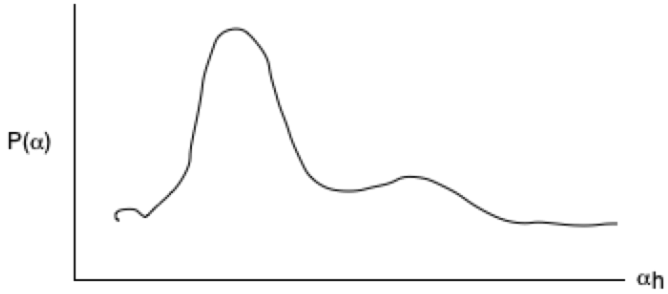
In this example, it is clear which phase to use, but a more objective criteria is required. This is done by combining phase distribution from each derivative by multiplication:

$$P(\alpha) = \prod_j P_j(\alpha) = P_{der1}(\alpha) \cdot P_{der2}(\alpha)...$$

Now: Lets take 2 different reflections:



Although both curves are unimodal, #1 is much sharper and is more likely to be correct. How can this impression be quantified? A related question is what phase do you use when the combined phase probability curve is still bimodal?



In this case, you might use α with the highest probability (the "most probable phase"). This is great if the phase is correct, but if not ..., there is a risk of a big error if the other peak (or another value) is correct. The most sensible approach would be to try some type of weighted average. Blow and Crick showed that the P(α) weighted average phase was the "best" phase in terms of minimizing the mean error taken over the entire electron density map. Here's how it works.

At the end of phasing, $|F_p|$ and P(α) are available, and you want to calculate the electron density as the Fourier transform of an amplitude and phase. For this, take the following average:

$$\left\langle \left| F_p \right| e^{i\alpha} \right\rangle = \left| F_p \right| \left\langle e^{i\alpha} \right\rangle$$

$$\left\langle e^{i\alpha} \right\rangle = \int_0^{2\pi} e^{i\alpha} P(\alpha) d\alpha \Big/ \int_0^{2\pi} P(\alpha) d\alpha$$

$$\equiv m e^{i\alpha_{best}}$$

where

$$m\cos\alpha_{best} = \int_0^{2\pi} \cos(\alpha) P(\alpha) d\alpha \Big/ \int_0^{2\pi} P(\alpha) d\alpha$$

$$m\sin\alpha_{best} = \int_0^{2\pi} \sin(\alpha) P(\alpha) d\alpha \Big/ \int_0^{2\pi} P(\alpha) d\alpha$$

$$\alpha_{best} = \tan^{-1}\left[\frac{m\sin\alpha_{best}}{m\cos\alpha_{best}} \right]$$

$$m = \left(m^2 \cos^2 \alpha_{best} + m^2 \sin^2 \alpha_{best} \right)^{1/2}$$

$$\equiv \text{ the figure of merit, with } 0 \le m \le 1$$

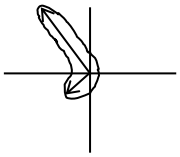
The figure of merit can be shown to equal the cosine of the average phase error, by redefining the origin of the phase circle to be $\alpha_{best} = 0$, so that m sin α_{best} =0, m cos α_{best} =m, and

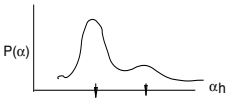
$$m = \int_0^{2\pi} \cos(\alpha - \alpha_{best}) P(\alpha) d\alpha \Big/ \int_0^{2\pi} P(\alpha) d\alpha$$

$$= \left\langle \cos(\alpha - \alpha_{best}) \right\rangle$$

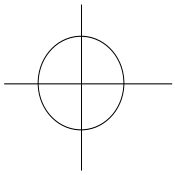
Graphically, this is equivalent to finding the centroid of the following figure, where the length of the vector = P(α). Hence, α_{best} is often called the "centroid phase". As a result of these considerations, the "proper" electron density maps to calculate are with Fourier coefficients $m\left|F_p\right| e^{i\alpha_{best}}$, the so-called figure of merit (fom) weighted map.

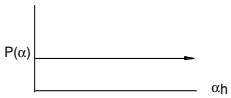
Some examples of phase probability curves:





high figure of merit





low figure of merit (m=0)

In addition to the figure of merit, other statistics that give information on the phase quality include the "phasing power" and “R cullis” (see H Ke, *Meth. Enzymol.* **276**, 448-461 1997)).

- the phasing power is defined as $<f_h/E>$, where f_h is the rms amplitude of the heavy atom scattering factor, and E is the rms lack of closure. The phasing power provides a measure of the signal to noise ratio; when it is < 1, this is bad, and if it is > 2, this is excellent. But, like all statistics (including m), one number can't always provide an objective assessment of how well things are (or aren't) working.
- R_{Cullis} is sort of like an R factor on the heavy atom f, and is defined as:

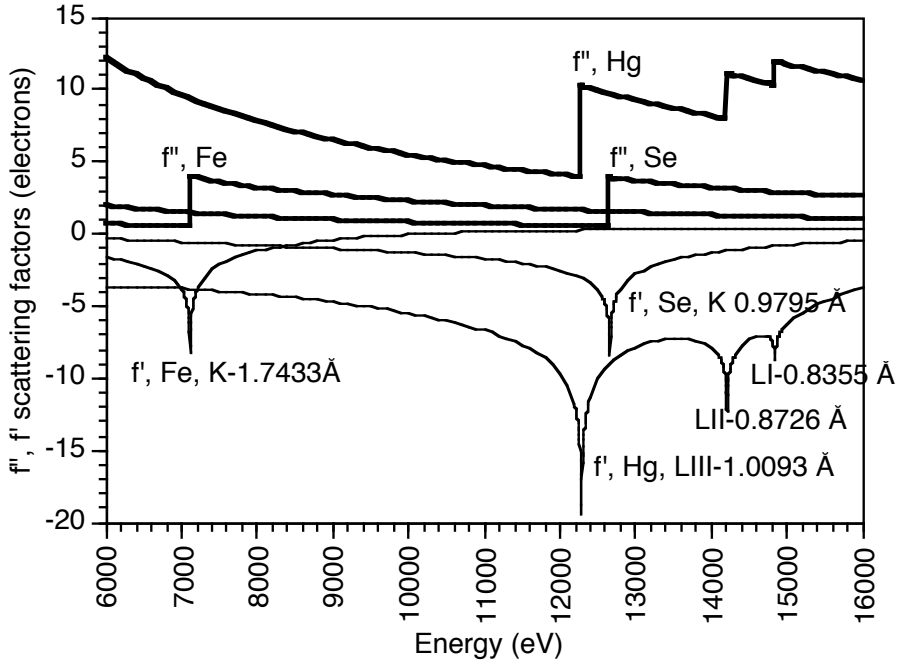
$$R_{Cullis} = \frac{\sum \left| \left(\left| F_{Ph} \right| - \left| F_P \right| \right) - \left| f_h \right| \right|}{\left| \left| F_{Ph} \right| - \left| F_P \right| \right|}$$

Strictly speaking, this should only be used with centric reflections, but it is often used for acentric reflections, as well. With perfect data and model, the ideal R_{Cullis} would be 0; for real data, values of 0.4-0.6 are excellent and higher values can still represent useful phasing information.

Anomalous dispersion

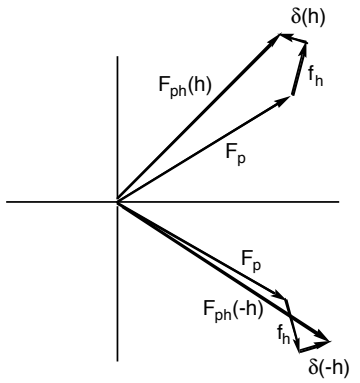
From the real nature of the electron density, we derived Friedel's law that gives |F(h)| = |F(-h)| and $\alpha(h) = -\alpha(-h)$. The hkl and -h,-k,-l reflections are known as Bijvoet pairs. For centric reflections, $\alpha = 0$ or 180˚ (at least when the center of inversion is at the origin), and so centric Bijvoet pairs have the same phase. Often, Friedel's law is an excellent approximation, but it is just that, an approximation. Under certain conditions, when the X-ray wavelength is near an absorption edge, Friedel's law breaks down, and these effects form the basis for obtaining both phase information and establishing absolute configuration.

Earlier, we discussed the X-ray absorption properties of elements. There are some analogies between the forced harmonic oscillator, and the interaction between the X-rays and electrons in an atom (bound to the nucleus). In the classical treatment of the forced oscillator, if the λ of the applied force is much longer than the native wavelength of the oscillator, λ_o, then the oscillator moves in phase with the force. If the forcing λ is much shorter than λ_o, however, then the oscillator moves 180˚ out of phase from the applied force. This is the typical case with X-rays, since the λ of the carbon K edge is 43Å, while the CuKα λ = 1.54Å. What happens when $\lambda \sim \lambda_o$? In this case, the oscillator moves 90˚ out of phase from the applied force, and energy is absorbed by the system. The 90˚ phase shift may be described by a complex scattering factor: $f = f_o + \Delta f' + i\Delta f''$. Δf' corresponds to a dispersion term related to the wavelength dependence of the refractive index (ORD), while Δf" corresponds to an absorption term related to the wavelength dependence of the extinction coefficient (CD). f_o is the "normal" scattering factor in the absence of anomalous effects. The wavelength dependences of these terms are shown below (generated from data at http://skuld.bmsc.washington.edu/scatter/):



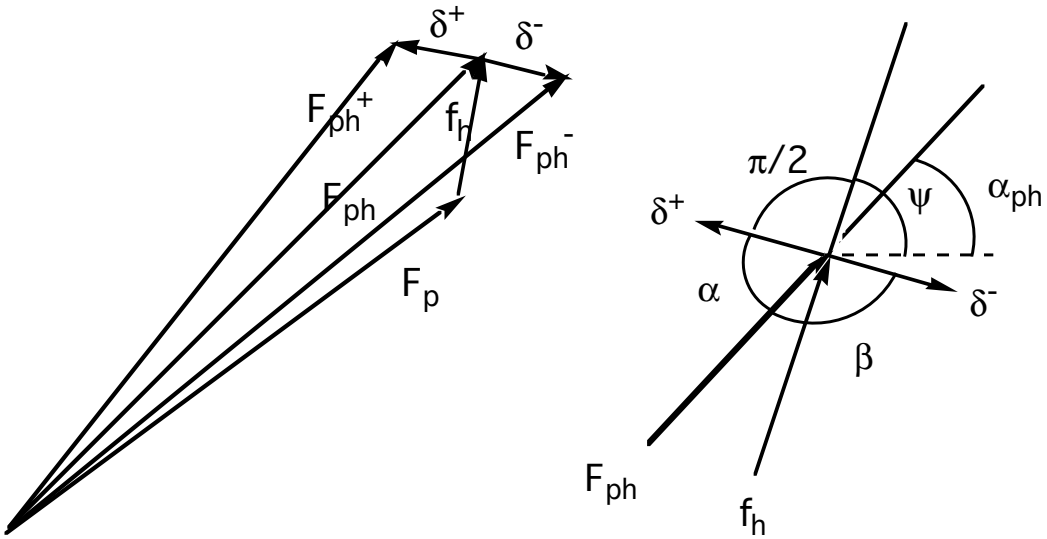
The position of the edge depends on the element and environment - if this is too anisotropic, there can be some problems with highly polarized radiation (as at synchrotrons).

As a consequence of the complex scattering factor, Friedel's law is no longer satisfied, as can be seen in the following example:



As a result, $\left|F(h)\right| \neq \left|F(\bar{h})\right|$. The magnitude of the difference, $\Delta_{ano} = \left|F(h)\right| - \left|F(\bar{h})\right|$, can be used to derive phase information that is complementary to MIR phase information obtained from the same heavy atom derivative. Additionally, and perhaps more importantly, at a tunable X-ray source (ie, synchrotron), these effects can be measured at multiple wavelengths, and in effect, each wavelength corresponds to a different derivative (so that new structures can be solved from one crystal).

Derivation of Bijvoet Differences



where ψ = the heavy atom phase
α_{ph} = phase of the normal scatterers
$|\delta|$ = magnitude of imaginary component

By the law of cosines

$$\left|F_{ph}^{+}\right|^2 = \left|F_{ph}\right|^2 + \left|\delta_h\right|^2 - 2\left|F_{ph}\right|\left|\delta_h\right|\cos\alpha$$

$$\left|F_{ph}^{-}\right|^2 = \left|F_{ph}\right|^2 + \left|\delta_h\right|^2 - 2\left|F_{ph}\right|\left|\delta_h\right|\cos\beta$$

$$= \left|F_{ph}\right|^2 + \left|\delta_h\right|^2 + 2\left|F_{ph}\right|\left|\delta_h\right|\cos\alpha \text{ (since } \cos\beta=\cos(\pi\text{-}\alpha)=\text{-}\cos\alpha)$$

subtract and factor

$$(\left|F_{ph}^{+}\right| - \left|F_{ph}^{-}\right|)(\left|F_{ph}^{+}\right| + \left|F_{ph}^{-}\right|) = -4\left|F_{ph}\right|\left|\delta_h\right|\cos\alpha$$

for small δ, $(\left|F_{ph}^{+}\right| + \left|F_{ph}^{-}\right|) \approx 2\left|F_{ph}\right|$

$$(\left|F_{ph}^{+}\right| - \left|F_{ph}^{-}\right|) = -2\left|\delta_h\right|\cos\alpha$$

Now, α equals π-$(\frac{\pi}{2} - (\psi - \alpha_{ph})) = \frac{\pi}{2} + \alpha_{ph} - \psi$

$$(\left|F_{ph}^{+}\right| - \left|F_{ph}^{-}\right|) = -2\left|\delta_h\right|\cos\left(\frac{\pi}{2} + \alpha_{ph} - \psi\right)$$

$$= 2\left|\delta_h\right|\sin(\alpha_{ph} - \psi) \text{ (since } \cos(\frac{\pi}{2} + \vartheta) = -\sin\vartheta)$$

$$\therefore \Delta_{ano} = (\left|F_{ph}^{+}\right| - \left|F_{ph}^{-}\right|) = -2\left|\delta_h\right|\sin(\psi - \alpha_{ph})$$

The largest anomalous differences occur when ψ and α_{ph} are orthogonal - this is complementary to isomorphous differences, where parallel ψ and α_{ph} give the largest isomorphous differences. An exception to the occurrence of anomalous differences are centric reflections, since ψ-α_{ph} = 0, 180˚, the anomalous difference vanishes (but, α_{ph} is no longer 0 or 180˚).

In terms of real and imaginary parts , with F_{ph} = A + iB and f_h = a + ib, Δ_{ano} may be expanded to:

$$\Delta_{ano} = -\frac{2\delta}{\left|F_{ph}\right|\left|f_h\right|}\left(bA - aB\right)$$

Anomalous differences can be used to generate Patterson functions and find heavy atoms, just like isomorphous differences. If both terms are available, an improved approximation can be calculated from $f_h^2 = \Delta_{iso}^2 + \frac{k^2}{4}\Delta_{ano}^2$, where $k = \left|f_h\right| / \left|\delta_h\right|$. Also, the anomalous differences can be used to calculate phases, in an analogous fashion to the MIR treatment. The main danger here is that the correct hand for the heavy atom coordinates, etc. must be used with anomalous data - otherwise, there are real problems (ie the *A. vinelandii* ferredoxin). This is because left handed coordinate systems interchange h and -h, and hence change the sign of Δ_{ano}.

There are advantages/disadvantages to both isomorphous and anomalous phasing.

1. With a native scatterer, the anomalous differences are measured from an exactly isomorphous crystal.

2. But, isomorphous differences are generally ~10x larger (80 electrons vs 8 electrons, unless a lucky metal/absorption edge can be found)

3. But, at a synchrotron, anomalous/dispersion differences can be measured at multiple wavelengths, permitting the solution from a single crystal. These can be done either with intrinsic metals, added metals, or Se incorporation into selenomethionine. The $\Delta f'$ changes correspond to isomorphous differences, while the $\Delta f''$ changes to anomalous differences.

4. Anomalous differences can be used to establish the absolute configuration - usually not a problem with proteins or nucleic acids, since handedness of structures is usually known (now). But, the first time, the expected anomalous differences could be calculated from the coordinates and compared to the measured values, to see if the hand is correct. The classic paper in this area is by J. Bijvoet, *Nature* **168**, 271 (1951) and the absolute configuration of tartaric acid (also, Trueblood and Glusker App. 10, pp. 217-218).

5. The correct hand is needed for anomalous phasing (see Kraut *JMB* **35**, 511-12); if the wrong hand is used then the resulting maps are garbage. For isomorphous data, the wrong hand leads to the enantiomeric structure, but otherwise everything is fine.

Multi-wavelength Anomalous Diffraction Experiments (MAD)

References
W.A. Hendrickson *Science* **254**, 51-58 (1991)
W.A. Hendrickson *Trans. Am. Crystallogr. Assoc.* **21**, 11-21 (1985)
Chapters 28-31 of *Methods in Enzymology*, vol. **276** (1997).
scattering factor tables are in http://skuld.bmsc.washington.edu/scatter/

The form (scattering) factor, *f,* of an atom is not a constant, but rather is a function of the X-ray wavelength, with significant changes evident near the absorption edges of an element. To describe these effects, *f* may be written in the form:

$$f(\lambda) = f_o + \Delta f'(\lambda) + i\Delta f''(\lambda)$$

The wavelength dependence of *f* was illustrated above for several different elements. Changes in $\Delta f'$ alter the real component of the scattering factor, which is equivalent to the types of changes produced by isomorphous heavy atom derivatives, while changes in $\Delta f''$ alter the imaginary component of the scattering factor, which is the basis of anomalous scattering effects. The maximal value of $\Delta f''$ occurs at the absorption maximum, while the maximum of $\Delta f'$ occurs at the inflection point of the absorption curve, which is at slightly lower energy. Apart from the energy of the absorption band, the K edges of all elements are essentially alike, and of all L edges are essentially alike. L_{III} edges, which are associated with the six 2p electrons, have anomalous scattering factor magnitudes ~3 times those of the K edges, which are associated with the two 1s electrons. "Typical" maximal values of $\Delta f''$ for K and L_{III} edges are ~5 and 15 electrons, respectively, while corresponding values of $\Delta f'$ for K and L_{III} edges are ~-10 and -25 electrons, respectively. At synchrotrons, wavelength windows of ~0.7 to 2 Å are available (at least at some sources and beamlines), which includes K edges for elements with Z = 24 to 42 (Cr to Mo) and L_{III} edges for a number of higher Z elements.

Ethan Merritt has posted wonderful website for anomalous scattering that tabulates edge positions and values for Δf' and Δf" as a function of energy for all elements. Wavelengths and energies (interconverted through the relationship E (keV) = 12.398/λ (Å)) for the elements depicted above, along with peak, etc. data for various MAD data sets we've collected are tabulated below:

Element	Edge	keV	Å
Hg	K	83.1023	0.1492
	L-I	14.8393	0.8355
	L-II	14.2087	0.8726
	L-III	12.2839	1.0093
	M1	3.5616	3.4811
	M2	3.2785	3.7817
	M3	2.8471	4.3548
	M4	2.3849	5.1987
	M5	2.2949	5.4026
Fe	K	7.1120	1.7433
Co	K	7.7089	1.6083
Zn	K	9.6586	1.2837
Se	K	12.6578	0.9795

wavelengths used in actual protein structure determinations from our group

protein/edge	high/remote	peak	inflection	low
Fe edge				
all ferrous FeP	1.378	1.737	1.741	
CODH	1.647	1.7394	1.742	1.846
FRD	1.65	1.74	1.741	
Aq Fd	1.692	1.739	1.742	1.744
CompA	1.692	1.7396	1.7419	
ISF	1.6531	1.7367	1.7423	1.771
hydrogenase	1.5498	1.7398	1.7418	
average (Å)		1.7388	1.7417	
average (keV)		7.1303	7.1183	
Co edge				
BtuF	1.03317	1.6027	1.6052	
keV		7.7358	7.7237	
Zn edge				
CAB (Zn)	1.0333	1.282	1.2832	
keV		9.6708	9.6618	
Se edge				
BtuCD	0.9184	0.9798	0.9800	
MscS	0.9184	0.9788	0.9791	
AfJAMM	0.9184	0.9790	0.9792	
average (Å)		0.9792	0.9794	
average (keV)		12.6614	12.6583	

note: the inflection tends to be near the edge position for the pure elemental form (although shifted to slightly higher energy). Variations in peak and inflection wavelengths between different samples reflect the effects of chemical environment, oxidation state and calibration errors in the monochromator.

In principle, data measured at two different wavelengths (including Bijvoet pairs) is sufficient to resolve the phase ambiguity (in the absence of additional information), since this would provide the equivalent of one isomorphous derivative (from the difference in amplitudes at the two wavelengths), as well as two sets of anomalous differences. In practice, since the signal is relatively weak (especially for K edges; less so for L_{III} edges), data is measured at 3 or more different wavelengths. These are generally selected at the absorption maximum, the inflection point, and one or more remote positions where both the dispersive and anomalous signals are weak (at higher and/or lower energies than the peak position).

Difference Fourier Methods

It is often difficult to identify "minor" heavy atom sites through Patterson methods, but these sites are important to obtain the most accurate phase information through the methods of isomorphous replacement and anomalous scattering. Using initial phases calculated from the "major" heavy atom sites, it is possible to detector minor sites using more sensitive difference Fourier methods. These occur in two basic flavors:

isomorphous difference Fourier maps, calculated with coefficients $\left(\left|F_{Ph}\right|-\left|F_{P}\right|\right)e^{i\alpha}$, where the initial MIR, etc. phases are used. To a first approximation, this is the difference between the electron density of the derivative, approximated by the Fourier transform of $\left|F_{Ph}\right|e^{i\alpha}$, and the electron density of the protein alone, approximated by the Fourier transform of $\left|F_{P}\right|e^{i\alpha}$. Consequently, the difference Fourier map should have positive peaks for heavy atoms that are not in the current model, but should be, and negative peaks for heavy atoms that are in the current model, but shouldn't be.

anomalous difference Fourier maps, calculated with coefficients $\left(\left|F\left(h\right)\right|-\left|F\left(\overline{h}\right)\right|\right)e^{i\left(\alpha-\frac{\pi}{2}\right)}$. The factor of $\pi/2$ subtracted from the phase is required to identify the "imaginary" component of the electron density which represents any anomalous scatterers (remember the $i\Delta$f" term in the scattering factor?). The original derivation of this was provided by Kraut (*JMB* **35**, 511-512 (1968)). If the phase term in the "normal" Fourier series is given by $e^{i\alpha}=\cos\alpha+i\sin\alpha\equiv A+iB$, then the Fourier coefficients of the anomalous difference Fourier are:

$$A'=\ \Delta_{anom}B$$

$$B'=-\Delta_{anom}A$$

Kraut also demonstrated that if the original set phases were calculated from the incorrect hand, the anomalous difference Fourier will have **negative** peaks at the **inverse** position. There can be complicated (but interesting) effects if anomalous data is used to calculate the phases for this map.

Chapter 4: Non-crystallographic Symmetry Averaging and Molecular Replacement

References

M.G. Rossmann & D.M. Blow "The detection of subunits within the crystallographic asymmetric unit" *Acta Cryst.* **15**, 24-31 (1962).
R.A. Crowther & D.M. Blow "A method of positioning a known molecule in an unknown crystal structure" *Acta Cryst.* **23**, 544-548 (1967).
F.M.D. Vellieux & R.J. Read "Noncrystallographic symmetry averaging in phase refinement and extension" *Meth. Enzymol.* **277**, 18-53 (1997).

An excellent introduction to this subject is given in:
G.J. Kleywegt & R.J. Read "Not your average density" *Structure* **5**, 1557-1569 (1997)

Introduction

Experimentally determined phases can be refined/supplemented/replaced by density modification and molecular replacement methods. There are several scenarios that we will mention (plus countless other variants that we won't). These include:

The really good:

- Multiple copies of the same molecule are present in the asymmetric unit, which permits the use of non-crystallographic symmetry (NCS) averaging to improve an initial set of phases.
- Multiple crystal forms of the same or related molecules are available. Averaging can take place between these forms, once some initial phases are available for one (or more) of the different forms.

The good:

- Density modification methods (solvent flattening or flipping), with or without NCS, can improve phases

The bad and the ugly

- If a similar structure has been solved, then one can calculate the transformation needed to orient this known molecule in the unknown cell by rotation/translation functions, allowing the initial calculation of phases from this molecular replacement model. Frankly, molecular replacement approaches should be outlawed, however, since there is always a problem of model phase bias. If combined with NCS averaging and density modifications, however, this may not be too bad. (Note: from the perspective of 2024, in an era of AlphaFold, superior data collection and refinement methods, and EM structures, this statement seems unduly harsh, but still, model bias is a real issue in X-ray crystallography).

Determination of NCS operations

An essential step in all of these problems is to determine the orientational and translational relationships between different molecules in the same crystal form, or in different crystal forms, or for a previously solved structure. Keep in mind, however, that there may be non-rigid body transformations taking place (ie "conformational changes"), so that the NCS may not be exact. These relationships may be established by some combination of:

(1) Rotation functions (RFs). RFs are used to establish the rotational relationships between different molecules. There are two types of RFs: self and cross. Self-rotation functions look for rotational relationships within a single crystal form, while cross rotation functions look for rotational relationships between crystal forms. RFs can be calculated using either intensity data, as in the first rotation functions introduced by Rossmann and Blow, and later implemented in the Fast Rotation Function of Crowther and Blow (now used in AMORE, POLARRF). RFs can also be calculated with peaks from Patterson maps (X-PLOR). These are (essentially) equivalent, and provide another illustration that (essentially) all crystallographic calculations can be performed in either real or reciprocal space.

Remember, rotation function interpretation is an experimental process; changes in variables like integration radius, shape of the integration volume, resolution and data set must be explored. It can also be important to exclude very low-resolution data, particularly when calculating RFs between an observed and calculated set of structure factors, as the calculated SFs will often be very large at very low resolution (since no solvent is present). The shape of the integration volume cannot typically be varied, but this could be important for certain types of problems (saucer-shaped or cigar shaped molecules, for example).

(2) Native Patterson functions. These are useful for identifying even fold NCS rotations that are parallel to even fold crystallographic rotation axes, or for molecules that are related by a translation. If the NCS relationships are approximate, characteristic peaks should be stronger at low resolution (8Å or so) than at higher resolutions. **Native Patterson maps and self-rotation functions should always be calculated if NCS is suspected.**

(3) Translation functions (TFs). TFs come in various flavors, depending on what sort of phase or model information is available for the various crystal forms. In the initial formulation by Crowther and Blow, molecules were positioned based on agreement between observed and calculated cross-vectors in Patterson maps. Phased translation functions can position molecules in electron density maps. These days, it is more likely that the translational parameters for known molecules will be determined by a brute force R-factor search, systematically trying all positions for a properly oriented molecule in the unit cell.

(4) Heavy atom positions. NCS parameters can be established from the positions of heavy atoms or native anomalous scatterers (but beware: they may not all obey the NCS.)

(5) Brute force search. If all else fails, one can systematically search rotation and translation space for NCS relationships in either electron density maps or with models.

(6) Initial parameters can be refined either based on Patterson maps, electron density maps, or rigid body refinement of coordinates.

Ideally, one or more of these approaches will succeed, and the NCS relationships can be determined and represented in the form:

$$x_2 = C\, x_1 + d$$

where x_1 and x_2 are the coordinates of equivalent positions in the NCS related molecules. C is the rotation matrix, and d is the translation vector. Rotation matrices are specified by 3 rotation angles, which are described in greater detail in the "Matrix Methods in Crystallography" handout. **Beware:** sometimes the transpose of the matrices and/or vectors are used, for reasons that are totally obscure to me.

Molecular Envelopes

Since these transformations are noncrystallographic, it is necessary to define a molecular envelope or mask where the NCS operations are valid. In a molecular replacement problem, the known structure can be used to construct an envelope. B.C. Wang (*Meth. Enzymol*. **115**, 90-112 (1985)) devised a very useful algorithm to determine a molecular envelope from experimental phases, that essentially looks for regions of high local density that likely correspond to protein. In this approach, a modified map is calculated in real space where the density at each grid point is proportional to the weighted sum of the positive electron density within a radius R from that grid point in the initial electron density map:

$$\rho'_j \sim \sum_i w_i \rho_i \qquad \begin{cases} w_i = 1 - r_{ij}/R, \text{ with } r_{ij} < R \text{ and } \rho_i > 0 \\ w_i = 0, \text{ with } r_{ij} > R \text{ or } \rho_i < 0 \end{cases}$$

where r_{ij} is the distance from output grid point to the surrounding grid points in the summation. The best radius for this summation varies with resolution, and has been found to be 9 and 12 Å for 3 and 6Å resolution data, respectively. The molecular boundary is established by setting a threshold density for the modified map such that the volume of density above the threshold corresponds to the fraction of protein in the map.

Andrew Leslie (*Acta Cryst.* **A43**, 124-136 (1987)) recognized that this modification is a convolution, and he devised a very efficient algorithm for calculating this modified map by taking the Fourier transform of the product of two functions in reciprocal space (yet another illustration that some calculations are easier in one space than the other).

There are many variants on the Wang/Leslie theme for envelope determination – maps can be first averaged by the NCS, and then a Wang-type algorithm used to define a mask, etc. etc.

After the initial mask is obtained, the mask should be examined in a graphics program for overlaps with crystallographically related molecules.

As the phases, NCS transformations, etc. are improved, the masks should be updated periodically.

NCS averaging

Once the NCS transformations, envelope and starting phases are available, averaging can begin. Typically, this is now performed with DM or SOLOMON. There are two basic operations involved with NCS averaging: averaging and treatment of the solvent. (Other constraints like histograph matching can also be incorporated into this process).

Averaging: just as the name implies – the densities of points related by NCS are averaged:

$$\rho(x) = \frac{1}{N}\sum_{i=1}^{N}\rho(C_i x + d_i)$$

Since they generally will not coincide with grid points, it is necessary to interpolate the corresponding electron density values from the neighboring grid points.

Solvent: The simplest thing to do with the solvent is to set all grid points outside the molecular envelopes to the average value of the solvent. In my experience, this value is very close to zero (assuming that the F_{000} term is excluded from the Fourier summation). A great improvement was the introduction of "solvent flipping" by Abrahams and Leslie (*Acta Cryst.* **D52**, 30-42 (1996)) where the solvent density at a point, ρ', is set equal to:

$$\rho'(x) = \rho_{avg} + k_{flip}\left(\rho - \rho_{avg}\right)$$

where ρ_{avg} is the average density of the solvent, and ρ is the density in the original map. k_{flip} is defined as $\gamma/(\gamma-1)$, where the so-called γ-correction $\gamma = U/V$ = molecular volume/asymmetric unit volume. For solvent flattening (no averaging) $\gamma \sim 0.5$, so $k_{flip} = -1$**;** this sets the output density to the negative of the density (relative to the average density) in the preceding map. This has been shown to accelerate convergence through overshifting of the changes in structure factors at each cycle of averaging/density modification (Abrahams, *Acta cryst.* **D53**, 371-376 (1997)).

Examples of NCS related manipulations

A. Sampling Theory

In real space, averaging involves equating the electron densities at points related by the NCS symmetry. Since the equivalent points related by the NCS are typically not all grid points, it is necessary to interpolate the electron density values from the neighboring grid points. A similar type of operation takes place in reciprocal space, where the presence of NCS imposes relationships on parts of the diffraction pattern related by these transformations. Typically, the parts of the diffraction pattern related to a given reflection by the NCS will not always coincide with reciprocal lattice points, so that it is necessary to interpolate the value of the diffraction pattern at these non-integral points. This operation involves the use of sampling theory, which is outlined below.

The molecular and crystal transforms of an object are given by:

$$F(S) = \int_{-1/2}^{1/2}\rho(x)e^{2\pi iSx}dx$$

$$F(h) = \int_{-1/2}^{1/2}\rho(x)e^{2\pi ihx}dx;\ h = \text{integer}$$

The crystal transform is given by the molecular transform sampled at reciprocal lattice points (convolution theorem).

By the inverse Fourier transform:

$$\rho(x) = \sum_h F(h) e^{-2\pi i h x} \quad \text{(neglecting the volume factor)}$$

$$F(S) = \int_{-1/2}^{1/2} \sum_h F(h) e^{-2\pi i h x} e^{2\pi i S x} dx$$

$$= \sum_h F(h) \int_{-1/2}^{1/2} e^{2\pi i (S-h)} dx$$

$$= \sum_h F(h) \frac{\sin \pi (S-h)}{\pi (S-h)}$$

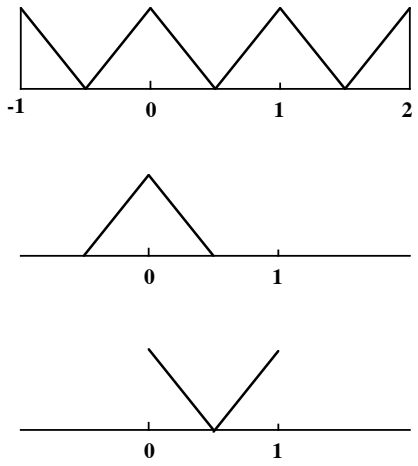
This sampling theorem permits reconstruction of the continuous molecular transform from the discrete, sampled crystal transform.

Now, for integer n, $\frac{\sin \pi n}{\pi n} = \delta(n)$, so that when S equals an integer h, $F(S) = F(h)$, and the value of this amplitude is independent of all other $F(h)$'s.

The exact form of the sampling theorem depends on the precise limits used in the integration. If instead of $-1/2 < x < 1/2$, the limits $0 < x < 1$ are used, then:

$$F(S) = \sum_h F(h) \frac{\sin \pi (S-h)}{\pi (S-h)} \left[e^{\pi i (S-h)} \right]$$

This reflects the different molecular transforms of the following objects, although they have the exact same crystal transform:



The term in brackets is related to the phase shift of the diffraction pattern associated with a real space translation.

B Structure factor relationships imposed by density modification

$$F(S) = \int_{-1/2}^{1/2} \rho(x) e^{2\pi i S x} dx$$

$$\text{if } \rho(x) = 0 \text{ when } \frac{a}{2} < |x| < \frac{1}{2} \text{ (note } \gamma \equiv \frac{U}{V} = a)$$

$$= \int_{-a/2}^{a/2} \rho(x) e^{2\pi i S x} dx$$

$$F(S) = \sum_h F(h) \int_{-a/2}^{a/2} \rho(x) e^{2\pi i (S-h) x} dx$$

$$= \sum_h F(h) \frac{\sin \pi (S-h) a}{\pi (S-h)}$$

for integer S $\equiv p$

$$F(p) = \sum_h F(h) \frac{\sin \pi (p-h) a}{\pi (p-h)}$$

In the case where $p=h$, the leading term in the summation is $aF(h) \cong \gamma F(h)$. In this case, other *F(h)*'s contribute to *F(p)* in addition to the term $p=h$. For example, when a=1/2, then for |p-h| = 0, 1, 2, 3, etc., the sinx/x term has the value 0.5, 0.319, 0, -0.106, etc., compared to the values 1, 0, 0, 0, ... when a =1, giving rise to the equation:

$$F(h) = \frac{1}{2} F(h) + 0.319 \left(F(h+1) + F(h-1) \right) + \ldots$$

In the notation of Kleywegt & Read, this may be written as:

$$F_{avg}(h) = \gamma F(h) + (1-\gamma) F_{new}(h) \quad \text{(Eq. 6) (with } \frac{\sin ax}{x} = a \equiv \gamma)$$

$F_{avg}(h)$ = the structure factor following density modification

$$F_{new}(h) = \frac{1}{(1-\gamma)} F_{avg}(h) + \frac{\gamma}{(1-\gamma)} F(h) \text{ (Eq. 7)}$$

$$F_{new}(h) = F(h) + \frac{1}{(1-\gamma)} \left(F_{avg}(h) - F(h) \right) \text{ (Eq. 8)}$$

taking the Fourier transform of this equation gives

$$\rho_{new}(x) = \rho(x) + \frac{1}{(1-\gamma)} \left(\rho_{avg}(x) - \rho(x) \right)$$

This interdependence of the structure factors permits the estimation and refinement of phase information, which is beautifully detailed in papers based on Crowther's thesis work (*Acta Crystallogr.* **22**, 758-764 (1967); *Acta Crystallogr.* **B25**, 2571-2580 (1969)), and by P. Main and M.G. Rossmann (*Acta Cryst.* **21**, 67-72 (1966))

As can be seen from this example, the dominant contribution to F(p) comes from itself (ie, F(h=p)). Separation of this term from the others forms the basis of Abrahams & Leslie's γ-correction which is at the foundation of the solvent-flipping algorithm discussed previously.

C. Application of the Sampling Theorem to the Rotation Function:

The Patterson functions of two crystals (possibly corresponding to the same crystal) are given by the expressions:

$$P_1(x) = \sum_h \left|F_h\right|^2 e^{-2\pi ihx}$$

$$P_2(y) = \sum_p \left|F_p\right|^2 e^{-2\pi ipy}$$

if $y = Cx$, where C is a rotation matrix, then

$$P_2(Cx) = \sum_p \left|F_p\right|^2 e^{-2\pi ipCx}$$

define the rotation function R (Rossmann and Blow, *Acta Crystallogr*. **15**, 24 (1962))

$$R(C) = \int_U P_1(x)P_2(Cx)dx$$

$$= \sum_h \sum_p \left|F_h\right|^2 \left|F_p\right|^2 \int_U e^{-2\pi i(h+pC)x} dx$$

$$\equiv \sum_h \sum_p \left|F_h\right|^2 \left|F_p\right|^2 G_{hp} = \sum_h \left|F_h\right|^2 \sum_p \left|F_p\right|^2 G_{hp}$$

where U is the molecular volume. G_{hp} is the sampling or interference function, and has a value near 0 unless $h \sim -pC$. G_{hp} interpolates the value of the diffraction pattern corresponding to a non-integral $-pC$ point.

With $H = pC-h$, expressions for G_{hp} for a sphere of radius R or a box with $|x| < a/2$ are given by:

$$\frac{3\left[\sin(2\pi HR) - (2\pi HR)\cos(2\pi HR)\right]}{(2\pi HR)^3}$$

and

$$\frac{\sin \pi Ha}{\pi H}$$

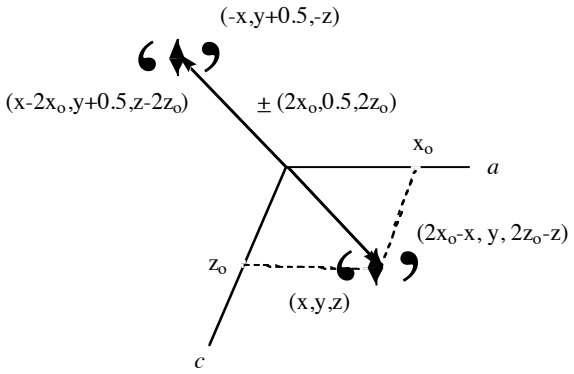
respectively.

For real molecules, the shape is much more complicated, so that G_{hp} will not be a simple function in reciprocal space which makes this calculation much more difficult. However, in real space the complexity of the calculation is not very sensitive to the details of the mask.

D. Interaction of crystallographic and NCS operations: Packing and Klug peaks

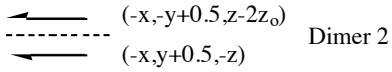
Packing peaks: Even fold non-crystallographic axes parallel to even-fold crystallographic axes give rise to large packing peaks in the native Patterson map, since the non-crystallographic symmetry in these cases can also be described as a translation. Assume a non-crystallographic two-fold axis is parallel to the b axis in space group $P2_1$, and passes through the x and z coordinates x_o and z_o. The coordinates of the packing peak are found as follows:

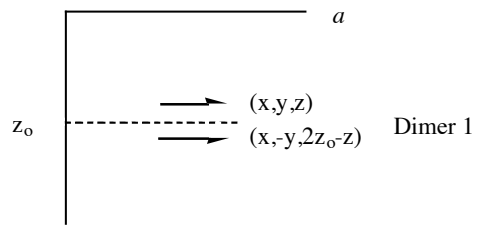
The ncs twofold along y that passes through (x_o, z_o) converts a point with (x, z) to -(x-x_o)+x_o, -(z-z_o)+z_o. The packing peak has coordinates $\pm$ ($2x_o$, 0.5, $2z_o$), which is the same as the self vector coordinates in the Harker section of a point at x_o ,z_o.



Klug peaks: One dimer per asymmetric unit crystallizes in space group $P2_1$, with cell angle β = 90.00˚. The dimer twofold axis is parallel to the a axis, and is defined by y and z coordinates y_o=0 and z_o. What does the two-fold rotation (κ=180˚) section of the self-rotation function look like for this crystal? The two dimers are related by both the crystallographic 2_1 screw operation, and a non-crystallographic symmetry operation. What is the non-crystallographic symmetry relationship between the two dimers?

The ncs two-fold parallel to a with y=0 and z=z_o relates the point (x,y,z) to the one with (x,-y,2z_o-z).





Applying the crystallographic symmetry generates Dimer 2 from Dimer 1. From the coordinates of the equivalent positions, the relationship between Dimer 1 and Dimer 2 corresponds to a non-crystallographic 2_1 screw axis parallel to *c* with a translational component along *z* of $-2z_o$, and that passes through the point $(x,y) = (0,0.25)$:

$$\begin{bmatrix} -x \\ 0.5-y \\ z-2z_0 \end{bmatrix}_{\text{Dimer 2}} = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} + \begin{bmatrix} 0 \\ 0.5 \\ -2z_o \end{bmatrix}$$

In this case, a self-rotation function will show three perpendicular two-fold axes, even though the molecule does not have 222 point group symmetry. There are no packing peaks in the native Patterson function.

E. Phased (Real Space) Translation Functions

Suppose you have:

- An experimentally determined electron density map, calculated with complex structure factors $\mathbf{F_o(h)} = A_o(\mathbf{h}) + iB_o(\mathbf{h})$.
- A properly oriented (but perhaps incorrectly positioned) model structure, which when placed in a P1 cell of the same dimensions as the unknown structure has the corresponding complex structure factors $\mathbf{F_m(h)} = A_m(\mathbf{h}) + iB_m(\mathbf{h})$.

The appropriate translation, **u**, needed to properly position the model in the unknown map may be determined by finding the **u** that maximizes the function T(**u**), which is the integral over the entire cell of the product between ρ_o and ρ_m, the observed and model electron densities, respectively:

$$T(u) = \int \rho_o(x)\rho_m(x+u)dx$$

Although T(**u**) could be evaluated by a brute force calculation (i.e. translate the model to $\mathbf{u_1}$, calculate T($\mathbf{u_1}$); translate to $\mathbf{u_2}$, calculate T($\mathbf{u_2}$); etc.), a faster and more elegant way to calculate T(**u**) for all **u** is with the following Fourier transform:

$$T(u) = \frac{1}{V^2}\sum_h \left[F_o(\overline{h})F_m(h)\right]e^{-2\pi ih\cdot u}$$

$$= \frac{1}{V^2}\sum_h \left[\{A_oA_m + B_oB_m\} + i\{A_oB_m - B_oA_m\}\right]e^{-2\pi ih\cdot u}$$

This relationship is derived as follows in one-dimension; the extension to three-dimensions is "straight-forward".

$$T(u) = \int_0^1 \rho_o(x)\,\rho_m(x+u)\,dx$$

with $\rho(x) = \frac{1}{V}\sum_h F(h) e^{-2\pi i h x}$

$$T(u) = \int_0^1 \left[\frac{1}{V}\sum_k F_o(k) e^{-2\pi i k x}\right]\left[\frac{1}{V}\sum_h F_m(h) e^{-2\pi i h(x+u)}\right] dx$$

$$= \frac{1}{V^2}\sum_k\sum_h F_o(k) F_m(h) e^{-2\pi i h u} \int_0^1 e^{-2\pi i (h+k) x}\,dx$$

the integral equals 0, unless $h = -k$, and then equals 1

$$\therefore T(u) = \frac{1}{V^2}\sum_h F_o(\bar{h}) F_m(h) e^{-2\pi i h u}$$

T(u) can also be calculated by a P1 Fourier series using coefficients A_c and B_c:

If $F_o(h) = A_o(h) + iB_o(h)$ and $F_m(h) = A_m(h) + iB_m(h)$

then $F_o(\bar{h})F_m(h) = A_c(h) + iB_c(h)$

where $A_c(h) = A_o(h)A_m(h) + B_o(h)B_m(h)$

and $B_c(h) = A_o(h)B_m(h) - B_o(h)A_m(h)$

This problem can also be solved by the convolution theorem (the Fourier transform of the product of two functions is the convolution of the Fourier transforms of the two functions).

This type of TF is known as a real space or phased translation function, which is described in more detail in the following references:

- P.M. Colman, H. Fehlhammer and K. Bartels in *Crystallographic Computing Techniques* (F.R. Ahmed, K. Huml and B. Sedlacek, eds), pp. 248-258, Copenhagen:Munksgaard (1976)
- X. Zhu, *et al. Science* **251**, 90-93 (1991)
- G.A. Bentley and A. Houdusse, *Acta cryst.* **A48**, 312-322 (1992)

The original translation function of Crowther and Blow was derived by a similar line of reasoning, except that instead of comparing two sets of electron densities, observed cross-peaks in a Patterson map were compared with the cross-peaks calculated from the known structure. In the final formulation, the Crother and Blow translation function calculates T(t), defined as:

$$T(t) = \int_V P_{01}(u,t) P(u)\,du$$

$P_{01}(u,t)$ gives the set of cross-vectors between the reference molecule (0) and a crystallographically symmetry related molecule (1) as a function of a translation t, and P(u) is the native Patterson function. T(t) can be written as a Fourier summation:

$$T(t) = \sum \left|F_{obs}(h)\right|^2 F_{\text{model}}(h) F^*_{\text{model}}(hC) \exp(-2\pi i h \cdot t)$$

the vector t is related to the real coordinate translation by the coordinates of the Patterson peak between the two crystallographically related molecules 0 and 1.

F. Brute force R-factor Searches

Another way to solve the translation problem is a brute force search of systematically placing a properly oriented molecule at each point in the unit cell, calculating structure factors (SFs) and comparing to the observed Fs. Rather than do a complete SF calculation at each position, the easy way is to calculate the SFs once, and then use the principle that a translation in real space is a phase shift in reciprocal space to calculate the translated SFs. This provides an easy way to calculate SFs from translated molecules.

The basic principle can be illustrating by showing that a translation of all atoms in a structure in real space corresponds to a phase shift of all reflections in reciprocal space. The relationship $F(h) = \sum_{\text{atoms } j} f_j e^{2\pi i h x_j}$ is useful.

$$F(h) = \sum_{\text{atoms } j} f_j e^{2\pi i h x_j} \quad \text{replace } x_j \text{ by } x_j + \Delta x$$

$$F'(h) = \sum_{\text{atoms } j} f_j e^{2\pi i h (x_j + \Delta x)}$$

$$= F(h) e^{2\pi i h \Delta x}$$

$$= F(h) e^{i \Delta \alpha_h}$$

G. Leslie's envelope algorithm

The Wang algorithm for mask calculate creates a modified map, ρ', from the original map ρ through an operation which is equivalent to a convolution of ρ with a weight function w:

$$\rho'(u) = \int \rho(x+u) w(u) du$$

By the convolution theorem, the Fourier transform of a convolution in real space is given by the product of the Fourier transforms of the two functions in reciprocal space, ie:

$$FT(\rho') = FT(\rho) FT(w) = F(h) FT(w)$$

Leslie (*Acta cryst.* **A43**, 124-136 (1987)) recognized this relationship and the computational enhancement it would permit. For the weighting function introduced by Wang, Leslie calculated:

$$FT(w) = Y(uR) - Z(uR)$$

$$u = 4\pi \sin\vartheta / \lambda$$

$$Y(x) = 3(\sin x - x\cos x)/x^3$$

$$Z(x) = 3(2x\sin x - (x^2 - 2)\cos x - 2)/x^4$$

Even a non-optimized implementation of this algorithm increased the speed of calculation by over 50-fold (from 35 hours(!) to 40 minutes).

Finally, with amplitudes and phases, an electron density map can be calculated and interpreted in terms of a molecular model. Models are generally built with interactive computer graphics programs such as O (Jones *et al., Acta crystallogr.* **A47**, 110-119 (1991)) or with an automated chain trace algorithm (wARP) (Perrakis *et al., Nature Str. Biol.* **6**, 458-463 (1999)), and the process can either be straightforward (as with a good electron density map or hopefully from a molecular replacement model), or more problematic. [Of course in 2024, O no longer exists and everyone uses Coot (Emsley and Cowtan, *Acta Cryst.* **D60**, 2126 (2004)].

How accurate are these coordinates? If the electron density map is calculated at a certain resolution, then the density values are typically calculated on a grid spacing of ~1/3 the resolution; ie a 1Å grid for 3 Å resolution. At 3 Å resolution, the density features for the protein are ~2 Å wide. Consequently, it is likely that the manually built model coordinates will be in error by at least 0.5-1.0Å. To improve the model, refinement methods are performed, based on least squares adjustment of the model to maximize the agreement between $|F_c|$ and $|F_o|$, or more precisely, to minimize $(|F_o|-|F_c|)^2$. (In principle, the refinement should be against the experimental observations (the intensities), but this is not usually done in practice).

First, some general background into least squares methods, starting with linear problems.

Linear Least Squares

n observations f_i, which depend on

m unknowns x_j.

$$f_1 = a_{11}x_1 + a_{12}x_2 + ... + a_{1m}x_m$$

.

.

$$f_n = a_{n1}x_1 + a_{n2}x_2 + ... + a_{nm}x_m$$

the a_{ij} are known coefficients

In matrix notation, these equations can be written

$$F = AX$$

$$F = \begin{pmatrix} f_1 \\ ... \\ f_n \end{pmatrix} \qquad A = \begin{pmatrix} a_{11} & ... & a_{1m} \\ ... & & ... \\ a_{n1} & & a_{nm} \end{pmatrix} \qquad X = \begin{pmatrix} x_1 \\ ... \\ x_m \end{pmatrix}$$

Let E=error vector $\equiv (F)_{obs}-(F)_{calc}$

$$E = F - AX$$

The best least squares solution for X minimizes $E^T E$

$$\Phi = E^T E = (F - AX)^T (F - AX)$$

when Φ is a minimum, $\partial\Phi / \partial X = 0$

$$\partial\Phi / \partial X = \partial / \partial X \left[X^T A^T AX - X^T A^T F - F^T AX - F^T F \right]$$

$$= A^T AX - A^T F = 0$$

$$\Rightarrow X = (A^T A)^{-1} A^T F \qquad \text{linear least squares solution}$$

Example: least squares scaling of two data sets

Given two sets of structure factor amplitudes F_i and G_i.

What is the best scale factor α that multiplies G_i?

Guess $\alpha = \sum F_i / \sum G_i$? The least squares solution is given by:

$$\Phi = \sum_i \left(F_i - \alpha G_i \right)^2$$

$$\partial\Phi / \partial\alpha = -2 \sum_i \left(F_i - \alpha G_i \right) G_i = 0$$

$$\alpha = \sum_i F_i G_i \Big/ \sum_i G_i^{\,2}$$

Linear problems are nice, but most situations are non-linear. To treat these problems, the problem is linearized by expanding the function f in a Taylor series, and truncating to first order.

$$f_i^{obs}\left(x_1^0, x_2^0, ...\right) = f_i^{calc}\left(x_1, x_2, ...\right) + \frac{\partial f_i}{\partial x_1}\left(x_1^0 - x_1\right) + \frac{\partial f_i}{\partial x_2}\left(x_2^0 - x_2\right) + ...$$

$x_1^0, ... = x$ values at the minimum of f

$x_1, ... =$ current x values

$$f_i^{obs} - f_i^{calc} = \sum_j \frac{\partial f_i}{\partial x_j}\left(x_j^0 - x_j\right)$$

$$\Delta f_i = \sum_j \frac{\partial f_i}{\partial x_j} \Delta x_j \quad \Rightarrow \text{this expression is linear in the shifts to } x$$

$F = AX$ matrix notation - equivalent to linear LSQ

The application to crystallographic refinement problems is as follows:

F = vector with $\left|F_o(h)\right| - \left|F_c(h)\right|$

A = matrix of derivatives $\dfrac{\partial\left|F_c(h)\right|}{\partial x_i}$

X = vector with shifts = $\left(A^T A\right)^{-1} A^T F$

where

$$\left(A^T A\right)_{ij} = \sum_h \frac{\partial\left|F_c(h)\right|}{\partial x_i}\frac{\partial\left|F_c(h)\right|}{\partial x_j}$$

$$\left(A^T F\right)_i = \sum_h \frac{\partial\left|F_c(h)\right|}{\partial x_i}\left(\left|F_o(h)\right| - \left|F_c(h)\right|\right)$$

The calculations of the necessary derivatives are performed as follows:

need $\dfrac{\partial\left|F_c(h)\right|}{\partial x_i}$

$$\begin{aligned} F_c(h) &= \left|F_c(h)\right| e^{i\alpha_h} \\ &= \left|F_c(h)\right|\cos\alpha_h + i\left|F_c(h)\right|\sin\alpha_h \\ &= A + iB \end{aligned}$$

Intensity $I = \left|F_c(h)\right|^2$

$$dI = 2\left|F_c(h)\right| d\left|F_c(h)\right|$$

$$\frac{dI}{2\left|F_c(h)\right|} = d\left|F_c(h)\right|$$

now: $I = A^2 + B^2$

$$dI = 2AdA + 2BdB$$

$$\begin{aligned} d\left|F_c(h)\right| &= \frac{dI}{2\left|F_c(h)\right|} = \frac{A}{\left|F_c(h)\right|}dA + \frac{B}{\left|F_c(h)\right|}dB \\ &= \cos\alpha_h \frac{dA}{dx_i} + \sin\alpha_h \frac{dB}{dx_i} \end{aligned}$$

Space group specific expressions for A and B are found in volume I of the International Tables. Generally, the parameters x_i to be refined include coordinates, temperature factors, scale factor, and occasionally, occupancy.

In a macromolecular refinement, often have > 10^4 or more parameters to refine, which means that the A^TA matrix is 10^4x10^4; this is difficult to invert, so many approximations are used to solve the least squares equation. Unlike the situation with small molecule crystallography, macromolecular crystal structure determinations are generally only marginally overdetermined (although this is starting to change with the refinement of macromolecular structures at resolutions near or below 1 Å). Hence, the least squares equations are supplemented with stereochemical terms to maintain good geometry. These types of refinements are called "restrained", and can

include target values for bond distances and angles, planes, non-bonded contacts, and torsion angles. This can introduce a bias into stereochemical parameters that must be kept in mind when analyzing structures, especially for parameters like metal-ligand bond distances.

Two general approaches are used for the LSQ of restrained equations - (1) try to solve the equations by a conjugate-gradient or related method (ie, move "downhill" to closest local minimum). This is found in programs like REFMAC, TNT or PROLSQ, or (2) use molecular dynamics to try to get out of local minimum and find deeper local minimum - maybe even the global minimum. This is the philosophy of X-PLOR/CNS.

The problem of local vs global minimum is serious - just because a model refines doesn't mean it is correct. An important question, without a foolproof answer, is how good is a model? In a real least squares refinement, the $(A^TA)^{-1}$ matrix provides estimates of the uncertainties in the refined parameters; but, this matrix is not calculated in macromolecular work, with the (current) exception of very high resolution refinements of structures with the SHELX package. An alternate approach (Luzzati, *Acta cryst.* **5**, 802 (1952)) is to use the agreement between $|F_o|$ and $|F_c|$, as provided by the so-called R-factor:

$$R = \sum_h \frac{\left||F_o(h)| - |F_c(h)|\right|}{|F_o(h)|}$$

R is similar, but not identical to what is minimized in the least squares refinement. A good R would obviously be 0, but what is a bad R, say for a random structure? This result was first derived by Wilson: for non-centric and centric structures, the random R (on F's) is 0.586 and 0.828, respectively, while the random R (on I's) is 1.00 and 1.27, respectively. In general, would expect R to be > 0 and < 0.586. R is resolution dependent, since the effects of coordinate errors are greater at higher resolution. For a "typical" refinement of a macromolecule, the final R is ~0.2, or less. **Coordinate errors are typically estimated as ~0.1 to 0.2 of the limiting resolution - ie, they are on the order of several tenths of Å, on average**. This has important consequences for determining accurate metal-ligand bond distances, for example.

Use of R_{free}

Brünger made a key contribution through the introduction into crystallography of the statistical method of cross-validation, most notably the R_{free} index, to monitor problems due to over-fitting of the experimental data during structure refinement. An extensive description of R_{free} is provided by A.T. Brünger (*Meth. Enzymol.* **277**, 366-396 (1997)). A test set of reflections, typically 5-10%, is extracted from the reflection file, and the structure refinement proceeds against the remaining, working set reflections. If the refinement is robust, then the R factor between the calculated and observed structure factors should be reduced for both the working and test set reflections; if over-fitting is occurring, the R_{free} will remain unchanged or even increase despite reductions in the regular R. R_{free} should be no lower than R; for initial molecular replacement solutions, R and R_{free} should be the same (within statistical variation), since the number of molecular replacement parameters is so small relative to the number of reflections that overfitting is very unlikely. Acceptable differences between R and R_{free} are not as well understood – for some structures, they are virtually the same, while for others there can be an appreciable variation. Brunger suggests that a threshold for R_{free} is 40%, which roughly would correspond to an overall average coordinate error of 1 Å. Ideally, we would like to see R below 25% and R_{free} below 30% (for comparison,

before R_{free}, essentially all structures were refined to R < 20% - often at the expense of (then unappreciated) overfitting).

R_{free} can also be used to optimize weights given to structure factors and geometrical terms during refinement. Particularly important applications involve the use of non-crystallographic symmetry restraints. We have noted that it is often necessary to use rather strict restraints at the beginning, and relax these only after refinement has proceeded to avoid overfitting; removing NCS restraints at the beginning can often lead to classic overfitting symptoms, with R rapidly decreasing and R_{free} rapidly decreasing. Interestingly, the geometry observed at the end of most macromolecular refinements is closer to ideal than observed in small molecule structures.

Brünger has suggested carrying out the refinement multiple times with different test sets to test for variations in R_{free}; provided sufficiently large reflection sets are used (> 500 reflections) the variation in R_{free} should be small, so that this type of analysis is typically not conducted.

When refining a series of isomorphous structures, it is essential that the same test set reflections are used in all refinements!! Otherwise, the working and test set reflections are not independent and the R_{free} is meaningless. Also, if Friedel mates are used in the refinement, then it is important that each pair be in either the test or work set, but not both.

Rebuilding Maps

Several types of electron density maps are useful for examining the quality of atomic models and identifying problems:

(a) difference Fourier maps, calculated with coefficients $\left(|F_o| - |F_c|\right)e^{i\alpha}$, where the phases are either calculated from the model, experimental or combined. To a first approximation, this is the difference between the true electron density, approximated by $|F_o|e^{i\alpha}$, and the calculated (model) electron density given by the Fourier transform of $|F_c|e^{i\alpha}$. Consequently, the difference Fourier map should have positive peaks for atoms omitted from the model, and negative for atoms incorrectly included in the model. From the gradient of this map, coordinate shifts can be determined.

(b) electron density maps calculated with Fourier coefficients $\left(2|F_o| - |F_c|\right)e^{i\alpha}$. Because α_{calc} is only an approximation to the true phases, the peak and troughs in an $|F_o|e^{i\alpha}$ electron density map corresponding to features either omitted or incorrectly included in the atomic model, respectively, are about 1/2 of the true value. Hence, a better approximation to the true electron density is given by the following type of map with Fourier coefficients $|F_o|e^{i\alpha} + \left(|F_o| - |F_c|\right)e^{i\alpha} = \left(2|F_o| - |F_c|\right)e^{i\alpha}$. Other, more complex types of amplitude weighting terms can be generated (Read, *Acta Cryst* **A42**, 140 (1987)).

These days, computational approaches for obtaining and refining models are extremely powerful - it is essential to remember that once a model is used for phasing, the electron density maps will always be biased towards those features - this makes it hard (almost impossible) to find errors. Like most computational processes - garbage in - garbage out! The three best ways to

avoid problems are to have good experimental data; to have good experimental data; and to have good experimental data.

Structure validation programs

General references

G.J. Kleywegt "Validation of protein crystal structures" *Acta crystallographica* **D56**, 249-265 (2000).
R.A. Laskowski, M.W. MacArthur and J.M. Thornton "Validation of protein models derived from experiment" *Curr. Op. Struct. Biol.* **8**, 631-639 (1998)
G.J. Kleywegt and T.A. Jones "Where freedom is given, liberties are taken" *Structure* **3**, 535-540 (1995).

Structure validation refers to the analysis of structures to detect problems in the crystallographic analysis. These can range from gross errors, with misconnections, incorrect folds, substantial register errors, etc. than can be detected from the behavior of R_{free}, problems with overall geometry (PROCHECK) and environmental profiles (VERIFY3D). These are becoming increasingly rare, particularly due to the use of R_{free}. Most common errors involve local problems with mainchain or sidechain fitting; these can be identified through real space correlation analysis (how the model fits the electron density (calculated with O or SFCHECK); local problems in the Ramachandran plot (PROCHECK); peptide flips; comparison of NCS related molecules; etc. Although we have not been using SFCHECK, this looks like an excellent program for identifying both global and local problems in protein structures.