Diffraction - Theory & Data Collection

Lewis & Clark Workshop
Macromolecular Crystallography
© Michael S. Chapman

X-RAYS

Why X-rays?
- Atoms scatter X-rays in all directions.
- Phase change is exactly π rad.
- Measure sum of scattering.
- Path length through 2nd atom differs.
  \[ \phi_{\text{atom } 1} \neq \phi_{\text{atom } 2} \rightarrow \text{interference.} \]

The optimal wavelength
- Make much longer
  \[ \lambda \sim \text{interatomic distances.} \]
- Choose 0.7 Å ≤ \( \lambda \) ≤ 1.8 Å
- Commonly \( \lambda \sim 1.5 \) Å ~ C–C bond.

X-rays have wavelengths close to optimal
- Max. interference: path length difference = \( \lambda \).
- \( \lambda \sim \text{interatomic distances.} \)

Electromagnetic waves.
- Amplitude (A) varies as cosine of distance from origin (O). |F|
- Wavelength = \( \lambda \), split into 360° or 2\( \pi \) radians.
- Phase (\( \phi \) or \( \alpha \))
  - Measured:
    * angle from +x-axis (anticlockwise)

Definitions.
- Electromagnetic waves.
- Amplitude (A) varies as cosine of distance from origin (O). |F|
- Wavelength = \( \lambda \), split into 360° or 2\( \pi \) radians.
- Phase (\( \phi \) or \( \alpha \))
  - Measured:
    * angle from +x-axis (anticlockwise)
Why Diffraction? -
Crystallography in a nutshell
- X-ray refractive indices close to 1.0.
- No refraction or reflection → No lenses.
- Uncharged → No electrostatic lenses like E.M.
- Left with diffraction!
- X-rays scattered & reconstructed computationally.

Atoms or Electron Density
- What is scattering the X-rays?
- Atoms → Not exactly – electron clouds
- Image is electron density
  - Atomic Structure is an interpretation...
  - Nuclear positions consistent with electron density.
  - Exp. error in density can → difficult interpretation.

Crystallography in a nutshell
- Microscopy without the lens...
- Lens would physically "sum" scattered waves.
- Crystallographers have to "sum" waves computationally.
- Measure intensity in each direction.
- Amplitudes not enough
  - Phases – synchronization of waves
    - How they line up, how far peaks lag behind each other.
    - Can't be measured directly - "Phase Problem" Challenge.

Conventional sources of radiation
- e− acceleration → X-rays.
- Fire electrons into target.
- High voltage filament
- Be window
- Copper target
- Water cooling
- X-rays

Synchrotrons
- e− accelerated in circular path
  - Relativistic speeds
- Energy (X-rays) tangential
- High intensity
  - More data before damage spreads
- Tunable wavelength
- Phasing
- Shared multi-user facilities

Preparation of X-rays
- Fine parallel beam
  - Collimator
  - Focusing mirrors
- Monochromatic (single wavelength)
  - 1.51 Å or 0.91 Å or adjusted 0.7 – 2.0 Å
- Monochromator
- Focusing mirrors
Interaction of X-rays with samples
- Absorption (fluorescence)
  - Mostly minor annoyances
- Scattering
  - Important - because no lenses for imaging
  - Need understanding of how atomic structure affects the scattering of materials
- Elastic (Coherent aka Thompson) scattering
  - No loss of energy ($\lambda$ unchanged; $\Delta \phi \approx \pi$)
  - Dominates diffraction.
- Inelastic (Incoherent aka Compton) scattering
  - Collision w/ e⁻ (energy transfer)...
  - Background scattering

Electromagnetic waves.
$$E = E_0 e^{i \omega t - k_0 \cdot r}$$
- $E_0$ is field @ time $t$, position, $x$ from origin $O$.
- $\omega \cdot \mathbf{r}$ represents a complex vector.
- Only amplitude observed, but complex form allows addition of waves.
- By de Moivre’s theorem ($e^{i \theta} = \cos \theta + i \sin \theta$):
  - Field accelerates a charged particle with frequency $\nu$.
  - Max. acceleration as particle passes node @ max $E_i$.
  - Thus (electron) particle displacement $\pi/2$ from $E_i$.
  - The accelerating orbital electron initiates a second electromagnetic wave with a 2nd phase change of $\pi/2$.

Interference of Scattered X-rays
- Scatter from $O$ and $O'$ related by vector $r$.
- Path-difference:
  - $|OA| + |OB| = |r| \cos \theta_O' + |r| \cos \theta_O'$
- $s_o$ unit vector along incident direction.
- $s$ unit vector along scattered direction.
- $\Delta \phi = (2\pi/\lambda) r \cdot (s - s_o)$

William & Lawrence Bragg
- Max von Laue demonstrated wave character of x-rays
  - Diffraction by crystals
- Braggs - father & son
  - 1913 - 1914
  - Realized that x-ray scattering could be imagined as reflection from planes
  - Solved NaCl.
  - Nobel prize 1915
Atomic Structure Factor eqn.

\[ F(r^*) = \sum_{j=1}^{N} A_j \exp 2\pi i r^* \cdot r_j \]

where

- \( A_j \) is the atomic scattering factor
- \( r_j \) is the position of the jth atom
- \( r^* = \frac{r}{\lambda} \)
- \( \lambda \) is the wavelength of the X-rays
- \( \rho_j \) is the electron density at \( r_j \)

From last lecture:

- Inverse Fourier transform:
  \[ \rho(r) = \frac{1}{V} \mathcal{F}^{-1}[F(r^*)] = \frac{1}{V} \int \ldots \int \mathcal{F}[F(r^*)] \exp -2\pi i r \cdot r^* \, dr^* \]

Diffraction of crystals:

- Only strong diffraction spots
- Insensitive intensity between spots

Tabulated theoretical or experimental values.

**Fourier Transforms (1)**

- If \( f(x) \) is piecewise integratable, it can be approximated by a Fourier transform (T):
  \[ f(x) = \int_{-\infty}^{\infty} F(x^*) \exp -2\pi i x x^* \, dx^* \]
  \[ F(x^*) = \int_{-\infty}^{\infty} f(x) \exp 2\pi i x x^* \, dx \]

- **Fourier Transform of a Step Function**
  \[ g \cdot x \geq 0 \Rightarrow f(x) = \begin{cases} \exp 2\pi i x x^* & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases} \]
  \[ F(x^*) = \int_{0}^{\infty} \exp 2\pi i x x^* \, dx = \frac{1}{2} \sin 2\pi x x^* \]

- **Scattering by elements of electron density**
  \[ F(r^*) = \sum_{j=1}^{N} A_j \exp 2\pi i r^* \cdot r_j \]
  \[ \text{Let } r_j \text{ be small element of electron density, } \rho \]
  \[ \text{Consider total scattering: } \]
  \[ F(r^*) = \int \ldots \int \mathcal{F}[F(r^*)] \exp -2\pi i r \cdot r^* \, dr^* \]

- **Infinite 1-D lattice -- Lessons**
  \[ \mathcal{F}[\text{lattice}] = \text{(lattice), in reciprocal space.} \]
  \[ \text{Spacing of reciprocal lattice inversely proportional to real lattice.} \]
  \[ \text{Diffraction of crystals } \rightarrow \text{relatively strong diffraction spots and insignificant intensity between spots.} \]
  \[ \text{Need only use lattice points } \rightarrow \text{discrete transform.} \]
Fourier Series

- $\mathcal{F}$ can approximate any piecewise integrable function.
- Coefficients each have amplitude and phase.
- Effect of truncation

Electron Density Equation

$$\rho(x) = \frac{1}{V} \sum F_h \exp(-2\pi i h \cdot x)$$

- $F_h$ are structure factor amplitudes (measured experimentally).
- $\phi$ difficult to determine $\Rightarrow$ "The phase problem"

Structure Factor Equation

$$F_h = \frac{1}{V} \sum y(x) \exp(2\pi i h \cdot x)$$

- $F_h = |F| \cos \phi$, $B = |F| \sin \phi$
- $F$ is structure factor amplitude
- Measured in fractional units
- Fraction of unit cell $(a,b,c)$
- $V = a \cdot b \cdot c$

Note symmetry with electron density equation

Atomic Structure Factor Equation

$$F_{hkl} = \sum_{j=1}^{N} f_{\text{at},j} \exp(2\pi i (hx + ky + lz))$$

- Summing over atoms instead of grid points
- Computationally less efficient if many atoms
- Structure factor equation can use FFT

Crystalline Lattices

- Previously learned...
  - Diffraction pattern = Molecular transform sampled at reciprocal lattice points.
  - Reciprocal lattice = FT or diffraction pattern from crystalline lattice
- Why understand real-space lattices?
  - Prediction of location of diffraction spots.
  - Interaction of neighboring molecules in crystal lattice
    - (molecular symmetry)
Lattices

- Grid: equally spaced || lines (planes).
- Choice of origin is arbitrary.

Unit Cells

- Axes labeled \((a, b, c)\) by convention.
- Right-handed system
- Specify 3 vectors
  - \(a, b, c\)
- or 6 parameters
  - 3 lengths + 3 angles
    - \(a, b, c\) \(\alpha, \beta, \gamma\)
    - \(\alpha\) = angle between \(b\) and \(c\) axes etc.
- Faces labeled \(A, B, C\)
  - \(A\) intersected by \(a\)-axis etc.

Lattice Planes in 2-D

- Line can be drawn between any 2 grid points.
- Parallel lines pass through every grid point.

Bragg's Law (Real space) 1

- Previously showed scattering like reflection from plane through scatterer.
- Path length same for all points on plane:
- All points scatter in phase.
- Planes containing many identical scatterers have strong scattering (Lattice planes).

Bragg's Law -- Real space 2

- Consider || planes \(P_1, P_2, \ldots, P_j, P_{j+1}, \ldots P_N\).
- Path differences: \(\Delta(P_{j+1} - P_j) = 2d \sin \theta\)
- \(\sum \text{planes scatter much larger when } \text{in phase}.\)
- \(\sum \text{... path difference } = 2d \sin \theta \text{=} n\lambda; \text{ } (n=1)\)

Bragg's Law -- Reciprocal lattice

- Diffraction only at integral path difference = \(2d \sin \theta \text{=} n\lambda\).
- Spots spaced regularly
- Lattice in 3-D
- This reciprocal lattice has spacing inversely proportional to crystal lattice spacing
- Directions of axes depend on crystal
  - Let \(d_{100}\) be unit vector || \(d^{*}_{100}\)
  - \(d^{*}_{100} = b \times c / (b |c| \sin \theta)\)
  - \(d^{*}_{100} = 1 / d_{100}\)
  - \(|d_{100}| = V / (b |c| \sin \theta) = 1 / d^{*}_{100}\)
  - \(a^{*} = b \times c / (d_{100} b |c| \sin \theta) = b \times c / V = a^{*}\)
- (Remember that we are using bold for vectors)
Bragg’s Law – Graphically -- Implications

- For any point, $P$, on the surface of the sphere, $OP = 2\sin \theta / \lambda$.
- $OP = d^*$ is a solution to Bragg’s law.
- Diffraction only at reciprocal lattice points.
- Scattering || $OP$ when sphere and lattice point superimpose.
- To see more spots, rotate crystal.

Resolution

- Let $D_{\text{max}}$ be the distance of furthest spot from the direct beam.
- Let $d_{\text{min}}$ be its interplanar spacing.
- $d_{\text{min}} = \lambda / (2 \sin (\theta_{\text{max}}) = \lambda / 2 \sin \left(\frac{1}{2} \tan^{-1}(D_{\text{max}}/l)\right))$
- $d_{\text{min}}$ is de facto resolution limit.
- Note $d_{\text{min}}$ reflection at max, i.e. farthest from beam.

Bragg’s Law Graphically – Ewald construction

- For given wavelength and crystal orientation...
  - Only those relps on surface of the “sphere of reflection” observed.
  - Diffraction only at reciprocal lattice points.
  - When sphere and lattice point superimpose.
  - Scattering || $OP$.

Diffraction Conditions

- Reminder: Bragg’s Law shows that diffraction only occurs when a reciprocal lattice point lies on the sphere of reflection.

DATA COLLECTION & PROCESSING

SYMMETRY

Topic to be skipped...
No reflections?
- No reflections on sphere of reflections if
  - Truly monochromatic radiation
  - and infinitely thin reciprocal lattice points
- Why do we see any reflections for still crystal?
  - Filters & Monochromators: \( \Delta \lambda \neq 0 \).
  - "Mosaic spread" \( \rightarrow \) finitely large reciprocal lattice points.

Mosaic Spread
- Crystal = a mosaic of sub-micron crystalline blocks
  - separated by fault lines or other defects.
  - The variation of orientation \( 0.15^\circ \).
- Filters & Monochromators: \( \Delta \lambda \neq 0 \).
- "Mosaic spread" \( \rightarrow \) finitely large reciprocal lattice points.

Still Photography
- Reciprocal lattice points on lattice planes.
- All diffracting relps at intersection of plane & Ewald sphere.
- Intersection is a circle
- Projected to film at an angle \( \rightarrow \) ellipse
  - If film \( \perp \) incident beam,
  - Reflections on ellipse.

Zones
- Not one, but a set of \( \parallel \) planes.
- \( \rightarrow \) concentric ellipses of reflections.
- Direct beam \( \rightarrow \) reflection F000; layer 0
- Layers \(-1, -2, \ldots \) inside
- Layers \(+1, 2, 3, \ldots \) outside

Zones - example
- Crystal rotated
  - (a little)
- Direct beam \( \rightarrow \) reflection F000; layer 0
- Layers \(-1, -2, \ldots \) inside
- Layers \(+1, 2, 3, \ldots \) outside

Moving the crystal
- Still crystal
  - few spots on each ellipse
- Rotate crystal
  - Red regions pass through Ewald sphere
  - Additional reflections seen
**Diffraction of moving crystal**
- All spots between ellipses between
  - Start of rotation
  - End
- Regions of reflections = "lunes"

**Lunar or spatial overlap**
- Wide rotation
- Fat ellipses
- Many reflections from each plane
- Planes project to some part of detector
  - Lunes overlap
  - Spots may overlap

**Avoiding Overlap**

**Oscillation method**
- Small rotations
- Overlap more with finely spaced planes
  - Large unit cells
  - Rotations < 0.5°
- Typical unit cells
  - Rotations: 2.5 to 5°
  - Also depends on
    - Resolution
    - Mosaic spread

**Precession method**
- Metal screen
  - Allows through x-rays of one plane only
- Circular aperture
- Large rotations in all directions
  - ~ All reflections in one plane
- Film rotated to be || to crystal plane
  - Undistorted image of lattice
  - "Precession"

**Rotation or Oscillation Photography**
- Crystal rotated - small angle, $\Delta \phi$ \perp beam; ...
- 10 fast oscillations better than 1 slow rotation:
  - Averages crystal decay & beam fluctuations
  - during angular sweep.

**Rotate the Crystal or the Generator?**
- Most prefer to rotate the crystal
  - Diagrams easier if rotate the laboratory
  - Equivalent

**Full and Partial Reflections**

- Reciprocal lattice points have finite size
  - Mosaic spread
- Whole reflections
  - Crossed completely by Ewald sphere w/in $\Delta \phi$.
- Partial reflections
  - On sphere at start or end of rotation
  - Only part of reflection recorded
  - Fraction recorded = "partiality"
Structure amplitudes from partials
- Measured intensity = $F^2$/partiality
- Partiality depends on experiment, not structure
- Need to correct
  - Make equivalent to full reflection
  - Partiality = 1
- Two approaches
  1. Series of contiguous oscillations
     - Add intensities from successive images
     - Errors - changing beam intensity etc.
  2. Estimate partiality \( \rightarrow \) mathematical correction
     - Errors - Requires accurate understanding of crystal orientation
- Partial reflections less accurate

Crystal Alignment
- Some orientations \( \rightarrow \) more efficient collection
  - Approximate alignment OK (not like precession)
  - Determine orientation then re-orient
     - Center zones from still images
     - Computer-auto-indexing (details later)
     - From 1 to 3 still images
- No longer necessary to precisely orient before data collection
- Some use the American Method
  - Shoot first - ask questions later
  - When crystal lifetime in beam is short

Freezing Crystals
- Vitreous ice (glass) OK
- Ice crystals:
  - Destroy protein crystal
  - Add ice diffraction
  - Must be avoided
  - Cryo-protection
    - Antifreeze
    - Glycerol, PEG, MPO...
    - Oil coating
    - Search for one that does no damage
    - Serial transfer
    - Reduce osmotic shock
- Flash freezing - 2 ways:
  - Plunge into liq N\(_2\), propane
    - Then put on x-ray camera
    - Keeping at 100K
  - Place on camera at 20º
    - Quickly start N\(_2\) stream
    - Maintain in cold N\(_2\) gas stream

Crystal Mounting
- Cryo-data collection
  - Drop of frozen mother liquor
  - Held in loop of fiber
  - Scooped up using surface tension
  - Open crystal \( \rightarrow \) better heat conduction to liq N\(_2\)
- Collection at 4 to 30ºC
  - In sealed capillary
  - Maintain humidity
  - Stop crystal drying
  - Capillary made of quartz
  - Minimize x-ray absorption

Data Collection Instrumentation - Synchrotron
Data Collection Instrumentation

Helium path to minimize air-absorption

Motor to rotate crystal

Detector (image plates + automatic scanner)

Crystal here (honestly!)

Thanks to Soma

Motor to rotate crystal

Detector (image plates + automatic scanner)

Crystal here (honestly!)

Generator

Detectors

| Type          | Advantages        | Disadvantages                  |
|---------------|-------------------|--------------------------------|
| Scintillation counter | Most precise      | One spot at a time (small unit cells) |
| Film          | Spatial resolution + 50 μm | Background, dynamic range, turnaround time |
| Multiwire     | Precise           | Spatial resolution, expense     |
| TV detector   |                   | Unstable, calibration, expense  |
| Image plate   | Size, Medium precision | Scan-time                      |
| CCD's         | Precise, dynamic range | Large ones very expensive      |

Data Processing - Overview

- Determine which region contains each reflection
  - "Indexing"
  - Orientational refinement
- Integrate intensity near reflection center
- Subtract background
- Scaling:
  - Correct for factors that don't depend on structure, but
    - Geometry of data collection
    - Shape and absorption of crystal
- Merging - symmetry equivalent measurements
  - Calculate quality statistics

Scaling - Introduction

- Calculate image scale constant
  - Intensities agree w/ symmetry equivalents
- Approx. correction for many factors:
  - Crystal Absorption:
    - Depends on path length
    - Depends on crystal orientation and individual reflection.
  - Other absorption: capillary, solvent -
    - Varies slowly, can be minimized.
  - Volume of crystal in beam:
    - Depends on φ, therefore image #
  - Decay due to radiation damage:
    - Depends on resolution and time (→ Δφ, image #).

Quality of Scaling

- Quality of data is now assessable with a scaling R-factor:
  \[ R_i = \frac{\sum_h |\langle I_h \rangle| - K_h \langle I_h \rangle}{\sum_h \langle I_h \rangle} \]
  - Compare this R-factor to the least-squares residual (or corresponding variance):
    - Both depend on magnitude of difference.
    - Residual is squared, so more sensitive to large differences.
    - R-factor is normalized, and expressed as decimal fraction or %.
  - \( R_i (R_{sym}) \) is calculated from intensities.
    - Most other R-factors are calculated from \( |F| = \sqrt{I} \).

Quality of Data

- Ri is used as a measure of data quality:
  - \( R_i \leq 3\%: \) excellent.
  - \( R_i \leq 5\%: \) typical average protein.
  - \( R_i \leq 9\%: \) typical large protein.
  - \( R_i \leq 13\%: \) typical virus capsid.
- Partial reflections are usually excluded
  - Underestimate of error, especially for large molecules (small \( Δφ \), many partials).
  - Inclusion of partials for virus → \( R_i = 18\% \).