X-Ray Crystallography

"If a picture is worth a thousand words, then a macromolecular structure is priceless to a physical biochemist." – van Holde

Topics:

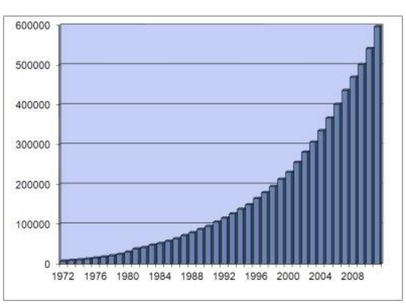
1. **Image Formation** *(optical illusions)*
   Resolution / Wavelength (Amplitude, Phase) / Light Microscopy / EM / X-ray / (NMR)
2. **Protein Data Bank (PDB)**
   Data mining and Protein Structure Analysis Tools

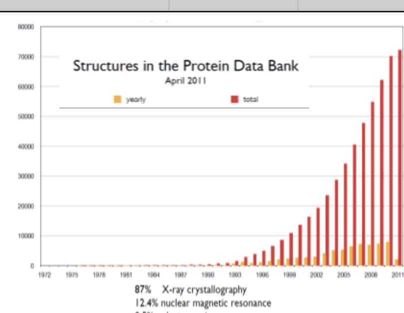
---

3. **X-Ray Crystallography**
   - **a) 100 years of X-ray Crystallography**
   - **b) Crystal Growth – Materials / Methods**
   - **c) Crystal Lattices - Lattice Constants / Space Groups / Asymmetric Unit**
   - ➡ **d) X-ray Sources – Sealed Tube / Rotation Anode / Synchrotron**
   - ➡ **e) Data Collection – Methods / Detectors / Structure Factors**
   - ➡ **f) Theory of Diffraction – Bragg's Law / Reciprocal Space**
   - ➡ **g) Structure Solution – Phase Problem: MIR / MR / MAD**
   - ➡ **h) Refinements and Models / Analysis and presentation of results**

A Half Century of Advances in Small Molecule Crystallography

Decade	1950's	1960's	1970's	1980's	1990's	2000's
Size	30 atoms	<100 non-H atoms	<200 non-H atoms	<400 non-H atoms	<600 non-H atoms	<1000 non-H atoms
Time / Structure	~ 1year	3-6 months	1 month	1 week	Few days	Few hours
# Structures	<500	~8300	32,000	95,000	229,000	528,000





What happened? **1913 – 1963 vs. 1963 - 2013**

1) **Sources**

2) **Instrumentation** -
   diffractometers / computers / **detectors**

3) Software / Computers -
   FORTRAN programming → SHELX, CCP4, Phenix, etc.

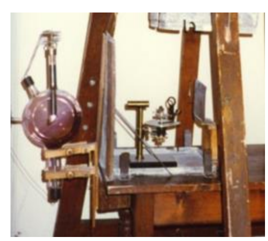
4) Molecular Biology - cloning / expression systems; sequencing

5) Automation - robotics

6) Methods -
   Phasing: Patterson / Heavy atom; MIR; SIR; MR; MAD; SAD
   Model Building: Contour tracing, Richards Box, FRODO, COOT
   Refinement and Validation

X-ray Sources:

X-ray tubes: the "sealed" tube



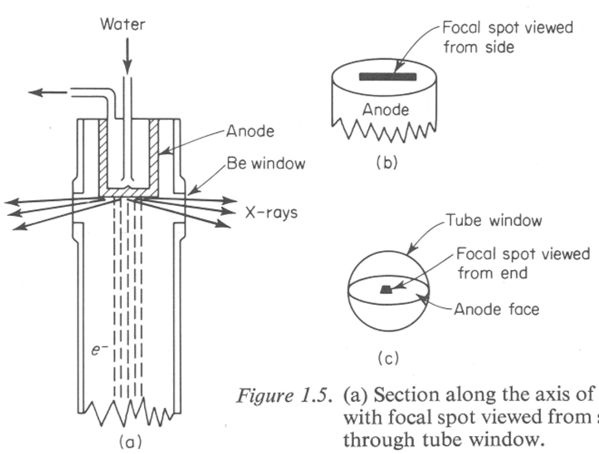


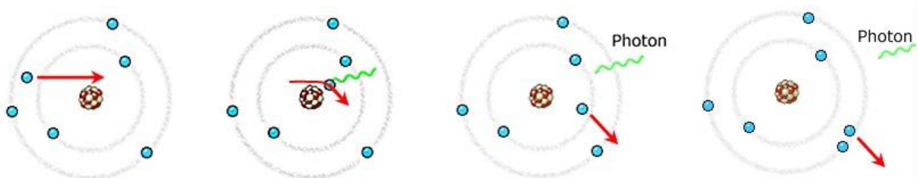
Figure 1.5. (a) Section along the axis of an X-ray tube. (b, Anode with focal spot viewed from side. (c) Focal spot viewed through tube window.

Origin of Non-characteristic X-rays

The electron (much lighter than the nucleus) comes very close to the nucleus and the electromagnetic interaction causes a deviation of the trajectory where the electron looses energy and an X-ray photon is emitted.

Bremsstrahlung X-rays

In an X-ray tube the electrons emitted from the anode are accelerated towards the metal target cathode by an accelerating voltage of typically 50 kV. The high energy electrons interact with the atoms in the metal target. Sometimes the electron comes very close to a nucleus in the target and is deviated by the electromagnetic interaction. In this process, which is called bremsstrahlung (braking radiation), the electron looses much energy and a photon (X-ray) is emitted. The energy of the emitted photon can take any value up to a maximum corresponding to the energy of the incident electron.



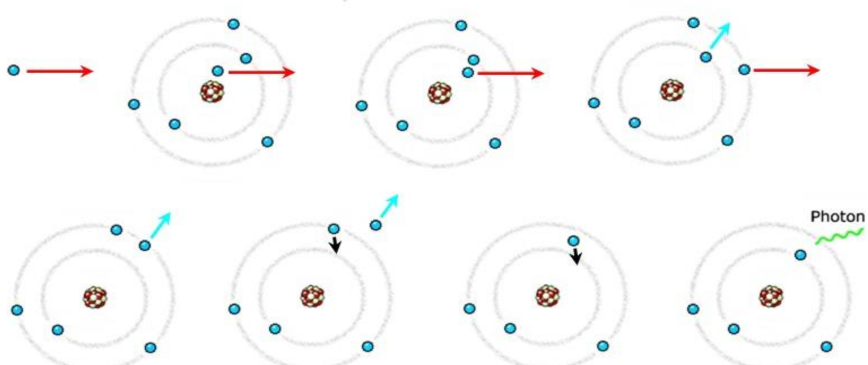
Origin of characteristic X-rays

Related Laureate

The Nobel Prize in Physics 1917 - Charles Glover Barkla »

Characteristic X-ray Lines

The high energy electron can also cause an electron close to the nucleus in a metal atom to be knocked out from its place. This vacancy is filled by an electron further out from the nucleus. The well defined difference in binding energy, characteristic of the material, is emitted as a monoenergetic photon. When detected this X-ray photon gives rise to a characteristic X-ray line in the energy spectrum. C. Barkla observed these lines in 1908-09 and was given the 1917 Nobel Prize for this discovery. He also made the first experiments suggesting that the X-rays are electromagnetic waves.



Characteristic X-rays arise from electronic transitions

$K_{\alpha_1}, K_{\alpha_2}$ $L \rightarrow K$

K_{β_1}, K_{β_2} $M \rightarrow K$

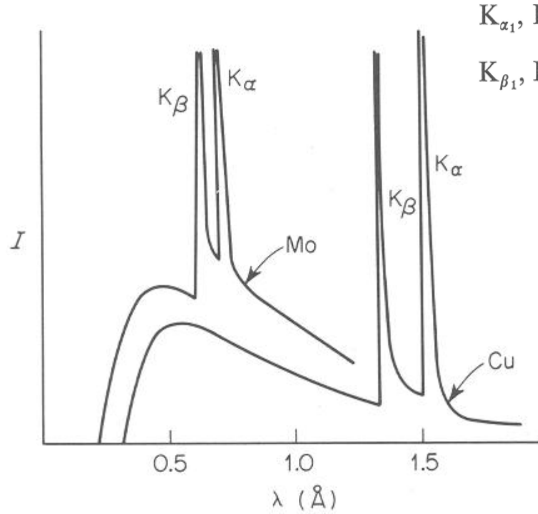
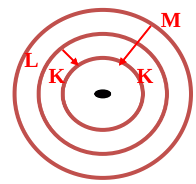


Figure 1.2. X-ray spectra with characteristic peaks: MoKα, 50 Kv; CuKα, 35 Kv.



Characteristic X-rays have defined λ

Table 1.1. Target Materials and Associated Constants

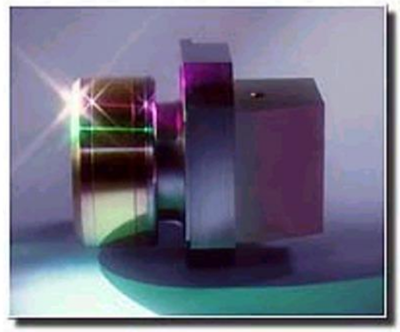
	Cr	Fe	Cu	Mo
Z	24	26	29	42
α_1, Å	2.2896	1.9360	1.5405	0.70926
α_2, Å	2.2935	1.9399	1.5443	0.71354
$\bar{\alpha}$,* Å	2.2909	1.9373	1.5418	0.71069
β_1, Å	2.0848	1.7565	1.3922	0.63225
β, filt.	V, 0.4 mil†	Mn, 0.4 mil	Ni, 0.6 mil	Nb, 3 mils
α, filt.	Ti	Cr	Co	Y
Resolution, Å	1.15	0.95	0.75	0.35
Critical potential, kV	5.99	7.11	8.98	20.0
Operating conditions, kV:	30–40	35–45	35–45	50–55
half- or full-wave-rectified, mA	10	10	20	20
constant potential, mA	7	7	14	14

* $\bar{\alpha}$ is the intensity-weighted average of α_1 and α_2 and is the figure usually used for the wavelength when the two lines are not resolved.

† 1 mil = 0.001 inch = 0.025 mm.

X-ray Generators

FR591



FR591 Rotating Anode X-ray Generator

The Nonius' FR591 rotating anode X-ray generator now has dramatically improved the performance of the anode, by a complete redesign.
We now have a static shaft and a rotating anode, instead of rotating both. The cooling water flow has also been redesigned to give much higher throughput, higher flow and higher turbulence, which results in better heat transfer and hence better cooling capacity.

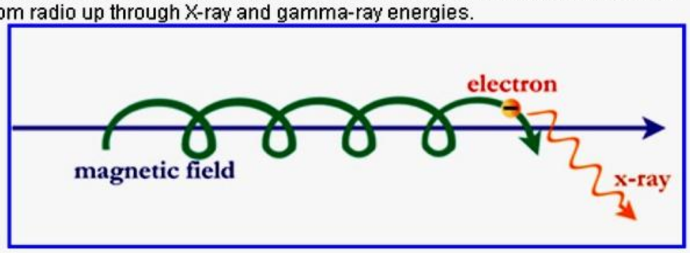
Now with the new ULTRA anode you can get 6 kW on a 0.3mm focus!

Another Source of "X-rays"

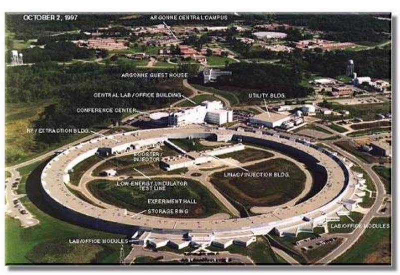
Synchrotron Radiation

X-ray photons can also be created under different conditions. When physicists were operating the first particle accelerators, they discovered that electrons can produce photons without colliding at all. This was possible because the magnetic field in the accelerators was causing the electrons to move in large spirals around magnetic field lines of force. This process is called synchrotron radiation.

In the cosmos particles such as electrons can be accelerated to high energies– near the speed of light– by electric and magnetic fields. These high-energy particles can produce synchrotron photons with wavelengths ranging from radio up through X-ray and gamma-ray energies.

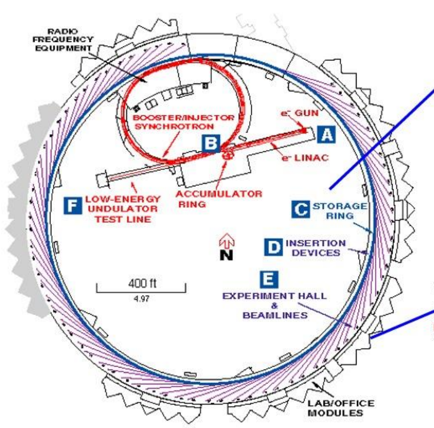


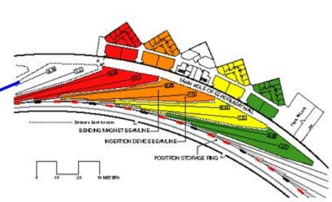
Synchrotron Radiation: Electrons moving in magnetic field radiate photons.



APS - Advanced Photon Source

Argonne National Laboratory





"X-ray" Sources: X-ray tubes

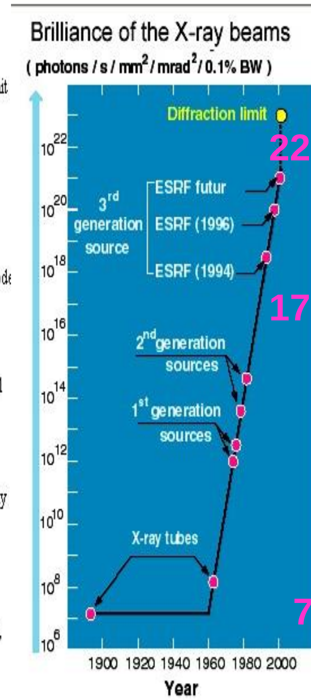
The **brilliance** of a light source is defined as the number of photons emitted per second, per unit source size, per unit space angle and for a bandwidth of 1/1000 of the photon energy

The Comparison between various sources of X-rays shows large differences in their brilliance.

X-ray tubes:
Wilhelm Conrad Röntgen discovered X-rays in 1895 whilst working with cathode-ray tubes. Using the principle of fast electrons hitting a metallic target, a first substantial gain in brilliance was not obtained until the introduction of rotating anode sources (~1960).

Synchrotron Radiation Facilities:
The progress of high energy physics, with the construction of powerful particle accelerators gave birth to what we now call *First* generation synchrotron sources (~1970). Using the deflection of high energy electrons by a magnetic field for the production of X-rays proved so promising that a number of dedicated *Second* generation sources were built (~1980). Relying on the combination of needle thin electron beams and Insertion Devices, *Third* generation synchrotron sources (~1995) are now emitting synchrotron X-ray beams that are a trillion (10^{12}) times more brilliant than those produced by X-ray tubes.

Free Electron X-ray Lasers:
Coupling electron and X-ray beams together, the Free Electron X-ray Lasers currently on the drawing boards could be the next generation of X-ray sources. While they promise to achieve an increase in peak brilliance by another factor of a trillion, the first prototypes may be operational around the year 2010.

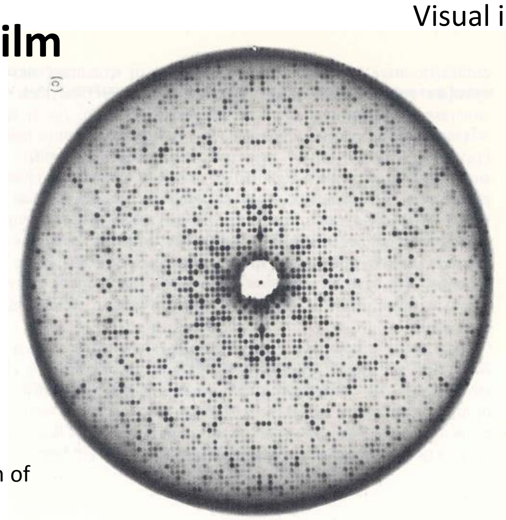


Advanced Photon Source (APS) synchrotron near Chicago.

Detectors

Visual intensity estimation

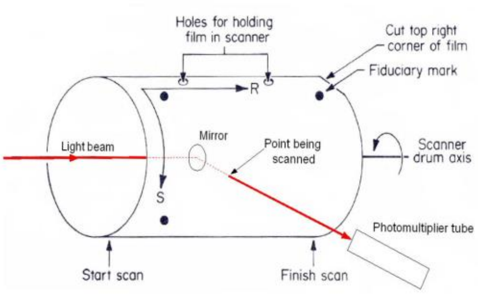
Film



Precession film of
Phosphorylase
From Blundell and Johnson p.156

Ron Hamlin – Supper Award talk

Concept drawing of a film scanner



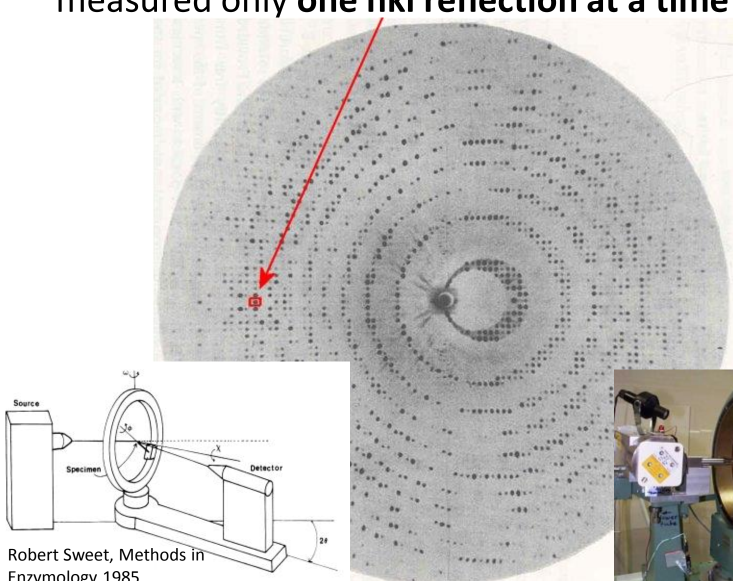
Rossmann, Methods in Enzymology 1985, p. 242

Optronics P-1000 film scanner

Ron Hamlin – Supper Award talk

Photo provided by Dieter Schneider, Brookhaven National Laboratory

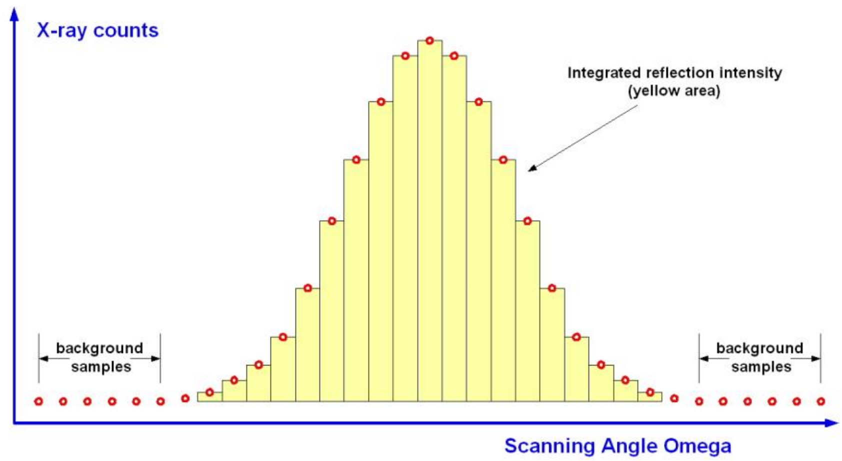
Diffractometer: automatic but measured only **one hkl reflection at a time**



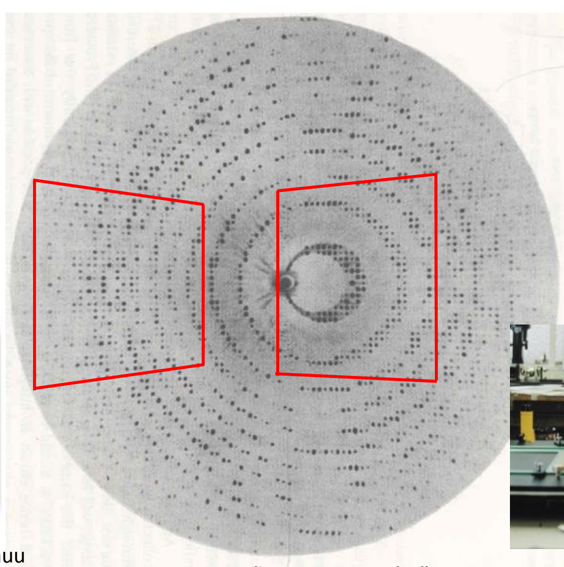
Robert Sweet, Methods in Enzymology 1985

Ron Hamlin – Supper Award talk

Omega scan of a single reflection

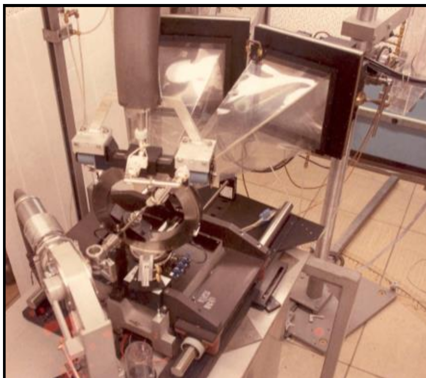


Area Detectors: Typical coverage of diffraction pattern by a pair of ADSC multiwire detectors



Xuong Nguyen-huu

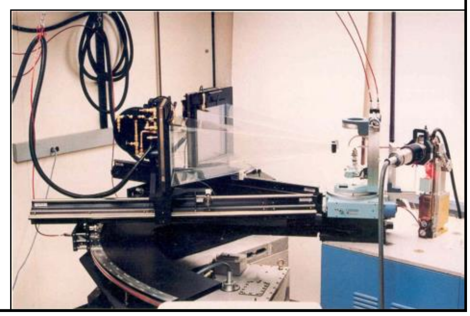
Ron Hamlin



Era of Multiwire Area Detectors

Parts of first ADSC multiwire counter system in Ron's living room in early 1984

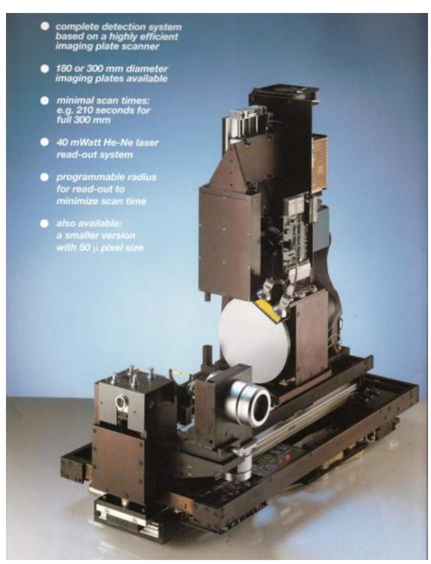
The two-detector Mark II system started operation in Xuong's lab in about 1982



Motor driven two theta table
U. Texas Austin 1988

Ron Hamlin – Supper Award talk

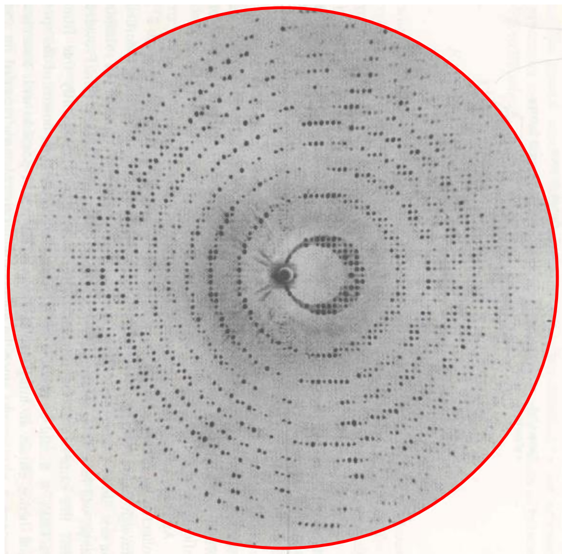
MAR 180 with cover removed



Ron Hamlin – Supper Award talk

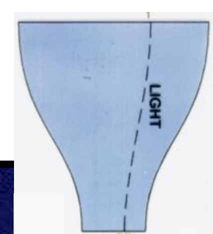
Image Plate Detectors

brute force solid angle coverage



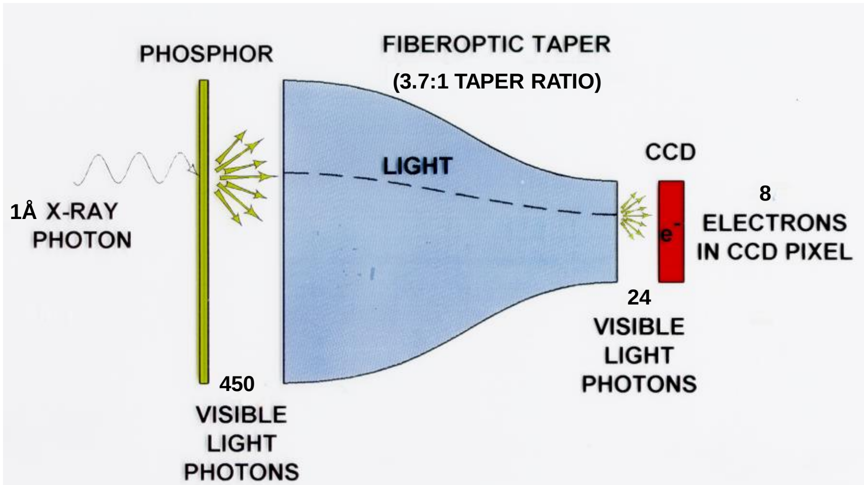
Ron Hamlin – Supper Award talk

Fiberoptic Tapers



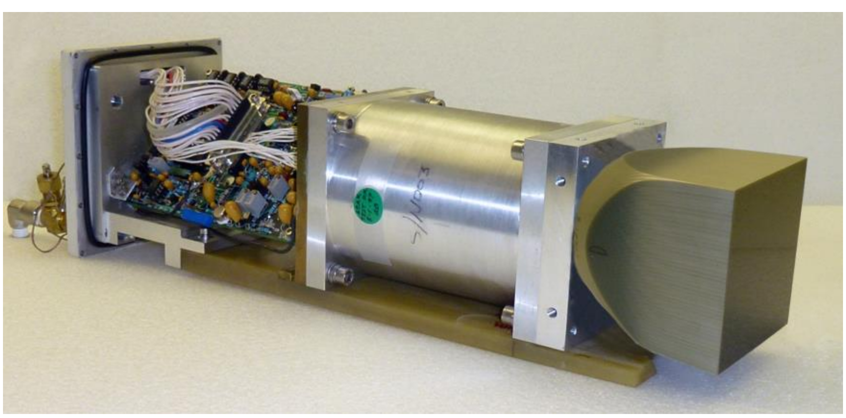
Ron Hamlin – Supper Award talk

Basic Principle of Operation



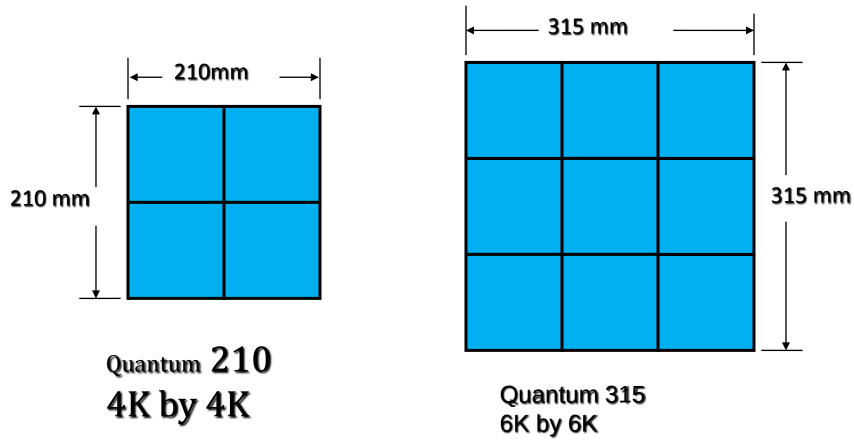
Ron Hamlin – Supper Award talk

Quantum 1 cover removed

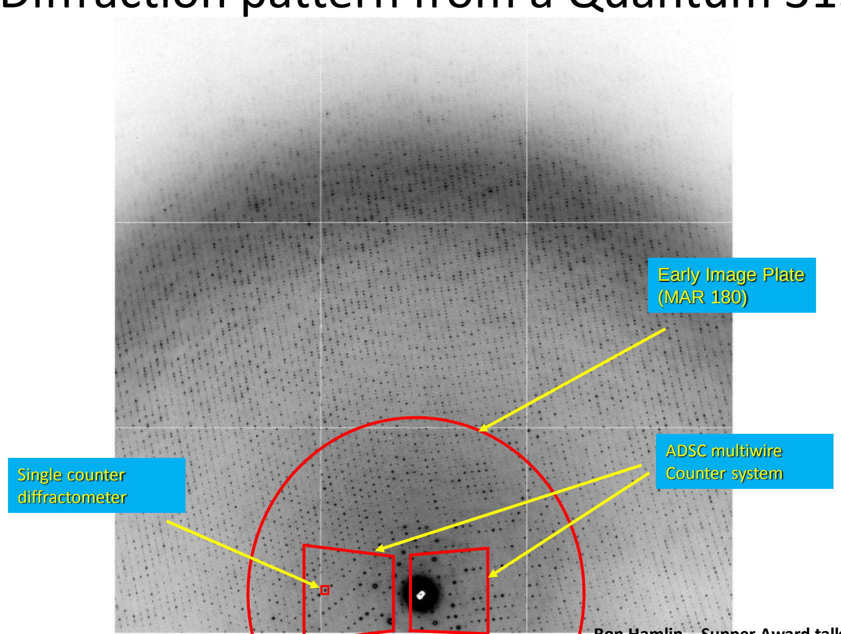


Ron Hamlin – Supper Award talk

The Quantum 315 uses 9 instead of 4 of exactly the same modules as used in the Quantum 210.



Diffraction pattern from a Quantum 315



X-Ray Crystallography

"If a picture is worth a thousand words, then a macromolecular structure is priceless to a physical biochemist." – van Holde

Topics:

1. **Image Formation** *(optical illusions)*

   Resolution / Wavelength (Amplitude, Phase) / Light Microscopy / EM / X-ray / (NMR)

2. **Protein Data Bank (PDB)**

   Data mining and Protein Structure Analysis Tools

---

3. **X-Ray Crystallography**

   a) 100 years of X-ray Crystallography

   b) Crystal Growth – Materials / Methods

   c) Crystal Lattices - Lattice Constants / Space Groups / Asymmetric Unit

   d) X-ray Sources – Sealed Tube / Rotation Anode / Synchrotron

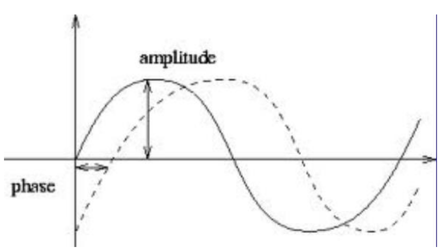
   e) Data Collection – Methods / Detectors / Structure Factors

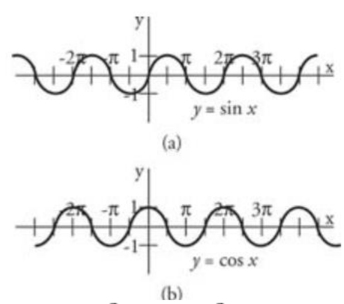
   → **f) Theory of Diffraction – Waves, Fourier / Bragg's Law / Reciprocal Space**

   → **g) Structure Solution – Phase Problem: MIR / MR / MAD**

   → **h) Refinements and Models / Analysis and presentation of results**

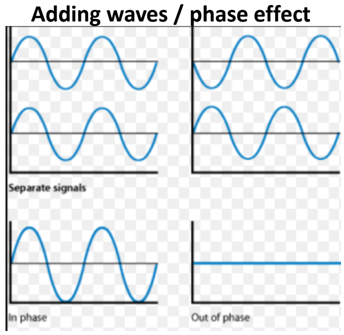
Waves (Amplitudes & Phases)

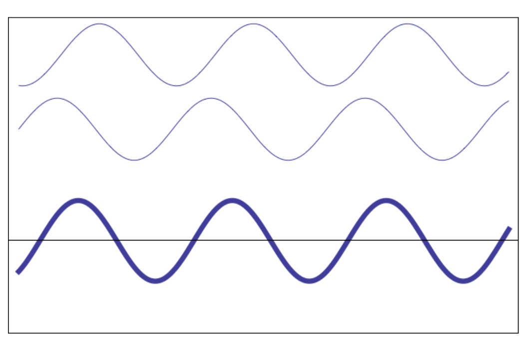


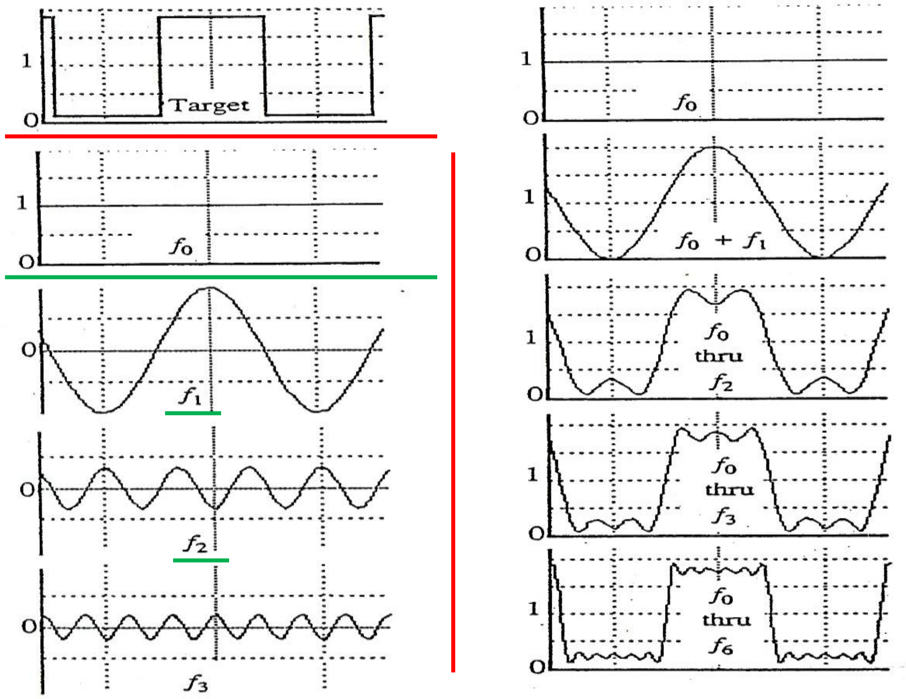


$$\sin^2\theta + \cos^2\theta = 1$$

Adding waves / phase effect







Joseph Fourier / Fourier Series ~1808

Fourier series are named in honor of Joseph Fourier (1768-1830), who made important contributions to the study of trigonometric series, after preliminary investigations by Euler, d'Alembert, and Bernoulli. He applied this technique to find the solution of the heat equation, publishing his initial results in 1807, and publishing his Théorie analytique de la chaleur in 1822

$$f(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos\frac{n\pi t}{L} + \sum_{n=1}^{\infty} b_n \sin\frac{n\pi t}{L}$$

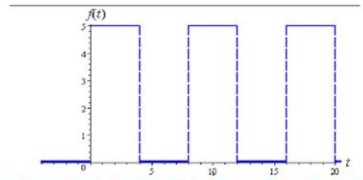
$$a_0 = \frac{1}{L}\int_{-L}^{L} f(t)\,dt$$

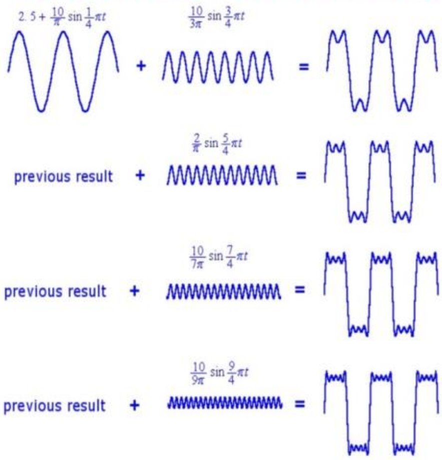
$$a_n = \frac{1}{L}\int_{-L}^{L} f(t)\cos\frac{n\pi t}{L}\,dt \quad b_n = \frac{1}{L}\int_{-L}^{L} f(t)\sin\frac{n\pi t}{L}\,dt$$

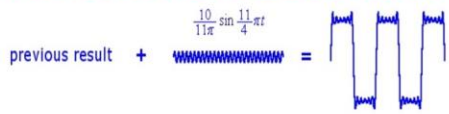
where $n = 1, 2, 3 \ldots$

Fourier Series

Example - Square Wave

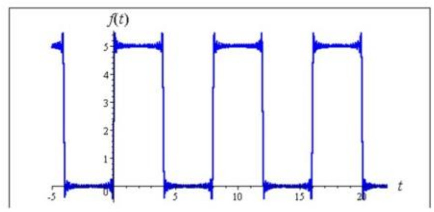






If we graph many terms, we see that our series is producing the required function. We graph the first 20 terms:

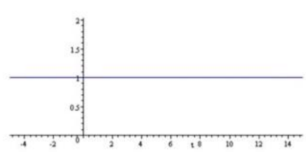
$$2.5 + \frac{10}{\pi}\sum_{n=1}^{20}\frac{1}{(2n-1)}\sin\frac{(2n-1)\pi t}{4}$$



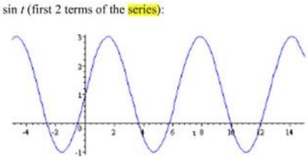
Example - Saw Tooth Function

$f(t) = 1 + 2\sin t - \sin 2t + \frac{2}{3}\sin 3t$

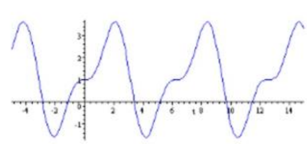
$f(t) = 1$ (first term of the series):

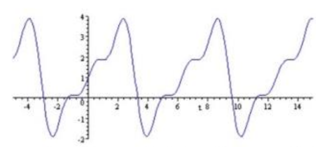


$f(t) = 1 + 2\sin t$ (first 2 terms of the series):

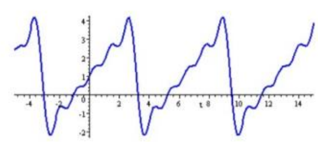


$f(t) = 1 + 2\sin t - \sin 2t$ (first 3 terms of the series):



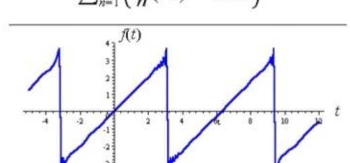


$f(x) = 1 + 2\sin t - \sin 2t + \frac{2}{3}\sin 3t - \frac{1}{2}\sin 4t + \frac{2}{5}\sin 5t + \ldots$



The graph of the first 40 terms is:

$$\sum_{n=1}^{40}\left(\frac{2}{n}(-1)^{n+1}\sin nt\right)$$



Fourier Series Applet http://www.falstad.com/fourier/

Sines / Cosines / and Expoentials

$$\exp(x) \equiv e^x \equiv \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

$$= 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \frac{x^4}{24} + \frac{x^5}{120} + \ldots$$

If we let x be *imaginary*, $x = i\theta$ (where θ is real), then this can be written

$$e^{i\theta} = 1 + i\theta - \frac{\theta^2}{2} - i\frac{\theta^3}{6} + \frac{\theta^4}{24} + i\frac{\theta^5}{120} - \ldots$$

recall

$$\cos\theta = 1 - \frac{\theta^2}{2} + \frac{\theta^4}{24} - \ldots$$

&

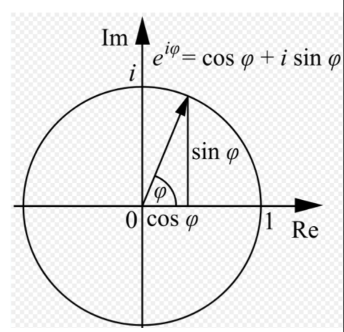
$$\sin\theta = \theta - \frac{\theta^3}{6} + \frac{\theta^5}{120} - \ldots$$

thus

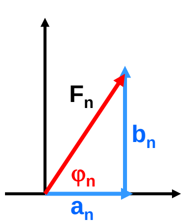
$$\boxed{e^{i\theta} = \cos\theta + i\,\sin\theta}$$

Euler's Equation

Euler's formula (Leonhard Euler, 1707-1783) gives the relationship between the complex exponential function and common trig terms. For any real number "φ"



Replace list of a_n / b_n with $|F_n|$ and φ_n



n	a	b	F	φ
1	7	0	7	0
2	0	8	8	90
3	5	5	7.1	??
4	8	6	??	36.9

Fourier Series / Fourier Transforms

$$f(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos\frac{n\pi t}{L} + \sum_{n=1}^{\infty} b_n \sin\frac{n\pi t}{L}$$

or

$$a_0 = \frac{1}{L}\int_{-L}^{L} f(t)\,dt$$

$$a_n = \frac{1}{L}\int_{-L}^{L} f(t)\cos\frac{n\pi t}{L}\,dt \quad b_n = \frac{1}{L}\int_{-L}^{L} f(t)\sin\frac{n\pi t}{L}\,dt$$

Now consider **electron density (as a function or a set of coefficients)**

$$\rho(\mathbf{x}) = \frac{1}{V}\sum_{\mathbf{h}} \mathbf{F}(\mathbf{h})\exp(-2\pi i\mathbf{h}\cdot\mathbf{x})$$

Real Space

or

$$\mathbf{F}(\mathbf{h}) = \int_{cell} \rho(\mathbf{x})\exp(2\pi i\mathbf{h}\cdot\mathbf{x})\,d\mathbf{v}$$

Reciprocal Space

AND – F_{hkl} can also be **calculated** as the resultant scattering or the **sum of the individual scattering atoms**!!

$$\mathbf{F}(hkl) = F(hkl)e^{i\alpha(hkl)} = \sum_{j=1}^{N'} \mathbf{f}_j(hkl) = \sum_{j=1}^{N'} f_j(hkl)e^{i\phi_j(hkl)}$$

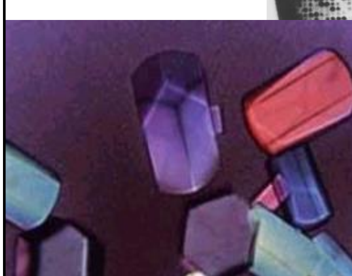
Object → **Transform** → **Image**

Transform / Reciprocal Space

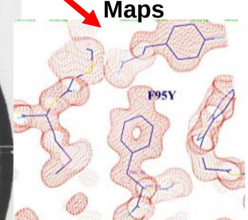
X-rays

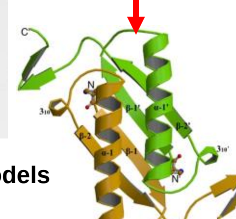
+

Object / Real Space



Electron Density Maps





Models

Diffraction: Scattering from (two) "atoms"

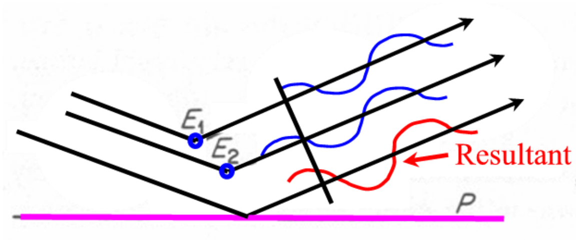
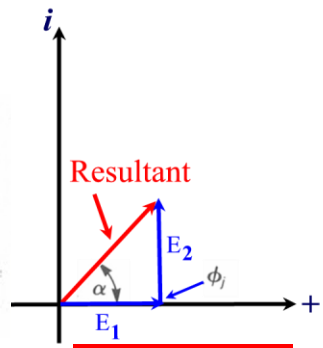


Figure 2.10. Diffraction from E_1 and E_2 as if reflected from plane *P*.



Represent a wave as an amplitude + phase.

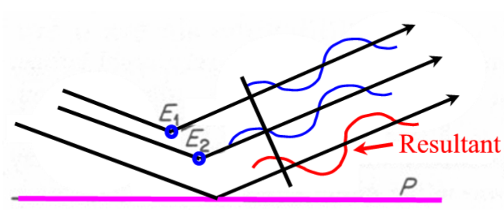
Scattering from "many atoms"

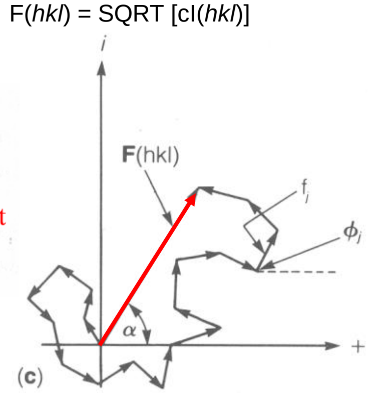
$$\mathbf{F}(hkl) = F(hkl)e^{i\alpha(hkl)} = \sum_{j=1}^{N'} \mathbf{f}_j(hkl) = \sum_{j=1}^{N'} f_j(hkl)e^{i\phi_j(hkl)}$$

Calculated

Experimental

F(*hkl*) = SQRT [cI(*hkl*)]

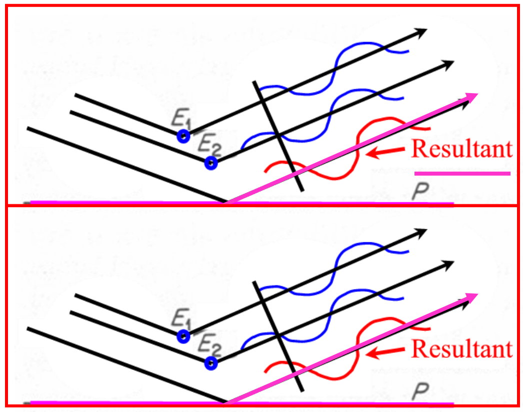




The structure factor for a reflection may be thought of as the vector sum of the x-ray scattering contributions from many atoms.

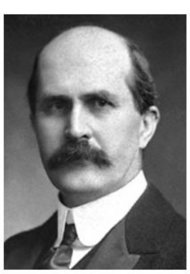
Each of the j contributions may be represented as a vector in the complex plane, with amplitude f_j and phase phi_j.

Scattering from "atoms in two unit cells"



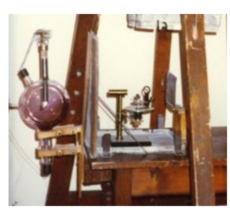
The Nobel Prize in Physics 1915
William Bragg, Lawrence Bragg

The Nobel Prize in Physics 1915

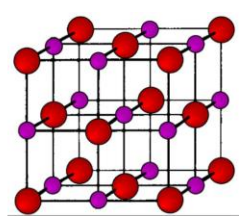


Sir William Henry Bragg

William Lawrence Bragg

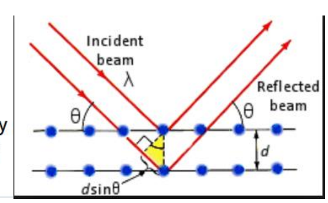


X-ray apparatus



Sodium Chloride (NaCl)

The Nobel Prize in Physics 1915 was awarded jointly to Sir William Henry Bragg and William Lawrence Bragg *"for their services in the analysis of crystal structure by means of X-rays"*

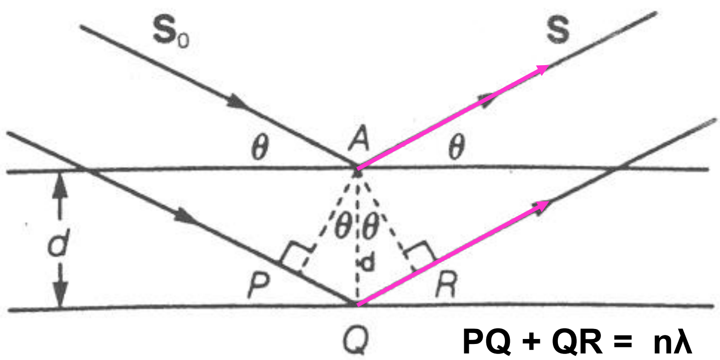


Bragg's Law ($n\lambda = 2d \sin\theta$)

Crystals: Scattering from "planes"

Resultant scattering of resultant scattering!

Bragg Equation $n\lambda = 2\, d \sin(\theta)$

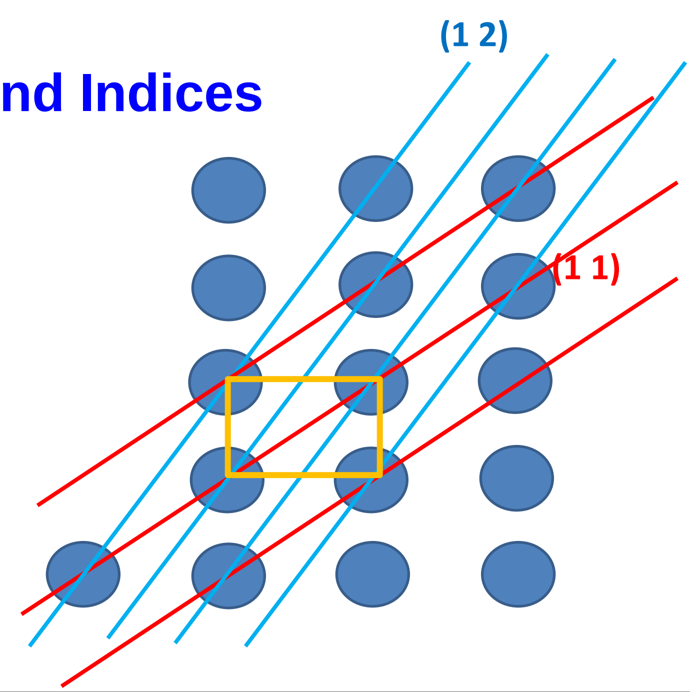


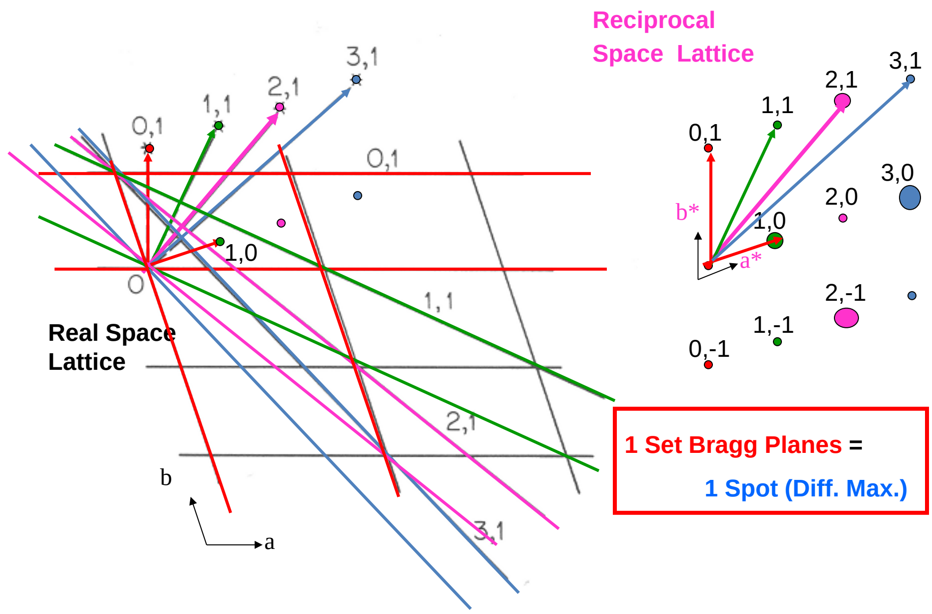
$PQ + QR = n\lambda$

→ **Scattering** will only be **"observed"** at discrete **Bragg angles(θ)**

The spacings of the Bragg reflections → Lattice Constants

Planes and Indices





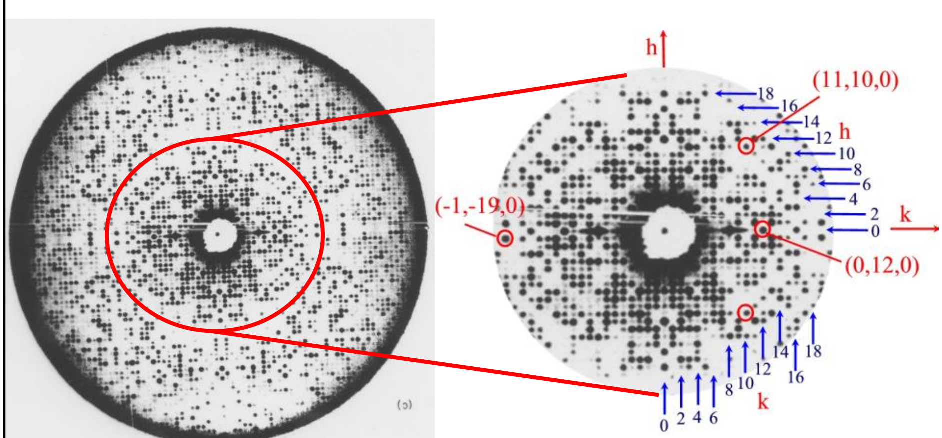
1 Set Bragg Planes =

1 Spot (Diff. Max.)

Planes in direct space represented by points in reciprocal space.

Electron Density Function

$$\rho(X,Y,Z) = \frac{1}{V}\sum_h\sum_k\sum_l \underline{F(hkl)}\ \underline{\exp[i\alpha(hkl)]}\exp[-2\pi i(hX + kY + lZ)]$$

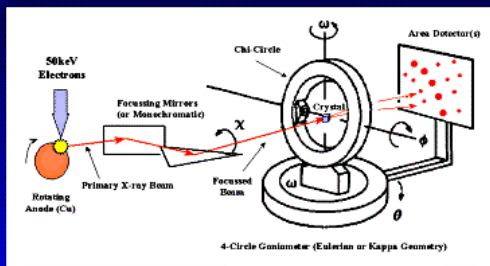


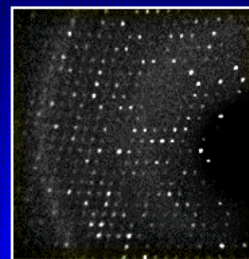
Measure thousands of **Amplitudes** - [Fhkl]'s - ?? How do we obtain **Phases** αhkl ??

→ **Phase Problem**

Advanced Methods in Modern Biomolecular Crystallography

The information we get from a single diffraction experiment......





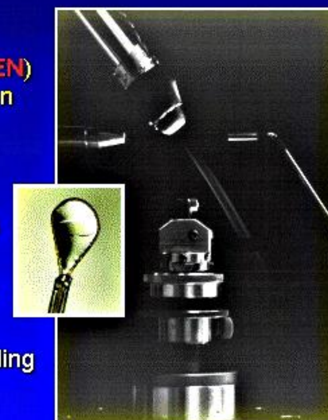
The reflections are indexed (consistent assignment of reciprocal cell indices h,k,l) and all we get for the money is a long list of intensities from several ten thousand reflections

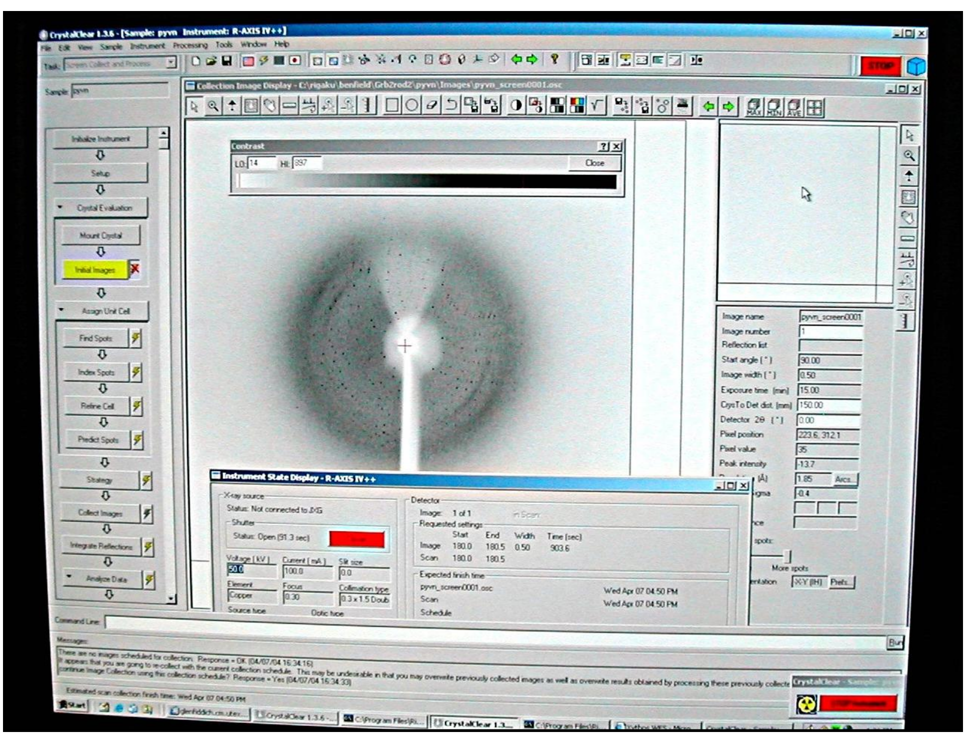
2	10	1	326.58
3	10	1	1644.72
4	10	1	3228.45
5	10	1	1279.83
6	10	1	320.48
7	10	1	775.63
8	10	1	1344.55
9	10	1	431.73
10	10	1	1760.14
11	10	1	709.18
12	10	1	20.37
13	10	1	408.72
14	10	1	51.36
15	10	1	114.72
16	10	1	776.26
17	10	1	87.57
18	10	1	30.93
0	11	1	99.30
1	11	1	2258.68
2	11	1	770.18

Advanced Methods in Modern Biomolecular Crystallography

Cryo-cooling efficiently improves data quality

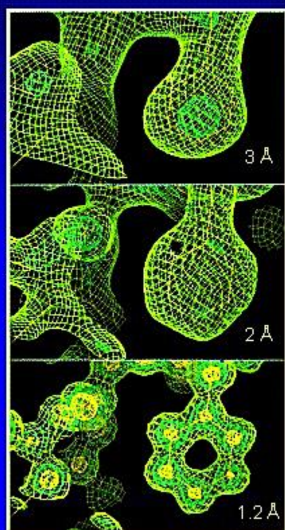
- Crystals are rapidly cooled (NOT FROZEN) to near liquid nitrogen temperature
- Reduced thermal vibrations
- Increased resolution
- Reduced disorder
- Eliminated radiation damage
- No merging and scaling errors

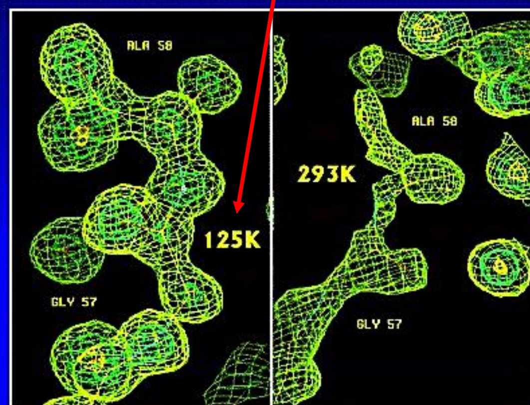




Effect of Resolution

Reduced Disorder at Lower Temperatures





Dramatic improvements in the overall structure are likely to result from better definition of disordered regions regardless of resolution

X-Ray Crystallography

"If a picture is worth a thousand words, then a macromolecular structure is priceless to a physical biochemist." – van Holde

Topics:

1. Image Formation *(optical illusions)*

Resolution / Wavelength (Amplitude, Phase) / Light Microscopy / EM / X-ray / (NMR)

2. Protein Data Bank (PDB)

Data mining and Protein Structure Analysis Tools

---

3. X-Ray Crystallography

a) 100 years of X-ray Crystallography

b) Crystal Growth – Materials / Methods

c) Crystal Lattices - Lattice Constants / Space Groups / Asymmetric Unit

d) X-ray Sources – Sealed Tube / Rotation Anode / Synchrotron

e) Data Collection – Methods / Detectors / Structure Factors

f) Theory of Diffraction – Waves, Fourier / Bragg's Law / Reciprocal Space

➡ **g) Structure Solution – Phase Problem: MIR / MR / MAD**

➡ **h) Refinements and Models / Analysis and presentation of results**

Solving the Phase Problem

Early Days:

Centric structures (all phases 0^o or 180^o) Heavy atom / Patterson method

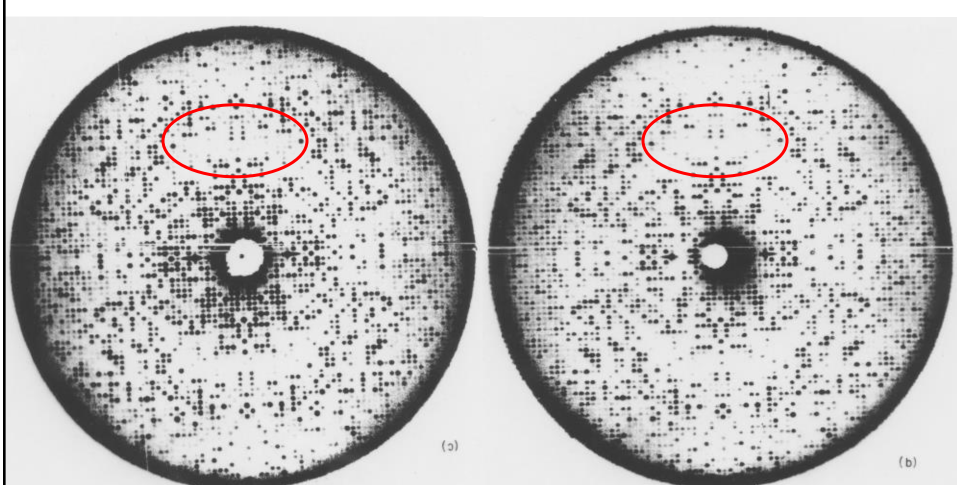
Macromolecular Crystallography

1. **MIR:** **Multiple Isomorphous Replacement (Heavy Atom)**
2. **MR:** **Molecular Replacement**
3. **MAD:** **multiwavelength anomolous dispersion**

****

- **Molecular Modeling (predicting starting structure from sequence alone)**

Use of Heavy Metal Ions for Phasing by MIR Methods

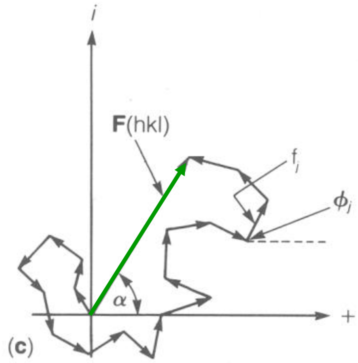


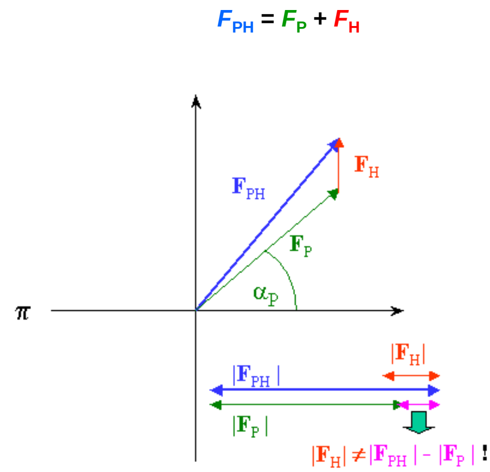
Native Phosphorylase

Phosphorylase + Ethyl Hg thiosalicylate

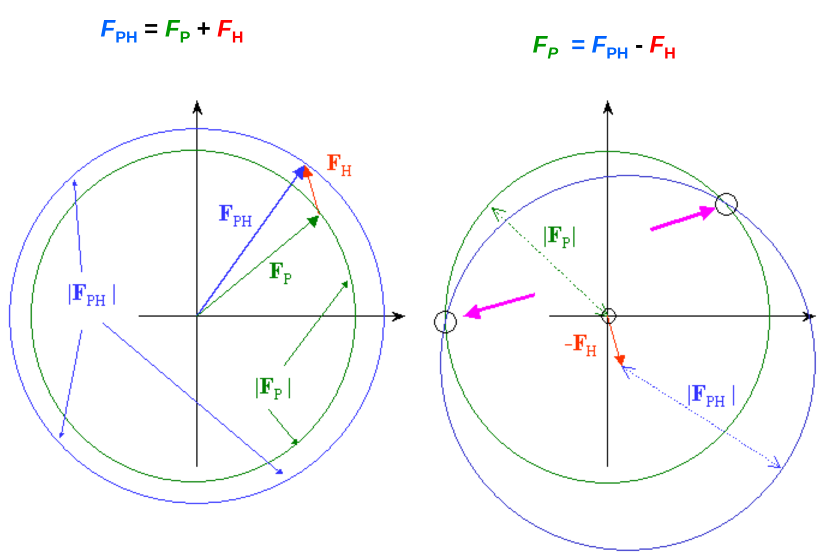
Effect of adding 1 "heavy" atom with lots of electrons!

(e.g. Hg, Pt)



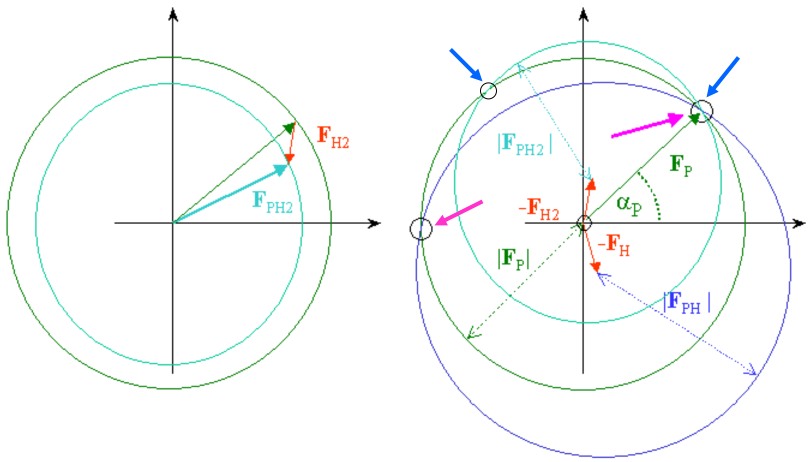


Multiple Isomorphous Replacement (MIR) method



Multiple Isomorphous Replacement (MIR) method

$F_P = F_{PH} - F_H$



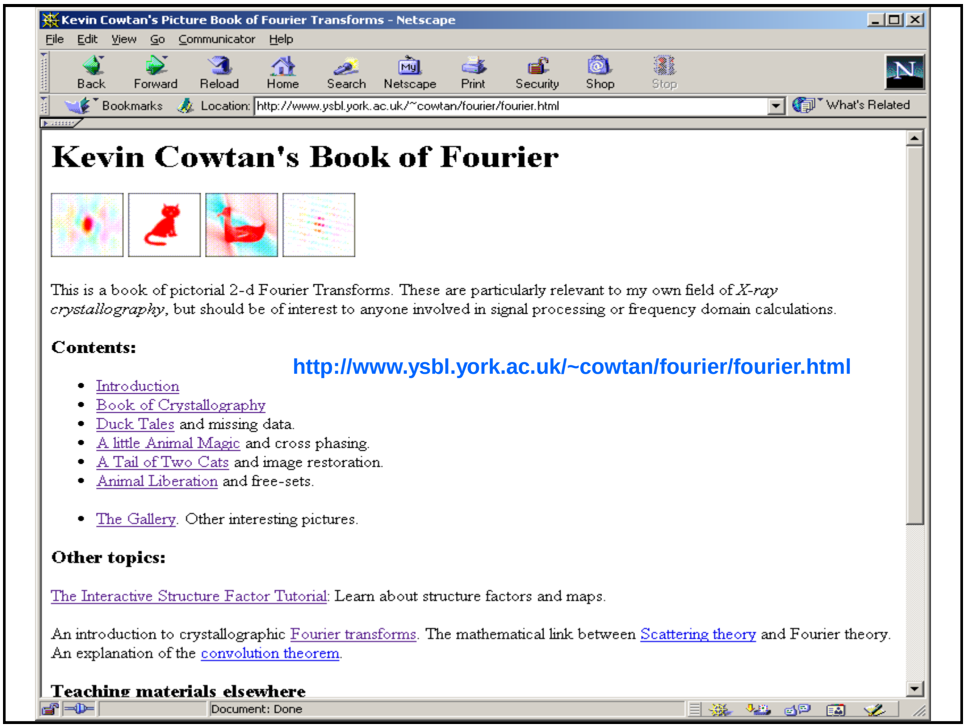
Solving the phase problem by "**Molecular Replacement**".

If an approximate model of the protein structure is known in advance, approximate phases can be guessed, and the unknown parts of the structure can be calculated in an iterative procedure.

No heavy atom derivative required.

BUT – need starting model and orientation (rotation and translation)

For example, molecular replacement can be used to determine the structure of an complex with inhibitor bound to an enzyme active site, if the structure of the enzyme itself is already known. Also, MR is often used to solve the structures of closely related proteins in a superfamily.



Kevin Cowtan's Picture Book of Fourier Transforms - Netscape

File Edit View Go Communicator Help

Back Forward Reload Home Search Netscape Print Security Shop Stop

Bookmarks Location: http://www.ysbl.york.ac.uk/~cowtan/fourier/fourier.html What's Related

Kevin Cowtan's Book of Fourier

This is a book of pictorial 2-d Fourier Transforms. These are particularly relevant to my own field of *X-ray crystallography*, but should be of interest to anyone involved in signal processing or frequency domain calculations.

Contents:

http://www.ysbl.york.ac.uk/~cowtan/fourier/fourier.html

- Introduction
- Book of Crystallography
- Duck Tales and missing data.
- A little Animal Magic and cross phasing.
- A Tail of Two Cats and image restoration.
- Animal Liberation and free-sets.

- The Gallery. Other interesting pictures.

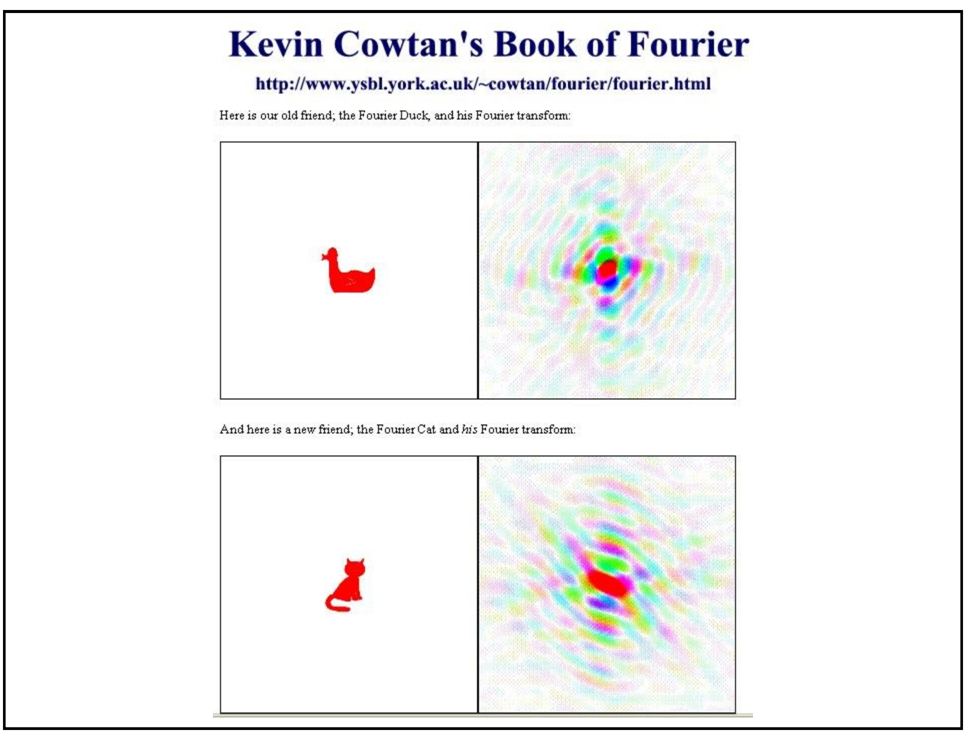
Other topics:

The Interactive Structure Factor Tutorial: Learn about structure factors and maps.

An introduction to crystallographic Fourier transforms. The mathematical link between Scattering theory and Fourier theory. An explanation of the convolution theorem.

Teaching materials elsewhere

Document: Done



Kevin Cowtan's Book of Fourier

http://www.ysbl.york.ac.uk/~cowtan/fourier/fourier.html

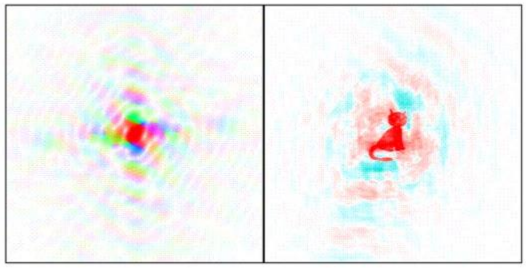
Here is our old friend; the Fourier Duck, and his Fourier transform:

And here is a new friend; the Fourier Cat and *his* Fourier transform:

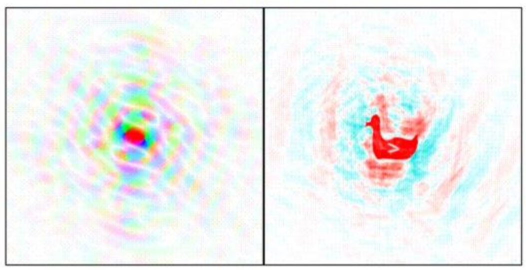
Kevin Cowtan's Book of Fourier

http://www.ysbl.york.ac.uk/~cowtan/fourier/fourier.html

Duck Transform Amplitudes + Cat Phases



Cat Transform Amplitudes + Duck Phases



X-Ray Crystallography

"If a picture is worth a thousand words, then a macromolecular structure is priceless to a physical biochemist." – van Holde

Topics:

1. **Image Formation** *(optical illusions)*

   Resolution / Wavelength (Amplitude, Phase) / Light Microscopy / EM / X-ray / (NMR)

2. **Protein Data Bank (PDB)**

   Data mining and Protein Structure Analysis Tools

---

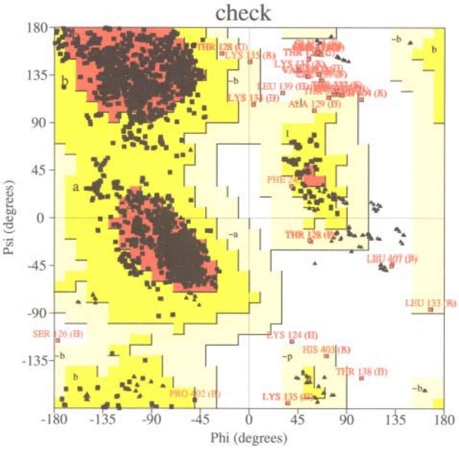
3. X-Ray Crystallography

- **a) 100 years of X-ray Crystallography**
- **b) Crystal Growth – Materials / Methods**
- **c) Crystal Lattices - Lattice Constants / Space Groups / Asymmetric Unit**
- **d) X-ray Sources – Sealed Tube / Rotation Anode / Synchrotron**
- **e) Data Collection – Methods / Detectors / Structure Factors**
- **f) Theory of Diffraction – Waves, Fourier / Bragg's Law / Reciprocal Space**
- **g) Structure Solution – Phase Problem: MIR / MR / MAD**
- ➡ **h) Refinements and Models / Analysis and presentation of results**

PROCHECK

Ramachandran Plot

check



Plot statistics

Residues in most favoured regions [A,B,L]	3149	85.9%
Residues in additional allowed regions [a,b,l,p]	460	12.6%
Residues in generously allowed regions [~a,~b,~l,~p]	24	0.7%
Residues in disallowed regions	31	0.8%
Number of non-glycine and non-proline residues	3664	100.0%
Number of end-residues (excl. Gly and Pro)	228	
Number of glycine residues (shown as triangles)	312	
Number of proline residues	168	
Total number of residues	4372	

Based on an analysis of 118 structures of resolution of at least 2.0 Angstroms and R-factor no greater than 20%, a good quality model would be expected to have over 90% in the most favoured regions.

check_01.ps

Crystal Structure of *M. tuberculosis* Alanine Racemase

Table 1: Data Collection and Processing Statistics for the MAD and Native Data Sets of Alr_{Mtb}

	MAD 1	MAD 2	MAD 3	MAD 4	native
λ (Å)	0.9788	0.9790	0.9562	0.9809	0.9160
resolution (Å)		2.20			1.80
mosaicity		0.50			0.65
no. of reflections observed > 1σ	432376	446744	431524	336135	779600
no. of unique reflections > 1σ	35817	37506	36020	36242	67592
$R_{merge}{}^{a}$ (%)	6.9	6.4	5.1	3.7	6.0 (67.2)
completeness (%)	91.8	95.8	92.1	92.1	99.3 (95.6)
$\langle I/\sigma \rangle$	30.3	34.3	41.6	50.9	34.5 (2.6)

[a] $R_{merge} = \sum|I_{obs} - I_{avg}|/\sum|I_{avg}|$.

Biochemistry **2005**, *44*, 1471−1481 1471

The 1.9 Å Crystal Structure of Alanine Racemase from *Mycobacterium tuberculosis* Contains a Conserved Entryway into the Active Site[†,‡]

Pierre LeMagueres,[§] Hookang Im,[§] Jerry Ebalunode,[§] Ulrich Strych,[§] Michael J. Benedik,[∥] James M. Briggs,[§] Harold Kohn,[⊥] and Kurt L. Krause[*,§,@]

Department of Biology and Biochemistry, University of Houston, Houston, Texas 77204-5001, Department of Biology, Texas A&M University, College Station, Texas 77843-3258, Division of Medicinal Chemistry and Natural Products, School of Pharmacy, University of North Carolina, Chapel Hill, North Carolina 27599-7360, and Section of Infectious Diseases, Department of Medicine, Baylor College of Medicine, Houston, Texas 77030

Received June 27, 2004; Revised Manuscript Received October 22, 2004

Table 2: Final Refinement Statistics for Alr_{Mtb} at 1.9 Å Resolution

R factor[a] (%)	20.4
R_{free} (%) (for 1747 reflections)	25.4
average B factor (Å^2)[b]	
main chain	25.5
side chain	31.5
PLP	21.9
waters	32.4
rms deviations	
bond lengths (Å)	0.006
bond angles (deg)	1.9
no. of reflections > 2σ	55001
no. of residues	722
no. of protein atoms	5360
no. of PLP atoms	30
no. of water molecules	350

[a] R-factor $= \sum|F_{obs} - F_{calc}|/\sum|F_{obs}|$. [b] All isotropic model.

Energy Refinement

(Simulated Annealing)

$$E_{TOTAL} = E_{EMPIRICAL} + E_{EFFECTIVE}$$

$$E_{EFFECTIVE} = E_{XREF} + E_{NOE} + E_{HARM} + E_{CDIH} + E_{NCS} + E_{DG} + E_{RELA} + E_{PLAN}$$

$$E_{\mathrm{EMPIRICAL}} = \Sigma^{N}_{p=l} [w^{p}_{BOND}E_{BOND} + w^{p}_{ANGL}E_{ANGL} + w^{p}_{DIHE}E_{DIHE} + w^{p}_{IMPR}E_{IMPR} + w^{p}_{VDW}E_{VDW} + w^{p}_{\mathrm{ELEC}}E_{ELEC} + w^{p}_{\mathrm{PVDW}}E_{PVDW} + w^{p}_{PELE}E_{PELE} + w^{p}_{HBON}E_{HBON}].$$

Difference Fourier

Obs. $$\rho_o(x, y, z) = \frac{1}{V}\sum_h \sum_k \sum_l F_{o,hkl}\, e^{-2\pi i(hx+ky+lz)} + R$$

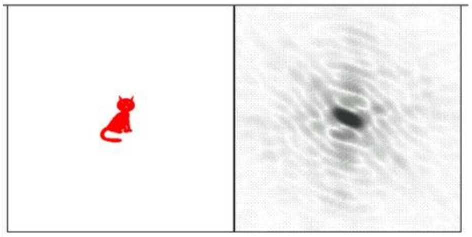
Calc. $$\rho_c(x, y, z) = \frac{1}{V}\sum_h \sum_k \sum_l F_{c,hkl}\, e^{-2\pi i(hx+ky+lz)} + R'$$

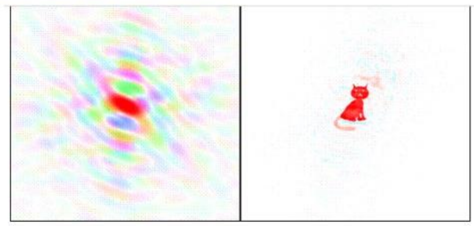
$$\rho_o(x, y, z) - \rho_c(x, y, z) = \frac{1}{V}\sum_h \sum_k \sum_l (F_o - F_c)_{hkl}\, e^{-2\pi i(hx+ky+lz)} + R - R'$$

$$\rho_o - \rho_c = \frac{1}{V}\sum_h \sum_k \sum_l \Delta F_{hkl}\, e^{-2\pi i(hx+ky+lz)}$$

Kevin Cowtan's Book of Fourier

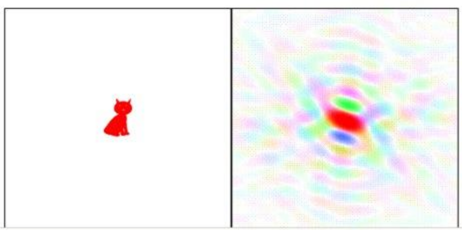
http://www.ysbl.york.ac.uk/~cowtan/fourier/fourier.html

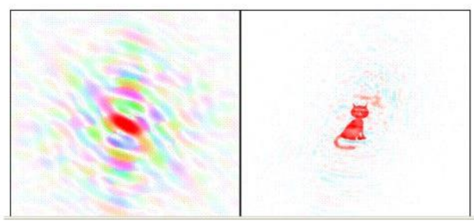


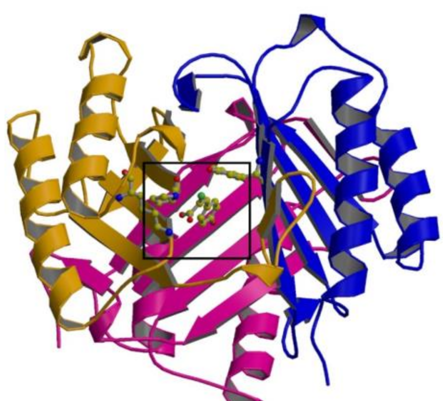


a) Cat - Cat Transform (Amplitudes only)
b) Manx (tailless) Cat - Manx Transform

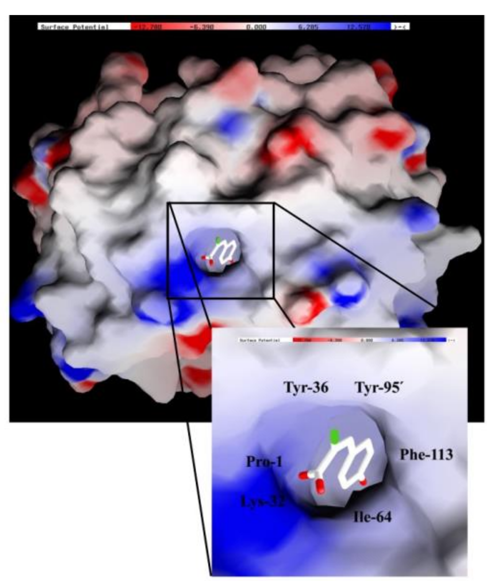
c) Cat Amplitudes + Manx Phases
d) [2x(Cat Amplitudes) - Manx Amplitudes] + Manx Phases





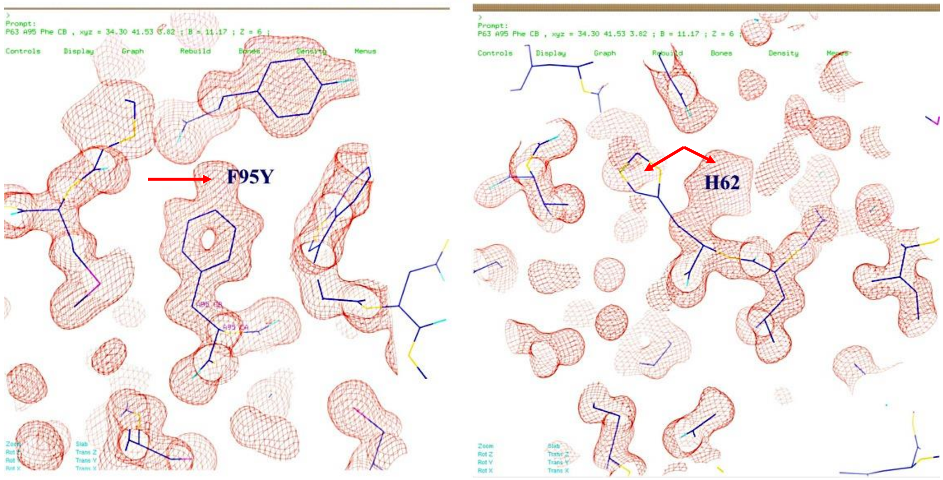


MIF + FHC

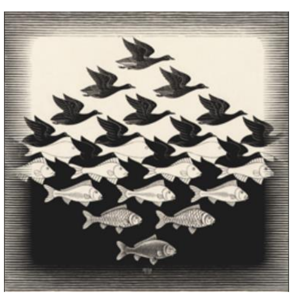


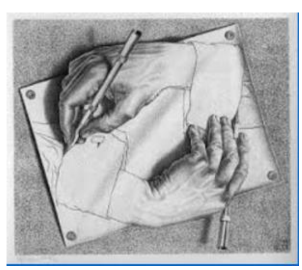
Examples of Difference Fouriers

MIF - 1.5A 2Fo-Fc

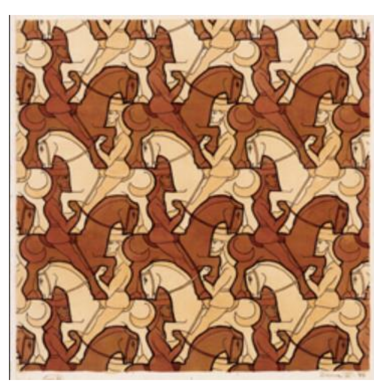


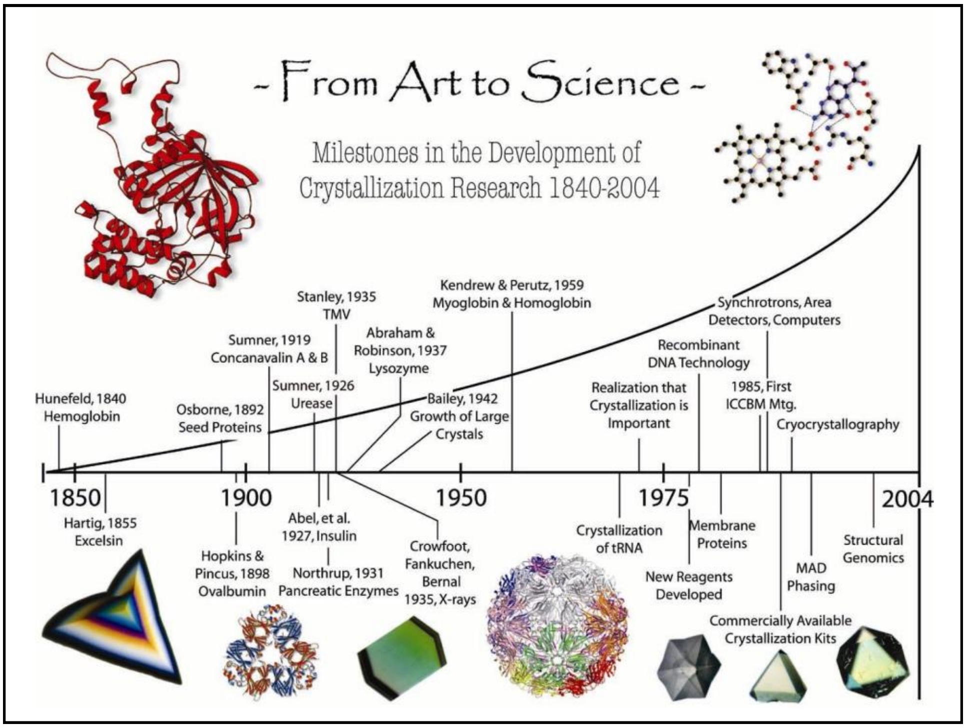
M.C. Escher





Escher sketch




- From Art to Science -
Milestones in the Development of
Crystallization Research 1840-2004
Hunefeld, 1840
Hemoglobin
Osborne, 1892
Seed Proteins
Sumner, 1919
Concanavalin A & B
Stanley, 1935
TMV
Sumner, 1926
Urease
Abraham &
Robinson, 1937
Lysozyme
Bailey, 1942
Growth of Large
Crystals
Kendrew & Perutz, 1959
Myoglobin & Homoglobin
Realization that
Crystallization is
Important
Recombinant
DNA Technology
Synchrotrons, Area
Detectors, Computers
1985, First
ICCBM Mtg.
Cryocrystallography
1850
1900
1950
1975
2004
Hartig, 1855
Excelsin
Hopkins &
Pincus, 1898
Ovalbumin
Abel, et al.
1927, Insulin
Northrup, 1931
Pancreatic Enzymes
Crowfoot,
Fankuchen,
Bernal
1935, X-rays
Crystallization
of tRNA
Membrane
Proteins
New Reagents
Developed
Commercially Available
Crystallization Kits
MAD
Phasing
Structural
Genomics


2014
International year of crystallography
IUCr
International Year of Crystallography – 2014
General Assembly
International Year of Crystallography
IYCr2014
A Celebration with an Emphasis on Education?
Have an idea? Share it!!
Support Bragg's Law
I ♥ Crystallography
Driven to Diffraction
** Nobel Laureates **