Intensities of the Reflections

With the help of Braggs law and the Ewald construction, we can calculate the place of a reflection on the detector, provided we know the unit cell dimensions. Indeed, the position of a spot is determined alone by the metric symmetry of the unit cell. The intensity of a spot, however, depends on the contents of the unit cell (and, of course, on exposure time, crystal size, *etc.*).

Reflections are weakened by the radius of the atoms (atomic form factor) and the thermal motion of the atoms (temperature factor U). Both these effects are stronger at high resolution.
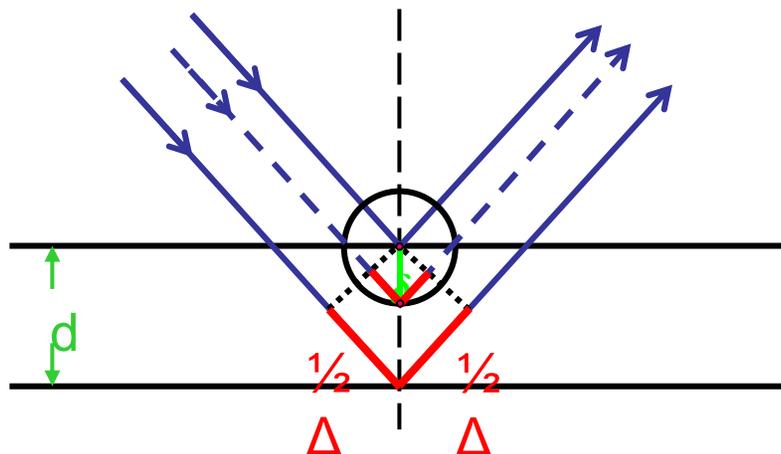
Removed due to copyright restrictions.

Please see:

Massa, Werner. *Crystal Structure Determination. 2nd ed.* Translated into English by R. O. Gould. New York, NY: Springer, 2004. ISBN: 3540206442.

Structure Factors

Every atom in the unit cell contributes to every reflection according to its chemical nature and its relative position. Owing to this shift in position relative to the other atoms, the photons contributed by each atom in the unit cell have a phase shift relative to those from other atoms.

$$\Phi_i = \Delta\Phi_{i(a)} + \Delta\Phi_{i(b)} + \Delta\Phi_{i(c)} = 2\pi\left(hx_i + ky_i + lz_i\right)$$
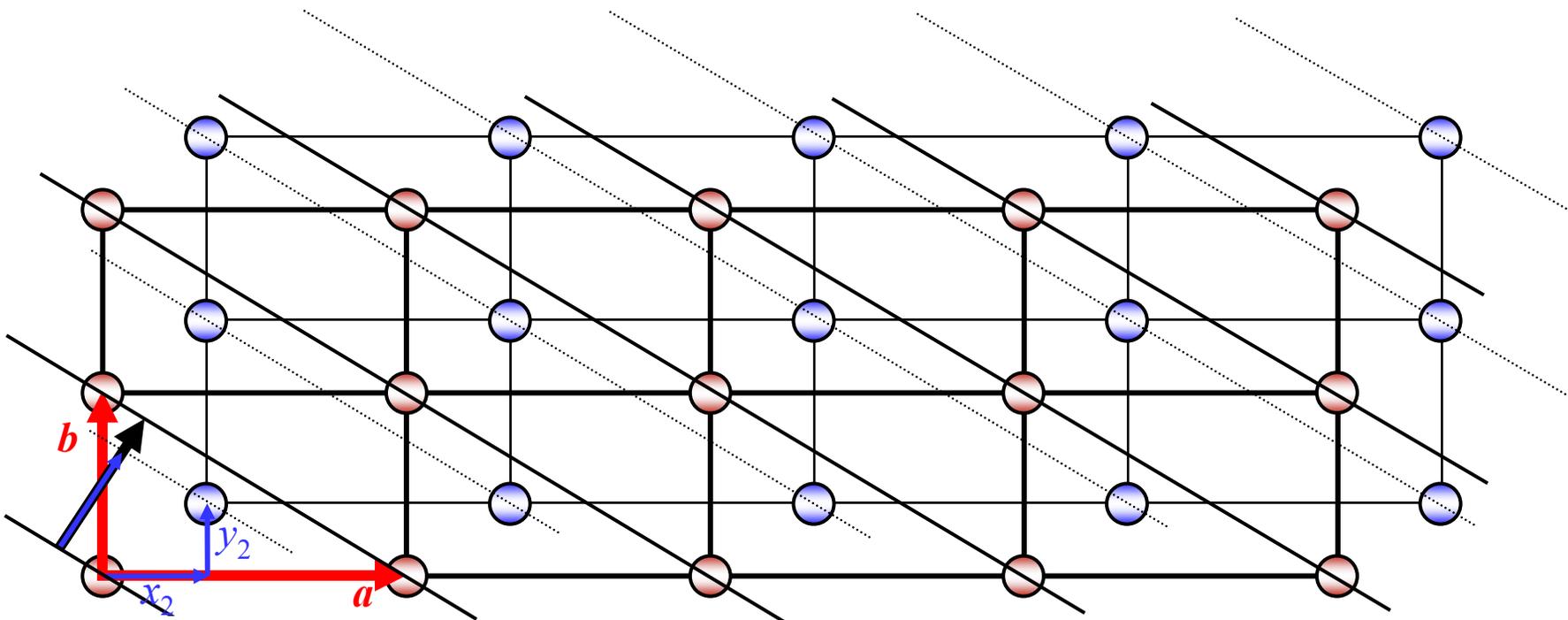
$$0 \leq \Delta\Phi \leq 2\pi$$

Structure Factors

$$\Phi_i = \Delta\Phi_{i(a)} + \Delta\Phi_{i(b)} + \Delta\Phi_{i(c)} = 2\pi\left(hx_i + ky_i + lz_i\right)$$

This makes the structure factor a complex number:

$$F_i = f_i \cdot \exp\left[i\Phi_i\right] = f_i\left(\cos\Phi_i + i\sin\Phi_i\right)$$

Every atom i in the unit cell contributes to every structure factor $F(hkl)$ (that is reflection) according to its position in the cell and its chemical nature (different values for f_i !):

$$F(hkl) = \sum_i f_i\left[\cos 2\pi\left(hx_i + ky_i + lz_i\right) + i\sin 2\pi\left(hx_i + ky_i + lz_i\right)\right]$$

Complex Numbers and the Argand Plane

In general: Complex numbers have both real and imaginary components: They have the general form $x = a + ib$, where $i = \sqrt{-1}$.
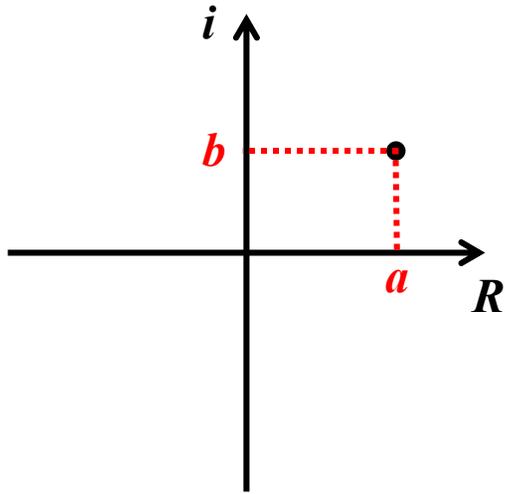
Euler's equation gives a different perspective for complex numbers:

$$e^{i\alpha} = \cos(\alpha) = i\sin(\alpha)$$

This equation can be verified by expanding the functions: $y(\alpha) = e^{i\alpha}$ and $y(\alpha) = \cos(\alpha) + i\sin(\alpha)$ in a Taylor's series about $\alpha = 0$ and matching the expressions term for term.

Another check is that both expressions obey the same differential equation: $y' = iy$.

The expression $e^{i\alpha}$ is apparently a complex number represented by a vector of unit length and angle α in the complex plane. The expression $re^{i\alpha}$ is a complex number of magnitude (length) r and angle α.
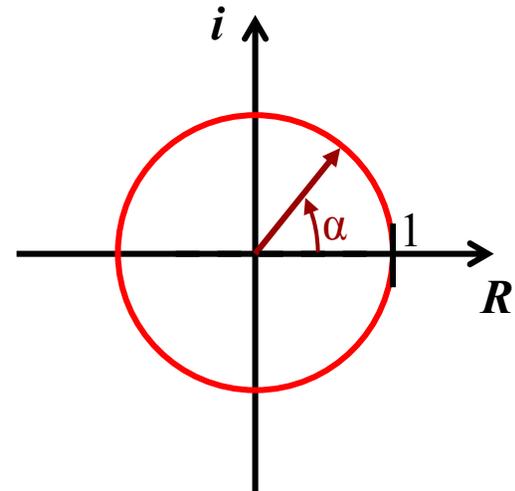
Complex Numbers and the Argand Plane

A given complex number may be specified by either its real and I maginary components (a and b) or its magnitude and phase (r and α), with the following relationships between them:

$$a = r\cos(\alpha) \qquad r = \left(a^2 + b^2\right)^{1/2}$$

$$b = r\sin(\alpha) \qquad \alpha = \tan^{-1}(b/a)$$
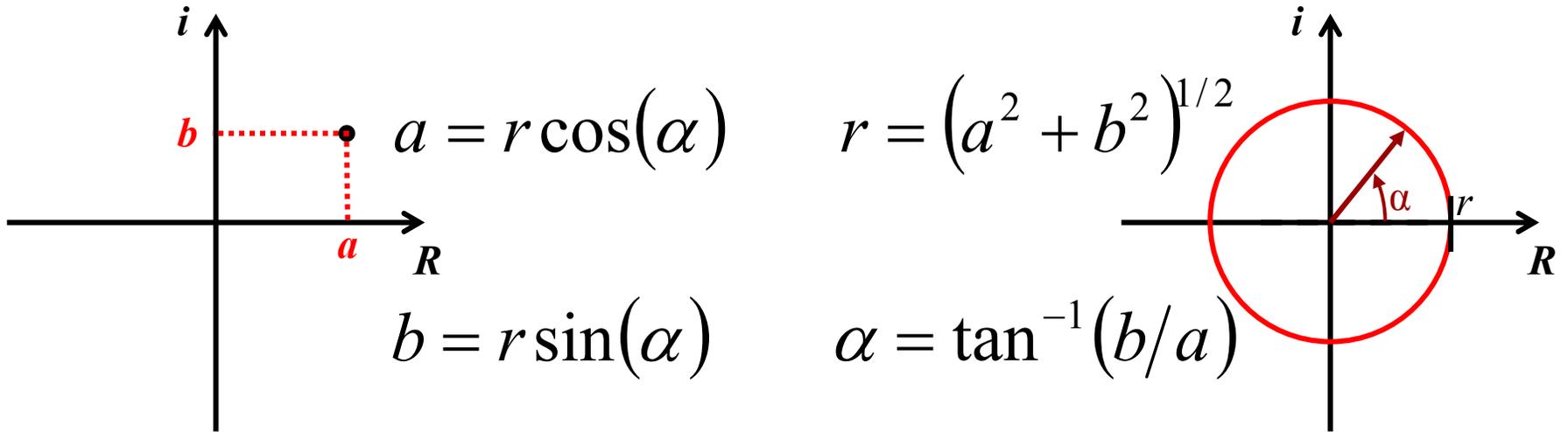
Structure factors $\vec{F}(hkl)$ are complex quantities. If the magnitude of a particular reflection is measured but the phase has not been determined yet, the possible values of that structure factor can be represented by a circle of radius $\left|\vec{F}\right|$ in the complex plane.

Structure Factors

$$F(hkl) = \sum_i f_i \left[\cos 2\pi \left(hx_i + ky_i + lz_i \right) + i \sin 2\pi \left(hx_i + ky_i + lz_i \right) \right]$$

$$= \sum_i f_i \exp\left[i\Phi_i \right]$$

Argand plane
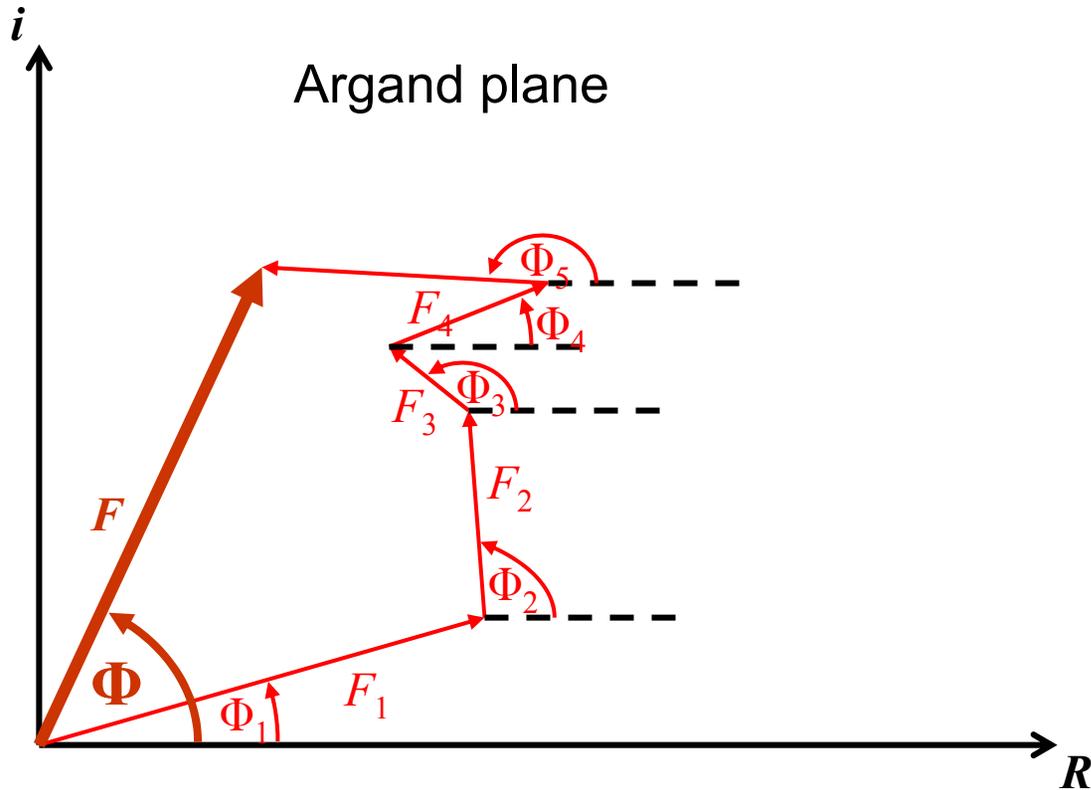
Electron Density

$$F(hkl) = \sum_i f_i \left[\cos 2\pi \left(hx_i + ky_i + lz_i \right) + i \sin 2\pi \left(hx_i + ky_i + lz_i \right) \right]$$

$$\big\downarrow \text{FT} \big\uparrow$$

$$\rho_{x,y,z} = \frac{1}{V} \sum_{hkl} F_{hkl} \cdot \exp\left[-i 2\pi \left(hx + ky + lz \right) \right]$$

The electron density at every given place in the unit cell (real space) can be calculated from the equation above.

Note that x_i, y_i, z_i in the structure factor equation refer to atomic coordinates, while x, y, z in the electron density equation refer to arbitrary places anywhere in the unit cell.

Electron Density

$$\rho_{x,y,z} = \frac{1}{V} \sum_{hkl} F_{hkl} \cdot \exp\left[-i2\pi(hx + ky + lz)\right]$$

1.7 Å
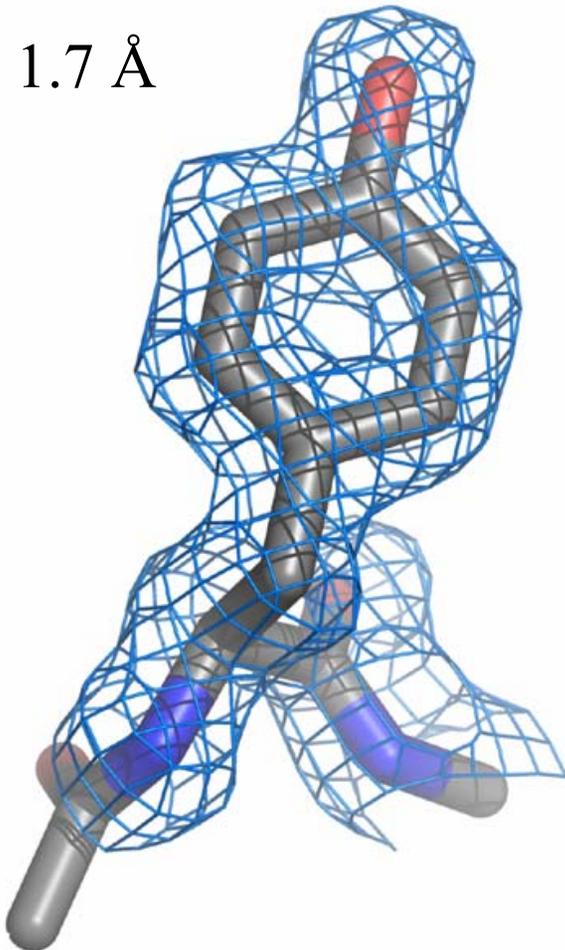
It is sufficient to calculate the electron density at a number of grid points within the unit cell and extrapolate between the points.

0.9 Å

N(3)   C(3)   N(2)   N(4)   C(4)   C(5)   C(1)   C(2)   O(1)   N(1)   S

Courtesy of George M. Sheldrick. Used with permission.
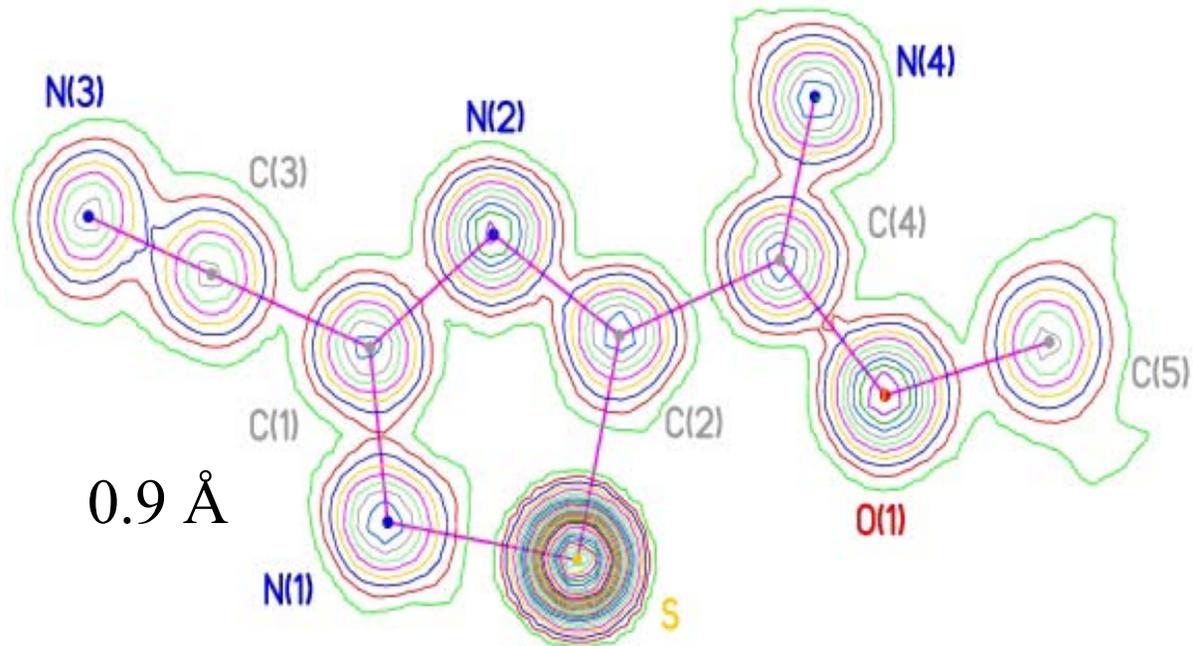
Calculating Electron Density

$$\rho_{x,y,z} = \frac{1}{V} \sum_{hkl} F_{hkl} \cdot \exp\left[-i2\pi\left(hx + ky + lz\right)\right]$$

To make this equation work we need the volume of the unit cell (easy), a dataset with the intensities for h, k and l (also easy) and the phase of every structure factor (not quite so easy).

We can measure the intensities and determine the unit cell from the locations of the reflections, but it is very difficult to actually measure the relative phase of the reflections. This is known as the "crystallographic phase problem".

The Phase Problem: Fourier Synthesis

Let's imagine a one dimensional crystal with a three-atomic molecule in the unit cell: two carbons and one oxygen. The electron density in this unit cell looks like this:
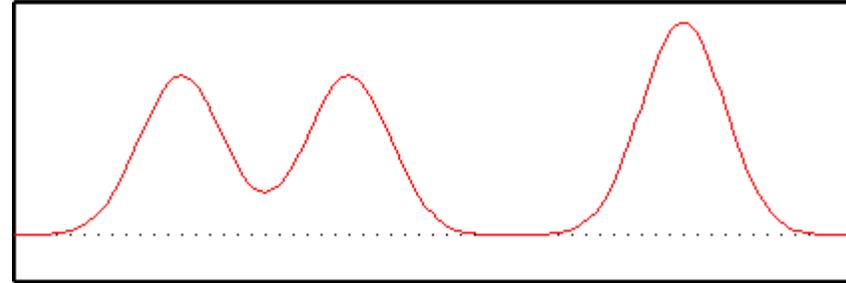
Try to represent this function in terms of sine waves. The first sine wave has a frequency of 2, that is there are two repeats of the wave across the unit cell. One peak represents the oxygen, and the other the two carbons:
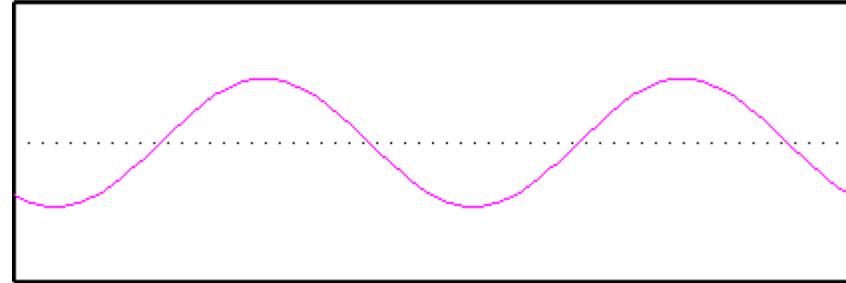
The second sine wave has a frequency of 3; three repeats of the wave across the unit cell. It has a different phase, in other words we start at a different place on the wave. The amplitude is also different:
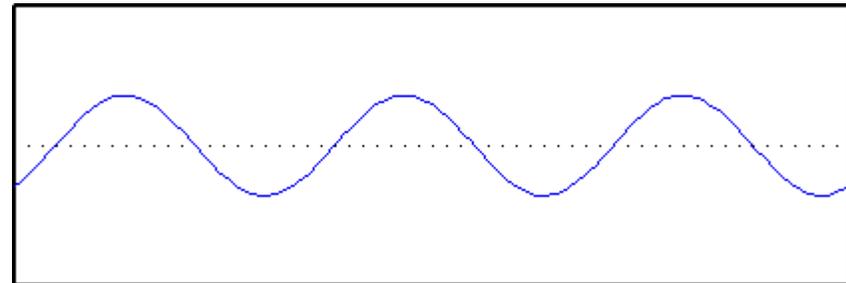
The Phase Problem: Fourier Synthesis

Let's imagine a one dimensional crystal with a three-atomic molecule in the unit cell: two carbons and one oxygen. The electron density in this unit cell looks like this:
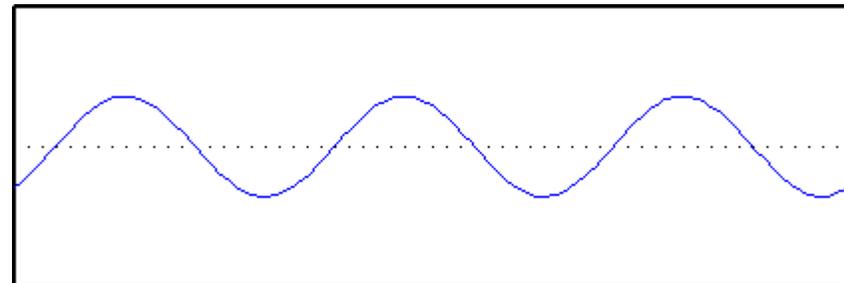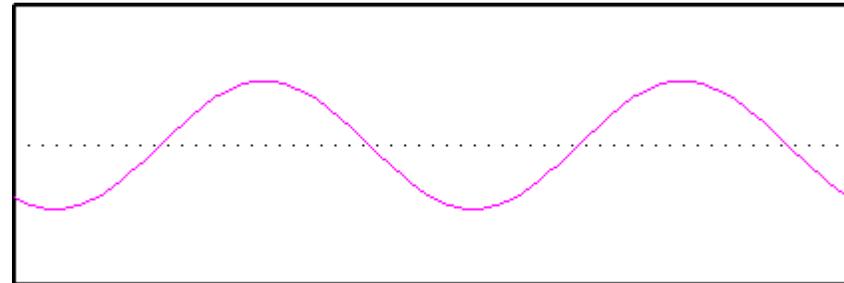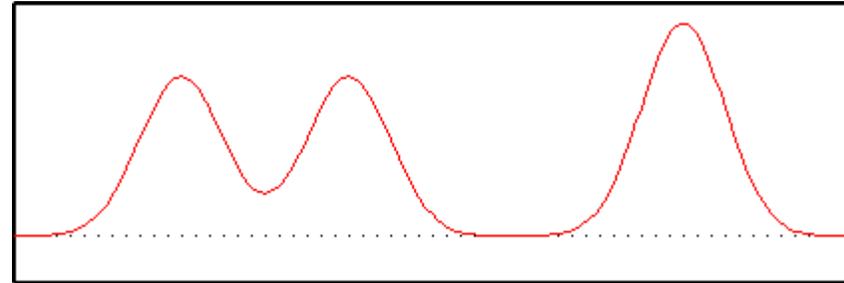
Finally, we introduce a sine wave with a frequency of 5. Two of the peaks of this wave are lined up with the carbon atoms:
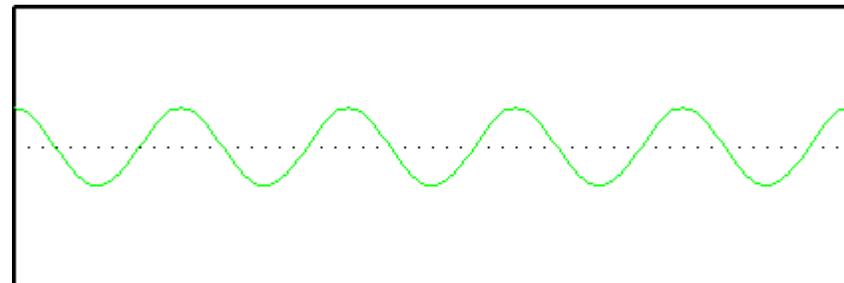
The Phase Problem: Fourier Synthesis

Let's imagine a one dimensional crystal with a three-atomic molecule in the unit cell: two carbons and one oxygen. The electron density in this unit cell looks like this:
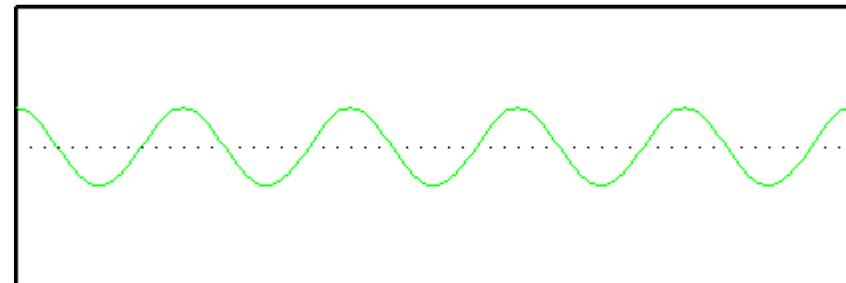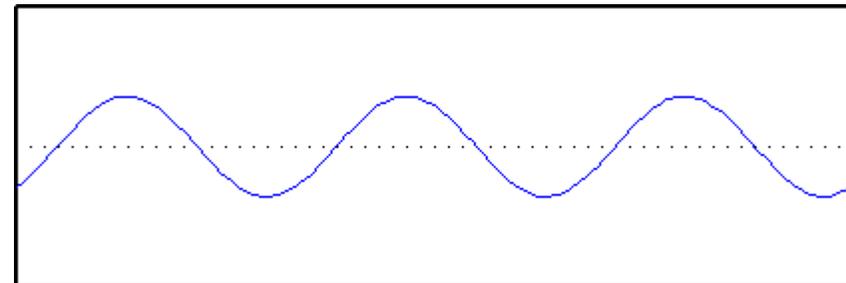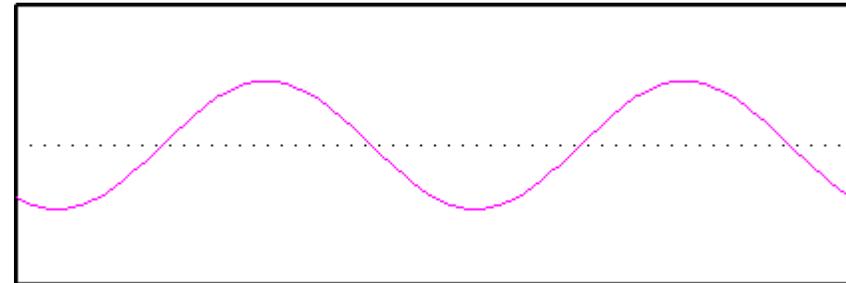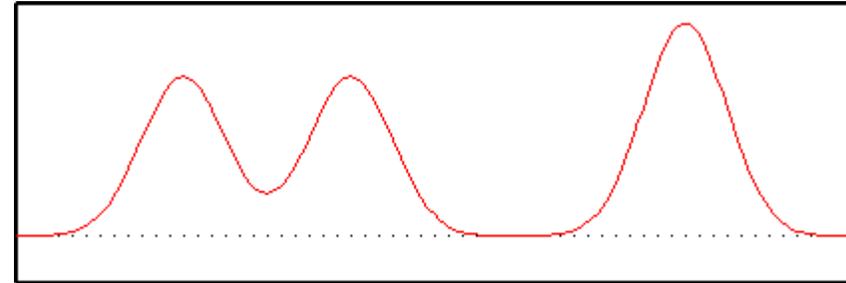
Now we add them all together:
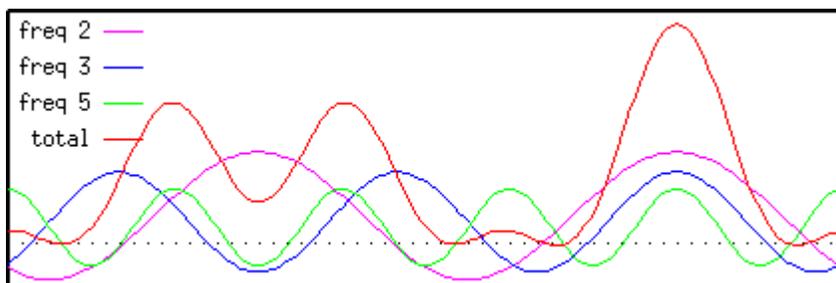
The Phase Problem: Fourier Synthesis
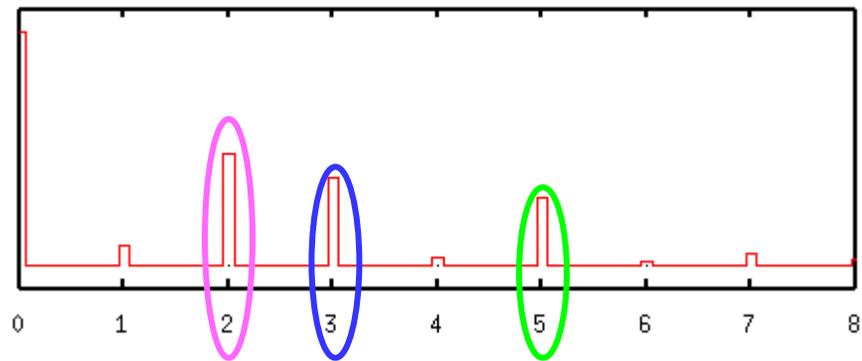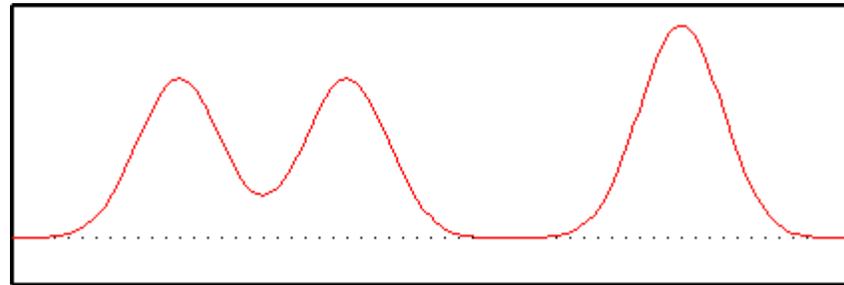
Now we will look at the Fourier Transform of the same unit cell

Now we add them all together:

FT

FT

freq 2
freq 3
freq 5
total
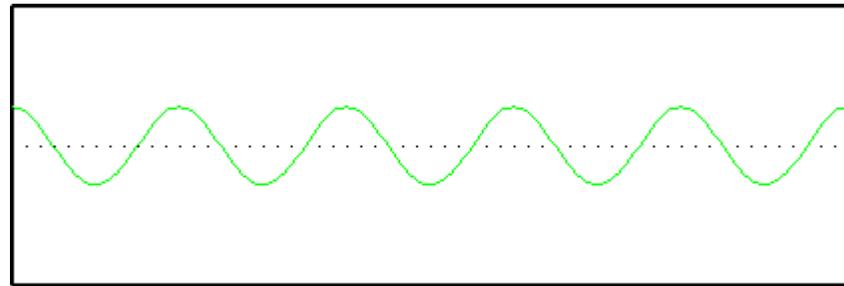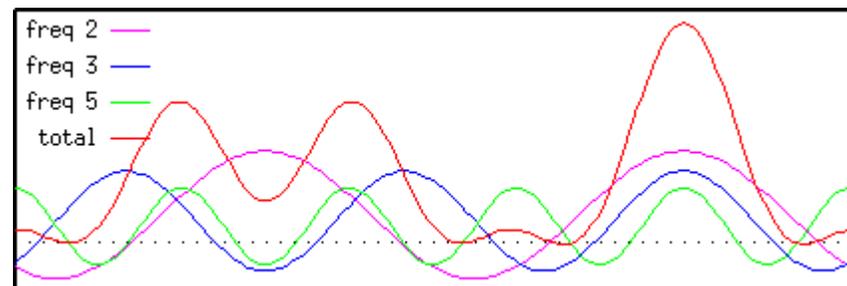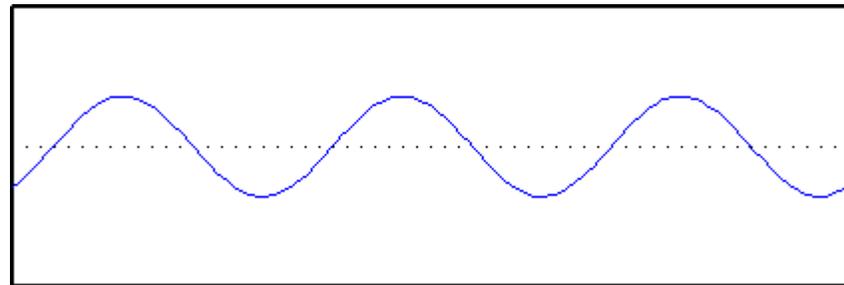
The Phase Problem: Animal Magic

A duck

FT →
← FT
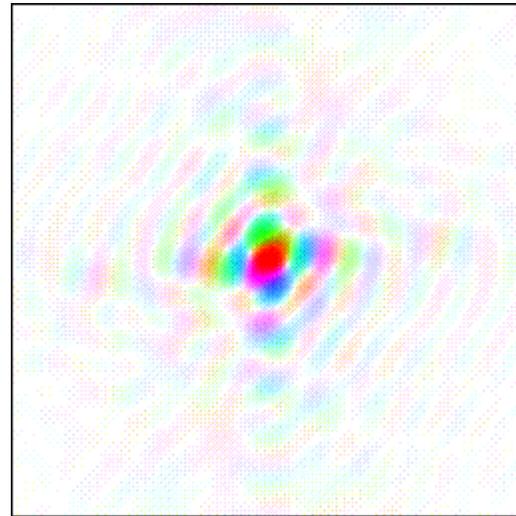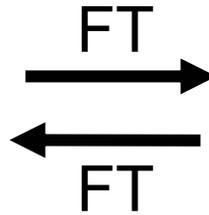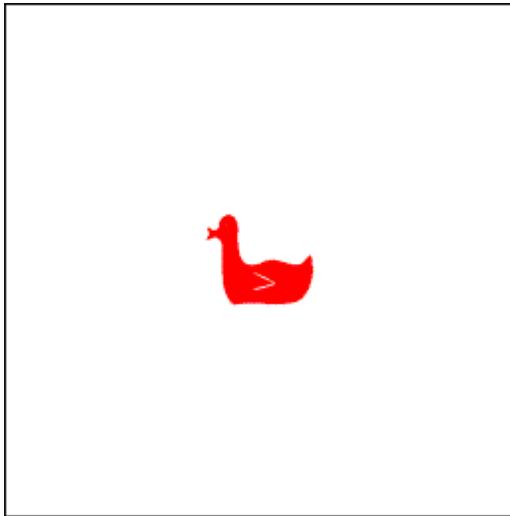
Fourier transform of a duck

A cat

FT →
← FT
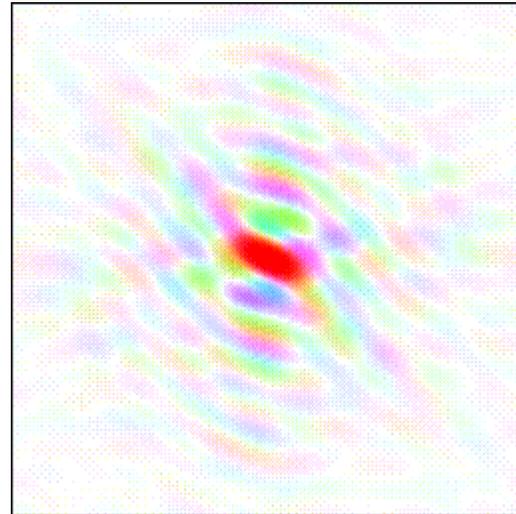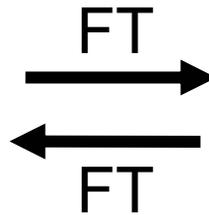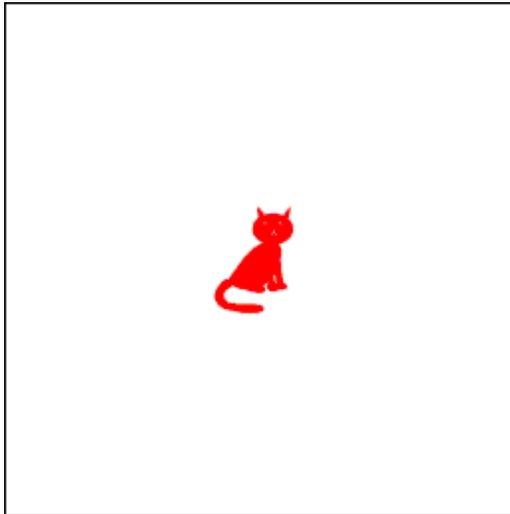
Fourier transform of a cat

The Phase Problem: Animal Magic

Combine the **magnitudes** from the **Duck** transform with the **phases** from the **Cat** transform
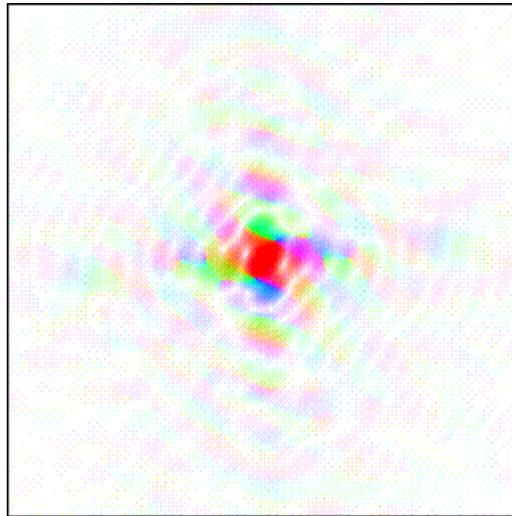
FT
FT
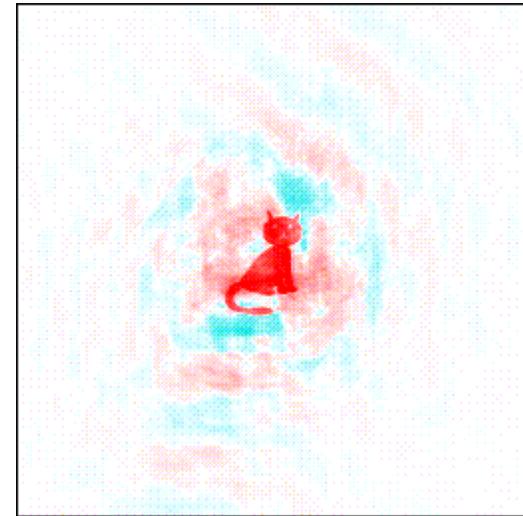
Combine the **magnitudes** from the **Cat** transform with the **phases** from the **Duck** transform
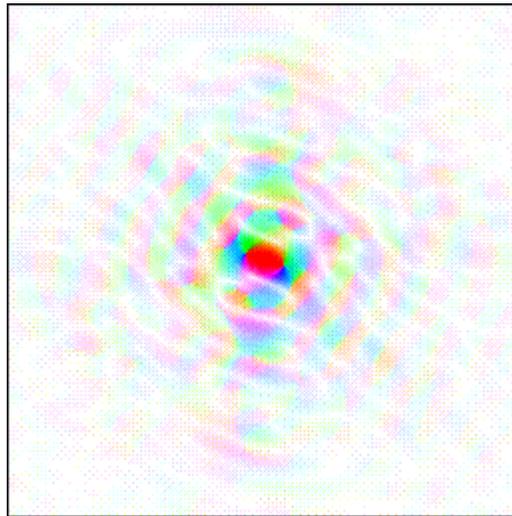
FT
FT
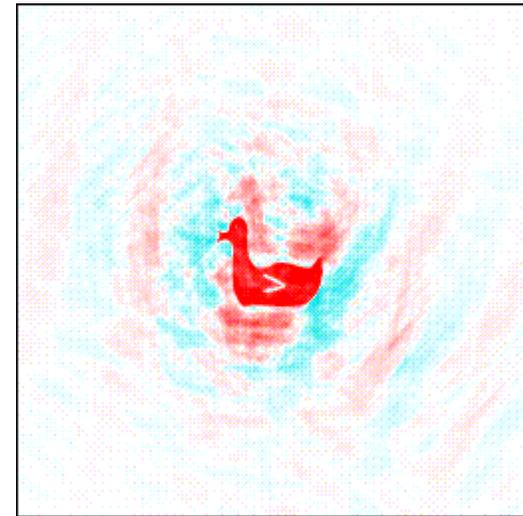
The phase contains the bulk of the information!

Symmetry in Reciprocal Space

Before we can talk about how to solve the phase problem, we need to talk about symmetry in reciprocal space.

Symmetry in the crystal (real space) influences the symmetry in reciprocal space. A twofold in real space causes a twofold in reciprocal space, a mirror, causes a mirror, a fourfold a fourfold, etc.

The reciprocal space, however, is always centrosymmetric, and translational symmetry has no effect on the symmetry in reciprocal space. That means a *P*2, *P*2$_1$, *P*2/m, *P*2$_1$/m, *P*2$_1$/c, *etc.* all correspond to the same symmetry group in reciprocal space.

The symmetry group in reciprocal space is called Laue group. There are eleven of them.

Symmetry in Reciprocal Space

Diffraction pattern of a cubic insulin crystal, observed with CuKα radiation.
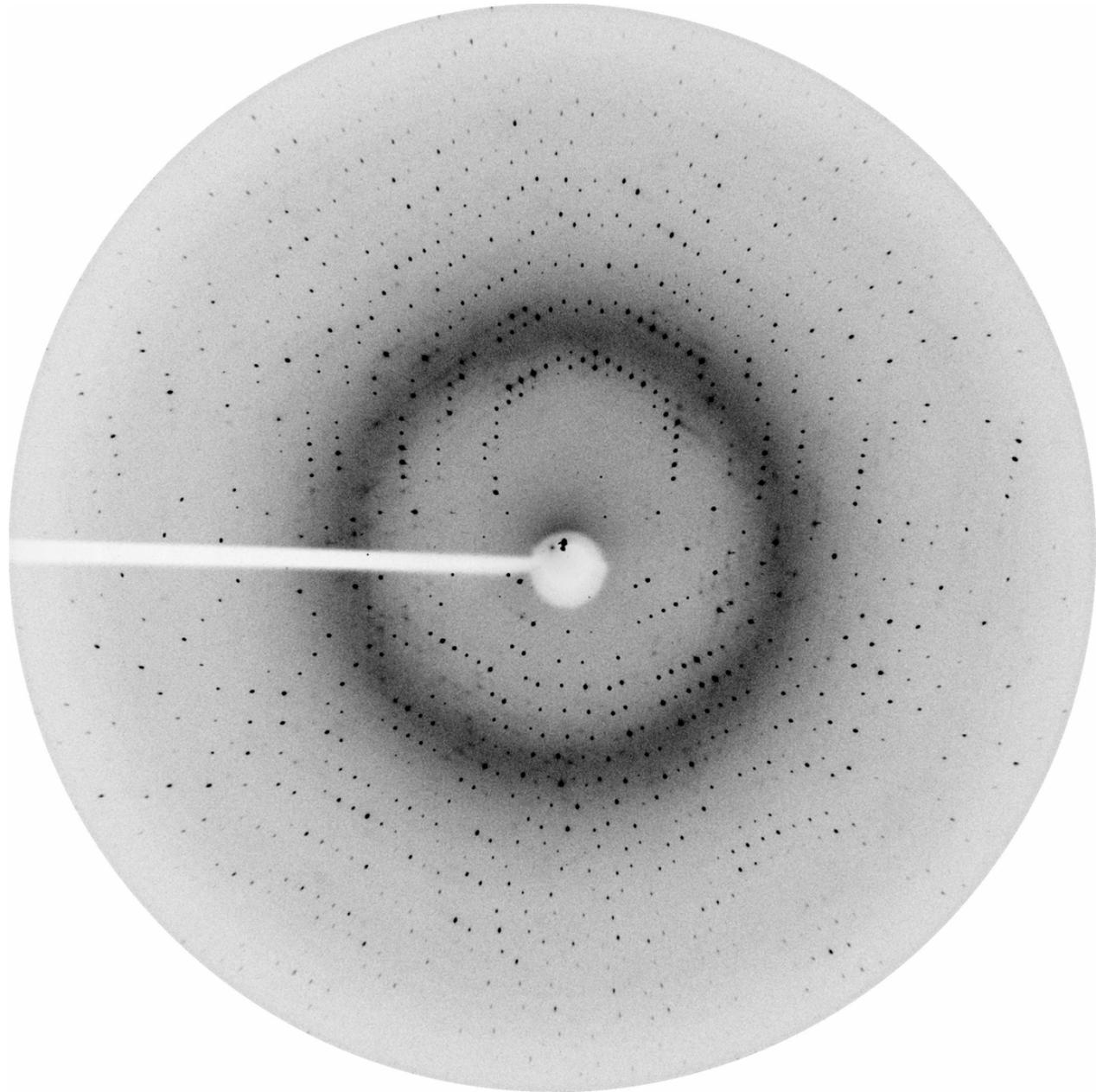
Symmetry in Reciprocal Space

Triclinic: hk0 layer.
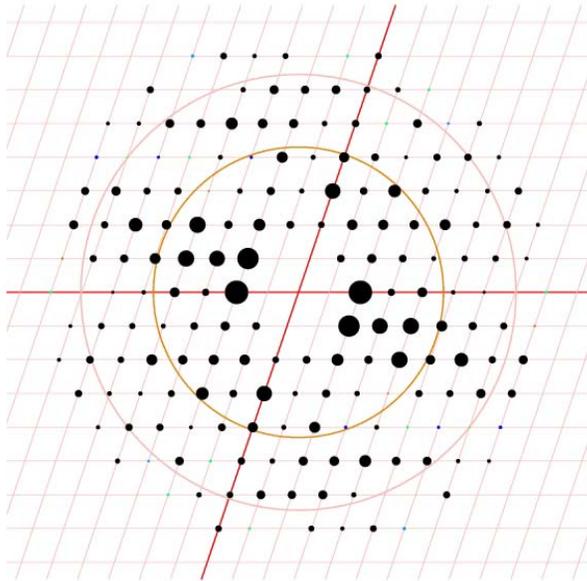
Triclinic: hk1 layer.
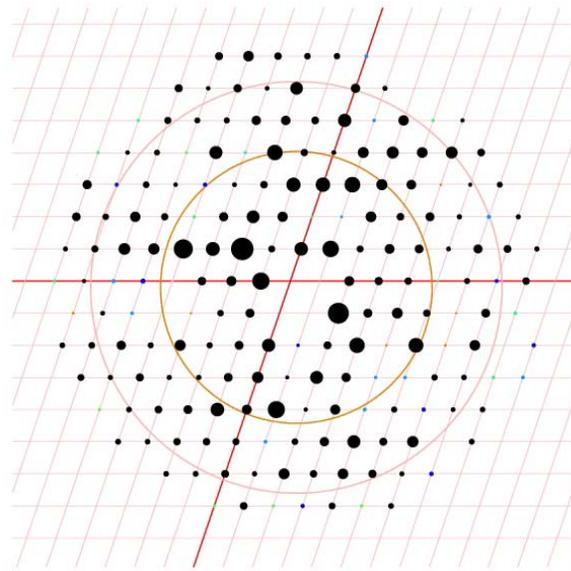
Orthorhombic: hk0 layer.
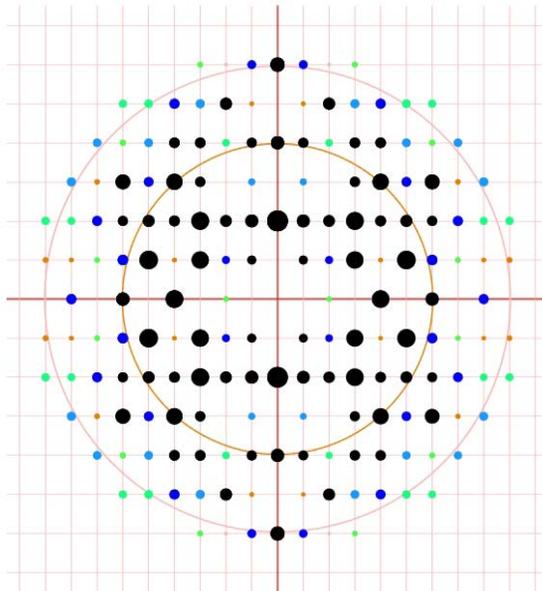
Orthorhombic: hk1 layer.
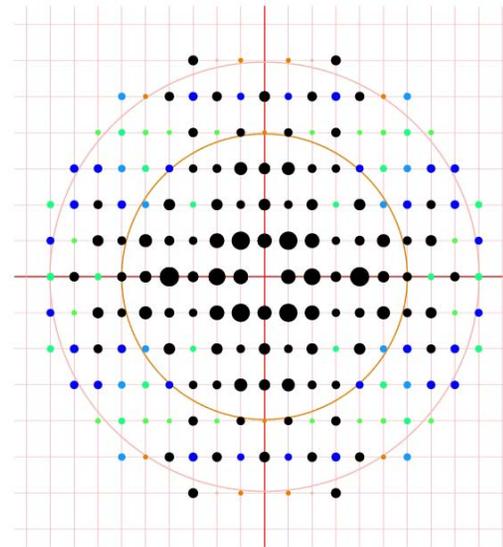
Symmetry in Reciprocal Space

The diffraction pattern is always centrosymmetric (at least in good approximation). Friedel's law: $I_{hkl} = I_{-h-k-l}$.

Fourfold symmetry in the diffraction pattern corresponds to a fourfold axis in the space group ($4, \bar{4}, 4_1, 4_2$ or 4_3), threefold to a threefold, etc.

If you take away the translational part of the space group symmetry and add an inversion center, you end up with the Laue group. The Laue group describes the symmetry of the diffraction pattern. **The Laue symmetry can be lower than the metric symmetry of the unit cell, but never higher.**

That means: A monoclinic crystal with $\beta = 90°$ is still monoclinic. The diffraction pattern from such a crystal will have monoclinic symmetry, even though the metric symmetry of the unit cell looks orthorhombic.

There are 11 Laue groups:
-1, 2/m, mmm, 4/m, 4/mmm, -3, -3/m, 6/m, 6/mmm, m3, m3m

Laue Symmetry

Crystal System	Laue Group	Point Group
Triclinic	-1	1, -1
Monoclinic	2/m	2, m, 2/m
Orthorhombic	mmm	222, mm2, mmm
Tetragonal	4/m	4, -4, 4/m
	4/mmm	422, 4mm, -42m, 4/mmm
Trigonal/ Rhombohedral	-3	3, -3
	-3/m	32, 3m, -3m
Hexagonal	6/m	6, -6, 6/m
	6/mmm	622, 6mm, -6m2, 6/mmm
Cubic	m3	23, m3
	m3m	432, -43m, m3m

Space Group Determination

The first step in the determination of a crystal structure is the determination of the unit cell from the diffraction pattern.

Second step: Space group determination.

From the symmetry of the diffraction pattern, we can determine the Laue group, which narrows down the choice quite considerably. Usually the Laue group and the metric symmetry of the unit cell match.

The $<| E^2-1 |>$ statistics, can give us an idea, whether the space group is centrosymmetric or acentric. Even thought the diffraction pattern is always centrosymmetric, the intensity distribution across the reciprocal space is much more even for a centrosymmetric space group.

From systematic absences, we can determine the lattice type as well as screw axes and glide planes.

This is usually enough to narrow down the choice to a very short list.

E^2-1 Statistics

We measure intensities I

$I \longrightarrow F^2$ F: structure factors

Normalized structure factors E:

$E^2 = F^2/<F^2>$ $<F^2>$: mean value for reflections at same resolution

$<E^2> = 1$

$< | E^2\text{-}1 | > =$ 0.736 for non-centrosymmetric structures
0.968 for centrosymmetric structures

Heavy atoms on special positions and twinning tend to lower this value.
Pseudo translational symmetry tend to increase this value.

E^2-1 Statistics

$< |\ E^2\text{-}1\ | > =$ 0.736 for non-centrosymmetric structures

0.968 for centrosymmetric structures

$<|E^2{-}1|> = 0.968$

$<|E^2{-}1|> = 0.736$
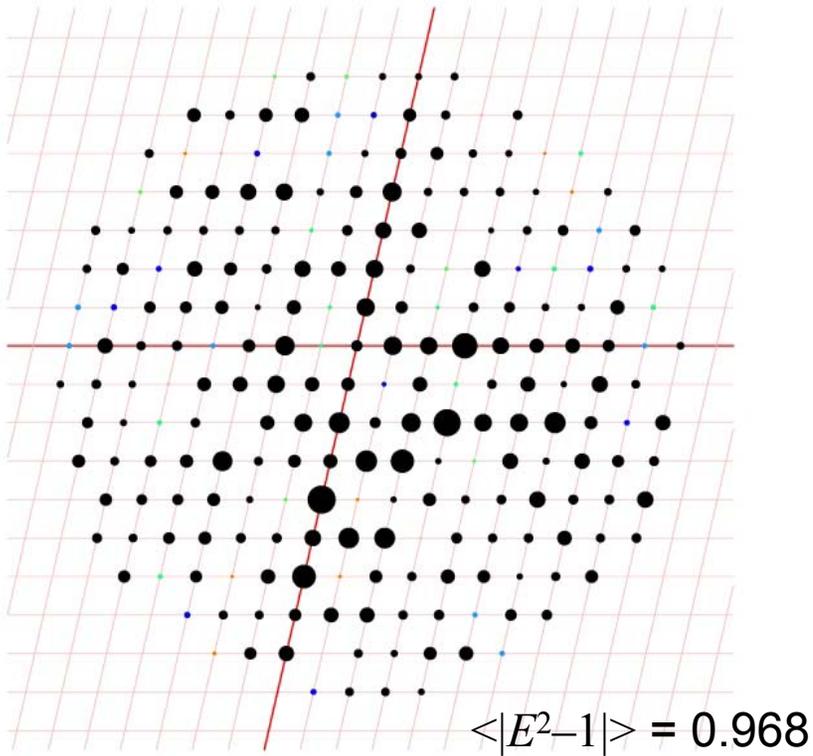
$2kl$ projection of the reflections of a structure in the space group P-1.
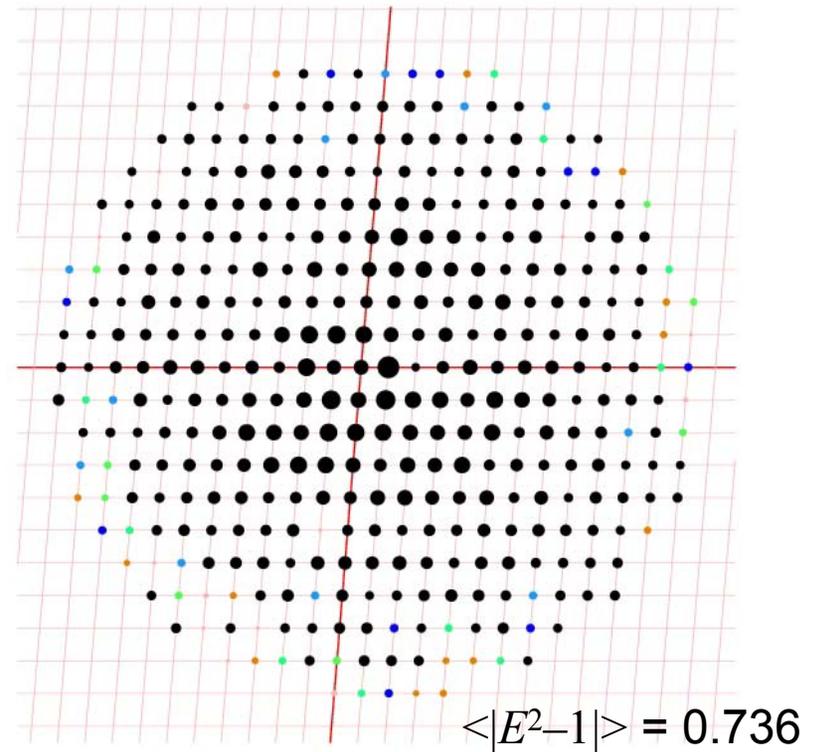
$2kl$ projection of the reflections of a structure in the space group $P1$.

Systematic Absences

Lattice centering and symmetry elements with translation (glide planes and screw axes) cause certain reflections to have zero intensity in the diffraction pattern. If, *e.g.*, ***all*** reflections $0, k, 0$ with odd values for k are absent, we know that we have a 2_1 axis along b.

Other example: if all reflections $h, 0, l$ with odd values for l are absent, we have a c glide plane perpendicular to b.

How come?

Systematic Absences

Monoclinic cell, projection along b with c glide plane (*e.g. Pc*).

(x, y, z)

(x, ½-y, ½+z)
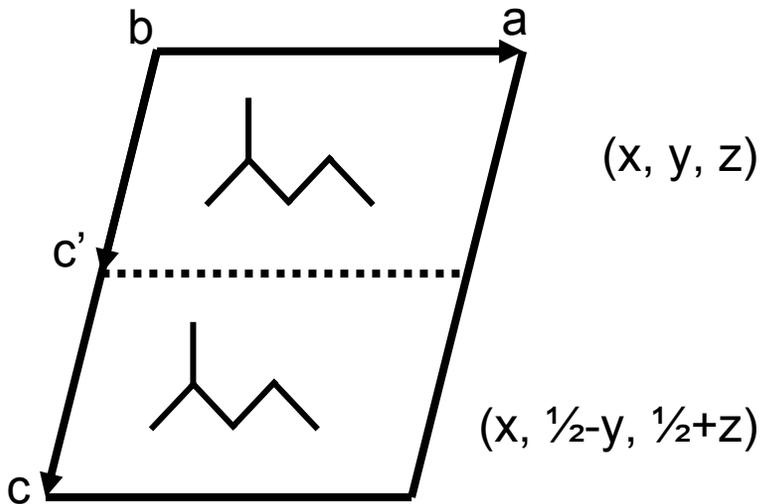
In this two 2D projection the structure is repeated at $c/2$. Thus, the unit cell *seems* to be half the size: $c' = c/2$ in this projection.

This doubles the reciprocal cell accordingly: $c^{*\prime} = 2c^*$. Therefore, the reflections *corresponding to this projection* $(h, 0, l)$ will be according to the larger reciprocal cell.

That means $h, 0, l$ reflections with $l \neq 2n$ are not observed.

Systematic Absences

Lattice centering

Reflections affected	Conditions for reflections	Symmetry element
hkl	none	P
	$h+k+l = 2n$	I
	$h+k = 2n$	C
	$k+l = 2n$	A
	$h+l = 2n$	B
	$-h+k+l = 3n$	R (obv.)
	$h-k+k = 3n$	R (rev.)

Systematic Absences

Glide Planes

Reflections affected	Conditions for reflections	Symmetry element
$0kl$	$k = 2n$	$b \llcorner a$
	$l = 2n$	$c \llcorner a$
	$k+l = 2n$	$n \llcorner a$
$h0l$	$l = 2n$	$c \llcorner b$
	$h+l = 2n$	$n \llcorner b$
$hk0$	$h = 2n$	$a \llcorner c$
	$k = 2n$	$b \llcorner c$
	$h+k = 2n$	$n \llcorner c$

Systematic Absences

Screw Axes

Reflections affected	Conditions for reflections	Symmetry element
$h00$	$h = 2n$	$2_1 \parallel a$
	$h = 4n$	$4_1, 4_3 \parallel a$
$0k0$	$k = 2n$	$2_1 \parallel B$
	$k = 4n$	$4_1, 4_3 \parallel b$
$00l$	$l = 2n$	$2_1, 4_2, 6_3 \parallel c$
	$l = 3n$	$3_1, 3_2, 6_2, 6_4 \parallel c$
	$l = 4n$	$4_1, 4_3 \parallel c$
	$l = 6n$	$6_1, 6_5 \parallel c$

Frequently Occurring Space Groups

Space group frequency in the Cambridge Structure Database (1990):

$P2_1/c$	39%
P-1	16%
$P2_12_12_1$	12%
$C2/c$	7%
$Pbca$	5%
Sum:	79%

Space group frequency in the Protein Data Bank (PDB):

$P2_12_12_1$	24%
$P3_121$ & $P3_221$	15%
$P2_1$	14%
$P4_12_12$ & $P4_32_12$	8%
$C2$	6%
Sum:	67%

The Triclinic, Monoclinic and Orthorhombic Space Groups

Crystal system	Laue group	Point group	Space group
Triclinic	$\bar{1}$	1	P1
		$\bar{1}$	P$\bar{1}$
Monoclinic	2/m	2	P2, P2$_1$, C2
		m	Pm, Pc, Cm, Cc
		2/m	P2/m, P2$_1$/m, C2/m, P2/c, <u>P2$_1$/c</u>, C2/c
Orthorhombic	mmm	222	P222, <u>P222$_1$</u>, <u>P2$_1$2$_1$2</u>, <u>P2$_1$2$_1$2$_1$</u>, C222, <u>C222$_1$</u>, I222, I2$_1$2$_1$2$_1$, F222
		$mm2$	Pmm2, Pmc2$_1$, Pcc2, Pma2, Pca2$_1$, Pnc2, Pmn2$_1$, Pba2, Pna2$_1$, Pnn2, Cmm2, Cmc2$_1$, Ccc2, Amm2, Abm2, Ama2, Aba2, Imm2, Iba2, Ima2, Fmm2, <u>Fdd2</u>
		mmm	Pmmm, <u>Pnnn</u>, Pccm, <u>Pban</u>, Pmma, <u>Pnna</u>, Pmna, <u>Pcca</u>, Pbam, <u>Pccn</u>, Pbcm, Pnnm, Pmmn, <u>Pbcn</u>, <u>Pbca</u>, Pnma, Cmcm, Cmca, Cmmm, Cccm, Cmma, <u>Ccca</u>, Immm, Ibam, <u>Ibca</u>, Imma, Fmmm, <u>Fddd</u>

Underlined: unambiguously determinable from systematic absences.
Red: chiral
Blue non-centrosymmetric
Black: centrosymmetric

5.069 Crystal Structure Analysis
Spring 2008