

MIT OpenCourseWare
<http://ocw.mit.edu>

5.069 Crystal Structure Analysis
Spring 2008

For information about citing these materials or our Terms of Use, visit: <http://ocw.mit.edu/terms>.

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 $\beta = 90^\circ$ 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 $\langle |E^2 - 1| \rangle$ statistics, can give us an idea, whether the space group is centrosymmetric or acentric. Even though 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.

E²-1 Statistics

We measure intensities I

$$I \longrightarrow F^2 \quad F: \text{structure factors}$$

Normalized structure factors E :

$$E^2 = F^2 / \langle F^2 \rangle \quad \langle F^2 \rangle: \text{mean value for reflections at same resolution}$$

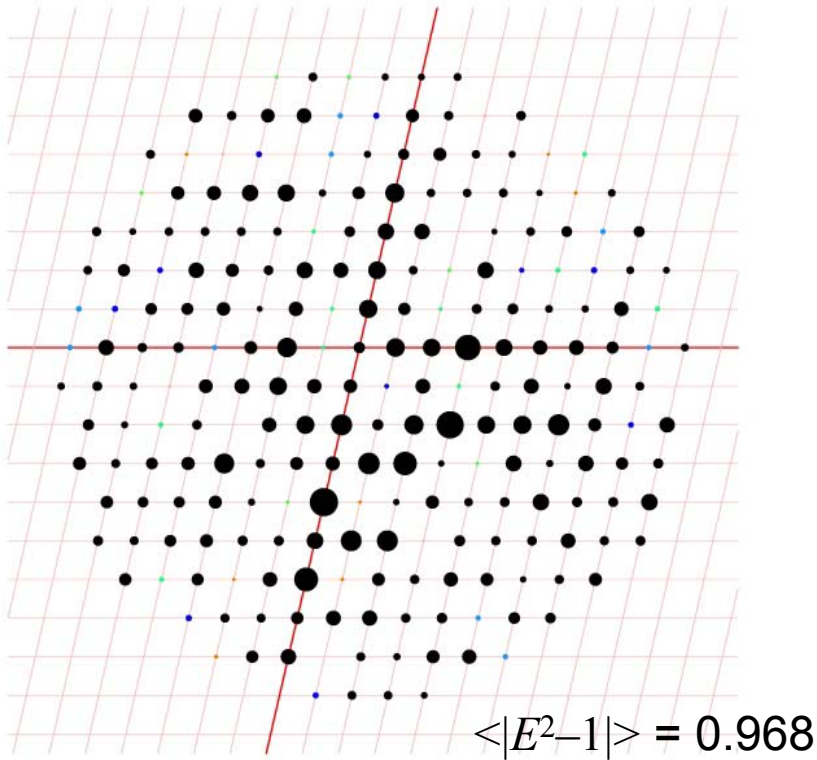
$$\langle E^2 \rangle = 1$$

$$\langle |E^2 - 1| \rangle = \begin{array}{l} 0.736 \text{ for non-centrosymmetric structures} \\ 0.968 \text{ for centrosymmetric structures} \end{array}$$

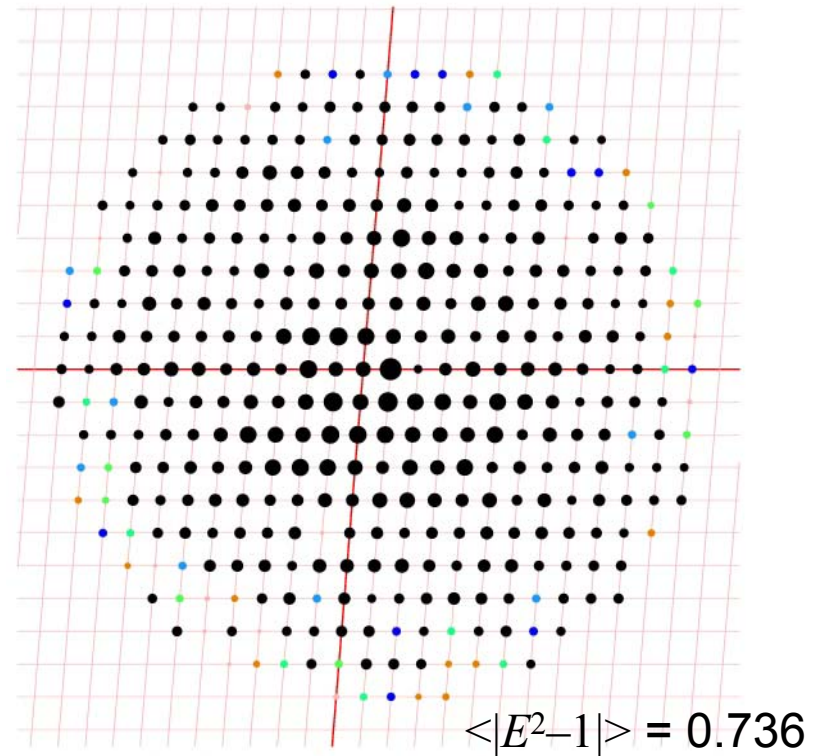
Heavy atoms on special positions and twinning tend to lower this value.
Pseudo translational symmetry tend to increase this value.

E²-1 Statistics

$\langle |E^2-1| \rangle =$ 0.736 for non-centrosymmetric structures
0.968 for centrosymmetric structures



$2kl$ 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

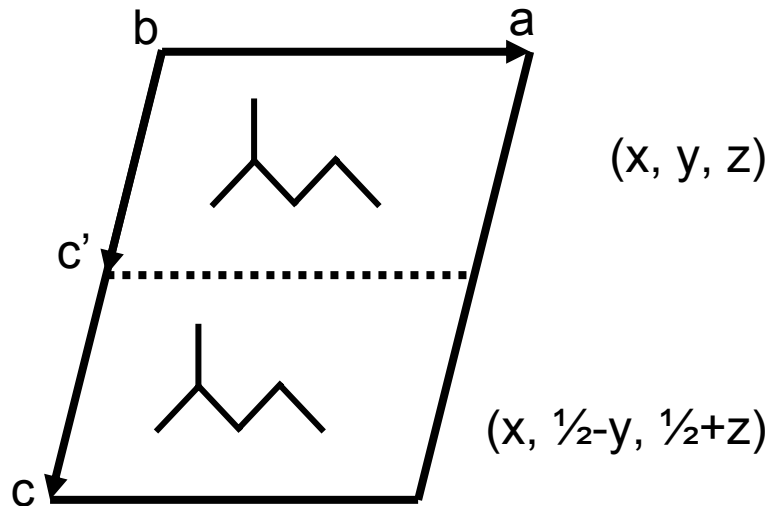
Lattice centering and symmetry elements with translation (glide planes and screw axes) cause certain reflections to have zero intensity in the diffraction pattern. If, e.g., **all** reflections $0, k, 0$ with odd values for k are absent, we know that we have a 2_1 axis along b .

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).



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: $c^{*'} = 2c^*$. 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.)
	$h-k+l = 3n$	R (rev.)

Systematic Absences

Glide Planes

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

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	-1	1	P1
		-1	P1
Monoclinic	2/m	2	P2, P2 ₁ , C2
		<i>m</i>	<i>Pm</i> , <i>Pc</i> , <i>Cm</i> , <i>Cc</i>
		2/m	P2/m, P2 ₁ /m, C2/m, P2/c, <u>P2₁/c</u> , C2/c
Orthorhombic	mmm	222	<u>P222</u> , <u>P222₁</u> , <u>P2₁2₁2</u> , <u>P2₁2₁2₁</u> , C222, <u>C222₁</u> , I222, I2 ₁ 2 ₁ 2 ₁ , F222
		<i>mm2</i>	<i>Pmm2</i> , <i>Pmc2₁</i> , <i>Pcc2</i> , <i>Pma2</i> , <i>Pca2₁</i> , <i>Pnc2</i> , <i>Pmn2₁</i> , <i>Pba2</i> , <i>Pna2₁</i> , <i>Pnn2</i> , <i>Cmm2</i> , <i>Cmc2₁</i> , <i>Ccc2</i> , <i>Amm2</i> , <i>Abm2</i> , <i>Ama2</i> , <i>Aba2</i> , <i>Imm2</i> , <i>Iba2</i> , <i>Ima2</i> , <i>Fmm2</i> , <u>Fdd2</u>
		<i>mmm</i>	<i>Pmmm</i> , <u><i>Pnnn</i></u> , <i>Pccm</i> , <u><i>Pban</i></u> , <i>Pmma</i> , <u><i>Pnna</i></u> , <i>Pmna</i> , <u><i>Pcca</i></u> , <i>Pbam</i> , <u><i>Pccn</i></u> , <i>Pbcm</i> , <i>Pnnm</i> , <i>Pmmn</i> , <u><i>Pbcn</i></u> , <u><i>Pbca</i></u> , <i>Pnma</i> , <i>Cmcm</i> , <i>Cmca</i> , <i>Cmmm</i> , <i>Cccm</i> , <i>Cmma</i> , <u><i>Ccca</i></u> , <i>Immm</i> , <i>Ibam</i> , <u><i>Ibca</i></u> , <i>Imma</i> , <i>Fmmm</i> , <u><i>Fddd</i></u>

Underlined: unambiguously determinable from systematic absences.

Red: chiral

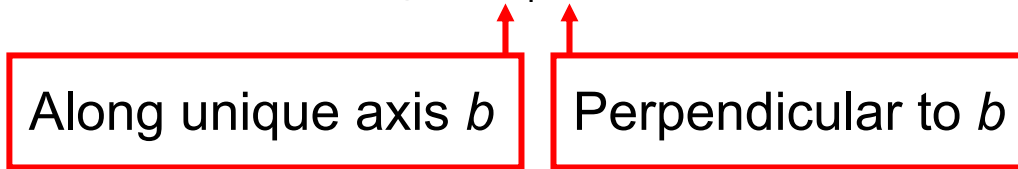
Blue non-centrosymmetric

Black: centrosymmetric

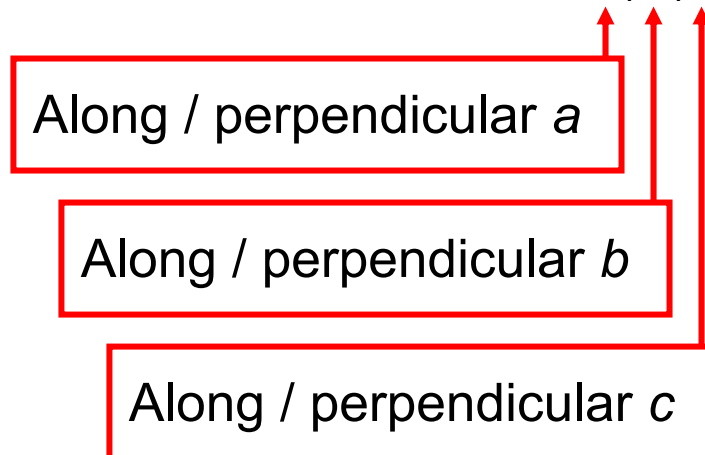
Crystallographic Directions

Triclinic: No unique directions, only two space groups, $P1$ and $P-1$

Monoclinic: b is unique. *E.g.* $P2_1/c$

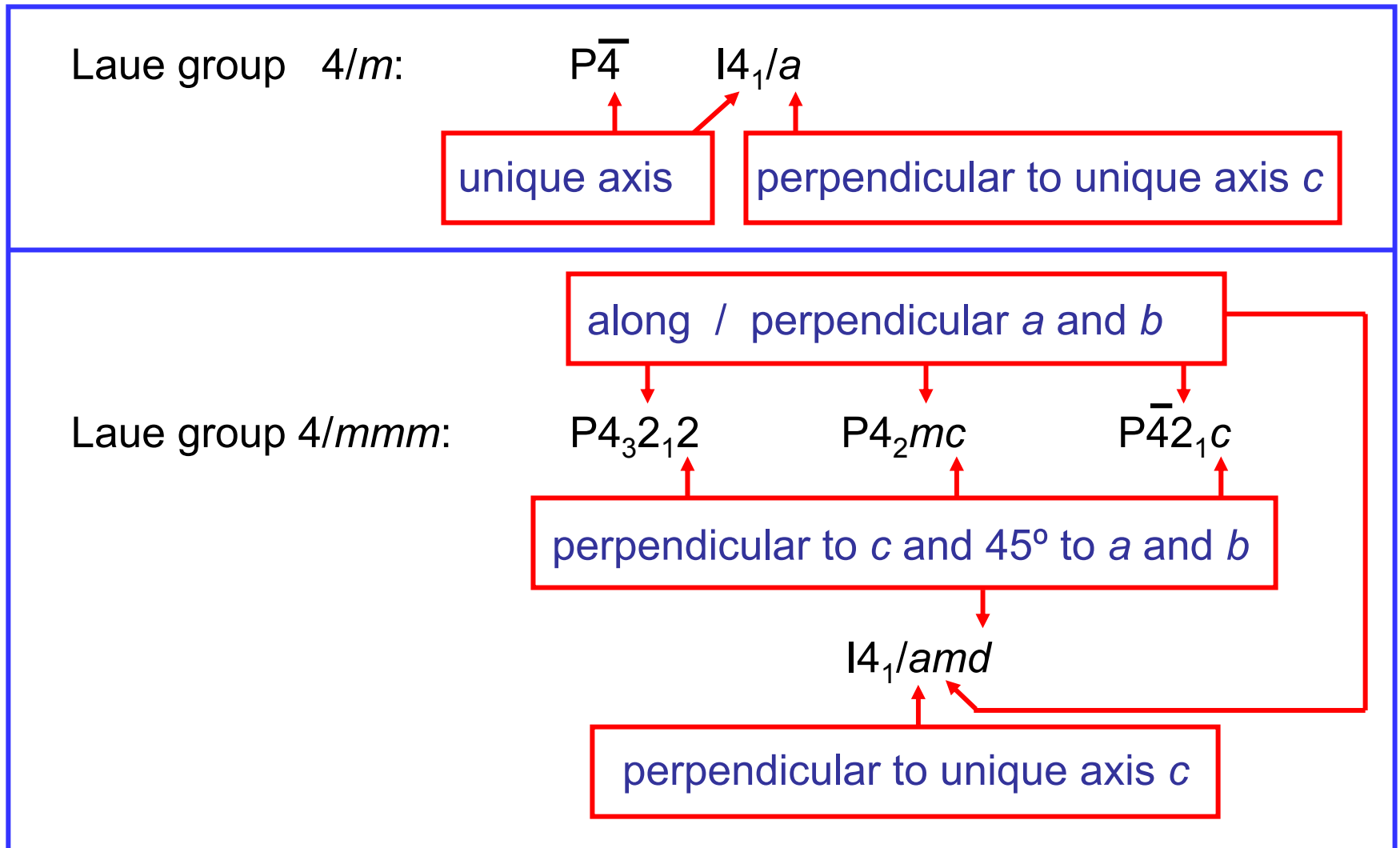


Orthorhombic: no unique directions. *E.g.* $P2_12_12_1$



Crystallographic Directions: Tetragonal Space Groups

There are two tetragonal Laue groups, $P4/m$ and $P4/mmm$. The unique axis is always c (that's where the 4-fold is). Space group symbols:



The Tetragonal Space Groups

Crystal system	Laue group	Point group	Space group
Tetragonal	$4/m$	4	$P4, P4_1, P4_2, P4_3, I4, \underline{I4_1}$
		$\bar{4}$	$P\bar{4}, I\bar{4}$
		$4/m$	$P4/m, P4_2/m, \underline{P4/n}, \underline{P4_2/n}, I4/m, \underline{I4_1/a}$
Tetragonal	$4/mmm$	422	$P422, \underline{P42_12}, P4_122, P4_12_12, \underline{P4_222}, \underline{P4_22_12}, P4_322, P4_32_12, I422, I4_122$
		$4mm$	$P4mm, P4bm, P4_2cm, P4_2nm, P4cc, P4nc, P4_2mc, P4_2bc, I4mm, I4cm, I4_1md, \underline{I4_1cd}$
		$\bar{4}m$	$P\bar{4}2m, P\bar{4}2c, P\bar{4}2_1m, P\bar{4}m2, P\bar{4}c2, \underline{P\bar{4}2_1c}, P\bar{4}b2, P\bar{4}n2, I\bar{4}m2, I\bar{4}c2, I\bar{4}2m, I\bar{4}2d$
		$4/mmm$	$P4/mmm, P4/mcc, \underline{P4/nbm}, \underline{P4/nnc}, P4/mbm, P4/mnc, \underline{P4/nmm}, \underline{P4/ncc}, P4_2/mmc, P4_2/mcm, \underline{P4_2/nbc}, \underline{P4_2/nnm}, P4_2/mbc, P4_2/mnm, \underline{P4_2/nmc}, \underline{P4_2/ncm}, I4/mmm, I4/mcm, \underline{I4_1/amd}, \underline{I4_1/acd}$

Underlined: unambiguously determinable from systematic absences.

Red: chiral **Blue** non-centrosymmetric **Black**: centrosymmetric

Courtesy of George M. Sheldrick.
Used with permission.

The Patterson Function

We measure intensities I . After applying several corrections, they translate into squared structure factors also known as structure factor amplitudes F^2 .

intensities I :

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

The Fourier transform of the structure factors (with phases) is the electron density function. The Fourier transform of the structure factor amplitudes (as measure, without the phases) is the Patterson function. The unit cell of the Patterson map is the same as that of the crystal structure.

The Patterson function has some very interesting features:

It consists of peaks, that have coordinates (u, v, w) and an intensity

The distance of each peak from the origin corresponds to an interatomic distance in real space.

The height of the peaks is proportional the involved electrons and proportional the multiplicity of the corresponding vector in real space.

These features make it possible to use the Patterson map to solve the phase problem

The Patterson Function

That means, a peak u, v, w in the Patterson map indicates that atoms exist in the unit cell at x_1, y_1, z_1 and x_2, y_2, z_2 such that

$$u = x_1 - x_2$$

$$v = y_1 - y_2$$

$$w = z_1 - z_2$$

Thus, for N atoms in a unit cell, the Patterson map will contain N^2 peaks. N out of these N^2 peaks will be of zero length from each atom to itself. All these zero vectors will fall on top of one another in the origin of the Patterson map, which generates a very high zero-peak.

Taking this into account, the Patterson map contains $N^2 - N + 1$ theoretically distinguishable peaks.

In real life many of those peaks, which are relatively broad, will overlap and be, in fact, not distinguishable from some neighbor peaks. Usually, the strongest peaks, corresponding to vectors between heavy atoms, are well resolved and usable.

The Patterson Function

As mentioned the relative height of a Patterson peak is proportional the atom number of the two atoms involved and also proportional the multiplicity of the corresponding distance:

$$H \propto m \cdot Z_i \cdot Z_j$$

The origin peak (its height is $H_{000} = \sum Z_i^2$) is usually arbitrarily scaled to 999. That means the height of the other peaks is calculated to

$$H = \frac{999 \cdot m \cdot Z_i \cdot Z_j}{\sum Z_i^2}$$

The sum is over all atoms in the unit cell.

One Heavy Atom in $P-1$

For every atom at x, y, z there is a symmetry equivalent atom at $-x, -y, -z$ and hence a Patterson peak at $2x, 2y, 2z$.

The compound $C_{32}H_{24}AuF_5P_2$ crystallizes in $P-1$ with two molecules per unit cell (that makes one per asymmetric unit). The two gold atoms are related by the inversion center. Besides the peak on the origin (height 999) there is a peak with height 374 at $u = 0.318 = 2x$, $v = 0.471 = 2y$, $w = 0.532 = 2z$, which is much higher than all the other peaks (≤ 145). Its height is consistent with the calculated value for a Au-Au vector (377).

To calculate the positions x, y, z of the gold atom in the unit cell, we can divide each of the Patterson coordinates by 2. However, we also need to take into account the fact that there is always a peak at $u+1, v+1$ and $w+1$ (corresponding to next unit cell in the crystal)! Thus:

$$x = 0.318/2 = \mathbf{0.159} \text{ or } 1.318/2 = \mathbf{0.659}$$

$$y = 0.471/2 = \mathbf{0.236} \text{ or } 1.471/2 = \mathbf{0.736}$$

$$z = 0.532/2 = \mathbf{0.266} \text{ or } 1.532/2 = \mathbf{0.766}$$

Those eight truly equivalent solutions correspond to the eight possible positions of the inversion center in the space group $P-1$.

Two Independent Heavy Atoms in $P-1$

The compound $[\text{C}_{24}\text{H}_{20}\text{S}_4\text{Ag}]^+ [\text{AsF}_6]^-$ crystallizes in $P-1$ with a unit cell volume of 1407 \AA^3 . This corresponds to two formula units per unit cell (one per asymmetric unit). We have two heavy atoms per formula unit (As and Ag), corresponding to the coordinates x_1, y_1, z_1 and x_2, y_2, z_2 . Their symmetry equivalents as generated by the inversion center in $P-1$ are at $-x_1, -y_1, -z_1$ and at $-x_2, -y_2, -z_2$. Let's generate all possible difference vectors between those four atoms (4 X 4 table).

	x_1, y_1, z_1	$-x_1, -y_1, -z_1$	x_2, y_2, z_2	$-x_2, -y_2, -z_2$
x_1, y_1, z_1	$0, 0, 0$	$-2x_1, -2y_1, -2z_1$	$x_2 - x_1, y_2 - y_1, z_2 - z_1$	$-x_1 - x_2, -y_1 - y_2, -z_1 - z_2$
$-x_1, -y_1, -z_1$	$2x_1, 2y_1, 2z_1$	$0, 0, 0$	$x_1 + x_2, y_1 + y_2, z_1 + z_2$	$x_1 - x_2, y_1 - y_2, z_1 - z_2$
x_2, y_2, z_2	$x_1 - x_2, y_1 - y_2, z_1 - z_2$	$-x_1 - x_2, -y_1 - y_2, -z_1 - z_2$	$0, 0, 0$	$-2x_2, -2y_2, -2z_2$
$-x_2, -y_2, -z_2$	$x_1 + x_2, y_1 + y_2, z_1 + z_2$	$x_2 - x_1, y_2 - y_1, z_2 - z_1$	$2x_2, 2y_2, 2z_2$	$0, 0, 0$

The **mixed peaks** have $m=2$, the **zero peaks** have $m=4$.

Courtesy of George M. Sheldrick. Used with permission.

Two Independent Heavy Atoms in $P-1$

$[\text{C}_{24}\text{H}_{20}\text{S}_4\text{Ag}]^+ [\text{AsF}_6]^-$ in $P-1$ with two formula units per unit cell. As and Ag at x_1, y_1, z_1 and x_2, y_2, z_2 as well as $-x_1, -y_1, -z_1$ and $-x_2, -y_2, -z_2$. Difference vectors in 4 X 4 table.

	x_1, y_1, z_1	$-x_1, -y_1, -z_1$	x_2, y_2, z_2	$-x_2, -y_2, -z_2$
x_1, y_1, z_1	0, 0, 0	$-2x_1, -2y_1, -2z_1$	$x_2-x_1, y_2-y_1, z_2-z_1$	$-x_1-x_2, -y_1-y_2, -z_1-z_2$
$-x_1, -y_1, -z_1$	$2x_1, 2y_1, 2z_1$	0, 0, 0	$x_1+x_2, y_1+y_2, z_1+z_2$	$x_1-x_2, y_1-y_2, z_1-z_2$
x_2, y_2, z_2	$x_1-x_2, y_1-y_2, z_1-z_2$	$-x_1-x_2, -y_1-y_2, -z_1-z_2$	0, 0, 0	$-2x_2, -2y_2, -2z_2$
$-x_2, -y_2, -z_2$	$x_1+x_2, y_1+y_2, z_1+z_2$	$x_2-x_1, y_2-y_1, z_2-z_1$	$2x_2, 2y_2, 2z_2$	0, 0, 0

The mixed peaks have $m=2$, the zero peaks have $m=4$.

Courtesy of George M. Sheldrick. Used with permission.

Calculating the peak heights: $Z(\text{Ag}) = 47$, $Z(\text{As}) = 33$, $\sum Z^2 = 11384$.

With $H = 999 m Z_i Z_j / \sum Z^2$:

Ag—As $m = 2$ height = 272;

Ag—Ag $m = 1$ height = 194;

As—As $m = 1$ height = 96.

Two Independent Heavy Atoms in $P-1$

Ag—As: height 272; Ag—Ag: height 194; As—As: height 96

#	u	v	w	height	explanation
1	0	0	0	999	Origin
2	0.765	0.187	0.974	310	$x(\text{Ag})+x(\text{As})$
3	0.392	0.099	0.325	301	$x(\text{Ag})-x(\text{As})$
4	0.159	0.285	0.298	250	$2x(\text{Ag})$
..
..
14	0.364	0.077	0.639	102	$2x(\text{As})$

One consistent solution is:

Courtesy of George M. Sheldrick. Used with permission.

Ag@ $x = 0.080, y = 0.143, z = 0.149$

As@ $x = 0.682, y = 0.039, z = 0.820$

There are 8 equivalent solutions for the first atom (Ag); you can divide one of the $2x, 2y, 2z$ peaks by 2. The second atom (As) needs to be consistent with the first one. You can subtract the coordinates of the first atom from one of the cross peaks. Always check whether your solution explains all peaks!

Two Independent Heavy Atoms in $P2_1$

Two heavy atoms, corresponding to the coordinates x_1, y_1, z_1 and x_2, y_2, z_2 . Their symmetry equivalents as generated by the 2_1 axis are at $-x_1, \frac{1}{2}+y_1, -z_1$ and at $-x_2, \frac{1}{2}+y_2, -z_2$. The corresponding 4 X 4 table:

	x_1, y_1, z_1	x_2, y_2, z_2	$-x_1, \frac{1}{2}+y_1, -z_1$	$-x_2, \frac{1}{2}+y_2, -z_2$
x_1, y_1, z_1	0, 0, 0	$-x_1+x_2, -y_1+y_2, -z_1+z_2$	$-2x_1, \frac{1}{2}, -2z_1$	$-x_1-x_2, \frac{1}{2}-y_1+y_2, -z_1-z_2$
x_2, y_2, z_2	$x_1-x_2, y_1-y_2, z_1-z_2$	0, 0, 0	$-x_1-x_2, \frac{1}{2}+y_1-y_2, -z_1-z_2$	$-2x_2, \frac{1}{2}, -2z_2$
$-x_1, \frac{1}{2}+y_1, -z_1$	$2x_1, \frac{1}{2}, 2z_1$	$x_1+x_2, \frac{1}{2}-y_1+y_2, z_1+z_2$	0, 0, 0	$x_1-x_2, -y_1+y_2, z_1-z_2$
$-x_2, \frac{1}{2}+y_2, -z_2$	$x_1+x_2, \frac{1}{2}+y_1-y_2, z_1+z_2$	$2x_2, \frac{1}{2}, 2z_2$	$-x_1+x_2, y_1-y_2, -z_1+z_2$	0, 0, 0

$+\frac{1}{2}$ and $-\frac{1}{2}$ are equivalent!

0 0 0	$m = 4$	origin
$\pm\{ 2x_1, \frac{1}{2}, 2z_1 \}$	$m = 1$	Harker-section at $y = \frac{1}{2}$
$\pm\{ 2x_2, \frac{1}{2}, 2z_2 \}$	$m = 1$	
$\pm\{ x_1-x_2, y_1-y_2, z_1-z_2 \}$	$m = 1$	cross vectors
$\pm\{ x_1-x_2, -y_1+y_2, z_1-z_2 \}$	$m = 1$	
$\pm\{ x_1+x_2, \frac{1}{2}+y_1-y_2, z_1+z_2 \}$	$m = 1$	
$\pm\{ x_1+x_2, \frac{1}{2}-y_1+y_2, z_1+z_2 \}$	$m = 1$	

Symmetry of the Patterson Function

For every vector $i \rightarrow j$ there is always a vector $j \rightarrow i$. Thus the Patterson is always centrosymmetric.

In general, the symmetry of the Patterson is determined by the symmetry of the diffraction pattern (also centrosymmetric). Glide planes and screw axes of the space groups correspond to mirror planes and normal rotation axes in reciprocal space. The Patterson has the same symmetry as the Laue group.

Harker Sections

Space group symmetry leads to accumulation of Patterson peaks in certain sections (planes or lines). *E.g.*

$P2_1$: atoms at x, y, z and $-x, \frac{1}{2}+y, -z$. \rightarrow Eigenvectors at $2x, \frac{1}{2}, 2z$;
Harker section at $v = \frac{1}{2}$.

$P2$: atoms at x, y, z and $-x, y, -z$. \rightarrow Eigenvectors at $2x, 0, 2z$;
Harker section at $v = 0$.

Pm : atoms at x, y, z and $x, -y, z$. \rightarrow Eigenvectors at $0, 2y, 0$;
Harker section at $u = 0, w = 0$.

Space groups $P2$ und Pm both have the same systematic absences (none), but they have different Harker sections.

Problems of the Patterson Method

Frequently localizing a single heavy atom in a structure is enough to find all other atoms by means of iterative difference Fourier synthesis with amplitudes $|F_o - F_c|$ and phases ϕ_c .

Problems arise in non-centrosymmetric space groups when the heavy atom substructure is centrosymmetric (e.g. only one heavy atom). In this case, all ϕ_c -values derived from the heavy atom positions are 0° or 180° and the calculated electron density is centrosymmetric, *i.e.* it shows a double-image.

Another problem is the atom-type assignment: Electron density and Patterson peak heights are proportional to atomic number, however only roughly and isoelectronic species are notoriously difficult to distinguish. This is a problem also with direct methods and sometimes even during structure refinement.

Direct Methods

Direct methods determine the phases directly from the diffraction pattern without any knowledge about the nature of the sample.

Several statistical equations relate the phase of a reflection to its intensity and the intensity and phase of other reflections in the dataset. This can be done with a certain probability. Many probability relations together with fast computers make it possible to determine the phases of many measured reflections with some accuracy.

Direct methods usually do not use structure factors but operate in E -space (remember the normalized structure factors from the E^2-1 statistics?). The advantage is, that the intensity of a normalized structure factor does not decline with the resolution; and direct methods assume atoms to be point-scatterers anyway.

Direct methods are used predominantly as “black-box” methods and we don’t really have the time to change that here.