CRYSTALLOGRAPHY

Crystal Structure Analysis

This course is a brief introduction to X-ray crystallography. It aims to help you understand:

- what the method can do for you;
- how it works;
- what its limitations are;
- getting started

X-ray Safety Course

Every one intending to carry out their own work (single crystal or powder) using X-rays must attend the Safety Office lecture course. . “Safe use of X-ray generators, sealed sources and accelerators”

Lecturer: Mark Bradley
Date: Wednesday 12th November 2013 (week 5), 2 - 4 PM
Location: Medical Sciences Teaching Centre.

This course is mandatory. Please take your university card along for identification purposes.

Health and Safety

- X-rays have the potential to cause serious radiation burns.
- The equipment is safety-interlocked to current UK standards.
- It is a criminal offence to tamper with the interlocks.
- The interlocks are checked weekly.
- **People wishing to use the equipment must be trained and registered.**
- Members of chemistry may **observe** the equipment in use without registering.

Why is crystallography important?

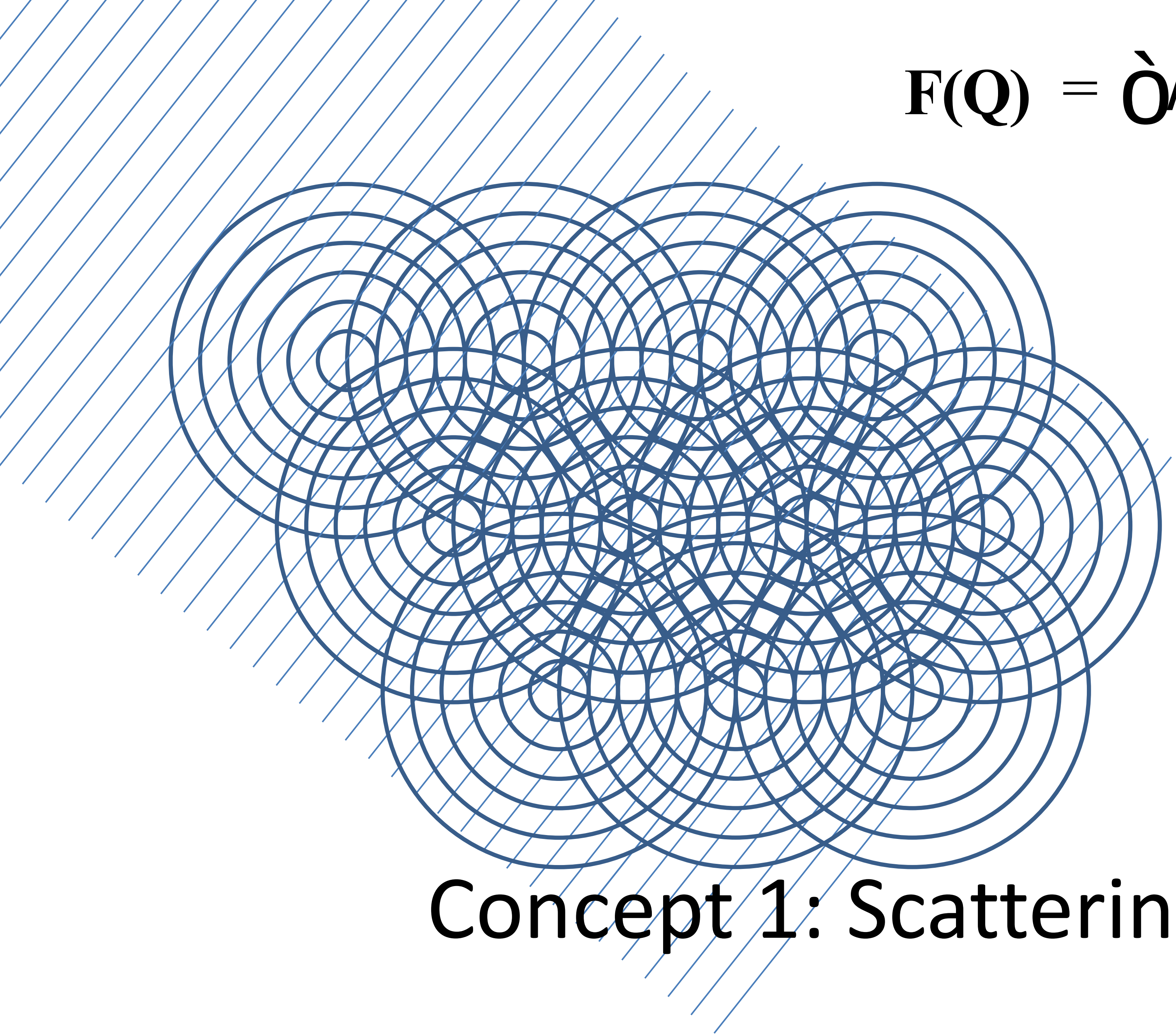
1. Connectivity – what does the material look like.
2. Geometry – what are the bond lengths and angles.
3. Chemical composition – what atoms are present.
4. Relative or absolute stereochemistry.
5. Electron density distribution.

Crystal diffraction in 6 concepts

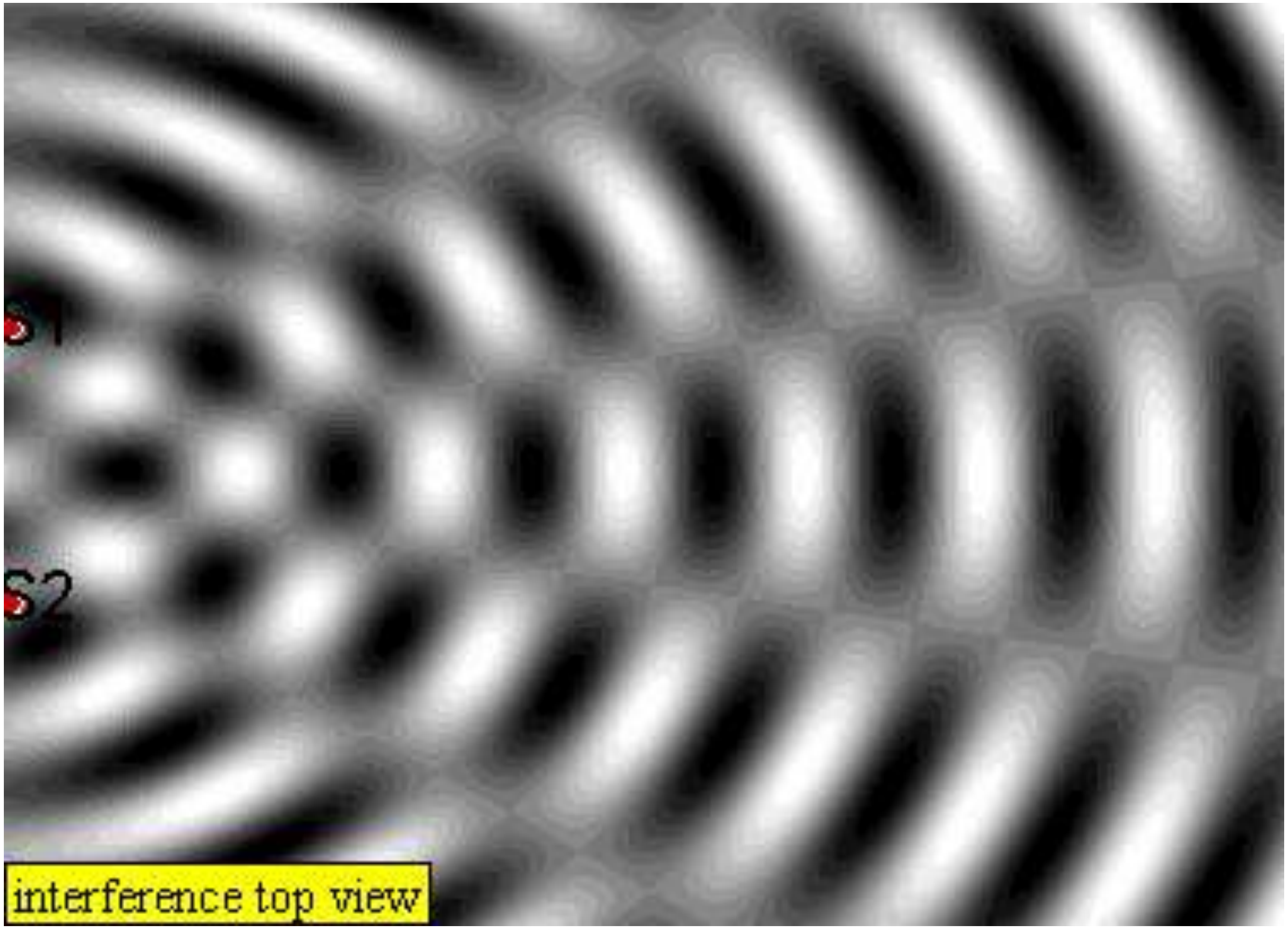
- Scattering
- Crystal
- Time/space average
- Mosaicity
- Twinning
- Reverse Fourier transform

Concept 1: Scattering

$$\mathbf{F(Q)} = \int \rho(\mathbf{r}) \cdot e^{i\mathbf{Q}\cdot\mathbf{r}}\, d\mathbf{r}$$



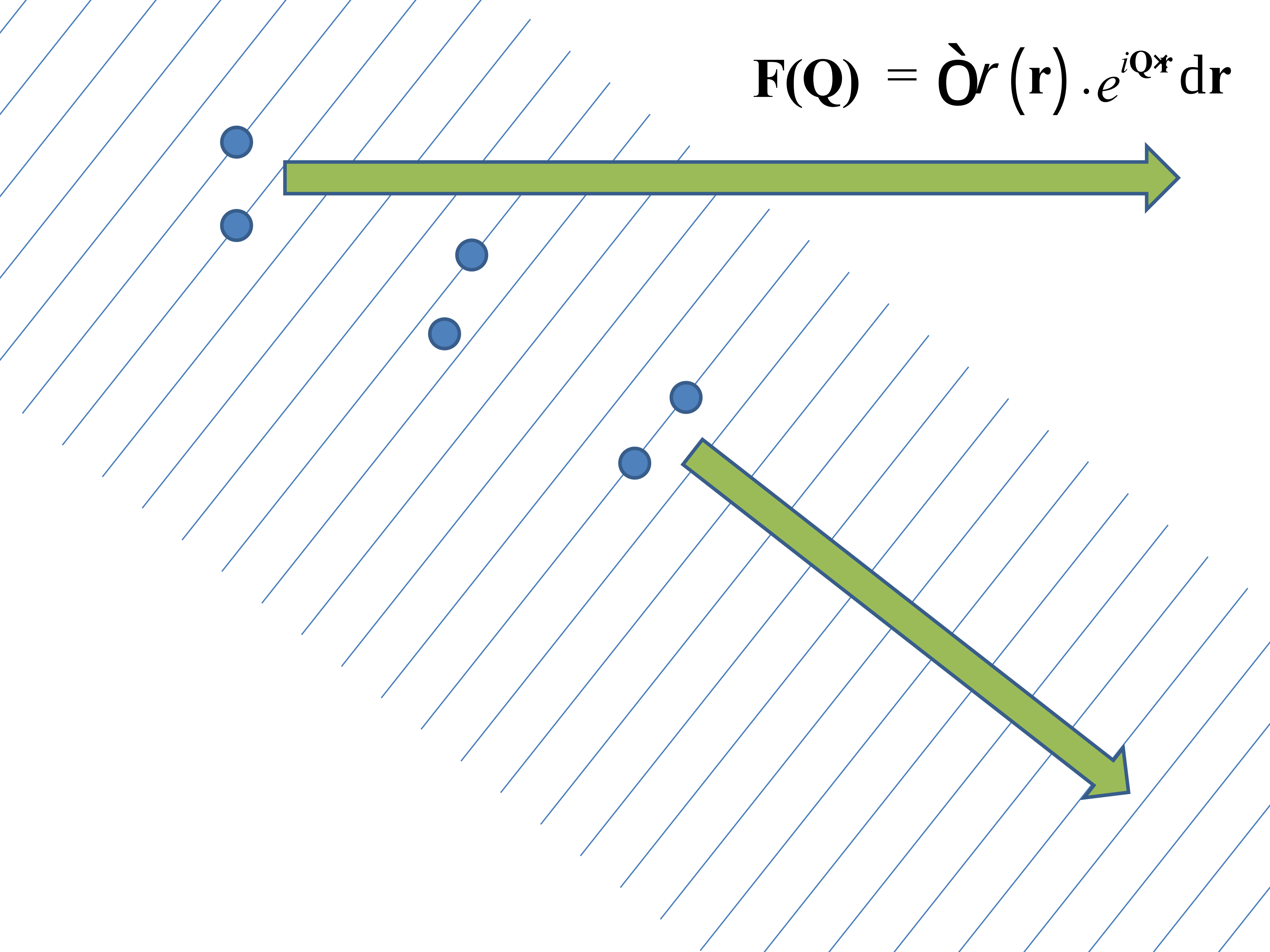
F is the amplitude and phase of the sum of the scattered waves in the direction ‘Q’.



The amplitude of the waves over here is just the sum of the waves from S1 and S2 at a given point, at time t.

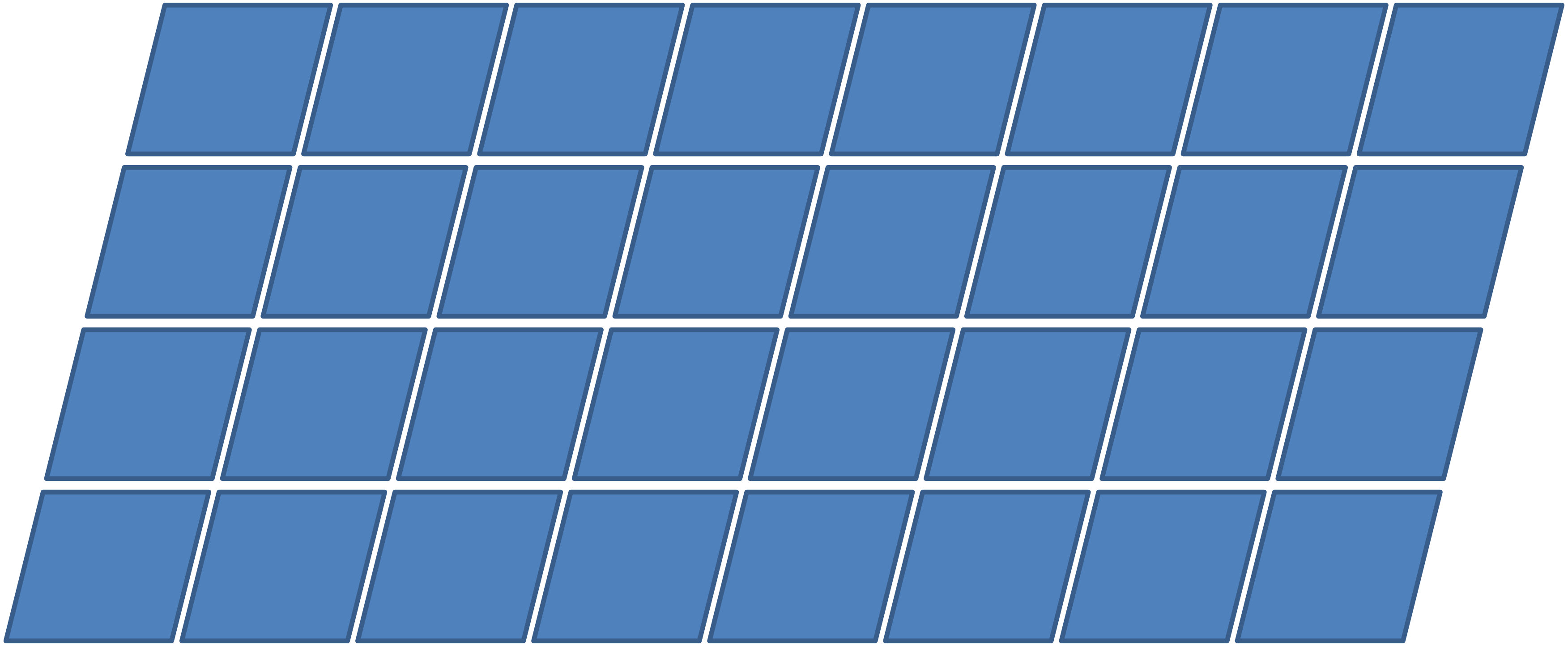
The waves have the same wavelength so the sum varies from ‘in-phase’ to ‘out-of-phase’.

$$\mathbf{F(Q)} = \int \rho(\mathbf{r}) . e^{i\mathbf{Q}\cdot\mathbf{r}} \, d\mathbf{r}$$



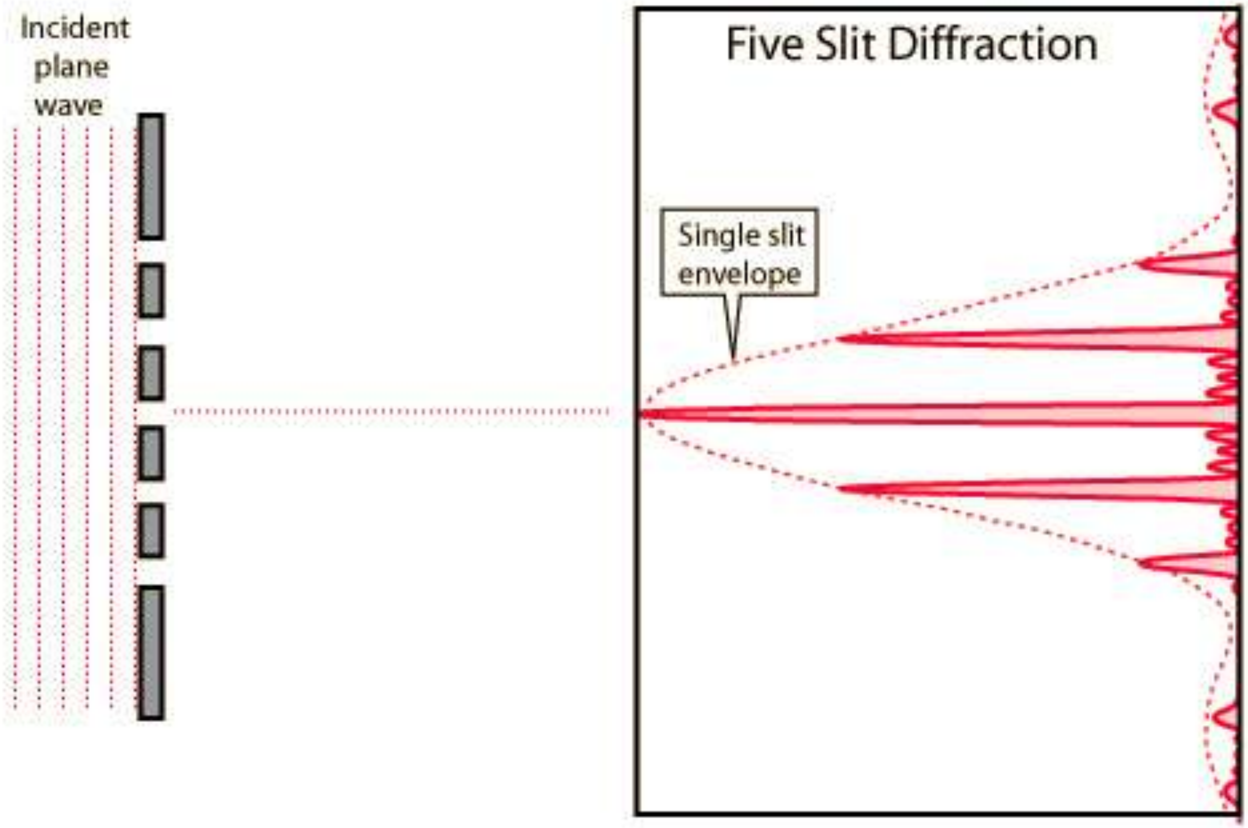
Concept 2: A crystal

$$\mathbf{F}(\mathbf{Q}) \approx \sum_j f_j \cdot e^{i\mathbf{Q}\cdot\mathbf{r}}$$



Q = hkl **r** = xyz

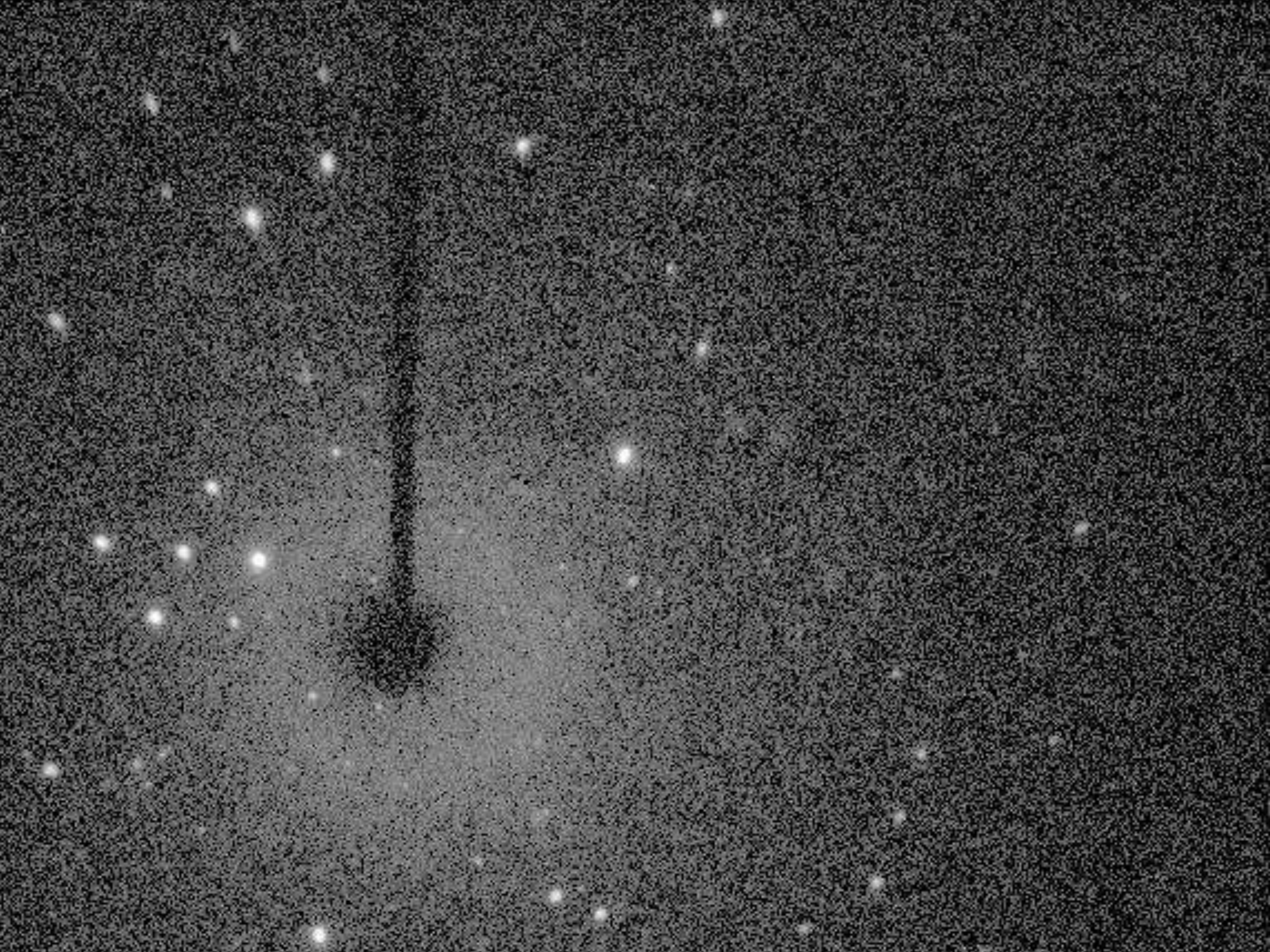
In the case of diffraction from a regular array of ‘things’, the scattering becomes very directional [Bragg peaks] with almost nothing in between.

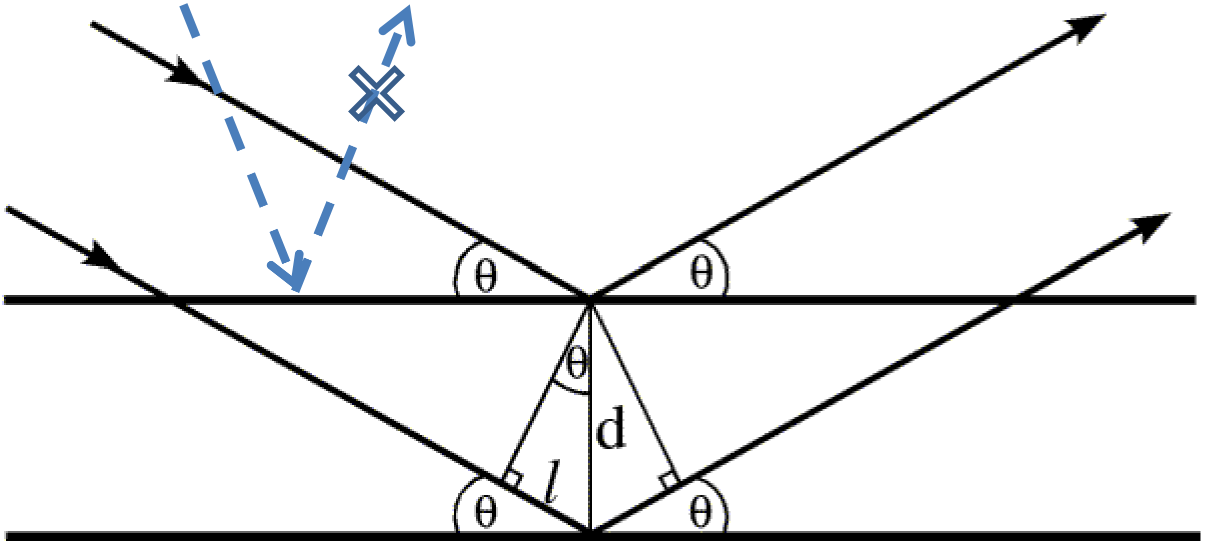


Bragg's Law

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

$$d^2 = \frac{a^2}{h^2 + k^2 + l^2}$$ for cubic crystals



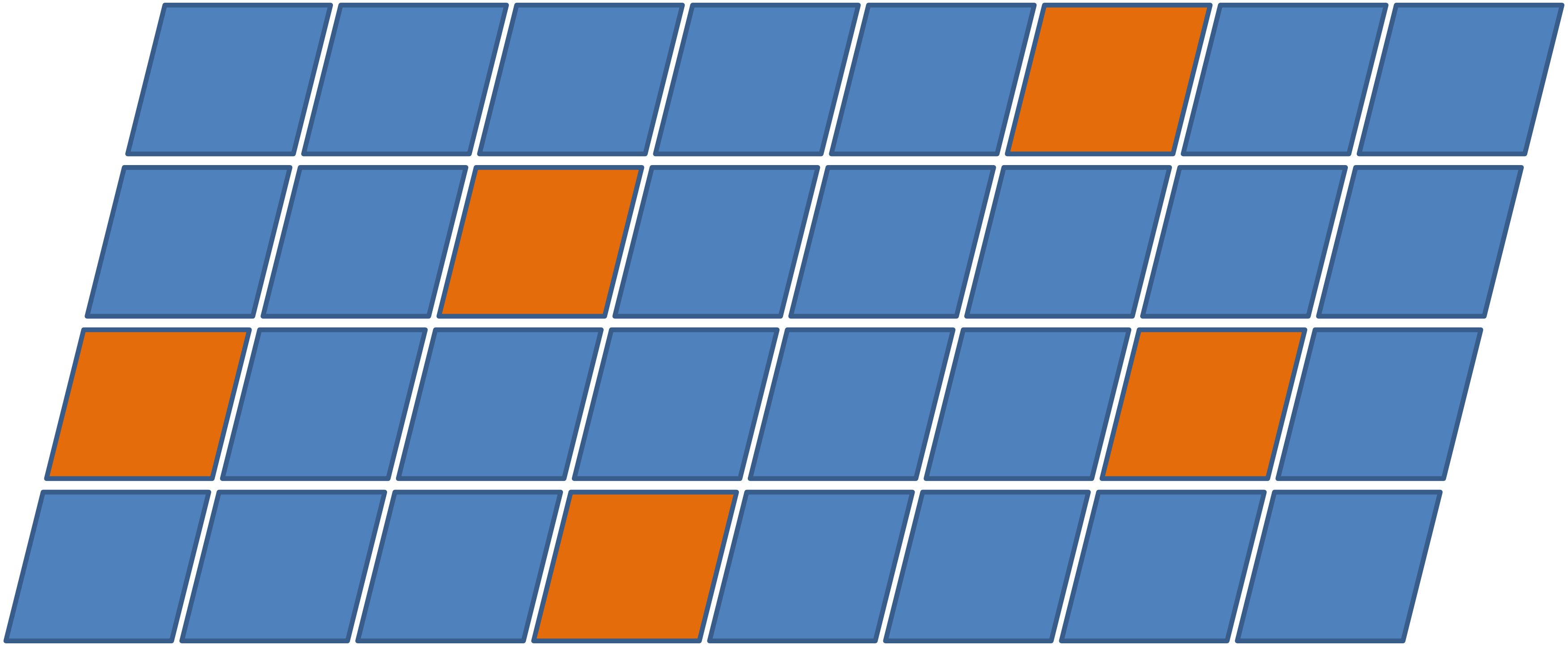


- Incoming and diffracted X-ray beams (solid lines).
- Dashed lines show an incident angle that does not obey Bragg’s law. No diffraction is observed.
- The lattice planes do not act as a crystal mirror, reflecting X-rays!

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

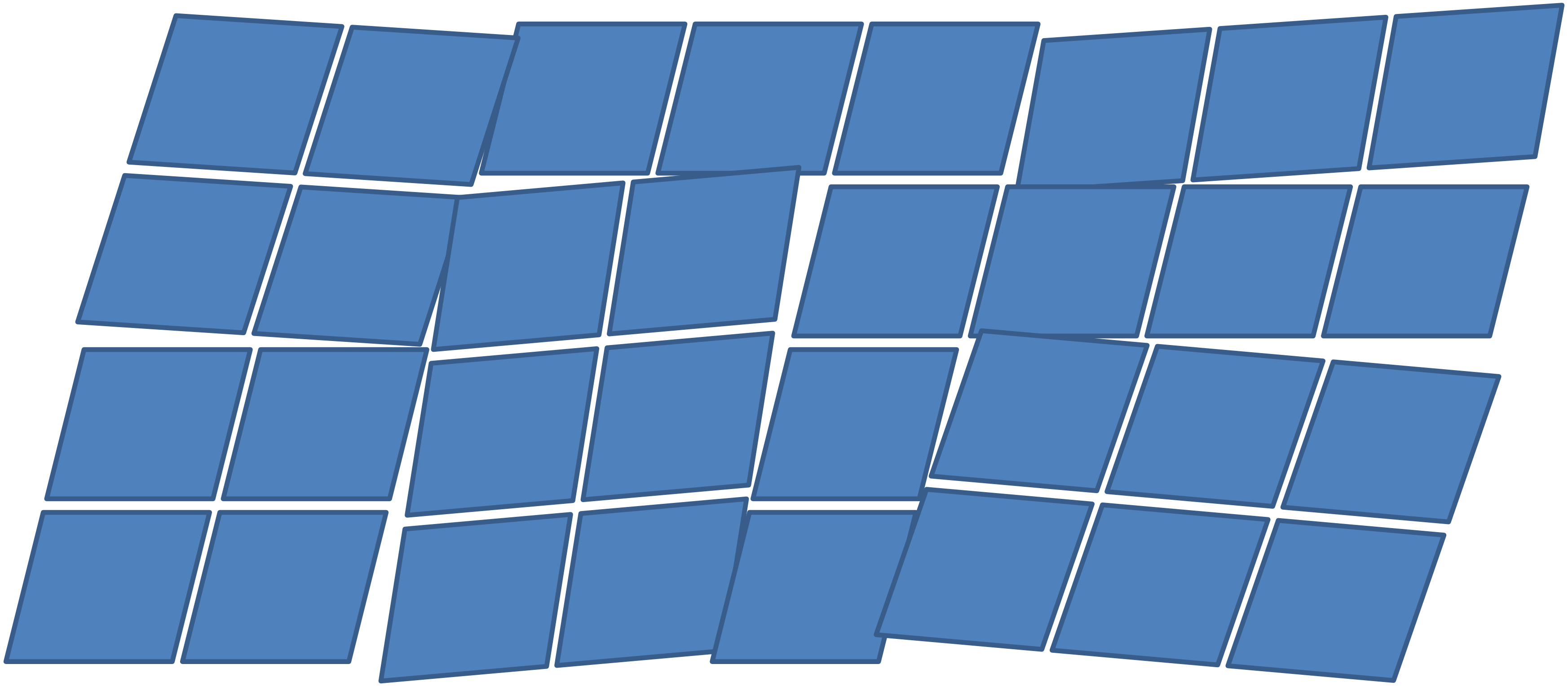
Concept 3: Time/Space Snapshot

$$\mathbf{F}(\mathbf{Q}) \approx \sum_j f_j \cdot e^{i\mathbf{Q}\cdot\mathbf{r}}$$



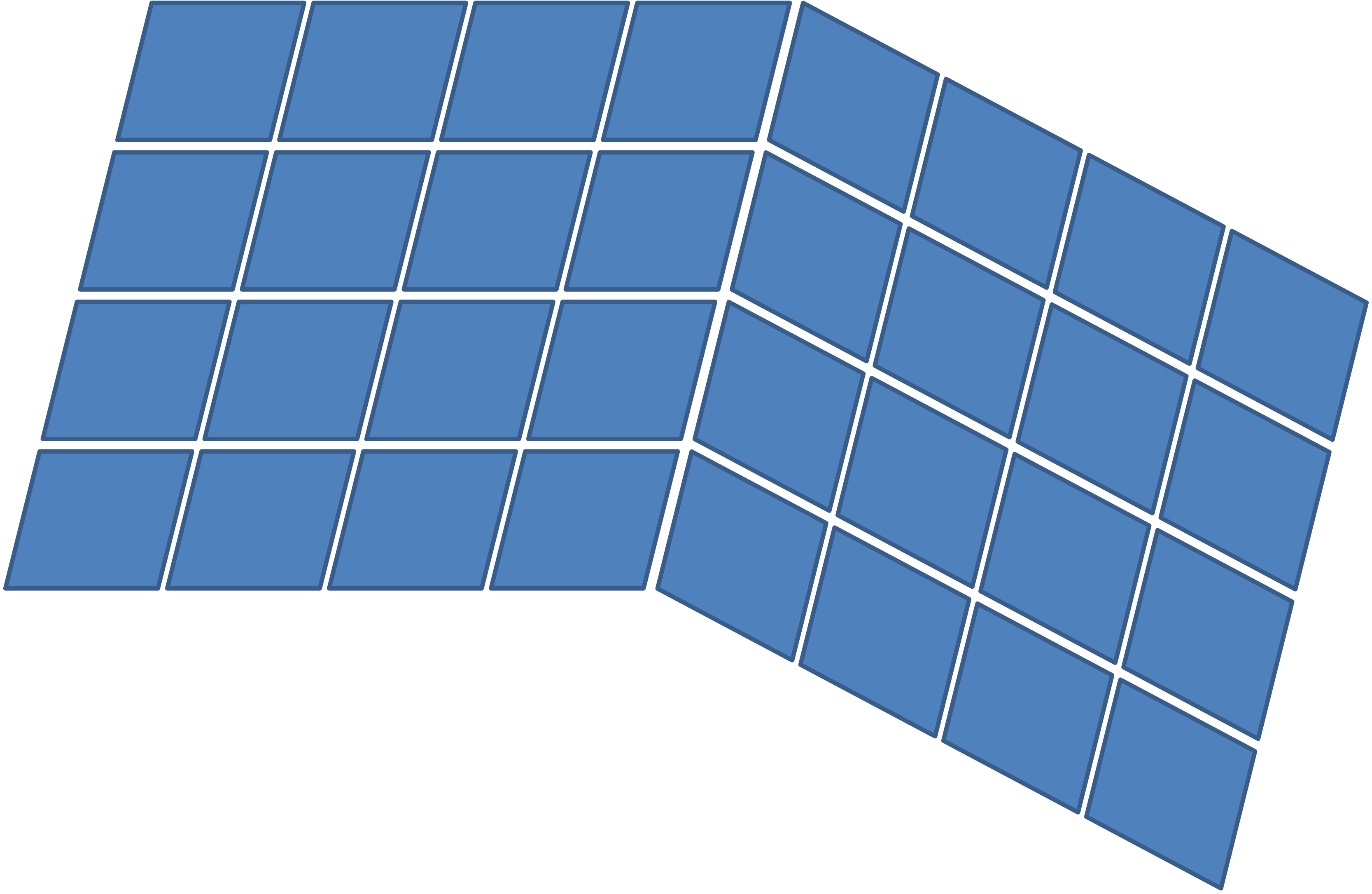
Mosaicity

$$\mathbf{F}^2(\mathbf{Q}) = \mathbf{F}^2(\mathbf{Q})_1 + \mathbf{F}^2(\mathbf{Q})_2 + \dots$$



Concept 5: Twinning

$$\mathbf{F}^2(\mathbf{Q}) = \mathbf{F}^2(\mathbf{Q}) + \mathbf{F}^2(\mathbf{RQ})$$

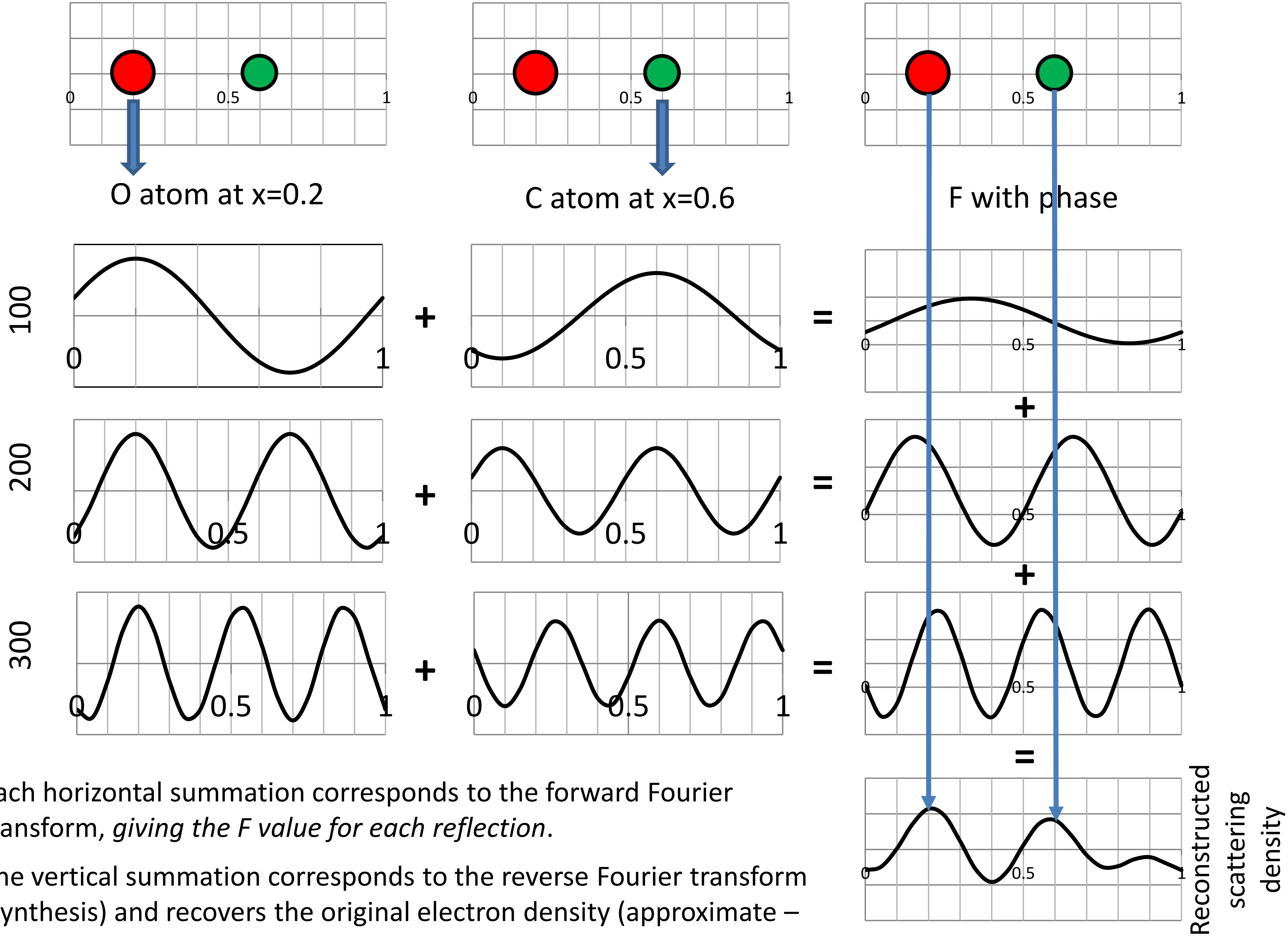


Concept 6: Reverse Fourier Transform

$$\mathbf{F}(\mathbf{Q}) \approx \sum_{j} f_{j} \cdot e^{i\mathbf{Q}\cdot\mathbf{r}}$$

Because the calulation of scattering amplitude is a sum of waves, it is analogous / identical to a Fourier transform:

1. The scattering amplitude is related to *spatial frequency* of electron density in the crystal.
2. There is a **reverse Fourier transform** to compute electron density from the scattering.

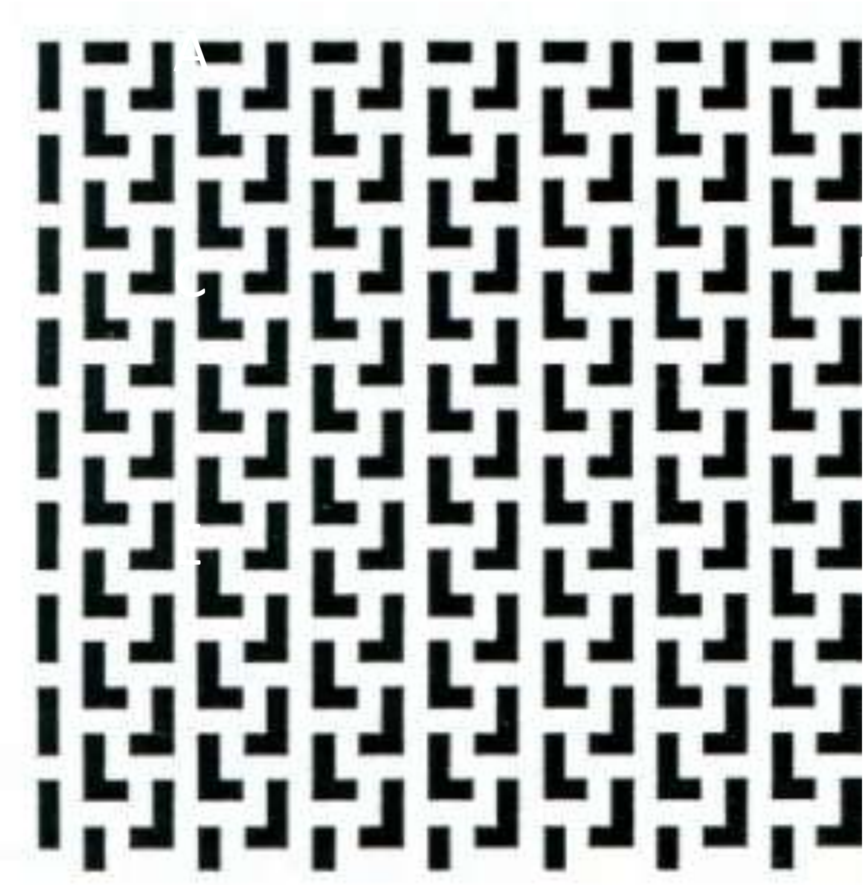


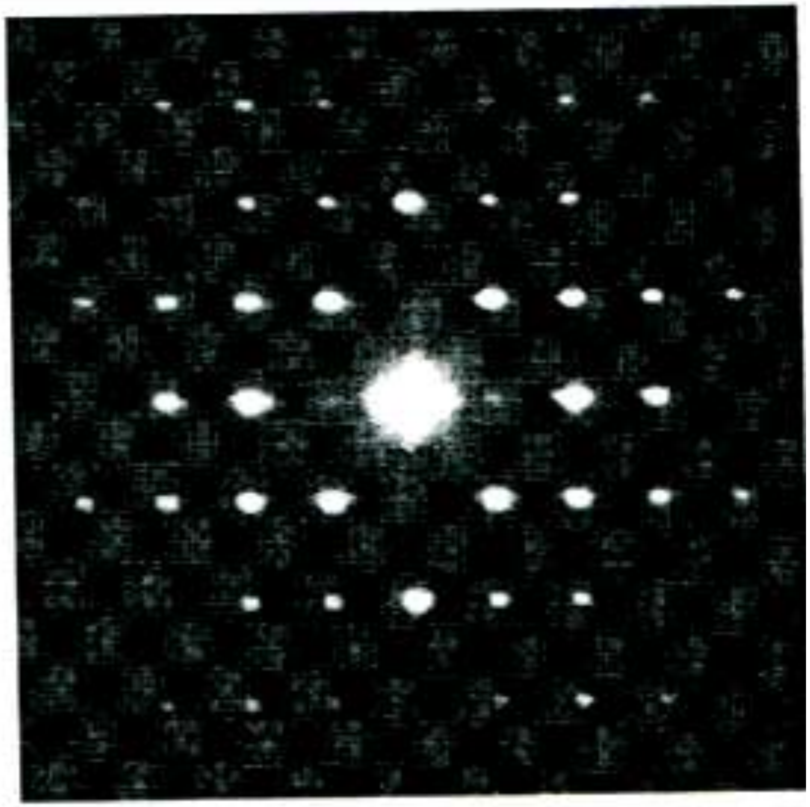
Each horizontal summation corresponds to the forward Fourier transform, *giving the F value for each reflection*.

The vertical summation corresponds to the reverse Fourier transform (synthesis) and recovers the original electron density (approximate – due to only three low resolution waves used).

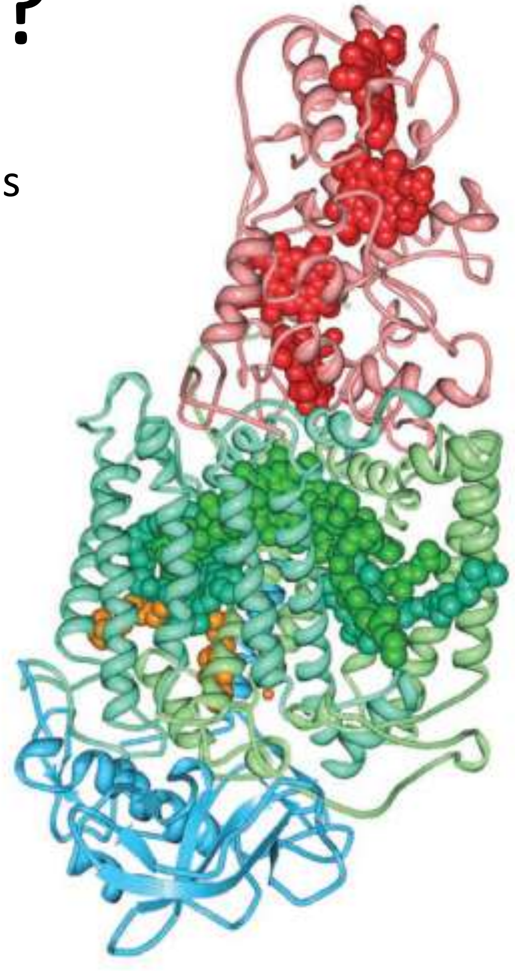
Lost phases

Could you predict the left-hand (2D) pattern from its (2D) diffraction?

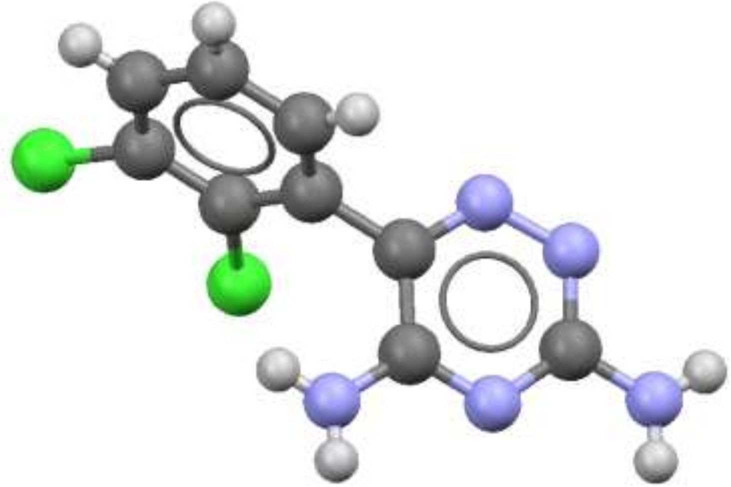




Who uses it today?



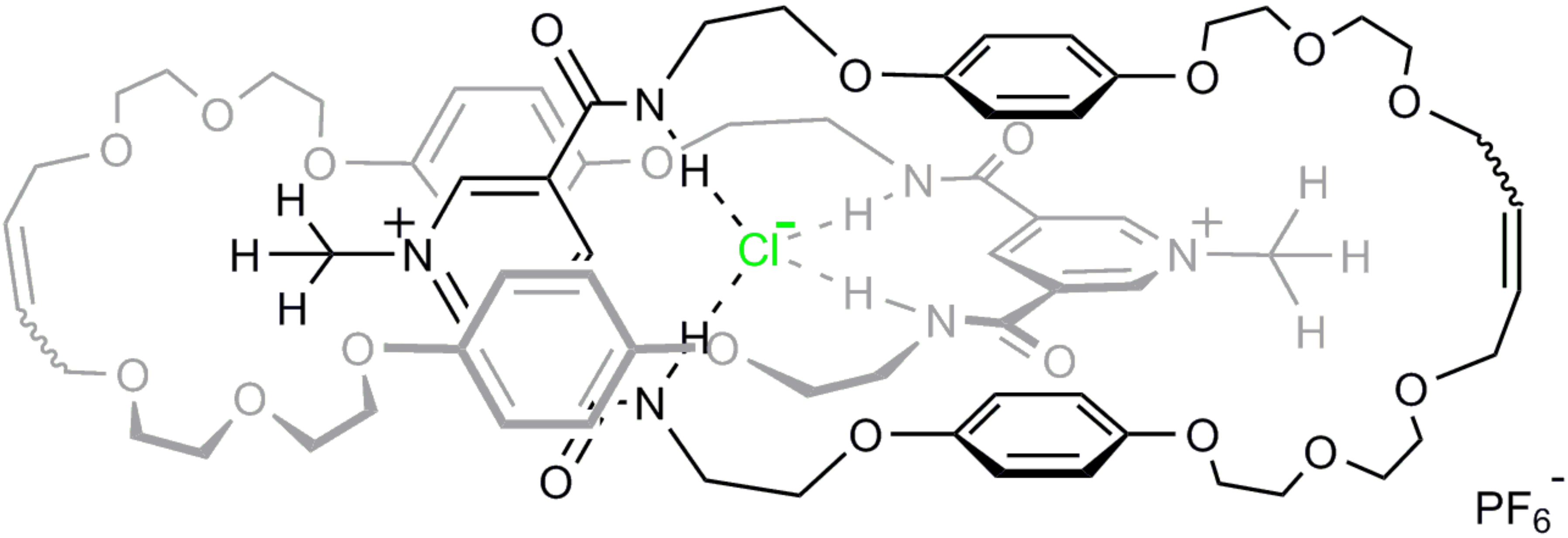
Over 66,000 protein complexes and DNA fragments crystal structures archived by the Protein Databank



592,938 small molecule crystal structures archived by the 2012 Cambridge Structural Database. >40,000 added per year.

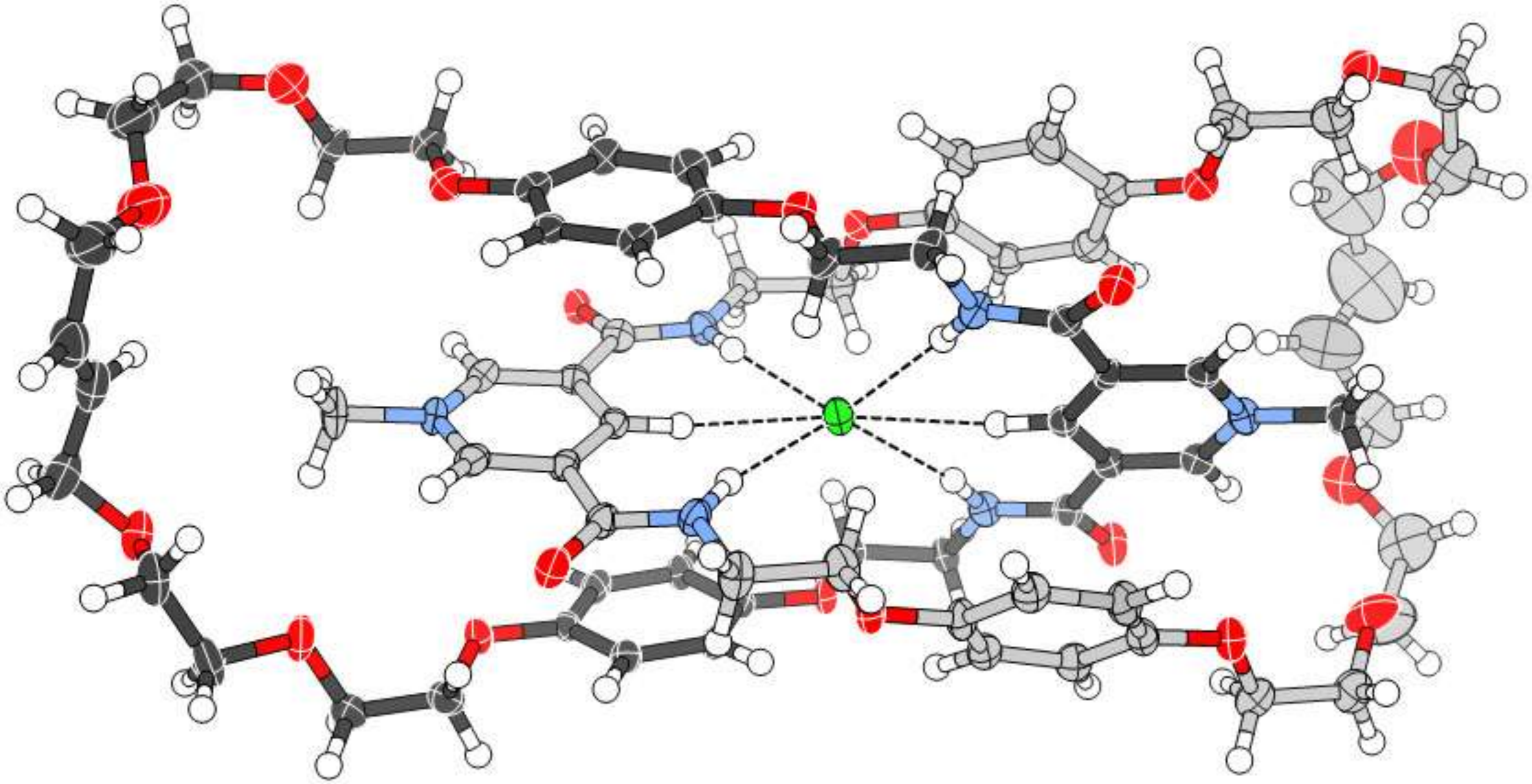
webcsd.ccdc.cam.ac.uk

What do Molecules Look Like?



Paul Beer Group

What do Molecules Look Like?

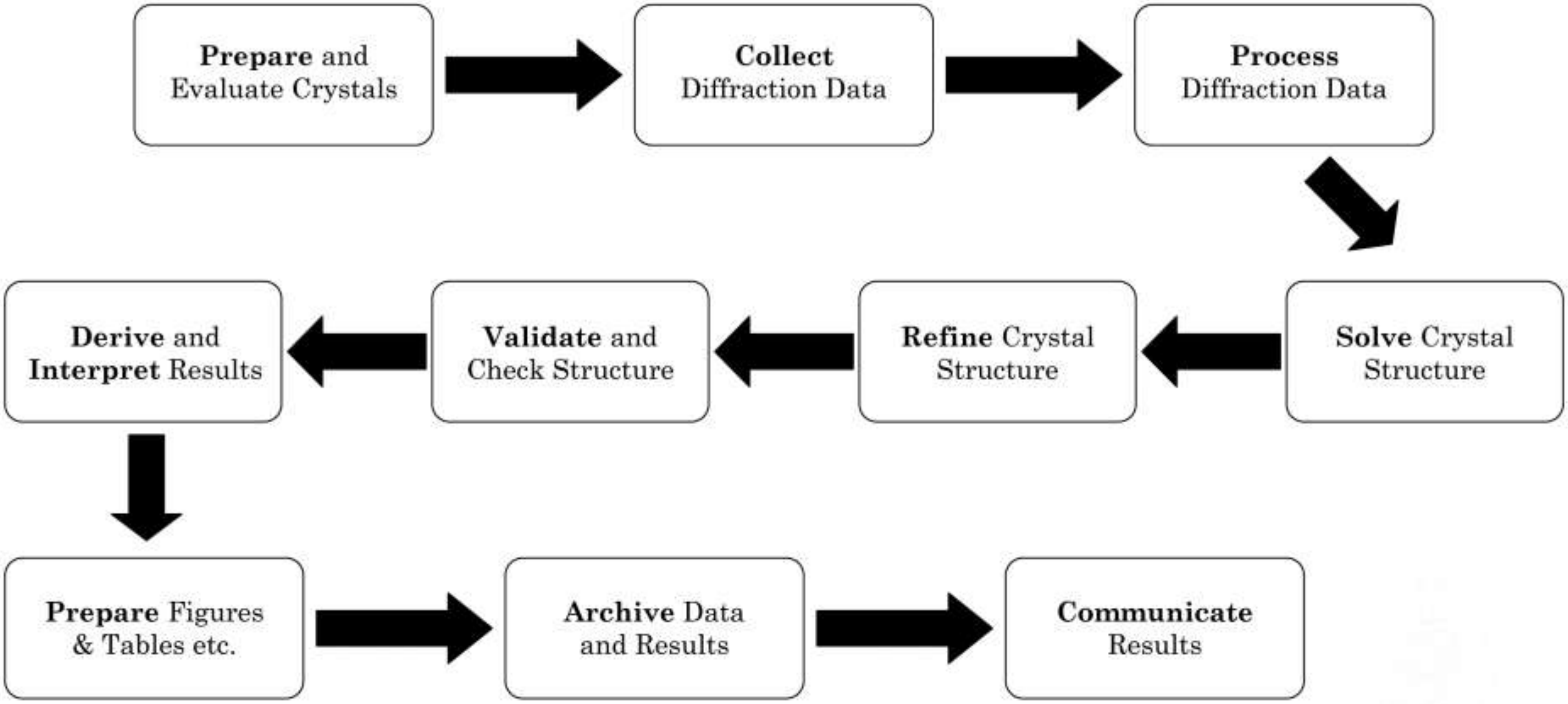


Beer Group Catenane.

The synthesis really did work!

Practical Crystal Structure Analysis

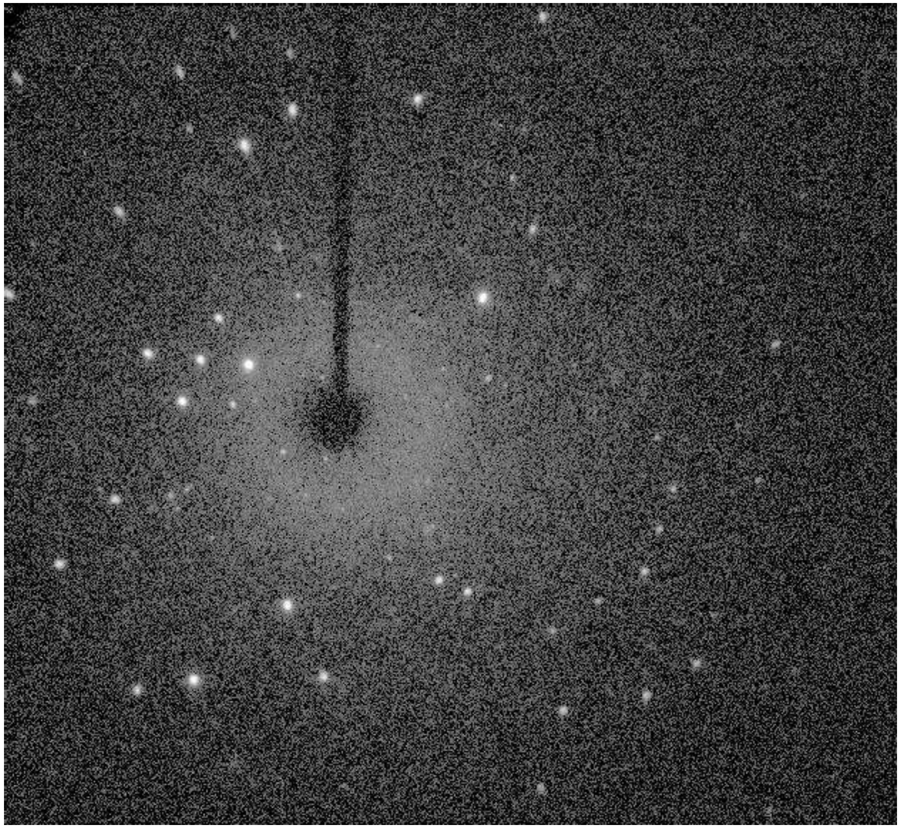
Diagram of a Typical Crystal Structure Analysis

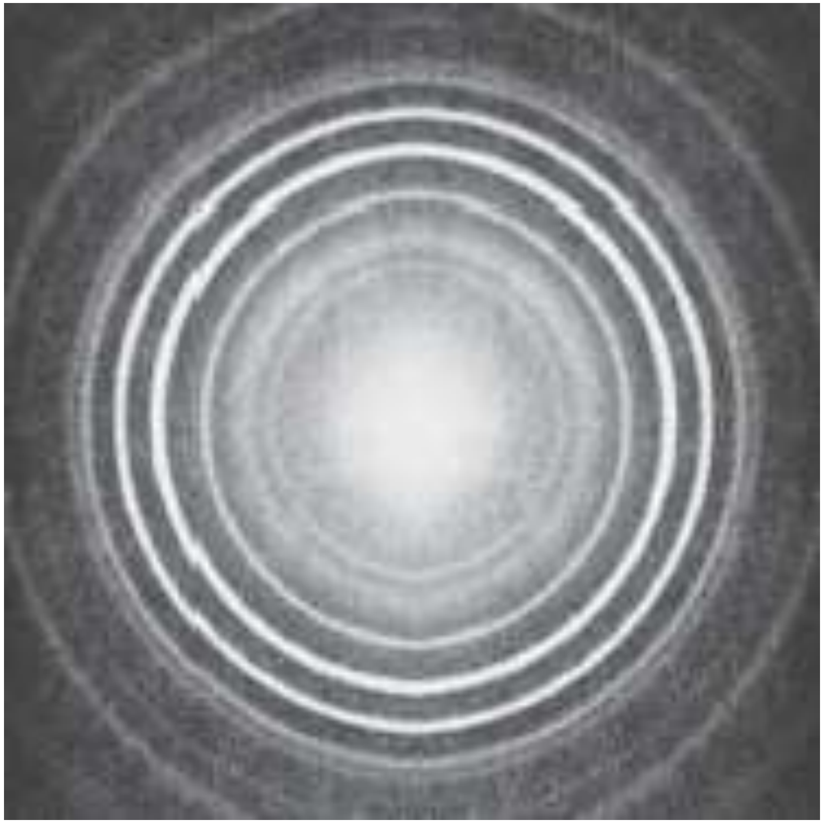


- **Prepare** and Evaluate Crystals
- **Collect** Diffraction Data
- **Process** Diffraction Data
- **Solve** Crystal Structure
- **Refine** Crystal Structure
- **Validate** and Check Structure
- **Derive** and **Interpret** Results
- **Prepare** Figures & Tables etc.
- **Archive** Data and Results
- **Communicate** Results

Diffraction Image

When you expose a crystal to a narrow beam of X-rays, a diffraction pattern is created.





X-rays and Crystals

The crystal both creates an X-ray shadow, but also acts as a 3-D diffraction grating

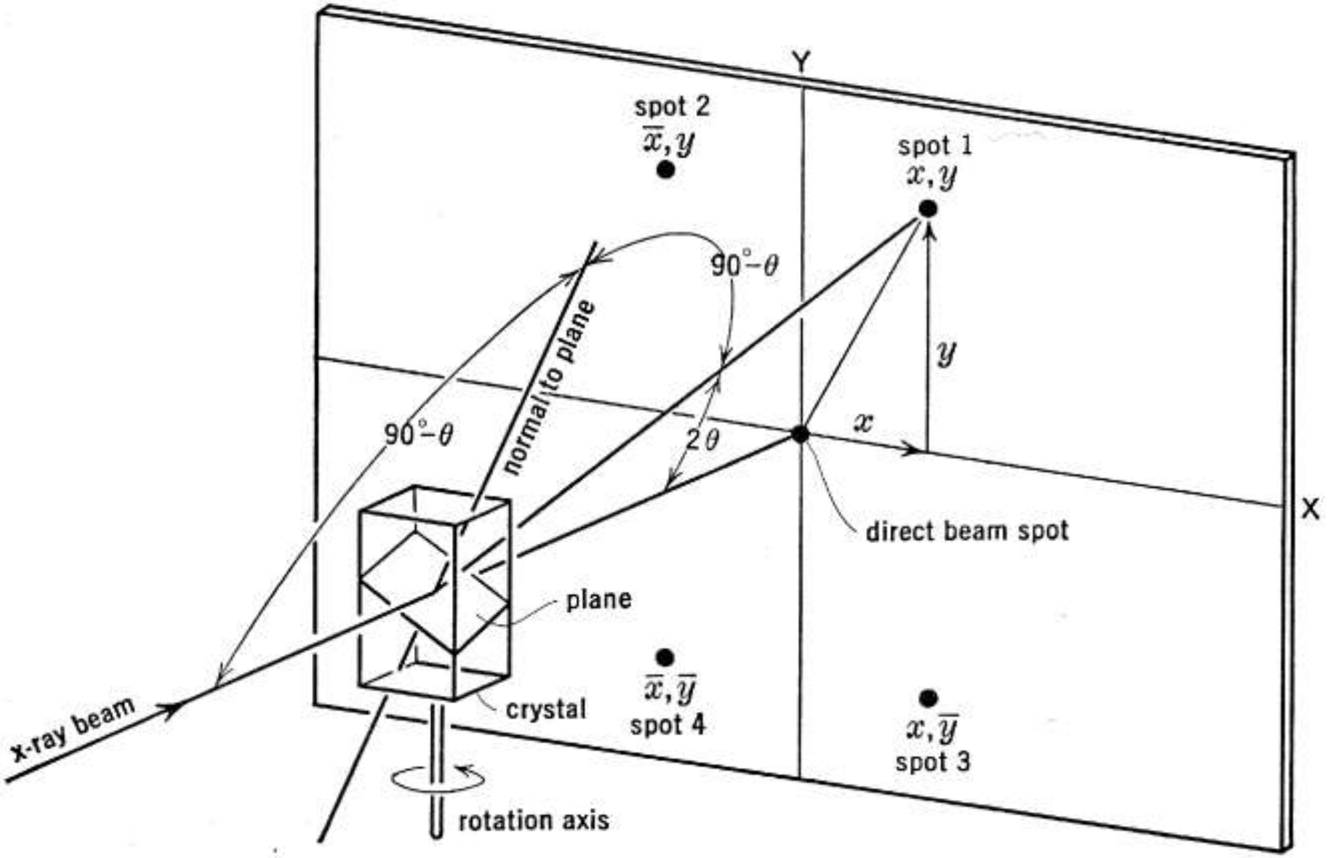
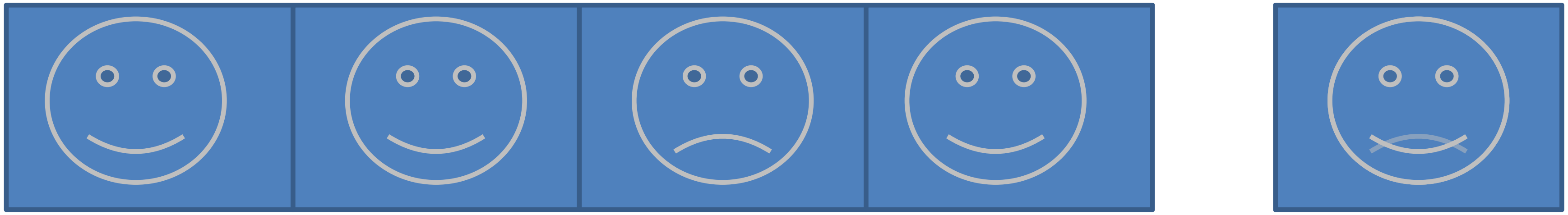


FIG. 57.

What are we trying to find?

X-rays tell us about the *continuous periodic electron distribution* throughout the crystal.

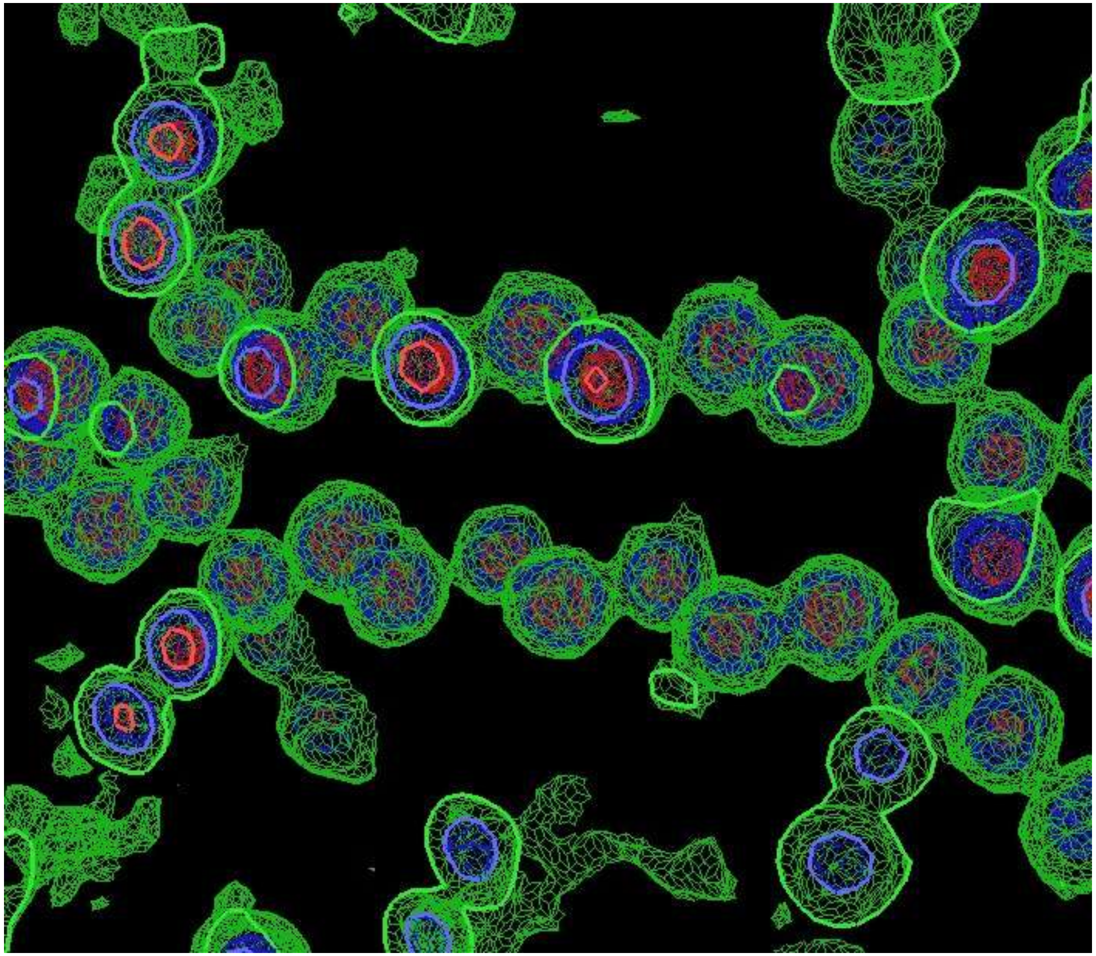
The experiment takes a long time compared to atomic motions, so we get a *time average*.



The experiment requires a large number of regularly ordered unit cells, so we also get a *space average.*

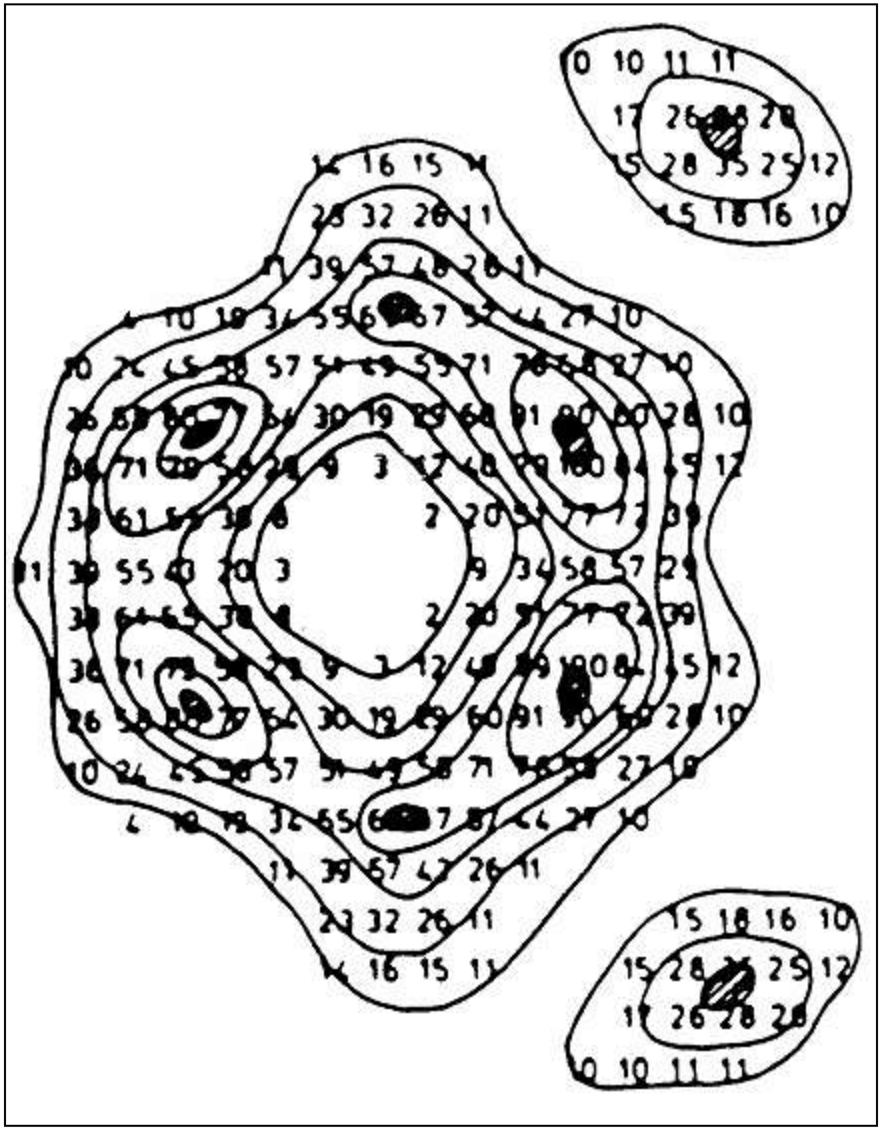
What do we find?

If all goes well, we can create an image of the periodic electron density.



Electron density displayed using Michael Husak's MCE viewer and CRYSTALS

Antique Manual Contouring



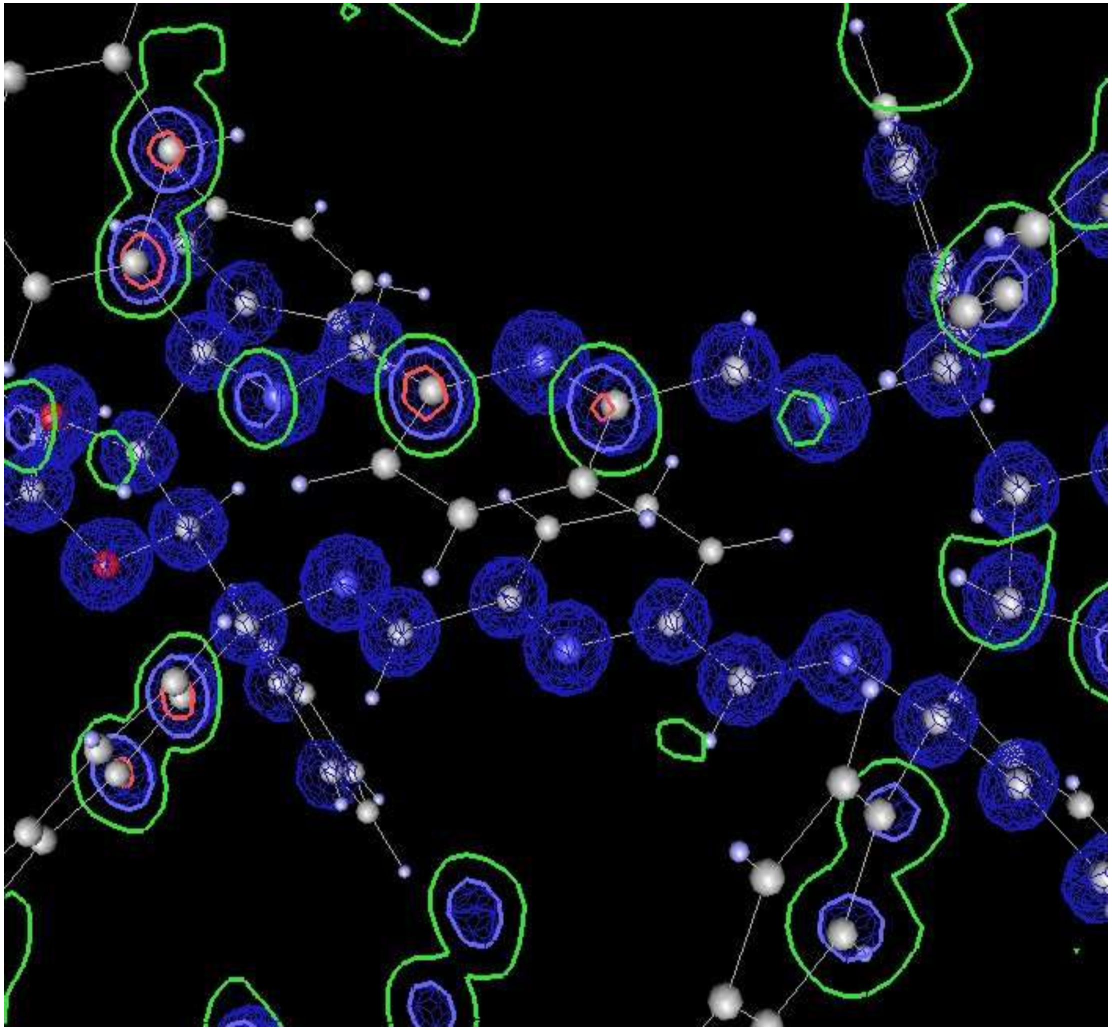
As in geographical contouring (where lines are drawn joining points with the same elevation), in crystallography lines are drawn joining points with the same electron density.

About 360 density points were needed for this diagram

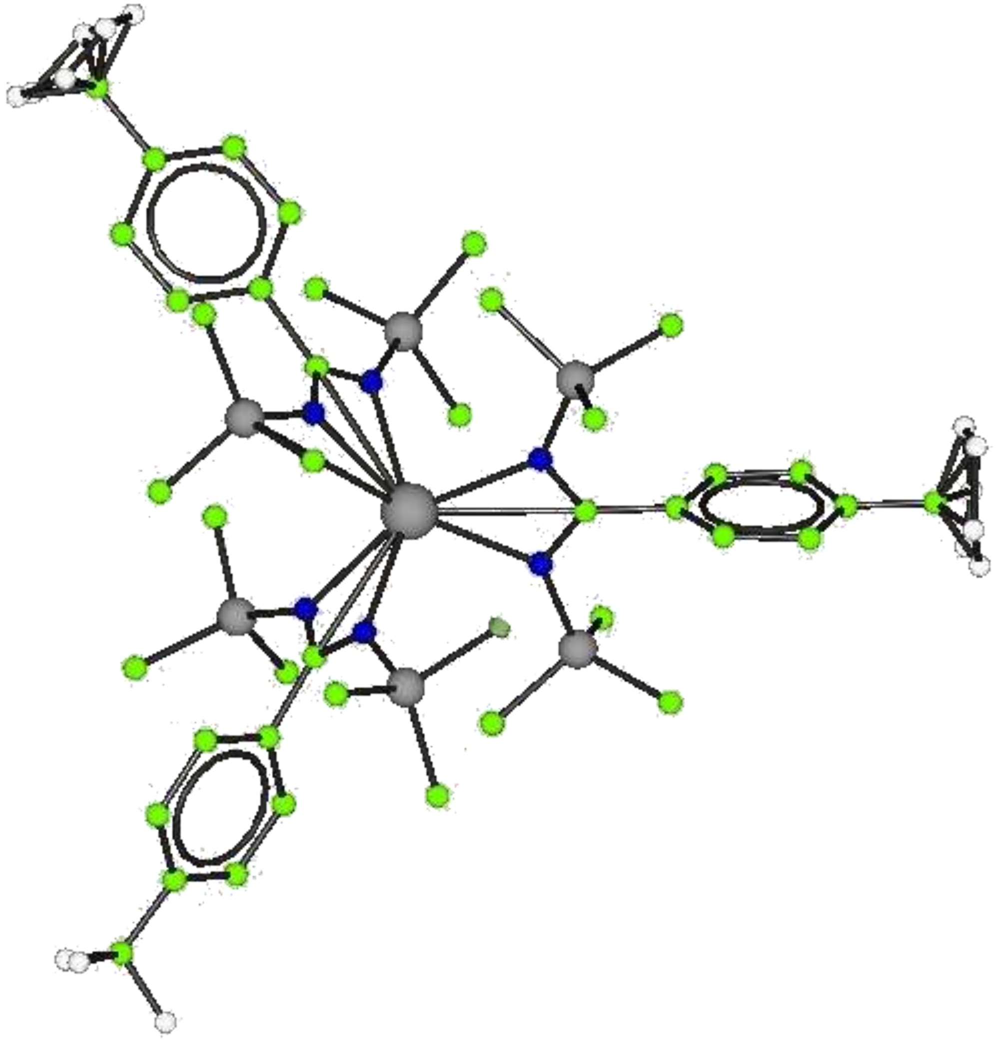
What Next?

Insert atoms into the electron density.

At a resolution of 3 points per Angstrom, a 10x10x10 Å cell would require 27,000 sampling points.

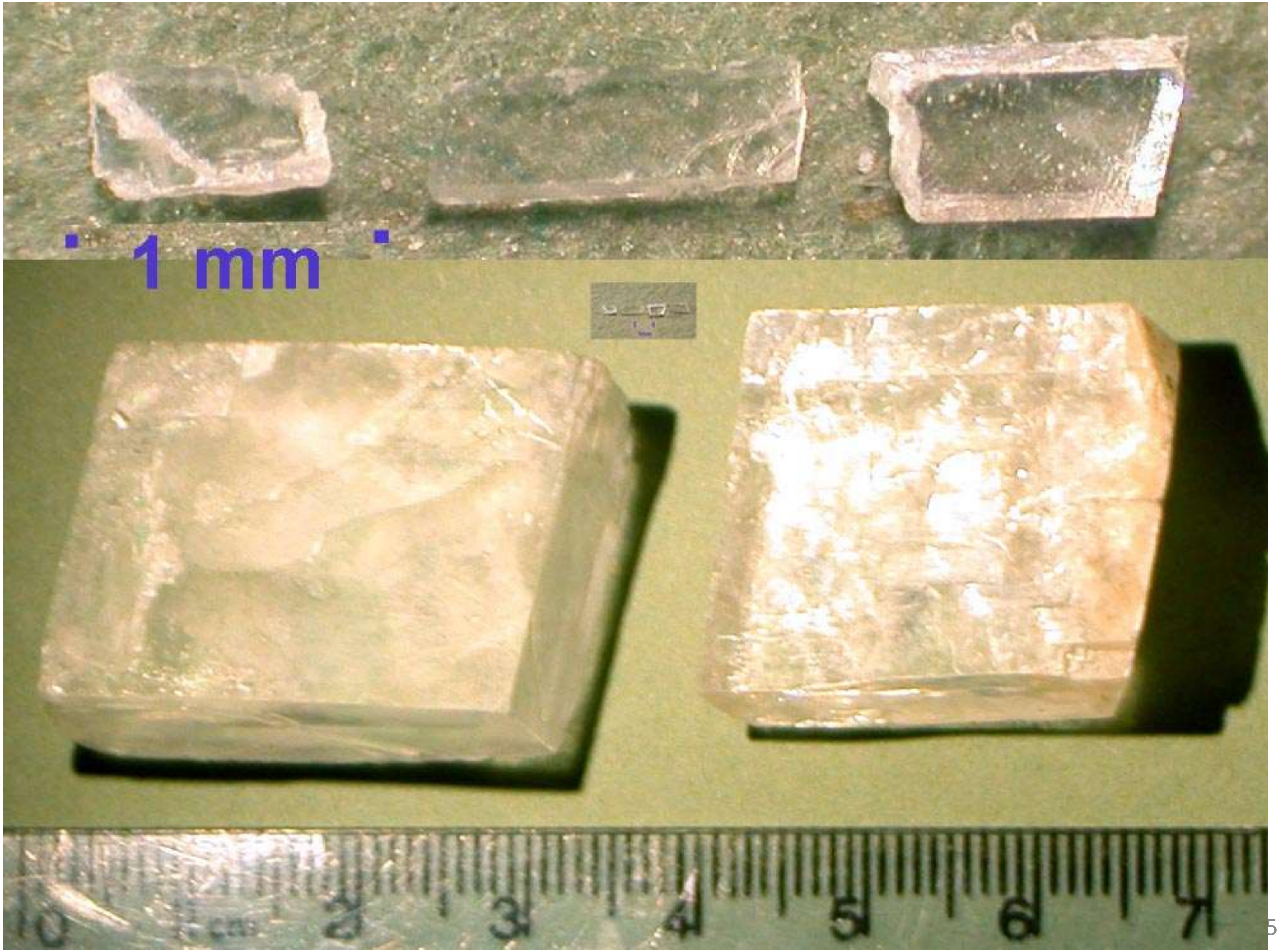


A ‘Crystal Structure’

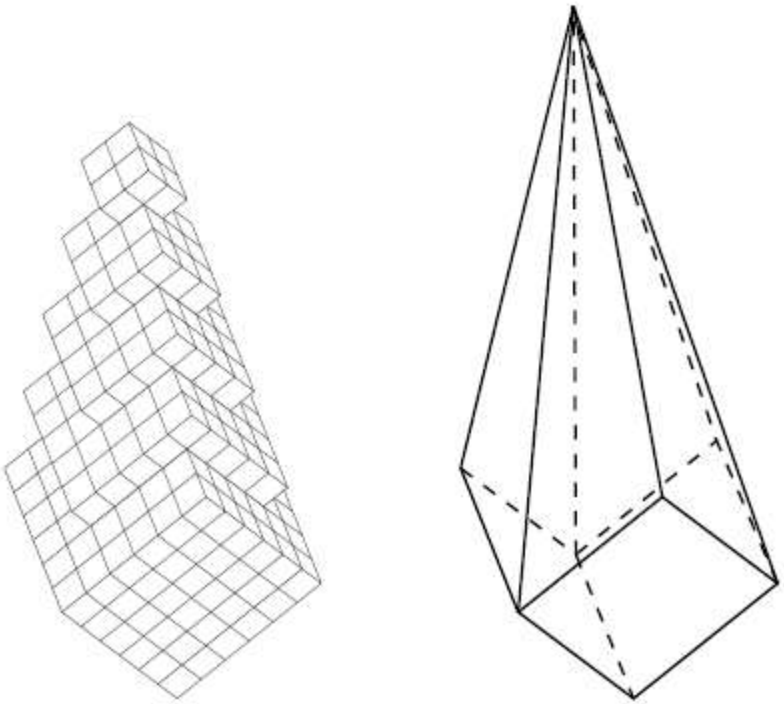


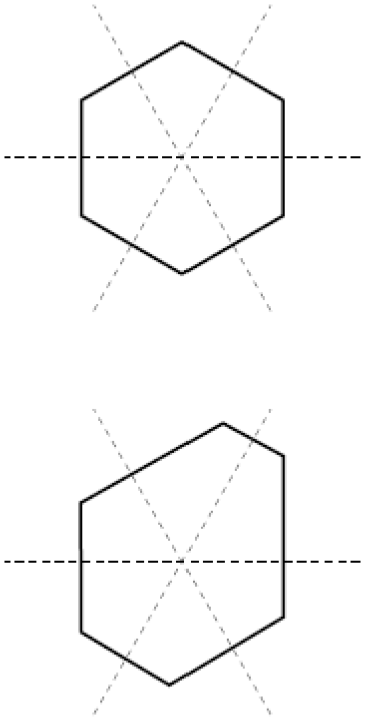
Once the atom sites have been located, the electron density is usually forgotten.

Crystals: Calcite Fragments



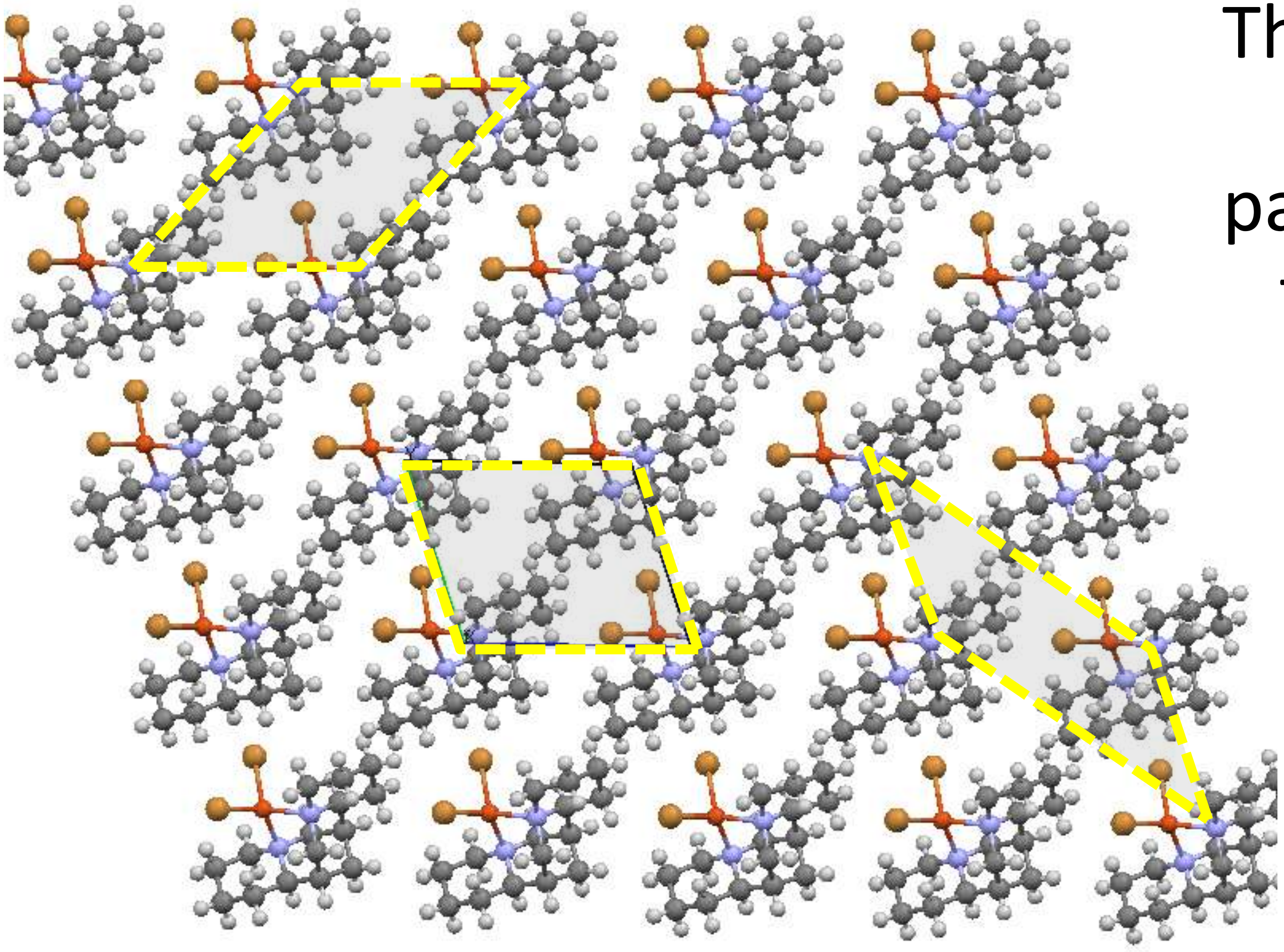
Cleaving Crystals





Constant interfacial angles

The Unit Cell

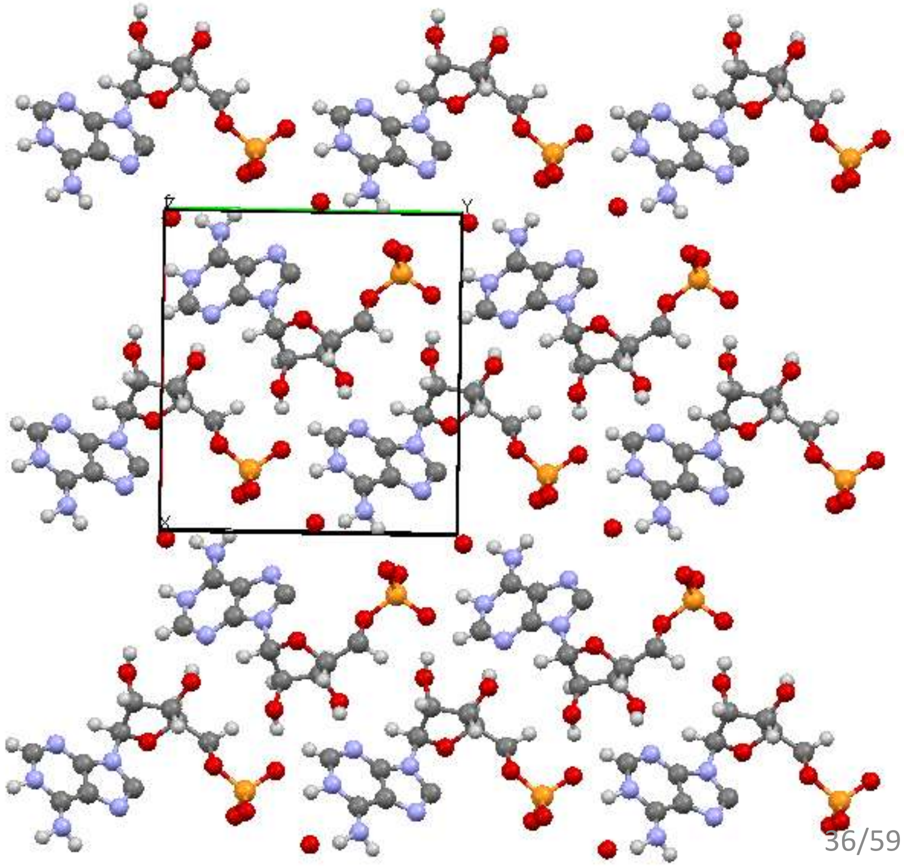


The unit cell is the smallest parallelepiped that displays the full properties of the whole crystal.

Symmetry Operators

Symmetry operations inside the unit cell mean that the contents of the cell can be represented by an asymmetric unit plus a list of symmetry operations.

Though this may seem complicated to a human, nature does it, and it greatly simplifies the computations.



Symmetry in the Unit Cell

The motifs within the cell may be related to each other by operations which are not simple translations.

These operations include:

Centre of Symmetry (inversion)

Rotation

Rotation-translation – called Screw operations

Reflection

Reflection-translation – called Glide operations

The Angstrom

1Å = 0.1 nm　　1Å = 0.0001 micron

1Å = 10^{-8} cm　　1Å = 10^{-10} m

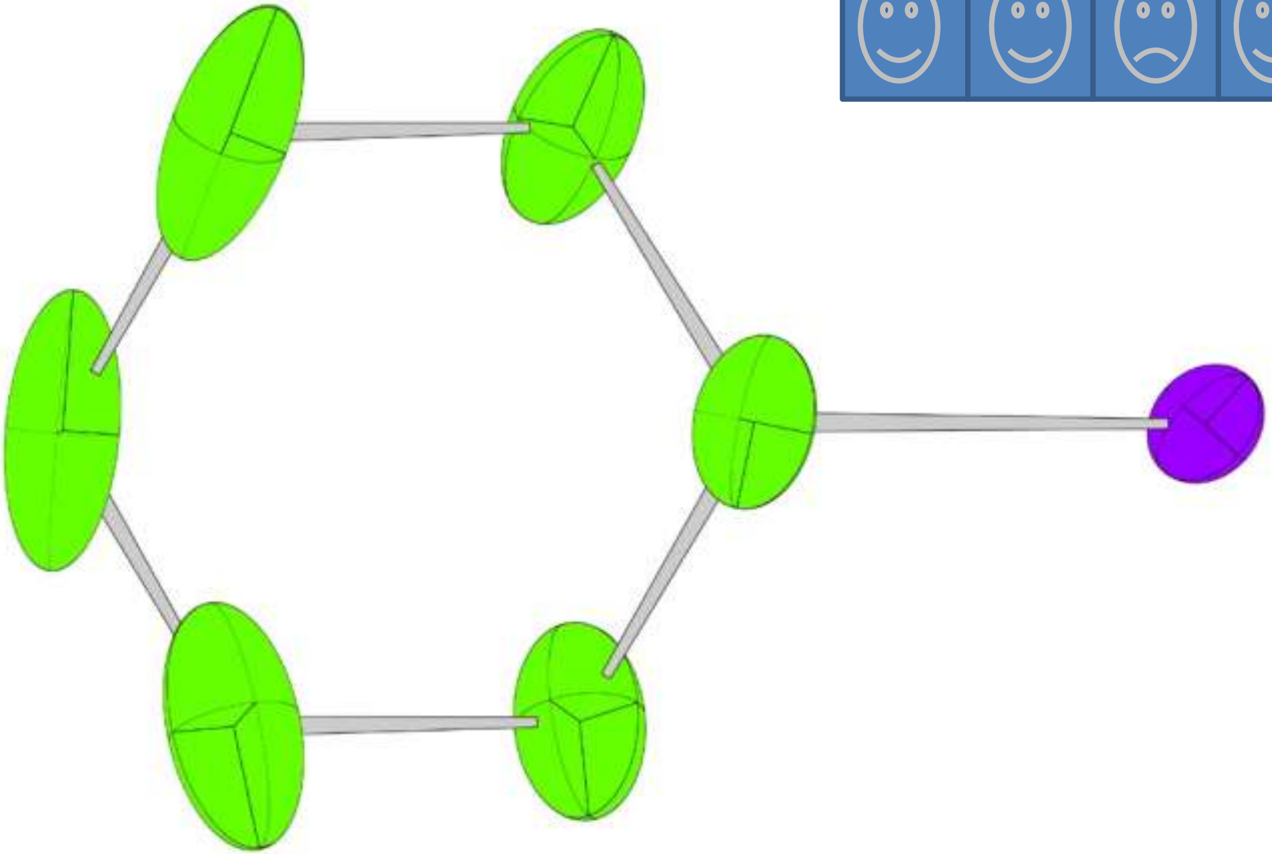
C-C bond lengths are about 1.5 Å.

Wavelength of laboratory X-rays are 0.7 or 1.5 Å

“Its use is officially discouraged by the International Committee for Weights and Measures and is not included in the European Union's catalogue of units of measure that may be used within its Internal Market”

Atomic Displacement Parameter

☺ ☺ ☹ ☺ 😐



Terminology

Cambridge Structural Database; CSD; *a database of all CH-containing crystal structures. Available as desktop installation with powerful search tools, or ‘lite’ version via webcsd.ccdc.cam.ac.uk*

CCDC; *the Cambridge Crystallographic Data Centre, maintainers of the CSD.*

Conquest; *CSD the search software.*

Mercury; *crystal structure visualisation software.*

Terminology

Unit cell *smallest volume (which shows off the space group symmetry elements) which can be used to reproduce the entire crystal by translation alone.*

Asymmetric unit *smallest volume which can reproduce the entire unit cell by application of space group symmetry elements.*

Terminology

hkl; *a triplet of numbers identifying a spot in the diffraction pattern (aka. a reflection).*

xyz; a fractional coordinate within a unit cell

restraint; *an extra equation to include in the fit of a crystallographic model*

Mo ***Kα;*** *characteristic wavelength X-rays produced by a molybdenum target*

e.s.d.'s or **s.u's**; *estimates of precision of parameters in a crystallographic model.*

Terminology

Special postions *positions on symmetry operators which will constrain an atoms position or motion*

General position *any position (x,y,z) which is not a special position. But... "the number of general positions" of a space group is the number of positions generated within a unit cell when an atom is placed at x,y,z. (E.g. In P-1 there are two general positions: x,y,z and -x,-y,-z).*

Terminology

Isotropic or **anisotropic refinement** *the model used for modeling atomic displacement from a mean position. (Simply: anisotropic = ellipsoid shaped probability surface, isotropic = spherical probability surface).*

50% probability level *there is a 50% chance that an atoms' electrons are within this surface on a time/space average.*

Riding atoms *usually hydrogen – atoms placed in a model relative to a connected atom.*

Example paper 1

Bis(5-amino-1H-1,2,4-triazol-3-yl-kN^4)acetato-kO]diaquanickel(II)

Acta Cryst (2014). E**70**, 286-289

Terminology

Crystal data

M_r = 412.99

Monoclinic, *P21/n*

a = 7.6270 (17) Å

b = 7.2603 (16) Å

c = 13.580 (2) Å

V = 751.6 (3) $Å^3$

Z = 2

Ag K radiation

μ = 0.72 mm^{-1}

T = 295 K

0.20 x 0.20 x 0.20 mm

Terminology

Data collection

Enraf-Nonius CAD-4

Absorption correction: phi-scan

$T_{min} = 0.945$, $T_{max} = 0.958$

1706 measured reflections

1640 independent reflections

1215 reflections with $I > 2\sigma(I)$

$R_{int} = 0.021$

$$R = \frac{\sum_{hkl} \sum_j |I_{hkl,j} - \langle I_{hkl} \rangle|}{\sum_{hkl} \sum_j I_{hkl,j}}$$

Terminology

Refinement

$R[F^2 > 2\sigma(F^2)] = 0.035$

$wR(F^2) = 0.077$

S = 1.02

1640 reflections

140 parameters

$\Delta\rho_{max} = 0.34$ eÅ^{-3}

$\Delta\rho_{min} = -0.31$ eÅ^{-3}

$$R = \frac{\sum_{hkl} |F_{hkl}^{obs} - F_{hkl}^{calc}|}{\sum_{hkl} F_{hkl}^{obs}}$$

$$wR = (\frac{\sum |w|Y_o - Y_c|^2|}{\sum |wY_o^2|})^{1/2}$$

The Stable Pentamethylcyclopentadienyl Cation

J.B.Lambert et al. Angew. Chem. Int. Ed. 2002, 41, 1429-1431

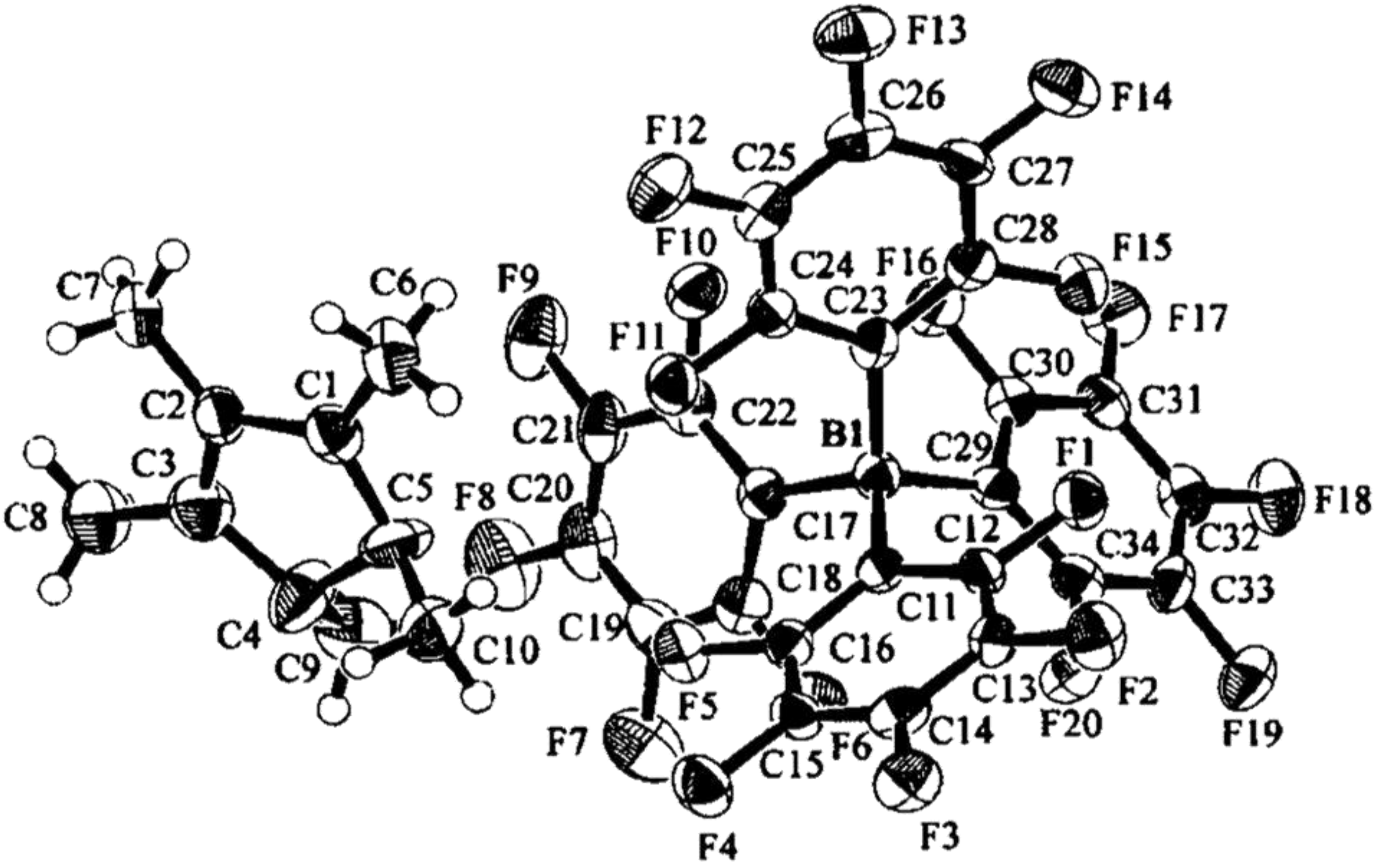
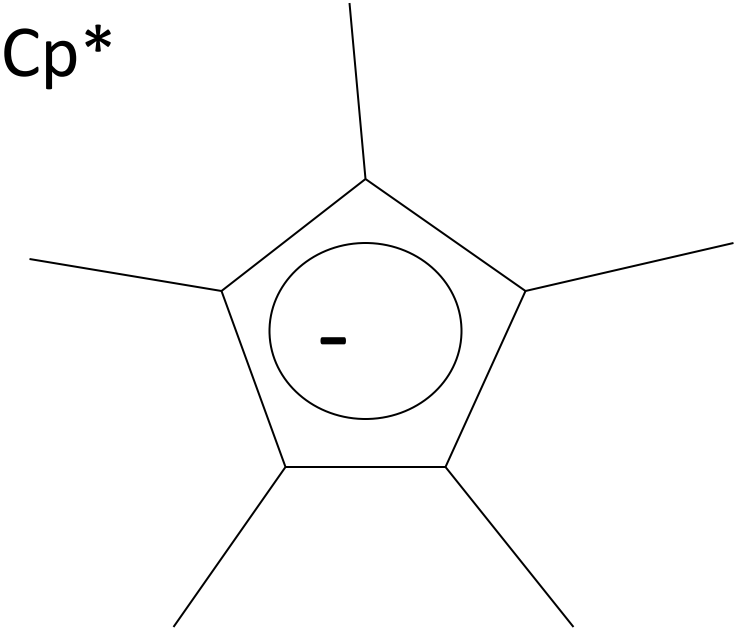


Figure 1. The crystal structure of pentamethylcyclopentadienyl tetrakis(pentafluorophenyl)borate. There is no covalent bonding between the cation on the left and the anion on the right.

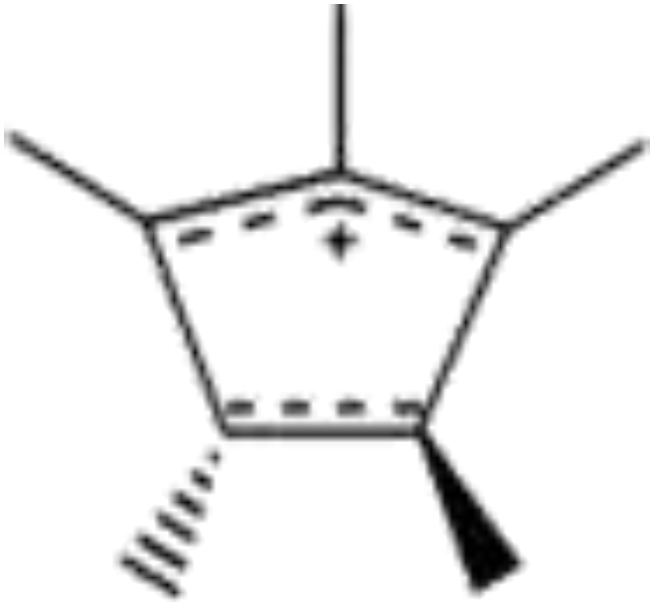
Pentamethylcyclopentadienyl

Common Anionic Form

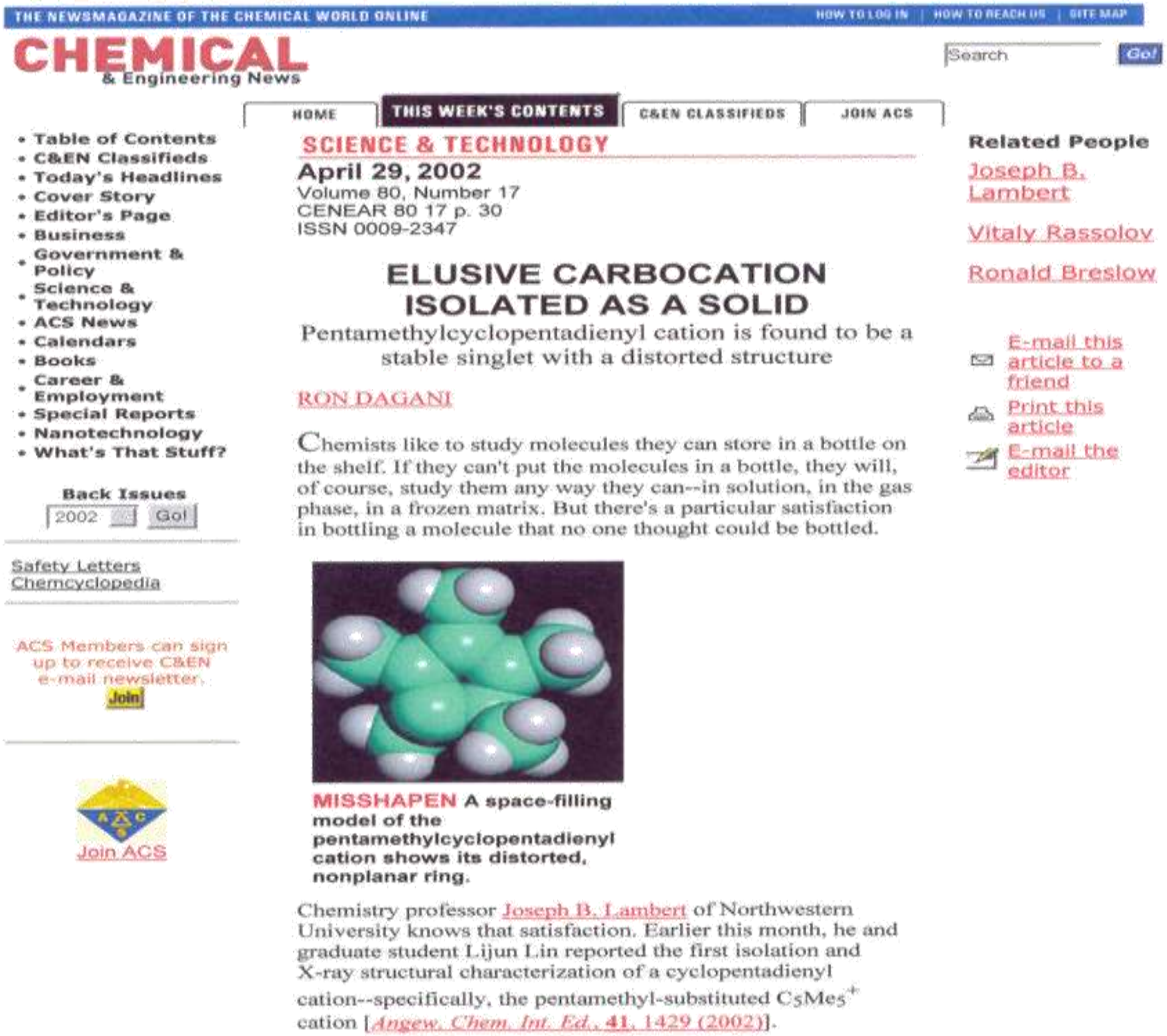
Cp*



"New" Cationic Form



Breakthrough



SCIENCE & TECHNOLOGY

April 29, 2002
Volume 80, Number 17
CENEAR 80 17 p. 30
ISSN 0009-2347

ELUSIVE CARBOCATION ISOLATED AS A SOLID

Pentamethylcyclopentadienyl cation is found to be a stable singlet with a distorted structure

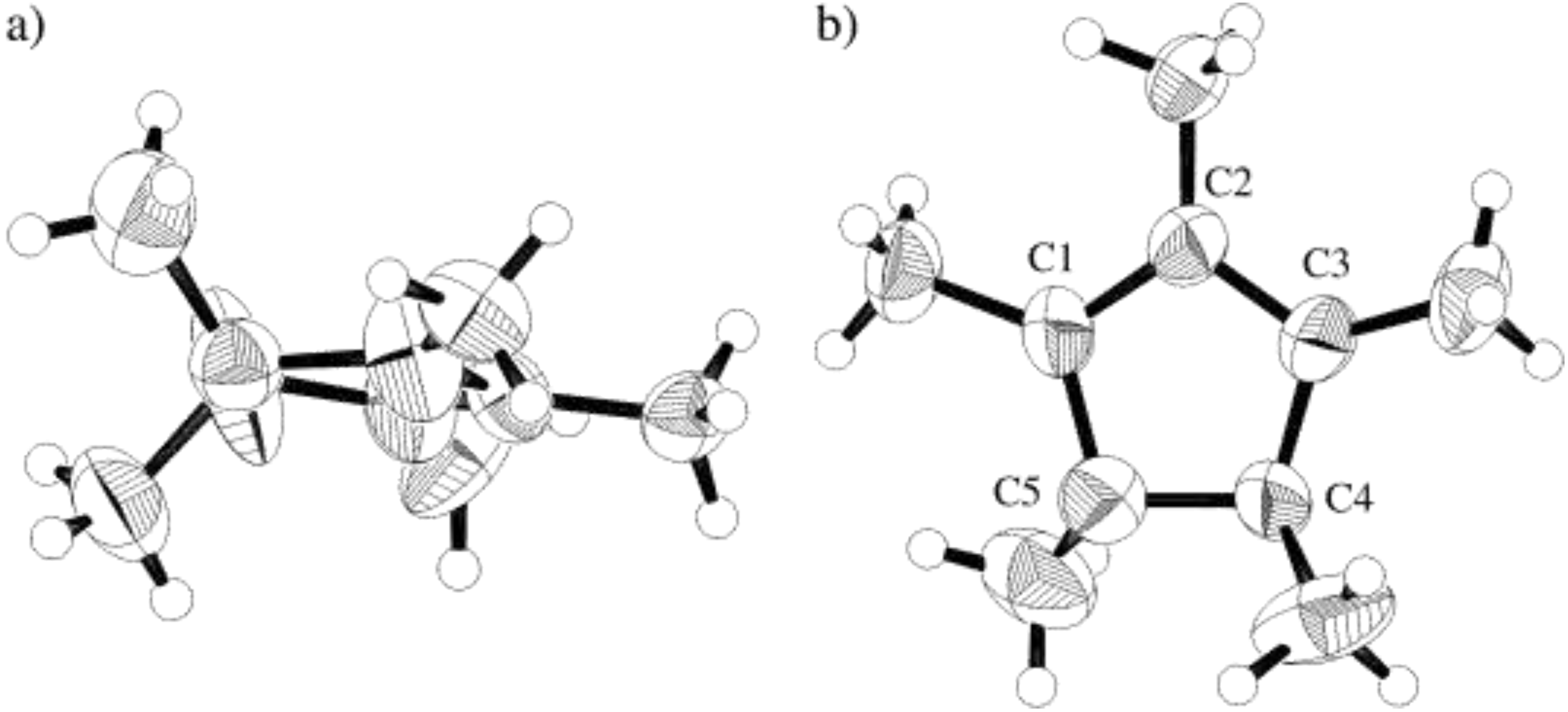
RON DAGANI

Chemists like to study molecules they can store in a bottle on the shelf. If they can't put the molecules in a bottle, they will, of course, study them any way they can--in solution, in the gas phase, in a frozen matrix. But there's a particular satisfaction in bottling a molecule that no one thought could be bottled.

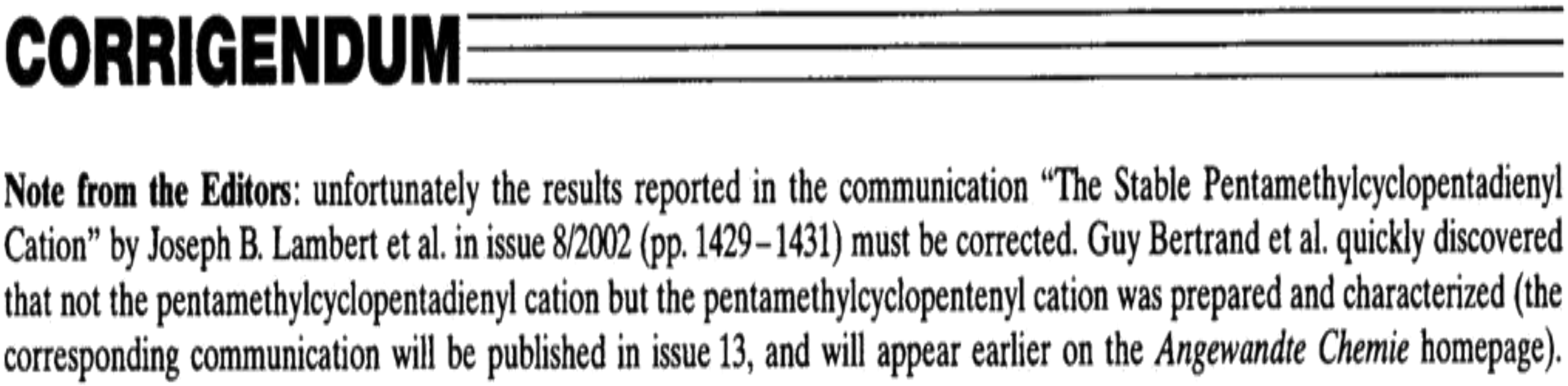
MISSHAPEN A space-filling model of the pentamethylcyclopentadienyl cation shows its distorted, nonplanar ring.

Chemistry professor Joseph B. Lambert of Northwestern University knows that satisfaction. Earlier this month, he and graduate student Lijun Lin reported the first isolation and X-ray structural characterization of a cyclopentadienyl cation--specifically, the pentamethyl-substituted $C_5Me_5^+$ cation [*Angew. Chem. Int. Ed.*, **41**, 1429 (2002)].

Puckered ring



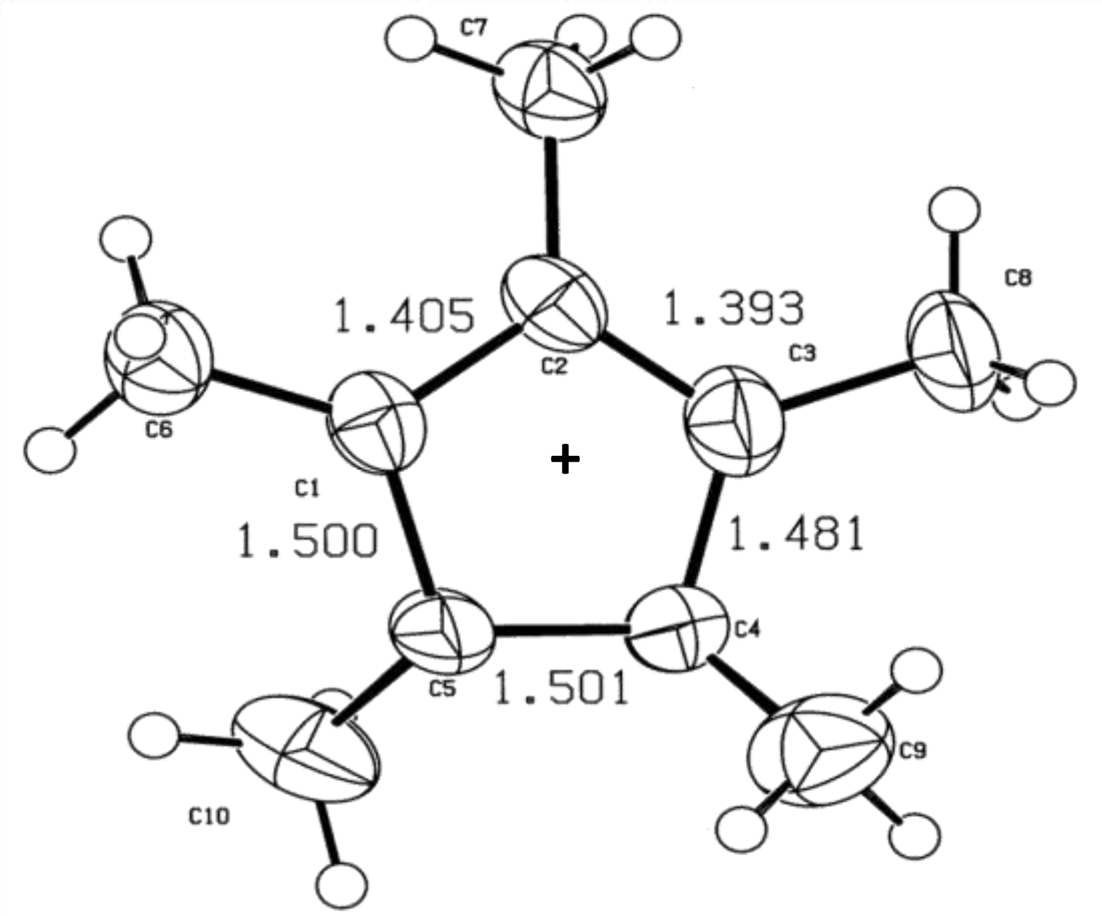
Uh-oh. Next issue:

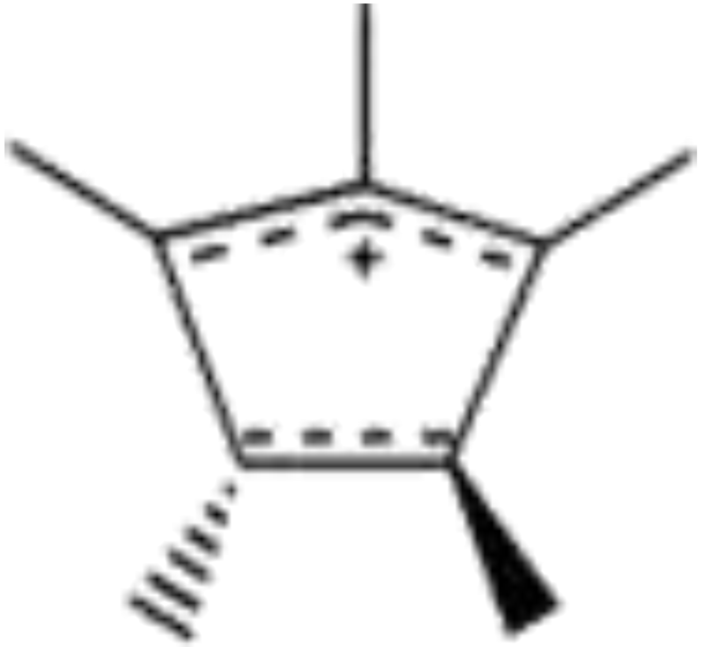


CORRIGENDUM

Note from the Editors: unfortunately the results reported in the communication “The Stable Pentamethylcyclopentadienyl Cation” by Joseph B. Lambert et al. in issue 8/2002 (pp. 1429–1431) must be corrected. Guy Bertrand et al. quickly discovered that not the pentamethylcyclopentadienyl cation but the pentamethylcyclopentenyl cation was prepared and characterized (the corresponding communication will be published in issue 13, and will appear earlier on the *Angewandte Chemie* homepage).

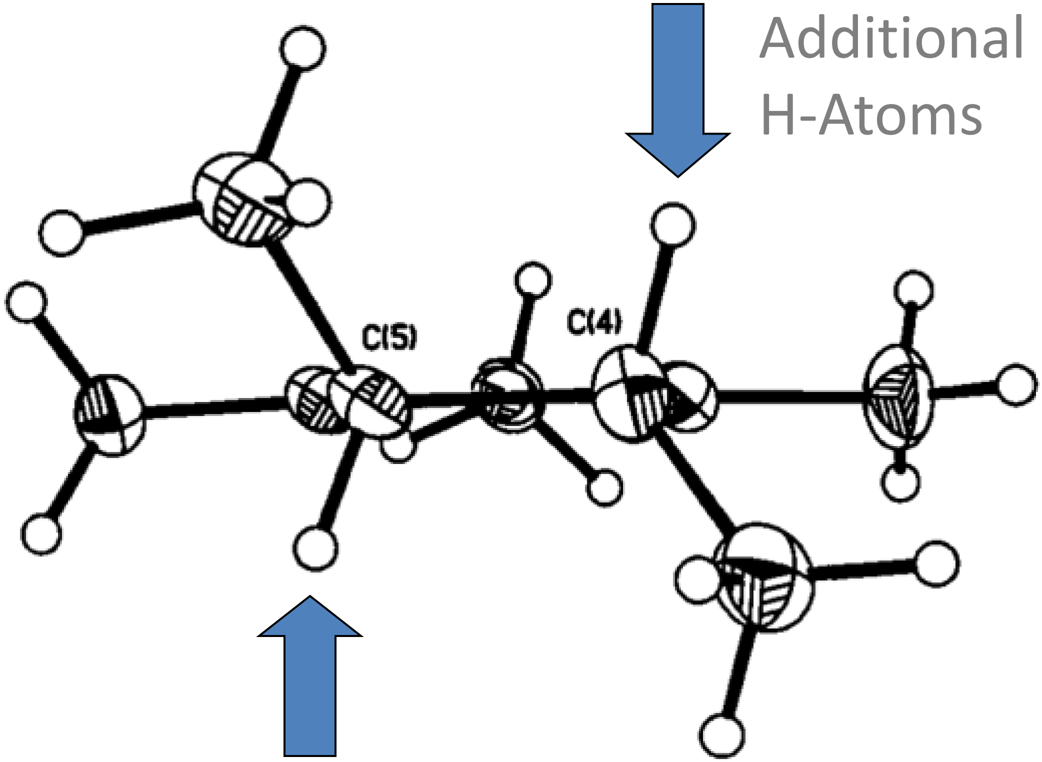
The warning signs





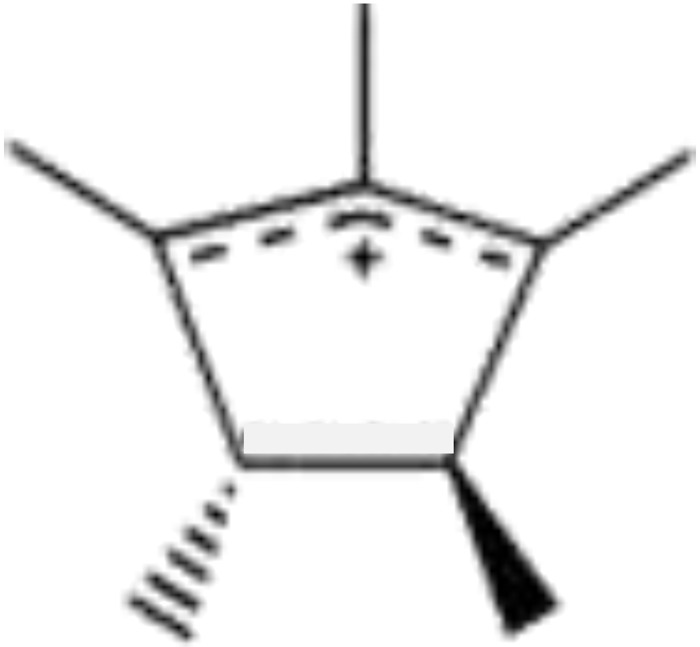
Corrected Structure

J.N. Jones et al.,Chem. Comm. 2002,1520-1521



Additional H-Atoms

Fig. 2 Side-on view of the pentamethylcyclopentenyl cation showing the hydrogen atoms attached to C(4) and C(5) and the *trans* arrangement of CH_3 groups attached to these carbon atoms.

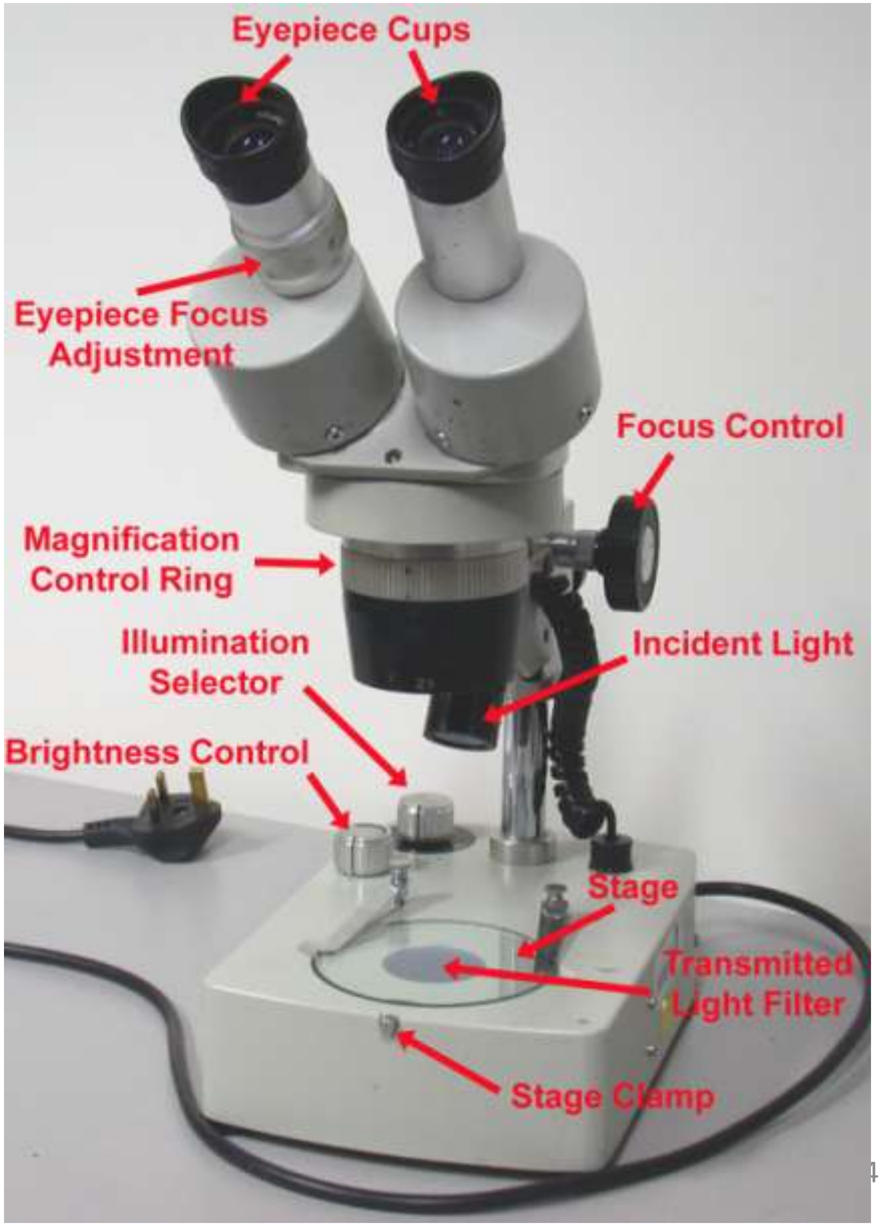


Over-enthusiastic interepretation

- First impression of the structure was likely based on a preliminary plot without H-atoms.
- Rushed as a short communication into the prestigious high impact journal Angewandte Chemie.
- Ignoring all counter theoretical and spectroscopic evidence (confirmation bias).
- No experienced crystallographer as referee?
- No *unbiased* automated CIF-validation.

The Optical Microscope

Structure Analysts usually use a dissecting microscope because it creates a stereoscopic image the right way up.



The Optical Microscope

Good crystals are glass-clear with well defined faces.

Small encrustations or surface damage may be acceptable.

The Optical Microscope

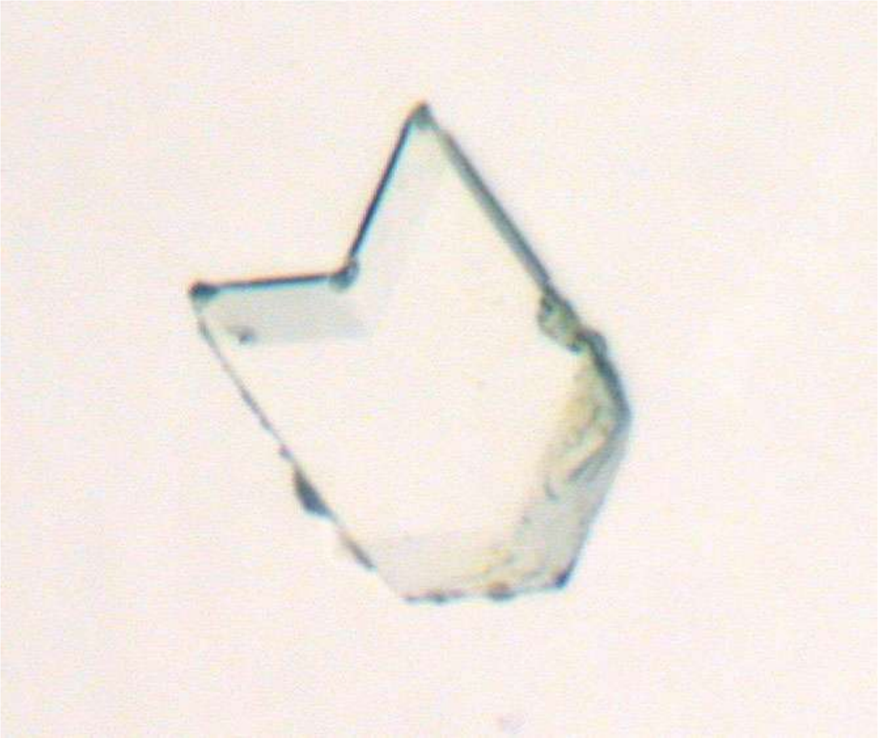
Polarised light may help to reveal crystal damage or twinning

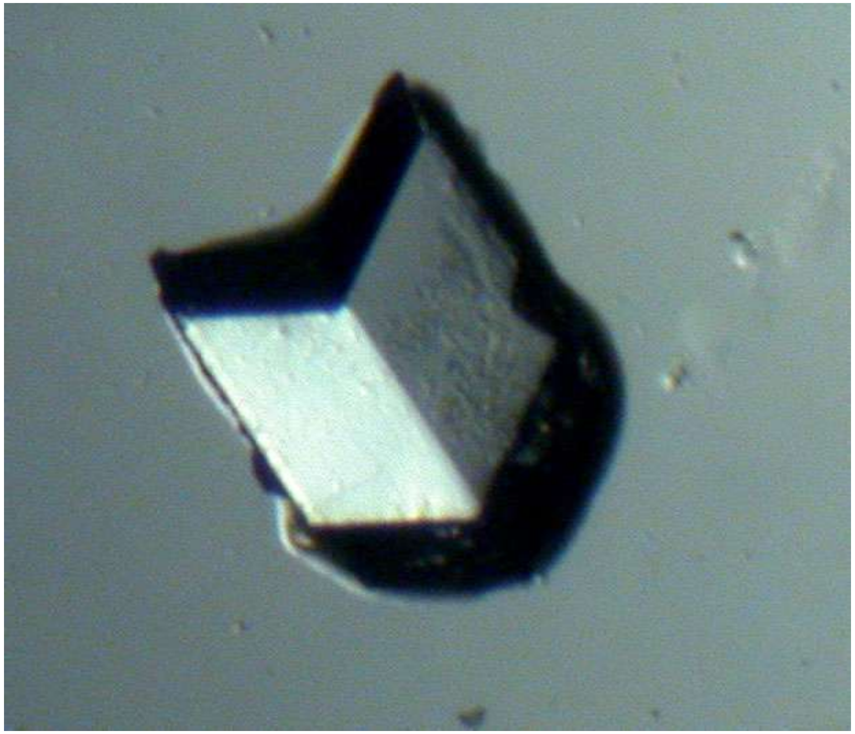
The Optical Microscope

Twinned Crystal.

Left: Normal light. Note ‘re-entrant angle’

Right: Polarised Light. Note one half of the crystal is darker than the other.



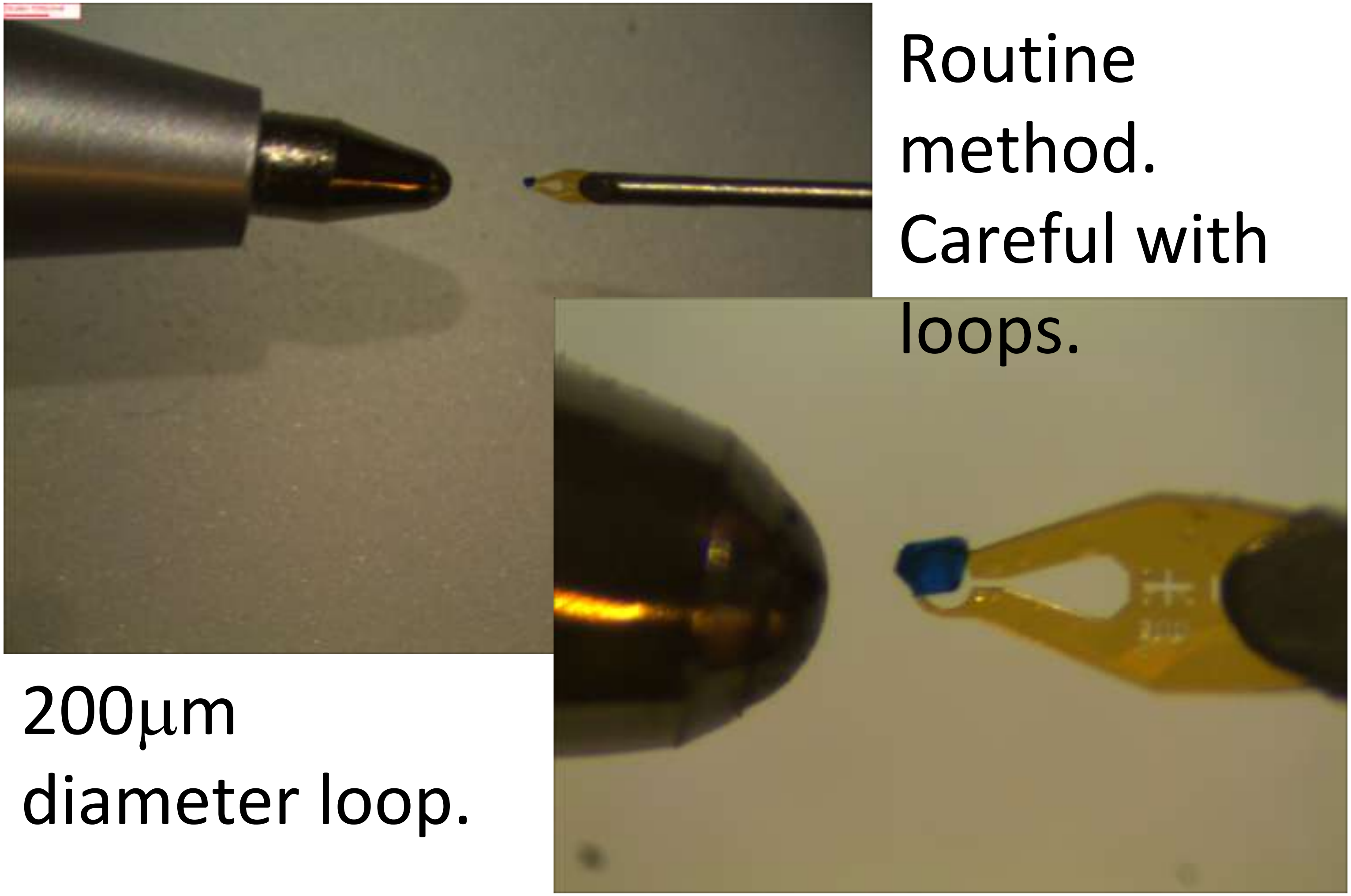


Ideal Crystal

- Glass-clear, without flaws

1. Cubic, about 0.5 mm
2. Rectangular prism
3. Lath
4. Plate

Mitegen Loop mounts

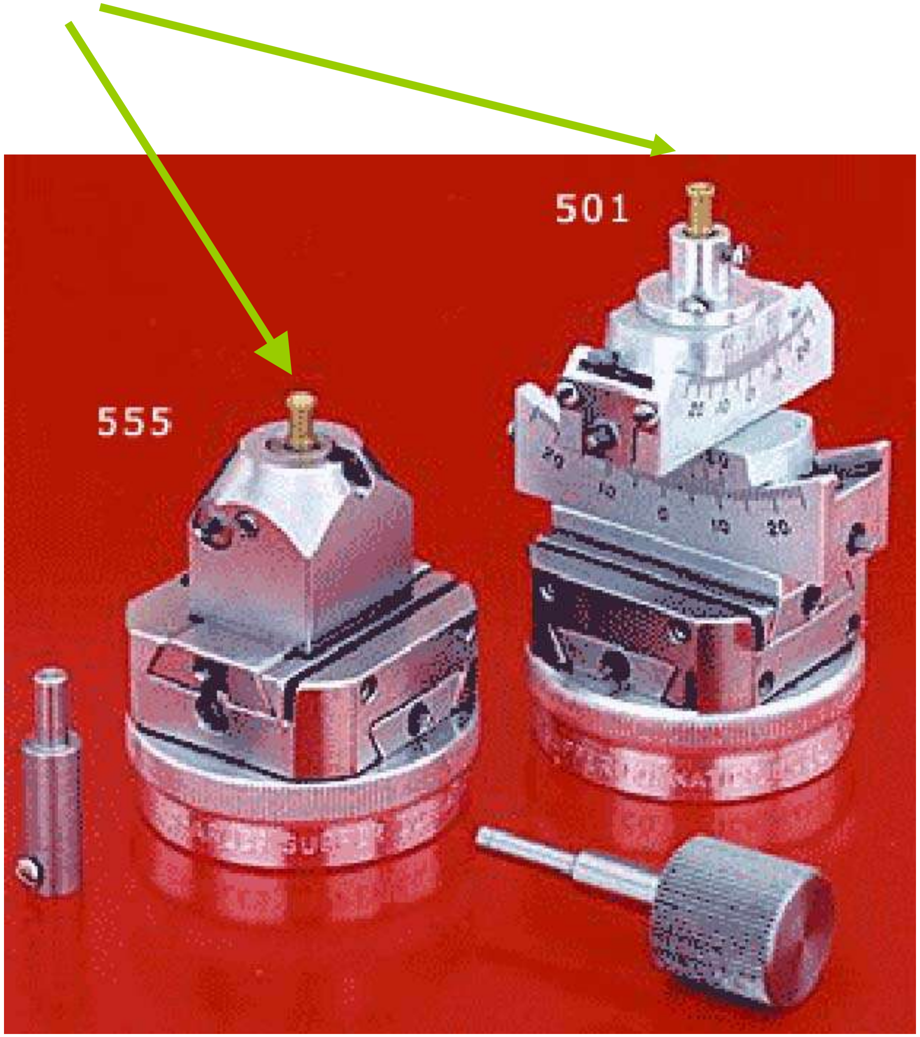


Routine method. Careful with loops.

200μm diameter loop.

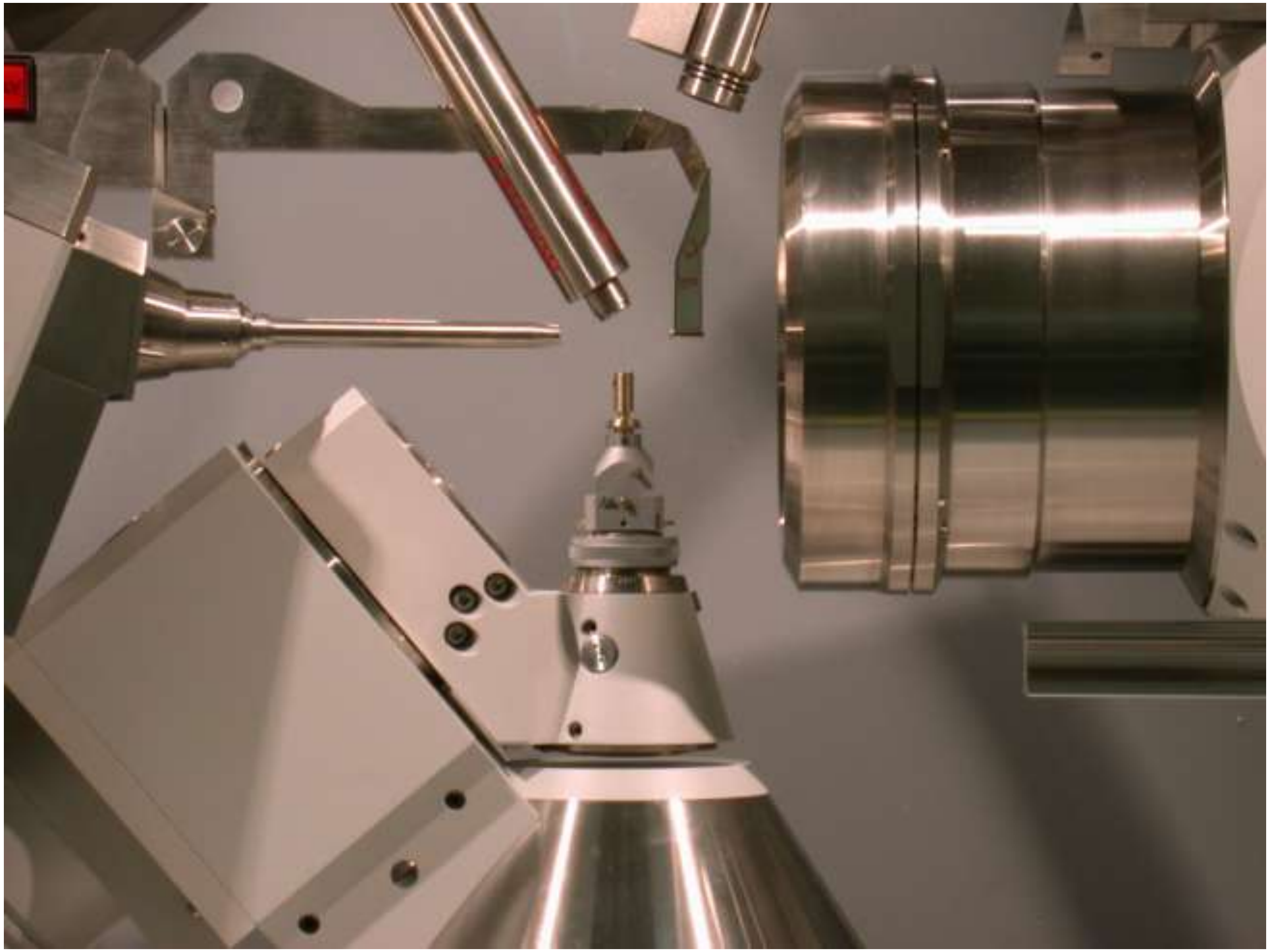
Goniometer Head

This is used to support the crystal and enable it to be located at the centre of the instrument.



Crystal is mounted here

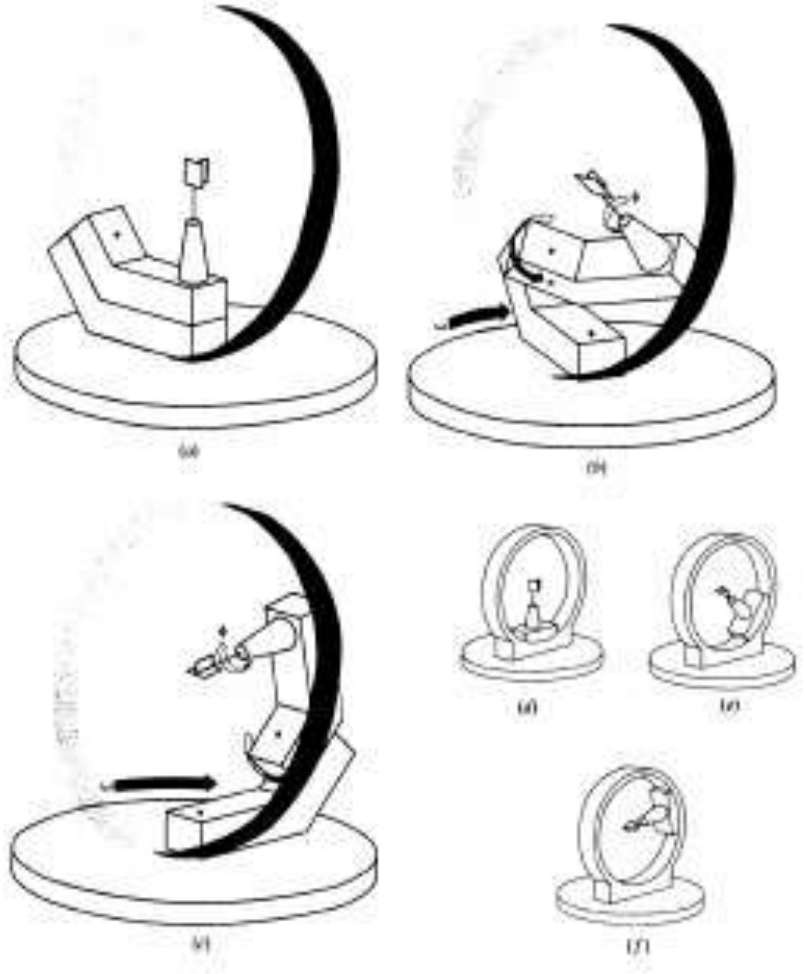
Close up of Crystal Mounted on a Gemini Diffractometer



Crystal Centring

The diffractometer consists of a number of rotateable axes which intersect at a point. The X-ray beam also goes through this point.

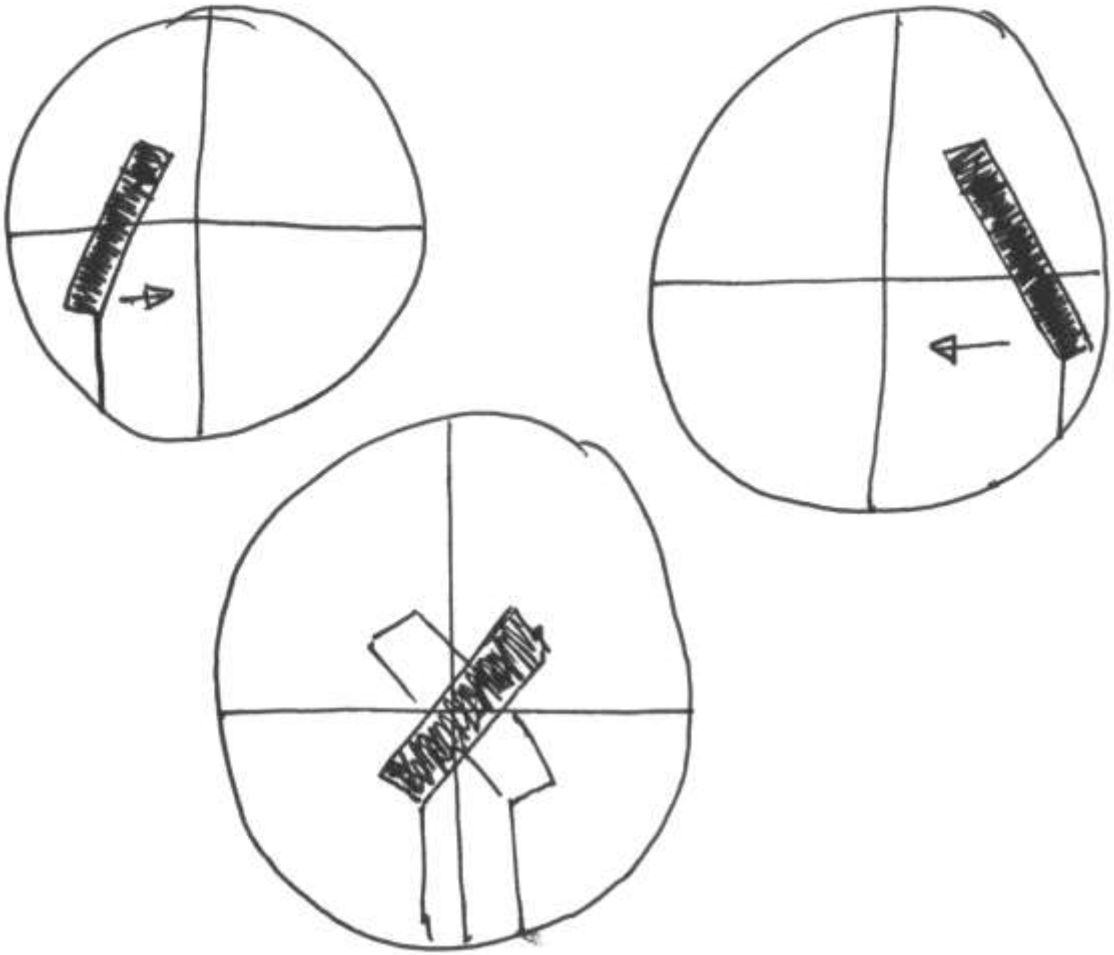
The crystal must be adjusted so that it lies at this intersection for all positions of the instrument



Crystal Centring

The Goniometer head supports the crystal and enables it to be adjusted to the centre of the instrument.

Crystal Centring



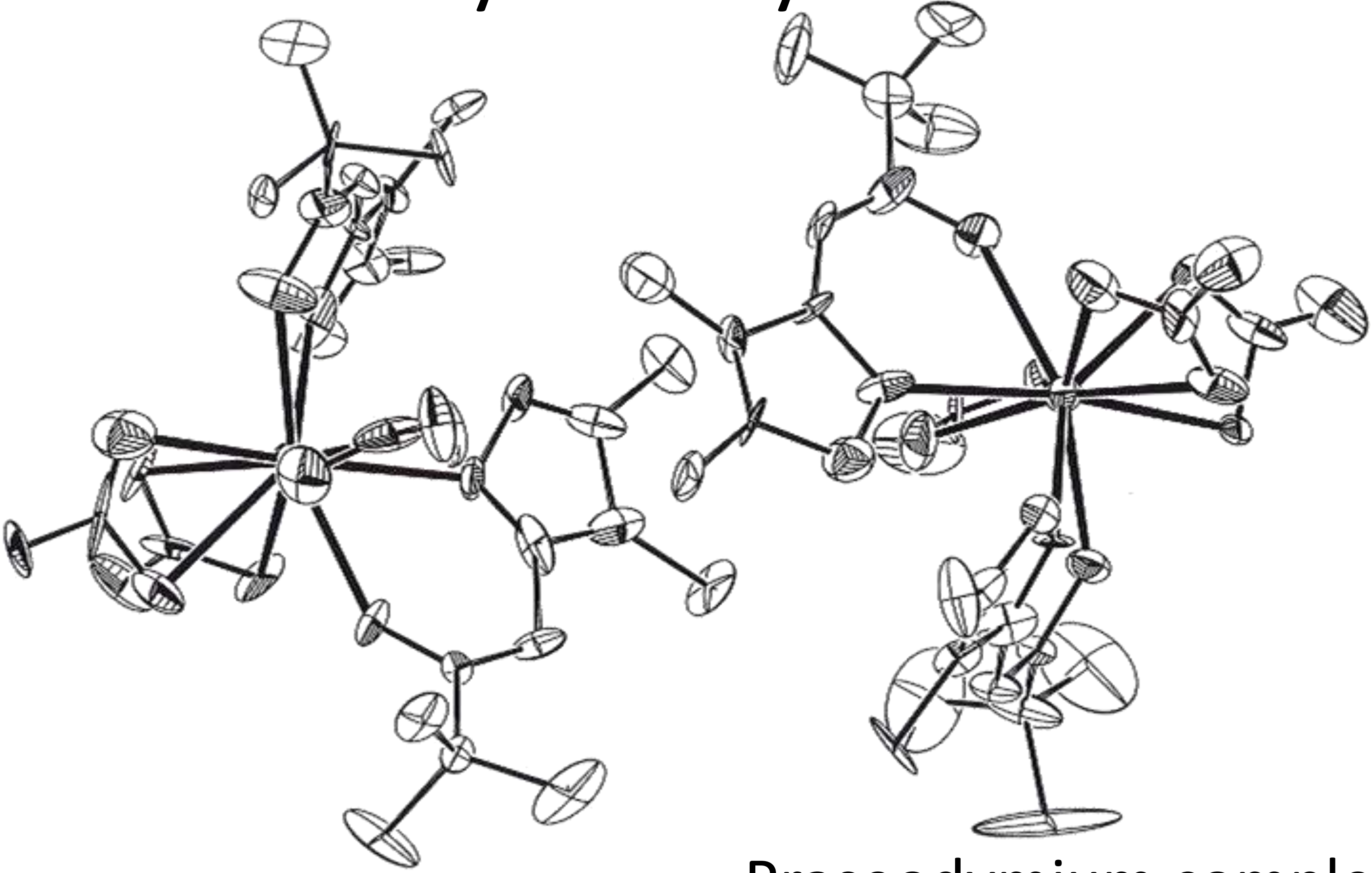
Crystal badly off-centre viewed at 0° and 180°.

Crystal reasonably well centred viewed at 0° and 180°.

Break: Read this

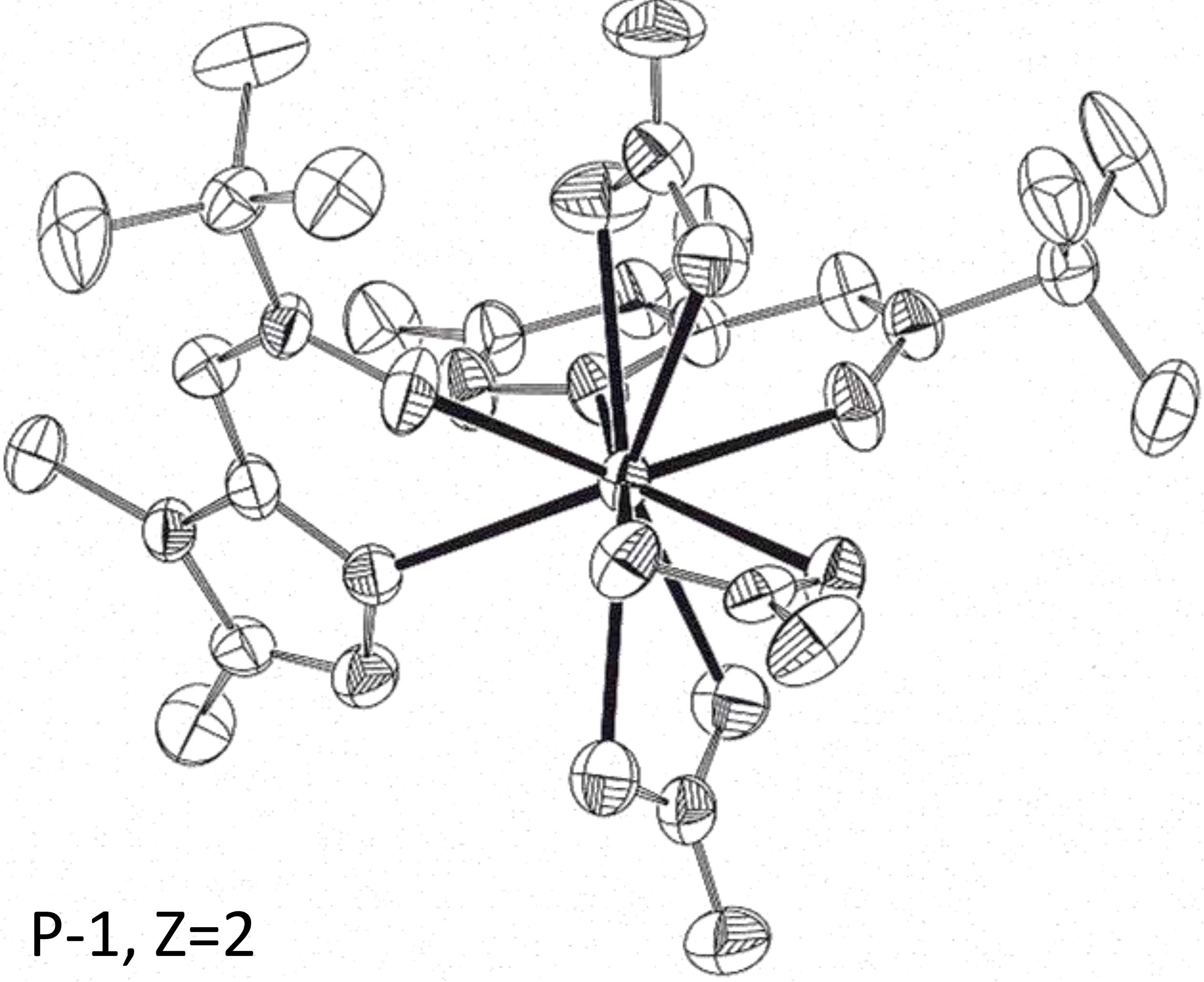
Praseodymium complex

Symmetry P1



Praseodymium complex
J.A.C.S. (2000),122,3413 – P1, Z = 2

Symmetry P-1



P-1, Z=2

Recap

- Scattering; waves; Fourier summations
- Bragg peaks; Bragg’s law; Diffraction patterns
- Time / space average; mosaic blocks; twins
- Reverse Fourier transform: miller index defines a wave vector, intensity gives its amplitude (squared). **Lost phases**.
- What do results look like?
- Nomenclature: unit cell; symmetry; hkl; xyz
- Example experiment (speeded up)
- What good samples look like.

1. Sample Prep
2. Symmetry
3. Structure Solution
4. Refinement
5. Problems: Disorder, Twinning

What are the difficult steps?

1. Grow a decent crystal (can be easy or hard – nature decides)
2. Determine the *Laue* symmetry (the symmetry of the diffraction pattern) to help collect the data.
3. Determine the *Space Group* symmetry to solve the structure.
4. Deal with disorder, twinning, solvent molecules, unexpected results.

Sample Prep

http://www.xtl.ox.ac.uk/oxonly/intro.html links to:

Paul D. Boyle's "Growing Crystals That Will Make Your Crystallographer Happy":

http://crystallography.usc.edu/growingcrystals.pdf

Slow evaporation; slow cooling; vapour diffusion; layering; convection; co-crystallants; counterions;

"Golden rules"

Symmetry

Symmetry

Types of symmetry
Space Groups
Effect on Structure Factor
Determine space group from absences
Why you should care

Understanding Symmetry

Crystallographic symmetry seems difficult.

This may be because it combines point symmetry (as used to describe isolated molecules) and translational symmetry.

Because of this, crystallographers have to use a special notation.

Some common Space Groups

P -1	a centre of symmetry
P 2_1	a two-fold screw axis
P 2_1 2_1 2_1	three perpendicular 2-fold screws
P 2_1/c	c-glide ⊥ to 2-fold screw axis
P b c a	three perpendicular glide planes

230 in total

Some common Space Groups

P -1 a centre of symmetry

$x, y, z;$ $-x, -y, -z;$

P 2_1 a two-fold screw axis

$x, y, z;$ $-x, y+½, -z;$

P 2_1 2_1 2_1 three perpendicular 2-fold screws

$x, y, z;$ $-x+½, -y, -z+½;$

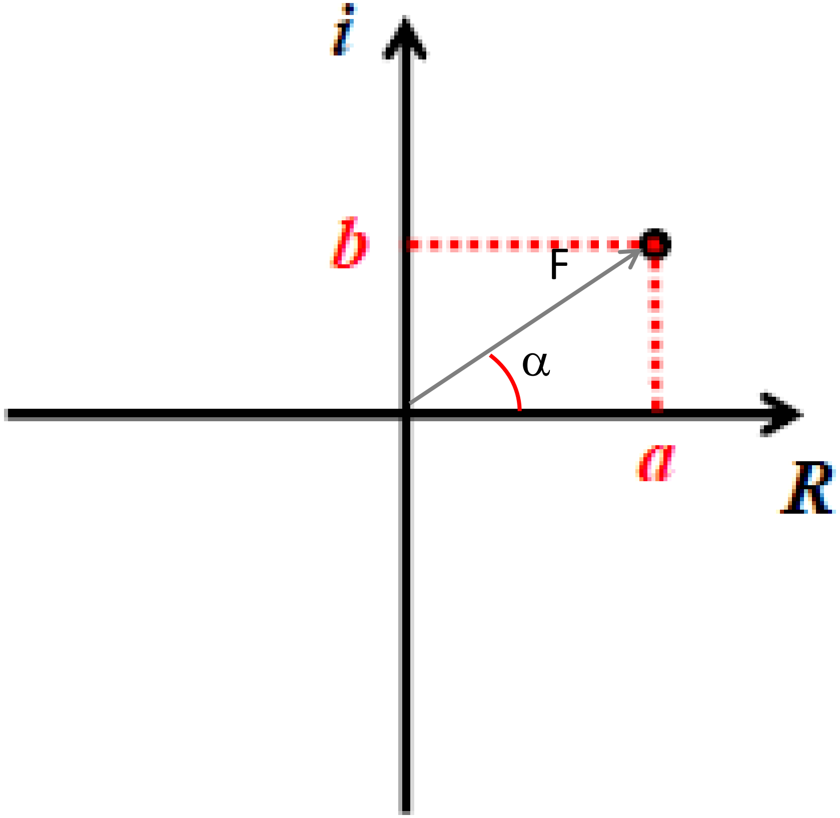
$-x, y+½, -z+½;$ $x+½, -y+½, -z;$

The Structure Factor

$$Fc_h = \sum_{j}^{allatoms} f_j \exp^{2\pi i\left(hx_j + ky_j + lz_j\right)}$$

$$A_h = \sum_{j}^{allatoms} f_j \cos 2\pi\left(hx + ky + lz\right)_j$$

$$B_h = \sum_{j}^{allatoms} f_j \sin 2\pi\left(hx + ky + lz\right)_j$$



$$F_c^{\,2} = A^2 + B^2 \qquad \tan\alpha = \frac{B}{A} = \frac{\sum \sin 2\pi hx}{\sum \cos 2\pi hx}$$

Symmetry

The structure factor calculation is a summation over all the atoms in the unit cell.

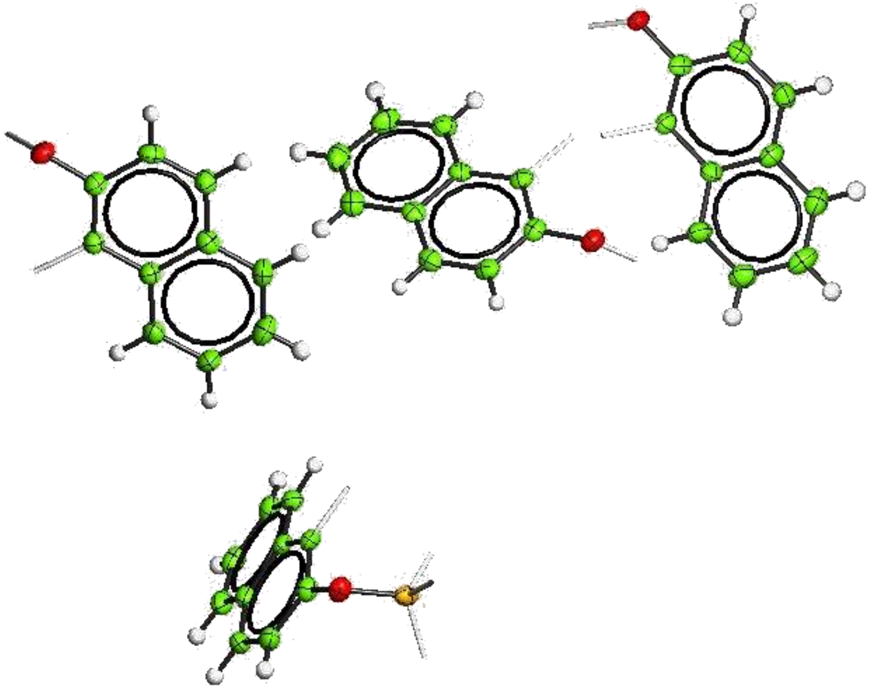
Because the *sines* and *cosines* are multiples of 2π, integers can be added to atomic coordinates without changing the calculation.

$$A_h = \sum_{j}^{allatoms} f_j \cos 2\pi\left(hx + ky + lz\right)_j$$

$$= \sum_{j}^{allatoms} f_j \cos 2\pi\left(h(x+n) + ky + lz\right)_j$$

Symmetry

However, usefully visualising a molecule requires the atoms to be put into positions which bring them closest to each other



Symmetry

In addition, there may be symmetry *inside* the unit cell.

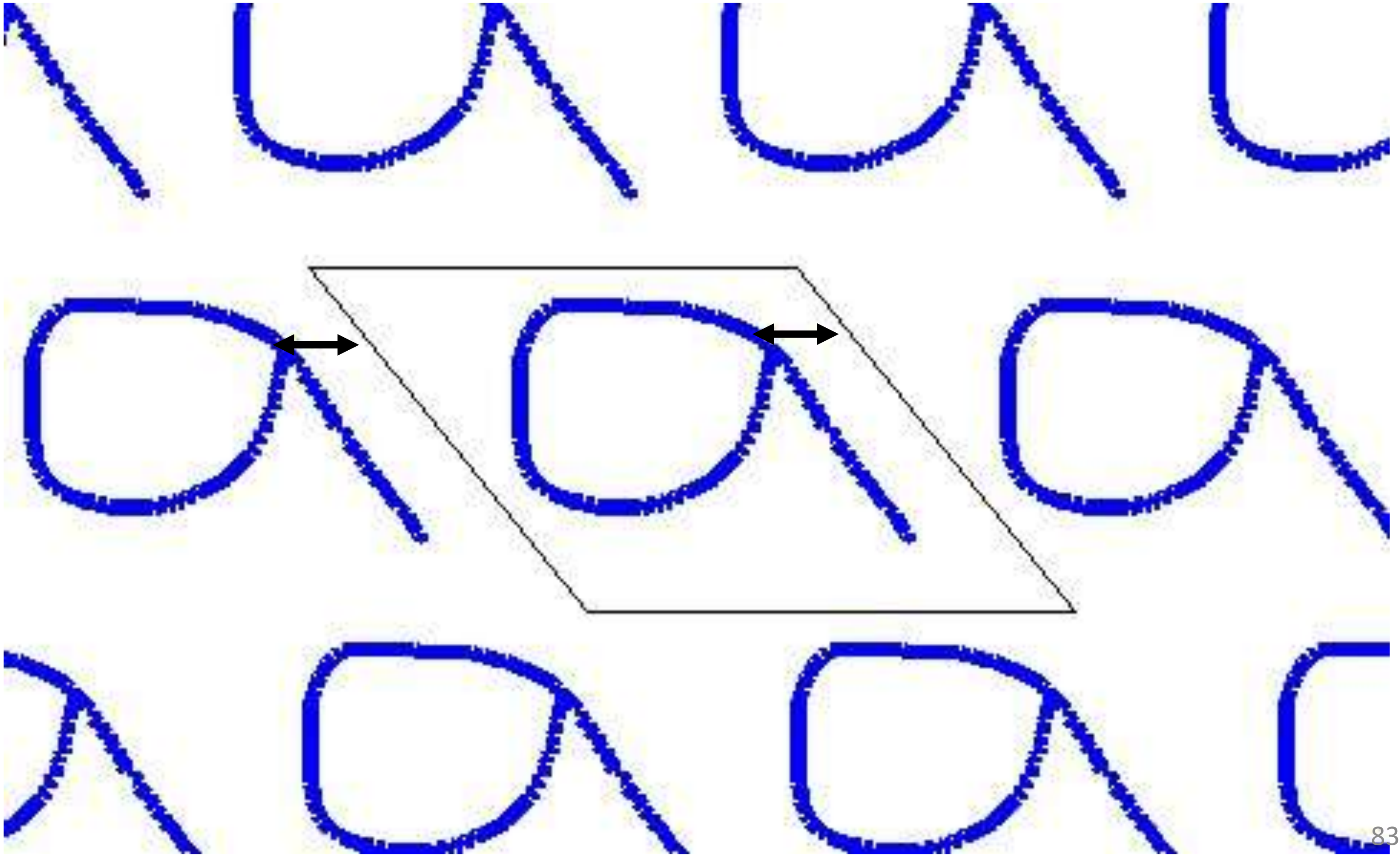
The structure factor calculation can be simplified so that it only computes terms for the *unique* atoms, and then modifies the terms to account for this symmetry.

Crystallographic drawing programs need to know about these symmetry considerations

Demo: Fragment

Triclinic P1.

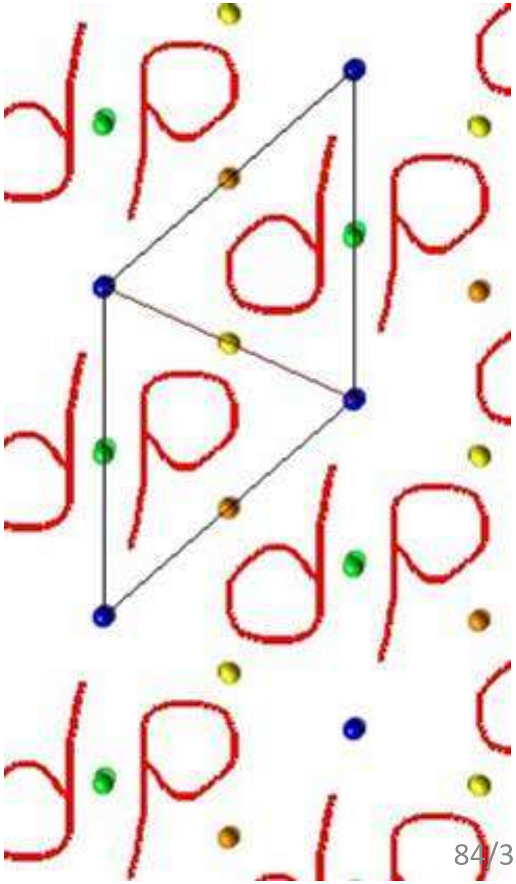
The whole environment is enclosed within the unit cell



Symmetry Operators

Symmetry operations inside the unit cell mean that the contents of the cell can be represented by an **asymmetric unit** plus a list of symmetry operations.

Though this may seem complicated to a human, Nature uses it, and it greatly simplifies the computations.



Symmetry in the Unit Cell

The motifs within the cell may be related to each other by operations which are not simple translations.

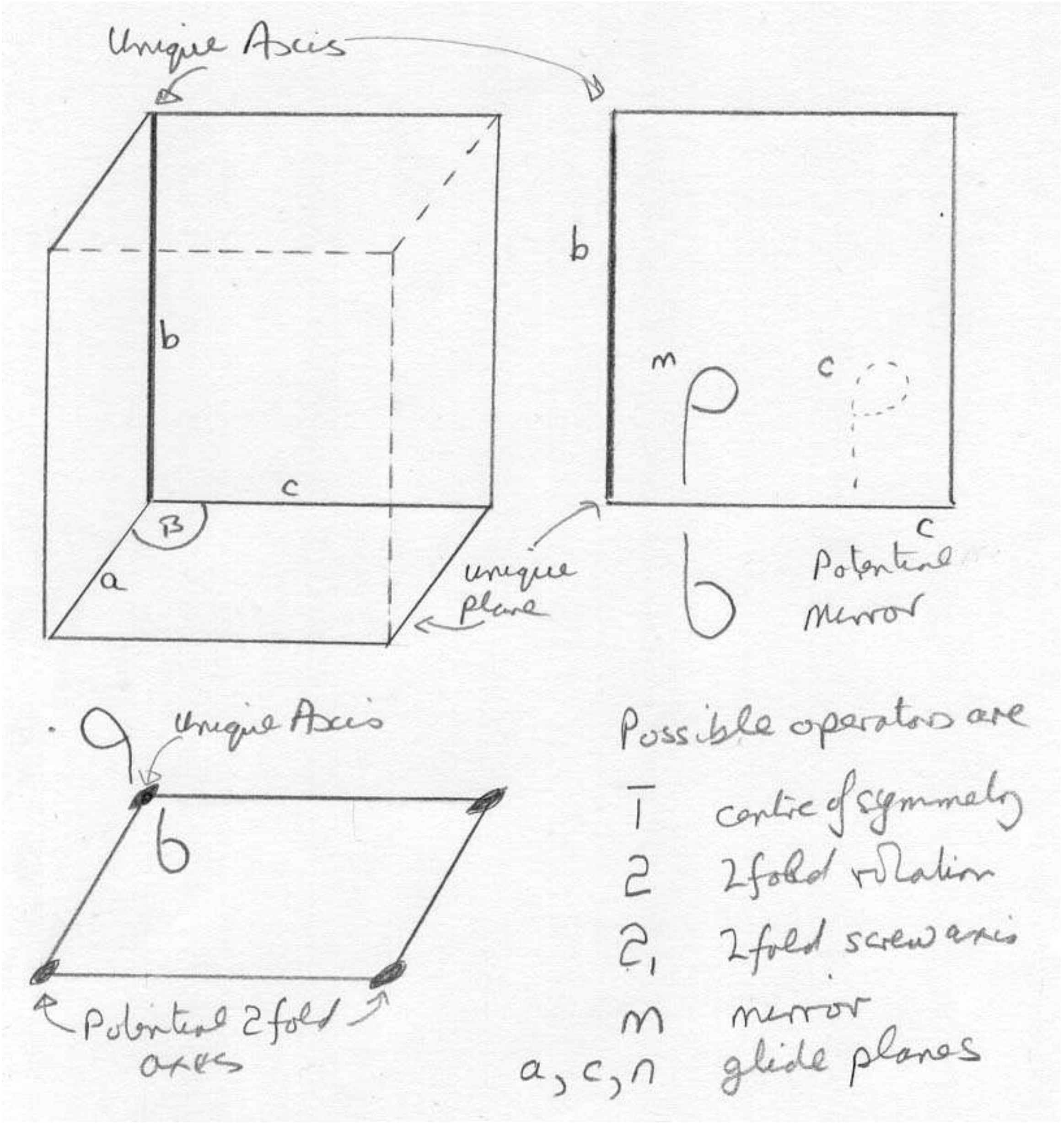
These operations include:

- Inversion
- Rotation
- Rotation-translation – called Screw operations
- Reflection
- Reflection-translation – called Glide operations

Permitted Symmetry

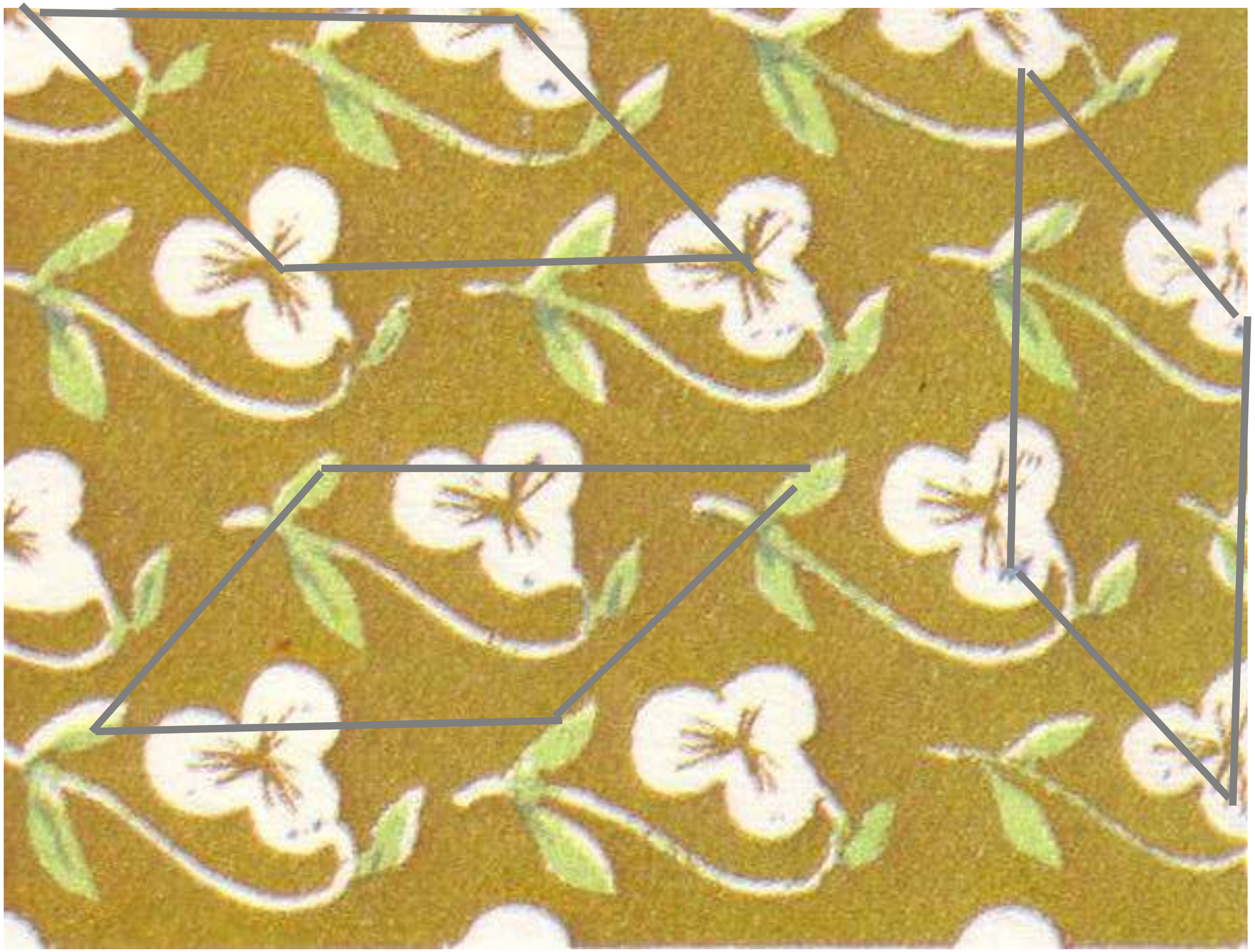
Operators can only act if permitted by the unit cell symmetry.

A monoclinic cell



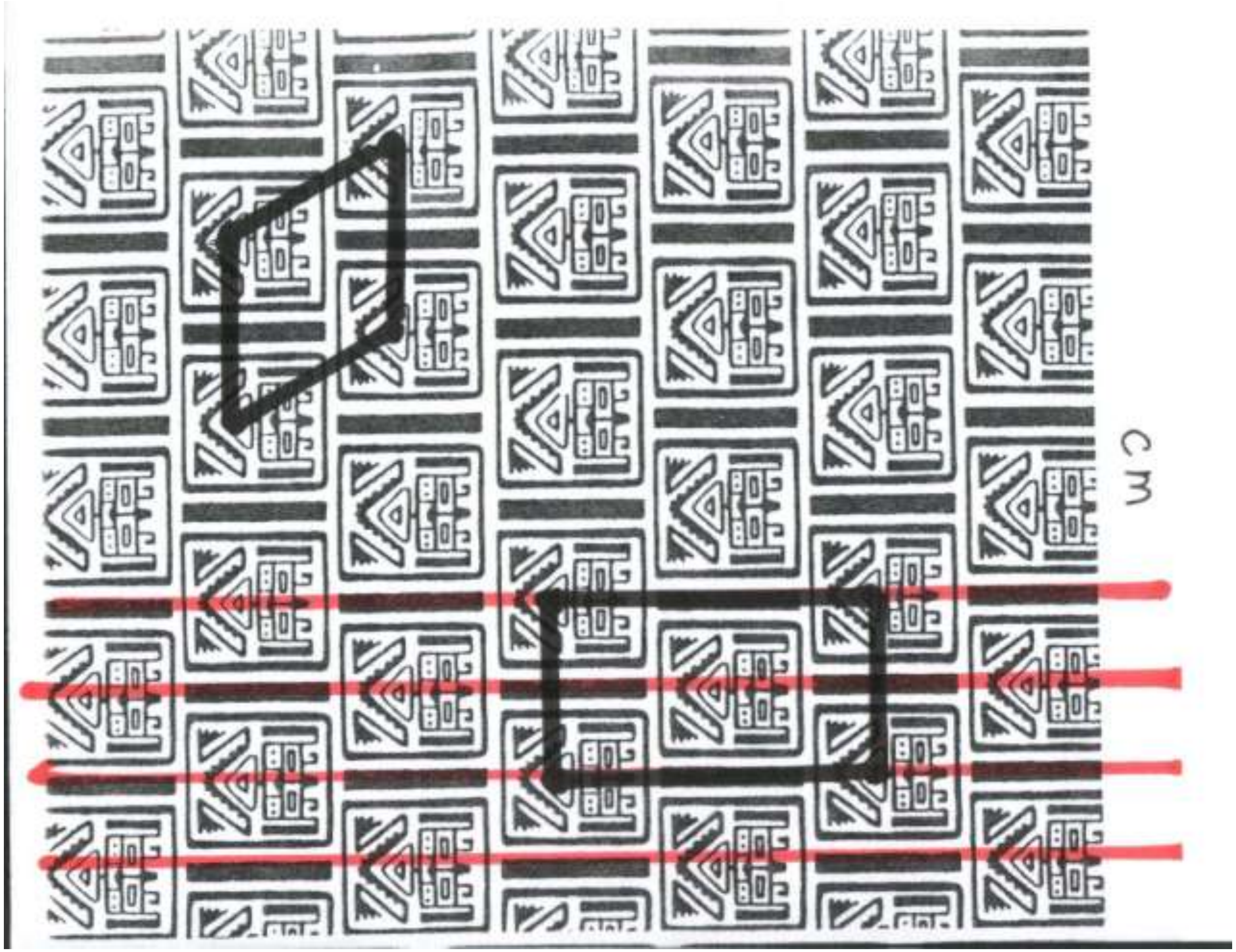
Symmetry

No symmetry except simple translation: P1



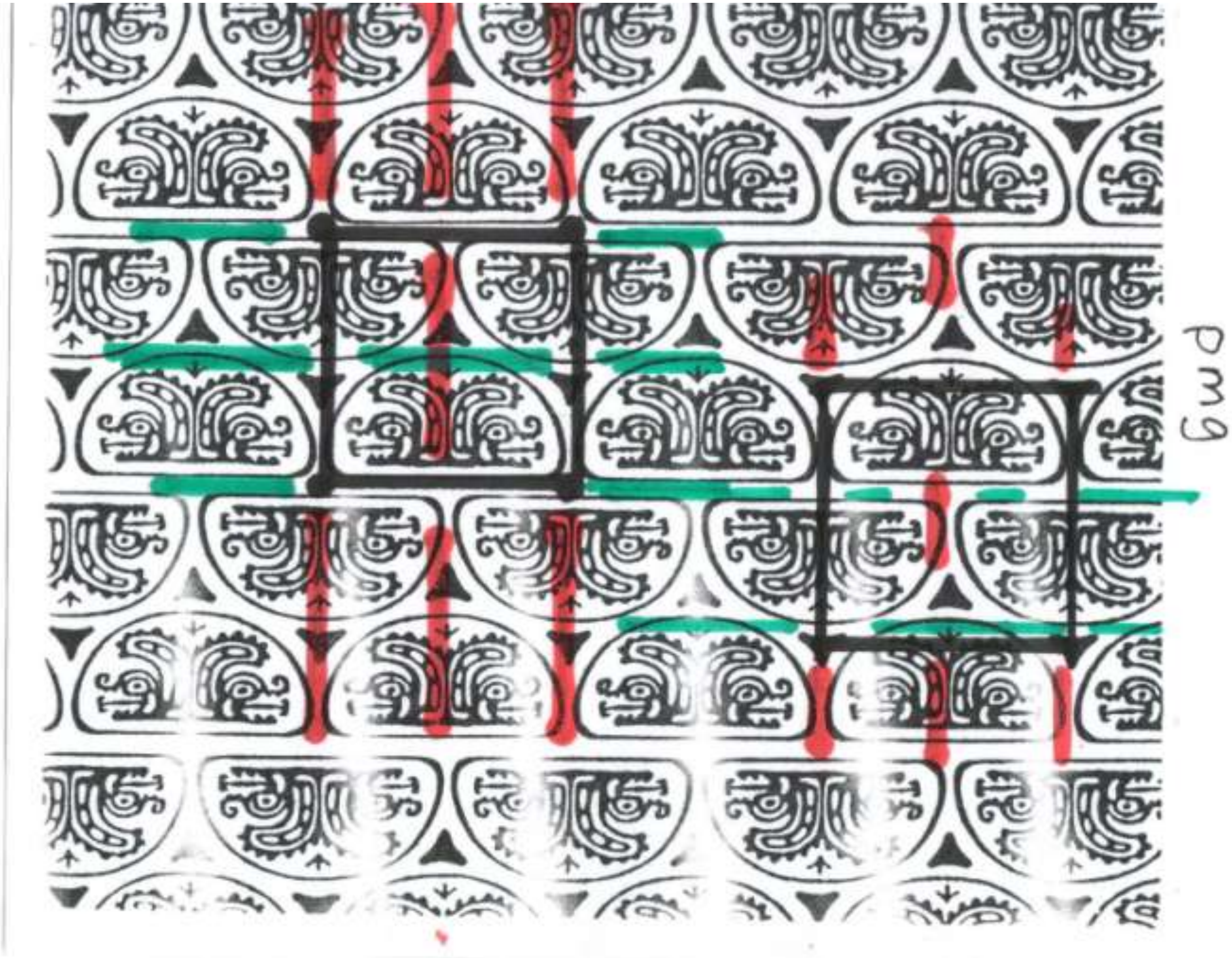
Symmetry

In this example, the most highly symmetric cell is rectangular, centred and contains a mirror.



Symmetry

Symmetry operators are mirrors and glide planes.
Choose the cell origin to lie on these operators



DEMO

Escher Web Sketch is a program to enable you to get an intuitive feeling for space groups by playing with them.

Hardaker, W. & Chapuis, G. (1997*b*). *Escher Web Sketch* [*http://escher.epfl.ch/escher/*](http://escher.epfl.ch/escher/)

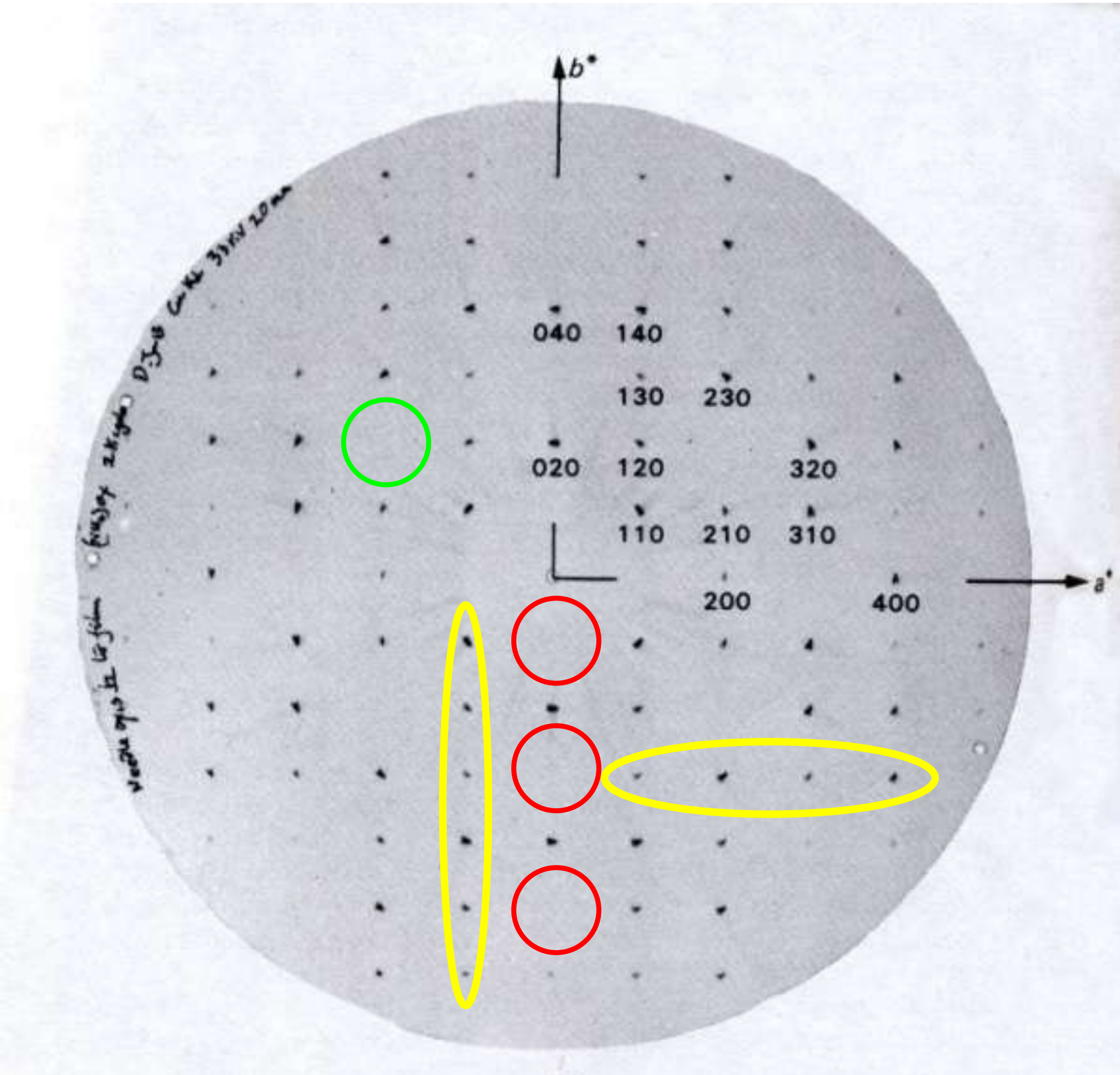
Link

Symmetry

The intensity of the diffraction spots tell us about the atom distribution in the cell.

A regular pattern of zero intensity reflections (**systematic absences**) tells us that there is a regular translational pattern to the density in the cell – crystallographic symmetry.

Reciprocal Lattice Image



(a)

The diffraction image is distorted because of the instrument geometry.

It can be transformed into an undistorted image.

Green ring – accidental absence.

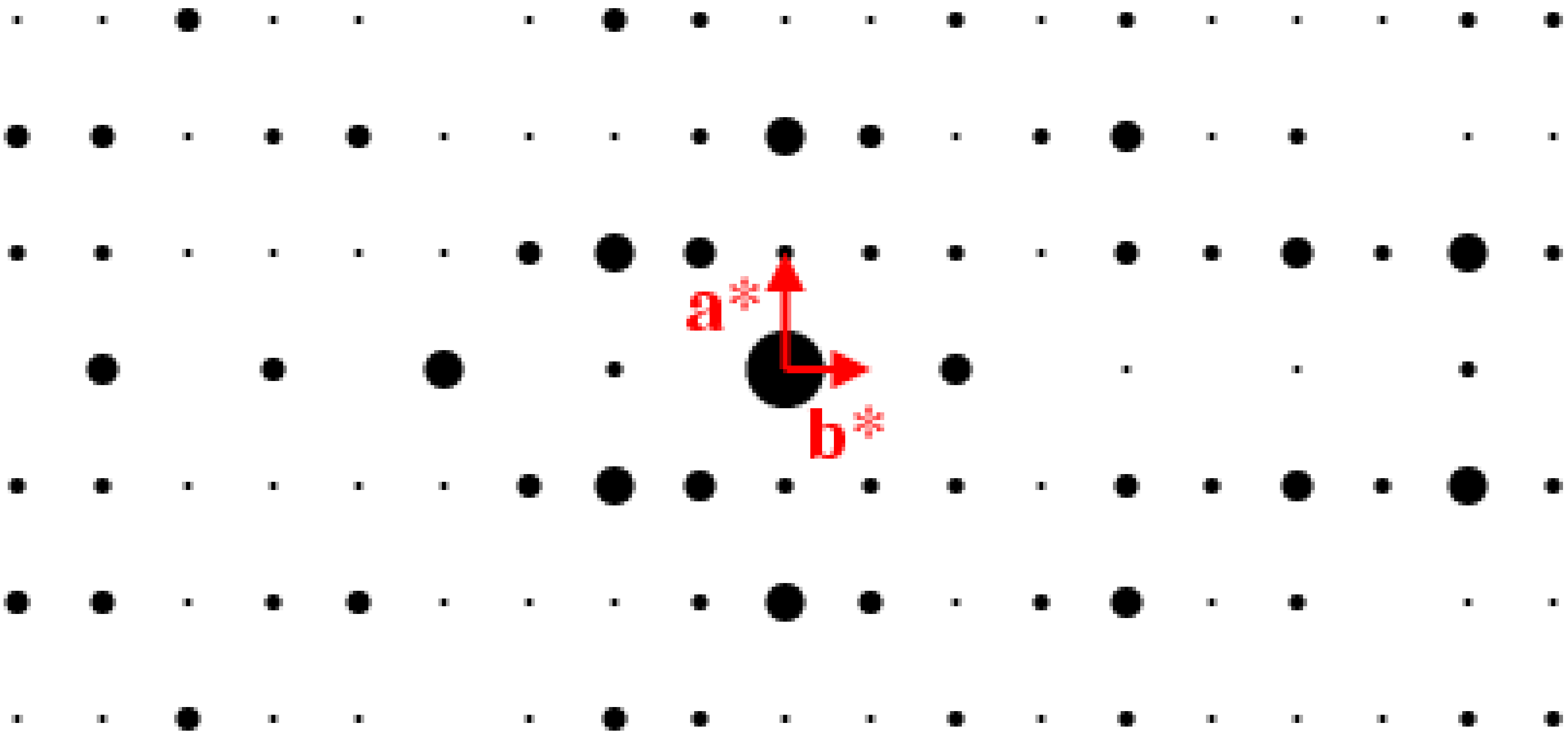
Red rings – systematic absences.

International Tables

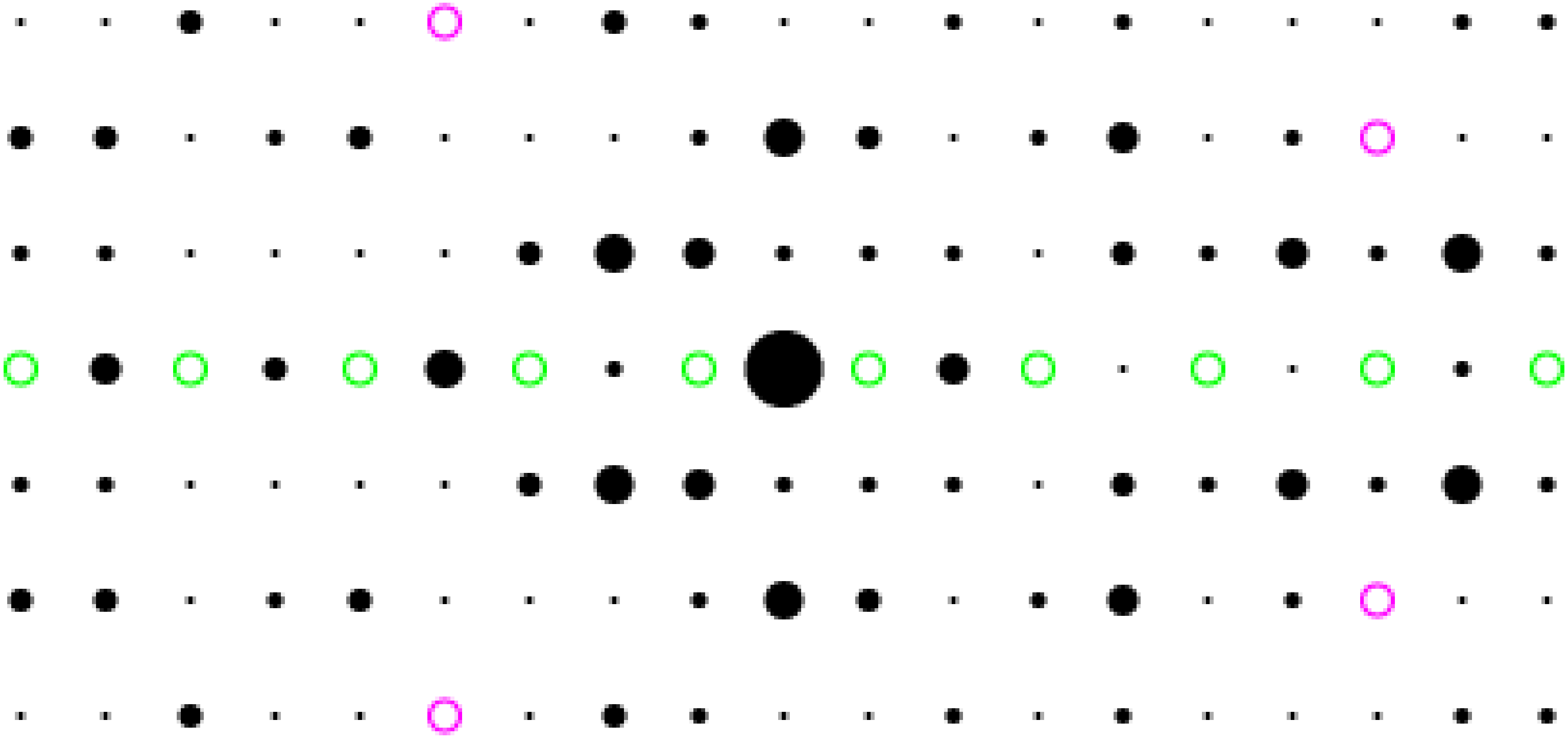
International Tables, Volume A, contains information about all the 230 possible 3D space groups, including diagrams, lists of geometric relationships, and the systematic absences.

It also includes a list to help determine the space group from the observed absences.

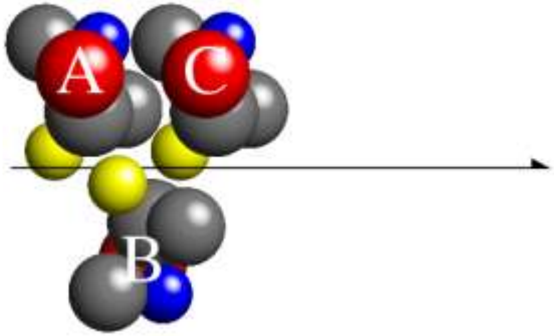
Absences



Absences



Screw axis



Absences in the *0k0* row are caused by the 2_1 axis parallel to b.

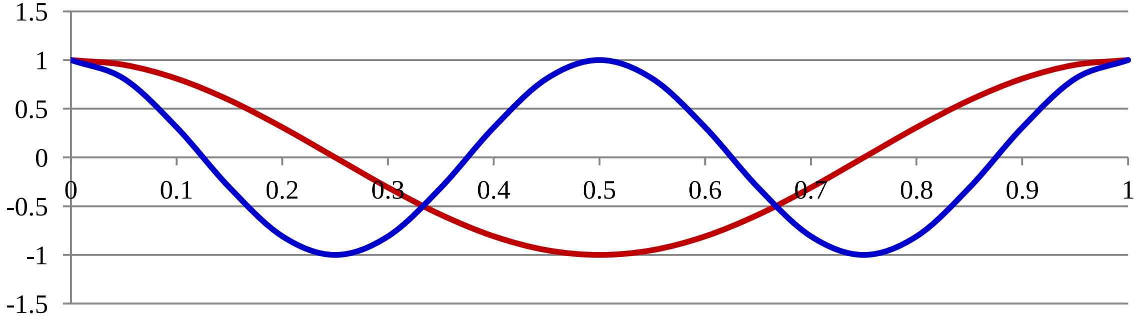
It's the translation of a half a unit cell that causes the absence.

$$Fc_{hkl} = \sum_{j}^{allatoms} f_j \exp^{2\pi i\left(hx_j+ky_j+lz_j\right)}$$

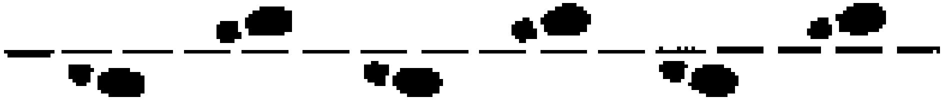
$$Fc_{0k0} = \sum_{j}^{allatoms} f_j \exp^{2\pi i\left(ky_j\right)}$$

$$= \sum_{j}^{allatoms} f_j\left(\cos 2\pi\left(ky_j\right)+i\sin 2\pi\left(ky_j\right)\right)$$

$$= \sum_{j}^{halfatoms} 2f_j \begin{pmatrix} \cos 2\pi k\left(y_j\right)+i\sin 2\pi k\left(y_j\right)+ \\ \cos 2\pi k\left(y_j+\frac{1}{2}\right)+i\sin 2\pi k\left(y_j+\frac{1}{2}\right) \end{pmatrix}$$



Glide plane



Absences in the *h0l* plane are caused by the c-glide axis perpendicular to b.

$$Fc_{hkl} = \sum_{j}^{allatoms} f_j \exp^{2\pi i\left(hx_j + ky_j + lz_j\right)}$$

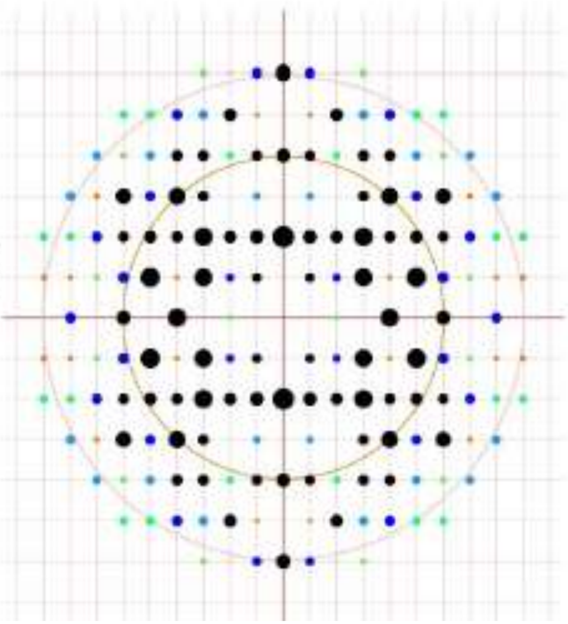
$$Fc_{h0l} = \sum_{j}^{allatoms} f_j \exp^{2\pi i\left(hx_j + lz_j\right)}$$

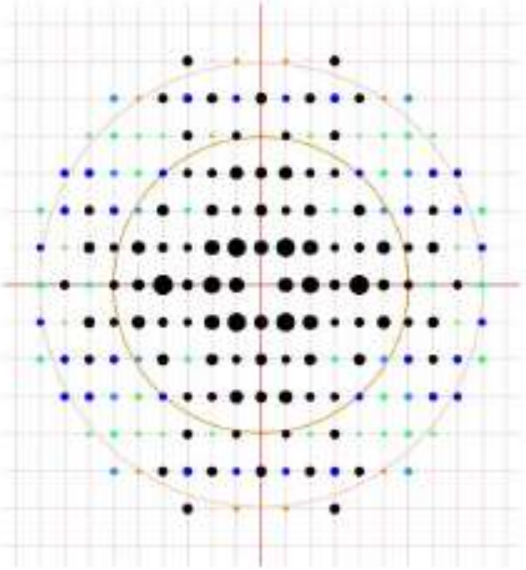
$$= \sum_{j}^{allatoms} f_j \left(\cos 2\pi\left(hx_j + lz_j\right) + i \sin 2\pi\left(hx_j + lz_j\right)\right)$$

Again it’s the translation of a half a unit cell that causes the absences.

$$= \sum_{j}^{halfatoms} 2 f_j \begin{pmatrix} \cos 2\pi\left(hx_j + lz_j\right) + \\ i \sin 2\pi\left(hx_j + lz_j\right) + \\ \cos 2\pi k\left(hx_j + l\left(z_j + \tfrac{1}{2}\right)\right) + \\ i \sin 2\pi k\left(hx_j + l\left(z_j + \tfrac{1}{2}\right)\right) \end{pmatrix}$$

P 2_1 2_1 2_1 absences





P 2_1/c Systematic absences

Reflection Conditions

$h0l$: $l = 2n$

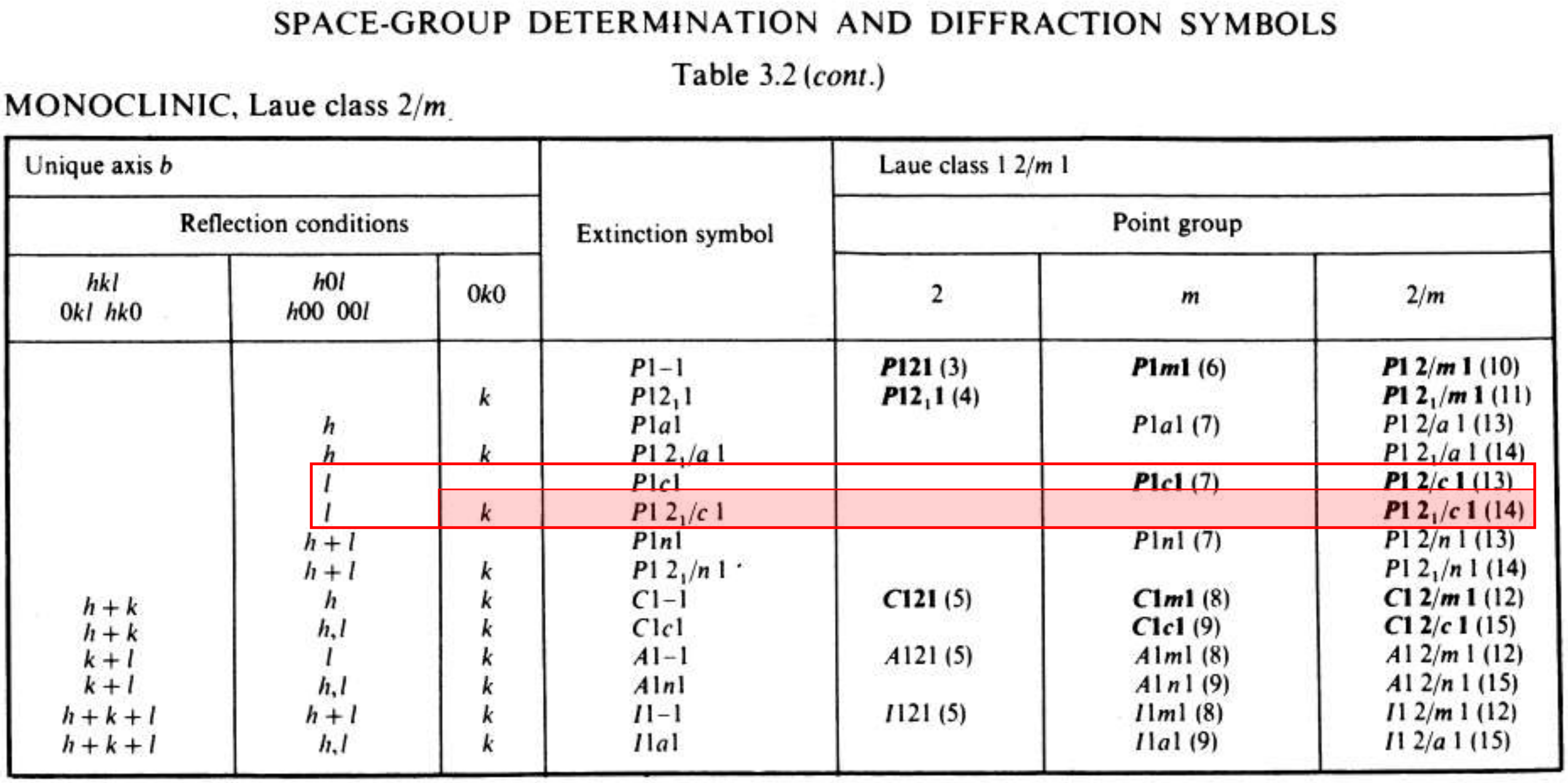
$0k0$: $k = 2n$

$00l$: $l = 2n$

2 non-zero indices - a zone. Tells us about 2D operators

1 non-zero index. Either a 'special case' of above, or tells us about 1D operators

International Tables



SPACE-GROUP DETERMINATION AND DIFFRACTION SYMBOLS

Table 3.2 (*cont.*)

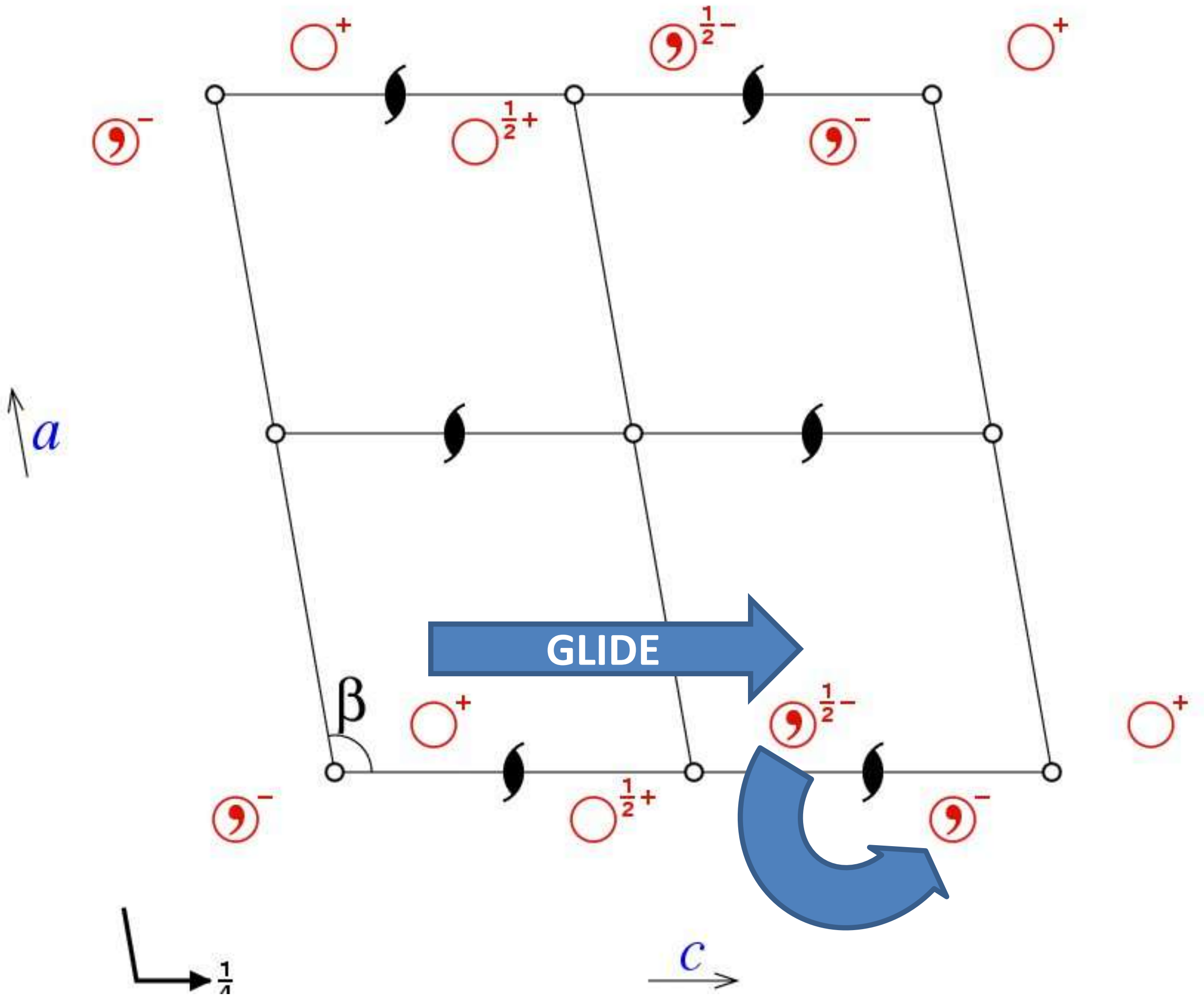
MONOCLINIC, Laue class $2/m$

Unique axis *b*				Laue class $1\ 2/m\ 1$		
Reflection conditions			Extinction symbol	Point group		
hkl *0kl hk0*	*h0l* *h00 00l*	*0k0*		2	*m*	$2/m$
			$P1-1$	**$P121$** (3)	**$P1m1$** (6)	**$P1\ 2/m\ 1$** (10)
		k	$P12_11$	**$P12_11$** (4)		**$P1\ 2_1/m\ 1$** (11)
	h		$P1a1$		$P1a1$ (7)	$P1\ 2/a\ 1$ (13)
	h	*k*	$P1\ 2_1/a\ 1$			$P1\ 2_1/a\ 1$ (14)
	l		$P1c1$		**$P1c1$** (7)	**$P1\ 2/c\ 1$** (13)
	l	*k*	$P1\ 2_1/c\ 1$			**$P1\ 2_1/c\ 1$** (14)
	$h+l$		$P1n1$		$P1n1$ (7)	$P1\ 2/n\ 1$ (13)
	$h+l$	*k*	$P1\ 2_1/n\ 1$			$P1\ 2_1/n\ 1$ (14)
$h+k$	*h*	*k*	$C1-1$	**$C121$** (5)	**$C1m1$** (8)	**$C1\ 2/m\ 1$** (12)
$h+k$	*h*,*l*	*k*	$C1c1$		**$C1c1$** (9)	**$C1\ 2/c\ 1$** (15)
$k+l$	*l*	*k*	$A1-1$	$A121$ (5)	$A1m1$ (8)	$A1\ 2/m\ 1$ (12)
$k+l$	*h*,*l*	*k*	$A1n1$		$A1n1$ (9)	$A1\ 2/n\ 1$ (15)
$h+k+l$	$h+l$	*k*	$I1-1$	$I121$ (5)	$I1m1$ (8)	$I1\ 2/m\ 1$ (12)
$h+k+l$	*h*,*l*	*k*	$I1a1$		$I1a1$ (9)	$I1\ 2/a\ 1$ (15)

h0l, l=2n; 0k0, k=2n; 00l, l=2n

Monoclinic, P 2_1/c

$P2_1/c$ $P\,1\,2_1/c\,1$ $2/m$



1 x, y, z

2 $\bar{x}, \frac{1}{2}+y, \frac{1}{2}-z$

3 $\bar{x}, \bar{y}, \bar{z}$

4 $x, \frac{1}{2}-y, \frac{1}{2}+z$

Orthorhombic example

Systematic Absences

h00, h=2n+1

0k0, k=2n+1

00l, l=2n+1

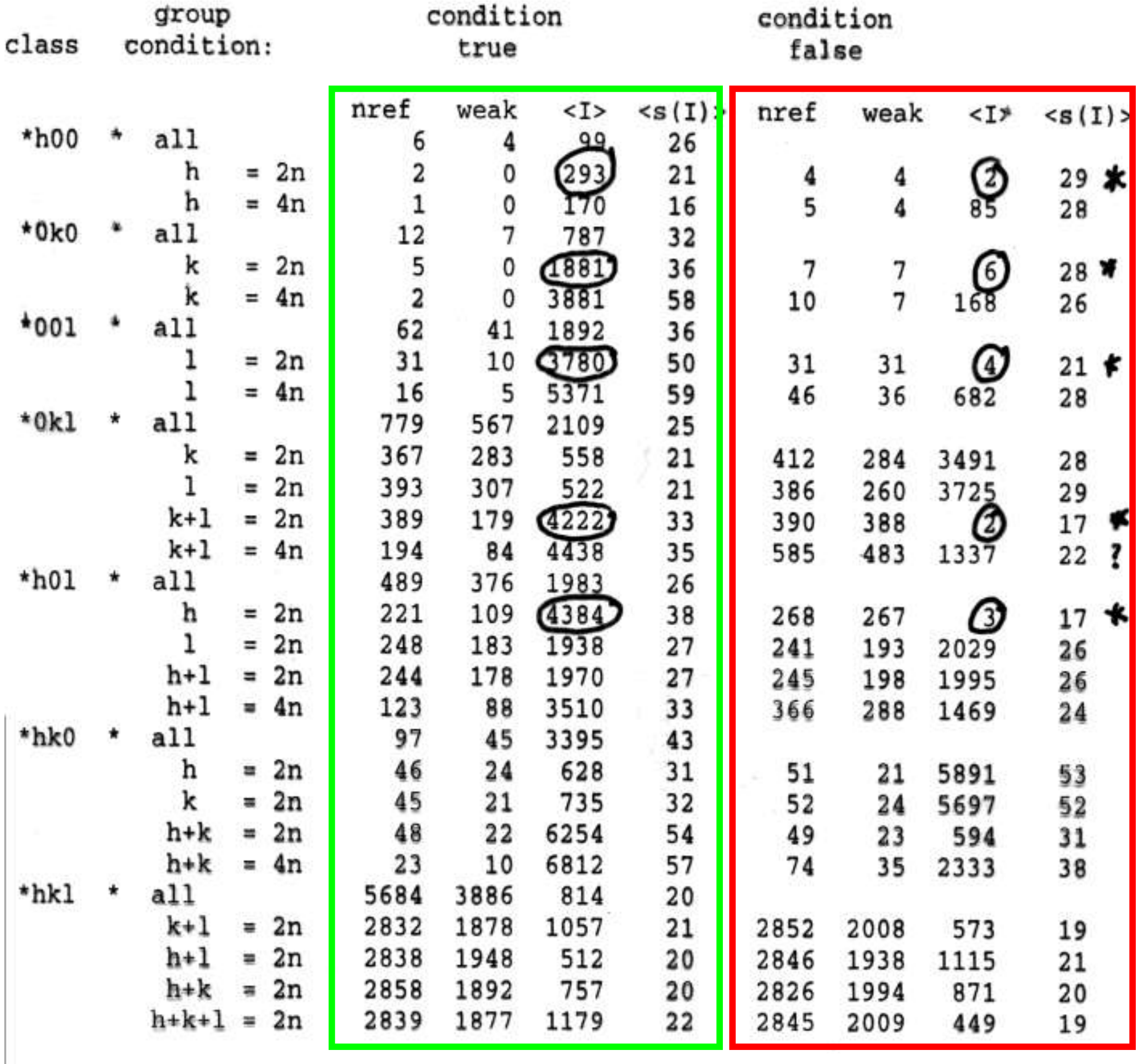
0kl, k+l=2n+1

h0l, h=2n+1

hk0, nothing

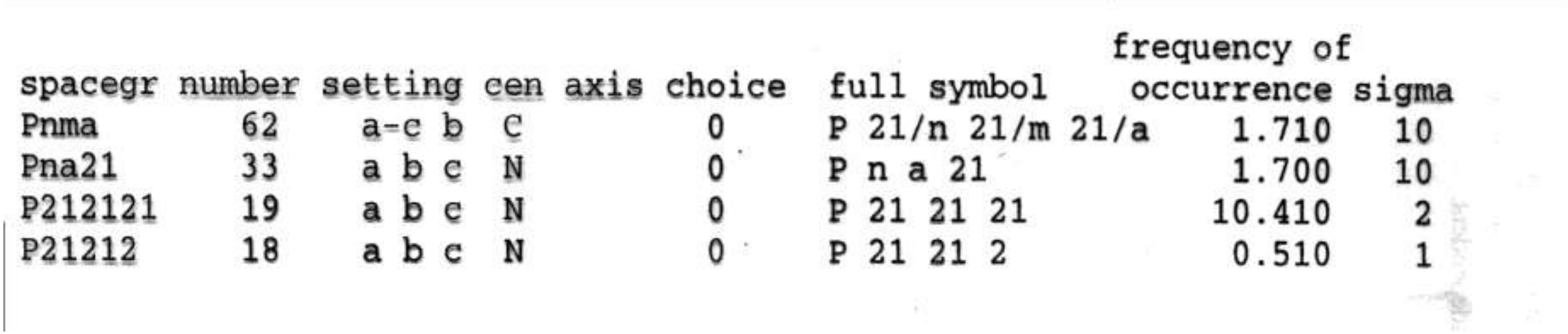
hkl, nothing

P n a m



class	group condition:	condition true nref	condition true weak	condition true <I>	condition true <s(I)>	condition false nref	condition false weak	condition false <I>	condition false <s(I)>	
*h00	* all	6	4	99	26					
	h = 2n	2	0	293	21	4	4	2	29	*
	h = 4n	1	0	170	16	5	4	85	28	
*0k0	* all	12	7	787	32					
	k = 2n	5	0	1881	36	7	7	6	28	*
	k = 4n	2	0	3881	58	10	7	168	26	
*00l	* all	62	41	1892	36					
	l = 2n	31	10	3780	50	31	31	4	21	*
	l = 4n	16	5	5371	59	46	36	682	28	
*0kl	* all	779	567	2109	25					
	k = 2n	367	283	558	21	412	284	3491	28	
	l = 2n	393	307	522	21	386	260	3725	29	
	k+l = 2n	389	179	4222	33	390	388	2	17	*
	k+l = 4n	194	84	4438	35	585	483	1337	22	?
*h0l	* all	489	376	1983	26					
	h = 2n	221	109	4384	38	268	267	3	17	*
	l = 2n	248	183	1938	27	241	193	2029	26	
	h+l = 2n	244	178	1970	27	245	198	1995	26	
	h+l = 4n	123	88	3510	33	366	288	1469	24	
*hk0	* all	97	45	3395	43					
	h = 2n	46	24	628	31	51	21	5891	53	
	k = 2n	45	21	735	32	52	24	5697	52	
	h+k = 2n	48	22	6254	54	49	23	594	31	
	h+k = 4n	23	10	6812	57	74	35	2333	38	
*hkl	* all	5684	3886	814	20					
	k+l = 2n	2832	1878	1057	21	2852	2008	573	19	
	h+l = 2n	2838	1948	512	20	2846	1938	1115	21	
	h+k = 2n	2858	1892	757	20	2826	1994	871	20	
	h+k+l = 2n	2839	1877	1179	22	2845	2009	449	19	

Possible Space Groups



```
                                                                    frequency of
spacegr number setting cen axis choice  full symbol       occurrence sigma
Pnma        62    a-c b  C         0    P 21/n 21/m 21/a       1.710    10
Pna21       33    a b c  N         0    P n a 21               1.700    10
P212121     19    a b c  N         0    P 21 21 21            10.410     2
P21212      18    a b c  N         0    P 21 21 2              0.510     1
```

Programs try to decide what classes of reflections are systematically absent, and then deduce the space group.

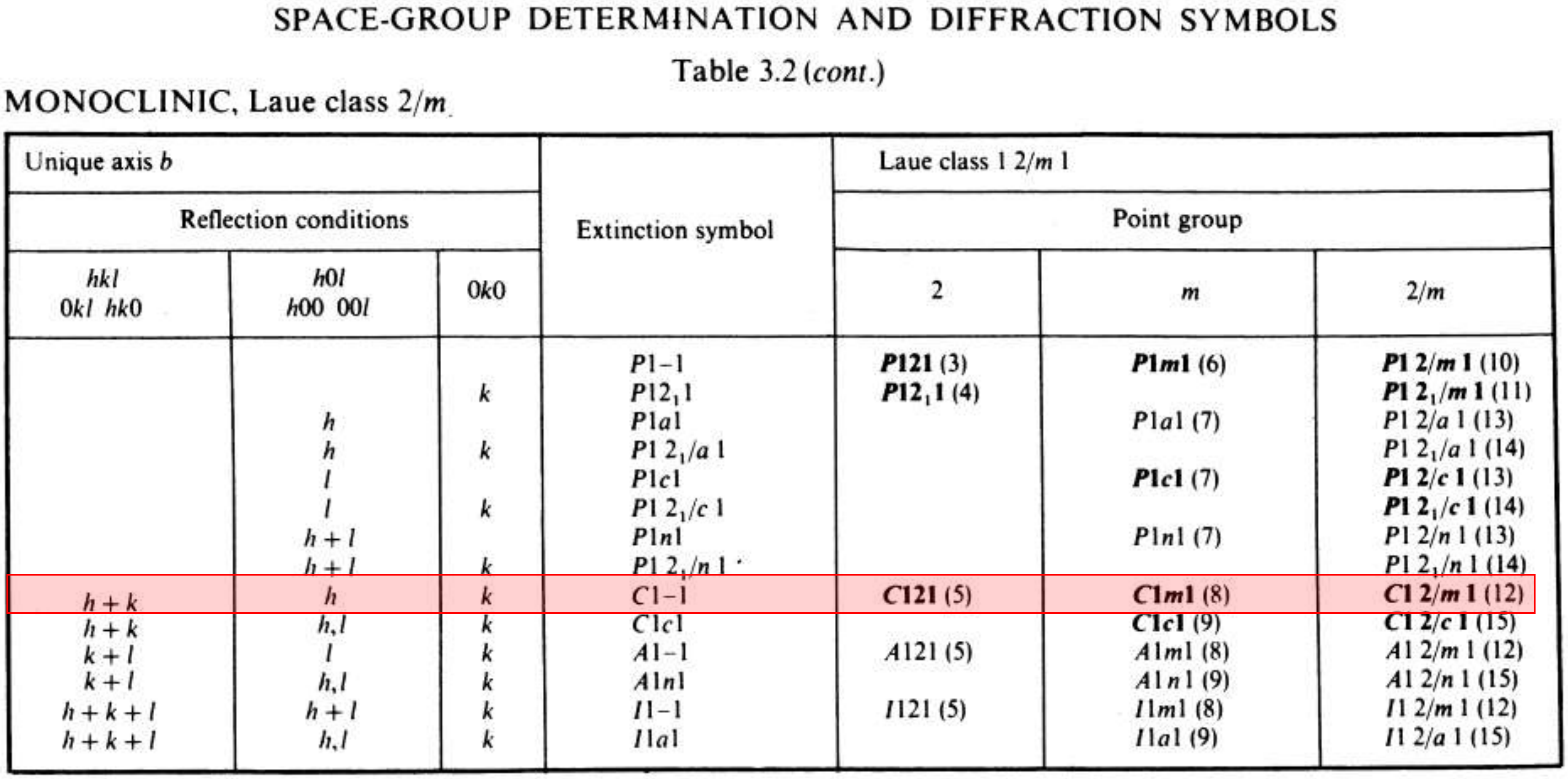
There are various sources of ambiguity.

Possible Space Groups

Sources of ambiguity include:

- The crystal quality may be poor so that some diffracted intensity occurs for ‘forbidden’ reflections (twinning?)
- The diffractometer may not have measured sufficient reflections to enable a sound decision to be made.
- More than one SG may have the same theoretical absences

International Tables



SPACE-GROUP DETERMINATION AND DIFFRACTION SYMBOLS

Table 3.2 (*cont.*)

MONOCLINIC, Laue class 2/*m*

Unique axis *b*			Extinction symbol	Laue class 1 2/*m* 1		
Reflection conditions				Point group		
hkl 0*kl* *hk*0	*h*0*l* *h*00 00*l*	0*k*0		2	*m*	2/*m*
			$P1{-}1$	**P121** (3)	**P1*m*1** (6)	**$P1\,2/m\,1$** (10)
		k	$P12_11$	**$P12_11$** (4)		**$P1\,2_1/m\,1$** (11)
	h		$P1a1$		$P1a1$ (7)	$P1\,2/a\,1$ (13)
	h	*k*	$P1\,2_1/a\,1$			$P1\,2_1/a\,1$ (14)
	l		$P1c1$		**P1*c*1** (7)	**$P1\,2/c\,1$** (13)
	l	*k*	$P1\,2_1/c\,1$			**$P1\,2_1/c\,1$** (14)
	$h+l$		$P1n1$		$P1n1$ (7)	$P1\,2/n\,1$ (13)
	$h+l$	*k*	$P1\,2_1/n\,1$			$P1\,2_1/n\,1$ (14)
$h+k$	*h*	*k*	$C1{-}1$	**C121** (5)	**C1*m*1** (8)	**$C1\,2/m\,1$** (12)
$h+k$	h,l	*k*	$C1c1$		**C1*c*1** (9)	**$C1\,2/c\,1$** (15)
$k+l$	*l*	*k*	$A1{-}1$	$A121$ (5)	$A1m1$ (8)	$A1\,2/m\,1$ (12)
$k+l$	h,l	*k*	$A1n1$		$A1n1$ (9)	$A1\,2/n\,1$ (15)
$h+k+l$	$h+l$	*k*	$I1{-}1$	$I121$ (5)	$I1m1$ (8)	$I1\,2/m\,1$ (12)
$h+k+l$	h,l	*k*	$I1a1$		$I1a1$ (9)	$I1\,2/a\,1$ (15)

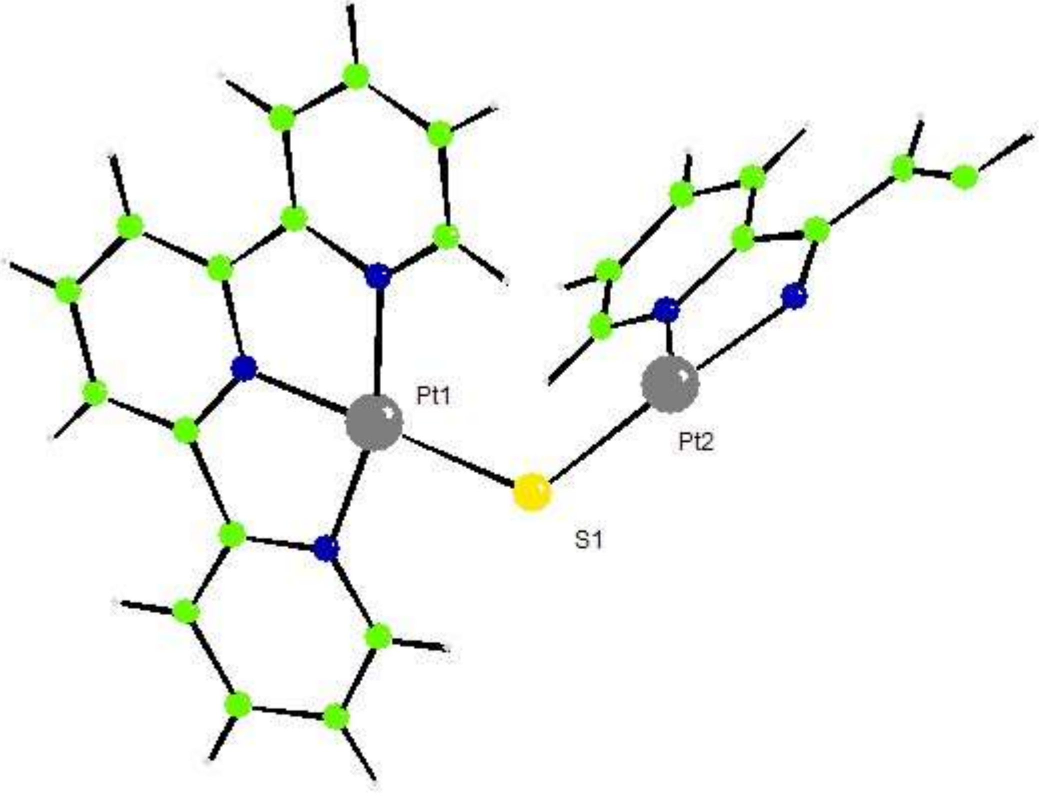
The highlighted systematic absences can be due to a choice of three space groups

Why does a chemist need to understand crystallographic symmetry?

1. To solve a crystal structure!
2. To see a whole structure when the asymmetric unit is only part of a molecule
3. To see the structure if the material is an extended lattice compound
4. To see interactions between molecules in the solid state, e.g. hydrogen bonding.

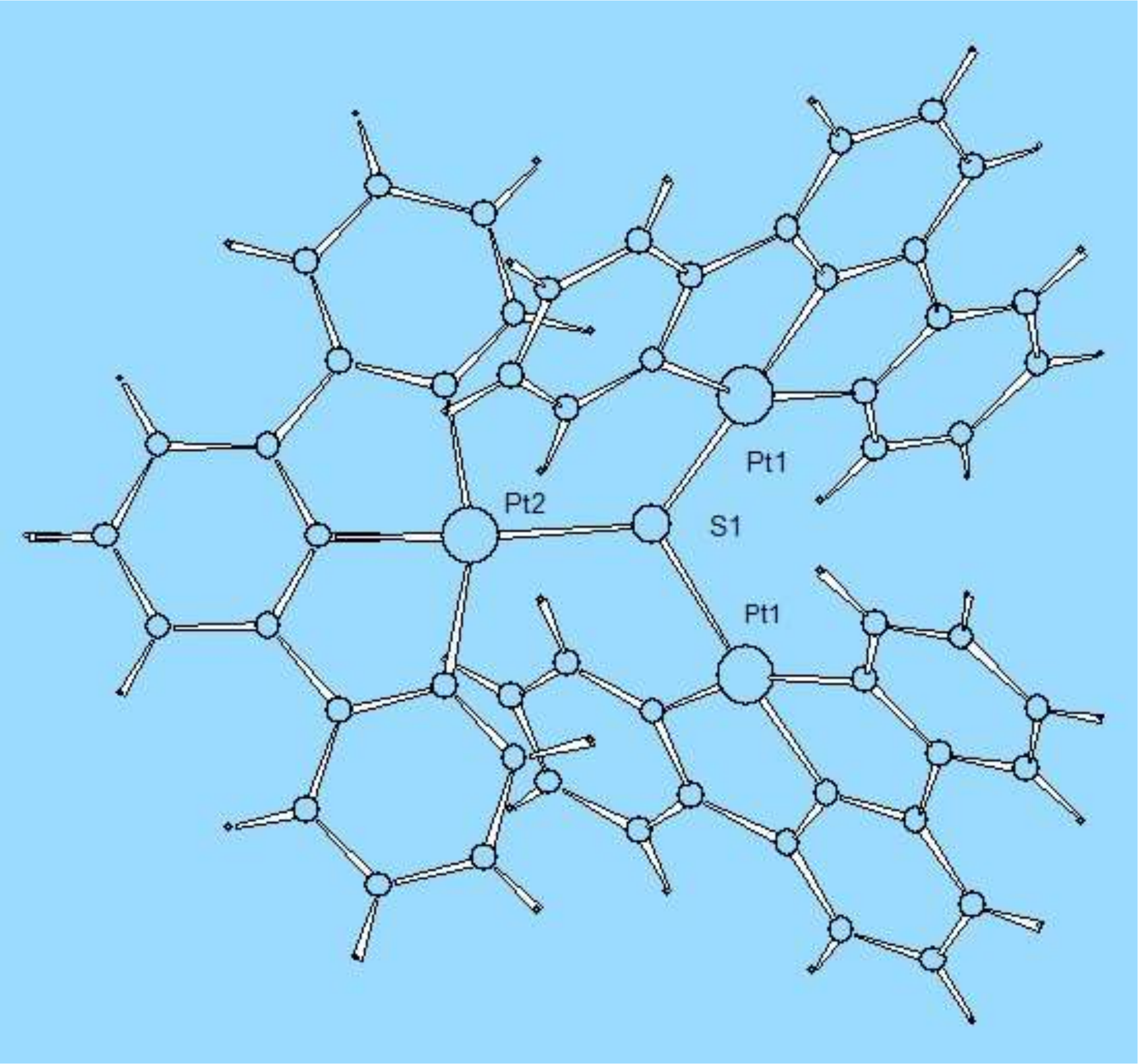
Example: ½ molecule in asymmetric unit

The crystallographic entity is ½ of the molecule. The second half (which shares the S) is generated by crystallographic symmetry.



Example: The full molecule

The completed molecule has a propeller shape



Extended Lattice Structures

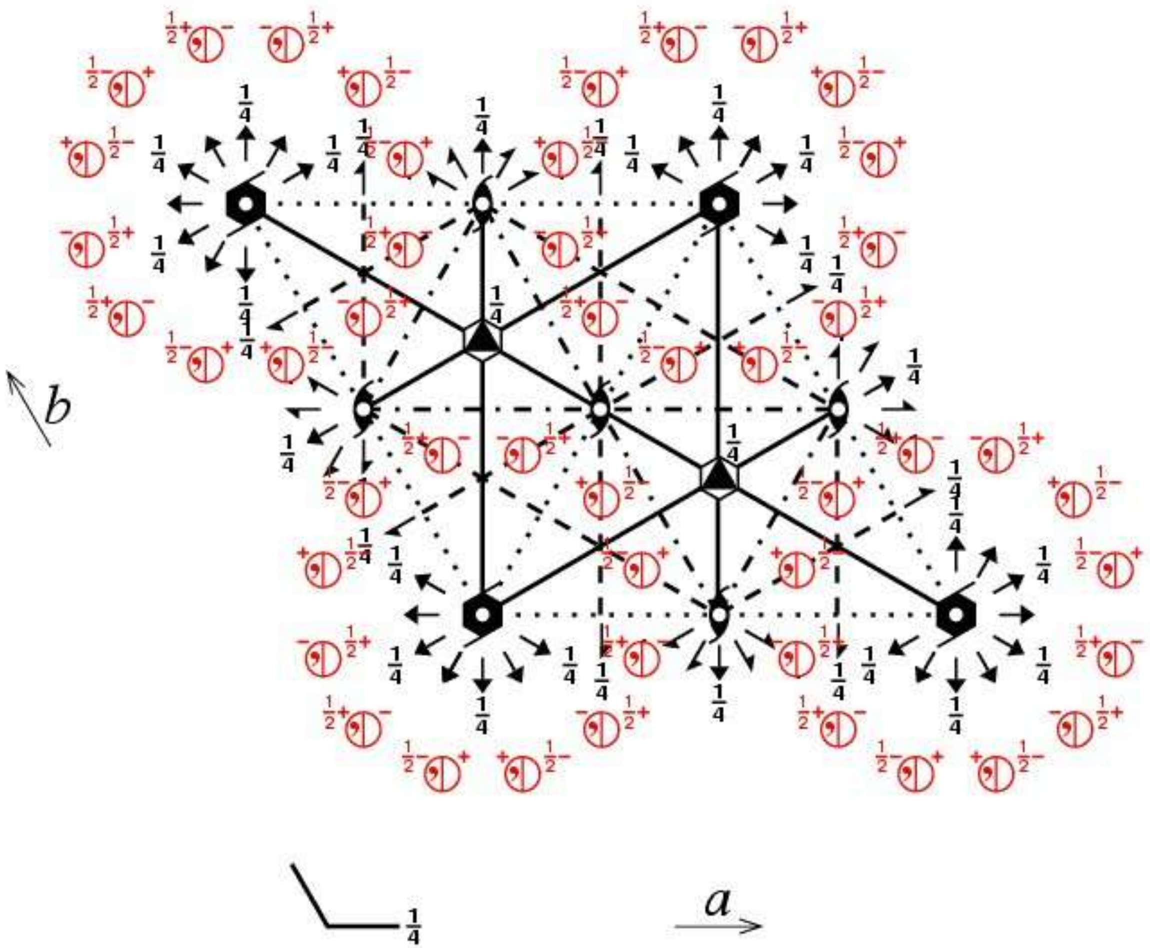
Minerals and synthetic mineral-like materials often have a very simple asymmetric unit, but a complicated space-group with very many internal symmetry operators. For example, the SG $\boldsymbol{F\,d\,\bar{3}\,c}$ (cubic, No 228) has 192 symmetry operators.

For every atom in the asymmetric unit, 191 others are generated by symmetry to fill the unit cell.

In general, software helps with the visualisation of these high symmetry structures.

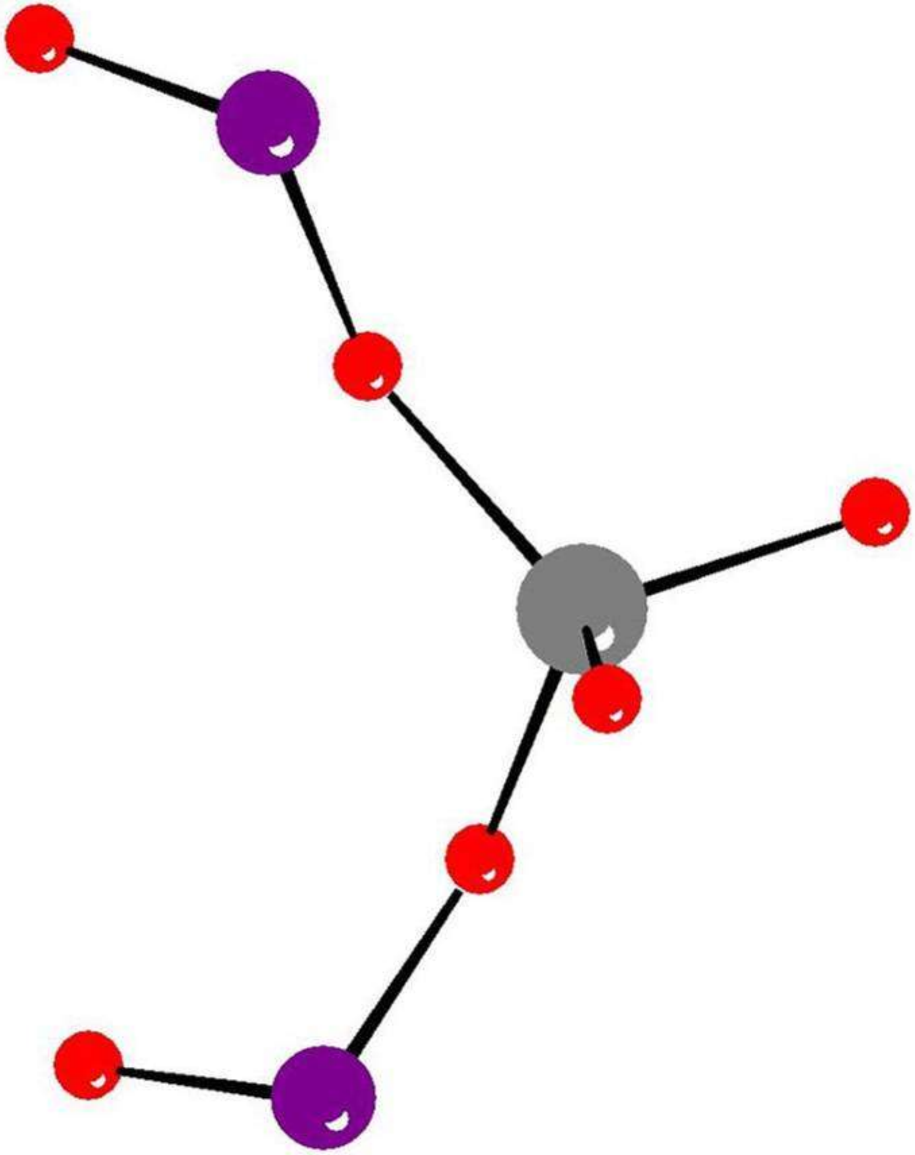
Hexagonal Symmetry

$P6_3/mmc$ $P\,6_3/m\;2/m\;2/c$ $6/mmm$ No. 194



1 x, y, z
2 $\bar{y}, x-y, z$
3 $\bar{x}+y, \bar{x}, z$
4 $\bar{x}, \bar{y}, \frac{1}{2}+z$
5 $x-y, x, \frac{1}{2}+z$
6 $y, \bar{x}+y, \frac{1}{2}+z$
7 $\bar{y}, \bar{x}, z$
8 $\bar{x}+y, y, z$
9 $x, x-y, z$
10 $y, x, \frac{1}{2}+z$
11 $x-y, \bar{y}, \frac{1}{2}+z$
12 $\bar{x}, \bar{x}+y, \frac{1}{2}+z$
13 $\bar{x}, \bar{y}, \bar{z}$
14 $y, \bar{x}+y, \bar{z}$
15 $x-y, x, \bar{z}$
16 $x, y, \frac{1}{2}-z$
17 $\bar{x}+y, \bar{x}, \frac{1}{2}-z$
18 $\bar{y}, x-y, \frac{1}{2}-z$
19 $y, x, \bar{z}$
20 $x-y, \bar{y}, \bar{z}$
21 $\bar{x}, \bar{x}+y, \bar{z}$
22 $\bar{y}, \bar{x}, \frac{1}{2}-z$
23 $\bar{x}+y, y, \frac{1}{2}-z$
24 $x, x-y, \frac{1}{2}-z$

Extended Lattice



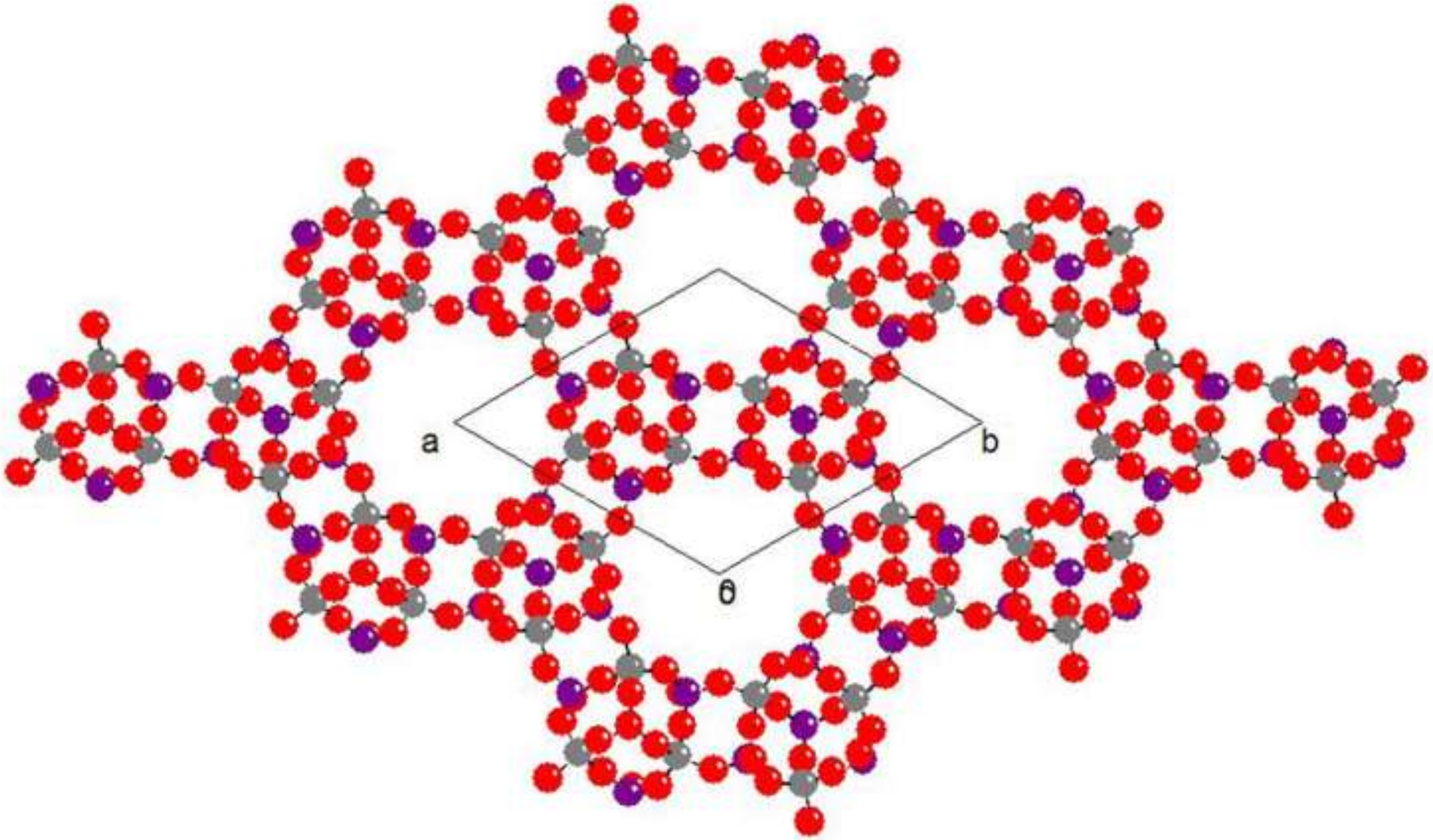
The asymmetric unit consists of Al O_6 P_2

The oxygen with only one bond are in fact bridging to other atoms in the unit cell.

Applying the SG symmetry operators fills out the structure.

Extended Lattice

The space group is $\boldsymbol{P\ \bar{3}}$. The ‘holes’ in the structure are in fact channels running all through the crystal, which can contain other molecules



Symmetry

Types of symmetry: point symmetry (restricted by unit cell requirements) + translational modifications

Space Groups: 230 possible combinations of symmetry in 3D space.

Effect on Structure Factor: absences in zones, rows.

Absences help to determine space group.

Why you should care: need SG to solve structure; intermolecular interactions, packing, extended solids.

Structure Solution

To find the atoms, we need to compute the electron density at every point in the cell

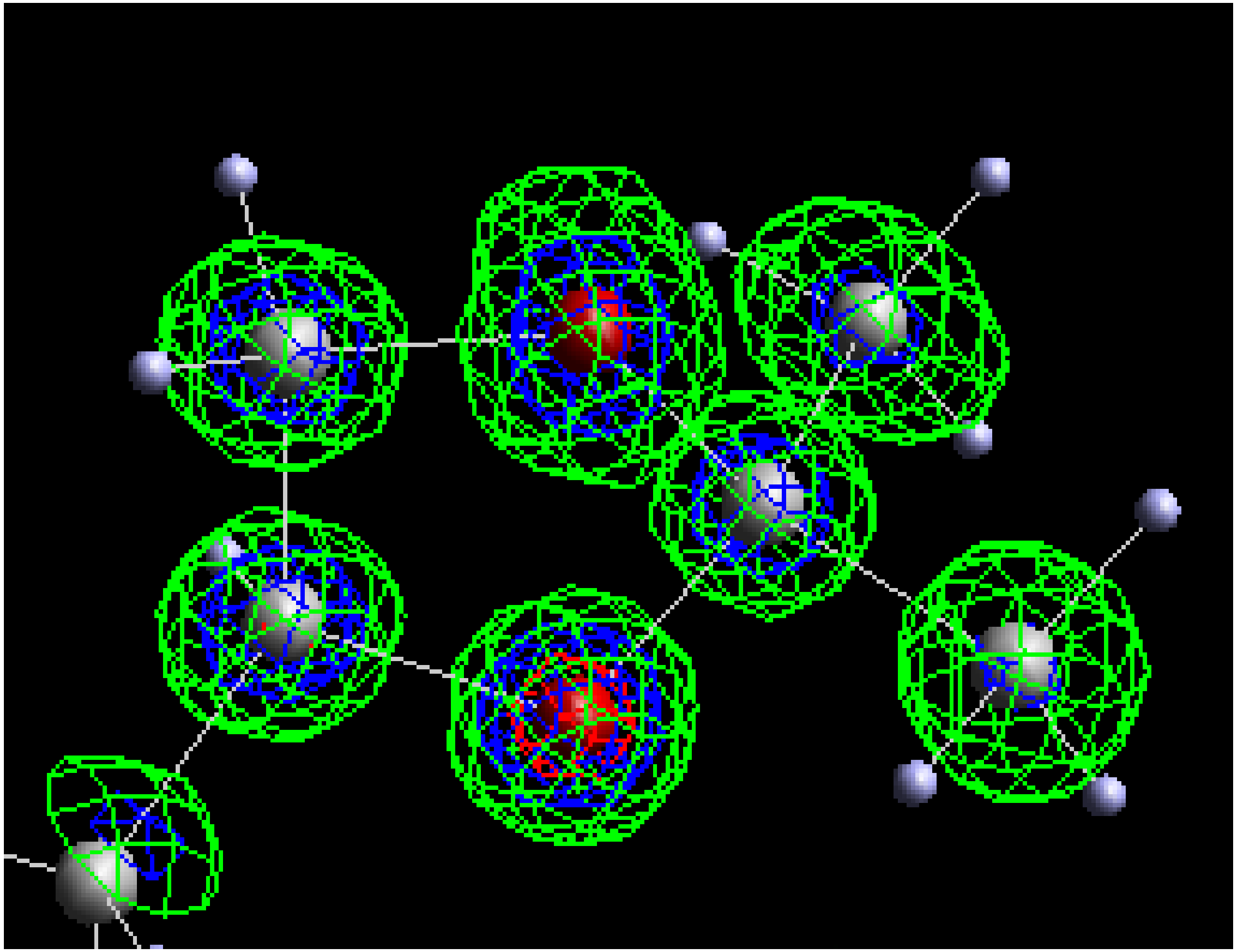
To compute the electron density, we need the phases.

$$\rho_{xyz} = \frac{1}{v}\sum\sum\sum |F|_{hkl}\, e^{-2\pi i(hx+ky+lz-\alpha_{hkl})}$$

We can observe the magnitudes, but not the phases.

The phases have to be estimated by some non-experimental method

Actual electron density.
Atoms are placed at regions of high density.



How do we know we have the right phases?

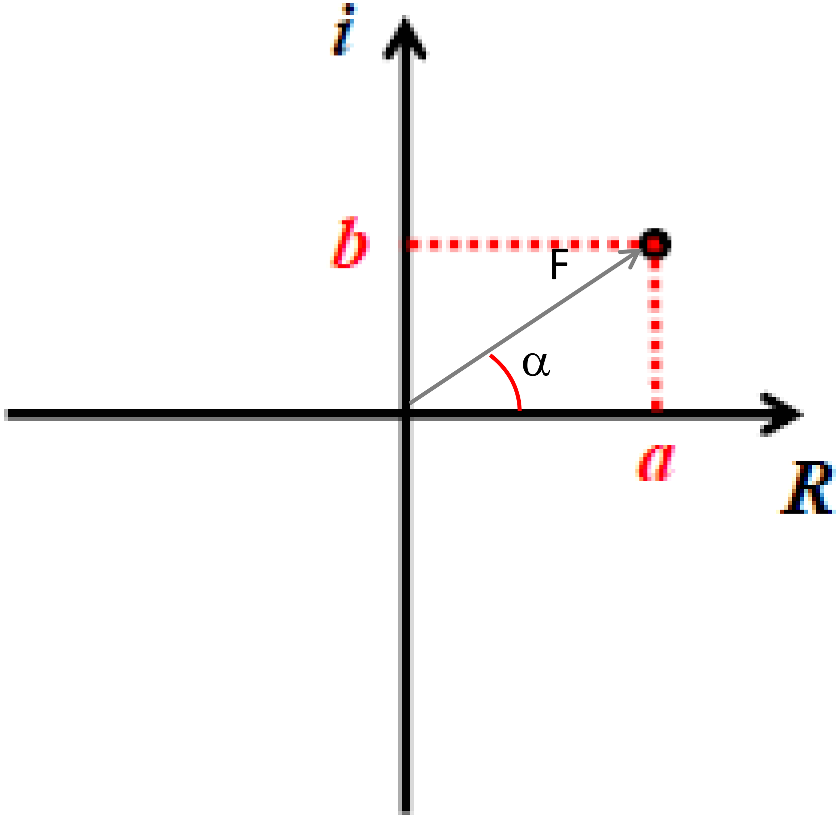
- The electron density map can be interpreted in terms of atoms with reasonable geometric relationships.
- A packing diagram contains neither molecular clashes nor voids.
- The structure factors (Fc) computed from the model are all similar to the observations (Fo).
- Low R factor. $R = \sum \left| Fo - Fc \right| / \sum Fo$
- The ‘figures of merit’ are good.

The Structure Factor

$$Fc_h = \sum_j^{allatoms} f_j \exp^{2\pi i\left(hx_j + ky_j + lz_j\right)}$$

$$A_h = \sum_j^{allatoms} f_j \cos 2\pi\left(hx + ky + lz\right)_j$$

$$B_h = \sum_j^{allatoms} f_j \sin 2\pi\left(hx + ky + lz\right)_j$$



$$F_c^{\,2} = A^2 + B^2 \qquad \tan\alpha = \frac{B}{A} = \frac{\sum \sin 2\pi hx}{\sum \cos 2\pi hx}$$

Direct Methods

Although the phase cannot be observed for individual reflections, the phases of groups of reflections are to some extent related.

These statistical relationships are derived from some loose constraints:

1. the electron density is never negative
2. the local peaks are isolated from each other
3. the density is distributed fairly randomly throughout the cell.

Tangent Formula

(in complex space – the complex plane)

Strong relationships exist between the phases of reflections which form a triplet.

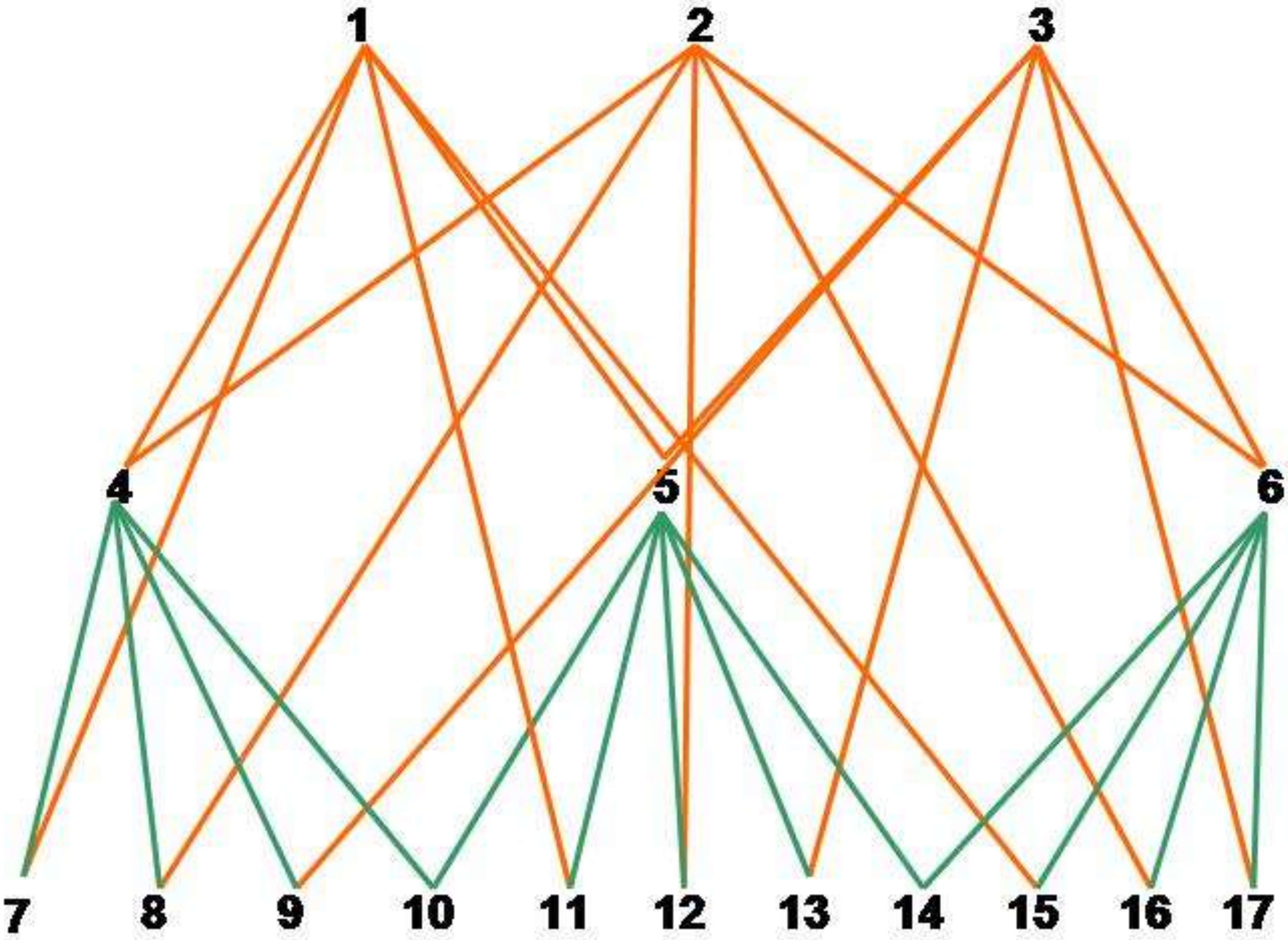
The Tangent Formula relates the phases of reflections in a triplet.

If we know the phase of h_1 and of h_2-h_1, we can compute the phase of h_2

$$\tan(\alpha_{h2}) \approx \frac{E_{h1}.E_{h2-h1}.\sin\left(\alpha_{h1}+\alpha_{h2-h1}\right)}{E_{h1}.E_{h2-h1}.\cos\left(\alpha_{h1}+\alpha_{h2-h1}\right)}$$

The **E** are the *normalised structure factors*

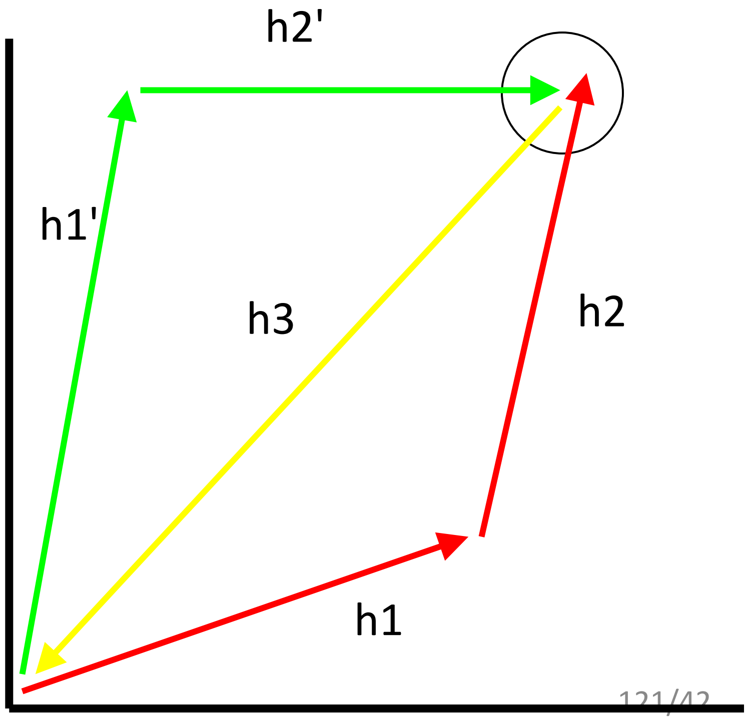
Phase Extension



Figures of Merit

Many pairs of reflections *h1* and *h2*, will give the phase of the reflection *h3*.

How well these different predictions agree can be used to assess the probability that the phase is correct. This information can be combined for all the reflections as a 'figure of merit' for the whole phase extension procedure.



Getting Started

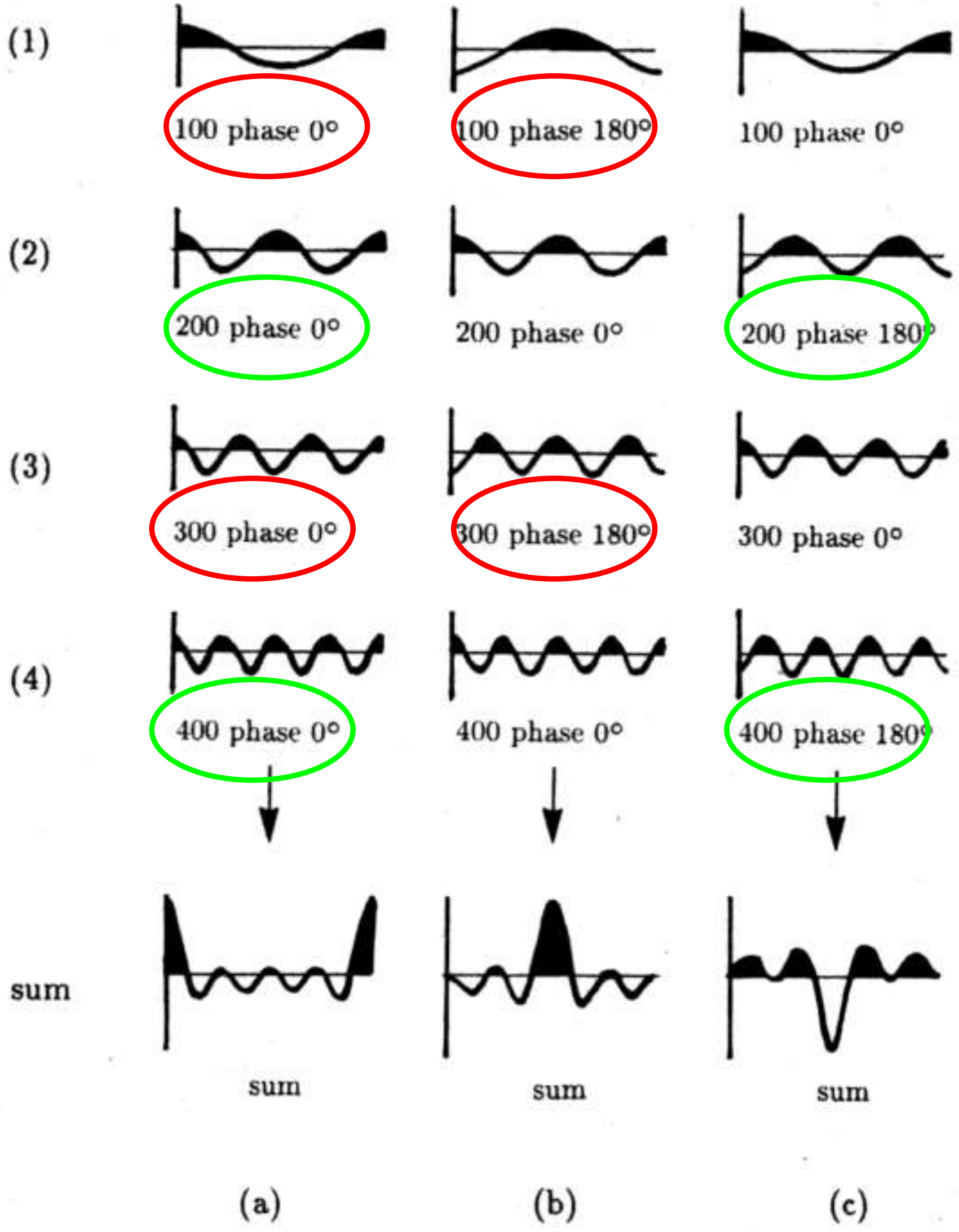
- Given the phase of reflections *h1* and *h2*, we can compute the phase of *h3*.
- The problem is to get these starting phases.
- In practice, a few reflections with large E-values are selected and given arbitrary phases. The Tangent formula is used to propagate these phases, and figures of merit are computed.
- A new set of phases are chosen, and the process repeated. The set with the best figure of merit is used to compute a Fourier Synthesis.

Phase Permutation

The phases of the starting reflections are permutated to cover all of the 'phase space'. This depends upon the space group.

Try	ϕ_1	ϕ_2	ϕ_3	fom
1	0	0	0	0.12
2	0	0	45	0.34
	---	---	---	
7	0	45	45	0.75
8	0	45	90	0.26
	---	---	---	
25	0	180	180	0.38
26	45	0	0	0.84
27	45	0	45	0.53
	---	---	---	
41	45	135	0	0.92
	---	---	---	
124	180	180	135	0.27
125	180	180	180	0.06

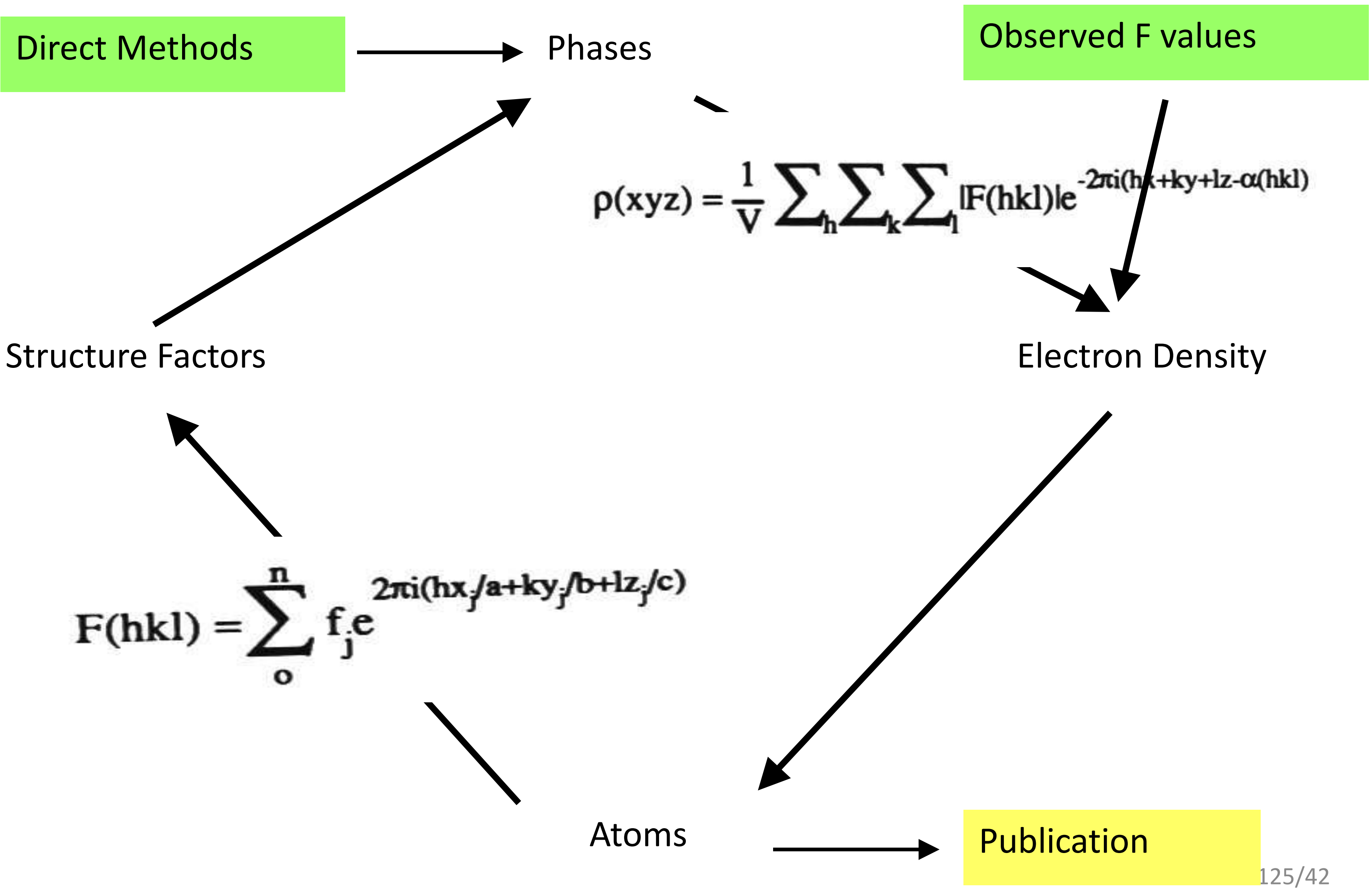
Random Phases



Sets (a) and (b) produce atoms.

Set (c) produces noise with a deep negative region.

Structure Development



Direct Methods

Phases

Observed F values

$$\rho(xyz) = \frac{1}{V}\sum_h \sum_k \sum_l |F(hkl)| e^{-2\pi i(hx+ky+lz-\alpha(hkl))}$$

Structure Factors

Electron Density

$$F(hkl) = \sum_{o}^{n} f_j e^{2\pi i(hx_j/a+ky_j/b+lz_j/c)}$$

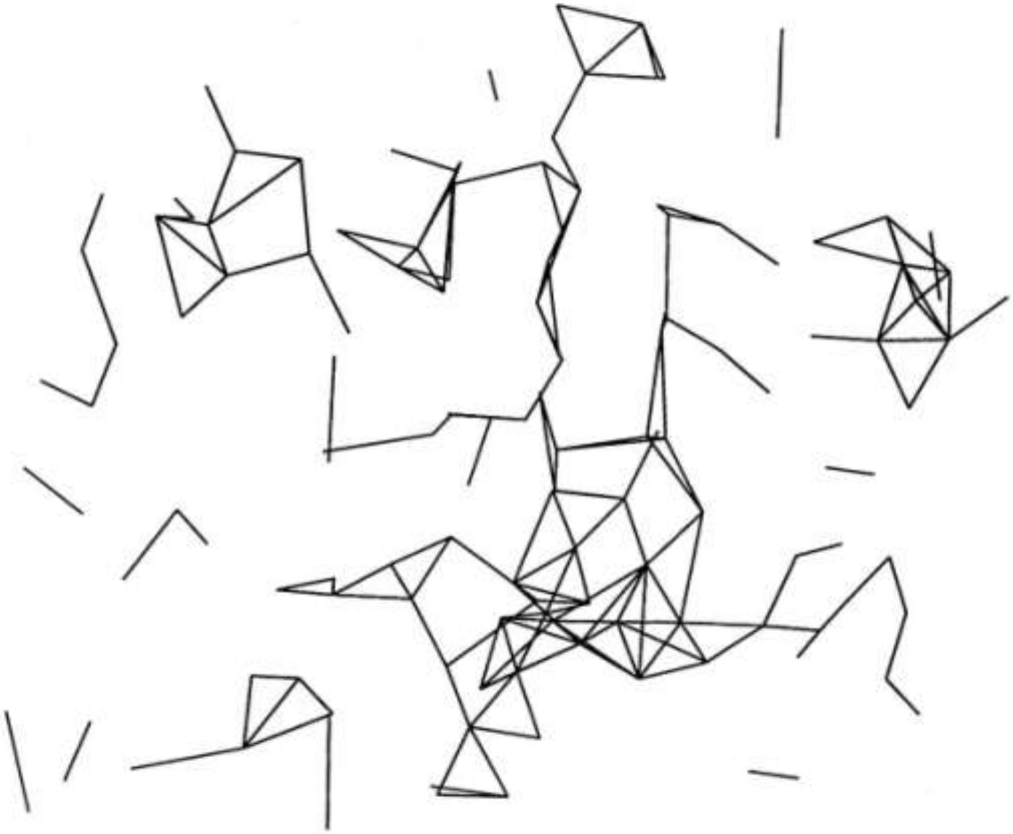
Atoms

Publication

Charge Flipping Demo

Penicillin-V

Initial atom map

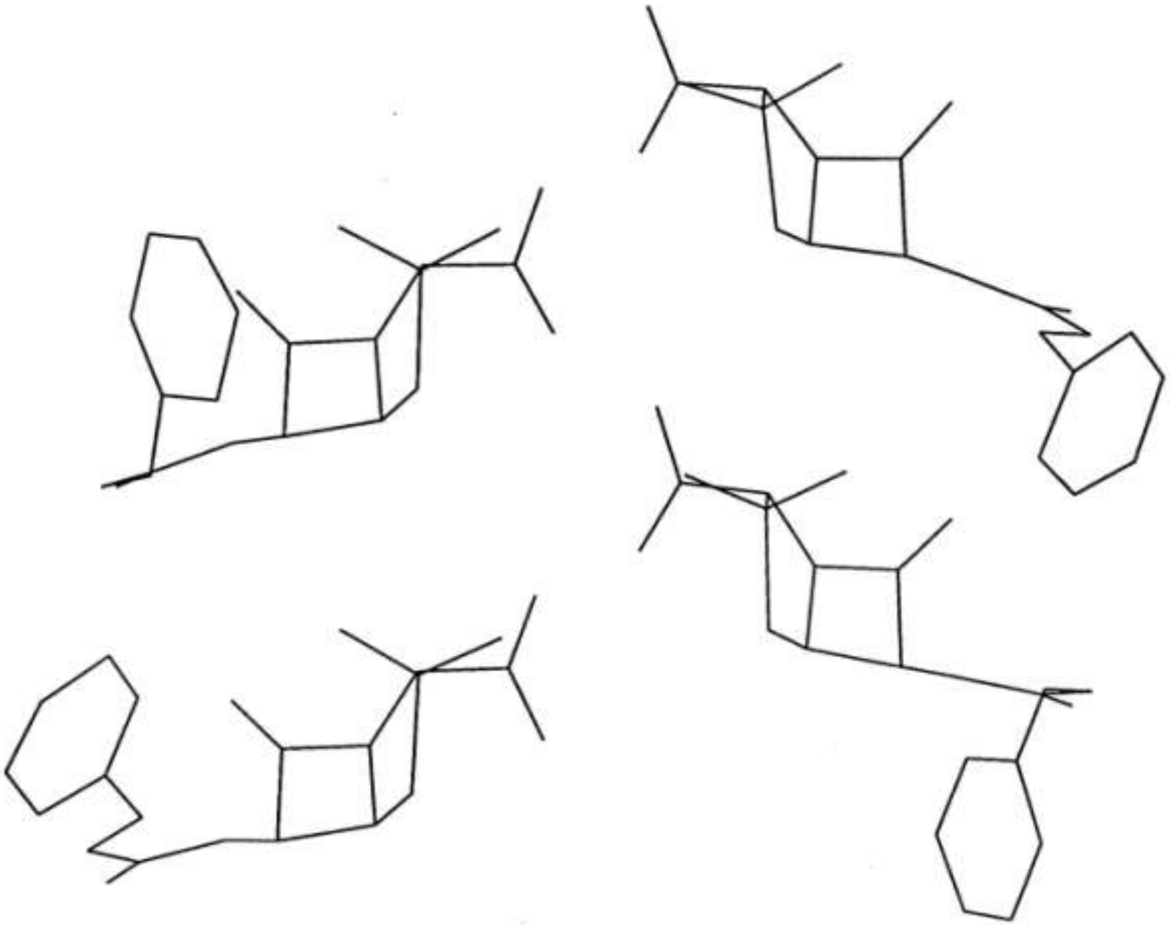


K Pen V - Direct Methods Structure

Penicillin-V

Final structure

Note the two levels of pseudo-symmetry

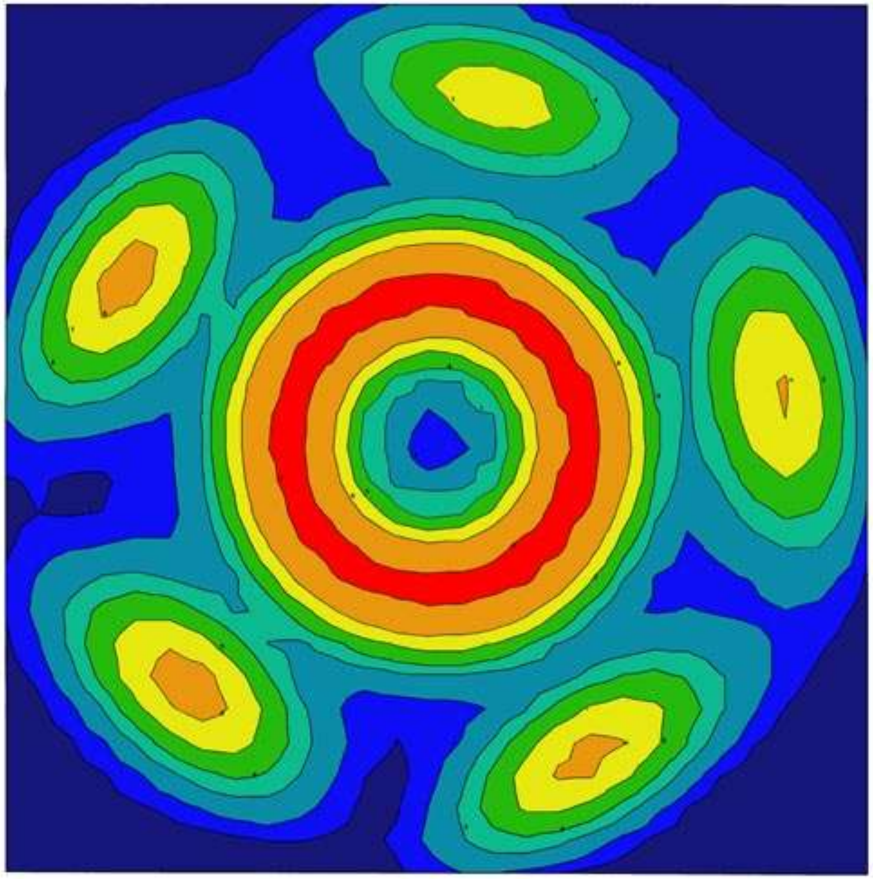


K Pen V - Final Structure

Refinement

The atoms can only be located approximately in the electron density map.

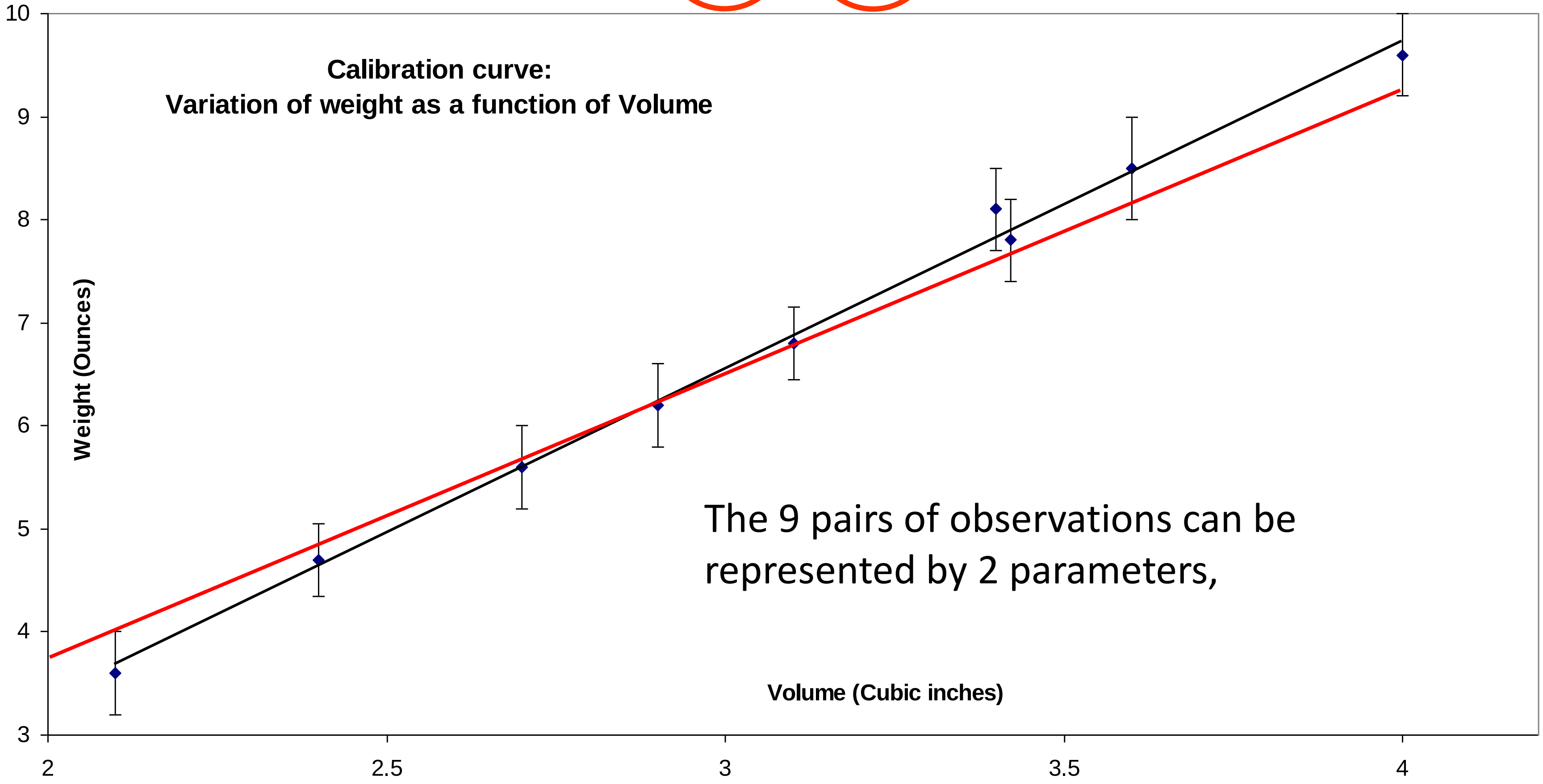
Their optimal positions (and adps) are determined by Least Squares



Straight Line

A straight trend line passes quite close to all the observations. A good digital representation of the data is

$$y = \boxed{3.2}x - \boxed{3.0}$$



The 9 pairs of observations can be represented by 2 parameters,

Linear Least Squares

$$a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = y_1$$

The observable quantity *y* is linearly dependent upon the unknown quantities *x.*

We make many observation of *y* under different conditions *a*.

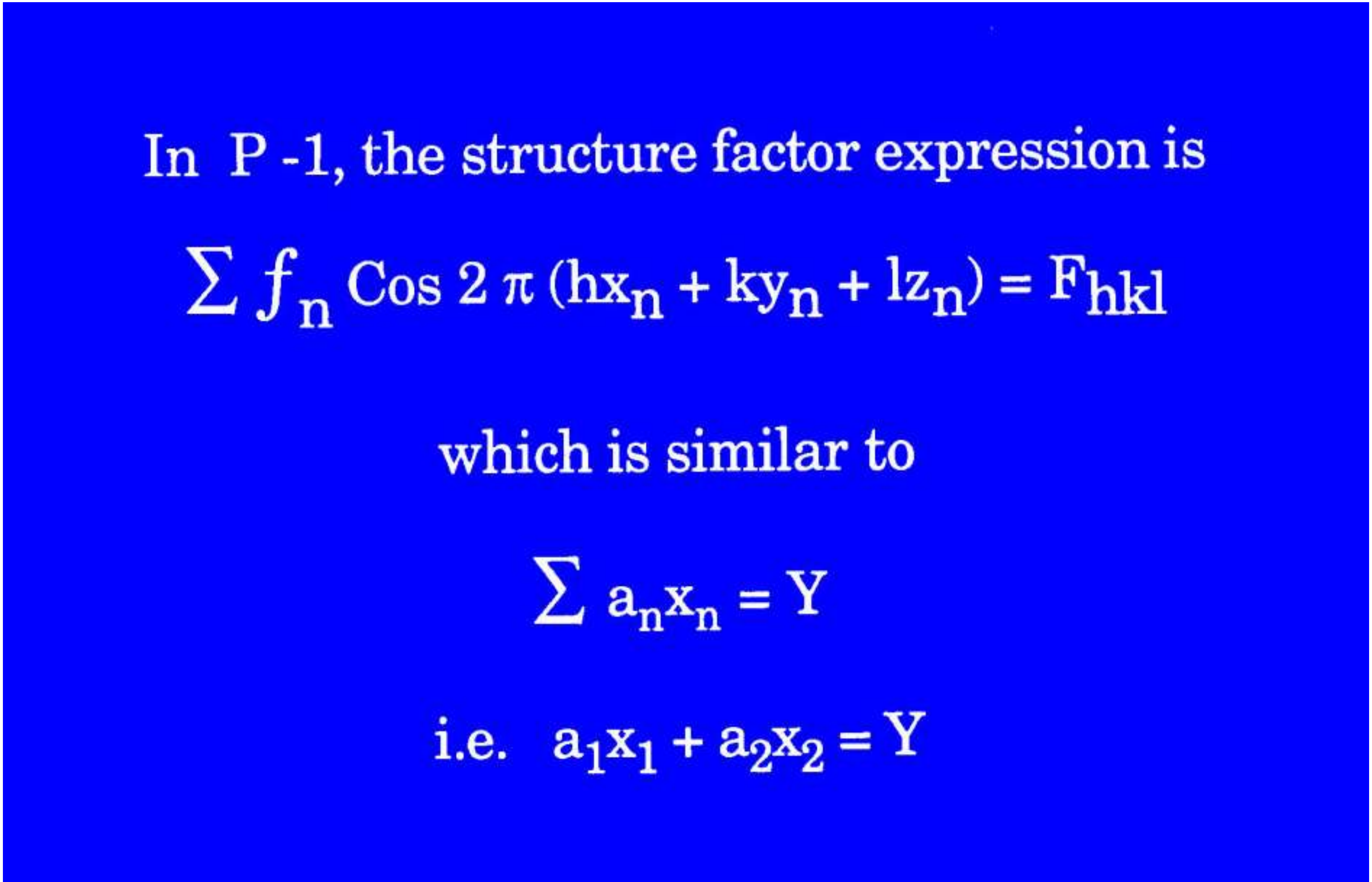
We can generally choose the values of *a.*

Linear Least Squares

$$a_{11}x_1 + a_{12}x_2 + a_{13}x_3 = y_{calc} \approx y_{obs}$$

We adjust the values of x_i to minimise the sum of the square of the differences between $y_{observed}$ and $y_{calculated}$ summed over all the observations

Crystallographic Least Squares



In P -1, the structure factor expression is

$$\sum f_n \cos 2\pi (hx_n + ky_n + lz_n) = F_{hkl}$$

which is similar to

$$\sum a_n x_n = Y$$

i.e. $a_1 x_1 + a_2 x_2 = Y$

Non Linear Least Squares

$$\sum_{i}^{unknowns} a_i \,.\, x_i = y$$

$$\sum^{atoms} f_i e^{-T} Cos\left(\underset{\sim}{h} \,.\, \underset{\sim}{x}\right) = Fc \approx Fo$$

Use Taylor Expansion

$$Fc + \sum_{i}^{parameters} \frac{\partial Fc}{\partial x_i} \,.\, \delta x_i = F'c \approx Fo$$

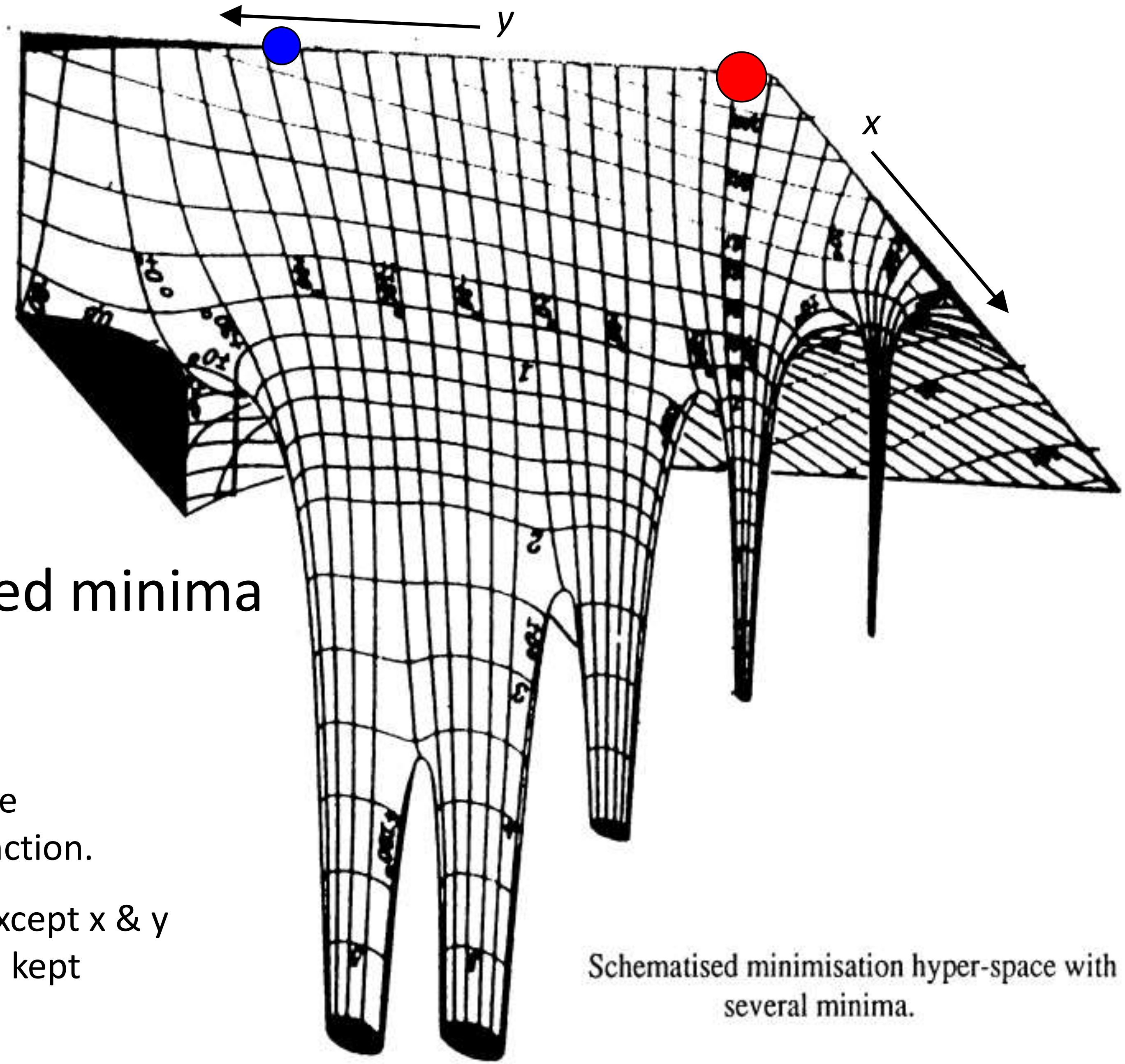
$$\sum_{i}^{parameters} \frac{\partial Fc}{\partial x_{\ i}} \,.\, \delta x_i \approx Fo - Fc$$

Note that in non-linear least squares, **we determine improvements in parameters**, but not the parameters themselves.

Problems with Least Squares

1. Refinement is to a local minimum (next slide)
2. We need 'good' estimates of the starting parameters
3. We compute improvements to the parameters
4. The shifts are only approximate, so it takes many cycles of refinement
5. We get no information about parameters that have been omitted from the model

Schematised minima



Schematised minimisation hyper-space with several minima.

Vertical axis is the minimisation function.

All parameters except x & y for one atom are kept constant.

Crystallographic Least Squares

We adjust the values of x_i to minimise the sum of the square of the differences between F_o and F_c, or between their squares.

F refinement

$$M_1 = \sum_{i}^{reflections} w_i \left(F_o - F_c\right)^2$$

F^2 refinement

$$M_2 = \sum_{i}^{reflections} w_i \left(F_o^2 - F_c^2\right)^2$$

R factors

The crystallographic R factors are one-parameter estimates of reliability.

$$R = \frac{\sum^{data}\left(\left|F_o\right| - \left|F_c\right|\right)}{\sum^{data}\left|F_o\right|} \qquad wR_2 = \sqrt{\frac{\sum^{data} w\left(F_o^{\,2} - F_c^{\,2}\right)^2}{\sum^{data} w\left(F_o^{\,2}\right)^2}}$$

R_2 is always much higher than R

Weights

The observations do not all have the same accuracy or precision.

We can decrease the significance of reflections we are unsure about by giving them a low weight in the refinement.

The weights may be related to the esds (sus) of the reflections, or simply be estimated mathematically

Goodness of Fit

At the end of a refinement, the ‘goodness of fit’, S, is estimated from

$$\sqrt{\frac{\sum w\left(F_o - F_c\right)^2}{n-m}} \approx 1.$$

w is usually adjusted to make S equal to unity.

Refineable Parameters

1. The overall scale of the diffraction intensities
2. *x,y,z* the positions of the atoms
3. Uiso or Uaniso, the atomic displacement parameters (adps)
4. Extinction (an attenuating effect)
5. The Flack parameter – this is concerned with the absolute configuration.
6. Twin fractions

Problem Structures

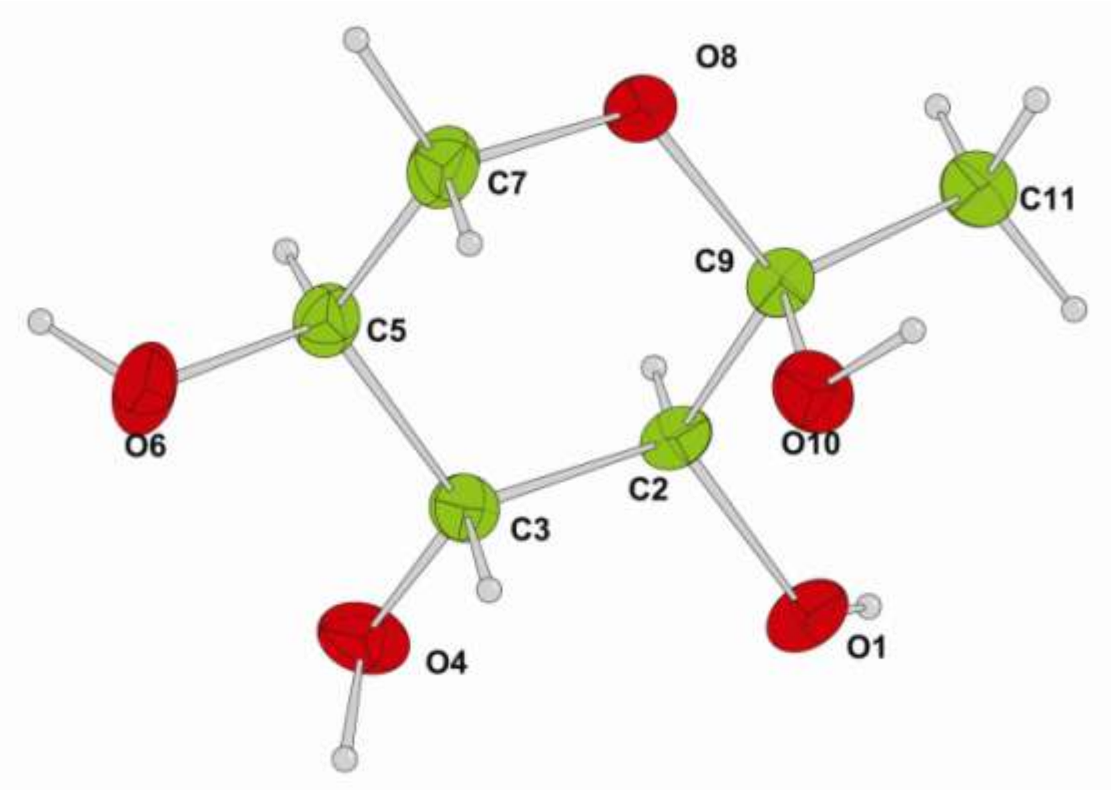
Somewhere between 20% and 80% of all structures collected have a problem which will need experience to resolve.

Problems may include:

- Poor quality or small crystals
- Disorder – the atoms are not precisely located
- Not the expected material
- Twinning
- Z' >1 structures

DEMO

Fleet Compound NAJ 251



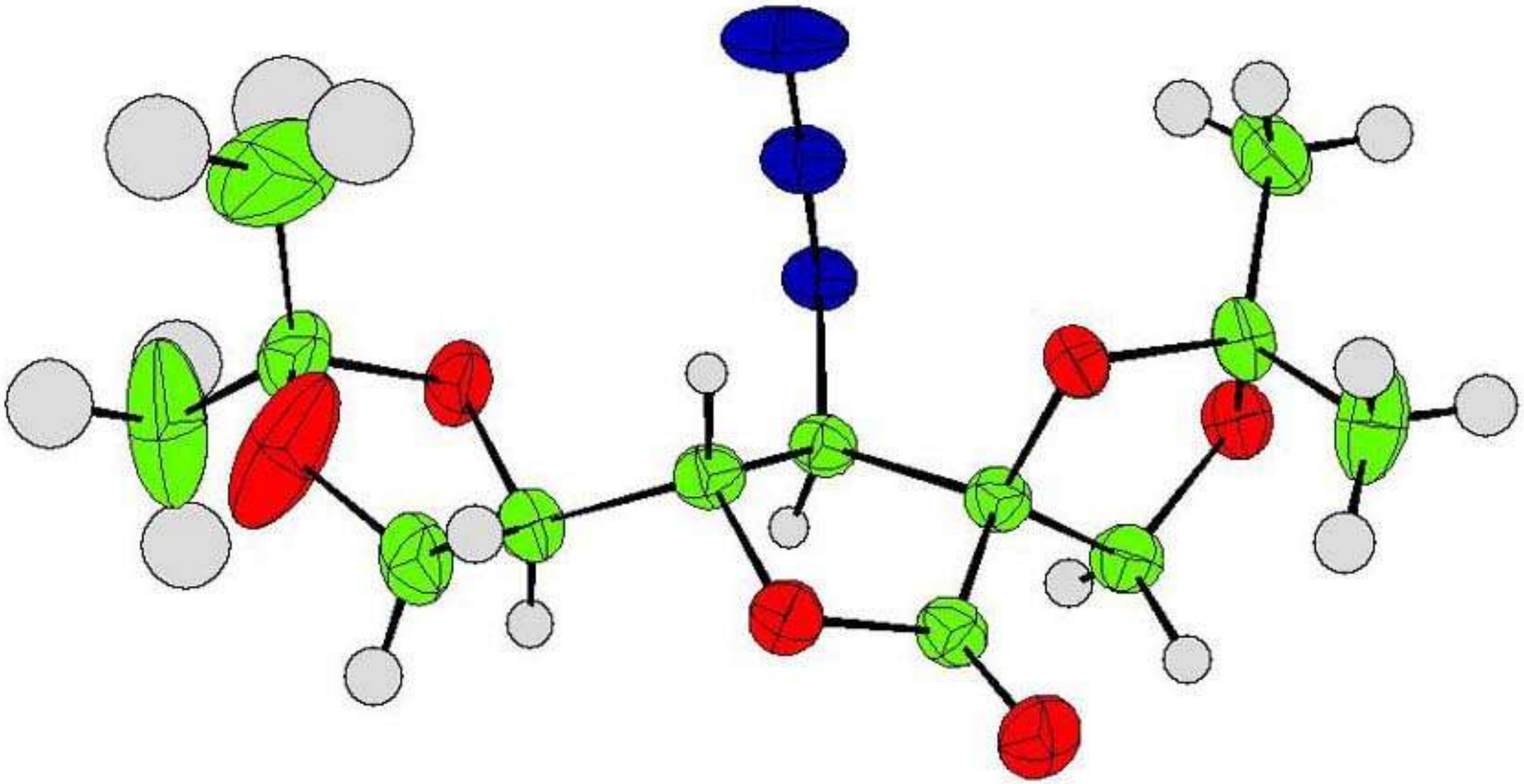
Examples of:

- Disorder
- Interpreting electron density
- Z’ > 1
- Hydrogen bonding
- Squeeze

Disordered Azide

(Soengas , Harding, Shallard-Brown)

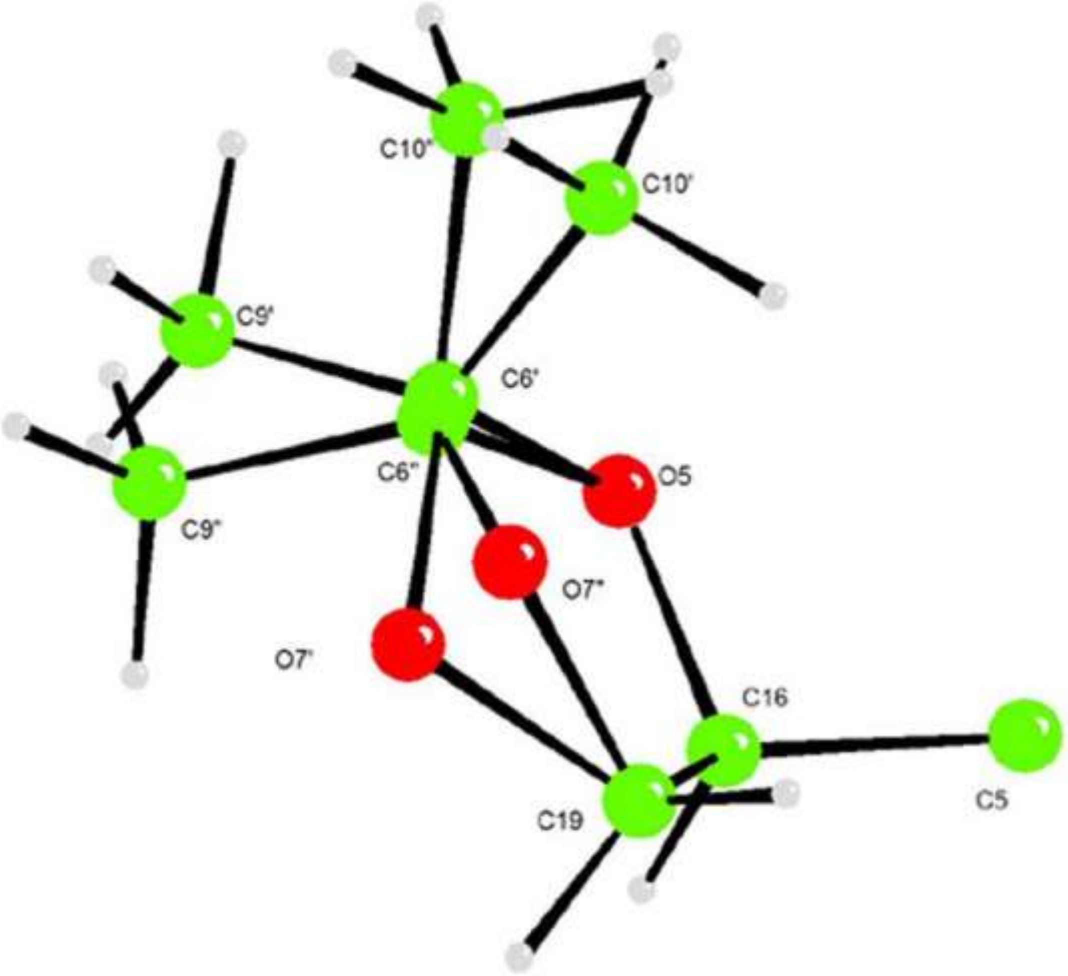
- Note that the left hand end has high adps



Disordered Azide

(Soengas , Harding, Shallard-Brown)

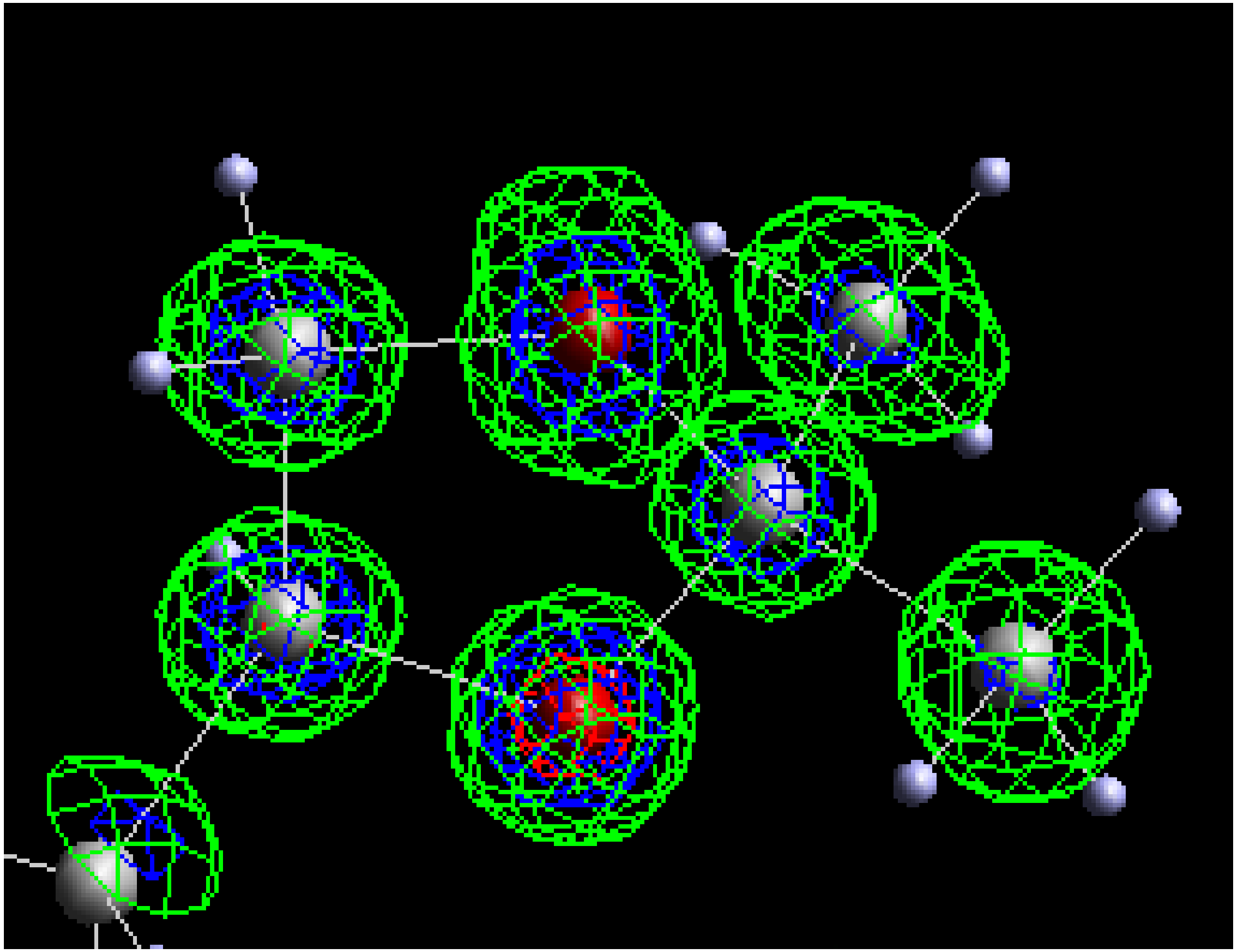
The disordered part modelled as ‘split’ atoms



Disordered Azide

(Soengas , Harding, Shallard-Brown)

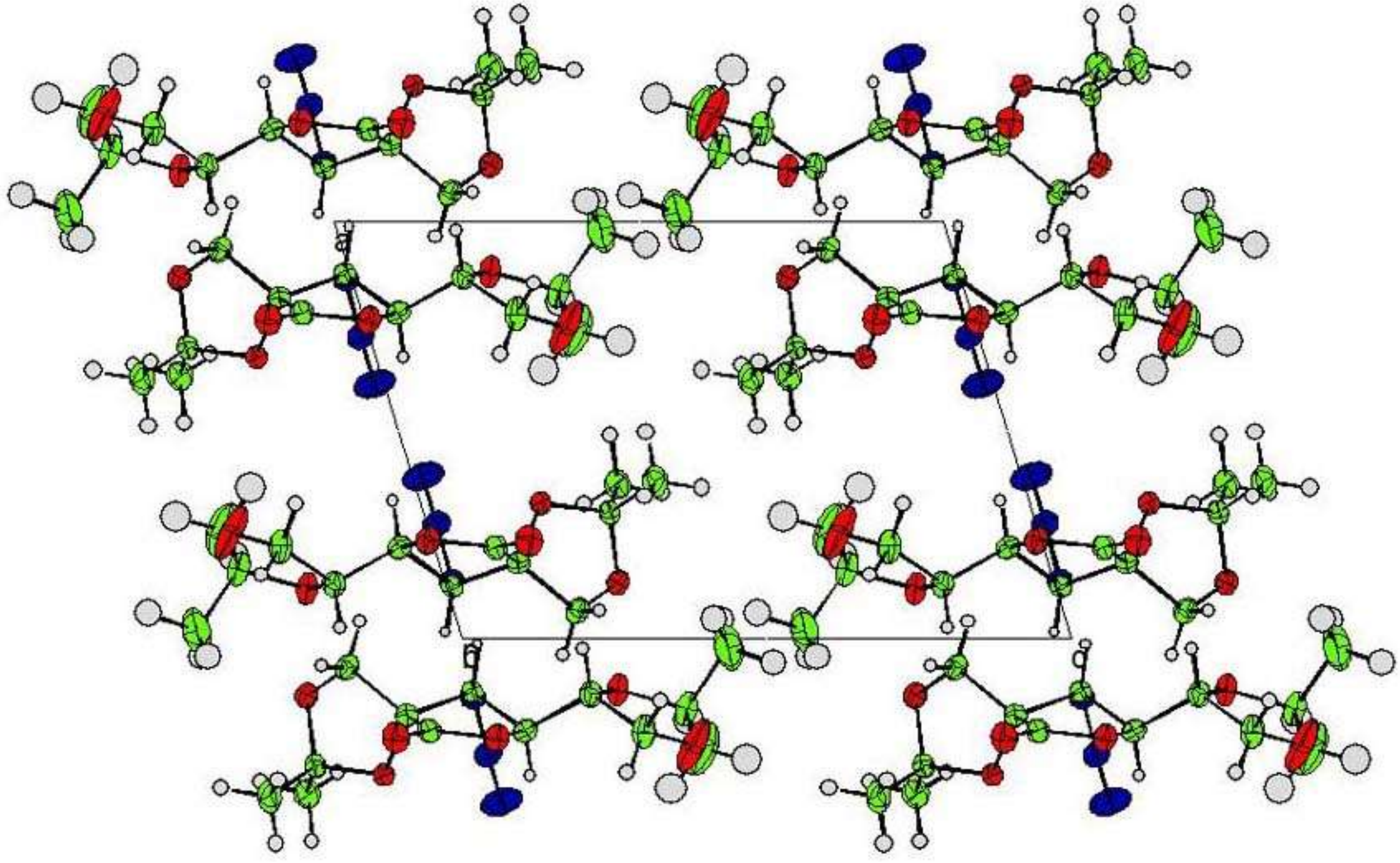
Actual electron density. Note elongations.



Disordered Azide

(Soengas , Harding, Shallard-Brown)

Note alternating ordered and disordered region.



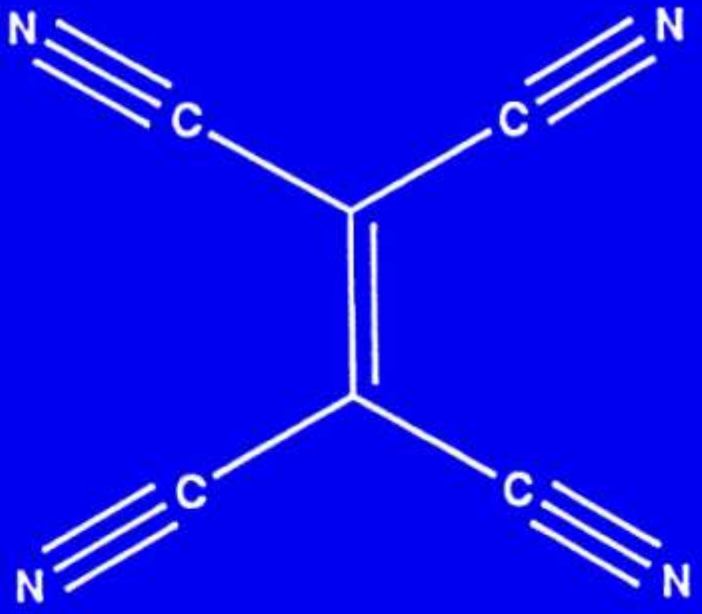
Interpretation Of Density Maps

Generally, crystallographers allow the computer to look for atoms in electron density maps (peak searching).

Occasionally, it is useful to look at maps manually

TCNE ligand

TCNE, tetracyanoethylene =



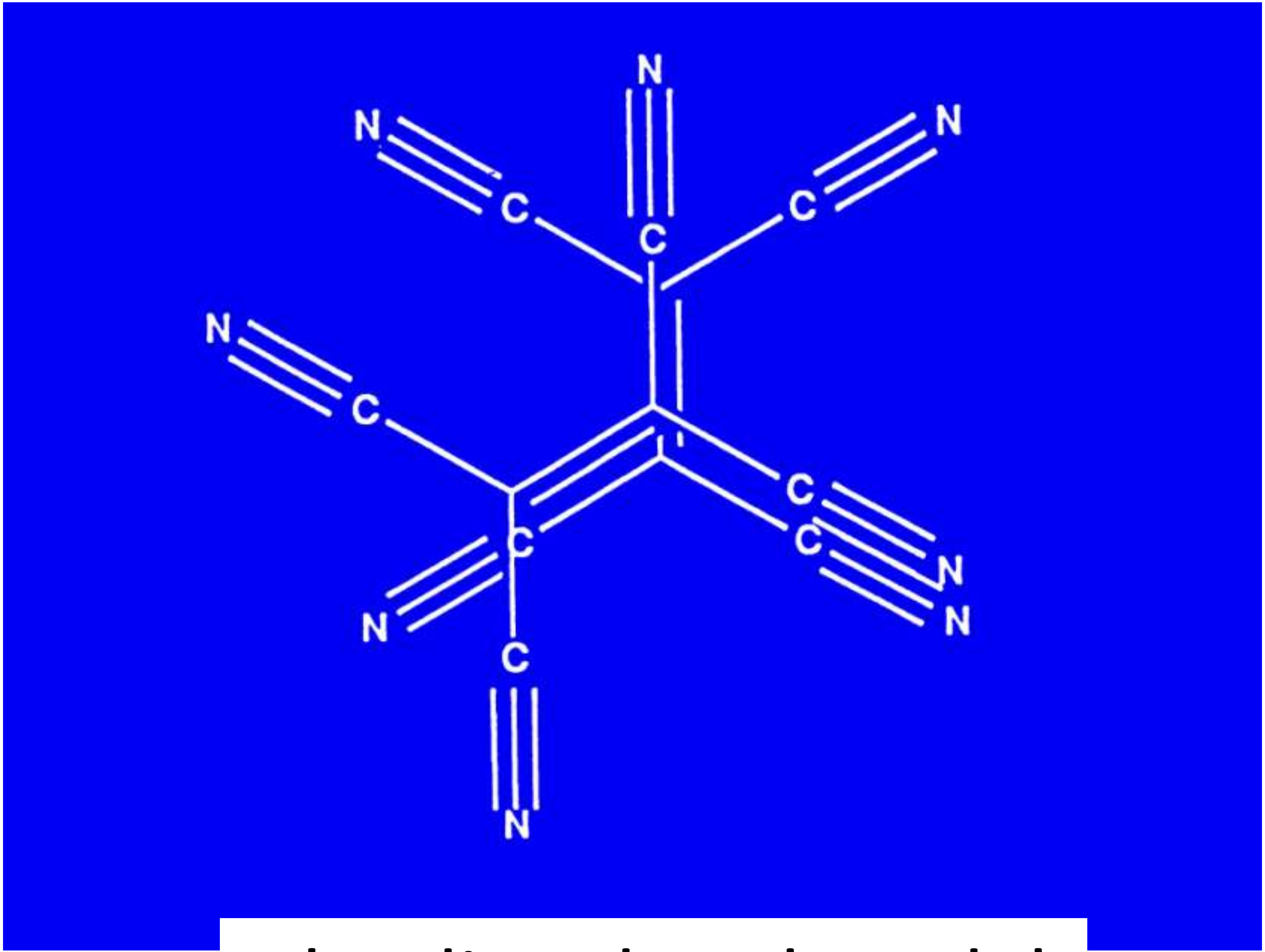
- one TCNE is "normal"
- second counter ion appeared "disordered":

TCNE Ligand

The chemist who made this material expected 2 molecules of the ligand.

When he tried to solve the structure, he used peak searches to find the atoms.

He decided to use a disordered model for one ligand, but found that the adps and site occupation factors were unreasonable.

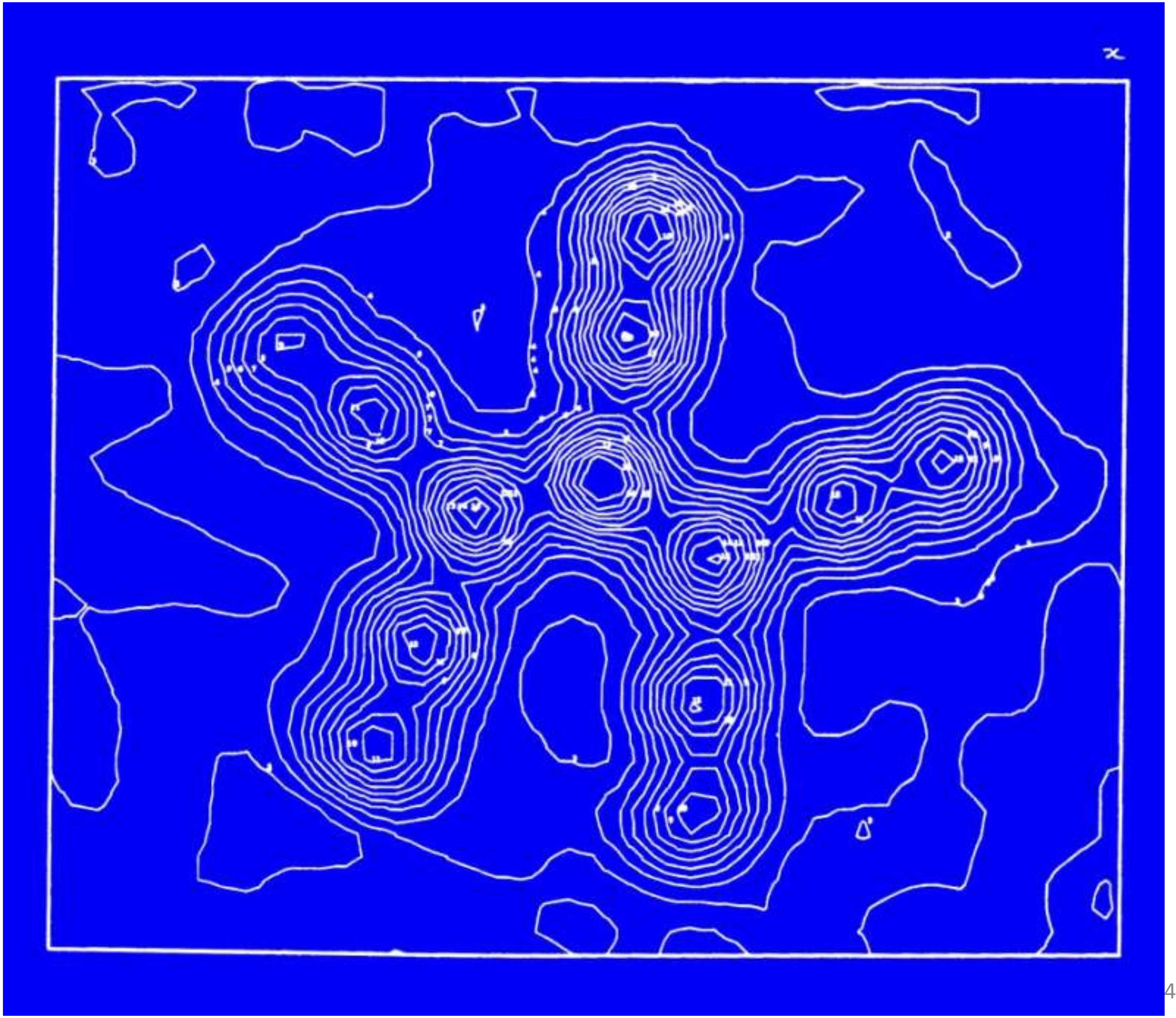


The disordered model

TCNE

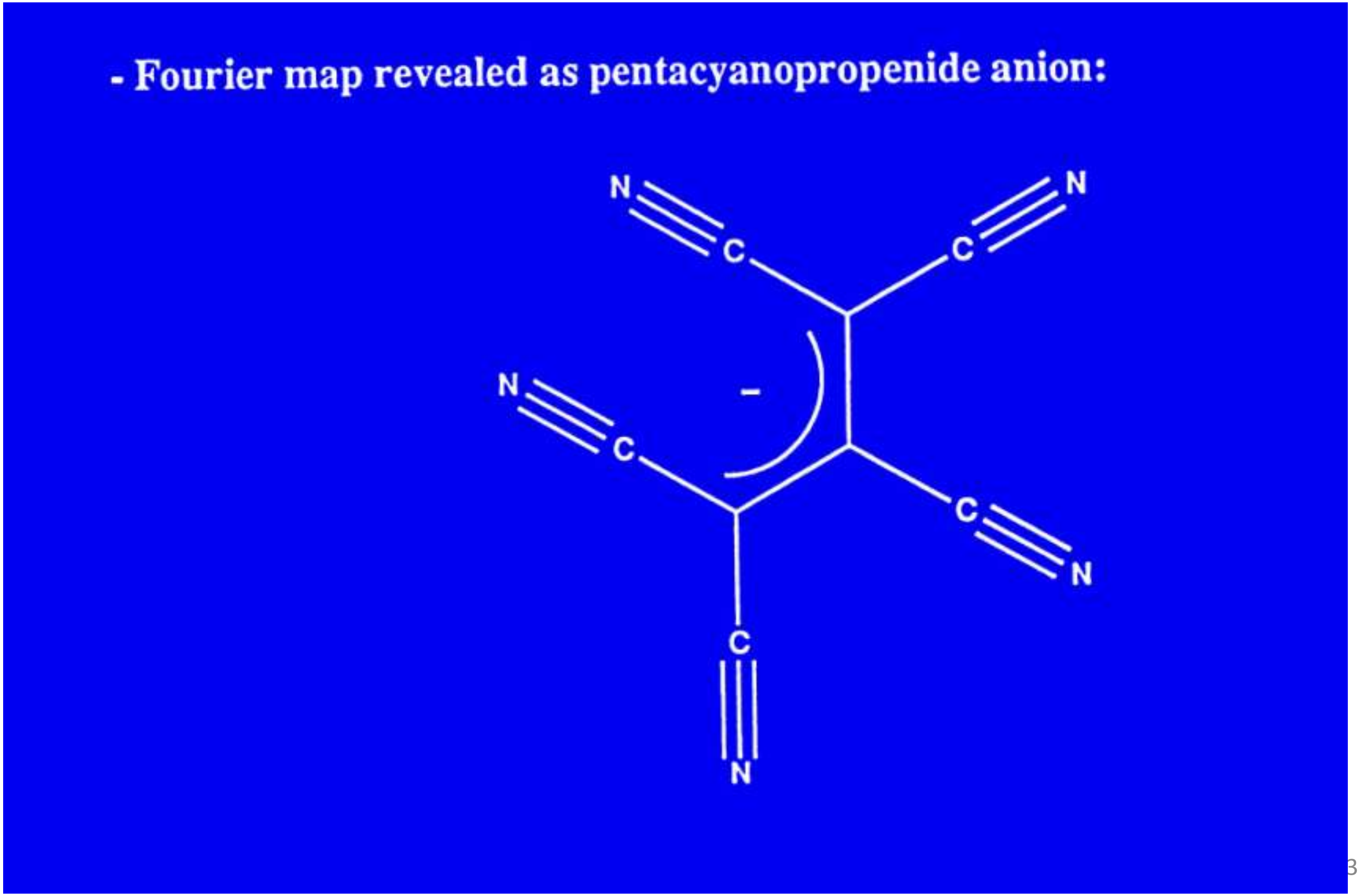
The Fourier map computed in the plane of the ligand was clean and sharp, with no ghost atoms.

A search of the CSD found several examples of pentcyanopropenide anions.



x

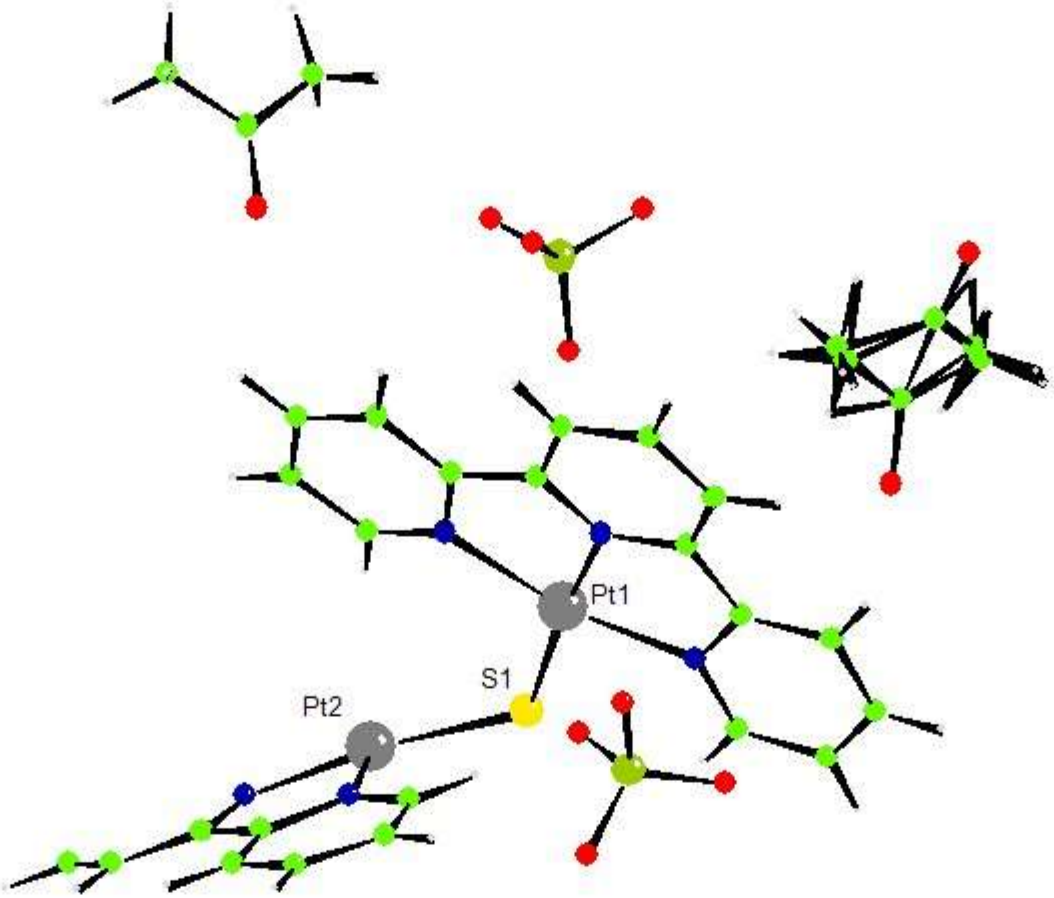
Final structure of the second ligand



- Fourier map revealed as pentacyanopropenide anion:

Example: Solvent of crystallisation

As well as the ClO_4 counter ions, there is acetone of solvation.



Nature’s Complexity – Z′ >1

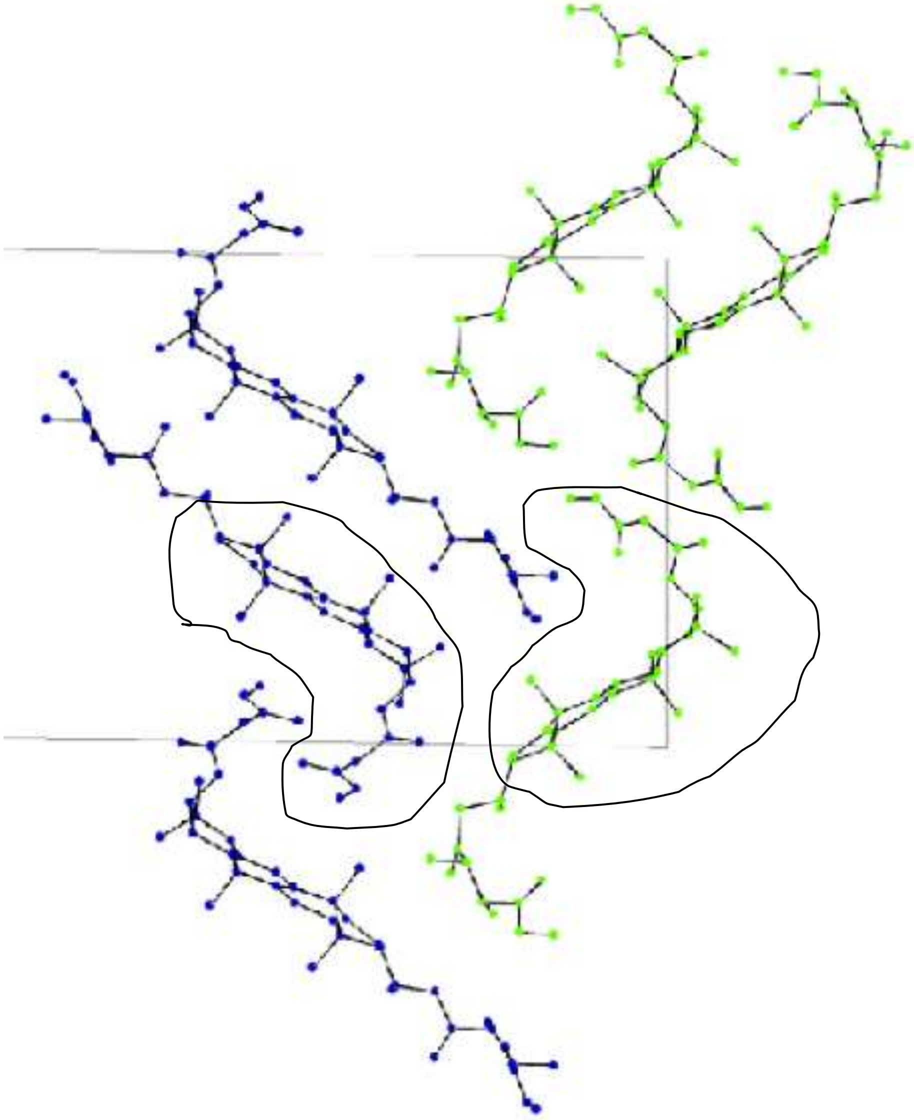
Z′ is the number of formula units in the unit cell. For organometallic and organic compounds, the formula unit is usually one molecule.

No one knows why the asymmetric unit sometimes contains several molecules – it is often observed in low temperature structure determinations.

Types of Z′ > 1

Different conformations.

- The ‘green’ molecules have the same connectivity as the ‘blue’, but have an extra kink.



Only one molecular species present

Hydrogen Bonding

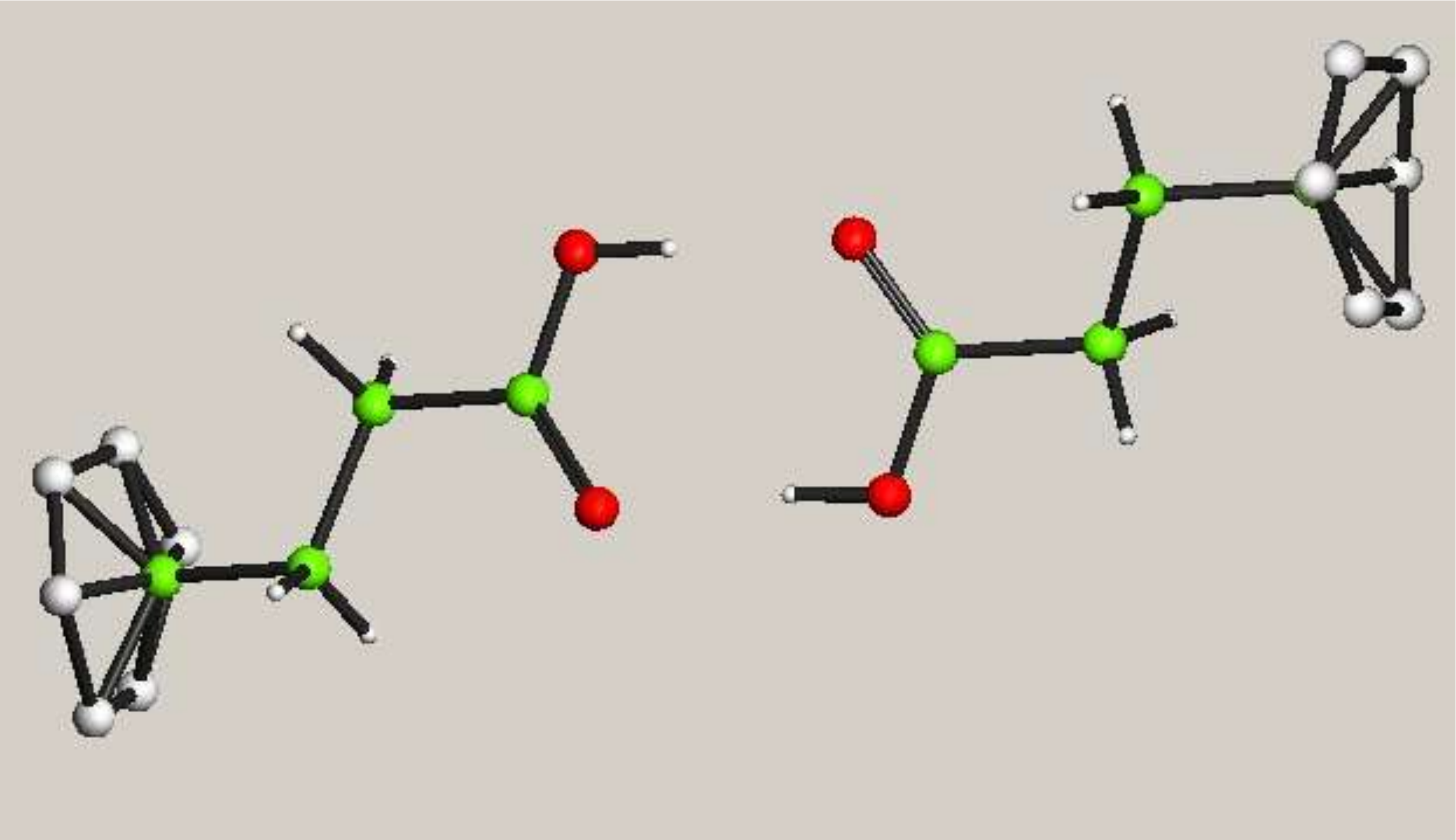
For many organic compounds, this provides a strong mechanism for holding the crystal together. Molecules can be linked into finite dimers, trimers etc, or into infinite ribbons or sheets.

Hydrogen Bonding is only of real interest to the chemist if there is a suspicion that it significantly changes the conformation of the individual molecules.

Example: tf-butenoic acid Dimer

There are 2 molecules in the asymmetric unit, forming a hydrogen bonded dimer.

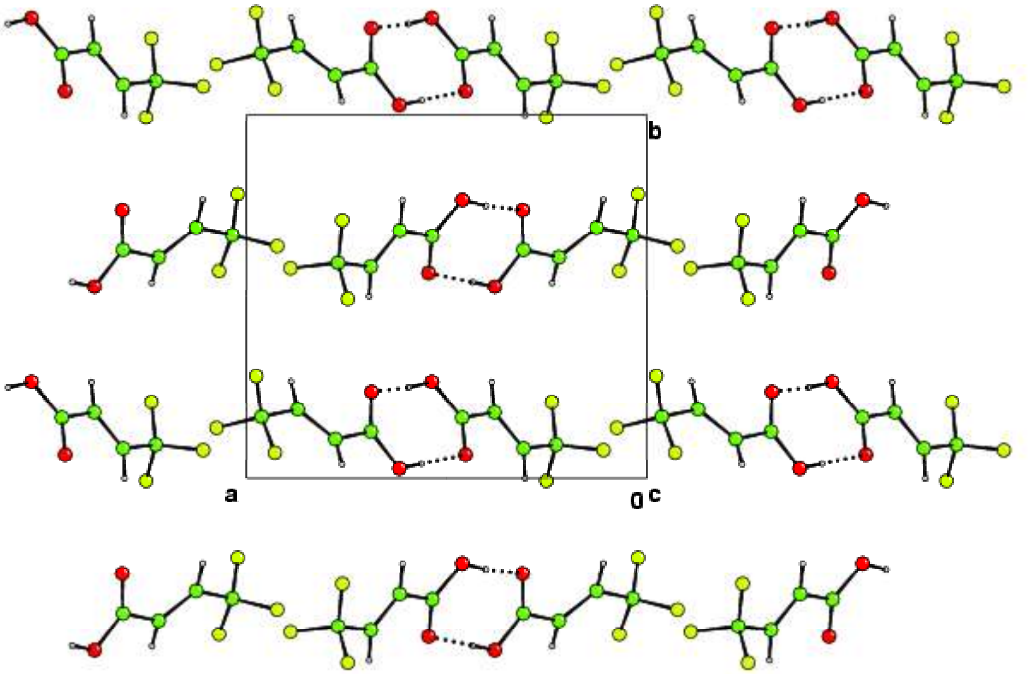
Note: both terminal CF_3 groups are disordered.



Types of Z′ > 1: tf-butenoic acid

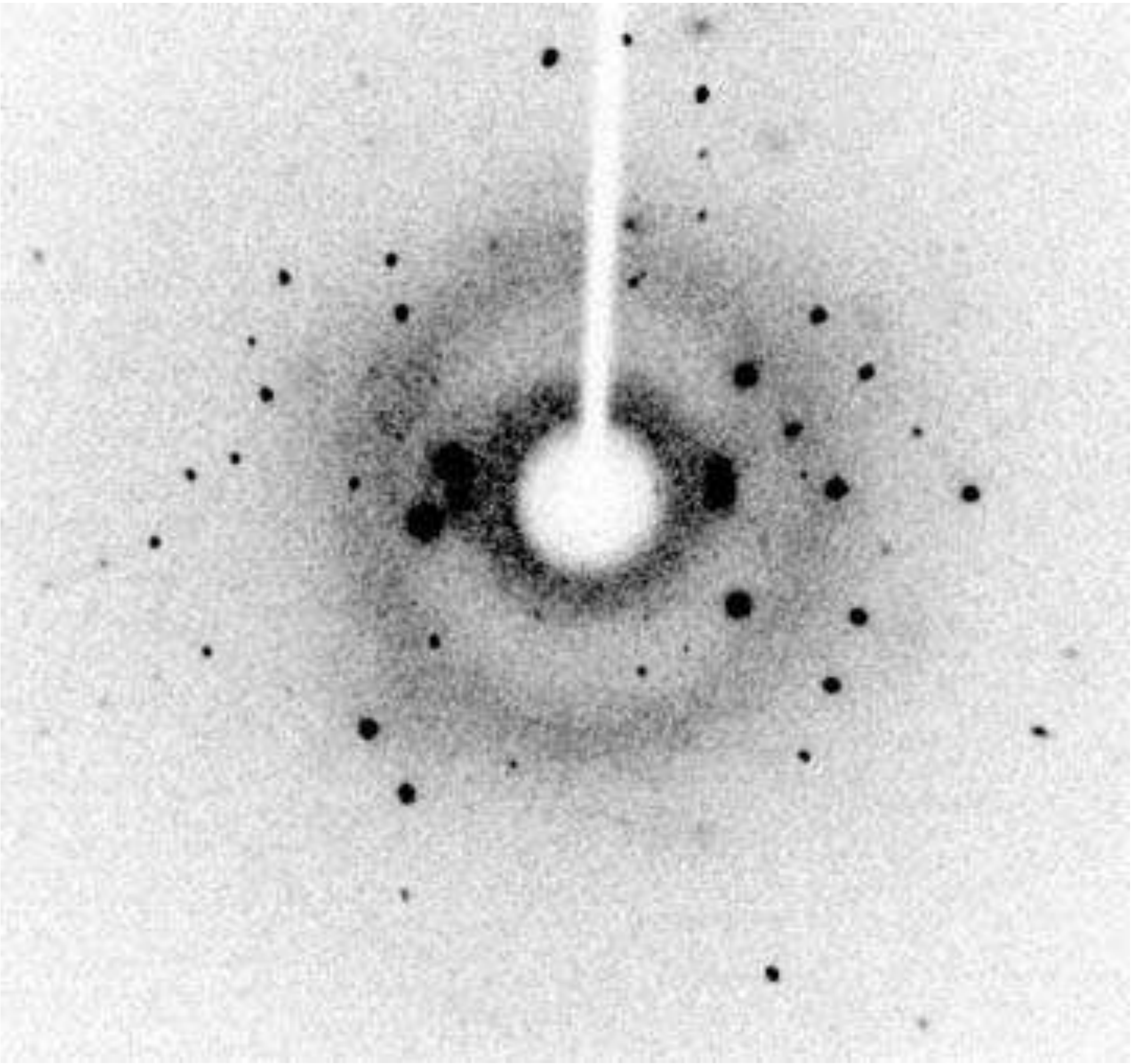
Pseudo-symmetry: The molecules are related by a non-crystallographic symmetry element. Usually the symmetry does not extend beyond the immediately neighbouring molecules.

There is no phase change in the range 220-260 K. The H-bonded dimer has a local centre of symmetry. The sg is P***21/c***



4,4,4-trifluoro-trans-2-butenoic acid

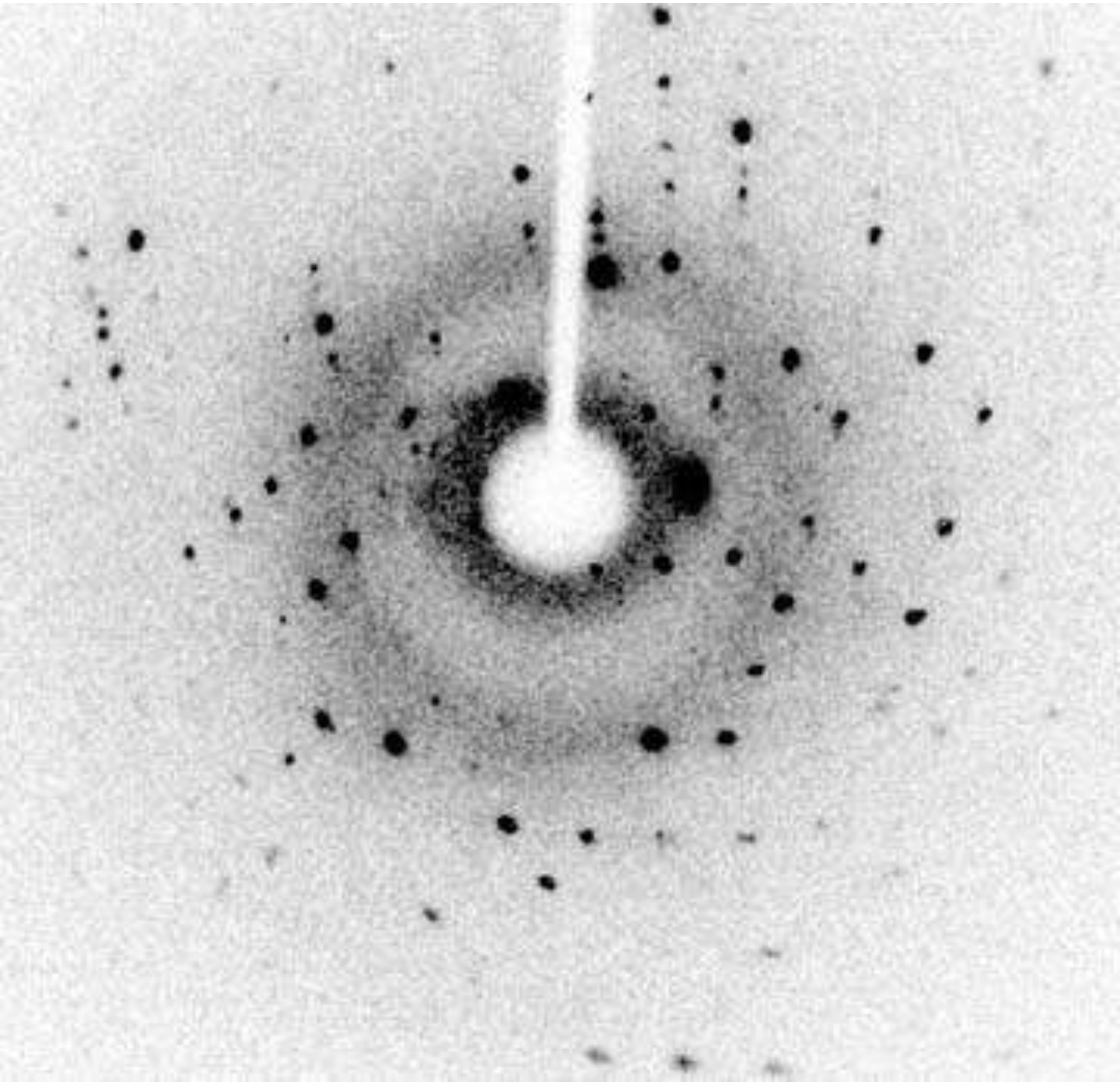
tf- butanoic acid High Temperature Form



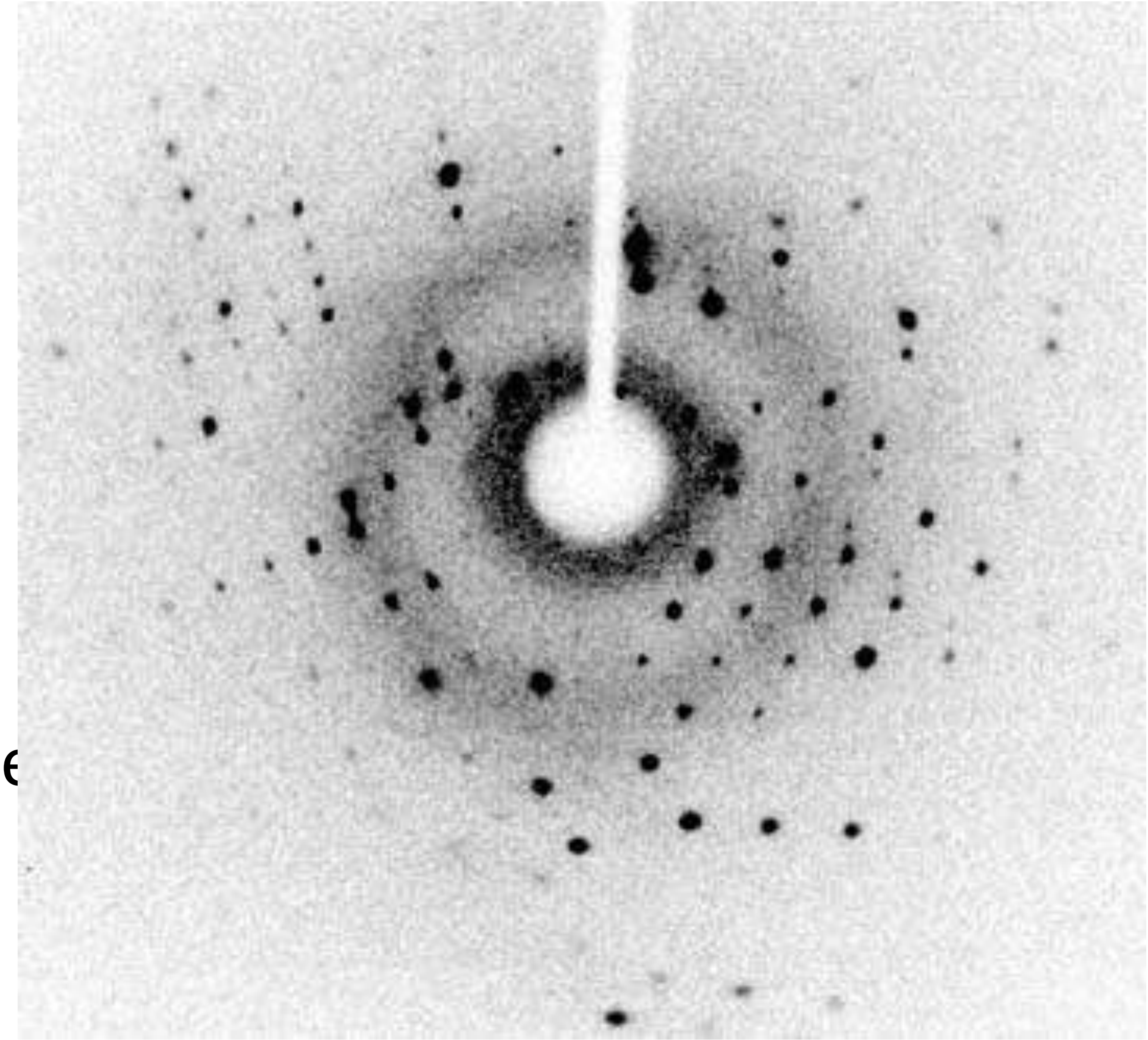
- Frame 74
- Time 3:43:57
- 230K

Transition

- Frame 76
- Time 4:44:47
- Interval : 50 sec
- 230K



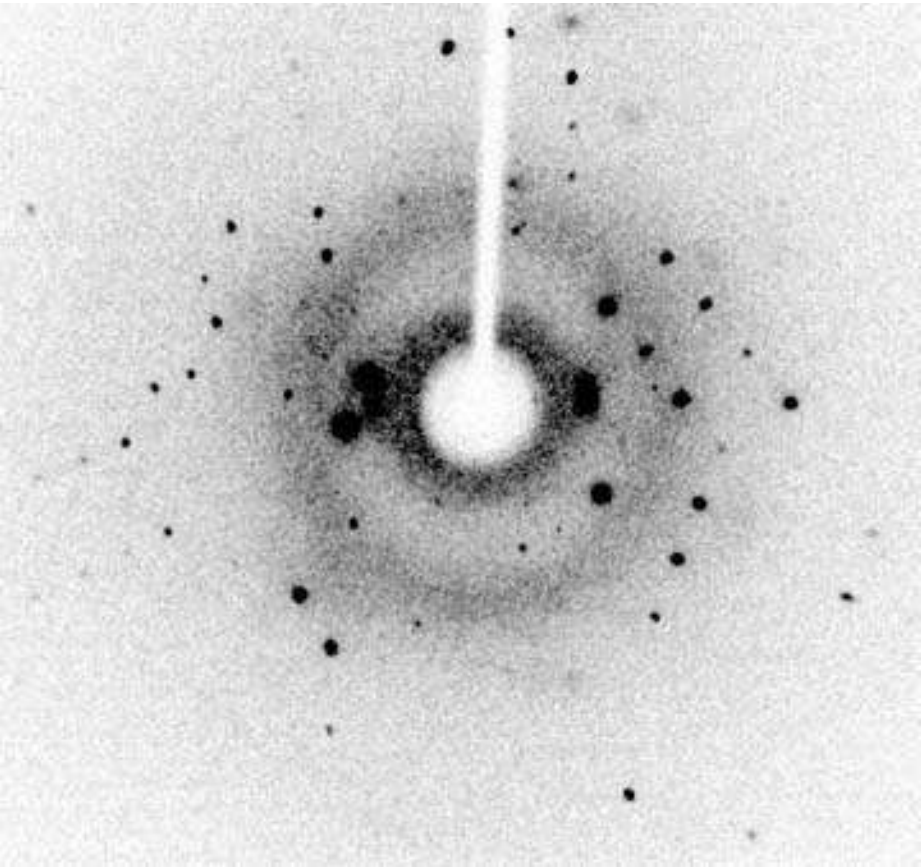
Low Temperature Form



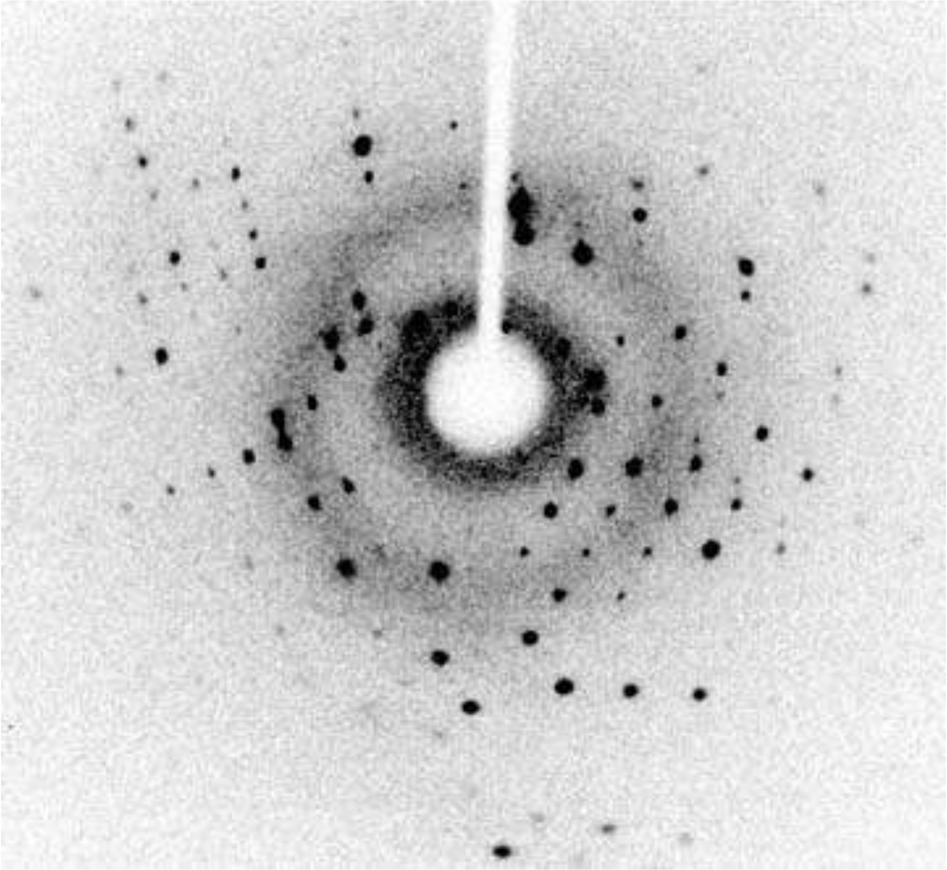
- Frame 79
- Time 4:46:02
- Interval : 75 se
- 230K

tf-butanoic acid

High temp form

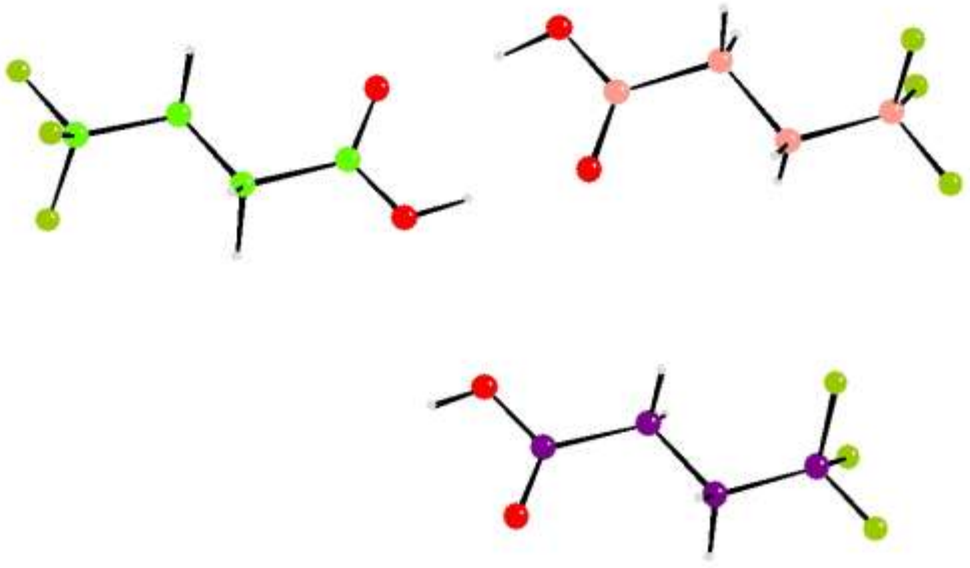


Low temp form



Trifluorobutanoic Acid – Z′=3

At 150K there are 3 molecules in the asymmetric unit

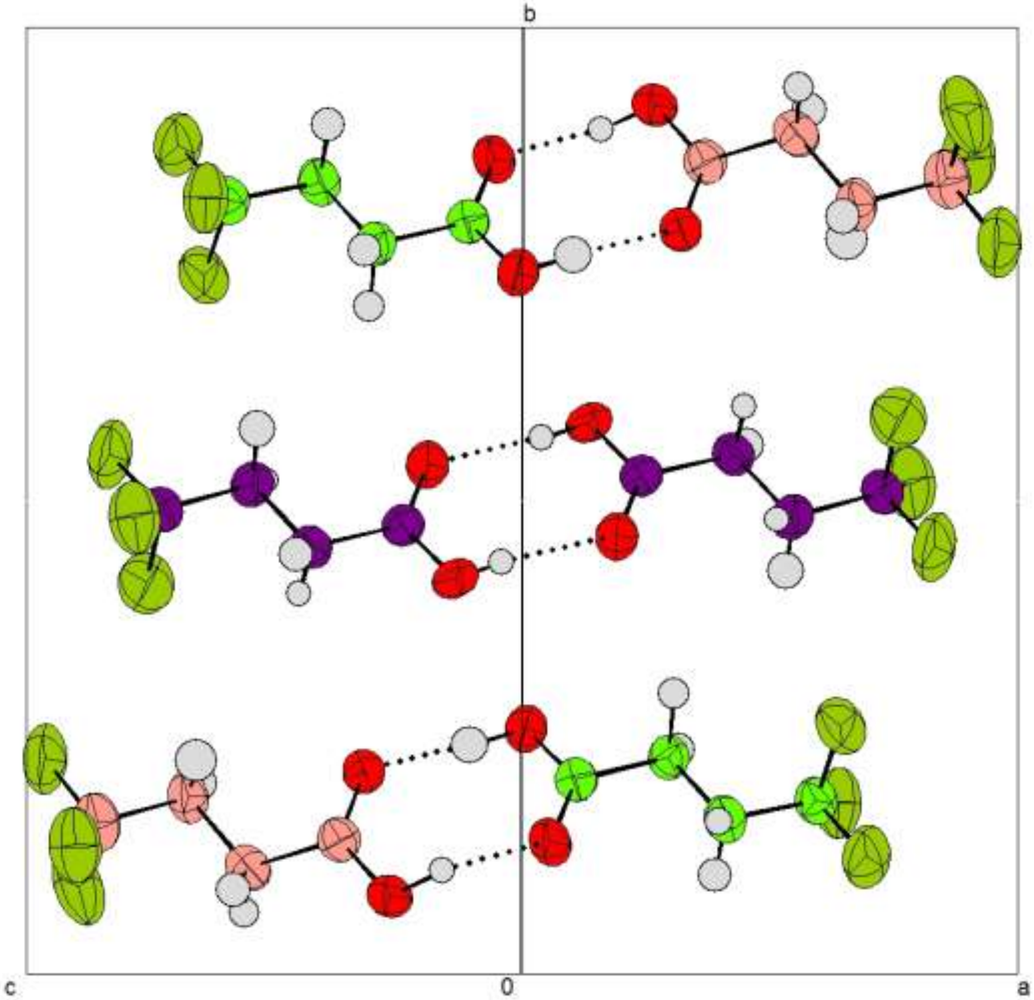


There is an evident hydrogen bonded dimer

(carbon atoms colour- coded by molecule)

Trifluorobutanoic Acid – Z′=3

The packing diagram shows that the isolated molecule also forms a dimer with a symmetry related molecule.



SQUEEZE and BYPASS

The structure factor is a complex number (has both magnitude and phase). The magnitude can be represented by:

$$F^2 = A^2 + B^2$$

where A is the real and B the imaginary part.

$$A_{hkl} = \int_V \rho_{xyz} \cos 2\pi(hx + ky + lz)\partial v$$

and

$$A_{hkl} = \sum_j f_j \cos 2\pi(hx + ky + lz)$$

with *B* **given by similar sin terms.**

SQUEEZE and BYPASS

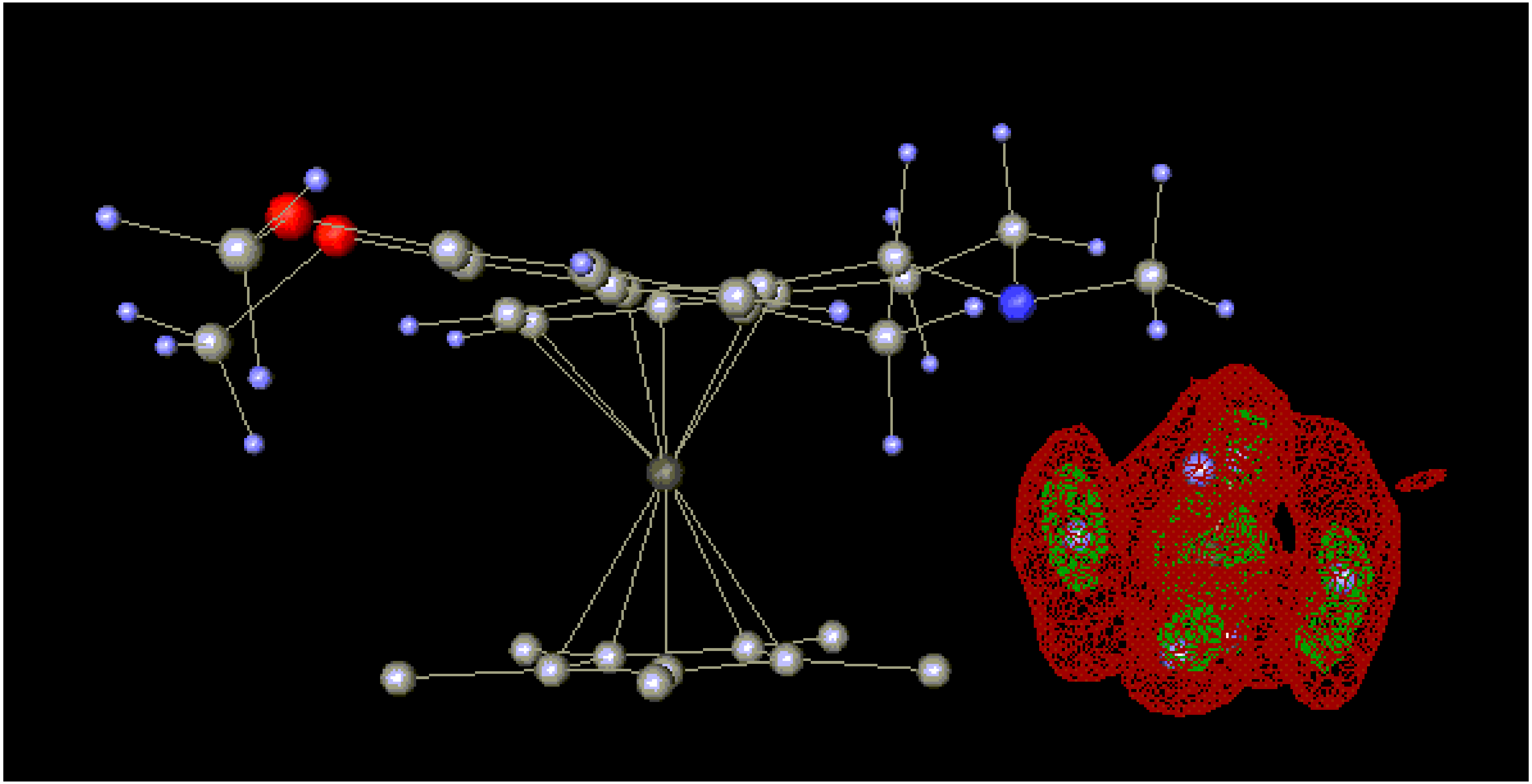
Ton Spek has tried to promote a hybrid structure factor expression:

$$A_{hkl} = \sum_j f_j \cos 2\pi\left(hx + ky + lz\right) + \int_v \rho_{xyz} \cos 2\pi\left(hx + ky + lz\right)\partial v$$

The first term is a summation over the resolved atoms. The integral in the second term is replaced by a summation over unresolved parts of the electron density map.

SQUEEZE and BYPASS

The PF_6 could be modelled by multiple disorder, or by its discrete transform



SQUEEZE and BYPASS

Problems:

- **The electron density used in the summation is computed from the current model, so will not be completely correct.**
- **The volume to be integrated over may be difficult to define.**

Crystallisation of a racemic mixture of chiral molecules.

Case 1: Achiral crystals

- The molecules associate in pairs across a symmetry operator of the 'second kind' *e.g.* centre of inversion, mirror, glide.
- The space group will be achiral and can be either centrosymmetric (e.g. $P2_1/c$) or non-centrosymmetric (e.g. Pc).
- The crystal will contain equal numbers of the R and S molecules
  - A solution made from an achiral crystal will not rotate polarised light.

Crystallisation of a racemic mixture of chiral molecules.

Case 2: Spontaneous resolution

- The molecules segregate themselves, with each crystal containing exclusively R or S molecules. This is called spontaneous resolution, and the resulting mixture of crystals is called a conglomerate. Both ‘hands’ of crystals in Pasteur’s resolved tartrate were in space group $P2_1\ 2_1\ 2$.
- The space group will be **chiral non-centrosymmetric**, *e.g.* $P2_1\ 2_1\ 2_1$. The space group of a crystal containing only R molecules will almost always be the same as that containing only S molecules. Exceptions are pairs like $P3_1$ and $P3_2$.
- The specific rotation of a solution made from a crystal containing R molecules will be equal and opposite to that made from a crystal containing S molecules

Crystallisation of a racemic mixture of chiral molecules.

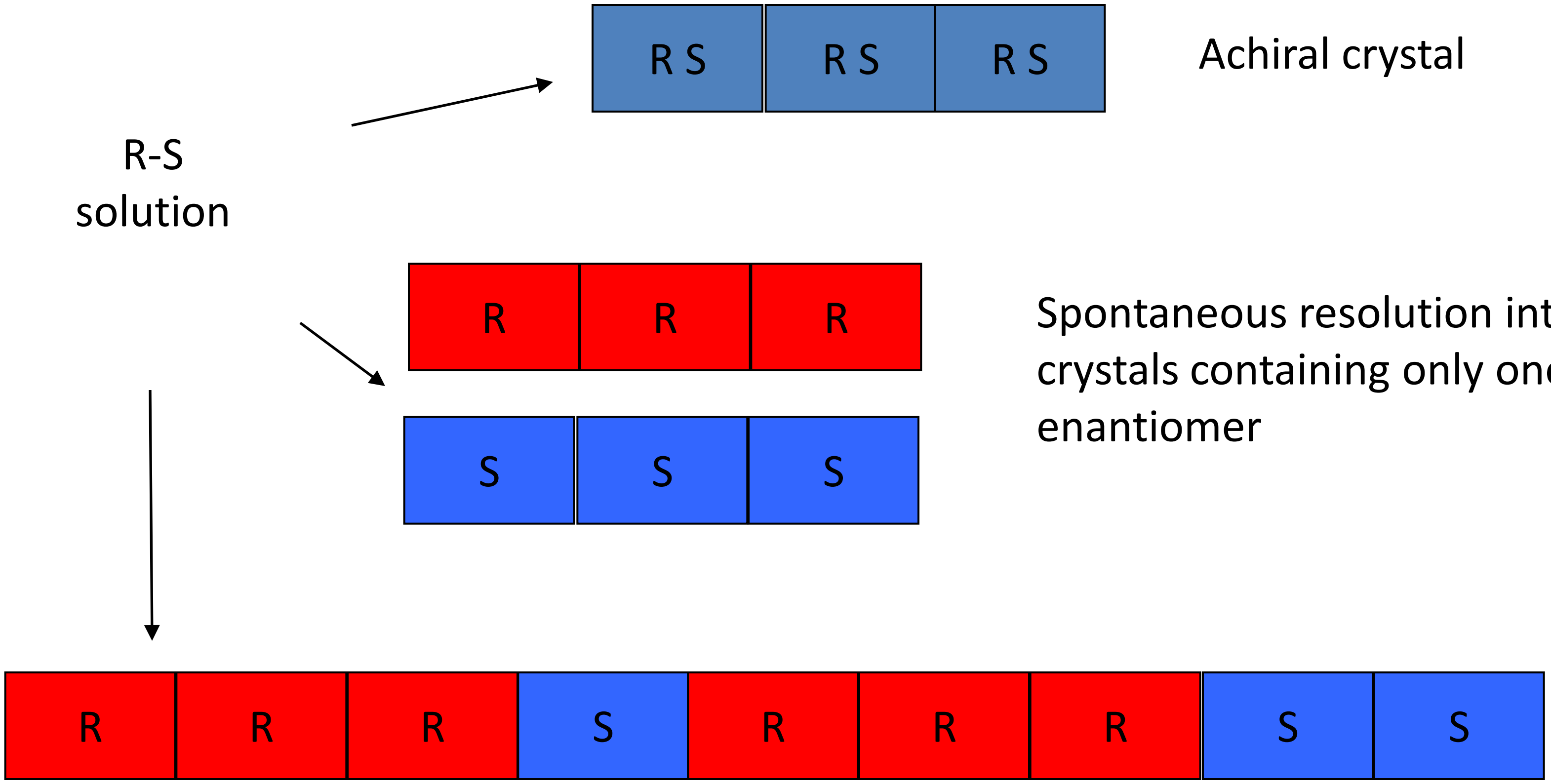
Case 3: Racemic twinning

- The crystals contain regions consisting exclusively of R molecules, and regions consisting exclusively of S molecules. These regions will have an exact crystallographic relationship with each other, but need not necessarily be present in equivalent quantities. The crystal is said to be twinned by inversion, often called racemic twinning.

- The space group will be chiral non-centrosymmetric, *e.g.* $P2_1\ 2_1\ 2_1$.

- The specific rotation of a solution made from a single crystal will be less than that from an un-twinned crystal.

Crystallisation of a racemic mixture of chiral molecules.



Achiral crystal

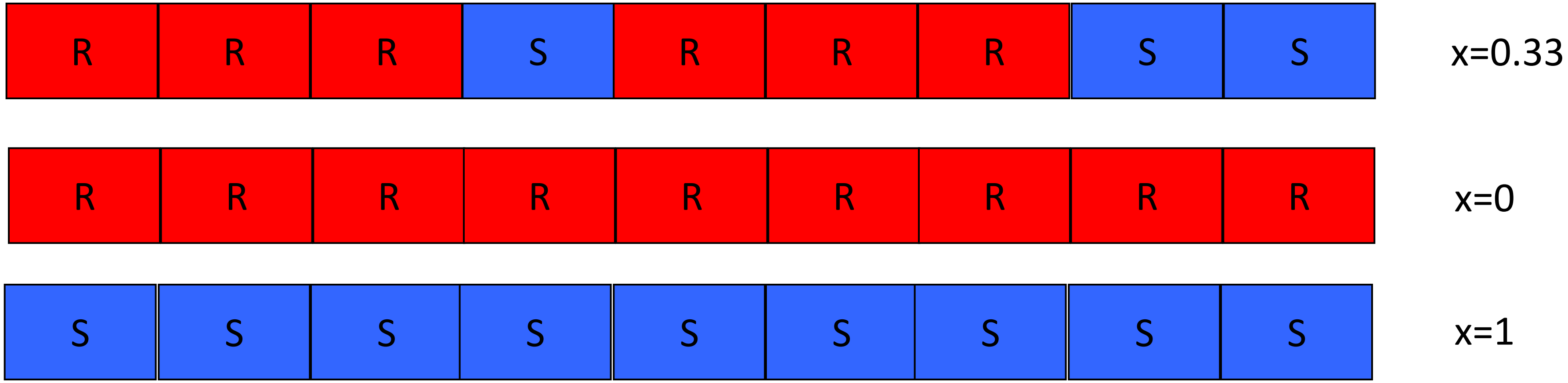
Spontaneous resolution into crystals containing only one enantiomer

Twin by inversion containing pure R and pure S regions.

$$C = (1 - x)R + xS$$

Flack parameter, x

$$F^2{}_{hkl} = (1-x)F^2_{hkl} + xF^2_{-h-k-l}$$

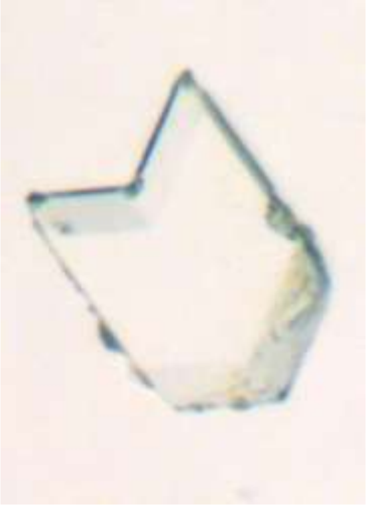


R	R	R	S	R	R	R	S	S	x=0.33
R	R	R	R	R	R	R	R	R	x=0
S	S	S	S	S	S	S	S	S	x=1

Refining the Flack parameter reveals the proportion of R and S enantiomers in an inversion twin. If x=0 or 1, the crystal contains only one enantiomer

Example x values for a model consisting of the R enantiomer

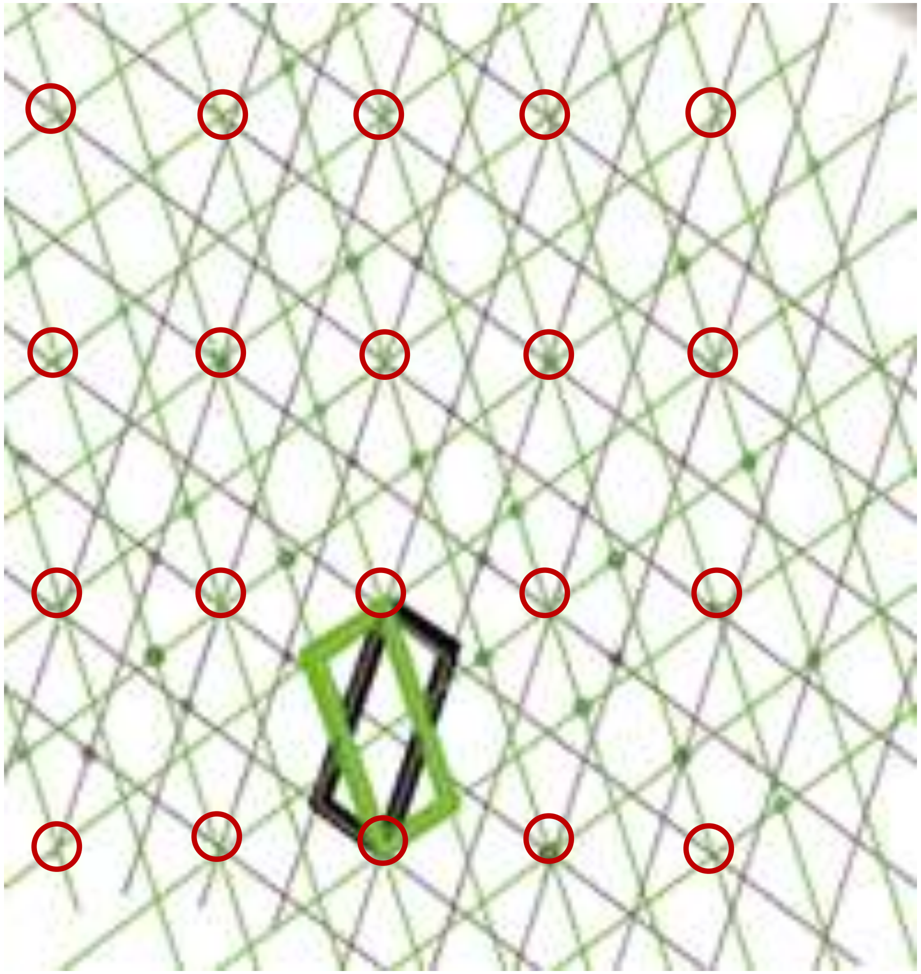
Selecting a good crystal: twinning



- Occasionally, something happens to cause a crystal that is happily growing in one direction to suddenly change direction. The new direction has a specific geometric relationship with the original direction. Crystals showing this phenomenon are called twins.

- Two crystals that just happen to be stuck together are not technically twins, although the undesirable effects on the diffraction pattern are the same.

Twinning: Effect on diffraction pattern

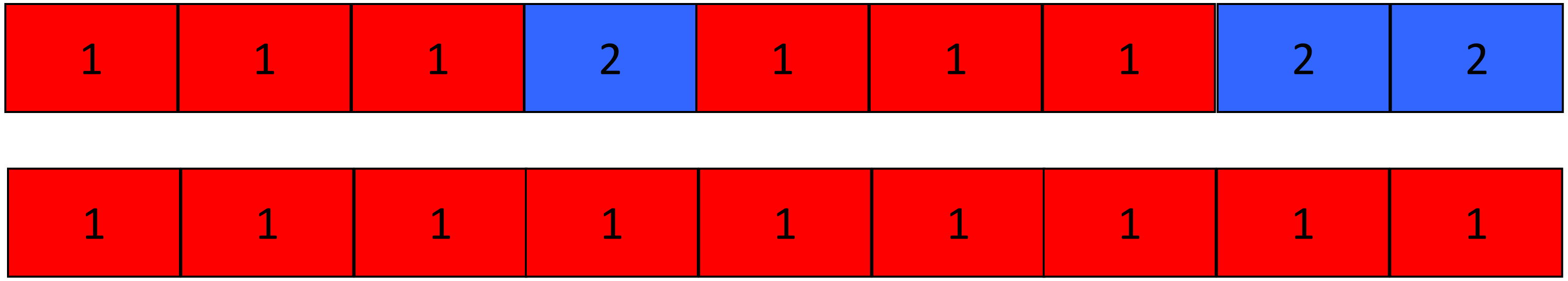


- Twinned crystals have two (or more) reciprocal lattices. These give rise to overlaid diffraction patterns which make it
  1. difficult to find a consistent cell and spacegroup.
  2. difficult to extract intensities (overlapped and nearly overlapped reflections).

Twin domains (generalisation of Flack)

$$[h'k'l'] = \mathbf{R}[hkl]$$

$$F^2{}_{hkl} = v_1 F^2_{hkl} + v_2 F^2_{h'k'l'}$$



$v_1/(v_1+v_2)$
$=0.66$

$v_1/(v_1+v_2)$
$=1$

Example proportions of v_1 for data indexed on domain 1.

Refining the twin proportions reveals the relative volume of each domain. Note that the Flack refinement is just a special case where R is the inversion matrix [-1 0 0; 0 -1 0; 0 0 -1]