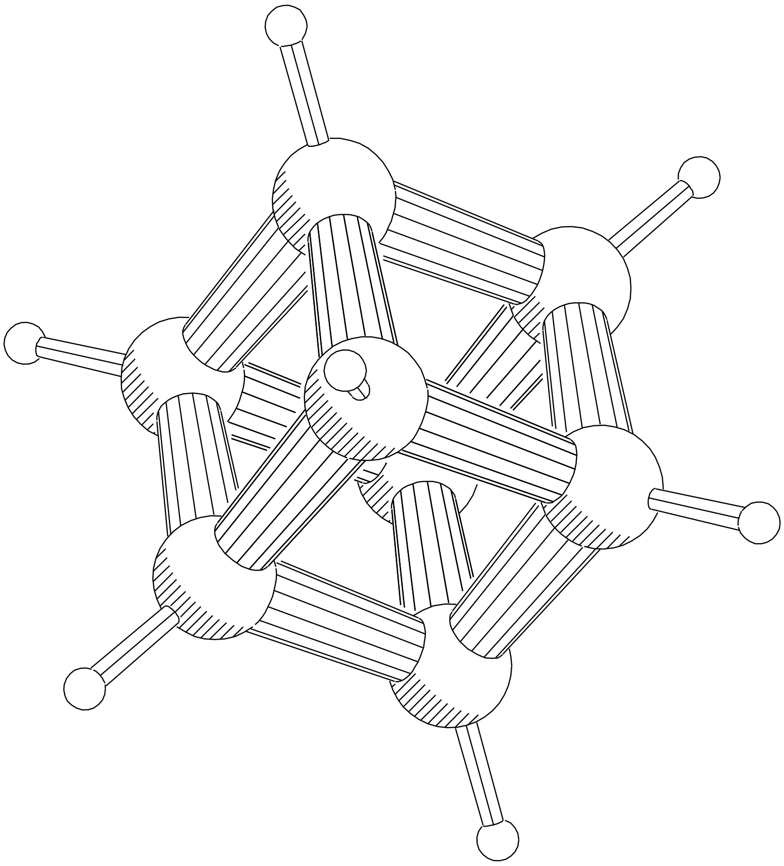
Interpretation of crystal structure determinations



Huub Kooijman
Bijvoet Center for Biomolecular Research
Crystal and Structural Chemistry, Utrecht University
Padualaan 8, 3584 CH Utrecht
The Netherlands
h.kooijman@chem.uu.nl
http://www.crystal.chem.uu.nl/people/huub.html
Version 2.3, January 2005

Purpose

This course offers

- an illustration of the basic principles of X-ray crystallography
- an overview of the successive stages of a crystal structure determination
- a description of the crystallographic terms found in the "experimental section" of a paper reporting crystal structures
- an extensive review of the geometrical properties of a crystal structure
- some of the basic formulae needed to calculate geometrical properties of a crystal structure
- an overview of the geometrical section of PLATON ("A multi-purpose crystallographic tool")
- a guide to the information found in a PLATON geometry listing
- an instruction on the use of PLUTON, the graphical section of PLATON

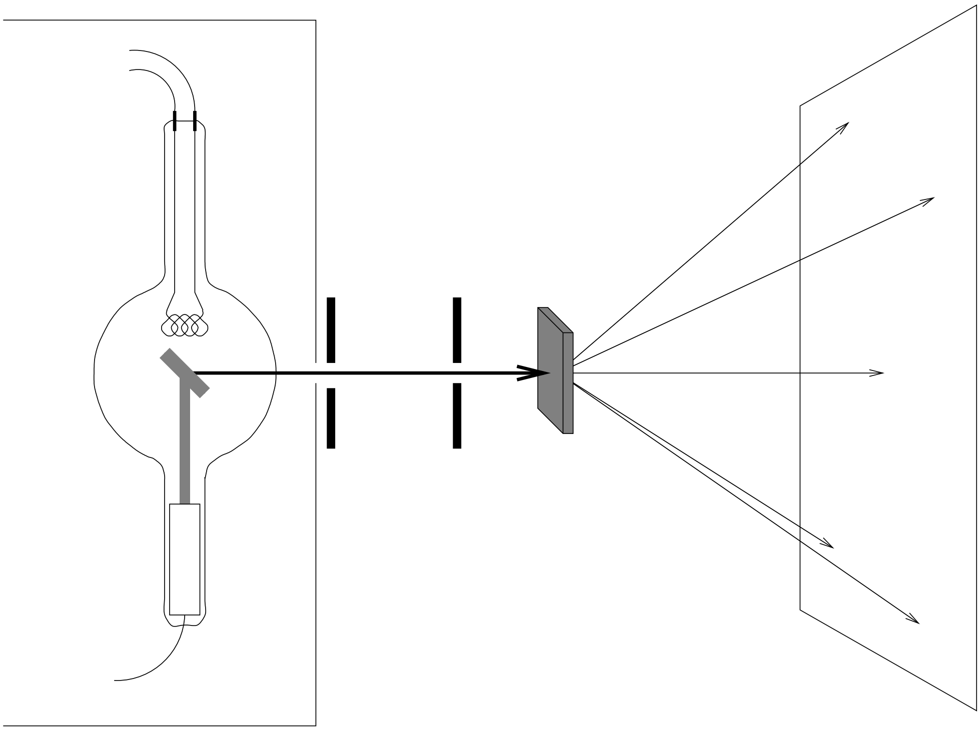
An index is given on page 169 ff.

Subjects

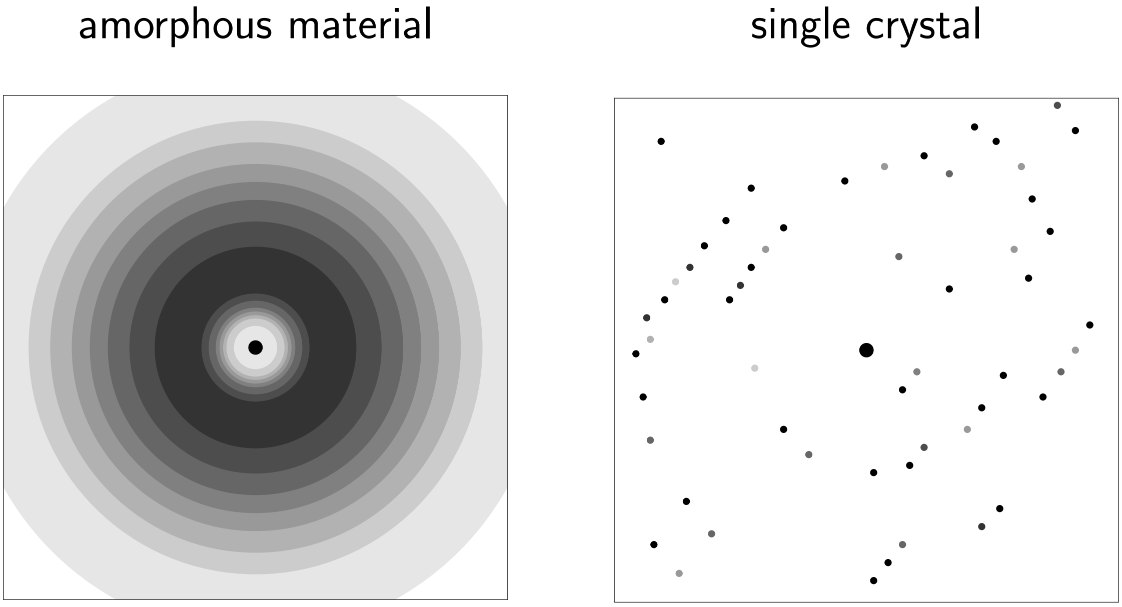
- crystallization (M. Lutz, not included here)
- diffraction
- structure determination
- symmetry
- intramolecular geometry, including thermal motion analysis
- graphics
- intermolecular geometry, including hydrogen bonding
- miscellaneous subjects, including powder diffraction, graph sets, twinning, absolute configuration, data bases, validation
- X-ray crystallography and chemical bonding (M. Lutz, not included here)

X-ray diffraction

The experiment:



The results:

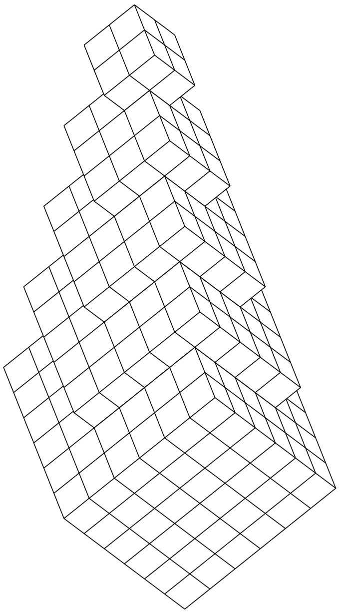

amorphous material

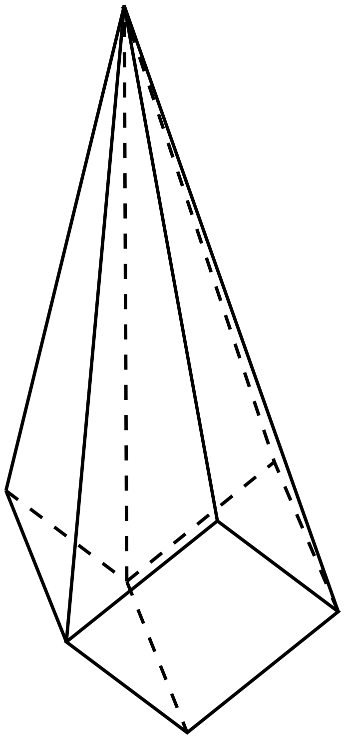
single crystal

(various intermediate forms are found for materials with short-range ordering)

The periodic nature of crystals

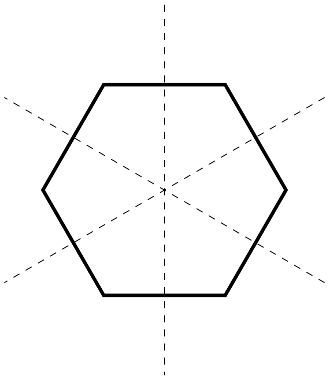
To explain the regular shape of a crystal Huygens (17th century) and Haüy (18th century) described crystals as a regular stackings of “elementary building blocks” or “molécules integrantes”.

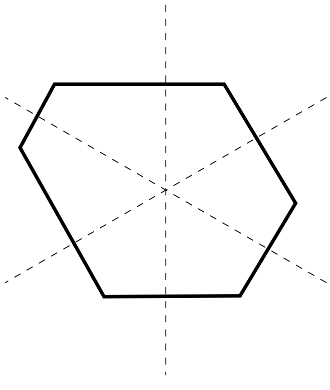




This hypothesis explains the “preservation of angles”:

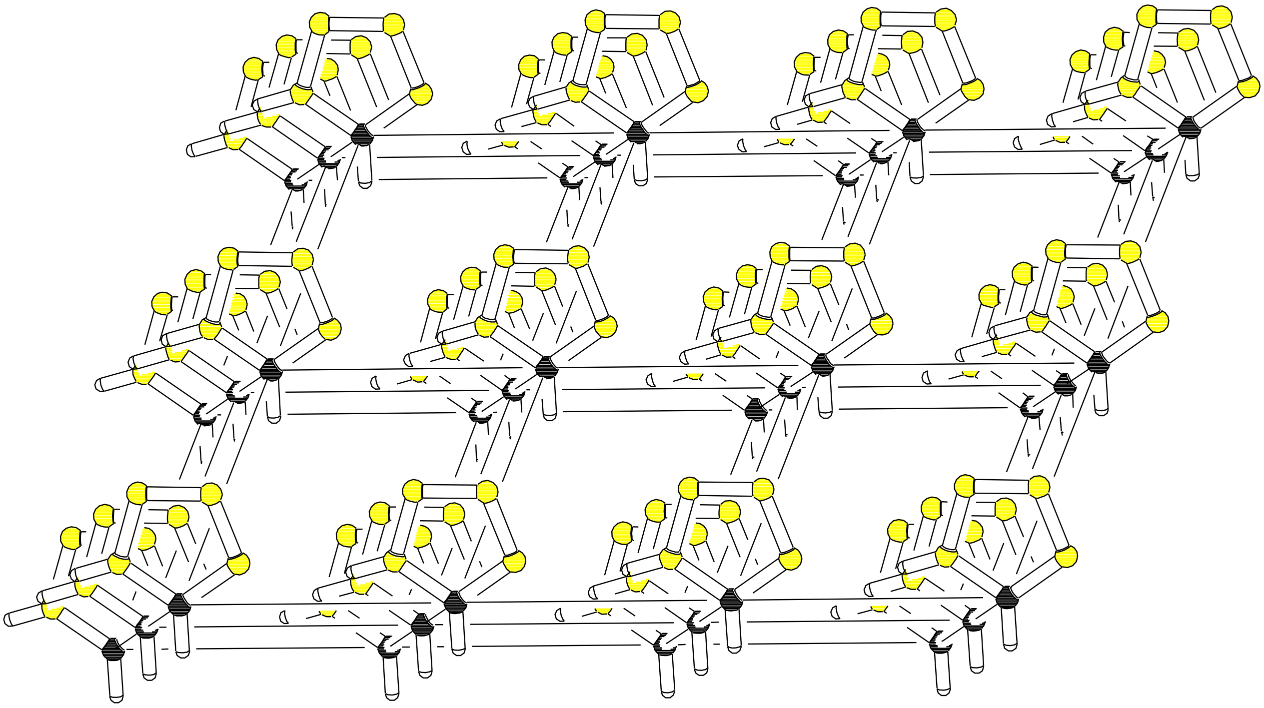
> The bundle of lines originating in any point in a crystal and normal to the faces of the crystal is an invariable characteristic of the crystalline species.



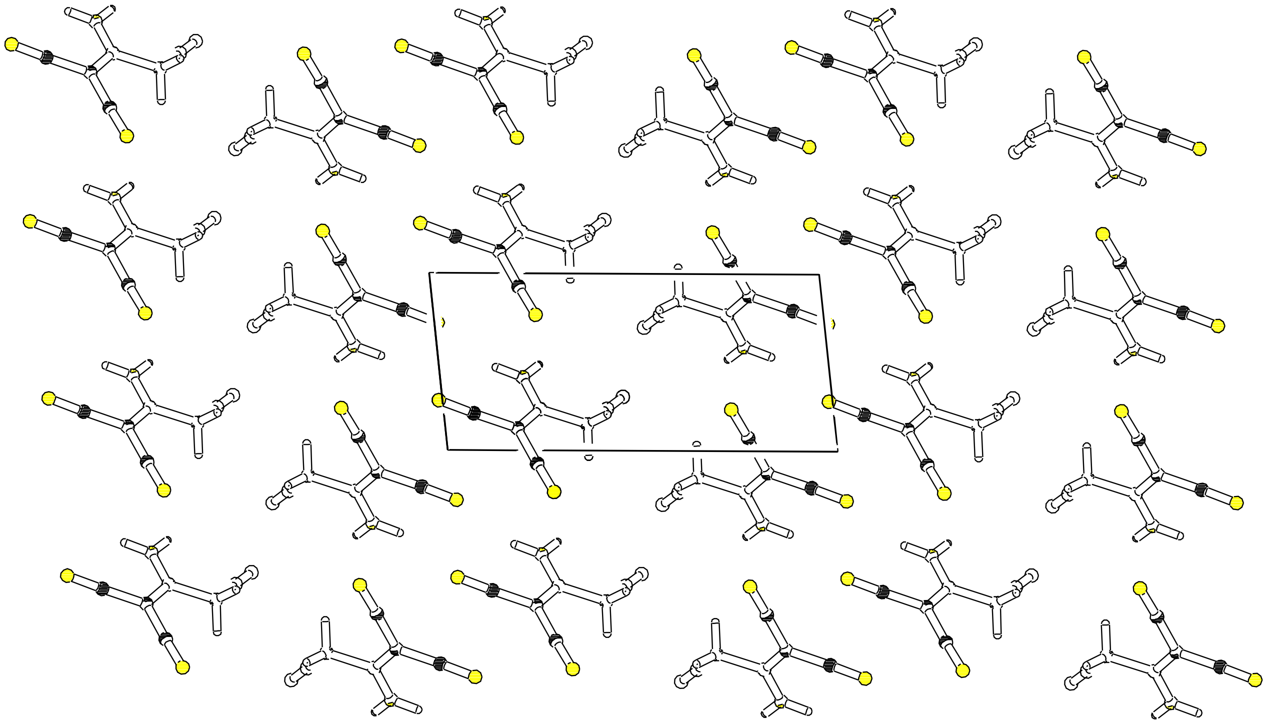


Translation symmetry in crystals

A crystal can be described as a unit cell on which translation symmetry has been applied. Two examples of the ordering of molecules in crystals:



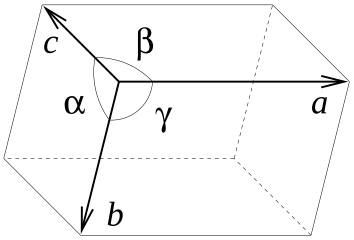
One molecule in each unit cell.



Two molecules, related by a rotation, in each unit cell

Description of the unit cell

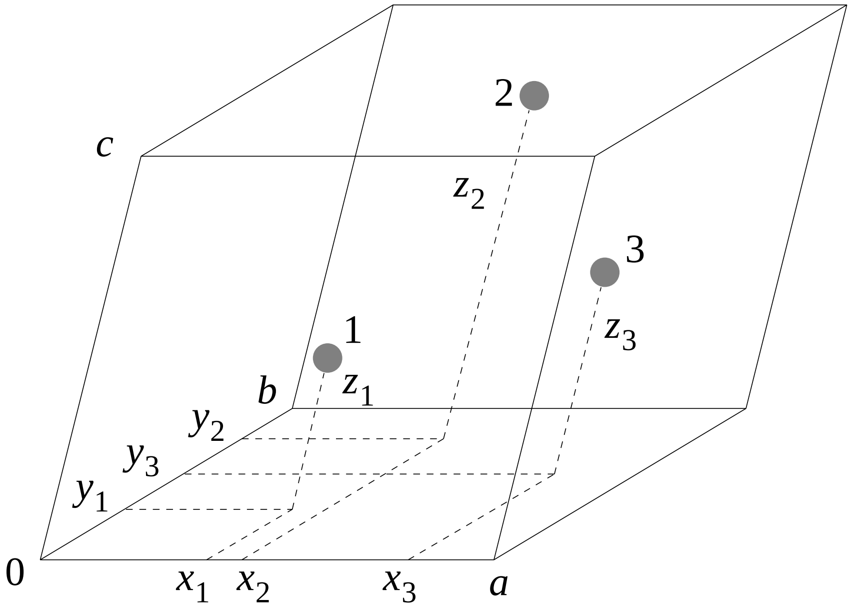
The unit cell is defined by three independent translation vectors $\vec{a}$, $\vec{b}$ and $\vec{c}$. Vector lengths (a, b, and c) and the inter-vector angles (α, β and γ) are used to characterize the cell.



The volume of the unit cell is given by

$$V = abc\sqrt{1 - \cos^2\alpha - \cos^2\beta - \cos^2\gamma + 2\cos\alpha\cos\beta\cos\gamma}$$

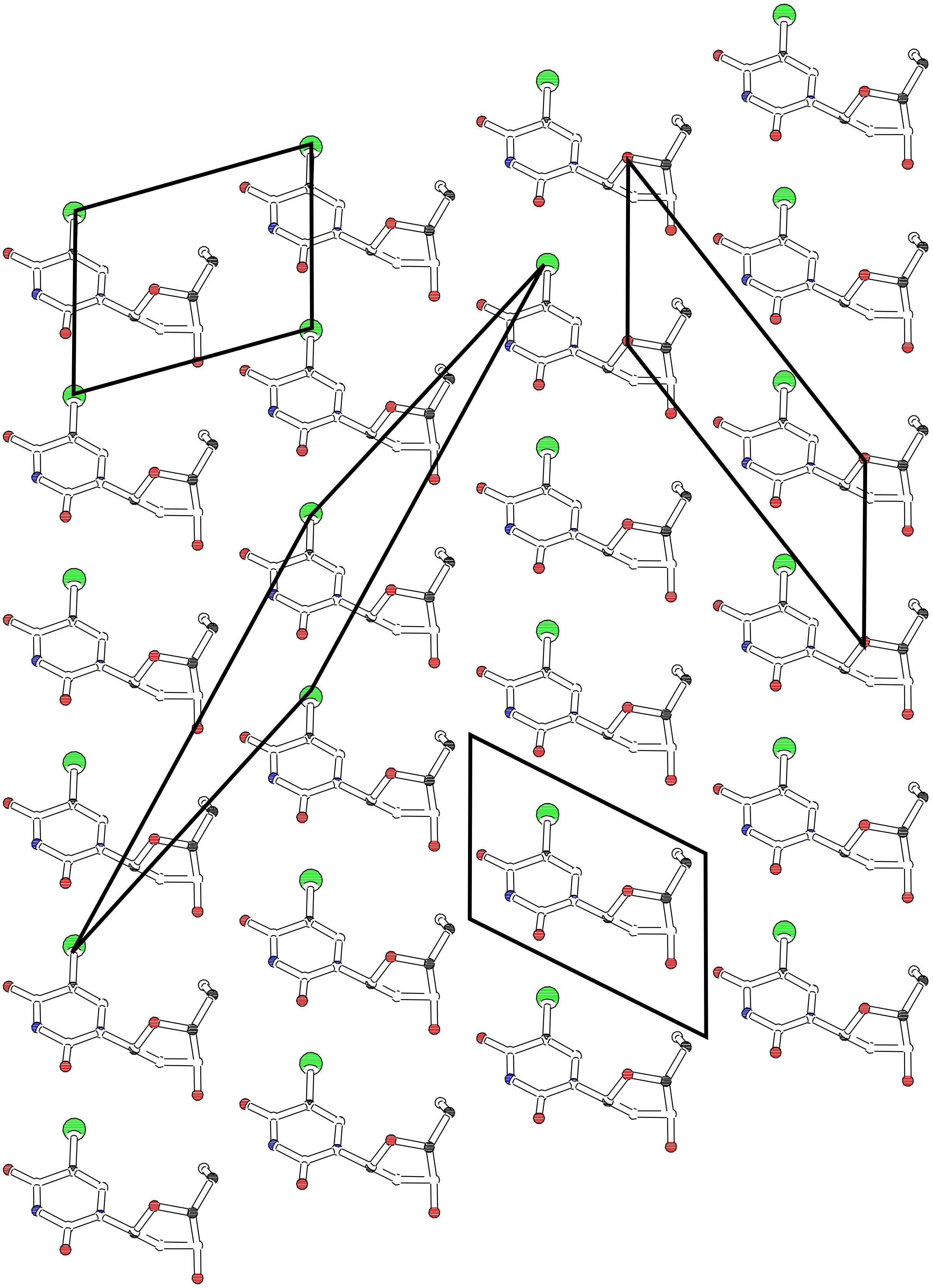
The position of atoms in the unit cell are indicated with fractional coordinates (x, y, z) of the atomic nuclei.



Atoms that lie within the boundaries of the unit cell have fractional coordinates in the range $0 < x, y, z < 1$. Applying translation symmetry amounts to adding or substracting integers.

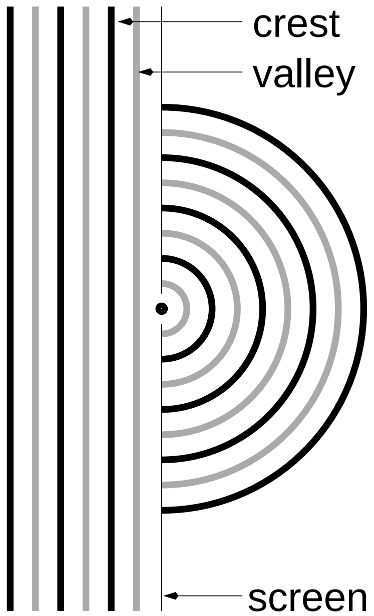
Choice of unit cell

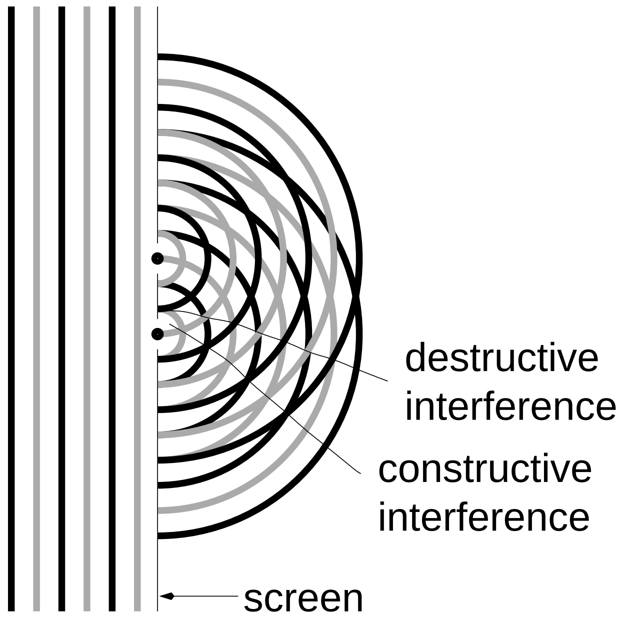
An infinite number of unit cells can be chosen; not all are practical to use. The choice of origin is free.



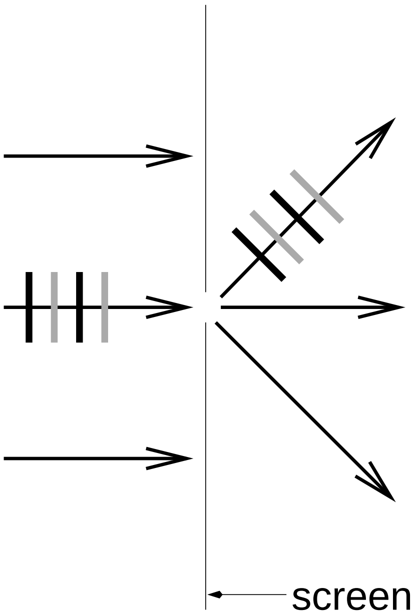
Interference of waves

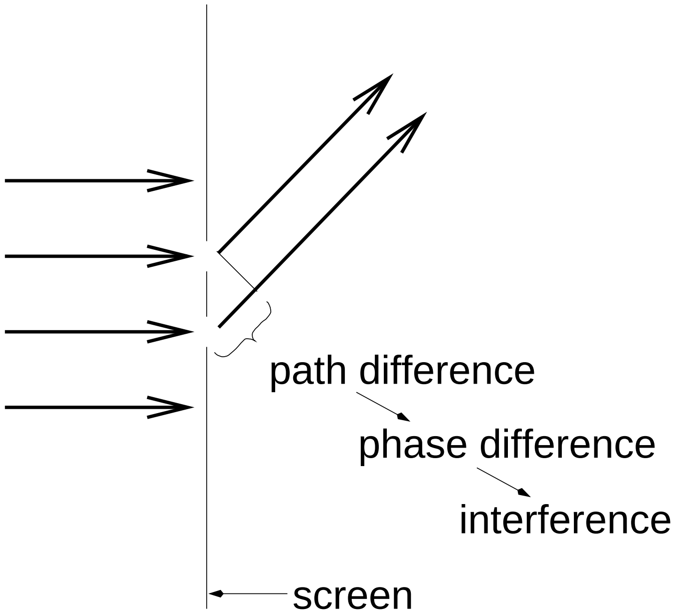
Wavefront model



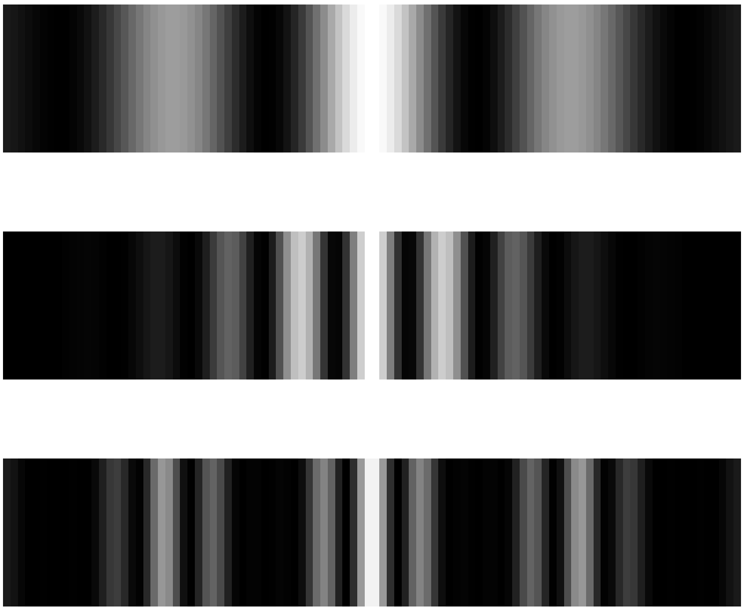


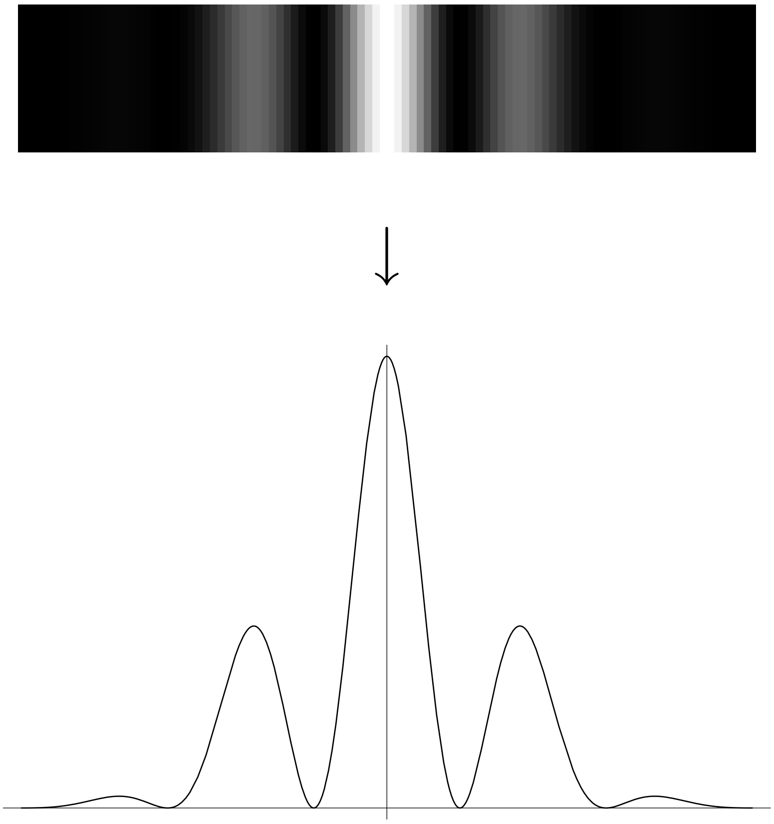
Wavevector model





Interference of light

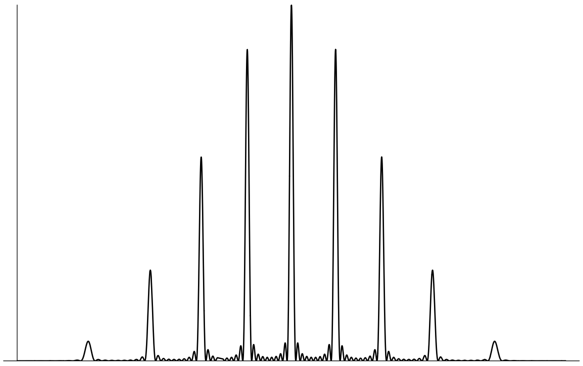




Diffraction at a 10-slit system

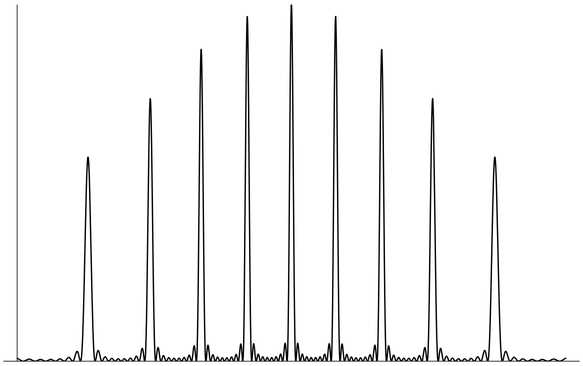
$r =$ slit width; $a =$ slit distance (or period)

(numerical values are in wavelength units)



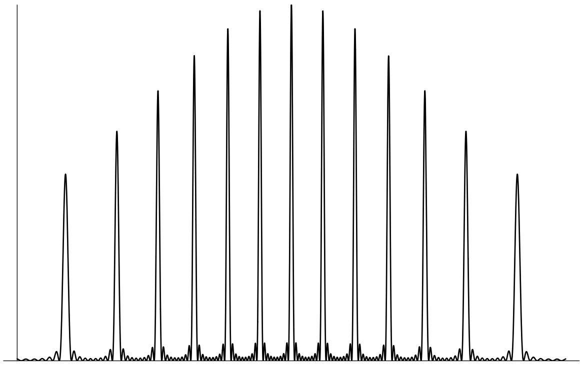
$r = 1.0$
$a = 5.0$

↑ change of slit



$r = 0.5$
$a = 5.0$

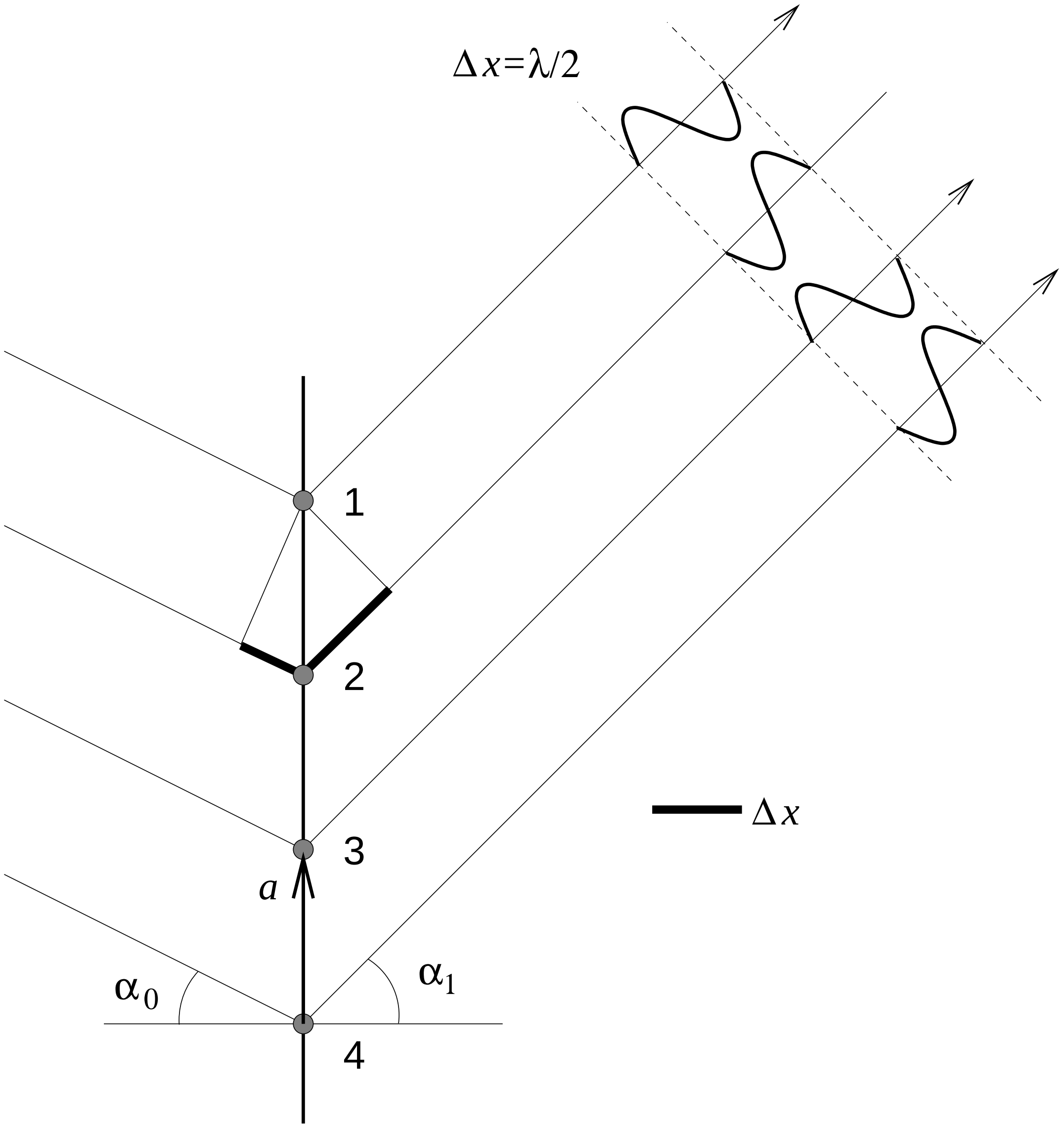
↓ change of period



$r = 0.5$
$a = 7.0$

X-radiation and atoms in a periodic row (1)

Consider an infinite array of atoms, related by translation symmetry, where each atom behaves as a point scatterer of X-radiation.

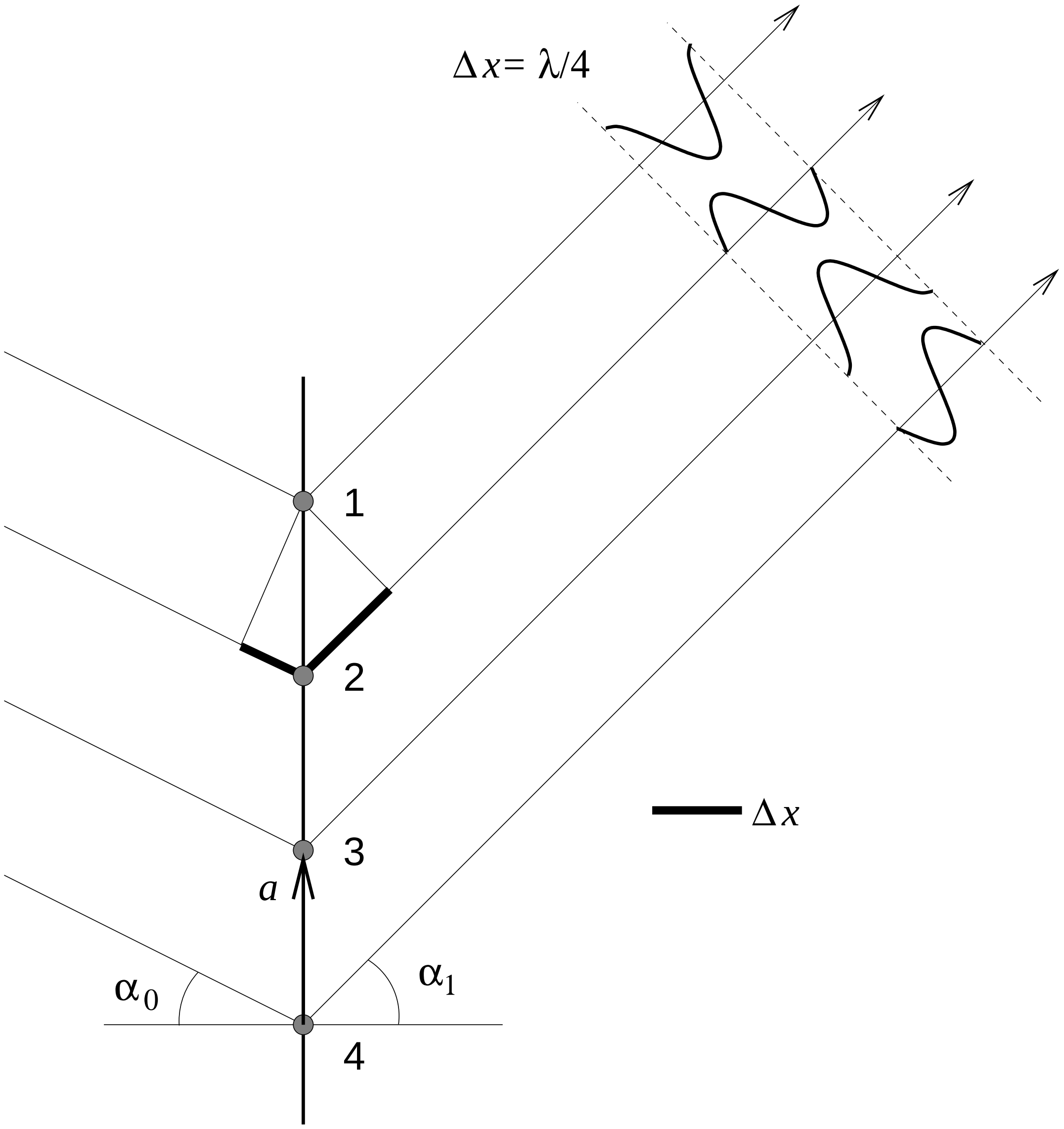


The path difference Δx between X-radiation diffracted by two atoms one translation period apart is determined by the period (a), the incoming angle (α_0) and the outgoing angle (α_1).

If $\Delta x = n\lambda$ $(n \in \mathbb{N})$ all waves reinforce each other.

If $\Delta x = (n + \frac{1}{2})\lambda$ $(n \in \mathbb{N})$ the waves diffracted by atom 1 and atom 2 extinguish each other, as do the waves diffracted by atom 3 and atom 4.

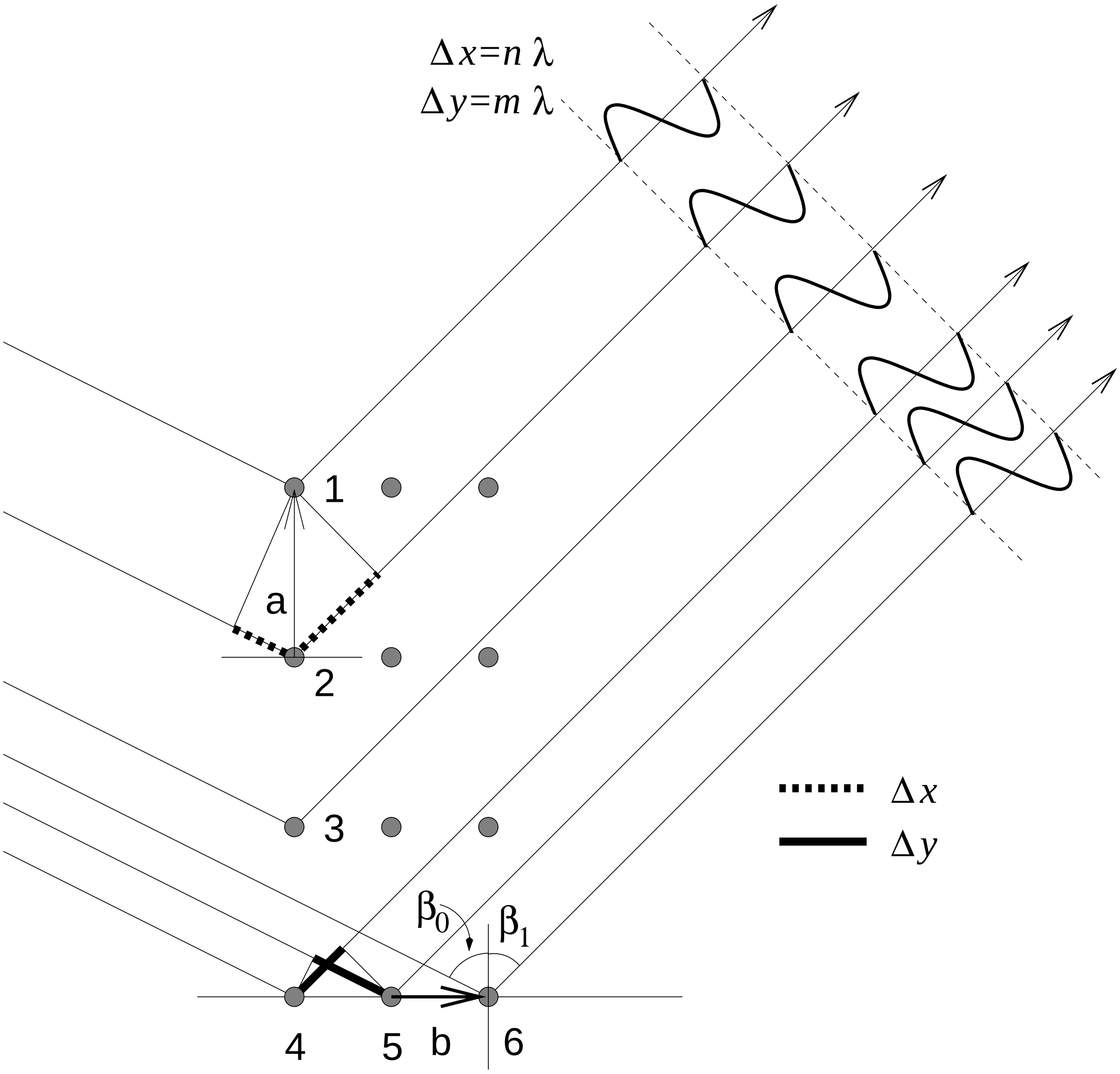
X-radiation and atoms in a periodic row (2)



If $\Delta x = (n + \frac{1}{4})\lambda$ the waves diffracted by atom 1 and atom 3 extinguish each other, as do the waves diffracted by atom 2 and atom 4

X-radiation and a 2-dimensional array of atoms

Consider an infinite 2-dimensional array of atoms, related by translation symmetry, where each atom behaves as a point scatterer of X-radiation.



Diffracted beams always extinguish each other, unless two two conditions are satisfied:

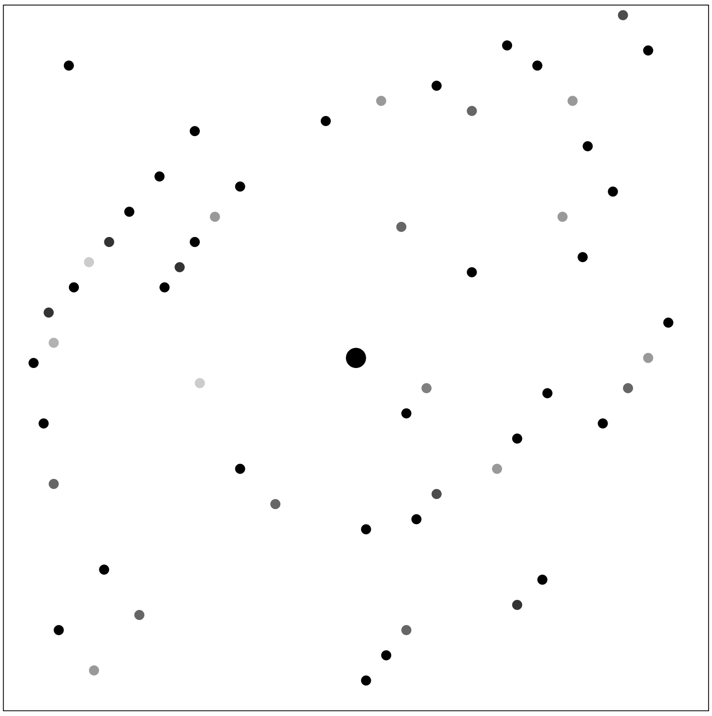
$$\Delta x = n\lambda \qquad (n \in \mathbb{Z})$$

$$\Delta y = m\lambda \qquad (m \in \mathbb{Z})$$

where Δy is determined by b, β_0, β_1

A 3-dimensional array of atoms imposes three conditions.

The information in a diffraction pattern



Each spot indicates a diffracted beam, also called reflection.

A spot is identified by three integer numbers, the Laue indices $h_1h_2h_3$, related to the path difference.

The position of a reflection is determined by the translation lattice of the crystal.

The intensity of a reflection is determined by the contents of the unit cell (type and relative position of the atoms).

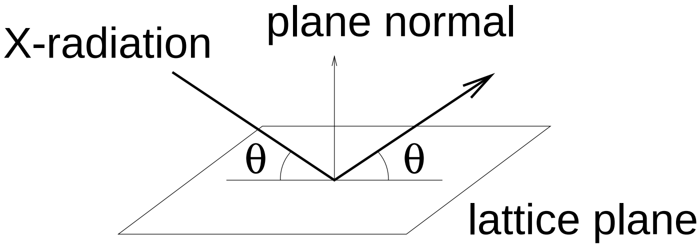
Direction of the diffracted beam

The directions of diffracted beams are determined by the crystal lattice translation vectors $\vec{a}$, $\vec{b}$, $\vec{c}$ and the wavelength of the X-radiation.

To predict the direction of a diffracted beam a transformation of the crystal lattice, the reciprocal lattice, is used:

$$\vec{a^*} = \frac{\vec{b} \times \vec{c}}{V} \qquad \vec{b^*} = \frac{\vec{c} \times \vec{a}}{V} \qquad \vec{c^*} = \frac{\vec{a} \times \vec{b}}{V}$$

A diffracted beam is considered to be a reflection against a lattice plane (= a plane through at least three non-colinear crystal lattice points):



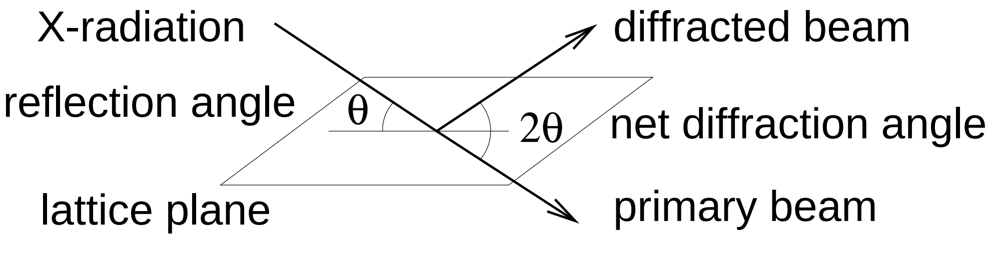
Each reflection is associated with a lattice point $\vec{H}$ in the reciprocal lattice:

$$\vec{H} = h_1\vec{a^*} + h_2\vec{b^*} + h_3\vec{c^*}$$

h_1, h_2 and h_3 are integers and are called the Laue indices (often written as h, k and l, respectively). $\vec{H}$ contains all information on the direction of the diffracted beam:

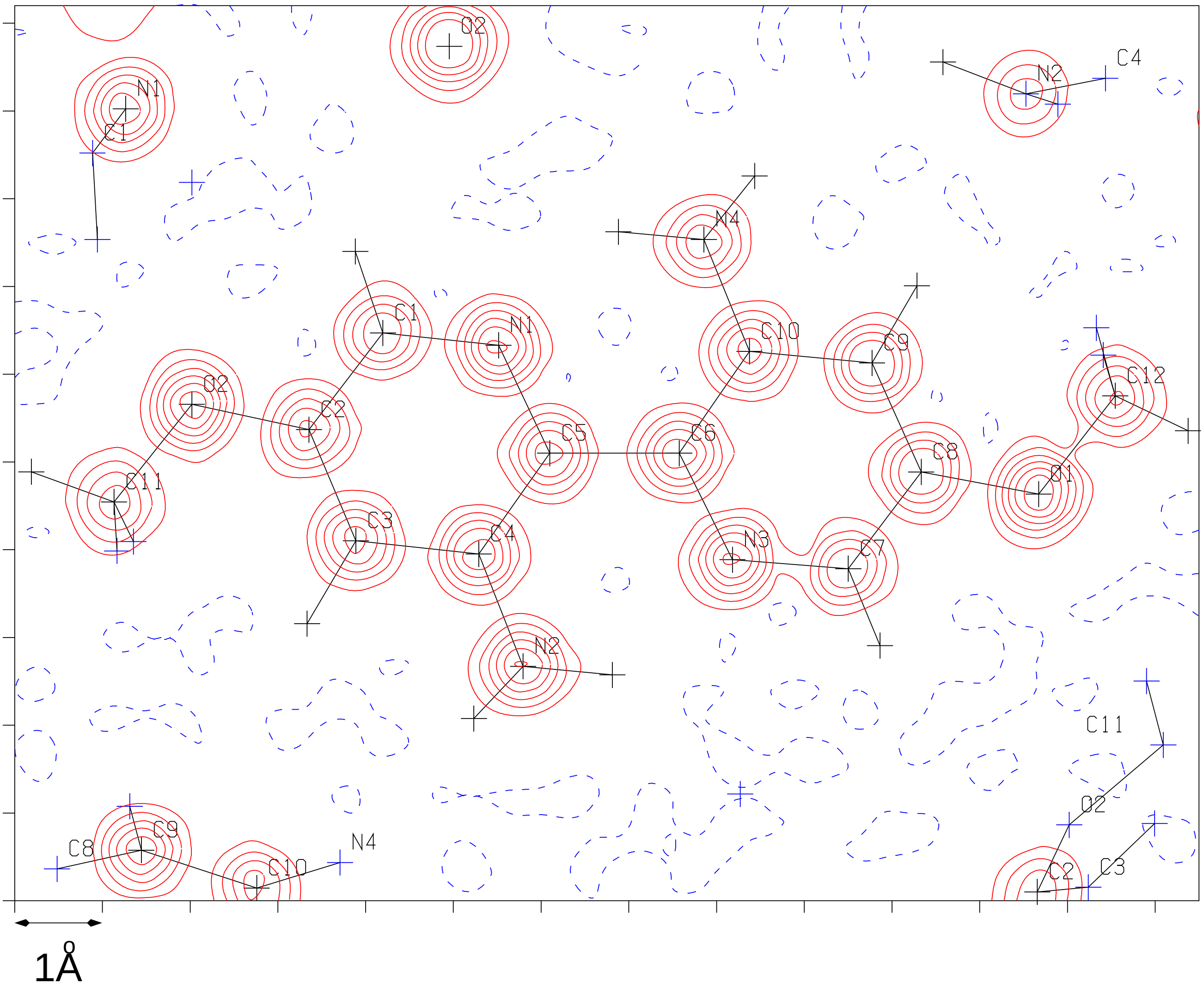
$$\vec{H} \parallel \text{plane normal}$$

$$|\vec{H}| = 2\sin\theta/\lambda$$



Electron density

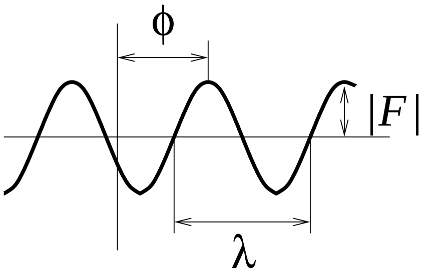
X-radiation is scattered by electrons. To calculate the intensities of the diffraction pattern the distribution of electrons, described with the electron density function $\rho(x, y, z)$, is needed .



Contour plot of the electron density in the least-squares plane through 5,5′-dimethoxy-3,3′-diamine-2,2′-bipyridine. Lines are drawn at intervals of 2.0 e Å^{-3}; the dashed line represent the level of 0.0 e Å^{-3}. A + indicates the position of an atom within 1.0 Å of the plane.

Adding waves

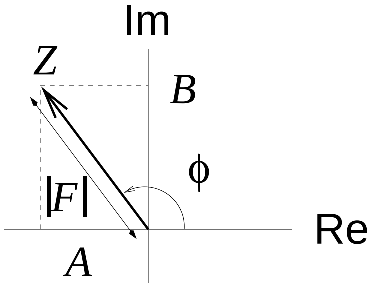
A wave of a given wavelength is characterized by its amplitude $|F|$ and its phase ϕ:



The wave can be represented by a complex number:

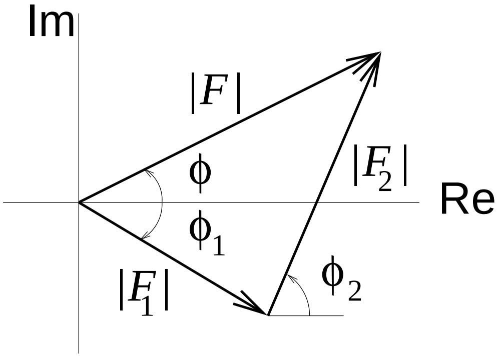
$$Z = A + iB = |F|e^{i\phi} = |F|\cos\phi + i|F|\sin\phi$$

where $i = \sqrt{-1}$



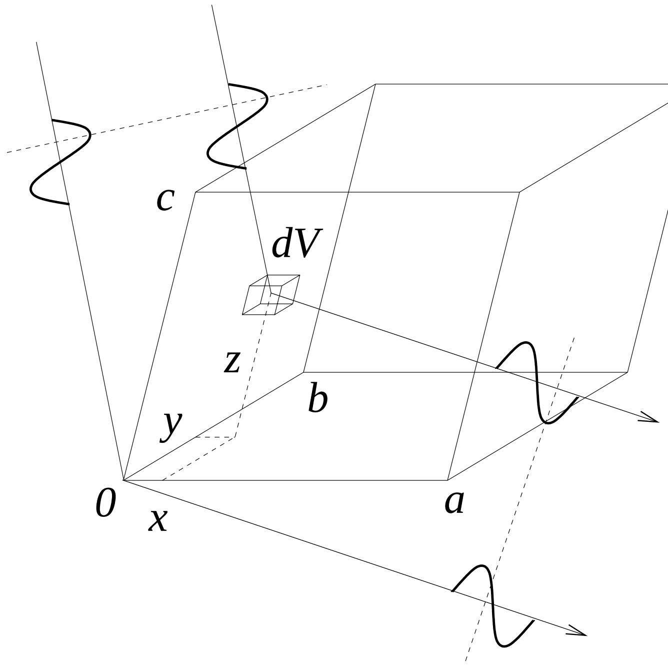
Adding waves amounts to adding complex numbers:

$$Z = Z_1 + Z_2 = |F_1|e^{i\phi_1} + |F_2|e^{i\phi_2} = |F|e^{i\phi}$$



Intensity of the diffracted beam

Consider a reflection $h_1h_2h_3$, with diffraction angle 2θ.



There is a phase difference between the waves diffracted by small volumes dV at $(0,0,0)$ and (x,y,z):

$$\Delta\phi = 2\pi(h_1x + h_2y + h_3z)$$

The amount of electrons in the volume element dV located at (x,y,z) is $\rho(x,y,z)dV$. The wave diffracted by the volume element is

$$\rho(x,y,z)e^{2\pi i(h_1x+h_2y+h_3z)}dV$$

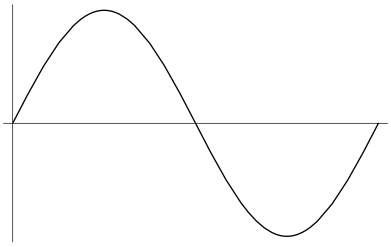
Adding all the waves diffracted by the different volume elements within the unit cell gives the wave diffracted by the whole unit cell:

$$F_{h_1h_2h_3} = V\int_0^1\int_0^1\int_0^1 \rho(x,y,z)e^{2\pi i(h_1x+h_2y+h_3z)}dxdydz$$

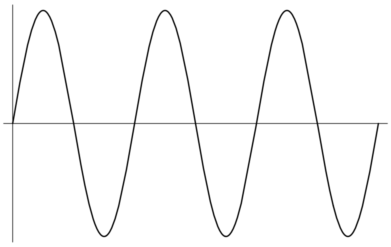
The amplitude of the wave is $|F_{h_1h_2h_3}|$. The intensity $I_{h_1h_2h_3}$ of the wave is proportional to $|F_{h_1h_2h_3}|^2$

Fourier transformation

Any periodic function can be described as the sum of a series of sine and cosine terms, differing in the order n. The coefficients c_n of the terms are given by the Fourier transform of the original function.



$(n = 1)$



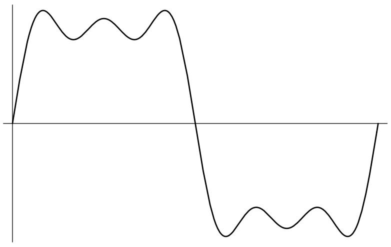
$(n = 2)$

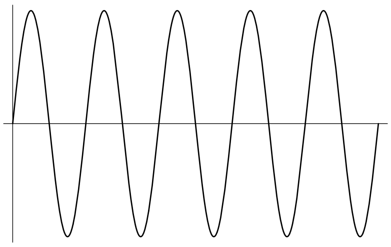
Fourier analysis

$\longleftarrow$

$\longrightarrow$

Fourier synthesis





$(n = 3)$

The Fourier transform of a periodic function $f(x)$ with period L is

$$\begin{aligned} c_n &= \int_0^L f(x) e^{2\pi i n x/L} dx \\ &= \int_0^L f(x)[\cos(2\pi n x/L) + i \sin(2\pi n x/L)] dx \end{aligned}$$

c_n is only non-zero if $n \in \mathbb{Z}$. The coefficient c_n is in principle complex; c_n has both an amplitude and a phase.

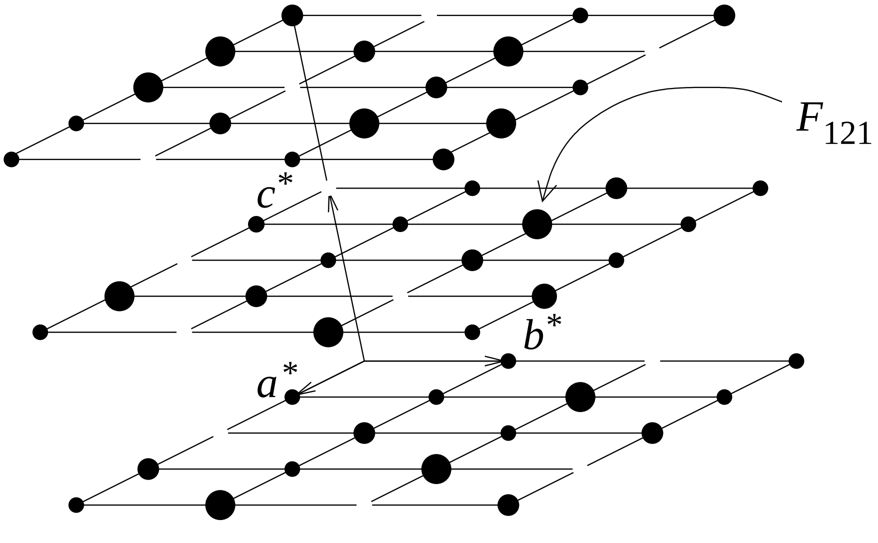
The Fourier transform of $\rho(x, y, z)$

The wave diffracted by the unit cell in the direction associated with $h_1h_2h_3$ is given by

$$F_{h_1h_2h_3} = V \int\limits_{\text{cell}} \rho(x, y, z) e^{2\pi i(h_1x+h_2y+h_3z)} dxdydz$$

$F_{h_1h_2h_3}$ is the 3-dimensional Fourier transform of the electron density $\rho(x, y, z)$. Since $\rho(x, y, z)$ is periodic, $F_{h_1h_2h_3}$ is only non-zero at lattice points.

The diffraction pattern can be represented as the weighted reciprocal lattice: the lattice point (h_1, h_2, h_3) takes the value of $F_{h_1h_2h_3}$.



According to the Fourier theorem $\rho(x, y, z)$ is the inverse Fourier transform of $F_{h_1h_2h_3}$:

$$\rho(x, y, z) = \frac{1}{V} \sum_{h_1} \sum_{h_2} \sum_{h_3} F_{h_1h_2h_3} e^{-2\pi i(h_1x+h_2y+h_3z)}$$

$\rho(x, y, z)$ can be calculated with a Fourier series (a summation, also called Fourier synthesis) because $F_{h_1h_2h_3}$ is only non-zero at lattice points.

Density waves

The electron density expression

$$\rho(x,y,z) = \frac{1}{V}\sum_{h_1h_2h_3} F_{h_1h_2h_3}e^{-2\pi i(h_1x+h_2y+h_3z)}$$

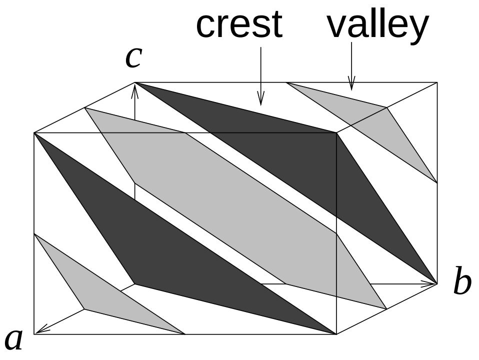
can be re-arranged to

$$\rho(x,y,z) = \frac{F_{000}}{V} + \frac{2}{V}\sum_{h'_1h'_2h'_3} |F_{h_1h_2h_3}|\cos[2\pi(h_1x+h_2y+h_3z) - \phi_{h_1h_2h_3}]$$

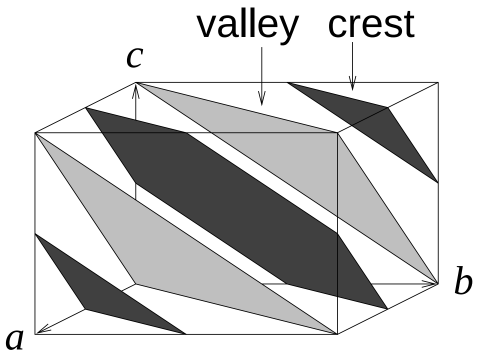
where $\sum_{h'_1h'_2h'_3}$ indicates that only one member of the centrosymmetric pair $(h_1,h_2,h_3),(-h_1,-h_2,-h_3)$ is included in the summation.

The interpretation of this equation is that the electron density can be constructed out of 3-dimensional density waves:

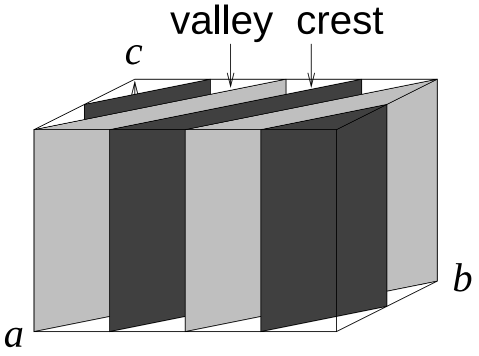
$1\ \bar{1}\ \bar{1}$ density wave, $\phi = 0^\circ$



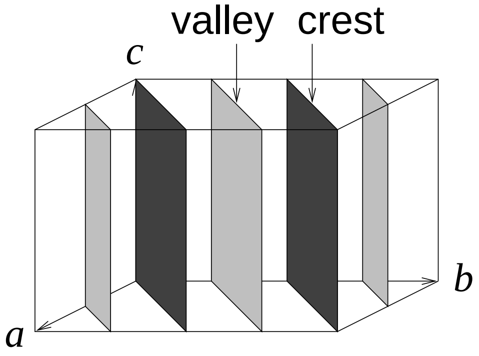
$1\ \bar{1}\ \bar{1}$ density wave, $\phi = 180^\circ$



$1\ 2\ 0$ density wave, $\phi = 180^\circ$

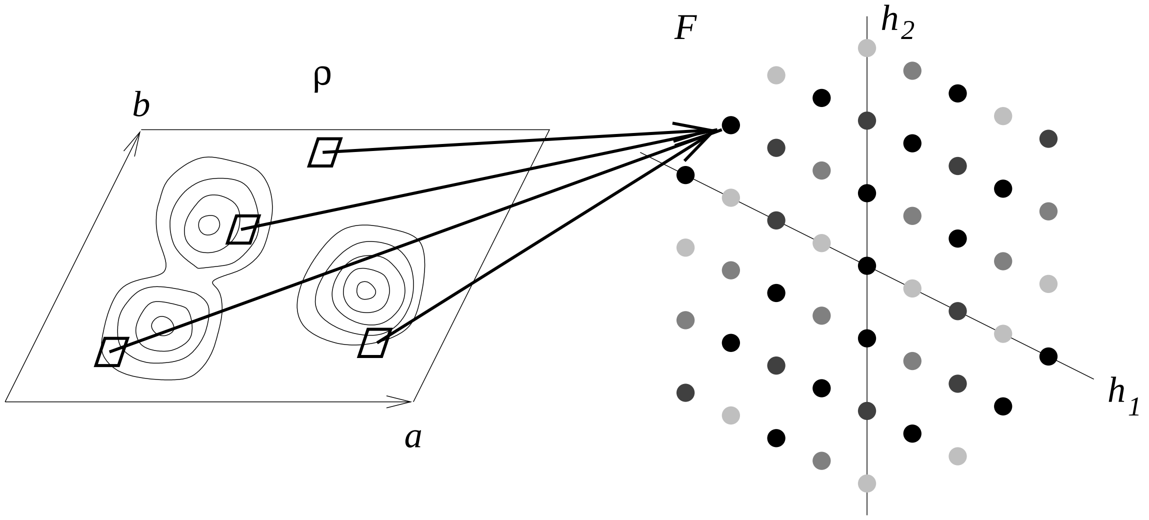


$1\ \bar{2}\ 0$ density wave, $\phi = 0^\circ$



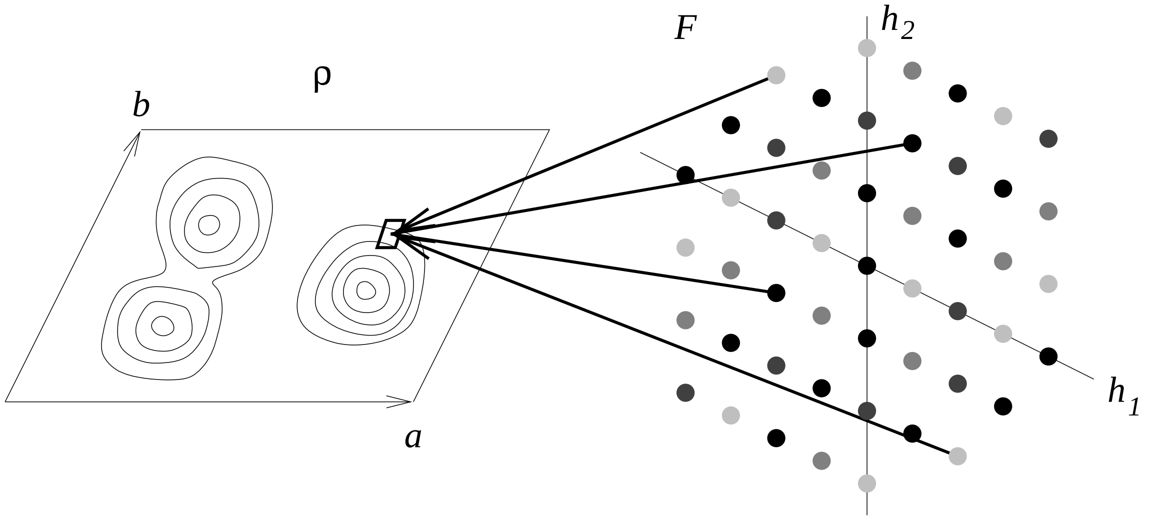
The relations between $\rho(x, y, z)$ and $F_{h_1h_2h_3}$

$$F_{h_1h_2h_3} = V \int_{\text{cell}} \rho(x, y, z) e^{2\pi i(h_1x+h_2y+h_3z)} dxdydz$$



The electron density at all elements $dxdydz$ contributes to a single structure factor $F_{h_1h_2h_3}$.

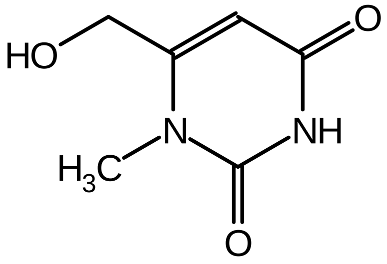
$$\rho(x, y, z) = \frac{1}{V} \sum_{h_1h_2h_3} F_{h_1h_2h_3} e^{-2\pi i(h_1x+h_2y+h_3z)}$$

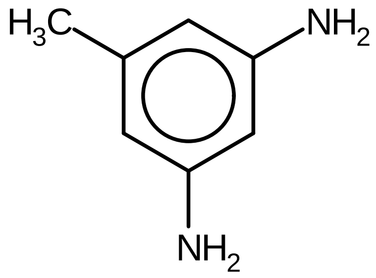


All structure factors $F_{h_1h_2h_3}$ contribute to the electron density at a particular point (x, y, z).

Resolution (1)

Consider the crystal structure of





H_2O

Resolution, expressed in Å, refers to the value of

$$\frac{\lambda}{2\sin\theta^{\mathsf{max}}} = \frac{1}{|\vec{H}|^{\mathsf{max}}}$$

where

$$\vec{H} = h_1\vec{a^*} + h_2\vec{b^*} + h_3\vec{c^*}$$

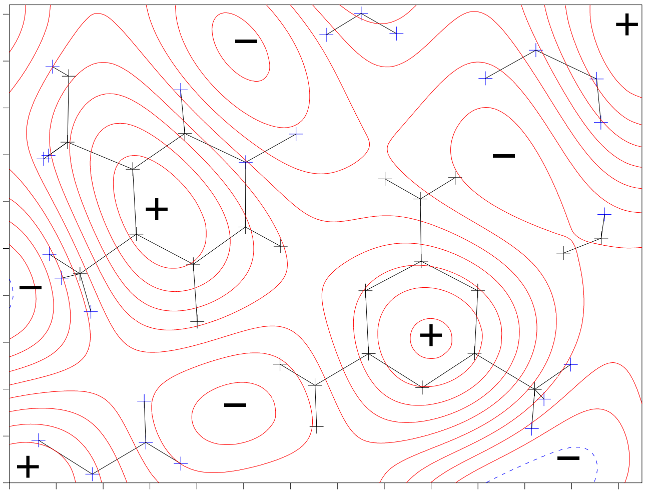
Other measures of resolution are $\sin(\theta^{\mathsf{max}})$ (wavelength dependent) and $\sin(\theta^{\mathsf{max}})/\lambda$ (wavelength independent).

Resolution (Å)	θ^{max} [Mo$K\alpha$] (°)	θ^{max} [Cu$K\alpha$] (°)	$\sin(\theta^{\mathsf{max}})/\lambda$	# reflections[a]
4.5	4.5	9.8	0.11	16
3.1	6.5	14.2	0.16	44
2.4	8.5	18.7	0.21	93
1.9	11.0	24.5	0.27	205
1.4	15.0	34.2	0.36	
1.0	20.0	47.9	0.40	1209
0.83	25.3	67.7	0.60[b]	2385
0.77	27.5	—	0.65	

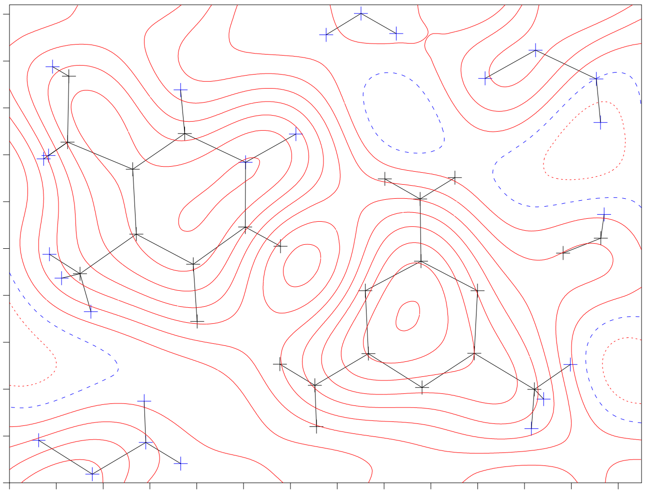
[a] for the given example

[b] Acta Crystallogr. norm

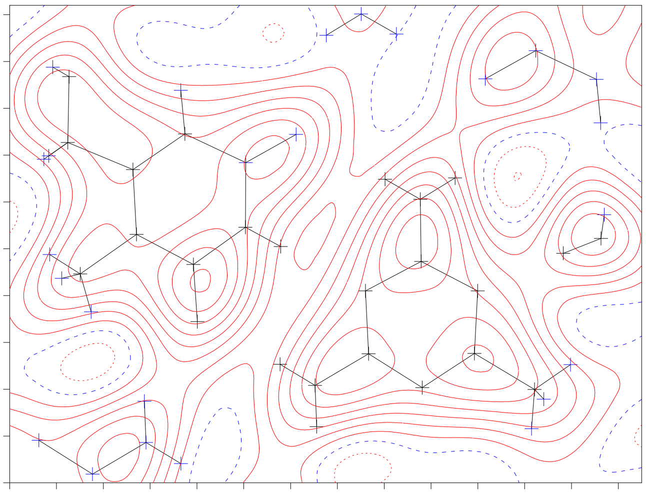
Resolution (2)



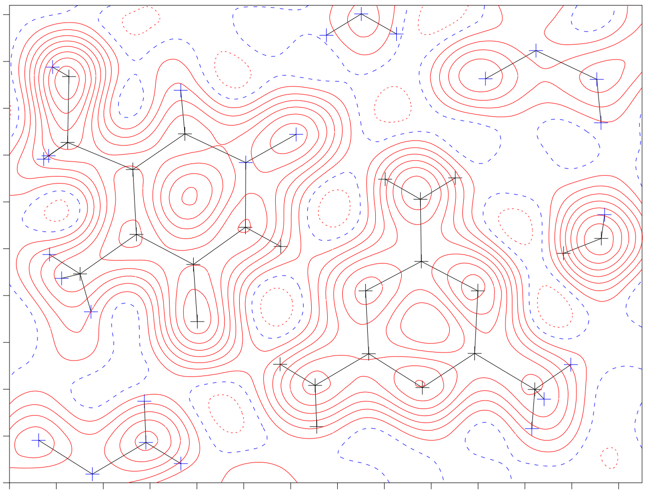
4.5 Å



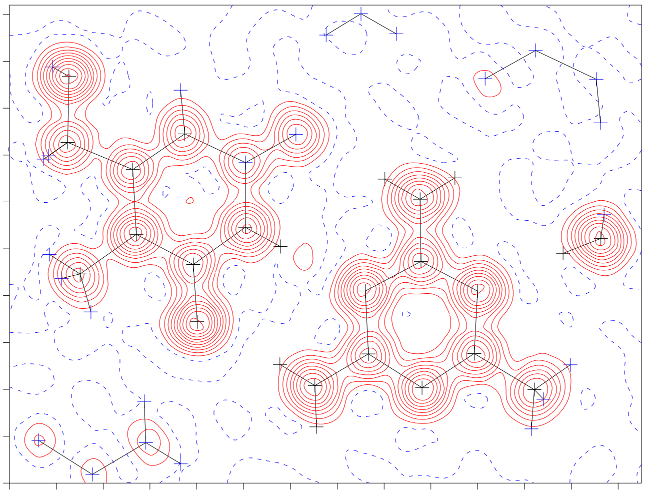
3.1 Å



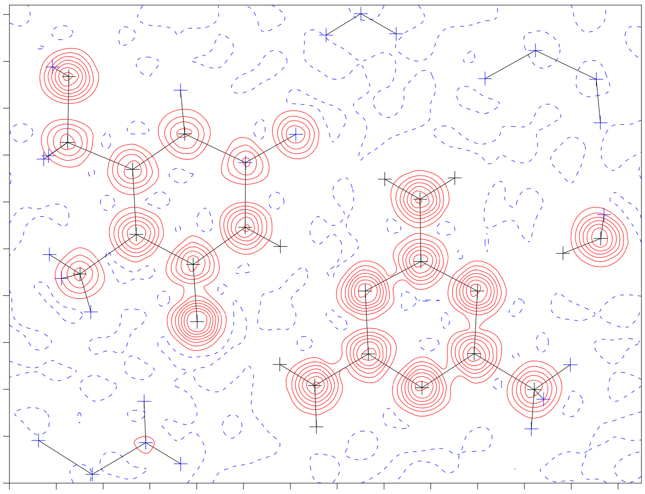
2.4 Å



1.9 Å



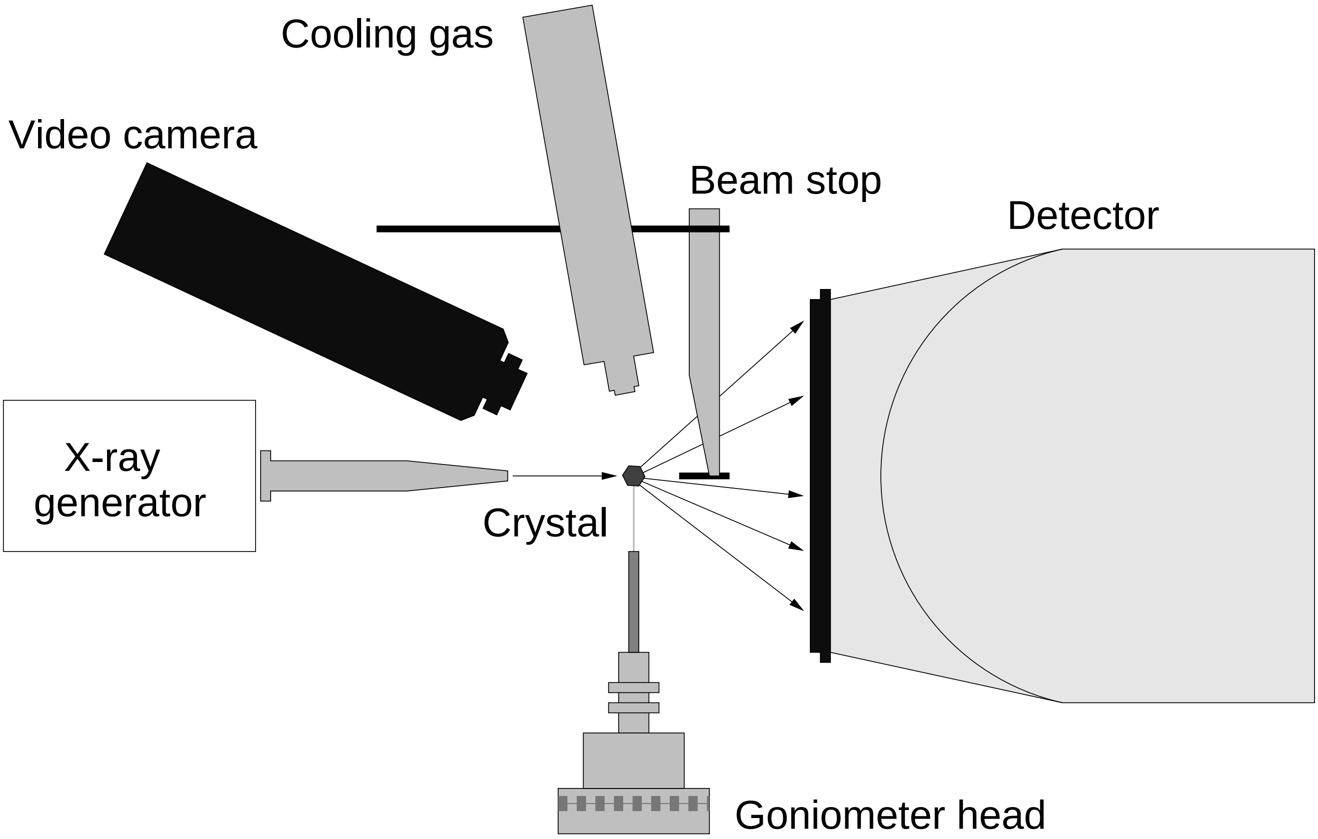
1.0 Å



0.83 Å

Experimental set-up

The National Service Facility uses a Nonius KappaCCD diffractometer on rotating anode as the main instrument.



X-ray generator: molybdenum rotating anode (60kV, 50 mA), fine focus primary beam (∅ 0.3 mm), wavelength $\lambda_{\mathrm{MoK}\bar{\alpha}} = 0.71073$ Å.

Cooling gas: Evaporated liquid nitrogen, heated to a specified temperature (working range: ca 100–360 K)

Detector 512×512 pixel image, recorded on CCD chip

Distance crystal to detector: 25–175 mm

Video camera for crystal monitoring

The phase problem

$\rho(x, y, z)$ can be calculated when $F_{h_1h_2h_3}$ is known:

$$\rho(x,y,z) = \frac{F_{000}}{V} + \frac{2}{V} \sum_{h_1'h_2'h_3'} \underbrace{|F_{h_1\vec{h}_2h_3}|}_{\text{known}} \cos[2\pi(h_1x + h_2y + h_3z) - \underbrace{\phi_{h_1h_2h_3}}_{\text{unknown}}]$$

The intensity of the diffracted beam is related to the structure factor:

$$\begin{aligned} I_{h_1h_2h_3} &\sim F_{h_1h_2h_3}F^*_{h_1h_2h_3} \\ &= |F_{h_1h_2h_3}|e^{i\phi_{h_1h_2h_3}}|F_{h_1h_2h_3}|e^{-i\phi_{h_1h_2h_3}} \\ &= |F_{h_1h_2h_3}|^2 \end{aligned}$$

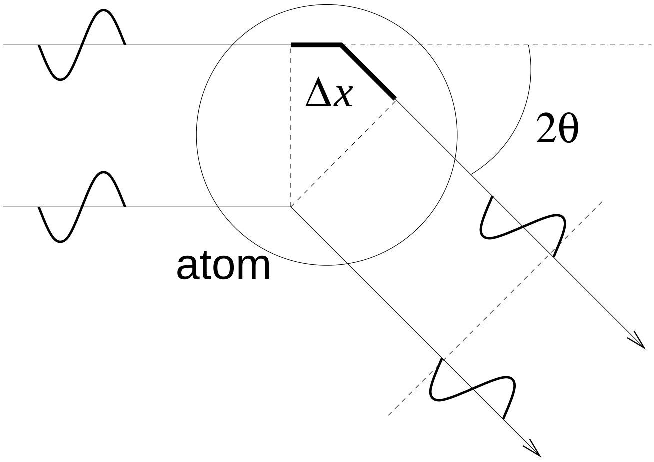
The phase information can therefore not be retrieved from the experiment.

In chemical crystallography the most important solutions to the phase problem are:

- Direct methods
- Patterson methods

Scattering factors (1)

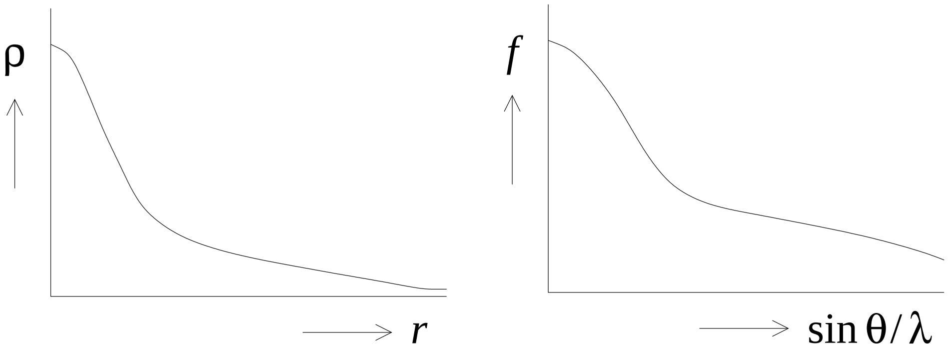
The dimensions of atoms are of the same order of magnitude as the wavelength of the X-radiation used in diffraction experiments. The intensity of a beam diffracted by a single atom is a function of the scattering angle 2θ.



The amount of radiation scattered in each direction is described with the scattering factor f, also called form factor.

The unit of f is the amount of radiation scattered by 1 free electron.

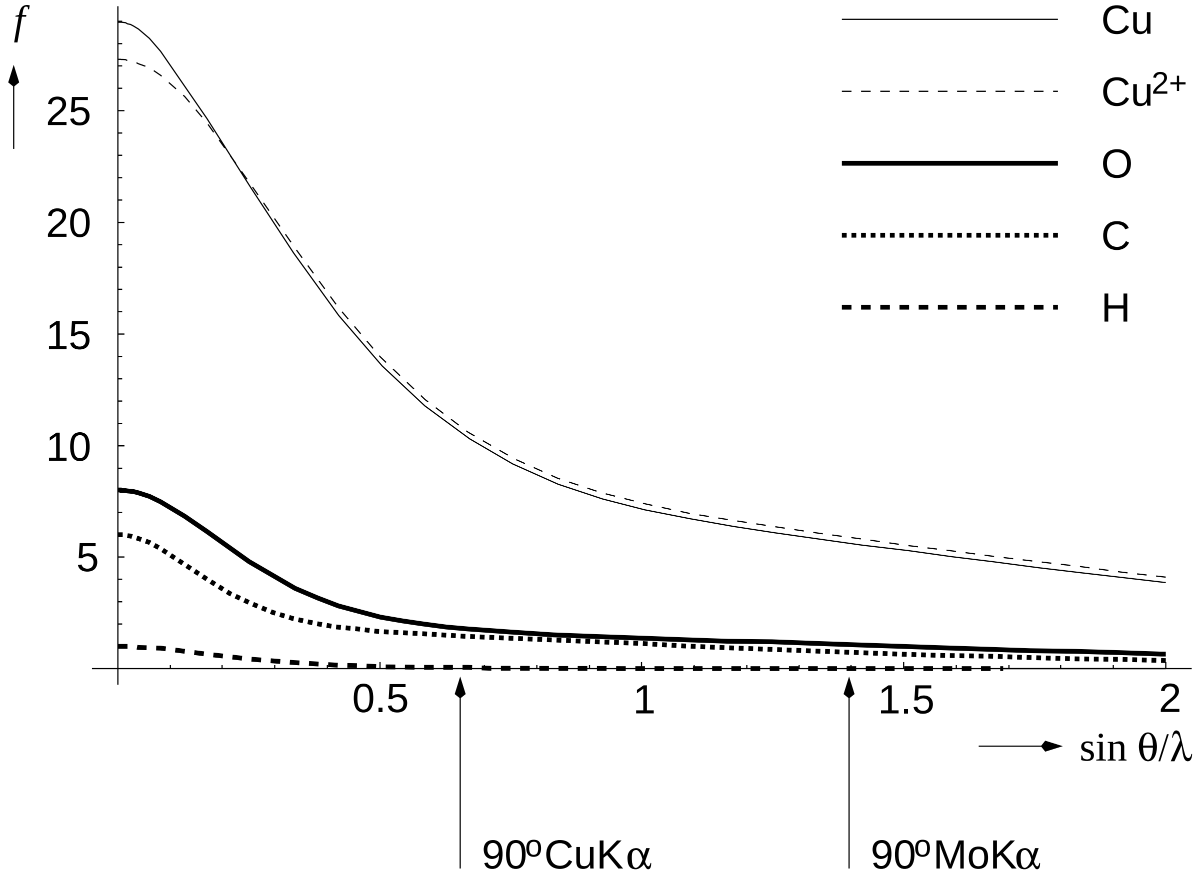
The scattering factor of an atom is the Fourier transform of the electron density.



For most known atoms and monatomic ions f is available for appropriate values of $\sin\theta/\lambda$ in tabular form and as a parameterized function.

Scattering factors (2)

The scattering factors of a few atoms and ions are given below:



The scattering factor at $\sin\theta/\lambda = 0°$ is equal to the number of electrons in the atom.

Hydrogen atoms scatter weakly and can only be detected with X-ray diffraction when accurate intensity measurements are used. Low order reflections contain most information about hydrogen atoms.

The scattering factors of Cu and Cu^{2+} show only small relative differences. In normal structure determinations the neutral atom scattering factors are used.

Structure factors and the atomic model

The structure factor is the Fourier transform of the electron density:

$$F_{h_1h_2h_3} = V \int_{\text{cell}} \rho(x,y,z) e^{2\pi i(h_1x+h_2y+h_3z)} dxdydz$$

The electron density can be approximated as a superposition of N non-interacting atoms located at positions

$$\vec{r}_j = x_j\vec{a} + y_j\vec{b} + z_j\vec{c}$$

$F_{h_1h_2h_3}$ can be expressed as the sum of the Fourier transforms of the atomic electron densities:

$$F_{h_1h_2h_3} = \sum_{j=1}^{N} f_j^{\sin\theta/\lambda} e^{2\pi i(h_1x_j+h_2y_j+h_3z_j)}$$

were f_j, the atomic scattering factor of atom j, is a function of $sin\theta/\lambda$.

This equation enables the calculation of structure factors belonging to an atomic model of the crystal structure.

A model can be checked by comparing calculated structure factor amplitudes with observed structure factor amplitudes.

The Patterson method

Principle

All data are available to calculate the following function:

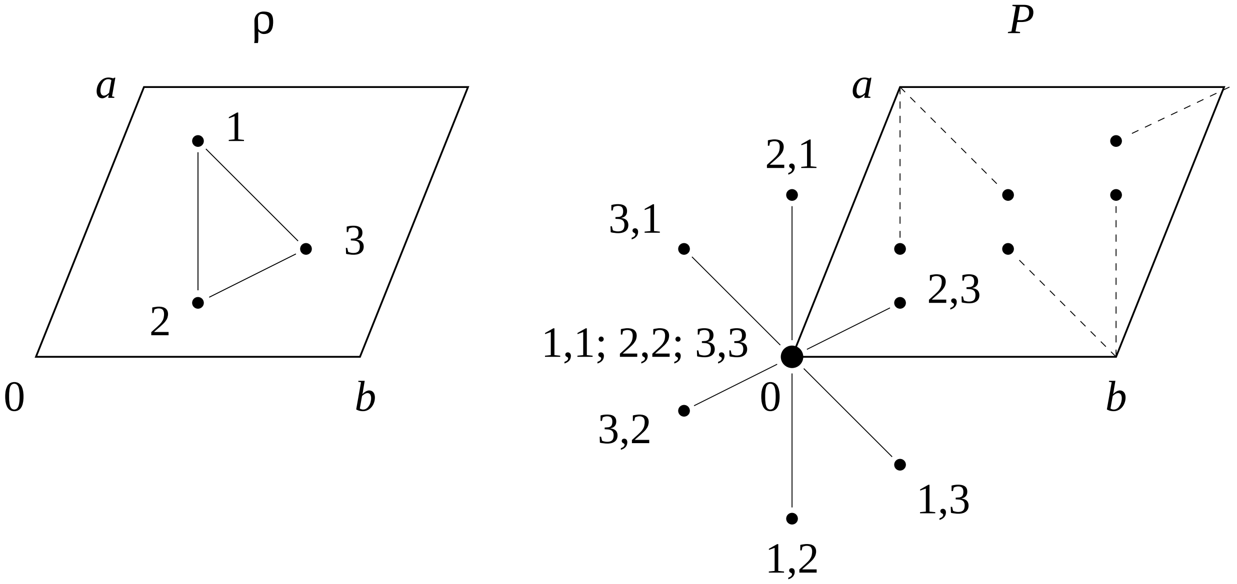
$$P(u,v,w) = \frac{1}{V} \sum_{h_1 h_2 h_3} |F_{h_1 h_2 h_3}|^2 e^{-2\pi i (h_1 u + h_2 v + h_3 w)}$$

which can be interpreted as:

$$P(u,v,w) = \int_{\text{cell}} \rho(x,y,z)\rho(x+u, y+v, z+w) dx dy dz$$

- $P(u,v,w)$ contains peaks for all $N(N-1)$ interatomic vectors
- the peak height is proportional to $Z_j \times Z_k$
  (e.g. C → C : 36, Fe → C : 156, Fe → Fe : 676)

Example



Application

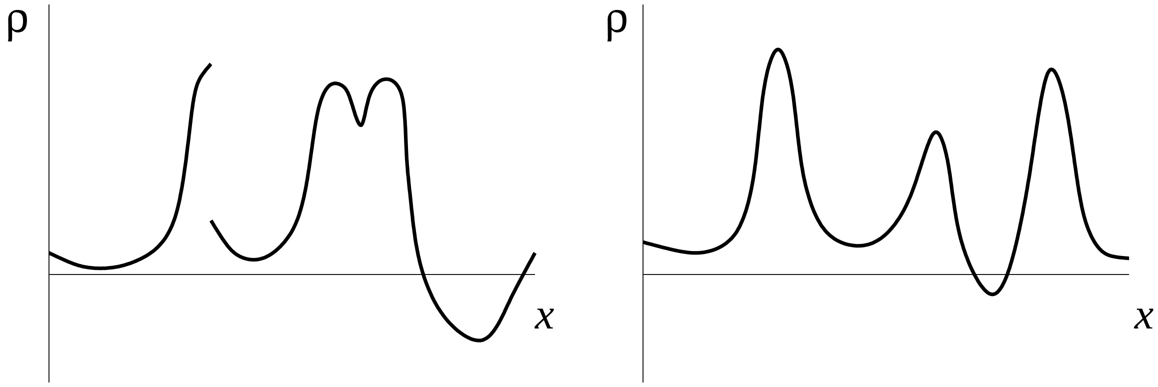
- Deduce heavy atom positions for structures with a few heavy atoms
- Locate fragments with known geometry

Direct methods

Principle

Statistical methods make it possible to derive information concerning the phases from the magnitudes of the structure factors.

The method is based on the assumption that $\rho(x, y, z)$ consists of atoms. If this assumption holds (b) is a reasonable electron density while (a) is not.



(a) (b)

Example

The triple product phase relationship

$$\phi_{h_1,h_2,h_3} + \phi_{k_1,k_2,k_3} + \phi_{-h_1-k_1,-h_2-k_2,-h_3-k_3} \approx 0$$

is a probability relation. The numerical value of the probability can be calculated. The probability becomes higher if the associated structure factor amplitudes become larger. The probability decreases as the number of atoms N increases.

Application

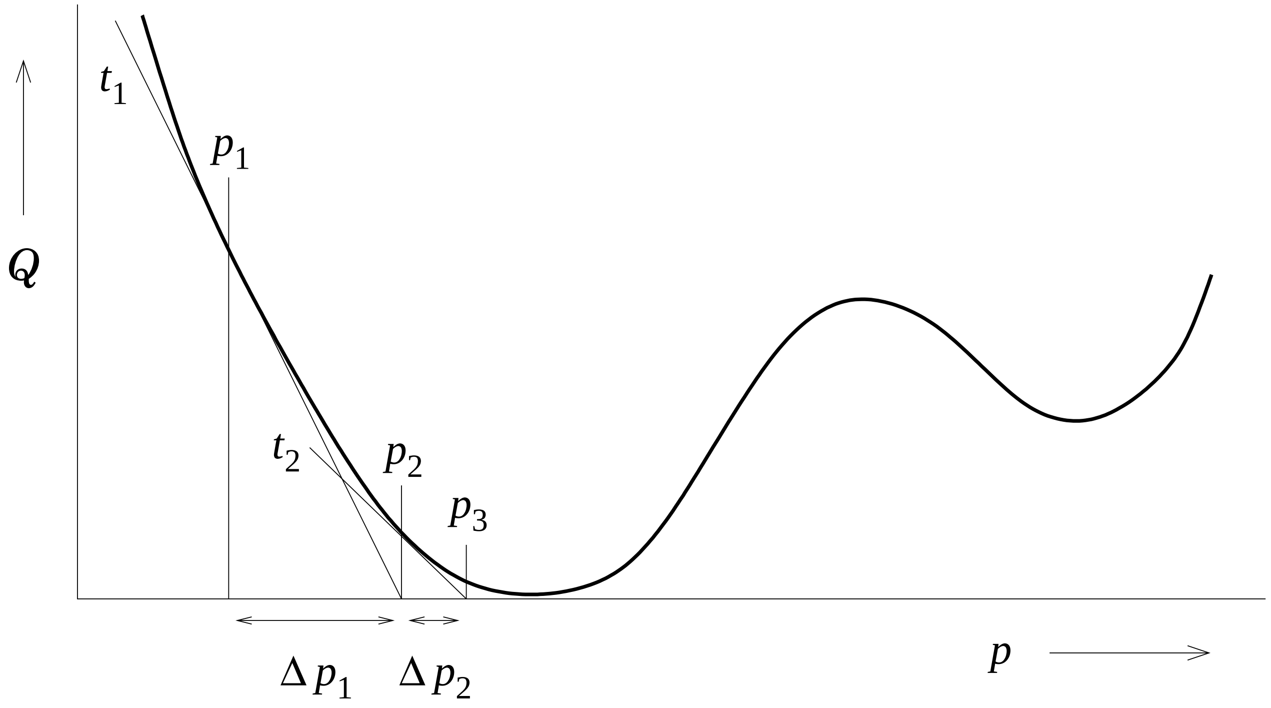
Determine phases for structures containing atoms with approximately equal atomic numbers Z.

Refinement of the model

Patterson and direct methods give a rough model of the crystal structure, expressed in atomic positions. The observed $|F^{\mathsf{obs}}_{h_1h_2h_3}|$ can now be compared to $|F^{\mathsf{calc}}_{h_1h_2h_3}|$, calculated from the atomic positions, with the function Q

$$Q = \sum_{h_1h_2h_3} w(|F^{\mathsf{obs}}_{h_1h_2h_3}|^2 - |F^{\mathsf{calc}}_{h_1h_2h_3}|^2)^2$$

Non-linear least-squares minimization of Q will result in parameter shifts, derived from tangents of Q, leading to a better model.



The calculation of Δp gives a standard uncertainty (s.u., also called estimated standard deviation, e.s.d.) for each p. The refinement continues until

$$\Delta p \ll \mathsf{s.u.}(p)$$

The initial value of the parameters should be within the convergence radius of the global minimum.

Model parameters

The refinement of the atomic model is based on the overdeterminancy of the problem. The ratio of observations ($|F^{\mathsf{obs}}_{h_1h_2h_3}|$, n) to parameters (p) should at least be 5:1. Most journals demand a higher ratio.

Parameters used in the model include:

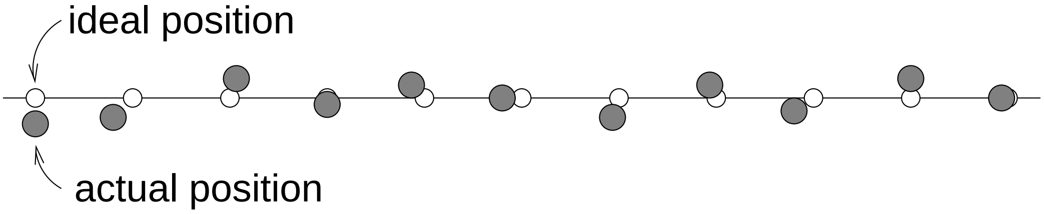
- atomic parameters
  - element type (scattering factor; not refined)
  - positional parameters (x, y, z)
  - displacement parameters (U_{ij})
  - occupancy parameter
- global parameters
  - overall scale factor
  - overall displacement parameter
  - parameterized empirical corrections (e.g. extinction)
  - chirality parameter

Improvement of the $n : p$ ratio:

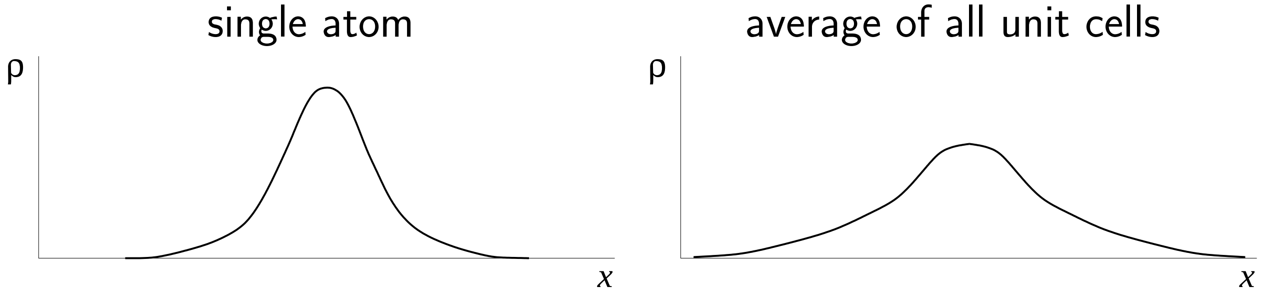
- constraints: reduce p
  Example: describe benzene ring as regular hexagon (6 parameters in stead of 18).
  Disadvantage: geometrical parameters are given ideal values, the "real" values can not be determined.
- restraints: increase n
  Example: add an equation (with a chosen weight) imposing a known bond length to the minimization function Q
  Disadvantage: derived geometrical data are not solely based on observations.

Atomic displacement

Atoms in a crystal lattice vibrate around their equilibrium positions. The vibration frequency is much smaller than the X-ray frequency. X-radiation therefore interacts with a lattice of stationary, randomly displaced atoms.



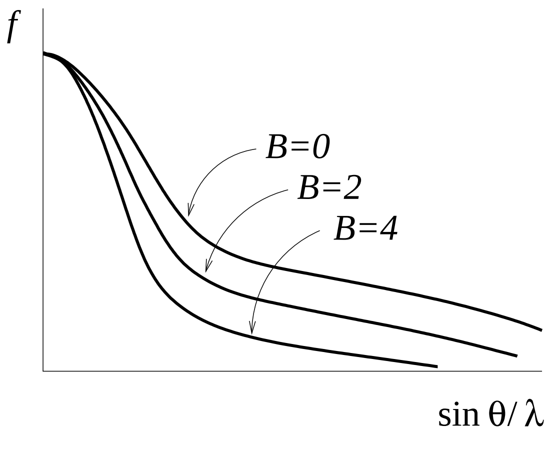
In the structure determination all unit cells are averaged. The result is a "smeared out" atom:



A displacement correction (formerly called temperature correction) has to be added to the scattering factor. Like f^0, the correction is a function of $\sin\theta/\lambda$:

$$f = f^0 T = f^0 e^{-B\sin^2\theta/\lambda^2}$$

where $B = 8\pi < u^2 >$, with u the vibration amplitude.



Displacement parameters

To account for atomic displacement a correction has to be applied on the scattering factor:

$$f = Tf^0$$

Isotropic displacement correction

$$T = e^{-B_{\mathsf{iso}} \sin^2\theta/\lambda^2} = e^{-8\pi^2 U_{\mathsf{iso}} \sin^2\theta/\lambda^2}$$

$$B_{\mathsf{iso}} = 8\pi^2 U_{\mathsf{iso}} = 8\pi^2 < u^2 >$$

$< u^2 >$ is the mean square displacement in Å^2.

In case of an overall isotropic displacement parameter, U_{iso} is set to the same value for all atoms.

Anisotropic displacement correction

The displacement amplitude can be different for each main direction:

$$T = e^{-h^2\beta_{11}+k^2\beta_{22}+l^2\beta_{33}+2hk\beta_{12}+2hl\beta_{13}+2kl\beta_{23}}$$

$$T = e^{-\frac{1}{4}(h^2B_{11}a^{*2}+k^2B_{22}b^{*2}+l^2B_{33}c^{*2}+2hkB_{12}a^*b^*+2hlB_{13}a^*c^*+2klB_{23}b^*c^*)}$$

$$T = e^{-2\pi^2(h^2U_{11}a^{*2}+k^2U_{22}b^{*2}+l^2U_{33}c^{*2}+2hkU_{12}a^*b^*+2hlU_{13}a^*c^*+2klU_{23}b^*c^*)}$$

$$\beta_{11} = \frac{1}{4}a^{*2}B_{11} \qquad B_{11} = 8\pi^2 U_{11} \qquad \beta_{11} = 2\pi^2 a^{*2} U_{11}$$

$$\beta_{12} = \frac{1}{4}a^*b^*B_{12} \qquad B_{12} = 8\pi^2 U_{12} \qquad \beta_{12} = 2\pi^2 a^*b^* U_{12}$$

U_1, U_2, U_3 are the main axes components of U_{ij}.

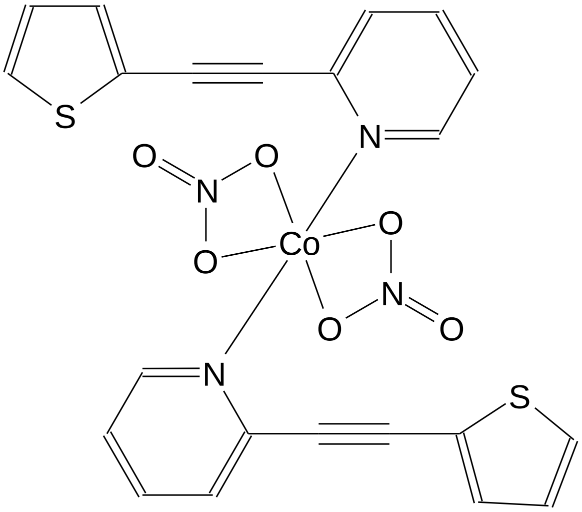
U_1/U_3 is a measure for the anisotropicity.

U_{eq} is the equivalent isotropic displacement parameter:

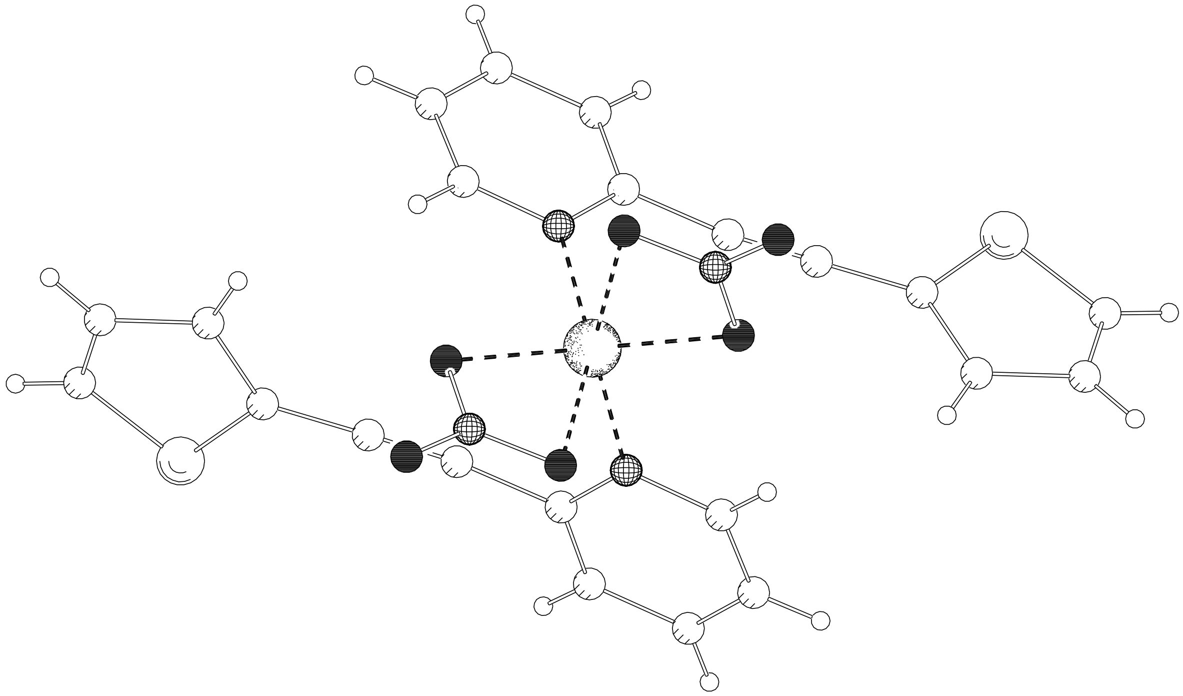
$$U_{\mathsf{eq}} = \frac{1}{3}\sum_{i=1}^{3}\sum_{j=1}^{3} U_{ij} a_i^* a_j^* \vec{a}_i \cdot \vec{a}_j$$

Disorder (1)

The complex



was expected to have the following structure:

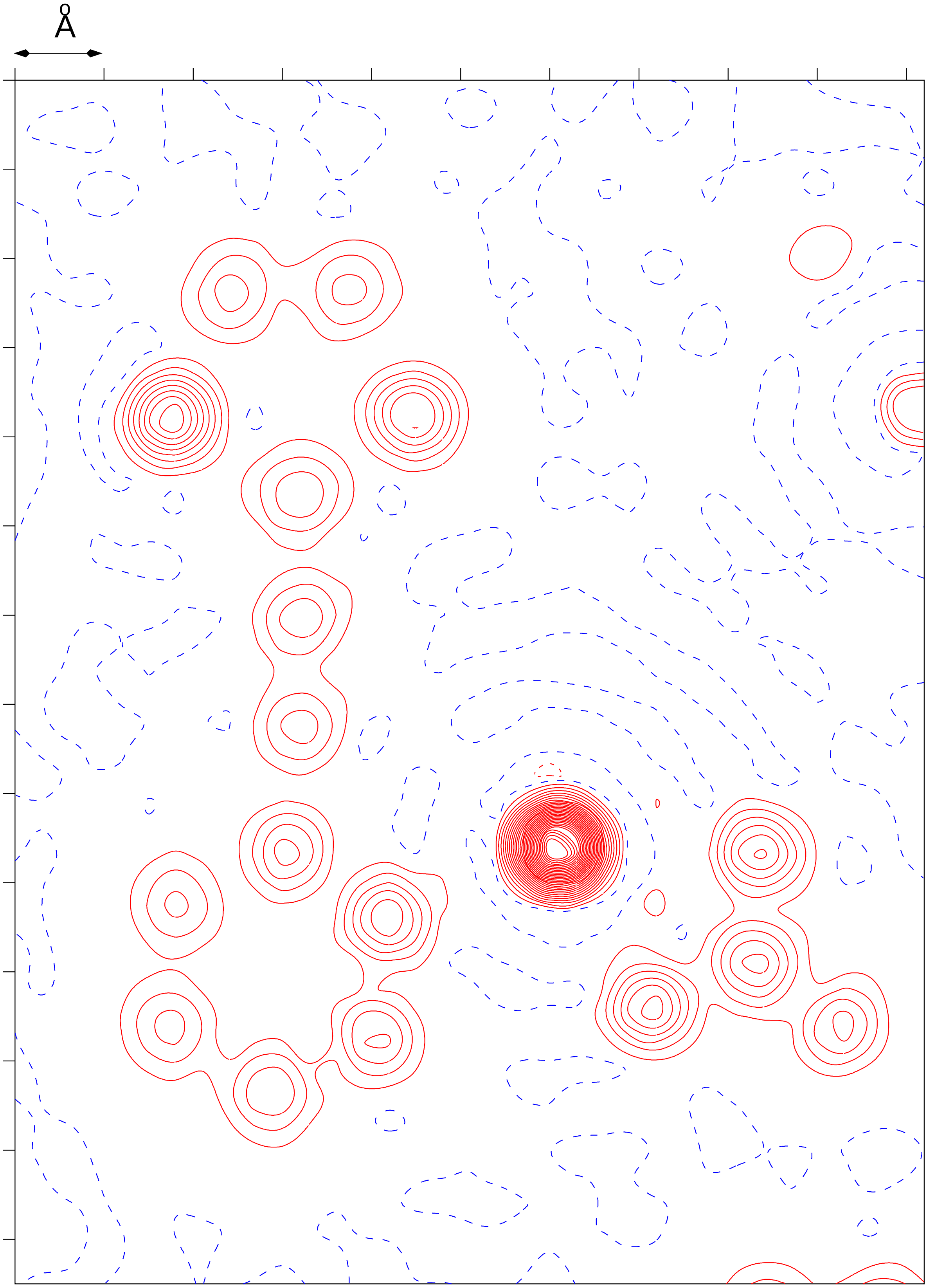


with all atoms more or less in one of two perpendicular planes.

Kooijman, Spek, Neenan & Driessen, *Acta Crystallogr.* **C52** (1996) 2191–2193

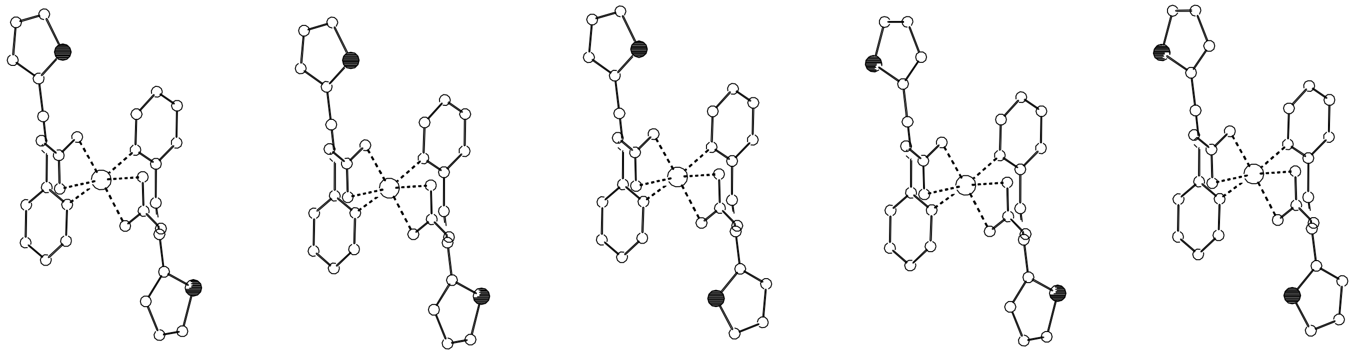
Disorder (2)

The following electron density was found in one of the planes:

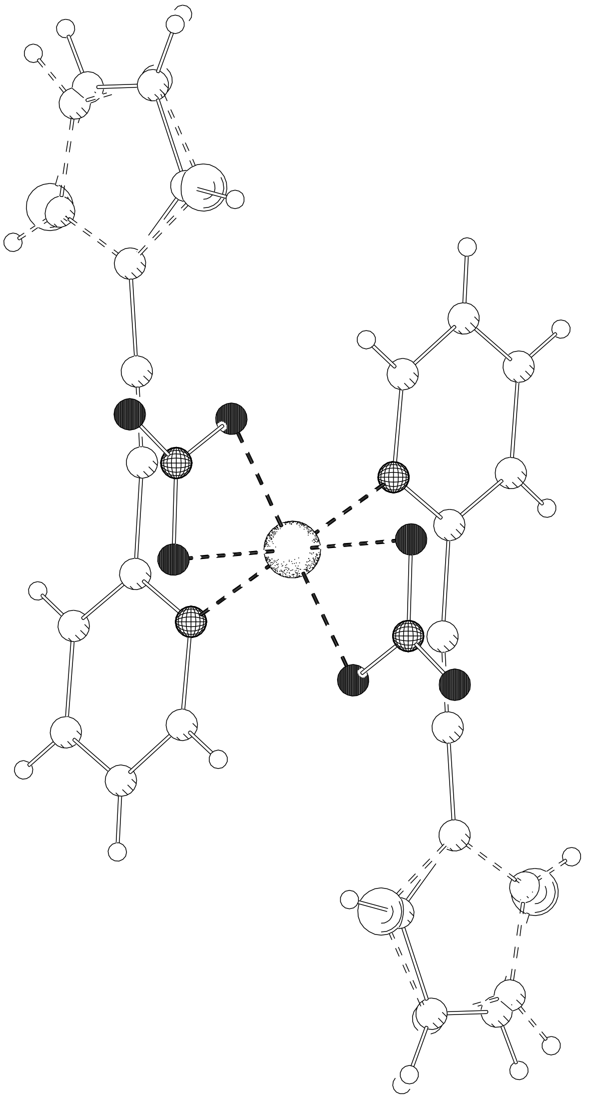


Disorder (3)

The crystal structure can differ slightly from unit cell to unit cell:

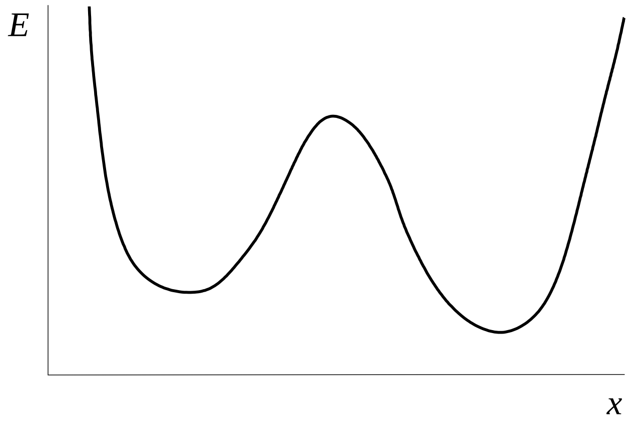


causing intensity at non-integer values of $h_1 h_2 h_3$. The calculation of the Fourier series enforces strict translation symmetry, amounting to averaging all unit cells, both in space and time. The following result is obtained:

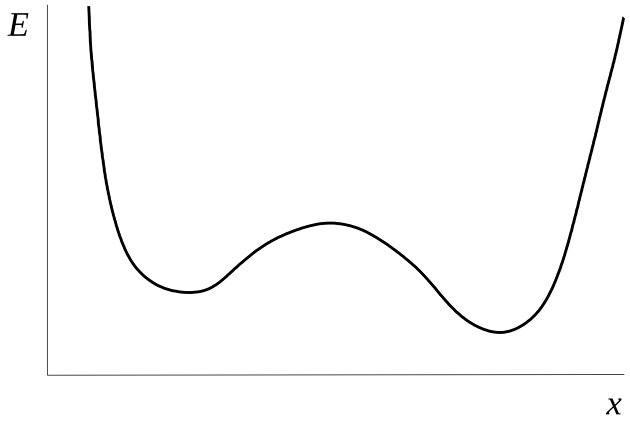


Disorder types

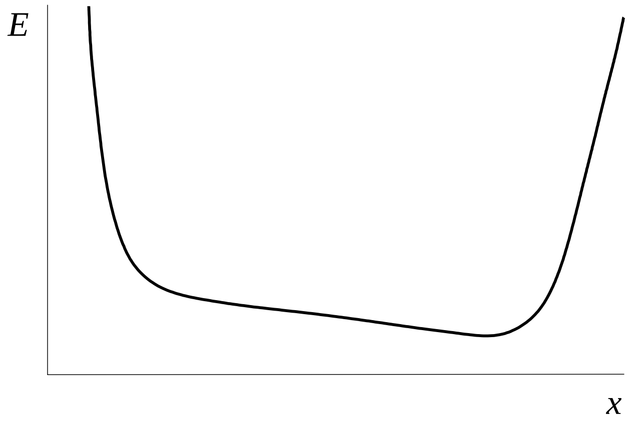
Disorder occurs when two or more sites are available for one atom. The following situations are possible:



- two distinct sites
- no interchange possible
- occupancy can be refined
- static disorder



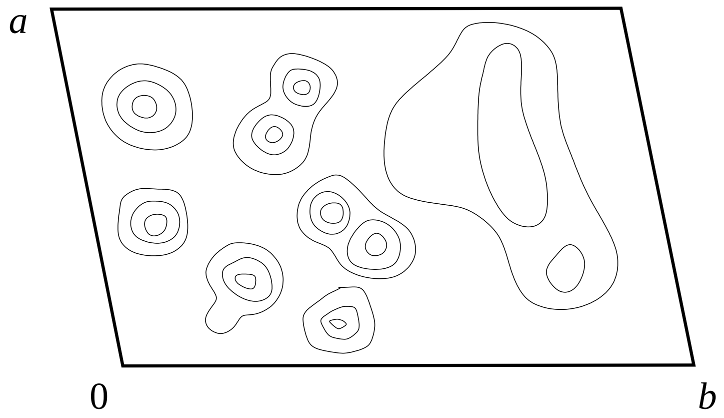
- two distinct sites
- interchange possible
- occupancy can be refined
- dynamic disorder



- no distinct sites
- averaged geometry
- high displacement parameter
- high anisotropicity

In case of dynamic disorder, the (refined) value of the occupancy is determined by ΔE. In case of static disorder the preferred conformation in solution can be captured in the crystal.

Uninterpretable disorder



$$\rho(x,y,z) = \rho^{\mathsf{ord}}(x,y,z) + \rho^{\mathsf{dis}}(x,y,z)$$

The PLATON/SQUEEZE procedure seperates ordered and disordered parts of the electron density.

$$\begin{aligned}
F_{h_1h_2h_3} &= \mathcal{F}[\rho(x,y,z)] \\
&= \mathcal{F}[\rho^{\mathsf{ord}}(x,y,z)] + \mathcal{F}[\rho^{\mathsf{dis}}(x,y,z)] \\
&= \sum_j f_j e^{2\pi i(h_1x+h_2y+h_3z)} + \mathcal{F}^{\mathsf{n}}[\rho^{\mathsf{dis}}(x,y,z)] \\
F'_{h_1h_2h_3} &= F_{h_1h_2h_3} - \mathcal{F}^{\mathsf{n}}[\rho^{\mathsf{dis}}(x,y,z)] \\
&= \sum_j f_j e^{2\pi i(h_1x+h_2y+h_3z)}
\end{aligned}$$

where $\mathcal{F}$ stands for an analytical Fourier transformation and $\mathcal{F}^{\mathsf{n}}$ for a numerical Fourier transformation.

Refinement against $F'_{h_1h_2h_3}$ improves the geometry of the ordered part of the structure.

Geometrical information of the disordered part is lost, only the volume and the number of electrons present are obtained.

The procedure can only be applied if the ordered and disordered parts can be clearly separated.

Identification of a SQUEEZE'd molecule

A heavily disordered part of the electron density can not always be identified. Candidates are solvents used in synthesis and crystallization, free ligand, counter ions, side products and water.

Solvent sites are not necessarily fully occupied, the observed number of electrons can therefore be misleading.

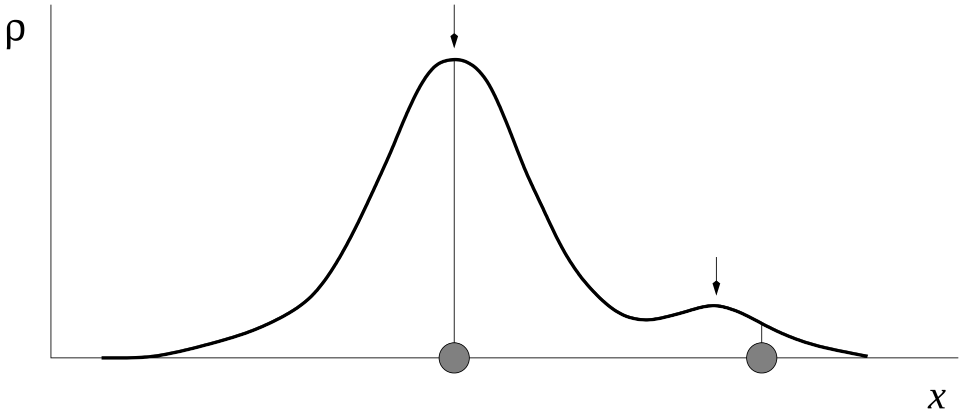
The volume of the disordered region should be comparable to the molecular volume in the liquid phase (V^{liq}).

The shape of the disordered region can also be useful in identification.

solvent/ion	formula	M_r (g)	V^{liq} (Å^3)	# e
water	H_2O	18.02	29.92	10
methanol	CH_4O	32.04	67.28	18
acetonitril	C_2H_3N	41.05	86.76	22
ethanol	C_2H_5OH	46.07	96.92	26
aceton	C_3H_6O	58.08	122.14	32
tetrahydrofuran	C_4H_8O	72.12	134.68	40
methylenechloride	CH_2Cl_2	84.93	106.31	42
diethylether	$C_4H_{10}O$	74.12	172.44	42
pentane	C_5H_{12}	72.15	191.33	42
benzene	C_6H_6	78.12	147.64	42
1,4-dioxane	$C_4H_8O_2$	88.11	141.54	48
hexane	C_6H_{14}	86.18	216.76	50
toluene	C_7H_8	92.15	176.52	50
chloroform	$CHCl_3$	119.38	133.66	58
perchlorate	ClO_4^-	99.45	—	50
tetrafluoroborate	BF_4^-	86.80	—	42
orthotriflate	$CF_3SO_3^-$	149.07	—	73

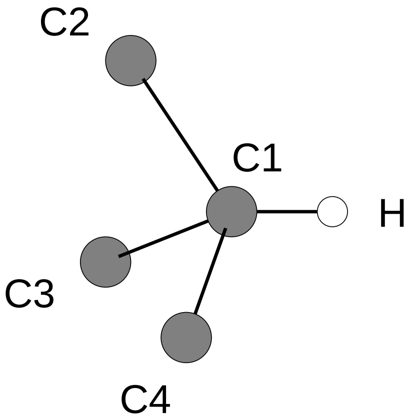
Hydrogen atoms (1)

An electron density maximum corresponding to an X—H hydrogen atom does not coincide with the position of the nucleus.



X—H bond lengths derived from crystal structure co-ordinates are therefore smaller than those derived by other techniques.

Hydrogen peaks often fall in the noise level of $\rho(x, y, z)$. In such cases hydrogen atoms can often be introduced at calculated positions, e.g.:

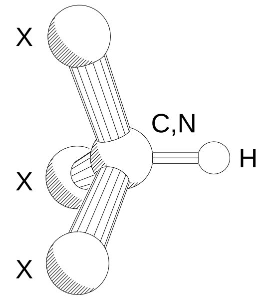


$$\vec{r}_{\mathrm{C1-H}} = \frac{\vec{r}_{\mathrm{C2-C1}} + \vec{r}_{\mathrm{C3-C1}} + \vec{r}_{\mathrm{C4-C1}}}{|\vec{r}_{\mathrm{C2-C1}} + \vec{r}_{\mathrm{C3-C1}} + \vec{r}_{\mathrm{C4-C1}}|} d^{\mathrm{lit}}_{\mathrm{C-H}}$$

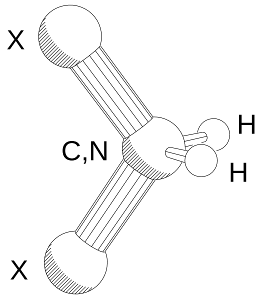
The length of the bond X—H is determined by the nature of X and the measurement temperature.

Hydrogen atoms (2)

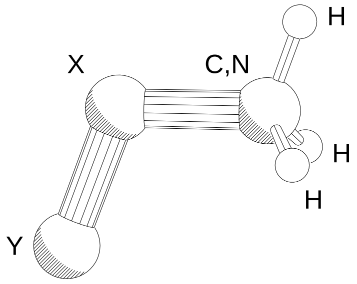
The following types of hydrogen atoms can be introduced at calculated positions (“fixed”) with the popular refinement program SHELXL:



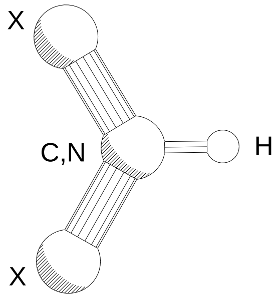
$X_3C(sp^3)H$
$X_3N^+(sp^3)H$



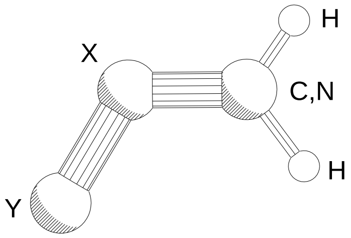
$X_2C(sp^3)H_2$
$X_2N^+(sp^3)H_2$



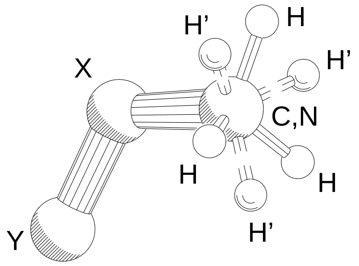
$Y{-}X{-}C(sp^3)H_3$
$Y{-}X{-}N^+(sp^3)H_3$
(staggered to shortest XY)



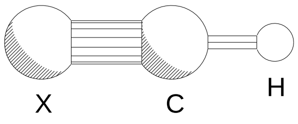
$X_2C(sp^2)H$
$X_2N^+(sp^2)H$
$X_2N(sp^2)H$



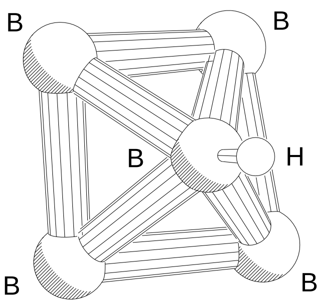
$Y{-}X{=}C(sp^2)H_2$
$Y{-}X{=}N^+(sp^2)H_2$
$Y{-}X{-}N(sp^2)H_2$



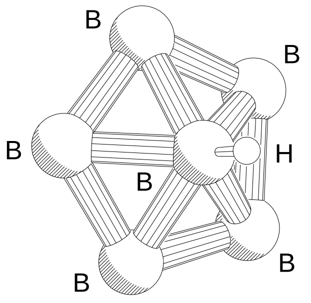
$Y{-}X{-}C(sp^3)H_3$
$Y{-}X{-}N^+(sp^3)H_3$
(ideally disordered)



$X{\equiv}C(sp)H$



B_4BH



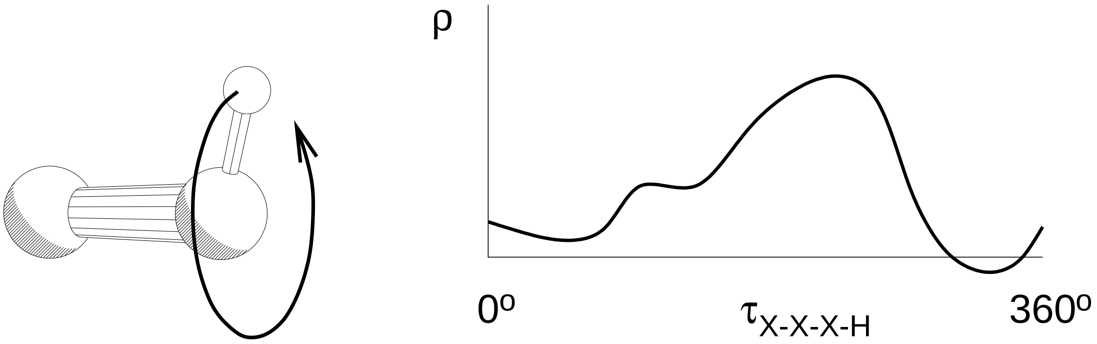
B_5BH

N.B. The hydrogens of O$-$H, sp^3 NH_2, sp^3 NH and $C{\equiv}C{-}CH_3$ can not be unambiguously fixed.

Hydrogen atom refinement strategies

Due to their low scattering power, H atoms can not always be refined. Several levels of approximation can be used:

- completely free: x, y, z, U_{iso} refined (anisotropic displacement parameters are not used for hydrogen atoms in normal structure determinations)
- displacement constrained: x, y, z refined, U_{iso} linked to U_{eq} of the carrier atom by a constant factor (default values: 1.5 for CH_3, OH; 1.2 for other types), or U_{iso} fixed at a standard value.
- idealized angles $\phi_{\text{Y}-\text{X}-\text{H}}$ and $\phi_{\text{H}-\text{X}-\text{H}}$; $d_{\text{X}-\text{H}}$ is refined
- CH_3 and OH: introduced as rigid group with ideal geometry ($d_{\text{X}-\text{H}}$, $\phi_{\text{Y}-\text{X}-\text{H}}$ and $\phi_{\text{H}-\text{X}-\text{H}}$). The torsion angle defining the conformation ($\tau_{\text{Z}-\text{Y}-\text{X}-\text{H}}$) is included as parameter in the refinement. The starting value of $\tau_{\text{Z}-\text{Y}-\text{X}-\text{H}}$ can be obtained from a Fourier map.



- idealized angles $\phi_{\text{X}-\text{X}-\text{H}}$ and $\phi_{\text{H}-\text{X}-\text{H}}$ and idealized bond length $d_{\text{X}-\text{H}}$ in riding mode refinement ($\Delta x_{\text{H}} = \Delta x_{\text{carrier}}$, etc.)

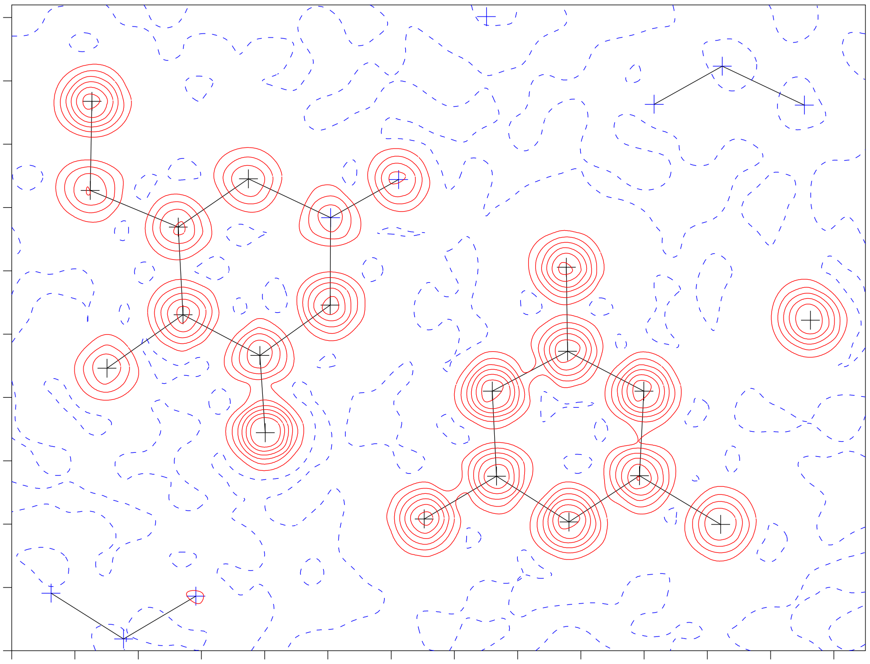
Often a mix of the mentioned strategies is applied. The "interesting" hydrogens (e.g. those donating hydrogen bonds) are freely refined, while the other hydrogen atoms are fixed.

The procedure of choice depends strongly on the quality of the diffraction data and on the type of compound.

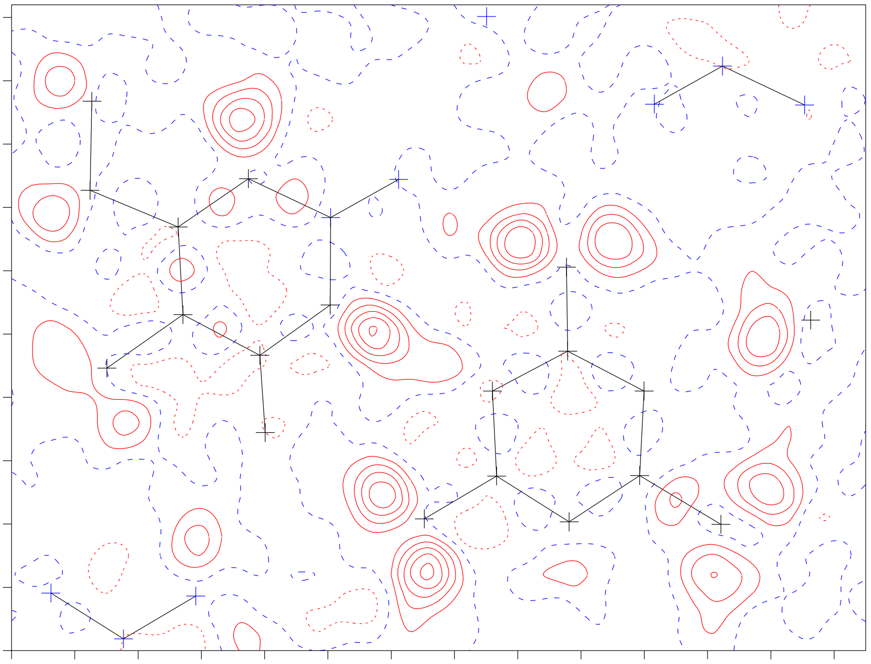
Residual density

A useful tool to determine the position of missing atoms is the difference Fourier function, also called residual density function.

$$\Delta\rho(x,y,z) = \frac{1}{V}\sum_{h_1,h_2,h_3} (|F^{\text{obs}}_{h_1h_2h_3}| - |F^{\text{calc}}_{h_1h_2h_3}|)e^{\phi^{\text{calc}}_{h_1h_2h_3}}e^{-2\pi i(h_1x+h_2y+h_3z)}$$



$\rho(x,y,z)$ for a model without hydrogen atoms (contour interval 2 e Å^{-3})



$\Delta\rho(x,y,z)$ for the same model (contour interval 0.15 e Å^{-3})

Figures of merit

- residual

$$wR2 = \left(\frac{\sum\limits_{h_1h_2h_3} w(|F^{\mathsf{obs}}_{h_1h_2h_3}|^2 - c^{-1}|F^{\mathsf{calc}}_{h_1h_2h_3}|^2)^2}{\sum\limits_{h_1h_2h_3} w(|F^{\mathsf{obs}}_{h_1h_2h_3}|^2)^2} \right)^{1/2}$$

- "classic" residual

$$R1 = \frac{\sum\limits_{h_1h_2h_3} ||F^{\mathsf{obs}}_{h_1h_2h_3}| - c^{-\frac{1}{2}}|F^{\mathsf{calc}}_{h_1h_2h_3}||}{\sum\limits_{h_1h_2h_3} |F^{\mathsf{obs}}_{h_1h_2h_3}|}$$

  summation over all $h_1h_2h_3$ for which

  $I_{h_1h_2h_3} > t\sigma(I_{h_1h_2h_3})$ ($t = 2$ or $t = 2.5$ in most cases).

- goodness of fit (GoF, in SHELXL: GooF)

$$S = \left(\frac{\sum\limits_{h_1h_2h_3} w(|F^{\mathsf{obs}}_{h_1h_2h_3}|^2 - c^{-1}|F^{\mathsf{calc}}_{h_1h_2h_3}|^2)^2}{n - p} \right)^{1/2}$$

  $n =$ number of reflections

  $p =$ number of parameters

- residual density extrema

$$\Delta\rho(x, y, z) = \frac{1}{V} \sum_{h_1h_2h_3} (|F^{\mathsf{obs}}_{h_1h_2h_3}| - |F^{\mathsf{calc}}_{h_1h_2h_3}|) e^{i\phi^{\mathsf{calc}}_{h_1h_2h_3}} e^{2\pi i(h_1x + h_2y + h_3z)}$$

- internal geometric consistency
- external geometric consistancy (crystallographic data bases)
- displacement parameters and their anistropicity

Potential problems during crystal structure determinations

synthesis	• PM
crystallization	• no crystals • insufficient quality of crystals (size, twinning)
data collection	• no diffraction (or too low $\theta^{\text{obs}}_{\text{max}}$) • insuficient reflection quality (split reflections, broad reflection profiles)
structure solution	• uninterpretable electron density (twin)

⇒ preliminary result

refinement	• unstable refinement • uninterpretable electron density in part of the cell (disorder) • unreasonable geometry

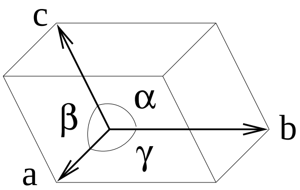
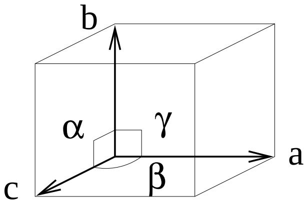
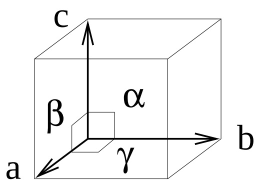
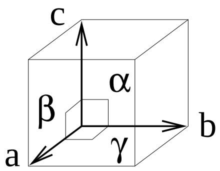
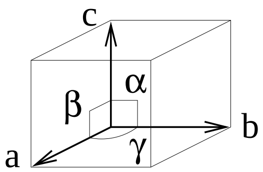
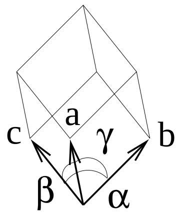
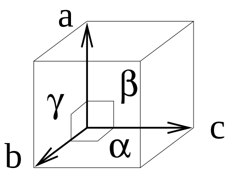
⇒ final result

interpretation	• insufficient precision • insufficient resolution • model interpretation problems

Unexpected crystal structures — some explanations

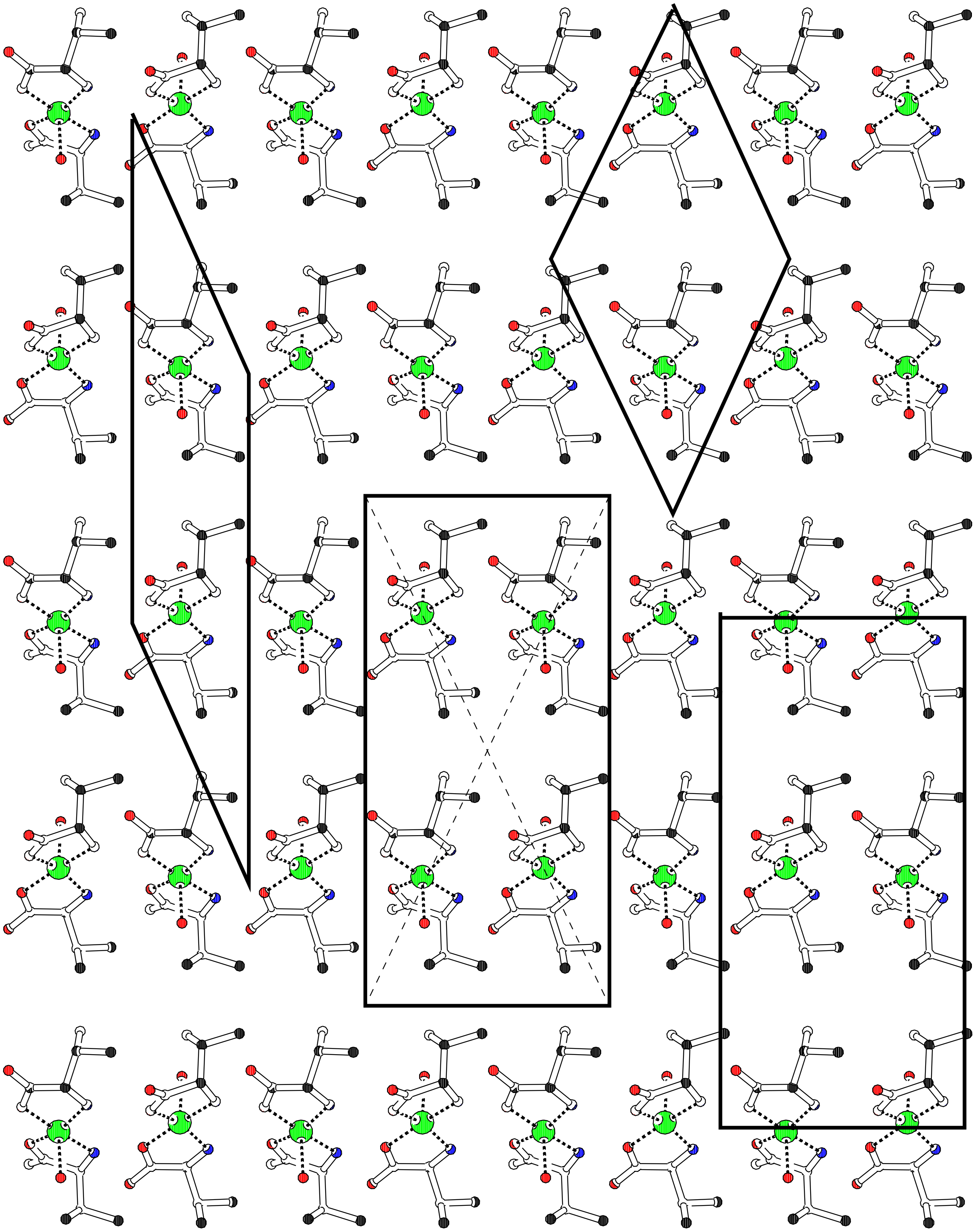
- There can always be differences between the conformation and/or configuration found in the crystal (solid phase) and that found in the liquid or gas phase.
- A crystal structure is not necessarily the minimum energy conformation of a compound, but it is always a low energy conformation (i.e. available at the temperature of crystallization).
- The substance crystallizing from a solution is not necessarily the main component.
- The measured crystal can differ from all the other crystals in a certain batch (different compound, co-crystallized solvent, polymorph, . . .).

Crystal systems

unit cell	name	geometry	symmetry
	triclinic (anorthic)	$a \neq b \neq c \neq a$ $\alpha \neq \beta \neq \gamma \neq \alpha$	1
	monoclinic	$a \neq b \neq c \neq a$ $\alpha = \gamma = 90^\circ$ $\beta \neq 90^\circ$	2
	orthorhombic	$a \neq b \neq c \neq a$ $\alpha = \beta = \gamma = 90^\circ$	222
	tetragonal	$a = b \neq c$ $\alpha = \beta = \gamma = 90^\circ$	4
	trigonal hexagonal	$a = b \neq c$ $\alpha = \beta = 90^\circ$ $\gamma = 120^\circ$	3 6
	rhombohedral	$a = b = c$ $\alpha = \beta = \gamma \neq 90^\circ$	3 on one body diagonal
	cubic	$a = b = c$ $\alpha = \beta = \gamma = 90^\circ$	3 on all body diagonals

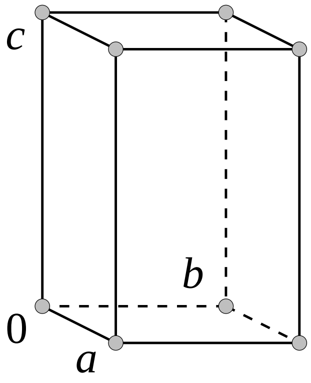
Centering of the unit cell (1)

Sometimes a non-primitive unit cell is preferable for easy description of the crystal.



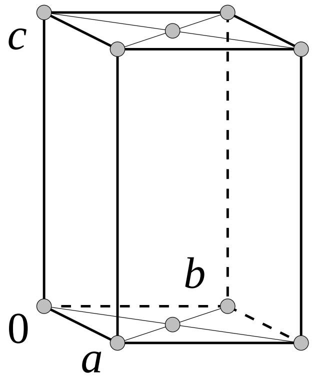
Centering of the unit cell (2)

Types of centering for 3-dimensional unit cells



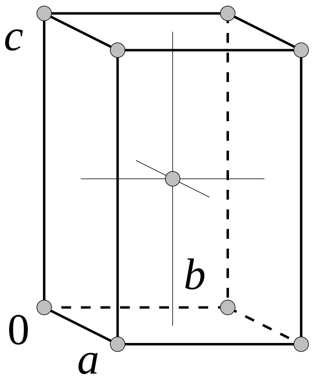
P

primitive



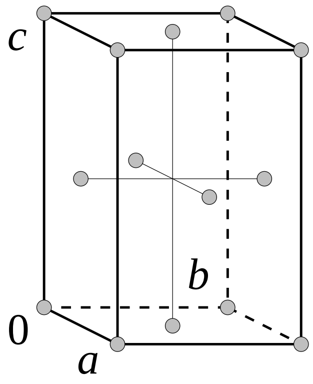
C

base centered (A, B also possible)



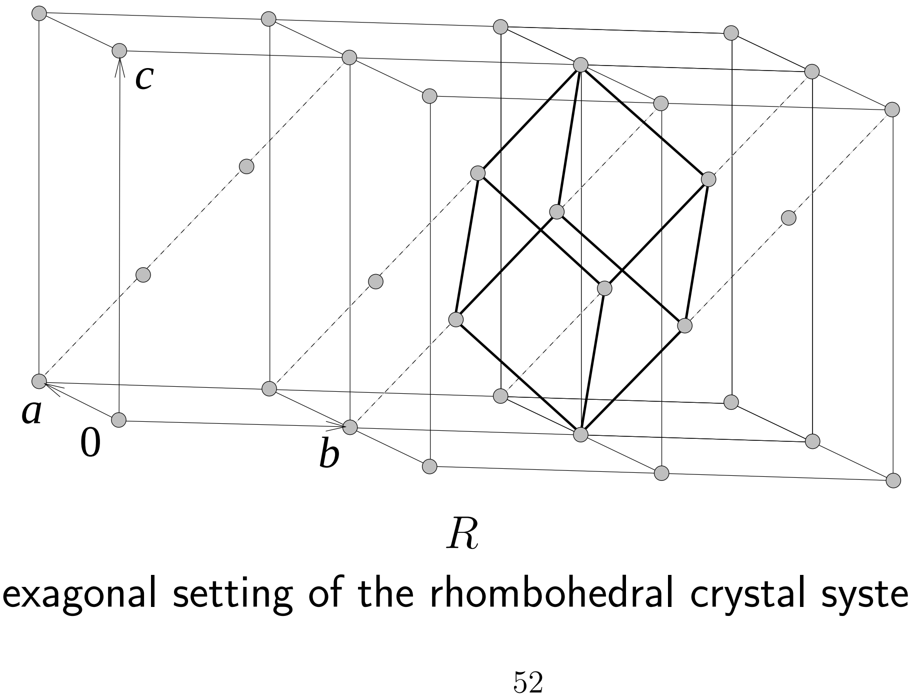
I

body centered



F

face centered

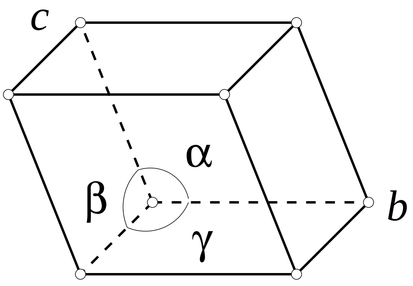


R

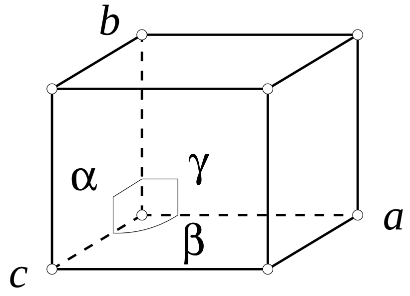
hexagonal setting of the rhombohedral crystal system

Bravais lattices

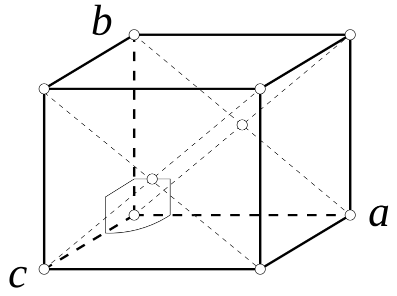
There are fourteen combinations of crystal systems and centering types.



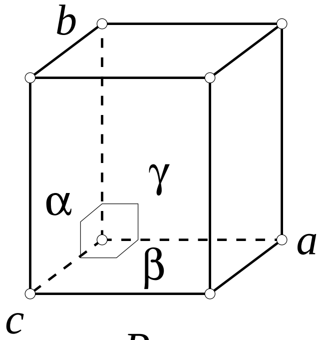
aP



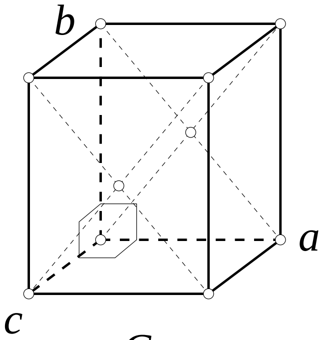
mP



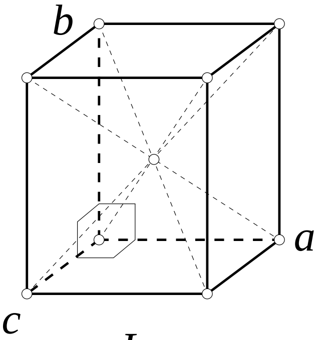
mC



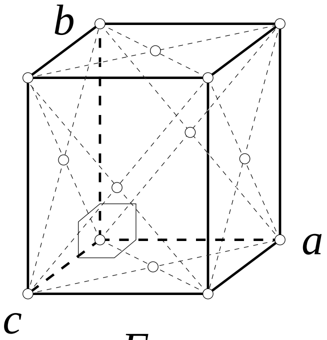
oP



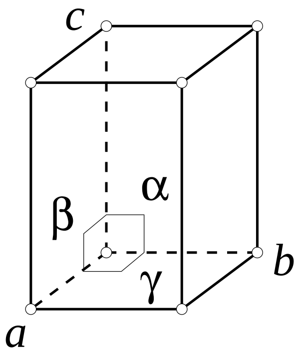
oC



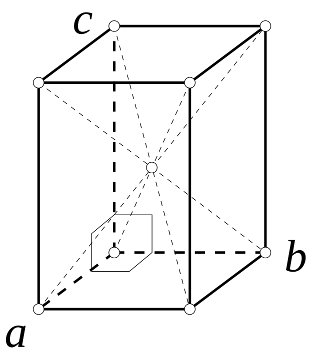
oI



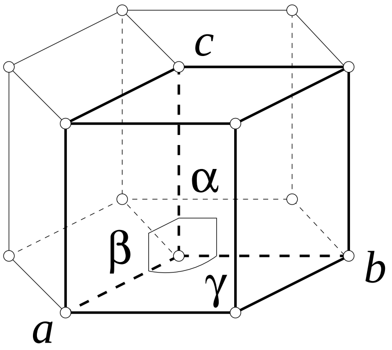
oF



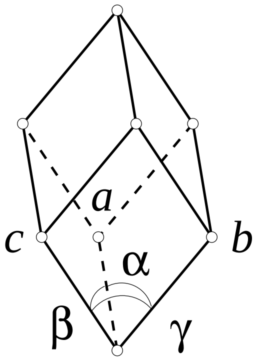
tP



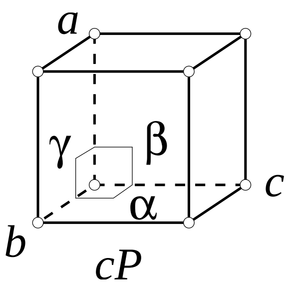
tI



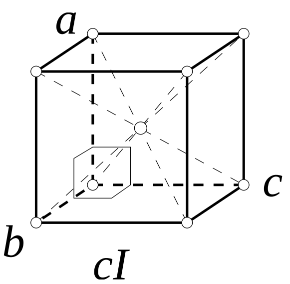
hP



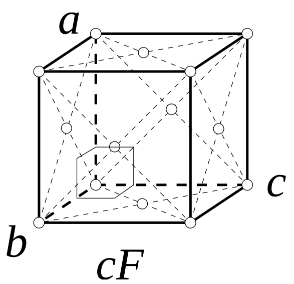
rP=R



cP



cI

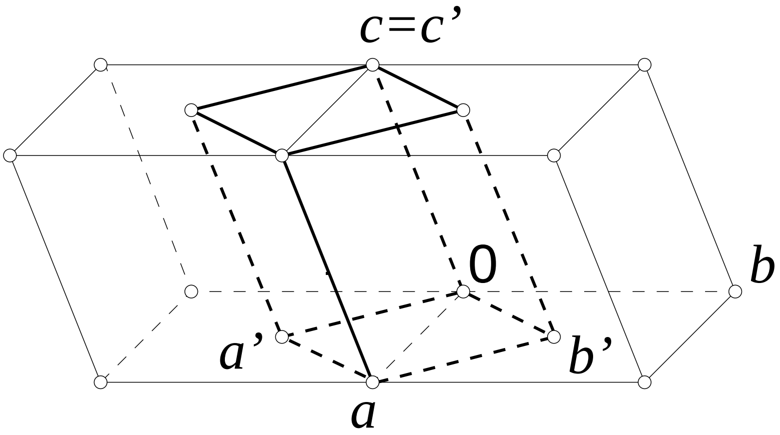


cF

Key: a = triclinic (anorthic), m = monoclinc, o = orthorhombic, t = tetragonal, h = hexagonal (trigonal), r = rhombohedral, c = cubic, P = primitive, C = base centered, I = body centered, F = face centered.

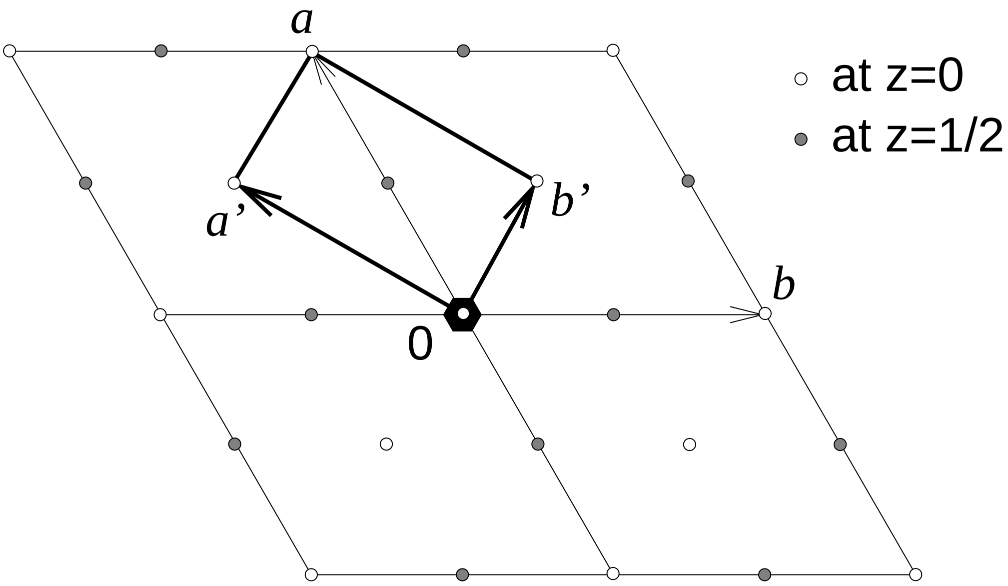
Invalid combinations of centering and crystal system

Triclinic C



The primitive cell $(\vec{a'}, \vec{b'}, \vec{c'})$ is triclinic; centering does not give a crystal system of higher symmetry. The same holds for triclinic I and triclinic F.

Hexagonal F

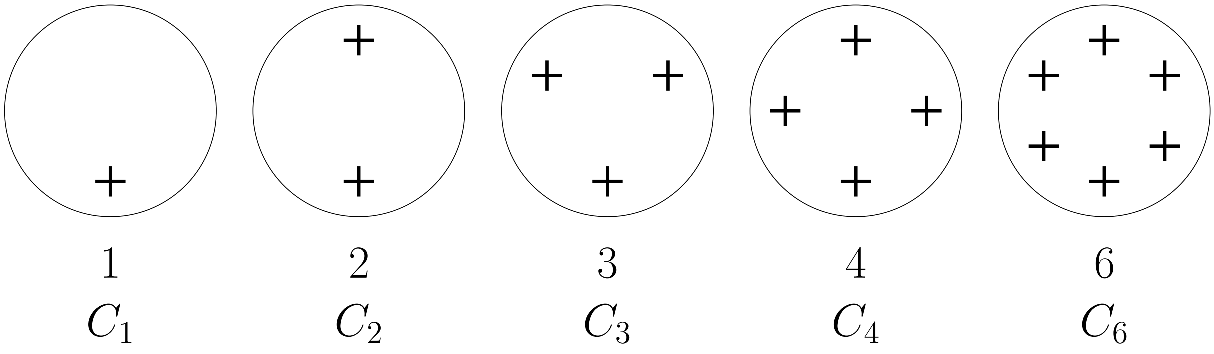


An F centered hexagonal lattice loses the sixfold symmetry of the $(\vec{a}, \vec{b})$ plane. Since the diagonals of a diamond are perpendicular, an orthorhombic I lattice can be formed from $(\vec{a'}, \vec{b'}, \vec{c})$.

Rotation symmetry in the crystal lattice (1)

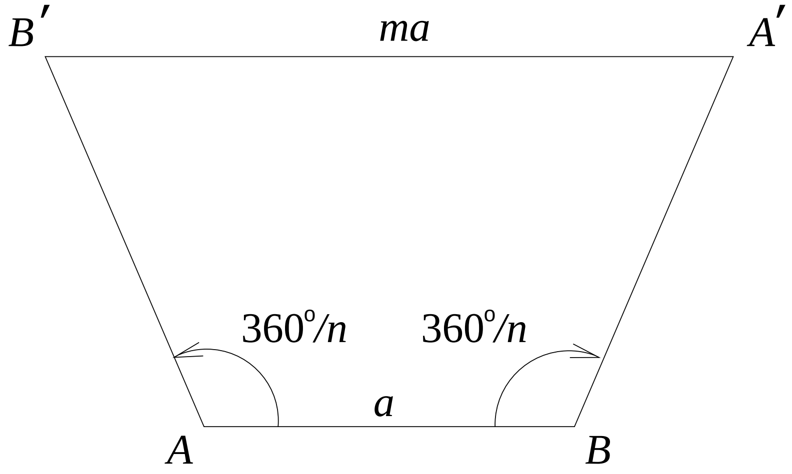
Only a limited number of symmetry elements can be embedded in a translation lattice in such a way that the operations are valid for the whole lattice and its contents.

Proper rotations



n: Hermann Mauguin notation; C_n: Schoenflies notation

Other rotations can not be combined with translation symmetry.



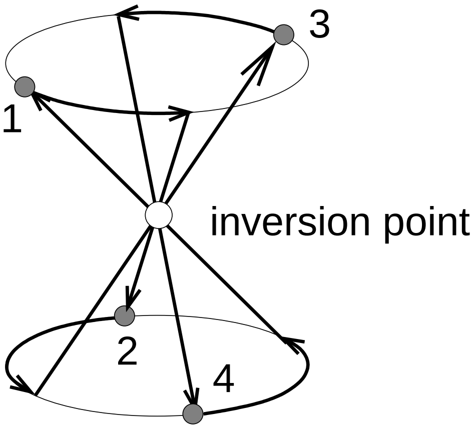
n	m
1	1
2	3
3	2
4	1
5	0.38
6	0
7	0.25
8	0.41

A' is the image of A, rotated 360°$/n$ around B
B' is the image of B, rotated $-$360°$/n$ around A
$AB \parallel A'B'$ therefore $d(A'B')$ should be equal to ma, with $m \in \mathbb{N}$.

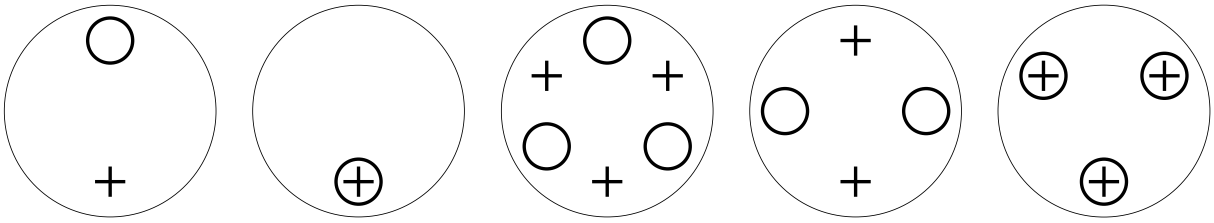
Rotation symmetry in the crystal lattice (2)

Improper rotations

A combination of a rotation operation with an inversion (used in crystallography) or a mirror operation (used in spectroscopy) into one new operation is called an improper rotation, e.g. the four-fold improper rotation:



In crystallography there are five improper rotations:



$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{4}$	$\bar{6}$
S_2	S_1	S_6	S_4	S_3

($+$ lies above the plane of projection, ○ below)

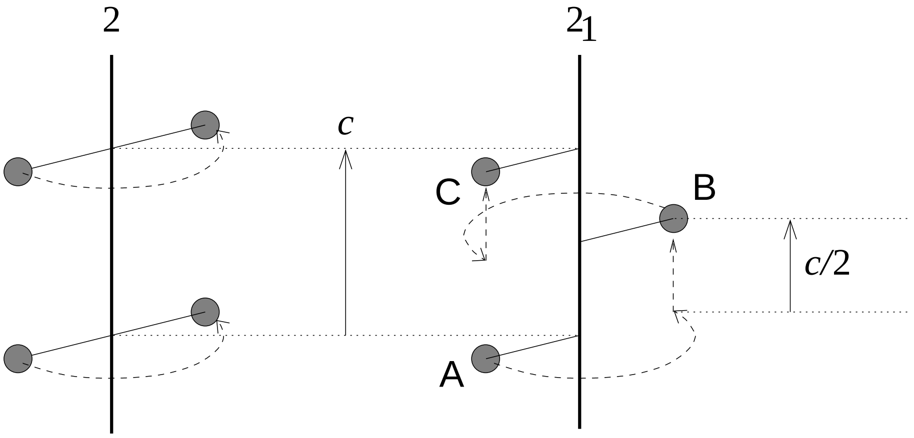
$\bar{n}$: Hermann Mauguin notation; S_n: Schoenflies notation

$\bar{1}$ is the inversion centre (inversion point), in the Schoenflies system usually denoted as C_i.
$\bar{2}$ is the mirror operation, in the Hermann Mauguin system usually denoted as m.

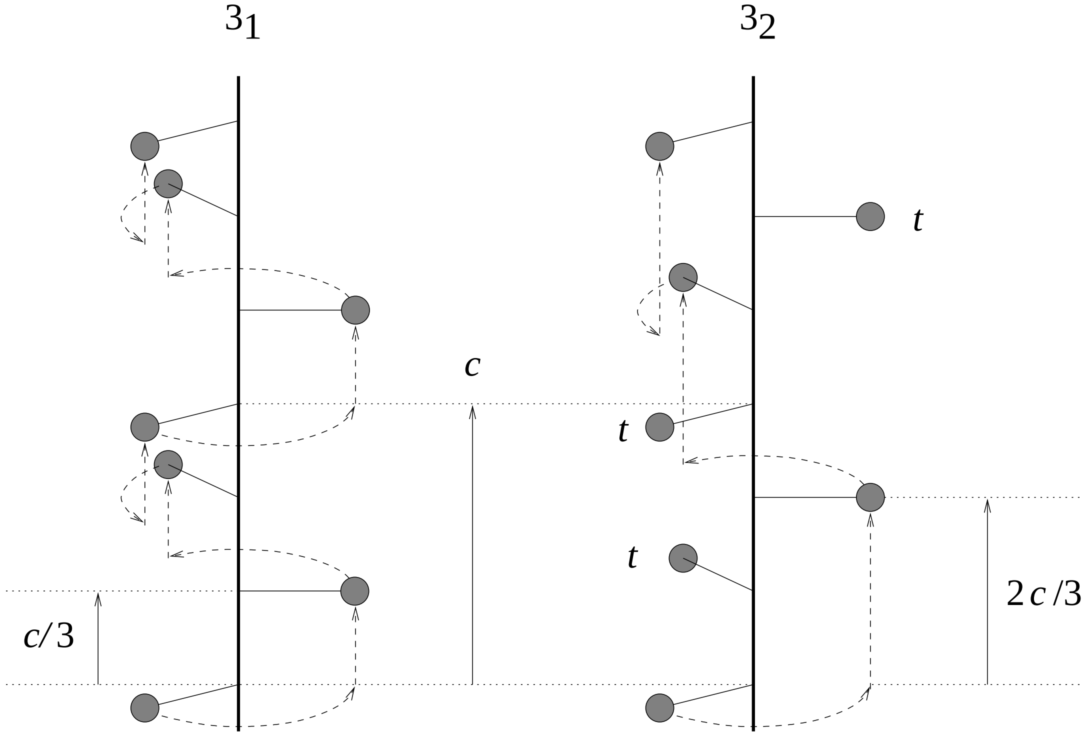
Screw axes

A screw axis combines rotation with fractional translation in one single new operation. The screw axis n_m performs a $360°/n$ rotation around the axis and a translation over m/n along the axis.

2 axis and 2_1 axis $\parallel \vec{a}$:



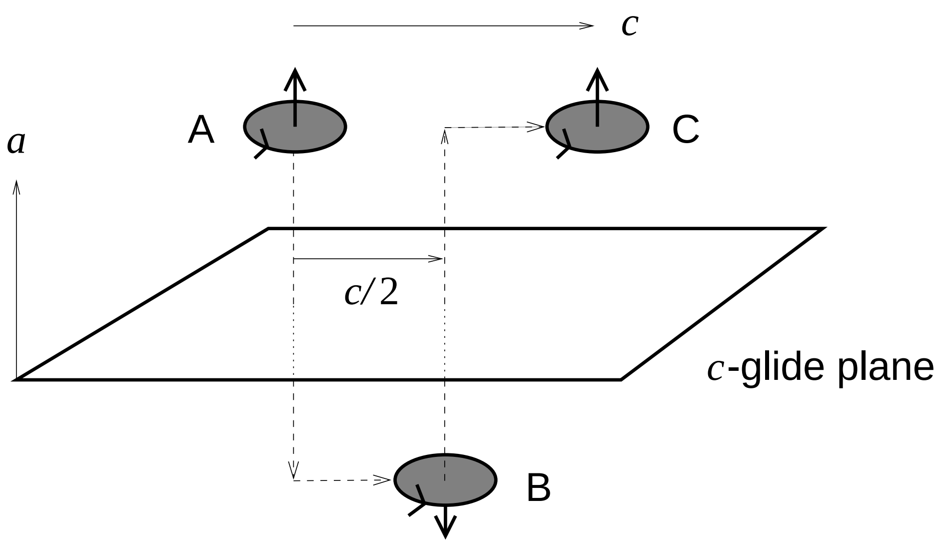
3_1 and 3_2 axes $\parallel \vec{a}$ are righthanded and left-handed screw axes, respectively (i.e. related by a mirror operation):



(t indicates a position obtained by unit cell translations)

Glide planes

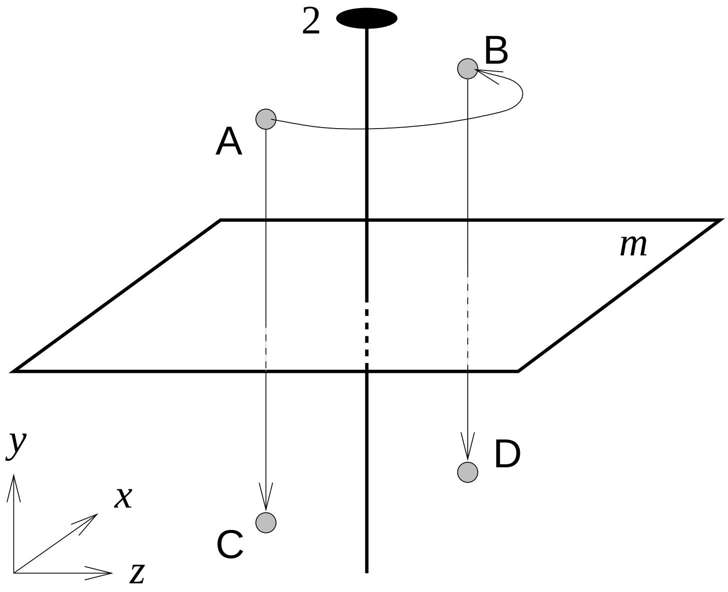
Glide planes combine a mirror operation with a fractional translation parallel to the mirror plane into one new combination. The symbol used indicates the translation direction.



glide plane	translation component	orientation of the glide plane
a	$\frac{1}{2}\vec{a}$	$\perp \vec{b}$ or $\perp \vec{c}$
b	$\frac{1}{2}\vec{b}$	$\perp \vec{a}$ or $\perp \vec{c}$
c	$\frac{1}{2}\vec{c}$	$\perp \vec{a}$ or $\perp \vec{b}$
n	$\frac{1}{2}(\vec{b}+\vec{c})$	$\perp \vec{a}$
	$\frac{1}{2}(\vec{a}+\vec{c})$	$\perp \vec{b}$
	$\frac{1}{2}(\vec{a}+\vec{b})$	$\perp \vec{c}$
d	$\frac{1}{4}(\vec{b}+\vec{c})$	$\perp \vec{a}$
	$\frac{1}{4}(\vec{b}-\vec{c})$	$\perp \vec{a}$
	$\frac{1}{4}(\vec{a}+\vec{b})$	$\perp \vec{b}$
	(etc.)	(etc.)
	$\frac{1}{4}(\vec{a}+\vec{b}+\vec{c})$	$\perp (\vec{a}+\vec{b}+\vec{c})$
	$\frac{1}{4}(\vec{a}+\vec{b}+\vec{c})$	$\perp (\vec{a}+\vec{b})$
	(etc.)	(etc.)

Combination of symmetry elements (1)

Consider a two-fold rotation axis (2) perpendicular to a mirror plane (m).

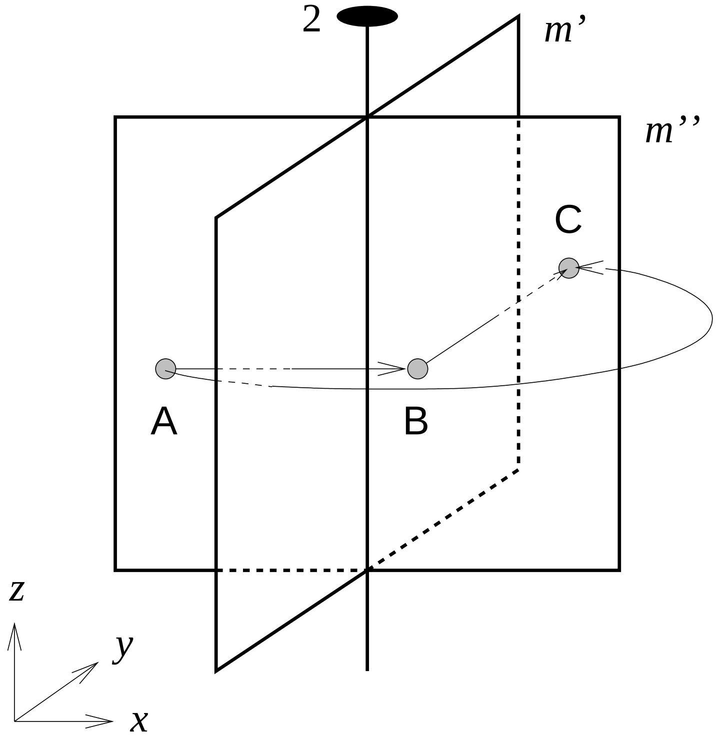


$$\mathsf{A}(x,y,z) \xrightarrow{2} \mathsf{B}(\bar{x},y,\bar{z})$$
$$\mathsf{A}(x,y,z) \xrightarrow{m} \mathsf{C}(x,\bar{y},z)$$
$$\mathsf{B}(\bar{x},y,\bar{z}) \xrightarrow{m} \mathsf{D}(\bar{x},\bar{y},\bar{z})$$
$$\mathsf{C}(x,\bar{y},z) \xrightarrow{2} \mathsf{D}(\bar{x},\bar{y},\bar{z})$$

$$\mathsf{A}(x,y,z) \xrightarrow{?} \mathsf{D}(\bar{x},\bar{y},\bar{z})$$

$?$ is $\bar{1}$ on $(0,0,0)$

In a similar way two perpendicular mirror operations combine to a two-fold rotation.



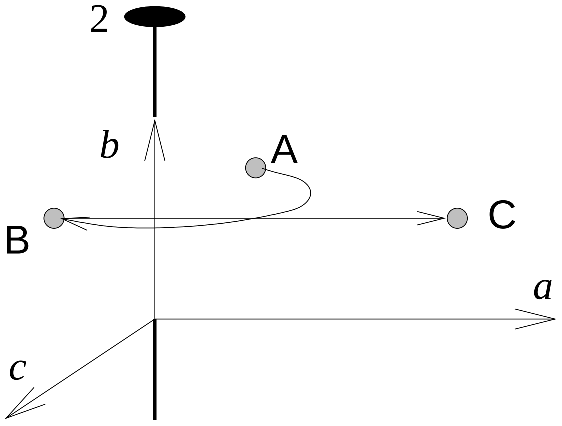
$$\mathsf{A}(x,y,z) \xrightarrow{m'} \mathsf{B}(\bar{x},y,z)$$
$$\mathsf{B}(\bar{x},y,z) \xrightarrow{m''} \mathsf{C}(\bar{x},\bar{y},z)$$

$$\mathsf{A}(x,y,z) \xrightarrow{?} \mathsf{C}(\bar{x},\bar{y},z)$$

$?$ is 2 on $(0,0,z)$

Combination of symmetry elements (2)

Consider a translation lattice with a two-fold rotation axis, parallel to $\vec{b}$.

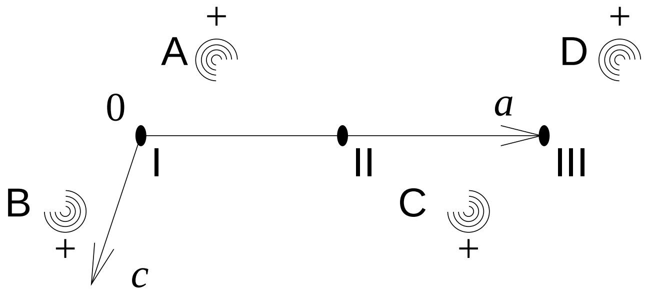


$$\mathrm{A}(x,y,z) \xrightarrow{2\|\vec{b}} \mathrm{B}(\bar{x},y,\bar{z})$$
$$\mathrm{B}(\bar{x},y,\bar{z}) \xrightarrow{\vec{a}} \mathrm{C}(\bar{x}+1,y,\bar{z})$$

$$\mathrm{A}(x,y,z) \xrightarrow{?} \mathrm{C}(\bar{x}+1,y,\bar{z})$$

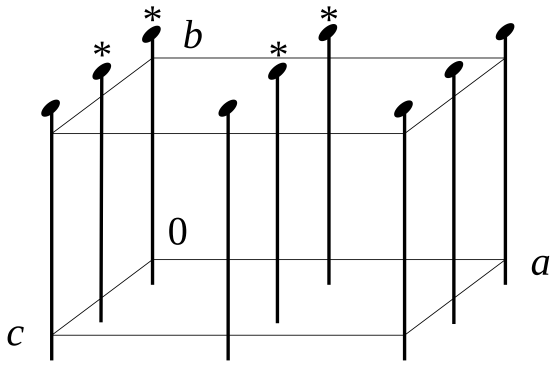
$?$ is $2 \parallel \vec{b}$ through $(\frac{1}{2},0,0)$

Projected down $\vec{b}$:



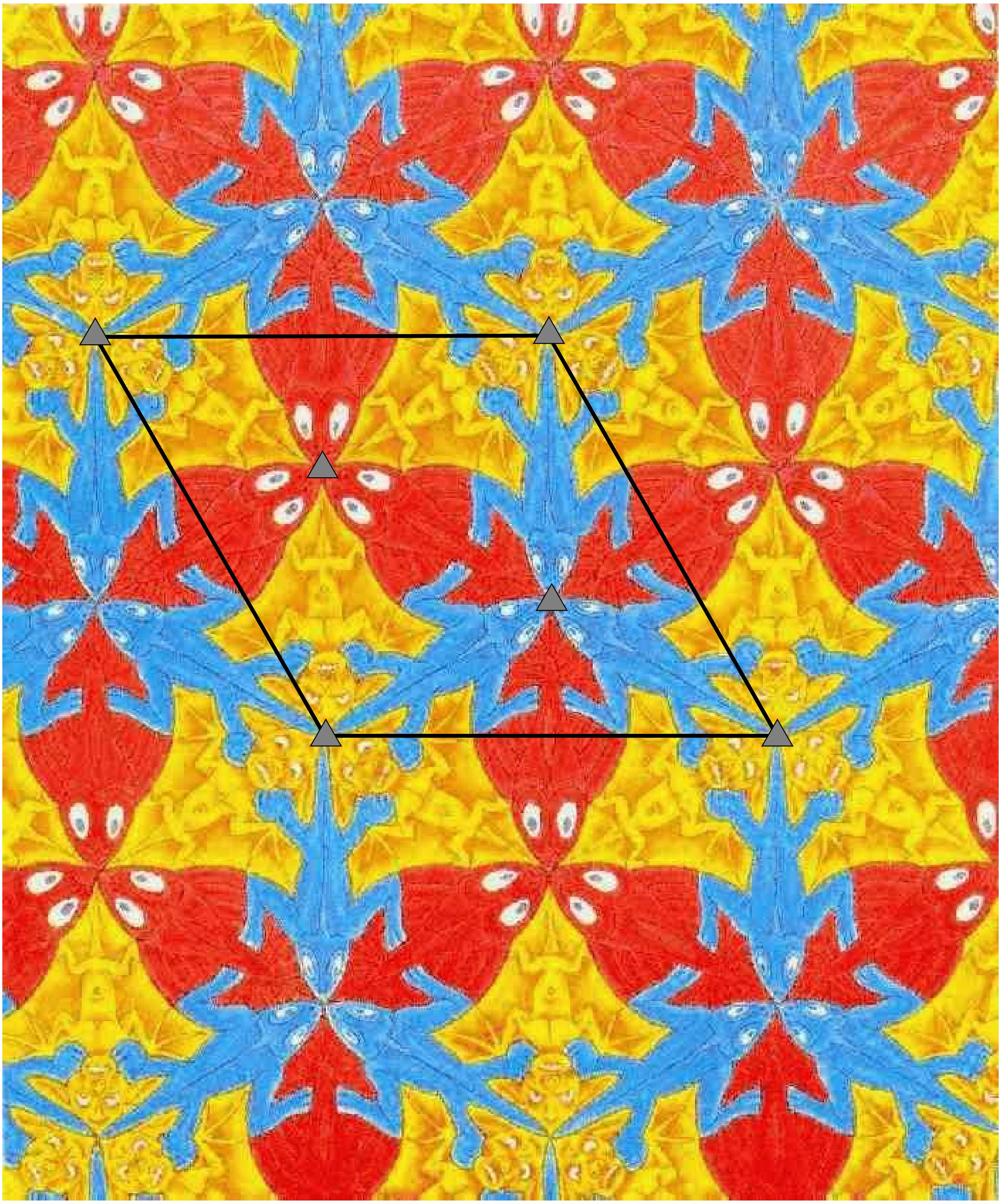
2^{III} is the translation image of 2^{I}, therefore the rotation axes I and III are identical. Rotation axis II is clearly distinguishable from I and II.

When other translations are applied, a total of four distinguishable two-fold rotation axes, marked by an *, are found:



Combination of symmetry elements (3)

The generation of extra symmetry elements can be seen in all periodic drawings, e.g. those of M.C. Escher:



Space group symbols

A space group is an infinite, but closed group of symmetry elements (proper and improper rotation axes, screw axes and glide planes) in a translation lattice (primitive or centered).

In contrast to point groups, the symmetry elements in a space group do not have one common point. There are 230 space groups.

A space group symbol consists of the Bravais symbol and up to three positions with (a combination of) symmetry elements. Each position indicates a different direction in the crystal lattice, depending on the crystal system.

system	positions			Examples
	1st	2nd	3rd	
triclinic	$(0,0,0)$	—	—	$P1$, $P\bar{1}$
monoclinic	$\vec{b}$	—	—	$P2$, $C2/c$
orthorhombic	$\vec{a}$	$\vec{b}$	$\vec{c}$	$P222_1$, $Pbca$, $Fmmm$
tetragonal	$\vec{c}$	$\vec{a}$, $\vec{b}$	$\vec{a}+\vec{b}$, $\vec{a}-\vec{b}$	$P42_12$, $I4$ $I4_1/a$
trigonal, hexagonal	$\vec{c}$	$\vec{a}$, $\vec{b}$, $\vec{a}+\vec{b}$	$\perp \vec{a}$, $\perp \vec{b}$, $\perp (\vec{a}+\vec{b})$	$P3_1$, $P6/mcc$, $P3_112$, $P3_121$
cubic	$\vec{a}$, $\vec{b}$, $\vec{c}$	$\vec{a}+\vec{b}+\vec{c}$, (etc.)	$\vec{a}+\vec{b}$, (etc.)	$Pm\bar{3}$, $I4_132$, $P23$, $F432$

The combination of a twofold rotation axis perpendicular to a mirror plane is denoted as $2/m$ ("two over em") in a single position.

An inversion point is only explicitly denoted in space group $P\bar{1}$.

Space group frequency

After the July 2004 update, the Cambridge Structural Data base contains 322 419 organic and organometallic crystal structures. Some space groups occur far more often than others:

	space group	# entries
	$P1$	3097
*	$P\bar{1}$	69805
	$P2$	74
*	$P2_1$	17904
	Pm	10
	Pc	1249
	$P2/m$	64
	$P2_1/m$	1935

	space group	# entries
	$P2/c$	1761
*	$P2_1/c$	112786
*	$C2/c$	24869
	$P222$	21
	$P222_1$	45
	$P2_12_12$	1422
*	$P2_12_12_1$	26655
	$Pbca$	11526

The 5 space groups (*i.e.* 2%) marked with a * contain 78% of the crystal structures reported in the CSD.

Screw axes and glide planes give in general a better (more dense) crystal packing than rotation axes and mirror planes.

Enantiopure compounds can not crystallize in space groups with an inversion operation ($\bar{n}$, glide planes).

Racemic compounds can crystallize in space groups without inversion operations.

Equivalent and special positions; asymmetric unit

Equivalent positions are the positions which can be derived from starting position (x, y, z) by successively applying all symmetry operations, except full unit cell translations. n is the number of equivalent positions.

An atom lying on a rotation axis, mirror plane or inversion centre coincides with its own image. These positions are called special positions.

- $P\bar{1}$, $n = 2$

equivalent	(x, y, z)	$(\bar{x}, \bar{y}, \bar{z})$
example	$(0.1, 0.4, 0.3)$	$(-0.1, -0.4, -0.3) \Rightarrow (0.9, 0.6, 0.7)$
special	$(0, 0, 0)$	$(0, 0, 0)$
	$(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$	$(-\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}) \Rightarrow (\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$
	(etc.)	

- $P222$, $n = 4$

equivalent	(x, y, z)	$(\bar{x}, \bar{y}, z)$	$(\bar{x}, y, \bar{z})$	$(x, \bar{y}, \bar{z})$
special	$(x, 0, 0)$	$(\bar{x}, 0, 0)$	$(\bar{x}, 0, 0)$	$(x, 0, 0)$
	$(0, y, 0)$	$(0, \bar{y}, 0)$	$(0, y, 0)$	$(0, \bar{y}, 0)$
	$(0, 0, z)$	$(0, 0, z)$	$(0, 0, \bar{z})$	$(0, 0, \bar{z})$

- $P2_1/c$, $n = 4$

equivalent	(x, y, z)	$(\bar{x}, \frac{1}{2} + y, \frac{1}{2} - z)$	$(\bar{x}, \bar{y}, \bar{z})$	$(x, \frac{1}{2} - y, \frac{1}{2} + z)$
special	$(0, 0, 0)$	$(0, \frac{1}{2}, \frac{1}{2})$	$(0, 0, 0)$	$(0, \frac{1}{2}, \frac{1}{2})$
	$(0, 0, \frac{1}{2})$	$(0, \frac{1}{2}, 0)$	$(0, 0, \frac{1}{2})$	$(0, \frac{1}{2}, 0)$
	(etc.)			

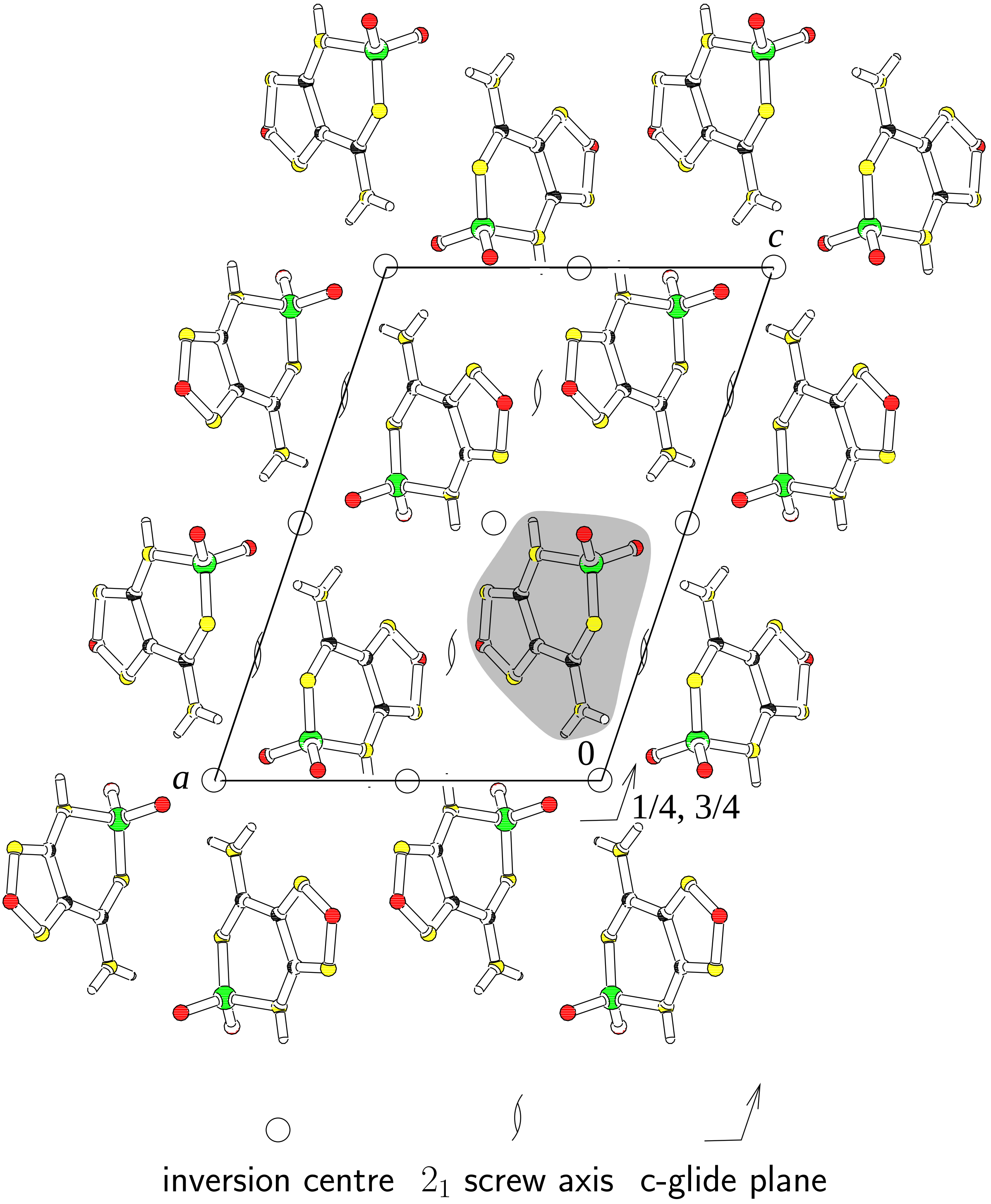
N.B. 2_1 and c do not pass through $(0, 0, 0)$, the position of $\bar{1}$.

The asymmetric unit is a section of the unit cell from which the complete unit cell can be generated by applying the space group symmetry operations.

If there are Z molecular units in the unit cell, the number of independent molecules is $Z/n = Z'$.

Independent molecules (1)

Projection down $\vec{b}$ of a structure in space group $P2_1/c$.

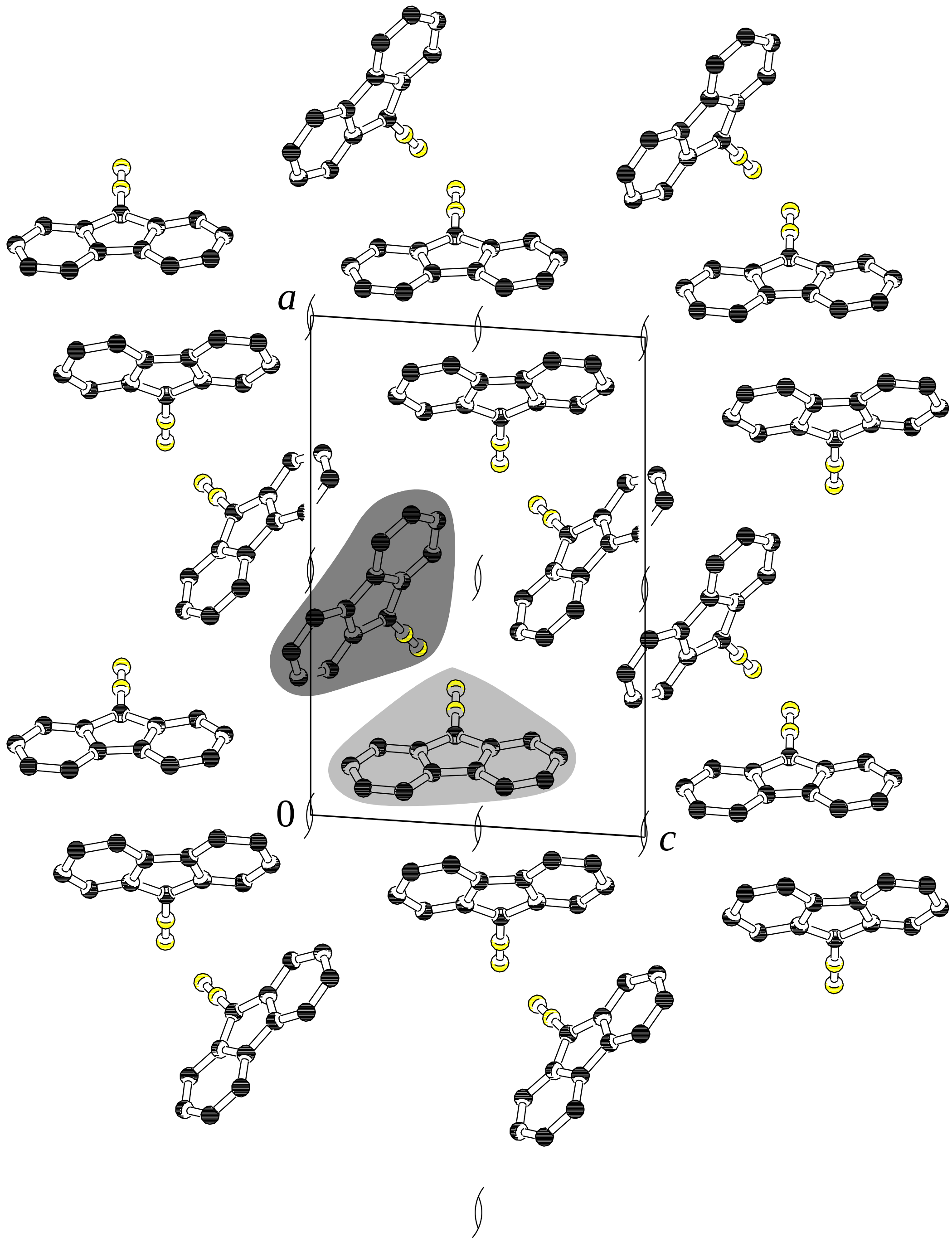


Space group multiplicity, number of molecules in the unit cell, and number of independent molecules (shaded in grey):

$$n = 4 \qquad Z = 4 \qquad Z' = Z/n = 1$$

Independent molecules (2)

Projection down $\vec{b}$ of a structure in space group $P2_1$.

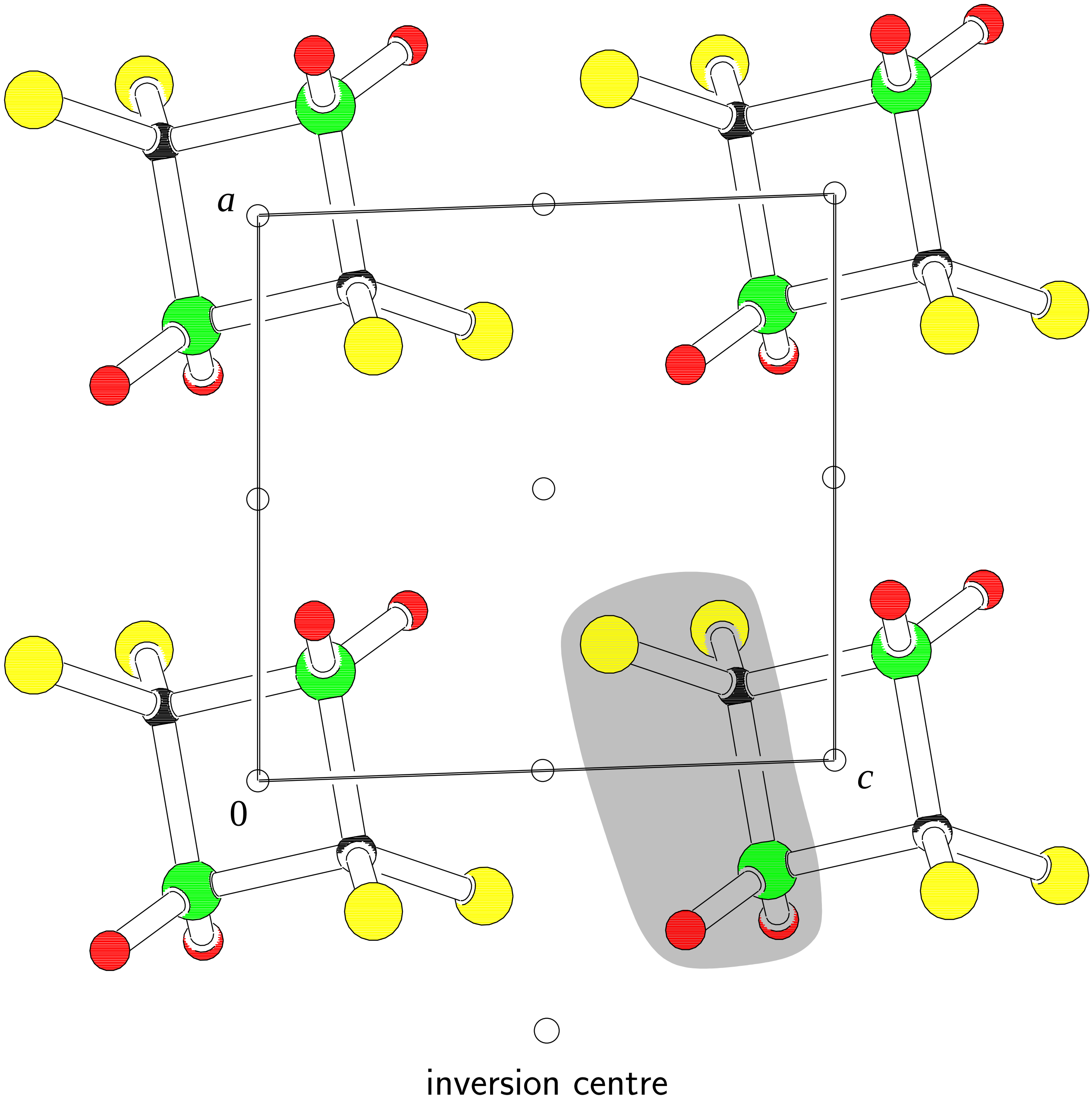


Space group multiplicity, number of molecules in the unit cell, and number of independent molecules (shaded in grey):

$$n = 2 \qquad Z = 4 \qquad Z' = Z/n = 2$$

Independent molecules (3)

Projection down $\vec{b}$ of a structure in space group $P\bar{1}$.

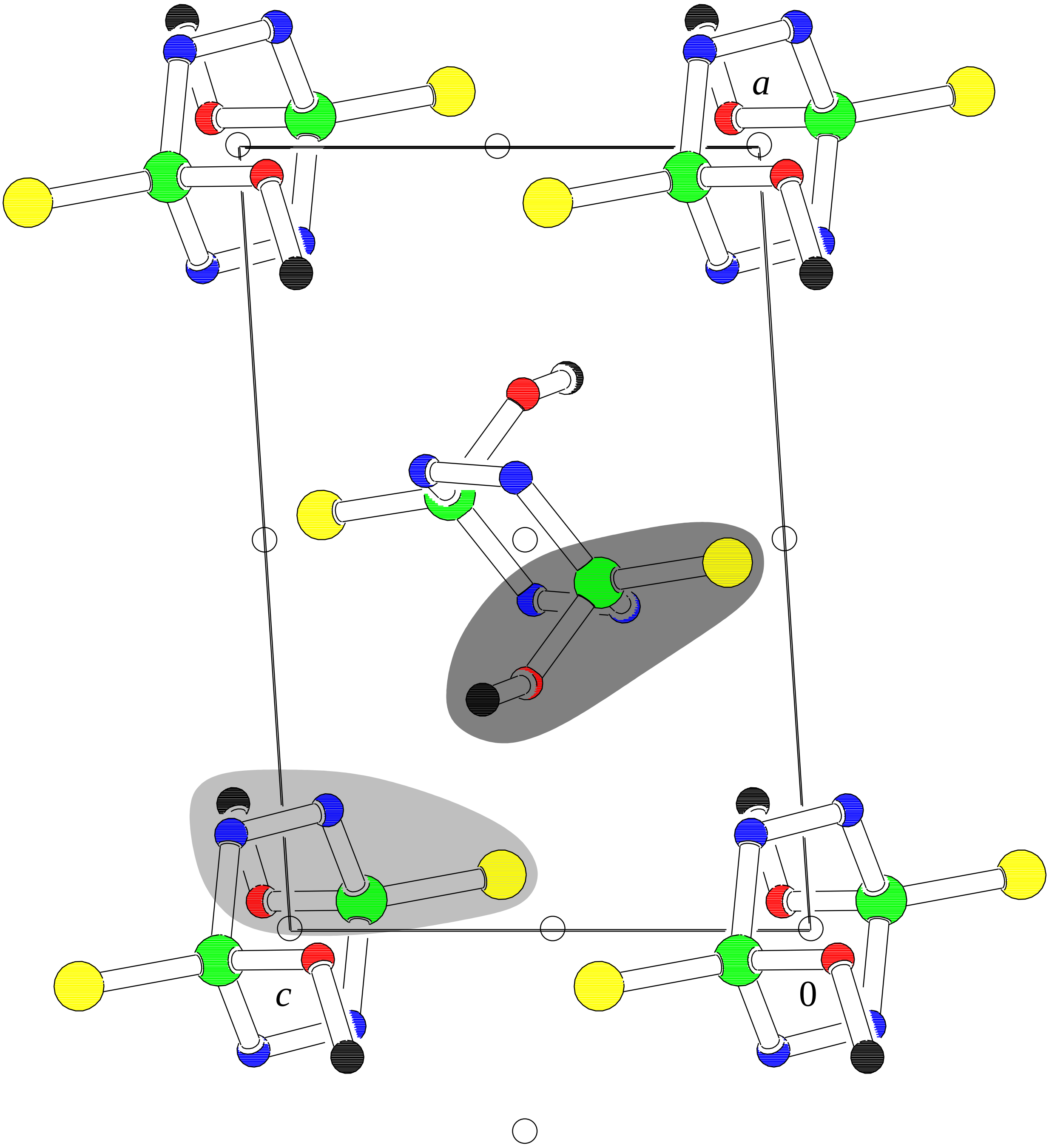


Space group multiplicity, number of molecules in the unit cell, and number of independent molecules (shaded in grey):

$$n = 2 \qquad Z = 1 \qquad Z' = Z/n = \frac{1}{2}$$

Independent molecules (4)

Projection down $\vec{b}$ of a structure in space group $P\bar{1}$.



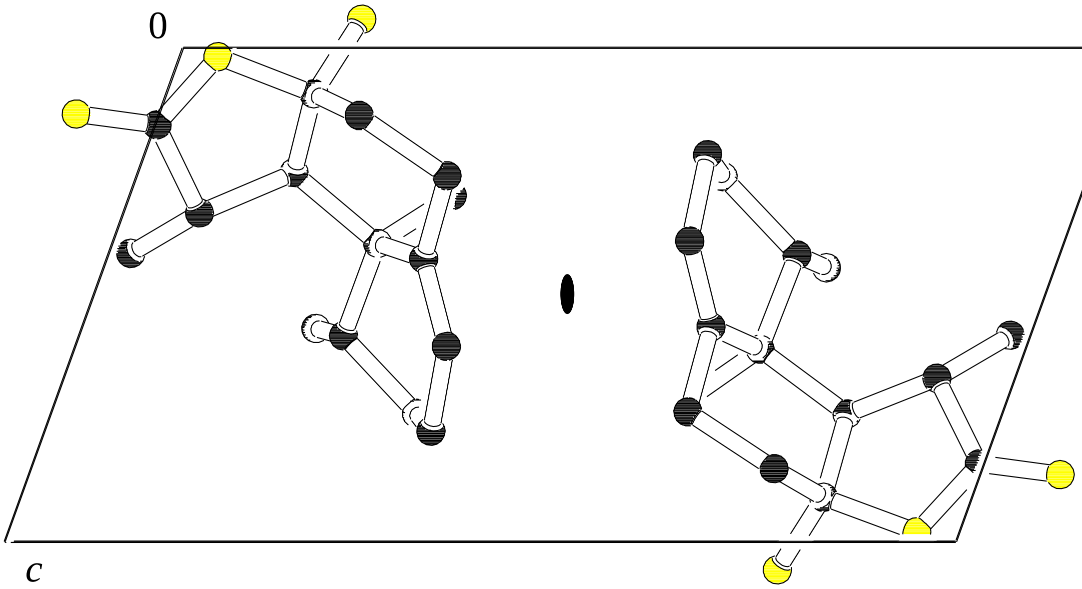
inversion centre

Space group multiplicity, number of molecules in the unit cell, and number of independent molecules (shaded in grey):

$$n = 2 \qquad Z = 2 \qquad Z' = Z/n = 1 = 2 \times \frac{1}{2}$$

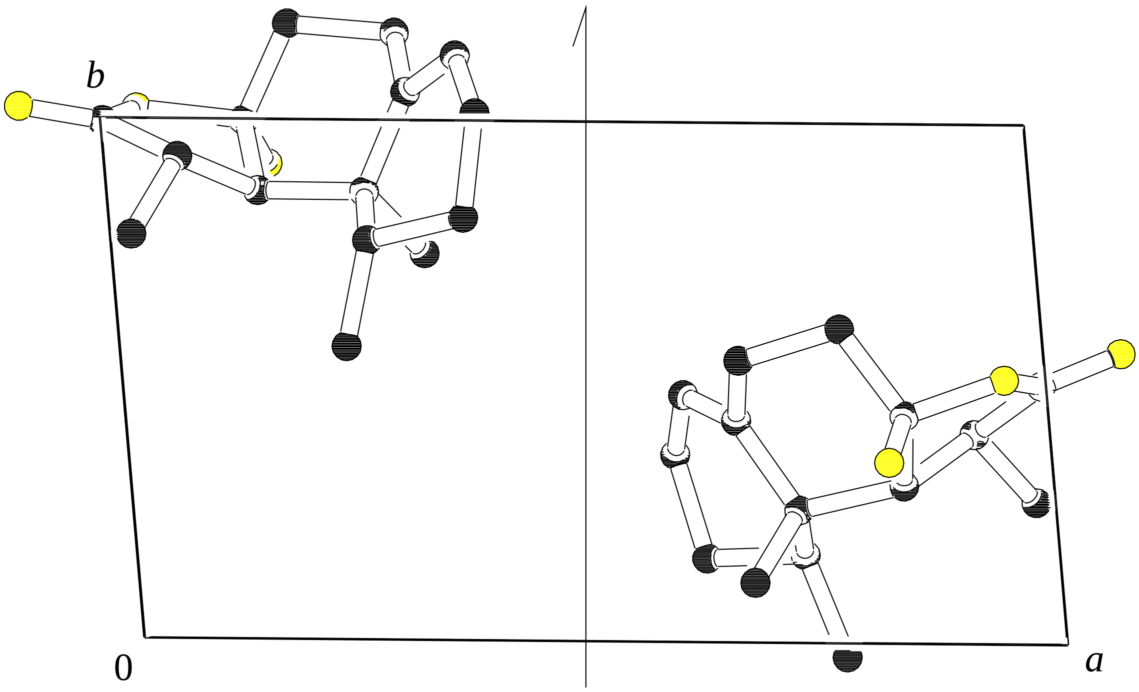
Local symmetry (1)

Projection down $\vec{b}$ of a structure in space group $P1$, with $Z = 1$, $n = 2$, $Z' = 2$.



A two-fold rotation axis appears to be present.

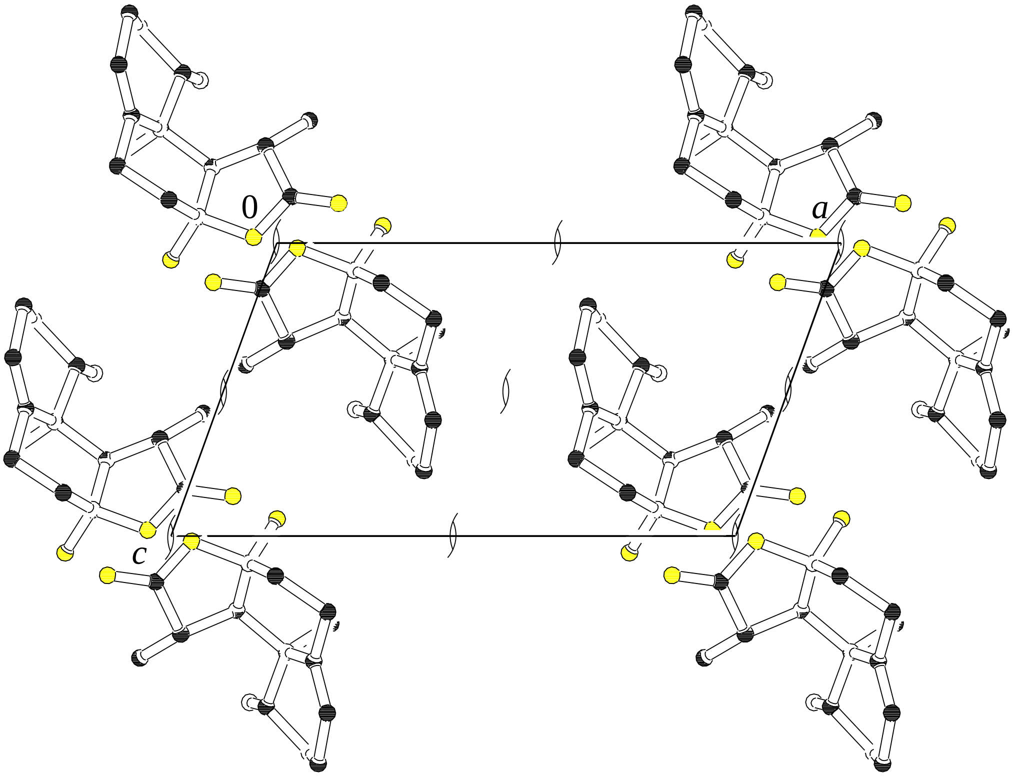
Projection down $\vec{c}$ of the same molecules.



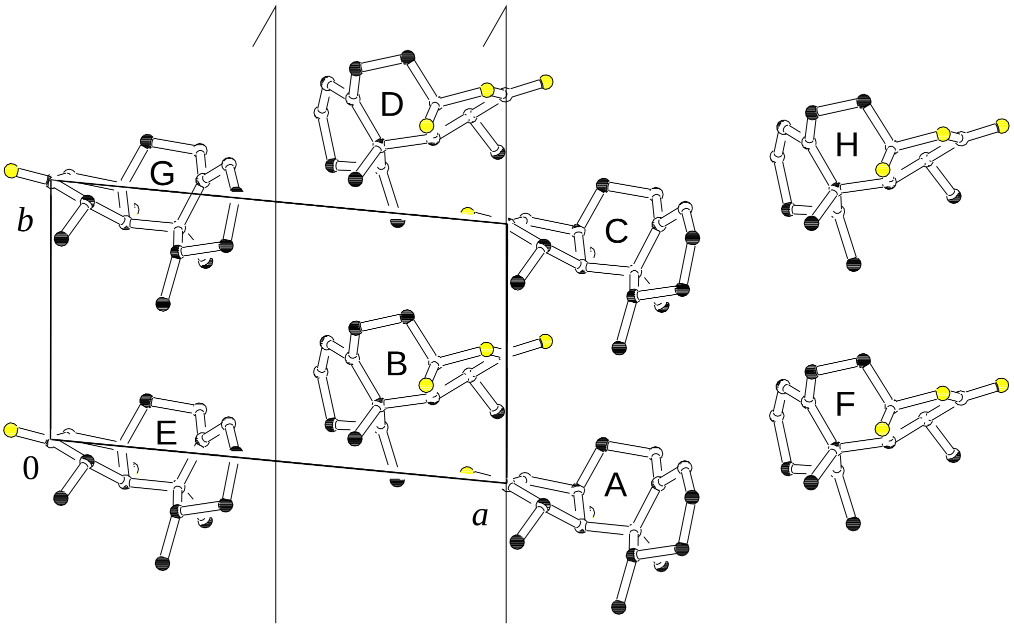
The axis appears to be a 2_1 screw axis.

Local symmetry (2)

The screw axis appears to be valid for the whole crystal when projected down $\vec{b}$

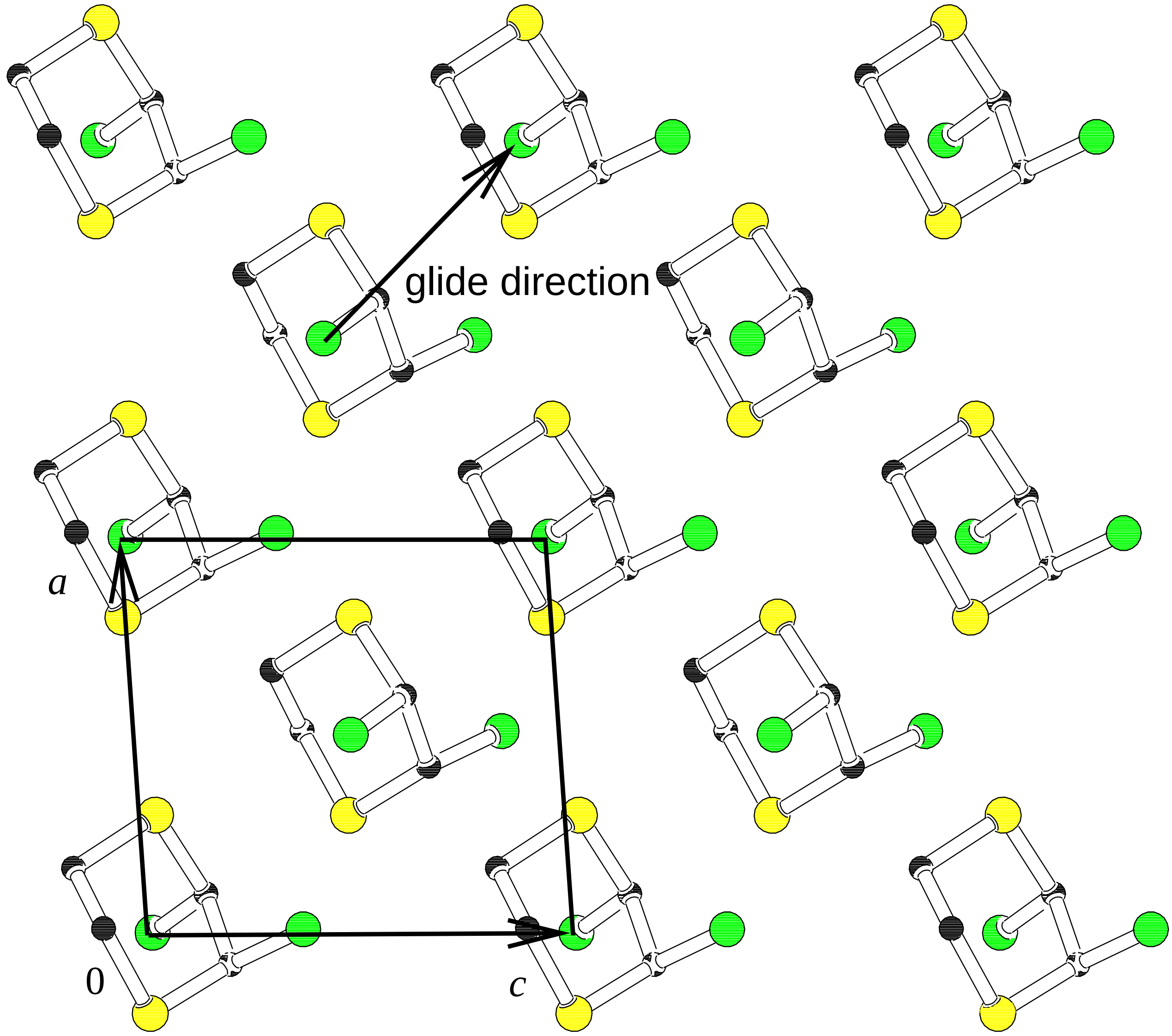


but not when projected down $\vec{c}$. 2_1 is only valid locally.



Non-standard space group settings (1)

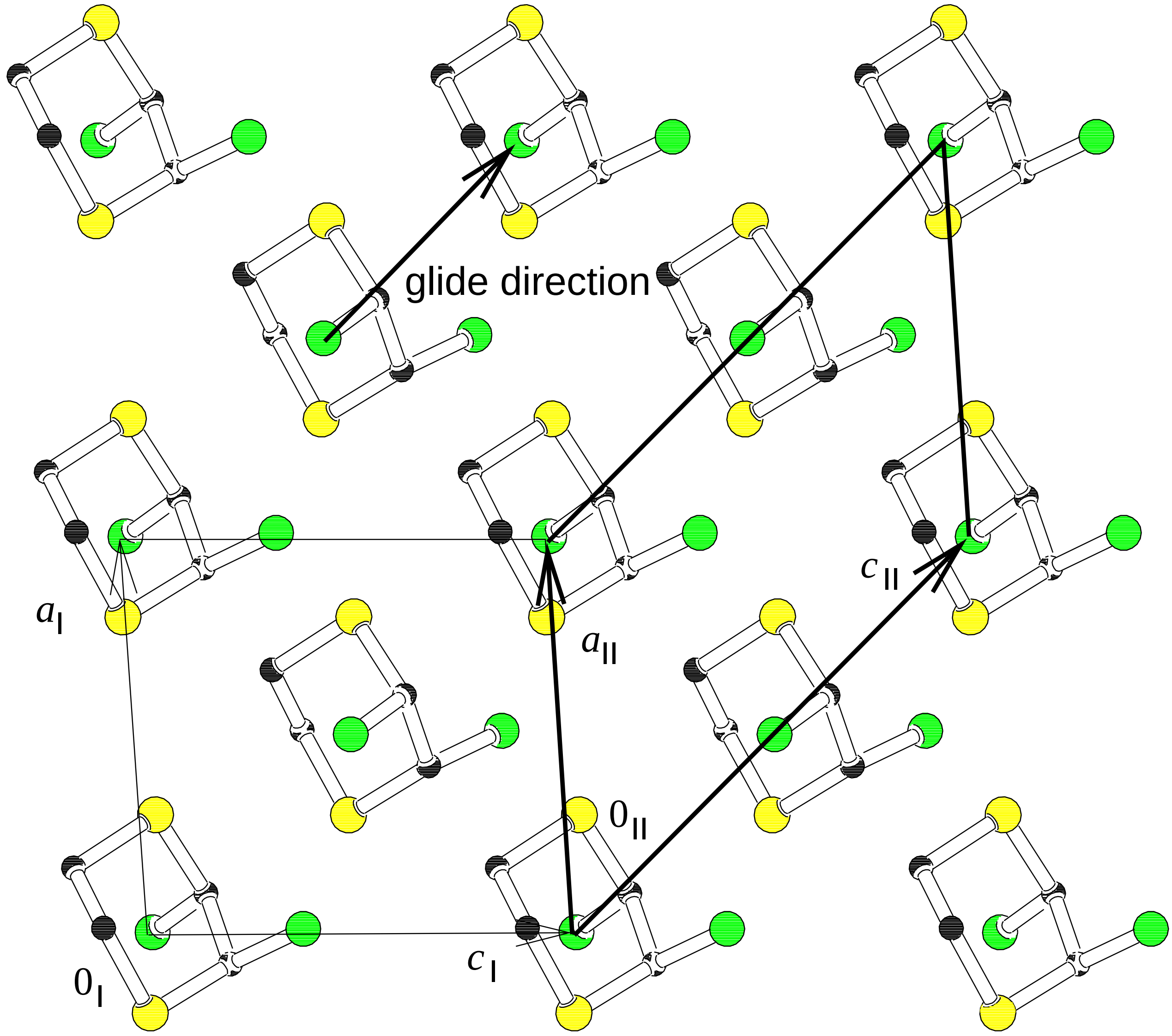
Projection down $\vec{b}$.



The translation component of the glide operation is $(\frac{1}{2}, 0, \frac{1}{2})$; the space group is therefore Pn.

Non-standard space group settings (2)

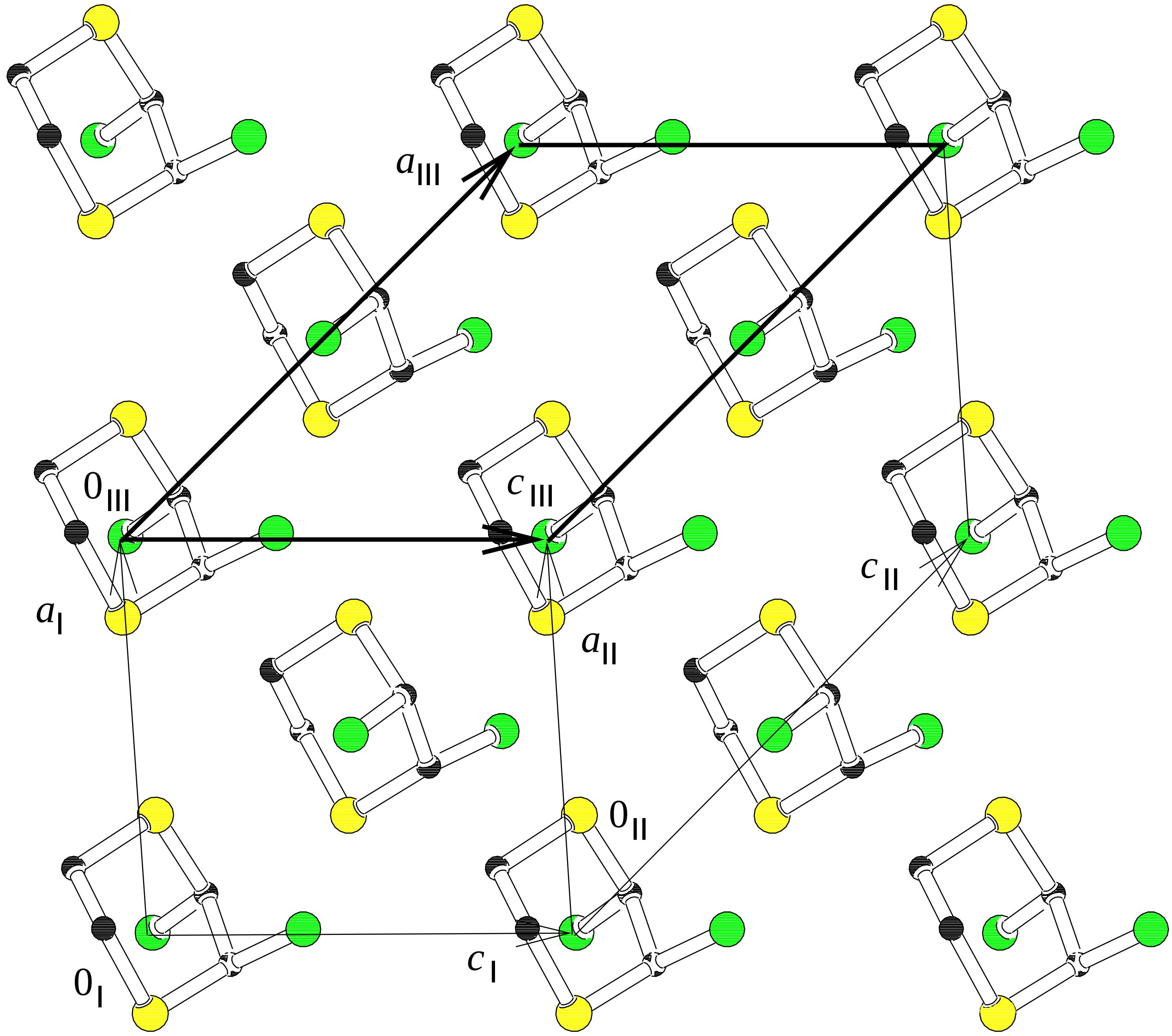
Projection down $\vec{b}$.



Another choice of unit cell (axes system II), leads to another expression for the translation component of the glide operation: $(0, 0, \frac{1}{2})$; the space group is now called Pc.

Non-standard space group settings (3)

Projection down $\vec{b}$.



unit cell choice	translation component	space group symbol
I	$(\frac{1}{2}, 0, \frac{1}{2})$	Pn
II	$(0, 0, \frac{1}{2})$	Pc
III	$(\frac{1}{2}, 0, 0)$	Pa

Pc is the standard setting.

Results of a structure determination

- unit cell parameters

  Obtained by fitting the cell parameters to the θ values (and other setting angles) of a number of reflections. For CCD data sets usually all reflections are used.

  s.u.'s are derived during the least-squares fitting.

- atomic co-ordinates, occupancy and displacement parameters

  derived from the intensities of the complete data set (i.e. measured to a certain resolution) and occasionally influenced by geometry information introduced through restraints.

  s.u.'s are derived during the least-squares refinement of the parameters. Constrained parameters are not refined and have therefore no s.u.

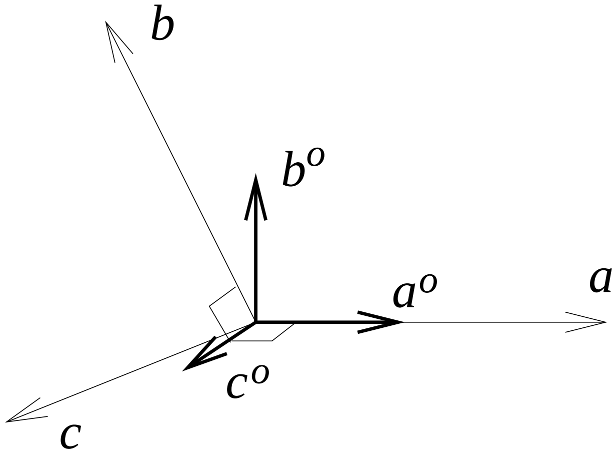
- space group

  Derived from systematic absences in the diffraction pattern, symmetry of the diffraction pattern, intensity distributions, general symmetry considerations and a thorough check of the final co-ordinate list.

Atomic co-ordinates

Atomic positions are given as fractional co-ordinates (x, y, z). For some purposes orthogonal (or more precise: orthonormal) co-ordinates (x^o, y^o, z^o), also called cartesian co-ordinates, are more convenient to work with.

There are many ways to transform fractional co-ordinates to cartesian co-ordinates. The following choice of an orthonormal axes system is widely used in current literature:



$$\vec{a^o} \parallel \vec{a} \qquad \vec{c^o} \parallel \vec{c^*} = \frac{[\vec{a} \times \vec{b}]}{V} \qquad \vec{b^o} = \vec{c^o} \times \vec{a^o}$$

These definitions imply that $\vec{b^o}$ lies in the $\vec{a}, \vec{b}$-plane.

During orthogonalization the information concerning the packing of asymmetric units is lost, unless symmetry-related co-ordinates are also transformed.

Fractional co-ordinates can always be transformed into a set of cartesian co-ordinates (if the unit cell is known).The reverse operation is only possible if the details of the orthogonalization are known (i.e. the orthogonalization matrix).

Co-ordinate file types

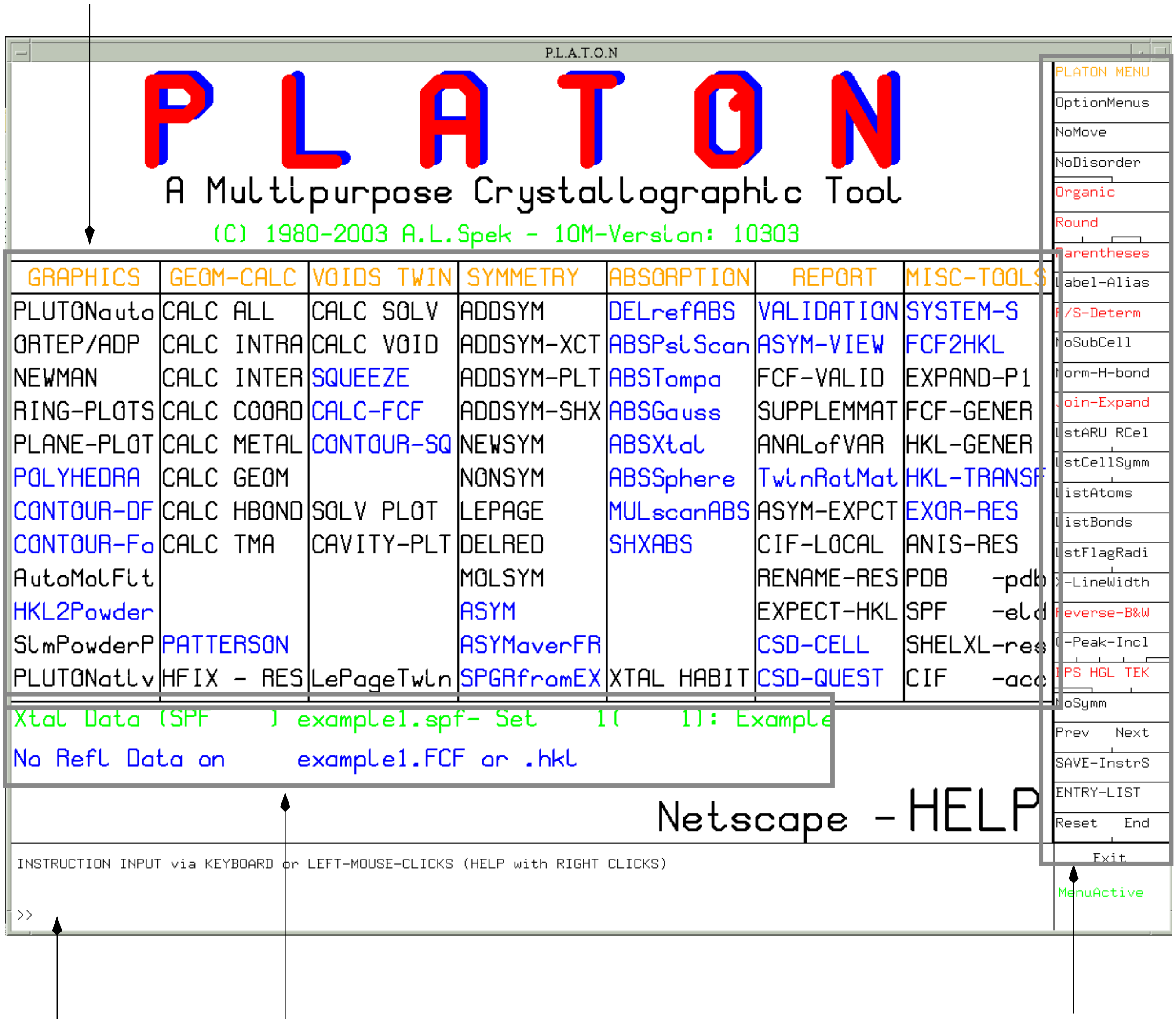
Several file types are used in crystallography:

- .spf Standard Parameter File (original PLATON format)
  Contains positional parameters; often also occupancy and displacement parameters and can contain s.u.'s for all reported parameters.
- .res SHELX result file
  Contains positional, occupancy and displacement parameters, but no s.u.'s.
- .pdb Protein Data Bank file
  Various formats possible, often only with orthogonal coordinates and without s.u.'s.
- .fdat Cambridge Structural Data base format
  Contains only positional parameters without s.u.'s, but with additional information (colour, R-values, etc.) in a compact but extremely human-unfriendly format.
- .cif Crystallographic Information File
  Can contain all the information concerning a structure determination, including data collection, procedures followed and the complete text and tables of an Acta Crystallographica Section C or E paper.

All these file types can be handled by PLATON

PLATON

main computing instructions (clickable menu)
options for which insufficient data are available are shown in blue



Xtal Data (SPF] example1.spf- Set 1(1]: Example
No Refl Data on example1.FCF or .hkl

INSTRUCTION INPUT via KEYBOARD or LEFT-MOUSE-CLICKS (HELP with RIGHT CLICKS)

names of input files

options to customize output
(clickable menu)

command line for keyboard instructions

File resolution:

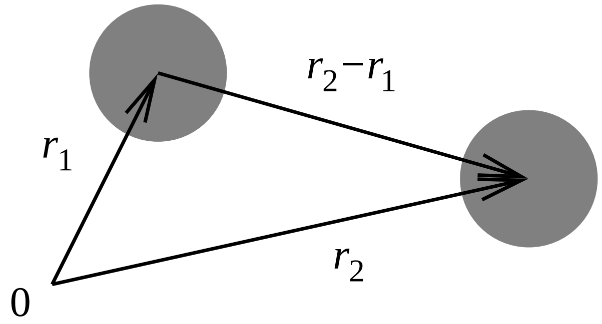
co-ordinate file (Xtal data)	reflection file 1st choice	reflection file 2nd choice
.spf	.hkl	.fcf
.res	.hkl	.fcf
.cif	.fcf	.hkl

Bond lengths and connectivity

Calculation of bond lengths in a system of fractional co-ordinates is performed with the expression:

$$\begin{aligned} d_{12}^2 &= |\vec{r}_2 - \vec{r}_1|^2 \\ &= (\vec{r}_2 - \vec{r}_1) \cdot (\vec{r}_2 - \vec{r}_1) \\ &= x_{12}^2 a^2 + y_{12}^2 b^2 + z_{12}^2 c^2 + 2x_{12}y_{12}ab\cos\gamma \\ &\quad +2x_{12}z_{12}ac\cos\beta + 2y_{12}z_{12}bc\cos\alpha \end{aligned}$$

. where $x_{12} = x_2 - x_1$, $y_{12} = y_2 - y_1$ and $z_{12} = z_2 - z_1$.



Establishing connectivity in a crystal structure is done on the basis of interatomic distances. Atoms i and j are considered bonded if

$$d_{ij} < r_i + r_j + t$$

where r_i is the covalent radius of atom i and t is the tolerance, an empirical parameter depending on the bond type, and the compound type (inorganic or organic). Some examples:

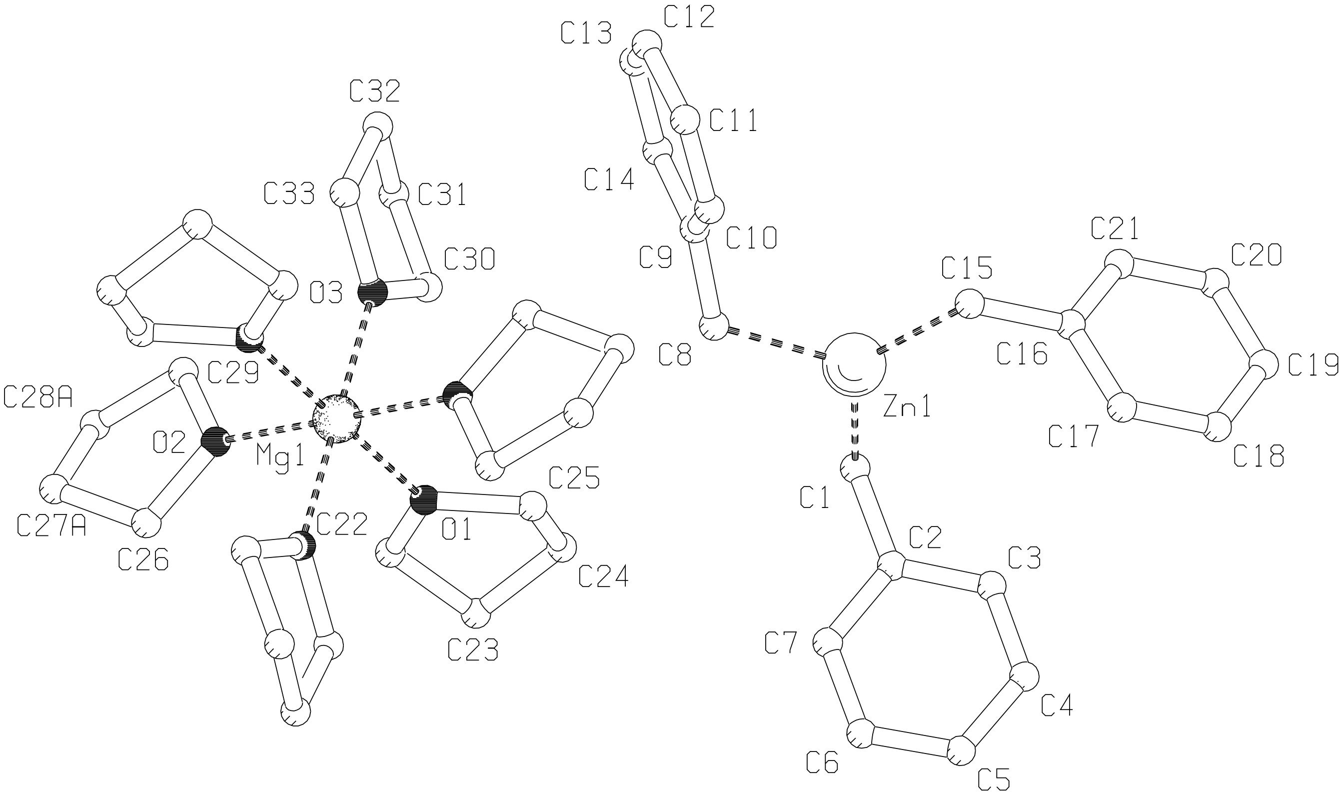
$t = 0.40$	default tolerance
$t = 0.00$	for metal—metal bonds
$t = 1.10$	for (earth)alkali—non-metal bonds
$t = 0.75$	for Cu,Mn$\cdots X$ contacts

A set of connected atoms is called a residue; it can be an isolated molecule, ion or a polymer.

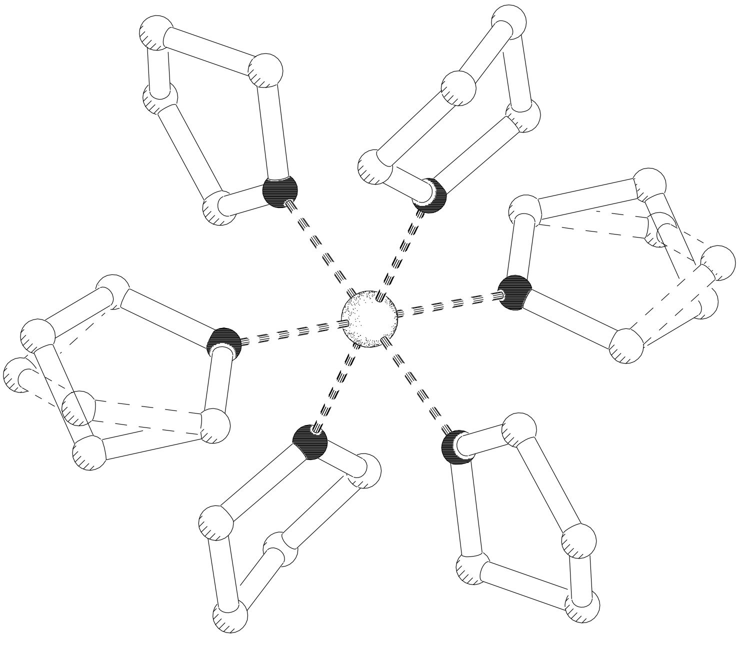
The Asymmetric Residue Unit (ARU) is that part of the residue which is crystallographically unique (contains no crystallographic symmetry).

Example 1

Magnesium bis(tribenzylzincate), $Mg(C_4H_8O)_6.[ZN(CH_2C_6H_5)_3]_2$. The Mg-containing residue is located on a crystallographic inversion centre.



The Mg(THF)$_6$ ion displays disorder in one of the unique THF moieties:



Symmetry coding in PLATON

Two notations are used to code symmetry operations in PLATON

- $n.ijk$
  n indicates equivalent position number of the current space group, according to the list given on page 1 of each PLATON listing
  i indicates a translation over $i-5$ in the $\vec{a}$ direction
  j indicates a translation over $j-5$ in the $\vec{b}$ direction
  k indicates a translation over $k-5$ in the $\vec{c}$ direction
  Used to report the transformation applied on input co-ordinates.
- $nijk.r$
  n, i, j, k as given above
  r residue number, as given in the co-ordinate list
  Used to refer to ARU's in the description of intermolecular interactions.

Example: Space group $P2_1/c$. Equivalent positions are

$$\begin{array}{lr} 1 & x, y, z \\ 2 & \bar{x}, \frac{1}{2}+y, \frac{1}{2}-z \\ 3 & \bar{x}, \bar{y}, \bar{z} \\ 4 & x, \frac{1}{2}-y, \frac{1}{2}+z \end{array}$$

Codes used in reporting transformations applied on input co-ordinates:
1.555 $\Rightarrow x, y, z$ (i.e. no transformation applied)
2.654 $\Rightarrow \bar{x}+1, \frac{1}{2}+y, \frac{1}{2}-z-1 = 1-x, \frac{1}{2}+y, -\frac{1}{2}-z$

Codes used in reporting intermolecular geometry:
1555.03 refers to ARU number 3, no symmetry operations applied, i.e. having the co-ordinates as reported at the begin of the PLATON listing.

4763.02 refers to ARU nr 2, on which operation $2+x, 1\frac{1}{2}-y, -1\frac{1}{2}+z$ has been applied to the co-ordinates reported in the PLATON listing (which may already have been transformed with respect to the input co-ordinates).

Incomplete atom lists

If atomic co-ordinates are not available for a part of the structure (e.g. because a solvent area was SQUEEZE'd) the information derived from the atom list **must** be adapted. Besides the chemical formula, the items concerned are:

- Formula mass (Formula weight)

$M_r = \sum_j n_j M_j^a$ sum over element types

M_j^a = atomic weight of element j

n_j = number of atoms of element j in the chemical unit

- Scattering power of the unit cell

$F_{000} = \sum_j n_j Z_j$ Z_j = atomic number of element j

- Crystallographic density

$$d_X = \frac{Z \sum_j n_j M_j^a}{V N_{\mathsf{Av}}}$$

V = unit cell volume

Z = number of chemical units in the unit cell (N.B. $Z \neq Z_j$)

N_{Av} = Avogadro's number

- Absorption coefficient

$$\mu = \frac{Z \sum_j n_j (\mu/\rho)_j M_j^a}{V N_{\mathsf{Av}}}$$

$(\mu/\rho)_j$ = mass absorption coefficient of j

Be careful when mixing units (cm, mm, Å)!

$$d_X[\text{g cm}^{-3}] = d_X[\text{Mg m}^{-3}] = \frac{Z \times \sum_j n_j \times M_j^a[\text{g mol}^{-1}]}{V[\text{Å}^{-3}] \times 0.6022}$$

$$\mu[\text{mm}^{-1}] = \frac{Z \times \sum_j n_j \times (\mu/\rho)_j[\text{cm}^{+2}\ \text{g}^{-1}] \times M_j^a[\text{g mol}^{-1}]}{V[\text{Å}^{-3}] \times 6.022}$$

Atomic contributions to M_r, d_X, F_{000} and μ

Atomic contributions to M_r, d_X, F_{000} and μ are summarized in the following table for the elements found most often in counterions and solvent molecules. (μ/ρ), the mass absorption coefficient or mass attenuation coefficient, is wavelength dependent. Comprehensive tables for various wavelengths can be found in the "International Tables for Crystallography", Volume C, (Kluwer, 2004), Table 4.2.4.3, page 230–236.

element	Z	M^a [g mol^{-1}]	μ/ρ[Mo$K\alpha$] [cm^{+2} g^{-1}]	μ/ρ[Cu$K\alpha$] [cm^{+2} g^{-1}]
H	1	1.008	0.373	0.391
B	5	10.81	0.368	2.31
C	6	12.01	0.576	4.51
N	7	14.01	0.845	7.44
O	8	16.00	1.22	11.44
F	9	19.00	1.63	15.78
Si	14	28.09	6.53	65.32
P	15	30.97	7.97	75.45
S	16	32.07	9.99	93.33
Cl	17	35.45	11.52	106.00
Se	34	78.96	68.82	82.89
Br	35	79.90	75.37	88.94
I	53	126.90	36.33	291.7

Standard uncertainty of a bond length

The error propagation formula

$$\sigma_f^2 = \sum_{j=1}^{N} \left(\frac{\partial f}{\partial x_j} \right)^2 \sigma_{x_j}^2$$

applied on the bond length expression (page 78) gives

$$\begin{aligned} \sigma_{d_{12}}^2 &= (\sigma_{x_1}^2 + \sigma_{x_2}^2)[(x_{12}a + y_{12}b\cos\gamma + z_{12}c\cos\beta)/d_{12}]^2 \\ &\quad +(\sigma_{y_1}^2 + \sigma_{y_2}^2)[(y_{12}b + x_{12}a\cos\gamma + z_{12}c\cos\alpha)/d_{12}]^2 \\ &\quad +(\sigma_{z_1}^2 + \sigma_{z_2}^2)[(z_{12}c + x_{12}a\cos\beta + y_{12}b\cos\alpha)/d_{12}]^2 \end{aligned}$$

which is only valid when the atomic co-ordinates involved are not correlated.

In practice, s.u.'s are calculated by substitution of numerically evaluated values of $(\partial f/\partial x_j)$ in the error propagation formula.

If the s.u. in the atomic co-ordinates of atom n is isotropic and can be represented by σ_n, the s.u. in the bond length is given by

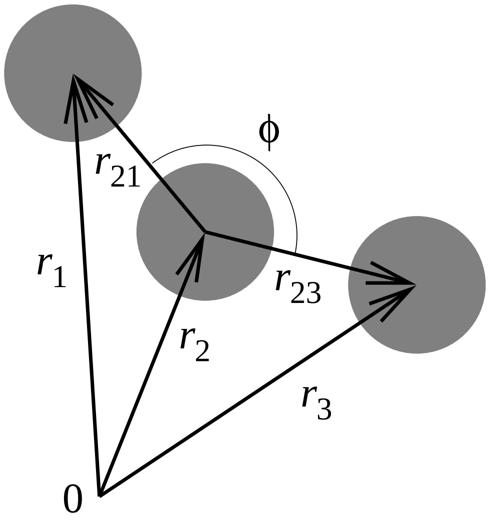
$$\sigma_{d_{12}}^2 = \sigma_1^2 + \sigma_2^2$$

S.u.'s in atomic co-ordinates are approximately proportional to $1/Z$. Therefore, a structure that consists of elements with large differences in atomic numbers, displays a large spread in bond length s.u.'s.

Example: A structure containing Zr ($Z = 40$), C ($Z = 6$) and H ($Z = 1$) could show the following values for the s.u. in the bond lengths of the following bond types:

$$\begin{aligned} \sigma(d_{\text{Zr}-\text{Zr}}) &= 0.0007\text{Å} \\ \sigma(d_{\text{Zr}-\text{C}}) &= 0.003\text{Å} \\ \sigma(d_{\text{C}-\text{C}}) &= 0.005\text{Å} \\ \sigma(d_{\text{C}-\text{H}}) &= 0.02\text{Å} \end{aligned}$$

Bond angles



The bond angle 1—2—3, also called valence angle, can be calculated with

$$\begin{aligned}\cos\phi &= \frac{\vec{r}_{21}\cdot\vec{r}_{23}}{|\vec{r}_{21}||\vec{r}_{23}|} \\ &= [x_{21}x_{23}a^2 + y_{21}y_{23}b^2 + z_{21}z_{23}c^2 \\ &\quad +(x_{21}y_{23}+y_{21}x_{23})ab\cos\gamma \\ &\quad +(x_{21}z_{23}+z_{21}x_{23})ac\cos\beta \\ &\quad +(y_{21}z_{23}+z_{21}y_{23})bc\cos\alpha]/(d_{12}d_{13})\end{aligned}$$

or with

$$\cos\phi = \frac{d_{12}^2 + d_{23}^2 - d_{13}^2}{2d_{12}d_{13}}$$

d can be calculated with the expression on page 78.

If the s.u. in the atomic co-ordinates of atom n is isotropic and can be represented by σ_n, the s.u. in the bond angle is given by

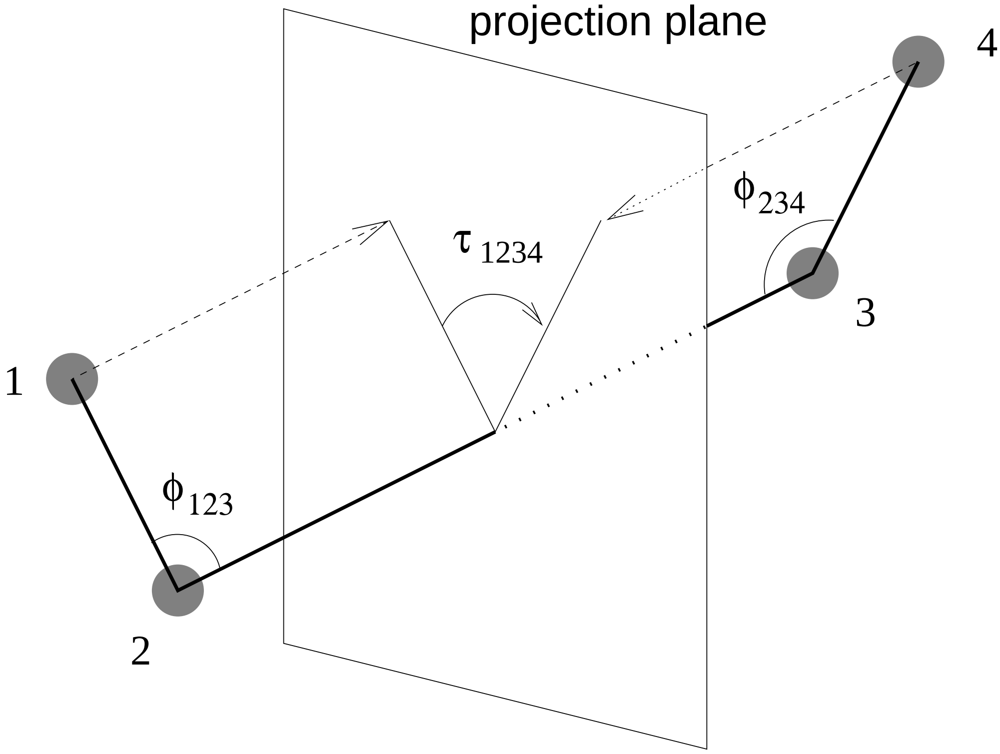
$$\sigma_\phi^2 = \frac{\sigma_1^2}{d_{21}^2} + \frac{\sigma_2^2 d_{13}^2}{d_{21}^2 d_{23}^2} + \frac{\sigma_3^2}{d_{23}^2}$$

σ_ϕ is dependent on the interatomic distances 1—2, 2—3 and 1—3, and therefore also on ϕ

The dependency of σ_n on Z also leaves its traces on σ_ϕ.

Torsion angles (1)

The torsion angle τ_{1234} is defined as the angle between the projections of the vectors $\vec{r}_{2\to1}$ and $\vec{r}_{3\to4}$ on the bisecting plane of the bond 2—3.



The torsion angle is calculated as the angle between the normal vectors to the plane through atoms 1, 2 and 3 and the plane through atoms 2, 3 and 4:

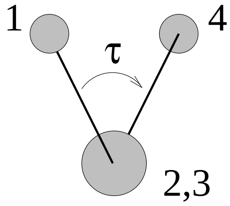
$$\tau_{1234} = \frac{(\vec{r}_{21} \times \vec{r}_{32}) \cdot (\vec{r}_{32} \times \vec{r}_{43})}{|\vec{r}_{21}||\vec{r}_{32}|^2|\vec{r}_{43}| \sin\phi_{123} \sin\phi_{234}}$$

Torsion angles with embedded bond angles larger than 160° become very susceptible for large s.u.'s and are therefore not calculated in PLATON.

σ_τ is also sensitive for the dependency of σ_n on Z.

Torsion angles (2)

τ is positive when the smallest rotation needed to superimpose 1 on 4 in the projection $2 \rightarrow 3$ is clockwise (Klyne & Prelog convention).

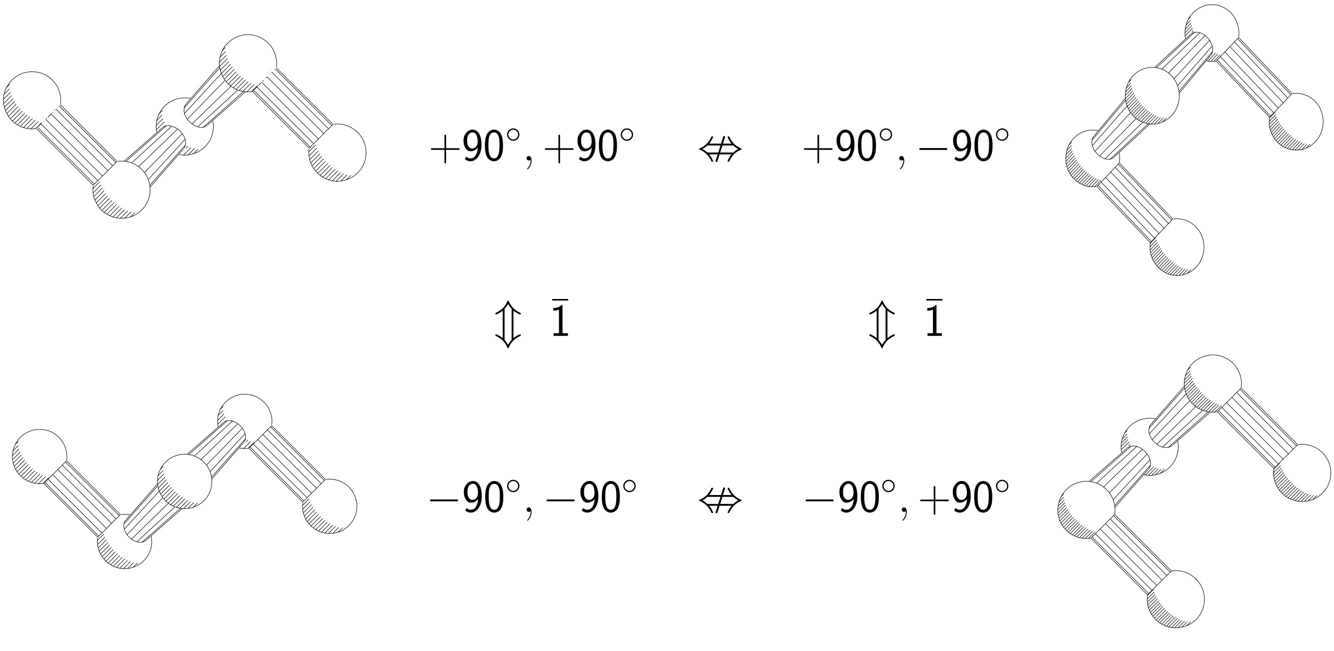


The sign of the torsion angle does not change when the order of the atoms is reversed: $\tau_{1234} = \tau_{4321}$.

The sign of the torsion angle does not change when a proper rotation, screw axis or translation is applied.

The sign of the torsion angle does change when an improper rotation ($\bar{1}$, m, $\bar{3}$, $\bar{4}$, $\bar{6}$) or a glide plane is applied.

If two torsion angles in a molecule, crystallizing in a centrosymmetric space group, have values $+\phi, +\psi$, there is also a symmetry-related molecule with torsion angles $-\phi, -\psi$. There are no molecules with torsion angles $+\phi, -\psi$ or $-\phi, +\psi$.



Torsion angle nomenclature

The term “dihedral angle” for τ is not correct. By definition the dihedral angle (the angle between two planes) is acute. The torsion angle can also be obtuse. The (mis)use of dihedral angle is fairly wide spread in literature.

Descriptors for torsion angles (numbers in degrees):

−150				−150
		(−ap)		
±180	antiperiplanar		trans	±180
		(+ap)		
+150				+150
+120	anticlinal	(+ac)		+120
+90				+90
+60	synclinal	(+sc)	+gauche	+60
+30				+30
		(+sp)		
0	synperiplanar			0
		(−sp)		
−30				−30
−60	synclinal	(−sc)	−gauche	−60
−90				−90
−120	anticlinal	(−ac)		−120
−150				−150
		(−ap)		
∓180	antiperiplanar		trans	∓180
		(+ap)		
+150				+150

Comparison of geometrical parameters (1)

Probability theory can be used to evaluate the significance of differences in observed geometrical parameters ($\Delta = g_1 - g_2$), assuming that the studied parameter (g) follows the normal distribution.

The standard uncertainty in the difference Δ equals

$$\sigma_\Delta = \sqrt{\sigma_{g_1}^2 + \sigma_{g_2}^2}$$

The difference Δ can now be expressed in units of σ_Δ:

$$\Delta = \lambda \sigma_\Delta$$

The probability that Δ only results from random errors (implying that g_1 and g_2 are measurements of the same quantity) can be derived from the following table:

p	λ
1.0	0.000
0.9	0.126
0.7	0.385
0.5	0.674
0.3	1.04
0.1	1.65
0.05	1.96
0.01	2.58
0.001	3.29
0.0001	3.89

The following arbitrary limits of significance are commonly used in statistics:

$p > 0.05$	no significant difference
$0.01 < p < 0.05$	possible significant difference
$p < 0.01$	significant difference

Comparison of geometrical parameters (2)

Example

Comparing the Zn—C bonds of Example listing 1.

bond	d [Å]
Zn1—C1	2.051(3)
Zn1—C8	2.0582(19)
Zn1—C15	2.032(2)

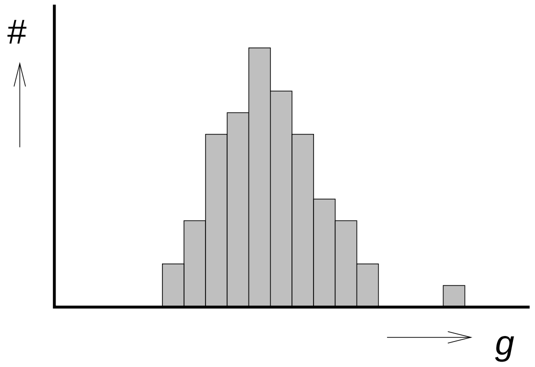
For each pair of bonds, the difference in bond distances ($\Delta_{12} = |d_1 - d_2|$) and its standard uncertainty ($\sigma_{\Delta_{12}}$) can be calculated. The bond lengths difference can be expressed in units of standard deviations using $\lambda = (\Delta_{12}/\sigma_{\Delta_{12}})$. The table on page 88 enables the determination of the probability level and the significance range.

bond 1	bond 2	Δ [Å]	σ_Δ [Å]	λ	p range
Zn1—C1	Zn1—C8	0.0072	0.0036	2.03	$0.01 < p < 0.05$
Zn1—C1	Zn1—C15	0.0190	0.0028	6.79	$p < 0.0001$
Zn1—C8	Zn1—C15	0.0320	0.0036	8.89	$p < 0.0001$

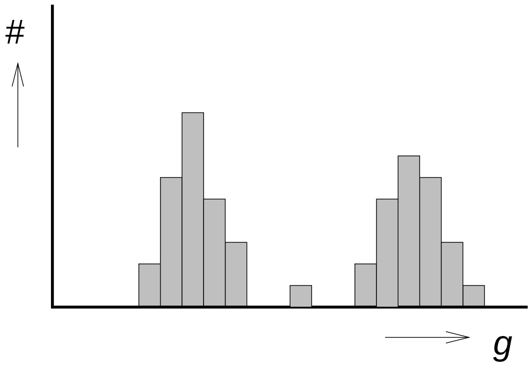
The observed difference in bond lengths Zn1—C1 and Zn1—C8 is possibly significant; the observed differences in the other bond lengths are significant.

Averaging of geometrical parameters (1)

Averaging of geometrical parameters is only allowed if they represent different measurements of the same quantity. If enough measurements are available, a distribution histogram of g can give an indication if this is the case.



averaging appears valid



averaging is not valid

The χ^2 distribution is used to test quantitatively wether a sample of N observed geometrical parameters g_i are measurements of the same quantity. The s.u. in the sample g_i is compared to the average s.u. of the individual measurements by calculating the ratio

$$R = \frac{\sum_{i=1}^{N} (g_i - \bar{g}_i)^2}{\left(\sum_{i=1}^{N} (\sigma_{g_i}/N) \right)^2}$$

This is only valid if σ_g shows a relatively small spread.

The probability that the sample (with $N-1$ degrees of freedom) represents a single quantity, can be obtained from the tabulated χ^2 distribution.

If the χ^2-test is failed, or the results are unclear, it is better to report a range of observed parameters than an averaged value.

Averaging of geometrical parameters (2)

The χ^2 distribution.

$N-1$	p							
	0.990	0.950	0.900	0.500	0.100	0.050	0.025	0.010
1	0.00	0.00	0.02	0.45	2.71	3.84	5.02	6.63
2	0.02	0.10	0.21	1.39	4.61	5.99	7.38	9.21
3	0.11	0.22	0.35	2.37	6.25	7.81	9.35	11.34
4	0.30	0.71	1.06	3.36	7.78	9.49	11.14	13.28
5	0.55	1.15	1.61	4.35	9.24	11.07	12.83	15.09
10	2.56	3.94	4.87	9.34	15.99	18.31	20.48	23.21
15	5.23	7.26	8.55	14.34	22.31	25.00	27.49	30.58
20	8.26	10.85	12.44	19.34	28.41	31.41	34.17	37.57
25	11.52	14.61	16.47	24.34	34.38	37.65	40.65	44.31
30	14.95	18.49	20.60	29.34	40.26	43.77	46.98	50.89
40	22.16	26.51	29.05	39.34	51.80	55.76	59.34	63.69
50	29.71	34.76	37.69	49.33	63.17	67.50	71.42	76.15
60	37.48	43.19	46.46	59.33	74.40	79.80	83.30	88.38
70	45.44	51.74	55.33	69.33	85.83	90.53	95.02	100.42
80	53.54	60.39	64.28	79.33	96.58	101.88	106.63	112.33
90	61.75	69.13	73.29	89.33	107.56	113.14	118.14	124.12
100	70.06	77.93	82.36	99.33	118.50	124.34	129.56	135.81

If the χ^2-test is satisfied the average and its s.u. can be calculated using:

$$\bar{g} = \sum_{i=1}^{N} g_i/N$$

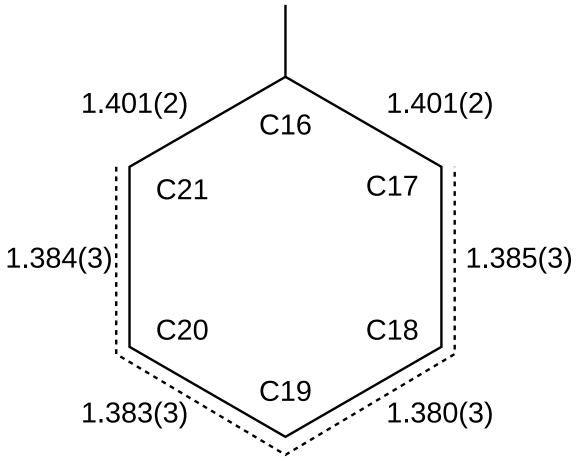
$$\sigma_{\bar{g}}^2 = \frac{\sum_{i=1}^{N}(g_i - \bar{g})^2}{N(N-1)}$$

The expression for $\sigma_{\bar{g}}$ is only valid if the measurements are independent.

N.B. The bond lengths of bonds that have one atom in common are not independent.

Averaging of geometrical parameters (3)

Example: Phenyl ring bond lengths (taken from example listing 1).



For all bond lengths in the phenyl ring

$$R = \frac{0.000470}{0.000007} = 67.1$$

The table on page 91 can be used to find that $p < 0.01$ for five degrees of freedom. The six bond lengths of this phenyl ring do not represent the same quantity; the bonds are not chemically equivalent.

When the test is done on the bonds indicated with the dotted lines, the following results are obtained:

$$R = \frac{0.000014}{0.000009} = 1.56$$

Using the table on page 91 we find $0.50 < p < 0.90$ for three degrees of freedom. These four bonds represent, with a large probability, the same quantity; they are chemically equivalent at the accuracy level where the crystal structure was determined. They can now be averaged:

$$\bar{d} = 1.3830(11)\text{Å}$$

N.B. The s.u. in the average bond length is calculated with the expression give before. This is in principle incorrect, since the bond lengths are not truely independent.

Least-squares planes (1)

A least-squares fit of a plane through four or more atoms gives the parameters for the plane equation:

$$px + qy + rz = s$$

(A plane through three atoms is defined by the same equation, but is determined exactly.)

The deviation $d_j$of an atom j is the shortest distance from this atom to the plane and can be calculated with

$$d_j = px_j + qy_j + rz_j - s$$

Least-squares planes are automatically calculated for rings (maximum number of atoms $N = 6$), residues, co-ordination planes, and planar fragments of at least four connected atoms. Additional planes can be specified with the LSPL keyboard instruction of PLATON.

In order to check planarity, PLATON calculates the following quantities:

$$\text{Sigref} = \sqrt{\sum_{j=1}^{N} \sigma_j^2/N}$$

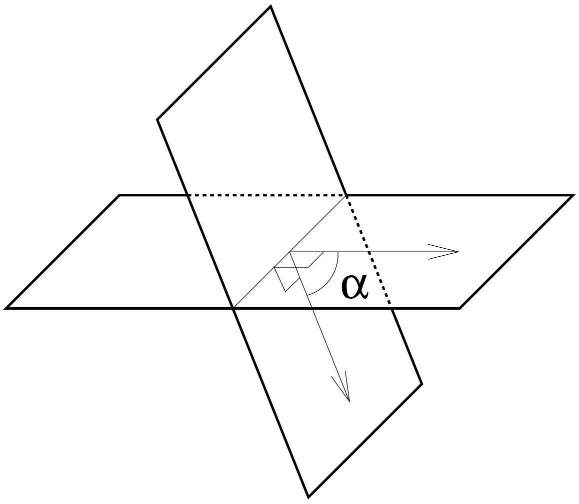
with σ_j the isotropic s.u. of atom j.

$$\text{Sigplan} = \sqrt{\sum_{j=1}^{N} d_j^2/(N-3)}$$

$$\text{Chisq} = N(\sum_{j=1}^{N} d_j^2)/(\sum_{j=1}^{N} \sigma_j^2)$$

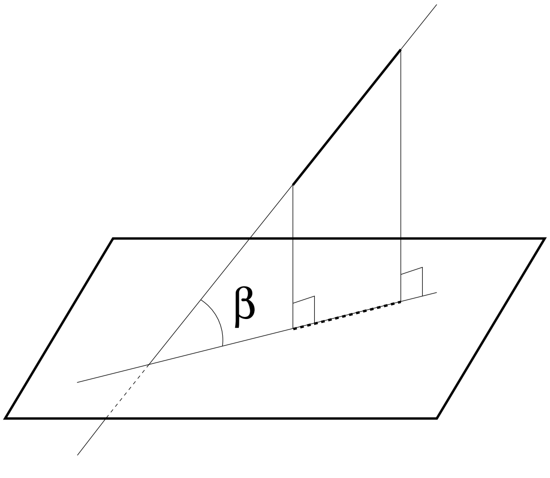
Least-squares planes (2)

Angle between planes (dihedral angles)



α is the acute angle between two lines, each in a different plane and both perpendicular to the intersection line. $0^\circ \leq \alpha \leq 90^\circ$, in contrast to torsion angles.

Angle between a bond and a plane



β is the acute angle between a line through the bond and the perpendicular projection of that bond on the plane. $0^\circ \leq \beta \leq 90^\circ$

Ring puckering—Cremer and Pople parameters

The geometrical centre C_g of an N-membered ring with cartesian co-ordinates $\vec{r}_j$ is placed on the origin:

$$\vec{R}_j = \vec{r}_j - \vec{C}_g = \vec{r}_j - \sum\nolimits_{j=1}^{N} \vec{r}_j / N$$

Two vectors define the reference plane:

$$\vec{R}_x = \sum\nolimits_{j=1}^{N} \vec{R}_j \cos[2\pi(j-1)/N]$$

$$\vec{R}_y = \sum\nolimits_{j=1}^{N} \vec{R}_j \sin[2\pi(j-1)/N]$$

This plane does not necessarily coincide with the least-squares plane. The unit vector perpendicular to the reference plane is:

$$\vec{n}_z = \vec{R}_x \times \vec{R}_y / |\vec{R}_x \times \vec{R}_y|$$

The deviations of each atom from the plane (or amplitudes) are

$$z_j = \vec{R}_j \cdot \vec{n_z}$$

Since C_g is located on the origin, $\sum_j z_j = 0$. The puckering parameters are now defined as:

$$Q_m \cos\phi_m = (2/N)^{1/2} \sum\nolimits_{j=1}^{N} z_j \cos[2\pi m(j-1)/N]$$

$$Q_m \sin\phi_m = -(2/N)^{1/2} \sum\nolimits_{j=1}^{N} z_j \sin[2\pi m(j-1)/N]$$

for integer m with $2 \leq m \leq (N-1)/2$. If N is even an additional parameter is defined:

$$Q_{N/2} = N^{-1/2} \sum\nolimits_{j=1}^{N} z_j \cos[\pi(j-1)]$$

The total puckering amplitude is

$$Q^2 = \sum\nolimits_{j=1}^{N} z_j^2 = \sum\nolimits_{m=1}^{N/2} Q_m^2$$

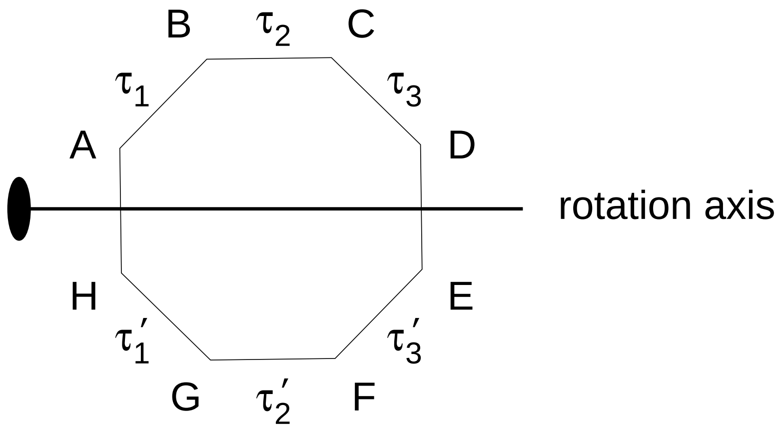
(Don't confuse Q^2 with Q_2.)

Ring puckering — asymmetry parameters

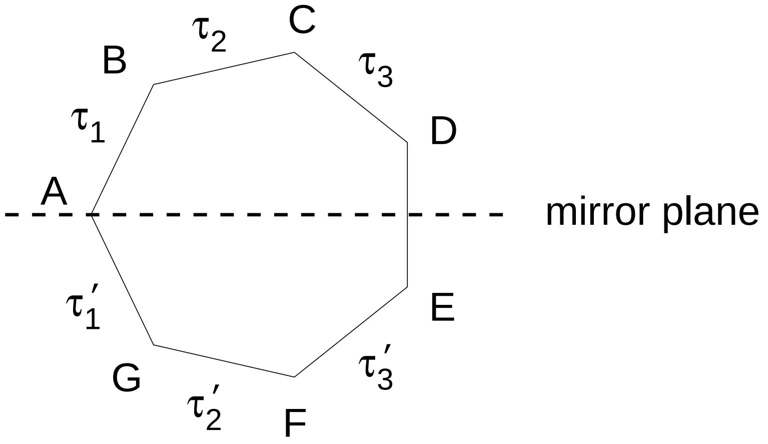
Ideal ring conformations are characterized by the presence of local symmetry elements in the ring. An asymmetry parameter quantifies the deviation from this ideal symmetry.

Asymmetry parameters are calculated for two-fold rotation axes lying within the ring plane and mirror planes perpendicular to the ring plane. The symmetry elements can run through atoms or through midpoints of bonds.

The calculation involves the comparison of the endocyclic torsion angles related by the local symmetry element.



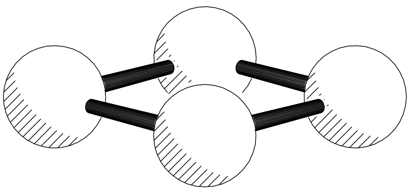
$$\Delta C_2(\mathsf{A}-\mathsf{H}) = \Delta C_2(\mathsf{D}-\mathsf{E}) = \sqrt{\frac{\sum_{i=1}^{m}(\tau_i - \tau_i')^2}{m}}$$

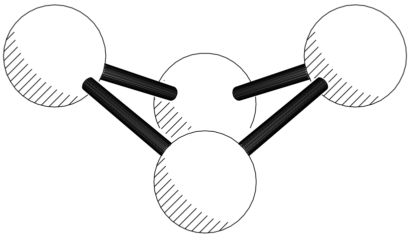


$$\Delta C_s(\mathsf{A}) = \Delta C_s(\mathsf{D}-\mathsf{E}) = \sqrt{\frac{\sum_{i=1}^{m}(\tau_i + \tau_i')^2}{m}}$$

4-Membered rings

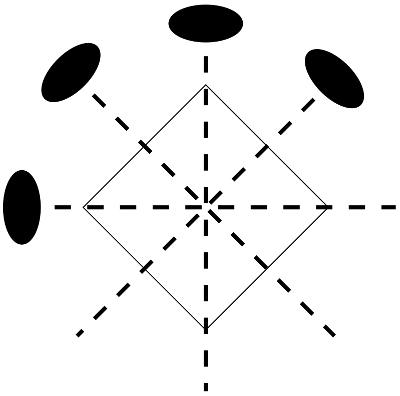
4-Membered rings are either flat or bent:



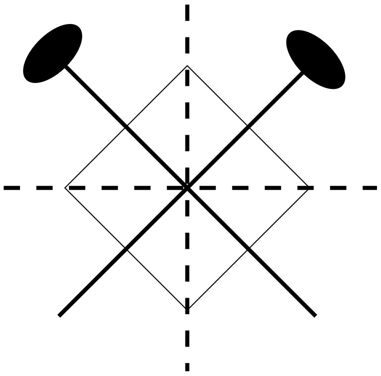


The Cremer and Pople puckering parameter Q_2 measures amplitude of the atoms (their deviation from the reference plane) and is therefore a clear indicator of the conformation.

The expected symmetry of the conformations is given below:

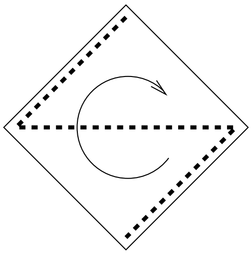


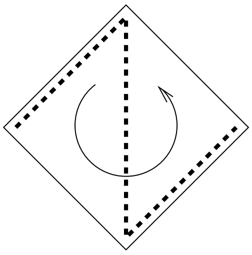
flat



bent

PLATON also reports two “improper” torsion angles in order to determine the conformation:

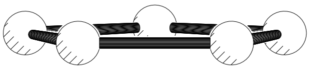




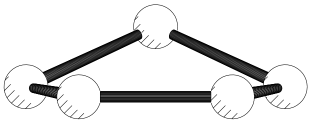
These torsions are 0° for planar rings.

5-Membered rings — conformations

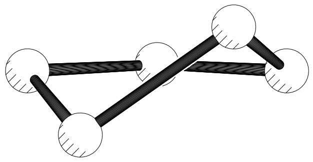
The three main conformations of a 5-membered ring, with their expected symmetry are:



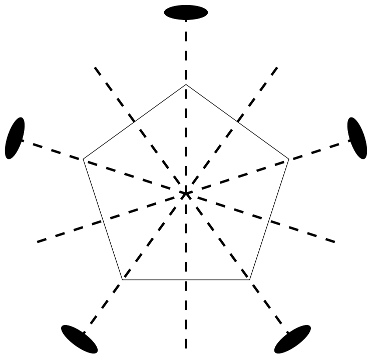
P
planar

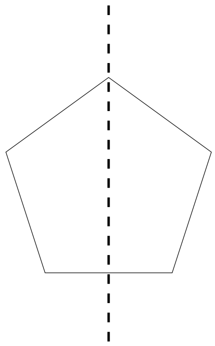


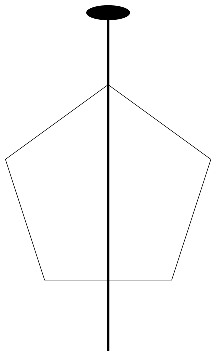
E
envelope



T
half-chair (twist)







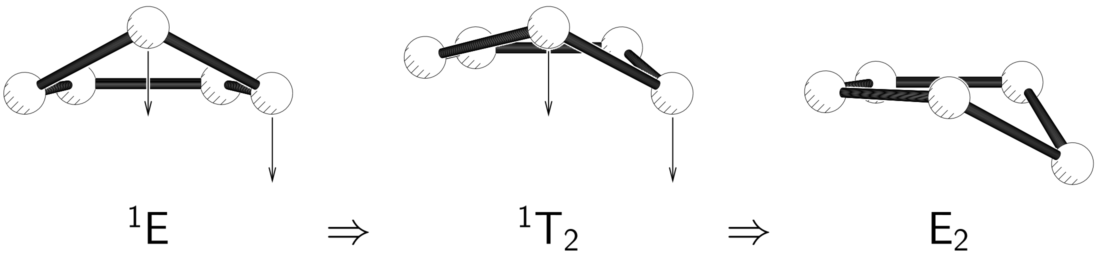
The Cremer & Pople Q_2 parameter specifies the puckering amplitude and can distinguish planar from non-planar conformations. The ϕ_2 parameter is $k \times 36^\circ$ for an envelope conformation and $k \times 36^\circ + 18^\circ$ for a half-chair conformation ($k \in \mathbb{Z}$).

Substituents are characterized by the angle ψ between the connecting bond to the ring and the normal to the Cremer & Pople reference plane:

ψ-range	description
$0^\circ < \psi < 30^\circ$	axial
$30^\circ < \psi < 60^\circ$	bisectional
$60^\circ < \psi < 90^\circ$	equatorial

5-Membered rings — pseudorotation (1)

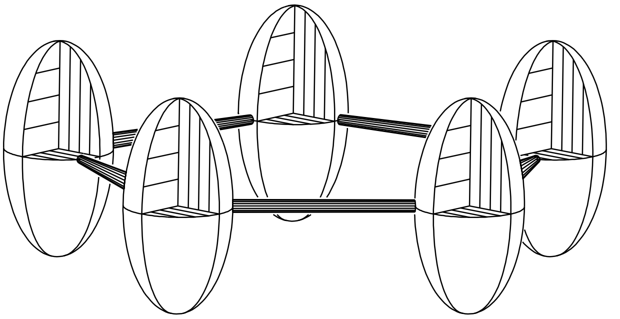
An envelope conformation can be brought into a half-chair conformation by applying relatively small atomic shifts. This half-chair conformation can be transformed into a new envelope conformation.



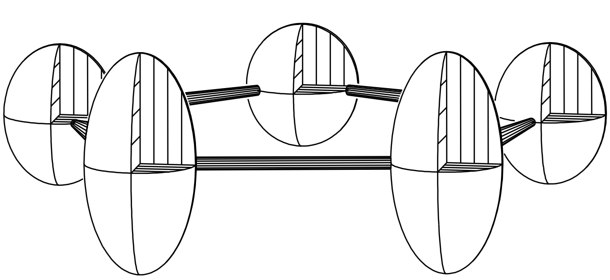
$^1T_2$ is a half-chair conformation with atom 1 above the Cremer & Pople reference plane, and atom 2 below the reference plane.

Since the protruding atom appears to be travelling around the ring, this process is called pseudorotation.

Due to disorder, caused by pseudorotation, 5-membered rings in crystal structures sometimes appear to be planar, with unusually short bond lengths. In these cases, the displacement parameters perpendicular to the ring plane are unusually large.



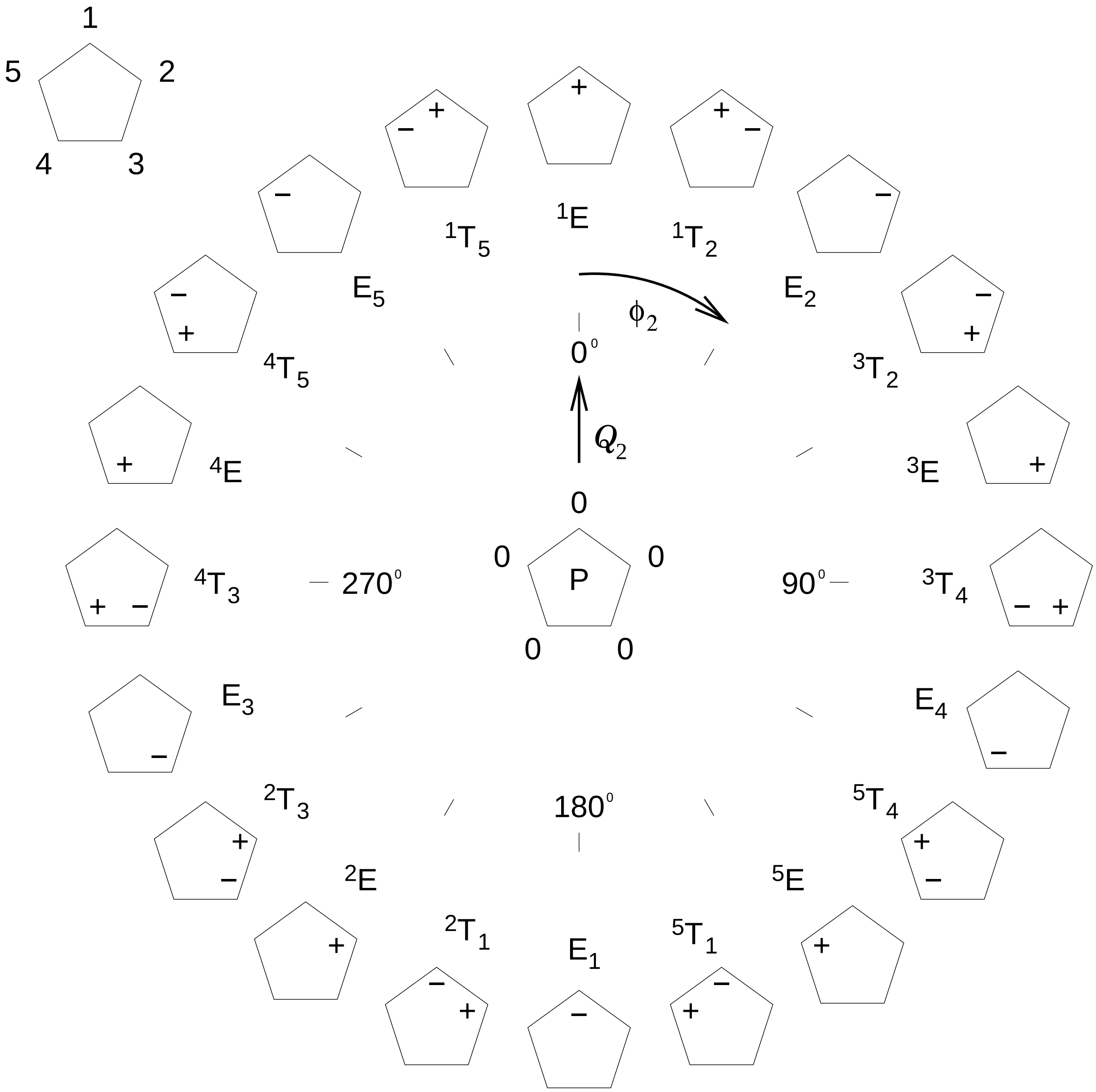
complete pseudorotational disorder



“flip-flop” disorder

5-Membered rings — pseudorotation (2)

The complete pseudorotation pathway of a 5-membered ring can be drawn in a polar co-ordinate diagram of Q_2 and ϕ_2:

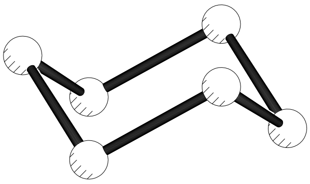


+ and − indicate positions above and below the plane of projection; atoms with no symbol or 0 lie in the plane of projection.

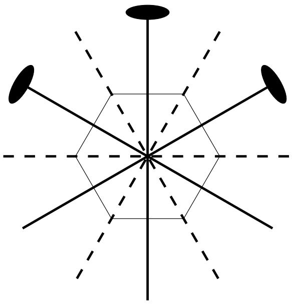
E-conformations (and T-conformations) can only be distinguished when hetero-atoms or substituents are present.

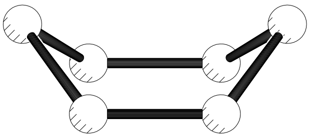
6-Membered rings — conformations

A 6-membered ring can adopt the planar conformation (P) or one of the following 6 puckered conformations:

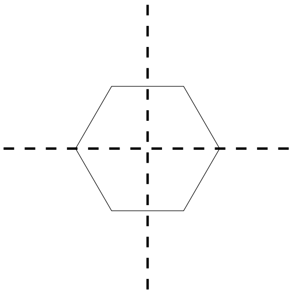


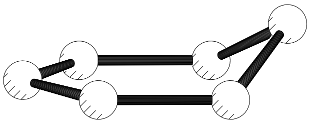
chair
C



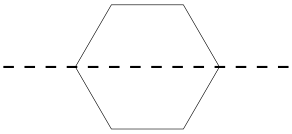


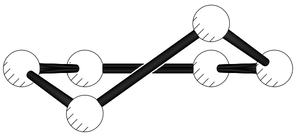
boat
B



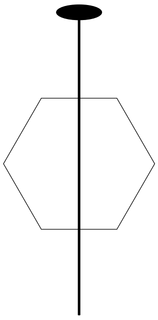


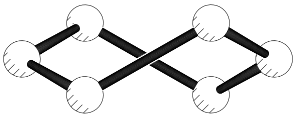
envelope
E
(sofa; half-boat HB)



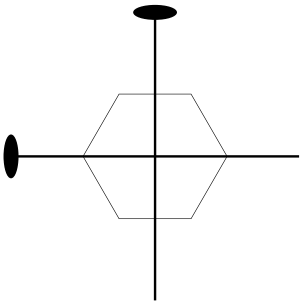


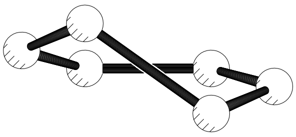
half-chair
H
(HC)



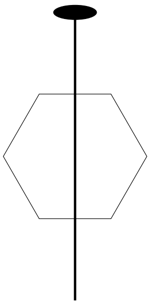


twist-boat
T
(TB; skew-boat)





screw-boat
S
(in PLATON versions before 10-V-99 also called skew-boat)



Asymmetry parameters can not distinguish H from S.

6-Membered rings — Cremer & Pople parameters

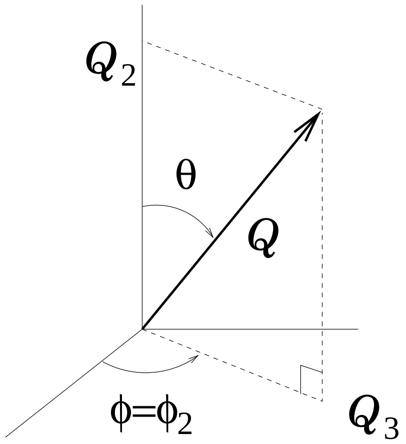
Analysis of the puckering of a 6-membered ring according to Cremer & Pople gives the parameters Q_2, Q_3 and ϕ_2. These cylindrical co-ordinates can be transformed to a polar co-ordinate system (Q, θ, ϕ) using

$$Q_2 = Q \sin\theta$$

$$Q_3 = Q \cos\theta$$

$$\phi_2 = \phi$$

where $0^\circ < \theta < 180^\circ$. Q is the total puckering amplitude ($Q_2^2 + Q_3^2 = Q^2$).



For a given puckering amplitude Q, conformations can be mapped on a sphere:

conformation	θ	ϕ
C	0° or 180°	$0^\circ < \phi < 360^\circ$
H	50.8° or 129.2°	$k \times 60^\circ + 30^\circ$
E	54.7° or 125.3°	$k \times 60^\circ$
S	67.5° or 112.5°	$k \times 60^\circ + 30^\circ$
B	90°	$k \times 60^\circ$
T	90°	$k \times 60^\circ + 30^\circ$

$(k \in \mathbb{Z})$

The planar conformation has $Q = Q_2 = Q_3 = 0^\circ$.

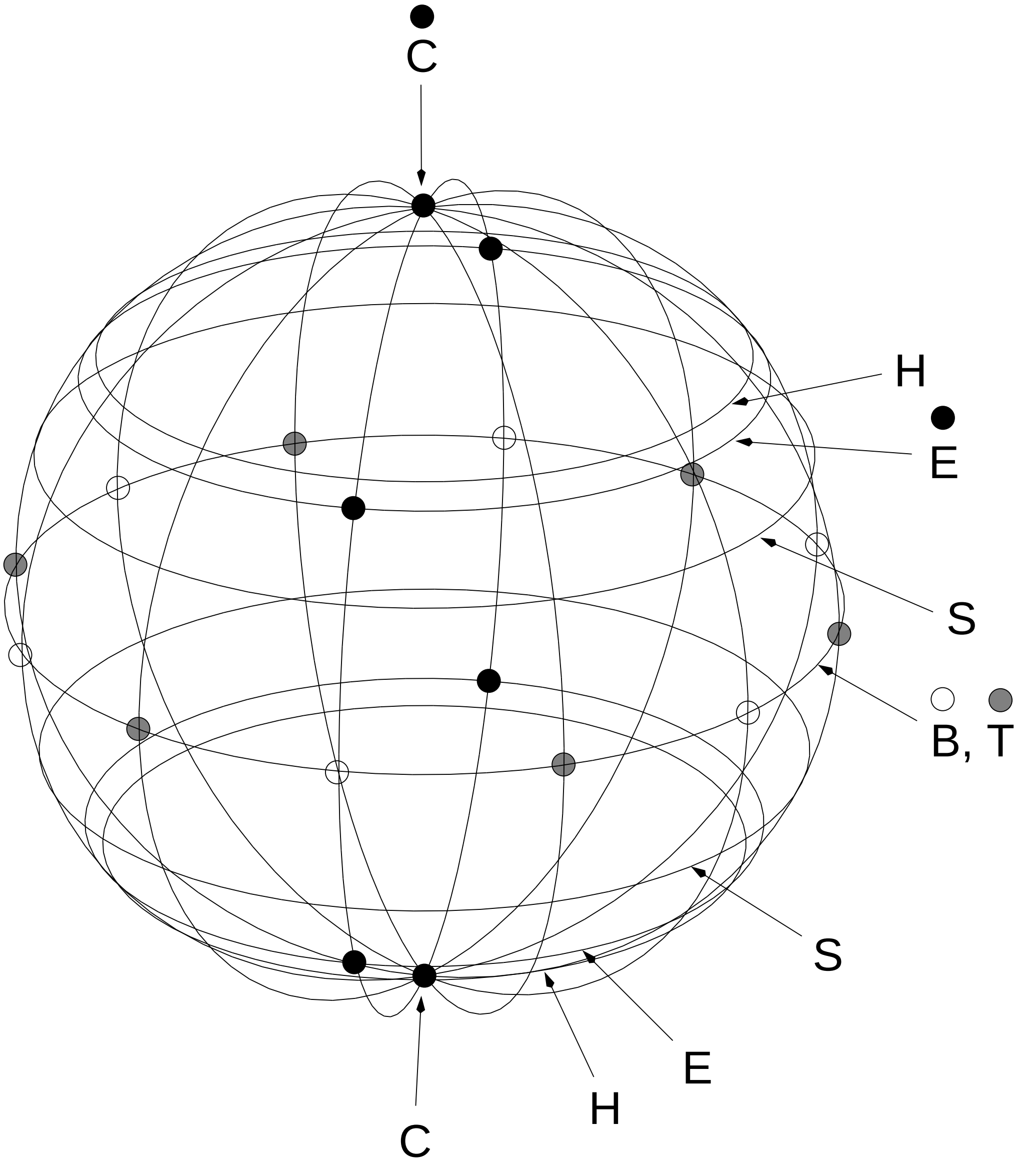
6-Membered rings — Pseudorotation pathways

Several pseudorotation pathways can be mapped on the sphere, e.g along a meridian:

$$^{1}C_{4} \rightarrow E_{4} \rightarrow B_{4,1} \rightarrow E_{1} \rightarrow {}^{4}C_{1} \rightarrow {}^{4}E \rightarrow {}^{1,4}B \rightarrow {}^{1}E \rightarrow {}^{1}C_{4}$$

or along the equator of the sphere:

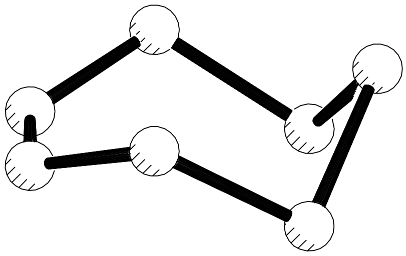
$$^{1,4}B \rightarrow {}^{4}T_{2} \rightarrow B_{2,5} \rightarrow {}^{6}T_{2} \rightarrow {}^{3,6}B \rightarrow {}^{3}T_{1} \rightarrow B_{4,1} \rightarrow$$
$$\rightarrow {}^{2}T_{4} \rightarrow {}^{2,5}B \rightarrow {}^{2}T_{6} \rightarrow B_{3,6} \rightarrow {}^{1}T_{3} \rightarrow {}^{1,4}B$$



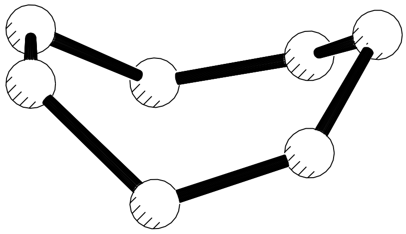
N.B. Even in the presence of heteroatoms and substituents, the symbols $^{1}C_{4}$, $^{2}C_{5}$ and $^{3}C_{6}$ refer to the same conformation.

7-Membered rings — Conformations

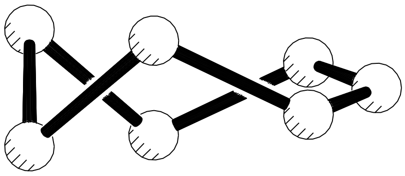
There are 6 main conformations for 7-membered rings:



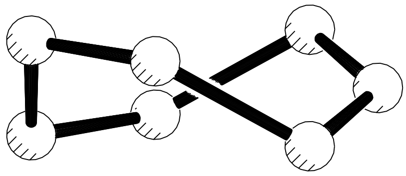
chair
C



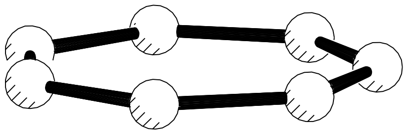
boat
B



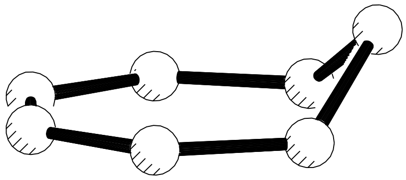
twist-chair
TC



twist-boat
TB



planar
P

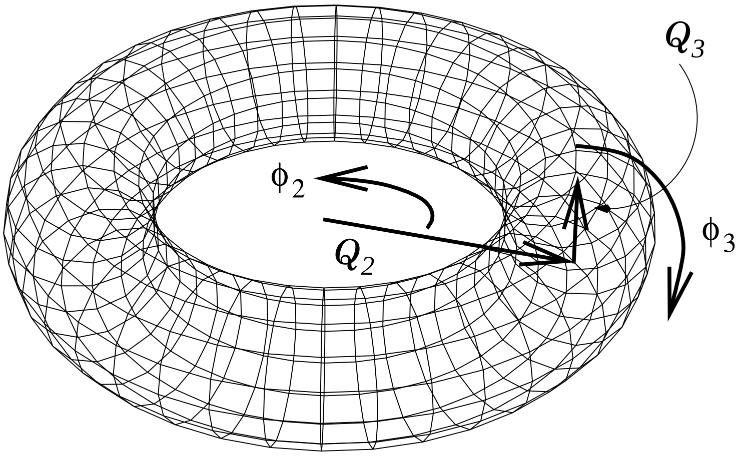


sofa
S

Occassionally the intermediate forms twist-sofa and boat-sofa are mentioned in literature.

7-Membered rings — Cremer & Pople parameters

The Cremer & Pople parameters Q_2, Q_3, ϕ_2 and ϕ_3 can be mapped on a torus:



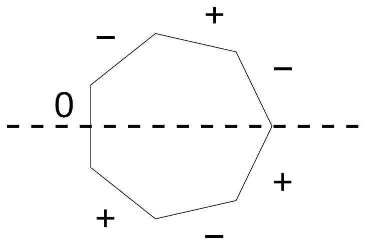
Characteristic parameters for the conformations mentioned earlier are:

conformation	ϕ_2	ϕ_3	Q_2	Q_3
C	$\frac{n6}{28} \times 360°$	$\frac{14+2n}{28} \times 360°$	—	—
B	$\frac{n6}{28} \times 360°$	$\frac{14+2n}{28} \times 360°$	—	0
TC	$\frac{3+n6}{28} \times 360°$	$\frac{15+2n}{28} \times 360°$	—	—
TB	$\frac{1+n6}{28} \times 360°$	$\frac{19+2n}{28} \times 360°$	—	—
S	$\frac{n6}{28} \times 360°$	$\frac{2n}{28} \times 360°$	—	—
P	—	—	0	0

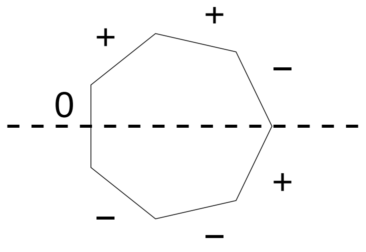
(parameters ϕ_2 and ϕ_3 are modulo $360°$;
n is equal for ϕ_2 and ϕ_3 in one conformation)

7-Membered rings — Asymmetry parameters and pseudorotation

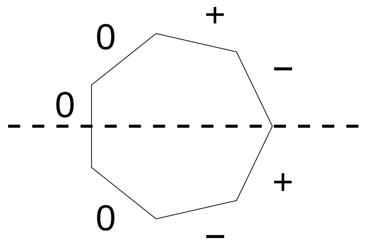
Asymmetry parameters can not be used to distinguish all main conformations. Inspection of the signs of the individual torsion angles is necessary for a complete identification.



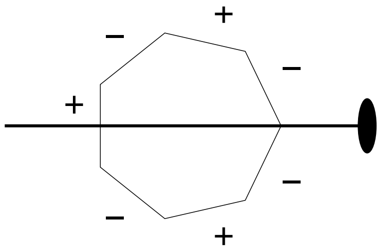
chair
C



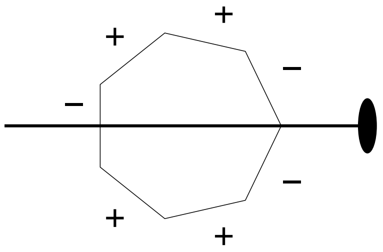
boat
B



sofa
S



twist-chair
TC



twist-boat
TB

In the planar form all asymmetry parameters are zero.

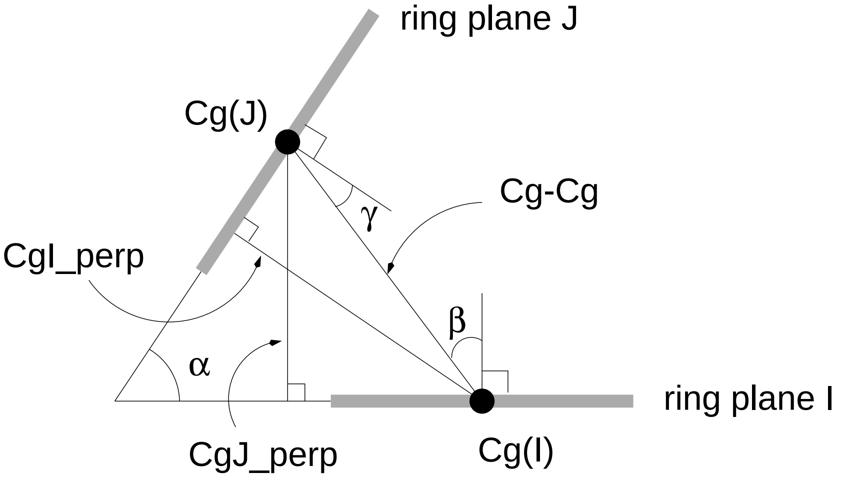
Several pseudorotation pathways spiral along the “puckering torus”. Examples of pseudorotation pathways are

$$^{1}\mathrm{C} \rightarrow {}^{4}\mathrm{TC} \rightarrow \mathrm{C}_7 \rightarrow \mathrm{TC}_3 \rightarrow {}^{6}\mathrm{C} \rightarrow \text{etc.}$$

$$^{1}\mathrm{B} \rightarrow {}^{2}\mathrm{TB} \rightarrow \mathrm{B}_3 \rightarrow \mathrm{TB}_4 \rightarrow {}^{5}\mathrm{B} \rightarrow \text{etc.}$$

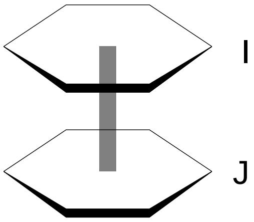
Analysis of short ring interactions

The following parameters are tabulated in the geometrical analysis of ring$\cdots$ring interactions:



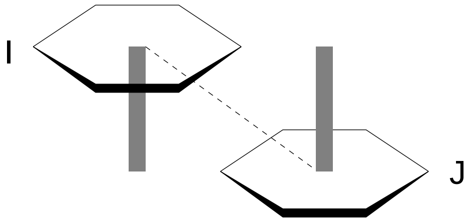
where Cg is the geometrical centre of the ring. The geometry is analyzed when Cg-Cg < 6.0 Å and $\beta < 60°$.

Examples



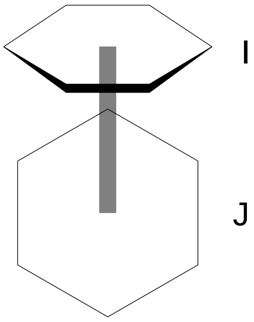
$$\alpha = \beta = \gamma = 0°$$

$$\text{Cg}-\text{Cg} = \text{CgI\_perp} = \text{CgJ\_perp}$$



$$0° = \alpha < \beta = \gamma$$

$$\text{Cg}-\text{Cg} > \text{CgI\_perp} = \text{CgJ\_perp}$$



$$\alpha = \gamma = 90°$$

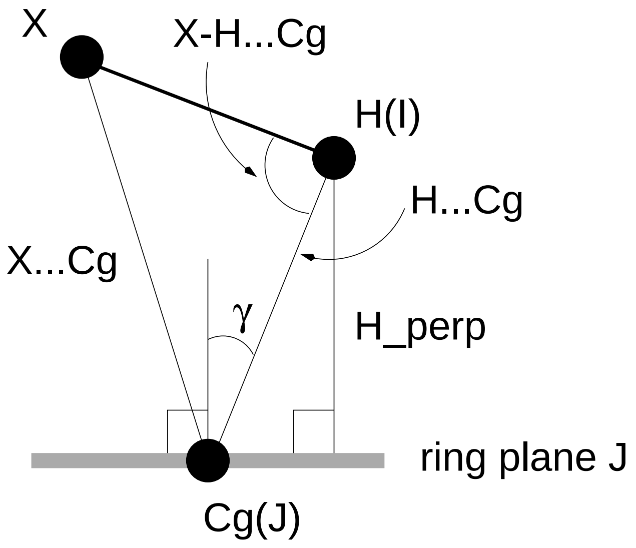
$$\beta = 0°$$

$$\text{Cg}-\text{Cg} = \text{CgJ\_perp}$$

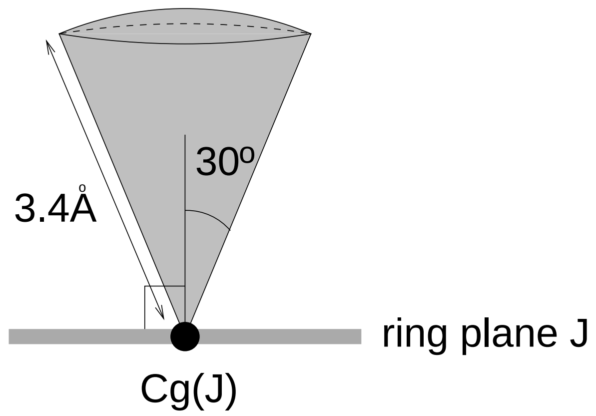
$$\text{CgI\_perp} = 0\text{Å}$$

Analysis of X—H$\cdots\pi$ interactions (1)

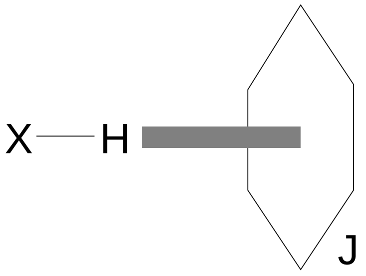
To identify potential X—H$\cdots\pi$ hydrogen bonds, the following parameters are tabulated:



Parameters are only tabulated if H...Cg < 3.4 Å and $\gamma < 30^\circ$.



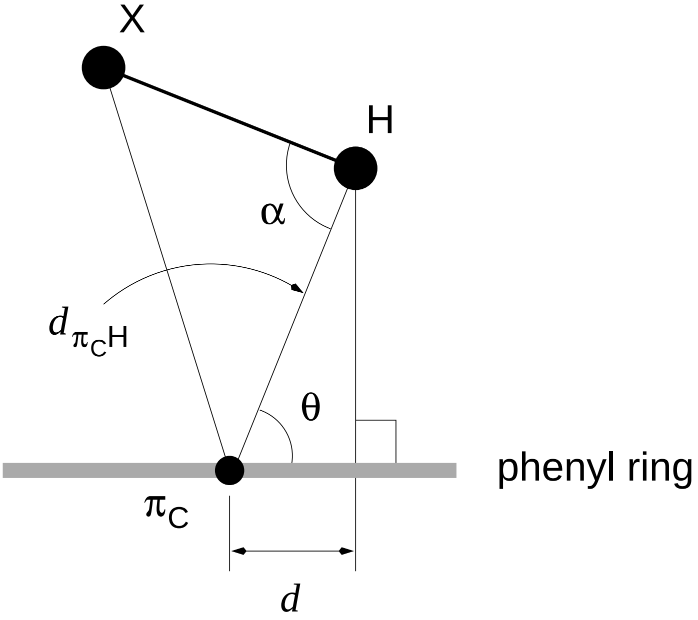
For an ideal linear X—H$\cdots\pi$ interaction, the following values are found:



$$X{-}H...Cg = 180^\circ$$
$$\gamma = 0^\circ$$
$$H\_perp = H...Cg$$

Analysis of X—H$\cdots\pi$ interactions (2)

Malone et al. introduced a classification of X—H$\cdots\pi$ (phenyl) interactions based on the following geometric parameters:



The relations to the parameters tabulated in PLATON are

$$\alpha = \text{X}-\text{H...Cg}$$

$$\theta = 90^\circ - \gamma$$

$$d_{\pi_C \text{H}} = \text{H...Cg}$$

$$d = \sqrt{(\text{H...Cg})^2 - (\text{H\_perp})^2} \qquad \text{or}$$

$$d = (\text{H...Cg}) \sin\gamma$$

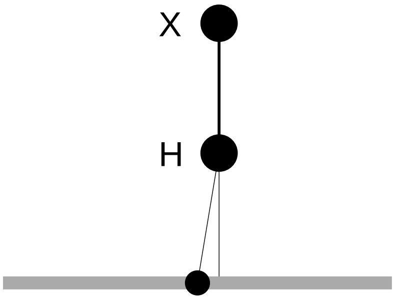
6 interaction types are distinguished on the basis of geometry.

N.B. The numbers given in the following table are only valid for X—H$\cdots\pi$ interaction with phenyl π systems.

Malone, Murray, Charlton, Docherty & Lavery. *J. Chem. Soc., Faraday Trans.* (1997) 3429–3436

Analysis of X—H$\cdots\pi$ interactions (3)

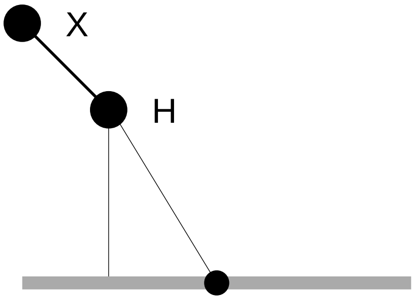
type I



$d_{\pi_\mathrm{C}\mathrm{H}} \leq 3.05\text{Å};\ d \leq 0.5\text{Å}$

$\theta \geq 53°;\ 150° \leq \alpha \leq 180°$

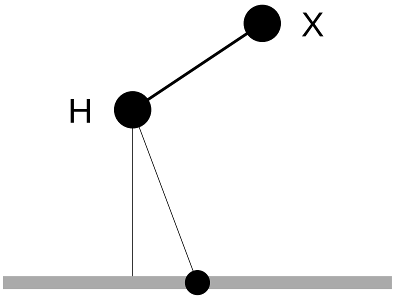
type II



$d_{\pi_\mathrm{C}\mathrm{H}} \leq 3.05\text{Å};\ 0.5\text{Å} < d \leq 1.4\text{Å}$

$\theta \geq 53°;\ 150° \leq \alpha \leq 180°$

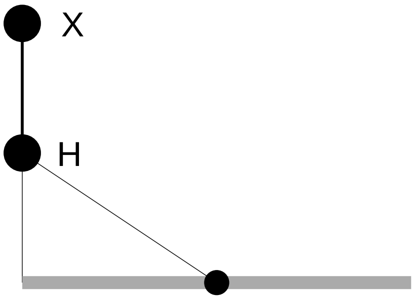
type III



$d_{\pi_\mathrm{C}\mathrm{H}} \leq 3.05\text{Å};\ d \leq 1.4\text{Å}$

$\theta \geq 53°;\ \alpha < 150°$

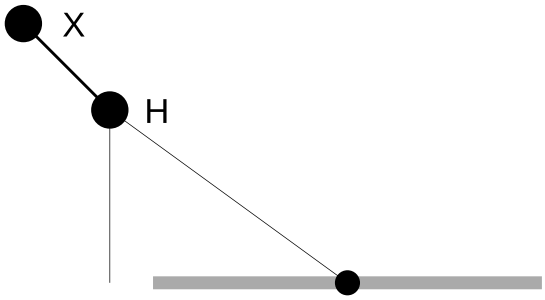
type IV



$d_{\pi_\mathrm{C}\mathrm{H}} \leq 3.05\text{Å};\ 1.4\text{Å} \leq d \leq 1.5\text{Å}$

$40° \leq \theta \leq 60°;\ 130° \leq \alpha \leq 150°$

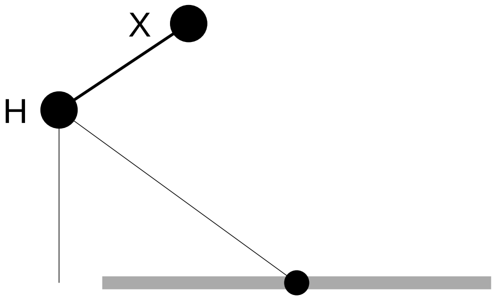
type V



$d_{\pi_\mathrm{C}\mathrm{H}} \leq 4.0\text{Å};\ d > 1.4\text{Å}$

$\theta \leq 90°;\ 90° \leq \alpha \leq 180°$

type VI



$d_{\pi_\mathrm{C}\mathrm{H}} \leq 4.0\text{Å};\ d > 1.4\text{Å}$

$\theta \leq 90°;\ \alpha \geq 90°$

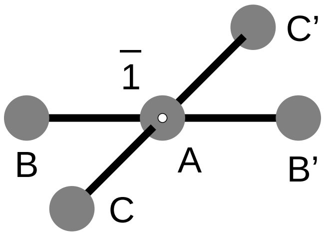
(Only valid when π_C is a phenyl ring.)

Geometry and symmetry

Crystallographic symmetry can force geometric parameters to adopt specific values. In such cases, s.u.'s are not always reported.

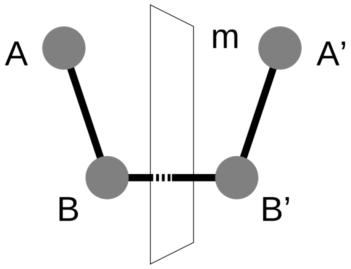
Examples

- AB_2C_2, located on an inversion centre



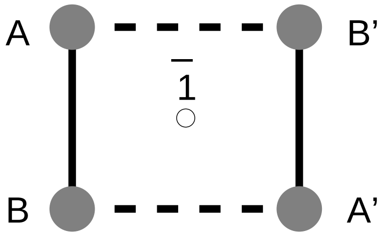
$$\phi_{B-A-B'} = 180^\circ$$
$$\phi_{C-A-C'} = 180^\circ$$
$$\phi_{B-A-C'} = 180^\circ - \phi_{B-A-C}$$

- A—B—B′—A′, with a mirror plane bisecting B–B



$$\tau_{A-B-B'-A'} = 0^\circ$$

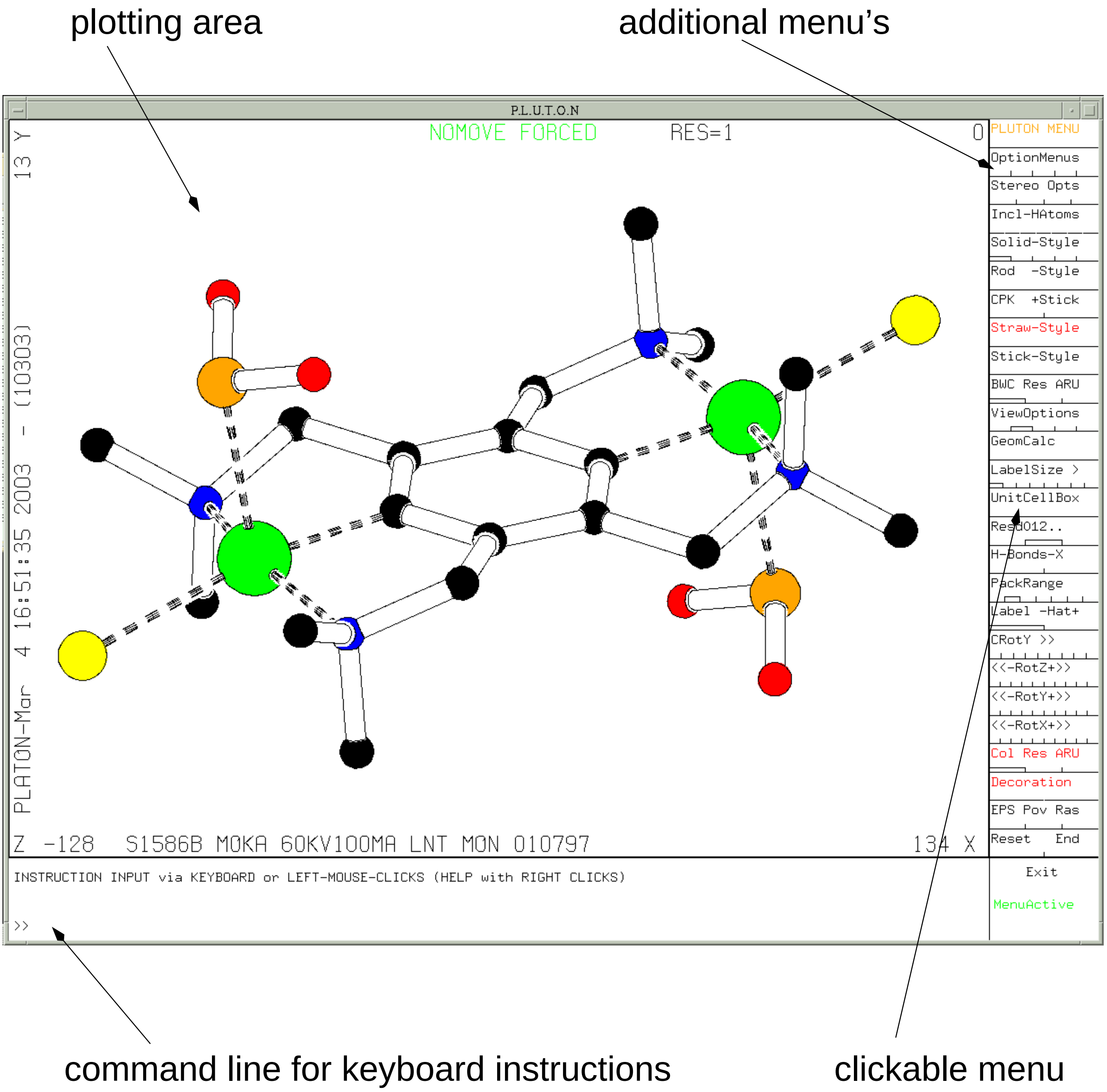
- A—B dimer, related by an inversion centre



A, A′, B and B′ lie exactly in one plane

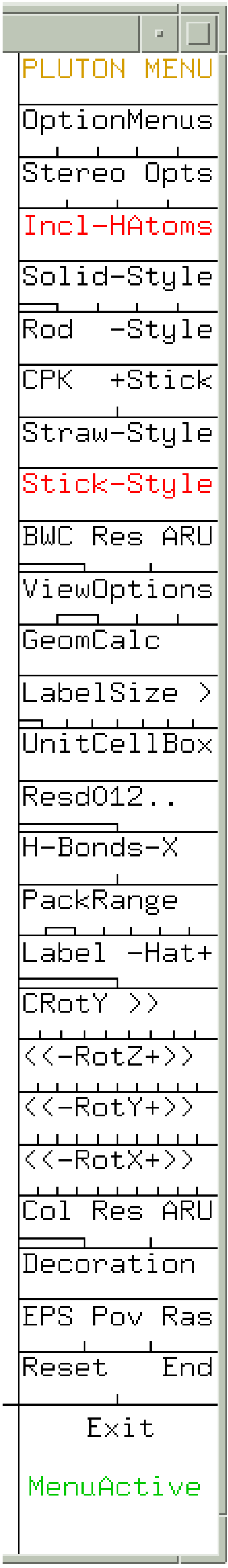
N.B. Atomic arrangements that do not fit the symmetry exactly can be disordered over such a symmetry site. This will be reflected in the displacement parameters. The symmetry-restrained values of the geometrical parameters have no physical meaning in these cases.

PLUTON

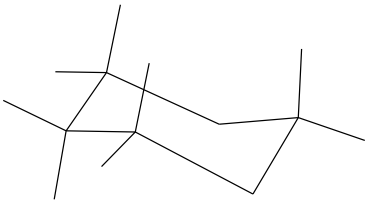


```
INSTRUCTION INPUT via KEYBOARD or LEFT-MOUSE-CLICKS (HELP with RIGHT CLICKS)
```

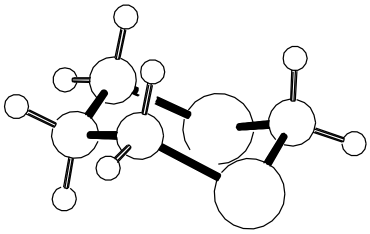
PLUTON — plot styles (1)



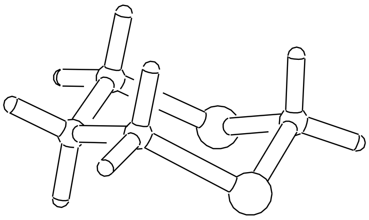
PLUTON main menu



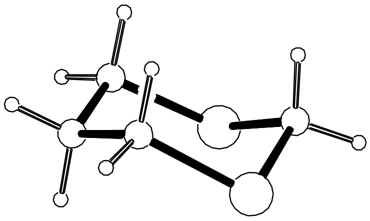
stick



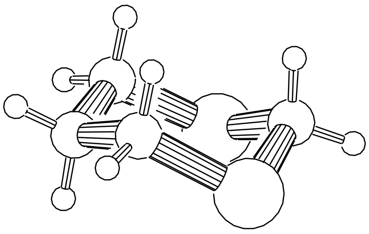
solid 1



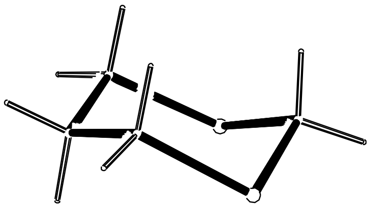
straw



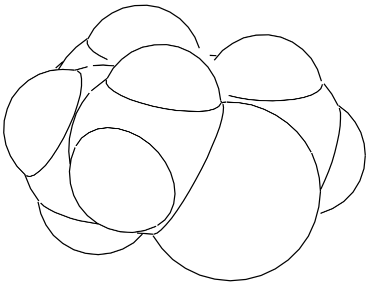
solid 3



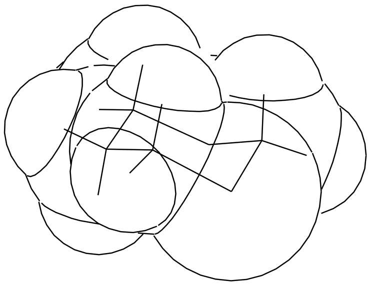
rod



solid 5



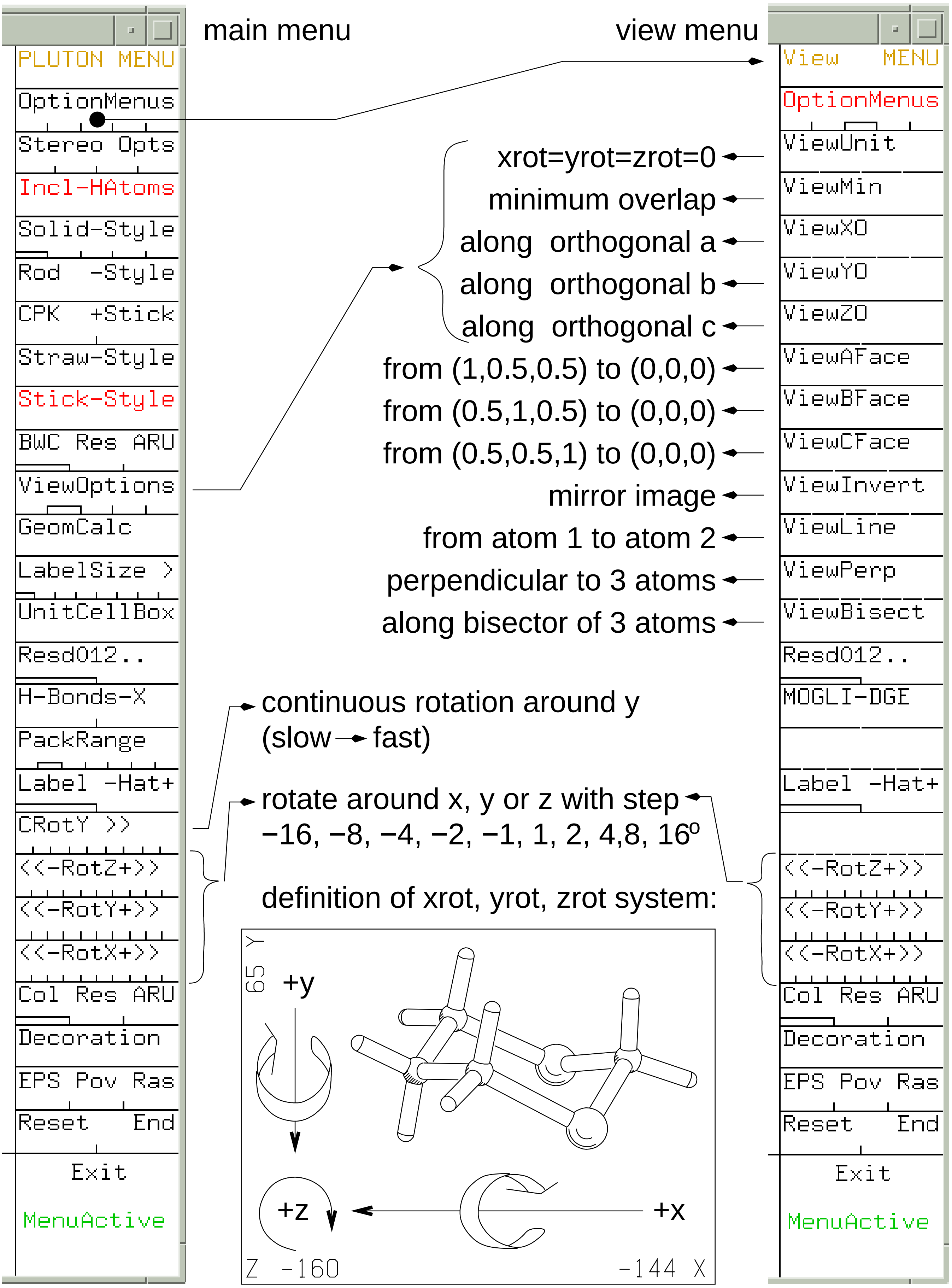
CPK



CPK+stick

BWC toggle hatching, COL toggle colour options: atom type, residue type or ARU

PLUTON — orientation



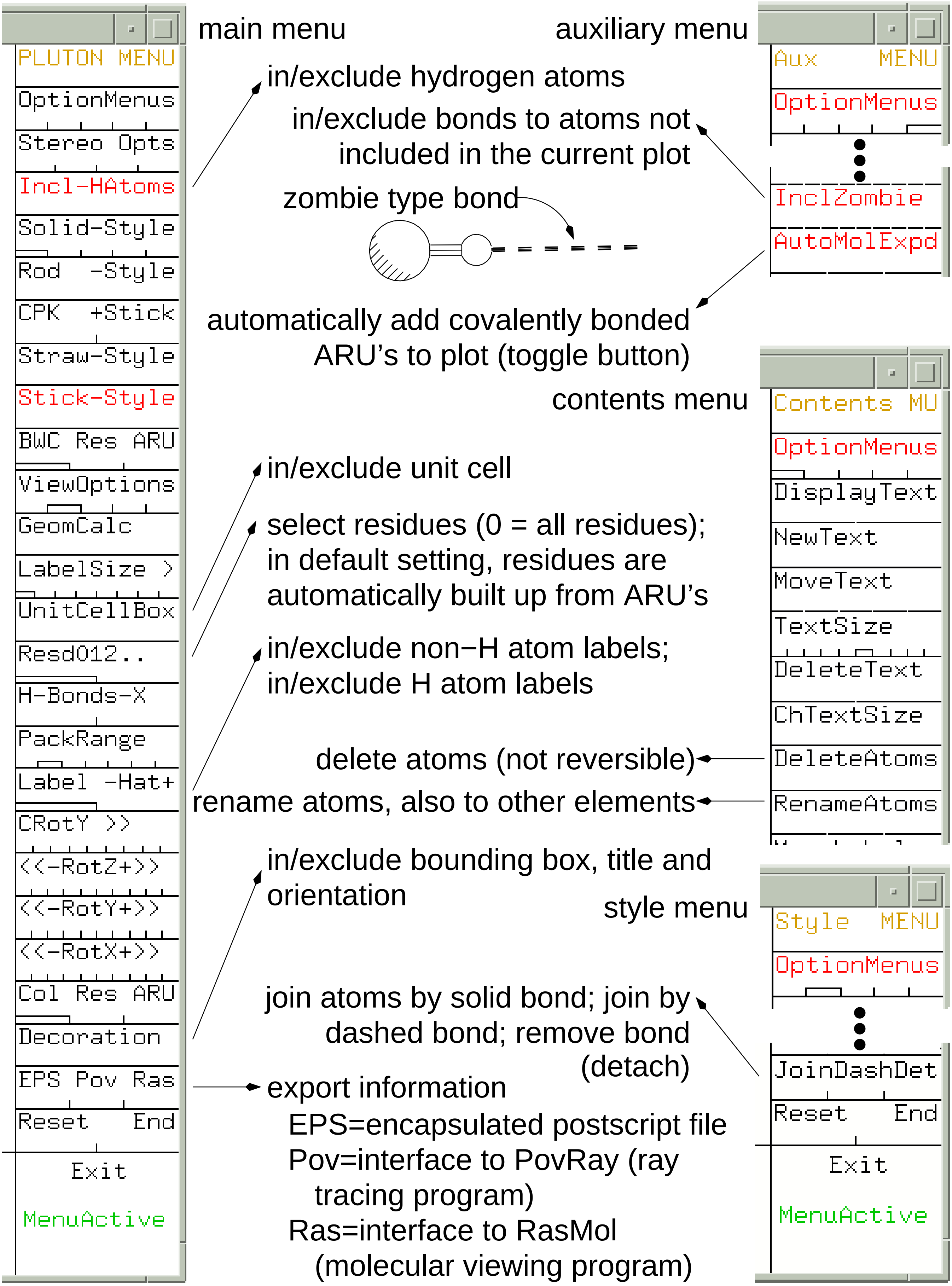
xrot=yrot=zrot=0
minimum overlap
along orthogonal a
along orthogonal b
along orthogonal c
from (1,0.5,0.5) to (0,0,0)
from (0.5,1,0.5) to (0,0,0)
from (0.5,0.5,1) to (0,0,0)
mirror image
from atom 1 to atom 2
perpendicular to 3 atoms
along bisector of 3 atoms

continuous rotation around y (slow → fast)

rotate around x, y or z with step −16, −8, −4, −2, −1, 1, 2, 4,8, 16°

definition of xrot, yrot, zrot system:

PLUTON — select plot items



in/exclude hydrogen atoms

in/exclude bonds to atoms not included in the current plot

zombie type bond

automatically add covalently bonded ARU’s to plot (toggle button)

in/exclude unit cell

select residues (0 = all residues); in default setting, residues are automatically built up from ARU’s

in/exclude non−H atom labels;
in/exclude H atom labels

delete atoms (not reversible)

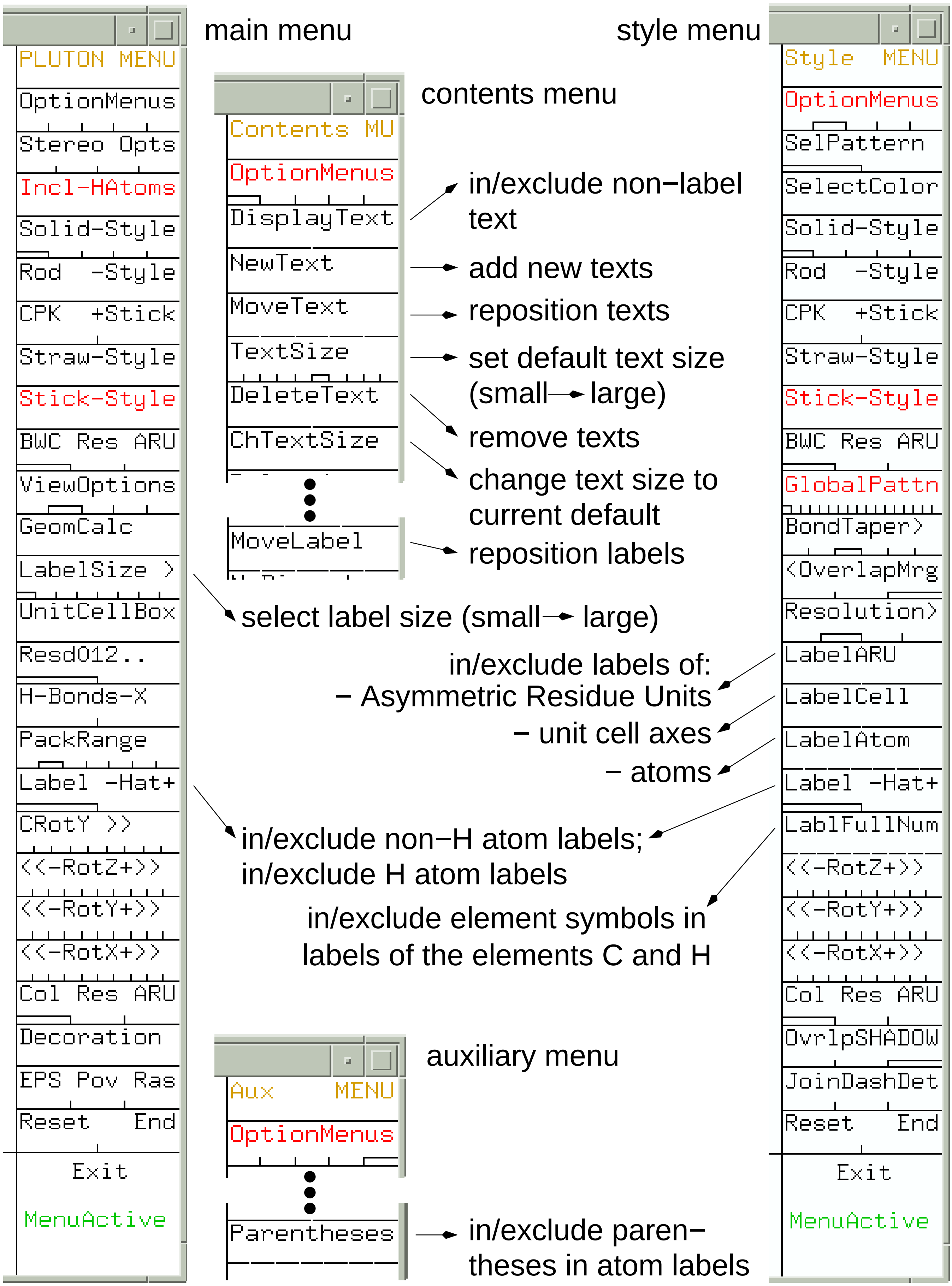
rename atoms, also to other elements

in/exclude bounding box, title and orientation

join atoms by solid bond; join by dashed bond; remove bond (detach)

export information
EPS=encapsulated postscript file
Pov=interface to PovRay (ray tracing program)
Ras=interface to RasMol (molecular viewing program)

PLUTON — labels and texts



main menu

style menu

contents menu

in/exclude non−label text

add new texts

reposition texts

set default text size (small → large)

remove texts

change text size to current default

reposition labels

select label size (small → large)

in/exclude labels of:
− Asymmetric Residue Units
− unit cell axes
− atoms

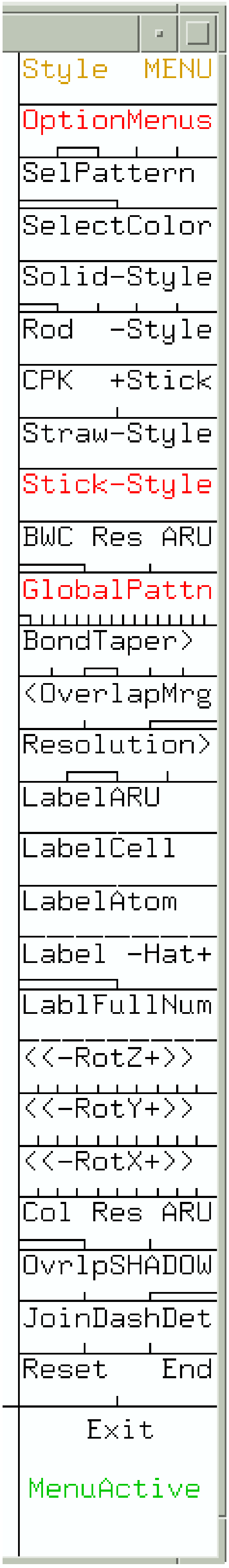
in/exclude non−H atom labels;
in/exclude H atom labels

in/exclude element symbols in labels of the elements C and H

auxiliary menu

in/exclude parentheses in atom labels

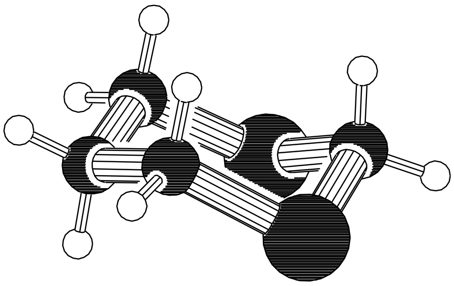
PLUTON — plot styles (2)



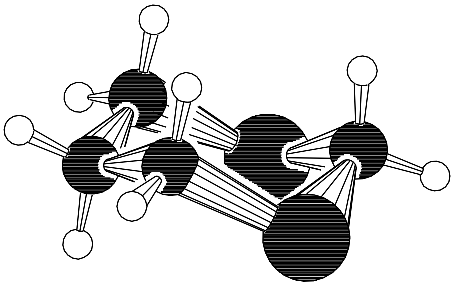
style menu

bond tapering

small

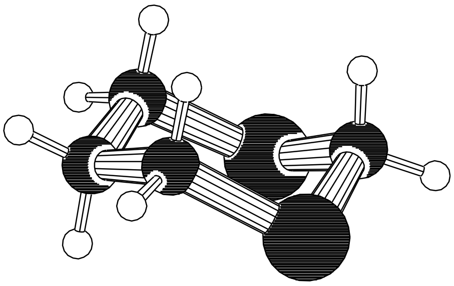


large

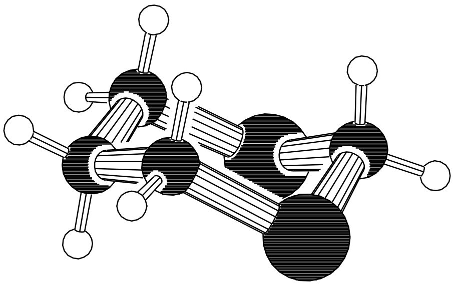


overlap margin

small

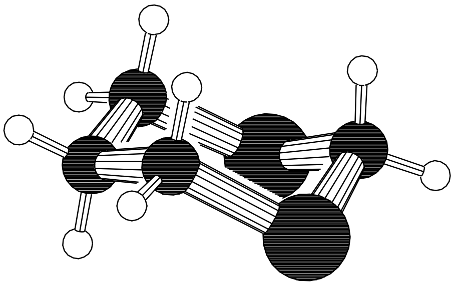


large

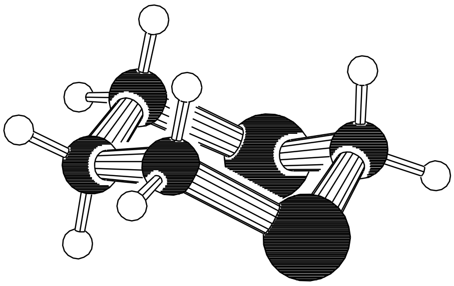


overlap shadow

small

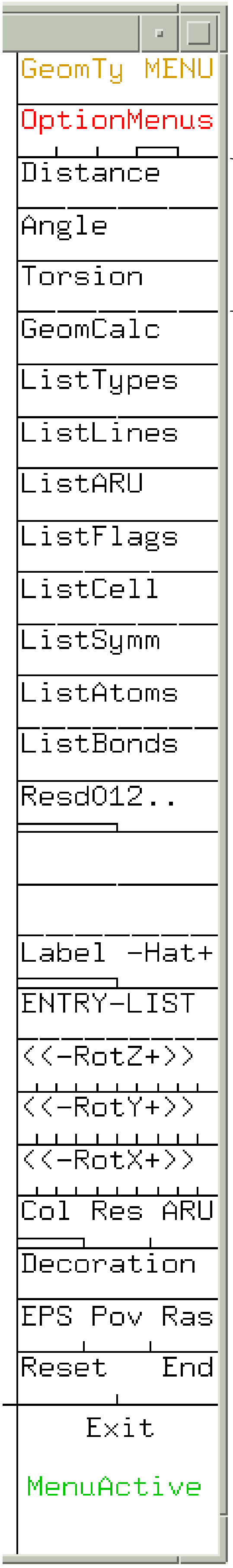


large



determines accuracy of begin and end point of a plotted line, at the expence of computing time

PLUTON — geometry calculations and lists



geometry/type menu

Distance, Angle, Torsion → calculate geometric parameters by clicking on atoms

GeomCalc → calculate geometry involving a pivot atom

ListTypes → list atom types (plot radius, colour, hatching)

ListLines → list bond types (nr of lines, radius, dashing)

ListARU → list Asymmetric Residue Units in the plot

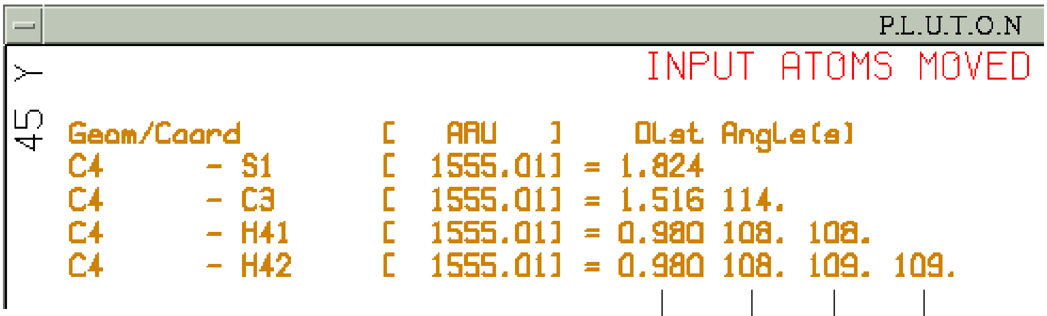
ListFlags → list atomic flags

ListCell → list cell parameters

ListSymm → list symmetry operations of the space group

ListAtoms → list atomic co−ordinates

ListBonds → list bonds



```
                                          INPUT ATOMS MOVED
Geom/Coord         [   ARU    ]    Dist  Angle(s)
C4        - S1     [  1555.01] = 1.824
C4        - C3     [  1555.01] = 1.516 114.
C4        - H41    [  1555.01] = 0.980 108.  108.
C4        - H42    [  1555.01] = 0.980 108.  109.  109.
```

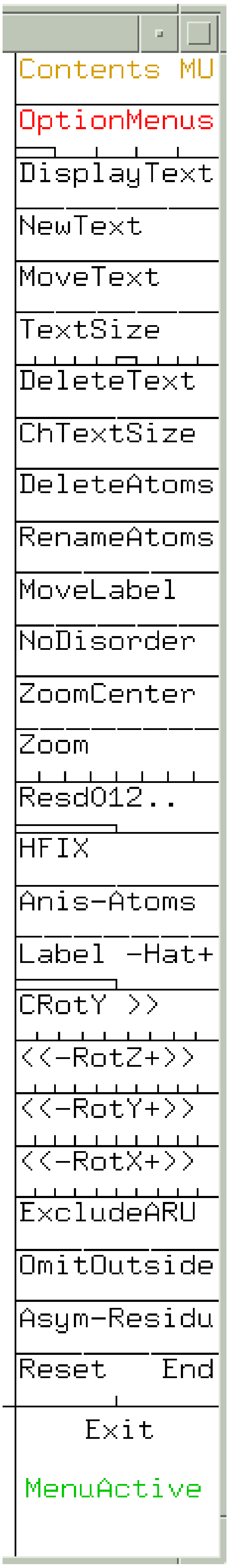
distance C4−H42

angle H42−C4−S1

angle H42−C4−C3

angle H42−C4−H41

PLUTON — miscellaneous options



contents menu

auxiliary menu

remove clickable menu's
undo with MENU ON command

distance for point−projected plot
close → far

visualization of isotropic displacement
(only for shelxl.res formatted files)

refresh plot after each change
(toggle button)

in/exclude disorder components

set zoom center

set zoom scale

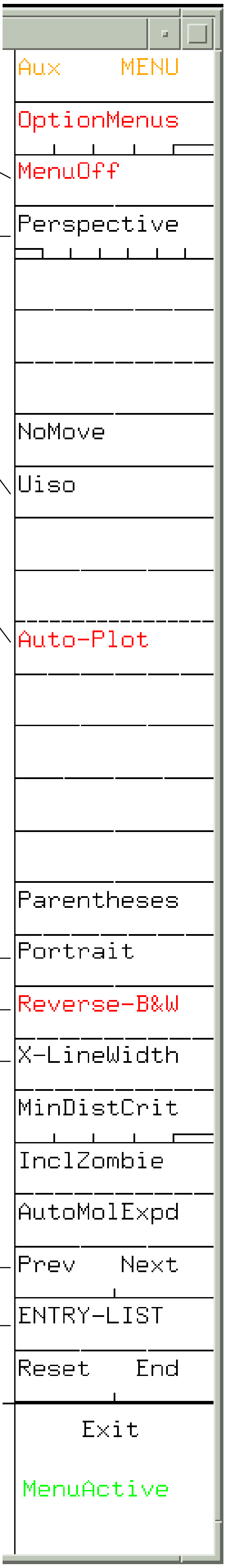
portrait/landscape mode

toggle background colour black/white

small/broad plot lines

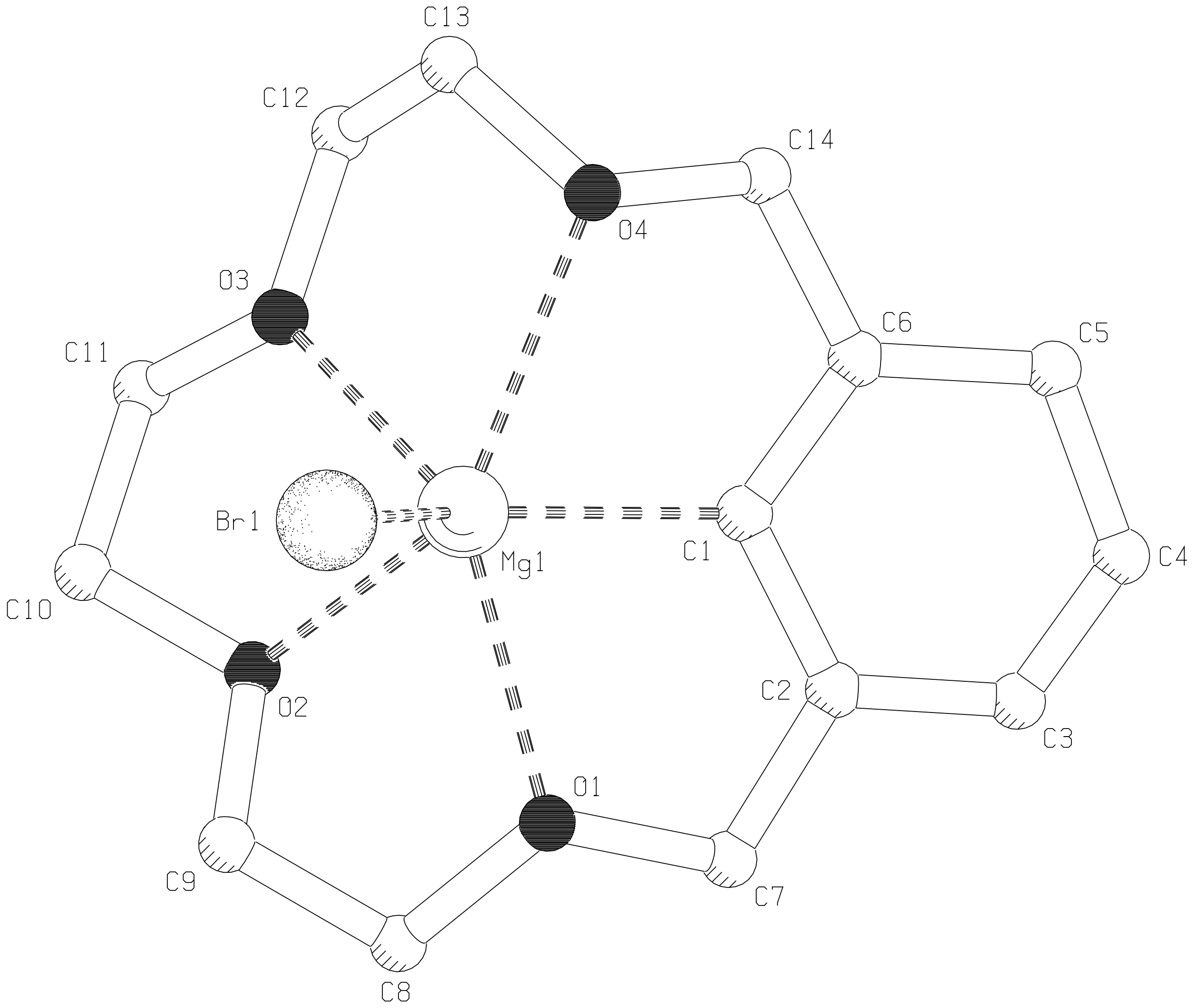
previous/next structure in CSD file

list structure names in CSD file



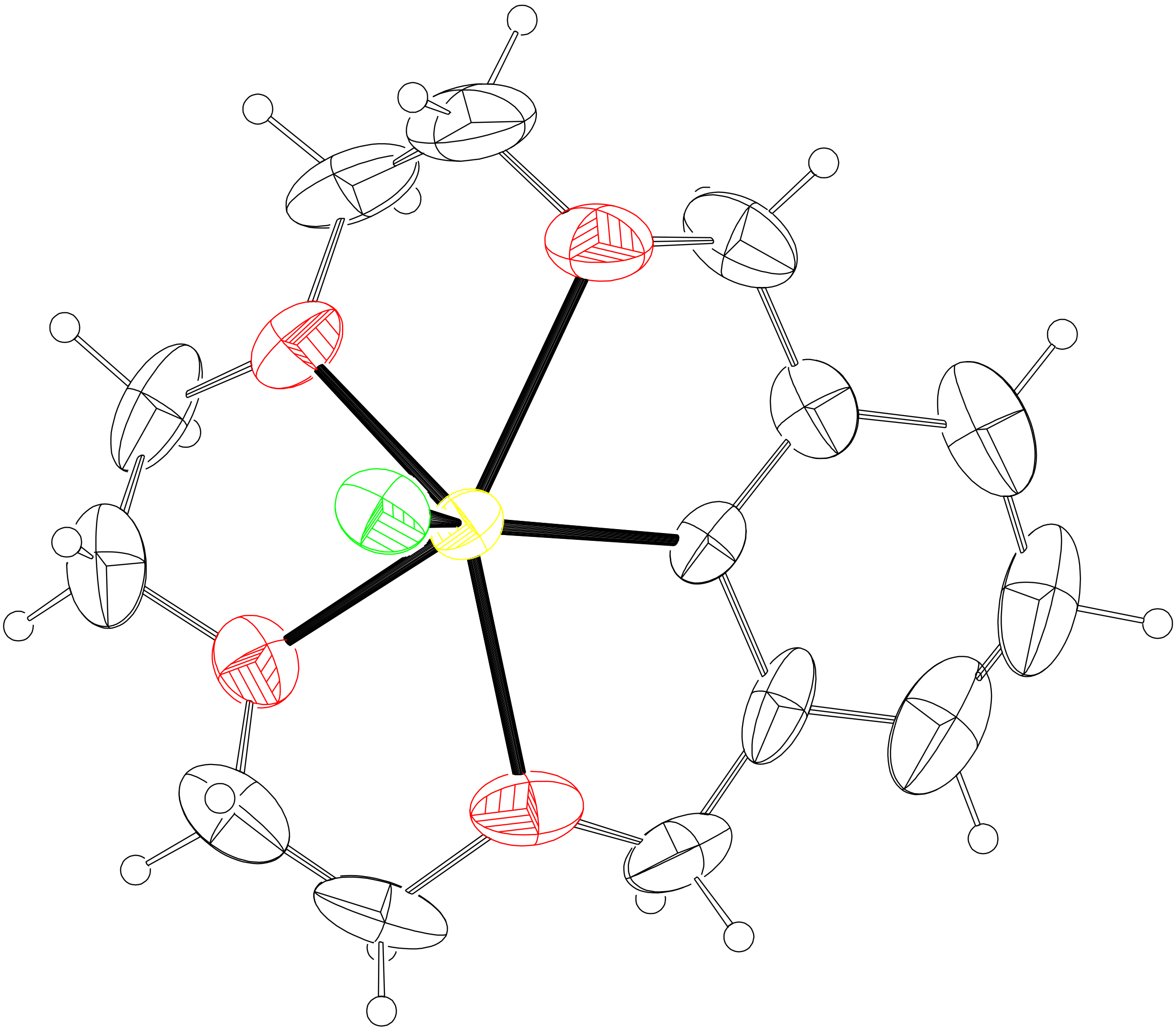
Example 2

2-(bromomagnesio)-1,3-xylyl-15-crown-4



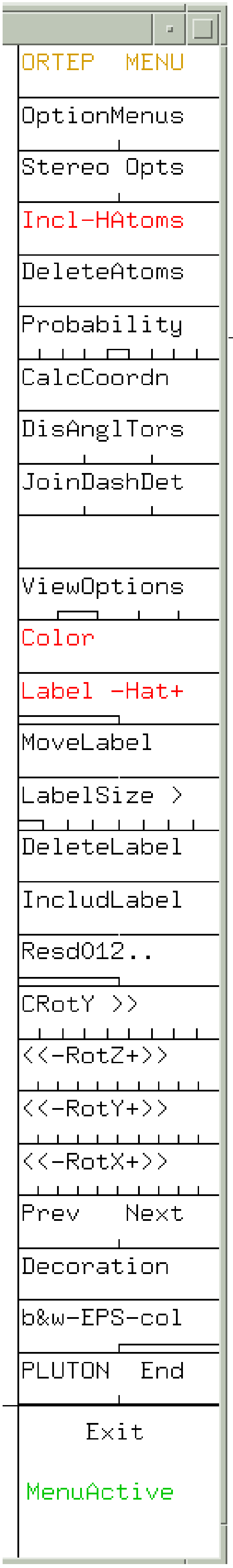
ORTEP of example 2

Anisotropic displacement parameters can be visualized with a displacement ellipsoid plot, also known as ORTEP (Oak Ridge Thermal Ellipsoid Plot).



drawn at 30% probability level, meaning that the chance of finding the atomic nucleus within the plotted ellipsoid is 30%.

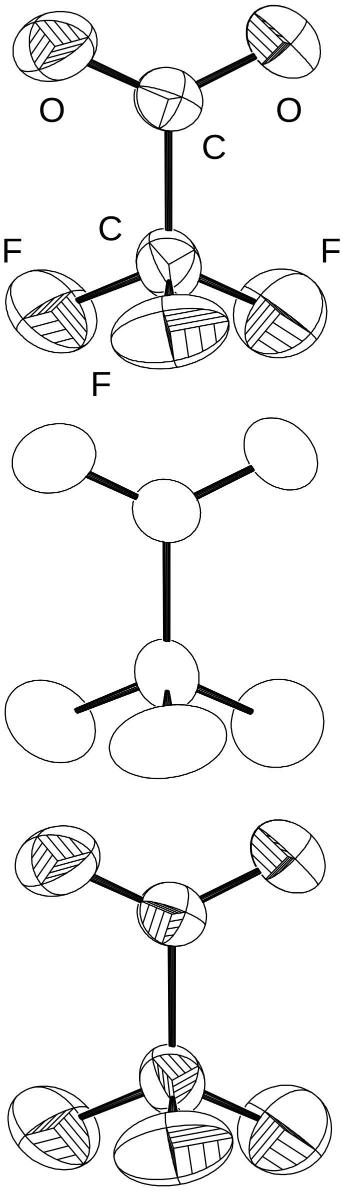
ORTEP — plot styles

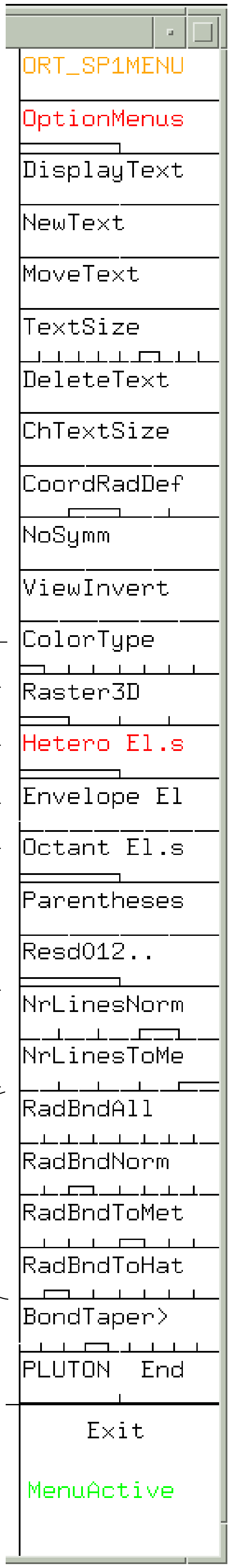


set probability level (10% → 90%)

interface to RasterGraphics

change color
type of
element





write encapsulated postscript
file; black & white or colour

set number of lines in normal
bonds or bonds to metals

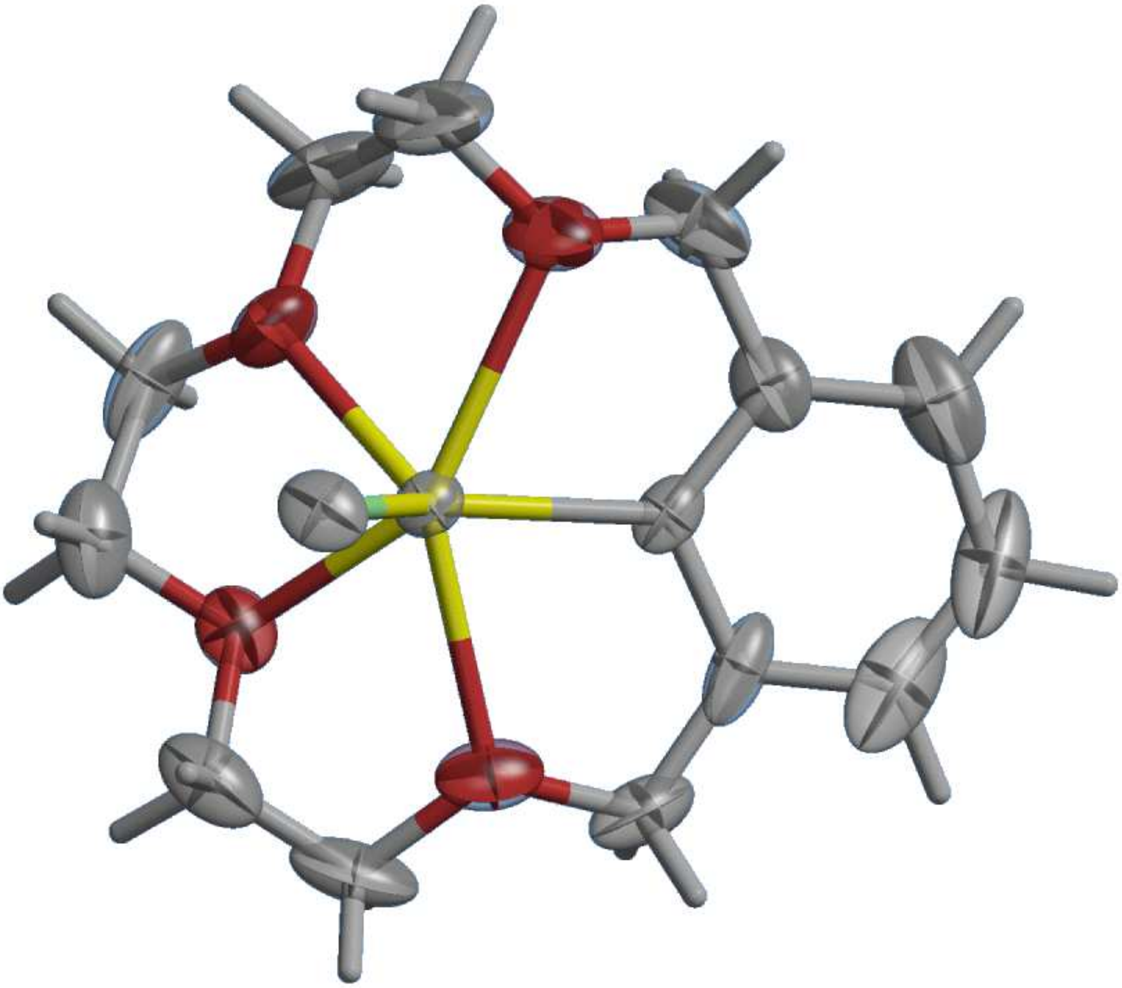
set plot radii of all bonds, normal bonds,
bonds to metal atoms, bonds to H atoms

other options as in PLUTON

ORTEP — RasterGraphics example

The ORTEP routine in PLATON contains an interface to RasterGraphics. Four styles are available:

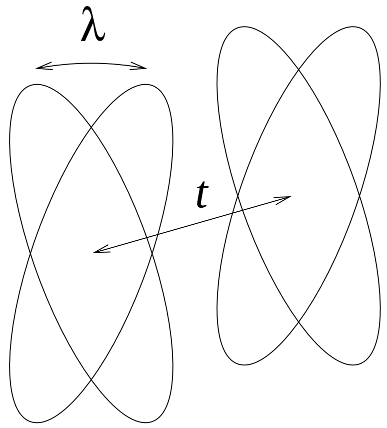
1 sticks
2 ellipsoids
3 intersecting ellipses
4 combination of ellipsoids with intersecting ellipses



Example 2, in RasterGraphics style 4, drawn at 30% probability level.

Rigid body motion analysis—the TLS model

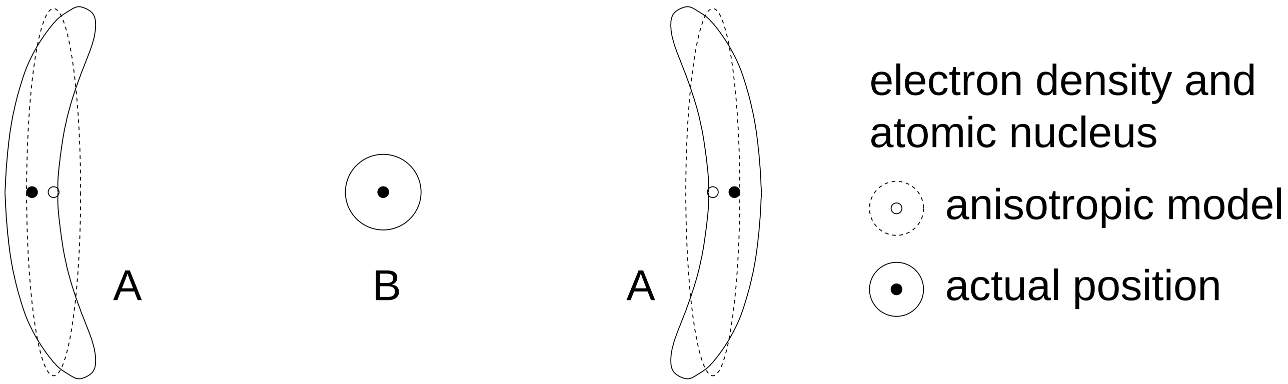
Atomic displacement is influenced by intramolecular forces. If the molecule can be considered as a rigid body, the atomic displacements can be described as a combination of a translation and a libration of the rigid body.



Librational and translational parameters are fitted to the observed U_{ij}^{obs} with a least-squares procedure. With the obtained translation (T), libration (L) and correlation (S) tensors, the TLS model values U_{ij}^{calc} can be calculated. The validity of the TLS model can be assessed by comparing U_{ij}^{obs} and U_{ij}^{calc} via R-values.

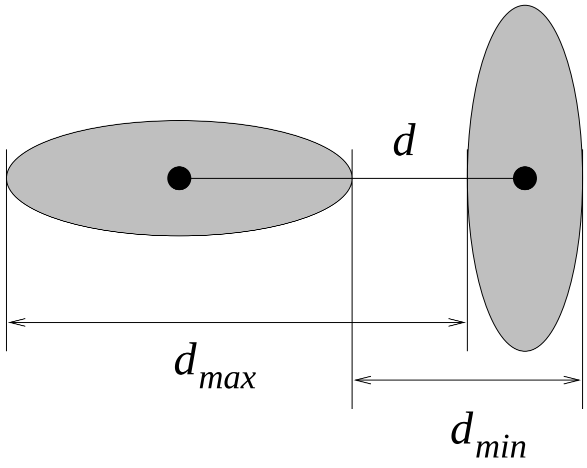
The libration tensor can be used to correct geometrical parameters for the effect of libration.

Example: The system A—B—A shows a large libration around B. In the anisotropic model, $d_{\mathsf{A-B}}$ is therefore too short.



Rigid bond test

Due to their high force constants, bonds in crystal structures can be considered to be rigid. Two bonded atoms will therefore have equal vibration amplitudes in the direction of the bond.



$$d_{min} \ll d \ll d_{max}$$

The Hirshfeld rigid bond test evaluates Δ_{ij}:

$$\begin{aligned}\Delta_{ij} &= \text{vibration}_{i\rightarrow j} - \text{vibration}_{i\leftarrow j} \\ &= \sqrt{\langle u^2_{i\rightarrow j}\rangle} - \sqrt{\langle u^2_{i\leftarrow j}\rangle}\end{aligned}$$

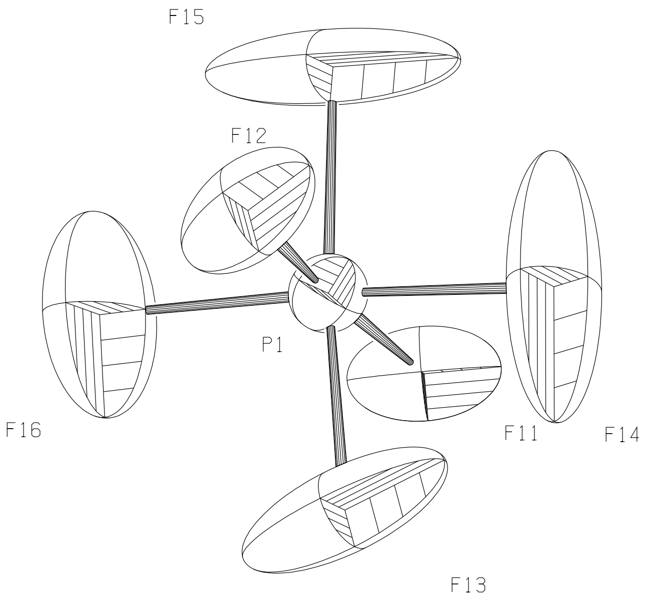
Δ_{ij} is expected to be 0 Å. PLATON gives a warning if $\Delta_{ij} > 5\sigma(\Delta_{ij})$

N.B. The refinement program SHELXL has an option to restrain Δ_{ij} to 0 (DELU instruction). When this restraint is applied, even at mild force, the Hirshfeld test looses its validity as a figure of merit. Wether the restraint is applied or not can not be concluded from the positional or displacement parameters. It should therefore always be reported in the "Experimental".

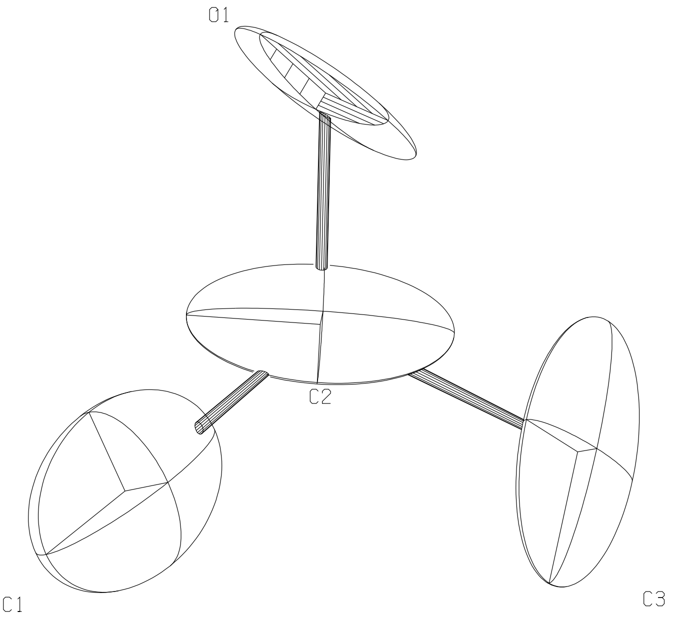
Using ORTEP plots as a figures of merit

Displacement ellipsoid plots can often be used as an indication of the quality of a crystal structure determination.

Examples:



High but physically realistic anisotropicity.



Unrealistic anisotropicity.

The plots are drawn at 30% probability level.

Intramolecular rotation barriers

Libration of a rigid group within a molecule is determined by the potential energy curve of the rigid group rotation.

The estimate of the libration amplitude obtained by TLS analysis can be used to calculate the libration force constant f:

$$f = \frac{kT}{\langle \phi^2 \rangle}$$

where ϕ is the libration amplitude, k is Boltzmann's constant, and T is the data collection temperature.

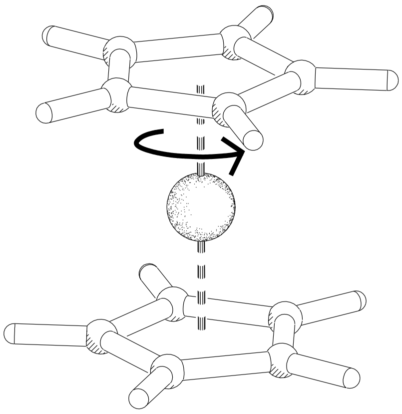
Based on a sinusoidal potential energy funtion with n maxima, the height B of the energy barrier can be calculated with

$$B = \frac{2RT}{n^2 \langle \phi^2 \rangle}$$

This formula is an approximation for small libration amplitudes.

N.B. The intramolecular force constant and energy barrier can be severely influenced by intermolecular interactions.

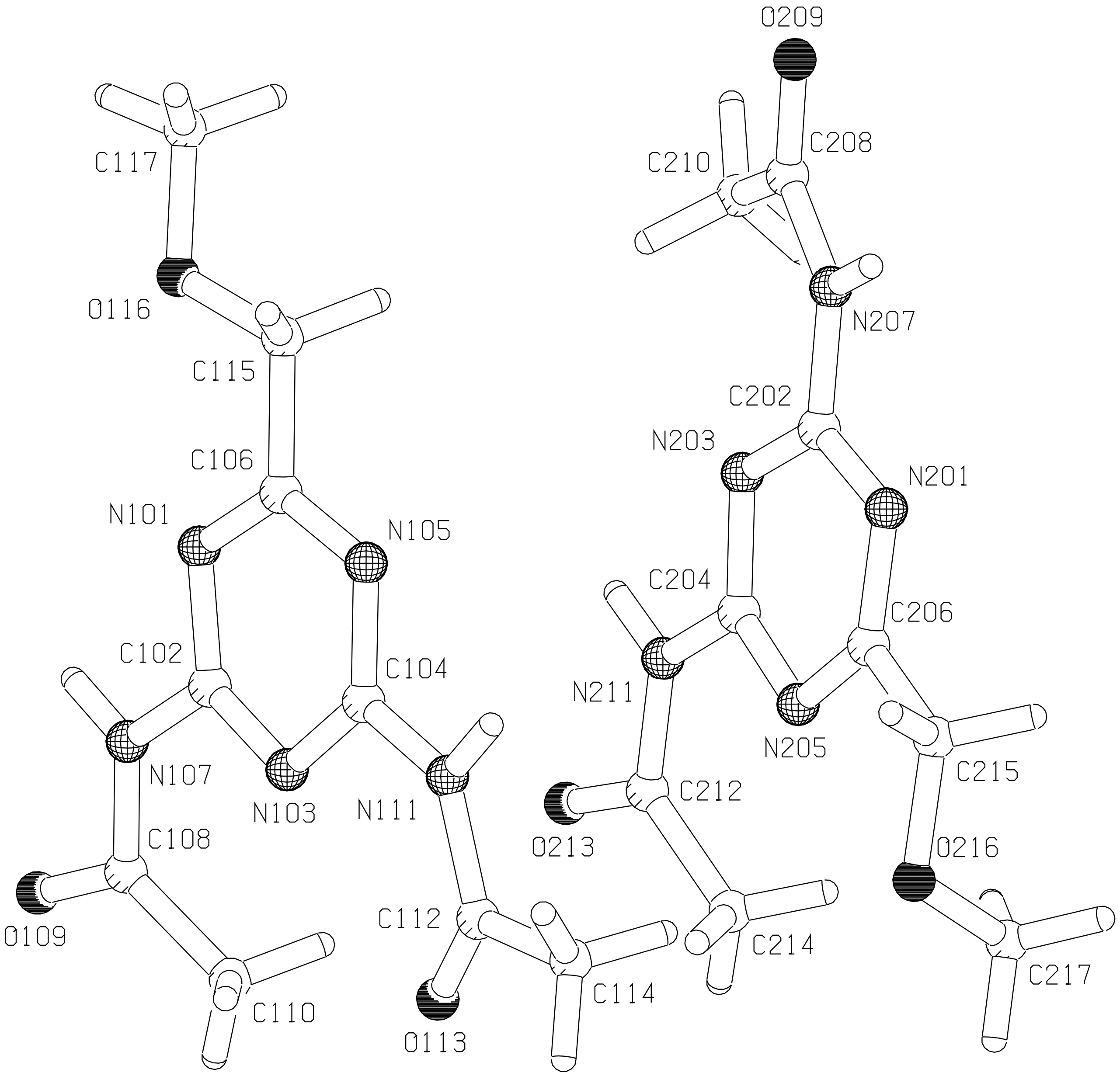
Example:



With this technique the cyclopentadiene rotation barrier around the (local) 5-fold axis in ferrocene was found to be in the range 7.6(2)–9(2) kJ mol^{-1} at 101K; in good agreement with results obtained by other techniques.

Example 3

2,4-di(acetylamino)-6-methoxymethyl-s-triazine



The asymmetric unit contains two independent molecules which differ in conformation.

Intra- and intermolecular contacts

The PLATON option "CALC INTER" performs a systematic search for all contacts of atoms I with J satisfying

$$d_{I-J} = r_I + r_J + t$$

with r_I the contact radius of atom I; t is the tolerance (default $t =$ 0.2 Å).

The search includes all symmetry-related atoms.

Intramolecular contacts are only reported if I and J are separated by more than three bonds. Contacts involving atoms of type "D/A" are always reported.

Intramolecular contacts are labelled with "Intra".

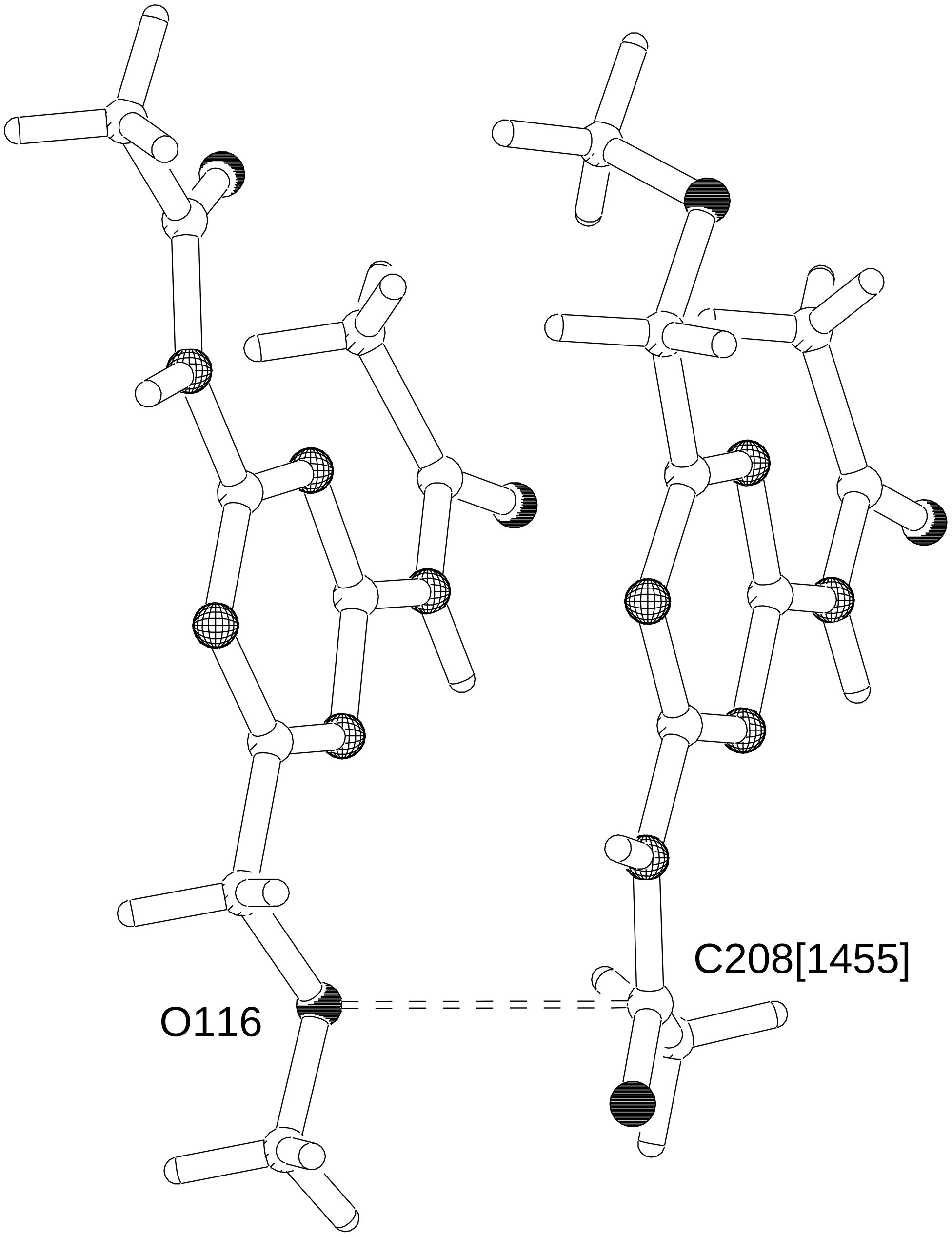
Short distances are marked by symbols:

$$\begin{array}{rl} \text{``}<\text{''} & \text{if } d_{I-J} - r_I - r_J < 0.0 \text{ Å} \\ \text{``}<<\text{''} & \text{if } d_{I-J} - r_I - r_J < -0.2 \text{ Å} \end{array}$$

If an atom X, bonded to atom I makes an angle $X - I \cdots J > 100°$, the label of atom X and the value of the angle are reported.

Contact analysis is performed for each residue. A table listing all neighbouring residues and the number of atomic contacts involved concludes the list of atom pairs.

Example 3 — Intermolecular contact



Hydrogen bonds

PLATON refers to a hydrogen bond as

donor—hydrogen$\cdots$acceptor

$$D\text{–}H \cdots A$$

Potential donor and acceptor atoms are N, O, F, Cl, S, Br. C—H$\cdots A$ interactions are included at the end of the list of hydrogen bonds.

The geometrical criteria for accepting a close contact as hydrogen bond are

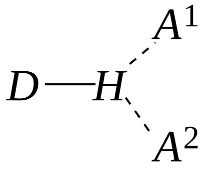
$$d_{D \cdots A} < r_D + r_A + 0.5$$

$$d_{H \cdots A} < r_H + r_A - 0.12$$

$$\phi_{D-H \cdots A} > 100^\circ$$

Since hydrogen positions are normally less well determined than D or A positions, $d_{D \cdots A}$ is usually a more reliable indicator for the hydrogen bond strength than $d_{H \cdots A}$.

A bifurcated hydrogen bond is formed when the H atom interacts with two acceptors



The angle $\phi_{A^1 \cdots H \cdots A^2}$ is included in the geometry table. In a truly bifurcated hydrogen bond the sum of the angles with hydrogen as central atom is 360°.

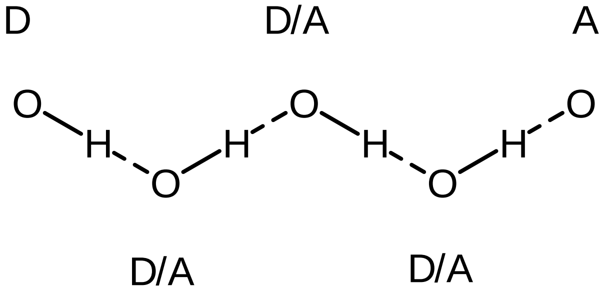
Hydrogen bond networks

Starting from each residue in the unit cell (and including all symmetry-related residues) PLATON builds aggregates of molecules that are joined by hydrogen bonds.

The aggregate is categorized in one of the following patterns

- isolated (finite) cluster of residues
- infinite 1-dimensional chain; the base vector is reported (in fractional co-ordinates)
- infinite 2-dimensional network; two base vectors are reported
- infinite 3-dimensional network; three base vectors are reported

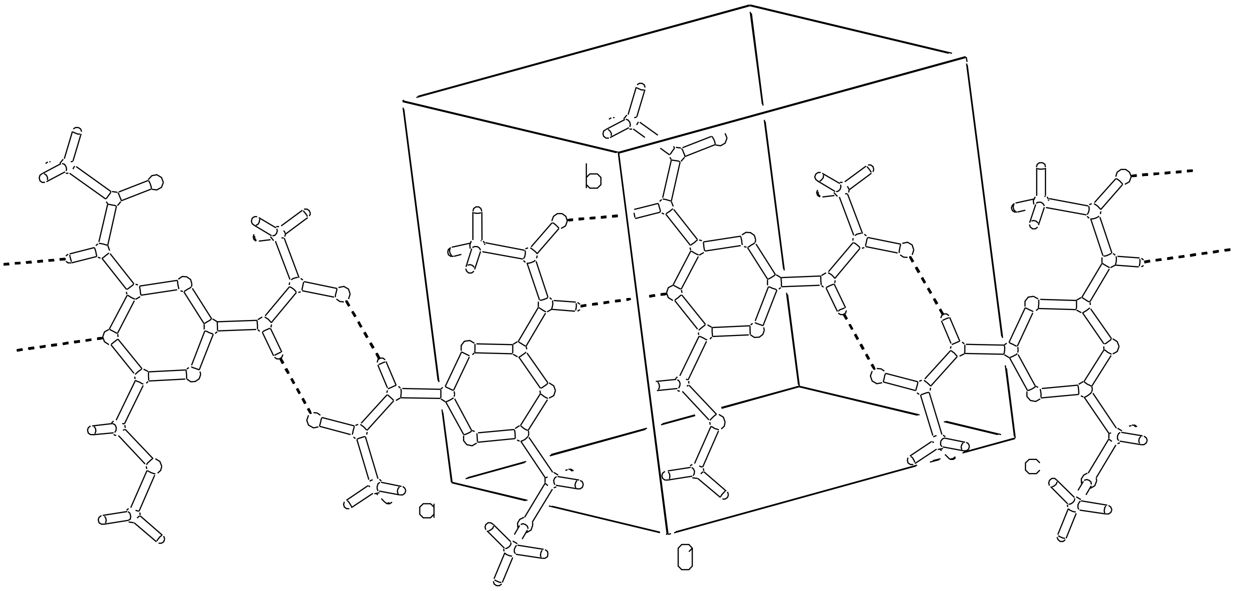
A co-operative effect is observed if the acceptor of a hydrogen bond donates a hydrogen bond to a new acceptor:



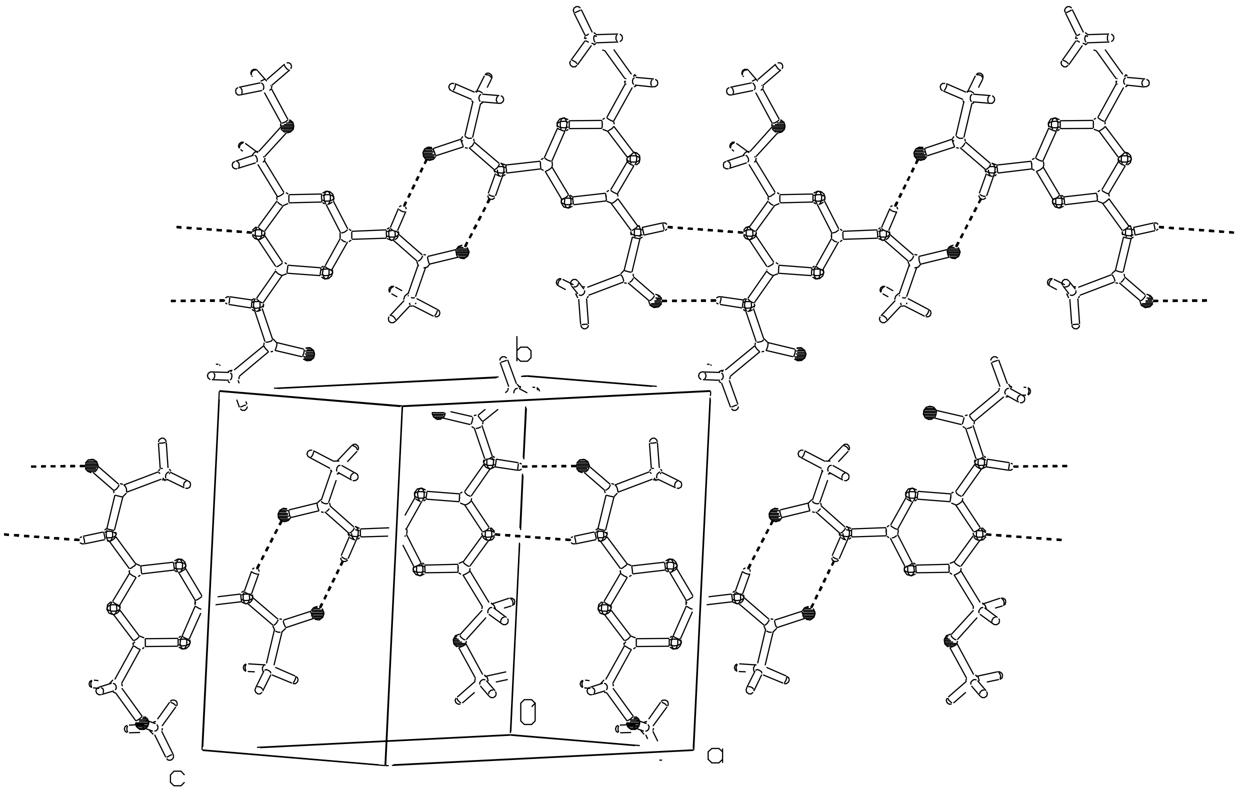
An analysis of co-operative networks can locate the following patterns:

- isolated chain
- closed ring; the number of members is reported (maximum 18)
- infinite chain; the base vector is reported (often associated with crystallographic symmetry elements containing a translation component: screw axes, glide planes)

Example 3 — Hydrogen bond networks (1)

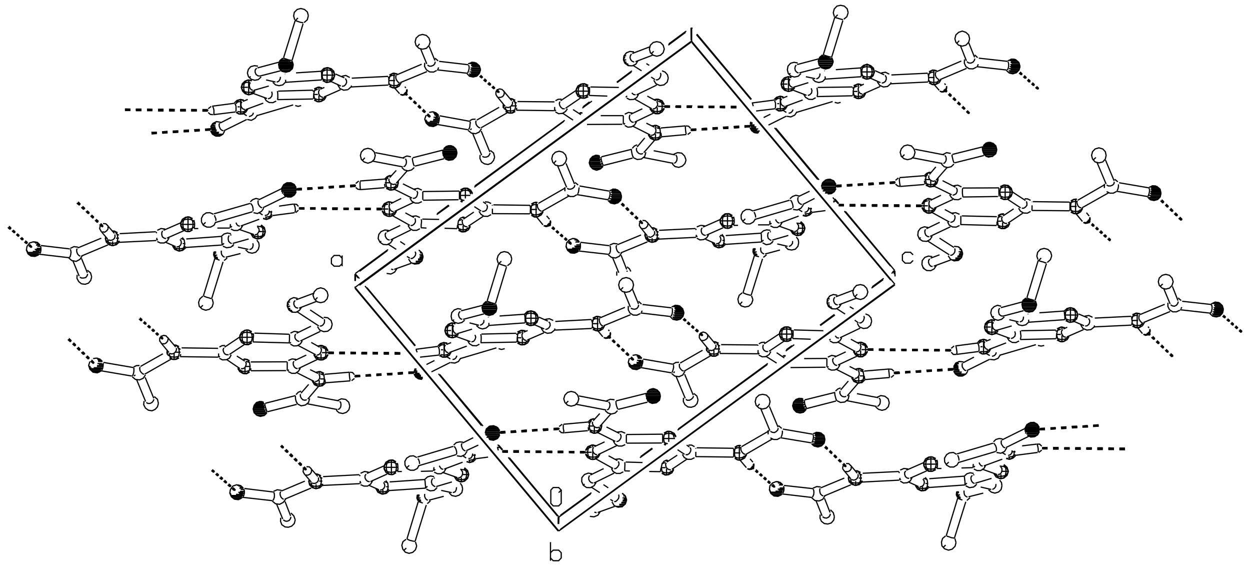


A single chain of hydrogen-bonded residues.

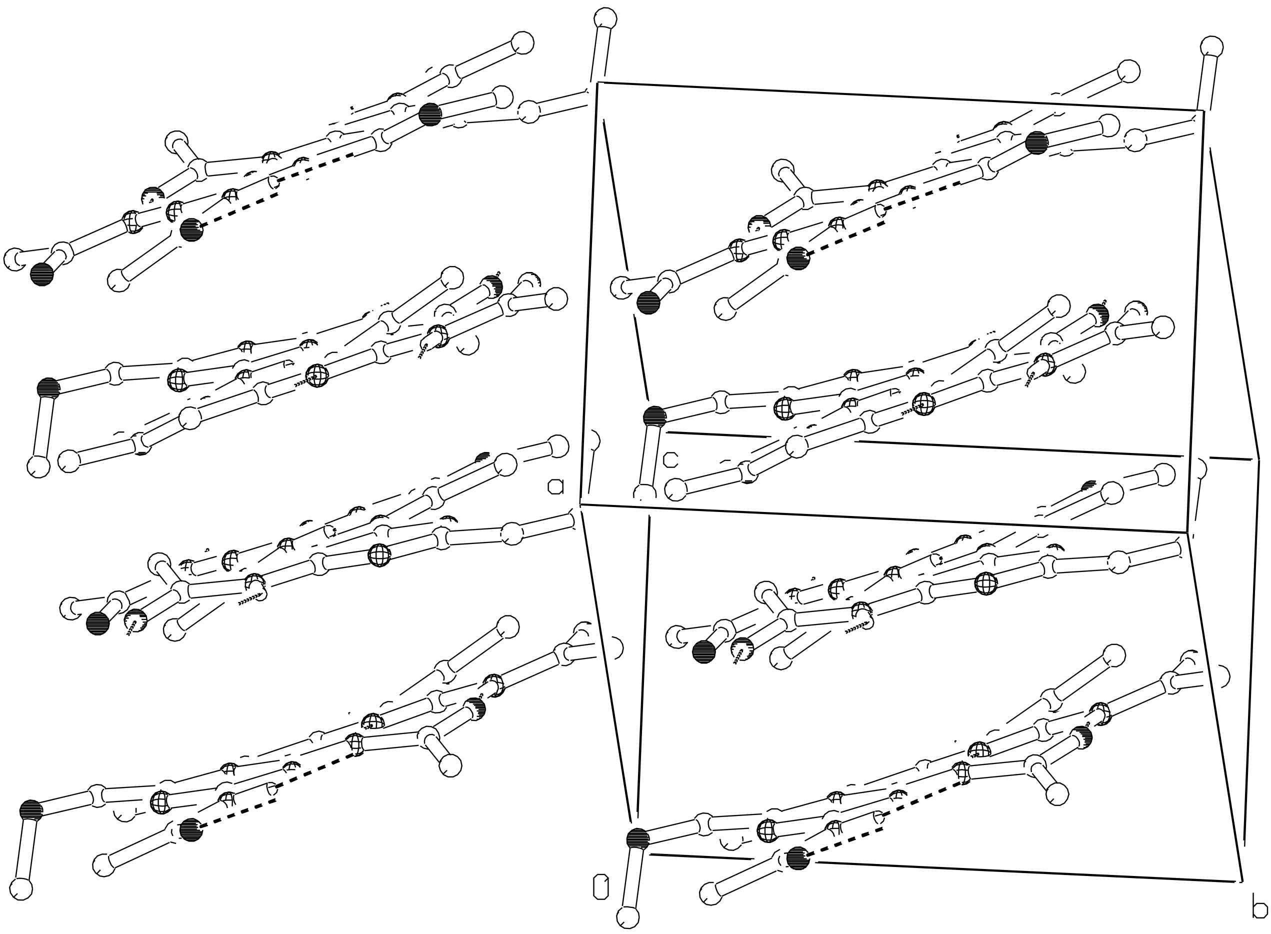


Two parallel chains of hydrogen-bonded residues.

Example 3 — Hydrogen bond networks (1)



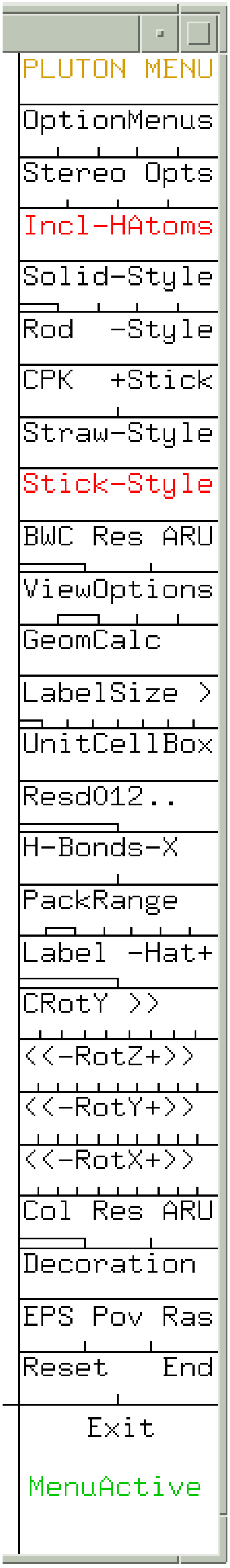
Packing of the unit cell, viewed perpendicular to the direction of the hydrogen-bonded chain. This view is approximately along $\vec{b}$.



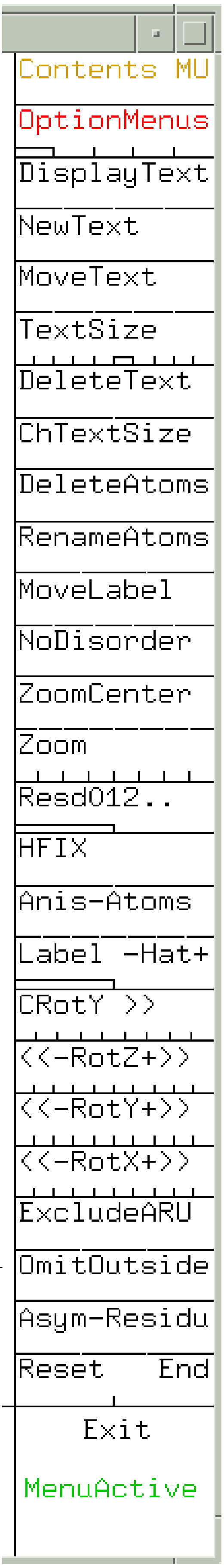
Packing of the unit cell, viewed along the direction of the hydrogen-bonded chains.

PLUTON — crystal packing

main menu



contents menu



Keyboard instructions:

```
ARU nijk.r
ARU −nijk.r
PACK RANGE sx ex sy ey sz ez
PACK PLANE h k l d1 d2 RANGE
      sx ex sy ey sz ez
OMIT OUTSIDE atom_name radius
OMIT OUTSIDE sx ex sy ey sz ez
```

in/exclude unit cell box

select residue (0=all)

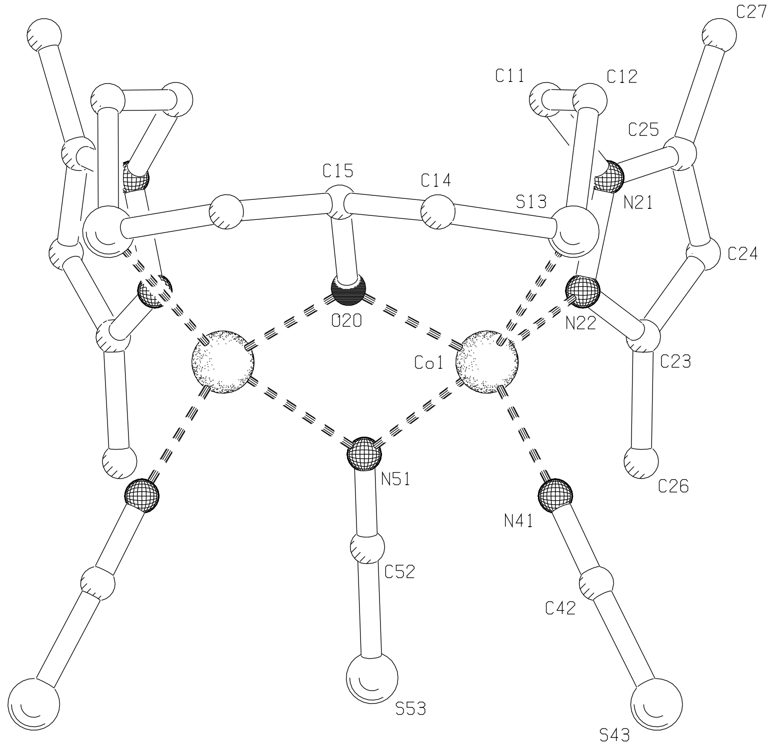
find and display D−H...A bonds;
find and display C−H...A bonds

plot symmetry−related residues
less → more

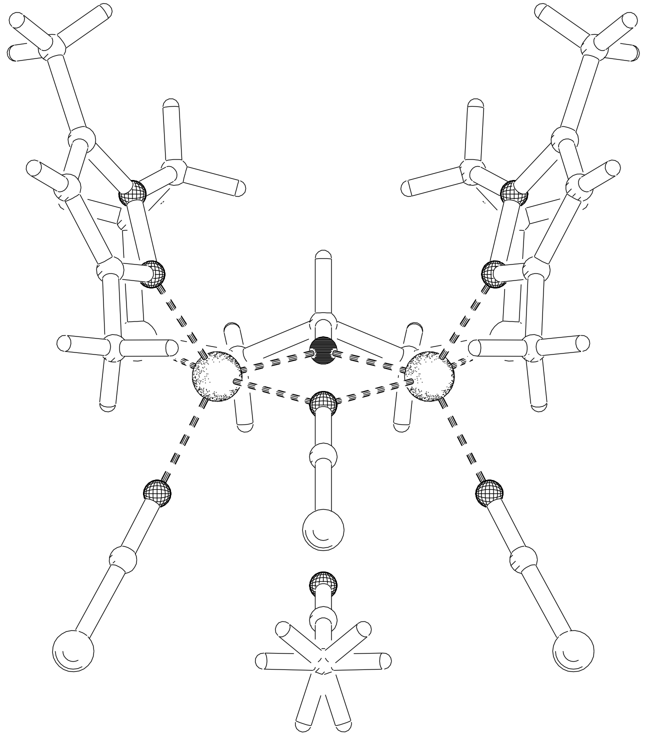
in/exclude atoms outside unit cell

Example 4

$[\mathrm{Co_2(bdnol)_3(NCS)_3}].\mathrm{MeCN}$
(bdnol is 5-dehydro-1,9-bis(3,5-dimethyl-1-pyrazolyl)-3,6-dithia-5-nonanol)



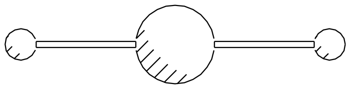
The molecule is located on a crystallographic mirrorplane.



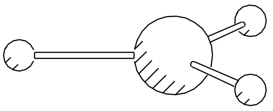
The cocrystallized acetonitril solvent molecules also lie on the crystallographic mirrorplane, displaying disordered hydrogens.

Co-ordination polyhedra

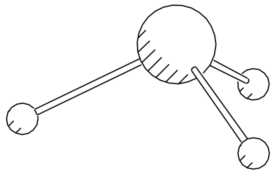
Co-ordination polyhedra, commonly used names and polyhedral symbols



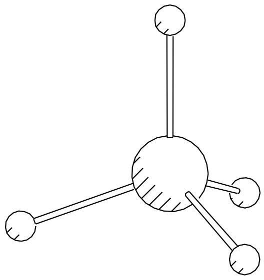
linear
L-2



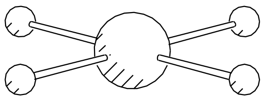
trigonal planar
TP-3



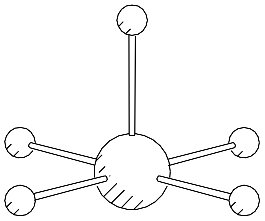
trigonal pyramid
TPY-3



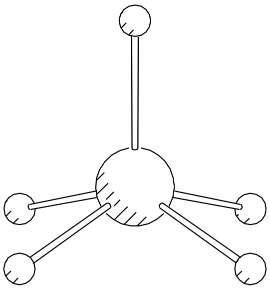
tetrahedron
T-4



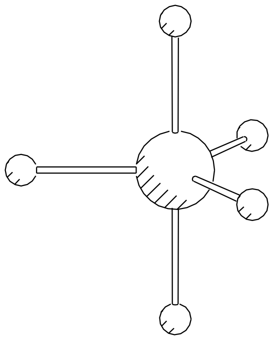
square planar
SP-4



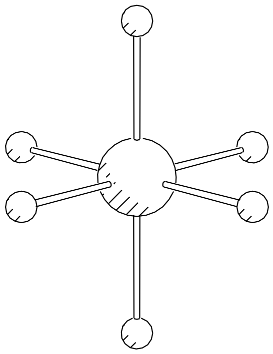
square pyramid
SPY-5



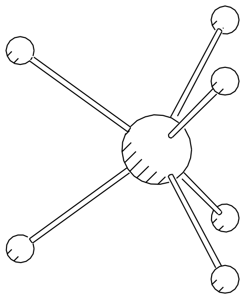
square pyramid
with elevated metal



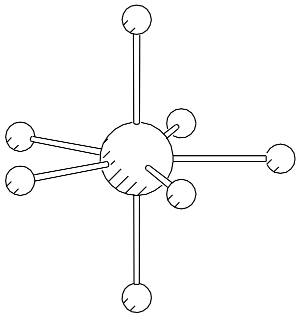
trigonal bipyramid
TBPY-5



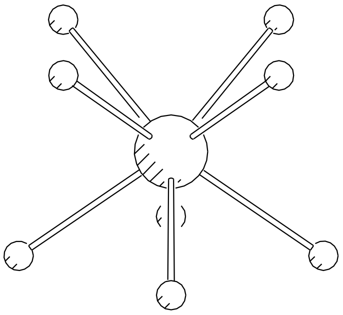
octahedron
OC-6



trigonal prism
TPR-6



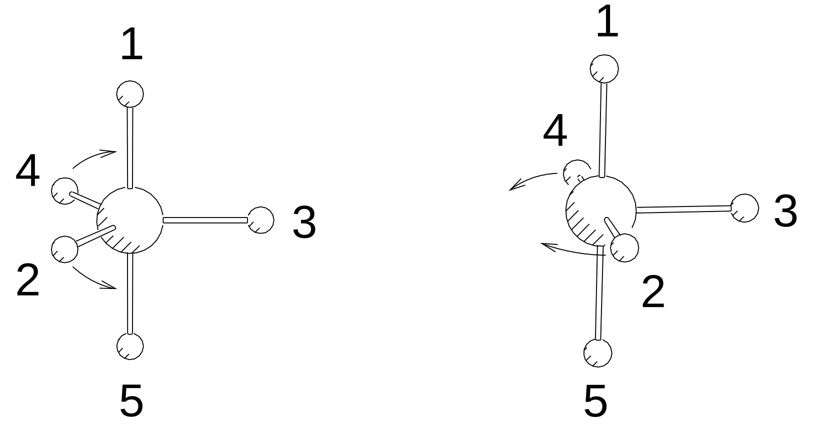
pentagonal bipyramid
PBY-7



square antiprism
SAPR-8

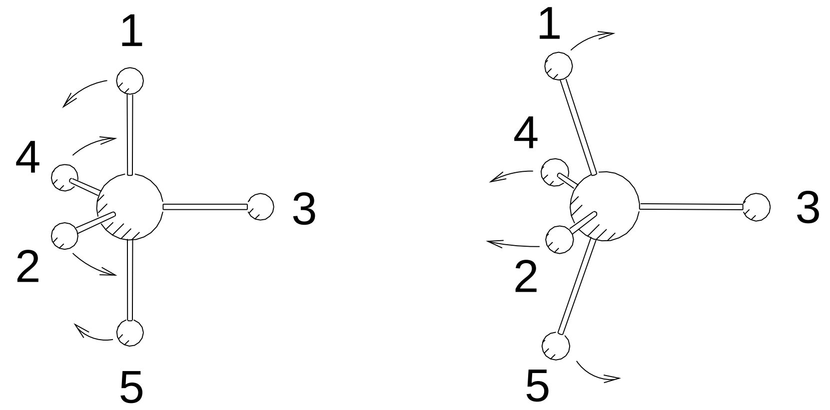
Berry pseudorotation (1)

The Berry pseudorotation transforms one ML_5 co-ordination sphere into another (atom 3 acts as pivot atom):



trigonal bipyramid ⟷ square pyramid

or



trigonal bipyramid ⟷ square pyramid
with elevated metal

The τ descriptor for five-co-ordination is the normalized difference between the largest trans-basal angle $\phi_{\mathrm{L-M-L}}$ (indicated with β) with the second-largest trans-basal angle (indicated with α):

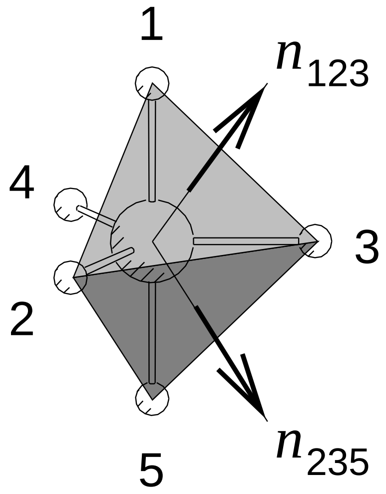
$$\tau = (\beta - \alpha)/60^\circ$$

For an ideal TBP co-ordination $\beta = 180^\circ$; $\alpha = 120^\circ$; $\tau = 1$
For an ideal SP co-ordination $\beta = 180^\circ$; $\alpha = 180^\circ$; $\tau = 0$.
For SP with elevated metal $\alpha = \beta \neq 180^\circ$; $\tau = 0$.

τ indicates the position on the Berry pseudorotation pathway SP $\rightarrow$ TBP.

Berry pseudorotation (2)

The Holmes analysis of five-co-ordination compares the observed angles between the normals $\vec{n}$ of the triangular faces of the co-ordination polyhedron with the ideal values of both trigonal bipyramid and square pyramid co-ordination.



$$\delta_{23} = \angle \vec{n}_{123}, \vec{n}_{235}$$

Two parameters are calculated for the analysis:

$$A = \sum_i |\delta_i^{\text{obs}} - \delta_i^{\text{TBP}}|$$

$$B = \sum_i |\delta_i^{\text{obs}} - \delta_i^{\text{SP}}|$$

For a true Berry pseudorotation $A + B = 217.7^\circ$. If A and $217.7^\circ - B$ are equal, the studied co-ordination sphere lies on the pseudorotation pathway.

The values of δ_i^{SP} are dependent on the transbasal angles 1—M—5 = 2—M—4. The results of the analysis are therefore also dependent on this angle. The default value of 150° can be changed with the keyboard instruction:

```
CALC COORDN element radius FIVE trans_basal_angle
```

The position on the pseudo-rotation pathway TBP → SP (reverse of τ-pathway) is calculated with $(217.7^\circ + A - B)/435.4^\circ$

PLATON — Customizing output (1)

Commands to adapt PLATON output to specific wishes. Most commands are also available from the customize menu. The keyboard instructions are given here.

- `ROUND (ON/OFF) range`
  determines the rounding of derived data
  `range`=1: 1-9 rule
  `range`=2: 1-19 rule (default)
  `range`=3: 1-29 rule
- `INCLUDE element_1 element_2 ...`
  `EXCLUDE element_1 element_2 ...`
  determines which element types are included in or excluded from the calculation. `element` can be an element symbol or MET. Should be given before any calculations are performed. In- and exclusion are not reversible once the calculations have started.
- `NOMOVE (ON/OFF)`
  `ON` keeps atoms at input positions, avoid repositioning. Equivalent to “Move” option in the customize menu.
- `PARENTHESES (ON/OFF)`
  switch the use of parentheses in atom labels on or off.
- `LSPL atom_name_1 atom_name_2 ...`
  define extra least-squares planes for analysis
- `LINE atom_name_1 atom_name_2 ...`
  adds a “bond” between non-bonded atoms to the bond—plane angle list of the least-squares planes analysis
- `RING atom_name_1 atom_name_2 ...`
  define extra rings (non-bonded rings are possible) for analysis

PLATON — Customizing output (2)

- `SET WINDOWS fraction`
  Adapt the size of the X-windows graphic interface.
- `SET PAR par_number new_value`
  `SET IPR ipr_number new_value`
  `SET GL gl_number new_value`
  resetting of internal parameters: real, integer and global parameter, respectively.
- `DOAC element_1 element_2 ...`
  Redefine potential hydrogen bond donor and acceptor atoms by element type. Default is N, O, F, S, Cl, Br.
- `HBOND p1 p2 p3`
  redefine criteria for the classification of a $D–H \cdots A$ system as a hydrogen bond:
  $d_{D \cdots A} < r_D + r_A + \text{p1}$
  $d_{H \cdots A} < r_H + r_A + \text{p2}$
  $\phi_{D-H \cdots A} > \text{p3}$
  Default values are p1 = +0.5 Å, p1 = –0.12 Å, p3 = 100°.
- `HBOND NORM (p1 p2 p3)`
  invokes normalization of d_{D-H} bond lengths and a subsequent hydrogen bond analysis. (p1, p2 and p3 as in the previous item). The default values for d_{D-H}, which are based on neutron diffraction data, can be changed by resetting of internal parameters.

$D-H$	d_{D-H}	instruction
C—H	1.083	`SET PAR 296 new_value`
N—H	1.009	`SET PAR 297 new_value`
O—H	0.983	`SET PAR 298 new_value`

PLUTON — Customizing output (1)

Most keyboard instructions can also be found in the menu bars.

- `EXCLUDE atom_name_1 atom_name_2 ...`
  `EXCLUDE element_1 element_2 ...`
  `INCLUDE atom_name_1 atom_name_2 ...`
  `INCLUDE element_1 element_2 ...`
  `INCLUDE ALL`
  Determines which atoms or elements are included in the plot. In- and exclusion are reversible during the run of the program.
- `JOIN atom_name_1 atom_name_2 (symmetry_operation)`
  Add (intermolecular) bonds to the plot. `symmetry_operation` is given as “nijk” or “nijk.r”. If a symmetry related atom is added to the bond list, a lowercase letter is attached to the label, e.g. C1_b, Fe2_a. Subsequent commands can refer to these labels as C(1)b and Fe(2)a.
- `DETACH atom_name_1 atom_name_2`
  Remove bonds from the plot.
- `DEFINE atom_name_1 TO atom_name_2 atom_name_3 ...`
  Replace the bonds 1—2, 1—3, . . . with a single bond 1—Cg, where Cg is the geometrical centre of atoms 2, 3, . . .
- `COL TYPE element_1 colour_1 element_2 colour_2 ...`
  Assign colours to element types. `colour` is one of: black (white), red, green, blue, yellow, orange, violet, brown, none.
- `BWC TYPE element_1 bwc_1 element_2 bwc_2 ...`
  Assign hatching patterns (“black and white colours”) to element types. `bwc` is one of: `bwc` is one of: contour, net, shade, segment, dots, black, cross, parallel, globe, meridian, horizontal, vertical, mesh, diagonal, slant, textile, void.
- `RADII ATOMS element_1 r_1 element_2 r_2 ...`
  Change plot radii of atoms, `r` is specified in Å.

PLUTON — Customizing output (2)

- `RADII BONDS (DASH) atom_name_1 atom_name_2 r n`
  `RADII BONDS (DASH) TO element r n`
  `RADII BONDS (DASH) INTER r n`
  `RADII BONDS (DASH) NORMAL r n`
  `RADII BONDS (DASH) ALL r n`
  Change drawing style of a bond. `r` is the bond radius; `n` is the number of lines drawn for each bond. The optional `DASH` parameter causes dashed lines to be drawn.
  "`TO element`" bonds are bonds to a certain element type, specified by chemical element symbol
  "`INTER`" bonds are intermolecular bonds and all hydrogen bonds (including intramolecular H-bonds)
  "`NORMAL`" bonds are all intramolecular covalent bonds not falling in a "`TO element`" category.
  "`ALL`" effects all bonds, overriding earlier instructions.

- `PUT atom_name_1 position_1 atom_name_2 position_2 ...`
  Change position of atomlabels. `position` is one of n, ne, e, se, s, sw, w, nw, nucl, auto. The symbols "n" through "nw" refer to the compass directions, with "n" above the atom, "w" to the left; "nucl" places a label at the centre of the sphere, "auto" calls for automatic placement, based on minimal overlap considerations.

- `TITL title`
  Change title included in plot.

- `VIEW UNIT XR xrotation YR yrotation ZR zrotation`
  Recall previously found rotation.

- `MONO PERSP distance`
  Make a perspective (i.e. point projected as opposed to plane projected) drawing, with the viewpoint `distance` cm from the screen. `distance` = 60 gives convenient perspective in most cases.

Graph sets — definitions

Graph sets can be used to identify hydrogen bonding patterns. The general notation is

$$G_d^a(n)$$

G is the descriptor and indicates the type of network:

S = intramolecular
D = finite, isolated patterns
R = rings
C = infinite chains

The superscript a indicates the number of acceptor atoms and the subscript d gives the number of donor atoms. If $a = d = 1$ they are not included in the graph set symbol.

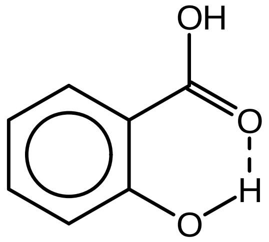
The degree n is the total number of atoms in the pattern, counting via the shortest route. It is not defined for patterns of type D.

A motif, or unitary graph set is a hydrogen-bonded set in which only one chemically distinct type of hydrogen bond is present.

A binary graph set is a descriptor of a pattern in which two types of hydrogen bonds are present.

Graph sets — examples

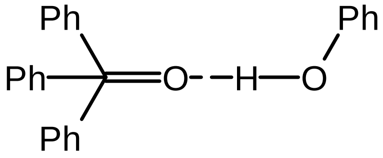
- Intramolecular hydrogen bond:



$S(6)$

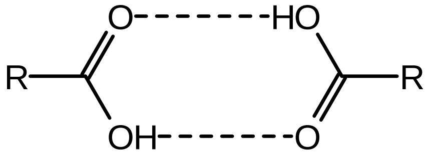
n is counted over the smallest ring formed by the hydrogen bond.

- Isolated hydrogen bond:



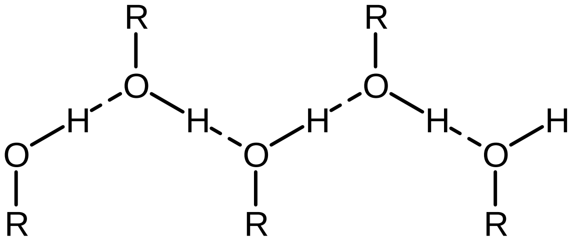
D

- Acid dimer:

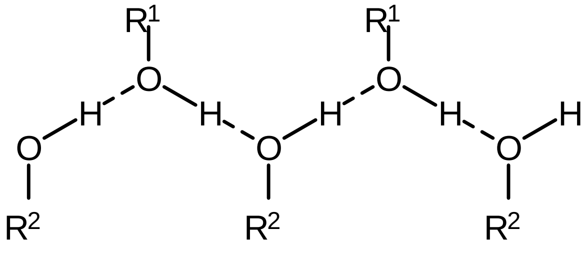


$R_2^2(8)$

- Co-operative hydrogen-bonded chain:



$C(2)$

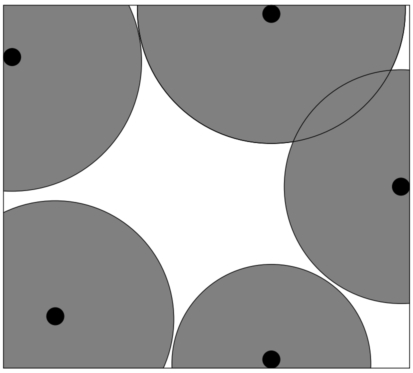


unitary DD
binary $C_2^2(4)$

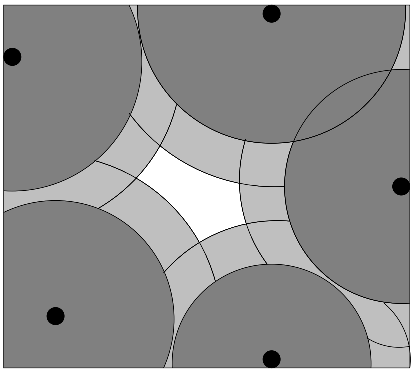
Solvent accessible voids

In most structures, 30% of the volume is outside the van der Waals spheres of the constituting atoms. Solvent accessible voids can be calculated with PLATON's “CALC SOLV” command and are defined as follows:

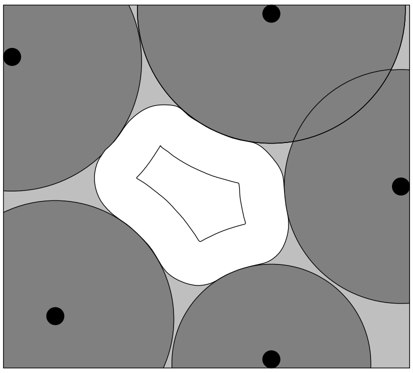
1. Exclude the volume inside the van der Waals spheres.



2. Add the radius of the smallest possible solvent (water, $r = 1.2$ Å) to the excluded volume of step 1. The remaining volume (white) is accessible for the nucleus of the solvent.

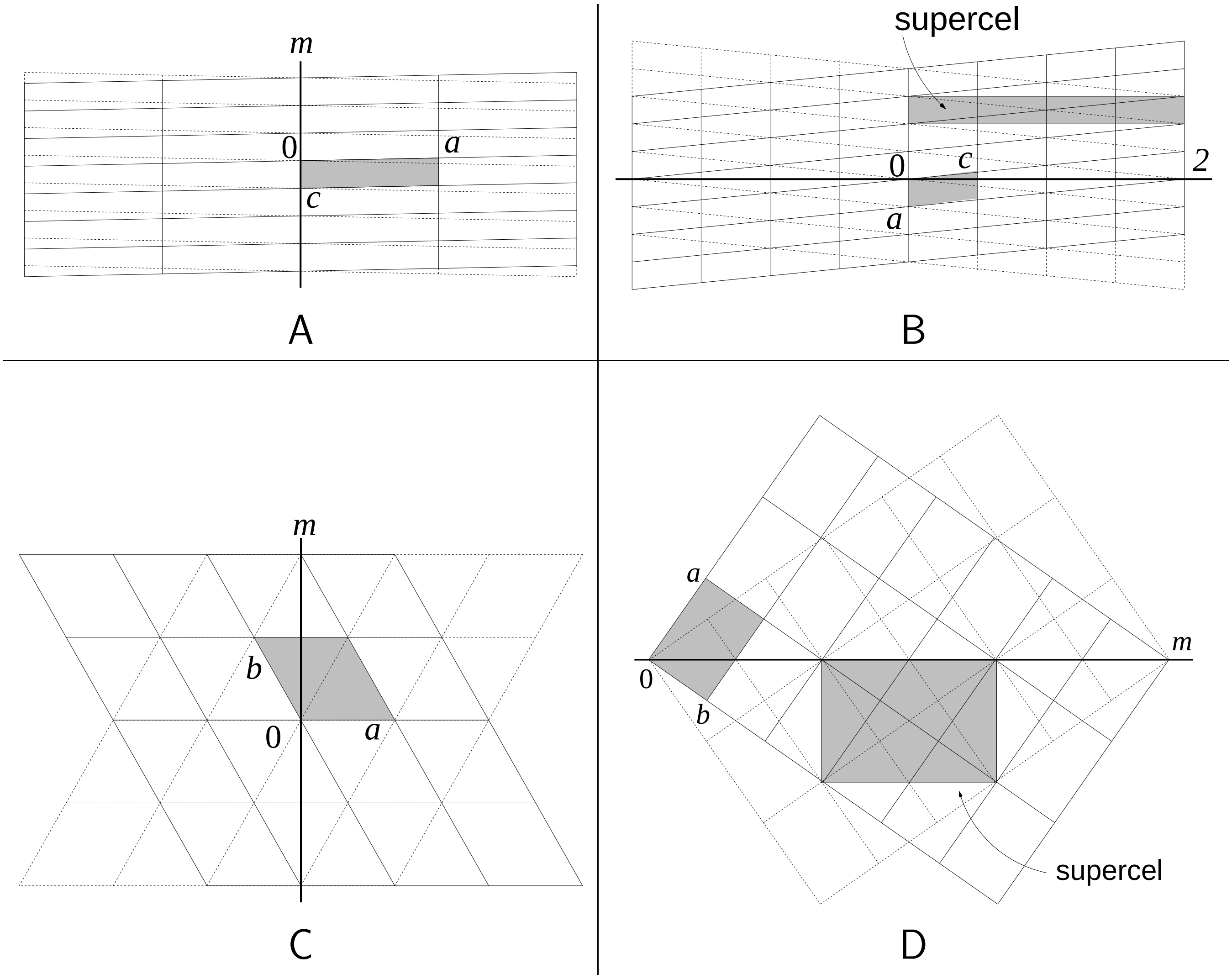


3. Add the radius of the smallest possible solvent to the remaining volume of step 2. The solvent accessible volume is indicated in white.



Twinning

Twins are crystals consisting of intergrown individuals. There are two (or more) translation lattices present in the sample, related by a simple symmetry operation, the twin law:



Lattice points of the individual lattices can overlap. Several situations are possible: hardly any overlap (A); complete overlap (C); overlap in zones of one index, e.g. $h_3 = 2n$ (B); overlap mimicking a larger unit cell or supercell (B, D).

The lattice points of the reciprocal lattice (i.e. reflections) can overlap in a similar way.

If the twin has been properly handled (de-twinned), there are no consequences for the interpretation of the crystal structure.

Twinning and disorder

Twinning and disorder have in common that not all unit cells can be exactly super-imposed by applying operations belonging to a translation lattice.

	disorder	twinning
unit cell orientation	equal	different
atomic coordinates	different (available for all components)	equal
nature of differences at molecular level	orientation configuration conformation chemistry	orientation configuration (racemic twin)
cell stacking	random	domains
reflection intensity	$I \sim (F_A + F_B)^2$	$I \sim F_A^2 + F_B^2$
area affected	part of the unit cell	complete unit cell

There is a “gray area” between twinning and disorder

Anomalous diffraction

X-radiation with a wavelength close to the absorption edge of one of the elements present in the crystal obtains an extra phase shift upon scattering. To account for this shift a real and imaginary dispersion correction must be added to the atomic scattering factor:

$$f^{\text{anom}}\left(\frac{\sin\theta}{\lambda}\right) = f^{0}\left(\frac{\sin\theta}{\lambda}\right) + f'(\lambda) + if''(\lambda)$$

N.B. f^0 and f^{anom} are functions of $\sin(\theta)/\lambda$; f' and f'' depend only on λ.

The values of f' and f'' for the elements present in a structure are tabulated in the PLATON listing.

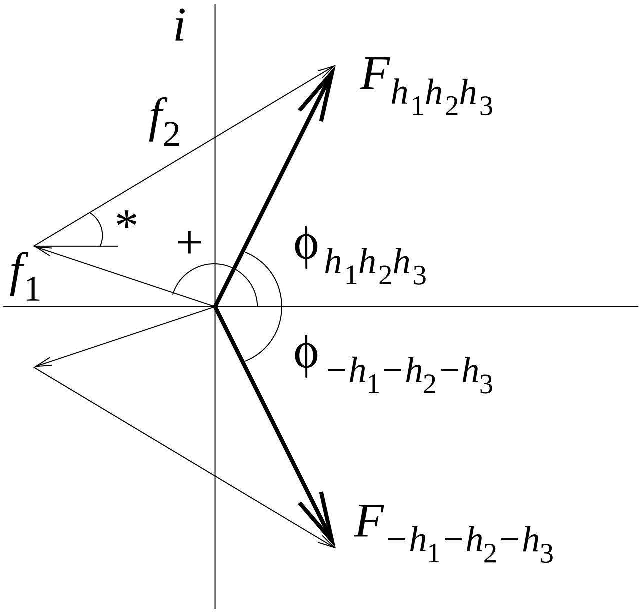
The contribution of anomalous diffraction to the intensity is dependent on the elements types present and the wavelength of the radiation:

		Mo$K\alpha$		Cu$K\alpha$	
	$f^0(0)$	f'	f''	f'	f''
H	1.000	0.000	0.000	0.000	0.000
C	5.999	0.003	0.002	0.018	0.009
O	7.999	0.011	0.006	0.049	0.032
Cl	17.001	0.148	0.159	0.364	0.702
Fe	25.990	0.346	0.844	−1.134	3.197
Br	34.993	−0.290	2.460	−0.676	1.281
Ru	43.995	−1.259	0.836	0.055	3.296
Hg	79.957	−2.389	9.227	−4.292	7.685

Normal diffraction and absolute configuration

When anomalous diffraction is ignored, the structure factors $F_{h_1h_2h_3}$ and $F_{-h_1-h_2-h_3}$ have the same amplitude, but opposite phases:

$$F_{h_1h_2h_3} = \sum_{j=1}^{N} f_j^0 T_j e^{2\pi i(h_1x_j+h_2y_j+h_3z_j)}$$



$$+ = 2\pi(h_1x_1 + h_2y_1 + h_3z_1) \quad * = 2\pi(h_1x_2 + h_2y_2 + h_3z_2)$$

$$|F_{h_1h_2h_3}| = |F_{-h_1-h_2-h_3}|$$

$$\phi_{h_1h_2h_3} = -\phi_{-h_1-h_2-h_3}$$

These relationships are known as Friedel's law. $F_{h_1h_2h_3}$ and $F_{-h_1-h_2-h_3}$ are called a Friedel pair.

Inverting the structure $[(x, y, z) \rightarrow (-x, -y, -z)]$ has the same effect on F as inverting the Laue indices $[h_1h_2h_3 \rightarrow -h_1 - h_2 - h_3]$.

Both operations do not change the appearance of the diffraction pattern (position and intensity of the peaks).

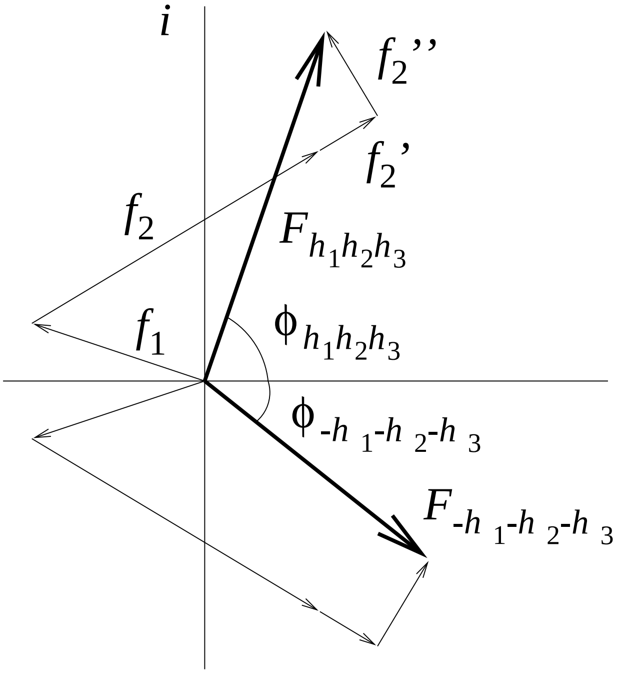
Using normal diffraction no distinction can be made between enantiomorphs.

Anomalous diffraction and absolute configuration

Taking anomalous diffraction into account results in different values for $F_{h_1h_2h_3}$ and $F_{-h_1-h_2-h_3}$

$$F_{h_1h_2h_3} = \sum_{j=1}^{N} (f_j^0 + f_j' + i f_j'') T_j e^{2\pi i(h_1x_j + h_2y_j + h_3z_j)}$$

Assume that only one atom displays significant anomalous scattering.



$$|F_{h_1h_2h_3}| \neq |F_{-h_1-h_2-h_3}|$$

$$\phi_{h_1h_2h_3} \neq -\phi_{-h_1-h_2-h_3}$$

Because of this difference the co-ordinate set (x, y, z) can be distinguished from the co-ordinate set $(-x, -y, -z)$, provided that co-ordinates and Laue indices are consistently defined in a right-handed axes system.

Determination of the absolute configuration (1)

The absolute structure can be established by co-crystallizing with an internal standard, i.e. a chiral centre of known configuration.

Prerequisites for establishing absolute configurations ab initio are:

- presence of significant anomalous scatterers (heavy atoms)
- precise intensity measurement (good crystal)

A simple way to establish absolute configuration is to compare the sign of the observed Bijvoet differences:

$$\triangle^{\mathsf{obs}}_{h_1 h_2 h_3} = |F^{\mathsf{obs}}_{h_1 h_2 h_3}|^2 - |F^{\mathsf{obs}}_{-h_1 -h_2 -h_3}|^2$$

with that of the calculated Bijvoet difference. If the signs agree the absolute structure has been correctly assigned.

A graph of $\triangle^{\mathsf{obs}}_{h_1 h_2 h_3}$ vs $\triangle^{\mathsf{calc}}_{h_1 h_2 h_3}$ can be made with the “BijvoetPair” option in PLATON.

Determination of the absolute configuration (2)

Flack's x-parameter can be used to determine the absolute structure. This method is based on the occurrence of twins with an inversion centre as twin operation (racemic twins). The observed structure factors G of these twins have contributions from both individuals:

$$G^2_{h_1h_2h_3} = (1-x)|F_{h_1h_2h_3}|^2 + x|F_{-h_1-h_2-h_3}|^2$$

The physical domain of the x-parameter is $0 \leq x \leq 1$.

x can be introduced as a parameter during structure refinement. The uncertainty u_x is a measure for the inversion-distinguishing power:

$u_x < 0.04$	strong inversion-distinguishing power
$u_x < 0.08$	enantiopure-sufficient inversion-distinguishing power
$u_x > 0.3$	weak inversion-distinguishing power

Enantiopure-sufficient inversion-distinguishing power is sufficient if a priori knowledge of the purity of the sample is available.

A structure with $|x| < 2u_x$ assures a valid absolute structure determination of an enantiopure crystal.

N.B. 1 x is sometimes not refined but calculated after the refinement of the structure (e.g. in SHELXL). This may lead to erroneous results. An x obtained this way should always be compared to the x calculated for the inverted structure.

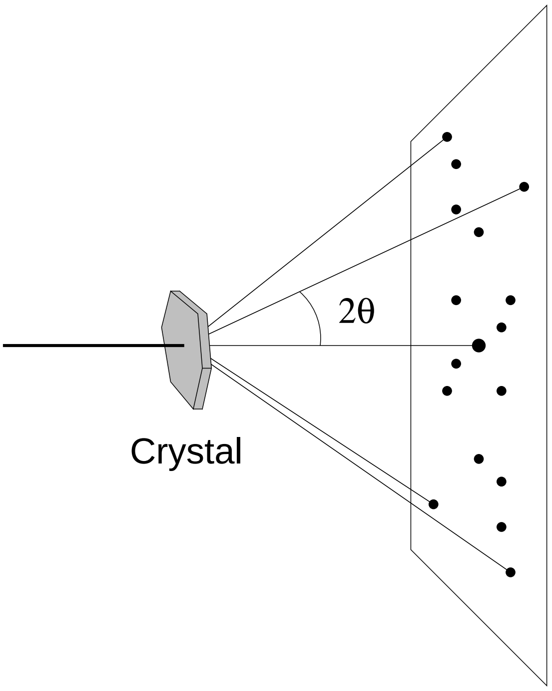
N.B. 2 The absolute structure determination is only an indication for the absolute structure of the measured crystal, not the whole batch. Some racemates crystallize as a mixture of enantiopure crystals.

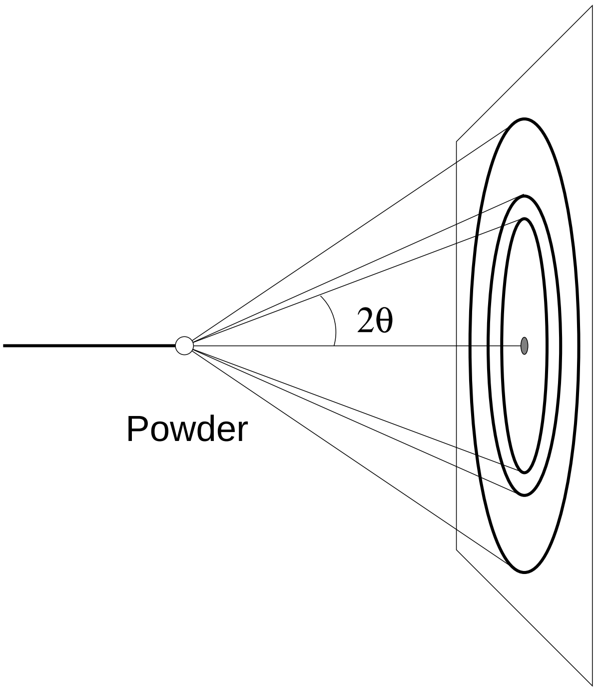
Powder diffraction (1)

A crystalline powder consist of small grains; each grain is a single crystal.

In general the orientation of the grains in a powder sample is random. Grains with anisotropic shapes tend to show some ordering.

The diffraction pattern consists of discrete cones, with top angle 4θ, dictated by Bragg's law. When projected on a plane perpendicular to the incident beam, a series of concentric circles is obtained.





The diffraction pattern of a single crystal will change when the sample is rotated. The diffraction pattern of a powder will keep the same appearance when the sample is rotated, although other powder grains will contribute to the pattern.

Powder diffraction (2)

In a powder diffraction pattern large numbers of reflections will overlap, causing problems with

- indexing (determination of the unit cell)
- intensity measurement

To obtain accurate powder diffraction patterns, the use of synchrotron data is advisable. The ab initio structure determination from powder data is far from routine.

Powder diffraction can be used to identify materials without solving the structure. Identification is performed by comparing the observed diffraction pattern with the pattern in a data base, the Powder Diffraction File.

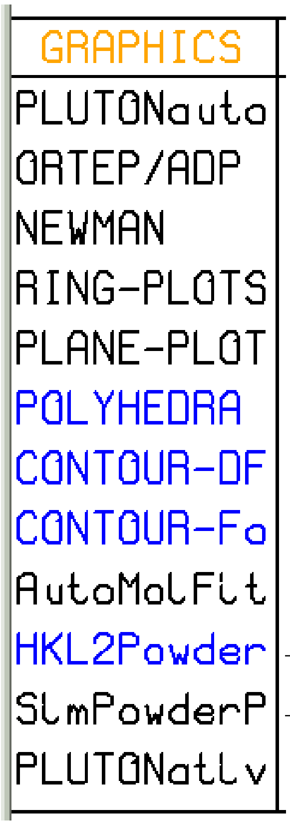
PLATON offers two options to calculate a powder diffraction pattern

- based on atomic co-ordinates: intensities and 2θ values are calculated
- based on (single crystal) diffraction data: 2θ values are calculated, intensities are read from file

A reliable diffraction pattern can only be obtained when a complete data set is available.

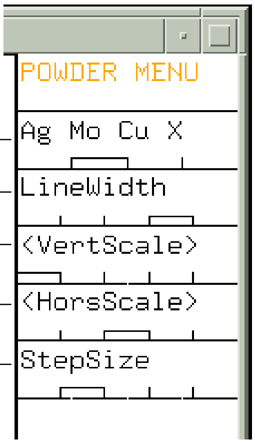
Simulated powder diffraction patterns

PLATON front page



HKL2Powder → based on observed structure factors

SimPowderP → based on atomic co−ordinates



radiation, X = specified on input file

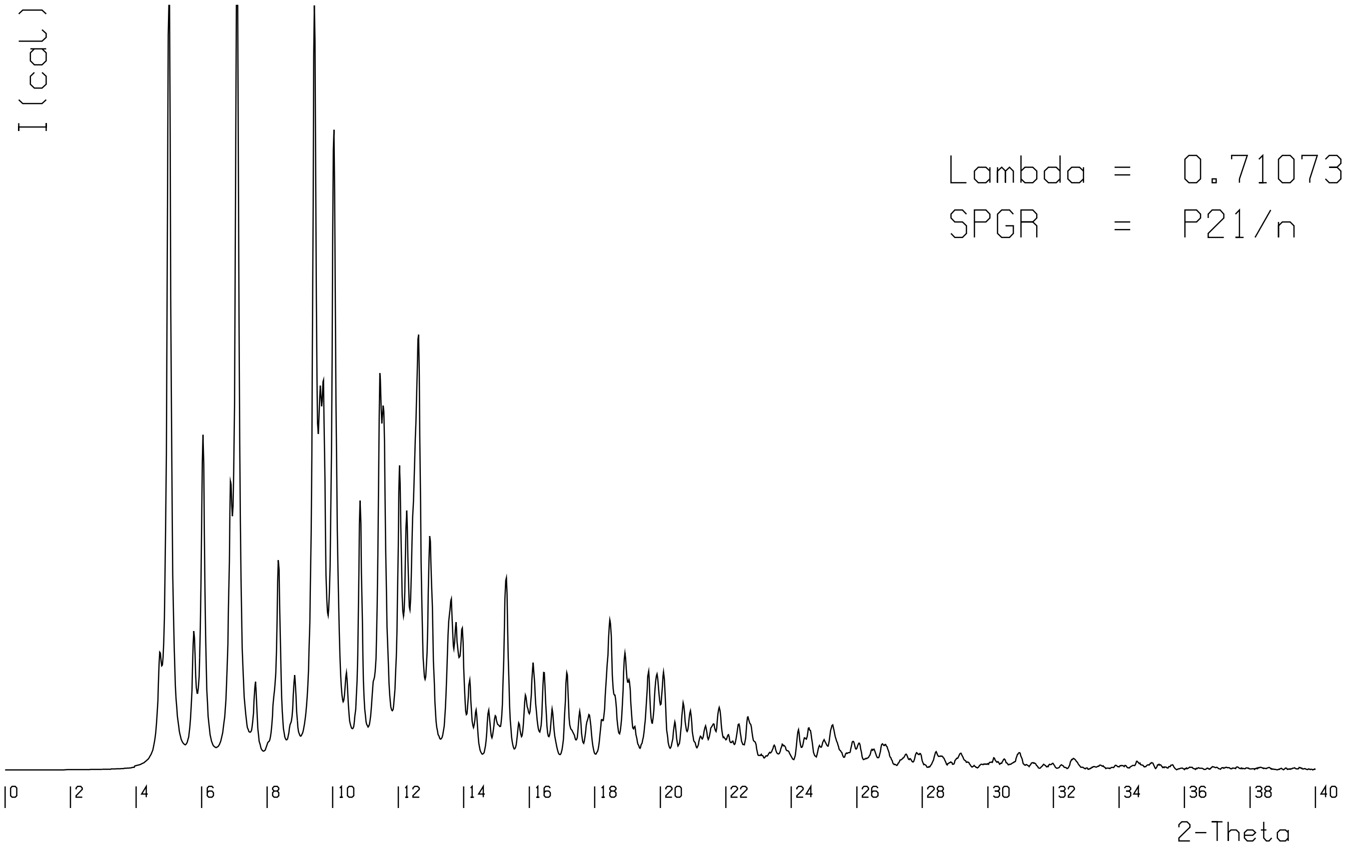
linewidth (narrow → wide)

vertical scale (left: scale max)

horizontal scale (2θ range)

number of points calculated

Powder diffraction pattern of Example 2, based on atomic co−ordinates



Fiber diffraction

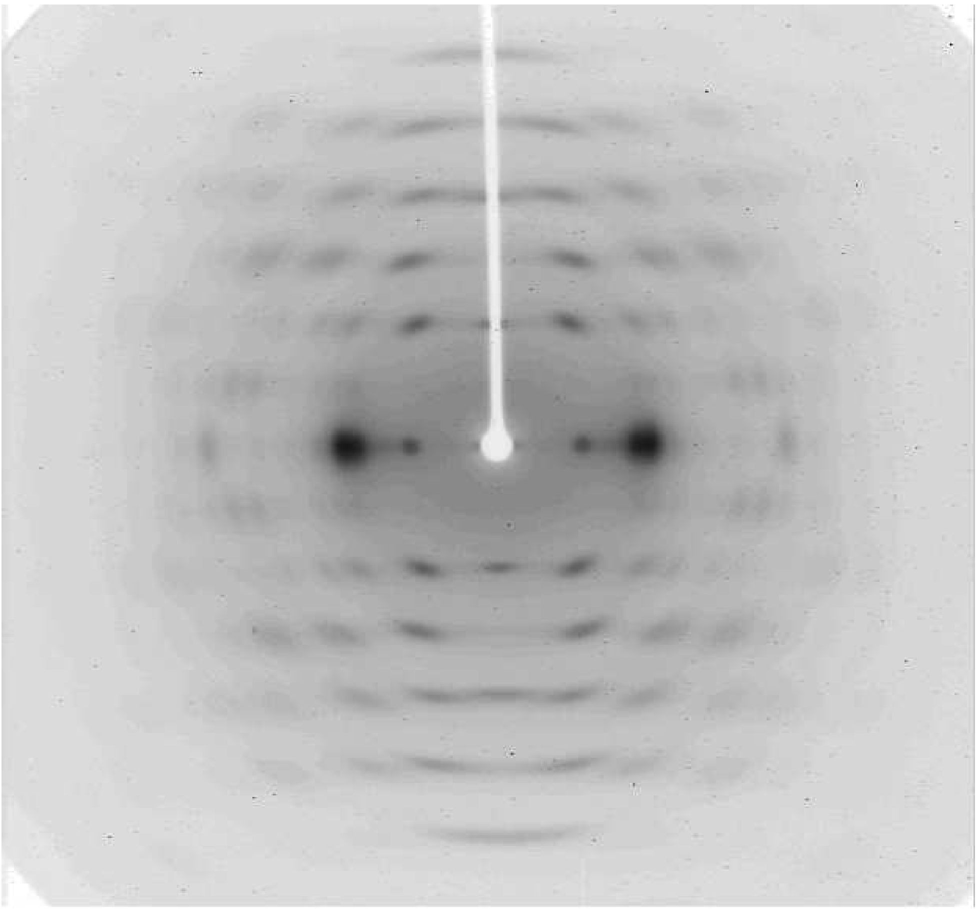
Most fibers possess crystalline domains.

Strongest ordering is along the fiber axis, with short range order (like in a crystalline powder) perpendicular to this axis.

Diffraction patterns of fibers are more diffuse than those of single crystals.

Reflections are grouped in layers with equal index l, associated with the fiber axis, but with significant overlap within the layer.

There are too few reflections for ab initio structure determination; confirmation of a structural model is in some cases possible.



Crystallographic data bases

Results of crystal structure determinations are gathered in a number of data bases.

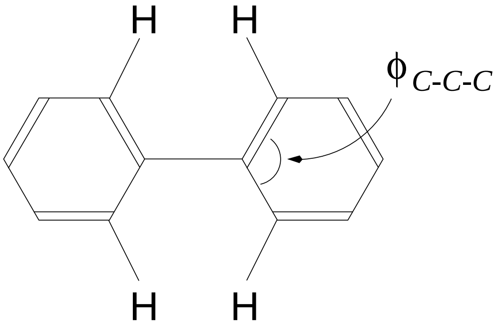
data base name	type of compounds	#entries	growth (#/year)
Cambridge Structural Data base	organic, organometallic, coordination compounds	325,709 (2004)	25,000
Inorganic Crystal Structure Data base	inorganic compounds, minerals	76,000 (2004)	2,000
NRCC Metals Crystallographic Data File	metals, alloys	60,000	— (closed)
Protein Data Bank	proteins, bio-macro-molecular complexes	25,000 (2004)	1,000
Powder Diffraction File	inorganic, organic, organometallic, coordination compounds	350,000 (2003)	20,000

The powder diffraction file also contains simulated patterns based on single crystal data from the CSD and ICSD.

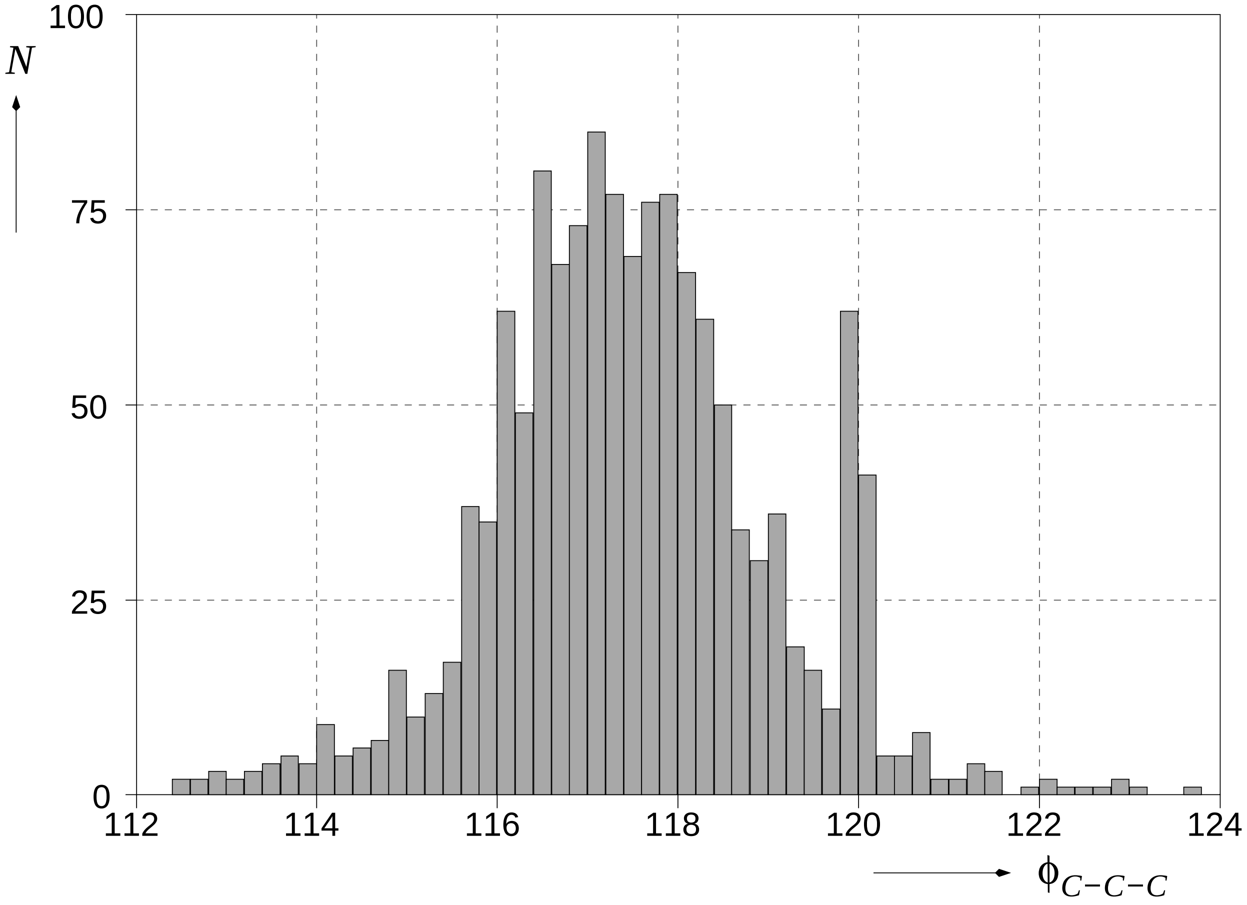
Data bases can contain structures published in a scientific journal as well as (unrefereed) "private communications".

Pollution of geometry statistics by constraints

Crystal structures containing the following fragment were extracted from the Cambridge Structural Database (February 2003 update, 281 936 entries).



812 structures, containing 1383 fragments were found. The values of the bond angles ϕ_{C-C-C} were calculated and a histogram was constructed from the results.

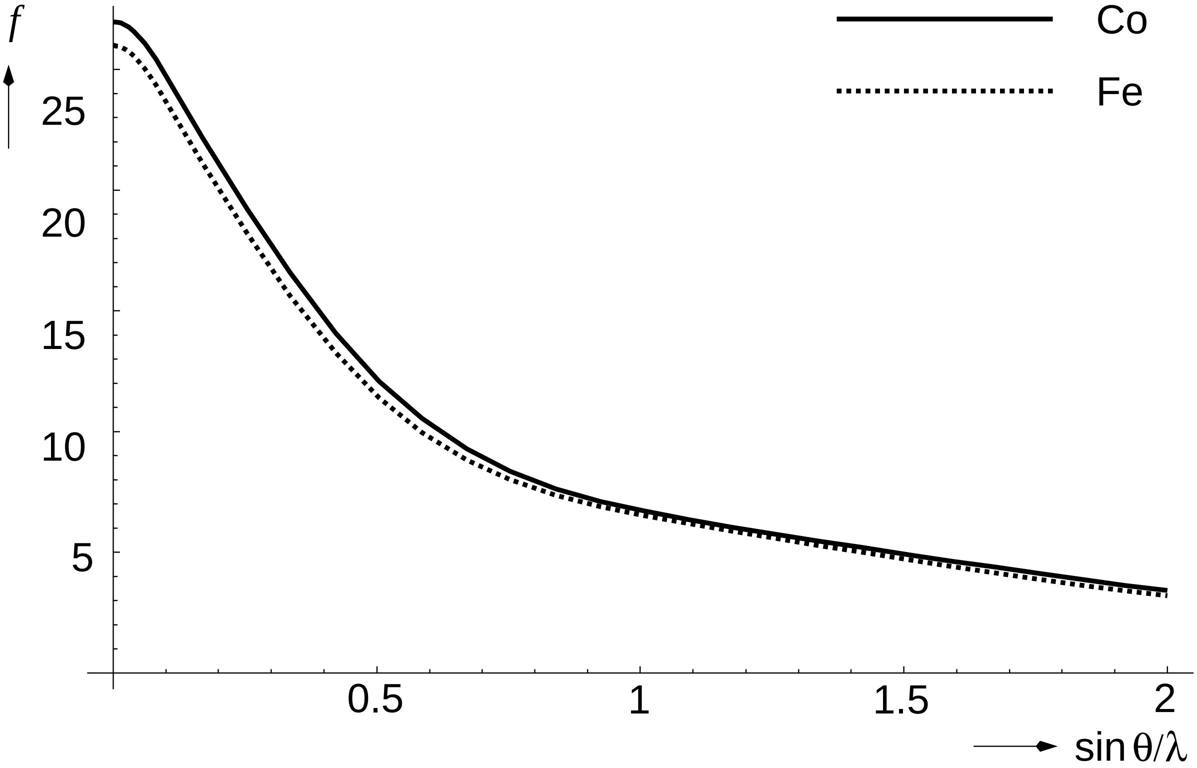


The sharp peak at $\phi_{C-C-C} = 120°$ contains those structures in which the phenyl group has been refined as a regular hexagon.

Wrong structures (1)

The fact that a structure is published in a scientific journal is no guarantee that it is correct. Some of the more common errors found in literature are listed below.

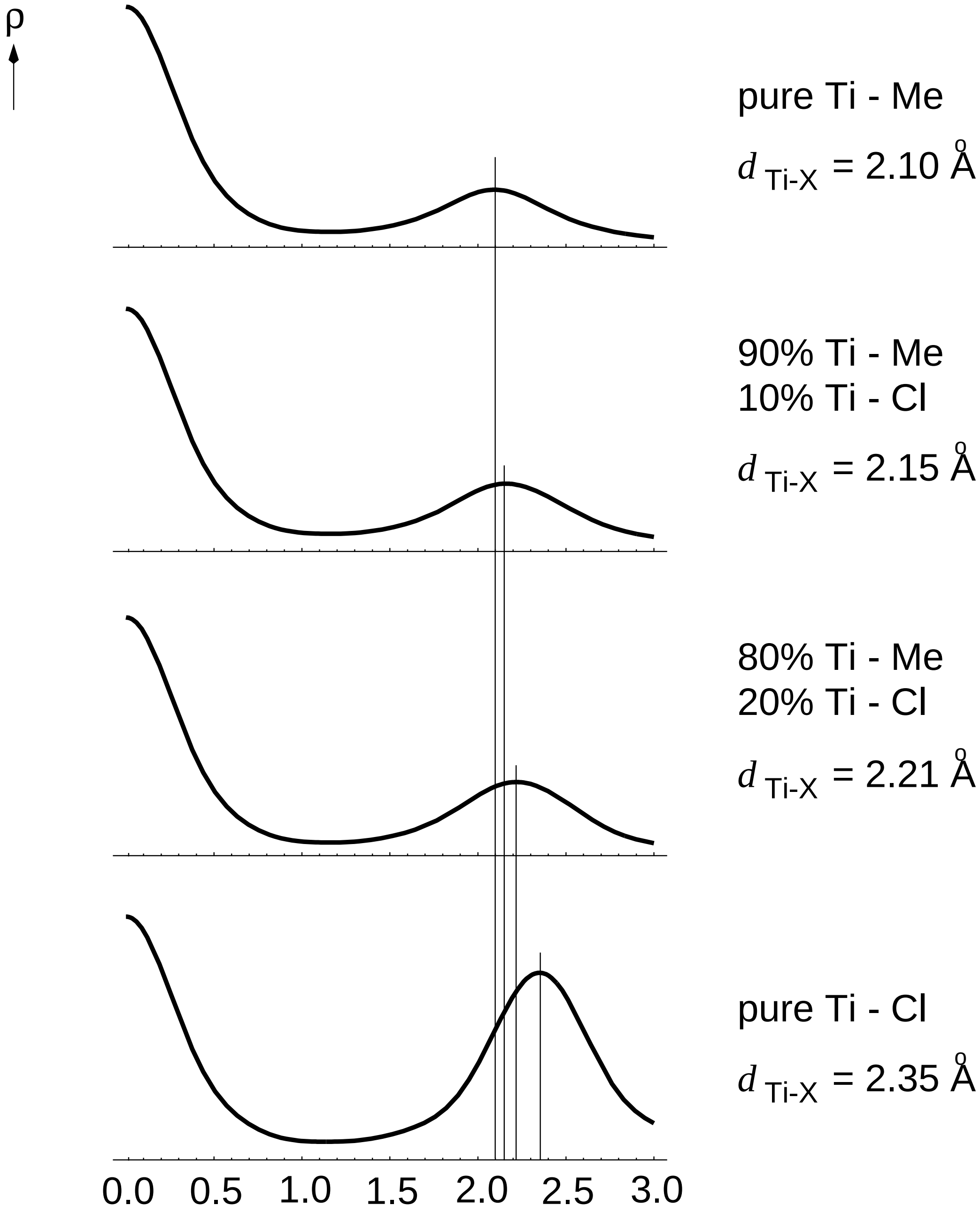
- Wrong assignment of scattering factors
  It can be difficult to discern between elements with approximately equal atomic numbers, e.g. N and C; Fe and Co or even Zn and Mo or Cu and Br.



- Refinement to a local minimum
  A local minimum with wrong, but not necessarily unrealistic geometries and good figures of merit can be reached during refinement. This is not easy to detect. Refinement against $|F|^2$ instead of $|F|$ reduces this risk.

Wrong structures (2)

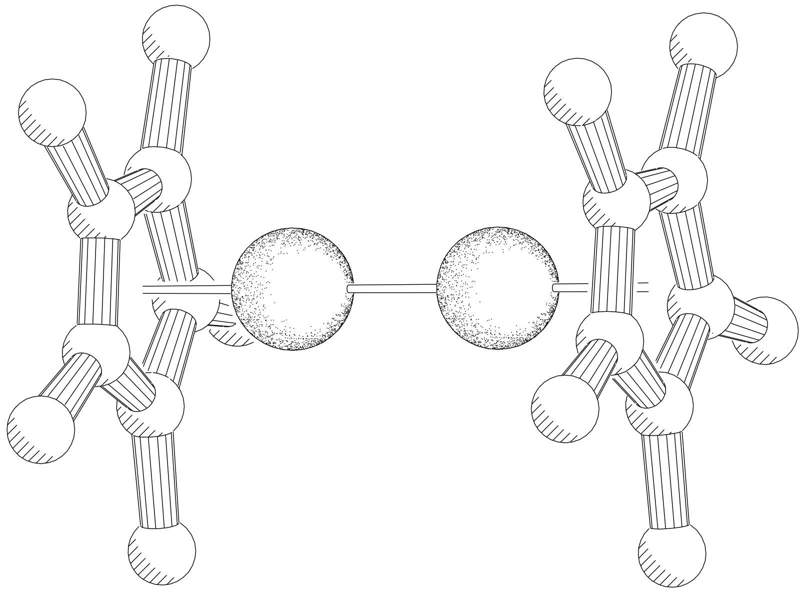
- Overlooked substitutional disorder
  This type of disorder is the cause of so-called bond stretch isomerism, the occurrence of different bond lengths in chemically equivalent systems.



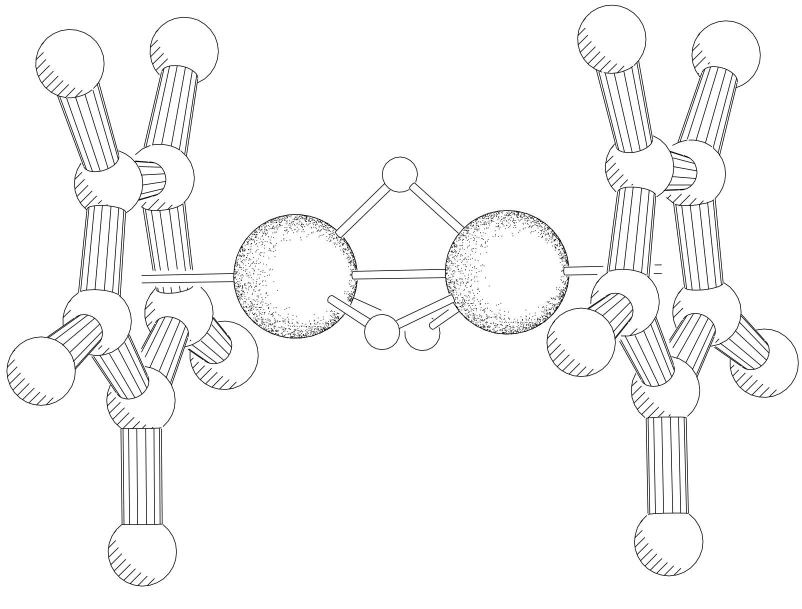
Wrong structures (3)

- Incorrect hydrogen atoms
  Due to the low value of the scattering factor, hydrogen atoms are normally introduced at calculated positions. In a number of cases too few or too many hydrogen atoms have been introduced, especially on nitrogen atoms. Even if the number of H atoms is correct, they can still be in the wrong position (hydroxyl groups).

  Hydrides are very easy to miss. The compound thought to be Cp*—Co=Co—Cp* (R1=0.040):



was later shown to be Cp*—Co(μ_2-H)$_3$Co—Cp* (R1=0.030):



Wrong structures (4)

- Wrong space group assignment
  - Missed crystallographic symmetry elements cause a large increase in the number of parameters needed to describe the structure. The data-parameter ratio is unnecessarily reduced. Mathematical problems (numerical instabilities) during least-squares refinement can lead to distorted structures.
  - Treating local symmetry as crystallographic symmetry amounts to the averaging of slightly miss-aligned molecules, resulting in distorted geometries and unrealistic anisotropic displacement parameters.
  - Interchange of crystallographic and local inversion centres (so-called pseudo inversion centres) or other symmetry elements can result in a disordered structure. This problem is sometimes referred to as a wrong origin choice. The interchange is made during structure determination, and can normally not be remedied by refinement.
- Overlooked twinning
  Some cases of disorder are in fact caused by twinning. A correct description of the twin (which is often more difficult to find than the disorder model) gives a more accurate, ordered structure, without unrealistic geometries and intermolecular interactions.

Validation

Automated validation checks can be performed on the Crystallographic Information File (.cif). Full validation also uses a listing of observed and calculated structure factors (.fcf, CIF format).

Tests are performed to check completeness, internal consistency and crystallographic validity of the CIF.

When tests are not satisfied, an alert is issued. An alert has format:

category_ALERT_type_level

item	class	description
category	0nn	General issues
	1nn	Cell and symmetry
	2nn	Anistropic displacement parameters
	3nn	Intramolecular geometry
	4nn	Intermolecular geometry
	5nn	Co-ordination geometry
	6nn	Void tests
	7nn	Varia
	8nn	Fatal Errors
	9nn	Reflection data issues
type	1	CIF construction/syntax error, inconsistent or missing data
	2	Indicator that the structure model may be wrong or deficient
	3	Indicator that the structure quality may be low
	4	Cosmetic improvement, methodology, query or suggestion
level	A	in general: serious problem
	B	potentially serious problem
	C	check and explain

Validation report

Validation can be performed with PLATON or on line via the International Union of Crystallography (www.iucr.org).

Example of a validation report:

```
#===============================================================================
# PLATON/CHECK-(110303) versus check.def version of 110303   for entry: 1
# Data From: m0981.cif - Data Type: CIF          Bond Precision   C-C = 0.0065 A
#
# Cell    8.137(2)    7.519(2)   31.439(8)           90    95.535(8)          90
# WaveLength 0.71073    Volume  Reported    1914.5(8)  Calculated    1914.5(8)
# SpaceGroup from Symmetry P 21          Hall: P 2yb
#                 Reported P 21                P 2yb
# MoietyFormula C23 H29 N O2
#      Reported C23 H29 N O2
#    SumFormula C23 H29 N O2
#      Reported C23 H29 N O2
# Mr        =    351.47[Calc],    351.47[Rep]
# Dx,gcm-3  =     1.219[Calc],     1.219[Rep]
# Z         =         4[Calc],         4[Rep]
# Mu (mm-1) =     0.077[Calc],     0.077[Rep]
# F000      =     760.0[Calc],     760.0[Rep]  or F000’ =    760.31[Calc]
# Calculated T limits: Tmin=0.984 Tmin’=0.977  Tmax=0.998
# Reported   Hmax=  9, Kmax=  9, Lmax= 37, Nref= 3787       , Th(max)= 25.42
# Calculated Hmax=  9, Kmax=  9, Lmax= 37, Nref= 3825( 7071), Ratio= 0.99( 0.54)
# R= 0.0574( 2770), wR2= 0.1561( 3787), S = 1.071, Npar= 469
#===============================================================================

          >>> The Following ALERTS were generated <<<
          -------------------------------------------
     Format: alert-number_ALERT_alert-type_alert-level text

028_ALERT_3_C _diffrn_measured_fraction_theta_max  Low .....       0.99
340_ALERT_3_C Low Bond Precision on C-C bonds (x 1000) Ang..          7
#===============================================================================

               ALERT_Level and ALERT_Type Summary
               ==================================
  2 ALERT_Level_C = Check & Explain

  2 ALERT_Type_3 Indicator that the Structure Quality may be Low.
#===============================================================================
```

Examples of validation alerts

Some examples of validation alerts. The column “codes” lists category, type and level of the alert.

codes	message
028,3,A	_diffrn_measured_fraction_theta_max Low 0.75
051,1,A	Mu(calc) and Mu(CIF) Ratio Differs from 1.0 by 35.25 Perc.
053,1,A	Minimum Crystal Dimension Missing (or Error) . ?
080,2,A	Maximum Shift/Error 1.36
306,2,A	Isolated Oxygen Atom (H-atoms Missing ?) ..... O(508)
601,2,A	Structure Contains Solvent Accessible VOIDS of 393.00 A**3
773,2,A	Suspect C-C Bond in CIF C522 - C525 = 1.74 Ang.
082,2,B	High R1 Value 0.16
094,2,B	Ratio of Maximum / Minimum Residual Density .. 4.74
241,2,B	Check High U(eq) as Compared to Neighbors .. O(510)
068,1,C	Reported F000 Differs from Calcd (or Missing). ?
120,1,C	Reported SPGR ? Inconsistent with Explicit P21/C
141,4,C	su on a - Axis Small or Missing (x 100000) ... 30 Ang.
230,2,C	Hirshfeld Test Diff for O(507) - C(538) = 5.13 su
430,2,C	Short Inter D...A Contact: O(506) .. O(508) = 2.88 Ang.

Currently, a CIF validation involves 314 different tests.

Structure factor validation

There is a routine ("FCF-VALID") to separately validate a structure factor file with PLATON.

Example of a structure factor validation report:

```
================================================================================
PLATON/ASYM-(Version  110303) FCF-FILE Validation for:global
================================================================================
Crystal Data From:m0981.cif
Fo/Fc   Data From:m0981.fcf FCF-TYPE=SHELXL
Space Group      :P21
Wavelength (Ang) :   0.71073
Unit Cell        :   8.1370   7.5190  31.4390   90.000   95.535   90.000

================================================================================
Reflections with Large abs((I(obs) - I(calc)) / sigma(I)) .GT. 10.0
================================================================================
   Nr       H    K    L  Theta       I(obs)      I(calc)    sig(I)   Ratio
--------------------------------------------------------------------------------
    1       3    1    0   8.04       984.39       269.59     54.83   13.04
    2       3    2    0   9.33       321.11       224.41      7.94   12.18
    3      -1    1    1   3.71     20695.81     24155.24    251.49  -13.76
    4       1    1    1   3.80     30991.49     35985.79    424.42  -11.77
    5       3    2    1   9.40       715.30       549.31     16.35   10.15
(...)
   21       0    1    4   3.76       461.76       525.82      5.62  -11.40
   22      -1    1    5   4.77       635.63       751.92      9.92  -11.72
   23       2    1    5   6.82      1421.55      1625.24     18.26  -11.15
   24      -1    2    5   6.70       698.03       786.23      8.17  -10.80
   25       2    1    7   7.62       306.45       362.95      5.23  -10.80
                                                                  --------
                                                         Average =    3.37

For I(calc) < 2 sig(I): <I(obs)> =      15.80 and <I(calc)> =      13.45

================================================================================
Missing Reflections (Asymmetric Reflection Unit) below sin(th)/lambda = 0.5
================================================================================
   Nr       H    K    L    sin(th)/lambda   Theta     I(calc)  I(calc)/I(max)
--------------------------------------------------------------------------------
    1       0    0    1        0.016         0.65        2.38      0.0001
    2      -8    0    5        0.493        20.50       53.41      0.0015
```

(continued on next page)

Structure factor validation (2)

(continued from previous page)

```
================================================================================
Resolution & Completeness Statistics (Cumulative)
================================================================================
Theta sin(th)/Lambda Complete  Expected Measured  Missing
--------------------------------------------------------------------------------
 20.82     0.500     0.999         2200     2198        2
 23.01     0.550     0.998         2910     2905        5
 25.24     0.600     0.997         3748     3737       11
 27.51     0.650     0.989         3829     3787       42

================================================================================
R-Value Statistics as a Function of Resolution (in Resolution Shell)
================================================================================
Theta sin(Th)/Lambda  #     R1    wR2      S av(I/SIG)     av(I) av(SIG)
--------------------------------------------------------------------------------
  4.08       0.10    23  0.070  0.140  7.220   45.07     3625.93   64.18
  6.12       0.15    50  0.077  0.111  5.830   46.21     3079.66   70.46
  8.17       0.20    83  0.059  0.097  5.534   49.79     1174.48   23.22
 10.23       0.25   141  0.039  0.079  4.273   47.63      826.35   17.02
 12.31       0.30   208  0.037  0.078  3.587   40.69      587.33   12.05
 14.40       0.35   278  0.043  0.077  2.402   25.50      242.35    7.67
 16.52       0.40   363  0.048  0.077  2.006   20.33      206.95    8.29
 18.65       0.45   471  0.047  0.075  1.719   17.58      227.66   10.45
 20.82       0.50   579  0.082  0.106  1.386    9.81      126.13   12.00
 23.01       0.55   707  0.151  0.162  0.980    4.28       66.54   15.25
 25.24       0.60   832  0.279  0.272  0.888    2.18       42.49   19.88
 27.51       0.65    50  0.451  0.372  0.893    1.35       38.20   31.27

================================================================================
Summary of Reflection Data in FCF
================================================================================
Total  (in FCF) ..............   3787  (Hmax =  9, Kmax =  9, Lmax = 37)
Actual Theta(max) (Deg.) .....  25.42  (Hmax =  9, Kmax =  9, Lmax = 37)
Actual Theta(min) (Deg.) .....   1.30

Unique (Expected) ............   3829  (HKL  3829, -H-K-L     0)
Unique (in FCF) ..............   3787  (HKL  3787, -H-K-L     0)
Observed [I .GT. 2 sig(I)] ...   2770  (HKL  2770, -H-K-L     0)
Less-Thans ...................   1017  (HKL  1017, -H-K-L     0)
Missing ......................     42  (HKL    42, -H-K-L     0)

Space Group Extinctions ......      5
```

Index