

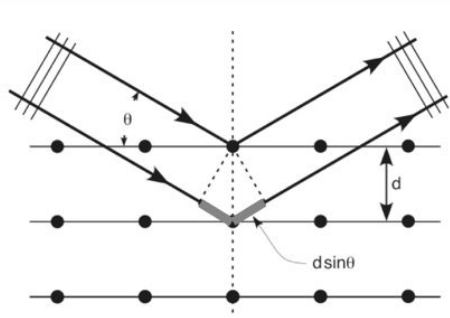
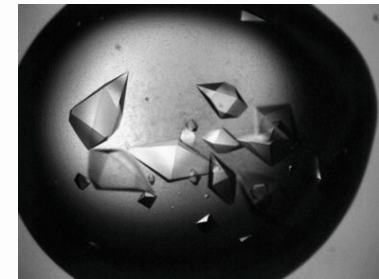
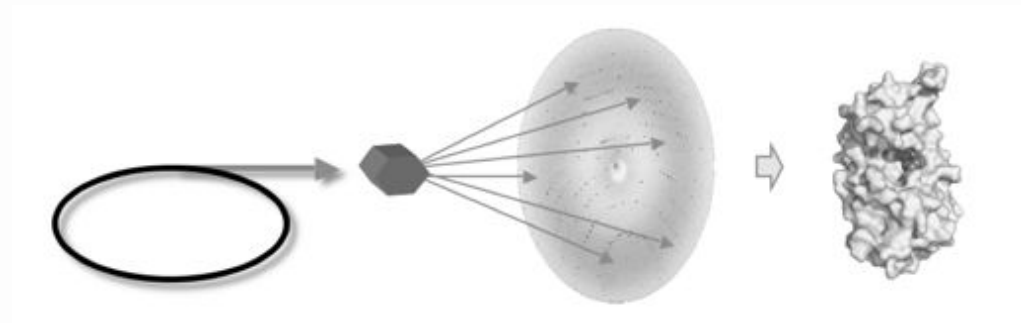
Biology II: MX – The Method

Thomas R. Schneider, EMBL Hamburg
30/6/2015

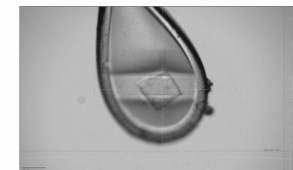
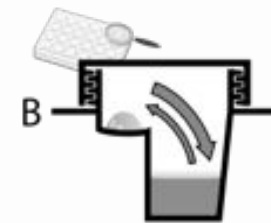
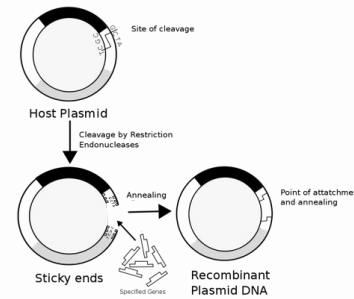
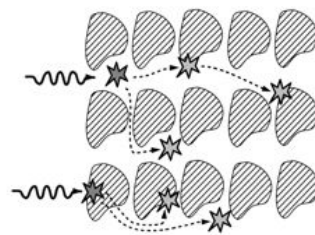
thomas.schneider@embl-hamburg.de

Recap

- MX
- Proteins
- Protein crystals
- Recombinant production of proteins
- Radiation damage and cryogenic cooling
- Bragg's law



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



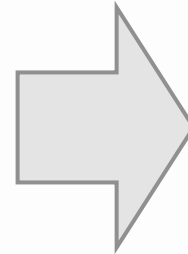
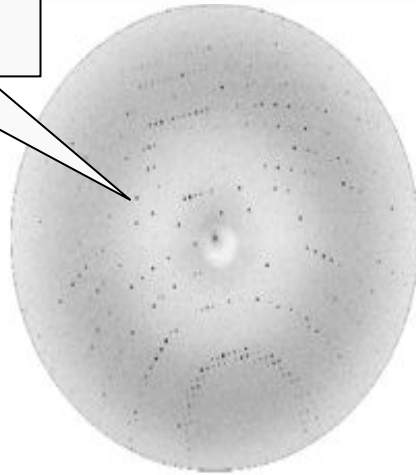
Today

- Electron densities and structure factors
/ real space and reciprocal space
 - Fourier transforms
 - Argand diagrams
- The crystallographic phase problem
- Symmetry in real and reciprocal space

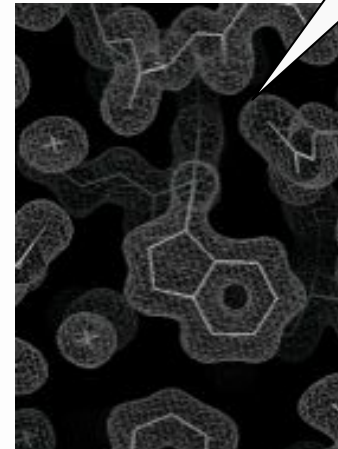
- Solving the crystallographic phase problem
- Anomalous scattering from a crystallographer's point of view

Calculating Electron density

Intensity I_{hkl} for some set of planes with Miller indices hkl



Electron density for all points in the unit cell



$$\rho_{xyz} = \sum_{hkl} |F_{hkl}| e^{-i\varphi_{hkl}} e^{-2\pi i(hx+ky+lz)}$$

Electron density at some point xyz in space

Sum over all Bragg reflections

Structure Factor Amplitude for each reflection hkl . $F_{hkl} \approx \sqrt{I_{hkl}}$

Structure Factor Phase for each reflection hkl .

Phase shift depending on hkl and position in space

Fourier Transforms

Fourier Transform

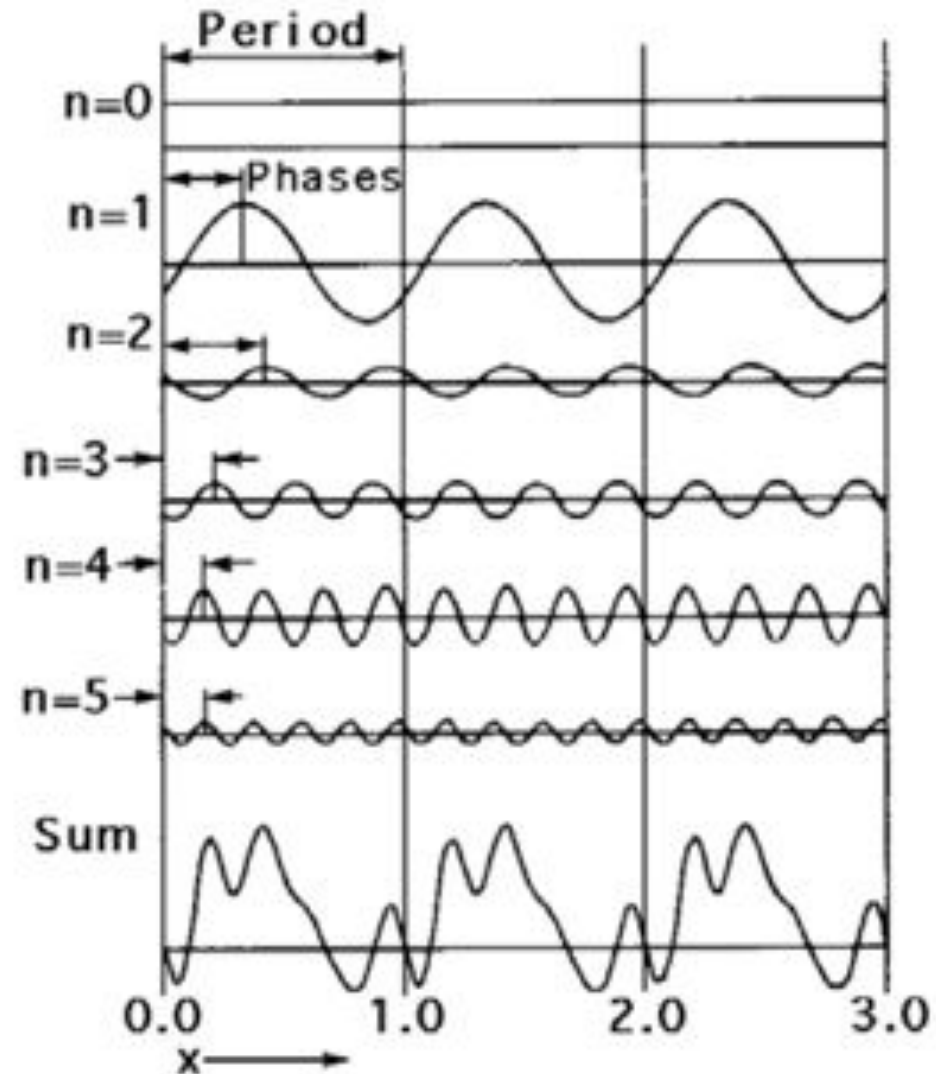
- Any periodic function can be considered as a sum of sinusoidal waves.

$$f(x) = \sum_{n=0}^{\infty} A_n \cos(nx) + B_n \sin(nx)$$

- Or:

$$f(x) = \sum_{n=0}^{\infty} C_n \cos(nx - \varphi_n)$$

- The above sum is called a 'Fourier series'
- For higher n (i.e. shorter period) finer features are generated.



<http://mathworld.wolfram.com/FourierSeries.html>

Fourier Transform

- Any periodic function can be considered as a sum of sinusoidal waves.

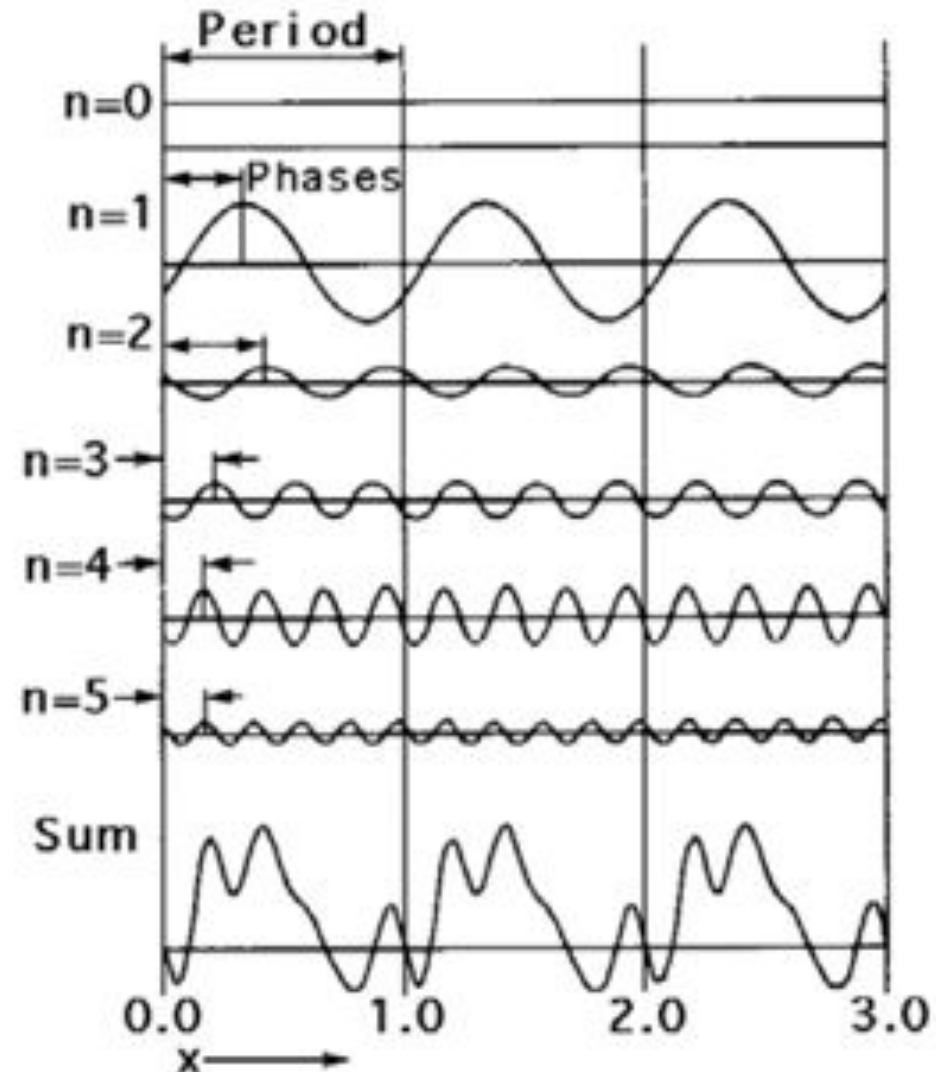
$$f(x) = \sum_n A_n \cos(nx) + B_n \sin(nx)$$

- To calculate the coefficients:

$$A_n = \int_{\text{period}} f(x) \cos(nx) dx$$

$$B_n = \int_{\text{period}} f(x) \sin(nx) dx$$

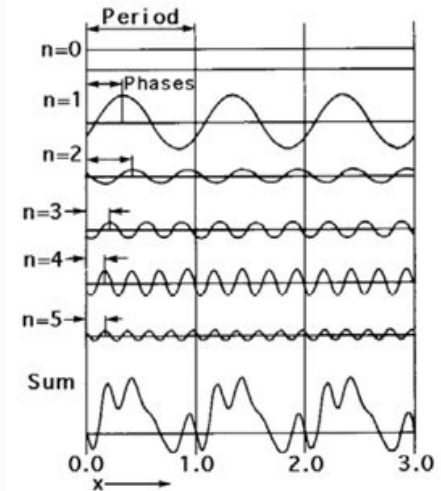
- The integral can be seen as the projection of $f(x)$ onto $\cos(nx)$ – 'scalar product'



Fourier Transform with complex numbers

- Fourier series using real numbers / function:

$$W(x) = \sum_{n=0}^{\infty} A_n \cos(nx) + B_n \sin(nx) = \sum_{n=0}^{\infty} C_n \cos(nx - \varphi_n)$$



- Periodic functions can be written more elegantly using complex numbers instead of sine and cosine functions

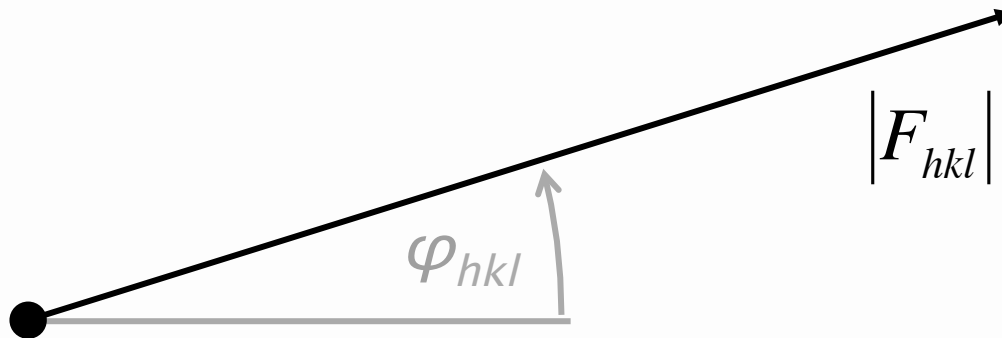
$$W(x) = \sum_{n=0}^{\infty} A_n \exp[i(2\pi nx + \varphi_n)] = \sum_{n=0}^{\infty} F_n \exp[i(2\pi nx)] = \sum_{n=0}^{\infty} |F_n| \exp[i\varphi_n] \exp[i(2\pi nx)]$$

Complex number

Electron density from Diffraction Data

- To evaluate the electron density in the crystal, we need to know the 'phase' φ_{hkl} for each reflection / complex structure factor

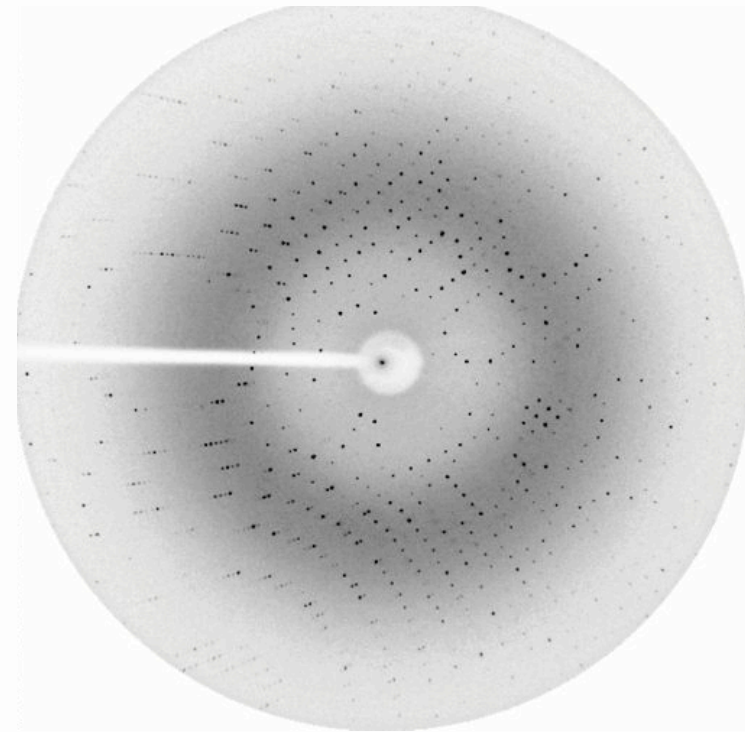
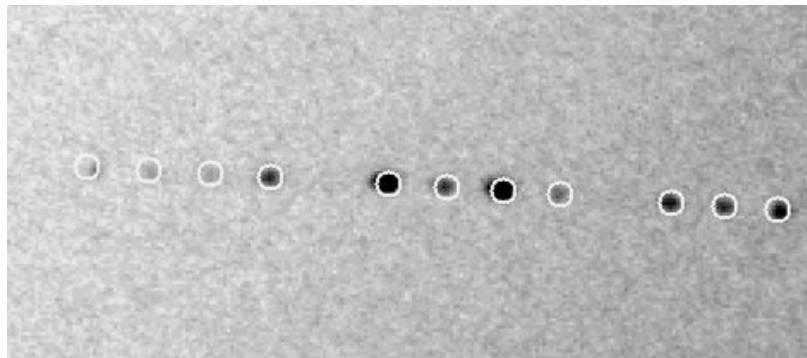
$$\rho_{xyz} = \sum_{hkl} |F_{hkl}| \exp(-i\varphi_{hkl}) \exp(-2\pi i(hx + ky + lz))$$



The Crystal as a Fourier-Analyzer

$$\rho_{xyz} = \sum_{hkl} |F_{hkl}| \exp(-i\varphi_{hkl}) \exp(-2\pi i(hx + ky + lz))$$

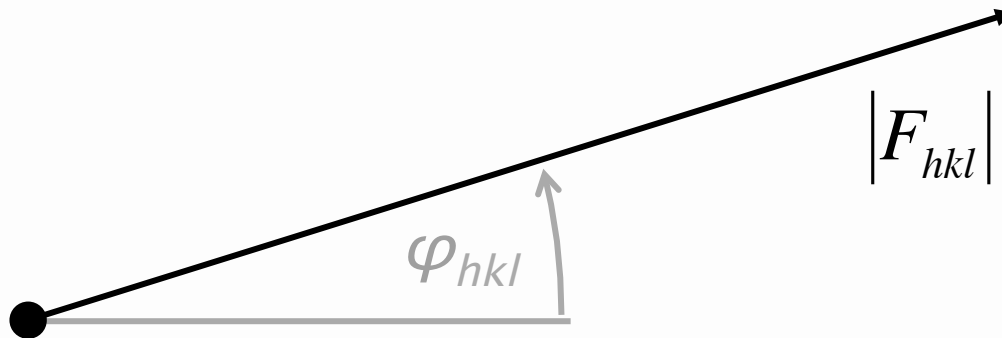
$$|F_{hkl}| = \sqrt{I_{hkl}}$$



Electron density from Diffraction Data

- To evaluate the electron density in the crystal, we need to know the 'phase' φ_{hkl} for each reflection / complex structure factor

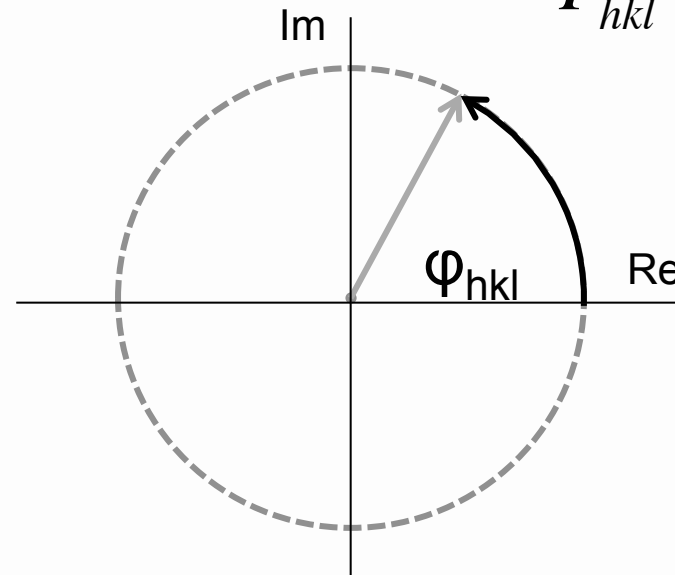
$$\rho_{xyz} = \sum_{hkl} |F_{hkl}| \exp(-i\varphi_{hkl}) \exp(-2\pi i(hx + ky + lz))$$



The Phase Problem - graphically

$$F_{hkl} = |F_{hkl}| e^{i\varphi}$$

h	k	l	$ F_{hkl} $	φ_{hkl}
0	20	35	4980.5	???
0	20	36	6906.6	???
0	20	37	8302.3	???
0	20	38	63209.5	???
0	20	39	459.6	???
0	20	40	1017.4	???
0	20	41	5.6	???
0	20	42	33.8	???
0	20	43	4545.7	???
0	20	44	0.2	???
0	20	45	808.8	???
0	20	46	72.2	???
0	20	47	792.3	???
.				
.				
.				

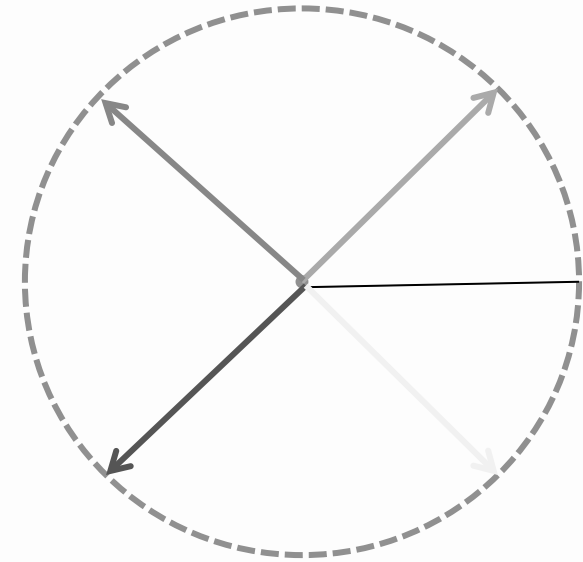


Argand Diagram

- The 'structure factor amplitude' $|F|$ can be anything positive
- The 'phase angle' φ varies between 0 and 360° .
- The pair of $|F|$ and φ can be collected in a complex number

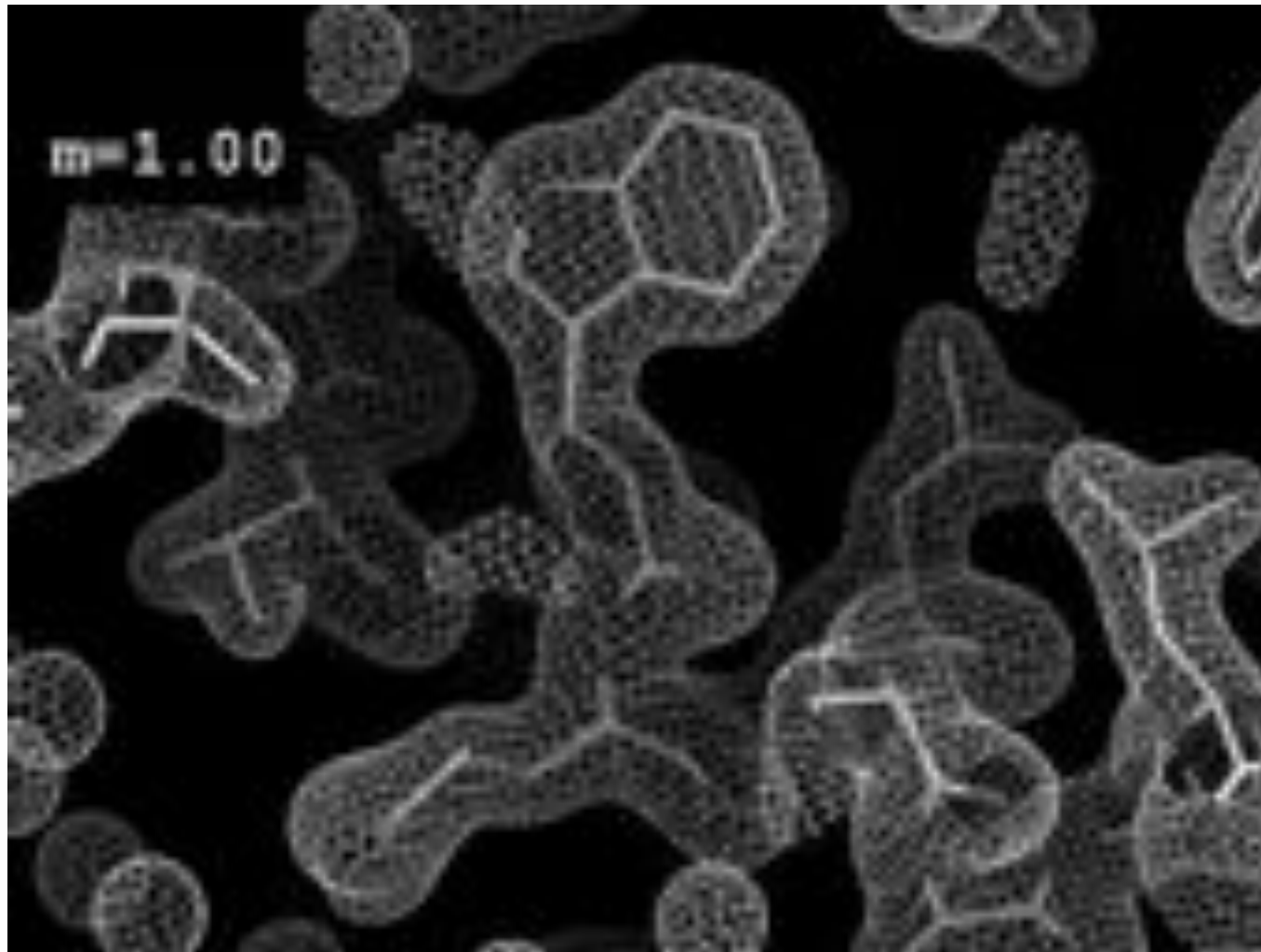
Let me do it ...

- Knowing the phase within 90 deg. is good enough
- 4 possibilities: 45, 135, 225, 315
- A piece of cake for our 1000 CPU cluster
- 200 reflections
- $4^{200} = 2.5 * 10^{120}$ possibilities
- Number of atoms in the universe 10^{80}
- Time since big bang $4.3 * 10^{17}$ sec
- 1 set per CPU cycle of 3 GHz = $12.9 * 10^{26}$ trials since the beginning

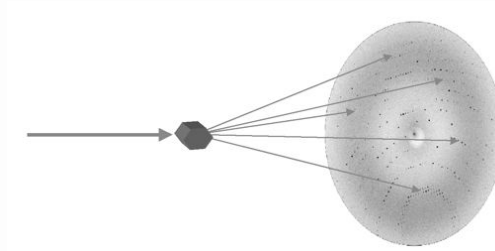


Phases

<http://ucxray.berkeley.edu/~jamesh/movies>



Diffraction from a Crystal



$$F_{hkl} = \sum_{j=1}^N f_j e^{2\pi i(hx_j + ky_j + lz_j)}$$

Complex Structure Factor for reflection hkl

Sum over all atoms in the unit cell

Scattering Factor of atom j. Depends on Z, B (and hkl)

Phase difference for the set of Bragg planes corresponding to reflection hkl

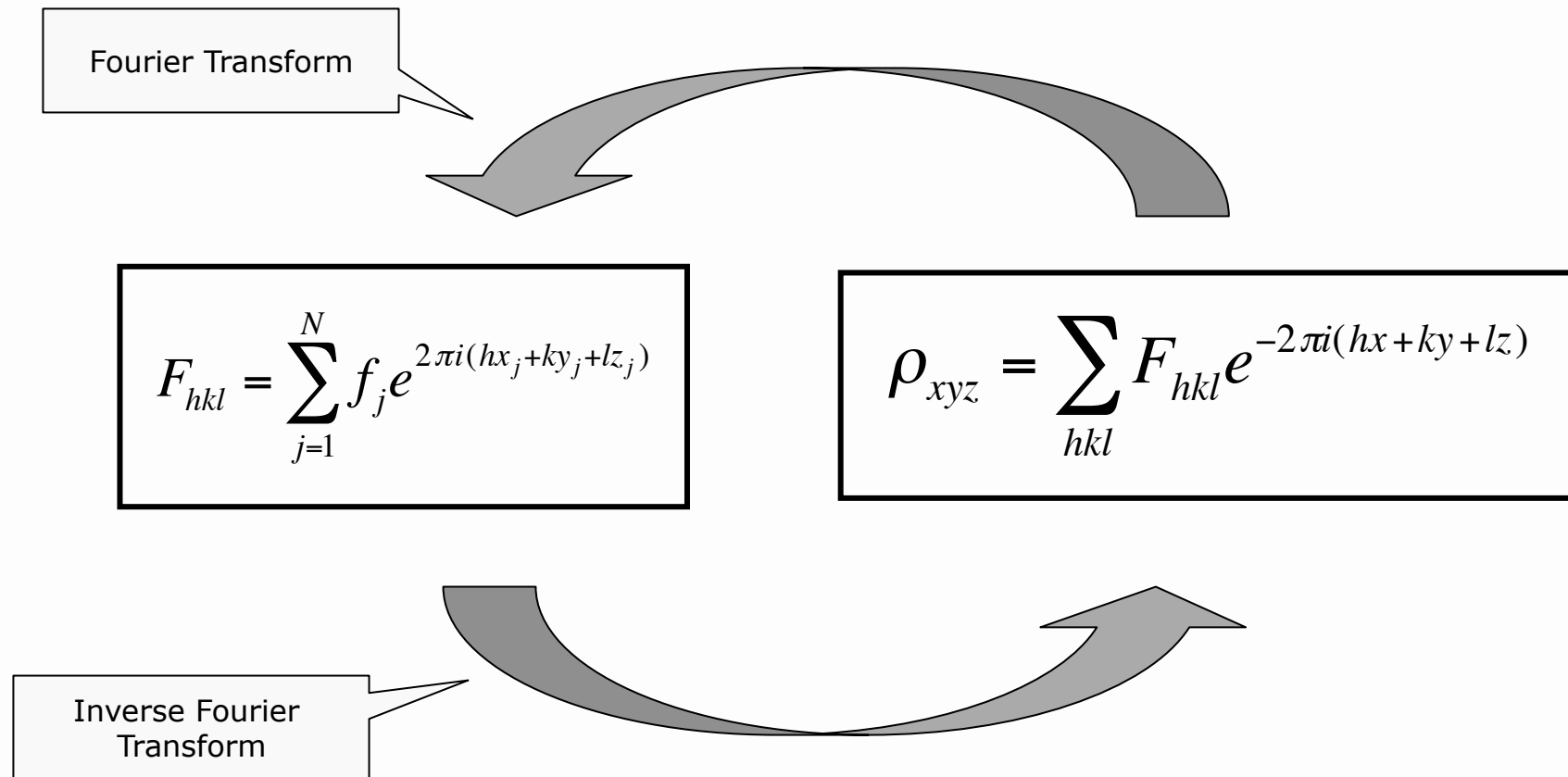
$$I_{hkl} = F_{hkl}^2$$

Real number

Source for Structure Factors: Stout and Jensen

Fourier Transforms

- Structure factors and electron density are related by FT and FT⁻¹



Fourier Transforms

- Equation for deriving diffraction from a crystal can be rewritten in terms of electron density

Fourier Transform



$$F_{hkl} = \int_V \rho_{xyz} e^{2\pi i(hx+ky+lz)} dV$$

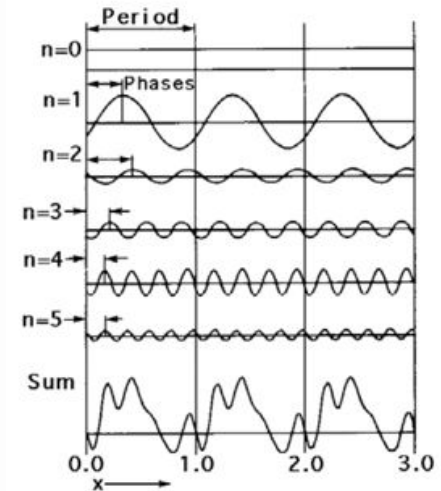
$$\rho_{xyz} = \sum_{hkl} F_{hkl} e^{-2\pi i(hx+ky+lz)}$$

Inverse Fourier Transform



Electron density and the Fourier transform

- The calculation of the electron density can be considered as adding 'base waves' with specified amplitude (structure factor amplitude) and phase:

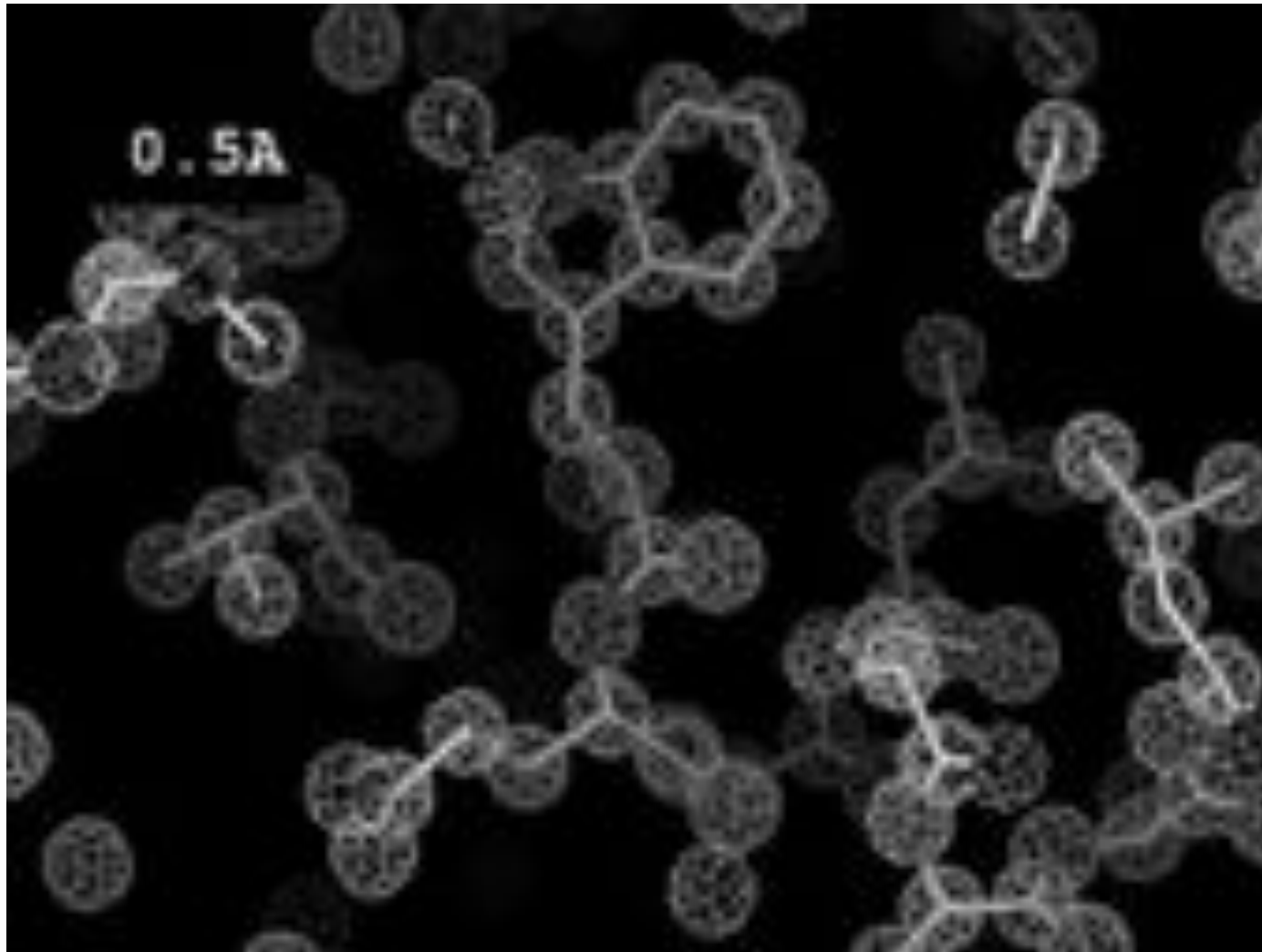


$$\rho_{xyz} = \sum_{hkl} F_{hkl} e^{-2\pi i(hx+ky+lz)} = \sum_{hkl} |F_{hkl}| e^{-i\varphi_{hkl}} e^{-2\pi i(hx+ky+lz)}$$

- Technically, this operation corresponds to an 'inverse' Fourier transform.
- The higher the resolution / higher the diffraction angle / smaller the lattice plane distance, the shorter is the wavelength of the corresponding hkl-term. Shorter wavelength correspond to finer features.

Resolution

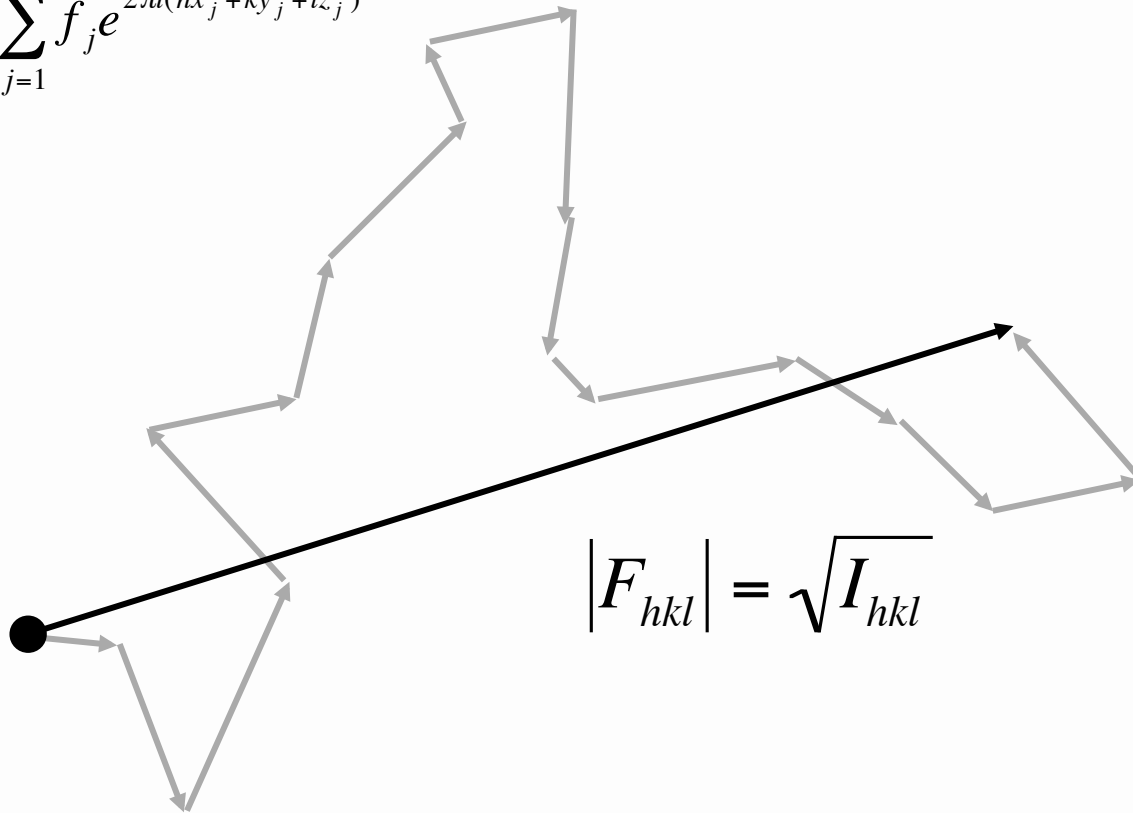
<http://ucxray.berkeley.edu/~jamesh/movies>



Diffraction from a protein (in vectors)

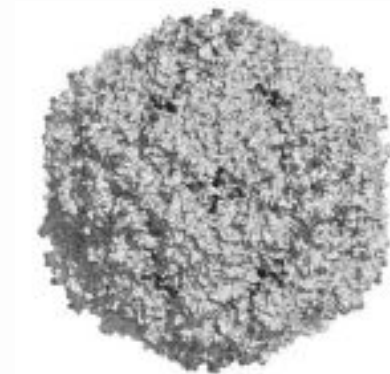
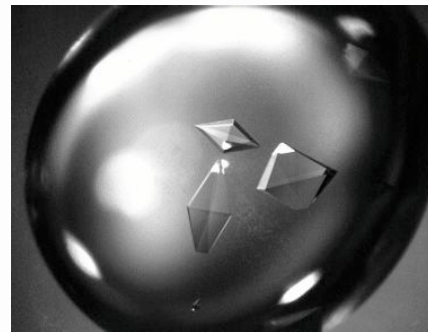
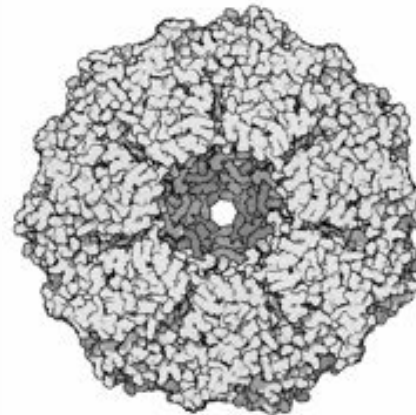
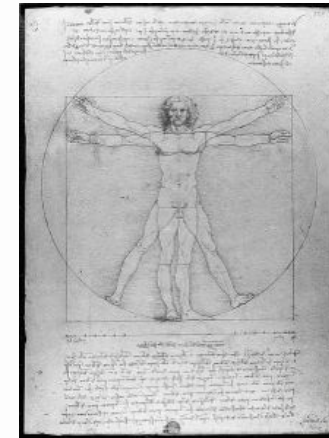
- To calculate the structure factor for a certain hkl based on the knowledge of the positions of the atoms in the unit cell, we sum over all atoms (here: 17 atoms):

$$F_{hkl} = \sum_{j=1}^N f_j e^{2\pi i(hx_j + ky_j + lz_j)}$$



Symmetry

- An object is symmetric if, after some operation has been carried out, the result is indistinguishable from the original object [Blow2007].



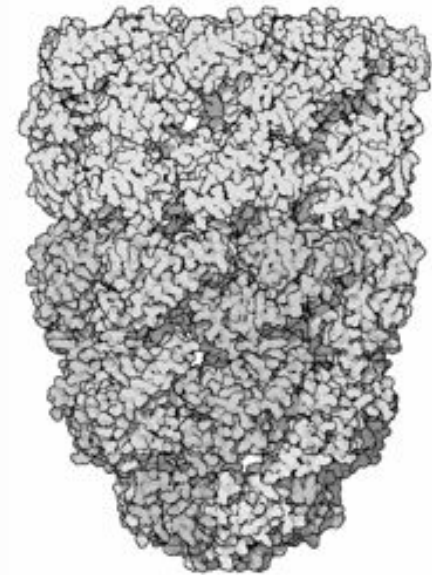
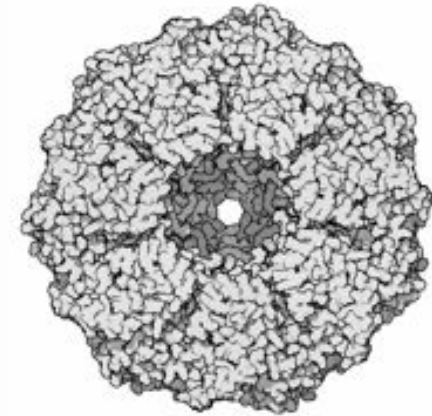
Rotational Symmetry

- After a rotation of 60 deg. about its centre, the picture of the snowflake is indistinguishable from the picture before the operation.
- The snowflake possesses 6-fold symmetry



<http://commons.wikimedia.org/wiki/File:Snowflake6.png>

Other n-fold symmetries



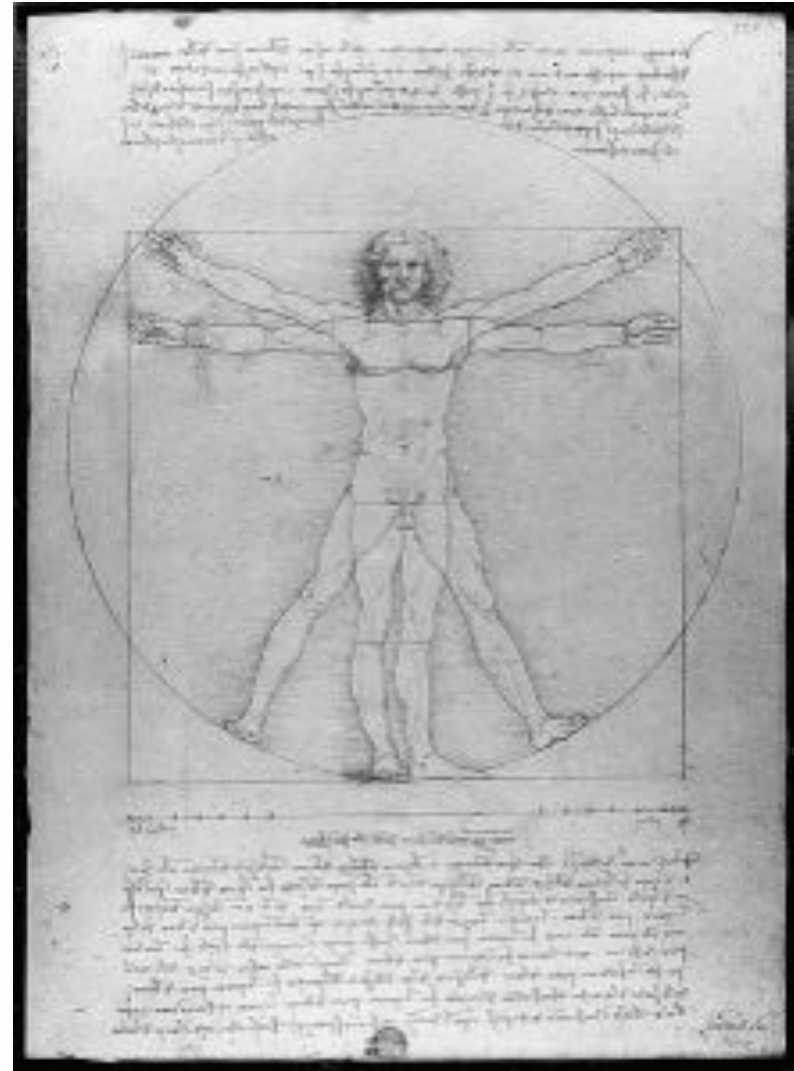
GROEL. <http://www.rcsb.org/pdb/101/motm.do?momID=32>

Mirror symmetry

- One version of the object is related to the other one by reflection through a plane (the mirror)



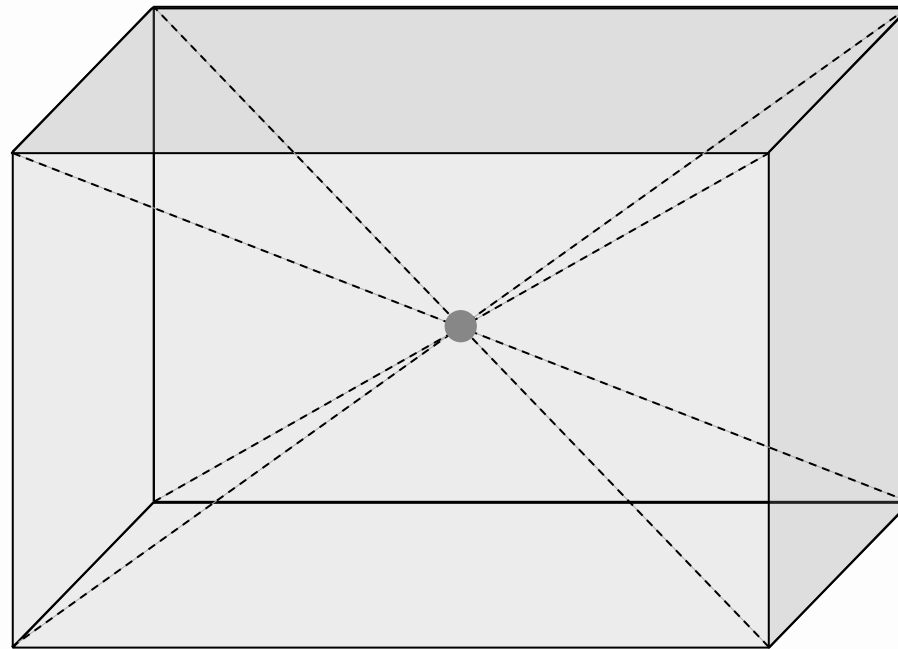
[http://commons.wikimedia.org/wiki/
File:Schmetterling_1a_neucc.jpg](http://commons.wikimedia.org/wiki/File:Schmetterling_1a_neucc.jpg)



[http://
commons.wikimedia.org/
wiki/
File:Da_Vinci_Vitruve_Luc_
Viatour.jpg](http://commons.wikimedia.org/wiki/File:Da_Vinci_Vitruve_Luc_Viatour.jpg)

Centrosymmetry

- When the object created by moving every point in the object along the line joining it to a chosen centre point, and continuing along this line until it is equally far from the centre on the other side, is identical to the original object, the object is centrosymmetric, or symmetric under inversion in the point.
- Example: rectangular block



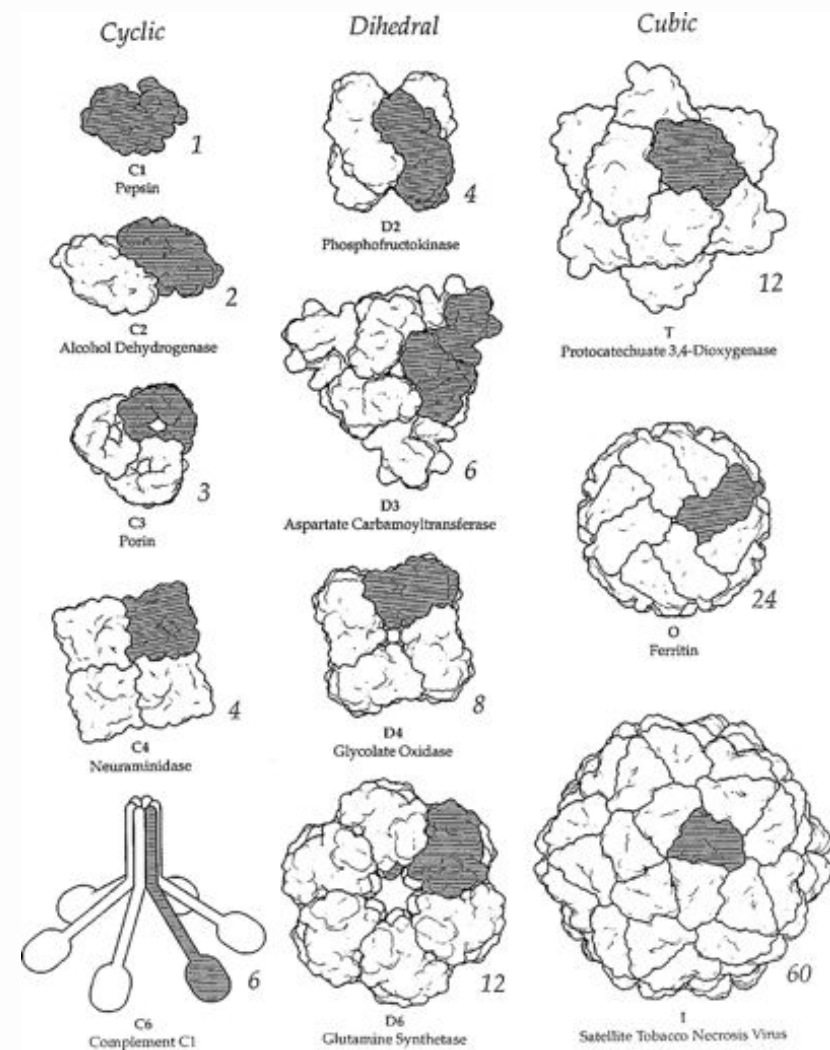
Centrosymmetry and Biology

- Would you expect centrosymmetric structures in biology?
- **No, because natural proteins are made (almost exclusively) of L-amino acids. By inversion, one would obtain D-amino acids.**

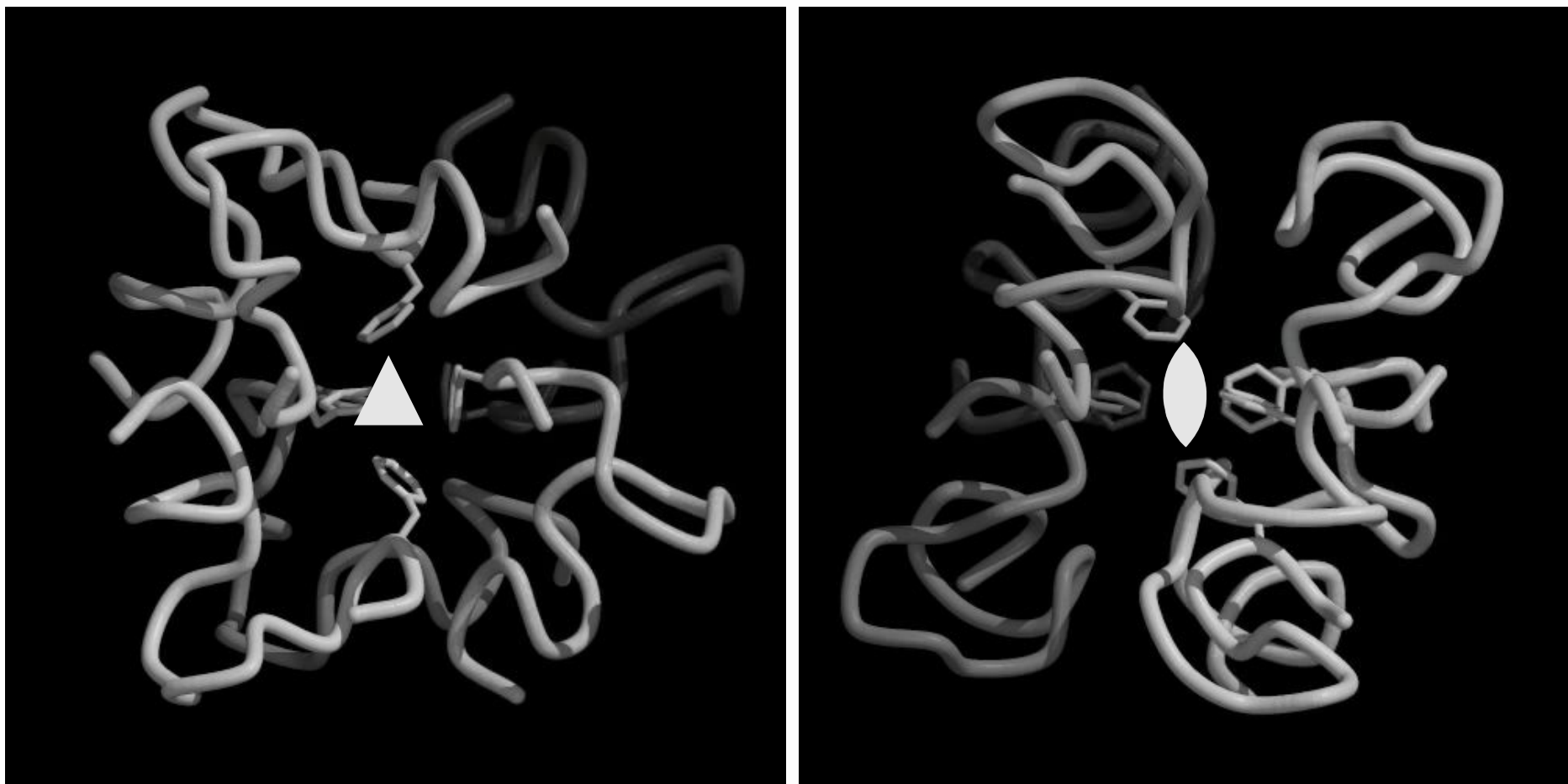
Point Groups

Goodsell (2000) Annu. Rev. Bioph. Biom. 29:105-153

- groups of symmetries that leave one point fixed are called 'point groups'
- rotations and mirrors are possible
- There are a total of 32 crystallographic point groups
- Local assemblies of objects (oligomers of protein molecules) can follow point group symmetry
- Examples: 1, 4, 422, 2, 222, 6, 622, 3, 322, 23, 432 [Blow2007, p27]
- These are the 11 enantiomorphic point groups accessible for chiral objects (such as proteins)

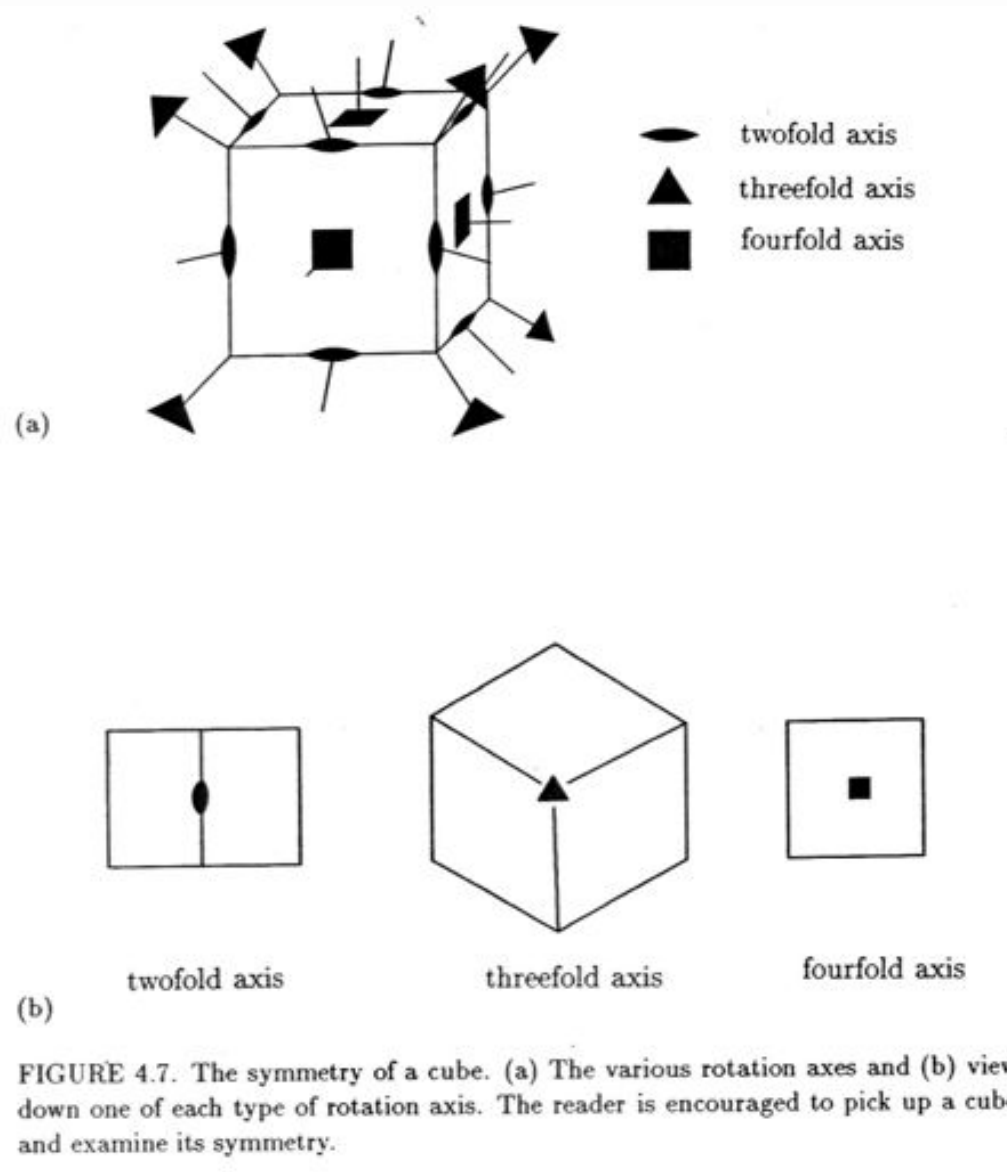


Mersacidin: NCS



Some objects ...

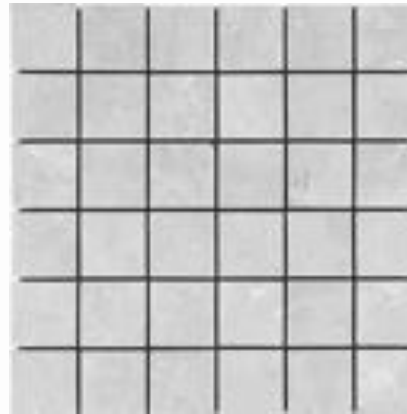
Symmetry of a cube



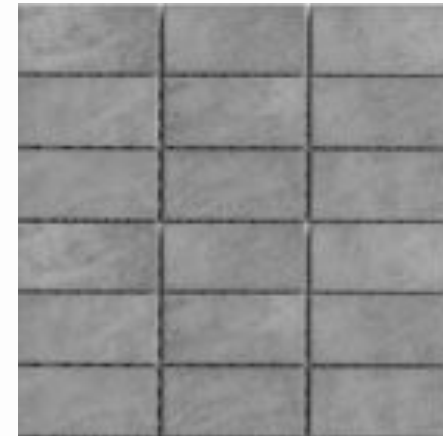
How can we build a crystal from these symmetric objects?

In 2D: Tiling

- Invariance of a system under translation
- Parallelogram is the general shape that can be used.



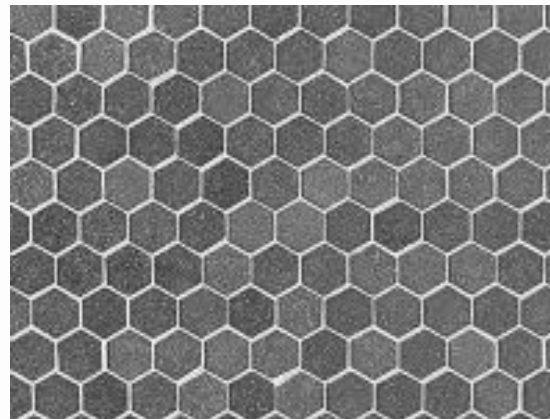
4 - square



2 - rectangular



6 - hexagonal

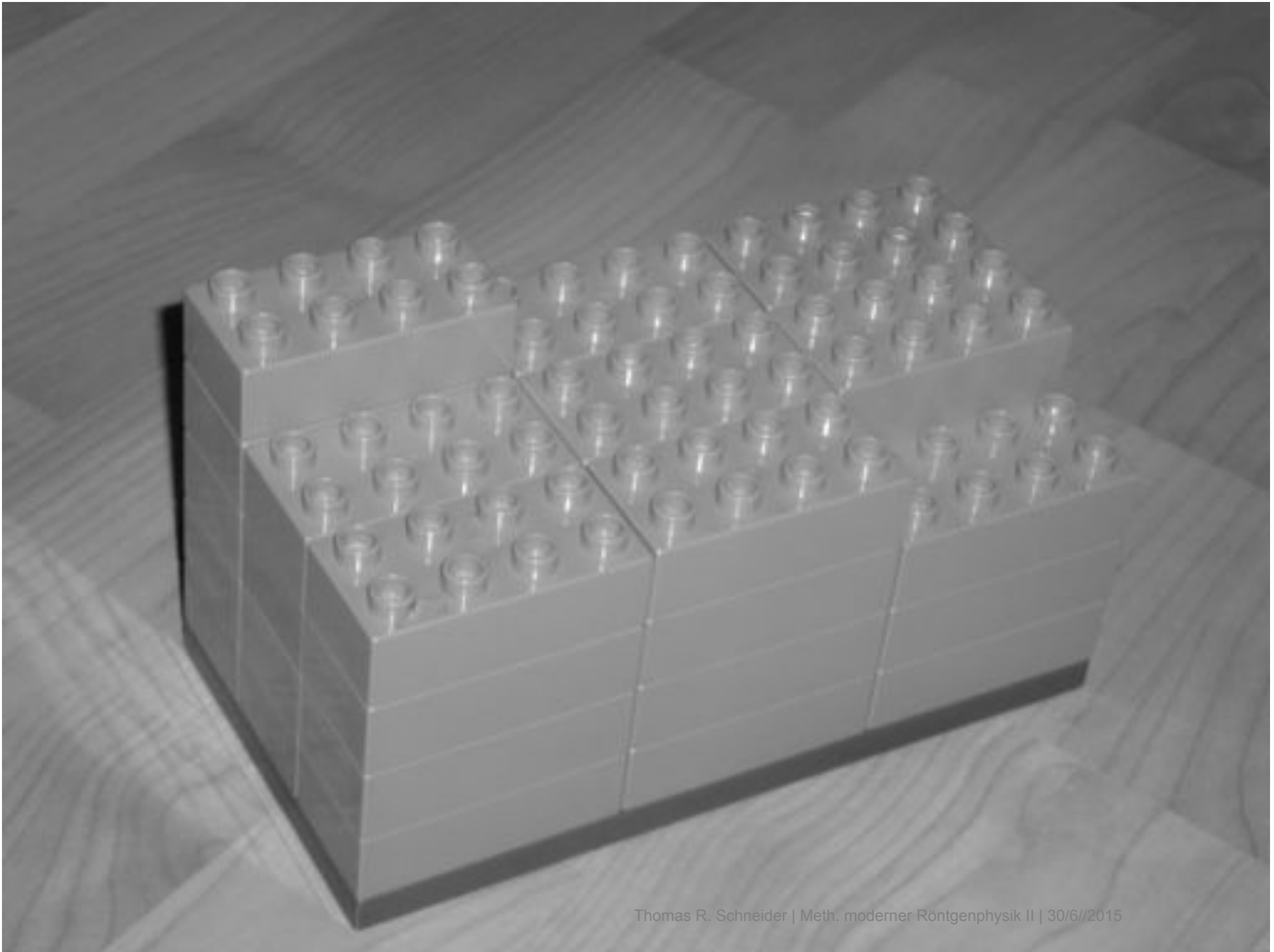


2 - rhombohedral

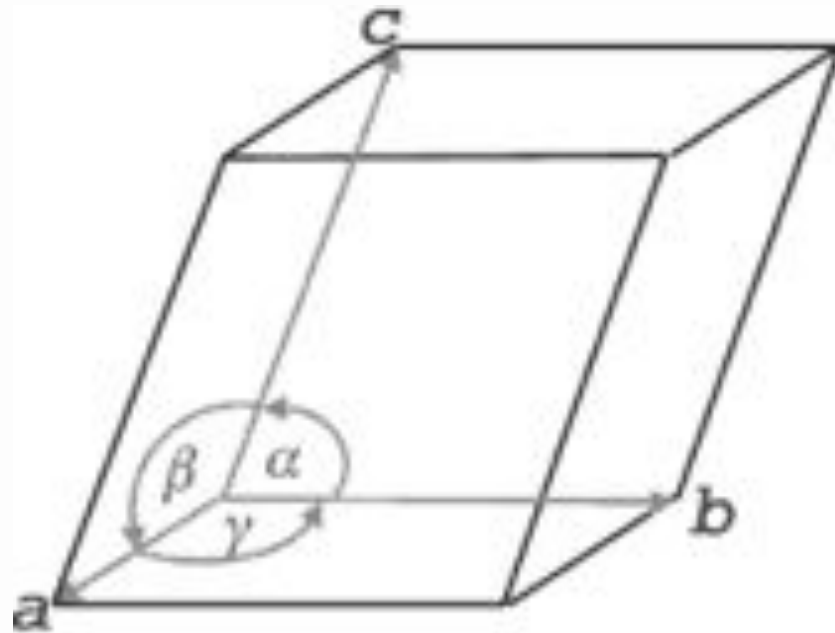
In 3D: Building crystals from unit cells

- A crystal is a homogeneous solid formed by a repeating three-dimensional pattern of atoms, ions, or molecules.
- In terms of translational symmetry, the repeating unit is the unit cell

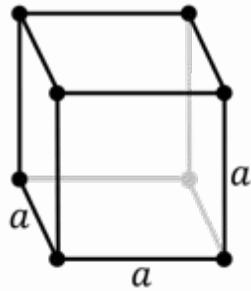




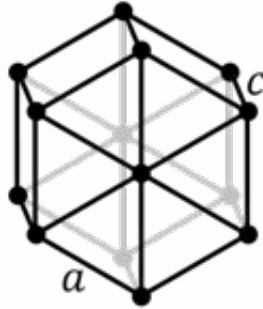
The general building block for a crystal is a parallelepiped



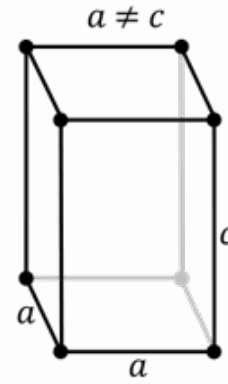
Lattice systems (primitive)



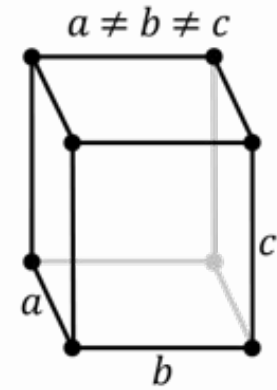
cubic



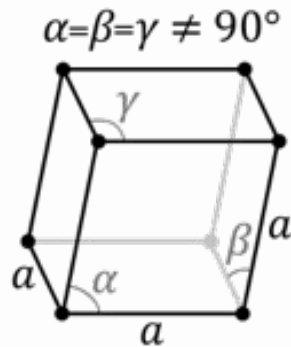
hexagonal



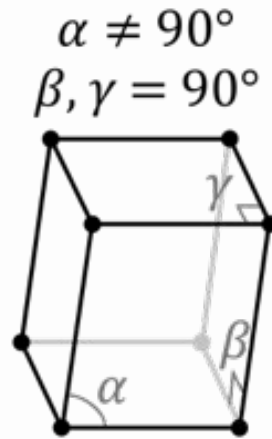
tetragonal



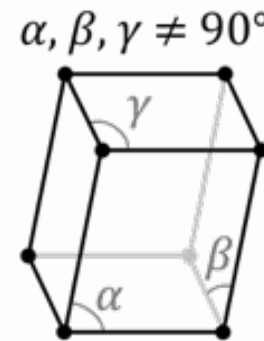
orthorhombic



rhombohedral



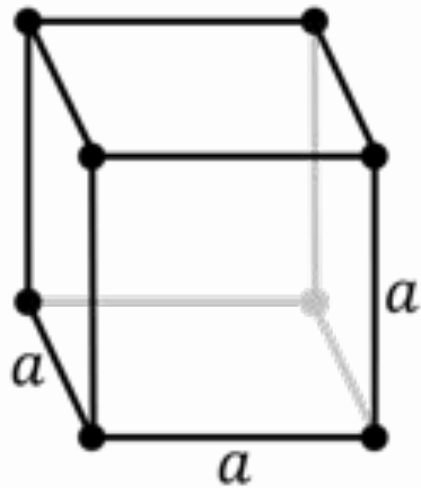
monoclinic



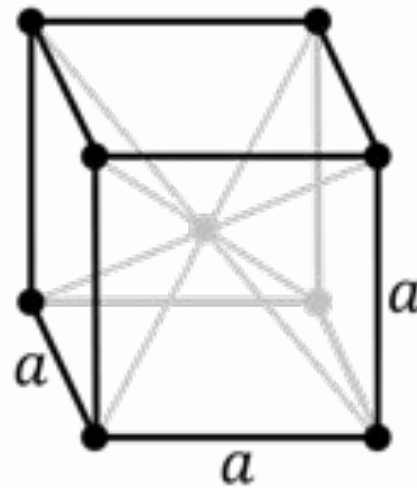
triclinic

Centered lattices

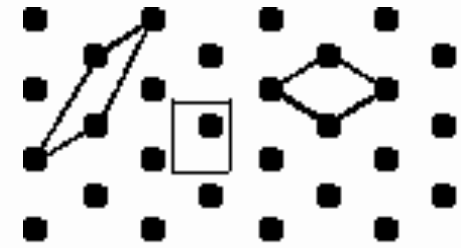
- Sometimes, it is more convenient to use a centered cell (lattice points on faces or inside the cell, to reflect the symmetry properly).
- This gives rise to a total of 14 Bravais lattices



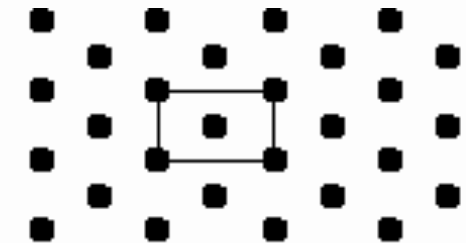
cubic (P432)



cubic body centered (I432)

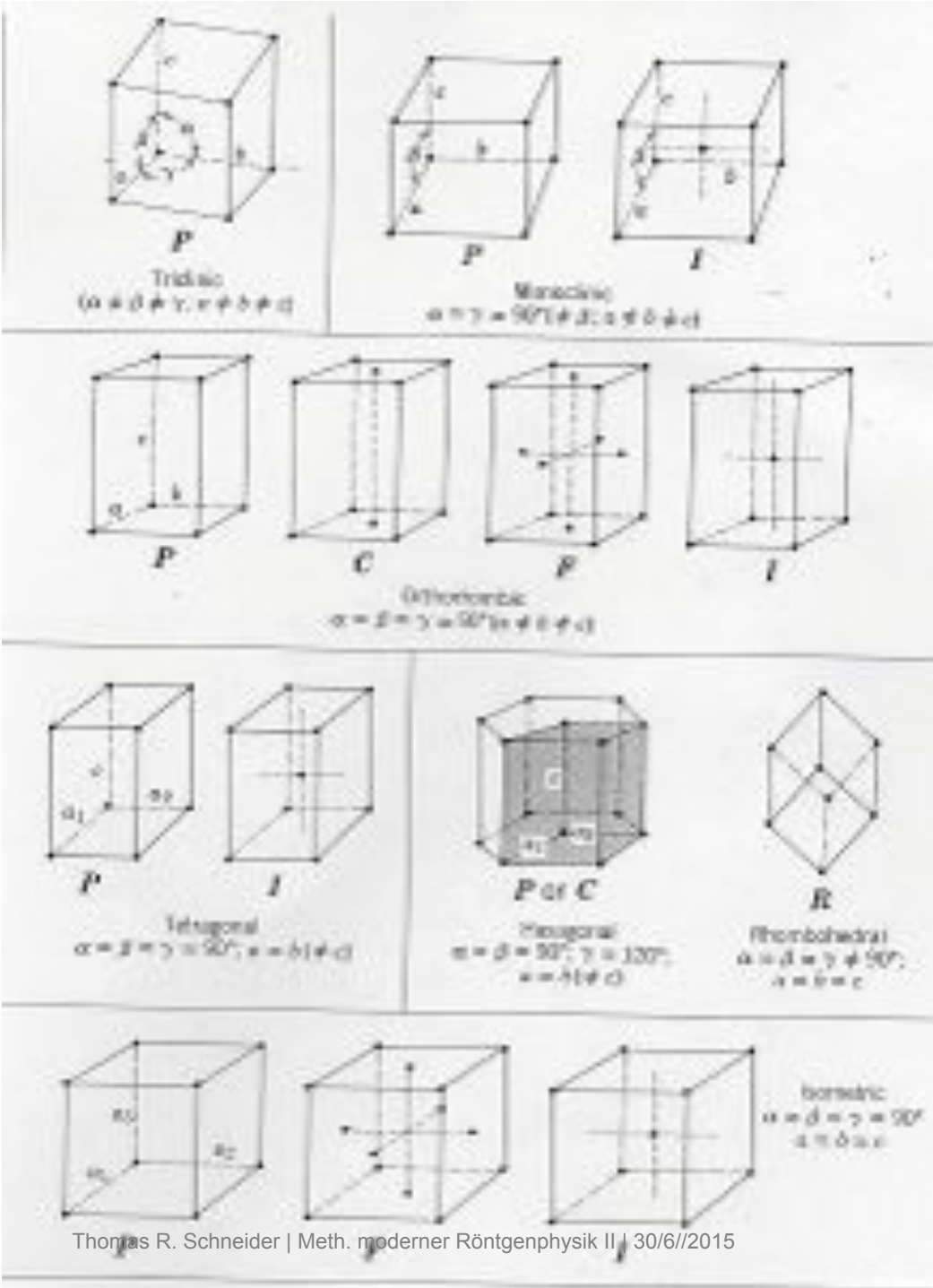


Primitive unit cell



Conventional unit cell

14 Bravais lattices



Space groups

- Only certain point groups (local symmetry) are compatible with the translational symmetry in a crystal as expressed by the Bravais lattice.
- 32 point groups combined with the 14 Bravais lattices give rise to a total of **230 space groups**
 - Fyodorov (1891)
 - Schönflies (1891)
 - Barlow (1894)
- Only 1,2,3,4,6-fold axis are in fact allowed; 5, 7 and higher folds are not compatible with an infinite translational repetition
- For chiral objects (such as protein molecules) only 65 space groups are possible

65 Space groups

That is all!

TRICLINIC

P1

MONOCLINIC

P2 P2₁ C2

ORTHORHOMBIC

P222 P222₁ P2₁2₁2 P2₁2₁2₁
 C222₁ C222 F222 I222 I2₁2₁2₁

TETRAGONAL

P4 P4₁ P4₂ P4₃
 I4 I4₁
 P422 P4₂12 P4₁22 P4₁2₁2 P4₂22
 P4₂2₁2 P4₃22 P4₃2₁2
 I422 I4₁22

TRIGONAL

P3 P3₁ P3₂ R3
 P312 P321 P3₁12 P3₁21
 P3₂12 P3₂21 R32

HEXAGONAL

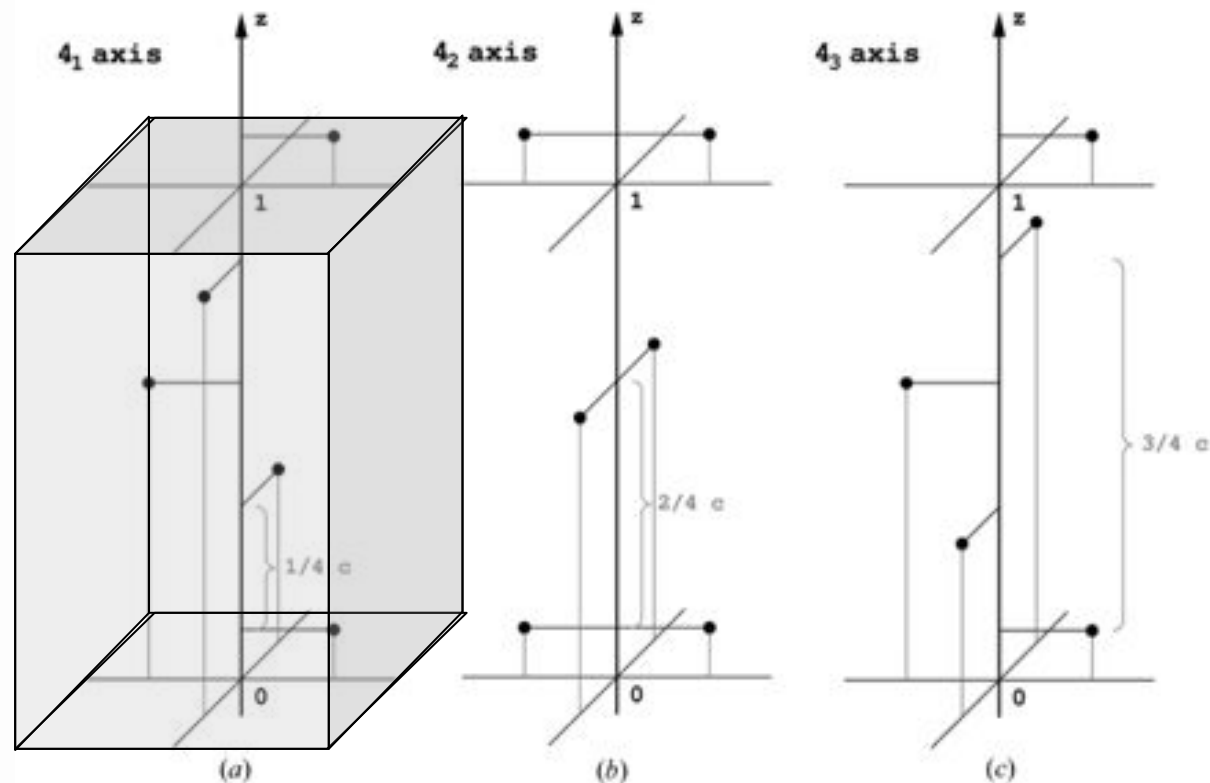
P6 P6₁ P6₅ P6₂ P6₄ P6₃
 P622 P6₁22 P6₅22 P6₂22 P6₄22 P6₃22

CUBIC (minus sign in front of triade optional)

P23 F23 I23 P2₁3 I2₁3
 P432 P₄232 F432
 F4₁32 I432 P4332 P4132 I4₁32

Understanding space group symbols

- **P222**: 'P-two-two-two', a primitive lattice with three 2-fold axis perpendicular to each other. Orthorhombic.
- **P4**: a primitive lattice with a 4 fold along the z-axis. tetragonal.
- **P4₁**: a primitive lattice with a 4-fold screw axis along the z-axis. tetragonal.



65 Space groups

That is all!

TRICLINIC

P1

MONOCLINIC

P2 P2₁ C2

ORTHORHOMBIC

P222 P222₁ P2₁2₁2 P2₁2₁2₁
C222₁ C222 F222 I222 I2₁2₁2₁

TETRAGONAL

P4 P4₁ P4₂ P4₃
I4 I4₁
P422 P4₂12 P4₁22 P4₁2₁2 P4₂22
P4₂2₁2 P4₃22 P4₃2₁2
I422 I4₁22

TRIGONAL

P3 P3₁ P3₂ R3
P312 P321 P3₁12 P3₁21
P3₂12 P3₂21 R32

HEXAGONAL

P6 P6₁ P6₅ P6₂ P6₄ P6₃
P622 P6₁22 P6₅22 P6₂22 P6₄22 P6₃22

CUBIC (minus sign in front of triade optional)

P23 F23 I23 P2₁3 I2₁3
P432 P₄232 F432
F4₁32 I432 P4332 P4132 I4₁32

International Tables Volume A



<http://it.iucr.org/A>

International Tables for Crystallography (2006). Vol. A, Space group 76, p. 333.

Tetragonal

4

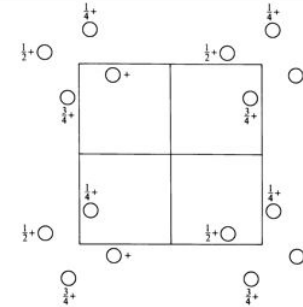
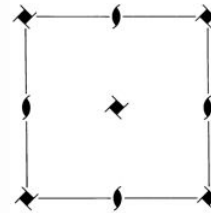
C_4

$P4_1$

Patterson symmetry $P4/m$

$P4_1$

No. 76



Origin on 4_1

Asymmetric unit $0 \leq x \leq \frac{1}{2}; 0 \leq y \leq \frac{1}{2}; 0 \leq z \leq 1$

Symmetry operations

(1) 1 (2) $2(0,0,\frac{1}{2})$ $0,0,z$ (3) $4^+(0,0,\frac{1}{4})$ $0,0,z$ (4) $4^-(0,0,\frac{3}{4})$ $0,0,z$

Generators selected (1); $r(1,0,0)$; $r(0,1,0)$; $r(0,0,1)$; (2); (3)

Positions

Multiplicity,
Wyckoff letter,
Site symmetry

Coordinates

Reflection conditions

4 a 1 (1) x,y,z (2) $\bar{x},\bar{y},z+\frac{1}{2}$ (3) $\bar{y},x,z+\frac{1}{2}$ (4) $y,\bar{x},z+\frac{1}{2}$

General:

$00l : l = 4n$

Symmetry of special projections

Along $[001]$ $p4$

$\mathbf{a}' = \mathbf{a}$ $\mathbf{b}' = \mathbf{b}$

Origin at $0,0,z$

Along $[100]$ $p1g1$

$\mathbf{a}' = \mathbf{b}$ $\mathbf{b}' = \mathbf{c}$

Origin at $x,0,0$

Along $[110]$ $p1g1$

$\mathbf{a}' = \frac{1}{2}(-\mathbf{a} + \mathbf{b})$ $\mathbf{b}' = \mathbf{c}$

Origin at $x,x,0$

Maximal non-isomorphic subgroups

I $[2]P2, (4) 1; 2$

IIa none

IIb none

Maximal isomorphic subgroups of lowest index

IIc $[2]C4, (\mathbf{a}' = 2\mathbf{a}, \mathbf{b}' = 2\mathbf{b})(P4_1, 76)$; $[3]P4_3, (\mathbf{c}' = 3\mathbf{c})(78)$; $[5]P4_5, (\mathbf{c}' = 5\mathbf{c})(76)$

Minimal non-isomorphic supergroups

I $[2]P4, 22(91)$; $[2]P4_1, 2, 2(92)$

II $[2]I4_1, (80)$; $[2]P4_2, (\mathbf{c}' = \frac{1}{2}\mathbf{c})(77)$

Mores space groups ...

International Tables for Crystallography (2006). Vol. A, Space group 76, p. 333.

Tetragonal 4 C_4^2 $P4_1$
 Patterson symmetry $P4/m$ $P4_1$ No. 76

Origin on 4₁
Asymmetric unit $0 \leq x \leq \frac{1}{2}; 0 \leq y \leq \frac{1}{2}; 0 \leq z \leq 1$
Symmetry operations
 (1) 1 (2) $2(0,0,\frac{1}{2})$ 0,0,z (3) $4^+(0,0,\frac{1}{2})$ 0,0,z (4) $4^-(0,0,\frac{1}{2})$ 0,0,z
Generators selected (1); $r(1,0,0)$; $r(0,1,0)$; $r(0,0,1)$; (2); (3)
Positions

Multiplicity	Wyckoff letter	Site symmetry	Coordinates	Reflection conditions
4	a	1	(1) x,y,z (2) $\bar{x},y,z+\frac{1}{2}$ (3) $\bar{y},x,z+\frac{1}{2}$ (4) $y,\bar{x},z+\frac{1}{2}$	00l : l = 4n

Symmetry of special projections
 Along [001] $p4$ $a' = a$ $b' = b$ $c' = c$ Origin at 0,0,z
 Along [100] $p1g1$ $a' = b$ $b' = c$ $c' = a$ Origin at x,0,0
 Along [110] $p1g1$ $a' = \frac{1}{2}(-a+b)$ $b' = c$ $c' = a$ Origin at x,x,0
Maximal non-isomorphic subgroups
I [2] $P2$, (4) 1; 2
IIa none
IIb none
Maximal isomorphic subgroups of lowest index
IIc [2] C_4 , ($a' = 2a, b' = 2b$) ($P4_1, 76$); [3] $P4$, ($c' = 3c$) (78); [5] $P4$, ($c' = 5c$) (76)
Minimal non-isomorphic supergroups
I [2] $P4, 22$ (91); [2] $P4, 2, 2$ (92)
II [2] $I4, 80$; [2] $P4, (c' = \frac{1}{2}c)$ (77)
 Copyright © 2006 International Union of Crystallography 333

International Tables for Crystallography (2006). Vol. A, Space group 91, pp. 366–367.

$P4_122$ D_4^3 422 **Tetragonal**
 No. 91 $P4_122$ Patterson symmetry $P4/mmm$

Origin on 2[010] at 4₁(1,2)1
Asymmetric unit $0 \leq x \leq 1; 0 \leq y \leq 1; 0 \leq z \leq \frac{1}{2}$
Symmetry operations
 (1) 1 (2) $2(0,0,\frac{1}{2})$ 0,0,z (3) $4^+(0,0,\frac{1}{2})$ 0,0,z (4) $4^-(0,0,\frac{1}{2})$ 0,0,z
 (5) 2 0,y,0 (6) 2 x,0, $\frac{1}{2}$ (7) 2 x,x, $\frac{1}{2}$ (8) 2 x, \bar{x} , $\frac{1}{2}$
 Copyright © 2006 International Union of Crystallography 366

Diffraction and Symmetry

- If a crystal has e.g. a four-fold symmetry axis along the rotation axis, the diffraction pattern will repeat itself every 90 deg. of rotation.
- That means, one needs to collect only 90 deg. of data to catch every plane at least once.
- In fact, due to Friedel's law, $I(hkl) = I(-h,-k,-l)$, only half of this rotation is needed.
- In reality, crystals end up in random orientations and it is not trivial to 'guess' how many degrees of data are needed to complete a data set -> strategy programs such as BEST.
- Radiation damage needs to be taken into account!
- If the crystal permits (in terms of radiation damage), symmetry equivalent reflections can be collected and compared.

Rejections

- When compared to symmetry equivalents, some reflections do not fit into the picture; these reflections are marked as outliers and will be rejected in subsequent rounds of scaling

From scale.log

Reflections printed out due to failing chi**2 test during scaling

H	K	L	Total	chi**2,	<I>	resol		
9	1	46		40.2	4736.3	2.94		
original	hkl	film#	dev./sigma	I,	%PROB,	sigma		
-9	-1	46 a+	147	0.7	4982.1	0	213.6	
-1	9	46 f+	160	-1.4	4406.7	0	305.0	
1	9	-46 a+	77	1.9	5124.8	0	166.4	
1	9	-46 a+	178	0.2	4861.6	0	188.9	
-1	-9	-46 a+	95	-1.3	4487.7	0	281.2	
-9	1	-46 f+	190	-5.4	3180.9	100	295.7	
-9	-1	-46 a-	178	-0.3	4742.5	0	365.1	
-1	9	-46 a-	27	-1.7	4531.7	0	188.7	
1	-9	-46 a-	81	0.4	5012.1	0	407.1	
1	9	46 f-	172	-0.3	4748.7	0	316.3	
-9	1	46 a-	127	1.0	5060.7	0	244.0	

Data Quality – R-factors

From output.sca (10 symmetry equivalent reflections for (9,1,46))

-9	-1	46	9	1	46	147	1	0	2	4973.3	215.4
-1	9	46	9	1	46	160	1	0	4	4402.4	306.1
1	9	-46	9	1	46	77	1	0	5	5121.4	168.9
1	9	-46	9	1	46	178	1	1	5	4855.8	191.1
-1	-9	-46	9	1	46	95	1	1	6	4490.0	282.9
-9	-1	-46	9	1	46	178	2	1	1	4736.8	365.9
-1	9	-46	9	1	46	27	2	0	3	4545.1	191.2
1	-9	-46	9	1	46	81	2	1	4	5125.4	416.2
1	9	46	9	1	46	172	2	0	6	4732.4	316.8
-9	1	46	9	1	46	127	2	0	8	5044.3	245.2
			9	1	46					4836.2	75.7

- The standard R-factor:

$$R_{lin} = \frac{\sum_{hkl} |I_{hkl} - \langle I \rangle_{hkl}|}{\sum_{hkl} I_{hkl}}$$

describes how well the measurements of symmetry equivalent reflections agree with their mean.

Scaling and merging

From output.sca (10 symmetry equivalent reflections for (9,1,46))

-9	-1	46	9	1	46	147	1	0	2	4973.3	215.4
-1	9	46	9	1	46	160	1	0	4	4402.4	306.1
1	9	-46	9	1	46	77	1	0	5	5121.4	168.9
1	9	-46	9	1	46	178	1	1	5	4855.8	191.1
-1	-9	-46	9	1	46	95	1	1	6	4490.0	282.9
-9	-1	-46	9	1	46	178	2	1	1	4736.8	365.9
-1	9	-46	9	1	46	27	2	0	3	4545.1	191.2
1	-9	-46	9	1	46	81	2	1	4	5125.4	416.2
1	9	46	9	1	46	172	2	0	6	4732.4	316.8
-9	1	46	9	1	46	127	2	0	8	5044.3	245.2
			9	1	46					4836.2	75.7

Scaled and merged intensity for (9,1,46)

Error estimate corrected for multiple measurements (\sqrt{N})

- For this unique reflection, the R-factor is 4.6%

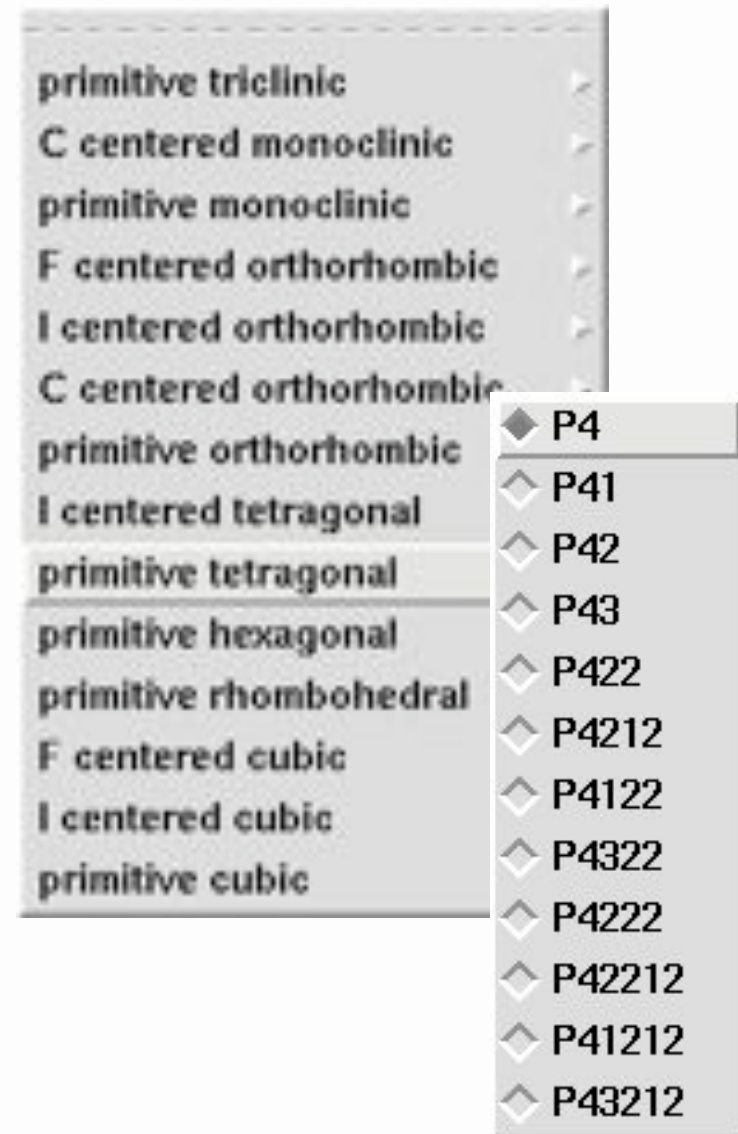
R-factors and Multiplicity

- The standard R-factor does not reflect the improvement in data quality when more measurements are made:

4973.3										4973.3	0.0000
4973.3	4402.4									4687.9	0.0609
4973.3	4402.4	5121.4								4832.4	0.0593
4973.3	4402.4	5121.4	4855.8							4838.2	0.0450
4973.3	4402.4	5121.4	4855.8	4490.0						4768.6	0.0541
4973.3	4402.4	5121.4	4855.8	4490.0	4736.8					4763.3	0.0462
4973.3	4402.4	5121.4	4855.8	4490.0	4736.8	4545.1				4732.1	0.0458
4973.3	4402.4	5121.4	4855.8	4490.0	4736.8	4545.1	5125.4			4781.3	0.0497
4973.3	4402.4	5121.4	4855.8	4490.0	4736.8	4545.1	5125.4	4732.4		4775.8	0.0453
4973.3	4402.4	5121.4	4855.8	4490.0	4736.8	4545.1	5125.4	4732.4	5044.3	4802.7	0.0461

Space group determination

- Combination of symmetry elements with the 14 Bravais lattices results in 230 space groups.
- Space groups can be identified by checking:
 - How well the data fit the symmetry (R_{merge})
 - The systematic absences
- Some space groups (e.g. enantiomorphic space groups) can only be distinguished by solving/refining the structure.



Space groups – Checking for symmetry

XPREP output

Identical indices and Friedel opposites combined before calculating R(sym)

Option	Space Group	No.	Type	Axes	CSD	R(sym)	N(eq)	Syst. Abs.	CFOM
[A]	P4	# 75	chiral	1	4	0.051	57875	0.0 / 11.0	23.80
[B]	P4(2)	# 77	chiral	1	7	0.051	57875	0.0 / 11.0	16.30
[C]	P4(1)	# 76	chiral	1	114	0.051	57875	0.0 / 11.0	4.67
[D]	P4(3)	# 78	chiral	1	114	0.051	57875	0.0 / 11.0	4.67
[E]	P422	# 89	chiral	1	4	0.056	72030	0.0 / 11.0	20.94
[F]	P42(1)2	# 90	chiral	1	4	0.056	72030	0.0 / 11.0	20.94
[G]	P4(2)22	# 93	chiral	1	4	0.056	72030	0.0 / 11.0	20.94
[H]	P4(2)2(1)2	# 94	chiral	1	20	0.056	72030	0.0 / 11.0	5.70
[I]	P4(1)22	# 91	chiral	1	8	0.056	72030	0.0 / 11.0	12.05
[J]	P4(3)22	# 95	chiral	1	8	0.056	72030	0.0 / 11.0	12.05
[K]	P4(1)2(1)2	# 92	chiral	1	245	0.056	72030	0.0 / 11.0	1.34
[L]	P4(3)2(1)2	# 96	chiral	1	245	0.056	72030	0.0 / 11.0	1.34

Summary - Symmetry

- All chiral objects can only crystallize in one of the 65 chiral space groups
- These space groups arise from the combination of the condition of rotational and translational invariance of the crystal with the fact that the repeating object is chiral.
- The symmetry of the crystal can be used to short-cut data collection.
- Collection of symmetry-related reflections allows to do statistics about the accuracy and precision of the measured diffraction intensities.
- **WARNING:** The packing introduced by the crystal symmetry does not always correspond to the biologically relevant oligomer.

Sources

- **Dauter1999.** Zbigniew Dauter (1999) 'Data-Collection Strategies' Acta Cryst. D55:1703-1717
- **Dauter2010.** Zbigniew Dauter & Mariusz Jaskolski (2010) 'How to read (and understand) Volume A of International Tables for Crystallography: an introduction for nonspecialists' Acta Cryst D43:1150-1171
- **Blow2007.** David Blow (2007): 'Outline of Crystallography for Biologists' Oxford University Press
- **Glusker1994.** Jenny Glusker, Mitchell Lewis, Miriam Rossi (1994): 'Crystal Structure Analysis for Chemists and Biologists' VCH Publishers