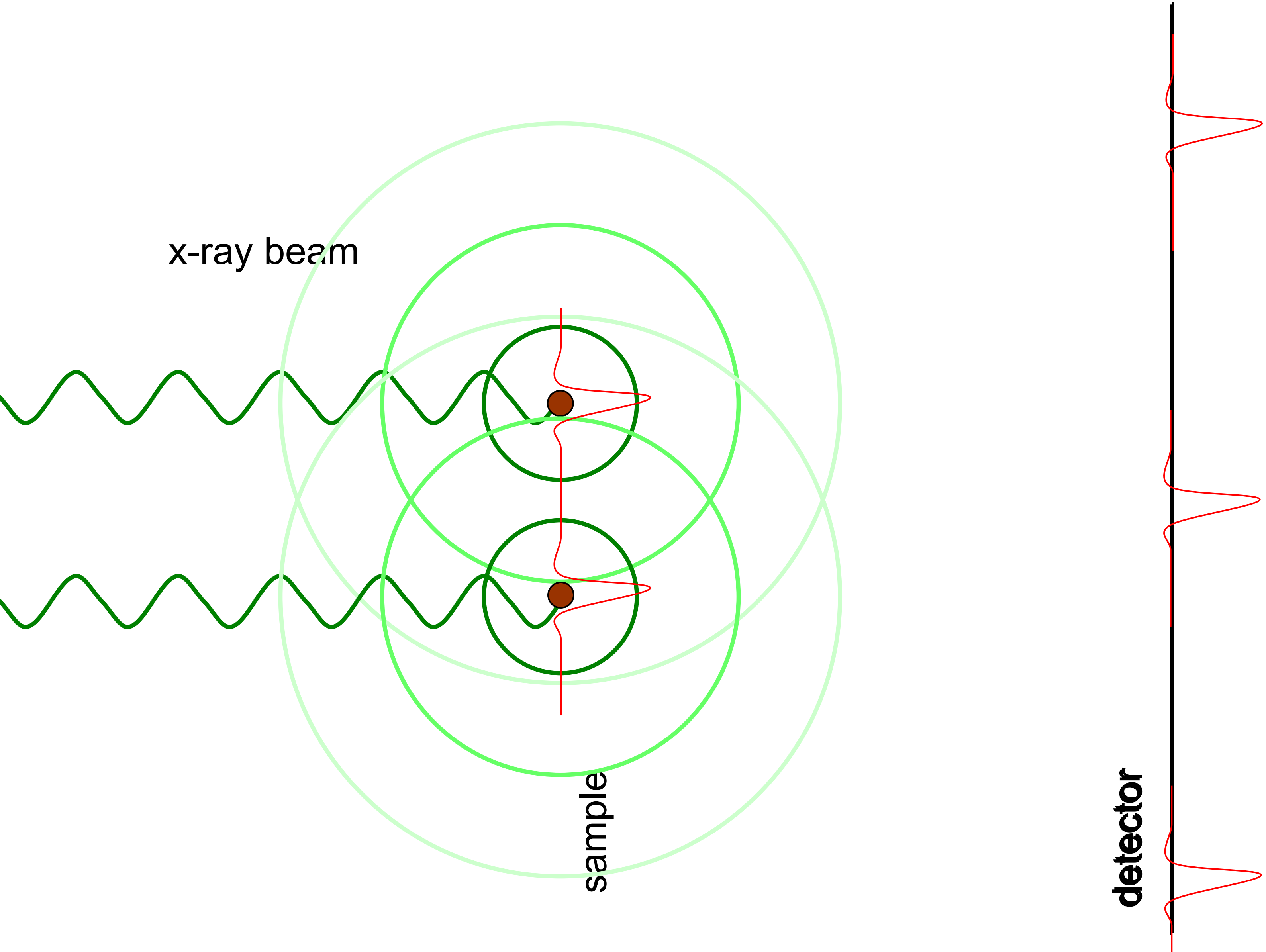
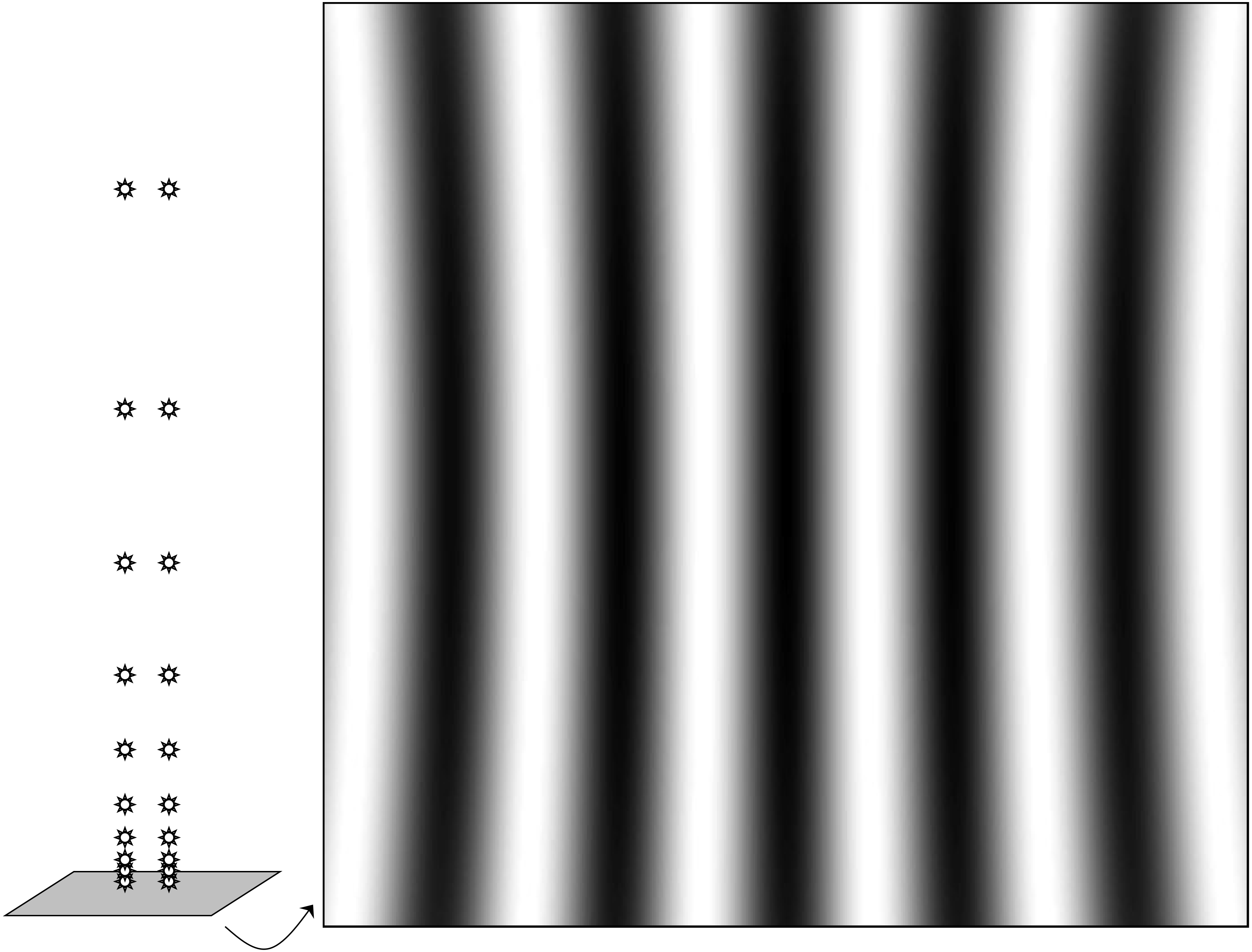
What is your
Space Group?

https://bl831.als.lbl.gov/
~jamesh/powerpoint/
fear_no_Spacegroups.pptx

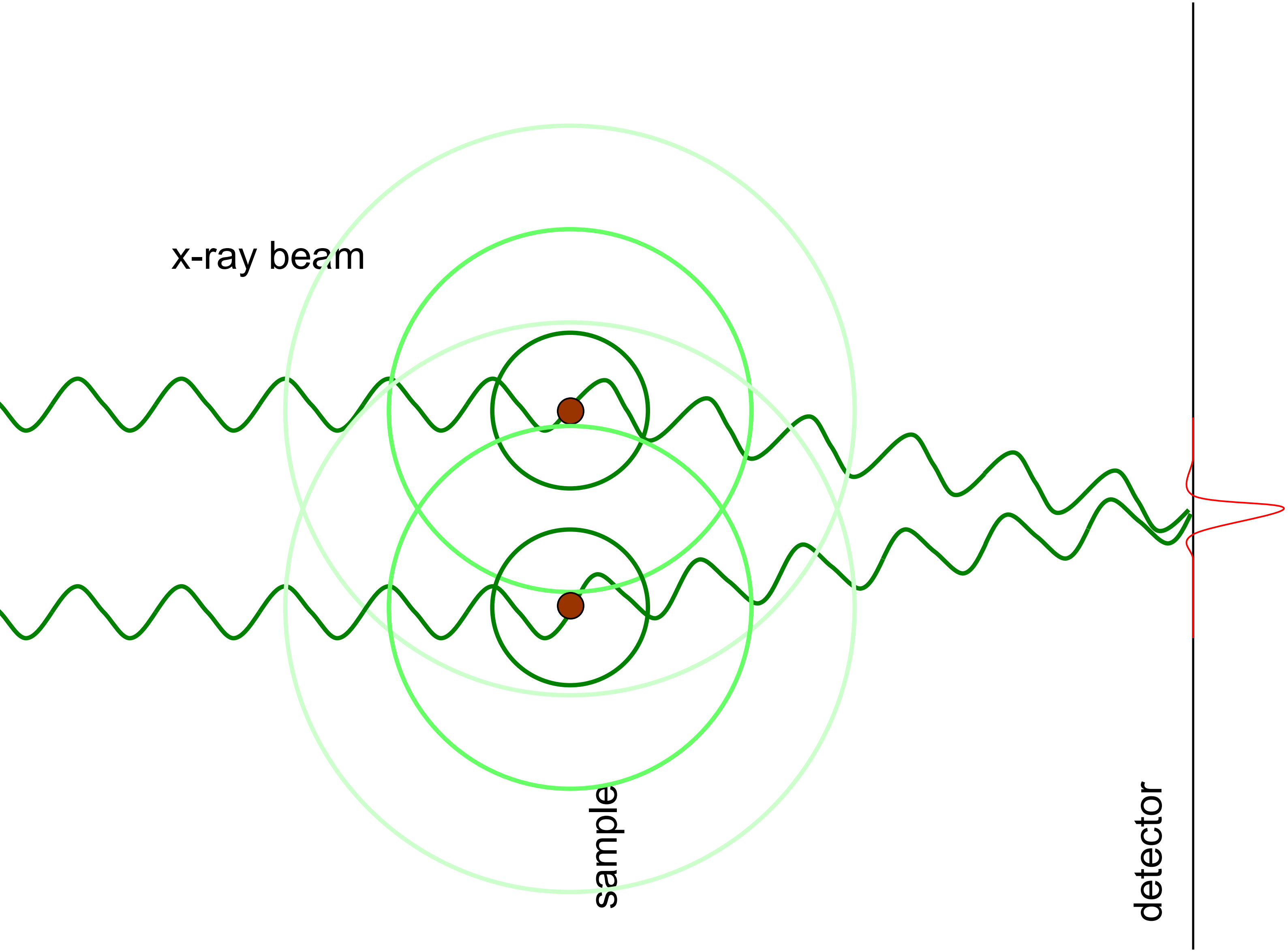
the first question asked and the last to be answered

scattering

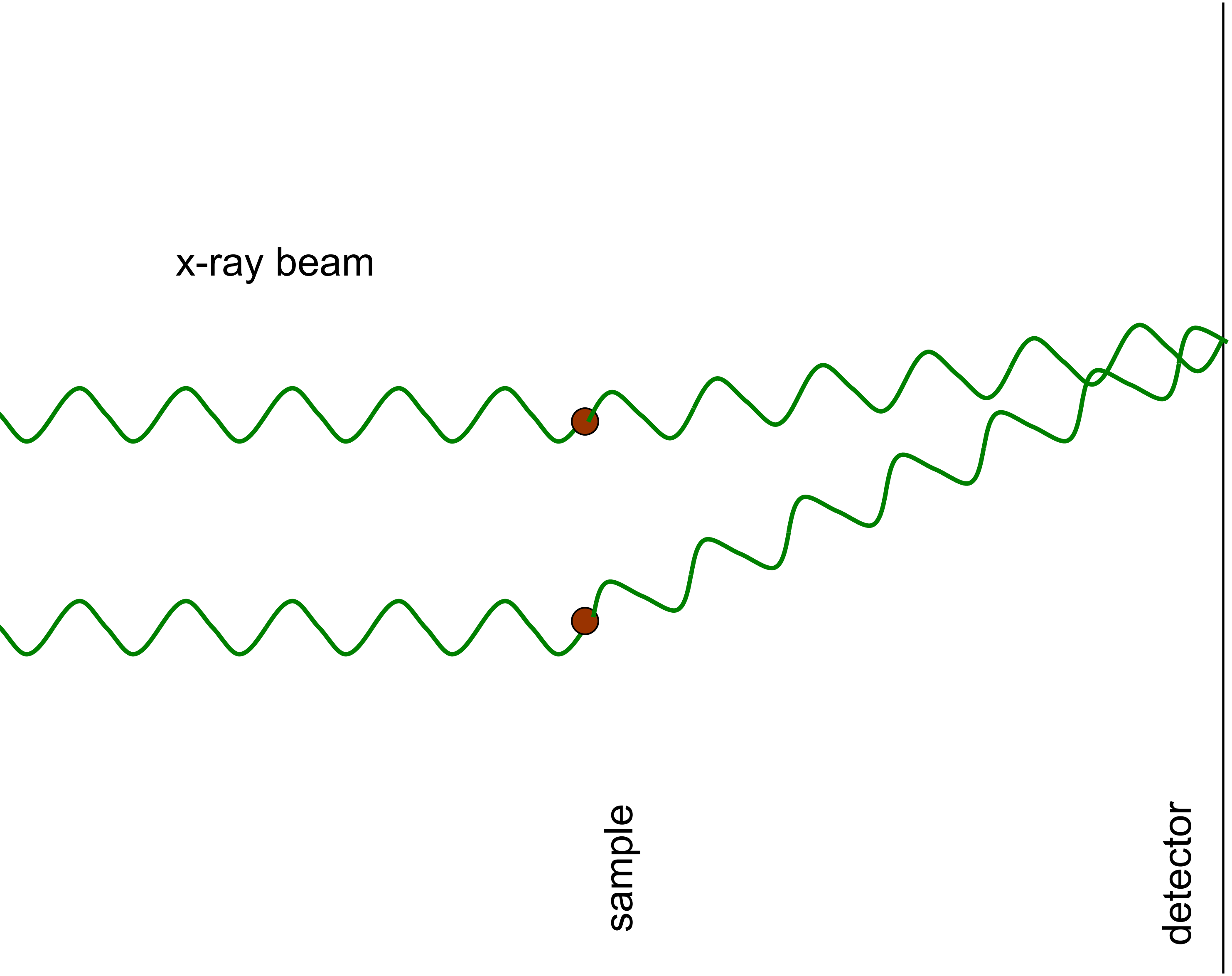




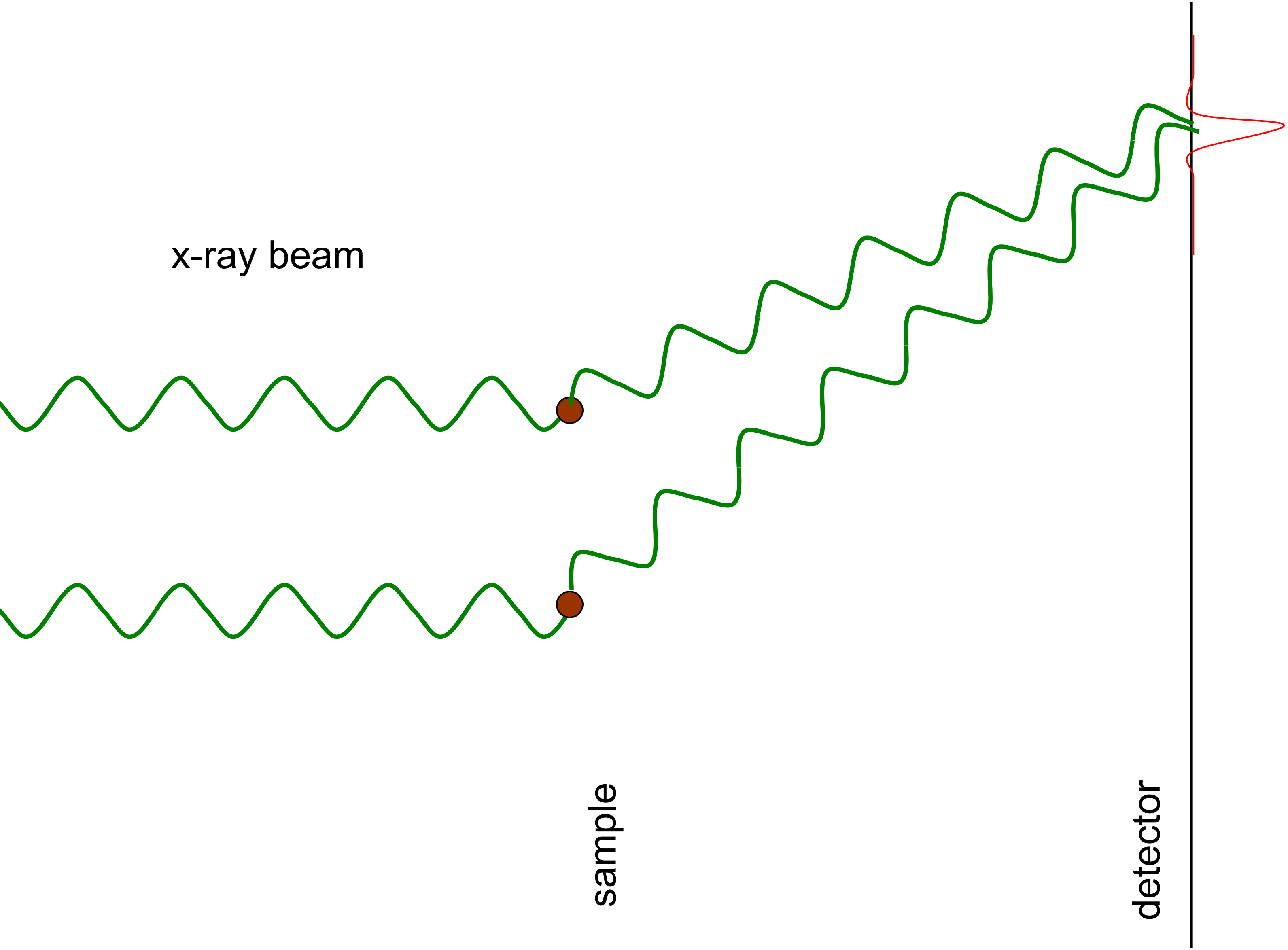
scattering



scattering

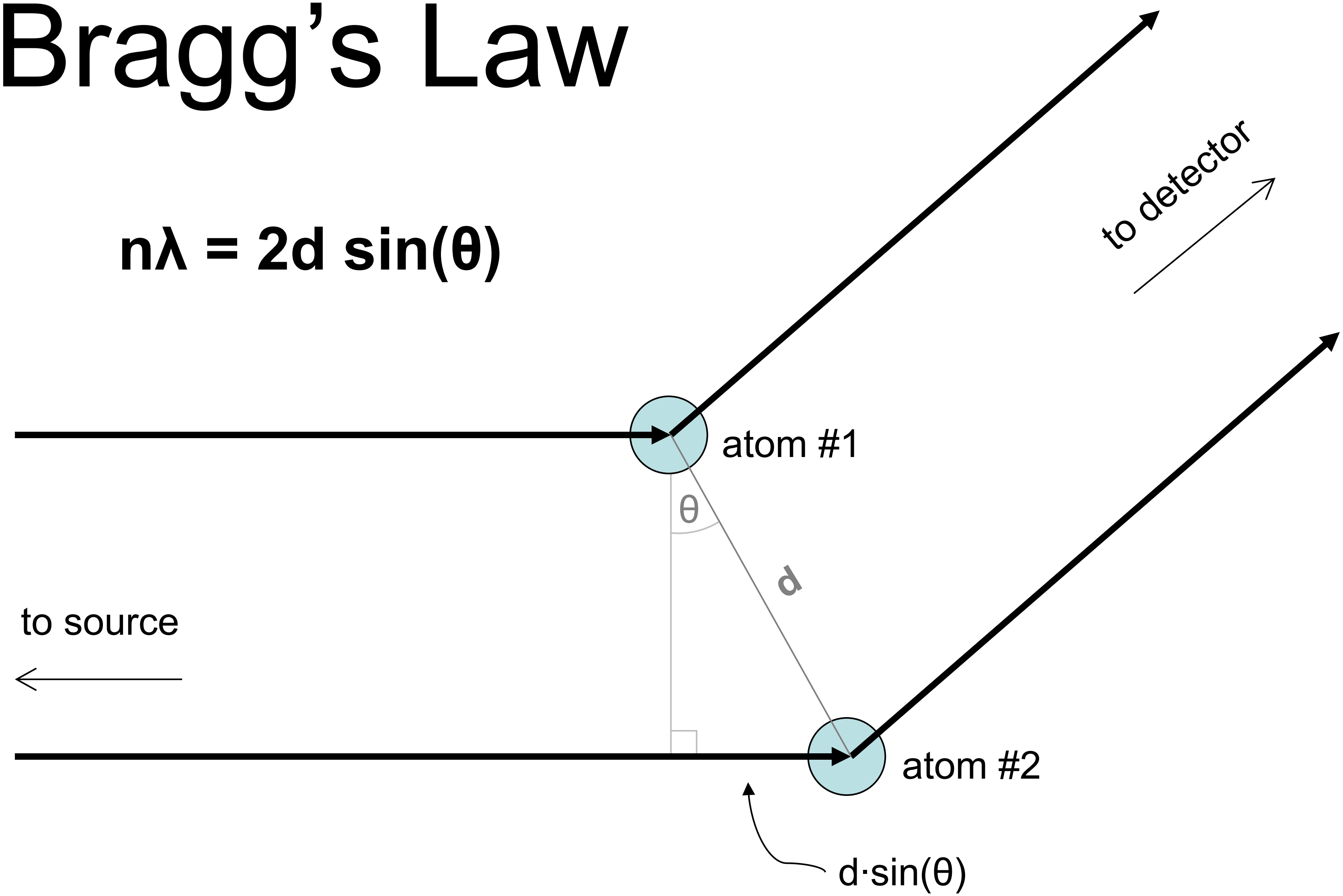


scattering



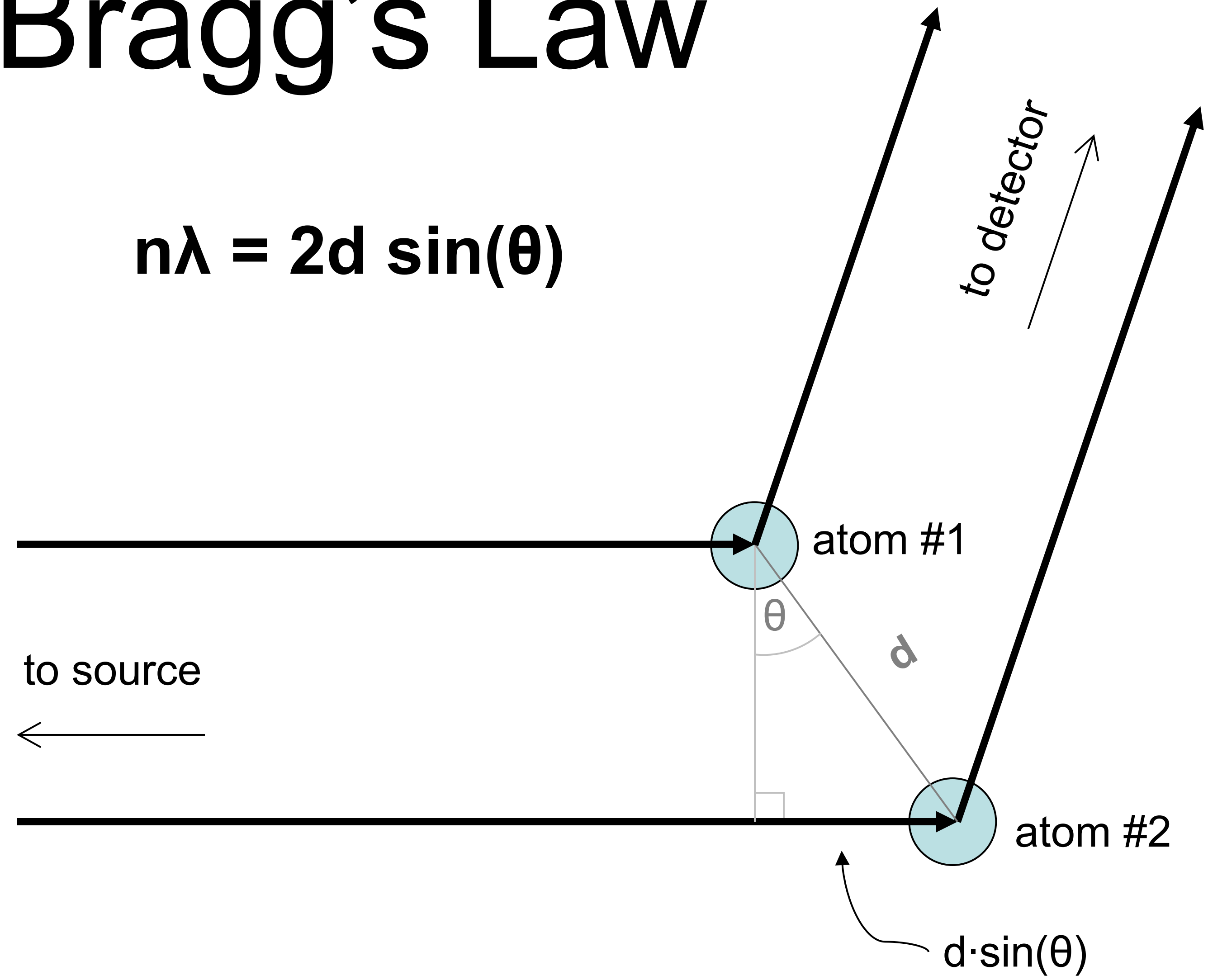
Bragg's Law

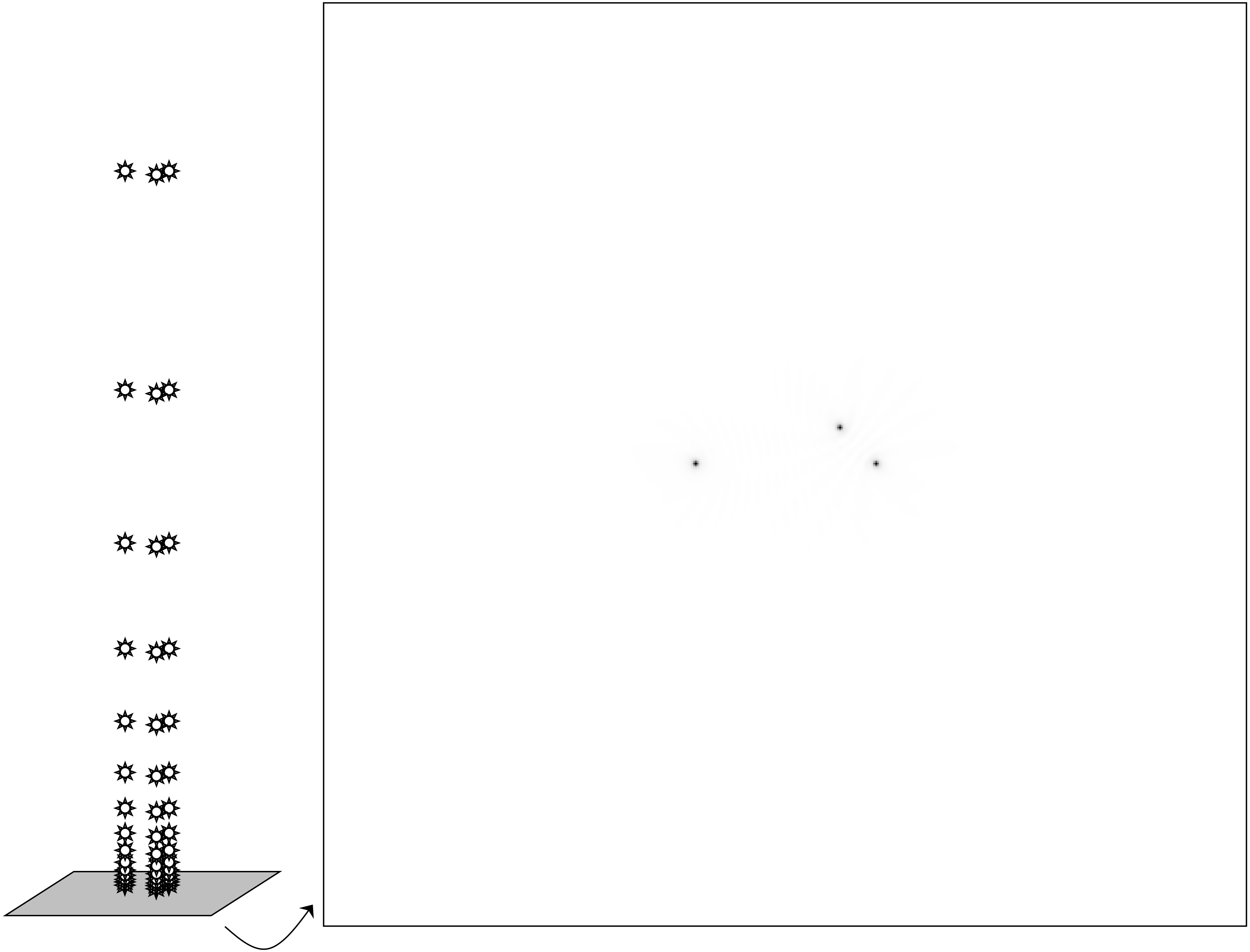
$$n\lambda = 2d \sin(\theta)$$



Bragg’s Law

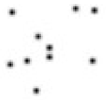
$n\lambda = 2d \sin(\theta)$





scattering from a structure

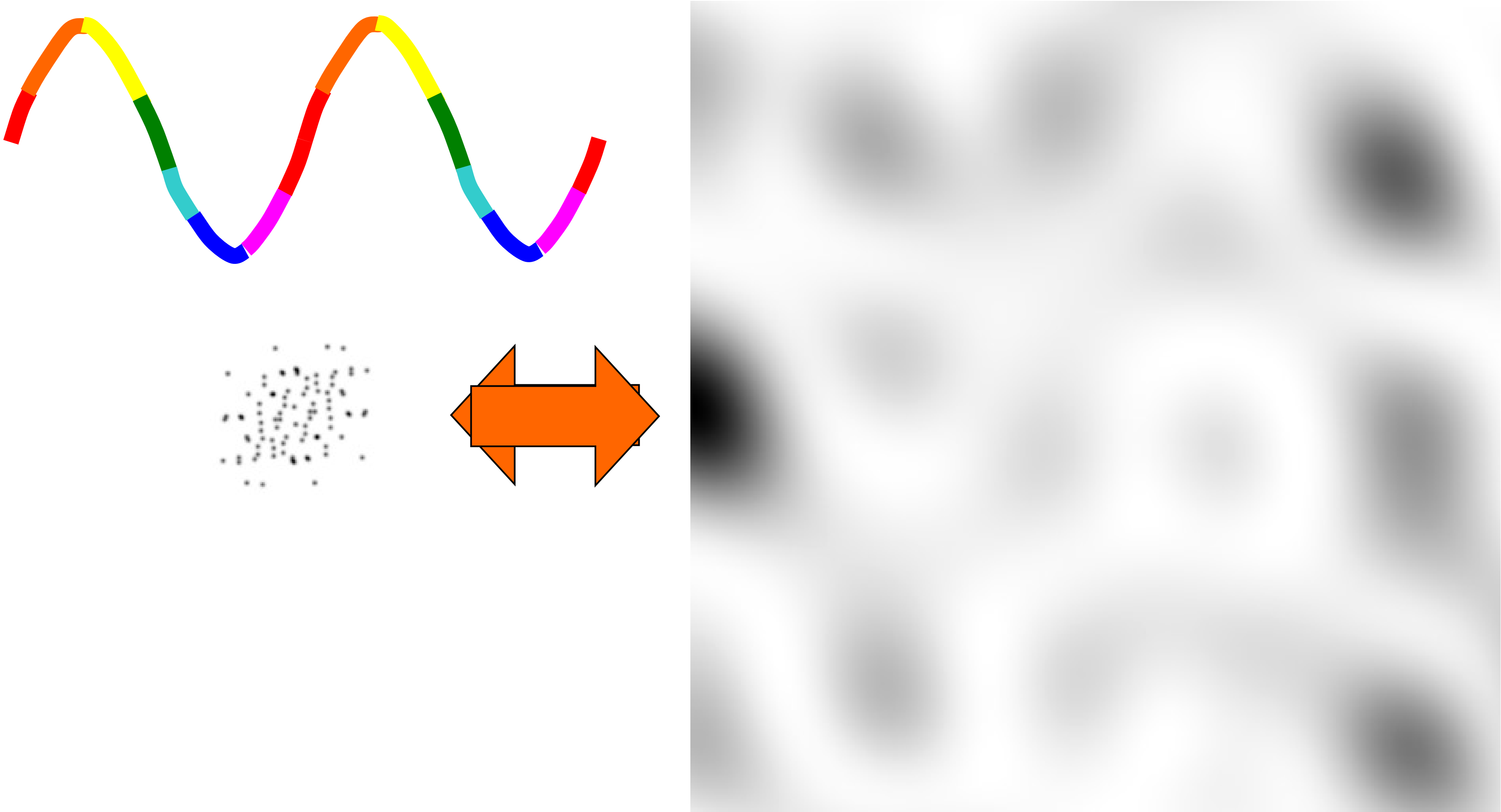
sample



detector

forward inverse Fourier Transform

no phase



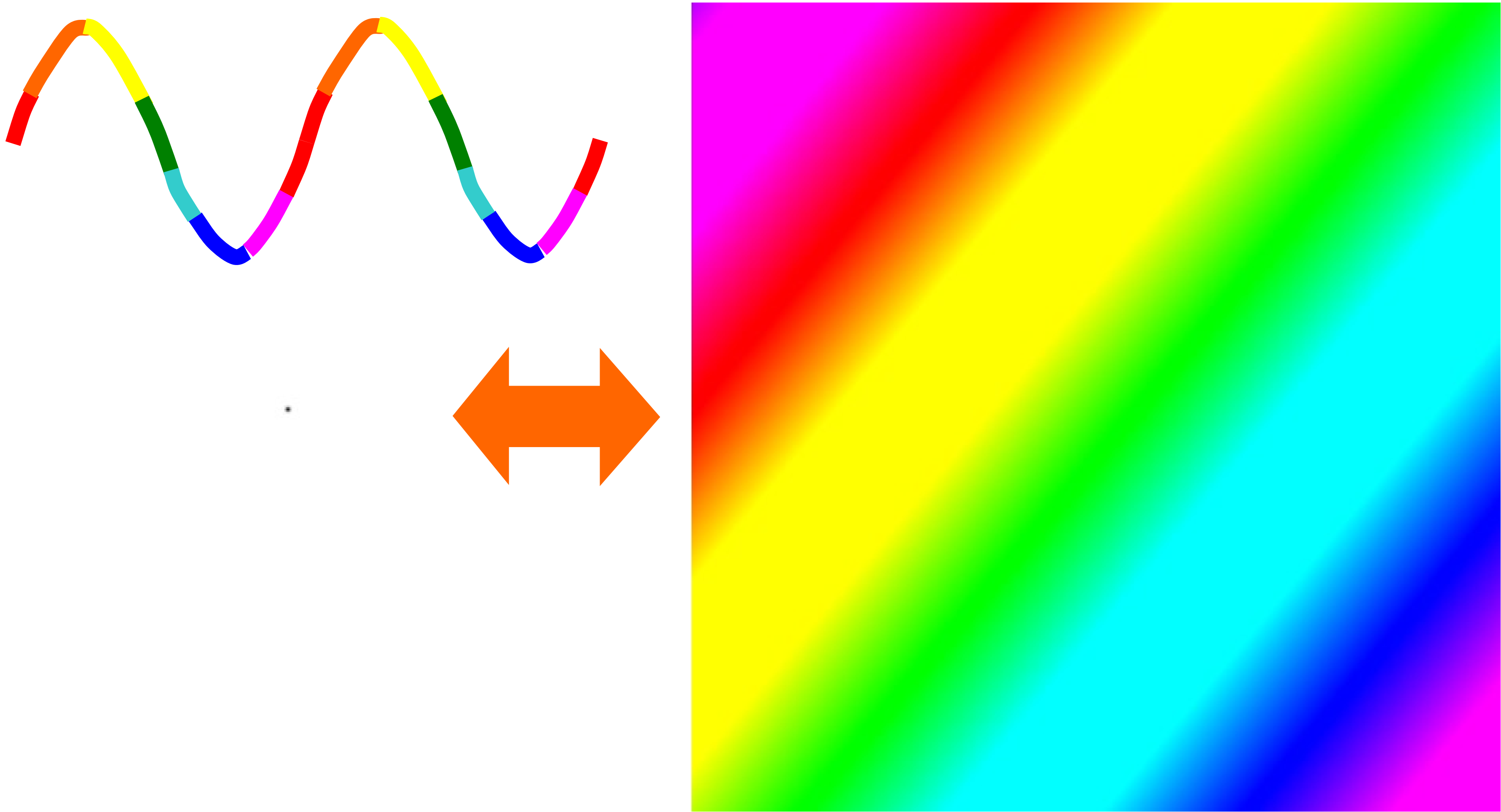
inverse Fourier Transform

scattering from a structure

colored by phase

sample

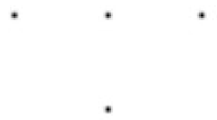
detector



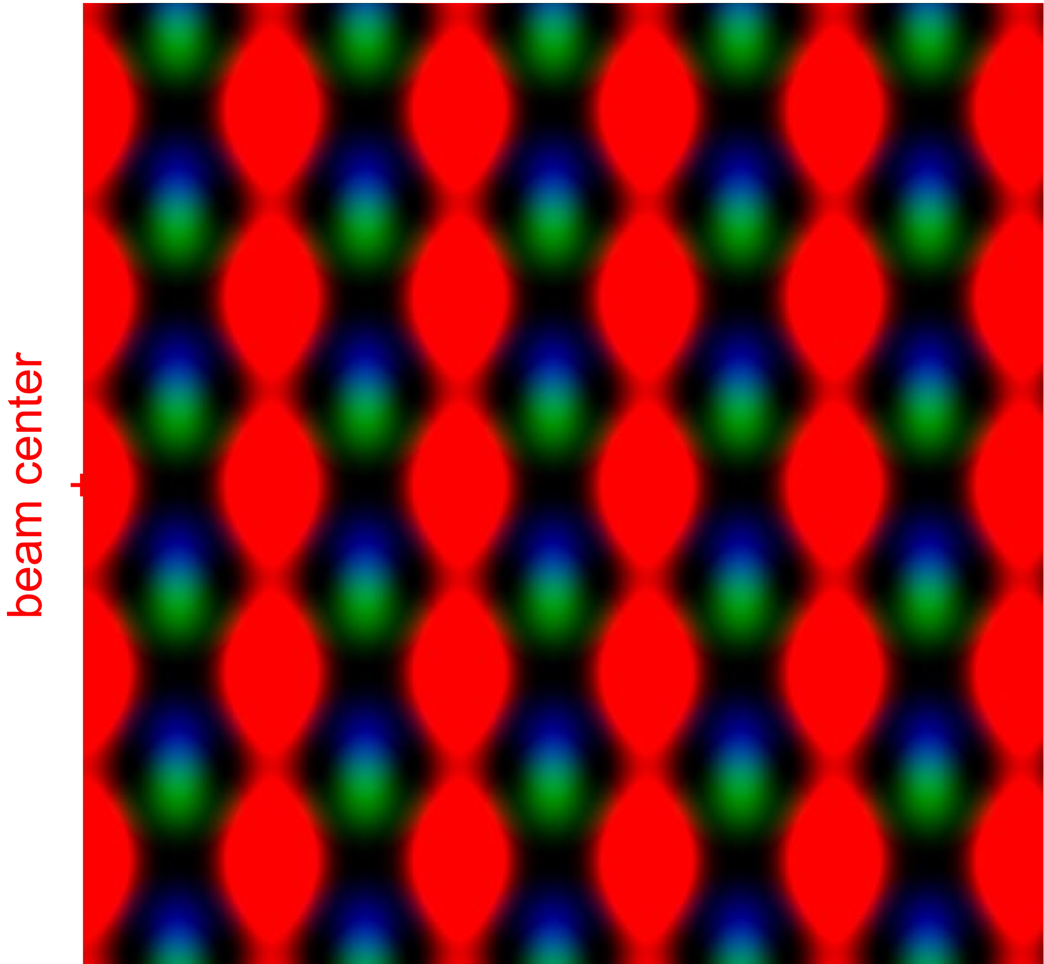
scattering from a lattice

colored by phase

sample

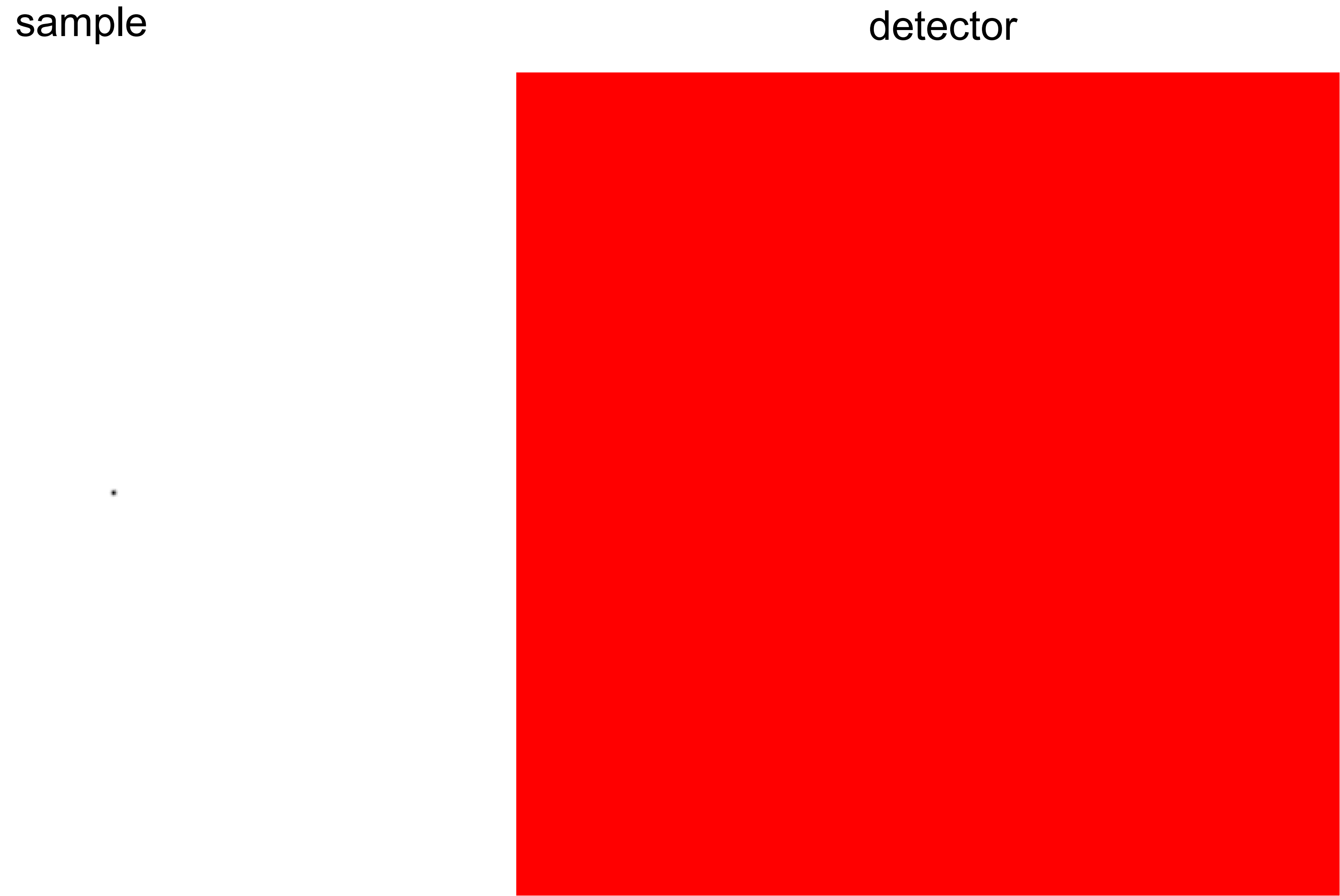


detector



scattering from a lattice

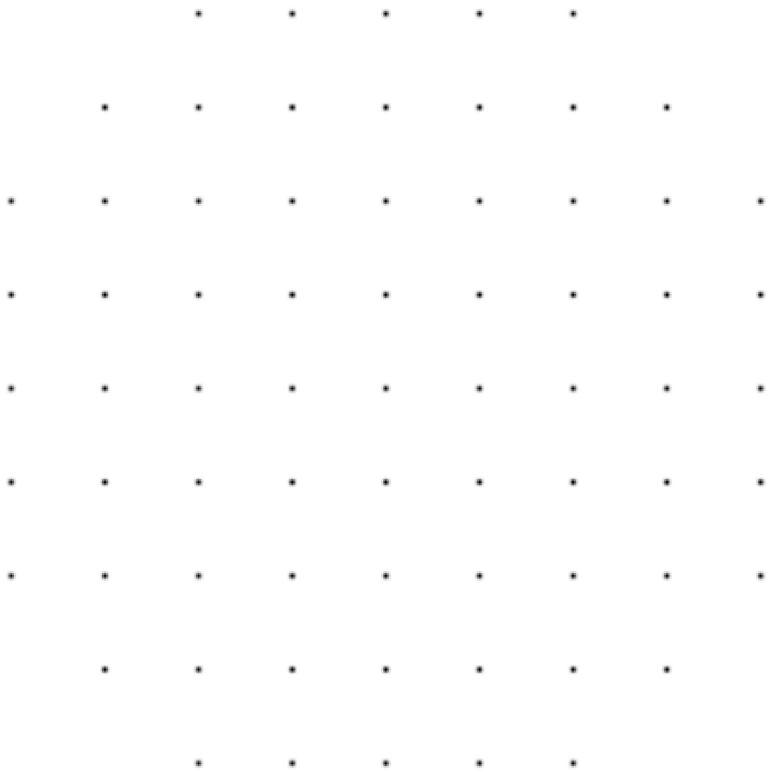
colored by phase



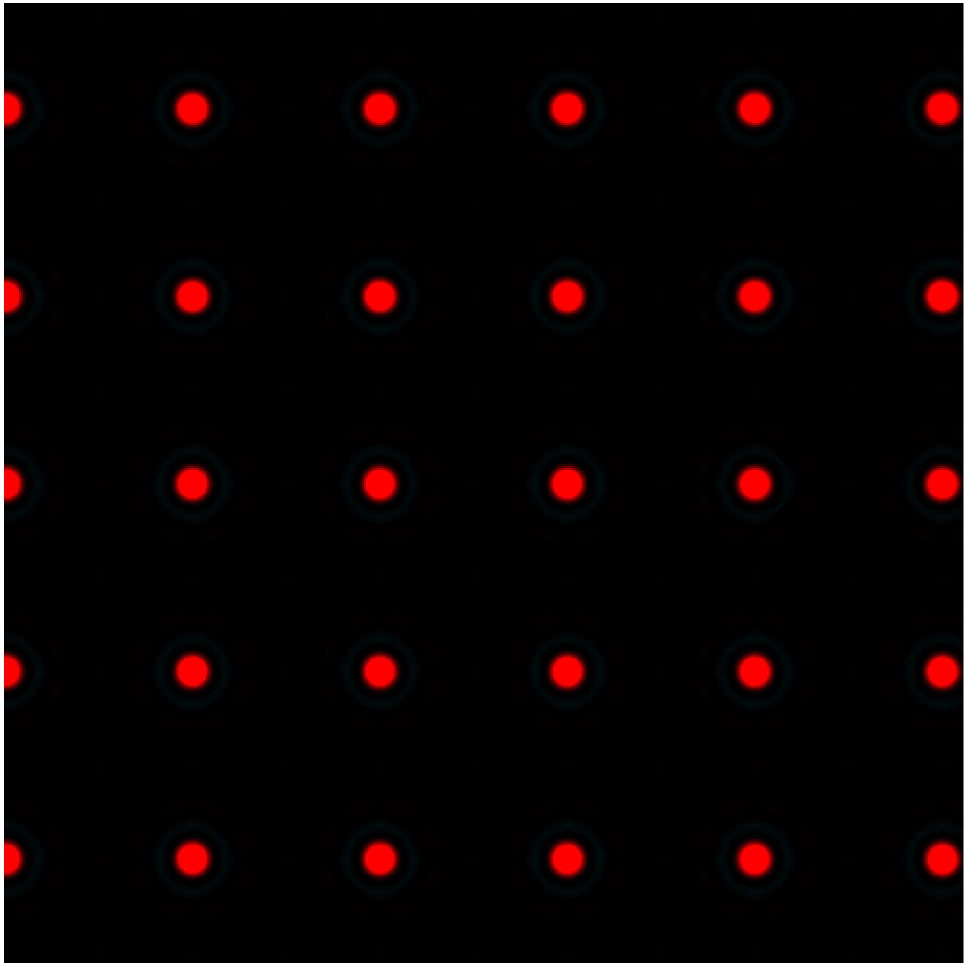
scattering from a lattice

colored by phase

sample



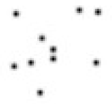
detector



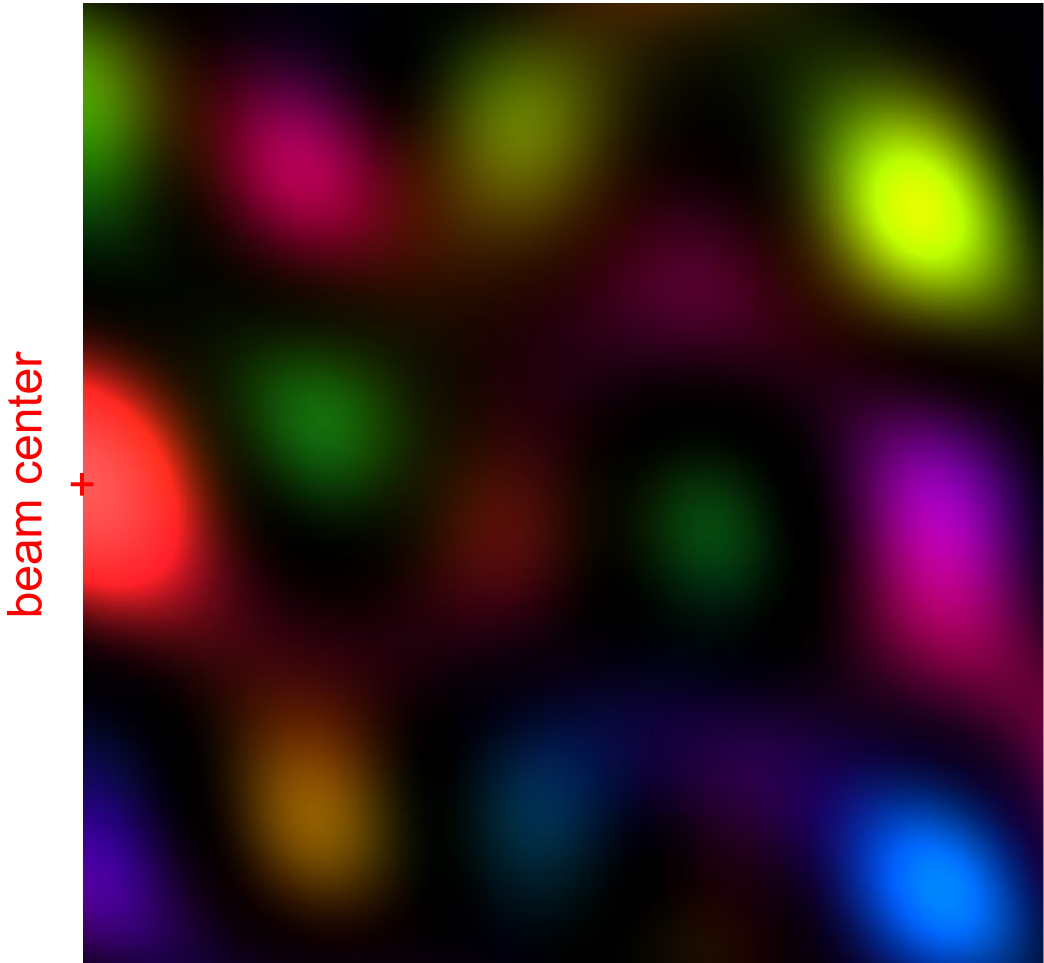
scattering from a crystal structure

colored by phase

sample



detector

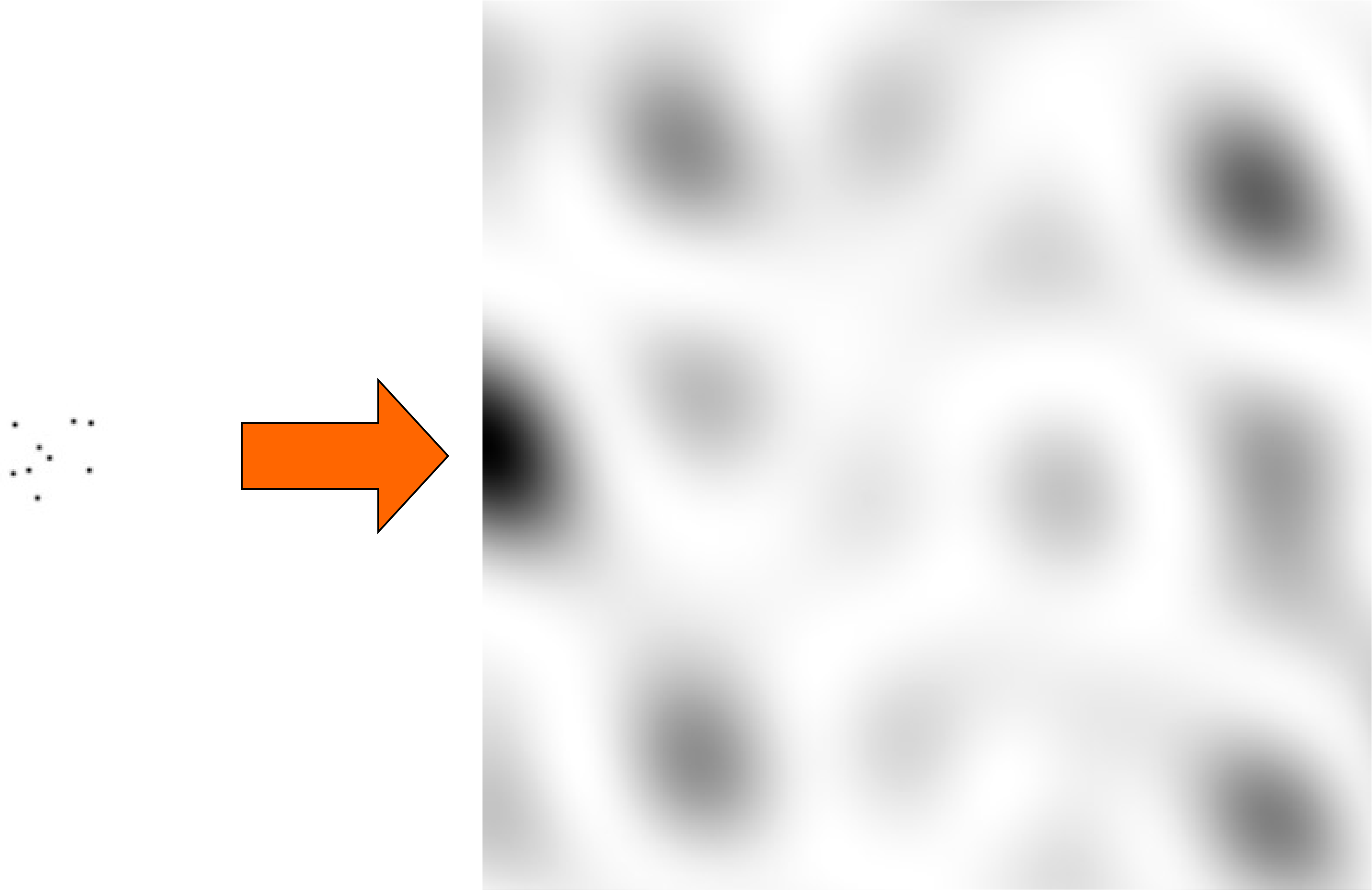


Major Phasing techniques

- Molecular Replacement
- Multiple Isomorphous Replacement
- Multiwavelength Anomalous Diffraction
- Single-wavelength Anomalous Diffraction

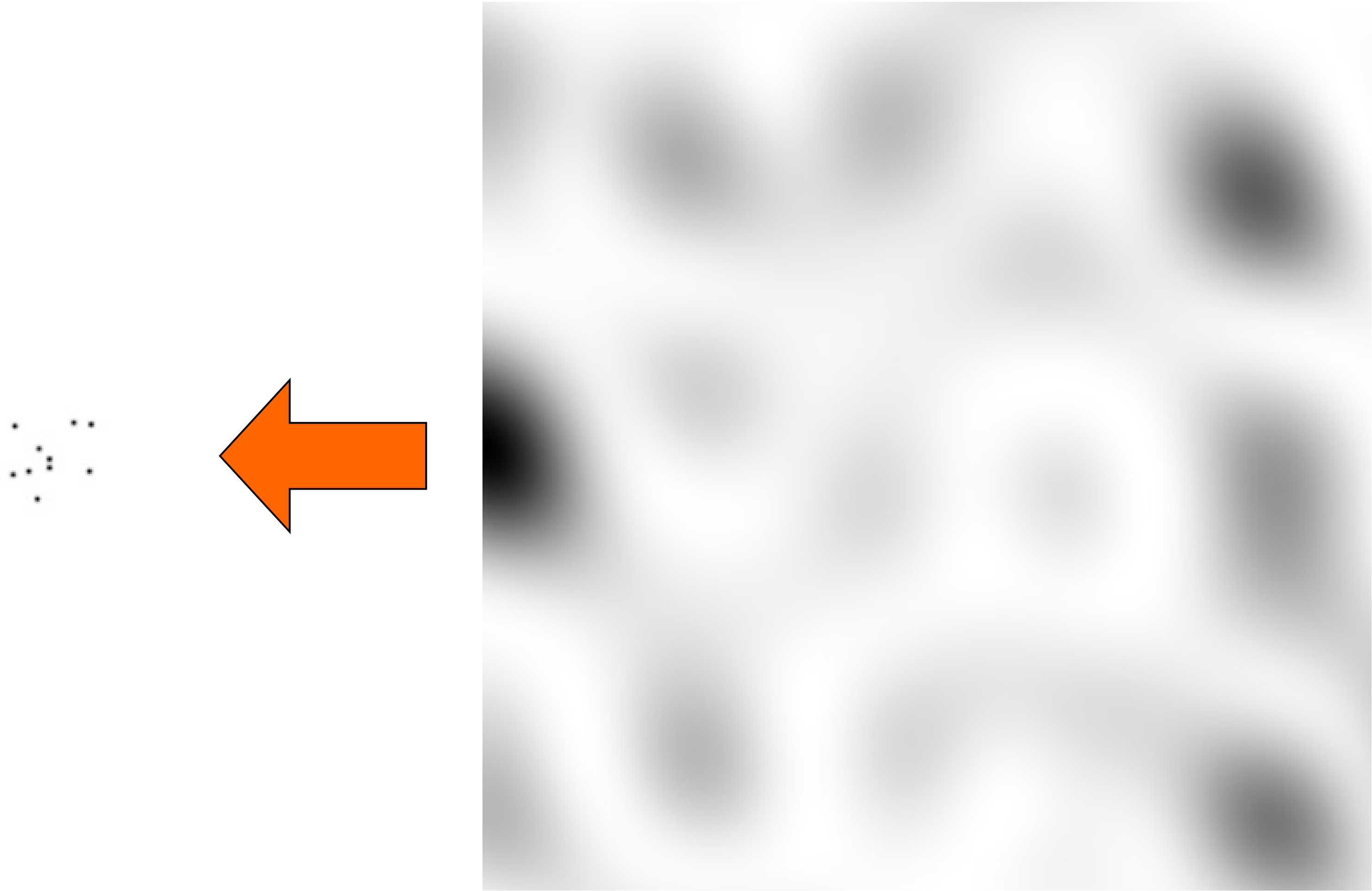
inverse Fourier Transform

no phase



inverse Fourier Transform

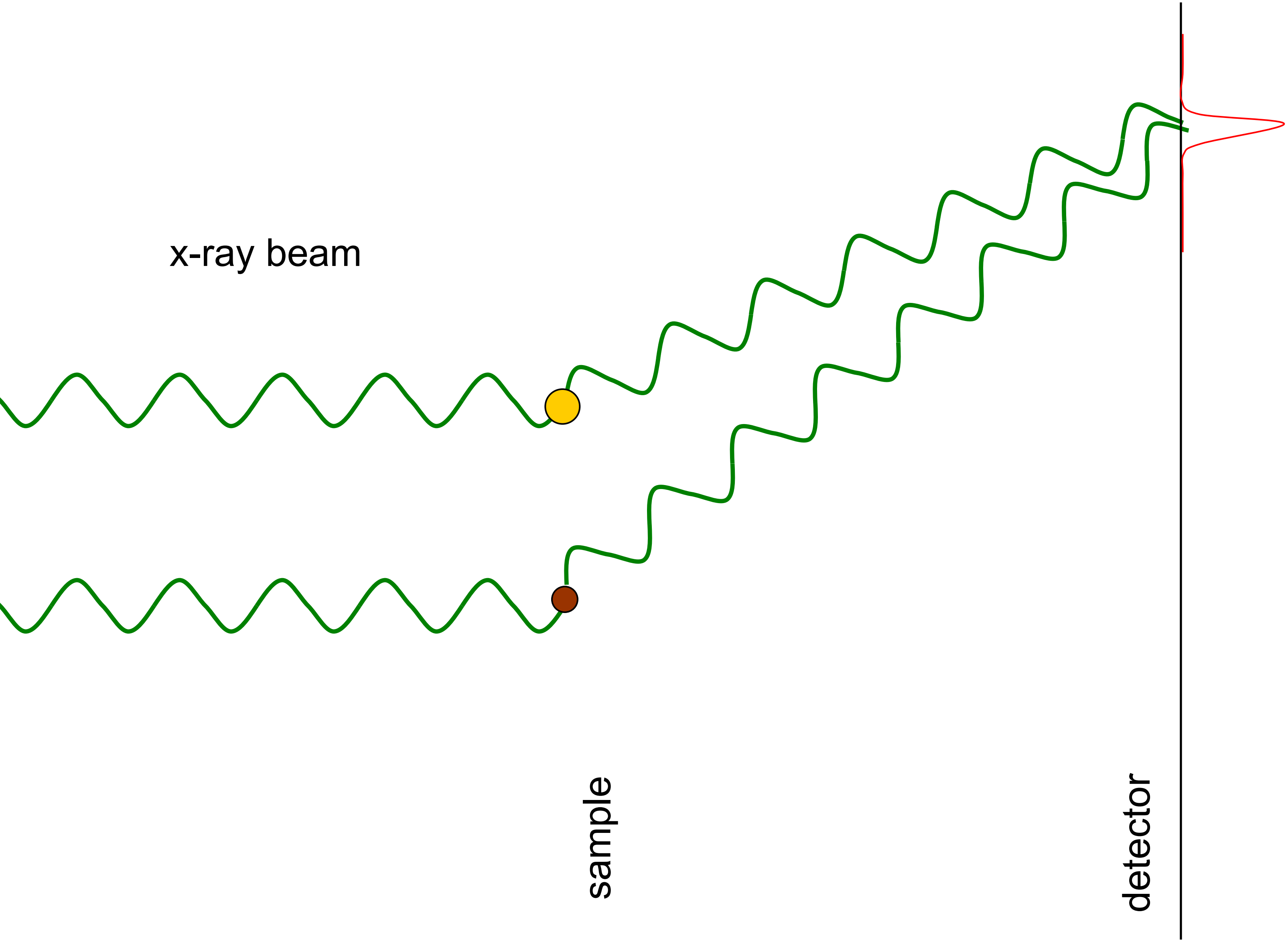
no phase



Major Phasing techniques

- Molecular Replacement
- Multiple Isomorphous Replacement
- Multiwavelength Anomalous Diffraction
- Single-wavelength Anomalous Diffraction

anomalous scattering

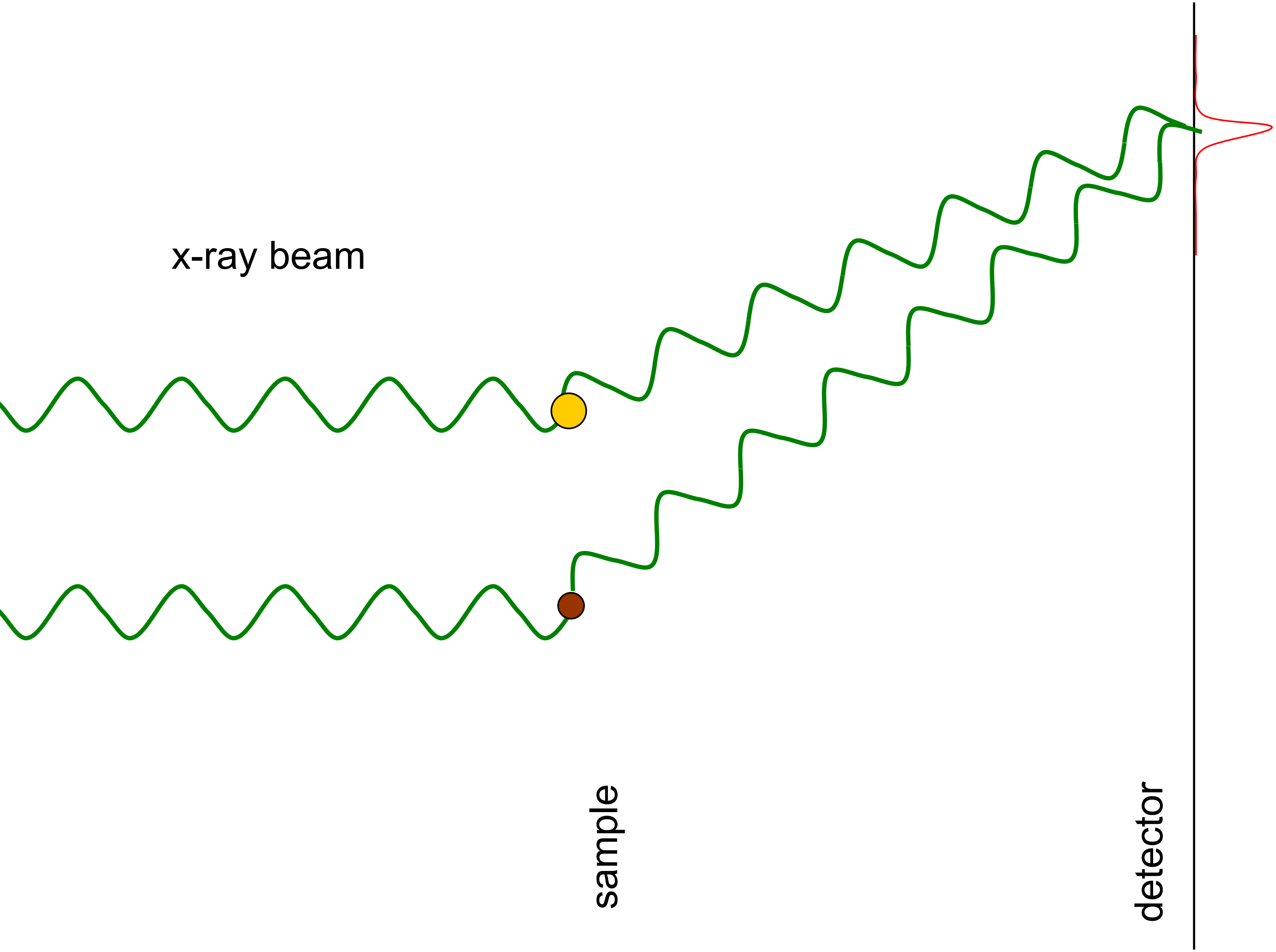


x-ray beam

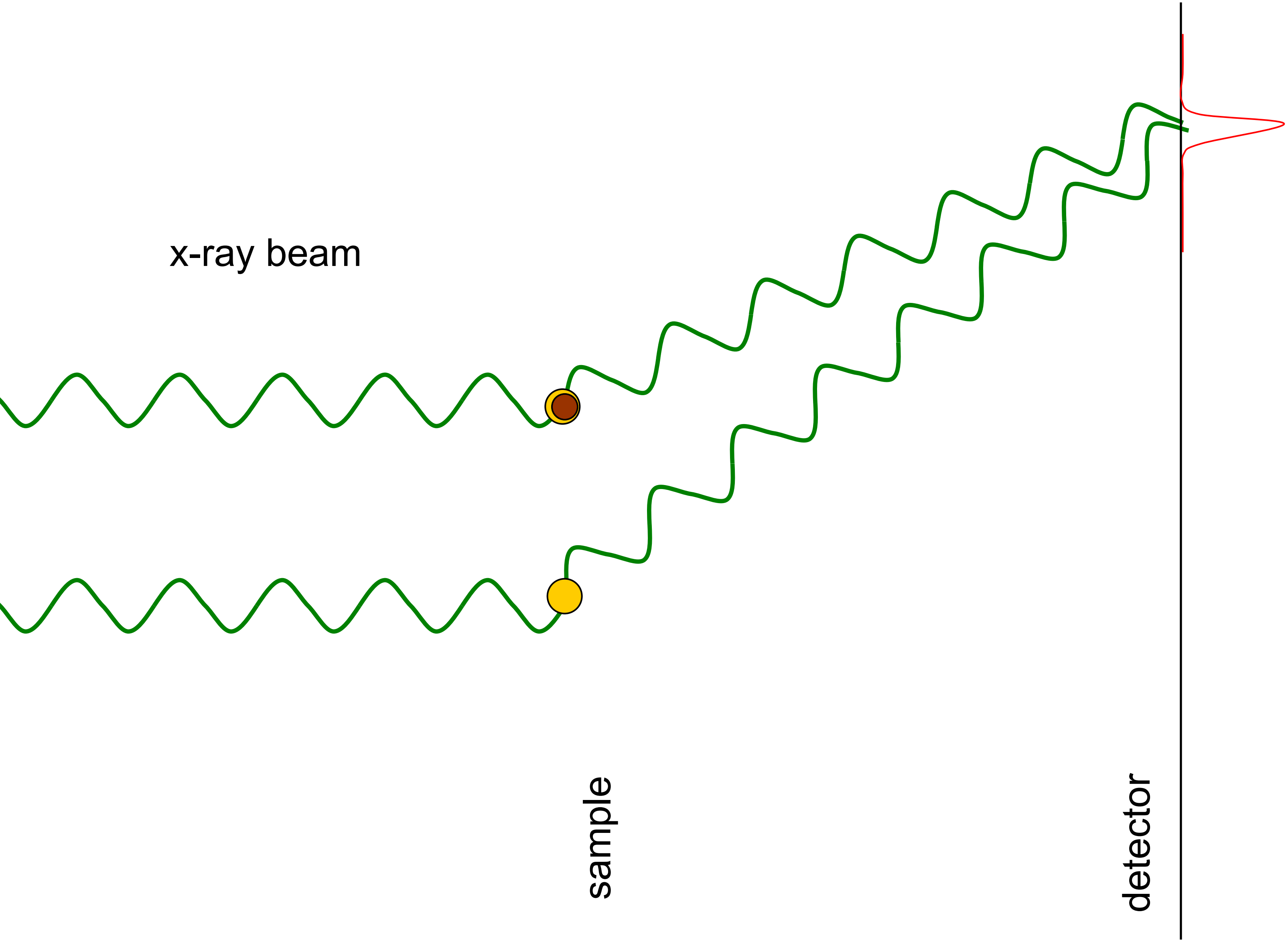
sample

detector

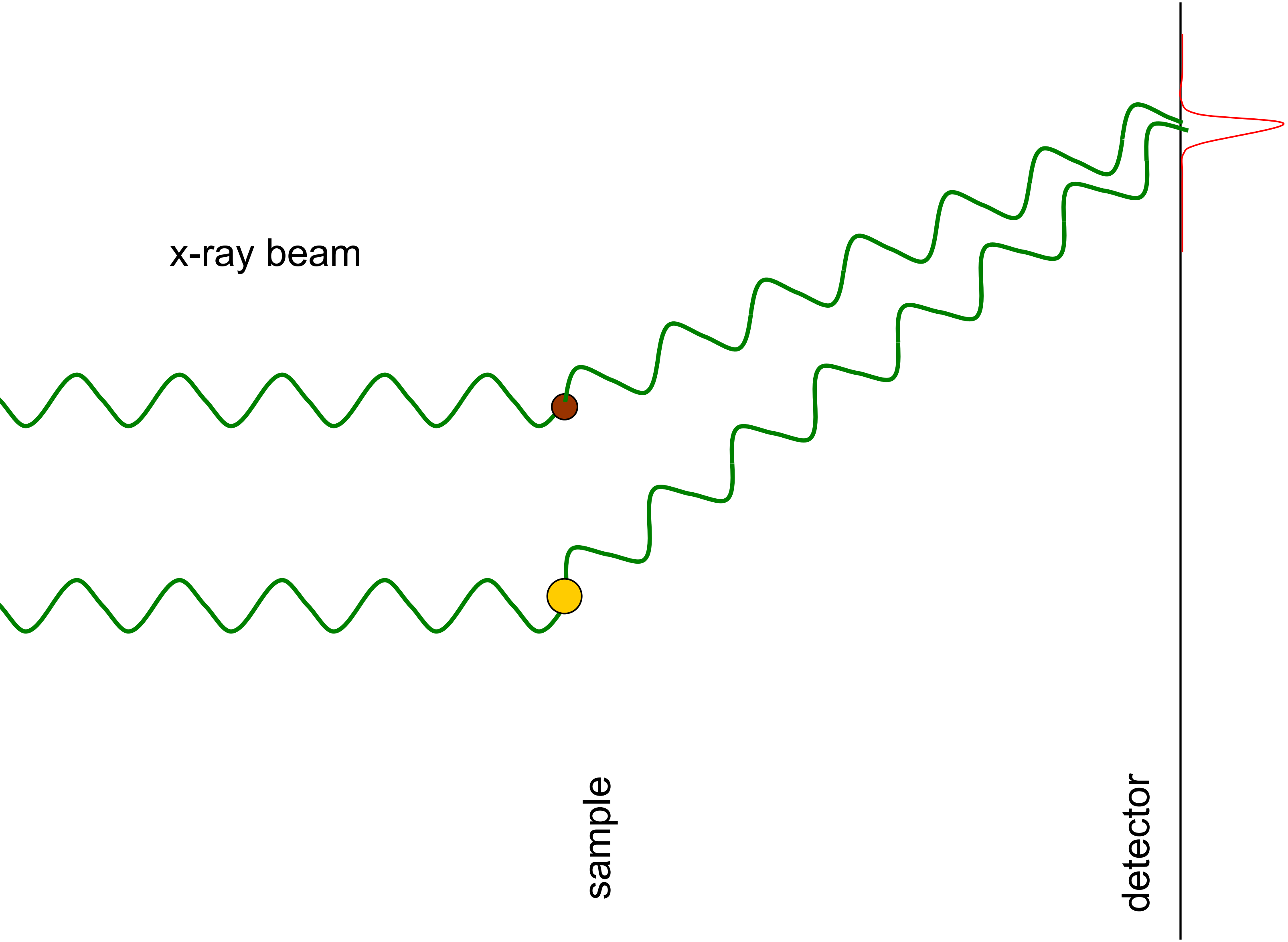
anomalous scattering



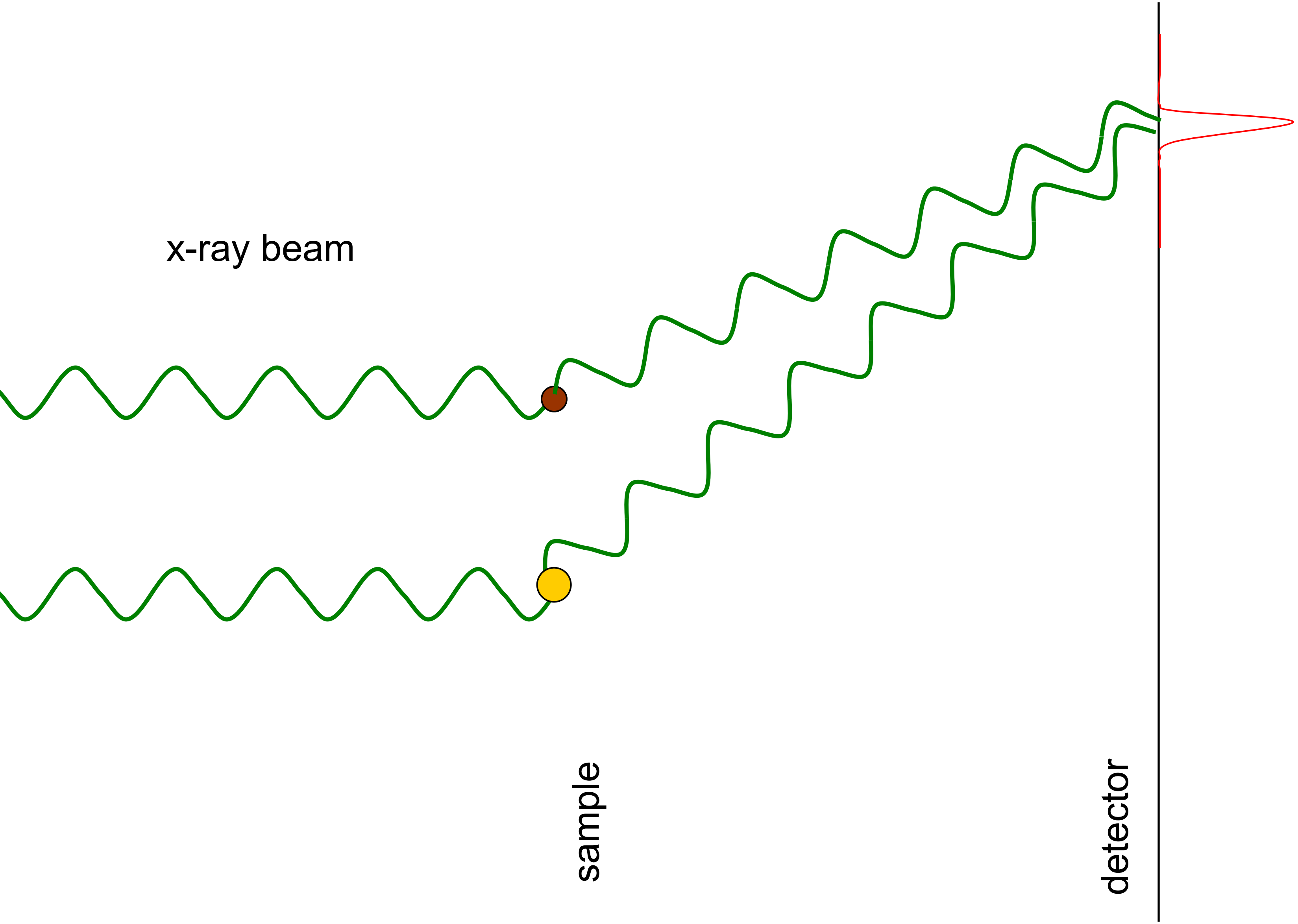
anomalous scattering



anomalous scattering



anomalous scattering

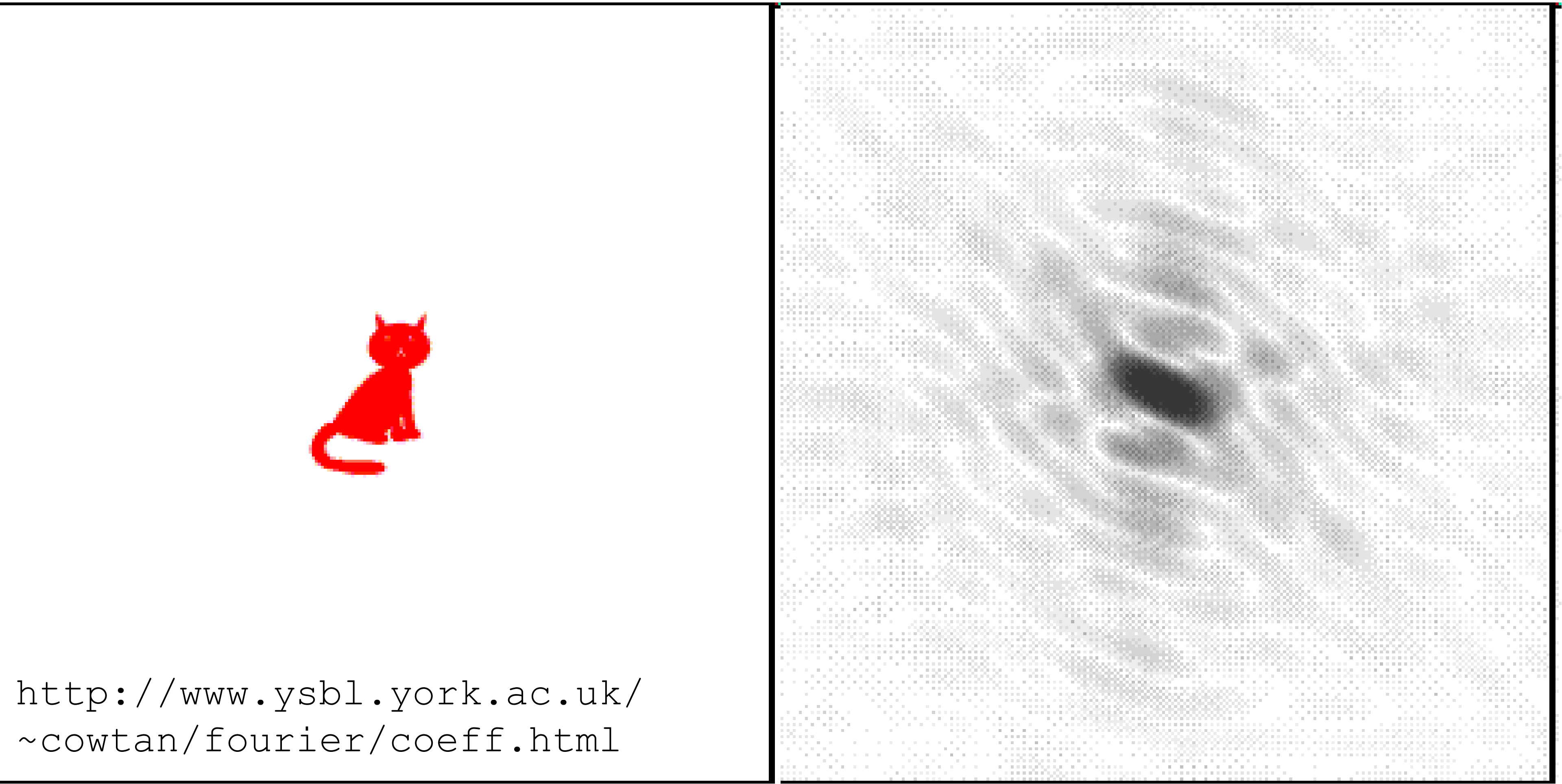


Major Phasing techniques

- Molecular Replacement
- Multiple Isomorphous Replacement
- Multiwavelength Anomalous Diffraction
- Single-wavelength Anomalous Diffraction

Molecular Replacement

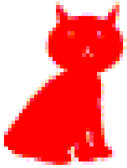
correct structure and intensities

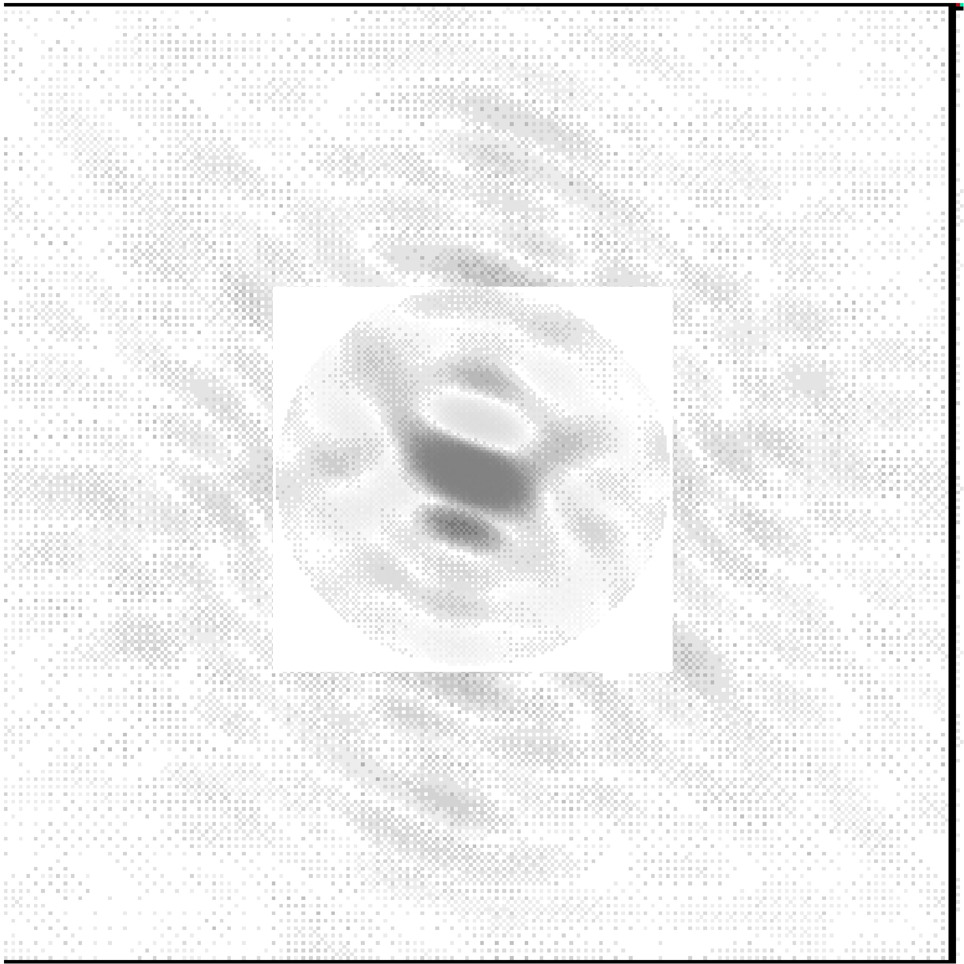


http://www.ysbl.york.ac.uk/~cowtan/fourier/coeff.html

Molecular Replacement

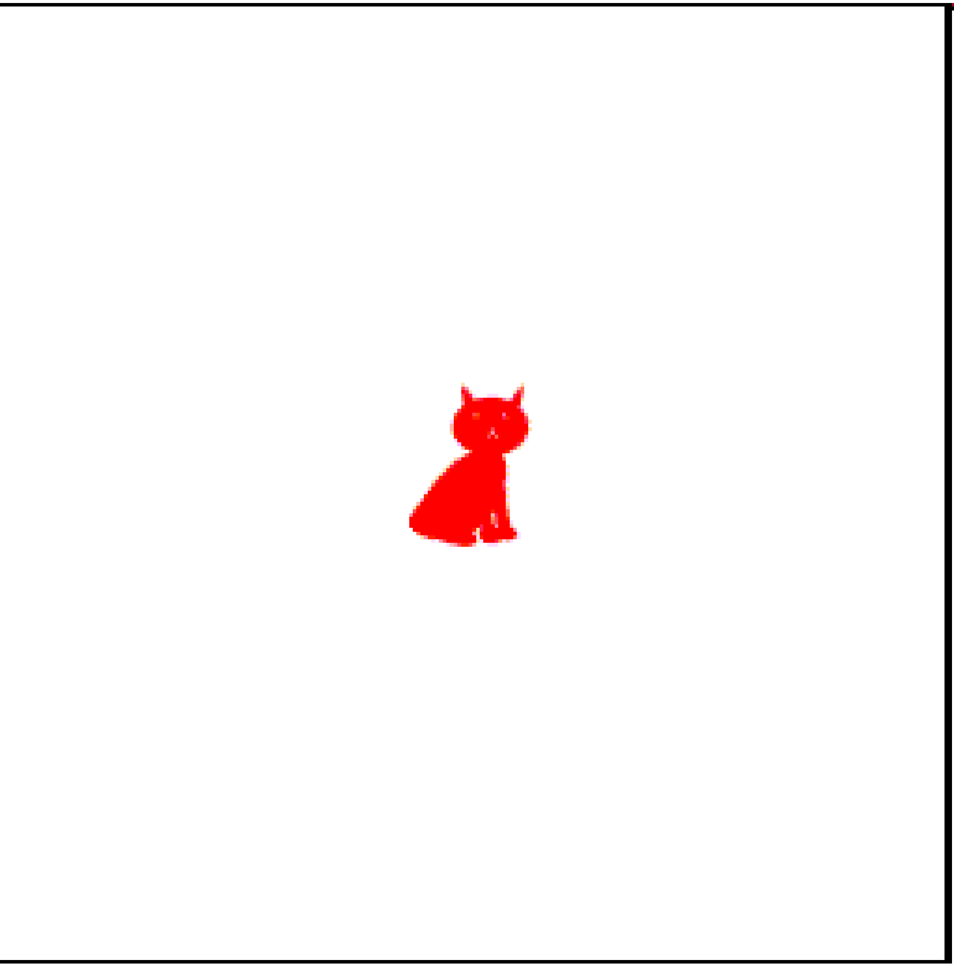
use something similar as a starting model

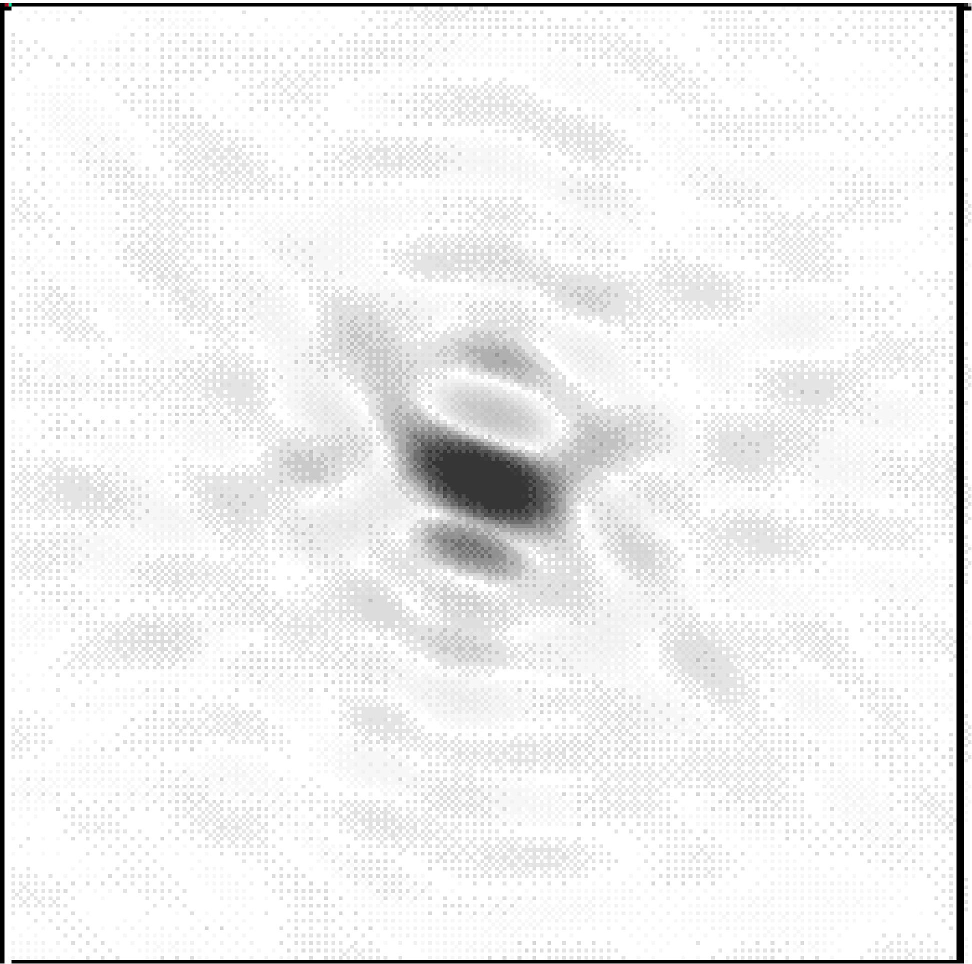




Model Building

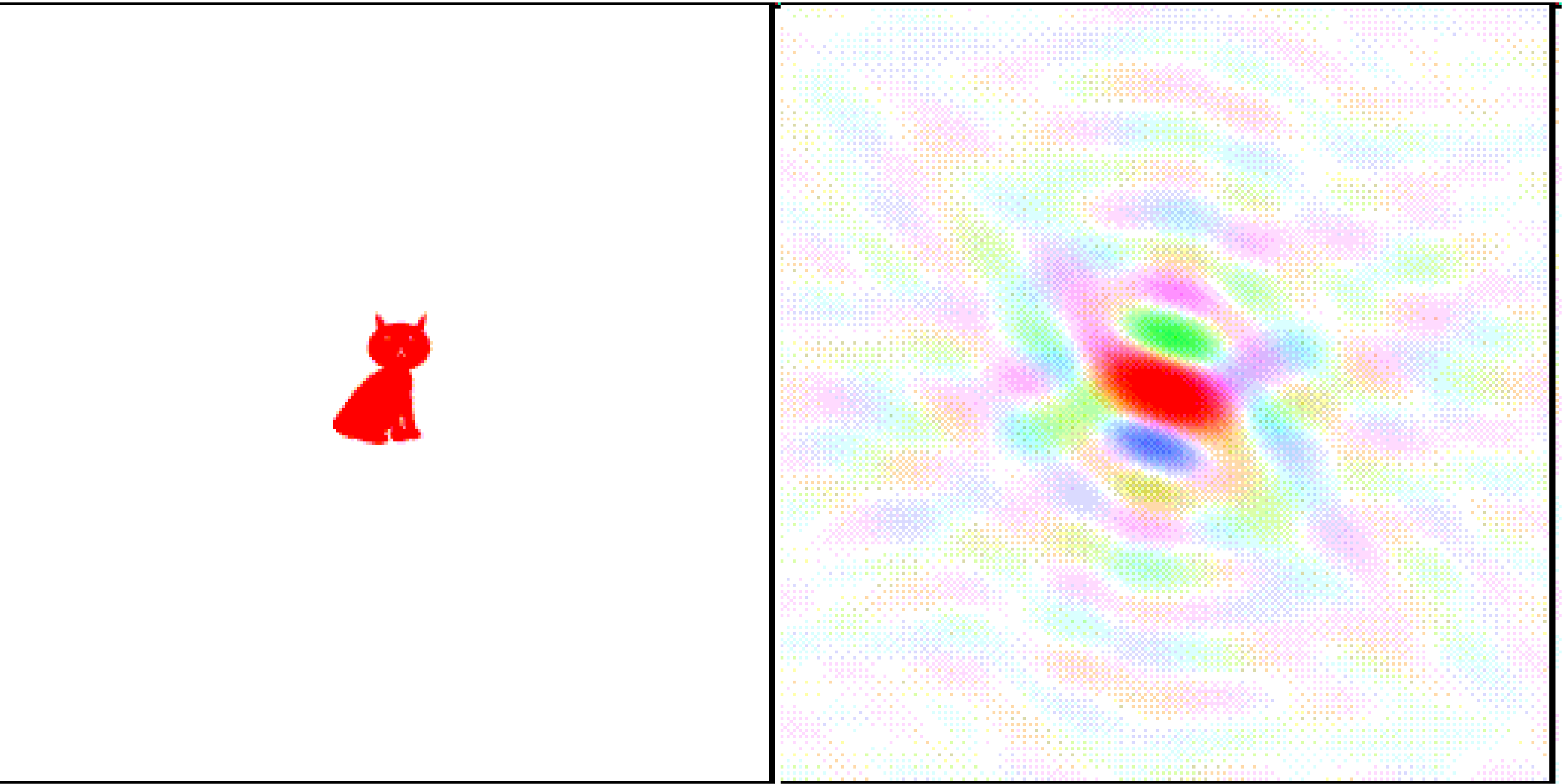
current model is missing something





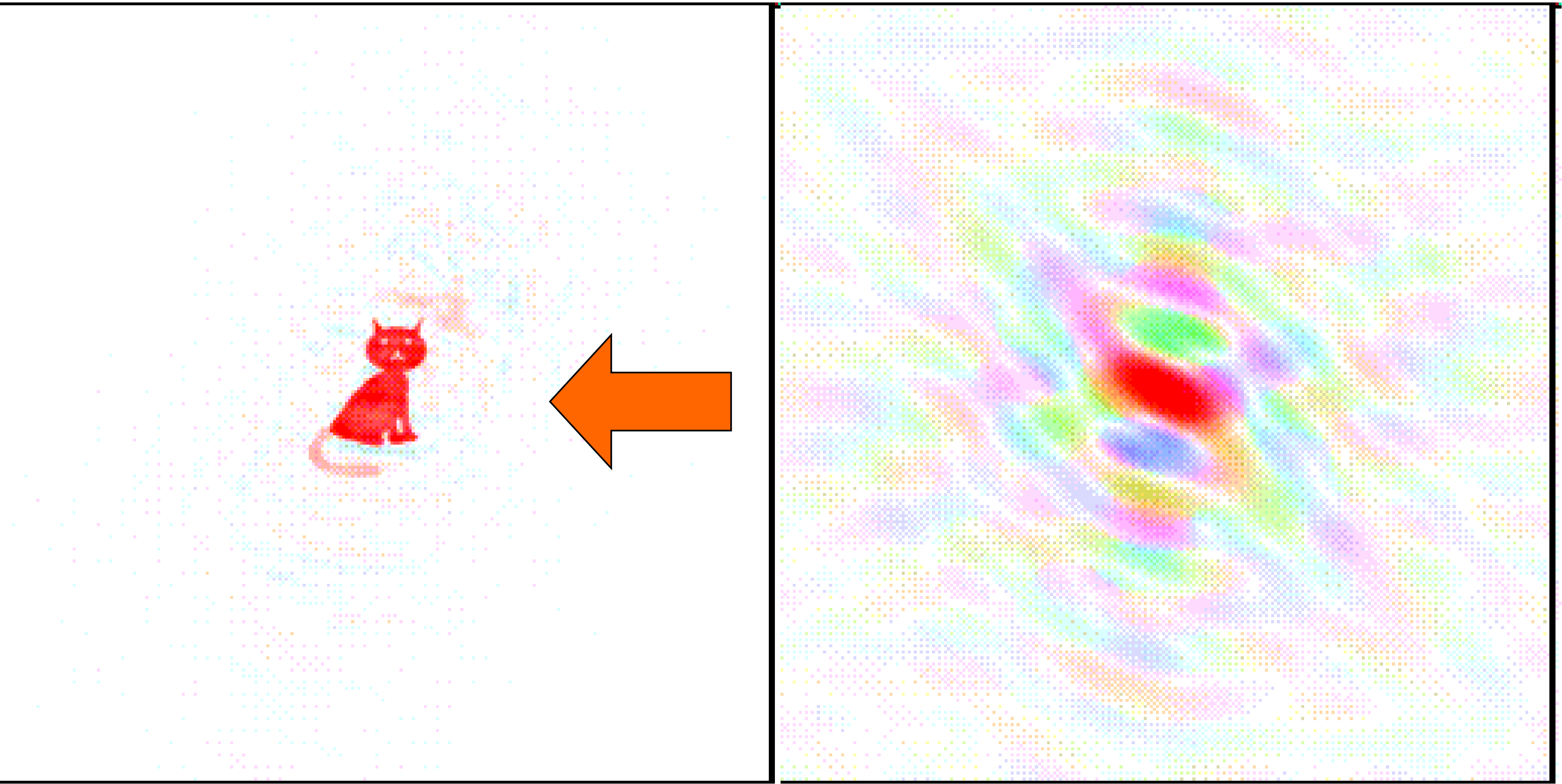
Model Building

phases from model



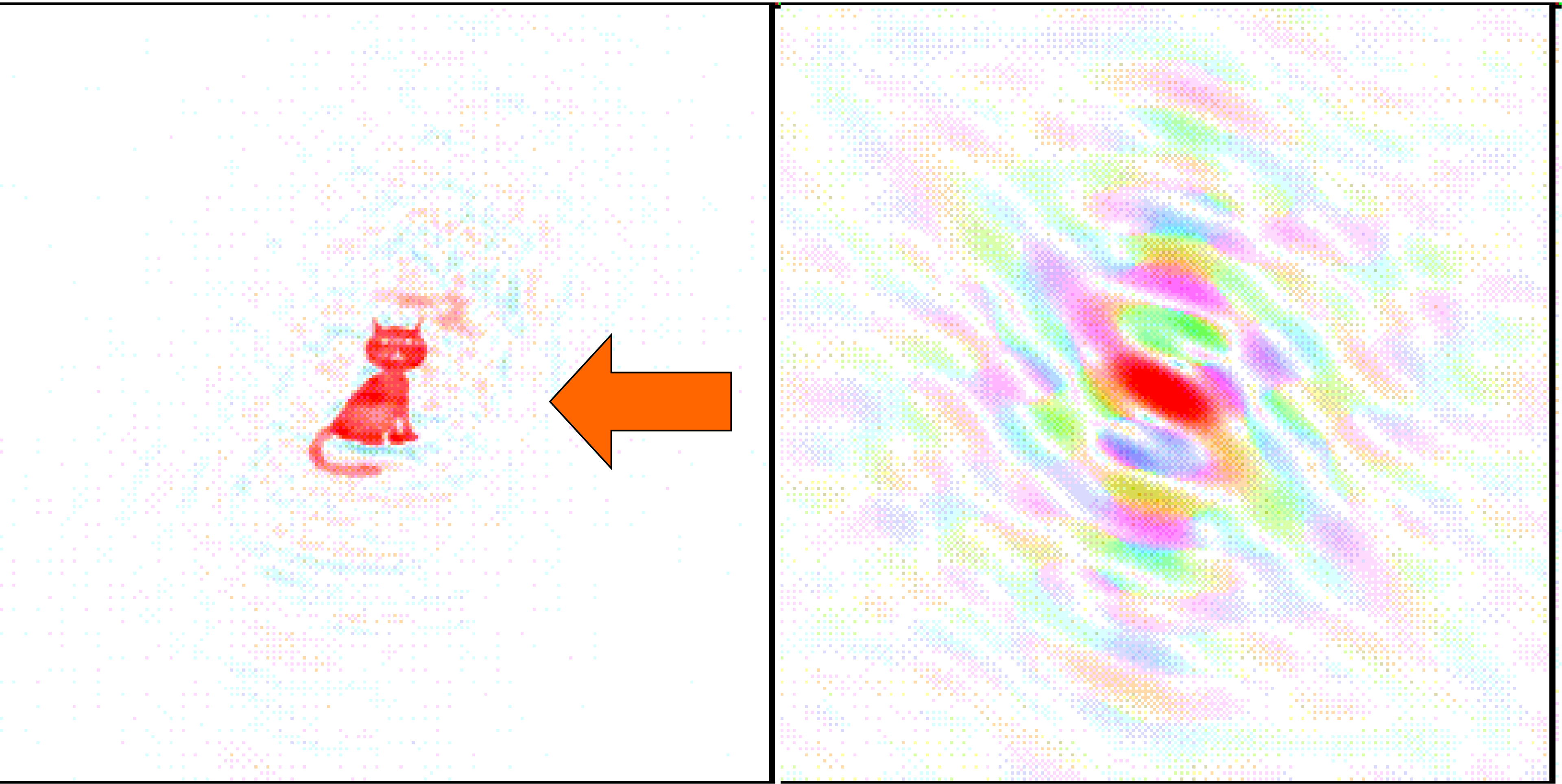
Model Building

missing bits show up in “difference map”



Model Building

missing bits show up better in $F_O + (F_O - F_C)$ map



Is the structure any good?

- Molecular replacement:
  - R/R_{free} (compare with R_{merge})
- MAD/SAD/MIR
  - R and FOM
- All cases
  - does the map look like the model?!

“R” factors

R = % error

“R” factors

$$R = \frac{\sum \left| F_{obs} - F_{calc} \right|}{\sum F_{obs}}$$

completely random: 0.59

starting MR solution: 0.4-0.55

something still wrong?: > 0.3

correct chain trace: < 0.2

small molecule: ~ 0.05

“R” factors

R_{cryst} (or R_{work}, or just “R”)

observed vs calculated data (Fs)

R_{free}

cross-check with “random” subset of data

should be < 0.3 and < R_{cryst} + 0.1

“R” factors

R_{cryst} (or just “R”)

observed vs calculated data (Fs)

R_{free}

cross-check with “random” subset of data
should be < 0.3 and < R_{cryst} + 0.1

R_{sym} = R_{merge} (self-consistency of data: Is)

R_{merge}

$$R_{merge} = \frac{\sum \left| I_{obs} - \langle I \rangle \right|}{\sum I_{obs}}$$

completely random:	**0.59**
weak data (high angle):	**0.7- ∞**
wrong symmetry choice?:	**~0.2-0.55**
small or disordered crystal:	**~0.1-0.2**
typical:	**~ 0.05**

via Parseval's Theorem

$$\frac{\rho}{\sigma(\rho)} \approx \frac{e^{-}}{7R}$$

e^{-}	number of electrons in peak
7	number of electrons in typical atom
R	R_{iso} or R_{free}
ρ	electron density peak
σ	rms deviation expected

The UNTANGLE Challenge

`best.pdb`

`refme.mtz`

1) Build better than the best:

as-good or better wE, R_{free} as best.pdb

non-equivalent ensemble

$1000

2) Recover best.pdb from refme.mtz

true phases, auto-building allowed

no cheating (don’t use best.pdb as guide)

$500

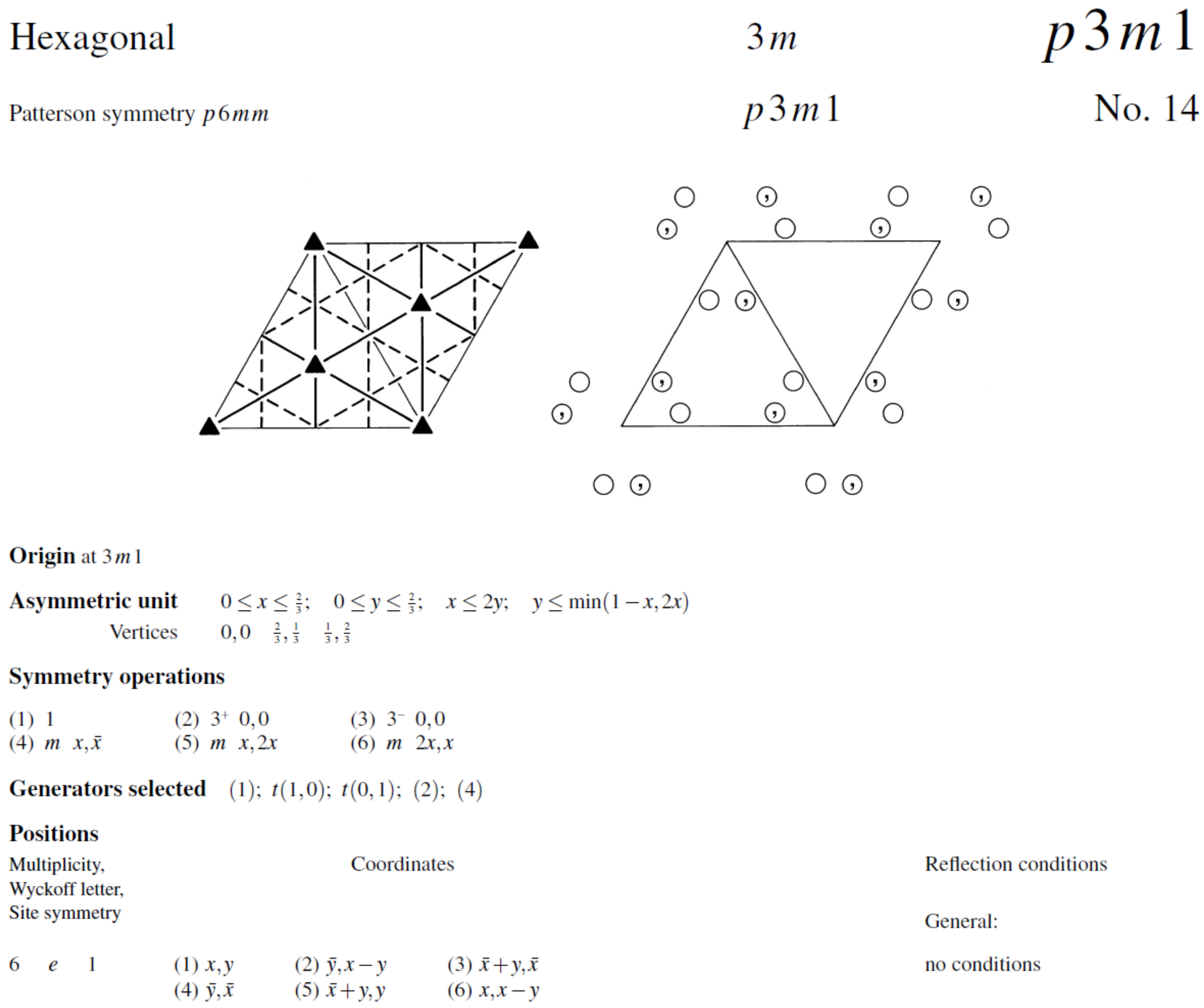
Prove: true ensemble is unique
and can be recovered

https://github.com/jmholton/UnTangle

What is your Space Group?

- What is a crystal?
- What is a lattice?
- What is a unit cell?
- What is a symmetry operator?
- What is an asymmetric unit?
- Pitfalls: the “screwies”, C vs P3, alternate index
- How will I know?

International Tables, Vol A



Hexagonal $3m$ $p3m1$

Patterson symmetry $p6mm$ $p3m1$ No. 14

Origin at $3m1$

Asymmetric unit	$0 \le x \le \frac{2}{3}$;	$0 \le y \le \frac{2}{3}$;	$x \le 2y$;	$y \le \min(1-x, 2x)$
Vertices	$0,0$	$\frac{2}{3},\frac{1}{3}$	$\frac{1}{3},\frac{2}{3}$	

Symmetry operations

(1) 1	(2) 3^{+} $0,0$	(3) 3^{-} $0,0$
(4) m $x,\bar{x}$	(5) m $x,2x$	(6) m $2x,x$

Generators selected (1); $t(1,0)$; $t(0,1)$; (2); (4)

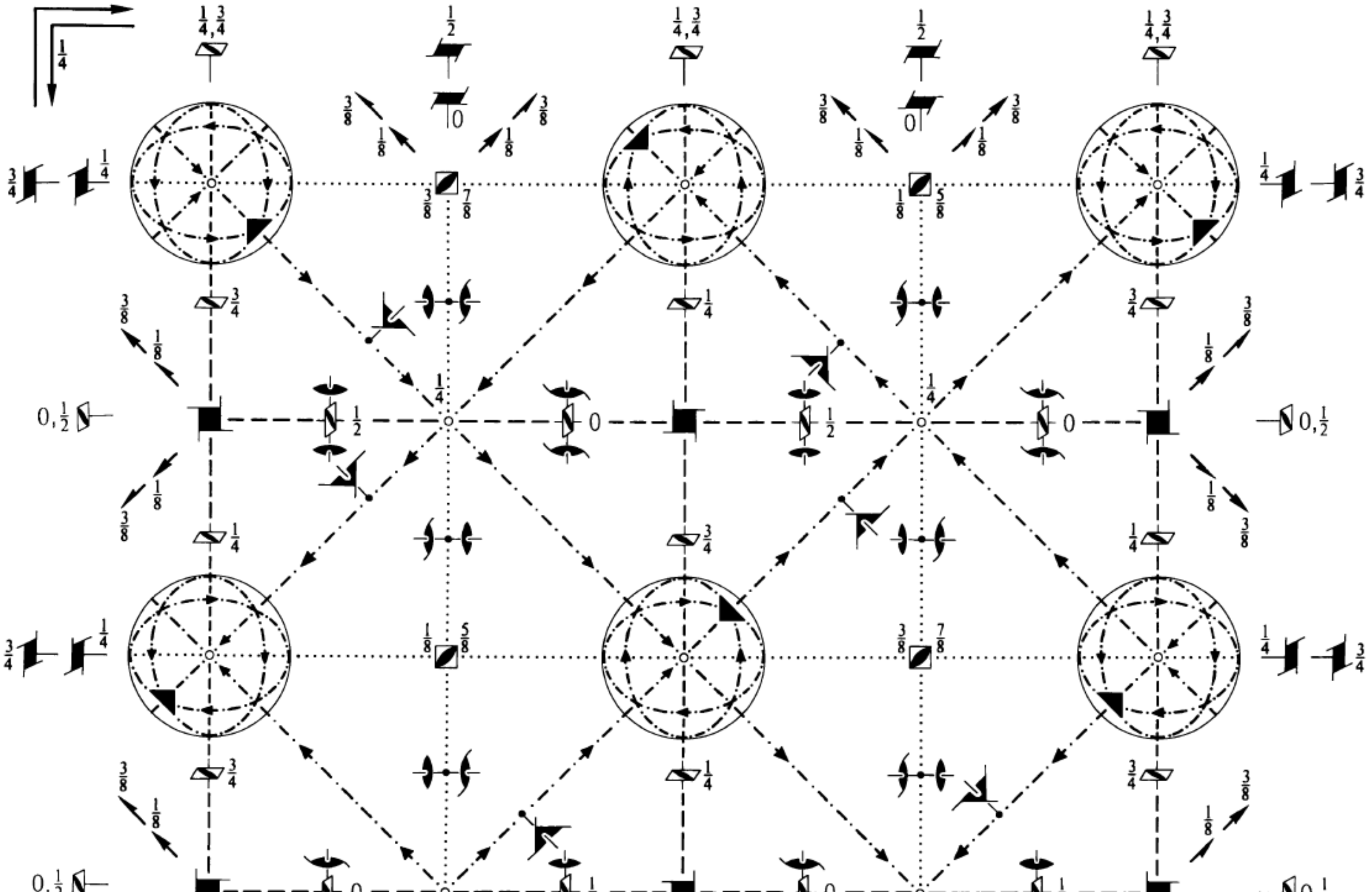
Positions

Multiplicity, Wyckoff letter, Site symmetry	Coordinates			Reflection conditions General:
6 e 1	(1) x,y	(2) $\bar{y},x-y$	(3) $\bar{x}+y,\bar{x}$	no conditions
	(4) $\bar{y},\bar{x}$	(5) $\bar{x}+y,y$	(6) $x,x-y$	

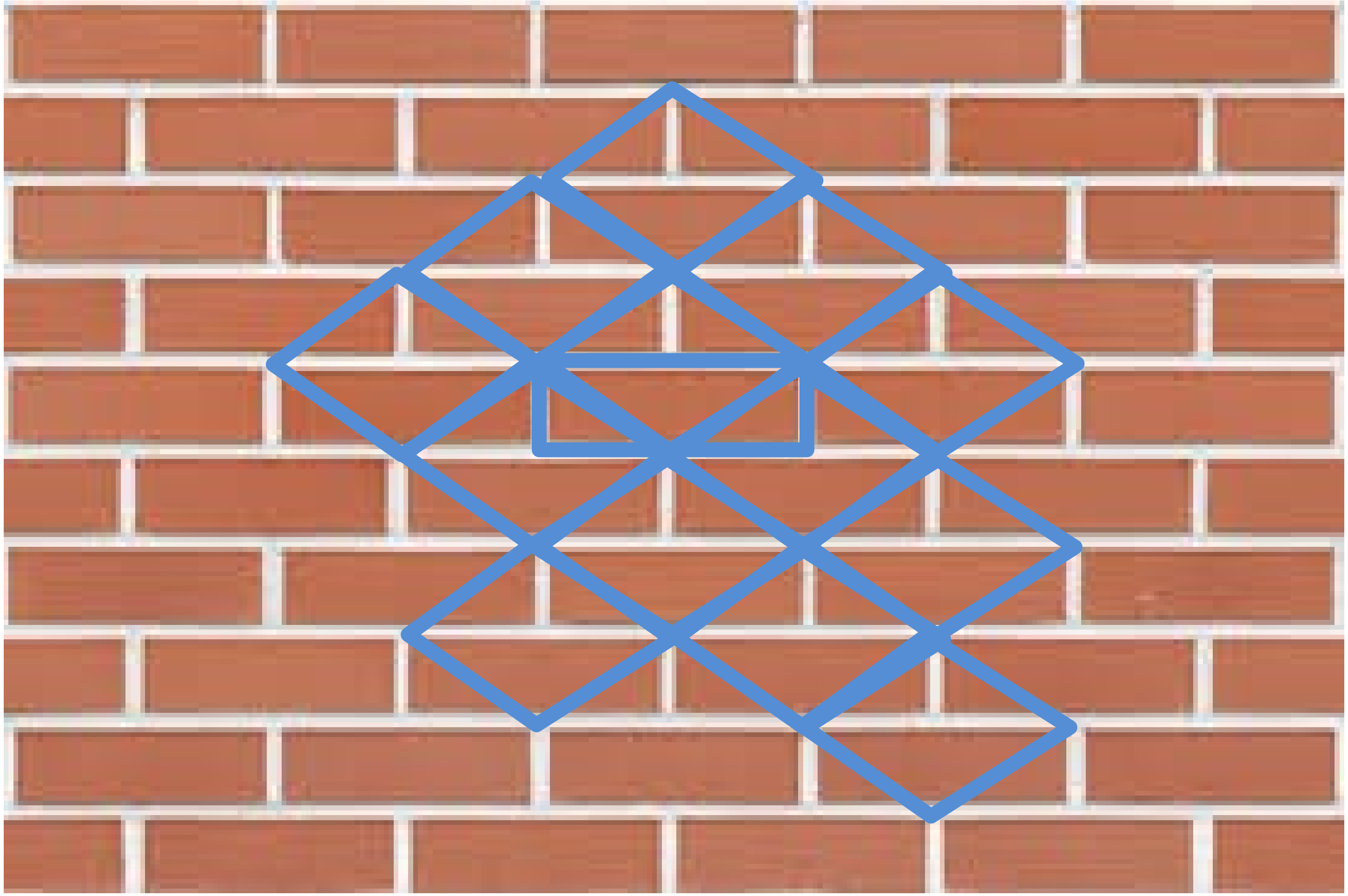
International Tables, Vol A

$Ia\bar{3}d$ O_h^{10} $m\bar{3}m$ Cubic

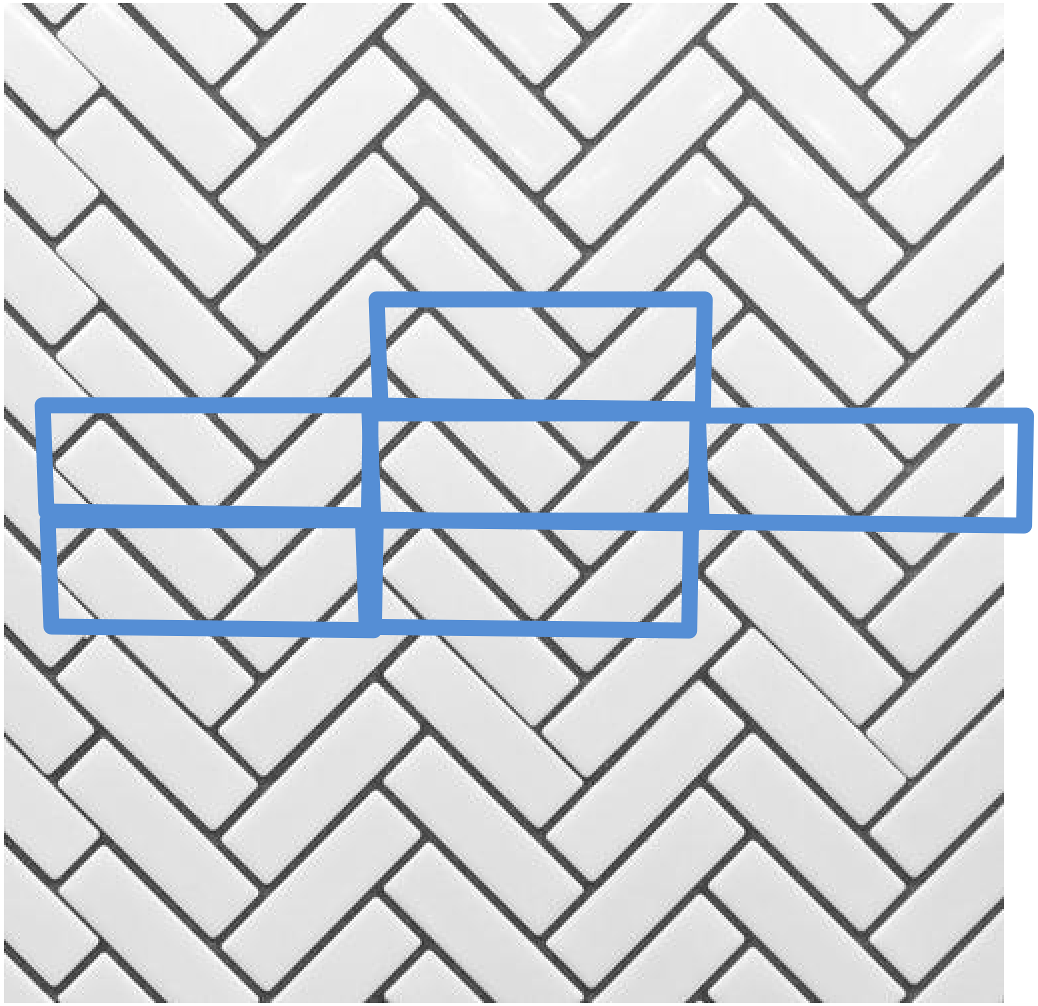
No. 230 $I\ 4_1/a\ \bar{3}\ 2/d$ Patterson symmetry $Im\bar{3}m$



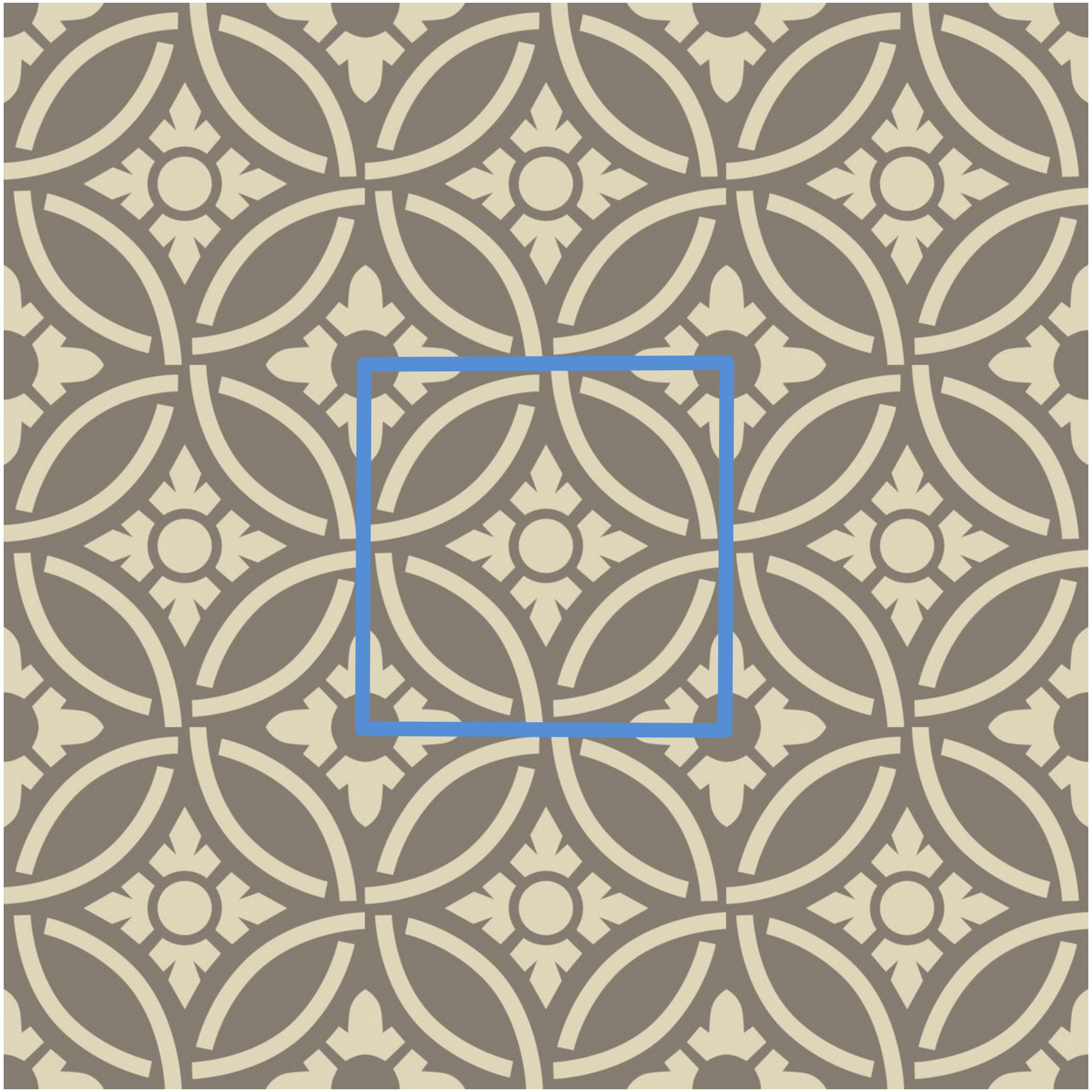
Unit cell?



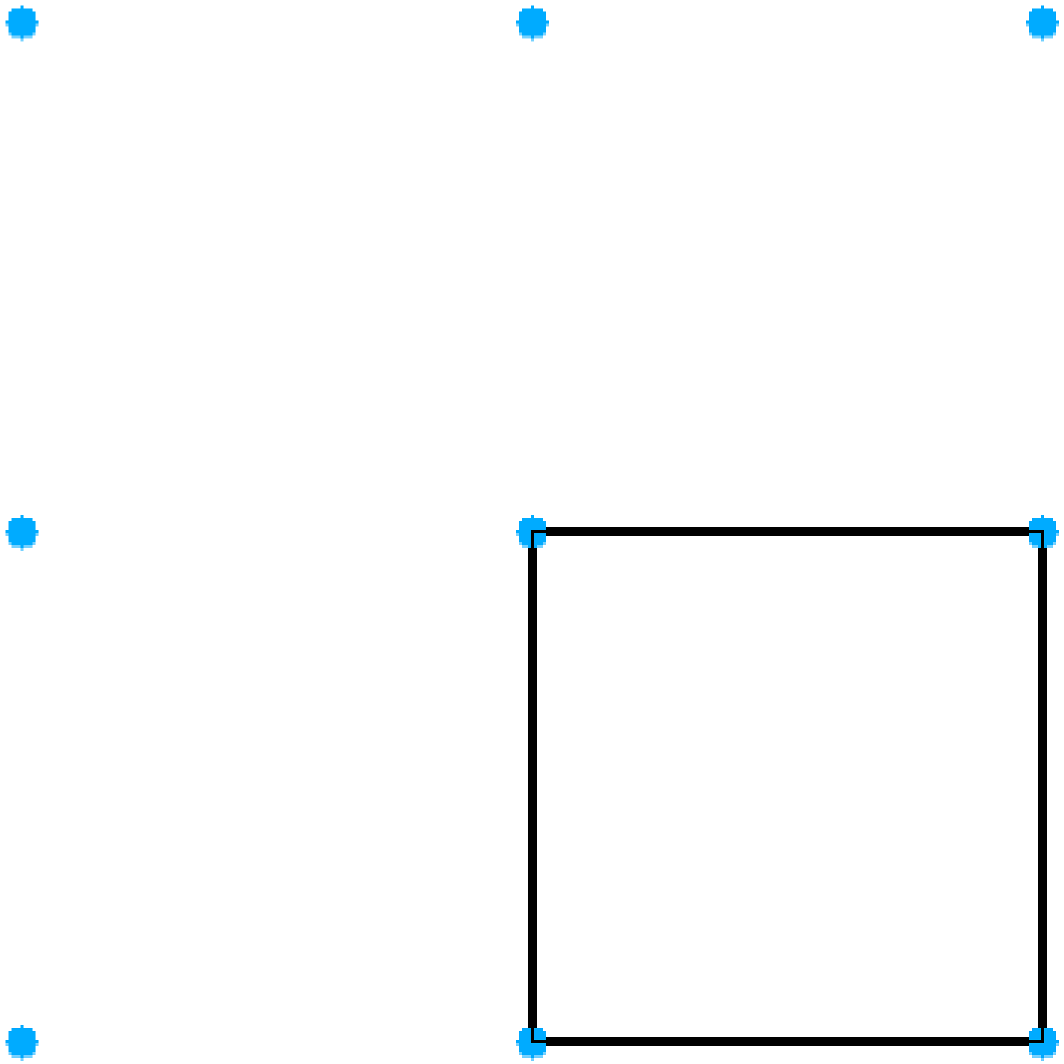
Where's the unit cell?



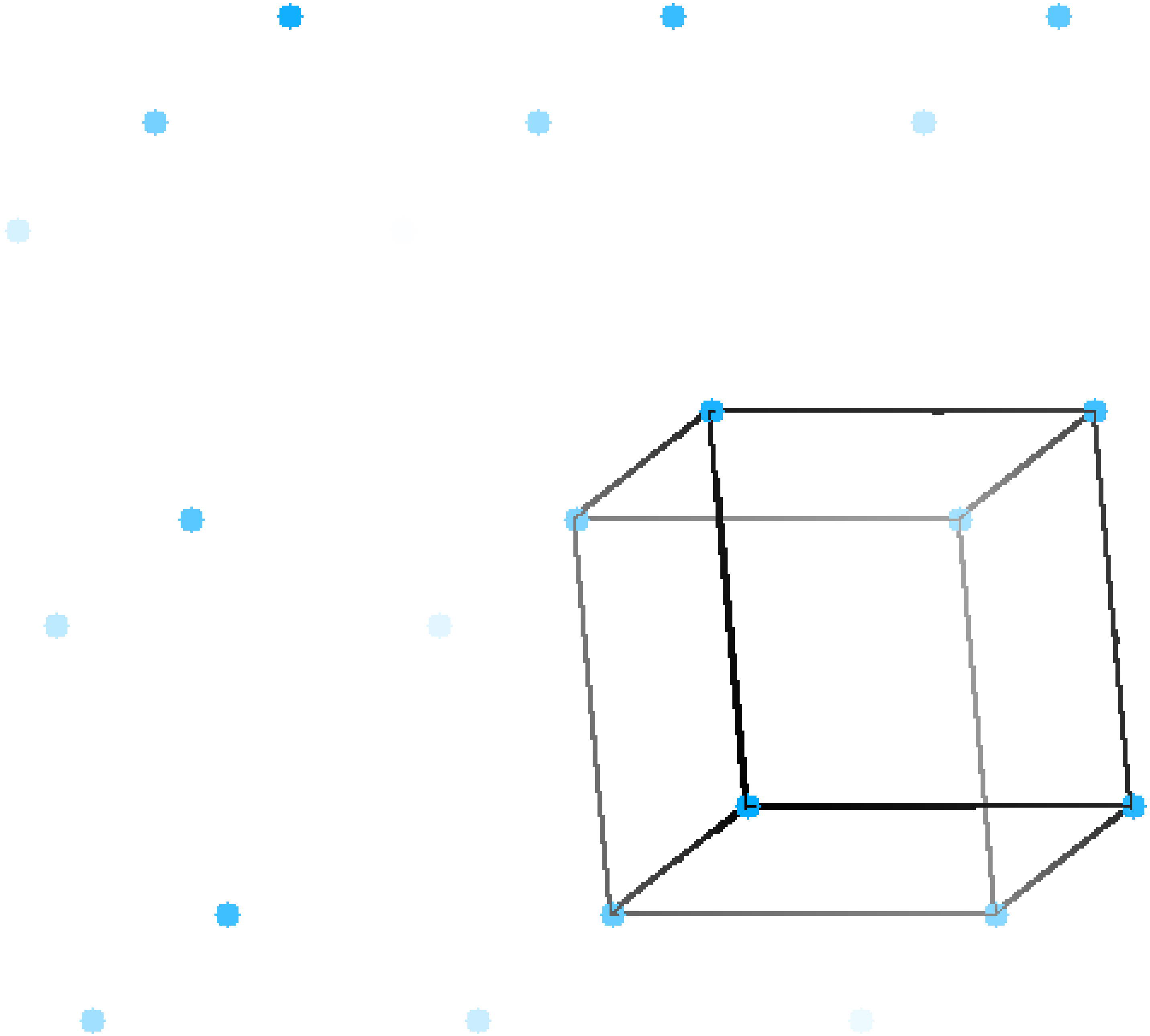
Where’s the unit cell?



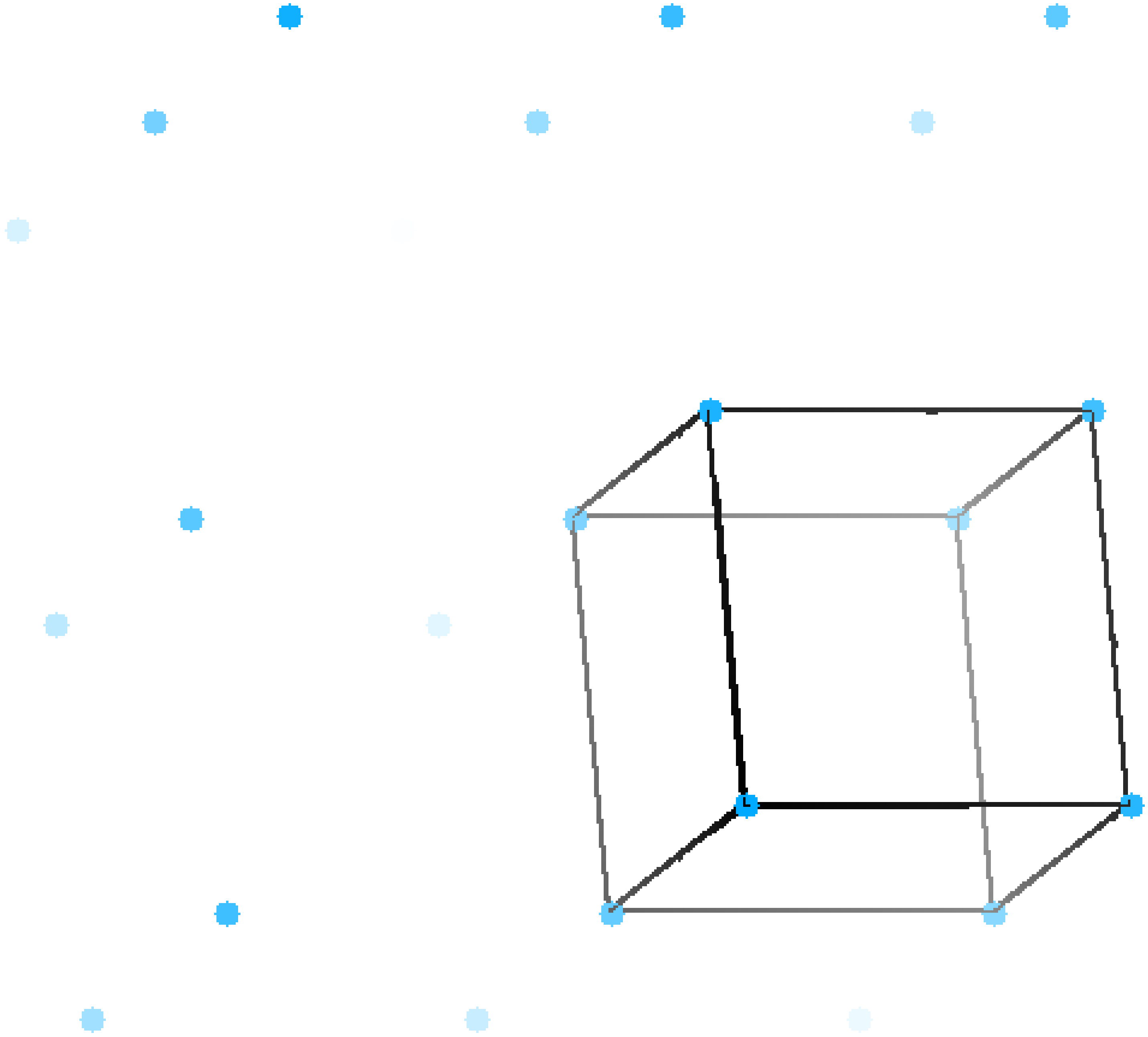
Primitive, cubic lattice



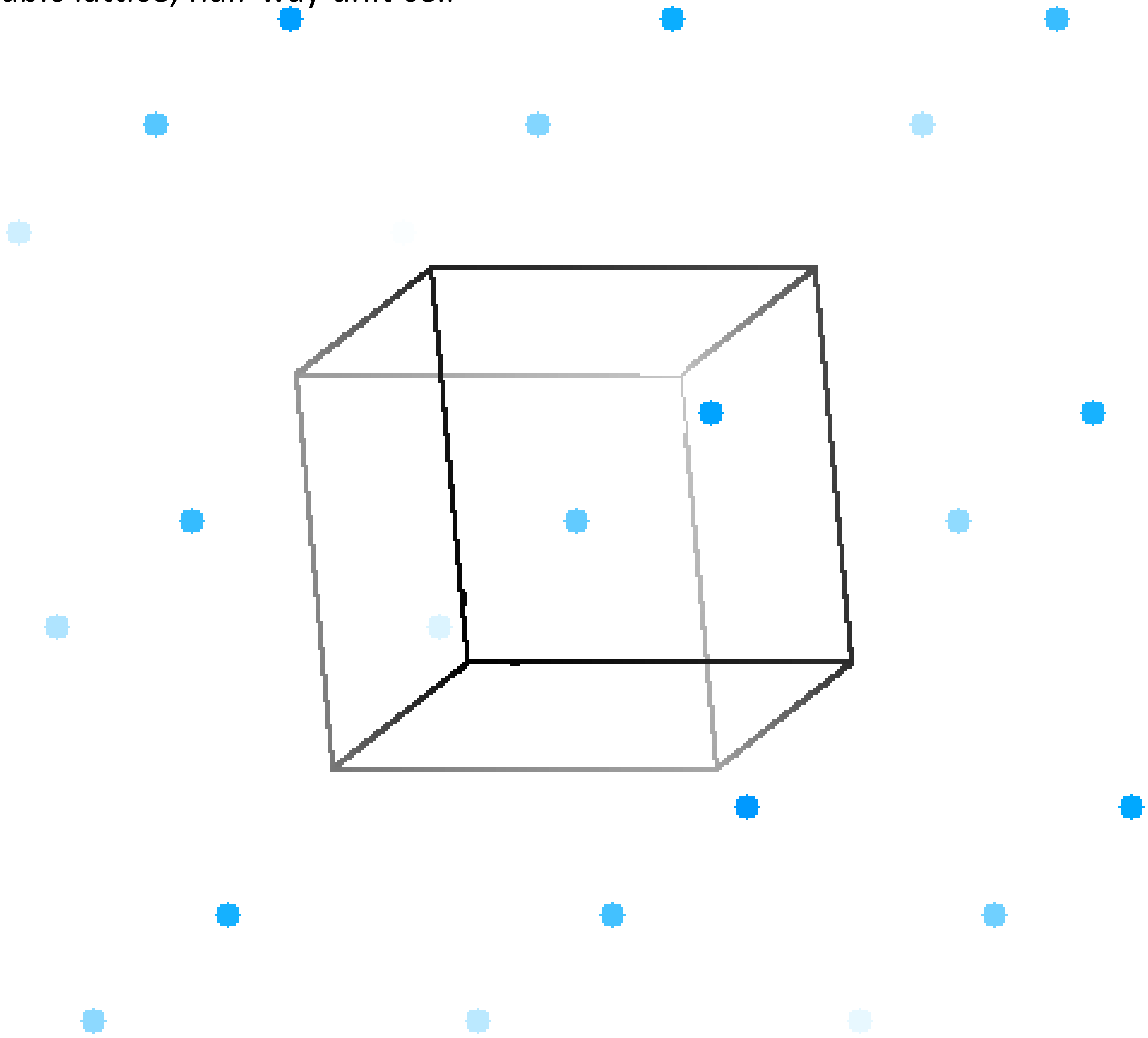
Primitive, cubic lattice



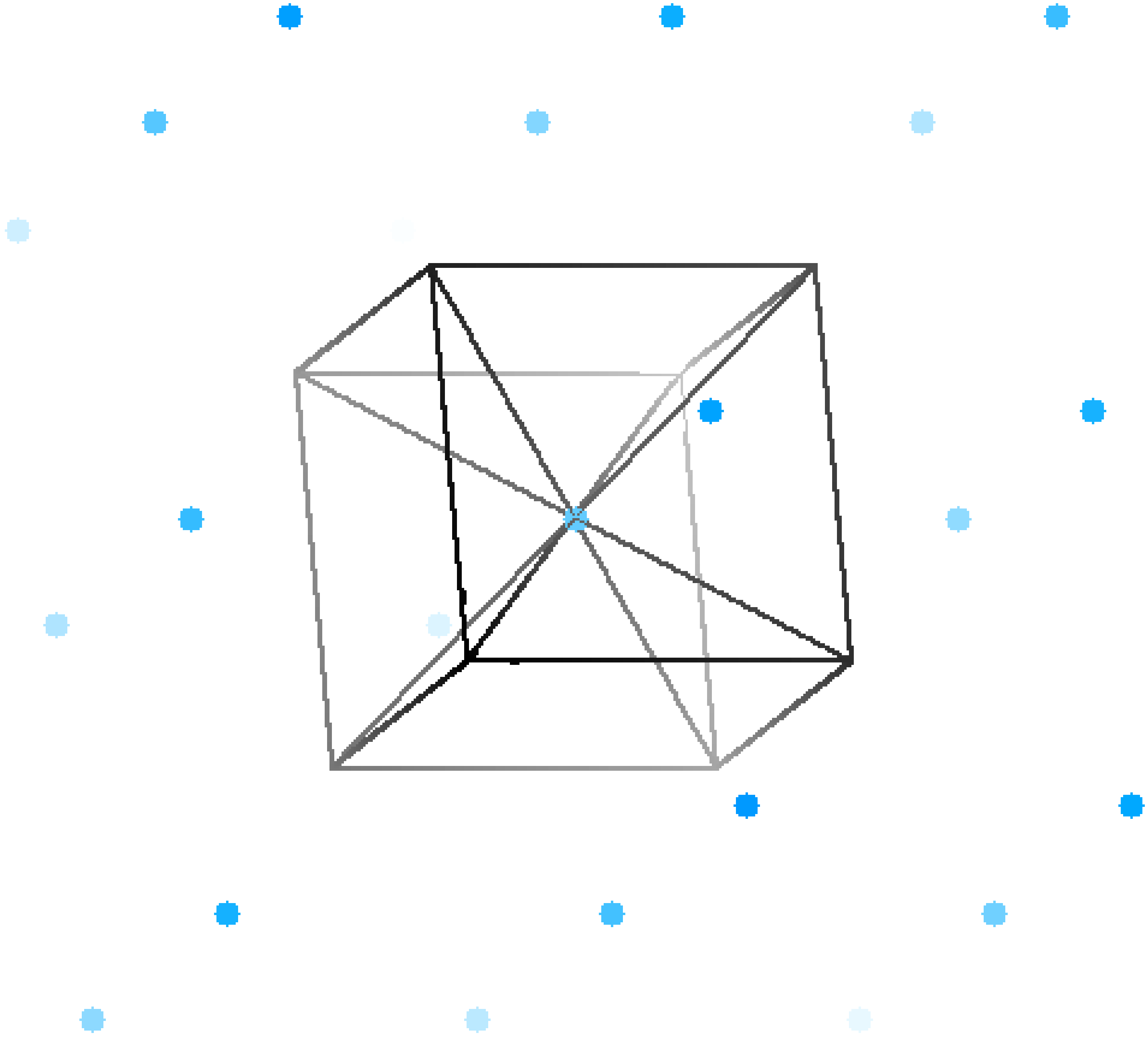
Primitive, cubic lattice



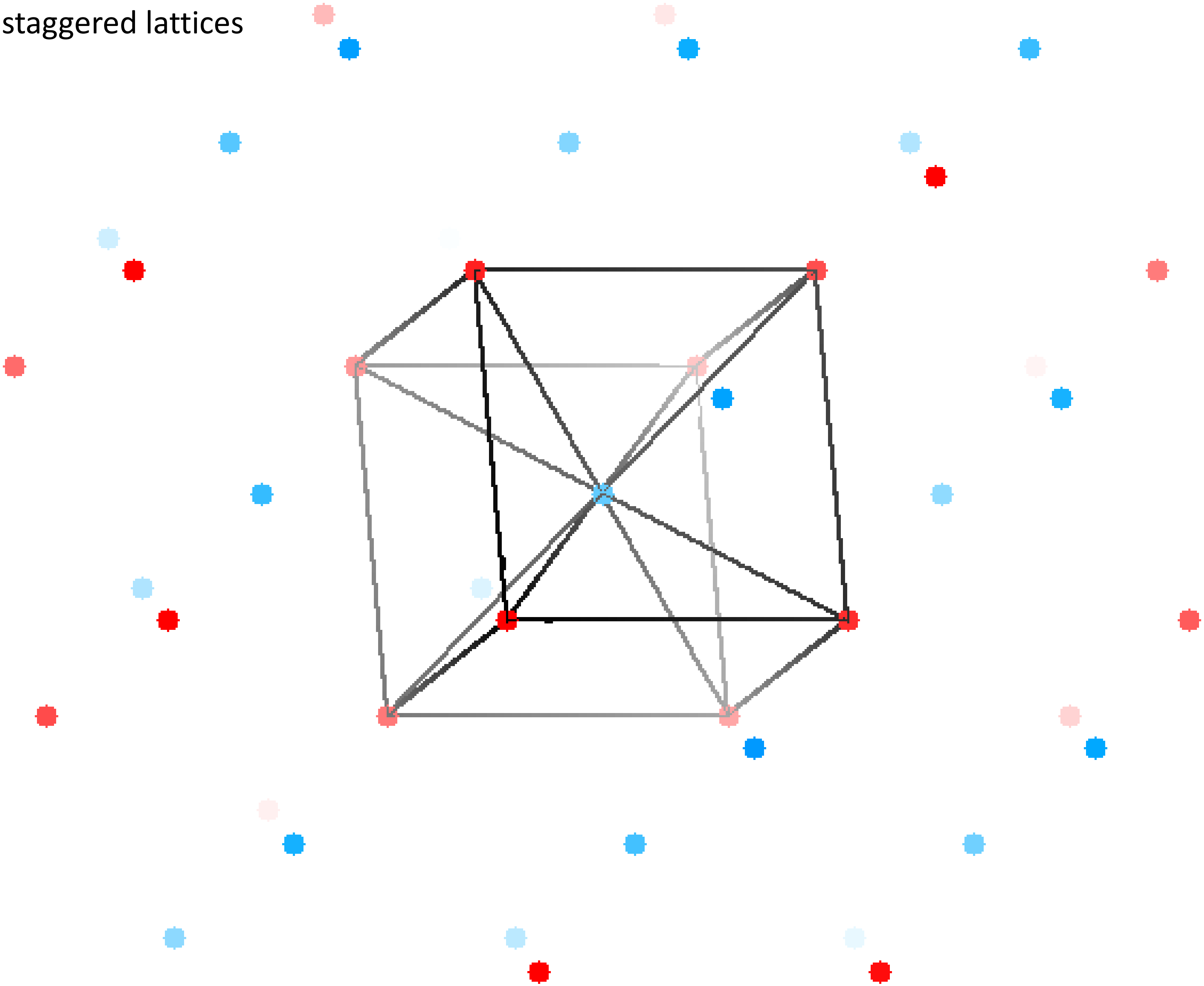
Primitive, cubic lattice, half-way unit cell



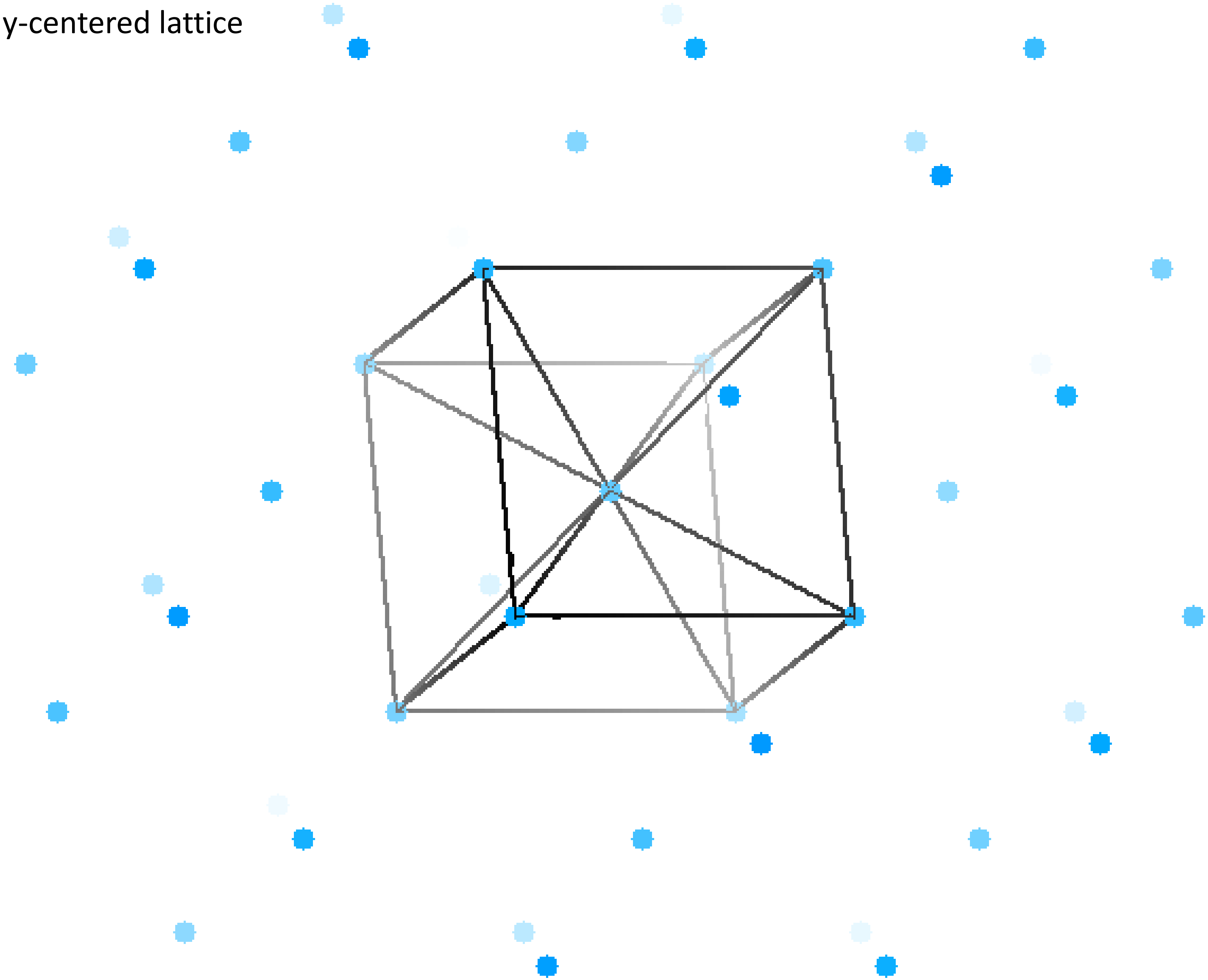
Body-centered cell



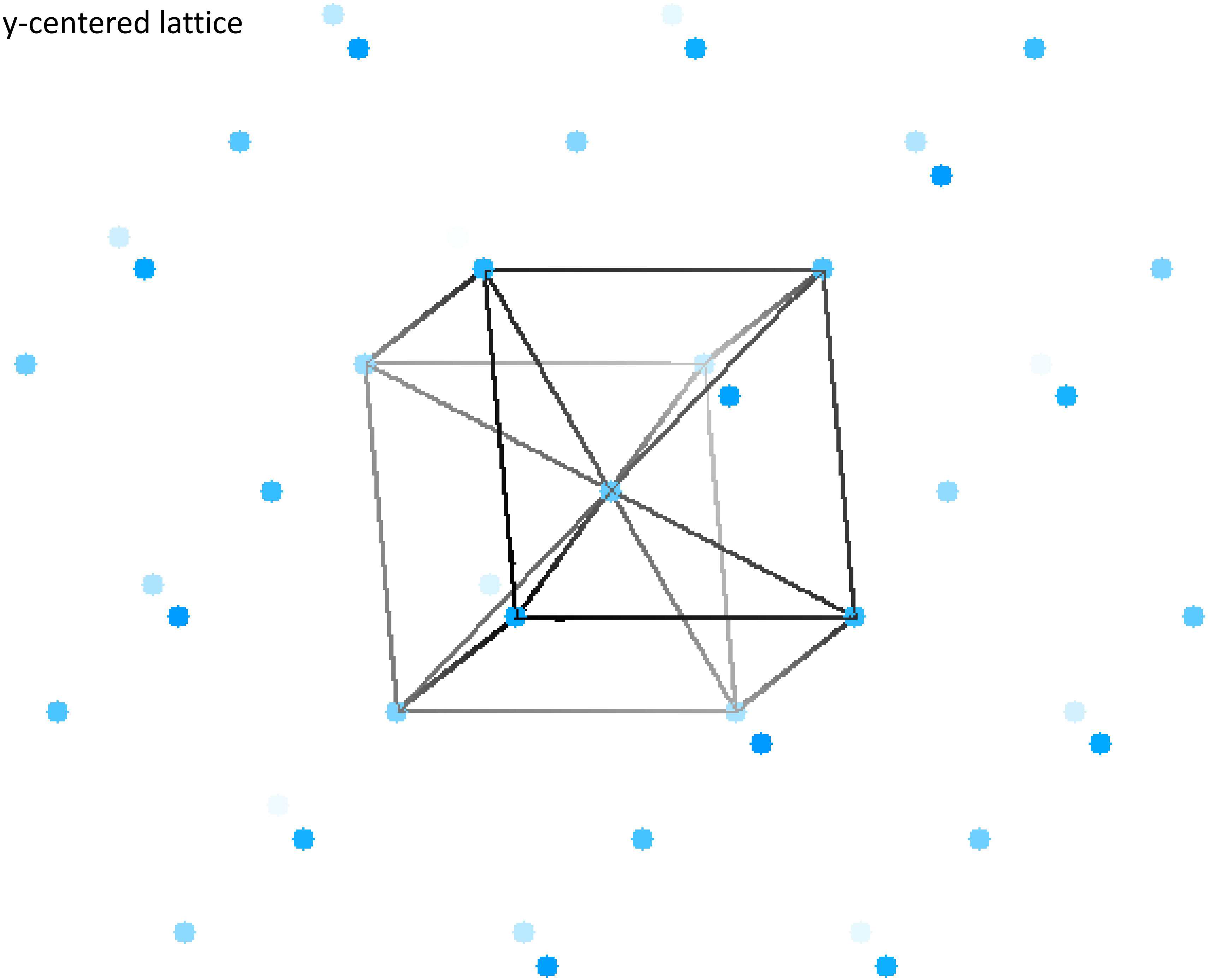
Two staggered lattices



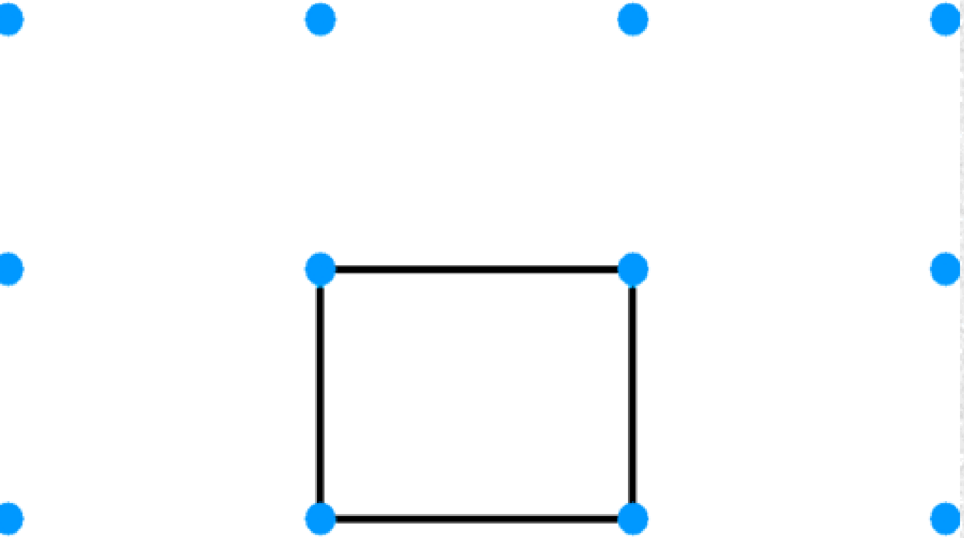
Body-centered lattice



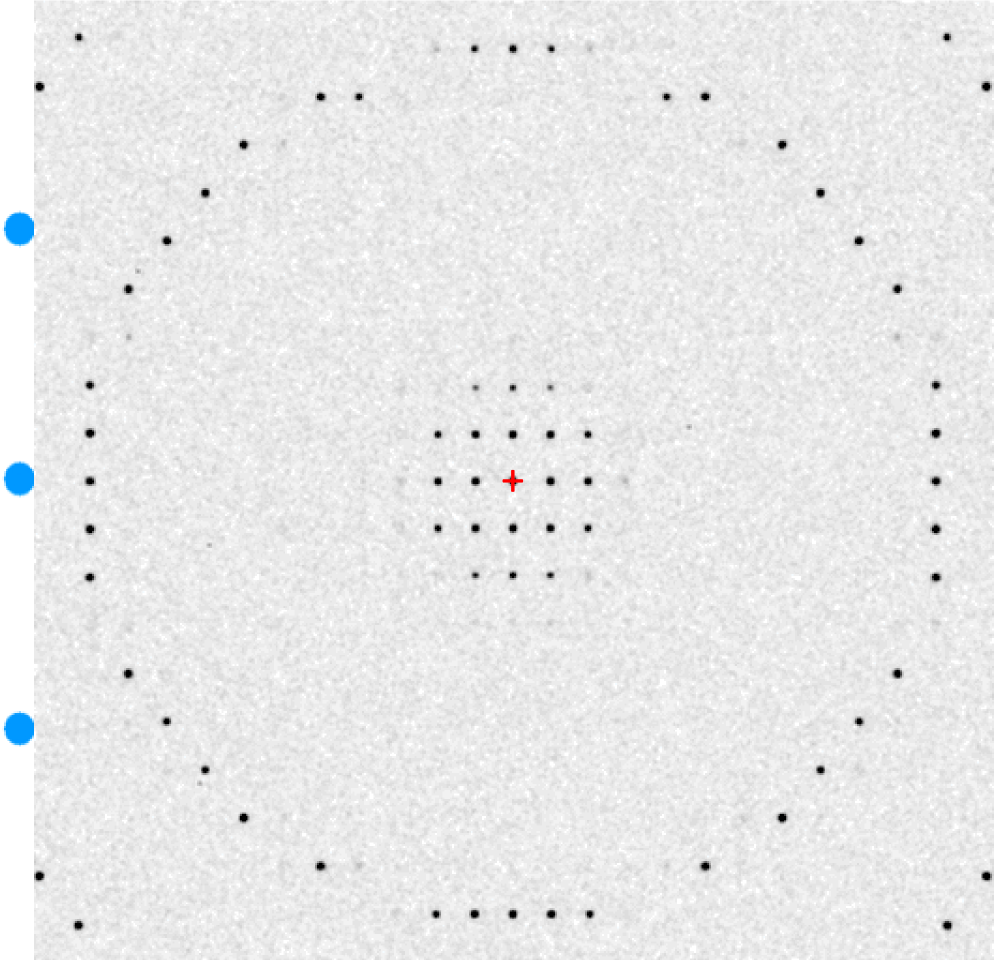
Body-centered lattice



Diffraction from a 3D lattice

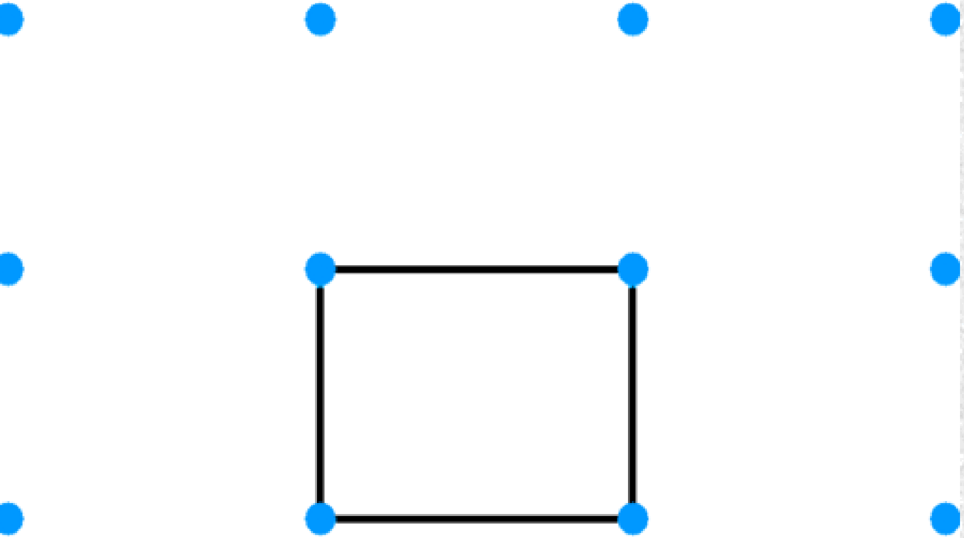


real space

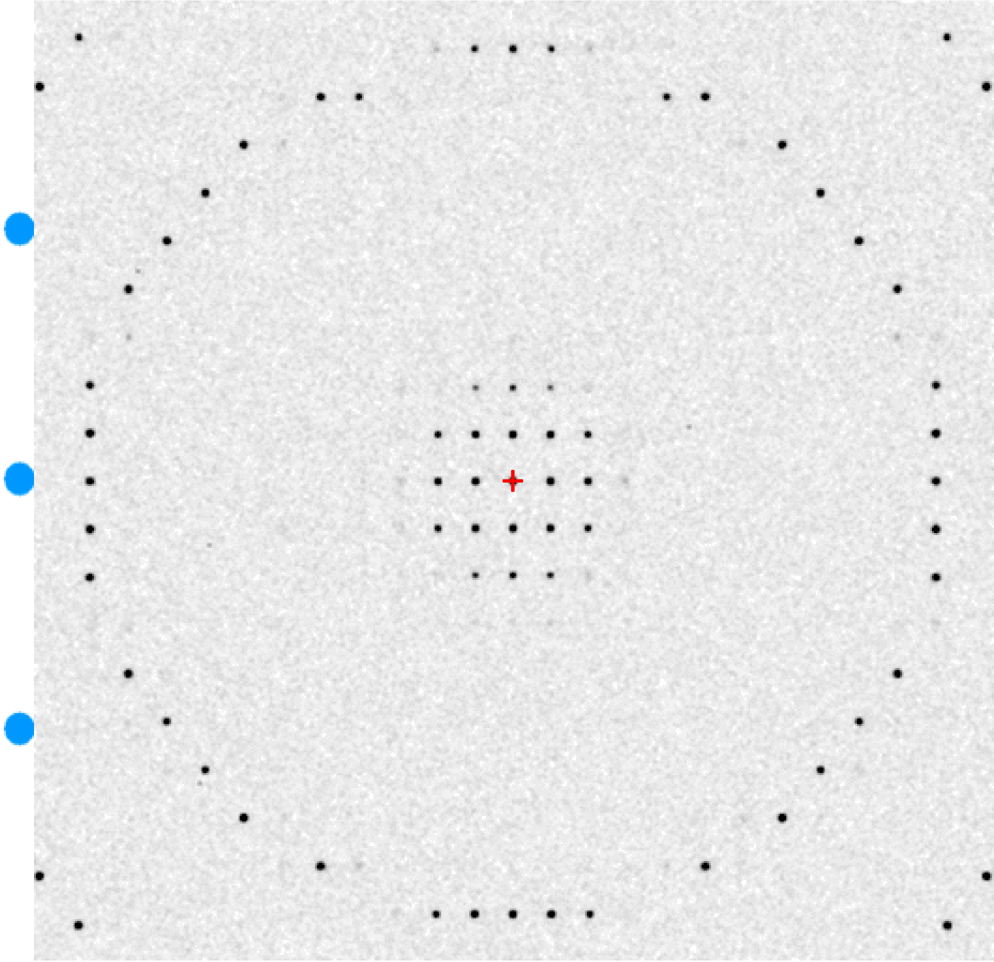


reciprocal space

Diffraction from a 3D lattice

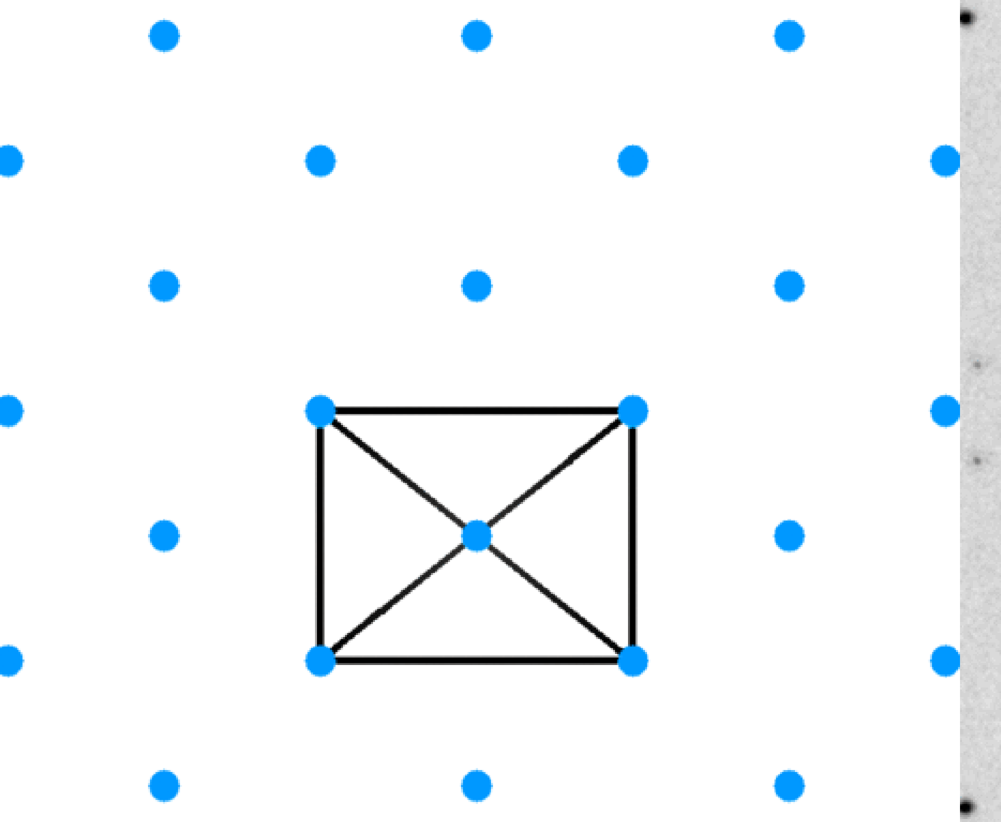


real space

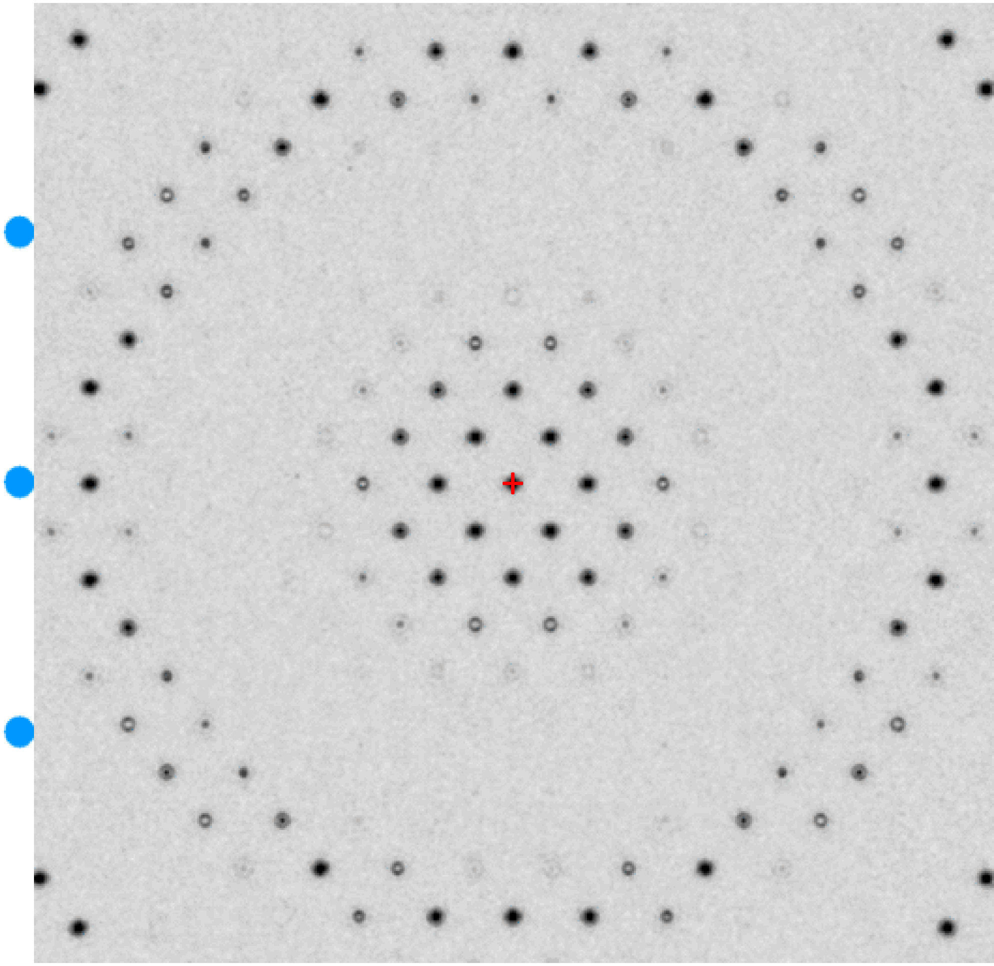


reciprocal space

Diffraction from a 3D lattice

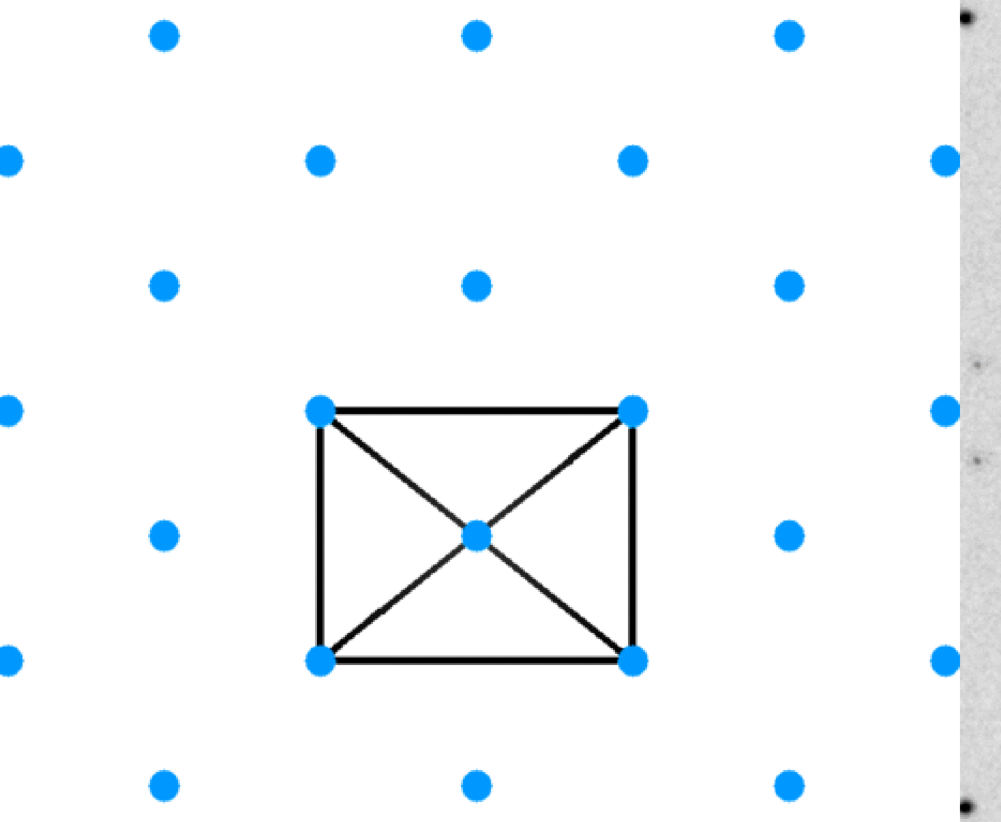


real space

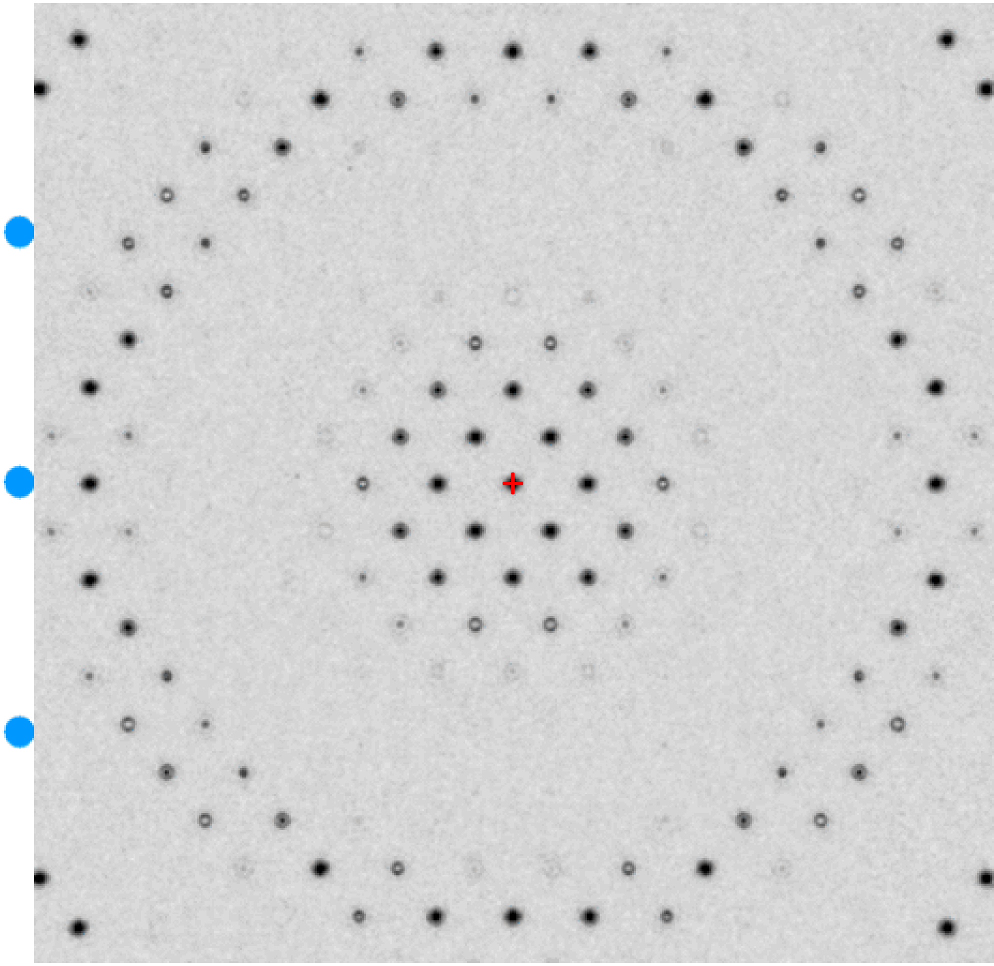


reciprocal space

Diffraction from a 3D lattice

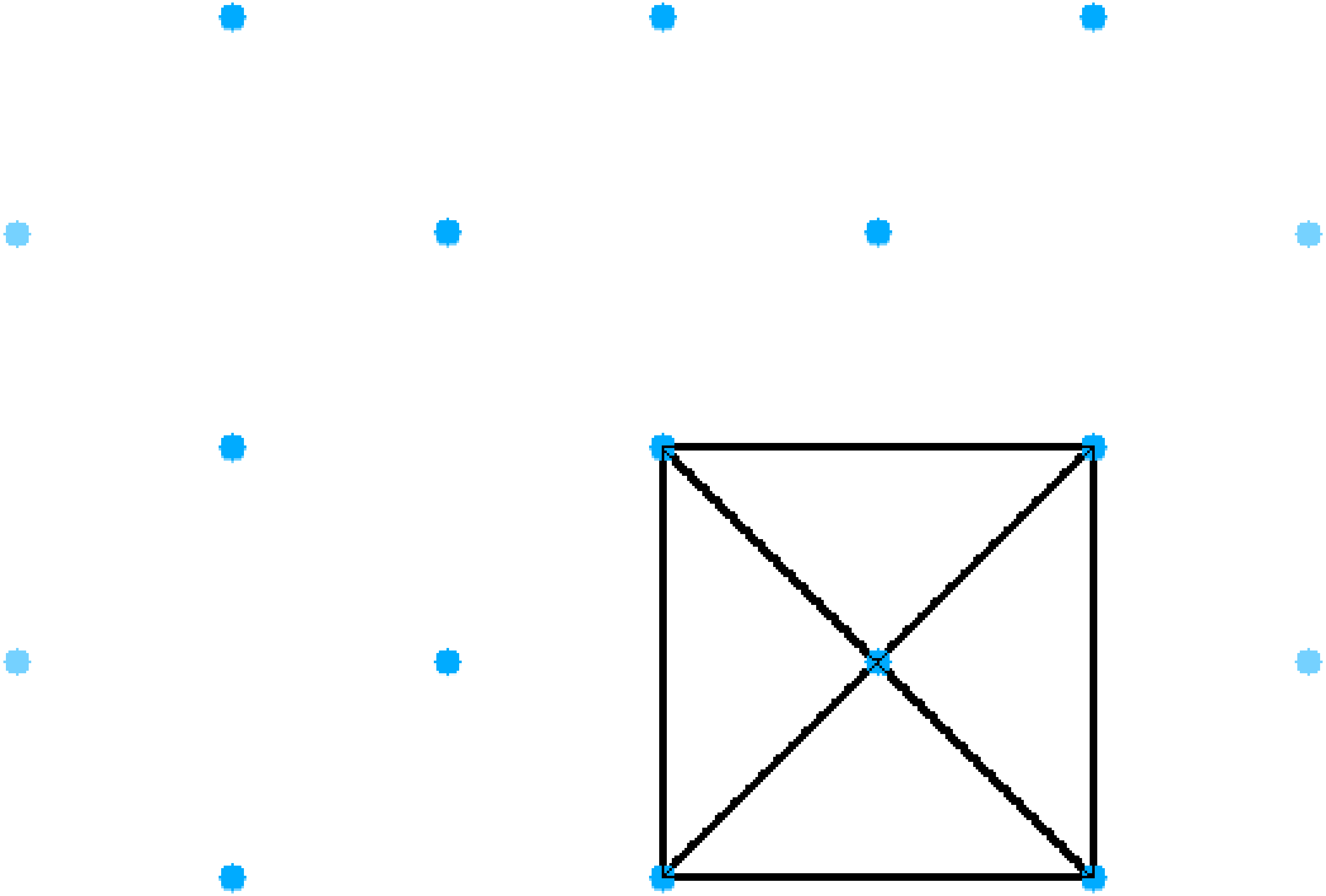


real space

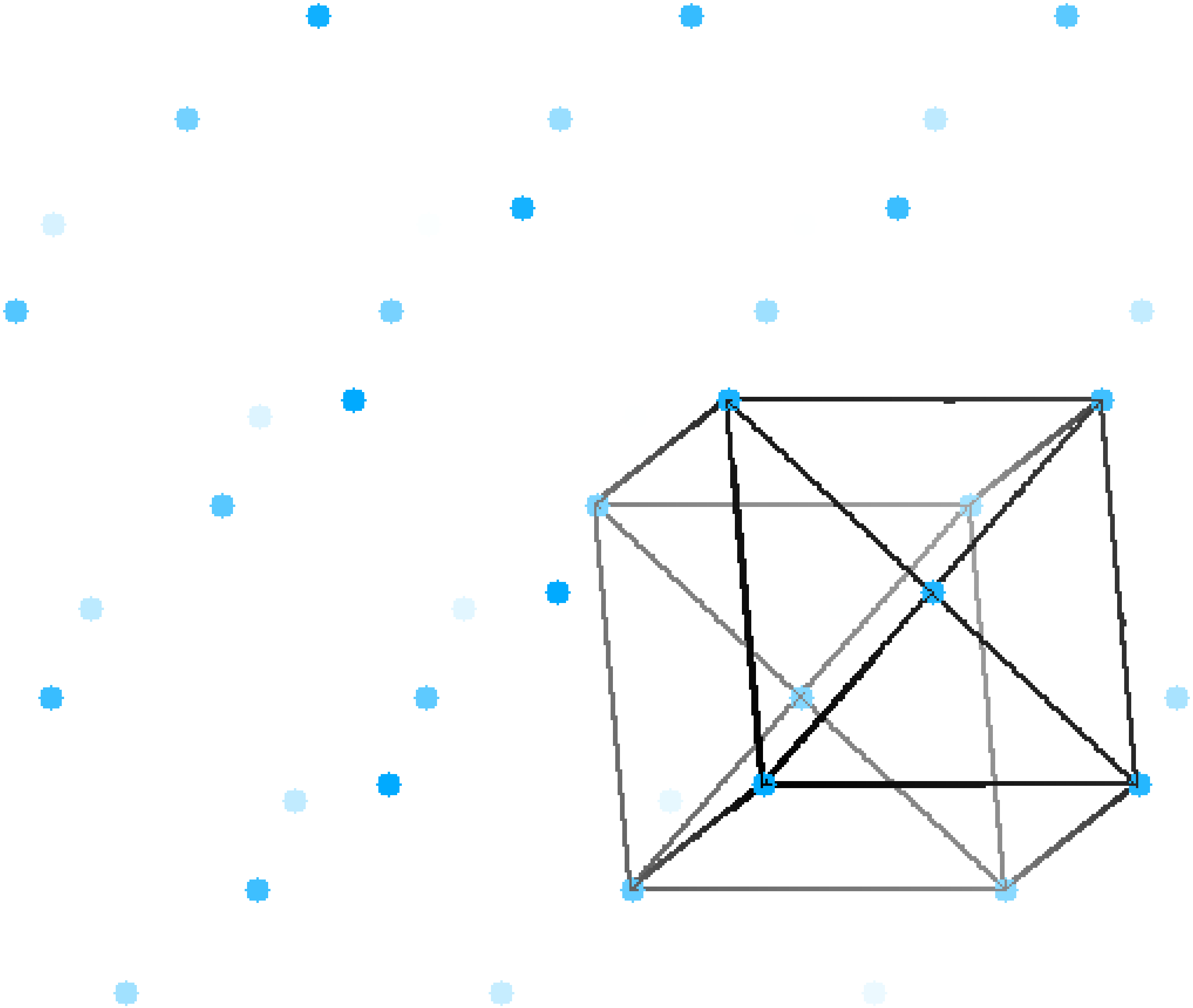


reciprocal space

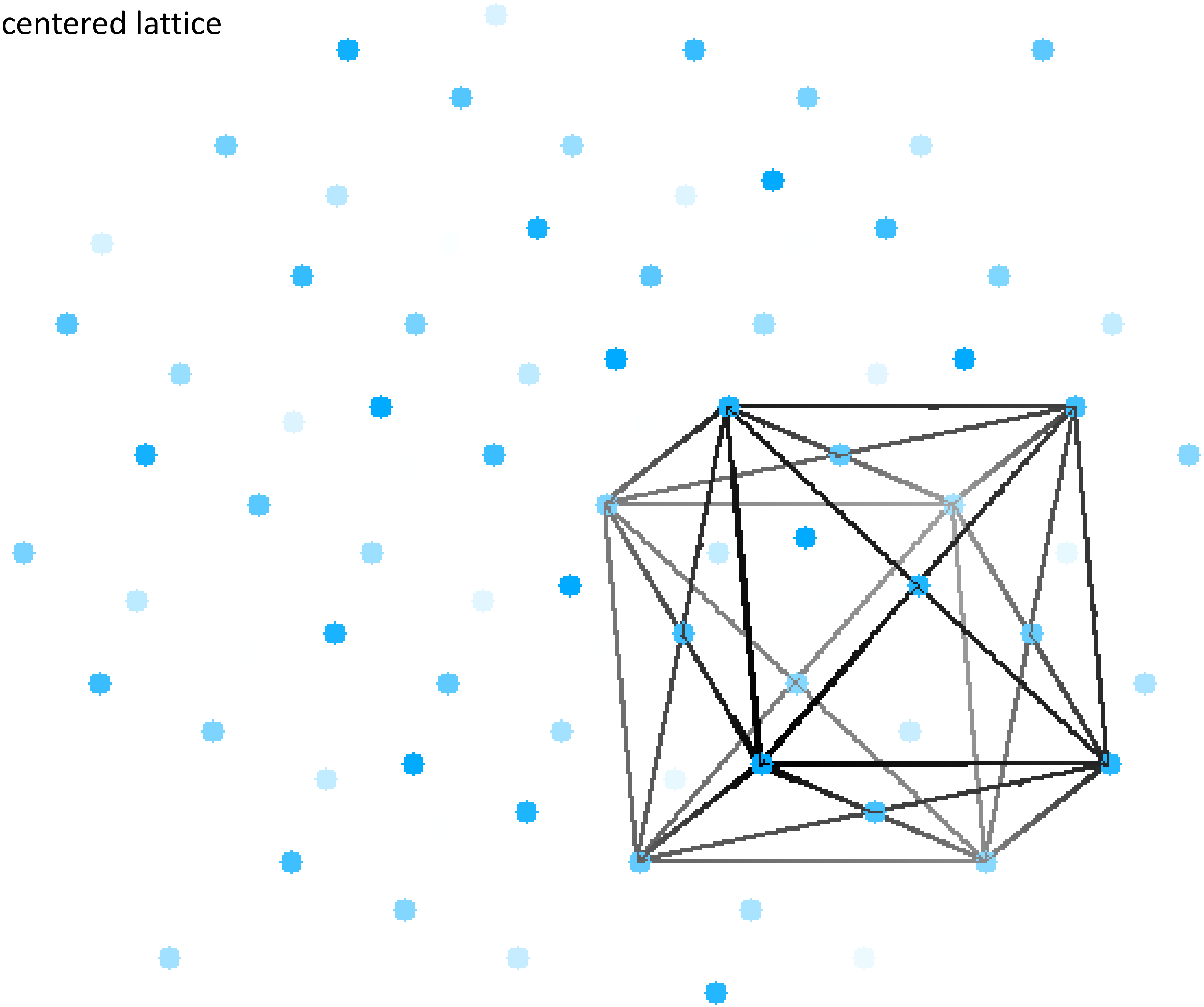
Base-centered lattice



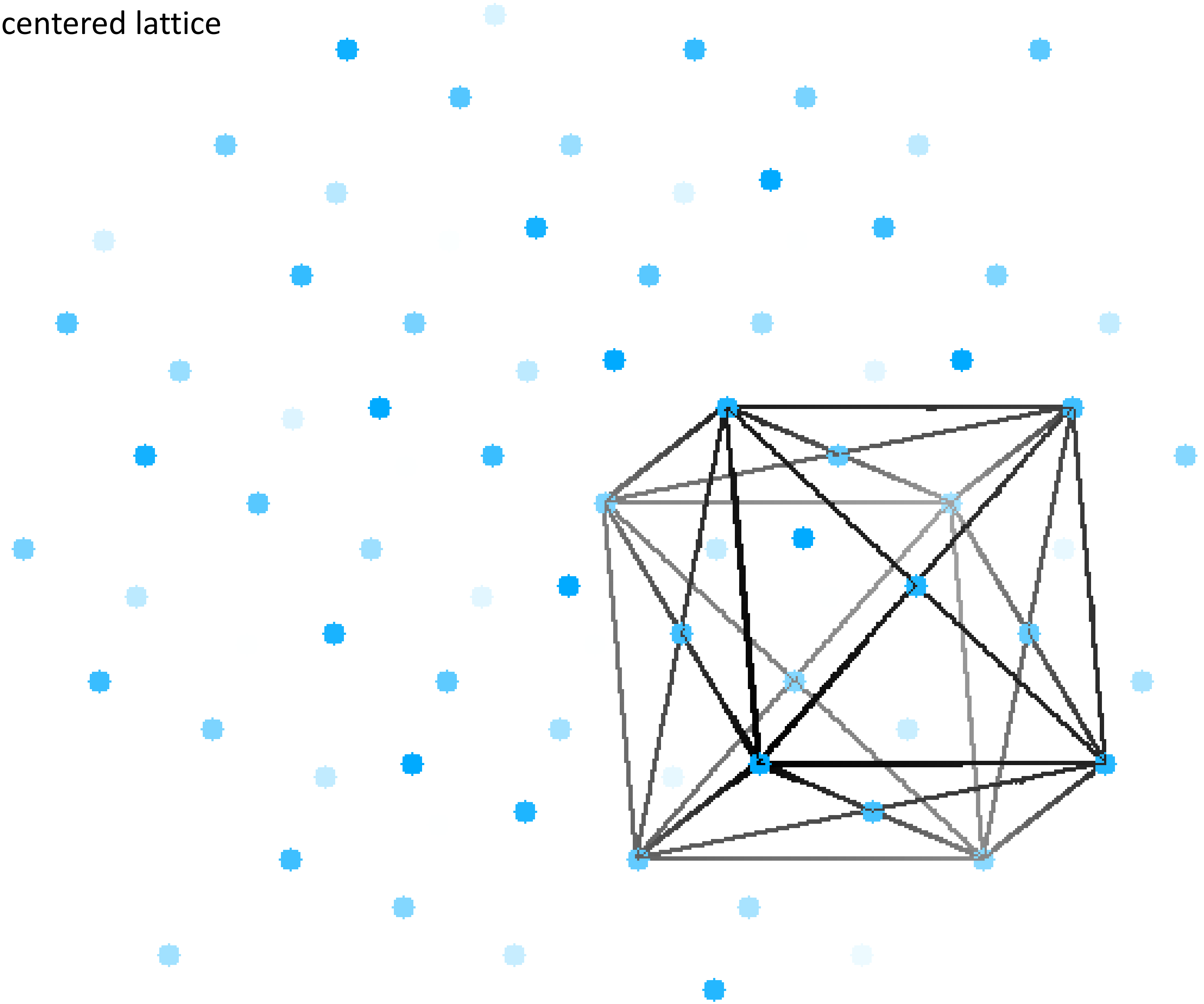
Base-centered lattice

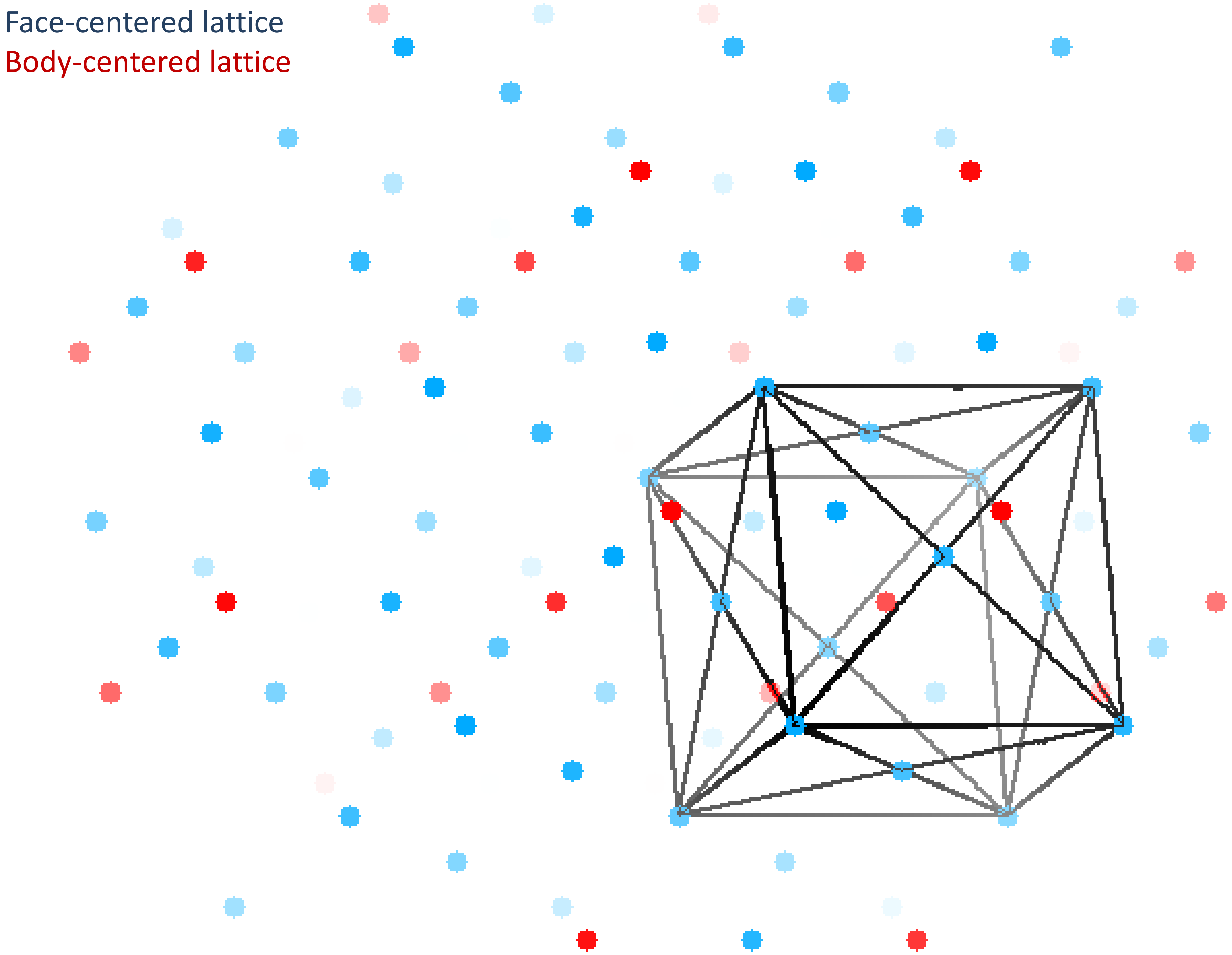


Face-centered lattice



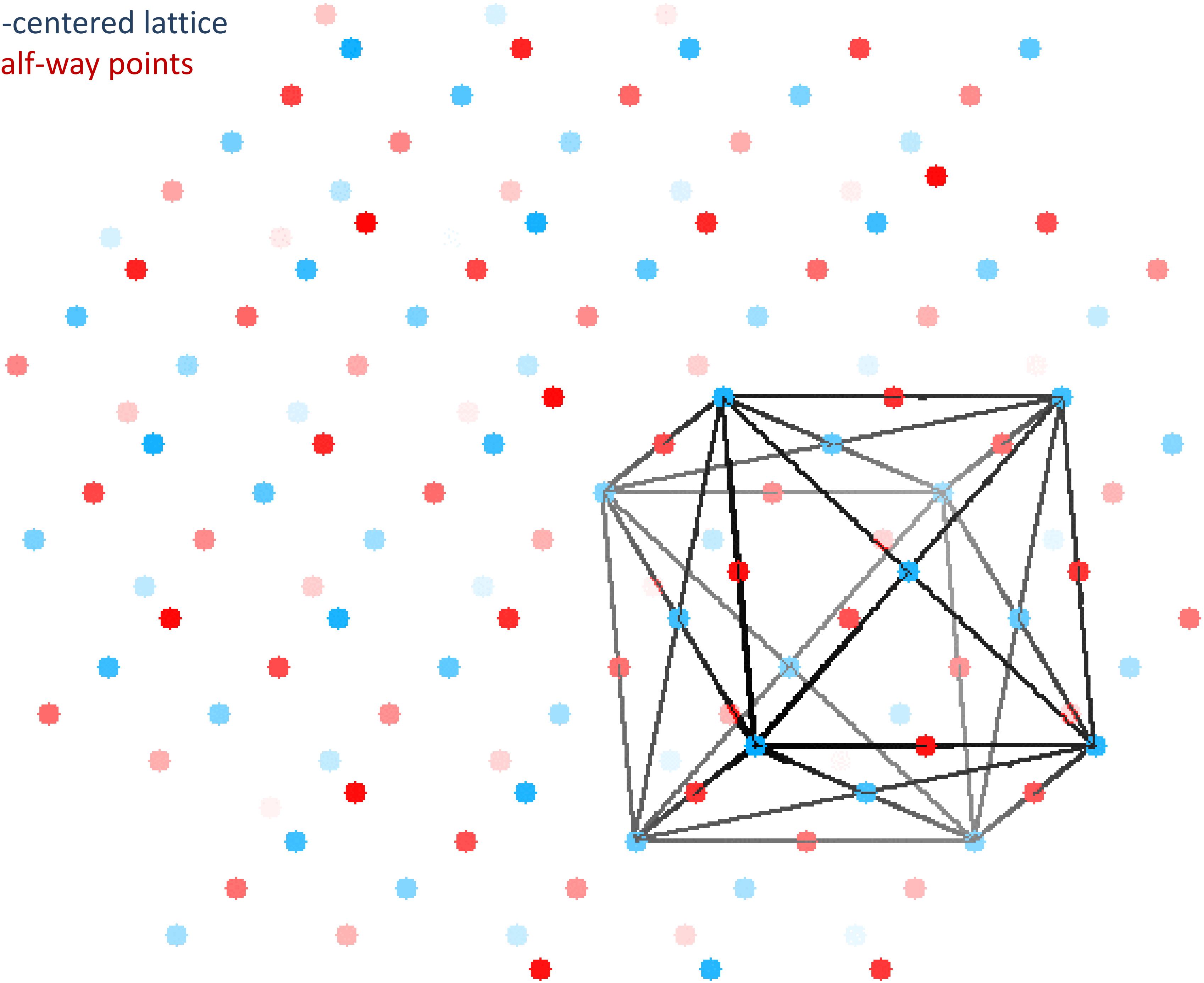
Face-centered lattice




Face-centered lattice
Body-centered lattice

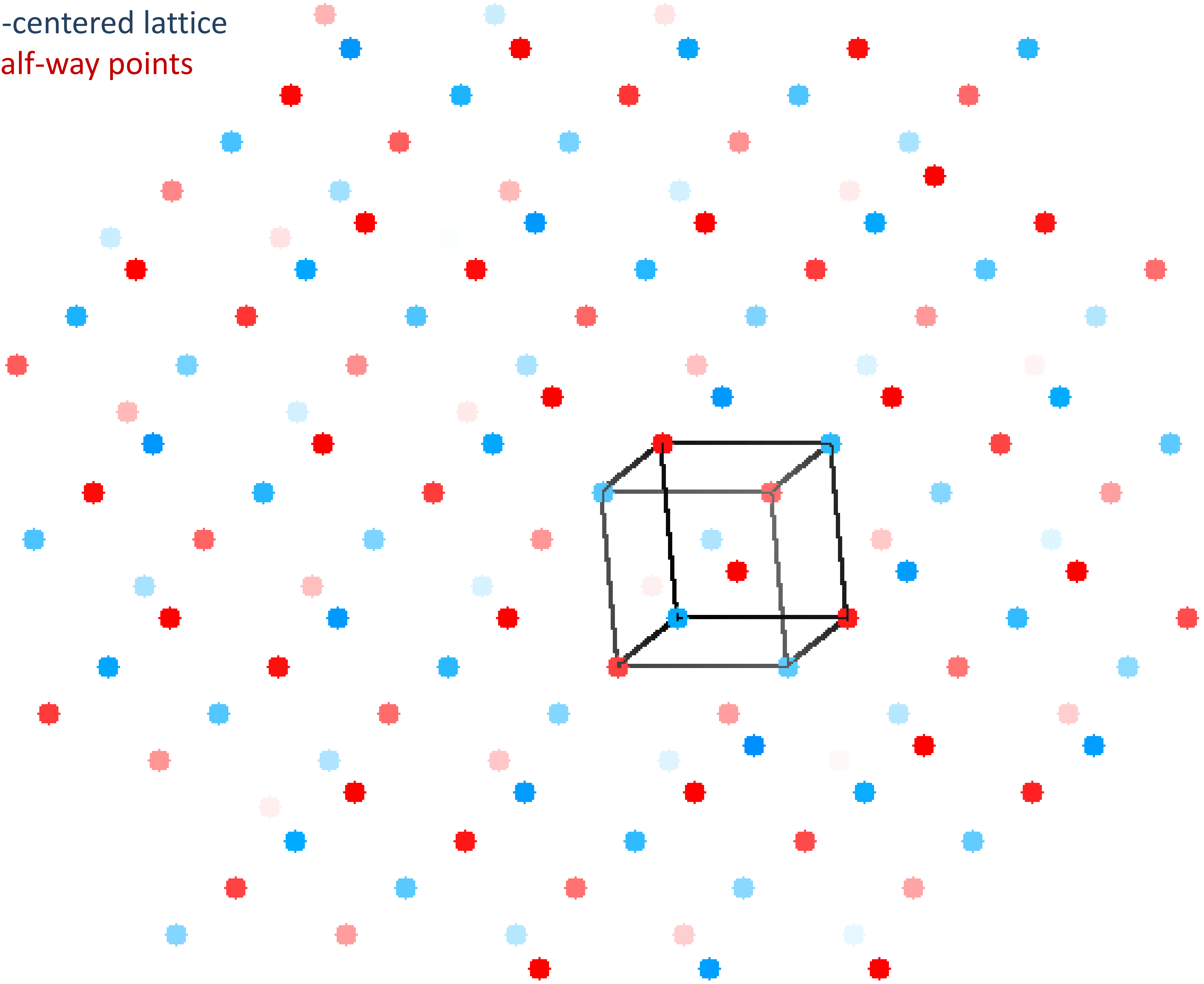
Face-centered lattice

All half-way points

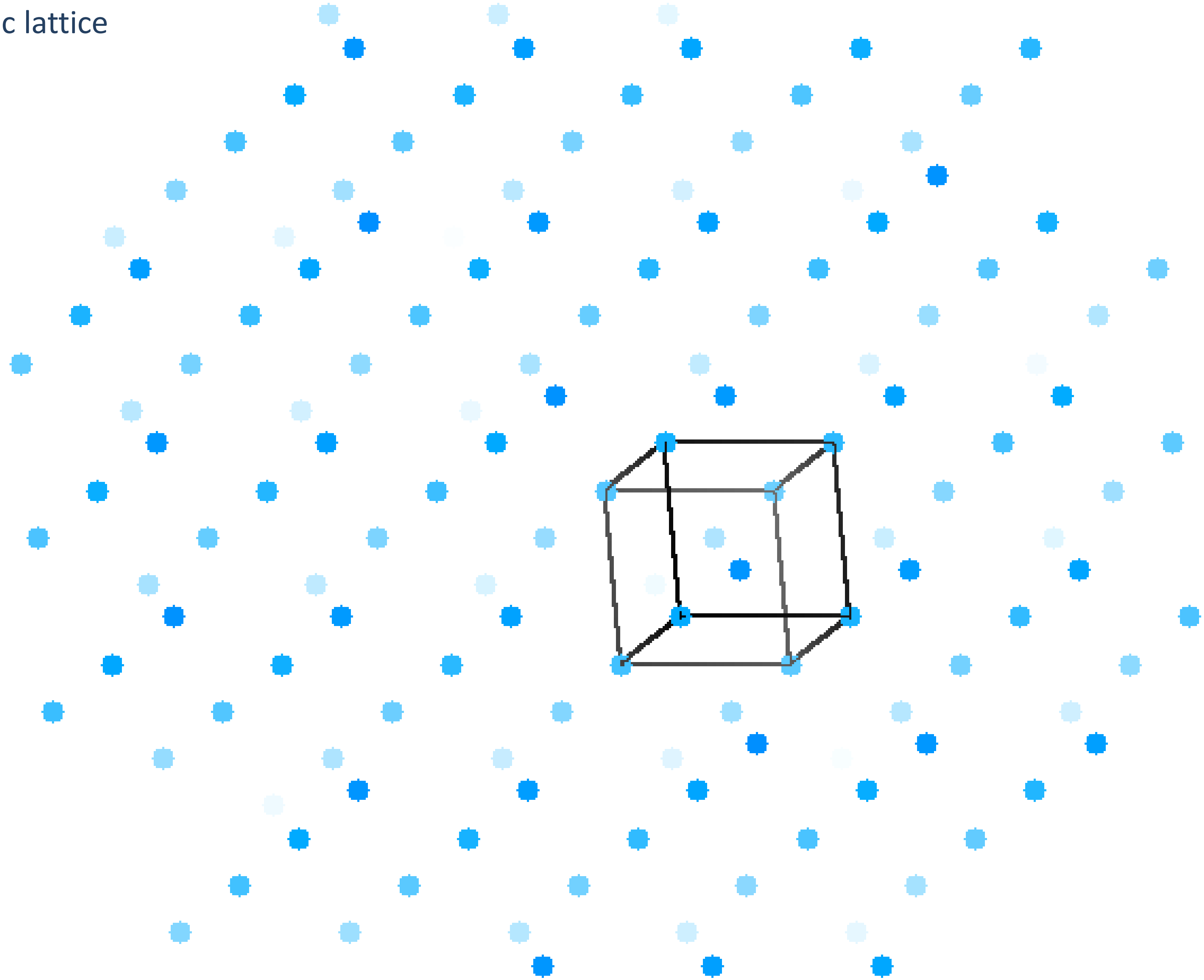


Face-centered lattice

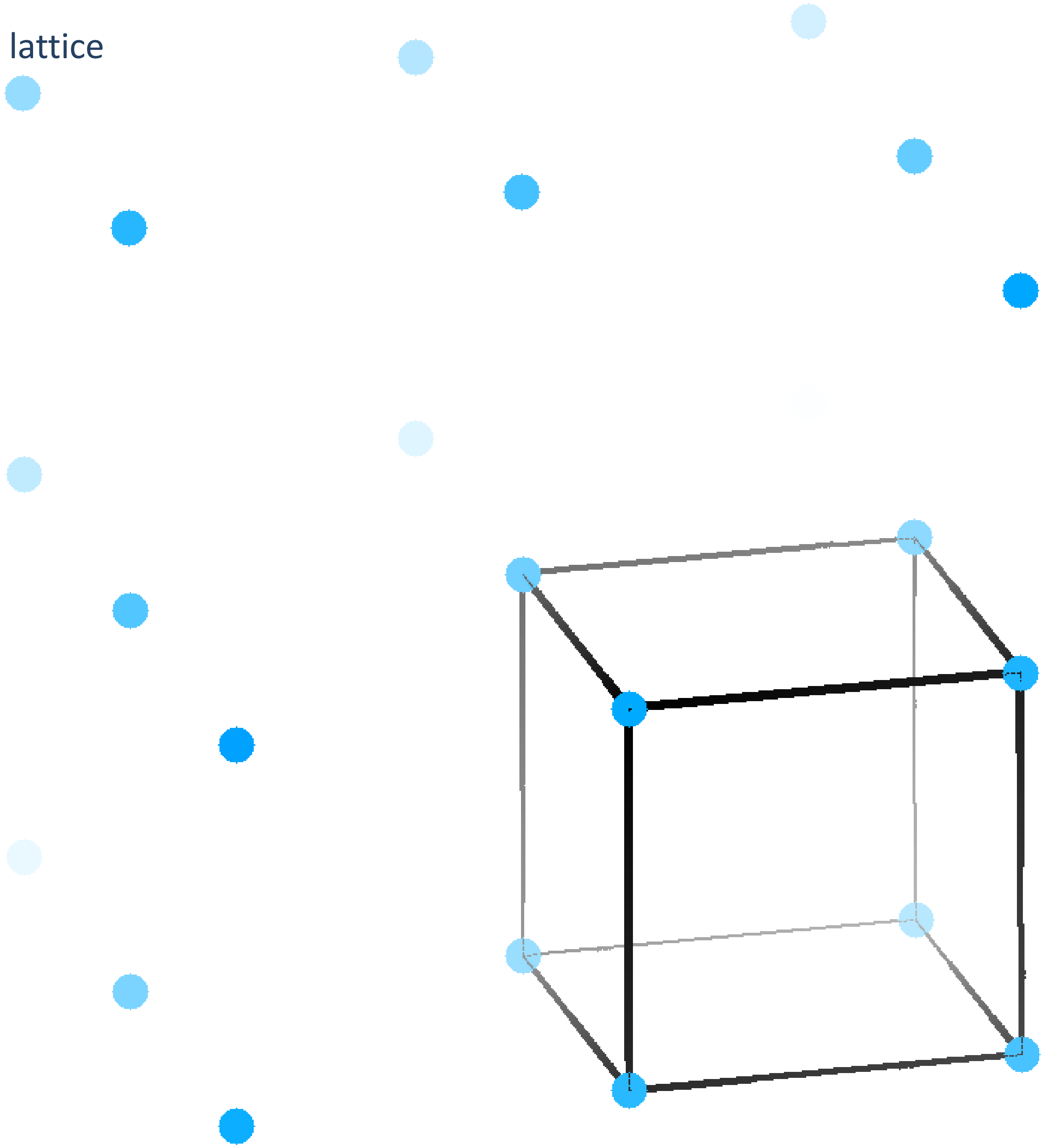
All half-way points



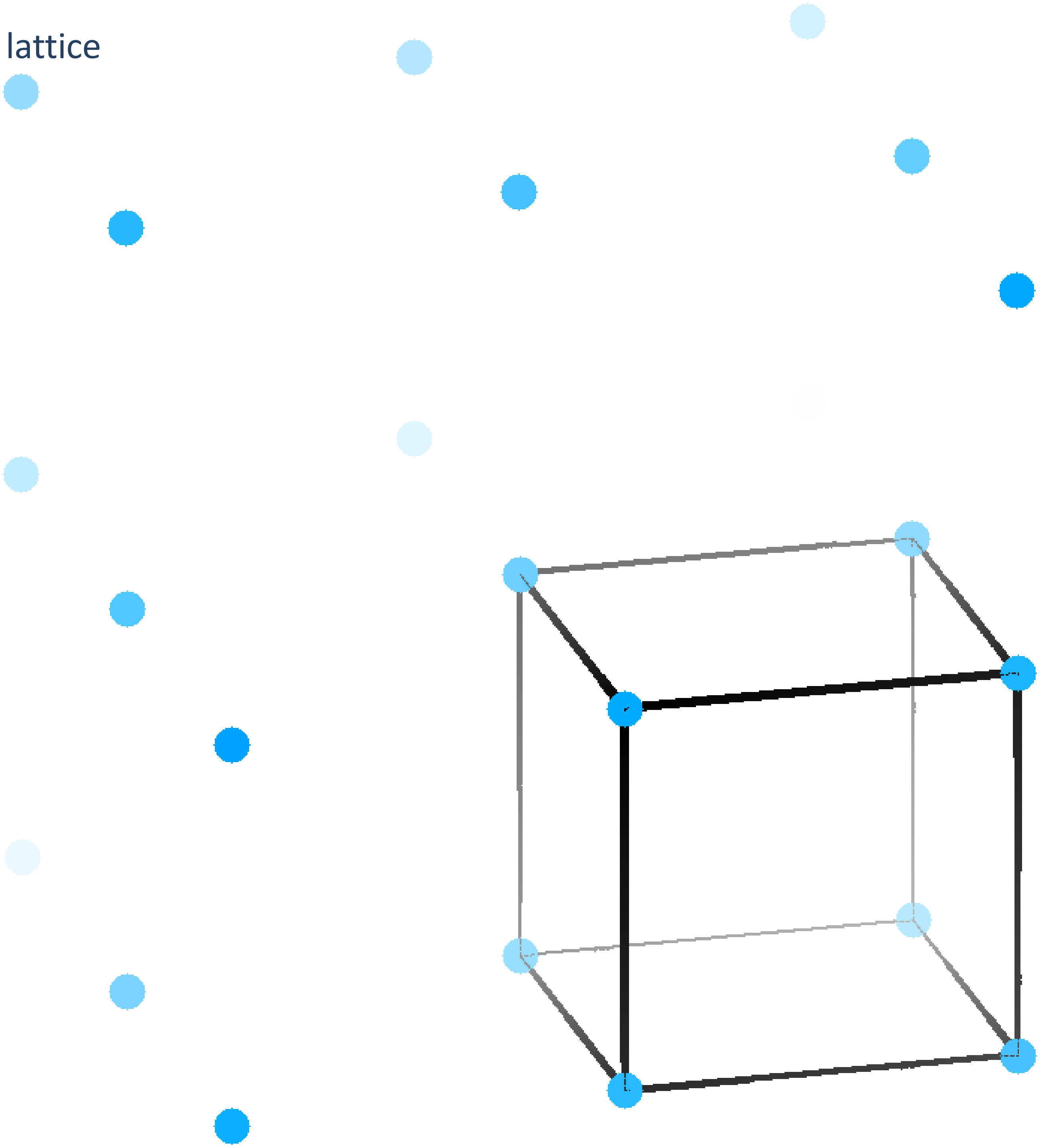
cubic lattice



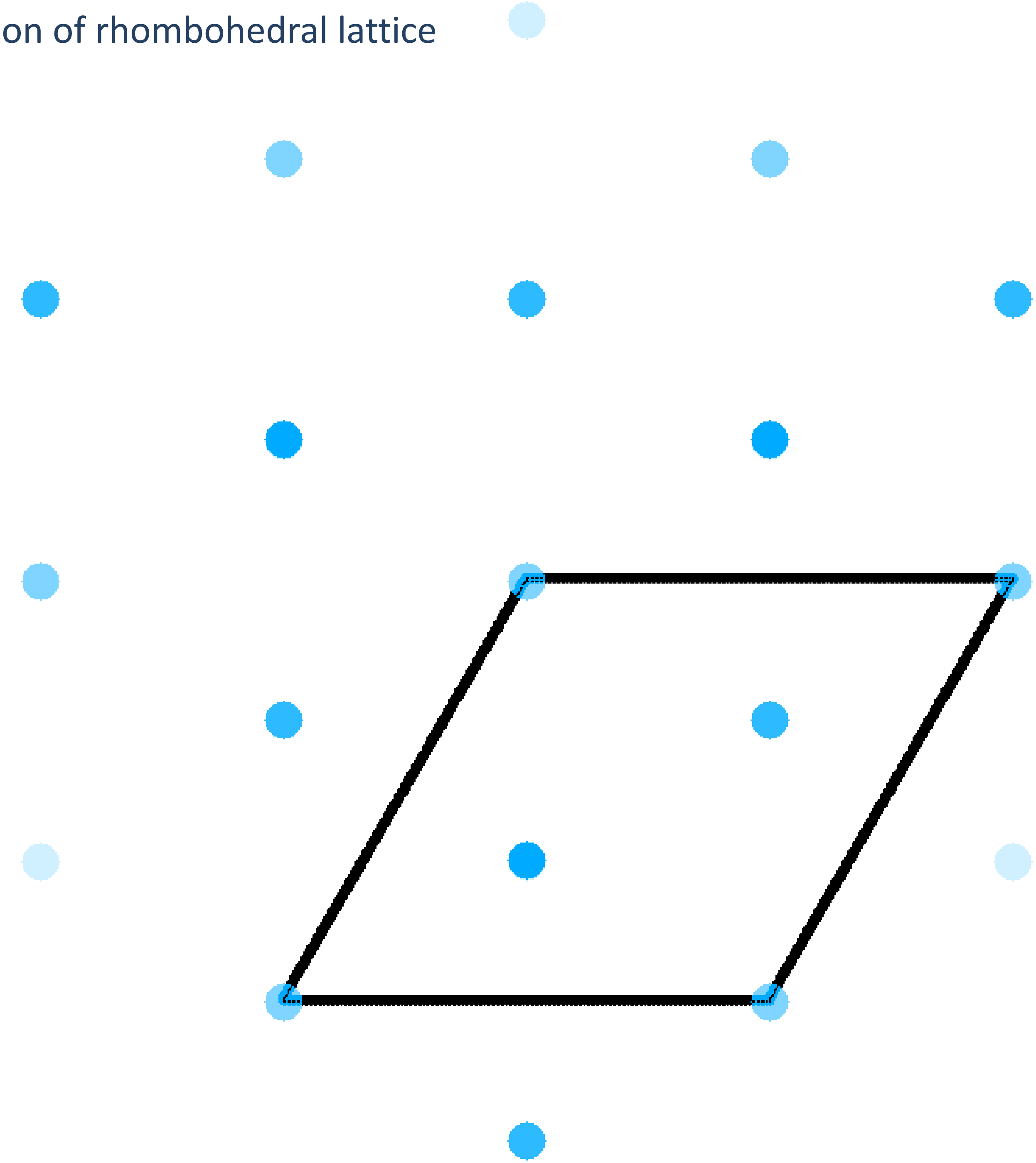
rhombohedral lattice



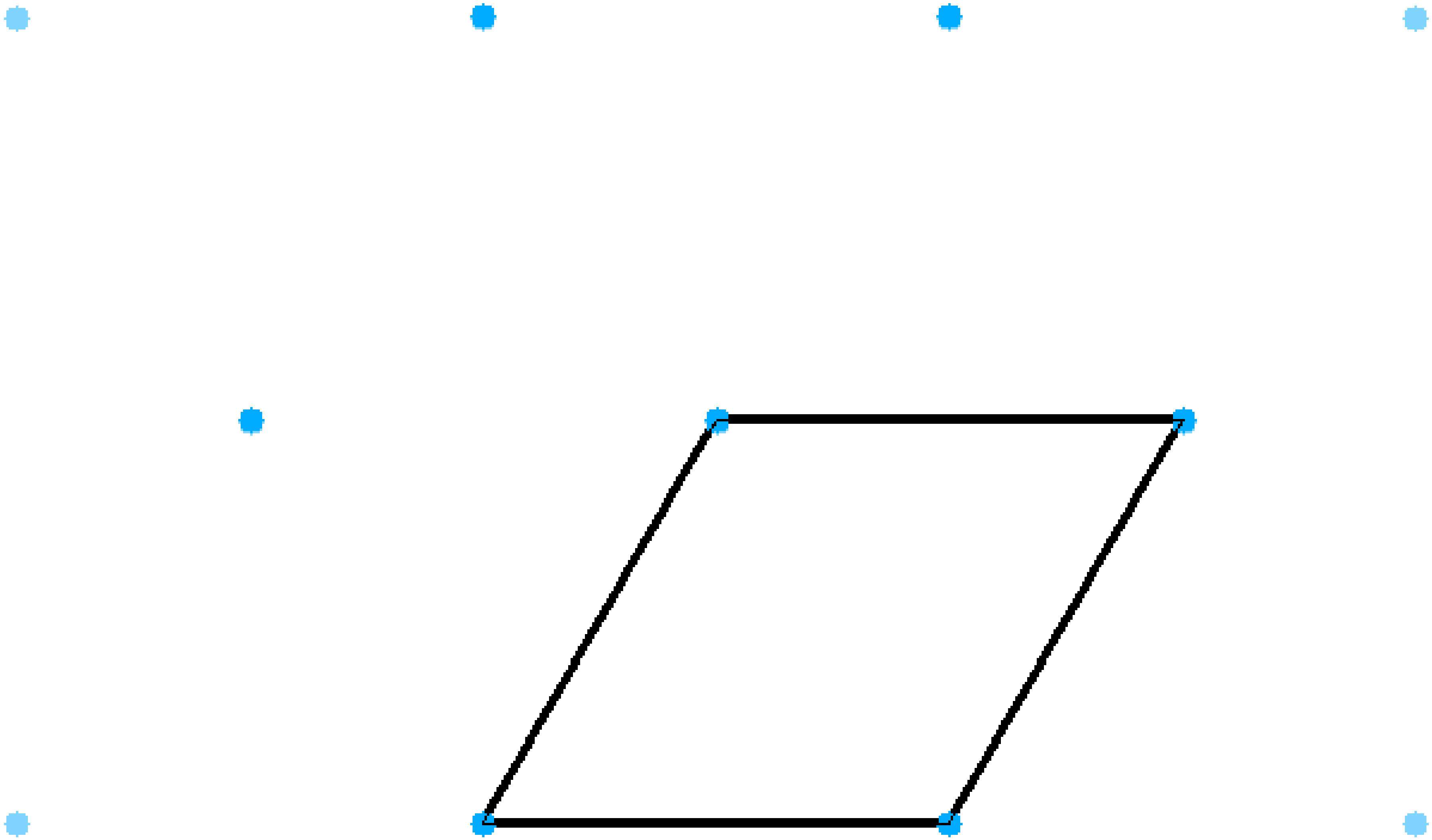
rhombohedral lattice



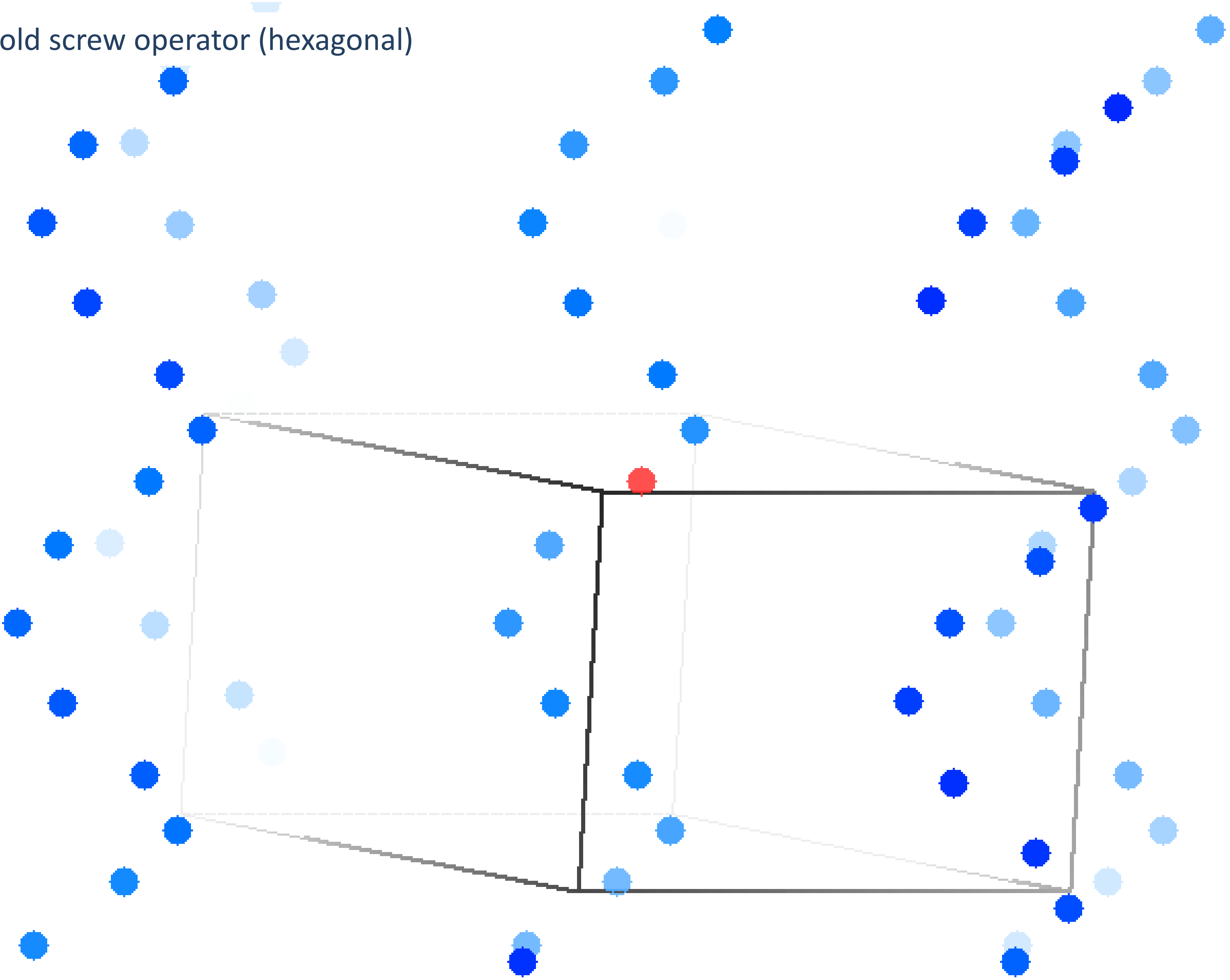
Hexagonal version of rhombohedral lattice



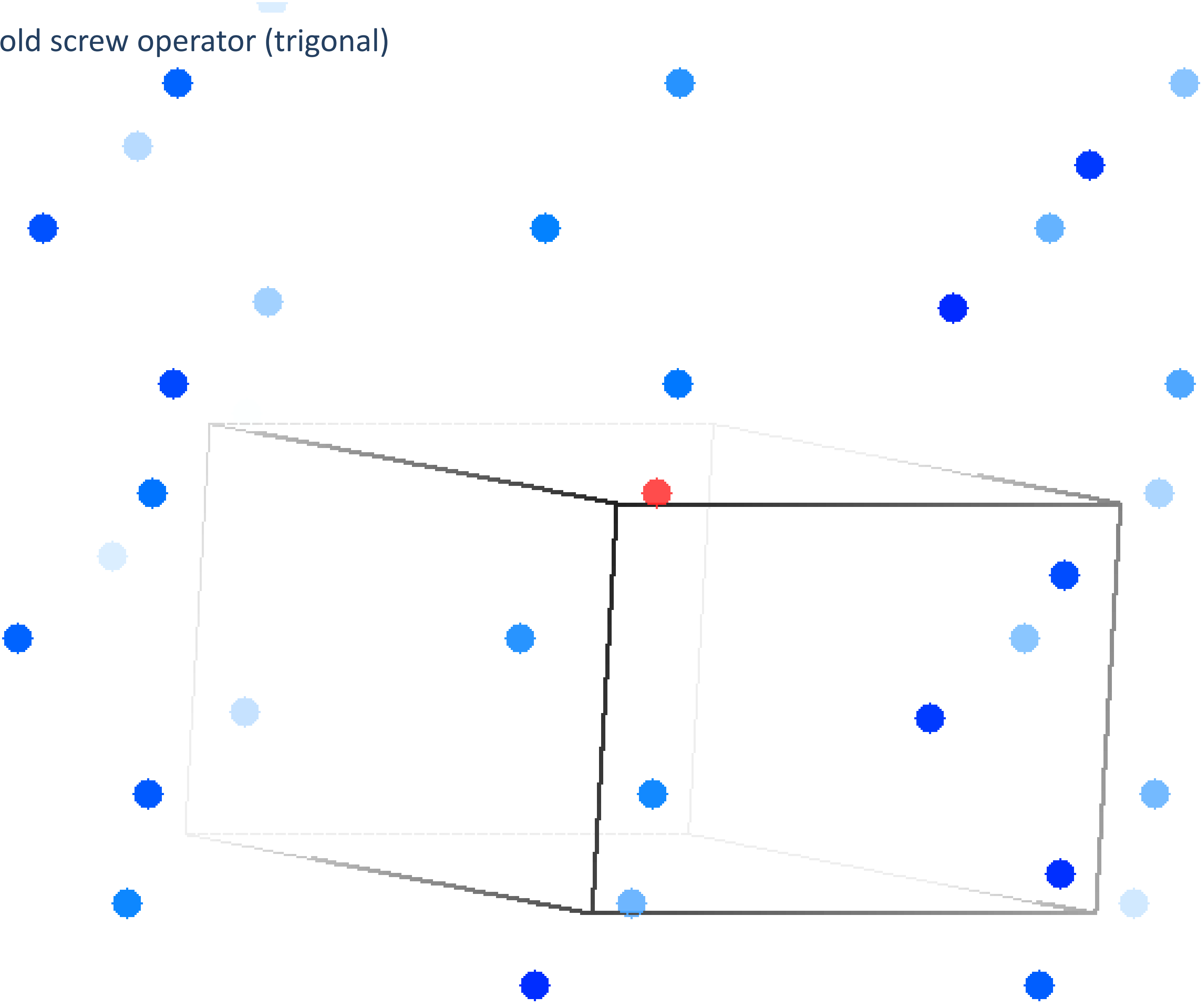
Hexagonal lattice



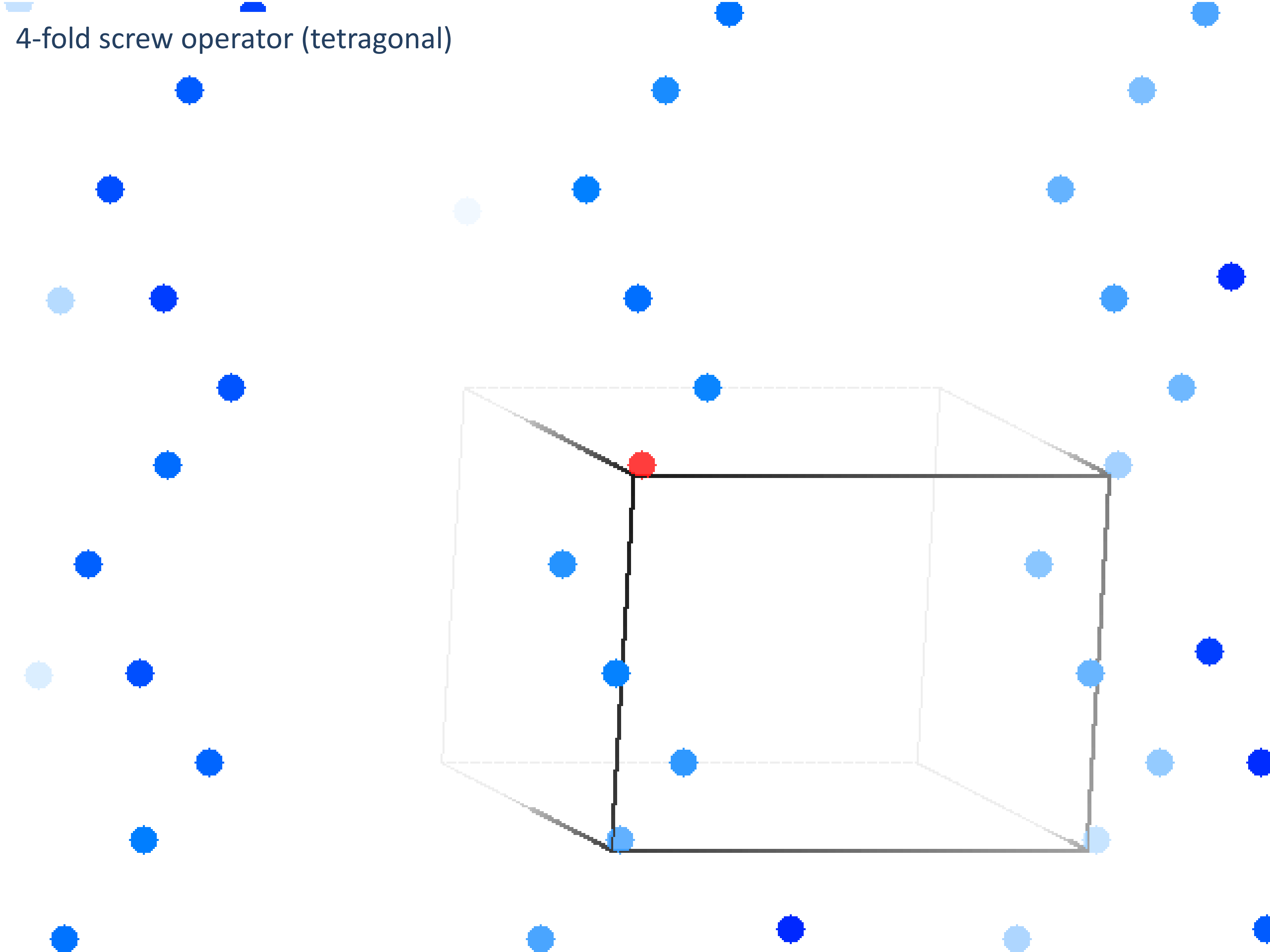
6-fold screw operator (hexagonal)



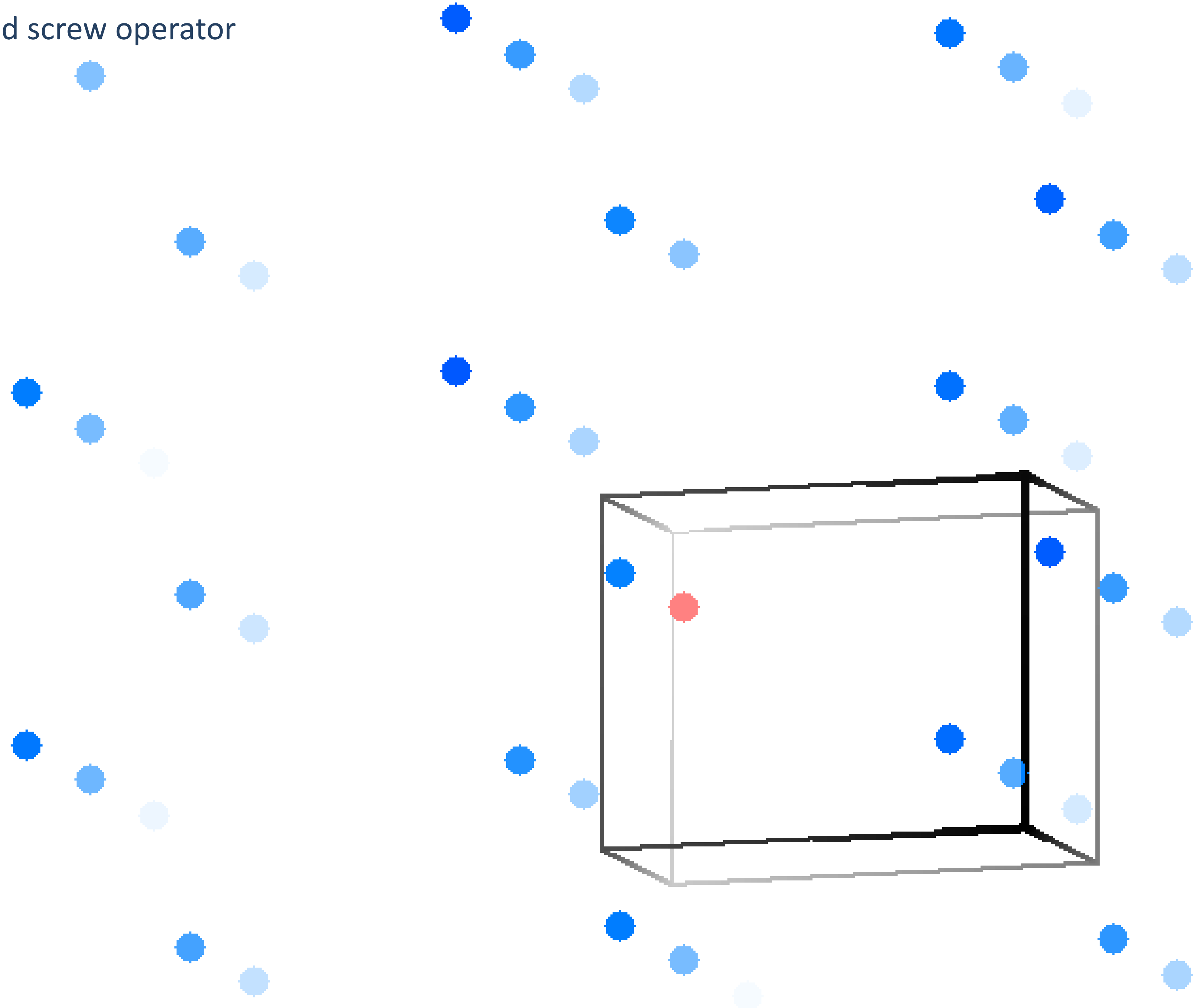
3-fold screw operator (trigonal)



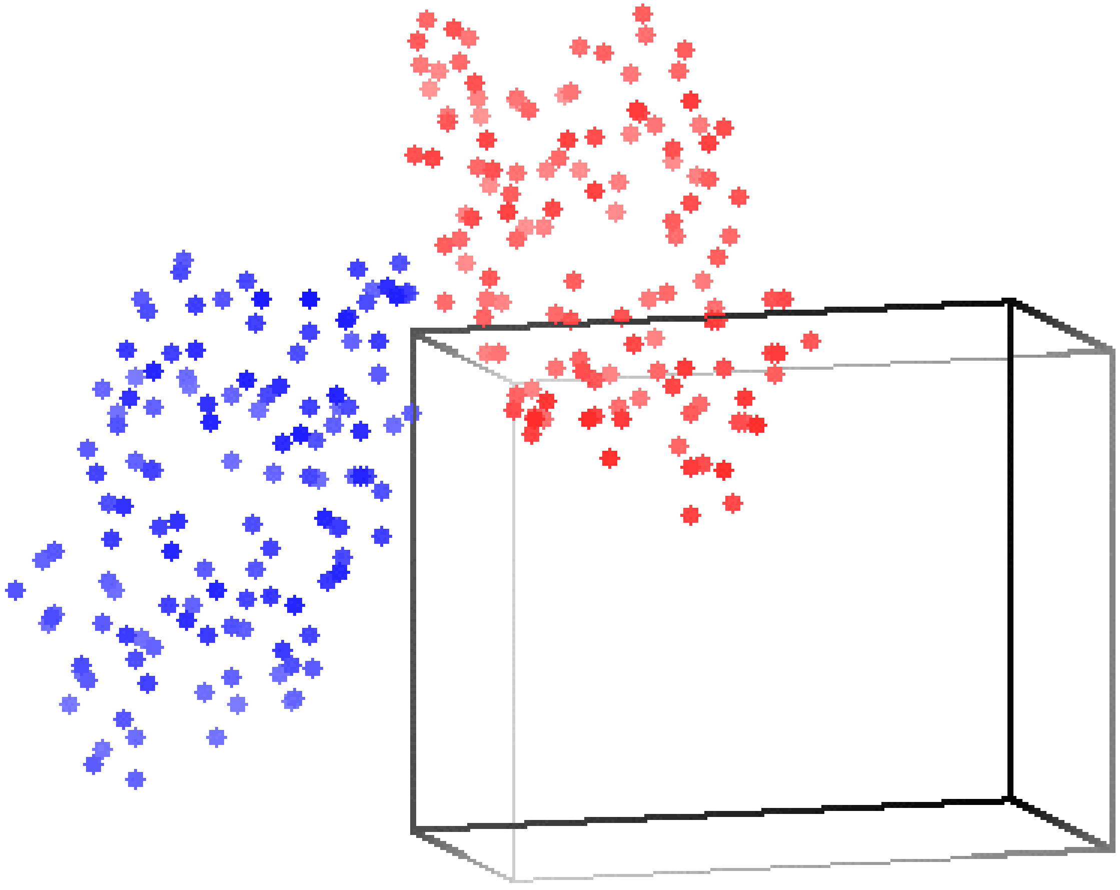
4-fold screw operator (tetragonal)



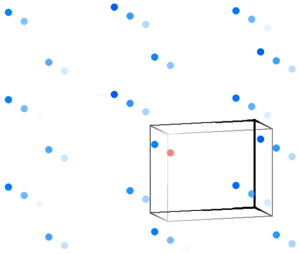
2-fold screw operator



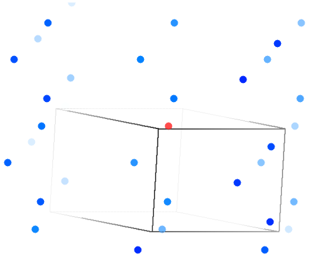
2-fold screw operator: monoclinic lysozyme



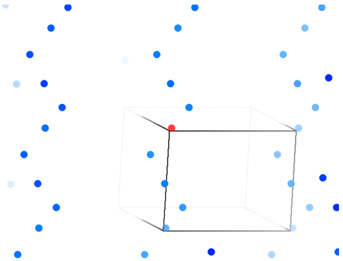
Only 4 kinds of rotation operators



2

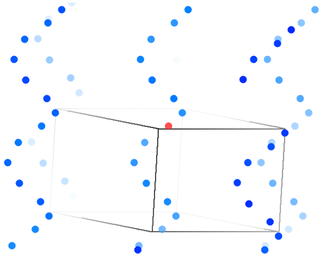


3



4

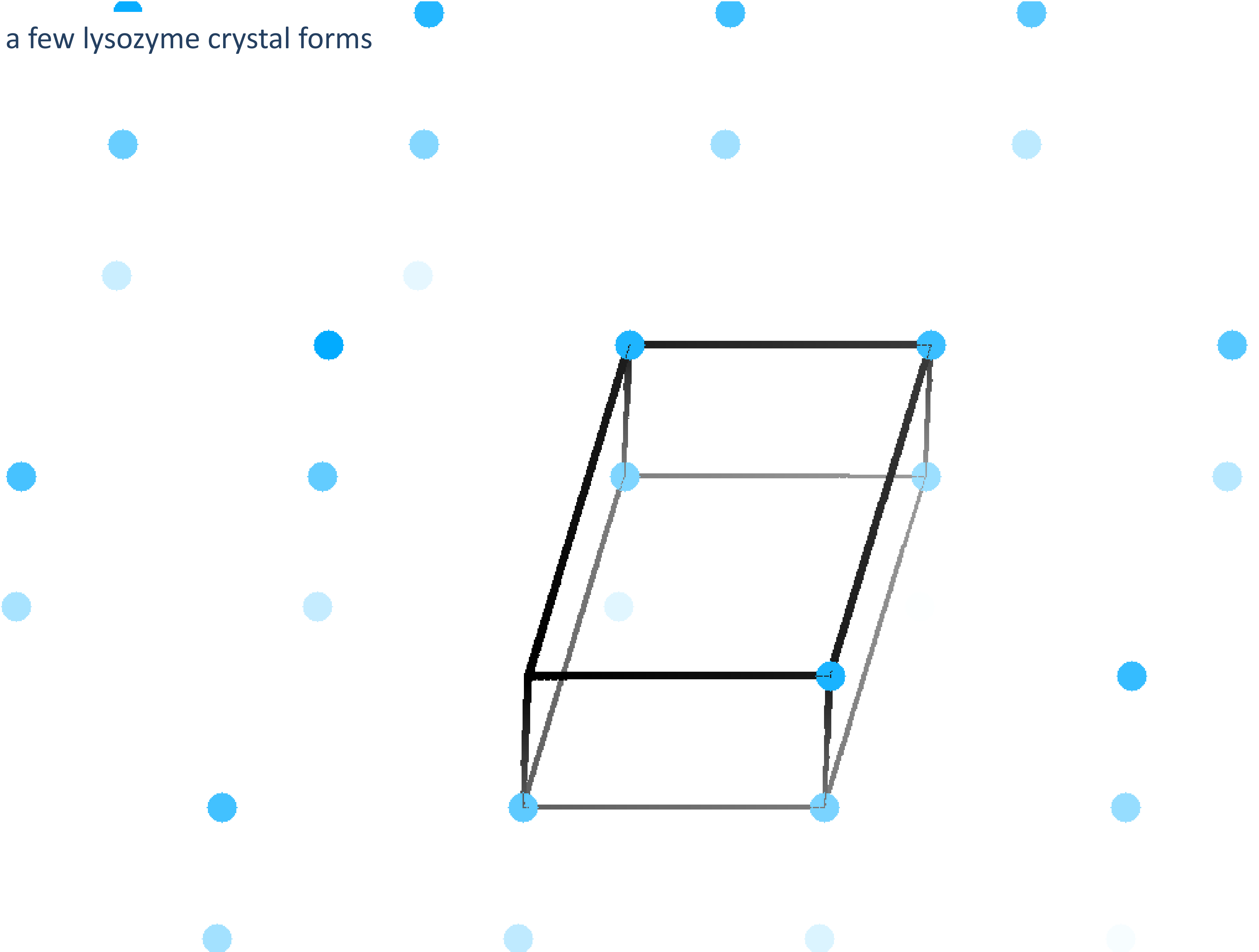
2*2



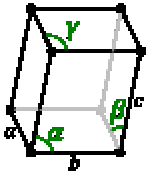
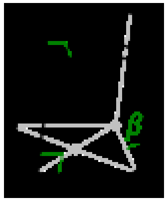
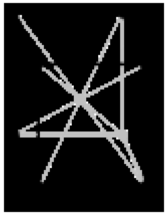
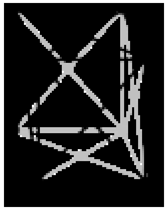
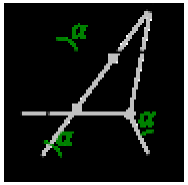
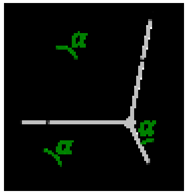
6

2*3

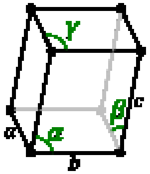
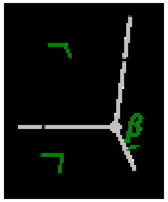
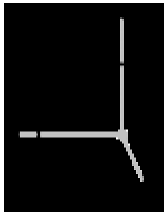
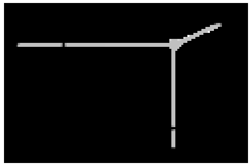
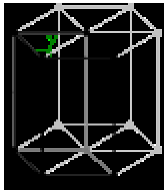
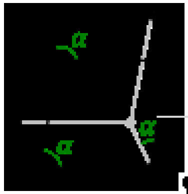
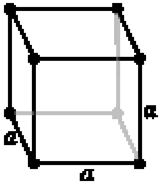
a few lysozyme crystal forms



Only 5 kinds of centering

None	Primitive		No restrictions
ABC	Base		at least two 90° angles
I	Body		All angles 90°
F	Face		All angles 90°
R	Rhombo		Also fits: 

Only 7 lattice systems

Triclinic		No restrictions
Monoclinic		2x 90° angles
Orthorhombic		All angles 90°
Tetragonal		All 90°; square top
Hexagonal		90° 90° 120°; a = b
Rhombohedral		angles same, a=b=c
Cubic		All 90°; a = b = c

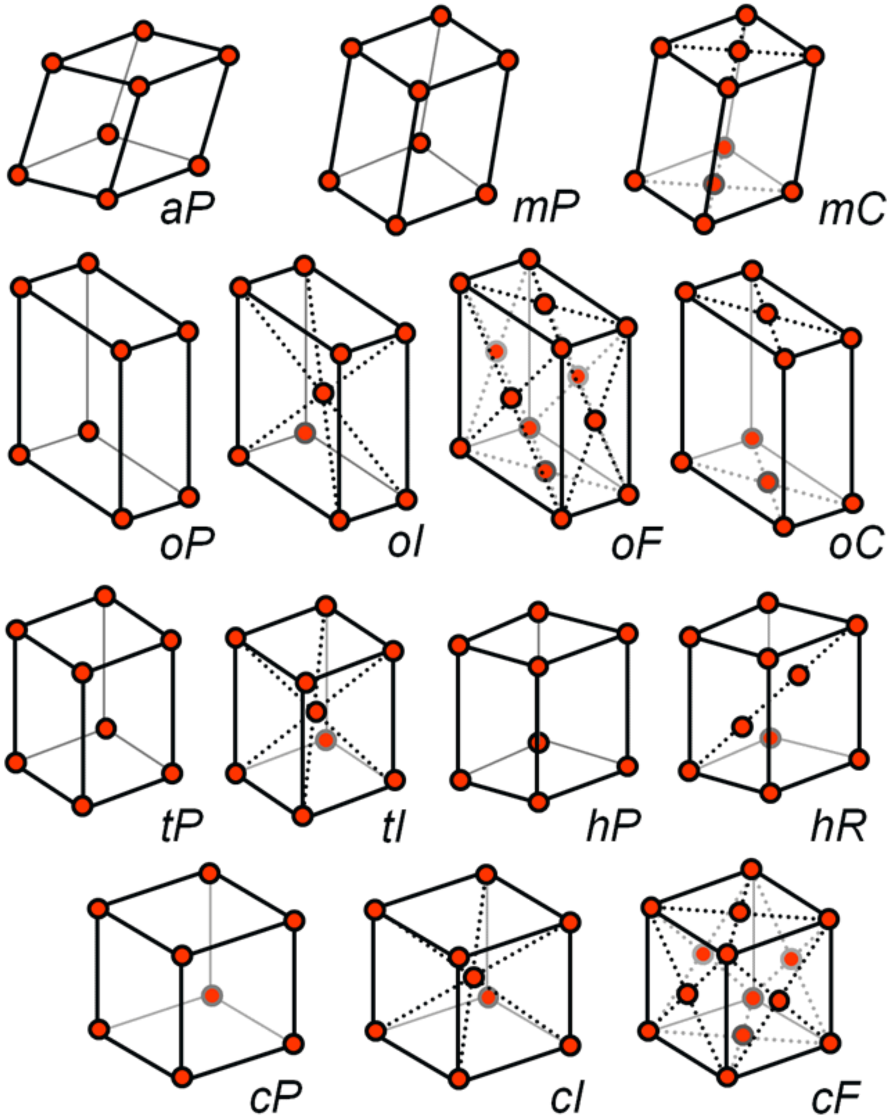
Only 14 lattice types in 3D

Lattice system

a	asymmetric
m	monoclinic
o	orthorhombic
t	tetragonal
h	hexagonal
c	cubic

Centering

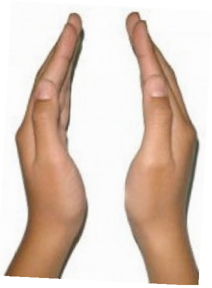
P	primitive (no centering)
C	base (A,B not used)
I	body (extra in middle)
F	face (extra on each)
R	rhombohedral (two diagonal extras)

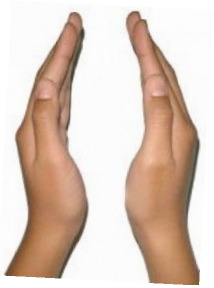


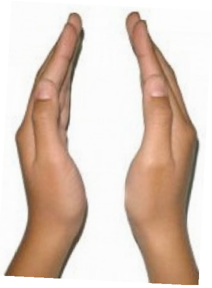
Bravais (1848)

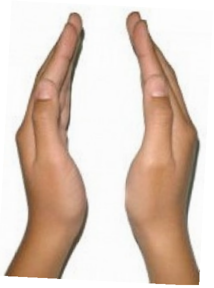
www.ruppweb.com

Enantiomorphic operators

3_1  3_2

6_1  6_5

4_1  4_3

6_2  6_4

Systematic absences

How to interpret:

Every 2nd

2_1 4_2 6_3

Every 3rd

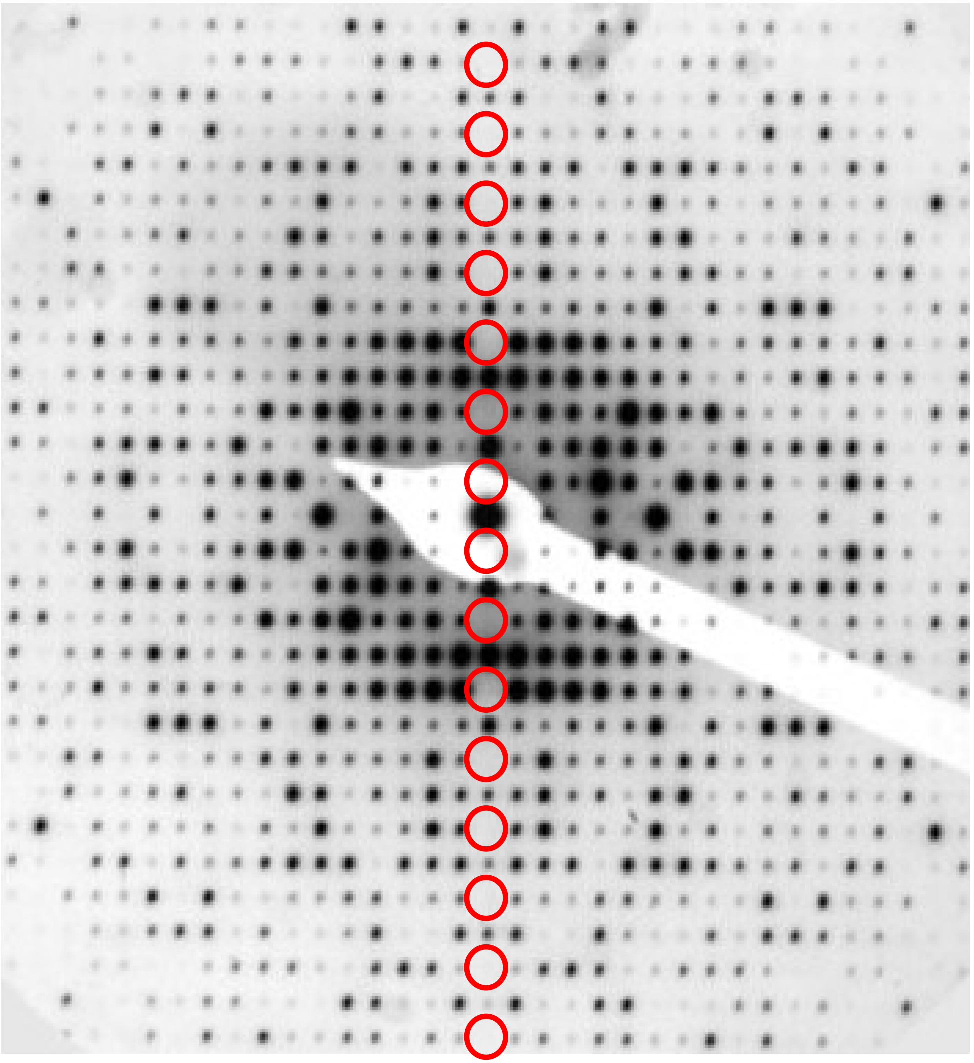
3_1 3_2

Every 4th

4_1 4_3

Every 6th

6_1 6_5



Hermann-Mauguin symbol

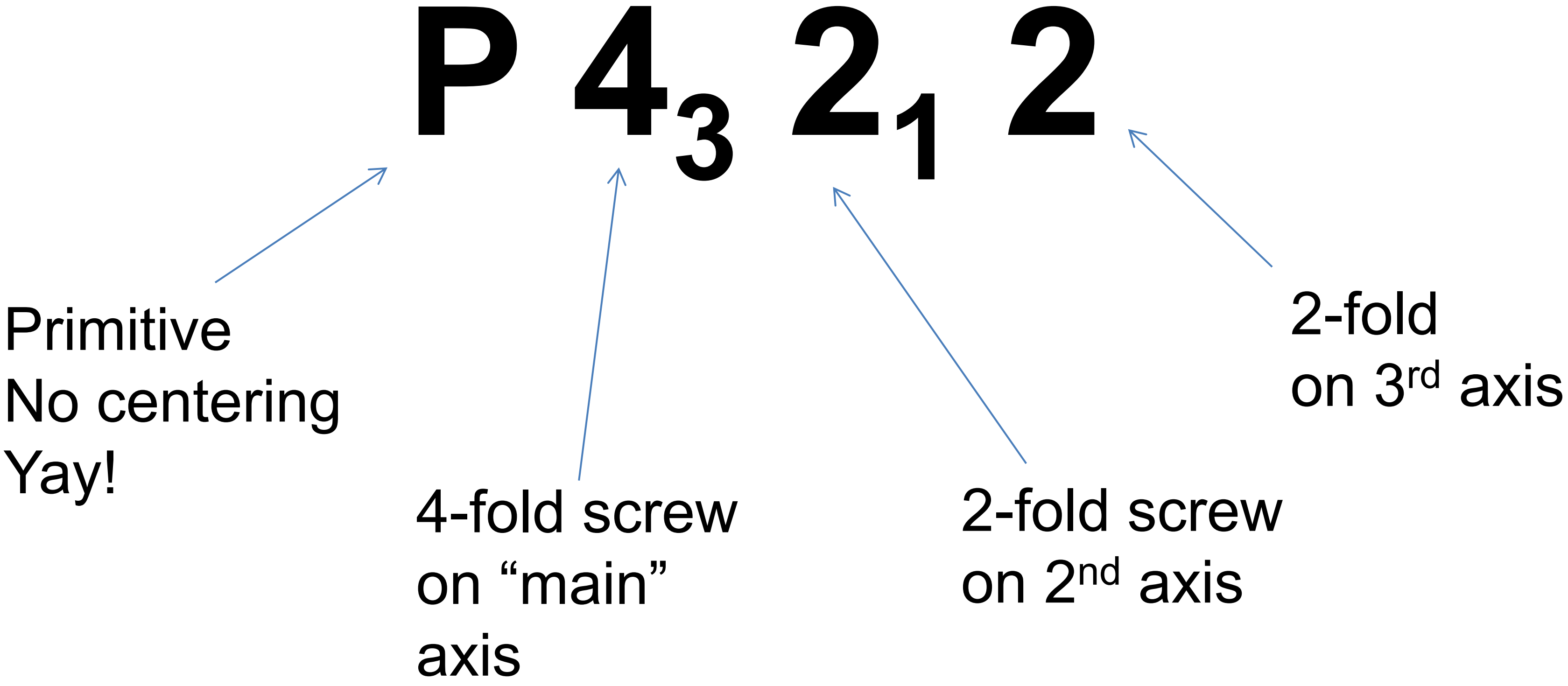
P 1

Primitive
No centering

And that's all

Also tells you: Bravais lattice = aP ; point group: 1
Unit cell can be anything

Hermann-Mauguin symbol



$$P\ 4_3\ 2_1\ 2$$

Primitive
No centering
Yay!

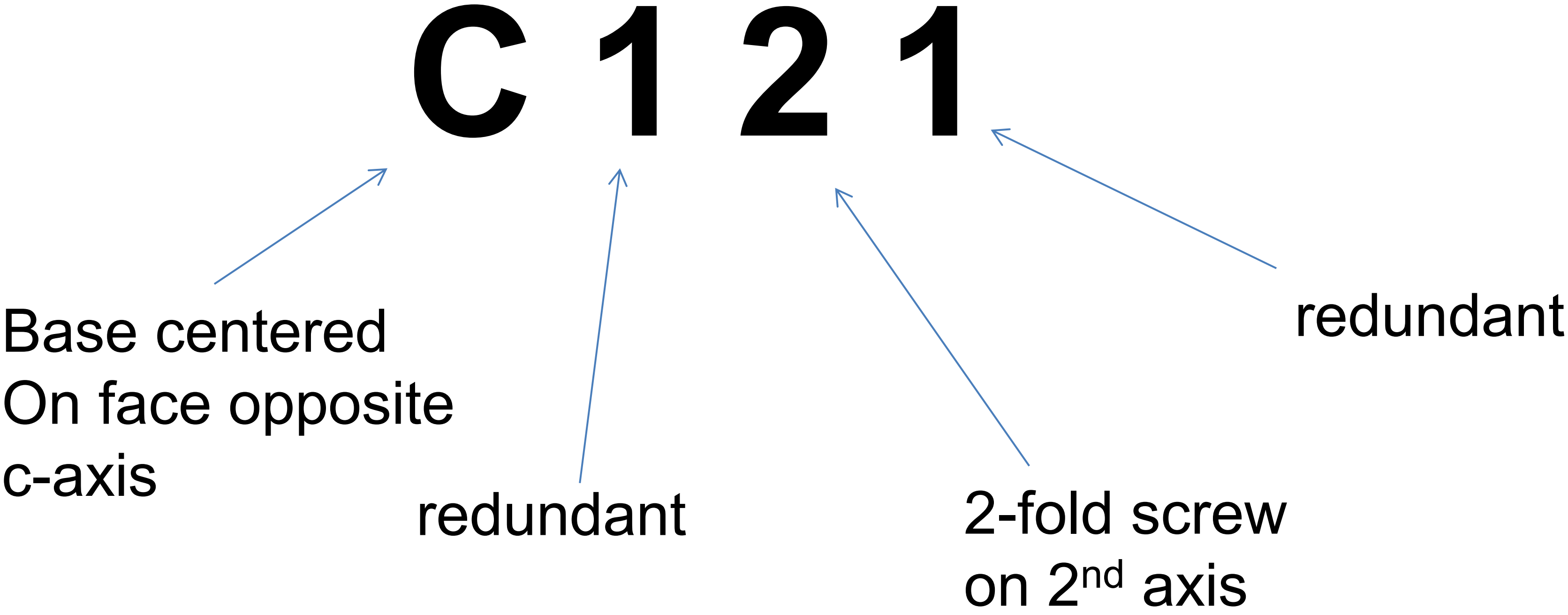
4-fold screw
on “main”
axis

2-fold screw
on 2nd axis

2-fold
on 3rd axis

Also tells you: Bravais lattice = tP ; point group: 422
tetragonal system a == b and a & b are equivalent

Hermann-Mauguin symbol



C 1 2 1

Base centered
On face opposite
c-axis

redundant

2-fold screw
on 2nd axis

redundant

Also tells you: Bravais lattice = mC ; point group: 2
Angles α & γ = 90°, β can be anything

Hermann-Mauguin symbol

C2

Base centered
On face opposite
c-axis

2-fold screw
on 2nd axis

Also tells you: Bravais lattice = mC ; point group: 2
Angles α & γ = 90°, β can be anything

Hermann-Mauguin symbol

F23

Face centered

2-fold

3-fold after
2-fold means
Cubic
(3-fold diagonal)

Also tells you: Bravais lattice = cF ; point group: 23
All edges same legnth, all angles α β γ = 90°

What is a “point group”?

Essentially:
A point group is a space group without all the subscripts

Pitfalls: the “screwies”

$P2_12_12$

$P2_122_1$

$P22_12_1$

$P222_1$

$P22_12$

$P2_122$

Pitfalls: sub-space groups

P622

What else?

$P6_122$ $P6_522$ $P6_222$ $P6_422$ $P6_322$

$P6$ $P6_1$ $P6_5$ $P6_2$ $P6_4$ $P6_3$

$P3_12$ $P321$ $P3_112$ $P3_121$ $P3_212$ $P3_221$

$P3$ $P3_1$ $P3_2$

$P2$ $P2_1$ $C2$ $C222$ $C222_1$

$P1$

Pitfalls: alternate indexing

`http://www.ccp4.ac.uk/html/reindexing.html`

Dangerous ones:

$P6_x$

$P4_x$ $I4_x$

$P3_x$ $R3_x$ $P3_xXX$

Trick Question:
What is the space group?

$a = 63$ $b = 63$ $c = 63$

$\alpha = 90$ $\beta = 90$ $\gamma = 90$

What you know when:

Autoindexing: Lattice (14)

Scaling: Point Group (43)

Refinement: Space Group (84)

Unit Cell

Smallest part of the crystal
that you can use to build up the rest of the crystal
with translation symmetry only

Asymmetric Unit (ASU)

Smallest part of the crystal
that you can use to build up the rest of the crystal
with symmetry operations

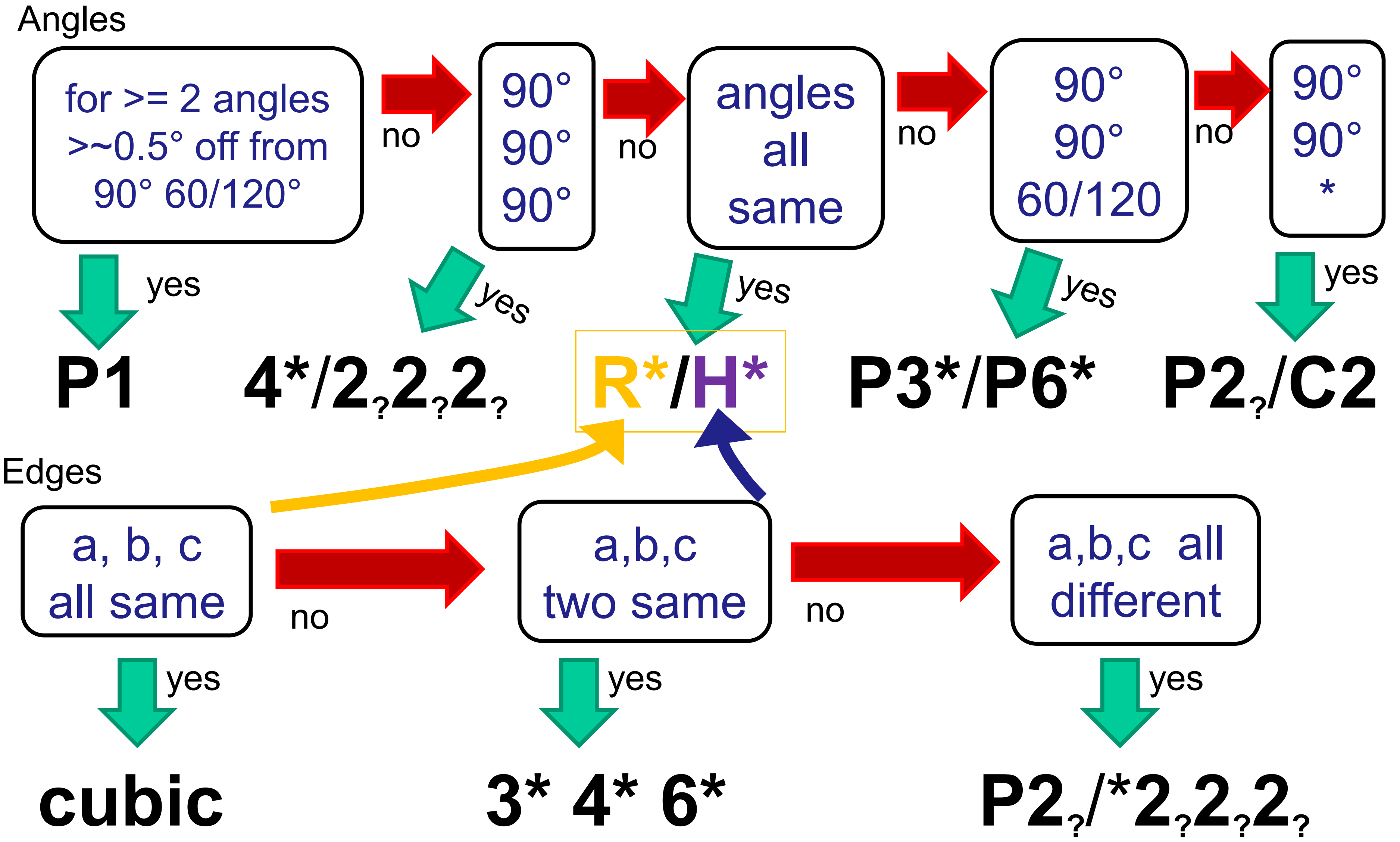
“NCS ASU”

Builds ASU with Non-crystallographic symmetry operations

Frequency of Space Groups in PDB

13007	$P2_12_12_1$	566	$R32$	191	$P6_322$	74	$P4_322$	36	$P3_212$
9629	$P2_1$	491	$P6_522$	187	$P2_13$	71	$P4_122$	30	$P622$
5300	$C2$	474	$P6_1$	161	$P6$	70	$P6_4$	24	$F4_132$
3021	$P2_12_12$	421	$P6_5$	144	$P42_12$	70	$P6_222$	22	$P4_332$
2543	$C222_1$	353	$P4_1$	142	$I23$	67	$P6_422$	22	$P422$
2109	$P1$	351	$P6_3$	132	$F432$	66	$F222$	22	$P23$
2030	$P4_32_12$	286	$P3_1$	124	$P321$	63	$P6_2$	19	$P3_112$
1565	$P3_221$	257	$I422$	115	$I4_1$	59	$P4_132$	18	$P4_222$
1359	$P4_12_12$	253	$P4_22_12$	103	$P3$	46	$P222_1$	18	$I4_132$
1198	$P3_121$	249	$I4$	99	$I2_13$	43	$P4$	16	$P432$
1147	$I222$	246	$P4_3$	94	$C222$	43	$I432$	8	$P4_232$
825	$R3$	239	$I4_122$	93	$P2$	40	$P4_2$	3	$P222$
620	$P6_122$	193	$P3_2$	82	$I2_12_12_1$	37	$F23$	2	$P312$

Space Group Thought Process



Angles

for >= 2 angles >~0.5° off from 90° 60/120° — yes → **P1**; no →

90° 90° 90° — yes → **$4^*/2_?2_?2_?$**; no →

angles all same — yes → **R*/H***; no →

90° 90° 60/120 — yes → **P3*/P6***; no →

90° 90° * — yes → **$P2_?/C2$**

Edges

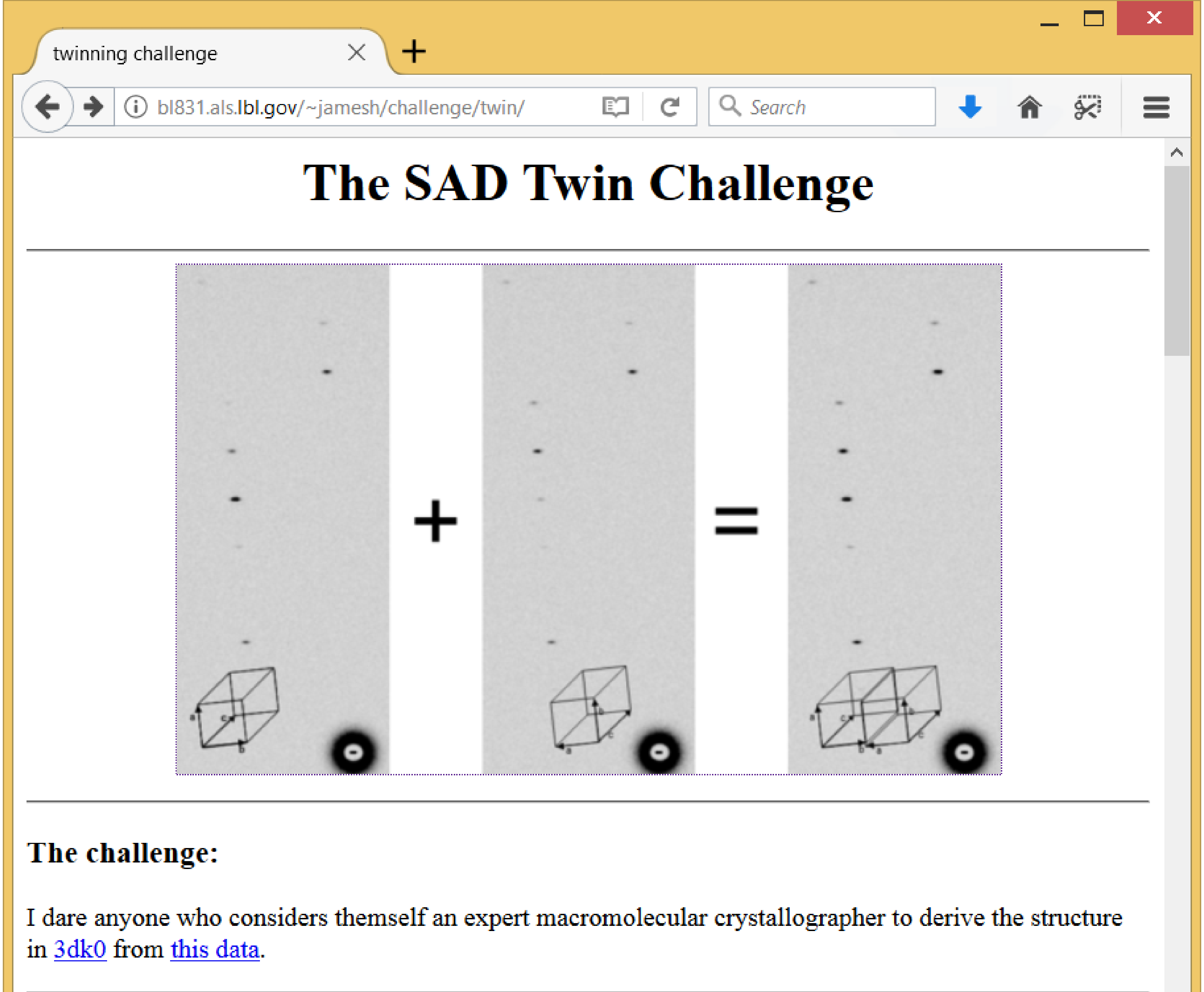
a, b, c all same — yes → **cubic**; yes → **R***; no →

a,b,c two same — yes → **3* 4* 6***; yes → **H***; no →

a,b,c all different — yes → **$P2_?/{}^*2_?2_?2_?$**

Evil Twinning?!?

and pseudo symmetry



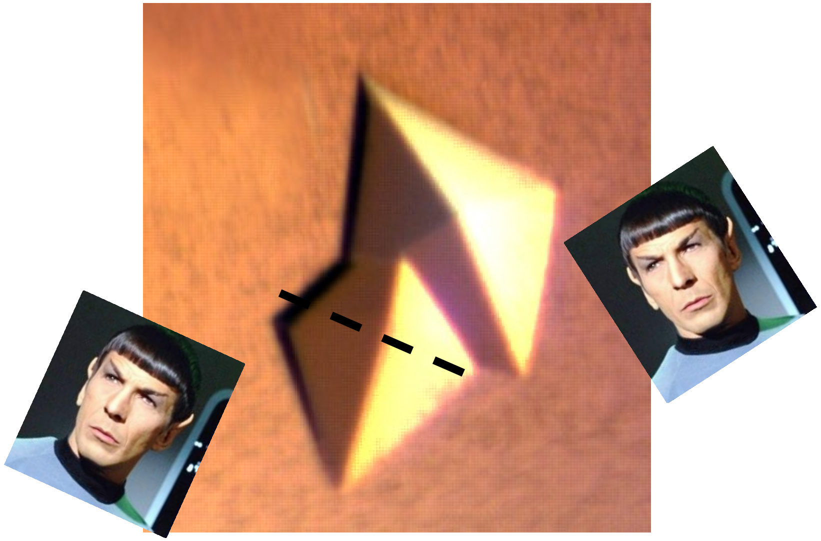
The SAD Twin Challenge

The challenge:

I dare anyone who considers themself an expert macromolecular crystallographer to derive the structure in 3dk0 from this data.

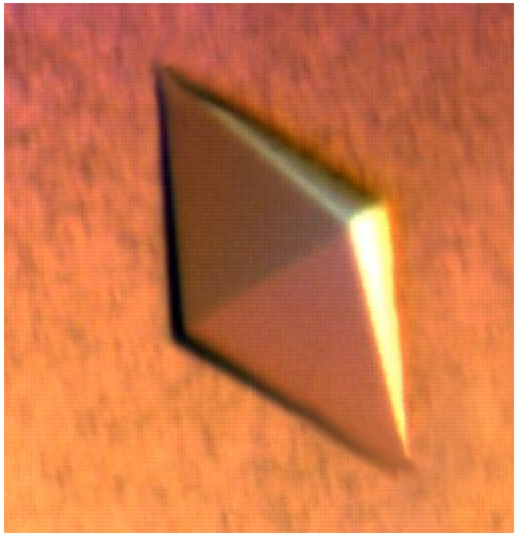
Can you tell by looking?

Macroscopic twin



Disecting apart might give a single crystal!

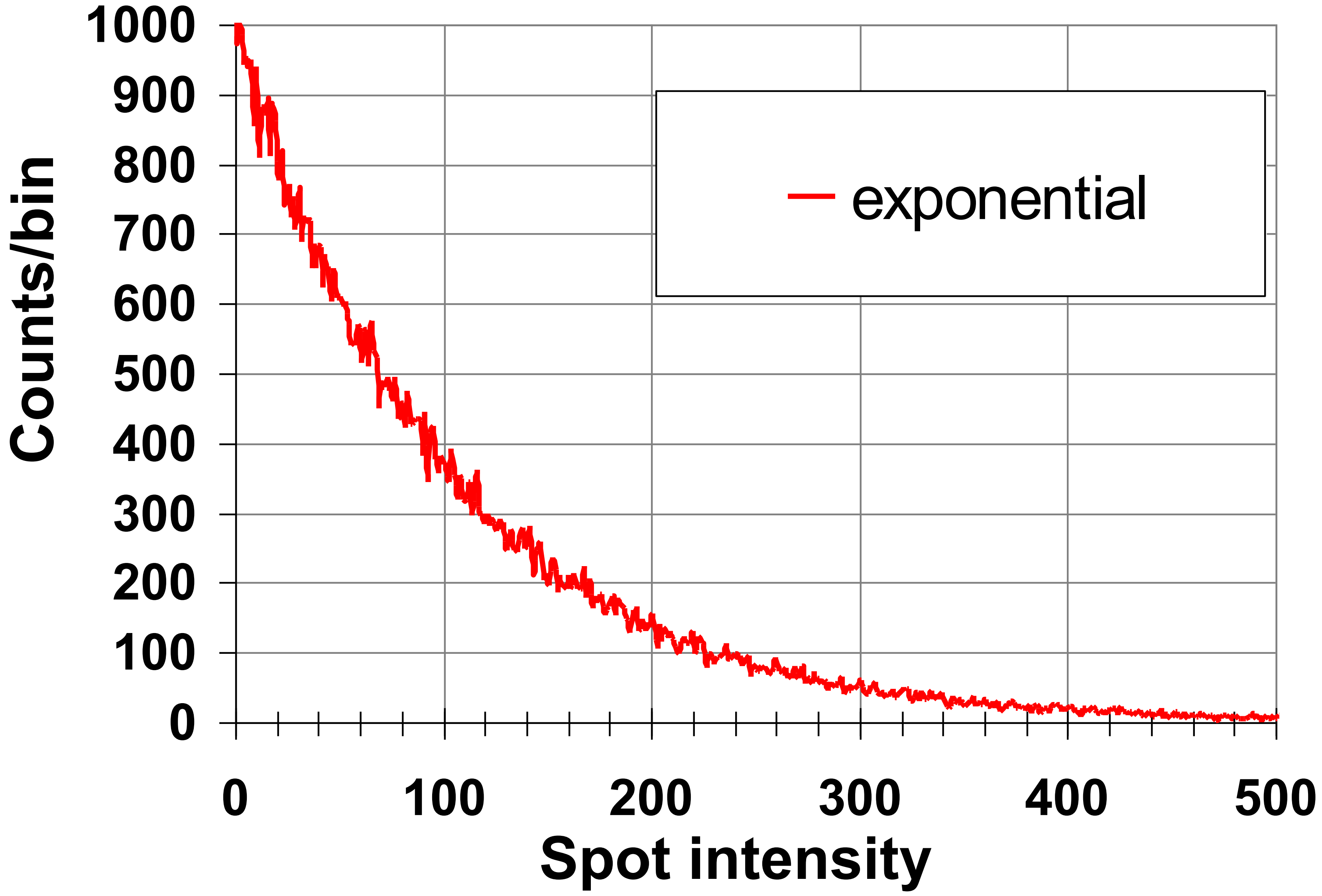
Microscopic twin



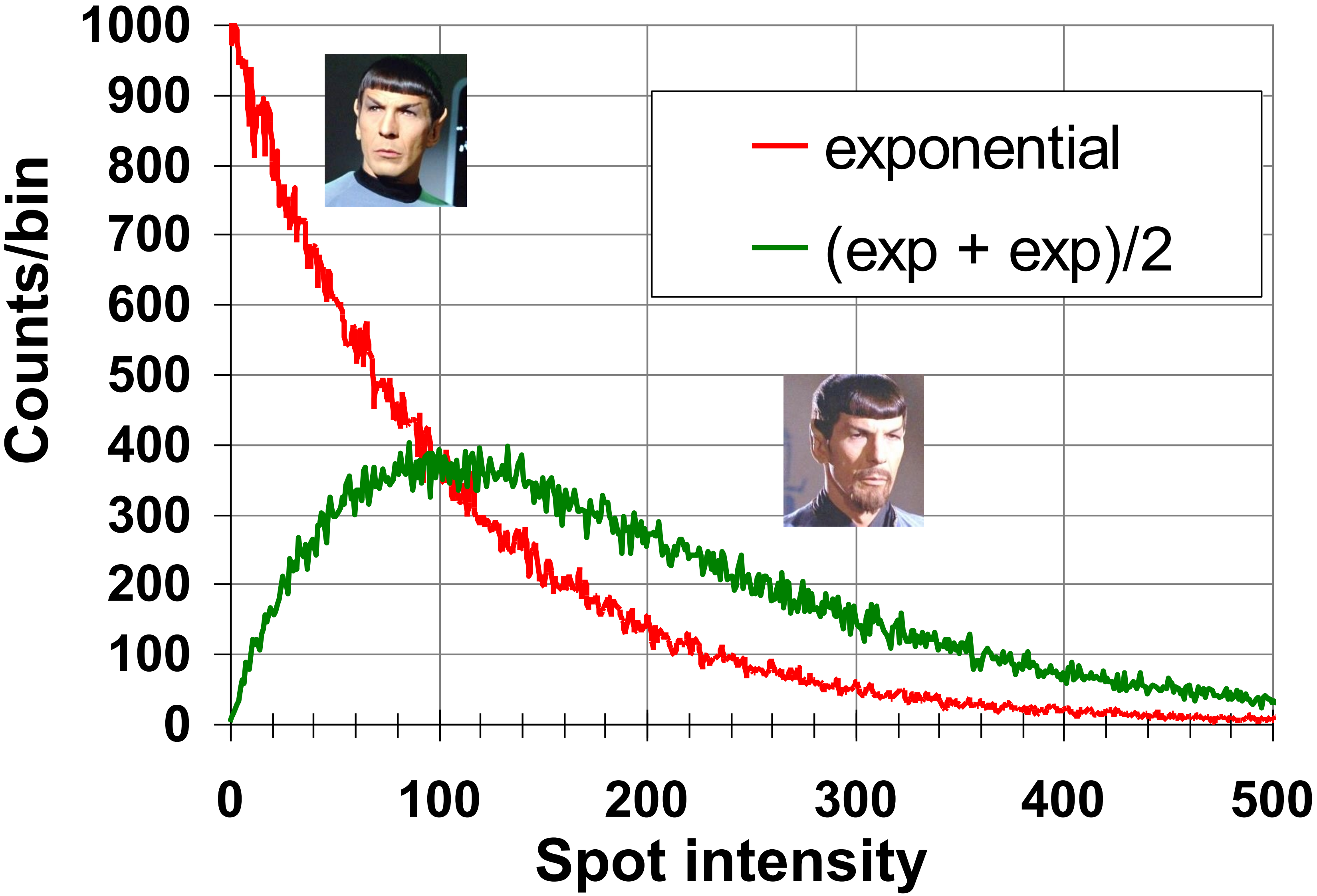
Crystal looks fine, but is twinned.

Images property of Andrea Thorn
and Claudia Egerer-Sieber

Make a histogram of your data



Make a histogram of your data



Twinning: how can you tell?

1. Intensity distribution is abnormal
   H test, L test
2. R_{merge} for higher symmetry
   < 50% → might be twin
   < 10% → could be 50:50 twin
3. one of the “usual suspects”
   P622, P6, P321, P312, P422, P43
   http://www.ccp4.ac.uk/html/twinning.html
4. can’t solve it

Find the twin law?

1. refine in REFMAC

   Automatically checks all possible twin laws!

2. copy twin law into other programs:

   phenix.refine, etc.

Pick R_{free} flags in highest symmetry

i.e. P622 if true SG is P3

Do not “twin refine” if not twinned

control: swap non-twin operator

Twinning: what do I do?

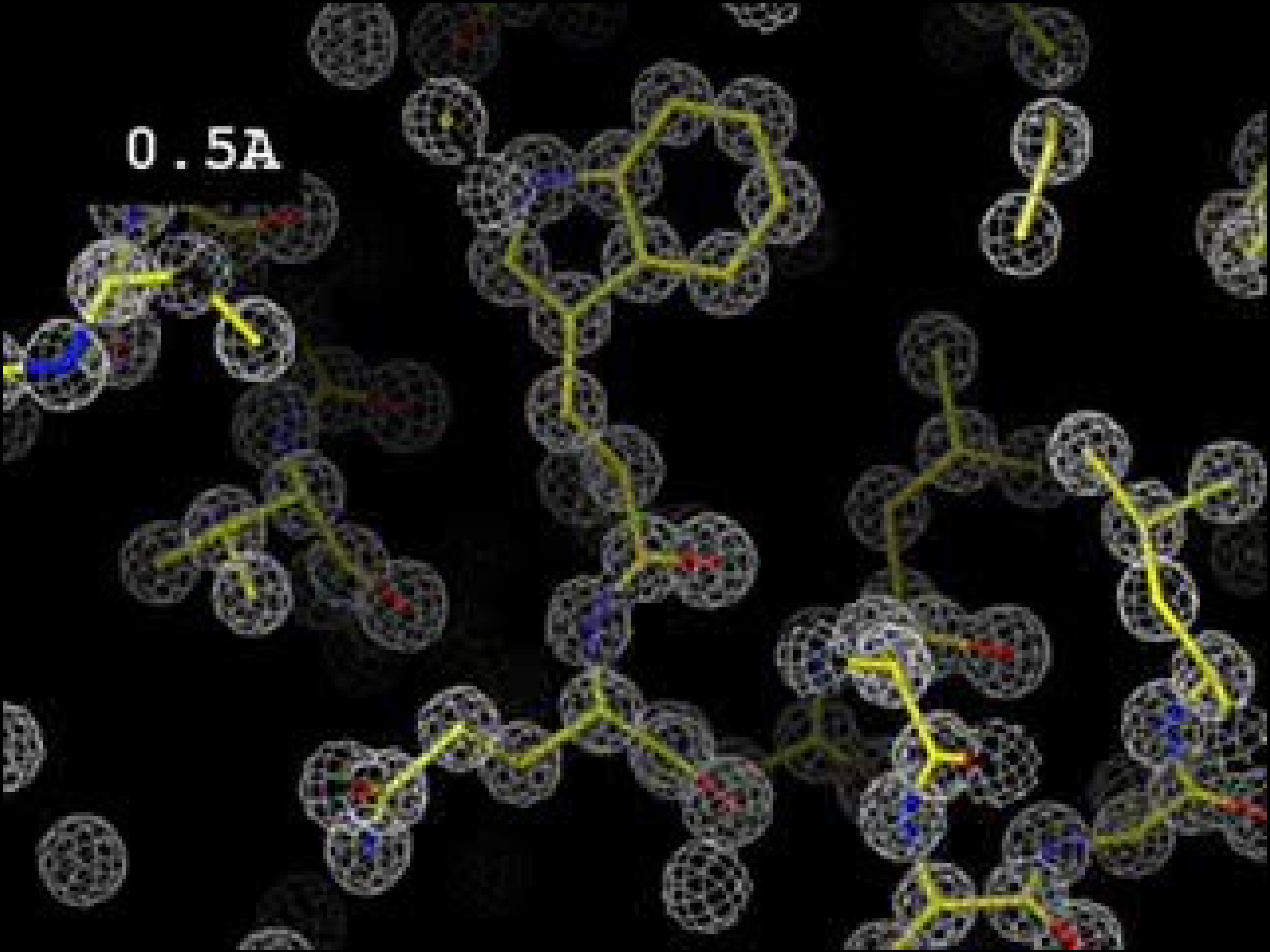
1. Integrate with low-symmetry SG
   Yes, all possible ones

2. try to solve it
   Yes, in all possible SGs

MR: good chance if copies = 1
MAD: maybe
S-SAD: no way
MIR: if you're lucky

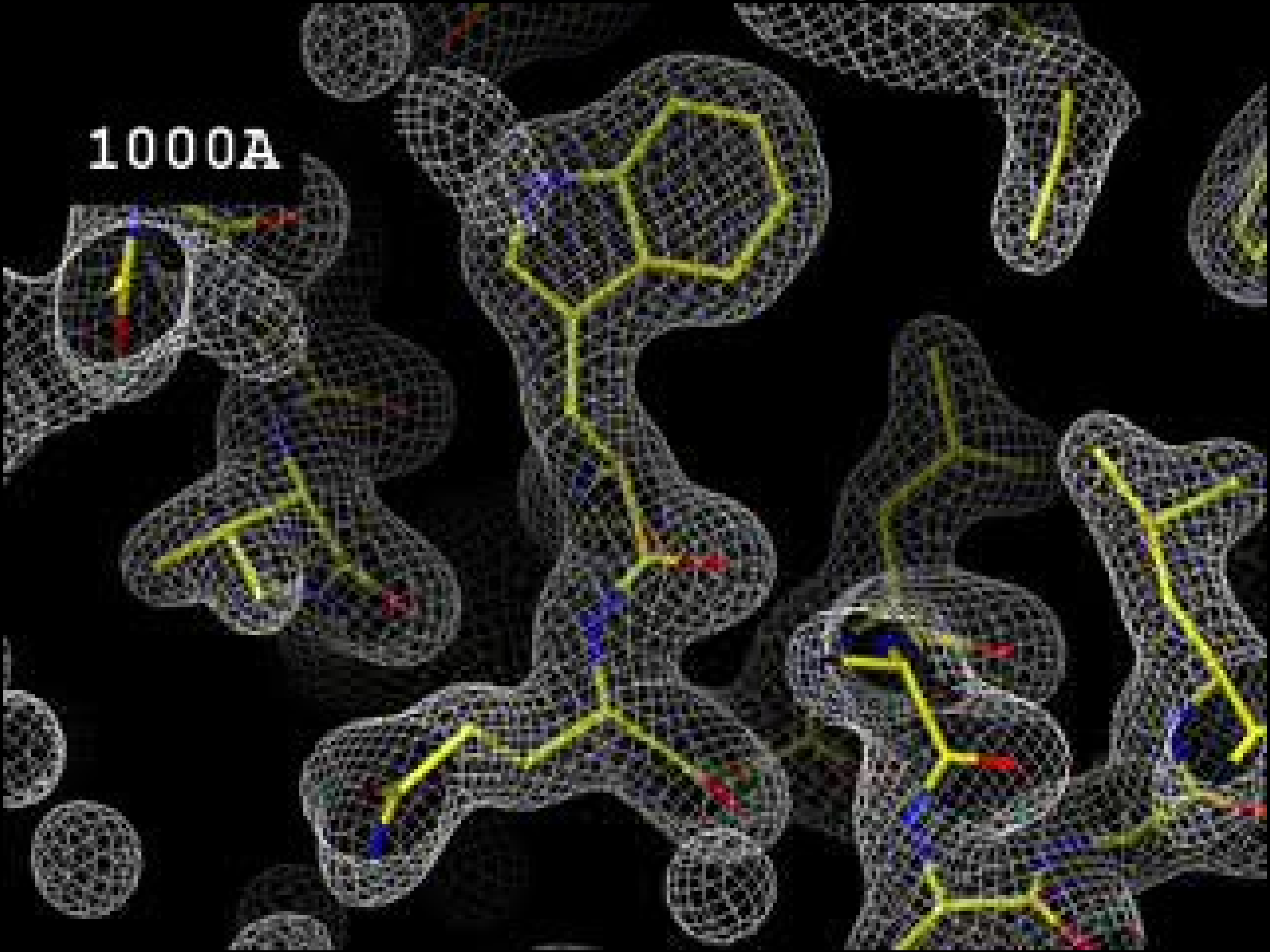
Thanks to: Andrea Thorn

Resolution



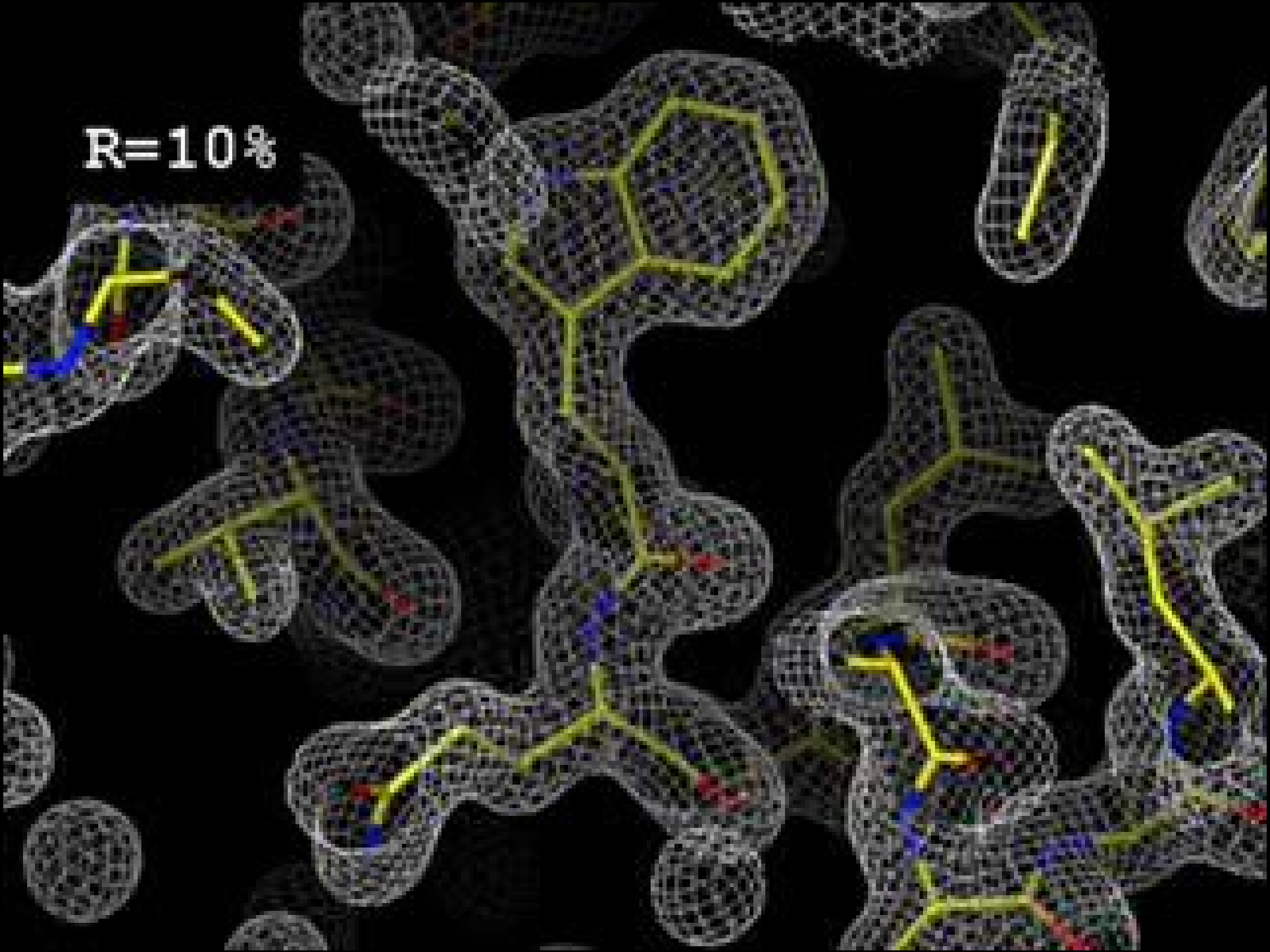
http://bl831.als.lbl.gov/~jamesh/movies/resolution.mpeg

Resolution: low-angle cutoff



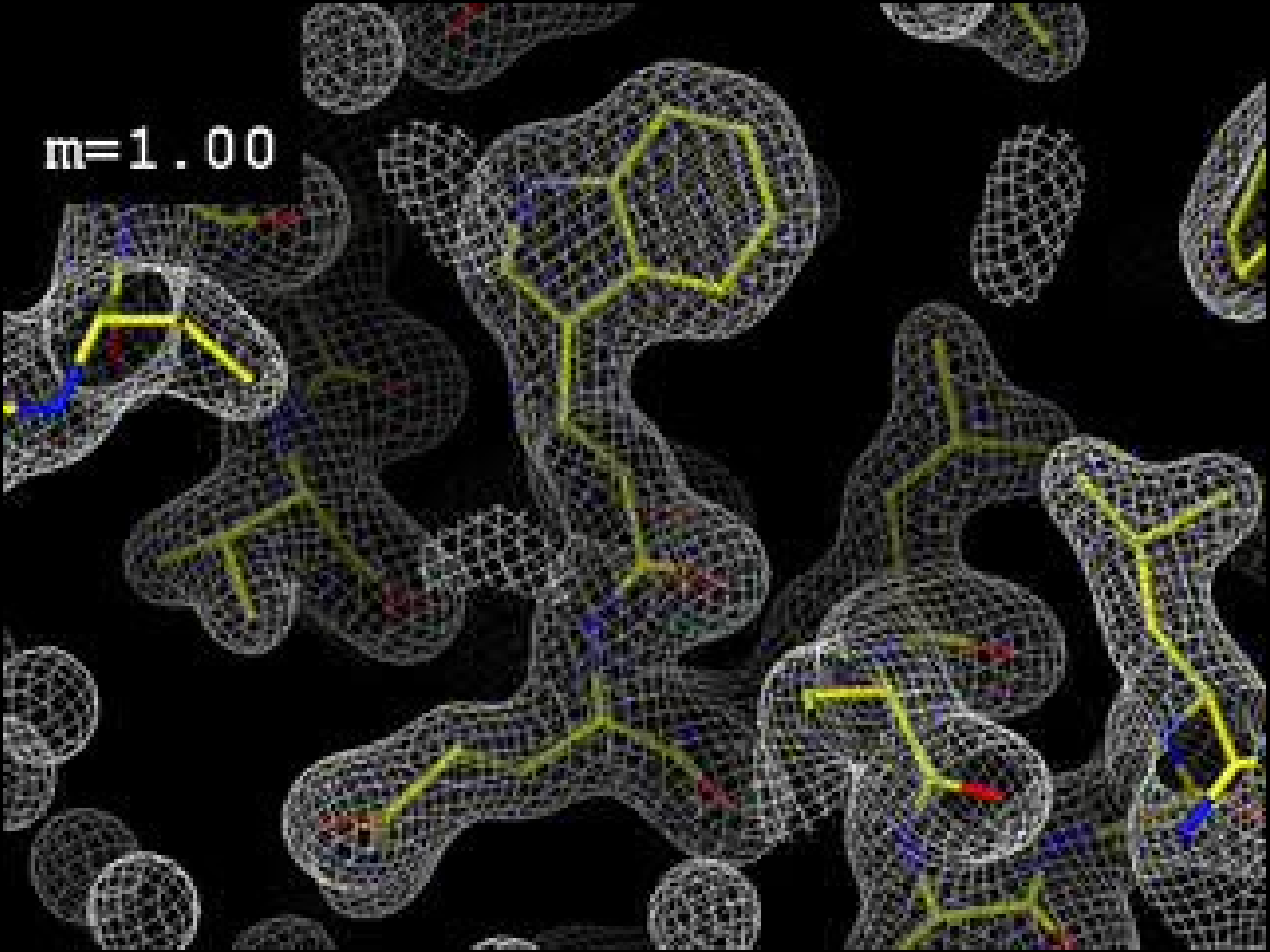
http://bl831.als.lbl.gov/~jamesh/movies/lo_cut.mpeg

R-factor



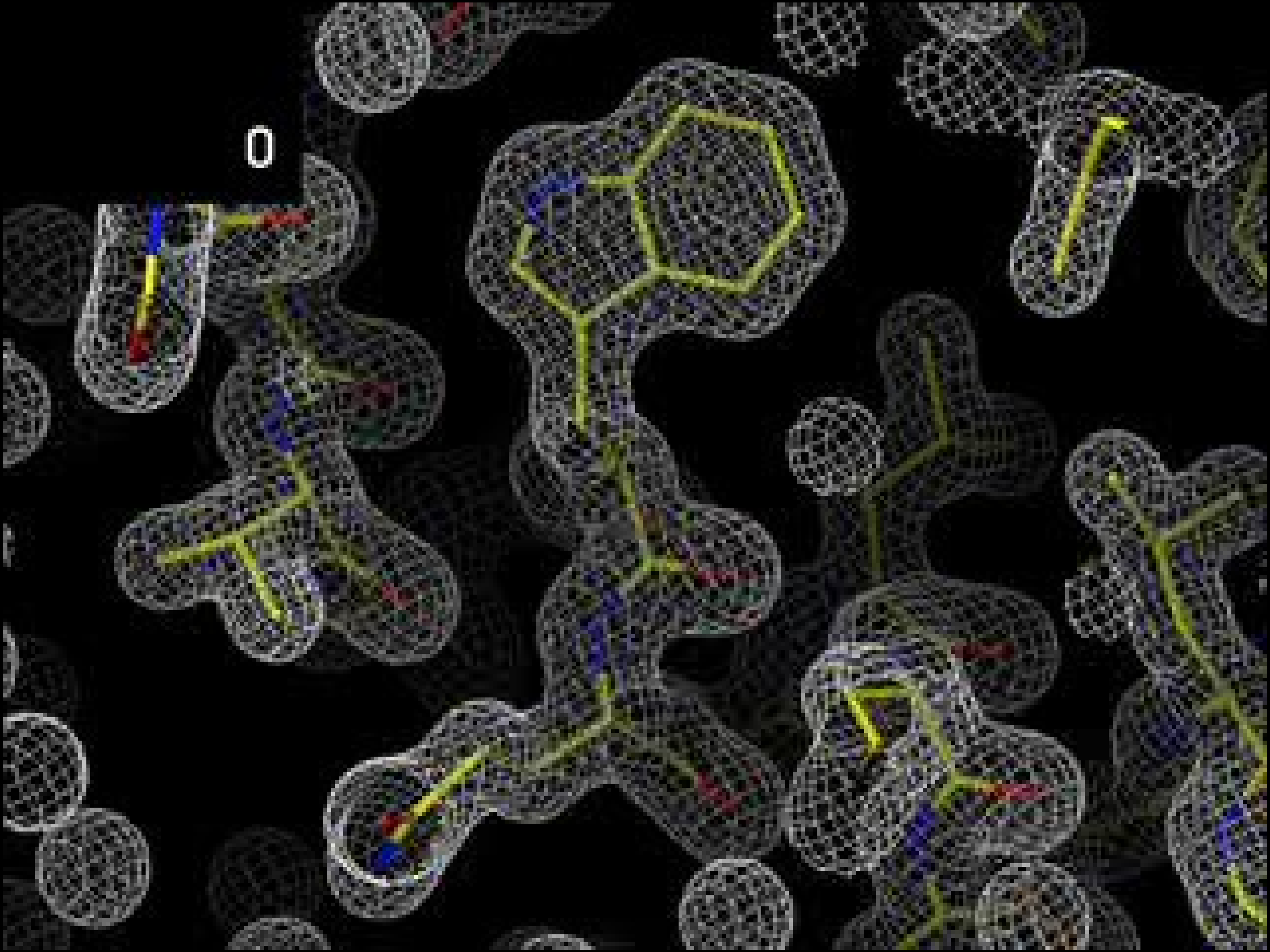
http://bl831.als.lbl.gov/~jamesh/movies/rfactor.mpeg

Figure of Merit



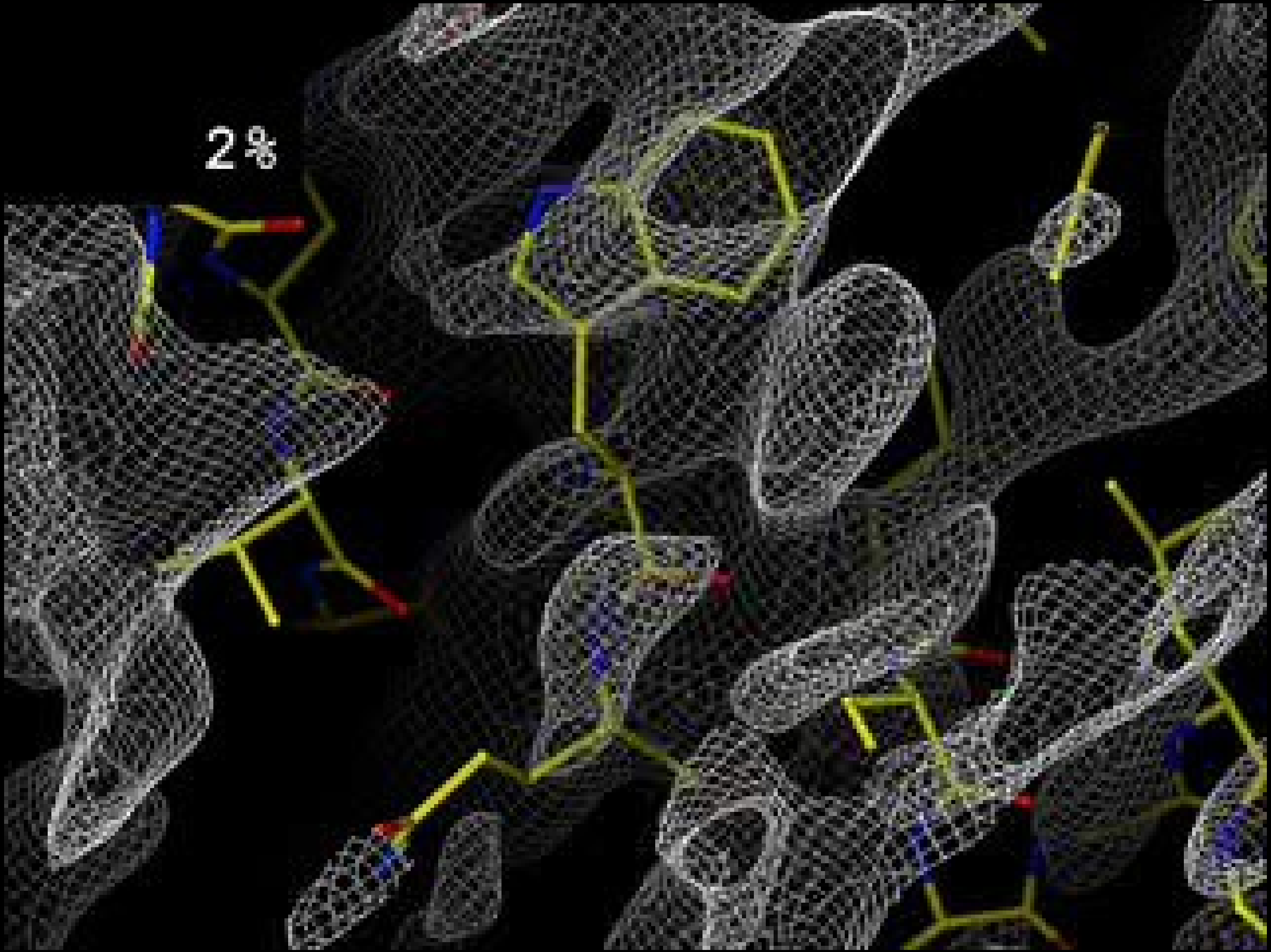
http://bl831.als.lbl.gov/~jamesh/movies/dephase.mpeg

Overloads



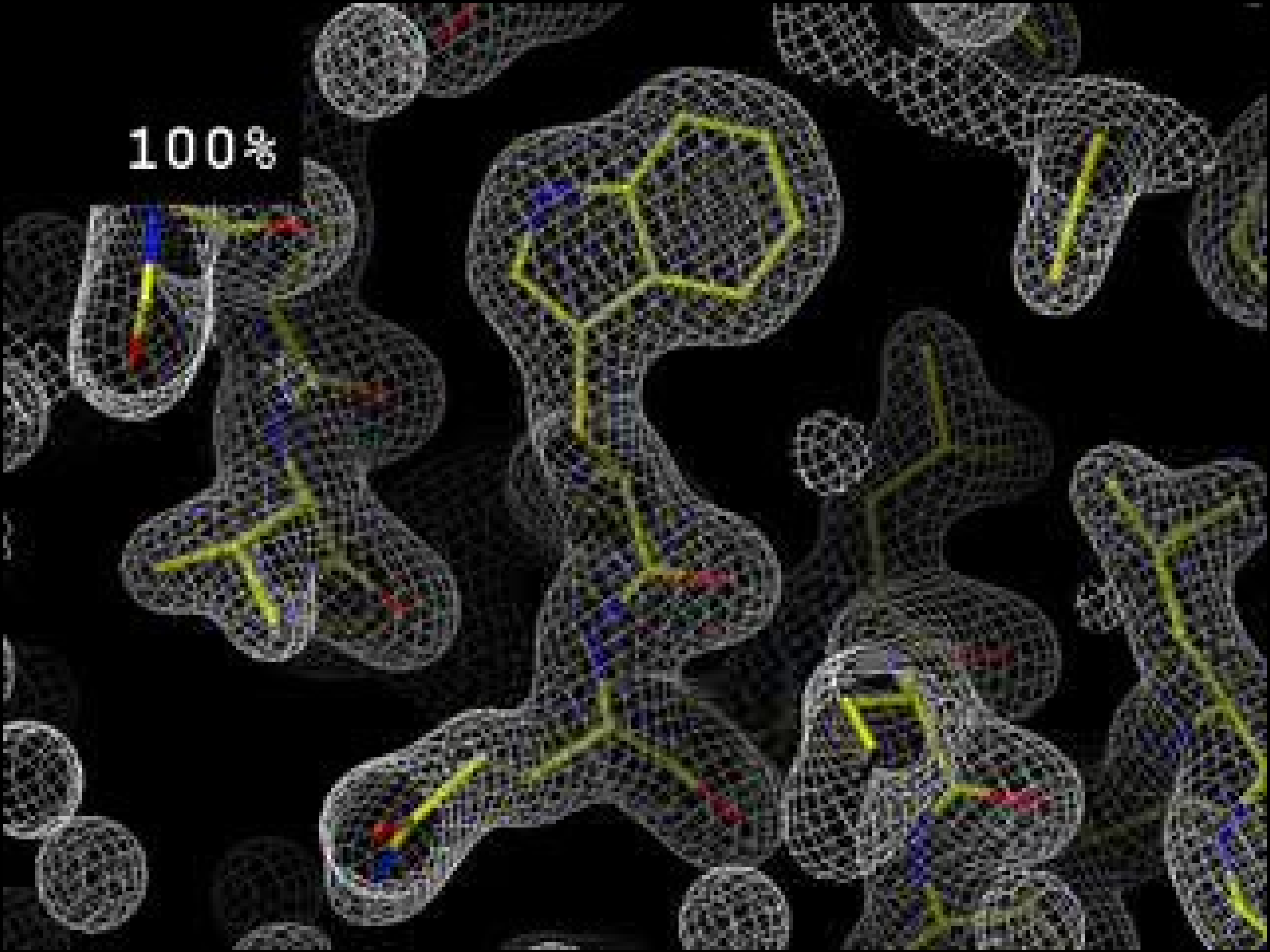
http://bl831.als.lbl.gov/~jamesh/movies/overloads.mpeg

Completeness: missing wedge



http://bl831.als.lbl.gov/~jamesh/movies/osc.mpeg

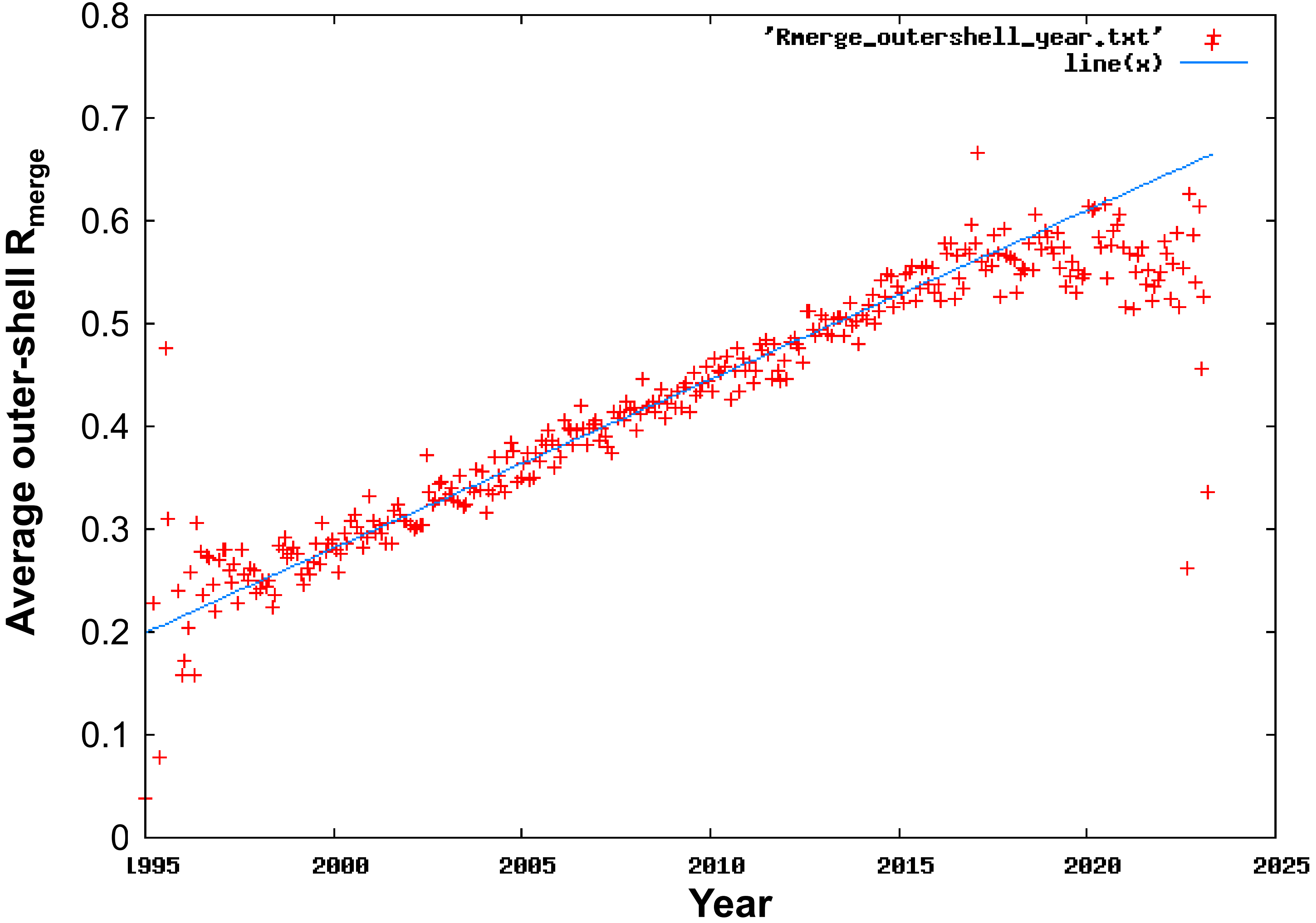
Completeness: random deletion



http://bl831.als.lbl.gov/~jamesh/movies/completeness.mpeg

Resolution cutoff?

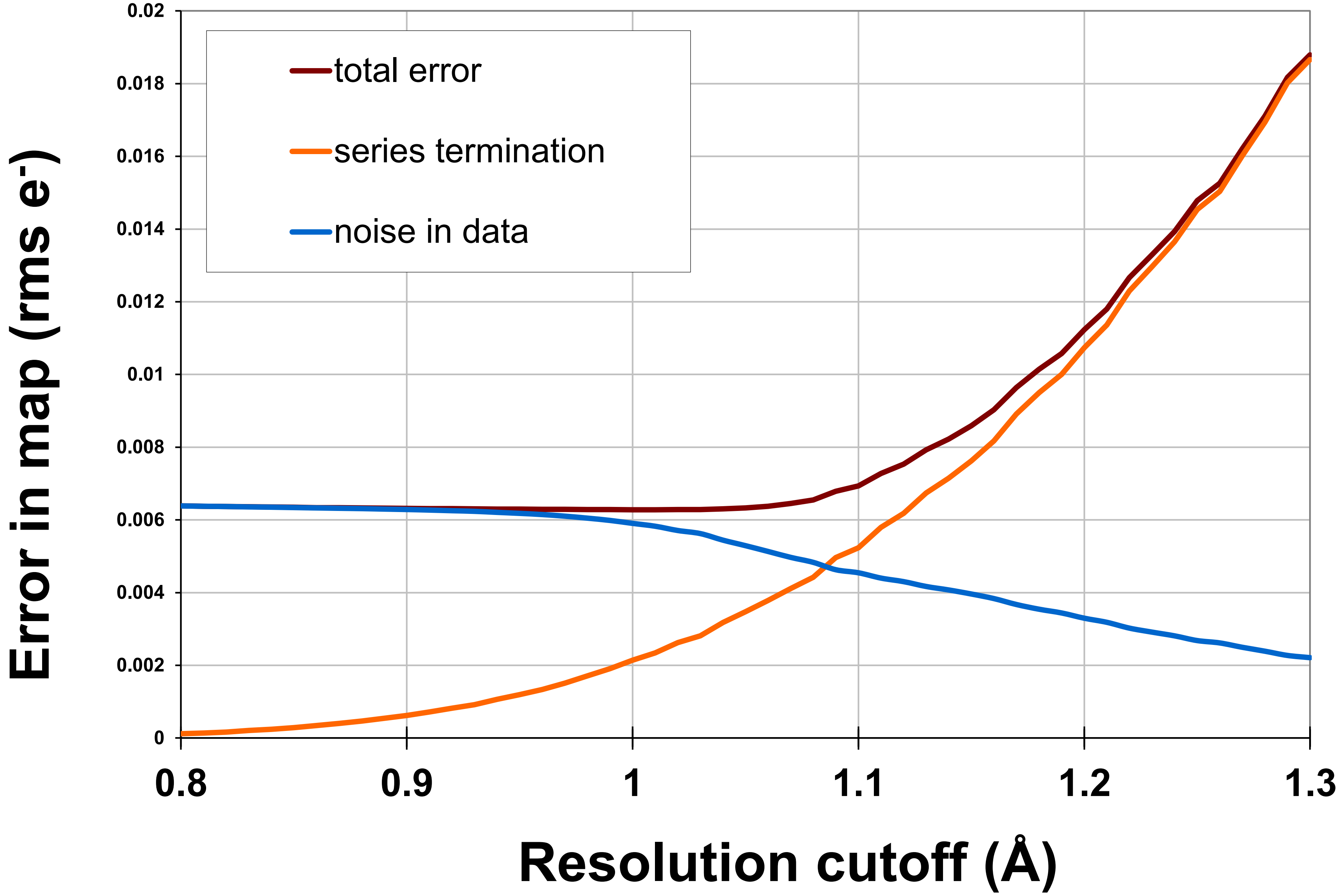
R_{merge} at the resolution limit in PDB



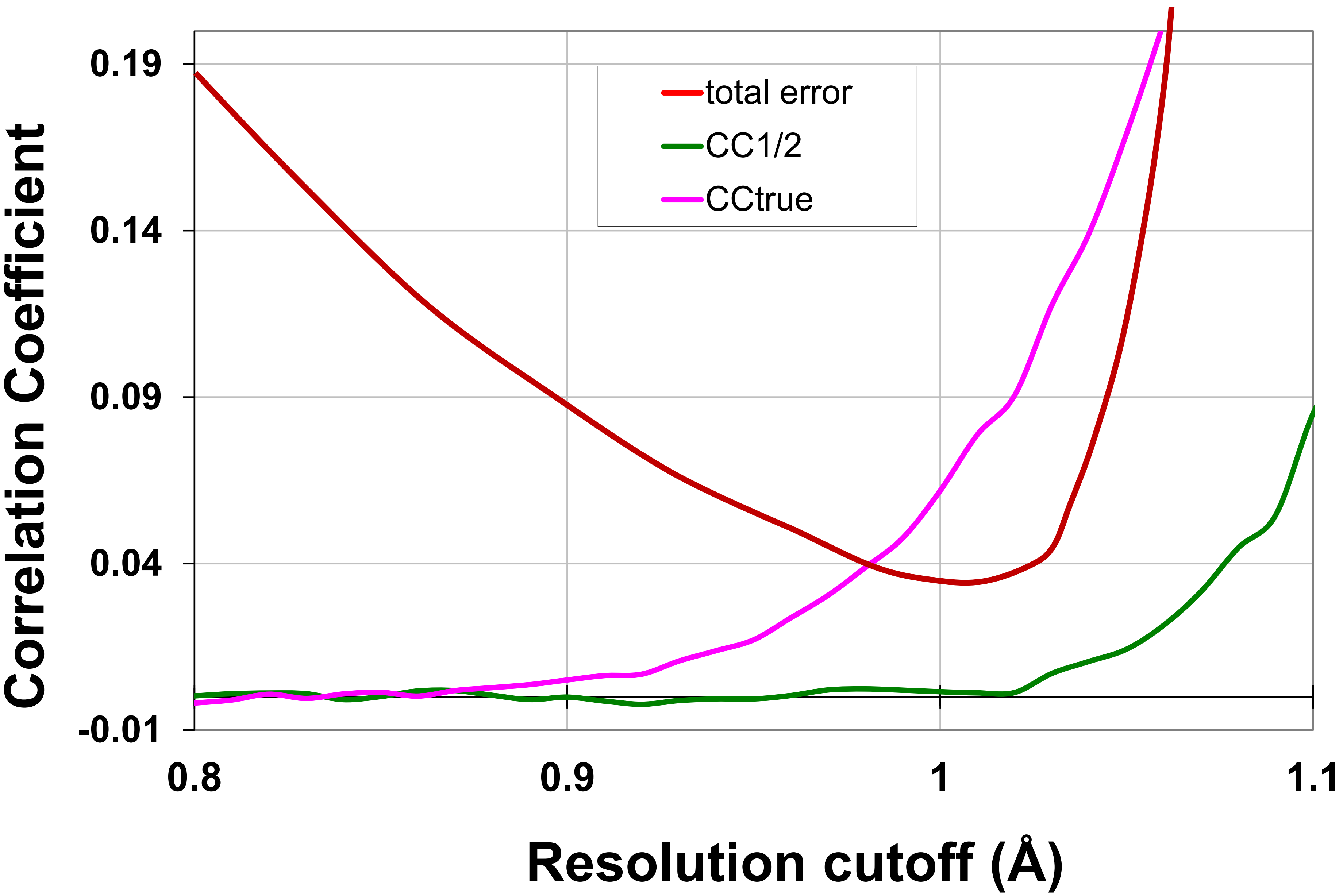
Optimum resolution cutoff is:

- Too optimistic: add nothing but noise
- Too pessimistic: series-termination error
- Happy medium?
- Simulate:
  - Random atoms, compute F^2
  - Add Gaussian noise, RMS = 1
  - Truncate
  - Subtract “right” map, RMS difference

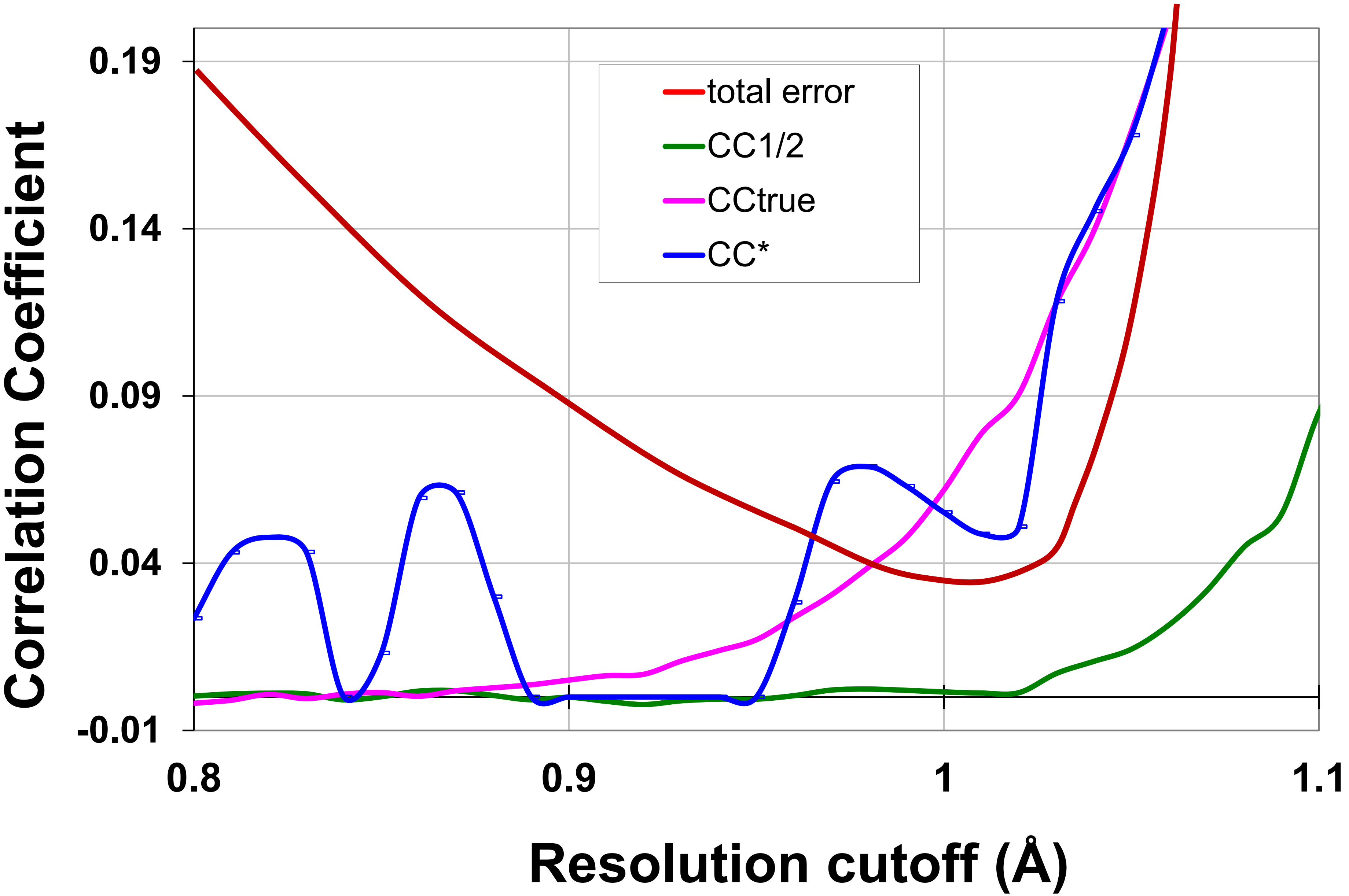
Optimal resolution cutoff



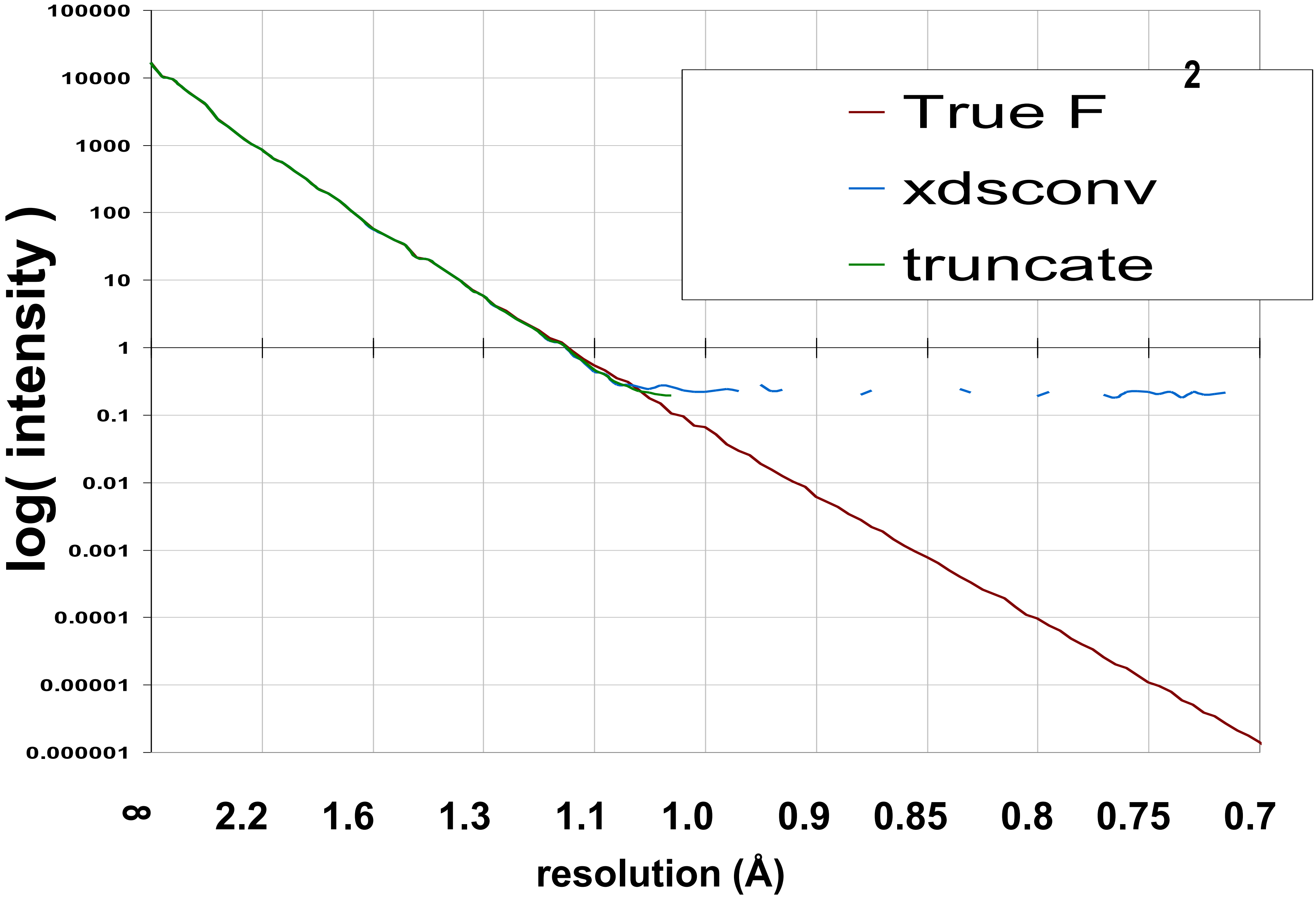
Optimal resolution cutoff



Optimal resolution cutoff



Wilson Plot


100000
10000
1000
100
10
1
0.1
0.01
0.001
0.0001
0.00001
0.000001
log(intensity)
True F^2
xdsconv
truncate
∞
2.2
1.6
1.3
1.1
1.0
0.9
0.85
0.8
0.75
0.7
resolution (Å)

Optimum resolution cutoff is:

0.0 Å

But my reviewer says…

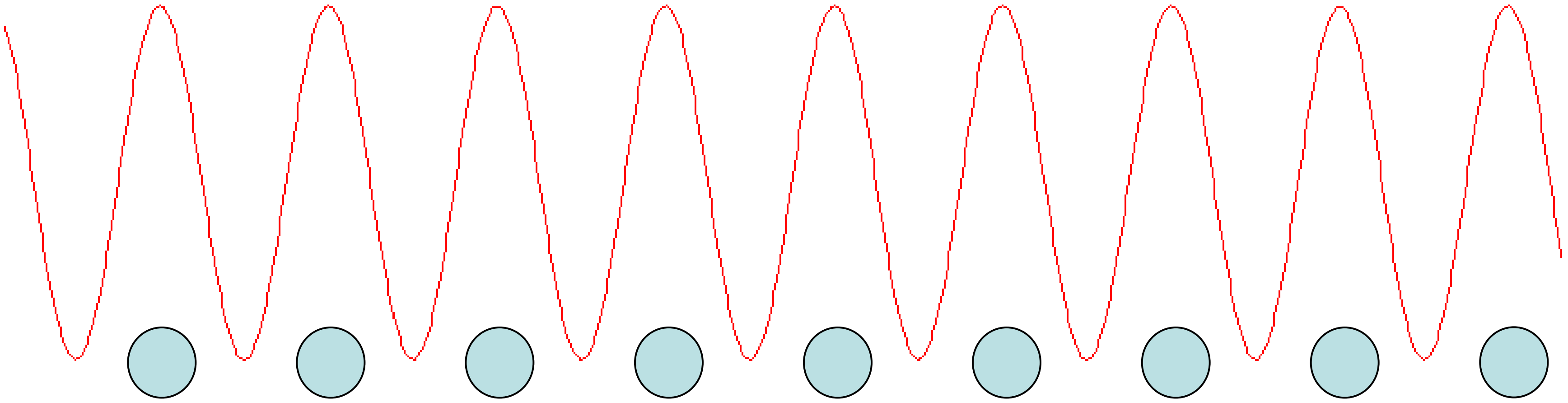
R_{merge} in high-resolution bin is >100% !

Answer: this is expected

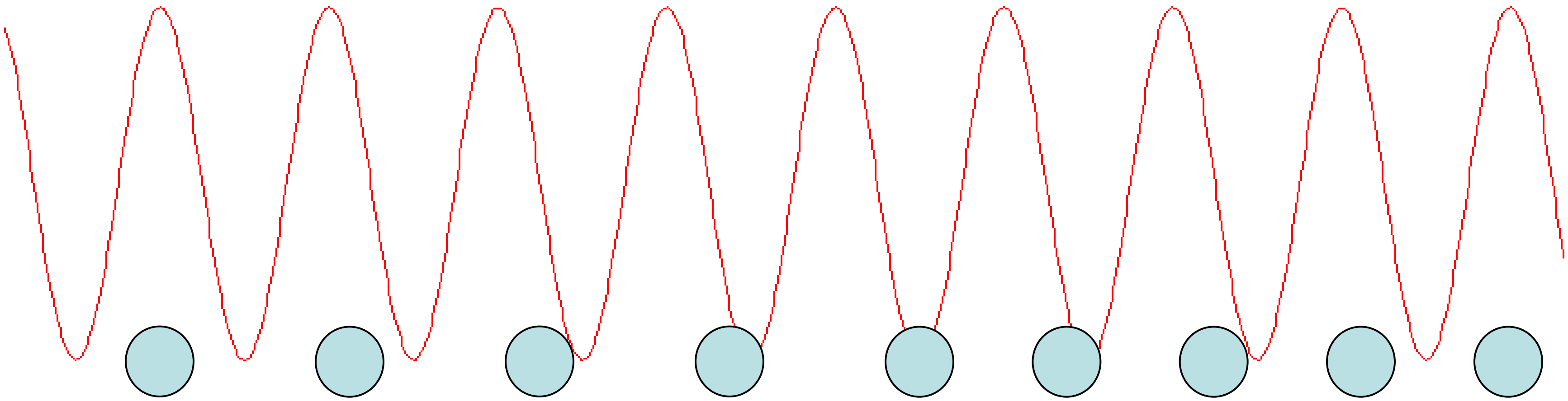
Expected R_{merge} as $I_{obs} \rightarrow 0$

$$R_{merge} = \frac{\sum \left| I_{obs} - \langle I \rangle \right|}{\sum I_{obs}}$$

strong reflection



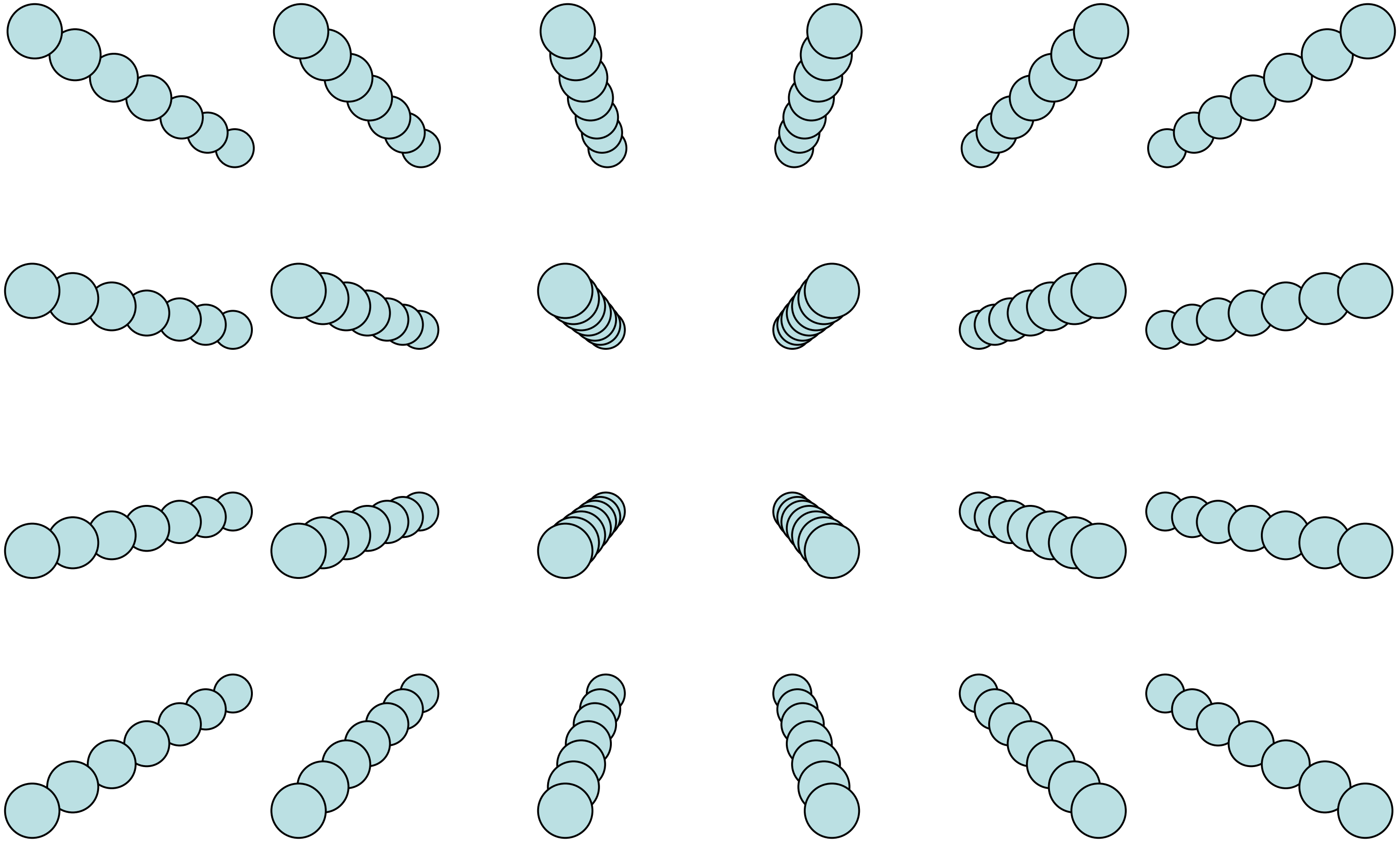
weak reflection



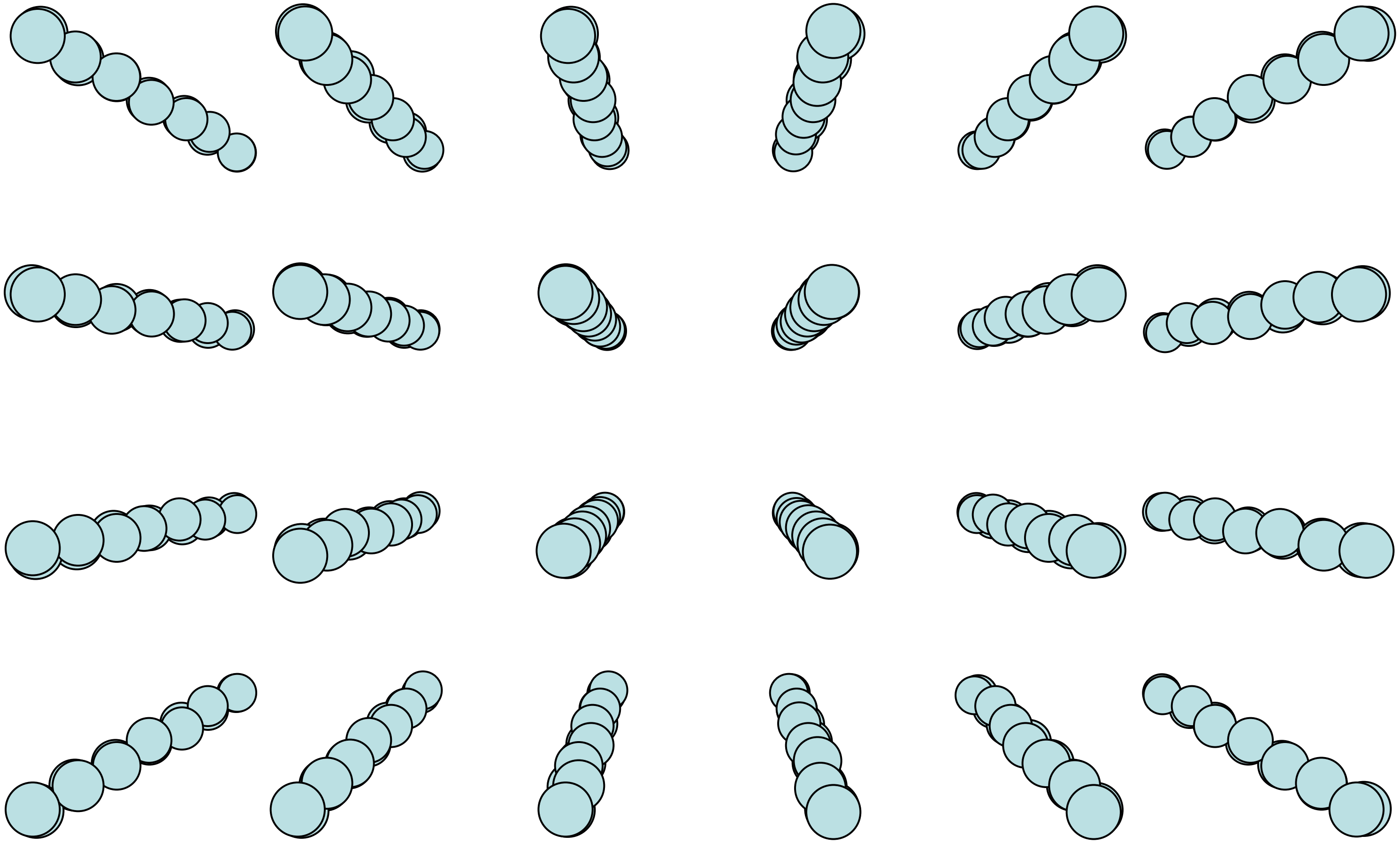
“B” factors

```
ATOM    122  N   LEU A  13      -3.244  25.808  19.998  1.00 16.96
ATOM    123  CA  LEU A  13      -2.877  25.448  21.355  1.00 15.29
ATOM    124  C   LEU A  13      -2.792  23.966  21.561  1.00 17.54
ATOM    125  O   LEU A  13      -1.814  23.493  22.143  1.00 16.35
ATOM    126  CB  LEU A  13      -3.907  26.164  22.268  1.00 18.72
ATOM    127  CG  LEU A  13      -3.577  25.982  23.738  1.00 21.19
ATOM    128  CD1 LEU A  13      -2.283  26.820  24.019  1.00 19.43
ATOM    129  CD2 LEU A  13      -4.702  26.474  24.639  1.00 24.65
ATOM    130  N   SER A  14      -3.677  23.149  20.979  1.00 15.96
ATOM    131  CA  SER A  14      -3.646  21.711  21.061  1.00 18.26
ATOM    132  C   SER A  14      -2.373  21.203  20.360  1.00 18.71
ATOM    133  O   SER A  14      -1.747  20.315  20.930  1.00 17.47
ATOM    134  CB  SER A  14      -4.875  21.077  20.419  1.00 17.62
ATOM    135  OG ASER A  14      -4.825  19.665  20.388  0.50 20.89
ATOM    136  OG BSER A  14      -6.027  21.408  21.164  0.50 18.67
ATOM    137  N   LYS A  15      -2.045  21.772  19.215  1.00 18.03
ATOM    138  CA  LYS A  15      -0.799  21.361  18.555  1.00 18.12
ATOM    139  C   LYS A  15       0.446  21.707  19.351  1.00 18.81
ATOM    140  O   LYS A  15       1.400  20.948  19.411  1.00 17.77
ATOM    141  CB  LYS A  15      -0.700  22.034  17.177  1.00 14.49
ATOM    142  CG  LYS A  15      -1.727  21.368  16.256  1.00 16.12
ATOM    143  CD  LYS A  15      -1.663  22.147  14.936  1.00 19.40
ATOM    144  CE ALYS A  15      -2.725  21.614  13.986  0.50 17.42
```

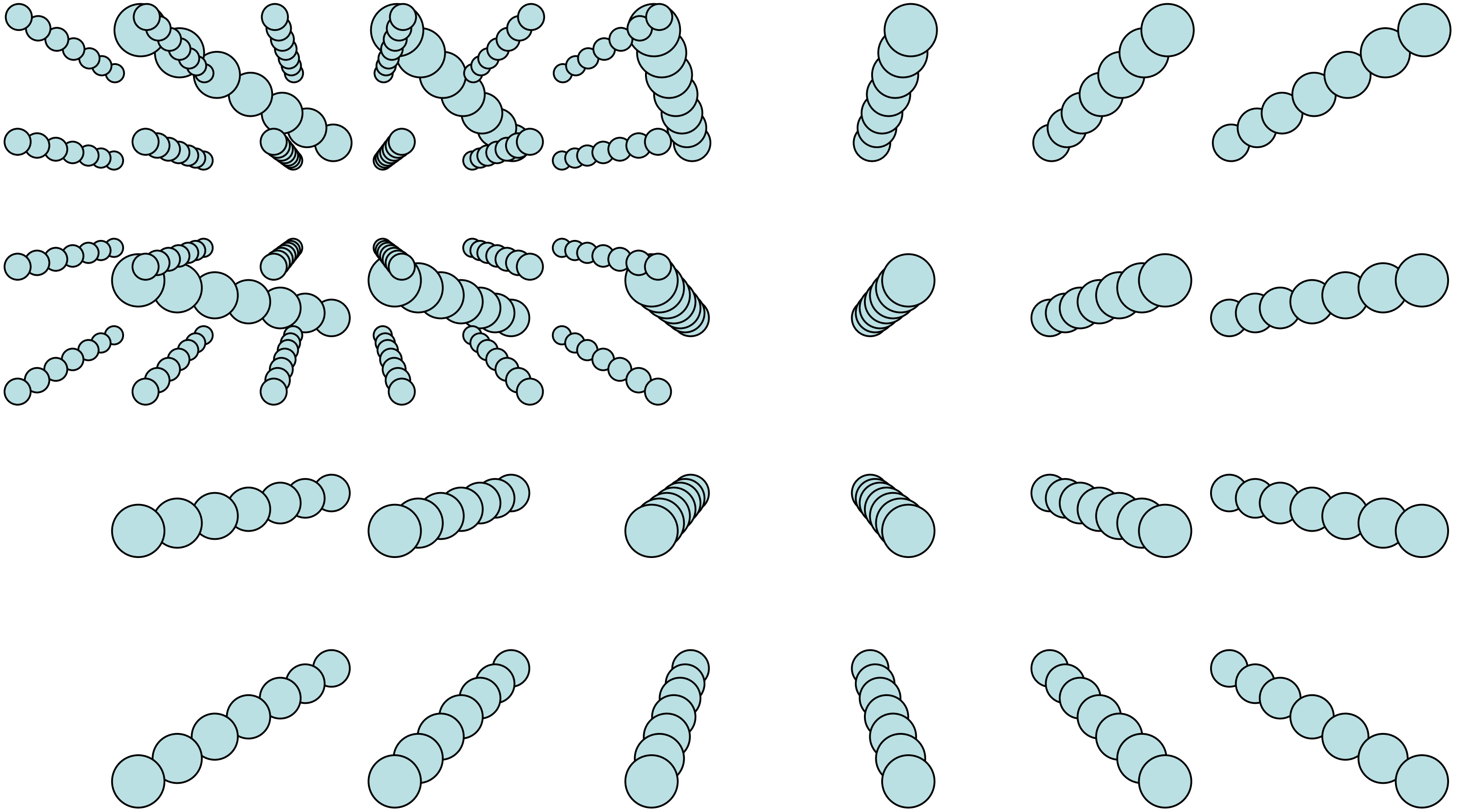
What is a "B factor"?



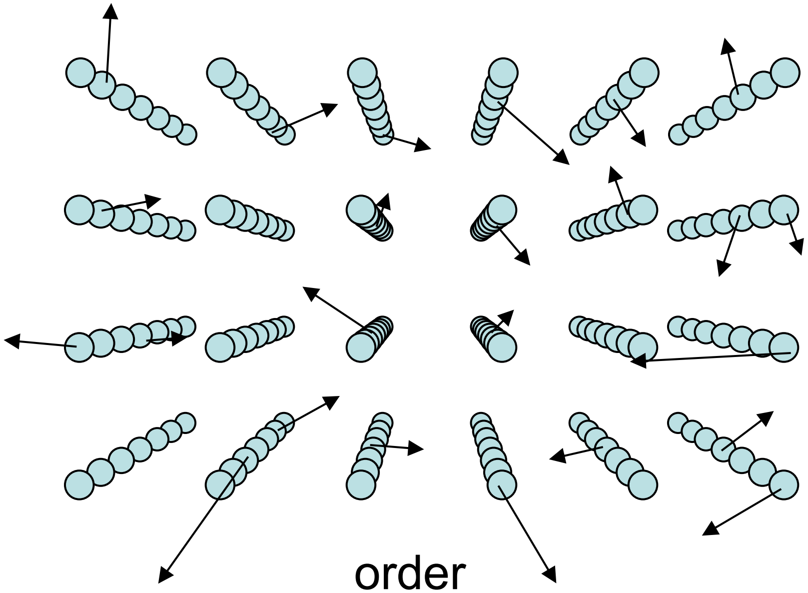
What is a "B factor"?



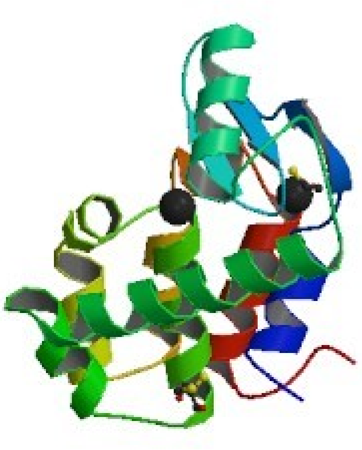
What is a "B factor"?



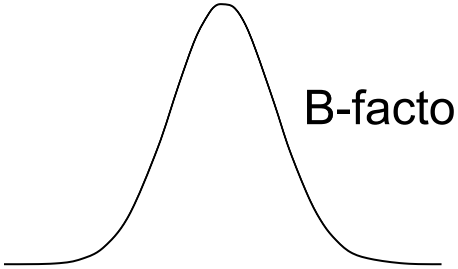
What is a "B factor"?



order



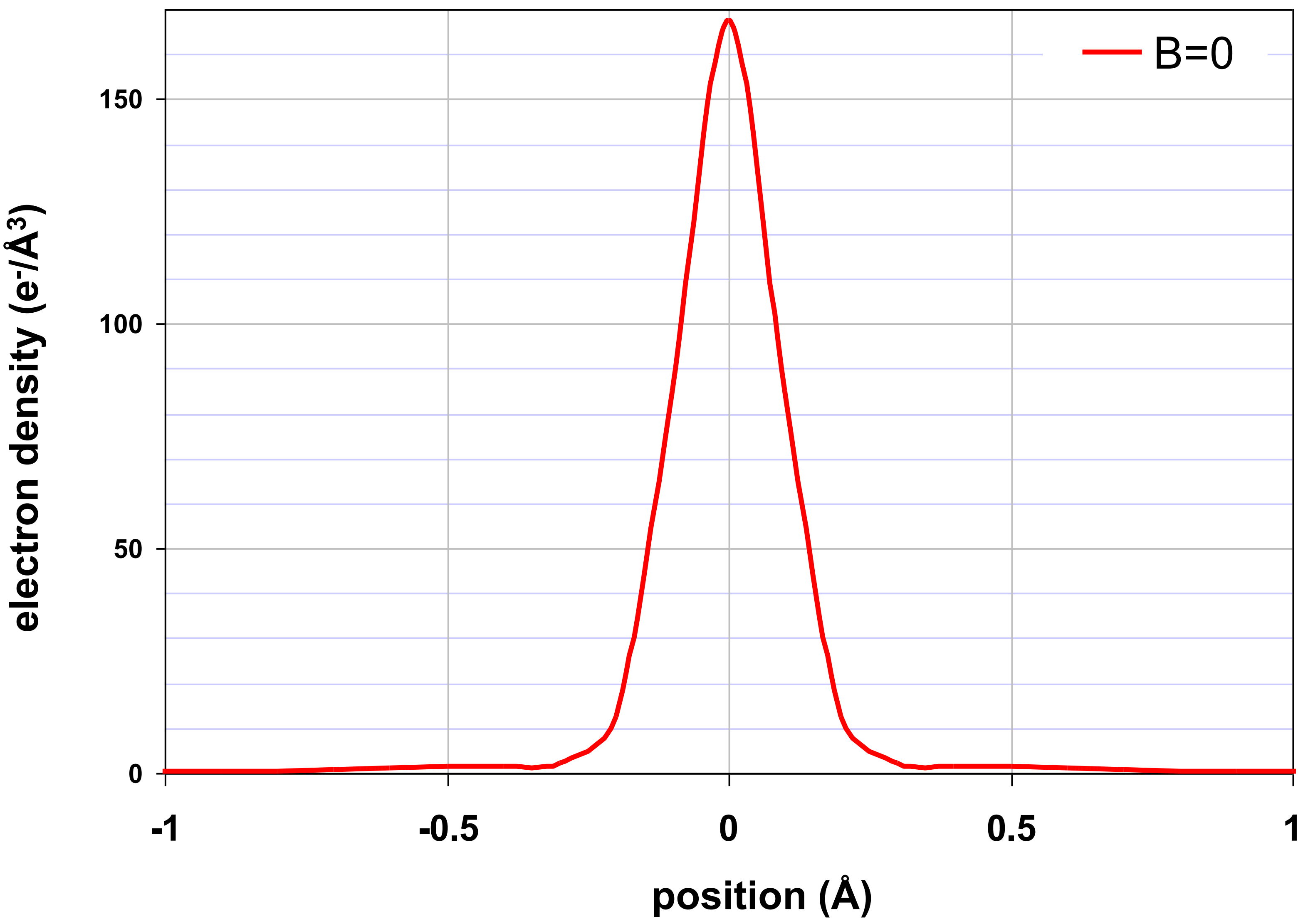
order



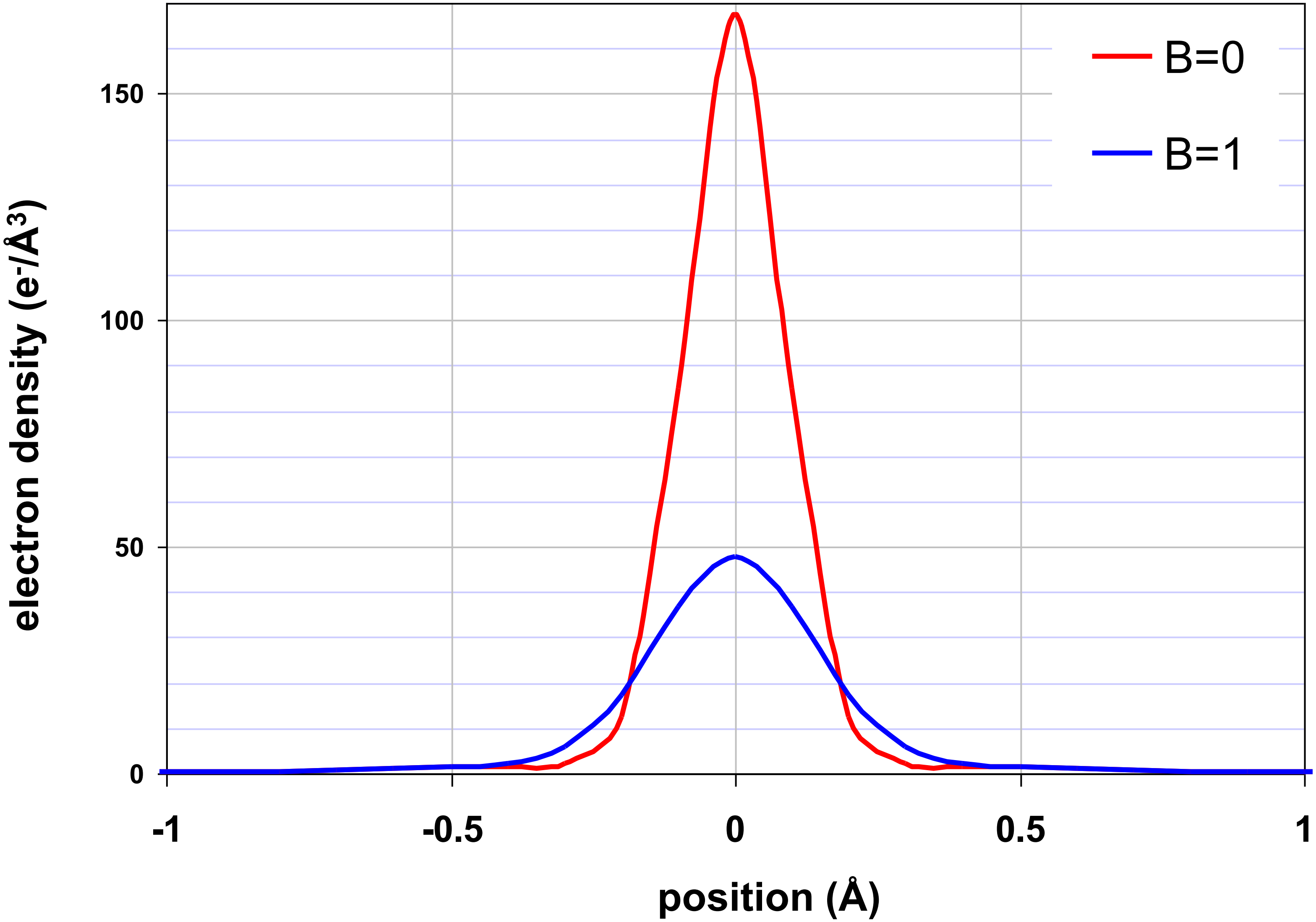
B-factor

disorder

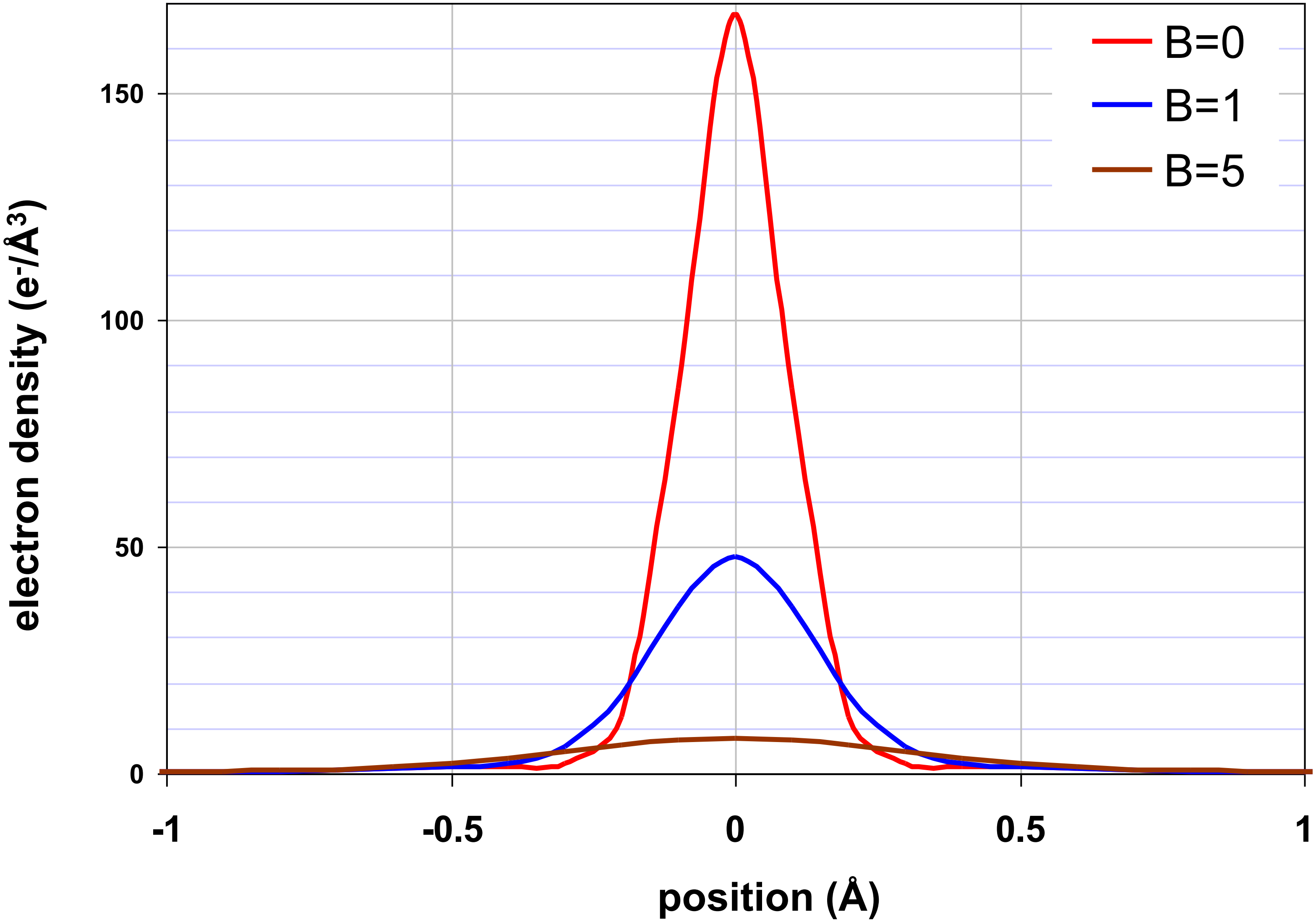
“B” factors



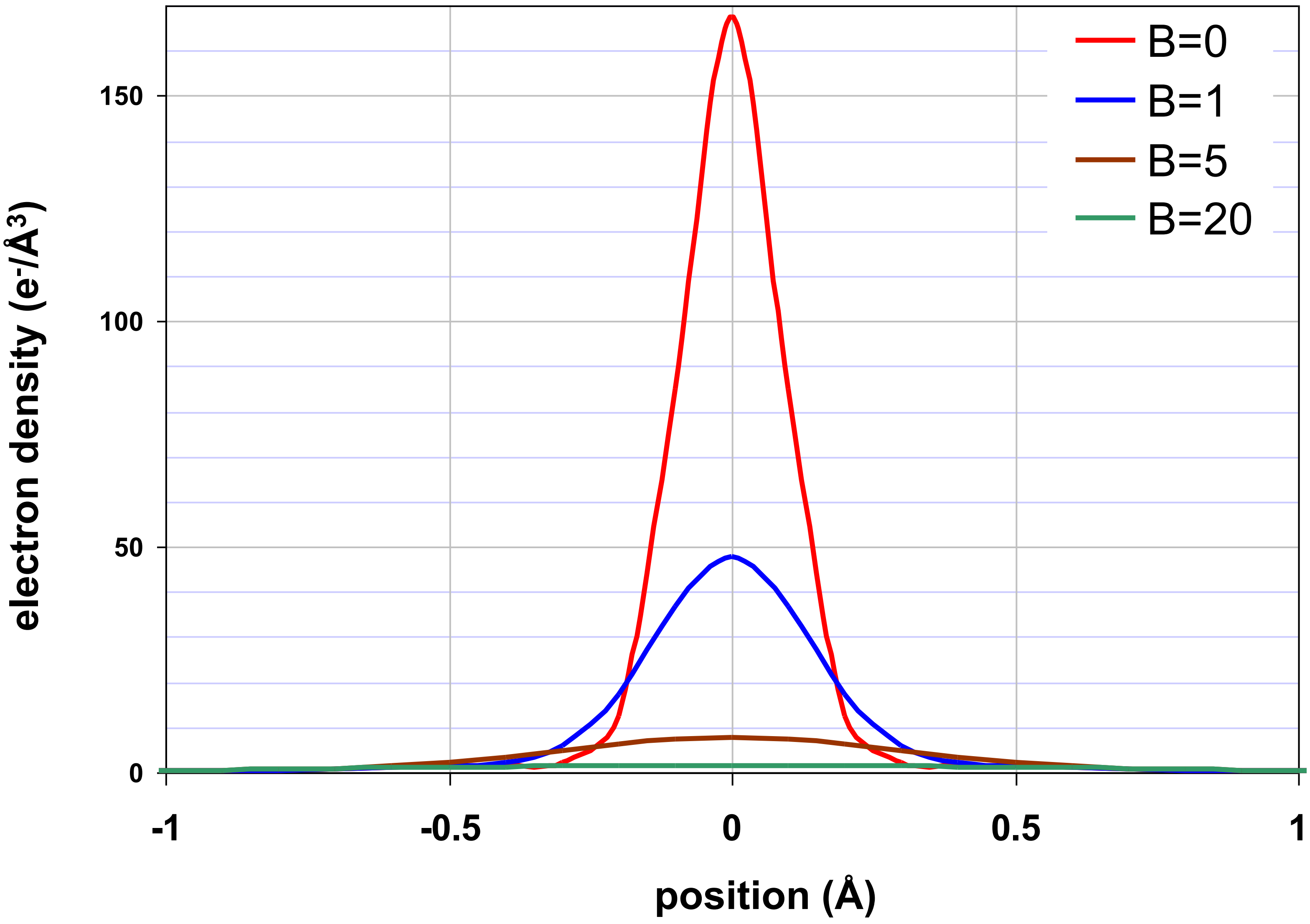
“B” factors



“B” factors

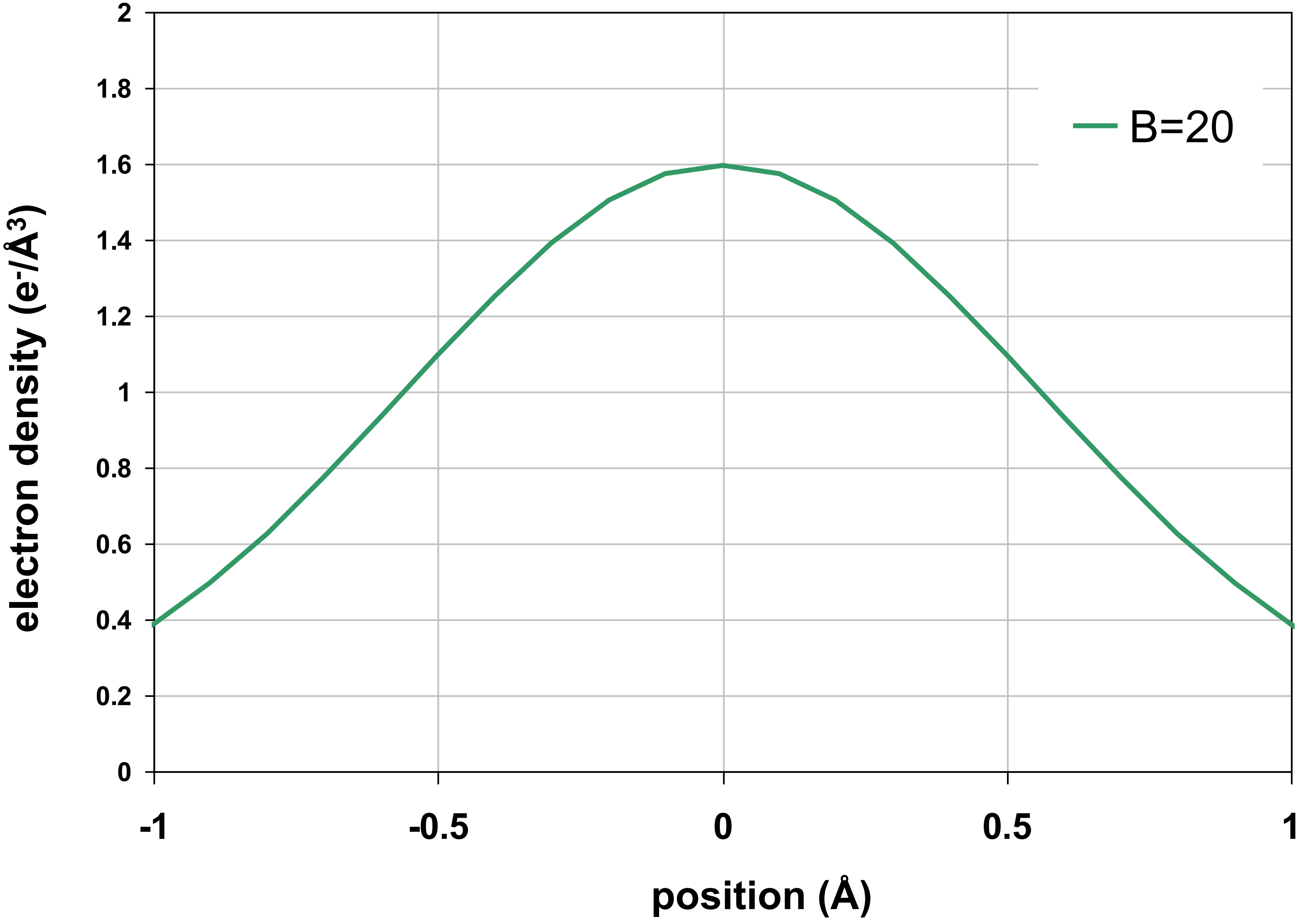


“B” factors

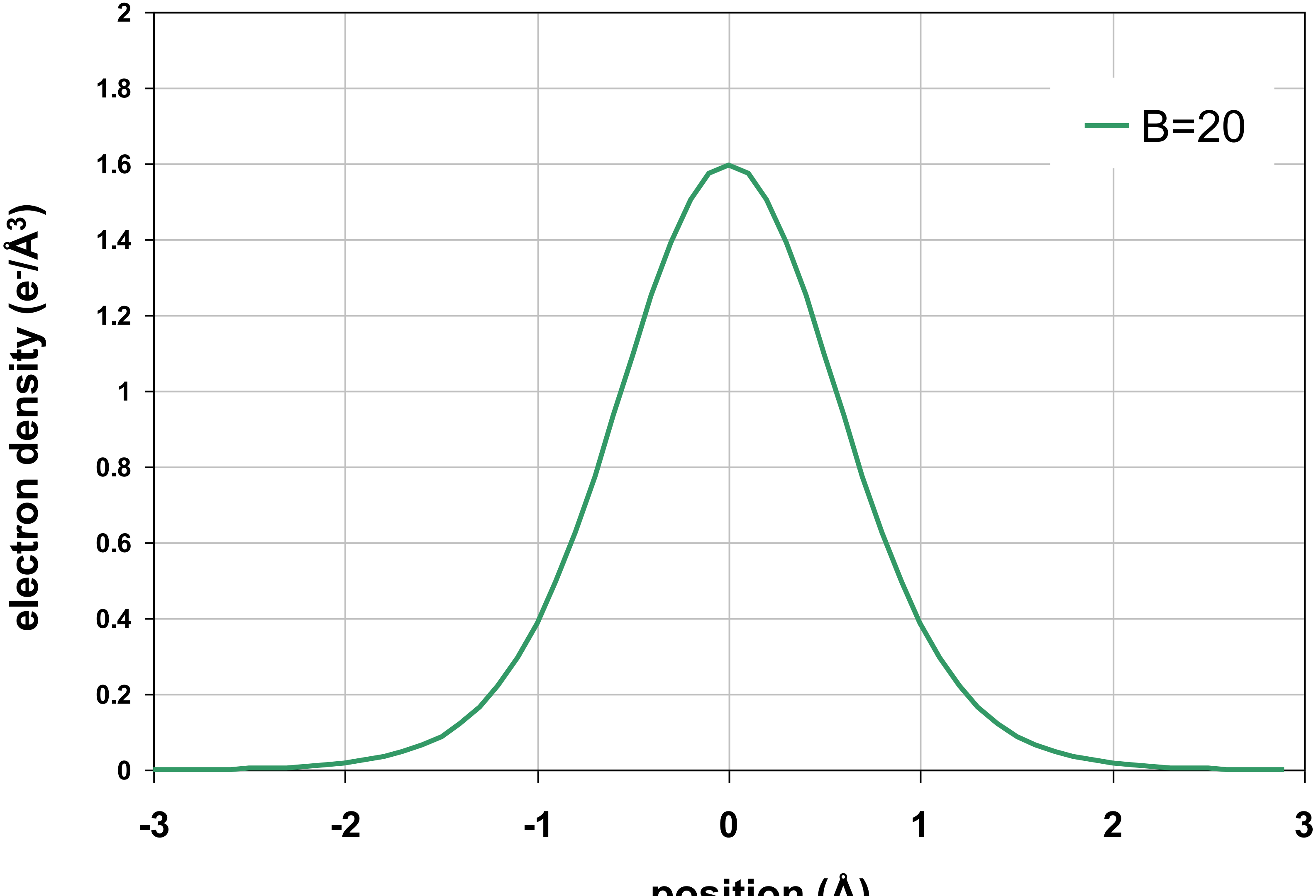


B=0
B=1
B=5
B=20
150
100
50
0
-1
-0.5
0
0.5
1
electron density (e-/Å^3)
position (Å)

“B” factors

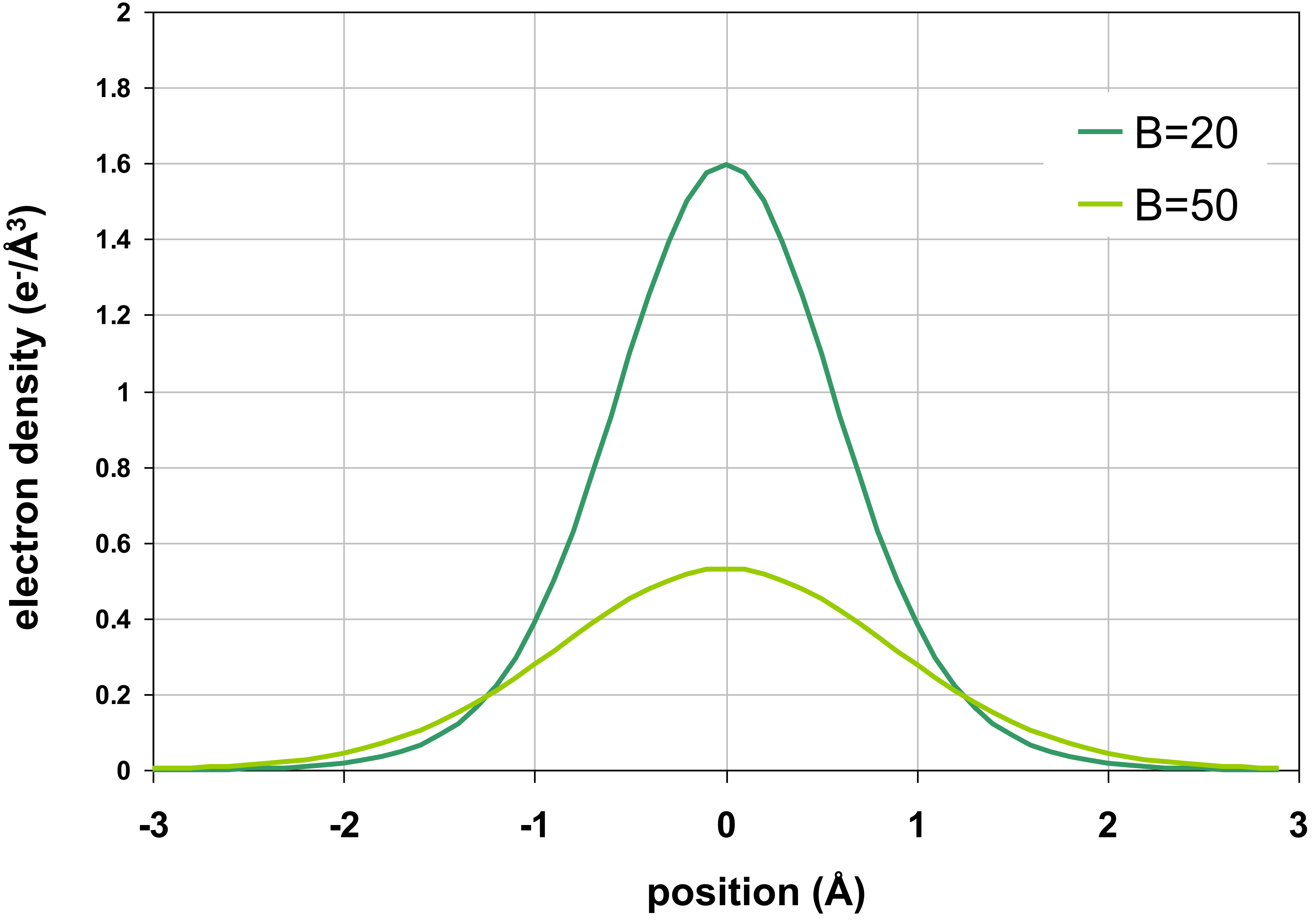

2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
electron density (e-/Å3)
-1
-0.5
0
0.5
1
position (Å)
B=20

“B” factors



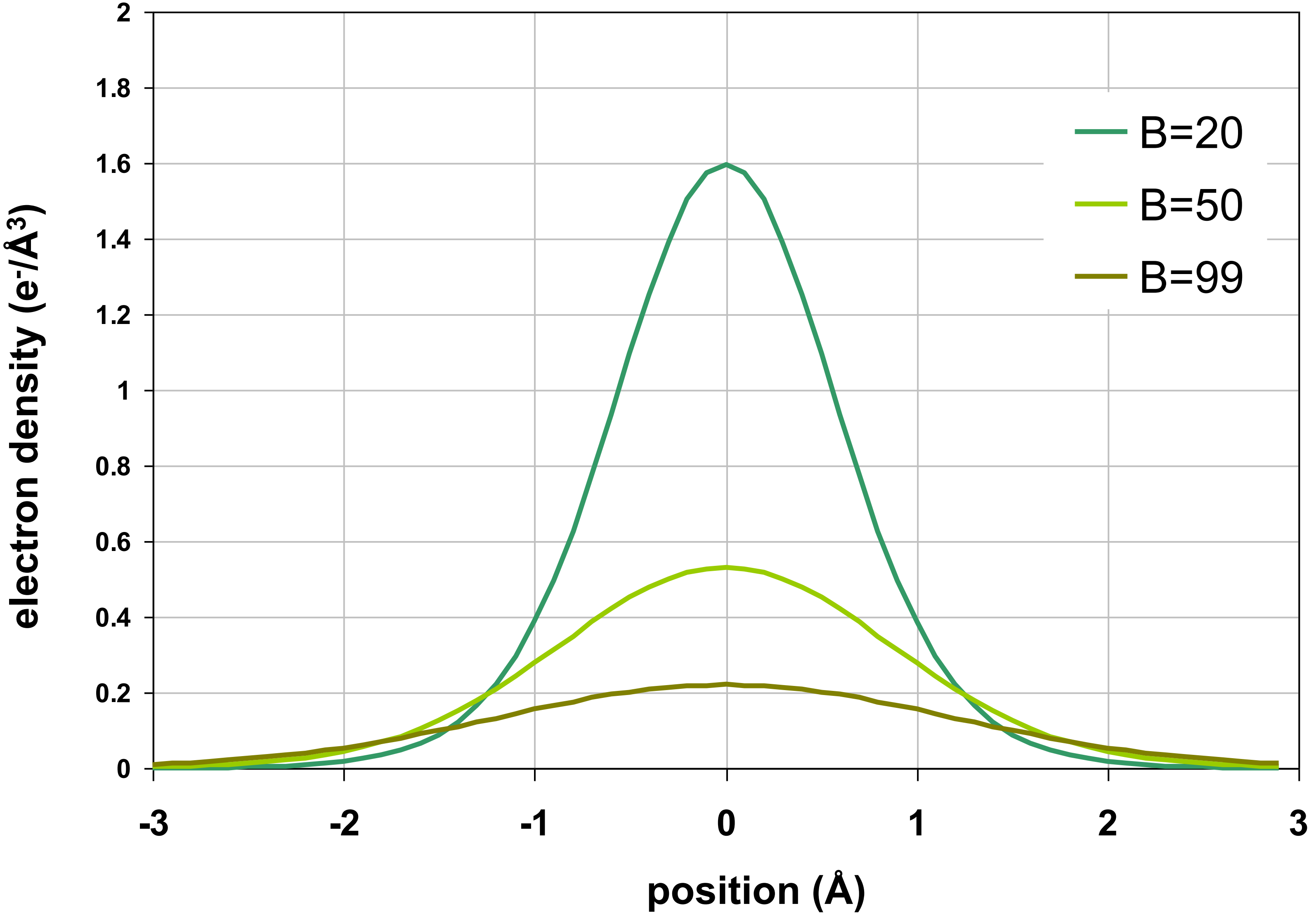
electron density (e⁻/Å³)
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
-3
-2
-1
0
1
2
3
position (Å)
B=20

“B” factors



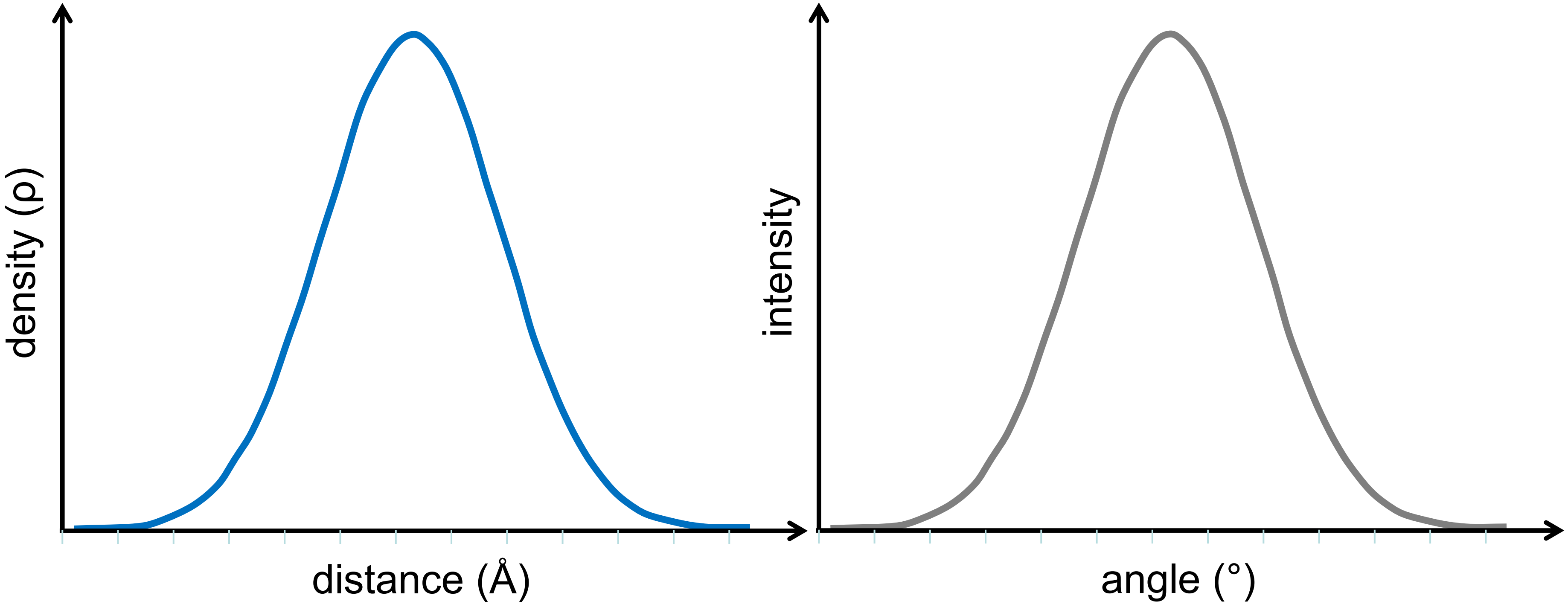
electron density (e⁻/Å³)
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
-3
-2
-1
0
1
2
3
position (Å)
B=20
B=50

“B” factors

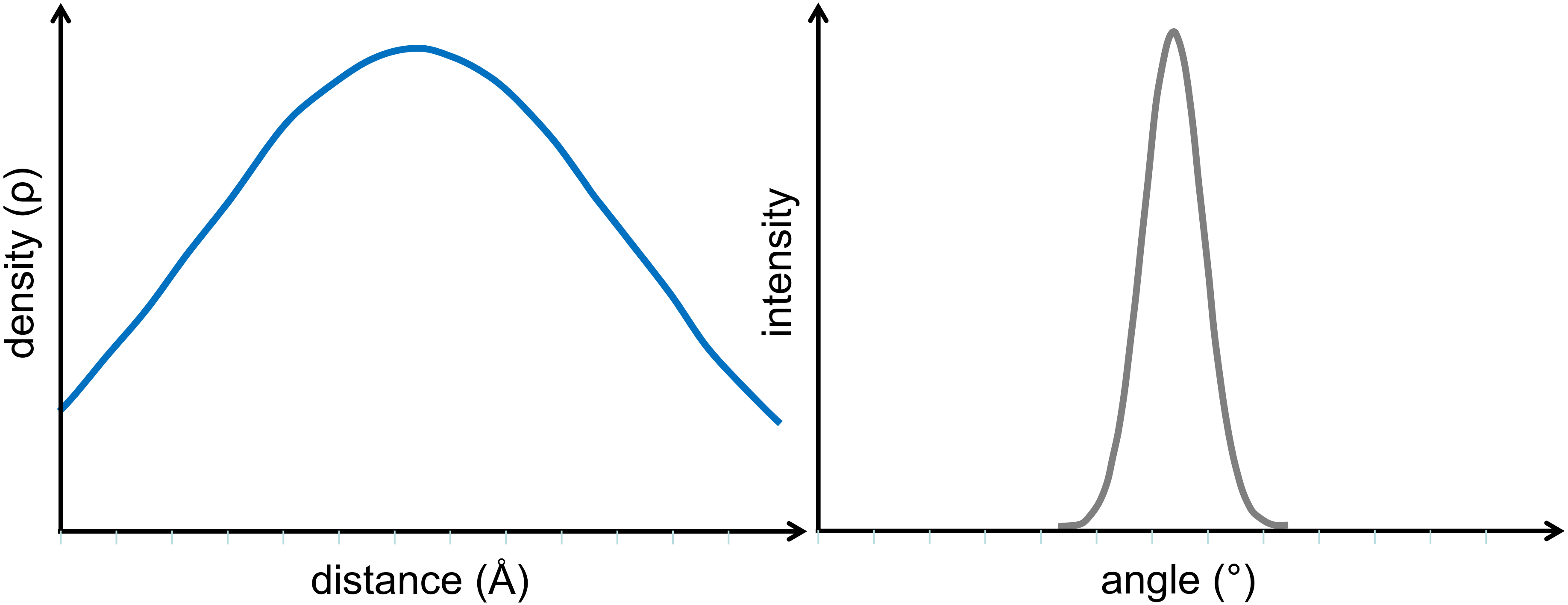


2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
electron density (e⁻/Å³)
-3
-2
-1
0
1
2
3
position (Å)
B=20
B=50
B=99

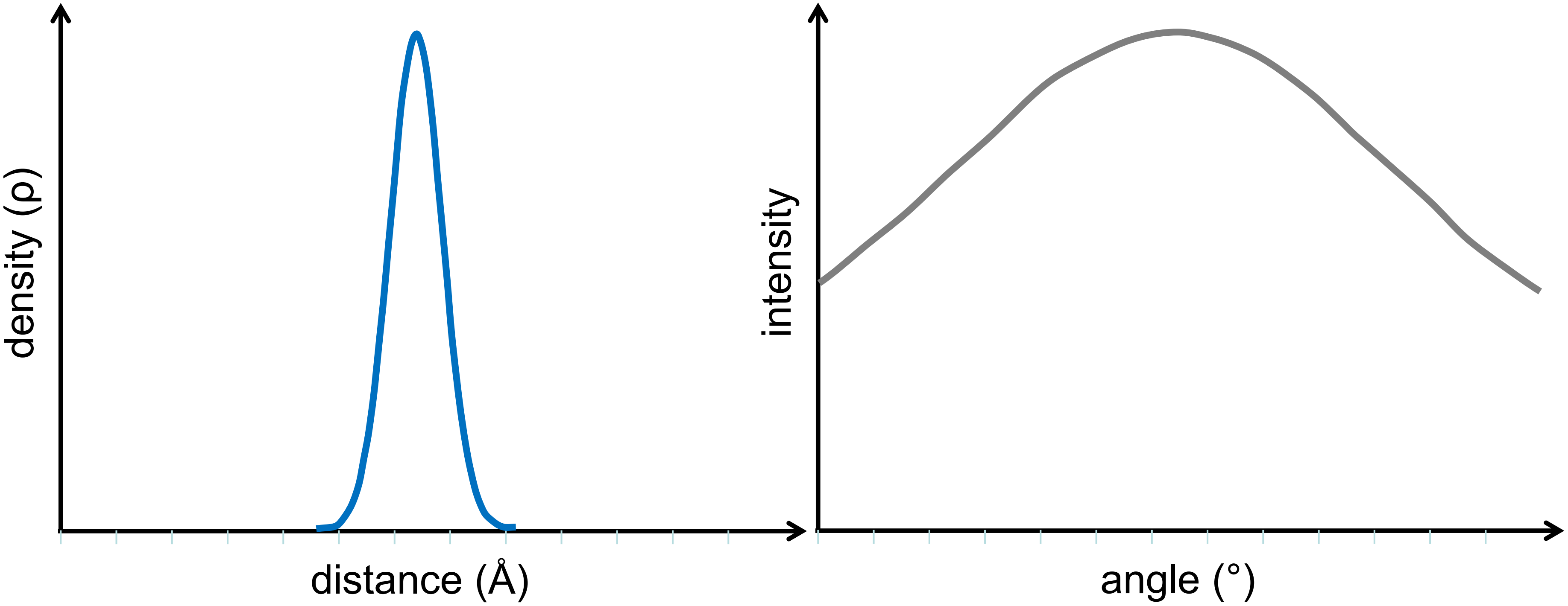
A Fourier view of scattering



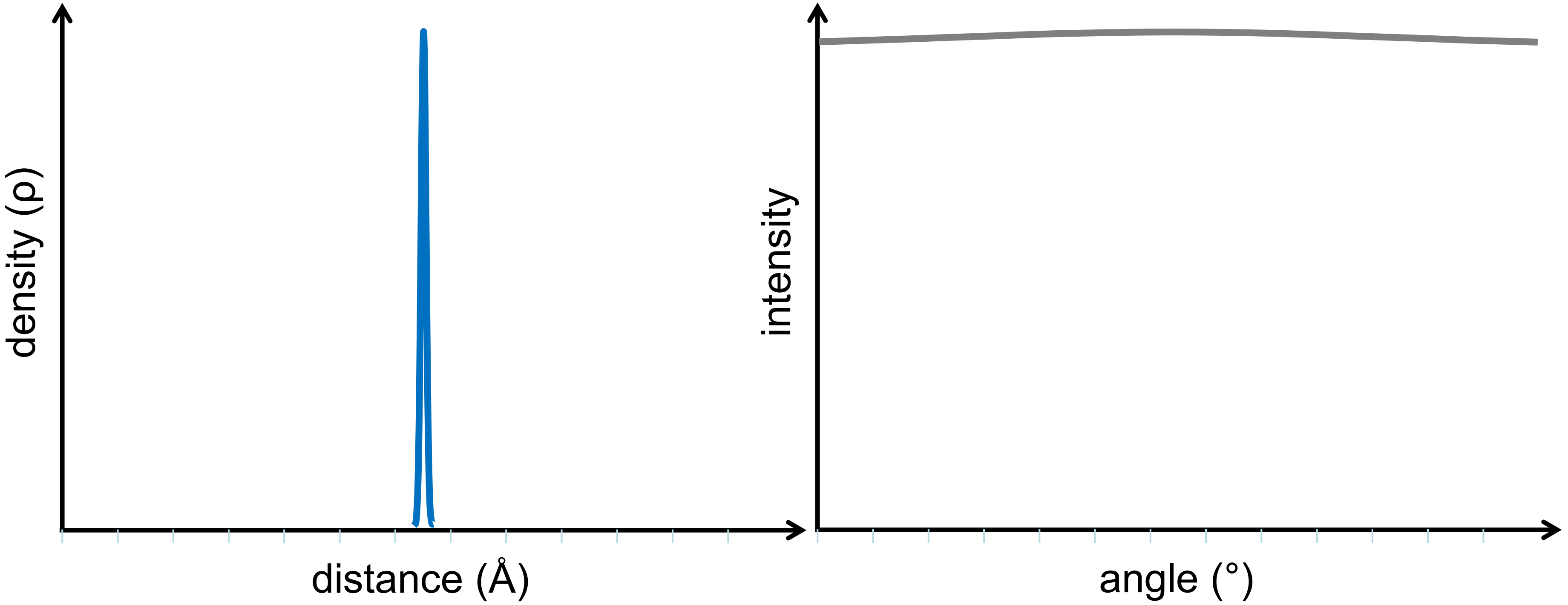
A Fourier view of scattering



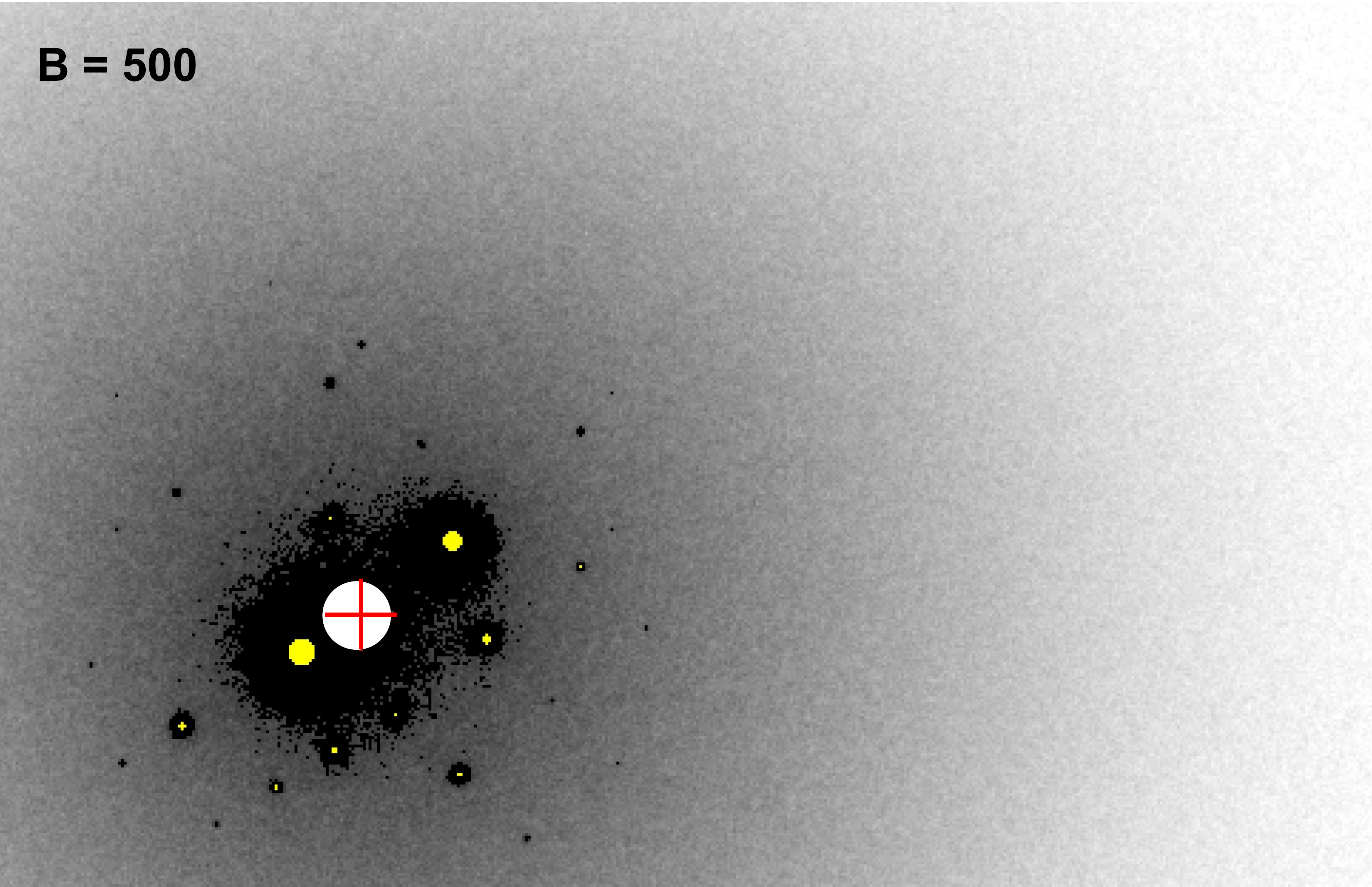
A Fourier view of scattering



A Fourier view of scattering



B factor from image analysis



B factor from image analysis