Rhombohedral Hexagonal Setting

Jon Cooper

Rhombohedral Hexagonal Setting

Jon Cooper (June - September 2021)

1 Background

I have quite often heard it said that structures which are rhombohedral are actually treated as being in a pseudo-hexagonal space group, called $H3$, *i.e.* it has 3-fold symmetry, rather than 6-fold which would be needed for it to be a proper hexagonal system. Thankfully, I have only had to work with one or two such proteins. Beyond seeing $H3$ as the space group in the data processing files or entering it as the space group when needed, I had very little idea what it meant. I had also heard it referred to as the obverse setting, that it was complicated, that it was done to allow the maps to be calculated by FFT (*fast Fourier transform*) and that you need triangular graph paper to understand it. But, that's about it!

I found an exercise on it in the book *Mathematical Crystallography* (Boisen and Gibbs, 1985, pp. 89-90) which got me interested because it has a figure which, although extremely helpful, was also extremely baffling and the legend, I believe, has a minor error in it.

The normal convention for rhombohedral is for the **a**, **b** and **c** vectors to be of equal magnitude and for the unit cell angles to be equal. Note that if the angles were all 90°, then having all cell dimensions equal, too, would normally indicate a cubic system. However, cubic space groups require additional symmetry and in principle there is nothing to stop a rhombohedral crystal having cell angles of 90°. With this in mind, at this point the best thing is to watch the following video by Frank Hoffmann (Hamburg):

https://youtu.be/Ks1TNkG6tXY

The video explains the system infinitely better than a thousand words ever could.

2 The rhombohedral lattice

As the video says, a rhombohedral cell is made by stretching or squashing a cube along one of its body-diagonals. This diagonal is then the axis of highest symmetry, which is a 3-fold. Since it is generally helpful to have the axis of

highest symmetry aligned with the crystallographic z axis, this body diagonal becomes the **c** unit cell vector of the pseudo-hexagonal cell, as shown in Fig. 1. Viewing the unit cell along this axis, the 3 lattice points one third of the way up (coloured pale green as in the video) appear to form an equilateral triangle and the 3 lattice points two thirds of the way up (pale blue as in the video) form another equilateral triangle that is rotated by 60 °around z relative to the first triangle.

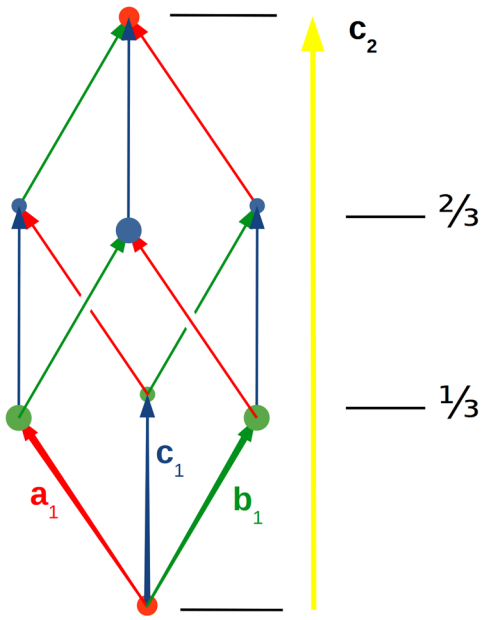

Figure 1: A rhombohedral unit cell with vectors $\mathbf{a}_1$, $\mathbf{b}_1$ and $\mathbf{c}_1$ drawn in red, green and blue, respectively, *i.e.* using the abbreviation RGB as an *aide memoire*. The vectors which are labelled are drawn with a perspective effect to indicate that $\mathbf{a}_1$ and $\mathbf{b}_1$ project towards the viewer whereas $\mathbf{c}_1$ points away. This is true for all the vectors shown although the remainder are just drawn as plane arrows. The unit cell vectors are all of equal amplitude, *i.e.* $a_1 = b_1 = c_1$ and they diverge from one another by an equal angle, *i.e.* $\alpha_1 = \beta_1 = \gamma_1$ (not shown). The yellow arrow indicates the direction in which the rhombohedral cell has been distorted from a cube. It is labelled $\mathbf{c}_2$ since it forms the **c** unit cell vector of the pseudo-hexagonal lattice. This vector is the body-diagonal of the rhombohedral cell and two layers of lattice points (shown pale green and blue with a perspective effect) divide it equally into thirds, as shown.

3 Making it hexagonal

If we imagine looking down the $\mathbf{c}_2$ unit cell vector in Fig. 1, the 3 green lattice points would from an equilateral triangle, as shown in Fig. 2. The $\mathbf{a}_1$, $\mathbf{b}_1$ and $\mathbf{c}_1$ appear to originate from the orange lattice point directly below the centroid of this triangle. This lattice point is also the origin of the pseudo-hexagonal cell which is shown in yellow along with the $\mathbf{a}_2$ and $\mathbf{b}_2$ unit cell vectors.

The hexagonal cell is centred, containing the pale green lattice point at the tip

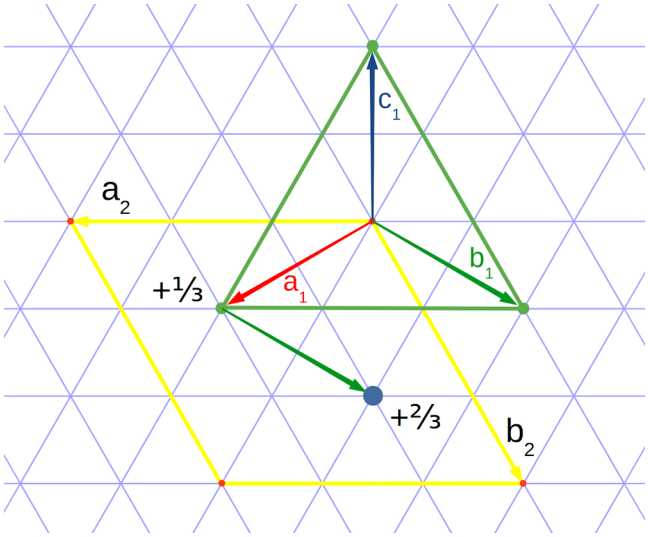

Figure 2: The rhombohedral and pseudo-hexagonal unit cell vectors viewed along the body-diagonal of the rhombohedral cell. This diagonal forms the 3-fold or $\mathbf{c}_2$ unit cell vector (not labelled) of the pseudo-hexagonal lattice. The rhombohedral unit cell vectors $\mathbf{a}_1$, $\mathbf{b}_1$ and $\mathbf{c}_1$ originate from an orange lattice point, as in Fig. 1, and point roughly towards the viewer. The choice of hexagonal cell is shown in yellow with the corresponding $\mathbf{a}_2$ and $\mathbf{b}_2$ unit cell vectors labelled. The hexagonal cell therefore possesses 3 lattice points: the small orange one at the origin, the pale green one at the tip of the $\mathbf{a}_1$ vector and the pale blue one formed by a subsequent translation along $\mathbf{b}_1$. As shown in Fig. 1, the pale green and pale blue lattice points occur at heights of $^1/_3$ $\mathbf{c}_2$ and $^2/_3$ $\mathbf{c}_2$, respectively.

of $\mathbf{a}_1$ and the pale blue lattice point, which is generated from the pale green one by application of the $\mathbf{b}_1$ unit cell vector. Therefore, in contrast to the rhombohedral cell, which is primitive, the pseudo-hexagonal cell contains 3 lattice points which lie along a body-diagonal. The volume of the pseudo-hexagonal cell is therefore 3 times that of the rhombohedral unit cell.

Anyhow, I am beginning to see why, in macromolecular crystallgraphy, space group $R3$ is usually referred to as $H3$ and $R32$ as $H32$. The notation $R3{:}h$ or $R32{:}h$ is also used to refer to the pseudo-hexagonal setting and $R3{:}r$ and $R32{:}r$ for the corresponding rhombohedral cell. As long as the unit cell is given (and

ideally the symmetry operators), there should be no ambiguity.

4 Obverse and reverse settings

The terms *obverse* and *reverse* refer to alternative choices of the hexagonal cell, in much the same way that they can refer to the alternative sides of a coin *i.e.* the 'heads' and 'tails' sides, respectively. The yellow hexagonal cell shown in Fig. 2 is in fact the obverse choice and we can see that for this setting the 3 lattice points within this cell occur at (0, 0, 0), (⅔, ⅓, ⅓) and (⅓, ⅔, ⅔) with respect to the $\mathbf{a}_2$, $\mathbf{b}_2$ and $\mathbf{c}_2$ axes.

However, there exists an alternative choice of the hexagonal cell which is obtained by rotating the yellow unit cell in Fig. 2 by 60° around the z axis, as shown in Fig. 3.

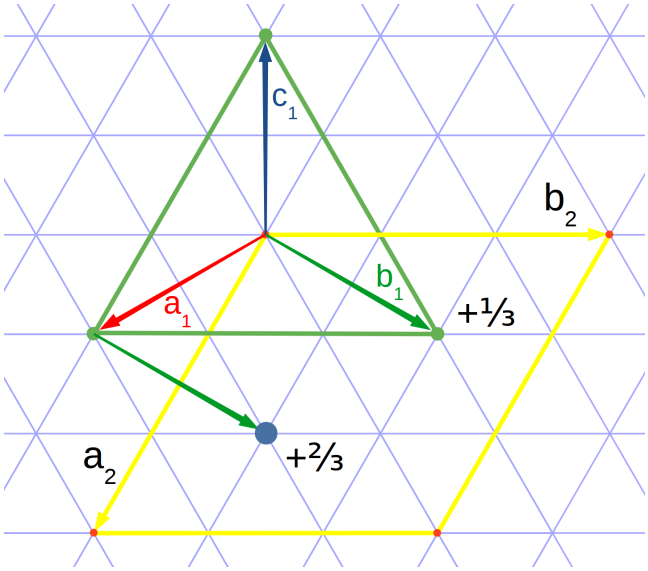

Figure 3: The reverse setting of the pseudo-hexagonal unit cell. As in Fig. 2 (which shows the obverse setting), we are looking down the body-diagonal of the rhombohedral cell such that its unit cell vectors $\mathbf{a}_1$, $\mathbf{b}_1$ and $\mathbf{c}_1$, point out of the page roughly towards the viewer. The alternative choice of hexagonal cell is shown in yellow with the corresponding unit cell vectors in the plane of the page, $\mathbf{a}_2$ and $\mathbf{b}_2$, labelled. The heights above the page of the pale green and pale blue lattice points, ⅓ $\mathbf{c}_2$ and ⅔ $\mathbf{c}_2$, are indicated.

The yellow hexagonal cell shown in Fig. 3 is the reverse choice and we can see that the 3 lattice points occur at (0, 0, 0), (⅔, ⅓, ⅔) and (⅓, ⅔, ⅓) with respect to the $\mathbf{a}_2$, $\mathbf{b}_2$ and $\mathbf{c}_2$ axes. This is where we differ from Boisen and Gibbs (*Mathematical Crystallography*, 1985, pp. 89-90), which I believe has a typo, since we are consistent with Wikipedia (the free encyclopedia) and the

International Tables for Crystallography.

Put simply, if we look first at the tip of the $\mathbf{a}_2$ vector in Fig. 2, we see that moving from the pale green centred lattice point to the pale blue one, we move away from the plane of the page in the obverse setting. In contrast, in the reverse setting, starting at the tip of $\mathbf{a}_2$, stepping from the pale blue lattice point to the pale green one, along the body diagonal, we move towards the plane of the page.

Note that the obverse setting is the one most commonly used *i.e.* if the setting is not stated, it can be assumed to be obverse. Note also that having a mixture of obverse and reverse lattices within a rhombohedral crystal is a common form of twinning known as *reticular merohedry* (see Andrea Thorne's slides from her 2011 IUCr Computing School presentation on Twinning).

5 Relating the two lattices

From Figs. 1 and 2, the following formulae given by Boisen and Gibbs (1985) for the obverse setting arise from vector addition:

$\mathbf{a}_2=\mathbf{a}_1-\mathbf{b}_1$
$\mathbf{b}_2=\mathbf{b}_1-\mathbf{c}_1$
$\mathbf{c}_2=\mathbf{a}_1+\mathbf{b}_1+\mathbf{c}_1$

and inspection of the triangular graph paper establishes that:

$\mathbf{a}_1= \frac{2}{3}\,\mathbf{a}_2 + \frac{1}{3}\,\mathbf{b}_2 + \frac{1}{3}\,\mathbf{c}_2$
$\mathbf{b}_1= -\frac{1}{3}\,\mathbf{a}_2 + \frac{1}{3}\,\mathbf{b}_2 + \frac{1}{3}\,\mathbf{c}_2$
$\mathbf{c}_1= -\frac{1}{3}\,\mathbf{a}_2 - \frac{2}{3}\,\mathbf{b}_2 + \frac{1}{3}\,\mathbf{c}_2$

Similarly, for the reverse setting:

$\mathbf{a}_2=\mathbf{a}_1-\mathbf{c}_1$
$\mathbf{b}_2=\mathbf{b}_1-\mathbf{a}_1$
$\mathbf{c}_2=\mathbf{a}_1+\mathbf{b}_1+\mathbf{c}_1$
$\mathbf{a}_1= \frac{1}{3}\,\mathbf{a}_2 - \frac{1}{3}\,\mathbf{b}_2 + \frac{1}{3}\,\mathbf{c}_2$
$\mathbf{b}_1= \frac{1}{3}\,\mathbf{a}_2 + \frac{2}{3}\,\mathbf{b}_2 + \frac{1}{3}\,\mathbf{c}_2$
$\mathbf{c}_1= -\frac{2}{3}\,\mathbf{a}_2 - \frac{1}{3}\,\mathbf{b}_2 + \frac{1}{3}\,\mathbf{c}_2$

6 Systematic absences

From the previous section we know that the additional lattice points for the obverse setting are at the following positions relative to the hexagonal axes:

$\mathbf{r}_1 = {}^2/_3\ \mathbf{a}_2 + {}^1/_3\ \mathbf{b}_2 + {}^1/_3\ \mathbf{c}_2$
$\mathbf{r}_2 = {}^1/_3\ \mathbf{a}_2 + {}^2/_3\ \mathbf{b}_2 + {}^2/_3\ \mathbf{c}_2$

If we consider the scattering from a molecule (or molecules) at a single lattice point, the diffraction amplitudes will be of the form:

$$F'_{hkl} = \sum_j f_j e^{2\pi i(hx_j + ky_j + lz_j)} \tag{1}$$

Hence, the total scattering from all three lattice points is given by:

$$F_{hkl} = F'(1 + e^{2\pi i(({}^2/_3)h + ({}^1/_3)k + ({}^1/_3)l)} + e^{2\pi i(({}^1/_3)h + ({}^2/_3)k + ({}^2/_3)l)}) \tag{2}$$

The systematic absences will, of course, arise when $F_{hkl} = 0$ and this can only happen when the terms making up the sum in brackets cancel out. This requires that the phases of the second and third terms are 120° and 240°, respectively, or -120° and -240°, respectively. In other words the three components of this sum form an equilateral triangle in an Argand diagram, as shown in Fig. 4. The triangle can be above or below the real axis, corresponding to the above two situations of the phases being either positive or negative. Of course, it is possible for multiples of 360° to be added to or subtracted from these phases and the effect would be exactly the same.

Given that 120° and 240° correspond to ${}^1/_3$ and ${}^2/_3$ of a full circle, for the upper triangle in Fig. 4, we can say that:

$$({}^2/_3)h + ({}^1/_3)k + ({}^1/_3)l = {}^1/_3 + p \tag{3}$$

and

$$({}^1/_3)h + ({}^2/_3)k + ({}^2/_3)l = {}^2/_3 + q \tag{4}$$

where p and q are integers. OK, so how can we solve these equations for h, k and l? One possibility might be to add them giving:

$$h + k + l = 1 + p + q \tag{5}$$

However, the right-hand side of this equation can be any integer (*e.g.* n) so we are left with:

$$h + k + l = n \tag{6}$$

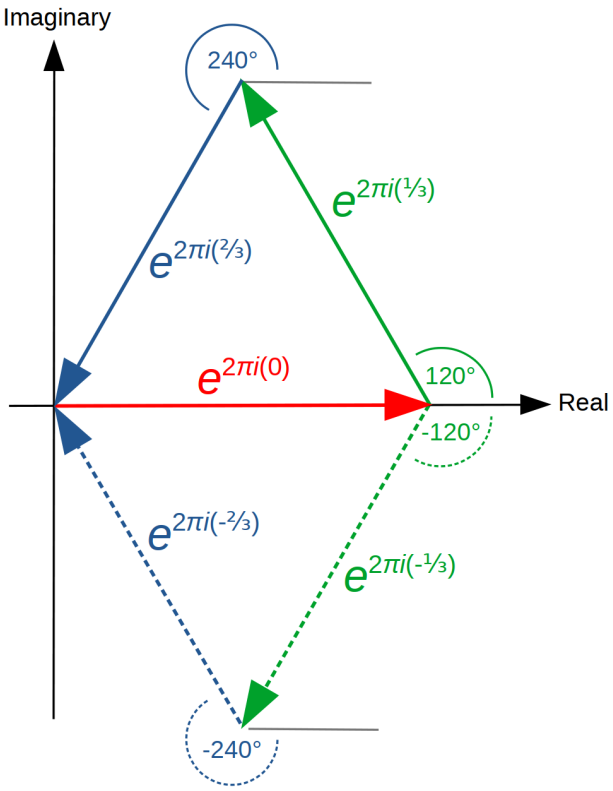

Figure 4: An Argand diagram showing how a systematic absence arises in the hexagonal setting. The three components of the bracketed term in the equation for F_{hkl} (equation 2) are shown. The horizontal red arrow of unit length represents scattering from the lattice point at the origin and the other two solid arrows show the additional scattering due to the centred lattice points at (2/3, 1/3, 1/3) in green and (1/3, 2/3, 2/3) in blue. The phases of the latter two components can be 120° and 240°, respectively, as shown by the upper equilateral triangle. It is also possible for the phases of these components to be -120° and -240°, respectively, and this is shown by the lower equilateral triangle with the arrows drawn dashed.

where n is any positive or negative integer. Hence, for a reflection to be a systematic absence, the symmetry of the pseudo-hexagonal setting does not place any restrictions on the value of $h + k + l$. If instead we subtract equation (4) from equation(3), we get the following:

$$(1/3)h - (1/3)k - (1/3)l = -1/3 + p - q \tag{7}$$

and multiplying by -3 gives:

$$-h + k + l = 1 + 3(q - p) \tag{8}$$

and since q-p can be any integer:

$$-h + k + l = 1 + 3n \tag{9}$$

Hence, the symmetry dictates that for a reflection to be systematically absent, its indices must obey the above rule *i.e.* there is a restriction on the value of $-h+k+l$ but not on the value of $h+k+l$. Interesting.

For the lower triangle in Fig. 4, we can say:

$$(2/3)h + (1/3)k + (1/3)l = -1/3 + p \tag{10}$$

and

$$(1/3)h + (2/3)k + (2/3)l = -2/3 + q \tag{11}$$

and subtracting these two equations gives:

$$(1/3)h - (1/3)k - (1/3)l = 1/3 + p - q \tag{12}$$

Multiplying by -3 again gives:

$$-h + k + l = -1 + 3(q - p) \tag{13}$$

Hence, another possible way that a reflection can be systematically absent is if its indices obey this rule:

$$-h + k + l = -1 + 3n \tag{14}$$

where n is any positive or negative integer. Combining the systematic absence rules for the upper and lower triangles in Fig. 4 gives us:

$$-h + k + l = 3n \pm 1 \tag{15}$$

as the general requirement for a reflection to systematically have zero amplitude in the pseudo-hexagonal setting. Hence, diffraction only occurs when the above rule is disobeyed, but how often does that happen? This can be seen by generating a short series of $3n\pm1$ numbers: 1, 2, 4, 5, 7, 8, 10, 11, 13, *etc*, which evidently includes every integer except those which are multiples of 3, *i.e.* $3n$, which has values of 0, 3, 6, 9, 12, *etc*. Hence, 2/3 of the reflections are systematically absent, which matches our expectations given that the hexagonal setting has two centred lattice points. The general arrangement of the systematic absences and presences is shown in Fig. 5.

It turns out that we can also simplify equation 2 for reflections which are systematically present. If we take the equation $-h+k+l=3n$ and add $3h$ to each side we get:

$$2h + k + l = 3(n + h) \tag{16}$$

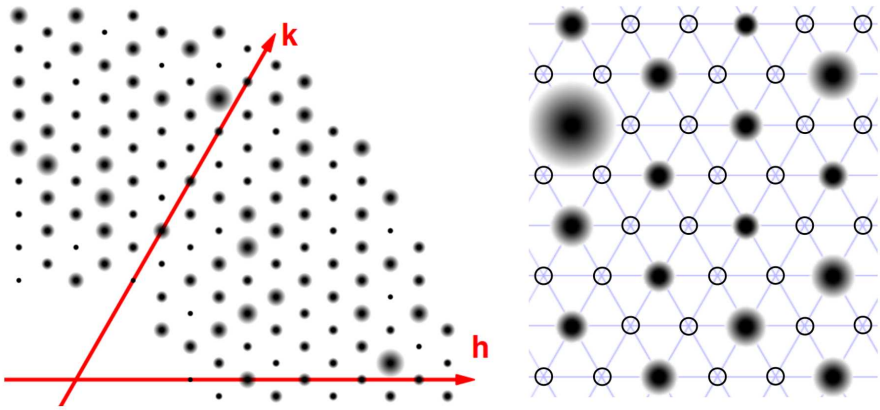

Figure 5: One layer of the diffraction pattern of a rhombohedral crystal indexed in the pseudo-hexagonal system. The left-hand picture shows the h and k axes, l being perpendicular to the page. The right-hand figure shows a zoomed-up region of the pattern with the underlying hexagonal symmetry of the reciprocal lattice drawn in pale grey lines. The hollow circles indicate the positions of systematic absences which outnumber the observed spots by a factor of 2.

Dividing by 3 we get part of the first exponential term of equation 2:

$$(2/3)h + (1/3)k + (1/3)l = n + h \tag{17}$$

which becomes:

$$e^{2\pi i(n+h)} = 1 \tag{18}$$

since any integral multiple of 2π is equivalent to a phase of zero.

If we take the equation $-h + k + l = 3n$, multiply through by -1 and add $3(k + l)$ we get: $h + 2k + 2l = 3(k + l - n)$ then dividing by 3 gives:

$$(1/3)h + (2/3)k + (2/3)l = k + l - n \tag{19}$$

This shows that the second exponential term of equation 2 also becomes 1 since $k + l - n$ is an integer giving a phase of zero to this component in the Argand diagram. Therefore, $F_{hkl} = 3F'_{hkl}$ when $-h + k + l = 3n$. Hence, whilst $2/3$ of the reflections are systematically absent in the hexagonal setting, the scattering power of all 3 lattice points in the unit cell goes into the $1/3$ of the reflections that are observable. As remarkable as the mathematics seems to be, this result is really only to be expected since the triply-centred hexagonal lattice is essentially just an artificial construction.

Now, for the reverse setting. In Section 4 we showed that the additional lattice points are at the following positions relative to the hexagonal axes:

$\mathbf{r}_1 = {}^1/_3\ \mathbf{a}_2 + {}^2/_3\ \mathbf{b}_2 + {}^1/_3\ \mathbf{c}_2$
$\mathbf{r}_2 = {}^2/_3\ \mathbf{a}_2 + {}^1/_3\ \mathbf{b}_2 + {}^2/_3\ \mathbf{c}_2$

from which we derive a structure factor equation:

$$F_{hkl} = F'(1 + e^{2\pi i(({}^1/_3)h + ({}^2/_3)k + ({}^1/_3)l)} + e^{2\pi i(({}^2/_3)h + ({}^1/_3)k + ({}^2/_3)l)}) \tag{20}$$

in which the two exponential terms must have phases of 120°and 240° in order to form the upper equilateral triangle in the Argand diagram (Fig. 4) and give zero diffraction amplitude. Hence:

$$({}^1/_3)h + ({}^2/_3)k + ({}^1/_3)l = {}^1/_3 + p \tag{21}$$

and

$$({}^2/_3)h + ({}^1/_3)k + ({}^2/_3)l = {}^2/_3 + q \tag{22}$$

Subtracting the first of these two equations from the second and multiplying by 3 gives us $h - k + l = 3n + 1$ and the same treatment for the lower triangle in Fig. 4 yields $h - k + l = 3n - 1$. Hence we can say the systematic absences will occur in the reverse setting when:

$$h - k + l = 3n \pm 1 \tag{23}$$

and reflections will be observable when:

$$h - k + l = 3n \tag{24}$$

Thus, it is possible to determine from the systematic absences whether the data have been indexed in the obverse or reverse setting.

This section was based entirely on J. W. Jeffery's book *Methods in X-ray Crystallography*, Academic Press, London (1971) *pp.* 343-4.

7 Notes on twinning

We touched on the subject of obverse/reverse twinning in rhombohedral crystals. This arises when regions or domains of the crystal differ in orientation in the same way as the two choices of pseudo-hexagonal cell shown earlier. This can happen when there is a pseudo 2-fold axis parallel either with the pseudo-hexagonal axis or with $\mathbf{a}_2$-$\mathbf{b}_2$ (Herbst-Irmer, R. & Sheldrick, G. M. (2002). *Acta Crystallogr. B* **58**, 477-481). With this type of twinning there are four types of reflection to think about. Firstly consider those with $-h + k + l = 3n$

and $h - k + l \neq 3n$. These reflections are systematically present for the obverse domain but are systematically absent for the reverse one, so only the obverse domain contributes to the diffraction. Then there are reflections with $-h + k + l \neq 3n$ and $h - k + l = 3n$ which are absent for the obverse domain but present for the reverse one. Reflections with $-h + k + l \neq 3n$ and $h - k + l \neq 3n$ are absent for both domains and reflections with $-h + k + l = 3n$ and $h - k + l = 3n$ have contributions from both. Quoting exactly from Herbst-Irmer & Sheldrick (2002): "Since only one third of the reflections (with $l = 3n$) are affected by the twinning, structure solution is normally not a severe problem, because two thirds of the reflections have contributions from only one domain and are often sufficient for structure solution." OK, so where does the $l = 3n$ come from? Adding together $-h + k + l = 3n$ and $h - k + l = 3n$ gives us $2l = 6n$ or $l = 3n$. Sorted. So the point is that only 1/3 of the reflections for an obverse-indexed dataset will obey this rule and will therefore be affected by the twinning, while the rest won't. Nice.

8 Worked example

The final part of the exercise in *Mathematical Crystallography* (Boisen and Gibbs, 1985, pp. 89-90) involves calculating the a_1, b_1, c_1, α_1, β_1 and γ_1 values of a rhombohedral unit cell given that $a_2 = b_2 = 15.951$ Å and c_2=7.24 Å.

Refering to Fig. 2, we can determine the projection of $\mathbf{a}_1$ onto the plane of $\mathbf{a}_2$ and $\mathbf{b}_2$, shown as OB in Fig. 6.

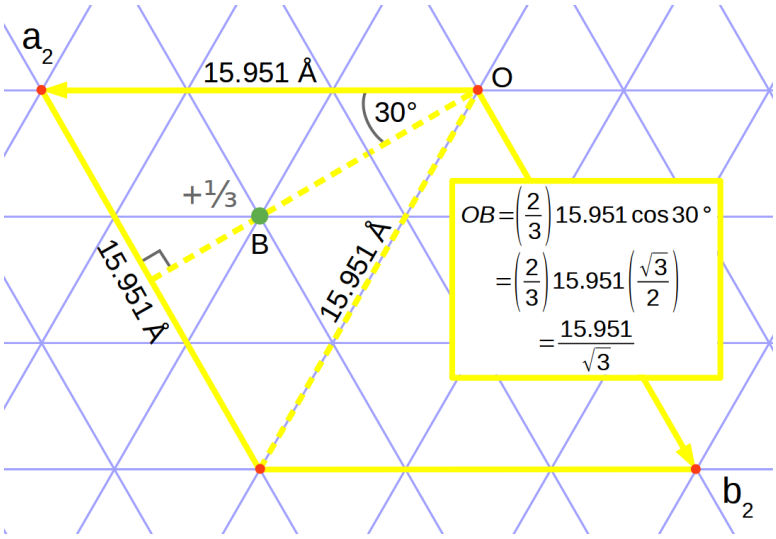

Figure 6: Deriving the length of the component of $\mathbf{a}_1$ in the horizontal plane.

The grey +1/3 symbol next to the pale green dot (B) in Fig. 6 indicates that

the tip of $\mathbf{a}_1$ vector is at a height of $(1/3)c_2$ above this point, as shown on the left hand side of Fig. 7. Basic trigonometry allows us to calculate a_1 (9.52 Å) and the angle shown as θ which is essentially the latitude of the point where $\mathbf{a}_1$ emerges from the sphere shown on the right of Fig. 7. The $\mathbf{b}_1$ vector emerges at the same latitude and the spherical triangle formed by these two points and the north pole allows γ_1 to be calculated.

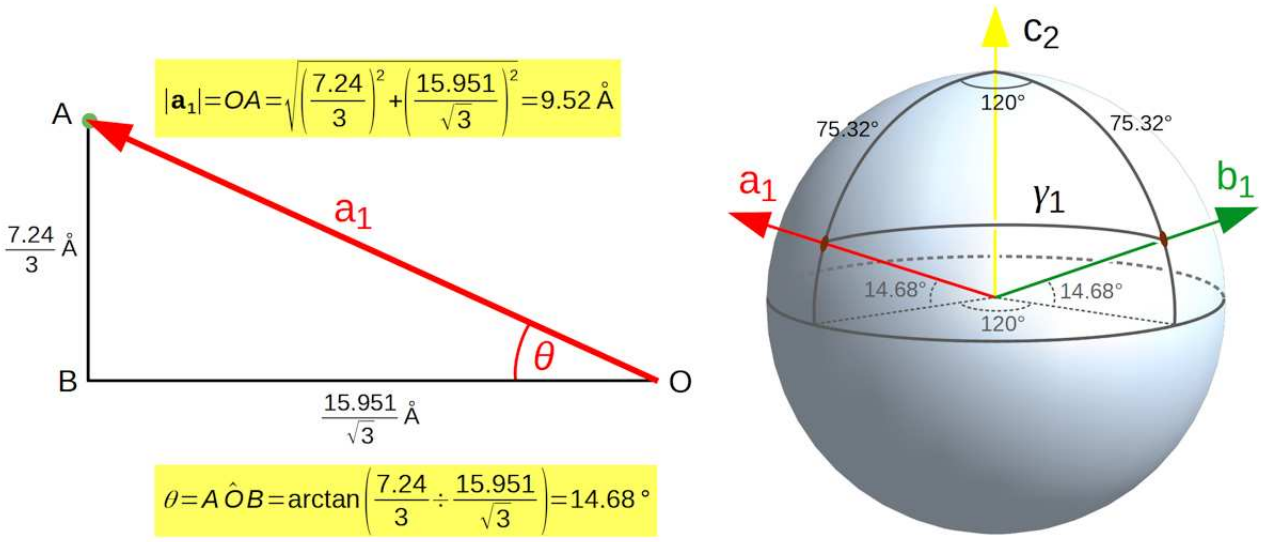

Figure 7: Deriving the amplitude of $\mathbf{a}_1$ and the angle it makes with the horizontal plane (θ) is shown on the left. The construction on the right shows how the γ_1 angle can be calculated using spherical trigonometry. The vectors $\mathbf{a}_1$ and $\mathbf{b}_1$ emerge from the sphere at the brown-coloured points which differ in longitude by 120° and are both at an angular distance of (90° - 14.68°) or 75.32° from the north pole, where the pseudo-hexagonal $\mathbf{c}_2$ axis emerges. All the arcs shown are great circles (*i.e.* their centres are at the centre of the sphere) so we can use the spherical cosine rule to calculate γ_1.

With reference to Fig. 8 the following equation is known as the spherical law of cosines:

$$\cos(a) = \cos(b)\cos(c) + \sin(b)\sin(c)\cos(A) \tag{25}$$

and we can identify $b = c = 75.32°$ and $A = 120°$. The unknown rhombohedral unit cell angle, γ_1 (or a in Fig. 8), can then be calculated as 113.8° which agrees with the answer in the book.

One final point, I don't think this is how the authors expect the calculation to be done because the preceding pages are about vector and matrix operations.

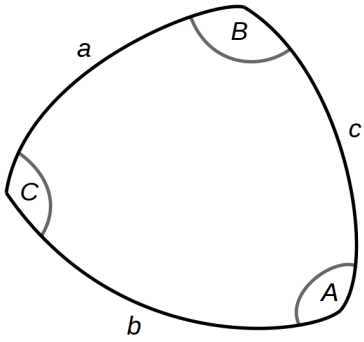

Figure 8: A general spherical triangle with sides a, b and c which are arcs formed by great circles. Note that a, b and c are actually the angles subtended by each arc at the centre of the sphere. A given arc defines a plane and the angles between these planes are shown in capital letters as A, B and C.

9 Bibliography

Boisen, M. B. & Gibbs, G. V. (1985). *Mathematical Crystallography*, (Ed. Ribbe, P. H.), *Reviews in Minerology*, **15**.

Herbst-Irmer, R. & Sheldrick, G. M. (2002). Refinement of obverse/reverse twins. *Acta Crystallogr. B* **58**, 477-481

Hoffmann, F. (2016). *The Fascination of Crystals and Symmetry.* https://www.youtube.com/c/FrankHoffmann1000

International Tables for Crystallography, Volume A. (2016). (Ed. M. I. Aroyo) IUCr.

Jeffery, J. W. (1971). *Methods in X-ray Crystallography*, Academic Press, London.

Thorne, A. (2011). Twinning. IUCr Computing School slides.

10 Acknowledgements

I am very grateful to Dr Ian Tickle (Global Phasing Ltd, Cambridge, UK) for commenting on this manuscript and explaining some key points.

Please send any further comments, corrections and suggestions to the author so I can correct any errors or omissions (e-mail: jbcooper@fastmail.net).

- Written with **Gummi** - the simple LaTeX editor -