1.6. Methods of space-group determination

U. Shmueli, H. D. Flack and J. C. H. Spence

1.6.1. Overview

This chapter describes and discusses several methods of symmetry determination of single-domain crystals. A detailed presentation of symmetry determination from diffraction data is given in Section 1.6.2.1, followed by a brief discussion of intensity statistics, ideal as well as non-ideal, with an application of the latter to real intensity data from a $P\overline{1}$ crystal structure in Section 1.6.2.2. Several methods of retrieving symmetry information from a solved crystal structure are then discussed (Section 1.6.2.3). This is followed by a discussion of chemical and physical restrictions on space-group symmetry (Section 1.6.2.4), including some aids in symmetry determination, and by a brief section on pitfalls in space-group determination (Section 1.6.2.5).

The following two sections deal with reflection conditions. Section 1.6.3 presents the theoretical background of conditions for possible general reflections and their corresponding derivation. A brief discussion of special reflection conditions is included. Section 1.6.4 presents an extensive tabulation of general reflection conditions and possible space groups.

Other methods of space-group determination are presented in Section 1.6.5. Section 1.6.5.1 deals with an account of methods of space-group determination based on resonant (also termed 'anomalous') scattering. Section 1.6.5.2 is a brief description of approaches to space-group determination in macromolecular crystallography. Section 1.6.5.3 deals with corresponding approaches in powder-diffraction methods.

The chapter concludes with a description and illustration of symmetry determination based on electron-diffraction methods (Section 1.6.6), and principally focuses on convergent-beam electron diffraction.

This chapter deals only with single crystals. A supplement (Flack, 2015) deals with twinned crystals and those displaying a specialized metric.

1.6.2. Symmetry determination from single-crystal studies

By U. Shmueli and H. D. Flack

1.6.2.1. Symmetry information from the diffraction pattern

The extraction of symmetry information from the diffraction pattern takes place in three stages.

In the first stage, the unit-cell dimensions are determined and analyzed in order to establish to which Bravais lattice the crystal belongs. A conventional choice of lattice basis (coordinate system) may then be chosen. The determination of the Bravais lattice[1] of the crystal is achieved by the process of cell reduction, in which the lattice is first described by a basis leading to a primitive unit cell, and then linear combinations of the unit-cell vectors are taken to reduce the metric tensor (and the cell dimensions) to a standard form. From the relationships amongst the elements of the metric tensor, one obtains the Bravais lattice, together with a conventional choice of the unit cell, with the aid of standard tables. A detailed description of cell reduction is given in Chapter 3.1 of this volume and in Part 9 of earlier editions (*e.g.* Burzlaff *et al.*, 2002). An alternative approach (Le Page, 1982) seeks the Bravais lattice directly from the cell dimensions by searching for all the twofold axes present. All these operations are automated in software. Regardless of the technique employed, at the end of the process one obtains an indication of the Bravais lattice and a unit cell in a conventional setting for the crystal system, primitive or centred as appropriate. These are usually good indications which, however, must be confirmed by an examination of the distribution of diffracted intensities as outlined below.

In the second stage, it is the point-group symmetry of the intensities of the Bragg reflections which is determined. We recall that the average reduced intensity of a pair of Friedel opposites (hkl and $\overline{hkl}$) is given by

$$\begin{aligned}|F_{\rm av}(\mathbf{h})|^2 &= \tfrac{1}{2}[|F(\mathbf{h})|^2 + |F(\overline{\mathbf{h}})|^2] \\ &= \sum_{i,j}[(f_i + f_i')(f_j + f_j') + f_i'' f_j'']\cos[2\pi\mathbf{h}(\mathbf{r}_i - \mathbf{r}_j)] \equiv A(\mathbf{h}),\end{aligned} \tag{1.6.2.1}$$

where the atomic scattering factor of atom j, taking into account resonant scattering, is given by

$$\mathbf{f}_j = f_j + f_j' + if_j'',$$

the wavelength-dependent components f_j' and f_j'' being the real and imaginary parts, respectively, of the contribution of atom j to the resonant scattering, $\mathbf{h}$ contains in the (row) matrix (1×3) the diffraction orders (hkl) and $\mathbf{r}_j$ contains in the (column) matrix (3×1) the coordinates (x_j, y_j, z_j) of atom j. The components of the $\mathbf{f}_j$ are assumed to contain implicitly the displacement parameters. Equation (1.6.2.1) can be found *e.g.* in Okaya & Pepinsky (1955), Rossmann & Arnold (2001) and Flack & Shmueli (2007). It follows from (1.6.2.1) that

$$|F_{\rm av}(\mathbf{h})|^2 = |F_{\rm av}(\overline{\mathbf{h}})|^2 \text{ or } A(\mathbf{h}) = A(\overline{\mathbf{h}}),$$

regardless of the contribution of resonant scattering. Hence the averaging introduces a centre of symmetry in the (averaged) diffraction pattern.[2] In fact, working with the average of Friedel opposites, one may determine the Laue group of the diffraction pattern by comparing the intensities of reflections which should be symmetry equivalent under each of the Laue groups. These are the 11 centrosymmetric point groups: $\overline{1}$, $2/m$, mmm, $4/m$, $4/mmm$, $\overline{3}$, $\overline{3}m$, $6/m$, $6/mmm$, $m\overline{3}$ and $m\overline{3}m$. For example, the reflections of which the intensities are to be compared for the Laue group $\overline{3}$ are: hkl, kil, ihl, $\overline{hkl}$, $\overline{kil}$ and $\overline{ihl}$, where $i = -h - k$. An extensive listing of the indices of symmetry-related reflections in all the point groups, including of course the Laue groups, is

[1] The Bravais lattice symbol consists of two characters. The first is the first letter of the name of a crystal family and the second is the centring mode of a conventional unit cell. For details see Tables 3.1.2.1 and 3.1.2.2.

[2] We must mention the well known Friedel's law, which states that $|F(\mathbf{h})|^2 = |F(\overline{\mathbf{h}})|^2$ and which is only a reasonable approximation for noncentrosymmetric crystals if resonant scattering is negligibly small. This law holds well for centrosymmetric crystals, independently of the resonant-scattering contribution.

given in Appendix 1.4.4 of *International Tables for Crystallography* Volume B (Shmueli, 2008).[3] In the past, one used to inspect the diffraction images to see which classes of reflections are symmetry equivalent within experimental and other uncertainty. Nowadays, the whole intensity data set is analyzed by software. The intensities are merged and averaged under each of the 11 Laue groups in various settings (*e.g.* 2/*m* unique axis *b* and unique axis *c*) and orientations (*e.g.* $\bar{3}m1$ and $\bar{3}1m$). For each choice of Laue group and its variant, an $R_{\rm merge}$ factor is calculated as follows:

$$R_{{\rm merge},i} = \frac{\sum_{\mathbf{h}} \sum_{s=1}^{|G|_i} |\langle |F_{\rm av}(\mathbf{h})|^2 \rangle_i - |F_{\rm av}(\mathbf{h}\boldsymbol{W}_{si})|^2|}{|G|_i \sum_{\mathbf{h}} \langle |F_{\rm av}(\mathbf{h})|^2 \rangle_i}, \qquad (1.6.2.2)$$

where $\boldsymbol{W}_{si}$ is the matrix of the *s*th symmetry operation of the *i*th Laue group, $|G_i|$ is the number of symmetry operations in that group, the average in the first term in the numerator and in the denominator ranges over the intensities of the trial Laue group and the outer summations $\sum_{\mathbf{h}}$ range over the *hkl* reflections. Choices with low $R_{{\rm merge},i}$ display the chosen symmetry, whereas for those with high $R_{{\rm merge},i}$ the symmetry is inappropriate. The Laue group of highest symmetry with a low $R_{{\rm merge},i}$ is considered the best indication of the Laue group. Several variants of the above procedure exist in the available software. Whichever of them is used, it is important for the discrimination of the averaging process to choose a strategy of data collection such that the intensities of the greatest possible number of Bragg reflections are measured. In practice, validation of symmetry can often be carried out with a few initial images and the data-collection strategy may be based on this assignment.

In the third stage, the intensities of the Bragg reflections are studied to identify the conditions for systematic absences. Some space groups give rise to zero intensity for certain classes of reflections. These 'zeros' occur in a systematic manner and are commonly called systematic absences (*e.g.* in the *h*0*l* class of reflections, if all rows with *l* odd are absent, then the corresponding reflection condition is *h*0*l*: $l = 2n$). In practice, as implemented in software, statistics are produced on the intensity observations of all possible sets of 'reflections conditions' as given in Chapter 2.3 (*e.g.* in the example above, *h*0*l* reflections are separated into sets with $l = 2n$ and those with $l = 2n + 1$). In one approach, the number of observations in each set having an intensity (I) greater than n standard uncertainties [$u(I)$] [*i.e.* $I/u(I) > n$] is displayed for various values of n. Clearly, if a trial condition for systematic absence has observations with strong or medium intensity [*i.e.* $I/u(I) > 3$], the systematic-absence condition is not fulfilled (*i.e.* the reflections are not systematically absent). If there are no such observations, the condition for systematic absence may be valid and the statistics for smaller values of n need then to be examined. These are more problematic to evaluate, as the set of reflections under examination may have many weak reflections due to structural effects of the crystal or to perturbations of the measurements by other systematic effects. An alternative approach to examining numbers of observations is to compare the mean value, $\langle I/u(I) \rangle$, taken over reflections obeying or not a trial reflection condition. For a valid reflection condition, one expects the former value to be considerably larger than the latter. In Section 3.1 of Palatinus & van der Lee (2008), real examples of marginal cases are described.

[3] The tables in Appendix 1.4.4 mentioned above actually deal with space groups in reciprocal space; however, the left part of any entry is just the indices of a reflection generated by the point-group operation corresponding to this entry.

Table 1.6.2.1
The ability of the procedures described in Sections 1.6.2.1 and 1.6.5.1 to distinguish between space groups

The columns of the table show the number of sets of space groups that are indistinguishable by the chosen technique, according to the number of space groups in the set, *e.g.* for Laue-class discrimination, 85 space groups may be uniquely identified, whereas there are 8 sets containing 5 space groups indistinguishable by this technique. The tables in Section 1.6.4 contain 416 different settings of space groups generated from the 230 space-group types.

	No. of space groups in set that are indistinguishable by procedure used					
	1	2	3	4	5	6
No. of sets for Laue-class discrimination	85	78	43	0	8	1
No. of sets for point-group discrimination	390	13	0	0	0	0

The third stage continues by noting that the systematic absences are characteristic of the space group of the crystal, although some sets of space groups have identical reflection conditions. In Chapter 2.3 one finds all the reflection conditions listed individually for the 230 space groups. For practical use in space-group determination, tables have been set up that present a list of all those space groups that are characterized by a given set of reflection conditions. The tables for all the Bravais lattices and Laue groups are given in Section 1.6.4 of this chapter. So, once the reflection conditions have been determined, all compatible space groups can be identified from the tables. Table 1.6.2.1 shows that 85 space groups may be unequivocally determined by the procedures defined in this section based on the identification of the Laue group. For other sets of reflection conditions, there are a larger number of compatible space groups, attaining the value of 6 in one case. It is appropriate at this point to anticipate the results presented in Section 1.6.5.1, which exploit the resonant-scattering contribution to the diffracted intensities and under appropriate conditions allow not only the Laue group but also the point group of the crystal to be identified. If such is the case, the last line of Table 1.6.2.1 shows that almost all space groups can be unequivocally determined. In the remaining 13 pairs of space groups, constituting 26 space groups in all, there are the 11 enantiomorphic pairs of space groups [($P4_1$–$P4_3$), ($P4_122$–$P4_322$), ($P4_12_12$–$P4_32_12$), ($P3_1$–$P3_2$), ($P3_121$–$P3_221$), ($P3_112$–$P3_212$), ($P6_1$–$P6_5$), ($P6_2$–$P6_4$), ($P6_122$–$P6_522$), ($P6_222$–$P6_422$) and ($P4_132$–$P4_332$)] and the two exceptional pairs of $I222$ & $I2_12_12_1$ and $I23$ & $I2_13$, characterized by having the same symmetry elements in a different arrangement in space. These 13 pairs of space groups cannot be distinguished by the methods described in Sections 1.6.2 and 1.6.5.1, but may be distinguished when a reliable atomic structural model of the crystal has been obtained. On the other hand, all these 13 pairs of space groups can be distinguished by the methods described in Section 1.6.6 and in detail in Saitoh *et al.* (2001). It should be pointed out in connection with this third stage that a possible weakness of the analysis of systematic absences for crystals with small unit-cell dimensions is that there may be a small number of axial reflections capable of being systematically absent.

It goes without saying that the selected space groups must be compatible with the Bravais lattice determined in stage 1, with the Laue class determined in stage 2 and with the set of space-group absences determined in stage 3.

We thank L. Palatinus (2011) for having drawn our attention to the unexploited potential of the Patterson function for the determination of the space group of the crystal. The discovery of

this method is due to Buerger (1946) and later obtained only a one-sentence reference by Rogers (1950) and by Rossmann & Arnold (2001). The method is based on the observation that interatomic vectors between symmetry-related (other than by inversion in a point) atoms cause peaks to accumulate in the corresponding Harker sections and lines of the Patterson function. It is thus only necessary to find the location of those Harker sections and lines that have a high concentration of peaks to identify the corresponding symmetry operations of the space group. At the time of its discovery, it was not considered an economic method of space-group determination due to the labour involved in calculating the Patterson function. Subsequently it was completely neglected and there are no recent reports of its use. It is thus not possible to report on its strengths and weaknesses in practical modern-day applications.

1.6.2.2. Structure-factor statistics and crystal symmetry

Most structure-solving software packages contain a section dedicated to several probabilistic methods based on the Wilson (1949) paper on the probability distribution of structure-factor magnitudes. These statistics sometimes correctly indicate whether the intensity data set was collected from a centrosymmetric or noncentrosymmetric crystal. However, not infrequently these indications are erroneous. The reasons for this may be many, but outstandingly important are (i) the presence of a few very heavy atoms amongst a host of lighter ones, and (ii) a very small number of nearly equal atoms. Omission of weak reflections from the data set also contributes to failures of Wilson (1949) statistics. These erroneous indications are also rather strongly space-group dependent.

The well known probability density functions (hereafter p.d.f.'s) of the magnitude of the normalized structure factor E, also known as ideal p.d.f.'s, are

$$p(|E|) = \begin{cases} \sqrt{2/\pi}\exp\left(-|E|^2/2\right) & \text{for } P\overline{1} \\ 2|E|\exp(-|E|^2) & \text{for } P1 \end{cases}, \qquad (1.6.2.3)$$

where it is assumed that all the atoms are of the same chemical element. Let us see their graphical representations.

It is seen from Fig. 1.6.2.1 that the two p.d.f.'s are significantly different, but usually they are not presented as such by the software. What is usually shown are the cumulative distributions of $|E|^2$, the moments: $\langle|E|^n\rangle$ for $n = 1, 2, 3, 4, 5, 6$, and the averages of low powers of $|E^2 - 1|$ for ideal centric and acentric distributions, based on equation (1.6.2.3). Table 1.6.2.2 shows the numerical values of several low-order moments of $|E|$ and that of the lowest power of $|E^2 - 1|$. The higher the value of n the greater is the difference between their values for centric and acentric cases. However, it is most important to remember that the influence of measurement uncertainties also increases with n and therefore the higher the moment the less reliable it tends to be.

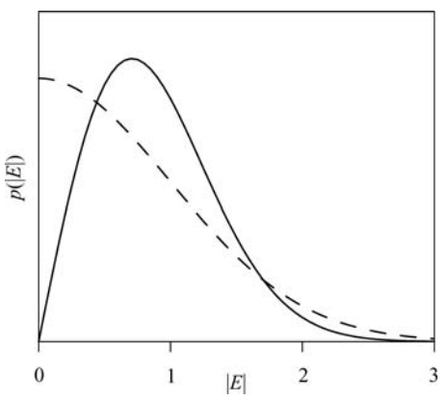

Figure 1.6.2.1
Ideal p.d.f.'s for the equal-atom case. The dashed line is the centric, and the solid line the acentric ideal p.d.f.

Table 1.6.2.2
The numerical values of several low-order moments of $|E|$, based on equation (1.6.2.3)

Moment	$P\overline{1}$	$P1$
$\langle\|E\|\rangle$	0.798	0.886
$\langle\|E\|^2\rangle$	1.000	1.000
$\langle\|E\|^3\rangle$	1.596	1.329
$\langle\|E\|^4\rangle$	3.000	2.000
$\langle\|E\|^5\rangle$	6.383	3.323
$\langle\|E\|^6\rangle$	15.000	6.000
$\langle\|E^2 - 1\|\rangle$	0.968	0.736

There are several ideal indicators of the status of centrosymmetry of a crystal structure. The most frequently used are: (i) the $N(z)$ test (Howells *et al.*, 1950), a cumulative distribution of $z = |E|^2$, based on equation (1.6.2.3), and (ii) the low-order moments of $|E|$, also based on equation (1.6.2.3). Equation (1.6.2.3), however, is very seldom used as an indicator of the status of centrosymmetry of a crystal stucture.

Let us now briefly consider p.d.f.'s that are valid for any atomic composition as well as any space-group symmetry, and exemplify their performance by comparing a histogram derived from observed intensities from a $P\overline{1}$ structure with theoretical p.d.f.'s for the space groups $P1$ and $P\overline{1}$. The p.d.f.'s considered presume that all the atoms are in general positions and that the reflections considered are general (see, *e.g.*, Section 1.6.3). A general treatment of the problem is given in the literature and summarized in the book *Introduction to Crystallographic Statistics* (Shmueli & Weiss, 1995).

The basics of the exact p.d.f.'s are conveniently illustrated in the following. The normalized structure factor for the space group $P\overline{1}$, assuming that all the atoms occupy general positions and resonant scattering is neglected, is given by

$$E(\mathbf{h}) = 2\sum_{j=1}^{N/2} n_j \cos(2\pi\mathbf{h}\boldsymbol{r}_j),$$

where n_j is the normalized scattering factor. The maximum possible value of E is $E_{\max} = \sum_{j=1}^{N} n_j$ and the minimum possible value of E is $-E_{\max}$. Therefore, $E(\mathbf{h})$ must be confined to the $(-E_{\max}, E_{\max})$ range. The probability of finding E outside this range is of course zero. Such a probability density function can be expanded in a Fourier series within this range (*cf.* Shmueli *et al.*, 1984). This is the basis of the derivation, the details of which are well documented (*e.g.* Shmueli *et al.*, 1984; Shmueli & Weiss, 1995; Shmueli, 2007). Exact p.d.f.'s for any centrosymmetric space group have the form

$$p(|E|) = \alpha\left\{1 + 2\sum_{m=1}^{\infty} C_m \cos(\pi m|E|\alpha)\right\}, \qquad (1.6.2.4)$$

where $\alpha = 1/E_{\max}$, and exact p.d.f.'s for any noncentrosymmetric space group can be computed as the double Fourier series

$$p(|E|) = \tfrac{1}{2}\pi\alpha^2|E|\sum_{m=1}^{\infty}\sum_{n=1}^{\infty} C_{mn} J_0[\pi\alpha|E|(m^2 + n^2)^{1/2}], \qquad (1.6.2.5)$$

where $J_0(X)$ is a Bessel function of the first kind and of order zero. Expressions for the coefficients C_m and C_{mn} are given by

references