

1.6. METHODS OF SPACE-GROUP DETERMINATION

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

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

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

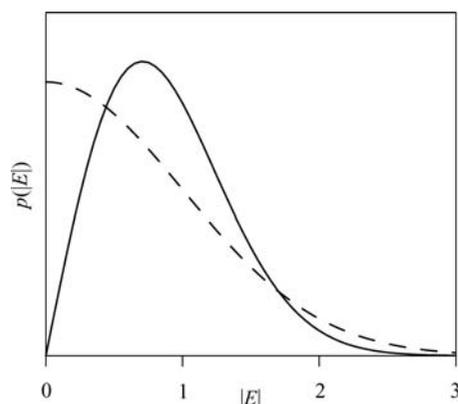


Figure 1.6.2.1
Ideal p.d.f.'s for the equal-atom case. The dashed line is the centric, and the solid line the acentric ideal p.d.f.

Table 1.6.2.2

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

Moment	$P\bar{1}$	$P1$
$\langle E \rangle$	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

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.

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 structure.

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 $P\bar{1}$ structure with theoretical p.d.f.'s for the space groups $P1$ and $P\bar{1}$. 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 $P\bar{1}$, assuming that all the atoms occupy general positions and resonant scattering is neglected, is given by

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

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

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

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

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

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

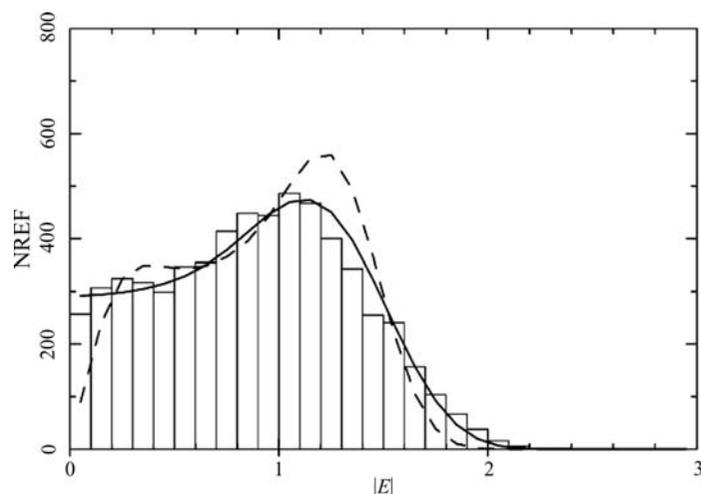


Figure 1.6.2.2

Exact p.d.f.'s for a crystal of [(Z)-ethyl *N*-isopropylthiocarbamato- κ S](tricyclohexylphosphine- κ P)gold(I) in the triclinic system. Solid curve: $P\bar{1}$, computed from (1.6.2.4); dashed curve: $P1$, computed from (1.6.2.5); histogram based on the data computed from all the reflections with non-negative reduced intensities. The height of each bin corresponds to the number of reflections (NREF) in its range of $|E|$ values. The p.d.f.'s are scaled up to the histogram.

Rabinovich *et al.* (1991) and by Shmueli & Wilson (2008) for all the space groups up to and including $Fd\bar{3}$.

The following example deals with a very high sensitivity to atomic heterogeneity. Consider the crystal structure of [(Z)-ethyl *N*-isopropylthiocarbamato- κ S](tricyclohexylphosphine- κ P)gold(I), published as $P\bar{1}$ with $Z = 2$, the content of its asymmetric unit being $\text{AuSPONC}_{24}\text{H}_{45}$ (Tadbuppa & Tiekink, 2010). Let us construct a histogram from the $|E|$ data computed from all the observed reflections with non-negative reduced intensities and compare the histogram with the p.d.f.'s for the space groups $P1$ and $P\bar{1}$, computed from equations (1.6.2.5) and (1.6.2.4), respectively. The histogram and the p.d.f.'s were put on the same scale. The result is shown in Fig. 1.6.2.2.

A visual comparison strongly indicates that the space-group assignment as $P\bar{1}$ was correct, since the recalculated histogram agrees rather well with the p.d.f. (1.6.2.4) and much less with (1.6.2.5). The ideal Wilson-type statistics incorrectly indicated that this crystal is noncentrosymmetric. It is seen that the ideal p.d.f. breaks down in the presence of strong atomic heterogeneity (gold among many lighter atoms) in the space group $P\bar{1}$. Other space groups behave differently, as shown in the literature (*e.g.* Rabinovich *et al.*, 1991; Shmueli & Weiss, 1995).

Additional examples of applications of structure-factor statistics and some relevant computing considerations and software can be found in Shmueli (2012) and Shmueli (2013).

1.6.2.3. Symmetry information from the structure solution

It is also possible to obtain information on the symmetry of the crystal after structure solution. The latter is obtained either in space group $P1$ (*i.e.* no symmetry assumed) or in some other candidate space group. The analysis may take place either on the electron-density map, or on its interpretation in terms of atomic coordinates and atomic types (*i.e.* chemical elements). The analysis of the electron-density map has become increasingly popular with the advent of dual-space methods, first proposed in the charge-flipping algorithm by Oszlányi & Sütő (2004), which solve structures in $P1$ by default. The analysis of the atomic coordinates and atomic types obtained from least-squares refinement in a candidate space group is used extensively in

structure validation. Symmetry operations present in the structure solution but not in the candidate space group are sought.

An exhaustive search for symmetry operations is undertaken. However, those to be investigated may be very efficiently limited by making use of knowledge of the highest point-group symmetry of the lattice compatible with the known cell dimensions of the crystal. It is well established that the point-group symmetry of any lattice is one of the following seven centrosymmetric point groups: $\bar{1}$, $2/m$, mmm , $4/mmm$, $\bar{3}m$, $6/mmm$, $m\bar{3}m$. This point group is known as the holohedry of the lattice. The relationship between the symmetry operations of the space group and its holohedry is rather simple. A rotation or screw axis of symmetry in the crystal has as its counterpart a corresponding rotation axis of symmetry of the lattice and a mirror or glide plane in the crystal has as its counterpart a corresponding mirror plane in the lattice. The holohedry may be equal to or higher than the point group of the crystal. Hence, at least the rotational part of any space-group operation should have its counterpart in the symmetry of the lattice. If and when this rotational part is found by a systematic comparison either of the electron density or of the positions of the independent atoms of the solved structure, the location and intrinsic parts of the translation parts of the space-group operation can be easily completed.

Palatinus and van der Lee (2008) describe their procedure in detail with useful examples. It uses the structure solution both in the form of an electron-density map and a set of phased structure factors obtained by Fourier transformation. No interpretation of the electron-density map in the form of atomic coordinates and chemical-element type is required. The algorithm of the procedure proceeds in the following steps:

- (1) The lattice centring is determined by a search for strong peaks in the autocorrelation (self-convolution, Patterson) function of the electron density and the potential centring vectors are evaluated through a reciprocal-space R value.
- (2) A complete list of possible symmetry operations compatible with the lattice is generated by searching for the invariance of the direct-space metric under potential symmetry operations.
- (3) A figure of merit is then assigned to each symmetry operation evaluated from the convolution of the symmetry-transformed electron density with that of the structure solution. Those symmetry operations that have a good figure of merit are selected as belonging to the space group of the crystal structure.
- (4) The space group is completed by group multiplication of the selected operations and then validated.
- (5) The positions of the symmetry elements are shifted to those of a conventional setting for the space group.

Palatinus & van der Lee (2008) report a very high success rate in the use of this algorithm. It is also a powerful technique to apply in structure validation.

Le Page's (1987) pioneering software *MISSYM* for the detection of 'missed' symmetry operations uses refined atomic coordinates, unit-cell dimensions and space group assigned from the crystal-structure solution. The algorithm follows all the principles described above in this section. In *MISSYM*, the metric symmetry is established as described in the first stage of Section 1.6.2.1. The 'missed' symmetry operations are those that are present in the arrangement of the atoms but are not part of the space group used for the structure refinement. Indeed, this procedure has its main applications in structure validation. The algorithm used in Le Page's software is also implemented in *ADDSYM* (Spek, 2003). There are numerous reports of successful applications of this software in the literature.