

## 3. METHODOLOGY

The method has the advantage of being very efficient for indexing low-symmetry patterns. The main disadvantage is its sensitivity to errors in the peak positions, particularly in the low  $2\theta$  region.

## 3.4.3.1.2. Shirley–Ishida–Watanabe (SIW) heuristic strategy

This needs only one single well-established zone, then it arbitrarily chooses the 001 line from the first-level lines. The indexing problem is thus lowered to two dimensions and an exhaustive search is carried out.

## 3.4.3.1.3. Index-heuristics strategy

The index-heuristics strategy searches for the correct cell *via* a trial-and-error approach, assigning tentative Miller indices to a few experimental peak positions (basis lines), usually belonging to the low  $2\theta$  region of the experimental pattern. It was first proposed by Werner (1964), then successively refined (Werner *et al.*, 1985) and made more robust and effective (Altomare *et al.*, 2000, 2008, 2009). This approach, which works in the index space, was defined by Shirley as semi-exhaustive (Shirley, 1980). The search starts from the highest-symmetry crystal system (cubic) and, if no plausible solution is found, it is extended to lower symmetry down to triclinic. The number of selected basis lines increases as the crystal symmetry lowers. A dominant zone occurs when one cell axis is significantly shorter than the other two; in this case most of the first observed lines (in terms of increasing  $2\theta_{hkl}$  values) can be indexed with a common zero Miller index. Special short-axis tests, aimed at finding two-dimensional lattices, have been proposed for monoclinic symmetry in order to detect the presence of dominant zones (Werner *et al.*, 1985). The index-heuristics method is based on the main indexing equation [see equation (3.4.2)] that can be rewritten (Werner *et al.*, 1985) as

$$Q(hkl) = h^2x_1 + k^2x_2 + l^2x_3 + h k x_4 + h l x_5 + k l x_6,$$

where  $\{x_i\} = X$  is the vector of unknown parameters, which are derived by solving a system of linear equations

$$\mathbf{M}\mathbf{X} = \mathbf{Q}, \quad (3.4.8)$$

where  $\mathbf{M}$  is a matrix of Miller indices and  $\mathbf{Q}$  is the vector of the selected  $Q(hkl)$  values corresponding to the basis lines. The dimensions of  $\mathbf{M}$ ,  $\mathbf{X}$  and  $\mathbf{Q}$  change according to the assumed symmetry. From the inverse matrix  $\mathbf{M}^{-1}$  the corresponding  $\mathbf{X}$  is obtained *via*  $\mathbf{X} = \mathbf{M}^{-1}\mathbf{Q}$ . In the case of monoclinic and higher symmetry,  $\{x_i\}$  are calculated by Cramer's rule. Different  $\mathbf{X}$  vectors are derived by using a different selection of basis lines. The possible solutions are checked by using the full list of peak positions (up to the first 25 experimental lines). The method is sensitive to errors on peak positions and to the presence of impurities (the presence of only one impurity peak is not critical). The correctness of the  $\{x_i\}$  strongly depends on the accuracy of the observed  $Q$  values, especially for low- $2\theta$  region lines, which are the most dominant ones for this indexing procedure. The possibility of testing different combinations of basis-line sets enables the correct cell to be found by bypassing the cases for which errors in the basis lines occur.

The method has been recently enhanced (Altomare *et al.*, 2000, 2009) by introducing new procedures that are able to increase the probability of successful indexing (see Section 3.4.4.2.1); among them are: (1) a correction for systematic errors in the experimental  $2\theta$  values (positive and negative trial  $2\theta$  zero shifts are taken into account); this correction should, in principle, describe a real diffractometer error; in practice, it also approximates the

specimen displacement error well (perhaps coupled with transparency for organic samples); (2) a more intensive search in solution space for orthorhombic and monoclinic systems; (3) an improvement of the triclinic search; (4) a new figure of merit, WRIP20, which is more powerful than  $M_{20}$  in identifying the correct solution among a set of possible ones (see Section 3.4.2.1); (5) a check for geometrical ambiguities; (6) an automatic refinement of the possible cells; and (7) a statistical study of the parity of the Miller indices, performed at the end of the cell refinement, aimed at detecting doubled axes or additional lattice points (for *A*-, *B*-, *C*-, *I*-, *R*- or *F*-centred cells) (such information is used in the successive steps).

## 3.4.3.1.4. Index-permutation strategy

This strategy was proposed by Taupin (1973), and is based on a systematic permutation of indices associated to observed lines for obtaining candidate cells. Because this trial-and-error strategy is similar to the index-heuristics approach, we do not describe it here.

## 3.4.3.1.5. Successive-dichotomy search method

The successive-dichotomy method, first developed by Louër & Louër (1972), is based on an exhaustive strategy working in direct space (except for triclinic systems, where it operates in reciprocal space) by varying the lengths of the cell axes and the interaxial angles within finite intervals. The search for the correct cell is performed in an  $n$ -dimensional domain  $D$  (where  $n$  is the number of cell parameters to be determined). If no solution belongs to  $D$ , the domain is discarded and the ranges for the allowed values of cell parameters are increased; on the contrary, if  $D$  contains a possible solution, it is explored further by dividing the domain into  $2^n$  subdomains *via* a successive-dichotomy procedure. Each subdomain is analyzed and discarded if it does not contain a solution. The method was originally applied to orthorhombic and higher-symmetry systems (Louër & Louër, 1972), but it has been successively extended to monoclinic (Louër & Vargas, 1982) and to triclinic systems (Boultif & Louër, 1991). The search can be performed starting from cubic then moving down to lower symmetries (except for triclinic) by partitioning the space into shells of volume  $\Delta V = 400 \text{ \AA}^3$ . For triclinic symmetry  $\Delta V$  is related to the volume  $V_{\text{est}}$  suggested by the method proposed by Smith (1977), which is able to estimate the unit-cell volume from only one line in the pattern:

$$V_{\text{est}} \simeq \frac{0.60d_N^3}{\frac{1}{N} - 0.0052},$$

where  $d_N$  is the value for the  $N$ th observed line; in the case  $N = 20$  the triclinic cell volume is  $V_{\text{est}} \simeq 13.39d_{20}^3$ .

Let us consider, as an example, the monoclinic case; in terms of direct cell parameters,  $Q(hkl)$  is given by (Boultif & Louër, 1991)

$$Q(hkl) = f(A, C, \beta) + g(B),$$

where  $f(A, C, \beta) = h^2/A^2 + l^2/C^2 - 2hl \cos \beta / (AC)$ ,  $A = a \sin \beta$ ,  $C = c \sin \beta$ ,  $g(B) = k^2/B^2$  and  $B = b$ . The search using the successive-dichotomy method is performed in a four-dimensional space that is covered by increasing the integer values  $i, l, m$  and  $n$  in the intervals  $[A_-, A_+] = [A_- = A_0 + ip, A_+ = A_- + p]$ ,  $[B_-, B_+] = [B_- = B_0 + lp, B_+ = B_- + p]$ ,  $[C_-, C_+] = [C_- = C_0 + mp, C_+ = C_- + p]$  and  $[\beta_-, \beta_+] = [\beta_- = 90 + n\theta, \beta_+ = \beta_- + \theta]$ , where the step values of  $p$  and  $\theta$  are  $0.4 \text{ \AA}$  and  $5^\circ$ , respectively, and  $A_0, B_0$  and  $C_0$  are the lowest values of  $A, B$  and  $C$  (based on the positions of the lowest-angle peaks), respectively. Each quartet of intervals