Enhancements of the ‘Auto-Indexing’ Method for Cell Determination in Four-Circle Diffractometry

BY WILLIAM CLEGG

Institut für Anorganische Chemie der Universität, Tammannstrasse 4, D-3400 Göttingen, Federal Republic of Germany

(Received 9 March 1984; accepted 15 May 1984)

Abstract
Enhancements in the auto-indexing procedure for cell determination improve its power and reliability. A proper selection of three basis reflections for construction of an initial subcell is essential. A standard cell-reduction procedure applied to the subcell, and the calculation of appropriate limits for the direct-lattice-vector generation integers, make a comprehensive coverage of the direct space efficient and effective. Refinement of the generated lattice vectors against all the available reflection data makes assignment of the correct unit cell easier and more reliable. Generation of some additional long vectors is recommended in special cases.

1. Introduction
Before intensities can be measured from a crystal of unknown lattice parameters and orientation matrix on a four-circle diffractometer, the $3 \times 3$ orientation matrix $UB$ (as defined by Busing & Levy, 1967) must be determined (Clegg, 1984). An automatic procedure involves first finding a set of general reflections (with the aid of photographs or by a general search), then generating a UB matrix, which assigns integral indices to the reflections. The UB matrix relates the indices $h$ of a reflection to the coordinates $x$ of the corresponding reciprocal-lattice vector referred to an orthogonal $\phi$-axis-fixed system: $x = UBh$.

Two basic strategies for the generation of the UB matrix have been described. One finds the three shortest non-coplanar reciprocal-lattice vectors which can be generated from the available list of reflections and assigns these as the reciprocal-cell axes (Hornstra & Vossers, 1974). The resulting unit cell is always primitive, and additional steps are required to establish the conventional cell and Bravais-lattice type (see § 3f). This method is used in the standard control software of Philips, Enraf–Nonius and older Stoe diffractometers.

The second method, named ‘auto-indexing’ by its originator (Sparks, 1976, 1982; Jacobson, 1976), and used for Syntex/Nicolet and more recent Stoe–Siemens diffractometers, generates direct-lattice vectors.

It has been described in algebraic terms by Sparks (1976, 1982). We present here an alternative explanation, which serves as a foundation for the enhancements described subsequently.

2. The auto-indexing method
From the available list of reflections, we choose three non-coplanar reciprocal-lattice vectors (usually the shortest, i.e., the reflections with lowest 20), and arbitrarily assign to them indices 100, 010 and 001. We can now generate an orientation matrix and direct unit cell corresponding to the three indexed reflections (Busing & Levy, 1967). Although this cell $(a', b', c')$ is not usually the correct one, it must be a subcell of the direct lattice: all vectors between points of the true crystal lattice are also vectors in the lattice described by the preliminary subcell, but the converse is not usually true. The auto-indexing method generates direct space vector $t = ua' + vb' + wc'$ ($u, v, w$ integral) and tests each one against all the observed reciprocal vectors (reflections) to see if this $t$ vector could be a true crystal lattice vector. From the fundamental relationship between the direct and reciprocal lattices (International Tables for X-ray Crystallography, 1972), a possible direct-lattice vector $t$ must satisfy the condition $t.x = integer$ for all the reflections. In practice, a small tolerance is allowed for experimental errors.

The auto-indexing procedure generates all integer triples $u, v, w$, up to a maximum value for each integer or for their sum; allowing negative values for two of the integers covers half the vector space, since $+t$ and $-t$ are bound to be equally good (or bad) solutions. Vectors satisfying the $t.x$ integer criterion are stored in a list. Their lengths and angles subtended by pairs of them can be printed, and three vectors can be selected, either manually or automatically, to form the required unit cell.

The commercially provided control software incorporating the auto-indexing method includes a valuable enhancement of the basic procedure: for any vector $t$, a specified maximum number of reflections is allowed to fail the $t.x$ test. For each vector accepted as a possible solution, the offending reflections are
flagged. One or two reflections that are inconsistent with most or all of the vectors produced could be derived from a secondary crystal fragment or an impurity, or from a machine fault. Even twinned crystals, with two non-coincident lattices, can be successfully investigated, although care is needed in selecting a self-consistent set of vectors for the unit-cell axes.

3. Enhancements

The auto-indexing method is the standard cell-determination procedure in our local control software for a Stoe–Siemens AED diffractometer (Clegg, 1981b). From experience of its application for the determination of unit cell and UB matrix for many hundreds of crystals in recent years, we have adapted and enhanced the basic routine to improve reliability. We describe here enhancements that have been found effective in the subcell generation, the integer triple generation and the treatment of the derived vectors.

(a) Selection of basis reflections

If the method is to succeed, obviously the initial subcell must provide a valid basis for the generation of lattice vectors. Selection of three reflections at lowest 2θ values has the advantage of not producing an unnecessarily small direct-lattice subcell and thus having to generate a large number of integer triples to reach the preset maximum vector length. If, however, the reflections are weak, they may be spurious or not very precisely located. In such cases, very few or no vectors will be found. We have found it advantageous to select instead the three most intense reflections, but also to allow a manual selection to override this automatic choice.

(b) Subcell generation and reduction

If the three basis reflections correspond to reciprocal-space vectors that are approximately orthogonal, the subcell axes will also display near orthogonality. In many cases, however, unless the angular distribution is taken into account when selecting the basis reflections, the subcell may have disparate axial lengths and interaxial angles far removed from 90°. In some directions in the crystal lattice small values of u, v or w correspond to very long vectors, while in other directions large values of the triples are required to generate even short vectors. Reduction of the subcell before generation of triples solves this problem. The exact nature of the reduction process is unimportant, since a truly reduced subcell is not the ultimate aim. We use an algorithm for Niggli reduction (Krivý & Gruber, 1976) that is part of the control software in any case (Clegg, 1981a). This subcell reduction, together with the triple generation method described below, guarantees that all possible vectors shorter than a specified maximum length will be generated without producing large numbers of unnecessarily long vectors.

(c) Integer triple generation

The Syntex/Nicolet auto-indexing procedure (Sparks, 1976, 1982) generates all possible triples with \( s = |u| + |v| + |w| = 1 \), before going on to \( s = 2, 3, \ldots \) up to a maximum specified \( s \). For a markedly non-reduced subcell, this generates much longer vectors in some directions than others, as noted above, and the techniques of gradually increasing \( s \) rather than covering the full range of triples in any other way does not necessarily produce vectors in anything like an increasing length order. A more even angular distribution of generated vectors is obtained if individual limits are set for the three integers \( u, v, w \). These limits are simply

\[
\begin{align*}
\text{for } a^* &= U_{\text{max}} = l_{\text{max}} a^*, \quad b^* = l_{\text{max}} b^*, \quad c^* = l_{\text{max}} c^*, \\
\text{where } a^*, b^*, c^* \text{ are reciprocal-subcell lengths derived from the reduced subcell, and } l_{\text{max}} \text{ is the maximum allowed vector length.}
\end{align*}
\]

If all triples are generated with \( 0 \leq u \leq u_{\text{max}} \), \( -r_{\text{max}} \leq v \leq r_{\text{max}} \), and \( -w_{\text{max}} \leq w \leq w_{\text{max}} \), no vectors shorter than \( l_{\text{max}} \) are missed, and generation of vectors longer than \( l_{\text{max}} \) is minimized; \( u_{\text{max}}, v_{\text{max}} \) and \( w_{\text{max}} \) are automatically set from the reduced subcell, and the only limit specified by the user for vector generation is \( l_{\text{max}} \). Thus, unnecessary computation is avoided, as are possible pitfalls due to uneven coverage of the vector space. It may be noted incidentally that a similar calculation, with direct and reciprocal-space exchanged, gives the maximum attainable reflection indices for a given maximum 2θ; this is useful in setting up for data collection: \( h_{\text{max}} = 2a(sin \theta_{\text{max}})/\lambda \), and similarly for \( k_{\text{max}} \) and \( l_{\text{max}} \).

(d) Vector refinement

Vectors are accepted as possible lattice vectors if \( t.x \) is close to an integer for all the reflections. This test is fulfilled exactly by the three basis reflections, but the remaining reflections will generally produce small deviations from integral results even for true lattice vectors. Thus, the vectors, although found to be consistent with all the reflections, are calculated in fact from only three of them. Once a vector has passed the \( t.x \) test, it is possible to refine it, so that the results represent a least-squares fit to all the available data:

\[
t = C^{-1} A
\]

where

\[
C_{ij} = \sum_{n=1}^{N} x_i x_j, \quad A_i = \sum_{n=1}^{N} x_i k_n
\]

for \( N \) reflections; \( k_n \) is the nearest integral value to \( t.x \) for reflection \( n \). This is a standard linear least-squares refinement requiring no iteration, and takes very little time. It is performed only for vectors that have already been accepted on the basis of the \( t.x \) test; any reflections flagged as inconsistent for a particular
vector are not used in the refinement of that vector. Refinement of the vectors is equivalent to a linear least-squares refinement of the UB matrix elements based on the reflections after the three cell axes have been selected from the vector list (Tichý, 1970; Shoemaker & Bassi, 1970), so that this subsequent matrix refinement is not necessary. Refining the vectors before selecting the cell axes rather than afterwards particularly improves the estimation of the inter-vector angles, especially in cases where the three basis reflections correspond to markedly non-orthogonal reciprocal-lattice vectors, and makes the recognition of orthogonal cell axes easier and more reliable.

(e) The generation of additional vectors

Sorting the generated vectors according to their length is a simple way of making the vector list easier to read and handle, if manual selection of axes is to be made. Further vectors can be added to the list in special cases. The sorted list is examined for vectors with approximately equal lengths, and extra vectors given by sums of such vector pairs and triples are added (Clegg, 1981a). These additional vectors, longer than the specified \( l_{\text{max}} \), may be long axes for centred unit cells, or a long hexagonal \( c \) axis for a rhombohedral cell with \( c \gg a \) (Sparks, 1982; Clegg, 1981a,c).

(f) Selection of cell axes

From the generated vector list, three axes can be chosen manually, particularly if the cell geometry is already known. Alternatively, the program can automatically select the three shortest non-coplanar vectors to form a primitive cell. Following cell reduction, a check can then be made for a possible higher symmetry cell. In all the commercial diffractometer systems, the method of examination of the reduced form is used (Mighell & Rodgers, 1980). We have already commented on the possible pitfalls of this method (Clegg, 1981a), and there have been abundant examples in the recent literature to underline them (e.g. Marsh & Herbstein, 1983, and references cited therein). An alternative method, which identifies two-fold axes in the lattice, has proved robust and reliable, and can be strongly recommended (Le Page, 1982).

4. Possible problems

In the majority of cases, the auto-indexing method, with the enhancements described here, leads automatically and rapidly to the correct metric cell and orientation matrix. Problems usually manifest themselves in a lack of \( t \) vectors: either no vectors at all, or only a few coplanar ones are found. There are various ways of dealing with this situation.

(a) Allow more reflections to fail the \( t \times x \) test. Normally a small number of reflections will be repeatedly flagged as inconsistent, and the correct cell will be found.

(b) Increase \( l_{\text{max}} \), if the true cell axis lengths are unknown.

(c) Select the three basis reflections manually, while simultaneously allowing some inconsistent reflections. The use of an inconsistent reflection as a basis reflection undermines the whole auto-indexing procedure! In cases of twinning, with strong reflections present from both crystal lattices, several attempts may have to be made, and the cell axes must also be selected manually to ensure that they all belong to the same twin component.

Problems may also occur if all reflections belong to a special subset. For example, if all the reflections have an even \( l \) index (this may be caused by heavy atoms lying on special positions and producing systematically weak reflections with odd \( l \)), the cell produced will have a halved \( c \) axis. Auto-indexing itself cannot avoid this problem, which requires further attention once the cell has been obtained, e.g. via oscillation photographs taken about the cell axes (Sparks, 1982; Clegg, 1984).

I thank Professor G. M. Sheldrick for helpful discussions and for his suggestion of the vector refinement, and my colleagues for providing a variety of problem cases.

References