
Trial and error indexing program for powder patterns of monoclinic substances. By F. Kohlbeck* and E. M. Hörl, Österreichische Studiengesellschaft für Atomenergie GmbH, Institut für Metallurgie, Forschungszentrum Seibersdorf, Austria

(Received 21 March 1977; accepted 28 September 1977)

A method for finding the unit cells from the d values of powder patterns of monoclinic substances is described. An appropriate computer program based on this method is developed and combined with an already existing program [Kohlbeck & Hörl, J. Appl. Cryst. (1976), 9, 28–33] which has treated monoclinic cells as triclinic. In this way, calculation times are obtainable for monoclinic cells approximately five times shorter than those obtained with the previous program alone.

Outline of method

By use of d values only the problem of indexing unknown monoclinic lattices can be stated by solving the following system of linear equations:

\[ |Q_i - Ah_i^2 - Bk_i^2 - Cl_i^2 - Dh_i k_i| \leq AQ_i \quad i = 1 \ldots n \]  

(1)

with the following abbreviations:

- \( Q_i = \frac{1}{d_i^2} \)
- \( A = a^* a^* \)
- \( C = c^* c^* \)
- \( B = b^* b^* \)
- \( D = 2a^* b^* \).  

(2)

(1) is written in the first setting; as are all other equations and deductions of this paper. \( AQ_i \) are the errors of the measured \( d_i \) resulting from the errors of the observed d values \( d_i \); n is the number of the observed lines and \( a^*, b^*, c^* \) are the reciprocal translation vectors of the unknown lattice. Solutions of (1) are never unique and should be examined by calculating the indices of reliability defined by de Wolff (1968).

We try to find four lines (called base lines), with indices \( r, s, t, u \), which lead to a non-singular system of equations:

\[
\begin{align*}
Q_1 &= Ah_1^2 + Bk_1^2 + Cl_1^2 + Dh_1 k_1, \\
Q_2 &= Ah_2^2 + Bk_2^2 + Cl_2^2 + Dh_2 k_2, \\
Q_3 &= Ah_3^2 + Bk_3^2 + Cl_3^2 + Dh_3 k_3, \\
Q_4 &= Ah_4^2 + Bk_4^2 + Cl_4^2 + Dh_4 k_4 \\
\end{align*}
\]  

(3)

with \( A, B, C, D \) being the correct, but unknown, parameters defined in (2). Selecting these lines is one of the main problems of the trial and error method. Choosing these lines randomly would, with high probability, give a singular system (3) which would be of no use.

After the base lines have been selected (3) is solved with indices \( h_i, k_i, l_i \) up to a preselected limit. The solutions are checked with the remaining \( n-4 \) equations (1) but with expanded error ranges \( AQ_i \) instead of the original \( AQ_i \). The new error ranges \( AQ_i \) are a consequence of the parameter errors resulting from (3). In the standard use of the program we choose \( AQ_i = 0.015 Q_i \). This is insufficient in most cases. However, if the matrix of (3) is ill-conditioned, this expansion of \( AQ_i \) will not be sufficient. Because the real errors \( AQ_i \) are not known, an exact calculation of \( AQ_i \) is not possible. Experience with the program has shown that it is of no use to select \( AQ_i > 0.03 Q_i \). In this case too many solutions of (1) would occur. If a parameter set fulfills all but one or two of the equations (1) within the errors \( AQ_i \), this solution is improved by minimization of the sum of the squares of the left sides. Afterwards, the index of reliability \( M_{20} \) is calculated, not taking into account systematic extinctions. With standard input parameters solutions with \( M_{20} \geq 7 \) are printed out.

Selection of base lines

Considering the reciprocal-lattice vectors corresponding to the four base lines \( r, s, t, u \), the most probable reasons for the matrix of the system (3) being singular are: (1) two reciprocal-lattice vectors have the same direction, (2) three vectors are lattice vectors of a rectangular net plane, (3) all four vectors are lattice vectors of a common net plane.

Taking into account only the first few observed lines, case (1) most probably will occur if the reduced cell is elongated in one direction. Cases (2) and (3) may be present when the cell is plate-like. The program chooses base-line index \( r \) equal to one. To avoid case (1) it selects the first 11 lines having no d value ratios of \( \frac{1}{2}, \frac{3}{2}, 2 \). These lines are referred to as set \( dx \) in the following. The second line of \( dx \) is chosen to be the base line with index \( s \). The next seven lines are checked with respect to cases (2) and (3) if they are assignable to common net planes with lines \( r \) and \( s \). The indices of the two-dimensional nets are restricted to be less than three. The set of four lines which has the lowest chance* of fulfilling conditions (2) and (3) is taken as base-line set. If a rectangular net is found which explains eight of the first nine lines of set \( dx \), the procedure of selecting the base lines is repeated.

The tolerances for indexing plane nets are reduced, and lines 2 to 11 of set \( dx \) are used for further base-line selecting.

Calculation of lattice parameters

It is assumed that the base lines have low indices when

* At present assistant at the Geophysics Institute of the Technical University of Vienna, Austria.

* No exact calculation in the sense of mathematical probability theory is carried out.
referred to a reduced cell. Without undue loss of generality we can choose the unit cell in such a way that one of the following two sets of conditions holds:

1st set: 
- $0 \leq h, l, n, k, l, l \leq 2$
- $|h| \leq 3$
- $0 \leq l, n, l, l \leq 3$
- $0 \leq h, l \leq 4$
- if $h, l \neq 0$: $k = 0$
- if $k \neq 0$: $h = 0$ \hspace{1cm} (4)

2nd set: 
- $(h, n, h, k, k,): \{(2002), (0020), (2000)\}$
- $k + k = 2n \hspace{1cm} n = \text{integer}$
- $-4 \leq h, l \leq 4$
- $0 \leq h, l \leq 4$
- $0 \leq h, l \leq 5$ \hspace{1cm} (5)

The second set is assumed to correspond to a sublattice with lattice parameters $A', B', C', D'$. From this lattice we deduce the actual lattice parameters:

- $A = A' + B' + D'$
- $B = A' + B' - D'$
- $C = C'$
- $D = (A' - B')/2$.

With the assumptions either of (4) or of (5), system (3) can be written as one of the following matrix equations:

$$
\begin{pmatrix}
Q_1 \\
Q_2 \\
Q_3 \\
Q_4
\end{pmatrix}
= 
\begin{pmatrix}
h_1^2 & 0 & l_1^2 & 0 \\
0 & k_2^2 & l_2^2 & 0 \\
h_3^2 & k_4^2 & l_4^2 & h_5 \\
h_6^2 & k_7^2 & l_7^2 & h_8
\end{pmatrix}
\begin{pmatrix}
A \\
B \\
C \\
D
\end{pmatrix}
$$

\hspace{1cm} (6)

$$
\begin{pmatrix}
Q_1 \\
Q_2 \\
Q_3 \\
Q_4
\end{pmatrix}
= 
\begin{pmatrix}
h_1^2 & 0 & l_1^2 & 0 \\
0 & k_2^2 & l_2^2 & 0 \\
h_3^2 & k_4^2 & l_4^2 & h_5 \\
h_6^2 & k_7^2 & l_7^2 & h_8
\end{pmatrix}
\begin{pmatrix}
A' \\
B' \\
C' \\
D'
\end{pmatrix}
$$

\hspace{1cm} (7)

If one solution of (6) or (7) is known, further solutions are obtained very quickly: with solutions

$$
X = \begin{pmatrix}
A \\
B \\
C \\
D
\end{pmatrix}
\quad \text{and} \quad
X' = \begin{pmatrix}
A' \\
B' \\
C' \\
D'
\end{pmatrix}
$$

\hspace{1cm} (8)

assumed for (6) or (7) and the corresponding equation with $h, l, k$ replaced by $h, l, l'$, respectively, there exists the relation

$$
X' = X - (S/\delta') \cdot W
$$

\hspace{1cm} (9)

with $S$ given by

$$
S = (h_1^2 - h_2^2)A + (l_1^2 - l_2^2)C.
$$

\hspace{1cm} (10)

$\delta$ and $\delta'$ are the determinants of the square matrices $M$ and $M'$ in (6) or (7) with indices $h, l, k$ and $h, l, l'$ respectively. They are related by:

$$
\delta' = \delta + (h_1^2 - h_2^2)w_1 + (l_1^2 - l_2^2)w_3.
$$

\hspace{1cm} (11)

$w_1$ and $w_3$ are elements of the vector $W$, which is the first column vector in the adjoint matrix $M^{-1}$ of $M$. The proof of (9) is straightforward.

**Verification of solutions**

After the lattice parameters have been obtained the volume of the unit cell is calculated. If this volume is too large or if it is not in accordance with the density of the substance the cell is eliminated. The maximum volume $V_{\text{max}}$ (de Wolff, 1961) is chosen to be

$$
V_{\text{max}} = 260 \, d_3^{2}\, .
$$

An estimate of $60 \, d_3^{2}$ would guarantee that not more than 60 lines of the calculated lattice exist having $Q$ values lower than $Q_{20}$. This would be an appropriate limit for a realistic cell dimension. The extra factor of two has to be added since one has to allow for centred lattices.

The next step in the calculation concerns the indexing of the remaining $(n-4)$ lines with the calculated lattice parameters. In order to keep computing time short, the following considerations are taken into account. Since

$$
Q + AQ \geq Ah^2 + Bk^2 + Cl^2 + Dhk \geq Q - AQ
$$

\hspace{1cm} (12)

has to be solved with $A, B, C > 0$ and unknowns $h, k, l$, one obtains

$$
Q + AQ - Bk^2 - Ah^2 - Dhk > Cl^2
$$

\hspace{1cm} (12)

With

$$
f_1 = (Q - Bk^2 - Ah^2 - Dhk + AQ)/C
$$

$$
f_2 = f_1 - 2AQ/C
$$

we obtain the condition

$$
f_1 \geq f_2 \geq f_3.
$$

\hspace{1cm} (13)

If $f_1$ and $f_2$ are already calculated, condition (13) can be checked with only three Fortran statements if the integer-squares $l^2$ are put into a field and the rounded values of $f_1$ and $f_2$ are used as searching indices. Therefore, we obtain the advantage of checking only a two-parameter number of relations (13) instead of a three-parameter number of possible relations (12).

**Testing of the program**

The program has been tested with the data of the monoclinic substances used by Kohlbeck & Hörbl (1976). On average, the computation time has been reduced by a factor of approximately five. Comparison of computation time in the various program parts has shown that more than 90% of the time is used to optimize the cell parameters and to calculate the $M_{20}$ values. Further development in trial and error indexing should therefore primarily concern these program parts.

The authors are indebted to the computer centre of the Technical University of Vienna for support. They further wish to thank Professor Dr F. Lihl, Professor Dr A. Scheidegger and Dozent Dr F. Koss for their continuous interest in developing the program.

**References**


