A novel powder diagrams indexing, using classical geometry

M.L. Ettorche,[1,a] M. Sebais,[1] and Z. Hammoudi[2]
[1]*Laboratoire de Cristallographie, Département de Physique, Faculté des Sciences, Université Mentouri Constantine, Route Ain El Bey, Constantine 25000, Algeria*
[2]*Laboratoire Signaux et Systèmes de Communication, Département d'Electronique, Faculté de Sciences de l'ingénieur, Université Mentouri Constantine, Route Ain El Bey, Constantine 25000, Algeria*

(Received 19 March 2012; accepted 7 September 2012)

Based only on a geometrical approach, we present a technique to index powder diffraction diagrams. This would allow us to find the cell parameters from the experimental data. It is well known that methods proposed in the literature make a direct use of the experimental data to build the cell, whereas our approach exploits them to calculate theoretical values, which could be multiples of two of the three vectors' lengths of the unit cell, and then uses them along with the experimental values. To show the effectiveness of the proposed algorithm, several examples, requiring only minor limitations in linear dimensions (<35 Å) and volume (<4500 Å^3), are treated. For all considered cases, except the triclinic symmetry that is time consuming, the corresponding FORTRAN routine is executed in a reasonable time (<3 min with a 3 GHz processor). © *2012 International Centre for Diffraction Data.*
[doi:10.1017/S088571561200070X]

Key words: powder diffraction, indexing

I. INTRODUCTION

In crystallography literature, whenever powder diffraction data is of interest, indexing is the first step that a scientist should be involved with. That is, since the 1960s, many researchers have been interested in trying to find methods of calculation, often analytical, to get efficiently through this step.

Among these methods, three have emerged and have been frequently used, namely: *ITO* (Visser, 1969), *TREOR* (Werner *et al.*, 1985) and *DICVOL*91 (Boultif and Louër, 1991). Nevertheless, these methods have shown their limits over time in meeting requirements which have not been foreseen or have needed more resources and efforts. For instance, we may cite the zero-point error, the absolute angular error of lines' positions, the presence of impurities, and so on. Therefore, improved versions are found in the literature. That is, in *DICVOL04* (Boultif and Louër, 2004; Louër and Boultif, 2006), the authors have added the correction of the zero-point error introduced in *DICVOL91*. In *DICVOL06* (Louër and Boultif, 2007), they have introduced a new strategy to limit the risk to miss a solution as in *DICVOL*04. On the other hand, in *N-TREOR* (Altomare *et al.*, 2000), the authors have improved *TREOR90*, by making it more exhaustive and capable of refining the selected unit cell automatically. Finally, in *N-TREOR09* (Altomare *et al.*, 2009), the authors have introduced a new figure of merit.

Moreover, the growth of the power of computing machines and the resolution and accuracy of recording devices, have yielded the emergence of other methods; all aiming to increase the probability of obtaining good indexing. To this end, *GAIN* (Kariuki *et al.*, 1999) used the genetic algorithm. It exploits the Le Bail method (Le Bail *et al.*, 1988) to fit the line profiles and *X-Cell* (Neumann, 2003) explores the successive dichotomies combined with the search of zero-point error by allowing unindexed lines. Furthermore, *McMAILLE* (Le Bail, 2004) makes use of the Monte–Carlo process and *CRYSFIRE* suite (Shirley, 1999) employs several programs in a complementary way, among them *ITO*, *DICVOL* and *TREOR*. At this stage, it is worth noting that these methods remain insufficient since the solution is not always unique, the figure of merit, exploited for identifying the correct cell, can be inefficient and it is up to the user to select the most suitable one. The number of solutions depends mainly on the quality of the data. It is clear that worse the data quality is, more solutions we have and more likely the user gets confused.

Unlike the existing methods that are computationally complex, here we present a method to solve the problem of indexing in terms of a simple geometrical approach. The multiple cases that we treated numerically, showed the efficiency of the proposed approach and its capability of resolving difficulties associated with the zero-point error, the absolute angular error and so on. Note that the proposed method treats all symmetries, except the rhombohedral one, which is included in the hexagonal symmetry.

The paper is organized as follows. In Sections II, III and IV, we present the theory of the considered approach, describe the implementation of the proposed method and treat the determination of the zero-point errors, respectively. In Section V, we discuss the problem inherent to impurities. In Section VI, we describe the proposed algorithm. Then, in Sections VII and VIII, we treat numerical examples. Finally, in Section IX, we end the paper with some concluding remarks.

II. THE METHOD

A. Principles

To clarify the basic idea used in this work, let us consider three distances p_1, p_2 and p_3. A triangle whose sides are p_1, p_2 and p_3 may then be constructed (see Figure 1).

[a] Author to whom correspondence should be addressed. Electronic mail: ettorche_lamine@yahoo.fr

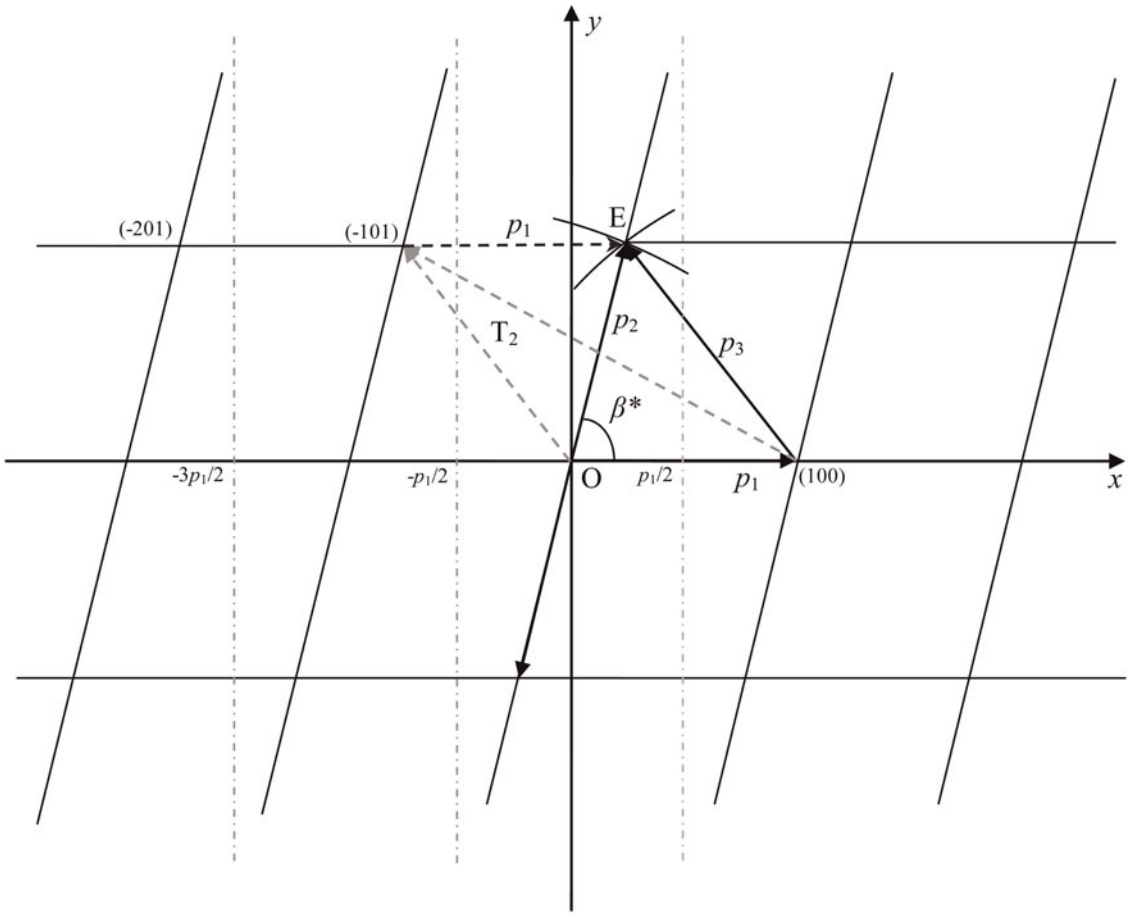

Figure 1. Description of a face of the reciprocal cell and the angle β^* at the origin.

Let (Oxy) be an orthonormal coordinate system in which p_1 is along the Ox axis. Therefore, point E, the intersection of p_2 and p_3, has coordinates (x, y) such that

$$x = (p_1^2 - p_3^2 + p_2^2)/2p_1, \quad (1)$$

$$y = \sqrt{p_2^2 - x^2}\ . \quad (2)$$

Note that x may have negative values.

On the other hand, it is well known that an angle between two sides may be easily derived geometrically. That is, if β^* is the angle between Ox and the line segment OE, then

$$cos\beta^* = x/p_2. \quad (3)$$

For instance, if p_1, p_2 and p_3 are the vectors in the reciprocal lattice corresponding, respectively, to the reflections of indices (100), (001) and (-101), then the triangle of Figure 1, would describe a face of the reciprocal cell and β^* the angle at the origin.

To construct the other faces of the cell, the above construction process is repeated for all reciprocal distances, allowing cell determination.

Note that the vector length $p_i = 1/d_{hkl}$, where d_{hkl} is the interplanar spacing.

B. Determination of powder diagram's new lines

We start by exploiting the experimental data at hand, noted set (A), which is used to build sets (B) and (C). These two sets may then contain values that could be the lengths of vectors b^* and c^*, respectively, and then, while carrying out the building of all possible faces, we determine the missing data values which could be necessary for the construction of the unit cell. In this way, we may build more triangles and choose those whose point E has abscissa x lying in the interval $[-3p_1/2, p_1/2]$. This interval has been chosen on the basis of the minimum angular error criterion. That is, compared with the reflections lying in the high 2θ region of the powder pattern, the distances that correspond to those that appear in the first part and the middle of the powder diagram, usually, satisfy this criterion. Note that this interval is considered to be sufficiently large such that the determination of the missing values is made possible.

We recall that in high symmetries, the reciprocal unit cell has two rectangular or square faces, whereas the third one could be neither rectangular nor square. Let a^*, b^* and c^* be the three vectors basis of the reciprocal unit cell. The rest of the paper assumes that the length differences of the basis vectors are not too large. Considering the vector a^* along Ox, we will first be interested in the construction of the cell's rectangular faces containing a vector b^* that is perpendicular to a^*. In the following, we assume that the distance p_1 of set (A), could be a multiple of the length of a^*.

Then, we begin by looking for triangles such that abscissa x of E′ is equal to 0 or $-p_1$, Figure 2. If x =0, then the length of OE' could be a multiple of the length of b^*. However, if $x = -p_1$, the ordinate y of E′ could be a multiple of the length

https://doi.org/10.1017/S088571561200070X Published online by Cambridge University Press

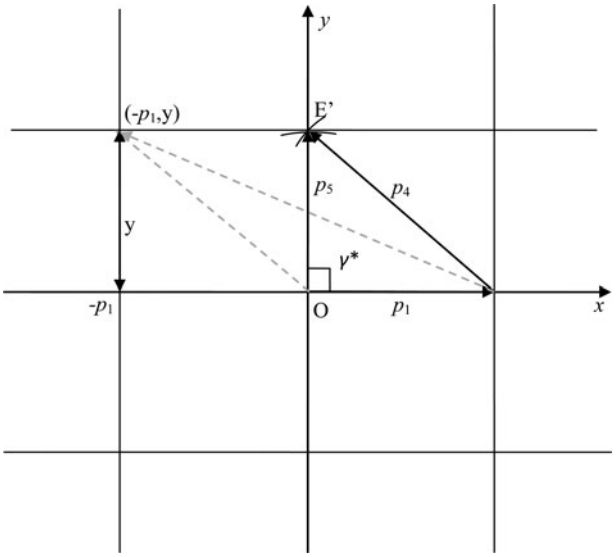

Figure 2. Description of a face of the reciprocal cell and the angle $\gamma^* = 90°$ at the origin.

of b^*. In this manner, we can construct set (B) from values of length of OE′ or from those of the ordinate y of E′.

To construct set (C), let us return to Figure 1. We can see that, among the values of angle β^*, the one that is the closest to 90° allows the nodes (001) and (00 − 1) to remain in the area bounded by the straight lines $x = p_1/2$ and $x = -p_1/2$ (i.e., $x \in [-p_1/2, +p_1/2]$). The identification of these nodes' positions yields to the search of triangles for which point E is situated between these two lines. In this case, we admit that the side OE is a multiple of the length of c^*. The periodic arrangement of the nodes in the reciprocal lattice shows that the other nodes of the same lines, i.e., (h01) and (h0−1), along Ox, are distant hp_1, from nodes (001) and (00 − 1), respectively; all within the area bounded by the straight lines $x = hp_1 + p_1/2$and $x = hp_1 - p_1/2$. In this way, the knowledge of one's position, determines the position of the other and vice versa. The construction of triangle T_2, whose distances are assumed to correspond to the reflections of indices (100), (−101) and (−201), allows us to know the position of node (−101), which in turn gives the position of node (001). This follows from the fact that if (x, y) are the coordinates of node (−101), then those nodes (001) are ($x + p_1$, y). Then, when $x \in [-3p_1/2, -p_1/2]$, i.e., $h = -1$, the side of the reciprocal lattice that could be a multiple of the length of c^*, is obtained by the following:

$$p' = \sqrt{(x + p_1)^2 + y^2}\ . \tag{4}$$

Note that the case $x = -p_1/2$ is included in the first interval (i.e., $[-p_1/2, +p_1/2]$) and therefore excluded from the second because, according to Eq. (4), if $x = -p_1/2$ then $p' = \sqrt{(p_1/2)^2 + y^2} = \text{OE}$, which again satisfies the hypothesis of the first interval (i.e., OE is a multiple of the length of c^*).

Finally, the selected values are gathered in set (C). In summary, for each value p_1 of set (A), we have built two sets (B) and (C). We should specify that, in the following, in addition to set (A), sets (B) and (C) are built for the high symmetries. However, only set (C) is necessary for triclinic symmetry. Regardless of the original value of p_1 (multiples of a^* or not), the procedure is repeated for p_1 divided by 2 and 3.

C. Cell construction

The construction of sets (B) and (C) was done upon the basis that there are two types of symmetries in the crystal system; namely, the one in which the cells have two rectangular or square faces and the other in which the cells possess neither a rectangular nor a square face.

In the first type of symmetry, we determine the triangle that defines the first face. For this, we use two elements p_1 and p_2 of set (A) and an element p_3' of set (C). This triangle has p_1 as a basis and p_2 and p_3' as sides that intersect at point E. β^* is the angle at the origin. It is obtained from Eqs (1) and (3).

As in this type of symmetry, α^* and γ^* are equal to 90°, the cell construction is completed once we place perpendicularly to this face, an element p_4'' of set (B), which cuts off p_1 and p_3' at point O.

In addition, this construction has also been realized using doubles and halves values of p_1 and p_3' in order to increase the probability of obtaining the cell of concern.

Since the nature of the elements (primitives or multiples) of sets (A), (B) and (C) is unknown, the unit cell is obtained by further dividing by n ($n = 1$, 2, 3 and 4), p_1, p_4'' and p_3'.

Finally, to fully exploit the experimental data and thus to have more chance of finding the solution, the above procedure is repeated by assigning successively to p_1, the first ten elements of set (A) and to p_2, p_3' and p_4'', all the elements of sets (A), (C) and (B), respectively.

D. Extension to triclinic cases

In the triclinic symmetry, the three angles are different from 90°. Then, we must determine the three faces of the structure. To do this, we first start by building the first two triangles. Knowing that these triangles must have a common side, we take three elements p_1, p_2 and p_3 of set (A), which we combine with two elements p_2' and p_5' of set (C) to obtain two triangles. Note that p_1 intersects with p_2' and p_5' at O and it is chosen to be a common side of the two triangles.

Using these triangles, we identify five parameters of the reciprocal cell; namely the three sides p_1, p_2' and p_5' and two angles α^* and γ^*, which are derived similarly to β^* by equations such as Eqs (1) and (3). Specifically, we use for α^*, in Eq. (1), p_2' and p_2 in place of p_2 and p_3, respectively, and in Eq. (3), p_2' in place of p_2, and for γ^*, in Eq. (1), p_5' and p_3 in place of p_2 and p_3, respectively, and in Eq. (3), p_5' in place of p_2.

Note that, these two triangles in the trihedron could be at the limit, either superposed; i.e., the angle between them is equal to 0°, or completely spread out, i.e., the angle between them is equal to 180°. As shown in Figure 3, according to the above limits, we deduce that the smallest angle between p_2' and p_5' is equal to $|\alpha^* - \gamma^*|$ (position 1) and the greatest angle is equal to $(\alpha^* + \gamma^*)$ (position 2), where |·| designates the absolute value.

The trihedron is defined completely only when we build the third triangle. This is achieved by using sides p_2', p_5' and an element p_4 of set (A) and taking into account that the angle β^* at O, between p_2' and p_5', must satisfy

$$\left|\alpha^* - \gamma^*\right| \le \beta^* \le \alpha^* + \gamma^*. \tag{5}$$

https://doi.org/10.1017/S088571561200070X Published online by Cambridge University Press

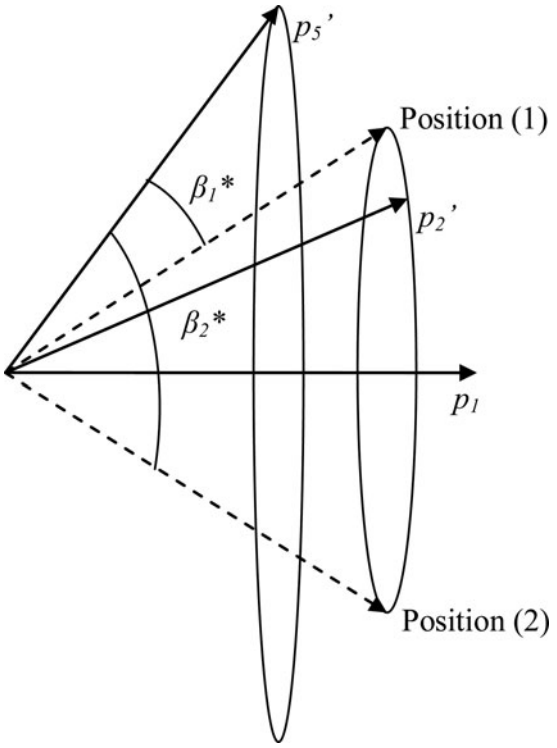

Figure 3. Illustration of positions (1) and (2) as extreme limits of p_2'.

To ensure that the reciprocal cell is elementary, the linear parameters are further divided by a factor n' ($n' = 1, 2$).

Note that by assigning all values of sets (A) and (C) to p_2, p_3 and p_4, and p_2' and p_5', respectively, we may construct several trihedra.

III. IMPLEMENTATION OF THE PROPOSED METHOD

To estimate the effectiveness of this method, a general program, named 'IGC' (indexation by geometrical construction) was written in FORTRAN. This program implements the techniques introduced above. It begins by building sets (B) and (C) [or only set (C), in the triclinic case]. Then, it combines their elements with the experimental values of set (A) to construct a model of the cell which can be a solution to the problem of indexation. When the data are affected by the zero-point error, they are corrected. Therefore, more accurate unit cell parameters are determined (see Section IV). Moreover, the quality factors M_{20} (de Wolff, 1968) and F_{20} (Smith and Snyder, 1979) are calculated to appreciate the validity of the proposed solution.

Many examples (up to 150) have been treated with this program. The results, similar to those obtained by other programs (*e.g.*, Louër and Vargas, 1982) showed that the small volume solution, provided by high quality data, is usually the correct solution. It comes out that the strategy of this program is the same as the one adopted by the majority of the existing programs. This requires a volume cut into slices whose volume is 400 Å^3. The unit cells whose volumes are less than a slice section belong to the slice in question. Such a strategy can speed up the search process until one (or more) solution(s) of smaller volume(s) is (are) found.

This time limitation is more perceptible when we deal with the case of high symmetry systems. Going from the lowest volume to the highest volume, it is known that solutions usually appear from the lowest symmetry to the highest symmetry. To have a wide field of search for solutions in the monoclinic system, the angle β is set up to 125°. Then, if the solution happens to be in a slice, the search of other solutions in larger volumes will continue simply by attributing to β the values 90° or $\gamma = 120°$ in the case of a hexagonal symmetry. This would enable solutions in the other symmetries to appear.

In practice, the first 20 lines are sufficient to index a powder diagram. It is recommended that the considered phase be pure and the absolute angular error of line positions be less than 0.03° in 2θ.

The tests were performed on examples taken from the Personal Computer Powder Diffraction File (PCPDF) using a 3 GHz processor. The success rate was 96% and the execution time for most of these examples did not exceed 3 min (except the triclinic symmetry that required more time). The treated examples were cells with linear parameters less than 35 Å and volumes not exceeding 4500 Å^3.

IV. ZERO-POINT ERROR

The proposed method also treats problems that may arise when the lines' positions of the powder diffraction diagram are not zero-point error free. This error appears clearly in the harmonic lines series. It is perceptible through a small displacement, in the same direction, of these lines.

In this work and, in a similar manner, by the reflection-pair method (Dong *et al.*, 1999), we determine the zero-point error by looking for all pairs of the experimental data which verify the following relation:

$$m_i\, p_m / n_i\, p_n \approx 1, \tag{6}$$

where p_m and p_n, elements of set (A), are lines corresponding to the diffraction angles θ_m and θ_n, respectively. The multipliers m_i and n_i are integers.

Equation (6) can be set exactly equal to '1' if we simply introduce a real value Zp, called the zero-point error, to yield

$$n_i \sin(\theta_n + Zp) = m_i \sin(\theta_m + Zp). \tag{7}$$

Following this reasoning, several values of Zp are obtained. However, as in 98% of the treated cases, the correction of zero-point error on θ has not exceeded 0.05°, i.e., $|Zp| \leq 0.05°$. We kept only those that were less than this limit. Note that since relation (6) holds for $p_m = p_n$, then a zero value of the zero-point error is also expected.

V. TREATMENT OF IMPURITIES

To have a reliable indexing, we consider that all the lines of the powder diffraction pattern are important and then hope that the measurements of lines positions are correct, the compound is pure, and that the phase is unique. Otherwise, in order to get around any of these problems, we allow Ni lines to be non-indexed.

In most indexed examples, $Ni \leq 2$. However, in some cases, the mixture of phases and the presence of impurities, cause the appearance of a large number of fake lines. Therefore, Ni may be greater than 2. To this effect, many solutions could be obtained. The risk of finding them in a slice of

https://doi.org/10.1017/S088571561200070X Published online by Cambridge University Press

volume less than the volume which contains the correct solution is considerable.

VI. PROPOSED ALGORITHM

As shown in Figure 4, the proposed algorithm works as follows. It starts reading the experimental data (N values) given as distances d_{hkl} or diffraction angles 2θ. Using these data and as explained earlier (Section II.B), it builds sets (B) and (C). Then, it constructs the cell that could be a solution. That is, using this cell, the algorithm calculates the theoretical diffraction angles that are to be compared with the experimental ones (Test (1))

$$|2\theta_{\text{Theoretical}} - 2\theta_{\text{Experimental}}| \leq \Delta(2\theta)_{\text{Global}}\,, \qquad (8)$$

where $\Delta(2\theta)_{\text{Global}} = \Delta(2\theta) + 2Zp$ and $\Delta(2\theta)$ is the absolute angular error. Test (1) assumes that $\Delta(2\theta)_{\text{Global}} > \Delta(2\theta)$. That is, at this stage, we used only the upper bound value of the zero-point error ($Zp = 0.05°$). Note that this test has considerably reduced the computation time. Once Test (1) is satisfied, the elements of (A), (B) and (C) which contributed to the construction of the candidate cell are corrected by adding Zp [calculated from Eq. (7)] to their respective diffraction angles. As long as values of Zp are not used up, new cell parameters and new diffraction angles are calculated and compared with the experimental angles (Test (2))

$$\left|2\theta_{\text{Corrected-Theoretical}} - 2\theta_{\text{Experimental}}\right| \leq \Delta(2\theta). \qquad (9)$$

Once Test (2) is satisfied, Test (3) is performed according to the well-known M_{20} and F_{20} figures of merit. Now, if Test (3) is satisfied, which means that M_{20} and F_{20} are, in general, greater than 10, then there exist at least one solution. This procedure is repeated until all the elements of sets (A), (B) and (C) are used up. Note that, if there is no solution in the high symmetries, the algorithm looks for solutions in triclinic symmetry. Finally, it ends either when there is no solution in both high symmetries and the triclinic one or when there is at least one solution in either one of them.

VII. EXPERIMENTAL RESULTS

In order to show the efficiency of the proposed algorithm, this section is devoted to examples of different symmetries that may occur in the construction of a powder pattern cell. That is, most of the treated cases were chosen according to the difficulties that may be encountered in the process of diagram indexation such as the presence of zero-point error, the absence of lines of indices (*0k0*) or (00 *l*) or both, the absolute angular error of lines' positions, and the cell volume. In the latter case, when we have a big cell volume, the powder

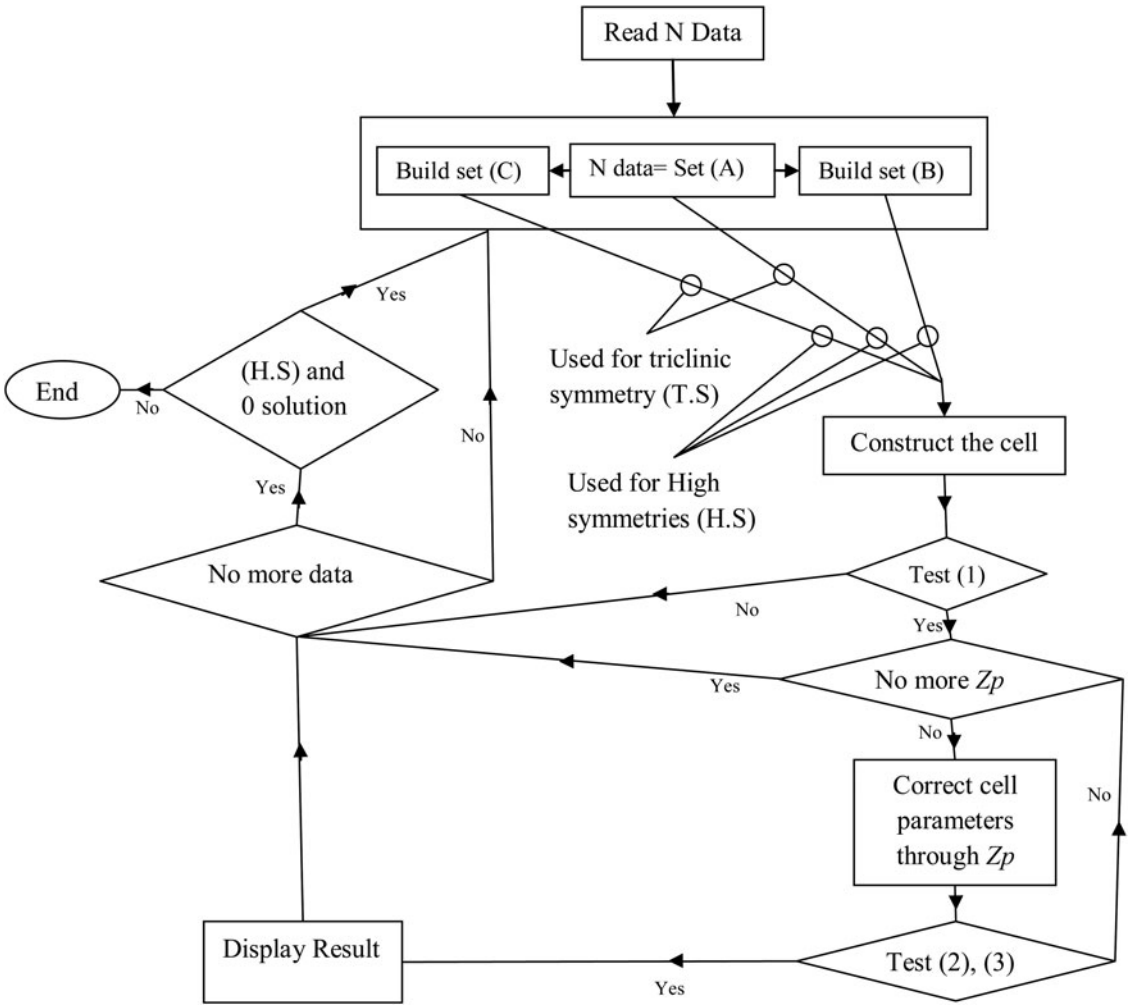

Figure 4. Flow chart of the proposed algorithm.

https://doi.org/10.1017/S088571561200070X Published online by Cambridge University Press

TABLE I. Unit cell parameters of $C_{26}H_{31}GeO_5P$, obtained by applying, DICVOL06, N-TREOR and our method.

Program	a (Å)	b (Å)	c (Å)	β (°)	M_{20}	F_{20}	$\Delta(2\theta)$	V (Å^3)	Zero-point (2θ)
DICVOL06	14.2237	8.0962	23.3309	103.697	10.7	25.8 (0.0125, 62)	0.045°	2610.32	0.0783°
N-TREOR	23.3197	8.077	14.26	103.707	11	27. (0.0119, 63)	–	2609.70	0.08°
I.G.C	14.248	8.101	23.444	103.725	7	18 (0.017, 63)	0.045°	2628.5	0.0404°
True cell	14.30	8.115	23.46	103.68	–	–	–	2646.55	–

pattern contains, generally, several lines, therefore, the space between these lines is very narrow, which makes the zero-point error and the absolute angular error of lines' positions very annoying. These conditions often permit the solutions of low volume to appear, and to nearly miss the solution that should appear in a slice of higher volume.

The aim of these examples is to find solutions that are closest, in terms of cell parameters and volume, to those given by the Joint Committee on Powder Diffraction Standards (JCPDS) file. Among these examples, there is one the processing of which has revealed all the potentials of the proposed procedure. The example is $C_{26}H_{31}GeO_5P$ (Diethyl 1 – (triphenylgermanyl)-3-Keto-2-oxapentylphophinate). It is known that this compound should be indexed in monoclinic symmetry. Using $\lambda = 1.5406$ Å, the powder diffraction did not show any reflections of indices (*0k0*). The program has to look for the reflections of indices (*0k0*) and those of indices (*00 l*) before starting any cell construction. Among several theoretical values of p_4'' and p_3', calculated in order to be associated with the experimental line p_1 ($1/p_1 = 13.929$ Å) indexed by (100), two values have led to the cell that constitutes a good approximation of the desired solution. The first value was $1/p_3' = 5.7084$ Å which is less than half of the experimental value (11.475 Å) indexed by (002), and the second was $1/p_4'' = 4.057\,67$ Å whose experimental line is totally absent. The unit cell parameters were $a_c = 14.248$ Å, $b_c = 8.101$ Å, $c_c = 23.444$ Å, $\beta_c = 103.725°$ and $V_c = 2628.50$ Å^3. Note that the value 5.7084 has been retained in cell construction while the value 11.475 has not.

Using this cell, 20 lines were indexed over the 20 initially used lines. The tolerated absolute angular error $\Delta(2\theta) = 0.045°$. The uniqueness of the solution and the values of the figures of merit ($M_{20} = 7$ and $F_{20} = 18$ (0.017, 63)) show that the solution is acceptable. Note the zero-point error $Zp = 0.020\,23$ degrees in θ.

The solution found in the JCPDS file is given by $a = 14.30$ Å, $b = 8.115$ Å, $c = 23.46$ Å, $\beta = 103.68°$ and $V = 2646.55$ Å^3. Although small, the disparities between these parameters and ours are, generally, owing to the absolute angular error of the lines' positions. Note that these disparities have an effect on the theoretical reflection lines calculated using our unit cell parameters. This could favor, for certain experimental lines, one indexing over another. For instance, this appears in lines 4.0153 and 3.9728 indexed, respectively, by (310) and (302), instead of being indexed, respectively, by (213) and (015), as given in the JCPDS file.

For comparison purposes, the above example has also been treated by programs, which are considered as benchmarks in the crystallography literature, such as *DICVOL06* and *N-TREOR*. Their results are reported in Table I. Although the result given by our approach is closer to the true solution in terms of volume and cell parameters, we agree that, if we confine ourselves to the given example and not others, the results given by *DICVOL06* and *N-TREOR* are better if we look at their figures of merit, which are higher compared with those obtained by our approach. However, if we compare the time and effort that have been used so far to set these programs, we may conclude that our approach needs to be improved and affined to attain such results.

We may observe that in this approach a lot of experimental data are directly or indirectly involved in the construction of the cell. Thus, the effectiveness of the proposed method is proven to be closely related to the accuracy of the measurements.

To extend the study to more interesting examples, Table II lists some of the results obtained from other compounds found

TABLE II. Examples of unit cell parameters obtained by applying the proposed method.

	Parameters calculated								True parameters				
Formula	a (Å)	b (Å)	c (Å)	β (°)	V (Å^3)	M_{20}	F_{20}	t (s)	at (Å)	bt (Å)	ct (Å)	βt (°)	Vt (Å^3)
$C_{12}H_{12}CaN_2O_{12}Zr\text{-}6H_2O^{(q)}$	10.92	–	18.81	90.00	2242.50	24	43(0.009, 51)	24	10.90	–	18.78	90.00	2232.61
$Ca_6Al_2(SeO_4)_3(OH)_{12}\text{-}26H_2O^{(H)}$	11.41	–	21.51	120.00	2426.45	32	69(0.009, 29)	32	11.39	–	21.46	120.00	2411.89
$K_3La_2(NO_3)_6^{(c)}$	13.67	–	–	90.00	2556.31	57	86(0.007, 31)	21	13.66	–	–	90.00	2549.29
$Ce\ Mg(NO_3)_6\text{-}6H_2O^{(c)}$	12.57	–	–	90.00	1986.48	107	182(0.004, 25)	25	12.57	–	–	90.00	1986.26
$C_{26}H_{31}GeO_5P^{(M)}$	14.25	8.10	23.44	103.72	2628.50	7	18(0.017, 63)	27	14.30	8.115	23.46	103.68	2646.55
$C_{19}H_{35}Cl_2N_2O_7P^{(O)}$	22.12	9.99	11.88	90.00	2624.48	11	26(0.016, 42)	52	22.17	10.01	11.97	90.00	2659.14
$C_{15}H_{20}ClN_3O^{(O)}$	18.30	11.06	15.38	90.00	3113.47	20	28(0.012, 52)	60	18.34	11.08	15.40	90.00	3132.44
$Ca(ReO_4)_2.H_2O^{(O)}$	23.19	11.30	11.93	90.00	3127.57	14	34(0.013, 41)	57	23.18	11.33	11.97	90.00	3144.77
$C_{24}H_{21}N_7\text{-}H_2O^{(M)}$	18.21	12.43	9.97	91.36	2257.06	11	30(0.014, 43)	39	18.23	12.30	9.973	91.34	2236.91
$C_{20}H_{20}As_2Fe_4O_{12}S_2^{(M)}$	17.20	13.80	14.47	112.50	3171.91	13	23(0.013, 63)	31	17.22	13.83	14.48	112.46	3190.15
$Al_{43}Cr_4Yb_6^{(H)}$	10.87	–	17.57	120.00	1797.09	29	40(0.013, 39)	26	10.86	–	17.57	120.00	1794.58
$C_{11}H_{14}ClO_2PS^{(O)}$	25.73	6.26	8.19	90.00	1319.21	16	32(0.016, 37)	34	25.86	6.276	8.202	90.00	1331.32
$C_4BaO_9Ti\text{-}4H_2O^{(M)}$	14.05	13.81	13.39	88.45	2598.60	48	160(0.003, 32)	24	14.04	13.81	13.38	91.48	2594.92
$C_{36}H_{48}FeO_{12}P_3^{(M)}$	25.37	8.55	19.3	98.12	4145.78	14	51(0.009, 41)	36	25.31	8.514	19.22	98.13	4102.81
$C_{46}H_{54}N_4O_4Sr^{(M)}$	18.16	19.89	12.59	108.41	4315.23	13	49(0.008, 48)	39	18.16	19.82	12.54	108.41	4287.92

https://doi.org/10.1017/S088571561200070X Published online by Cambridge University Press

in the JCPDS files, which could be considered as a sample of all the examples we have treated so far. The choices were made on the basis of the difficulties encountered with diagram indexation. Referring to Table II, the compounds were indexed; some completely and others partially, in a maximum time of 60 s. The number of non-indexed lines $Ni \leq 2$. The used lines are 20, the absolute angular error $\Delta(2\theta) = 0.045°$ or 0.06°, $\beta_{(max)} = 125°$ and the zero-point error $|Zp| \leq 0.05°$. The symbols $^{(c)}$, $^{(q)}$, $^{(H)}$, $^{(O)}$ and $^{(M)}$ designate cubic, tetragonal, hexagonal, orthorhombic and monoclinic symmetries, respectively.

In Table II, we show the best results according to their M_{20} and F_{20}. These solutions are not always unique. Although, they do not often appear, at the head of the list of the results, they are usually among the top five. Note that we did not adopt any criterion to rank the solutions.

For each compound of monoclinic symmetry, the number of solutions that appear is less than six. These solutions are not very different neither in dimensions of the cell parameters nor in values of merit factors M_{20} and F_{20}. The gap between their merit factors is less than 3, except in the case of compound $C_4BaO_9Ti\text{-}4H_2O$, where the difference is considerable between the merit factors of the cell ($a = 14.05$ Å, $b = 13.81$ Å, $c = 13.39$ Å, $\beta = 88.45°$, $V = 2598.60$ Å^3, $M_{20} = 48$, $F_{20} = 160$ (0.003,32)), considered to be the best and which appears in the second place in the list of solutions and the solution that has the lowest values of M_{20} and F_{20} ($a = 14.1$ Å, $b = 13.86$ Å, $c = 13.44$ Å, $\beta = 91.69°$, $V = 2624.62$ Å^3, $M_{20} = 17$, $F_{20} = 57$ (0,011, 33)) and which appears in the first position. This shows that the program has used different lines to determine the first and second solutions.

As, going from the lowest volume to the highest volume, the solutions usually appear from the lowest symmetry to the highest symmetry. Consequently, the strategy followed in our program is to let the solutions appear in their natural order. Therefore, the solutions of compounds of high symmetries (except the monoclinic symmetries), usually appear alone in their volume section; making them easy to detect.

Finally, we should point out that the execution time is acceptable. It could be further shortened if we reduce the number of experimental values of p_1. In effect, to increase the probability of having one of the lines (*h00*), the first ten values of set (A) are assigned to p_1. Knowing that, in the high symmetry case, one of the (*h00*) lines may be present in the first five experimental lines, then, except for the monoclinic and triclinic symmetries, the execution time could be divided by 2.

VIII. BETHANECHOL CHLORIDE BENCHMARKS

To test the robustness and the limits of our program, we investigated more interesting operational cases through the use of the 'Bethanechol Chloride ($C_7H_{17}ClN_2O_2$) benchmarks' (Bergmann *et al.*, 2004). The two entries ICDD PDF 43-1748 and 46-1964 were the subject of the following tests. These tests are presented in the form of 10 different lists of which each one contains 20 lines. They are listed as follows:

A. Indexing the raw data. A(1) for ICDD PDF entry 43-1748 and A(2) for 46-1964.
B. Indexing the data with $I \geq 5\%$ (I/I_{max}). B(1) and B(2) as above.

TABLE III. Results for the bethanechol chloride benchmarks.

	A(1)		A(2)		B(1)		B(2)		C(1)		C(2)		D(1)		D(2)		E(3)		F(4)		Note		Global
Program	Def-man		Def-man		Def-man		Def-man		Def-man		Def-man		Def-man		Def-man		Def-man		Def-man		Def-man		Note
ITO13	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	−1	+1	+1	+1	−8	−6	−14
DICVOL91	−1	−1	−1	−1	−1	−1	−1	+1	−1	−1	−1	−1	−1	−1	−1	+1	+1	+1	+1	+1	−6	−2	−8
DICVOL04[a]	−1	−1	−1	−1	−1	0_2	−1	+1	−1	−1	+1	−1	−1	−1	−1	+1	+1	+1	+1	+1	−6	+1	−5
TREOR90	−1	−1	−1	−1	−1	−1	+1	+1	−1	−1	−1	−1	−1	−1	+1	+1	+1	+1	+1	+1	−2	−2	−4
DICVOL04[b]	−1	+1	−1	+1	−1	+1	−1	+1	−1	0_7	−1	+1	−1	+1	−1	+1	+1	+1	+1	+1	−6	+9	+3
McMaille	−1	0_6	+1	+1	+1	+1	−1	+1	−1	−1	−1	+1	−1	+1	−1	+1	+1	+1	+1	+1	−2	+7	+5
I.G.C	Na	−1	na	−1	na	+1	na	+1	na	−1	na	+1	na	0_2	na	+1	na	+1	na	+1	na	+3	−7

(1) ICDD PDF entry 43-1748, λ = 1.5418 Å.
(2) ICDD PDF entry 46-1964, λ = 1.5418 Å.
(3) Conventional X-ray data, λ = 1.540 56 Å.
(4) Synchrotron data, λ = 0.6995 Å.
na, not applicable, there is only a manual mode (taken as −1 when calculating the totals).
0_n means that it was the *n*th cell proposal.
DICVOL04[a]: fast default and manual tests made by A. Le Bail.
DICVOL04[b]: manual tests made by D. Louër.

https://doi.org/10.1017/S088571561200070X Published online by Cambridge University Press

TABLE IV. Unit cell parameters of example F(4) obtained by different programs.

Program	a (Å)	b (Å)	c (Å)	β (°)	V (Å^3)	M_{20}	F_{20}
ITO13	8.876	16.409	7.135	93.829	1036.78	159.8	–
DICVOL91	8.855	16.408	7.135	93.828	1036.86	150.2	862.3 (0.0007, 31)
DICVOL04[a]	7.137	16.411	8.876	93.825	1037.46	147.6	858.2 (0.0008, 31)
TREOR90	8.875	16.408	7.135	93.829	1036.87	157	880.(0.0007, 30)
DICVOL04[b]	7.134	16.409	8.875	93.830	1036.73	161.6	856.2 (0.0008, 31)
McMAILLE	7.134	16.409	8.875	93.830	1036.78	167.9	888.18 (0.0008, 30)
I.G.C	8.875	16.411	7.134	93.837	1036.78	161	828 (0.0008, 31)

DICVOL04[a]: fast default and manual tests made by A. Le Bail.
DICVOL04[b]: manual tests made by D. Louër.

C. Indexing data corrected for zeroshift. C(1) and C(2) as above.
D. Indexing data corrected for zeroshift and having $I \geq 5\%$ (I/I_{max}). D(1), D(2) as above.

To these 2×4 tests, are added two further and considerably easier tests:

E. Indexing new laboratory X-ray data.
F. Indexing synchrotron data.

These conditions correspond probably to more than 50% of the crystal structures stored in the ICSD and CSD database. The indexing programs should be applied to the following series of 10 data sets in automatic and manual modes. According to these benchmarks, the results of some of these indexing programs are reported in Table III as follows: '1' point means that the correct cell was found in the first FoM position among the proposals; '0' means that the correct cell is mixed with the incorrect ones, not at the head of the list, but listed among the first ten; The '−1' means that the correct cell was not found at all, or at a position larger than 10 in the lists.

These tests show that our approach gives results that are comparable to those given by other routines. The examples in which our routine has failed are examples A(1), A(2) and C(1). In A(1), the number of impurity lines was 9. Thus, the number of lines of the phase that we wanted to index (11) was not sufficient to compute the missing lines such as (*0k0*) and (*00 l*); in A(2) and C(1), the real solutions were found but appeared after the tenth. The merit factors of the other examples obtained by our method were, sometimes, smaller than those found by others. These drawbacks will be taken into consideration in future works to improve the proposed approach.

The global note in the last column in Table III is obtained by adding the 20 results (1, 0 or −1). Note that 'na' is taken as −1 when computing the totals.

For the easiest example *F*(4) given in Table IV, we remark that, as far as the cell parameters obtained by the different approaches are concerned, the difference is not noticeable. Nevertheless, it appears important in terms of figures of merit, but not to the point to doubt about one method or another.

IX. CONCLUSION

In this study, we have presented a technique to index powder diffraction diagrams. Unlike the existing methods that are computationally complex, we have introduced a method to solve the problem of indexing in terms of a simple geometrical approach. To show the effectiveness of the proposed algorithm, several examples have been treated. We have shown that the first example has treated the problems inherent to the absence of lines (*0k0*) and (*00 l*) and the zero-point error, by introducing two new calculated values and correcting the zero-point error. Regarding the simplicity of the proposed algorithm, we should point out that the obtained results are comparable to those given in the JCPDS files with an acceptable execution time. Finally, we think that our method can be improved if some crystallographic aspects and other mathematical tools are further exploited.

ACKNOWLEDGEMENTS

The authors would like to thank Pr. Armel Le Bail from University of Maine and Dr Toufik Laroussi from University Mentouri at Constantine, for their valuable assistance.

Altomare, A., Giacovazzo, C., Guagliardi, A., Moliterni, A. G. G., Rizzi, R., and Werner, P.-E. (**2000**). "New techniques for indexing: *N-TREOR* in *EXPO*," J. Appl. Crystallogr. **33**, 1180–1186.

Altomare, A., Campi, G., Cuocci, C., Eriksson, L., Giacovazzo, C., Moliterni, A., Rizzi, R., and Werner, P.-E. (**2009**). "Advances in powder diffraction pattern indexing: *N-TREOR09*," J. Appl. Crystallogr. **42**, 768–775.

Bergmann, J., Le Bail, A., Shirley, R., and Zlokazov, V. (**2004**). "Renewed interest in powder diffraction data indexing," Z. Kristallogr. **219**, 783–790.

Boultif, A. and Louër, D. (**1991**). "Indexing of powder diffraction patterns for low-symmetry lattices by the successive dichotomy method," J. Appl. Crystallogr. **24**, 987–993.

Boultif, A. and Louër, D. (**2004**). "Powder Pattern Indexing with the dichotomy method," J. Appl. Crystallogr. **37**, 724–731.

De Wolff, P. M. (**1968**). "A simplified criterion for the reliability of a powder pattern Indexing," J. Appl. Crystallogr. **1**, 108–113.

Dong, C., Wu, F., and Chen, H. (**1999**). "Correction of zero shift in powder diffraction patterns using the reflection-pair method," J. Appl. Crystallogr. **32**, 850–853.

Kariuki, B. M., Belmonte, S. A., McMahon, M. I., Johnston, R. L., Harris, K. D. M., and Nelmes, R. J. (**1999**). "A new approach for indexing powder diffraction data based on whole-profile fitting and global optimization using a genetic algorithm," J. Synchrotron Radiat. **6**, 87–92.

Le Bail, A. (**2004**). "Monte Carlo indexing with McMaille," Powder Diffr. **19**,, 249–254.

Le Bail, A., Duroy, H., and Fourquet, J. L. (**1988**). "Ab-initio structure determination of LiSbWO6 by X-ray powder diffraction," Mater. Res. Bull. **23**, 447–452.

Louër, D. and Vargas, R. (**1982**). "Indexation automatique des diagrammes de poudre par dichotomies successives," J. Appl. Crystallogr. **15**, 542–545.

Louër, D. and Boultif, A. (**2006**). "Indexing with the successive dichotomy method, DICVOL04," Z. Kristallogr. Suppl. **23**, 225–230.
Louër, D. and Boultif, A. (**2007**). "Powder pattern indexing and the dichotomy algorithm," Z. Kristallogr. Suppl. **26**, 191–196.
Neumann, M. A. (**2003**). "X-Cell: a novel indexing algorithm for routine tasks and difficult cases," J. Appl. Crystallogr. **36**, 356–365.
Shirley, R. (**1999**). http://www.ccp14.ac.uk/tutorial/crys/program/crysfire.htm
Smith, G. S. and Snyder, R. L. (**1979**). "F_N: A criterion for rating powder diffraction patterns and evaluating the reliability of powder-pattern indexing," J. Appl. Crystallogr. **12**, 60–65.
Visser, J. W. (**1969**). "A fully automatic program for finding the unit cell from powder data," J. Appl. Crystallogr. **2**, 89–95.
Werner, P.-E., Eriksson, L., and Westdahl, M. (**1985**). "*TREOR*, a semi-exhaustive trial-and-error powder indexing program for all symmetries," J. Appl. Crystallogr. **18**, 367–370.

https://doi.org/10.1017/S088571561200070X Published online by Cambridge University Press