

## A Simplified Criterion for the Reliability of a Powder Pattern Indexing

BY P. M. DE WOLFF

Laboratorium voor Technische Natuurkunde, Technische Hogeschool, Lorentzweg 1, Delft, The Netherlands

(Received 9 April 1968)

A figure of merit is defined for an indexing of a given pattern as  $M_{20} = Q_{20}/2\bar{\epsilon}N_{20}$ . Here  $N_{20}$  is the number of different calculated  $Q$  values up to  $Q_{20}$ , which is the  $Q$  value for the 20th observed and indexed line;  $\bar{\epsilon}$  is the average discrepancy in  $Q$  for these 20 lines. From a number of indexings which have been disproved by single-crystal analysis, the conclusion is drawn that  $M_{20} < 6$  must give rise to considerable doubt about the result. A number of confirmed indexings shows values of 20–60 for good routine work on pure, well-crystallized samples, and down to 6 for retrievable correct indexings of less accurate data. If the number of unindexed lines below  $Q_{20}$  is not more than two, a value  $M_{20} > 10$  guarantees that the indexing is substantially correct.

### Introduction

In a previous paper (de Wolff, 1961*b*) a procedure has been proposed for estimating the reliability of a unit cell derived exclusively from a given powder pattern. Application of this test has proved to be cumbersome. Since anyhow a rigorous statistical treatment of the problem is next to impossible – because of the intermixing of continuous variables (the diffraction angles) and completely unknown integers (the indices) – it seemed worth while to look for a simplified criterion. The present paper describes such a criterion in the form of a ‘figure of merit’.\* This figure  $M$  is closely related to the quantities discussed in the former paper, yet it can be understood independently, as shown in the next section.

### One-dimensional model

In order to understand the definition of  $M$ , consider a hypothetical crystal for which the possible diffraction angles  $2\theta$ , when plotted on an appropriate scale, yield multiples of a single parameter. This is, for instance, the case for a cubic lattice if the lines are plotted on a  $1/d^2$  or  $Q$  scale. We assume, however, that systematic vacancies (such as the well-known cases  $h^2 + k^2 + l^2 = 7, 15$  etc. in the cubic pattern) do not occur.

Let us call the variable, as a function of which the lines are mapped,  $\varphi$ . The period of the equidistant calculated lines is called  $p$ , so that  $\varphi_{\text{calc}} = hp$ , with an integer ‘index’  $h$ . Let the problem be to determine the parameter  $p$  from a set of measured angles, which represents a series of random  $h$  values and has random errors of measurement.

In Fig. 1, the situation is illustrated for a number of observed values of  $\varphi$ . Supposing a certain value of  $p$ , say  $p = p'$ , each line can be ‘indexed’ by assigning to it the index  $h_i$  of its nearest neighbour in the series

\* Substantially the same criterion was formulated earlier by de Wolff (1966*a*), but a severe error slipped into that definition,  $N$  being called the number of observed, instead of calculated, lines.

of multiples of  $p'$ . For each observed line  $\varphi_i$ , we shall call  $|\varphi_i - h_i p'|$  the discrepancy,  $\epsilon_i$ . Another value  $p''$  of  $p$  yields a different set of discrepancies.

The indexing procedure – by whatever method – aims at finding a value of  $p$  for which the discrepancies are very small. How small? This is a very pertinent question, since the  $\epsilon_i$  can be reduced indefinitely by choosing ever smaller values of  $p$ . As a matter of fact, each  $\epsilon_i$  is always smaller than  $\frac{1}{2}p$  because of the above mentioned identification with the closest multiple of  $p$ .

Here we hit on an important notion: even arbitrary values of  $p$  yield limited discrepancies. Being distributed evenly between 0 and  $\frac{1}{2}p$ , their average value is  $\frac{1}{4}p$ . Since we use this average value quite generally, we shall call it  $\bar{\epsilon}_a$ , and we have for the linear model

$$\bar{\epsilon}_a = \frac{1}{4}p. \quad (1)$$

The index  $a$  means that  $\bar{\epsilon}_a$  refers to an arbitrary value of  $p$ . Indeed, if  $p$  has its true value the discrepancies can be much smaller, provided the measurement has been accurate enough. This immediately answers the above question: a reliable value of  $p$  is indicated by discrepancies much smaller than  $\bar{\epsilon}_a$ . Again, however, a question comes up: how much smaller? We shall put this question in a numerical way by introducing the ratio  $M$ :

$$M(p) = \bar{\epsilon}_a(p)/\bar{\epsilon}(p), \quad (2)$$

where  $\bar{\epsilon}_a$  is given by (1), and  $\bar{\epsilon}(p)$  is the average value of the actual discrepancies for the parameter value  $p$ .

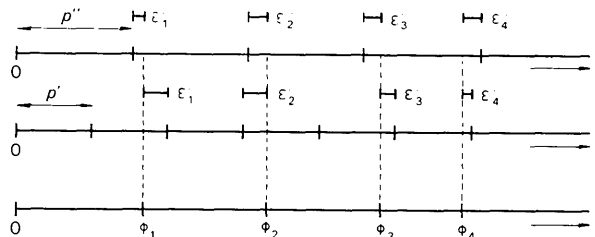


Fig. 1. Schematic illustration of discrepancies for a given one-dimensional pattern (lower scale), for two values  $p'$  and  $p''$  of the parameter.

In Fig. 2,  $M$  is shown as a function of  $p$  for two synthetic patterns (Table 1) in which arbitrary errors and vacancies were introduced.

Table 1. 'Observed' values of  $\varphi$  for two synthetic one-dimensional patterns:  $p_{\text{true}} = 1$

| (a) $\bar{\varepsilon} = 0.05$ |       | (b) $\bar{\varepsilon} = 0.14$ |       |
|--------------------------------|-------|--------------------------------|-------|
| 1.98                           | 19.92 | 1.93                           | 19.75 |
| 2.94                           | 25.04 | 2.82                           | 25.13 |
| 3.88                           | 26.02 | 3.65                           | 26.07 |
| 11.05                          | 26.94 | 11.14                          | 26.82 |
| 13.05                          | 27.91 | 13.16                          | 27.73 |
| 14.00                          | 32.00 | 13.99                          | 31.99 |
| 14.99                          | 35.03 | 14.98                          | 35.10 |
| 16.10                          | 36.01 | 16.31                          | 36.03 |
| 17.96                          | 38.04 | 17.87                          | 38.13 |
| 18.97                          | 39.04 | 18.92                          | 39.12 |

They differ in the average discrepancy which is 0.05 and 0.14 times  $p_{\text{true}}$ , respectively. Since for arbitrary  $p$  the average discrepancy is  $\bar{\varepsilon}_a$ , the curves oscillate around a mean value  $M=1$ . For the true value  $p=1$ ,  $\bar{\varepsilon}$  has a sharp minimum. Equally deep dips in  $\bar{\varepsilon}$  occur at  $p=p_{\text{true}}/m$ ,  $m=\text{integer}$ . Of the corresponding peaks in  $M$ , however, the one at  $p=p_{\text{true}}$  is by far the largest since the numerator of (2) is inversely proportional to  $m$ . Moreover, the pseudo-solutions with  $m > 1$  will be-

tray themselves by the conspicuous property that all their indices are multiples of  $m$ .

We see, therefore, that the peak for the true value of  $p$  stands out above the pseudo-solution peaks. However trivial this result may seem, it illustrates the fundamental difference between  $M$  and the standard statistical criteria. The latter compare the discrepancies  $\varepsilon$  with an expected value  $\sigma$  of the measuring error of each line. Therefore they assign exactly equal reliability to the solutions  $p=p_{\text{true}}$ ,  $\frac{1}{2}p_{\text{true}}$ ,  $\frac{1}{3}p_{\text{true}}$  etc. This is inevitable because it is impossible to formulate an *a priori* probability distribution of  $p$ .

The quantity  $M$ , on the other hand, takes account of the amount of 'coverage', that is, the fraction of calculated lines actually observed. For a given pattern, the higher this fraction, the higher is  $p$ ; and  $M$  is proportional to  $p$  through (1) and (2). That the magnitude of the errors, too, profoundly influences  $M$  is clear from (2) as well as from Fig. 2.

A less fundamental difference between (2) and statistical criteria is the use of  $\bar{\varepsilon}$ , instead of the root mean square discrepancy. In our model the r.m.s. error might be preferable; in actual indexing, however, several circumstances (variation of  $\sigma$  as  $\theta$  changes, systematic errors, discarding of unindexable lines as impurities) tend to make the extra complication useless.

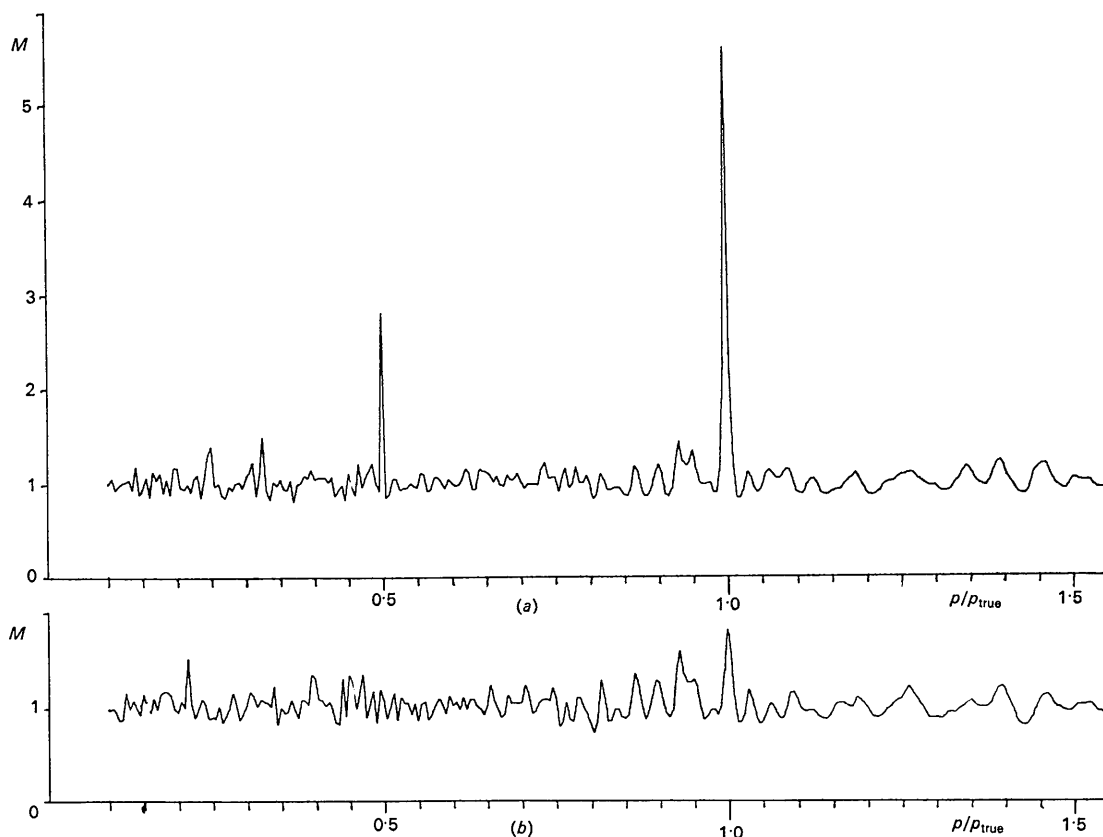


Fig. 2. The quantity  $M$  as a function of  $p$  for the synthetic one-dimensional patterns given by (a) Table 1(a) and (b) Table 1(b)

We shall not pursue the theory of the one-dimensional model any further. The question of how large  $M$  should be for a reliable indexing could be answered for this model, but we prefer to discuss it for the actual indexing later on.

### Role of $\sigma$ , the error of measurement

It may seem amazing that a statistical criterion can be formed without  $\sigma$  in it. However, the practice of indexing introduces whatever is known of  $\sigma$  in an implicit way. As a matter of fact, indexings with discrepancies much larger than  $\sigma$  (or what one believes to be an acceptable error) are either discarded as a whole, or explained by leaving out the unindexable lines and ascribing these to impurities. Thus, in practice, any indexing will automatically be such that  $\bar{\epsilon}$  is of the order of  $\sigma$ , the 'expected' error. Since, however,  $\sigma$  is hardly ever known very well – especially in its dependence on  $\theta$ , line width and intensity – it seems better to stick to the actual discrepancies in defining  $M$ .

Disregarding a few unexplained lines is sometimes condemned in principle. We think that its virtues depend upon the case; indeed experience has justified it very often, both in the literature and in our own work. The reliability of the unit cell depends, of course, on the number of such unindexable lines. In the present paper we shall express this by stating  $X_{20}$ : the number of unindexable lines occurring up to the 20th observed and indexed line.

### Definition of $M_{20}$

In actual indexing, the calculated line positions are no longer equidistant. The interval between two successive calculated lines (or intrinsic multiplets, such as  $hkl$  and  $h\bar{k}l$  in the monoclinic system) has a distribution which closely approximates the exponential distribution of free path lengths (de Wolff, 1961*b*). It is characterized by a single parameter, for which one may take the average interval length  $\bar{p}$ . Moreover,

this distribution is valid only in a narrow range of  $\theta$ : the parameter of the exponential function is itself a function of  $\theta$ .

If we neglect the second circumstance – or if we imagine the lines to be mapped on a  $\varphi$  scale, such that the parameter  $\bar{p}$  is independent of  $\varphi$  – there remains the statistical fluctuation of the intervals. Because of this, the average discrepancy for an arbitrary unit cell is now equal not to  $\frac{1}{4}\bar{p}$  but to  $\frac{1}{2}\bar{p}$ . The extra factor 2 is a well-known paradoxical result of statistics, explained by the larger chance for an observed line to sit in a large interval as compared with sitting in a small one.

To carry out the mapping just mentioned would be impractical; a different  $\varphi$  scale would be needed for each choice of unit-cell parameter. However, mapping is not needed if we choose a scale such that the variation of  $\bar{p}$  is not too large. (This is the case for the quantity  $Q = 1/d^2$ , somewhat less so for  $1/d$  and for  $\theta$ , but not at all for the  $d$  scale. Hence the reliability should on no account be judged from discrepancies in  $d$ -value.) Then  $\bar{p}$  can be defined as the mean interval in the range of observed lines. Taking the  $Q$  scale, this means that, if the first 20 observed lines are considered,

$$\bar{\epsilon}_a = \frac{1}{2}\bar{p} = Q_{20}/2N_{20},$$

where  $Q_{20}$  is the value of  $Q$  for the 20th observed line (not counting unexplained lines), and  $N_{20}$  is the number of different calculated  $Q$  values up to  $Q_{20}$ . It follows from (2) that for these 20 lines the value of  $M$  is

$$M_{20} = Q_{20}/2\bar{\epsilon}N_{20}. \quad (3)$$

The 'coverage' mentioned before as an implicit factor of  $M$  is clearly apparent in (3). Actually the observed fraction of calculated lines is  $20/N_{20}$ .

The number 20 merely serves to make estimates of reliability of different patterns comparable. This restriction, is, of course, arbitrary. We found 20 to be a useful range. The number of observed lines should be reasonably large in relation to the number (up to 6)

Table 2. Examples of (a) correct, (b) incorrect and (c) unconfirmed indexings

| No.             | Compound  | $10^4 \cdot Q_{20}$ | $10^4 \cdot \bar{\epsilon}$ | $N_{20}$ | $X_{20}$ | $M_{20}$ |
|-----------------|---|---------------------|-----------------------------|----------|----------|----------|
| (a) Correct     |   |                     |                             |          |          |          |
| 1               | $\alpha$ -Li <sub>4</sub> B <sub>2</sub> O <sub>5</sub> | 1680                | 0.5                         | 26       | –        | 60       |
| 2               | UO <sub>2</sub> WO <sub>4</sub>                         | 1304                | 0.6                         | 35       | 3        | 33       |
| 3               | $\gamma$ -Cd(OH) <sub>2</sub>                           | 3669                | 4.0                         | 25       | 0        | 18       |
| 4               | K <sub>2</sub> RuCl <sub>5</sub> ·NO                    | 1271                | 1.6                         | 30       | 2        | 13       |
| 5               | $\alpha$ -Li <sub>4</sub> B <sub>2</sub> O <sub>5</sub> | 2128                | 2.0                         | 40       | 5        | 13       |
| 6               | $\beta$ -Li <sub>4</sub> B <sub>2</sub> O <sub>5</sub>  | 2993                | 1.9                         | 70       | 0        | 10       |
| 7               | NaB <sub>5</sub> O <sub>8</sub> ·5H <sub>2</sub> O      | 1082                | 1.6                         | 50       | 0        | 6.7      |
| 8               | Li <sub>6</sub> B <sub>4</sub> O <sub>9</sub>           | 1703                | 2.6                         | 63       | 2        | 5.3      |
| (b) Incorrect   |   |                     |                             |          |          |          |
| 9               | Li <sub>6</sub> B <sub>4</sub> O <sub>9</sub>           | 1703                | 3.0                         | 52       | 2        | 5.4      |
| 10              | $\gamma$ -UO <sub>3</sub>                               | 1357                | 1.3                         | 180      | 1        | 3.0      |
| 11              | Li <sub>3</sub> BO <sub>3</sub>                         | 2510                | 2.3                         | 210      | 0        | 2.6      |
| 12              | $\gamma$ -Cd(OH) <sub>2</sub>                           | 3911                | 4.4                         | 310      | 3        | 1.5      |
| (c) Unconfirmed |   |                     |                             |          |          |          |
| 13              | $\alpha$ -Ta <sub>2</sub> O <sub>5</sub>                | 1621                | 0.6                         | 38       | 0        | 36       |
| 14              | KTa <sub>5</sub> O <sub>13</sub>                        | 1484                | 0.6                         | 33       | 0        | 37       |

Table 2 (cont.)

| No. | Data*  |   | Indexed with constants |         |         | $\beta$ | by*         |
|-----|--------|---|------------------------|---------|---------|---------|-------------|
|     |        |   | $a$                    | $b$     | $c$     |         |             |
| 1   | Vi     | Orthorhombic primitive  | 10.20 Å                | 17.58 Å | 4.730 Å |         | s.c.        |
| 2   | Vi     | Monoclinic primitive  | 13.60                  | 5.489   | 7.22    | 104.49° | Vi          |
| 3   | 12-62  | Monoclinic <i>I</i> -centred  | 5.67                   | 10.22   | 3.407   | 91.52   | H, W        |
| 4   | 15-46  | Orthorhombic primitive  | 10.36                  | 13.30   | 6.90    |         | V, W        |
| 5   | 12-130 | Hexagonal (cf. No. 1)   | 20.32                  |         | 4.74    |         | V, W        |
| 6   | 12-128 | Monoclinic <i>I</i> -centred  | 10.23                  | 4.698   | 8.78    | 93.54   | H, V        |
| 7   | 12-265 | Monoclinic <i>A</i> -centred  | 13.63                  | 16.44   | 11.09   | 112.73  | V, W        |
| 8   | 12-129 | Monoclinic primitive  | 9.18                   | 23.41   | 3.32    | 92.68   | s.c.        |
| 9   | 12-129 | Orthorhombic <i>B</i> -centred  | 12.30                  | 13.37   | 11.87   |         | W           |
| 10  | See C  | Monoclinic  |                        |         |         |         | C           |
| 11  | 12-127 | Triclinic   |                        |         |         |         | unpublished |
| 12  | 12-62  | Orthorhombic  |                        |         |         |         | unpublished |
| 13  | 9-360  | Pseudo-tetragonal <i>I</i> -centred<br>actually triclinic,<br>$\alpha = 90.10^\circ$ and $\gamma = 90.01^\circ$ | 3.78                   | 3.80    | 35.75   | 90.90   | V, W        |
| 14  | 12-90  | Orthorhombic primitive  | 10.71                  | 16.79   | 5.651   |         | W           |

\* The numbers under *Data* are card numbers in the Powder Data File of the ASTM. The other symbols mean:

|         |   |
|---------|---|
| C       | Connolly (1959)                                   |
| Vi      | Visser (1966)                                     |
| H, V, W | Contributions to a 'computer indexing project' by |
| H       | Professor H. Haendler, Durham N.H., U.S.A.        |
| V       | Dr J. Villadsen, Vedbaek, Denmark                 |
| W       | Present author                                    |
| s.c.    | Indexing derived from single-crystal data         |

#### Remarks on Table 2

- (1) and (5) The indexing of No. 5 is interesting: the unit cell is not pseudo-hexagonal but it was indexed as hexagonal because the difference between  $b$  and  $a/\sqrt{3}$  is nil. Apart from this systematic error, the indexing of No. 5 was correctly found, though the second strongest line of the pattern belongs to a contamination! No. 1 is a pattern measured by Dr J. W. Visser after he had established the true unit cell by single-crystal work; therefore the number  $X_{20}$  is immaterial. The result is listed as an example of routine accuracy and is to be compared with No. 5, data taken from the literature.
- (2) Diffractometer data of an impure phase, indexed by Visser (1966), and later identified with the phase described with single-crystal data by Juenke & Bartram (1964).
- (3) and (12) Data originally given as diffraction angles to  $0.1^\circ$ . Indexing (12) was obtained with a computer program not able to cope with monoclinic cells. No. 3 was confirmed by structure analysis (de Wolff, 1966*b*) and single-crystal work (Oswald, 1964).
- (4) The first two lines are from a contamination. The cell was later confirmed by the single-crystal work of Khodashova & Boky (1960).
- (6) Confirmed by single-crystal work by Visser (1966).
- (7) After indexing, Villadsen discovered a single-crystal structure analysis by Sabelli (1962), agreeing with his and our cell constants.
- (8) and (9) are, respectively, the correct and an incorrect indexing of the same data, listed in Table 3. The correct unit cell was derived from single-crystal work by Visser (1966).
- (10) The correct unit cell for the same data gives  $M_{20} = 8$  (de Wolff, 1961*a*). It was confirmed by structure analysis (Engmann & de Wolff, 1963).
- (11) The true, monoclinic primitive unit cell ( $a = 8.337$ ,  $b = 9.179$ ,  $c = 3.260$ ,  $\beta = 101.59^\circ$ ) was not found from the data (for which it yields  $M_{20} = 10$ ,  $X_{20} = 0$ ) but from single-crystal work by Visser (1966).

of unit-cell parameters. All the same, including many more lines does not raise the significance of  $M$  very much. This is so because, if we keep to the  $Q$  scale,  $\bar{\epsilon}_a$  always decreases with increasing  $Q$ . In moderately complex patterns, even the medium angle region tends to be overcrowded with calculated lines. On the other hand, the error of measurement tends to increase with increasing  $Q$ . In other words, the indexing of high-angle lines is always much easier – and therefore adds much less to the reliability – than for low-angle lines.

#### The critical value of $M_{20}$

A hypothesis such as the correctness of a unit cell can never be verified with absolute certainty. In testing hypotheses which are more amenable to theoretical

treatment, limits for certain criteria can be set in relation to a 'percentage confidence'. The present problem is too complex for such a treatment. Therefore we can rely only on experience; this has now accumulated to a degree where it may yield some idea of a 'critical' value. The following discussion refers to Table 2, where the values of  $Q_{20}$ ,  $\bar{\epsilon}$ ,  $N_{20}$  and of  $M_{20}$  and  $X_{20}$  are listed for a number of cases.

#### (a) Correct indexings

The term 'correct' is used for results which have been confirmed by single-crystal investigations. The values of  $M_{20}$  and  $X_{20}$  are listed for a number of such cases in Table 2. They show, for instance, the high values obtainable with routine diffractometer or Guinier-camera measurements. On the other hand, lower figures occur

which are interesting in so far as these patterns have been solved correctly (without knowledge of the single-crystal data) by several authors. Of course these lower values of  $M$  have no relation to the critical limit for reliability. In theory,  $M_{20}$  could be as low as 1 for a correctly indexed pattern – only it is highly improbable that the correct solution would ever be derived from so inaccurate and/or incomplete data. The lower limit of  $M_{20}$  for an indexing which is both correct and retrievable seems to lie around  $M_{20}=6$ , and for  $M_{20}=13$  even quite a few impurity lines ( $X_{20}=5$ ) appear not to render the solution impossible.

(b) *Incorrect indexings*

Here 'incorrect' means disproved by single-crystal investigation and/or successful structure determination. Now the range of  $M_{20}$  is a straightforward warning signal: any indexing falling within this range is untrustworthy. Again we must observe, however, that there is no absolute confidence or inconfidence. For instance, it is difficult to see why incorrect indexings could not occasionally have very high  $M_{20}$  values. In practice, fortunately, it appears that most of them do not exceed  $M_{20}=3$ . The value of 5 for  $\text{Li}_6\text{B}_4\text{O}_9$  is quite an exception. It is therefore described in full in Table 3. Just as with several other incorrect indexings of Table 2, there is not the remotest analogy between correct and incorrect reciprocal lattice; there does not even exist a common zone. It is remarkable that correct indexing of the same data yields almost the same value of  $M_{20}$ , and the same value of  $X_{20}$ . The intensities, however, gave a clear indication because the incorrect indexing misses the strongest line of the pattern! A similar failure to explain a strong line weighed heavily against example No. 10, which was eventually disproved – but a strong line can also be missed by a correct indexing, as shown by No. 5.

Finally, Table 2 includes some (as yet) unconfirmed results with very high values of  $M_{20}$ . We thought it worth while to point to the existence in the literature of such excellent patterns, indexable with a precision and completeness which in itself is fully convincing. The only remaining uncertainty in such cases is a possible oversight of pseudo-extinctions, if the structure happens to have a predominant subcell. This, however, is also true for the corresponding reflexions in single-crystal work, so the difference from the results of the latter is only one of degree. In the higher symmetries, another kind of misinterpretation is possible, of which No. 5 of Table 2 is an example (*cf. Remarks*).

Apart from such 'systematic' errors, it seems safe to conclude that a value of  $M_{20}=10$  or higher will guarantee the essential correctness of the indexing,\* pro-

\* This conclusion receives additional support from a private communication by Dr Villadsen, which he kindly permitted us to quote: 'In applying this criterion to very many indexings of patterns taken from the literature, no case has yet been observed where an indexing proved to be wrong with  $M_{20}$  larger than 10'.

Table 3. Powder data for  $\text{Li}_6\text{B}_4\text{O}_9$   
(*cf.* Table 2, Nos. 8 and 9)

| $10^4 Q$ from<br>card 12-129 | Correct<br>indexing |            | Incorrect<br>indexing |            |
|------------------------------|---------------------|------------|-----------------------|------------|
|                              | $hkl$               | $10^4 Q_c$ | $hkl$                 | $10^4 Q_c$ |
| 192.9                        | 120                 | 191.9      | 111                   | 193        |
| 286.3                        | 130                 | 283.1      | 002                   | 284        |
| 358.7                        | –                   | –          | 121                   | 362        |
| 548.5                        | 220                 | 548.7      | 201                   | 547        |
| 644.4                        | 230                 | 640.0      | 131                   | 642        |
| 776                          | 160                 | 775.6      | 222                   | 772        |
| 886                          | –                   | –          | 321                   | 888        |
| 924                          | 011                 | 924.5      | 123                   | 930        |
| 995                          | 10 $\bar{1}$        | 994.5      | –                     | –          |
| 1061                         | 101                 | 1055.8     | 232                   | 1052       |
| 1075                         | 11 $\bar{1}$        | 1074.1     | –                     | –          |
| 1132                         | 260                 | 1132.5     | 004                   | 1136       |
| 1161                         | 13 $\bar{1}$        | 1158.7     | 240                   | 1163       |
| 1233                         | 330                 | 1234.7     | 303                   | 1233       |
| 1287                         | 180                 | 1286.4     | 313                   | 1289       |
| 1366                         | 340                 | 1362.4     | 024                   | 1361       |
| 1393                         | 22 $\bar{1}$        | 1393.6     | 412                   | 1393       |
| 1455                         | 15 $\bar{1}$        | 1450.5     | 214                   | 1455       |
| 1560                         | 061                 | 1562.9     | 422                   | 1561       |
| 1606                         | 231                 | 1607.5     | 143                   | 1607       |
| 1646                         | 280                 | 1643.2     | 034                   | 1645       |
| 1703                         | 161                 | 1712.5     | 501                   | 1715       |

vided there are few spurious lines ( $X_{20}$  not above 2). On the other hand, values below 6 must be regarded with suspicion and below 3 as hardly significant at all, though, again, this does not mean that such an indexing could not be correct!

**Other factors influencing the reliability**

(a)  $X_{20}$  is obviously a very rough measure of the contamination. To what extent it discredits the indexing depends also on the intensity of the unexplained lines, as discussed in the foregoing section.

On the other hand, complete or even partial explanation of these lines by the pattern of an acceptable impurity may lend much weight to the indexing – again, however, not expressible quantitatively.

(b) The discrepancies – and therefore the value of  $M$  – depend to some extent upon the exact values chosen for the cell parameters. For the first 20 lines, this effect is usually not very strong. Indeed, refinement of parameters will appreciably raise  $M_{20}$  if it is already high, but it will hardly improve doubtful values, as will be evident by inspecting Fig. 2 for the one-dimensional model.

(c) Indices triples falling under obvious general extinctions (*e.g.* for centred lattices) should of course be disregarded in counting  $N_{20}$ . Special extinctions (mainly for glide planes) play quite a different rôle because their selective nature makes the indexing more convincing. A conservative way to express this is to subtract special extinctions, too, from the triplets to be counted.

(d) Normally the number of close doublets among the  $N_{20}$  calculated lines is small. In cases of pseudo-symmetry, those doublets which would be unobservable as such may be counted as one line. If doublets are recognizable in the pattern, e.g. by systematic intensity ratios, a corresponding indices relation adds a good deal to the reliability.

(e) The same is true for a good fit with an observed density (provided  $Z$  is evident) and for several other circumstances, such as good agreement with texture effects. Sometimes this circumstantial evidence is so strong as to make other criteria superfluous.

Still there remain quite a number of investigations (e.g. those carried out at non-room temperatures and/or pressures) where such extraneous support is usually not available. In computer indexing, too, one cannot always account for non-powder data, but a criterion like  $M_{20}$  can easily be programmed. It has, indeed, proved to be a useful test in an indexing program written and frequently applied by Visser (1968).

The author is much indebted to Drs Visser (T.P.D., Delft) for providing many of the data used; to Ir W.

Peterse of this laboratory for preparing the plotting program for Fig. 1; and to Professor Haendler (Durham) and Dr Villadsen (Vedbaek, Denmark) for permission to use their contribution to a 'computer indexing project' in which they participated in 1964.

#### References

- CONNOLLY, D. E. (1959). *Acta Cryst.* **12**, 949.  
 ENGMANN, R. & DE WOLFF, P. M. (1963). *Acta Cryst.* **16**, 993.  
 JUENKE, E. F. & BARTRAM, S. E. (1964). *Acta Cryst.* **17**, 618.  
 KHODASHOVA, T. S. & BOKY, G. B. (1960). *Zh. Strukt. Khim.* **1**, 151.  
 OSWALD, H. R. (1964). Private communication.  
 SABELLI, C. (1962). *Atti Soc. Toscana Sci. Nat. Pisa, Proc. Verbali*, Mens. Ser. A **69** (2), 299.  
 VILLADSEN, J. (1968). Private communication.  
 VISSER, J. W. (1966). Private communication of data, some to be published in the X-Ray Powder Data File of ASTM.  
 VISSER, J. W. (1968). To be published.  
 WOLFF, P. M. DE (1961a). *Acta Cryst.* **14**, 322.  
 WOLFF, P. M. DE (1961b). *Acta Cryst.* **14**, 579.  
 WOLFF, P. M. DE (1966a). *Acta Cryst.* **21**, A11.  
 WOLFF, P. M. DE (1966b). *Acta Cryst.* **21**, 432.

*J. Appl. Cryst.* (1968). **1**, 113

## Etude Cristallographique de Quelques Fluorures Complexes de Terres Rares de Formule $A_2NaTF_6$

PAR SUZANNE ALÉONARD ET CHANTAL POUZET

Laboratoire d'Electrostatique et de Physique du Métal, Grenoble, Isère, France

(Reçu le 7 avril 1968)

$Rb_2NaTF_6$  and  $Cs_2NaTF_6$  fluorides, with  $T=Yb, Er, Ho, Y, Tb$  and  $Sm$ , belong to the cubic system [space group  $Fm\bar{3}m (O_h^h)$ ]. Lattice parameters (near to 9 Å), indices, intensities and spacings are tabulated for the whole series. They are ordered perovskites. The cations are in parameterless positions. The parameters of the  $F^-$  ions have been determined in  $Rb_2NaErF_6$  and  $Cs_2NaErF_6$ .

### Introduction

Récemment, lors d'une étude sur les fluorures complexes du cérium, Besse & Capestan (1968) signalent l'existence du composé  $Cs_2NaCeF_6$ : selon ces auteurs,

sa maille cubique, de paramètre  $a=9,26$  Å, appartiendrait au groupe d'espace  $Pa\bar{3} (T_h^6)$ . A notre connaissance, c'est le seul composé  $A_2NaTF_6$  étudié jusqu'ici.

Etant donné que les fluorures  $NaTF_4$  présentent à haute température la structure fluorine cubique (Hund,

Tableau 1. Caractéristiques cristallographiques des composés  $Rb_2NaTF_6$  et  $Cs_2NaTF_6$

| Formules      | Paramètres de mailles,<br>$a$ | Volume moléculaire      | Densité,<br>$d$ |
|---------------|-------------------------------|-------------------------|-----------------|
| $Rb_2NaYbF_6$ | $8,824 \pm 0,002$ Å           | $171,76$ Å <sup>3</sup> | 4,65            |
| $Rb_2NaErF_6$ | $8,867 \pm 0,001$             | 174,28                  | 4,53            |
| $Rb_2NaHoF_6$ | $8,881 \pm 0,001$             | 175,11                  | 4,49            |
| $Rb_2NaYF_6$  | $8,8693 \pm 0,0005$           | 174,40                  | 3,78            |
| $Rb_2NaTbF_6$ | $8,9208 \pm 0,0005$           | 177,48                  | 4,37            |
| $Rb_2NaSmF_6$ | $8,988 \pm 0,001$             | 181,52                  | 4,19            |
| $Cs_2NaYbF_6$ | $9,022 \pm 0,001$             | 183,59                  | 5,21            |
| $Cs_2NaErF_6$ | $9,061 \pm 0,001$             | 185,98                  | 5,09            |
| $Cs_2NaYF_6$  | $9,056 \pm 0,001$             | 185,67                  | 4,39            |
| $Cs_2NaTbF_6$ | $9,107 \pm 0,001$             | 188,82                  | 4,94            |
| $Cs_2NaSmF_6$ | $9,163 \pm 0,002$             | 192,33                  | 4,78            |