RIFRAN 85 — A set of programmable calculator programs for powder diffraction phase identification

ᵃT. HAVLÍK, ᵃM. ŠKROBIAN, and ᵇM. HAVLÍK

ᵃ*Department of Nonferrous Metallurgy, Metallurgical Faculty of the Technical University, CS-043 85 Košice*

ᵇ*Research Laboratory of Metallic Materials, Technical University, CS-043 85 Košice*

A small, effective and practical identification system, RIFRAN 85, has been developed and implemented on a programmable calculator EMG 666/B (product of Hungary) with 8 kbytes of operational memory. These interactive calculator programs have been used for qualitative phase analysis of X-ray diffraction patterns using standard files created by the authors, which were collected from various published data. The computer programs are written in the EMG assembly language.

A specialized standard database stored on magnetic compact cassette tape is used for phase identification. Each standard pattern comprises 35 diffraction lines represented by 2Θ and I_r values. The identification procedure is based on the comparison of diffraction patterns of standard and unknown sample involving the chemical and match criteria.

Была разработана небольшая, эффективная и практическая идентификационная система RIFRAN 85, которая была применена на настольном программируемом калькуляторе EMG 666/B (Венгрия) с объемом оперативной памяти 8 кбайт. Эти интерактивные калькуляторные программы используются для качественного дифракционного рентгеновского фазового анализа с применением стандартной датабазы, разработанной авторами по различным литературным источникам. Программы написанные кодом машины EMG.

Используемая специализированная стандартная датабаза уложена на магнитофонной ленте. Каждый стандарт описан 35 дифракционными линиями, представленными величинами 2Θ и I_r. Идентификация основана на сравнении дифрактограмм стандартов и неизвестного образца, а также на химическом и критериальном анализе.

At present, the reference database of X-ray powder diffraction patterns comprises large number of data, its effective utilization requires therefore data-processing methods.

Frevel [1] and *Nichols* [2] were among the first to publish the programs for qualitative phase identification on the basis of X-ray powder diffraction patterns using, at an early stage, only a subset of the file of standards. However, the newly developed search/match methods pioneered by *Johnson* and *Vand* [3] have soon allowed the users to go through the whole database. The effort of numerous researchers dealing with phase identification resulted in the programs which with more or less success meet the user's requirements [4]. The programs were published by *Huang* and *Parrish* [5], *Johnson* [6, 7], *Lin Tian-Hui et al.* [8], *O'Connor* and *Bagliani* [9], *Fiala* [10], *Schreiner et al.* [11], and *Goehner* and *Garbauskas* [12]. *Nichols* and *Johnson* [13] have compared several programs whereas the results of round robin test have been published by *Jenkins* and *Hubbard* [14]. Step by step, with the appearance of minicomputers, the individual programs have been tailored and improved so that they could be implemented in this type of computers.

A perfect computer is not always available and running such a computer is usually expensive. Microcomputers and programmable calculators, on the other hand, have become available in every research laboratory. Furthermore, highly specialized research centres do not always need a complete file of standards, a minifile comprising only the selected standards can fully satisfy their requirements. The above considerations have led to the development of identification system for less sophisticated computers, such as the programs by *Edmonds* [15], *Hare et al.* [16], *Lindsey et al.* [17], *Škrobian et al.* [18]. These systems are easy to transfer and sufficiently flexible and may therefore be run on calculators installed directly almost in all laboratories.

The number of standards that we routinely encounter in our research is limited to the inorganic substances related to extractive metallurgy. With regard to the fact that our research laboratory is equipped with a programmable calculator EMG 666/B with 8 kbytes of the memory, the program RIFRAN was developed in which the likely phases were selected on the basis of a complete match of four strongest peaks of the unknown diffraction pattern and the standard pattern [18]. This version was subsequently developed into search/ /match system RIFRAN 85, which is described below.

Standard database

Due to rather limited memory capacity of the calculator, which corresponds to maximum of 1000 digits or 500 ASCII registers, the standard database has been divided into several subfiles and recorded on a magnetic tape. Each subfile comprised twenty standards. Every standard occupied 200 bytes and comprised the following data (Fig. 1).
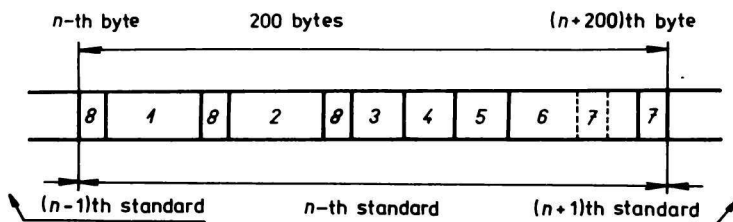
122

Fig. 1. Standard data sequence on magnetic tape.
1. Identification code; 2. name of mineral, data quality, lattice type; 3. chemical formulae; 4. characteristic reflections; 5. atomic numbers; 6. 2Θ and I_r values; 7. terminal characters of datafile; 8. format codes.

Identification code

With regard to the fact that RIFRAN 85 is written mainly for our use, its database contains, apart from the ASTM standards, other inorganic substances the powder diffraction data of which were published in different journals or specialized monographs [19—22]. The identification code of a standard consists of a capital letter, identifying the reference, and six digits, indicating the origin of the reference, *e.g.*

A 25 983   (ASTM 25 983) [19]
M 310   (*Mikheev*, Table 310) [22]
B 71.80   (*Brindley* and *Brown*, page 71, 1980) [21]

The database compiled in this manner represents the RIFRAN 85 minifile system. The entire database of standards is divided into several subfiles according to their relevance to a particular research topic. Thus, we have created subfiles of standards relevant to extractive metallurgy of copper, iron, nickel, aluminium, steelmaking, *etc.* The subfiles may obviously be arbitrarily enlarged or changed while the rest of data remain unchanged.

Name of mineral

This part of information identifies the mineral or chemical compound, crystallographic system (lattice type) and finally the data quality in agreement with the Joint Committee for Powder Diffraction Standards (JCPDS), *e.g.*
PERICLASE (*)   (cubic)
TITANIUM HYDRIDE (I)   (tetragonal b.c.)

Chemical formulae

This information is very useful in case of nonstoichiometric compounds and solid solutions as it may serve as a basis for a priori elimination of some standards from among the potential candidates, *e.g.* TiH_2, $TiH_{1.924}$, $TiH_{1.723}$.

Identification peaks

The first fundamental criterion, on the basis of which the standard may be rejected from subsequent search/match data, *i.e.* from the list of likely phases, is the complete match of one or two characteristic peaks of the standard and sample patterns. Two strongest peaks have been selected as identification peaks and their position within the file is given in degrees of 2Θ, each peak occupying two bytes. In the first byte the integer part of 2Θ value is stored while the second contains the decimal fraction (if any) made integer by multiplying by 100.

In this way the round-off error is kept to be less than $0.01°\ 2\Theta$. The maximum error in *d*-spacing for CuKα radiation varies with 2Θ angles in a manner shown in Fig. 2.
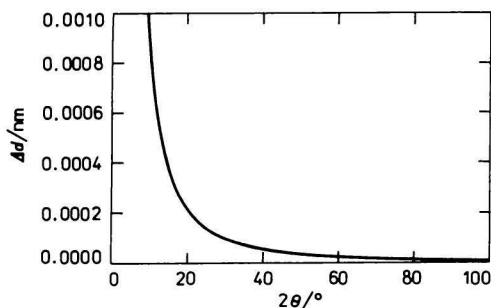

Fig. 2. Error *d vs.* 2Θ plot.

Atomic numbers

Another important criterion which is decisive in consideration of the likelihood of the standard being present in an unknown sample, is the discrepancy between the elemental composition of the standard and sample. Elemental restrictions can be imposed on the basis of preliminary chemical analysis or an a priori assumptions. The chemical composition of each standard is described by the atomic numbers of its constituent elements.

Positions of the peaks and their relative intensities

In order to determine further criteria and hence to increase the accuracy of phase analysis, it is necessary to know the position and the relative intensities of as many peaks as possible. A maximum of thirty-five peaks have been used for identification with their relative intensities ranging from zero to 100. If the standard contains more than 35 peaks, arbitrary peaks may be ignored, *e.g.* the weakest, depending on the type of standard.

Terminal characters of the datafile

As not every standard pattern comprised 35 peaks (*e.g.* metals with a cubic lattice), the datafile of each standard pattern ends with three zeroes which ensure the termination of search.

Format codes

These are inherent to the computer language, *i.e.* machine code.

Searching

A fundamental criterion that the standard pattern must fulfil in order to become a possible candidate is a complete match of one or two characteristic peaks, according to decision of the user. This method was with more or less success used in the original version of the program RIFRAN [18]. If the first criterion is fulfilled, the chemical composition of the likely standards is then checked by elemental restrictions imposed in terms of absent elements (the so-called "negative elements" [3]). Possible discrepancies in chemical composition are not the reason for excluding a standard from further match process but a warning is always given (Table 1). The standards with correct elemental composition of course pass on to further match test.

Match method

Two match criteria have been used. The first ($P/\%$), was the fraction of the matching peaks of the total possible number of matches within the experimental window. The fraction of the sum of relative intensities of matched standard and sample peaks of the sum of possible relative intensities of all standard peaks within the experimental window, $I/\%$, is another criterion. None of the above match criteria is sufficient [10] but their combination enables us to make a final decision concerning the likelihood of the standard being present in the unknown sample.

An illustration of how $P/\%$ match criterion may fail, if it is considered alone, is Cristoballite low, JCPDS Powder Diffraction File (PDF) index 11-695 [19]. Within the 2Θ range 20—90°, Cu$K\alpha$ radiation, the standard comprises 35 peaks with the sum of available intensities SAI = 272. Consider the case that cristoballite is a minor component and let the strongest peaks be searched, *i.e.* $d_1 = 0.405$ nm, $I_{1,r} = 100$ and $d_2 = 0.2485$ nm, $I_{2,r} = 20$. The $P/\%$ value will be then equal to 5, *i.e.* two from 35 possible matches, although the $I/\%$ value will be 44, *i.e.* 120 from 272 possible matches.

It is obvious that the final decision depends on the user, the decision, however, must be substantiated by significant values of $P/\%$, $I/\%$, their combination and the user's experience with the program. This type of complex evaluation is, however, unfeasible because of small memory capacity.

Table 1

Output of identification

——————————————— * RIFRAN 85 * ———————————————

PARAMETERS OF IDENTIFICATION:

ERROR WINDOW:	1.5
BLOCK OF STANDARDS:	2
NEGATIVE ELEMENTS:	Mg Cr Co Ni S C
2TH(OBS)MIN:	12.00
2TH(OBS)MAX:	65.00

IF AGREEMENT OF SAMPLE AND STANDARD IS ONLY IN ONE PEAK
— NO OUTPUT!

INPUT DATA:

LABEL:　　　BAUXITE SEPARATE

2TH(OBS)	D(OBS)/nm	2TH − (OBS)	2TH + (OBS)
12.28	0.720	12.15	12.41
14.45	0.612	14.32	14.58
18.35	0.483	18.22	18.49
19.90	0.446	19.77	20.04
20.60	0.431	20.47	20.74
21.35	0.416	21.22	21.49
24.95	0.357	24.82	25.09
25.25	0.352	25.12	25.39
28.15	0.317	28.01	28.29
35.00	0.256	34.86	35.14
36.00	0.249	35.86	36.14
37.70	0.238	37.56	37.84
38.26	0.235	38.12	38.40
39.40	0.229	39.26	39.54
40.40	0.223	40.26	40.54
41.74	0.216	41.60	41.88
44.50	0.203	44.36	44.64
45.65	0.199	45.51	45.79
47.82	0.190	47.68	47.97
48.15	0.189	48.01	48.30
48.90	0.186	48.76	49.05
49.20	0.185	49.05	49.35
50.89	0.179	50.74	51.04
51.55	0.177	51.40	51.70

Table 1 (Continued)

2TH(OBS)	*D*(OBS) / nm	2TH − (OBS)	2TH + (OBS)
52.50	0.174	52.35	52.65
54.90	0.167	54.75	55.05
55.25	0.166	55.10	55.40
60.55	0.153	60.40	60.70
62.25	0.149	62.10	62.41
62.47	0.149	62.32	62.63
64.10	0.145	63.94	64.26

OUTPUT DATA:

PDF	MINERAL		(SYSTEM)				FORMULA
	NIP	TAP	*P*/%	SII	SAI	*I*/%	
A211307	BOEHMITE (*)		(ORTHORHOMBIC)				AlOOH GAMMA
	10	11	91	323	333	97	
A 12460	GIBBSITE (*)		(MONOCLINIC)				Al(OH)3 ALPHA
	11	25	44	198	322	61	
A211272	ANATASE (*)		(TETRAGONAL)				TiO2
	4	9	44	159	233	68	
A 60221	KAOLINITE 1MD (*)		(PSEUDOMONOCLINIC)				Al2Si2O5(OH)4
	8	14	57	620	890	70	
A 15580	GARNIERITE ( )		(MONOCLINIC)				(Ni,Mg)Si2O5(OH)4
	CHEMICALLY INCORRECT						

———————— END OF IDENTIFICATION ————————

Input data

The first step is entering the input data and parameters of identification. The observed values of peaks positions may be expressed either in terms of interplanar spacing, *d*, or 2Θ angles. The incident Cu*K*α radiation is assumed, but this value can be easily changed, if necessary. The end of the input data of an unknown pattern is indicated by the end code (zero in this case). The next step is the determination of the experimental window (in terms of $2\Theta_{min}$ to $2\Theta_{max}$) followed by the sample label comprising an arbitrary text with maximum 16 ASCII characters.

Identification parameters, *i.e.* the total number of tested subfiles and the ordinal number of the subfile are given in the same sequence as recorded on the

cassette. Commercial compact cassette CC60 with given record density is equivalent to approximately 120 kbytes, *e.g.* 30 subfiles, each containing 20 standards.

The next step is the selection of the error limit expressed in terms of Δd at a certain d to define the windows for matching patterns of a standard to the unknown sample. This parameter, called the error window, EW, is taken into account in the program considering the equation

$$\Delta d = d + d^2 \cdot \text{EW} \cdot 10^{-3} \tag{1}$$

The magnitude of EW depends on the precision and accuracy of the sample data (2Θ or d observed), in our case the majority of values were in the range between 1 and 2, which for the angles between 10 to 90° 2Θ corresponded to approximately 0.1—0.2° 2Θ.

Furthermore, atomic numbers of absent elements are entered. The list of absent elements may comprise up to seven items. Finally, the decision is made whether the standards with only one matching peak will be selected or rejected from the subfile of potential candidates. The error windows for all observed peaks are then calculated from eqn (*1*) and the Bragg equation. Subsequently the first subfile is transferred from the cassette to the memory. Each standard subfile is searched sequentially. Each standard pattern is examined by selected pre-screens, *i.e.* the characteristic peaks, elemental composition, $P/\%$ and $I/\%$.
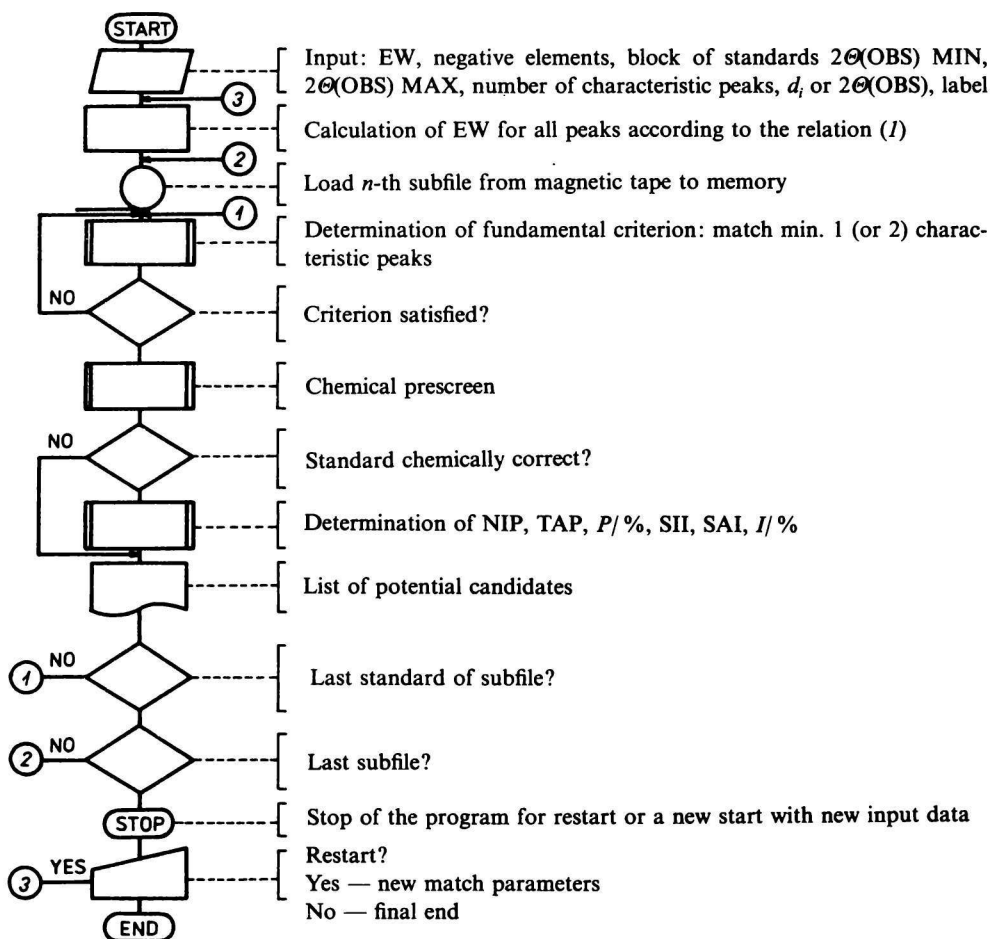
Output data

At the completion of search/match process the calculator prints the output starting with the identification parameters and the input data (sample number and error limit). Rejection of standard with only one matching line is signalled by the text. The output data are printed as the list of potential candidates. The first line of the table contains the PDF identification number followed by the name of mineral, accuracy of standard pattern, lattice type, and chemical formulae. The second line gives the number of matches in a given experimental window (TAP), $P/\%$, the sum of the relative intensities of possible matches (SAI), and $I/\%$. Incorrectness of the elemental composition is indicated by printing. Termination of identification process is indicated by the final text.

Hardware and software limitations

Due to considerable limitations imposed by the small capacity of the memory of the calculator, the data were stored and handled on the byte level. Also the

access to the data (characters and numbers) has been made *via* bytes. This type of data processing enabled us to process 4000 bytes of data in one cycle although the calculator comprised only 1008 memory registers, a considerable portion of which is occupied by the program, input data, and identification parameters. Searching and matching of a single standard preselected for the list of potential candidates on the basis of four or less matches out of 35 characteristic peaks, takes about 50 s. The first cycle of identification, *i.e.* searching, takes several seconds. Of similar duration is also the chemical prescreen. Operations involving cassette unit and typewriter (200 characters per min) are the slowest steps of identification process.

Fig. 3. Flowchart of the program.

Input: EW, negative elements, block of standards $2\Theta(OBS)$ MIN, $2\Theta(OBS)$ MAX, number of characteristic peaks, d_i or $2\Theta(OBS)$, label

Calculation of EW for all peaks according to the relation (*1*)

Load *n*-th subfile from magnetic tape to memory

Determination of fundamental criterion: match min. 1 (or 2) characteristic peaks

Criterion satisfied?

Chemical prescreen

Standard chemically correct?

Determination of NIP, TAP, $P/\%$, SII, SAI, $I/\%$

List of potential candidates

Last standard of subfile?

Last subfile?

Stop of the program for restart or a new start with new input data

Restart?
Yes — new match parameters
No — final end

Searching and matching of several subfiles (according to user's requirements) is carried out in such a way that the standard data from the first subfile are transferred from the cassette to the given registers of the memory of the calculator. The search/match process is then carried out together with necessary printout and criteria analysis of potential candidates. Potential candidates are listed in the same sequence as they have been identified, regardless of $P/\%$ and $I/\%$ values, but they are not sorted. After that the data from the second subfile of standards are transferred from the cassette to the same registers of the operational memory thus replacing the data used in the first cycle. Again, the identification process is carried out and the list printed. This is repeated with each standard subfile until the entire process of identification is completed.

The entire process of identification of a sample may be restarted with different parameters of identification selected by the user. This may be done at the completion of each identification cycle. Error window, EW, and absent elements may be changed, or the user may decide whether matching of the standard with only one matching peak should be carried out. It is not, however, necessary to reenter the input data and in accordance to the user's decision the input data may not be printed out. In case of choosing the no-print-out-of-input-data option at the restart the identification process is speeded up.

Fig. 3 presents the flowchart of the program.

Example of use

Program RIFRAN 85 was used to carry out the phase analysis of the sample taken from bauxite ore, appearing as white stone. Absent elements were chosen on the basis of preliminary chemical analysis.

Results of phase analysis are presented in the output of calculator (Table 1). The output comprises the parameters of indentification, input data, and the list of potential candidates, according to which the sample consisted of boehmite, gibbsite, anatase, and kaolinite. With regard to high $P/\%$ and $I/\%$ values it may be concluded that boehmite is the major constituent. Eight matching peaks out of 14 possible, giving $P/\% = 57$ have been found for kaolinite. The $I/\%$ criterion being equal to 70 is rather high, suggesting that the strongest peaks with high I_r values have been matched. This may suggest that kaolinite is less abundant than boehmite for which almost all characteristic peaks have been matched (10 out of 11 possible) with high $I/\%$ values. Obviously, this consideration is justified only if the standard pattern does not comprise low number of strong peaks out of possible number of peaks (as for instance cristoballite) and, moreover, if the considered standard patterns have approximately equivalent I_{abs} values.

130

Similar considerations may be applied when scrutinizing the presence of gibbsite and anatase. A measure of the presence of a given standard in the sample may be the product of $P/\%$ and $I/\%$. The presence of garnierite has been rejected on the basis of chemical incorrectness. Its identification was the result of fortuitous overlapping of peaks, but if it were not so it would be necessary to consider the presence of a phase having similar structure as garnierite. We may then suggest that another phase yet not having been included in the database is present in the sample. This phase, however, would have similar lattice type and similar lattice parameters as the rejected standard.

A comparison of the overall number of input peaks with the identified peaks indicates that all peaks may have been identified, although some of them may have been fortuitously superposed.

Results of the above phase analysis have been confirmed by "hand" search//match method.

Conclusion

Preliminary experience with the identification system RIFRAN 85 has shown that the system is effective and capable of providing reliable results. Of the great value is not only the list of potential candidates present in the analyzed sample, but also the results of criteria analysis, as this enables us to estimate the semiquantitative composition of the sample in terms of individual phases. The role of user, however, cannot be underestimated.

An indisputable advantage of the described search/match system is its flexibility together with its simplicity and speed. It may be used directly in laboratories instead of troubling the computer centre and thus speeding up the entire process of phase analysis. Due to the limited capacity of the memory of calculator there are only very few possibilities for further improvement of the current version of RIFRAN 85. However, a new, sophisticated version of the system preserving its effectiveness will be prepared for a personal computer.

References

1. Frevel, L. K., *Anal. Chem. 37,* 471 (1965).
2. Nichols, M. C., *UCRI 700 78.* Lawrence Livermore Lab., 1966.
3. Johnson, G. G. and Vand, V., *Ind. Eng. Chem. 59,* 18 (1967).
4. Stewart, J. M., in *Computing in Crystallography,* IUCr Summer School. (Schenk, H., Editor.) P. 3. Delft University Press, 1978.

5. Huang, T. C. and Parrish, V., *Advan. X-Ray Anal. 25*, 213 (1982).
6. Johnson, G. G., *Acta Crystallogr. A37*, C-277 (1981).
7. Johnson, G. G., in *Laboratory Systems and Spectroscopy*, p. 76. M. Dekker, New York, 1977.
8. Lin Tian-Hui, Zhang Sai-Zhu, Chen Li-Jun, and Cai Xin-Xing, *J. Appl. Crystallogr. 16*, 150 (1983).
9. O'Connor, B. H. and Bagliani, F., *J. Appl. Crystallogr. 9*, 419 (1976).
10. Fiala, J. and Fialová, J., *Chem. Listy 79*, 48 (1965).
11. Schreiner, W. N., Surdukowski, C., and Jenkins, R., *J. Appl. Crystallogr. 15*, 513 (1982).
12. Goehner, R. P. and Garbauskas, M. F., *X-Ray Spectrometry 13*, 172 (1984).
13. Nichols, M. C. and Johnson, Q., *Advan. X-Ray Anal. 23*, 273 (1980).
14. Jenkins, R. and Hubbard, C. R., *Advan. X-Ray Anal. 22*, 133 (1979).
15. Edmonds, J. W., *J. Appl. Crystallogr. 13*, 191 (1980).
16. Hare, T. M., Russ, J. C., and Lanzo, M. J., *Advan. X-Ray Anal. 25* 237 (1982).
17. Lindsey, J. S., Christenson, C. P., and Henslee, W. W., *Advan. X-Ray Anal. 23*, 151 (1980).
18. Škrobian, M., Havlík, T., and Havlík, M., *Hutn. Listy 5*, 341 (1985).
19. *Powder Diffraction File Set 25* (1975). The Joint Committee for Powder Diffraction Standards, Swarthmore, Pennsylvania.
20. Heller, B. A. and Taylor, H. F. W., *Crystallographic Data for Calcium Silicates*. Her Majesty's Stationary Office, London, 1956.
21. Brindley, G. W. and Brown, G., *Crystal Structures of Clay Minerals*. Mineralogical Society, London, 1980.
22. Mikheev, V. I., *Rentgenometricheskii opredeliteľ mineralov*. Gosudarstvennoe nauchno-tekhnicheskoe izdateľstvo literatury po geologii i okhrane nedr, Moscow, 1957.

Translated by L. Komorova