Use of Full Diffraction Spectra, Both Experimental and Calculated, in Quantitative Powder Diffraction Analysis*

Deane K. Smith,[A] *Gerald G. Johnson, Jr*[A] *and Andrew M. Wims*[B]

[A] Department of Geosciences and Materials Research Laboratory,
The Pennsylvania State University,
University Park, PA 16802, U.S.A.
[B] General Motors Research Laboratories,
Warren, MI 48060, U.S.A.

Abstract

The use of the full powder diffraction trace over a selected diffraction range for quantitative analysis has advantages over using single peaks in that it compensates for the effects of peak overlap and low levels of preferred orientation. Using a data base composed of experimental and calculated traces, the phase composition of an unknown may be determined by determining the least-squares best-fit sum of the appropriate data base patterns to the pattern of the unknown. Weight fractions are calculated from the pattern weighting factors using the reference-intensity-ratio method.

1. Introduction

The use of X-ray powder diffraction for quantitative analysis (QXRPD) is of considerable interest because it is the most direct way to analyse for phase composition. Its implementation is beset with problems, including insufficient particles in the irradiated volume to allow random orientation to be achieved due to particle size, and preferred orientation of the crystallites in the sample due to the shape of the grains, resolution of appropriate diffraction peaks for determining the intensity of the diffraction component of each phase, locating appropriate standards to represent the phases in the samples to be studied, and intensity variations due to extinction and microabsorption. The basic theory of X-ray quantitative analysis was developed by Alexander and Klug (1948) and extended to the problems of overlapped peaks by Copeland and Bragg (1958), deWolff and Visser (1964), Karlak and Burnett (1966), and Chung (1974). Each of these approaches used specific peaks or integrated ranges and the measurement of their intensities. Use of the whole diffraction trace over a range which covers all the significant peaks of all the phases present allows each point in the trace to be considered as a separate intensity in the estimation of phase content. Smith *et al.* (1983) have shown that this approach helps minimise some of the sample and data reduction problems and leads to an increase in accuracy in the weight-fraction estimates.

The use of the full pattern as the basis of a quantitative analysis procedure grew out of an extension of the Karlak and Burnett (1966) procedure in the study of oil shales (Smith *et al.* 1985). In this earlier technique 20 integration ranges were defined, and

---

these same ranges were measured for every standard and for every unknown to be analysed. The method allowed for overlap and up to eight phases could be detected quantitatively in a single sample; however, when a new mineral with peaks outside the selected ranges had to be added to the data base of reference patterns, it was necessary to build a completely new data base by remeasuring all the patterns. It was recognised that if the whole pattern had been stored initially, then selected ranges could be defined as desired when the analysis was to be done. Finally, it was decided to use the full trace on a point-by-point matching scheme which proved to be highly successful.

The accuracy of the method using the full diffraction trace depends on the accuracy of the reference patterns used for the data base. Ideally, experimental patterns collected with a statistical accuracy greater than the patterns to be analysed should be used whenever possible. It is not always possible to isolate the desired phase, however, and it is sometimes necessary to use calculated diffraction traces for reference patterns. Two types of calculated patterns are available. One is a theoretical pattern calculated from the crystal structure by POWD (Smith *et al.* 1983), and the other is one derived from a set of *d* and *I* values such as are in the Powder Diffraction File (PDF) using program SIMUL (Wims *et al.* 1988).

2. Fitting the Full-trace Patterns

The concept of the full-trace approach is to find the weighted sum of selected reference patterns which best matches the trace of the diffraction pattern of the unknown sample. Obtaining this best-fit solution of selected data base patterns to the unknown is to minimise the set of expressions (Wims *et al.* 1988)

$$\delta I(\Delta 2\theta) = I_{unk}(\Delta 2\theta) - \sum_p W_p \mathrm{RIR}_p(\Delta 2\theta),$$

where $I_{unk}(\Delta 2\theta)$ and $I_p(\Delta 2\theta)$ are the diffraction intensities in the interval $\Delta 2\theta$ for the unknown and each of the reference phases respectively, W_p is the weight fraction of phase p, and RIR_p is its reference intensity ratio.* The W_p are the unknowns in the expression. The $\Delta 2\theta$ interval for the calculation need not be the same as used for the data collection. In fact, in practice, $\Delta 2\theta$ has been set to $0 \cdot 10°2\theta$ and may be set wider if desired.

The first step is to select the appropriate data base patterns. Usually, the user knows which phases to select, and it is straightforward to designate the choices. However, two other situations may occur. The user may not know all the phases present and under designate the phase selection, or the user may think a phase is present which is not and over designate the selection. Usually, the former may be detected when the solution is examined and some peaks display a poor fit while the rest of the pattern fits well. The latter case is indicated when one of the selected phases shows a zero or negative weight fraction. Experience has shown that correcting the phase designation, adding the missing phase or deleting the over-designated phase, leads to a better final fit because some of the peaks from the erroneous phase overlap other peaks in the data set and interact in the calculation.

---

* The reference intensity ratio was first defined by deWolff and Visser (1964). It is the ratio of the intensity of the strongest peak of a material to the strongest peak of a reference material in the diffraction pattern of a 50:50 mixture.

The best-fit results may be selected by comparing the pattern R factor for several calculations using different reference patterns or pattern shifts

$$R_{\text{pat}} = \left(\sum | I_{\text{unk}}(\Delta 2\theta) - I_{\text{calc}}(\Delta 2\theta)| \right) \Big/ \sum I_{\text{unk}}(\Delta 2\theta),$$

and by displaying the patterns graphically and examining the residual trace.

(a) Least-squares Solution

The least-squares calculation is carried out by a subroutine LSEI written by Haskell and Hansen (1978, 1979). This routine was first used for QXRPD by Snyder *et al.* (1983). It uses Householder functions and allows the user to include equality and inequality conditions on the solution. Careful use of condition restraints can be valuable to finding logical solutions to the data fitting, but over restraining the results can be detrimental.

The most obvious restraints involve the weight fractions. In a totally crystalline sample, the weight fractions of all the phases must sum to unity. Individual weight fractions must lie between zero and one. If the user does not analyse for all the phases present or if there is an amorphous component to the sample, the maxima must be less than unity. These restraints are most useful when the number of phases to be analysed is large and when the user is not sure of the phase composition and may include a non-existent phase in the input parameter.

The other restraints involve the chemical content if known. Some analytical schemes (Slaughter 1985; Johnson *et al.* 1985) consider the chemical information primary and use the X-ray information to determine the structure types for distribution of the elements. The technique described in this paper, like most QXRPD techniques, considers the X-ray information primary and the chemical information secondary. It is not uncommon to have a partial or total chemical analysis for the bulk sample. If the compositions of the phases in the sample are known and are fixed, the chemical information may be incorporated as part of the phase analysis with a high degree of confidence. If the phases have variable compositions, then the use of the chemical information as restraints becomes more questionable. In fact, the interest in powder diffraction as the method of quantitative analysis is because of the chemical variability problems or due to polymorphism where the chemical analysis is insufficient to determine the phase composition by itself.

Because chemical analyses are experimental data, the actual values have experimental uncertainties. These uncertainties preclude chemical restraints being used as equality restraints under any conditions. The user must allow for the uncertainties and set up ranges for the restraints. The variability of the phase chemistry complicates the situation considerably. Ideally, the sums of the products of the weight fractions of a given element in a phase times the weight fraction of the phase in the sample should equal the chemical analysis of that element in the sample. The range of an element in a given phase along with the analytical uncertainties can be used to set maximum and minimum limits to the restraints. This situation is severe with many common minerals and solid solutions.

(b) Reference Intensity Ratios

The reference-intensity-ratio method (Hubbard *et al.* 1976) is very useful with the full-trace technique, but its use is equivalent to using the peak-intensity ratios

rather than integrated-intensity ratios. In the procedures for determining RIR values (Morris *et al.* 1985; Davis 1986), the RIR is defined as the ratio of the integrated intensity of the strongest peak of the phase to the strongest peak of a standard (usually α-Al_2O_3) in a 50:50 mixture. Ratios of peak heights are frequently used as estimates of the ratios of integrated intensities. The procedure, established by C. R. Hubbard at the National Bureau of Standards, for measuring the RIR values included in the PDF is to measure the integrated intensities of the three strongest resolved lines in the pattern for both the phase and the standard. Dividing each measurement by the relative intensity of the corresponding line in the pure phase pattern yielded three estimates of the strongest intensity. Averaging each set of intensities prior to taking the ratio provides a better estimate of the RIR value.

The RIR in the full-trace method is an extension of this approach of using multiple intensity measurements to define the RIR value. There are two differences. First, every measured point along the trace is used in the estimate, although very low intensities have very little influence on the RIR. This measurement means there are many more points than used in the International Centre Associateship method. Second, the integration interval is much smaller than that used for integrating specific peaks and is much less than the width of the peak profile. Consequently, it is this width of integration which defines the effect of the normalisation, and the RIR is essentially a peak ratio rather than an integrated ratio.

The RIR value is most effectively determined by treating the reference/standard mixtures as unknowns to the quantitative procedure and determining the RIR value which yields the phase ratio of the prepared mixture. Three mixtures, 25:75, 50:50, and 75:25, allow three estimates of the RIR. Either the average of the RIR values found or the slope of a regression line in a plot of normalised peak ratios versus known composition yields useful estimates of the true RIR.
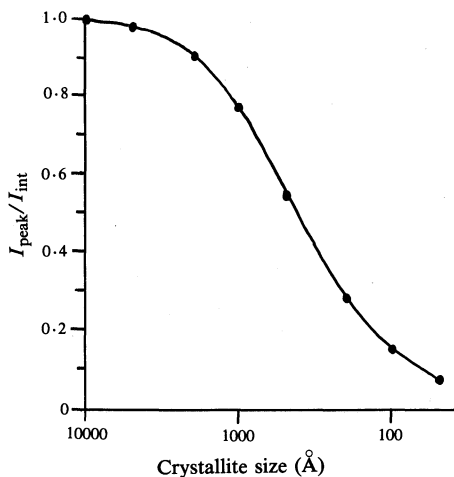

Fig. 1. Variation of the ratio of peak height to integrated intensity as a function of crystallite size.

The RIR for the full-trace method is very sensitive to the crystallinity of the sample because it is a peak ratio. As the peaks broaden with the decrease in crystallite size, the intensity level of the pattern diminishes and so does the RIR value. The effect of crystallite-size peak broadening on the RIR can be modelled by convoluting a gaussian distribution function with a half width corresponding to the Scherrer broadening for a selected crystallite size. Such a broadening effect has been included

in the POWD calculation (Bennett and Smith 1988), and the dependence of I_{peak}/I_{int} on the crystallite size is plotted in Fig. 1.

In practice there are two options to compensate for the crystallinity effects: expand the integration step or use a reference pattern whose crystallinity is similar to the phase to be analysed. If the user elects to expand the integration step size, the RIR approaches that of an integrated ratio. The step size should be at least five times the half width of the profile of the phase. Note that the step size of the data collection remains small; it is only the interval of the integration which is increased. If available, reference material of the phase being analysed should be used to generate the data base pattern used for the analysis. The RIR value determined from this reference will certainly be more appropriate than any which could be calculated.

(c) Using an Internal Standard

The most common method in QXRPD is the internal standard method where a pre-determined weight fraction of a standard is added prior to the analysis. If the standard is part of the unknown, the method is called spiking or addition. Though α-Al_2O_3 is a common standard, many other materials are available (Standard Reference Materials 675 from the Office of Standard Reference Materials, U.S. National Bureau of Standards).

This method may be used with the full-trace method as well as with the earlier methods as long as the standard has a pattern in the data base. Use of the internal standard allows independent estimates of the total crystallinity of the unknown, and confirms the presence of an amorphous component or if a phase was omitted in the designation of phases. It is a simple calculation to normalise the results to the absence of the standard and total the residual weight fractions. If the total is less than one, something is missing; if the total is greater than one, there is a problem with the data or the phase designations.

3. Building the Data Base

The heart of this quantitative analysis technique are the patterns which comprise the data base. To be effective, the patterns must be representative of the phases which occur in the samples to be analysed. Each data set should cover the useful range of the diffraction peaks and be collected with sufficient statistical accuracy to yield a reasonable pattern match. The step size and the range need to be selected for the data base in general, whereas the count time can vary as desired for the magnitude of the diffracted intensities and the accuracy needed. It is probably best to store raw data in a master data base and build a derived data base with each background removed and patterns normalised. Such a strategy allows for future modifications of processing the patterns.

Data base parameters were selected in consultation with many potential users including the sponsors of this project. For simplicity in pattern manipulation, the parameters for all patterns were kept the same in the data base. All data were collected on a Philips diffractometer with a variable divergence slit using Cu radiation. The range was chosen as 5° to 75° 2θ with a step size of 0·02° 2θ. This range should include all but a very few high angle peaks that would ever be used for quantitative analysis. The step size is a trade-off between the desire to have all the details of the profile shapes yet minimise the storage requirements. Data taken at 0·02° step size can be combined into larger intervals for the calculation of the best-fit solutions

but retained at 0·02° for display. Patterns for the unknowns can be collected and matched over ranges smaller than that used for the data base patterns, but the reverse cannot be true. In practice, it is common to start scans just below the first peak in the pattern rather than at $5° 2\theta$.

There are three sources for data base patterns. The best source is to collect an experimental pattern on a phase-pure sample and to determine the RIR on a mixture of the same material mixed with the standard. Alternatively, the user may generate a calculated trace from a known crystal structure using program POWD (Smith *et al.* 1983) and calculating the RIR value, or the user may simulate a trace from the *d–I* data in the PDF or other source using program SIMUL (Wims *et al.* 1988). Determining the appropriate RIR for a SIMULated trace requires an actual sample of known composition. Such a sample need not be phase pure if the other phases have known RIR values. It is best to use experimental patterns where possible, but calculated patterns provide traces free of impurity lines which may prove to be the only way to get some reference sets.

Experimental patterns require further processing to remove the background and any other artifacts. If the background is not removed from the reference data, it will be over-emphasised when several reference patterns are added during the pattern fitting procedure. During this processing, unwanted second phase peaks may be removed from the trace. Quartz is a very common contaminant in most clay concentrates, and it is often easier to strip the 101 peak mathematically than to beneficiate the sample to eliminate the last trace of the quartz. The subroutine for processing was the same for the reference data and for the unknown data. In processing the unknown data, any region could be zeroed. The quantitative program is designed to ignore the zeroed region in the least-squares best-fit process. Thus, if the user wanted to use only the strongest peaks of each phase, all the other regions could be zeroed, and the calculation would be equivalent to the classical method of integrating individual peaks to determine each phase.

Selecting the appropriate background is critical to achieving the best solution. There is no good fundamental criterion for defining the background due to the long tails of the diffraction peaks. For this program, two basic assumptions are made: the background is a smooth function with only small changes in slope, and it follows the lower edge of the experimental trace rather than through the average of the peak free portions of the trace. This selection is made to under-subtract the tails rather than over-subtract them. Large regions in simple patterns where no peaks are known to occur may be zeroed rather than background fit.

Calculated patterns require no further data processing because they lack any background. The key in their calculation is to select the proper profile shape and allow for the change in half-width with diffraction angle in order to match the experimental data most closely. The normal Cauchy (Lorentzian) profile has proved to be the most useful to date. A half-width versus two-theta table can be generated for any diffractometer using a well crystallised sample with well resolved peaks. Sodalite, $Na_4Al_3Si_3O_{12}Cl$, is one useful material for the half-width determination because of its primative cubic symmetry and large cell size, $a = 8·878$ Å.

Obtaining patterns for stoichiometric compounds is usually not difficult, but finding appropriate reference samples for compounds which show variations in composition and/or crystallinity is more of a problem. For compounds which exhibit solid solution, it may be necessary to include patterns for several compositions in the data

base and determine *a posteriori* which pattern yields the best solution. An example of this problem is dolomite–ankerite, $Ca(Mg, Fe)(CO_3)_2$–$Ca(Fe, Mg)(CO_3)_2$. The Fe for Mg substitution affects the intensities significantly and a ferroan dolomite pattern is measurably different from a pure dolomite. Multiple patterns are necessary to cover the range of compositions to be expected. A similar approach may be used for compounds which display variable crystallinities such as kaolinite, $Al_2Si_2O_5(OH)_4$. Patterns of different crystallinities are included in the data base, and the most appropriate one is determined by trial-and-error. Compounds which show more than one type of variation are even more troublesome. The feldspar albite, $(Na_{1-x}, Ca_x)Al_{1+x}Si_{3-x}O_8$, is an example which shows both structural and chemical differences. The best way to handle this problem is to use several references and find the best pattern *a posteriori*. Although these problems complicate finding accurate weight fractions by the X-ray powder method, they are even more difficult to solve by any other method.

4. Results and Discussion

The evaluation of the accuracy of any quantitative analysis scheme can be tailored by selecting well behaved analytes, and this method is no different. Up to eight phases have been determined in a single sample with an absolute accuracy of 2%. The error increases with the complexity of the compounds including chemical variability and crystallinity problems. Detectability limits may be as low as 0·5%, and even lower with exceptional materials like quartz, SiO_2. Compounds like the feldspars have proved to be the most difficult to achieve adequate accuracy.

One difficulty commonly encountered is the shifting of the patterns on the 2θ scale due to sample displacement and beam penetration errors. It has proved necessary to allow the unknown pattern to be shifted with respect to the reference patterns to find the best-fit final solution. Even better fits would be obtained by shifting the reference patterns to their correct position prior to matching the unknown pattern. Ideally, it would be advisable to include a d-spacing standard in all patterns, both reference and unknown, and use its characteristic lines to align the patterns. In practice, this approach would also require that the d standard be treated as a phase in the quantitative solution. Sometimes there is a phase in the unknown which can be used as a d standard. Alternatively, the unknown can be shifted in several steps, and the best fit can be selected as the one with the lowest R factor or the one with the best residual trace. One limitation of the present data base is the step size of 0·02° which prevents alignment of patterns to closer intervals.

For the purposes of this paper, it will be of interest to compare the results obtained using experimental patterns to those obtained with calculated reference patterns.

(a) Example 1—Two Components

The first example is a binary mixture of Al_2O_3 and ZnO. Three cases are compared: (1) the reference patterns are experimental patterns of the same substances used in the mixture, (2) the reference patterns are calculated from the known crystal structures, and (3) the reference patterns are simulated from the d–I listings in the PDF. The comparison that is important is the perfection of the fit of the patterns of the experimental mixture and the best-fit composite pattern found by the computer. The estimate of the weight fraction depends on the value assigned to the RIRs. The
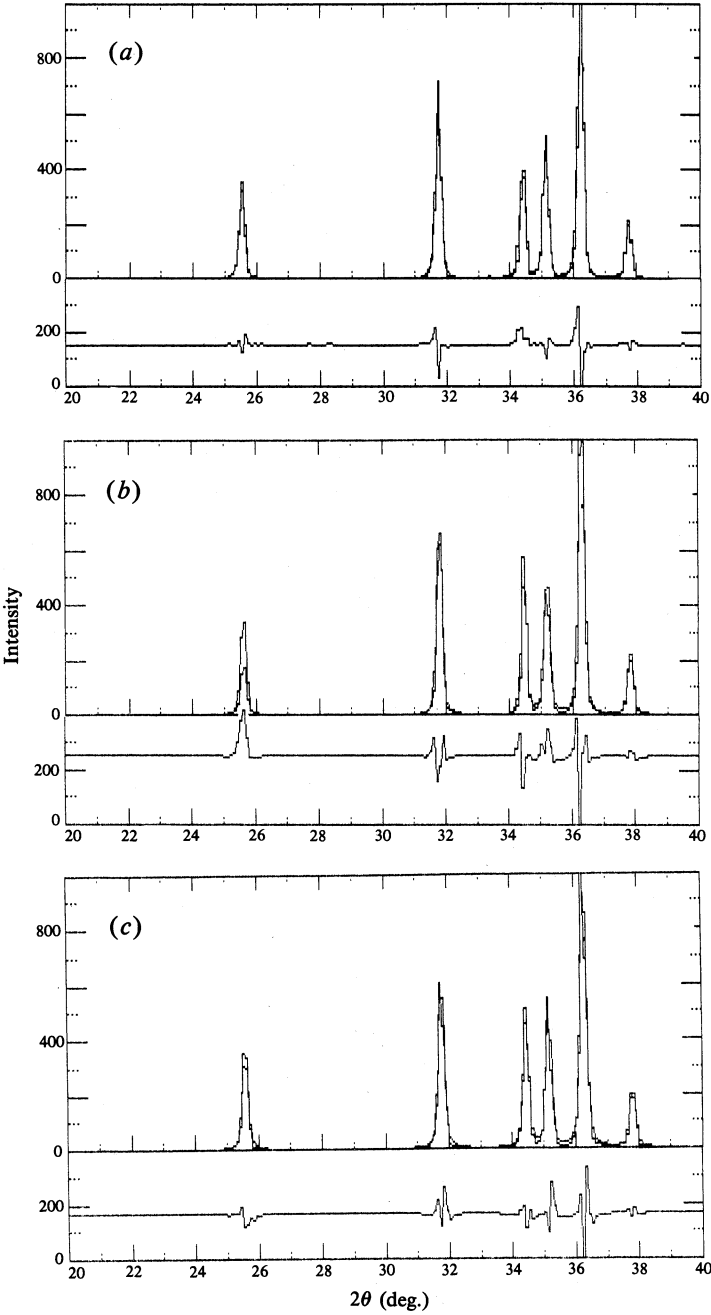
Fig. 2. Pattern match for a two-phase unknown using (*a*) experimental reference patterns, (*b*) POWD calculated reference patterns and (*c*) PDF simulated reference patterns.
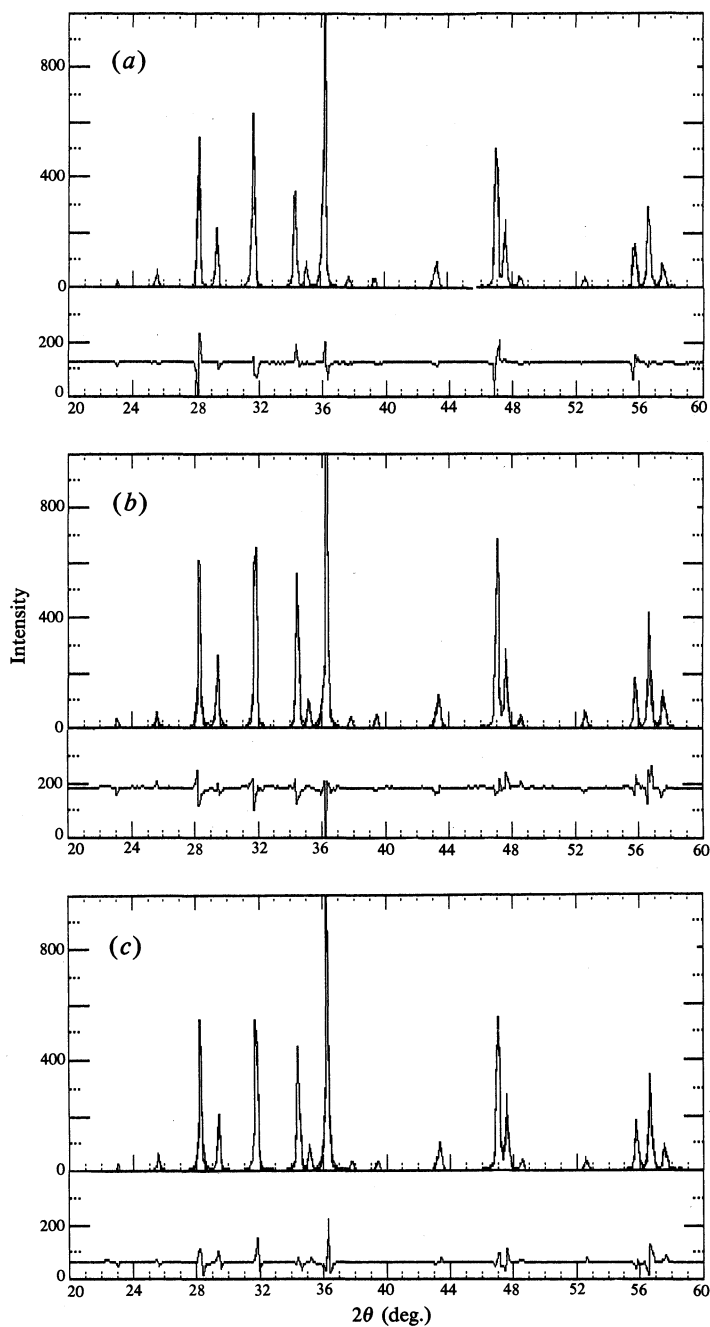
Fig. 3.   Pattern match for a four-phase unknown using (*a*) experimental reference patterns, (*b*) POWD calculated reference patterns and (*c*) PDF simulated reference patterns.

comparison can be made by using the pattern R factor and by examining the plot of the differences between the composite pattern and the experimental trace.

Fig. 2*a* shows the superposition of a portion of the calculated and experimental traces of case 1 and their differences. The difference curve indicates that the major discrepancy is the slight 2θ offset of some of the peaks. Because $\Delta 2\theta$ is $0\cdot02°$, this offset cannot be matched perfectly and contributes significantly to the residual. The R factor for this case is $0\cdot16$.

Fig. 2*b* shows the results of case 2 for the calculated reference patterns. The R factor for this case is $0\cdot26$. The difference trace is plotted on the same scale as the previous plot. The difference trace suggests that the major discrepancy is in the relative intensity of the first Al_2O_3 peak and in the narrower profile selected for the calculated patterns. There is a slight indication of offset of the pattern for the ZnO.

Fig. 2*c* shows the results of case 3 for the reference patterns simulated from the PDF. The R factor for this case is $0\cdot19$. The intensities are a closer match than for case 2, but there is still a mismatch in the widths assigned to the peaks. It is interesting that the agreement is better than for case 2. Simulation of patterns is thus a useful alternative when the reference material cannot be collected in a pure state. Determining the RIR for this case is most difficult in that it still requires an experimental mixture with a known composition to set the value.

(b) Example 2—Four Components

The second example is a mixture of Al_2O_3, ZnO, CaF_2 and $CaCO_3$. Again three cases are compared. Fig. 3*a* shows the results for the experimental reference patterns. The R factor is $0\cdot14$. The major discrepancy is due to the offsets of a few of the peaks. The peak heights show a very good match. Fig. 3*b* is for the case with the POWD calculated patterns. The width of the peak and some peak shifts contribute to the discrepancy and to the higher R of $0\cdot22$. Fig. 3*c* shows the case with the PDF simulations. The R factor is intermediate at $0\cdot19$, and the major discrepancy is the mismatch of some of the peak widths.

Table 1. Comparative results from a four-phase unknown (with ESDs in parentheses)

Phase	Experimental references	POWD references	PDF references	Actual
CaF_2	$30\cdot8(0\cdot7)$w/o	$32\cdot6(0\cdot4)$w/o	$31\cdot0(0\cdot3)$	30w/o
$CaCO_3$	$12\cdot8(0\cdot3)$	$9\cdot0(0\cdot4)$	$10\cdot4(0\cdot4)$	10
Al_2O_3	$19\cdot2(0\cdot7)$	$16\cdot7(0\cdot7)$	$20\cdot4(0\cdot9)$	20
ZnO	$37\cdot3(0\cdot2)$	$41\cdot9(0\cdot3)$	$38\cdot3(0\cdot2)$	40

Table 1 shows the weight per cent (w/o) estimates obtained from the three runs. Experimental RIRs are used with the experimental reference data. The RIRs calculated by POWD were used with both the POWD references and the PDF simulated references, and consequently the estimated standard deviations are a measure of the precision of the data fit and not of the accuracy of the weight fraction estimates. Adjustments of the RIR values can make all three results agree with the weight fractions of the mixture, but there is no justification to make the adjustments. This situation illustrates that the bigger problem in QXRPD analysis is not the pattern fitting, but the determination of the appropriate RIR value.

5. Conclusions

Pattern summation employing the whole diffraction trace and using experimental data was shown to be an effective method for QXRPD by Smith *et al.* (1987). This study shows that calculated powder patterns based on well determined crystal structures and simulated patterns based on accurate *d–I* data sets are viable alternatives to experimental patterns for quantitative reference data. There are two limitations: (1) the determination of the appropriate reference intensity ratio, and (2) the matching of the shape of the peak profiles.

References

Alexander, L. E., and Klug, H. P. (1948). *Anal. Chem.* **20**, 886.

Bennett, H. D., and Smith, D. K. (1988). *Adv. X-ray Anal.* **31**, (in press).

Chung, F. H. (1974). *J. Appl. Cryst.* **7**, 519–25.

Copeland, L. E., and Bragg, R. H. (1958). *Anal. Chem.* **30**, 196.

Davis, B. L. (1986). Reference intensity method of quantitative analysis. South Dakota School of Technology, Institute for Atmospheric Sciences.

deWolff, P. M., and Visser, J. W. (1964). Technique for obtaining absolute intensities for powder samples. Report 641.109, University of Delft.

Haskell, K. H., and Hanson, R. J. (1978). An algorithm for linear least squares problems with equality and nonnegativity constraints. Sandia National Laboratories, SAND77-0552.

Haskell, K. H., and Hanson, R. J. (1979). Selected algorithms for linearly constrained least squares problems—A users guide. Sandia National Laboratories, SAND78-1290.

Hubbard, C. R., Evans, E. H., and Smith, D. K. (1976). *J. Appl. Cryst.* **9**, 353–62.

Hubbard, C. R., Robbins, C. R., and Snyder, R. L. (1983). *Adv. X-ray Anal.* **26**, 149–56.

Johnson, L. J., Chu, C. H., and Hussey, G. A. (1985). *Clays Clay Miner.* **33**, 107–17.

Karlak, R. F., and Burnett, D. S. (1966). *Anal. Chem.* **38**, 1741.

Morris, M. C., McMurdie, H. F., Evans, E. H., Paretzkin, B., Parker, H. S., Wong-ng, W., Gladhill, D. M., and Hubbard, C. R. (1985). 'Standard X-ray Diffraction Powder Patterns', NBS Monograph 25, Section 21 (US Department of Commerce: Washington).

Slaughter, M. (1985). Int. Clay Conference, Denver (Ed. D. E. Pevcar) (Clay Minerals Society: Bloomington, IN).

Smith, D. K., Johnson, G. G., Jr, Scheible, A., Wims, A. M., Johnson, J. L., and Milburn, M. V. (1987). *Powder Diffraction* **2**, 73–7.

Smith, D. K., Kerrick, D. M., Sterner, S. M., and Meddaugh, W. S. (1985). *Adv. X-ray Anal.* **28**, 169–74.

Smith, D. K., Nichols, M. C., and Zolensky, M. E. (1983). POWD10—A FORTRAN IV program for calculating X-ray powder diffraction patterns, Version 10, Pennsylvania State Univ., College of Earth and Mineral Sciences.

Wims, A. M., Johnson, J. L., Milburn, M. V., Smith, D. K., Johnson, G. G., Jr, and Scheible, A. (1988). *Adv. X-ray Anal.* **31**, (in press).