Aust. J. Phys., 1988, **41**, 229–36

Learned Peak Shape Functions for Powder Diffraction Data*

A. Hepp and Ch. Baerlocher

Institut für Kristallographie und Petrographie,
ETH-Z, CH-8092 Zürich, Switzerland.

Abstract

An algorithm is described for the determination of an experimental (learned) peak shape function, which has been used succesfully in crystal structure refinements from powder data. The function gives an optimal fit to almost any peak shape since it is not based on an analytical expression. It is determined from a single peak in a pattern by first fitting this peak with the proposed algorithm which ensures that the function is smooth and has only one maximum and two inflection points. The learned function is then normalised and decomposed into a symmetric and an asymmetric part. These are stored in tabulated form, permitting linear interpolation. As with an analytical function, a FWHM and asymmetry function describing the 2θ dependence of the peak shape can be applied.

1. Introduction

In powder structure determinations and refinements it is important to have a peak shape function which gives an optimal description of the observed peak. The crucial parts of a peak which make fitting particularly difficult are the peak 'tails' and the asymmetry often observed in low angle peaks. The closer the agreement between the assumed and the observed peak shape the more correct will be the distribution of the integrated intensities among overlapping peaks and thus the better will be the refinement. This applies to the Rietveld technique, the integrated intensity method of structure refinement and of structure determination by direct methods. A better description of the peak shape will also lead to more accurate lattice parameters and improve the results of quantitative analysis of multiple-phase mixtures.

It has not been easy to find analytical functions to describe the peak shapes found in X-ray powder data. A review of a number of such functions used has been given by Young and Wiles (1982), who point out that there is still a need for a better function. The best choice appears to be either a Pearson VII [$f(x) = 1/(1+kx^2)^m$ where m is a variable to be optimised], or a Voigt which is a convolution of a Lorentzian and Gaussian or a pseudo-Voigt (Hecq 1981). All these functions are symmetric and some kind of asymmetry must be introduced to model a real X-ray peak. To allow for this asymmetry Parrish *et al.* (1980) used three overlapping Lorentzians per peak in their profile fitting procedure.

* Paper presented at the International Symposium on X-ray Powder Diffractometry, held at Fremantle, Australia, 20–23 August 1987.

0004-9506/88/020229$03.00

We have taken a different approach to finding the best peak shape function; namely to use the experimentally observed peak profile. Of course, this cannot be used as measured because of the scattered nature of observed peaks. Some kind of fitting or smoothing has to be done. Such experimentally determined shape functions have, for instance, been applied by Mortier and Costenoble (1973) who used a Fourier series. Baerlocher and Hepp (1980) determined a standard peak shape function by fitting an expression involving a quotient of two polynomials to a single non-overlapping peak. Pyrros and Hubbard (1983) proposed similar rational functions as profile models in powder diffraction. Although polynomials generally give a good fit, they are not flexible enough to describe all observed peak shapes equally well, or else have a large number of parameters. Fourier series as well as more general polynomials tend to introduce additional wiggles, especially when strongly scattered data have to be used.

These problems can be overcome with a special algorithm. This algorithm fits a smooth well-behaved curve to any kind of peak, whether it is strongly asymmetric, Gaussian or mainly Lorentzian in shape. The procedure, described in the next section, also works with badly scattered, low count rate data. A FORTRAN program, PEAK, has been written with this algorithm. It produces, in numerical form, a continuous function for which the second derivatives exist. From this curve, a normalised numerical function is generated which can be used like an analytical function. The form of this function is very suitable for powder patterns since it allows the half-widths at half-maximum as well as the peak asymmetry to vary as a function of 2θ. The values of the function and its derivatives at an arbitrary point are rapidly obtained by a linear interpolation. In this paper we briefly describe the ideas behind the peak fitting algorithm, give the definition for the peak shape function and show some examples of its application.

2. Procedure for Numerical Fitting of an Observed Peak

An idealised experimental peak shape $y(2\theta)$ has the following characteristics:

(i) the function value is always positive;

(ii) it increases monotonically to a maximum at $2\theta_0$ and then decreases again monotonically; and

(iii) the curvature is positive at the beginning, changes sign at $2\theta_a$, becomes negative around the maximum, and becomes positive again at $2\theta_b$.

For practical purposes, the region of the peak is limited by the lower and upper boundaries $2\theta_l$ and $2\theta_u$. In addition to the trivial condition

$$2\theta_l < 2\theta_a < 2\theta_0 < 2\theta_b < 2\theta_u,$$

these characteristics lead to the following special conditions:

(i) $y(2\theta) > 0;$

(ii) $y'(2\theta) > 0 \quad \text{for} \quad 2\theta_l \leqslant 2\theta < 2\theta_0$

$y'(2\theta) = 0 \quad \text{for} \quad 2\theta = 2\theta_0$

$y'(2\theta) < 0 \quad \text{for} \quad 2\theta_0 < 2\theta \leqslant 2\theta_u;$

(iii) $y''(2\theta) > 0$ for $2\theta_l \leqslant 2\theta < 2\theta_a$ and

$2\theta_b < 2\theta \leqslant 2\theta_u$

$y''(2\theta) = 0$ for $2\theta = 2\theta_a$ and $2\theta = 2\theta_b$

$y''(2\theta) < 0$ for $2\theta_a < 2\theta < 2\theta_0$ and

$2\theta_0 < 2\theta < 2\theta_b$

$y''(2\theta) \leqslant 0$ for $2\theta = 2\theta_0$.

A curve $y(2\theta)$ which obeys* these—and only these—conditions is fitted to the observed data $y^o(2\theta_i)$. This is achieved through a series of deformations of a starting curve which already possesses the required characteristics. These deformations are such that the values of the function, the derivatives and the zeros are changed without violating the above conditions.

To begin, additional 'observed' data points are generated by interpolation with rational splines (Späth 1978) in order to obtain a better description of the peak shape, especially in the strongly curved regions. From a statistical viewpoint, these data points are of equal quality to the measured ones.

The choice of the starting curve is not critical. The program chooses a function which is somewhat slimmer than the observed one. Three types of deformation are used in an iterative procedure to fit the starting curve $y_1(2\theta)$ to the observed data $y^o(2\theta_i)$. The deformations of type 2 and 3 are almost trivial and pose no problems regarding the required conditions. They are respectively the translation along the 2θ axis to adjust $2\theta_0$ and the scaling of the $y(2\theta)$ values to fit the peak height. With a type 1 deformation the 'shape' of the starting curve is adjusted.

To ensure that the above conditions are always fulfilled, the type 1 deformation is performed according to the algorithm illustrated in Fig. 1. For simplicity the discussion is limited to the high-angle half of the peak. A similar procedure, with the signs reversed, is followed for $2\theta < 2\theta_0$. Starting at the maximum $2\theta_0$ the curve $y_k(2\theta)$ is deformed into $y_{k+1}(2\theta)$ at successive points $2\theta_j$, distributed in proportion to the magnitude of the curvature according to

$$y_{k+1}(2\theta) = y_k(t(2\theta)).$$

The function $t(2\theta)$, which may be understood as a scaling of the 2θ axis, has the form

$$\begin{aligned} t(2\theta) &= 2\theta && \text{for } 2\theta < 2\theta_j - \delta \\ &= 2\theta + a_3\xi^3 + a_4\xi^4 && \text{for } 2\theta_j - \delta \leqslant 2\theta \leqslant 2\theta_j + \delta \\ &= 2\theta + a(2\theta - 2\theta_j) && \text{for } 2\theta > 2\theta_j + \delta, \end{aligned}$$

where δ is the smallest step size used and

$$\xi = 2\theta - (2\theta_j - \delta), \qquad a_3 = a/4\delta^2, \qquad a_4 = -a/16\delta^3.$$

* The curve $y(2\theta)$ is generated in numerical form, but the omission of the subscript on 2θ should indicate that it can be calculated at any point by spline-interpolation.

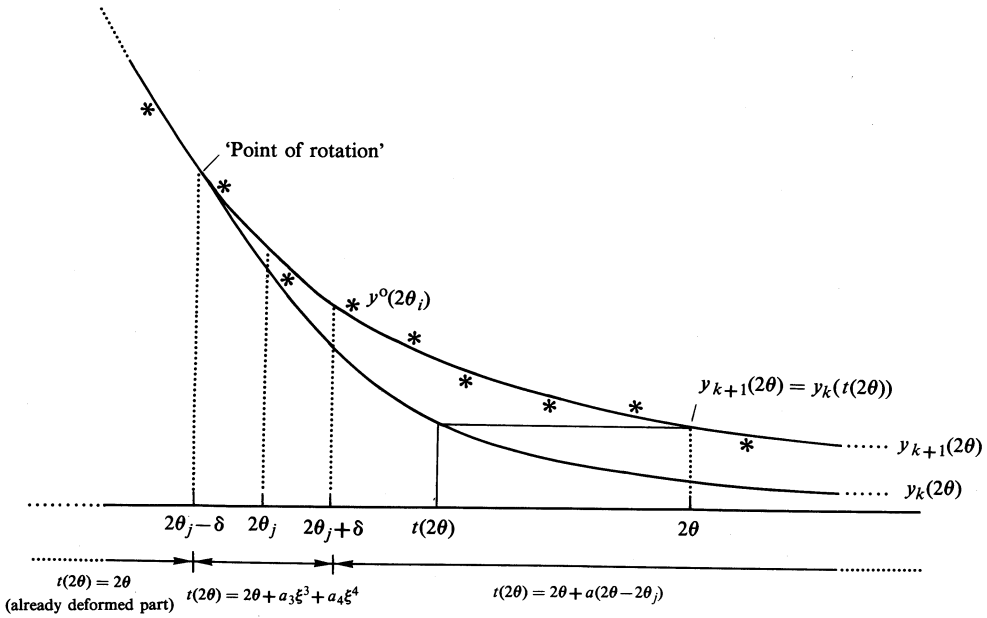

Fig. 1. Deformation of $y_k(2\theta)$ into $y_{k+1}(2\theta)$ (see Section 2).

Thus, the 2θ axis remains unchanged on the low-angle side of the 'point of rotation' $2\theta_j-\delta$ and it is stretched or compressed by the factor $1/(1+a)$ for $2\theta > 2\theta_j+\delta$. (This implies that $a > -1$, a condition which can also be deduced from the requirements for the first derivative.) The more complicated function of a in the interval $2\theta_j-\delta \leqslant 2\theta \leqslant 2\theta_j+\delta$ preserves the required properties of the curve. The parameter a is determined by using a weighting scheme which takes into account the relative slope of the curve at the point $2\theta_j$:

$$a = wa_y+(1-w)a_t \qquad (0 \leqslant w \leqslant 1).$$

The components a_y and a_t are calculated such that they minimise the deviations between $y_{k+1}(2\theta)$ and the observed data $y^o(2\theta_i)$ in the direction of y and 2θ respectively. The conditions (iii) above restrict the change allowed in the second derivatives $y''(2\theta)$ during the deformation. [The derivatives $y'(2\theta)$ and $y''(2\theta)$ are obtained as a spin-off from the spline-interpolation of $y(2\theta)$.] This again leads to certain restrictions on a. If these restrictions are violated, a is reduced systematically in magnitude until they are fulfilled. [For further details regarding the weight w, the step sizes used, and the derivation of the restrictions on a etc., see Hepp (1981).]

In an iterative procedure using the three deformation types in turn, a calculated curve $y(2\theta)$ is obtained which gives a very good fit after only a few cycles (see Figs 2*a* and 3).

3. Numerical Peak Shape Function Φ

The function Φ describes the intensity distribution of a reflection obtained with monochromatic radiation. Similar to most other profile functions it is a function of

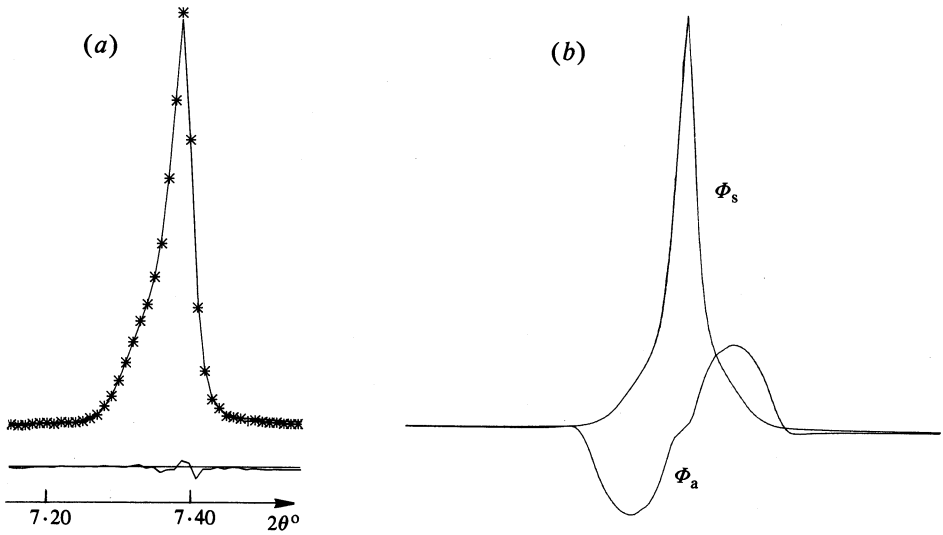

Fig. 2. (*a*) Observed (stars) and fitted curve of a strongly asymmetric X-ray peak. The difference is plotted below the curve. (*b*) Symmetric Φ_s and asymmetric Φ_a functions for the fitted curve shown in Fig. 2*a* (not to scale).

the difference $\Delta 2\theta = 2\theta - 2\theta_0$, the position of the maximum $2\theta_0$ and the half-width at half-height H. Since X-ray diffraction lines, especially at low angles, are asymmetric, Φ is also dependent on the degree of asymmetry A. Therefore Φ is assumed to have the following general form:

$$\Phi = \Phi(\Delta 2\theta;\, H, A) = \frac{1}{H}\, \Phi_s(r)\{1 - A\Phi_a(r)\}\,, \tag{1}$$

where Φ_s and Φ_a are the symmetric and asymmetric parts of the function. Both are tabulated as functions of r, where $r = \Delta 2\theta / H$. The function is normalised so that for any value of H and A

$$\int_{-\infty}^{\infty} \Phi(\Delta 2\theta;\, H, A)\, \mathrm{d}(\Delta 2\theta) = 1 \tag{2a}$$

with the following conditions: For the symmetric part

$$\Phi_s(-r) = \Phi_s(r); \qquad \Phi_s(1) = \tfrac{1}{2}\Phi_s(0), \qquad \int_{-\infty}^{\infty} \Phi_s(r)\, \mathrm{d}r = 1\,, \tag{2b}$$

and for the asymmetric part

$$\Phi_a(-r) = -\Phi_a(r), \qquad \Phi_a(1) = 1\,. \tag{2c}$$

The standard line profile is thus represented by the two functions $\Phi_s(r)$ and $\Phi_a(r)$. They are calculated from the fitted curve $y(2\theta)$ described in the previous section. This is done by decomposing $y(2\theta)$ into the symmetric part $y_s(\Delta 2\theta)$ and the asymmetric

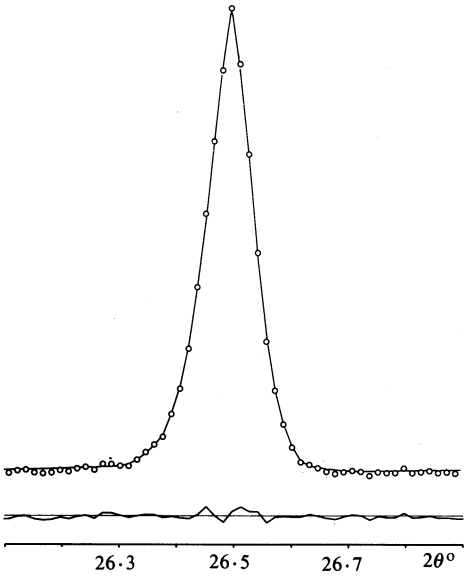

Fig. 3. Observed (open circles) and fitted standard peak for hydrated zeolite Rho (peak maximum is 4320 counts and $R = 0{\cdot}053$).

part $y_a(\Delta 2\theta)$ according to

$$y_s(\Delta 2\theta) = \tfrac{1}{2}\{y(2\theta_0+\Delta 2\theta)+y(2\theta_0-\Delta 2\theta)\}, \tag{3a}$$

$$y_a(\Delta 2\theta) = \frac{y_s(\Delta 2\theta)-y(2\theta_0+\Delta 2\theta)}{y_s(\Delta 2\theta)+\epsilon y(2\theta_0)}. \tag{3b}$$

The small number $\epsilon y(2\theta_0)$ forces $y_a(\Delta 2\theta)$ to converge towards zero at both ends of the interval thus becoming insensitive to fluctuations of the data.

From this, the symmetric and asymmetric functions Φ_s and Φ_a are calculated according to equations (2). Their values, together with their first derivatives, are stored in a table which permits linear interpolation. Fig. 2*b* shows an example of a symmetric and an asymmetric part obtained from the peak fit in Fig. 2*a*. These functions can now be used to fit all peaks of a pattern by varying the half-width at half-maximum H and the peak asymmetry A.

4. Results and Discussion

Two examples are given to illustrate the power and the versatility of the procedure described. In Fig. 2*a* a strongly asymmetric low-angle peak of a new zeolite-like material is shown. The data were collected with synchrotron radiation at HASYLAB, DESY, Hamburg in steps of $0{\cdot}01°2\theta$, using the Ge(111) plane as the analyser crystal. The peak maximum has 12700 counts. After seven cycles, the fitting procedure converged with an R value of $0{\cdot}023$. The remaining differences which can be seen below the curve in Fig. 2*a* are mainly due to counting statistics. It is not possible to fit these data—admittedly an extreme case—with any of the commonly used analytical functions. Even taking only the symmetric part Φ_s of this peak from Fig. 2*b* and fitting it with a pseudo-Voigt function did not give a satisfactory result.

A more common peak shape is shown in Fig. 3. This is the peak fit used to determine the learned or standard peak shape function for the refinement of hydrated zeolite Rho (McCusker and Baerlocher 1984). Here the peak count is somewhat lower

(4320 counts) and the R value obtained after only three cycles was 0·053. Despite the larger scatter, no wiggles are introduced. To demonstrate how successfully the learned peak shape function also models the changing peak shapes with 2θ, sections of the refined pattern of hydrated zeolite Rho are shown in Fig. 4. Since this zeolite is cubic, the peaks are all single or exactly overlapping. They have been fitted in a Rietveld refinement using the learned peak shape function Φ (equation 1) and Φ_s and Φ_a determined from Fig. 3.

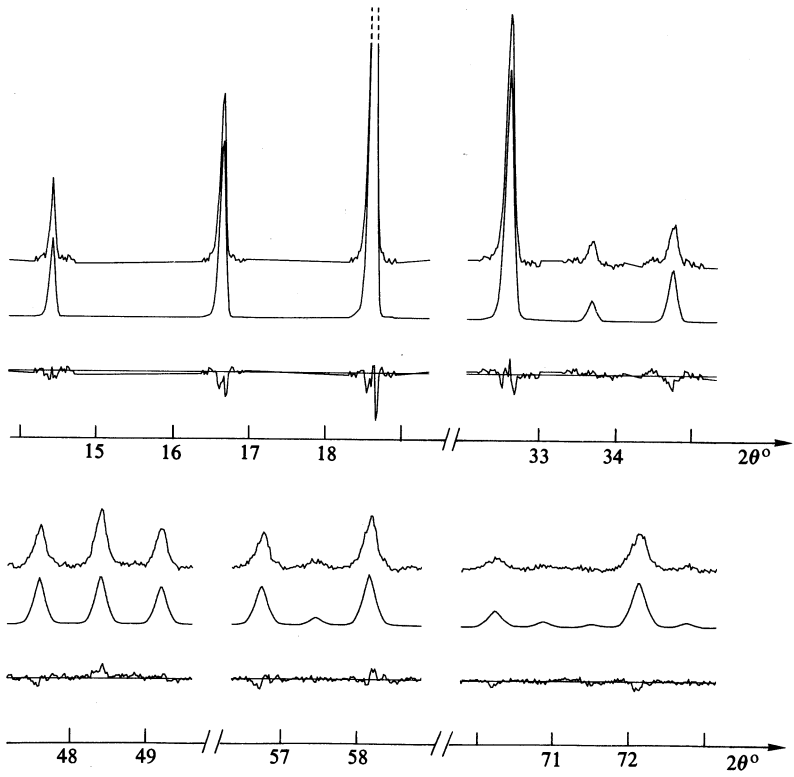

Fig. 4. Sections of the Rietveld refinement of hydrated zeolite Rho using the peak shape function determined from Fig. 3. The function can easily cope with the changing peak width and asymmetry.

The FORTRAN program PEAK, based on the algorithm described, is a stand-alone program which is simple to use and is available on request from the authors. It does not require any input other than the observed peak profile. The generated function can, in principle, be used with any Rietveld or pattern decomposition program. It has been incorporated in the X-ray Rietveld system (Baerlocher 1982) and has now been successfully applied to data from various sources. Examples of its application can be found in the review article by Baerlocher (1986, and references therein). The procedure described requires a suitable single peak in the pattern. In most cases, such a peak can be found. However, should it prove to be a problem, a peak from a related phase measured under similar conditions can be used. This is probably still more accurate than an assumed analytical function. In cases where not only the half-width at half-maximum and the peak asymmetry changes with 2θ but also the *shape*, the

following procedure could be used. Two peak shape functions, one at a lower and one at a higher angle, could be determined and then mixed in a way similar to the mixing of a Gaussian and a Lorentzian in the pseudo-Voigt. The mixing parameter would then be dependent on 2θ. So far, such an approach has not been used and indeed has not been found to be necessary. However, two different peak shape functions, one for the low angle and one for the high angle peaks, have been used.

There is of course no theoretical or physical basis for such a learned function. We have used a purely practical approach to find the best function for a structure refinement. Nonetheless, it is still possible to gain information about the physical state of the sample. The procedure described finds the width at half-maximum, the height and area of each peak, and these parameters can be used for any of the normal applications of powder diffraction, including estimates of size and strain.

We believe that this function has been one of the crucial factors in the successful Rietveld refinements of complicated zeolite structures. The most recent example is the refinement of monoclinic ZSM-5 (Baerlocher and Schicker 1987). Up to $80°2\theta$ ($\lambda = 1{\cdot}54$ Å) this sample has 2950 reflections with a correspondingly high number of overlapping reflections (up to 96 at a single point). Only with a very good model for the peak shape, such as the learned peak shape function, will the intensity be correctly distributed among the overlapping reflections and thus allow a satisfactory refinement.

Acknowledgments

We thank Professor W. M. Meier for many helpful discussions and Dr G. Harvey for critically reading the manuscript. This work has been supported in part by grants from the Swiss National Science Foundation.

References

Baerlocher, Ch. (1982). 'The X-ray Rietveld System' (Institut für Kristallographie und Petrographie: Zürich).

Baerlocher, Ch. (1986). *Zeolites* **6**, 325–33.

Baerlocher, Ch., and Hepp, A. (1980). Symp. on Accuracy in Powder Diffraction, NBS Special Publication No. 567 (Eds S. Block and C. R. Hubbard), p. 165 (NBS: Washington, D.C.).

Baerlocher, Ch., and Schicker, P. (1987). *Acta Crystallogr.* A **43**, *Suppl.* C-233.

Hecq, M. (1981). *J. Appl. Crystallogr.* **14**, 60–1.

Hepp, A. (1981). Ph.D. Thesis, University of Zürich.

McCusker, L. M., and Baerlocher, Ch. (1984). Proc. 6th Int. Conf. on Zeolites, Reno, U.S.A. (Eds D. Olson and A. Bisio), pp. 812–22 (Butterworths: London).

Mortier, W. J., and Costenoble, M. L. (1973). *J. Appl. Crystallogr.* **6**, 7–11.

Parrish, W., Huang, T. C., and Ayers, G. L. (1980). *Trans. Am. Cryst. Assoc.* **12**, 55–73.

Pyrros, N. P., and Hubbard, C. R. (1983). *J. Appl. Crystallogr.* **16**, 289–94.

Späth, H. (1978). 'Spline-Algorithmen zur Konstruktion glatter Kurven und Flächen', p. 103 (Oldenburg Verlag: München).

Young, R. A., and Wiles, D. B. (1982). *J. Appl. Crystallogr.* **15**, 430–8.

Manuscript received 21 August, accepted 30 October 1987