

*Acta Cryst.* (1989). **A45**, 213–217

## Explicit Fourier Representations of Non-Ideal Hypercentric P.D.F.'s of $|E|$

BY URI SHMUELI

*School of Chemistry, Tel Aviv University, Ramat Aviv, 69 978 Tel Aviv, Israel*

GEORGE H. WEISS

*Division of Computer Research and Technology, Laboratory of Physical Sciences, National Institute of Health, Bethesda, Maryland 20892, USA*

AND ARTHUR J. C. WILSON

*Crystallographic Data Centre, University Chemical Laboratory, Lensfield Road, Cambridge CB2 1EW, England*

(Received 12 August 1987; accepted 19 September 1988)

### Abstract

Explicit non-ideal hypercentric distributions of the magnitude of the normalized structure factor have been derived and investigated for the space group  $P\bar{1}$ . One of the distribution types investigated arises when the asymmetric unit of  $P\bar{1}$  consists of several *identical* centrosymmetric motifs, interrelated by additional non-crystallographic centres of symmetry. The ideal version of such distributions was studied by Rogers & Wilson [*Acta Cryst.* (1953), **6**, 439–449]. The other distribution type studied originates from atomic arrangements in which the asymmetric unit is composed of several unrelated centrosymmetric fragments that may differ in their sizes and chemical compositions. Explicit non-ideal probability density functions (p.d.f.'s) of the magnitude of the normalized structure factor were formulated for both above types of distribution as Fourier series, were evaluated numerically and were compared with appropriately simulated distributions of  $|E|$ . The computations were carried out for a range of atomic compositions, and numbers of unrelated centrosymmetric fragments that comprise the asymmetric unit; both of these factors have a significant influence on the deviation of the hypercentric p.d.f.'s from the values predicted on the basis of the central-limit theorem approximation.

### Introduction

Ideal distributions (*i.e.* those based on the central limit theorem) which account for the presence of non-crystallographic symmetry were first proposed by Lipson & Woolfson (1952) and were further generalized by Rogers & Wilson (1953). Shmueli & Wilson (1982) mention a study of the non-ideal bicentric distribution, in which several explicit orthogonal

polynomials that have the Rogers & Wilson (1953) probability density function (p.d.f.) as their weight function were derived. Another approximate generalization of the Rogers & Wilson ideal p.d.f., to all centrosymmetric space groups and heterogeneous atomic compositions, was put forward by Ghosh & Nigam (1983). These approximate non-ideal p.d.f.'s are characterized by an unusual complexity, and the chance of success for deriving non-ideal hypercentric p.d.f.'s, which lend themselves to accurate computation, looked remote. The introduction of the Fourier method into crystallographic intensity statistics (Shmueli, Weiss, Kiefer & Wilson, 1984) led to a relatively straightforward derivation of an analytically exact and accurately computable bicentric p.d.f. (Shmueli, Weiss & Kiefer, 1985). A major role, in this novel approach, is played by the characteristic function corresponding to the required p.d.f. In fact, the knowledge of this function suffices for an exact and explicit formulation of the coefficients of the Fourier series which represents the required p.d.f. Similar investigations of hypercentric distributions of higher orders can now be readily undertaken, at least for the space groups leading to a simple functional form of the structure factor.

The purpose of the present work is (i) to derive exact and explicit functional forms of two types of non-ideal hypercentric p.d.f.'s based on an asymmetric unit of the space group  $P\bar{1}$ , which (*a*) consists of several identical centrosymmetric subunits, of arbitrary composition, that are interrelated by additional centres of symmetry [as in Rogers & Wilson (1953)], or (*b*) is built up from several unrelated centrosymmetric subunits of arbitrarily different sizes and chemical compositions, and (ii) to compute the p.d.f.'s, for a range of subunit numbers, sizes and compositions, to good accuracy and to test the results against histograms of appropriately simulated  $|E|$  values.

### Derivation of multicentric p.d.f.'s in $P\bar{1}$

#### Case 1: Identical centrosymmetric fragments related by centres of symmetry

In this case we seek to determine the p.d.f. for a  $P\bar{1}$  structure, the asymmetric unit of which is composed of  $2^{n-2}$  identical centrosymmetric motifs, related by  $n-1$  non-crystallographic centres of symmetry (Rogers & Wilson, 1953). In contrast to that reference, the fundamental motif is allowed to have any chemical composition. The non-crystallographic centres of symmetry are located at the rationally independent positions  $\mathbf{d}_i$ ,  $i=2, \dots, n$ . Note that the integer  $n$  refers, in the present case, to *independent* centres of symmetry, including the crystallographic one; thus  $n=2$  corresponds to the bicentrosymmetric arrangement.

If we follow the derivation of hypercentric distributions given by Rogers & Wilson (1953), we find that the normalized structure factor  $E$  for the above arrangement is given by

$$E = 2^n \cos \psi_2 \cos \psi_3 \dots \cos \psi_n \sum_{j=1}^{m/2} n_j \cos \theta_j, \quad (1)$$

where  $\psi_k = 2\pi \mathbf{h} \cdot \mathbf{d}_k$  and  $\theta_j = 2\pi \mathbf{h} \cdot \mathbf{r}_j$ ,  $\mathbf{h}$  being a reciprocal-lattice vector, and  $\mathbf{r}_j$  the atomic position vector of the  $j$ th atom within the asymmetric subunit;  $\mathbf{d}_k$ ,  $k=2, \dots, n$ , are the position vectors of the independent non-crystallographic centres of symmetry,  $n_j$  is the normalized scattering factor of the  $j$ th atom, and  $m$  is the number of atoms in the fundamental centrosymmetric fragment.

The required characteristic function is given by

$$C(\omega_1) = \langle \exp(i\omega_1 E) \rangle, \quad (2)$$

which, upon substitution of  $E$  from (1) and integration over the angular variables  $\theta_j$ ,  $j=1, \dots, m/2$ , becomes

$$C(\omega_1) = (2/\pi)^{n-1} \int_0^{\pi/2} \dots \int_0^{\pi/2} d\psi_2 \dots d\psi_n \times \left[ \prod_{j=1}^{m/2} J_0(2^n \omega_1 n_j \cos \psi_2 \dots \cos \psi_n) \right] \quad (3)$$

and the Fourier coefficient for the p.d.f. of  $|E|$ , to be used with (10), is obtained from (3) as indicated below.

The limiting form of the above p.d.f., for a structure composed of a large number of nearly equal atoms in the unit cell, is given by equation (12) of Rogers & Wilson (1953). The latter equation can be rewritten in terms of the magnitude of the normalized structure factor as

$$p(|E|) = [2^n / \pi^{(2n-1)}]^{1/2} \times \int_0^{\pi/2} \dots \int_0^{\pi/2} \exp[-(E^2/2^n) \sec^2 \psi_2 \dots \sec^2 \psi_n] \times \sec \psi_2 \dots \sec \psi_n d\psi_2 \dots d\psi_n. \quad (4)$$

Equation (12) of Rogers & Wilson (1953) has recently been represented in terms of Meijer's  $G$  function (Wilson, 1987), thereby eliminating the multiple integration. The present comparative calculations involve a direct use of (4) in evaluating the tricentric p.d.f. for the equal-atom case and its comparison with the result obtained from (3), as used with the Fourier p.d.f. (10).

#### Case 2: $n-1$ unrelated centres in asymmetric unit

We now wish to find the p.d.f. of  $|E|$  for a  $P\bar{1}$  structure, the asymmetric unit of which consists of  $n-1$  unrelated centrosymmetric subunits, internally related by  $n-1$  non-crystallographic centres of symmetry located at  $\mathbf{d}_i$ ,  $i=2, \dots, n$ . The  $n-1$  non-crystallographic centres of symmetry are assumed to occupy rationally independent positions, and the subunits may differ in their sizes and chemical compositions. Clearly, for  $n=2$  we obtain the bicentrosymmetric arrangement that leads to the non-ideal distribution. Let the  $i$ th subunit be centred at  $\mathbf{d}_i$ , and contribute to the unit-cell contents  $N_i$  atoms. Its contribution to the normalized structure factor is given by

$$E^{(i)}(\mathbf{h}) = 4 \sum_{j=1}^{N_i/4} n_j \cos(2\pi \mathbf{h} \cdot \mathbf{d}_i) \times \cos[2\pi \mathbf{h} \cdot (\mathbf{r}_j - \mathbf{d}_i)], \quad (5)$$

and has the same functional form as that of the structure factor for the bicentrosymmetric arrangement (Shmueli, Weiss & Kiefer, 1985). Since the normalized structure factor for the above-defined  $n$ -centrosymmetric arrangement is just

$$E(\mathbf{h}) = \sum_{i=2}^n E^{(i)}(\mathbf{h}), \quad (6)$$

the characteristic function for this arrangement is a product of  $n-1$  characteristic functions corresponding to terms of the form of (5). The latter characteristic function has the same form as that of the bicentric p.d.f. (Shmueli, Weiss & Kiefer, 1985), and hence

$$C(\omega_1) = \langle \exp(i\omega_1 E) \rangle \quad (7)$$

$$= \prod_{i=2}^n C^{(i)}(\omega_1), \quad (8)$$

where

$$C^{(i)}(\omega_1) = (2/\pi) \int_0^{\pi/2} \left[ \prod_{j=1}^{N_i/4} J_0(4\omega_1 n_j^{(i)} \cos \theta) \right] d\theta, \quad (9)$$

where  $n_j^{(i)}$  is the normalized scattering factor of the  $j$ th atom in the asymmetric part of the  $i$ th centrosymmetric fragment of the structure.

The general expression for the non-ideal centric Fourier p.d.f. of  $|E|$  is given by

$$p(|E|) = \alpha \left[ 1 + 2 \sum_{k=1}^{\infty} C(\pi \alpha k) \cos(\pi \alpha k |E|) \right], \quad (10)$$

where  $\alpha$  is the reciprocal of the maximum value of  $|E|$ , and  $C(\pi\alpha k)$  - the value of the characteristic function at the point  $\pi\alpha k$  - is the composition- and symmetry-dependent Fourier coefficient [see, for example, Shmueli & Weiss (1985, 1987)]. In the present work, the functional forms of  $C$  are given by (3) and (7)-(9) for cases 1 and 2, respectively.

### Illustrative examples

The examples below involve simulation of the distributions, and their comparison with the theoretical p.d.f.'s. For case 1 above, (1) served the purpose of the simulation, and the theoretical p.d.f. for this case is computed from (10), with coefficients derived from the characteristic function (3). In distributions pertaining to case 2 above, (5) has been used for the simulation, and the theoretical p.d.f.'s were computed from (10) with coefficients given by (7)-(9) as described above. The simulation of the distributions of  $|E|$  in both cases outlined above consisted of sampling the structure-factor expressions while replacing the scalar products  $\mathbf{h} \cdot \mathbf{d}_i$  and  $\mathbf{h} \cdot \mathbf{r}_j$  by computer-generated random numbers, uniform in (0, 1). The values of 3000  $|E|$ 's have so been simulated in each calculation, and their histogram was constructed in the  $[0, 3]$  range of  $|E|$ , in thirty equal channels. For the purpose of the comparison, the values of the theoretical p.d.f., taken at the midpoints of the histogram channels, were scaled to the histogram (Shmueli *et al.*, 1984).

The numerical integrations which were carried out in this work were found to be most conveniently done by employing Romberg's adaptive procedure (Davis & Rabinowitz, 1967). This method seems to overcome some difficulties associated with the presence of oscillatory functions in the integrands. The subroutine *DCADRE* of the popular mathematical program library IMSL was employed for this purpose.

Fig. 1 shows distributions belonging to case 1. This is a Rogers & Wilson type of hypercentrosymmetry, with a heterogeneous atomic composition displayed in Fig. 1(a), and the equal-atom case shown in Fig. 1(b). Both figures contain the simulated histograms, the non-ideal p.d.f.'s computed from (10), with coefficients derived from (3), and the ideal Rogers & Wilson (1953) p.d.f.'s computed directly from (4) (for  $n=3$ ) by numerical integration. The Fourier p.d.f.'s in Fig. 1 show some convergence problems, and the series-truncation ripples are more marked in the equal-atom case than for the heterogeneous composition. Such ripples can be smoothed out using, for example, the Lanczos sigma factors (Hamming, 1973), but it was decided not to apply this device since the weighting factors tend to distort the p.d.f. at its low end.

There is very good agreement between the histograms in Fig. 1 and the corresponding Fourier p.d.f.'s. In the equal-atom case (Fig. 1a) there is also a very

good agreement between the ideal and the non-ideal p.d.f.'s, as could have been expected. In the presence of heavy atoms (Fig. 1b) there is a significant discrepancy between the ideal and the non-ideal p.d.f.'s. However, the discrepancy is much smaller than in the case of the bicentric distribution (Shmueli, Weiss & Kiefer, 1985). This is not surprising since the

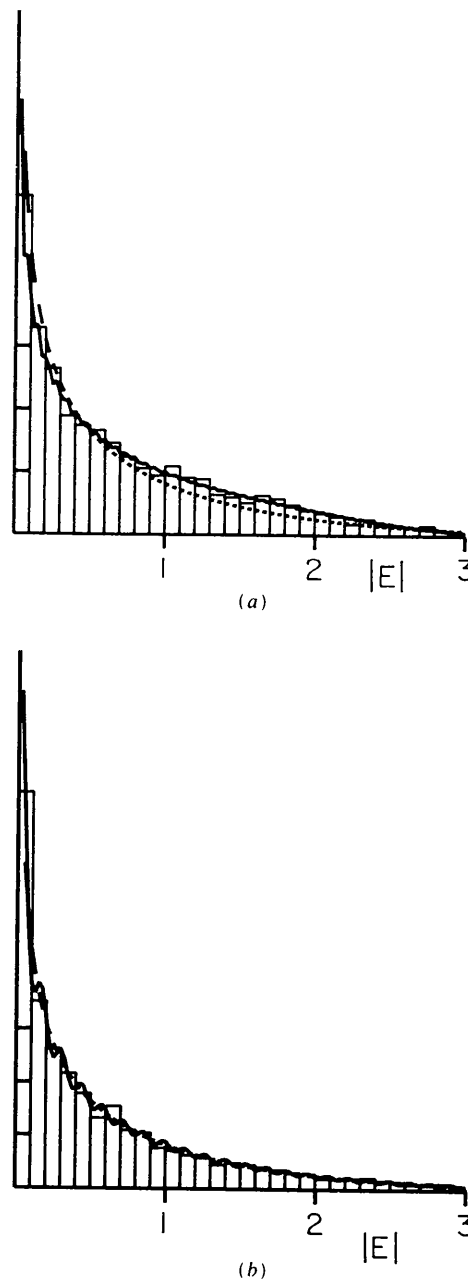


Fig. 1. Histograms, ideal and non-ideal p.d.f.'s for hypercentric models of case 1. The construction of the histograms is outlined in the text. (a)  $C_{14}U$  subunit,  $n=3$ ; solid: equations (10) and (3); dashed: Rogers & Wilson (1953) ideal hypercentric p.d.f. (4). (b)  $C_{15}$  subunit,  $n=3$ ; solid: equations (10) and (3); dashed: Rogers & Wilson (1953) ideal hypercentric p.d.f. (4).

asymmetric unit in the tricentrosymmetric arrangement contains two centrosymmetric fragments and its total composition is  $C_{56}U_4$ . This degree of heterogeneity is lower than that of  $C_{28}U_2$ , which was the model asymmetric unit in the study of the bicentric case.

Fig. 2 illustrates distributions belonging to case 2 (two unrelated centrosymmetric motifs in the asym-

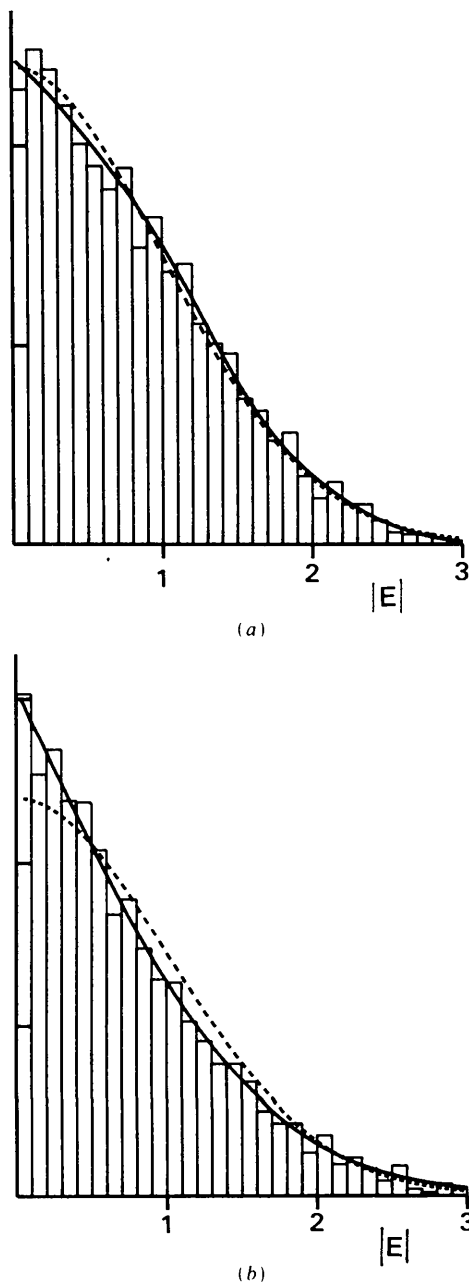


Fig. 2. Histograms, ideal and non-ideal p.d.f.'s for hypercentric models of case 2. The construction of the histograms is outlined in the text. (a)  $C_{14}U$  subunit,  $n = 3$ ; solid: equations (10) and (7)-(9); dashed: Wilson (1949) centric p.d.f. (b)  $C_{15}$  subunit,  $n = 3$ ; solid: equations (10) and (7)-(9); dashed: Wilson (1949) centric p.d.f.

Table 1. Discrepancy measures for comparison of simulated and theoretical distributions

The assumed composition of the asymmetric subunit is  $C_{14}X$ ,  $X = C$  or  $U$ , and all the subunits are taken to be identical, even in case 2 calculations. The effective number of histogram channels is 29 or 30 in all simulations that are referred to in the table. The quantity  $n - 1$  denotes the number of independent non-crystallographic centres of symmetry in case 1 calculations, and the number of unrelated centrosymmetric fragments in case 2 calculations. The subscript 'centric' on  $\chi^2$  refers to the comparison of the simulated histogram with the Wilson (1949) p.d.f., and the subscript 'eq. (4)' on  $\chi^2$  refers to the comparison of the simulated histogram with the Rogers & Wilson (1953) ideal hypercentric p.d.f. In the calculation of  $\chi^2$  for case 1 the first bin of the histogram was omitted, since the p.d.f.'s tend to infinity at the origin, and at the midpoint of the first bin their values are unusually large.

Case 1			
$n - 1$	Composition	$\chi^2$	$\chi_{eq(4)}^2$
2	$C_{15}$	32.28	19.15
2	$C_{14}U$	26.00	133.18
Case 2			
$n - 1$	Composition	$\chi^2$	$\chi_{centric}^2$
2	$C_{15}$	26.68	72.35
5	$C_{15}$	16.01	18.93
2	$C_{14}U$	17.77	26.47
5	$C_{14}U$	16.01	15.48

metric unit of  $P\bar{1}$ ); Fig. 2(a) displays the effect of atomic heterogeneity, and Fig. 2(b) pertains to the equal-atom case. The dashed curve in both figures is the Wilson (1949) centric p.d.f. It is seen that the distribution arising from two centrosymmetric unrelated units is remarkably close to the ideal centric one, although the heterogeneous composition gives rise to a better agreement with the centric p.d.f. than does the equal-atom composition in case 2. In the heterogeneous composition we have four equal heavy atoms in the asymmetric unit of  $P\bar{1}$ , and since the centrosymmetric subunits were assumed to be unrelated - and the heavy atoms taken as uranium - we have (very roughly) a  $P\bar{1}$  structure having four equal atoms in the asymmetric unit. When the centrosymmetric subunits are chosen to be different, and there is appreciable heterogeneity, the deviation from the centric p.d.f. is more significant. The small but significant deviation from the centric p.d.f. shown by the equal-atom case 2 distribution is the effect of non-crystallographic centrosymmetry on the distribution. When the number of the unrelated centrosymmetric subunits is increased to five, the agreement with the centric p.d.f. becomes very good. The geometry of the case 2 model is related to that underlying the sesquicentric distribution (Wilson, 1956).

In all above cases, the agreement between the Fourier p.d.f.'s and the simulated histograms is (visually) very close, and is characterized by good  $\chi^2$  values, shown in Table 1. However, it should be pointed out that for the compositions chosen the ideal p.d.f.'s also perform quite well, except in case 1 for extreme heterogeneity and in case 2 for two equal-atom frag-

ments. When the number of unrelated fragments increases the Fourier p.d.f.'s and the histograms are seen to approach the ideal centric p.d.f. for homogeneous as well as heterogeneous compositions.

We wish to thank Mrs Zafra Stein for computing and programming assistance. All the computations were carried out on the Cyber 170-855 computer at the Tel Aviv University Computation Centre. This research was supported in part by grant No. 84-00076 from the United States-Israel Binational Science Foundation (BSF), Jerusalem, Israel.

#### References

DAVIS, P. J. & RABINOWITZ, P. (1967). *Numerical Integration*. London: Blaisdel.

GHOSH, S. & NIGAM, G. D. (1983). *Z. Kristallogr.* **163**, 61-68.  
 HAMMING, R. W. (1973). *Numerical Methods for Scientists and Engineers*. New York: McGraw-Hill.  
 LIPSON, H. & WOOLFSON, M. M. (1952). *Acta Cryst.* **5**, 680-682.  
 ROGERS, D. E. & WILSON, A. J. C. (1953). *Acta Cryst.* **6**, 439-449.  
 SHMUELI, U. & WEISS, G. H. (1985). In *Structure and Statistics in Crystallography*, edited by A. J. C. WILSON, pp. 53-66. Guilderland: Adenine Press.  
 SHMUELI, U. & WEISS, G. H. (1987). *Acta Cryst.* **A43**, 93-98.  
 SHMUELI, U., WEISS, G. H. & KIEFER, J. E. (1985). *Acta Cryst.* **A41**, 55-59.  
 SHMUELI, U., WEISS, G. H., KIEFER, J. E. & WILSON, A. J. C. (1984). *Acta Cryst.* **A40**, 651-660.  
 SHMUELI, U. & WILSON, A. J. C. (1982). In *Conformation in Biology*, edited by R. SRINIVASAN & R. H. SARMA, pp. 383-388. Guilderland: Adenine Press.  
 WILSON, A. J. C. (1949). *Acta Cryst.* **2**, 318-321.  
 WILSON, A. J. C. (1956). *Acta Cryst.* **9**, 143-144.  
 WILSON, A. J. C. (1987). *Acta Cryst.* **A43**, 554-556.

*Acta Cryst.* (1989). **A45**, 217-227

## Géométrie des Relations d'Orientation dans la Symétrie Hexagonale. Dimension de la Coïncidence

PAR S. HAGÈGE\* ET G. NOUET

LERMAT, ISMRa Université, 14032 Caen CEDEX, France

(Reçu le 4 décembre 1987, accepté le 22 août 1988)

#### Abstract

The main characteristics of the orientation relationship between two hexagonal lattices are simply denoted by four parameters ( $M$ ,  $U$ ,  $V$ ,  $W$ ). In the case of a tridimensional coincidence orientation relationship, for a rational value of  $(c/a)^2$ , these four parameters become four prime integers ( $m$ ,  $un$ ,  $vn$ ,  $wn$ ). The equivalence class of the orientation relationship may be represented by 12 equivalent descriptions for which the indices of the rotation axes are noted on the basis of one crystal. This original representation can lead to the concept of spatial distribution of the equivalent rotation axes, a distribution which is strongly related to the  $c/a$  ratio. In real materials ( $c/a$  irrational), bidimensional coincidence orientation relationships can describe a large number of grain boundaries.

#### 1. Introduction

Les études sur la géométrie des orientations de coïncidence pour la symétrie hexagonale ont vraiment débuté par les travaux de Bruggeman, Bishop & Hartt (1972). Ils ont recherché de façon systématique, mais non exhaustive, les coïncidences qui apparaissaient

par rotation autour d'un axe (simple) choisi *a priori*. C'était l'application de la notion de fonction génératrice que Ranganathan (1966) avait développée pour la symétrie cubique. Une deuxième approche, beaucoup plus générale, a été établie par Warrington & Bufalini (1971) et Warrington (1975); elle considère la matrice de rotation décrivant l'orientation de coïncidence; ils ont montré qu'une telle matrice, écrite sur la base du crystal, doit avoir tous ses éléments rationnels. Ceci a été démontré la suite par Grimmer (1976). Dès lors, il a été évident que les coïncidences tridimensionnelles ne pouvaient exister que pour des rapports  $(c/a)^2$  rationnels ou que pour des cas particuliers, indépendants du rapport  $c/a$  (rotations d'axe  $c$  ou rotations  $180^\circ$  d'axes dans le plan de base). Parallèlement depuis le début des années 1970, Grimmer a développé le concept de quaternion (quatre nombres entiers, premiers entre eux) pour caractériser la rotation de coïncidence. D'abord pour la symétrie cubique (Grimmer, 1973), puis en généralisant aux autres symétries (Grimmer, 1980), cet auteur a montré la validité et l'utilité de ce concept. Bonnet (1980) a aussi utilisé ce concept pour décrire les orientations de coïncidence entre deux réseaux quelconques. Une autre approche a été proposée par Bonnet & Cousineau (1977) et Bonnet, Cousineau & Warrington (1981); elle est associée à une méthode numérique pour la détermination directe de la maille

\* En détachement provisoire à l'Institut des Sciences Industrielles de l'Université de Tokyo, Japon.