

The Calculation of the Second Moments for the Values of Fourier Syntheses with Random Structure Factors

BY V. YU. LUNIN

Research Computing Centre, USSR Academy of Sciences, 142292, Pushchino, Moscow Region, USSR

(Received 20 May 1988; accepted 26 January 1989)

Abstract

Formulas to calculate variances and correlations of the values of Fourier syntheses at individual points are derived for an arbitrary symmetry group when structure factors contain independent random errors. It is shown that variance at a given point in real space depends on the values the weighted Patterson function takes at the points corresponding to Harker interatomic vectors.

1. Introduction

It rarely occurs in practice that one knows accurate values of structure factors F_s while calculating the Fourier synthesis

$$\rho_r = \rho(\mathbf{r}) = (1/|V|) \sum_{s \in S} F_s \exp[-2\pi i(\mathbf{s}, \mathbf{r})], \quad (1)$$

where S is a set of points in the reciprocal-space lattice and $|V|$ is the volume of the unit cell V . As a rule, information on the structure-factor phases φ_s exists as distributions of probabilities $P_s(\varphi_s)$, which means some uncertainty in the values of phases. The moduli of structure factors may also contain errors or even be wholly undetermined. In this situation ρ_r can take a spectrum of values. The situation may be modeled by a set $\{\rho_r\}_{r \in V}$ regarded as a random field defined by (1) with random $\{F_s\}_{s \in S}$.

An estimate of ρ_r is usually its mean:

$$\langle \rho_r \rangle = (1/|V|) \sum_{s \in S} \langle F_s \rangle \exp[-2\pi i(\mathbf{s}, \mathbf{r})]. \quad (2)$$

Here angle brackets $\langle \rangle$ denote the mean of a random value.

If structure-factor moduli $|F^o(\mathbf{s})|$ are assumed to be accurate and phase values are characterized by the probability distributions $P_s(\varphi_s)$, then (2) is the 'best' synthesis:

$$\begin{aligned} \langle \rho_r \rangle &= (1/|V|) \sum_{s \in S} m(\mathbf{s}) |F^o(\mathbf{s})| \\ &\quad \times \exp[i\varphi^{\text{best}}(\mathbf{s})] \exp[-2\pi i(\mathbf{s}, \mathbf{r})], \end{aligned}$$

(Blow & Crick, 1959), in which

$$\begin{aligned} m(\mathbf{s}) \exp[i\varphi^{\text{best}}(\mathbf{s})] &= \langle \exp(i\varphi_s) \rangle \\ &= \int_0^{2\pi} \exp(i\varphi) P_s(\varphi) d\varphi. \end{aligned}$$

To estimate how close $\langle \rho_r \rangle$ is to the real value of the synthesis at a given point \mathbf{r} , the common practice is to calculate the root mean square deviation as

$$\sigma_r = \langle (\rho_r - \langle \rho_r \rangle)^2 \rangle^{1/2}.$$

This value may be interpreted as the level of noise in the synthesis (2) at point \mathbf{r} .

For the case where the values $|F^o(\mathbf{s})|$ are regarded as accurate, the value of σ_r^2 averaged over the unit cell is known to be

$$(1/|V|) \int_V \sigma_r^2 dV_r = (1/|V|)^2 \sum_{s \in S} [1 - m^2(\mathbf{s})] |F^o(\mathbf{s})|^2$$

(Blundell & Johnson, 1979).

The last equation gives an average estimate for the spread in ρ_r values about the mean. In this paper we derive formulas to estimate the variance σ_r^2 at individual points for any crystallographic group. In a more general formulation, we derive expressions to calculate covariances

$$\text{cov}(\rho_r, \rho_u) = \langle (\rho_r - \langle \rho_r \rangle)(\rho_u - \langle \rho_u \rangle) \rangle,$$

for the case when the coefficients F_s are independent random values, in the sense we define below. At $\mathbf{u} = \mathbf{r}$ the latter expression gives the formula for σ_r^2 .

2. Main results

We have assumed that the values $\{\rho_r\}_{r \in V}$ are defined by (1) where $\{F_s\}_{s \in S}$ are complex valued and random. We now assume that the first and the second moments (mean and covariances) are known for each of the random values F_s , so, in particular, the following values are known:

$$\begin{aligned} A_s &= \langle |F_s|^2 \rangle - |\langle F_s \rangle|^2, \\ B_s &= \langle F_s^2 \rangle - \langle F_s \rangle^2. \end{aligned} \quad (3)$$

If the function $\rho(\mathbf{r})$ is real and has the symmetry of the group

$$\Gamma = \{g_\nu\}_{\nu=1}^n = \{(G_\nu, \mathbf{t}_\nu)\}_{\nu=1}^n,$$

such that

$$\rho(\mathbf{r}_\nu) = \rho[g_\nu(\mathbf{r})] = \rho(G_\nu \mathbf{r} + \mathbf{t}_\nu), \quad \nu = 1, \dots, n, \quad (4)$$

then the structure factors F_s have the corresponding symmetry

$$F_{G_s^T} = \exp[-2\pi i(\mathbf{s}, \mathbf{t}_\nu)] F_s, \quad F_{-s} = \overline{F_s}. \quad (5)$$

Here the overbar indicates complex conjugation.

In particular, the values A_s and B_s have the symmetry

$$\begin{aligned} A_{G_s^T} &= A_s, & A_{-s} &= A_s, \\ B_{G_s^T} &= \exp[-2\pi i(\mathbf{s}, 2\mathbf{t}_\nu)] B_s, & B_{-s} &= \overline{B_s}. \end{aligned} \quad (6)$$

For example, if $\rho(\mathbf{r})$ has the symmetry group $P2_12_12_1$, then the values A_s correspond to the symmetry group $Pmmm$, and the values B_s to the symmetry group $P222$.

Equation (5) shows that the random values $\{F_s\}_{s \in S}$ cannot be regarded as independent. We may, however, assume independence of random values $\{F_s\}_{s \in Sa}$ with indexes s lying in the 'asymmetric part' of the set S ; these are values without symmetry relations. We define the 'asymmetric part' Sa of the set S to be a set of grid points of the reciprocal-space lattice such that:

(a) each point s in S may be found from a point s' in Sa via one of the symmetry transforms of the type $\pm G_\nu^T$;

(b) these transforms, when applied to the points of Sa , do not yield points which are outside S ;

(c) no points in Sa have symmetry relations of the type $\pm G_\nu^T$, excepted for the identical transform.

For a point s we introduce the multiplicity $\tau(s)$ with respect to the group $\{\pm G_\nu^T\}$ in the form

$$\begin{aligned} \tau(s) &\text{ is the number of points of the type } \pm G_\nu^T s, \\ (\nu &= 1, \dots, n) \text{ coinciding with } s. \end{aligned} \quad (7)$$

For example, for the $P2_12_12_1$ group the asymmetric part may be given by $h \geq 0$, $k \geq 0$ and $l \geq 0$. In this case $\tau(s) = 2^\omega$ where ω is the number of zero indexes among h , k and l .

Now we can formulate our main result.

Theorem 1

Assume the real function $\rho(\mathbf{r})$ has the symmetry of crystallographic group Γ (4) and is given by (1) where $\{F_s\}_{s \in Sa}$ are independent, complex valued and random, and the symmetry (5) is involved. Then for arbitrary points \mathbf{r} and \mathbf{u}

$$\text{cov}(\rho_r, \rho_u) = \sum_{\kappa=1}^n P[\mathbf{r} - \mathbf{g}_\kappa(\mathbf{u})] + \sum_{\kappa=1}^n Q[\mathbf{r} + \mathbf{g}_\kappa(\mathbf{u})], \quad (8)$$

where the functions $P(\mathbf{r})$ and $Q(\mathbf{r})$ are determined by

$$\begin{aligned} P(\mathbf{r}) &= (1/|V|) \sum_{s \in S} [\tau(s)|V|]^{-1} A_s \\ &\quad \times \exp[-2\pi i(\mathbf{s}, \mathbf{r})], \end{aligned} \quad (9)$$

and

$$\begin{aligned} Q(\mathbf{r}) &= (1/|V|) \sum_{s \in S} [\tau(s)|V|]^{-1} B_s \\ &\quad \times \exp[-2\pi i(\mathbf{s}, \mathbf{r})], \end{aligned} \quad (10)$$

with the coefficients A_s , B_s and $\tau(s)$ found by (3) and (7).

Consequence. The mean square deviation of ρ_r can be given by

$$\sigma_r^2 = \sum_{\kappa=1}^n P[\mathbf{r} - \mathbf{g}_\kappa(\mathbf{r})] + \sum_{\kappa=1}^n Q[\mathbf{r} + \mathbf{g}_\kappa(\mathbf{r})].$$

Remark 1. The symmetry conditions (5) may either make some of the reflections vanish [$F_s = 0$ if for some $G_\nu^T s = s$ and $(\mathbf{s}, \mathbf{t}_\nu) \neq 0 \pmod{1}$] or locate them symmetrically about the centre (if $G_\nu^T s = -s$, then either $\varphi_s = \varphi^0 = \pi(\mathbf{s}, \mathbf{t}_\nu)$, or $\varphi_s = \varphi^0 + \pi$). We assume that the relevant probability distributions for F_s meet these requirements.

Theorem 2

The variance averaged over the unit cell should be given by

$$(1/|V|) \int_V \sigma_r^2 dV_r = (1/|V|)^2 \sum_{s \in S} A_s. \quad (11)$$

Remark 2. If the group Γ contains the transform $g_\nu(\mathbf{r}) = -\mathbf{r}$, which means $\rho(\mathbf{r})$ with the symmetry centre, it is easy to see that F_s and $A_s = B_s = \langle F_s^2 \rangle - \langle F_s \rangle^2$ are real, and (8) has the form

$$\text{cov}(\rho_r, \rho_u) = 2 \sum_{\kappa=1}^n P[\mathbf{r} - \mathbf{g}_\kappa(\mathbf{r})].$$

3. Examples

3.1. Analysis of Fourier syntheses with undetermined phases

In order that equations (9)–(10) have a more obvious sense, we want to examine a particular case. Assume that the structure-factor moduli $|F^o(\mathbf{s})|$ are accurate and that each of the phases φ_s may with equal probability be either $\varphi^o(\mathbf{s}) + \Delta(\mathbf{s})$ or $\varphi^o(\mathbf{s}) - \Delta(\mathbf{s})$. This example illustrates the use of the SIR method to determine phases in protein crystallography. It is easy to see that in this case

$$\begin{aligned} \langle \rho_r \rangle &= (1/|V|) \sum_{s \in S} \cos \Delta(\mathbf{s}) |F^o(\mathbf{s})| \\ &\quad \times \exp[i\varphi^o(\mathbf{s})] \exp[-2\pi i(\mathbf{s}, \mathbf{r})], \\ A_s &= \sin^2 \Delta(\mathbf{s}) |F^o(\mathbf{s})|^2, \\ B_s &= -\sin^2 \Delta(\mathbf{s}) \exp[i2\varphi^o(\mathbf{s})] |F^o(\mathbf{s})| \end{aligned}$$

and covariances can be determined by (8) with

$$\begin{aligned} P(\mathbf{r}) &= (1/|V|) \sum_{s \in S} [\tau(s)|V|]^{-1} \sin^2 \Delta(\mathbf{s}) |F^o(\mathbf{s})|^2 \\ &\quad \times \exp[-2\pi i(\mathbf{s}, \mathbf{r})] = \sigma(\mathbf{r}) * \sigma(-\mathbf{r}), \end{aligned}$$

$$Q(\mathbf{r}) = -\sigma(\mathbf{r}) * \sigma(\mathbf{r})$$

where

$$\begin{aligned} \sigma(\mathbf{r}) &= (1/|V|) \sum_{s \in S} [\tau(s)|V|]^{-1/2} |\sin \Delta(\mathbf{s})| |F^o(\mathbf{s})| \\ &\quad \times \exp[i\varphi^o(\mathbf{s})] \exp[-2\pi i(\mathbf{s}, \mathbf{r})]. \end{aligned}$$

In this way, $P(\mathbf{r})$ is the Patterson synthesis corresponding to the function $\sigma(\mathbf{r})$. The function $Q(\mathbf{r})$ is the self-convolution of $\sigma(\mathbf{r})$. It is easy to see that functions $\langle \rho_r \rangle$ and $\sigma(\mathbf{r})$ bear a relation to the 'phase-error function' $f(\mathbf{r})$ (Silva & Viterbo, 1980). To an accuracy of a multiplier,

$$f(\mathbf{r}) * \rho^\circ(\mathbf{r}) = \langle \rho_r \rangle + i\sigma(\mathbf{r}),$$

where

$$\begin{aligned} \rho^\circ(\mathbf{r}) &= (1/|V|) \sum_{\mathbf{s} \in S} |F^\circ(\mathbf{s})| \\ &\times \exp [i\varphi^\circ(\mathbf{s})] \exp [-2\pi i(\mathbf{s}, \mathbf{r})]. \end{aligned}$$

The function $P(\mathbf{r})$ may be assumed to be the weighted Patterson synthesis in the more general case when the probability distributions $P_s(\varphi_s)$ are arbitrary. Then

$$A_s = [1 - m^2(\mathbf{s})] |F^\circ(\mathbf{s})|^2.$$

The function $Q(\mathbf{r})$ is in this case more complex.

3.2. Analysis of Patterson syntheses

Now we want to describe a situation when the phases in (1) are accurate and the moduli are random. This is the case of the Patterson synthesis

$$p(\mathbf{r}) = (1/|V|) \sum_{s \leq s_{\max}} |F^\circ(\mathbf{s})|^2 \exp [-2\pi i(\mathbf{s}, \mathbf{r})], \quad (12)$$

in which intensities $I^\circ(\mathbf{s}) = |F^\circ(\mathbf{s})|^2$ for some s have not been measured. In this case the unknown intensities $I(\mathbf{s})$ may be regarded as random values falling under the Wilson statistics, so that for non-centrosymmetric reflections

$$P_s(I) = (1/\Omega) \exp (-I/\Omega),$$

where $\langle I \rangle = \Omega$ (Srinivasan & Parthasarathy, 1976).

The 'best' synthesis in this case will be

$$p_{\text{best}}(\mathbf{r}) = (1/|V|) \sum_{s \leq s_{\max}} \langle |F_s|^2 \rangle \exp [-2\pi i(\mathbf{s}, \mathbf{r})], \quad (13)$$

with

$$\langle |F_s|^2 \rangle = \begin{cases} I^\circ(\mathbf{s}) & \text{if the experimental value} \\ & \text{of } I^\circ \text{ is known,} \\ \Omega(\mathbf{s}) & \text{if it is missing.} \end{cases}$$

Let us stress that for the unknown $I^\circ(\mathbf{s})$ the 'best' synthesis (13) involves the non-zero estimate $\Omega(\mathbf{s})$. In particular, in the limit case with all $I^\circ(\mathbf{s})$ lacking, the 'best' Patterson synthesis consists of one central Gaussian peak only [if $\Omega(\mathbf{s})$ is assumed to be Gaussian].

Variance for (12) is given by

$$\sigma_r^2 = 2 \sum_{\kappa=1}^n (1/|V|)^2 \sum_{s \leq s_{\max}} A_s \exp \{-2\pi i[\mathbf{s}, \mathbf{r} - \mathbf{g}_\kappa(\mathbf{r})]\}$$

where

$$A_s = \begin{cases} 0 & \text{if } I \text{ is known,} \\ \langle I^2 \rangle - \langle I \rangle^2 = \Omega^2(\mathbf{s}) & \text{if not.} \end{cases}$$

3.3. Nonlinear filtration of a noisy synthesis

In their papers Urzhumtsev (1985) and Wang (1985) suggested a method of bounding the region of a macromolecule in a noisy electron-density synthesis. The first step of the method is the search for 'basic points', *i.e.* points which, with great certainty, can be ascribed to the region of the molecule (Urzhumtsev, Lunin & Luzyanina, 1989). In the simplest case the choice may follow from the comparison of the synthesis value at such a point with a threshold. A more accurate procedure implies introducing a probability that every point in the space belongs to the region (Urzhumtsev, Lunin & Luzyanina, 1986). In this case not only the value of the synthesis at the point is required, but also the 'level of noise', that is the spread in values about the mean. In the present paper we give formulas to calculate statistical characteristics of Fourier syntheses.

4. Proofs

4.1. Proof of theorem 1

Let us transform the right-hand side of (1) to the sum over the asymmetric part of the set S . Here we take account of the symmetry (5) and of the fact that points $\pm G_\nu^T \mathbf{s}$ include $\tau(\mathbf{s})$ points coinciding with \mathbf{s} . Then

$$\begin{aligned} \rho_r &= (1/|V|) \sum_{\nu=1}^n \sum_{\mathbf{s} \in Sa} [\tau(\mathbf{s})]^{-1} \\ &\times \{F_{G_\nu^T \mathbf{s}} \exp [-2\pi i(G_\nu^T \mathbf{s}, \mathbf{r})] \\ &+ F_{-G_\nu^T \mathbf{s}} \exp [-2\pi i(-G_\nu^T \mathbf{s}, \mathbf{r})]\} \\ &= (2/|V|) \operatorname{Re} \left\{ \sum_{\mathbf{s} \in Sa} [\tau(\mathbf{s})]^{-1} F_s \right. \\ &\quad \left. \times \sum_{\nu=1}^n \exp [-2\pi i(\mathbf{s}, G_\nu \mathbf{r} + \mathbf{t}_\nu)] \right\} \\ &= \sum_{\mathbf{s} \in Sa} \operatorname{Re} [F_s \theta_s(\mathbf{r})] \end{aligned}$$

where

$$\theta_s(\mathbf{r}) = [\tau(\mathbf{s})|V|]^{-1} \sum_{\nu=1}^n \exp \{-2\pi i[\mathbf{s}, \mathbf{g}_\nu(\mathbf{r})]\}.$$

Since the values $\{F_s\}_{\mathbf{s} \in Sa}$ are independent,

$$\operatorname{cov}(\rho_r, \rho_u) = \sum_{\mathbf{s} \in Sa} \operatorname{cov}[\operatorname{Re} \theta_s(\mathbf{r}) F_s, \operatorname{Re} \theta_s(\mathbf{u}) F_s]. \quad (14)$$

To calculate the sum (14), we use the following statement, which can easily be checked straightforwardly:

If z is complex valued and random, and θ_1 and θ_2 are complex values, then

$$\operatorname{cov}(\operatorname{Re} \theta_1 z, \operatorname{Re} \theta_2 z) = \frac{1}{2} \operatorname{Re} \{ \theta_1 \bar{\theta}_2 A + \theta_1 \theta_2 B \}, \quad (15)$$

where

$$A = \langle |z|^2 \rangle - \langle z \rangle^2, \quad B = \langle z^2 \rangle - \langle z \rangle^2.$$

Since Γ is a group, we can write

$$\begin{aligned} \theta_s(\mathbf{r}) \overline{\theta_s(\mathbf{u})} &= \sum_{\mu, \nu=1}^n [\tau(\mathbf{s}) |V|]^{-2} \\ &\quad \times \exp \{-2\pi i[\mathbf{s}, \mathbf{g}_\nu(\mathbf{r}) - \mathbf{g}_\mu(\mathbf{u})]\} \\ &= 4[\tau(\mathbf{s}) |V|]^{-2} \\ &\quad \times \sum_{\nu, \kappa=1}^n \exp \{-2\pi i[\mathbf{s}, \mathbf{g}_\nu(\mathbf{r}) - \mathbf{g}_\nu[\mathbf{g}_\kappa(\mathbf{u})]]\} \\ &= 4[\tau(\mathbf{s}) |V|]^{-2} \\ &\quad \times \sum_{\nu, \kappa=1}^n \exp \{-2\pi i[G_\nu^T \mathbf{s}, \mathbf{r} - \mathbf{g}_\kappa(\mathbf{u})]\} \quad (16) \end{aligned}$$

and

$$\begin{aligned} \theta_s(\mathbf{r}) \theta_s(\mathbf{u}) &= 4[\tau(\mathbf{s}) |V|]^{-2} \sum_{\nu, \kappa=1}^n \exp \{-2\pi i(\mathbf{s}, 2\mathbf{t}_\nu)\} \\ &\quad \times \exp \{2\pi i[G_\nu^T \mathbf{s}, \mathbf{r} + \mathbf{g}_\kappa(\mathbf{u})]\}. \quad (17) \end{aligned}$$

Then, using (14)-(17), we get the expression

$$\begin{aligned} \text{cov}(\rho_r, \rho_u) &= \sum_{\kappa=1}^n \left(2(1/|V|)^2 \text{Re} \sum_{\mathbf{s} \in S_a} [1/\tau(\mathbf{s})]^2 A_s \right. \\ &\quad \times \sum_{\nu=1}^n \exp \{-2\pi i[G_\nu^T \mathbf{s}, \mathbf{r} - \mathbf{g}_\kappa(\mathbf{u})]\} \\ &\quad + (1/|V|)^2 \text{Re} \sum_{\mathbf{s} \in S_a} [1/\tau(\mathbf{s})]^2 \\ &\quad \times \sum_{\nu=1}^n B_s \exp \{-2\pi i(\mathbf{s}, 2\mathbf{t}_\nu)\} \\ &\quad \left. \times \exp \{-2\pi i[G_\nu^T \mathbf{s}, \mathbf{r} + \mathbf{g}_\kappa(\mathbf{r})]\} \right). \end{aligned}$$

Now the symmetry (6) allows us to pass in the last equation to the summation over the whole set S so that

$$\begin{aligned} &2(1/|V|)^2 \text{Re} \sum_{\mathbf{s} \in S_a} (1/\tau(\mathbf{s}))^2 A_s \\ &\times \sum_{\nu=1}^n \exp \{-2\pi i(G_\nu^T \mathbf{s}, \mathbf{r})\} \\ &= (1/|V|) \sum_{\mathbf{s} \in S_a} \sum_{\nu=1}^n [1/\tau(\mathbf{s})] \{[\tau(G_\nu^T \mathbf{s}) |V|]^{-1} \\ &\quad \times A_{G_\nu^T \mathbf{s}} \exp \{-2\pi i(G_\nu^T \mathbf{s}, \mathbf{r})\} + [\tau(-G_\nu^T \mathbf{s}) |V|]^{-1} \\ &\quad \times A_{-G_\nu^T \mathbf{s}} \exp \{-2\pi i(-G_\nu^T \mathbf{s}, \mathbf{r})\}\} \\ &= (1/|V|) \sum_{\mathbf{s} \in S} [\tau(\mathbf{s}) |V|]^{-1} A_s \exp \{-2\pi i(\mathbf{s}, \mathbf{r})\} \\ &= P(\mathbf{r}). \end{aligned}$$

Analogously,

$$\begin{aligned} &2(1/|V|)^2 \text{Re} \sum_{\mathbf{s} \in S_a} [1/\tau(\mathbf{s})]^2 \sum_{\nu=1}^n B_s \\ &\quad \times \exp \{-2\pi i(\mathbf{s}, 2\mathbf{t}_\nu)\} \exp \{-2\pi i(G_\nu^T \mathbf{s}, \mathbf{r})\} = Q(\mathbf{r}) \end{aligned}$$

where $P(\mathbf{r})$ and $Q(\mathbf{r})$ are functions defined by (9) and (10).

4.2. Proof of theorem 2

Note that

$$\begin{aligned} &(1/|V|) \int_V \exp \{-2\pi i[\mathbf{s}, \mathbf{r} - \mathbf{g}_\nu(\mathbf{r})]\} dV_r \\ &= \exp [2\pi i(\mathbf{s}, \mathbf{t}_\nu)] (1/|V|) \\ &\quad \times \int_V \exp \{-2\pi i[(E - G_\nu^T) \mathbf{s}, \mathbf{r}]\} dV_r \\ &= \begin{cases} 0 & \text{if } G_\nu^T \mathbf{s} \neq \mathbf{s}; \\ \exp \{-2\pi i(\mathbf{s}, \mathbf{t}_\nu)\} & \text{if } G_\nu^T \mathbf{s} = \mathbf{s}. \end{cases} \end{aligned}$$

Since the condition $2\pi(\mathbf{s}, \mathbf{t}_\nu) \neq 0 \pmod{2\pi}$ for $G_\nu^T \mathbf{s} = \mathbf{s}$ gives zero reflections, then

$$\begin{aligned} &(1/|V|) \int_V \sum_{\nu=1}^n P[\mathbf{r} - \mathbf{g}_\nu(\mathbf{r})] dV_r \\ &= (1/|V|)^2 \sum_{\mathbf{s} \in S} [1/\tau(\mathbf{s})] A_s (1/|V|) \\ &\quad \times \int_V \exp \{-2\pi i[\mathbf{s}, \mathbf{r} - \mathbf{g}_\nu(\mathbf{r})]\} \\ &= (1/|V|)^2 \sum_{\mathbf{s} \in S} \tau_+(\mathbf{s}) / \tau(\mathbf{s}) \quad (18) \end{aligned}$$

where $\tau_+(\mathbf{s})$ is the number of points $G_\nu^T \mathbf{s}$ coinciding with \mathbf{s} . Analogously,

$$\begin{aligned} &(1/|V|) \int_V \exp \{-2\pi i[\mathbf{s}, \mathbf{r} + \mathbf{g}_\nu(\mathbf{r})]\} dV_r \\ &= \begin{cases} 0 & \text{if } -G_\nu^T \mathbf{s} \neq \mathbf{s}; \\ \exp \{-2\pi i(\mathbf{s}, \mathbf{t}_\nu)\} & \text{if } -G_\nu^T \mathbf{s} = \mathbf{s}. \end{cases} \end{aligned}$$

The condition $-G_\nu^T \mathbf{s} = \mathbf{s}$ yields centrosymmetric structure factors that may have phases with either $\varphi_s = \varphi^0 = \pi(\mathbf{s}, \mathbf{t}_\nu)$ or $\varphi_s = \varphi^0 + \pi$. In this case $F_s = t_s \exp [i\varphi^0(\mathbf{s})]$ where t_s is a real random value and

$$A_s = \langle t_s^2 \rangle - \langle t_s \rangle^2,$$

$$B_s = \exp (i2\varphi^0) \{ \langle t_s^2 \rangle - \langle t_s \rangle^2 \} = \exp (i2\varphi^0) A_s.$$

Hence,

$$\begin{aligned} &(1/|V|) \int_V \sum_{\nu=1}^n Q[\mathbf{r} + \mathbf{g}_\nu(\mathbf{r})] dV_r \\ &= (1/|V|)^2 \sum_{\mathbf{s} \in S, -G_\nu^T \mathbf{s} = \mathbf{s}} [1/\tau(\mathbf{s})] B_s \exp (-i2\varphi^0) \\ &= (1/|V|)^2 \sum_{\mathbf{s} \in S} [\tau_-(\mathbf{s}) / \tau(\mathbf{s})] A_s, \quad (19) \end{aligned}$$

where $\tau_-(\mathbf{s})$ is the number of points $-G_\nu^T \mathbf{s}$ coinciding with \mathbf{s} . Since it is likely that $\tau(\mathbf{s}) = \tau_-(\mathbf{s}) + \tau_+(\mathbf{s})$, expressions (18) and (19) yield (11).

5. Concluding remarks

Ambiguity in some of the structure-factor values can, in its turn, lead to ambiguous values of the function of electron-density distribution calculated as the sum (1). In this case the most representative (*i.e.* giving the least r.m.s. error) is the 'mean' synthesis (2) which is the general form of the best synthesis of Blow & Crick (1959). However, the possible deviation from the mean may vary for different points in the unit cell and is characterized by r.m.s. error σ_r .

Formulas (8)–(10) estimate the individual values σ_r for the case when the errors in the structure factors are regarded as independent and their spread is known. [This spread is characterized by A_s and B_s in (3)]. The values σ_r are closely related to Harker peaks at weighted Patterson syntheses. The derived formulas may be used by various approaches where knowledge of individual values ρ_r is required.

Acta Cryst. (1989). **A45**, 505–523

Coincidence Orientations of Grains in Rhombohedral Materials

BY HANS GRIMMER

Paul Scherrer Institute, Laboratory of Materials Science, CH-5232 Villigen PSI, Switzerland

(Received 22 July 1988; accepted 15 February 1989)

Abstract

In experimental investigations and computer simulations of the structure and properties of grain boundaries, the results are usually discussed with reference to the special case of coincidence boundaries, where the two neighbouring grains have a three-dimensional lattice of symmetry translations in common. For historical reasons this lattice is called the coincidence site lattice or CSL. A systematic determination of CSL's for the case of grains with a lattice of rhombohedral Bravais type is presented. It is shown that a number of investigations of the structure of grain boundaries in alumina (α -Al₂O₃) have to be reinterpreted in the light of the present results. A central result is the Σ -rhomb theorem, which expresses the ratio Σ of unit-cell volumes of the CSL and the rhombohedral crystal lattice in terms of four integral parameters that describe the axis and angle of the rotation connecting the rhombohedral lattices of the two neighbouring grains and in terms of their axial ratio c/a . Two types of coincidence rotations, *i.e.* of rotations generating CSL's, may be distinguished, *viz* common rotations, which generate CSL's with the same Σ for every value of c/a , and specific rotations, which generate CSL's with a low value of Σ only for a few values of the axial ratio. The Σ -rhomb theorem makes it possible to determine systematically not only

The author is grateful to O. M. Liguinchenko for her help in preparing the manuscript.

References

- BLOW, D. M. & CRICK, F. H. C. (1959). *Acta Cryst.* **12**, 794–802.
 BLUNDELL, T. L. & JOHNSON, L. N. (1979). *Protein Crystallography*. New York: Academic Press.
 SILVA, A. M. & VITERBO, D. (1980). *Acta Cryst.* **A36**, 1065–1070.
 SRINIVASAN, R. & PARTHASARATHY, S. (1976). *Some Statistical Applications in X-ray Crystallography*. New York: Pergamon Press.
 URZHUMTSEV, A. G. (1985). *Use of Local Averaging in the Analysis of Macromolecular Images in Electron Density Maps*. Preprint, USSR Academy of Sciences, Pushchino.
 URZHUMTSEV, A. G., LUNIN, V. YU. & LUZYANINA, T. B. (1986). Tenth Eur. Crystallogr. Meet., Wrocław, Poland. Collected Abstracts, pp. 51–52.
 URZHUMTSEV, A. G., LUNIN, V. YU. & LUZYANINA, T. B. (1989). *Acta Cryst.* **A45**, 34–39.
 WANG, B. C. (1985). *Methods Enzymol.* **115**, 90–112.

all common rotations with Σ up to a given maximum value Σ_c but also all specific rotations with $\Sigma \leq \Sigma_c$ and with c/a in any given interval about the experimental value of c/a for the material in question. It is shown that the multiplicities of the CSL's generated by a given rotation in a hexagonal and in a rhombohedral lattice with the same value of c/a differ by at most a factor 3.

1. Introduction

Metals and ceramics are used in polycrystalline form for most of their applications. The boundaries between the crystallites often control mechanical and corrosion properties of the materials to a large extent. For this reason, great efforts are taken in the production and heat treatment of modern engineering materials to optimize the size of the grains and the impurity content of the boundaries between them as well as the distribution of additional phases. Significant improvements have been obtained in this way, *e.g.* in the toughness and strength of steels or in the tensile strength of ceramics.

Boundaries between regions with the same crystal structure will be considered in the present work. They will be called *grain boundaries* and include the special case of twin boundaries.