Acta Cryst. (1977). A33, 527–531

Strengthening of the Triplet Relationship. II. A New Probabilistic Approach in $P1$

BY C. GIACOVAZZO

Istituto di Mineralogia e Petrografia, Università degli Studi, Bari, Italy

(*Received* 10 *September* 1976; *accepted* 8 *October* 1976)

The probabilistic approach recently described in $P\bar{1}$ [Giacovazzo, *Acta Cryst.* (1976), A32, 967–976] is generalized to non-centrosymmetric space groups. The method enables us to derive the expected value of $\cos(\varphi_{\mathbf{H}_1}+\varphi_{\mathbf{H}_2}-\varphi_{\mathbf{H}_1+\mathbf{H}_2})$ and its variance, given $|E_{\mathbf{H}_1}|$, $|E_{\mathbf{H}_2}|$, $|E_{\mathbf{H}_1+\mathbf{H}_2}|$ and one or more triples ($|E_{\mathbf{K}}|$, $|E_{\mathbf{H}_1+\mathbf{K}}|$, $|E_{\mathbf{H}_1+\mathbf{H}_2+\mathbf{K}}|$).

Preliminary remarks

In a recent paper (Giacovazzo, 1976*a*; from now on denoted as paper I) a probabilistic approach is described in $P\bar{1}$ which gives, from the triples ($|E_{\mathbf{K}}|$, $|E_{\mathbf{H}_1+\mathbf{K}}|$, $|E_{\mathbf{H}_1+\mathbf{H}_2+\mathbf{K}}|$), the probability of the sign of $E_{\mathbf{H}_1}E_{\mathbf{H}_2}E_{\mathbf{H}_1+\mathbf{H}_2}$. The procedure is able to correct the positivity required by the Cochran–Woolfson formula so that negative triplets may in principle be identified.

In this paper the method is generalized to non-centrosymmetric space groups. Unlike preceding approaches (*e.g.* Karle & Hauptman, 1958; Hauptman, 1964; Hauptman, Fisher, Hancock & Norton, 1969; Hauptman, 1970), the triples ($|E_{\mathbf{K}}|$, $|E_{\mathbf{H}_1+\mathbf{K}}|$, $|E_{\mathbf{H}_1+\mathbf{H}_2+\mathbf{K}}|$) will be explored in order to give the expected value of $\cos(\varphi_{\mathbf{H}_1}+\varphi_{\mathbf{H}_2}-\varphi_{\mathbf{H}_1+\mathbf{H}_2})$ and its variance rather than the 'exact' value of the cosine. The general background is described in the first paragraph of paper I, to which the reader is referred.

The joint probability distribution $P(R_{\mathbf{H}_1},R_{\mathbf{H}_2},R_{\mathbf{H}_1+\mathbf{H}_2},R_{\mathbf{K}},R_{\mathbf{H}_1+\mathbf{K}},R_{\mathbf{H}_1+\mathbf{H}_2+\mathbf{K}},\varphi_{\mathbf{H}_1},\varphi_{\mathbf{H}_2},\ldots,\varphi_{\mathbf{H}_1+\mathbf{H}_2+\mathbf{K}})$

Suppose that a crystal structure consists of N identical atoms per unit cell in the space group $P1$. We assume that the reciprocal vectors $\mathbf{H}_1$, $\mathbf{H}_2$, $\mathbf{H}_1+\mathbf{H}_2$, $\mathbf{K}$, $\mathbf{H}_1+\mathbf{K}$, $\mathbf{H}_1+\mathbf{H}_2+\mathbf{K}$ are fixed and the atomic coordinates are the primitive random variables. We introduce the abbreviation

$$E_1=R_1\exp(i\varphi_1)=A_1+iB_1=E_{\mathbf{H}_1},$$
$$E_2=R_2\exp(i\varphi_2)=\ldots\ldots\ldots=E_{\mathbf{H}_2},$$
$$E_6=R_6\exp(i\varphi_6)=\ldots\ldots\ldots=E_{\mathbf{H}_1+\mathbf{H}_2+\mathbf{K}}.$$

By means of the probabilistic approach described in paper I for centrosymmetric crystals and by Giacovazzo (1976*b*) for the non-centrosymmetric crystals, the joint probability distribution function, correct up to and including terms of order $1/N\sqrt{N}$, is

$$P(R_1,\ldots,R_6,\varphi_1,\ldots,\varphi_6)$$
$$\simeq\frac{1}{\pi^6}R_1R_2\ldots R_6\exp(-R_1^2-R_2^2-\ldots-R_6^2)$$
$$\times\Big\{1+\frac{2}{\sqrt{N}}[R_1R_2R_3\cos(\varphi_1+\varphi_2-\varphi_3)$$
$$+R_1R_4R_5\cos(\varphi_1+\varphi_4-\varphi_5)$$
$$+R_2R_5R_6\cos(\varphi_2+\varphi_5-\varphi_6)$$
$$+R_3R_4R_6\cos(\varphi_3+\varphi_4-\varphi_6)]$$
$$-\frac{1}{N}[(1-R_1^2)(1-R_2^2)(1-R_3^2)$$
$$-R_1^2R_2^2R_3^2\cos(2\varphi_1+2\varphi_2-2\varphi_3)$$
$$+(1-R_1^2)(1-R_4^2)(1-R_5^2)$$
$$-R_1^2R_4^2R_5^2\cos(2\varphi_1+2\varphi_4-2\varphi_5)$$
$$+(1-R_2^2)(1-R_5^2)(1-R_6^2)$$
$$-R_2^2R_5^2R_6^2\cos(2\varphi_2+2\varphi_5-2\varphi_6)$$
$$+(1-R_3^2)(1-R_4^2)(1-R_6^2)$$
$$-R_3^2R_4^2R_6^2\cos(2\varphi_3+2\varphi_4-2\varphi_6)]$$
$$-\frac{2}{N}[R_2R_3R_4R_5(1-R_1^2)\cos(\varphi_2-\varphi_3-\varphi_4+\varphi_5)$$
$$+R_1R_3R_5R_6(1-R_2^2)\cos(\varphi_1-\varphi_3-\varphi_5+\varphi_6)+\ldots]$$
$$-\frac{1}{4N}[R_1^4+R_2^4+\ldots+R_6^4-4(R_1^2+\ldots+R_6^2)+12]$$
$$+\frac{2}{N}[R_1R_2R_4R_6\cos(\varphi_1+\varphi_2+\varphi_4-\varphi_6)$$
$$+R_2R_3R_4R_5\cos(\varphi_2-\varphi_3-\varphi_4+\varphi_5)$$
$$+R_1R_3R_5R_6\cos(\varphi_1-\varphi_3-\varphi_5+\varphi_6)]$$
$$-\frac{1}{4N\sqrt{N}}[R_1^3R_2R_3\cos(3\varphi_1-\varphi_2+\varphi_3)$$
$$-3R_1(2-R_1^2)R_2R_3\cos(\varphi_1+\varphi_2-\varphi_3)$$
$$+R_1^3R_4R_5\cos(3\varphi_1-\varphi_4+\varphi_5)$$
$$-3R_1(2-R_1^2)R_4R_5\cos(\varphi_1+\varphi_4-\varphi_5)+\ldots]$$
$$+\frac{1}{N\sqrt{N}}[\tfrac{1}{3}R_1^3R_2^3R_3^3\cos(3\varphi_1+3\varphi_2-3\varphi_3)$$
$$-R_1R_2R_3(2-R_1^2)(2-R_2^2)(2-R_3^2)\cos(\varphi_1+\varphi_2-\varphi_3)$$

$$+\tfrac{1}{3}R_1^3R_4^3R_5^3\cos(3\varphi_1+3\varphi_4-3\varphi_5)$$
$$-2R_1R_2R_3(2-R_3^2)(1-R_4^2)(1-R_6^2)\times\cos(\varphi_1+\varphi_2-\varphi_3)$$
$$-R_1R_4R_5(2-R_1^2)(2-R_4^2)(2-R_5^2)\times\cos(\varphi_1+\varphi_4-\varphi_5)$$
$$-2R_1R_2R_3(2-R_2^2)(1-R_5^2)(1-R_6^2)\times\cos(\varphi_1+\varphi_2-\varphi_3)$$
$$-2R_1R_2R_3(2-R_1^2)(1-R_4^2)(1-R_5^2)\times\cos(\varphi_1+\varphi_2-\varphi_3)$$
$$+2R_1R_2R_3(R_4^2-1)(R_5^2-1)(R_6^2-1)\times\cos(\varphi_1+\varphi_2-\varphi_3)]$$
$$-\frac{1}{2N\sqrt{N}}R_1R_2R_3[(R_1^4-6R_1^2+6)+(R_2^4-6R_2^2+6)+(R_3^4-6R_3^2+6)+(R_4^4-4R_4^2+2)+(R_5^4-4R_5^2+2)+(R_6^4-4R_6^2+2)]\cos(\varphi_1+\varphi_2-\varphi_3)+\ldots$$
$$+\frac{2}{N\sqrt{N}}R_1R_2R_3[(1-R_4^2)(1-R_5^2)+(1-R_4^2)(1-R_6^2)+(1-R_5^2)(1-R_6^2)]\cos(\varphi_1+\varphi_2-\varphi_3)+\ldots\}. \quad (1)$$

Several terms not essential for our purpose are omitted from (1).

The conditional expected value of $\cos(\varphi_{\mathbf{H}_1}+\varphi_{\mathbf{H}_2}-\varphi_{\mathbf{H}_1+\mathbf{H}_2})$, given $R_{\mathbf{H}_1}, R_{\mathbf{H}_2}, R_{\mathbf{H}_1+\mathbf{H}_2}, R_{\mathbf{K}}, R_{\mathbf{H}_1+\mathbf{K}}, R_{\mathbf{H}_1+\mathbf{H}_2+\mathbf{K}}$

Let us define $\varphi=\varphi_1+\varphi_2-\varphi_3$. Of considerable importance is the expected value of $\cos\varphi$ when the magnitudes of the structure factors are known but the phases are not. From (1)

$$P(\varphi|R_1,\ldots,R_6)=\frac{1}{2\pi}+\frac{1}{2\pi(1+q/N)}\frac{2R_1R_2R_3}{\sqrt{N}}\cos\varphi\{1+\frac{1}{N}[\tfrac{1}{8}(R_1^2+R_2^2+R_3^2)-3+\tfrac{1}{2}(R_1^2-2)(R_2^2-2)(R_3^2-2)+(R_4^2-1)(R_5^2-1)(R_6^2-1)+(R_1^2-1)(R_4^2-1)(R_5^2-1)+(R_2^2-1)(R_5^2-1)(R_6^2-1)+(R_3^2-1)(R_4^2-1)(R_6^2-1)-\tfrac{1}{4}(R_1^4-4R_1^2+2)-\ldots-\tfrac{1}{4}(R_6^4-4R_6^2+2)]\}$$
$$+\frac{1}{2\pi(1+q/N)}\left(\frac{1}{N}R_1^2R_2^2R_3^2\cos 2\varphi+\frac{1}{3N\sqrt{N}}R_1^3R_2^3R_3^3\cos 3\varphi\right), \quad (2)$$

where

$$q=\frac{1}{N}[(R_1^2-1)(R_2^2-1)(R_3^2-1)+(R_1^2-1)(R_4^2-1)(R_5^2-1)+(R_2^2-1)(R_5^2-1)(R_6^2-1)+(R_3^2-1)(R_4^2-1)(R_6^2-1)]-\frac{1}{4N}[(R_1^4-4R_1^2+2)+\ldots+(R_6^4-4R_6^2+2)].$$

Then

$$\langle\cos\varphi|R_1,\ldots,R_6\rangle\simeq\frac{R_1R_2R_3}{\sqrt{N}}\left\{1+\frac{[p+(R_4^2-1)(R_5^2-1)(R_6^2-1)]/N}{1+q/N}\right\}, \quad (3)$$

where

$$p=\tfrac{1}{8}(R_1^2+R_2^2+R_3^2)+\tfrac{1}{2}(R_1^2-2)(R_2^2-2)(R_3^2-2)-(R_1^2-1)(R_2^2-1)(R_3^2-1)-3.$$

The conditional expected value of $\cos(\varphi_{\mathbf{H}_1}+\varphi_{\mathbf{H}_2}-\varphi_{\mathbf{H}_1+\mathbf{H}_2})$, given $R_{\mathbf{H}_1}, R_{\mathbf{H}_2}, R_{\mathbf{H}_1+\mathbf{H}_2}, R_{\mathbf{K}_1}, R_{\mathbf{H}_1+\mathbf{K}_1}, R_{\mathbf{H}_1+\mathbf{H}_2+\mathbf{K}_1}, R_{\mathbf{K}_2}, R_{\mathbf{H}_1+\mathbf{K}_2},\ldots$

For a given phase $\varphi_{\mathbf{H}_1}+\varphi_{\mathbf{H}_2}-\varphi_{\mathbf{H}_1+\mathbf{H}_2}$ several triples $(R_{\mathbf{K}}, R_{\mathbf{H}_1+\mathbf{K}}, R_{\mathbf{H}_1+\mathbf{H}_2+\mathbf{K}})$ in general exist in the set of measured reflexions. The more complex distribution function

$$P(\varphi, R_{\mathbf{H}_1},\ldots,R_{\mathbf{K}_1}, R_{\mathbf{H}_1+\mathbf{K}_1}, R_{\mathbf{H}_1+\mathbf{H}_2+\mathbf{K}_1}, R_{\mathbf{K}_2}, R_{\mathbf{H}_1+\mathbf{K}_2}, R_{\mathbf{H}_1+\mathbf{H}_2+\mathbf{K}_2},\ldots)$$

must then be explored in order to obtain a more accurate estimation of φ. By a procedure similar to that described in paper I we obtain

$$\langle\cos\varphi|\ldots\rangle\simeq\frac{R_1R_2R_3}{\sqrt{N}}\left(1+\frac{A/N}{1+B/N+C/4N}\right), \quad (4)$$

where

$$A=\sum_{\mathbf{K}}(R_{\mathbf{K}}^2-1)(R_{\mathbf{H}_1+\mathbf{K}}^2-1)(R_{\mathbf{H}_1+\mathbf{H}_2+\mathbf{K}}^2-1), \quad (5)$$

$$B=(R_{\mathbf{H}_1}^2-1)\sum_{\mathbf{K}}(R_{\mathbf{K}}^2-1)(R_{\mathbf{H}_1+\mathbf{K}}^2-1)+(R_{\mathbf{H}_2}^2-1)\sum_{\mathbf{K}}(R_{\mathbf{H}_1+\mathbf{K}}^2-1)(R_{\mathbf{H}_1+\mathbf{H}_2+\mathbf{K}}^2-1)+(R_{\mathbf{H}_1+\mathbf{H}_2}^2-1)\sum_{\mathbf{K}}(R_{\mathbf{K}}^2-1)(R_{\mathbf{H}_1+\mathbf{H}_2+\mathbf{K}}^2-1), \quad (6)$$

$$C=-\sum_{\mathbf{K}}(R_{\mathbf{K}}^4-4R_{\mathbf{K}}^2+2)+(R_{\mathbf{H}_1+\mathbf{K}}^4-4R_{\mathbf{H}_1+\mathbf{K}}^2+2)+(R_{\mathbf{H}_1+\mathbf{H}_2+\mathbf{K}}^4-4R_{\mathbf{H}_1+\mathbf{H}_2+\mathbf{K}}^2+2)]. \quad (7)$$

The variance is given by

$$V=\langle\cos^2\varphi|\ldots\rangle-\langle\cos\varphi|\ldots\rangle^2\simeq\tfrac{1}{2}+\left[\frac{1}{4(1+B/N+C/4N)}-\left(1+\frac{A/N}{1+B/N+C/4N}\right)^2\right]\frac{R_1^2R_2^2R_3^2}{N}. \quad (8)$$

It should be noted that the term p which occurs in (3) does not appear in (4) and (8). In fact, if N is large enough and $\mathbf{K}$ is allowed to vary over a sufficiently large number of reciprocal vectors,

$$p \ll 1, \quad \text{or (and)}$$
$$p \ll A.$$

(3), (4) and (8) are in accordance with the formula

$$P(\varphi) \simeq \frac{1}{2\pi}\left(1+2\sum_{1n}^{\infty}\langle \cos n\varphi\rangle \cos n\varphi\right)$$

obtained by Naya, Nitta & Oda [1965, equation (10)].

When a large number of triples $(R_{\mathbf{K}}, R_{\mathbf{H}_1+\mathbf{K}}, R_{\mathbf{H}_1+\mathbf{H}_2+\mathbf{K}})$ is explored the expected value of $\cos\varphi$ may strongly differ from the value deduced when only the terms $R_{\mathbf{H}_1}$, $R_{\mathbf{H}_2}$, $R_{\mathbf{H}_1+\mathbf{H}_2}$ are known. In fact, according to whether

$$S=1+\frac{A/N}{1+B/N+C/4N} \gtreqless 0$$

the expected value of the cosine will assume positive or negative values. (8) tells us that the reliability of the cosine values provided by (4) increases with S.

Expected value of cos ($\varphi_{\mathbf{H}_1}+\varphi_{\mathbf{H}_2}-\varphi_{\mathbf{H}_1+\mathbf{H}_2}$) when the exponential form of the probability density is used

When only $R_{\mathbf{H}_1}$, $R_{\mathbf{H}_2}$, $R_{\mathbf{H}_1+\mathbf{H}_2}$ are known, (4) reduces to

$$\langle\cos\varphi\rangle \simeq \frac{R_{\mathbf{H}_1}R_{\mathbf{H}_2}R_{\mathbf{H}_1+\mathbf{H}_2}}{\sqrt{N}}.$$

Under this condition it has been shown (Cochran, 1955) that φ obeys the von Mises (1918) probability distribution of density function

$$P(\varphi)=\frac{1}{2\pi I_0(G)}\exp(G\cos\varphi), \tag{9}$$

with parameter

$$G=\frac{2R_{\mathbf{H}_1}R_{\mathbf{H}_2}R_{\mathbf{H}_1+\mathbf{H}_2}}{\sqrt{N}}.$$

From (9)

$$\langle\cos\varphi\rangle=\frac{I_1(G)}{I_0(G)}. \tag{10}$$

A result formally similar to (9) may be achieved by transforming our series expansion of the joint probability distribution to an exponential form, as from the application of the central-limit theorem (Bertaut, 1960*a,b*). In particular, the transformation assures the general positivity both of the probability density and of the variance values. So, when one or more triples $(R_{\mathbf{K}}, R_{\mathbf{H}_1+\mathbf{K}}, R_{\mathbf{H}_1+\mathbf{H}_2+\mathbf{K}})$ are known, (10) should be still valid, but

$$G=\frac{2R_{\mathbf{H}_1}R_{\mathbf{H}_2}R_{\mathbf{H}_1+\mathbf{H}_2}}{\sqrt{N}}\left(1+\frac{A/N}{1+B/N+C/4N}\right). \tag{11}$$

The estimation of cos ($\varphi_{2\mathbf{H}}-2\varphi_{\mathbf{H}}$)

A reliable estimate of $\cos(\varphi_{2\mathbf{H}}-2\varphi_{\mathbf{H}})$ may be particularly useful in the first stages of the phase-determination process. In fact, if $\varphi_{\mathbf{H}}$ is known, the value of $\varphi_{2\mathbf{H}}$ may under favourable circumstances be determined by the value of $\langle\cos(\varphi_{2\mathbf{H}}-2\varphi_{\mathbf{H}})\rangle$. This expected value and its variance may be derived from the study of the distribution

$$P(R_{\mathbf{H}}, R_{2\mathbf{H}}, R_{\mathbf{K}}, R_{\mathbf{H}+\mathbf{K}}, R_{2\mathbf{H}+\mathbf{K}}, \varphi_{\mathbf{H}}, \ldots, \varphi_{2\mathbf{H}+\mathbf{K}}).$$

If the series expansion of the distribution function is transformed to an exponential form, we obtain

$$\langle\cos(\varphi_{2\mathbf{H}}-2\varphi_{\mathbf{H}})\rangle \simeq \frac{I_1(G)}{I_0(G)},$$

$$\operatorname{var}[\cos(\varphi_{2\mathbf{H}}-2\varphi_{\mathbf{H}})] \simeq 1-\frac{I_1(G)}{GI_0(G)}-\frac{I_1^2(G)}{I_0^2(G)}, \tag{12}$$

where

$$G\simeq\frac{R_{\mathbf{H}}^2R_{2\mathbf{H}}}{\sqrt{N}}\left(1+\frac{2A/N}{1+B/N+C/4N}\right),$$

$$A=\sum_{\mathbf{K}}(R_{\mathbf{K}}^2-1)(R_{\mathbf{H}+\mathbf{K}}^2-1)(R_{2\mathbf{H}+\mathbf{K}}^2-1),$$

$$B=(R_{\mathbf{H}}^2-1)\sum_{\mathbf{K}}[(R_{\mathbf{K}}^2-1)(R_{\mathbf{H}+\mathbf{K}}^2-1)+(R_{\mathbf{H}+\mathbf{K}}^2-1)(R_{2\mathbf{H}+\mathbf{K}}^2-1)]+(R_{2\mathbf{H}}^2-1)\sum_{\mathbf{K}}(R_{\mathbf{K}}^2-1)(R_{2\mathbf{H}+\mathbf{K}}^2-1),$$

$$C\simeq-[\sum_{\mathbf{K}}(R_{\mathbf{K}}^4-4R_{\mathbf{K}}^2+2)+(R_{\mathbf{H}+\mathbf{K}}^4-4R_{\mathbf{H}+\mathbf{K}}^2+2)+(R_{2\mathbf{H}+\mathbf{K}}^4-4R_{2\mathbf{H}+\mathbf{K}}^2+2)].$$

When no triple $(R_{\mathbf{K}}, R_{\mathbf{H}+\mathbf{K}}, R_{2\mathbf{H}+\mathbf{K}})$ is known, G reduces to $R_{\mathbf{H}}^2R_{2\mathbf{H}}/\sqrt{N}$: unlike the $P\bar{1}$ case, $\langle\cos(\varphi_{2\mathbf{H}}-2\varphi_{\mathbf{H}})\rangle$ is always positive, whatever $R_{\mathbf{H}}$ may be. If triples $(R_{\mathbf{K}}, R_{\mathbf{H}+\mathbf{K}}, R_{2\mathbf{H}+\mathbf{K}})$ are tested, (12) is in principle able to give expected negative values of the cosine.

The role of A, B and C

In order to obtain some insight into the role played by A, B and C in (4) or (10) we derive in this paragraph their conditional expected values. From the distribution

$$P(R_{\mathbf{K}}, R_{\mathbf{H}_1+\mathbf{K}}, R_{\mathbf{H}_1+\mathbf{H}_2+\mathbf{K}}, \ldots | R_{\mathbf{H}_1}, R_{\mathbf{H}_2}, R_{\mathbf{H}_1+\mathbf{H}_2}, \varphi)$$

we obtain

$$\langle A\rangle \simeq \frac{1}{Q}\,\frac{2}{N\sqrt{N}}R_{\mathbf{H}_1}R_{\mathbf{H}_2}R_{\mathbf{H}_1+\mathbf{H}_2}\cos\varphi, \tag{13}$$

$$\langle B\rangle \simeq \frac{1}{8NQ}\left(1+\frac{2}{\sqrt{N}}R_{\mathbf{H}_1}R_{\mathbf{H}_2}R_{\mathbf{H}_1+\mathbf{H}_2}\cos\varphi\right)\times[(R_{\mathbf{H}_1}^2-1)^2+(R_{\mathbf{H}_2}^2-1)^2+(R_{\mathbf{H}_1+\mathbf{H}_2}^2-1)^2], \tag{14}$$

$$C=+\frac{1}{8NQ}\,\frac{9}{8}\left(1+\frac{2}{N}R_{\mathbf{H}_1}R_{\mathbf{H}_2}R_{\mathbf{H}_1+\mathbf{H}_2}\cos\varphi\right). \tag{15}$$

(13) tells us that the expected sign of A coincides with the sign of the triplet cosine. Thus the negative cosine invariants may be in principle singled out. However,

positive values of A should be expected when the triplet is a 'bad' one but has a small positive cosine. In this case the expected cosine value as calculated after the exploration of the terms $(R_{\mathbf{K}}, R_{\mathbf{H}_1+\mathbf{K}}, R_{\mathbf{H}_1+\mathbf{H}_2+\mathbf{K}})$ will probably be larger than that calculated *via* the sole factor $R_{\mathbf{H}_1}R_{\mathbf{H}_2}R_{\mathbf{H}_1+\mathbf{H}_2}/\sqrt{N}$. This behaviour is mainly due to the fact that our probability density function of the triplet cosine has its maximum always in 0 or π.

From (14) and (15) one derives that, if N is large enough, positive values of B and C will be expected whatever the sign of the cosine may be. The absolute values of B and C increase on average with $\cos\varphi$. For the triplets commonly used in the procedures for crystal structure solution $\langle C\rangle$ seems negligible on average in comparison with B, so that its calculation may be omitted. Furthermore one may expect values of B (which is a term of order $1/N$) larger than the corresponding value of A (which is a term of order $1/N\sqrt{N}$).

It is anticipated that an overall rescaling of A and B terms is in general advisable in the practical procedures for the crystal structure solution. These aspects will be described in subsequent papers.

The 'exact' value of the cosine invariant: a comparison with preceding approaches

A form of $B_{3,0}$ formula valid in centrosymmetric crystals [equation 2.1.3 of Hauptman & Karle (1958)] and in non-centrosymmetric crystals [equation 2.1.3 of Karle & Hauptman (1958)] is

$$|R_{\mathbf{H}_1}R_{\mathbf{H}_2}R_{\mathbf{H}_1+\mathbf{H}_2}|\cos(\varphi_{\mathbf{H}_1}+\varphi_{\mathbf{H}_2}-\varphi_{\mathbf{H}_1+\mathbf{H}_2}) \simeq Y\langle(R_{\mathbf{K}}^2-1)(R_{\mathbf{H}_1+\mathbf{K}}^2-1)(R_{\mathbf{H}_1+\mathbf{H}_2+\mathbf{K}}^2-1)\rangle_{\mathbf{K}}+L \quad (16)$$

where L is a negligible correction term, $Y=N^{3/2}/8$ and $N^{3/2}/2$ for centro- and non-centrosymmetric crystals respectively. The constant Y as given in the original theory turned out not to be suitable for application to real structures. Different procedures for obtaining the rescaling of Y have been presented.

From the result that the number of induced or chance interactions causes the average values of $(R_{\mathbf{K}}^2-1)^2$ and $(R_{\mathbf{K}}^2-1)^3$ to increase, Hauptman (1964) proposed

$$Y=\frac{(N-1)(N-2)}{\sqrt{N}\langle(R_{\mathbf{K}}^2-1)^3\rangle_{\mathbf{K}}}.$$

Later, Hauptman, Fisher, Hancock & Norton (1969) presented a procedure in which the parameter Y was determined for each group of triplets for which the values of $R_{\mathbf{H}_1}$, $R_{\mathbf{H}_2}$, $R_{\mathbf{H}_1+\mathbf{H}_2}$ were essentially constant. Y is such that the resulting frequency distribution of $\cos(\varphi_{\mathbf{H}_1}+\varphi_{\mathbf{H}_2}-\varphi_{\mathbf{H}_1+\mathbf{H}_2})$ agrees, for fixed $R_{\mathbf{H}_1}R_{\mathbf{H}_2}R_{\mathbf{H}_1+\mathbf{H}_2}$, with the known theoretical conditional probability distribution.

Karle (1970) presented a modification of (16) which eliminates the constant Y, substituting instead a ratio of averages over the set of triples $(R_{\mathbf{K}}, R_{\mathbf{H}_1+\mathbf{K}}, R_{\mathbf{H}_1+\mathbf{K}_2+\mathbf{K}})$ really used in the procedure.

A further modification is that proposed by Hauptman (1970, 1972):

$$\cos(\varphi_{\mathbf{H}_1}+\varphi_{\mathbf{H}_2}-\varphi_{\mathbf{H}_1+\mathbf{H}_2}) \simeq \frac{N^{3/2}}{2R_{\mathbf{H}_1}R_{\mathbf{H}_2}R_{\mathbf{H}_1+\mathbf{H}_2}(R_{\mathbf{K}}^2-1)(R_{\mathbf{H}_1+\mathbf{K}}^2-1)} \times[\langle(R_{\mathbf{H}_1+\mathbf{H}_2+\mathbf{K}}^2-1)|R_{\mathbf{K}},R_{\mathbf{H}_1+\mathbf{K}}\rangle-\langle(R_{\mathbf{H}_1+\mathbf{H}_2+\mathbf{K}}^2-1)|R_{\mathbf{K}}\rangle - \langle(R_{\mathbf{H}_1+\mathbf{H}_2+\mathbf{K}}^2-1)|R_{\mathbf{H}_1+\mathbf{K}}\rangle]. \quad (17)$$

All the above-mentioned formulae which modify (16) readily permit the use of a limited subset of the experimental data in obtaining values of $\cos\varphi$. In this connexion closely related to (17) are the $(D-S)/S$ formula (Hauptman, 1970) and $M(D-KS)$ formula (Fisher, Hancock & Hauptman, 1970).

Because of its generality our approach is able to provide the 'exact' value of the triplet cosine invariant. By a combination of suitable conditional distribution functions we obtain

$$\cos(\varphi_{\mathbf{H}_1}+\varphi_{\mathbf{H}_2}-\varphi_{\mathbf{H}_1+\mathbf{H}_2}) \simeq \frac{N^{3/2}}{2R_{\mathbf{H}_1}R_{\mathbf{H}_2}R_{\mathbf{H}_1+\mathbf{H}_2}(R_{\mathbf{K}}^2-1)(R_{\mathbf{H}_1+\mathbf{K}}^2-1)} \times\{\langle(R_{\mathbf{H}_1+\mathbf{H}_2+\mathbf{K}}^2-1)|R_{\mathbf{K}},R_{\mathbf{H}_1+\mathbf{K}}\rangle \times[1+(R_{\mathbf{H}_1}^2-1)(R_{\mathbf{K}}^2-1)(R_{\mathbf{H}_1+\mathbf{K}}^2-1)/N^*] - \langle R_{\mathbf{H}_1+\mathbf{H}_2-\mathbf{K}}^2-1|R_{\mathbf{K}}\rangle-\langle(R_{\mathbf{H}_1+\mathbf{H}_2+\mathbf{K}}^2-1)|R_{\mathbf{H}_1+\mathbf{K}}\rangle\}. \quad (18)$$

We have marked with an asterisk a term in (18) not given by (17).

Conclusions

In contrast to previous distributions in which only the single vector $\mathbf{K}$ had been assumed to be the primitive random variable, in the present paper the joint probability distribution

$$P(E_{\mathbf{H}_1},E_{\mathbf{H}_2},E_{\mathbf{H}_1+\mathbf{H}_2},E_{\mathbf{K}_1},E_{\mathbf{H}_1+\mathbf{K}_1}, E_{\mathbf{H}_1+\mathbf{H}_2+\mathbf{K}_1},E_{\mathbf{K}_2},E_{\mathbf{H}_1+\mathbf{K}_2},\ldots)$$

had been derived in $P1$ on the basis that the atomic positional parameters are the primitive random variables. The approach enables us to calculate the expected value of the cosine invariants in addition to their 'exact' values, thus allowing their immediate use in the tangent procedures.

This work was supported by the Consiglio Nazionale delle Ricerche (grant No. 75.00034.05.115.4593).

References

BERTAUT, E. F. (1960*a*). *Acta Cryst.* **13**, 546–552.
BERTAUT, E. F. (1960*b*). *Acta Cryst.* **13**, 643–649.
COCHRAN, W. (1955). *Acta Cryst.* **8**, 473–478.
FISHER, J., HANCOCK, H. & HAUPTMAN, H. (1970). Naval Res. Lab. Rep. No. 7132, Washington, D.C.
GIACOVAZZO, C. (1976*a*). *Acta Cryst.* A**32**, 967–976.
GIACOVAZZO, C. (1976*b*). *Acta Cryst.* A**32**, 91–99.

HAUPTMAN, H. (1964). *Acta Cryst.* **17**, 1421–1433.
HAUPTMAN, H. (1970). Abstract B8, New Orleans Meeting of the Amer. Cryst. Assoc., March 2.
HAUPTMAN, H. (1972). *Z. Kristallogr.* **135**, 1–17.
HAUPTMAN, H., FISHER, J., HANCOCK, H. & NORTON, D. A. (1969). *Acta Cryst.* B**25**, 811–814.
HAUPTMAN, H. & KARLE, J. (1958). *Acta Cryst.* **11**, 149–157.
KARLE, J. (1970). *Acta Cryst.* B**26**, 1614–1617.
KARLE, J. & HAUPTMAN, H. (1958). *Acta Cryst.* **11**, 264–269.
MISES, R. VON (1918). *Phys. Z.* **19**, 490–500.
NAYA, S., NITTA, I. & ODA, T. (1965). *Acta Cryst.* **19**, 734–747.

Acta Cryst. (1977). A**33**, 531–538

A Probabilistic Theory of the Coincidence Method. I. Centrosymmetric Space Groups

BY C. GIACOVAZZO

Institute of Mineralogy and Petrography, University of Bari, Bari, Italy

(*Received* 28 *September* 1976; *accepted* 17 *December* 1976)

A general probabilistic theory of sign coincidences is described which is valid in all the centrosymmetric space groups. The theory makes full use of space-group symmetry by means of suitable joint probability distribution functions. The phase relationships obtained are in some way space-group dependent: their estimation requires an appropriate use of space-group algebra.

1. Introduction

The method of coincidences was described by Grant, Howells & Rogers (1957) for the application of the Sayre relation $S(\mathbf{H}_1)S(\mathbf{H}_2)\simeq S(\mathbf{H}_1+\mathbf{H}_2)$ to projections with symmetry *pgg*, *pmg*, and *p*4*g*. In their terminology, if **A** and **B** are reciprocal vectors, the product $S(\mathbf{A})\times S(\mathbf{B})$ can enter into a relation with several different third terms $S(\mathbf{C}_1)$, $S(\mathbf{C}_2)$, *etc.*, whose signs are said to coincide:

$$[S(\mathbf{C}_1)\simeq S(\mathbf{C}_2)\ldots]_{A,B}.$$

A particular coincidence case is obtained for centrosymmetric space groups when $|E_{\mathbf{H}_1}|$, $|E_{\mathbf{H}_2}|$, $|E_{\mathbf{H}_1+\mathbf{H}_2}|$, $|E_{\mathbf{H}_1-\mathbf{H}_2}|$ are sufficiently large: in fact

$$S(\mathbf{H}_1)S(\mathbf{H}_2)\simeq S(\mathbf{H}_1+\mathbf{H}_2), \quad (1a)$$

$$S(\mathbf{H}_1)S(\mathbf{H}_2)\simeq S(\mathbf{H}_1-\mathbf{H}_2), \quad (1b)$$

from which

$$S(\mathbf{H}_1+\mathbf{H}_2)\simeq S(\mathbf{H}_1-\mathbf{H}_2). \quad (2)$$

Debaerdemaeker & Woolfson (1972) have extended the idea of coincidence to non-centrosymmetric space groups in which there are translational elements of symmetry. According to them, from a $\sum_2$ listing one can deduce the existence of pairs of phase relations of the general form

$$\varphi_p+B_{prs}(\varphi_r+\varphi_s)+b_{prs}\simeq 0, \quad (3a)$$

$$\varphi_q+B_{qrs}(\varphi_r+\varphi_s)+b_{qrs}\simeq 0, \quad (3b)$$

where the *b*'s are constant angles arising because of the translational symmetry, and the *B*'s can be ± 1. By a combination of (3*a*) and (3*b*) one obtains a general relation between φ_p and φ_q:

$$\varphi_p\pm\varphi_q\pm(b_{prs}\pm b_{qrs})\simeq 0. \quad (4)$$

Debaerdemaeker & Woolfson (1972) supplied in their paper a probabilistic theory of the coincidence phase relations. In accordance with them, we refer to $P2_12_12_1$: let us consider the relations

$$\varphi_p\simeq\eta$$
$$\zeta_q=b\varphi_q-m\pi\simeq\eta,$$

where $b=\pm 1$ and $m=0{\cdot}1$. P_1 and P_2 are the probability distributions for φ_p and ζ_q, derivable from the Cochran (1955) theory:

$$P_{1,2}(\varphi)=[2\pi I_0(G)]^{-1}\exp[G\cos(\varphi-\eta)],$$

where (Karle & Karle, 1966) $G=2\sigma_3\sigma_2^{-3/2}|E_1E_2E_3|$.

According to Debaerdemaeker & Woolfson (1972), the strength of a coincidence phase relation depends on the probability distribution P of the quantity $\theta=\varphi_p-\zeta_q$:

$$P=\int_{-\pi}^{\pi}P_1(x)P_2(x+\theta)\mathrm{d}x. \quad (5)$$

In centrosymmetric space groups (5) is equivalent to the statement: if P_1 and P_2 are respectively the probabilities of the relations (1*a*) and (1*b*), the probability of (2) is given by (Woolfson, 1961)

$$P=P_1P_2+(1-P_1)(1-P_2). \quad (6)$$

Giacovazzo (1974*a*) showed that (6) is in contrast with the Harker–Kasper inequalities, and that results are misleading when $|E_{\mathbf{H}_1}|$ or $|E_{\mathbf{H}_2}|$ is small. We will show in part II of the present paper that (5) may also lead to wrong phase relations when $|E_{\mathbf{H}_1}|$ and (or) $|E_{\mathbf{H}_2}|$ are small. The phase information deduced from weak reflexions $\mathbf{H}_1$ and (or) $\mathbf{H}_2$ is particularly important in symmorphic space groups because it can lead to relations such as

$$\varphi_p-\zeta_q\simeq\pi. \quad (7)$$

Debaerdemaeker & Woolfson's (1972) approach im-