On the other hand, they can be chosen to have the form

$$\mathbf{v}_i=\tfrac{1}{4}\sum_{j\neq i}a_{ij}\mathbf{e}_j \quad \text{where} \quad \sum_{j\neq i}a_{ij}=0 \bmod 2\,.$$

A calculation shows that

$$\mathbf{t}_{ij}=\tfrac{1}{2}a_{ij}\mathbf{e}_j-\tfrac{1}{2}a_{ji}\mathbf{e}_i$$

and therefore $a_{ij}=a_{ji} \bmod 2$. The graph representing $\mathscr{G}$ (Fig. 3) is obtained by joining the ith vertex to the jth vertex by an edge if and only if a_{ij} is an odd integer.

Corollary: The number of such groups is bounded below by $2^{(n-1)(n-2)/2}/n!$ and for $n=2,3,4,5,6$ is 1,2,3,7,16 [for diagrams of the various graphs see Appendix 1 of Harary (1969)].

Theorem 4.2. The groups of arithmetic crystal class $Um\ldots m$ ($n-1$ factors) are in one–one correspondence with the (non-directed) graphs with $n-1$ vertices.

Proof: The above description shows that each group determines a graph with n vertices, one of which is distinguished. An even number of edges end at each vertex. Removing all edges which end at the distinguished vertex we obtain the required graph with $n-1$ vertices.

Corollary: The number of such groups is bounded below by $2^{(n-1)(n-2)/2}/(n-1)!$ and for $n=2,3,4,5,6$ is 1,2,4,11,34.

I wish to thank Dr L. L. Boyle and Dr M. S. Paterson for several helpful discussions.

References

BELOV, N. V. (1951). *Tr. Inst. Krist. Akad. Nauk SSSR* **6**, 25–62. Engl. Trans.: *Proc. Leeds Phil. Lit. Soc.* **8**, 1–46.

BELOV, N. V. & KLEVTSOVA, R. F. (1959). *Kristallografiya*, **4**, 289–292, 473–476.

BÜLOW, R., NEUBÜSER, J. & WONDRATSCHEK, H. (1971). *Acta Cryst.* A**27**, 517–535.

BURCKHARDT, J. J. (1966). *Die Bewegungsgruppen der Kristallographie.* 2nd ed. Basel, Stuttgart: Birkhäuser Verlag.

HARARY, F. (1969). *Graph Theory.* Reading, Mass.: Addison Wesley.

MAXWELL, G. (1975). *J. Algebra*, **35**, 159–177.

SCHWARZENBERGER, R. L. E. (1974). *Proc. Camb. Phil. Soc.* **76**, 23–32.

Acta Cryst. (1976). A**32**, 359

About Quartets — Relation with the Invariant Phases of Triplets

BY BERNARD BUSETTA

Laboratoire de Cristallographie et de Physique Cristalline associé au CNRS, Université de Bordeaux I, 351 *Cours de la Libération*, 33405 *Talence, France*

(*Received* 6 *October* 1975; *accepted* 9 *October* 1975)

On the basis that quartet invariants are the difference between triplet invariants, we have determined a theoretical distribution of the quartet invariant phases. New formulae to estimate triplet invariant cosines are described and the results they give for a known test structure are compared .

Recently, Hauptman (1974) gave an estimate of the invariant cosine of the sum of phases of four linear dependent reflexions l,m,n,p such that the sum over each set of three indices is zero ($l+m+n+p=0$). Schenk (1973) had already spoken of such invariants as quartets of the second kind, and showed that they are obtained by constructing the difference of the phases of two invariant triplets relative to the same reflexion H, *e.g.*

$$\varphi_{\bar{H}}+\varphi_K+\varphi_{H-K}=\alpha_{H,K}\,,$$

and

$$\varphi_{\bar{H}}+\varphi_L+\varphi_{H-L}=\alpha_{H,L}\,.$$

from which is derived the quartet $\varphi_K+\varphi_{H-K}+\varphi_{\bar{L}}+\varphi_{L-H}$ (the sum of indices of the four reflexions $K, H-K, \bar{L}$, $L-H$ is actually equal to zero) with a phase equal to $(\alpha_{H,K}-\alpha_{H,L})$.

The value of the invariant cosine $\cos(\varphi_K+\varphi_{H-K}+\varphi_{\bar{L}}+\varphi_{L-H})$ may be estimated from the moduli of seven structure factors $E_K, E_{H-K}, E_L, E_{L-H}$ and also E_H, E_{K-L} and E_{K+L-H}. Such an estimation is more accurate than the estimation of the phases $\alpha_{H,K}$ or $\alpha_{H,L}$ of the generating triplets each of which is computed from only three structure factors E_H, E_K and E_{H-K} or E_H, E_L and E_{H-L}.

Furthermore, different algebraic formulae have been described to compute the invariant cosine of the phase of a triplet from the moduli of the structure factors of the whole reciprocal space: (1) triple product formula (Hauptman, Fisher, Hancock & Norton, 1969) and (2) MDKS formula (Fisher, Hancock & Hauptman, 1970).

We intend, here, to compare results provided respectively by the estimation of quartet phases and by al-

gebraic computation of the invariant triplet phases. However, we first point out some theoretical implications which result from the dependence of the quartets on their generating triplets.

I. A conditional distribution function of invariant phases of quartets

The theoretical distribution of the phases x of invariant triplets corresponding to a given A value [$A=(2\sigma_3/\sigma_2^{3/2})|E_H E_K E_{H-K}|$] may be defined by its probability density

$$\varrho(x)=\frac{1}{2\pi I_0(A)}\exp(A\cos x)$$

in the range $(-\pi,\pi)$ and extended by translation as a periodic function of period 2π.

To integrate the exponential, it is generally easier to use a development in Bessel functions with imaginary argument

$$\exp(A\cos x)=I_0(A)+2\sum_{n=1}^{+\infty}I_n(A)\cos nx\,.$$

The distribution function $\psi(t)=\int_0^t\varrho(x)\mathrm{d}x$ is therefore an odd function. As we intend to use values of x varying from $-\pi$ to $+\pi$, we normalize $\psi(t)$ on a period of 2π instead of π as was done by Hauptman *et al.* (1969); then we have $\psi(\pi)=\frac{1}{2}$.

In this paragraph, we intend to compute the probability

$$\text{Prob}\,(|\alpha_2-\alpha_1|<a)=\text{Prob}\,[\cos(\alpha_2-\alpha_1)<\cos a]$$

where a is a constant such that $0\le a\le\pi$. α_1 and α_2 are the phases of two invariant triplets corresponding to A values equal to A_1 and A_2. If these two triplets are relative to the same reflexion H then the desired probability is the cumulative probability of the phase of invariant quartets built from this reflexion.

Let us call $P(\alpha_1)$ the conditional probability

$$P(\alpha_1)=\text{Prob}\,(|\alpha_2-\alpha_1|<a)\,,$$

when α_1 is fixed.

As $P(\alpha_1)$ and $\varrho(\alpha_1)$ are even functions, the final integration may be computed in the range $|0,\pi|$:

$$\text{Prob}\,(|\alpha_2-\alpha_1|<a)=2\int_0^\pi P(\alpha_1)\varrho(\alpha_1)\mathrm{d}\alpha_1\,.$$

From the particular way that $\varrho(x)$ is defined, it is necessary to compute $P(\alpha_1)$ in two separate ranges for α_1, $|0,\pi-\alpha|$ and $|\pi-a,\pi|$. With the previous notation, we have:

if $0<\alpha_1<\pi-a$

$$P(\alpha_1)=\text{Prob}\,(\alpha_1-a<\alpha_2<\alpha_1+a)=\psi(\alpha_1+a)-\psi(\alpha_1-a)$$

and if $\pi-a<\alpha_1<\pi$

$$P(\alpha_1)=[\tfrac{1}{2}-\psi(\alpha_1-a)]+[\tfrac{1}{2}-\psi(2\pi-\alpha_1-a)]\,.$$

After integration, we obtain in the two cases the same expression for $P(\alpha_1)$

$$P(\alpha_1)=\frac{1}{2\pi I_0(A_2)}\left\{2aI_0(A_2)+2\sum_{n=1}^{+\infty}\frac{I_n(A_2)}{n}[\sin n(\alpha_1+a)-\sin n(\alpha_1-a)]\right\}$$

which may be also written, from the addition formula of sine functions,

$$P(\alpha_1)=\frac{1}{2\pi I_0(A_2)}\left[2aI_0(A_2)+4\sum_{n=1}^{+\infty}\frac{I_n(A_2)}{n}\sin na\cos n\alpha_1\right].$$

To compute $1/[2\pi I_0(A_1)]\int_0^\pi P(\alpha_1)\exp(A_1\cos\alpha_1)\mathrm{d}\alpha_1$, we have to integrate terms of the form

$$\int_0^\pi\exp(A_1\cos\alpha_1)\cos n\alpha_1\mathrm{d}\alpha_1=\int_0^\pi\left[I_0(A_1)+2\sum_{n'=1}^{\infty}\frac{I_{n'}(A_1)}{n'}\cos n'\alpha_1\right]\cos n\alpha_1\mathrm{d}\alpha_1\,.$$

If $n\neq n'$ the corresponding integral is zero and if $n=n'$ it is equal to $\pi/2$.

Finally, we have

$$\text{Prob}\,[|\alpha_2-\alpha_1|<a]=\frac{a}{\pi}+\frac{2}{\pi I_0(A_2)I_0(A_1)}\sum_{n=1}^{+\infty}\frac{I_n(A_2)I_n(A_1)}{n^2}\sin na\,.$$

II. The estimation of the phase of triplet invariants from the corresponding estimation for quartets

Now we consider a particular triplet invariant relative to the reflexion H. Its phase is $\alpha_{H,K}=\alpha_0$. We intend to compute the mean value m of the cosine of the difference $\alpha-\alpha_0$ between the considered triplet and all the other triplets relative to the reflexion H for a given value of A. We have

$$m=\frac{1}{\pi I_0(A)}\int_0^\pi\cos(\alpha-\alpha_0)\exp(A\cos\alpha)\mathrm{d}\alpha\,.$$

It is always possible to change the variable of integration; we may suppose also that α_0 varies and α is fixed. Then, the constant A may be considered as the corresponding A value of the initially considered triplet.

After integration we have

$$m=\cos\alpha_0\,.\,\frac{I_1(A)}{I_0(A)}\,.$$

But m is also the mean value of the cosine of different quartet invariants built from a given triplet and all the other triplets relative to the same reflexion H.

Then from the different estimations $\cos(\alpha_1-\alpha_0)$ of quartet invariants, we estimate the cosine of triplet invariants by

$$\cos\alpha_0=\langle\cos(\alpha_i-\alpha_0)\rangle_i\cdot\frac{I_0(A)}{I_1(A)}\ . \qquad (1)$$

As the number of quartets used to compute each triplet cosine invariant is always small (200 to 400), it is important to know if the estimated cosine $\cos(\alpha_i-\alpha_0)$ corresponds to a positive or a negative $\cos\alpha_i$. Indeed, if equation (1) involves only α_i triplet phase invariants near to π, a negative cosine $\cos\alpha_0$ would be assumed spuriously as positive. To avoid this drawback, we use a two-stage estimation of the triplet cosine invariants: in the first stage, temporary cosine invariants $\cos\alpha_i'$ are computed from (1); in the second, the final invariant cosines are computed from

$$\cos\alpha_0=\langle\cos\alpha_i'\cos(\alpha_i-\alpha_0)\rangle_i\cdot\frac{I_0(A)}{I_1(A)}\ . \qquad (1')$$

If such a computation is not sufficiently accurate to provide the correct values of triplet cosine invariants, a rough sorting between the different cosine invariants may be achieved as in the case of the modified triple product formula. Then a correct estimate of these cosine invariants is obtained from the theoretical distribution of triplet phases (Busetta & Comberton, 1974).

III. Fast computation of the MDKS formula

Equation (1) is now compared with the classical modified triple product formula (Hauptman *et al.*, 1969)

$$|E_HE_KE_{H-K}|\cos\alpha_0=\langle\varepsilon_L\varepsilon_{L-(H-K)}\varepsilon_{L-H}\rangle_L\ ,$$

where

$$\varepsilon_L=\sqrt{|E_L|}-\langle\sqrt{|E_L|}\rangle_L$$

and which may be also written (Busetta & Comberton, 1974)

$$|E_HE_KE_{H-K}|\cos\alpha_0=\langle\varepsilon_L\varepsilon_{L-H}(\varepsilon_{L-K}+\varepsilon_{L-(H-K)})\rangle_{L>H/2}\ .$$

In (1) quartets are constructed only from the strongest reflexions ($E>E_0$), and, among the seven structure factors used to estimate each quartet cosine invariant, only the terms $|E_{L-K}|$ and $|E_{L-(H-K)}|$ may be less than E_0. From these considerations, (1) appeared strongly related to the D term of the MDKS formula (Fisher *et al.*, 1970)

$$D=\langle\varepsilon_{L-I}\mid|E_L|>E_0\ ,\quad|E_{L+J}|>E_0\rangle_L\ , \qquad (2)$$

where $I=\bar{H},K,H-K$ when $J=K,H-K,\bar{H}$ respectively.

It is possible to verify that (1) and (2) involve exactly the same structure factors; the only difference between them is the mathematical function used to represent the quartet cosine invariant; in (1) it corresponds to the exact mathematical estimation given by Hauptman (1974), that is to say a series involving imaginary Bessel functions; in (2) it corresponds to $(\varepsilon_{L-K}+\varepsilon_{L-(H-K)})$.

Both representations have roughly the same behaviour; they give negative values if the two structure factors involved are weak and positive values if they are large.

In addition the consideration of quartet invariants which are differences between triplet invariants involving the same reflexion, provides a way to compute the D term of the MDKS formula by a fast process which was not obvious at the beginning. As soon as the $\sum_2$ relations involving the strongest E factors are obtained, the triplets related to the same reflexions H are combined two by two, for instance, the H, K triplet with the H,L triplet; the ε_{L-K} and $\varepsilon_{L-(H-K)}$ terms are directly available, and may be added simultaneously to the corresponding terms $D_{H,K}$ and $D_{H,L}$.

As the same operation is done with triplet invariants involving the reflexions K and $H-K$, the $D_{H,K}$ term will be computed on average for a number of ε_{L-I} contributors equal to

$$2(n_H+n_K+n_{H-K}-3)\ ,$$

where n_H,n_K,n_{H-K} are the number of $\sum_2$ relations involving respectively the three reflexions $H,K,H-K$.

Then the mean number of contributors used in the estimation of the triplet cosine invariants may be estimated from the mean number of $\sum_2$ relations involving the same reflexion.

In practice a good estimate of the triplet cosine invariant is obtained as soon as the number of considered $\sum_2$ relations is five times the number of reflexions used. In this way, the estimation of triplet cosine invariants is less time consuming than the search of triplets themselves (that is to say the construction of the set of $\sum_2$ relations) and thus may be used in crystal structure determination, even for large molecules.

Finally one notices that, as previously, it may be better to use a two-stage estimation of the triplet cosine invariant, where the final D term would be obtained from

$$D_{H,K}=\langle\varepsilon_{L-I}D'_{L,J}\mid|E_L|>E_0,|E_{L+J}|>E_0\rangle_L\ , \qquad (2')$$

where $D'_{L,J}$ represents a first computation of $D_{L,J}$ from (2).

IV. Experimental results

The triplet cosine invariants were computed for an orthorhombic structure ($Pna2_1$, $C_{20}Si_2H_{28}$, $Z=4$) by the classical modified triple product formula, from the estimation of quartet cosine invariants (1'), by the semi-empirical equation (2'), and then compared them with their real values.

For this purpose, the triplet invariants are arranged in groups of 100 elements in which the value of $A_{H,K}$ may be considered as constant. Then they are sorted

Table 1. *Comparison of the* 1500 *triplet cosine invariants of the test structure*

Theoretical cosine invariant $\frac{I_1(A)}{I_0(A)}$	Experimental mean cosine invariant	Mean cosine invariant of the 50 strongest cosines in each group: Estimated triple product formula	Equation (1′)	Equation (2′)
0·8027	0·8673	0·8975	0·9061	0·9066
0·7599	0·8122	0·8947	0·8748	0·8821
0·7350	0·7594	0·8417	0·8244	0·8511
0·7148	0·8428	0·9050	0·8811	0·8984
0·6975	0·7310	0·7575	0·7667	0·7771
0·6842	0·8077	0·8588	0·9121	0·8888
0·6692	0·7922	0·8897	0·8903	0·8710
0·6561	0·8093	0·8994	0·9159	0·9213
0·6435	0·8260	0·9141	0·8749	0·8770
0·6321	0·7547	0·8895	0·8775	0·8858
0·6212	0·7880	0·8118	0·8678	0·8313
0·6102	0·7948	0·8939	0·8044	0·8167
0·6011	0·7541	0·8946	0·8029	0·8293
0·5891	0·7391	0·7616	0·7889	0·8159
0·5789	0·6488	0·6607	0·7280	0·7912

in decreasing order of the estimated cosines. The ability of the different formulae to sort the triplet cosine invariants may be estimated by comparing, in each group, the mean cosine invariant of the 50 first sorted triplets to the mean cosine invariant in the whole group. The results, obtained with our test structure, are given in Table 1.

The results provided by the three formulae appear very similar, though (1′) gives slightly less good results than the other formulae. They allow the determination of invariant cosines near +1·0 with good accuracy and may be used equally to select invariants for which the relation $\varphi_{\bar{H}}+\varphi_K+\varphi_{H-K}=0$ may be considered as nearly realized.

The invariant cosines computed between 0 and −1 are less accurate, because some positive invariant cosines are computed spuriously as negative. Since for these latter a wrong negative estimation is obtained for all the quartets constructed with the other triplet invariants, it seems that these problems depend on peculiarities of the structure, which cannot be allowed for in a general formula.

As the number of contributors in (1′) or (2′) is small, we wanted to see if the observed errors depend on specific reflexions. For that purpose, when a reflexion interfered in (1′) or (2′), we summarized the involved error observed in the estimated triplet cosine invariant and then determined, for each reflexion, a mean involved error. This test showed us that no particular reflexion involved significant errors, and that involved errors are independent of the resolving power of the considered reflexions. This is important if we intend to estimate triplet cosine invariants for large molecules for which high resolution cannot be assumed; it indicated that the accuracy of the estimation will depend only on the number of contributors used in (1′) or (2′). Then a good estimation may be expected if the conditions set up at the end of § III are obeyed.

Equation (2′) has been used successfully to estimate the triplet cosine invariants in the determination of luteoskyrin ($P2_12_12$, $Z=4$) $C_{30}O_{12}H_{22}.2(\frac{1}{2}CH_3\text{–}CO\text{–}CH_3)$ and nogalamycin ($P2_12_12_1$, $Z=8$) $2\times(C_{39}NO_{16}H_{48}.CH_3OH)$, the crystal structures of which will be published later.

The author thanks Professor H. Hauptman for valuable discussions and his critical reading of this paper.

References

BUSETTA, B. & COMBERTON, G. (1974). *Acta Cryst.* A**30**, 564–568.

FISHER, J., HANCOCK, J. & HAUPTMAN, H. (1970). Naval Research Laboratory Report 7132.

HAUPTMAN, H. (1974). *Acta Cryst.* A**30**, 822–829.

HAUPTMAN, H., FISHER, J., HANCOCK, H. & NORTON, D. A. (1969). *Acta Cryst.* B**25**, 811–814.

SCHENK, H. (1973). *Acta Cryst.* A**29**, 77–82.