

completely overwhelms  $f_1^V$  for  $q=0$  and 1, thus falls off very rapidly with  $q$ . If the cubic or icosahedral p.g. is purely rotational, 432 or 532, then at  $q=V/2$  for any polyhedron (or assembly of points) of this symmetry with no vertex (point) on a rotation axis  $\frac{1}{2}(1-V)\log 2 - (3 \log e)/4V + \log k \leq \log R < \frac{1}{2}(1-V)\log 2 + \log k$ , where  $k=9$  for 432 and 15 for 532. The linearity of  $\log R$  improves asymptotically as  $V$  increases. This expression shows that the effect of symmetry on  $N$  tends to vanish even in point groups of high symmetry if  $V$  is sufficiently large. For  $V=240$  and p.g. 532, for example,  $\log R \sim -34.80$ , which means that  $N(A_{120}B_{120}) \sim C(240, 120) \sim 10^{72}$ , so that the effect of symmetry relative to  $N$  is completely negligible.

We are indebted to Professor J. D. H. Donnay for a number of constructive comments on the draft manuscript. The costs of computing were defrayed by the National Research Council of Canada through a grant in aid of research, and by a grant from IBM Canada Ltd.

#### References

- BRITTON, D. & DUNITZ, J. D. (1973). *Acta Cryst.* **A29**, 362–371.
- CUNDY, H. M. & ROLLETT, A. P. (1961). *Mathematical Models*. 2nd ed. Oxford Univ. Press.
- FREUDENTHAL, H. & VAN DER WAERDEN, B. L. (1947). *Simon Stevin*, **25**, 115–121.
- GRÜNBAUM, B. (1967). *Convex Polytopes*. London: Interscience.
- HARARY, F. & PALMER, E. M. (1973). *Graphical Enumeration*. New York, London: Academic Press.
- HILL, T. L. (1943). *J. Chem. Phys.* **11**, 294–297.
- International Tables for X-ray Crystallography* (1952). Vol. I. Birmingham: Kynoch Press.
- KASPER, J. S. (1956). In *Theory of Alloy Phases*, p. 264. Cleveland, Ohio: American Society for Metals.
- KENNEDY, B. A., MCQUARRIE, D. A. & BRUBAKER, C. H. JR (1964). *Inorg. Chem.* **3**, 265–268.
- LUNN, A. C. & SENIOR, J. K. (1929). *J. Phys. Chem.* **33**, 1027–1079.
- MAIN SMITH, J. D. (1924). *Chemistry and Atomic Structure*, p. 97. London: Benn.
- MARCHI, L. E., FERNELIUS, W. C. & McREYNOLDS, J. P. (1943). *J. Amer. Chem. Soc.* **65**, 329–333.
- MILLER, J. C. P. (1954). *Table of Binomial Coefficients. Royal Society Mathematical Tables*, Vol. III. Cambridge Univ. Press.
- NIGGLI, P. (1941). *Lehrbuch der Mineralogie und Kristallchemie*, 3rd ed., Part 1. Berlin: Gebr. Borntraeger.
- NIGGLI, P. (1945). *Grundlagen der Stereochemie*. Basel: Birkhäuser.
- NOWACKI, W. (1933). *Z. Kristallogr.* **86**, 19–31.
- PÓLYA, G. (1936). *Z. Kristallogr.* **93**, 415–443.
- PÓLYA, G. (1937). *Acta Math.* **68**, 145–253.
- REDFIELD, J. H. (1927). *Amer. J. Math.* **49**, 433–455.
- SALTHOUSE, J. A. & WARE, M. J. (1972). *Point Group Character Tables and Related Data*, p. 84. Cambridge Univ. Press. p. 84.
- SCHULTZ, H. P. (1965). *J. Org. Chem.* **30**, 1361–1364.
- TRIMBLE, R. F. JR (1954). *J. Chem. Educ.* **31**, 176–179.
- WELLS, A. F. (1956). *The Third Dimension in Chemistry*. Oxford Univ. Press.
- WHITE, P. S. (1972). Ph.D. thesis. Dalhousie Univ., Halifax, N.S., Canada.

*Acta Cryst.* (1975). **A31**, 472

## Use of Negative Quartet Cosine Invariants as a Phasing Figure of Merit: NQUEST

BY GEORGE T. DE TITTA, JAMES W. EDMONDS,\* DAVID A. LANGS AND HERBERT HAUPTMAN

*Medical Foundation of Buffalo, 73 High Street, Buffalo, New York 14203, U.S.A.*

(Received 4 October 1974; accepted 5 February 1975)

Recent theoretical advances in the identification of those cosine invariants  $\cos(\varphi_h + \varphi_k + \varphi_l + \varphi_m)$  which are probably negative suggest algorithms for the calculation of a figure of merit which is sensitive to the integrity of a phase set. The negative quartet figure of merit, NQUEST, defined here is of particular utility in conjunction with fast multi-solution tangent formula techniques. Development of the methods and applications to both known and unknown crystal structures are presented.

#### Introduction

A general methodology of crystal structure determination which has found wide application in one form or

another is the multi-solution tangent refinement technique. Although the actual procedures employed within the general framework of the method may vary widely, the use of the tangent formula (Karle & Hauptman, 1956) to extend and refine a number of plausible basis sets of phases is a common feature to all. On one end of the spectrum are those procedures which introduce

\* Present Address: Dow Chemical U.S.A., Analytical Laboratories, Midland, Michigan 48640.

a number of ambiguities at the onset, assign them initial numerical values, and immediately proceed with tangent refinement and extension. These methods are characterized by the small size of the basis set of phases available at the commencement of tangent refinement and the rather large number of plausible phase sets which must be generated to insure that at least one correct phase set is reached. Such methods lend themselves to automation, and excellent program packages such as *MULTAN* (Germain, Main & Woolfson, 1971) which employ this approach are widely used. On the other end of the spectrum are those methods which seek to broaden the basis set of phases, introducing symbolic ambiguities when necessary, before proceeding with tangent refinement. Such methods tend to minimize the number of ambiguities by minimizing the number of 'conflicts' in phase assignment, such as in the symbolic addition procedure (Zachariasen, 1952; Karle & Karle, 1966). These methods do not lend themselves as easily to automation.

Both methods eventually produce a number of plausible phase sets which then must be tested for correctness. If only a small number of phase sets need be tested, then an inspection of the resulting  $E$  maps poses no great problem. If, however, a large number of phase sets must be tested, the calculation and inspection of the resulting  $E$  maps may present a real problem in terms of computing time and possible inspection errors. Ideally one would like to know *a priori* whether a plausible phase set is likely to yield a solution without calculating and inspecting its resulting  $E$  map. Various figures of merit, such as the 'absolute figure of merit' (Germain, Main & Woolfson, 1971), the  $\Psi_0$  test (Cochran & Douglas, 1955), and the residual  $R_K$  (Karle & Karle, 1966) have been proposed to rank phase sets in the order of their plausibility. However it has been found by experience that none of these figures of merit consistently discriminates against hopelessly incorrect phase sets, nor can they be depended upon to indicate the presence of a correct phase set. Declercq, Germain, Main & Woolfson (1973), recognizing the poor performance of these figures of merit, have incorporated a fast Fourier transform program into the most recent *MULTAN* package which automatically calculates  $E$  maps for all phase sets and displays them in an easily usable form. Nevertheless, it would be clearly desirable to construct a figure of merit which is reliable, absolute, and easily calculated. Recently secured estimates for the cosine invariants,  $\cos(\varphi_h + \varphi_k + \varphi_l + \varphi_m)$ , make it possible to propose such a figure of merit, NQUEST.

### Method

The quartets are the four-phase cosine invariants  $\cos(\varphi_h + \varphi_k + \varphi_l + \varphi_m)$  where  $h+k+l+m=0$ . Hauptman (1974a,b) has shown, for sufficiently large  $B = (2/N) |E_h E_k E_l E_m|$ , where  $N$  is the number of assumed

identical atoms in the unit cell, that

$$\cos(\varphi_h + \varphi_k + \varphi_l + \varphi_m) \simeq \begin{cases} +1 & \text{if } |E_{h+k}|, |E_{h+l}|, |E_{h+m}| \\ & \text{are all large} \quad (1a) \\ -1 & \text{if } |E_{h+k}|, |E_{h+l}|, |E_{h+m}| \\ & \text{are all small.} \quad (1b) \end{cases}$$

We define a PQ (positive quartet) to be a quartet with 'cross terms'  $\{|E_{h+k}|, |E_{h+l}|, |E_{h+m}|\}$  all large and a NQ (negative quartet) to be one with cross terms all small. Note that a PQ is not necessarily positive nor is a NQ necessarily negative.

We will restrict our subsequent attention to the NQ because they form the basis of the required figure of merit, but it should be apparent that the PQ are equally important in their own right. The PQ are not, however, suitable as the basis for the required figure of merit because they are insensitive to the integrity of a phase set derived from tangent formula procedures. The reason for this is that tangent procedures generally insure that cosine invariants  $\cos(\varphi_h + \varphi_k + \varphi_{-h-k}) \simeq +1$ , for large  $A = (2/N^{1/2}) |E_h E_k E_{-h-k}|$ , that is, the strongest phase relationships are generally not violated by the tangent formula procedure. It can be easily shown that, for a given quartet invariant, if any one of the conditions:

$$\{\cos(\varphi_h + \varphi_k + \varphi_{-h-k}) = +1 \\ \text{and } \cos(\varphi_l + \varphi_m + \varphi_{h+k}) = +1\},$$

or

$$\{\cos(\varphi_h + \varphi_l + \varphi_{-h-l}) = +1 \\ \text{and } \cos(\varphi_k + \varphi_m + \varphi_{h+l}) = +1\},$$

or

$$\{\cos(\varphi_h + \varphi_m + \varphi_{-h-m}) = +1 \\ \text{and } \cos(\varphi_k + \varphi_l + \varphi_{h+m}) = +1\} \quad (2)$$

is fulfilled, then necessarily  $\cos(\varphi_h + \varphi_k + \varphi_l + \varphi_m) = +1$ .

The cosine invariants of equation (2) are the six cosine triples associated with each quartet; *i.e.*, the associated triples. For PQ the  $A$  values of the associated triples are large, by equation (1a), and for NQ the  $A$  values of the associated triples are small. Therefore we expect the phase sets generated by tangent formula procedures to identify the PQ well, regardless of the correctness of the set. However, the tangent refinement procedure is ordinarily designed to avoid the associated triples of the NQ and therefore to avoid biasing the proposed figure of merit.

For some specific values of  $B_{\min}$  and  $E_{\text{cross}}^*$  construct  $n$  quartets such that  $B \geq B_{\min}$  and  $\{|E_{h+k}|, |E_{h+l}|, |E_{h+m}|\} \leq E_{\text{cross}}$ . In addition, for a specific value of  $E_{\text{main}}$ , insure that only quartets of interest will be constructed by imposing the condition that the 'main terms' satisfy  $\{|E_h|, |E_k|, |E_l|, |E_m|\} \geq E_{\text{main}}$ .<sup>\*</sup> This latter condition limits the quartets generated to those with each main term greater than or equal to  $E_{\text{main}}$ , the

\*  $E_{\text{cross}}$  is defined as the least upper bound of the cross terms and  $E_{\text{main}}$  as the greatest lower bound of the main terms.

smallest normalized structure factor phased in the tangent refinement procedure. Then an estimate, NQUEST, of the integrity of the phase set with respect to the NQ is defined by

$$\text{NQUEST} = \frac{\sum_{h,k,l}^n B \cos(\varphi_h + \varphi_k + \varphi_l + \varphi_m)}{\sum_{h,k,l}^n B}$$

Clearly, NQUEST ranges from  $-1$  to  $+1$  and, in view of equation (1b), is expected to be negative for a correct set of phases provided that  $E_{\text{cross}}$  is sufficiently small and  $B_{\text{min}}$  is sufficiently large.

### Applications

A preliminary analysis of the negative quartet cosine invariants for an artificial 29 equal-atom structure in  $P1$  suggested the simple form of a figure of merit based on the NQ. An inspection, Fig. 1, of the observed conditional average cosine

$$\langle \cos \Phi_4 \rangle \equiv \langle \cos(\varphi_h + \varphi_k + \varphi_l + \varphi_m) \{ |E_{h+k}|, |E_{h+l}|, |E_{h+m}| \} \leq E_{\text{cross}} \rangle_{h,k,l}$$

for various small values of  $E_{\text{cross}}$  as a function of  $B$  indicates that  $\langle \cos \Phi_4 \rangle$  is everywhere negative and approaches  $-1$  in the limit of large  $B$ . Additionally, the standard deviation of the cosine,  $\sigma(\cos \Phi_4)$ , decreases as  $B$  increases, indicating that the estimate for  $\cos(\varphi_h + \varphi_k + \varphi_l + \varphi_m)$  given in equation (1b) becomes increasingly reliable at large  $B$  values. These results suggest that, for a cross term cutoff which is small, say  $E_{\text{cross}} \leq 0.85$ , a simple average cosine weighted on  $B$  would be sensitive to the correctness of a phase set. NQUEST, by definition, is normalized to lie in the range  $\{-1, +1\}$ , thereby eliminating the dependence of the figure of merit on the number of NQ generated.

In order to determine the effect of varying  $B_{\text{min}}$  and  $E_{\text{cross}}$  on NQUEST, negative quartet cosine invariants were constructed for the known structure  $d$ -(+)-biotin (DeTitta, Edmonds, Stallings & Donohue, 1975). NQUEST was calculated using the refined ( $R=0.055$ ) phases and the criteria in Table 1. In each case NQUEST is negative. As the restrictions on  $E_{\text{cross}}$  and  $B_{\text{min}}$  are relaxed; *i.e.*, as  $E_{\text{cross}}$  increases or  $B_{\text{min}}$  decreases, the total number of NQ increases and NQUEST becomes less negative. In actuality, these variations are in complete accord with the theoretical predictions which state that, for fixed small values of  $|E_{h+k}|, |E_{h+l}|, |E_{h+m}|$ , the larger the value of  $B$  the more negative the

cosine is likely to be and that, for a fixed value of  $B$ , the smaller the values of  $|E_{h+k}|, |E_{h+l}|, |E_{h+m}|$ , the more negative the cosine will be. Qualitatively we observe that  $E_{\text{cross}} > 1.0$  is too large to be considered 'small' for normal values of  $B_{\text{min}}$ , and  $B_{\text{min}} < 0.25$  is too small to be considered 'sufficiently large' for the conditions in equation (1b) to hold. In practice we have found  $E_{\text{cross}} \sim 0.7$  to be a good starting value with  $B_{\text{min}}$  chosen such that a sufficient number ( $\sim 100$ – $300$ ) of NQ are constructed. An alternative method might employ a sliding scale for  $E_{\text{cross}}$  as a function of  $B_{\text{min}}$  in an attempt to make NQUEST uniform for all ranges of  $B$  considered.

Table 1. Variation of NQUEST with  $B_{\text{min}}$  and  $E_{\text{cross}}$  for  $d$ -biotin

The number in brackets refers to the total number of NQ

$E_{\text{cross}}$	0.50	0.75	1.00
$B_{\text{min}}$			
0.25	-0.47 [717]	-0.33 [3519]	$\geq 3519$ *
0.50	-0.69 [39]	-0.54 [137]	-0.29 [421]
0.75	-0.99 [3]	-0.83 [10]	-0.71 [17]

\* The NQUEST value for  $B_{\text{min}}=0.25, E_{\text{cross}}=1.00$  could not be readily calculated because of insufficient computer storage. We estimate the number of NQ for this particular set of  $B_{\text{min}}$  and  $E_{\text{cross}}$  values to be in excess of 40000.

The total number of NQ generated for specific  $B_{\text{min}}, E_{\text{main}}$ , and  $E_{\text{cross}}$  seems to be strongly structure dependent. Negative quartet invariants were generated using the criteria of Table 2 for the refined crystal structures prostaglandin  $A_1$ : monoclinic form (PGA<sub>1</sub>M DeTitta, Edmonds & Duax, 1975), and orthorhombic form (PGA<sub>1</sub>O, Duax & Edmonds, 1973; Edmonds & Duax, 1975), prostaglandin  $E_2$  (PGE<sub>2</sub>, Edmonds & Duax, 1975), indomethacin (INDOM, Kistenmacher

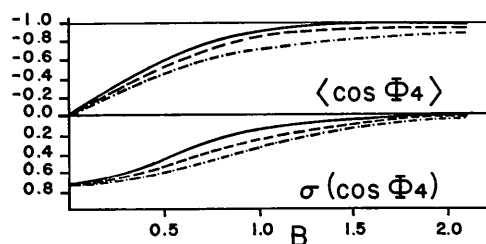


Fig. 1. Plots of  $\langle \cos \Phi_4 \rangle \equiv \langle \cos(\varphi_h + \varphi_k + \varphi_l + \varphi_m) \{ |E_{h+k}|, |E_{h+l}|, |E_{h+m}| \} \leq E_{\text{cross}} \rangle_{h,k,l}$  and  $\sigma(\cos \Phi_4)$  as functions of  $B$  at values of  $E_{\text{cross}}=0.35$  (—),  $0.50$  (---) and  $0.85$  (-·-·-) for a 29 equal-atom artificial structure in  $P1$ .

Table 2. Variation of NQUEST with structure

See text for explanation of the abbreviations.

Structure	Space group	$N$	$\langle ( E ^2 - 1)^2 \rangle$	$E_{\text{main}}$	$B_{\text{min}}$	$E_{\text{cross}}$	$\# \text{NQ}$	NQUEST
Biotin	$P2_12_12_1$	64	1.115	1.5	0.5	0.7	137	-0.55
CLACM	$P2_1$	54	1.360	1.5	0.5	0.7	961	-0.41
PGA <sub>1</sub> O	$P2_12_12_1$	96	1.627	1.5	0.5	0.7	84	-0.36
PGE <sub>2</sub>	$P1$	25	1.893	1.5	0.5	0.7	154	-0.35
PGA <sub>1</sub> M	$P2_1$	48	2.485	1.75	1.0	0.7	123	-0.64
INDOM	$P1$	50	2.969	1.75	1.0	0.7	280	-0.66





Table 6 (cont.)

(b) 64 Phase Sets		RESID		NQUEST		Set #		ABSFOM		RESID		NQUEST		Set #		ABSFOM		RESID		NQUEST	
Set #	ABSFOM	RESID	NQUEST	Set #	ABSFOM	RESID	NQUEST	Set #	ABSFOM	RESID	NQUEST	Set #	ABSFOM	RESID	NQUEST	Set #	ABSFOM	RESID	NQUEST		
1	0.02	33.3	0.01	17	0.93	29.8	0.19	33	0.95	29.7	0.16	49	0.94	30.9	0.10						
2	0.88	32.0	0.11	18	0.96	32.3	0.12	34	0.97	31.4	0.14	50	0.90	32.6	0.17						
3	0.88	31.9	0.06	19	0.06	30.6	0.09	35	0.98	30.7	0.10	51	0.91	32.1	0.14						
4	0.92	32.8	0.09	20	0.93	30.3	0.14	36	0.96	29.9	0.15	52	0.97	30.2	0.06						
5	0.98	27.5	-0.34	21	0.97	31.6	0.14	37	0.88	34.3	0.31	53	0.95	29.7	0.16						
6	0.99	27.5	-0.34	22	0.95	29.7	0.16	38	0.88	34.3	0.31	54	0.97	31.6	0.14						
7	0.99	27.2	-0.38	23	0.95	29.9	0.15	39	0.87	34.3	0.30	55	0.98	30.8	0.09						
8	0.99	27.2	-0.38	24	0.98	30.8	0.09	40	0.88	34.1	0.25	56	0.95	39.8	0.15						
9	0.88	32.0	0.11	25	0.90	32.6	0.17	41	0.97	31.4	0.14	57	0.96	32.2	0.12						
10	0.92	33.3	0.01	26	0.94	30.9	0.10	42	0.95	29.7	0.16	58	0.93	29.8	0.19						
11	0.92	32.8	0.10	27	0.97	30.2	0.06	43	0.96	29.9	0.15	59	0.93	30.2	0.14						
12	0.88	31.9	0.06	28	0.91	32.1	0.14	44	0.98	30.7	0.10	60	0.96	30.6	0.09						
13	0.96	30.9	0.05	29	0.99	27.9	-0.29	45	0.93	30.0	0.18	61	0.92	32.0	-0.07						
14	0.96	30.9	0.05	30	0.92	32.0	-0.07	46	0.92	29.4	0.19	62	0.99	27.9	-0.29						
15	0.97	30.6	0.07	31	0.92	32.6	-0.05	47	0.93	29.7	0.15	63	1.00	27.4	-0.28						
16	0.97	30.6	0.07	32	1.01	27.4	-0.28	48	0.94	30.4	0.11	64	0.92	32.6	-0.05						

culated for that phase set revealed all but one atomic position,  $\frac{24}{25}$  atoms being the strongest peaks in the map.

### Results

In the test cases using known crystal structures it has been shown that, if  $B_{\min}$  is sufficiently large and  $E_{\text{cross}}$  sufficiently small, NQUEST is always negative. Specifically, if  $B_{\min} \approx 0.5$  and  $E_{\text{cross}} \approx 0.7$ , then NQUEST will lie in the range  $-0.6$  to  $-0.3$  for the refined set of phases. Confronted with the choice of lowering the  $B_{\min}$  cutoff or raising the  $E_{\text{cross}}$  cutoff to increase the number of NQ generated, the former has been found to be preferable. Although the probability that  $\cos(\varphi_h + \varphi_k + \varphi_l + \varphi_m) < 0$  is a function of both  $B$  and the magnitudes of the cross terms  $|E_{h+k}|, |E_{h+l}|, |E_{h+m}|$ , the dependence on  $B$  is easily taken into account in NQUEST while the dependence on the cross terms is much more complicated (Hauptman, 1974a). Until additional tests of NQUEST on solved crystal structures are made, only plausible conclusions regarding the dependence of NQUEST and the total number of NQ generated on the cutoff criteria,  $B_{\min}$ ,  $E_{\text{cross}}$ ,  $E_{\text{main}}$ , and on the basic characteristics of the structure, such as  $N$ , space group and the Patterson overlap function  $\langle (|E|^2 - 1)^2 \rangle$ , can be made. However, the initial results are encouraging in that they fit the theoretical predictions.

In the test cases using unknown crystal structures the results have been very satisfying. Not only has NQUEST performed better than the more traditional figures of merit but the expected independence of NQUEST from the tangent refinement procedure has also been borne out. In the case of  $\text{PGA}_1\text{M}$ , ABSFOM and RESID were completely unable to rank the sets in the order of their plausibility. The  $\Psi_0$  test proved more sensitive although it might be argued that the difference of  $\Psi_0 = 0.98$  for a correct phase set and  $\Psi_0 = 1.02$  for an incorrect phase set is, at best, on the fringe of significance. On the other hand, NQUEST proved to be highly sensitive; the four phase sets with  $\text{NQUEST} \approx -0.5$  containing two correct solutions and two correctly oriented but translated solutions, with all other phase sets having  $\text{NQUEST} \geq 0.6$ .

Of perhaps equal significance is the 'absoluteness' of NQUEST as an estimate of the correctness of a phase set. By this is meant that NQUEST not only points out which of  $n$  plausible phase sets is *most* likely to be correct; *i.e.* it is a ranking procedure; but it also decides whether *any* of  $n$  phase sets is likely to be correct. For example, in the case of  $\text{INDOM}$  it is doubtful whether or not any  $E$  maps would have been calculated when  $\text{MULTAN}$  was permitted to generate eight and 16 plausible phase sets, Tables 7(a) and (b), because in these instances the best NQUEST values were  $+0.03$  and  $-0.03$  respectively. It may be premature to make definitive statements about what value of NQUEST implies that a given phase set is correct but it seems safe to conclude that NQUEST cannot be positive for a

correct set of phases when calculated with the previously suggested cutoff criteria. Borderline cases, where NQUEST is slightly negative (for example, sets 7 and 16 for CLACM, Table 6(a),  $NQUEST = -0.24$ ) and partial structure is revealed in the  $E$  map, point out the need to optimize the cross term, main term, and  $B$  criteria to insure a clearer indication in these situations.

In terms of computing time, NQUEST represents a substantial saving compared with the calculation of all  $E$  maps, even when a fast Fourier transform program is available. The generation of 154 NQ from a set of 170 normalized structure factors used as 'main terms' for  $PGE_2$  consumes approximately the same computing time as the synthesis of a single  $E$  map from the same 170 terms in conjunction with a fast Fourier program. The generation of the NQ represents the major portion of the combined NQ/NQUEST operation; the time required for the evaluation of NQUEST for the phase sets is trivial. Thus the use of NQUEST is justified in any situation which suggests that more than one  $E$  map may have to be calculated in the normal course of multi-solution tangent refinement analysis.

### Discussion

Schenk (1973) recognized the utility of the quartet cosine invariants which are negative in a multi-solu-

tion symbolic addition procedure and has applied a somewhat analogous figure of merit based upon the special negative quartet invariants where  $h=k$  (Schenk & de Jong, 1973) which were developed from the Harker-Kasper inequalities. More recently, Schenk (1974) has discussed the use of negative quartets and explored the strength of the NQ's relative to the  $\Sigma_2$  relation. On the basis of a failure frequency distribution study Schenk has concluded that the number of NQ which are reliably negative (*versus* the number of  $\Sigma_2$  which are reliably positive) is so small as to render the NQ's virtually useless for the evaluation of phases. This pessimism is considered to be unwarranted for the following reasons: (1) it is probably incorrect to expect the estimation of a negative cosine invariant involving four phases,  $\cos(\varphi_h + \varphi_k + \varphi_l + \varphi_{-h-k-l}) \simeq -1$ , to be directly comparable to the estimation of a positive cosine invariant involving three phases,  $\cos(\varphi_h + \varphi_k + \varphi_{-h-k}) \simeq +1$ ; (2) a comparison of the failure frequency distributions of NQ's and  $\Sigma_2$ 's at values of  $B=A$  implies that  $B$  and  $A$  are comparable quantities whereas recent theoretical results (Hauptman, 1975) imply that the comparable quantities are  $A$  and  $2B$ ; (3) the cutoff criterion for the cross terms suggested by Schenk; *i.e.* that the average value of the cross terms be less than a specified value, is not only less restrictive than the condition that all three cross terms be smaller than a specified value, but recent theoretical

Table 7. *Figures of merit for INDOM*

(a) 8 Phase sets				(c) 32 Phase sets			
Set #	ABSFOM	RESID	NQUEST	Set #	ABSFOM	RESID	NQUEST
1	1.16	28.8	1.00	1	1.15	17.1	1.00
2	0.87	41.3	0.42	2	0.91	29.1	-0.15
3	0.73	44.4	0.15	3	0.71	37.4	-0.15
4	0.91	39.7	0.47	4	0.69	37.6	0.08
5	0.63	48.0	0.03	5	0.81	32.0	0.05
6	0.77	45.2	0.24	6	0.74	33.6	-0.15
7	0.92	39.2	0.44	7	0.70	37.7	0.09
8	0.71	44.8	0.21	8	0.67	38.9	0.10
				9	0.66	37.3	-0.18
				10	0.78	34.3	0.34
				11	0.84	31.8	0.16
				12	0.84	31.1	-0.13
				13	0.71	35.9	-0.14
				14	0.82	33.9	0.41
				15	0.90	28.0	0.02
				16	0.90	31.1	0.46
				17	0.80	32.6	-0.11
				18	0.74	37.8	0.36
				19	0.86	30.3	-0.26
				20	0.86	31.1	0.18
				21	0.71	36.0	-0.11
				22	0.72	38.4	0.24
				23	0.91	31.5	0.42
				24	0.79	34.8	0.11
				25	0.81	32.3	0.25
				26	0.79	33.8	0.10
				27	0.66	37.9	0.11
				28	0.71	37.0	0.16
				29	0.97	28.4	0.39
				30	0.99	24.1	-0.55
				31	0.68	38.4	0.00
				32	0.78	32.9	0.53

(b) 16 Phase sets			
Set #	ABSFOM	RESID	NQUEST
1	1.15	17.1	1.00
2	0.89	29.2	0.35
3	0.70	40.5	0.14
4	0.80	33.6	-0.03
5	0.68	37.3	0.31
6	0.82	33.6	0.49
7	0.84	32.3	0.30
8	0.79	33.7	-0.02
9	0.71	38.3	0.17
10	0.75	35.2	0.27
11	0.91	31.5	0.43
12	0.91	30.1	0.54
13	0.97	29.0	0.39
14	0.91	28.9	0.10
15	0.82	32.4	0.51
16	0.77	34.0	0.25

results (Hauptman, 1975) also show that under certain conditions the expected value of  $\cos(\varphi_h + \varphi_k + \varphi_l + \varphi_m)$  can be positive if two cross terms are moderately small and the third cross term is moderately large. Additionally it should be emphasized that the identification of even a small number of negative cosine invariants, whether quartets or triples, is extremely helpful in the solution of crystal structures by direct methods.

Schenk (1974) has proposed a practical symbolic addition procedure for symmorphic space groups which employs a figure of merit based on the NQ. As a modification, an analogous scenario for more difficult problems is proposed:

(1) Construct  $\sum_2$  and NQ lists. Select origin, enantiomorph, and suitable basis vectors to produce  $P$  plausible phase sets.  $P$  may well be of the order of  $2^{10}$ .

(2) Initiate tangent refinement and extension for the  $P$  plausible phase sets. Stop when phase values have been obtained for some 30 or more NQ; and calculate NQUEST over the  $P$  plausible phase sets.

(3) Continue tangent extension for, let us say, the  $P/2$  plausible phase sets with the best NQUEST until phase values have been obtained for an additional 40–50 NQ, and calculate NQUEST over these phase sets.

(4) Continue tangent extension for, let us say, the  $P/8$  plausible phase sets with the best NQUEST until phase values have been obtained for an additional 40–50 NQ.

The threshold values of  $P/2$  and  $P/8$  in steps 3 and 4 have been arbitrarily selected to indicate that the rejection criteria may become more selective as larger numbers of NQ are reached through the phase extension.

This process is repeated until a specified small number of plausible phase sets is reached and then tangent refinement for these phase sets is completed. This procedure insures that only that specific small fraction of the  $P$  plausible basis sets with the best NQUEST values will be extended to the limit of the tangent refinement. The initial weeding of the  $P$  ambiguities might well be accomplished at a point where the tangent refinement procedure has extended the basis set to as few as 50 phases. Since the time required for successive cycles of tangent refinement increases rapidly, a relatively short time may be spent on steps 2, 3 and 4. This procedure permits the luxury of a large initial basis set, without the attendant expense of a full tangent refinement calculation over all plausible phase sets. Necessarily this procedure is

recommended only if less time-consuming procedures such as *MULTAN*/NQUEST fail to yield the desired results.

In summary then, experience has shown that the traditional figures of merit often do not discriminate strongly between correct and incorrect phase sets and can even, on occasion, be grossly misleading. NQUEST, on the other hand, consistently and accurately discriminates between the two.

Our special thanks are extended to Steve Potter whose help to one of us (GDT) in coding the programs to generate negative quartet cosine invariants was indispensable. Additional thanks are extended to Dr R. Parthasarathy for calculating the  $\Psi_0$  results for  $PGA_1M$  and to Dr William Duax for helpful discussions. This work was supported by N.I.H. Research Grant No. HL-15378.

#### References

- COCHRAN, W. & DOUGLAS, A. S. (1955). *Proc. Roy. Soc. A* **227**, 486–500.
- DECLERQ, J. P., GERMAIN, G., MAIN, P. & WOOLFSON, M. M. (1973). *Acta Cryst. A* **29**, 231–234.
- DE TITTA, G. T., EDMONDS, J. W. & DUAX, W. L. (1975). *Prostaglandins*. In the press.
- DE TITTA, G. T., EDMONDS, J. W., STALLINGS, W. & DONOHUE, J. (1975). *J. Amer. Chem. Soc.* Submitted.
- DUAX, W. L. & EDMONDS, J. W. (1973). *Prostaglandins*, **3**, 201–208.
- EDMONDS, J. W. & DUAX, W. L. (1974). *Prostaglandins*, **5**, 275–281.
- EDMONDS, J. W. & DUAX, W. L. (1975). *J. Amer. Chem. Soc.* **97**, 413–417.
- GERMAIN, G., MAIN, P. & WOOLFSON, M. M. (1971). *Acta Cryst. A* **27**, 368–376.
- HAUPTMAN, H. (1964). *Acta Cryst.* **17**, 1421–1433.
- HAUPTMAN, H. (1974a). *Acta Cryst. A* **30**, 472–476.
- HAUPTMAN, H. (1974b). *Acta Cryst. A* **30**, 822–829.
- HAUPTMAN, H. (1975). *Acta Cryst. A*. In the press.
- KARLE, J. & HAUPTMAN, H. (1956). *Acta Cryst.* **9**, 635–651.
- KARLE, J. & KARLE, I. (1966). *Acta Cryst.* **21**, 849–859.
- KISTENMACHER, T. J. & MARSH, R. E. (1972). *J. Amer. Chem. Soc.* **94**, 1340–1345.
- ROHRER, D. C., DUAX, W. L. & WOLFF, M. (1974). Personal communication.
- SCHENK, H. (1973). *Acta Cryst. A* **29**, 480–482.
- SCHENK, H. (1974). *Acta Cryst. A* **30**, 477–481.
- SCHENK, H. & DE JONG, J. G. H. (1973). *Acta Cryst. A* **29**, 31–34.
- ZACHARIASEN, W. H. (1952). *Acta Cryst.* **5**, 68–73.