lead to virtually coherent areas of domains with the same phase state. Thus the line width is a function of the statistics of the domain configurations and the strength of the interaction which depends on temperature and the characteristics of the domain boundary.

We wish to acknowledge the financial support by the Deutsche Forschungsgemeinschaft (project Ja 15/32).

References

ADLHART, W., FREY, F. & JAGODZINSKI, H. (1980). Acta Cryst. A36, 450–460.

BROWN, W. L., HOFFMANN, W. & LAVES, F. (1963). Naturwissenschaften, 50, 221.

CZANK, J., VAN LANDUYT, J., SCHULZ, H., LAVES, F. & AMELINCKX, S. (1973). Z. Kristallogr. 138, 403–418.

FOIT, F. F. JR & PEACOR, D. R. (1973). Am. Mineral. 58, 665–675.

HEUER, A. H. & NORD, G. L. JR (1976). Polymorphic Phase Transitions in Minerals. In Electron Microscopy, edited by H. R. WENK, Chapter 5.1. Berlin: Springer.

HEUER, A. H., NORD, G. L. JR, LALLY, J. S. & CHRISTIE, J. M. (1976). Origin of the (c) Domains in Anorthite. In Electron Microscopy, edited by H. R. WENK, Chapter 5.7. Berlin: Springer.

JAGODZINSKI, H. & KOREKAWA, M. (1973). Geochim. Cosmochim. Acta, Suppl. 4, Vol. 1, 933–951.

JAGODZINSKI, H. & KOREKAWA, M. (1976). Z. Kristallogr. 143, 239–277.

LAVES, F., CZANK, M. & SCHULZ, H. (1970). Schweiz. Mineral. Petrogr. Mitt. 50, 519–525.

MCLAREN, A. C. & MARSHALL, D. B. (1974). Contrib. Mineral. Petrol. 44, 237–249.

MÜLLER, W. F., WENK, H. R., BELL, W. L. & THOMAS, G. (1973). Contrib. Mineral. Petrol. 40, 63–74.

MÜLLER, W. F., WENK, H. R. & THOMAS, G. (1972). Contrib. Mineral. Petrol. 34, 304–314.

ONSAGER, L. (1944). Phys. Rev. 65, 117–149.

The Use of Higher Invariants in MULTAN

BY A. A. FREER AND C. J. GILMORE

Department of Chemistry, University of Glasgow, Glasgow G12 8QQ, Scotland

Abstract

A method of using quartets and quintets in the direct-methods program MULTAN is described with several successful applications.

Introduction

MULTAN is the most widely used direct-methods computer program. Several extensions to the system have been described recently in which magic integer/ψ map and random phase set/linear equation algorithms have been employed (Declercq, Germain & Woolfson, 1979). These developments undoubtedly enhance the power of the program. However, they are confined to the use of three-phase invariants. Recently, there has been considerable activity in deriving formulae for estimating the magnitudes of four- and five- phase structure invariants (Hauptman, 1977a,b; Giacovazzo, 1976a,b; van der Putten & Schenk, 1977). These relationships contain new phase information, and it is therefore a logical extension of the MULTAN procedure

to incorporate them into this program and determine any advantages that thereby accrue.

Two modes of usage of higher invariants must be distinguished:

(a) the active mode in which the invariants are used to generate new phase information;

(b) the passive mode where the invariants are used only for figures of merit for selecting the most probable phase set.

Of course, the use of (a) does not preclude the use of (b) and both techniques will be discussed in this paper.

Quartets

The MULTAN 78 program was modified to use the seven-magnitude, second neighbourhood $P_{1/7}$ and the 13-magnitude, third neighbourhood $P_{1/13}$ joint conditional probability distributions of Hauptman (1977a) for the non-centrosymmetric case and the corresponding $P_7^{\pm}$ and $P_{13}^{\pm}$ formulae for centrosymmetric space groups (Hauptman, 1977b). These formulae give reliable invariant estimates for all combinations of the

principal and cross terms that comprise the quartet. The seven-magnitude formulae are straightforward in their estimation but $P_{1/13}$ and $P_{13}^{\pm}$ require further comment. These formulae use the third neighbourhood of the quartet. For the invariant

$$\varphi_h + \varphi_k + \varphi_l + \varphi_m \simeq \varPhi_4, \qquad (1)$$

where $h + k + l + m = 0$, a second invariant is constructed:

$$\varphi_h + \varphi_k + \varphi_p + \varphi_q \simeq \varPhi_4', \qquad (2)$$

where $h + k + p + q = 0$ and $|E_p|$ and $|E_q|$ are large but otherwise unrestricted. The magnitude of $\varPhi_4$ now depends on the first and second neighbourhoods ($|E_h|$, $|E_k|$, $|E_l|$, $|E_m|$; $|E_{h+k}|$, $|E_{k+l}|$, $|E_{l+h}|$) and the third neighbourhood ($|E_{h+p}|$, $|E_{k+p}|$, $|E_{l-p}|$, $|E_{m-p}|$, $|E_p|$, $|E_q|$) (Hauptman, 1977c). The floating vectors p and q result in a multiplicity of third neighbourhoods each giving rise to an estimate of $|\varPhi_4|$ and its associated variance V_{13}. Because more E magnitudes are involved, V_{13} can be significantly lower than that derived from the seven-magnitude formula (V_7). This can be especially useful in larger structures where the N^{-1} dependence of the quartet distributions can make it difficult to find reliable invariants. (N is the number of non-hydrogen atoms in the unit cell.)

The multiple estimates for $|\varPhi_4|$ can also be useful. From observations of several known structures, it has been found that the most reliable quartets are those in which all the individual third-neighbourhood estimates for $|\varPhi_4|$ agree. Where any discrepancies occur between these estimates, there is a greater likelihood of the individual $|\varPhi_4|$ estimates being incorrect (Freer & Gilmore, 1980). This is equivalent to searching for discrepant trio relationships (De Titta, Edmonds, Langs & Hauptman, 1976). Accordingly all such discrepant quartets are excluded from active use in the phasing procedures. In a similar way quartets which exhibit discrepancies between the second and third neighbourhood $|\varPhi_4|$ estimates are also excluded.

The third neighbourhood is also useful when working with poor quality data sets in which a large proportion of the data are weak or missing. Multiple indications from $P_{1/13}$ or $P_{13}^{\pm}$ can often find one or more reliable $|\varPhi_4|$ estimates where the second-neighbourhood formula may have difficulty (Gilmore, Hardy, MacNicol & Wilson, 1977). The penalty for this additional information is a considerable increase in the amount of computer time needed although this can be alleviated to some extent by severely restricting the magnitudes of $|E_p|$ and $|E_q|$. The use of the 13-magnitude formulae is thus an option in the program. When it is used a single estimate of V_{13} or $P_{13}^{\pm}$ is required for later use; the best value encountered in the third-neighbourhood search is used.

Missing cross terms in the second neighbourhood (but not the third) are also permitted as an option. This can be useful also for poor-quality data sets. The distributions derived by Heinerman (1978) are used.

Quintets

As with quartets, there are several quintet distributions available employing the first and second neighbourhoods. Of the formulae we have tested, the $P_{1/15}$ distribution of van der Putten & Schenk (1977) gives reliable quintet estimates for the non-centrosymmetric case whilst the $P_{15}^{\pm}$ formulae of Fortier & Hauptman (1977) is used for centrosymmetric space groups. These formulae are thus available as a user option in the enhanced $MULTAN$ 78 system. As a further check on reliability, the discriminant, Δ, as defined by Fortier & Hauptman is employed, where

$$\Delta = \frac{2\sigma_3^3}{\sigma_2^{9/2}} \sum_{15} R_{12}^2 R_{34}^2 - \frac{2\sigma_3}{\sigma_2^{9/2}} (3\sigma_3^2 - \sigma_2 \sigma_4) \sum_{10} R_{12}^2$$
$$+ \frac{2}{\sigma_2^{9/2}} (15\sigma_3^3 - 10\sigma_2 \sigma_3 \sigma_4 + \sigma_2^2 \sigma_5), \qquad (3)$$

where $\sigma_n = \sum_{j=1}^{N} Z_j^n$ and R_{12}, R_{34}, etc. are the magnitudes of the cross terms of the quintet. Δ is clearly a function of these cross terms, but they are used in a different way from that in the distributions mentioned above and thus Δ can be used as an additional check of quintet reliability.

Quintets are expensive to generate. In this program only the negative quintet subset is calculated. As an empirical observation, quintets having $\Delta \leq -1\cdot0$ and which also have an estimated magnitude of $180°$ are the most reliably negative and only these invariants are used. Furthermore, they are utilized only in the passive mode, since their information content for one individual phase is quite low and they can readily give rise to large accumulated errors when used to generate new phase angles especially when magic integers are employed as initial representations of the phases.

Active use of quartets

In its original form, $MULTAN$ 78 uses only triplets. As a measure of the reliability of each three-phase invariant an associated variable κ_{hkl} is used where

$$\kappa_{hkl} = 2\sigma_3 \sigma_2^{-3/2} |E_h E_k E_l|. \qquad (4)$$

E_h, E_k, E_l are the three E magnitudes involved in the triplet. This variable is used throughout the convergence mapping and tangent refinement procedures. In order to mix the quartets with these relationships it is

necessary to use the same scale of reliability. For the non-centrosymmetric case, this is carried out as follows.

(a) For each quartet the relevant joint conditional probability distribution $P(\Phi)$ [where $P(\Phi)$ is either $P_{1/7}$ or $P_{1/13}$] is calculated in 45° intervals. The mode, $|\Phi_m|$, is found.

(b) The distribution is normalized *via* numerical integration using Simpson's rule, such that

$$\int_0^{2\pi} P(\Phi)\,\mathrm{d}\Phi = 1 \qquad (5)$$

(c) The associated variance, V, is also derived *via* numerical integration of the normalized distribution,

$$V = \int_0^{2\pi} (\Phi - |\Phi_m|)^2\,P(\Phi)\,\mathrm{d}\Phi. \qquad (6)$$

(d) Each quartet is assigned an equivalent κ value κ^{eq} related to V (in degrees^2) by an empirically derived equation.

$$\kappa^{eq}_{hklm} \simeq 5583/(V + 255). \qquad (7)$$

Only quartets for which $\kappa^{eq}_{hklm} > 0.6$ are accepted.

A similar procedure is used in the centrosymmetric case, but here the probability P^+ is converted directly to κ^{eq}_{hklm} *via* the relationship.

$$\kappa^{eq}_{hklm} = |\log_e (P^+/1 - P^+)|. \qquad (8)$$

In practice this procedure is readily automated and does not necessitate the use of large amounts of computer time. The three- and four-phase invariants can now be freely mixed together throughout the convergence mapping and tangent refinement routines.

For the latter a version of the formula of van der Putten & Schenk (1979) is used:

$$\tan \varphi_h = \frac{\text{sintri} + \text{sinqua}}{\text{costri} + \text{cosqua}} = \frac{T_h}{B_h}, \qquad (9)$$

where

$$\text{sintri} = \sum_k w_k\,w_l\,\kappa_{hkl} \sin (\varphi_k + \varphi_l)$$

and

$$\text{sinqua} = \sum_k \sum_l w_k\,w_l\,w_m\,\kappa^{eq}_{hklm}$$
$$\times \sin (\Phi_k + \Phi_l + \Phi_m + s|\Phi_4|)$$

with corresponding cosine expressions for costri and cosqua;

$$w_h = \tanh [\sigma_3\,\sigma_2^{-3/2}\,(T_h^2 + B_h^2)^{1/2}]. \qquad (10)$$

The variable $S = \pm 1$ and is chosen such that $\varphi_h = \varphi_k + \varphi_l + \varphi_m + s|\Phi_4|$ is closest to zero. This is only relevant

for enantiomorph-sensitive quartets, which can thus only be used when a value, albeit approximate, can be assigned to Φ_4.

It must be emphasized that (9) assumes that quartets and triplets are independent. In the case of triplets and negative quartets this is substantially true since the latter utilizes small E magnitudes in the cross terms. This mixture of quartets and triplets is one option in the program.

For the strongly positive quartets (those with a zero mode and a low variance), one or more of the cross terms will involve large E magnitudes. If these E's are also used in the triplets then the independence of the three- and four-phase invariants is lost. As the number of such cross terms in the quartet increases, this correlation will also increase until a point is reached where all three cross terms involve large E's. Under these circumstances, the quartet can be considered as an overlap of triplets with common phase angles. Thus the triplet and quartet contain similar information although it is used in a different way (Giacovazzo, 1980). In the absence of a theoretical estimate of the covariances of these relationships a simple, linear weighting scheme is employed in which κ^{eq}_{hklm} for the quartet is modified to give κ'_{hklm}:

$$\kappa'_{hklm} = \kappa^{eq}_{hklm} (1 - n/3), \qquad (11)$$

where n is the number of cross terms in the second neighbourhood for which the corresponding phase angle has been determined. The triplets keep the same weight. Thus at the beginning of phase determination where very few phase angles are known, most quartets have their full weight in the tangent formula, but as the phasing procedure continues this weight is progressively reduced to zero for the strongly positive quartets whilst the best negative invariants maintain their full weight.

This dynamic use of κ poses problems during convergence mapping, since it is no longer possible to predict an *a priori* value of κ. In this case, the quartets are still included but they are given the minimum value of κ that they are likely to achieve during phase expansion and refinement. This procedure seems satisfactory.

Figures of merit

MULTAN 78 uses three figures of merit – ABSFOM, ψ_0 and R_{Karle}. The availability of negative invariants allows two further figures of merit to be included.

(a) NQEST (De Titta, Edmonds, Langs & Hauptman, 1975) in the modification described by Gilmore (1977). This employs negative quartets:

$$\text{NQEST} = \sum_{h,k,l,m} w_{hklm} \cos(\varphi_h + \varphi_k + \varphi_l + \varphi_m)/\sum_{h,k,l,m} w_{hklm},$$

$$(12)$$

where the summation is taken over all those quartets for which the mode is predicted to be 180°. For the non-centrosymmetric case $w_{hklm} = 1/V_{hklm}$ whilst for the centrosymmetric case $w_{hklm} = 1 - 2P^+$.

(b) A similar negative quintet figure of merit (NQINT) in which the quartet terms in NQEST are replaced by quintet terms and the weights are modified by the discriminant Δ such that $w_{hklm} = -\Delta$. V_{hklm} applies to the non-centrosymmetric case and $-\Delta(1 - 2P^+)$ to the centrosymmetric situation. Only those quintets having $\Delta \leq 1 \cdot 0$ are included. These two figures of merit can be especially useful in symmorphic space groups.

NQEST and NQINT use information contained in the small E magnitudes. ψ_0 uses similar information but in a different way. It is useful at this stage to compare some values of ψ_0 with the negative invariant figures of merit to investigate their dependence. The results for seven structures are shown in Table 1.

It can be seen that NQEST and NQINT offer figures of merit substantially independent of ψ_0. NQEST and NQINT are highly correlated. Since NQINT uses negative quintets which are expensive to generate its use remains optional. All the available figures of merit are collated to give a single combined figure of merit CFOM by an obvious extension of the MULTAN 78 concept.

Applications in the centrosymmetric case

Relatively simple structures in centrosymmetric space groups can sometimes prove resistant to direct methods if a large portion of the molecule is substantially planar or exhibits some form of pseudo-symmetry. Our first example, $C_{19}H_{26}N_2O_6$ (unpublished) belongs to the former category, containing a 19 atom, planar, 3,5-dinitrobenzoate ester function. Several runs of MULTAN 78 produced confused E maps in which multiple, misplaced images of the planar portion of the molecule were visible. Inclusion of the 3,5-dinitrobenzoate function into MULTAN as a randomly oriented molecular fragment did not cure the problem. Negative quartets were generated via P_{13}^+ for the top 100 E magnitudes and included in an active and passive way in MULTAN. The solution with the best CFOM

(3·42 out of a possible 4·0, NQEST = −0·84) produced an E map in which the top 19 peaks formed the expected 19-atom planar fragment. The remaining atoms were easily located from a weighted Fourier synthesis.

For our second example we take the unpublished structure of a novel clathrate comprising a $C_{72}H_{90}S_6$ host and $\frac{1}{2}(C_{30}H_{50})$ guest in the asymmetric unit (93 non-hydrogen atoms). The space group was $P\bar{1}$. Symmorphic space groups often give difficulties with direct methods, and indeed routine applications of MULTAN 78 were unsuccessful. Accordingly all the quartets and a subset of negative quintets were generated for the top 120 E magnitudes. Routine inclusion of all the quartets in active mode into MULTAN 78 gave eight possible solutions. The set with the highest CFOM (4·45 out of a possible 5·0, NQEST = −0·83, NQINT = −0·88) produced an E map in which 65 atoms were clearly visible. The structure was completed by Fourier techniques.

This problem clearly shows the potential cost effectiveness of quartets if they are properly used. A simple MULTAN 78 run on this structure required 270 s of computer time on an ICL 2976 (roughly equivalent to an IBM 370/168). A re-run of the package generating and using quartets added only 120 s to the total time. This compares favourably in cost with alternative strategies that involve altering the input options for MULTAN and trying a new phasing path with triplets alone.

Applications in non-centrosymmetric space groups

For the first non-centrosymmetric example we take a coral extract using data provided by J. P. Declerq and G. Germain. This structure is resistant to solution via simple runs of MULTAN 78 and was solved using the MAGIC program (Declerq, Germain & Wolfson, 1979). It therefore represents a good test of the viability of quartet procedures.

All the quartets were generated for the top 100 E magnitudes. All these invariants were included in an otherwise routine run of MULTAN 78 using the dynamic weights of (11). A solution with CFOM =

Table 1. A comparison of ψ_0, NQEST and NQINT

No. of atoms in asymmetric unit	Z	Space group	Best solution						No. of solutions
			ψ_0	Ranked	NQEST	Ranked	NQINT	Ranked	
62	4	$P2_1$	1·19	2	−0·38	1	−	−	48
72	8	$C2$	1·65	10	−0·03	1	−0·28	1	50
27	4	$P2_12_12_1$	1·10	20	−0·84	1	−	−	64
93	2	$P\bar{1}$	1·56	1	−0·83	1	−0·88	1	8
94	2	$P2_1$	1·84	4	−0·13	1	−0·39	2	40
19	4	$P2_12_12_1$	1·28	1	−0·36	1	−	−	12
24	4	$P2_12_12_1$	1·18·	10	−0·10	1	−	−	16

4·00 (the best possible value; NQEST = −0·1) clearly revealed the complete structure. A similar run using triplets and negative quartets was equally successful. A solution having CFOM = 3·91 also revealed the complete molecule. The computer time needed compared favourably with that required for a run of the *MAGIC* system with which it was initially solved.

The next example is another clathrate in which the asymmetric unit has the formula $C_{68}H_{70}O_8S_6$ (unpublished). The space group is $P2_1$ with $Z = 2$, thus there are 82 non-hydrogen atoms in the asymmetric unit. The negative quartets and quintets were generated for the top 100 E magnitudes. A routine run of *MULTAN* 78 incorporating these invariants produced 40 solutions of which only 14 were unique. The solution with CFOM = 4·73 (NQEST = −0·13, NQINT = −0·39) revealed the positions of 32 atoms. The remaining 50 atoms were located *via* Fourier techniques.

The third example also crystallizes in space group $P2_1$. It is an alkaloid with the empirical formula $C_{19}H_{20}N_2O_7$, $Z = 4$. There is a molecule of ethyl acetate associated with each pair of alkaloid molecules so that the asymmetric unit contains 62 non-hydrogen atoms. Application of *MULTAN* 78 using only triplets produced 25 unique solutions. The best of these (CFOM = 2·20) gave an E map that shows almost one complete molecule in the asymmetric unit but with the addition of numerous spurious peaks coupled with an unusable fragment of the second molecule. When the negative quartets were generated for the top 150 E magnitudes and included in both an active and passive way in *MULTAN* 78, the best solution (CFOM = 3·86; NQEST = −0·38) produced an E map in which only the solvent and two methoxy carbon atoms were missing. Obviously the problem could have been solved using triplets alone, but the inclusion of quartets made the process much simpler since only one Fourier synthesis was needed to complete the structure.

The most difficult structure which we have solved using higher invariants is a sesquiterpenoid lactone of formula $C_{25}H_{32}O_{11}$ which crystallizes in space group $C2$ with $Z = 8$ (unpublished). Thus there are 72 non-hydrogen atoms in the asymmetric unit. There is a high degree of pseudosymmetry in this structure in which the two molecules in the asymmetric unit are related by a pseudotranslation of approximately $a/2$. Numerous applications of *MULTAN* 78 were unsuccessful including the incorporation of negative quartets. The problem seemed to be a loss of enantiomorph definition. The *MAGIC* procedure (Declercq, Germain & Woolfson, 1979) can be very useful in these cases, and this program was therefore run between the convergence map and tangent refinement sections of *MULTAN* 78 using triplets alone. Fifteen primary reflections were represented by the elements of {8 11 13 14 15}(*xyz*) and there were also 22 secondary reflections. The 37 reflections of the

combined primary and secondary sets were linked by 14 triplets and a ψ map was calculated from these. The top 50 peaks were refined by a parameter-shift method in which the shifts were restricted to be ≤15°. This was necessary to prevent very large parameter shifts. These solutions were then refined using the tangent formula. Quartets and quintets were used in the final figures of merit. The solution with the highest CFOM (3·22 with a theoretical maximum of 5·00) revealed the positions of 32 atoms in the asymmetric unit. The structure was completed with difficulty using Fourier techniques. The incorporation of quartets and quintets into CFOM was essential for the success of this technique since the correct solution was ranked 10th in ψ_0, 23rd in ABSFOM and 22nd in R_{Karle} and would not normally have been investigated. However, this solution was the *only* one with both a negative NQEST (−0·03) and NQINT (−0·28).

Our final example concerns the lactam $C_{21}H_{16}ClNO$ reported by Collens, Declercq, Germain, Putzeys & van Meerssche (1974). This crystallizes in space group $Pca2_1$ with $Z = 8$, thus there are two molecules in the asymmetric unit. This structure is cited by Lessinger (1976) as an example of a situation in which the correct phase angles are unstable in standard weighted tangent refinement. Negative quartets were generated for the top 100 E magnitudes and incorporated into the tangent refinement. Under these circumstances the tangent formula remains stable and there is no consistent drift of the phase angles from their true values. However, the structure is still difficult to solve. The best E maps generated using quartets in an active mode in attempts to solve this structure in an *a priori* manner revealed only the Cl atom positions.

Comments and the future

Our experience has shown that the use of higher invariants in both active and passive roles in the *MULTAN* program provides a useful alternative when standard runs of the package using triplets alone are unsuccessful. In this respect the technique can be compared with the alternatives *MAGIC* and *YZARC* suggested by Declercq, Germain & Woolfson (1979). In themselves none of these developments is a panacea when difficulties are encountered in direct methods, but they offer different methods of attacking otherwise intractable problems with good prospects of success. The use of negative quartets in particular never seems to degrade the performance of *MULTAN* and often enhances it considerably, especially in symmorphic space groups. Positive quartets pose more difficult problems because of correlation with triplets, and whereas the weighting scheme proposed here is successful it probably does not fully exploit the phase information contained in these invariants, nor does it

solve the difficulties inherent in the correlations between the quartets themselves. These are problems of a theoretical nature that need to be resolved.

A logical extension to these applications of four- and five-phase invariants is to include them in both an active and passive way in the magic integer/ψ map program *MAGIC* and the random phase set/linear equation system *YZARC* where they offer the potential of further enhancing these techniques.

Quartets and quintets can be expensive to calculate in terms of computer time especially when the third-neighbourhood formulae are used. The search for these invariants and their neighbourhoods lends well to the parallel techniques of vector and array processors. The advent of cheaper processors of this kind may well make fully exhaustive searches for quartets and quintets a routine computer operation. We intend to explore this line of development.

We wish to thank the Computing Service of the University of Glasgow for the provision of facilities for carrying out this work.

References

COLLENS, A., DECLERCQ, J. P., GERMAIN, G., PUTZEYS, J. P. & VAN MEERSSCHE, M. (1974). *Cryst. Struct. Commun.* **3**, 119–122.

DECLERCQ, J. P., GERMAIN, G. & WOOLFSON, M. M. (1979). *Acta Cryst.* A**35**, 622–626.

DE TITTA, G. T., EDMONDS, J. W., LANGS, D. A. & HAUPTMAN, H. (1975). *Acta Cryst.* A**31**, 472–479.

DE TITTA, G. T., EDMONDS, J. W., LANGS, D. A. & HAUPTMAN, H. (1976). Am. Crystallogr. Assoc. Summer Meeting, Pennsylvania State Univ., Abstract D3.

FORTIER, S. & HAUPTMAN, H. (1977). *Acta Cryst.* A**33**, 829–833.

FREER, A. A. & GILMORE, C. J. (1980). In preparation.

GIACOVAZZO, C. (1976a). *Acta Cryst.* A**32**, 91–99.

GIACOVAZZO, C. (1976b). *Acta Cryst.* A**32**, 100–104.

GIACOVAZZO, C. (1980). *Acta Cryst.* A**36**, 74–82.

GILMORE, C. J. (1977). *Acta Cryst.* A**33**, 712–716.

GILMORE, C. J., HARDY, A. D. U., MACNICOL, D. D. & WILSON, D. R. (1977). *J. Chem. Soc. Perkin Trans.* 2, pp. 1427–1434.

HAUPTMAN, H. (1977a). *Acta Cryst.* A**33**, 556–564.

HAUPTMAN, H. (1977b). *Acta Cryst.* A**33**, 565–568.

HAUPTMAN, H. (1977c). *Acta Cryst.* A**33**, 553–555.

HEINERMAN, J. J. L. (1978). In *Direct Methods in Crystallography*, edited by H. A. HAUPTMAN, pp. 177–194. Pittsburgh: Polycrystal.

LESSINGER, L. (1976). *Acta Cryst.* A**32**, 538–550.

PUTTEN, N. VAN DER & SCHENK, H. (1977). *Acta Cryst.* A**33**, 856–858.

PUTTEN, N. VAN DER & SCHENK, H. (1979). *Acta Cryst.* A**35**, 381–387.

Crystallographic Literature:
A Bibliometric and Citation Analysis*

BY DONALD T. HAWKINS

Bell Laboratories, Murray Hill, New Jersey 07974, USA

Abstract

The literature of crystallography was studied by bibliometric analysis using the Lockheed DIALOG® on-line information retrieval system and the *Chemical Abstracts* database. In the period 1972 to 1976, the publication of journal articles on crystallography remained approximately constant at 9550 papers per year. Major journals publishing articles on crystallography were identified and ranked; ten journals devote more than half their contents to the subject. Twenty journals account for half the papers on crystallography. Use of Bradford's Law along with ranking by percentage of crystallographic papers showed that there are 22 core journals. The 18 that had citation data available were ranked on an 'influence' basis. The most influential journals, as identified by citation data, are large well known chemistry and physics journals. The citation linkages between the field of crystallography and other closely related fields of

---

* *Editorial note:* A preliminary investigation of the extent to which *Acta Crystallographica* attracts and publishes the most important papers on crystallography was made at the request of the Chairman of the Commission on Journals for presentation at an Open Meeting of the Commission in Warsaw, 11 August 1978. Objective examination, from time to time, of how well a journal serves its community of readers, authors and subscribers may be beneficial both to that community and to its publishers. Interest in the relationships discerned among the various journals publishing crystallographic papers led to subsequent analysis. The present paper, following its subjection to the usual acceptance process, makes the results of this analysis available to the crystallographic community and also provides a basis for future assessment.