

Table 5. Homomorphisms of reducible space groups onto layer and rod groups with respect to Z reductions: trigonal groups

			σ_{ab}	σ_c	\mathfrak{q}
C_3	C_3^4	$R3$	$\hat{p}_{1/3}^3$	$\#_{e/3}^3$	$0, a/3, 2a/3$
C_{3i}	C_{3i}^2	$R\bar{3}$	$\hat{p}_{1/3}^{\bar{3}}$	$\#_{e/3}^{\bar{3}}$	$0, c/6, c/3$
D_3	D_3^2	$R32$	$\hat{p}_{1/3}^3 312$	$\#_{e/3}^3 321$	$\left\{ \begin{array}{l} 0, a/3, 2a/3, c/6, \\ c/3, a/3+c/6, \\ a/3+c/3, 2a/3+c/6, \\ 2a/3+c/3 \end{array} \right.$
C_{3v}	C_{3v}^5	$R3m$	$\hat{p}_{1/3}^3 31m$	$\#_{e/3}^3 3m1$	$\left\{ \begin{array}{l} 0, a/3, 2a/3 \\ 0, a/3, 2a/3, c/6, \\ c/3, a/3+c/6, \\ a/3+c/3, 2a/3+c/6, \\ 2a/3+c/3 \end{array} \right.$
	C_{3v}^6	$R3c$	$\hat{p}_{1/2}^3 31m$	$\#_{e/3}^3 3c1$	
D_{3d}	D_{3d}^5	$R\bar{3}m$	$\hat{p}_{1/3}^{\bar{3}} 312/m$	$\#_{e/3}^{\bar{3}} 3m1$	$\left\{ \begin{array}{l} 0, a/3, 2a/3, c/6, \\ c/3, a/3+c/6, \\ a/3+c/3, 2a/3+c/6, \\ 2a/3+c/3 \end{array} \right.$
	D_{3d}^6	$R\bar{3}c$	$\hat{p}_{1/3}^{\bar{3}} 312/m$	$\#_{e/3}^{\bar{3}} 3c1$	

national Tables for Crystallography as well as in the 1952/1969 editions of *International Tables for X-ray Crystallography*; and the reverse setting used in the first edition of Vol. I of *Internationale Tabellen zur Bestimmung von Kristallstrukturen* (1935). The rotation of a group by $\pi/3$ around the hexagonal axis sends it from one setting to the other. It leaves both homomorphic projections (layer and rod groups) invariant, so that they are the same for both obverse and inverse settings. They are listed in Table 5.

7. Concluding remarks

Let us briefly overview what we have described in the two papers and what still should be done. We have determined homomorphic projections of reducible plane and space groups onto their factor groups with respect to all types of Z reductions and Z decompositions. The projections connected with Z decompositions were used to introduce the nomenclature and symbols of frieze, layer and rod groups compatible with symbols for plane and space groups. The

determination of factor groups is also equivalent to the classification of reducible plane and space groups into pairs of frieze-group classes and into pairs of layer- and rod-group classes, respectively. We must say, however, that this is not the end of the problem up to three dimensions. An exact and complete solution must also involve the problem of origin choices since there are cases where we may list a layer or rod group as a projection of different space groups using the same symbol for this layer and/or rod group, even though these projections have different locations in space, if we accept that the location of the space group is given by its diagram. There is only one remedy for this: we have to fix the Hermann–Mauguin symbols of space groups and of frieze, layer and rod groups to certain standard origin choices and use modified symbols if the groups are shifted in space. The addition to the standard symbol of a shift vector in parentheses would give a simple and clear convention for this purpose.

References

ASCHER, E. & JANNER, A. (1965). *Helv. Phys. Acta*, **38**, 551–572.
 ASCHER, E. & JANNER, A. (1968/1969). *Commun. Math. Phys.* **11**, 138–167.
 BROWN, H., BÜLOW, R., NEUBÜSER, J., WONDRATSCHEK, H. & ZASSENHAUS, H. (1978). *Crystallographic Groups of Four-Dimensional Space*. New York: Wiley.
 CURTIS, W. C. & REINER, I. (1966). *Representation Theory of Finite Groups and Associative Algebras*. New York: Wiley-Interscience.
 HIRSHFELD, F. L. (1968). *Acta Cryst.* **A24**, 301–311.
Internationale Tabellen zur Bestimmung von Kristallstrukturen (1935). Bd 1. Berlin: Borntraeger.
International Tables for Crystallography (1987). Vol. A. Dordrecht: Kluwer Academic Publishers.
International Tables for X-ray Crystallography (1952, 1969). Vol. I. Birmingham: Kynoch Press.
 KOPSKÝ, V. (1988). *Czech. J. Phys.* **B38**, 945–967.
 KOPSKÝ, V. (1989a). *Acta Cryst.* **A45**, 805–815.
 KOPSKÝ, V. (1989b). *Acta Cryst.* **A45**, 815–823.
 KOPSKÝ, V. (1992). *J. Math. Phys.* Submitted.
 KOPSKÝ, V. (1993). *Acta Cryst.* **A49**, 269–280.

Acta Cryst. (1993). **A49**, 287–293

Automatic Determination of Crystal Structures using Karle–Hauptman Matrices

BY R. DE GELDER, R. A. G. DE GRAAFF AND H. SCHENK

X-ray Department, Gorlaeus Laboratories, PO Box 9502, 2300 RA Leyden, The Netherlands

(Received 22 January; accepted 23 June 1992)

Abstract

This paper describes a method for automatic structure determination by the application of Karle–Hauptman

matrices to the phase problem. A new method, the common-minor strategy, is used to combine the information contained in several Karle–Hauptman matrices. Sets of phases large enough to define the

structure are obtained. If the matrices are linked properly, phases contained in a larger array of matrices can be adequately restricted to a common origin. The partial solutions obtained on Fourier transformation are extended to the complete structure in a fully automatic manner. A new algorithm is presented for the maximization of the determinant of a Karle-Hauptman matrix as a function of the phases.

Introduction

In a previous paper, which described a new algorithm for the construction of separate Karle-Hauptman matrices (de Gelder, de Graaff & Schenk, 1990), results were presented for phase sets obtained from single matrices. Phase sets, obtained by the maximization of the determinant as a function of the phases, contained low phase errors but were too small to solve the structure. Extension of these sets *via* tangent refinement, using *FASTAN* from the *MULTAN* system (Declercq, Germain & Woolfson 1975), proved difficult. To extend a phase set without using the tangent formula, Tsoucaris (1970) used the maximum-determinant rule to determine phases of additional columns, starting from Karle-Hauptman matrices containing known phases. Application to protein structures led to interesting results (de Rango, Mauguen & Tsoucaris, 1975; de Rango, Mauguen, Tsoucaris, Dodson, Dodson & Taylor, 1979). However, for small structures it is almost always impossible to build a Karle-Hauptman matrix containing only known reflections.

The generalized maximum-determinant rule, suggested by Tsoucaris (1970) and mathematically proved by Heinerman, Kroon & Krabbendam (1979) seemed to overcome this problem since *ab initio* phase information was no longer needed. Matrices with trial phases obtained by maximization of the determinant could in principle be used as a starting matrix for Tsoucaris's method. The process of maximizing the determinant becomes unstable if the maximum value of the Karle-Hauptman matrix is too small, which is the case for matrices of increasing order. Therefore, this way of extending phase sets would fail for many small structures.

To overcome these problems, a multimatrix method - the concurrent maximization of a number of interdependent matrices - was proposed by de Graaff & Vermin (1982). This method forces a set of reflections to occur in all resulting matrices although the individual matrices are built one after another. Linking different matrices this way might cause the phases to refine to a common origin, possibly resulting in sets of phases large enough to define the structure (at least partly). The method gave satisfactory results in a few cases but needed tangent refinement to optimize and extend the phase sets further.

In the current paper, a new method is described for linking Karle-Hauptman matrices, which gives satisfactory results for all nine structures tested. After constructing a large Karle-Hauptman matrix, by use of the algorithm described in a previous paper (de Gelder *et al.*, 1990), several smaller overlapping matrices can be extracted from the large matrix. The construction procedure, called the common-minor strategy (CMS), will be described in detail. Karle-Hauptman matrices built using this method are all strongly linked. This is illustrated by the quality of the large phase sets calculated. After Fourier transformation, the partial models obtained are of sufficient quality to permit the automatic extension of the map obtained to the complete structure. For this extension, the program *AUTOFOUR* (Kinneking & de Graaff, 1984) was used.

The implementation of the common-minor strategy and the automatic Fourier transformation program *AUTOFOUR* into a set of programs, *CRUNCH*, led to a system capable of the automatic solution of crystal structures. A new method of maximizing determinants was implemented. The results are compared here to the results obtained using the earlier method (de Graaff & Vermin, 1982). The new maximization procedure is described in detail. Both the old and the new procedures gradually change the values of the phases in the matrix to maximize the value of the determinant. Because of this, the maximization methods are only applicable to noncentrosymmetric structures.

Karle-Hauptman matrices

The following abbreviations and notations will be used throughout this paper:

KH	Karle-Hauptman
A	the KH matrix with elements $E(H)$
m	the order of A
B	the inverse of A
a_{ij}	an element of A ($i, j = 1, 2, \dots, m$)
α_{ij}	the phase of a_{ij} ($i, j = 1, 2, \dots, m$)
b_{ij}	an element of B ($i, j = 1, 2, \dots, m$)
β_{ij}	the phase of b_{ij} ($i, j = 1, 2, \dots, m$)
α	a vector of elements $\alpha_1, \alpha_2, \dots, \alpha_n$
β	a vector of elements $\beta_1, \beta_2, \dots, \beta_n$
n	the number of independent reflections in A
H_{ij}	the reciprocal-lattice vector associated with a_{ij} , e.g. $H_j - H_i$
N	the number of atoms in the unit cell
π	a vector of elements π
$\det A$	the value of the determinant of A

Karle-Hauptman matrices can be constructed using ordinary, normalized or unitary structure factors. We use normalized structure factors and thus an element a_{ij} of a Karle-Hauptman matrix A is defined as $a_{ij} = E(H_{ij} - H_{i_1})$, where H_{ij} and H_{i_1} are

the reciprocal-lattice vectors associated with the first-row elements i and j . Of course, all elements in a Karle-Hauptman matrix are uniquely defined by the first-row elements.

Important properties of Karle-Hauptman matrices are:

- (1) A is Hermitian, *i.e.* $a_{ij} = a_{ji}^*$.
- (2) A is positive semidefinite, *i.e.* all eigenvalues are greater than or equal to zero (Goedkoop, 1952).
- (3) $\det A > 0$ if $m \leq N$, $\det A = 0$ if $m > N$ (Goedkoop, 1952; Kitaigorodsky, 1950).
- (4) The maximum-determinant rule (Tsoucaris, 1970) holds: when a given KH matrix A of order m , containing structure factors with known phases, is enlarged to a KH matrix A_1 of order $m+1$ by adding a row and a column containing nonzero elements with unknown phases, the most probable set of phases in the new column maximizes $\det A_1 / \det A$.
- (5) The generalized maximum-determinant rule (Tsoucaris, 1970) holds: for a given KH matrix A the most probable set of phases will maximize $\det A$. Heinerman, Kroon & Krabbendam (1979) have been able to prove this generalized rule if terms of fifth and higher order in the expansion of the determinant are negligible.
- (6) A simple expression exists for the derivative of $\det A$ with respect to the phase of element ij :

$$\delta \det A / \delta \alpha_{ij} = 2|a_{ij}| |b_{ij}| \sin(\beta_{ij} - \alpha_{ij}) \det A$$

$$(i, j = 1, 2, \dots, m) \quad (1)$$

(de Graaff & Vermin, 1982).

For a KH matrix to be useful in *ab initio* phase determination, the main criterion is that the matrix A must have large off-diagonal elements (thus decreasing $\det A$) to increase the selectivity of the local maxima. In practice this means that the average $|E|$ value of the elements in the matrix A must be as large as possible (de Gelder *et al.*, 1990).

The order of a KH matrix must be less than the number of atoms in the unit cell. However, it will be shown that the influence of the order of the matrices on the quality of the phase sets is not drastic.

Construction of interdependent matrices: the common-minor strategy

Determination of a set of phases large enough to define the structure implies the construction of a set of matrices that together contain a sufficient number of independent reflections. However, the matrices must also have a number of reflections in common, to ensure concurrent refinement of the phases leading to a common origin (in real space) for all the matrices used. These two requirements impose conflicting constraints on the construction process.

If each matrix is constructed by bordering a minor of the previous one, a method is obtained that guar-

antees a strong connection between neighbouring matrices. This strategy was proposed by Kinnegeing (1986), who called it the panorama strategy (Fig. 1).

Experience showed that greater overlap between successive matrices improves the stability of the process of phase determination. However, the amount of overlap between different matrices needed to ensure stability of the refinement process proved to be prohibitive. Unrealistically high numbers of matrices are necessary to obtain an adequate set of phases. Even so, the link between the first and the last matrix of the chain proved to be tenuous at best. The conclusion must be that determination of larger sets of phases by the panorama strategy is not possible in practice. An alternative procedure, based on the algorithm proposed in an earlier paper, is given below.

A large Karle-Hauptman matrix A of order L , the overall matrix, is constructed using the procedure proposed by de Gelder *et al.* (1990). The first row of this matrix, excluding $E(000)$, is divided into T equal parts. Smaller matrices are constructed by combining any two of these to form the first row of a Karle-Hauptman matrix with elements related to those found in the large matrix A . Permutation allows $\frac{1}{2}(T^2 - T)$ different matrices of order $2(L-1)/T+1$ to be constructed. Fig. 2 is an illustration of the effects of this procedure. In this case, the first row of a large matrix A is divided into three parts ($T=3$). The matrices obtained by combining the resulting parts are drawn. The different shading illustrates the overlap between the three smaller matrices: the common minors.

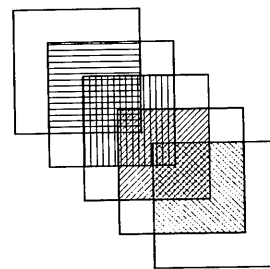


Fig. 1. The panorama strategy.

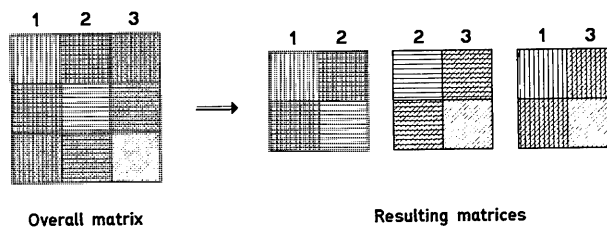


Fig. 2. The common-minor strategy: the construction of three interdependent matrices.

The relation between the order of the overall matrix, L , and the number of small matrices, M , given the order of the small matrices, S , is given by

$$L = [1 + (1 + 8M)^{1/2}](S - 1)/4 + 1. \quad (2)$$

Only values of M for which the term $(1 + 8M)^{1/2}$ is an integer are allowed. Both S and L must be odd and $T(S - 1)/2 (= L - 1)$ must be an integer.

The number of small matrices, M , given the values of L and S , is

$$8M = (4L - S - 3)^2 / (S - 1)^2 - 1. \quad (3)$$

If the values of L and S or M are known, T can be calculated using the simple relations

$$T = 2(L - 1) / (S - 1) \quad (4)$$

or

$$T = [1 + (1 + 8M)^{1/2}] / 2. \quad (5)$$

From Fig. 2, it is clear that in this case all three matrices are connected in pairs having one minor in common. In general, each matrix is connected by a common minor to $2(T - 2)$ other matrices. In a set of M matrices there are $(T^3 - 3T^2 + 2T)/2$ connections of this type, henceforth called first-order links. Each matrix is not linked directly to every other matrix. However, each matrix is connected to all other matrices *via* two first-order links. In the set of M matrices, $2T^4 - 11T^3 + 19T^2 - 10T$ second-order links are present.

The first-order links in the case of $M = 6$ are drawn in Fig. 3. Each dot represents a submatrix formed by the combination of two parts from the first row of the overall matrix A . The meaning of the numbers is as follows. Matrix I/J is constructed by the combination of the I th and J th parts from the first row of the overall matrix.

Clearly, M cannot take on any integer value; only certain integers result in integer values for L and S . The number of matrices one can extract from the overall matrix must be in the sequence 1, 3, 6, 10, 15, 21, 28, ... *etc.*;

$$M_n = \sum_{j=1}^n j \quad (6)$$

i.e.

$$M_n = \frac{1}{2}n(n + 1). \quad (7)$$

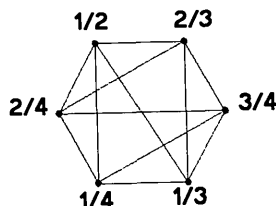


Fig. 3. First-order links for $M = 6$.

Given the value of M_n , n can be calculated from the formula

$$n = -\frac{1}{2} + (\frac{1}{4} + 2M_n)^{1/2}. \quad (8)$$

In practice, one takes a value for M that results in suitable values for S to ensure successful phase refinement in the individual submatrices. The order of the overall matrix (L) depends on the number of independent (strong) reflections needed to solve the structure using automatic Fourier extension.

The distribution of strong reflections in reciprocal space determines the properties of the overall matrix obtained by the algorithm used for the construction. For each structure, different values for L and M will be needed to get a phase set large enough to solve the structure.

Phase refinement

The derivative of $\det A$ with respect to a given phase in the matrix A is given by (1). Near the maximum, the derivative must approach zero, which implies a relation between α and β (the phases in the matrix and in its inverse, respectively). These partial derivatives cannot be exactly zero because of symmetry relations present in A . By use of the function $\tau(\mathbf{r})$, given by Knossow, de Rango, Manguen, Sarrazin & Tsoucaris (1977), which has maxima at the atomic positions of the actual structure, one can conclude that, near the maximum of $\det A$,

$$\beta_{ij} \approx \alpha_{ij} + \pi \quad (i, j = 1, 2, \dots, m). \quad (9)$$

de Graaff & Vermin (1982) proposed an iterative phase refinement based on these relations. The main objection to the scheme is that correlations between the independent phases are not taken into account while the shifts are being calculated. Nevertheless, satisfactory results were obtained.

A much simpler and more straightforward procedure, also based on the relations given above, is the following: instead of changing the phases α in A in the direction suggested by (9), modify the phases β of A^{-1} to fit the set of relations, then calculate the inverse of the modified inverse and use the phases in the resulting matrix as the next estimate of α . The iteration is carried out until convergence. In this way, changes in any phase β_i are reflected in all elements of α . There are a few obvious problems here: if one changes the phases in A^{-1} it may not remain positive definite and if one substitutes the phases obtained by inverting the modified inverse of A into A this may result in the same situation for A .

In practice, these problems are easily solved. Rather than a substitution of the values $\alpha + \pi$ into A^{-1} , the new values of β are calculated from

$$\beta'_{ij} = \tan^{-1} \left\{ \frac{\sin(\alpha_{ij} + \pi) / F + F \sin \beta_{ij}}{\cos(\alpha_{ij} + \pi) / F + F \cos \beta_{ij}} \right\}. \quad (10)$$

F is a damping factor that may be decreased or increased as required. The new values of β_{ij} are close either to $\alpha_{ij} + \pi$ or to their old values, depending on the current value of F . The new estimate of α is substituted into A in a similar way. In both cases, the damping factors are chosen so that all matrices remain positive definite.

Convergence is generally smooth; the radius of convergence seems to be somewhat greater than that of the earlier alternative. In the section discussing the results obtained using the package *CRUNCH*, a comparison is given between the two methods of phase refinement.

Automatic Fourier extension

After the reflections are phased in the matrices constructed using the common-minor strategy, a Fourier summation is calculated by *EXFFT80* from all independent reflections in the matrices. From the E maps obtained, the largest fragment found by *SEARCH80* is used as input to the program *AUTOFOUR*. Owing to the small number of phases used to calculate the E maps, the fragment that is to be extended to the full structure is bound to contain atomic positions that are incorrect, even if the input phases are completely correct. *AUTOFOUR* can cope with this. If the fragment contains enough information about the structure, *AUTOFOUR* is capable of generating the complete model of the structure with very small errors.

CRUNCH: the automatic determination of structures

Construction and CMS linking of matrices, maximization of the determinants, automatic Fourier methods and several auxiliary programs are implemented in the procedure *CRUNCH*. *CRUNCH* automatically determines suitable parameters for the construction and maximization of the matrices, starts the calculation of the E map and feeds the results of the peak search into *AUTOFOUR*. On completion of the extension process – successful or otherwise – the results are checked. If no solution has been found, the process is repeated with different random starting phases put into the maximization cycle *etc.* The files used in *CRUNCH* are compatible with the *MULTAN* system (Declercq, Germain & Woolfson, 1975).

CRUNCH consists of the following steps:

(1) A set of linked Karle-Hauptman matrices is constructed using the common-minor strategy. The overall matrix is constructed using the method described in an earlier paper (de Gelder *et al.*, 1990). The order O of the small matrices is determined by the function $O = \text{MIN}(15, N/4)$, where N is the number of atoms in the cell. The number of small matrices is determined by the minimum number of

strong reflections needed to yield a reasonable E map. This number was set to $2N_a$, where N_a is the number of atoms in the asymmetric unit.

(2) From random initial phases, the product of the determinants of the matrices is maximized as a function of the phases (procedure *DETER*).

(3) An E map is calculated using all reflections present in the matrices (procedure *EXFFT80*).

(4) The strongest peaks present in the E map are located and the largest molecular fragment is sought (procedures *SEARCH* and *FRAGMENT*).

(5) The fragment is fed into the automatic Fourier transform procedure to extend the model and to delete wrong atoms (procedure *AUTOFOUR*).

Steps (2) to (5) are repeated until step (5) converges to a model resulting in an acceptable R_2 value.

Input to *CRUNCH* consists of:

- (i) space group;
- (ii) normalized structure factors;
- (iii) cell parameters;
- (iv) cell contents;
- (v) scale;
- (vi) overall temperature factor.

For the normalization of the structure factors, the program *NORMAL80* is used. The output from *CRUNCH* is the complete model of the structure. Fig. 4 shows a flow diagram of the procedure *CRUNCH*.

Test structures

CRUNCH was tested by trying to solve the phase problem for nine test structures. All the test structures were noncentrosymmetric. Two test structures were heavy-atom structures that were easily solved from the Patterson function. These structures were included in the test to find out whether *CRUNCH* is effective for non-equal-atom structures. The other structures consisted of easy as well as difficult equal-atom structures. For all test structures, real data sets were used and no information about the structures, except the cell contents, was used as input to *CRUNCH*.

The test compounds were:

BP40: $C_{12}H_{18}Cu_2N_6O_{12}S_2$ (van Koningsbruggen, Haasnoot & Reedijk, 1993), $N = 68$, space group $P2_1$, $Z = 2$, $a = 9.973$, $b = 14.710$, $c = 7.403$ Å, $\beta = 96.867^\circ$. Solved by Patterson methods.

DMADMB: $C_{11}H_{14}N_2$ (Gorter, 1988), $N = 52$, space group $P2_12_12_1$, $Z = 4$, $a = 7.801$, $b = 7.800$, $c = 17.016$ Å. Solved by *MULTAN*.

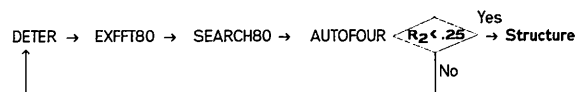


Fig. 4. Schematic overview of the *CRUNCH* procedure.

K2RUO4K: $K_2RuO_3(OH)_2$ (Elout, Haije & Maasant, 1988), $N = 32$, space group $P2_12_12_1$, $Z = 4$, $a = 8.012$, $b = 10.588$, $c = 6.687$ Å. Solved by Patterson methods.

PYROC: $C_{35}H_{46}N_2O_6$ (de Kok & Romers, 1975), $N = 86$, space group $P2_1$, $Z = 2$, $a = 11.28$, $b = 11.27$, $c = 12.537$ Å, $\beta = 93.25^\circ$. Solved by Patterson search methods.

ISOPYR: $C_{35}H_{46}N_2O_6$ (de Kok, Romers & Hoogendorp, 1975), $N = 86$, space group $P2_1$, $Z = 2$, $a = 18.364$, $b = 5.955$, $c = 14.568$ Å, $\beta = 94.03^\circ$. Solved by Patterson search methods.

TOXISTEROL: $C_{34}H_{46}N_2O_6$ (de Kok, Boomsma & Romers, 1976), $N = 84$, space group $P2_1$, $Z = 2$, $a = 8.912$, $b = 7.3136$, $c = 24.889$ Å, $\beta = 98.93^\circ$. Solved by Patterson search methods.

TRIGAL: $C_{38}H_{54}O_{24}$ (Hoogendorp & Romers, 1983), $N = 124$, space group $P2_1$, $Z = 2$, $a = 12.480$, $b = 8.821$, $c = 21.18$ Å, $\beta = 98.46^\circ$. Solved by Patterson search methods.

GLUCOPYR: $C_{20}H_{22}N_2O_{14}$ (Koeners, de Kok, Romers & van Boom, 1980), $N = 144$, space group $P2_12_12_1$, $Z = 4$, $a = 8.16$, $b = 17.05$, $c = 17.35$ Å. Solved by Patterson search methods.

RN001: $C_{21}H_{28}O_3$ (Gorter & Brussee, 1992), $N = 48$, space group $P2_1$, $Z = 2$, $a = 10.646$, $b = 6.462$, $c = 13.522$ Å, $\beta = 95.218^\circ$. Solved by *MULTAN*.

Results and concluding remarks

All test structures were solved by *CRUNCH* using default settings. Tables 1 and 2 show the results found, using the old and the new phase-refinement procedures, respectively. The number of cycles, matrices and orders are given together with the final R and R_2 values, both based on $|E|$ values. The c.p.u. time on an IBM Risc 6000/320 computer is also given. *CRUNCH* proved to be a reliable and fast system for solving the test structures in a fully automatic way. While a considerable number of cycles was needed to solve the test structures TRIGAL and GLUCOPYR, they had not previously been solved by direct methods. Determination of the test structure ISOPYR ended with large values for R and R_2 . In this case, *CRUNCH* did not yield a completely correct model of the structure ISOPYR although the E map contained a large fragment of the molecule. From Tables 1 and 2, it should be concluded that neither refinement procedure is significantly better; the new refinement method is merely a (good) alternative to the old one.

Tables 3 and 4 show results obtained using non-default settings of *CRUNCH* for RN001 and TOXISTEROL, respectively, using the old refinement procedure. From these tables, it may be concluded that the influences of the order and the number of matrices are not dramatic: for all settings, *CRUNCH*

Table 1. Results with test structures using the old phase-refinement method

Compound: name of the compound.
Cycles: number of *CRUNCH* cycles.
Matrices: number of matrices used.
Order: order of the matrices used.
C.p.u.: c.p.u. time needed for *CRUNCH* on an IBM Risc 6000/320 computer (in format hh: mm: ss).

Compound	Space group	Cycles	Matrices	Order	Final R	Final R_2	C.p.u.
BP40	$P2_1$	1	3	15	2.63	9.26	0:57
DMADMB	$P2_12_12_1$	7	3	13	5.56	12.80	9:34
GLUCOPYR	$P2_12_12_1$	17	21	15	3.16	7.60	1:08:24
ISOPYR	$P2_1$	3	10	15	6.93	27.06	8:33
K2RUO4K	$P2_12_12_1$	2	3	7	5.36	14.10	1:08
PYROC	$P2_1$	7	15	15	5.57	12.69	36:13
RN001	$P2_1$	4	10	11	7.08	16.34	6:33
TOXISTEROL	$P2_1$	37	15	15	3.26	8.34	1:52:00
TRIGAL	$P2_1$	40	15	15	3.41	9.86	5:22:54

Table 2. Results with test structures using new phase-refinement method

For explanation of headings see Table 1.

Compound	Space group	Cycles	Matrices	Order	Final R	Final R_2	C.p.u.
BP40	$P2_1$	1	3	15	2.65	9.33	1:00
DMADMB	$P2_12_12_1$	10	3	13	5.54	12.78	13:53
GLUCOPYR	$P2_12_12_1$	18	21	15	4.88	10.94	1:07:48
ISOPYR	$P2_1$	11	10	15	6.74	26.09	27:37
K2RUO4K	$P2_12_12_1$	3	3	7	5.36	14.10	1:22
PYROC	$P2_1$	8	15	15	6.04	13.45	43:54
RN001	$P2_1$	4	10	11	5.19	11.74	7:19
TOXISTEROL	$P2_1$	9	15	15	6.66	14.74	23:35
TRIGAL	$P2_1$	57	15	15	4.97	13.09	7:36:24

Table 3. Results with non-default settings for test structure RN001

For explanation of headings see Table 1.

Space group	Cycles	Matrices	Order	Final R	Final R_2	C.p.u.
$P2_1$	2	10	3	5.01	11.70	3:12
$P2_1$	4	10	5	5.09	11.88	5:30
$P2_1$	1	10	7	4.89	11.64	1:30
$P2_1$	1	10	9	5.01	11.51	1:41
$P2_1$	4	10	11	7.08	16.35	6:33
$P2_1$	6	10	13	5.32	12.96	10:24
$P2_1$	9	10	15	4.81	11.33	16:24
$P2_1$	8	10	17	5.20	11.67	15:30
$P2_1$	6	6	11	6.63	15.33	9:29
$P2_1$	3	15	11	6.69	14.26	4:22
$P2_1$	3	21	11	4.85	11.05	4:31
$P2_1$	4	28	11	6.23	14.23	7:10

Table 4. Results with non-default settings for test structure TOXISTEROL

For explanation of headings see Table 1.

Space group	Cycles	Matrices	Order	Final R	Final R_2	C.p.u.
$P2_1$	95	15	13	3.34	8.48	4:37:18
$P2_1$	37	15	15	3.26	8.34	1:52:01
$P2_1$	52	15	17	3.37	8.53	2:34:03
$P2_1$	33	15	19	3.28	8.23	1:50:35
$P2_1$	17	15	21	3.24	8.22	59:40
$P2_1$	85	10	15	3.36	8.54	4:10:45
$P2_1$	53	21	15	3.39	8.60	2:37:52
$P2_1$	1	28	15	3.25	8.20	3:17

could solve the structures. Clearly, the default settings are not optimal for each individual structure.

References

- DECLERCQ, J.-P., GERMAIN, G. & WOOLFSON, M. M. (1975). *Acta Cryst.* **A31**, 367–372.
- ELOUT, M. O., HAIJE, W. G. & MAASKANT, W. J. A. (1988). *Inorg. Chem.* **27**, 610–614.
- GELDER, R. DE, DE GRAAFF, R. A. G. & SCHENK, H. (1990). *Acta Cryst.* **A46**, 688–692.
- GOEDKOOP, J. A. (1952). *Computing Methods and the Phase Problem in X-ray Crystal Analysis*, edited by R. PEPINSKY, pp. 61–83. The Pennsylvania State College.
- GORTER, S. (1988). *Z. Kristallogr.* **185**, 216.
- GORTER, S. & BRUSSEE, J. (1992). *Acta Cryst.* **C48**, 344–347.
- GRAAFF, R. A. G. DE & VERMIN, W. J. (1982). *Acta Cryst.* **A38**, 464–470.
- HEINERMAN, J. J. L., KROON, J. & KRABBENDAM, H. (1979). *Acta Cryst.* **A35**, 105–107.
- HOOGENDORP, J. & ROMERS, C. (1983). *Carbohydr. Res.* **114**, 169–180.
- KINNEGING, A. J. (1986). Thesis, Univ. of Leiden, The Netherlands.
- KINNEGING, A. J. & DE GRAAFF, R. A. G. (1984). *J. Appl. Cryst.* **17**, 364–366.
- KITAIGORODSKY, A. I. (1950). *X-ray Structure Analysis*, Vol. III, p. 61. Moscow: Gostekhizdat.
- KNOSSOW, M., DE RANGO, C., MAUGUEN, Y., SARRAZIN, M. & TSOUCARIS, G. (1977). *Acta Cryst.* **A33**, 119–125.
- KOENERS, H. J., DE KOK, A. J., ROMERS, C. & VAN BOOM, J. H. (1980). *Recl Trav. Chim. Pays-Bas*, pp. 355–362.
- KOK, A. J. DE, BOOMSMA, F. & ROMERS, C. (1976). *Acta Cryst.* **B32**, 2492–2496.
- KOK, A. J. DE & ROMERS, C. (1975). *Acta Cryst.* **B31**, 1535–1542.
- KOK, A. J. DE, ROMERS, C. & HOOGENDORP, J. (1975). *Acta Cryst.* **B31**, 2818–2823.
- KONINGSBRUGGEN, P. J. VAN, HAASNoot, J. G. & REEDIJK, J. (1993). In preparation.
- RANGO, C. DE, MAUGUEN, Y. & TSOUCARIS, G. (1975). *Acta Cryst.* **A31**, 227–233.
- RANGO, C. DE, MAUGUEN, Y., TSOUCARIS, G., DODSON, G. G., DODSON, E. J. & TAYLOR, D. J. (1979). *J. Chim. Phys. Phys. Chim. Biol.* **76**, 811–812.
- TSOUCARIS, G. (1970). *Acta Cryst.* **A26**, 492–494.

Acta Cryst. (1993). **A49**, 293–300

The Enumeration and Symmetry-Significant Properties of Derivative Lattices. II. Classification by Colour Lattice Group

BY JOHN S. RUTHERFORD

Department of Chemistry, University of Transkei, Private Bag X1, Umtata, Transkei, South Africa

(Received 3 January 1992; accepted 15 July 1992)

Abstract

Dirichlet-series generating functions may be constructed to enumerate the number of colour lattice groups of any order in the triclinic case. Appropriate factorization of the previously known lattice-enumerating functions gives the number of derivative lattices belonging to each of these lattice groups. These numbers are tabulated for all indices up to 20. Based on these Dirichlet functions, asymptotic estimates of the average values of the corresponding arithmetic functions may be made; these are 1.977 for the three-dimensional colour lattice groups of order n and $1.823gh^2$ for the derivative lattices having group structure $C_{fgh} \otimes C_{fg} \otimes C_f$. Such estimates can also be made for the relative abundance of groups with different numbers of cycles in their structure; a single-cycle structure occurs for roughly 92% of all derivative lattices. A similar argument shows that, in over 98% of cases, one properly chosen co-opted term suffices to ensure primitivity in direct methods.

Introduction

The rapidly expanding field of mathematical chemistry, that is the application of graph theory and

combinatorics to chemistry, has had considerable impact on organic chemistry and even in inorganic chemistry has developed enough to warrant a recent major review (King, 1992). Despite this, there have been few attempts to apply such approaches to solid-state and crystal chemistry. This series of papers, together with some parallel work on graph-theory approaches to the bond-valence distribution in solids (Rutherford, 1990, 1992a), represents an attempt to redress this situation and explore the potential that mathematical chemistry holds for the enrichment of crystallography.

One important concept of combinatorics is the generating function, where the number of distinct objects with a given property is simply the coefficient of one term in the expansion of that function. The application of power-series generating functions to isomer-counting problems in chemistry derives from Cayley (1874). Their advantages, besides elegance and compactness, lie in their usefulness in deriving statistical information on, and asymptotic estimates of, the number of isomers (or other geometric objects) involved in the enumeration.

Derivative lattices (Billiet & Bertaut, 1983) arise in practice both as real lattices (commensurate super-