DECLERCQ, J. P., PIRET, P. & VAN MEERSSCHE, M. (1971). *Acta Cryst.* B**27**, 1276.
DUAX, W. & HAUPTMAN, H. (1972). *Acta Cryst.* B**28**, 2912–2916.
EDMONDS, J. W. & DUAX, W. L. (1974*a*). *Prostaglandins*, **5**, 275–281.
EDMONDS, J. W. & DUAX, W. L. (1974*b*). Abstracts. Am. Crystallogr. Assoc., Spring Meeting, p. 34.
GERMAIN, G. (1978). Private communication.
GERMAIN, G., MAIN, P. & WOOLFSON, M. M. (1970). *Acta Cryst.* B**26**, 274–285.
GERMAIN, G., MAIN, P. & WOOLFSON, M. M. (1971). *Acta Cryst.* A**27**, 368–376.
GILMORE, C. J. (1977). *Acta Cryst.* A**33**, 712–716.
HARK, TH. E. M. VAN DEN, PRICK, P. & BEURSKENS, P. T. (1976). *Acta Cryst.* A**32**, 816–821.
HAUPTMAN, H. (1972). *Crystal Structure Determination: The Role of the Cosine Seminvariants*, pp. 232–236. New York: Plenum.
HAUPTMAN, H., FISHER, J., HANCOCK, H. & NORTON, D. (1969). *Acta Cryst.* B**25**, 811–814.
HULL, S. E. (1978). *Acta Cryst.* A**34**, 38–42.
HULL, S. E., LEBAN, I., MAIN, P., WHITE, P. S. & WOOLFSON, M. M. (1976). *Acta Cryst.* B**32**, 2374–2381.
KARLE, J. (1968). *Acta Cryst.* B**24**, 182–186.
KARLE, J. (1976). *Crystallographic Computing Techniques*, edited by F. R. AHMED, pp. 155–164. Copenhagen: Munksgaard.
KARLE, J. & HAUPTMAN, H. (1956). *Acta Cryst.* **9**, 635–651.
KARLE, J. & KARLE, I. L. (1966). *Acta Cryst.* **21**, 849–859.
LESSINGER, L. (1976). *Acta Cryst.* A**32**, 538–550.
MAIN, P. (1977). Private communication.
MAIN, P. (1978). *MULTAN 78. A Program System for the Automatic Solution of Crystal Structures from X-ray Diffraction Data*. Univ. of York, England.
SAYRE, D. (1952). *Acta Cryst.* **5**, 60–65.
SCHENK, H. (1972). *Acta Cryst.* A**28**, 412–422.
SHELDRICK, G. M. (1976). *SHELX*. Program for Crystal Structure Determination. Univ. of Cambridge, England.
SIM, G. A. (1959). *Acta Cryst.* **12**, 813–815.
SIM, G. A. (1961). *Computing Methods and the Phase Problem in X-Ray Crystal Analysis*, edited by R. PEPINSKY, J. M. ROBERTSON & J. C. SPEAKMAN, pp. 227–235. Oxford: Pergamon Press.

Acta Cryst. (1978). A**34**, 870–883

On the Application of Phase Relationships to Complex Structures. XV. Magic Determinants

BY D. J. TAYLOR, M. M. WOOLFSON AND P. MAIN

*Department of Physics, University of York, Heslington, York YO*1 5*DD, England*

(*Received* 15 *March* 1978; *accepted* 2 *May* 1978)

A magic determinant is a Karle–Hauptman determinant in which the phases of reflexions forming the elements are expressed in magic-integer representation. By means of appropriate algorithms, which are described, magic determinants may be found, up to order 30, which have a high content of large E's. The value of the magic determinant depends on the values of the independent variables (usually two) in terms of which the magic-integer-represented phases are defined. Maxima are sought in a map of magic-determinant values and these give trial values for the constituent phases. These phases may be refined by maximizing one or more of the largest eigenvalues of the matrix of elements. Applications to five different structures are described and the usefulness and limitations of the magic-determinant concept are discussed. It is concluded that Karle–Hauptman determinants, chosen in the way described, provide a very discriminating figure of merit. However, the structure factors they contain usually provide a rather poor base for subsequent phase development and this limits the usefulness of the magic-determinant approach.

Introduction

Phase-determining methods based on magic integers were described in papers VII and VIII of this series (White & Woolfson, 1975; Declercq, Germain & Woolfson, 1975). Further developments in the use of magic integers were described in paper X where a general overview of the *MAGLIN* system was presented (Woolfson, 1977). A theory of magic integers and a critical examination of the use of magic integers in phase determination were given by Main (1977, 1978) in papers XI and XII respectively. This paper describes in detail the part of *MAGLIN* which makes use of the properties of Karle–Hauptman determinants. Magic integers are used to represent the phases of determinant elements creating the so-called 'magic determinants'. The discriminating power of the Karle–Hauptman determinant, combined with magic-integer representation, enables large starting sets of reflexions to be phased for use in other phase-development procedures.

Karle–Hauptman determinants

Karle & Hauptman (1950) showed that for a positive electron density the determinant:

$$D_m = \begin{vmatrix} E_0 & E_{\mathbf{h}_1} & E_{\mathbf{h}_2} & \cdots & E_{\mathbf{h}_{m-1}} \\ E_{\bar{\mathbf{h}}_1} & E_0 & E_{\mathbf{h}_2-\mathbf{h}_1} & \cdots & E_{\mathbf{h}_{m-1}-\mathbf{h}_1} \\ E_{\bar{\mathbf{h}}_2} & E_{\mathbf{h}_1-\mathbf{h}_2} & E_0 & \cdots & E_{\mathbf{h}_{m-1}-\mathbf{h}_2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ E_{\bar{\mathbf{h}}_{m-1}} & E_{\mathbf{h}_1-\mathbf{h}_{m-1}} & E_{h_2-\mathbf{h}_{m-1}} & \cdots & E_0 \end{vmatrix} \geq 0,$$

where $\mathbf{h}_1, \mathbf{h}_2 \ldots \mathbf{h}_{m-1}$ correspond to different indices but may be symmetry-related reflexions. Since $E_{\bar{\mathbf{h}}} = E^*_{\mathbf{h}}$ (* = complex conjugate), the array of structure factors is an Hermitian matrix and the value of the determinant is therefore real. $U_{\mathbf{h}}$ may replace $E_{\mathbf{h}}$ in the determinant with U_0 ($= 1$) in place of E_0 ($\simeq N^{1/2}$ where N is the number of atoms in the unit cell). Karle & Hauptman (1950) and Goedkoop (1958) showed that for $m > N$, $D_m = 0$.

If each column of the determinant is multiplied by the complex conjugate of the leading element and the same is done for each row, then each element of the determinant is converted into a structure invariant of the form $E_{\mathbf{h}_p}E_{-\mathbf{h}_q}E_{\mathbf{h}_q-\mathbf{h}_p}$. The determinant as a whole has been multiplied by $E_0^2 \, \Pi_{i=1}^m |E_{\mathbf{h}_i}|^2$ and so it is clear that D_m is itself a structure invariant (Kitaigorodsky, 1961). Since the maximum value of the determinant of an Hermitian matrix occurs when the phase of each element is zero, it follows that if the magnitudes of all elements are large, the set of phases that makes all the triple-phase invariants close to zero will maximize the determinant.

From the maximum-determinant rule (de Rango, 1969; Tsoucaris, 1970) it may be inferred that for a determinant of order less than N, the most probable set of phases for the included E's, assuming they are large in magnitude, is that which gives the determinant a maximum value. The present work is based on the properties of Karle–Hauptman determinants where magic integers are used to represent the phases of determinant elements.

Selection of determinants

Requirements

Suitable determinants should contain reflexions of large $|E_{\mathbf{h}}|$ that occur repeatedly. Messager & Tsoucaris (1972) have shown that for a determinant of order 4 the larger the $|E_{\mathbf{h}}|$'s in the top row, the higher will be $\overline{|E_{\mathbf{h}}|}$ over the whole determinant. These findings may be extrapolated to higher-order determinants but no proof exists in such cases. Two approaches have been tried to obtain suitable determinants.

Method 1

(*a*) A convergence map is produced for the structure by *MULTAN*.

(*b*) A top row is selected to contain origin-defining and enantiomorph-fixing reflexions and some large $|E_{\mathbf{h}}|$'s from near the foot of the convergence map.

(*c*) The content of the determinant is then examined for the number of large $|E_{\mathbf{h}}|$'s contained.

Stages (*b*) and (*c*) are repeated with various selections of large $|E_{\mathbf{h}}|$'s until a satisfactory matrix is obtained that contains as many large $|E_{\mathbf{h}}|$'s as possible occurring repeatedly.

Method 2

The original algorithm for this procedure was developed and programmed by Main (1975). Some modifications have been made to the original process and the steps of the current version are described below.

(*a*) A large Karle–Hauptman matrix is constructed with the top row chosen from all available large $|E_{\mathbf{h}}|$'s using all symmetry-related reflexions including Friedel opposites. (The central region of reciprocal space is used with the aim of reducing the occurrence of reflexions outside the data set which are given zero magnitudes.)

(*b*) The row and column with the smallest $\sum |E_{\mathbf{h}}|$ are eliminated.

(*c*) Step (*b*) is repeated until $\overline{|E_{\mathbf{h}}|}$ over the matrix is greater than some preselected minimum value.

(*d*) A second large matrix is then constructed as in (*a*) but with the top row taken from reflexions which appear frequently in the (*c*) matrix.

(*e*) Steps (*b*) and (*c*) are repeated until a matrix of desired order (often 100) is obtained.

(*f*) A figure of merit is computed for each column. Two types have been used successfully: either $\overline{|E_{\mathbf{h}}|}$ or $\overline{||E_{\mathbf{h}}|^2 - 1|}$ divided by the number of unique strong reflexions.

(*g*) The column and corresponding row with the lowest figure of merit are eliminated unless the column contains less than or equal to k known strong reflexions

(k is usually 1 or 2) in the reduced matrix, in which case it becomes a column in the starting matrix.

(h) The procedure is repeated from step (f) until only the starting matrix remains. The elimination procedure ensures that the origin and enantiomorph are still defined by the strong reflexions contained in the starting matrix.

It has become clear that the selection of suitable matrices by this method for use with magic determinants or in other phase-development procedures is very structure dependent. In particular, any bias or slight pseudosymmetry in the data is amplified by the selection procedure and easily leads to the exclusion of whole classes of reflexions from the order 100 matrix of step (e). It has also been found that it is difficult to obtain small starting matrices without both extensive adjustment to the program control parameters and repeated trials. No general principle has emerged that works in most cases.

Nevertheless, suitable large matrices for several structures have been selected automatically by this method and the results obtained are discussed in a later section.

Representation of phases by magic integers

A typical element of an Hermitian matrix $|E_{\mathbf{h}}| \exp(i\varphi_{\mathbf{h}})$ can be expressed in magic-integer form as $|E_{\mathbf{h}}| \exp(2\pi i n x)$ where $0 \le x < 1$ and n is one integer of a magic-integer sequence. For example, the phases of the six independent reflexions in a D_4 determinant may be represented (in cycles) by

$$\begin{vmatrix} 0 & 3x & 3y & 4x \\ -3x & 0 & 5x & 4y \\ -3y & -5x & 0 & 5y \\ -4x & -4y & -5y & 0 \end{vmatrix}.$$

The value of the determinant with phases so expressed is $D_4(x,y)$. To find the most probable set of phases in a determinant the following procedure is used.

(1) Independent unknown phases in a D_m determinant are represented in magic-integer notation with origin and enantiomorph phases given fixed values.

(2) For x and y as the independent variables D_m is computed for a grid of points covering the x,y domain. Intervals of 1/(4 × maximum integer) have been found sufficient.

(3) Values of (x,y) corresponding to peaks in the map are recorded.

(4) Each peak is translated to the phases that it represents, giving a number of possible phase sets.

Step 1

The magic-integer sequences used to represent phases have been optimized for a uniform distribution of errors applying the rules given by Main (1977) in paper XI of this series. The grid interval used at present is $\frac{1}{100}$ in x and y and this puts an upper limit of about 25 on the highest integer in a magic-integer sequence. This in turn restricts the maximum length of an 'efficient' sequence to about 12. The r.m.s. errors for the magic-integer sequences used in this work are shown in Table 1.

When the number of phases to which magic integers are to be assigned exceeds 24, magic-integer symbols are allocated to the reflexions using the primary–secondary concept developed in paper VIII of this series (Declercq, Germain & Woolfson, 1975). An algorithm has been developed to select which reflexions it is most useful to place in the primary set. The procedure has been adapted from the algorithm described by Main (1978) in paper XII of this series for selecting primary sets that have a large number of associated secondary reflexions.

In this adaptation, the primary set is chosen such that all reflexions in the determinant other than primary reflexions can be defined by a triple-phase invariant

Table 1. *Magic-integer sequences used to represent phases in magic determinants*

The r.m.s. error was determined using 500 trials and a Monte Carlo approach.

Length	Sequence	Lower bound (°)	r.m.s. error (°)
1	1	0·0	0·0
2	2 3	20·4	20·4
3	4 6 7	26·2	26·6
4	8 12 14 15	29·5	29·9
5	12 17 20 22 23	34·4	35·9
6	13 18 21 23 24 25	40·5	41·6
7	13 17 20 22 23 24 25	45·9	47·9
8	15 19 22 24 26 27 28 29	49·2	50·8
9	14 17 19 21 23 24 25 26 27	53·3	55·2
10	15 18 20 22 24 25 26 27 28 29	55·8	57·6
11	16 19 21 23 25 26 27 28 29 30 31	57·9	59·7
12	15 17 19 21 22 23 24 25 26 27 28 29	60·5	62·0

involving two members of the primary set. With the above restriction the primary set is selected as in Main's procedure. As Karle–Hauptman determinants have a proportion of reflexions with small magnitudes, care is taken to avoid the inclusion of such reflexions in the primary set.

Once the primary set has been selected, secondaries are then defined by the strongest available triple-phase invariant but with the following two restrictions. Firstly, no secondary reflexion may be defined by two origin- and enantiomorph-defining reflexions which have fixed phase values. Secondly, wherever possible, different pairs of primaries are used to define the secondaries. These two restrictions allow maximum variability to the phase representations for different secondaries in the determinant.

The dominant variable term in the values of the Karle–Hauptman determinants of low order used in this work arises from the triple-phase invariants involving the large magnitudes contained within the determinant. Quartets and higher-order invariants do contribute to the value of the determinant but are reduced by a factor $|E_{\mathbf{h}}|/|E_0|$ for every additional phase contained in the invariant.

It is therefore essential to keep the size of the magic integers generated for the triple-phase invariants as small as possible. To achieve this, a correlation procedure is employed that takes account of the magnitude and sign of all the interactions of the primary reflexions with each other in these invariants. A triple-phase invariant of the form

$$\varphi_1 + \varphi_2 + \varphi_3 + b \simeq 0$$

is represented in the extreme case (three secondary phases) in terms of primaries alone by the equation

$$\pm P_1 \pm P_2 \pm P_3 \pm P_4 \pm P_5 \pm P_6 + b' \simeq 0$$

where

$$\varphi_1 = \pm P_1 \pm P_2 + b_1$$
$$\varphi_2 = \pm P_3 \pm P_4 + b_2$$
$$\varphi_3 = \pm P_5 \pm P_6 + b_3$$
$$b' = b + b_1 + b_2 + b_3.$$

In this representation an 'interaction' is defined by the presence of two primaries in a triple-phase invariant and is described as positive if the two primaries have the same sign. Clearly, positive interactions between primaries represented in magic-integer notation by the same variable will generate large integers and this type of interaction must therefore be avoided. This can be done by separating the primaries into two sets and by reversing the signs of some of the magic integers within a set. The following procedure has been adopted to minimize positive interactions and hence to keep the integers defining the triple-phase invariants to a minimum.

(1) Those triple-phase invariants not defining secondary reflexions are searched for interactions between the primary reflexions, the accumulated interactions being displayed in the form of a matrix (see Table 2). The numbers relate to the detailed example for lithocholic acid (Arora, Germain & Declercq, 1976) given later in this paper. The upper triangle contains positive interactions and the lower triangle negative interactions for later use. Each interaction is given a weight proportional to the κ value of the invariant in which it occurs, where

$$\kappa = \frac{2}{E_0}\left|E_{\mathbf{h}}E_{\mathbf{h}'}E_{\bar{\mathbf{h}}-\mathbf{h}'}\right|.$$

(2) Positive and negative interactions are then added together (see Table 3) and the primary reflexions divided into two sets so as to minimize the total interaction between primaries in the same set. This is carried

Table 2. *Matrix of the interactions between primary reflexions in triple-phase invariants*

Positive interactions

	P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8	P_9	P_{10}	P_{11}	P_{12}	P_{13}
P_1		60	0	18	0	0	0	7	0	0	9	2	0
P_2	0		17	0	12	0	3	0	0	13	2	2	10
P_3	65	70		33	0	0	0	0	6	30	27	35	1
P_4	0	41	6		0	7	0	20	15	12	0	8	8
P_5	22	0	49	30		0	0	6	0	0	0	0	0
P_6	0	31	42	22	0	(4)	12	0	0	7	3	4	0
P_7	2	0	8	47	17	18		0	0	0	0	4	0
P_8	0	42	32	0	0	14	0		0	0	11	0	0
P_9	0	0	5	15	6	0	15	28		0	0	0	0
P_{10}	0	0	0	0	0	0	0	13	5		0	0	0
P_{11}	0	26	1	30	4	0	17	0	0	0		29	0
P_{12}	0	0	0	0	0	0	0	0	0	6	0		6
P_{13}	1	0	0	0	0	0	8	11	14	6	28	0	

Negative interactions

Table 3. *Matrix of combined interactions of primary reflexions in the triple-phase invariants*

	P_1	P_2	P_3	P_4	P_5	P_6	P_7	P_8	P_9	P_{10}	P_{11}	P_{12}	P_{13}
P_1		60	65	18	22	0	2	7	0	0	9	2	1
P_2			87	41	12	31	3	42	0	13	28	2	10
P_3				39	49	42	8	32	11	30	28	35	1
P_4					30	29	47	20	30	12	30	8	8
P_5						0	17	6	6	0	4	0	0
P_6							30	14	0	7	3	4	0
P_7								0	15	0	17	4	8
P_8									28	13	11	0	11
P_9										5	0	0	14
P_{10}											0	6	6
P_{11}												29	28
P_{12}													6
P_{13}													

out using the following algorithm. (*a*) The largest total interaction between two primaries, P_2–P_3, is found and the primaries involved assigned to different sets (see the ringed entry in Table 3). (*b*) The interactions of these two primaries with those still unassigned are shown in columns (I*a*) and (II*a*) of Table 4. Columns (III*a*) and (III*b*) show the magnitudes of the differences between columns (I*a*) and (II*a*) for the unassigned primaries. The values in columns (III*a*) and (III*b*) indicate the advantage to be gained by assigning a particular primary to the same set as P_2 or P_3. It is clear in this case that the best selection is to assign P_5 to the set containing P_2 (see $\underline{37}$ in column III*a*). (*c*) The interactions of this chosen primary, P_5, with those remaining unassigned are recorded and added to the list of interactions for the unassigned primaries in column (I*b*). (*d*) The new differences in columns (III*c*) and (III*d*) indicate which primary should next be selected and so on. Further primaries are selected as above ensuring that the two sets of primaries develop evenly until all the primaries have been assigned. The procedure is shown in detail in Table 4.

(3) All possible pair swaps between primaries in the two groups are investigated to determine whether the remaining interactions can be further reduced.

(4) The signs of primaries within a set are then permuted until negative interactions are maximized. The procedure employed is shown in Table 5. All interactions between members of the same primary set are displayed as the two matrices at the head of the table. The sum of positive and negative interactions for each primary with all the other members of its set are then computed and recorded as the first two rows of numbers, (I*a*) and (I*b*). The differences between these two sums for each primary are shown in row (I*c*). To maximize negative interactions, a primary is chosen where the positive interactions exceed the negative ones by the greatest amount (see the ringed entries for P_{12} and P_4 in row I*c*). The sign of this primary is then reversed and the sums of positive and negative interactions for the primaries re-computed and recorded as in rows (II*a*) and (II*b*). Further primaries have their signs reversed in like manner until the sums of negative interactions exceed the sums of positive interactions for all primaries in the two sets (see Table 5).

(5) Finally, primary reflexions with the largest total interactions remaining, including at this stage self interactions in $\sum_1$-like terms, are assigned the smallest integers in the given sequences (see the last part of Table 5).

This procedure keeps the weighted r.m.s. coefficient for a triple-phase invariant approximately equal to the r.m.s. average integer of the magic-integer sequence

Table 4. *The procedure dividing the primary list of reflexions into two sets so as to minimize the total interaction between primaries in triple-phase invariants*

	Columns (I) First set of primaries						Columns (II) Second set of primaries						Columns (III) These values indicate the net interaction saved by placing an unassigned primary into a particular set																					
	(Table of interactions)																																	
	a	*b*					*a*	*b*					*a*	*b*	*c*	*d*																		
															P_5		P_5 P_4		P_{12} P_5 P_4		$P_{12}P_{13}$ P_5 P_4		P_9 $P_{12}P_{13}$ P_5 P_4		P_9 P_8 $P_{12}P_{13}$ P_5 P_4		P_6 P_9 P_8 $P_{12}P_{13}$ P_5 P_4		P_6 P_7 P_9 P_8 $P_{12}P_{13}$ P_5 P_4		P_{11} P_6 P_7 P_9 P_8 $P_{12}P_{13}$ P_5 P_4		$P_{11}P_1$ P_6 P_7 P_9 P_8 $P_{12}P_{13}$ P_5 P_4	
	P_2	P_5	P_{12}	P_9	P_6	P_{11}	P_3	P_4	P_{13}	P_8	P_7	P_1	P_2	P_3	P_2	P_3	P_2	P_3	P_2	P_3	P_2	P_3	P_2	P_3	P_2	P_3	P_2	P_3	P_2	P_3	P_2	P_3	P_2	P_3
Unassigned primaries																																		
P_1	60	22	2	0	0	9	65	18	1	7	2		5			17	1			1	0	0	0	0	7		7		9		0	$\underline{0}$		
P_4	41	30					39							2		$\underline{32}$																		
P_5	12						49						$\underline{37}$																					
P_6	31	0	4	0			42	29	0	14			11		11		40		36		36		35		$\underline{50}$									
P_7	3	17	4	15	30		8	47	8	0			5			12	35		31		39		24		24			$\underline{6}$						
P_8	42	6	0	28			32	20	11					10		16		4	4		15			$\underline{13}$										
P_9	0	6	0				11	30	14				11		5		35		35		$\underline{49}$													
P_{10}	13	0	6	5	7	0	30	12	6	13	0	0	17		17		29		23		29		24		37		30		30		30		$\underline{30}$	
P_{11}	28	4	29	0	3		28	30	28	11	17		0	0		4	26			3	25		25		36		33		$\underline{50}$					
P_{12}	2	0					35	8					33		33		$\underline{41}$																	
P_{13}	10	0	6				1	8						9		9	1			$\underline{7}$														

Table 5. *Permutation of the signs of primary reflexions to maximize negative interactions; assignment of magic integers to the primary reflexions*

Negative \ Positive	P_2	P_5	P_6	P_9	P_{10}	P_{11}	P_{12}
P_2		12	0	0	13	2	2
P_5	0		0	0	0	0	0
P_6	31	0		0	7	3	4
P_9	0	6	0		0	0	0
P_{10}	0	0	0	5		0	0
P_{11}	26	4	0	0	0		29
P_{12}	0	0	0	0	6	0	

Negative \ Positive	P_1	P_3	P_4	P_7	P_8	P_{13}
P_1		0	18	0	7	0
P_3	65		33	0	0	1
P_4	0	6		0	20	8
P_7	2	8	47		0	0
P_8	0	32	0	0		0
P_{13}	1	0	0	8	11	

	P_2	P_5	P_6	P_9	P_{10}	P_{11}	P_{12}
(I*a*) $\Sigma(+)$	29	12	14	0	20	34	35
(I*b*) $\Sigma(-)$	57	10	31	11	11	30	6
(I*c*) $\Sigma(+)\ \Sigma(-)$	−28	2	−17	−11	9	4	㉙
Sign of P_{12} is reversed.							
(II*a*) $\Sigma(+)$	27	12	10	0	26	5	6
(II*b*) $\Sigma(-)$	59	10	35	11	5	59	35
(II*c*) $\Sigma(+)-\Sigma(-)$	−32	2	−25	−11	21	−54	−29
Sign of P_{10} is reversed.							
$\Sigma(+)$	14	12	3	5	5	5	0
$\Sigma(-)$	72	10	42	6	26	59	41
$\Sigma(+)-\Sigma(-)$	−58	2	−39	−1	−21	−54	−41
Sign of P_5 is reversed.							
$\Sigma(+)$	2	10	3	11	5	9	0
$\Sigma(-)$	84	12	42	0	26	55	41
$\Sigma(+)-\Sigma(-)$	−82	−2	−39	11	−21	−46	−41
Sign of P_9 is reversed.							
$\Sigma(+)$	2	4	3	0	0	9	0
$\Sigma(-)$	84	18	42	11	31	55	41
$\Sigma(+)-\Sigma(-)$	−82	−14	−39	−11	−31	−46	−41

	P_1	P_3	P_4	P_7	P_8	P_{13}
$\Sigma(+)$	25	34	79	0	27	9
$\Sigma(-)$	68	111	53	65	43	20
$\Sigma(+)-\Sigma(-)$	−43	−77	㉖	−65	−16	−11
Sign of P_4 is reversed.						
$\Sigma(+)$	7	7	53	47	7	1
$\Sigma(-)$	86	138	79	18	63	28
$\Sigma(+)-\Sigma(-)$	−79	−131	−26	29	−56	−27
Sign of P_7 is reversed.						
$\Sigma(+)$	9	15	6	18	7	9
$\Sigma(-)$	84	130	126	47	63	20
$\Sigma(+)-\Sigma(-)$	−75	−115	−120	−29	−56	−11

(*x*)

	P_2	$-P_5$	P_6	$-P_9$	$-P_{10}$	P_{11}	$-P_{12}$
$\Sigma(+)$	2	4	3	0	0	9	0
$\Sigma(-)$	84	18	42	11	31	55	41
Σ_1	0	0	4	0	0	0	0
Total	86	22	49	11	31	64	41
Symbol	13*x*	24*x*	20*x*	25*x*	23*x*	17*x*	22*x*

(*y*)

	P_1	P_3	$-P_4$	$-P_7$	P_8	P_{13}
$\Sigma(+)$	9	15	6	18	7	9
$\Sigma(-)$	84	130	126	47	63	20
Σ_1	0	0	0	0	0	0
Total	93	145	132	65	70	29
Symbol	21*y*	13*y*	18*y*	24*y*	23*y*	25*y*

used. Given a list of the reflexions contained in the determinant and selected origin-defining and enantiomorph-fixing phases, a suitable magic-integer representation is devised automatically by a computer program incorporating the procedures described above.

Step 2

The determinant as a function of x and y is then computed at intervals 0·01 from $0 \le (x,y) < 1$. For determinants of order 20 or less it is possible to reduce the number of points computed by the following technique. Initially a grid of only 0·02 in x and y is used. This is followed by a local search about points with high positive values of the determinant at intervals of 0·01 in x and y. This reduces the number of determinants evaluated from 10 000 to approximately 3500. For determinants of larger order the best results have been obtained when the full grid has been computed at intervals of 0·01, as a better resolution is needed to detect the peaks.

The determinants are evaluated by a computer algorithm that performs a Cholesky decomposition of the matrix and thus makes full use of the inherent

Hermitian symmetry. The time required for computing a determinant is proportional to the cube of the order and this is the most time-consuming aspect of the whole procedure. On a DEC System 10 machine about 0·2 s is needed to evaluate one determinant of order 30; therefore 35 minutes of computer time are required to evaluate an order 30 magic determinant over a grid of 10 000 points.

Steps 3 *and* 4

The peaks are examined in order of descending magnitude and an attempt is made to centre the peak using a nine-point Gaussian ellipsoid approximation around the highest grid value (Dawson, 1961). If a larger interpolated magnitude than the grid point is obtained then this new peak position is used in step 4 and in the subsequent refinement of phase sets. If not, then the peak position is taken as that of the grid point.

In step 4 the highest peaks – usually between 20 and 30 in number – are stored and each peak in turn is subjected to the next stage of the process – refinement of the phase set that it represents.

Refinement of phases

Phase sets are refined by a method due to Main (1975). The technique is based on matrix algebra and simultaneously modifies all phases in the determinants by maximizing the largest eigenvalues. For magic determinants the top two eigenvalues λ_1, λ_2 are always maximized but further eigenvalues are included if $\lambda_n > 0{\cdot}5\lambda_1$. Initially only λ_1 is maximized, iteration being continued until the change in magnitude and phase in any component in the corresponding eigenvector is below a set limit. The matrix is then reconstructed from this eigenvector and its corresponding eigenvalue [see equations (34) and (35), Main, 1975]. The reconstructed matrix is then examined for the consistency of the phase indications for reflexions that occur repeatedly. A calculated magnitude and phase for such a reflexion is obtained by vector summation over all appearances in the determinant:

$$E_{\mathbf{h}} = \left(\sum_{\text{matrix}} E_{\mathbf{h}}\right) \Big/ (\text{number of appearances}). \qquad (1)$$

It therefore follows that where conflicting estimates for a particular phase are obtained the calculated magnitude is considerably reduced. The phases for special reflexions are projected on to the appropriate direction and the reduced calculated magnitude is noted.

The observed magnitudes and those calculated from the reconstructed matrix are then placed on the same scale using a scale factor, s, computed using the formula:

$$s = \frac{\sum\limits_{\text{matrix}} |E_{\text{obs}}||E_{\text{calc}}|}{\sum\limits_{\text{matrix}} |E_{\text{calc}}|^2}. \qquad (2)$$

A residual is also computed using the formula:

$$R = \frac{\sum ||E_{\text{obs}}| - s|E_{\text{calc}}||}{\sum |E_{\text{obs}}|}. \qquad (3)$$

A weighting scheme given by

$$w = 3r^2 - 2r^3 \qquad (4)$$

is employed for the reflexions where $r = s|E_{\text{calc}}|/|E_{\text{obs}}|$. If r is greater than 1·0, then its reciprocal is used in equation (4) to calculate the weight. A weighted average error given by

$$\overline{\Delta\varphi_{\mathbf{h}}^{w}} = \frac{\sum\limits_{\mathbf{h}} w_{\mathbf{h}} \Delta\varphi_{\mathbf{h}} |E_{\mathbf{h}}|}{\sum\limits_{\mathbf{h}} w_{\mathbf{h}} |E_{\mathbf{h}}|} \qquad (5)$$

calculated after every cycle of refinement has been used to check the change in phase error for known structures. The weights $w_{\mathbf{h}}$ are also applied to the next cycle of matrix refinement, the observed magnitudes $|E_{\mathbf{h}}|$ being replaced by the values $w_{\mathbf{h}}|E_{\mathbf{h}}|$.

Further cycles of matrix refinement are then performed with two or more eigenvalues until the absolute magnitude of λ_1 increases by less than 1%. When the matrix is reconstructed using equations (34) and (35) of Main (1975), the eigenvectors are given relative weights of $\lambda^{3/2}$ as this is the theoretical weight derived from the maximum-determinant rule (de Rango, 1969; Tsoucaris, 1970). To enhance the rate of convergence of the eigenvalues a reduced value of E_{000} is employed during matrix refinement that takes account of the limited fraction of the data set that is contained in the Karle–Hauptman determinant. The empirical value used is given by

$$(E_{000})_{\text{red}} = E_{000} \left(\frac{\sum\limits_{\text{matrix}} |E_{\mathbf{h}}|^2}{\sum\limits_{\text{data set}} |E_{\mathbf{h}}|^2} \right)^{1/2}.$$

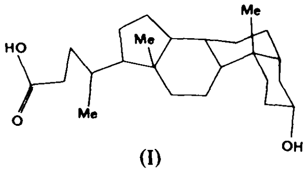
Three quantities have been used as figures of merit to rank the sets of phases obtained following matrix refinement. Firstly the value of the determinant D_{ref} obtained for the refined phases should be as large as possible, secondly the residual R [equation (3)] should be as small as possible and thirdly λ_1 should be a maximum. When the order of the determinant is a significant fraction of the number of atoms in the unit cell ($N/3$ to $N/2$) errors in the observed magnitudes have the effect of allowing negative eigenvalues to occur in the matrix constructed with the refined phases. The magnitude of the determinant can thus be negative if there is an odd number of negative eigenvalues and can no longer be used as a figure of merit for the selection of good phase sets.

Experimental results

This method of developing large phase sets by magic determinants has been tested on several structures in space group $P2_12_12_1$ and two in $P2_1$. A detailed example is given below for the D_8 determinant of lithocholic acid used as an example in the *MAGLIN* paper (Woolfson, 1977). The new magic-integer representation assigned automatically by the computer program illustrates the use of optimized magic-integer sequences and correlation procedures. The list of peaks with their associated figures of merits shows how clearly the best set of phases is indicated (see Table 8).

Detailed example

The D_8 determinant for lithocholic acid (I) ($C_{24}H_{40}O_3$, $P2_12_12_1$, $Z = 4$, $E_0 = 10{\cdot}60$) conveniently illustrates the way in which magic determinants are computed. The top row was selected by method 2. The figures and tables used earlier in this paper to clarify the correlation processes necessary for assignment of magic integers have been taken from this determinant.



(I)

Table 6. *Magic-integer representation of the 13 reflexions in the primary set*

$\|E_h\|$	h	k	l	Symbol	$\|E_h\|$	h	k	l	Symbol
2·28	2	4	11	$13x$	2·99	2	4	1	$13y$
3·11	1	1	12	$17x$	2·01	1	1	11	$-18y$
1·31	1	4	6	$20x$	2·15	0	0	10	$21y$
1·37	3	3	11	$-22x$	2·77	2	3	17	$23y$
1·35	2	3	5	$-23x$	2·38	2	1	6	$-24y$
2·16	1	1	1	$-24x$	1·78	1	0	6	$25y$
1·65	1	2	17	$-25x$					

The origin and enantiomorph were defined as follows

E_h	h	k	l	Phase (φ_h)°
2·55	3	5	0	90
1·73	1	0	5	90
2·54	0	2	23	180
1·61	0	1	6	90.

The primary set was chosen to contain all reflexions with magnitudes greater than 1·3. These 13 reflexions are represented by two magic-integer sequences of lengths 7 and 6 respectively. The correlation procedures described above result in the assignment of magic integers as shown in Table 6. The remaining reflexions, the secondary set, are expressed in magic-integer form in Table 7. The determinant with these phase representations is shown in Fig. 1. The peaks produced for this determinant as a function of x and y and the associated figures of merit are shown in Table 8.

Tests on known structures

Cholest-4-en-3-one (*steron*)(II) (Sheldrick, Oeser, Caira, Nassimbeni & Pauptit, 1976), $C_{27}H_{44}O$, space group $P2_1$, $a = 14{\cdot}634$, $b = 7{\cdot}862$, $c = 10{\cdot}674$ Å, $\beta = 105{\cdot}09°$, $Z = 2$, $N = 56$, $E_{000} = 9{\cdot}21$, $250 \|E_h\| > 1{\cdot}319$.

0,0,0 10·6 0	$0,0,\overline{10}$ 2·2 $-21y$	$\bar{2},4,\overline{11}$ 2·3 $180+13x$	$1,\bar{1},\overline{11}$ 2·0 $-18y$	$\bar{1},0,\bar{5}$ 1·7 270	$\bar{2},4,1$ 3·0 $-13y$	0,1,6 1·6 90	1,1,12 3·1 $17x$
	0,0,0 10·6 0	$\bar{2},4,\bar{1}$ 3·0 $180+13y$	$1,\bar{1},\bar{1}$ 2·2 $-24x$	$\bar{1},0,5$ 2·2 90	$\bar{2},4,11$ 2·3 $-13x$	0,1,16 0·8 $-10y+180$	1,1,22 0·4 $17x+21y$
		0,0,0 10·6 0	$3,\bar{5},0$ 2·6 90	$1,\bar{4},6$ 1·3 $-20x$	0,0,12 1·2 $-13x+13y+180$	$2,\bar{3},17$ 2·8 $-23y$	$3,\bar{3},23$ 1·1 $180+4x$
			0,0,0 10·6 0	$\bar{2},1,6$ 2·4 $180+24y$	$\bar{3},5,12$ 1·3 $5y$	$\bar{1},2,17$ 1·7 $180+25x$	0,2,23 2·5 180
				0,0,0 10·6 0	$\bar{1},4,6$ 1·3 $180-20x$	1,1,11 2·0 $-18y$	2,1,17 1·2 $-24y$
					0,0,0 10·6 0	$2,\bar{3},5$ 1·4 $23x$	$3,\bar{3},11$ 1·4 $22x$
						0,0,0 10·6 0	1,0,6 1·8 $25y$
							0,0,0 10·6 0

Fig. 1. Karle–Hauptman determinant of order 8 for lithocholic acid. Each element shows the indices, value of $|E_h|$ and representation of phase in magic-integer form or in degrees.

Table 7. *Magic-integer representation of the six reflexions in the secondary set*

$\|E_h\|$	$h\,k\,l$	$h\,k\,l$	$h\,k\,l$	Symbol
1·28	$\varphi(3,5,12) \simeq$	$\varphi(2,4,1)(13y)$	$+\ \varphi(1,1,11)(-18y)$	$-5y$
1·18	$\varphi(0,0,12) \simeq$	$\varphi(\bar{2},4,11)(-13x)$	$+\ \varphi(2,\bar{4},1)(-13y+180°)$	$-13x-13y+180°$
1·15	$\varphi(2,1,17) \simeq$	$\varphi(0,2,23)(180°)$	$+\ \varphi(2,\bar{1},\bar{6})(-24y-180°)$	$-24y$
1·05	$\varphi(3,3,23) \simeq$	$\varphi(2,4,11)(13x)$	$+\ \varphi(1,\bar{1},12)(-17x+180°)$	$-4x+180°$
0·84	$\varphi(0,1,16) \simeq$	$\varphi(\bar{2},4,\bar{1})(13y-180°)$	$+\ \varphi(2,\bar{3},17)(-23y)$	$-10y+180°$
0·40	$\varphi(1,1,22) \simeq$	$\varphi(0,0,10)(21y)$	$+\ \varphi(1,1,12)(17x)$	$17x+21y$

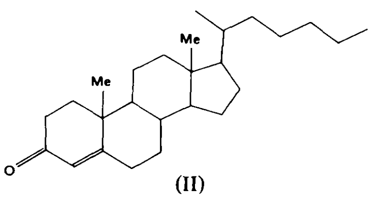
Table 8. *Lists of peaks and associated figures of merit for the lithocholic acid D_8 determinant*

x	y	D ($\times 10^{-8}$) (initial)	$\overline{\Delta\varphi_h^w}$ (initial)	$\overline{\Delta\varphi_h^w}$ (final)	R (%)	λ_1 ($\times 10^{-4}$)	D ($\times 10^{-8}$) (refined)
10	21	0·6887	85·4°	83·3°	28·4	0·1335	0·7242
29	3	0·6893	52·0	56·1	28·4	0·1370	0·7572
56	83	0·6927	79·5	81·0	29·5	0·1339	0·7521
84	1	0·6747	76·7	80·6	35·6	0·1325	0·7224
54	73	0·6692	47·8	61·5	26·4	0·1384	0·7651
83	41	0·6702	65·8	62·8	38·9	0·1262	0·7102
1	23	0·6606	75·5	79·4	32·1	0·1330	0·7286
1	85	0·6625	81·5	81·0	32·5	0·1311	0·7189
42	8	0·6653	71·8	80·9	34·3	0·1289	0·7267
87	33	0·6674	83·6	95·5	39·5	0·1253	0·6607
10	59	0·6574	74·9	78·0	24·0	0·1370	0·7381
24	35	0·6659	27·2	22·8	24·7	0·1394	0·7657
56	6	0·6528	77·1	73·5	33·3	0·1300	0·7100
81	54	0·6571	86·5	91·4	41·1	0·1265	0·7139
82	25	0·6487	79·6	73·1	27·0	0·1375	0·7633
36	52	0·6446	66·5	63·2	36·5	0·1270	0·7112
52	53	0·6457	58·6	62·3	42·4	0·1234	0·6775
56	22	0·6427	85·1	82·2	30·7	0·1378	0·7519
60	52	0·6451	60·7	62·0	29·7	0·1318	0·7233
82	3	0·6462	69·9	60·6	37·8	0·1268	0·7189

Table 9. *Magic-determinant trials for steron*

Order	$(E_{000})_{\text{red}}$	Number of unique E_h	Number of $\lvert E_h\rvert \geq 1·3$	Magic integers: Sequence length x	y	r.m.s. integer in triple-phase invariants x	y	Known reflexions (origin + enantiomorph, others)	Best results for all $\lvert E_h\rvert \geq 1·3$: $\overline{\Delta\varphi_h^w}$	Number with $w \geq 0·50$	Number of peaks selected	Position of D_{init} in ranking order for chosen peak	Position of D_{ref}, $\lambda_{\max}$ in ranking order for chosen peak after refinement
8	3·45	19	18	5	5	17·16	16·60	4,1	12·3	18/18	–	–	–
Most solutions are equivalent with very little difference in the three figures of merit D_{ref}, $\lambda_{\max}$ and R.													
20	7·89	72	51	7	7	19·94	18·96	4,1	12·6	50/51	22	5	(5),2
				12	12	20·81	20·50	4,1	13·1	50/51	21	17	(6),2
				12	12	19·43	20·81	3,1 (enantiomorph not fixed)	15·2	51/51	20	16	(7),1
R and λ_1 are very discriminating; D_{ref} is less than D_{init} by a factor of 10^3.													
30	9·21	133	68	10	9	22·15	20·21	4,1	22·4	66/68	20	4	(–),1
				12	12	20·92	21·93	4,1	10·5	66/68	20	10	(–),5
									27·4	66/68		19	(–),1
(new weighting scheme)				12	12	20·92	21·93	4,1	11·3	66/68	20	10	(–),1

At this order of determinant for this small structure, the determinant value is not a good figure of merit.



(II)

A Karle–Hauptman determinant for this structure was selected by Method 2. An initial determinant of order 500 was constructed using reflexions for which $\sin\theta/\lambda < 0·45$ Å^{-1}. Rows and columns were then eliminated down to order 250. A second determinant of order 250 was then constructed using reflexions that occurred at least 10 times in the order 250 determinant produced in the first step. Rows and columns were then eliminated down to order 100. The top row of this determinant up to order 30 is as follows:

$\bar{1}\bar{1}\bar{1}$, $93\bar{5}$, $\bar{6}3\bar{1}$, 261, 060, $\bar{2}0\bar{2}$, $\bar{2}6\bar{2}$, 631, 202, $1\bar{1}1$, $\bar{9}35$, 262, 201, $\bar{5}38$, $\bar{7}36$, 403, $\bar{4}30$, $\bar{2}32$, $\bar{7}37$, $\bar{4}31$, $\bar{2}6\bar{1}$, $\bar{2}0\bar{1}$, $\bar{4}0\bar{3}$, $\overline{11},3,3$, $\bar{3}06$, $\bar{9}\bar{3}5$, $43\bar{1}$, $\bar{5}04$, $23\bar{2}$.

Origin-defining reflexions	201	360°
	$15,0,\bar{4}$	360
	111	19
Enantiomorph fixed by	$93\bar{5}$	225
Permuted phase	202	180, 360.

The magic determinant was evaluated at orders 8, 20 and 30. The results obtained are shown in Table 9. The best sets of phases were extended using the tangent-

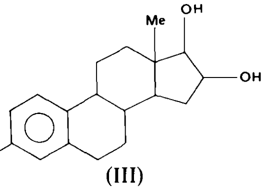
Table 10. *Magic-determinant trials for estriol*

Order	$(E_{000})_{\rm red}$	Number of unique $E_{\mathbf{h}}$	Number of $\|E_{\mathbf{h}}\| \geq 1{\cdot}3$	Magic integers Sequence length x	y	r.m.s. integer in triple-phase invariants x	y	Known reflexions (origin + enantiomorph, others)	Best results for all $\|E_{\mathbf{h}}\| \geq 1{\cdot}3$ $\overline{\Delta\varphi_{\mathbf{h}}^{w}}$	Number with $w \geq 0{\cdot}50$	Number of peaks selected	Position of $D_{\rm init}$ in ranking order for chosen peak	Position of $D_{\rm ref}, \lambda_{\max}$ in ranking order for chosen peak after refinement
8	2·65	19	17	7	6	16·83	18·16	3,1	28·5	17/17	20	6	(–),(–)
Most solutions are equivalent with very little difference in the three figures of merit. Invariants in the matrix are close to 0° except for those linking the reflexions 2,10,1 and 0,10,1 with other reflexions in the matrix.													
20	5·79	73	46	12	12	19·46	19·75	3,1	40·0	46/46	21	8	6,1
The figures of merit are more discriminating at this order of determinant.													
30	8·04	152	87	12	12	21·37	21·87	3,1	41·6	82/87	20	20	5,1
									37·0	81/87		10	3,3

$$\textit{FASTAN}\text{ phase development of 81 phases }\left\{\begin{array}{ll}\overline{\Delta\varphi_{\mathbf{h}}} = 0° & 41/42\\ \overline{\Delta\varphi_{\mathbf{h}}} = 44{\cdot}0° & 32/42\\ \overline{\Delta\varphi_{\mathbf{h}}} = 36{\cdot}9° & 30/42\end{array}\right\}\text{ atoms located in the } E \text{ map.}$$

formula part of the *MULTAN* system and *E* maps calculated using the resulting phases. About half the steron molecule could be located from such *E* maps. When *MULTAN* itself was used in the normal way with a convergence map based on origin + enantiomorph + two other reflexions, the set with the highest combined figure of merit out of the twelve sets produced an *E* map with 25 out of the 28 atoms correctly located. For this structure the small value of *N* (56) produces negative values for the determinants at order 30 even for refined sets of phases. The value of the determinant cannot therefore be used as a figure of merit in this case.

Estriol (III) (Hauptman, Fisher, Hancock & Norton, 1969), $C_{18}H_{24}O_3$, space group $P2_1$, $a = 9{\cdot}261$, $b = 23{\cdot}025$, $c = 7{\cdot}561$ Å, $\beta = 111{\cdot}00°$, $Z = 4$, $N = 84$, $E_{000} = 10{\cdot}61$, $412|E_{\mathbf{h}}| > 1{\cdot}36$.



(III)

A Karle–Hauptman determinant for this structure was selected by method 2. An initial determinant of order 500 was constructed using reflexions for which $\sin\theta/\lambda < 0{\cdot}42$ Å^{-1}. Rows and columns were then eliminated down to order 250. A second determinant of order 250 was then constructed using reflexions that occurred at least 12 times in the order 250 determinant produced in the first step. Rows and columns were then eliminated down to order 100. The top row of this determinant up to order 30 is as follows:

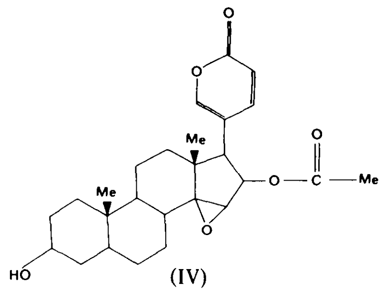
$13\bar{3}$, $\bar{1}3\bar{3}$, 172, $32\bar{4}$, 372, $2,10,\bar{1}$, 200, $0,10,\bar{1}$, $\bar{1}\bar{7}\bar{2}$, $0,\overline{10},\bar{1}$, $1\bar{3}3$, $2,\overline{10},1$, $\bar{2}00$, $\bar{1}\bar{3}3$, $\bar{1},\overline{13},4$, 245, $4\bar{1}\bar{1}$, $2\bar{1}\bar{1}$, 045, $\bar{2}56$, $\bar{3}\bar{2}4$, $\bar{2}46$, $2\bar{4}\bar{6}$, $\bar{4}11$, $\bar{2}11$, 124, $\bar{1}\bar{2}4$, $\bar{1},18,2$, $1,\overline{18},\bar{2}$.

There are very few *h*0*l* reflexions in this structure as a result of the length of the *b* axis. Origin and enantiomorph were therefore defined using general reflexions.

Origin and enantiomorph definition	133	130°
	$32\bar{4}$	123
	045	333
Permuted phase	200	180, 360.

Magic determinants were evaluated at orders 8, 20 and 30. The results obtained are shown in Table 10. The best sets of phases were extended using the tangent-formula part of the *MULTAN* system and *E* maps calculated using the resulting phase sets. 32 out of the 42 atoms could be located from *E* maps generated using the approximate phases from the order 30 determinant. (When the correct values for the phases in the order 30 determinant were used as the basis for tangent-formula phase development then 41 out of 42 atoms were correctly located in the resulting *E* map.)

Cinobufagin (IV) (Declercq, Germain & King, 1977), $C_{26}H_{34}O_6$, space group $P2_12_12_1$, $a = 7{\cdot}663$, $b = 15{\cdot}900$, $c = 19{\cdot}291$ Å, $Z = 4$, $N = 128$, $E_{000} = 12{\cdot}90$, $320|E_{\mathbf{h}}| > 1{\cdot}41$.



(IV)

Method 1 was used to select an order 25 Karle–Hauptman determinant. A *MULTAN* convergence map on to six reflexions was used to indicate which

Table 11. *Magic-determinant trials for cinobufagin*

Order	$(E_{000})_{\text{red}}$	Number of unique $E_{\mathbf{h}}$	Number of $\lvert E_{\mathbf{h}}\rvert \geq 1{\cdot}3$	Magic integers Sequence length x	 Sequence length y	r.m.s. integer in triple-phase invariants x	r.m.s. integer in triple-phase invariants y	Known reflexions (origin + enantiomorph)	Best results for all $\lvert E_{\mathbf{h}}\rvert \geq 1{\cdot}3$ $\overline{\Delta\varphi_{\mathbf{h}}^{w}}$	 Number with $w \geq 0{\cdot}50$	Number of peaks selected	Position of D_{init} in ranking order for chosen peak	Position of $D_{\text{ref}}, \lambda_{\text{max}}$ in ranking order for chosen peak after refinement
15	(12·90)	65	42	9	9	22·70	21·15	4	30·2	39/42	39	35	5,1
	(only one eigenvalue used, all three figures of merit are discriminating)												
	(12·90)	(two eigenvalues used, relative weight of eigenvalue = $\lambda^{3/2}$)							29·2	40/42	39	35	3,1
	(12·90)	(three eigenvalues used)							28·0	39/42	39	35	2,1
	3·51								27·8	38/42	39	35	6,4
	3·51	(correct scaling of observed and calculated $\lvert E_{\mathbf{h}}\rvert$ values)							27·9	38/42	39	35	5,3
	The above results compare different refinement procedures for the same initial peak in the determinant map.												
25	(12·90)	155	73	9	9	22·85	24·20	44	42·9	67/73	22	19	9,6
									50·2	71/73		4	3,3
									46·3	72/73		1	2,2
	7·67			12	12	22·81	23·18	4	41·6	65/73	21	6	1,1
									45·8	67/73		10	2,2
	7·67	(new correlation method)		8	8	24·18	26·01	4	43·6	69/73	21	14	1,1
30	(12·90)	94	40	9	9	31·76	22·68	4	33·8	35/40	15	15	4,4
	8·92			12	12	21·09	25·87	4	52·4	35/40	22	1	1,10
									50·8	36/40		4	2,1
	(several other comparable solutions, some on other origins)								33·5	31/40		9	3,9
	8·92			7	7	23·64	22·53	4	52·5	37/40	20	20	2,1
									53·1	32/40		5	1,2
	8·92			12	12	25·19	24·31	4	49·8	32/40	22	7	6,5
									52·2	36/40		16	2,1

reflexions might be suitable for inclusion in the top row of the determinant. The top row of this determinant up to order 25 is:

$0,12,1,\ 5,12,0,\ \bar{1},15,0,\ 4,7,15,\ 4,5,14,\ \bar{1},13,\bar{1},\ 1\bar{1}\bar{2},$
$1\bar{1}2,\ \bar{1}1\bar{2},\ 0,12,\bar{1},\ 41\bar{1},\ 301,\ 4\bar{1}1,\ 30\bar{1},\ 501,\ 50\bar{1},$
$2,7,13,\ 1,15,2,\ 1,15,\bar{2},\ 4,13,\bar{2},\ 1,13,1,\ 4,13,2,$
$2,13,\bar{4},\ 2\bar{2}\bar{4}.$

Magic determinants were evaluated at orders 15 and 25 for the above top row.

Origin and enantiomorph for the order 25 determinant were defined as follows:

Origin-defining reflexions	0,12, 1	360°
	5,12, 0	90
	1,15, 0	270
Enantiomorph fixed by	4, 5,14	135.

Another Karle–Hauptman determinant was selected by method 2. An initial determinant of order 500 was constructed using reflexions for which $\sin\theta/\lambda < 0{\cdot}45$ Å^{-1}. Rows and columns were then eliminated down to order 100. A second determinant of order 500 was then constructed using reflexions that occurred at least three times in the order 100 determinant produced in the first step. Rows and columns were then eliminated down to order 100. The top row of this determinant up to order 30 is as follows:

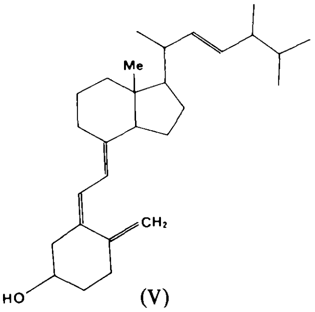
$\bar{1},13,1,\ \bar{4}11,\ 301,\ 07\bar{6},\ \bar{1},13,1,\ 411,\ \bar{3}0\bar{1},\ \bar{1},13,\bar{3},$
$1,13,\bar{3},\ \bar{2}00,\ 200,\ \bar{1},13,\bar{1},\ 1,13,\bar{1},\ \bar{1}1\bar{2},\ 11\bar{2},\ \bar{4}1\bar{1},\ \bar{1}12,$
$\bar{1},13,3,\ 1,13,3,\ 112,\ 41\bar{1},\ 301,\ \bar{3}01,\ 0,12,\bar{1},\ 0,12,1,$
$1\bar{1}\bar{2},\ \bar{1}1\bar{2},\ \bar{2}1\bar{1},\ 21\bar{1}.$

Origin and enantiomorph for the order 30 determinant were defined as follows:

Origin-defining reflexions	5,12,0	90°
	301	90
	076	90
Enantiomorph fixed by	1,13,1	225.

Magic determinants were evaluated at order 30 for the above top row. The results for cinobufagin determinants are shown in Table 11. The best sets of phases were extended using the tangent-formula part of the *MULTAN* system and *E* maps calculated using the resulting phase sets. Only a small number of atoms could be located from *E* maps generated using the approximate phases from magic determinants at orders 25 and 30. In both cases when correct phase values were used as the basis for the tangent-formula phase development all the atoms were correctly located in the resulting *E* maps. Two hypotheses may be put forward as to why the sets of magic-determinant phases are not good enough to provide a starting set from which a solution can be obtained. Firstly, the determinants themselves do not contain a good representative subset of the complete data set. Secondly, the large number of special reflexions included lead to sets of phases that are very close to a centrosymmetric structure and cannot be developed to a noncentrosymmetric phase set by the tangent formula.

Ergocalciferol (V) (Hull, Leban, Main, White & Woolfson, 1976), $C_{28}H_{44}O$, space group $P2_12_12_1$, $a = 21\cdot695$, $b = 6\cdot857$, $c = 35\cdot320$ Å, $Z = 8$, $E_{000} = 16\cdot93$, $412|E_{\mathbf{h}}| > 1\cdot43$, $N = 232$.



(V)

(*a*) Method 1 was used to select an order 8 determinant. A *MULTAN* convergence map on to six reflexions was used to indicate which reflexions might be suitable for inclusion in the top row of the determinant. The top row of this order 8 determinant is as follows:

$40\bar{1}$, 7,0,29, 013, 2,1,29, 6,1,28, 6,1,30, $4\bar{1}2$.

Origin-defining reflexions	401	90°
	013	90
	7,0,29	90
Enantiomorph fixed by	2,1,29	315.

(*b*) Method 2 was used to select an order 14 determinant. In this case $\overline{||E_{\mathbf{h}}|^2 - 1|}$ was used as a figure of merit in the final elimination of rows and columns from order 100. The top row of this determinant is as follows:

$40\bar{1}$, $13\bar{1}$, $0\bar{1}\bar{3}$, $03\bar{3}$, $\bar{5}14$, $\bar{4}0\bar{1}$, 530, 401, $0\bar{1}3$, 131, $\bar{4}01$, 033, $\bar{5}14$.

Origin-defining reflexions	401	90°
	013	90
	530	90
Enantiomorph fixed by	527	135.

(*c*) Method 2 was used to select this determinant. An initial determinant of order 500 was constructed and rows and columns were eliminated down to order 100. The determinant was not set up a second time. The top row of this determinant up to order 29 is:

$\bar{5}10$, $\bar{1}1\bar{1}$, 2,1,14, $\bar{5}0\bar{3}$, $\bar{5}\bar{1}0$, $\bar{1}\bar{1}\bar{1}$, 020, $0\bar{2}0$, $1\bar{1}1$, 111, 2,$\bar{1}$,14, 503, 510, $5\bar{1}0$, $50\bar{3}$, $\bar{5}03$, $11\bar{1}$, $1\bar{1}\bar{1}$, 3,0,15, $\bar{3}$,0,$\overline{15}$, $\bar{1}11$, $\bar{1}\bar{1}1$, $\bar{3}$,0,15, 3,0,$\overline{15}$, $\bar{2}$,1,14, $\bar{2}$,$\bar{1}$,14, 401, $\bar{4}01$.

Origin-defining reflexions	401	90°
	510	90
	013	90
Enantiomorph fixed by	3,0,15	270.

Magic determinants were evaluated at orders 19, 25 and 29.

(*d*) This determinant was also selected by method 2. An initial determinant of order 1500 was constructed using reflexions for which $\sin\theta/\lambda < 0\cdot42$ Å^{-1}. Rows and columns were then eliminated down to order 750. A second determinant of order 750 was then constructed using reflexions that occurred at least 40 times in the order 750 determinant produced in the first step. Rows and columns were then eliminated down to order 150. The top row of this determinant up to order 30 is as follows:

$01\bar{3}$, $\bar{5}10$, $\bar{3}$,$\bar{1}$,$\overline{17}$, $\bar{4}0\bar{1}$, 3,$\bar{1}$,13, $\bar{5}0\bar{3}$, $0\bar{1}\bar{3}$, $\bar{5}\bar{1}0$, $\bar{3}$,1,$\overline{17}$, $50\bar{3}$, $03\bar{3}$, $0\bar{3}3$, $40\bar{1}$, 3,1,13, $\bar{1}$,1,14, $\bar{1}$,$\bar{1}$,14, $\bar{4}01$, $\bar{1}11$, $\bar{1}\bar{1}1$, $3\bar{1}2$, 312, 401, 0,1,$\overline{16}$, 0,$\bar{1}$,$\overline{16}$, $\bar{6}$,0,$\overline{15}$, 6,0,$\overline{15}$, $\bar{3}$,1,13, $\bar{3}$,$\bar{1}$,13, 0,2,$\overline{14}$.

Origin-defining reflexions	401	90°
	510	90
	013	90
Enantiomorph fixed by	3,1,17	135
Permuted phase	6,0,30	360, 180.

Ergocalciferol is an extremely difficult structure to solve. No set of approximate phases from any of the above magic determinants has yielded interpretable *E* maps following extension of the phase set by the tangent formula. Even when the exact phases for the magic-determinant phases are used no recognizable fragments are obtained in the *E* maps. A combination of the contents of determinants (*a*) and (*b*) given exact phases and used as a starting set for *MULTAN* produced an *E* map in which 38 out of the 58 atoms were correctly located ($\overline{\Delta\varphi_{\mathbf{h}}} = 44\cdot4°$ for $412E_{\mathbf{h}}$'s).

The results obtained for all the above magic determinants are shown in Table 12.

Conclusions

Several important improvements to magic-integer methods of phase determination have been made as a result of this work. The new optimized magic-integer sequences, described in paper XI of this series, have been used with success enabling large primary sets to

Table 12. *Magic-determinant trials for ergocalciferol*

				Magic integers					Best results for all $\|E_{\mathbf{h}}\| \geq 1{\cdot}3$				
Order	$(E_{000})_{\text{red}}$	Number of unique $E_{\mathbf{h}}$	Number of $\|E_{\mathbf{h}}\| > 1{\cdot}3$	Sequence length x	Sequence length y	r.m.s. integer in triple-phase invariants x	r.m.s. integer in triple-phase invariants y	Known reflexions (origin + enantiomorph, others)	$\overline{\Delta\varphi_{\mathbf{h}}^{w}}$	Number with $w \geq 0{\cdot}50$	Number of peaks selected	Position of D_{init} in ranking order for chosen peak	Position of $D_{\text{ref}}, \lambda_{\text{max}}$ in ranking order for chosen peak after refinement
8	(16·93)	23	20	5	5			4,0	16·1	20/20			
	4·12			3	3	4·62	3·91	4,0	16·3	20/20		(Many peaks give the same solution after refinement)	
	4·12			8	8	18·57	15·22	4,0	22·25	20/20			
14	(16·93)	36	22	5	5			4,0	43·2	22/22			
	(16·93)			7	7			4,0	33·9	21/22			
	6·0			5	5	18·89	18·84	4,0	34·9	20/22	20	16	19,5
19	(16·93)	44	23	5	5	22·72	19·18	4,0	45·0	21/23	17	5	4,3
				10	10	26·55	24·0	4,0	45·0	21/23	17	2	2,1
25	(16·93)	61	30	9	9	26·16	21·56	4,0	50·1	28/30	15	2	2,1
						Several other peaks refine to $\overline{\Delta\varphi_{\mathbf{h}}} = 50°$ for 30 phases.							
29	(16·93)	77	35	9	9	25·45	22·39	4,0	49·8	33/35	26	22	1,1
	10·61			12	12	23·01	21·73	4,0	47·1	25/35	22	5	4,1
	10·61			7	7	22·52	23·34	4,0	46·4	31/35	20	5	1,1
	10·61			12	12	23·58	23·43	4,0	42·6	21/35	20	3	5,6
30	9·96	156	68	12	12	26·63	25·51	4,1	47·1	53/68	20	7	1,1

be represented efficiently by comparatively small integers. To enhance still further the power of magic integers, several correlation procedures were investigated in an attempt to minimize the occurrence of large coefficients in triple-phase invariants represented by magic integers. The most powerful method found to date is the one described above under '*Step* 1'.

In general, the values of the Karle–Hauptman determinants used in this work have been found to be very discriminating as figures of merit. The results presented in Tables 8–12 indicate that it is only necessary to consider the top 20–30 peaks in the determinant map in order to obtain a satisfactory set of phases. When the determinant is used as a figure of merit to rank the phase sets following matrix refinement it is even more discriminating and the best sets of phases are usually found in the top six values. However, when the order of the determinant, m, is $\simeq N/3$ to $N/2$, the presence of negative eigenvalues in the eigenvalue spectrum precludes the use of the determinant values for this purpose. This is most noticeable with small structures, *e.g.* steron, and under these circumstances λ_{max} and R must be used to rank the phase sets. It can be seen from Tables 8–12 that λ_{max} is in fact as discriminating a figure of merit for ranking the refined phase sets as is the determinant value and may therefore be used instead of the determinant values where necessary.

The sets of phases obtained from magic determinants for the test structures appear fairly good until the extension of the phase sets using *FASTAN* is examined. Early experiments with lithocholic acid data showed that it was easy to extend a phase set by *FASTAN* and thus solve the structure. The detailed results for the four test structures show that this early promise has not been maintained. The main cause for these disappointing results appears to lie with the nature of the matrix of structure factors used to derive the starting sets. As mentioned above, it is extremely difficult to select a matrix of structure factors by a systematic process so as to include reflexions with large magnitudes occurring repeatedly. In addition the elimination process encourages the selection of reflexions that represent only a sub-unit of the structure. The effect of this is clearly seen in the case of steron where a phase set, developed by *FASTAN* from 66 magnitudes with $\overline{\Delta\varphi_{\mathbf{h}}^{w}} = 11{\cdot}3°$, showed only $\sim\frac{1}{2}$ the molecule in the resulting E map. In this case, the starting set of 66 phases is too close to a centrosymmetric sub-unit and the final E map shows a centrosymmetric pattern of peaks. In an effort to solve this problem, work is currently in progress on the production of more suitable matrices for use as magic determinants.

A major drawback to the determinant approach is the large time taken to produce a determinant map, especially when compared with the time required to produce a ψ map (White & Woolfson, 1975) using FFT's in more straightforward applications of the magic-integer approach. Since the time taken for determinant evaluation is proportional to m^3, there is a practical upper limit of $m = 30$ for the computing facilities currently available to the authors. This is unfortunate since for large structures it is desirable to have as large a starting set as possible and matrices of higher order would enable more reflexions to be included in the matrix. However, for these larger matrices some triple-phase invariants involving the large magnitudes do not occur in the matrix and the proportion increases with the number of large magnitudes in the matrix. This extra information can only be used if a function based on a Fourier summation is employed instead of a determinant to derive positions in (x,y) space. Tests are now being made using a 2D FFT summation that includes contributions from these omitted triple-phase invariants.

In conclusion, although the value of a Karle–Hauptman determinant is a discriminating figure of merit for selecting good phase sets, when used in conjunction with magic integers in a phase-determining role, it is not of sufficient power to justify the amount of computer time required to evaluate the very large number of determinants involved.

Two of us (MMW and DJT) gratefully acknowledge the assistance of the Science Research Council for sponsorship of a project to develop the use of the magic-integer concept.

References

ARORA, S. K., GERMAIN, G. & DECLERCQ, J.-P. (1976). *Acta Cryst.* **B32**, 415–419.

DAWSON, B. (1961). *Acta Cryst.* **14**, 999–1000.

DECLERCQ, J.-P., GERMAIN, G. & KING, G. S. D. (1977). Fourth European Crystallographic Meeting, Oxford. Abstracts *A*, pp. 279–280.

DECLERCQ, J.-P., GERMAIN, G. & WOOLFSON, M. M. (1975). *Acta Cryst.* **A31**, 367–372.

GOEDKOOP, J. A. (1950). *Acta Cryst.* **3**, 374–378.

HAUPTMAN, H., FISHER, J., HANCOCK, H. & NORTON, D. (1969). *Acta Cryst.* **B25**, 811–814.

HULL, S. E., LEBAN, I., MAIN, P., WHITE, P. S. & WOOLFSON, M. M. (1976). *Acta Cryst.* **B32**, 2374–2381.

KARLE, J. & HAUPTMAN, H. (1950). *Acta Cryst.* **3**, 181–187.

KITAIGORODSKY, A. I. (1961). *Theory of Crystal Structure Analysis*. New York: Consultants Bureau.

MAIN, P. (1975). *Crystallographic Computing Techniques*, edited by F. R. AHMED, pp. 165–175. Copenhagen: Munksgaard.

MAIN, P. (1977). *Acta Cryst.* **A33**, 750–757.

MAIN, P. (1978). *Acta Cryst.* **A34**, 31–38.

MESSAGER, J. C. & TSOUCARIS, G. (1972). *Acta Cryst.* **A28**, 482–484.

RANGO, C. DE (1969). Thesis, Paris.

SHELDRICK, G. M., OESER, E., CAIRA, M. R., NASSIMBENI, L. R. & PAUPTIT, R. A. (1976). *Acta Cryst.* **B32**, 1984–1987.

TSOUCARIS, G. (1970). *Acta Cryst.* **A26**, 492–499.

WHITE, P. S. & WOOLFSON, M. M. (1975). *Acta Cryst.* **A31**, 53–56.

WOOLFSON, M. M. (1977). *Acta Cryst.* **A33**, 219–225.

Acta Cryst. (1978). A**34**, 883-892

On the Application of Phase Relationships to Complex Structures. XVI. A Random Approach to Structure Determination

BY R. BAGGIO* AND M. M. WOOLFSON

Department of Physics, University of York, Heslington, York YO1 5DD, England

AND J. P. DECLERCQ AND G. GERMAIN

Laboratoire de Chimie Physique et de Cristallographie, Université de Louvain, Bâtiment Lavoisier, 1 place Louis Pasteur, 1348 Louvain-la-Neuve, Belgium

(*Received* 15 *March* 1978; *accepted* 26 *May* 1978)

With triple-phase relationships treated as linear equations it is possible to refine a set of phases from given initial values. Phases so obtained are better than those found by refining to self-consistency with the tangent formula. An investigation of the radius of convergence of the least-squares refinement process showed that a substantially correct solution may often be found even starting with random phases. Systems containing up to 300 phases have been investigated and the results and their implications are discussed. It is concluded that the random approach can, at the very least, be used to obtain 70–100 phases as a good starting point for phase development. There is also the possibility of obtaining a sufficient number of phases directly to define a reasonably complex structure, especially with a computer augmented by an array processor. A problem which can arise with linear equations, as with the tangent formula, is that the phases obtained do not adequately define the enantiomorph and give an E map with a pseudo centre of symmetry. Two methods of overcoming this problem are described.

---

* Present address: Comision Nacional de Energia Atomica, Av. Libertador 8250, Buenos Aires, Argentina.