Acta Cryst. (1965). **18**, 908

A Systematic Method for Unravelling a Periodic Vector Set

BY MASAYASU TOKONAMI AND SUKEAKI HOSOYA

Institute for Solid State Physics, University of Tokyo, Azabu, Minato-ku, Tokyo, Japan

(*Received* 14 *May* 1964)

A new method is presented for unravelling a periodic vector set. This method, being entirely different from the image seeking method or its variations, makes full and systematic use of periodic characters involved in a periodic vector set. The details of unravelling procedures are systematically described. The procedures are illustrated in a simple case, and further it is shown that the proposed procedures are really effective for unravelling vector sets corresponding to several hypothetical and real structures. In some examples, the computations including logical operations are carried out on an electronic computer. Some characteristics of the present method are discussed. General aspects of the phase problem are also described from a standpoint reached through the present work.

Introduction

The Patterson function, $P(\mathbf{r})=\varrho(\mathbf{r})*\varrho(\mathbf{r})$, is a self-convolution, or a spatial autocorrelation function, of a periodic electron density $\varrho(\mathbf{r})$ and can be derived directly from a set of absolute values of the crystal structure factors, $|F(hkl)|$, which are obtained from observed intensities. The periodic electron density $\varrho(\mathbf{r})$ is expressed by:

$$\varrho(\mathbf{r})=\sum_t \delta(\mathbf{r}-\mathbf{r}_t)*\varrho_{\text{unit cell}}, \qquad (1)$$

where δ is Dirac's delta function, $\mathbf{r}_t$ means the position vector for the origin of the tth unit cell, and the sum is taken over all unit cells. When all atoms are identical and are supposed to be point atoms, the density in a unit cell is essentially given by

$$\varrho_{\text{unit cell}}(\mathbf{r})=\sum_{i=1}^{N}\delta(\mathbf{r}-\mathbf{r}_i)\,, \qquad (2)$$

where N is the number of point atoms in a unit cell. It is important to study how to find out $\varrho_{\text{unit cell}}(\mathbf{r})$ from $P(\mathbf{r})$ for case (2), because such a study may help us to solve the problem for a real crystal. In this special case, $P(\mathbf{r})$ is called a periodic vector set and $\varrho(\mathbf{r})$ a periodic fundamental set. When, on the other hand, $\varrho(\mathbf{r})$ is not periodic and is expressed as

$$\varrho(\mathbf{r})=\sum_{i=1}^{N}\delta(\mathbf{r}-\mathbf{r}_i)\,, \qquad (3)$$

the functions $P(\mathbf{r})$ and $\varrho(\mathbf{r})$ are called a finite vector set and a finite fundamental set respectively.

Let us consider a finite fundamental set consisting of N points in an m-dimensional space, and a periodic fundamental set derived from the former by introducing periodicity. The finite and the periodic vector sets derived respectively from the above two fundamental sets differ from each other as follows. The finite vector set which consists of N^2 points including N points at the origin is found within a space of volume 2^m times that of a unit cell, while the periodic vector set includes N^2 points per unit cell, corresponding to the superposition of the above finite vector sets with periodic lattice translations.

The image seeking method for unravelling the finite vector set was first devised by Wrinch (1939), and later by Buerger (1950, 1959) and many others. Wrinch and Buerger suggested that a similar method might be used also to unravel a periodic vector set. In order to unravel a periodic vector set by this method, however, the patterns in a unit cell are not enough and those in 2^m unit cells are required. If a certain point takes part in image construction, 2^m-1 points which are equivalent to this point but with shifts corresponding to the lattice translations cannot take part. It is not always easy to decide which point should be chosen out of 2^m equivalent points at every stage.

In the present paper, a new method is proposed for unravelling a periodic vector set. This method, being entirely different from the image seeking method or its variations, makes full and systematic use of all periodic characters involved in a periodic vector set. Merits and demerits of the both methods will be compared after the new method is described for hypothetical and real structures. General aspects of the phase problem will also be discussed from a standpoint reached through the present work.

Theory

A periodic fundamental set and the relevant periodic vector set can be specified respectively by fractional coordinates of each point and by corresponding components of interatomic vectors. The former will hereafter be expressed as $\mathbf{v}_i(x_i, y_i, z_i)$ with small letters, and the latter as $\mathbf{V}_{ij}(X_{ij}, Y_{ij}, Z_{ij})$ with capital letters, where $\mathbf{V}_{ij}$ means a vector from the ith to the jth point of the fundamental set. In a one-dimensional case, only x (or X) coordinates, and in a two-dimensional case, x (or X) and y (or Y) coordinates are to be taken into account. However, general descriptions will hereafter be given in a three-dimensional case, if not otherwise stated.

For **V** vectors, the lattice translations along Y and Z directions are ignored throughout the present work. In other words, only fractional parts are taken into consideration for Y and Z components of a vector, whereas its X component holds its usual meaning in full. According to this definition, it follows that vectors, for example, $\mathbf{V}(X, Y, Z)$, $\mathbf{V}'(X, Y, Z+1)$ and $\mathbf{V}''(X, Y+1, Z)$ are taken to be identical with one another but to be different from a vector $\mathbf{V}'''(X+1, Y, Z)$.

Let us specify N points of a periodic fundamental set, in a unit cell lying in $0 \leq x < 1$, by numbers $0, 1, 2, \ldots, N-1$ in the order of increasing x coordinates, In a similar way, N points in a unit cell lying in $1 \leq x < 2$ are specified by numbers $N, N+1, \ldots, 2N-1$. Points p and $p+N$ are correlated with each other by a lattice translation **a** in the x direction. If two or more points happen to have the same x coordinate, their order should be decided according to their y and z coordinates.

The vector $\mathbf{V}_{p,\,p+n}$ from a point p to a point $p+n$ is called *an nth nearest neighbour vector* (hereafter abbreviated to an nth n. n. vector), though the meaning of the word *nearest* is here limited to the plus x direction. This vector will be often denoted as $\mathbf{V}_p(n)$, while the notation, $\mathbf{V}(n)$ without suffix, means an nth n. n. vector in general.

Every $\mathbf{V}_p(n)$ vector is a member of the periodic vector set which includes N^2 vectors per unit cell. Therefore, these N^2 vectors, say in a region $0 \leq X < 1$, can be classified into N groups of nth n. n. vectors according as n ranges from 0 to $N-1$, and each group contains N member vectors.

The numbering of **V** vectors, and accordingly the classification of **V** vectors, depends upon the choice of the coordinate system. In any case, however, the following relations are valid among the vectors $\mathbf{V}_p(n)$.

(A) A vector $\mathbf{V}_i(n)$ with $n>1$ can be decomposed as

$$\begin{aligned}\mathbf{V}_i(n) &= \mathbf{V}_i(1)+\mathbf{V}_{i+1}(n-1)\\ &= \mathbf{V}_i(n-1)+\mathbf{V}_{i+n-1}(1)\\ &= \mathbf{V}_i(1)+\mathbf{V}_{i+1}(1)+\cdots\cdots+\mathbf{V}_{i+n-1}(1)\,.\end{aligned}$$

(A') A vector $\mathbf{V}_i(1)$ cannot be decomposed any further, so long as trivial zero vectors $\mathbf{V}(0)$ are disregarded, although some $\mathbf{V}(1)$ vectors may happen to coincide with the sum of other vectors.

(B) The total sum over all members of $\mathbf{V}(n)$ is identical with n times the unit translation along the X axis.

(B') As a special case of (B), $\sum_p \mathbf{V}_p(1)=(1, 0, 0)$.

The problem of unravelling a given vector set is to determine indices p and n of all $\mathbf{V}_p(n)$ vectors. The indexing of p and n can be done by use of the recurrence relations (A). In reality, it is sufficient to determine the index p of all $\mathbf{V}_p(1)$ vectors, because it is possible to deduce the indices p of $\mathbf{V}_p(n)$ for $n>1$ from those of $\mathbf{V}_p(1)$. The indexing of p for $\mathbf{V}_p(1)$ can be done by successively utilizing the recurrence formulae for $\mathbf{V}(n)$ up to a certain value $n=q$. This critical q value depends upon each vector set. It is generally as small as 2 as shown later in Examples 2 and 3 but sometimes as large as 7 as shown in Example 4.

In the following, systematic procedures for unravelling are described in a general form, though there may be a more efficient strategy with minor modifications depending upon each problem. It will be shown that the above general relations (A), (A'), (B) and (B') are a set of clues satisfactory enough to unravel a given periodic vector set, after vectors being arranged in the order of increasing X components.

(I) *The way of determining the* $\mathbf{V}(1)$ *vectors*

(I*a*) *Picking up* $\mathbf{V}(1)$ *vectors.* There are N $\mathbf{V}(0)$ vectors at the origin, which need not be searched for. Apart from them, the $\mathbf{V}(1)$ vectors will be found mostly among those with small X components. The relation (A') gives the useful information that if a member of a vector set proves to be different from the sum of any non-zero vectors, this member is certainly a $\mathbf{V}(1)$ vector. Therefore, the two vectors with smallest X components are certainly $\mathbf{V}(1)$ vectors. Other $\mathbf{V}(1)$ vectors can be picked up by testing one by one whether each vector is different from the sum of any other vectors with smaller X components.

In a favourable case, all $\mathbf{V}(1)$ vectors may be picked up uniquely in this way. If N $\mathbf{V}(1)$ vectors thus picked up prove to satisfy the relation (B'), we can go on to procedure (II).

(I*b*) *Searching for hidden* $\mathbf{V}(1)$ *vectors.* In an unfavourable case, it may happen that some genuine $\mathbf{V}(1)$ vectors are hidden by accidental coincidences with the sum of other vectors, as was mentioned in the relation (A'). If this is the case, the number of $\mathbf{V}(1)$ vectors which can be picked up in the above way is less than N, and naturally the X component of their total sum is less than unity. Since, however, the number of hidden $\mathbf{V}(1)$ vectors is known, they can be picked up from **V**'s in the vector set by testing all possible combinations satisfying the relation (B'). If it happens that more than one set passed through this test, then, all sets found out are to be subject to the following procedures.

(II) *Finding out* $\mathbf{V}(2)$ *vectors by a consistency check*

(II*a*) *Picking up* $\mathbf{V}(2)$ *vectors.* When a vector is found to be the sum of two $\mathbf{V}_p(1)$ vectors, it is clear from relation (A) that the former is one of $\mathbf{V}(2)$ vectors, provided there is no haphazard coincidence, and, at the same time, that the latter two vectors are neighbours among the $\mathbf{V}(1)$ vectors.

(II*b*) *Check by the number of vectors picked up.* If the $\mathbf{V}(2)$ vectors picked up in the above are N in number, we can go on to procedure (II*c*). It is, however, likely that there will be more than N because of

haphazard coincidences. In such cases, all possible combinations of N **V**(2) vectors are chosen from them, and every set thus obtained is subject to procedure (II*c*). If, on the contrary, there are less than N, the relevant set of **V**(1) vectors is considered to be an unreasonable one and should be discarded.

(II*c*) *Check by the sum of all vectors.* A set of N **V**(2) vectors should be discarded if it does not satisfy relation (*B*). If a set survives, it should be subjected to procedure (II*d*).

(II*d*) *Check of consistency.* Information about every neighbouring pair of **V**(1) vectors is already known in procedure (II*a*). In a favourable case, this information is enough to determine uniquely a serial arrangement of all **V**(1) vectors making 'a closed ring', in which each **V**(1) vector takes part once. In this case, we go on to procedure (IV), skipping procedure (III). When, on the other hand, a surviving set is easily found to make no closed ring, such a set can naturally be discarded.

In case a vector set includes many vectors with the same components, the set often has the possibility of generating various closed rings. In such an unfavourable case, it is better to go on to procedure (III) rather than examine whether each candidate set can really form any closed ring.

(III) *Finding out and checking with* **V**(n) *vectors where* $n>2$.

The procedure similar to (II) can be used for finding out **V**(3) vectors. In the procedure corresponding to (II*a*), the **V**(3) vectors are found as sums of **V**(2) and **V**(1), and this information gives neighbouring relations of three successive **V**(1) vectors. While the procedures analogous to (II*b*), (II*c*) and (II*d*) are carried out, some sets of **V**(1) vectors are discarded, and others employed. The latter is to be subjected either to the procedure (IV) or to the similar procedure as above but with **V**(4) vectors. Further procedures should be repeated with **V**(n) of higher and higher degree as needed. Thus some sets are discarded, and others are sooner or later subjected to procedure (IV).

(IV) *Construction of a fundamental set*

Coordinates for members $\mathbf{v}_i(x_i, y_i, z_i)$ of a fundamental set are now easily obtained by adding **V**(1) vectors one by one in the order in which a closed ring is formed. If, as a whole, only one fundamental set is obtained, it is certainly a right solution. When more than one fundamental set is obtained, however, it may be preferable to test further whether each relevant vector set is consistent with all **V**(n) vectors of higher degree, by reproducing a vector set from an obtained fundamental set and by comparing it with the given one. If more than one set should survive all through the tests to the end, it means that one of these homometric solutions is what is really being sought. The unravelling is thus finished. The coordinate origin of the fundamental set obtained does not matter. As is easily done, however, it may be preferable to choose the origin at the special position according to the symmetry of the fundamental set obtained.

Application

In this section, it is shown by several examples of application that the above mentioned procedures are really effective in unravelling various types of vector sets. Example 1 deals with a very simple two-dimensional model which indicates how each procedure works. Other examples are more complicated. Among them, Examples 2 and 3 are concerned with three-dimensional vector sets, the former with hypothetical vector sets and the latter with a real one. Example 4 is concerned with a real one-dimensional vector set which seems difficult to unravel by the usual image seeking method because of heavy superposition. In Examples 2 and 3, most of the procedures were carried out on an electronic computer.

Example 1. *Illustration of the present method by a hypothetical two-dimensional example*

(*a*) *Diagrammatical procedures*

Let us consider a periodic fundamental set with five points in a unit cell as shown in Fig. 1(*a*). **V**(1) and **V**(2) vectors are shown in Fig. 1(*b*) and (*c*) respectively, where the horizontal axis is taken as the special x axis. The problem is to unravel the corresponding vector set shown in Fig. 2(*a*) where $20=5(5-1)$ field peaks are found. All five **V**(1) vectors are uniquely found out by procedure (I*a*), as shown as **A**, **B**, **C**, **D** and **E** in Fig. 2(*b*). It may be unnecessary to confirm that they satisfy condition (*B'*). Five **V**(2) vectors are then found out by procedure (II*a*); these are shown as **F**, **G**, **H**, **I** and **J** in Fig. 2(*c*). This has been done because they satisfy the property (*A*): namely, there exist relations $\mathbf{F}=\mathbf{A}+\mathbf{B}$, $\mathbf{G}=\mathbf{B}+\mathbf{D}$, $\mathbf{H}=\mathbf{C}+\mathbf{D}$, $\mathbf{I}=\mathbf{C}+\mathbf{E}$ and $\mathbf{J}=\mathbf{A}+\mathbf{E}$. This information already uniquely gives a closed ring (**ABDCE**) without taking **V**(n) vectors with $n\geq 3$ into consideration. Thus the ordering of **V**(1) vectors has already been determined. Finally, if only we combine these **V**(1) vectors one after another corresponding to procedure (IV), then the fundamental set shown in Fig. 1(*a*) is reconstructed.

(*b*) *Numerical procedures*

The unravelling procedures on an electronic computer will be used for Examples 2 and 3. Though the present example is naturally too simple to be unravelled on a computer, an outline of the programming to be used for complicated cases will be easily understood by the following numerical illustration.

The fundamental set in the present example is expressed, for brevity, in two decimal digits, as follows:

·05, ·15 ·52, ·28 ·15, ·70 ·30, ·35 ·70, ·85

From these values, a vector set is produced as follows:

0 0	·47, ·13	·10, ·55	·25, ·20	·65, ·70
·53, ·87	0 0	·63, ·42	·78, ·07	·18, ·57
·90, ·45	·37, ·58	0 0	·15, ·65	·55, ·15
·75, ·80	·22, ·93	·85, ·35	0 0	·40, ·50
·35, ·30	·82, ·43	·45, ·85	·60, ·50	0 0

All members of the vector set except **V**(0) vectors are arranged in the order of increasing X components as follows:

·10, ·55	·53, ·87
·15, ·65	·55, ·15
·18, ·57	·60, ·50
·22, ·93	·63, ·42
·25, ·20	·65, ·70
·35, ·30	·75, ·80
·37, ·58	·78, ·07
·40, ·50	·82, ·43
·45, ·85	·85, ·35
·47, ·13	·90, ·45

Five **V**(1) vectors can be found out by procedure (I*a*) as follows:

·10, ·55 ·15, ·65 ·18, ·57 ·22, ·93 ·35, ·30

All pairs of successive **V**(1) vectors are found by the relations as follows:

(·10, ·55)+(·15, ·65)=(·25, ·20)
(·15, ·65)+(·22, ·93)=(·37, ·58)
(·22, ·93)+(·18, ·57)=(·40, ·50)
(·18, ·57)+(·35, ·30)=(·53, ·87)
(·35, ·30)+(·10, ·55)=(·45, ·85)

This information uniquely determines a closed ring [procedure(II*d*)] as follows:

(·10, ·55)−(·15, ·65)−(·22, ·93)
−(·18, ·57)−(·35, ·30)−(·10, ·55)

Apart from a relative shift (·05, ·15), the fundamental set can be reconstructed (procedure (IV)) as follows:

0 0 ·10, ·55 ·25, ·20 ·47, ·13 ·65, ·70

Example 2. *Unravelling of a hypothetical three-dimensional vector set by use of a computer.*

In order to test the practicability of the present method in more complicated cases, some hypothetical three-dimensional vector sets were unravelled on an

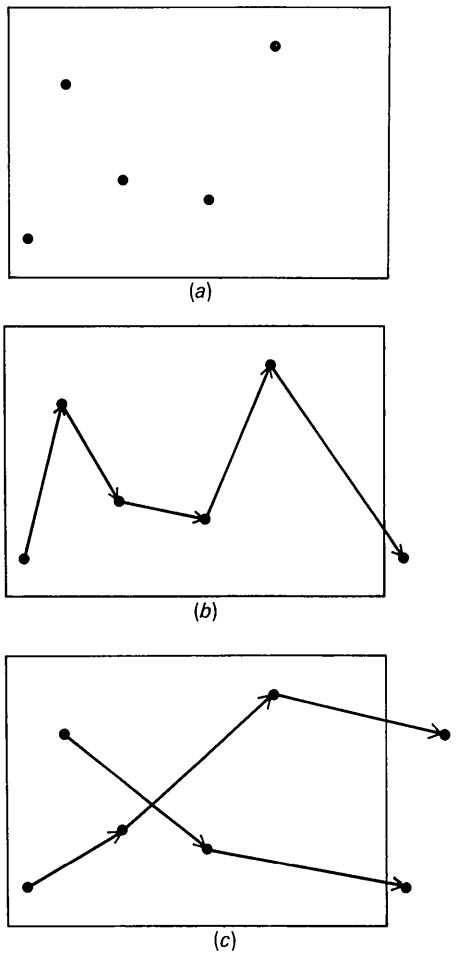

Fig. 1. Fundamental set.

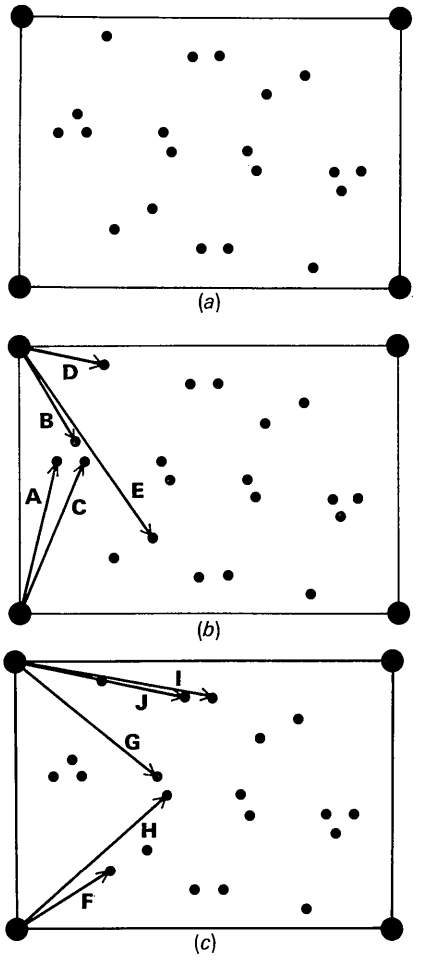

Fig. 2. Vector set.

electronic computer. The computer used was FACOM-202 which has about 8000 core memories and operation time of about 60 μsec. Atoms in the fundamental sets assumed in these examples were 15 to 55 in number, so that the relevant vector sets contained 225 to 3025 points per unit cell. Coordinate values of atoms in the fundamental sets were taken from a table of random numbers and were expressed by 4 decimal digits.

(*a*) A program was devised such that the computer at first produced a vector set from a given fundamental set, then unravelled the vector set obtained, following the above-mentioned procedures. Because errors were not taken into account and, in addition, the coordinates were randomly given, it was expected that only a few, if any, haphazard coincidences would occur. Therefore the program was as simple as follows.

The program was to find **V**(1) vectors by the relation (*A'*) [procedure (I*a*)], to take all vectors satisfying (*A*) as **V**(2)'s [procedure (II*a*)], then to reconstruct a fundamental set [procedure (IV)]. More details of the program will be understood from the illustration in Example 1(*b*).

It was confirmed that the given sets were successfully unravelled by use of this program. Naturally the necessary machine time including printing depended upon the number of atoms, as shown in Table 1.

Table 1. *Machine time for unravelling a vector set without errors*

Number of point atoms	25	30	40	50	55
Machine time in minutes	0·5	1	4	6	7

(*b*) Since the example in (*a*) above without errors might be too hypothetical, similar computations were carried out by allowing for a certain amount of error in the vector components.

A program was devised which first produced the components for each vector in a vector set, then introduced in each component an error. Four sets were derived, each with an error as indicated in Table 2, and they were subjected to procedures (I*a*), (I*b*), (II*a*) and (II*d*). When a closed ring could be easily formed at procedure (II*d*), procedure (IV) immediately followed. In order to find other candidate sets, the similar procedures were repeated from (I*b*), until possible combinations of vectors for the hidden $V(1)$ vectors were exhausted. Procedure (III) was entirely omitted.

The machine time for each case is shown in Table 2, where the entries marked with * mean that the number of trials prohibitively increased because of coincidences due to errors introduced. On the other hand, for the cases corresponding to the filled entries, the machine time was not very much longer than in the case without errors. This is mainly because the program used was more elaborate than in (*a*).

Every time more than N successive pairs of **V**(1) vectors were found by procedure (II*a*) in the above, all pairs were printed out, and were later examined by hand to ascertain whether N pairs from among them could form a closed ring. This procedure, even in the most complicated case, took about 10 minutes until the set proved either to form uniquely a closed ring or to form none.

Table 2. *Machine time for unravelling a vector set with errors*

Number of point atoms		15	20	30	40	50	55
Machine time in minutes	error ±1/64†	0·5	*				
	error ±1/128	0·3	1·3	*			
	error ±1/256			2·5	*		
	error ±1/512				5·5	10	17·3

† Each figure means the range of errors randomly introduced into components of member vectors.

Example 3. *Unravelling of the vector set for a real structure.*

It was attempted to unravel a vector set based on a known structure in contrast with a hypothetical one, because random values of coordinates used in Example 2 might make the problem unduly easy. In order to make this point clear, (+)-hetisine hydrobromide, $C_{20}H_{27}O_3N$. HBr (Przybylska, 1963) was chosen as a real example, because it has a relatively low symmetry and a suitable number of atoms. When hydrogen atoms are disregarded, this crystal with the space group $P2_1$ includes two molecules of $C_{20}O_3NBr$ in a unit cell, which are related with each other by a screw axis. In the present treatment, however, the symmetry 2_1 was not taken into account and each of 50 point atoms in a unit cell was assumed to have an equal weight.

The same program as was used in Example 2(*a*) was applied to this problem, namely the machine at first produced a vector set from the above simplified structure, then unravelled it. Two tries were carried out, each by taking either the original a or b axis as the special x axis. Although the b axis is parallel to the 2_1 axis and the a axis is not, the machine time for unravelling was about 7 minutes in each case.

In the former case, **V**(1) vectors include two vectors of $(X, \frac{1}{2}, Z)$ type and 24 pairs of (X, Y, Z) and $(X, -Y, Z)$. In the latter case, **V**(1) vectors are 25 pairs of (X, Y, Z) and $(-X, Y, -Z)$ type. When the space group is known beforehand and these features are taken into account in a program, the time for unravelling a vector set will be more or less reduced.

Example 4. *An analysis of a real structure*, 96*R*−SiC

One of the authors has recently solved rhombohedral SiC with 96 layers in a unit cell (Tokonami, 1965). Example 4 deals with the unravelling of a vector set corresponding to this one-dimensional structure. As is well known, SiC with a long period is composed of identical SiC layers whose mode of stacking is expressed by the *ABC* notation. Since, in the present case, 32 layers are cyclically repeated three times accor-

Table 3. *A list of a vector set*

Vectors with component 49–95 are omitted, because there are two centers of symmetry at vectors with component 0 and 48.

Component	Number of vectors*	Component	Number of vectors	Component	Number of vectors
0	32	17	8	33	17
1	0	18	18	34	10
2	10	19	6	35	10
3	12	20	8	36	12
4	10	21	14	37	13
5	4	22	10	38	0
6	26	23	10	39	23
7	0	24	14	40	6
8	12	25	9	41	10
9	12	26	6	42	12
10	10	27	15	43	12
11	6	28	10	44	2
12	22	29	10	45	23
13	3	30	12	46	4
14	10	31	15	47	11
15	13	32	0	48	12
16	10				

* Number of vectors corresponds to a peak height in the Patterson map.

ding to its rhombohedral symmetry, the structure is determined by the position of, say, *A* layers only. Therefore, it is sufficient to consider a hypothetical linear crystal of 96 periodic sites, among which 32 sites are occupied by point atoms. The weight of each vector in the relevant vector set given in Table 3 was obtained from the fundamental set based on the solved structure. The figures listed mean, therefore, idealized heights of the sharpened Patterson peaks corresponding to point atoms.

The vector set consists of 1024 vectors in a unit cell, including 32 **V**(0) vectors. It is noted here that a notation **n** will hereafter be used instead of a vector with component $n/96$.

(1) *Determination of* **V**(1) *vectors*

All of ten **2** and twelve **3** vectors, which cannot be decomposed as the sum of shorter vectors, are obviously **V**(1) vectors [procedure (I*a*)]. The rest of **V**(1) vectors should be $32-10-12=10$ in number, and their sum must be $96-(2\times 10+3\times 12)=40$. Therefore, ten **4** vectors were uniquely identified to be $V(1)$'s. These results are listed in Table 4.

Table 4. **V**(1) *vectors*

Component	Number of vectors
2	10
3	12
4	10

(2) *Determination of* **V**(2) *vectors*

There are six possibilities of combining two **V**(1) vectors as follows [procedure (II*a*)]:

$$\mathbf{2}+\mathbf{2}=\mathbf{4}, \quad \mathbf{2}+\mathbf{3}=\mathbf{5}, \quad \mathbf{2}+\mathbf{4}=\mathbf{6},$$
$$\mathbf{3}+\mathbf{3}=\mathbf{6}, \quad \mathbf{3}+\mathbf{4}=\mathbf{7}, \quad \mathbf{4}+\mathbf{4}=\mathbf{8}.$$

Let numbers of the **V**(2) vectors corresponding to these combinations be $a(\mathbf{4})$, $b(\mathbf{5})$, $c(\mathbf{6})$, ..., $f(\mathbf{8})$ respectively. Now that the shortest 32 vectors next to **V**(1) vectors in Table 3, namely all **5** and **6** vectors and two out of twelve **8** vectors, make just 2×96 in total, they are **V**(2) vectors. It follows from Tables 3 and 4 that

$$a(\mathbf{4})=0,\ b(\mathbf{5})=4,\ c(\mathbf{6})+d(\mathbf{6})=26,$$
$$e(\mathbf{7})=0 \text{ and } f(\mathbf{8})=2.$$

Since 32 **V**(1) vectors should be included in 32 **V**(2) vectors twice for each, it holds that

$$2a(\mathbf{4})+b(\mathbf{5})+c(\mathbf{6})=10\times 2,\ b(\mathbf{5})+2d(\mathbf{6})+e(\mathbf{7})$$
$$=12\times 2 \text{ and } c(\mathbf{6})+e(\mathbf{7})+2f(\mathbf{8})=10\times 2.$$

These eight simultaneous linear indeterminate equations should have non-negative integral roots. The solutions are:

$$a(\mathbf{4})=0,\ b(\mathbf{5})=4,\ c(\mathbf{6})=16,\ d(\mathbf{6})=10,$$
$$e(\mathbf{7})=0 \text{ and } f(\mathbf{8})=2,$$

as shown in Table 5.

Table 5. **V**(2) *vectors*

Component	Composition	Number of vectors
5	**2+3**	4
6	**2+4**	16
6	**3+3**	10
8	**4+4**	2

(3) *Determination of* **V**(3) *vectors*

Each **V**(3) vector can be considered as a combination of two **V**(2) vectors which include a **V**(1) vector in common. Thus, if we combine two vectors, for instance **2+3** and **2+4**, so as to make a vector **3+2+4**, then the latter vector is a **V**(3) vector. Two vectors **2+3** make either **2+3+2** or **3+2+3**. In this way, it was easily found from the composition in Table 5 that the following nine combinations are possible for **V**(3) vectors:

2+3+2=7, 3+2+3=8, 2+4+2=8,
2+3+3=8, 3+2+4=9, 3+3+3=9.
4+2+4=10, 2+4+4=10, 4+4+4=12.

Let the numbers of these combinations be $a', b', \ldots, i'$ respectively. Now the sum of the shortest 32 vectors next to **V**(2) vectors is just 3×96, therefore, these are **V**(3) vectors. From Tables 3, 4 and 5, it follows that

$$a'=0,\ b'+c'+d'=10,\ e'+f'=12,\ g'+h'=10 \quad \text{and } i'=0.$$

On the other hand, each **V**(2) vector must appear twice. Therefore,

$$2a'+2b'+d'+e'=4\times 2,$$
$$2c'+e'+2g'+h'=16\times 2,$$
$$d'+2f'=10\times 2,$$
$$\text{and } h'+2i'=2\times 2.$$

Two sets of solutions were obtained from these nine equations, and are shown in Table 6.

Table 6. **V**(3) *vectors*

Component	Composition	Number of vectors	
8	3+2+3	0	3
8	2+4+2	6	7
8	2+3+3	4	0
9	3+2+4	4	2
9	3+3+3	8	10
10	4+2+4	6	6
10	2+4+4	4	4

(4) *Determination of* **V**(4) *vectors*

There are the following ten possible combinations for **V**(4) vectors:

2+3+3+2, 2+3+3+3, 3+3+2+3,
2+4+2+3, 3+3+2+4, 2+4+2+4,
3+3+3+3, 2+4+4+2, 3+2+4+4,
4+4+2+4.

Possible sets of **V**(4) vectors satisfying the relation (*A*) are 6, 22, 2, 2, 0 or 6, 22, 3, 0, 1, in number, of the vectors **11**, **12**, **13**, **14** and **15** respectively. However, the latter cannot be allowed because the longest vector should be **14** in the above listed ten. Therefore, the indeterminate equations for the former are to be solved in both cases corresponding to the two sets of **V**(3) vectors. In the first case no solution was found, and the second case gave only one solution shown in Table 7.

Table 7. **V**(4) *vectors*

Component	Composition	Number of vectors
11	2+3+3+3	4
11	2+4+2+3	2
12	3+3+2+4	4
12	2+4+2+4	10
12	3+3+3+3	6
12	2+4+4+2	2
13	3+2+4+4	2
14	4+4+2+4	2

(5) *Consistency among higher order vectors*

V(5), **V**(6) and **V**(7) vectors were obtained in a similar way, as shown in Tables 8, 9 and 10, where only the sets marked with * could survive. Finally, **V**(7) vectors uniquely determined the following arrangement of 32 layers:

2+3+3+3+2+4+2+4+2+4+2+4+4+2+3
+3+3+3+3+3+3+3+2+4+4+2+4+2
+4+2+4.

Table 8. **V**(5) *vectors*

Component	Composition	Number of vectors *				
13	2+3+3+3+2	1	1	1	1	1
14	2+3+3+3+3	2	2	2	2	2
14	2+4+2+3+3	0	1	2	3	4
14	2+4+2+4+2	6	5	4	3	2
15	3+3+3+2+4	4	4	4	4	4
15	3+2+4+2+4	4	3	2	1	0
15	3+3+3+3+3	5	5	5	5	5
15	2+4+4+2+3	0	1	2	3	4
16	3+3+2+4+4	4	3	2	1	0
16	4+2+4+2+4	2	3	4	5	6
16	2+4+2+4+4	0	1	2	3	4
16	2+4+4+2+4	4	3	2	1	0

Table 9. **V**(6) *vectors*

Component	Composition	Number of vectors *		
17	2+3+3+3+2+4	2	2	2
17	2+3+3+3+3+3	2	2	2
17	2+4+2+3+3+3	4	3	2
17	2+4+2+4+2+3	0	1	2
18	3+3+3+3+2+4	2	2	2
18	3+3+2+4+2+4	0	1	2
18	2+4+2+4+2+4	8	7	6
18	3+3+3+3+3+3	4	4	4
18	2+4+4+2+3+3	4	3	2
18	2+4+2+4+4+2	0	1	2
19	3+2+4+2+4+4	4	2	0
19	3+2+4+4+2+4	0	1	2
19	3+3+3+2+4+4	0	1	2
20	4+2+4+2+4+4	0	1	2
20	4+2+4+4+2+4	2	1	0

Table 10. **V**(7) *vectors*

Component	Composition	Number of vectors *	
19	2+3+3+3+2+4+2	2	2
20	2+3+3+3+3+3+3	2	2
20	3+3+3+3+2+4+2	0	1
20	3+3+2+4+2+4+2	2	3
20	2+4+2+4+2+4+2	2	0
21	4+2+3+3+3+2+4	1	1
21	3+3+3+3+3+2+4	2	2
21	3+3+3+2+4+2+4	2	1
21	3+2+4+2+4+2+4	2	1
21	3+3+3+2+4+4+2	2	3
21	3+3+3+3+3+3+3	3	3
21	3+2+4+4+2+4+2	2	3
22	3+3+3+3+2+4+4	2	1
22	4+2+4+2+4+2+4	2	4
22	2+4+2+4+2+4+4	2	3
22	3+3+2+4+4+2+4	2	1
22	4+2+4+2+4+4+2	2	1

It was confirmed that this structure satisfies the relation (*A*) for $\mathbf{V}(n)$ of higher order and also that this can reproduce the assumed fundamental set. This confirmation means that the vector set in Table 3 was right or at least self-consistent. No homometric mate was found in this example.

Characteristics of the present method

In this section, merits and demerits of the present method are described.

(*a*) If the problem is concerned with unravelling a vector set to get a fundamental set, the following characteristics are to be noted.

The image seeking method makes use of a geometrical pattern of a vector set. The present method, on the other hand, is an entirely analytical one which is more suitable for putting in a computer.

While the image seeking method may be conveniently applied to two-dimensional work, the present method can be applied even more effectively to a three-dimensional case without any additional complexity. *Y* and *Z* components can obviously make each unravelling procedure very efficient, as if they gave a triple precision to an *X*-component of each vector.

The present method allows us to choose a special axis *x* or *X* in different ways. This axis has only to be any crystallographic zone axis. In case the coordinates of vectors include errors, it is quite possible to facilitate the procedures of unravelling by trying several ways of choosing the *X*-axis.

If a vector set to be unravelled includes points with heavy coincidences as in Example 4, the present method is far more powerful than the image seeking method and its variations.

Although information about a partial structure, for example, about a benzene ring, any radical or about an interatomic distance between heavy atoms, can be a valuable clue for unravelling a structure by the image seeking method, it cannot readily be utilized in the present method. This method, however, is effective for structures such as contain no heavy atoms or have an unknown molecular form.

Because all atoms in a unit cell were taken to be independent in the present work, it does not seem convenient to make full use of the information known about the symmetry elements of the relevant space group. However, the procedures will probably be accelerated to some extent, if the general program is modified so as to take this elaboration into account.

(*b*) The following comments should be made as to the applicability of the present method to a Patterson map, although these will be discussed later from a wider point of view.

When there is a broad peak consisting of many unresolved peaks, only the information about the number of individual peaks is sufficient to unravel the vector set. For the present method to be effective, however, all true peaks in a given map should be found out correctly, though coordinates of every point are allowed to have errors to some extent. Even a few false peaks mistaken for true ones will make the unravelling almost impossible. On the other hand, the image seeking method is not much disturbed by so few mistakes.

(*c*) Lastly, it should be mentioned that the present method may be quite useful to check quickly, preferably on a computer, whether a solved structure has any homometric mate or not.

Theoretical aspects of the phase problem

In this section, some general aspects of the phase problem as seen from the standpoint of the present work, will be mentioned.

The procedure of the structure analysis is outlined by a scheme as shown in Fig. 3, where solid arrows represent operations which can be readily carried out and broken arrows those which can not. The Patterson function $P(\mathbf{r})$ is readily obtained from all observed data, $|F(\mathbf{h})|^2$, and the electron distribution $\varrho(\mathbf{r})$ can be readily obtained from $F(\mathbf{h})$ values including their phases. In order to get $\varrho(\mathbf{r})$, it is necessary either to find phases of $F(\mathbf{h})$'s by any of the direct methods or to unravel the $P(\mathbf{r})$ function. The former procedure is the phase problem in a narrower sense. The latter, however, naturally is equivalent to the former.

The electron density $\varrho(\mathbf{r})$ in a crystal is usually expressed by a sum of electron clouds around centers of atoms. The structure factor can, therefore, be expressed as $F(\mathbf{h})=\sum_j f_j \exp(2\pi i \mathbf{r}_j \cdot \mathbf{h})$ where f_j is the atomic structure factor of the jth atom including its temperature factor, and $\mathbf{r}_j$ is the relevant position vector. Since the function f_j is more or less proportional to an average structure factor f, $F(\mathbf{h})$ may be approximated as

$$F(\mathbf{h})=f\sum_j n_j \exp(2\pi i \mathbf{r}_j \cdot h) \qquad (4)$$

where n_j means the weight of the nth atom.

Let us consider in the following the case in which the use of the unitary structure factor, $U(\mathbf{h})=F(\mathbf{h})/f\sum_j n_j$, can be justified as in many direct methods. As shown by the scheme in Fig. 4, the Fourier trans-

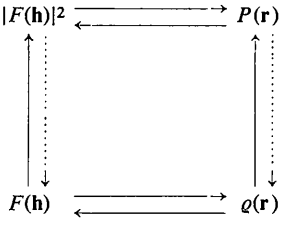

Fig. 3. Structure analysis procedure.

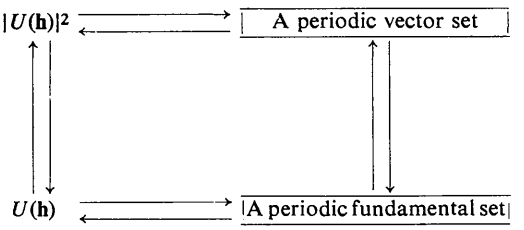

Fig. 4. Fourier tranforms of $U(\mathbf{h})$ and $|U(\mathbf{h})|^2$.

form of the $U(\mathbf{h})$'s is a periodic fundamental set including weights for atoms and that of the $|U(\mathbf{h})|^2$'s is the relevant periodic vector set.

In the present paper, it has been shown at least in principle that a periodic vector set can be unravelled even if it involves heavy coincidences. This situation does not change even if point atoms in a fundamental set are different in weight. Therefore, the operation to obtain the fundamental set from a vector set can be represented by a solid arrow in Fig. 4. In other words, the operation to obtain $U(\mathbf{h})$'s from $|U(\mathbf{h})|$'s should, in principle, give no difficulty.

Practical limitations, therefore, come from the following two points:

(1) The values of $|U(\mathbf{h})|$ can be obtained only around the origin of reciprocal space. Even inside the limiting sphere, $|F(\mathbf{h})|$'s may be too small to be observed.

(2) The values $|F(\mathbf{h})|$ experimentally obtained are inevitably influenced by a certain number of errors. Besides, the expression (4) can never exactly hold. The form factor f_j cannot always be satisfactorily accurate. Even the same kind of atom may have different temperature factors, and some form factors may have different shapes or even anisotropy. For these reasons, the values of $|U(\mathbf{h})|$ bear a fairly large amount of error.

The limitation (1) is essentially concerned with the problem of the resolution due to the termination of the Fourier or Patterson synthesis. It is generally believed that the resolution in the Fourier map and the Patterson map is of the order of the wavelength of the radiation used, and that over-sharpening is harmful because of increasing ripples. However, a knowledge of the shape of electron distribution in each atom is usually available beforehand with reasonable accuracy. Thus, if we take this knowledge into account, the resolution can be substantially higher than usually considered. As will be discussed elsewhere in detail, the limitation (1) does not seem to be so severe.

As regards limitation (2), the problem will be understood in the following way. A set of accurate values of $|U(\mathbf{h})|$ in the observable region usually offer redundant information for uniquely determining the structure. If errors accompanying them become larger, the redundancy becomes smaller. If errors as a whole exceed critical values, a set of $|U(\mathbf{h})|$'s either has no plausible solution, or occasionally give a false structure, so that the structure analysis becomes impossible. However, it is difficult to assess the critical values of errors.

In favourable cases, it is possible to recover the true values of $|U(\mathbf{h})|$ from the observed $|U(\mathbf{h})|$'s which are disturbed by errors. In the case of 96*R*-SiC shown in Example 4, SiC layers have fractional coordinates of special values, multiples of 1/96, and the $U(\mathbf{h})$ values of this crystal are therefore periodic in the reciprocal space. In such a special case, it can be shown that there is a way to recover the true values of $|U(\mathbf{h})|$ even if the errors are rather large. The details will be reported in another paper (Tokonami, 1965).

It will be understood how the problem is made easier by information about electron distribution around each atom or about the atomic structure factor as well as by accurate observed values of $|F(\mathbf{h})|$. It may be worth while to investigate under what conditions the crystal analysis becomes impossible because of the two limitations.

The authors are very much indebted to Prof. Miyake for his interest and encouragement.

References

BUERGER, M. J. (1950). *Acta Cryst.* **3**, 87.
BUERGER, M. J. (1959). *Vector Space*. New York: Wiley.
PRZYBYLSKA, M. (1963). *Acta Cryst.* **16**, 871.
TOKONAMI, M. (1965). To be published in *J. Min. Soc. Japan.*
WRINCH, D. M. (1939). *Phil. Mag.* **27**, 98.