Acta Crystallographica Section A
Foundations of Crystallography
ISSN 0108-7673

Received 8 October 2004
Accepted 23 December 2004

On the algebra of binary codes representing close-packed stacking sequences

Ernesto Estevez-Rams,[a]* Cristy Azanza-Ricardo,[a] Jorge Martínez García[a] and Beatriz Aragón-Fernández[b]

[a]Institute for Materials and Reagents (IMRE), University of Havana, San Lazaro y L, CP 10400, C. Habana, Cuba, and [b]Universidad de Ciencias Informáticas, Carretera San Antonio de los Baños, km 2 1/2, Torrens, Boyeros, C. Habana, Cuba. Correspondence e-mail: estevez@imre.oc.uh.cu

A systematic use of binary codes derived from the Hägg symbol are used to study close-packed polytypes. Seitz operators acting over the corresponding binary codes are defined and used. The number of non-equivalent polytypes of a given length are calculated through the use of the Seitz operators. The same procedure is applied to the problem of counting the number of polytypes complying with a given symmetry group. All counting problems are reduced to an eigenvector problem in the binary code space. The symmetry of the binary codes leads to the different space groups to which polytypes can belong.

1. Introduction

The densest packing of spheres in three dimensions is known to be (and recently proved by Hales, 2002) the so-called close packing of spheres. In a close-packing arrangement, hexagonal layers of spheres are stacked in one direction in such a way that the spheres on one layer rest over the voids of the layers below and above it (Pandey & Krishna, 2004).

In crystallography, the interest in close packing derives from the fact that many structures can be described as such. The simplest examples of periodic close-packed structure are the face-centred cubic and the hexagonal close packed. Both arrangements are also the only ones where every layer is crystallographically equivalent to any other layer in the stack.

The close-packed hexagonal layer can be stacked in more complex arrangements than the face-centred cubic and the hexagonal close packed. For a given number of layers P, there will be a finite number of possible, crystallographically non-equivalent, stackings. We will call this number $\Gamma(P)$. Each stacking sequence counted in $\Gamma(P)$ will represent a different close-packed polytype.

The problem of finding $\Gamma(P)$ has been solved by a number of authors by exhaustive generation and counting (Brown & Bailey, 1962; Lima de Faria & Figueiredo, 1969; Smith & Rinaldi, 1962; Thompson & Downs, 2001). Exhaustive generation is an unrealistic approach for large periodicities as the number of trial structures grows as 2^P. Other methods avoiding direct enumeration have been used, most notably by Iglesias (1981) through the use of combinatorial techniques and McLarnan (1981) through group-theoretical methods.

It is also known that all monoatomic close-packed arrangements can be classified into eight possible space groups: $P3m1$, $P\bar{3}m1$, $P\bar{6}m2$, $P6_3mc$, $P6_3/mmc$, $R3m$, $R\bar{3}m$ and $Fm3m$. The deduction of the possible space groups is usually referred to the classical paper of Belov (taken from Pandey & Krishna, 2004, and references therein).

Close-packed arrangements can be coded into binary sequences in different ways (Verma & Krishna, 1966). The two most used ones are the following.

(i) To code each layer according to its surroundings. According to the nearest neighbours, two surroundings are possible, one called the hexagonal surrounding, where the layers above and below are in equivalent positions, and the cubic surrounding, where the layers above and below are in different positions.

(ii) The binary code is derived from the Hägg symbol, where to any pair of layers a $+$ or $-$ code can be assigned if the pair forms a 'forward' sequence or a 'backward' sequence.

In any case, the coding of close-packed arrangements into a binary sequence allows one to translate the study of close packing into the realms of coding theory (and related fields of number theory and combinatorics) and to define a proper algebra acting over the binary codes.

In this paper, we will explore the symmetry and other properties of periodic close packing through the use of its binary code. Seitz operators will be defined acting over the binary codes taken as vectors. The introduction of the Seitz operators will allow calculation of $\Gamma(P)$ as an eigenvector problem in a straight-forward manner. The possible space groups of close-packed sequences will be viewed as a result of the possible symmetries of the binary codes. We will also make use of the Seitz operators to calculate the number of polytypes complying with a specific symmetry group. The examples studied in the paper will prove the power of the binary coding and its algebraic approach to the study and understanding of close packing.

2. Close packing

In a close-packed structure, hexagonal layers can occupy three different locations, usually labelled A, B and C (Verma & Krishna, 1966). The stacking arrangement can be coded indicating for each layer which position is occupied (*e.g.* ABACABACABAC). In the case of periodic stacking, it will suffice if we represent the code of a repeating unit of the stack (for the above example ABAC). Translational periodicity also implies that where we start the code is irrelevant, so two distinctive codes that can be brought to coincidence by a cyclic rotation of one of the codes represent the same arrangement (*e.g.* ABAC $\rightarrow$ ACAB). In terms of the letter code, the close-packed condition is the impossibility of having the same letter consecutively in the code. The close-packed condition together with the translational periodicity also impedes a code starting and ending with the same letter.

As the labels are irrelevant as long as consistency is kept, any permutation of two letters will give rise to the same arrangement, so a family of codes of the same length that can be brought into coincidence by a permutation of letters will represent the same stacking (*e.g.* ABCB $\rightarrow$ BACA).

The Hägg code is a binary notation that can be derived from the ABC notation (Pandey & Krishna, 2004). In terms of coding, we can translate the letter code to a Hägg code following the rule: AB, BC and CA to $+$ and AC, CB, BA to $-$. In such a way, a binary code can be constructed for any stacking sequence (*e.g.* ACABC $\rightarrow -++++$). Instead of using $+$ and $-$ signs, the binary code can be replaced by 0's and 1's (*e.g.* $-++++ \rightarrow$ 01111). This binary coding has been used for Monte Carlo modelling of polytype phase transformation (see for example Shrestha & Pandey, 1996, 1997). In this paper, the binary code will be represented between $|>$ (*e.g.* $|01111>$) or, in the case that the binary representation is too long, by its decimal equivalent, stating explicitly the length of the code in case ambiguity arises *e.g.* $|01111> \rightarrow |15>_5$.

The equivalence conditions given by the letter permutations or the cyclic shift can now, in terms of the binary code, be given by the following rules:

1. Two binary codes are equivalent if they can be brought into coincidence by permuting the two digits (*e.g.* $|101100> \hat{=} |010011>$, where $\hat{=}$ means crystallographically equivalent). This symmetry operation will be called a negation and be represented by the symbol $-\hat{P}$, where P is the length of the code over which the operation is acting. ($-\hat{6}|101100> = |010011>$).

2. Two binary codes are equivalent if they can be brought into coincidence by a reversion of the code (*e.g.* $|101100> \hat{=} |001101>$). This symmetry operation will be represented by the symbol $\hat{P}I$ (*e.g.* $\hat{6}I|101100> = |001101>$).

3. Two binary codes are equivalent if they can be brought into coincidence by a cyclic shift of the code (*e.g.* $|101100> \hat{=} |011001>$). This symmetry operation will be represented by the symbol $\hat{n}$, where $\hat{n}$ is an integer number representing the number of places shifted to the left (*e.g.* $\hat{3}|101100> = |100101>$).

Sometimes is useful to know the number of 'forward' displacements or the number of 1's in a code; we will denote such a number by $\#_1$, conversely the number of 0's or 'backward' displacements will be denoted by $\#_0$, obviously $\#_1 + \#_0 = P$, where P is the number of layers in the repeating unit, or number of symbols in the code.

Binary codes can be further compacted by a run length encoding procedure known in crystallography as the Zhdanov symbol (Pandey & Krishna, 2004). We will not dwell on the properties of Zhdanov symbols in this paper.

3. The neutrality condition

The close-packed condition, together with the translational periodicity, constrains the valid binary codes to those that comply with

$$\#_s = \#_1 - \#_0 = 0 \bmod 3. \tag{1}$$

This condition will be called the neutrality condition. The neutrality condition has been explained by Verma & Krishna (1966). This condition implies that, from the set of 2^P possible binary codes of length P, only a subset will represent valid polytypes. The number of elements $\Omega(P)$ in such a set can be found if we rewrite the neutrality condition as

$$P - 2\#_1 = 0 \bmod 3. \tag{2}$$

According to the above equation, the neutrality condition is equivalent to finding these codes with

$$\#_1 = (P - 3s)/2, \tag{3}$$

where s is an integer number such that $0 < \#_1 < P$. This last condition limits the value of s to $-[P/3] < s < [P/3]$, where $[x]$ represents the integer part of x.

As $\#_1$ is an integer number, (3) also implies that $P - 3s$ must be an even number and, therefore, P and s must have the same parity.

(i) P and $s = 2s'$ are even.

In this case,

$$\#_1 = 3s' - 3[P/6] + P/2, \tag{4}$$

where $0 < s' < 2[P/6]$. For each value of s', the number of neutral codes will be given by the binomial coefficient

$$\binom{P}{3s' - 3[P/6] + P/2}.$$

The number of neutral codes of length P will then be given by

$$\Omega(P) = \sum_{s=0}^{2[P/6]} \binom{P}{3s - 3[P/6] + P/2}. \tag{5}$$

(ii) P and $s = 2s' + 1$ are odd.

In this case,

$$\#_1 = \frac{P+3}{2} - 3\left[\frac{[P/3]+1}{2}\right] + 3s', \tag{6}$$

where $0 < s' < 2[([P/3] + 1)/2]$ and as in the even case we will then have

$$\Omega(P) = \sum_{s=0}^{2[([P/3]+1)/2]} \binom{P}{3s + \frac{P+3}{2} - 3\left[\frac{[P/3]+1}{2}\right]}. \quad (7)$$

It is also useful to know the number of sequences of length P that do not comply with the neutrality condition; two cases can be found.

1. $\#_1 - \#_0 = 1 \bmod 3$.

We will call the number of polytypes of length P complying with the above condition $\Omega_1(P)$ and, following the same reasoning that led to (5) and (7), the following expression can be deduced:

$$\Omega_1(P) = \begin{cases} \sum_{p=0}^{G_1(P)} \binom{P}{\frac{P-6p+[(P+4)/6]-4}{2}} & 2|P \\ \sum_{p=0}^{G_2(P)} \binom{P}{\frac{P-6p+[(P+1)/6]-1}{2}} & 2\not|P, \end{cases} \quad (8)$$

where $G_1(P) = [(P-4)/6] + [(P+4)/6]$ and $G_2(P) = [(P-1)/6] + [(P+1)/6]$.

In (8), $x|y$ ($x\not|y$) reads x is (not) a divisor of y.

2. $\#_1 - \#_0 = 2 \bmod 3$.

We will call the number of polytypes of length P complying with this condition $\Omega_2(P)$. Taking into account that

$$\Omega(P) + \Omega_1(P) + \Omega_2(P) = 2^P, \quad (9)$$

we will have

$$\Omega_2(P) = 2^P - \Omega(P) - \Omega_1(P). \quad (10)$$

4. Seitz operators for close-packed binary codes

A Seitz operator (Arnold, 2002), commonly used in crystallography, acting over a vector $|v\rangle$ is an operator $\hat{\mathbf{S}} = (\mathbf{S}, |\mathbf{s}\rangle)$, where $\mathbf{S}$ is a $P \times P$ square matrix and $|\mathbf{s}\rangle$ is a vector of length P such that $\hat{\mathbf{S}}|\mathbf{v}\rangle = \mathbf{S}|\mathbf{v}\rangle + |\mathbf{s}\rangle$.

If we now consider the binary code of length P as a vector $|\mathbf{b}\rangle_P$, we can also define a Seitz operator acting over $|\mathbf{b}\rangle_P$ as

$$\hat{\mathbf{S}}|\mathbf{b}\rangle_P = \mathbf{S}|\mathbf{b}\rangle_P + |\mathbf{s}\rangle_P \bmod 2. \quad (11)$$

Again, $\mathbf{S}$ is a $P \times P$ matrix and $|\mathbf{s}\rangle_P$ is a binary vector of length P. The need for the mod 2 operation arises from the fact that we are dealing with a binary representation of the stacking sequence. The matrix component of the Seitz operator represents a permutation, while the vector part could represent a negation or a null vector.

In terms of Seitz operators, the negation of a code will be given simply by

$$-\hat{\mathbf{P}} = (\mathbf{E}, |2^P - 1\rangle). \quad (12)$$

E is the $P \times P$ identity matrix and $|2^P - 1\rangle$ will be a vector of length P composed only of 1's.

A cyclic rotation of p places will be given by

$$\hat{\mathbf{p}} = (\tilde{\mathbf{p}}, |0\rangle_P). \quad (13)$$

$\tilde{\mathbf{p}}$ is the cyclic shift matrix given by

$$\tilde{\mathbf{p}} = \begin{pmatrix} 0 & 0 & \ldots & 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & \ldots & 0 & 0 & 1 & \ldots & 0 \\ . & . & \ldots & . & . & . & \ldots & . \\ 0 & 0 & \ldots & 0 & 0 & 0 & \ldots & 1 \\ 1 & 0 & \ldots & 0 & 0 & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 & 0 & 0 & \ldots & 0 \\ . & . & \ldots & . & . & . & \ldots & . \\ 0 & 0 & \ldots & \underbrace{1 \;\; 0 \;\; 0}_{p} & & & \ldots & 0 \end{pmatrix}. \quad (14)$$

The reversion operation in terms of the Seitz operator will be

$$\hat{P}I = (\tilde{I}, |0\rangle_P), \quad (15)$$

where

$$\tilde{I} = \begin{pmatrix} 0 & 0 & 0 & . & . & . & 0 & 0 & 1 \\ 0 & 0 & 0 & . & . & . & 0 & 1 & 0 \\ 0 & 0 & 0 & . & . & . & 1 & 0 & 0 \\ . & . & . & . & . & . & . & . & . \\ 0 & 0 & 1 & . & . & . & 0 & 0 & 0 \\ 0 & 1 & 0 & . & . & . & 0 & 0 & 0 \\ 1 & 0 & 0 & . & . & . & 0 & 0 & 0 \end{pmatrix}. \quad (16)$$

The reversion operation, together with the negation and the $\hat{1}$ rotation, are the generators of a group in the mathematical sense, of which (12), (13) and (15) form the basis of a representation. We will call such a group $\mathcal{CP}_P$.

The action of $\mathcal{CP}_P$ over the set of all possible codes of length P will classify each code into a family of orbits of codes, each family representing the same polytype.

The reversion operation $\hat{P}I$ will form a cyclic group of order 2, the same will be true for the negation operation $-\hat{P}$. The rotation operations form a cyclic group of order P with generator $\hat{1}$. The set of P rotations forms an invariant subgroup of $\mathcal{CP}_P$. The order of one operation $\hat{p}$ will be $P/\text{g.c.d.}(P, p)$, where g.c.d.(x, y) stands for the greatest common divisor of x and y.

The complete group of close-packed symmetries $\mathcal{CP}_P$ will contain, additionally to the reversion, negation and rotation operators; the reversion–rotation operators $\hat{p}I$, negation–rotation operators $-\hat{p}$ and negation–reversion–rotation operators $-\hat{p}I$. For polytypes of length P, the symmetry group $\mathcal{CP}_P$ will have order $4P$.

The set of P reversion–rotation $\{\hat{p}I\}$ do not form a subgroup of $\mathcal{CP}_P$, yet the set of all rotation operations together with the set of reversion–rotation operations do form a group which will be denoted as $\{\hat{p}I\}^2$. The $\{\hat{p}I\}^2$ group is not Abelian and $\hat{p}I = I(\widehat{P+p})$. The order of an operation $\hat{p}I$ will be 2.

The negation element $-\hat{P}$ forms a cyclic subgroup $\mathcal{CP}_P$, this group will be invariant, and $\mathcal{CP}_P$ can be written as the direct product of $\{\hat{p}I\}^2$ and $-\hat{P}$. The set of operations $\{-\hat{p}\}$ do not form a subgroup of $\mathcal{CP}_P$, yet the set $\{-\hat{p}\}$ together with the set $\{\hat{p}\}$ do form a normal subgroup of $\mathcal{CP}_P$, which will be denoted by $\{\pm\hat{p}\}$. The order of a $-\hat{p}$ operation is $P/\text{g.c.d.}(P, p)$ if such division is even, and $2P/\text{g.c.d.}(P, p)$ otherwise.

5. The number of non-equivalent polytypes of length *P*: $\Gamma(P)$

In order to count the number of distinctive polytypes (one of each orbit) of length P, use will be made of the Burnside lemma (Cauchy–Frobenius lemma) (McLarnan, 1981).

Let G be a finite group of order $\#_G$ permuting a finite set K of elements, then

$$\Gamma(K) = \frac{1}{\#_G} \sum_{g \in G} \#_{K_g}, \qquad (17)$$

where $\Gamma(K)$ is, under the action of G, the number of non-equivalent members of K. K_g is the set of all elements of K left fixed by g and $\#_{K_g}$ is the order of K_g.

Following the Burnside lemma, the problem of counting the number of distinctive polytypes of length P involves finding a formula for $\#_{K_g}$, for all $g \in \mathcal{CP}_P$, where now K stands for all neutral codes of length P.

$$\Gamma(P) = \frac{1}{4P} \sum_{p=1}^{P} \{\#_{K_{\hat{p}}} + \#_{K_{-\hat{p}}} + \#_{K_{\hat{p}I}} + \#_{K_{-\hat{p}I}}\} - \sum_{d_i} \Gamma(P/d_i). \qquad (18)$$

The last term in (18) sums over all d_i divisors of P, and this term avoids counting polytypes with period a divisor of P as polytypes of length P.

We will now be calculating the number of neutral binary codes that remain unchanged under a $\mathcal{CP}_P$ operator. As is common in algebra, these vectors can be obtained by solving the eigenvector equation. In this case,

$$\hat{S}|\mathbf{b}> \, = |\mathbf{b}> \bmod 2. \qquad (19)$$

As already pointed out, of all the solutions of (19) only those complying with the neutrality condition will be valid.

Let us consider an eigenvector $|\mathbf{b}>$ formed by P characters. We will denote by $\{f\}$ the set of characters left independent by (19). $\#_f$ characters will be free to take either the value 0 or the value 1 and, once the choice is made for each $\{f\}$ character, the remaining $P - \#_f$ will be fixed by (19). Let $\#_m(f_i)$ be the number of characters fixed by the character f_i. In the case of $\mathcal{CP}_P$, all solutions to (19) relate one character b_i from $|b>$ with one character f_i, but the relation is not a one-to-one correspondence. The neutrality condition will then be written as

$$\sum_{i}^{\#_f} \#_m(f_i) f_i = 0 \bmod 3. \qquad (20)$$

In general, several f_i can have the same multiplicity $\#_m(f_i)$. The neutrality condition will be governed by the combination of terms in (20).

Let us consider some examples.

For the operation $\hat{3}$ acting over codes of length 15, the eigenvector equation reads

$$\hat{3}|b>_{15} = |b>_{15} \bmod 2. \qquad (21)$$

Solutions to (21) will be of the type

$$|b>_{15} = |b_1 b_2 b_3 b_1 b_2 b_3 b_1 b_2 b_3 b_1 b_2 b_3 b_1 b_2 b_3>.$$

The independent characters will then be $\{b_1, b_2, b_3\}$ and $\#_f = 3$. The multiplicity of all the independent characters is $\#_m = 5$ and the neutrality equation will be

$$5(b_1 + b_2 + b_3) = 0 \bmod 3. \qquad (22)$$

In order to satisfy (22), we will need

$$(b_1 + b_2 + b_3) = 0 \bmod 3$$

and therefore there will be $\Omega(3)$ possible combinations.

When all the independent characters $\{f_i\}$ have the same multiplicity $\#_m$, the neutrality equation will be

$$\#_m \sum_{i}^{\#_f} f_i = 0 \bmod 3 \qquad (23)$$

and two cases are possible:

(i) $3|\#_m$: the neutrality condition is assured by the multiplicity, and the $\#_f$ characters can take any value, there will be $2^{\#_f}$ possibilities;

(ii) $3\not|\#_m$: the neutrality condition has to be fulfilled by the sum of the $\#_f$ characters, and the number of possible combinations will then be $\Omega(\#_f)$.

Let us take a second example. Consider the operation $\hat{7}I$ acting over the codes of length 16. The eigenvector equation is then

$$\hat{7}I|b>_{16} = |b>_{16} \bmod 2. \qquad (24)$$

The solution will be of the type

$$|b>_{16} = |b_1 b_2 b_3 b_4 b_5 b_4 b_3 b_2 b_1 b_6 b_7 b_8 b_9 b_8 b_7 b_6>.$$

$\{f_i\}$ will be $\{b_1, b_2, b_3, b_4, b_5, b_6, b_7, b_8, b_9\}$ and $\#_f = 9$. The multiplicity of $\{b_1, b_2, b_3, b_4, b_6, b_7, b_8\}$ is 2 and the multiplicity of $\{b_5, b_9\}$ is 1, the neutrality condition will then be

$$2(b_1 + b_2 + b_3 + b_4 + b_6 + b_7 + b_8) + (b_5 + b_9) = 0 \bmod 3 \qquad (25)$$

and the following possibilities arise.

1. $(b_1 + b_2 + b_3 + b_4 + b_6 + b_7 + b_8) = 0 \bmod 3$ and $(b_5 + b_9) = 0 \bmod 3$, we will then have $\Omega(7)\Omega(2)$ combinations.
2. $(b_1 + b_2 + b_3 + b_4 + b_6 + b_7 + b_8) = 1 \bmod 3$ and $(b_5 + b_9) = 1 \bmod 3$, we will then have $\Omega_1(7)\Omega_1(2)$ combinations.
3. $(b_1 + b_2 + b_3 + b_4 + b_6 + b_7 + b_8) = 2 \bmod 3$ and $(b_5 + b_9) = 2 \bmod 3$, we will then have $\Omega_2(7)\Omega_2(2)$ combinations.

Adding the three cases, we finally get

$$\Omega(7)\Omega(2) + \Omega_1(7)\Omega_1(2) + \Omega_2(7)\Omega_2(2) \qquad (26)$$

possible neutral vectors $|b>_{16}$ complying with (24).

If the solution to (19) leads to a neutrality equation of the type

$$\#_{m_1} \sum_{j}^{\#_{f_1}} {}_1b_j + \#_{m_2} \sum_{j}^{\#_{f_2}} {}_2b_j = 0 \bmod 3, \qquad (27)$$

then the possible cases are

$\#_{m_1} \bmod 3$	$\#_{m_2} \bmod 3$	No. of neutral codes
0	0	$2^{\#_{f_1}+\#_{f_2}}$
0	$\neq 0$	$2^{\#_{f_1}}\,\Omega(\#_{f_2})$
$\neq 0$	0	$2^{\#_{f_2}}\,\Omega(\#_{f_1})$
1	1	$\Omega_2(\#_{f_1})\Omega_1(\#_{f_2})$
2	2	$+\Omega_1(\#_{f_1})\Omega_2(\#_{f_2}) + \Omega(\#_{f_1})\Omega(\#_{f_2})$
2	1	$\Omega_2(\#_{f_1})\Omega_2(\#_{f_2})$
1	2	$+\Omega_1(\#_{f_1})\Omega_1(\#_{f_2}) + \Omega(\#_{f_1})\Omega(\#_{f_2})$

When the operator includes the negation operation, the eigenvector equation does not necessarily have a solution. If (19) has a solution, the independent characters come in pairs of character-negated character, this guarantees the neutrality condition, and there will be $2^{\#_f}$ possible eigenvectors.

The examples above exhaust all possibilities for $\mathcal{CP}_P$.

Table 1 shows the values of $\Gamma(P)$ for values of P up to 25, calculated by the above procedure using the eigenvector equation (19); they are identical with those reported by McLarnan (1981).

Table 1
Number of non-equivalent close packings $\Gamma(P)$ for different periodic length P.

Numbers are identical to those reported by McLarnan (1981).

$\Gamma(P)$	P	$\Gamma(P)$	P
1	0	21	8492
2	1	22	16409
3	1	23	30735
4	1	24	59290
5	1	25	112530
6	2	26	217182
7	3	27	415620
8	6	28	803076
9	7	29	1545463
10	16	30	2990968
11	21	31	5778267
12	43	32	11201472
13	63	33	21702686
14	129	34	42140890
15	203	35	81830744
16	404	36	159139498
17	685	37	309590883
18	1343	38	602935713
19	2385	39	1174779333
20	4625	40	2290915478

6. The symmetry of close-packed binary codes

Up to now, we have considered symmetry relations that resulted from the equivalence conditions in the stacking sequences. These conditions can be said to be based on the 'interpretation' of the codes as a representation of stacking sequences and not based on the structure of the code itself. According to this equivalence, two binary codes different in appearance codify the same stacking arrangement. In this section, we will deal with other types of symmetries, those that result from the structure of the code itself.

Fig. 1 shows the symmetry elements in a single close-packed plane, the hexagon, black triangle and inverted triangle symbols represent a sixfold, threefold and threefold axis, respectively. Each threefold axis represented in Fig. 1 intersects only one of the two different types of interstitial sites present in the layer. For the stacking arrangement, the existence of a sixfold axis and a threefold axis perpendicular to the plane will be important as they can derive into 6_k- and 3_k-fold screw axes when the stacking arrangement is taken into account. Additional to these two symmetry operations, the appearance, as a result of the stacking arrangement, of mirror planes perpendicular to the stacking direction, and of inversion centres, will have to be considered.

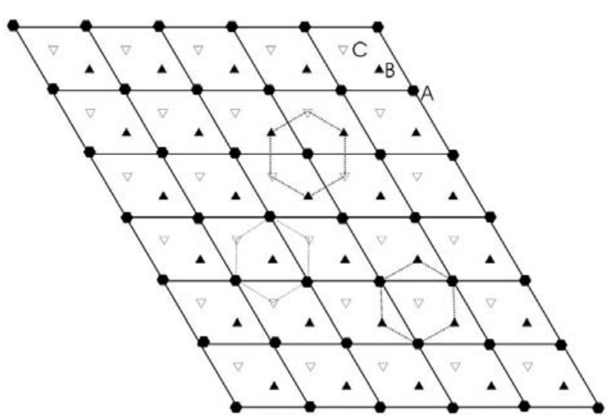

Figure 1
Close-packed layers and the sixfold and threefold axes. The hexagon, black triangle and inverted triangle symbols represent a sixfold, threefold and threefold axis, respectively. If we take the layer to be in an A position, then hexagon corresponds to the A location (0, 0), black triangle to the B (1/3, 2/3) location and inverted triangle (2/3, 1/3) to the C location.

As in the previous section, if we wish to count the number of polytypes complying with a given symmetry group, we can make use of the Burnside lemma. The set of codes $\{K\}$ over which the $\mathcal{CP}_P$ group acts will now all be neutral codes that also comply with the restriction imposed by the additional symmetries. Therefore we will be solving eigenvector equations of the type (19) together with additional equations of the type

$$\hat{G}|\mathbf{b}\rangle = |\mathbf{b}\rangle \bmod 2, \tag{28}$$

where $\hat{G}$ will be the additional symmetry operator.

Patterson & Kasper (1959) have described the effect of the different symmetry elements in the appearance of the Hägg symbol which we will discuss in what follows.

6.1. Codes containing a 3_k symmetry operation

The threefold axis in the close-packed plane brings an A (B, C) position to coincidence with itself. The operation will change a 1 (0) character to the same character 1 (0), it can be considered as the identity operation acting over the code characters. As a consequence of the threefold symmetry, a 3_k-fold screw axis will be equivalent to a code $|b\rangle_P$ formed by the repetition k times of a code $|r\rangle$ of length (P/k):

$$\boxed{\triangleright}\boxed{\triangleright}\cdots\boxed{\triangleright}$$

where $\triangleright$ represents the $|r\rangle$ code. $|r\rangle$ cannot be neutral as it will then be the representative code of the stacking sequence and not $|b\rangle$. We will then have for the neutrality condition

$$k\#_s(|r\rangle) = 0 \bmod 3,$$

where $\#_s$ can take values of 1 or 2 mod 3. k will have to be the smallest non-zero integer such that

$$k = 0 \bmod 3. \quad (29)$$

The only solution to (29) is $k = 3$. If a code $|b\rangle$ is formed by repeating identical blocks, then the only possibility is

$$\boxed{\triangleright}\boxed{\triangleright}\boxed{\triangleright}$$

The corresponding symmetry operator will be

$$\hat{R} = ((\widetilde{P/3}), |0\rangle), \quad (30)$$

which also corresponds to a 3_1-fold axis. It is clear that this symmetry will only appear for codes of length a multiple of 3.

From the discussion above, the $\hat{R}$ symmetry corresponds to a stacking sequence formed by one block of layers repeating itself along **c**. Each block will be shifted with respect to the previous block by a vector given by the $\#_s$ value of the block, which can lead to a $\frac{1}{3}\mathbf{a} + \frac{2}{3}\mathbf{b}$ or a $\frac{2}{3}\mathbf{a} + \frac{1}{3}\mathbf{b}$ shift (see Fig. 1). This sequence corresponds to a rhombohedral cell.

To count the number of non-equivalent sequences with such symmetry, equation (28) reduces to

$$\hat{R}|b\rangle = |b\rangle \bmod 2. \quad (31)$$

6.2. Codes containing a 6_k symmetry operation

The sixfold axis perpendicular to the stacking direction and passing through a layer position for example A (B, C), will bring the B (C, A) layer to coincidence with a C layer (A, B). The operation will then change a 1 (0) character to a 0 (1) character. The 6_k-fold screw axis acting over a code will result in a code of the form:

$$\boxed{\triangleright}\boxed{\blacktriangleright}\cdots\boxed{\triangleright}\boxed{\blacktriangleright}$$

where $\blacktriangleright$ represents the negated $\triangleright$ code. It is immediately clear from the expression above that the only choice compatible with the neutrality condition and the minimum-length-code criteria will be 6_3, and the code will be of the form

$$\boxed{\triangleright}\boxed{\blacktriangleright}$$

The corresponding Seitz operator will be given by

$$\hat{6}_3 = ((\widetilde{P/2}), |1\rangle) \quad (32)$$

and the eigenvector equation (28),

$$-(\widehat{P/2})|b\rangle = |b\rangle. \quad (33)$$

The number of codes with this symmetry will be different from zero only for even values of P.

6.3. Codes containing a $\bar{1}$ symmetry operation

A $\bar{1}$ symmetry operation at the centre of the code will bring a 1 (0) character in a $+z$ position of the code to a 1 (0) character in a $-z$ position in the code. In this case, the code will have the structure

$$\boxed{\triangleright}\boxed{\triangleleft}$$

and $\triangleleft$ is the reversed $\triangleright$ code. The Seitz operator will be

$$\bar{1} = (\widetilde{PI}, |0\rangle) \quad (34)$$

and the eigenvector equation to be solved is

$$\widehat{PI}|b\rangle = |b\rangle. \quad (35)$$

6.4. Codes containing a $\bar{2}$ (mirror) symmetry operation

The codes containing a mirror operation at the centre of the code will be of the form

$$\boxed{\triangleright}\boxed{\blacktriangleleft}$$

the symmetry operator will be represented by

$$\bar{2} = (\widetilde{PI}, |1\rangle) \quad (36)$$

and the eigenvector equation to be solved for counting purposes is

$$-\widehat{PI}|b\rangle = |b\rangle. \quad (37)$$

Although the symmetry operation given by (36) exists for any value of P, the structure of the code formed by two blocks, one the negated of the other, will force the codes with such symmetry to have length an even number P.

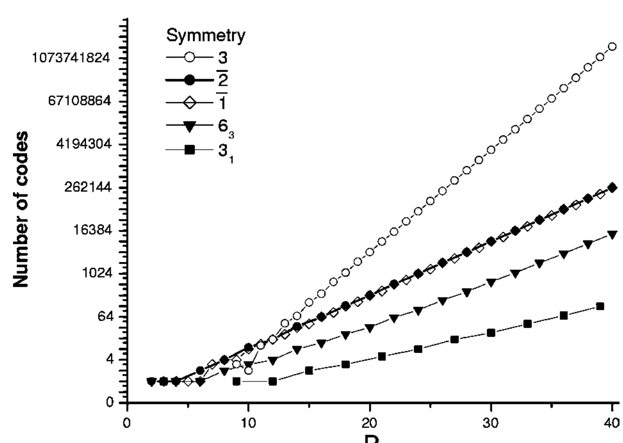

Figure 2
Number of non-equivalent sequences of length P with a single symmetry operation other than the identity (only values different from zero are shown). Scale is $\log_2$.

The appearance of only one of the symmetry operations described above, together with the symmetries describing the hexagonal-close-packed planes, will lead to the space groups $R3m$ for 3_1, $P6_3mc$ for 6_3, $P\bar{3}m1$ for $\bar{1}$ and $P\bar{6}m2$ for $\bar{2}$. When only the identity symmetry is present (threefold axis), we will be dealing with a code belonging to space group $P3m1$.

Fig. 2 shows the number of polytypes with only one symmetry operation for increasing length P, calculated from the corresponding eigenvector problem. From the figure, it can be observed that the 3_1 operation has the smallest number of codes for large values of P. It can then be said that the 3_1 symmetry is the symmetry operation that imposes the largest restriction over the code structure. In decreasing order of restriction, after the 3_1 operation follows the 6_3 symmetry. The $\bar{1}$ and $\bar{2}$ operations impose the same degree of restriction, while the least restrictive is of course the 3 (identity) symmetry. The fact that most codes do not possess any additional symmetry at all points to the known result in Kolmogorov complexity analysis that, for a given code length, the majority of the codes will be completely random strings (Li & Vitányi, 1993).

6.5. Codes containing a combination of symmetry operations

The combination of the symmetry elements discussed above will lead to the remaining space groups:

• Symmetry operations 3_1 and $\bar{1}$ ($R\bar{3}m$):

The codes must comply at the same time with the structures

▷	▷	▷

and

▷	◁

which immediately leads to codes of the type

◁▷	◁▷	◁▷

The $|111>$ code is a special case of this symmetry. If the metric tensor of the arrangement meets the cubic constraints, this sequence leads to the $Fm\bar{3}m$ space group, which still contains the 3_1 and $\bar{1}$ symmetry. This code, on the other hand, is only composed of three characters and each character is a block by itself, symmetrically equivalent to the other two. This last property contrasts with any other vector in the same group, which will have blocks composed of several characters; characters within a block do not have to be symmetrically equivalent with each other. The special condition that each character is a block by itself, symmetry equivalent to the other characters in the vector, is pointing to the fact that in this arrangement every layer is symmetry equivalent to any other layer in the stack.

• Symmetry operations 6_3 and $\bar{2}$ ($P6_3/mmc$):

The codes must comply at the same time with the structures

▷	▶

and

▷	◀

which leads to

◁▷	◀▶

This code can be rotated 1/4 of the periodicity and written as

▷◀	▶◁

If we compare this last code structure to the structure of the codes with $\bar{1}$ symmetry, the appearance of the inversion symmetry is immediately apparent, as a result of the combination of 6_3 and $\bar{2}$ operations. The $\bar{1}$ operation appears at 1/4 of the mirror operation. The $P6_3/mmc$ space group is centrosymmetric. We will represent the $\bar{1}$ operation by a dashed vertical line.

◁┆▷	◀┆▶

The $|10>$ code belonging to the hexagonal-close-packed sequence belongs to this space group. Again, as in the face-centred cubic code, the hexagonal-close-packed code is composed of two characters, each being a separate block symmetrically equivalent to the other. In this arrangement too, every layer is symmetrically equivalent to any other layer in the stack.

• Symmetry operations 3_1 and $\bar{2}$:

This will lead to a code of the type

▷◀	▷◀	▷◀

which will have a representative code

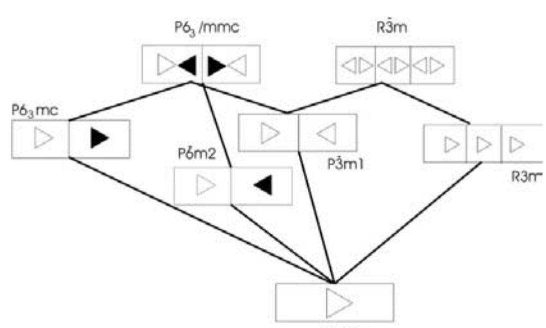

Figure 3
Group–subgroup relations among the close-packed codes. Subgroups appear below the supergroup and are connected by a solid line

▷	◀

the same as the $P\bar{6}m2$ space group.

- Symmetry operations 3_1 and 6_3:

This will lead to a code of the type

▷▶	▷▶	▷▶

which will have as representative code

▷	▶

and we are dealing with the $P6_3mc$ space group.

This will exhaust all possible combinations.

The group–subgroup relations can be immediately deduced by inspection of the code structures and this is shown in Fig. 3

The procedure for counting the number of polytypes belonging to each space group is straight forward: for each of the symmetry elements present in the symmetry group, several of the eigenvector equation (31), (33), (35) or (37) will hold. The simultaneous compliance with the symmetry eigenvector equations together with (19) will lead to the number of codes belonging to the corresponding space group, after subtracting redundant sequences.

The best procedure for counting is to start with the higher supergroups and proceed along the subgroup relation. In this way, we can take into account that the eigenvector solution contains the number of codes belonging to the supergroup, already counted and therefore redundant.

Fig. 4 shows the number of polytypes belonging to each space group for increasing P values.

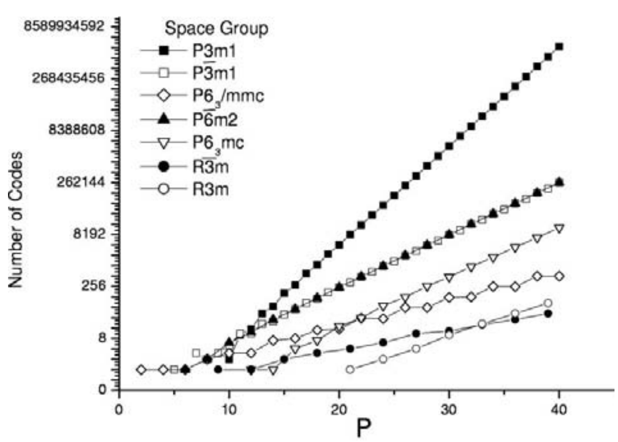

Figure 4
Number of non-equivalent sequences of length P and given space group (only values different from zero are shown). Scale is $\log_2$.

7. Conclusions

The systematic use of binary coding together with Seitz operators acting over the binary field can be used to study close-packed sequences. The algebraic approach, as well as that of McLarnan (1981), is of more general use in enumeration problems than that of Iglesias (1981). The method presented here avoids the necessity of involved group-theoretical considerations to solve counting problems in polytypes. The coding approach also allows the symmetry group of stacking sequences to be derived as a natural result of code symmetry. The examples presented in the paper should show the power of this approach. Further developments can include the extension to non-periodic stacking sequences.

One of the authors (EER) acknowledges support from the Humboldt Foundation through an AvH fellowship. Part of this project was carried out under an Alma Mater grant from the University of Havana.

References

Arnold, H. (2002). *International Tables for Crystallography*, Vol. A, 5th ed., edited by Th. Hahn, pp. 77–89. Dordrecht: Kluwer Academic Publishers.
Brown, B. E. & Bailey, S. W. (1962). *Am. Mineral.* **47**, 819–850.
Hales, T. C. (2002). Proceedings of the ICM, Beijing, 2002, Vol. 3, pp. 795–804.
Iglesias, E. (1981). *Z. Kristallogr.* **155**, 121–127.
Li, M. & Vitányi, P (1993). *An Introduction to Kolmogorov Complexity and its Applications*, ch. 1. Berlin: Springer Verlag.
Lima de Faria, J. & Figueiredo, M. O. (1969). *Z. Kristallogr.* **26**, 54–67.
McLarnan, T. J. (1981). *Z. Kristallogr.* **155**, 269–291.
Pandey, D. & Krishna, P. (2004). *International Tables for Crystallography*, Vol. C, pp. 752–759. Dordrecht: Kluwer Academic Publishers.
Patterson, A. L. & Kasper, J. S. (1959). *International Tables for Crystallography*, Vol. II, pp. 342–354. Birmingham: Kynoch Press.
Shrestha, S. P. & Pandey, D. (1996). *Europhys. Lett.* **34**, 269–274.
Shrestha, S. P. & Pandey, D. (1997). *Proc. R. Soc. London Ser. A*, **453**, 1311–1330.
Smith, J. V. & Rinaldi, F. (1962). *Mineral. Mag.* **33**, 202–212.
Thompson, R. M. & Downs, R. T. (2001). *Acta Cryst.* B**57**, 766–771.
Verma, A. R. & Krishna, P. (1966). *Polymorphism and Polytypism in Crystals.* New York: Wiley.