Acta Crystallographica Section A
Foundations of Crystallography
ISSN 0108-7673

Received 27 March 2008
Accepted 14 May 2008

The symmetry of HK codes representing close-packed structures and the efficient generation of non-equivalent polytypes of a given length

Ernesto Estevez-Rams[a]* and Jennifer Martinez-Mojicar[b]

[a]Institute of Materials Research and Engineering (IMRE), University of Havana, San Lazaro y L. CP 10400 C., Habana, Cuba, and [b]Universidad de Oriente, Santiago de Cuba, Cuba. Correspondence e-mail: estevez@imre.oc.uh.cu

The HK representation of close-packed polytypes is studied as a binary code. It is shown that the HK code can be seen as operators forming a group. The neutrality condition is then translated to HK sequences that result in the identity operator. The symmetry of an HK word can be related to the space-group symmetry of the corresponding polytype. All HK code types corresponding to all possible close-packed space groups are reported. From a coding perspective, equivalent HK codes correspond to bracelet equivalent classes. An efficient algorithm with execution time constant per generated object is modified to generate all non-equivalent polytypes of a given length.

1. Introduction

It is known that close-packed crystals can be described by the local environment of each layer in the stack (Verma & Krishna, 1966). This description results, for nearest neighbors, in the so-called HK code, introduced by Jagodzinski (1949), where a letter h is assigned to each layer having the same type of layer at each side (*e.g.* AXA, BXB, CXC), and a letter k is assigned otherwise. The HK code emphasizes the fact that the letter assigned to each layer position is irrelevant as long as consistency is kept, while the first-neighborhood environment has a real physical meaning. This description has been less studied than the Hägg coding, which is the basis of the Zhdanov symbols of polytypes (Zhdanov, 1945); a more recent and in-depth discussion of Zhdanov symbols can be found in Iglesias (2006*a*). Both descriptions, the HK and the Hägg codings, are bijective between them and with respect to the single-layer description $\ldots ABC \ldots$.

For a given length N, there will be in general 2^N different binary sequences, yet, if these sequences are to represent close-packed polytypes, then symmetry will significantly reduce the number of non-equivalent codes (Iglesias, 1981; McLarnan, 1981; Estevez-Rams, Azanza-Ricardo, Martínez García & Aragón-Fernández, 2005). To effectively count the number $\Gamma(N)$ of non-equivalent polytypes of length N, the naive approach of generating all possible sequences and then classifying equivalent ones becomes intractable due to its exponential explosion with N. Smarter counting procedures that avoid the necessity of generating all the sequences have been reported (McLarnan, 1981; Iglesias, 1981; Estevez-Rams, Azanza-Ricardo, Martínez García & Aragón-Fernández, 2005; Estevez-Rams, Azanza-Ricardo & Aragon-Fernandez, 2005). Recently, Iglesias (2006*b*) obtained closed expressions, much simpler than previous approaches, that allow the number of polytypes of a given space group to be counted in an efficient way.

The problem of generating all non-equivalent polytypes of a given length shares common problems with the counting problem. As there are 2^N possible sequences, the naive approach will also explode its execution time per object with N.

In this paper, the complete symmetry description of the HK code corresponding to each close-packed polytype space group is reported. Also, an efficient algorithm for generating all non-equivalent polytypes of a given length will be presented. Efficient will mean constant amortized time (CAT), which is proportional to $\Gamma(N)$, the number of objects generated. The algorithms are based on a previously reported CAT algorithm for bracelet generation (Sawada, 2001) modified to take into account the neutrality of the generated codes. The paper is organized as follows: we first introduce the HK code as non-commuting operators forming a group and, from there, discuss the neutrality condition that any code describing a valid polytype must obey. The symmetry of the HK codes will then be studied and its relation to the space group of the corresponding polytype. The generation problem of non-equivalent polytypes will be built considering each code as a bracelet word in coding theory. The algorithm will then be described and reported.

2. The HK coding and the neutrality condition

The HK code takes symbols from the alphabet $\Sigma = \{h, k\}$, and every finite string of concatenated symbols drawn from Σ will be called a word.

For a periodic arrangement of periodicity N, the corresponding word of length N will be sufficient to describe the whole arrangement, for example the periodic stacking $\ldots ABCBAB|ABCBAB|ABCBAB|ABCBAB\ldots$, of period 6, can be described by the word *hkhkhh* of the same length. The first and last symbol in the HK word is a consequence of the stacking periodicity. The close-packed condition means that nowhere in the stacking sequence can two equal consecutive layers (AA, BB or CC) happen, this in turn implies for a periodic sequence that the repeating unit cannot start and end with the same letter.

For every polytype of a given length, more than one HK code or word can be constructed. From the definition above, and the periodicity of the polytype, it is clear that any cyclic rotation of the word will represent the same polytype (*e.g.* *hkhkhh*, *khkhhh*, *hkhhhk*, *khhhkh*, *hhhkhk*, *hhkhkh* are all equivalent), the reversion of the word (*hkhkhh* and *hhkhkh*) will also represent the same the polytype, as the reversed word translates into observing the same stacking arrangement from the other 'end'. Therefore, all words that can be brought to coincidence by a cyclic shift, a reversion or a combination of both operations will represent the same polytype: they belong to the same equivalence class or orbit. Any member of the equivalence class will be equally good for representing the whole orbit.

A further consideration is important: periodicity implies that the repeating unit must start with the same layer and therefore any valid word must comply with this restriction. For example, the code *hkkh* will describe a sequence of the type $\ldots A|BACB|C\ldots$, where the supposed periodic unit is between the vertical bars. It is clear that it is not periodic and furthermore $BACB$, taken as a periodic unit, violates the close-packed condition where two consecutive layers of the same type are forbidden; words that lead to this type of inconsistency are called non-neutral codes, we will also introduce them as charged codes. On the contrary, the word *hkhk* represents the sequence $\ldots A|BACA|B\ldots$, which is clearly periodic and does not violate the close-packed restriction. The HK codes that result in a valid polytype sequence will be called neutral. A similar concept has been discussed for the Hägg code (Estevez-Rams, Azanza-Ricardo & Aragon-Fernandez, 2005; Estevez-Rams, Azanza-Ricardo, Martínez García & Aragón-Fernández, 2005). When a code is not neutral, repeating the code a number of times will eventually result in a neutral code (more on this below), thus, a non-neutral code represents the repeating unit of the complete valid HK code representing a polytype. This fact was used in the original Jagodzinski notation to further shorten the notation. For example, a valid neutral code like *hhhkhkk hhhkhkk hhhkhkk* was shortened to $[hhhkhkk]_3$, indicating that the non-neutral code *hhhkhkk* has to be expanded three times to get a valid non-neutral code representing a polytype. In what follows, we will use the full code for representing the polytypes. This makes it easier to understand some discussions; the notation is clear enough to go from the original Jagodzinski notation to the expanded one and so the original code will be used when no confusion arises.

In order to derive an expression for neutrality that allows one to discriminate valid from invalid words, each symbol in the HK code can be viewed as an operator acting on a pair of ABC symbols and giving as output again a pair of ABC symbols, where the last symbol is the next symbol to the input sequence. Let us take an XY pair, where X and Y stand for A, B or C, the h operator will then be defined by the following relation:

$$h(XY) = YX, \tag{1}$$

for example, $h(AB) = BA$. The k operator will be defined by the relation

$$k(XY) = YZ, \tag{2}$$

for example, $k(AB) = BC$. The multiplication of the defined operators will result in the following four new operators:

$$\begin{aligned} k^2(XY) &= k(k(XY)) = k(YZ) = ZX \\ hk(XY) &= h(k(XY)) = h(YZ) = ZY \\ kh(XY) &= k(h(XY)) = k(YX) = XZ \end{aligned} \tag{3}$$

and the identity

$$e(XY) = XY. \tag{4}$$

The reader can verify the following multiplication table

	h	k	k^2	kh	hk	e
h	e	hk	kh	k^2	k	h
k^2	hk	e	k	h	kh	k^2
k	kh	k^2	e	hk	h	k
kh	k	h	hk	e	k^2	kh
hk	k^2	kh	h	k	e	hk
e	h	k	k^2	kh	hk	e

It follows that $\{h, k\}$ are the generators of a group with members $\{h, k, k^2, hk, kh = hk^2, e\}$ isomorphic with the S_3 permutation group and the $3m$ point group. From these isomorphisms, the sets $\{k, k^2\}$ and $\{h, hk, kh\}$ each form conjugate classes. In the first conjugate class, the operators are of order 3, while in the latter, they are of order 2. The alternating cyclic permutation subgroup $\{e, k, k^2\}$ is invariant. From the isomorphism, the following matrix representation can be ascribed to each operator:

$$\begin{aligned} &h = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad k = \begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix}, \quad k^2 = \begin{pmatrix} -1 & 1 \\ -1 & 0 \end{pmatrix}, \\ &kh = \begin{pmatrix} -1 & 0 \\ -1 & 1 \end{pmatrix}, \quad hk = \begin{pmatrix} 1 & -1 \\ 0 & -1 \end{pmatrix}, \quad e = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \end{aligned} \tag{5}$$

The neutrality constraint for a given HK code is now reduced to the condition that the multiplication of the operators in the code must result in the identity operator. For example,

$$hkkhhhkhhk = hkkh(hh)khhk = hk(khkh)hk = (hkhk) = e.$$

If a code does not result in the identity operator, the code is not neutral. The charge of a code will be labeled by the resulting operator [*e.g.* the code *khkkhhkh* will have charge k resulting from $khkk(hh)kh = kh(kkk)h = k(hh) = k$, that is, deleting any subword which fulfils the neutrality will not

Table 1
The equivalence between the expanded code and the original HK notation of Jagodzinski (1949).

An HK word w with the given charge results in an expanded and short valid code as given by the table.

Charge	Expanded code	Short code
e	w	w
h	ww	$[w]_2$
k	www	$[w]_3$
k^2	www	$[w]_3$
hk	ww	$[w]_2$
kh	ww	$[w]_2$

change the charge of the word]. From the ongoing discussion, three times a code of charge k or k^2 will result in a neutral code, while the same will happen for two times a code of charge h, kh or hk. (As kindly pointed out by one of the referees, this last rules allows us to recover the original Jagodzinski notation and non-neutral codes can be considered valid with periodicity given by the order of the charge, see Table 1.) Reversion of the code will keep the charge constant for k, k^2 and h while, for the charges hk and kh, a transposition of the two charges will occur.

3. Relation between the symmetry of the polytype and the symmetry of the HK coding

The discussion in what follows will be building on the description of symmetry in Hägg codes described by Estevez-Rams, Azanza-Ricardo, Martínez García & Aragón-Fernández (2005) and specially on the thorough discussion of Zhdanov symbols and cyclotomic sets by Iglesias (2006*a*). The reader is referred to these two articles for a discussion of the symmetry of the stacking sequence, keeping in mind that the relationship between Hägg codes and HK codes follow two rules:

(i) every time a $+-$ or a $-+$ sequence appears in a Hägg code, the corresponding HK word will have a letter h in the last position;

(ii) every time two equal consecutive symbols $++$ or $--$ appear in a Hägg code, the corresponding HK word will have a letter k in the last position.

According to the above rules, the following Hägg and HK codes of a periodic sequence are equivalent:

$$\begin{array}{ccccccccc} + & - & + & + & + & - & + & + & - \\ h & h & h & k & k & h & h & k & h. \end{array}$$

Five symmetry operations (Patterson & Kasper, 1959; Iglesias, 2006*a*), which are relevant for the stacking arrangement, can act over the close-packed layers: the 3_1 and the 6_3 screw axes perpendicular to the layer plane; the mirror plane m parallel to the layer plane; the inversion center which can be located in two different sites, over the spheres of a close-packed layer or in the octahedral sites between two close-packed layers; and, finally, the identity or a 3 axis perpendicular to the layers.

3.1. Polytypes with only one symmetry operator acting over the stacking arrangement

3.1.1. The 3 axis (*P*3*m*1). The 3 axis over a sphere position is equivalent to the identity operation, therefore there is no restriction to the HK code, and the only condition that it has to comply with is neutrality (Fig. 1*a*). The first code belonging to this symmetry is $[h^4kh^2k^2] = |hhhhkhhkk|$.

3.1.2. The 6_3 axis ($P6_3mc$). The 6_3 screw axis passing through the sphere of a close-packed layer will result in a layer of the same type displaced by one half of the stacking periodicity, while layers in the two other alternative sites will result in transposed layers with the same displacement. Consider for example the 6_3 axis passing through the A sites of a layer, then this will result in an A layer displaced by one half from the first A-layer position, the B (C) layers will result in C (B) layers also displaced by one half. Therefore, the 6_3 axis divides the stacking sequence into two blocks displaced by one half of the stacking period from each other. According to this rule, as one symbol is kept and the two others are transposed, the nearest neighbor of each layer in the first block is kept in the displaced one. For example, if the first block has the sequence $ABCB$, the second block will be $ACBC$ and finally the whole periodic sequence will be $ABCB|ACBC$, which corresponds to the HK code $kkhk|kkhk$. The HK code corresponding to this sequence will then consist of two identical blocks (Fig. 1*b*). The HK code will have a periodicity of one half of the periodicity of the stacking sequence. The periodicity of the stacking sequence will therefore be a even number. As the whole code must be neutral, each half must have a charge of order two which corresponds to h, kh or hk.

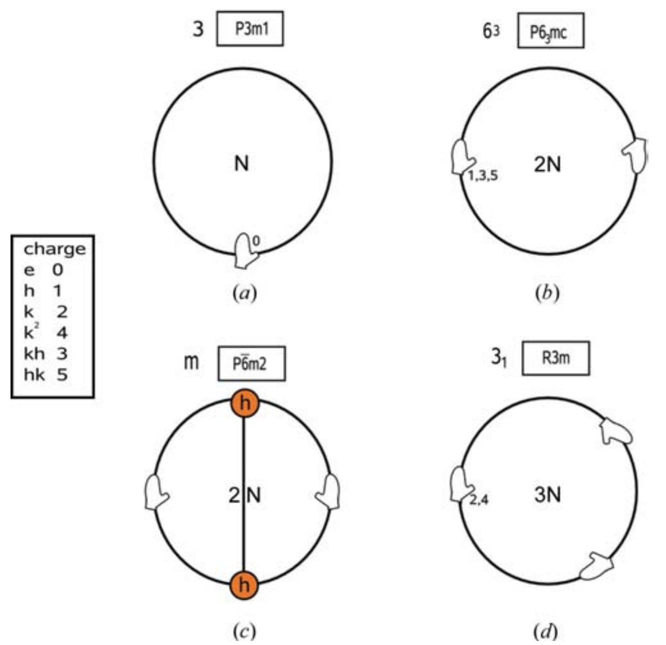

Figure 1
HK-code symmetry for a stacking arrangement with only one symmetry operation: (*a*) the identity (3), (*b*) a sixfold screw axis (6_3), (*c*) a mirror (m) and (*d*) a threefold screw axis (3_1). The X block is represented by a left hand and the reverse code $\tilde{X}$ by a right hand, the charge of the block is indicated by a number defined in the panel on the left.

The first code belonging to this symmetry is $[h^2khk^2]_2 = |hhkhkk|hhkhkk|$.

3.1.3. The *m* mirror plane (*P*$\bar{6}$*m*2). The close-packed restriction forces any m symmetry parallel to the layers to be contained within a close-packed layer; from this, the periodicity of the stacking sequence must be an even number. The symmetry will result in the same type of layer above and below the mirror plane, again the sequence will consist of two blocks divided by the mirror-plane layer, for example, $\underline{C}ABCB\underline{A}BCBA|\underline{C}$, where mirror planes are indicated by the underlined letters and the HK word is left of the vertical bar. From the symmetry rule, it is clear that the environment of the mirror plane can only be h and the blocks at each side of the mirror plane are reversed with respect to each other. The symmetry of the HK code is graphically shown in Fig. 1(*c*). As the whole code must be neutral, from the same Fig. 1(*c*), the following equation results:

$$Xh\tilde{X}h = e, \quad (6)$$

where X is the charge of one block and $\tilde{X}$ is the charge of the reversed block. The equation is satisfied for all possible X charges.

The first code belonging to this symmetry is $[(hk)^2h^2] = |hk\,h\,kh\,h|$.

3.1.4. The 3_1 axis (*R*3*m*). The 3 axis over a sphere position is equivalent to the identity operation, therefore the 3_1 operator results in a block repeated three times (Fig. 1*d*). The periodicity of the HK word is one third of the periodicity of the stacking sequence. The periodicity of the stacking sequence must be an integer multiple of three. The neutrality condition in this case leads to a charge of the repeating block equal to k or k^2.

The first code belonging to this symmetry is $[h^3khk^2]_3 = |hhhkhkk\,hhhkhkk\,hhhkhkk|$.

3.1.5. The $\bar{1}$ inversion center (*P*$\bar{3}$*m*1). Three possibilities now arise: the inversion center is at the spheres in a layer, this position is denoted by the letter S; the inversion center is at the octahedral site between two layers, this position is denoted by the letter O; and finally, inversion centers are at both sites already described, this position is denoted by SO.

$\bar{1}(S)$. Consider that the inversion center is at the sphere center in an A layer. Then, a layer B (C) on one side of the inversion layer will result in a C (B) layer at the same distance from the inversion center on the other side of the inversion layer. A layers on one side will still be A layers on the other side. The resulting stacking sequence will be made up of two blocks separated by a k corresponding to the environment of the inversion layer, the blocks will be a reverse pair (Fig. 2*a*). If the only inversion point is at the sphere center in a layer, the periodicity will be an even number (McLarnan, 1981).

The neutrality condition will be given by the equation

$$Xk\tilde{X}k = e. \quad (7)$$

This equation is satisfied by a block X of charge h and k^2. The first code belonging to this symmetry is $[h^2k^5h^2k] = |hhkk\,k\,kkhh\,k|$.

$\bar{1}(O)$. Consider that the inversion center is at the octahedral sites between two layers A and B. Then the inversion center will be over the C sites. The layer transformation rules will be $A \leftrightarrow B$, $C \leftrightarrow C$. The HK code will be formed by a pair of reversed related blocks (Fig. 2*b*). The neutrality condition will result in both blocks having charge h. Again, if the only inversion point is at the octahedral site in a layer, the periodicity will be an even number (McLarnan, 1981).

The first code belonging to this symmetry is $[(hk)^2(kh)^2] = |hkhk\,khkh|$.

$\bar{1}(SO)$. This case is the combination of the two previous cases but the number of layers in the periodic unit will be an odd number. It should be noticed that the $\bar{1}(S)$ centers are in one layer and the $\bar{1}(O)$ centers are in the octahedral site of any pair of layers not including the layer with the inversion centers (an exception to this is the kkk code belonging to $Fm\bar{3}m$). The resulting HK code is made by a k symbol at one side separating a pair of reversion-related blocks (Fig. 2*c*).

The neutrality condition will be given by the equation

$$Xk\tilde{X} = e. \quad (8)$$

This equation is satisfied by a block X of charge k and hk.

The first code belonging to this symmetry is $[h^2k^3] = |hk\,k\,kh|$.

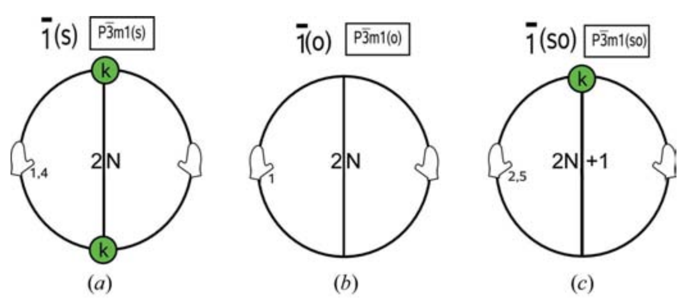

Figure 2
HK-code symmetry for a stacking arrangement with inversion center at (*a*) the sphere (*s*), (*b*) octahedral sites (*o*) and (*c*) both (*so*). The numbers representing the charge of the X blocks follow the same convention as in Fig. 1.

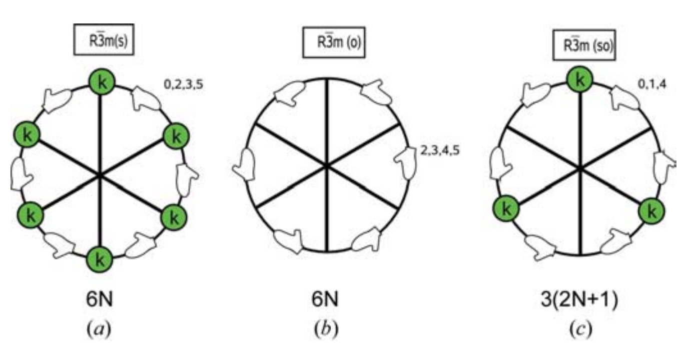

Figure 3
HK-code symmetry for a stacking arrangement with the combination of a 3_1 screw axis and an inversion center at (*a*) the sphere (*s*), (*b*) octahedral sites (*o*) and (*c*) both (*so*). The numbers representing the charge of the X blocks follow the same convention as in Fig. 1.

3.2. Combination of symmetry operations

3.2.1. The 3_1 and $\bar{1}$ operations ($R\bar{3}m$). The three possibilities for the inversion center position are consistent with the 3_1 operation resulting in three configurations.

$\bar{1}(S)$. This case will be a combination of the diagrams in Figs. 1(*d*) and 2(*a*), the resulting diagram is shown in Fig. 3(*a*). The periodicity will be a multiple of 6. The neutrality equation will be

$$(kXk\tilde{X})^3 = e \qquad (9)$$

but one sequence $kXk\tilde{X}$ must be non-neutral, otherwise it will be the smallest periodic unit. So the following constraint must be added:

$$kXk\tilde{X} \neq e. \qquad (10)$$

These equations are satisfied by a block X of charge e, k, kh and hk. The first code belonging to this symmetry is $[hk^3hk]_3 = |hk\,k\,kh\,k|hk\,k\,kh\,k|hk\,k\,kh\,k|$.

$\bar{1}(O)$. This case will be a combination of the diagrams in Figs. 1(*d*) and 2(*b*), the resulting diagram is shown in Fig. 3(*b*). The periodicity will be a multiple of 6. The neutrality equation will be

$$(X\tilde{X})^3 = e \qquad (11)$$

and

$$X\tilde{X} \neq e. \qquad (12)$$

These equations are satisfied by a block X of charge k, k^2, kh and hk.

The first code belonging to this symmetry is $[h^2k^2]_3 = |hk\,kh|hk\,kh|hk\,kh|$.

$\bar{1}(SO)$. This case will be a combination of the diagrams in Figs. 1(*d*) and 2(*c*), the resulting diagram is shown in Fig. 3(*c*). The periodicity will be a multiple of 3 but of the type $3(2N+1)$ where N is an integer number.

The neutrality equation will be

$$(Xk\tilde{X})^3 = e \qquad (13)$$

and

$$Xk\tilde{X} \neq e. \qquad (14)$$

These equations are satisfied by a block X of charge e, h, k^2.

The first code belonging to this symmetry is $[k]_3 = |k|k|k|$, which actually belongs to the special case of cubic symmetry $Fm\bar{3}m$ due to the ideal metric of the close-packed arrangement ($R\bar{3}m$ is a subgroup of $Fm\bar{3}m$).

3.2.2. The 6_3 and $\bar{1}$ operations ($P6_3/mmc$). Two possibilities arise, inversion center only at the center of spheres or only at the octahedral sites, $\bar{1}(SO)$ is excluded (Iglesias, 2006*a*):

$\bar{1}(S)$. This case will be a combination of Figs. 1(*b*) and 2(*a*), the resulting diagram is shown in Fig. 4(*a*). The periodicity will be a multiple of 4. In this case, the neutrality equation will be

$$(hXk\tilde{X})^2 = e \qquad (15)$$

and

$$hXk\tilde{X} \neq e, \qquad (16)$$

which allows all charges for the X blocks. The first code belonging to this symmetry is $[hk]_2 = |hk\,hk|$.

$\bar{1}(O)$. This case will be a combination of the diagrams in Figs. 1(*b*) and 2(*b*), the resulting diagram is shown in Fig. 4(*b*). The periodicity will be a even number of the type $2(2N+1)$, where N is an integer. The neutrality equation will be

$$(Xh\tilde{X})^2 = e \qquad (17)$$

and

$$Xh\tilde{X} \neq e, \qquad (18)$$

which allows all charges for the X blocks. The first code belonging to this symmetry is $[h]_2 = |h|h|$.

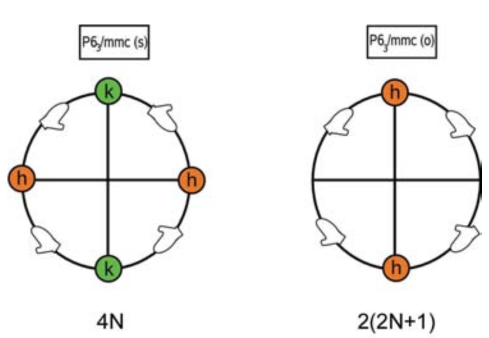

Figure 4
HK-code symmetry for a stacking arrangement with the combination of a 6_3 screw axis and an inversion center at (*a*) the sphere (*s*), (*b*) octahedral sites (*o*).

4. Efficient generation of polytypes

A bijection between the set of HK binary codes and the non-negative integer numbers can be made, and it is useful to define a lexicographic order over all possible HK words:

$0 \to$ (empty code)
$1 \to h$
$2 \to k$
$3 \to hh$
$4 \to hk$
$5 \to kh$
$6 \to kk$
$7 \to hhh$
$8 \to hhk$
$9 \to hkh$
$10 \to hkk$
$11 \to khh$
$12 \to khk$
$13 \to kkh$
$14 \to kkk.$

Table 2
The non-equivalent close-packing polytypes up to a length of 12 in binary notation.

0 corresponds to *h* and 1 to *k*. Lexicographic order (LO) and space group (SG) are also given.

LO	SG	Sequence	LO	SG	Sequence
3	$P6_3/mmc(o)$	0 0	14	$Fm\bar{3}m(so)$	1 1 1
20	$P6_3/mmc(s)$	0 1 0 1	38	$P\bar{3}m1(so)$	0 0 1 1 1
68	$P\bar{6}m2$,	0 0 0 1 0 1	90	$P6_3/mmc(o)$	0 1 1 0 1 1
134	$P\bar{3}m1(so)$	0 0 0 0 1 1 1	146	$P\bar{3}m1(so)$	0 0 1 0 0 1 1
174	$P\bar{3}m1(so)$	0 1 0 1 1 1 1	260	$P\bar{6}m2$	0 0 0 0 0 1 0 1
272	$P6_3/mmc(s)$	0 0 0 1 0 0 0 1	282	$P\bar{6}m2$	0 0 0 1 1 0 1 1
300	$P\bar{3}m1(o)$	0 0 1 0 1 1 0 1	318	$P\bar{3}m1(o)$	0 0 1 1 1 1 1 1
374	$P6_3/mmc(s)$	0 1 1 1 0 1 1 1	518	$P\bar{3}m1(so)$	0 0 0 0 0 0 c 1 1
530	$P3m1$	0 0 0 0 1 0 0 1 1	558	$P3m1$	0 0 0 1 0 1 1 1 1
584	$R\bar{3}m(so)$	0 0 1 0 0 1 0 0 1	598	$P3m1$	0 0 1 0 1 0 1 1 1
618	$P\bar{3}m1(so)$	0 0 1 1 0 1 0 1 1	734	$P\bar{3}m1(so)$	0 1 1 0 1 1 1 1 1
1028	$P\bar{6}m2$	0 0 0 0 0 0 0 1 0 1	1040	$P\bar{6}m2$	0 0 0 0 0 1 0 0 0 1
1050	$P\bar{6}m2$	0 0 0 0 0 1 1 0 1 1	1068	$P\bar{3}m1(o)$	0 0 0 0 1 0 1 1 0 1
1086	$P\bar{3}m1(o)$	0 0 0 0 1 1 1 1 1 1	1098	$P3m1$	0 0 0 1 0 0 1 0 1 1
1108	$P\bar{6}m2$	0 0 0 1 0 1 0 1 0 1	1122	$P6_3/mmc(o)$	0 0 0 1 1 0 0 0 1 1
1142	$P\bar{6}m2$	0 0 0 1 1 1 0 1 1 1	1182	$P\bar{3}m1(s)$	0 0 1 0 0 1 1 1 1 1
1188	$P6_3/mmc(o)$	0 0 1 0 1 0 0 1 0 1	1210	$P3m1$	0 0 1 0 1 1 1 0 1 1
1230	$P\bar{3}m1(o)$	0 0 1 1 0 0 1 1 1 1	1370	$P\bar{6}m2$	0 1 0 1 0 1 1 0 1 1
1406	$P\bar{3}m1(s)$	0 1 0 1 1 1 1 1 1 1	1518	$P6_3/mmc(o)$	0 1 1 1 1 0 1 1 1 1
2054	$P\bar{3}m1(so)$	0 0 0 0 0 0 0 0 1 1 1	2066	$P3m1$	0 0 0 0 0 0 1 0 0 1 1
2094	$P3m1$	0 0 0 0 0 1 0 1 1 1 1	2114	$P\bar{3}m1(so)$	0 0 0 0 1 0 0 0 0 1 1
2120	$P\bar{3}m1(so)$	0 0 0 0 1 0 0 1 0 0 1	2134	$P3m1$	0 0 0 0 1 0 1 0 1 1 1
2154	$P\bar{3}m1(so)$	0 0 0 0 1 1 0 1 0 1 1	2190	$P\bar{3}m1(so)$	0 0 0 1 0 0 0 1 1 1 1
2204	$P3m1$	0 0 0 1 0 0 1 1 1 0 1	2214	$P3m1$	0 0 0 1 0 1 0 0 1 1 1
2226	$P3m1$	0 0 0 1 0 1 1 0 0 1 1	2270	$P3m1$	0 0 0 1 1 0 1 1 1 1 1
2346	$P3m1$	0 0 1 0 0 1 0 1 0 1 1	2386	$P\bar{3}m1(so)$	0 0 1 0 1 0 1 0 0 1 1
2414	$P3m1$	0 0 1 0 1 1 0 1 1 1 1	2428	$P\bar{3}m1(so)$	0 0 1 0 1 1 1 1 1 0 1
2486	$P3m1$	0 0 1 1 0 1 1 0 1 1 1	2558	$P\bar{3}m1(so)$	0 0 1 1 1 1 1 1 1 1 1
2734	$P\bar{3}m1(so)$	0 1 0 1 0 1 0 1 1 1 1	2774	$P\bar{3}m1(so)$	0 1 0 1 1 0 1 0 1 1 1
3006	$P\bar{3}m1(so)$	0 1 1 1 0 1 1 1 1 1 1	4100	$P\bar{6}m2$	0 0 0 0 0 0 0 0 0 1 0 1
4112	$P\bar{6}m2$	0 0 0 0 0 0 0 1 0 0 0 1	4122	$P\bar{6}m2$	0 0 0 0 0 0 0 1 1 0 1 1
4140	$P\bar{3}m1(o)$	0 0 0 0 0 0 1 0 1 1 0 1	4158	$P\bar{3}m1(o)$	0 0 0 0 0 0 1 1 1 1 1 1
4160	$P6_3/mmc(s)$	0 0 0 0 0 1 0 0 0 0 0 1	4170	$P3m1$	0 0 0 0 0 1 0 0 1 0 1 1
4180	$P\bar{6}m2$	0 0 0 0 0 1 0 1 0 1 0 1	4194	$P\bar{6}m2$	0 0 0 0 0 1 1 0 0 0 1 1
4214	$P\bar{6}m2$	0 0 0 0 0 1 1 1 0 1 1 1	4236	$P3m1$	0 0 0 0 1 0 0 0 1 1 0 1
4254	$P3m1$	0 0 0 0 1 0 0 1 1 1 1 1	4260	$P\bar{3}m1(o)$	0 0 0 0 1 0 1 0 0 1 0 1
4282	$P3m1$	0 0 0 0 1 0 1 1 1 0 1 1	4302	$P3m1$	0 0 0 0 1 1 0 0 1 1 1 1
4326	$P\bar{3}m1(o)$	0 0 0 0 1 1 1 0 0 1 1 1	4372	$P\bar{3}m1(s)$	0 0 0 1 0 0 0 1 0 1 0 1
4386	$P\bar{3}m1(o)$	0 0 0 1 0 0 1 0 0 0 1 1	4392	$P\bar{6}m2$	0 0 0 1 0 0 1 0 1 0 0 1
4406	$P3m1$	0 0 0 1 0 0 1 1 0 1 1 1	4442	$P3m1$	0 0 0 1 0 1 0 1 1 0 1 1
4460	$P\bar{6}m2$	0 0 0 1 0 1 1 0 1 1 0 1	4478	$P3m1$	0 0 0 1 0 1 1 1 1 1 1 1
4502	$P3m1$	0 0 0 1 1 0 0 1 0 1 1 1	4522	$P\bar{6}m2$	0 0 0 1 1 0 1 0 1 0 1 1
4550	$P6_3/mmc(s)$	0 0 0 1 1 1 0 0 0 1 1 1	4590	$P\bar{6}m2$	0 0 0 1 1 1 1 0 1 1 1 1
4686	$P\bar{3}m1(o)$	0 0 1 0 0 1 0 0 1 1 1 1	4700	$P\bar{3}m1(s)$	0 0 1 0 0 1 0 1 1 1 0 1
4710	$P3m1$	0 0 1 0 0 1 1 0 0 1 1 1	4762	$P\bar{6}m2$	0 0 1 0 1 0 0 1 1 0 1 1
4780	$P3m1$	0 0 1 0 1 0 1 0 1 1 0 1	4798	$P3m1$	0 0 1 0 1 0 1 1 1 1 1 1
4810	$P6_3mc$	0 0 1 0 1 1 0 0 1 0 1 1	4854	$P3m1$	0 0 1 0 1 1 1 1 0 1 1 1
4914	$R\bar{3}m(o)$	0 0 1 1 0 0 1 1 0 0 1 1	4958	$P3m1$	0 0 1 1 0 1 0 1 1 1 1 1
4986	$P\bar{3}m1(o)$	0 0 1 1 0 1 1 1 1 0 1 1	5038	$P3m1$	0 0 1 1 1 0 1 0 1 1 1 1
5494	$P\bar{6}m2$	0 1 0 1 0 1 1 1 0 1 1 1	5562	$P\bar{3}m1(s)$	0 1 0 1 1 0 1 1 1 0 1 1
5886	$P\bar{3}m1(o)$	0 1 1 0 1 1 1 1 1 1 1 1	6110	$P6_3/mmc(s)$	0 1 1 1 1 1 0 1 1 1 1 1

In coding theory, the quasilexicographically smallest word among all the rotationally equivalent words (that is under all cyclic shifts of the word) is called a necklace (Lothaire, 1983). For example, from the equivalence class under rotation formed by the words

$$S1 = \{hkhkhh, hhkhkh, hhhkhk, khhhkh, hkhhhk, khkhhh\},$$

hhhkhk will be a necklace. If the necklace itself is not periodic, then it is called a Lyndon word. The set of all necklaces of length N will be denoted by $\mathrm{Nc}(N)$. If we add to our equivalence operations the reversion, then the lexicographically smallest word among all rotationally and reverse equivalent words is called a bracelet (Lothaire, 1983). For example, to the set analyzed above, the following reverse set must be added,

$$S2 = \{hhkhkh, hkhkhh, khkhhh, hkhhhk, khhhkh, hhhkhk\},$$

to form the equivalence class under rotations and reversion, and the unique bracelet will still be given by the word *hhhkhk*. The set of all bracelets of length N is denoted by $\mathrm{Br}(N)$. Owing to the symmetry of the HK code, the set $\mathrm{Plt}(N)$ of all non-equivalent polytypes of length N, described in the HK coding, will be the subset of $\mathrm{Br}(N)$ for which the neutrality condition holds. Formally,

$$\mathrm{Plt}(N) = \{\alpha \in \mathrm{Br}(N) \mid \alpha \text{ is neutral}\}. \quad (19)$$

An algorithm to exhaustively generate bracelets has been reported by Sawada (2001). The algorithm builds necklaces with an additional checking to discard those which will not result in bracelets. The algorithm runs in constant amortized time (CAT). An algorithm is said to be CAT if the computation is proportional to the number of objects generated. A brief description of the main logic behind this algorithm will be presented.

Given a code α formed by two words u and v by concatenation, $\alpha = uv$, u will be called a prefix of α. If α happens to be a necklace then u is called a prenecklace. The set of all binary prenecklaces of length P will be denoted by $P\mathrm{Nc}(P)$:

$$P\mathrm{Nc}(P) = \{u \mid uv \in \mathrm{Nc}(P+M), M \geq 0, u, v \text{ are words}\}. \quad (20)$$

The algorithm for generating necklaces is due to Cattel *et al.* (2000) and recursively adds a new symbol x to a prenecklace β of length P, such that βx is still a prenecklace. If $l = \mathrm{lyn}(\beta)$ is the length of the longest Lyndon word prefix of β, then $x = k$ if $\beta_{P-l} = k$ and l is unchanged, otherwise x can take any value h or k and $l = P + 1$. If $P + 1$ is divisible by l, then the generated word βx is a necklace, if $N + 1 = l$, then the necklace will be a Lyndon word. The algorithm starts with the simple word $\beta = h$ and generates necklaces in lexicographic order.

Sawada (2001) modified the above algorithm to spot, as soon as it was generated, if a prenecklace could not give rise to a bracelet. The idea is to avoid comparison of the generated prenecklace with all rotations of the reverse code which will add a heavy toll of $O(P^2\mathrm{Nc}(p))$ running time to the necklace CAT algorithm. In order to avoid this naive approach, some heuristic code was added to the necklace algorithm that checks if the prenecklace is valid as a prebracelet. The total time for this checking is proportional to the number of objects generated. The resulting algorithm is still CAT. To this code we now add an additional predicate to check if the resulting necklace is neutral. In order to do so, a parameter is added to the bracelet algorithm that keeps track of the charge of the generated word, in this way no additional complexity is added to the execution time of the algorithm. When the algorithm reaches the required length, if it is a neutral Lyndon word, the code is a valid HK code; if it is not aperiodic and the obtained code is formed by two identical blocks, then each block must be non-neutral, that is with charge h, hk or kh, and the whole code will be neutral and a valid HK code; finally, if the bracelet is neutral and formed by three identical blocks, then each block must be non-neutral, that is with charge belonging to the conjugate class $\{k, k^2\}$.

In Fig. 5, the plot of the execution time (in arbitrary units) per generated object against the polytype length is shown where, for sufficiently large word length L, the CAT behavior can be recognized. Table 2 shows the first polytypes up to length 12; the space group was identified using the results obtained for each symmetry as discussed above.

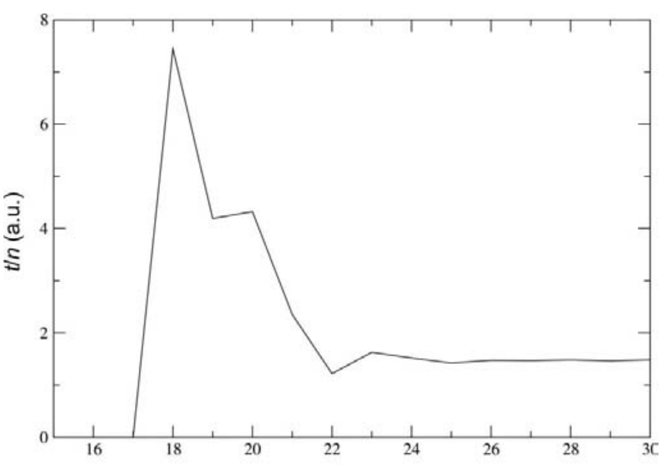

Figure 5
Execution time per number of objects generated (t/n) in arbitrary units against the length of the code (L). CAT behavior is easily recognized for L values larger than 22.

5. Conclusions

The HK coding of close-packed polytypes has been discussed. The symmetry of the HK code has been related to the symmetry of the polytype. The whole chart of HK code symmetries has been presented. An efficient algorithm with CAT execution time has been described. The algorithm allows all valid HK codes of a given length to be generated.[1]

Further research on a CAT algorithm to generate polytypes by space group is under way and will be reported elsewhere.

One of the authors (EER) acknowledges support from the Humboldt Foundation through an AvH equipment donation. JMM thanks the Ministry of Higher Education (MES) for a doctoral study grant. Both anonymous referees are thanked for pointing out several useful ideas that improved the article. We would like especially to thank one of the referees who made a thorough comment of the paper and discussed several issues of the HK coding concerning symmetry, the group operators and the use of notation, sharing with the authors their experience on the subject.

References

Cattel, K., Ruskey, F., Sawada, J., Serra, M. & Miers, C. R. (2000). *J. Algorithms*, **37**, 267–282.
Estevez-Rams, E., Azanza-Ricardo, C. & Aragon-Fernandez, B. (2005). *Z. Kristallogr.* **220**, 592.
Estevez-Rams, E., Azanza-Ricardo, C., Martínez García, J. & Aragón-Fernández, B. (2005). *Acta Cryst.* A**61**, 201–208.
Iglesias, J. E. (1981). *Z. Kristallogr.* **155**, 121–127.

[1] The C code of the polytype generation algorithm is available from the IUCr electronic archives (Reference: ZM5042). Services for accessing these data are described at the back of the journal.

Iglesias, J. E. (2006*a*). *Acta Cryst.* A**62**, 195–200.
Iglesias, J. E. (2006*b*). *Z. Kristallogr.* **221**, 237–245.
Jagodzinski, H. (1949). *Acta Cryst.* **2**, 201–207.
Lothaire, M. (1983). *Combinatorics on Words.* Cambridge University Press.
McLarnan, T. J. (1981). *Z. Kristallogr.* **155**, 269–291.
Patterson, A. L. & Kasper, J. S. (1959). *International Tables for Crystallography*, Vol. II. Birmingham: Kynoch Press.
Sawada, J. (2001). *SIAM J. Comput.* **31**, 259–268.
Verma, A. R. & Krishna, P. (1966). *Polymorphism and Polytypism in Crystals.* New York: Wiley.
Zhdanov, G. S. (1945). *C. R. (Dokl.) Acad. Sci. URSS*, **48**, 39–42.