

## Introduction to a general crystallography

**A. Janner**

Copyright © International Union of Crystallography

Author(s) of this paper may load this reprint on their own web site provided that this cover page is retained. Republication of this article or its storage in electronic databases or the like is not permitted without prior permission in writing from the IUCr.

## Introduction to a general crystallography

A. Janner

Institute for Theoretical Physics, University of Nijmegen, Toernooiveld, 6525 ED Nijmegen, The Netherlands. Correspondence e-mail: alo@sci.kun.nl

The definition of an extended crystallographic group is given, based on an  $n$ -dimensional Euclidean space, carrier of a faithful integral representation of a permutation group of atomic positions. The Euclidean crystallography of normal crystals and the higher-dimensional one applied to incommensurately modulated crystals, intergrowth crystals and quasicrystals are special cases of a general crystallography. The same is true for the multimetric crystallographic characterization of ice and of snow crystals. This approach can also be applied to single molecules, leading to what may be denoted as molecular crystallography. It possibly allows non-trivial structural relations between atomic positions belonging to the asymmetric unit of the molecular point group to be obtained. Two simple molecules, polycyclic aromatic hydrocarbons, are treated as illustrative examples.

© 2001 International Union of Crystallography  
Printed in Great Britain – all rights reserved

## 1. Preliminary remarks

The need to extend crystallography beyond Euclidean crystallography was already apparent with the superspace approach for incommensurate crystal phases which restores lattice symmetry in a higher-dimensional space only (Janner & Janssen, 1980*a,b*; Janssen *et al.*, 1999). The discovery of quasicrystals with scaling invariant diffraction peak positions (Ostlund & Wright, 1986; Shechtman *et al.*, 1984) also requires a suitable non-Euclidean extension because scaling does not leave the Euclidean distance invariant. Consider, as illustration, the one-dimensional Fibonacci chain (observed in many structures), which is characterized by a sequence of two intervals  $L$  and  $S$  in the golden mean ratio  $L/S = \tau = (1 + 5^{1/2})/2$ , invariant with respect to the scale transformation:  $S \rightarrow L$  and  $L \rightarrow L + S$ . This transformation is expressible as a two-dimensional invertible integral matrix and the square of it is a hyperbolic rotation in the superspace leaving invariant a square lattice. The Fibonacci scale transformation is of such fundamental importance for the structural properties of a Fibonacci chain that it justifies the necessity to extend Euclidean crystallography to include hyperbolic rotations. The appearance of circular and hyperbolic rotations in the same group is reflected in the name *multimetric crystallography* (Janner, 1991, 1995) we have adopted. Multimetric crystallography can also be applied to normal crystals. In particular, the atomic structure of ice is left invariant by a multimetric space group. Its point group, of infinite order, allows the interpretation of structural features in snow crystals (Janner, 1997).

The possible relevance of an analogous multimetric point group for sixfold helical nucleic acid molecules is revealed by hexagonal scaled forms similar to those observed in snow crystals, but now occurring at the microscopic molecular level.

In the molecules involved, a number of atomic positions in the asymmetric unit of the Euclidean line symmetry group, if considered modulo translations along the helical axis, can be related by elements of this extended point group. This observation opened the door for a crystallographic approach applied to (idealized) molecular structures occurring, in particular, in biomacromolecules with a given axial rotation symmetry (Janner, 2000, 2001).

A formal definition of an *extended crystallographic group* given in §2 applies to all the various crystallographic approaches mentioned above, as shown in §3. Before considering the technical details, one should keep in mind that crystallography (general or not) always describes structural relations of a Euclidean system, even if possibly expressed by non-Euclidean transformations. The Euclidean character of the system implies that the Euclidean group is the covariance group, which is the symmetry group of Euclidean geometry, the invariance transformations, however, need not be Euclidean.

In the course of the exploration of biomacromolecules, the conviction grew that, to reach a better understanding of what is going on at the physical level, one should consider generic molecules and not biomacromolecules only. In other words, it is a question of physics and of chemistry and not of biochemistry only. To justify this expectation, illustrative molecular examples reported in §4 are supramolecules, having benzene as modular repeat unit (Berresheim *et al.*, 1999). In these cases, it is the sixfold symmetry of benzene that allows one to adopt the same point group of infinite order, as for snow flakes, but at the same atomic level as for hexagonal nucleic acids.

The question of the physical relevance of the present approach is still open. In any case, its physical foundation is ensured because all the structures involved occur in nature

and obey, therefore, the laws of physics (and of chemistry). From this general statement, one cannot expect that one is also able to indicate in each concrete case the specific connections between structure and physics due to the new observed structural relations.

Last but not least a warning: the present formulation is at a working level only, even if considered by the author to be essentially correct. He is aware that a meaningful generalization of the classical crystallographic framework requires the involvement of many more scientists than himself only, and that future work can lead to important conceptual changes.

## 2. Basic concepts

Crystal structures are considered as point sets in a way as described in the book *Crystallographic and Metacrystallographic Groups* by W. Opechowski, warmly recommended to readers interested to know more (Opechowski, 1986). There, invariance groups of functions defined on crystal structures were considered. Here, the concept of a crystallographic group itself is extended to include non-Euclidean transformations also, and the crystallographic structures considered are more general than those of crystals in a point-atom approximation but the aim is the same: to convey the underlying mathematical ideas not only to crystallographers but to physicists, chemists and biologists as well.

### 2.1. Model and real structures

Considered are model structures  $X$  specified in terms of positions given as a discrete set of points  $x$  in the three-dimensional Euclidean space:

$$X = \{x | x \in \mathcal{E}(3)\}. \quad (1)$$

Real structures  $Y$  are given in the same way, together with a correspondence  $\rho$  between model and real positions:

$$\rho : X_0 \longrightarrow Y_0 \quad \text{for} \quad \rho(X_0) = Y_0 \subseteq Y, \quad X_0 = \rho^{-1}(Y_0) \subseteq X \quad (2)$$

having the following properties:

(i) the model structure is an approximation of the real one within a given upper bound  $\varepsilon$

$$|x - y| \leq \varepsilon \quad \text{for} \quad y = \rho(x);$$

(ii) the definition domain of  $\rho$  is in general a subset  $X_0$  of the model structure  $X$  (typically, the correspondence between model positions and real structure involves a finite number of atoms, whereas the model set can be infinite);

(iii) the mapping  $\rho$  is injective (it is thus a monomorphism):  $\rho(x) \neq \rho(x')$  implies  $x \neq x'$  and conversely;

(iv) the image of  $\rho$  is in general a subset  $Y_0$  of the whole structure  $Y$  (a number of atoms are possibly disregarded).

At the microscopic level, the positions are atomic positions. For macroscopic structures, positions may represent inclusions or averaged structures. For macromolecules, different hierarchical levels are possible, where positions are associated

with single atoms or with atomic clusters of increasing complexity. In the case of crystals, the restrictions given above are considered to be self-evident without any conceptual consequences. The deviations between real and determined structures are expressed in terms of an  $R$  factor,  $\varepsilon$  being the resolution. One mostly forgets about defects and interstitial atoms. A periodic boundary condition allows one to bridge the gap between finite structure and infinite crystal.

In the case of aperiodic crystals, the same restrictions require an additional change of perspective and new concepts. For modulated and intergrowth crystals, a higher-dimensional description in a Euclidean *superspace* is adopted leading to *modulation lines* and to *atomic surfaces* instead of point atoms. For quasicrystals, an analogous superspace approach implies, within the *cut-projection method*, a projection in real space of positions occurring in the *cut region*.

Finally, in the biomacromolecular case, the concept of partial structure arises naturally when speaking of *backbone* and that of *non-crystallographic symmetry* takes into account a misfit between the structural properties of molecules, or of molecular fragments, considered locally, and those compatible with the crystallography of their crystalline order (Rossmann & Arnold, 1993). *Pseudosymmetry* is used when a non-negligible part of the atoms approximately has a higher symmetry than the whole structure (Viterbo, 1992). Here, we avoid speaking of non-crystallographic symmetries because the aim is to show that these symmetries are crystallographic, even if in a more general sense only. We adopt instead and generalize the term pseudosymmetry, which reflects the spirit of the present approach.

All this leads to the specification of a generalization of the meaning of crystallographic structure and of a crystallographic group common for all the above structures.

### 2.2. Crystallographic structures and groups

Crystallography is based on an interplay between geometry, number theory (arithmetic in particular) and algebra. These three aspects remain essential in the extension given here. As already pointed out, a crystallographic structure obeys the laws of Euclidean geometry, so that the Euclidean group is the *covariance group* of the system (an *invariance group* leaves the system invariant, whereas a covariance group leaves invariant the laws of the system). The equivalent positions involved can be labeled by sets of rational integers, the *indices*, because of underlying lattices (direct and reciprocal) and there is a non-trivial invariance group of the indexed positions. These requirements, which strictly speaking only apply to ideal structures, are reflected in the definition of an extended crystallographic group. A structure having such a group as invariance group is then also called an (extended) crystallographic structure.

**Definition 1.** A group  $G$  of permutations of a discrete set of points  $X = \{x | x \in \mathcal{E}(3)\}$  is called an *extended crystallographic group* if there is an  $n$ -dimensional Euclidean representation

space  $\mathcal{E}(n)$  carrier of faithful linear integral representations  $\Gamma(G)$  and  $\Gamma^*(G)$ .

This implies:

(i)  $GX = X$  with  $G \subseteq S(X)$ , where  $S$  is the symmetric group of the set  $X$ ;

(ii) there are in  $\mathcal{E}(n)$  dual bases  $a = \{a_1, \dots, a_n\}$  and  $a^* = \{a_1^*, \dots, a_n^*\}$  such that:  $\Gamma(g) = g(a) \in Gl(n, \mathbb{Z})$  and  $\Gamma^*(g) = g(a^*) \in Gl(n, \mathbb{Z})$  for any  $g \in G$ , where  $Gl(n, \mathbb{Z})$  is the group of the invertible  $n$ -dimensional integral matrices;

(iii)  $\Gamma(G) \simeq \Gamma^*(G) \simeq G$ , with  $\simeq$  denoting group isomorphism.

Accordingly, one has:

$$\Gamma(G) \subseteq Gl(n, \mathbb{Z}) \quad \text{and} \quad \Gamma^*(G) \subseteq Gl(n, \mathbb{Z}). \quad (3)$$

The Euclidean structure of the representation space is needed to allow one to derive from the dual representations  $\Gamma$  and  $\Gamma^*$  the corresponding relations between atomic positions and diffraction spots and/or delimiting planes. In general, however,  $\Gamma(G)$  and  $\Gamma^*(G)$  need not leave invariant the Euclidean metric of  $\mathcal{E}(n)$ .

**Definition 2.** An *extended crystallographic structure* is a discrete point set  $X$  whose invariance group is an (extended) crystallographic group  $G$ . This group is a *pseudosymmetry group* (as defined further on) of the corresponding real structure. The *symmetry group* of the structure is a Euclidean subgroup  $G_0$  of  $G$ .

In many cases, but not always,  $G_0$  is the maximal Euclidean subgroup of  $G$ , so that  $G_0 = G \cap E(3)$ . Normally, the point set  $X$  is infinite; it can also be finite. The discreteness is essential: According to this definition, the charge-density distribution of a crystal is not a crystallographic structure.

The model crystallographic structure  $X$  is partitioned into sets of  $G$ -equivalent points. Each set forms an *orbit*, generated by  $G$  from one point of the orbit. The group  $G$  is not an invariance group of the real structure  $Y$  and, in general, not even of the set  $X_0 = \rho^{-1}(Y_0)$  of the corresponding model positions. This  $X_0$  is, however, also partitioned into sets of  $G$ -equivalent points, belonging thus to one orbit, but not forming, in general, a complete orbit. The group elements involved in this equivalence imply structural relations for  $X_0$  and thus also for the real structure, even if in an approximate way only.

**Definition 3.** An extended crystallographic group  $G$  is a *pseudosymmetry group* of a real structure  $Y$  if for  $\rho$  the mapping considered above from the corresponding model denoted by  $X$ , for given small deviations with an  $\varepsilon$  upper bound, one has:

$$X_0 \subseteq GX_0 \quad \text{for} \quad X_0 = \rho^{-1}(Y_0), \quad X_0 \subseteq X \quad \text{and} \quad Y_0 \subseteq Y$$

and the asymmetric unit with respect to the pseudosymmetry group  $G$  is smaller than the one with respect to the (Euclidean) symmetry group  $G_0$ , subgroup of  $G$ .

While this definition takes into account the approximate character admitted for a pseudosymmetry in the usual sense,

no condition is imposed here on 'a non-negligible part' of the atoms involved.

## 2.3. Crystallographic forms

The laws of crystallography have been discovered phenomenologically by looking at *crystal growth forms*. The theoretical idea of an underlying lattice of symmetry translations became accepted on the basis of two experimental properties of growth forms (Friedel, 1911):

(a) the *law of rational indices*, which allows assignment to flat facets a set of three indices  $h, k, l$  (small integers), interpreted as components of a reciprocal-lattice vector  $(h, k, l) = ha_1^* + ka_2^* + la_3^* \in \Lambda^*$ , perpendicular to the facet;

(b) the classification of all observed growth forms into 32 crystal classes, obtained by imposing on conjugated finite subgroups of  $O(3)$  the condition to leave invariant the corresponding dual set of direct-lattice points  $[u, v, w] = ua_1 + va_2 + wa_3 \in \Lambda$ .

Accordingly, a crystal growth form is delimited by one or more sets of point-group-equivalent lattice planes. From crystal morphology, one only gets ratios between lattice vectors, so that nothing can be said about their absolute lengths. Therefore, the lattice can equally well be mesoscopic or microscopic. On the basis of X-ray diffraction, one identifies  $\Lambda$  and  $\Lambda^*$  with the microscopic lattice of symmetry translations and of the positions of the Bragg peaks, respectively.

In snow flakes, however, macroscopic lattices  $\Lambda$  and  $\Lambda^*$  both have a morphological meaning. Indeed, snow crystals are characterized by *scaled growth forms* (Janner, 1997). As shown further on, these scaled forms allow a double set of rational indices: one for the planes (lines in the two-dimensional description) and one for points (at water droplets or at dendritic bifurcation points). Scaled forms of the same type enclose hexagonal biomacromolecules like poly(C) (Janner, 2000, 2001), but now for molecular indexing lattices (to be distinguished from the translation lattice of the corresponding crystalline phase).

The *indexing lattices*  $\Lambda$  and  $\Lambda^*$  are generated by the bases of the  $n$ -dimensional representations  $\Gamma$  and  $\Gamma^*$ , respectively, defined above for (extended) crystallographic structures. In particular, for scaled forms one has:

**Definition 4.** A *scaled form* consists of point-group-equivalent planes through points equivalent by the same extended crystallographic point group  $K$ .

Accordingly, these scaled planes  $p$  can be indexed as:

$$p = (h_1, \dots, h_n)[m_1, \dots, m_n] \quad \text{with} \quad (h_1, \dots, h_n) \in \Lambda^*, \quad [m_1, \dots, m_n] \in \Lambda \quad (4)$$

with  $\Lambda^*$  and  $\Lambda$  freely generated by  $a^* = \langle a_1^*, \dots, a_n^* \rangle$  and by  $a = \langle a_1, \dots, a_n \rangle$  bases of the representations  $\Gamma^*(K)$  and  $\Gamma(K)$ , respectively.

There are two possible equivalence relations for scaled planes, which are compatible with the definition given above. A *strong equivalence* requires:

**Table 1**

Comparative characterization of (extended) crystallographic forms.

| Physical system                                                                                         | Crystals                            |                                                               | Aperiodic crystals<br>incommensurate and<br>quasicrystals<br>$n > 3$ | Molecules with a crystallographic<br>pseudosymmetry group<br>$n \geq 3$ |
|---------------------------------------------------------------------------------------------------------|-------------------------------------|---------------------------------------------------------------|----------------------------------------------------------------------|-------------------------------------------------------------------------|
|                                                                                                         | Normal case<br>$n = 3$              | Snow flakes<br>$n = 2(3)$                                     |                                                                      |                                                                         |
| Geometric form $F$                                                                                      | Growth form                         | Scaled growth form                                            | Growth form                                                          | Scaled enclosing form                                                   |
| Point group $K$<br>$\Gamma(K) \subseteq Gl(n, \mathbb{Z})$<br>$\Gamma^*(K) \subseteq Gl(n, \mathbb{Z})$ | Finite                              | Infinite                                                      | Finite                                                               | Infinite                                                                |
| Lattice group $\Lambda$<br>$\Lambda \simeq \mathbb{Z}^n$<br>$K\Lambda = \Lambda$                        | Translations                        | Indexing of plane<br>positions                                | Quasitranslations<br>$\mathbb{Z}$ -module of rank $n$                | Indexing of plane<br>positions                                          |
| Reciprocal lattice $\Lambda^*$<br>$\Lambda^* \simeq \mathbb{Z}^n$<br>$K\Lambda^* = \Lambda^*$           | Indexing of plane<br>orientations   | Indexing of plane<br>orientations                             | Indexing of plane<br>orientations                                    | Indexing of plane<br>orientations                                       |
| $F$ delimiting planes                                                                                   | $\Lambda$ -orbit planes<br>$F = KF$ | $\Lambda$ -orbit planes = $K$ -orbit planes<br>$F \subset KF$ | $\mathbb{Z}$ -module orbit planes<br>$F = KF$                        | $K$ -orbit planes<br>$F \subset KF$                                     |

Remarks: [1]  $n$  is the dimension of the (Euclidean) representation space of a linear integral faithful representation of the point group  $K$ . [2] The quasitranslations are orthogonal projections, into the space of the crystal, of  $n$ -dimensional lattice translations in the *superspace*, which is a Euclidean representation space. Self-similarity of quasicrystal structures has been disregarded here. [3]  $F$ -planes delimiting holes do not only occur in molecular enclosing forms, but have also been observed in quasicrystals (Janot *et al.*, 1999; Mancini *et al.*, 1998).

$$\begin{aligned}
 p' \approx p &\longleftrightarrow (h'_1, \dots, h'_n)[m'_1, \dots, m'_n] \\
 &= \alpha(a^*)(h_1, \dots, h_n)\alpha(a)[m_1, \dots, m_n] \\
 &\quad \text{for } \alpha \in K,
 \end{aligned} \quad (5)$$

whereas for the *weak equivalence* it is sufficient to have:

$$\begin{aligned}
 p' \sim p &\longleftrightarrow (h'_1, \dots, h'_n)[m'_1, \dots, m'_n] \\
 &= \mathcal{O}_{(h_1, \dots, h_n)}(K)\mathcal{O}_{[m_1, \dots, m_n]}(K)
 \end{aligned} \quad (6)$$

for  $\mathcal{O}$  an orbit of  $K$  in the representation space.

Experimental data seem to favor the weak equivalence, but it makes sense to consider both. Scaled forms are one of several possible crystallographic forms, whose characteristic features are summarized in Table 1.

### 3. Crystallographic approaches in a nutshell

The aim is to show how the various approaches mentioned in the first section satisfy the same set of definitions given in §2 and are, therefore, special cases of a general crystallography.

#### 3.1. Euclidean crystallography

Discussed is the three-dimensional crystallography, which is the normal case. For other dimensions, the properties are correspondingly the same.

A model crystal  $X$  is characterized by lattice periodicity:

$$\Lambda X = X \subset \mathcal{E}(3) \quad \text{for } \Lambda = \langle a_1, a_2, a_3 \rangle \simeq \mathbb{Z}^3. \quad (7)$$

Here, a lattice is identified with the group of lattice translations. Let us now consider typical cases.

As extended crystallographic group, the lattice translation group  $\Lambda$  can be associated with a four-dimensional repre-

sentation space spanned by  $a_0, a_1, a_2, a_3$  with  $a_0$  perpendicular to the space of the crystal.

$$\Lambda \simeq \Gamma(\Lambda) \subset Gl(4, \mathbb{Z}) \quad \text{with} \quad \Gamma(t) = \left( \begin{array}{c|c} \mathbb{1} & t(a) \\ \hline 0 & 1 \end{array} \right), \quad t \in \Lambda, \quad (8)$$

with  $t(a)$  an integral column vector expressing  $t$  with respect to the lattice basis  $a = \{a_1, a_2, a_3\}$ , and  $\mathbb{1}$  is the three-dimensional unit matrix. One clearly also has

$$\begin{aligned}
 \Lambda^* &\simeq \Gamma^*(\Lambda^*) \subset Gl(4, \mathbb{Z}) \\
 &\quad \text{with} \quad \Gamma^*(t) = \left( \begin{array}{c|c} \mathbb{1} & t(a^*) \\ \hline 0 & 1 \end{array} \right), \quad t \in \Lambda^*
 \end{aligned} \quad (9)$$

with  $t$  now expressed with respect to the dual basis  $a^*$ .

A point group  $K$  leaving  $\Lambda$  invariant is not only crystallographic but also extended crystallographic. The representation space is three-dimensional and identified with the space of the crystal.

$$K \simeq \Gamma(K) \subset Gl(3, \mathbb{Z}), \quad \Gamma(\alpha) = \alpha(a) \quad \text{for } \alpha \in K \subset O(3), \quad (10)$$

where  $\alpha(a)$  denotes the orthogonal transformation  $\alpha$  expressed in the lattice basis  $a$ . As

$$\alpha(a^*) = \tilde{\alpha}^{-1}(a), \quad (11)$$

where  $\sim$  denotes transposition, one also has  $K \simeq \Gamma^*(K) \subset Gl(n, \mathbb{Z})$ .

The point-group representation space used in *International Tables for Crystallography* (Hahn, 1992) is not always three-dimensional. Indeed, the four-index notation adopted for the trigonal and hexagonal systems defines a four-dimensional representation space, as discussed by Frank (1965).

A space group is also an extended crystallographic group. A non-symmorphic space group is always a subgroup of a symmorphic one, where the non-primitive translations of the subgroup are lattice translations (Ascher & Janner, 1968). Therefore, it is sufficient to consider a symmorphic space group  $G$ . In this case, writing

$$g = \{\alpha|t\} \in G \quad \text{with} \quad \alpha \in K, \quad t \in \Lambda, \quad (12)$$

one has  $\Gamma(G) \simeq G$  defined by

$$\Gamma(g) = g(a) = \begin{pmatrix} \alpha(a) & t(a) \\ 0 & 1 \end{pmatrix} \in Gl(4, \mathbb{Z}), \quad g = \{\alpha|t\} \in G. \quad (13)$$

Again,  $\Gamma^*(G)$  is also crystallographic.

### 3.2. Superspace crystallography

Superspace crystallography characterizes the symmetry of aperiodic crystals like incommensurately modulated, intergrowth or composite crystals (Janssen *et al.*, 1999) and of quasicrystals (Janssen, 1986; Yamamoto, 1996). The representation space is  $n$ -dimensional Euclidean and called *superspace*.

To begin with, let us consider a  $d$ -dimensional incommensurately modulated crystal in three dimensions  $X \subset \mathcal{E}(3)$ . The Euclidean asymmetric unit of  $X$  is infinite and in one-to-one correspondence with the finite asymmetric unit of a  $(3+d)$ -dimensional crystal with respect to a space group  $G_s$  (a superspace group, more precisely). This higher-dimensional space group induces in three-space an isomorphic group of permutations  $G$  of the atomic positions in  $X$ .

$$G \subset S(X), \quad G \simeq G_s \simeq \Gamma(G), \\ G_s \subset E(3+d), \quad \Gamma(G) \subset Gl(4+d, \mathbb{Z}). \quad (14)$$

The symmetry group  $G$  of a modulated crystal is a non-Euclidean permutation group in three dimensions, but is isomorphic to a Euclidean space group  $G_s$  in  $(3+d)$  dimensions. The superspace can thus be considered as a representation space for the extended crystallographic group  $G$ , in the same way as in (8) for the three-space.

For quasicrystals, the situation is similar but not exactly the same because of the existence of a one-to-one correspondence (given by a projection  $\pi$ ) between a basis  $a_s = a_{1s}, \dots, a_{ns}$  of a lattice  $\Lambda_s$  in the  $n$ -dimensional superspace and a projected basis  $a = a_1, \dots, a_n$  spanning the three-dimensional space of the quasicrystal, linearly independent of the rational numbers and generating a free  $\mathbb{Z}$ -module  $M$  of rank  $n$ . For a modulated crystal, this is only the case for the dual  $\Lambda^*$  and  $M^*$  in the corresponding reciprocal space because some of the elements of the basis  $a_s$  can project in three-space to zero. In particular, and in the standard basis representation, the modulation wave translations are along a direction perpendicular to the three-dimensional space. For quasicrystals, the relation between the two bases  $a$  and  $a_s$  (and correspondingly  $a^*$  and  $a_s^*$ ) can be found in the paper by Yamamoto quoted above (Yamamoto, 1996).

The dimension of the representation space depends on the point group  $K$ , which is a finite subgroup of  $O(3)$  and has always a faithful integral representation  $\Gamma(K)$  of minimal dimension  $n$  with basis  $a$  of the three-space as above. In the particular case of a cyclic point group of order  $m$ , and for  $m \geq 3$ , the minimal dimension  $n$  is given by  $\varphi(m) + 1$ , where  $\varphi$  is the Euler function. The function  $\varphi(m)$  counts the number of integers  $i$  not exceeding  $m$  and coprime with  $m$  (see *e.g.* Grosswald, 1966). For  $m$  prime,  $\varphi(m) = m - 1$  and the  $\mathbb{Z}$ -module basis  $a$  is given by

$$a_1, \alpha a_1, \dots, \alpha^{m-2} a_1, a_3 \\ \text{with} \quad \alpha a_3 = a_3 \perp a_1, \quad \alpha^m = 1, \quad \alpha \in O(3), \quad (15)$$

where  $\alpha$  is the generating rotation and  $m \geq 3$ . As  $\Gamma(K)$  is a finite subgroup of  $Gl(n, \mathbb{Z})$ , it can be considered as a group of orthogonal transformations  $K_s$  in the superspace, expressed with respect to the lattice basis  $a_s$ . The projected lattice translations are then possible translational symmetries of the quasicrystals and are called *quasilattice translations*. One has:

$$KM = M \simeq \mathbb{Z}^n, \quad K_s \Lambda_s = \Lambda_s, \quad K \simeq K_s \simeq \Gamma(K) \subset Gl(n, \mathbb{Z}) \quad (16)$$

with  $M = \pi \Lambda_s$  and  $\pi$  is the projection considered above. Point group and (quasi)lattice translations can then be combined in the usual way to form an  $n$ -dimensional space group. The orthogonal character of the point group ensures the validity of corresponding dualities, which in the superspace are given by the Euclidean duality between the direct and the reciprocal spaces. Accordingly, one can conclude that the symmetry group of a quasicrystal is an extended crystallographic one.

### 3.3. Multimetric crystallography

The crystallographic approaches described so far have been extended to multimetric ones in order to include self-similarity in the crystallography of quasicrystals. Indeed, a self-similarity  $\sigma$  of the module  $M$  of quasitranslations leads in the superspace to a hyperbolic rotation leaving invariant the corresponding  $n$ -dimensional lattice  $\Lambda_s$  of symmetry translations. It is, therefore, an element of an orthogonal group  $O(n - q, q)$ , for  $q$  a positive integer smaller than  $n$ , leaving an indefinite metric invariant and having an integral matrix representation  $\Gamma(M) = \Gamma(\Lambda)$ . The point group  $K$  consists of scale-rotation transformations. In the simplest case, one has:

$$K = \langle \alpha, \dots, \sigma, \dots \rangle \simeq \Gamma(K) \subseteq Gl(n, \mathbb{Z}) \quad (17)$$

for  $KM = M$  together with a superspace representation

$$\Gamma(\alpha) = \alpha(a_s) \in O(n) \cap Gl(n, \mathbb{Z}) \\ \Gamma(\sigma) = \sigma(a_s) \in O(n - q, q) \cap Gl(n, \mathbb{Z}) \quad (18)$$

defining a superspace point group  $K_s$  of hyperbolic and circular rotations leaving  $\Lambda_s$  invariant. The non-trivial property, important from the physical point of view, is that the same  $K_s$  also leaves invariant the lattice  $\Lambda_s^*$  reciprocal to  $\Lambda_s$  with respect to the Euclidean metric in the superspace. For the corresponding reciprocal basis  $a_s^*$ , one indeed has

$$\begin{aligned}\Gamma^*(\alpha) &= \alpha(a_s^*) = \tilde{\alpha}^{-1}(a_s) \\ \Gamma^*(\sigma) &= \sigma(a_s^*) = \tilde{\sigma}(a_s)\end{aligned}\quad (19)$$

for  $\alpha$  a circular rotation and  $\sigma$  a hyperbolic rotation, with  $\sim$  denoting transposition. The more general case is a combination of these two pure cases. The same relations then follow in space for the dual modules  $M$  and  $M^*$  and the corresponding dual  $Z$ -module bases  $a$  and  $a^*$ . As one also has the possibility of defining lattice translations in the superspace and corresponding (projected) quasitranslations in space, one derives multimetric space groups  $G_s$  in the usual way. Moreover, the Euclidean subgroup  $G_{0s}$  is the symmetry group of the quasicrystal when self-similarity is not taken into account.

For  $n = 3$ , the superspace can be identified with the physical space and one can consider a multimetric space group leaving a given (normal) crystal invariant. In this case, strictly speaking, there are no scaling transformations (which never leave a lattice invariant). Scaling appears in a more subtle way, macroscopically through scaled growth forms and microscopically through properties of the Wyckoff positions. Indeed, Wyckoff positions for the multimetric group consist of complete sets of Wyckoff positions for the Euclidean subgroup and occur in scaled related families (Janner, 1995).

The multimetric invariance space group for the ice structure has been determined. Its point group, of infinite order, has  $6/mmm$  as Euclidean subgroup, leaves the hexagonal lattice invariant, is generated by

$$K = \langle L_z, R_z, m_z, m_x \rangle = (4)6/mmm \quad (20)$$

and has a faithful matrix representation  $\Gamma(K)$  defined by the matrices

$$\begin{aligned}R_z(a) &= \begin{pmatrix} 1 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, & L_z(a) &= \begin{pmatrix} 3 & 1 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \\ m_z(a) &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, & m_x(a) &= \begin{pmatrix} -1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.\end{aligned}\quad (21)$$

This point group has allowed us to interpret the hexagonal scaled growth forms observed in snow crystals (Janner, 1997).

What has been presented so far indicates how one can show that multimetric crystallography is another special case of the general crystallography defined in §2, even if in this paper we do not give a full treatment.

### 3.4. Molecular crystallography

The extended symmetry of a molecule can be approached in the same way as for quasicrystals. Indeed, let  $X_0$  describe the positions of the atoms (irrespective of atom type) of a molecule with a point-group symmetry  $K_0$ , which is a finite subgroup of  $O(3)$ . As in the case of quasicrystals, there is a set of vectors  $a_1, \dots, a_n$  in space, linearly independent of the rational numbers, spanning a faithful integral representation  $\Gamma(K_0)$ . This matrix group can be regarded as a subgroup of

the orthogonal group  $O(n)$  expressed with respect to a basis  $a_s = \langle a_{1s}, \dots, a_{ns} \rangle$  of a Euclidean  $n$ -dimensional space (the superspace), generating a lattice  $\Lambda_s$ , whereas the basis  $a$  generates a three-dimensional  $Z$ -module  $M$  of rank  $n$ . The basis  $a_s$  can be oriented to give by a one-to-one projection in space the  $Z$ -module basis  $a$ . The dual basis  $a_s^* = \langle a_{1s}^*, \dots, a_{ns}^* \rangle$  then generates the reciprocal lattice  $\Lambda_s^*$ , defines the representation  $\Gamma^*(K) \subset Gl(n, \mathbb{Z})$  and projects in three-space to a basis  $a^*$  of the dual  $Z$ -module  $M^*$ .

One can always take as first basis vector  $a_1$  the vector pointing to a freely chosen atomic position  $x_0 \in X_0$  considered from the origin left invariant by  $K_0$ , which then is the projection of  $a_{1s} = [1, 0, \dots, 0]$ . The points belonging to the  $K_0$  orbit of  $x_0$ , when expressed with respect to the basis  $a$ , have integral coordinates, which are the indices of the corresponding points of the orbit. So far, nothing is gained, as it is simply a relabeling.

Interesting and not trivial is the case when there is an invertible linear transformation  $\sigma$  not belonging to  $K_0$ , leaving both  $Z$ -modules  $M$  and  $M^*$  of the molecule invariant ( $\sigma$  is then an automorphism of  $M$  and of  $M^*$ , respectively). In this case, the point group  $K_0$  can be extended to  $K$  by including  $\sigma$ , still keeping its (extended) crystallographic property. Indeed, for  $K = \langle K_0, \sigma \rangle$ , one has

$$\sigma(a_s) \in \Gamma(K) \subseteq Gl(n, \mathbb{Z}), \quad K \simeq \Gamma(K), \quad KM = M \quad (22)$$

and corresponding relations for  $\sigma(a_s^*)$ ,  $\Gamma^*(\sigma)$  and  $M^*$ . This procedure can possibly be repeated. Eventually, the group  $K$  one obtains is a pseudosymmetry group for  $X_0$ , if it implies non-trivial structural relations expressible by elements of  $K$  among the positions belonging to the Euclidean asymmetric unit of the symmetry group  $K_0$  of the molecule. Note that the point group  $K$  is a group of permutations of *positions*, not of atoms of the same type, so that structural relations are also possible between atoms of different species. This fact has been one of the main surprising results obtained from a first phenomenological exploration of scale rotations in nucleic acids and proteins.

In general,  $K$  does not leave  $X_0$  invariant. The orbits of  $K$  from  $X_0$  form the larger set  $X$  of model positions. The need to make a distinction between real and model positions arises from the fact that a non-trivial extension is very often only possible for a more regular structure than the real one.

In specific cases, one meets not only the difficulty of determining a suitable model (idealized) to be mapped into the real structure, but also the fact that the infinite subgroups of  $Gl(n, \mathbb{Z})$  are badly known. Moreover, for  $n > 3$ , the orbit points in three-space are dense and can have arbitrarily large indices, whereas, for the law of rational indices, only small integers make sense. As already pointed out, for biomacromolecules it is not even required that the positions one has to consider correspond to single atoms but can refer to still badly defined atomic clusters.

Under these circumstances, it is highly surprising that there are molecules for which one may recognize non-trivial pseudosymmetry groups.

#### 4. Illustrative molecular examples

The examples given in this section are very simple and have been taken from a review paper on polyphenylene nanostructures (Berresheim *et al.*, 1999) based on benzene units. Two cases are considered of polycyclic aromatic hydrocarbons (PAH) with a two-dimensional architecture, approached here from the point of view of molecular crystallography. The first case, a planar one, is  $C_{54}H_{22}$  on graphite with  $D_{2h}$  point-group symmetry (Dias, 1993; Halleux *et al.*, 1958). The second case, a twisted one, is hexaphenylbenzene in a propeller configuration with approximate sixfold symmetry (Bart, 1968).

The first step in a crystallographic characterization of a given molecule requires the distinction between real and model structures. In the case of oligophenylene, the radius of a benzene ring is about 1.4 Å and the length of an inter-ring single bond 1.5 Å (Berresheim *et al.*, 1999). For the model structure, one keeps the same hexagonal arrangement of the benzene rings and the same distances between the corresponding centers, but with an equal value for both the radius of a benzene ring and for the inter-ring distance, which is then 1.433 Å. This choice represents the best fit of the model with the observed structure.

##### 4.1. $C_{54}H_{22}$ on graphite

This molecule consists of nine benzene rings (Dias, 1993; Halleux *et al.*, 1958). The planar structure allows the two-dimensional description adopted here, with a point group  $K_0 = 2mm$  for both sets of real and model positions defined by  $Y_0$  and  $X_0$ , respectively, in the notation of definition 3 given in §2, where in  $X_0$  and in  $Y_0$  the H atoms are omitted. We recall that  $X_0$  denotes the occupied part of the full model structure  $X$  and  $Y_0$  the corresponding real structure. The fit between  $X_0$  and  $Y_0$  considered above and based on common benzene ring centers leads to a maximal distance deviation  $\varepsilon = 0.033$  Å between corresponding C-atom positions (Fig. 1).

All model C positions are on points  $n_1, n_2$  of a two-dimensional hexagonal lattice  $\Lambda$  and this allows their indexing. In the present case, therefore, the superspace (carrier of the integral point-group representation) is two-dimensional and can be identified with the space of the molecular structure. In the notation of §2, one has:  $\Lambda = \Lambda_s = M$ .

Disregarding the hydrogen atoms, the molecular form of  $C_{54}$ , delimited by the set of four planes  $(\pm 10)$ ,  $(\pm 1\mp 1)$  is of the mid-edge type, as in Fig. 4 of (Janner, 1997). This implies scaling steps by a radial factor 1/2. In this particular case, the ratio between the largest rhombic form (forming the envelope) and the central smallest one (the hole) is 1/4. The intermediate form, following from a first scaling step, also has a structural meaning (Fig. 2). The scaled planes involved have a double indexing given by  $(10)[40]$ ,  $(10)[10]$ ,  $(10)[20]$ , respectively, together with the  $2mm$ -equivalent ones.

There are 15 atoms in the  $K_0$  asymmetric unit of  $C_{54}$ , with 12 C atoms in general positions and three C atoms on a mirror plane. Two non-trivial extensions of  $K_0$  by automorphisms of the hexagonal lattice  $\Lambda$  give rise to structural relations among these inequivalent positions:

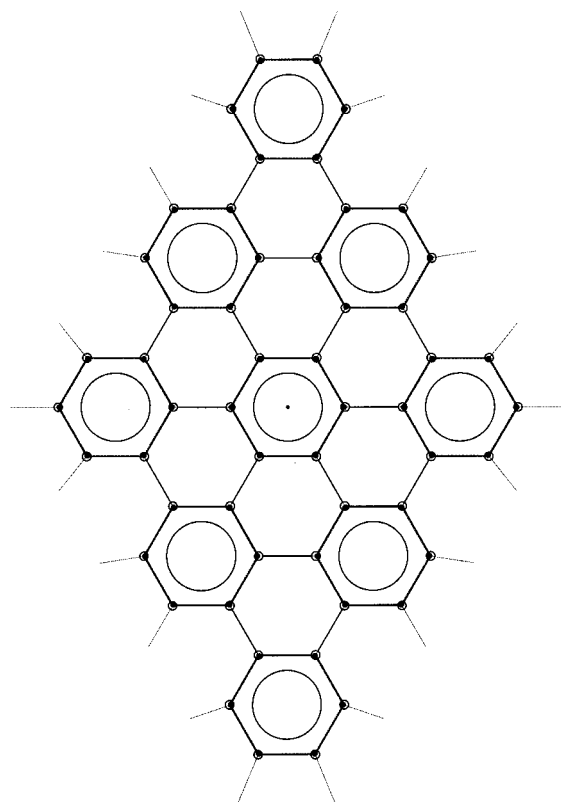
- (i) a sixfold rotation  $R_6$ ;
- (ii) a trace-four hyperbolic rotation  $L_4$ .

These additional transformations generate non-occupied equivalent positions only outside the rhombic enclosing form. The sixfold rotation is an expression of pseudosymmetry, owing to the central benzene ring. The hyperbolic rotation  $L_4$  takes properly into account the additional benzoic structure of the C atoms, which in the  $K_0$  asymmetric unit have highly correlated positions. The extended point group  $K$  generated accordingly,

$$K = \langle L_4, R_6, K_0 \rangle = \langle L_4, R_6, m_x \rangle = (4)6mm, \quad (23)$$

is a pseudosymmetry group for the indexed  $C_{54}$  molecule. There are four inequivalent C-atom positions in the  $K$  asymmetric unit, with orbits  $\mathcal{O}_{[10]}(K)$ ,  $\mathcal{O}_{[20]}(K)$ ,  $\mathcal{O}_{[40]}(K)$  and  $\mathcal{O}_{[46]}(K)$  having a molecular multiplicity of 38, 6, 6 and 4, respectively (Fig. 2). A molecular multiplicity indicates the number of orbit points within the molecular form.

The crystallographic character of this extended point group is ensured by the two-dimensional faithful integral representations  $\Gamma(K)$  and  $\Gamma^*(K)$ , obtained by expressing the element of  $K$  with respect to the hexagonal lattice bases  $a_1 = a_0(1, 0)$ ,  $a_2 = a_0(-1/2, 3^{1/2}/2)$  and  $a_1^* = a_0^{-1}(1, 1/3^{1/2})$ ,  $a_2^* = a_0^{-1}(0, 2/3^{1/2})$  with, in our case,  $a_0 = 1.433$  Å:



**Figure 1**

Structure of the planar polyphenoid  $C_{54}H_{22}$ , with C atoms indicated by filled dots, assuming 2.8 Å for the diameter of the benzene ring and 1.5 Å for the inter-ring single bond. In the model structure, indicated with open circles, all C—C distances have the same 1.433 Å value. This ensures that the centers of the benzene rings in the model and in the molecule coincide.

$$\Gamma(R_6) = R_6(a) = \begin{pmatrix} 1 & \bar{1} \\ 1 & 0 \end{pmatrix}, \quad \Gamma(L_4) = L_4(a) = \begin{pmatrix} 3 & 1 \\ 2 & 1 \end{pmatrix},$$

$$\Gamma(m_x) = m_x(a) = \begin{pmatrix} \bar{1} & 1 \\ 0 & 1 \end{pmatrix} \quad (24)$$

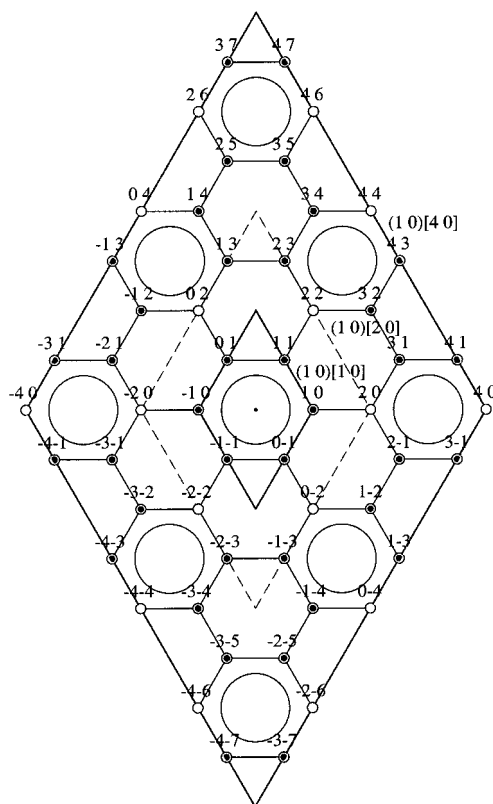
and correspondingly for  $\Gamma^*(K)$ :

$$\Gamma^*(R_6) = R_6(a^*) = \begin{pmatrix} 0 & \bar{1} \\ 1 & 1 \end{pmatrix}, \quad \Gamma^*(L_4) = L_4(a^*) = \begin{pmatrix} 3 & 2 \\ 1 & 1 \end{pmatrix},$$

$$\Gamma^*(m_x) = m_x(a^*) = \begin{pmatrix} \bar{1} & 0 \\ 1 & 1 \end{pmatrix}. \quad (25)$$

The relations between  $\Gamma$  and  $\Gamma^*$  given in (19) are indeed satisfied.

The pseudosymmetry point group  $(4)6mm$  is the same as for snow crystals (Janner, 1997) and for Z-DNA (Janner, 2000, 2001). However, to investigate the possible physical consequences of this extended group on a self-consistent molecular potential and on the electronic wave functions involved should be easier for  $C_{54}H_{22}$  than for Z-DNA or for snow crystals.



**Figure 2**  
Indexed C-atom positions of  $C_{54}H_{22}$  with the enclosing rhombic scaled form, delimited by planes, which are  $2mm$ -equivalent to  $(10)$  through  $[40]$ ,  $[20]$  and  $[10]$ , respectively. Two radial scaling steps are involved, leading from the external boundaries to intermediate ones and to what can be considered to be the internal hole.

## 4.2. Hexaphenylbenzene

The interest of this molecule is due to the combination of a two-dimensional architecture with a three-dimensional structure, having a propeller configuration of a type observed in other highly symmetrical hexasubstituted benzenes (Biemans, 1997). The molecular data have been taken from the X-ray diffraction crystal structure determination by J. C. Bart, omitting the H atoms (Bart, 1968). The approximate sixfold symmetry suggests a hexagonal point group for the model molecule should be adopted. This is, however, a fairly poor characterization. A better one follows from the 1999 paper of Berresheim *et al.* (1999), where it is observed that a number of polyphenylene molecules share the motifs one finds in cubic graphite (Gibson *et al.*, 1946). Indeed, in cubic graphite one finds propeller-shaped local atomic arrangements involving seven graphite rings: a central one, surrounded by six others with ring planes perpendicular to cubic body diagonals. To get a model for the chiral structure of hexaphenylbenzene, one has to decompose the local achiral motif mentioned above into two trigonal chiral components, and then to recombine elements of the same chirality, related by a  $60^\circ$  rotation around the hexagonal axis. The f.c.c. lattice of cubic graphite then allows an indexing of the atomic positions of each subsystem. An appropriate indexing for the whole is also possible, along the lines indicated in the previous section, but it requires a five-dimensional superspace. The essential elements of the present approach, however, already follow from a three-dimensional treatment, which is much more convenient. The higher-dimensional description will only be mentioned, very briefly, at the end.

Let us consider the first trigonal subsystem formed by the central hexagon and the ones attached to C(2), C(4) and C(6), respectively, and labeled by Bart as *B*, *D* and *F*. The second subsystem is then given by the same central ring together with the peripheral rings *A*, *C* and *E*. The two sets of indexed positions then follow by putting at  $[-101]$  the atoms C(1) and C(2), respectively. The corresponding model and real coordinates are indicated in Table 2(*a,b*), with respect to an orthonormal coordinate system. The numerical values indicated in brackets are obtained from those given by Bart by an overall shift of  $\Delta x = -2.397$  and  $\Delta y = 0.004$  (which puts the origin at the center of the molecule) followed by a  $-17.73^\circ$  rotation  $\rho_0$  around the hexagonal axis (corresponding to the  $[111]$  direction of the f.c.c. lattice) for the first subsystem, and by one of  $42.27^\circ$  for the second subsystem.

In Fig. 3, the model structure is compared graphically with the experimental one. The molecular scaled enclosing form is hexagonal and involves two successive radial scalings, by a factor of  $1/2$  each, as for  $C_{54}H_{22}$ , so that the central form is  $1/4$  of the external one, as indicated in Fig. 4. It is not evident that this enclosing form is physically relevant but it has to be considered in the general context of scaled enclosing forms. The two sets of index positions for the two trigonal subsystems of the (ideal) model structure are indicated in Fig. 5, together with one of the enclosing cubes. In Fig. 4, the other enclosing cube also appears in a view along the fourfold axis.

**Table 2**

Cubic indexed positions and comparison between model and real coordinates of a first and a second trigonal subsystem of hexaphenylbenzene (Bart, 1968).

*(a)* First trigonal subsystem

| Trigonal subsystem                                                                    | Atom           | Indices (basis <i>a</i> ) | Cartesian coordinates model (real structure) (Å) |                                    |                                    |
|---------------------------------------------------------------------------------------|----------------|---------------------------|--------------------------------------------------|------------------------------------|------------------------------------|
|                                                                                       |                |                           | <i>x</i> ( <i>x</i> <sub>0</sub> )               | <i>y</i> ( <i>y</i> <sub>0</sub> ) | <i>z</i> ( <i>z</i> <sub>0</sub> ) |
| Cubic basis parameters: <i>a</i> <sub>0</sub> = 1.0 Å, ρ <sub>0</sub> (111) = −17.73° |                |                           |                                                  |                                    |                                    |
| 0, <i>B</i> , <i>D</i> , <i>F</i>                                                     | C(1)           | 101                       | −1.000 (−1.000)                                  | 0.000 (−0.032)                     | 1.000 (0.981)                      |
|                                                                                       | C(2)           | 011                       | 0.000 (0.018)                                    | −1.000 (−0.997)                    | 1.000 (0.990)                      |
|                                                                                       | C(3)           | 110                       | 1.000 (0.997)                                    | −1.000 (−0.975)                    | 0.000 (−0.011)                     |
|                                                                                       | C(4)           | 101̄                      | 1.000 (0.972)                                    | 0.000 (−0.004)                     | −1.000 (−0.991)                    |
|                                                                                       | C(5)           | 011̄                      | 0.000 (0.004)                                    | 1.000 (1.005)                      | −1.000 (−0.999)                    |
|                                                                                       | C(6)           | 110̄                      | −1.000 (−0.991)                                  | 1.000 (1.002)                      | 0.000 (0.029)                      |
|                                                                                       | C(1 <i>B</i> ) | 022̄                      | 0.000 (0.025)                                    | −2.000 (−2.047)                    | 2.000 (2.041)                      |
|                                                                                       | C(2 <i>B</i> ) | 123̄                      | 1.000 (1.085)                                    | −2.000 (−2.001)                    | 3.000 (3.002)                      |
|                                                                                       | C(3 <i>B</i> ) | 134̄                      | 1.000 (1.034)                                    | −3.000 (−3.030)                    | 4.000 (4.047)                      |
|                                                                                       | C(4 <i>B</i> ) | 044̄                      | 0.000 (−0.014)                                   | −4.000 (−3.948)                    | 4.000 (4.070)                      |
|                                                                                       | C(5 <i>B</i> ) | 143̄                      | −1.000 (−1.023)                                  | −4.000 (−3.977)                    | 3.000 (3.078)                      |
|                                                                                       | C(6 <i>B</i> ) | 132̄                      | −1.000 (−0.992)                                  | −3.000 (−2.998)                    | 2.000 (2.085)                      |
|                                                                                       | C(1 <i>D</i> ) | 202̄                      | 2.000 (1.962)                                    | 0.000 (−0.104)                     | −2.000 (−2.155)                    |
|                                                                                       | C(2 <i>D</i> ) | 312̄                      | 3.000 (3.005)                                    | 1.000 (0.806)                      | −2.000 (−2.259)                    |
|                                                                                       | C(3 <i>D</i> ) | 413̄                      | 4.000 (3.959)                                    | 1.000 (0.673)                      | −3.000 (−3.281)                    |
|                                                                                       | C(4 <i>D</i> ) | 404̄                      | 4.000 (3.827)                                    | 0.000 (−0.406)                     | −4.000 (−4.164)                    |
|                                                                                       | C(5 <i>D</i> ) | 314̄                      | 3.000 (2.730)                                    | −1.000 (−1.319)                    | −4.000 (−4.092)                    |
|                                                                                       | C(6 <i>D</i> ) | 213̄                      | 2.000 (1.816)                                    | −1.000 (−1.154)                    | −3.000 (−3.065)                    |
|                                                                                       | C(1 <i>F</i> ) | 220̄                      | −2.000 (−1.975)                                  | 2.000 (2.110)                      | 0.000 (0.097)                      |
|                                                                                       | C(2 <i>F</i> ) | 231̄                      | −2.000 (−1.959)                                  | 3.000 (2.983)                      | 1.000 (1.185)                      |
|                                                                                       | C(3 <i>F</i> ) | 341̄                      | −3.000 (−2.856)                                  | 4.000 (4.046)                      | 1.000 (1.266)                      |
|                                                                                       | C(4 <i>F</i> ) | 440̄                      | −4.000 (−3.834)                                  | 4.000 (4.206)                      | 0.000 (0.243)                      |
|                                                                                       | C(5 <i>F</i> ) | 431̄                      | −4.000 (−3.877)                                  | 3.000 (3.296)                      | −1.000 (−0.827)                    |
|                                                                                       | C(6 <i>F</i> ) | 321̄                      | −3.000 (−2.935)                                  | 2.000 (2.272)                      | −1.000 (−0.910)                    |

*(b)* Second trigonal subsystem

| Trigonal subsystem                                                                   | Atom           | Indices (basis <i>a</i> ) | Cartesian coordinates model (real structure) (Å) |                                    |                                    |
|--------------------------------------------------------------------------------------|----------------|---------------------------|--------------------------------------------------|------------------------------------|------------------------------------|
|                                                                                      |                |                           | <i>x</i> ( <i>x</i> <sub>0</sub> )               | <i>y</i> ( <i>y</i> <sub>0</sub> ) | <i>z</i> ( <i>z</i> <sub>0</sub> ) |
| Cubic basis parameters: <i>a</i> <sub>0</sub> = 1.0 Å, ρ <sub>0</sub> (111) = 42.27° |                |                           |                                                  |                                    |                                    |
| 0, <i>A</i> , <i>C</i> , <i>E</i>                                                    | C(1)           | $\bar{1}10$               | −1.000 (−1.015)                                  | 1.000 (0.966)                      | 0.000 (−0.002)                     |
|                                                                                      | C(2)           | $\bar{1}01$               | −1.000 (−0.982)                                  | 0.000 (−0.011)                     | 1.000 (1.004)                      |
|                                                                                      | C(3)           | $0\bar{1}1$               | 0.000 (0.018)                                    | −1.000 (−0.989)                    | 1.000 (0.982)                      |
|                                                                                      | C(4)           | $1\bar{1}0$               | 1.000 (0.976)                                    | −1.000 (−0.988)                    | 0.000 (−0.011)                     |
|                                                                                      | C(5)           | $10\bar{1}$               | 1.000 (1.006)                                    | 0.000 (0.003)                      | −1.000 (−0.998)                    |
|                                                                                      | C(6)           | $01\bar{1}$               | 0.000 (−0.002)                                   | 1.000 (1.018)                      | −1.000 (−0.975)                    |
|                                                                                      | C(1 <i>A</i> ) | $\bar{2}20$               | −2.000 (−2.138)                                  | 2.000 (1.972)                      | 0.000 (−0.053)                     |
|                                                                                      | C(2 <i>A</i> ) | $\bar{2}31$               | −2.000 (−2.283)                                  | 3.000 (2.956)                      | 1.000 (0.896)                      |
|                                                                                      | C(3 <i>A</i> ) | $\bar{3}41$               | −3.000 (−3.364)                                  | 4.000 (3.863)                      | 1.000 (0.788)                      |
|                                                                                      | C(4 <i>A</i> ) | $\bar{4}40$               | −4.000 (−4.263)                                  | 4.000 (3.751)                      | 0.000 (−0.280)                     |
|                                                                                      | C(5 <i>A</i> ) | $\bar{4}31$               | −4.000 (−4.109)                                  | 3.000 (2.758)                      | −1.000 (−1.206)                    |
|                                                                                      | C(6 <i>A</i> ) | $\bar{3}21$               | −3.000 (−3.041)                                  | 2.000 (1.866)                      | −1.000 (−1.112)                    |
|                                                                                      | C(1 <i>C</i> ) | $0\bar{2}2$               | −0.000 (0.056)                                   | −2.000 (−2.056)                    | 2.000 (2.053)                      |
|                                                                                      | C(2 <i>C</i> ) | $1\bar{2}3$               | 1.000 (1.142)                                    | −2.000 (−2.118)                    | 3.000 (2.937)                      |
|                                                                                      | C(3 <i>C</i> ) | $1\bar{3}4$               | 1.000 (1.181)                                    | −3.000 (−3.082)                    | 4.000 (3.923)                      |
|                                                                                      | C(4 <i>C</i> ) | $0\bar{4}4$               | −0.000 (0.110)                                   | −4.000 (−3.980)                    | 4.000 (4.085)                      |
|                                                                                      | C(5 <i>C</i> ) | $1\bar{4}3$               | −1.000 (−0.986)                                  | −4.000 (−3.906)                    | 3.000 (3.215)                      |
|                                                                                      | C(6 <i>C</i> ) | $1\bar{3}2$               | −1.000 (−1.013)                                  | −3.000 (−2.945)                    | 2.000 (2.198)                      |
|                                                                                      | C(1 <i>E</i> ) | $20\bar{2}$               | 2.000 (2.050)                                    | 0.000 (0.039)                      | −2.000 (−2.037)                    |
|                                                                                      | C(2 <i>E</i> ) | $31\bar{2}$               | 3.000 (2.996)                                    | 1.000 (1.062)                      | −2.000 (−2.058)                    |
|                                                                                      | C(3 <i>E</i> ) | $41\bar{3}$               | 4.000 (3.959)                                    | 1.000 (1.121)                      | −3.000 (−3.049)                    |
|                                                                                      | C(4 <i>E</i> ) | $40\bar{4}$               | 4.000 (4.044)                                    | 0.000 (0.128)                      | −4.000 (−3.999)                    |
|                                                                                      | C(5 <i>E</i> ) | $31\bar{4}$               | 3.000 (3.112)                                    | −1.000 (−0.917)                    | −4.000 (−4.009)                    |
|                                                                                      | C(6 <i>E</i> ) | $21\bar{3}$               | 2.000 (2.109)                                    | −1.000 (−0.964)                    | −3.000 (−3.014)                    |

As already pointed out, the main aim of the new approach is the derivation of a pseudosymmetry group because it implies more structural relations than the Euclidean symmetry group. In the present case, we only show the existence of such a group by means of an extended cubic point group leaving the f.c.c.

lattice invariant. This group does the job, but is certainly too large for the hexaphenylbenzene molecule.

Let us start from the three-dimensional arithmetic group  $Gl(3, \mathbb{Z})$ , which is obtained by adding a parabolic generator  $P_y$  to those of the cubic point group  $O_h = m\bar{3}m$ , with elements

expressed with respect to the cubic lattice basis  $c = \{c_1, c_2, c_3\}$ . As indicated by Coxeter & Moser (1957, p. 92), one can choose as generators for  $Gl(3, \mathbb{Z})$  the matrices  $P = -m_{xy}(c)$ ,  $O = 2_x(c)$ ,  $Q = R_3(c)$  and  $U = P_y(c)$  with

$$\begin{aligned} m_{xy}(c) &= \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, & 2_x(c) &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \\ R_3(c) &= \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}, & P_y(c) &= \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \end{aligned} \quad (26)$$

With respect to the f.c.c. lattice basis  $f = \{f_1 = (0, \frac{1}{2}, \frac{1}{2}), f_2 = (\frac{1}{2}, 0, \frac{1}{2}), f_3 = (\frac{1}{2}, \frac{1}{2}, 0)\}$ , the transformations of  $m\bar{3}m$  are unimodular, but not the parabolic transformation  $P_y$ , only the squared one  $P_y^2(c)$ :

$$P_y(f) = \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{2} & -\frac{1}{2} \\ 0 & \frac{1}{2} & \frac{3}{2} \end{pmatrix}, \quad P_y^2(f) = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & -1 \\ 0 & 1 & 2 \end{pmatrix}. \quad (27)$$

This defines a three-dimensional integral representation  $\Gamma(K)$  with, in particular,  $\Gamma(P_y^2) = P_y^2(f)$ . One also has to check that the dual representation  $\Gamma^*(P_y^2) = P_y^2(f^*)$  is integral. Indeed, for  $f^* = \{f_1^* = (\bar{1}, 1, 1), f_2^* = (1, \bar{1}, 1), f_3^* = (1, 1, \bar{1})\}$ , one finds:

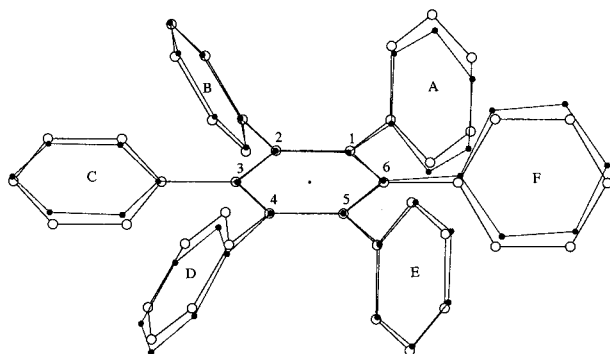
$$P_y^2(f^*) = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 1 \\ 1 & \bar{1} & 2 \end{pmatrix}. \quad (28)$$

Accordingly, the group

$$K = \langle m\bar{3}m, P_y^2 \rangle \quad (29)$$

is an extended crystallographic point group that leaves the f.c.c. lattice invariant.

One then verifies that the following atomic position of the trigonal asymmetric unit of the first subsystem of hexaphenylbenzene C(1), C(2), C(2B), C(3B), C(5B) and C(6B) all belong to the same orbit  $\mathcal{O}_{[\bar{1}01]}(K)$ , as indicated in Fig. 5. So, for example, one finds the following structural relations:

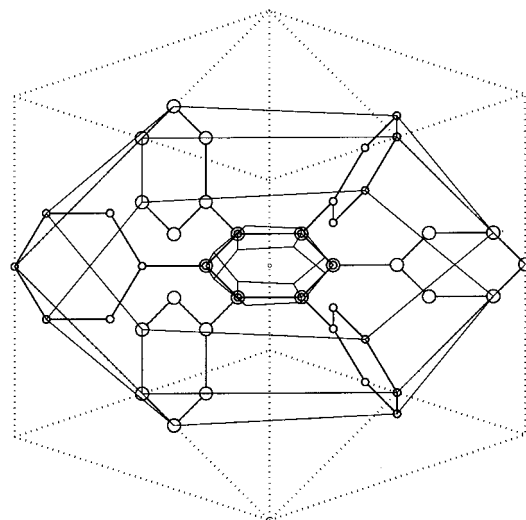


**Figure 3**  
Comparison between the molecular structure of hexaphenylbenzene, as observed in an orthorhombic crystal (black dots) and the corresponding ideal model based on cubic graphite motifs (open circles). The experimental twisting angle is about 65° and approximates the model one of 70°32' between the body-diagonal directions in a cube.

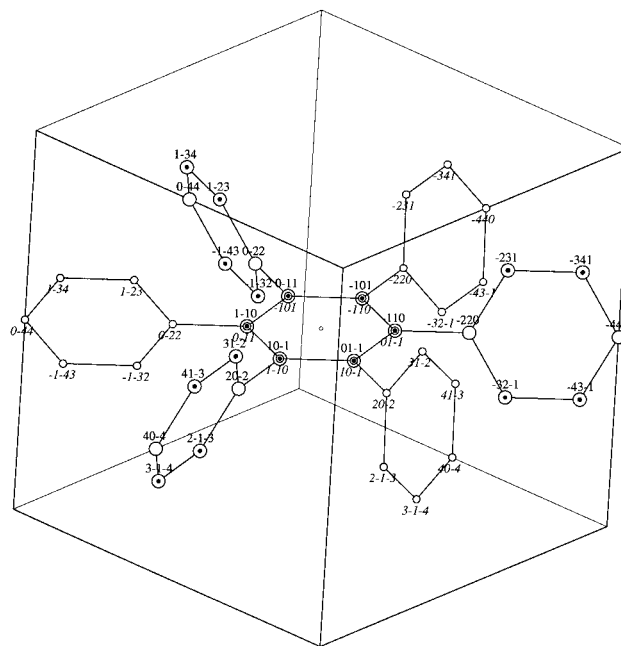
$$C(3B)[\bar{1}34] \xrightarrow{\alpha_1} C(6B)[\bar{1}33] \xrightarrow{\alpha_2} C(2)[0\bar{1}1] \xrightarrow{\alpha_3} C(1)[\bar{1}01] \quad (30)$$

with  $\alpha_1 = R_3^{-1}2_x m_{xy} P_y^2 m_{xy} R_3$ ,  $\alpha_2 = P_y^2 R_3 P_y^2 R_3$  and  $\alpha_3 = m_{xy}$ .

In the same way, C(1), C(2), C(2A), C(3A), C(5A) and C(6A) of the second trigonal subsystem, also belong to the same orbit, but now with  $[\bar{1}01]$  expressed with respect to the



**Figure 4**  
Hexagonal scaled enclosing form of hexaphenylbenzene. The internal shape (at the central ring) is scaled by a factor 1/4 from the external one formed by the peripheral rings. The cubes enclosing the two trigonal subsystems are indicated by dotted lines.



**Figure 5**  
The model C-atom positions of hexaphenylbenzene are indexed for each of the trigonal subsystems by corresponding f.c.c. lattice points (indicated by the two subsystems (implying a double indexing for the C atoms), the two sets of three peripheral rings are related by a 60° rotation around the hexagonal axis (and so also the two f.c.c. lattices around their common  $[111]$  direction).

60°-rotated basis  $c' = \{R_6c_1, R_6c_2, R_6c_3\}$ . This shows that  $K$  is a pseudosymmetry group for hexaphenylbenzene.

Finally, for taking into account the hexagonal symmetry of hexaphenylbenzene, one has to consider the three-dimensional  $Z$ -module  $M$  freely generated by the two cubic bases  $c$  and  $c'$ . Because of the integral linear relation  $c_1 + c_2 + c_3 = c'_1 + c'_2 + c'_3$ , the rank of this module is five and not six. There is then a five-dimensional representation space for the superspace point group  $K_s = \langle m\bar{3}m, P_y^2, R_6 \rangle$ , which is a pseudosymmetry group such that the atomic positions indicated above become equivalent.

This paper has been written taking into account the need expressed by the two referees of a paper on nucleic acids submitted to *J. Mol. Biol.* to have a formal conceptual frame. Stimulating discussions with Annalisa Fasolino are gratefully acknowledged. Thanks are expressed to R. de Gelder for his suggestion to look for examples in supramolecular chemistry and to J. van Hest for drawing the author's attention to the PhD work of H. A. M. Biemans, kindly made available by his thesis supervisor E. W. Meijer. The careful reading of the first manuscript by a referee and by R. de Gelder, together with their pertinent remarks and suggestions, have greatly helped the author to improve the presentation.

## References

- Ascher, E. & Janner, A. (1968). *Commun. Math. Phys.* **11**, 138–167.
- Bart, J. C. J. (1968). *Acta Cryst.* **B24**, 1277–1287.
- Berresheim, A. J., Müller, M. & Müllen, K. (1999). *Chem. Rev.* **99**, 1747–1785.
- Biemans, H. A. M. (1997). *Multi-Chromophoric Arrays in Molecules of High Symmetry*. PhD thesis, University of Eindhoven, The Netherlands.
- Coxeter, H. S. M. & Moser, W. O. J. (1957). *Generators and Relations for Discrete Groups*. Berlin: Springer.
- Dias, J. R. (1993). *J. Chem. Inf. Comput. Sci.* **33**, 117–127.
- Frank, F. C. (1965). *Acta Cryst.* **18**, 862–866.
- Friedel, G. (1911). *Leçons de Cristallographie*. Paris: Hermann.
- Gibson, J., Holohan, M. & Riley, H. L. (1946). *J. Chem. Soc.* pp. 456–461.
- Grosswald, E. (1966). *Topics from the Theory of Numbers*. New York: Macmillan.
- Hahn, T. (1992). Editor. *International Tables for Crystallography*, Vol. A. *Space-Group Symmetry*. Dordrecht/Boston/London: Kluwer Academic Publishers.
- Halleux, A., Martin, R. H. & King, G. S. D. (1958). *Helv. Chim. Acta*, **41**, 1177–1183.
- Janner, A. (1991). *Phys. Rev. B*, **43**, 13206–13214.
- Janner, A. (1995). *Acta Cryst.* **B51**, 386–401.
- Janner, A. (1997). *Acta Cryst.* **A53**, 615–631.
- Janner, A. (2000). *Cryst. Eng.* Indaba III Conference Special Edition.
- Janner, A. (2001). In preparation.
- Janner, A. & Janssen, T. (1980a). *Acta Cryst.* **A36**, 399–408.
- Janner, A. & Janssen, T. (1980b). *Acta Cryst.* **A36**, 408–415.
- Janot, C., Loreto, L., Farinato, R., Mancini, L., Baruchel, J. & Gastaldi, J. (1999). *Mater. Res. Soc. Proc.* **553**, 55–66.
- Janssen, T. (1986). *Acta Cryst.* **A42**, 261–271.
- Janssen, T., Janner, A., Looijenga-Vos, A. & de Wolff, P. M. (1999). *International Tables for Crystallography*, Vol. C. *Mathematical, Physical and Chemical Tables*, 2nd ed., edited by A. J. C. Wilson & E. Prince, pp. 899–937. Dordrecht/Boston/London: Kluwer Academic Publishers.
- Mancini, L., Janot, C., Loreto, L., Farinato, R., Gastaldi, J. & Baruchel, J. (1998). *Philos. Mag. Lett.* **78**, 159–167.
- Opechowski, W. (1986). *Crystallographic and Metacrystallographic Groups*. Amsterdam: North-Holland.
- Ostlund, S. & Wright, D. C. (1986). *Phys. Rev. Lett.* **56**, 2068–2071.
- Rossmann, M. G. & Arnold, E. (1993). *International Tables for Crystallography*, Vol. B. *Reciprocal Space*, edited by U. Shmueli, pp. 230–263. Dordrecht/Boston/London: Kluwer Academic Publishers.
- Shechtman, D., Blech, I., Gratias, D. & Cahn, J. W. (1984). *Phys. Rev. Lett.* **53**, 1951–1953.
- Viterbo, D. (1992). *Fundamentals of Crystallography*, edited by C. Giacovazzo, pp. 319–401. Oxford University Press.
- Yamamoto, A. (1996). *Acta Cryst.* **A52**, 509–560.