_wave_vector_seq_id** is a pointer to the description of the separate modulation wave vectors and must match one of the identifiers **_atom_site_Fourier_wave_vector_seq_id** listed separately in the ATOM_SITE_FOURIER_WAVE_VECTOR category. Likewise, the ***_atom_site_label** data item must match a value of **_atom_site_label** in the main list of atom positions. This is how the modulation is linked to the atom list. The item **_atom_site_displace_Fourier_id** is the formal key for the ATOM_SITE_DISPLACE_FOURIER category. It is used to locate the matching Fourier coefficients in the ATOM_SITE_DISPLACE_FOURIER_PARAM category. The coefficients may be reported in a sine–cosine (**_atom_site_displace_Fourier_param_sin**, ***_cos**) or modulus–argument (_mod**, ***_phase**) representation.

Where a group of atoms is treated as a rigid group, the categories above describe only the translational part of the positional distortion. ATOM_SITE_ROT_FOURIER and ATOM_SITE_ROT_FOURIER_PARAM are used to describe the rotational components.

ATOM_SITE_OCC_FOURIER and ATOM_SITE_U_FOURIER, and their associated *_PARAM categories, are the analogous categories for the modulation of site occupation and thermal parameters.

All the categories above describe the properties of individual atom sites. Larger-scale descriptions of the displacive modulation or of the rotational component of a rigid group are covered by the categories ATOM_SITES_DISPLACE_FOURIER and ATOM_SITES_ROT_FOURIER, each of which at present contains one descriptive data item.

The ATOM_SITES_MODULATION category contains data items describing the initial phases of the modulation waves, which are essential for determining the space group of the commensurate superstructure. More details are given in the dictionary.

3.4.3.3.3. *Special modulation functions*

Data items in these categories are as follows:

(*a*) ATOM_SITE_DISPLACE_SPECIAL_FUNC

```
• _atom_site_displace_special_func_atom_site_label
        → _atom_site_label
  _atom_site_displace_special_func_sawtooth_ax
  _atom_site_displace_special_func_sawtooth_ay
  _atom_site_displace_special_func_sawtooth_az
  _atom_site_displace_special_func_sawtooth_c
  _atom_site_displace_special_func_sawtooth_w
```

(*b*) ATOM_SITE_OCC_SPECIAL_FUNC

```
• _atom_site_occ_special_func_atom_site_label
        → _atom_site_label
  _atom_site_occ_special_func_crenel_c
  _atom_site_occ_special_func_crenel_w
```

The bullet (•) *indicates a category key. The arrow* (→) *is a reference to a parent data item.*

Several data items cover modulation functions that are not expressed as Fourier expansions. The examples in the current msCIF dictionary are restricted to the one-dimensional modulations (sawtooth displacive and occupational crenel functions) implemented in the program *JANA2000* (see Section 3.4.2).

3.4.3.3.4. *Molecular or packing geometry*

New data items in these categories are as follows:

(*a*) GEOM_ANGLE

```
_geom_angle_av
_geom_angle_max
_geom_angle_min
_geom_angle_site_ssg_symmetry_1
_geom_angle_site_ssg_symmetry_2
_geom_angle_site_ssg_symmetry_3
```

(*b*) GEOM_BOND

```
_geom_bond_distance_av
_geom_bond_distance_max
_geom_bond_distance_min
_geom_bond_site_ssg_symmetry_1
_geom_bond_site_ssg_symmetry_2
```

(*c*) GEOM_CONTACT

```
_geom_contact_distance_av
_geom_contact_distance_max
_geom_contact_distance_min
_geom_contact_site_ssg_symmetry_1
_geom_contact_site_ssg_symmetry_2
```

(*d*) GEOM_TORSION

```
_geom_torsion_av
_geom_torsion_max
_geom_torsion_min
_geom_torsion_site_ssg_symmetry_1
_geom_torsion_site_ssg_symmetry_2
_geom_torsion_site_ssg_symmetry_3
_geom_torsion_site_ssg_symmetry_4
```

For each of the geometry categories, there are two groups of extensions. One set covers maximum, minimum and average values of bonds, contact distances, angles and torsion angles. The other extends the symmetry-operation code used in geometry listings in the core CIF dictionary (see Section 3.2.4.3.2) to the higher-dimensional superspace form.

3.4.3.3.5. *Symmetry information*

New data items in these categories are as follows:

(*a*) SPACE_GROUP

```
_space_group_ssg_IT_number
_space_group_ssg_name
_space_group_ssg_name_IT
_space_group_ssg_name_WJJ
_space_group_ssg_WJJ_code
```

(*b*) SPACE_GROUP_SYMOP

```
_space_group_symop_ssg_id
_space_group_symop_ssg_operation_algebraic
```

At present, the msCIF dictionary extends the core CIF dictionary symmetry categories to describe superspace groups for one-dimensional modulated structures in four ways: as the superspace-group number in Janssen *et al.* (2004) (**_space_group_ssg_IT_number**), as the *International Tables* superspace-group symbol (***_ssg_name_IT**), as one of the notations from de Wolff *et al.* (1981) (***_ssg_name_WJJ**, ***_ssg_WJJ_code**), or in some other formalism (***_ssg_name**). At present, superspace-group names for higher dimensions can only be indicated using **_space_group_ssg_name**.

Symmetry operations in the superspace group are specified in the SPACE_GROUP_SYMOP category by an obvious extension to the method used in the core dictionary. These items must always be present in a CIF corresponding to a modulated or composite structure.

3.4.3.4. File metadata

The categories modified in the msCIF dictionary to formalize the construction of a multi-block description of modulated or composite structures are as follows:

AUDIT group

AUDIT ¶

AUDIT_LINK ¶

Categories marked with ¶ *are already defined in the core CIF dictionary.*

Data items revised in these categories are as follows:

(*a*) AUDIT
`_audit_block_code`
(*b*) AUDIT_LINK
`_audit_link_block_code`

The core dictionary definitions of these items are revised in order to formalize the relationships between multiple data blocks representing reference and modulated structures. Guidance is provided in the msCIF dictionary on how to label data blocks in a way that makes their mutual relationships clear.

3.4.4. Use of the msCIF dictionary

In this section, some of the capabilities of the dictionary will be demonstrated using simple examples. More detailed examples can be found at http://www.iucr.org/iucr-top/cif/ms and on the CD-ROM accompanying this volume.

3.4.4.1. Description of reciprocal space

Modulated and composite structures need more than three reciprocal vectors in order to index the whole set of reflections with integer numbers. Hence a diffraction vector is written as

$$\mathbf{H} = h\mathbf{a}^* + k\mathbf{b}^* + l\mathbf{c}^* + m_1\mathbf{q}_1 + \ldots + m_d\mathbf{q}_d, \qquad (3.4.4.1)$$

where the notation has been chosen according to the core CIF dictionary. In the case of a modulated structure, $\mathbf{a}^*$, $\mathbf{b}^*$ and $\mathbf{c}^*$ are the reciprocal vectors of the reference structure (and therefore h, k and l index the main reflections). $\mathbf{q}_1, \ldots, \mathbf{q}_d$ are the modulation wave vectors. They are three-dimensional vectors with some irrational component (if the modulated structure is incommensurate) in the lattice spanned by $\mathbf{a}^*$, $\mathbf{b}^*$ and $\mathbf{c}^*$. d is the dimension of the modulation. In the case of composite structures, the diffraction pattern can be indexed using $3 + d$ (arbitrarily selected) vectors $\mathbf{a}_k^*$ ($k = 1, \ldots, 3+d$). $\mathbf{a}_1^*$ ($\equiv \mathbf{a}^*$), $\mathbf{a}_2^*$ ($\equiv \mathbf{b}^*$) and $\mathbf{a}_3^*$ ($\equiv \mathbf{c}^*$) normally span the reciprocal lattice of the main reflections of one of the substructures (notice that this is only one particular, but highly intuitive, choice). The remaining d vectors with $k = 4, \ldots, d$ are the wave vectors of the modulation [$\mathbf{q}_1, \ldots, \mathbf{q}_d$ in equation (3.4.4.1)].

In a composite structure, the $(3 + d)$-dimensional reciprocal basis of the subsystem ν is determined by a $(3 + d) \times (3 + d)$ matrix W^ν [see van Smaalen (1995) and references therein]:

$$\mathbf{a}_i^{*\nu} = \sum_{k=1}^{3+d} W_{ik}^\nu \mathbf{a}_k^*, \quad i = 1, \ldots, 3 + d, \qquad (3.4.4.2)$$

where the subscripts $i = 1$, 2 and 3 label the reciprocal vectors $\mathbf{a}^{*\nu}$, $\mathbf{b}^{*\nu}$ and $\mathbf{c}^{*\nu}$, and $i = 4, \ldots, d$ label the wave vectors of the modulation expressed as linear combinations of $\mathbf{a}^{*\nu}$, $\mathbf{b}^{*\nu}$ and $\mathbf{c}^{*\nu}$.

The simplest case corresponds to a one-dimensional ($d = 1$) modulated structure. Consider for example the incommensurate phase of K_2SeO_4. The wave vector of the modulation can be chosen to be $\mathbf{q}_1 = \alpha\mathbf{a}^*$. Relevant information about the diffraction pattern of this compound is expressed using both the core CIF and msCIF dictionaries as shown in Example 3.4.4.1.

A more complicated example is the composite structure $(LaS)_{1.14}NbS_2$. The two mutually incommensurate subsystems (along the a axis) are (van Smaalen, 1991) NbS_2 ($\nu = 1$) and LaS ($\nu = 2$). The reciprocal basis can be chosen to be $\mathbf{a}_1^* = \mathbf{a}^{*1}$, $\mathbf{a}_2^* = \mathbf{b}^{*1}$, $\mathbf{a}_3^* = \mathbf{c}^{*1}$ and $\mathbf{a}_4^* = \mathbf{a}^{*2}$. For this particular choice, the two W matrices [see equation (3.4.4.2)] are

Example 3.4.4.1. *msCIF description of the diffraction pattern of a one-dimensional modulated structure.*

```
_exptl_crystal_type_of_structure         'mod'
_cell_reciprocal_basis_description
; a*,b*,c* (reciprocal basis spanning the lattice of
  main reflections), q modulation wave vector.
;

_diffrn_symmetry_description
; The whole diffraction pattern shows orthorhombic
  symmetry. The following extinction rules were
  detected:
     0kl0    k+l=odd
     h0lm    h+m=odd
     hk0m    m=odd
     h00m    h,m=odd
  Superspace group: P:Pnam:-1ss
;

_diffrn_reflns_satellite_order_max            1

_diffrn_reflns_theta_max                  40.14
_diffrn_reflns_theta_min                   3.32
_diffrn_reflns_limit_h_max                    8
_diffrn_reflns_limit_k_max                   18
_diffrn_reflns_limit_l_max                   10
_diffrn_reflns_limit_index_m_1_max            1
_diffrn_reflns_limit_h_min                    0
_diffrn_reflns_limit_k_min                    0
_diffrn_reflns_limit_l_min                    0
_diffrn_reflns_limit_index_m_1_min           -1

# Modulation wave vector
  loop_
     _cell_wave_vector_seq_id
     _cell_wave_vector_x
         1        0.318(5)
```

$$W^1 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad W^2 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

This information is transcribed to CIF format as shown in Example 3.4.4.2. (Note that the default values for the wave vector components and the elements of W are 0.)

3.4.4.2. Description of symmetry

The symmetry of a modulated or composite structure is described by a superspace group which leaves the $(3 + d)$-dimensional embedding of the structure invariant. Superspace is built of two orthogonal subspaces and both of them are kept invariant separately by the superspace symmetry operations. (In reciprocal space this means that, for these structures, main reflections and satellite reflections are never transformed into one another by superspace symmetry operations.) Consequently, superspace groups are not general $(3 + d)$-dimensional space groups. The standard notation for superspace groups only covers the one-dimensional superspace groups, which are listed in Janssen *et al.* (2004). As a consequence, msCIFs must include a list of all the symmetry operations in an (x, y, z) format (using as symbols $x_1 \ldots x_{3+d}$) similar to that used in the core CIF dictionary. Superspace-group names for one-dimensional structures can be expressed either according to Janssen *et al.* (2004) or according to the original notation of de Wolff *et al.* (1981). Alternative names or higher-dimensional superspace groups can also be included, but not parsed.

references