3.5. Classification and use of electron density data

BY P. R. MALLINSON AND I. D. BROWN

3.5.1. Introduction

Modern high-resolution X-ray diffraction methods enable the electron density distributions in many crystals to be accurately determined, allowing one to explore the interactions that hold crystals and molecules together (Coppens, 1997; Koritsanszky & Coppens, 2001). For example, Bader's (1990) theory of atoms in molecules gives a description of chemical bonding in terms of the topology of the electron density, but it requires an accurate knowledge of the electron density. [Note that the terms 'electron density' (or distribution) and 'charge density' tend to be used interchangeably in the literature. The term 'electron density' is used here since 'charge density' strictly includes both positive (nuclear) and negative (electronic) charge. Properties such as electrostatic potential, electric fields and their gradients, and molecular electric moments are derived from the charge density distribution, which includes the contribution of the nuclear charge (Spackman & Brown, 1994).] The advent of area detectors, more powerful laboratory computing facilities and developments in software have combined to produce a rapid increase in the number of accurate electron density determinations now available. While such studies are not yet routine, they are sufficiently common to justify defining a structure for a secure and systematic archive. This in turn requires a method of representing the electron density acceptable to the crystallographic community.

The CIF standard which has been adopted by the IUCr and is in common use in macromolecular and small-molecule crystallography is the obvious choice for such a file structure. At the Sagamore XIII conference held at Stare Jablonki, Poland, in 2000, the IUCr Commission on Charge, Spin and Momentum Densities decided to prepare a CIF dictionary for reporting electron densities in crystals. The result is the rhoCIF dictionary described in this chapter. It was prepared in conjunction with members of the Commission and other members of the community. [The name rhoCIF for this dictionary was chosen over the more obvious choice of edCIF to avoid any possible confusion with electron diffraction. The symbol $\rho(\mathbf{r})$ conventionally represents the electron density in real space at the point defined by the position vector $\mathbf{r}$.]

The distribution of electron density in a molecule or crystal can also be calculated theoretically, and the results are usually compared with experiment. The Hohenberg–Kohn theorem (Hohenberg & Kohn, 1964) shows that the ground-state energy is a unique functional of the electron density, thus providing a theoretical justification for relating chemical and physical properties directly to the observable electron density. Density functional theory has been applied to chemical problems for some time (see, for example, Labanowski & Andzelm, 1991; Ziegler, 1991; Kohn *et al.*, 1996; Nagy, 1998). Although the current version of the rhoCIF dictionary is intended to record the results of experimental measurements, it can also be used to represent a theoretical electron density that has been fitted with an appropriate model.

3.5.2. Dictionary design considerations

The rhoCIF dictionary contains the definitions of data items used to report accurate electron densities in crystals. The structure of these definitions conforms to the CIF rules and conventions, and selected data items reflect those currently needed in the field. In any field of science, the nature of the data used to describe the experiment and the results changes with time, and accurate electron density studies are no exception. The items defined in this dictionary describe the density representation approach used most often today – the atom-centred multipole expansion model introduced by Stewart (1973) and refined by Hansen & Coppens (1978). In this model, the electron density is expressed as the sum of three components. Two are the partial contributions of the core and valence shell of the spherical free atom, usually taken from the paper by Clementi & Roetti (1974). The third component is an aspherical multipole expansion of the remaining electron density. This is an atom-centred function, flexible enough to describe most electron density distributions likely to be encountered in a crystal, and with components that have the same symmetries as the atom-centred wave functions (orbitals) frequently used by chemists when describing chemical bonding.

3.5.3. Classification of data definitions

The rhoCIF dictionary is intended to be used in combination with the core CIF dictionary, or, after adaptation to DDL2, the mmCIF dictionary. The two dictionaries can be combined to form a composite dictionary by concatenation. Alternatively, the merging procedures described in Section 3.1.9 can be used to create a virtual dictionary at run time. Because the dictionary is not self-contained, its structure can be kept very simple, comprising only two categories: one to define a local axis system around each atom and one to contain the various parameters of the Hansen & Coppens (1978) multipole function.

The two categories in the electron density CIF dictionary lie within the single ATOM category group and relate to the structural model. They are as follows:

Specification of local axes at each atom (§3.5.3.1)
   ATOM_LOCAL_AXES
Atom-centred multipole expansion functions (§3.5.3.2)
   ATOM_RHO_MULTIPOLE

3.5.3.1. Specification of local axes at each atom

Data items in this category are as follows:

ATOM_LOCAL_AXES
- **_atom_local_axes_atom_label**
      → **_atom_site_label**
  _atom_local_axes_atom0
  _atom_local_axes_atom1
  _atom_local_axes_atom2
  _atom_local_axes_ax1
  _atom_local_axes_ax2

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

Affiliations: PAUL R. MALLINSON, Department of Chemistry, University of Glasgow, Glasgow G12 8QQ, Scotland; I. DAVID BROWN, Brockhouse Institute for Materials Research, McMaster University, Hamilton, Ontario, Canada L8S 4M1.

Example 3.5.3.1. *Definition of the local Cartesian axis system for the atoms in a proton sponge complex.*

```
loop_
_atom_local_axes_atom_label
_atom_local_axes_atom0
_atom_local_axes_ax1
_atom_local_axes_atom1
_atom_local_axes_atom2
_atom_local_axes_ax2
N(1)      C(1)      X    N(1)      N(2)      Y
N(2)      C(8)      X    N(2)      N(1)      Y
C(1)      N(1)      X    C(1)      C(2)      Y
C(2)      C(1)      X    C(2)      C(3)      Y
C(3)      C(2)      X    C(3)      C(4)      Y
C(4)      C(10)     X    C(4)      C(3)      Y
C(5)      C(10)     X    C(5)      C(6)      Y
C(6)      C(5)      X    C(6)      C(7)      Y
C(7)      C(8)      X    C(7)      C(6)      Y
C(8)      C(9)      X    C(8)      N(2)      Y
C(9)      C(10)     X    C(9)      C(1)      Y
C(10)     C(9)      X    C(10)     C(5)      Y
C(11)     N(1)      X    C(11)     H(111)    Y
C(14)     N(2)      X    C(14)     H(141)    Y
CL(1A)    C(2A)     X    CL(1A)    CL(2A)    Y
CL(2A)    C(3A)     X    CL(2A)    CL(1A)    Y
O(1A)     C(1A)     X    O(1A)     H(100)    Y
O(2A)     C(1A)     X    O(2A)     CL(1A)    Y
O(3A)     C(4A)     X    O(3A)     CL(2A)    Y
O(4A)     C(4A)     X    O(4A)     H(100)    Y
C(1A)     O(1A)     X    C(1A)     O(2A)     Y
C(2A)     C(1A)     X    C(2A)     CL(1A)    Y
C(3A)     C(2A)     X    C(3A)     CL(2A)    Y
C(4A)     O(3A)     X    C(4A)     O(4A)     Y
H(1NN)    N(1)      X    H(1NN)    N(2)      Y
H(100)    O(1A)     X    H(100)    O(4A)     Y
H(111)    C(11)     X    H(111)    H(112)    Y
H(112)    C(11)     X    H(112)    H(111)    Y
H(113)    C(11)     X    H(113)    H(111)    Y
H(141)    C(14)     X    H(141)    H(142)    Y
H(142)    C(14)     X    H(142)    H(141)    Y
H(143)    C(14)     X    H(143)    H(141)    Y
H(2)      C(2)      X    H(2)      H(3)      Y
H(3)      C(3)      X    H(3)      H(2)      Y
H(4)      C(4)      X    H(4)      H(3)      Y
H(5)      C(5)      X    H(5)      H(4)      Y
H(6)      C(6)      X    H(6)      H(7)      Y
H(7)      C(7)      X    H(7)      H(6)      Y
```

The ATOM_LOCAL_AXES category is used to define a local Cartesian axis system for each atom. Data items in this category are looped, with the category key (`_atom_local_axes_atom_label`) being a pointer to the `_atom_site_label` defined in the ATOM_SITE category of the core CIF dictionary. The orientation of the Cartesian basis is defined in terms of vectors between atoms, and any arbitrary orientation can be described by including dummy atoms if necessary. The spherical coordinate system used in the description of the multipoles is defined in terms of the local Cartesian basis: z is the polar axis from which the angle θ is measured, φ is measured from the x axis in the xy plane with the y axis having a value of $\varphi = +90°$.

Example 3.5.3.1 shows how the category is used in part of an electron density CIF relating to a charge density study of the complex of 1,8-bis(dimethylamino)naphthalene with 1,2-dichloromaleic acid (a 'proton sponge') (Mallinson *et al.*, 1997).

3.5.3.2. Atom-centred multipole expansion functions

Data items in this category are as follows:

ATOM_RHO_MULTIPOLE
- `_atom_rho_multipole_atom_label`
  → `_atom_site_label`
  `_atom_rho_multipole_coeff_Pc`
  `_atom_rho_multipole_coeff_Pv`
  `_atom_rho_multipole_coeff_P00`

```
_atom_rho_multipole_coeff_P1-1
_atom_rho_multipole_coeff_P10
_atom_rho_multipole_coeff_P11
_atom_rho_multipole_coeff_P2-2
_atom_rho_multipole_coeff_P2-1
_atom_rho_multipole_coeff_P20
_atom_rho_multipole_coeff_P21
_atom_rho_multipole_coeff_P22
_atom_rho_multipole_coeff_P3-3
_atom_rho_multipole_coeff_P3-2
_atom_rho_multipole_coeff_P3-1
_atom_rho_multipole_coeff_P30
_atom_rho_multipole_coeff_P31
_atom_rho_multipole_coeff_P32
_atom_rho_multipole_coeff_P33
_atom_rho_multipole_coeff_P4-4
_atom_rho_multipole_coeff_P4-3
_atom_rho_multipole_coeff_P4-2
_atom_rho_multipole_coeff_P4-1
_atom_rho_multipole_coeff_P40
_atom_rho_multipole_coeff_P41
_atom_rho_multipole_coeff_P42
_atom_rho_multipole_coeff_P43
_atom_rho_multipole_coeff_P44
_atom_rho_multipole_configuration
_atom_rho_multipole_core_source
_atom_rho_multipole_kappa
_atom_rho_multipole_kappa_prime0
_atom_rho_multipole_kappa_prime1
_atom_rho_multipole_kappa_prime2
_atom_rho_multipole_kappa_prime3
_atom_rho_multipole_kappa_prime4
_atom_rho_multipole_radial_function_type
_atom_rho_multipole_radial_slater_n0
_atom_rho_multipole_radial_slater_n1
_atom_rho_multipole_radial_slater_n2
_atom_rho_multipole_radial_slater_n3
_atom_rho_multipole_radial_slater_n4
_atom_rho_multipole_radial_slater_zeta0
_atom_rho_multipole_radial_slater_zeta1
_atom_rho_multipole_radial_slater_zeta2
_atom_rho_multipole_radial_slater_zeta3
_atom_rho_multipole_radial_slater_zeta4
_atom_rho_multipole_scat_core
_atom_rho_multipole_scat_valence
_atom_rho_multipole_valence_source
```

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

Data items in this category are also looped. The category key (`_atom_rho_multipole_atom_label`) is also a child of `_atom_site_label` in the ATOM_SITE category, thereby linking the electron density multipole coefficients to a specific atom site in the list of atom coordinates. This category contains all the parameters of the Hansen & Coppens (1978) function

$$\rho(\mathbf{r}) = P_c \rho_{\text{core}}(\mathbf{r}) + P_v \kappa^3 \rho_{\text{valence}}(\kappa r) \\ + \sum_{0 \le l \le l_{\max}} \{\kappa'(l)^3 R(\kappa'(l), l, \mathbf{r})\} \sum_{-l \le m \le l} \{P(l,m)d(l,m,\theta,\varphi)\},$$

where $\mathbf{r}$ is a position vector with respect to the atomic nucleus, and $\rho_{\text{core}}(\mathbf{r})$ and $\rho_{\text{valence}}(\kappa r)$ are the spherical core and valence electron densities, respectively. They are obtained from atomic orbital analytic wavefunctions such as those tabulated by Clementi & Roetti (1974). They are also the Fourier transforms of the X-ray scattering factors given in `_atom_rho_multipole_scat_core` and `_atom_rho_multipole_scat_valence`.

P_c is the weight applied to the free-atom core (`_atom_rho_multipole_coeff_Pc`), P_v is the weight applied to the free-atom valence shell (`_atom_rho_multipole_coeff_Pv`), $P(0,0)$ is the number of remaining electrons (`_atom_rho_multipole_coeff_P00`) and $P_c + P_v + P(0,0) = Z$ (the atomic number) for

Example 3.5.3.2. *Multipole expansion for an atom in the proton sponge complex of Example 3.5.3.1.*

```
loop_
_atom_rho_multipole_atom_label
_atom_rho_multipole_coeff_Pv
_atom_rho_multipole_coeff_P00
_atom_rho_multipole_coeff_P11
_atom_rho_multipole_coeff_P1-1
_atom_rho_multipole_coeff_P10
_atom_rho_multipole_coeff_P20
_atom_rho_multipole_coeff_P21
_atom_rho_multipole_coeff_P2-1
_atom_rho_multipole_coeff_P22
_atom_rho_multipole_coeff_P2-2
_atom_rho_multipole_coeff_P30
_atom_rho_multipole_coeff_P31
_atom_rho_multipole_coeff_P3-1
_atom_rho_multipole_coeff_P32
_atom_rho_multipole_coeff_P3-2
_atom_rho_multipole_coeff_P33
_atom_rho_multipole_coeff_P3-3
_atom_rho_multipole_coeff_P40
_atom_rho_multipole_coeff_P41
_atom_rho_multipole_coeff_P4-1
_atom_rho_multipole_coeff_P42
_atom_rho_multipole_coeff_P4-2
_atom_rho_multipole_coeff_P43
_atom_rho_multipole_coeff_P4-3
_atom_rho_multipole_coeff_P44
_atom_rho_multipole_coeff_P4-4
_atom_rho_multipole_kappa
_atom_rho_multipole_kappa_prime0
_atom_rho_multipole_kappa_prime1
_atom_rho_multipole_kappa_prime2
_atom_rho_multipole_kappa_prime3
_atom_rho_multipole_kappa_prime4
_atom_rho_multipole_configuration
_atom_rho_multipole_radial_slater_n0
_atom_rho_multipole_radial_slater_zeta0
_atom_rho_multipole_radial_slater_n1
_atom_rho_multipole_radial_slater_zeta1
_atom_rho_multipole_radial_slater_n2
_atom_rho_multipole_radial_slater_zeta2
_atom_rho_multipole_radial_slater_n3
_atom_rho_multipole_radial_slater_zeta3
_atom_rho_multipole_radial_slater_n4
_atom_rho_multipole_radial_slater_zeta4
_atom_rho_multipole_core_source
_atom_rho_multipole_valence_source
N(1)    2.63(5)
  0.00       -0.037(17)  0.062(14)  0.00      -0.084(18)
  0.00        0.00      -0.027(15) -0.048(13)  0.00
 -0.098(16) -0.063(14)  0.00        0.00       0.082(14)
 -0.037(14)  0.00        0.00        0.00       0.00
  0.00        0.00        0.00        0.00       0.00
0.992(8)      0.80(4)     0.80       0.80      0.80     0.80
;
1S 2S 3S 4S 2P 3P 4P 3D 4D 4F 5S 5P 6S 6P 5D 7S 6D 5F
 2 -2  0  0 -3  0  0  0  0  0  0  0  0  0  0  0  0  0
;
   2   7.2553  2    7.2553  2    7.2553  3    7.2553
   4   7.2553
'Clementi & Roetti, 1974'
'Clementi & Roetti, 1974'

# Data for other atoms have been omitted for brevity
```

$d(l, m, \theta, \varphi)$ is the spherical harmonic of order l, m at the angle (θ, φ). The summations are performed over the index ranges $0 \leq l \leq l_{max}$, $-l \leq m \leq l$, where l_{max} is the highest order of multipole applied.

Example 3.5.3.2 demonstrates how the category is used in the proton sponge complex of Example 3.5.3.1. Only the first atom is shown in the example.

3.5.4. Development of the dictionary and supporting software

The first implementation of the dictionary appeared in a version of the *XD* program package (Koritsanszky *et al.*, 2003). Thus *XD* can read and write CIFs which include data items in the two categories described above (as well as, of course, items in the core CIF dictionary). It is envisaged that future developments to the rhoCIF dictionary will add features that are relevant to other widely used program packages. In the *XD* implementation, use was made of the *CIFtbx* library of Fortran functions for programming CIF applications (Hall & Bernstein, 1996).

References

Bader, R. F. W. (1990). *Atoms in molecules: a quantum theory.* Oxford University Press.

Clementi, E. & Roetti, C. (1974). *Roothan–Hartree–Fock atomic wavefunctions. Basis functions and their coefficients for ground and certain excited states of neutral and ionized atoms. At. Data Nucl. Data Tables,* **14**, 177–478.

Coppens, P. (1997). *X-ray charge densities and chemical bonding.* Oxford University Press.

Hall, S. R. & Bernstein, H. J. (1996). *CIF applications. V. CIFtbx2: extended tool box for manipulating CIFs. J. Appl. Cryst.* **29**, 598–603.

Hansen, N. K. & Coppens, P. (1978). *Testing aspherical atom refinements on small-molecule data sets. Acta Cryst.* A**34**, 909–921.

Hohenberg, P. & Kohn, W. (1964). *Inhomogeneous electron gas. Phys. Rev. (Second Series),* **136**, B864–B871.

Kohn, W., Becke, A. D. & Parr, R. G. (1996). *Density functional theory of electronic structure. J. Phys. Chem.* **100**, 12974–12980.

Koritsanszky, T., Richter, T., Macchi, P., Volkov, A., Gatti, C., Howard, S., Mallinson, P. R., Farrugia, L., Su, Z. & Hansen, N. K. (2003). *XD. Computer program package for multipole refinement and topological analysis of electron densities from diffraction data.* http://xd.chem.buffalo.edu.

Koritsanszky, T. S. & Coppens, P. (2001). *Chemical applications of X-ray charge density analysis. Chem. Rev.* **101**, 1583–1627.

Labanowski, J. K. & Andzelm, J. W. (1991). *Density-functional methods in chemistry.* New York: Springer-Verlag.

Mallinson, P. R., Wozniak, K., Smith, G. T. & McCormack, K. L. (1997). *A charge density analysis of cationic and anionic hydrogen bonds in a 'proton sponge' complex. J. Am. Chem. Soc.* **119**, 11502–11509.

Nagy, A. (1998). *Density functional. Theory and application to atoms and molecules. Phys. Rep.* **298**, 2–79.

Spackman, M. A. & Brown, A. S. (1994). *Charge densities from X-ray diffraction data. R. Soc. Chem. Ann. Rep. Sect. C,* pp. 175–212.

Stewart, R. F. (1973). *Electron population analysis with generalized X-ray scattering factors: higher multipoles. J. Chem. Phys.* **58**, 1668–1676.

Ziegler, T. (1991). *Approximate density functional theory as a practical tool in molecular energetics and dynamics. Chem. Rev.* **91**, 651–667.

a neutral atom. κ is the valence electron expansion factor (`_atom_rho_multipole_kappa`), $R(\kappa'(l), l, \mathbf{r})$ is the radial function (Slater or equivalent) (`_atom_rho_multipole_radial_*`), $\kappa'(l)$ is the multipole function expansion factor (`_atom_rho_multipole_kappa_prime[l]`), $P(l, m)$ are the spherical harmonic coefficients (`_atom_rho_multiple_coeff_P[lm]`) and