

Chapter 11

XD Utility programs

11.1 XDVIB1 - A Program to Calculate the Mean-Square Displacement Amplitudes from Harmonic Vibrational Frequencies and Normal Modes.

11.1.1 Overview

The XDVIB programs described in this chapter were developed to improve the significance of the anisotropic displacement parameters (adp's) obtained by LS refinement against X-ray data. This can be done by incorporating independent information based on spectroscopic and/or theoretical calculations into the refinement in terms of constraints or restraints.

A full description of nuclear motion in molecular crystals, within the mean-field approximation, is an M -parameter problem, where M is the number of elements of the symmetric mean square displacement amplitude matrix to be derived ($M=3N(3N+1)/2$, N being the number of nuclei in the molecule). This is a hopeless task, given the fact that even in an optimal case, only $6N$ adp's are available from a diffraction experiment. Consequently, further approximations and independent data are needed. A feasible approximation is to neglect the coupling between relative motion of the nuclei (internal modes) and the overall motion of the molecule in the lattice (external modes, 3 translation and 3 rotation) yielding a reduction in the number of free parameters by $6n$, where n is the internal degree of freedom ($n=3N-6$ for nonlinear and $3N-5$ for linear molecules). The knowledge of harmonic frequencies alone would further reduce the number of unknowns by N . A complete knowledge of the internal vibrational modes, including the frequencies, [$n(n+1)/2$ elements of the corresponding MSDA matrix] leads to $m=21$ ($6 \times 7/2$) parameters to be derived ($m < 21$ from molecules composed of $N < 8$ atoms). However, to estimate $6N$ adp's (the diagonal, symmetric blocks of the total MSDA matrix) only $6N-m$ independent parameters, associated with internal modes are needed. But how to choose these remaining parameters?

When neglecting internal-external coupling, each adp can be given as a sum of two terms:

$$\mathbf{U}_a = \mathbf{V}_a + \mathbf{W}_a$$

where \mathbf{V}_a and \mathbf{W}_a are the MSDA tensors (adp's) of atom **a** corresponding to the internal and external modes, respectively. The latter term can be adequately accounted for in terms of the m parameters of the **T**, **L** and **S** tensors of the rigid-body model [1]:

$$\mathbf{W}_a = \mathbf{R}'_a \mathbf{L} \mathbf{R}_a + \mathbf{R}'_a \mathbf{S} + \mathbf{S}' \mathbf{R}_a + \mathbf{T}$$

where \mathbf{R}_a is an antisymmetric tensor representing the vector product

$$\mathbf{R}_a = \begin{pmatrix} 0 - y_a z_a \\ y_a 0 - x_a z_a \\ -z_a x_a 0 \end{pmatrix}$$

with x_a , y_a and z_a being the Cartesian coordinates of nucleus **a** situated at position \mathbf{r}_a . Since \mathbf{V}_a and \mathbf{W}_a are additive, they cannot be obtained independently via an LS refinement, but their different behaviour upon a special transformation can be used as extra

information. It can be easily shown that the difference MSDA corresponding to \mathbf{W} (the difference of the components of \mathbf{W}_a and \mathbf{W}_b along the internuclear vector $\mathbf{d}_{ab} = \mathbf{r}_b - \mathbf{r}_a$) vanishes for all **a-b** links:

$$\Delta(\mathbf{W})_{ab} = \mathbf{d}_{ab}' (\mathbf{W}_a - \mathbf{W}_b) \mathbf{d}_{ab} = 0$$

but does not necessarily for \mathbf{V} . This property of the internal ADPs can be imposed as a restriction in the LS fit. Let us suppose we know \mathbf{V}_a for all atoms, start the refinement with these values and constrain the shifts in the ADPs according to the above requirement for a certain number of independent links:

$$\Delta(\delta\mathbf{U})_{ab} = 0$$

After the k -th cycle we obtain

$$\mathbf{U}_a^{(k)} = \mathbf{V}_a + \delta\mathbf{U}_a^{(k)}, \quad \mathbf{U}_b^{(k)} = \mathbf{V}_b + \delta\mathbf{U}_b^{(k)} \quad \text{with} \quad \Delta(\mathbf{U}^{(k)})_{ab} = \Delta(\mathbf{V})_{ab}$$

Thus, the information on the intermolecular vibrations are preserved during the refinement. There are $l=N(N-1)/2$ links. Singular value decomposition of the matrix of constraints eliminates linear dependencies, yielding, in general, $6N-m$ independent constraints.

XDVIB1 calculates MSDAs (the total $3N \times 3N$ \mathbf{V} matrix) from n frequencies and normal modes obtained by the *Gaussian98/03* program packages. Details of the vibrational analysis implemented in *Gaussian98/03* are described in reference [2]. Users should note that it is necessary to run the *Gaussian98/03* frequency analysis with the IOP(7/33=1) option to obtain the normal coordinates in a form readable by **XDVIB1**. It is also imperative that the "opt" option is not used in the Gaussian job. There is an important point to make concerning these calculations. Harmonic vibrational analyses based on theoretical force fields are valid only if the gradient of the electronic energy with respect of nuclear coordinates vanishes (equilibrium geometry) and the force-constant matrix (the Hessian of the energy or the second derivatives) is taken at the equilibrium configuration. Thus the molecular geometry used in the vibrational analysis must be optimized at the same level of theory, and utilizing the same basis set that the derivation of the frequencies is based on.

The *Gaussian98/03* programs work in terms of mass-weighted generalized coordinates to separate the internal (vibration) and external (translation and rotation) motion. The internal coordinates \mathbf{v} (n -vector) are generated to be orthogonal to the external ones, \mathbf{w} (6-vector related to the Eckart coordinates). The output displays the wavenumber (ν_k), the reduced mass (μ_k) and the normalized Cartesian displacement vector for each normal mode (\mathbf{q}_k).

Important The six external modes calculated by Gaussian are often far from zero, even if the opt=VeryTight option is used. This seems to result in normal modes which are not orthogonal to each other. As a consequence, the MSDA matrix calculated by XDVIB may have some 'contamination' from external (artefact) modes (especially if default opt is used). This problem is under investigation.

The MSDA matrix is diagonal in terms of normal coordinates [3]

$$\langle \mathbf{q} \mathbf{q}' \rangle = \text{diag}(\delta_1, \delta_2, \delta_3, \dots, \delta_n)$$

where

$$\delta_k = \frac{h}{8\pi^2 c \nu_k} \coth\left(\frac{h c \nu_k}{2kT}\right)$$

while in Cartesian representation

$$\langle \mathbf{v} \mathbf{v}' \rangle = \mathbf{V} = \mathbf{L} \langle \mathbf{q} \mathbf{q}' \rangle \mathbf{L}' \quad , \quad \mathbf{q} = \mathbf{L} \mathbf{u}$$

11.1.2 Files used and created by XDVIB1

Input: `xd.mas`, *Gaussian98/03* output file
 Output: `xd_vib1.par`, `xd_vib1_sig.dat` `xd_vib1.out`

11.1.3 Input instructions for XDVIB1**11.1.3.1 SELECT**

SELECT temp 100.0 **scale** 1.0 **(*)nlinear linear**

temp *temperature*

The temperature (K) maintained during the data collection.

scale *scalefactor*

A number used to scale the calculated frequencies in wavenumbers [cm^{-1}], printed in the *Gaussian* output file, `gaussian.out`. Frequencies calculated at different level of theory and basis set are scaled by empirical factors to eliminate known systematic errors. Selected values taken from reference [2] are listed below. If experimental normal frequencies (corrected for anharmonicity) are available, one can scale the calculated ones directly to those observed.

Method/basis	Scale factor	Method/basis	Scale factor
HF/3-21G	0.9085	HF/631-G(d)	0.8928
MP2(Full)/6-31G(d)	0.9427	MP2(FC)/6-31G(d)	0.9434
SVWN/6-31G(d)	0.9833	BLYP/6-31G(d)	0.9940
B3LYP/6-31G(d)	0.9613		

(*)nlinear linear

The molecule must be described either as a nonlinear or a linear molecule.

11.1.3.2 MODES

MODES (*)all frqmin *frqmin* **frqmax** *frqmax*

MODES include 1 2 3 ...

MODES exclude 1 2 3 ...

By these commands the vibrational modes used in the ADP calculation are selected. Only one type of **MODES** command is allowed.

all

The default option includes all normal modes.

frqmin *frqmin* **frqmax** *frqmax*

If the command is starred, wavenumbers in the range of $\text{frqmin} < \nu < \text{frqmax}$ are included. The default upper cutoff value of 1500 cm^{-1} limits the calculation to "soft" (large- amplitude) modes.

include 1 2 3 ...

The normal modes are listed in the `gaussian.out` file in order of increasing wavenumbers. The **include** command allows the user to select a set of normal modes according to their order number on the list. Those not listed in the command line will be ignored.

exclude 1 2 3 ...

The listed modes will be excluded. For the **MODES** command with **include** or **exclude** options more than one line can be entered (but not mixed!).

11.3.2.3 DATAFILE

DATAFILE ***gaussian** *filename* **orient** ***standard**

This command will be used to identify different files containing the calculated vibrational data.

***gaussian** *filename*

The only type of data file allowed in the present version of **XDVIB1** is a *Gaussian98/03* output file. Its name is given as a character string *filename*.

orient ***standard** **input**

Atomic positional coordinates corresponding to the standard orientation of the molecule (the origin is placed at the center of molecular nuclear charge) are used. If the coordinates are read from the checkpoint file of a previous geometry optimization, they are listed as "Input Orientation" in the `gaussian.out` file of the frequency job, and in this case the **input** option should be starred.

11.2 XDVIB2 – A Program to Transform ADPs from Cartesian Systems to the Crystal Frame.

11.2.1 Overview

As mentioned above, the internal adp's (\mathbf{V}_a) calculated by **XDVIB1** refer to the equilibrium molecular geometry. The conformation of the optimized molecule can considerably differ from that found in the crystal (experimental conformation). However, the comparison of Δ_{ab} values corresponding to short links in structurally analogous molecules reveals only a moderate conformational dependence. For bonds formed by atoms of comparable nuclear mass, Δ is invariant under rotation about the bond vector. 1-3 links show a similar transferability. This local symmetry of the adp's is made use of when they are transformed from Cartesian (optimized molecule) to the crystal (experimental frame) coordinate system.

The procedure starts with generation of atomic local frames defined in the same way as for the spherical harmonics in the multipole model. **XDVIB2** does that automatically, using the atomic connectivity of the input molecule or fragment. The calculated adp's are then transformed into the local frames

$$\overline{\mathbf{V}}_a = \mathbf{M}_a \mathbf{V}_a \mathbf{M}_a'$$

where \mathbf{M}_a is an orthogonal 3x3 matrix whose column vectors are the components of the local basis vectors in the Cartesian system. The same transformation is applied to the experimental adp's :

$$\overline{\mathbf{U}}_a = \mathbf{N}_a \mathbf{O} \hat{\mathbf{U}}_a \mathbf{O}^{-1} \mathbf{N}_a'$$

where \mathbf{N}_a is based on the local connectivity of the experimental geometry transferred to a Cartesian system via the matrix \mathbf{O} . In the last step the calculated adp's are transferred to the crystal frame

$$\hat{\mathbf{V}}_a = \mathbf{P}_a^{-1} \overline{\mathbf{V}}_a \mathbf{P}_a \quad \mathbf{P}_a = \mathbf{N}_a \mathbf{O}$$

The program prints out the full $\Delta(\mathbf{V})$ matrices in both representations (Cartesian and crystal). Their comparison can suggest a possible model for the segmentation that is applicable to the molecule during the LS refinement.

If

$$\Delta_{ab}(\hat{\mathbf{V}}) \cong \Delta_{ab}(\mathbf{V})$$

for all links, the application of a full set of independent rigid-link constraints is feasible. Otherwise, the user should try to identify rigid groups and limit the constraints to intra-group links. The former approach corresponds to a rigid-body, while the latter to a segmented rigid-body model refinement.

For a recent publication describing the use of the **XDVIB** programs see reference 4.

11.2.2 Files used and created by XDVIB2

Input: `xd.mas, xd_vib2.inp, xd.inp`
 Output: `xd_vib2.res, xd_vib2.out`

11.2.3 Input instructions for XDVIB2

There are no input instructions for **XDVIB2**. The program reads the file containing the Cartesian atomic parameters (`xd_vib1.par` as calculated and output by **XDVIB1** and renamed as `xd_vib2.inp`) for as many molecules or atomic groups as many is needed to build the unit cell. The molecules are separated by a **GROUP** command line. Let us suppose there are two independent molecules in the unit cell and we completed the two *Gaussian* calculations, each followed by an **XDVIB1** run. The `xd_vib1.par` files obtained for the two calculations can then be merged to create the `xd_vib2.inp` files. Since the Cartesian-crystal transformation is done through the local systems, the parameters of the different groups can be given in different (but orthogonal) frames. This makes it easy to build a database of calculated adp's.

Bibliography

1. V. Schomaker and K. N. Trueblood, K.N. *Acta Cryst.* **B54**, 507 (1968)
2. J. W. Ochterski, Vibrational Analysis in *Gaussian* (October, 1999) PDF file available from : http://gaussian.com/g_whitepap/vib/vib.pdf
3. S. J. Cyvin, *Molecular Vibration and Mean Square Amplitudes*, Amsterdam: Elsevier, 1968.
4. A. E. Whitten and M. A. Spackman *Acta Cryst.* **B62**, 875-888 (2006)

11.3 XDPDF - A Program to Calculate the Probability Density Function of an Atom.

The program XDPDF calculates the vibrational probability density function (pdf) of a single atom using the least-squares estimated thermal tensors up to the fourth order Gram-Charlier expansion [1-3]. The pdf is obtained either as a 3-D grid in sections parallel to the real cell axes (or the U_{ij} basis vectors), or as a 2-D grid in a specified plane. XDPDF also provides an estimation [2] of the data resolution required for satisfactory refinement of anharmonic coefficients.

As outlined in the Introduction, in the harmonic approximation, the vibrational pdf of the nuclear displacement vector \mathbf{u} , taken with respect to the equilibrium position ($\mathbf{u} = \mathbf{q} - \mathbf{q}_0$), is a normal trivariate Gaussian distribution:

$$P_o(\mathbf{u}) = (2\pi)^{-3/2} (\det \mathbf{U})^{-1/2} \exp[-(\mathbf{u}^T \mathbf{U}^{-1} \mathbf{u}) / 2] \quad [1]$$

where \mathbf{U} is the mean-square displacement amplitude (MSDA) matrix.

The corresponding atomic anisotropic temperature factor is the Fourier transform of $P_o(\mathbf{u})$:

$$t_o(\mathbf{h}) = \exp(-2\pi^2 \mathbf{h}^T \mathbf{U} \mathbf{h}) \quad [2]$$

In XD, the anharmonic model used is the Gram-Charlier expansion [1], in which the anharmonic pdf is approximated in terms of zero and higher derivatives of the normal distribution:

$$P(\mathbf{u}) = (1 + \frac{1}{3!} C_{jkl} H_{jkl} + \frac{1}{4!} D_{jklm} H_{jklm} + \dots) P_o \quad [3]$$

where H_{jkl} , H_{jklm} are three dimensional Hermite polynomials being functions of \mathbf{U} and \mathbf{u} , while the coefficients C_{jkl} , D_{jklm} are the quasi-moments being related to the moments of the pdf. The advantage of this form is that its Fourier transform is reduced to a simple power series expansion about the harmonic temperature factor:

$$T(H) = (1 - \frac{4}{3} \pi^3 i C_{jkl} h_j h_k h_l + \frac{2}{3} \pi^4 D_{jklm} h_j h_k h_l h_m + \dots) T_o(H) \quad [4]$$

It is always important to ascertain whether the thermal parameters obtained by the least-squares refinement represent physically realistic atomic motion. For harmonic anisotropic motion this is straight-forward - if any of the eigenvalues of the U_{ij} tensor are negative (*i.e.* the tensor is non-positive definite), this implies an imaginary displacement. When anharmonic motion is involved however, it is less obvious, since the Gram-Charlier expansion coefficients C_{jkl} and D_{jklm} have (in general) no direct physical meaning [3]. The best way is to examine carefully the derived pdf function. A 3D map encompassing the full pdf should be examined (*cell option). Negative regions of the pdf are physically meaningless - however, due to the limited accuracy of the refined coefficients, the pdf may have some negative features far from the center, with small absolute values. If negative regions are found close to the center of the pdf, especially if they are surrounded by positive areas, it is likely that the refined anharmonic coefficients are meaningless. The program reports the ranges in the pdf for each section, and also writes an XD grid file. To allow for better inspection of any negative regions, the pdf may be visualised in XDGRAPH using an iso-surface plot or a relief or contour plot (see below).

It should always be remembered that there is a close mathematical resemblance between the formalisms used to describe non-Gaussian pdf's and aspherical electron densities, see for example Mallinson *et al* [4]. There are likely to be high correlations between the Gram-Charlier coefficients and the multipole populations for the same atom in the least-squares refinement, and great care should be taken to ensure that a true minimum is reached. Kuhs [2] has provided an estimation of the minimum required resolution in the data Q_n for

a satisfactory least-squares refinement for each order n of the anharmonic term. XDPDF prints a table of these Q_n values for all anisotropic atoms in the parameter file. The program expects the parameter file to contain the refined C^{jkl} and D^{ijklm} Gram-Charlier coefficients from recent versions (4.12 or later) of XDLSM. If a parameter file from an earlier version of XD is supplied, the program assumes that the anharmonic parameters are of the form U^{jkl} and U^{ijklm} , and makes the necessary conversion. Note that the reported C^{jkl} and D^{ijklm} coefficients in XD are multiplied by 10^3 and 10^4 respectively, following the standard convention.

11.3.1 Files used and created by XDPDF

Input:	xd.res, xd.mas
Optional input	xd.inp (if xd.res does not exist)
Output :	xd_pdf.out, xd_pdf.grd

11.3.2 Input instructions for XDPDF

Input instructions for XDPDF are given in the master file xd.mas and are placed between the MODULE *XDPDF and END XDPDF lines. The program writes two files, the list file xd_pdf.out and the grid file xd_pdf.grd. The grid files may be visualised using either XDGRAPH or external programs such as Moliso [5] or MolecoolQT [6] (see Figure 11-1).

11.3.2.1 SELECT

SELECT atom atom(1) scale 1.0 (*)angstrom

atom atom(1)

The label of the atom for which the pdf is to be calculated. If this label does not correspond to a non-isotropic atom in the parameter file, the pdf calculation is halted.

scale scalefactor

The pdf is multiplied by this scalefactor.

(*)angstrom

The pdf is normally calculated using the unitless coefficients β_{ij} , C^{jkl} and D^{ijklm} , in sections parallel to the real axial directions. The displacement vector \mathbf{u} is in unitless cell fractions along this basis and the resultant pdf is therefore unitless (effectively per unit cell). If ***angstrom** is selected, then the pdf is computed in units of \AA^{-3} . The coefficients U^{ij} , U^{jkl} , U^{ijklm} are used in expressions 1 and 2 above, as these have units of \AA^2 , \AA^3 and \AA^4 respectively. In this case the displacement vector \mathbf{u} is in units of \AA along the basis vector ($a^*\mathbf{a}$, $b^*\mathbf{b}$, $c^*\mathbf{c}$) – see ref 3. Only for orthorhombic, tetragonal and cubic unit cells will this basis form an orthogonal set. For these unit cells only, the unitless pdf can be converted to units of \AA^{-3} , by dividing by the cell volume in \AA^3 . For all other cells, this is not the case and the view of the grid file xd_pdf.grd in XDGRAPH will be distorted, since the latter program expects an orthogonal isometric grid.

11.3.2.2 ADDPDF

ADDPDF atom(1), atom(2), atom(3)

As well as the central atom selected above, the pdf's of the atoms atom(1)... etc will also be included in the calculation of the grid file. Isotropic atoms are not allowed in this list. Their contributions are not reported in the output file xd_pdf.out, which just lists the pdf for the central atom.

11.3.2.3 QMORD

QMORD (*)second (*)third (*)fourth

This selects which order of quasi-moments are to be included in the calculation of the pdf. If only ***second** is selected, then the Gaussian anisotropic (harmonic) pdf is calculated, while selecting ***third** or ***fourth** includes also the third and fourth order Gram-Charlier anharmonic components. Any combination of quasi-moments is acceptable - if none are selected, the harmonic pdf is calculated. The previous keyword **CUMORD** is also acceptable.

11.3.2.3 GRID**GRID (*)3-points *cryst**

If ***3-points** is selected, then a single section of the pdf is calculated through a plane described by three points. The first point is *atom(1)*, the origin of the plot plane. The other two points are given by two subsequent **ATOM** instructions, which specify respectively the x-direction (*atom(1) → atom(2)*) and the x-y plane (*atom(1) → atom(3)*)

There are 2 types of format for **ATOM** instructions

1. **ATOM label atom (symm symop) (trans tx ty tz)**
atom must be identical to an atom label given in the parameter file. **NOTE : symm and trans are included for future enhancement and are not operative in the current version of XDPDF.**
2. **ATOM xyz x y z**
the position of a point in crystallographic fractional coordinates

If ***cryst** is selected (the default), then the pdf is calculated in sections parallel to the real cell axes **a, b, c** (unitless pdf) or to the basis vector (**a**, **b**, **c**) (pdf in units of Å⁻³), controlled by the **LIMITS** instructions.

11.3.2.4 LIMITS

This command defines the limits of the calculated grid file. For the ***cryst** option, a 3-D grid is obtained when all three axes have more than one grid point.

LIMITS (keyword value) ...

Where *keyword* is any of the following (default values in brackets):

xmin [-0.8]	xmax [0.8]	nx [50]
ymin [-0.8]	ymax [0.8]	ny [50]
zmin [-0.8]	zmax [0.8]	nz [50]

If the ***cryst** option is used, *?min* and *?max* are the minimum and maximum displacements in Å from the equilibrium position of the atom along the respective basis. *n?* is the number of grid points in total in this direction.

If the ***3-points** option is used, the values of *?min* and *?max* are the minimum and maximum fractional displacements (in arbitrary units) from the equilibrium position of the atom along x- and y-directions in the plot plane. The values for **zmin**, **zmax** and **nz** are ignored. It is recommended to use equal fractions and numbers of grid points along the axes. The atom indicated in the **SELECT atom** instruction is placed in the center of the plot.



Fig 11-1 Illustrative iso-surface plots of the density modification from the 3rd and 4th order anharmonic components to a pdf, showing the positive (cyan) and negative (magenta) regions. The graphical programs used were (a) MCE [6] and (b) XDGRAPH.

Conclusion

In conclusion, to establish whether a set of refined Gram-Charlier coefficients represent a physically reasonable model.

(a) the total pdf should be calculated, using all the refined Gram-Charlier coefficients. The total pdf is best visualized as an isosurface plot, which includes both positive and negative isosurfaces.

(b) ideally the pdf should be positive everywhere. Negative parts of the pdf are physically non-meaningful. In most cases the changes due to anharmonicity are only a few percent of the total probability density, and the gaussian distribution from the harmonic part is not visibly affected.

(c) due to the limited accuracy of the refined coefficients, there may be small volumes of negative pdf, distant from the nucleus.

(d) large volumes of very negative pdf, especially near the nucleus should be regarded with the greatest suspicion. Very probably the computed anharmonic coefficients are physically meaningless. It should be realized however that the pdf close to the nucleus is very sensitive to experimental error and to the limitations of the anharmonic expansion, particularly for heavier elements. These cases need to be considered on an individual basis.

Bibliography

1. C.K. Johnson and H.A. Levy, *International Tables for X-ray Crystallography*, Vol. IV, 311-336, Birmingham: Kynoch Press, 1974.
2. W. F. Kuhs *Acta Cryst.* **A48**, 80-98 (1992).
3. K. N. Trueblood, H.-B. Bürgi, H. Burzlaff, C. M. Gramaccioni, H. H. Schultz, U. Shmueli, S. C. Abrahams *Acta Cryst.* **A52**, 770-781 (1996).
4. P. R. Mallinson, T. Koritsansky, E. Elkaim, N. Li, P. Coppens *Acta Cryst.* **A44**, 336-342 (1988).
5. C. Hübschle, P. Luger, *J. Appl. Cryst.* **38**, 901 (2006).
6. C. Hübschle, B. Dittrich, *J. Appl. Cryst.* **44**, 238 (2011).

11.4 XDCIF - A Program to Create an Archive CIF

XDCIF is a program to combine the CIFs produced by XD into an archive CIF called `xd_archive.cif`, which is suitable for submission to journals or deposition databases. Currently XD programs write the following CIFs - `xd_lsm.cif`, `xd_fft.cif` and `xd_geo.cif`. XDCIF will load these files (if found) and will also look for the files `xd_fou.cif` and `xd_pro.cif`. These latter files are not currently produced by XD, but are included for future enhancements.

In addition, a file `xd_dat.cif` is also sought. This last file is not produced by the XD programs, but must be supplied by the user. It should contain those details of the study which are not accessible in any of the other CIFs. Such information could include unit cell dimensions and errors, crystal size, crystal colour, space group symbols, data reduction details *etc.* A suitable file could be that from a SHELX refinement or a WinGX archive CIF, but of course with all the details of the refinement and structural geometry removed.

XDCIF also needs to read a request file called `xdcif.dat` which must be placed in the directory pointed to by the environment variable `XD_DATADIR` (normally also the location of the XD databank files). If the environment variable `XD_DATADIR` is not set, the program will halt. This request file may be edited by the user, and it should include all the CIF data items which the user wishes to include in the final `xd_archive.cif`. Details of the syntax for these entries is given in the header to the default version of this file which is included in the XD release. **Note :** it is not important if a data item in this request list is not found in any CIF, but if a data item is not present in the request list, then it cannot be included in `xd_archive.cif`.

The program differs from other XD programs in being interactive. It will first ask the user if all the CIFs found in the working directory should be included in the output CIF. It then cycles through the items in the request list. If a data item is not found in any CIF, the user is given the opportunity of entering the data value manually. Alternatively, the user may respond "a" (for automatic mode), when the program will continue without further prompts.

11.4.1 Files used and created by XDCIF

```
Input:      xd_*.cif (* = fft, lsm, geo, dat, fou, pro)
Output:     xd_archive.cif
```

11.5 XDWTAN - A Program to Analyse the Weighting Scheme

XDWTAN analyses the structure factor file `xd.fco` written by XDLSM and provides a listing of discrepant reflections and tables of $\langle u(F_o - F_c)^2 \rangle$ and R values as functions of hkl indices, index parities, F_o and $\sin\theta/\lambda$. It is based on the GX program WTANAL [1].

The reader should note that :

- the XDLSM refinement may be based on either F or F^2 (a user choice).
- regardless of the choice, the `xd.fco` file always contains F_o^2 , F_c^2 and $\sigma(F_o^2)$.
- the weighting analysis is always based on F_o , F_c and $\sigma(F_o)$.

The calculated weights used in XDWTAN are based on these considerations, but may not necessarily exactly correspond with those used by XDLSM, due to approximations used in the program.

The problems of choosing the correct weights for the observations in the least squares procedures are well known and of course not restricted to multipole refinements. See reference [2] for a discussion of this topic. The weighting scheme used by XDLSM is the

same as that utilized in SHELXL [3] when refinement is based on F^2 , and the weight is based on this when refining against F (see Section 4.6.6). Ideally the weights should be chosen such that $\langle(\Delta/\sigma)^2\rangle$ is unity ($\Delta=Y_o - Y_c$), but this can rarely be achieved. This condition corresponds to $\langle(w\Delta^2)\rangle \sim 1$, where w is the statistical weight equal to $1/\sigma^2(Y)$ and the so-called goodness of fit parameter (Gof or S) defined as

$$\text{Gof}(S) = [\sum w(F_o - F_c)^2 / (\text{nobs} - \text{npar})]^{1/2}$$

in XDWTAN provides a measure of this. The Gof value is generally greater than unity and this may be due either to an inadequacy of the model or to an underestimation of the observational errors. In XDWTAN, for the listing of discrepant reflections with $\Delta/\sigma(F)$ greater than 3.0, the overall Gof is used as an effective scale factor for the weights. Ideally there should be little variation of $\langle(w\Delta^2)\rangle$ with hkl indices, index parities, F_o and $\sin\theta/\lambda$. The analysis is carried out twice: firstly using all reflection data and secondly using only those reflections which were included in the least-squares refinement.

11.5.1 Files used and created by XDWTAN

Input: `xd.fco`, `xd.mas`
Output: `xd_wta.out`

Bibliography

1. P. R. Mallinson and K. W. Muir, *J. Appl. Cryst.* **18**, 51 (1985).
2. J. D. Dunitz, *X-Ray Analysis and the Structures of Organic Molecules*, VCH Weinheim, pp 213-222 (1995).

11.6 XDHKL - A Program to Analyse the Reflection Data

XDHKL analyses the reflection file `xd.hkl` (by default) or SHELX format files (if the file name is given on the command line). The file format is automatically determined by the program, which

- examines the lattice centering absences and suggests an appropriate lattice type
- prints a summary of the reflection intensity statistics
- checks for systematic absences according to the space group implied in `xd.mas` and reports whether any are found, and
- writes out a merged reflection file, containing just the unique data and with all systematic absences removed. There are no arbitrary limits on the number of reflections handled, other than the system memory on the computer.

Note that the merging process involves a simple averaging, and so is *quite inappropriate* for experimental data, where a much more sophisticated program such as SORTAV [1] should be used. The program is, however, useful for merging synthetic reflection data, *e.g.* as obtained from quantum chemical programs such as CRYSTAL [2], where there are no statistical issues involved. If the file `xd.hkl` contains phase angles (format NDAT -7, see Table 2-2), these are transferred to the merged reflection file.

11.6.1 Files used and created by XDHKL

Input: `xd.hkl`, `xd.mas`
Output: `xd_hkl.out`, `xd.hkl.merged`

Bibliography

1. R. H. Blessing, *J. Appl. Cryst.* **22**, 396 (1989).
2. <http://www.crystal.unito.it/>

11.7 AIM2TAB

The program AIM2TAB is interfaced to XD, but is distributed separately. Please send an email to Dr. Anatoliy Volkov (volkov@chem.buffalo.edu) to request the program.

AIM2TAB reads TOPOND9x, TOPXD or AIMPAC atomic integration files and calculates total molecular moments in original Cartesian coordinate system and Gaussian9x-like system using integrated atomic moments. AIM2TAB also prints out extensive information on other integrated properties, like atomic volumes, moments and integrated Lagrangians, etc. The program requires the file `files.dat` with names and locations of TOPOND/TOPXD/AIMPAC output files (1 line per filename).

There are several optional input files:

`trans.dat` to transform the coordinates (and atomic dipoles) of integrated atoms using rotation matrices and translation vectors in crystal coordinate system (1 line per atom):
 $R_{11} R_{12} R_{13} \quad R_{21} R_{22} R_{23} \quad R_{31} R_{32} R_{33} \quad T_1 T_2 T_3$

`symm.dat` to generate symmetry-equivalents of integrated atoms - 1 line per each new symmetry-equivalent atom in the following format:
`<atom label> R11 R12 R13 R21 R22 R23 R31 R32 R33 T1 T2 T3`
 where `<atom label>` is the atomic label of the integrated "parent" atom, and R_{ij} and T_i are the rotations and translations in the crystal coordinate system, respectively.

`new.dat` with additional information on atomic site occupations and number of electrons of the free atom (in the current version of AIM2TAB only atomic charges and volumes will be rescaled, NOT the higher moments). The format is (1 record per each atom):
`<atom label> <occupation> <n. of electrons in free atom>`

11.7.1 Files used and created by AIM2TAB

Input: TOPOND/TOPXD/AIMPAC outputfiles, files.dat,
 (trans.dat, symm.dat, new.dat)
 Output: aim2tab.out

11.8 LSDB

This is a program [1] for automatic setup of the atomic local coordinate systems and chemical constraints starting from SHELX [2], PLATON [3] or XD structure files. It links into the UBDB (the Buffalo Aspherical Atom Database). The program is designed for interactive use and is self-explanatory. The atomic local coordinate systems can then be visualized in PLATON. Current versions of the program for Linux and Windows and the UBDB are freely available for academic users from

<http://harker.chem.buffalo.edu/group/databank.html>

Bibliography

1. (a) A. Volkov, X. Li, T. Koritsanszky, P. Coppens *J. Phys. Chem* (2004) 108, 4283-4300 (b) P. M. Dominiak, A. Volkov, X. Li., M. Messerschmidt, P. Coppens *J. Chem. Theo. Comp.* (2006) in press.

2. G. M. Sheldrick, SHELX97. *Programs for Crystal Structure Analysis (Release 97-2)*. University of Göttingen, Germany (1997)
3. A. L. Spek *PLATON, A Multipurpose Crystallographic Tool* Utrecht University, Utrecht, The Netherlands (2003)

11.9 ADDGRID, SCALEGRID

These are two simple interactive routines to add and scale grid files produced by other XD sections. In particular, **ADDGRID** allows summing, subtracting, multiplying, dividing, scaling and applying exponents to a given number of grids. In this way, many properties derived (exactly or empirically) from the electron density and/or its derivatives can be visualized.

11.10 RAYS2POV

This program provides another way of visualising the inter-atomic surface file `rays.dat`, which is obtained when the ***SaveSurf** option is used in TOPXD (see Section 10.8). It writes an output file for the well known ray-tracing program POV-Ray, obtainable for both Windows™ and Unix/Linux platforms.

It is invoked by the command line : **rays2pov** *<input_file>* *<output_file>*

The command line parameters *<input_file>* *<output_file>* are optional, available if the user wishes to change the defaults. The default input file is `rays.dat` and the default output POV-Ray scene description file is `xd_ias.pov`. The program is controlled by the instruction file `rays2pov.ini`, which the program looks for first in the local working directory, and then in the directory pointed to by the environment variable `XD_DATADIR`. If this file is not found, the program defaults will be used. A default version of `rays2pov.ini`, as supplied with the installation, is shown below.

```
# colors - only colors given in colors.inc are acceptable (case sensitive !)  
  
SurfaceColor=Copper  
AtomColor=SteelBlue  
BondColor=SteelBlue  
PointColor=Black  
CP_Color=Red  
TextColor=Black  
  
# textures  
  
SurfaceTexture=Metal  
AtomTexture=Shiny  
BondTexture=Shiny  
PointTexture=Shiny  
CP_Texture=Shiny  
  
# numerical parameters - all in Angstroms except rCutoff in a.u.  
  
rCutOff=3.0  
BondTol=0.2  
TextOffset=0.1  
MinimumBond=0.8  
MaximumBond=3.0  
TextWidth=0.1  
TextScale=0.1  
AtomRadius=0.15  
BondRadius=0.02  
PointRadius=0.005  
CP_Radius=0.1  
X_Rotation=0.0  
Y_Rotation=0.0  
Z_Rotation=0.0  
X_Translation=0.0  
Y_Translation=0.0  
Z_Translation=0.0  
ZoomFactor=1.0  
LabelAtoms=no  
RunPovray=yes  
ShowIntersections=yes
```

The meaning of the parameters are :

SurfaceColor	Colour of IAS - any legal POV-Ray colour
AtomColor	Colour of atoms - any legal POV-Ray colour
BondColor	Colour of bonds - any legal POV-Ray colour
PointColor	Colour of points where rays intersect the IAS - any legal POV-Ray colour
CP_Color	Colour of bond critical points - any legal POV-Ray colour
TextColor	Colour of labels - any legal POV-Ray colour
SurfaceTexture	Any legal POV-Ray texture
AtomTexture	Any legal POV-Ray texture
BondTexture	Any legal POV-Ray texture
PointTexture	Any legal POV-Ray texture
CP_Texture	Any legal POV-Ray texture
rCutOff	Cutoff for IAS - in atomic units !
BondTol	bond is drawn if distance between A and B is $< (R_a + R_b + \text{bontol})$
TextOffset	offset from atom position
MinimumBond	minimum distance for bond to be drawn
MaximumBond	maximum distance for bond to be drawn
TextWidth	
TextScale	
AtomRadius	
BondRadius	
PointRadius	
CP_Radius	
X_Rotation	
Y_Rotation	
Z_Rotation	
X_Translation	
Y_Translation	
Z_Translation	
ZoomFactor	
LabelAtoms	whether to label the atoms - yes/no
RunPovray	whether to run POV-Ray after writing output file - yes/no
ShowIntersections	whether to show the points where the rays intersect the IAS - yes/no

The output file `xd_ias.pov` can be very large and take a significant time for POV-Ray to parse. The structure is intended for easy manual editing. For instance, if labelling is chosen it will probably be necessary to manually edit the individual positions of the labels.

NOTE: The option TOPINT in XDPROP also write out a rays.dat file, but these files cannot be processed by the stand-alone version of RAYS2POV.