User Manual of CLaUDe Program

Version 2.2, May 19, 2019

Contents

1 Introduction

The Suite is used to build a database coding sampled interatomic distances (*.smp files) for a population of crystalline nanoparticles of suitable shape and increasing size. Additional utilities are included such as the calculation of the powder diffraction pattern of a single nanoparticle and the calculation of the G(r) function.

2 filename.ddb

2.1 General Info

1 FILE TYPE: `form=formatted, access=sequential, recordlength=256` (max.).

2 READING METHOD: each record is read into a 256-places character variable `rline`.

3 COMMENTS: marked by a `'!'` or a > in the first character of `rline`.

4 IDENTIFIER: the first characters until the colon in `rline`

5 BUFFER: the characters following the colon in `rline`

6 MANDATORY/OPTIONAL FLAG: a symbol ! or (M) specifies that a given entry is always mandatory; * that it is always optional; ? that it is conditionally optional (depending on the value of other entries, as specified).

7 FILES AND PATHNAMES: when we refer to a file name, unless otherwise specified, to the complete (pathname-inclusive) name, as and when required by the operating system

8 SECTIONS: The file `filename.ddb` is organized in three consecutive main sections, in the order:

I the **PHASE** Section (Sec. 2.2);

II the **SHAPE/SIZE** Section (Sec. 2.3);

III the **SAMPLING** Section (Sec. 2.4).

2.2 PHASE

This section of `filename.ddb` contains the directives for reading the structural information of the building database (*.smp files) .

ident	buffer content	
`Phase_Name`	**PHASE** is the name of the file containing the phase information, provided either as asymmetric unit **PHASE.pha** or list of cartesian coordinates **PHASE.xyz**.	!
`Spacegroupnumber_orig`	#␣o# the Space group number and, when necessary the origin choice or the unique axis (the full list of Space groups is avalable in `DebUsSy-Suite_vX`/ext_database/SpaceGroups/).	!
`Atomic Species No.`	# (Integer) Number of atomic species as in **PHASE**.	!
`Cell Origin`	#␣#␣# Three real numbers. Shift to be applied to the cell origin.	!

Continues. . .

ident	buffer content	
Pearson Symbol	$\alpha\alpha$**#** Two letters followed by an integer. The Pearson symbol.	!
constr	α One letter: either **P** or **S** for resp. disabling or enabling the (optional) construction of the most symmetric unit cell around a suitably chosen origin starting from the user-supplied asymmetric unit.	*
Tolerance minimal distance	# One real number. A (small) length, in Ångstrœm, representing a tolerance margin on the interatomic distances, so that two atoms that are "too close" by a small amount, which is probably due to some small errors in the coordinates' source, are NOT considered "too close" by the distance generation algorithm.	?
Density (g/cm^3)	# One real number. The mass density of the substance.	?

2.3 SHAPE/SIZE

This section of filename.ddb contains the directives for reading the nanocrystal shape and size information of the building database.

ident	buffer content	
Shape of Clusters	**SHAPE** Three letters identifying the shape: SPH (sphere), QBE (cube), PAR (parallelepiped), CYL (cylinder), HEX (hexagonal).	!
Diam_max of SPH (nm)	**D** Real number in nm. Diameter/side of the largest *spherical*/*cubic* cluster.	?

Continues...

ident	buffer content	
N_max of SPH	**N** (Integer) Number of shells used to build the largest *spherical*/*cubic* cluster.	?
D_max of PAR/CYL/HEX (nm)	**D** Real number in nm. Diameter of the circle of equivalent area to the base (*ab*-plane) of the largest *PAR/CYL/HEX* cluster.	?
L_max of PAR/CYL/HEX (nm)	**L** Real number in nm. Length (along *c*-axis) of the largest *PAR/CYL/HEX* cluster.	?
N1_max of PAR/CYL/HEX	**N1** (Integer) Number of layers used to build the largest cluster in the *ab*-plane.	?
N2_max of PAR/CYL/HEX	**N2** (Integer) Number of layers used to build the largest cluster along the *c*-direction.	?
TODO	**TODO** Three cases : 1 **TODO** = all_clusters: all the clusters will be built 2 **TODO** = largest_only: for each value of **D** (or **L** or **N2**) only the clusters with maximum value of **D** or **N1** will be built. 3 **TODO** = all_clusters␣4: Make all clusters with all possible heights (**L** or **N2**) and with many possible base sizes (**D** or **N1**) up to the maximum, but (in number of cells) the base will increase with step 4 cells instead of (default) 1 cell.	!

Continues...

ident	buffer content

PARA

PARACRYSTALLINITY MODEL *

Paracrystalline disordered model.
Two cases :

1 **MODEL** = WelbAnys #␣#␣#␣#␣#␣#␣ ␣#␣: Paracrystalline two-dimensional disorder. 9 numbers following, the first 7 are used. They are: r_a, s_a, r_b, s_b, c_{xy}, σ_a, σ_b. The latter two define the variance at infinity as $2\sigma_a^2$ or $2\sigma_b^2$ along a or b,respectively.

2 **MODEL** = WelbIsot #␣#␣#␣#␣#␣#␣ ␣#␣: Paracrystalline isotropic disorder. 9 numbers following, the first 2 are used. They are: σ_∞, L_{decorr}. The former defines the variance at infinity as $2\sigma_\infty^2$ or $2\sigma_b^2$. The second is a length (decorrelation length) defining how fast the asymptotic variance is approached. The law of increase of the variance with the distance d is

$$\sigma^2 = 2\sigma_\infty^2\left(1 - \mathrm{e}^{-d/L}\right).$$

OCC1

OCCUPANCY *

Occupation site factors flag
Two cases :

1 **OCCUPANCY** = y, Y, 1: Site occupation factors will be forced to 1.0 for each atomic specie (independently on the value set in the .pha file)

2 **OCCUPANCY** = n, N, 0: Site occupation factors will be set to the value given in the **PHASE** file

Continues...

ident	buffer content	
XYZ?	**XYZ** Real space coordinates flag Two cases : 1 **XYZ** = y, Y, 1: In addition to the sampled-distances clusters database also the xyz-clusters will be generated according to the value of the **TODO** variable 2 **XYZ** = n, N, 0: The xyz-clusters will not be generated	*

2.4 SAMPLING

This section of filename.ddb contains the directives for reading the sampling information of the building database.

ident	buffer content	
Sampling	**SAMPLING** Two cases : 1 **SAMPLING** = one: Inter-atomic distances will be sampled by using one sampling step (in the range 0.3- 96.0 $\AA$), derived according to **LAMBDA** and $2\theta_{\max}$ values supplied in the next lines. 2 **SAMPLING** = all: If all is found, all sampling steps will be used, producing 32 sets of sampled distances.	!
Wavelength	**LAMBDA** The wavelength (in Å) used for data collection. If zero, a working wavelength 0.1477211 Å is assumed.	!
2-Theta Max	$2\theta_{\max}$ Real number. Max 2θ angle (in degrees) of collected data. If zero, $2\theta_{\max}$=160.0 deg is assumed.	

3 PHASE

The phase information can be provided either as an asymmetric unit file **.pha** or as a list of cartesian coordinates **.xyz** file.
If a Crystallographic Information File, `.cif`, file is available the **.pha** file can be conveniently generated using the utility program `CIF2PHA` which is a part of CLaUDe. `CIF2PHA` converts .cif files into the internal standard **.pha** file format; presently all the conventional .cif formats, downloadable from the crystallographic databases (CSD, ICSD, American Mineralogist, COD), are managed. To get the **.pha** from the `.cif` one using the `CIF2PHA` program it is sufficient to type the following command in the terminal window:

```
<path_to_program>/CIF2PHA <filename.cif> <filename.pha>
```

Such operation is automatized within the GUI.

3.1 PHASE.pha

3.1.1 General Info

1 FILE TYPE: `form=formatted, access=sequential, recordlength=256` (max.).

2 READING METHOD: each record is read into a 256-places character variable `rline`.

3 COMMENTS: marked by a `'!'` or a > in the first character of `rline`.

4 IDENTIFIER: the first five characters in `rline`

5 BUFFER: the characters following 5th in `rline(5:)`

6 MANDATORY/OPTIONAL FLAG: a symbol `!` or (M) specifies that a given entry is always mandatory; `*` that it is always optional; `?` that it is conditionally optional (depending on the value of other entries, as specified).

7 FILES AND PATHNAMES: when we refer to a file name, unless otherwise specified, to the complete (pathname-inclusive) name, as and when required by the operating system

8 SECTIONS: The file `PHASE.pha` contains only one section:

I the **PHASE** Section (Sec. 3.1.2);

3.1.2 PHASE.pha file

ident	buffer content
Title	**PHASE** ! is the name of the phase.
Cell	$a\ b\ c\ \alpha\ \beta\ \gamma$! Six real numbers. Lattice parameters and angles.
Space	**SPACE_GROUP** ! The space group number and, when necessary the origin choice or the unique axis.
Coord	**X**␣#␣#␣#␣#␣#␣# ! Atomic symbol, atomic species identifier, x-, y-, z- coordinates, isotropic thermal parameter, site occupation factor.

3.2 PHASE.xyz

3.2.1 General Info

1 FILE TYPE: `form=formatted, access=sequential, recordlength=256` (max.).

2 READING METHOD: each record is read into a 256-places character variable `rline`.

3 COMMENTS: marked by a `'!'` or a > in the first character of `rline`.

4 SECTIONS: The file `PHASE.xyz` is a one section file.

3.2.2 PHASE.xyz file

buffer content
! Number of atoms listed in the files.
PHASE.xyz ! The phase name.

Continues...

`buffer` content	
string␣x␣y␣z The atom symbol followed by the three cartesian coordinates. The atoms should be listed in order of increasing atomic number.	!

4 diffractor.inp

4.1 General Info

1 FILE TYPE: `form=formatted, access=sequential, recordlength=256` (max.).

2 READING METHOD: each record is read into a 256-places character variable `rline`.

3 COMMENTS: marked by a `'!'` or a > in the first character of `rline`.

4 IDENTIFIER: the first four characters in `rline`

5 BUFFER: the characters following 4th in `rline(4:)`

6 MANDATORY/OPTIONAL FLAG: a symbol `!` or (M) specifies that a given entry is always mandatory; `*` that it is always optional; `?` that it is conditionally optional (depending on the value of other entries, as specified).

7 FILES AND PATHNAMES: when we refer to a file name, unless otherwise specified, to the complete (pathname-inclusive) name, as and when required by the operating system

8 SECTIONS: The file `diffractor.inp` contains only one section:

I the **DIFFRACTOR** Section (Sec. 4.2);

4.2 DIFFRACTOR

The `diffractor.inp` file contains the directives for calculating the diffraction pattern from files containing the sampled interatomic distances of a nanocrystalline cluster.

`ident` `buffer` content

`VARX` **VARX** *

Two cases :

1 **VARX** = `twotheta`: Default. Generates a powder pattern at constant 2θ step as specified later.

2 **VARX** = `q`: Generates a powder pattern at constant $q = 2\sin(\theta)/\lambda$ step, specified later. Does not correct for the variable solid angle!!!

`WLEN` **WLEN** !

Two cases :

1 **WLEN** = λ [real number Å]: The wavelength.

2 **WLEN** = λ [real number, Å], `istep` [integer, multiple of 30, mÅ]: The wavelength and the sampling step Δ (e.g. istep=120 $\rightarrow \Delta = 0.12$ Å). Two numbers (real, integer) separated by space(s).

`RANG` **min␣max␣step** !

Three real numbers. Minimum, maximum , step for the **VARX** variable 2θ, q will be recalculated accordingly.

`QRAN` **min␣max␣step** !

Three real numbers. Minimum, maximum , step for the **VARX** variable q, 2θ will be recalculated accordingly.

Continues. . .

ident	buffer content	
RAYS	**flag_rad** indicating the radiation used. Three cases : 1 **RAYS** = x: X-Rays 2 **RAYS** = n: Neutrons 3 **RAYS** = e: Electrons	?
SOFQ	**i3** Two cases : 1 **i3** = 0: the full powder pattern is calculated; 2 **i3** = 1: $S(q) - 1$ is calculated instead [meaning that the zero-distance contribution is subtracted and an appropriate scaling factor is applied].	*

Continues...

ident	buffer content	

DIVI **DIVI** *

Six cases :

1 **DIVI** = f2a: The scaling factor applied to $S(q) - 1$ (if the SOFQ option is activated) is $\langle f^2(q) \rangle$ (Xrays/electrons) or $\langle b^2 \rangle$ (neutrons);

2 **DIVI** = b2a: The scaling factor applied to $S(q) - 1$ (if the SOFQ option is activated) is $\langle f^2(q) \rangle$ (Xrays/electrons) or $\langle b^2 \rangle$ (neutrons);

3 **DIVI** = Z2a: The scaling factor applied to $S(q) - 1$ (if the SOFQ option is activated) is $\langle Z^2 \rangle$ (Xrays/electrons) or $\langle b^2 \rangle$ (neutrons);

4 **DIVI** = fa2: The scaling factor applied to $S(q) - 1$ (if the SOFQ option is activated) is $\langle f(q) \rangle^2$ (Xrays/electrons) or $\langle b \rangle^2$ (neutrons);

5 **DIVI** = ba2: The scaling factor applied to $S(q) - 1$ (if the SOFQ option is activated) is $\langle f(q) \rangle^2$ (Xrays/electrons) or $\langle b \rangle^2$ (neutrons);

6 **DIVI** = Za2: The scaling factor applied to $S(q) - 1$ (if the SOFQ option is activated) is $\langle Z \rangle^2$ (Xrays/electrons) or $\langle b \rangle^2$ (neutrons).

IMAX **Y** *

Real number > 0. The pattern is scaled so that its maximum is equal to **Y**. If **Y**=0, then no scale factor is applied. Disabled when SOFQ activated.

HKLS **hkls** *

Logical (True/False/1/0), default 0. Whether to calculate the positions of HKL reflections according to **WLEN** and the lattice constants and space group information encoded in the **fname** file. The output is stored in a file `phase`.hkl where `phase` is **fname** stripped of the extension, shape and size codes.

Continues...

ident	buffer content	
CMPT	**add_compton** Binary variable. Whether to add Compton scattering component for X-ray scattering.	*
OUTP	**io** Integer number. If **io**=0, only 5 col.s are in output (the first 3 are 2θ, q, $I_{calc}(q)$ [or $S(q)-1$ if SOFQ activated]; then we have $I_0(q)$ and $K(q)$, those two being the term that is (or would be) subtracted and the scale factor, respectively, to convert $I_{calc}(q)$ to $S(q)-1$. If **io**=1, a lot of other columns appear after the first 5. See the code.	*
PATH	**path** Character string. The path (relative or absolute) to the folder where the SAMPTO### folder containing the .smp and .smp_INFO files is located. Note that ### is a 3-digit integer representing the sampling step (see WLEN above).	!
FILE	**fname** Character string. The name of the .smp file containing the sampled interatomic distances. Extension has to be included, e.g. **fname**=<filename>.smp.	!
NATO	**nat** Integer number. The number of atom species. It must be the same as in the .smp_INFO file.	!

Properties of the atoms: each line input is now a vector of dimension **nat**.

ZELE	**Z(nat)** Integer vector. The Z's of the atom species. Can be the same or not from the .smp_INFO file.	!

Continues...

ident	buffer content	
BATO	**bato(nat)** Real vector. The isotropic Debye-Waller parameters of the atom species.	!
ATOC	**occ(nat)** Real vector. The occupancies of the atom species. Can be the same or not as in the `.smp_INFO` file.	*
FPRI	**f'(nat)** Real vector. The f' anomalous scattering factors.	*
FDPR	**f"(nat)** Real vector. The f" anomalous scattering factors.	*
NSCL	**nscl(nat)** Real vector. The neutron scattering lengths of the atom species. Can be any value. If `NSCL` is not present, the internally tabulated scattering lengths of the elements are used.	*

5 dopdf.inp

5.1 General Info

1 FILE TYPE: `form=formatted, access=sequential, recordlength=256` (max.).

2 READING METHOD: each record is read into a 256-places character variable `rline`.

3 COMMENTS: marked by a `'!'` or a > in the first character of `rline`.

4 IDENTIFIER: the first four characters in `rline`

5 BUFFER: the characters following 4th in `rline(4:)`

6 MANDATORY/OPTIONAL FLAG: a symbol ! or (M) specifies that a given entry is always mandatory; * that it is always optional; ? that it is conditionally optional (depending on the value of other entries, as specified).

7 FILES AND PATHNAMES: when we refer to a file name, unless otherwise specified, to the complete (pathname-inclusive) name, as and when required by the operating system

8 SECTIONS: The file `dopdf.inp` contains only one section:

the **DOPDF** Section (Sec. 5.2);

5.2 DOPDF

The `dopdf.inp` file contains the directives for calculating the pair distribution function from files containing the calculated powder diffraction pattern of a nanocrystalline cluster and/or experimental data.

`ident`	`buffer` content	
`NFIL`	**nfil** Integer number. The number of files to be processed.	!
`FILE`	**fname** Character string. The name of the file containing the sampled interatomic distances.	!
`NCOLS`	**ncols** Integer number. The number of columns of **fname**. Default is **ncol** = 2.	*

Continues. . .

ident	buffer content	
TTCOL	**ttcol** Either a single integer number or two integers followed by two reals. In the first case the integer is the column number of the 2θ values to be read in **fname**, default is **ttcol** = 1. In the second case two integers denote two column numbers and the reals are the coefficients for the linear combination of the two. Note that column number equal to 0 means a constant value of 1.0.	*
I_COL	**icol** Either a single integer number or two integers followed by two reals. In the first case the integer is the column number of the intensity values to be read in **fname**, default is **icol** = 2. In the second case two integers denote two column numbers and the reals are the coefficients for the linear combination of the two. Note that column number equal to 0 means a constant value of 1.0.	*
E_COL	**ecol** Either a single integer number or two integers followed by two reals. In the first case the integer is the column number of the sigma values to be read in **fname**, default is **ecol** = 3. In the second case two integers denote two column numbers and the reals are the coefficients for the linear combination of the two. Note that column number equal to 0 means a constant value of 1.0.	*
RAYS	**flag_rad** indicating the radiation used. Three cases : 1 **RAYS** = x: X-Rays 2 **RAYS** = n: Neutrons 3 **RAYS** = e: Electrons	?

Continues...

ident	buffer content	
WLEN	**WLEN** = λ [real number Å]: The wavelength.	!
POLA	**eccentricity**, **phi** In case **RAYS** = X, two real numbers: **eccentricity** ratio b/a of electric field; **phi** angle ϕ (deg) of the closest ellipse axis (def. as b) to the scattering plane. For $s-$polarised Synchrotron radiation **eccentricity**≈0.01, **phi**=0.0, for circularly polarised laboratory X-ray source without monochromator **eccentricity**=1.0, **PHI**=0.0.	*
ARANG	**tt0 tt1 dtt** Three real numbers. Minimum 2θ, maximum 2θ, step in 2θ.	!
RRANG	**rmin rmax rstep** Three real numbers. Minimum r, maximum r, step in r.	!
NQCUT	**ncut_Q**, **is_cut_rel** = Integer indicating the number of Q cutoffs, integer flag indicating whether the cutoff is relative (fraction, **is_cut_rel**=1) or absolute (Q-value, **is_cut_rel**=0). Cutoff value(s) is(are) given at QCUTV.	!
QCUTV	**Qcut** = Real array with dimension **ncut_Q**, containing the Q cutoffs values.	!
QMIN	**qmin** = Real value for the minimum cutoff value of Q, the effective Q_{min} is the maximum between **qmin** and **Q0**= $4\pi \sin(\textbf{tt0}/2)/\textbf{WLEN}$.	*

Continues...

ident	buffer content	
DONOR	**norm_mode** indicating the normalization option Four cases : 1 **DONOR** = fa2: The scaling factor applied to $S(q)-1$ is $\langle f(q)\rangle^2$ (Xrays/electrons) or $\langle b\rangle^2$ (neutrons) – this is the default choice; 2 **DONOR** = f2a: The scaling factor applied to $S(q)-1$ is $\langle f^2(q)\rangle$ (Xrays/electrons) or $\langle b^2\rangle$ (neutrons); 3 **DONOR** = Za2: The scaling factor applied to $S(q)-1$ is $\langle Z\rangle^2$ (Xrays/electrons) or $\langle b\rangle^2$ (neutrons); 4 **DONOR** = Z2a: The scaling factor applied to $S(q)-1$ is $\langle Z^2\rangle$ (Xrays/electrons) or $\langle b^2\rangle$ (neutrons).	*

Continues...

ident	buffer content	
SCALE	**extscale**, **scale_val** Character (length=1) string and a real number (conditionally optional) . Five cases : 1 **extscale** = `m` manual scale, followed by the scale value **scale_val**; 2 **extscale** = `t` tailing scale, followed by the scalar **tail_fraction**; 3 **extscale** = `p` scale factor such that $\langle S(Q)\rangle = 1$, **scale_val** ignored; 4 **extscale** = `z` followed by the optional scalar **tail_fraction**. A scale factor s and a background b are calculated such that $S(Q) = 1 + \frac{sI(Q)-b-\langle I_i+\langle f^2\rangle\rangle}{\langle f\rangle^2}$, with constraints: $S(Q \to 0) \to 0$, $S(Q \to +\infty) \to +1$; 5 **extscale** = `w` followed by the optional scalar **tail_fraction**. Like with `z` with the additional constraint such $\int_{Q_{\min}}^{Q_{\max}} F(Q)\mathrm{d}Q = \int_{Q_{\min}}^{Q_{\max}} Q[S(Q) - 1]\mathrm{d}Q \approx 0$.	*
BROAD	**ubroad** Real number. Additional artificial uniform Gaussian broadening, used to reduce truncation ripples. **ubroad** is the added r.m.s. displacement.	*
VALEN	**valence** Logical (True/False/1/0), default True. Whether to use modified scattering form factors for elements with $5 \leq Z \leq 9$.	*
INCOH	**use_incoh** for incoherent scattering correction, integer flag (1: True (default), 0: False).	!

Continues...

ident	buffer content	
NATO	**nat** Integer number. The number of atom species.	!
ZELE	**Z(nat)** Integer vector. The Z's of the atomic species.	!
CHEM	**x(nat)** Integer vector. The fractional composition.	!

6 Examples

6.1 filename.ddb examples

```
!
! PHASE SECTION
!
Phase_Name (.pha) (M) :   magnetite.pha
Spacegroupnumber_orig (M): 227 o1
Atomic Species No. (M):   3
Cell Origin (M): 0.0 0.0 0.0
Pearson Symbol (M) (max 4 ch.): cF56
Constr :   P
!
! SHAPE/SIZE SECTION
!
Shape of Clusters (M) : SPH
Diam_max of SPH (nm) : 15.0
N_max of SPH : 0.0
D_max of PAR/CYL/HEX (nm) : 0
L_max of PAR/CYL/HEX (nm): 0
N1_max of PAR/CYL/HEX: 0
N2_max of PAR/CYL/HEX: 0
TODO all_clusters
!PARA WelbAnys 0.999 0.99 0.99  0.99 0.99 0.99 0.99 0.99 0.99
!
! SAMPLING SECTION
!
Sampling (M): one
Wavelength (M): 0.77482143
2-Theta Max (M): 140.0
```

Figure 1: filename.ddb input file used to build a spherical database of magnetite. The asymmetric unit is contained in the magnetite.pha file (see Section 6.2).

```
!
! PHASE SECTION
!
Phase_Name (.pha) (M) :   anatase.pha
Spacegroupnumber_orig (M): 141 o2
Atomic Species No. (M):   2
Cell Origin (M): 0.0 0.0 0.0
Pearson Symbol (M) (max 4 ch.): tI12
!
! SHAPE/SIZE SECTION
!
Shape of Clusters (M) : PAR
Diam_max of SPH (nm) : 0
N_max of SPH : 0.0
D_max of PAR/CYL/HEX (nm) : 0
L_max of PAR/CYL/HEX (nm): 0
N1_max of PAR/CYL/HEX: 20
N2_max of PAR/CYL/HEX: 20
TODO all_clusters
!PARA WelbAnys 0.999 0.99 0.99  0.99 0.99 0.99 0.99 0.99 0.99
!
! SAMPLING SECTION
!
Sampling (M): one
Wavelength (M): 0.729568
2-Theta Max (M): 0.0
```

Figure 2: `filename.ddb` input file used to build a spherical database of anatase. The asymmetric unit is contained in the anatase.pha file (see Section 6.2).

```
!
! PHASE SECTION
!
Phase_Name (.pha) (M) :   sio2_amorphous.pha
Spacegroupnumber_orig (M):
Atomic Species No. (M):   2
Cell Origin (M): 0.0 0.0 0.0
Pearson Symbol (M) (max 4 ch.): aP00
Constr :   P
0.20000000
dens 2.138382221476936
!
! SHAPE/SIZE SECTION
!
Shape of Clusters (M) :
Diam_max of SPH (nm) : 0
N_max of SPH : 0.0
D_max of PAR/CYL/HEX (nm) : 0
L_max of PAR/CYL/HEX (nm): 0
N1_max of PAR/CYL/HEX: 0
N2_max of PAR/CYL/HEX: 0
TODO all_clusters
!
! SAMPLING SECTION
!
Sampling (M): one
Wavelength (M): 0.49592
2-Theta Max (M): 140.0
```

Figure 3: filename.ddb input file used to generate the sampled distances file from the list of cartesian coordinates of a cluster of amorphous silica.

6.2 PHASE.pha examples

```
Title  magnetite
Cell  8.3457 8.3457 8.3457  90.00  90.00  90.00
Space 227 o1
>
Coord Fe 1  0.0          0.0         0.0        0.5   1.00
Coord Fe 2  0.625        0.625       0.625      0.5   1.00
Coord O  3  0.37968      0.37968     0.37968    0.5   1.00
```

Figure 4: magnetite.pha file.

```
Title  anatase
Cell  3.7994 3.7993 9.4980 90.0 90.0 90.0
Space  141 o2
>
Coord Ti  1    0.0    0.250    0.375    0.39    1.00
Coord O   2    0.0    0.250    0.16686  0.61    1.00
```

Figure 5: anatase.pha file.

6.3 diffractor.inp examples

```
STEP  twotheta
RAYS  x
WLEN  0.77482143
TWOT  5.0  120.0 0.02
IMAX  0.0
PATH  DISTANCES/SAMPTO
FILE  magnetite_r010_SPH.smp
NATO  3
ZELE  26  26    8
ATOC  1.0 1.0 1.0
BATO  0.1 0.1 0.1
FPRI  0.0 0.0 0.0
FDPR  0.0 0.0 0.0
```

Figure 6: diffractor.inp file for a magnetite sampled distances cluster file.

7 Higher-level inputs

7.1 PHASE.cel

Input file for most programs used to generate a file of interatomic sampled distances (*.smp) for nanocrystals of any shape.

7.1.1 PHASE.cel example

```
 cF56  8.390000      8.390000      8.390000      90.000000      90.000000      90.000000        3
          8            16           32
         26            26            8
Fe       0.125                 0.125                 0.125                 1.                    0.7
Fe       0.625                 0.125                 0.625                 1.                    0.7
Fe       0.125                 0.625                 0.625                 1.                    0.7
Fe       0.875                 0.375                 0.375                 1.                    0.7
Fe       0.875                 0.875                 0.875                 1.                    0.7
Fe       0.625                 0.625                 0.125                 1.                    0.7
Fe       0.375                 0.875                 0.375                 1.                    0.7
Fe       0.375                 0.375                 0.875                 1.                    0.7
Fe       0.5                   0.5                   0.5                   1.                    0.7
Fe       0.25                  0.75                  0.                    1.                    0.7
Fe       0.75                  0.                    0.25                  1.                    0.7
Fe       0.5                   0.                    0.                    1.                    0.7
Fe       0.                    0.5                   0.                    1.                    0.7
Fe       0.                    0.25                  0.75                  1.                    0.7
Fe       0.75                  0.25                  0.                    1.                    0.7
Fe       0.25                  0.25                  0.5                   1.                    0.7
Fe       0.75                  0.75                  0.5                   1.                    0.7
Fe       0.25                  0.                    0.75                  1.                    0.7
Fe       0.75                  0.5                   0.75                  1.                    0.7
Fe       0.5                   0.25                  0.25                  1.                    0.7
Fe       0.5                   0.75                  0.75                  1.                    0.7
Fe       0.                    0.75                  0.25                  1.                    0.7
Fe       0.                    0.                    0.5                   1.                    0.7
Fe       0.25                  0.5                   0.25                  1.                    0.7
O        0.2546791666666667    0.2546791666666667    0.2546791666666667    1.                    0.8
O        0.4953208333333333    0.9953208333333333    0.7546791666666667    1.                    0.8
O        0.9953208333333333    0.7546791666666667    0.4953208333333333    1.                    0.8
O        0.0046791666666667    0.5046791666666667    0.2453208333333333    1.                    0.8
O        0.7453208333333333    0.7453208333333333    0.7453208333333333    1.                    0.8
O        0.2546791666666667    0.7546791666666667    0.7546791666666667    1.                    0.8
O        0.7546791666666667    0.2546791666666667    0.7546791666666667    1.                    0.8
O        0.7546791666666667    0.4953208333333333    0.9953208333333333    1.                    0.8
O        0.5046791666666667    0.0046791666666667    0.2453208333333333    1.                    0.8
O        0.4953208333333333    0.4953208333333333    0.2546791666666667    1.                    0.8
O        0.9953208333333333    0.9953208333333333    0.2546791666666667    1.                    0.8
O        0.2453208333333333    0.0046791666666667    0.5046791666666667    1.                    0.8
O        0.5046791666666667    0.2453208333333333    0.0046791666666667    1.                    0.8
O        0.0046791666666667    0.2453208333333333    0.5046791666666667    1.                    0.8
O        0.9953208333333333    0.2546791666666667    0.9953208333333333    1.                    0.8
O        0.4953208333333333    0.7546791666666667    0.9953208333333333    1.                    0.8
O        0.7453208333333333    0.0046791666666667    0.0046791666666667    1.                    0.8
O        0.7453208333333333    0.5046791666666667    0.5046791666666667    1.                    0.8
O        0.2453208333333333    0.5046791666666667    0.0046791666666667    1.                    0.8
O        0.7546791666666667    0.9953208333333333    0.4953208333333333    1.                    0.8
O        0.2546791666666667    0.4953208333333333    0.4953208333333333    1.                    0.8
O        0.7546791666666667    0.7546791666666667    0.2546791666666667    1.                    0.8
O        0.2546791666666667    0.9953208333333333    0.9953208333333333    1.                    0.8
O        0.9953208333333333    0.4953208333333333    0.7546791666666667    1.                    0.8
O        0.0046791666666667    0.7453208333333333    0.0046791666666667    1.                    0.8
O        0.5046791666666667    0.7453208333333333    0.5046791666666667    1.                    0.8
O        0.4953208333333333    0.2546791666666667    0.4953208333333333    1.                    0.8
O        0.0046791666666667    0.0046791666666667    0.7453208333333333    1.                    0.8
O        0.5046791666666667    0.5046791666666667    0.7453208333333333    1.                    0.8
O        0.2453208333333333    0.7453208333333333    0.2453208333333333    1.                    0.8
O        0.7453208333333333    0.2453208333333333    0.2453208333333333    1.                    0.8
O        0.2453208333333333    0.2453208333333333    0.7453208333333333    1.                    0.8
```

Figure 7: PHASE.cel file example..

Line 1: Pearson symbol or space group number (string, length=4), cell parameters a, b, c, α, β, γ, number of *distinct atomic species*.
Line 2: For each species, count of atomic sites. Arbitrary order.
Line 3: For each species, atomic number Z, same order as line 2 above.

Following lines: For each atomic species, a group of lines containing chemical symbol, x, y, z fractional (cell) coordinates, *sof* and B (crystallographic Debye-Waller coefficient). One line for each atomic site. Atoms of the same species are expected to have the same chemical symbol (strictly required) and of course the same *sof* and B. In this example, there are three species, represented respectively by 8, 16, 32 atomic sites; for the first two species, $Z = 26$ (both are Fe atoms) and for the third, $Z = 8$ (oxygen). Following there are 8 lines with the symbol, coordinates, *sof* and B of the 8 atoms of the first species; 16 lines with the same information for the second species and 32 lines likewise for the third species.

7.2 PHASE.cely

Input file (alternative to .cel) for most programs used to generate a file of interatomic sampled distances (*.smp) for nanocrystals of any shape. Especially thought for molecular compounds.

7.2.1 PHASE.cely example

```
ZnbpbOH tP34 full - OH
CELL        13.2795  13.2795  6.98214      90.000000      90.000000      90.000000
NSUBUNITS          2
NSUBUNIT_MEMBERS  2  2
SUBUNIT_LOOP:
  SUBU 1                          ! subunit counter
  SCALABLE                        ! subunit keyword
  1                               ! number atomic species
  SPECIES_LOOP: !  i_at, z_at, n_at, xn_at, AvOccup, B; xn_at = n_at * Occup; 0<Occup<1
   1 30  1  1.0 1.00 2.0
     ATOM_LOOP: C
!    Sym      x                       y                       z          Occ (optional,not read)  B (optional,not read)
     Zn 1     0.                      0.                      0.25
     MEMBERS_LOOP:
       MMB 1                                ! member counter
       TRC  0.0 0.0 0.0                     ! TRC means the translation vector in crystal coord. is given
       RTE  0.0 0.0 0.0                     ! RTE means Euler angles are given
       MMB 2                                ! member counter
       TRC  0.0 0.0 0.5                     ! TRC means the translation vector in crystal coord. is given
       RTE  90.0 0.0 0.0           ! RTE means Euler angles (degrees) are given
  SUBU 2                          ! subunit counter
  SCALABLE                        ! subunit keyword
  2                               ! number atomic species
  SPECIES_LOOP: !  i_at, z_at, n_at, xn_at, AvOccup, B; xn_at = n_at * Occup; 0<Occup<1
   2  6  12 12.0 1.00 4.0
   3  7  4   4.0 1.00 4.0
     ATOM_LOOP: C
!    Sym      x                       y                       z          Occ (optional,not read)  B (optional,not read)
      C   2  0.6046726  0           0
      C   2  0.7138635  0           0
      C   2  0.7740606  0           0.1575825
      C   2  0.7740606  0          -0.1575825
      C   2  0.5523363  0.06779394  0.1144517
      C   2  0.4476637  0.06779394  0.1144517
      C   2  0.5523363 -0.06779394 -0.1144517
      C   2  0.4476637 -0.06779394 -0.1144517
      C   2  0.3953274  0           0
      C   2  0.2861365  0           0
      C   2  0.2259394  0           0.1575825
      C   2  0.2259394  0          -0.1575825
      N   3  0.8714616  0           0.09739134
      N   3  0.8714616  0          -0.09739134
      N   3  0.1285384  0           0.09739134
      N   3  0.1285384  0          -0.09739134
     MEMBERS_LOOP:
       MMB 1                                ! member counter
       TRC  0.0 0.0 0.0                     ! TRC means the translation vector in crystal coord. is given
       RTE  0.0 0.0 0.0                     ! RTE means Euler angles (degrees) are given
       MMB 2                                ! member counter
       TRC  0.0 0.0 0.5                     ! TRC means the translation vector in crystal coord. is given
       RTE  90.0 0.0 0.0                    ! RTE means Euler angles are given
```

Figure 8: PHASE.cel file example.

Lines - or parts thereof - starting with "!" are comments and will be ignored hereafter.

Line 1: Descriptive or simply poetic string (title)

Line 2: Flagword CELL; cell parameters a, b, c, α, β, γ

Line 3: Flagword NSUBUNITS; number of different basic subunits (*e.g.* molecules, ...) that constitute the unit cell, and each of them can be present in many translated and rotated instances in one unit cell.

Line 4: Flagword NSUBUNIT_MEMBERS; number of members (instances, copies, ...) present for each subunit in the unit cell.

Line 5: Flagword SUBUNIT_LOOP: marks the start of a loop over the different subunits, whose number is given above (line 3). A following block of lines describe the first subunit, then subsequently similar blocks describe all given subunits. We now number the lines in the first block as roman uppercase numbers.

Line I: Flagword SUBU; subunit counter

Line II: Flagword SCALABLE (optional); defines if the subunit in question is rigid or can at least be expanded.

Line III: No flagword. number of atomic species in the current subunit.

Line IV: Flagword SPECIES_LOOP: marks the start of a loop over the different atomic species in the current subunit. A following block of lines (one for each atomic species) describe their characteristics.

After IV: No flagword, just six numbers (counter of atomic species, atomic number, number of sites, sum of all relevant *sof*'s, average *sof*, crystallographic Debye-Waller coefficient B.

Next: Flagword ATOM_LOOP: followed by indicator character (possible values C or A). It marks the start of a loop over the different atomic sites in the current subunit. A following block of lines (one for each atomic site) describe their coordinates, species, *sof*,... C ("Cell") means that the atomic coordinates will be given as fractional cell coordinates; A ("Absolute") means that the atomic coordinates will be given with respect to a chosen Cartesian frame, in Ångstroem

Next lines: No flagword, just for each atom of the current subunit: chemical symbol; species index [IMPORTANT: the species index tells which species it is within the GLOBAL atomic species' list, not just in the current subunit's list!!!]; fractional (C) or absolute (A) coordinates x, y, z.

Next: Flagword MEMBERS_LOOP: marks the start of a loop over the different members (instances, copies, ...) of the current subunit present for each subunit in the unit cell. Following, for each member, it will be specified a translation and a rotation that yield the actual coordinates of the member with respect to the coordinates given above.

Next lines: Grouped in 3-lines blocks, one for each member, as follows:

1 of three: Flagword MMB followed by a counter (of the current member);

2 of three: Flagword TRC or TRA: translation vector, three real numbers, in fractional (TRC) or absolute (TRA) coordinates, as above.

3 of three: Flagword RTE or RTM: rotation Euler angles, in degrees (case RTE), three real numbers; or rotation matrix (case RTM), nine real numbers, line after line.

8 Secondary input files - .ini files

This section contains a detailed description of the actual input files (*.ini) for the various MK_* programs contained in the CLaUDe suite. These files are automatically generated by the DB_PHA_CLU program.

8.1 General Info

1 FILE TYPE: form=formatted, access=sequential, recordlength=256 (max.).

2 READING METHOD: each record is read into a 256-places character variable rline.

3 COMMENTS: marked by a '!' or a > in the first character of rline.

4 MANDATORY/OPTIONAL FLAG: a symbol ! or (M) specifies that a given entry is always mandatory; * that it is always optional; ? that it is conditionally optional (depending on the value of other entries, as specified).

5 FILES AND PATHNAMES: when we refer to a file name, unless otherwise specified, to the complete (pathname-inclusive) name, as and when required by the operating system

6 SECTIONS: The files *.ini contain only one section.

8.2 mkcell.ini

Input file for the MK_CELL program used to generate the unit cell file (given the filename.pha file, the output will be filename.cel).

buffer content	
filename.pha the .pha filename.	!
filename.grp the .grp filename.	!
# Integer number. The number of atomic 'species' in the .pha file. Species means atoms with the same chemical symbol, occupancy and Debye Waller factor.	!
#␣#␣# Three real numbers (crystal coordinates). Origin shift to be considered when building up NCs.	!

Continues...

`buffer` content

constr␣string *
Single character string, may be **P** or **S**, for disabling or enabling the surface broadening option, respectively.

string !
Pearson symbol.

8.2.1 `mkcell.ini` example

```
magnetite.pha
SPG_grp/SG_Nr_227_o1.grp
    3
      0.000000       0.000000       0.000000
constr  P
cF56
```

Figure 9: mkcell.ini file.

8.3 clumk.ini

Input file for the MK_RODS program used to generate files of the Cartesian coordinates with extension .xyz of PAR/CYL/HEX shaped nanoparticles.

`buffer` content

filename.cel !
The .cel filename.

string␣#␣# !
Two cases:

1 **N** followed by two integer numbers: the maximum number of shells used to build the nanoparticles in the *ab*-plane and along the *c*-axis, respectively.

2 **D** followed by two real numbers: the maximum diameter and length of the nanoparticles in the *ab*-plane and along the *c*-axis, respectively.

Continues...

buffer content

SHAPE !
See 2.3.

8.3.1 clumk.ini example

```
magnetite.cel
N   12  12
PAR
```

Figure 10: clumk.ini file.

8.4 clumkQ.ini

Input file for the MK_LAYER_GEN program used to generate a database (*.smp files) of PAR/CYL/HEX shaped nanoparticles.

buffer content

filename.cel !
The .cel filename.

string␣#␣# !
Two cases:

1 **N** followed by two integer numbers: the maximum number of shells used to build the nanoparticles in the *ab*-plane and along the *c*-axis, respectively.

2 **D** followed by two real numbers: the maximum diameter and length of the nanoparticles in the *ab*-plane and along the *c*-axis, respectively.

SHAPE !
See 2.3.

TODO␣**TODO**
The variable **TODO** content is described in Sec. 2.3 !

PARA␣**PARACRYSTALLINITY MODEL**
The variable **PARACRYSTALLINITY MODEL** value is described in Sec. 2.3 *

Continues...

buffer content	
OCC1␣**OCCUPANCY** The variable **OCCUPANCY** value is described in Sec. 2.3	*
SAMP␣**SAMPLING**␣**LAMBDA**␣**2**$\theta_{\max}$ The variables values are described in Sec. 2.4.	!

8.4.1 clumkQ.ini example

```
magnetite.cel
N  12 12
PAR
TODO all_clusters
OCC1  y
SAMP one 0.62001905 120.00000
```

Figure 11: clumkQ.ini file.

8.5 clumkS.ini

Input file for the MK_SPHERE program used to generate files of the Cartesian coordinates with extension .xyz of SPH shaped nanoparticles.

buffer content	
filename.cel The .cel filename.	!
D␣**#** real number. Maximum diameter of the nanoparticle.	!
SHAPE See 2.3.	!

8.6 sphmkQ.ini

Input file for the MK_BALL program used to generate a database (*.smp files) of SPH shaped nanoparticles.

buffer content

filename.xyz !
The phase filename (containing the list of cartesian coordinates).

string␣#␣# !
Two cases:

1 **N** followed by one integer number: the maximum number of shells used to build the nanoparticles.

2 **D** followed by one real number: the maximum diameter of the nanoparticles.

SHAPE !
See 2.3.

TODO␣**TODO**
The variable **TODO** content is described in Sec. 2.3 !

PARA␣**PARACRYSTALLINITY MODEL**
The variable **PARACRYSTALLINITY MODEL** value is described in Sec. 2.3 *

OCC1␣**OCCUPANCY**
The variable **OCCUPANCY** value is described in Sec. 2.3 *

XYZ?␣**XYZ**
The variable **XYZ** value is described in Sec. 2.3 *

SAMP␣**SAMPLING␣LAMBDA␣2**$\theta_{\max}$
The variables values are described in Sec. 2.4. !

8.6.1 sphmkQ.ini example

```
magnetite.cel
D  6.00000
SPH
TODO largest_only
OCC1  y
XYZ?  y
SAMP one 1.00000000 120.00000
```

Figure 12: sphmkQ.ini file.

8.7 qbemkQ.ini

Input file for the MK_QBE program used to generate a database (*.smp files) of QBE shaped nanoparticles.

buffer content	
filename.cel The .cel filename.	!
string␣#␣# Two cases: 1 **N** followed by one integer number: the maximum number of shells used to build the nanoparticles. 2 **D** followed by one real number: the maximum side of the nanoparticles.	!
SHAPE See 2.3.	!
TODO␣**TODO** The variable **TODO** content is described in Sec. 2.3	!
PARA␣**PARACRYSTALLINITY MODEL** The variable **PARACRYSTALLINITY MODEL** value is described in Sec. 2.3	*
OCC1␣**OCCUPANCY** The variable **OCCUPANCY** value is described in Sec. 2.3	*
XYZ?␣**XYZ** The variable **XYZ** value is described in Sec. 2.3	*
SAMP␣**SAMPLING␣LAMBDA␣2**$\theta_{\max}$ The variables values are described in Sec. 2.4.	!

8.7.1 qbemkQ.ini example

```
magnetite.cel
D  6.00000
QBE
TODO largest_only
OCC1  y
XYZ?  y
SAMP one 1.00000000 120.00000
```

Figure 13: qbemkQ.ini file.

8.8 setmkQ.ini

Input file for the MK_LATSET program used to generate a (single-file) database (*.smp files) of SET shaped nanoparticles. SET is a special case where an atomic cluster (a unit cell, a molecule, ...) is present as a set of multiple identical copies, which are translated with respect to each other by a set of (separately given) translation vectors. The latter may or may not be a subset of a periodic lattice. It is useful to simulate special cases of defective crystals.

---

buffer content

---

filename.cel !

The .cel filename. This contains the atomic coordinates of the cluster that is translated around in space by the translation vectors set. The format is the same as an usual .cel file (see Sec. 7.1); note that the unit cell parameters (first row, Sec. 7.1.1) are now superfluous and their values should be set to $a, b, c, \alpha, \beta, \gamma$ = `1.0 1.0 1.0 90.0 90.0 90.0`. Accordingly, the atomic coordinates x, y, z should be given in Ångstrœm, in a fixed Cartesian frame.

filename set coordinates !

The name of the file containing the list of translations. File ending is free. The format is simply four columns of real numbers, each row containing the coordinates x, y, z and the probability o of each translation vector. The atomic coordinates x, y, z should be given in Ångstrœm, in a fixed Cartesian frame. Several thousands of translations may be contained in such a file (after a certain number you'll run out of RAM, however). Example with only 5 translations:

```
 0.000   0.000   0.000 0.850
 5.431   0.000   0.000 1.000
10.862   0.000   0.000 1.000
10.862   5.431   0.000 1.000
10.862  10.862   0.000 0.850
```

This shows five translation, describing an L-shaped five-cells nanocrystal (!), whose extremes are slightly less likely to be occupied (85%) than the central ones (100%). This file could be called, for instance, `ell_shaped.trcoo` .

string␣#␣# !

This input string is dead, kept for compatibility. Its value should simply be set to

```
N  1  1
```

Continues...

buffer content
SHAPE ! The value must be SET
SAMP␣**SAMPLING␣LAMBDA␣2**$\theta_{\max}$ The variables values are described in Sec. 2.4. !

8.8.1 setmkQ.ini example

```
calcium_hypozedite.cel
ell_shaped.trcoo
N    1 1
SET
SAMP one   0.6199d0 120.d0
```

Figure 14: setmkQ.ini file.

8.9 molmkd.ini

Input file for the MK_MOLEC program used to generate a file of interatomic sampled distances (*.smp) from a cartesian coordinate files of any phase.

---

buffer content

---

filename.cel !
The .xyz filename, containing the list of cartesian coordinates.

SAMPLING␣LAMBDA␣2$\theta_{\max}$!
The variables values are described in Sec. 2.4.

TOL. MIN. DIST. *
Real value. a (small) length, in Ångstrœm, representing a tolerance margin on the interatomic distances, so that two atoms that are "too close" by a small amount, which is probably due to some small errors in the coordinates' source, are NOT considered "too close" by the distance generation algorithm.

DENS *
Flagword dens followed by the mass density of the substance.

8.9.1 molmkd.ini example

```
sio2_amorphous.xyz
one  0.49592 150.0
0.20000000
dens 2.138382221476936
```

Figure 15: molmkd.ini file.