

DEBUSSY: a Debye user system for nanocrystalline materials

Antonio Cervellino, a Cinzia Giannini b and Antonietta Guagliardi b,c,*

aSwiss Light Source, Paul Scherrer Institute, CH-5232 Villigen, Switzerland, bIstituto di Cristallografia, CNR, Via G. Amendola 122/O, I-70126 Bari, Italy, and cUniversità dell’Insubria, Dipartimento di Scienze Chimiche e Ambientali, Via Valleggio 11, I-22100 Como, Italy. Correspondence e-mail: antonella.guagliardi@ic.cnr.it

DEBUSSY is a new free open-source package, written in Fortran95 and devoted to the application of the Debye function analysis (DFA) of powder diffraction data from nanocrystalline, defective and/or non-periodic materials through the use of sampled interatomic distance databases. The suite includes a main program, taking the name of the package, DEBUSSY, and dealing with the DFA of X-ray, neutron and electron experimental data, and a suite of 11 programs, named CLAUDE, enabling users to create their own databases for nanosized crystalline materials, starting from the list of space-group generators and the asymmetric unit content. A new implementation of the Debye formula is adopted in DEBUSSY, which makes the approach fast enough to deal with the pattern calculation of hundreds of nanocrystals, to sum up their contributions to the total pattern and to perform iterative algorithms for optimizing the parameters of the pattern model. The package strategy uses the sampled-distance database(s) created previously by CLAUDE and combines, for any phase, a log-normal or a bivariate log-normal function to deal with the sample-size distribution; four different functions are implemented to manage possible lattice expansions/contractions as a function of crystal size. A number of output ASCII files are produced to supply some statistics and data suitable for graphical use. The databases of sampled interatomic non-dimensional distances for cuboctahedral, decahedral and icosahedral structure types, suitable for dealing with noble metal nanoparticles, are also available.

1. Introduction

The Debye function (DF) (Debye, 1915; Cervellino et al., 2003), like the more popular pair distribution function (PDF) (Debye & Menke, 1930; Egami & Billinge, 2003; Juhás et al., 2006), is a total scattering approach which is nowadays acknowledged to be more effective than standard Rietveld-like methods (Rietveld, 1969; Cheary & Coelho, 1992) for extracting relevant (micro)structural details of nanocrystalline systems from their powder diffraction data (Hall, 2000; Hall et al., 2000; Kaszur, 2000; Palosz et al., 2000; Cervellino et al., 2004; Cozzoli et al., 2006; Guagliardi et al., 2010). The short length scale of the domains, sometimes coupled with structural defects and/or distortions caused by surface effects, results in some peculiar features in the diffraction pattern (Warren, 1990; Neder & Proffen, 2008; Neder, 2010). Conventional powder diffraction approaches, mainly developed for microcrystalline specimens and relying strongly on instrumental broadening as the dominant contribution to the experimental pattern, can fail in modelling such sample effects or at best reach this goal through a phenomenological model which allows the indirect estimation of physical parameters. For both isotropically and anisotropically shaped nanoparticles with very small coherent domains and/or a high defect concentration, the instrumental broadening usually becomes negligible with respect to the finite size, shape and structure-based sample contributions. Furthermore, domains in real samples are likely distributed in size or shape along one (the isotropic shape case), two or even three growth directions.

These sample features can be described through a physically based model within the Debye function approach as it works in direct space. In addition, unlike the PDF, the DF does not transform the experimental data, thus keeping the advantage of simultaneously dealing with both reciprocal and direct spaces. In reciprocal space, the experimental data can (in principle) be almost exactly reproduced by the formula, provided that extra-sample contributions to the background, such as air and sample-holder scattering, are properly accounted for and as long as a number of requirements are achieved in direct space. These are (within the averaged-volume approximation of the powder diffraction technique) as follows: (a) a reliable structural model (ordered, disordered or defective); (b) a suitable model for the shape and size of nanocrystals (NCs) and their possible distributions, according to a physically and chemically sound description.

The main advantage of using the Debye formula when investigating short-range ordered materials such as nanocrystalline compounds is the ability to model simultaneously both Bragg (if any) and diffuse sample scattering, all arising from the distribution of the interatomic distances within the sample and not requiring periodicity and order.

Considering the general case of a particle containing $N_a$ atomic species and $N_b$ atoms of the $s$th species ($s = 1 \ldots N_b$), each atom of the same species is identified by a mean position vector $\mathbf{r}_j$ ($j = 1 \ldots N_s$), a mean occupancy $o_s$, a scattering length $b_s$ and an isotropic Debye–Waller thermal factor $T_D$. The diffracted intensity is obtained...
by spherically averaging over all possible orientations of the particle [or of the scattering vector \( q = |\mathbf{q}| = 2\sin(\theta)/\lambda \), where \( 2\theta \) is the diffraction angle and \( \lambda \) the wavelength of the incident radiation],

\[
I(q) = \sum_{j=1}^{N_v} N_{v}|\mathbf{b}_j|^2 \quad \text{term (1)}
\]

\[
+ 2 \sum_{s=1}^{N_s} N_s^2 b_j^2 T_s^2 \sum_{j=1}^{N_v} \left( 2\pi g j_{k,s} \right) \quad \text{term (2)}
\]

\[
+ 2 \sum_{s=1}^{N_s} N_s^2 \delta_{s k} N_v \sum_{j=1}^{N_v} \left( 2\pi g j_{k,s} \right) \quad \text{term (3)}
\]

where \( \sin(x) = \sin(x)/x \). The equation has been split into three terms to highlight the contributions of the zero distances between an atom and itself (term 1) and of the nonzero interatomic distances (term 2) and of the zero distances between an atom and all the tricks mentioned above are now introduced in the suite description: purpose and strategy

Although the relevance of this formula has been clear since it was presented in 1915, the Debye function approach has suffered from low popularity over the years because of the unfeasible computational times that are intrinsic to its formulation (the number of interatomic distances increases with the square of the NC size for ordered materials, and with up to its sixth power for disordered ones). This aspect has remained an unsolved drawback for the application of Debye function analysis (DFA) to real cases until recently, so that, despite the progress in computing resources over recent decades, its use has been limited mostly to the calculation of single simulations (Neder & Proffen, 2008; Oddershede et al., 2008), discouraging the development of dedicated software packages. A very up-to-date and promising way of dealing with the computing problems of the DF seems to rely on the use of graphics processing units (Gelisio et al., 2010). Based on a different and original approach, many tricks have been suggested recently in order to turn the DF method into an efficient and viable structural and microstructural data analysis tool. These tricks are now systematically organized and implemented in the suite of programs presented here. Among them, besides grouping the same distances in a single term and evaluating its multiplicity, we cite the following: (i) exploiting the crystal symmetry in order to reduce the number of interatomic distances (Grover & McKenzie, 2001; Cervellino & Guagliardi, 2010; Thomas, 2010); (ii) sampling the NC interatomic distances (Hall & Monot, 1991; Cervellino et al., 2006), so that the number of terms in the DF is reduced by orders of magnitude without losing accuracy in the pattern calculation; (iii) using the recursive properties of Chebyshev polynomials of the second kind for implementing a fast modified formula for calculating the diffraction profile, taking advantage of the constant step used for encoding the sampled distances (Navaza, 2002; Cervellino et al., 2006; Cervellino & Guagliardi, 2010).

The first case of a fast Debye formula implementation exploiting the trick of a previously computed sampled-distances database dates back to the early 2000s (Cervellino et al., 2004), when the analysis of synchrotron data from mixtures of face-centred cubic (f.c.c.) and multiple-twinned gold nanoparticles (icosahedral and decahedral) made it possible to extract information about the structure types, their concentration and their size distribution. The same approach has been practically applied and all the tricks mentioned above are now introduced in the DEBUSSY package, the goals and general features of which are described in the next section.

2. Suite description: purpose and strategy

The DEBUSSY package is mainly aimed at the DFA of nanocrystalline materials. The first release described here is currently able to deal with three types of (texture-free) nanostructures and their relative size/shape distributions. The first type (case 1) includes monoatomic f.c.c. crystalline (cuboctahedral) and f.c.c.-derived noncrystallographic multiple-twinned particles (decahedral and icosahedral structures), typically found in very small clusters of noble metal samples (Au, Pd, Ag etc.) (Ino & Ogawa, 1967; Ino, 1969; Vogel, 1998; Kaszkar, 2000; Cervellino et al., 2003). The other two types of nanostructures refer to defect-free crystalline materials featured in unit- (case 2) or bivariate (case 3) families of particles of increasing size. Cases (2) and (3) are suitable for describing many materials, such as metals, oxides, semiconductors etc., of spherical, rod and plate-like shapes. The general strategy of the DEBUSSY’s suite is to carry out the analysis according to a double-step approach. The first step concludes with the creation of a database containing the sampled interatomic distances of NCs of increasing size. The second step performs the DFA of experimental diffraction data, collected on both pure phases and mixtures, using the previously created database(s). For case (1), the first step is skipped, as the databases of sampled interatomic non-dimensional distances for cubocta-, icosahedral- and decahedral clusters up to about 50 nm in diameter are directly provided to users. For cases (2) and (3), the suite enables users to build their own databases, executing a suite of three or five programs (step-by-step procedure) or of five or seven programs (automatic procedure), depending on the selected shape. The unit cell is used as the NC building block; spherical and rod/plate-like shapes can be chosen to build up the nanocrystal family.

Spherical NCs are grown according to a concentric shell model. Discrete clusters of linearly increasing radii are obtained by applying the relationship

\[
r_k = r_0 + k\Delta R,
\]

where \( k \geq 0 \), \( r_0 = \Delta R \) and \( \Delta R \) is a function of the cube root of the unit-cell volume \( (a) \). An additional feasible requirement is that every \( k \)th added shell contains an integer number of primitive unit cells, so that the compound stoichiometry is preserved. For primitive cubic \((5)\), f.c.c. \((6)\) and body-centred cubic \((b.c.c.) \((7)\)\) lattices, both requirements are achieved by choosing

\[
\Delta R = a \left[ \frac{3}{2\pi} \right]^{1/3},
\]

\[
\Delta R = a \left[ \frac{6}{\pi} \right]^{1/3},
\]

\[
\Delta R = a \left[ \frac{3}{2\pi} \right]^{1/3}.
\]

Once the radii are fixed, every \( k \)th cluster is built up by selecting all lattice nodes that are within \( r_k \) of an appropriately chosen centre. For every chosen lattice node, the content of a unit cell is added. Suitable provisions for near-surface nodes or atoms may be necessary.

Rod- and plate-like shapes are obtained by combining the set of Patterson vectors \( P = \{\pm \mathbf{d}_{sl} = \pm(\mathbf{r}_i - \mathbf{r}_j), f > k = 1 \ldots N_c \} \) (\( N_c \) is the number of atoms in the unit cell) with the set of lattice vector translations of the unit cell \( \mathbf{L} = n_1 \mathbf{a} + n_2 \mathbf{b} + n_3 \mathbf{c} \) (\( n_1, n_2, n_3 \) are integers, and \( \mathbf{a}, \mathbf{b} \) and \( \mathbf{c} \) are the unit-cell defining vectors), according to a layer-by-layer construction (see Cervellino & Guagliardi, 2010, for more details) and under the assumption, which holds for biaxial systems, of two kinetically different growth directions, one along the three-, four- or sixfold crystallographic axis and the second in the orthogonal plane.
The suite (containing a total of 11 programs) enabling users to build their own databases is named **CLAUDE** (the acronym stands for crystalline layered user database).

The second step of the analysis is performed by the program **DEBUSSY** on single- or multiple-phase specimens, trying to fit one or more experimental data sets. **DEBUSSY** uses the previously created database(s) and carries out the fast calculation of the diffraction profile of each NC that contributes to the total pattern, according to a suitable and adjustable size-distribution function. Different options are also available to deal with lattice expansions or contractions variable with the crystal size. Both Chebyshev polynomials and experimentally collected data can be used to manage the background component. Synchrotron or laboratory X-ray, neutron and electron beams can be used to collect data; for all radiation types, the scattering amplitudes/factors of atomic species are encoded into the program.

### 2.1. **CLAUDE** suite

The tasks performed by **CLAUDE** are described graphically in Fig. 1, along with the names of the programs: (a) calculating the symmetry operators of the space group (**MK_GROUP**) starting from the corresponding list of generators, available at the Bilbao Crystallographic Server (http://www.cryst.ehu.es/); (b) expanding the asymmetric unit into the entire atomic cell content (**MK_CELL**); (c) building up uni- (**MK_SPHERE**) or bivariate (**MK_RODS**) families of atomistic models of nanocrystals of different shapes using the unit cell as the ‘building block’; (d) calculating the interatomic distances of each NC (**MK_DIST**); (e) sampling the interatomic distances and encoding these values in a suitable database (**MK_SAMP**).

This sequence of programs creates three types of databases (DBs): (i) the DB collecting the atomistic models of all NCs and including files of the Cartesian coordinates with extension .xyz (output of **MK_SPHERE** and **MK_RODS** programs). Plot files for **Jmol**, an open-source Java viewer for chemical structures in three dimensions (http://www.jmol.org/), are also created; (ii) the DB collecting the files of the interatomic distances of each NC with extension .dist; (iii) the DB collecting the files of the sampled interatomic distances of each NC with extension .smp.

For a bivariate size/shape family, a much faster program (**MK_LAYERS**) is supplied, which creates only the .smp database. Using **MK_LAYERS** is highly recommended in cases of large families, to save computing time and disk space. The entire sequence works with fixed-name files to input and transfer information throughout the suite. This step-by-step **CLAUDE** execution is described in detail in the manual, which is delivered as a .pdf file.

However, managing all the files and steps of the procedure (which also includes the creation of suitable folders) can be quite a tedious operation. For this reason, an automatic run (depicted in Fig. 2) has been organized as follows: (a) all the input information is collected in two files of fixed name and format, **DB_Phase_Info.inp** and **DB_CLU_Info.inp**, grouping all details related to the phase and the NC family that the user wishes to create, respectively; (b) three additional programs have been introduced as utilities for managing the input, creating the files used for the transfer of information and launching the program of the appropriate step according to the flow diagram depicted in Fig. 1; (c) the automatic run is dealt with through a shell script.

It may be useful to supply some information about the computing time and disk space required by the automatic **CLAUDE** procedure. As an example, we give here the values for a family of rod-shaped TiO$_2$ NCs (anatase polymorph) with sizes of up to about 9 nm in the $ab$ plane and 19 nm along the $c$ axis. Running the **MK_LAYERS** program for creating the final database (.smp), the whole suite took 15 min on a Mac OS X 10.6.4 machine using a 3.2 GHz Intel Core i3

---

**Figure 1**
A flow chart of the **CLAUDE** suite, illustrating how users can build their own databases for crystalline materials according to a step-by-step procedure.

**Figure 2**
A flow chart of the **CLAUDE** suite, illustrating how users can build their own databases for crystalline materials according to the automatic procedure.
processor, and 20 min on a Linux machine using a 3.0 GHz Intel Core Duo processor. The same run required a disk space of 22.7 MB to store the .amp database, while 1.8 GB were necessary to store all three databases (.xyz, .dist and .amp) described above.

Once the final database of sampled distances is available, an additional program (MK_PATTERN) enables the calculation of the powder diffraction simulation from a single atomic cluster, a useful utility for checking the correctness of the final output before running the DEBussy program. An example of MK_PATTERN output is shown in Fig. 3 for titanium dioxide, anatase polymorph, for a spherical (bottom) and a rod-like (top) NC.

2.2. DEBussy program

DEBussy performs DFA of powder diffraction patterns collected on nanosized crystalline and periodic but partly disordered materials. A schematic view of the program flow chart is given in Fig. 4. Input requirements are as follows: (i) the experimental data (one or more files in the case of multiple data sets) and the working conditions (wavelength, radiation, background info etc.); (ii) the sampled distance database path (one for each phase), identified by a suitable code to distinguish the different cases; (iii) starting values for the uni- (two parameters) or bivariate (five parameters) log-normal size-distribution function (Sampson & Siegel, 1985) and for size-dependent lattice expansion or contraction (up to three parameters).

All the information and instructions are provided through an input file, organized in three sections (data set, structure, simulation/refinement/output). Any extension is accepted for this file but we refer to it here using the extension .dwa to identify it. The program can be used in simulation or refinement mode. An external file (required extension .par) collects, for each structure, the starting value and the upper/lower limit of the parameters, the former to be used for simulations and the latter for refinement runs, along with user-managed refinement codes supplied through an additional file (required extension .ref). Simplex (Nelder & Mead, 1965) and simulated annealing (Metropolis et al., 1958) algorithms are implemented to optimize the starting model, and multiple refinement stages can be encoded in the .ref file.

Each DEBussy run creates (or overwrites) a number of ASCII files. Some of these are used to supply sample statistics (average particle size and dispersion for uni- and bivariate distributions, weight fractions) and data suitable for graphical use (calculated total and single-phase patterns; values of the size, number and mass distributions, mass, volume, distance deformation etc. for each nanoparticle), while others are used to save the parameters corresponding to the best solution of a refinement run, to be used as the input of a new execution. The user manual, available as a .pdf file, describes the main DEBussy inputs and outputs and provides some useful examples.

An example of the performance of DEBussy is presented here for a sol–gel-synthesized anatase powder and can be seen in Fig. 5, where the starting and final pattern models are matched to the experimental synchrotron data in the angular range 8–115°. The measurements were carried out at the PSI-SLS MS4-Powder beamline (Switzerland) using a Debye–Scherrer capillary transmission set up, an energy beam of 17 keV and a Mythen detector (Bergamaschi et al., 2010), allowing the full pattern to be collected in a few seconds. The following parameters were refined: the pair average/standard deviation of the bivariate log-normal size distributions along the two growth directions ($a_1$ and $s_1$ in the $ab$ plane, and $a_2$ and $s_2$ along the $c$ axis) and their correlation angle ($\psi$); the Ti occupancy factor; three coefficients that linearly correct the isotropic atomic Debye–Waller factor as a function of size along the two growth directions; scale and Chebyshev polynomial coefficients (12) modelling the background. Starting values (in terms of the number of stacked cells) of 7.0, 2.0, 7.0, 2.0 and 0.0 were refined to 8.525, 6.111, 9.0, 4.833 and –9.171 for $a_1$, $s_1$, $a_2$, $s_2$ and $\psi$, respectively. The Ti occupancy refined to an average value of 0.92, indicating the occurrence of vacancy defects, as
perviously reported in sol–gel-prepared anatase (Grey & Wilson, 2007). The refinement was performed following a multi-stage user-defined approach. The best-fit agreement indices are $R_w = 4.26$, goodness of fit = 3.30. The data analysis provided an average rod-shaped nanocrystal of 3.25 ± 1.79 nm in the $ab$ plane (diameter of the equivalent circle) and 8.05 ± 3.67 nm along the c axis.

3. Supported platforms and distribution

The DEBussy suite can presently run on Linux and Mac OS X platforms. Binary files for Linux 32-bit, Linux 64-bit, and Mac OS X 10.5 and 10.6 are available upon request. Binaries and source codes will soon be accessible at the dedicated URL http://debussy.sourceforge.net. For program distribution and any support for the suite, contact A. Guagliardi (antonella.guagliardi@ic.cnr.it) or A. Cervellino (antonio.cervellino@psi.ch).

Partial financial support by Fondazione Cariplo (project No. 2009-2446) is acknowledged. The X-ray powder diffraction data of TiO$_2$ were recorded at the MS4-Powder beamline of the SLS synchrotron, Villigen, Switzerland. The authors thank Dr Mario Oriani for his technical support.

References


