Published in American Mineralogist, 96, 437-444 (2011)

The WURM project – a freely available web-based repository of computed physical data for minerals

Razvan Caracas[1,2] and Ema Bobocioiu[2]

1. Centre National de la Recherche Scientifique, razvan.caracas@ens-lyon.fr
2. Laboratoire de Sciences de la Terre, Ecole Normale Supérieure de Lyon, 46 allée d'Italie, 69364 Lyon cedex 07

Abstract

The WURM project is a database of computed Raman and infrared spectra and other physical properties for minerals. The calculations are performed within the framework of the density-functional theory and the density-functional perturbation theory. The database is freely-available for teaching and research purposes and is presented in a web-based format, hosted on the http://www.wurm.info website. It provides the crystal structure, the parameters of the calculations, the dielectric properties, the Raman spectra with both peak positions and intensities and the infrared spectra with peak positions for minerals. It shows the atomic displacement patterns for all the zone-center vibrational modes and the associated Raman tensors. The web presentation is user friendly and highly oriented towards the end user, with a strong educational component in mind. A set of visualization tools ensures the observation of the crystal structure, the

vibrational pattern, and the different spectra. Further developments include elastic and optical properties of minerals.

Introduction

The last decades saw the rise of computational physics as a new branch of physics intermediate between the two traditional main branches: theoretical physics and experimental physics. This evolution was triggered by the exponential rise of the computational power, together with the development of performant algorithms and the larger availability of software implementations. Nowadays, numerical physics evolves into an investigation tool, which is fully complementary to experiments; the use of these techniques lately increased in the Earth sciences and the mineral physics, too.

Today theoretical results are usually in good agreement with experimental data (on the order of a few percent for interatomic distances and a few degrees on the interatomic angles). The vibrational spectra should be considered more carefully as they can differ as much as a few tens of wavenumbers from experimental data. Usually the shifts are consistent within chemical variations along a solid solution or mineralogical group. The relative intensities of the major Raman peaks are in good agreement with respect to experiment. However nowadays mineral identification cannot be done solely based on theoretical spectra. Nevertheless the latter can be used to understand and to interpret the measured spectra and to assist the identification based on experimental data. The first-principles calculations have another advantage, in that they offer the ability to easily study materials in thermodynamic and/or thermochemical conditions that are hard to reach in experiments. These make the computational methods an important and useful tool in modern science.

In parallel with this effort, gathering of physical data for minerals advanced at a fast pace on the experimental side. Efforts to synthesize this information and to classify it matured with the recent formation of several public databases, notably of Raman spectra: the RRUFF project (Downs 2006) of the University of Arizona, USA; the Raman Spectra Database of the Dipartimento di Scienza della Terra in Siena, Italy; the Raman Spectra Database of Minerals and Inorganic Materials (RASMIN) of the National Institute of Advanced Industrial Science and Technology, Japan; the Handbook of Mineral Raman Spectra of the ENS Lyon, France; the Mineral Raman Database of the University of Parma, Italy; the Database of Mineral Raman Spectra of the SFMC, France; etc.

However till now there were no similar efforts on the theoretical side. The WURM project comes to fill in this gap: the first database of computed physical properties for minerals that are freely available through the Internet, hosted on the http://www.wurm.info website. The WURM project offers a complementary numerical alternative to the RRUFF project. It is addressed to the community of mineralogists and crystallographers in a broad sense, both for research and teaching purposes.

Database overview

The aim of the database is to provide computed values for different physical properties of minerals: crystal structure, dielectric tensors, dynamical atomic charges, Raman spectra with both peak position and intensities, Raman tensors, frequencies and vibrational pattern for all zone-center modes.

By mineral we understand the definition adopted by the International Mineralogical Association (Nickel 1995): “a crystalline inorganic chemical which can be naturally found …” but we extend it to synthetic analogs to known minerals, planetary ices, aerosols and any other inorganic chemicals that are related to the mineral realm known or susceptible to naturally exist on Earth or in another place of the Solar System.

The WURM project has a similar philosophy as the RRUFF project, run by Robert Downs. All the information contained in the database is public, distributed under the Creative Commons license. The database is hosted at the http://www.wurm.info website. A series of javascripts and php scripts ensure a highly interactive way of displaying the information where the users are able to personalize the content they are viewing. The website is in a continuous up-date state with new minerals and new features added on a regular basis, following a flexible structure, detailed below. The website features an internal search engine, a full index of spectra as well as general information about the project and the data hosted.

Methodology

Three stages need to be successfully completed to add a new entry in the database: a decisional stage, a processing stage and a post-processing stage, as summarized in Figure 1.

In the decisional stage we select, because of current implementation restrictions, a new mineral that is not metallic or magnetic. We use its experimental structure, but also compute two theoretical structures: one at experimental density, i.e. experimental

volume, and one at theoretical 0GPa pressure. During these structural relaxations we minimize the energy, the residual forces on the atoms and the non-hydrostatic stresses on the unit cell. This stage is synthesized in the left part of the scheme in Figure 1.

In the next stage we perform the *ab initio* calculations using each of the three structures. The calculations are done using the density functional theory (Hohenberg and Kohn 1964; Kohn and Sham 1965) in the ABINIT implementation with planewaves and pseudopotentials (Gonze et al. 2002; 2005; 2009). The calculations take several steps, each yielding specific physical properties and/or energy or wavefunctions derivatives. Eventually we get the dynamical matrices, which are used to extract the entire vibrational information. The detailed procedure and the theory are described in the appendix. This stage is synthesized in the central part of scheme in Figure 1.

The complete dynamical matrices are then diagonalized in the third stage. We obtain the atomic charges and the dielectric tensors, and for each phonon its frequency, atomic eigendisplacements and Raman tensor. If there are optical phonons with imaginary frequencies, we have an indication about a possible phase transition at low temperatures. We then keep the results for a further analysis of the instabilities. If all the optical phonons have positive frequencies the structure is dynamically stable. Then we build an xml file combining the structural information and the parameters of the calculation with the dielectric and the vibrational properties. Eventually the xml file is uploaded on the WURM website and the latter updated.

Database content

In the present state of advancement of the database there are four main categories of information that can be accessed: the structural data, the key parameters of the calculation, the dielectric properties and the vibrational properties.

Structural information

For each mineral analyzed the database provides all the structural information relevant to the calculations: the lattice parameters and the angles of the (primitive) unit cell and the reduced coordinates of all atoms in the primitive unit cell. The content of the primitive unit cell is summarized. Orientative experimental values of the lattice parameters and unit cell angles are listed to provide a basis to assess the precision of the theoretical structure.

The structure can be visualized in the embedded Jmol applet (Jmol). The size of the supercell is variable and user-defined, with a default value of 2x2x2 unit cells.

Parameters of the calculations

This section covers the minimal information required for a user to assess the quality of the calculations and if needed to reproduce them. It includes the grid of **k**-points (Monkhorst and Pack 1976) and the kinetic energy cut-off (Payne et al. 1992). Together with the structural information these parameters can help rebuilding the input file.

A brief description of the pseudopotentials used in the calculations is also given.

Dielectric properties

The first set of perturbational properties is obtained from the derivatives of the energy with respect to electric fields. These yield the Born effective charges and the dielectric tensors.

The Born effective charges, also called dynamical charges, are tensors that correspond to the energy derivative with respect to atomic displacements and electric fields or, equivalently, to the change in atomic force due to an electric field. The sum of the Born effective charges of all nuclei in one cell must vanish, element by element, along each of the three directions of the space. The value of the Born effect charges is an indication about the amount of hybridization and charge transfer occurring during atomic vibrations: larger is the deviation of these charges with respect to the nominal charge values larger is the charge transfer between neighboring vibrating atoms.

The dielectric tensors represent the derivatives of the energy with respect to two electric fields. They also relate the induced polarization to the external electric field. The square root of the electronic component of the dielectric tensor gives the refractive index (Baroni et al. 2001; Gonze et al. 2005a; Caracas and Gonze 2010).

Spectroscopic properties

The vibrational part is the main focus of the database.

The Raman spectra for powders are plotted with both peak position and peak intensities. The main Raman peaks are labeled with their frequency (Figure 2) and

character, if determined. The user can finely tune the plot of the Raman spectra: vertical and horizontal zoom help visualizing regions of the spectrum. The default visualization is for spectra on powders with unpolarized lasers. The contributions from either parallel or perpendicular laser polarizations can be shown.

The phonon modes in the center of the Brillouin zone are listed with their frequencies, their characters and, for the Raman active modes, their Raman intensities. The transverse (TO) and the longitudinal (LO) components of all modes are listed, the latter ones along each of the three Cartesian directions. The cells in the four columns with the mode frequencies, i.e. the TO and the three LO components, are all clickable: at click an external browser window with the Jmol applet opens and the vibration can be visualized dynamically (Figure 3). The size of the vibrating structure can be tuned, with a default value of 2x2x2 supercell. Arrows are attached to the vibrating atoms.

Database structure

The results of the calculations are stored on the website in the form of xml files. This solution was found as the most reliable, most flexible and most appropriate for data exchange. The xml files have a tree-like structure that follows the different property types displayed on the website. This structure can be used as a model for storing data for other mineral properties.

The main branches of the xml files are: header, structure, parameters, dielectric, and dynamics. The header branch contains comment fields, the name of the mineral phase and its chemistry. The structure branch contains the unit cell information, its content, the list of atoms and the full description of the symmetry. Experimental data

about the unit cell parameters are also stored here. The parameters branch contains the details of the calculations and the information about the pseudopotentials. The dielectric branch contains the Born effective charge tensors, and the dielectric tensors, including the refractive index tensor, all with eigenvalues. The Born tensors are listed for all the atoms in the primitive unit cell, regardless of the symmetry.

The dynamics branch contains the phonon information. For each vibrational mode the current number, its character and the Raman/infrared activity are listed, followed by the TO and the three LO components. Each component sub-branch contains the mode frequency, its eigendisplacements, namely the atomic displacement along each of the three Cartesian directions, and the Raman tensors.

A glance at the internal machinery of the website

The list with the mineral data (stored in the xml files) is centralized in a main xml index file. This is accessed by the search engine.

Then a php script reads the structure and the content of the individual xml files and initializes the different internal variables necessary for building the webpages. The different actions initiated by the user activate hidden scripts and routines that process the raw data and generate the final images or pages. The data are passed between routines and applets through temporary ascii files. For example the crystal structure is stored as a matrix with the atomic positions. For the visualization of the structure, a javascript routine reads this matrix, the unit cell parameters and the angles and builds a temporary xyz file, which is then passed to the Jmol applet. For the visualizing of the phonon vibrations, the mode eigendisplacements are used together with the crystal structure data

to build the temporary the xyz files, which are then passed to the Jmol applets. The Raman intensities are built in a similar way. These scripts and routines run on the host server.

Further developments

We are continuously increasing the size of the database with new spectra at a rate of about a few each week. This rate is imposed by computing limitations and by variation of the computing requirements from one mineral to the other. New visualization tools and new mineral properties will be added in the foreseeable future.

The first main update will consist of displaying the infrared spectra. This will be followed by the dynamic visualization of the single-crystal spectra that will allow for observing the effect of the sample orientation.

In terms of mineral properties the first basis have been already put to include the elasticity. The computation of the elastic constants tensor is also done in framework of the DFPT. The elastic constants are expressed as the derivative of the energy with respect to strains. The terms used in the computation of the phonons are used to correct for the couplings with the atomic vibrations and the electric fields (Hamman et al., 2005).

Acknowledgements

We thank our generous sponsor Michael Scott whose donation made possible the realization of this project. We acknowledge help from Robert Downs who runs the experimental Raman RRUFF website on the http//www.rruff.info. We would like to

thank Geroge Rossman, James Kubicki and an anonymous reviewer for useful suggestions and discussions.

The calculations are performed using the ABINIT package for first-principles calculations based on density-functional theory in a planewave-pseudopotential implementation. The calculations are performed on our local resources, on the machines of the Pole de Simulation et Modelisation Numerique (PSMN) of the Ecole Normale Superieure de Lyon and on the SGI Altix machine JADE of the Centre Informatique National de l'Enseignement Superieur (CINES) through DARI grants stl2816.

The website was built in collaboration with Samuel Vernon Design.

If interested, there are several ways of how anybody can get involved and provide support to our efforts: donate manpower, money, computing time and expertise.

References

Baroni, S., and Resta, R. (1986), Ab initio calculation of the low-frequency Raman cross section in silicon, Phys. Rev. B, 33, 5969– 5971.

Baroni, S., Giannozzi, P., and Testa, A. (1987) Green's-function approach to linear response in solids. Phys. Rev. Lett. 58, 1861–1864.

Baroni, S., de Gironcoli, S., Dal Corso, A., and Giannozzi, P. (2001) Phonons and related crystal properties from density-functional perturbation theory. Rev. Mod. Phys., 73, 515–562.

Bellaiche, L., and Vanderbilt, D. (2000) Virtual crystal approximation revisited: application to dielectric and piezoelectric properties of perovskites. Phys. Rev. B 61, 7877–7882.

Caracas, R. (2007) Raman spectra and lattice dynamics of cubic gauche nitrogen, J. Chem. Phys. 127, 144510.

Caracas, R., and Cohen, R.E. (2006) Theoretical determination of the Raman spectra of $MgSiO_3$ perovskite and post-perovskite at high pressure, Geophys. Res. Lett., 33, L12S05, doi:10.1029/2006GL025736.

Caracas, R., and Cohen, R.E. (2007) Post-perovskite phase in selected sesquioxides from density-functional calculations. Phys. Rev. B 76, 184101.

Caracas, R., and Banigan, E. (2009) Elasticity and Raman and infrared spectra of $MgAl_2O_4$ spinel from density-functional perturbation theory. Phys. Earth Planet. Inter. 174, 113-121.

Caracas, R., and Gonze, X. (2010) Lattice Dynamics and Thermodynamical Properties. In Chaplot, S.L., Mittal, R., and Choudhury, N. (Eds.) Thermodynamic

Properties of Solids: Experiment and Modeling. Pp. 291-315. WILEY-VCH Verlag, Weinheim.

Cardona, M. (1982), Light Scattering in Solids II: Basic Concepts and Instrumentation, vol. 50, Topics in Applied Physics, edited by M. Cardona and G. Güntherodt, pp. 19– 168, Springer, New York.

Cohen, M.L. (2007) Quantum alchemy. In: Keinan, E., and Schechter, I. (Eds.), Chemistry for the 21st century. Pp. 247-270 Wiley-VCH.

Downs, R.T. (2006) The RRUFF project: an integrated study of ht echemistry, crystallography, Raman and infrared spectroscopy of minerals. Program and Abstracts of the 19th General Meeting of the International Mineralogical Association in Kobe, Japan, O03-13.

Ghosez, Ph., Michenaud, J.-P., and Gonze, X. (1998) Dynamical atomic charges: the case of ABO_3 compounds. Phys. Rev. B58, 6224-6240.

Giantomassi, M., Boeri, L., and Bachelet, G.B. (2005) Electrons and phonons in the ternary alloy $CaAl_{2-x}Si_x$ as a function of composition. Phys. Rev. B 72, 224512.

Gonze, X., and Vigneron, J.-P. (1989) Density-functional approach to nonlinear-response coefficients of solids. Phys. Rev. B49 13120–13128.

Gonze, X., Allan, D. C., and Teter, (1992) M. P. Dielectric tensor, effective charges and phonon in a-quartz by variational density-functional perturbation theory. Phys. Rev. Lett., 68, 3603–3606.

Gonze, X., and Lee C. (1997) Dynamical matrices, Born effective charges, dielectric permittivity tensors, and interatomic force constants from density-functional perturbation theory. Phys. Rev. B, 55, 10355–10368.

Gonze, X., Beuken, J.-M., Caracas, R., Detraux, F., Fuchs, M., Rignanese, G.-M., Sindic, L., Verstraete, M., Zerah, G., Jollet, F., Torrent, M., Roy, A., Mikami, M., Ghosez, Ph., Raty, J.-Y., and Allan, D. C. (2002) First-principle computation of material properties the ABINIT software project. Comput. Materials Science, 25, 478–492. [http//www.abinit.org]

Gonze, X., Rignanese, G.-M., Verstraete, M., Beuken, J.-M., Pouillon, Y., Caracas, R., Jollet, F., Torrent, M., Zerah, G., Mikami, M., Ghosez, Ph., Veithen, M., Raty, J.-Y., Olevano, V., Bruneval, F., Reining, L., Godby, R., Onida, G., Hamann, D. R., and Allan, D. C. (2005) A brief introduction to the ABINIT software package. Z. Kristallogr., 220, 558–562.

Gonze, X., Rignanese, G.-M., Caracas, R. (2005a) First-principles studies of the lattice dynamics of crystals, and related properties. Z. Kristallogr. 220, 458–472.

Gonze, X., Amadon, B., Anglade, P.-M., Beuken, J.-M., Bottin, F., Boulanger, P., Bruneval, F., Caliste, D., Caracas, R., M. Côté, Deutsch, T., Genovesi, L., Ghosez, Ph., Giantomassi, M., Goedecker, S., Hamann, D.R., Hermet, P., Jollet, F., Jomard, G., Leroux, S., Mancini, M., Mazevet, S., Oliveira, M.J.T., Onida, G., Pouillon, Y., Rangel, T., Rignanese, G.-M., Sangalli, D., Shaltaf, R., Torrent, M., Verstraete, M.J., Zerah G., and Zwanziger, J.W. (2009) ABINIT: First-principles approach to material and nanosystem properties, Comp. Phys. Commun 180, 2582-2615.

Hamann, D., Wu, X., Rabe, K.M., and Vanderbilt, D. (2005) Metric tensor formulation of strain in density-functional perturbation theory. Phys. Rev. B 71, 035117.

Hayes, W., and Loudon, R., (1978) Scattering of Light by Crystals. John Wiley, Hoboken, New Jersey, 113 pp.

Hermet, P., Goffinet, M., Kreisel, J., and Ghosez, Ph. (2007) Raman and infrared spectra of multiferroic bismuth ferrite from first-principles, Phys. Rev. B 75, 220102(R).

Hohenberg, P., and Kohn, W. (1964) Inhomogeneous electron gas. Phys. Rev. 136, B864-B871.

Jmol: an open-source Java viewer for chemical structures in 3D. http://www.jmol.org/

Kohn, W., and Sham, L. J. (1965) Self-consistent equations including exchange and correlation effects. Phys. Rev. 140, A1133–A1138.

Lazzeri, M., and Mauri, F. (2003), First-principles calculation of vibrational raman spectra in large systems: Signature of small rings in crystalline SiO2, Phys. Rev. Lett., 90, 36– 401.

Monkhorst, H.J., and J.D. Pack (1976) Special points for Brillouin-zone integrations. Phys. Rev. B13, 5188-5192.

Nickel, E.H. (1995) The definition of a mineral. Canad. Mineral. 33, 689-690.

Payne, M.C., M.P. Teter, D.C. Allan, T.A. Arias, and J.D. Joannopoulos (1992) Iterative minimization techniques for ab initio total-energy calculations: molecular dynamics and conjugate gradients. Rev. Mod. Phys., 64, 1045-1097.

Placzek, G. (1934), Rayleigh-Streuung und Raman-Effekt, in Handbuch der Radiologie, vol. 6, edited by G. Marx, pp. 205–374, Akademische, Frankfurt-Main, Germany.

Prosandeev, S. A., Waghmare, U., Levin, I., and Maslar, J. (2005) First-order Raman spectra of $AB'_{1/2}B''_{1/2}O_3$ double perovskites. Phys. Rev. B, 71, 214–307.

Scandolo, S., Giannozzi, P., Cavazzoni, C., de Gironcoli, S., Pasquarello, A., and Baroni, S. (2005) First-principles codes for computational crystallography in the Quantum-ESPRESSO package. Z. Kristallogr. 220 574–579. [http//www.pwscf.org]

Teter, M. (1993) Additional condition for transferability in pseudopotentials, Phys. Rev. B 48, 5031-5041.

Troullier, N., and Martins, J.L. (1991) Efficient pseudopotentials for plane-wave calculations. Phys Rev B 43, 1993-2006.

Veithen, M., Gonze, X., and Ghosez, Ph. (2005) Non-linear optical susceptibilities, Raman efficiencies and electrooptic tensors from first-principles density functional perturbation theory. Phys. Rev. B71, 125107.

Wang, Z., Zhao, Y., Zha, C.-s., Xue ,Q., Downs, R.T., Duan, R.G., Caracas, R., and Liao, X. (2008) X-Ray Induced Synthesis of 8H Diamond, Adv. Mater. 20, 3303-3307.

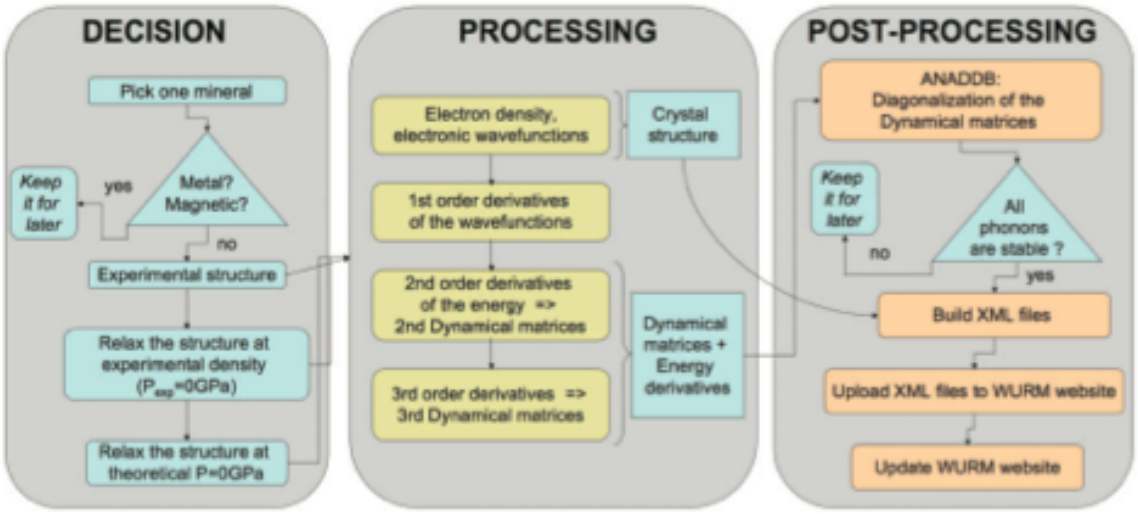

Figure 1. Schematic flowchart with all the different steps for obtaining a new entry in the WURM website. There are three main phases corresponding to the three main panels. Only insulating non-magnetic phases are considered for the time being, because of implementation limitations. The calculations are performed using the density-functional theory and the density-functional perturbation theory in the ABINIT implementation. A set of perl scripts transforms the output of the calculations to xml files that are uploaded on the WURM website.

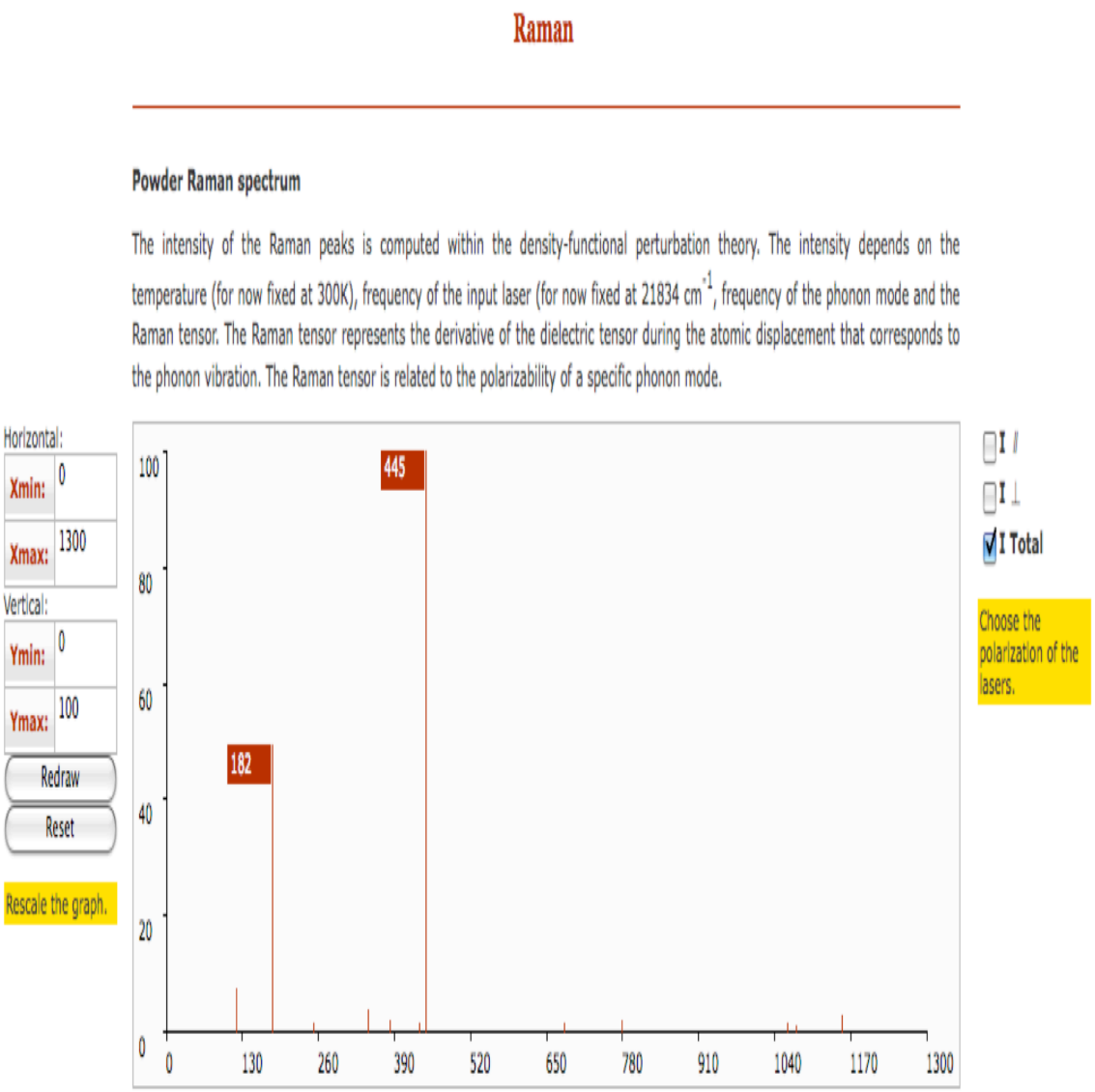

Figure 2. The Raman spectra are reported with both peak position and intensities. The user can finely tune the plots and see the powder spectra as with perpendicular and/or parallel lasers. The main peaks are labeled.

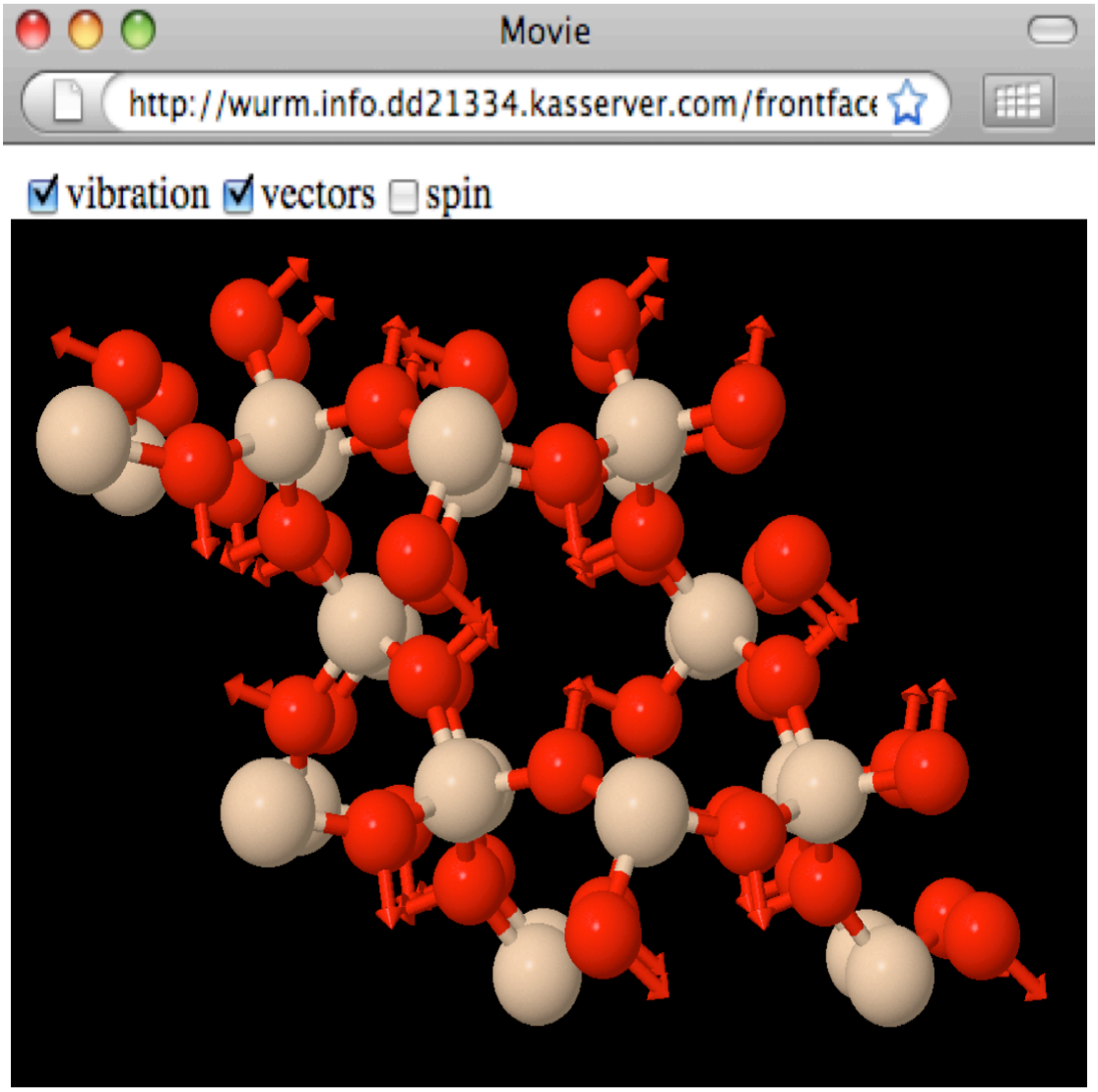

Figure 3. A Jmol applet is loaded in a pop-up window and allows the dynamic visualization of the phonon vibration.

Appendix – Detailed methodology

Computational details

We perform first-principles calculations using the local density approximation of density functional theory (Hohenberg and Kohn 1964; Kohn and Sham 1965) as implemented in the ABINIT package, based on planewaves and pseudopotentials (Gonze et al. 2002; 2005; 2009). We sample the reciprocal space using grids of equally distanced **k** points according to the Monkhorst-Pack scheme (Monkhorst and Pack 1976). The kinetic energy cut-off is related to the number of planewaves used to describe the electronic wavefunctions. As usual with planewave basis sets, the numerical accuracy of the calculation can be systematically improved by increasing the cut-off kinetic energy of the planewaves and the density of the sampling of the Brillouin zone (Payne et al. 1992). Usually we aim at a convergence in total energy better than 1mHa, corresponding roughly to 0.0005% of the total energy or less. In terms of energy derivative relative to volume this corresponds to a convergence on the order of a few kbars.

The calculations are performed using Troullier-Martins scheme (Troullier and Martins 1991) for pseudopotentials. For some particular cases we chose to employ equivalent Teter norm-conserving pseudopotentials (Teter 1993).

Crystal structures

We always start with the experimental structures and consider three possible options. First we can use the experimental structure without performing any structural

optimization. Often in this case some phonon modes are artificially softened and may even become imaginary. Second we allow the full relaxation of the unit cell parameters and the atomic internal positions with the constraint of preserving the value of the experimental density (i.e. unit cell volume) at ambient conditions. In this case we minimize the residual forces on the atoms and we eliminate (i.e. minimize) the non-hydrostatic stresses on the unit cell, but we keep the total volume fixed. Usually the lattice bears a finite hydrostatic pressure. Third we allow the full relaxation of the unit cell parameters and the atomic internal positions. In this case we minimize the residual forces on the atoms and we minimize both the non-hydrostatic and the hydrostatic stresses on the unit cell.

Because the electronic wavefunctions have the periodicity of the lattice, all the calculations must be performed in the smallest unit cell – the primitive unit cell. This restriction avoids the phase-corrections of the wavefunctions and of the phonons that would otherwise contribute from the neighboring Brillouin zone by folding back to the original zone center. The same restriction is highly useful as it ensures the minimum computational effort by reducing the size of the unit cell.

There are two ways to treat solid solutions. The classical method is the cluster expansion model. According to this approach the solid solution is described using supercells with different representative atomic configurations and making use of statistical physics. This model is straightforward and it can be applied to any type of disordered structure or solid solution. However the results are highly dependent on the quality of the sampling. Each individual calculation is more tedious because of larger system sizes and possible losses of symmetry due to the specific atomic configuration.

An alternative approach is the virtual crystal approximation (Bellaiche and Vanderbilt 2000). For each of the crystallographic sites affected by the solution we construct one alchemical pseudopotential mixing the atoms that participate in the solid solution in the appropriate amounts (Giantomassi et al. 2005; Cohen 2007). This approach uses small simulation cells and preserves the symmetry, but the mixing should be realized between chemically similar atoms. In practice the alchemical pseudopotentials are obtained as follows: the local potentials of the original pseudopotentials are linearly mixed; the form factors of the non-local projectors are all preserved, and all considered to generate the alchemical potential; the scalar coefficients of the non-local projectors are multiplied proportionally; and the core charge cut-off radius and functions are linear combinations of respectively the radii and the functions of the original pseudopotentials. This second approach has been successfully applied for example to compute the Raman spectra of disordered spinel (Caracas and Banigan 2009).

Before the start of the WURM project we performed calculations on a variety of minerals, with various structural types and atomic compositions (Caracas and Cohen 2006; 2007; Caracas, 2007; Hermet et al. 2007; Wang et al., 2008; Caracas and Banigan 2009). Most of these minerals were also characterized experimentally in Raman spectroscopy and we could use this as a basis for comparison. In most of the cases the theoretical structure is the one that yields the best agreement between theory and experiment. For minerals with anisotropic structures or with structures containing large void spaces the best agreement is obtained if the Raman calculation is performed at experimental density.

Density functional perturbation theory

We employ the density functional perturbation theory (DFPT) to compute the vibrational properties (Baroni et al 2001; Gonze et al 2005a; Caracas and Gonze 2010). In DFPT one expresses the energy of the crystal, F_{e+i}, as a Taylor expansion with respect to infinitesimal perturbations, λ_i:

$$F_{e+i}[\lambda] = F_{e+i}^{(0)}[\lambda] + \sum_i \left(\frac{\partial F_{e+i}}{\partial \lambda_i} \right) \lambda_i + \frac{1}{2} \sum_{ij} \left(\frac{\partial^2 F_{e+i}}{\partial \lambda_i \partial \lambda_j} \right) \lambda_i \lambda_j + \frac{1}{6} \sum_{ij} \left(\frac{\partial^3 F_{e+i}}{\partial \lambda_i \partial \lambda_j \partial \lambda_k} \right) \lambda_i \lambda_j \lambda_k + \ldots \quad (1)$$

The above equation can be solved analytically using the Hellman-Feynman description of the forces that allows us to obtain the first derivatives of the wavefunctions and energy and the 2n+1 theorem which allows us to compute the $(2n+1)^{th}$ order derivatives and their respective terms of the Taylor expansion knowing the n^{th} order derivatives of the wavefunctions.

Considering the atomic displacements, $\boldsymbol{\tau}$, and the electric fields, $\boldsymbol{\varepsilon}$, as perturbations, then each of the different terms of the full expansion up to the third order corresponds to a (measurable) physical property (derived in Veithen et al. 2005; Gonze et al. 2005a; discussed in more detail in Caracas and Gonze 2010):

$$
\begin{aligned}
&F_{e+i}[\mathrm{R}_k, E] = F_{e+i}[\mathrm{R}_k, E] \\
&-\Omega_0 \sum_{\alpha} \underbrace{P_\alpha^s}_{\text{polarization}} \varepsilon_\alpha - \sum_\alpha \sum_k \underbrace{F_\alpha^0}_{\text{forces}} \tau_{\kappa\alpha} \\
&-\Omega_0/2 \sum_{\alpha\beta} \underbrace{\chi_{\alpha\beta}^{\infty(1)}}_{\text{dielectric tensor}} \varepsilon_\alpha \varepsilon_\beta - \sum_{\alpha\beta} \sum_\kappa \underbrace{Z_{\kappa,\alpha\beta}^*}_{\text{dynamical charges}} \tau_{\kappa\alpha} \varepsilon_\beta + 1/2 \sum_{\alpha\beta} \sum_{\kappa_1\kappa_2} \underbrace{C_{\alpha\beta}(\kappa_1,\kappa_2)}_{\text{interatomic force cts.}} \tau_{\kappa\alpha} \tau_{\kappa'\beta} \\
&-\Omega_0/3 \sum_{\alpha\beta} \underbrace{\chi_{\alpha\beta\gamma}^{\infty(2)}}_{\text{Non-linear opt. coeff.}} \varepsilon_\alpha \varepsilon_\beta \varepsilon_\gamma - \Omega_0/2 \sum_\kappa \sum_{\alpha\beta} \underbrace{\frac{\partial \chi_{\alpha\beta}^{\infty(1)}}{\partial \tau_{\kappa\gamma}}}_{\text{Raman coeff.}} \varepsilon_\alpha \varepsilon_\beta \tau_{\kappa\gamma} \\
&-1/2 \sum_{\kappa_1\kappa_2} \sum_{\alpha\beta} \frac{\partial Z_{\kappa_1,\alpha\beta}^*}{\partial \tau_{\kappa_2\gamma}} \tau_{\kappa_1\alpha} \tau_{\kappa_2\gamma} \varepsilon_\beta + 1/3 \sum_{\kappa_1\kappa_2\kappa_3} \sum_{\alpha\beta} \underbrace{\Xi(\kappa_1,\kappa_2,\kappa_3)}_{\text{anharmonicities}} \tau_{\kappa_1\alpha} \tau_{\kappa_2\beta} \tau_{k_3\gamma} + \ldots.
\end{aligned}
\tag{2}
$$

α, β, γ stand for Cartesian directions, κ, κ', κ" stand for translation vectors of the unit cells and Ω_0 is the unit cell volume. The complexity of the exact expression for each higher-order term of the Kohn-Sham equations and consequently the complexity of the implementation increase with the order of the derivation.

By inverting this equation one can obtain the different physical properties of interest. The first order properties are vectorial properties and all the higher order properties are tensors. In the following we briefly comment each of the terms in the Equation 2:

The first order terms correspond to the polarization, P, and forces, F, defined as the derivative of the energy respectively to one electric field and to the one atomic displacement. The centro-symmetric have a nil residual spontaneous polarization and the relaxed structures have nil forces.

The second order terms consist of the first order dielectric tensor, $\chi^{\infty(1)}$, which is the derivative of the energy with respect to two electric fields; the Born effective charges,

Z^*, defined as the mixed derivative to atomic displacements and electric fields; and the interatomic force constants tensor, C, obtained as the change in energy due to two atomic displacements in two unit cells separated by the phonon wavevector $q=\kappa-\kappa'$. The dielectric tensors can be static or frequency dependent. The square roots of the static dielectric tensor give the refractive index tensor. The Born effective charges can also be seen as the change in polarization due to atomic displacements and can be related to the amount of charge exchanged by neighboring atoms during vibrations. The interatomic force constants when normalized with the masses yield the dynamical matrices.

The phonons frequencies are obtained by diagonalizing these dynamical matrices. The Raman and infrared modes are obtained form the dynamical matrices in the zone center, characterized by q=0. There are selection rules based on the crystal symmetry and on its effect on the vibrational pattern associated to a specific mode to determine whether that mode is Raman-active, infrared-active or silent.

The third order terms in the Equation 2 correspond to non-linear effects: the non-linear optical coefficients, $\chi^{\infty(2)}$, are the derivative of the energy with respect to three electric fields, and the Raman coefficients are the derivative of the energy with respect to two electric field and one atomic displacement. The third term of the third order derivative of the energy, i.e. with respect to one electric field and two atomic displacements participates to the width of the infrared and Raman peaks, while the very last term from the expression yields anharmonicities, related to the phonon time of life, the line broadening of Raman and infrared spectra and thermal transport due to phonon scattering.

Currently there are two available implementations for the Raman tensors within DFPT: one in the espresso package (Baroni and Resta 1986; Baroni et al. 1987; Baroni et al. 2001; Lazzeri and Mauri 2003; Scandolo et al. 2005) and a second one in the abinit package (Gonze and Vigneron 1989; Gonze et al. 1992; Gonze and Lee 1997; Gonze et al. 2002; 2005; 2005a; 2009; Veithen et al. 2005).

Raman spectra

The Raman intensity for an xy-oriented single-crystal depends on the frequency of the incident laser ω_L, on temperature, T, which are user-defined parameters, on the frequency of the Raman mode ω_i and on the Raman tensor α_{xy} (Cardona 1982):

$$I_{xy} \approx C\alpha_{xy}^2 \tag{3}$$

where the prefactor *C* is

$$C \approx (\omega_L - \omega_i)^4 \frac{1 + n(\omega_i)}{30\omega_i} \tag{4}$$

The temperature dependence is given by the Bose occupancy factor, n(ω):

$$1 + n(\omega_i) = [1 - \exp(-\hbar\omega_i / kT)]^{-1} \tag{5}$$

The Raman tensors are defined as:

$$\alpha_{xy} = \sqrt{\Omega_0} \sum_{\alpha\tau} \frac{\partial \chi_{xy}}{\partial r_{\alpha\tau}} \tag{6}$$

where Ω_0 is the unit cell volume, the atom α moves along direction τ, and χ_{xy} is the macroscopic dielectric tensor:

$$\chi_{xy} = \frac{\varepsilon_{xy}^{\infty} - \delta_{xy}}{4\pi} \tag{7}$$

The derivative of the macroscopic dielectric tensor is computed with respect to the atomic displacements, $r_{\alpha\tau}$, corresponding to the phonon eigenvector.

For non-centrosymmetric crystals, some modes are active in both Raman and infrared. In this case, the same formalism as above holds for the TO component, while a supplementary correction is needed for the modes in LO geometry that accounts for the coupling between the incident laser and the mode dipole.

We next need to build the Raman spectra for crystalline powders. For this we neglect the surface reconstruction and the grain size effects and we perform integrals over all possible orientation of ideal bulk crystals. These integrals are reduced to the following three rotational invariants (Placzek 1934; Hayes and Loudon 1978; Cardona 1982; Prosandeev et al. 2005; Caracas and Cohen 2006; Caracas 2007; Caracas and Gonze 2010):

$$
\begin{aligned}
G^{(0)} &= \sum_{ii,i=x,y,z} (\alpha_{ii})^2 / 3 \\
G^{(1)} &= \sum_{ii,i,j=x,y,z} (\alpha_{ij} - \alpha_{ji})^2 / 2 \\
G^{(2)} &= \sum_{ii,i,j=x,y,z} (\alpha_{ij} + \alpha_{ji})^2 / 2 + (\alpha_{ii} - \alpha_{jj})^2 / 3
\end{aligned}
\tag{8}
$$

The intensities of the two polarized components of the powder spectra, parallel and perpendicular, and the resulting total powder spectra are:

$$
\begin{aligned}
I_{\parallel}^{powder} &= C(10G^{(0)} + 4G^{(2)}) \\
I_{\perp}^{powder} &= C(5G^{(1)} + 3G^{(2)}) \\
I_{total}^{powder} &= I_{\parallel}^{powder} + I_{\perp}^{powder}
\end{aligned}
\tag{9}
$$

The single-crystal treatment of the Raman spectra, containing the dependence of the Raman spectra with the sample orientation, will be addressed in a future development of the database.