On the use of ab initio methods for the interpretation of the experimental IR reflectance spectra of crystalline compounds

Marco De La Pierre,1 Cédric Carteret,2 Roberto Orlando,1 and Roberto Dovesi1

1 Dipartimento di Chimica, Università di Torino and NIS (Nanostructured Interfaces and Surfaces) Centre of Excellence, Via P. Giuria 7, 10125 Torino, Italy
2 Laboratoire de Chimie Physique et Microbiologie pour l’Environnement (LCPME), UMR 7564, Université de Lorraine - CNRS, 405 rue de Vandœuvre, 54601 Villers-lès-Nancy, France

(Dated: March 1, 2013)

Abstract

It is shown that ab initio simulation can be used as a powerful complementary tool in the interpretation of the experimental reflectance spectra \( R(\nu) \) of crystalline compounds. Experimental frequencies and intensities are obtained from a best fit of \( R(\nu) \) with a set of damped harmonic oscillators, whose number and initial position in frequency can dramatically influence the final results, as the parameters are strongly correlated. Computed ab initio values for frequencies and intensities are accurate enough to represent an excellent starting point for the best fit process. Moreover, at variance with respect to experiment, simulation permits to identify all the symmetry allowed modes, also when characterized by low intensity or when close to a very intense peak. Overall, simulation-aided analysis of experimental spectra prevents from classifying combination modes as fundamental modes, and permits to discard artifacts due to superposition of bands, background and noise. Finally, it allows to (almost) completely characterize the set of fundamental modes.

Keywords: IR reflectance spectra, experiment, best fit, frequencies, intensities, ab initio simulation, CRYSTAL code
I. INTRODUCTION

A reflectance infrared (IR) experiment provides a spectrum $R(\nu)$ that is defined in terms of the complex dielectric function $\epsilon(\nu)$ and of the angle $\theta$ between the incident beam and the normal to the surface:\footnote{1}

$$R(\nu) = \left| \frac{\sqrt{\epsilon(\nu) - \sin^2(\theta) - \cos(\theta)}}{\sqrt{\epsilon(\nu) - \sin^2(\theta) + \cos(\theta)}} \right|^2 \quad (1)$$

For ease of notation the directional nature of $R(\nu)$ and the tensor nature of $\epsilon(\nu)$ are disregarded at this point. They will be taken into account later on.

$R(\nu)$ does not directly provide the quantities of interest for each $n$th IR active mode, namely the vibrational frequency $\nu_{TO,n}$ of the transverse optical (TO) mode and the corresponding intensity, usually expressed in terms of the oscillator strength $f_n$. A model must be adopted instead, that establishes the link between $R(\nu)$ and the set \{$\nu_{TO,n}; f_n$\}.

$\epsilon(\nu)$ is obtained as the sum of the high frequency dielectric contribution $\epsilon_\infty$ (due to the electrons) and of a ionic contribution. As regards the latter, different models have been proposed. The most common one is the Drude-Lorentz model\textsuperscript{1–4}, which consists in a superposition of Lorentzian curves $L_n(\nu)$, each one corresponding to a damped harmonic oscillator:

$$\epsilon(\nu) = \epsilon_\infty + \sum_n L_n(\nu) = \epsilon_\infty + \sum_n \frac{f_n \nu_{TO,n}^2}{\nu_{TO,n}^2 - \nu^2 - i\nu \gamma_n} \quad (2)$$

Each oscillator is then characterized by three parameters, namely the frequency $\nu_{TO,n}$ of the TO mode, its oscillator strength $f_n$ and the damping factor $\gamma_n$. Note that also the optical dielectric constant $\epsilon_\infty$ must be regarded as a parameter.

Once a model for the dielectric function has been chosen, its parameters can be obtained by best fit, through minimization of the chi square ($\chi^2$) between the calculated (reconstructed) spectrum $R^{calc}(\nu)$ (Eq. (1)) and the experimental one $R^{exp}(\nu)$, over a set of points on which the spectrum has been digitalized ($\sigma^{exp}$ is the experimental standard deviation):

$$\chi^2 = \sum_{i=1}^{M} \left( \frac{R(\nu_i)^{exp} - R(\nu_i)^{calc}}{\sigma_i^{exp}} \right)^2 \quad (3)$$

Typically\textsuperscript{5–11} the spectra cover a range of about 1000-1500 cm$^{-1}$ in steps of 0.25 to 4 cm$^{-1}$; then they are composed of a number of points $M$ in the order of thousands.
This best fit procedure is not straightforward at all, for various reasons: \( R(\nu) \) is non-linear with respect to the parameters; the \( \chi^2 \) function presents a large number of local minima; parameters are strongly correlated (\( f_n \) and \( \gamma_n \) in particular); the number of parameters can be very high. As regards the latter point, let us consider the case of ortho-enstatite\(^{11}\) (MgSiO\(_3\)) as an example. There are three polarization directions (with B\(_{1u}\), B\(_{2u}\) and B\(_{3u}\) symmetry), so the above discussion applies to three \( R_i(\nu) \) functions, one for each polarization direction \( i \). According to standard group theory analysis, the number of active modes along each direction is 29. This implies that for each direction the fitting function of Eq. (3) contains as many as 3\( \cdot 29 + 1 = 88 \) parameters.

One of the crucial points is then a proper definition of the number, position (on the frequency axis) and intensity of the oscillators to be used in the best fit; in other words, a set of quite reasonable initial parameters is required.

For this purpose, a Kramers-Kronig analysis is often applied to \( R(\nu) \) (see for example Hofmeister\(^{5-7}\)). This method provides the dielectric function \( \epsilon(\nu) \), which contains estimated values for the TO frequencies \( \nu_{TO,n} \) (they are the maxima of the imaginary part of \( \epsilon(\nu) \)), the longitudinal optical (LO) frequencies \( \nu_{LO,n} \) (minima of the imaginary part of \( 1/\epsilon(\nu) \)), the oscillator strengths \( f_n \) (integrated areas) and the damping factors \( \gamma_n \) (full widths at half height).

However, the Kramers-Kronig analysis requires an integration of the reflectance function extended to the full frequency axis, and this in general requires extrapolations at the two extremes. Typical sources of noise at low frequencies (i.e. \(< 100-200 \text{ cm}^{-1}\)) include sample size, beam quality, setup for low temperature measurement. Moreover, the identification of the modes contributing to the spectrum can be tricky, for many reasons (superposition of peaks, low intensity, noise, combination modes). As a matter of fact, the number of peaks to be looked for, their position and intensity are still largely undefined at the beginning of the best fit process.

As regards the number of IR modes to be included in the model, two extreme strategies can be found in the literature:

- As the symmetry analysis indicates that \( N \) modes must be active, then \( N \) harmonic oscillators are used in the fitting process. This is the attitude usually adopted by Hofmeister, that in the case of garnets\(^{5-7}\) tries in all cases to identify the \( N=17 \) oscillators suggested by symmetry. To reach the desired number, minor features of the
spectrum are also considered, that may correspond or not to fundamental frequencies.

- A harmonic oscillator is assigned only to the spectral features that show a good signal to noise ratio in \( R(\nu) \), i.e. one oscillator per one band. This is the attitude usually adopted by Suto and collaborators (private communication\(^{12} \), and Refs. 8–11). In this case oscillators that are IR allowed, but whose IR intensity is very low, might be excluded from the analysis. Oscillators with very close frequencies such that the convolution of the two peaks appears as a single peak are attributed to a single oscillator (i.e. no deconvolution process is applied to the broad features). As a consequence, the list of observed frequencies and intensities is shorter than that proposed by symmetry analysis. For example, in the case of ortho-enstatite\(^{11} \) this approach allows Suto and collaborators to identify only 65 peaks out of 87.

In recent years quantum mechanical computer simulation of the electronic properties of crystalline compounds has become a common practice (an incomplete list of software packages includes CRYSTAL\(^{13} \), Quantum Espresso\(^{14} \), SIESTA\(^{15} \), VASP\(^{16} \)). The vibrational frequencies and, in most of the cases, also the intensities are generated in a relatively standard way with these codes (see, for example, Refs. 17–21). Excellent results have been obtained in the calculation of the vibrational properties of ionic and semi-ionic compounds\(^{22–32} \). These codes are publicly available and their price, if any, is a small fraction of the cheapest spectrometer; then experimental groups have now the possibility to use \textit{ab initio} simulation as a complementary tool in the interpretation of the IR reflectance spectra.

The simulated frequencies and intensities can strongly depend on the level of theory, the variational basis set and, last but not least, the adopted code. However the features that can bias the simulated spectrum are different (say “orthogonal”) from the ones that make the interpretation of the experimental spectrum difficult. Let us mention some of these issues: a) the simulated spectrum (SS) contains only the fundamental frequencies; in the experimental spectrum (ES) combination modes, overtones and Fermi resonances can appear; b) in the SS the composition is defined a priori; in ES purity is a target not always easily reached; c) in SS the sample is infinite and perfect, in ES the size of the sample may be a problem; d) the SS refers to 0 K, the experiment can explore a wide range of temperatures; e) isotopic substitution is easy and cheap in SS, and the regular distribution of the isotopes can be imposed a priory, whereas this is not the case in ES; f) the SS is affected by model limi-
tations (harmonic approximation, choice of the hamiltonian and basis set), whereas the ES can be affected by instrumental limitations (signal-to-noise ratio, frequency ranges difficult to explore). As a result, the synergic use of the two approaches can play a crucial role in understanding the vibrational features of the compound under study.

In the present paper we explore the possibilities offered by quantum-mechanical simulation in the interpretation of the experimental reflectance spectra of crystalline compounds, including symmetry analysis of the modes, prediction of frequencies and intensities, identification of fundamental and combination modes. We will use a wide set of $R(\nu)$ spectra (3 for forsterite, 1 for grossular, 3 for enstatite) measured by different research groups (Hofmeister, Suto), to provide evidence of the usefulness of the synergic approach here presented.

The paper is organized as follows. Section II describes the adopted methods: simulation (IIA), reflectance experiments (IIB), best fit of the spectra (IIC). Section III presents some examples of reflectance spectra, analyzes them and discusses the contribution provided by simulation. The main conclusions are drawn in Section IV.

II. METHOD

A. Computational Details

Simulations were performed with the CRYSTAL09 periodic ab initio code\textsuperscript{13,33}, by using all-electron Gaussian-type basis sets and the B3LYP hybrid functional\textsuperscript{34–36}. For a detailed discussion of the computational details, see the Supplementary Material and Refs. 11,30,31. We here outline the calculation of the vibrational frequencies, IR intensities and dielectric tensors, as implemented in the CRYSTAL code\textsuperscript{33}.

The calculation of the transverse optical (TO) vibrational frequencies $\nu_{TO,n}$ at the $\Gamma$ point is performed within the harmonic approximation. The key quantity is the matrix $H$ of the second derivatives of the total energy $V$ with respect to the atomic cartesian coordinates $u$, which is constructed by numerical differentiation of the analytical gradient vector $v$:

$$H_{\alpha i, \beta j} = \frac{1}{2} \left[ \frac{\partial^2 V}{\partial u_{\alpha i} \partial u_{\beta j}} \right]_0 = \frac{1}{2} \left[ \frac{\partial v_{\alpha i}}{\partial u_{\beta j}} \right]_0$$

$$\approx \frac{1}{2} \frac{v_{\alpha i}(0, \cdots, u_{\beta j}, 0, \cdots) - v_{\alpha i}(0, \cdots, -u_{\beta j}, 0, \cdots)}{2u_{\beta j}}$$

(4)
where greek and latin indices refer to atoms and atomic cartesian coordinates, respectively. Note that the calculated (optimized) equilibrium geometry is taken as a reference. The TO frequencies are then obtained as the eigenvectors of the mass-weighted Hessian matrix $W$ at the $\Gamma$ point:

$$W_{\alpha i,\beta j} = \frac{H_{\alpha i,\beta j}}{\sqrt{M_\alpha M_\beta}}$$  \hspace{1cm} (5)

where $M_\alpha$ and $M_\beta$ are the masses of atoms $\alpha$ and $\beta$.

The longitudinal optical (LO) frequencies $\nu_{LO,n}$ can be calculated by considering that the dynamical matrix in ionic and semi-ionic compounds is a sum of two terms:

$$W_{\alpha i,\beta j}(k \rightarrow 0) = W_{\alpha i,\beta j}(k = 0) + W^{NA}_{\alpha i,\beta j}(k \rightarrow 0)$$  \hspace{1cm} (6)

The first term is the analytical part and corresponds to the quantity calculated through Equation (5). The latter term, the non-analytical contribution, depends on the momentum ($k$) exchanged with the radiation field (see Ref. 2 sections 5, 10, 34, 35; and Ref. 37):

$$W^{NA}_{\alpha i,\beta j}(k \rightarrow 0) = \frac{4\pi}{\Omega} \frac{\sum_m k_m Z^{*}_{\alpha,mi}(\sum_n k_n Z^{*}_{\beta,nj})}{\sqrt{M_\alpha M_\beta}(\sum_{m,n} k_m \epsilon_{\infty} k_n)}$$  \hspace{1cm} (7)

This equation shows that the non-analytic correction to the dynamical matrix originates from parallel polarization and phonon wave vector. The related modes are LO modes. Their frequencies are then obtained through diagonalization of the complete (i.e. analytical plus non-analytical) dynamical matrix.

Once the Hessian matrix $H$ is calculated, frequency shifts due to isotopic substitutions can be obtained at zero computational cost, by changing the masses in formula (5). This technique represents a useful tool to evaluate the participation of the various atoms to the vibrational modes in the different frequency regions of the spectra. Further details on the calculation of vibrational frequencies can be found in Refs. 17 and 18.

The oscillator strengths $f_n$ are computed for each $n$-th mode by means of the mass-weighted effective mode Born charges $\bar{Z}_n$:

$$f_{n,ij} = \frac{1}{4\pi\epsilon_0} \frac{4\pi}{\Omega} \frac{\bar{Z}_{n,i} \bar{Z}_{n,j}}{\nu_{TO,n}^2}$$  \hspace{1cm} (8)

where $\epsilon_0$ is the vacuum dielectric permittivity ($1/4\pi\epsilon_0 = 1$ atomic unit), $\Omega$ is the unit cell volume, $i$ and $j$ refer to the cartesian components, $t_{n,\alpha i}$ is an element of the eigenvectors...
matrix $T$ of the mass-weighted Hessian matrix $W$, that transforms the cartesian atomic directions into the $n$–th normal coordinate directions; finally, $Z_{\alpha,ij}$ is the Born effective charge tensor associated with atom $\alpha$, which is evaluated through a Berry phase approach\textsuperscript{19–21}. The ionic components to the static dielectric tensor are evaluated as the sum of the oscillator strengths: $F_{ij} = \sum_n f_{n,ij}$. The electronic high frequency components $\epsilon_{\infty,ij}$ are calculated through the Coupled-Perturbed KS/HF (Kohn-Sham/Hartree-Fock) scheme\textsuperscript{40–44}. Note that $\epsilon_\infty$ is almost independent from frequency in the IR range, as electronic transition energies are very large compared to vibrational energies.

**B. Experimental Setup**

Digitalized spectra were provided by Hofmeister\textsuperscript{5–7} and Suto\textsuperscript{9–11}, that we carefully acknowledge.

Suto and coworkers\textsuperscript{8–11} collected polarized infrared reflectance spectra $R_{exp}(\nu)$ of single crystals at variable temperatures, namely in the ranges 295-50 K for forsterite and 300-60 K for enstatite. Spectra were recorded from 80 to 2500 cm$^{-1}$ with resolution between 0.25 and 1 cm$^{-1}$. The beam incident angle to the sample was 10$^\circ$. KRS-5 and Polyethylene Polarizers were used in the spectral range 2500-500 and 500-80 cm$^{-1}$, respectively. The s-polarized geometry was employed because it has been demonstrated to reduce the contamination from other crystallographic directions\textsuperscript{45,46}. Reflections with electric field polarization parallel to the axis $a$ were measured on (100) surface, axis $b$ on both (100) and (001), axis $c$ on both (100) and (010). Data were reduced to the final spectrum selecting those with the higher reflectivities among the measurements and eliminating those with anomalous features. The reflection spectra of several measurements were averaged for each axis.

As regards Hofmeister and coworkers\textsuperscript{5–7}, they collected infrared reflectance spectra $R_{exp}(\nu)$ for grossular mineral using a FTIR microscope equipped with a reflectance accessory. Far-IR (70÷500 cm$^{-1}$) and mid-IR (400÷1100 cm$^{-1}$) spectra were measured with a resolution of 4 and 1 cm$^{-1}$, respectively, at normal incidence and ambient temperature on polished randomly oriented crystals around 500 $\mu$m size. Far-IR files were scaled to match mid-IR files throughout the region of overlap (400÷500 cm$^{-1}$), and then the files were merged.
C. Best fit of the reflectance spectra

We focus here on the case of orthorhombic or higher symmetry, which includes the compounds to be discussed in the following. In this case, the complex dielectric tensor $\epsilon$ turns out to be diagonal. If we adopt the three-parameter Drude-Lorentz model, which describes each vibrational mode as a damped harmonic oscillator, then the tensor components are defined as follows (compare Eq. (2) in Section I for the scalar case):

$$\epsilon_{ii}(\nu) = \epsilon_{\infty,ii} + \sum_n L_n,ii(\nu)$$

$$= \epsilon_{\infty,ii} + \sum_n \frac{f_{n,ii} \nu^2_{TO,n}}{\nu^2_{TO,n} - \nu^2 - i\nu \gamma_{n}}$$

(10)

where the $ii$ subscript indicates the crystallographic direction of the normal to the analyzed surface. Note that the modes contributing to a given direction $ii$ are the ones whose oscillator strengths $f_n$ have a non-zero $ii$-th component.

Equation (1) for the reflectance curve takes the vector form:

$$R_i(\nu) = \left| \frac{\sqrt{\epsilon_{ii}(\nu) - \sin^2(\theta) - \cos(\theta)}}{\sqrt{\epsilon_{ii}(\nu) - \sin^2(\theta) + \cos(\theta)}} \right|^2$$

(11)

An alternative Four-Parameter Semi-Quantum model to describe the dielectric function is due to Berreman and Unterwald; it was applied for the first time by Gervais and Piriou. This model is able to take into account the asymmetry of the reflectance peaks, by attributing distinct values to the damping factors of the TO and LO modes. As a result, an expression containing four parameters for each oscillator is obtained:

$$\epsilon_{ii}(\nu) = \epsilon_{\infty,ii} \prod_n \frac{\nu^2_{LO,n}}{\nu^2_{TO,n} - \nu^2 - i\nu \gamma_{LO,n}}$$

(12)

where $\nu_{TO,n}$ and $\nu_{LO,n}$ are the TO and LO frequencies, and $\gamma_{TO,n}$ and $\gamma_{LO,n}$ are their corresponding damping factors. Note that, in this formula, the fact that modes contributing to a given direction $ii$ must have a non-zero component of $f_n$ is not explicit. The oscillator strengths can be calculated by comparison of Equations (10) and (12) in which the damping is neglected:

$$f_{n,ii} = \epsilon_{\infty,ii} \left( \frac{\nu^2_{LO,n}}{\nu^2_{TO,n}} - 1 \right) \prod_{k \neq n} \frac{\nu^2_{LO,k} - \nu^2_{TO,n}}{\nu^2_{TO,k} - \nu^2_{TO,n}}$$

(13)

where the $k$ product extends to all the modes corresponding to the $ii$ direction.

In the present study, best fit procedures against the Drude-Lorentz and Four-Parameter
Semi-Quantum models were performed with the RefFIT interactive program written by Alexey Kuzmenko\textsuperscript{50}. The simulated frequencies, intensities and high frequency dielectric tensors were used as an initial guess for the fits. First, damping factors were determined; then, individual mode parameters were optimized iteratively within a finite window around the simulated data (30 cm\textsuperscript{-1} for frequencies, 30 \% for intensities). In the case of modes with very low intensities, this window was reduced by a factor 3 to 4. Finally, a constrained global minimization on all the parameters was performed.

III. RESULTS

In the following we will examine a set of 7 IR reflectance spectra $R(\nu)$ (3 for forsterite, 1 for grossular, 3 for enstatite) taken from the literature. We will focus on the analysis of the experimental reflectance curves, by systematically employing simulation as a complementary tool able to provide the full set of mode frequencies and intensities, distinguish between fundamental and combination modes, analyze minor features. All spectra are fitted by using both the three-parameter Drude-Lorentz (DL) and the Four-Parameter Semi-Quantum (FPSQ) models. The full sets of computed and fitted quantities (frequencies, oscillator strengths, damping factors, dielectric tensors) are available as Supplementary Material.

A. Forsterite

Olivine forsterite, an orthosilicate with chemical formula Mg\textsubscript{2}SiO\textsubscript{4}, is the first example here considered. Its crystal structure is orthorhombic, which implies that three distinct reflectance spectra can be collected, along the $a$, $b$ and $c$ directions, with 13, 13 and 9 IR active modes, respectively (the total number of modes is 78, as the number of atoms is 27). Suto \textit{et al.}\textsuperscript{10} have identified 11, 12 and 8 modes along the three axes, respectively. The availability of three independent spectra for a single compound permits a quite rich analysis. On the other hand the limited number of modes per direction (13 or 9) simplifies the fitting process and the interpretation of the spectral features. Figures 1(a), 2(a) and 3(a) show the three spectra in the range 100-1200 cm\textsuperscript{-1}. They are characterized by broad intense bands (for example the one centered at about 600 cm\textsuperscript{-1} for
the $a$ axis), narrow intense peaks (e.g. the one at 200 cm$^{-1}$ for the $a$ axis), low-intensity peaks and side-peaks, the latter two not easy to identify at this spectral scale. Experimental spectra along the $b$ and $c$ axes are quite noisy up to 250 cm$^{-1}$; this does not permit to identify any peak in this range, except for the intense one at about 150 cm$^{-1}$ for the $b$ axis. Obviously intense peaks are easily localized in the fitting process. Thus, the following discussion will mainly focus on the low intensity peaks.

Selected zooms are proposed in Figures 1(b), 2(b) and 3(b). Let us start from the $a$ axis, for which the 350-600 cm$^{-1}$ range is shown (Figure 1(b)). Compared to the previous investigation by Suto et al.\textsuperscript{10}, our B3LYP simulation indicates additional modes at 475.9 and 540.0 cm$^{-1}$. As regards the first mode, a small peak appears in fact in the experimental spectrum at about the simulated frequency. Using the simulated data as an initial guess, both DL and FPSQ fitting models localize an oscillator in that frequency range, at 476.1 (DL) and 476.6 (FPSQ) cm$^{-1}$. Let us consider now the wide band between 500 and 550 cm$^{-1}$: it has a fine structure, and the peak splits on the top. When an oscillator is added at 540.0 cm$^{-1}$ (the value from simulation), it remains almost unchanged (542.7 cm$^{-1}$) if the DL model is used. The small shoulder around 550 cm$^{-1}$ is properly described, but the double peak structure does not appear. When the FPSQ model is used, on the contrary, the added oscillator moves to 532.3 cm$^{-1}$. The damping factors of the TO and LO modes differ significantly (5.7 and 9.4 cm$^{-1}$, whereas the DL model provides a unique TO value of 10.0 cm$^{-1}$). These adjustments permit to describe the double feature, still properly depicting the shoulder, thus improving the overall structure of the band. On the whole, the combined use of simulation and of two fitting models of increasing complexity does permit to recover all the fundamental peaks in the spectrum along the $a$ axis, and to describe its minor features.

The full spectrum along the $b$ axis is shown in Figure 2(a). There is one missing mode, which the B3LYP simulation locates at 637.6 cm$^{-1}$, with an almost null oscillator strength of only $1.5 \cdot 10^{-4}$. Due to this very low intensity, even a rescaling of the reflectance axis (not shown) does not permit to identify any peak from the background noise. Figure 2(b) gives a zoom of the spectrum along the $b$ axis in the 350-600 cm$^{-1}$ window. It is interesting to compare the behavior of the two fitting models. For example, the leftmost band in the Figure (between 350 and 380 cm$^{-1}$) is composed by only one oscillator. The two models locate the TO and LO frequencies nearly at the same values: 352.2 and 376.9 cm$^{-1}$ (DL) and 351.7 and 376.7 cm$^{-1}$ (FPSQ). The same behavior holds for the oscillator strengths,
that are 1.243 and 1.246, respectively. Nonetheless the Figure shows that the latter model achieves a significantly better agreement with the experiment. The reason is the availability of two distinct values for the TO and LO damping factors, that permit to take into account the peak asymmetry: the unique DL value is 1.2 cm\(^{-1}\), whereas the FPSQ model yields 2.8 (TO) and 1.0 (LO) cm\(^{-1}\); as a result, the fitting is much better. A similar behavior is clearly visible also in the case of the band at 530-580 cm\(^{-1}\) (see Figure), which is again correlated to one oscillator. Both the TO and LO frequencies differ by less than 2 cm\(^{-1}\) between the two fitting models, and the oscillator strength by only 0.006. Also in this case then the large difference is due to the damping factors: 7.7 cm\(^{-1}\) when DL is used, while the FPSQ model gives 4.6 and 8.3 cm\(^{-1}\). The band at 450-500 cm\(^{-1}\) would be an additional example of the fact that the FPSQ model on the one hand permits to achieve an impressive fit to the experimental spectrum, and on the other hand that this model, due to the higher number of fitting parameters, could not be applied without the availability of an accurate guess as provided by \textit{ab initio} simulation.

As anticipated above, the large experimental noise in the 100-270 cm\(^{-1}\) spectral range of the \textit{c} axis spectrum (Figure 3(a)) hinders the identification of the computed peak at 206.5 cm\(^{-1}\) (indeed of quite low intensity, \(f\) being only 2.0 \(\times\) 10\(^{-3}\)). Let us concentrate again in the spectral range 350-600 cm\(^{-1}\), shown in Figure 3(b). The single-oscillator band at 500-600 cm\(^{-1}\) shows an asymmetry similar to the one just discussed in the case of the \textit{b} axis. The most interesting feature of this spectrum is related to the broad band at 420-460 cm\(^{-1}\). The B3LYP simulation indicates one fundamental oscillator at 427.9 cm\(^{-1}\) (with \(f = 0.30\)). However, the inspection of the band in the experimental spectrum clearly shows a double feature on the band top. We performed a fit including one more oscillator, to account for one additional mode; the DL model yields frequencies at 418.7 and 425.6 cm\(^{-1}\), with intensities \(f\) of 0.73 (fundamental) and 7.0 \(\times\) 10\(^{-2}\) (additional), respectively; the FPSQ model gives very similar values. As simulation provides the complete list of fundamental IR modes, all those experimental features that do not have a corresponding simulated mode, as in this case, must be attributed to combination modes, overtones, or Fermi resonances.
B. Grossular

Grossular garnet (Ca$_3$Al$_2$Si$_3$O$_{12}$) is an orthosilicate with cubic structure, providing a single reflectance spectrum with 17 IR active modes. Hofmeister et al.$^5$ have identified 14 out of 17 modes in the reflectance spectrum; two additional modes (at 356 and 430 cm$^{-1}$) could be identified in the same study by performing additional absorption measurements on thin films.

Figure 4 shows the full reflectance spectrum in the range 100-1500 cm$^{-1}$ (a), as well as a zoom in the restricted range 325-650 cm$^{-1}$ (b). Let us start our analysis from the small peak at about 350 cm$^{-1}$. Our simulation locates a mode at 355.5 cm$^{-1}$; with this guess, the two fitting models identify a peak at 355.0 (DL) and at 355.1 (FPSQ) cm$^{-1}$. A second computed mode at 404.1 cm$^{-1}$ cannot be identified in the experimental spectrum (see Figure 4(b)). Indeed, this mode has an almost null intensity (the computed oscillator strength is 1.8·$10^{-3}$ only). The third extra-simulated mode is at 421.7 cm$^{-1}$ ($f=1.2·10^{-2}$). A shoulder is visible in this spectral range; however the almost null experimental (fitted) intensity does not permit to locate the corresponding mode frequency (the χ$^2$ of the fit is independent from its value).

It is worth discussing other two features in the zoomed range 325-650 cm$^{-1}$ (Figure 4(b)). Remarkably, the use of the FPSQ model permits to significantly improve the fitting of the peak at 617.8 cm$^{-1}$ with respect to the DL model. This is due to a modification of almost all the parameters of the oscillator: $\nu_{LO}$ (from 627.3 to 623.2 cm$^{-1}$), $f$ (from 0.048 to 0.025), and $\gamma$, from 28.4 cm$^{-1}$ (TO only) to 10.9 (TO) and 11.8 (LO) cm$^{-1}$. On the other hand, both fitting models fail in properly fitting the band at 470 cm$^{-1}$ with a single oscillator. This suggests the possible presence of additional modes; however, the large noise in the experimental data hinders in this case the addition of more oscillators to the fitting model. A new registration of the spectrum after more than 20 years from the pioneering paper of Hofmeister et al.$^5$ would certainly help in clarifying the residual doubts about this spectrum (as well as the related ones of the other members of the garnet family).
C. Enstatite

Ortho-enstatite (MgSiO$_3$) is a pyroxene inosilicate with orthorhombic crystal structure. There are 80 atoms in the unit cell, and then 237 vibrational modes. Among them, 87 are IR active. Three distinct IR reflectance spectra can be collected along the three crystallographic directions $a$, $b$, and $c$; each spectrum contains 29 fundamental IR modes. The resulting high number of peaks is evident when looking at Figures 5(a), 6(a) and 7(a). The large number of modes increases the complexity of the analysis, and makes particularly challenging the combined approach here presented.

Some of the present authors have already performed a joint investigation of the IR spectra of ortho-enstatite, which is published in Demichelis et al.$^{11}$ In that study, simulations and experiments have been carried out separately by Demichelis and Suto, respectively, the two sets of results being compared “a posteriori”. As regards the experimental outcome, Suto has identified 22, 22 and 21 modes (out of 29) in the spectra along the $a$, $b$, and $c$ axes, respectively.

Let us now proceed with our combined analysis. Small differences between the data obtained with the two fitting models are found only in a limited set of cases; they are related to either asymmetric peaks (already discussed for forsterite) or fitting correlation effects between $f$ parameters of superposed peaks (that however are small and do not affect the shape of the reconstructed spectra). In the following, we will discuss the results obtained with the FPSQ model only.

Let us consider the $a$ axis first. The full scale spectrum (150-1200 cm$^{-1}$) is shown in Figure 5(a), while two zooms in the ranges 150-400 and 600-1000 cm$^{-1}$ are reported in Figures 5(b) and 5(c). In the 150-400 cm$^{-1}$ range there are three peaks that have not been identified by Suto$^{11}$, and that on the contrary we were able to identify through fitting, thanks to the guess provided by the B3LYP simulation. The three small peaks are at 236.9, 264.3 and 283.2 cm$^{-1}$. There are probably two reasons why they have been missed by Suto: a) the low signal-to-noise ratio characterizing the spectrum (see Figure); b) the many contributions to the best fit process (22 oscillators and 67 parameters); in this case also a peak with non-negligible intensity has little influence on the $\chi^2$ index of the best fit, and is easily lost, if it is not looked for explicitly. Along this axis there are four more simulated modes that could not be identified neither by Suto nor by us through the best fit process. Two of them are at 86.1
cm$^{-1}$ and 145.8 cm$^{-1}$, and have very low computed intensities ($f$ are $0.1\cdot10^{-5}$ and $1.0\cdot10^{-3}$, respectively). As the three spectra are remarkably noisy below 150 cm$^{-1}$, only very intense features can be identified in this zone of the spectrum. The other two missing modes are at 870.8 and 928.6 cm$^{-1}$. The former has $f$ as small as $5.5\cdot10^{-4}$, and cannot be distinguished from the (slightly noisy) background (see Figure 5(c)). The latter is very close to another intense mode (the computed distance being as small as 0.9 cm$^{-1}$), while the $f$ ratio is about $1/17$; this makes impossible to resolve the two peaks through best fit. A look at Figure 5(c) confirms that only one large band is visible, and the two modes cannot be distinguished. The spectrum along the $b$ axis is represented in Figures 6(a) (spectral range 150-1200 cm$^{-1}$) and 6(b) (300-550 cm$^{-1}$). As in the case of $a$ axis, two low frequency modes at 109.9 and 190.5 cm$^{-1}$ could not be identified due to noise. As shown in the zoomed view, our best fit identifies four peaks that could not be identified by Suto. The first one is at 354.8 cm$^{-1}$, and is a shoulder of the intense peak at 350.3 cm$^{-1}$. Let us consider now the large band at 475-550 cm$^{-1}$, that is rich in spectral features. Suto has inserted only one oscillator in this zone. Thanks to the guess of the B3LYP simulation, we were able to identify three additional fundamental modes in a relatively narrow interval: 488.2, 499.4 and 514.8 cm$^{-1}$. In spite of the addition of these three peaks, there are features of this band that persistently are not described by the fitted curve; this suggests the possible presence of combination modes. An excellent agreement was in fact obtained by adding three extra modes in the best fit process at 507.1, 508.1 and 524.0 cm$^{-1}$. As they are not fundamental modes, they might be an artefact of the fitting process due to background noise, or correspond to combinations/overtones. We are not in a position to discriminate among these two hypotheses. Also the 850-950 cm$^{-1}$ interval deserves some attention (Figure 6(a), zoom not shown). Two bands can be clearly identified (they have been listed also by Suto) with frequencies of 889.8 and 919.1 cm$^{-1}$. However, the B3LYP simulation permits to identify through fitting also a shoulder mode at 905.6 cm$^{-1}$, its $f$ value being $1/26$ of the major mode at 889.8 cm$^{-1}$.

Finally we consider the $c$ axis: Figures 7(a) and 7(b) show the ranges 125-1200 and 250-500 cm$^{-1}$, respectively. Remarkably, in the region below 200 cm$^{-1}$ there are three modes that are intense enough to be distinguished from the large noise: 139.4 cm$^{-1}$ (the fitted oscillator strength $f$ is $8.9\cdot10^{-2}$), 147.1 cm$^{-1}$ ($f=5.5\cdot10^{-2}$) and 196.4 cm$^{-1}$ ($f=8.3\cdot10^{-2}$). Now, if we look at the zoom in Figure 7(b) our best fit identifies as many as six peaks that Suto
has not characterized; they are at 297.3, 349.1, 391.0, 429.6, 458.1 and 475.0 cm$^{-1}$. The first four are indeed not easy to distinguish in the spectrum, as the corresponding spectral range (see Figure) is rich in small features that could be interpreted as background noise or combination modes. The calculated mode at 458.1 cm$^{-1}$ is quite intense (fitted $f$ of 0.15). As for the mode at 283.2 cm$^{-1}$ along the $a$ axis, this mode is not easily identified due to the difficulties of the fitting process when a large number of parameters is present. The peak at 475.0 cm$^{-1}$ has a moderate intensity (fitted $f$ of 1.4·10$^{-2}$), and is clearly visible in the spectrum between the two large bands at 460 and 500 cm$^{-1}$. In this case it is then more difficult to understand why it has been lost by Suto. Finally, two modes could not be identified through fitting, in spite of the initial B3LYP guess, due to low signal-to-noise ratio; they are at 170.3 cm$^{-1}$ (computed $f$ of 5.7·10$^{-3}$) and 224.6 cm$^{-1}$ (1.7·10$^{-2}$).

Overall, the enstatite case brings two additional comments to our analysis: i) spectra with a high number of modes (29 in this case) are rich in minor features, among which it becomes extremely hard to distinguish from fundamentals, combinations and noise; ii) simulation in this respect has a major role in enriching the description of the spectrum; for enstatite, 14 more fundamental modes, out of 87, were identified, as well as 3 possible combination modes.

IV. DISCUSSION AND CONCLUSIONS

A) When approaching the analysis of the IR spectrum from the experimental side, a first issue to tackle is the identification of the set of peaks to be considered as fundamental modes. Intense features are easy to label as IR modes; however, when the signal-to-noise ratio is not that high, or in the case of superposing peaks, the analysis is ambiguous.

A computational approach turns out to be decisive in this respect. Beside providing the symmetry analysis of the modes at zero cost, it produces the full set of frequencies for all the fundamental modes of the system, including the IR active ones. The frequencies have usually an accuracy of a few cm$^{-1}$, so that they can be effectively used to locate IR fundamental modes in the experimental spectrum.

B) Once a set of modes is identified, a second task is to perform a best fit against the experimental curve. In this way, the frequency values are refined. Additional information is then obtained as a by-product: either oscillator strengths and damping factors (Drude-Lorentz
model) or LO frequencies and TO/LO damping factors (Four-Parameter Semi-Quantum model). Also this step presents many difficulties: the number of parameters to be optimized is high (for a system with N IR active modes, 3N+1 or 4N+1 parameters, depending on the adopted model), and correlation among them is large. As a consequence, the fitting procedure is cumbersome, and the obtained numbers can be affected by large error bars. In this respect, the advantage brought by the simulation is that it provides not only the IR active frequencies, but also the oscillator strengths and the LO frequencies. This set of quantities can be used as a starting point for the parameters of the fit, reducing the number of unknowns and providing a guess for the final values. Moreover, the values provided by simulation are quite accurate, which implies a well defined direction in the minimization process, as the sets of initial and final values are quite close to each other.

C) As a result of the fitting process, a calculated spectrum is obtained, that can be superposed and compared to the experimental one. What information is possible to extract, concerning the nature of the spectral features? A first point concerns the model adopted for the best fit. Drude-Lorentz describes properly narrow peaks or symmetric bands. On the other hand, the Four-Parameter Semi-Quantum model permits to improve the description of asymmetric spectral features, thanks to the use of distinct values for TO and LO damping factors (we must underline that the Four-Parameter Semi-Quantum fit, due to the larger number of parameters, can present higher difficulty when used without an appropriate initial guess).

By using simulation, an improved classification of spectral features can be performed. First, the quantum mechanical approach ensures that all the IR fundamental modes are identified. Then, by using the computed values as a starting point, both Drude-Lorentz and Four-Parameter Semi-Quantum fits can be performed, so that features of the fundamental peaks are properly described. Finally, all the additional spectral features, identified through best fit, can be labeled, by exclusion, as combination modes, overtones, or Fermi resonance modes. Therefore, a robust interpretation scheme is provided by simulation, able to unveil the nature of almost all the spectral features.
V. ACKNOWLEDGEMENTS

The authors acknowledge Compagnia di San Paolo for financial support (Progetti di Ricerca di Ateneo-Compagnia di San Paolo-2011-Linea 1A, progetto ORTO11RRT5) and the CINECA supercomputing centre for computer support. A. M. Hofmeister and H. Suto are carefully acknowledged for providing us with the experimental spectra. F. Pascale is acknowledged for its support. CC would like to thank F. Brehat and B. Wyncke for fruitful discussions and continuous support.
FIG. 1: Reflectance spectra along the $a$ axis of forsterite, in the ranges (a) 100-1200 cm$^{-1}$ and (b) 350-600 cm$^{-1}$: experimental (circles), fitted with the DL model (dashed line, Eq. (10) in Section II A), fitted with the FPSQ model (solid line, Eq. (12) in Section II.A). Additional peaks (compared to previous experimental investigations) identified through simulation are indicated by arrows, together with the corresponding frequencies as obtained with the FPSQ model.
FIG. 2: Reflectance spectra along the $b$ axis of forsterite, in the ranges (a) 100-1200 cm$^{-1}$ and (b) 350-600 cm$^{-1}$. See caption to Figure 1 for more details.
FIG. 3: Reflectance spectra along the $c$ axis of forsterite, in the ranges (a) 100-1200 cm$^{-1}$ and (b) 350-600 cm$^{-1}$. See caption to Figure 1 for more details.
FIG. 4: Reflectance spectra of grossular, in the ranges (a) 100-1500 cm\(^{-1}\) and (b) 325-650 cm\(^{-1}\).

See caption to Figure 1 for more details.
FIG. 5: Reflectance spectra along the $a$ axis of ortho-enstatite, in the ranges (a) 150-1200 cm$^{-1}$, (b) 150-400 cm$^{-1}$ and (c) 600-1000 cm$^{-1}$. See caption to Figure 1 for more details.
FIG. 6: Reflectance spectra along the $b$ axis of ortho-enstatite, in the ranges (a) 150-1200 cm$^{-1}$ and (b) 300-550 cm$^{-1}$. See caption to Figure 1 for more details.
FIG. 7: Reflectance spectra along the $c$ axis of ortho-enstatite, in the ranges (a) 125-1200 cm$^{-1}$ and (b) 250-500 cm$^{-1}$. See caption to Figure 1 for more details.
In the fitting analysis on the experimental data, we tried to minimize the number of the vibration modes, and adopt one mode to each feature which has good signal to noise ratio. For the broad features which may contain multiple modes, we did not try deconvolve them into multiple modes, and applied one set of parameters (oscillation wavenumber, damping factor, strength).
this is because within the accuracy of our measurements, we achieved the reasonable fitting results adopting single mode for each feature and also we do not have the clear clues on the broad features of how to deconvolve them.


21 Y. Noël, C. M. Zicovich-Wilson, B. Civalleri, Ph. D’Arco, and R. Dovesi. Polarization properties of ZnO and BeO: an ab initio study through the Berry phase and Wannier functions approaches.


43 M. Ferrero, M. Rérat, R. Orlando, and R. Dovesi. *Coupled Perturbed Hartree-Fock Calculation*


50 RefFIT, a software to fit optical spectra, by Alexey Kuzmenko. [http://optics.unige.ch/alexey/reffit.html](http://optics.unige.ch/alexey/reffit.html).