

## Investigation of the LO-TO splitting in complex binary crystals

This article has been downloaded from IOPscience. Please scroll down to see the full text article.

1993 J. Phys.: Condens. Matter 5 6165

(<http://iopscience.iop.org/0953-8984/5/34/005>)

View [the table of contents for this issue](#), or go to the [journal homepage](#) for more

Download details:

IP Address: 171.66.16.159

The article was downloaded on 12/05/2010 at 14:21

Please note that [terms and conditions apply](#).

# Investigation of the LO–TO splitting in complex binary crystals

H Schober†† and D Strauch§

† Kernforschungszentrum Karlsruhe, D-76021 Karlsruhe, Federal Republic of Germany

‡ Institut Laue–Langevin, F-38042 Grenoble Cédex, France

§ Institut für Theoretische Physik, Universität, D-93053 Regensburg, Federal Republic of Germany

Received 22 April 1993, in final form 7 June 1993

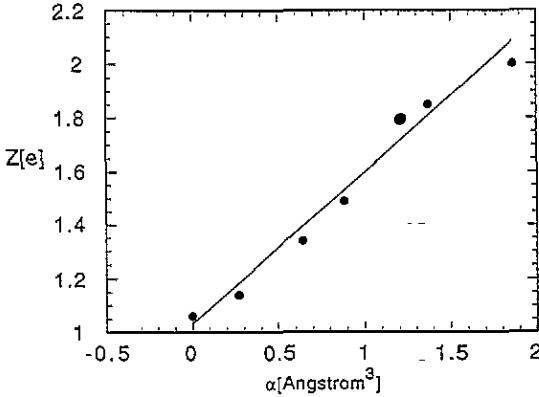
**Abstract.** The trace of the irregular part of the dynamical matrix for  $q = 0$  constitutes a measure  $\Omega^2$  of the LO–TO splitting. Assuming that the lattice dynamics of the crystals under investigation can be described by shell models, analytical expressions for  $\Omega$  as well as the high-frequency dielectric constants are derived for the large class of binary crystals with only one polarizable ion type. The general theory is applied to the concrete cases of  $\alpha$ -quartz ( $\text{SiO}_2$ ) and sapphire ( $\text{Al}_2\text{O}_3$ ). Using the analytical expressions, it is shown how the LO–TO splitting imposes strong correlations on the model parameters. Particular attention is given to the scaling of the ionic charge and high-frequency dielectric constant with the polarizability. This scaling behaviour is detected, when comparing different shell models, used to describe the lattice dynamics of  $\alpha$ -quartz and sapphire.

## 1. Introduction

When trying to describe the experimentally determined dispersion curves of a crystal by phenomenological models one often encounters the problem that several sets of parameters lead to comparable results. This is particularly true in the case of  $\alpha$ -quartz ( $\text{SiO}_2$ ). A recent study of the lattice dynamics of this crystal [1] showed that a large number of shell models describe the dispersion curves about equally well [1]. The existence of this model variety generally makes the interpretation of the model parameters a very tedious task and thus prevents us from gaining deeper physical insight into the dynamics of the crystal under investigation. This is, however, only true if the model parameters are uncorrelated. If, to the contrary, correlations between the model parameters exist, the fact that there is a large variety of models turns into an advantage, as it helps us to determine the nature of these correlations, i.e. to establish mathematical relations in between the model parameters, on the one hand, and specific aspects of the fitted data on the other hand, which permits us to explain the observed correlations. Both in order to make sure that the encountered correlations are statistically significant and also in order to determine the underlying mathematical relations the largest possible number of models is desirable. As the determined relations are a common property of the whole set of models and not only of a single member they constitute very reliable pieces of information concerning the dynamics of the crystal.

In the case of  $\alpha$ -quartz the most interesting parameters in the above-described context are the ionic charge  $Z$  and the polarizability  $\alpha$  of the *hypothetically* free oxygen ion. Depending on the short-range interactions, included in the shell-models, the values for  $Z$  and  $\alpha$  range from  $1.4e$  to  $2.0e$  and  $0.8 \text{ \AA}^3$  to  $1.9 \text{ \AA}^3$  respectively. As can be seen in figure 1, the ionic charge  $Z$  does not, however, vary randomly, but there is a strong nearly linear correlation with the polarizability  $\alpha$ . A similar behaviour is found for the high-frequency dielectric

constant (see section 8). The aim of this paper is to describe the origin of these correlations by giving analytical expressions for the functional dependence of the parameters and to demonstrate how these analytical expressions can be used to obtain physical estimates of the parameter values themselves.



**Figure 1.** Ionic charge  $Z$  as a function of the polarizability of the free oxygen ion  $\alpha$  for various shell models of  $\alpha$ -quartz. The full line is a least-squares fit through the data points. The slope of the line is  $0.57e \text{ \AA}^{-3}$  and  $Z(\alpha = 0) = 1.03e$ .

As we will show, the central property of the dispersion relations in crystals like  $\alpha$ -quartz responsible for the correlation of  $Z$  and  $\alpha$  is the strong LO–TO splitting. Its investigation constitutes the main subject of this paper. In order to get an analytical expression for the functional dependence of the ionic charge  $Z$  on the polarizability  $\alpha$  we quantify the LO–TO splitting for the general case of a crystal with several polar modes (section 2). This quantification is achieved through the trace  $\Omega^2$  of the irregular part of the dynamical matrix. As the irregular part of the dynamical matrix can easily be separated for crystals with non-polarizable ions,  $\Omega^2$  will first be calculated for rigid-ion models (section 3) yielding its connection with the plasma frequency. In section 4 an analytical expression for  $\Omega^2$  will be developed for crystals with polarizable ions, starting from the concept of lattice dynamical shell models. The analytical expression will be investigated (section 5) in order to explain the scaling of the ionic charge with the polarizability. The connection with the effective charges (section 6) and high-frequency dielectric tensor (section 7) will be demonstrated. The results will then be applied to the specific case of  $\alpha$ -quartz (section 8) and sapphire (section 9). We would like to stress that the obtained results are not limited to quartz and sapphire, but are applicable to a large class of binary crystals.

## 2. Definition of a general measure of the LO–TO splitting

In order to introduce notation, we briefly recall the origin of the LO–TO splitting. The Fourier transform of the Coulomb potential on a crystal lattice is given by

$$C_{\alpha\beta}(\kappa\kappa'|\mathbf{q}) = \sum_{\mathbf{G}} (\mathbf{q} + \mathbf{G})_{\alpha} (\mathbf{q} + \mathbf{G})_{\beta} \left( \frac{4\pi e^2}{v} \frac{1}{(\mathbf{q} + \mathbf{G})^2} \right) \exp(i\mathbf{G}[\mathbf{d}(\kappa) - \mathbf{d}(\kappa')]). \quad (1)$$

In (1)  $\mathbf{G}$  denotes the vectors of the reciprocal lattice,  $v$  the volume of the primitive cell and  $\mathbf{d}(\kappa)$  the position of the ion  $\kappa$  within the primitive cell. Due to the term with  $\mathbf{G} = 0$  expression (1), and therefore the dynamical matrix, is not well-defined at the zone centre, i.e. at  $\mathbf{q} = 0$ . This singularity leads to the splitting of the polar modes, which is known under the names of LO–TO splitting or Lyddane–Sachs–Teller splitting. We will not go into a discussion of the splitting itself, since this has been done extensively elsewhere [2].

Through the following equation we define a matrix  $\mathbf{C}^*$  which we call the irregular part of the Coulomb matrix†. The regular part, i.e.  $\mathbf{C} - \mathbf{C}^*$ , is denoted by  $\bar{\mathbf{C}}$ .

$$C_{\alpha\beta}^*(\kappa, \kappa') := \frac{4\pi e^2 q_\alpha q_\beta}{v q^2}. \quad (2)$$

Having defined the regular part of the Coulomb matrix  $\bar{\mathbf{C}}$ , the regular part of the dynamical matrix  $\mathbf{D}$  is obtained by replacing  $\mathbf{C}$  by  $\bar{\mathbf{C}}$  in the definition of  $\mathbf{D}$ .

In order to measure the LO-TO splitting quantitatively we need a scalar that unambiguously characterizes the crystal. In particular, this scalar should not be a function of the wave vector  $\mathbf{q}$ . As the LO-TO splitting has its origin in the irregular part of the dynamical matrix we are looking for a  $\mathbf{q}$ -independent linear mapping from the vector space of  $(3r \times 3r)$  matrices into the real numbers, where  $r$  denotes the number of atoms in the primitive cell of the crystal. As we will now show, such a mapping is achieved through the trace of the irregular part of the dynamical matrix.

Let us arrange the ensemble of frequencies, belonging to a specific  $\mathbf{q}$ -vector, in such a way that they can be considered elements  $\omega(\mathbf{q})$  of a  $3r$ -dimensional complex vector-space. Due to the irregularity of the dynamical matrix  $\omega(\mathbf{q})$  is not defined at  $\mathbf{q} = 0$ .

Considering first the case of non-polarizable ions, the dynamical matrix  $\mathbf{D}$  can be written as a superposition of a *mechanical* part, not involving long-range forces, and a *purely electrostatic* part. As the trace of the electrostatic part vanishes as a consequence of the Poisson equation [3], the modulus of  $\omega(\mathbf{q})$  only depends on the mechanical part of  $\mathbf{D}$  and therefore shows no singularity at the zone centre.

Separating the dynamical matrix  $\mathbf{D}$  into its regular and irregular parts, as outlined above, we define the frequency vector  $\omega^{\text{reg}}(\mathbf{q})$  which consists of the eigenvalues of only the regular part of  $\mathbf{D}$ . As this vector is well defined at  $\mathbf{q} = 0$  the following expression is a scalar quantity independent of  $\hat{\mathbf{q}}$  and can therefore serve as a measure of the LO-TO splitting:

$$\Omega^2 := \sum_{i=1}^{3r} (\omega_i(\mathbf{q})^2 - \omega_i^{\text{reg}}(\mathbf{q})^2) \quad |\mathbf{q}| \rightarrow 0. \quad (3)$$

This quantity, which is identical to the trace of the irregular part of the dynamical matrix, can be determined without diagonalization. An analytic treatment of  $\Omega^2$ , i.e. the LO-TO splitting, therefore becomes possible even for rather complicated structures. On the other hand  $\Omega^2$  is easily accessible experimentally.

$$\Omega^2 = \sum_{i=1}^{3(r-1)} (\omega_i^{\text{LO}}(\mathbf{q})^2 - \omega_i^{\text{TO}}(\mathbf{q})^2) \quad |\mathbf{q}| \rightarrow 0. \quad (4)$$

For its evaluation one simply has to determine the longitudinal and transverse optic frequencies for a chosen  $\mathbf{q}$  direction, e.g. by infrared spectroscopy‡.

† Note that we may always add a regular matrix to an irregular matrix and still end up with an irregular matrix. Calling  $\mathbf{C}^*$  the irregular part of the Coulomb matrix is therefore only one of many possibilities. From a physical point of view our choice (expression (2)) stands out as it can be cast into the form of a macroscopic field (see [2]). It is this later property of  $\mathbf{C}^*$  which motivates our definition.

‡ Another way of characterizing the LO-TO splitting for general structures is by means of the generalized Lyddane-Sachs-Teller relation [4,5]

$$\prod_{i=1}^{3(r-1)} (\omega_i^{\text{LO}}/\omega_i^{\text{TO}})^2 = \epsilon_0/\epsilon_\infty.$$

It is unfit for our purpose because it contains ratios of frequencies making it necessary to diagonalize the dynamical matrix.

As the above arguments are based on the harmonic and adiabatic approximation  $\Omega$  is independent of  $\hat{q}$  only if the optical modes show no damping. The damping of the modes can be determined by fitting classical oscillators to the infrared reflectance spectra of the crystal [6].

Damping is, however, not the only mechanism leading to a  $\hat{q}$  dependence of  $\Omega$ . In most crystals the main contribution to the  $\hat{q}$  dependence will be due to the polarizability of the ions. The polarizability couples the electrostatic forces to the short-range forces in such a way that it is impossible to write the dynamical matrix  $\mathbf{D}$  as a superposition of a mechanical and electrostatic part. Due to this coupling the short-range forces indirectly contribute to the irregular part of the dynamical matrix. As a consequence, not only the dynamical matrix  $\mathbf{D}$  but also the trace of  $\mathbf{D}$  becomes ill-defined at the zone-centre.

If  $\Omega$  as defined in (3) loses its scalar properties and becomes a function of  $\hat{q}$  it can no longer be employed as a measure of the LO-TO splitting. There is, however, a natural extension of (3) which is based on the fact that  $\Omega^2(\hat{q})$  is a quadratic form in the components of  $\mathbf{q}$  (see section 6):

$$\Omega^2(\hat{q}) = \frac{1}{q^2} \sum_{\alpha, \beta=1}^3 q_\alpha \Omega_{\alpha\beta}^2 q_\beta. \quad (5)$$

The tensor  $\Omega_{\alpha\beta}$  has the same symmetry as the dielectric constant. As the trace of this tensor does not depend on the choice of the coordinate system it can be used as a measure of the LO-TO splitting. It can be determined by applying (3) to the three orthogonal directions of  $\hat{q}$ :

$$\Omega^2 := \frac{1}{3} \sum_{\alpha=1}^3 \Omega_{\alpha\alpha}^2. \quad (6)$$

### 3. Determination of $\Omega$ for rigid-ion models

For rigid-ion models the irregular part of the dynamical matrix is given by

$$D_{\alpha, \beta}^{\text{irreg}}(\kappa, \kappa') = \frac{4\pi e^2}{v} \frac{Z(\kappa)Z(\kappa')}{\sqrt{M(\kappa)M(\kappa')}} \frac{q_\alpha q_\beta}{q^2} \quad (7)$$

and therefore

$$\Omega^2 = \frac{4\pi e^2}{v} \sum_{\kappa=1}^r \frac{Z(\kappa)^2}{M(\kappa)}. \quad (8)$$

If there are only two symmetrically distinct lattice sites, as in the case of quartz, we can rewrite (8) as follows:

$$\Omega^2 = 4\pi \frac{\xi^2}{\mu} \quad (9a)$$

with

$$\mu := \frac{1}{v} \frac{r_1 M_1 r_2 M_2}{r_1 M_1 + r_2 M_2} \quad (9b)$$

$$\xi := \frac{1}{v} r_1 |Z_1 e| = \frac{1}{v} r_2 |Z_2 e| \quad (9c)$$

and  $r_i$  denoting the multiplicity of the  $i$ th site in the primitive cell of the crystal.  $\mu$  is the reduced mass density of the system and  $\xi^2$  is proportional to the mean squared charge.

The result obtained for  $\Omega$  is identical to the one obtained for the plasma frequency of the free electron gas, if we replace the mass density of the electron by  $\mu$  and the electronic charge density by  $\xi$ . This is not surprising as we know that the frequencies of the transverse modes are zero in the case of a free electron gas, due to the absence of restoring forces for transverse motions. Therefore, in the case of a free electron gas,  $\Omega$  is identical to the frequency of the longitudinal vibration, i.e. the plasma frequency.

If  $\Omega$  is known experimentally we can immediately determine the ionic charge if we assume that our system can be described well by a rigid-ion model [7]. Using an alternative approach and making additional assumptions expression (8) has been used by Scott [8] as well as Gervais [9] to predict effective charges in ternary oxide compounds.

#### 4. Determination of $\Omega$ for shell models

To evaluate  $\Omega$  for shell models we have to isolate the irregular part of the dynamical matrix. Since this involves the inversion of the matrix  $\mathbf{K} + \mathbf{S} + \mathbf{YCY}$  (see (10)) describing the shell-shell interactions, it can in general not be done analytically.

$$\mathbf{MD} = (\mathbf{ZCZ} + \mathbf{R}) - (\mathbf{T} + \mathbf{ZCY})(\mathbf{K} + \mathbf{S} + \mathbf{YCY})^{-1}(\mathbf{T} + \mathbf{ZCY})^+ \quad (10)$$

Here  $\mathbf{M}$  stands for the mass tensor and  $\mathbf{R}$ ,  $\mathbf{T}$ ,  $\mathbf{S}$  for the core-core, core-shell and shell-shell short-range matrices respectively.  $\mathbf{Z}$  is the ionic and  $\mathbf{Y}$  the shell charge tensor;  $\mathbf{K}$  denotes the interaction of a core with its own shell. For a more detailed explanation of the single terms of (10) see [2] and in particular [10]. Special attention must be given to the correct inclusion of the respective self-terms.

An approximate expression in closed form for  $\Omega^2$  can be obtained for a large class of crystals defined by the following properties. (i) The structure possesses only two symmetrically distinct lattice sites. (ii) Only one ion type occupying these sites is polarizable and the effective polarizability of this ion type is not strongly anisotropic. We will give a precise mathematical definition of what we mean by effective polarizability in the following discussion.

Examples of crystals belonging to this class are quartz ( $\text{SiO}_2$ ) and sapphire ( $\text{Al}_2\text{O}_3$ ).

The analytic calculation of the trace of the irregular part of  $\mathbf{D}$ , for the above defined class of crystals, is made possible by the fact that the sums over matrix elements occurring in the course of the calculation can be related to the self-terms of the respective matrices. The self-terms in turn can be expressed in closed form as functions of the model parameters. The details of this procedure are rather involved and are therefore given in an appendix.

The final result is summarized in (11):

$$\Omega^2 = 4\pi \frac{\xi^2}{\mu} \left\{ 1 - \left[ \left( \frac{4\pi}{v} r_p + 2v \right) \alpha_{\text{eff}} - (\gamma \alpha_{\text{eff}})^2 \right] \left( \frac{1}{1 + (4\pi/v) r_p \alpha_{\text{eff}}} \right) \right\} \quad (11)$$

or in a more compact form

$$\Omega^2 = 4\pi \frac{\xi^2}{\mu} \left( \frac{1 - 2v\alpha_{\text{eff}} + \gamma^2 \alpha_{\text{eff}}^2}{1 + (4\pi/v) r_p \alpha_{\text{eff}}} \right) \quad (12)$$

The quantities  $\xi$  and  $\mu$  have already been defined in (9);  $r_p$  denotes the number of polarizable ions per primitive cell. The term  $\alpha_{\text{eff}}$  can be considered an effective polarizability and its connection with the free polarizability is given by

$$\alpha_{\text{eff}} := (\alpha^{-1} + \phi)^{-1} \quad (13a)$$

with

$$\phi := \frac{1}{3} \text{Tr}(\mathbf{Y}^{-1} [\mathbf{F}_0^T(\kappa, \kappa) + [\mathbf{Y}\bar{\mathbf{C}}\mathbf{X}]_0(\kappa, \kappa)] \mathbf{Y}^{-1}). \quad (13b)$$

The  $(3 \times 3)$  matrices  $\mathbf{F}_0^T(\kappa, \kappa)$  and  $[\mathbf{Y}\bar{\mathbf{C}}\mathbf{X}]_0(\kappa, \kappa)$  are the self-terms of the short-range and Coulomb shell-core interactions respectively.

The parameters  $\nu$  and  $\gamma$  are simple scalar functions related to the trace and quadratic norm of the self-terms of the shell-core interactions respectively (see (A37) and (A52)):

$$\nu := \frac{1}{3ZYe^2} \sum_{\kappa'\alpha} [T + X\bar{C}Y]_{\alpha\alpha}(\kappa, \kappa') \quad \kappa \in [r - r_p + 1, r] \quad (13c)$$

$$\gamma^2 := \frac{\mu}{3r\nu Y^2 e^2 \xi^2} \sum_{\alpha\beta\kappa} \left( \sum_{\kappa'} [T + X\bar{C}Y]_{\alpha\beta}(\kappa, \kappa') \right)^2 \quad (13d)$$

where  $Z$ ,  $Y$  and  $X$  are the ionic, shell and core charge of the polarizable ion. Note that the ions are labelled such that the ones belonging to the non-polarizable type come first;  $\nu$  and  $\gamma$  can easily be calculated knowing the Coulomb coefficients and short-range force constants. We want to point out that this is a necessary condition for the later application to concrete crystals.

When deriving expression (11), we have substituted the matrix  $\mathbf{Y}\mathbf{N}^{-1}\mathbf{Y}$ , within the subspace of the polarizable ions, by  $\alpha_{\text{eff}}$  multiplied by the unit matrix:

$$[\mathbf{Y}\mathbf{N}^{-1}\mathbf{Y}]_{\alpha\beta}(\kappa, \kappa') \mapsto \alpha_{\text{eff}} \delta_{\alpha\beta} \delta_{\kappa\kappa'}. \quad (14)$$

Expression (11) is therefore only exact if the matrix

$$\mathbf{N} = \mathbf{K} + \mathbf{F}_0^T + [\mathbf{Y}\bar{\mathbf{C}}\mathbf{X}]_0 \quad (15)$$

is isotropic in this subspace, as in this case

$$[\mathbf{Y}\mathbf{N}^{-1}\mathbf{Y}]_{\alpha\beta}(\kappa, \kappa') = \alpha_{\text{eff}} \delta_{\alpha\beta} \delta_{\kappa\kappa'}. \quad (16)$$

As long as the polarizability of the *hypothetically* free ions  $\alpha$  is small and not anisotropic the matrix  $\mathbf{N}$  will be dominated by the isotropic matrix  $\mathbf{K}$  and condition (16) can be assumed fulfilled in all cases. With increasing  $\alpha$  the off-diagonal elements of  $\mathbf{N}$  become more important and (11) will deviate more and more from the exact numerical values.

We now investigate briefly the single contributions to  $\Omega$ . The first contribution is identical to the one we obtained for rigid-ion models and just describes the unscreened LO-TO splitting arising from the ionic charges. The second term is, due to its sign, a pure screening term, i.e. reducing the LO-TO splitting. It depends only on the ionic charge and the effective polarizability.

The third contribution differs from the second only in that the factor  $r_p/\nu$  in the nominator is replaced by  $2\nu$ . In contrast to  $r_p/\nu$ , which is a constant wholly determined by the crystal structure,  $\nu$  depends on the interaction potentials through the short-range part of  $[\mathbf{F}^T + \mathbf{X}\bar{\mathbf{C}}\mathbf{Y}]_0$ . The long-range part gives only a constant contribution to  $\nu$ , determined by

the crystal structure through the Coulomb coefficients. If we compare different models for the same crystal, giving all comparable descriptions of the dispersion curves, it is reasonable to assume that the sum of the short-range forces in between the core of an unpolarizable ion and the shells of the polarizable ions scales approximately with the corresponding sum of the Coulomb forces, i.e. with  $Z \cdot Y$ . This is justified by the argument that short-range and Coulomb forces are more or less independent of each other as they have a different  $q$  dependence. In this case  $\nu$  does not depend on the model but only on the crystal structure and the ratio of short-range to long-range forces. The sign of  $\nu$  is not generally positive, so that we can not say *a priori* whether we are dealing with a screening or anti-screening contribution, i.e. with decreasing or increasing  $\Omega^2$ .

The above argument can be taken over directly for the final contribution. It implies that  $\gamma$  also is more or less independent from the specific model-setup. The main difference compared to the preceding contributions is its more complicated dependence on  $\alpha$ . Besides that, it stands out due to its sign:  $\gamma$ , being always positive, has an anti-screening effect, i.e. it increases  $\Omega^2$ .

### 5. Scaling of ionic charge with polarizability

Besides constants depending on the details of the crystal structure  $\Omega^2$ , as given by (11), is determined on the one hand by  $\phi$ ,  $\nu$  and  $\gamma$  and on the other hand by the polarizability  $\alpha$  and the ionic charge  $Z$ . As outlined above,  $\nu$  and  $\gamma$  cannot be changed appreciably without changing the description of the dispersion curves as they are determined by the ratio of short-range to Coulomb interactions in the system. Treating  $\phi$ ,  $\nu$  and  $\gamma$  as parameters and not as variables we invert

$$\Omega^2 = \Omega^2(Z, \alpha; \phi, \nu, \gamma) \quad (17)$$

to get

$$Z = Z(\alpha; \Omega_{\text{exp}}^2, \phi, \nu, \gamma) = Z(\alpha) \quad (18)$$

where  $\Omega_{\text{exp}}^2$  denotes the experimental value of  $\Omega^2$ . This means that we are looking for the pairs of  $Z$  and  $\alpha$  which give for a special set of  $\phi$ ,  $\nu$ ,  $\gamma$  the  $\Omega^2$  determined by experiment. As (18) involves only quantities defined by the crystal structure and the general ratio of short-range to Coulomb interactions it explains the high correlation of  $Z$  and  $\alpha$  observed when fitting shell models to experimental data. Examples of  $Z(\alpha)$  are given in figure 2. The calculation is based on the structure of  $\alpha$ -quartz [1] and the values of  $\phi$ ,  $\nu$  and  $\gamma$  are the ones found for the shell-models of  $\alpha$ -quartz (see table 1). All the curves show a more or less extended region where  $Z$  scales linearly with  $\alpha$ . The slope is given by the derivative of  $Z(\alpha)$  at  $\alpha = 0$ :

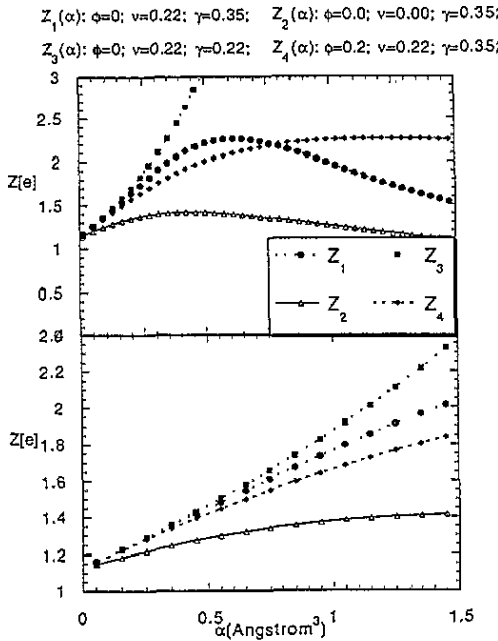
$$Z'(0) = \frac{1}{2}Z(0)[(4\pi/\nu)r_p + 2\nu]. \quad (19)$$

### 6. Connection with the effective charges

Using the dielectric function approach the irregular part of the dynamical matrix for a cubic system can be expressed in the form [11]

$$(MD)_{\alpha\beta}^{\text{irreg}}(\kappa, \kappa') = \frac{4\pi e^2}{\nu q^2} \left( \sum_{\gamma} q_{\gamma} Z_{\gamma\alpha}^{\text{eff}}(\kappa) \right)^* \epsilon_{\infty}^{-1}(\mathbf{q}, \mathbf{q}) \left( \sum_{\gamma} q_{\gamma} Z_{\gamma\beta}^{\text{eff}}(\kappa') \right). \quad (20)$$





**Figure 2.**  $Z(\alpha)$  for selected values of the parameters  $\phi$ ,  $\nu$  and  $\gamma$ . The parameters used for curve  $Z_4$  correspond to the lattice dynamical shell models of  $\alpha$ -quartz [1]. The region close to the origin is shown enlarged in the lower figure.

It has been shown by Sham [11] that the components of the charge tensor fulfill the following sum rule, which also constitutes a necessary condition for the existence of the acoustic branches:

$$\sum_{\kappa} Z_{\alpha\beta}^{\text{eff}}(\kappa) = 0. \quad (21)$$

Calculating the trace of (20) using (21) we get for binary crystals

$$\Omega^2 = \frac{4\pi e^2 r_p}{v^2 \mu} (Z^{\text{eff}})^2 \epsilon_{\infty}(q, q)^{-1} \quad (22)$$

with  $Z^{\text{eff}}$  the effective charge of the polarizable ion.

As  $\Omega$  and  $\epsilon$  are known experimentally (22) immediately yields the effective charge for binary systems.

For more complicated structures where  $\epsilon_{\alpha\beta}$  is not isotropic (22) has to be replaced by

$$\Omega^2 = \frac{4\pi e^2}{3v} \text{Tr}(\epsilon^{-1}(\infty)) \sum_{\kappa} \frac{1}{M(\kappa)} \sum_{\alpha\sigma} (Z_{\sigma\alpha}^{\text{eff}}(\kappa))^2. \quad (23)$$

Although the sum rule holds also for these crystals the trace cannot be expressed in the simple form of (22) due to the off-diagonal elements of the effective charges.

## 7. High-frequency dielectric constant

As the effective charges depend in a complicated way on the model parameters, (22) cannot be used to determine the relation of the high-frequency dielectric constant  $\epsilon(\infty)$  and the model parameters. As the high-frequency dielectric tensor for a shell model is given by

$$\epsilon_{\alpha\beta}(\infty) = 1 + \frac{4\pi e^2}{v} \sum_{\kappa\kappa'} [Y(S + Y\bar{C}Y)^{-1}Y]_{\alpha\beta}(\kappa\kappa') \quad (24)$$

it can, however be calculated directly, analogous to  $\Omega^2$ . If the substitution (14) is a valid approximation, this leads us to

$$\epsilon(\infty) = 1 + \frac{4\pi}{v} r_p \alpha_{\text{eff}}. \quad (25)$$

In the case of a cubic diatomic system with no short-range forces acting upon the shells the effective polarizability is given by  $\alpha_{\text{eff}} = (\alpha^{-1} - 4\pi/3v)^{-1}$ . Assuming that the shell charge  $Y$  is equal to the ionic charge  $Z$ , (25) reduces to the well known Clausius-Mossotti relation:

$$\frac{\epsilon(\infty) - 1}{\epsilon(\infty) + 2} = \frac{4\pi}{3v} \alpha. \quad (26)$$

As the effective polarizability is a scalar it is immediately obvious from (25) that in our approximation the dielectric tensor has to be isotropic. The experimentally determined anisotropy of  $\epsilon_\infty$  must therefore be very small for our results to be applicable. A zero anisotropy of  $\epsilon_\infty$ , on the other hand, is not sufficient to conclude that (14) is valid.

## 8. Application to quartz

We will now apply the above developed formalism to the lattice dynamics of  $\alpha$ -quartz.  $\alpha$ -quartz is a particularly good candidate as its dispersion relations can be very well described by various shell models differing appreciably in their short-range as well as Coulomb interactions.

In table 1 we list in columns 2-6 the values of  $\nu$ ,  $\phi$ ,  $\gamma$  together with the charge and free polarizability of the models. The rest of the parameters for the models SM(1) to SM(4) can be found in [1]. In columns 7 and 8 we compare the value of  $Z_a^{\text{RI}}(\Omega)$  as determined by using the analytical expression (11) to the exact value of  $Z_c^{\text{RI}}(\Omega)$  as determined after diagonalization of the dynamical matrix.  $Z^{\text{RI}}(\Omega)$  is defined through (9), i.e. it is the charge required to yield  $\Omega^2$  in the case of a rigid-ion model. In column 9 we give the value of  $\epsilon(\infty)$ .

Table 1. Parameters for the models SM(1) to SM(6).

| Model | $\nu(\text{\AA})^{-3}$ | $\phi(\text{\AA})^{-3}$ | $\gamma(\text{\AA})^{-3}$ | $Z(e)$ | $\alpha(\text{\AA})^3$ | $Z_a^{\text{RI}}(e)$ | $Z_c^{\text{RI}}(e)$ | $\epsilon(\infty)$ |
|-------|------------------------|-------------------------|---------------------------|--------|------------------------|----------------------|----------------------|--------------------|
| SM(1) | 0.21                   | -0.04                   | 0.44                      | 1.49   | 0.88                   | 1.04                 | 1.06                 | 1.62               |
| SM(2) | 0.23                   | 0.08                    | 0.44                      | 1.85   | 1.39                   | 1.17                 | 1.12                 | 1.96               |
| SM(3) | 0.22                   | 0.07                    | 0.41                      | 1.79   | 1.20                   | 1.15                 | 1.14                 | 1.85               |
| SM(4) | 0.22                   | 0.09                    | 0.41                      | 2.00   | 1.86                   | 1.17                 | 1.13                 | 2.36               |
| SM(5) | 0.29                   | 0.28                    | 0.66                      | 1.32   | 0.64                   | 1.04                 | 1.08                 | 1.39               |
| SM(6) | 0.22                   | 0.07                    | 0.42                      | 1.80   | 1.22                   | 1.15                 | 1.14                 | 1.86               |

By comparing columns 7 and 8 we realize that the analytical results are in close agreement with the numerical ones. This means that the error induced by (14) is very small and the results of sections 4-7 can be fully applied to the lattice dynamical investigations of  $\alpha$ -quartz.

Concerning the parameters  $\nu$ ,  $\phi$ ,  $\gamma$  we see that they are more or less model-independent as expected. Concentrating on  $\nu$  this implies that for all our models the short-range and Coulomb interactions between oxygen shells and silicon ions always contribute in the same ratio  $c_\nu$  to the trace of the self-term of the oxygen shell. This entails a strong correlation of

the product of ionic charge  $Z$  times shell charge  $Y$  on the one hand, and the longitudinal and transverse force constants in between silicon ions and oxygen shells on the other hand. Knowing the Coulomb coefficients for the quartz structure and averaging over all our models we get  $c_\nu = -1.92 \pm 0.05$ .

The fact that  $\nu$  and  $\gamma$  are more or less independent of the model and the rather small value of  $\phi$  explains the linear scaling of the ionic charge  $Z$  with the polarizability of the free ion  $\alpha$  (see figure 1) as outlined in section 5. The slope of  $Z(\alpha)$  is correctly predicted by (19), using the experimental value for  $Z^{\text{RI}} = 1.15e$  and the value of  $c_\nu$  given above. This means that the LO-TO splitting in  $\alpha$ -quartz is reproduced correctly only by shell models having a well defined  $Z(\alpha)$  determined basically by the constant  $c_\nu$ . All other parameters have a minor influence on  $\Omega^2$ . The interactions between the oxygen shells, for example, do not enter at all the analytic expression derived for  $\Omega^2$ .

The investigation of the LO-TO splitting alone is insufficient to predict the values of the ionic charge  $Z$  and polarizability  $\alpha$ . If we require, however, that the model also comes up with the correct value for the high-frequency dielectric constants we can determine both the ionic charge  $Z$  and polarizability  $\alpha$  through (12) and (25). With the exception of  $\epsilon(\infty)$  for  $\alpha = 1.86 \text{ \AA}^3$ , (25) is very well satisfied by the dielectric constants derived from our models, as can be seen in figure 3, even if we replace  $\alpha_{\text{eff}}$  by  $\alpha$ . The latter is not surprising given the small value of  $\phi$  (see (13)). Using (25) in conjunction with the experimental value for  $\epsilon(\infty) = 2.37$  [12] we predict that the effective polarizability  $\alpha_{\text{eff}}$  in  $\alpha$ -quartz has a value of  $1.8 \pm 0.2 \text{ \AA}^3$ . Taking the average value for  $\phi$  this corresponds to a polarizability  $\alpha$  of  $2.0 \pm 0.2 \text{ \AA}^3$ . Having determined the value of  $\alpha_{\text{eff}}$  we can use (12) to predict the charge of the oxygen ion. The value obtained lies very close to  $2e$ . It is actually the above analysis leading to model SM(4) that we consider to best describe the overall lattice dynamics of  $\alpha$ -quartz while at the same time reproducing the experimental value of the dielectric constant.

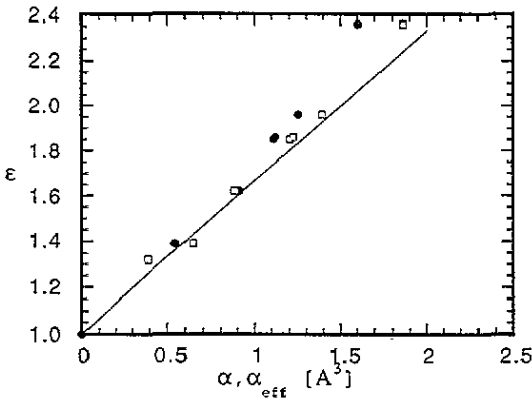


Figure 3.  $\epsilon_{zz}(\infty)$  as a function of  $\alpha_{\text{eff}}$  (full circles) and of  $\alpha$  (open squares) for the shell models of  $\alpha$ -quartz [1].  $\epsilon^\infty(\alpha_{\text{eff}})$  according to (25) is given by the full line.

Our results do not support the inclusion of an anisotropic free polarizability into our models as this would in most cases lead to an anisotropy in the dielectric tensor which is not observed experimentally. In fact, no substantial improvement of the fits could be achieved by allowing for an anisotropic free polarizability.

## 9. Application to sapphire

In the case of sapphire ( $\text{Al}_2\text{O}_3$ ) all the shell models we fitted to the experimental dispersion

curves gave about the same value for the ionic charge,  $Z = 1.5 \pm 0.1e$ , and free polarizability,  $\alpha = 0.7 \pm 0.1 \text{ \AA}^3$  [13]. These models do not, however, reproduce the experimental value of  $\epsilon_{xx}(\infty) = 3.2$  and  $\epsilon_{zz}(\infty) = 3.1$  [14].

By sacrificing a best fit to the dispersion curves we optimized various models with the value of the free polarizability of the oxygen ion  $\alpha$  fixed in between  $0.7 \text{ \AA}^3$  and  $2.0 \text{ \AA}^3$ . Investigating these models, we find that the anisotropy of the self-term  $F_0^T(\kappa, \kappa) + [\bar{Y}\bar{C}\bar{X}]_0(\kappa, \kappa)$  is very small for all of them and the parameters  $\nu$  and  $\gamma$  entering (12) are practically zero. The analytical results of sections 4–7 are therefore fully applicable to sapphire. The zero values for  $\nu$  and  $\gamma$  imply that the short-range contributions to the self-term arising from the interactions in between aluminium cores and oxygen shells cancel the corresponding Coulomb contributions. It should be noted that this cancellation is not required by symmetry and therefore reflects the distribution of forces acting in the crystal. Using the terminology of the preceding section, we get  $c_\nu = -1.0 \pm 0.05$  for sapphire in comparison to  $c_\nu = -1.92 \pm 0.05$  for  $\alpha$ -quartz. As  $\nu$  is vanishing,  $\phi$  is solely determined by the Coulomb coefficients ( $\phi = 0.2955X/Y$ ) and therefore negative as long as the core charge  $X$  of the oxygen ion has the opposite sign of the shell charge  $Y$ . This means according to (13) that the effective polarizability  $\alpha_{\text{eff}}$  is larger than the free polarizability  $\alpha$ , in contrast to what we have found in  $\alpha$ -quartz.

As can be seen from figure 4, (25) is well satisfied by the dielectric constants  $\epsilon(\infty)$  calculated from the models. In order for the models to reproduce the experimental value of  $\epsilon(\infty)$ , the effective polarizability  $\alpha_{\text{eff}}$  must be close to  $2.7 \pm 0.2 \text{ \AA}^3$ . Inserting this result into (14) and using the experimentally determined value of  $\Omega^2$  we get for the ionic charge of the oxygen ion in sapphire  $Z = 2.0 \pm 0.2e$ . As the shell charge for all our models  $Y = 2.9 \pm 0.1e$  we can also determine the free polarizability of the oxygen ion required to reproduce both the right magnitude of the LO-TO splitting as well as the right dielectric constant to be  $\alpha = 1.9 \pm 0.2 \text{ \AA}^3$ .

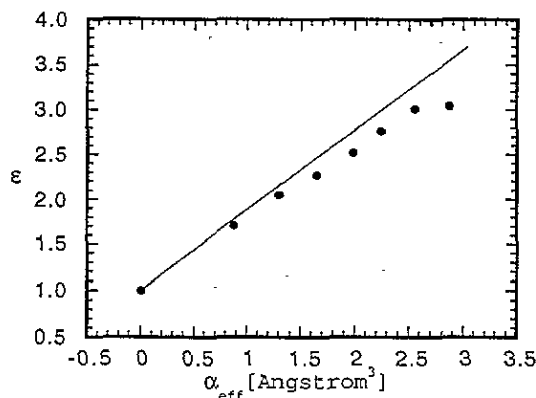


Figure 4.  $\epsilon_{zz}(\infty)$  as a function of  $\alpha_{\text{eff}}$  for the shell models of sapphire [13].  $\epsilon_{\infty}(\alpha_{\text{eff}})$  according to (25) is given by the full line.

The results for the free polarizability and the ionic charge of the oxygen ion in sapphire are identical to the ones obtained for  $\alpha$ -quartz. The large difference in the dielectric constants of the two crystals is solely due to the different structures and short-range interionic potentials. These results compare also very well with other oxides, like the perovskites (see [15] and references therein).

## 10. Conclusion

We have shown that an analytical treatment of the LO-TO splitting and the high-frequency dielectric constants in complex crystals like  $\alpha$ -quartz and sapphire is possible and constitutes a powerful tool to extract information about charges and polarizabilities as well as the correlations of the shell model parameters. This information is the more valuable as it is not based on a single lattice dynamical model but on their common characteristics.

It turns out that the self-term associated with the shell-core interactions plays the dominant role in the determination of the LO-TO splitting.

In this way we have found that both in sapphire as well as in  $\alpha$ -quartz the ionic charge of the oxygen ion is very close to  $2e$  paired with a free polarizability of about  $1.9\text{\AA}^3$ . The charge as well as the polarizability are strongly correlated with the model parameters entering the self-term of the core-shell interactions and an analytical expression for this correlation is given. The results are based on the assumption that the lattice dynamics of quartz and sapphire can be described by shell models, however, without reference to any particular model.

## Acknowledgments

We are grateful to Dr G Eckold for sending a version of the UNISOFT computer program package [10, 16] prior to publication.

## Appendix. Derivation of (11)

In order to facilitate notation we rewrite the dynamical matrix given in (10) in the form

$$\mathbf{D} = (\mathbf{Z}\mathbf{C}^*\mathbf{Z} + \mathbf{R}) - (\mathbf{T} + \mathbf{Z}\mathbf{C}^*\mathbf{Y})[\mathbf{K} + \mathbf{S} + \mathbf{Y}\mathbf{C}^*\mathbf{Y}]^{-1}(\mathbf{T} + \mathbf{Z}\mathbf{C}^*\mathbf{Y})^+ \quad (\text{A1})$$

where we have redefined the matrixes  $\mathbf{R}$ ,  $\mathbf{T}$  and  $\mathbf{S}$  by including the regular part of the corresponding Coulomb interactions:

$$\mathbf{R} := \mathbf{R} + \mathbf{Z}\bar{\mathbf{C}}\mathbf{Z} \quad (\text{A1a})$$

$$\mathbf{T} := \mathbf{T} + \mathbf{X}\bar{\mathbf{C}}\mathbf{Y} \quad (\text{A1b})$$

$$\mathbf{S} := \mathbf{S} + \mathbf{Y}\bar{\mathbf{C}}\mathbf{Y} \quad (\text{A1c})$$

where  $\bar{\mathbf{C}}$  denotes the regular and  $\mathbf{Y}\mathbf{C}^*$  the irregular part of the Coulomb matrix. The gauge we choose such that the regular part of the dynamical matrix is given by

$$\mathbf{D}^{\text{reg}} = \mathbf{R} - \mathbf{T}(\mathbf{K} + \mathbf{S})^{-1}\mathbf{T}^+ \quad (\text{A2})$$

as only in this case is (3) is identical to (4). Instead of treating the irregular part of the dynamical matrix as a whole we split it up into three parts defined as follows:

$$\mathbf{D}^{\text{I}} := \mathbf{Z}\mathbf{C}^*\mathbf{Z} - (\mathbf{Z}\mathbf{C}^*\mathbf{Y})[\mathbf{K} + \mathbf{S} + \mathbf{Y}\mathbf{C}^*\mathbf{Y}]^{-1}(\mathbf{Y}\mathbf{C}^*\mathbf{Z}) \quad (\text{A3a})$$

$$\mathbf{D}^{\text{II}} := -\mathbf{T}[\mathbf{K} + \mathbf{S} + \mathbf{Y}\mathbf{C}^*\mathbf{Y}]^{-1}(\mathbf{Y}\mathbf{C}^*\mathbf{Z}) + \text{HC} \quad (\text{A3b})$$

$$\mathbf{D}^{\text{III}} := -\mathbf{T}[(\mathbf{K} + \mathbf{S} + \mathbf{Y}\mathbf{C}^*\mathbf{Y})^{-1} - (\mathbf{K} + \mathbf{S})^{-1}]\mathbf{T}^+. \quad (\text{A3c})$$

The calculation of the trace of these expressions can generally not be done analytically as it involves the inversion of  $(\mathbf{K} + \mathbf{S} + \mathbf{Y}\mathbf{C}^*\mathbf{Y})$ . We will show that under certain conditions in order to calculate the trace of expressions (A3) the matrix  $(\mathbf{K} + \mathbf{S} + \mathbf{Y}\mathbf{C}^*\mathbf{Y})^{-1}$  can simply be replaced by  $(\mathbf{N} + \mathbf{Y}\mathbf{C}^*\mathbf{Y})^{-1}$  and that the matrix  $(\mathbf{N} + \mathbf{Y}\mathbf{C}^*\mathbf{Y})^{-1}$  can be calculated in closed form:

$$\mathbf{N} = \mathbf{K} + \mathbf{F}_0^T + [\mathbf{Y}\bar{\mathbf{C}}\mathbf{X}]_0 \tag{A4}$$

with

$$[\mathbf{Y}\bar{\mathbf{C}}\mathbf{X}]_{0\alpha,\beta}(\kappa, \kappa') = -\delta(\kappa, \kappa') \sum_{(\kappa'', l'') \neq (\kappa, l)} Y(\kappa) \Phi_{\alpha,\beta}^{\bar{\mathbf{C}}}(\kappa, \kappa'', l, l'') X(\kappa''). \tag{A5}$$

where  $\mathbf{F}_0^T$  denotes the self-term of the short-range interactions between shells and cores.

The starting point of our calculations is the identity

$$\begin{aligned} (\mathbf{K} + \mathbf{S} + \mathbf{Y}\mathbf{C}^*\mathbf{Y})^{-1} &= (\mathbf{N} + \mathbf{S}^S + \mathbf{Y}\mathbf{C}^*\mathbf{Y})^{-1} \\ &= (\mathbf{N} + \mathbf{Y}\mathbf{C}^*\mathbf{Y})^{-1} - (\mathbf{N} + \mathbf{Y}\mathbf{C}^*\mathbf{Y})^{-1} \mathbf{S}^S [\mathbf{N} + \mathbf{Y}\mathbf{C}^*\mathbf{Y} + \mathbf{S}^S]^{-1} \end{aligned} \tag{A6}$$

with

$$\mathbf{S}^S := \mathbf{S} - \mathbf{F}_0^T - [\mathbf{Y}\bar{\mathbf{C}}\mathbf{X}]_0. \tag{A7}$$

Let us first investigate the matrix  $(\mathbf{N} + \mathbf{Y}\mathbf{C}^*\mathbf{Y})^{-1}$ . As we are treating only the case of one polarizable ion type  $(\mathbf{N} + \mathbf{Y}\mathbf{C}^*\mathbf{Y})^{-1}$  is of the form

$$(\mathbf{N} + \mathbf{Y}\mathbf{C}^*\mathbf{Y})^{-1} = \begin{pmatrix} 0 & 0 \\ 0 & \mathbf{M}^{-1} \end{pmatrix} \tag{A8}$$

with  $\mathbf{M}^{-1}$  being a matrix of rank  $r_p$ . It is therefore sufficient to consider only  $\mathbf{M}^{-1}$ . If the terms in (A4) are isotropic, and if we do not include an anisotropic free polarizability in our models, then

$$\mathbf{Y}^{-1} \mathbf{N} \mathbf{Y}^{-1} = \begin{pmatrix} 0 & 0 \\ 0 & x^{-l} \mathbf{E}_{3r_p} \end{pmatrix} \tag{A9}$$

where we denote the unit matrix of rank  $l$  by  $\mathbf{E}_l$ .

As the elements of  $\mathbf{Y}\mathbf{C}^*$

$$a := C_{\alpha\beta}^*(\kappa, \kappa') = \frac{4\pi e^2}{3v} \tag{A10}$$

do not depend on  $\kappa$  it is straightforward to calculate  $\mathbf{M}^{-1}$ :

$$M_{i,j}^{-1} = \frac{1}{\text{Det}(\mathbf{M})} (-1)^{i+j} \text{Det}(\mathbf{M}_{j,i}) \tag{A11}$$

with

$$\text{Det}(\mathbf{M}) = x^{-3r_p} + 3r_p a x^{-3r_p+1} \tag{A12a}$$

$$\text{Det}(\mathbf{M}_{ii}) = x^{-3r_p+1} + 3(r_p - 1) a x^{-3r_p+2} \tag{A12b}$$

$$\text{Det}(\mathbf{M}_{ij}) = (-1)^{i+j+1} a x^{-3r_p+2} \quad i \neq j \tag{A12c}$$

leading to

$$\begin{aligned}
 \mathbf{M}^{-1} &= \frac{1}{x^{-3r_p} + 3r_p a x^{-3r_p+1}} \\
 &\times \begin{pmatrix} x^{-3r_p+1} + (3r_p - 1)ax^{-3r_p+2} & -ax^{-3r_p+2} & & -ax^{-3r_p+2} \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & x^{-3r_p+1} + (3r_p - 1)ax^{-3r_p+2} \end{pmatrix} \\
 &= \frac{1}{1 + r_p a x} \begin{pmatrix} x + (3r_p - 1)ax^2 & & & -ax^2 \\ -ax^2 & x + (3r_p - 1)ax^2 & & -ax^2 \\ \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & x + (3r_p - 1)ax^2 \end{pmatrix}
 \end{aligned}
 \tag{A13}$$

where for convenience we have chosen a coordinate system such that  $\hat{q} = 1/\sqrt{3}(1, 1, 1)$ .

If  $\mathbf{N}$  is not of the form (A9), the result can be generalized in such a way that the form of  $\mathbf{M}^{-1}$  is preserved, i.e.  $\mathbf{M}^{-1}$  can still be written as

$$\mathbf{M}^{-1} = \mathbf{X} + \mathbf{H}
 \tag{A14}$$

with all the elements of the  $(3r_p \times 3r_p)$ -matrix  $\mathbf{H}$  being identical and with  $\mathbf{X}$  being diagonal in  $\kappa$ . The  $(3 \times 3)$ -submatrices of  $\mathbf{X}$  are, however, no longer multiples of the unit matrix as is the case for expression (A13).

We will use (A13) to calculate the trace of  $\mathbf{D}^I$ , which for this purpose we transform as follows:

$$\begin{aligned}
 \mathbf{D}^I &= \mathbf{ZC}^* \mathbf{Z} - (\mathbf{ZC}^* \mathbf{Y})(\mathbf{N} + \mathbf{YC}^* \mathbf{Y})^{-1} \mathbf{YC}^* \mathbf{Z} \\
 &\quad + (\mathbf{ZC}^* \mathbf{Y})(\mathbf{N} + \mathbf{YC}^* \mathbf{Y})^{-1} \mathbf{S}^S [\mathbf{N} + \mathbf{YC}^* \mathbf{Y} + \mathbf{S}^S]^{-1} \mathbf{YC}^* \mathbf{Z}.
 \end{aligned}
 \tag{A15}$$

As the elements of  $\mathbf{YC}^*$  do not depend upon  $\kappa$  the product  $\mathbf{YC}^* \mathbf{A}$  is zero for all  $(3r \times 3r)$ -matrices  $\mathbf{A}$  fulfilling the condition

$$\sum_{\kappa'=1}^r A_{\alpha\beta}(\kappa, \kappa') = 0.
 \tag{A16}$$

Therefore

$$\mathbf{YC}^* \mathbf{Y} \mathbf{S}^S = 0.
 \tag{A17}$$

In order to apply (A17) to (A15) we have to commute  $(\mathbf{N} + \mathbf{YC}^* \mathbf{Y})^{-1}$  either with  $(\mathbf{ZC}^* \mathbf{Y})$  or with  $\mathbf{S}^S$ . As  $\mathbf{YC}^* \mathbf{Z}$  is of the form

$$\mathbf{YC}^* \mathbf{Z} = \begin{pmatrix} 0 & 0 \\ \mathbf{B}_1 & \mathbf{B}_2 \end{pmatrix}
 \tag{A18}$$

with both  $\mathbf{B}_1$  as well as  $\mathbf{B}_2$  not equal to zero it does not commute with  $(\mathbf{N} + \mathbf{YC}^* \mathbf{Y})^{-1}$ . We are therefore obliged to calculate the commutator

$$\mathbf{C}_{SN} = [\mathbf{S}^S, (\mathbf{N} + \mathbf{YC}^* \mathbf{Y})^{-1}].
 \tag{A19}$$

As

$$\mathbf{S}^S = \begin{pmatrix} 0 & 0 \\ 0 & \mathbf{S}^S \end{pmatrix}
 \tag{A20}$$

where  $\mathbf{S}^S$  commutes with  $\mathbf{N}^{-1}$ , if  $\mathbf{S}^S$  commutes with  $\mathbf{M}^{-1}$ . Using (A14), and knowing that  $\mathbf{H}$  commutes with any symmetrical matrix, we get

$$\mathbf{C}_{SN} = [\mathbf{S}^S, \mathbf{X}]. \tag{A21}$$

As  $\mathbf{S}^S$  depends on the specific model used  $\mathbf{C}_{SN}$  in general is only zero if  $\mathbf{X}$  is a multiple of unity due to the symmetry of the crystal. In all other cases we get a result in closed form only by working in the following approximation.

Let

$$\alpha_{\text{eff}}^{-1} := \frac{1}{3r_p} \text{Tr}(\mathbf{Y}^{-1} \mathbf{M} \mathbf{Y}^{-1}). \tag{A22}$$

As  $\alpha_{\text{eff}}^{-1}$  does not depend on  $\kappa$

$$\alpha_{\text{eff}}^{-1} = \frac{1}{3} \text{Tr}(\mathbf{Y}^{-1} \mathbf{M}(\kappa, \kappa) \mathbf{Y}^{-1}) = (\alpha^{-1} + \phi) \tag{A23}$$

with an arbitrary  $\kappa$  chosen among  $[r - r_p, r]$ ;  $\phi$  is given by

$$\phi := \frac{1}{3} \text{Tr}(\mathbf{Y}^{-1} [\mathbf{F}_0^T(\kappa, \kappa) + [\mathbf{Y} \bar{\mathbf{C}} \mathbf{X}]_0(\kappa, \kappa)] \mathbf{Y}^{-1}). \tag{A24}$$

By replacing  $\mathbf{N}$  with  $\mathbf{Y} \alpha_{\text{eff}}^{-1} \mathbf{Y}$  the matrix  $(\mathbf{N} + \mathbf{Y} \mathbf{C}^* \mathbf{Y})^{-1}$  is given by expression (A13) if we identify  $x$  with  $\alpha_{\text{eff}}$ . In principle we can always choose an anisotropic free polarizability  $\alpha'_{\alpha\beta}(\kappa)$  such that

$$\mathbf{Y} \alpha'^{-1} \mathbf{Y} + \mathbf{N} = \mathbf{Y} \alpha_{\text{eff}}^{-1} \mathbf{E}_{3r} \mathbf{Y}. \tag{A25}$$

For this kind of model the analytical expressions we are going to derive are correct.

Summarizing our results so far, we have shown that the third term of (A15)

$$(\mathbf{Z} \mathbf{C}^* \mathbf{Y})(\mathbf{N} + \mathbf{Y} \mathbf{C}^* \mathbf{Y})^{-1} \mathbf{S}^S [\mathbf{N} + \mathbf{Y} \mathbf{C}^* \mathbf{Y} + \mathbf{S}^S \mathbf{J}]^{-1} \mathbf{Y} \mathbf{C}^* \mathbf{Z} \tag{A26}$$

vanishes if  $\mathbf{C}_{SN}$  is zero. As the first term

$$\mathbf{Z} \mathbf{C}^* \mathbf{Z}_{\alpha,\beta}(\kappa, \kappa') = \frac{4\pi e^2}{v} \frac{Z(\kappa) Z(\kappa')}{\sqrt{M(\kappa) M(\kappa')}} \frac{q_\alpha q_\beta}{q^2} \tag{A27}$$

is known explicitly we are left with the expression

$$(\mathbf{Z} \mathbf{C}^* \mathbf{Y})(\mathbf{N} + \mathbf{Y} \mathbf{C}^* \mathbf{Y})^{-1} \mathbf{Y} \mathbf{C}^* \mathbf{Z}. \tag{A28}$$

As in the chosen coordinate system the elements of  $\mathbf{C}^*$  are all identical to  $a$  (A10)

$$\mathbf{C}^* \mathbf{A} \mathbf{C}^* = a^2 y \mathbf{H} \tag{A29}$$

with

$$y = \sum_{i,j} A_{i,j}. \tag{A30}$$

Taking into consideration the special form of  $\mathbf{Y} \mathbf{C}^* \mathbf{Z}$  (A18) as well as the atomic masses we get for (A28) with the help of (A13) and (A29)

$$\begin{aligned} D_{\alpha\beta}^I(\kappa, \kappa') &= \frac{4\pi e^2}{9v} \left( \frac{Z(\kappa) Z(\kappa')}{\sqrt{M(\kappa) M(\kappa')}} - (4\pi/v) \sum_{\kappa''} \frac{Z(\kappa)}{\sqrt{M(\kappa)}} \frac{\alpha_{\text{eff}}}{1 + \frac{4\pi}{v} r_p \alpha_{\text{eff}}} \frac{Z(\kappa')}{\sqrt{M(\kappa')}} \right) \\ &= \frac{4\pi e^2}{9v} \frac{Z(\kappa) Z(\kappa')}{\sqrt{M(\kappa) M(\kappa')}} \left( 1 - \frac{(4\pi/v) r_p \alpha_{\text{eff}}}{1 + (4\pi/v) r_p \alpha_{\text{eff}}} \right). \end{aligned} \tag{A31}$$



Calculating the trace yields

$$\Omega_I^2 = 4\pi \frac{\xi^2}{\mu} \left( 1 - \frac{(4\pi/v)r_p \alpha_{\text{eff}}}{1 + (4\pi/v)r_p \alpha_{\text{eff}}} \right). \quad (\text{A32})$$

The second contribution to the irregular part of the dynamical matrix, i.e. (A3b), in an analogous fashion to (A3a) can be cast into the form

$$\mathbf{D}^{\text{II}} = -\mathbf{T}[\mathbf{N} + \mathbf{Y}\mathbf{C}^*\mathbf{Y}]^{-1}(\mathbf{Y}\mathbf{C}^*\mathbf{Z}) + \text{H.C.} \quad (\text{A33})$$

Calculating the product of the two matrices to the right and using of (A13) and (A18) we get

$$[(\mathbf{N} + \mathbf{Y}\mathbf{C}^*\mathbf{Y})^{-1}(\mathbf{Y}\mathbf{C}^*\mathbf{Z})]_{\alpha\beta}(\kappa, \kappa') = \aleph(\kappa)Z(\kappa')\delta_{\alpha\kappa}\delta_{\beta\kappa'} \quad (\text{A34})$$

with

$$\begin{aligned} \aleph(\kappa) &= 0 & \kappa \in [1, r - r_p] \\ \aleph(\kappa) &= \frac{(4\pi/v)\alpha_{\text{eff}}}{1 + (4\pi/v)r_p \alpha_{\text{eff}}} Y(\kappa)^{-1} & \kappa \in [r - r_p + 1, r]. \end{aligned} \quad (\text{A35})$$

where we have changed the coordinate system such that  $\hat{q} = \hat{z}$ . Equation (A35) substituted into (A33) leaves us with

$$D_{\alpha\beta}^{\text{II}}(\kappa, \kappa') = -\frac{4\pi}{v} \left( Z(\kappa') \sum_{\kappa''} T_{\alpha\kappa}(\kappa, \kappa'') \aleph(\kappa'') \delta_{z\beta} \right) + \text{H.C.} \quad (\text{A36})$$

As we are only interested in the trace of  $\mathbf{D}^{\text{II}}$  it is sufficient to carry out the summation in (A36) for the diagonal elements. This task is again facilitated by the translational symmetry of the system. Let us first investigate the case when  $\kappa$  represents an ion of type I. As the ions of type I are non-polarizable they possess no shells and the sum in (A36) involves only the interactions of the cores of type I with the shells of type II. The sum in (A36) is therefore proportional to the negative self-term of  $\mathbf{F}^{\text{T}} + \mathbf{X}\mathbf{C}\mathbf{Y}$  for this special  $\kappa$ . In the case of  $\kappa$  representing an ion of type II there are contributions to the sum from the interactions of the corresponding core with all the shells, plus the interactions of the corresponding shell with the rest of the shells. But as in this case the summation also covers the diagonal element  $T_{\alpha\alpha}(\kappa\kappa)$  all the off-diagonal contributions to the sum are cancelled by the corresponding ones of the self-term  $[\mathbf{F}^{\text{T}} + \mathbf{Y}\mathbf{C}\mathbf{X}]_0$ . Only the contributions to  $[\mathbf{F}^{\text{T}} + \mathbf{Y}\mathbf{C}\mathbf{X}]_0$  arising from the interaction of the shell  $\kappa$  with the cores of all the non-polarizable ions remain.

As the trace

$$\text{Tr} \sum_{\kappa'} T(\kappa, \kappa') = \sum_{\kappa'\alpha} T_{\alpha,\alpha}(\kappa, \kappa')$$

is invariant under orthogonal transformations it depends only on the ion type but not on  $\kappa$ . This allows us to define a constant  $\nu$  such that

$$\frac{1}{3} \sum_{\kappa'\alpha} T_{\alpha\alpha}(\kappa, \kappa') =: -\frac{r_p}{r - r_p} ZY e^2 \nu \quad \kappa \in [1, r - r_p] \quad (\text{A37a})$$

$$\frac{1}{3} \sum_{\kappa'\alpha} T_{\alpha\alpha}(\kappa, \kappa') =: ZY e^2 \nu \quad \kappa \in [r - r_p + 1, r]. \quad (\text{A37b})$$

Here  $Z$  and  $Y$  are the ionic and shell charge of the polarizable ion respectively. The definition of  $\nu$  through (A37a) and (A37b) is unique due to the identity

$$\sum_{\kappa=1}^{r-r_p} \sum_{\kappa'=r-r_p+1}^r T_{\alpha\beta}(\kappa, \kappa') = \sum_{\kappa=r-r_p+1}^r \sum_{\kappa'=1}^{r-r_p} T_{\beta\alpha}(\kappa, \kappa'). \quad (\text{A37c})$$

Averaging over the cartesian coordinates leads to the final result

$$\Omega_{\text{II}}^2 = 4\pi \frac{\xi^2}{\mu} \left( \frac{-2\nu\alpha_{\text{eff}}}{1 + (4\pi/\nu)r_p\alpha_{\text{eff}}} \right). \quad (\text{A38})$$

In order to calculate the third contribution to  $\Omega$

$$\mathbf{D}^{\text{III}} = \mathbf{T}[(\mathbf{K} + \mathbf{S} + \mathbf{Y}\mathbf{C}^*\mathbf{Y})^{-1} - (\mathbf{K} + \mathbf{S})^{-1}]\mathbf{T}^{\text{III}} \quad (\text{A39})$$

we use an identity analogous to (A6):

$$(\mathbf{K} + \mathbf{S} + \mathbf{Y}\mathbf{C}^*\mathbf{Y})^{-1} = (\mathbf{N} + \mathbf{S}^{\text{S}} + \mathbf{Y}\mathbf{C}^*\mathbf{Y})^{-1} = \mathbf{N}^{-1} - \mathbf{N}^{-1}(\mathbf{S}^{\text{S}} + \mathbf{Y}\mathbf{C}^*\mathbf{Y})[\mathbf{N} + \mathbf{Y}\mathbf{C}^*\mathbf{Y} + \mathbf{S}^{\text{S}}]^{-1}. \quad (\text{A40})$$

Equation (A6) applied to (A40) gives

$$(\mathbf{K} + \mathbf{S} + \mathbf{Y}\mathbf{C}^*\mathbf{Y})^{-1} = -\mathbf{N}^{-1}(\mathbf{S}^{\text{S}} + \mathbf{Y}\mathbf{C}^*\mathbf{Y})(\mathbf{N} + \mathbf{Y}\mathbf{C}^*\mathbf{Y})^{-1} + \mathbf{N}^{-1}(\mathbf{S}^{\text{S}} + \mathbf{Y}\mathbf{C}^*\mathbf{Y})(\mathbf{N} + \mathbf{Y}\mathbf{C}^*\mathbf{Y})^{-1} \times \mathbf{S}^{\text{S}}[\mathbf{N} + \mathbf{Y}\mathbf{C}^*\mathbf{Y} + \mathbf{S}^{\text{S}}]^{-1} + \mathbf{N}^{-1}. \quad (\text{A41})$$

If  $\mathbf{C}_{\text{SN}}$  is zero this expression can be simplified using (A17):

$$(\mathbf{K} + \mathbf{S} + \mathbf{Y}\mathbf{C}^*\mathbf{Y})^{-1} = -\mathbf{N}^{-1}(\mathbf{S}^{\text{S}} + \mathbf{Y}\mathbf{C}^*\mathbf{Y})(\mathbf{N} + \mathbf{Y}\mathbf{C}^*\mathbf{Y})^{-1} + \mathbf{N}^{-1}\mathbf{S}^{\text{S}}(\mathbf{N} + \mathbf{Y}\mathbf{C}^*\mathbf{Y})^{-1}\mathbf{S}^{\text{S}}[\mathbf{N} + \mathbf{Y}\mathbf{C}^*\mathbf{Y} + \mathbf{S}^{\text{S}}]^{-1} + \mathbf{N}^{-1}. \quad (\text{A42})$$

Applying again (A6) we find

$$(\mathbf{K} + \mathbf{S} + \mathbf{Y}\mathbf{C}^*\mathbf{Y})^{-1} = -\mathbf{N}^{-1}(\mathbf{Y}\mathbf{C}^*\mathbf{Y})(\mathbf{N} + \mathbf{Y}\mathbf{C}^*\mathbf{Y})^{-1} + \mathbf{N}^{-1}\mathbf{S}^{\text{S}}[\mathbf{N} + \mathbf{Y}\mathbf{C}^*\mathbf{Y} + \mathbf{S}^{\text{S}}]^{-1} + \mathbf{N}^{-1}. \quad (\text{A43})$$

Transforming the second part of (A43) according to

$$(\mathbf{N} + \mathbf{Y}\mathbf{C}^*\mathbf{Y} + \mathbf{S}^{\text{S}})^{-1} = (\mathbf{N} + \mathbf{S}^{\text{S}})^{-1} + [(\mathbf{N} + \mathbf{Y}\mathbf{C}^*\mathbf{Y} + \mathbf{S}^{\text{S}})^{-1} - (\mathbf{N} + \mathbf{S}^{\text{S}})^{-1}] \quad (\text{A44})$$

and using

$$(\mathbf{N} + \mathbf{Y}\mathbf{C}^*\mathbf{Y} + \mathbf{S}^{\text{S}})^{-1} - (\mathbf{N} + \mathbf{S}^{\text{S}})^{-1} = -(\mathbf{Y}\mathbf{C}^*\mathbf{Y})[(\mathbf{N} + \mathbf{S}^{\text{S}})(\mathbf{N} + \mathbf{Y}\mathbf{C}^*\mathbf{Y} + \mathbf{S}^{\text{S}})]^{-1} \quad (\text{A45})$$

together with (A17) leads to

$$\mathbf{N}^{-1} + \mathbf{N}^{-1}\mathbf{S}^{\text{S}}[\mathbf{N} + \mathbf{Y}\mathbf{C}^*\mathbf{Y} + \mathbf{S}^{\text{S}}]^{-1} = \mathbf{N}^{-1} - \mathbf{N}^{-1}\mathbf{S}^{\text{S}}[\mathbf{N} + \mathbf{S}^{\text{S}}]^{-1} = (\mathbf{N} + \mathbf{S}^{\text{S}})^{-1} = (\mathbf{K} + \mathbf{S})^{-1}. \quad (\text{A46})$$

Therefore

$$(\mathbf{K} + \mathbf{S} + \mathbf{Y}\mathbf{C}^*\mathbf{Y})^{-1} = -\mathbf{N}^{-1}(\mathbf{Y}\mathbf{C}^*\mathbf{Y})(\mathbf{N} + \mathbf{Y}\mathbf{C}^*\mathbf{Y})^{-1} + (\mathbf{K} + \mathbf{S})^{-1} \quad (\text{A47})$$

substituted into (A39) gives

$$\mathbf{D}^{\text{III}} = -\mathbf{T}\mathbf{N}^{-1}(\mathbf{Y}\mathbf{C}^*\mathbf{Y})(\mathbf{N} + \mathbf{Y}\mathbf{C}^*\mathbf{Y})^{-1}\mathbf{T}^+. \quad (\text{A48})$$

With the help of (A34) we calculate

$$[N^{-1}(\mathbf{Y}\mathbf{C}^*\mathbf{Y})(\mathbf{N} + \mathbf{Y}\mathbf{C}^*\mathbf{Y})^{-1}]_{\alpha\beta}(\kappa, \kappa') = \sqrt{\aleph(\kappa)\aleph(\kappa')} \delta_{\alpha z} \delta_{\beta z} \quad (\text{A49})$$

with now

$$\begin{aligned} \aleph(\kappa) &:= \frac{(4\pi/v)\alpha_{\text{eff}}^2}{1 + (4\pi/v)r_p\alpha_{\text{eff}}} Y^{-2} & \kappa \in [r - r_p + 1, r] \\ \aleph(\kappa) &:= 0 & \text{otherwise.} \end{aligned} \quad (\text{A50})$$

For the elements of the matrix (A48) we get in the coordinate system with  $\hat{q} = \hat{z}$

$$\left( \sum_{\kappa''=1}^r T_{\alpha z}(\kappa, \kappa'') \sqrt{\aleph(\kappa'')} \right) \left( \sum_{\kappa''=1}^r \sqrt{\aleph(\kappa'')} T_{z\beta}^+(\kappa'', \kappa') \right). \quad (\text{A51})$$

The terms arising when calculating the trace of (A51) are therefore proportional to

$$\gamma_1^2 := \frac{1}{3Z^2 Y^2 e^4} \sum_{\alpha\beta} \left( \sum_{\kappa'} T_{\alpha\beta}(\kappa, \kappa') \right)^2 \quad \kappa \in [1, r - r_p] \quad (\text{A52a})$$

$$\gamma_2^2 := \frac{1}{3Z^2 Y^2 e^4} \sum_{\alpha\beta} \left( \sum_{\kappa'} T_{\alpha\beta}(\kappa, \kappa') \right)^2 \quad \kappa \in [r - r_p + 1, r] \quad (\text{A52b})$$

where we have again averaged over three orthogonal directions;  $\gamma_1$  and  $\gamma_2$  do not depend on  $\kappa$  as the sums on the right are nothing else but the squares of the canonical norm of  $(3 \times 3)$  matrices and are therefore invariant under orthogonal transformations. Note that the ratio of  $\gamma_1$  to  $\gamma_2$  depends on the specific setup of the model. To get a formally similar result to  $\Omega_{\text{I}}$  and  $\Omega_{\text{II}}$  we define  $\gamma = f(\gamma_1, \gamma_2)$  such that

$$\Omega_{\text{III}}^2 = 4\pi \frac{\xi^2}{\mu} (\gamma \alpha_{\text{eff}})^2 \left( \frac{1}{1 + (4\pi/v)r_p\alpha_{\text{eff}}} \right). \quad (\text{A53})$$

Summarizing the partial results (A32), (A38) and (A53) we get the final expression

$$\Omega^2 = 4\pi \frac{\xi^2}{\mu} \left\{ 1 - \left[ \left( \frac{4\pi}{v} r_p + 2v \right) \alpha_{\text{eff}} - (\gamma \alpha_{\text{eff}})^2 \right] \left( \frac{1}{1 + (4\pi/v)r_p\alpha_{\text{eff}}} \right) \right\}. \quad (\text{A54})$$

## References

- [1] Schober H, Strauch D, Nützel K and Dörner B 1993 *J. Phys.: Condens. Matter* **5** 6155
- [2] Sinha S K 1973 *CRC Crit. Rev. Solid State Sci.* **4**
- [3] Rosenstock H B 1962 *Phys. Rev.* **129** 1959
- [4] Kurosawa T 1961 *J. Phys. Soc. Japan* **16** 1298
- [5] Cochran W and Cowley R A 1962 *J. Phys. Chem. Solids* **23** 447
- [6] Seinosuke O, Toshihiro A and Keiei K 1977 *Phys. Rev. B* **16** 1717
- [7] Lucovsky G, Martin R M and Burstein E 1971 *Phys. Rev. B* **4** 1367
- [8] Scott J F 1971 *Phys. Rev. B* **4** 1360
- [9] Gervais F 1973 *Solid State Commun.* **18** 197
- [10] Eckold G, Stein-Arsic M and Weber H J 1986 *Bericht der KFA, Jülich* **366**
- [11] Sham L J 1969 *Phys. Rev.* **188** 1431
- [12] Gervais F and Piriou B 1975 *Phys. Rev. B* **11** 3944
- [13] Schober H, Strauch D and Dörner B 1993 to be published
- [14] Barker A S Jr 1963 *Phys. Rev.* **132** 1474
- [15] Kress W, Schröder U, Prade J, Kulkarni A D and deWette F 1988 *Phys. Rev. B* **38** 2906
- [16] Eckold G, Stein-Arsic M and Weber H J 1987 *J. Appl. Cryst.* **20** 134