Systematic second-order perturbation theory for displacements, strains, and electric fields

Notes relevant to the `ANADDB` module of `ABINIT`

D. Vanderbilt
August 5, 2017

These notes were originally written to explain the background, establish the notation, and provide an explanation of certain features added to the `ANADDB` [1] package in an upgrade implemented by Xifan Wu in January 2004, providing for improved flexibility in the computation of dielectric, elastic, piezoelectric, and related tensors. However, I believe they may prove helpful more generally, e.g., as a starting point for the definition and computation of higher-order derivatives, or of properties computed under conditions of finite electric fields.

These notes will normally reside in the file *'vanderbilt-anaddb-notes.pdf'* in the *'/Infos'* subdirectory of the `ABINIT` distribution [1]. Regarding the definitions of elastic tensors, especially under conditions of finite pressure or stress, please also consult the notes *'elasticity-oganov.pdf'* [2] written by A. Oganov and located in the same subdirectory.

I would like to thank X. Wu, D.R. Hamann, and K.M. Rabe for input and helpful comments.

1 Introduction

The purpose of these notes is to provide a systematic framework for the discussion of response tensors that can be defined in terms of three kinds of homogeneous perturbations in crystalline insulators:

- Zone-center phonons

- Homogeneous electric fields

- Homogeneous strains

In particular, the notes provide a set of definitions of many of the elastic, dielectric, piezoelectric, and other tensors that can be defined in terms of the system response to these perturbations, and clarifies the connections between them. In addition, the notes make a direct connection with the quantities (i.e., the matrices of second derivatives) that are calculated and stored in the "derivative database" (`DDB`) of the `ABINIT` computer package, and with the analysis of these carried out by the `ANADDB` package of `ABINIT` [1].

2   Units

SI units are used throughout. This is the convention of most textbooks and review articles such as Lines and Glass [3], Nye [4], Ballato [5], etc. However, Waghmare's article [6], and much of the usual electronic structure literature, uses Gaussian units. We recall here that in SI units, potential $= (1/4\pi\epsilon_0)\, q_1 q_2/r$, energy density $= (\epsilon_0/2)\mathcal{E}^2$, and $D = \epsilon_0\, \mathcal{E} + P$.

3   Notation

Consider an insulating crystal with N atoms per primitive cell. We choose a reference state in which the lattice vectors are $\mathbf{a}_1$, $\mathbf{a}_2$, and $\mathbf{a}_3$ and the atomic coordinates are $R_m^{(0)}$. Here m is a composite label (atom and displacement direction) running over $1, ..., 3N$.

$$
\begin{aligned}
u_m &= \text{displacements of atoms from positions } R_m^{(0)}\ (m = 1, ..., 3N) \\
\eta_j &= \text{strain in Voigt notation } (j = 1, ..., 6) \\
\mathcal{E}_\alpha &= \text{electric field } (\alpha = x, y, z)
\end{aligned}
\tag{1}
$$

Cartesian coordinates are used throughout, with the exception of some `ABINIT` internal representations discussed in Sec. 4. I will try to consistently use m, n, ... for the $3N$ atomic displacement labels (e.g., forces will be F_m); j, k, ... for Voigt labels (e.g., stresses will be σ_j); and α, β, ... for Cartesian labels (e.g., polarizations will be P_α). Also note that the Voigt notation is a bit tricky for shear strains and stresses, e.g., $\sigma_6 = \sigma_{xy}$ but $\eta_6 = 2\eta_{xy}$, etc. This is explained more carefully in Sec. 6 (see also Nye, Ref. [4]).

Throughout these notes, we are only going to be interested in atomic displacements that preserve the primitive unit cell. In the phonon language, this means we are considering only phonon modes with q-vectors at the Brillouin zone (BZ) center.

Also, we will restrict ourselves entirely to zero temperature. Thus, entropy will never enter, and the distinction between thermodynamic functions written in terms of temperature vs. in terms of entropy will never arise.

The cell volume in the reference configuration (i.e., at zero strain) is by definition $\Omega_0 = \mathbf{a_1} \cdot \mathbf{a_2} \times \mathbf{a_3}$, so that the cell volume at strain η_j is $\Omega = \Omega_0\, \det[\eta]$ (interpreted as the determinant of the Cartesian $\eta_{\alpha\beta}$ corresponding to Voigt η_j).

Energies E (and other related thermodynamic energy functions) will be understood to be defined as the *energy per unit undeformed volume.* That is, the energy E is really an energy density, with units of J/m^3, defined as the energy per primitive cell of the *strained* crystal divided by the volume of the *unstrained* crystal. This is a tricky point that is often swept under the rug in many textbooks, leaving confusion. You can find an explicit discussion in Landau and Lifshitz, Theory of Elasticity (3rd Edition), p. 8: "The following remark..." (Many texts will say that E is the energy per unit volume, but then this turns out to be inconsistent with what they do later. Think about a crystal at its equilibrium volume, where the stress should be zero. The meaning of equilibrium is really that the energy *per unit cell* (let's call it $\bar{E}_c$) is

stationary. But since the volume is *not* stationary with respect to strains, defining σ as the derivative of the quantity energy per unit volume is incorrect.) In summary, H, H_0, etc. will have units of J/m^3, as though it were an energy per unit volume, but it is really an energy per *undeformed* unit volume.

We may sometimes use a collective notation

$$x_a = (\, u_m \,,\, \eta_j \,,\, \mathcal{E}\alpha \,)$$

where a runs over the $(3N + 6 + 3)$ coordinates describing the complete set of displacements, the strain, and the electric field. Then, for example, we may write an expansion

$$H = H_0 + A_a\, x_a + \tfrac{1}{2} B_{ab}\, x_a\, x_b \tag{2}$$

(implied sum notation) which is short-hand for

$$
\begin{aligned}
H = H_0 \;+\;& A_m\, u_m + A_j\, \eta_j + A_\alpha\, \mathcal{E}_\alpha \\
+\;& \tfrac{1}{2} B_{mn}\, u_m\, u_n + \tfrac{1}{2} B_{jk}\, \eta_j\, \eta_k + \tfrac{1}{2} B_{\alpha\beta}\, \mathcal{E}_\alpha\, \mathcal{E}_\beta \\
+\;& B_{mj}\, u_m\, \eta_j + B_{m\alpha}\, u_m\, \mathcal{E}_\alpha + B_{j\alpha}\, \eta_j\, \mathcal{E}_\alpha \;\;.
\end{aligned}
\tag{3}
$$

In this expansion, the first-order coefficients have the meaning

$$
\begin{aligned}
A_m &= -F_m/\Omega_0, \quad F_m = \text{force (N)} \\
A_j &= +\sigma_j, \quad \sigma_j = \text{stress } (J/m^3) \\
A_\alpha &= -P_\alpha, \quad P_\alpha = \text{polarization } (C/m^2) \;\;.
\end{aligned}
\tag{4}
$$

The second-order coefficients are the three diagonal-block tensors

$$
\begin{aligned}
B_{mn} &= K_{mn}/\Omega_0, \quad K_{mn} = \text{force-constant matrix} \\
B_{jk} &= C_{jk}, \quad C_{jk} = \text{elastic-constant matrix} \\
B_{\alpha\beta} &= -\chi_{\alpha\beta}, \quad \chi_{\alpha\beta} = \text{dielectric susceptibility matrix}
\end{aligned}
\tag{5}
$$

and the three off-diagonal-block tensors

$$
\begin{aligned}
B_{mj} &= -\Lambda_{mj}/\Omega_0, \quad \Lambda_{mj} = \text{``internal-strain tensor''} \\
B_{m\alpha} &= -Z_{m\alpha}/\Omega_0, \quad Z_{m\alpha} = \text{Born dynamical charge tensor} \\
B_{j\alpha} &= -e_{j\alpha}, \quad e_{j\alpha} = \text{piezoelectric tensor} \;\;.
\end{aligned}
\tag{6}
$$

In other words, the "big gradient vector" and "big Hessian matrix" are

$$
A = \begin{pmatrix} -F/\Omega_0 \\ \sigma \\ -P \end{pmatrix}
\qquad
B = \begin{pmatrix} K/\Omega_0 & -\Lambda/\Omega_0 & -Z/\Omega_0 \\ -\Lambda^T/\Omega_0 & C & -e \\ -Z^T/\Omega_0 & -e^T & -\chi \end{pmatrix} \;\;.
$$

The quantity F above is the force computed at $u_m = \eta_j = \mathcal{E}_\alpha = 0$, but the variation of the force with these variables is given by

$$F_m(u_m, \eta_j, \mathcal{E}\alpha) = -\Omega_0 \left(A_m + B_{mn} u_n + B_{mj} \eta_j + B_{m\alpha} \mathcal{E}_\alpha \right) \;\;. \tag{7}$$

Defining ΔF_m to be the change in the force from the reference crystal, and doing a similar analysis for $\Delta\sigma_j$ and ΔP_α, we find

$$
\begin{aligned}
\Delta F_m &= -K_{mn}\,u_n + \Lambda_{mj}\,\eta_j + Z_{m\alpha}\,\mathcal{E}_\alpha \\
\Delta\sigma_j &= -\Omega_0^{-1}\Lambda_{mj}\,u_m + C_{jk}\,\eta_k - e_{j\alpha}\,\mathcal{E}_\alpha \\
\Delta P_\alpha &= -\Omega_0^{-1}Z_{m\alpha}\,u_m + e_{j\alpha}\,\eta_j + \chi_{\alpha\beta}\,\mathcal{E}_\beta \ .
\end{aligned}
\tag{8}
$$

The following points should be kept in mind regarding the definitions of the various quantities above:

- All the second-derivative tensors K, C, χ, Λ, Z and e are "bare" quantities calculated *at fixed u, η, and $\mathcal{E}$.* The formulation of the "relaxed-ion" or "dressed" quantities will be discussed later.

- The terminology regarding the "internal-strain tensor" does not seem to be well established in the literature. Our Λ_{mj} is the tensor that expresses the first-order change in the force on an atom resulting from a first-order strain; for this reason, we will sometimes refer to it as the "force-response internal-strain tensor" when confusion might otherwise arise. Similarly, there is a "displacement-response internal-strain tensor" Γ_{mj} expressing the first-order change of the relaxed atomic displacements resulting from a first-order strain; this will be discussed in Sec. 5.2.

- Regarding the piezoelectric coefficient, our definition is really the transpose of the one most commonly found in the literature; that is, what we call 'e_{51}' is more commonly referred to as 'e_{15}' in the literature.

- All quantities defined here and calculated by `ABINIT` use the Voigt notation for stress- or strain-linked indices. However, all such quantities appearing above – that is, C, Λ, e, and σ – have the property that no factors of 2 are needed to relate them to true tensor quantities. That is, $\sigma_4 = \sigma_{yz}$, $C_{14} = C_{xx,yz}$, $C_{44} = C_{yz,yz}$, etc. This is explained further in Sec. 6, where it becomes evident that any object that can be defined as a (first, second, or higher) derivative of H with respect to strain is free of such conversion factors.

4   Connection to `ABINIT`

4.1   Introduction

`ABINIT` stores the following information in the `DDB` database:

- Geometry of reference structure: $\{R_m^{(0)}\}$, $\mathbf{a}_1$, $\mathbf{a}_2$, $\mathbf{a}_3$.

- Masses M_m and bare ionic charges Z_m^{ion}.

- Hessian tensors entering B: K, C, χ, Λ, Z, and e.

- Gradients F and σ are *not* included. See discussion below.

4

- The polarization $\mathbf{P}$ is currently not included, but we may wish to think about revising `ABINIT` so that it is included in the future.

Regarding F and σ, the philosophy is that the structural degrees of freedom, including strain, should have been relaxed already in the main run. Thus, any remaining forces and strains are assumed to be "computational noise" associated with numerical roundoff and non-zero tolerances. So, the job of `ANADDB` should be to compute the properties *as though* all the forces and stresses are zero.

Other information is stored in the `DDB` as well. The header part of the file includes various information to help identify the run that produced it, and some other information, such as force-constant matrix elements at non-zero q and certain third derivatives, is also included in the main part. For the present purposes, we will ignore this other information.

4.2   Detailed description of `ABINIT` inputs

In this section, we attempt to specify precisely how the second-derivative information is stored in the `DDB` file. This information corresponds to partial second derivatives of the energy with respect to displacements, electric fields, and strains, and thus there is a direct correspondence with the information in the tensors K, C, χ, Λ, Z and e. However, the conventions for these quantities in the `DDB` file differ from those in the present notes by factors associated with units conversions, reduced vs. Cartesian coordinates, cell volume factors, etc.

4.2.1   Lattice vectors and cell volume

The real space lattice vectors are $\mathbf{a}_\mu$ and henceforth μ, ν, ... will be used to label the three lattice directions. The reciprocal-space lattice vectors are defined via $\mathbf{G}_\mu \cdot \mathbf{a}_\nu = 2\pi\,\delta_{\mu\nu}$, but we introduce instead the "reduced" reciprocal-space lattice vectors $\mathbf{G}'_\mu = \mathbf{G}_\mu/2\pi$, so that

$$\mathbf{G}'_\mu \cdot \mathbf{a}_\nu = \delta_{\mu\nu} \tag{9}$$

Recall that $\Omega_0 = \mathbf{a_1} \cdot \mathbf{a_2} \times \mathbf{a_3}$. We will use the subscript "c" to denote an energy per unit cell. Thus,

$$H_{\rm c} = \Omega_0\,H \tag{10}$$

4.2.2   Reduced displacements and forces

In Sec. 3, the displacements were defined in Cartesian coordinates as u_m, $m = 1, 3N$. Recall that m is a composite index, $m \to \tau\alpha$, where $\tau = 1, N$ is a label for the atom in the unit cell and α is a Cartesian axis label.

The `ABINIT` code internally uses "reduced" coordinates (also sometimes called "lattice-vector" or "canonical" or "internal" coordinates) for displacements, defined via $s_{\tau\mu} = \mathbf{G}'_\mu \cdot \mathbf{u}_\tau$ and $\mathbf{u}_\tau = s_{\tau\mu}\mathbf{a}_\mu$ or, using index notation,

$$s_{\tau\mu} = G'_{\mu\alpha}\,u_{\tau\alpha} \tag{11}$$

and

$$u_{\tau\alpha} = s_{\tau\mu}\, a_{\mu\alpha} \tag{12}$$

Application of the chain rule shows that derivatives behave like

$$\frac{\partial}{\partial s_{\tau\mu}} = a_{\mu\alpha}\, \frac{\partial}{\partial u_{\tau\alpha}} \quad . \tag{13}$$

Thus, if $F_{\tau\alpha} = -\partial H_{\mathrm{c}}/\partial u_{\tau\alpha}$ is the force vector in Cartesian coordinates, then the "force vector in reduced coordinates" is

$$f_{\tau\mu} = -\frac{\partial H_{\mathrm{c}}}{\partial s_{\tau\mu}} = a_{\mu\alpha}\, F_{\tau\alpha} \tag{14}$$

where f has dimensions of energy.

4.2.3 Reduced electric fields and polarizations

Let

$$\varepsilon_\mu = |q_e|\, \mathcal{E} \cdot \mathbf{a}_\mu \quad , \tag{15}$$

be the electric field in reduced coordinates, where $|q_e|$ is the proton charge. Thus, ε_μ, having dimensions of energy, is the work done on a charge quantum $|q_e|$ to translate it by a lattice vector $\mathbf{a}_\mu$. Similarly, we can introduced the "reduced polarization"

$$p_\mu = |q_e|^{-1}\, \Omega_0\, \mathbf{G}'_\mu \cdot \mathbf{P} \quad . \tag{16}$$

Note that p_μ (dimensionless) is defined modulo 1 (that is, the "quantum of polarization" is just unity for each component p_μ), and that the electronic contribution to it is just

$$p_\mu^{\mathrm{elec}} = -(2\pi)^{-1} \sum_n \phi_{n,\mu} \tag{17}$$

where $\phi_{n,\mu}$ is the Berry phase of band n (defined modulo 2π) as in Ref. [17]. The energy per primitive cell arising from the $\mathcal{E}$-field is converted from Cartesian to internal coordinates as

$$\Omega_0\, \mathcal{E} \cdot \mathbf{P} = \varepsilon_\mu\, p_\mu \tag{18}$$

where both sides have dimensions of energy.

4.2.4 Definitions of perturbations in `ABINIT`

`ABINIT` computes and stores the derivatives of the wavefunctions with respect to *reduced* displacements, *reduced* electric fields, and *Cartesian* strains, and its energy units are Hartree (atomic units). The wavefunctions are expanded as $\psi_{n\mathbf{k}}(\mathbf{r}) = \sum_l c_{n\mathbf{k},l}\, \exp[i(\mathbf{G}_l + \mathbf{k}) \cdot \mathbf{r}]$ (where l represents a triplet of integers) and the $c_{n\mathbf{k},l}$ are dimensionless, so what is stored is

- For displacements: $dc_{n\mathbf{k}}/ds_\mu$ (dimensionless)

- For $\mathcal{E}$-fields: $dc_{n\mathbf{k},l}/d\varepsilon_\mu$ (inverse energy)

- For strains: $dc_{n\mathbf{k},l}/d\eta_j$ (Cartesian Voigt; dimensionless)

Similarly, forces are initially computed in reduced form as $f_{\tau\mu} = -\,dH_{\mathrm{c}}/ds_{\tau\mu}$, while stresses are computed as $\sigma_j = \Omega_0^{-1}\,dH_{\mathrm{c}}/d\eta_j$.

4.2.5 Second derivative information in the DDB

Similarly, the quantities that are stored in the second-derivative DDB are as follows.

- Displacement–and–Displacement: The quantity that is computed and stored in the DDB is

$$\frac{d^2 H_{\mathrm{c}}}{ds_{\tau\mu}\,ds_{\tau'\nu}} \tag{19}$$

  in terms of which the force-constant matrix is

$$K_{\tau\alpha,\tau'\beta} = \; G'_{\mu\alpha}\,G'_{\nu\beta}\,\frac{d^2 H_{\mathrm{c}}}{ds_{\tau\mu}\,ds_{\tau'\nu}}\;. \tag{20}$$

- Strain–and–Strain: The quantity that is computed and stored in the DDB is

$$\frac{d^2 H_{\mathrm{c}}}{d\eta_j\,d\eta_k} \tag{21}$$

  in terms of which the elastic-constant tensor is

$$C_{jk} = \Omega_0^{-1}\,\frac{d^2 H_{\mathrm{c}}}{d\eta_j\,d\eta_k} \tag{22}$$

  (recall that $H_{\mathrm{c}} = \Omega_0\,H$).

- Field–and–Field: The quantity that is computed and stored in the DDB is

$$\frac{d^2 H_{\mathrm{c}}}{d\varepsilon_\mu\,d\varepsilon_\nu} \tag{23}$$

  having dimensions of inverse Hartree, in terms of which the susceptibility is

$$\chi_{\alpha\beta} = -\,q_e^2\,\Omega_0^{-1}\,a_{\mu\alpha}\,a_{\nu\beta}\,\frac{d^2 H_{\mathrm{c}}}{d\varepsilon_\mu\,d\varepsilon_\nu} \tag{24}$$

  and $\epsilon_{\alpha\beta} = \delta_{\alpha\beta} + 4\pi\,\chi_{\alpha\beta}$.

- Displacement–and–Strain: The quantity that is computed and stored in the `DDB` is

$$\frac{d^2 H_\mathrm{c}}{ds_{\tau\mu}\, d\eta_j} \tag{25}$$

in terms of which the force-response internal-strain tensor is

$$\Lambda_{\tau\alpha,j} = -\, G'_{\mu\alpha}\, \frac{d^2 H_\mathrm{c}}{ds_{\tau\mu}\, d\eta_j} \quad . \tag{26}$$

- Displacement–and–Field: The quantity that is computed and stored in the `DDB` is the purely electronic contribution

$$\frac{d^2 H_\mathrm{c}}{ds_{\tau\mu}\, d\varepsilon_\nu} \tag{27}$$

to the dynamical charge tensor. The ionic contribution must be added to it. Thus, the full Born dynamical charge tensor is

$$Z_{\tau\alpha,\beta} = -\, |q_e|\, G'_{\mu\alpha}\, a_{\nu\beta}\, \frac{d^2 H_\mathrm{c}}{ds_{\tau\mu}\, d\varepsilon_\nu} + Z^{\mathrm{ion}}_{\tau}\, \delta_{\alpha\beta} \tag{28}$$

which has dimensions of charge.

- Strain–and–Field: The quantity that is computed and stored in the `DDB` is

$$\frac{d^2 H_\mathrm{c}}{d\eta_j\, d\varepsilon_\nu} \tag{29}$$

in terms of which the piezoelectric tensor is

$$e_{j\alpha} = -\, |q_e|\, \Omega_0^{-1}\, a_{\nu\alpha}\, \frac{d^2 H_\mathrm{c}}{d\eta_j\, d\varepsilon_\nu} \tag{30}$$

Note that $dH_\mathrm{c}/d\varepsilon_\nu$ has the interpretation of being a reduced polarization p^{elec}_ν in the sense of Eqs. (17-18). Thus, the quantity in Eq. (29) has the interpretation of $dp^{\mathrm{elec}}_\nu/d\eta_j$, i.e., a derivative of the *reduced polarization* with respect to strain. This corresponds [17] to the "proper" and not the "improper" piezoelectric tensor. Thus, it is (a reduced version of) the "proper" piezoelectric tensor that is stored in the `DDB`.

Note that a finite-difference calculation of derivatives of unreduced polarization **P** with respect to strain would yield the "improper" piezoelectric tensor. On the other hand, the quantity in Eq. (29) corresponds (up to factors of -2π, see Eq. (17)) to numerical derivatives of the *Berry phases* with respect to strain components. As explained in Ref. [17], when converted via an equation like Eq. (30), these yield the "proper" piezoelectric tensor.

Thankfully, the ionic contribution to the "proper" piezoelectric tensor is identically zero [17], so we don't have to include it in the `DDB`.

4.3 Electric field

Extensions to `ABINIT` to allow non-zero electric fields are in progress, based on the approach of Souza, Íñiguez and Vanderbilt [16]. For the time being, however, we will assume that the reference calculation is always at $\mathcal{E} = 0$, i.e., in vanishing electric field.

4.4 Symmetry reduction

Often it may be possible to block the vectors A_a and the tensors B_{ab} into irreducible representations ("irreps") according to the crystal symmetry. In this case, it may be possible to focus on just one irrep and calculate all the derived quantities just for that irrep. However, at least for now, the philosophy of the ANADDB implementation is to compute the full tensors, and to handle any symmetry reduction by hand at the stage of analyzing the results.

5 Computation of derived quantities

We are in the process of implementing, in the `ANADDB` module of `ABINIT`, a package that will be capable of computing and reporting a large number of derived quantities based on the input of information from the `DDB`. In this section, we list some of these quantities and give the formulas necessary for computing them.

The starting point will be the bare quantities K, C, χ, Λ, Z and e. It will be assumed that the information needed to construct these tensors has been read from the `DDB` file, and any needed conversions (units, reduced vs. Cartesian, cell volume factors, etc.) have already been applied as described in Sec. 4.2.

5.1 Elastic compliance tensor

The clamped-ion elastic compliance tensor S is simply obtained by inverting the elastic-tensor matrix C:

$$S_{jk} = (C^{-1})_{jk}. \tag{31}$$

Warning messages are issued if the matrix C_{jk} is found to be nearly singular.

5.2 Internal-strain tensors

As noted at the end of Sec. 3, Λ_{mj} is the "force-response internal-strain tensor" giving the force on atom m due to strain j. If instead one wants the "displacement-response internal-strain tensor," which gives displacement of atom m due to strain j, it is given by the product

$$\Gamma_{mj} = (K^{-1})_{mn} \Lambda_{nj}. \tag{32}$$

9

5.3 Relaxed-ion quantities

5.3.1 Formulation

The tensors C, χ, and e are "clamped-ion" quantities defined at fixed u. We define the corresponding "relaxed-ion" or "dressed" response tensors $\widetilde{C}$, $\widetilde{\chi}$, and $\widetilde{e}$ as follows. Let

$$\widetilde{H}(\eta_j, \mathcal{E}_\alpha) = \min_{\{u_m\}} H(u_m, \eta_j, \mathcal{E}_\alpha) \ . \tag{33}$$

Referring back to Eq. (3) and setting $\partial H / \partial u_n = 0$, and assuming that the reference configuration is one in which the forces vanish, we find

$$0 = B_{nm}\, u_m + B_{nj}\, \eta_j + B_{n\alpha}\, \mathcal{E}_\alpha$$

from which it follows that

$$u_m = -(B^{-1})_{mn}\, [\, B_{nj}\, \eta_j + B_{n\alpha}\, \mathcal{E}_\alpha]$$

or

$$\widetilde{H}(\eta_j, \mathcal{E}_\alpha) = \tfrac{1}{2}\widetilde{B}_{jk}\, \eta_j\, \eta_k + \tfrac{1}{2}\widetilde{B}_{\alpha\beta}\, \mathcal{E}_\alpha\, \mathcal{E}_\beta + \widetilde{B}_{j\alpha}\, \eta_j\, \mathcal{E}_\alpha \tag{34}$$

where

$$
\begin{aligned}
\widetilde{B}_{jk} &= B_{jk} - B_{mj}\,(B^{-1})_{mn}\, B_{nk} \\
\widetilde{B}_{\alpha\beta} &= B_{\alpha\beta} - B_{m\alpha}\,(B^{-1})_{mn}\, B_{n\beta} \\
\widetilde{B}_{j\alpha} &= B_{j\alpha} - B_{mj}\,(B^{-1})_{mn}\, B_{n\alpha} \ .
\end{aligned}
\tag{35}
$$

Using the definitions given in Eqs. (5) and (6), and making use of the fact that B, and thus $\widetilde{B}$, are real symmetric matrices, these become

$$
\begin{aligned}
\widetilde{C}_{jk} &= C_{jk} - \Omega_0^{-1}\, \Lambda_{mj}\,(K^{-1})_{mn}\, \Lambda_{nk} \tag{36} \\
\widetilde{\chi}_{\alpha\beta} &= \chi_{\alpha\beta} + \Omega_0^{-1}\, Z_{m\alpha}\,(K^{-1})_{mn}\, Z_{n\beta} \tag{37} \\
\widetilde{e}_{j\alpha} &= e_{j\alpha} + \Omega_0^{-1}\, \Lambda_{mj}\,(K^{-1})_{mn}\, Z_{n\alpha} \tag{38}
\end{aligned}
$$

Note that $\widetilde{C}$ and $\widetilde{e}$ are the "physical" elastic constant and piezoelectric tensors, respectively, since it is virtually impossible to measure these quantities without allowing the atomic coordinates to respond. The first and second terms on the right-hand sides of the above equations are usually referred to as the "clamped-ion" (or "purely electronic") and "relaxed-ion" (or "lattice") contributions, respectively.

5.3.2 Pseudo-inverse of the force-constant matrix

The force-constant matrix K should have three vanishing eigenvalues as a result of translational symmetry. Moreover, in soft-mode systems, other eigenvalues may be close to zero. Thus, for the many operations that use K^{-1}, there is a danger of numerical instabilities arising from attempting to invert a singular, or nearly singular, matrix. For this reason, we have implemented a careful procedure for obtaining the

"pseudo-inverse" of K; throughout these notes, whenever we refer to K^{-1}, we really mean the pseudo-inverse.

Briefly, we begin by identifying the three-dimensional space of acoustic modes (i.e., uniform translations), and project the matrix K onto the complementary subspace of dimension $3N - 3$. We then invert in this subspace, reporting a warning message if any eigenvalues are very small, and then transform back to the full space. Thus, by construction, the resulting pseudo-inverse is zero in the subspace of translational modes, and is the inverse of the original matrix in the complementary subspace.

As a result, any time K^{-1} is multiplied by another tensor, a pre-projection onto the complementary subspace of dimension $3N - 3$ is effectively carried out. In other words, "the acoustic sum rule is effectively enforced" in any operation involving K^{-1}.

5.3.3   Dielectric tensor

The dielectric response can be measured in three frequency regimes:

- At frequencies much greater than any phonon frequency (but less than electronic frequencies). In this case only the electrons respond and

$$\epsilon_{\alpha\beta}^{(\infty)} = \delta_{\alpha\beta} + 4\pi\chi_{\alpha\beta} \ . \tag{39}$$

- At frequencies low compared to any phonon frequency, but above mechanical resonance frequencies,

$$\epsilon_{\alpha\beta}^{(0)} = \delta_{\alpha\beta} + 4\pi\widetilde{\chi}_{\alpha\beta} \ . \tag{40}$$

- At frequencies low compared to mechanical resonance frequencies of the macroscopic sample, the strain can also relax,

$$\epsilon_{\alpha\beta}^{(\sigma)} = \delta_{\alpha\beta} + 4\pi\widetilde{\chi}_{\alpha\beta}^{(\sigma)} \tag{41}$$

as discussed below and in Sec. 5.6.

Note that the `ANADDB` module of `ABINIT` has been capable of computing $\epsilon^{(0)}$ for some time. Actually, it is capable of computing $\epsilon_{\alpha\beta}(\omega)$ in the regime of lattice frequencies, and in particular, if this is evaluated for $\omega = 0$, $\epsilon^{(0)}$ results. However, the revised version of the `ANADDB` is being written to compute $\epsilon^{(0)}$ and $\epsilon_{\alpha\beta}^{(\sigma)}$ independently, using the pseudo-inverse of Sec. 5.3.2.

The derivation for the free-stress susceptibility $\epsilon^{(\sigma)}$ starts from the free-stress energy functional

$$\widetilde{H}^{(\sigma)}(\mathcal{E}_\alpha) = \min_{\{\eta_j\}} \widetilde{H}(\eta_j, \mathcal{E}_\alpha) \tag{42}$$

Following a line of reasoning similar to what was done for the elimination of displacement degrees of freedom in Sec. 5.3.1, we obtain

$$\widetilde{\chi}_{\alpha\beta}^{(\sigma)} = \widetilde{\chi}_{\alpha\beta} + \widetilde{e}_{j\alpha} (\widetilde{C}^{-1})_{jk} \widetilde{e}_{k\beta} \tag{43}$$

in terms of which the free-stress dielectric tensor $\epsilon_{\alpha\beta}^{(\sigma)}$ is given as in Eq. (41). Note that Eq. (43) is equivalent to Eq. (59) appearing later in Sec. 5.6.

file response.pdf was missing in abinit source

Figure 1: Definitions of response functions.

5.4   Change of notation

For the remainder of Sec. 5, all quantities will be assumed to be relaxed-ion quantities. Up to now such quantities have been denoted by a tilde (e.g., $\widetilde{\chi}$), but henceforth the tildes will be dropped.

5.5   Elastic tensors

From Eq. (8), the elastic tensor C is defined as the partial derivative of σ_j with respect to η_k at fixed macroscopic electric field $\mathcal{E}$; to emphasize this, we will optionally write a superscript $(\mathcal{E})$:

$$C_{jk}^{(\mathcal{E})} = \frac{\partial \sigma_j}{\partial \eta_k}\bigg|_{\mathcal{E}} \ . \tag{44}$$

For some purposes, we may be interested instead in the elastic constant tensor at *fixed displacement field D*:

$$C_{jk}^{(D)} = \frac{\partial \sigma_j}{\partial \eta_k}\bigg|_{D} \ . \tag{45}$$

The inverse of the elastic constant tensor is the compliance tensor. It can be defined either at fixed $\mathcal{E}$ or D:

$$S_{jk}^{(\mathcal{E})} = \frac{\partial \eta_j}{\partial \sigma_k}\bigg|_{\mathcal{E}} \ , \tag{46}$$

$$S_{jk}^{(D)} = \frac{\partial \eta_j}{\partial \sigma_k}\bigg|_{D} \ . \tag{47}$$

The definitions of C and S are shown schematically on the left-hand side of Fig. 1. The meaning of the arrow pointing from η to σ and labeled by C is that

$$\delta\sigma_j = C_{jk}\,\delta\eta_k$$

(either at fixed $\mathcal{E}$ or fixed D), and similarly the downgoing arrow labeled S indicates that

$$\delta\eta_j = S_{jk}\,\delta\sigma_k \ .$$

The elastic tensors and compliances are related as inverses,

$$S^{(\mathcal{E})} = (C^{(\mathcal{E})})^{-1} \tag{48}$$

$$S^{(D)} = (C^{(D)})^{-1} \ . \tag{49}$$

12

The relation between $C^{(D)}$ and $C^{(\mathcal{E})}$, and between $S^{(D)}$ and $S^{(\mathcal{E})}$, will be given at the end of Sec. 5.6.

Note that the compliance matrices are defined above in Voigt notation, and in this case *there are factors of 2 and 4* needed to make connection with true tensor quantities: $S_{14} = 2S_{xx,yz}$, $S_{44} = 4S_{yz,yz}$, etc. This is explained more fully in Sec. 6.

Finally, note that various different definitions can be given of the elastic constants under conditions of nonzero hydrostatic pressure or, more generally, under nonzero stress. In this case, the experimentally relevant tensors (e.g., for seismic waves in the interior of the earth) *do not* necessarily correspond to the ones computed directly by ABINIT. For a discussion of these issues, please see [2].

5.6 Piezoelectric tensors

The piezoelectric tensors are defined schematically by the lines crossing horizontally and diagonally in Fig. (1). (All piezoelectric tensors are the "proper" ones – see the last part of Sec. 4.2.5 and Ref. [17] for a discussion.) The interpretation of the arrows is

$$
\begin{aligned}
\delta\eta_j &= d_{j\alpha}\,\delta\mathcal{E}_\alpha \\
\delta D_\alpha &= d_{j\alpha}\,\delta\sigma_j \\
\delta\sigma_j &= -e_{j\alpha}\,\delta\mathcal{E}_\alpha \\
\delta D_\alpha &= e_{j\alpha}\,\delta\eta_j \ .
\end{aligned}
\tag{50}
$$

The first of the four equations above is sometimes said to describe the "converse" piezoelectric effect, while the second one describes the "direct" piezoelectric effect, and the third and/or the fourth describe the "inverse" piezoelectric effect. Thus, one sometimes refers to d as the coefficient of the "direct piezoelectric effect" while e (denoted as c almost as often as e – these notations are equivalent) is the coefficient of the "inverse piezoelectric effect." Restating, the piezoelectric coefficients may be defined via

$$
e_{j\alpha} = \left.\frac{\partial D_\alpha}{\partial\eta_j}\right|_{\mathcal{E}} = \left.\frac{\partial P_\alpha}{\partial\eta_j}\right|_{\mathcal{E}} \qquad \text{or} \qquad e_{j\alpha} = -\left.\frac{\partial\sigma_j}{\partial\mathcal{E}_\alpha}\right|_{\eta}
\tag{51}
$$

and

$$
d_{j\alpha} = \left.\frac{\partial\eta_j}{\partial\mathcal{E}_\alpha}\right|_{\sigma} \qquad \text{or} \qquad d_{j\alpha} = \left.\frac{\partial D_\alpha}{\partial\sigma_j}\right|_{\mathcal{E}} = \left.\frac{\partial P_\alpha}{\partial\sigma_j}\right|_{\mathcal{E}} \ .
\tag{52}
$$

The equivalence between the two expressions for e, and similarly between the two expressions for d, comes from thermodynamic relations as discussed in the Appendix of Ballato [5] and in Nye [4].

The e and d tensors are related by

$$
e_{j\alpha} = C_{jk}^{(\mathcal{E})}\,d_{k\alpha} \qquad \text{or, equivalently,} \qquad d_{j\alpha} = S_{jk}^{(\mathcal{E})}\,e_{k\alpha} \ .
\tag{53}
$$

One also sometimes defines tensors g and h via

$$
\begin{aligned}
\delta\eta_j &= g_{j\alpha}\,\delta D_\alpha \\
\delta\mathcal{E}_\alpha &= -g_{j\alpha}\,\delta\sigma_j \\
\delta\sigma_j &= -h_{j\alpha}\,\delta D_\alpha \\
\delta\mathcal{E}_\alpha &= -h_{j\alpha}\,\delta\eta_j \ .
\end{aligned}
\tag{54}
$$

(I am not aware of any standard names for these piezoelectric tensors.) They are related to e and d via

$$
g_{j\alpha} = \beta^{(\sigma)}_{\alpha\beta}\,d_{j\beta}
\tag{55}
$$

$$
h_{j\alpha} = \beta^{(\eta)}_{\alpha\beta}\,e_{j\beta} \ .
\tag{56}
$$

The Voigt notation introduces no factors of 2 for shear components of e or h, but there *are* such factors for d and g: $d_{51} = 2d_{xz,x}$, etc. (see Sec. 6). Also, note that it is more common to find the indices reversed in the literature; e.g., this piezoelectric component is more usually referred to as 'd_{15}'.

The relations between the elastic tensors defined at fixed $\mathcal{E}$ and fixed $\mathbf{D}$ in Sec. 5.5 are

$$
C^{(D)}_{jk} = C^{(\mathcal{E})}_{jk} + h_{j\alpha}\,e_{k\alpha}
\tag{57}
$$

$$
S^{(D)}_{jk} = S^{(\mathcal{E})}_{jk} - g_{j\alpha}\,d_{k\alpha}
\tag{58}
$$

and between the dielectric tensors defined at fixed η and fixed σ in Sec. 5.3.3 are

$$
\epsilon^{(\sigma)}_{\alpha\beta} = \epsilon^{(\eta)}_{\alpha\beta} + e_{j\alpha}\,d_{j\beta}
\tag{59}
$$

$$
\beta^{(\sigma)}_{\alpha\beta} = \beta^{(\eta)}_{\alpha\beta} - g_{j\alpha}\,h_{j\beta} \ .
\tag{60}
$$

Note that Eq. (59) is equivalent to Eq. (43).

6 Voigt notation

6.1 Basic formulation

There are two systems that are commonly in use to index quantities that depend on strains or stresses: the "true tensor notation" in which all Cartesian indices are written explicitly, and the "Voigt notation" in which a reduced index is used: $xx \to 1$, $yy \to 2$, $zz \to 3$, $yz \to 4$, $xz \to 5$, $xy \to 6$. The stress elements are defined simply by making this index replacement:

$$
\begin{array}{ll}
\sigma_1 = \sigma_{xx} & \sigma_4 = \sigma_{yz} \\
\sigma_2 = \sigma_{yy} & \sigma_5 = \sigma_{xz} \\
\sigma_3 = \sigma_{zz} & \sigma_6 = \sigma_{xy}
\end{array}
\tag{61}
$$

However, the *strain elements have factors of two inserted in the definition of the shear elements*:

$$\eta_1 = \eta_{xx} \qquad \eta_4 = 2\eta_{yz}$$
$$\eta_2 = \eta_{yy} \qquad \eta_5 = 2\eta_{xz} \qquad (62)$$
$$\eta_3 = \eta_{zz} \qquad \eta_6 = 2\eta_{xy}$$

The context here is that η is a symmetric tensor, i.e., $\eta_{xy} = \eta_{yx}$, defined via $\eta_{\alpha\beta} = \frac{1}{2}(u_{\alpha,\beta} + u_{\beta,\alpha})$. Here $u_{\alpha,\beta} = \partial u_\alpha/\partial x_\beta$ is the unsymmetrized tensor defined in term of spatial derivatives $\partial/\partial x_\beta$ of the medium displacement u_α. Thus, we could alternatively have written $\eta_6 = \eta_{xy} + \eta_{yx}$, etc.

The reason for the introduction of these factors of 2 is explained very nicely in the book by Nye, Ref. [4] (see, e.g., Secs. VII.2-3 and VIII.2 therein). I give a brief discussion based on the thermodynamic energy function H of Sec. 3. The change dH of the energy should be the same in both frameworks. In the true tensor framework, H is a a function of 9 variables, and we can write

$$dH = \sigma_{xx}\, d\eta_{xx} + ... + \sigma_{yz}\, d\eta_{yz} + ... + \sigma_{zy}\, d\eta_{zy} + ... \qquad (63)$$

where each '...' indicates two more terms obtained by cyclic permutation of (xyz). On the other hand, in the Voigt framework, the same dH can be written as a sum of only six terms as

$$dH = \sigma_1\, d\eta_1 + ... + \sigma_4\, d\eta_4 + ... \qquad (64)$$

where each '...' indicates two more terms obtained by cyclic permutation (123) or (456). Comparing these two equations, which must be equal, and using that $\eta_{yz} = \eta_{zy}$, it is clear that a factor of two must be inserted into the defining connection either between σ_{yz} and σ_4, or else between η_{yz} and η_4. The Voigt notation arises from making the second choice, i.e., Eqs. (61) and (62).

In practice, this means that if you want to introduce a Voigt strain $\eta_6 = t$ into an `ABINIT` calculation (say in order to check, by finite differences, the computation of the stress or elastic tensor), then one should set $\eta_{xy} = \eta_{yx} = t/2$ when constructing the unit cell. For example, if the original cell vectors describe a simple cubic lattice, then the new lattice vectors could be $\mathbf{a}'_1 = a(1, t/2, 0)$, $\mathbf{a}'_2 = a(t/2, 1, 0)$, and $\mathbf{a}'_3 = a(0, 0, 1)$; or they could be $\mathbf{a}'_1 = a(1, t, 0)$, $\mathbf{a}'_2 = a(0, 1, 0)$, and $\mathbf{a}'_3 = a(0, 0, 1)$; etc. (In the first case, $u_{x,y} = u_{y,x} = t/2$; in the second case, $u_{x,y} = 0$ and $u_{y,x} = t$; in both cases, $\eta_{xy} = t/2$ and $\eta_6 = t$.) Then, if H is expanded in powers of t, the coefficient of the linear term is just σ_6, the coefficient of the quadratic term is just C_{66}, etc.

6.2 Systematics

We can understand when to insert, or not to insert, factors of 2 in the relations between Voigt and Cartesian notations by following the following rules:

- Any inserted factor of two makes the Voigt object larger: $Q_{...4...} = 2Q_{...yz...}$

- If there is a factor of η in the numerator of a derivative, *do insert* a factor of 2 for each shear component.

- If there is a factor of σ in the numerator of a derivative, no factors are needed.

- If there is a derivative with respect to η, no factors are needed.

- If there is a derivative with respect to σ, *do insert* a factor of 2 for each shear component.

Thus, there is no need to insert factors of two when interpreting most of the objects we have introduced, including σ, C, e, Λ, and Γ. However, we *do* need factors for S; since $S_{ij} = d\eta_i/d\sigma_j$, we need $S_{14} = 2S_{xx,yz}$, $S_{44} = 4S_{yz,yz}$, etc. As for the piezoelectric tensors, it is clear from Eqs. (51-56) and the rules given above that factors of 2 *are* needed for d and g, but not for e and h.

Finally, in the linear-response calculation, one needs to calculate and store objects such the first-order derivatives of Bloch wavefunctions with respect to strains, i.e., $d|u_{n\mathbf{k}}\rangle/d\eta_j$. This is a derivative with respect to strain, so by the above rules, there is no need to worry about a factor of two in the definition of this object.

6.3   A word about terminology

Nye (Ref. [4]) refers to the arrays of Voigt elements as "matrices" while the arrays of Cartesian-labeled elements are called "tensors." In his terminology, the term "tensor" is reserved for objects that transform under rotations by the application of the associated 3×3 rotation matrix to each Cartesian index. Thus, for example, the rotated piezoelectric tensor is given by

$$e'_{\beta\gamma,\alpha} = R_{\beta\mu} R_{\gamma\nu} R_{\alpha\tau} e_{\mu\nu,\tau} \tag{65}$$

where R is the rotation matrix and there are implicit sums over repeated indices. Since the the 6×3 matrix of $e_{j\alpha}$ elements in the Voigt notation does *not* transform in a similar way, Nye takes pains to call the Voigt $e_{j\alpha}$ a "matrix" and not a "tensor."

While Nye's point is very well taken, the habit of referring to the Voigt elements $e_{j\alpha}$ as elements of the "piezoelectric tensor" is by now rather widely ingrained, and in my opinion it is rather too pedantic to insist on the narrow definition. In these notes, therefore, I generally use the term "tensor" to refer indiscriminately to either the Voigt or the fully Cartesian notations. However, in deference to Nye, I do sometimes use the extended phase "true tensor notation" to refer to the fully Cartesian notation.

7   Summary

As of January 2004, the calculation of the following quantities has been implemented for the proposed update of `ANADDB` to version 4.3:

- All six "bare" tensors K, C, χ, Λ, Z and e, Sec. 3.

- The clamped-ion compliance tensor S of Sec. 5.1.

- The displacement-response internal-strain tensor Γ of Sec. 5.2.

- The relaxed-ion elastic, dielectric, and piezoelectric tensors ($\widetilde{C}$, $\widetilde{\chi}$, and $\widetilde{e}$) of Sec. 5.3.1, and the relaxed-ion compliance tensor $S^{(\mathcal{E})}$ of Sec. 5.5.

It is our intention to include the ability to calculate additional tensors, including many of the ones defined in Secs. 5.5 and 5.6, in a future release of the `ANADDB` module of `ABINIT`.

References

[1] ABINIT is a common project of the Université Catholique de Louvain, Corning Incorporated, and other contributors (http://www.abinit.org). X. Gonze, J.-M. Beuken, R. Caracas, F. Detraux, M. Fuchs, G.-M. Rignanese, L. Sindic, M. Verstraete, G. Zerah, F. Jollet, M. Torrent, A. Roy, M. Mikami, Ph. Ghosez, J.-Y. Raty, D.C. Allan, Comput. Mater. Sci. **25**, 478-492 (2002).

[2] A set of notes on elastic constants can be found in the file *'elasticity-oganov.pdf'* written by A. Oganov and located in the *'/Infos'* subdirectory of the ABINIT distribution.

[3] M. E. Lines and A. M. Glass, *Principles and Applications of Ferroelectrics and Related Materials,* (Clarendon Press, Oxford, 1977).

[4] J. F. Nye, *Physical properties of crystals* (Oxford U.P., Oxford 1985).

[5] A. Ballato, IEEE Transac. Ultrason. Ferro. and Freq. Control **42**, 916 (1995).

[6] U. Waghmare, unpublished.

[7] S. Baroni, P. Giannozzi and A. Testa, Phys. Rev. Lett. **78**, 1861 (1987).

[8] P. Giannozzi, S. de Gironcoli, P. Pavone, and S. Baroni, Phys. Rev. B **43**, 7231 (1991).

[9] X. Gonze, Phys. Rev. A **52**, 1096 (1995).

[10] X. Gonze, Phys. Rev. B **55**, 10337 (1997).

[11] X. Gonze and C. Lee, Phys. Rev. B **55**, 10355 (1997).

[12] S. Baroni, S. de Gironcoli, A. Dal Corso, and P. Giannozzi, Rev. Mod. Phys. 73, 515 (2001).

[13] R.D. King-Smith and D. Vanderbilt, Phys. Rev. B **47**, 1651 (1993).

[14] R. Resta, Rev. Mod. Phys. 66, 899 (1994).

[15] R. W. Nunes and X. Gonze, Phys. Rev. B **63**, 155107 (2001).

[16] I. Souza, J. Íñiguez, and D. Vanderbilt, Phys. Rev. Lett. **89**, 117602 (2002).

[17] D. Vanderbilt, J. Phys. Chem. Solids **61**, 147 (2000).