

Adiabatic density-functional perturbation theory

Xavier Gonze

*Unité de Physico-Chimie et de Physique des Matériaux, Université Catholique de Louvain,
B-1348 Louvain-la-Neuve, Belgium*

(Received 12 December 1994)

The treatment of adiabatic perturbations within density-functional theory is examined, at arbitrary order of the perturbation expansion. Due to the extremal property of the energy functional, standard variation-perturbation theorems can be used. The different methods (Sternheimer equation, extremal principle, Green's function, and sum over state) for obtaining the perturbation expansion of the wave functions are presented. The invariance of the Hilbert space of occupied wave functions with respect to a unitary transformation leads to the definition of a "parallel-transport-gauge" and a "diagonal-gauge" perturbation expansion. Then, the general expressions are specialized for the second, third, and fourth derivative of the energy, with an example of application of the method up to third order.

PACS number(s): 31.15.Ew, 02.30.Mv, 02.70.-c, 71.10.+x

I. INTRODUCTION

The density-functional theory (DFT), introduced by Hohenberg, Kohn, and Sham nearly 30 years ago [1–3], is a powerful approach to the many-body problem in quantum mechanics. Applications of this formalism range from atomic [3], molecular [3], condensed matter [3], and nuclear [4] physics to materials science [5]. In essence, the theorems proved by Hohenberg and Kohn allow one to replace the computation for the complicated many-body ground-state fermionic wave function by the search for the ground-state density, with a tremendous reduction of the number of degrees of freedom. The subsequent Kohn and Sham construction [2], thanks to the use of a noninteracting kinetic energy functional, restricts our ignorance of the explicit expression for the energy to a small part of it, the exchange-correlation energy. The latter is found to be adequately mimicked by a local density approximation [3]. In this framework, the bond lengths between atoms in molecules or in a condensed phase are found within a few percent of the experimental data.

The total energy and the charge density are basic quantities of DFT, and give access to a wide number of experimental observables. In the present paper, the focus will be on the quantities obtained as derivatives of the total energy (or the density) with respect to a small change of the potential applied to the fermionic system. As an example, the force exerted on a nucleus is equal to minus the derivative of the total energy of the system with respect to the displacement of the atom. More examples will be given in the next section (force constants, polarizability, etc.).

The calculation of such energy derivatives can be done by finite-difference methods: the total energy is obtained for slightly different values of the applied external field, then the derivative of the total energy curve with respect to the small disturbance is extracted numerically. Although it is a very convenient method (since virtually no extra coding effort is required for specifically dealing with

derivatives), the recent practice has shown that perturbation techniques within DFT are more powerful. Such techniques, in the case of the calculation of the first-order derivative of wave functions, were discovered or rediscovered independently by many groups, working in atomic, molecular, or solid-state physics [6–14]. They are based either on the Sternheimer equation, Green's functions, sum-over-states techniques, or on the Hylleraas variational technique (see Sec. II). Applications of these perturbation techniques [15] were numerous [16–37].

A generalization to arbitrary order of perturbation, based on the powerful $2n+1$ theorem of perturbation theory [38] and on generalized Sternheimer equations, was proposed a few years ago [39], and applied recently, up to third order [40–43]. It is the purpose of the present paper to build upon these arbitrary order theoretical results, and to systematically derive new formulas from a more general framework.

At the center of the present approach lies the *variational* principle (actually, a minimum principle, see p. 1050 of Ref. [44]) obeyed by the DFT energy functional. In 1961, Sinanoglu [45] noticed that the perturbative expansion of a variational principle gives an interesting mathematical structure, with a generic $2n+1$ theorem, and variational properties of even derivatives of the energy.

In Sec. II the basic definitions needed for dealing with perturbation theory are recalled, as well as some example of perturbations and related observables. The previous theoretical works in density-functional perturbation theory (DFPT) are briefly described. The mathematical framework obtained from the consideration of perturbation theory applied to a variational principle [46,47], following Sinanoglu, is summarized. This section, written for internal completeness, could be skipped by the reader seeking new results.

Section III gives the application of perturbation theory to the Kohn-Sham energy to arbitrary order of perturbation. New explicit expressions for the $2n+1$ theorem and the variational principle for the even orders of per-

turbation in DFT are obtained. The invariance of the energy with respect to a “gauge” unitary transform within the occupied valence states is studied in detail. It gives the connection between the $2n + 1$ theorem for DFT as derived previously [39] and the present approach. The Hylleraas variational approach [46] to the determination of wave function derivatives and the Sternheimer equation [48] approach to them are also linked.

Section IV focuses on the lowest-order results, and gives detailed analysis of derivatives up to fourth order, as derived from the arbitrary order expressions. An application of the second- and third-order formulas is given, and compared with the finite-difference results.

Finally, it is worthwhile to mention that the treatment of perturbations within the Hartree-Fock approximation gives a formalism (the coupled-perturbed Hartree-Fock theory [49]), proposed almost 35 years ago, that exhibits some similarities to the present one. This formalism has been implemented by many different groups, and other quantum chemistry methods have also been adapted to the treatment of perturbations [49].

II. BACKGROUND

A. Definitions

Let us consider one perturbation, associated with a small parameter λ . For a generic observable of the system $X(\lambda)$, one writes the perturbation series as follows:

$$X(\lambda) = X^{(0)} + \lambda X^{(1)} + \lambda^2 X^{(2)} + \lambda^3 X^{(3)} + \dots \quad (1)$$

The expansion coefficients are not the derivatives of $X(\lambda)$ with respect to λ , but are related to them by a simple numerical coefficient:

$$X^{(n)} = \frac{1}{n!} \left. \frac{d^n X}{d\lambda^n} \right|_{\lambda=0} \quad (2)$$

The applied potential $V(r)$, external to the electronic system, is supposed known through all orders, while it is our aim to calculate the perturbation expansion of other quantities, such as the energy E or the density $n(r)$.

B. Typical perturbations of molecular systems

Two kinds of perturbations have a particular significance for molecular systems: small displacements of atoms and applied electromagnetic fields.

The perturbation corresponding to small displacements of atoms, written τ_μ , is such that the nucleus labeled τ , located in R_τ , with (positive) charge Z_τ , is displaced by the small amount λ_{τ_μ} along the direction μ (the unit vector along μ is written \mathbf{e}_μ). The change of potential energy that affects an electron located at \mathbf{r} (x, y, z) (negative charge—so the potential energy is the opposite of the electrostatic potential) is a function of λ_{τ_μ} :

$$\Delta V(\mathbf{r}) = -\frac{Z_\tau}{|\mathbf{r} - (\mathbf{R}_\tau + \lambda_{\tau_\mu} \mathbf{e}_\mu)|} + \frac{Z_\tau}{|\mathbf{r} - \mathbf{R}_\tau|} \quad (3)$$

Expanding this equation to first order gives

$$V^{(1)}(\mathbf{r}) = \frac{-[\mathbf{e}_\mu \cdot (\mathbf{r} - \mathbf{R}_\tau)] Z_\tau}{|\mathbf{r} - \mathbf{R}_\tau|^3} \quad (4)$$

Higher-order derivatives can easily be derived from Eq. (3).

An applied external potential or field is another type of perturbation. Such a perturbation, although restricted to be static (or adiabatic) in the present study, could be time dependent, [14,16,19,34,50]. Usually, one thinks of a homogeneous electric field, but fields that vary at the atomic length scale can be considered, as well as dipolar or quadrupolar fields. One can also consider vector potentials (needed for the description of magnetic fields).

In the case of a homogeneous electric field, the potential that affects an electron can be written as

$$\Delta V(\mathbf{r}) = \sum_{\alpha=1}^3 E_\alpha r_\alpha \quad (5)$$

for which a three-dimensional vector of small parameters E_α , defining the direction of the electric field, has to be specified. Let us choose the electric field $(E, 0, 0)$, with the single small parameter E , then

$$V^{(1)}(\mathbf{r}) = x \quad (6)$$

Note that for this perturbation ΔV is linear in the small parameter E , while the potential change due to an atomic displacement Eq. (3) has nonlinear contributions.

For all these cases, the perturbation treatment competes with a “direct” approach, in which the total energy change is computed in the unperturbed and perturbed cases, separately, and then compared.

C. Typical perturbations of periodic systems

In the crystalline case, changes of microscopic potentials will be either commensurate with the crystalline periodicity, or incommensurate. For example, phonons with incommensurate wave vectors can be considered. A finite-difference approach is unable to treat such perturbations [51] while perturbation theory easily overcomes the problem [39]. Homogeneous deformations of the whole solid are needed to investigate elastic properties of solids [17]. As a last example, in the framework of the pseudopotential approximation, the “alchemical perturbation” [23] transmutes the pseudopotential of one ionic species into the pseudopotential of another ionic species. This allows a convenient study of alloys or impurities, vacancies or interstitials.

D. Response functions

Let us now examine the properties associated with derivatives of the total energy with respect to the “atomic displacement” and “homogeneous electric field” pertur-

bations.

Minus the first derivative of the total energy with respect to a nuclei displacement is the force exerted on this nuclei:

$$F_{\tau_\mu} = -\frac{\partial E}{\partial \lambda_{\tau_\mu}}. \quad (7)$$

Minus the first derivative of the total energy with respect to a homogeneous electric field is the dipole moment in the case of a molecule,

$$p_\alpha = -\frac{\partial E}{\partial E_\alpha}. \quad (8)$$

Physical properties connected with second derivatives of total energy are often referred to as linear-response coefficients. The connection between these two different views on the same property is not difficult to establish. As an example, suppose that an electric field is applied to a molecule that has no spontaneous dipole moment. At the lowest order, the induced dipole moment is

$$p_\alpha(\mathbf{E}) = \sum_\beta \alpha_{\alpha\beta} E_\beta + \dots \quad (9)$$

The second-rank tensor $\alpha_{\alpha\beta}$ is the polarizability tensor. It describes the linear response of the molecule. But the dipole moment of a molecule is connected to the derivative of the energy with respect to an electric field, by Eq. (8), so that

$$\alpha_{\alpha\beta} = -\frac{\partial^2 E}{\partial E_\alpha \partial E_\beta} \quad (10)$$

and the polarizability tensor is indeed a second-order derivative of the energy.

The description of the molecular vibrations, in the harmonic approximation, is based on the knowledge of the “interatomic force constants,” second derivatives of the total energy with respect to the change of nuclei positions, describing the forces that appear when a nucleus is displaced from its equilibrium position, in the linear regime.

Going in the nonlinear regime, one has higher-order derivatives, such as the nonlinear dielectric tensor (especially interesting for nonlinear optics [52]) and anharmonic force constants (allowing one to predict vibrational properties beyond the harmonic approximation [53]). In the study of solid-state phase transitions, derivatives up to the eighth order could be needed [54].

Also of interest are the mixed derivatives of the total energy with respect to different perturbations, but these will not be treated in the present paper. They are addressed in Refs. [10–13,20–39].

E. Sum-over-states approach, Green’s functions, Sternheimer equation, and Hylleraas variational principle

Different techniques have been proposed to address the treatment of perturbations within DFT. In order to un-

derstand their differences, one can apply them to the simpler one-body Schrödinger equation, at first order. What follows is a summary of this analysis.

The Schrödinger equation

$$[H(\lambda) - E_\alpha(\lambda)]|\Phi_\alpha(\lambda)\rangle = 0, \quad (11)$$

with the normalization condition

$$\langle \Phi(\lambda) | \Phi(\lambda) \rangle = 1, \quad (12)$$

is expanded at first order. This gives the Sternheimer equation [48]

$$(H - E_\alpha)^{(0)}|\Phi_\alpha^{(1)}\rangle = -(H - E_\alpha)^{(1)}|\Phi_\alpha^{(0)}\rangle \quad (13)$$

and the constraint

$$\langle \Phi_\alpha^{(0)} | \Phi_\alpha^{(1)} \rangle + \langle \Phi_\alpha^{(1)} | \Phi_\alpha^{(0)} \rangle = 0. \quad (14)$$

Because of the phase freedom

$$|\tilde{\Phi}(\lambda)\rangle = e^{if(\lambda)}|\Phi(\lambda)\rangle \quad (15)$$

of Eqs. (11) and (12), one can choose to work with the stronger constraint [compare with Eq. (14)]

$$\langle \Phi_\alpha^{(0)} | \Phi_\alpha^{(1)} \rangle = 0. \quad (16)$$

The first-order Hamiltonian is supposed to be known, while our aim is to determine the first-order wave function and energies.

The premultiplication of the Sternheimer equation (13) by the wave function $\langle \Phi_\alpha^{(0)} |$ gives the first-order eigenenergy

$$E_\alpha^{(1)} = \langle \Phi_\alpha^{(0)} | H^{(1)} | \Phi_\alpha^{(0)} \rangle. \quad (17)$$

Thanks to this result (actually, the Hellmann-Feynman theorem [55]), the right-hand side of the Sternheimer equation is known, and the only unknown is the first-order wave function contained in the left-hand side. Thus the Sternheimer equation is a *nonhomogeneous* linear differential equation that can be solved by different algorithms. This is especially easy for atomic problems, the original domain of application of this technique [6–8,14].

Another method appears if the Green’s function in the subspace perpendicular to $|\Phi_\alpha^{(0)}\rangle$ is explicitly known (this occurs in one-dimensional geometries [6]). The Green’s function is defined as

$$G_{\perp\alpha}(E_\alpha) = [P_{\perp\alpha}(E_\alpha - H)^{(0)}P_{\perp\alpha}]^{-1}, \quad (18)$$

with $P_{\perp\alpha}$ being the projector on the subspace orthogonal to $|\Phi_\alpha^{(0)}\rangle$. Then, the Sternheimer equation becomes

$$|\Phi_\alpha^{(1)}\rangle = G_{\perp\alpha}(E_\alpha)H^{(1)}|\Phi_\alpha^{(0)}\rangle, \quad (19)$$

which can be the starting point of a numerical algorithm.

In an eventual further step, the Green’s function can be expressed as a sum over states as usual in perturbation

theory, so that

$$|\Phi_\alpha^{(1)}\rangle = \sum_{\beta \neq \alpha} |\Phi_\beta^{(0)}\rangle \frac{1}{E_\alpha - E_\beta} \langle \Phi_\beta^{(0)} | H^{(1)} | \Phi_\alpha^{(0)} \rangle \quad (20)$$

but the summation, over an infinite number of states, must usually be truncated, in practical problems (note that if a finite basis set is used, the truncation is naturally imposed).

Finally, the Hylleraas variational principle [46], derived from the perturbation expansion of the Rayleigh-Ritz variational principle, can as well be used:

$$\begin{aligned} E_0^{(2)} &\leq \hat{E}_{\text{constrained}}^{(2)}[\Phi_0^{(0)}, \Phi_{0,\text{trial}}^{(1)}] \\ &= \langle \Phi_{0,\text{trial}}^{(1)} | H^{(1)} | \Phi_0^{(0)} \rangle \\ &\quad + \langle \Phi_{0,\text{trial}}^{(1)} | (H - E_0)^{(0)} | \Phi_{0,\text{trial}}^{(1)} \rangle \\ &\quad + \langle \Phi_0^{(0)} | H^{(2)} | \Phi_0^{(0)} \rangle + \langle \Phi_0^{(0)} | H^{(1)} | \Phi_{0,\text{trial}}^{(1)} \rangle. \end{aligned} \quad (21)$$

Note that the functional \hat{E} and the value that this functional can take \mathbf{E} are distinguished in Eq. (21) as well as in what follows. Equation (21) is a variational expression for $E_0^{(2)}$ with respect to $\Phi_{0,\text{trial}}^{(1)}$ under the constraint Eq. (14). When the latter is equal to $\Phi_0^{(1)}$, the variational expression is minimum. So, minimization algorithms (like the conjugate-gradient algorithm [56]) can be used in this case in order to find the first-order wave function. Although the Sternheimer equation, Green's function method, the sum-over-states approach, and the Hylleraas variational principle lead to the same $\Phi_0^{(1)}$, their algorithmic implementation can differ widely, and, depending on the system or the basis set, one method might be preferable to another.

F. Development of DFPT

1. Linear responses

In the atomic physics community, the perturbative calculation of the linear response to homogeneous electric field perturbations within DFT was introduced [58] by Zangwill and Soven [7], Stott and Zaremba [6], and Mahan [8].

These authors relied on a generalization of the Sternheimer equation, or on the Green's function technique. More information can be found in the book by Mahan [14]. Another approach, advocated by Ghosh and Deb [9], relies on the Hylleraas variational principle, and has also been used extensively [16].

The study of other perturbations was considered in the solid-state and quantum chemistry community, especially the response to displacement of atoms. Independently, Zein [10] and Baroni, Giannozzi, and Testa [11] have developed a Sternheimer equation approach to the calculation of different properties of the solid state, such as phonon frequencies, elastic constants, dielectric tensor, effective charges, and piezoelectric tensors [17,18,20–23,25–30]. Variations on a similar theme using sum over states can be found in Refs. [12,31,32] (development of the formalism for molecular calculations). The Hylleraas variational technique has been adapted by

Gonze, Allan, and Teter [13] for second-order DFPT, also with numerous applications [24].

When the (finite) basis set used for the representation of the wave function depends on the perturbation [58], “incomplete basis set” corrections must be introduced, in the spirit of “Pulay forces” within the Hellmann-Feynman theorem [59]. This problem has also been addressed by different groups [12,31,32,35,37] (see Sec. III E).

2. Higher-order responses

The theory of higher-order derivatives has been studied by Gonze and Vigneron [39]. They have shown how the $2n + 1$ theorem of perturbation theory generalizes to the DFPT. For example, it is possible to calculate easily the third-order derivative of the total energy by using the first-order derivative of the wave functions. Successive orders of perturbation are obtained iteratively, the solution of one given order allowing one to begin solving for the next higher order, for which a generalized Sternheimer equation is available.

The $2n + 1$ theorem and Sternheimer equations were (implicitly) expressed in that paper in the diagonal gauge (see Sec. III). The use of the parallel transport gauge is more convenient, as shown in Sec. III of the present paper for arbitrary order of perturbation, and independently by Debernardi and Baroni [41] for the third-order case.

Connected work has been presented recently by Dal Corso and Mauri [42], who work in a generalization of the Kohn-Sham functional to nonorthogonal orbitals. As in the present work, they rely on the variational principle.

G. Perturbation theory and the variational principle

When perturbation theory is applied to a quantity for which a variational principle holds, an interesting mathematical structure is observed: variational principles for even orders of perturbation, a “ $2n + 1$ theorem,” as well as “interchange theorems.” The early review of Hirschfelder, Byers, Brown, and Epstein [38] mentions results obtained for unconstrained functionals, while the extension to constrained functionals can be found in Refs. [61,47]. The notation of Ref. [47] will be used in what follows.

1. Expansion of an unconstrained variational principle

Let $E_0(\lambda)$ be an energy function of λ , obtained by minimization of the functional $\hat{E}_{(\lambda)}[\Phi_{\text{trial}}]$ in the space of possible wave functions Φ_{trial} :

$$E_0(\lambda) = \min_{\Phi_{\text{trial}}} \{ \hat{E}_{(\lambda)}[\Phi_{\text{trial}}] \}. \quad (22)$$

The value of Φ_{trial} at the minimum, for a given value of λ , is written $\Phi_0(\lambda)$.

As a first interesting result, the knowledge of the expansion of $\Phi_0(\lambda)$ up to order n allows one to determine the expansion of $E_0(\lambda)$ up to order $2n + 1$. Explicitly

$$E_0^{(2n+1)} = \left\{ \hat{E}_{(\lambda)} \left[\sum_{i=0}^n \lambda^i \Phi_0^{(i)} \right] \right\}^{(2n+1)}. \quad (23)$$

At order $2n$, one even finds a variational (minimum) principle:

$$E_0^{(2n)} = \min_{\delta\Phi_t} \left\{ \left(\hat{E}_{(\lambda)} \left[\sum_{i=0}^n \lambda^i \Phi_0^{(i)} + \lambda^n \delta\Phi_t \right] \right)^{(2n)} \right\}. \quad (24)$$

At the minimum, $\delta\Phi_t$ is equal to $\Phi_0^{(n)}$, and this fact can be the basis of a method to find $\Phi_0^{(n)}$, in the spirit of the Hylleraas variational principle. So, for each even derivative of the energy, there is a variational principle.

$$E_0^{(2n+1)} = \left\{ \hat{E}_{(\lambda)} \left[\sum_{i=0}^n \lambda^i \Phi_0^{(i)} \right] - \sum_{\alpha=1}^N \Lambda_{\alpha,0}(\lambda) \hat{C}_{\alpha,(\lambda)} \left[\sum_{i=0}^n \lambda^i \Phi_0^{(i)} \right] \right\}^{(2n+1)} \quad (27)$$

and the even-order variational principles become

$$E_0^{(2n)} = \min_{\delta\Phi \text{ satisfying Eq. (29)}} \left\{ \left(\hat{E}_{(\lambda)} \left[\sum_{i=0}^{n-1} \lambda^i \Phi_0^{(i)} + \lambda^n \delta\Phi_t \right] - \sum_{\alpha=1}^N \Lambda_{\alpha,0}(\lambda) \hat{C}_{\alpha,(\lambda)} \left[\sum_{i=0}^{n-1} \lambda^i \Phi_0^{(i)} + \lambda^n \delta\Phi_t \right] \right)^{(2n)} \right\}, \quad (28)$$

where

$$\left\{ \hat{C}_{\alpha,(\lambda)} \left[\sum_{i=0}^{n-1} \lambda^i \Phi_0^{(i)} + \lambda^n \delta\Phi_t \right] \right\}^{(n)} = 0. \quad (29)$$

Again, the knowledge of the expansion of $\Phi_0(\lambda)$ up to order n allows one to know the expansion of $E_0(\lambda)$ up to order $2n + 1$. More explicit expressions corresponding to Eqs. (27) and (28) are given in Ref. [47].

To summarize, with or without constraints, a $2n + 1$ theorem is derived, as well as variational principles for even-order derivatives of the energy. The expressions that appear in the case with constraints include the functional and its derivative, but also the Lagrange multiplier and its derivative, while, in that case, the variational principles for even derivatives of the energy are obtained under constraints.

III. PERTURBATIVE TREATMENT OF THE KOHN-SHAM ENERGY FUNCTIONAL

A. Variational formulation of the Kohn-Sham energy functional

First, the basic equations of density-functional formalism are written in a form convenient for application of perturbation theory. The energy of the electronic system [61] is given by

2. Expansion of a variational principle valid under constraints

The theory of Ref. [47] was given for a minimization problem under a single constraint, and is now trivially generalized to minimizations under many constraints. The Lagrange multiplier method is central to this approach, and is recalled in the Appendix. The minimization of $\hat{E}_{(\lambda)}[\Phi_{\text{trial}}]$ is performed under a set of N constraints,

$$\hat{C}_{\alpha,(\lambda)}[\Phi_{\text{trial}}] = 0 \quad \text{for } \alpha = 1, \dots, N \quad (25)$$

so that

$$E_0(\lambda) = \min_{\hat{C}_{\alpha,(\lambda)}[\Phi_{\text{trial}}]=0} \{ \hat{E}_{(\lambda)}[\Phi_{\text{trial}}] \}. \quad (26)$$

The $2n + 1$ theorem becomes

$$E[\Phi_\alpha] = \sum_{\alpha=1}^N \langle \Phi_\alpha | T + v | \Phi_\alpha \rangle + E_{\text{Hxc}}[n]. \quad (30)$$

The ‘‘interaction energy’’ $E_{\text{Hxc}}[n]$ is the sum of Hartree and exchange-correlation parts [3],

$$E_{\text{Hxc}}[n] = E_H[n] + E_{\text{xc}}[n], \quad (31)$$

that depend on the Kohn-Sham (KS) orbitals $|\psi_\alpha\rangle$ through the density

$$n(\mathbf{r}) = \sum_{\alpha=1}^N \Phi_\alpha^*(\mathbf{r}) \Phi_\alpha(\mathbf{r}). \quad (32)$$

The expression Eq. (30) must be minimized with respect to variations of the wave functions, subject to the following orthonormalization constraints [62]:

$$\langle \Phi_\alpha | \Phi_\beta \rangle = \delta_{\alpha\beta}, \quad (33)$$

valid for all ‘‘occupied’’ orbitals (α and β run from 1 to N , where N is the number of orbitals, equal to half the number of electrons if the system is spin degenerate). Equation (33) easily generalizes to the spin DFT [3].

The Lagrange multiplier method, applied to this minimization problem, allows one to recover the usual Kohn-Sham equations, modulo a unitary transform within the space of occupied orbitals, as follows. One Lagrange multiplier for each orthonormalization constraint is introduced, such that

$$F[\Phi_\alpha] = \sum_{\alpha}^N \langle \Phi_\alpha | T + v | \Phi_\alpha \rangle + E_{\text{Hxc}}[n] - \sum_{\alpha, \beta=1}^N \Lambda_{\beta\alpha} [\langle \Phi_\alpha | \Phi_\beta \rangle - \delta_{\alpha\beta}] . \quad (34)$$

The corresponding Lagrange-Euler equation is

$$H|\Phi_\alpha\rangle = \sum_{\beta=1}^N \Lambda_{\beta\alpha} |\Phi_\beta\rangle , \quad (35)$$

where the Hamiltonian

$$H = T + v_{\text{KS}} = T + v + v_{\text{Hxc}} \quad (36)$$

is self-consistently determined: it includes the Hartree and exchange-correlation potential (a local potential, but a nonlocal functional of the density)

$$v_{\text{Hxc}}(\mathbf{r}) = \frac{\delta E_{\text{Hxc}}[n]}{\delta n(\mathbf{r})} . \quad (37)$$

Equation (35), valid for $\alpha = 1, \dots, N$, obviously differs from the Kohn-Sham equation

$$H|\Phi_\alpha\rangle = \varepsilon_\alpha |\Phi_\alpha\rangle . \quad (38)$$

The difference stems from the invariance of the total energy and density under a unitary transform among the occupied wave functions (gauge freedom). Indeed, if U is an $(N \times N)$ unitary matrix:

$$[U^{-1}]_{\alpha\beta} = [U^\dagger]_{\alpha\beta} \quad \text{or} \quad \delta_{\alpha\beta} = \sum_{\gamma=1}^N U_{\alpha\gamma}^* U_{\beta\gamma} = \sum_{\gamma=1}^N U_{\gamma\alpha}^* U_{\gamma\beta} \quad (39)$$

and if another set of wave functions $|\Phi'_\alpha\rangle$ is defined as

$$|\Phi'_\alpha\rangle = \sum_{\gamma=1}^N U_{\gamma\alpha} |\Phi_\gamma\rangle , \quad (40)$$

then both sets of wave functions $|\Phi_\alpha\rangle$ and $|\Phi'_\alpha\rangle$ give the same density, electronic energy, and fulfill the orthonormalization constraints Eq. (33). This gauge freedom [group $U(N)$] is much wider than the one obtained by simply changing the phase of each Kohn-Sham orbital.

Now, let us suppose that the set of wave functions $|\Phi_\alpha\rangle$ is the solution of the Kohn-Sham equations (38), and $|\Phi'_\alpha\rangle$ is the solution of Eq. (35). It can be checked that the unitary matrix U needed in Eq. (40) is the one that diagonalizes the matrix of Lagrange parameters:

$$\sum_{\beta, \gamma=1}^N [U]_{\alpha\beta} \Lambda_{\beta\gamma} [U^\dagger]_{\gamma\eta} = \delta_{\alpha\eta} \varepsilon_\alpha . \quad (41)$$

The Lagrange parameters are obtained by taking the scalar product of Eq. (35) with an occupied wave function, which gives

$$\Lambda_{\beta\alpha} = \langle \Phi_\beta | H | \Phi_\alpha \rangle . \quad (42)$$

So the Lagrange parameters are the matrix elements of

the Hamiltonian between the two corresponding wave functions.

Note the important point that the eigenvalues, the Lagrange parameters, and the energy of the system are all different quantities, while in the similar derivation of the one-body Schrödinger equation from the Ritz variational principle all are identical.

B. Derivatives of energy, density, and wave functions

Now, we proceed to the perturbation expansion of these basic equations. We do not impose a definite gauge (this problem will be addressed later), and work with Lagrange multipliers. The perturbation expansion of Eqs. (32), (33), (35)–(37), and (42) is as follows.

(1) The density formula Eq. (32) becomes, at order i ,

$$n^{(i)}(\mathbf{r}) = \sum_{j=0}^i \sum_{\alpha=1}^N \Phi_\alpha^{(j)*}(\mathbf{r}) \Phi_\alpha^{(i-j)}(\mathbf{r}) . \quad (43)$$

(2) The orthonormalization conditions Eq. (33) become, at order i ,

$$\sum_{j=0}^i \langle \Phi_\alpha^{(j)} | \Phi_\beta^{(i-j)} \rangle = 0 \quad \text{for } i \geq 1 . \quad (44)$$

(3) Equation (35) becomes, at order i ,

$$\sum_{j=0}^i H^{(j)} |\Phi_\alpha^{(i-j)}\rangle = \sum_{j=0}^i \sum_{\beta=0}^N \Lambda_{\beta\alpha}^{(j)} |\Phi_\beta^{(i-j)}\rangle . \quad (45)$$

(4) The expression for the Hamiltonian Eq. (36) becomes, at order i ,

$$H^{(i)} = T^{(i)} + v_{\text{KS}}^{(i)} = T^{(i)} + v^{(i)} + v_{\text{Hxc}}^{(i)} . \quad (46)$$

(5) The Hartree and exchange-correlation potential (37) at order i is given by

$$v_{\text{Hxc}}^{(i)}(\mathbf{r}) = \left\{ \frac{\delta E_{\text{Hxc}}[n(\lambda)]}{\delta n(\mathbf{r})} \right\}^{(i)} = \frac{1}{i!} \frac{d^i}{d\lambda^i} \left[\frac{\delta E_{\text{Hxc}}[n(\lambda)]}{\delta n(\mathbf{r})} \right] \Bigg|_{\lambda=0} \quad (47)$$

$$= \frac{1}{i!} \frac{d^i}{d\lambda^i} \left[\frac{\delta E_{\text{Hxc}} \left[\sum_{j=0}^i \lambda^j n^{(j)} \right]}{\delta n(\mathbf{r})} \right] \Bigg|_{\lambda=0} . \quad (48)$$

(6) The Lagrange multiplier Eq. (42) becomes, at order i ,

$$\Lambda_{\beta\alpha}^{(i)} = \sum_{j=0}^i \sum_{k=0}^i \langle \Phi_\beta^{(j)} | H^{(i-j-k)} | \Phi_\alpha^{(k)} \rangle . \quad (49)$$

Finally, Eq. (27) or (28), applied to the variational expression Eq. (40), gives (for $m = 2n$ or $m = 2n + 1$)

$$\begin{aligned}
E^{(m)} &= \sum_{\alpha=1}^N \sum_{j=0}^n \sum_{k=0}^m \sum_{l=0}^n \delta(m-j-k-l) \langle \Phi_{\alpha}^{(j)} | (T+v)^{(k)} | \Phi_{\alpha}^{(l)} \rangle \\
&\quad + \frac{1}{m!} \frac{d^m}{d\lambda^m} E_{\text{Hxc}} \left[\sum_{\alpha=1}^N \left(\sum_{j=0}^n \lambda^j \Phi_{\alpha}^{(j)*}(\mathbf{r}) \right) \left(\sum_{k=0}^n \lambda^k \Phi_{\alpha}^{(k)}(\mathbf{r}) \right) \right] \Big|_{\lambda=0} \\
&\quad - \sum_{\alpha,\beta=1}^N \sum_{j=0}^n \sum_{k=0}^m \sum_{l=0}^n \delta(m-j-k-l) \Lambda_{\beta\alpha}^{(k)} \langle \Phi_{\alpha}^{(j)} | \Phi_{\beta}^{(l)} \rangle + \sum_{\alpha,\beta=1}^N \Lambda_{\beta\alpha}^{(m)} \delta_{\alpha\beta} .
\end{aligned} \tag{50}$$

Thanks to Eqs. (43), (44), (46), and (48), this equation can be worked out to give

$$\begin{aligned}
E^{(m)} &= \sum_{\alpha=1}^N \sum_{j=0}^n \sum_{k=0}^{m-n-1} \sum_{l=0}^n \delta(m-j-k-l) \langle \Phi_{\alpha}^{(j)} | H^{(k)} | \Phi_{\alpha}^{(l)} \rangle - \sum_{\alpha,\beta=1}^N \sum_{j=0}^n \sum_{k=0}^{m-n-1} \sum_{l=0}^n \delta(2n-j-k-l) \Lambda_{\beta\alpha}^{(k)} \langle \Phi_{\alpha}^{(j)} | \Phi_{\beta}^{(l)} \rangle \\
&\quad + \sum_{\alpha=1}^N \sum_{j=0}^n \sum_{k=m-n}^m \sum_{l=0}^n \delta(m-j-k-l) \langle \Phi_{\alpha}^{(j)} | (T+v)^{(k)} | \Phi_{\alpha}^{(l)} \rangle + \frac{1}{m!} \frac{d^m}{d\lambda^m} E_{\text{Hxc}} \left[\sum_{j=0}^n \lambda^j n^{(j)} \right] \Big|_{\lambda=0} .
\end{aligned} \tag{51}$$

As promised, only wave functions up to order n are needed, and, for $m = 2n$, this functional is variational (actually minimal) with respect to the n th-order wave functions, under the constraints Eq. (44). Although the expansions of the kinetic and potential operators up to order m appear in this expression, these are supposed to be known, while the expansions of the Hamiltonian, the density, or the exchange-correlation potential that are not an input of the problem, only appear up to order $m - n - 1$ (that is, either $n - 1$ or n).

Equations (43) and (46)–(48) were already obtained in Ref. [39], while Eqs. (44), (45), and (49)–(51) differ from their counterparts in the same paper. In particular, the minimal property of Eq. (51) is not shared by Eq. (A3) of Ref. [39].

At this level, we have an iterative scheme for treating perturbations at arbitrary order: suppose that the wave function density and Hamiltonian have been calculated up to order $n - 1$, then the variational principle Eq. (51) for order $2n$ can be used to generate the n th-order derivative of the wave functions, concurrently with the n th-order derivative of the Hamiltonian, density, or exchange-correlation potential.

As an alternative scheme, note that Eq. (45) can be cast into the form of a generalized Sternheimer equation [see next section, Eq. (54)], that can also be solved self-consistently with Eqs. (43), (44), (46), (48), and (49).

C. The parallel-transport gauge

Now, we discuss the choice of the gauge. There is no reason not to take zero-order wave functions that satisfy the Kohn-Sham equation (38). This means

$$\Lambda_{\beta\alpha}^{(0)} = \delta_{\beta\alpha} \varepsilon_{\alpha}^{(0)} \tag{52}$$

and

$$U_{\beta\alpha}^{(0)} = \delta_{\beta\alpha} . \tag{53}$$

What about higher-order wave functions? Let us first transform Eq. (45) to a modified Sternheimer equation (it is called [8] “modified” because of the self-consistency),

$$\begin{aligned}
(H^{(0)} - \varepsilon_{\alpha}^{(0)}) | \Phi_{\alpha}^{(i)} \rangle &= - \sum_{j=1}^i \left(H^{(j)} | \Phi_{\alpha}^{(i-j)} \rangle \right. \\
&\quad \left. - \sum_{\beta=0}^N \Lambda_{\beta\alpha}^{(j)} | \Phi_{\beta}^{(i-j)} \rangle \right) .
\end{aligned} \tag{54}$$

The projection of this equation on the subspace of unoccupied unperturbed wave functions gives

$$P_c (H^{(0)} - \varepsilon_{\alpha}^{(0)}) P_c | \Phi_{\alpha}^{(i)} \rangle = - \sum_{j=1}^i P_c H^{(j)} | \Phi_{\alpha}^{(i-j)} \rangle , \tag{55}$$

where P_c is the projector on that subspace. The part of the wave functions that is contained in the subspace spanned by the occupied (valence) wave functions is to be determined by the expansion of the orthonormalization condition Eq. (44) and the choice of gauge. The first one fixes the symmetric combination

$$\langle \Phi_{\alpha}^{(0)} | \Phi_{\beta}^{(i)} \rangle + \langle \Phi_{\alpha}^{(i)} | \Phi_{\beta}^{(0)} \rangle = \begin{cases} - \sum_{j=1}^{i-1} \langle \Phi_{\alpha}^{(j)} | \Phi_{\beta}^{(i-j)} \rangle & \text{for } i > 1 \\ 0 & \text{for } i = 1 . \end{cases} \tag{56}$$

The parallel-transport gauge is defined by

$$\langle \Phi_{\alpha}^{(0)} | \Phi_{\beta}^{(i)} \rangle - \langle \Phi_{\alpha}^{(i)} | \Phi_{\beta}^{(0)} \rangle = 0 , \tag{57}$$

so that, in this gauge,

$$\langle \Phi_{\alpha}^{(0)} | \Phi_{\beta}^{(i)} \rangle = \begin{cases} - \frac{1}{2} \sum_{j=1}^{i-1} \langle \Phi_{\alpha}^{(j)} | \Phi_{\beta}^{(i-j)} \rangle & \text{for } i > 1 \\ 0 & \text{for } i = 1 . \end{cases} \tag{58}$$

$$\tag{59}$$

It is the gauge that gives the smallest admixture of unperturbed occupied wave functions in higher-order derivatives of wave functions. Equations (56), (58), and (59) are sufficient for a complete determination of the valence subspace projection of the higher-order wave functions. Note that the *first*-order derivatives of the wave functions are perpendicular to the ground-state occupied wave functions, while this is no longer true for higher orders. Nevertheless, the part of the higher-order wave functions that belongs to the valence subspace is entirely fixed by Eq. (58), which proves useful for practical implementation. This result was not obtained in Ref. [39].

Combined with the Sternheimer equation (55) or with the variational principle Eq. (51), these two equations determine the higher-order wave functions completely.

As for the case of the one-body Hamiltonian, there are now four different ways of calculating the higher-order wave functions: either minimize Eq. (51), for $m = 2n$, under the constraints Eqs. (56), (58), and (59), or self-consistently solve the systems of Eqs. (43), (46), (48), (49), (56), (57), and either the Sternheimer equation (55), or a Green's function version of it [in the spirit of Eq. (22)] or a sum-over-states version of it [in the spirit of Eq. (23)].

When the wave functions are known through order n , the energy can be calculated up to order $2n + 1$ thanks to Eq. (51).

D. The diagonal gauge

In some cases, one is interested in the derivative of the Kohn-Sham eigenenergies. In the parallel-transport gauge, the Lagrange multiplier matrix is not diagonal. So, the explicit condition on the wave function Eqs. (58) and (59) of the parallel-transport gauge should be replaced by the requirement that the nondiagonal elements of the Lagrange multiplier matrix be zero at all orders. Let us call this gauge the "diagonal" gauge. In that gauge, the diagonal Lagrange multipliers are equal to the Kohn-Sham eigenenergies.

This condition is more difficult to implement in practical calculation than the parallel-transport gauge. Moreover, this requirement is not enough to fix the wave functions: first, each wave function still has a phase indeterminacy, as in the one-body case, second, in the case of degeneracy, there is still another unitary symmetry present beyond the "diagonal-gauge" requirement.

In practice, one can consider that the perturbation expansion has been solved in the parallel-transport gauge (so the expansion of the Lagrange multiplier matrix is considered known up to the needed order), and that the transformation to the diagonal gauge is performed afterwards. In order to do this, one needs to build the perturbation expansion of the unitary matrix connecting the different wave functions in Eq. (40). Let us define $|\Phi_{\parallel,\alpha}\rangle$ to be the wave functions in the parallel-transport gauge, and $|\Phi_{d,\alpha}\rangle$ to be the wave functions in the diagonal gauge. The unitary matrix U is defined as

$$|\Phi_{\parallel,\alpha}\rangle = \sum_{\gamma=1}^N U_{\gamma\alpha} |\Phi_{d,\gamma}\rangle \quad (60)$$

or

$$|\Phi_{d,\alpha}\rangle = \sum_{\gamma=1}^N U_{\alpha\gamma}^* |\Phi_{\parallel,\gamma}\rangle. \quad (61)$$

The basic results, Eqs. (39) and (41), for the unitary matrix were already mentioned. From these equations, one can also derive

$$\sum_{\gamma=1}^N \Lambda_{\beta\gamma} U_{\eta\gamma}^* = \varepsilon_{\eta} U_{\eta\beta}^*, \quad (62)$$

whose perturbation expansion gives

$$\sum_{j=0}^i \sum_{\gamma=1}^N \Lambda_{\beta\gamma}^{(j)} U_{\eta\gamma}^{(i-j)*} = \sum_{j=0}^i \varepsilon_{\eta}^{(j)} U_{\eta\beta}^{(i-j)*} \quad (63)$$

or

$$\begin{aligned} & \left(\sum_{\gamma=1}^N \Lambda_{\beta\gamma}^{(0)} U_{\eta\gamma}^{(i)*} \right) - \varepsilon_{\eta}^{(0)} U_{\eta\beta}^{(i)*} \\ &= - \sum_{j=1}^i \left[\left(\sum_{\gamma=1}^N \Lambda_{\beta\gamma}^{(j)} U_{\eta\gamma}^{(i-j)*} \right) - \varepsilon_{\eta}^{(j)} U_{\eta\beta}^{(i-j)*} \right]. \quad (64) \end{aligned}$$

The zero-order Lagrange multiplier matrix is known [see Eq. (52)], which gives

$$\begin{aligned} & (\varepsilon_{\beta}^{(0)} - \varepsilon_{\eta}^{(0)}) U_{\eta\beta}^{(i)*} \\ &= - \sum_{j=1}^i \left[\left(\sum_{\gamma=1}^N \Lambda_{\beta\gamma}^{(j)} U_{\eta\gamma}^{(i-j)*} \right) - \varepsilon_{\eta}^{(j)} U_{\eta\beta}^{(i-j)*} \right]. \quad (65) \end{aligned}$$

When β and η are different, we get

$$\begin{aligned} U_{\eta\beta}^{(i)*} &= - \frac{1}{\varepsilon_{\beta}^{(0)} - \varepsilon_{\eta}^{(0)}} \sum_{j=1}^i \left[\left(\sum_{\gamma=1}^N \Lambda_{\beta\gamma}^{(j)} U_{\eta\gamma}^{(i-j)*} \right) \right. \\ & \quad \left. - \varepsilon_{\eta}^{(j)} U_{\eta\beta}^{(i-j)*} \right], \quad (66) \end{aligned}$$

an equation for the nondiagonal part of the unitary matrix of order i in terms of unitary matrices of order less than i , and eigenvalues of order up to i .

Note that there can be some problem with this equation when the denominator vanishes (this happens when valence eigenenergies are degenerate). Actually, one could then define the diagonal gauge by making use of degenerate perturbation theory, or use a parallel-transport gauge within the space of degenerate wave functions. Leaving this problem for further studies, we continue the analysis.

In order to obtain the eigenvalues of order i , β and η are taken equal in Eq. (65), as follows:

$$0 = \sum_{j=1}^i \left[\left(\sum_{\gamma=1}^N \Lambda_{\beta\gamma}^{(j)} U_{\beta\gamma}^{(i-j)*} \right) - \varepsilon_{\beta}^{(j)} U_{\beta\beta}^{(i-j)*} \right] \quad (67)$$

and by Eq. (53),

$$\varepsilon_{\beta}^{(i)} = \begin{cases} \Lambda_{\beta\beta}^{(i)} + \sum_{j=1}^{i-1} \left[\left(\sum_{\gamma=1}^N \Lambda_{\beta\gamma}^{(j)} U_{\beta\gamma}^{(i-j)*} \right) - \varepsilon_{\beta}^{(j)} U_{\beta\beta}^{(i-j)*} \right] & \text{for } i > 1 \\ \Lambda_{\beta\beta}^{(i)} & \text{for } i = 1 . \end{cases} \quad (68)$$

$$\Lambda_{\beta\beta}^{(i)} \text{ for } i = 1 . \quad (69)$$

The eigenvalues of order i are given in terms of unitary matrices of order less than i , and Lagrange multipliers of order up to i . We still need an equation for the diagonal part of the unitary matrix. It is easily obtained from the perturbation expansion of the unitary condition Eq. (39), for $\alpha = \beta$,

$$1 = \sum_{\gamma=1}^N U_{\alpha\gamma}^* U_{\alpha\gamma} \quad (70)$$

giving

$$U_{\beta\beta}^{(i)*} + U_{\beta\beta}^{(i)} = \begin{cases} -\sum_{j=1}^{i-1} \left(\sum_{\gamma \neq \beta}^N U_{\gamma\beta}^{(j)*} U_{\gamma\beta}^{(i-j)} \right) & \text{for } i > 1 \\ 0 & \text{for } i = 1 . \end{cases} \quad (71)$$

$$(72)$$

The imaginary part of the diagonal elements of the unitary matrix is not determined by this equation, as a consequence of the phase freedom of each wave function. One can use the following gauge:

$$U_{\beta\beta}^{(i)*} - U_{\beta\beta}^{(i)} = 0 . \quad (73)$$

The relationship between perturbation expansion of wave functions in the parallel-transport gauge and the diagonal gauge is worked out from Eq. (61), and one obtains

$$|\Phi_{d,\alpha}^{(i)}\rangle = \sum_{j=0}^i \sum_{\gamma=1}^N U_{\alpha\gamma}^{(j)*} |\Phi_{\parallel,\gamma}^{(i-j)}\rangle . \quad (74)$$

It is possible to show that the results of Ref. [39] were actually obtained in the diagonal gauge, but the corresponding constraints were not made explicit.

E. Incomplete basis set corrections

Until now, it has been supposed that the ground-state wave function and its perturbation expansion can be determined with infinite accuracy. This is, of course, not the case. In particular, the representation of the wave function is often done in terms of a finite set of M basis functions, with M sufficiently large. Within this basis set, the variational principle Eq. (34) will be used in order to determine the optimal set of coefficients describing the wave function:

$$\Phi(\mathbf{r}) = \sum_{p=1}^M c_p f_p(\mathbf{r}) . \quad (75)$$

Often, the basis functions depend on the perturbation (the typical case being a basis of Gaussian functions attached to the atomic locations, while the perturbation is an atomic displacement). Equation (74) implies

$$\Phi^{(i)}(\mathbf{r}) = \sum_{p=1}^M \sum_{j=0}^i c_p^{(j)} f_p^{(i-j)}(\mathbf{r}) . \quad (76)$$

So, the change in the approximate wave function, at order i , must be computed by taking into account not only changes in the coefficient that multiply the basis functions, but also changes in the basis functions. The latter are usually known, so that the coefficients are unknown, as one could have expected. When the present scheme is adapted for such an incomplete basis set, the quantities to be varied are no longer the wave functions, but the coefficients describing it within the basis set. In terms of the expansion of these coefficients, it will be possible to find a $2n + 1$ theorem, as well as variational principles for the even-order terms of the energy expansion. The adaptation of the scheme presented here is not difficult, from a technical point of view. See Ref. [35] for the corresponding second-order expression.

Also, the representation of other quantities, such as densities, could depend on basis functions, in which case similar corrections must be taken into account [12,31,32]. But when the basis functions do not depend on the perturbation (the case of a homogeneous electric field, or atomic displacements when plane-wave expansions are used), this complication does not appear. The application presented in Sec. IV J will make use of this fact.

F. Influence of the perturbation on the form of kinetic energy, Hartree functional, and exchange-correlation functional

In Secs. III B–III D, very general formulas have been derived for an arbitrary interaction energy or kinetic operator that could depend on the smaller parameter of the perturbation in an arbitrary way. It is now in order to discuss cases in which this full dependence is not found, and the subsequent simplifications that occur.

The kinetic energy operator, and the form of the Hartree energy are, in most of the cases, insensitive to the applied perturbation (they are insensitive to the small parameter of all the perturbation mentioned in Sec. II B). An exception occurs when an expansion or contraction is applied to a solid [17]. By contrast, for all the perturbations mentioned in Secs. II B and II C, the local part of the potential will be modified.

The form of the exchange-correlation energy can be affected by the perturbation in a few cases: when, for

numerical reasons, the density is fitted by a few functions that depend on the atomic positions [12,31–33], or when a nonlinear exchange-correlation core correction is present [63]; also, when an expansion or contraction is applied to a solid [17,64].

For the consideration of low-order expansions, presented in the next section, the following notations are introduced:

$$E_{\text{Hxc},0}^{(i)} = \left\{ E_{\text{Hxc}}[n^{(0)}] \right\}^{(i)} = \left. \frac{1}{i!} \frac{d^i}{d\lambda^i} E_{\text{Hxc}}[n^{(0)}] \right|_{\lambda=0}, \quad (77)$$

$$v_{\text{Hxc},0}^{(i)}(\mathbf{r}) = \left\{ v_{\text{Hxc}}[n^{(0)}](\mathbf{r}) \right\}^{(i)} = \left. \frac{1}{i!} \frac{d^i}{d\lambda^i} \frac{\delta E_{\text{Hxc}}[n^{(0)}]}{\delta n(\mathbf{r})} \right|_{\lambda=0}, \quad (78)$$

$$K_{\text{Hxc},0}^{(i)}(\mathbf{r}, \mathbf{r}') = \left. \frac{1}{i!} \frac{d^i}{d\lambda^i} \frac{\delta E_{\text{Hxc}}[n^{(0)}]}{\delta n(\mathbf{r}) \delta n(\mathbf{r}')} \right|_{\lambda=0}. \quad (79)$$

If the form of the exchange-correlation functional does not depend on the perturbation, all the quantities Eqs. (77)–(79) vanish for i different than 0.

Taking now the most used approximation to the exchange-correlation functional, the local density approximation (LDA), one has

$$E_{\text{xc,LDA}}[n] = \int e_{\text{xc}}(n(\mathbf{r})) n(\mathbf{r}) d\mathbf{r}, \quad (80)$$

where $e_{\text{xc}}(n)$ is the exchange-correlation energy per particle in a homogeneous electron gas [3]. Associated derivatives, Eqs. (82) and (83), for $i = 0$, become

$$v_{\text{xc,LDA}}(\mathbf{r}) = \mu_{\text{xc}}(n(\mathbf{r})), \quad (81)$$

where

$$\mu_{\text{xc}}(n) = \frac{d(ne_{\text{xc}})}{dn} \quad (82)$$

and

$$K_{\text{xc,LDA}}(\mathbf{r}, \mathbf{r}') = \left. \frac{d\mu_{\text{xc}}}{dn} \right|_{n(\mathbf{r})} \delta(\mathbf{r} - \mathbf{r}'). \quad (83)$$

In the generalized gradient approximation (GGA),

$$E_{\text{xc,GGA}}[n] = \int f_{\text{xc}}(n(\mathbf{r}), g(\mathbf{r})) d\mathbf{r}, \quad (84)$$

the function f_{xc} depends on the density and its gradient,

$$g(\mathbf{r}) = |\nabla n(\mathbf{r})| \quad (85)$$

and can take many different forms [65]. Derivatives can be found using either the calculus of variation, as in Refs. [26] or [65], or a more elaborate treatment, following Ref. [33].

IV. THE LOWEST ORDERS OF PERTURBATION IN DFPT

The general expressions of the preceding section are now specialized to the first-, second-, third-, and fourth-order derivatives of the energy, the most important for practical applications. The case in which the perturbation does not affect the kinetic energy and the Hartree and exchange-correlation energy is treated explicitly, as well as the case in which the perturbation takes the form of a simple change of local potential. The latter form is found when an electric field is applied to a finite system, or when an atomic displacement is treated in an all-electron framework, with the Coulombic potential. In a pseudopotential framework, an atomic displacement will usually lead to a nonlocal change in the potential, but the form of the kinetic and interaction energies will not change, except if a nonlinear core correction is taken into account.

A. The first-order energy

The first-order derivative of the electronic energy is

$$E^{(1)} = \left. \sum_{\alpha=1}^N \langle \Phi_{\alpha}^{(0)} | (T + v)^{(1)} | \Phi_{\alpha}^{(0)} \rangle + \frac{d}{d\lambda} E_{\text{Hxc}}[n^{(0)}] \right|_{\lambda=0}. \quad (86)$$

This formula is the translation of the Hellmann-Feynman theorem to the density-functional formalism [3]. The second part will be nonzero only if the interaction functional of the density has an explicit dependence on λ . No knowledge of the perturbed wave functions is needed in Eq. (86), which is a considerable advantage.

If the kinetic and interaction energies do not depend explicitly on the perturbation, one has

$$E^{(1)} = \sum_{\alpha=1}^N \langle \Phi_{\alpha}^{(0)} | v^{(1)} | \Phi_{\alpha}^{(0)} \rangle \quad (87)$$

and if, moreover, the change in the potential is local, one has

$$E^{(1)} = \int v^{(1)}(\mathbf{r}) n^{(0)}(\mathbf{r}) d\mathbf{r}, \quad (88)$$

a particularly simple formula, in which only the ground-state density and first-order external potential change appear. Note that Eq. (87) opens the way to a state-by-state decomposition of $E^{(1)}$, since one can define

$$E_{\alpha}^{(1)} = \langle \Phi_{\alpha}^{(0)} | v^{(1)} | \Phi_{\alpha}^{(0)} \rangle \quad (89)$$

such that

$$E^{(1)} = \sum_{\alpha=1}^N E_{\alpha}^{(1)}. \quad (90)$$

This can prove useful in quantifying (to first order) the orbital origin of changes in energy.

B. The second-order energy: Variational form

The following expression is variational (a minimum principle) with respect to $|\Phi_{\alpha}^{(1)}\rangle$:

$$\begin{aligned} E^{(2)} = & \sum_{\alpha=1}^N [\langle \Phi_{\alpha}^{(1)} | (T + v)^{(1)} | \Phi_{\alpha}^{(0)} \rangle + \langle \Phi_{\alpha}^{(1)} | (H - \varepsilon_{\alpha})^{(0)} | \Phi_{\alpha}^{(1)} \rangle + \langle \Phi_{\alpha}^{(0)} | (T + v)^{(2)} | \Phi_{\alpha}^{(0)} \rangle \\ & + \langle \Phi_{\alpha}^{(0)} | (T + v)^{(1)} | \Phi_{\alpha}^{(1)} \rangle] + \frac{1}{2} \int \int \frac{\delta^2 E_{\text{Hxc}}[n^{(0)}]}{\delta n(\mathbf{r}) \delta n(\mathbf{r}')} n^{(1)}(\mathbf{r}) n^{(1)}(\mathbf{r}') d\mathbf{r} d\mathbf{r}' \\ & + \int \frac{d}{d\lambda} \frac{\delta E_{\text{Hxc}}[n^{(0)}]}{\delta n(\mathbf{r})} \Bigg|_{\lambda=0} n^{(1)}(\mathbf{r}) d\mathbf{r} + \frac{1}{2} \frac{d^2}{d\lambda^2} E_{\text{Hxc}}[n^{(0)}] \Bigg|_{\lambda=0}, \end{aligned} \quad (91)$$

with

$$n^{(1)}(\mathbf{r}) = \sum_{\alpha=1}^N \Phi_{\alpha}^{(0)*}(\mathbf{r}) \Phi_{\alpha}^{(1)}(\mathbf{r}) + \Phi_{\alpha}^{(1)*}(\mathbf{r}) \Phi_{\alpha}^{(0)}(\mathbf{r}). \quad (92)$$

The functional defined in Eq. (91) is minimal with respect to the first-order wave functions, subject to the following constraints:

$$\langle \Phi_{\alpha}^{(0)} | \Phi_{\beta}^{(1)} \rangle + \langle \Phi_{\alpha}^{(1)} | \Phi_{\beta}^{(0)} \rangle = 0. \quad (93)$$

In the parallel-transport gauge, one has

$$\langle \Phi_{\alpha}^{(0)} | \Phi_{\beta}^{(1)} \rangle = 0. \quad (94)$$

The latter requirement means that the projection of the first-order wave functions on the valence bands vanishes, in the parallel-transport gauge.

As promised, only zeroth- and first-order wave functions are needed in Eq. (91). The first-order Lagrange multipliers are not needed.

The terms in Eq. (91) have either quadratic or linear dependency on the first-order wave functions or are independent of them. Let us rewrite Eq. (91) using this information, and the definitions given in Sec. III F:

$$E^{(2)} = E_{\text{quad}}^{(2)} + E_{\text{lin}}^{(2)} + E_{\text{ind}}^{(2)}, \quad (95)$$

with

$$\begin{aligned} E_{\text{quad}}^{(2)} = & \sum_{\alpha=1}^N \langle \Phi_{\alpha}^{(1)} | (H - \varepsilon_{\alpha})^{(0)} | \Phi_{\alpha}^{(1)} \rangle \\ & + \frac{1}{2} \int \int K_{\text{Hxc},0}^{(0)}(\mathbf{r}, \mathbf{r}') n^{(1)}(\mathbf{r}) n^{(1)}(\mathbf{r}') d\mathbf{r} d\mathbf{r}', \end{aligned} \quad (96)$$

$$\begin{aligned} E_{\text{lin}}^{(2)} = & \sum_{\alpha=1}^N \langle \Phi_{\alpha}^{(1)} | (T + v + v_{\text{Hxc},0})^{(1)} | \Phi_{\alpha}^{(0)} \rangle \\ & + \langle \Phi_{\alpha}^{(0)} | (T + v + v_{\text{Hxc},0})^{(1)} | \Phi_{\alpha}^{(1)} \rangle, \end{aligned} \quad (97)$$

$$E_{\text{ind}}^{(2)} = \sum_{\alpha=1}^N \langle \Phi_{\alpha}^{(0)} | (T + v)^{(2)} | \Phi_{\alpha}^{(0)} \rangle + E_{\text{Hxc},0}^{(2)}. \quad (98)$$

It is now easy to see the consequence of the minimum principle property of Eq. (91): its quadratic part $E_{\text{quad}}^{(2)}$ has a positive-definite kernel. This quadratic part is made of two terms. The first,

$$\sum_{\alpha=1}^N \langle \Phi_{\alpha}^{(1)} | (H - \varepsilon_{\alpha})^{(0)} | \Phi_{\alpha}^{(1)} \rangle, \quad (99)$$

is a sum the index of which runs on all valence band wave functions: for each α , the difference between the expectation values of $H^{(0)}$ and $\varepsilon_{\alpha}^{(0)}$ is computed. Because the first-order wave functions entering this expression are constrained to be orthogonal to the valence Hilbert space [due to Eq. (94)], the matrix element of the Hamiltonian will always be larger or equal to the matrix element of the lowest unoccupied state eigenvalue:

$$\begin{aligned} & \langle \Phi_{\alpha}^{(1)} | (H - \varepsilon_{\alpha})^{(0)} | \Phi_{\alpha}^{(1)} \rangle \\ & \geq (\varepsilon_{\text{conduction}} - \varepsilon_{\alpha})^{(0)} \langle \Phi_{\alpha}^{(1)} | \Phi_{\alpha}^{(1)} \rangle, \end{aligned} \quad (100)$$

so that the kernel of Eq. (99) is obviously positive definite.

The second part of the quadratic term in Eq. (84) is

$$\frac{1}{2} \int \int K_{\text{Hxc},0}^{(0)}(\mathbf{r}, \mathbf{r}') n^{(1)}(\mathbf{r}) n^{(1)}(\mathbf{r}') d\mathbf{r} d\mathbf{r}'. \quad (101)$$

Note that $n^{(1)}(\mathbf{r})$ is a linear functional of the first-order wave functions, see Eq. (92). Introducing again the distinction between the Hartree and the exchange-correlation part, the kernel $K_{\text{Hxc},0}^{(0)}(\mathbf{r}, \mathbf{r}')$ is given by

$$K_{\text{Hxc},0}^{(0)}(\mathbf{r}, \mathbf{r}') = \frac{1}{|\mathbf{r} - \mathbf{r}'|} + \frac{\delta^2 E_{\text{xc}}[n^{(0)}]}{\delta n(\mathbf{r}) \delta n(\mathbf{r}')} \Bigg|_{\lambda=0}. \quad (102)$$

The Hartree part in this equation is positive definite. In

order to analyze the exchange-correlation part, we need an explicit expression. Working in the LDA, we find that this part is negative definite:

$$\frac{\delta^2 E_{\text{xc}}[n^{(0)}]}{\delta n(\mathbf{r})\delta n(\mathbf{r}')} \Bigg|_{n=n^{(0)}(\mathbf{r})} = \frac{d\mu_{\text{xc}}}{dn} \Bigg|_{n=n^{(0)}(\mathbf{r})} \delta(\mathbf{r} - \mathbf{r}') \quad (103)$$

since the derivative of the local exchange-correlation potential with respect to the density is always negative [3].

Although the exchange-correlation part does not have the same sign as the one-body and Hartree kernels, the positive-definite character of the quadratic part Eq. (96) is established through the minimal property of Eq. (91). This result makes sense, in that the exchange-correlation part is usually much smaller than either the kinetic or the Hartree energy.

It is also possible to gain knowledge of the sign of the different terms in Eq. (95) after the minimization procedure. Note that the quadratic part, when considered alone, is minimal when the first-order wave functions vanish (its value is zero). But, the part of Eq. (95) depending linearly on the first-order wave functions, Eq. (97), can become negative when the wave functions do not vanish, and gives a “force,” driving the first-order wave functions away from zero in the minimization process. In this case, the sum of these two terms will decrease down to a minimum, negative, value. So, at the minimum,

$$0 \geq E_{\text{quad}}^{(2)} + E_{\text{lin}}^{(2)}. \quad (104)$$

As a consequence,

$$E_{\text{ind}}^{(2)} \geq E^{(2)}. \quad (105)$$

Finally, if the perturbation does not affect the explicit form of the kinetic energy, and Hartree and exchange-correlation energy, Eqs. (97) and (98) reduce to

$$E_{\text{lin}}^{(2)} = \sum_{\alpha=1}^N \langle \Phi_{\alpha}^{(1)} | v^{(1)} | \Phi_{\alpha}^{(0)} \rangle + \langle \Phi_{\alpha}^{(0)} | v^{(1)} | \Phi_{\alpha}^{(1)} \rangle, \quad (106)$$

$$E_{\text{ind}}^{(2)} = \sum_{\alpha=1}^N \langle \Phi_{\alpha}^{(0)} | v^{(2)} | \Phi_{\alpha}^{(0)} \rangle. \quad (107)$$

Moreover, if the perturbation affects only the local potential, one gets

$$E_{\text{lin}}^{(2)} = \int v^{(1)}(\mathbf{r}) n^{(1)}(\mathbf{r}) d\mathbf{r}, \quad (108)$$

$$E_{\text{ind}}^{(2)} = \int v^{(2)}(\mathbf{r}) n^{(0)}(\mathbf{r}) d\mathbf{r}. \quad (109)$$

The formulas contained in the present subsection were applied in Refs. [9,13,16,24].

C. The third-order energy

The third-order electronic energy is given by

$$\begin{aligned} E^{(3)} = & \sum_{\alpha=1}^N [\langle \Phi_{\alpha}^{(1)} | (T + v)^{(2)} | \Phi_{\alpha}^{(0)} \rangle + \langle \Phi_{\alpha}^{(1)} | H^{(1)} | \Phi_{\alpha}^{(1)} \rangle + \langle \Phi_{\alpha}^{(0)} | (T + v)^{(3)} | \Phi_{\alpha}^{(0)} \rangle \\ & + \langle \Phi_{\alpha}^{(0)} | (T + v)^{(2)} | \Phi_{\alpha}^{(1)} \rangle] - \sum_{\alpha,\beta=1}^N \Lambda_{\beta\alpha}^{(1)} \langle \Phi_{\alpha}^{(1)} | \Phi_{\beta}^{(1)} \rangle \\ & + \frac{1}{6} \int \int \int \frac{\delta^3 E_{\text{Hxc}}[n^{(0)}]}{\delta n(\mathbf{r})\delta n(\mathbf{r}')\delta n(\mathbf{r}'')} n^{(1)}(\mathbf{r}) n^{(1)}(\mathbf{r}') n^{(1)}(\mathbf{r}'') d\mathbf{r} d\mathbf{r}' d\mathbf{r}'' \\ & + \frac{1}{2} \int \int \frac{d}{d\lambda} \left[\frac{\delta^2 E_{\text{Hxc}}[n^{(0)}]}{\delta n(\mathbf{r})\delta n(\mathbf{r}')} \right]_{\lambda=0} n^{(1)}(\mathbf{r}) n^{(1)}(\mathbf{r}') d\mathbf{r} d\mathbf{r}' \\ & + \frac{1}{2} \int \frac{d^2}{d\lambda^2} \left[\frac{\delta E_{\text{Hxc}}[n^{(0)}]}{\delta n(\mathbf{r})} \right]_{\lambda=0} n^{(1)}(\mathbf{r}) d\mathbf{r} + \frac{1}{6} \frac{d^3}{d\lambda^3} E_{\text{Hxc}}[n^{(0)}] \Bigg|_{\lambda=0}. \end{aligned} \quad (110)$$

Again, only zero- and first-order wave functions are needed in Eq. (110).

If the perturbation does not affect the explicit form of the kinetic energy and Hartree and exchange-correlation energy, one gets [41]

$$\begin{aligned} E^{(3)} = & \sum_{\alpha=1}^N (\langle \Phi_{\alpha}^{(1)} | v^{(2)} | \Phi_{\alpha}^{(0)} \rangle + \langle \Phi_{\alpha}^{(1)} | H^{(1)} | \Phi_{\alpha}^{(1)} \rangle + \langle \Phi_{\alpha}^{(0)} | v^{(3)} | \Phi_{\alpha}^{(0)} \rangle + \langle \Phi_{\alpha}^{(0)} | v^{(2)} | \Phi_{\alpha}^{(1)} \rangle) \\ & - \sum_{\alpha,\beta=1}^N \Lambda_{\beta\alpha}^{(1)} \langle \Phi_{\alpha}^{(1)} | \Phi_{\beta}^{(1)} \rangle + \frac{1}{6} \int \int \int \frac{\delta^3 E_{\text{xc}}[n^{(0)}]}{\delta n(\mathbf{r})\delta n(\mathbf{r}')\delta n(\mathbf{r}'')} n^{(1)}(\mathbf{r}) n^{(1)}(\mathbf{r}') n^{(1)}(\mathbf{r}'') d\mathbf{r} d\mathbf{r}' d\mathbf{r}'', \end{aligned} \quad (111)$$

Note that if the Hartree term (a quadratic form of the density) does not depend explicitly on the perturbation, its

third-order derivative vanishes. Consequently, the H subscript in Eq. (111) has been omitted.

Moreover, if the perturbation affects only the local potential part of the energy expression, one gets

$$E^{(3)} = \left(\sum_{\alpha=1}^N \langle \Phi_{\alpha}^{(1)} | H^{(1)} | \Phi_{\alpha}^{(1)} \rangle \right) - \left(\sum_{\alpha, \beta=1}^N \Lambda_{\beta\alpha}^{(1)} \langle \Phi_{\alpha}^{(1)} | \Phi_{\beta}^{(1)} \rangle \right) + \frac{1}{6} \int \int \int \frac{\delta^3 E_{\text{xc}}[n^{(0)}]}{\delta n(\mathbf{r}) \delta n(\mathbf{r}') \delta n(\mathbf{r}'')} n^{(1)}(\mathbf{r}) \\ \times n^{(1)}(\mathbf{r}') n^{(1)}(\mathbf{r}'') d\mathbf{r} d\mathbf{r}' d\mathbf{r}'' + \int v^{(2)}(\mathbf{r}) n^{(1)}(\mathbf{r}) d\mathbf{r} + \int v^{(3)}(\mathbf{r}) n^{(0)}(\mathbf{r}) d\mathbf{r} . \quad (112)$$

D. The fourth-order energy: Variational form

The complete fourth-order expression derived from Eq. (57) will not be written here [66]. If the kinetic operator and the form of the exchange-correlation energy do not depend upon the small parameter, and if the change in the potential is linear in this small parameter, the fourth-order expression simplifies as

$$E^{(4)} = \sum_{\alpha=1}^N [\langle \Phi_{\alpha}^{(2)} | (H - \varepsilon_{\alpha})^{(0)} | \Phi_{\alpha}^{(2)} \rangle + \langle \Phi_{\alpha}^{(2)} | H^{(1)} | \Phi_{\alpha}^{(1)} \rangle + \langle \Phi_{\alpha}^{(1)} | H^{(1)} | \Phi_{\alpha}^{(2)} \rangle] \\ - \sum_{\alpha, \beta=1}^N \Lambda_{\beta\alpha}^{(1)} [\langle \Phi_{\alpha}^{(2)} | \Phi_{\beta}^{(1)} \rangle + \langle \Phi_{\alpha}^{(1)} | \Phi_{\beta}^{(2)} \rangle] + \frac{1}{2} \int \frac{\delta^2 E_{\text{Hxc}}[n^{(0)}]}{\delta n(\mathbf{r}) \delta n(\mathbf{r}')} \Big|_{\lambda=0} n^{(2)}(\mathbf{r}) n^{(2)}(\mathbf{r}') d\mathbf{r} d\mathbf{r}' \\ + \frac{1}{2} \int \frac{\delta^3 E_{\text{Hxc}}[n^{(0)}]}{\delta n(\mathbf{r}) \delta n(\mathbf{r}') \delta n(\mathbf{r}'')} \Big|_{\lambda=0} n^{(1)}(\mathbf{r}) n^{(1)}(\mathbf{r}') n^{(2)}(\mathbf{r}'') d\mathbf{r} d\mathbf{r}' d\mathbf{r}'' \\ + \frac{1}{24} \int \frac{\delta^4 E_{\text{Hxc}}[n^{(0)}]}{\delta n(\mathbf{r}) \delta n(\mathbf{r}') \delta n(\mathbf{r}'') \delta n(\mathbf{r}''')} \Big|_{\lambda=0} n^{(1)}(\mathbf{r}) n^{(1)}(\mathbf{r}') n^{(1)}(\mathbf{r}'') n^{(1)}(\mathbf{r}''') d\mathbf{r} d\mathbf{r}' d\mathbf{r}'' d\mathbf{r}''' , \quad (113)$$

with

$$n^{(2)}(\mathbf{r}) = \sum_{\alpha=1}^N \Phi_{\alpha}^{(0)*}(\mathbf{r}) \Phi_{\alpha}^{(2)}(\mathbf{r}) + \Phi_{\alpha}^{(1)*}(\mathbf{r}) \Phi_{\alpha}^{(1)}(\mathbf{r}) \\ + \Phi_{\alpha}^{(2)*}(\mathbf{r}) \Phi_{\alpha}^{(0)}(\mathbf{r}) . \quad (114)$$

The functional defined in Eq. (113) is minimal with respect to the first-order wave functions, subject to the following constraints:

$$\langle \Phi_{\alpha}^{(0)} | \Phi_{\beta}^{(2)} \rangle + \langle \Phi_{\alpha}^{(1)} | \Phi_{\beta}^{(1)} \rangle + \langle \Phi_{\alpha}^{(2)} | \Phi_{\beta}^{(0)} \rangle = 0 . \quad (115)$$

In the parallel-transport gauge, one has

$$\langle \Phi_{\alpha}^{(0)} | \Phi_{\beta}^{(2)} \rangle = -\frac{1}{2} \langle \Phi_{\alpha}^{(1)} | \Phi_{\beta}^{(1)} \rangle . \quad (116)$$

The latter requirement fixes the projection of the first-order wave functions on the valence bands.

As promised, only zeroth-, first-, and second-order wave functions are needed in Eq. (113). Note that the second-order Lagrange multipliers are not needed.

In order to make apparent the variational property of the fourth-order electronic energy with respect to second-order wave functions, following Eqs. (95)–(98), one can write the fourth-order derivative of the total energy, Eq. (113), as

$$E^{(4)} = E_{\text{quad}}^{(4)} + E_{\text{lin}}^{(4)} + E_{\text{ind}}^{(4)} , \quad (117)$$

with

$$E_{\text{quad}}^{(4)} = \sum_{\alpha=1}^N \langle \Phi_{\alpha}^{(2)} | (H - \varepsilon_{\alpha})^{(0)} | \Phi_{\alpha}^{(2)} \rangle \\ + \frac{1}{2} \int \int K_{\text{Hxc},0}^{(0)}(\mathbf{r}, \mathbf{r}') \overline{n^{(2)}}(\mathbf{r}) \overline{n^{(2)}}(\mathbf{r}') d\mathbf{r} d\mathbf{r}' , \quad (118)$$

while

$$\overline{n^{(2)}}(\mathbf{r}) = \sum_{\alpha=1}^N \Phi_{\alpha}^{(0)*}(\mathbf{r}) \Phi_{\alpha}^{(2)}(\mathbf{r}) + \Phi_{\alpha}^{(2)*}(\mathbf{r}) \Phi_{\alpha}^{(0)}(\mathbf{r}) . \quad (119)$$

The other parts of Eq. (117) are easily derived, and will not be mentioned here for the sake of brevity.

As in the second-order expression Eq. (95), Eq. (117) is a quadratic form on the second-order wave function. Note that the kernel of the quadratic part, Eq. (118), is *exactly* the same as the kernel of the quadratic part of the second-order expression, Eq. (96). Hence the analysis of this kernel that was done in Sec. IV B is also valid for this fourth-order expression [67].

E. Calculation of the first-order wave functions and Lagrange multipliers in the parallel-transport gauge

In Eqs. (110)–(113) and (119), the first-order Lagrange multipliers are needed. These can be obtained by

$$\Lambda_{\beta\alpha}^{(1)} = \langle \Phi_{\beta}^{(0)} | H^{(1)} | \Phi_{\alpha}^{(0)} \rangle . \quad (120)$$

The first-order Hamiltonian is given by

$$H^{(1)} = T^{(1)} + v_{\text{KS}}^{(1)} = T^{(1)} + v^{(1)} + v_{\text{Hxc}}^{(1)} , \quad (121)$$

with

$$v_{\text{Hxc}}^{(1)}(\mathbf{r}) = \int \frac{\delta^2 E_{\text{Hxc}}[n^{(0)}]}{\delta n(\mathbf{r}) \delta n(\mathbf{r}')} n^{(1)}(\mathbf{r}') d\mathbf{r}' + \left. \frac{d}{d\lambda} \frac{\delta E_{\text{Hxc}}[n^{(0)}]}{\delta n(\mathbf{r})} \right]_{\lambda=0} \quad (122)$$

$$= \int K_{\text{Hxc},0}^{(0)}(\mathbf{r}, \mathbf{r}') n^{(1)}(\mathbf{r}') d\mathbf{r}' + v_{\text{Hxc},0}^{(1)}(\mathbf{r}) . \quad (123)$$

If one does not want to directly minimize Eq. (95) (or simply the combination of its quadratic and linear part) in order to find the first-order wave functions, one can use the modified Sternheimer equation, derived from Eq. (55):

$$P_c(H - \varepsilon_{\alpha})^{(0)} P_c | \Phi_{\alpha}^{(1)} \rangle = -P_c H^{(1)} | \Phi_{\alpha}^{(0)} \rangle . \quad (124)$$

F. The second-order energy: Nonvariational forms

Using Eq. (124), one can find alternate expressions for the second-order energy. From Eqs. (95), (122), (124), and their Hermitian conjugates, we derive

$$E^{(2)} = \sum_{\alpha=1}^N \langle \Phi_{\alpha}^{(1)} | (T + v)^{(1)} | \Phi_{\alpha}^{(0)} \rangle + \sum_{\alpha=1}^N \langle \Phi_{\alpha}^{(0)} | (T + v)^{(2)} | \Phi_{\alpha}^{(0)} \rangle + E_{\text{Hxc},0}^{(2)} , \quad (125)$$

$$E^{(2)} = \sum_{\alpha=1}^N \langle \Phi_{\alpha}^{(0)} | (T + v)^{(1)} | \Phi_{\alpha}^{(1)} \rangle + \sum_{\alpha=1}^N \langle \Phi_{\alpha}^{(0)} | (T + v)^{(2)} | \Phi_{\alpha}^{(0)} \rangle + E_{\text{Hxc},0}^{(2)} , \quad (126)$$

or

$$E^{(2)} = \frac{1}{2} \sum_{\alpha=1}^N \langle \Phi_{\alpha}^{(1)} | (T + v)^{(1)} | \Phi_{\alpha}^{(0)} \rangle + \frac{1}{2} \sum_{\alpha=1}^N \langle \Phi_{\alpha}^{(0)} | (T + v)^{(1)} | \Phi_{\alpha}^{(1)} \rangle + \sum_{\alpha=1}^N \langle \Phi_{\alpha}^{(0)} | (T + v)^{(2)} | \Phi_{\alpha}^{(0)} \rangle + E_{\text{Hxc},0}^{(2)} . \quad (127)$$

If the perturbation does not affect the explicit form of the kinetic energy and Hartree and exchange-correlation energy, these equations reduce to

$$E^{(2)} = \sum_{\alpha=1}^N \langle \Phi_{\alpha}^{(1)} | v^{(1)} | \Phi_{\alpha}^{(0)} \rangle + \langle \Phi_{\alpha}^{(0)} | v^{(2)} | \Phi_{\alpha}^{(0)} \rangle \quad (128)$$

$$= \sum_{\alpha=1}^N \langle \Phi_{\alpha}^{(0)} | v^{(1)} | \Phi_{\alpha}^{(1)} \rangle + \langle \Phi_{\alpha}^{(0)} | v^{(2)} | \Phi_{\alpha}^{(0)} \rangle \quad (129)$$

$$= \sum_{\alpha=1}^N \frac{1}{2} (\langle \Phi_{\alpha}^{(1)} | v^{(1)} | \Phi_{\alpha}^{(0)} \rangle + \langle \Phi_{\alpha}^{(0)} | v^{(1)} | \Phi_{\alpha}^{(1)} \rangle) + \langle \Phi_{\alpha}^{(0)} | v^{(2)} | \Phi_{\alpha}^{(0)} \rangle , \quad (130)$$

while, if the perturbation reduces to a change of local potential, the result is

$$E^{(2)} = \frac{1}{2} \int v^{(1)}(\mathbf{r}) n^{(1)}(\mathbf{r}) d\mathbf{r} + \int v^{(2)}(\mathbf{r}) n^{(0)}(\mathbf{r}) d\mathbf{r} . \quad (131)$$

In this remarkably simple expression, only the change in the density is needed, and the change in wave functions is not needed. Unfortunately, the associated functional is *not* variational.

G. The first-order eigenenergies: The gauge freedom

Now, we come to the connection between the parallel-transport gauge and the diagonal gauge, at first order.

From Eqs. (69) and (120), one gets

$$\varepsilon_{\alpha}^{(1)} = \Lambda_{\alpha\alpha}^{(1)} = \langle \Phi_{\alpha}^{(0)} | H^{(1)} | \Phi_{\alpha}^{(0)} \rangle . \quad (132)$$

From Eqs. (71)–(73), the diagonal part of the unitary matrix vanishes:

$$U_{\beta\beta}^{(1)} = 0 , \quad (133)$$

while the nondiagonal part of the unitary matrix is derived from Eq. (66):

$$U_{\eta\beta}^{(1)*} = -\frac{\Lambda_{\beta\eta}^{(1)}}{\varepsilon_{\beta}^{(0)} - \varepsilon_{\eta}^{(0)}} \text{ for } \eta \neq \beta . \quad (134)$$

Thus the wave functions in the diagonal gauge can be obtained as [35]

$$|\Phi_{d,\alpha}^{(1)}\rangle = |\Phi_{\parallel,\alpha}^{(1)}\rangle - \sum_{\gamma \neq \alpha}^N \frac{\Lambda_{\gamma\alpha}^{(1)}}{\varepsilon_{\gamma}^{(0)} - \varepsilon_{\alpha}^{(0)}} |\Phi_{\parallel,\gamma}^{(1)}\rangle , \quad (135)$$

where the first-order Lagrange multiplier is given by Eq. (120).

In the preceding section the problem of degenerate eigenvalues appearing in the denominator was mentioned. At first order, the diagonal gauge can be defined properly as follows. Because there is an indeterminacy in the choice of the degenerate wave functions at zero order, one can choose the *zero-order* wave functions to diagonalize the *first-order* Lagrange submatrix related to the degenerate valence states. Indeed, from Eq. (120), the first-order Lagrange parameters depend on the zero-order

wave functions and the first-order Hamiltonian, but the latter is invariant under a gauge transformation. So, it is possible to get, for α and β labeling degenerate states,

$$\Lambda_{\alpha\beta}^{(1)} = \varepsilon_{\alpha}^{(1)} \delta_{\alpha\beta} . \quad (136)$$

Then, one can check that the contributions causing the vanishing of the denominator in Eq. (135) can be ignored.

A state-by-state decomposition of the second-derivative of the total energy, in the spirit of Eq. (89), must allow mixing between occupied states due to the perturbation. As a consequence, the results from the parallel-transport gauge would be wrong, and the diagonal gauge must be used, starting from the nonvariational Eqs. (125)–(127), where it is considered that the kinetic operator and the form of the exchange-correlation functional are not affected by the perturbation:

$$E_{\alpha}^{(2)} = \langle \Phi_{\parallel,\alpha}^{(1)} | v^{(1)} | \Phi_{\alpha}^{(0)} \rangle + \langle \Phi_{\alpha}^{(0)} | v^{(2)} | \Phi_{\alpha}^{(0)} \rangle \quad (137)$$

such that

$$E^{(2)} = \sum_{\alpha=1}^N E_{\alpha}^{(2)} . \quad (138)$$

H. Calculation of the second-order wave functions

In order to avoid the direct minimization of Eq. (117), the following modified Sternheimer equations can be used:

$$P_c(H - \varepsilon_{\alpha})^{(0)} P_c | \Phi_{\alpha}^{(2)} \rangle = -P_c(H^{(2)} | \Phi_{\alpha}^{(0)} \rangle + H^{(1)} | \Phi_{\alpha}^{(1)} \rangle) , \quad (139)$$

where the second-order Hamiltonian is given by

$$H^{(2)} = T^{(2)} + v_{\text{KS}}^{(2)} = T^{(2)} + v^{(2)} + v_{\text{Hxc}}^{(2)} , \quad (140)$$

with

$$\begin{aligned} v_{\text{Hxc}}^{(2)}(\mathbf{r}) &= v_{\text{Hxc},0}^{(2)}(\mathbf{r}) + \int K_{\text{Hxc},0}^{(0)}(\mathbf{r}, \mathbf{r}') n^{(2)}(\mathbf{r}') d\mathbf{r}' \\ &+ \int K_{\text{Hxc},0}^{(1)}(\mathbf{r}, \mathbf{r}') n^{(1)}(\mathbf{r}') d\mathbf{r}' \\ &+ \int \frac{\delta^3 E_{\text{Hxc}}[n^{(0)}]}{\delta n(\mathbf{r}) \delta n(\mathbf{r}') \delta n(\mathbf{r}'')} n^{(1)}(\mathbf{r}') \\ &\times n^{(1)}(\mathbf{r}'') d\mathbf{r}' d\mathbf{r}'' . \end{aligned} \quad (141)$$

I. The second-order eigenenergies: The gauge freedom

Now, we come to the connection between the parallel-transport gauge and the diagonal gauge, at second order. From Eqs. (68) and (134), one gets

$$\varepsilon_{\alpha}^{(2)} = \Lambda_{\alpha\alpha}^{(2)} - \sum_{\gamma=1}^N \frac{|\Lambda_{\gamma\alpha}^{(1)}|^2}{\varepsilon_{\gamma}^{(0)} - \varepsilon_{\alpha}^{(0)}} . \quad (142)$$

The second-order Lagrange multipliers are obtained by

$$\begin{aligned} \Lambda_{\beta\alpha}^{(2)} &= \langle \Phi_{\beta}^{(0)} | H^{(2)} | \Phi_{\alpha}^{(0)} \rangle + \langle \Phi_{\beta}^{(1)} | H^{(1)} | \Phi_{\alpha}^{(0)} \rangle \\ &+ \langle \Phi_{\beta}^{(0)} | H^{(1)} | \Phi_{\alpha}^{(1)} \rangle + \langle \Phi_{\beta}^{(1)} | H^{(0)} | \Phi_{\alpha}^{(1)} \rangle \\ &- \frac{1}{2} \langle \Phi_{\beta}^{(1)} | \Phi_{\alpha}^{(1)} \rangle (\varepsilon_{\alpha}^{(0)} + \varepsilon_{\beta}^{(0)}) . \end{aligned} \quad (143)$$

From Eqs. (71) and (73), the diagonal part of the unitary matrix is given by

$$U_{\beta\beta}^{(2)} = -\frac{1}{2} \sum_{\gamma \neq \beta}^N |U_{\gamma\beta}^{(1)}|^2 , \quad (144)$$

while the nondiagonal part of the unitary matrix is derived from Eq. (66):

$$\begin{aligned} U_{\eta\beta}^{(2)*} &= -\frac{1}{\varepsilon_{\beta}^{(0)} - \varepsilon_{\eta}^{(0)}} \left[\left(\sum_{\gamma=1}^N U_{\eta\beta}^{(1)*} \Lambda_{\beta\gamma}^{(1)} \right) + \Lambda_{\beta\eta}^{(2)} \right. \\ &\left. - \varepsilon_{\eta}^{(1)} U_{\eta\beta}^{(1)*} - \varepsilon_{\eta}^{(2)} \delta_{\eta\beta} \right] \text{ for } \eta \neq \beta . \end{aligned} \quad (145)$$

Thus the wave functions in the diagonal gauge are obtained, from Eq. (74), as

$$| \Phi_{d,\alpha}^{(2)} \rangle = | \Phi_{\parallel,\alpha}^{(2)} \rangle + \sum_{\gamma=1}^N U_{\alpha\gamma}^{(1)*} | \Phi_{\parallel,\gamma}^{(1)} \rangle + \sum_{\gamma=1}^N U_{\alpha\gamma}^{(2)*} | \Phi_{\gamma}^{(0)} \rangle . \quad (146)$$

J. An example

In this subsection, an example of calculation of second and third derivatives of the energy is given [40]. Results obtained from finite differences and using the perturbation theory developed in the present paper are compared. The results are shown to be essentially identical.

Two germanium atoms, with reduced coordinates (0 0 0) and (x 0 0) are placed in a repeated cubic supercell of size 10 bohrs. The total energy of this system, per supercell, is monitored as a function of the parameter x (with $0 < x < 1$), governing the distance between the two atoms (and their replicas in other cells). The size of the supercell is such that the system is made of noninteracting linear chains of germanium atoms. The evaluation of the DFT total energy is performed in the local density approximation [3], using either no exchange-correlation energy (for testing purposes), or the Ceperley-Alder exchange-correlation energy [68]. A conjugate-gradient algorithm [44] is used for the minimization of the energy. The ionic potential is replaced by a simple local pseudopotential of the Starkloff-Joannopoulos form [70]. The wave functions are expanded in terms of plane waves of kinetic energy less than 1.2 hartrees. The Brillouin zone is sampled by four points, located in $(\frac{3}{8} 0 0)$, $(\frac{1}{8} 0 0)$, $(-\frac{1}{8} 0 0)$, and $(-\frac{3}{8} 0 0)$ in reduced coordinates. The total energy curve is shown in Fig. 1, obtained with the Ceperley-Alder exchange-correlation energy.

The energy is now expanded as a function of x , around $x = 0.3$. For this, two methods have been followed.

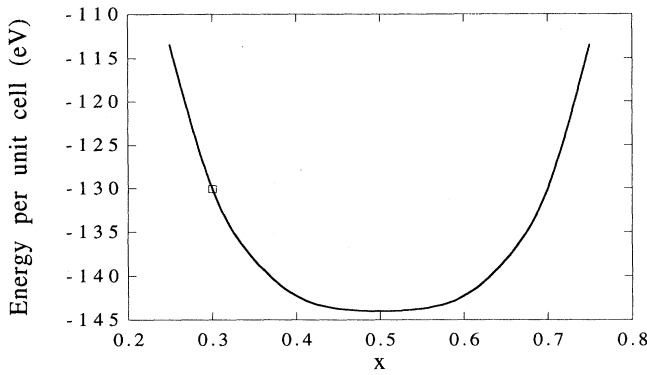


FIG. 1. The energy per unit cell of a model system made of two germanium atoms in a cubic cell of size 10 bohrs, one located at (0 0 0) and the other at (x 0 0) in reduced coordinates, as a function of x . The energy derivatives around 0.3 will be evaluated both by finite-difference techniques and by perturbation theory. More details on this system are given in the text.

First, the total energy is evaluated independently at the points $x = 0.300, 0.301, 0.302, 0.304, 0.299, 0.298, 0.296$. A fit of these values is realized using a sixth-order polynomial. First-, second-, and third-order derivatives of the total energy are extracted by using numerical finite-difference formulas.

In the second method, the second-order expression Eq. (95) is minimized using a similar conjugate-gradient algorithm as for the ground-state calculation [13,70]. The trial first-order wave functions are refined iteratively, until the accuracy level required is attained. The second-order derivative of the energy can be evaluated at each iteration, using the variational expression that is minimized, Eq. (95), or another, nonvariational expression, Eq. (100). The third-order derivative of the energy can also be evaluated at each iteration, since it only uses the first-order wave functions and the first-order Lagrange multipliers [see Eq. (112)]. Figure 2 shows the convergence of the different expressions towards their accurate value. It is clear that the variational expression for the second-order derivative of the total energy performs much better than the corresponding nonvariational

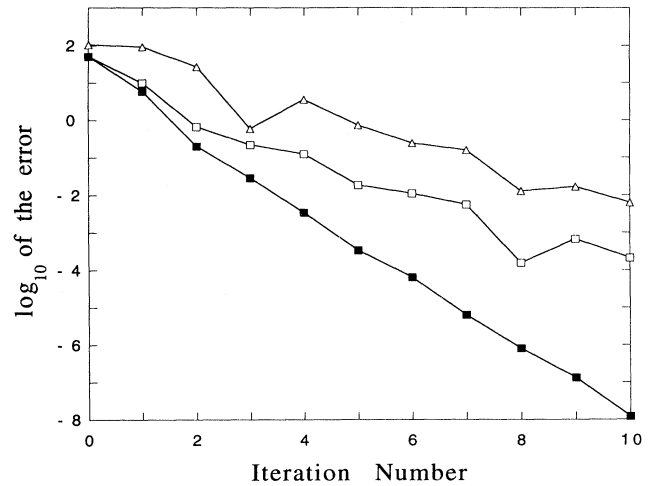


FIG. 2. Decimal logarithm of the absolute difference (in hartrees) between the value obtained after a given number of iterations of the conjugate-gradient algorithm and the converged value. Filled squares: the second derivative of the energy evaluated using the variational expression Eq. (95); empty squares: the second derivative of the energy evaluated using the nonvariational expression Eq. (100); triangles: the third derivative of the energy evaluated using Eq. (112). The energy calculation refers to the model system discussed in the text and for which the energy curve is shown, Fig. 1.

expression. The third-order expression is unfortunately nonvariational as well. The results are shown in Table I and compared with the values obtained previously, by use of perturbation theory.

The results clearly show that the finite-difference and perturbation approaches are in excellent agreement with each other. Without the exchange and correlation contribution, there is no difference, up to the numerical accuracy guaranteed by the finite-difference formulas. When the exchange and correlation energy is taken into account, very small differences appear, tentatively ascribed to the finiteness of the exchange and correlation grid (a programming error cannot be excluded, of course).

TABLE I. Derivatives of the energy per unit cell with respect to x , where the energy refers to the model system presented in Sec. IV J, and x to the coordinate of the second atom in this model. The upper part of the table refers to calculations done without exchange and correlation energy, while the lower part refers to calculation done with the LDA. In each part, the upper line refers to the values obtained using DFPT, while the lower line has been obtained by derivation of a polynomial fitted to a few energies, as explained in the text. All quoted figures are significant.

Calculation	dE/dx	d^2E/dx^2	d^3E/dx^3
Without xc energy			
DFPT	-8.559 221 877	118.029 86	-1346.67
Finite differences	-8.559 221 877	118.029 86	-1346.67
With xc energy			
DFPT	-7.375 153 878	116.534 58	-1377.19
Finite differences	-7.375 153 820	116.534 60	-1376.94

V. SUMMARY AND PERSPECTIVES

In this paper, I have examined rather thoroughly the schemes that allow one to calculate the different derivatives of the DFT electronic energy with respect to arbitrary perturbations.

From the wave functions up to order n , it is possible to construct the derivatives of the energy up to order $2n + 1$. The search for wave functions of order n can be done either by minimization of a functional of order $2n$, under constraints, or by resolution of modified Sternheimer equations, by Green's function techniques, or by sum over states. The invariance of the energy expression with respect to the unitary transformation of the wave functions brings specific problems that are solved by fixing the gauge. The parallel-transport gauge is the most convenient for numerical applications. On the other hand, the diagonal gauge is sometimes needed, for derivatives of the eigenenergies. The formulas needed for making the gauge transformation have been derived.

Equations derived from this general formalism have been considered up to the fourth derivative of the energy, and up to second-order derivative of the wave function, density, and eigenenergy. A simple application has shown the validity of this approach, when compared to finite-difference estimation of energy derivatives.

Generalizations of the present theory have to be worked out in the case of time-dependent perturbations (see Ref. [34] for a second-order example), as well as in the case of mixed derivatives of the energy (see Ref. [39] for a third-order example).

ACKNOWLEDGMENTS

I am grateful to J. -C. Charlier, Ph. Ghosez, and G.-M. Rignanese for useful discussions, as well as a care-

ful reading of the manuscript and the lecture notes (UCL-PHYS 3915) on which this paper was based. I thank D. Vanderbilt for an important remark on the manuscript, and D. C. Allan and C. Lee for their interest in this work. I acknowledge financial support from the National Fund for Scientific Research (Belgium).

APPENDIX: LAGRANGE MULTIPLIER METHOD

The minimization of $\hat{E}_{(\lambda)}[\Phi_{\text{trial}}]$ should be performed under a set of N constraints,

$$\hat{C}_{\alpha,(\lambda)}[\Phi_{\text{trial}}] = 0 \quad \text{for } \alpha = 1, \dots, N \quad (\text{A1})$$

so that

$$E_0(\lambda) = \min_{\hat{C}_{\alpha,(\lambda)}[\Phi_{\text{trial}}]=0} \{ \hat{E}_{(\lambda)}[\Phi_{\text{trial}}] \} . \quad (\text{A2})$$

This procedure, in the Lagrange multiplier method, is replaced by the minimization without constraint of an auxiliary functional $\hat{F}_{\Lambda_\alpha}[\Phi_{\text{trial}}]$ (the λ dependence has been omitted, for reasons of clarity):

$$\hat{F}_{\Lambda_\alpha}[\Phi_{\text{trial}}] = \hat{E}[\Phi_{\text{trial}}] - \sum_{\alpha=1}^N \Lambda_\alpha \hat{C}_\alpha[\Phi_{\text{trial}}] , \quad (\text{A3})$$

for all possible values of the Λ_α , called the Lagrange multipliers. Afterwards, the "correct" value of $\Lambda_{\alpha,0}$ is such that the corresponding Φ_0 fulfill

$$\hat{C}_\alpha[\Phi_0] = 0 . \quad (\text{A4})$$

-
- [1] P. Hohenberg and W. Kohn, Phys. Rev. B **136**, 864 (1964).
- [2] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).
- [3] See the review by R. O. Jones and O. Gunnarsson, Rev. Mod. Phys. **61**, 689 (1989).
- [4] I. Zh. Petkov and M. V. Stoitsov, *Nuclear Density Functional Theory* (Clarendon, Oxford, 1991).
- [5] T. A. Arais and J. D. Joannopoulos, Phys. Rev. Lett. **73**, 680 (1994).
- [6] M. J. Stott and E. Zaremba, Phys. Rev. A **21**, 12 (1980).
- [7] A. Zangwill and P. Soven, Phys. Rev. A **21**, 1561 (1980).
- [8] G. D. Mahan, Phys. Rev. A **22**, 1780 (1980).
- [9] S. K. Ghosh and B. M. Deb, Chem. Phys. **71**, 295 (1982).
- [10] N. E. Zein, Fiz. Tverd. Tela (Leningrad) **26**, 3028 (1984) [Sov. Phys. Solid State **26**, 1825 (1984)].
- [11] S. Baroni, P. Giannozzi, and A. Testa, Phys. Rev. Lett. **58**, 1861 (1987).
- [12] R. Fournier, J. Chem. Phys. **92**, 5422 (1990).
- [13] X. Gonze, D. C. Allan, and M. P. Teter, Phys. Rev. Lett. **68**, 3603 (1992).
- [14] G. D. Mahan, *Local Density Theory of Polarisability* (Plenum, New York, 1990).
- [15] Following Refs. [13,20–25], the application of these results to DFT is referred to as "density-functional perturbation theory" (DFPT). The same framework, in the context of *linear* responses to *time-dependent* electric fields, is usually referred to as "time-dependent DFT" [14,16,19] in the atomic physics community. A "time-dependent local density approximation" is subsequently performed. In the present paper, the emphasis is not laid on the time-dependent characteristics, but on high-order expansion, so that the term density-functional perturbation theory is more appropriate.
- [16] L. J. Bartolotti, J. Chem. Phys. **80**, 5687 (1984); J. Phys. Chem. **90**, 5518 (1986); L. J. Bartolotti and Q. Xie, Theor. Chim. Acta **77**, 239 (1990); K. D. Sen and L. J. Bartolotti, Phys. Rev. A **45**, 2076 (1992).
- [17] S. Baroni, P. Giannozzi, and A. Testa, Phys. Rev. Lett. **59**, 2662 (1987).
- [18] Z. H. Levine and D. C. Allan, Phys. Rev. Lett. **63**, 1719 (1989).
- [19] A. Zangwill, *Atomic Physics 8*, edited by I. Lindgren, A. Rosen, and S. Svanberg (Plenum, New York, 1983), p. 339; Z. H. Levine and P. Soven, Phys. Rev. Lett. **50**,

- 2074 (1983); A. Zangwill and D. A. Liberman, *J. Phys. B* **17**, L253 (1984); Z. H. Levine and P. Soven, *Phys. Rev. A* **29**, 625 (1984); Z. Crljen, J. Luberek, G. Wendin, and Z. H. Levine, *ibid.* **50**, 3529 (1994).
- [20] M. Buongiorno Nardelli, S. Baroni, and P. Giannozzi, *Phys. Rev. Lett.* **60**, 1959 (1992); S. de Gironcoli and S. Baroni, *ibid.* **69**, 1069 (1992); A. Dal Corso, S. Baroni, R. Resta, and S. de Gironcoli, *Phys. Rev. B* **47**, 3588 (1993); S. Guha *et al.*, *Phys. Rev. Lett.* **72**, 3359 (1994); S. de Gironcoli, *Phys. Rev. B* **46**, 2412 (1992); P. Giannozzi and S. Baroni, *J. Chem. Phys.* **100**, 8537 (1994); N. Marzari, S. de Gironcoli, and S. Baroni, *Phys. Rev. Lett.* **72**, 4001 (1994); M. Peressi and S. Baroni, *Phys. Rev. B* **49**, 7490 (1994).
- [21] P. Giannozzi, S. de Gironcoli, P. Pavone, and S. Baroni, *Phys. Rev. B* **43**, 7231 (1990).
- [22] S. de Gironcoli, S. Baroni, and R. Resta, *Phys. Rev. Lett.* **24**, 2853 (1989).
- [23] S. de Gironcoli, P. Giannozzi, and S. Baroni, *Phys. Rev. Lett.* **66**, 2116 (1991).
- [24] X. Gonze, D. C. Allan, and M. P. Teter, in *Phonon Scattering in Condensed Matter VII*, edited by M. Meissner and R. O. Pohl (Springer-Verlag, Berlin, 1993), p. 511; Ph. Ghosez, X. Gonze, and J.-P. Michenaud, *Ferroelectrics* **153**, 91 (1994); **164**, 113 (1995); C. Lee and X. Gonze, *Phys. Rev. Lett.* **72**, 1686 (1994); *Phys. Rev. B* **49**, 14730 (1994); *J. Phys. Condens. Matter* **7**, 3693 (1995); X. Gonze, J.-C. Charlier, D. C. Allan, and M. P. Teter, *ibid.* **50**, 13035 (1994); C. Lee, Ph. Ghosez, and X. Gonze, *ibid.* **50**, 13379 (1994).
- [25] P. Pavone *et al.*, *Phys. Rev. B* **48**, 3156 (1993); W. Windl *et al.*, *ibid.* **48**, 3164 (1993); W. Windl *et al.*, *ibid.* **49**, 8764 (1994); O. Schütt *et al.*, *ibid.* **50**, 3746 (1994).
- [26] A. Dal Corso, S. Baroni, and R. Resta, *Phys. Rev. B* **49**, 5323 (1994).
- [27] A. A. Quong and M. R. Pederson, *Phys. Rev. B* **46**, 12906 (1992); A. A. Quong and A. G. Eguluz, *Phys. Rev. Lett.* **70**, 3955 (1993).
- [28] J. Fritsch, P. Pavone, and U. Schröder, *Phys. Rev. Lett.* **71**, 4194 (1993); F. Ancilotto *et al.*, *Phys. Rev. B* **43**, 8930 (1991); S. Baroni, S. de Gironcoli, and P. Giannozzi, *Phys. Rev. Lett.* **65**, 84 (1990); S. Baroni, P. Giannozzi, and E. Molinari, *Phys. Rev. B* **41**, 3870 (1990); A. Fleszar and X. Gonze, *Phys. Rev. Lett.* **64**, 2961 (1990).
- [29] (a) D. Kh. Blat, N. E. Zein, and V. I. Zinenko, *J. Phys. Condens. Matter* **3**, 5515 (1991); (b) N. E. Zein, *Phys. Lett. A* **161**, 526 (1992).
- [30] Z. H. Levine, *Phys. Rev. B* **42**, 3567 (1990); Z. H. Levine and D. C. Allan, *Phys. Rev. Lett.* **66**, 41 (1991); *Phys. Rev. B* **43**, 4187 (1991); **44**, 12781 (1991); H. Zhong, Z. H. Levine, D. C. Allan, and J. W. Wilkins, *Phys. Rev. Lett.* **69**, 379 (1992); *Phys. Rev. B* **48**, 1384 (1993); Z. H. Levine and D. C. Allan, *ibid.* **48**, 7783 (1993); Z. H. Levine, *ibid.* **49**, 4532 (1994).
- [31] B. I. Dunlap and J. Andzelm, *Phys. Rev. A* **45**, 81 (1992); N. C. Handy, D. J. Tozer, C. W. Murray, G. J. Larning, and R. D. Amos, *Isr. J. Chem.* **33**, 331 (1993).
- [32] A. Komornicki and G. Fitzgerald, *J. Chem. Phys.* **98**, 1398 (1993).
- [33] B. G. Johnson and M. J. Fisch, *J. Chem. Phys.* **100**, 7429 (1994).
- [34] A. Pasquarello and A. Quattropani, *Phys. Rev. B* **48**, 5090 (1993).
- [35] S. Yu Savrasov, *Phys. Rev. Lett.* **69**, 2819 (1992); S. Yu Savrasov, D. Yu Savrasov, and O. K. Andersen, *ibid.* **72**, 372 (1994).
- [36] S. Yu Savrasov, *Solid State Commun.* **74**, 69 (1990).
- [37] Ricci Yu and H. Krakauer, *Phys. Rev. B* **49**, 4467 (1994); C.-Z. Wang, Rici Yu, and H. Krakauer, *Phys. Rev. Lett.* **72**, 368 (1994).
- [38] J. O. Hirschfelder, W. Byers Brown, and S. T. Epstein, *Advances in Quantum Chemistry* (Academic, New York, 1964), Vol. 1, pp. 267, 288, and 300.
- [39] X. Gonze and J.-P. Vigneron, *Phys. Rev. B* **39**, 13120 (1989).
- [40] X. Gonze (unpublished).
- [41] A. Debernardi and S. Baroni, *Solid State Commun.* **91**, 813 (1994).
- [42] A. Dal Corso and F. Mauri, *Phys. Rev. B* **50**, 5756 (1994).
- [43] A. Debernardi, S. Baroni, and E. Molinari (unpublished).
- [44] M. C. Payne, M. P. Teter, D. C. Allan, T. A. Arias, and J. D. Joannopoulos, *Rev. Mod. Phys.* **64**, 1045 (1992).
- [45] O. Sinanoglu, *J. Chem. Phys.* **34**, 1237 (1961).
- [46] E. A. Hylleraas, *Z. Phys.* **65**, 209 (1930).
- [47] X. Gonze, preceding paper, *Phys. Rev. A* **52**, 1086 (1995).
- [48] R. M. Sternheimer, *Phys. Rev.* **96**, 951 (1954); see also R. M. Sternheimer, *ibid.* **84**, 244 (1951); and R. M. Sternheimer and H. M. Foley, *ibid.* **92**, 1460 (1953).
- [49] See the introductory section of Ref. [29(b)].
- [50] P. W. Langhoff, S. T. Epstein, and M. Karplus, *Rev. Mod. Phys.* **44**, 602 (1972).
- [51] Although see R. D. King-Smith, R. J. Needs, V. Heine, and M. J. Hodgson, *Europhys. Lett.* **10**, 569 (1989); R. D. King-Smith and R. J. Needs, *J. Phys. Condens. Matter* **2**, 3431 (1990).
- [52] See, for example, J. L. Brédas, *Science* **263**, 487 (1994).
- [53] See, for example, J. M. L. Martin, P. R. Taylor, and T. J. Lee, *Chem. Phys. Lett.* **205**, 535 (1993).
- [54] See, for example, K. M. Rabe and J. D. Joannopoulos, *Phys. Rev. Lett.* **59**, 570 (1987).
- [55] H. Hellmann, *Einführung in die Quantenchemie* (Deuticke, Leipzig, 1937), pp. 61 and 285; R. P. Feynman, *Phys. Rev.* **56**, 340 (1939).
- [56] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes. The Art of Scientific Computing* (Cambridge University Press, Cambridge, England, 1986).
- [57] Note that perturbative X_α calculations, with a mathematical structure similar to DFT, were performed by Ahlberg and Goscinski previously: R. Ahlberg and O. Goscinski, *J. Phys. B* **6**, 2254 (1973); **8**, 2149 (1975).
- [58] A. C. Hurley, *Proc. R. Soc. London, Ser. A* **226**, 179 (1954); S. T. Epstein, *Theor. Chim. Acta* **55**, 251 (1980).
- [59] P. Pulay, in *Modern Theoretical Chemistry*, edited by H. F. Schaefer (Plenum, New York, 1977), Vol. 4, Chap. 4, p. 153.
- [60] S. T. Epstein, *Chem. Phys. Lett.* **70**, 311 (1980).
- [61] The classical ion-ion Coulomb energy (whose derivatives are usually straightforward, if not easy, to calculate) is not treated here.
- [62] Recently, unconstrained functionals have been proposed that at their minimum are equal to the usual Kohn-Sham functional, see F. Mauri, G. Galli, and R. Car, *Phys. Rev. B* **50**, 4316 (1994). These functionals could also be used to derive a density-functional perturbation theory.
- [63] S. G. Louie, S. Froyen, and M. L. Cohen, *Phys. Rev. B*

- 26**, 1738 (1982).
- [64] A. Dal Corso and R. Resta, *Phys. Rev. B* **50**, 4327 (1994).
- [65] C. Filippi, C. J. Umrigar, and M. Taut, *J. Chem. Phys.* **100**, 1290 (1994).
- [66] The complete fourth-order expression is available from the author.
- [67] Actually, the kernel of the quadratic part of all even-order energy expressions is the same.
- [68] D. M. Ceperley and B. J. Alder, *Phys. Rev. Lett.* **45**, 566 (1980), parametrized by M. P. Teter (unpublished).
- [69] Th. Starkloff and J. D. Joannopoulos, *Phys. Rev. B* **16**, 5212 (1977).
- [70] X. Gonze (unpublished).