

Gaussian basis density functional theory for systems periodic in two or three dimensions: Energy and forces

John E. Jaffe and Anthony C. Hess

Citation: *J. Chem. Phys.* **105**, 10983 (1996); doi: 10.1063/1.472866

View online: <http://dx.doi.org/10.1063/1.472866>

View Table of Contents: <http://jcp.aip.org/resource/1/JCPSA6/v105/i24>

Published by the [American Institute of Physics](#).

Additional information on J. Chem. Phys.

Journal Homepage: <http://jcp.aip.org/>

Journal Information: http://jcp.aip.org/about/about_the_journal

Top downloads: http://jcp.aip.org/features/most_downloaded

Information for Authors: <http://jcp.aip.org/authors>

ADVERTISEMENT



**ACCELERATE COMPUTATIONAL CHEMISTRY BY 5X.
TRY IT ON A FREE, REMOTELY-HOSTED CLUSTER.**

[LEARN MORE](#)

Gaussian basis density functional theory for systems periodic in two or three dimensions: Energy and forces

John E. Jaffe and Anthony C. Hess

Environmental Molecular Sciences Laboratory, Pacific Northwest Laboratory, Richland, Washington 99352

(Received 9 October 1995; accepted 20 September 1996)

We describe a formulation of electronic density functional theory using localized Gaussian basis functions for systems periodic in three dimensions (bulk crystals) or two dimensions (crystal slabs terminated by surfaces). Our approach generalizes many features of molecular density functional methods to periodic systems, including the use of an auxiliary Gaussian basis set to represent the charge density, and analytic gradients with respect to nuclear coordinates. Existing quantum chemistry routines for analytic and numerical integration over basis functions can be adapted to our scheme with only slight modifications, as can existing extended Gaussian basis sets. Such basis sets permit accurate calculations with far fewer basis functions (and hence much smaller matrices to diagonalize) than plane-wave based methods, especially in surface calculations, where in our approach the slab does not have to repeat periodically normal to the surface. Realistic treatment of molecule-surface interactions is facilitated since both molecule and surface can be treated at the same level of theory. Our real-space method also offers opportunities to exploit matrix sparsity, since in a large unit cell many pairs of basis functions will be essentially nonoverlapping and noninteracting. Longer-ranged Coulomb interactions are summed by a form of the Ewald technique that guarantees absolute convergence. We also give a straightforward extension to periodic systems (both two- and three-dimensional) of the usual molecular formalism for analytic nuclear first derivatives (forces). © 1996 American Institute of Physics. [S0021-9606(96)02148-4]

I. INTRODUCTION

Density functional theory¹ (DFT) has for many years been the dominant approach to predicting the electronic structure of solids,² and has recently become a major technique in quantum chemistry,³ especially for large molecules and clusters. Although the same fundamental theory¹ underlies the density functional approach in both solid state physics and quantum chemistry, and the same local⁴ or nonlocal⁵ functionals can be used in both disciplines, the detailed implementation is usually quite different. Since the electronic charge density of an isolated molecule is necessarily localized in a finite region of space, the usual quantum chemical method is to expand the electron wave functions in a basis of localized functions, usually centered on the atomic nuclei and often constructed of linear combinations ("contractions") of Gaussian functions. In contrast, the tradition in solid state physics is to use plane waves as a basis set for expanding the wave functions. These can be used in their pure form in the first-principles pseudopotential method,⁶ or modified near the atomic cores (augmented plane waves,⁷ etc.) in all-electron calculations, or supplemented by a limited number of localized functions in mixed-basis methods.⁸ The use of plane waves began in early studies of simple metals where plane waves are a reasonable first approximation to conduction band eigenfunctions, and continues today since plane-wave basis sets permit many formal simplifications and computational economies. Nevertheless, plane-wave methods have the disadvantage that very localized or inhomogeneous systems may require excessive numbers of waves for their representation. This problem can be especially serious in surface calculations, where the requirement

of periodicity in three dimensions leads to a model with slabs repeating periodically along the surface normal. If the space between the slabs is wide enough to make interactions between them negligible, the "lattice constant" normal to the surface has to be quite large and the corresponding reciprocal lattice vector quite short, resulting in a large number of plane waves within a given kinetic energy cutoff and hence a relatively expensive calculation. The distance between the slabs needs to be even larger if adsorbate molecules are to be added to the surface, since these molecules also should not interact across the space between the slabs.

Localized-basis approaches to periodic systems do not suffer from this disadvantage. Very localized states, including core states if desired, can be represented by a suitable choice of atom-centered basis functions. The electronic charge density goes naturally to zero in regions of space where there are no atoms, and the lack of basis functions in such regions is not a problem unless a detailed description of energetic excitations, scattering states, or tunneling is required. For surface problems there is no need for periodicity normal to the surface; a single slab of (sufficient) finite thickness can be used, since these methods can be formulated equally easily for systems periodic in two or three dimensions. This facilitates the accurate treatment of molecule-surface interactions, especially when the molecules are relatively far from the surface. Also, if the basis functions⁹ are carefully constructed and optimized for the crystalline environment,¹⁰ it is possible to represent the occupied electron eigenstates with a relatively small number of basis functions: just enough to accommodate all the electrons, plus a few more functions to give the Kohn-Sham orbitals suffi-

cient variational freedom. Thus, only a few tens of basis functions per atom are typically needed, versus hundreds per atom in typical plane-wave approaches. (The auxiliary basis sets⁹ needed to represent the charge density and, if desired, the exchange-correlation potential are also fairly modest in size.) Once the basis set has been constructed and the needed overlap and Hamiltonian matrices formed, the small (compared to plane wave) size of the basis can greatly reduce the computational cost of solving the Kohn–Sham equations. Also, especially in systems with large unit cells, we can exploit sparsity in these matrices, since basis functions on distant centers will have negligible overlap and interaction (except for certain long-range Coulomb multipole interactions which can be summed according to Ewald’s convention, see Sec. III below). Of course, a drawback of the local basis approach¹¹ is that very large numbers of multicenter integrals may sometimes have to be computed. Advanced techniques borrowed from computational chemistry can be applied to mitigate this task, but the local basis approach may still prove impractical for some cases such as alkali metals or close-packed materials at high pressure.

Below we present our formalism for implementing DFT for periodic systems using a basis set of contracted Gaussian-type orbitals. Our presentation is motivated by the fact that while several related implementations exist, their formal structure has not to our knowledge been described in the literature in explicit detail. Also, our approach differs in several details from the others (see Sec. VIII below). Our discussion should be of interest both to scientists who are actually developing local-basis periodic density functional codes and to those who merely wish to use such codes and understand their formal basis, including their connection with molecular methods. In Sec. II we will review the basic formulas of molecular DFT and extend them to periodic density functional theory, including a periodic generalization of the schemes used in the molecular theory to represent the charge density and (possibly) the exchange-correlation potential with auxiliary Gaussian basis sets. Thus, we define a unified approach to both finite and extended systems. Section III specifies in some detail the multicenter integrals needed in our theory and describes our use of the Ewald potential function to obtain absolutely convergent lattice sums over Coulomb interactions. Section IV discusses the straightforward modifications of the three-dimensionally periodic theory, including the treatment of Coulomb interactions, that are needed to produce a correct formulation for systems periodic in two dimensions, i.e., slabs. In Sec. V we will present a periodic generalization of techniques for calculating forces on nuclei, including corrections for basis set incompleteness,¹² while Sec. VI briefly discusses some issues of computational cost and efficiency. The approach described in Secs. II–VI has recently been implemented in the program GAPSS (for Gaussian Approach to Polymers, Surfaces, and Solids) and Sec. VII gives some numerical results from this implementation, showing close agreement with other (mostly plane wave) implementations of DFT. Finally, Sec. VIII contains a comparison of our method to some other local-basis periodic formulations, and a summary.

II. MOLECULAR AND PERIODIC DENSITY FUNCTIONAL THEORY

A. Molecular theory

To establish our notation let us first review the standard formulation of density functional theory for a finite system such as a molecule. The electron orbital basis functions, which are contractions (linear combinations) of Gaussians, are written as $\chi_i(\mathbf{r})$ where $i=(1,\dots,N)$ runs over all the orbital basis functions in the molecule or cluster. Alternatively we may think of i as a compound subscript containing an index α labeling all the nuclear centers, and a set of “atomic” quantum numbers for each atom. (Spin labels are not shown since we give the formulation for closed shell systems only, but the changes needed for more general cases are straightforward; also the basis functions are assumed to be real, but they could be made complex with only trivial changes in the formulas.) We expand the eigenfunctions (molecular orbitals) as

$$\psi_n(\mathbf{r}) = \sum_i c_{ni} \chi_i(\mathbf{r}). \quad (1)$$

The Kohn–Sham equation, after substitution, is

$$\sum_j H_{ij} c_{nj} = \epsilon_n \sum_j S_{ij} c_{nj}, \quad (2)$$

where

$$H_{ij} = \int \chi_i(\mathbf{r}) \mathbf{H} \chi_j(\mathbf{r}) d^3r, \quad S_{ij} = \int \chi_i(\mathbf{r}) \chi_j(\mathbf{r}) d^3r. \quad (3)$$

Solution for the c ’s yields the electron density in terms of the density matrix P_{ij} :

$$\begin{aligned} \rho(\mathbf{r}) &= \sum_n^{\text{occ}} \psi_n^*(\mathbf{r}) \psi_n(\mathbf{r}) \\ &= \sum_{ij} \chi_i(\mathbf{r}) \chi_j(\mathbf{r}) \sum_n^{\text{occ}} c_{ni}^* c_{nj} \equiv \sum_{ij} \chi_i(\mathbf{r}) \chi_j(\mathbf{r}) P_{ij}, \end{aligned} \quad (4)$$

where “occ” limits the sum to occupied eigenstates, and the sum over n implicitly includes spins. We use this density to construct the next iteration of H_{ij} :

$$\begin{aligned} H_{ij} &= -\frac{1}{2} \int \chi_i(\mathbf{r}) \nabla^2 \chi_j(\mathbf{r}) d^3r \\ &+ \int \chi_i(\mathbf{r}) \sum_{\alpha} \frac{-Z_{\alpha}}{|\mathbf{r} - \mathbf{r}_{\alpha}|} \chi_j(\mathbf{r}) d^3r \\ &+ \int \chi_i(\mathbf{r}_1) \chi_j(\mathbf{r}_1) \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} \rho(\mathbf{r}_2) d^3r_1 d^3r_2 \\ &+ \int \chi_i(\mathbf{r}) \chi_j(\mathbf{r}) V_{\text{xc}}[\rho(\mathbf{r})] d^3r, \end{aligned} \quad (5)$$

where \mathbf{r}_{α} is nuclear coordinate and Z_{α} a nuclear charge. For basis functions constructed from Gaussians the first three integrals can be done analytically, but the exchange-correlation

potential functional $V_{\text{XC}}[\rho]$ is nonlinear and perhaps nonlocal⁵ so the fourth term requires numerical integration. This is done on a grid of $G(\sim 10^3$ per atom) mesh points \mathbf{r}_m with weights w_m

$$\int \chi_i \chi_j V_{\text{XC}}[\rho] d^3r \approx \sum_m \chi_i(\mathbf{r}_m) \chi_j(\mathbf{r}_m) V_{\text{XC}}[\rho(\mathbf{r}_m)] w_m. \quad (6)$$

Various techniques can reduce the computational cost of the Coulomb and XC terms in Eq. (5). The Coulomb term is expensive if the density is represented as in Eq. (4) due to the need to perform on the order of N^4 four-center integrals. This cost can be reduced to order N^3 by fitting the density¹³ to an auxiliary basis of $M \approx 2N$ functions g such that

$$\rho(\mathbf{r}) \approx \tilde{\rho}(\mathbf{r}) = \sum_p a_p g_p(\mathbf{r}). \quad (7)$$

In our notation the fitting coefficients are given by

$$a_p = \sum_{p'} (V^{-1})_{pp'} (t_{p'} + \lambda n_{p'}), \quad (8)$$

where

$$\begin{aligned} V_{pp'} &= \int g_p(1) \frac{1}{r_{12}} g_{p'}(2) d^3r_1 d^3r_2, \\ t_p &= \int \rho(1) \frac{1}{r_{12}} g_p(2) d^3r_1 d^3r_2, \\ \lambda &= \frac{Q - \sum_{pp'} n_{pp'} (V^{-1})_{pp'} t_{p'}}{\sum_{pp'} n_{pp'} (V^{-1})_{pp'} n_{p'}}, \\ n_p &= \int g_p(\mathbf{r}) d^3r, \end{aligned} \quad (9)$$

and Q is the total number of electrons. Note that the ‘‘overlap’’ and projection integrals now have the form of Coulomb interactions (with $V_{pp}=1$ a convenient normalization) and λ is a Lagrange multiplier that enforces charge conservation $Q = \sum_p a_p n_p$. This scheme variationally minimizes the fitting error in the Coulomb part of the total energy when this energy is taken to be

$$\begin{aligned} E_{\text{coul}} &= \int \rho(\mathbf{r}_1) \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} \tilde{\rho}(\mathbf{r}_2) d^3r_1 d^3r_2 \\ &\quad - \frac{1}{2} \int \tilde{\rho}(\mathbf{r}_1) \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} \tilde{\rho}(\mathbf{r}_2) d^3r_1 d^3r_2. \end{aligned} \quad (10)$$

If the auxiliary basis is made up of Gaussians then all the integrals in the density fitting scheme can be done analytically.

The XC matrix elements in Eq. (6) are also costly since the expression must be evaluated for each basis function i, j leading to $N^2 G$ evaluations of the integrand per SCF cycle. A possible reduction in cost¹³ can be achieved by representing V_{XC} with a basis of $M \sim 2N$ auxiliary functions G_q

$$V_{\text{XC}}(\mathbf{r}) \approx \tilde{V}_{\text{XC}}(\mathbf{r}) = \sum_q b_q G_q(\mathbf{r}). \quad (11)$$

The fitting coefficients are then found from

$$\begin{aligned} b_q &= \sum_{q'} (S_{\text{XC}}^{-1})_{qq'} \int V_{\text{XC}}[\rho(\mathbf{r})] G_{q'}(\mathbf{r}) d^3r \\ &\approx \sum_{q'} (S_{\text{XC}}^{-1})_{qq'} \sum_m V_{\text{XC}}[\rho(\mathbf{r}_m)] G_{q'}(\mathbf{r}_m) w_m, \end{aligned} \quad (12)$$

where the overlap matrix in the exchange-correlation (XC) fitting basis is

$$S_{qq'}^{\text{XC}} = \int G_q(\mathbf{r}) G_{q'}(\mathbf{r}) d^3r. \quad (13)$$

This auxiliary basis overlap matrix is evaluated and inverted only once for each nuclear geometry, and the fitting coefficients are evaluated only once per SCF cycle. After that the H_{ij} integrals can be done analytically, leading to a cost on the order of $2N^2 M$ operations for the XC terms each cycle. However, when we adopt an effective scheme for deleting i, j pairs that are essentially nonoverlapping there is little advantage in fitting the XC potential; see Sec. VI below.

B. Periodic theory

Instead of a cluster or molecule let us now consider a macroscopic crystal (effectively infinite in three dimensions) composed of a primitive unit cell which reproduces the entire periodic crystal¹⁴ by repeated translations through lattice vectors \mathbf{R} . Basis functions in the home cell ($\mathbf{R}=0$) are still designated by $\chi_i(\mathbf{r})$ where now the index i indicates a particular basis function on a particular atom in the unit cell. The corresponding function in a cell displaced from the home cell by \mathbf{R} is $\chi_i(\mathbf{r}-\mathbf{R})$. We define Bloch functions

$$\phi_{\mathbf{k}i}(\mathbf{r}) = \sum_{\mathbf{R}} e^{i\mathbf{k} \cdot \mathbf{R}} \chi_i(\mathbf{r}-\mathbf{R}), \quad (14)$$

which are the basis in which we expand the eigenfunctions (crystal orbitals)

$$\psi_{\mathbf{k}n}(\mathbf{r}) = \sum_i c_{\mathbf{k}ni} \phi_{\mathbf{k}i}(\mathbf{r}), \quad (15)$$

where this n is a band index, and \mathbf{k} is a wave vector in the first Brillouin zone which is continuous in principle,¹⁵ but which can be taken in practice to be, for example, a Monkhorst–Pack¹⁶ sampling point. The Kohn–Sham eigenvalue or band structure equation is now

$$\mathbf{H} \psi_{\mathbf{k}n}(\mathbf{r}) = \epsilon_{\mathbf{k}n} \psi_{\mathbf{k}n}(\mathbf{r}). \quad (16)$$

After substitution we have

$$\sum_j H_{ij}(\mathbf{k}) c_{\mathbf{k}nj} = \epsilon_{\mathbf{k}n} \sum_j S_{ij}(\mathbf{k}) c_{\mathbf{k}nj}, \quad (17)$$

where

$$H_{ij}(\mathbf{k}) = \sum_{\mathbf{R}} e^{i\mathbf{k}\cdot\mathbf{R}} \int d^3r \chi_i(\mathbf{r}) \mathbf{H} \chi_j(\mathbf{r}-\mathbf{R}), \quad (18)$$

$$S_{ij}(\mathbf{k}) = \sum_{\mathbf{R}} e^{i\mathbf{k}\cdot\mathbf{R}} \int d^3r \chi_i(\mathbf{r}) \chi_j(\mathbf{r}-\mathbf{R}).$$

The solution for the c coefficients determines the charge density for the next iteration according to

$$\rho(\mathbf{r}) = \sum_{\mathbf{k}} w_{\mathbf{k}} \sum_n^{\text{occ}} \psi_{\mathbf{k}n}^*(\mathbf{r}) \psi_{\mathbf{k}n}(\mathbf{r}), \quad (19)$$

where the \mathbf{k} points are again Monkhorst–Pack or similar special points¹⁶ and the w 's are the associated weights ($\sum_{\mathbf{k}} w_{\mathbf{k}} = 1$ in this formulation). (For metals, the sum on \mathbf{k} must also be limited to occupied states, i.e., those within the Fermi surface, for those bands which intersect the Fermi energy.) The density can then be expressed as

$$\rho(\mathbf{r}) = \sum_{ij\mathbf{R}\mathbf{R}'} P_{ij}(\mathbf{R}) \chi_i(\mathbf{r}-\mathbf{R}') \chi_j(\mathbf{r}-\mathbf{R}-\mathbf{R}'), \quad (20)$$

where the density matrix is now

$$P_{ij}(\mathbf{R}) = \sum_{\mathbf{k}} w_{\mathbf{k}} \sum_n^{\text{occ}} e^{i\mathbf{k}\cdot\mathbf{R}} c_{\mathbf{k}ni}^* c_{\mathbf{k}nj}. \quad (21)$$

We can then build the Hamiltonian for the next cycle according to the prescriptions given above.

Next we must determine how to fit the charge density and (possibly) the XC potential in the periodic case. We can construct fitted approximations which are manifestly periodic by summing localized Gaussian fitting functions g, G of the same type as before but centered in all the unit cells of the crystal:

$$\tilde{\rho}(\mathbf{r}) = \sum_{\mathbf{R}} \sum_p a_p g_p(\mathbf{r}-\mathbf{R}),$$

$$\tilde{V}_{\text{XC}}(\mathbf{r}) = \sum_{\mathbf{R}} \sum_q b_q G_q(\mathbf{r}-\mathbf{R}). \quad (22)$$

For the XC potential the fitting coefficients can be found easily from

$$b_q = \sum_{q'} (S_{\text{XC}}^{-1})_{qq'} \int V_{\text{XC}}(\mathbf{r}) G_{q'}(\mathbf{r}) d^3r, \quad (23)$$

where

$$S_{qq'}^{\text{XC}} = \sum_{\mathbf{R}} \int G_q(\mathbf{r}) G_{q'}(\mathbf{r}-\mathbf{R}) d^3r. \quad (24)$$

The integral in Eq. (23) must be done by numerical quadrature, while Eq. (24) can be integrated analytically. The range of both integrals is over all space, but since their integrands are periodic functions times a localized Gaussian, only a finite region of space actually contributes to them. In contrast, the long-ranged nature of the Coulomb interaction causes convergence problems if we attempt a naïve periodic generalization of Dunlap's¹³ charge density fitting scheme. For example, a periodic generalization of $V_{pp'}$ such as

$$V_{pp'} = \sum_{\mathbf{R}} \int g_p(\mathbf{r}_1) \frac{1}{r_{12}} g_{p'}(\mathbf{r}_2-\mathbf{R}) d^3r_1 d^3r_2$$

clearly diverges at large R since the number of lattice points at a distance R from the origin is proportional to R^2 , and this factor overwhelms $(r_{12})^{-1} \cong R^{-1}$. To obtain a convergent sum we must rely on the charge neutrality of the system; that is, we must include nuclear as well as electronic charges in the sum.¹⁷ We do this by introducing a set of modified density basis functions that include a share of the total nuclear charge in the unit cell, such that each function contains zero net charge. As before, let α count the nuclei at points \mathbf{r}_{α} in the home cell, with charges Z_{α} such that $\sum_{\alpha} Z_{\alpha} = Q$. Defining the basis function

$$h_p(\mathbf{r}) = g_p(\mathbf{r}) - n_p \sum_{\alpha} \frac{Z_{\alpha}}{Q} \delta(\mathbf{r}-\mathbf{r}_{\alpha}), \quad (25)$$

the total fitted charge density, now including the nuclear charges, will be

$$\tilde{\rho}_{\text{tot}}(\mathbf{r}) = \sum_{p,\mathbf{R}} a_p h_p(\mathbf{r}-\mathbf{R}). \quad (26)$$

Minimizing the self-interaction of the charge density error as before¹³ leads again to

$$a_p = \sum_{p'} (V^{-1})_{pp'} (t_{p'} + \lambda n_{p'}), \quad (27)$$

where now

$$V_{pp'} = \int h_p(\mathbf{r}_1) \frac{1}{r_{12}} \sum_{\mathbf{R}} h_{p'}(\mathbf{r}_2-\mathbf{R}) d^3r_1 d^3r_2,$$

$$t_p = \int h_p(\mathbf{r}_1) \frac{1}{r_{12}} \rho_{\text{tot}}(\mathbf{r}_2) d^3r_1 d^3r_2,$$

$$\lambda = \frac{Q - \sum_{pp'} n_p (V^{-1})_{pp'} t_{p'}}{\sum_{pp'} n_p (V^{-1})_{pp'} n_{p'}},$$

$$n_p = \int g_p(\mathbf{r}) d^3r. \quad (28)$$

The equations for \mathbf{V} and \mathbf{t} involve localized distributions interacting with periodic ones (ρ_{tot} without the tilde is the exact total density); each includes nuclear–nuclear, electron–electron, and electron–nucleus interactions, and each is now conditionally convergent, i.e., the series may sum to a finite value but the exact results will depend on the order in which the terms are added.¹⁸ Further manipulations to obtain absolutely and rapidly convergent sums will be described in Sec. III below.

We can write the Kohn–Sham Hamiltonian as

$$H_{ij}(\mathbf{k}) = \sum_{\mathbf{R}} e^{i\mathbf{k}\cdot\mathbf{R}} [H_{ij}^{\text{kin}}(\mathbf{R}) + H_{ij}^{\text{coul}}(\mathbf{R}) + H_{ij}^{\text{XC}}(\mathbf{R})], \quad (29)$$

where

$$H_{ij}^{\text{kin}}(\mathbf{R}) = \int \chi_i(\mathbf{r}) \left(-\frac{1}{2} \nabla^2\right) \chi_j(\mathbf{r}-\mathbf{R}) d^3r. \quad (30)$$

If we next define an interaction integral between a product charge and a periodic array of neutralized charge fitting functions

$$C_{ijp}(\mathbf{R}) = \int \chi_i(\mathbf{r}_1) \chi_j(\mathbf{r}_1 - \mathbf{R}) \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} \times \sum_{\mathbf{R}'} h_p(\mathbf{r}_2 - \mathbf{R}') d^3 r_1 d^3 r_2, \quad (31)$$

then the Coulomb term in Eq. (29) is given by

$$H_{ij}^{\text{coul}}(\mathbf{R}) = \sum_p a_p C_{ijp}(\mathbf{R}). \quad (32)$$

The charge density coefficient a_p is given by Eq. (27), where now

$$t_p = T_p + \sum_{ij\mathbf{R}} P_{ij}(\mathbf{R}) C_{ijp}(\mathbf{R}) \quad (33)$$

and

$$T_p = \sum_{\alpha\mathbf{R}} \int \frac{-Z_\alpha}{|\mathbf{r} - \mathbf{r}_\alpha|} h_p(\mathbf{r} - \mathbf{R}) d^3 r. \quad (34)$$

The exchange-correlation (XC) term in the effective Hamiltonian is found by numerical integration:

$$H_{ij}^{\text{XC}}(\mathbf{R}) = \int \chi_i(\mathbf{r}) \chi_j(\mathbf{r} - \mathbf{R}) V_{\text{XC}}[\rho(\mathbf{r})] d^3 r. \quad (35)$$

The density and hence V_{XC} have the periodicity of the lattice, but the range of the integral can be restricted in practice to the region (possibly quite small) in which the product of basis functions is significantly different from zero. If we choose to also fit V_{XC} we can express Eq. (35) in terms of b_q [Eq. (23)] and a three-center overlap of orbital and fitting functions in analogy to Eq. (32).

The total energy per primitive unit cell is given by

$$E_{\text{tot}} = E_{\text{kin}} + E_{\text{XC}} + E_{\text{coul}}, \quad (36)$$

where

$$E_{\text{kin}} = \sum_{ij\mathbf{R}} P_{ij}(\mathbf{R}) H_{ij}^{\text{kin}}(\mathbf{R}). \quad (37)$$

For the exchange-correlation part of the total energy, we need to find the electronic part of the charge density and from it evaluate the exchange-correlation energy functional $\epsilon_{\text{XC}}[\rho]$ (recall that the exchange-correlation potential is related to the energy functional by $V_{\text{XC}} = \epsilon_{\text{XC}} + \delta\epsilon_{\text{XC}}/\delta\rho$, where ρ is the electronic density.) Then the XC energy is found by numerical integration as in Eq. (35):

$$E_{\text{XC}} = \sum_{ij\mathbf{R}} P_{ij}(\mathbf{R}) \int \chi_i(\mathbf{r}) \chi_j(\mathbf{r} - \mathbf{R}) \epsilon_{\text{XC}}[\rho(\mathbf{r})] d^3 r. \quad (38)$$

The XC energy can also be expressed in terms of a fit to the functional ϵ_{XC} analogous to Eq. (23). Finally, the Coulomb term in the total energy can be found by substituting Eqs. (20) and (26) into Eq. (10) and using Eqs. (28a) and (31):

$$E_{\text{coul}} = \sum_{ijp\mathbf{R}} a_p P_{ij}(\mathbf{R}) C_{ijp}(\mathbf{R}) + \sum_p a_p T_p - \frac{1}{2} \sum_{pp'} a_p V_{pp'} a_{p'}.$$

The first two terms here are equivalent to Eq. (33) multiplied by a_p and summed on p , while the third term can be simplified by matrix multiplying Eq. (27) by \mathbf{V} , leading to the compact expression

$$E_{\text{coul}} = \frac{1}{2} \sum_p a_p t_p - \frac{1}{2} \lambda Q. \quad (39)$$

III. MULTICENTER INTEGRALS AND COULOMB SUMMATION

Next we must find absolutely convergent expressions for the sums over integrals defined in Eqs. (28a), (31), and (34). We use a version of a recently proposed technique¹⁸ for calculating the Coulomb potential in periodic Hartree-Fock theory in which interactions of point multipoles are subtracted from the original expression and then added again, and the added terms are summed by the use of the Ewald convention.¹⁹ Related techniques have also been discussed by Saunders.²⁰ We first expand the integrals defined in Eq. (31) according to

$$C_{ijp}(\mathbf{R}) = \sum_{\mathbf{R}'} [C_{ijp}^{(1)}(\mathbf{R}, \mathbf{R}') + C_{ijp}^{(2)}(\mathbf{R}, \mathbf{R}') - C_{ijp}^{(3)}(\mathbf{R}, \mathbf{R}') - C_{ijp}^{(4)}(\mathbf{R}, \mathbf{R}')] + C_{ijp}^{\text{Ewald},3}(\mathbf{R}) + C_{ijp}^{\text{Ewald},4}(\mathbf{R}). \quad (40)$$

The various terms have the following meanings: $C^{(1)}$ contains three-center electron repulsion integrals between the product charge $\chi_i \chi_j$ and the fitting function g_p ; $C^{(2)}$ contains an electronuclear attraction integral between the product charge $\chi_i \chi_j$ and the compensating δ functions in h_p ; $C^{(3)}$ contains interactions between point objects having the same charge, dipole, and quadrupole moments as $\chi_i \chi_j$ and g_p , so that the \mathbf{R}' sum over $C^{(1)} - C^{(3)}$ is absolutely convergent; and $C^{(4)}$ contains interactions between a point object having the same charge, dipole, and quadrupole moment as $\chi_i \chi_j$ and the compensating δ functions from h_p . The sum over $C^{(2)} - C^{(4)}$ is then also absolutely convergent.

The Ewald terms represent the same interactions as $C^{(3)}$ and $C^{(4)}$ summed according to the Ewald convention.¹⁹ These terms must be added to compensate for $C^{(3)}$ and $C^{(4)}$ which were subtracted to improve the convergence of the \mathbf{R}' sum, which naively would only contain the conditionally convergent sum over $C^{(1)} + C^{(2)}$.

Next we work out the explicit forms. First we have

$$C_{ijp}^{(1)}(\mathbf{R}, \mathbf{R}') = \int \chi_i(\mathbf{r}_1) \chi_j(\mathbf{r}_1 - \mathbf{R}) \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} \times g_p(\mathbf{r}_2 - \mathbf{R}') d^3 r_1 d^3 r_2. \quad (41)$$

This is just a standard molecular three-center repulsion integral with basis function centers shifted by the indicated lattice vectors. Next

$$C_{ijp}^{(2)}(\mathbf{R}, \mathbf{R}') = -n_p \sum_{\alpha} \frac{Z_{\alpha}}{Q} \int \chi_i(\mathbf{r}_1) \chi_j(\mathbf{r}_1 - \mathbf{R}) \\ \times \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} \delta(\mathbf{r}_2 - \mathbf{r}_{\alpha} - \mathbf{R}') d^3 r_1 d^3 r_2.$$

Integrating out the delta function and shifting the origin in the integral, we have

$$C_{ijp}^{(2)}(\mathbf{R}, \mathbf{R}') = -n_p \sum_{\alpha} \frac{Z_{\alpha}}{Q} \int \chi_i(\mathbf{r} + \mathbf{R}') \chi_j(\mathbf{r} + \mathbf{R}' - \mathbf{R}) \\ \times \frac{1}{|\mathbf{r} - \mathbf{r}_{\alpha}|} d^3 r. \quad (42)$$

The integral here is a standard molecular electronuclear attraction between an orbital product charge and a nucleus at \mathbf{r}_{α} . Note that only the factor n_p depends on p so we need to evaluate the rest of the expression only once for each set of $ij\mathbf{R}$ before entering the loop on p . It is essential to proceed this way to preserve the formal N^3 scaling of the theory; if the integral is repeated for every set ijp , the sum on α increases the cost to order N^4 .

Now we consider $C^{(3)}$ and $C^{(4)}$. Since they involve only interactions of point objects they contain no multicenter integrals, just various inverse powers of the distance between the objects, and some moments which are found by one-center integration. Let “1” designate the product charge and “2” the fitted charge, and let μ, ν be indices running over the Cartesian x, y, z axes. Also let \mathbf{r}_1 represent the assumed location of the point multipoles representing the product charge. Then

$$q_1 = \int \chi_i(\mathbf{r}) \chi_j(\mathbf{r} - \mathbf{R}) d^3 r, \\ \mathbf{p}_1 = \int (\mathbf{r} - \mathbf{r}_1) \chi_i(\mathbf{r}) \chi_j(\mathbf{r} - \mathbf{R}) d^3 r, \quad (43) \\ Q_{1\mu\nu} = \int [3(\mathbf{r} - \mathbf{r}_1)_{\mu}(\mathbf{r} - \mathbf{r}_1)_{\nu} - \delta_{\mu\nu} |\mathbf{r} - \mathbf{r}_1|^2] \chi_i(\mathbf{r}) \\ \times \chi_j(\mathbf{r} - \mathbf{R}) d^3 r.$$

Here \mathbf{Q}_1 is the traceless Cartesian quadrupole tensor of the product charge. We can write similar expressions for the moments of the fitting charge:

$$q_2 = \int g_p(\mathbf{r}) d^3 r = n_p, \\ \mathbf{p}_2 = \int (\mathbf{r} - \mathbf{r}_{\alpha(p)}) g_p(\mathbf{r}) d^3 r, \quad (44) \\ \mathbf{Q}_2 = \int [3(\mathbf{r} - \mathbf{r}_{\alpha(p)})(\mathbf{r} - \mathbf{r}_{\alpha(p)}) - \mathbf{1} |\mathbf{r} - \mathbf{r}_{\alpha(p)}|^2] g_p(\mathbf{r}) d^3 r,$$

where we have written the quadrupole tensor of the fitting charge in a more compact tensor notation ($\mathbf{1}$ is the unit tensor

or identity matrix) and $\mathbf{r}_{\alpha(p)}$ represents the position of the nucleus on which the fitting function $g_p(\mathbf{r})$ is centered. Application of electrostatic theory now permits us to write

$$C_{ijp}^{(3)}(\mathbf{R}, \mathbf{R}') \\ = \frac{q_1 q_2}{d} - \frac{(q_1 \mathbf{p}_2 - q_2 \mathbf{p}_1) \cdot \mathbf{n}}{d^2} \\ + \frac{\frac{1}{2} \mathbf{n} \cdot (q_1 \mathbf{Q}_2 + q_2 \mathbf{Q}_1) \cdot \mathbf{n} + \mathbf{p}_1 \cdot \mathbf{p}_2 - 3(\mathbf{p}_1 \cdot \mathbf{n})(\mathbf{p}_2 \cdot \mathbf{n})}{d^3}, \quad (45)$$

where $\mathbf{d} = \mathbf{r}_{\alpha(p)} + \mathbf{R}' - \mathbf{r}_1$, $d = |\mathbf{d}|$, and $\mathbf{n} = \mathbf{d}/d$. We will discuss the choice of \mathbf{r}_1 below, but note here that \mathbf{r}_1 and hence \mathbf{d} depend on positions $\mathbf{r}_{\alpha(i)}$ and $\mathbf{r}_{\alpha(j)} + \mathbf{R}$ of the centers of the orbital basis functions, and possibly also on their exponents, angular momentum type and contraction scheme. When Eq. (45) is subtracted from Eq. (41), the difference decays with distance at least as rapidly as d^{-4} , which is why the sum over these differences converges absolutely.

Using Eq. (43) we can also find the fourth term:

$$C_{ijp}^{(4)}(\mathbf{R}, \mathbf{R}') = -n_p \sum_{\alpha} \frac{Z_{\alpha}}{Q} \left\{ \frac{q_1}{d} + \frac{\mathbf{p}_1 \cdot \mathbf{n}}{d^2} + \frac{\mathbf{n} \cdot \mathbf{Q}_1 \cdot \mathbf{n}}{2d^3} \right\}. \quad (46)$$

Here $\mathbf{d} = \mathbf{r}_{\alpha} + \mathbf{R}' - \mathbf{r}_1$ which depends on the nuclear index α and so must remain inside the sum on α . Note, though, that as with $C^{(2)}$ only the prefactor depends on p , so the sum over α can be done once for each i, j, \mathbf{R} , and \mathbf{R}' outside the loop on p . In fact we can presubtract this sum from the corresponding sum in $C^{(2)}$, determine the effective range of the \mathbf{R}' sum for $C^{(2)} + C^{(4)}$, do that sum, and store the result before looping over p . Again, Eq. (42) minus Eq. (46) decays at least as rapidly as d^{-4} .

Next we must consider the last two terms in Eq. (40), which sum the long-ranged Coulomb interactions according to the Ewald convention. First let us consider the expression for the electrostatic potential at a point \mathbf{d} due to a set of unit point charges at the points \mathbf{R} of a Bravais lattice (we assume the origin is one of these \mathbf{R} points). The potential of a uniform neutralizing background is implicitly included, but cancels when we sum over the components of an electroneutral system. If point \mathbf{d} does not coincide with any lattice point \mathbf{R} then this Ewald potential function is¹⁹

$$\Phi(\mathbf{d}) = -\frac{\pi \eta^2}{V} + \sum_{\mathbf{R}} \frac{\text{erfc}(|\mathbf{d} - \mathbf{R}|/\eta)}{|\mathbf{d} - \mathbf{R}|} \\ + \frac{4\pi}{V} \sum_{\mathbf{K} \neq 0} \frac{e^{i\mathbf{K} \cdot \mathbf{d} - K^2 \eta^2/4}}{K^2}, \quad (47)$$

where η is a constant length scale of somewhat arbitrary value, a good choice²⁰ being $\eta = 0.36V^{1/3}$. \mathbf{K} is a reciprocal lattice vector, $K = |\mathbf{K}|$ and V is the direct lattice primitive cell volume. This function may be thought of as a lattice generalization of the quantity $1/d$ which is finite (both \mathbf{R} and \mathbf{K} sums are rapidly convergent) and periodic [$\Phi(\mathbf{d} + \mathbf{R}) = \Phi(\mathbf{d})$].

We can construct the Ewald terms $C^{\text{Ewald},3}$ and $C^{\text{Ewald},4}$ by expressing $C^{(3)}$ and $C^{(4)}$ in terms of gradients (with respect to \mathbf{d}) of $1/d$, and then replacing $1/d$ by $\Phi(\mathbf{d})$. The resulting expression for $C^{\text{Ewald},3}$ is

$$C_{ijp}^{\text{Ewald},3}(\mathbf{R}) = q_1 q_2 \Phi(\mathbf{d}) + (q_1 \mathbf{p}_2 - q_2 \mathbf{p}_1) \cdot \nabla \Phi(\mathbf{d}) - (\mathbf{p}_1 \cdot \nabla)(\mathbf{p}_2 \cdot \nabla) \Phi(\mathbf{d}) + \frac{1}{6} (q_1 \mathbf{Q}_2 + q_2 \mathbf{Q}_1) : \nabla \nabla \Phi(\mathbf{d}), \quad (48)$$

where $\mathbf{d} = \mathbf{r}_{\alpha(p)} - \mathbf{r}_1$ and the following notation for differential operators is used:

$$(\mathbf{p}_1 \cdot \nabla)(\mathbf{p}_2 \cdot \nabla) = \sum_{\mu, \nu=1}^3 p_{1\mu} p_{2\nu} \frac{\partial^2}{\partial d_\mu \partial d_\nu},$$

$$\mathbf{Q}_1 : \nabla \nabla = \sum_{\mu, \nu=1}^3 Q_{1\mu\nu} \frac{\partial^2}{\partial d_\mu \partial d_\nu}.$$

Similarly,

$$C_{ijp}^{\text{Ewald},4}(\mathbf{R}) = -n_p \sum_\alpha \frac{Z_\alpha}{Q} \left\{ q_1 \Phi(\mathbf{d}) - \mathbf{p}_1 \cdot \nabla \Phi(\mathbf{d}) + \frac{1}{6} \mathbf{Q}_1 : \nabla \nabla \Phi(\mathbf{d}) \right\}, \quad (49)$$

where now $\mathbf{d} = \mathbf{r}_\alpha - \mathbf{r}_1$ (and hence depends on α , so it must stay inside the sum.) As with $C^{(2)}$ and $C^{(4)}$, only the prefactor depends on p ; hence the rest of the expression only needs to be evaluated once for all p (for a given i, j , and \mathbf{R}).

To calculate these quantities we need suitable expressions for the Ewald potential and its first and second derivatives. Using the fact that the reciprocal lattice sum contains $-\mathbf{K}$ for every \mathbf{K} , the potential is, for $\mathbf{d} \neq 0$,

$$\Phi(\mathbf{d}) = \sum_{\mathbf{R}} \frac{\text{erfc}(|\mathbf{d} - \mathbf{R}|/\eta)}{|\mathbf{d} - \mathbf{R}|} + \frac{4\pi}{V} \sum_{\mathbf{K} \neq 0} \frac{\cos(\mathbf{K} \cdot \mathbf{d})}{K^2} e^{-K^2 \eta^2/4} - \frac{\pi \eta^2}{V}. \quad (50)$$

The gradient and Hessian of this expression with respect to \mathbf{d} can now be taken straightforwardly (see Appendix A) and corresponding expressions for the case where $\mathbf{d} = 0$ can be found by an appropriate limiting process (Appendix B). This is all the information needed (at least in principle) to calculate $C_{ijp}(\mathbf{R})$. One should bear in mind that $C^{(1)}$ and $C^{(3)}$ should be computed and subtracted at fixed \mathbf{R}' before summing on \mathbf{R}' out to distances where the difference is small enough to neglect. The procedure is similar for $C^{(2)}$ and $C^{(4)}$ except that the sum is done independently of p before looping over p to prevent recomputation of the sum on α for every p .

The calculation of the quantities $V_{pp'}$ is similar to that of $C_{ijp}(\mathbf{R})$. From Eq. (28a) we have

$$V_{pp'} = \int h_p(\mathbf{r}_1) \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} \sum_{\mathbf{R}} h_{p'}(\mathbf{r}_2 - \mathbf{R}) d^3 r_1 d^3 r_2.$$

Proceeding by analogy to the C_{ijp} calculation,

$$V_{pp'} = \sum_{\mathbf{R}} [V_{pp'}^{(1)}(\mathbf{R}) + V_{pp'}^{(2)}(\mathbf{R}) - V_{pp'}^{(3)}(\mathbf{R}) - V_{pp'}^{(4)}(\mathbf{R})] + V_{pp'}^{\text{Ewald},3} + V_{pp'}^{\text{Ewald},4} + V_{pp'}^{\text{Ewald},NN}, \quad (51)$$

where

$$V_{pp'}^{(1)}(\mathbf{R}) = \int g_p(\mathbf{r}_1) \frac{1}{|\mathbf{r}_1 - \mathbf{r}_2|} g_{p'}(\mathbf{r}_2 - \mathbf{R}) d^3 r_1 d^3 r_2, \quad (52)$$

which is an electronic repulsion integral between fitting functions $g_p, g_{p'}$ centered at $\mathbf{r}_{\alpha(p)}$ and $\mathbf{r}_{\alpha(p')} + \mathbf{R}$. Next

$$V_{pp'}^{(2)}(\mathbf{R}) = -n_p \sum_\alpha \frac{Z_\alpha}{Q} \int \frac{1}{|\mathbf{r}_\alpha - \mathbf{r}_2|} g_{p'}(\mathbf{r}_2 - \mathbf{R}) d^3 r_2 - n_{p'} \sum_\alpha \frac{Z_\alpha}{Q} \int g_p(\mathbf{r}_1) \frac{1}{|\mathbf{r}_1 - \mathbf{r}_\alpha - \mathbf{R}|} d^3 r_1 = - \sum_\alpha \frac{Z_\alpha}{Q} \int \frac{1}{|\mathbf{r} - \mathbf{r}_\alpha|} \{ n_p g_{p'}(\mathbf{r} - \mathbf{R}) + n_{p'} g_p(\mathbf{r} + \mathbf{R}) \} d^3 r. \quad (53)$$

The integrals in this expression correspond to electronuclear attractions of fitting functions centered at $\mathbf{r}_{\alpha(p')} + \mathbf{R}$ and $\mathbf{r}_{\alpha(p)} - \mathbf{R}$ with a nucleus at \mathbf{r}_α .

The next terms $V^{(3)}$ and $V^{(4)}$ are pointlike approximations to $V^{(1)}$ and $V^{(2)}$, respectively. The moments are derived from the fitting functions exactly as Eq. (44), except that we now use ‘‘1’’ to denote the unprimed function and ‘‘2’’ the primed function. Then the expression for $V^{(3)}$ formally becomes the same as Eq. (41) for $C^{(3)}$:

$$V_{pp'}^{(3)}(\mathbf{R}) = \frac{q_1 q_2}{d} - \frac{(q_1 \mathbf{p}_2 - q_2 \mathbf{p}_1) \cdot \mathbf{n}}{d^2} + \frac{\frac{1}{2} \mathbf{n} \cdot (q_1 \mathbf{Q}_2 + q_2 \mathbf{Q}_1) \cdot \mathbf{n} + \mathbf{p}_1 \cdot \mathbf{p}_2 - 3(\mathbf{p}_1 \cdot \mathbf{n})(\mathbf{p}_2 \cdot \mathbf{n})}{d^3}, \quad (54)$$

where $\mathbf{d} = \mathbf{r}_{\alpha(p')} + \mathbf{R} - \mathbf{r}_{\alpha(p)}$. Also, $V^{(4)}$ is the sum of two terms that closely resemble Eq. (46) for $C^{(4)}$:

$$V_{pp'}^{(4)}(\mathbf{R}) = - \sum_\alpha \frac{Z_\alpha}{Q} \left\{ q_1 \left[\frac{q_2}{d} + \frac{\mathbf{p}_2 \cdot \mathbf{n}}{d^2} + \frac{\mathbf{n} \cdot \mathbf{Q}_2 \cdot \mathbf{n}}{2d^3} \right] + q_2 \left[\frac{q_1}{d'} + \frac{\mathbf{p}_1 \cdot \mathbf{n}'}{d'^2} + \frac{\mathbf{n}' \cdot \mathbf{Q}_1 \cdot \mathbf{n}'}{2d'^3} \right] \right\}, \quad (55)$$

where the multipoles are again defined as in Eq. (44), $\mathbf{d} = \mathbf{r}_\alpha - \mathbf{r}_{\alpha(p')} - \mathbf{R}$ and $\mathbf{d}' = \mathbf{r}_\alpha - \mathbf{r}_{\alpha(p)} + \mathbf{R}$ ($d' = |\mathbf{d}'|$ and $\mathbf{n}' = \mathbf{d}'/d'$).

The Ewald terms in V are closely analogous to those in C as well. We have

$$V_{pp'}^{\text{Ewald},3} = q_1 q_2 \Phi(\mathbf{d}) + (q_1 \mathbf{p}_2 - q_2 \mathbf{p}_1) \cdot \nabla \Phi(\mathbf{d}) - (\mathbf{p}_1 \cdot \nabla) \times (\mathbf{p}_2 \cdot \nabla) \Phi(\mathbf{d}) + \frac{1}{6} (q_1 \mathbf{Q}_2 + q_2 \mathbf{Q}_1) : \nabla \nabla \Phi(\mathbf{d}), \quad (56)$$

where now $\mathbf{d} = \mathbf{r}_{\alpha(p')} - \mathbf{r}_{\alpha(p)}$ and the moments are the same as in Eq. (53). Also,

$$V_{pp'}^{\text{Ewald},4} = - \sum_{\alpha} \frac{Z_{\alpha}}{Q} \{ q_1 [q_2 \Phi(\mathbf{d}) - \mathbf{p}_2 \cdot \nabla \Phi(\mathbf{d}) + \frac{1}{6} \mathbf{Q}_2 : \nabla \nabla \Phi(\mathbf{d})] + q_2 [q_1 \Phi(\mathbf{d}') - \mathbf{p}_1 \cdot \nabla \Phi(\mathbf{d}') + \frac{1}{6} \mathbf{Q}_1 : \nabla \nabla \Phi(\mathbf{d}')] \}, \quad (57)$$

where $\mathbf{d} = \mathbf{r}_{\alpha} - \mathbf{r}_{\alpha(p')}$, $\mathbf{d}' = \mathbf{r}_{\alpha} - \mathbf{r}_{\alpha(p)}$. Finally,

$$V_{pp'}^{\text{Ewald},NN} = n_p n_{p'} \sum_{\alpha\alpha'} \frac{Z_{\alpha} Z_{\alpha'}}{Q^2} \Phi(\mathbf{r}_{\alpha} - \mathbf{r}_{\alpha'}). \quad (58)$$

The only other quantity requiring the use of Ewald techniques to guarantee convergence is the “projection” integral T_p . From Eq. (34)

$$T_p = - \sum_{\alpha} Z_{\alpha} \int \frac{1}{|\mathbf{r} - \mathbf{r}_{\alpha}|} \sum_{\mathbf{R}} h_p(\mathbf{r} - \mathbf{R}) d^3 r$$

we expand

$$T_p = \sum_{\mathbf{R}} [T_p^{(1)}(\mathbf{R}) - T_p^{(2)}(\mathbf{R})] + T_p^{\text{Ewald},2} + T_p^{\text{Ewald},NN}, \quad (59)$$

where

$$T_p^{(1)}(\mathbf{R}) = - \sum_{\alpha} Z_{\alpha} \int \frac{1}{|\mathbf{r} - \mathbf{r}_{\alpha}|} g_p(\mathbf{r} - \mathbf{R}) d^3 r. \quad (60)$$

This is a fitted electronuclear attraction integral of the same type as those appearing in Eq. (53). It is compensated by a term in which the fitting function is replaced by point multipoles, namely

$$T_p^{(2)}(\mathbf{R}) = - \sum_{\alpha} Z_{\alpha} \left[\frac{q}{d} + \frac{\mathbf{p} \cdot \mathbf{n}}{d^2} + \frac{\mathbf{n} \cdot \mathbf{Q} \cdot \mathbf{n}}{2d^3} \right], \quad (61)$$

where $\mathbf{d} = \mathbf{r}_{\alpha} - \mathbf{r}_{\alpha(p)} - \mathbf{R}$ and the moments are defined as in Eq. (44). Using these definitions, the Ewald terms for T_p are

$$T_p^{\text{Ewald},2} = - \sum_{\alpha} Z_{\alpha} \{ q \Phi(\mathbf{d}) - \mathbf{p} \cdot \nabla \Phi(\mathbf{d}) + \frac{1}{6} \mathbf{Q} : \nabla \nabla \Phi(\mathbf{d}) \}, \quad (62)$$

where $\mathbf{d} = \mathbf{r}_{\alpha} - \mathbf{r}_{\alpha(p)}$, and

$$T_p^{\text{Ewald},NN} = \frac{n_p}{Q} \sum_{\alpha\alpha'} Z_{\alpha} Z_{\alpha'} \Phi(\mathbf{r}_{\alpha} - \mathbf{r}_{\alpha'}). \quad (63)$$

Recall that the scalar Q is the total number of electrons (or nuclear charges) per primitive cell, not to be confused with the traceless quadrupole tensor \mathbf{Q} .

Many of the quantities appearing in these expressions for $V_{pp'}$ and T_p are interrelated. For example if we define the self-energy of the lattice of nuclei (with a uniform compensating charge background) as

$$E_{\text{nuc}} = \sum_{\alpha\alpha'} Z_{\alpha} Z_{\alpha'} \Phi(\mathbf{r}_{\alpha} - \mathbf{r}_{\alpha'}), \quad (64)$$

then we have

$$T_p^{\text{Ewald},NN} = \frac{n_p}{Q} E_{\text{nuc}}, \quad V_{pp'}^{\text{Ewald},NN} = \frac{n_p n_{p'}}{Q^2} E_{\text{nuc}}. \quad (65)$$

Also, we can easily show that

$$V_{pp'}^{(2)}(\mathbf{R}) = \frac{n_p}{Q} T_{p'}^{(1)}(\mathbf{R}) + \frac{n_{p'}}{Q} T_p^{(1)}(-\mathbf{R}),$$

$$V_{pp'}^{(4)}(\mathbf{R}) = \frac{n_p}{Q} T_{p'}^{(2)}(\mathbf{R}) + \frac{n_{p'}}{Q} T_p^{(2)}(-\mathbf{R}), \quad (66)$$

$$V_{pp'}^{\text{Ewald},4} = \frac{n_p}{Q} T_{p'}^{\text{Ewald},2} + \frac{n_{p'}}{Q} T_p^{\text{Ewald},2}.$$

Before closing this section we need to define the orbital product charge center \mathbf{r}_1 appearing in the multipole moments of Eq. (43) and the displacement \mathbf{d} in Eqs. (45)–(48). This definition is somewhat arbitrary (as long as it is the same in the multipoles and in \mathbf{d}) but its precise choice will affect the convergence of the \mathbf{R}' sum in Eq. (40) and hence will affect the accuracy and efficiency of the entire scheme. We have obtained good convergence (see Sec. VII) using a “weighted midpoint” for the charge center

$$\mathbf{r}_1 = \frac{a_i^{\min} \mathbf{r}_{\alpha(i)} + a_j^{\min} (\mathbf{r}_{\alpha(j)} + \mathbf{R})}{a_i^{\min} + a_j^{\min}}, \quad (67)$$

where α_i^{\min} is the smallest exponent in the contraction of the i th orbital basis function. The sum on \mathbf{R}' is usually taken over a roughly spherical set of lattice points centered on the origin, but when \mathbf{R} is so large that the closest lattice point to \mathbf{r}_1 is not the origin, we have found it helpful to center the sum around that other lattice point instead. For those cases where $q_1 \neq 0$ in Eq. (43), a possible further refinement is to shift the center \mathbf{r}_1 so as to make the dipole moment \mathbf{p}_1 vanish; however, we have not found this to be helpful in practice.

We thus have a practical and mathematically well-behaved scheme for calculating the quantities $C_{ijp}(\mathbf{R})$, $V_{pp'}$, and T_p needed for fitting the charge density and computing the Coulomb contributions to the Hamiltonian and total energy. Examples of the actual numerical performance of this scheme are given in Sec. VII below.

IV. TWO-DIMENSIONAL PERIODICITY

We next briefly consider how to treat systems which are periodic in two directions and finite in the third (the z direction.) In this case the direct lattice vectors \mathbf{R} , reciprocal lattice vectors \mathbf{K} , and special \mathbf{k} sampling points have nonzero components only in the xy plane. However real-space integrals, whether analytic or numerical, are still standard three-dimensional volume integrals since the atoms are still three-dimensional objects; and the nuclear coordinates \mathbf{r}_{α} still have nonzero z components z_{α} in general, since we will be treating slabs that are several layers thick. When these changes

are made, it turns out that all expressions from Eq. (14) to (67) are formally correct for the two-dimensionally periodic case, with the exception that the expressions for the Ewald potential and its derivatives take different forms. Also, it is no longer absolutely necessary to add and subtract the charge–quadrupole and dipole–dipole interactions in Eqs. (40), (51), and (59), since only the charge–charge and charge–dipole interactions give conditionally convergent sums. However, it is still desirable to add and subtract these higher order terms to obtain faster convergence of the series and to allow the use of the same formalism for all lattice dimensionalities. The expression for the Ewald potential when $\mathbf{d} \neq 0$ is now²¹

$$\Phi(\mathbf{d}) = \Phi_1(\mathbf{d}) + \Phi_2(\mathbf{d}) + \Phi_3(\mathbf{d}), \quad (68)$$

where

$$\begin{aligned} \Phi_1(\mathbf{d}) &= \sum_{\mathbf{R}} \frac{\text{erfc}(|\mathbf{d} - \mathbf{R}|/\eta)}{|\mathbf{d} - \mathbf{R}|}, \\ \Phi_2(\mathbf{d}) &= \frac{\pi}{A} \sum_{\mathbf{K} \neq 0} \frac{\cos(\mathbf{K} \cdot \mathbf{d})}{K} \left[e^{Kd_z} \text{erfc}\left(\frac{K\eta}{2} + \frac{d_z}{\eta}\right) \right. \\ &\quad \left. + e^{-Kd_z} \text{erfc}\left(\frac{K\eta}{2} - \frac{d_z}{\eta}\right) \right] \\ \Phi_3(\mathbf{d}) &= -\frac{\pi}{A} \left[\frac{2\eta}{\sqrt{\pi}} e^{-(d_z/\eta)^2} + 2d_z \text{erfc}(d_z/\eta) \right]. \end{aligned} \quad (69)$$

Φ_3 replaces the $\mathbf{K}=0$ term in Φ_2 and corresponds to the constant term in Eq. (50). Derivatives of these expressions can easily be found (Appendix A), and the case $\mathbf{d}=0$ is treated by the appropriate limiting process (Appendix B). When we use Eqs. (68) and (69) instead of (52) for the Ewald potential and interpret direct and reciprocal lattice vectors (and special \mathbf{k} points, but not necessarily nuclear position vectors) as two-dimensional vectors, the entire formalism developed above gives a correct density functional theory for slabs of any finite thickness.

V. ENERGY GRADIENTS (FORCES)

We next consider how to calculate analytic first derivatives of the total energy in the periodic case. We assume in this section that the XC potential will not be fitted to an auxiliary basis, since even if this fitting saves some computer time and is acceptably accurate for the total energy, the lack of an XC energy variational principle (in contrast to the density fit) is likely to cause unacceptable errors in the XC contribution to the force.¹² Also, we assume that the SCF process has been iterated to self-consistency at the atomic geometry under consideration; if this is not the case, as in a Car–Parinello²² type of approach, additional forces to correct for non-self-consistency will appear.

Let us first return to the molecular case¹² to express the usual results in our notation. Let y denote a general nuclear coordinate, e.g., a particular Cartesian component of a particular nuclear coordinate \mathbf{r}_α in a cluster or molecule. Also let us define the molecular two and three-center integrals

$$\begin{aligned} H_{ij}^{\text{kin}} &= \int \chi_i(\mathbf{r})(-\nabla^2/2)\chi_j(\mathbf{r})d^3r, \\ H_{ij}^{\text{nuc}} &= \int \chi_i(\mathbf{r})\left(\sum_{\alpha} \frac{-Z_{\alpha}}{|\mathbf{r}-\mathbf{r}_{\alpha}|}\right)\chi_j(\mathbf{r})d^3r \\ C_{ijp} &= \int \chi_i(\mathbf{r}_1)\chi_j(\mathbf{r}_1) \frac{1}{|\mathbf{r}_1-\mathbf{r}_2|} g_p(\mathbf{r}_2)d^3r_1 d^3r_2. \end{aligned} \quad (70)$$

Then the gradient of the energy works out to be

$$\begin{aligned} \frac{dE}{dy} &= \sum_{ij} P_{ij} \left\{ \frac{dH_{ij}^{\text{kin}}}{dy} + \frac{dH_{ij}^{\text{nuc}}}{dy} + \sum_p \frac{dC_{ijp}}{dy} \right. \\ &\quad \left. + \int V_{\text{xc}}[\rho] \frac{d}{dy} [\chi_i(\mathbf{r})\chi_j(\mathbf{r})]d^3r \right\} \\ &\quad - \frac{1}{2} \sum_{pp'} a_p V_{pp'} a_{p'} + \frac{1}{2} \frac{d}{dy} \sum_{\alpha \neq \alpha'} \frac{Z_{\alpha} Z_{\alpha'}}{|\mathbf{r}_{\alpha} - \mathbf{r}_{\alpha'}|} \\ &\quad - \sum_{ij} P_{ij}^{\epsilon} \frac{dS_{ij}}{dy}, \end{aligned} \quad (71)$$

where P_{ij}^{ϵ} is the energy-weighted density matrix given by

$$P_{ij}^{\epsilon} = \sum_n^{\text{occ}} \epsilon_n c_{ni}^* c_{nj}, \quad (72)$$

and the integral over V_{xc} is again to be done by numerical quadrature. As usual¹² the derivatives act on the multicenter integrals through the implicit dependence of each basis function on the position of the nucleus where it is centered; we also have derivatives of explicit nuclear coordinates in the denominators of the H^{nuc} and nuclear repulsion terms which lead to the usual Helman–Feynman force, the other terms being the Pulay corrections.¹²

In the periodic case, the derivative of the energy of the entire crystal with respect to a coordinate of one nucleus in one cell can easily be shown to be the same, to the lowest order in the infinitesimal displacement of the nucleus, as the derivative of the crystal energy *per cell* with respect to a coordinate of the α_{th} nucleus in *every* cell, with all of them displaced by the same infinitesimal distance. The former is the negative of the force on the atom, but the latter is what we can calculate directly. Thus, we differentiate the cellular energy considered before [Eq. (36)] with respect to a component of the nuclear coordinate \mathbf{r}_{α} , again called y . Three kinds of terms result: (i) terms involving derivatives of fitting coefficients such as a_p ; (ii) terms involving nuclear derivatives of integrals such as $C_{ijp}(\mathbf{R})$; and (iii) terms containing nuclear derivatives of the density matrix. Terms of type (i) can be expressed in terms of types (ii) and (iii) by systematically differentiating the expressions for the fitting coefficients. However, when this is done for a_p , its contribution vanishes due to the variational nature of the charge density fitting procedure, as it also does in the molecular case. Since a_p enters the total energy only through E_{coul} , we can write

$$\frac{dE}{dy} = \sum_p \frac{\partial E_{\text{coul}}}{\partial a_p} \frac{da_p}{dy} + \left[\frac{dE}{dy} \right]_{a_p = \text{const}}. \quad (73)$$

In general, $da_p/dy \neq 0$, but when a_p obeys the charge conservation constraint we have $\partial E_{\text{coul}}/\partial a_p = 0$, so the first term in Eq. (73) vanishes, eliminating all fitting coefficient derivatives from the force expression.

The terms of the type (iii) (density matrix derivatives with respect to nuclear coordinates) can be brought together in an expression of the form

$$\begin{aligned} \sum_{ij\mathbf{R}} H_{ij}(\mathbf{R}) \frac{dP_{ij}(\mathbf{R})}{dy} \\ = \sum_{ij\mathbf{R}} \sum_{\mathbf{k}n_{\text{occ}}} w_{\mathbf{k}} H_{ij}(\mathbf{R}) e^{i\mathbf{k}\cdot\mathbf{R}} \left(c_{\mathbf{k}ni}^* \frac{dc_{\mathbf{k}nj}}{dy} + c_{\mathbf{k}nj} \frac{dc_{\mathbf{k}ni}^*}{dy} \right), \end{aligned} \quad (74)$$

where $H_{ij}(\mathbf{R})$ is the real-space representation of the Kohn–Sham effective Hamiltonian given by the term in brackets on the right-hand side of Eq. (29), and we have taken the derivative of Eq. (21) on the right-hand side of Eq. (74). Just as in the molecular case, our goal is to reduce Eq. (74) to a term of type (ii), specifically an orbital overlap derivative, since that is much easier to compute. Using Eq. (29) we transform Eq. (74) to reciprocal space:

$$\begin{aligned} \sum_{ij\mathbf{R}} H_{ij}(\mathbf{R}) \frac{dP_{ij}(\mathbf{R})}{dy} \\ = \sum_{ij} \sum_{\mathbf{k}n_{\text{occ}}} w_{\mathbf{k}} H_{ij}(\mathbf{k}) \left(c_{\mathbf{k}ni}^* \frac{dc_{\mathbf{k}nj}}{dy} + c_{\mathbf{k}nj} \frac{dc_{\mathbf{k}ni}^*}{dy} \right). \end{aligned} \quad (75)$$

Next we use the eigenvalue equation (17) and change the order of summation to obtain

$$\begin{aligned} \sum_{ij\mathbf{R}} H_{ij}(\mathbf{R}) \frac{dP_{ij}(\mathbf{R})}{dy} \\ = \sum_{\mathbf{k}n_{\text{occ}}} w_{\mathbf{k}} \epsilon_{\mathbf{k}n} \sum_{ij} S_{ij}(\mathbf{k}) \left(c_{\mathbf{k}ni}^* \frac{dc_{\mathbf{k}nj}}{dy} + c_{\mathbf{k}nj} \frac{dc_{\mathbf{k}ni}^*}{dy} \right). \end{aligned} \quad (76)$$

To relate the density matrix derivative to the overlap derivative we consider the normalization integral for the wave function of Eqs. (14) and (15),

$$\int d^3r \psi_{kn}^*(\mathbf{r}) \psi_{kn}(\mathbf{r}) = N_{\text{cell}} \sum_{ij\mathbf{R}} e^{i\mathbf{k}\cdot\mathbf{R}} c_{\mathbf{k}ni}^* c_{\mathbf{k}nj} S_{ij}(\mathbf{R}), \quad (77)$$

in which (according to the normalization convention we are using) the integral runs over the whole crystal, N_{cell} is the number of primitive cells in the crystal, and the real-space overlap matrix [compare Eq. (18)] is

$$S_{ij}(\mathbf{R}) = \int d^3r \chi_i(\mathbf{r}) \chi_j(\mathbf{r}-\mathbf{R}). \quad (78)$$

This integral on the left-hand side of (77) is a constant (equal to N_{cell} in our convention) so its derivative with respect to nuclear coordinates is zero. Thus, differentiating the right-hand side of (77) and setting the result equal to zero, we get

$$\begin{aligned} \sum_{ij\mathbf{R}} e^{i\mathbf{k}\cdot\mathbf{R}} S_{ij}(\mathbf{R}) \left(c_{\mathbf{k}ni}^* \frac{dc_{\mathbf{k}nj}}{dy} + c_{\mathbf{k}nj} \frac{dc_{\mathbf{k}ni}^*}{dy} \right) \\ = - \sum_{ij\mathbf{R}} e^{i\mathbf{k}\cdot\mathbf{R}} c_{\mathbf{k}ni}^* c_{\mathbf{k}nj} \frac{dS_{ij}(\mathbf{R})}{dy}, \end{aligned}$$

or, in purely reciprocal space terms,

$$\begin{aligned} \sum_{ij} S_{ij}(\mathbf{k}) \left(c_{\mathbf{k}ni}^* \frac{dc_{\mathbf{k}nj}}{dy} + c_{\mathbf{k}nj} \frac{dc_{\mathbf{k}ni}^*}{dy} \right) \\ = - \sum_{ij} c_{\mathbf{k}ni}^* c_{\mathbf{k}nj} \frac{dS_{ij}(\mathbf{k})}{dy}. \end{aligned} \quad (79)$$

The left-hand side of (79) is identical to the sum on ij on the right-hand side of Eq. (76). Substituting the right-hand side of (79) into (76) and changing the order of sums, we get

$$\begin{aligned} \sum_{ij\mathbf{R}} H_{ij}(\mathbf{R}) \frac{dP_{ij}(\mathbf{R})}{dy} \\ = - \sum_{ij} \sum_{\mathbf{k}n_{\text{occ}}} w_{\mathbf{k}} \epsilon_{\mathbf{k}n} c_{\mathbf{k}ni}^* c_{\mathbf{k}nj} \frac{dS_{ij}(\mathbf{k})}{dy}. \end{aligned} \quad (80)$$

Finally, we use Eq. (18) to reintroduce the real-space overlap matrix of Eq. (78), and define the energy-weighted density matrix

$$P_{ij}^{\epsilon}(\mathbf{R}) = \sum_{\mathbf{k}n_{\text{occ}}} w_{\mathbf{k}} \epsilon_{\mathbf{k}n} e^{i\mathbf{k}\cdot\mathbf{R}} c_{\mathbf{k}ni}^* c_{\mathbf{k}nj}, \quad (81)$$

so that the terms of type (iii) (density matrix derivatives) can all be combined into

$$\sum_{ij\mathbf{R}} H_{ij}(\mathbf{R}) \frac{dP_{ij}(\mathbf{R})}{dy} = - \sum_{ij\mathbf{R}} P_{ij}^{\epsilon}(\mathbf{R}) \frac{dS_{ij}(\mathbf{R})}{dy}. \quad (82)$$

The type (iii) terms thus reduce to terms of type (ii), i.e., derivatives of multicenter integrals. When all the type (ii) terms are brought together, the full expression for the derivative becomes

$$\begin{aligned} \frac{dE}{dy} = \sum_{ij\mathbf{R}} P_{ij}^{\epsilon}(\mathbf{R}) \left\{ \frac{dH_{ij}^{\text{kin}}(\mathbf{R})}{dy} + \sum_p a_p \frac{dC_{ijp}(\mathbf{R})}{dy} \right. \\ \left. + \int V_{\text{xc}}[\rho] \frac{d}{dy} [\chi_i(\mathbf{r}) \chi_j(\mathbf{r}-\mathbf{R})] d^3r \right\} \\ + \sum_p a_p \frac{dT_p}{dy} - \frac{1}{2} \sum_{pp'} a_p \frac{dV_{pp'}}{dy} a_{p'} \\ - \sum_{ij} P_{ij}^{\epsilon}(\mathbf{R}) \frac{dS_{ij}(\mathbf{R})}{dy}. \end{aligned} \quad (83)$$

If, for example, the symbol y stands for the x coordinate of the a th atom in the home cell, then the negative of Eq. (83) gives the x component of the force on that atom. These forces can then be used in geometry optimization, molecular dynamics simulations, calculations of elastic properties, phonon spectra, etc.

We will not consider the actual formulas for the integral derivatives appearing in (83) in the present work, except to

note that the same convergence issues discussed in Sec. III apply to the Coulomb terms. Also, the issue of correction forces arising from the finite accuracy of the numerical integration mesh will not be discussed here. We close this section by noting that the expressions for derivatives given above are also formally correct for systems periodic in two dimensions, provided that, as usual, the lattice and reciprocal vectors are interpreted as two-dimensional vectors and the appropriate form of the Ewald function [Eqs. (68) and (69)] is used. The force on an atom in a slab computed from (83) may, of course, have a nonzero component normal to the plane of the slab.

VI. COMPUTATIONAL COST OF THE METHOD

So far we have shown that the cost of a total energy calculation by the methods described here should scale like N^3 where N is the number of orbital basis functions, provided we fit the electronic density to the auxiliary basis $g_p(\mathbf{r})$ and handle the Coulomb part of the problem correctly as described in Sec. V. We would now like to show how this scaling behavior can be improved considerably if we exploit the sparsity of overlaps and interactions between localized functions. Our discussion will be somewhat general since the details of program implementation are beyond the scope of this paper.

First consider the overlap array $S_{ij}(\mathbf{R})$ of Eq. (78). In a solid with a large unit cell, the i th and j th basis functions in the home cell will often be far enough apart that their overlap is negligible, so that $S_{ij}(\mathbf{0})=0$. Except for a few cases where the centers of the two functions are near the cell boundary on opposite sides, we will also then have $S_{ij}(\mathbf{R})=0$ for all other lattice vectors as well. Thus, of the N^2 pairs i,j we can see that we only have to perform calculations for approximately n_0N of them, where n_0 is the average number of basis functions close enough to significantly overlap with a given basis function, and is independent of unit cell size. Thus, when the unit cell is large, the cost of forming $S_{ij}(\mathbf{k})$ and $H_{ij}^{\text{kin}}(\mathbf{k})$ will scale linearly with cell size. The cost of forming $C_{ijp}(\mathbf{R})$ and $V_{pp'}$ will scale like the square of the unit cell size. If storage permits, all of these quantities can be computed once for each geometry and then reused throughout the SCF process.

The cost of the numerical integrations for the XC potential and energy needs special attention. From a naïve viewpoint the cost of evaluating Eq. (20) for the density would scale like N^2 for each integration grid point. The numerical integration of Eq. (35) requires G evaluations of ρ and V_{XC} for each of the matrix elements needed, where G is the number of points in the integration mesh. By storing the results for ρ and V_{XC} across the grid as vectors of length G , we can avoid repeating their evaluation for each of the n_0N matrix elements, but the cost still scales as N^3 since G scales as N , and with a large coefficient. Some improvement could be achieved by using the fitted density [Eq. (22a)] as the argument of V_{XC} , and also by fitting V_{XC} as in Eq. (23), but with possibly serious losses of accuracy in both cases (and, in the latter case, the additional cost of inverting \mathbf{S}_{XC}). A better approach is to fully exploit the sparsity of the problem. Only

those basis functions within a certain radius of a given mesh point will contribute significantly to the density there given by Eq. (20); hence, for large cells the number of terms that need to be included in (20) becomes independent of cell size. Likewise, for a given pair of basis functions i,j the numerical integration of Eq. (35) only needs to include those mesh points at which the product charge is non-negligible, and the number of such points is also independent of cell size. Thus, the cost of all the numerical integration can be reduced to an amount that scales linearly with the size of the unit cell, and there is no longer a significant cost advantage in fitting V_{XC} . Therefore, we compute its matrix elements by numerical integration of Eq. (35) without resort to fitting procedures.

We see that for sufficiently large unit cells the main computational bottlenecks are the inversion of the matrix $V_{pp'}$ and the diagonalization of $H_{ij}(\mathbf{k})$ [in a generalized eigenproblem with overlap $S_{ij}(\mathbf{k})$]. Both of these operations normally scale like N^3 in cost. However, the number of significantly nonzero elements in $\mathbf{H}(\mathbf{k})$ and $\mathbf{S}(\mathbf{k})$ should again scale as N rather than N^2 , so that the cost of the SCF process can be reduced by one order in N (except for the orthogonalization of the eigenvectors which is still of order N^3 but with a very small coefficient.) It is harder to reduce the cost of inverting²³ \mathbf{V} , but we may be able to find some sparsity in it to exploit, and as already noted it does not change during the SCF process, so it only needs to be computed and inverted once. Also, we note again that the basis set size N is roughly an order of magnitude smaller than in plane wave calculations.

VII. NUMERICAL EXAMPLES

We now consider some numerical results from our computer program GAPSS, which is an implementation of the equations presented above. Since this paper is mainly concerned with problems involving the Coulomb terms in the Hamiltonian and total energy, we will discuss our approach to numerical integration of the XC potential and energy functionals in a separate publication.²⁴ Briefly, however, we decompose the multicenter integrals into sums of one-center numerical integrals on radial and angular grids centered on the atomic nuclei, and adapt the integration points and weights to the functional behavior of the charge density near the nuclei. A standard SCF procedure was used in which the generalized eigenproblem of Eq. (17) was solved at each iteration, and the new charge density [Eq. (20)] was mixed with the preceding one (30% old for the first seven iterations, 55% old for subsequent ones.) The starting electron density was formed from a superposition of isolated atomic densities, and the SCF process²⁵ was stopped when the total energy changed by less than 10^{-6} hartrees, requiring typically 12–15 iterations. For our two test cases, bulk MgO (rocksalt structure) and diamond, the Brillouin zone sampling was equivalent to four special \mathbf{k} points in the irreducible wedge of the zone. We used basis sets derived from those of Godbout *et al.*⁹ by the deletion of exponents below 0.1 a.u.⁻² (to avoid linear dependence problems from long-range interatomic overlap) and of d -type orbital and fitting basis func-

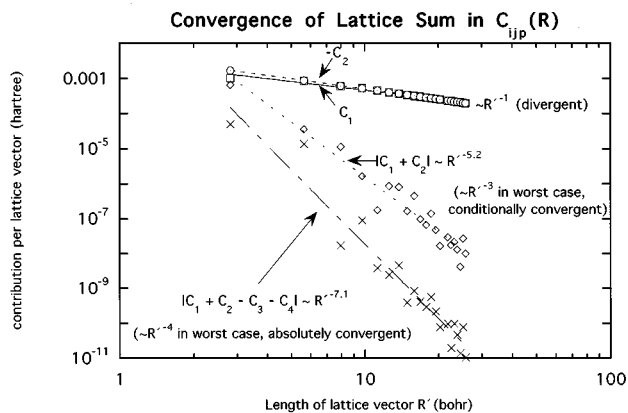


FIG. 1. Convergence test of the two-electron Coulomb series. In this example i represents an oxygen $2s$ orbital basis function, j is a Mg $2s$ orbital function, p is an uncontracted s -type density fitting function on Mg, and $\mathbf{R}=0$.

tions (which we expect to be unnecessary due to the cubic symmetry of these test cases.) For the XC calculation we used the Vosko–Wilk–Nusair²⁶ parameterization of the local density approximation (LDA).

It is instructive to first consider some aspects of the convergence of Coulomb series such as those summed in Eqs. (40) for $C_{ijp}(\mathbf{R})$ and (51) for $V_{pp'}$. In Fig. 1 we show various components,²⁷ e.g., $C_{ijp}^{(1)}(\mathbf{R}, \mathbf{R}')$, and combinations of components versus $|\mathbf{R}'|$. We see that $C^{(1)}$ and $C^{(3)}$ separately decay only as $|\mathbf{R}'|^{-1}$ while their sum $C^{(1)} + C^{(2)}$ decays somewhat faster, but only the fully compensated expression $C^{(1)} + C^{(2)} - C^{(3)} - C^{(4)}$ decays fast enough to allow practical numerical summation. The sum converges faster than its theoretical worst-case behavior due to partial cancellations between \mathbf{R}' vectors in different directions. The exact behavior depends on which basis functions are used, but we find that the sum in Eq. (40) converges fast enough for direct summation in all cases tested thus far. The situation with $V_{pp'}$ is more complex. The sum in Eq. (51) converges exponentially for s - and p -type basis functions, but for d -type fitting functions the $V^{(1)} - V^{(3)}$ part of the sum drops off too slowly to sum directly in a practical calculation. For this term, we have succeeded in restoring exponential convergence by including point dipole–quadrupole and quadrupole–quadrupole interactions²⁸ in Eq. (54), and putting the corresponding terms involving third and fourth derivatives of the Ewald potential function into Eq. (56). The quantity T_p as given by Eq. (59) converges exponentially for s -, p -, and d -type fitting functions alike.

Our results for some physical properties of MgO are shown in Table I. The equilibrium lattice constant and bulk modulus were obtained by calculating the self-consistent ground-state total energy per formula unit (neglecting nuclear zero-point motion) over a range of values of the lattice constant, and fitting the resulting energy-versus-volume curve to the Murnaghan equation. Our LDA values for the optical band gap and valence band width were obtained from the one-electron Kohn–Sham eigenvalues at the Γ point ($\mathbf{k}=0$) in the usual way. Table I shows that the calculated

TABLE I. Physical properties of MgO.

	Present work	Ref. 29	Ref. 30	Experiment ^a
a (Å)	4.1299	4.191	4.167	4.211
B (GPa)	174	146	172	155–162
Gap (eV)	4.73	4.50	5.0	7.77
Valence band width (eV)	17.23	17.14	16.7	21.0

^aCited in Ref. 29.

properties are in good agreement with other well-converged LDA calculations based on *ab initio* pseudopotential plane wave²⁹ and linear augmented plane wave³⁰ techniques. The lattice properties are also in good agreement with experiment,²⁶ while the band width and gap differ from spectroscopic data in ways typical of the LDA, i.e., excitation energies are underestimated.

Our predictions for some physical properties of diamond are shown in Table II. In addition to the properties tabulated for MgO, we have included the indirect gap $\Gamma'_{25v} - \Delta_c$ since it is the lowest gap in diamond. Our results agree very well with other published LDA calculations^{31–33} of lattice and electronic properties, as well as with experimental³⁴ lattice properties. Thus, for diamond, as for MgO, our method leads to results closely consistent with other, in some cases quite different, implementations of density functional theory. A more precise test of our method would require a controlled approach to the limit of basis set completeness, but since such a test would require a substantial effort in basis set design and optimization, it is outside the scope of the present paper. In closing this section, we also briefly note that GAPSS has given physically correct results in preliminary tests on the MgO (100) surface.

VIII. SUMMARY

In this paper we have presented an approach to density functional theory for periodic systems using localized basis functions, particularly Gaussians. A major element is our treatment (in Secs. II and III) of the variationally stationary¹³ and mathematically convergent fitting of the charge density to an auxiliary basis set. This fitting, and related problems in evaluating the Coulomb contributions to the total energy and Hamiltonian, were treated by adapting a real-space Ewald-type method that was originally formulated for periodic Hartree–Fock theory¹⁸ together with the basic idea¹⁷ of mak-

TABLE II. Physical properties of diamond.

	Present work	Ref. 31	Ref. 32	Ref. 33	Experiment ^a
a (Å)	3.5719	3.570	3.602	3.560	3.5667
B (GPa)	431	418	433	437	426
Indirect gap (eV)	4.38	4.18	4.1	5.05	5.48
Direct gap (eV)	5.66	5.8	...	6.02	6.5
Valence band width (eV)	21.83	21.4	...	21.03	24.2

^aCited in Ref. 34.

ing the fitting functions formally charge neutral. We have also discussed the extension of our approach to two-dimensionally periodic systems and to the analytic evaluation of forces on nuclei. Let us now compare the approach to periodic LCAO-DFT described here with several others implemented in the past to highlight the similarities and differences. We limit our discussion to methods designed for Gaussian-type orbitals (GTOs), although numerical basis LCAO schemes³⁵ also have certain points in common with ours.

One of the earliest approaches employing Gaussian basis sets was the “extended tight binding” method of Lin and co-workers³⁶ in which the wave functions were expanded in a modified form of Huzinaga’s³⁷ contracted GTO basis sets, multicenter integrals were done analytically, and a mixed representation (real and reciprocal space) was used for the potential. These early calculations were not self-consistent and yielded no total energies. The first self-consistent periodic LCGTO-DFT method known to us was developed by Callaway and Wang.³⁸ It used contracted Gaussians to expand the wave functions and plane waves to represent the density and potential. Another method developed by Hamann *et al.*³⁹ used a similar orbital basis but employed uncontracted Gaussians (both atom centered and interstitial) to fit the electron density and potential. The density fit was a simple (not variational¹³ in the energy) analytic least squares fit, while the XC potential was fitted numerically on a mesh. The long-ranged parts of the Coulomb interaction were treated in reciprocal space, and the method was readily extended to give accurate total energies and forces. Methods similar to that of Ref. 39, but treating the Coulomb potential as well as V_{XC} on a numerical mesh, were also developed by Mednick and Kleinman⁴⁰ and (in a pseudopotential⁶ framework and without fitting functions) by Schröer *et al.*⁴¹ and also by Chen *et al.*⁴² The approach of Chelikowsky and Louie,³³ on the other hand, used an uncontracted orbital basis, a Gaussian fit of their “input” potential, and a non-self-consistent density functional similar to the Harris⁴³ functional, with total energy evaluated mainly in reciprocal space. We also note that a self-consistent DFT version of the Hartree–Fock code mentioned in Ref. 10 exists.

Another line of development in periodic LCGTO methods leading in a very similar direction to what we propose started with the work of Mintmire *et al.*⁴⁴ in 1982. That work treated two-dimensionally periodic systems with separate uncontracted Gaussian basis sets for the wave functions, density, and XC potential. The density fitting functions were neutralized with nuclear charges as discussed here and in Ref. 17. For two dimensions their formulation required no further special techniques for absolute convergence in the variational¹³ density fit, secular matrix elements, and total energy. Boettger and Trickey⁴⁵ made several refinements to this approach to 2D systems including calculation of V_{XC} from the *fitted* density. At nearly the same time, Callaway *et al.*⁴⁶ published an improved version of their earlier³⁸ LCGTO approach to bulk solids that was also soon further refined by Boettger and Trickey.⁴⁷ This approach did not use a fitting basis and calculated the components of the total

energy mostly in reciprocal space. Ewald techniques were used to compensate for \mathbf{K} -space cutoffs [in contrast to our method which basically uses Ewald techniques to allow a finite cutoff of the *real*-space sums on \mathbf{R}' in Eq. (40) and \mathbf{R} in Eqs. (51) and (59)]. The method of Refs. 47 also used a perturbative technique to update V_{XC} with changes in the density in the SCF process. Very recently these same authors have published a treatment⁴⁸ of bulk aluminum using a method that combines the best features of Refs. 45 and 47 in that it generalizes the variational fitting scheme of Ref. 13 to systems periodic in as many as three dimensions. Since only a general description of this method⁴⁹ has been published to date it is difficult to compare it in detail to ours, but it appears to differ mainly in the use of the fitted density in the XC calculation, the fitting of the XC potential to a second auxiliary basis, and the treatment of only charges and dipoles (not quadrupoles as in our method) by the Ewald convention.

We close by pointing out that with the growth of interdisciplinary fields such as surface science, there is an expanding interest in techniques such as periodic LCAO-DFT which span traditional areas of quantum chemistry and solid-state theory. We hope that our approach will serve to advance theoretical efforts in this area. Applications of our method to a variety of solids and surfaces, together with basis sets optimized for solids and more details of our implementation, will be reported in future publications.

APPENDIX A

The gradient of the Ewald function for systems periodic in three dimensions is, for $\mathbf{d} \neq 0$,

$$\begin{aligned} \nabla \Phi(\mathbf{d}) = & - \sum_{\mathbf{R}} \frac{\mathbf{d} - \mathbf{R}}{|\mathbf{d} - \mathbf{R}|^2} \left\{ \frac{\operatorname{erfc}(|\mathbf{d} - \mathbf{R}|/\eta)}{|\mathbf{d} - \mathbf{R}|} \right. \\ & \left. + \frac{2e^{-(|\mathbf{d} - \mathbf{R}|/\eta)^2}}{\eta\sqrt{\pi}} \right\} \\ & - \frac{4\pi}{V} \sum_{\mathbf{K} \neq 0} \mathbf{K} \frac{\sin(\mathbf{K} \cdot \mathbf{d})}{K^2} e^{-K^2 \eta^2/4}. \end{aligned} \quad (\text{A1})$$

For the second derivatives we obtain in our notation

$$\begin{aligned} \frac{\partial^2 \Phi(\mathbf{d})}{\partial d_\mu \partial d_\nu} = & \sum_{\mathbf{R}} \left\{ \frac{-\delta_{\mu\nu}}{|\mathbf{d} - \mathbf{R}|^2} \left[\frac{\operatorname{erfc}(|\mathbf{d} - \mathbf{R}|/\eta)}{|\mathbf{d} - \mathbf{R}|} \right. \right. \\ & \left. \left. + \frac{2e^{-(|\mathbf{d} - \mathbf{R}|/\eta)^2}}{\eta\sqrt{\pi}} \right] \right. \\ & \left. + \frac{(\mathbf{d} - \mathbf{R})_\mu (\mathbf{d} - \mathbf{R})_\nu}{|\mathbf{d} - \mathbf{R}|^2} \left[\frac{3 \operatorname{erfc}(|\mathbf{d} - \mathbf{R}|/\eta)}{|\mathbf{d} - \mathbf{R}|^3} \right. \right. \\ & \left. \left. + \frac{2e^{-(|\mathbf{d} - \mathbf{R}|/\eta)^2}}{\eta\sqrt{\pi}} \left(\frac{3}{|\mathbf{d} - \mathbf{R}|^2} + \frac{2}{\eta^2} \right) \right] \right\} \\ & - \frac{4\pi}{V} \sum_{\mathbf{K} \neq 0} K_\mu K_\nu \frac{\cos(\mathbf{K} \cdot \mathbf{d})}{K^2} e^{-K^2 \eta^2/4}. \end{aligned} \quad (\text{A2})$$

The third derivatives (needed in force calculations) are

$$\begin{aligned} \frac{\partial^3 \Phi(\mathbf{d})}{\partial d_\mu \partial d_\nu \partial d_\lambda} = \sum_{\mathbf{R}} \left\{ \left[\frac{\delta_{\mu\nu}(\mathbf{d}-\mathbf{R})_\lambda + \delta_{\mu\lambda}(\mathbf{d}-\mathbf{R})_\nu + \delta_{\nu\lambda}(\mathbf{d}-\mathbf{R})_\mu}{|\mathbf{d}-\mathbf{R}|^2} - \frac{5(\mathbf{d}-\mathbf{R})_\mu(\mathbf{d}-\mathbf{R})_\nu(\mathbf{d}-\mathbf{R})_\lambda}{|\mathbf{d}-\mathbf{R}|^4} \right] \left[\frac{3 \operatorname{erfc}(|\mathbf{d}-\mathbf{R}|/\eta)}{|\mathbf{d}-\mathbf{R}|^3} \right. \right. \\ \left. \left. + \frac{2e^{-(|\mathbf{d}-\mathbf{R}|/\eta)^2}}{\eta\sqrt{\pi}} \left(\frac{3}{|\mathbf{d}-\mathbf{R}|^2} + \frac{2}{\eta^2} \right) \right] - \frac{8(\mathbf{d}-\mathbf{R})_\mu(\mathbf{d}-\mathbf{R})_\nu(\mathbf{d}-\mathbf{R})_\lambda}{|\mathbf{d}-\mathbf{R}|^2} \frac{e^{-(|\mathbf{d}-\mathbf{R}|/\eta)^2}}{\eta^5\sqrt{\pi}} \right\} \\ + \frac{4\pi}{V} \sum_{\mathbf{K} \neq 0} K_\mu K_\nu K_\lambda \frac{\sin(\mathbf{K} \cdot \mathbf{d})}{K^2} e^{-K^2 \eta^2/4}. \end{aligned} \quad (\text{A3})$$

For systems periodic in only two dimensions the Ewald function is given by (68) and (69) with first derivatives

$$\begin{aligned} \frac{\partial \Phi_1(\mathbf{d})}{\partial d_\mu} &= - \sum_{\mathbf{R}} \frac{(\mathbf{d}-\mathbf{R})_\mu}{|\mathbf{d}-\mathbf{R}|^2} \left[\frac{\operatorname{erfc}(|\mathbf{d}-\mathbf{R}|/\eta)}{|\mathbf{d}-\mathbf{R}|} \right. \\ &\quad \left. + \frac{2e^{-(|\mathbf{d}-\mathbf{R}|/\eta)^2}}{\eta\sqrt{\pi}} \right], \\ \frac{\partial \Phi_1(\mathbf{d})}{\partial d_z} &= - \sum_{\mathbf{R}} \frac{d_z}{|\mathbf{d}-\mathbf{R}|^2} \left[\frac{\operatorname{erfc}(|\mathbf{d}-\mathbf{R}|/\eta)}{|\mathbf{d}-\mathbf{R}|} \right. \\ &\quad \left. + \frac{2e^{-(|\mathbf{d}-\mathbf{R}|/\eta)^2}}{\eta\sqrt{\pi}} \right], \end{aligned} \quad (\text{A4})$$

$$\begin{aligned} \frac{\partial \Phi_2(\mathbf{d})}{\partial d_\mu} &= - \frac{\pi}{A} \sum_{\mathbf{K} \neq 0} \frac{K_\mu \sin(\mathbf{K} \cdot \mathbf{d})}{K} \\ &\quad \times \left[e^{Kd_z} \operatorname{erfc}\left(\frac{K\eta}{2} + \frac{d_z}{\eta}\right) \right. \\ &\quad \left. + e^{-Kd_z} \operatorname{erfc}\left(\frac{K\eta}{2} - \frac{d_z}{\eta}\right) \right], \\ \frac{\partial \Phi_2(\mathbf{d})}{\partial d_z} &= \frac{\pi}{A} \sum_{\mathbf{K} \neq 0} \cos(\mathbf{K} \cdot \mathbf{d}) \left[e^{Kd_z} \operatorname{erfc}\left(\frac{K\eta}{2} + \frac{d_z}{\eta}\right) \right. \\ &\quad \left. - e^{-Kd_z} \operatorname{erfc}\left(\frac{K\eta}{2} - \frac{d_z}{\eta}\right) \right], \end{aligned} \quad (\text{A5})$$

and

$$\frac{\partial \Phi_3(\mathbf{d})}{\partial d_z} = - \frac{2\pi}{A} \operatorname{erf}\left(\frac{d_z}{\eta}\right), \quad \frac{\partial \Phi_3(\mathbf{d})}{\partial d_\mu} = 0, \quad (\text{A6})$$

where $\mu=x$ or y . The second derivatives in two dimensions are

$$\begin{aligned} \frac{\partial^2 \Phi_1(\mathbf{d})}{\partial d_\mu \partial d_\nu} &= \sum_{\mathbf{R}} \left\{ \frac{-\delta_{\mu\nu}}{|\mathbf{d}-\mathbf{R}|^2} \left[\frac{\operatorname{erfc}(|\mathbf{d}-\mathbf{R}|/\eta)}{|\mathbf{d}-\mathbf{R}|} \right. \right. \\ &\quad \left. \left. + \frac{2e^{-(|\mathbf{d}-\mathbf{R}|/\eta)^2}}{\eta\sqrt{\pi}} \right] + \frac{(\mathbf{d}-\mathbf{R})_\mu(\mathbf{d}-\mathbf{R})_\nu}{|\mathbf{d}-\mathbf{R}|^2} \right. \\ &\quad \times \left[\frac{3 \operatorname{erfc}(|\mathbf{d}-\mathbf{R}|/\eta)}{|\mathbf{d}-\mathbf{R}|^3} + \frac{2e^{-(|\mathbf{d}-\mathbf{R}|/\eta)^2}}{\eta\sqrt{\pi}} \right. \\ &\quad \left. \left. \times \left(\frac{3}{|\mathbf{d}-\mathbf{R}|^2} + \frac{2}{\eta^2} \right) \right] \right\}, \end{aligned} \quad (\text{A7})$$

$$\begin{aligned} \frac{\partial^2 \Phi_1(\mathbf{d})}{\partial d_\mu \partial d_z} &= \sum_{\mathbf{R}} \frac{(\mathbf{d}-\mathbf{R})_\mu d_z}{|\mathbf{d}-\mathbf{R}|^2} \left[\frac{3 \operatorname{erfc}(|\mathbf{d}-\mathbf{R}|/\eta)}{|\mathbf{d}-\mathbf{R}|^3} \right. \\ &\quad \left. + \frac{2e^{-(|\mathbf{d}-\mathbf{R}|/\eta)^2}}{\eta\sqrt{\pi}} \left(\frac{3}{|\mathbf{d}-\mathbf{R}|^2} + \frac{2}{\eta^2} \right) \right], \end{aligned} \quad (\text{A8})$$

$$\begin{aligned} \frac{\partial^2 \Phi_1(\mathbf{d})}{\partial d_z^2} &= \sum_{\mathbf{R}} \frac{1}{|\mathbf{d}-\mathbf{R}|^2} \left\{ \left(\frac{3d_z^2}{|\mathbf{d}-\mathbf{R}|^2} - 1 \right) \right. \\ &\quad \times \frac{\operatorname{erfc}(|\mathbf{d}-\mathbf{R}|/\eta)}{|\mathbf{d}-\mathbf{R}|} \\ &\quad \left. + \left[d_z^2 \left(\frac{3}{|\mathbf{d}-\mathbf{R}|^2} + \frac{2}{\eta^2} \right) - 1 \right] \right. \\ &\quad \left. \times \frac{2e^{-(|\mathbf{d}-\mathbf{R}|/\eta)^2}}{\eta\sqrt{\pi}} \right\}, \end{aligned} \quad (\text{A9})$$

$$\begin{aligned} \frac{\partial^2 \Phi_2(\mathbf{d})}{\partial d_\mu \partial d_\nu} &= - \frac{\pi}{A} \sum_{\mathbf{K} \neq 0} \frac{K_\mu K_\nu \cos(\mathbf{K} \cdot \mathbf{d})}{K} \\ &\quad \times \left[e^{Kd_z} \operatorname{erfc}\left(\frac{K\eta}{2} + \frac{d_z}{\eta}\right) \right. \\ &\quad \left. + e^{-Kd_z} \operatorname{erfc}\left(\frac{K\eta}{2} - \frac{d_z}{\eta}\right) \right] \end{aligned}$$

$$\begin{aligned} \frac{\partial^2 \Phi_2(\mathbf{d})}{\partial d_\mu \partial d_z} &= - \frac{\pi}{A} \sum_{\mathbf{K} \neq 0} K_\mu \sin(\mathbf{K} \cdot \mathbf{d}) \left[e^{Kd_z} \operatorname{erfc}\left(\frac{K\eta}{2} + \frac{d_z}{\eta}\right) \right. \\ &\quad \left. - e^{-Kd_z} \operatorname{erfc}\left(\frac{K\eta}{2} - \frac{d_z}{\eta}\right) \right] \end{aligned}$$

$$\begin{aligned} \frac{\partial^2 \Phi_2(\mathbf{d})}{\partial d_z^2} &= \frac{\pi}{A} \sum_{\mathbf{K} \neq 0} \cos(\mathbf{K} \cdot \mathbf{d}) \left\{ K \left[e^{Kd_z} \operatorname{erfc}\left(\frac{K\eta}{2} + \frac{d_z}{\eta}\right) \right. \right. \\ &\quad \left. \left. + e^{-Kd_z} \operatorname{erfc}\left(\frac{K\eta}{2} - \frac{d_z}{\eta}\right) \right] \right. \\ &\quad \left. - \frac{4}{\sqrt{\pi}\eta} e^{-K^2 \eta^2/4} e^{-(d_z/\eta)^2} \right\}, \end{aligned} \quad (\text{A10})$$

and

$$\frac{\partial^2 \Phi_3(\mathbf{d})}{\partial d_z^2} = -\frac{4\sqrt{\pi}}{A\eta} e^{-(d_z/\eta)^2}, \quad \frac{\partial^2 \Phi_3(\mathbf{d})}{\partial d_\mu \partial d_\nu} = \frac{\partial^2 \Phi_3(\mathbf{d})}{\partial d_\mu \partial d_z} = 0. \quad (\text{A11})$$

Expressions for the third derivatives of $\Phi_1(\mathbf{d})$ can be taken from the real-space sum in Eq. (A3) by setting, when needed, $R_z=0$ and $\delta_{\mu z}=\delta_{\nu z}=0$. For the reciprocal space part we have

$$\begin{aligned} \frac{\partial^3 \Phi_2(\mathbf{d})}{\partial d_\mu \partial d_\nu \partial d_\lambda} &= \frac{\pi}{A} \sum_{\mathbf{K} \neq 0} \frac{K_\mu K_\nu K_\lambda \sin(\mathbf{K} \cdot \mathbf{d})}{K} \\ &\quad \times \left[e^{Kd_z} \operatorname{erfc}\left(\frac{K\eta}{2} + \frac{d_z}{\eta}\right) \right. \\ &\quad \left. + e^{-Kd_z} \operatorname{erfc}\left(\frac{K\eta}{2} - \frac{d_z}{\eta}\right) \right], \\ \frac{\partial^3 \Phi_2(\mathbf{d})}{\partial d_\mu \partial d_\nu \partial d_z} &= -\frac{\pi}{A} \sum_{\mathbf{K} \neq 0} K_\mu K_\nu \cos(\mathbf{K} \cdot \mathbf{d}) \\ &\quad \times \left[e^{Kd_z} \operatorname{erfc}\left(\frac{K\eta}{2} + \frac{d_z}{\eta}\right) \right. \\ &\quad \left. - e^{-Kd_z} \operatorname{erfc}\left(\frac{K\eta}{2} - \frac{d_z}{\eta}\right) \right], \\ \frac{\partial^3 \Phi_2(\mathbf{d})}{\partial d_\mu \partial d_z^2} &= -\frac{\pi}{A} \sum_{\mathbf{K} \neq 0} K_\mu \sin(\mathbf{K} \cdot \mathbf{d}) \left\{ K \left[e^{Kd_z} \operatorname{erfc}\left(\frac{K\eta}{2} \right. \right. \right. \\ &\quad \left. \left. + \frac{d_z}{\eta} \right) + e^{-Kd_z} \operatorname{erfc}\left(\frac{K\eta}{2} - \frac{d_z}{\eta} \right) \right] \\ &\quad \left. - \frac{4}{\sqrt{\pi}\eta} e^{-K^2\eta^2/4} e^{-(d_z/\eta)^2} \right\}, \\ \frac{\partial^3 \Phi_2(\mathbf{d})}{\partial d_z^3} &= \frac{\pi}{A} \sum_{\mathbf{K} \neq 0} \cos(\mathbf{K} \cdot \mathbf{d}) \left\{ K^2 \left[e^{Kd_z} \operatorname{erfc}\left(\frac{K\eta}{2} + \frac{d_z}{\eta} \right) \right. \right. \\ &\quad \left. \left. - e^{-Kd_z} \operatorname{erfc}\left(\frac{K\eta}{2} - \frac{d_z}{\eta} \right) \right] \right. \\ &\quad \left. + \frac{8d_z}{\sqrt{\pi}\eta^3} e^{-K^2\eta^2/4} e^{-(d_z/\eta)^2} \right\}, \end{aligned} \quad (\text{A12})$$

and finally the third derivatives of $\Phi_3(\mathbf{d})$ are all zero except for

$$\frac{\partial^3 \Phi_3(\mathbf{d})}{\partial d_z^3} = \frac{8\sqrt{\pi}d_z}{A\eta^3} e^{-(d_z/\eta)^2}. \quad (\text{A13})$$

APPENDIX B

When the displacement \mathbf{d} appearing in the Ewald potential or its derivatives is zero, we do not subtract [in the associated point interactions such as Eq. (45)] nor add (in the Ewald term) the singular $1/d$ terms. However, it is not correct to simply omit the $\mathbf{R}=0$ term from the direct lattice sum

in Eq. (50) or (69), rather we expand this term in the limit $d \rightarrow 0$ and remove only the singular $1/d$ part, obtaining

$$\Phi(0) = \sum_{\mathbf{R} \neq 0} \frac{\operatorname{erfc}(R/\eta)}{R} + \frac{4\pi}{V} \sum_{\mathbf{K} \neq 0} \frac{e^{-K^2\eta^2/4}}{K^2} - \frac{2}{\sqrt{\pi}\eta} - \frac{\pi\eta^2}{V}. \quad (\text{B1})$$

The gradient of the Ewald potential at $\mathbf{d}=0$ is

$$\nabla \Phi(0) = \sum_{\mathbf{R} \neq 0} \frac{\mathbf{R}}{R^2} \left[\frac{\operatorname{erfc}(R/\eta)}{R} + \frac{2e^{-R^2/\eta^2}}{\sqrt{\pi}\eta} \right] = 0, \quad (\text{B2})$$

where we use the fact that the Bravais lattice has inversion symmetry (even if the full crystal structure does not). The expression for the second derivatives at $\mathbf{d}=0$ is

$$\begin{aligned} \frac{\partial^2 \Phi(0)}{\partial d_\mu \partial d_\nu} &= \sum_{\mathbf{R} \neq 0} \frac{1}{R^2} \left\{ -\delta_{\mu\nu} \left[\frac{\operatorname{erfc}(R/\eta)}{R} + \frac{2e^{-(R/\eta)^2}}{\eta\sqrt{\pi}} \right] \right. \\ &\quad \left. + R_\mu R_\nu \left[\frac{3 \operatorname{erfc}(R/\eta)}{R^3} + \frac{2e^{-(R/\eta)^2}}{\eta\sqrt{\pi}} \right. \right. \\ &\quad \left. \left. \times \left(\frac{3}{R^2} + \frac{2}{\eta^2} \right) \right] \right\} - \frac{4\pi}{V} \sum_{\mathbf{K} \neq 0} \frac{K_\mu K_\nu}{K^2} e^{-K^2\eta^2/4} \\ &\quad + \frac{4\eta^{-3}}{3\sqrt{\pi}} \delta_{\mu\nu}. \end{aligned} \quad (\text{B3})$$

Note that R_μ , etc. are Cartesian components of lattice vectors, not lattice components, e.g., $R_x \neq n_1|\mathbf{a}_1|$ in general. Also, the third derivatives all vanish at $\mathbf{d}=0$ by symmetry.

In the limit $\mathbf{d}=0$ the 2D Ewald potential becomes [combining all three terms from Eq. (69)]

$$\Phi(0) = \sum_{\mathbf{R} \neq 0} \frac{\operatorname{erfc}(R/\eta)}{R} - \frac{2}{\sqrt{\pi}\eta} + \frac{2\pi}{A} \sum_{\mathbf{K} \neq 0} \frac{\operatorname{erfc}(K\eta/2)}{K} - \frac{2\sqrt{\pi}\eta}{A}. \quad (\text{B4})$$

Also

$$\frac{\partial \Phi(0)}{\partial d_\mu} = \frac{\partial \Phi(0)}{\partial d_z} = \frac{\partial^2 \Phi(0)}{\partial d_\mu \partial d_z} = 0 \quad (\text{B5})$$

and

$$\begin{aligned} \frac{\partial^2 \Phi(0)}{\partial d_z^2} &= -\sum_{\mathbf{R} \neq 0} \frac{1}{R^2} \left[\frac{\operatorname{erfc}(R/\eta)}{R} + \frac{2}{\sqrt{\pi}\eta} e^{-(R/\eta)^2} \right] \\ &\quad + \frac{4}{3\sqrt{\pi}\eta^3} + \frac{\pi}{A} \sum_{\mathbf{K} \neq 0} \left[2K \operatorname{erfc}(K\eta/2) \right. \\ &\quad \left. - \frac{4}{\sqrt{\pi}\eta} e^{-K^2\eta^2/4} \right] - \frac{4\sqrt{\pi}}{\eta A}, \end{aligned} \quad (\text{B6})$$

and finally

$$\begin{aligned} \frac{\partial^2 \Phi(0)}{\partial d_\mu \partial d_\nu} = & \sum_{\mathbf{R} \neq 0} \left\{ -\frac{\delta_{\mu\nu}}{R^2} \left[\frac{\operatorname{erfc}(R/\eta)}{R} + \frac{2}{\sqrt{\pi}\eta} e^{-(R/\eta)^2} \right] \right. \\ & + \frac{R_\mu R_\nu}{R^2} \left[\frac{3 \operatorname{erfc}(R/\eta)}{R^3} + \frac{2}{\sqrt{\pi}\eta} \left(\frac{3}{R^2} + \frac{2}{\eta^2} \right) \right. \\ & \left. \left. \times e^{-(R/\eta)^2} \right] \right\} + \frac{4\delta_{\mu\nu}}{3\sqrt{\pi}\eta^3} \\ & - \frac{2\pi}{A} \sum_{\mathbf{K} \neq 0} K_\mu K_\nu \frac{\operatorname{erfc}(K\eta/2)}{K}. \quad (\text{B7}) \end{aligned}$$

Again, the third derivatives vanish at $\mathbf{d}=0$.

- ¹P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964); W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).
- ²R. O. Jones and O. Gunnarsson, Rev. Mod. Phys. **61**, 689 (1989).
- ³For a recent review, see *Density Functional Methods in Chemistry*, edited by J. K. Labanowski and J. W. Andzelm (Springer-Verlag, New York, 1991).
- ⁴E. P. Wigner, Trans. Faraday Soc. **34**, 678 (1938); L. Hedin and S. Lundqvist, J. Phys. (France) **33**, C3-73 (1972); D. M. Ceperley and B. J. Alder, Phys. Rev. Lett. **45**, 566 (1980).
- ⁵A. D. Becke, J. Chem. Phys. **98**, 5648 (1993); J. P. Perdew and W. Yue, Phys. Rev. B **33**, 8800 (1986); J. P. Perdew, Phys. Rev. B **33**, 8822 (1986); L. C. Wilson and M. Levy, Phys. Rev. B **41**, 12930 (1990).
- ⁶D. R. Hamann, M. Schlüter, and C. Chiang, Phys. Rev. Lett. **43**, 1494 (1979).
- ⁷E. Wimmer, H. Krakauer, M. Weinert, and A. J. Freeman, Phys. Rev. B **24**, 864 (1981).
- ⁸P. Bendt and A. Zunger, Phys. Rev. B **26**, 3114 (1982).
- ⁹N. Godbout, D. R. Salahub, J. Andzelm, and E. Wimmer, Can. J. Chem. **70**, 560 (1992); J. Andzelm, E. Radzio, and D. R. Salahub, J. Comp. Chem. **6**, 520 (1985); K. Eichkorn, O. Treutler, H. Öhm, M. Häser, and R. Ahlrichs, Chem. Phys. Lett. **240**, 283 (1995). These and other molecular basis sets require modification for use in solids. Gaussians with exponents smaller than 0.1 a.u.⁻² typically have to be deleted, and exponents smaller than about 0.5 a.u.⁻² need to be reoptimized. For an example of a study of Gaussian basis set effects in periodic DFT, see J. Z. Wu, S. B. Trickey, J. R. Sabin, and J. C. Boettger, Phys. Rev. B **51**, 14576 (1995).
- ¹⁰R. Dovesi, C. Roetti, C. Freyria Fava, E. Aprà, V. R. Saunders, and N. M. Harrison, Philos. Trans. R. Soc. **341**, 203 (1992), and references therein. These papers deal with the reoptimization of molecular basis sets for periodic calculations in the context of Hartree-Fock theory, but similar considerations will apply in our case as far as the orbital basis is concerned. See Ref. 9 for discussions of density fitting basis sets.
- ¹¹Another problem with this approach is that localized functions may be a poor basis for conduction band states, except possibly those just above the gap in an insulator or semiconductor.
- ¹²These corrections are often called "Pulay forces" after their original derivation in the molecular Hartree-Fock context; see P. Pulay, Mol. Phys. **17**, 197 (1969). See also B. I. Dunlap and J. Andzelm, Phys. Rev. A **45**, 81 (1992), A. Komornicki and G. Fitzgerald, J. Chem. Phys. **98**, 1398 (1993), and references therein.
- ¹³B. I. Dunlap, J. W. D. Connolly, and J. R. Sabin, J. Chem. Phys. **71**, 3396 (1979).
- ¹⁴For a general discussion of crystal lattices and single-electron states in crystals, see N. W. Ashcroft and N. D. Mermin, *Solid State Physics* (Saunders College, Philadelphia, 1976), Chaps. 4–8.
- ¹⁵Strictly speaking, if the crystal is of macroscopic but still finite size then the \mathbf{k} points are a dense mesh (the Born-Von Karman points) rather than continuous. In practice, however, the sum over these points is always formally converted to an integral (see Ref. 14) which is then approximated by a sum on a coarser mesh (see Ref. 16 below.) Also, it is customary to normalize the eigenfunctions over the volume of the whole crystal rather than over a single unit cell as we have done, but the effect of this difference is absorbed into the sum on \mathbf{k} when we adopt our normalization convention $\sum_{\mathbf{k}} w_{\mathbf{k}} = 1$.
- ¹⁶H. J. Monkhorst and J. D. Pack, Phys. Rev. B **13**, 5188 (1976).

- ¹⁷Such a technique is discussed by J. W. Mintmire (Ref. 3, *op. cit.*, p. 125) in the context of one-dimensional periodicity (polymers) where the additional convergence techniques described in our Sec. III are unnecessary.
- ¹⁸S. Hammes-Schiffer and H. C. Andersen, J. Chem. Phys. **101**, 375 (1994).
- ¹⁹P. P. Ewald, Ann. Phys. **64**, 253 (1921).
- ²⁰V. R. Saunders, C. Freyria-Fava, R. Dovesi, L. Salasco, and C. Roetti, Mol. Phys. **77**, 629 (1992).
- ²¹D. E. Parry, Surf. Sci. **49**, 433 (1975); **54**, 195 (1976). Parry derives these expressions from the three-dimensional Ewald function by taking the limit of a periodic stack of slabs becoming infinitely separated. We have also derived them (J. E. Jaffe, unpublished) directly from first principles, but Parry's limiting process has the advantage that it carries over the rigorous justification of the Ewald potential (Refs. 18–20) from three- to two-dimensional periodicity.
- ²²R. Car and M. Parinello, Phys. Rev. Lett. **55**, 2471 (1985).
- ²³In some implementations, such as on massively parallel computers, it may be desirable to recast the inversion of the matrix \mathbf{V} as a solution of a system of linear equations (both problems scale as the cube of the basis set size but the latter parallelizes better.) This can be done as follows: define a vector \mathbf{m} as the solution of the linear system $\sum_p V_{pp'} m_{p'} = n_p$. Then the Lagrange multiplier λ of Eq. (28) is given by $\lambda = (Q - \sum_p m_p t_p) / \sum_p m_p n_p$ and the fitting coefficients a_p of Eq. (27) are found by solution of the linear system $\sum_p V_{pp'} a_{p'} = t_p + \lambda n_p$.
- ²⁴Z. Lin and A. C. Hess (in preparation).
- ²⁵This simple SCF process was chosen mainly for its ease of programming, since the code GAPSS is still under development. Our general formalism is readily compatible with more sophisticated convergence schemes such as conjugate gradients. See M. C. Payne, M. P. Teter, D. C. Allan, T. A. Arias, and J. D. Joannopoulos, Rev. Mod. Phys. **64**, 1045 (1992) and references therein.
- ²⁶S. H. Vosko, L. Wilk, and M. Nusair, Can. J. Phys. **58**, 1200 (1980).
- ²⁷In the test case shown, i and j were 2s contracted orbital basis functions on Mg and O respectively in bulk MgO, $\mathbf{R}=0$ and p was a diffuse s -type fitting function on Mg.
- ²⁸Recently we have found that including higher-order interactions of this type in Eqs. (45) and (48) can also greatly improve the practical convergence of Eq. (40).
- ²⁹K. J. Chang and M. L. Cohen, Phys. Rev. B **30**, 4774 (1984). This paper also contains references to the experimental data.
- ³⁰M. J. Mehl, R. E. Cohen, and H. Krakauer, J. Geophys. Res. **93**, 8009 (1988).
- ³¹D. R. Hamann, Phys. Rev. B **51**, 9508 (1995).
- ³²M. T. Yin and M. L. Cohen, Phys. Rev. B **24**, 6121 (1981).
- ³³J. R. Chelikowsky and S. G. Louie, Phys. Rev. B **29**, 3470 (1984).
- ³⁴*Semiconductors: Group IV Elements and III-V Compounds*, edited by O. Madelung (Springer-Verlag, Berlin, 1991) pp. 5–7, and references therein.
- ³⁵A. Zunger and A. J. Freeman, Phys. Rev. B **15**, 4716 (1977); B. Delley, *DSolid User Documentation* (Biosym Corporation, 1995); (in preparation).
- ³⁶R. C. Chaney, C. C. Lin, and E. E. Lafon, Phys. Rev. B **3**, 459 (1971), and references therein.
- ³⁷S. Huzinaga, J. Chem. Phys. **42**, 1243 (1965).
- ³⁸J. Callaway and C. S. Wang, Phys. Rev. B **7**, (1973); C. S. Wang and J. Callaway, Comput. Phys. Commun. **14**, 327 (1978).
- ³⁹J. A. Applebaum and D. R. Hamann, Solid State Commun. **27**, 881 (1978); P. J. Feibelman, J. A. Applebaum, and D. R. Hamann, Phys. Rev. B **20**, 1433 (1979); B. N. Harmon, W. Weber, and D. R. Hamann, *ibid.* **25**, 1109 (1982). This approach is a periodic generalization of the molecular method of H. Sambe and R. H. Felton, J. Chem. Phys. **62**, 1122 (1975).
- ⁴⁰K. Mednick and L. Kleinman, Phys. Rev. B **22**, 5768 (1980); D. M. Bylander and L. Kleinman, *ibid.* **27**, 3152 (1983).
- ⁴¹P. Schröer, P. Krüger, and J. Pollmann, Phys. Rev. B **47**, 6971 (1993).
- ⁴²X. Chen, J.-M. Langlois, and W. A. Goddard, Phys. Rev. B **52**, 2348 (1995); X. Chen, X. Hua, J. Hu, J.-M. Langlois, and W. A. Goddard, Phys. Rev. B **53**, 1377 (1996).
- ⁴³J. Harris, Phys. Rev. B **31**, 1770 (1985).
- ⁴⁴J. W. Mintmire, J. R. Sabin, and S. B. Trickey, Phys. Rev. B **26**, 1743 (1982).
- ⁴⁵J. C. Boettger and S. B. Trickey, Phys. Rev. B **32**, 1356 (1985).
- ⁴⁶J. Callaway, X. Zou, and D. Bagayoko, Phys. Rev. B **27**, 631 (1983).
- ⁴⁷J. C. Boettger and S. B. Trickey, Phys. Rev. B **29**, 6425, 6434 (1985).
- ⁴⁸J. C. Boettger and S. B. Trickey, Phys. Rev. B **51**, 15623 (1995).
- ⁴⁹J. C. Boettger, Int. J. Quantum Chem. Symp. **29**, 197 (1995).