

Kohn-Sham density-functional theory within a finite basis set

John A. Pople, Peter M.W. Gill and Benny G. Johnson

Department of Chemistry, Carnegie Mellon University, Pittsburgh, PA 15312, USA

Received 3 August 1992; in final form 14 September 1992

The Kohn-Sham self-consistent equations, using a finite orbital basis expansion, are formulated for exchange-correlation functionals which depend on local densities and their gradients. It is shown that these can be solved iteratively without evaluation of density Hessians. A general expression is given for the energy gradient (with respect to nuclear motion) after self-consistency has been achieved.

In the Kohn-Sham formulation of density functional theory (DFT) [1-3], the electronic energy is separated into parts,

$$E = E^T + E^V + E^J + E^{XC}. \quad (1)$$

Here E^T and E^V are the kinetic and electron-nuclear interaction energies, E^J is the Coulomb self-interaction of the electron density ρ and E^{XC} is the remaining (exchange-correlation) part of the electron-electron repulsion energy, also treated as a functional of the density ρ . Adopting a spin-unrestricted format, α - and β -electrons are assigned to sets of orthonormal orbitals ψ_i^α ($i=1, \dots, n_\alpha$) with ψ_i^β ($i=1, \dots, n_\beta$), respectively. Corresponding α -, β - and total densities are

$$\rho_\alpha = \sum_i^{n_\alpha} |\psi_i^\alpha|^2, \quad \rho_\beta = \sum_i^{n_\beta} |\psi_i^\beta|^2, \quad (2)$$

$$\rho = \rho_\alpha + \rho_\beta. \quad (3)$$

E^T , E^V and E^J are given by

$$E^T = \sum_i^{n_\alpha} (\psi_i^\alpha | -\frac{1}{2}\nabla^2 | \psi_i^\alpha) + \sum_i^{n_\beta} (\psi_i^\beta | -\frac{1}{2}\nabla^2 | \psi_i^\beta), \quad (4)$$

$$E^V = - \sum_A^{\text{nucl}} Z_A \int \rho(r) |r - r_A|^{-1} dr, \quad (5)$$

$$E^J = \frac{1}{2} \iint \rho(r_1) |r_1 - r_2|^{-1} \rho(r_2) dr_1 dr_2. \quad (6)$$

Once an approximate form for the exchange-correlation functional is specified, the Kohn-Sham differential equations for the orbitals ψ_i^α , ψ_i^β are obtained by minimization of the total energy (1).

Many approximate functionals $E^{XC}[\rho]$ have the local ^{#1} gradient-corrected form [4-13]

$$E^{XC} = \int f(\rho_\alpha, \rho_\beta, \gamma_{\alpha\alpha}, \gamma_{\alpha\beta}, \gamma_{\beta\beta}) dr, \quad (7)$$

$$\gamma_{\alpha\alpha} = |\nabla\rho_\alpha|^2, \quad \gamma_{\alpha\beta} = \nabla\rho_\alpha \cdot \nabla\rho_\beta,$$

$$\gamma_{\beta\beta} = |\nabla\rho_\beta|^2, \quad (8)$$

where f is a function only of the local ρ_α , ρ_β and their gradient invariants. The one-electron potentials corresponding to (7) can be obtained by the calculus of variations [3] and are

$$V_\alpha^{XC} = \frac{\partial f}{\partial \rho_\alpha} - 2\nabla \cdot \left(\frac{\partial f}{\partial \gamma_{\alpha\alpha}} \nabla \rho_\alpha \right) - \nabla \cdot \left(\frac{\partial f}{\partial \gamma_{\alpha\beta}} \nabla \rho_\beta \right), \quad (9)$$

the expression for V_β^{XC} being similar. This expression can be developed further but here we only note

^{#1} We will describe as "local" any functional which can be written as the expectation value of a one-electron operator. Although this definition includes gradient-corrected functionals, which have traditionally been termed "non-local" in the DFT literature, we prefer to restrict the term "non-local" to intrinsically multi-electron functionals such as (6).

Correspondence to: P.M.W. Gill, Department of Chemistry, Carnegie Mellon University, Pittsburgh, PA 15312, USA.

that second-derivative (Hessian) matrices of ρ_α and ρ_β are involved.

In practical computations, as in Hartree-Fock theory, it is convenient to write the orbitals as finite expansions in a basis set ϕ_μ ($\mu=1, \dots, N$)

$$\psi_i^\alpha = \sum_\mu c_{\mu i}^\alpha \phi_\mu, \quad \psi_i^\beta = \sum_\mu c_{\mu i}^\beta \phi_\mu, \quad (10)$$

so that

$$\rho_\alpha = \sum_\mu^N \sum_\nu^N \sum_i^{n_\alpha} (c_{\mu i}^\alpha)^* c_{\nu i}^\alpha \phi_\mu \phi_\nu = \sum_{\mu\nu} P_{\mu\nu}^\alpha \phi_\mu \phi_\nu, \quad (11)$$

$$\nabla \rho_\alpha = \sum_{\mu\nu} P_{\mu\nu}^\alpha \nabla(\phi_\mu \phi_\nu) \quad (12)$$

and similarly for ρ_β . By substituting these expressions into the energy (1) and then minimizing with respect to the unknown coefficients $c_{\mu i}^\alpha$, $c_{\mu i}^\beta$ (subject to the orthonormality of the ψ_i) we obtain a finite set of algebraic equations for canonical orbitals,

$$\sum_\nu^N (F_{\mu\nu}^\alpha - \epsilon_i^\alpha S_{\mu\nu}) c_{\nu i}^\alpha = 0 \quad (13)$$

and similarly for $c_{\nu i}^\beta$. The quantities ϵ_i^α , ϵ_i^β are one-electron eigenvalues for the occupied orbitals. The Fock-type matrices $F_{\mu\nu}$ are given by

$$F_{\mu\nu}^\alpha = H_{\mu\nu}^{\text{core}} + J_{\mu\nu} + F_{\mu\nu}^{\text{XC}\alpha}, \quad (14a)$$

$$F_{\mu\nu}^\beta = H_{\mu\nu}^{\text{core}} + J_{\mu\nu} + F_{\mu\nu}^{\text{XC}\beta}. \quad (14b)$$

Here $S_{\mu\nu}$ and $H_{\mu\nu}^{\text{core}}$ are the overlap and bare-nucleus Hamiltonian matrices, respectively, and $J_{\mu\nu}$ is the Coulomb matrix

$$J_{\mu\nu} = \sum_{\lambda\sigma}^N P_{\lambda\sigma}(\mu\nu|\lambda\sigma), \quad (15)$$

where $P_{\lambda\sigma}$ is the total density matrix ($P_{\lambda\sigma}^\alpha + P_{\lambda\sigma}^\beta$) and a conventional notation is used for two-electron repulsion integrals. The exchange-correlation parts of the Fock matrices are given by

$$F_{\mu\nu}^{\text{XC}\alpha} = \int \left[\frac{\partial f}{\partial \rho_\alpha} \phi_\mu \phi_\nu + \left(2 \frac{\partial f}{\partial \gamma_{\alpha\alpha}} \nabla \rho_\alpha + \frac{\partial f}{\partial \gamma_{\alpha\beta}} \nabla \rho_\beta \right) \cdot \nabla(\phi_\mu \phi_\nu) \right] d\mathbf{r} \quad (16)$$

and similarly for $F_{\mu\nu}^{\text{XC}\beta}$. Clearly, (16) and the matrix elements of the potential (9) are identical, using integration by parts.

The equations (13) are analogues of the Roothaan-Hall equations [14,15] (closed-shell with $c_{\mu i}^\alpha = c_{\mu i}^\beta$) and the Pople-Nesbet equations [16] (unrestricted open-shell) of conventional Hartree-Fock theory, the only difference being that the usual exchange matrix $K_{\mu\nu}$ is replaced by $F_{\mu\nu}^{\text{XC}\alpha}$ or $F_{\mu\nu}^{\text{XC}\beta}$. Once the equations have been solved by an iterative self-consistent procedure, the Kohn-Sham energy is obtained from (1) as

$$E = \sum_{\mu\nu}^N P_{\mu\nu} H_{\mu\nu}^{\text{core}} + \frac{1}{2} \sum_{\mu\nu\lambda\sigma}^N P_{\mu\nu} P_{\lambda\sigma}(\mu\nu|\lambda\sigma) + E^{\text{XC}}, \quad (17)$$

E^{XC} being given by (7).

For most functions $f(\rho_\alpha, \rho_\beta, \gamma_{\alpha\alpha}, \gamma_{\alpha\beta}, \gamma_{\beta\beta})$, the integrations in (7) and (16) have to be carried out by numerical quadrature. Here it is important to note that the integrands can be obtained *without* evaluation of the Hessians of the densities, even though these appear in the general potential (9). We have developed an efficient Kohn-Sham implementation [17] in which the cost of the exchange-correlation part of the calculation scales as $\mathcal{O}(N)$. The dominant step in this procedure is the evaluation of the spin densities and their gradients. Thus, avoidance of the density Hessians is a major computational advantage. It is worth noting that the exchange-correlation contribution to the Fock matrix can be constructed more efficiently from (9) than from (16) once the densities and their derivatives are available since the basis function gradients are not involved in the matrix elements of (9). However, this does not offset the additional expense of evaluating the Hessians.

The energy gradient with respect to nuclear coordinates is easily derived by differentiation of (1). Since the energy is already optimized with respect to the orbital coefficients $c_{\mu i}$, it is not necessary to determine changes in these coefficients (other than that implied by changes in the overlap matrix). The analysis follows closely that of gradients in Hartree-Fock theory [18], the final result for the gradient with respect to nucleus A being

$$\begin{aligned}
\nabla_A E = & \sum_{\mu\nu}^N P_{\mu\nu} (\nabla_A H_{\mu\nu}^{\text{core}}) + \frac{1}{2} \sum_{\mu\nu\lambda\sigma}^N P_{\mu\nu} P_{\lambda\sigma} \nabla_A (\mu\nu|\lambda\sigma) \\
& - \sum_{\mu\nu}^N W_{\mu\nu} (\nabla_A S_{\mu\nu}) - 2 \sum_{\mu}^N \sum_{\nu}^N \int \left[\frac{\partial f}{\partial \rho_{\alpha}} \phi_{\nu} \nabla \phi_{\mu} \right. \\
& \left. + X_{\mu\nu} \left(2 \frac{\partial f}{\partial \gamma_{\alpha\alpha}} \nabla \rho_{\alpha} + \frac{\partial f}{\partial \gamma_{\alpha\beta}} \nabla \rho_{\beta} \right) \right] d\mathbf{r} \\
& - 2 \sum_{\mu}^N \sum_{\nu}^N P_{\mu\nu}^{\beta} \int \left[\frac{\partial f}{\partial \rho_{\beta}} \phi_{\nu} \nabla \phi_{\mu} \right. \\
& \left. + X_{\mu\nu} \left(\frac{\partial f}{\partial \gamma_{\alpha\beta}} \nabla \rho_{\alpha} + 2 \frac{\partial f}{\partial \gamma_{\beta\beta}} \nabla \rho_{\beta} \right) \right] d\mathbf{r}. \quad (18)
\end{aligned}$$

(Note that primed sums are restricted to only those μ centered on nucleus A.) Here $W_{\mu\nu}$ is the energy-weighted density matrix

$$W_{\mu\nu} = \sum_i^{n_a} \epsilon_i^{\alpha} c_{\mu i}^{\alpha} c_{\nu i}^{\alpha} + \sum_i^{n_b} \epsilon_i^{\beta} c_{\mu i}^{\beta} c_{\nu i}^{\beta}, \quad (19)$$

and the matrix $X_{\mu\nu}$ is given by

$$X_{\mu\nu} = \phi_{\nu} \nabla (\nabla \phi_{\mu})^{\dagger} + (\nabla \phi_{\mu}) (\nabla \phi_{\nu})^{\dagger}. \quad (20)$$

We note that the integrands in (18) *do* require evaluation of the density Hessians. However, the gradient is only found once, after self-consistency has been achieved, whereas the integrations in (7) and (16) are required at each iteration.

In this manuscript, we have derived the basic equations directly from first principles. However, note should be taken of the previous publications which have treated parts of the problem. The Kohn–Sham equations for Gaussian basis sets and local density functionals (without gradient corrections) were considered originally by Sambe and Felton [19] and later by Dunlap et al. and Sabin [20]. The latter introduced projections of the density and exchange-correlation potentials onto additional bases, a technique which we have not considered here. The gradients of such procedures (again without gradient-corrected functionals) were first considered by Satoko [21] and, later in greater detail, by Fournier et al. [22] and Andzelm et al. [23]. The Kohn–Sham equations for gradient-corrected functionals have been derived by Perdew and Wang [5], Krijn and Feil [24], Kutzler and Painter [25], Mlynarski and Salahub [26] and Fan and Ziegler [27,28], but all

in forms which explicitly demand the evaluation of the density Hessians.

The self-consistent and gradient equations proposed in this paper have been coded in a modified version of the GAUSSIAN 92 program [29] and will be used in future studies of various functionals and gaussian orbital basis sets.

Acknowledgement

We thank Professor N.C. Handy for useful comments on this manuscript. This research was partly supported by the National Science Foundation (Grant No. 8918623). BGJ thanks the Mellon College of Science for a Graduate Fellowship.

References

- [1] P. Hohenberg and W. Kohn, *Phys. Rev. B* 136 (1964) 864.
- [2] W. Kohn and L.J. Sham, *Phys. Rev. A* 140 (1965) 1133.
- [3] R.G. Parr and W. Yang, *Density-functional theory of atoms and molecules* (Oxford Univ. Press, Oxford, 1989).
- [4] F. Herman, J.P. van Dyke and I.B. Ortenburger, *Phys. Rev. Letters* 22 (1969) 807.
- [5] J.P. Perdew and W. Yue, *Phys. Rev. B* 33 (1986) 8800.
- [6] J.P. Perdew, *Phys. Rev. B* 33 (1986) 8822.
- [7] S.K. Ghosh and R.G. Parr, *Phys. Rev. A* 34 (1986) 785.
- [8] A.D. Becke, *J. Chem. Phys.* 84 (1986) 4524.
- [9] A.D. Becke, *J. Chem. Phys.* 85 (1986) 7184.
- [10] A.D. Becke, *Phys. Rev. A* 38 (1988) 3098.
- [11] A.D. Becke, *J. Chem. Phys.* 88 (1988) 1053.
- [12] A. Cedillo, J. Robles and J.L. Gazquez, *Phys. Rev. A* 38 (1988) 1697.
- [13] C. Lee, W. Yang and R.G. Parr, *Phys. Rev. B* 37 (1988) 785; B. Miehlich, A. Savin, H. Stoll and H. Preuss, *Chem. Phys. Letters* 157 (1989) 200.
- [14] C.C.J. Roothaan, *Rev. Mod. Phys.* 23 (1951) 69.
- [15] G.G. Hall, *Proc. Roy. Soc. A* 205 (1951) 541.
- [16] J.A. Pople and R.K. Nesbet, *J. Chem. Phys.* 22 (1954) 571.
- [17] P.M.W. Gill, B.G. Johnson, J.A. Pople and M.J. Frisch, to be published.
- [18] P. Pulay, in: *Modern theoretical chemistry*, Vol. 4, ed. H.F. Schaefer III (Plenum Press, New York, 1977).
- [19] H. Sambe and R.H. Felton, *J. Chem. Phys.* 62 (1975) 1122.
- [20] B.I. Dunlap, J.W.D. Connolly and J.R. Sabin, *J. Chem. Phys.* 71 (1979) 3396.
- [21] C. Satoko, *Chem. Phys. Letters* 83 (1981) 111.
- [22] R. Fournier, J. Andzelm and D.R. Salahub, *J. Chem. Phys.* 90 (1989) 6371.

- [23] J. Andzelm, E. Wimmer and D.R. Salahub, The challenge of d and f electrons: theory and computation, ACS Symp. Ser. 394, eds. D.R. Salahub and M.C. Zerner (Am. Chem. Soc., New York, 1989) p. 228.
- [24] M.P.C.M. Krijn and D. Feil, Chem. Phys. Letters 150 (1988) 45.
- [25] F.W. Kutzler and G.S. Painter, Phys. Rev. B 37 (1988) 285.
- [26] P. Mlynarski and D.R. Salahub, Phys. Rev. B 43 (1991) 1399.
- [27] L. Fan and T. Ziegler, J. Chem. Phys. 94 (1991) 6057.
- [28] L. Fan and T. Ziegler, J. Chem. Phys. 95 (1991) 7401.
- [29] M.J. Frisch, G.W. Trucks, M. Head-Gordon, P.M.W. Gill, M.W. Wong, J.B. Foresman, B.G. Johnson, H.B. Schlegel, M.A. Robb, E.S. Replogle, R. Gomperts, J.L. Andres, K. Raghavachari, J.S. Binkley, C. Gonzalez, R.L. Martin, D.J. Fox, D.J. DeFrees, J. Baker, J.J.P. Stewart and J.A. Pople, GAUSSIAN 92 (Gaussian, Inc., Pittsburgh PA, 1992).