Journal of Computational Electronics X:YYY-ZZZ,200?

Convergence of Density Functional Iterative Procedures with a Newton-Raphson Algorithm

J.W. JEROME
Department of Mathematics, Northwestern University, Evanston, IL 60208 USA
jwj@math.northwestern.edu

P.R. SIEVERT, L.-H. YE, I.G. KIM, and A.J. FREEMAN
Department of Physics and Astronomy, Northwestern University, Evanston, IL 60208 USA
art@freeman.phys.northwestern.edu

Abstract. State of the art first-principles calculations of electronic structures aim at finding the ground state electronic density distribution. The performance of such methodologies is determined by the effectiveness of the iterative solution of the nonlinear density functional Kohn-Sham equations. We first outline a solution strategy based on the Newton-Raphson method. A form of the algorithm is then applied to the simplest and earliest density functional model, i. e., the atomic Thomas-Fermi model. For the neutral atom, we demonstrate the effectiveness of a charge conserving Newton-Raphson iterative method for the computation, which is independent of the starting guess; it converges rapidly, even for a randomly selected normalized starting density.

Keywords: Density functional theory, Kohn-Sham system, Newton-Raphson algorithm, Thomas-Fermi model

1. Introduction

Traditional mixing methods for density functional theory calculations may not be appropriate for the larger complex systems of current technological interest. Mixing employs successive approximation iterates of a fixed point mapping. Such iterates are often found to converge very slowly or not at all. Currently, mixing is employed at the level of the nonlinear discretization equations (see Ref. [1] for a discussion of Anderson mixing, Broyden updates, and the relation between them). Here an operator version of the Newton-Raphson method, applied to the Kohn-Sham system of density functional theory, is suggested.

2. Density Functional Theory

We are interested in the application of density functional theory to periodic structures in solid-state physics. A development of the Kohn-Sham system, as derived from the Hohenberg-Kohn theorem, is presented in Ref. [2]. The theory represents a many-electron system in terms of non-interacting effective particles; the inter-electron effects are transferred to the exchange-correlation potential, expressed as a functional of the electron density ρ. The Coulomb potential is also determined in terms of ρ and the fixed nuclear charge. The consolidated potential is the effective potential V_{eff}. This leads to the Hamiltonian H and its associated Kohn-Sham orbitals. We thus obtain the operator representation of the Kohn-Sham system as described specifically below. The operator (structure map) is linearized in our application of the Newton-Raphson method, and can be incorporated into existing electronic structure codes (see Ref. [3]).

3. Structure Map

The structure map may be represented schematically as a mapping P for the self-consistent Kohn-Sham equations. It is defined on possible electron density configurations, satisfying charge conservation laws. A solution of the Kohn-Sham system is a fixed point ρ, so that $P\rho = \rho$. The mapping P can be decomposed into $P = R \circ H \circ V$ with

$$\begin{aligned} V &: \rho \mapsto V_{\text{eff}} = V_C + \mu_{\text{exc}}, \\ H &: V_{\text{eff}} \mapsto (\Lambda, u) = (\lambda_1, \ldots, \lambda_{N_e}, \psi_1, \ldots, \psi_{N_e}), \\ R &: (\Lambda, u) \mapsto \tilde{\rho} = \sum_i w_i |\psi_i|^2. \end{aligned}$$

Here, the Coulomb potential V_C is determined as the solution of a Poisson equation subject to a periodic charge distribution and μ_{exc} is the exchange correlation potential, which is typically represented by the local density approximation (LDA): Ref. [4], and generalized gradient approximation (GGA): Ref. [5]. The ψ_i are Kohn-Sham orbitals in this theory with Kohn-Sham eigenvalues λ_i. The occupation numbers are customarily expressed in terms of the Fermi-Dirac distribution: $w_i = [\exp(\beta(\lambda_i - \mu_{\text{ec}})) + 1]^{-1}$ for $i = 1, \ldots, N_e$, satisfying $\sum w_i = N$ (electron number). The electrochemical potential μ_{ec} is determined by this condition. $\beta = (k_B T)^{-1}$, for the temperature T and Boltzmann's constant k_B. For a related study, see Ref. [6].

4. Variational Calculus

We write P_ρ as the functional derivative with respect to ρ, and consider its chain rule representation:

$$P_\rho(\rho) = R'(\Lambda, u) \circ H'(V) \circ V'(\rho).$$

The objective is to find a feasible $\delta\rho$ via the Newton-Raphson algorithm. An additional constraint is that of *charge conservation*: $\int \delta\rho = 0$, which requires that the effective variation of P does not include δw_i. This is equivalent to an operator projection. In consequence, the variation of R depends only upon δu, but not upon $\delta\Lambda$. We find:

- As a functional of ρ, the variation of V is:

$$\delta V = V'(\rho)\delta\rho = \delta V_C + \delta\mu_{\text{exc}},$$

where

- δV_C is the solution of the linear Poisson equation

$$-\nabla^2 \delta V_C = 4\pi\delta\rho.$$

Derivatives of μ_{exc} will depend upon representations of the exchange and correlation components (see Ref. [3]). For derivatives of H we compute: $H'(V)\delta V = (\delta\Lambda, \delta u)$, where

- $\delta\Lambda = (\delta\lambda_1, \ldots, \delta\lambda_{N_e}), \delta\lambda_\nu = (\delta V \psi_\nu, \psi_\nu)$; and,

- $\delta u = (\delta\psi_1, \ldots, \delta\psi_{N_e})$ is defined implicitly by:

$$[H - \lambda_\nu]\delta\psi_\nu = \mathcal{P}_\nu(\delta\lambda_\nu - \delta V)\psi_\nu = -\mathcal{P}_\nu(\delta V \psi_\nu),$$

where $\mathcal{P}_\nu$ is the projection onto the orthogonal complement of the spaces $\mathcal{M}_1, \ldots, \mathcal{M}_\nu$. Note that the Rayleigh quotient characterization of eigenvalues and eigenvectors has been employed in determining H'.

- As observed before, the variation of R depends only on δu:

$$R'(\Lambda, u)(\delta u) = 2 \sum_i w_i \ \text{Re} \ \{\psi_i^* \delta\psi_i\}.$$

This completes the determination of P_ρ.

5. The Newton Step

The Newton step is based on the map $F = I - P$, so that a fixed point for P is converted into a zero of F. The exact Newton step follows from: $F'(\rho)\delta\rho = -F(\rho)$. This means that $\delta\rho$ is *implicitly* defined by the relation $(I - P_\rho(\rho))\delta\rho = -(I - P)(\rho)$ for a given ρ. We then update ρ: $\tilde{\rho} = \rho + \delta\rho$. What has been developed is valid at the level of the differential system. Galerkin's method is now invoked, for the basis functions employed in the algorithm. The key step is to form the Jacobian matrix approximation $J(\rho) \approx F'(\rho)$. The implementation involving the generalized minimum residual (GMRES) may be carried out as described by Kerkhoven and Saad: Ref. [7, p. 537; Appendix C]. It may also be combined with projection methods: Ref. [8] for the eigenvalue problem.

6. Thomas-Fermi Model

The earliest density functional approach, the Thomas-Fermi model, avoids the complexity inherent in the Hamiltonian mapping within the Kohn-Sham theory, and facilitates the Newton step. This model has been extensively studied (see Ref. [9] for an exposition of the model, and Ref. [10] for rigorous mathematical statements, including the existence of a unique energy minimizer for the neutral atom). Ref. [10] includes discussion of the Dirac and von Weizsäcker extensions of the model; these are not incorporated here.

After the functional minimization of the energy, we obtain, for the atomic number Z, the following mapping for the nonnegative density ρ in SI units (see Ref. [9]):

$$\frac{5\kappa}{3}\rho^{2/3}(\mathbf{r}) - \frac{e^2}{4\pi\epsilon_0}\frac{Z}{|\mathbf{r}|} + \frac{e^2}{4\pi\epsilon_0}\int \mathrm{d}^3r' \, \frac{\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} = \mu, \quad (1)$$

where $\kappa = 3^{5/3}\pi^{4/3}\hbar^2/(10m)$. The Lagrange multiplier μ, which fixes the number of electrons, may be set to zero, as appropriate for the neutral atom (see Ref. [10]). After assuming spherical symmetry, we introduce $r = |\mathbf{r}|$ and the Bohr radius,

$$a_0 = \frac{4\pi\epsilon_0\hbar^2}{me^2}. \quad (2)$$

This allows the introduction of a dimensionless variable x, such that $bx = r$ where,

$$b = \left(\frac{3^2\pi^2}{2^7 Z}\right)^{1/3} a_0. \quad (3)$$

However, we are faced with the problem that the charge density ρ is singular at the origin. The singularity can be removed by transforming to a dimensionless variable Q which is a function of x:

$$\rho(r) \mapsto A\left(\frac{Q(x)}{x^{3/2}}\right), \quad (4)$$

where

$$A = \frac{2^5 Z^2}{3^2(a_0\pi)^3}. \quad (5)$$

Under this scaling transformation our mapping then has the form

$$Q^{\mathrm{out}} = R \circ P\left[Q^{\mathrm{in}}\right], \quad (6)$$

where

$$P[\bullet] \equiv \left[\int_x^\infty \mathrm{d}x' \; \bullet \left(\frac{x'-x}{\sqrt{x'}}\right)\right]^{3/2} \quad (7)$$

and

$$R[\bullet] = \frac{1}{\int_0^\infty \mathrm{d}x' \; \bullet \; \sqrt{x'}} \, \bullet \quad (8)$$

is a normalization mapping, which enforces charge conservation for the range of $R \circ P$.

7. Results and Discussion

To apply the Newton-Raphson method, we introduce the mapping $Y = I - R \circ P$. In terms of this mapping, the iterates are given by $Q^{\mathrm{new}} = Q^{\mathrm{old}} + \delta Q$, where δQ is determined by the implicit relation

$$Y'\left[Q^{\mathrm{old}}\right]\delta Q = -Y\left[Q^{\mathrm{old}}\right]. \quad (9)$$

The chain rule of the operator calculus must be utilized, as above in the case of the Kohn-Sham mapping. The format of R yields two terms for its functional derivative. Note that the argument appears in two distinct places in Eq. (8). We observed from the numerical computation, upon introducing numerical error $\epsilon > 0$, that one can enforce $\int_0^\infty \mathrm{d}x \delta Q(x)\sqrt{x} < \epsilon$, which suggests that the Newton-Raphson algorithm satisfies charge conservation automatically. The figures are based upon simulations involving a Newton-Raphson implementation of Eq. (9) with trapezoidal integration rules. The x-axis grid was defined via a uniform grid on the t-axis, via the transformation $t = \ln x$. Typical ranges for t were: $-20 \le t \le 10$. It was determined that the tail calculation associated with the improper integral (7) is critical to the convergence of the iteration. In order to incorporate the tail, we used the known asymptotic relation: $Q(x) \doteq (\frac{144}{x^3})^{3/2}$, $x \to \infty$, and combined this with a splitting of (x,∞) into $(x, x_{\max}] \cup (x_{\max}, \infty)$. The cutoff $x_{\max}$ terminates the discretization, whereas the integral over the second interval admits of exact evaluation as a linear polynomial function of x: $\frac{1728}{x_{\max}^3}\left\{\frac{1}{3} - \frac{x}{4x_{\max}}\right\}$. For purposes of comparison, we introduce the variable $\chi = Q^{2/3}$, where χ is the unique bounded solution, positive for all nonnegative x, of the boundary-value problem (see Ref. [9]),

$$\begin{aligned} \chi''(x) &= \frac{\chi^{3/2}}{x}, \\ \chi(0) = 1, &\qquad \chi(\infty) = 0. \end{aligned} \quad (10)$$

The boundary problem (10) describes the potential in the Thomas-Fermi model for the neutral atom. Two additional solutions for χ exist which are excluded: an unbounded solution corresponding to the case $N > Z$, not permitted for a consistent Thomas-Fermi model; and the (positive) ionic case, for which $\chi(x) = 0,\ x > x_0$. We solved the boundary problem (10) to great precision via a shooting method. In doing so, we relied upon the known value of $\chi'(0)$. We then employed three starting 'guesses' for the Newton-Raphson iteration, subject to charge normalization:

(1) $Q_0 = \chi^{3/2}$; (2) $Q_0(x) = 4\pi \exp(-4^{1/3}\pi x)$; (3) $Q_0(x) = \mathrm{Random}(x)$. Here, the random function has been normalized to have charge one. Figure 1 displays the iteration histories for each of the starting 'guesses' (1,2,3). Each history is in the form of the ℓ^1-norm of δQ versus the iteration number. It is found that convergence is rapid in each case, where the stopping criterion is: ℓ^1-distance $= \sum |\delta Q(x_i)| < 10^{-10}$.

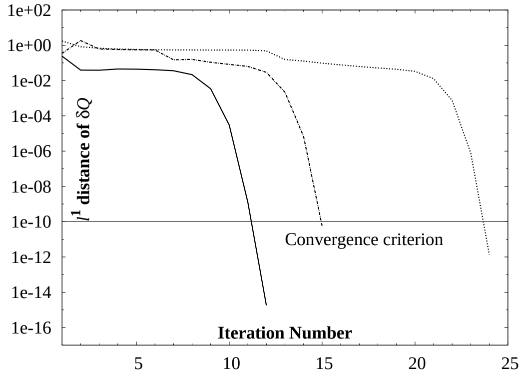

Figure 1: The ℓ^1 distance of δQ from the origin versus iteration number. Vertical axis log-scaled. Solid: $Q_0 = \chi^{3/2}$; Dash-Dot: $Q_0(x) = 4\pi e^{-4^{1/3}\pi x}$; Dots: $Q_0(x) = \mathrm{Random}(x)$.

To confirm the stability of the fixed point Q, we indicate in Figure 2 the rapidly converging iteration starting with the function $\mathrm{Random}(x)$. The ninth, thirteenth, and twenty-second iterations are shown in log-scale. As the iteration proceeds, the Q-tail, and the Q-function itself, converge toward the asymptotically decaying solution obtained with the shooting method.

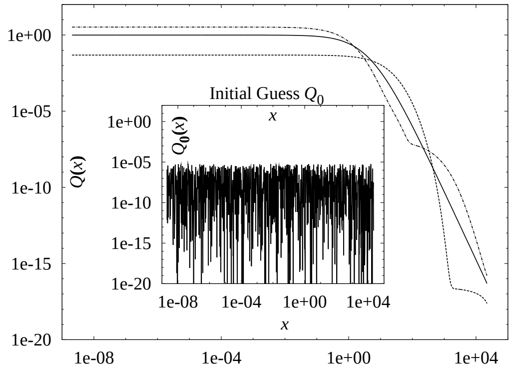

Figure 2: Convergence of the Newton-Raphson algorithm, starting with normalized random numbers. Axes log-scaled. Inset: $\mathrm{Random}(x)$; Dash: Iteration 9; Dash-Dot: Iteration 13; Solid: Iteration 22.

8. Summary

We have outlined an acceleration strategy, based upon the Newton-Raphson iteration method, for solution of the Kohn-Sham equations. As test case, we have analyzed the earliest density functional theory model, the Thomas-Fermi atomic model, which can be formulated without reference to orbitals. We have shown that the density corresponding to the neutral atom can be computed by a stable Newton-Raphson iteration, independent of normalized starting guess, as applied to an operator fixed point mapping. Accuracy is corroborated by reference to literature benchmarks.

Acknowledgment

This work was supported by an ONR-DARPA grant (Grant No. N00014-05-C-0241) and NSF through its MRSEC program in the Materials Science Research Center. IGK was supported by the Korea Science and Engineering Foundation (KOSEF).

References

1. V. Eyert, A comparative study on methods for convergent acceleration of iterative vector sequences, Journal of Computational Physics **124**, 271–285 (1996).
2. W. Kohn and P. Vashista, *Ch. 2, Theory of the Inhomogeneous Electron Gas* (Plenum Press, New York, 1983).
3. A. J. Freeman and E. Wimmer, Density functional theory as a major tool in computational materials science, Annual Reviews of Material Science **25**, 7–36 (1995).
4. L. Hedin and B.I. Lundqvist, Explicit local exchange correlation potentials, J.Phys. C **4** (14), 2064–2083 (1971).
5. J.P. Perdew,J.A. Chevay,S.H. Vosko,K.A. Jackson,M.R. Pederson, D.J. Singh, and C. Fiolhais, Atoms, molecules, solids, and surfaces, Phys. Rev. B **46**, (11), 6671–6687 (1992).
6. E. Prodan and P. Nordlander, On the Kohn-Sham equations with periodic background potentials, Journal of Statistical Physics **111**, 967–992 (2003).
7. T. Kerkhoven and Y. Saad, On acceleration methods for coupled nonlinear elliptic systems Numer. Math. **60**, 525–548 (1992).
8. Y. Huang, D.K. Hoffman, and D.J. Kouri, A minimal subspace residual method for large scale eigenvalue problems, Journal of Chemical Physics **110**, 8303–8308 (1999).
9. N. H. March, *Ch. 1, Theory of the Inhomogeneous Electron Gas* (Plenum Press, New York, 1983).
10. E. H. Lieb, Thomas-Fermi and related theories of atoms and molecules, Reviews of Modern Physics **54**, 603–641 (1981).