Journal of Computational Physics 231 (2012) 4515–4529

Contents lists available at SciVerse ScienceDirect

Journal of Computational Physics

journal homepage: www.elsevier.com/locate/jcp

Optimized local basis set for Kohn–Sham density functional theory

Lin Lin [a,*,1], Jianfeng Lu [b], Lexing Ying [c], Weinan E [d,e]

[a] Program in Applied and Computational Mathematics, Princeton University, Princeton, NJ 08544, USA
[b] Courant Institute of Mathematical Sciences, New York University, New York, NY 10012, USA
[c] Department of Mathematics and ICES, University of Texas at Austin, Austin, TX 78712, USA
[d] Department of Mathematics and PACM, Princeton University, Princeton, NJ 08544, USA
[e] Beijing International Center for Mathematical Research, Peking University, Beijing 100871, China

ARTICLE INFO

Article history:
Received 4 August 2011
Received in revised form 10 February 2012
Accepted 13 March 2012
Available online 2 April 2012

Keywords:
Electronic structure
Kohn–Sham density functional theory
Optimized local basis set
Discontinuous Galerkin
Trace minimization
Molecular dynamics
Pulay force
GMRES
Preconditioning

ABSTRACT

We develop a technique for generating a set of optimized local basis functions to solve models in the Kohn–Sham density functional theory for both insulating and metallic systems. The optimized local basis functions are obtained by solving a minimization problem in an admissible set determined by a large number of primitive basis functions. Using the optimized local basis set, the electron energy and the atomic force can be calculated accurately with a small number of basis functions. The Pulay force is systematically controlled and is not required to be calculated, which makes the optimized local basis set an ideal tool for ab initio molecular dynamics and structure optimization. We also propose a preconditioned Newton–GMRES method to obtain the optimized local basis functions in practice. The optimized local basis set is able to achieve high accuracy with a small number of basis functions per atom when applied to a one dimensional model problem.

1. Introduction

In scientific computation of systems with large number of degrees of freedom, an efficient choice of basis functions becomes desirable in order to reduce the computational cost. In this paper, we focus on the choice of efficient basis sets for the Kohn–Sham density functional theory (KSDFT) [1,2], which is the most widely used electronic structure theory for condensed matter systems. The methods and concepts illustrated here are also useful for other applications.

In KSDFT, the quantities of interest are the electron energy $E(R)$ and the atomic force $F(R)$. Here we denote by $R = (R_1, R_2, \ldots, R_{N_A})^T$ the atomic positions, where N_A is the number of atoms. The atomic force is expressed in terms of the derivatives of the electron energy with respect to the atomic positions as $F(R) = -\frac{\partial E(R)}{\partial R}$. This is an important quantity in many applications including structure optimization and first principle molecular dynamics. The electron energy is a functional of a set of Kohn–Sham orbitals $\{\psi_i\}_{i=1}^N$ where N is the number of electrons in the system. To illustrate the idea with minimal technicality, let us consider for the moment a system of non-interacting electrons at zero temperature. The energy functional for non-interacting electrons takes the form

$$E\left(\{\psi_i(x)\}_{i=1}^N; R\right) = \frac{1}{2}\sum_{i=1}^N \int |\nabla\psi_i(x)|^2 \mathrm{d}x + \int V(x;R)\sum_{i=1}^N |\psi_i(x)|^2 \mathrm{d}x. \qquad (1)$$

---

[*] Corresponding author.

E-mail addresses: linlin@lbl.gov (L. Lin), jianfeng@cims.nyu.edu (J. Lu), lexing@math.utexas.edu (L. Ying), weinan@math.princeton.edu (W. E).

[1] Present address: Computational Research Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA.

0021-9991/$ - see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.jcp.2012.03.009

The first term and the second term in (1) are the kinetic energy and the potential energy of the system, respectively. The energy $E(R)$ as a function of atomic positions is given by the following minimization problem

$$\begin{aligned} &E(R) = \min_{\{\psi_i(x)\}_{i=1}^N} E\left(\{\psi_i(x)\}_{i=1}^N; R\right), \\ &\text{s.t.} \quad \int \psi_i^*(x)\psi_j(x)\mathrm{d}x = \delta_{ij}, \quad i,j = 1,\ldots,N. \end{aligned} \tag{2}$$

We denote by $\{\psi_i(x;R)\}_{i=1}^N$ the minimizer. It can be readily shown that $\{\psi_i(x;R)\}_{i=1}^N$ are the lowest N eigenfunctions of the Hamiltonian operator $H(R)$, which takes the form

$$H(R) = -\frac{1}{2}\Delta_x + V(x;R). \tag{3}$$

Using the Hamiltonian operator, the electron energy has an alternative expression without the explicit dependence on the orbitals $\{\psi_i\}_{i=1}^N$:

$$E(R) = \mathrm{Tr}[H(R)\chi(H(R) - \mu(R))] \equiv \mathrm{Tr}[g_0(H(R))], \tag{4}$$

where $\chi(x) = 1$ if $x < 0$ and is 0 otherwise. $\mu(R)$ is the chemical potential, which takes value between the Nth and $(N+1)$th eigenvalues of H to control the number of electrons. Here we assume there is a positive gap between the Nth eigenvalue and the $N+1$th eigenvalue corresponding to the Hamiltonian $H(R)$.

Since all the quantities depend on the atomic positions R, to simplify the notation we drop the dependence on R unless otherwise specified. If we approximate the eigenfunctions $\{\psi_i\}_{i=1}^N$ by linear combination of a set of basis functions $\Phi = (\phi_1,\ldots,\phi_{N_b})$, the Hamiltonian operator H is discretized into a finite dimensional matrix $\Phi^{\mathrm{T}}H\Phi$ (here and in the following, we will use the linear algebra notation: $\phi_i^{\mathrm{T}}H\phi_j = \langle\phi_i|H|\phi_j\rangle$). The number of basis functions N_b is therefore called the *discretization cost*. The electron energy and the force can be expressed in terms of the discretized Hamiltonian operator as

$$\begin{aligned} &E_\Phi = \mathrm{Tr}\left[g_0\left(\Phi^{\mathrm{T}}H\Phi\right)\right], \\ &F_{\Phi,I} = -\frac{\partial E_\Phi}{\partial R_I} = -\mathrm{Tr}\left[g_0'\left(\Phi^{\mathrm{T}}H\Phi\right)\Phi^{\mathrm{T}}\frac{\partial H}{\partial R_I}\Phi\right] - 2\mathrm{Tr}\left[g_0'\left(\Phi^{\mathrm{T}}H\Phi\right)\Phi^{\mathrm{T}}H\frac{\partial \Phi}{\partial R_I}\right]. \end{aligned} \tag{5}$$

$F_{\Phi,I}$ is the Ith component of the force. In what follows the second equation in (5) is also written in a compact form as

$$F_\Phi = -\frac{\partial E_\Phi}{\partial R} = -\mathrm{Tr}\left[g_0'\left(\Phi^{\mathrm{Tr}}H\Phi\right)\Phi^{\mathrm{Tr}}\frac{\partial H}{\partial R}\Phi\right] - 2\mathrm{Tr}\left[g_0'\left(\Phi^{\mathrm{Tr}}H\Phi\right)\Phi^{\mathrm{Tr}}H\frac{\partial \Phi}{\partial R}\right]. \tag{6}$$

Choosing basis functions Φ adaptively with respect to the atomic positions R has obvious computational advantages, as it allows the possibility to reduce the discretization cost by a significant amount while maintaining the accuracy for the evaluation of the electron energy and atomic forces. Since the electron energy is defined variationally as in (2), an accurate basis set should minimize the electron energy. However, choosing the basis functions adaptively gives rise to some difficulties in the evaluation of the force (5) which requires the calculation of $\frac{\partial \Phi}{\partial R}$. In electronic structure theory, the contribution from $\frac{\partial \Phi}{\partial R}$ is referred to as the Pulay force [3]. We will henceforth adopt this terminology. The Pulay force originates from the incompleteness of the basis set, and has been found to be important to obtain the force with reliable accuracy for structure optimization or first principle molecular dynamics [3,4]. The calculation of the Pulay force can be quite expensive even if the basis functions Φ have analytical expressions, and the calculation of the Pulay force becomes almost intractable if the basis functions are defined implicitly such as in the adaptive mesh method [5–8]. We would like to systematically reduce the Pulay force so that the approximation

$$\frac{\partial E_\Phi}{\partial R} \approx \mathrm{Tr}\left[g_0'\left(\Phi^{\mathrm{T}}H\Phi\right)\Phi^{\mathrm{T}}\frac{\partial H}{\partial R}\Phi\right] \tag{7}$$

becomes adequate.

The key observation in this paper is that minimizing the electron energy and reducing the Pulay force can be simultaneously achieved by the following optimization procedure

$$\min_{\Phi\subset\mathcal{V},\Phi^{\mathrm{T}}\Phi=I} E_\Phi = \min_{\Phi\subset\mathcal{V},\Phi^{\mathrm{T}}\Phi=I} \mathrm{Tr}\left[g_0\left(\Phi^{\mathrm{T}}H\Phi\right)\right]. \tag{8}$$

Here $\mathcal{V}$ is an admissible subset of the space spanned by a set of *primitive basis functions* which are independent of R. Later $\mathcal{V}$ will be referred to as the *admissible set*. We select from $\mathcal{V}$ a small number of R-dependent *optimized basis functions* $\Phi = (\Phi_1,\ldots,\Phi_{N_b})$ which give rise to the lowest electron energy in $\mathcal{V}$. The Euler–Lagrange equation for the minimization problem (8) reads

$$\begin{cases} H\Phi g_0'(\Phi^{\mathrm{T}}H\Phi) = \Phi\Lambda, \\ \Phi^{\mathrm{T}}\Phi = I, \end{cases} \tag{9}$$