Journal of Computational Mathematics, Vol.27, No.2-3, 2009, 360–387.

DIRECT MINIMIZATION FOR CALCULATING INVARIANT SUBSPACES IN DENSITY FUNCTIONAL COMPUTATIONS OF THE ELECTRONIC STRUCTURE*

Reinhold Schneider and Thorsten Rohwedder
Institute for Mathematics, Technical University Berlin, Straße des 17. Juni 135, 10623 Berlin, Germany
Email: schneidr@math.tu-berlin.de, rohwedde@math.tu-berlin.de
Alexey Neelov
Institute of Physics, University of Basel, Klingelbergstrasse 82, 4056 Basel, Switzerland
Email: alexey.neelov@unibas.ch
Johannes Blauert
Institute for Mathematics, Technical University Berlin, Straße des 17. Juni 135, 10623 Berlin, Germany

Abstract

In this article, we analyse three related preconditioned steepest descent algorithms, which are partially popular in Hartree-Fock and Kohn-Sham theory as well as invariant subspace computations, from the viewpoint of minimization of the corresponding functionals, constrained by orthogonality conditions. We exploit the geometry of the admissible manifold, i.e., the invariance with respect to unitary transformations, to reformulate the problem on the Grassmann manifold as the admissible set. We then prove asymptotical linear convergence of the algorithms under the condition that the Hessian of the corresponding Lagrangian is elliptic on the tangent space of the Grassmann manifold at the minimizer.

Mathematics subject classification: 65Z05, 58E50, 49R50.
Key words: Eigenvalue computation, Grassmann manifolds, Optimization, Orthogonality constraints, Hartree-Fock theory, Density functional theory, PINVIT.

1. Introduction

On the length-scale of atomistic or molecular systems, physics is governed by the laws of quantum mechanics. A reliable computation required in various fields in modern sciences and technology should therefore be based on the first principles of quantum mechanics, so that *ab initio* computation of the electronic wave function from the stationary electronic Schrödinger equation is a major working horse for many applications in this area. To reduce computational demands, the high dimensional problem of computing the wave function for N electrons is often, for example in Hartree-Fock and Kohn-Sham theory, replaced by a nonlinear system of equations for a set $\Phi = (\varphi_1, \cdots, \varphi_N)$ of single particle wave functions $\varphi_i(\mathbf{x}) \in V = H^1(\mathbb{R}^3)$. This ansatz corresponds to the following abstract formulation for the minimization of a suitable energy functional $\mathcal{J}(\Phi)$

Problem 1. Minimize

$$\mathcal{J} : V^N \to \mathbb{R}, \quad \mathcal{J}(\Phi) \;=\; \mathcal{J}(\varphi_1, \cdots, \varphi_N) \;\longrightarrow\; \min, \tag{1.1}$$

* Received April 20, 2008 / Revised version received July 11, 2008 / Accepted July 23, 2008 /

with V^N defined via V as above and $\mathcal{J}$ is a sufficiently often differentiable functional which is

(i) invariant with respect to unitary transformations, i.e.,

$$\mathcal{J}(\Phi) = \mathcal{J}(\Phi\mathbf{U}) = \mathcal{J}\Big(\Big(\sum_{j=1}^{N} u_{i,j}\phi_j\Big)_{i=1}^{N}\Big), \tag{1.2}$$

for any orthogonal matrix $\mathbf{U} \in \mathbb{R}^{n\times n}$, and

(ii) subordinated to the orthogonality constraints

$$\langle \varphi_i, \varphi_j \rangle := \int_{\mathbb{R}^3} \varphi_i(x)\varphi_j(x)dx = \delta_{i,j}. \tag{1.3}$$

In the present article, we shall be concerned with minimization techniques for $\mathcal{J}$ along the admissible manifold characterized by (1.3). The first step towards this will be to set up the theoretical framework of the *Grassmann manifold* to be introduced in Section 2, reflecting the constraints (i) and (ii) imposed on the functional $\mathcal{J}$ and the minimizer Φ, respectively. In applications in electronic structure theory, formulation of the first order optimality (necessary) condition for the problem (1.1) results in a nonlinear eigenvalue problem of the kind:

$$A_\Phi\varphi_i = \lambda_i\varphi_i, \quad \lambda_1 \le \lambda_2 \le \cdots \le \lambda_N \tag{1.4}$$

for N eigenvalues λ_i and the corresponding solution functions assembled in Φ. In these equations, the operator A_Φ, is a symmetric bounded linear mapping

$$A_\Phi : V = H^1(\mathbb{R}^3) \to V' = H^{-1}(\mathbb{R}^3)$$

depending on Φ, so that we are in fact faced with a nonlinear eigenvalue problem. A_Φ is called the *Fock operator* in Hartree-Fock theory, and *Kohn-Sham* Hamiltonian in *density functional theory* (DFT) respectively. We will illustrate the relation between (1.4) and the minimization task above in further detail in Section 3. In this work, our emphasis will rather be on the algorithmic approximation of the minimizer of $\mathcal{J}$, i.e. an invariant subspace $\mathrm{span}[\Phi] := \mathrm{span}\{\varphi_1, \cdots, \varphi_N\}$, of (1.4), in the corresponding energy space V^N than on computation of the eigenvalues $\lambda_1, \cdots, \lambda_N$.

One possible procedure for computing the minimum of $\mathcal{J}$ is the so-called *direct minimization*, utilized e.g. in DFT calculation, which performs a steepest descent algorithm by updating the gradient of $\mathcal{J}$, i.e. the Kohn-Sham Hamiltonian or Fock operator, in each iteration step. Direct minimization, as proposed in [2], is prominent in DFT calculations if good preconditioners are available and the systems under consideration are large, e.g. for the computation of electronic structure in bulk crystals using plane waves, finite differences [7] and the recent wavelet code developed in the BigDFT project (see [45]). In contrast to the direct minimization procedure is the *self consistent field iteration (SCF)*, which keeps the Fock operator fixed until convergence of the corresponding eigenfunctions and updates the Fock operator thereafter, see Section 3.

In the rest of this article, we will pursue different variants of projected gradient algorithms to be compiled in Section 4. In addition, we will (for the case where the gradient $\mathcal{J}'(\Phi)$ can be written as an operator A_Φ applied to Φ, as it is the case in electronic structure calculation) investigate an algorithm based on [4] following a preconditioned steepest descent along geodesics on the manifold. so that no re-projections onto the admissible manifold are required. It turns