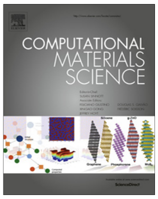
Computational Materials Science 133 (2017) 71–81

Contents lists available at ScienceDirect

Computational Materials Science

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



On-the-fly parameterization of internal coordinate force constants for quasi-Newton geometry optimization in atomistic calculations

Christoph Freysoldt

Max-Planck-Institut für Eisenforschung, Max-Planck-Str. 1, 40237 Düsseldorf, Germany

ARTICLE INFO

Article history:
Received 29 November 2016
Received in revised form 28 February 2017
Accepted 2 March 2017

Keywords:
Geometry optimization
Redundant internal coordinates
Quasi Newton

ABSTRACT

A variant of quasi-Newton geometry optimization in Cartesian coordinates for atomistic calculations is proposed. Like other schemes, it starts from an approximate Hessian diagonal in redundant internal coordinates (bond lengthes, bond angles) which is improved by Broyden-Fletcher-Goldfarb-Shanno (BFGS) updates. The key idea is to parameterize the diagonal elements of the starting Hessian on the fly. By automatically classifying all interatomic bonds according to bond length and elements involved, and treating all similar bonds equal, one arrives at a very small number of parameters that can be determined from few displacements, often only one. The superior performance for supercell calculations compared to a standard Cartesian-coordinate optimizer is demonstrated.

1. Introduction

Geometry optimization, i.e., determining the (local) energy minimum of a given system of atoms by varying the atomic coordinates, is an essential and recurring task ever since the early days of atomistic calculations (see [1] for a recent review, and references therein). From a very abstract point of view, this is a non-linear minimization problem in a 3N-dimensional space, where N is the number of atoms. For such a non-linear minimization, there are many general-purpose iterative schemes available [2]. Each scheme probes the energy $E(\{\mathbf{R}_I\})$ for a given configuration (atomic positions) $\{\mathbf{R}_I\}$ and possibly its first, second, etc. derivatives with respect to atomic positions. It then proposes from the data acquired so far a new configuration to probe until the change in energy and/or the forces fall below a given threshold. The schemes differ in the required input data: For instance, the downhill-simplex method [3] only requires the energy. Gradient-based schemes are steepest descent, conjugate-gradient [4,5], direct inversion in the iterative subspace (DIIS) [6,7] or other Krylov subspace methods. Last, Newton or quasi-Newton schemes also require the second derivative (or an approximation to it). The schemes further differ in their robustness against numerical noise, in their memory demands for the implementation, and last, but certainly not least in their efficiency in arriving at the desired minimum. The calculation of the energy, as well as its first and possibly second derivative with respect to the atomic coordinates can be computationally very demanding. Of course, this depends on the underlying theory for the potential energy surface, e.g., empirical interatomic potentials, semi-empirical electronic structure theory, tight-binding [8], bond-order potentials [9], density-functional theory (DFT) or other variants of *ab initio* electronic structure theory. It is therefore desirable to employ efficient methods that use a minimal number of steps, possibly exploiting additional physical and chemical knowledge that is not available to general-purpose minimizers. Some of the most common minimizers have been recently benchmarked [10].

The focus in this work is on local minimum searches in density-functional theory calculations with periodic boundary conditions. Such calculations are routinely employed in solid-state research, with typically up to several hundred or even a few thousand atoms. In this case, memory and computational demands for the geometry optimization are negligible compared to the underlying calculation. Forces $\mathbf{f}_I = -\partial E/\partial \mathbf{R}_I$ are readily available, while analytic second derivatives are prohibitive compared to the advantage that they may have. A common (if not *the* most common) approach for this situation is the quasi-Newton scheme – and also the clear winner for local geometry optimization in the above-mentioned benchmark [10]. The quasi-Newton scheme can be summarized as follows. A necessary condition for being in the energy minimum is that the forces vanish, i.e.,

$$f_i = -\frac{\partial E}{\partial R_i} = 0 \tag{1}$$

for all coordinates i. Given the second derivative (the Hessian) $H_{ij} = \partial^2 E/(\partial R_i \partial R_j)$, the optimal step $\Delta R_i = R_i^{(n+1)} - R_i^{(n)}$ to take from the n-th configuration to the next one is

---

E-mail address: freysoldt@mpie.de

http://dx.doi.org/10.1016/j.commatsci.2017.03.001
0927-0256/

$$\Delta R_i = \sum_j \left\{H^{-1}\right\}_{ij} f_j. \tag{2}$$

For a purely parabolic potential energy surface, this Newton step directly jumps to the minimum if the Hessian is known exactly. In more realistic cases, there is a weak non-linear (anharmonic) dependence of the forces on the displacement ΔR_i. Moreover, the Hessian used in Eq. (2) is often only approximate (which makes the scheme *quasi*-Newton). In these cases, the errors in positions, forces, energy, etc. are reduced on average by a constant factor in each step.

This can be understood from the general theory of Krylov subspace methods (see Appendix A for details). The rate of convergence is determined by the condition number (the ratio of the largest to the smallest eigenvalue of the Hessian). The quasi-Newton scheme can be interpreted as a steepest-descent scheme with the approximate Hessian serving as a preconditioner. The convergence rate then depends on how well the approximate Hessian is able to compress the eigenvalue spectrum of the true Hessian. Therefore, quasi-Newton schemes continuously improve the approximate Hessian based on the observed change in the forces upon displacing the atoms. In the Broyden-Fletcher-Goldfarb-Shanno (BFGS) variant [11–14] (the most common one in geometry optimization of local minima [1,10]), the inverse of the Hessian, B, is updated as

$$\begin{aligned} B_{ij}^{(n+1)} =& \sum_{i'j'} \left(\delta_{ii'} - \gamma s_i^{(n)} y_{i'}^{(n)}\right) B_{i'j'}^{(n)} \\ & \times \left(\delta_{j'j} - \gamma y_{j'}^{(n)} s_j^{(n)}\right) \\ & + \gamma s_i^{(n)} s_j^{(n)} \end{aligned} \tag{3}$$

with

$$s_i^{(n)} = R_i^{(n+1)} - R_i^{(n)} \tag{4}$$

$$y_i^{(n)} = -\left(f_i^{(n+1)} - f_i^{(n)}\right) \tag{5}$$

and

$$\gamma = 1 / \sum_i s_i^{(n)} y_i^{(n)}. \tag{6}$$

The update defined by Eqs. (3)–(6) garantuees that (a) the force change in the last step is correctly predicted by the new Hessian, and (b) the Hessian is otherwise changed minimally. Please note that there is no need to perform a line search if the potential energy surface is close to quadratic: If the force changes are linear in the displacements, the step proposed by the updated Hessian removes the remaining force along the previous step, see Appendix B.

As higher-order contributions to the potential energy surface become increasingly important for *large* steps, most algorithms limit the step size. This ensures that the effective quadratic contributions captured by the quasi-Newton scheme dominate the change in forces. Limiting the step size has an additional advantage in electronic-structure calculations. The wavefunctions and electronic density are typically computed iteratively from some initial guess. At each step of geometry optimization, the converged wavefunction and density of the current atomic configuration are then extrapolated to the new structure. The smaller the change in atomic positions, the better this new initial guess for the electronic loop becomes. Hence, several small atomic steps in the same direction may be quite comparable in total run time to one big step. In our case, the maximum displacement is limited to 0.3 bohr.

The initial Hessian is typically taken diagonal, and it is here where most of the trouble starts in Cartesian coordinate optimization. In the absence of external fields, atomic forces depend on interatomic distances, developing notably between nearest neighbors. The structure of the Hessian, i.e., its dominant entries, is therefore determined by the nature of the chemical bond between atoms, and their connectivity of course. Neglecting this fundamental principle gives a lousy initial spectrum. It is (unfortunately) frequent practice in periodic-boundary DFT codes to initialize diagonal Hessians for Cartesian-coordinate optimizations with a single value on the diagonal [10,15]. For a diagonal matrix, the eigenvalues are given by the entries on the diagonal. However, the true spectrum in a solid ranges from very hard local modes, that affect a few interatomic distances in a narrow spatial region, to very soft, elastic-wave like modes. Recovering the spectrum by BFGS updates even qualitatively takes $O(3N)$ steps. This has of course been recognized early on in the molecular electronic structure community. One solution proposed is to perform the optimization not in Cartesian coordinates, but in redundant internal coordinates [16,17]. Transformation from Cartesian coordinates to redundant internal coordinates is straight-forward, and so is the projection of forces [17]. The reverse procedure of translating redundant internal coordinates to Cartesian ones is approximate (because there may not exist a configuration that yields *all* redundant coordinates at the desired value), and iterative. It is also possible to set up the initial Hessian for Cartesian coordinate optimization from a (diagonal) Hessian in internal coordinates. While this improves the quality of the initial guess, other advantages of internal coordinates are lost. For instance, if a fragment of the structure rotates in space without changes in internal coordinates within the fragment, we may expect that the Hessian remains largely unchanged. In Cartesian coordinates, however, the rotation affects all entries.

Redundant internal coordinates obviously succeed in taking atom connectivity into account. However, the initial guess still requires reasonable start values. In molecular systems, or more generally: covalently bonded systems of main-group elements, the ‘usual’ internal coordinates (bond lengthes, bond angles, torsion angles, etc.) and empirical rules for force constants are well established. For instance, the bond length force constant (bond distance d) often follow the empirical Badger rule [16,18,19]

$$D = \frac{A}{(d - B)^3} \tag{7}$$

with constants A and B depending on the involved chemical elements. For main group elements of the 1st, 2nd, and 3rd row in the periodic table, these constants depend mostly on the row of the corresponding element [16,18]. Other rules for deriving force constants have been suggested, too, e.g. by Fischer [20] and Lindh [21].

To summarize, a quasi-Newton scheme consists of (1) the coordinate system, (2) the initial Hessian, (3) the updating scheme for the Hessian, and (4) the step size control. These ingredients are interconnected and all influence the number of steps taken by the algorithm to reach the desired minimum.

Despite the enormous advantages that have been demonstrated for internal coordinates in molecular [1,16,17,22] and covalently bonded periodic systems [23–25], they have never become the coordinate system of choice for other materials, e.g., ionic and metallic ones. This might be related to the difficulty in finding an appropriate set of internal coordinates and rules for the initial force constant values for non-covalent bonding. For instance, Waser and Pauling investigated the validity of the Badger rule for a wide range of solid materials from elements throughout the periodic table [19]. They found that the rule is valid for subgroups of pure elements, but that notably for compounds no reasonably general trend was visible. In view of the wide range of bonding situations in solids – ranging from metallic over covalent to ionic – the failure to find a single rule for a wide range of materials is not surprising.