Broyden's method for self-consistent field convergence acceleration

You may also be interested in:

Self-consistent pseudopotential calculations of the equilibrium bulk properties of diamond-type semiconductors
G P Srivastava

A simple improved iteration scheme for electronic structure calculations
H Akai and P H Dederichs

Self-consistent non-local pseudopotential calculations for the ground-state properties of -Sn
G P Srivastava

Electronic structure and total energy of Si, Ge and -Sn by the self-consistent local pseudopotential method
G P Srivastava

Phase diagram of a five-state spin system
F Y Wu

A new method for diagonalising large matrices
D M Wood and A Zunger

The electronic band structures for zincblende and wurtzite BeO
K J Chang, S Froyen and M Cohen

Stress theorem in the determination of static equilibrium by the density functional method
P Gomes Dacosta, O H Nielsen and K Kunc

The atomic geometry of the ZnSe (110) surface: determination by total-energy methods
A C Ferraz and G P Srivastava

LETTER TO THE EDITOR

Broyden's method for self-consistent field convergence acceleration

G P Srivastava

Physics Department, New University of Ulster, Coleraine, N Ireland BT52 1SA, UK

Abstract. Broyden's method for self-consistent field convergence acceleration, recently used by Bendt and Zunger in the band-structure calculation of solids, is reviewed and a simpler computational scheme is presented.

Recently, Bendt and Zunger (1982) have derived a variational principle for obtaining the minimum of the density-functional total energy of solids and molecules. For their variational principle to be equivalent to the variational principle derived by Kohn and Sham (1965) these authors have obtained a sufficient condition:

$$F(\mu) \equiv \nu - \mu = 0, \tag{1}$$

where μ is an independent vector variable representing the electronic screening potential and $\nu \equiv \nu(\mu)$ is the output screening potential to be determined from μ. Equation (1) is also the condition for the self-consistent band-structure calculation of the system for which the total energy is required.

The system of simultaneous nonlinear equations $F = 0$ in equation (1) provides a cherished topic in numerical analysis (see e.g. Byrne and Hall 1973). Of several iterative methods available to solve systems of nonlinear equations, the most sophisticated one is Broyden's updated version of the quasi-Newton–Raphson iterative procedure (see, for example, Ralston and Rabinowitz 1978). In this procedure one approximates $F(\mu)$ at various iterations, writing for the mth iteration

$$F(\mu) \simeq F(\mu^{(m)}) + J^{(m)}(\mu - \mu^{(m)}). \tag{2}$$

Further, one assumes that for the left-hand side to vanish the right-hand side must also vanish, which yields the following quasi-Newton–Raphson method for the desired solution

$$\mu^{(m+1)} = \mu^{(m)} - [J^{(m)}]^{-1} F(\mu^{(m)}). \tag{3}$$

It is very well known that the success of this iterative procedure depends entirely on how close the initial approximation for the Jacobian $J^{(1)} = \partial F / \partial \mu^{(1)}$ is to the true solution: if the initial approximation is not quite close to the solution, often the iteration will not converge. Finding a good initial approximation to the Jacobian is a serious problem. For $\mu = \{\mu_1, \mu_2, \ldots, \mu_N\}$, with a large value of N (which is usually the case in the context of the present discussion), an explicit evaluation and inversion of the initial Jacobian may itself involve excessive computational effort and storage requirements.

Broyden's modification (Broyden 1965) of the quasi-Newton–Raphson iterative procedure is extremely helpful in achieving a fast convergence rate with a moderately good initial guess for the Jacobian. According to this scheme one generates an approximation $\mathcal{J}^{(1)}$ to the initial Jacobian $J^{(1)}$ and for $m > 1$ it is required that $\mathcal{J}^{(m)}$ satisfies the following two conditions

(i) $\qquad \boldsymbol{g} \equiv \boldsymbol{F}^{(m)} - \boldsymbol{F}^{(m-1)} - \mathcal{J}^{(m)}(\boldsymbol{\mu}^{(m)} - \boldsymbol{\mu}^{(m-1)}) = 0$ $\qquad\qquad\qquad$ (4)

(ii) $\qquad Q \equiv \|\mathcal{J}^{(m)} - \mathcal{J}^{(m-1)}\|^2$ be minimised. $\qquad\qquad\qquad\qquad$ (5)

The consideration (ii), subject to (i), can be made by using the method of Lagrange multipliers. Consider a functional

$$K = Q + \boldsymbol{\lambda}^T \boldsymbol{g}, \qquad\qquad\qquad (6)$$

where $\boldsymbol{\lambda}$ is a Lagrange multiplier vector. Then the variational principle $\partial K / \partial \mathcal{J}^{(m)} = 0$ gives

$$\mathcal{J}^{(m)} = \mathcal{J}^{(m-1)} + \tfrac{1}{2}\boldsymbol{\lambda}^T(\boldsymbol{\mu}^{(m)} - \boldsymbol{\mu}^{(m-1)}). \qquad\qquad\qquad (7)$$

When compared with equation (4), this gives the desired result for $\boldsymbol{\lambda}$, which when used in equation (7) yields the following updating procedure for $\mathcal{J}^{(m)}$

$$\mathcal{J}^{(m)} = \mathcal{J}^{(m-1)} + \frac{[\boldsymbol{F}^{(m)} - \boldsymbol{F}^{(m-1)} - \mathcal{J}^{(m-1)}(\boldsymbol{\mu}^{(m)} - \boldsymbol{\mu}^{(m-1)})](\boldsymbol{\mu}^{(m)} - \boldsymbol{\mu}^{(m-1)})^T}{(\boldsymbol{\mu}^{(m)} - \boldsymbol{\mu}^{(m-1)})^T(\boldsymbol{\mu}^{(m)} - \boldsymbol{\mu}^{(m-1)})}. \qquad (8)$$

This is called a rank-one updating procedure, as the rank of $\mathcal{J}^{(m)} - \mathcal{J}^{(m-1)}$ is one. Thus equation (3) is used with equation (8) to achieve a fast rate of iterative convergence.

The initial guess $\mathcal{J}^{(1)}$ can be made in accordance with the problem at hand. In the present context of band-structure calculation, it can be shown that the Jacobian is the negative of the dielectric kernel. Thus a simple guess for $\mathcal{J}^{(1)}$ may be the diagonal form of the Thomas–Fermi screening

$$\mathcal{J}^{(1)} = -(1 + k^2/G^2)\hat{I} = -(1/\alpha(G))\hat{I} \qquad\qquad\qquad (9)$$

where $k = (4k_F/\pi)^{1/2}$, k_F being the Fermi momentum, and G is a reciprocal lattice vector. An even simpler guess would be a diagonal constant matrix

$$\mathcal{J}^{(1)} = -(1/\alpha)\hat{I} \qquad\qquad\qquad (10)$$

where $0 < \alpha < 1$. The particular choices in equations (9) and (10) without the updating scheme of equation (8) are, respectively, Kerker's scheme (Kerker 1981) and the potential mixing scheme (Bendt and Zunger 1982, Dunlap 1982, Dederichs and Zeller 1983). The choice of α in equations (9) and (10) should be such that $\Sigma_{j=1}^{N} [F_j^{(2)}]^2 < \Sigma_{j=1}^{N} [F_j^{(1)}]^2$ is satisfied (Powell 1970).

Bendt and Zunger have used Broyden's scheme to achieve rapid self-consistency in the band-structure calculation of a few semiconductors. However, in many applications, when N is very large, the updating scheme of equation (8) is not desirable on two counts: (i) storage of N^2 elements of $\mathcal{J}^{(m)}$ may be excessive, and furthermore, (ii) the use of equation (3) requires inverting the huge matrix $\mathcal{J}^{(m)}$. Fortunately, both these problems can be dealt with if the following computational scheme is adopted.

Let $\mathcal{G}^{(m)}$ be the inverse of $\mathcal{J}^{(m)}$. Then using Broyden's second method, the updating formula for $\mathcal{G}^{(m)}$ is

$$\mathcal{G}^{(m)} = \mathcal{G}^{(m-1)} + \frac{[\boldsymbol{\mu}^{(m)} - \boldsymbol{\mu}^{(m-1)} - \mathcal{G}^{(m-1)}(\boldsymbol{F}^{(m)} - \boldsymbol{F}^{(m-1)})](\boldsymbol{F}^{(m)} - \boldsymbol{F}^{(m-1)})^T}{(\boldsymbol{F}^{(m)} - \boldsymbol{F}^{(m-1)})^T(\boldsymbol{F}^{(m)} - \boldsymbol{F}^{(m-1)})}. \qquad (11)$$

This equation, with the help of equation (9) or (10), can be expressed as

$$\mathscr{G}^{(1)} = D = -\beta \hat{I} \tag{12}$$

$$\mathscr{G}^{(m)} = D + \sum_{i=2}^{m} u^{(i)} v^{T(i)}. \tag{13}$$

With this equation (3) can be expressed as the following vector equation

$$\mu^{(m+1)} = \mu^{(m)} + \eta^{(m)}, \tag{14}$$

where

$$\eta^{(m)} = \beta F^{(m)} - \sum_{i=2}^{m} c_{mi} u^{(i)}. \tag{15}$$

In equations (13) and (15) we have

$$u^{(i)} = \beta(F^{(i)} - F^{(i-1)}) + (\mu^{(i)} - \mu^{(i-1)}) + \sum_{j=2}^{i-1} a_{ij} u^{(j)}$$

$$a_{ij} = v^{T(j)}(F^{(i)} - F^{(i-1)})$$

$$c_{mi} = v^{T(i)} F^{(m)}$$

$$v^{T(i)} = \frac{(F^{(i)} - F^{(i-1)})^T}{(F^{(i)} - F^{(i-1)})^T (F^{(i)} - F^{(i-1)})}. \tag{16}$$

Thus the iterative solution of $F(\mu) = 0$ can be obtained by using Broyden's method which in the present scheme involves only storing a few vectors of length N.

To test the convergence rate of Broyden's method we have calculated the electronic band structure and the total crystal energy of InP using a local version of the *ab initio* pseudo-potentials of Zunger and Cohen (1978). Exchange and correlation were treated in the local density formalism using the formulation of Hedin and Lundqvist (1971).

Figures 1(*a*), (*b*) show the results of convergence test for, respectively, the screening potential V_{scr} and the total energy E_{tot} with the fixed Jacobian schemes of equations (9) and (10). It is clearly seen that both V_{scr} and E_{tot} diverge if no mixing of input and output potentials is used: i.e. when $\mathscr{G}^{(m)} = \mathscr{G}^{(1)} = -\hat{I}$. The results start to converge, though slowly, even for a small mixing of input and output potentials. In the present case of study it is found that $\alpha = 0.5$ (a 50%–50% mixing of the potentials) gives a very fast convergence. The Thomas–Fermi approximation for the Jacobian gives a better convergence rate than using the simple mixing case with $\alpha = 0.8$ (20% input mixing) but shows somewhat slower convergence rate than using $\alpha = 0.5$. Very recently, Dederichs and Zeller (1983) have rigorously analysed the simple mixing procedure and have obtained limits on α for this procedure to converge. An optimum value of α may depend on how good the starting input potential μ is.

Broyden's method of updating the Jacobian was found to accelerate the convergence rates in both V_{scr} and E_{tot} for all the cases studied in figure 1 (except $\alpha = 1$). As can be inferred from equations (14) and (15), acceleration of the convergence will start at the third iteration. In figure 2 we show the comparison of Broyden's method with the potential mixing schemes using the fixed Jacobians $\mathscr{G}_{TF}^{(m)}$ and $\mathscr{G}^{(m)}$ ($\alpha = 0.8$). For a stability of 10^{-4} Ryd in E_{tot} Broyden's method saves three iterations in this 'easy' case of two atoms/cell. For difficult cases (e.g. in slab surface calculations) this method must prove very time saving on the computer. However, if a value of α has been
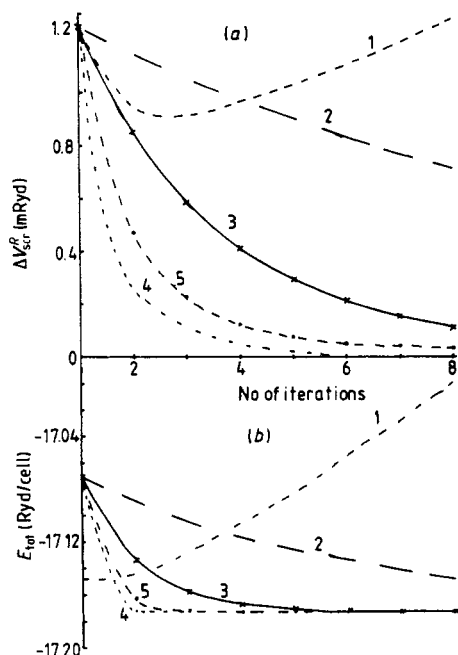
Fig. 1. Convergence studies on (*a*) the screening potential and (*b*) the total energy of InP using the fixed Jacobian scheme. Curves 1–4 use equation (10) with $\alpha = 1.0$, 0.95, 0.8 and 0.5 respectively. Curve 5 uses equation (9)
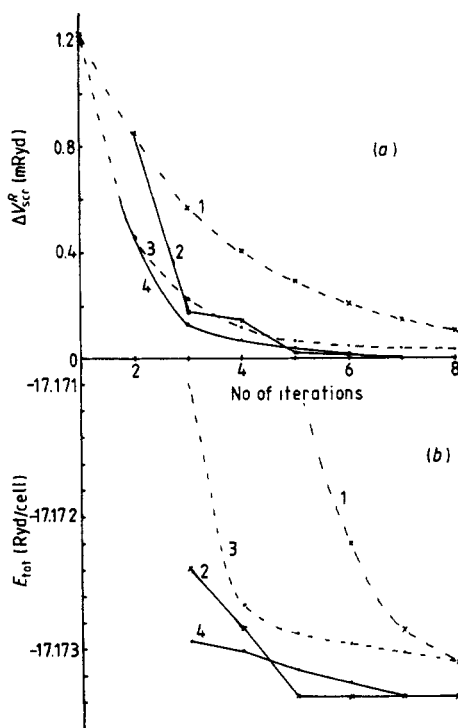
Figure 2. Convergence acceleration using Broyden's method. Curves 2 and 4 use equation (10) with $\alpha = 0.8$ and equation (9), respectively, for $\mathcal{J}^{(1)}$. Curves 1 and 3 correspond to curves 2 and 4 without Jacobian update.

found for which the simple mixing scheme converges rapidly, then Broyden's method may not offer much saving in computer time. Such is the case with $\mathcal{J}^{(1)}$ ($\alpha = 0.5$) in the present study, where Broyden's method saves only one (two) iteration(s) for a convergence of 10^{-5} (10^{-6}) Ryd in E_{tot}.

We believe that Broyden's method is superior to the recently proposed perturbative approach of Dunlap and the accelerated schemes presented in Dederichs and Zeller. Pulay (1980, 1982) has proposed an alternative approach for accelerating the convergence rate of the quasi-Newton–Raphson procedure.

References

Bendt P and Zunger A 1982 *Phys. Rev.* B **26** 3114
Broyden C G 1965 *Math. Comput.* **19** 577
Byrne G and Hall C A 1973 *Numerical Solution of Systems of Nonlinear Algebraic Equations* (New York: Academic)
Dederichs P H and Zeller R 1983 *Phys. Rev.* B **28** 5462

Dunlap B I 1982 *Phys. Rev.* A **25** 2847

Hedin L and Lundqvist B I 1971 *J. Phys. C: Solid State Phys.* **4** 2064

Kerker G P 1981 *Phys. Rev.* B **23** 3082

Powell M J D 1970 *Numerical Methods for Nonlinear Algebraic Equations* ed P Rabinowitz (New York: Gordon and Breach) p 87

Pulay P 1980 *Chem. Phys. Lett.* **73** 393

—— 1982 *J. Compt. Chem.* **3** 556

Ralston A and Rabinowitz P 1978 *A First Course in Numerical Analysis, 2nd edn* ch 8 (London: McGraw-Hill Kogakusha)

Zunger A and Cohen M L 1978 *Phys. Rev.* B **18** 5449