MATHEMATICS OF COMPUTATION, VOLUME 26, NUMBER 117, JANUARY, 1972

A Quasi-Newton Method with No Derivatives

By John Greenstadt

Abstract. The Davidon formula and others of the "quasi-Newton" class, which are used in the unconstrained minimization of a function f, provide a (generally) convergent sequence of approximations to the Hessian of f. These formulas, however, require the independent calculation of the gradient of f. In this paper, a set of new formulas is derived—using a previously described variational approach—which successively approximates the gradient as well as the Hessian, and uses only function values. These formulas are incorporated into an algorithm which, although still crude, works quite well for various standard test functions. Extensive numerical results are presented.

1. Introduction. The so-called variable-metric method for minimizing functions, which was discovered by Davidon [1] and developed by Fletcher and Powell [2], has been so successful that it has attracted a great deal of interest. Various theoretical studies, as well as new, related algorithms, have appeared in the literature ([3]–[6], among many others).

So far, all but one* of these variants of the DFP (Davidon-Fletcher-Powell) method have required the explicit evaluation, at each step, of the gradient of the function f to be minimized. From these computed gradients, the inverse of the Hessian matrix is gradually constructed, and the Newton formula (which is used to compute the next step direction) becomes gradually more accurate.

In a previous publication [7], it was shown how DFP-like formulas could be derived by solving a certain variational problem. In this paper, the same method will be applied to finding quasi-Newton** formulas which do not involve the explicit calculation of gradients. Clearly, since the gradient is needed in the Newton formula, the new algorithm will have to estimate it—as well as the Hessian— in the same way as the inverse Hessian is estimated in the DFP method.***

The basic notation to be used is as follows: $f(x)$ is the function of the variables $(x_1, x_2, \cdots, x_N)$ in R_N which is to be minimized; $\bar{g}$ and $\bar{G}$ are the gradient and Hessian of f, respectively. In the course of the work, certain estimates of these quantities will be discussed; these will be denoted by g and G (without bars). Further, $H \equiv G^{-1}$. At certain stages, vectors specifying directions for line searches are introduced; the letter d is used to denote these. When a direction vector d has been normalized (in a sense to be outlined later), the normalized direction is denoted by the letter s. Using a starting point x_0 and a unit direction s, a straight line in R_N may be expressed parametrically as follows:

Received July 27, 1970, revised July 26, 1971.

AMS 1970 *subject classifications*. Primary 90C30.

Key words and phrases. Quasi-Newton methods, gradient-free nonlinear optimization.

* This is the method of Stewart [4] which, however, computes the gradient by finite differences.

** The term "variable-metric" is reserved by convention for those methods in which the Hessian remains positive-definite (and hence can be regarded as a "metric" tensor).

*** A method due to Fiacco and McCormick [16] also estimates the gradient and Hessian using only function values. A comparison is made in Appendix B.

$$x(\alpha) = x_0 + \alpha s, \tag{1.1}$$

where α is a parameter which measures the distance from x_0 to $x(\alpha)$. If a line search along such a line has terminated at a certain value α_1, the displacement vector $x(\alpha_1) - x_0$ will be denoted by σ, so that $\sigma = \alpha_1 s$.

At appropriate places, subscripts may be appended to any of these symbols, to label the various steps with which they are associated. At other places, the context permitting, the subscripts will be dropped.

2. The Role of the Constraint in the DFP Case. In the DFP procedure, after the kth step from x_k to x_{k+1}, a new estimate H_{k+1} to the inverse Hessian is sought, which is to replace the current estimate H_k. This new estimate is required to satisfy the quasi-Newton condition (also known as the DFP condition):

$$H_{k+1}\bar{y}_k = \sigma_k, \tag{2.1}$$

where $\bar{y}_k$ is defined as $(\bar{g}_{k+1} - \bar{g}_k)$.

Where does this constraint come from? Basically, it is an *identity* which holds for *quadratic* functions. At the beginning of the kth iteration, we have a quadratic approximation to $f(x)$, say:

$$Q_k(x) = a_k + b_k^T x + \tfrac{1}{2}x^T G x, \tag{2.2}$$

(where the superscript T denotes the vector transpose) and, during this iteration, we make a step from the point x_k to a point x_{k+1}. At these two points, we have evaluated the exact gradient vectors:

$$\bar{g}_k \equiv \nabla f(x_k); \qquad \bar{g}_{k+1} \equiv \nabla f(x_{k+1}). \tag{2.3}$$

A new, improved quadratic approximation $Q_{k+1}(x)$ is now forced to fit $f(x)$ at these points, in the sense that the gradients calculated from $Q_{k+1}(x)$ match the exact ones:

$$g_{k+1}(x_k) = b_{k+1} + G_{k+1}x_k = \bar{g}_k, \tag{2.4a}$$

$$g_{k+1}(x_{k+1}) = b_{k+1} + G_{k+1}x_{k+1} = \bar{g}_{k+1}. \tag{2.4b}$$

It follows that the new G_{k+1} satisfies the condition:

$$\bar{g}_{k+1} - \bar{g}_k = G_{k+1}(x_{k+1} - x_k) \equiv G_{k+1}\sigma_k \tag{2.5}$$

which is equivalent to Eq. (2.1).

The method used in [7] to derive correction formulas was briefly as follows: The correction to H_k was written as:

$$H_{k+1} = H_k + E_k \tag{2.6}$$

and a quadratic norm of E_k was minimized subject to (2.1). (This amounts to a constraint on E_k.) In addition, it was required that E_k be symmetric so as to preserve the symmetry of H_{k+1}, given that of H_k. This amounts to another (linear) constraint on E_k. This constrained variational problem was solved, leading to a class of correction formulas. These formulas resemble the DFP formula, and it was, in fact, shown by D. Goldfarb [13] that the variationally derived class contains the DFP formula.

3. Constraints in the Derivativeless Case. We now have the task of translating the variational procedure to the case when there is no independently calculated gradient. The first thing we must do is to find an appropriate constraint corresponding to the QN condition.

Clearly, the new condition cannot contain $\bar{g}$ explicitly, since $\bar{g}$ cannot be independently computed. Hence, the only admissible ingredients are the values of f at various points.

As in all treatments of quasi-Newton methods, we assume $f(x)$ to be approximated by a quadratic function (as indicated previously). The approximation for f is Q, given in (2.2). If we replace b in favor of g, we obtain:

$$Q = a + g^T x - \tfrac{1}{2} x^T G x. \tag{3.1}$$

This form for $Q(x)$ has turned out, in practice, to be more convenient (less subject to rounding error) than that in (2.2), but it must be remembered that g depends on x.

Let us now assume that we are at some point x_0 and do a line search along some s (with length parameter α) for the minimum of f. For any x on this line ($x = x_0 + \alpha s$), we have for the estimate g, based on Q:

$$\begin{aligned} g &= b + Gx = b + G(x_0 + \alpha s) \\ &= (b + Gx_0) + G(\alpha s) = g_0 + \alpha G s. \end{aligned} \tag{3.2}$$

Correspondingly, for Q:

$$\begin{aligned} Q &= a + (g_0 + \alpha G s)^T (x_0 + \alpha s) - \tfrac{1}{2}(x_0 + \alpha s)^T G (x_0 + \alpha s) \\ &= (a + g_0^T x_0 - \tfrac{1}{2} x_0^T G x_0) + (g_0^T s)\alpha + \tfrac{1}{2}(s^T G s)\alpha^2. \end{aligned} \tag{3.3}$$

At some value α_1, we find the minimum value f_1. The corresponding x value is x_1 $(= x_0 + \alpha_1 s)$.

The spirit of the QN condition in the DFP case is to require that the estimated set of "parameters" $\{H_{ik}\}$ be such as to make the quadratic representation Q "fit" the independently computed *gradients*. What corresponds in the present case is to require the "parameters" g_0 and G to be such as to make the function $Q(x)$ "fit" the independently computed *values of* f. Thus, we shall require for our next estimates, g_0^* and G^*, say:

$$Q_0 \equiv Q(0) = a + g_0^{*T} x_0 - \tfrac{1}{2} x_0^T G^* x_0 = f_0, \tag{3.4a}$$

$$Q_1 \equiv Q(\alpha_1) = f_1. \tag{3.4b}$$

As in the DFP method, we eliminate what amounts to an additive constant (viz., a) by taking differences:

$$\Delta f = \Delta Q = Q_1 - Q_0 = (g_0^{*T} s)\alpha_1 + \tfrac{1}{2}(s^T G^* s)\alpha_1^2. \tag{3.5}$$

There is another independent constraint, based on the fact that f is a minimum at α_1. Hence, the derivative of Q, with respect to α, is forced to vanish at α_1:

$$\left(\frac{dQ}{d\alpha}\right)_{\alpha_1} = g_0^{*T} s + (s^T G^* s)\alpha_1 = s^T(g_0^* + \alpha_1 G^* s) = s^T g_1^* = 0. \tag{3.6}$$

Thus, we have *two* "QN conditions" at each step. Other combinations are possible, of course, such as fitting $Q(\alpha)$ to f at three distinct points along s. (This would also lead to two conditions.)

For reasons which will be apparent later, it is not feasible to attempt to correct g_0 and G after only one step. We therefore take more steps than one in each "correction cycle", and distinguish between a *minor step*, involving a line search along a single direction, and a *major step*, which will be a sequence of such minor steps.

In what follows, we shall suppress the major step index k, and concentrate on the set of minor steps which constitute a major step.

Starting from x_0 (the starting point of a major step), the first minor step direction d_1 is calculated by Newton's formula, using the current estimates g_0 and G:

$$d_1 = -G^{-1}g_0 \tag{3.7}$$

and d_1 is then normalized with respect to a positive-definite matrix L, to be chosen later. This gives the unit vector s_1, defined as follows:

$$s_1 \equiv d_1/(d_1^T L d_1)^{1/2}. \tag{3.8}$$

Note that it is necessary to solve a simultaneous linear system for d_1, since G^{-1} will not be directly estimated, as in the DFP method. The reason for this is that G is involved in Eqs. (3.5) in such a way, that replacing it by H^{-1} would unavoidably lead to a nonlinear constraint on H, thus rendering the variational problem intractable.

After the line search along s_1, yielding α_1 and f_1, the direction of the next minor step may be generated by combining s_1 with some other direction. A simple choice is one of the coordinate directions, say e_1. Then

$$d_2 = e_1 + \rho_1 s_1 \tag{3.9}$$

with ρ_1 chosen so as to make d_2 orthogonal to s_1, in the sense that $d_2^T L s_1 = 0$. d_2 is then normalized to give s_2, and a line search is performed, yielding α_2 and f_2. Next, a new direction d_3 is found by combining e_2, s_1 and s_2 linearly, and requiring d_3 to be orthogonal to s_1 and s_2 (with respect to L). d_3 is then normalized, etc.

If it should happen that one of the coordinate directions is a linear combination of the already computed direction vectors, it is simply dropped. In all, a total of N minor steps are attempted. In what follows, the index i will be a label for the minor steps within a major step.

If we denote the ith minor step by σ_i, we have:

$$x_i = x_{i-1} + \sigma_i. \tag{3.10}$$

τ_i is next defined as the total displacement from x_0 to x_i:

$$\tau_i \equiv x_i - x_0 = \sum_{j=1}^{i} \sigma_j. \tag{3.11}$$

Then, based on (3.2), we will impose the condition:

$$g_i^* = g_0^* + G^*\tau_i = g_{i-1}^* + G^*\sigma_i \tag{3.12}$$

and, corresponding to (3.6), we satisfy:

$$\sigma_i^T g_i^* = \sigma_i^T(g_0^* + G^*\tau_i) = 0 \tag{3.13}$$

for each τ_i.

Corresponding to (3.5), we have:

$$
\begin{aligned}
\Delta f_i &= f_i - f_{i-1} = Q_i - Q_{i-1} \\
&= g_i^{*T}(x_{i-1} + \sigma_i) - (g_i^* - G^*\sigma_i)^T x_{i-1} \\
&\quad - \tfrac{1}{2}(x_{i-1} + \sigma_i)^T G^*(x_{i-1} + \sigma_i) + \tfrac{1}{2}x_{i-1}^T G^* x_{i-1} \\
&= g_i^{*T}\sigma_i - \tfrac{1}{2}\sigma_i^T G^*\sigma_i = -\tfrac{1}{2}\sigma_i^T G^*\sigma_i,
\end{aligned} \tag{3.14}
$$

the last equation resulting from (3.13).

In summary, our constraints are:

$$\Delta f_i + \tfrac{1}{2}\sigma_i^T G^*\sigma_i = 0, \tag{3.15a}$$

$$\sigma_i^T g_0^* + \sigma_i^T G^*\tau_i = 0. \tag{3.15b}$$

It is important to note that the only independently computed functional quantities here are the $\{\Delta f_i\}$.

We are now going to consider the major step as an independent cycle, and make the corrections to our old estimates, g_0 and G, at the end of it. The corrections will be denoted by γ and Γ, so that the corrected values g_0^* and G^* will be:

$$g_0^* = g_0 + \gamma, \tag{3.16a}$$

$$G^* = G + \Gamma. \tag{3.16b}$$

Then the constraints (3.15), considered to apply to the new estimates g_0^* and G^*, are translated into constraints on γ and Γ as follows:

$$\tfrac{1}{2}\sigma_i^T\Gamma\sigma_i = -\{\Delta f_i + \tfrac{1}{2}\sigma_i^T G\sigma_i\} \equiv \rho_i, \tag{3.17a}$$

$$\sigma_i^T\gamma + \sigma_i^T\Gamma\tau_i = -\{\sigma_i^T g_0 + \sigma_i^T G\tau_i\} \equiv \epsilon_i. \tag{3.17b}$$

Now, there are N parameters in g_0 and $\tfrac{1}{2}N(N + 1)$ in G to be estimated. But in each major step, we have at most $2N$ constraints. Hence, when $N > 1$, there are fewer constraints than parameters; so that one major step does not determine all the parameters. Since each major step is treated independently of the others, any method based on these constraints will not necessarily be an "N-step" method. In fact, the formulas to be derived need not necessarily generate the exact G, even for quadratic functions. This is not to say, however, that it is impossible to construct "N-step" formulas (by other means).

4. The Variational Procedure for the Derivativeless Case. We now have the problem of setting up a functional to minimize, which somehow embodies, the norms of γ and of Γ. The most obvious norms to choose, which are quadratic, are:

$$||\gamma||^2 \equiv \gamma^T V\gamma, \tag{4.1a}$$

$$||\Gamma||^2 \equiv \mathrm{Tr}(W\Gamma W\Gamma^T), \tag{4.1b}$$

where V and W are positive-definite matrices of some sort.

A difficulty arises in somehow combining these norms in a natural manner. One wishes to have a quadratic function of the elements of γ and Γ which is also positive-definite. These two quantities are not really comparable, since it is easy to construct functions for which they have arbitrary values. The obvious device of simply adding them leads to the problem of insuring that their "units" are consistent. This might be accomplished, for example, by taking $W = G^{-1}$ and $V = ||g_0||^{-2}I$, where I is the unit matrix.

The most practical form, which was found after some trials, was the most obvious one, viz., a simple sum:

$$\Phi_0 = \tfrac{1}{2}\gamma^T V\gamma + \tfrac{1}{2}\,\mathrm{Tr}(W\Gamma W\Gamma^T), \tag{4.2}$$

and a large number of numerical trials, wherein various forms of V and W were chosen, seemed to indicate that the choices $V = I$, $W = \nu I$ (where ν is some arbitrary number) worked best in practice. However, we shall defer this specialization to a later section, but leave V and W arbitrary so as to show the general form of the corrections.

Incorporating the constraints (3.17) into the functional via the Lagrange multipliers $\{\eta_i\}$ and $\{\theta_i\}$ gives:

$$\begin{aligned}\Phi = \Phi_0 &- \sum_i \eta_i(\tfrac{1}{2}\sigma_i^T\Gamma\sigma_i - \rho_i)\\ &- \sum_i \theta_i(\sigma_i^T\gamma + \sigma_i^T\Gamma\tau_i - \epsilon_i).\end{aligned} \tag{4.3}$$

We should add to this the additional constraint $\Gamma^T = \Gamma$, but will dispense with doing this explicitly, and simply indicate the change in the formula for Γ, necessary to include this requirement.

The necessary conditions for a stationary Φ are obtained by differentiating, as follows:

$$\frac{\partial\Phi}{\partial\gamma} = V\gamma - \sum_i \theta_i\sigma_i = 0, \tag{4.4a}$$

$$\frac{\partial\Phi}{\partial\Gamma} = W\Gamma W - \sum_i \eta_i\cdot\tfrac{1}{2}\sigma_i\sigma_i^T - \tfrac{1}{2}\sum_i \theta_i(\sigma_i\tau_i^T + \tau_i\sigma_i^T) = 0. \tag{4.4b}$$

(The symmetrizing of the $\sigma_i\tau_i^T$ term is a result of taking account of the symmetry condition on Γ.)

If we define $\Lambda \equiv V^{-1}$, $M \equiv W^{-1}$, we have:

$$\gamma = \Lambda \sum \theta_i\sigma_i, \tag{4.5a}$$

$$\Gamma = \tfrac{1}{2}M\{\sum \eta_i\sigma_i\sigma_i^T + \sum \theta_i(\sigma_i\tau_i^T + \tau_i\sigma_i^T)\}M. \tag{4.5b}$$

We now solve for the Lagrange multipliers $\{\eta_i\}$ and $\{\theta_i\}$ by applying the constraints to γ and Γ. The resulting equations are rather complicated, but they reduce to the following (in matrix form):

$$A\theta + B\eta = \epsilon, \qquad B^T\theta + C\eta = \rho, \tag{4.6}$$

where

$$\epsilon \equiv \{\epsilon_i\}, \qquad \rho \equiv \{\rho_i\}, \tag{4.7}$$

$$A_{ij} \equiv \lambda_{ij} + \tfrac{1}{2}\{\mu_{ij}^{(1)}\mu_{ij}^{(3)} + \mu_{ij}^{(2)}\mu_{ji}^{(2)}\}, \tag{4.8a}$$

$$B_{ij} \equiv \tfrac{1}{2}\mu_{ij}^{(2)}\mu_{ij}^{(3)}, \tag{4.8b}$$

$$C_{ij} \equiv \tfrac{1}{4}\mu_{ij}^{(3)}\mu_{ij}^{(3)} \tag{4.8c}$$

and

(4.9a) $$\lambda_{ij} \equiv \sigma_i^T \Lambda \sigma_j,$$

(4.9b) $$\mu_{ij}^{(1)} \equiv \tau_i^T M \tau_j,$$

(4.9c) $$\mu_{ij}^{(2)} \equiv \tau_i^T M \sigma_j,$$

(4.9d) $$\mu_{ij}^{(3)} \equiv \sigma_i^T M \sigma_j;$$

i and j run from 1 to N and are not summed in (4.8).

If M and Λ are now chosen to be proportional to L, we gain a great simplification in the formulas for γ and Γ. We set (as suggested previously):

(4.10) $$W = \nu V; \quad \text{or} \quad M = \frac{1}{\nu} \Lambda$$

and, in addition:

(4.11) $$\Lambda = L, \quad \text{so that } M = \frac{1}{\nu} L.$$

We then have, since $\{s_i\}$ is now an orthonormal set with respect to L:

(4.12a) $$\lambda_{ij} = \sigma_i^T \Lambda \sigma_j = |\sigma_i|\,|\sigma_j|\, s_i^T L s_j = \sigma_i^2 \delta_{ij}$$

and, similarly:

(4.12b) $$\mu_{ij}^{(3)} = \sigma_i^T M \sigma_j = \frac{1}{\nu} \sigma_i^T L \sigma_j = \frac{1}{\nu} \sigma_i^2\, \delta_{ij},$$

so that $\{\lambda_{ij}\}$ and $\{\mu_{ij}^{(3)}\}$ are diagonal. Since, from Eq. (3.11), $\tau_i = \sum_{p=1}^{i} \sigma_p$, we have:

(4.13)
$$\begin{aligned} \mu_{ij}^{(2)} &= \frac{1}{\nu} \sum_{p=1}^{i} \sigma_p^T L \sigma_j = \frac{1}{\nu} \sum_{p=1}^{i} \sigma_p^2\, \delta_{pj} \\ &= \sigma_j^2/\nu, \quad \text{if } i \geqq j, \\ &= 0, \qquad \text{if } i < j, \end{aligned}$$

so that $\{\mu_{ij}^{(2)}\}$ is a lower triangular matrix.

Bearing in mind that the products in Eq. (4.8) are *not* matrix products, but element-by-element products, we see that:

1. $\{\mu_{ij}^{(1)} \mu_{ij}^{(3)}\}$ is diagonal because $\{\mu_{ij}^{(3)}\}$ is;
2. $\{\mu_{ij}^{(2)} \mu_{ji}^{(2)}\}$ is diagonal because $\{\mu_{ij}^{(2)}\}$ is triangular;
3. $\{\mu_{ij}^{(2)} \mu_{ij}^{(3)}\}$ is diagonal because $\{\mu_{ij}^{(3)}\}$ is.

Hence, A_{ij}, B_{ij} and C_{ij} all form diagonal matrices, and have the values:

(4.14a) $$A_{ij} = \left\{\sigma_i^2 + \frac{1}{2\nu^2} (\tau_i^2 \sigma_i^2 + \sigma_i^4)\right\} \delta_{ij},$$

(4.14b) $$B_{ij} = \frac{1}{2\nu^2} \sigma_i^4\, \delta_{ij},$$

(4.14c) $$C_{ij} = \frac{1}{4\nu^2} \sigma_i^4\, \delta_{ij},$$

where

$$\tau_i^2 = \sum_{p=1}^{i} \sigma_p^2, \tag{4.15}$$

all of which follows from the orthonormality of $\{s_i\}$ with respect to L.

The solution of Eq. (4.6) has the form:

$$\theta = (A - BC^{-1}B^T)^{-1}(\epsilon - BC^{-1}\rho), \tag{4.16a}$$

$$\eta = C^{-1}(\rho - B^T\theta) \tag{4.16b}$$

and these expressions may be easily evaluated because all the matrices are diagonal. The result is (by components):

$$\theta_i = \frac{2\nu^2(\epsilon_i - 2\rho_i)}{\sigma_i^2(2\nu^2 + \tau_i^2 - \sigma_i^2)}, \tag{4.17a}$$

$$\eta_i = \frac{4\nu^2}{\sigma_i^4}\rho_i - 2\theta_i, \tag{4.17b}$$

so that the evaluation of γ and Γ does not really involve any matrix inversions.

The algorithm now runs as follows:

1. Assume $G = I$, and estimate g_0 at the starting point by first differences. (See explanation in Section 5.)
2. To start a major step, compute a direction s_1 from Eqs. (3.7), (3.8).
3. Do a line search for a minimum of f along s (for each minor step).
4. Save σ, τ, ρ and ϵ as defined in Section 3. If a total of N independent directions have been generated, skip to step 6.
5. Form a new direction from the previous step directions plus a new linearly independent direction, and orthonormalize. Go to step 3.
6. Compute θ and η from Eqs. (4.17).
7. Compute γ and Γ from (4.5).
8. Correct g_0 and G (Eq. (3.16)) to form g_0^* and G^*.
9. Translate g_0^* using $g_0^{**} = g_0^* + G^*\tau_N$ (referring to Eq. (3.12), since the new x_0^* is $x_0 + \tau_N$).

This completes a major step.

10. Test for termination ($\|g_0^{**}\| <$ threshold, say). Otherwise, go back to step 2.

There are the usual complications in the program for this algorithm, mostly as a result of rounding error. These have not been described here.

5. Computational Experience. This method was programmed in the APL language for the IBM 360 computer and a good many trials were run on a few test functions. There was a good deal of tinkering necessary to get the method to converge reliably and reasonably efficiently, but the most effective choice of various arbitrary quantities turned out to be one of the simplest.

The worst difficulty with this method is that the successive estimates of G are not necessarily positive-definite. This precludes setting $L = G$ (hence $\Lambda = G$ and $M = G/\nu$) since minimizing a quadratic form with an indefinite metric can (and did!) yield very large, unstable corrections γ or Γ. The choice $L = I$ turned out to be the most stable (and the simplest) choice, and almost always led to the fastest convergence.

The best choice of ν turned out to be 0! Of course, one cannot simply set $\nu = 0$, and evaluate γ and Γ, since Eqs. (4.6) become singular for $\nu = 0$. It is possible, of course, to find the limiting solution as $\nu \to 0$, and this is described in Appendix A.

In many instances, the correction computed in this way caused G to become indefinite. This is easily detectable in those cases when a diagonal element becomes negative. This was cleared up in most instances by letting $\nu \to \infty$, instead of $\nu \to 0$. (The former case is analyzed in Appendix A.) When this device did not help, the indefiniteness was allowed to remain, and the next major step was begun. Near the point of convergence, this pathological effect nearly always disappeared; however, it did have the effect of slowing down convergence.

As will be seen from the printouts of some of the examples shown, the convergence does seem to be superlinear in many cases. This has not been proven and may not even always be true.

There is certainly no assurance that a variational derivation will yield formulas having the most desirable properties. It is likely that a deeper theoretical analysis of this type of QN method will yield better procedures with better properties (such as positive-definite G's).

As in the DFP method, the unit matrix was taken as a starting value for G. For a starting value of g_0, there is no "natural" vector, although, in principle, it is possible to start with any vector. When this was done (for example, by taking $g_0 = (10000\cdots)$ or $g_0 = (111\cdots)$, the method converged, but often with great difficulty. Ultimately, a rough estimate of g_0 was computed at the outset (by simple forward differences), and this stabilized matters quite considerably.

6. **Numerical Examples.** Tables 1–3 following are printouts generated at a terminal by the APL program. The entries are as follows:

NSTEP	The major step number.
P	The number of minor steps in the major step; in these tables, $P = N$ in all cases, except when some minor steps are too small. (The formulas for θ and η remain the same, except that N is replaced by P.)
NFUNC	The total number of evaluations of f after each major step.
F	The value of $f(x)$.
X	The position vector.

In these printouts, g_0 is denoted by *GZ* and G is denoted by *GG*. When G is found to be indefinite, the notation: *IG* (indefinite G) with the major step number is printed. The value of ν is then changed from 0 to ∞. When this still gives a detectably indefinite G, the same notation is printed again. The entire process was regarded as having converged when $\|g_0\| < 10^{-5}$, or, failing this, that no minor step $> 10^{-7}$ was possible. If the size of the major step falls below 10^{-6}, the notation "*SPF*" is printed, and the iteration terminated.

The functions tested were as follows: (The starting values in each case are listed on the first line with $NSTEP = 0$.)

(a) *Quadratic Function* 1.

$$f = x_1^2 + 100(x_2 - 1)^2 + (x_3 - 2)^2$$

whose Hessian is equal to:

$$G_1 = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 200 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

and minimum at (0, 1, 2). Various starting values were used.

(b) *Quadratic Function* 2.

$$f = (x_1 + x_2 - 2)^2 + 10^4(x_1 - x_2)^2$$

with Hessian:

$$G_2 = \begin{bmatrix} 20002 & -19998 \\ -19998 & 20002 \end{bmatrix}$$

and minimum at (1, 1).

(c) *Quadratic Function* 3.

$$f = (x_1 + 2x_2 + 3x_3)^2 + 100(x_2 - 1)^2 + (x_3 - 2)^2,$$

$$G_3 = \begin{bmatrix} 2 & 4 & 6 \\ 4 & 208 & 12 \\ 6 & 12 & 20 \end{bmatrix}$$

and minimum at (−8, 1, 2).

(d) *Rosenbrock's Function* [8].

$$f = 100(x_2 - x_1^2)^2 + (1 - x_1)^2,$$

$$G_{\mathrm{Ros}} = \begin{bmatrix} 802 & -400 \\ -400 & 200 \end{bmatrix} \quad \text{at } (1, 1).$$

(e) *Beales's Function* [9].

$$f = \sum_{i=1}^{3} [c_i - x_1(1 - x_2^i)]^2; \qquad \{c_i\} = \{1.5, 2.25, 2.625]$$

(Hessian not computed independently).

(f) *Powell's Function No.* 1 [10].

$$f = (x_1 + 10x_2)^2 + 5(x_3 - x_4)^2 + (x_2 - 2x_3)^4 + 10(x_1 - x_4)^4$$

(Hessian not computed independently).

(g) *Powell's Function No.* 2 [11].

$$f = [1 + (x_1 - x_2)^2]^{-1} + \sin(\tfrac{1}{2}\pi x_2 x_3) + \exp\left\{-\left(\frac{x_1 + x_3}{x_2} - 2\right)^2\right\}$$

(Hessian not computed independently).

(h) *Cube* [12].

$$f = 100(x_2 - x_1^3)^2 + (1 - x_1)^2,$$

$$G_{\mathrm{CUBE}} = \begin{bmatrix} 1802 & -600 \\ -600 & 200 \end{bmatrix} \quad \text{at } (1, 1).$$

(i) *Random Trigonometric Function* [2].

$$f = \sum_{i=1}^{N} \left\{ E_i - \sum_{j=1}^{N} (A_{ij} \sin x_j + B_{ij} \cos x_j) \right\}^2$$

(with A_{ij}, B_{ij}, and E_i randomly generated).

G_{RT} is variable and the solution is "*XNULL*", which is precomputed.

(j) *Helical Valley* [2].

$$f = 100[(x_3 - 10\theta)^2 + (r - 1)^2] + x_3^2$$

with

$$\theta = \tan^{-1}(x_2/x_1);\ r = (x_1^2 + x_2^2)^{1/2}.$$

Solution: (1, 0, 0).

(k) *Wood's Function* [15].

$$\begin{aligned} f = {} & 100(x_2 - x_1^2)^2 + (1 - x_1)^2 + 90(x_4 - x_3^2)^2 \\ & + (1 - x_3)^2 + 10.1\{(x_2 - 1)^2 + (x_4 - 1)^2\} + 19.8(x_2 - 1)(x_4 - 1). \end{aligned}$$

Solution: (1, 1, 1, 1).

It will be seen that various interesting (some good and some bad) things occur in these problems:

(1) The convergence near the solution is often clearly superlinear (even quadratic at times), but breaks down for functions which do not have a quadratic minimum (e.g., Powell 1).

(2) When G at the solution is singular, there is a good deal of difficulty with indefinite intermediate G's, which slows the convergence drastically.

(3) This method is not as speedy as several others (Simplex, Powell's, Rosenbrock's) but compares well in some cases.

(4) The successive estimates of G have been printed for Quadratic Function No. 3; evidently, a good value is generated very soon, which explains the quite rapid convergence in the quadratic cases. (A similar study of what happens to g_0 has not been made.)

(5) When P, the number of minor steps per major step is restricted to be $<N$, the convergence is slowed considerably. (These cases are not shown.) When $P = 1$, the correction to g_0 tends to make it vanish altogether, thus providing no direction for the next Newton step. (This was the reason for introducing additional minor steps in the first place.)

In Table 4 is shown a comparison with other methods for those test functions for which information is available. The starting points for all comparison functions are the "standard" ones, i.e., those used most in the literature.

The entries in Table 4 are as follows:

QNWD	stands for "Quasi-Newton Without Derivatives"
H-J	stands for "Hooke and Jeeves"
Ros	stands for "Rosenbrock"
SPLX	stands for "Simplex"
Pow	stands for "Powell"
Stew	stands for "Stewart".

For each case, the number of function evaluations needed to get the function

Table 1

```
QUADR. FUNC 2

NSTEP  P  NFUNC      F             X→
  0    0     3   1.0001E08    1.0000E02   0.0000E00
  1    2    15   9.3134E¯01   1.0096E00   9.9993E¯01
  2    1    25   2.3272E¯09   1.0000E00   1.0000E00
  3    1    35   8.9420E¯17   1.0000E00   1.0000E00
   CONVERGED
  3    0    45   8.9414E¯17   1.0000E00   1.0000E00
 GZ
¯0.47993  0.47984

 GG

  20002      ¯19998
 ¯19998       20002

QUADR. FUNC. 3

NSTEP  P  NFUNC      F             X→
  0    0     4   2.0100E02     3.0000E00   2.0000E00   1.0000E00
  1    3    15   1.2868E01    ¯4.4480E00   1.1904E00   1.6976E00
NIN[91]
    2.0573     3.8072      6.0731
    3.8072   208          12.065
    6.0731    12.065      19.942
  2    3    25   1.4396E¯02   ¯7.6544E00   9.9721E¯01  1.9123E00

    1.9107     4.4548      5.8024
    4.4548   208.09       11.949
    5.8024    11.949      20.003
  3    3    34   1.9062E¯05   ¯7.9869E00   1.0000E¯00  1.9956E00

    2.0031     4.0165      6.0058
    4.0165   208          12.056
    6.0058   .12.056      19.997
  4    2    47   1.1494E¯11   ¯8.0000E00 · 1.0000E¯00  2.0000E00

    1.9982     4.0165      5.9977
    4.0165   208          12.056
    5.9977    12.056      20.002
  5    1    60   3.4065E¯17   ¯8.0000E00   1.0000E00   2.0000E00
•SPF 5
   CONVERGED
  5    1    60   3.4065E¯17   ¯8.0000E00   1.0000E00   2.0000E00
 GZ
¯4.3438E¯9  ¯2.1134E¯7  8.6236E¯10

 GG

    1.9985     4.0165      5.9976
    4.0165   208          12.056
    5.9976    12.056      20.002


QUADR. FUNC. 3

NSTEP  P  NFUNC      F             X→
  0    0     4   8.6440E03    ¯1.0000E01   1.0000E01  ¯1.0000E01
  1    3    15   2.1194E02    ¯8.2552E00   2.4072E00   8.0563E¯02
  2    3    25   4.3574E¯03   ¯7.8743E00   1.0018E00   1.9458E00
  3    3    34   6.0971E¯04   ¯7.9218E00   9.9985E¯01  1.9766E00
  4    3    43   3.0676E¯11   ¯8.0000E00   1.0000E¯00  2.0000E00
  5    1    56   1.0791E¯14   ¯8.0000E00   1.0000E¯00  2.0000E00
   CONVERGED
  5    1    56   1.0791E¯14   ¯8.0000E00   1.0000E¯00  2.0000E00
 GZ
7.2411E¯8  1.6262E¯6  ¯1.9152E¯8

 GG

    1.9978     3.8979      5.9977
    3.8979   208          11.659
    5.9977    11.659      20.002


QUADR. FUNC. 3

NSTEP  P  NFUNC      F             X→
  0    0     4   1.0104E04     1.0000E02   0.0000E00   0.0000E00
  1    3    18   2.8819E03     4.9717E01   4.2341E¯01  9.3294E¯01
  2    3    31   8.3338E¯03   ¯8.0426E00   1.0002E00   1.9841E00
  3    3    40   1.3427E¯06   ¯8.0037E00   1.0000E00   2.0011E00
  4    2    54   2.0308E¯13   ¯8.0000E00   1.0000E00   2.0000E00
   CONVERGED
  4    0    66   1.3073E¯13   ¯8.0000E00   1.0000E00   2.0000E00
 GZ
2.7719E¯7  1.8613E¯5  ¯6.874E¯8

 GG

    1.9998     3.9947      5.9997
    3.9947   208          12.011
    5.9997    12.011      20
```

TABLE 1 (*continued*)

```
  QUADR. FUNC. 1

NSTEP  P  NFUNC      F              X+
  0    0     4    1.1000E02     3.0000E00    2.0000E00    1.0000E00
  1    3    14    8.7319E¯01    2.6728E¯01   1.0891E00    1.9109E00
  2    2    31    1.4443E¯10   ¯9.4633E¯06   1.0000E¯00   2.0000E00
  3    2    49    2.5580E¯12   ¯1.5982E¯06   1.0000E00    2.0000E00
*SPF 3
    CONVERGED
  3    2    49    2.5580E¯12   ¯1.5982E¯06   1.0000E00    2.0000E00
   GZ
1.8011E¯5  0.00065156  5.7706E¯10

   GG

  2.0000E0   ¯1.4511E¯4   1.8140E¯4
 ¯1.4511E¯4   2.0000E2    1.3655E¯4
  1.8140E¯4   1.3655E¯4   2.0000E0


  QUADR. FUNC. 1

NSTEP  P  NFUNC      F              X+
  0    0     4    8.3440E03    ¯1.0000E01    1.0000E01   ¯1.0000E01
  1    3    15    6.8539E00    ¯2.8660E¯01   1.2579E00    1.6561E00
  2    1    27    1.9813E¯11    2.7736E¯06   1.0000E¯00   2.0000E00
  3    1    39    4.9839E¯14   ¯2.8514E¯10   1.0000E00    2.0000E00
*SPF 3
    CONVERGED
  3    1    39    4.9839E¯14   ¯2.8514E¯10   1.0000E00    2.0000E00
   GZ
¯4.3749E¯6  0.00039373  ¯5.2498E¯6

   GG

  2.0000E0    1.3838E¯5  ¯7.0444E¯7
  1.3838E¯5   2.0000E2    1.6035E¯5
 ¯7.0444E¯7   1.6035E¯5   2.0000E0


  QUADR. FUNC. 1

NSTEP  P  NFUNC      F              X+
  0    0     4    1.0104E04     1.0000E02    0.0000E00    0.0000E00
  1    3    15    9.3271E03     9.6079E01    3.9261E¯02   7.8474E¯02
  2    3    29    2.7746E¯06    1.1907E¯03   9.9999E¯01   1.9988E00
  3    1    43    1.6889E¯15    2.3114E¯09   1.0000E00    2.0000E00
    CONVERGED
  3    1    43    1.6889E¯15    2.3114E¯09   1.0000E00    2.0000E00
   GZ
7.8167E¯10  2.6501E¯7  ¯5.2693E¯10

   GG

  1.9998E0    2.9797E¯5   2.2098E¯4
  2.9797E¯5   2.0000E2   ¯6.6800E¯4
  2.2098E¯4  ¯6.6800E¯4   2.0007E0


   QUADR. FUNC 2

NSTEP  P  NFUNC      F              X+
  0    0     3    3.2405E02     1.0000E01    1.0001E01
  1    2     9    3.2367E02     9.9948E00    9.9958E00
  2    2    23    1.3747E00     1.5862E00    1.5862E00
  3    1    43    6.9562E¯18    1.0000E00    1.0000E00
    CONVERGED
  3    1    43    6.9562E¯18    1.0000E00    1.0000E00
   GZ
9.6235E¯7  ¯9.6235E¯7

   GG

    20002       ¯19998
   ¯19998        20002


   QUADR. FUNC 2

NSTEP  P  NFUNC      F              X+
  0    0     3    4.0000E06    ¯1.0000E01    1.0000E01
  1    2    12    2.2491E¯06    9.9999E¯01   1.0000E00
  2    1    20    1.3917E¯13    1.0000E00    1.0000E00
  '  CONVERGED
  2    0    37    1.3970E¯15    1.0000E00    1.0000E¯00
   GZ
0.0089104  0.01089

   GG

    20002       ¯19998
   ¯19998        20002
```

TABLE 2

ROSENBROCKS FUNCTION

NSTEP	P	NFUNC	F	X+	
0	0	3	2.4200E01	¯1.2000E00	1.0000E00
1	2	13	3.6298E00	¯8.9067E¯01	7.6981E¯01
2	2	21	3.4233E00	¯8.4879E¯01	7.2767E¯01
3	2	37	1.8095E00	¯3.3952E¯01	1.2760E¯01
4	2	49	1.6006E00	¯2.5919E¯01	7.9437E¯02
5	2	57	1.1819E00	¯7.8469E¯02	1.9865E¯02
6	2	65	8.8453E¯01	6.0462E¯02	7.8961E¯03
7	2	73	7.1738E¯01	1.5304E¯01	2.4091E¯02
8	2	81	5.6197E¯01	2.5041E¯01	6.1746E¯02
9	2	90	4.2998E¯01	3.4429E¯01	1.1797E¯01
10	2	99	3.1698E¯01	4.3715E¯01	1.9244E¯01
11	2	108	2.2182E¯01	5.3139E¯01	2.8710E¯01
12	2	119	1.3396E¯01	6.3433E¯01	4.0080E¯01
13	2	127	1.0543E¯01	6.7550E¯01	4.5744E¯01
14	2	140	4.4759E¯02	7.8846E¯01	6.2137E¯01
15	2	150	3.0951E¯02	8.2677E¯01	6.8662E¯01
16	2	163	4.4350E¯03	9.3398E¯01	8.7320E¯01
17	2	173	2.2415E¯03	9.5434E¯01	9.1203E¯01
18	2	183	2.9801E¯05	9.9458E¯01	9.8924E¯01
19	2	191	6.0225E¯06	9.9756E¯01	9.9515E¯01
20	2	199	4.2134E¯09	9.9994E¯01	9.9987E¯01
21	1	208	8.0028E¯11	9.9999E¯01	9.9998E¯01
CONVERGED					
21	1	208	8.0028E¯11	9.9999E¯01	9.9998E¯01

GZ

5.5108E¯8 ¯2.7734E¯8

GG

798.76	¯401.73
¯401.73	202.7

BEALES FUNCTION

NSTEP	P	NFUNC	F	X+	
0	0	3	1.4203E01	0.0000E00	0.0000E00
1	2	25	3.4995E¯01	2.1250E00	2.0891E¯01
2	2	33	8.7130E¯02	2.4689E00	3.2620E¯01
3	2	48	5.3580E¯04	2.9462E00	4.8493E¯01
4	2	57	1.2898E¯04	2.9728E00	4.9262E¯01
5	2	65	9.6146E¯07	2.9976E00	4.9939E¯01
6	2	71	2.2789E¯09	2.9999E00	4.9997E¯01
7	2	77	9.0499E¯13	3.0000E00	5.0000E¯01
CONVERGED					
7	2	77	9.0499E¯13	3.0000E00	5.0000E¯01

GZ

¯6.8265E¯6 ¯1.6972E¯6

GG

3.9448	¯12.984
¯12.984	45.54

POWELLS FUNCTION 2

NSTEP	P	NFUNC	F	X+		
0	0	4	¯1.5000E00	0.0000E00	1.0000E00	2.0000E00
1	3	20	¯2.4976E00	2.1958E¯01	8.6335E¯01	1.6225E00
2	3	33	¯2.8698E00	4.0748E¯01	7.7388E¯01	1.1885E00
3	3	42	¯2.8974E00	4.7880E¯01	8.0091E¯01	1.1542E00
4	3	57	¯2.9999E00	1.0192E00	1.0079E00	9.9288E¯01
5	3	74	¯3.0000E00	1.0051E00	1.0023E00	9.9848E¯01
6	3	85	¯3.0000E00	1.0044E00	1.0017E00	9.9850E¯01
*IG 6						
7	3	96	¯3.0000E00	1.0033E00	1.0019E00	9.9876E¯01
8	3	108	¯3.0000E00	1.0034E00	1.0015E00	9.9900E¯01
*IG 8						
9	3	119	¯3.0000E00	1.0026E00	1.0012E00	9.9936E¯01
10	3	133	¯3.0000E00	1.0000E00	1.0001E00	1.0000E00
11	3	144	¯3.0000E00	1.0001E00	1.0000E00	1.0000E¯00
*IG 11						
12	3	153	¯3.0000E0C	1.0000E00	1.0000E00	1.0000E¯00
13	3	175	¯3.0000E00	1.0000E00	1.0000E00	9.9999E¯01
*SPF 13						
CONVERGED						
13	3	175	¯3.0000E00	1.0000E00	1.0000E00	9.9999E¯01

GZ

2.0848E¯5 1.7105E¯5 ¯3.1785E¯5

GG

4.1302	¯8.0504	1.8054
¯8.0504	9.0137	¯0.57887
1.8054	¯0.57887	7.7914

TABLE 2 (*continued*)

```
   POWELLS FUNCTION  1

NSTEP  P  NFUNC       F             X+
  0    0     5    2.1500E02      3.0000E00  ¯1.0000E00   0.0000E00   1.0000E00
  1    4    27    3.0957E00      6.4603E¯01 ¯1.5640E¯01  5.2225E¯01  7.0824E¯01
  2    4    44    5.3349E¯01     5.8315E¯01 ¯6.4841E¯03  3.3092E¯01  3.0313E¯01
  3    4    68    1.9547E¯01     1.8400E¯01 ¯1.3570E¯02  3.1948E¯01  3.4914E¯01
*IG 3
  4    4    83    6.2625E¯02     2.1017E¯01 ¯1.8541E¯02  2.3993E¯01  2.4747E¯01
*IG 4
  5    4   100    2.9855E¯02     1.5177E¯01 ¯1.9142E¯02  1.9380E¯01  2.0664E¯01
*IG 5
  6    4   115    1.4494E¯02     1.3998E¯01 ¯1.2674E¯02  1.5501E¯01  1.8126E¯01
*IG 6
  7    4   131    1.1635E¯02     1.3174E¯01 ¯1.1921E¯02  1.5335E¯01  1.6854E¯01
*IG 7
  8    4   146    7.8424E¯03     1.2730E¯01 ¯1.2215E¯02  1.4251E¯01  1.4106E¯01
*IG 8
  9    4   164    5.2363E¯03     1.1401E¯01 ¯1.2318E¯02  1.2767E¯01  1.2962E¯01
*IG 9
 10    4   184    4.3417E¯04     1.1878E¯01 ¯1.1483E¯02  4.9608E¯02  5.3873E¯02
*IG 10
 11    4   203    2.9105E¯04     1.1095E¯01 ¯1.1066E¯02  5.1263E¯02  5.2515E¯02
*IG 11
 12    4   217    2.7144E¯04     1.0805E¯01 ¯1.0775E¯02  5.2268E¯02  5.2864E¯02
*IG 12
 13    4   233    2.6617E¯04     1.0789E¯01 ¯1.0772E¯02  5.1243E¯02  5.2414E¯02
 14    4   248    2.5260E¯04     1.0601E¯01 ¯1.0577E¯02  5.1421E¯02  5.2296E¯02
 15    4   265    2.3636E¯04     1.0319E¯01 ¯1.0284E¯02  5.1608E¯02  5.1714E¯02
*IG 15
 16    4   279    2.3047E¯04     1.0303E¯01 ¯1.0259E¯02  5.0944E¯02  5.1592E¯02
*IG 16
 17    4   294    2.2543E¯04     1.0257E¯01 ¯1.0235E¯02  5.0364E¯02  5.1364E¯02
*IG 17
 18    4   319    2.0639E¯04     1.0171E¯01 ¯1.0145E¯02  4.5911E¯02  4.7415E¯02
*IG 18
 19    4   335    1.8336E¯04     9.8776E¯02 ¯9.8570E¯03  4.5914E¯02  4.6433E¯02
*IG 19
 20    4   353    1.7972E¯04     9.7878E¯02 ¯9.6105E¯03  4.5585E¯02  4.6349E¯02
*IG 20
 21    4   372    1.7149E¯04     9.6105E¯02 ¯9.7499E¯03  4.5862E¯02  4.5903E¯02
*IG 21
 22    4   389    1.6443E¯04     9.5874E¯02 ¯9.5636E¯03  4.4902E¯02  4.5793E¯02
*IG 22
 23    4   404    1.6076E¯04     9.5388E¯02 ¯9.5188E¯03  4.4701E¯02  4.5438E¯02
 24    4   420    1.2269E¯04     8.9529E¯02 ¯8.9421E¯03  3.9303E¯02  4.0342E¯02
*IG 24
 25    4   438    1.0511E¯04     8.6293E¯02 ¯8.6240E¯03  3.8419E¯02  3.9032E¯02
*IG 25
 26    4   453    1.0091E¯04     8.5387E¯02 ¯8.5260E¯03  3.8288E¯02  3.8825E¯02
 27    4   468    9.6162E¯05     8.4156E¯02 ¯8.4060E¯03  3.8470E¯02  3.9061E¯02
 28    4   484    6.5522E¯05     7.1816E¯02 ¯7.1699E¯03  3.9382E¯02  3.9515E¯02
 29    4   499    5.7095E¯05     6.2700E¯02 ¯6.2611E¯03  3.9800E¯02  3.9842E¯02
 30    4   513    5.6054E¯05     6.0344E¯02 ¯5.9848E¯03  3.9858E¯02  3.9744E¯02
 31    4   526    5.3594E¯05     5.9125E¯02 ¯5.9026E¯03  3.9437E¯02  3.9761E¯02
 32    4   542    5.0555E¯05     5.9026E¯02 ¯5.9094E¯03  3.8668E¯02  3.9105E¯02
 33    4   560    4.7046E¯05     5.8801E¯02 ¯5.8814E¯03  3.8031E¯02  3.8211E¯02
 34    4   579    3.8032E¯05     5.6198E¯02 ¯5.6099E¯03  3.5961E¯02  3.6201E¯02
 35    4   594    2.3221E¯05     5.0366E¯02 ¯5.0289E¯03  3.0844E¯02  3.1497E¯02
 36    4   610    4.0940E¯06     3.6028E¯02 ¯3.6013E¯03  1.9613E¯02  1.9598E¯02
*IG 36
 37    3   633    4.0560E¯06     3.6040E¯02 ¯3.6026E¯03  1.9532E¯02  1.9563E¯02
*IG 37
 38    4   650    3.6604E¯06     3.5698E¯02 ¯3.5732E¯03  1.8702E¯02  1.8759E¯02
 39    4   665    2.5530E¯06     3.3991E¯02 ¯3.3985E¯03  1.4181E¯02  1.4209E¯02
 40    4   678    2.3721E¯06     3.2601E¯02 ¯3.2598E¯03  1.1885E¯02  1.1897E¯02
 41    4   690    1.8674E¯06     2.9048E¯02 ¯2.9048E¯03  8.7539E¯03  8.8118E¯03
 42    4   705    1.5262E¯06     2.2480E¯02 ¯2.2508E¯03  2.7218E¯03  2.7756E¯03
 43    4   725    1.2772E¯06     1.2637E¯02 ¯1.2636E¯03 ¯6.2196E¯03 ¯6.2068E¯03
 44    4   741    1.2302E¯06     1.0228E¯02 ¯1.0229E¯03 ¯8.2679E¯03 ¯8.2748E¯03
 45    4   755    7.7647E¯07     8.0290E¯03 ¯8.0304E¯04 ¯8.3232E¯03 ¯8.3057E¯03
 46    4   771    2.9589E¯07     3.2318E¯03 ¯3.2251E¯04 ¯8.7766E¯03 ¯8.7689E¯03
 47    4   787    1.6264E¯07    ¯6.4599E¯04  6.4978E¯05 ¯9.1030E¯03 ¯9.1040E¯03
 48    3   804    1.4365E¯07    ¯2.5199E¯03  2.5194E¯04 ¯9.2401E¯03 ¯9.2441E¯03
 49    4   817    1.4007E¯07    ¯3.6216E¯03  3.6255E¯04 ¯9.3060E¯03 ¯9.3051E¯03
 50    3   835    1.3972E¯07    ¯3.6343E¯03  3.6329E¯04 ¯9.3017E¯03 ¯9.3009E¯03
 51    3   851    1.3941E¯07    ¯3.6243E¯03  3.6229E¯04 ¯9.2958E¯03 ¯9.2952E¯03
 52    3   866    1.3909E¯07    ¯3.5903E¯03  3.5889E¯04 ¯9.2881E¯03 ¯9.2874E¯03
 53    3   885    1.3714E¯07    ¯3.8647E¯03  3.8633E¯04 ¯9.2749E¯03 ¯9.2743E¯03
 54    4   899    1.3393E¯07    ¯4.3938E¯03  4.3925E¯04 ¯9.2451E¯03 ¯9.2446E¯03
 55    4   913    1.2288E¯07    ¯6.2465E¯03  6.2464E¯04 ¯9.0374E¯03 ¯9.0374E¯03
 56    2   935    1.0949E¯07    ¯1.2087E¯02  1.2085E¯03 ¯8.4541E¯03 ¯8.4562E¯03
 57    3   958    1.0930E¯07    ¯1.1855E¯02  1.1854E¯03 ¯8.4712E¯03 ¯8.4713E¯03
 58    3   978    1.0906E¯07    ¯1.1852E¯02  1.1852E¯03 ¯8.4662E¯03 ¯8.4665E¯03
*SPF 58
     CONVERGED
 58    3   978    1.0906E¯07    ¯1.1852E¯02  1.1852E¯03 ¯8.4662E¯03 ¯8.4665E¯03
   GZ
¯5.8332E¯5  ¯0.00023982  ¯5.7645E¯5  ¯3.1733E¯6

   GG

   2.9281      20.314      ¯1.6214       0.69141
  20.314      198.08       ¯4.1385      10.418
  ¯1.6214      ¯4.1385      9.8807     ¯13.095
   0.69141     10.418     ¯13.095        9.4135
```

TABLE 3

```
    CUBE

NSTEP ·P NFUNC      F            X+
  0    0    3    7.4904E02    ¯1.2000E00   1.0000E00
  1    2   14    1.2587E00    ¯1.2067E¯01  3.5058E¯03
  2    2   24    1.1820E00    ¯8.2517E¯02  9.5268E¯03
*IG 2
  3    2   31    1.0345E00    ¯1.5647E¯02  5.4047E¯03
  4    2   42    5.3092E¯01    2.7137E¯01  2.0427E¯02
  5    2   53    4.1518E¯01    3.5567E¯01  4.4593E¯02
  6    2   62    3.1294E¯01    4.4062E¯01  8.4976E¯02
  7    2   71    2.4084E¯01    5.0925E¯01  1.3220E¯01
  8    2   81    1.8317E¯01    5.7253E¯01  1.8978E¯01
  9    2   90    1.3738E¯01    6.3374E¯01  2.6022E¯01
 10    2  102    8.6609E¯02    7.0583E¯01  3.5079E¯01
 11    2  112    7.5453E¯02    7.2538E¯01  3.8226E¯01
 12    2  126    4.4514E¯02    7.8903E¯01  4.9146E¯01
 13    2  138    3.4141E¯02    8.1536E¯01  5.4136E¯01
 14    2  148    2.4578E¯02    8.4331E¯01  5.9923E¯01
 15    2  158    1.6041E¯02    8.7337E¯01  6.6595E¯01
 16    2  168    1.0048E¯02    8.9977E¯01  7.2830E¯01
 17    2  178    5.8923E¯03    9.2324E¯01  7.8687E¯01
 18    2  188    3.1131E¯03    9.4421E¯01  8.4175E¯01
 19    2  198    1.4142E¯03    9.6239E¯01  8.9137E¯01
 20    2  208    5.0608E¯04    9.7750E¯01  9.3402E¯01
 21    2  217    1.2327E¯04    9.8890E¯01  9.6706E¯01
 22    2  226    1.5146E¯05    9.9611E¯01  9.8837E¯01
 23    2  234    4.8050E¯07    9.9931E¯01  9.9792E¯01
 24    2  242    1.9537E¯09    9.9996E¯01  9.9987E¯01
 25    1  254    5.2986E¯15    1.0000E¯00  1.0000E¯00
     CONVERGED
 25    1  254    5.2986E¯15    1.0000E¯00  1.0000E¯00
    GZ
2.9499E¯8  ¯9.8342E¯9

    GG

    1800.2      ¯600.08
    ¯600.08      200.25



   RANDOM TRIG. FUNCTION
    XNULL        ¯0.48662  ¯1.3003  ¯0.89519

NSTEP  P NFUNC     F            X+
  0    0    4   5.7102E02   ¯4.3043E¯01 ¯1.3236E00  ¯6.1594E¯01
  1    3   13   9.9780E00   ¯4.9704E¯01 ¯1.4301E00  ¯7.7209E¯01
  2    3   27   8.0360E00   ¯4.9460E¯01 ¯1.3932E00  ¯7.8594E¯01
*IG 2
  3    3   38   4.5219E00   ¯4.9302E¯01 ¯1.3809E00  ¯8.1216E¯01
  4    3   48   3.6388E00   ¯4.9345E¯01 ¯1.3715E00  ¯8.2182E¯01
  5    3   58   1.4680E00   ¯4.9145E¯01 ¯1.3433E00  ¯8.5029E¯01
  6    3   68   1.1072E¯02  ¯4.8701E¯01 ¯1.3039E00  ¯8.9160E¯01
  7    3   77   4.9282E¯03  ¯4.8688E¯01 ¯1.3027E00  ¯8.9272E¯01
  8    3   94   1.8877E¯04  ¯4.8668E¯01 ¯1.3008E00  ¯8.9471E¯01
  9    3  105   7.4673E¯07  ¯4.8662E¯01 ¯1.3004E00  ¯8.9516E¯01
 10    3  115   1.2834E¯08  ¯4.8662E¯01 ¯1.3004E00  ¯8.9519E¯01
 11    1  130   1.5718E¯09  ¯4.8662E¯01 ¯1.3003E00  ¯8.9519E¯01
*SPF 11
    CONVERGED
 11    1  130   1.5718E¯09  ¯4.8662E¯01 ¯1.3003E00  ¯8.9519E¯01
   GZ
¯0.00076759  0.00072181  0.00047837

   GG

   30483       ¯2505.7       891.99
   ¯2505.7      21567       15387
     891.99     15387       13091
```

TABLE 3 (*continued*)

HELICAL VALLEY

NSTEP	P	NFUNC	F	X+		
0	0	4	2.5000E03	¯1.0000E00	0.0000E00	0.0000E00
1	3	28	2.2316E01	¯9.5036E¯01	6.7678E¯01	4.1635E00
2	3	39	1.6428E01	¯8.3528E¯01	6.2382E¯01	4.0148E00
*IG 2						
3	3	49	1.5238E01	¯7.8195E¯01	6.3790E¯01	3.9014E00
4	3	60	1.3986E01	¯6.9259E¯01	7.0984E¯01	3.7381E00
5	3	78	7.4964E00	¯1.5041E¯01	9.7597E¯01	2.7332E00
6	3	93	6.5887E00	¯5.0008E¯02	1.0018E00	2.5518E00
7	3	115	2.7227E00	5.2671E¯01	8.6135E¯01	1.6409E00
8	3	125	2,5228E00	5,5464E¯01	8,5487E¯01	1,5731E00
*IG 8						
9	3	143	2.1290E00	6.3925E¯01	7.9759E¯01	1.4370E00
10	3	154	1.7566E00	6.7228E¯01	7.4012E¯01	1.3253E00
11	3	168	1.3967E00	7.3603E¯01	6.4335E¯01	1.1547E00
12	3	190	4.9120E¯01	9.1941E¯01	4.2433E¯01	6.8933E¯01
13	3	207	2.1926E¯01	9.6011E¯01	2.8562E¯01	4.6523E¯01
14	3	217	1.2787E¯01	9.8189E¯01	2.1753E¯01	3.3623E¯01
15	3	235	7.3891E¯02	9.8975E¯01	1.6897E¯01	2.6874E¯01
16	3	246	1.6776E¯02	9.9673E¯01	8.1695E¯02	1.2904E¯01
17	3	258	9.9479E¯03	9.9816E¯01	6.2932E¯02	9.9053E¯02
18	3	268	3.6942E¯03	9.9912E¯01	3.8264E¯02	6.0732E¯02
19	3	280	1.3546E¯03	9.9984E¯01	2.2736E¯02	3.6579E¯02
20	3	291	3.8223E¯04	9.9393E¯01	1.2320E¯02	1.9538E¯02
21	3	304	6.4540E¯05	9.9999E¯01	5.0592E¯03	8.0306E¯03
22	3	313	9.3914E¯06	1.0000E¯00	1.9152E¯03	3.0616E¯03
23	3	325	3.0591E¯06	1.0000E¯00	1.1044E¯03	1.7414E¯03
24	3	335	9.5342E¯07	1.0000E¯00	6.1509E¯04	9.7598E¯04
25	3	346	2.9451E¯07	1.0000E00	3.3988E¯04	5.4246E¯04
26	3	358	4.5557E¯08	1.0000E00	1.3473E¯04	2.1288E¯04
27	3	369	6.4842E¯09	1.0000E¯00	5.0713E¯05	8.0496E¯05
28	3	378	1.1147E¯09	1.0000E¯00	2.0959E¯05	3.3386E¯05
29	2	394	3.5719E¯10	1.0000E¯00	1.1934E¯05	1.8812E¯05
30	1	412	8.2403E¯12	1.0000E¯00	1.7881E¯06	2.8380E¯06
CONVERGED						
30	0	424	7.9056E¯12	1.0000E¯00	1.7248E¯06	2.7451E¯06

GZ

2.4249E¯7 1.1199E¯5 ¯6.7899E¯6

GG

200.04	¯118.89	71.555
¯118.89	520.95	¯322.86
71.555	¯322.86	202.25

WOODS FUNCTION

NSTEP	P	NFUNC	F	X+			
0	0	5	1.9192E04	¯3.0000E00	¯1.0000E00	¯3.0000E00	¯1.0000E00
1	4	26	3.6916E01	1.3853E¯01	3.6686E¯01	¯2.2318E¯01	1.9226E¯01
2	4	52	1.0818E01	2.0079E¯01	1.2592E¯01	¯1.1184E00	1.1970E00
*IG 2							
3	4	71	8.3398E00	3.6494E¯01	1.8699E¯01	¯1.1208E00	1.2626E00
*IG 3							
4	4	85	7.7425E00	4.3428E¯01	2.2502E¯01	¯1.1071E00	1.2945E00
*IG 4							
5	4	99	7.0012E00	5.0987E¯01	2.5782E¯01	¯1.1190E00	1.3027E00
6	4	116	6.2782E00	5.2354E¯01	3.0096E¯01	¯1.1571E00	1.3530E00
7	4	130	5.9895E00	5.4113E¯01	3.1761E¯01	¯1.1778E00	1.3892E00
*IG 7							
8	4	144	5.8308E00	5.7212E¯01	3.2293E¯01	¯1.1742E00	1.4075E00
*IG 8							
9	4	159	5.7411E00	5.6489E¯01	3.3502E¯01	¯1.1856E00	1.4150E00
*IG 9							
10	4	179	5.4569E00	6.0393E¯01	3.5410E¯01	¯1.2355E00	1.5222E00
*IG 10							
11	4	193	5.4006E00	6.0036E¯01	3.6966E¯01	¯1.2370E00	1.5422E00
12	4	206	5.3824E00	6.0618E¯01	3.7343E¯01	¯1.2358E00	1.5387E00
13	4	223	5.3669E00	6.1359E¯01	3.7910E¯01	¯1.2386E00	1.5365E00
14	4	237	5.3471E00	6.1642E¯01	3.8586E¯01	¯1.2331E00	1.5300E00
*IG 14							
15	4	252	5.3216E00	6.2758E¯01	3.9249E¯01	¯1.2312E00	1.5211E00
16	4	269	5.2798E00	6.4026E¯01	4.1279E¯01	¯1.2324E00	1.5231E00
17	4	286	5.2274E00	6.7152E¯01	4.4786E¯01	¯1.2351E00	1.5274E00
18	4	298	5.2080E00	6.7482E¯01	4.5675E¯01	¯1.2308E00	1.5243E00
19	4	317	4.9765E00	7.2372E¯01	5.2878E¯01	¯1.1497E00	1.3258E00
20	4	341	3.9000E00	1.0574E00	1.1229E00	¯7.7958E¯01	6.5765E¯01
21	4	360	3.6285E00	1.1004E00	1.2100E00	¯7.3603E¯01	5.5792E¯01
22	4	377	3.2538E00	1.1631E00	1.3806E00	¯6.5712E¯01	4.4784E¯01
23	4	389	3.1268E00	1.1777E00	1.3958E00	¯6.9536E¯01	4.9994E¯01
24	4	401	3.0794E00	1.1905E00	1.4191E00	¯6.9126E¯01	4.9051E¯01
25	4	417	2.9547E00	1.2413E00	1.5458E00	¯6.6256E¯01	4.4919E¯01
26	4	435	2.8254E00	1.2839E00	1.6382E00	¯6.0128E¯01	3.7144E¯01
27	4	452	2.5623E00	1.3573E00	1.8338E00	¯4.5276E¯01	2.2204E¯01
28	4	468	2.5010E00	1.3441E00	1.8076E00	¯4.5745E¯01	2.0866E¯01
29	4	486	1.9926E00	1.3835E00	1.9158E00	¯2.1881E¯01	5.6012E¯02
30	4	510	1.7747E00	1.3980E00	1.9570E00	¯1.0740E¯01	2.2494E¯02
31	4	526	1.5361E00	1.4182E00	2.0041E00	2.3178E¯02	3.9660E¯03
32	4	539	1.3258E00	1.4139E00	1.9999E00	1.3001E¯01	2.1081E¯02
33	4	555	1.1472E00	1.4104E00	1.9909E00	2.3657E¯01	5.2709E¯02
34	4	574	1.1332E00	1.4111E00	1.9813E00	2.4626E¯01	5.6185E¯02

TABLE 3 (*continued*)

*IG 34							
35	4	596	1.0277E00	1.3837E00	1.9162E00	2.8151E¯01	6.5163E¯02
36	4	613	8.2455E¯01	1.3378E00	1.7916E00	3.7191E¯01	1.4946E¯01
37	4	632	7.5916E¯01	1.2964E00	1.6777E00	4.4681E¯01	2.0138E¯01
38	4	649	6.7945E¯01	1.2711E00	1.6346E00	4.9781E¯01	2.5375E¯01
39	4	674	4.6775E¯01	1.1978E00	1.4478E00	6.4818E¯01	4.1768E¯01
40	4	686	3.4407E¯01	1.2373E00	1.5324E00	6.3870E¯01	4.1517E¯01
41	4	700	3.3566E¯01	1.2476E00	1.5547E00	6.4013E¯01	4.0894E¯01
42	4	715	3.2760E¯01	1.2469E00	1.5565E00	6.4635E¯01	4.1351E¯01
*IG 42							
43	4	729	3.2277E¯01	1.2468E00	1.5570E00	6.4623E¯01	4.1788E¯01
*IG 43							
44	4	745	3.1248E¯01	1.2469E00	1.5558E00	6.5221E¯01	4.2778E¯01
45	4	763	2.7200E¯01	1.2417E00	1.5429E00	6.9021E¯01	4.7460E¯01
46	4	779	2.6988E¯01	1.2409E00	1.5366E00	6.8921E¯01	4.7337E¯01
47	4	794	2.6741E¯01	1.2391E00	1.5351E00	6.8879E¯01	4.7470E¯01
48	4	809	2.5585E¯01	1.2326E00	1.5199E00	6.9361E¯01	4.8172E¯01
49	4	823	2.2524E¯01	1.2223E00	1.4948E00	7.1959E¯01	5.1652E¯01
50	4	845	1.4381E¯01	1.1829E00	1.3943E00	7.9962E¯01	6.3764E¯01
51	4	861	1.3394E¯01	1.1752E00	1.3786E00	8.1302E¯01	6.6176E¯01
52	4	884	1.2542E¯01	1.1725E00	1.3698E00	8.0759E¯01	6.5058E¯01
53	4	896	1.2228E¯01	1.1701E00	1.3700E00	8.0565E¯01	6.4742E¯01
*IG 53							
54	4	913	1.2130E¯01	1.1702E00	1.3672E00	8.0594E¯01	6.4921E¯01
*IG 54							
55	4	929	1.2049E¯01	1.1689E00	1.3669E00	8.0493E¯01	6.4795E¯01
*IG 55							
56	4	944	1.1956E¯01	1.1688E00	1.3655E00	8.0645E¯01	6.4932E¯01
*IG 56							
57	4	959	1.1804E¯01	1.1673E00	1.3633E00	8.0736E¯01	6.5041E¯01
*IG 57							
58	4	977	1.1710E¯01	1.1673E00	1.3610E00	8.0857E¯01	6.5203E¯01
*IG 58							
59	4	993	1.1155E¯01	1.1631E00	1.3533E00	8.1279E¯01	6.6029E¯01
60	4	1007	1.0744E¯01	1.1602E00	1.3465E00	8.1577E¯01	6.6444E¯01
61	4	1023	8.4772E¯02	1.1341E00	1.2857E00	8.3330E¯01	6.9430E¯01
62	4	1036	7.6021E¯02	1.1143E00	1.2491E00	8.5059E¯01	7.2376E¯01
*IG 62							
63	4	1052	5.6929E¯02	1.1166E00	1.2471E00	8.6497E¯01	7.4832E¯01
64	4	1064	5.3193E¯02	1.1112E00	1.2352E00	8.6917E¯01	7.5685E¯01
65	4	1078	4.6602E¯02	1.0950E00	1.1985E00	8.8152E¯01	7.7748E¯01
66	4	1096	3.6398E¯02	1.0962E00	1.2024E00	8.9760E¯01	8.0689E¯01
67	4	1112	3.5651E¯02	1.0961E00	1.2003E00	8.9972E¯01	8.0923E¯01
*IG 67							
68	4	1131	3.5075E¯02	1.0950E00	1.1995E00	9.0028E¯01	8.0990E¯01
69	4	1160	3.1375E¯02	1.0882E00	1.1844E00	9.0142E¯01	8.1318E¯01
70	4	1175	2.3260E¯02	1.0656E00	1.1352E00	9.1898E¯01	8.4496E¯01
71	4	1193	1.3796E¯02	1.0609E00	1.1246E00	9.3851E¯01	8.8045E¯01
72	4	1207	9.5139E¯03	1.0461E00	1.0938E00	9.6061E¯01	9.2317E¯01
73	4	1221	7.6878E¯03	1.0457E00	1.0932E00	9.5537E¯01	9.1252E¯01
74	4	1237	6.8799E¯03	1.0432E00	1.0884E00	9.5757E¯01	9.1666E¯01
75	4	1257	5.9964E¯03	1.0388E00	1.0793E00	9.5788E¯01	9.1758E¯01
76	4	1271	3.1094E¯03	1.0292E00	1.0592E00	9.7055E¯01	9.4190E¯01
77	4	1288	3.5876E¯04	1.0074E00	1.0149E00	9.9491E¯01	9.8904E¯01
78	4	1301	3.8049E¯05	1.0032E00	1.0063E00	9.9671E¯01	9.9343E¯01
79	4	1315	3.0320E¯05	1.0029E00	1.0058E00	9.9730E¯01	9.9459E¯01
80	4	1330	1.7288E¯05	1.0019E00	1.0037E00	9.9858E¯01	9.9718E¯01
81	4	1344	1.8491E¯06	9.9994E¯01	9.9988E¯01	1.0003E00	1.0005E00
82	4	1356	2.2074E¯08	9.9998E¯01	9.9996E¯01	1.0000E00	1.0001E00
83	4	1378	1.1622E¯08	9.9997E¯01	9.9993E¯01	1.0000E00	1.0001E00
84	1	1396	6.1348E¯09	9.9996E¯01	9.9992E¯01	1.0000E00	1.0001E00
85	1	1415	2.5295E¯09	9.9998E¯01	9.9997E¯01	1.0000E00	1.0000E00
86	1	1432	1.0726E¯09	9.9999E¯01	9.9998E¯01	1.0000E00	1.0000E00
CONVERGED							
86	0	1454	8.9280E¯10	9.9999E¯01	9.9998E¯01	1.0000E00	1.0000E00

GZ

¯2.3498E¯5   2.6154E¯5   0.00010617   ¯4.9179E¯5

GG

789.6	¯380.23	51.112	¯18.938
¯380.23	205.12	¯23.93	31.517
51.112	¯23.93	727.31	¯378.26
¯18.938	31.517	¯378.26	224.25

TABLE 4
Comparison with Other Methods
(Figures taken from [14])

Method	QNWD	H-J	Ros	SPLX	Pow	Stew
Function						
Rosenbrock	208(−11)	250(−8)	200(−6)	200(−8)	151(−10)*	163(−12)**
Beale	77(−13)	100(−∞)	130(−7)	100(−8)		
Powell 1	978(−7)				433(−13)*	407(−10)**
Cube	254(−15)		200(−∞)	140(−7)		
Box	191(−11)	100(−∞)		290(−5)		
RTF(3)***	130−284				96−120	
	Av. = 189				Av. = 108	
RTF(5)***	312−406				166 − 167	
	Av. = 370				Av. = 166	

* These figures come from [10].
** These figures come from [4].
*** These are Random Trigonometric Functions of dimension 3 and 5. The accuracy criterion used is that the maximum error in any x-component is $<10^{-7}$. The smallest and largest numbers of evaluations taken are listed, as well as the averages.

down to a certain value is listed. The number in parentheses is the exponent, to base 10, of the least calculated function value. The value "$-\infty$" indicates that f was reduced to zero.

7. Acknowledgments. I am especially indebted to Dr. P. G. Comba, whose suggestions and criticisms sowed the seeds for many of the ideas in this work. I am also grateful to Drs. D. Goldfarb and Y. Bard for very helpful discussions, and to Jean-Claude Cohen for his help in setting up the program.

Thanks are also due to M. J. D. Powell whose criticisms resulted in a simplification and clarification of the presentation.

The bulk of this work was done at the IBM New York Scientific Center.

Appendix A.

Limiting Cases of $\nu \to 0$ *and* $\nu \to \infty$.

Case 1: $\nu \to 0$. If ν is set to zero in Eq. (4.17a), the formula for θ_1 is not defined, since $\tau_1^2 = \sigma_1^2$. Therefore, we must consider θ_1 (and η_1) separately. The formula for θ_1 is:

$$\theta_1 = (\epsilon_1 - 2\rho_1)/\sigma_1^2 \tag{A1}$$

and for η_1, we have:

$$\eta_1 = 4\nu^2\rho_1/\sigma_1^4 - 2\theta_1. \tag{A2}$$

When $i \neq 1$, we have:

$$\theta_i = \frac{2\nu^2(\epsilon_i - 2\rho_i)}{\sigma_i^2(\tau_i^2 - \sigma_i^2)} + O(\nu^4), \tag{A3}$$

(A4) $$\eta_i = \frac{4\nu^2}{\sigma_1^4}\rho_i - 2\theta_i.$$

When $\nu \to 0$, every term in formula (4.5a) goes to zero, except the first term (for $i = 1$). The result for γ is (also replacing Λ by L):

(A5) $$\gamma \to L\theta_1\sigma_1.$$

For Γ, we must be more careful. When we replace M by L/ν, we have a denominator which converges to 0, whereas θ_1 and η_1 do not. However, if we evaluate the terms in the brace in formula (4.5b) for $i = 1$, we obtain:

(A6) $$\{\ \}_{i=1} = \eta_1\sigma_1\sigma_1^T + 2\theta_1\sigma_1\sigma_1^T$$

since $\tau_1 = \sigma_1$. Replacing η_1 by expression (A2), we then have:

(A7) $$\{\ \}_{i=1} = \frac{4\nu^2}{\sigma_1^4}\rho_1\sigma_1\sigma_1^T - 2\theta_1\sigma_1\sigma_1^T + 2\theta_1\sigma_1\sigma_1^T,$$

so that all we have left is the first term. There is no difficulty with the rest of the terms in Eq. (4.5b).

For convenience, we define:

(A8) $$\tilde{\theta}_i \equiv \frac{2(\epsilon_i - 2\rho_i)}{\sigma_i^2(\tau_i^2 - \sigma_i^2)}; \qquad i \neq 1,$$

(A9) $$\tilde{\eta}_1 \equiv \frac{4\rho_1}{\sigma_1^4},$$

(A10) $$\tilde{\eta}_i \equiv \frac{4\rho_i}{\sigma_i^4} - 2\tilde{\theta}_i; \qquad i \neq 1,$$

so that

$$\theta_i/\nu^2 = \tilde{\theta}_i + O(\nu^2) \quad \text{and} \quad \eta_i/\nu^2 = \tilde{\eta}_i + O(\nu^2).$$

Then Γ becomes (replacing M by L/ν):

(A11) $$\begin{aligned}\Gamma &= \frac{1}{2\nu^2} L\left\{\frac{4\nu^2\rho_1\sigma_1\sigma_1^T}{\sigma_1^4} + \sum_{i\neq 1}[\eta_i\sigma_i\sigma_i^T + \theta_i(\sigma_i\tau_i^T + \tau_i\sigma_i^T)]\right\}L \\ &= \frac{1}{2} L\left\{\tilde{\eta}_1\sigma_1\sigma_1^T + \sum_{i\neq 1}[\tilde{\eta}_i\sigma_i\sigma_i^T + \tilde{\theta}_i(\sigma_i\tau_i^T + \tau_i\sigma_i^T)]\right\}L + O(\nu^2)\end{aligned}$$

and when $\nu \to 0$, the last term vanishes.

Clearly, this limiting procedure has the effect of correcting g_0 from the results of the first minor step only, and of removing part of the first minor step discrepancy from the correction to G.

Case 2: $\nu \to \infty$. In this case, there is no need to separate out the first minor step. The limit for θ_i is:

(A12) $$\theta_i \to (\epsilon_i - 2\rho_i)/\sigma_i^2,$$

but η_i still contains a multiple of ν^2. The formula for γ remains the same as (4.5a), but that for Γ becomes:

(A13) $$\Gamma = \frac{1}{2} L\left\{\sum_i \frac{4\rho_i}{\sigma_i^4} \sigma_i\sigma_i^T\right\}L + O\left(\frac{1}{\nu^2}\right)$$

and the last term vanishes for $\nu \to \infty$. In this case, g_0 is corrected in terms of all the minor steps, but the G-correction does not contain the θ's.

In the program used to run the test problems, L was set equal to the unit matrix I, as mentioned in the text.

Appendix B.

Comparison with Fiacco-McCormick Method. The method described by Fiacco and McCormick in their book [16] is based largely on a relation identical with Eq. (3.15a). Let a step σ be made up of a linear combination of at most two coordinate directions, viz.:

(B1) $$\sigma_{(ij)} = \alpha_i e_i + \alpha_j e_j.$$

That is, let the direction S_{ij} be specified in terms of coordinate directions e_i and e_j, and do a line search for the minimum of f along that direction, starting at a point x_0. Then the minimum is found at x_1 ($\equiv x_0 + \sigma_{(ij)}$) and the difference between starting and minimum values of f is denoted by $\Delta f_{(ij)}$. We then have, rewriting (3.15a):

(B2) $$\Delta f_{(ij)} = -\tfrac{1}{2}\sigma_{(ij)}^T G^* \sigma_{(ij)}$$

and, replacing $\sigma_{(ij)}$ according to (B1), we obtain:

(B3) $$\Delta f_{(ij)} = -\tfrac{1}{2}\{\alpha_i^2 e_i^T G^* e_i + 2\alpha_i\alpha_j e_i^T G^* e_j + \alpha_j^2 e_j^T G^* e_j\}$$

(remembering that G^* is symmetric). But, because the coordinate-direction vector e_i has the structure: $e_i = (0, 0, \cdots, 0, 1, 0, \cdots, 0)$—where the 1 is in the ith position—each of the products singles out a component of G^*. Thus, e.g.,

(B4) $$e_i^T G^* e_j = G_{ij}^*,$$

so that (B3) becomes:

(B5) $$\Delta f_{(ij)} = -\tfrac{1}{2}\{\alpha_i^2 G_{ii}^* + 2\alpha_i\alpha_j G_{ij}^* + \alpha_j^2 G_{jj}^*\}.$$

Now, we choose the first set of directions for σ so that they lie along the coordinates. Then, we have:

(B6) $$\Delta f_{(ii)} = -\tfrac{1}{2}\alpha_i^2 G_{ii}^*,$$

from which we can solve for the diagonal elements G_{ii}^*.

Next, we arrange that $\alpha_i = \alpha_j$ (and denote them both by α_{ij}), i.e., we search in a direction (always starting at x_0, as before) which bisects the right angle between e_i and e_j. We then have:

(B7) $$\Delta f_{(ij)} = -\tfrac{1}{2}\alpha_{ij}^2(G_{ii}^* + G_{jj}^* + 2G_{ij}^*),$$

from which we can solve for G_{ij}^*, since everything else is known. Clearly, since G^* is symmetric, we need only have done $\frac{1}{2} N(N + 1)$ line searches.

Once we have estimated G^* in this way, we make use of Eq. (3.15b), using the results of the searches along the coordinate directions. (τ is, of course, the same as σ for a single line search.) We then have:

$$\alpha_i g_{0i}^* + \alpha_i^2 G_{ii}^* = 0, \tag{B8}$$

from which we solve for $\{g_{0i}^*\}$. We may then translate g^* to any other point, using (3.12).

The main differences between this method and the QN method outlined in this paper are:

F-M	QN
1. $\frac{1}{2} N(N+1)$ line searches	N line searches
2. Complete estimate of g_0 and G (exact for a quadratic function)	Incomplete estimate of g_0 and G
3. Completely new estimate at next major step	Improvement of previous estimates at next major step

IBM Scientific Center
2670 Hanover Street
Palo Alto, California 94304

1. W. C. Davidon, *Variable Metric Method for Minimization*, Argonne National Lab., ANL-5990 Rev., 1959.
2. R. Fletcher & M. J. D. Powell, "A rapidly convergent descent method for minimization," *Comput. J.*, v. 6, 1963/64, pp. 163–168. MR **27** #2096.
3. C. G. Broyden, "Quasi-Newton methods and their application to function minimization," *Math. Comp.*, v. 21, 1967, pp. 368–381. MR **36** #7317.
4. G. W. Stewart III, "A modification of Davidon's minimization method to accept difference approximations of derivatives," *J. Assoc. Comput. Mach.*, v. 14, 1967, pp. 72–83. MR **35** #7566.
5. D. Goldfarb, "Sufficient conditions for the convergence of a variable-metric algorithm," in *Optimization*, R. Fletcher (Editor), Academic Press, New York, 1969.
6. G. P. McCormick & J. D. Pearson, "Variable metric methods and unconstrained optimization," in *Optimization*, R. Fletcher (Editor), Academic Press, New York, 1969.
7. J. Greenstadt, "Variations on variable-metric methods," *Math. Comp.*, v. 24, 1970, pp. 1–22. MR **41** #2895.
8. H. H. Rosenbrock, "An automatic method for finding the greatest or least value of a function," *Comput. J.*, v. 3, 1960/61, pp. 175–184. MR **24** #B2081.
9. E. M. L. Beale, *On an Iterative Method for Finding a Local Minimum of a Function of More than One Variable*, Stat. Tech. Res. Grp., Technical Report #25, Princeton University, Princeton, N. J., 1958.
10. M. J. D. Powell, "An iterative method for finding stationary values of a function of several variables," *Comput. J.*, v. 5, 1962, pp. 147–151.
11. M. J. D. Powell, "An efficient method for finding the minimum of a function of several variables without calculating derivatives," *Comput. J.*, v. 7, 1964, pp. 155–162. MR **32** #4828.
12. A. Leon, "A comparison among eight known optimizing procedures," in *Recent Advances in Optimization Techniques*, A. Lavi and T. P. Vogl (Editors), Wiley, New York, 1960.
13. D. Goldfarb, "A family of variable-metric methods derived by variational means," *Math. Comp.*, v. 24, 1970, pp. 23–26. MR **41** #2896.
14. J. Kowalik & M. R. Osborne, *Methods for Unconstrained Optimization Problems*, American Elsevier, New York, 1968.
15. A. R. Colville, *A Comparative Study of Non-Linear Programming Codes*, IBM Tech. Rep. #320-2949, 1968.
16. A. V. Fiacco n G. P. McCormick, *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*, Wiley, New York, 1968, p. 175 et seq. MR **39** #5152.