LOCATION OF STABLE AND UNSTABLE EQUILIBRIUM CONFIGURATIONS USING A MODEL TRUST REGION QUASI-NEWTON METHOD AND TUNNELLING

by

Hee-Yuen Herbert Kwok

Thesis submitted to the Faculty of the

Virginia Polytechnic Institute and State University

in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

in

Computer Science and Applications

APPROVED:

Layne T. Watson

Donald C.S. Allison

Werner E. Kohler

M. P. Kamat

September, 1983

Blacksburg, Virginia

LOCATION OF
STABLE AND UNSTABLE EQUILIBRIUM CONFIGURATIONS
USING
A MODEL TRUST REGION QUASI-NEWTON METHOD
AND TUNNELLING

by

Hee-yuen Herbert Kwok

(ABSTRACT)

A hybrid method consists of a quasi-Newton method and a homotopy method for locating multiple equilibrium configurations has been proposed recently. The hybrid method combined the efficiency of a quasi-Newton method capable of locating stable and unstable equilibrium solutions with a robust homotopy method capable of tracking equilibrium paths with turning points and exploiting sparsity of the Jacobian matrix at the same time. A quasi-Newton method in conjunction with a deflation technique is proposed here as an alternative to the hybrid method. The proposed method not only exploits sparsity and symmetry, but also represents an improvement in efficiency. Limit points and nearby equilibrium solutions, either stable or unstable, can be accurately located with the use of a modified pseudoinverse based on

the singular value decomposition. This pseudoinverse modification destroys the Jacobian matrix sparsity, but is invoked only rarely (at limit and bifurcation points where the Jacobian matrix is singular).

ACKNOWLEDGEMENTS

I wish to thank the members of my committee for their guidance to this work. A special thanks is owed to Dr. Layne Watson for all his time, effort and generous help. I would also like to thank Dr. Donald Allison for his consistent support during my entire graduate career in Virginia Tech.

Possibly, I would have never gone to graduate school if it were not from the enthusiastic encouragements from my good friend, Arthur Lee, whom I had lost in 1980. I would also like to thank Janice Tong who encouraged, inspired, supported and helped me throughout my graduate career.

In particular, I wish to thank my parents, brothers and sister for their love, patience, support and encouragements through my endeavor.

TABLE OF CONTENTS

Chapter I

INTRODUCTION

In predicting response of structures susceptible to limit and bifurcation point instabilities, previous techniques, as in [8], [42], [47] and [59] suffered serious difficulties in the vicinity of limit points. The present algorithm is proposed to overcome these difficulties, and is successful in locating equilibrium solutions in the vicinity of a limit point to great accuracy. The algorithm extends a quasi-Newton method with a deflation technique to solve the system of nonlinear equilibrium equations directly; multiple equilibrium solutions (stable or unstable), if they exist, can be located efficiently.

From a recent evaluation by Kamat, Watson, and Venkayya [30], the globally convergent quasi-Newton method of Dennis and Schnabel [16], although quite efficient, has its limitations in the vicinity of limit and bifurcation points and along unloading portions of the equilibrium curve, especially when used in the context of energy minimization. Since the Jacobian matrix of the system of nonlinear equilibrium equations is nearly singular in the vicinity of a critical (limit or bifurcation) point. quasi-Newton iterations encounter serious numerical difficulties. Gay [23] suggests using a modified pseudoinverse in place of the inverse of the Jacobian matrix in the Newton iteration to maintain numerical stability.

After locating the first equilibrium solution at a certain fixed load level by the proposed model trust region quasi-Newton method, deflation is used to locate multiple stable and unstable equilibrium solutions which may exist. The equilibrium solutions already found are used to deflate a nonlinear least squares function, which is used with a model trust region quasi-Newton algorithm to find another equilibrium solution (stable or unstable). If another equilibrium solution exists, it can be located in a finite number of iterations. In minimizing the deflated function, the quasi-Newton method can proceed using the Jacobian matrix of the previous (undeflated) nonlinear least squares function while continuing to exploit sparsity and symmetry.

By means of the matrix factorization LDL^t, the algorithm checks the condition number of the Jacobian matrix of the system of nonlinear equations at every iteration to see if the Jacobian matrix is ill-conditioned. The Jacobian matrix becoming ill-conditioned is generally an indication of entering the vicinity of a critical point, in which case Gay's modification [23] is used to perturb the Jacobian matrix into a better conditioned one to maintain numerical stability.

From any starting point, the globally convergent quasi-Newton method may converge to a local minimum or fail to make reasonable progress. The proposed algorithm detects this situation from the small gradient norm, or from the fact that the algorithm is not making reasonable progress with good directions (i.e. the trust region radius is small). Local minima are used to deflate the nonlinear least-squares function by a procedure known as tunnelling [35-39], and the ultimate result of tunnelling yields an equilibrium solution of the load level. So, the proposed method usually is "globally convergent".

Chapter II

BACKGROUND

The best known tool to solve the simultaneous nonlinear equations or the unconstrained minimization problem is Newton's method. After the introduction of Newton's method, researchers introduced new methods which although still iterative in nature, were quite different from Newton's method. Davidon [9,10] introduced the variable metric minimization method, while Broyden [4] introduced a class of methods for solving simultaneous nonlinear equations. In 1963, Fletcher and Powell [20] further modified and clarified Davidon's work, and together they formed a new class of methods called quasi-Newton methods.

Since then, there has been a proliferation of quasi-Newton methods for unconstrained minimization. Researchers, such as Fletcher [21], Gill and Murray [24], Greenstadt [26], Huang [28], Pearson [43], Powell [45] and Goldfarb [25], developed many different kinds of quasi-Newton algorithms for unconstrained minimization. However, the only quasi-Newton method that was used in solving simultaneous nonlinear equations is the one proposed by Broyden [4].

Since Newton's method required a costly Jacobian evaluation at every iteration, a number of update formulas for the Jacobian martix have been developed for the quasi-Newton methods. They included BFGS update [5,21,25,50], Toint's update [52], Davidon-Fletcher-Powell update [9,20], Powell-symmetric-Broyden update [44], least square sparse secant update [48], Fiacco and McCormick update [19], and Dennis's least square secant update [38]. Watson [58] had an accessment of quasi-Newton sparse update. These update formulas were important in speeding up the quasi-Newton methods. Another disadvantage of Newton's method was the requirement of a good starting point for convergence. Special techniques, such as Powell [44] and Dennis [11,12] were used to achieve global convergence.

However, due to modifications in the quasi-Newton methods, the rate of convergence for these methods was no longer quadratic as in Newton's method. Broyden [5,6], Dennis [11,12], Dennis and More [13] and Powell [46] have investigated the rate of convergence for different quasi-Newton methods.

Another class of methods, which are based upon the approximation of a curve by means of simplices or numerical solution of an initial value problem, also solves the simultaneous nonlinear equations. They are called the simplicial and continuation methods. Frequently, they use the fixed point theorems such as those of Banach, Brouwer, Kakutani, Schauder or Leray-Schauder as a crucial device in asserting the existence of a solution to a problem.

The activity in constructive approximation of fixed points for general mappings was initiated by Scarf [48]. He developed a simplicial method which yielded approximations to solutions by means of simplices. Hansen [27] and Kuhn [44,33] further polished Scarf's algorithm by giving an efficient and compact pivoting procedure for the simplices. These algorithms were quickly superceded by the algorithms of Eaves [17], Eaves and Saigal [18], and Kuhn and MacKinnon [34] which allowed finer approximation.

Eaves [17] also initiated the use of homotopies in simplicial algorithms to approximate fixed points. The homotopy device has since been a very important tool for simplicial and continuation methods. The main disadvantage of the simplicial method is the complexity of triangulations in higher dimensions. In 1976, Kellogg, Li and Yorke gave a novel algorithm for Brouwer's theorem. It was referred to as the homotopy continuation method since it was based on tracking a smooth curve related to a homotopy map. Here, no simplicial decompositions are used and differentiability is required. The continuation methods are improvements in several respects over the simplicial methods. The starting procedure of continuation methods is simpler and more flexible, and they can be applied to a larger class of problems.

Kellogg, Li and Yorke [31] have proposed in abstract terms an algorithm for computing Brouwer fixed points of C^2 maps. Li [40], Smale [51], Alexander and Yorke [2], Chow, Mallet-Paret and Yorke [7] and Alexander [1] give numerical implementations of that algorithm. Watson's algorithm [53-57] was based on Chow and Yorke's initial ideas [7].

However, Watson's algorithm remains well conditioned near limit points whereas the Kellog-Li-Yorke algorithm becomes infinitely ill-conditioned.

Chapter III

MODEL TRUST REGION QUASI NEWTON METHOD

Let R^n denote n-dimensional Euclidean space, and

$$F : R^n \dashrightarrow R^n$$

be twice continuously differentiable. The problem is to find $X_* \in R^n$ such that

$$F(X_*) = 0. \qquad (1)$$

A double dogleg strategy is applied to minimize the nonlinear least squares function:

$$f(X) = 1/2\ F^t(X)F(X). \qquad (2)$$

Of course, a local minimum of the nonlinear least squares function may not be a solution of the simultaneous nonlinear equations. Special techniques, such as tunnelling, have to be used to force the solution of the nonlinear least squares function away from local minima.

A quadratic model

$$m_c(X) = f(X_c) + \nabla f(X_c)^t(X-X_c) + 1/2(X-X_c)^t H_c(X-X_c) \qquad (3)$$

is built around the current estimate X_c. A step $S = (X_+ - X_c)$ is taken to minimize $m_c(X)$ within a region of radius δ_c, where the quadratic model can be "trusted". H_c is the Hessian at X_c and is approximated by J^tJ, where J is the Jacobian matrix of F evaluated at X_c.

If the Newton step

$$S_c^N = -H_c^{-1}\nabla f(X_c) \qquad (4)$$

is within the trust region, then $X_+ = X_c + S_c^N$ is taken as the next point since it is the global minimum of the model. Otherwise, the minimizer of the quadratic model occurs for

$$S = S(\mu) \equiv -(H_c+\mu I)^{-1}\nabla f(X_c),\ \mu \geq 0,$$

$$\text{such that } ||S(\mu)||_2 = \delta_c.$$

The $S(\mu)$ curve, as shown in Figure 1, runs smoothly from the Newton step when $\mu = 0$ to the steepest descent direction

$$S(\mu) \approx -1/\mu\ \nabla f(X_c)$$

for large μ.

The double dogleg strategy is to approximate the $S(\mu)$ curve by the double dogleg arc which connects the "Cauchy point" to a point $\hat{X}_N$ in the Newton direction for m_c, and to choose X_+ to be the point on this arc such that $||X_+ - X_c||$ $= \delta_c$. The strategy looks in the steepest descent direction when δ_c is small and more and more towards the Newton direction as δ_c increases.

The Cauchy point is the minimizer of the quadratic model in the steepest descent direction (which is the same for both $f(X)$ and $m_c(X)$) and is given by

$$X_{cp} = X_c + S_{cp} \tag{5}$$

where $$S_{cp} = -\lambda \nabla f(X_c),$$

and $$\lambda = \frac{||\nabla f(X_c)||^2}{\nabla f(X_c)^t H_c \nabla f(X_c)} .$$

If $\delta_c < ||S_{cp}||$, then X_{cp} is taken instead as

$$X_{cp} = X_c + \delta_c\, S_{cp} \,/\, ||S_{cp}|| . \tag{6}$$

The point in the Newton direction on the double dogleg arc is

$$\hat{X}_N = X_c + \eta S_c^N , \qquad (7)$$

where (as recommended by Dennis and Schnabel [16])

$$\eta = 0.8\nu + 0.2,$$

ανϕ n σατισ£ιεσ

$$||S_{cp}|| \leq \nu||S_c^N|| \leq ||S_c^N||,$$

and is defined by

$$\nu = \frac{||\nabla f(X_c)||^4}{(\nabla f(X_c)^t H_c \nabla f(X_c))(\nabla f(X_c)^t H_c^{-1} \nabla f(X_c))} .$$

Then X_+ is the point along the line joining X_{cp} and $\hat{X}_N$ such that $||X_+ - X_c|| = \delta_c$, i.e.,

$$X_+ = X_c + S_{cp} + \theta(\hat{X}_N - X_{cp}), \qquad (8)$$

where θ is chosen such that

$$||S_{cp} + \theta(\hat{X}_N - X_{cp})|| = \delta_c .$$

To ensure that the algorithm is making good progress, not only $f(X_+) < f(X_c)$ must be satisfied, but also the criterion

$$f(X_+) \leq f(X_c) + \alpha \nabla f(X_c)(X_+ - X_c) \quad (9)$$

(with $\alpha = 10^{-4}$). This guarantees that successive improvements in the function value do not become arbitrarily small. If this condition is not satisfied, then the model function $m_c(X)$ is not representing the true function $f(X)$ well, and the trust region must be reduced. The reduction factor is determined by a backtracking strategy utilizing $f(X_c)$, $f(X_+)$, and the directional derivative $\nabla f(X_c)^t(X_+ - X_c)$ to find a parabola interpolating these data. The new trust region radius δ_{new} corresponds to the minimum of this parabola, and is given by

$$\delta_{new} = - \delta_c \frac{\nabla f(X_c)^t (X_+ - X_c)}{2[f(X_+) - f(X_c) - \nabla f(X_c)^t (X_+ - X_c)]} . \quad (10)$$

The entire double dogleg step and trust region radius calculation (10) are repeated until the acceptability criterion (9) is satisfied.

If X_+ is acceptable, a check is made to see how well m_c has predicted changes in the function f. If the prediction is good, extend the trust region, otherwise, reduce the trust region. The new trust region radius δ_+ is determined as follows:

$$\text{if } |\Delta f| \geq 0.75\ |\Delta f_{pred}|, \text{ set } \delta_+ = 2\delta_c \ ;$$
$$\text{if } |\Delta f| < 0.1\ |\Delta f_{pred}|, \text{ set } \delta_+ = 1/2\delta_c \ ;$$
$$\text{otherwise leave } \delta_+ = \delta_c,$$

where

$$\Delta f_{pred} = m_c(X_+) - f(X_c),$$
$$\Delta f = f(X_+) - f(X_c).$$

The double dogleg strategy requires the Jacobian matrix of the system of nonlinear equations which may be singular or ill-conditioned at each iteration. If the Jacobian matrix J is ill-conditioned, the direction $-J^{-1}(X)F(X)$ would be misleading, and the whole strategy would fall apart. The next section discusses a method for dealing with ill-conditioned Jacobian matrices.

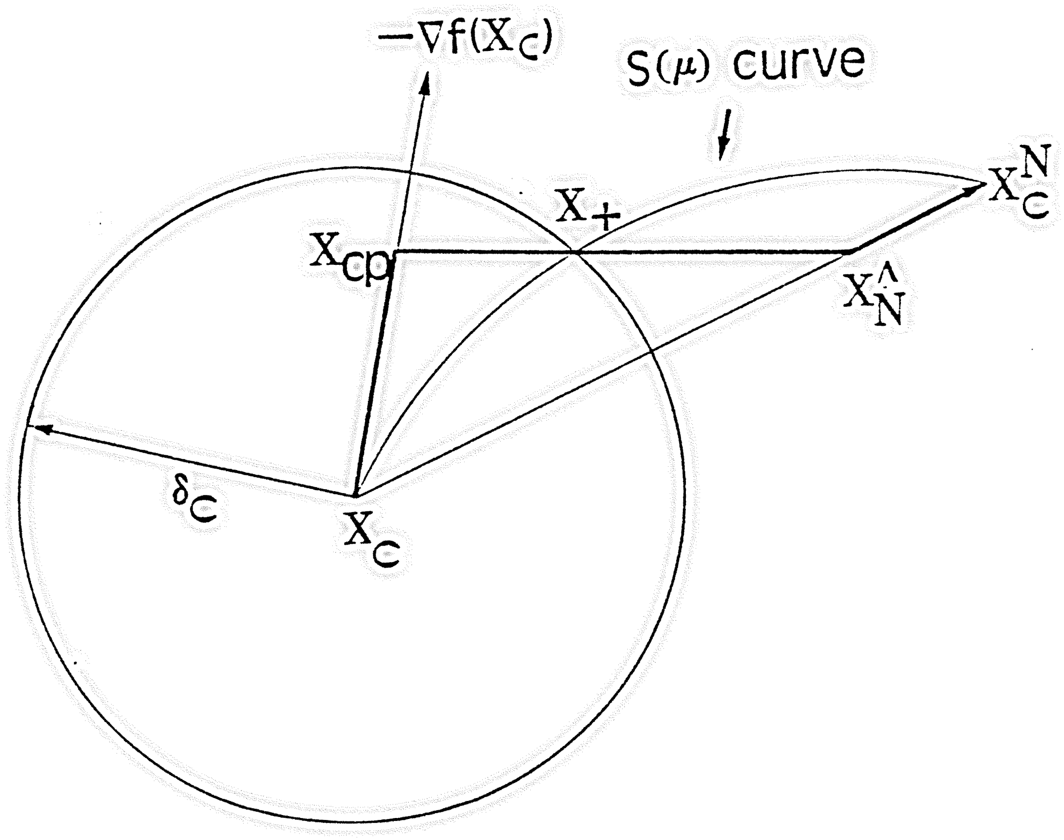

Figure 1. Geometry of the Double Dogleg Step

Chapter IV

GAY'S MODIFIED PSEUDOINVERSE ALGORITHM

Gay [23] shows that if a certain nondegeneracy assumption holds, a modified Newton iteration will converge to a solution of the system of nonlinear equations $F(X) = 0$, whose Jacobian matrix exists and is continuous but may be singular at solutions. Consider the Newton iteration

$$X_{k+1} = X_k - J(X_k)^{-1}F(X_k). \tag{11}$$

For $J(X_k)$ singular, $J^{-1}(X_k)$ is undefined, and for $J(X_k)$ nearly singular, a straightforward numerical implementation of (11) encounters serious difficulties. Gay [23] suggests replacing the inverse of the Jacobian matrix in $J^{-1}(X_k)F(X_k)$ by a modified pseudoinverse $\hat{J}(X_k)^{+}$ in order to maintain numerical stability. Let $J(X_k)$ have the singular value decomposition

$$J(X_k) = USV^t,$$

where U and V are orthogonal matrices, and S is a diagonal matrix with diagonal elements $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n \geq 0$. The Moore-Penrose pseudoinverse of $J(X_k)$ is given by

$$J^{+}(X_k) = VS^{+}U^t, \tag{12}$$

where S^+ is a diagonal matrix with diagonal elements σ_i^+ given by

$$\sigma_i^+ = \begin{cases} 1/\sigma_i & \text{if } \sigma_i > 0 \\ 0 & \text{if } \sigma_i = 0 . \end{cases}$$

For the Newton iteration using $J^+(X_k)$ in place of $J^{-1}(X_k)$ to be well-defined and convergent, $J^+(X_k)$ has to be continuous and bounded. However, $J^+(X_k)$ is discontinuous at -- and unbounded near -- points X_k where $J(X_k)$ changes rank. Modifications have to be made to produce a continuous substitute $\hat{J}(X)^+$ for $J^+(X)$.

Let $A \in R^{nxn}$ have singular value decomposition USV^t, $S = \mathrm{diag}(\sigma_1, \ldots , \sigma_n)$. Denote the modified singular values by

$\hat{\sigma}_i$, and define

$$\hat{S} = \mathrm{diag}(\hat{\sigma}_1, \ldots , \hat{\sigma}_n),$$
$$\hat{A} = U\hat{S}V^t.$$

Fix some tolerance $\varepsilon > 0$. The modified singular values

$\hat{\sigma}_i$ are to be chosen such that the following conditions hold for all $\delta > 0$, any matrix $A' = U'S'V'^t$, $||A - A'|| \leq \delta$, and all j, k, $1 \leq j,k \leq n$:

$$0 \leq \hat{\sigma}^+_j \leq 1/\varepsilon, \qquad (13.1)$$

$$|\hat{\sigma}^+_j - \hat{\sigma}'^+_k| = O(\delta + |\sigma_j - \sigma'_k|), \qquad (13.2)$$

$$\sigma_j = \sigma_k \Rightarrow \hat{\sigma}^+_j = \hat{\sigma}^+_k, \qquad (13.3)$$

$$\hat{\sigma}^+_j = O(\sigma_j), \qquad (13.4)$$

$$0 \leq \sigma_j \hat{\sigma}^+_j \leq 1. \qquad (13.5)$$

Under these conditions the modified pseudoinverse

$$\hat{A}^+ = V\hat{S}^+U^t \qquad (14)$$

of A is continuous and bounded in a neighborhood of A, where

$$\hat{S}^+ = \mathrm{diag}(\hat{\sigma}^+_1, \ldots, \hat{\sigma}^+_n),$$

and

$$\hat{\sigma}^+_i = \begin{cases} 1/\hat{\sigma}_i, & \hat{\sigma}_i > 0 \\ 0, & \hat{\sigma}_i = 0 . \end{cases} \qquad (15)$$

A technical requirement for the local convergence of the Newton iteration

$$X_{k+1} = X_k - \hat{J}^{+}(X_k)F(X_k) \qquad (16)$$

to a zero of F(X) is that

$$F(X)^t J(X)\hat{J}(X)^+ F(X) \geq \theta \; ||F(X)||^2 \qquad (17)$$

for some fixed $\theta > 0$ and all relevant $X \in R^n$, where $\hat{J}(X)^+$ is defined by (13-15). Gay [23] has proved that under the nondegeneracy assumption (17) on a C^1 function F(X), the modified Newton iteration (16), where $\hat{J}(X_k)^+$ is the pseudo-inverse of $J(X_k)$ modified according to (13-15), converges locally to a zero X_* of F(X), whether or not the Jacobian matrix $J(X_*)$ is singular. The requirement (17) is roughly that (2) has no nonzero local minima. Note that even though Gay's modification provides a numerically stable algorithm in the vicinity of critical points, it may fail if (2) has nearby local minima. The modification also destroys the sparsity of $J(X_k)$, and is only locally convergent in the form of equation (16).

Possible choices of $\hat{\sigma}^+$ satisfying the conditions (13) are

$$\hat{\sigma}^{+} = \min\left[\, \sigma/\varepsilon^{2} ,\ 1/\sigma\right],$$

or $$\hat{\sigma}^{+} = \sigma/[\sigma^{2} + \varepsilon^{2}/4],$$

or $$\hat{\sigma}^{+} = \sigma/[\, \sigma^{2} + \max[0 ,\ \varepsilon^{2}-\sigma_{n}^{2}]\ ],$$

where σ_n is the minimum of the singular values.

The numerical results show that Gay's modified pseudo-inverse does indeed handle the numerical instability near critical points, and when used judiciously as part of a model trust region strategy permits accurate calculation of equilibrium solutions at and near critical points.

Chapter V

TUNNELLING

When the globally convergent quasi-Newton method converges to a local minimum, tunnelling [35-39] is applied to tunnel under all irrelevant local minima, and the method approaches equilibrium solutions in an orderly fashion.

The tunnelling algorithm is designed to achieve a "generalized descent property", that is, to find sequentially local minima of f(X) at X_i^*, i = 1, 2, ... , G, such that

$$f(X_i^*) \geq f(X_{i+1}^*), \quad i = 1, 2, \ldots , G-1, \tag{18}$$

until f(X) = 0, thus avoiding all local minima that have functional values higher than $f(X_i^*)$.

The tunnelling algorithm is composed of a sequence of cycles. Each cycle consists of two phases, a minimization phase, and a tunnelling phase. The algorithm starts with the minimization phase to find a local minimum. If the local minimum is not an equilibrium solution, the tunnelling phase is entered to obtain a good starting point for minimization in the next cycle. The process is carried on until an equilibrium solution is located.

In the tunnelling phase, the local minimum point X^* is used as a pole. A root X_o of the tunnelling function

$$T(X,\Gamma) = \frac{f(X) - f(X^*)}{[\,(X - X^*)^t(X - X^*)\,]^\lambda} \qquad (19)$$

is sought. Γ denotes the set of parameters $(X^*, f(X^*), \lambda)$, where X^* is the current local minimum point, $f(X^*)$ is the functional value of the current local minimum, and λ is the pole strength at X^*. Starting with $\lambda = 1$, the pole strength λ at X^* is increased by 0.1 until $T(X,\Gamma)$ decreases away from X^*.

The tunnelling function $T(X,\Gamma)$ itself may have a lot of relative local minima where its gradient is zero. A stabilized Newton method (e.g., model trust region quasi-Newton algorithm) is used in the tunnelling phase to find a X_r such that the gradient $T_x(X_r,\Gamma)$ is equal to zero. If the method converges to a singular point X_m, a movable pole with a pole strength of η is introduced at X_m to cancel the singularity. The tunnelling function becomes

$$T(X,\Gamma) = \frac{f(X) - f(X^*)}{[\ [(X-X^*)^t(X-X^*)]^\lambda [(X-X_m)^t(X-X_m)]^\eta\ ]} . \qquad (20)$$

The tunnelling phase stops when $T(X,\Gamma) \leq 0$. Otherwise, X_m is moved to the most recently found relative local minimum X_1. Starting with $\eta = 0$, the pole strength η of the movable pole is increased (in increments of 0.1) to enforce a descent property on the nonlinear least squares function of the system $T_x(X,\Gamma)$.

The tunnelling phase is continued until a point X_o such that $T(X_o,\Gamma)$ is not a local minimum and $T(X_o,\Gamma) \leq 0$ is found. Then X_o is used as the starting point for the next minimization phase. If there is no $X_o \neq X^*$ such that $f(X_o) \leq f(X^*)$, and $f(X^*) \neq 0$, then there is no equilibrium solution at that given load level.

5.1 DEFLATION

Deflation, as a special case of tunnelling, looks for multiple solutions at a given load level. If X_o^* is an equilibrium solution, another equilibrium solution can be found by locating a zero of the tunnelling function

$$T(X,\Gamma) = [f(X)-f(X_o^*)] \,/\, [(X-X_o^*)^t(X-X_o^*)]^{\lambda}.$$

With the pole strength λ set to 1, the tunnelling function is the same as the deflated function

$$f^*(X) = f(X) \,/\, [(X-X_o^*)^t(X-X_o^*)], \tag{21}$$

since $f(X_o^*)$ is zero. The deflated function is minimized with an initial guess $X_1 = (1+\xi)X_o^*$ where ξ is a given perturbation. If a second equilibrium solution X_1^* does exist, the nonlinear least squares function $f(X)$ is deflated with both X_0^* and X_1^*. The deflated function

$$f^*(X) = f(X) \,/\, [(X-X_o^*)^t(X-X_o^*)][(X-X_1^*)^t(X-X_1^*)] \tag{22}$$

is minimized to see if a third equilibrium solution exists. The deflation process is continued until no more equilibrium solutions are found.

Some details on minimizing the deflated function follow [29]: In minimizing $f^*(X)$ using the double dogleg strategy, the directions $J^{*t}(X)F^*(X)$ and $-J^{*-1}(X)F^*(X)$ are required, where $J^*(X)$ is the Jacobian of $F^*(X)$. We assumed that minimizing $f^*(X)$ is equivalent to solving the system of nonlinear equations

$$F^*(X) = 0,$$

where $f^*(X) = 1/2\ F^{*t}(X)F^*(X)$.

For the single deflation case,

$$F^*(X) = \frac{F(X)}{||X - X_o^*||}, \tag{23}$$

for which the Jacobian matrix of $F^*(X)$ is given by

$$J^*(X) = \frac{J(X) + uv^t}{||X - X_o^*||}, \tag{24}$$

where

$$u = -F^*(X), \qquad \text{and} \qquad v = \frac{(X - X_o^*)}{||X - X_o^*||}.$$

$J^{*-1}(X)$ can be obtained explicitly as :

$$||X - X_o^*|| \; [\; J^{-1}(X) - 1/\beta \; J^{-1}(X)uv^tJ^{-1}(X) \;], \qquad (25)$$

where

$$\beta \quad = \quad 1 - v^tp \; ,$$

and $$p \quad = \quad J^{-1}(X)F^*(X) \; .$$

Thus

$$-J^{*-1}(X)F^*(X) \quad = \quad -||X - X_o^*|| \; (1/\beta) \; p \; , \qquad (26)$$

and

$$J^{*t}(X)F^*(X) \quad = \quad \frac{J^t(X)F^*(X) + vu^tF^*(X)}{||X - X_o^*||} \; . \qquad (27)$$

Similarly, for the double deflation case,

$$-J^{*-1}(X)F^*(X) = -||X - X_o^*|| \; ||X - X_1^*|| \; (1/\beta) \; p, \qquad (28)$$

and

$$J^{*t}(X)F^*(X) \quad = \quad \frac{J^t(X)F^*(X) + vu^tF^*(X)}{||X - X_o^*|| \; ||X - X_1^*||} \; , \qquad (29)$$

$$\text{with } u \quad = \quad -F^*(X) \quad = \quad \frac{-\;F(X)}{||X - X_o^*|| \; ||X - X_1^*||} \; ,$$

$$v = \frac{||X-X_o^*||^2(X-X_1^*) + ||X-X_1^*||^2(X-X_o^*)}{||X - X_o^*||\ ||X - X_1^*||} ,$$

$$\beta = 1 - v^t p ,$$

and

$$p = J^{-1}(X)F^*(X) .$$

With the above formulas, the quasi-Newton method with the double dogleg strategy can be implemented with deflation using a modified Jacobian matrix while continuing to exploit sparsity and symmetry.

Chapter VI

THE OVERALL ALGORITHM

To solve the system of nonlinear equations

$$F(X_*) = 0,$$

where $F : R^n \text{-->} R^n$, and $X_* \in R^n$, the following algorithm is applied to minimize the nonlinear least squares function

$$f(X) = 1/2\ F^t(X)F(X).$$

(1) Start out with an initial tolerance (TOL), an initial guess (X_c), an initial trust region radius (δ_c), and a maximum number of function/Jacobian evaluations (IEVAL).

(2) Calculate the Jacobian matrix $J(X_c)$ from $F(X)$. If the algorithm is minimizing the deflated function, use the modified Jacobian matrix instead.

(3) If the number of function/Jacobian evaluations exceeds IEVAL or $||\nabla f(X_c)|| < \text{TOL}$, go to step (12).

(4) Calculate the condition number of the Jacobian matrix. If the Jacobian matrix is ill-conditioned, Gay's modification is applied to perturb the Jacobian matrix into a better conditioned one.

(5) Build a quadratic model m_c around the current estimate X_c.

(6) Calculate the next step $S = S(\mu) = -(H + \mu I)^{-1}\nabla f(X_c)$ such that $||S(\mu)|| \leq \delta_c$ by the double dogleg strategy to minimize the nonlinear least squares function $f(X)$.

(7) Calculate $F(X_c + S)$. If the number of function/Jacobian evaluations exceeds IEVAL, go to step (12). If the step S is acceptable, go to step (8). Otherwise go to step (9).

(8) The step S is acceptable: Set $X_+ := X_c + S$. If $S=S(0)$ (the trust region includes the quasi-Newton point), go to step (11). Otherwise go to step (10).

(9) The step S is not acceptable. If the algorithm is not trying to take a bigger step, then reduce the trust region radius by a factor determined from a backtracking strategy, and go back to step (6). Otherwise, restore the good X_+ and $f(X_+)$ that was saved in step (10) before, and go to step (11).

(10) If the actual reduction and the predicted reduction are in good agreement or the reduction in the true function is large, then save the X_+ and $f(X_+)$. Go back to step (6) to try a bigger step by doubling the trust region radius.

(11) If $||F(X_+)|| > TOL$, update the trust region according to the prediction of the function $f(X)$ by the model function $m_c(X)$. Set $X_c := X_+$, and go back to step (2). Otherwise, go to step (13).

(12) $X_o = X_c$ is a local minimum or the algorithm has failed to make significant progress. Tunnelling is applied to find a X_o^+ such that $T(X_o^+,\Gamma) < 0$. If such a X_o^+ exists, reset IEVAL := 0, reset the initial guess X_c :=

X_c^+, and go back to step (2). Otherwise, $f(X_c)$ is the global minimum, and the algorithm stops since there are no more equilibrium solutions at the given load level.

(13) $X^* = X_+$ is an equilibrium solution. If more equilibrium solutions at that load level are desired, deflate the nonlinear least squares function with the solution X^*, reset IEVAL := 0, reset the initial guess X_c := $(1+\xi)X^*$, and go back to step (2). Otherwise, the algorithm stops.

In the implemention of the proposed method, two poles, one at the most recently found local minimum (X_{lm}), and the other at the most recently found equilibrium solution (X_{es}), are introduced in the nonlinear least squares function f(X) to form the deflated function $f^*(X)$ that is minimized. If the method converges to a new local minimum or equilibrium solution, the corresponding pole is moved to that local minimum or equilibrium solution. The process is carried on until either the desired number of equilibrium solutions is found, there are no more equilibrium solutions (see step (12)), or a limit on the number of function/Jacobian evaluations is exceeded. In the tunnelling phase, instead of a stabilized Newton method, a quasi-Newton method with the double dogleg strategy is used. The quasi-Newton method finds an X_o such that the tunnelling function $T(X,\Gamma)$ is less than zero, then X_o is used as the starting point in the next minimization phase, and the algorithm proceeds.

Chapter VII

ILLUSTRATION OF THE PROPOSED METHOD

The proposed method was first validated on the snap-through response of a shallow arch shown in Figure 2. The load deflection curve of the crown of the arch has two limit points, one at 1773.00 lb, and the other at 3064.18 lb. The task was to locate all equilibrium solutions at every load level. The load deflection curve of the crown of the arch was found by tracking the curve with the homotopy method of Kamat, Watson and Venkayya [30]. At each load level, the equilibrium solutions that were located by the proposed method are denoted by a '*'. Figure 3 shows that all equilibrium solutions at each level were located successfully by the proposed method.

After the first equilibrium solution was found, deflation was applied to locate the second and third equilibrium solutions (if they exist). The method was initiated with a load of 500 lb, and an increment of 500 lb for the next load steps. The first equilibrium solution of a given load step was used as an initial guess for the next load step. As shown in Figure 3, the first three load steps had only one equilibrium solution, the fourth to sixth load steps had three, and there was only one equilibrium solution each for the seventh and eighth load steps. At the seventh load step (3500 lb), when the first equilibrium solution of the sixth load step was used as the initial guess, the proposed method converged to a local minimum. Tunnelling was applied, and the distant equilibrium solution was located. When the load steps were close to limit loads, Gay's modification was applied to perturb the Jacobian matrix into a better conditioned one to accelerate convergence to the equilibrium solution.

The proposed method was compared to a classical Newton method and a quasi-Newton method using the double dogleg strategy (QNM-DDL) but without Gay's modification in the vicinity of limit points. The classical Newton method diverged, while QNM-DDL without Gay's modification failed to locate the equilibrium solutions near limit points. Both methods failed in the vicinity of limit points due to the ill-conditioning of the Jacobian matrix. For larger load steps the classical Newton method diverged.

The proposed method checks the condition number of the Jacobian matrix at every iteration. If it is necessary to locate equilibrium solutions in the vicinity of critical points, Gay's modification is invoked to perturb the Jacobian matrix into a better conditioned one, since the Jacobian matrix is ill-conditioned near critical points. Otherwise, it uses a standard quasi-Newton method with the double dogleg strategy to locate a minimum of the least squares function (2).

The method was also validated on the snap-through response of a shallow dome under a concentrated load at the center, as shown in Figure 4. The load response curve of the crown of the shallow dome is quite complex, having many equilibrium solutions at each load level. Since there are many bifurcation points, only a portion of the load response curve is shown in Figure 5. The equilibrium solutions that were located by the proposed method are indicated by a '*' in Figure 5. As shown in Figure 5, multiple equilibrium solutions at each load level were located to illustrate the success of the proposed method.

The proposed method only deflates using at most two poles: one at the most recently found local minimum point, and the other at the most recently found equilibrium solution (if they exist). However, since the proposed method may converge back to a recently found equilibrium solution or local minimum, and the application of deflation may virtually destroy some nearby minima, it is not guaranteed that all equilibrium solutions can be located. Of course, we could deflate with more than two poles, but then the algorithm quickly becomes unwieldy, and there is still no guarantee of success. Here we only try to illustrate the success of the proposed method in finding multiple (but not necessarily all) equilibrium solutions.

In applying the tunnelling algorithm, the initial guess in the tunnelling phase has to be far away from the current local minimum point to prevent getting back this same local minimum point as the next starting point. Since the local minimum point that was found is only an approximation to the true local minimum point, if we start the tunnelling phase with an initial guess close to the (computed) local minimum point, it may converge to another approximation of the <u>same</u> local minimum point but with a lower functional value. In this case the minimization phase and the next tunnelling phase produce again the same approximate local minimum point, and no progress is being made. However, if the initial guess is far away from the current local minimum point, there is a chance that some equilibrium solutions will be missed.

The curve in Figure 5 was generated by starting the homotopy method of Kamat, Watson and Venkayya [30] with different starting points and an accuracy of 10^{-10}. Due to the complicated response of the structure to the loads, there are many bifurcation points along the curve, which the homotopy method is not designed to handle. Neither the homotopy nor the quasi-Newton algorithm by itself could have produced all the branches shown in Figure 5. The curve in Figure 5 was generated by starting the homotopy method at zero first (the solid line), and then from the equilibrium solutions that were located by the proposed model trust region quasi-Newton method (the dashed, dotted, and dashed-dotted lines). For simplicity only portions of the equilibrium curves are shown in Figure 5. Some points on the curves that were not located by the quasi-Newton method were validated by starting the quasi-Newton method nearby, and the quasi-Newton method converged to the same equilibrium solutions computed by the homotopy method.

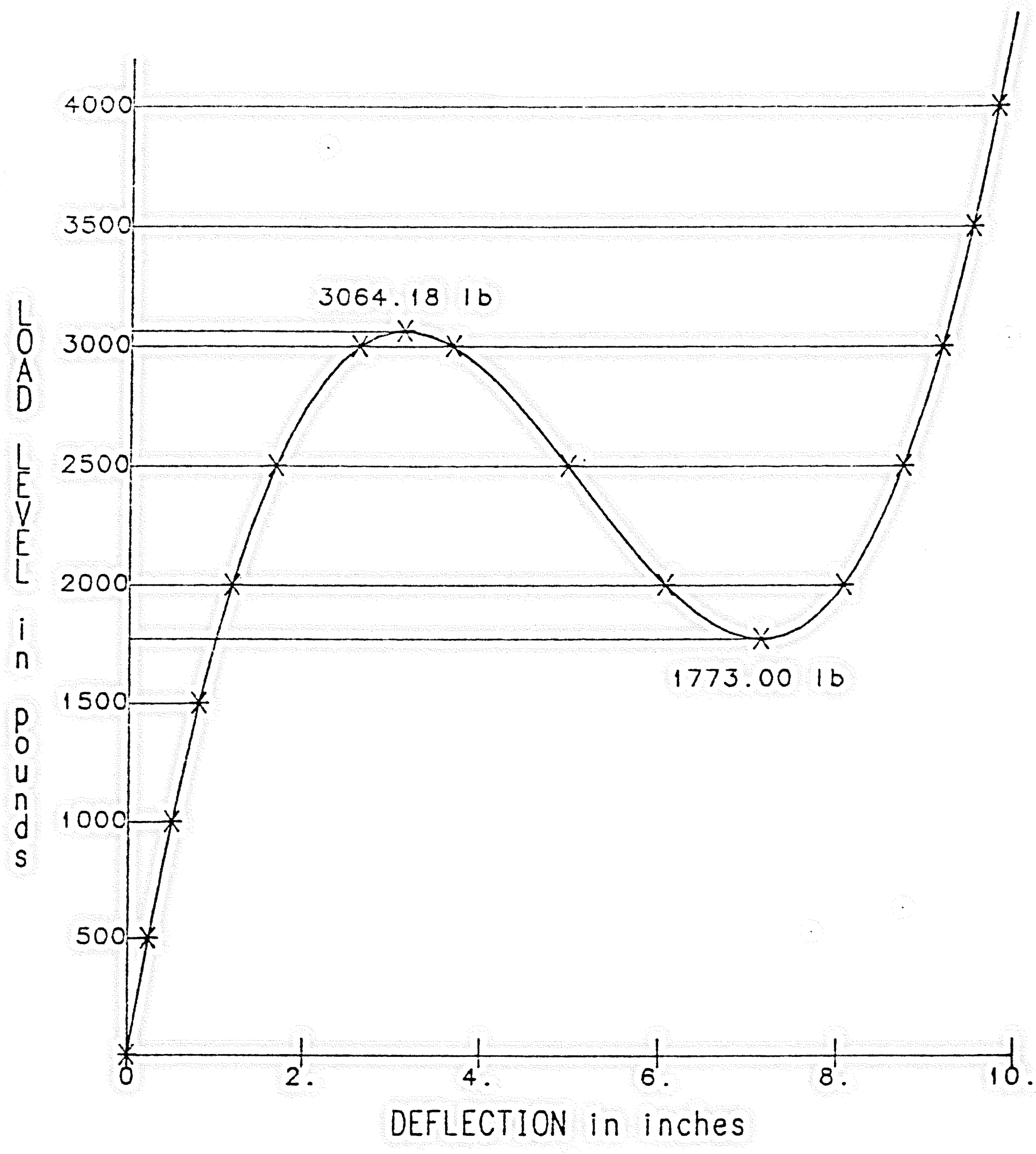

Figure 3. Load Deflection Curve of Shallow Arch Crown

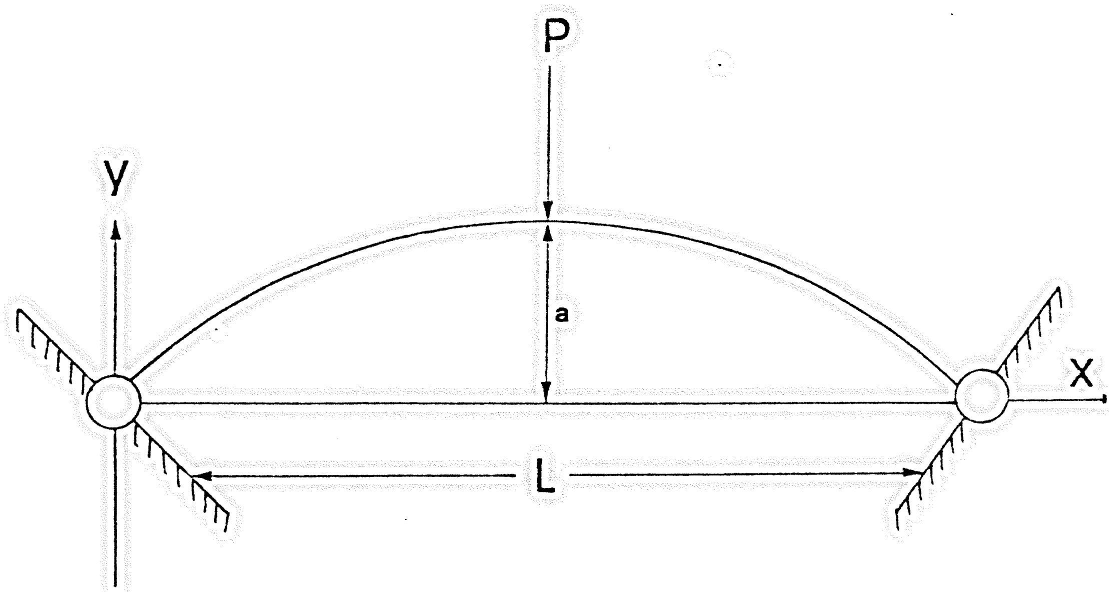

Figure 2. Shallow Arch (29 degrees of freedom)

Coordinates of the Node Points of Dome Structure.

Node	X	Y	Z
1	0.0	0.0	6.0
3	-15.0	25.9807	4.5
4	-30.0	0.0	4.5
9	0.0	60.0	0.0
10	-30.0	51.9615	0.0
11	-51.9615	30.0	0.0
12	-60.0	0.0	0.0

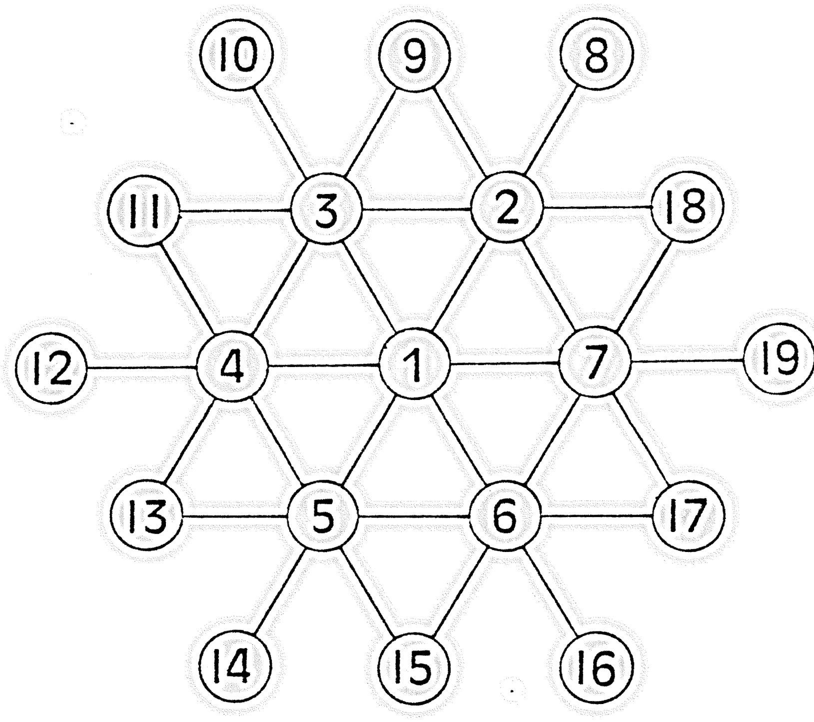

Figure 4. Shallow Dome (21 degrees of freedom)

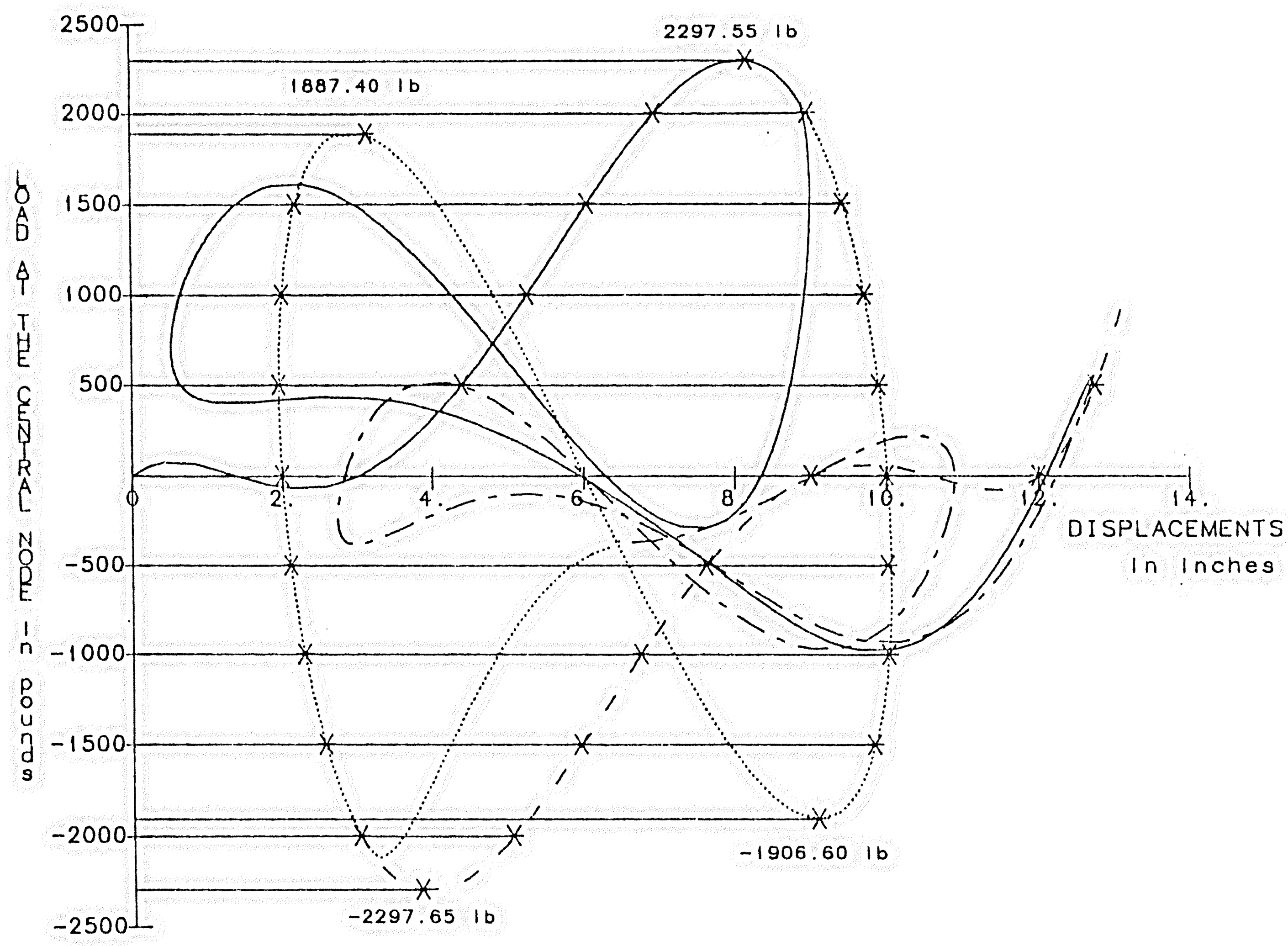

Figure 5. Load Deflection Curve of Shallow Dome Crown

Chapter VIII

CONCLUSIONS

The proposed method, model trust region quasi-Newton algorithm and tunnelling, works extremely well in locating multiple equilibrium solutions, either stable or unstable. Although the use of Gay's modification in the vicinity of critical points destroys sparsity and symmetry, it is only invoked when equilibrium solutions in the vicinity of critical points are needed.

As an alternative to the hybrid method of Kamat and Watson [29], the model trust region quasi-Newton method with tunnelling is a more efficient method for locating a few equilibrium solutions. With the use of deflation, multiple solutions can be located. If equilibrium solutions in the vicinity of critical points are desired, they can be located by using Gay's modified pseudoinverse. With the use of a skyline structure to store the Jacobian matrix, the method exploits sparsity and symmetry. Update formulas for the Jacobian matrix, such as Toint's update [52], can be used in the future to save the costly Jacobian evaluations at every iteration. On the other hand, the proposed algorithm may present a fragmented picture. For example, Figure 5 would have been very difficult to obtain without the homotopy method. Although more has to be done in the future to make the method robust, the preliminary results are promising.

REFERENCES

[1] J.C. Alexander, The Topological Theory of an Embedding Method, Continuation Methods, H. Wacker, ed., Academic Press, New York (1978).

[2] J.C. Alexander and J.A. Yorke, A Numerical Continuation Method that Works Generically, University of Maryland, Dept. of Math., MD77-9-JA, TR77-9 (1977).

[3] E. Allgower and K. Georg, Simplicial and Continuation Methods for Approximating Fixed Points and Solutions to Systems of Equations, SIAM Review 22, No. 1, (1980) 28-85.

[3] C.G. Broyden, A Class of Methods for Solving Nonlinear Simultaneous Equations, Math. Comp., 19, (1965) 577-593.

[5] C.G. Broyden, The Convergence of a Class of Double Rank Minimization Algorithms, Parts I II, Inst. Math. Appl. 6 (1971) 76-90, 222-236.

[6] C.G. Broyden, The Convergence of an Algorithm for Solving Sparse Nonlinear System, Math. Comp. 25 (1971) 285-294.

[7] S.N. Chow, J. Mallet-Paret and J.A. Yorke, Finding Zeros of Maps: Homotopy Methods that are Constructive with Probability One, Math. Comput., 32, (1978) 887-899.

[8] M.A. Crisfield, A Fast Incremental/Iterative Solution Procedure that Handles "Snap-Through", Computers and Structures 13 (1981) 55-62.

[9] W.C. Davidon, Variable Metric Method for Minimizations, Rep. ANL-5990 Rev. Argonne National Laboratories, Argonne, III (1959).

[10] W.C. Davidon, Optimally Conditioned Optimzation Algorithms without Line Searches, Math. Programming 9, (1975) 1-30.

[11] J.E. Dennis Jr., On the Convergence of Broyden's Method for Nonlinear Systems of Equations, Math. Comp. 25, (1971) 559-567.

[12] J.E. Dennis Jr., A Brief Survey of Convergence Results for Quasi-Newton Methods, Nonlinear Programming edited by R. Cottle and C. Lemke, SIAM-AMS, Proc. Vol. IX (1976).

[13] J.E. Dennis and J.J. More, A Characterization of Superlinear Convergence and Its Application to Quasi-Newton Methods, Math. Comp. 28, (1974) 549-560.

[14] J.E. Dennis Jr. and J.J. More, Quasi-Newton Methods: Motivation and Theory, SIAM Review 19 (1979) 46-89.

[15] J.E. Dennis Jr. and R.B. Schnabel, Least Change Secant Updates for Quasi-Newton Methods, SIAM Review 21 (1979) 443-459.

[16] J.E. Dennis Jr and R.B. Schnabel, Numerical Methods for Unconstrained Optimization and Nonlinear Equations Prentice-Hall, Englewood Cliffs, N.J. (1983).

[17] B.C. Eaves, Homotopies for Computation of Fixed Points, Ibid. 3, 1, (1972) 1-22.

[18] B.C. Eaves and R.Saigal, Homotopies for Computation of Fixed Points on Unbounded Regions, Math. Programming 3, 2, (1972) 225-237.

[19] A.V. Fiacco and G.P. McCormick, Nonlinear Programming: Sequential Unconstrained Minimization techniques, John Wiley, New York (1968).

[20] R. Fletcher and M.J.D. Powell, A Rapidly Convergent Descent Method for Minimization, Comput. J. 6, (1963) 163-168.

[21] R. Fletcher, A New Approach to Variable Metric Algorithms, Ibid. 13, (1970) 317-322.

[22] R. Fletcher, Methods for Solving Non-Linearly Unconstrained Optimization Problems, State of the Art in Numerical Analysis edited by D. Jacobs, Academic Press (1977).

[23] D. Gay, Modifying Singular Values: Existence of Solutions to Systems of Nonlinear Equations Having a Possible Singular Jacobian Matrix, Mathematics of Computation 31 (1977) 962-973.

[24] P.E. Gill and W. Murray, Quasi-Newton Methods for Unconstrained Optimization, J. Inst. Math. Appl. 9, (1972) 91-108.

[25] D. Goldfarb, A Family of Variable-Metric Methods Derived by Variational Means, Math. Comp. 24 (1970) 23-26.

[26] J. Greenstadt, Variations on Variable-Metric Methods, Math. Comp. 24, (1970) 1-18.

[27] T. Hansen, On the Approximation of Nash Equilibrium Points in an N-person Noncooperative Game, SIAM J. Appl. Math. 26, 3, (1968) 622-637.

[28] H.Y. Huang, Unified Approach to Quadratically Convergent Algorithms for Function Minimization, J. Optimization Theroy Appl. 5, (1970) 405-423.

[29] M.P. Kamat, L.T. Watson, and J.L. Junkins, A Robust and Efficient Hybrid Method for Finding Multiple Equilibrium Solutions, Proc. Third International Symposium on Numerical Methods in Engineering, Paris, France, Vol. II, (1983) 799-807.

[30] M.P. Kamat, L.T. Watson and V.B.Venkayya, A Quasi-Newton versus a Homotopy Method for Nonlinear Structural Analysis, Journal of Computers and Structures 17 (1983) 579-585.

[31] R.B. Kellogg, T.Y. Li and J. Yorke, Constructive Proof of the Brouwer Fixed Point Theorem and Computational Results, SIAM J. Numer. Anal. 4, (1976) 473-383.

[32] H.W. Kuhn, Simplicial Approximation of Fixed Points, Ibid. 61, (1968) 1238-1242.

[33] H.W. Kuhn, Approximate Search for Fixed Points, Computing Methods in Optimization Problems 2 Academic Press, New York (1969).

[34] H. Kuhn and J.G. MacKinnon, Sandwich Method for Finding Fixed Points, J. Optimization Theory Appl. 17, (1975) 189-204.

[35] A.V. Levy and S. Gomez The Tunnelling Algorithm for the Global Optimization of Constrained Functions, Communicaciones Tecnicas (1980), Serie Naranja, No. 231, IIMAS-UNAM.

[36] A.V. Levy and A. Montalvo, The Tunnelling Algorithm for the Global Minimization of Functions, Dundee Biennal Conference on Numerical Analysis (1977), University of Dundee, Scotland.

[37] A.V. Levy and A. Montalvo, Algoritmo de Tunelizacion para la Optimacion Global de Funciones, Communicaciones Tecnicas (1979), Serie Naranja, No. 204, IIMAS-UNAM.

[38] A.V. Levy and A. Calderon, A Robust Algorithm for Solving Systems of Non-linear Equations, Dundee Biennal Conference on Numerical Analysis (1979), University of Dundee, Scotland.

[39] A.V. Levy and A. Montalvo, A Modification to the Tunnelling Algorithm for Finding the Global Minima of an Arbitrary One Dimensional Function, Communicaciones Tecnicas (1980), Serie Naranja, No. 240, IIMAS-UNAM.

[40] T.Y. Li, Computing the Brouwer Fixed Point by Following the Continuation Curve, Proc. Sem., Dalhousie Univ., Halifax, Academic Press, New York, (1976) 131-135.

[41] E.S. Marwil, Exploiting Sparsity in Newton-Like Methods, Ph.D. Thesis, Cornell University, Ithaca, NY (1978).

[42] J. Padovan, Self-Adaptive Predictor-Corrector Algorithm for Static Nonlinear Structural Analysis, NASA CR-165410 (1981).

[43] I.D. Pearson, Variable Metric Methods of Minimization, Comput. J. 12, (1969) 171-178.

[44] M.J.D. Powell, A Hybrid Method for Nonlinear Equations, Numerical Methods for Nonlinear Algebraic Equations edited by P. Rabinowitz, Gordon and Breach, London (1970).

[45] M.J.D. Powell, A New Algorithm for Unconstrained Optimization, Nonlinear Programming J.B. Rosen, O.L. Mangasarian and K. Ritter, eds., Academic Press, New York (1970).

[46] M.J.D. Powell, Some Global Convergence Properties of a Variable Metric Algorithm for Minimization without Exact Line Searches, Nonlinear Programming, edited by R. Cottle and C. Lemke, SIAM-AMS, Proc, Vol IX (1975).

[47] E. Riks, An Incremental Approach to the Solution of Snapping and Buckling Problems, International Journal of Solids and Structures 15 (1979) 529-551.

[48] H. Scarf, The Approximation of Fixed Points of Continuous Mapping, SIAM J. Appl. Math., 15, 5, (1967) 1328-1343.

[49] L.K. Schubert, Modification of a Quasi-Newton Method for Nonlinear Equations with a Sparse Jacobian, Math. Comp. 24 (1970) 27-30.

[50] D.F. Shanno, Conditioning of Quasi-Newton Methods for Function Minimization, Math. Comp. 24 (1970) 647-656.

[51] S. Smale, A Convergent Process of Price Adjustment and Global Newton Methods, J. Math. Econ. 3, (1976) 1-14.

[52] Ph.L. Toint, On Sparse and Symmetric Matrix Updating Subject to a Linear equation, Math. Comp. 31 (1977) 954-961.

[53] L.T. Watson, Solving the Nonlinear Complementarity Problem by a Homotopy Method, SIAM J. Control and Optimization 17 (1979) 36-46.

[54] L.T. Watson, A Globally Convergent Algorithm for Computing Fixed Points of C^2 Maps, Appl. Math. Comp. 5 (1979) 297-311.

[55] L.T. Watson, Solving Finite Difference Approximations to Nonlinear Two-point Boundary Value Problems by a Homotopy Method, SIAM J. Sci. Stat. Comput. 1, No 4, (1980) 467-480.

[56] L.T. Watson and W.H. Yang, Optimal Design by a Homotopy Method, Applicable Analysis 10 (1980) 275-284.

[57] L.T. Watson and D. Fenner, Chow-Yorke Algorithm for Fixed Points or Zeros of C^2 Maps, ACM Transactions on Mathematical Software 6, No 2, (1980) 252-259.

[58] L.T. Watson, M.P. Kamat and D.J. VandenBrink, An Assessment of Quasi-Newton Sparse Update Techniques for Nonlinear Structural Analysis, Computer Methods in Applied Mechanics and Engineering 26 (1981) 363-375.

[59] G.A. Wempner, Discrete Approximations Related to Nonlinear Theories of Solids, International Journal of Solids and Structures 7 (1971) 1581-1599.

CAPTIONS OF FIGURES

Figure 1. Geometry of the Double Dogleg Step :

X_c to X_{cp} to $\hat{X}_N$ to X_c^N.

Figure 2. Shallow Arch (29 degrees of freedom) :

$y = a \sin(\pi x/L)$,

$a = 5$ in.,

$L = 100$ in.,

$A = 0.32$ in^2.,

$I = 1$ in^4.,

$E = 10^7$ psi,

10 frame elements for 1/2 span.

Figure 3. Load Deflection Curve of Shallow Arch Crown.

Figure 4. Shallow Dome (21 degrees of freedom) :

$A_i = 0.1$ in^2., $i = 1, 2, \ldots\ldots, 30$,

nodes 1 to 7 are free,

nodes 8 to 19 are pinned.

Figure 5. Load Deflection Curve of Shallow Dome Crown.

The vita has been removed from the scanned document