COMMUNICATIONS IN APPLIED NUMERICAL METHODS, Vol. 3, 581–585 (1987)

A NOTE ON THE OPTIMUM CHOICE FOR PENALTY PARAMETERS

BAHRAM NOUR-OMID*

Mechanics and Materials Engineering Lab., Lockheed Palo Alto Research Lab., Palo Alto, CA 94304, U.S.A.

AND

PETER WRIGGERS†

Institut fur Baumechanik und numeriche Mechanik, Universität Hannover, West Germany

SUMMARY

This paper looks at the numerical characteristics of penalty methods for evaluating the solution of symmetric systems of equations with imposed constraints. The sources of error resulting from this approach are identified and an estimate for the penalty parameter that minimizes this error is obtained. The results of the error analysis and the effect of penalty parameter on the accuracy and rates of convergence of the solution algorithm are demonstrated with the aid of some numerical examples.

INTRODUCTION

Most problems in structural and continuum mechanics may be formulated as the minimization of the potential energy subject to certain constraint conditions.[1] At present, there are three different procedures that can be employed for evaluating the solution to such problems; the Lagrange multiplier method,[2] penalty methods[3] and direct elimination of dependent variables. Boundary conditions are the most common form of constraints that occur in finite element analysis and are usually satisfied by direct elimination. However, more complex conditions such as incompressibility or contact constraints require more sophisticated methods.[4] In this paper we focus attention on the use of penalty methods[3,5,6] for the solution of such problems.

The main advantage of penalty methods is their simplicity and the fact that they can easily be implemented in a finite element program. They are particularly attractive for inequality constraint problems such as those arising in contact problems which are our focus here. However, the penalty approach results in solutions that satisfy the constraint equations only approximately. The accuracy of this approximation depends strongly on the penalty parameter. The correct choice for this parameter is the essence of the algorithm. Here we look at the errors that occur at intermediate steps of the algorithm and, furthermore, evaluate their influence on the accuracy of the computed solution. We then derive an expression for an appropriate measure of the error as a function of the penalty parameter. The minimization of this error results in the optimum penalty parameter. Finally, analytical results are compared with those obtained by numerical experiments, and good agreement is obtained.

STATEMENT OF THE PROBLEM

Consider the potential energy associated with a linear discretized structure

$$\pi(\mathbf{u}) = \frac{1}{2}\mathbf{u}^{\mathrm{T}}\mathbf{K}\mathbf{u} - \mathbf{f}^{\mathrm{T}}\mathbf{u} \tag{1}$$

* Visiting scholar, center for Pure and Applied Mathematics, University of California, Berkeley CA 94720, U.S.A.
† Visiting scholar, Department of Civil Engineering, University of California, Berkeley CA 94720, U.S.A.

0748–8025/87/060581–05$05.00

Received 30 August 1985
Revised 3 April 1987

where $\mathbf{K}$ is the $n \times n$ stiffness matrix obtained by a finite element discretization of the continuum, $\mathbf{f}$ is the applied load vector, and $\mathbf{u}$ is the vector of nodal displacements. The problem is to obtain the $\mathbf{u}$ that minimizes π and satisfies the linear equality constraint equations

$$\mathbf{b}_i^T \mathbf{u} = \gamma_i \qquad i = 1,2, \cdots, m \tag{2}$$

where $\mathbf{b}_i$ is the vector of coefficients for the ith constraint equation. Equation (2) may be the set of active constraints arising from inequality constraints such as those in contact problems. In most problems in structural mechanics $\mathbf{b}_i$ contains only a few non-zero terms. For example, when (2) is due to an incompressibility condition, then $\mathbf{b}_i^T\mathbf{u}$ is a linear function of the components of $\mathbf{u}$ that belong to the ith element and γ_i is simply zero. Equation (2) can be written in the more compact form

$$\mathbf{B}^T\mathbf{u} = \mathbf{g} \tag{3}$$

where $\mathbf{B}$ is an $n \times m$ matrix with $\mathbf{b}_i$ as its ith column, and $\mathbf{g}^T = (\gamma_1, \gamma_2, \cdots, \gamma_m)$.

The penalty function method, generally attributed to Courant,[3] incorporates the constraint equation (3) into the potential energy functional and therefore reduces the above constraint minimization problem to an unconstraint one. This is achieved by adding a fictitious energy term, also called penalty term, to equation (1) to get

$$\bar{\pi}(\mathbf{u}) = \pi(\mathbf{u}) + \frac{\kappa}{2}[(\mathbf{B}^T\mathbf{u} - \mathbf{g})^T\mathbf{W}(\mathbf{B}^T\mathbf{u} - \mathbf{g})] \tag{4}$$

where $\mathbf{W}$ is a symmetric positive definite $m \times m$ weighing matrix and κ is the penalty parameter. For simplicity, in many applications, $\mathbf{W}$ is chosen to be a diagonal matrix, and often the identity matrix (for more details see Reference 4). Physically, the penalty term is the energy associated with fictitious linear springs that are introduced to enforce the constraints. For example, in contact problems the springs are connected to each pair of points being in contact.

Remark. The penalty function method can also be derived from a perturbed Lagrangian functional.[7] Then, the elimination of the Lagrange parameters will lead directly to equation (4), where the penalty term depends on the perturbation in the potential energy for the Lagrange formulation.

The variation of $\bar{\pi}$ leads to

$$[\mathbf{D}\bar{\pi}(\mathbf{u})]^T\delta\mathbf{u} = [\mathbf{D}\pi(\mathbf{u})]^T\delta\mathbf{u} + \frac{\kappa}{2}[\mathbf{u}^T\mathbf{B}\mathbf{W}\mathbf{B}^T - \mathbf{g}^T\mathbf{W}\mathbf{B}^T]\delta\mathbf{u} = 0 \tag{5}$$

By the fundamental theory of variation, equation (5) becomes

$$[\mathbf{K} + \kappa\mathbf{B}\mathbf{W}\mathbf{B}^T]\mathbf{u} = \mathbf{f} + \kappa\mathbf{B}\mathbf{W}\mathbf{g} \tag{6}$$

The coefficient matrix remains symmetric positive definite but its profile structure may be different from that for $\mathbf{K}$.

ERROR ANALYSIS FOR PENALTY METHOD

There are two sources of error that affect the accuracy of analysis based on the penalty method. Both errors depend strongly on the penalty parameter, but in two completely different ways. The first error is due to large perturbation in the system of equations (6) that results from a small penalty parameter. This is due to the fact that the solution obtained by the penalty method is only exact in the limit $\kappa \to \infty$. Applying the Sherman–Morrison formula to obtain an expression for the inverse of the coefficient matrix in (6), we get a relation between the approximate solution, $\mathbf{u}$, and the penalty parameter, κ. Accordingly

$$\mathbf{u} = [\mathbf{K}^{-1} - \kappa\mathbf{K}^{-1}\mathbf{B}(\mathbf{W}^{-1} + \kappa\mathbf{B}^T\mathbf{K}^{-1}\mathbf{B})^{-1}\mathbf{B}^T\mathbf{K}^{-1}](\mathbf{f} + \kappa\mathbf{B}\mathbf{W}\mathbf{g}) \tag{7}$$

As the penalty parameter, $\kappa \to \infty$, $\mathbf{u}$ approaches the exact solution

$$\mathbf{u}_E = [\mathbf{K}^{-1} - \mathbf{K}^{-1}\mathbf{B}(\mathbf{B}^T\mathbf{K}^{-1}\mathbf{B})^{-1}\mathbf{B}^T\mathbf{K}^{-1}]\mathbf{f} + \mathbf{K}^{-1}\mathbf{B}(\mathbf{B}^T\mathbf{K}^{-1}\mathbf{B})^{-1}\mathbf{g} \tag{8}$$

Retaining only terms of order $1/\kappa$, the error in $\mathbf{u}$ becomes

$$\mathbf{u} - \mathbf{u}_E \approx \frac{1}{\kappa}\mathbf{K}^{-1}\mathbf{B}(\mathbf{B}^T\mathbf{K}^{-1}\mathbf{B})^{-1}\mathbf{W}^{-1}(\mathbf{B}^T\mathbf{K}^{-1}\mathbf{B})^{-1}(\mathbf{B}^T\mathbf{K}^{-1}\mathbf{f} - \mathbf{g}) \tag{9}$$

Taking norms, we get

$$\|\mathbf{u}-\mathbf{u}_E\| \approx \frac{1}{\kappa}\|\mathbf{K}^{-1}\mathbf{B}(\mathbf{B}^T\mathbf{K}^{-1}\mathbf{B})^{-1}\mathbf{W}^{-1}(\mathbf{B}^T\mathbf{K}^{-1}\mathbf{B})^{-1}(\mathbf{B}^T\mathbf{K}^{-1}\mathbf{f}-\mathbf{g})\|$$

$$\leq \frac{\|(\mathbf{B}^T\mathbf{K}^{-1}\mathbf{B})^{-1}\mathbf{W}^{-1}\|}{\kappa}\|[\mathbf{K}^{-1}\mathbf{B}(\mathbf{B}^T\mathbf{K}^{-1}\mathbf{B})^{-1}(\mathbf{B}^T\mathbf{K}^{-1}\mathbf{f}-\mathbf{g})]\|$$

The term in the square brackets is the contribution to $\mathbf{u}_E$ due to the constraint conditions (e.g. contact forces) and we may assume that this term is equal to $c\|\mathbf{u}_E\|$. Here, c is a constant that is close to unity in most cases. Then

$$\frac{\|\mathbf{u}-\mathbf{u}_E\|}{\|\mathbf{u}_E\|} \leq \frac{c\|(\mathbf{B}^T\mathbf{K}^{-1}\mathbf{B})^{-1}\mathbf{W}^{-1}\|}{\kappa} \tag{10}$$

The second source of error is due to the loss of information when a large quantity is added to a small one in the computer. For example, consider an environment where 8 digits of accuracy are used in all computations (the unit round-off error, $\epsilon = 10^{-8}$). Then a stiffness coefficient $k = \frac{1}{3}$ is represented as 0·33333333. If a penalty parameter, $\kappa = 10^3$, with an associated weight of unity, is added to this term the result will be $0{\cdot}10003333 \times 10^4$. Note that half of the digits in k are lost. Such errors were considered in Reference 8 and are bounded by

$$\frac{\|\mathbf{u}-\mathbf{u}_E\|}{\|\mathbf{u}_E\|} \leq n\epsilon\frac{\kappa}{k_{min}} \tag{11}$$

where k_{min} is the smallest stiffness coefficient that is modified by κ.

Adding the contributions from the two error bounds in (10) and (11) with $\mathbf{W}$ in (9) chosen to be the identity we obtain

$$\rho = n\epsilon\frac{\kappa}{k_{min}} + \frac{c\|(\mathbf{B}^T\mathbf{K}^{-1}\mathbf{B})^{-1}\|}{\kappa} \tag{12}$$

where ρ is the relative error bound for the solution $\mathbf{u}$ and represents the accuracy that can be attained with a given penalty parameter κ. ρ is a minimum when

$$\kappa = \kappa_{op} = \left[\frac{ck_{min}\|(\mathbf{B}^T\mathbf{K}^{-1}\mathbf{B})^{-1}\|}{n\epsilon}\right]^{1/2} \tag{13}$$

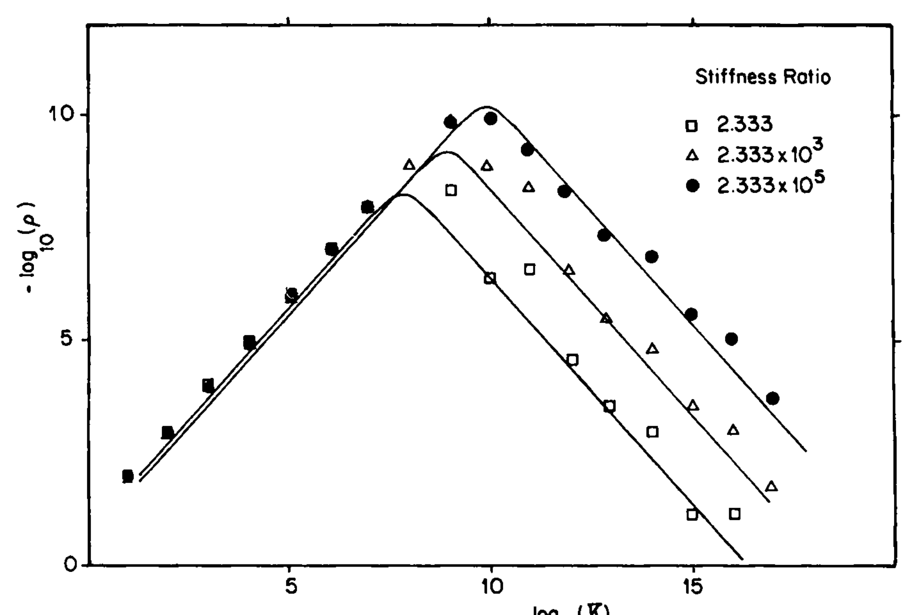

Figure 1. Effect of penalty parameter on the accuracy of the solution

$\|(\mathbf{B}^T\mathbf{K}^{-1}\mathbf{B})^{-1}\|$ is a measure of the stiffness of the parts of the structure affected by the constraint conditions. A lower bound to this quantity is k_{min}. Using this in (13), a lower bound for the optimum κ is

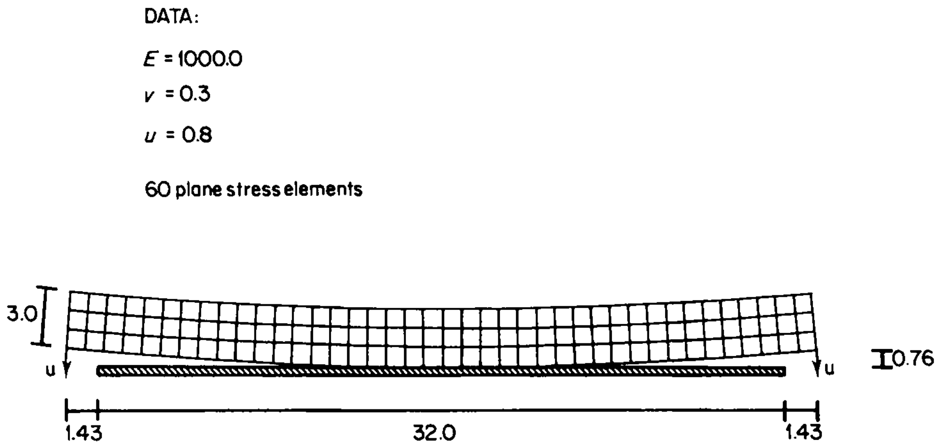

Figure 2. Geometry and data of the circular beam problem

$$\bar{\kappa} = k_{\min} \sqrt{\left(\frac{c}{n\epsilon}\right)} < \kappa_{\mathrm{op}} \tag{14}$$

In Figure 1 we plot $-\log_{10}(\rho)$ against $\log_{10}(\kappa)$ for a 2 degree-of-freedom contact problem. On the same plot we present the result of actual numerical experiments performed on a VAX 11/780 computer with $\epsilon \approx 10^{-17}$ (double precision).

Remark. Note that in this experiment we used $\mathbf{W} = \mathbf{I}$. In general, $\mathbf{W}$ must be chosen such that the errors that occur in each equation are of the same order. When $\mathbf{B}$ is an orthogonal matrix (i.e. $\mathbf{b}_i^{\mathrm{T}}\mathbf{b}_j = 0$, $i \neq j$), then a good choice for $\mathbf{W}$ would be a diagonal matrix with $(\mathbf{b}_i^{\mathrm{T}}\mathbf{K}\mathbf{b}_i)/(\mathbf{b}_i^{\mathrm{T}}\mathbf{b}_i)$ as the ith diagonal entry.

A NUMERICAL EXAMPLE

When inequality constraint conditions are involved, as in the case of contact problems, the system of equations resulting from (5) is nonlinear which may be solved using the Newton–Raphson method. In these cases the penalty parameter also affects the total number of Newton–Raphson iterations and therefore the cost of the analysis. As an example, we consider a circular beam in contact with a rigid foundation. The purpose of this example is to demonstrate the influence of the penalty parameter on the number of iterations. Due to the geometry of the problem (Figure 2), the middle of the beam will lift up. Using the symmetry condition, the beam is discretized by 60 plane stress elements. Imposed displacements were specified at the two ends of the beam. The deformed shape of the beam is presented in Figure 3. This problem is solved with a range of penalty parameters and the results are presented in Table I.

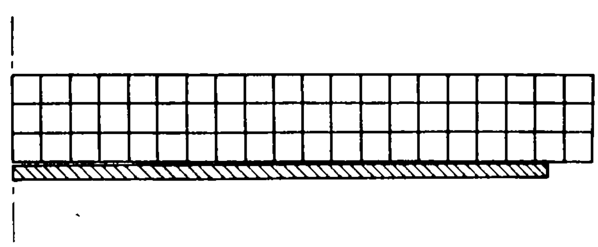

Figure 3. Deformed configuration of the circular beam

The optimum penalty parameter obtained from equation (6) is $\kappa_{\mathrm{opt}} = 10^8$. Here, optimality refers to accuracy in the constraint condition. The number of iterations for penalty parameters larger than κ_{opt} is the same as the number of iterations for κ_{opt}. For many applications lower accuracy may be sufficient. Then, underestimating the penalty parameter may translate into fewer iterations, as can be seen in Table I. In Reference 9 similar numerical results are used to arrive at suitable penalty parameters for the solution of contact problems.

Table I. Influence of penalty parameter on convergence

Penalty parameter	No. of iterations	Maximum penetration
10	4	$5{\cdot}27 \times 10^{-2}$
10^2	5	$2{\cdot}73 \times 10^{-2}$
10^3	6	$7{\cdot}09 \times 10^{-3}$
10^4	8	$9{\cdot}24 \times 10^{-4}$
10^6	8	$9{\cdot}28 \times 10^{-6}$
10^8	8	$9{\cdot}29 \times 10^{-8}$

The above example demonstrates the trade-off between accuracy and cost that exists when using penalty function methods. However, what is clear is that there is little to be gained with penalty parameters larger than κ_{opt}.

ACKNOWLEDGEMENTS

The first author gratefully acknowledges partial support by the Office of Naval Research under Contract N00014-76-C-0013, and the second author gratefully acknowledges support by the Deutsche Forschungsgemeinschaft.

REFERENCES

1. O. C. Zienkiewicz, *The Finite Element Method*, 3rd edn, McGraw-Hill, London, 1977.
2. B. M. Irons and K. J. Draper, 'Lagrange multiplier techniques in structural analysis', *A.I.A.A. J.*, **3**, *pp.* 1172–1174 (1965).
3. R. Courant, 'Variational methods for the solution of problems of equilibrium and vibrations', *Bull. Am. Math. Soc.*, **49**, 1–23 (1948).
4. D. G. Luenberger, *Linear and Nonlinear Programming*, 2nd edn, Addison-Wesley, Reading, Massachusetts, 1984.
5. C. A. Felippa, 'Error analysis of penalty function techniques for constraint definition in linear algebraic systems', *Int. j. numer. methods eng.* **11**, 709–728 (1977).
6. C. A. Felippa, 'Iterative procedures for improving penalty function solutions of algebraic systems', *Int. j. numer. methods eng.*, **12**, 821–836 (1978).
7. J. T. Oden, 'Exterior penalty methods for contact problems in elasticity', in *Nonlinear Finite Element Analysis in Structural Mechanics* (Eds. Wunderlich, Stein, Bathe), Springer-Verlag, New York, 1980.
8. J. H. Wilkinson, *Rounding Errors in Algebraic Problems*, Prentice-Hall, Englewood Cliffs, N.J., 1963.
9. J. O. Hallquist, 'NIKE2D: an implicit, finite deformation, finite element code for analyzing the static and dynamic response of two dimensional solids', *Rep. UCRL-52678*, Lawrence Livermore National Laboratory, Univ. of California (1982).
10. B. Nour-Omid and P. Wriggers, 'A two-level iteration method for solution of contact problems', *Comp. Meth. Appl. Mech. Eng.*, **54**, No. 2, 131–144 (1986).
11. P. Wriggers and B. Nour-Omid, 'Solution methods for contact problems', *Report No. UCB/SESM 84-09*, Univ. of California, Berkeley (1984).
12. P. Wriggers, 'Zur Berechnung von Stoss- und Kontaktproblemen mit Hilfe der Finite- Element Methode', *Bericht Nr. F81/1*, Forschungs- und Seminarberichte aus dem Bereich der Mechanik der Universitaet Hannover, Hannover (1981).