INTERNATIONAL JOURNAL FOR NUMERICAL METHODS IN FLUIDS, VOL. 2, 151–171 (1982)

ON PENALTY FUNCTION METHODS IN THE FINITE-ELEMENT ANALYSIS OF FLOW PROBLEMS†

J. N. REDDY

Department of Engineering Science and Mechanics, Virginia Polytechnic Institute and State University, Blacksburg, Va. 24061, U.S.A.

SUMMARY

In this paper the penalty function method is reviewed in the general context of solving constrained minimization problems. Mathematical properties, such as the existence of a solution to the penalty problem and convergence of the solution of a penalty problem to the solution of the original problem, are studied for the general case. Then the results are extended to a penalty function formulation of the Stokes and Navier–Stokes equations. Conditions for the equivalence of two penalty-finite element models of fluid flow are established, and the theoretical error estimates are verified in the case of Stokes's problem.

KEY WORDS Penalty method Incompressible Flow Finite Elements Convergence Existence

INTRODUCTION

In 1941, R. Courant[1,2] suggested a novel method of obtaining better convergence (of the derivatives of the solution) in the Rayleigh–Ritz method. The method, as applied to the equilibrium problem for a membrane ($\nabla^2 u = f$ in Ω and $u = 0$ on the boundary Γ of Ω) under external pressure f, can be described as follows: Instead of considering the usual variational problem of minimizing the functional

$$I(u) = \frac{1}{2}\int_\Omega \left[\left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2 + 2uf\right] dx\, dy,\ u = 0 \text{ on } \Gamma \tag{1}$$

the method seeks the minimum of a modified functional obtained from the original functional by adding terms of higher order which vanish for the actual solution u:

$$I_p(u, \alpha) = I(u) + \frac{1}{2}\int_\Omega \alpha(\nabla^2 u - f)^2\, dx\, dy \tag{2}$$

where α is an arbitrary (preselected) positive constant or function. Courant termed the functional in (2) a 'sensitized' functional since it is more sensitive to variations of u without changing the solution. Another example of the use of this idea is provided by the inclusion of the essential boundary condition in the Dirichlet problem. The modified functional is given

† Dedicated to my *guru*, Professor John Tinsley Oden on his 46th birthday.

0271-2091/82/020151-21$02.10

Received 15 August 1980
Revised 21 August 1980

by

$$I_p(u, \gamma) = I(u) + \frac{\gamma}{2} \int_\Gamma u^2 \, ds \tag{3}$$

For sufficiently large values of γ, the boundary value problem corresponding to the functional in (3) is almost equivalent to that associated with the functional in (1).

Although the idea was motivated by physical considerations, its value as a technique for transforming a given constrained minimization problem into a (sequence of) unconstrained minimization problem(s) was not recognized immediately. The idea was apparently not rigorously pursued for over a decade. In 1954 there was renewed interest in the *penalty function*† method[3] as a computational device in mathematical programming (see, for example, the 'logarithmic potential method' of Frisch,[4] and the 'inverse penalty function technique' of Carroll[5]). In 1956 Moser[2] proved convergence of the solution of the penalty problem to the solution of the original problem. In 1957 a very significant contribution was made by Rubin and Unger[6] which took the original technique of Courant out of the realm of conjecture for a much wider class of problems. They generalized Courant's technique to multiple variables and multiple equality constraints, and provided a convergence proof and a proof of the existence of Lagrange multipliers. Apart from these results, there was no significant theoretical development of the technique for a long time. However, the penalty function technique was often used as a computational device to approximate solutions of variational problems. There have been numerous papers devoted to various modifications of the method and their applications to particular problems; see References 7 and 8 for applications in mathematical programming and optimization, and Reference 9 for an application in meteorology.

The first use of the penalty function method in the finite-element analysis of a constrained minimization problem was apparently due to Babuska,[10] who proved the existence and uniqueness of the finite element solution to the penalty-function formulation of the Dirichlet problem for Poisson's equation (i.e. the finite-element formulation of the variational problem associated with the functional in equation (3)).

Despite its wide use in mathematical programming and optimization, the penalty function method was not regarded, until recently, as a powerful computational device. This is mainly due to two shortcomings: (i) the technique was used in connection with the approximate solution of variational problems by Rayleigh–Ritz type methods, which were themselves never regarded as competitive when compared to the traditional finite difference methods; (ii) in the practical application of the penalty function method, the penalty terms 'misbehave' without proper selection of the approximation functions or integration. These two shortcomings were overcome by the finite element method (and *reduced integration* techniques). In 1973, the penalty function method was introduced into the finite-element analysis of fluid flow problems by Zienkiewicz.[11] However, the second shortcoming was not overcome until Zienkiewicz, Taylor and Too[12] devised, rather ingeniously, the so-called reduced integration technique, which was later used by Zienkiewicz and his colleagues[13,14] in the numerical integration of the penalty terms. Thus, over three decades after the original idea was suggested, the penalty function method was brought into the realm of computational mechanics (especially into finite element analysis) where it now serves as a simple yet

† The word 'penalty function' was first used by T. N. Edelbaum in Chapter 1 of the book on optimization techniques edited by Leitmann.[3]

effective computational technique of handling physical as well as mathematical constraints. Exploitation of further generalizations and extensions of the technique in the finite-element solution of a variety of engineering problems is the current state of the technique.[15–24]

Following this introduction to the historical development of the penalty function method, the basic idea of the method and the mathematical properties are reviewed for the abstract (linear) problem of determining the minimum of a *quadratic functional* subjected to an equality constraint. As a specific example of the abstract problem the Stokes problem associated with the slow laminar motion of an incompressible fluid is presented. The abstract problem enables one to apply the penalty method to any constrained minimization problem. Subsequent sections are devoted to the application of the penalty function method to the stationary Navier–Stokes equations. Existence and uniqueness of solutions to the penalty function formulation of the Navier–Stokes equations are discussed and error estimates are given. Conditions for the equivalence of two penalty models of fluid flow are established, and the influence of the penalty parameter on the solution is investigated; the theoretical error estimates are verified numerically for the Stokes problem.

THE PENALTY FUNCTION METHOD

Consider the following variational problem: Find the minimum of the functional,

$$I(u) = \int_\Omega F(x, y, u, u_x, u_y)\, dx\, dy \tag{4}$$

in a Hilbert space H_1, subject to the constraint,

$$G(u) = 0 \tag{5}$$

where G is, in general, a non-linear operator from H_1 into some Hilbert space H_2. The solution u belongs to a subspace of H_1.

The problem is ordinarily solved by means of the Lagrange multiplier method (saddle-point problem), which seeks to find the stationary values (u, λ) of the modified functional,

$$L(u, \lambda) = I(u) + \int_\Omega \lambda G(u)\, dx\, dy \tag{6}$$

on the product space $H_L = H_1 \times H_2$. Here λ denotes the Lagrange multiplier.

The penalty function method reduces problems of conditional (or constrained) extremum to problems without constraints by the introduction of a penalty on the infringement of constraints. Instead of solving the original problem, the minimum of the functional

$$J_n(u) = I(u) + \tfrac{1}{2}\alpha_n \|G(u)\|^2_{H_2} \tag{7}$$

on the whole of H_1 is sought for some $\alpha_n > 0$. Here $\|\cdot\|_{H_2}$ denotes the norm in H_2.

The following theorem (see Polyak[25]) guarantees the existence of solution (u_n, λ_n) to the penalty problem.

Theorem 1

Let the following assumptions be satisfied:
(i) There exists a local point of minimum u_0 in H_1 of $I(u)$; that is,

$$\begin{aligned} &\text{(a) } I(u_0) \leq I(u), \text{ for every } u \text{ in } H_1 \\ &\text{(b) } G(u_0) = 0 \\ &\text{(c) if } G(u) = 0, \text{ then } \|u - u_0\|_{H_1} \leq \varepsilon,\ \varepsilon > 0 \end{aligned} \tag{8}$$

(ii) In an ε-neighborhood of u_0 the first and second (Gateaux) derivatives of I and G exist, and the second derivatives satisfy the Lipschitz conditions

$$\begin{aligned} \|\delta^2 I(u; \eta, \xi) - \delta^2 I(v; \eta, \xi)\|_{H_2} &\leq C_1 \|u - v\|_{H_1} \\ \|\delta^2 G(u; \eta, \xi) - \delta^2 G(v; \eta, \xi)\|_{H_2} &\leq C_1 \|u - v\|_{H_1} \end{aligned} \tag{9}$$

(iii) The adjoint A^* of the linear operator $A \equiv \delta G(u_0, \cdot)$ has a continuous inverse on H_2,

$$\|A^* q\| \geq \gamma \|q\|_{H_2}, \qquad \gamma > 0, \text{ for all } q \text{ in } H_2 \tag{10}$$

Then the Lagrange multiplier, P_0, exists in H_2 such that

$$\delta_u L(u_0, P_0) = \delta I(u_0) + A^*(P_0) \tag{11}$$

where

$$L(u, P) = I(u) + (P, G(u)) \tag{12}$$

Here, $(\cdot, \cdot)$ denotes inner product in H_2.

In addition to above assumptions, let the following assumption be satisfied:

(iv) The linear self-adjoint operator, $\Lambda \equiv \delta_u^2 L(u_0, P_0; \cdot, \cdot)$ is positive-definite in the sense that there exists $M > 0$ such that

$$(\Lambda\xi, \xi) \equiv (\delta^2 I(u_0, \xi, \xi), \xi) + (P_0, \delta^2 G(u_0, \xi, \xi)) \geq M \|\xi\|^2 \tag{13}$$

for every ξ in H_1.

Then there exists a u_n which is the unique point of local minimum of $J_n(u)$ in an ε-neighborhood of u_0, and an approximation to the Lagrange multiplier,

$$\lambda_n \equiv \alpha_n G(u_n) \tag{14}$$

such that

$$\|u_n - u_0\|_{H_1} \leq \frac{C_1}{2\alpha_n} \|\lambda_0\|_{H_2} \tag{15}$$

$$\|\lambda_n - \lambda_0\|_{H_2} \leq \frac{C_1}{2\alpha_n} \|\lambda_0\|_{H_2} \tag{16}$$

where the constant C_1 is independent of α_n.

An elegant proof of the theorem is given by Polyak;[25] see also References 26 and 27. In the next section we consider a special case of the abstract problem presented here.

APPLICATION TO STOKESIAN FLOWS

Here we consider an example of the theory presented in the previous section. We consider the Stokes problem, which consists of determining the solution (u, v, P) to the equations governing the slow, two-dimensional flow of a viscous incompressible fluid:

$$2\mu\frac{\partial^2 u}{\partial x^2}+\mu\frac{\partial}{\partial y}\left(\frac{\partial u}{\partial y}+\frac{\partial v}{\partial x}\right)-\frac{\partial P}{\partial x}=f_x \tag{17}$$

$$2\mu\frac{\partial^2 v}{\partial y^2}+\mu\frac{\partial}{\partial x}\left(\frac{\partial u}{\partial y}+\frac{\partial v}{\partial x}\right)-\frac{\partial P}{\partial y}=f_y, \text{ in } \Omega$$

$$\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}=0 \tag{18}$$

Here (u, v) denote the velocity components, P the pressure, μ the viscosity, and f_x and f_y denote the body force components. The velocity field must also satisfy certain boundary conditions of the problem. For simplicity we assume that $u = v = 0$ on the boundary $\partial\Omega$ of Ω. The problem can be viewed as one of seeking the solution (u, v) in $H_1(\Omega) = H_0^1(\Omega) \times H_0^1(\Omega)$ such that equation (18) is satisfied and the functional

$$I(u, v)=\int_\Omega \left\{\mu\left[\left(\frac{\partial u}{\partial x}\right)^2+\left(\frac{\partial v}{\partial y}\right)^2+\frac{1}{2}\left(\frac{\partial u}{\partial y}+\frac{\partial v}{\partial x}\right)^2\right]+f_x u+f_y v\right\} dx\, dy \tag{19}$$

is minimized. The pressure drops out of the functional owing to the fact that the velocity field satisfies the incompressibility condition (18) identically. Once the velocity field is known, the pressure can be calculated from equation (17) (or from the Poisson equation for pressure). Clearly, the Stokes problem has the same form as the abstract problem in equations (4)–(5), with

$$G(u, v)\equiv\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}, \qquad G: [H_0^1(\Omega)]^2 \to L_2(\Omega) \tag{20}$$

The functional in the Lagrange multiplier method is given by

$$L(u, v, \lambda)=I(u, v)+\int_\Omega \lambda G(u, v)\, dx\, dy \tag{21}$$

Computing the first variation of $L(\cdot)$ and determining the Euler equations, one observes (by comparing the Euler equations with (17)–(18) that the Lagrange multiplier $\lambda \in L_2(\Omega)$ is indeed the negative of the pressure,

$$\lambda = -P \tag{22}$$

The penalty functional in equation (7) becomes[28,29]

$$J_n(u_n, v_n)=I(u_n, v_n)+\frac{\alpha_n}{2}\int_\Omega [G(u_n, v_n)]^2\, dx\, dy \tag{23}$$

The Lagrange multiplier $\lambda = -P$ is given by (it can also be verified by comparing the Euler equations of J with those of L),

$$P_n = -\alpha_n\left(\frac{\partial u_n}{\partial x}+\frac{\partial v_n}{\partial y}\right) \tag{24}$$

Note that equation (24) is deduced from the penalty functional, J_n; it does not form the basis of the penalty method, as implied in a number of papers on the subject, but is a consequence of equation (23). If one starts with equation (24) to describe the penalty function method, not only is the reader misinformed but he is also confused because there is no obvious reason to assume that the pressure P is related to the velocity field through equation (24).

An alternative, but equivalent, penalty function formulation of the Stokes problem is possible (only in retrospect!). If we replace the continuity equation (18) by equation (24) (and replacing P in equation (17) by P_n), we obtain the functional (compare this with the Lagrange multiplier functional),

$$\hat{J}_n(u_n, v_n, P_n) = L(u_n, v_n, -P_n) - \frac{1}{2\alpha_n}\int_\Omega P_n^2\,\mathrm{d}x\,\mathrm{d}y \tag{25}$$

Note that as the limit $n \to \infty$ (i.e., $\alpha_n \to \infty$), the functional $\hat{J}_n$ approaches L with $(u_n, v_n, P_n) \to (u, v, P)$. The error estimate in equations (15) and (16) becomes

$$[\|u_n - u\|_1^2 + \|v_n - v\|_1^2]^{\frac{1}{2}} \leq \frac{C_1}{2\alpha_n}\|P\|_0$$

$$\|P_n - P\|_0 \leq \frac{C_1}{2\alpha_n}\|P\|_0$$

These error estimates imply that the solution (u_n, v_n) to the penalty problem corresponding to the Stokes problem converges to the true solution (u, v), and that the error is proportional to $(1/\alpha_n)$. In the section on the numerical results we will show that this theoretical estimate is also confirmed by the numerical experiment.

It should be pointed out that Theorem 1 holds only for linear problems which can be cast as one of minimizing a functional subject to a constraint. In the case of Navier–Stokes (N–S) equations, Theorem, 1 does not hold and therefore an independent result is needed. In the next section we investigate the questions of existence and uniqueness of solutions to the steady Navier–Stokes equations for two-dimensional incompressible flows, and then establish that the penalty solution converges to the true solution (existence of the true solution has been already established by Girault and Raviart[30]).

EXISTENCE AND UNIQUENESS OF SOLUTIONS TO N–S EQUATIONS

Here we present the existence and uniqueness (under certain conditions) of the solution to the penalty problem associated with the Navier–Stokes equations. Since the alternative penalty formulation (see equation (25)) resembles the Lagrange multiplier (or mixed) formulation, the results of Girault and Raviart[30] for the mixed formulation can easily be extended to the penalty problem. First certain mathematical preliminaries are in order.

Mathematical preliminaries

Let Ω be a bounded domain in the two-dimensional Euclidean space $\mathbb{R}^2$ with a Lipschitz-continuous boundary $\partial\Omega$. A typical point in Ω will be denoted by $x=(x_1, x_2)$, and let $L_2(\Omega)$ be the space of all square integrable functions defined on Ω. The inner product and the norm in $L_2(\Omega)$ are given by[31]

$$(u, v)_0 = \int_\Omega u(x)v(x)\,\mathrm{d}(x), \qquad \|u\|_{0,\Omega} = \sqrt{(u, u)} \tag{26}$$

For any integer $m \geq 0$ the Hilbert space of order m on Ω, $H^m(\Omega)$, is defined by[31]

$$H^m(\Omega) = \{u: u \in L_2(\Omega),\ D^\alpha u \in L_2(\Omega),\ |\alpha| \leq m\} \tag{27}$$

equipped with the inner product and norm,

$$(u, v)_m = \sum_{|\alpha|\leq m} (D^\alpha u, D^\alpha v)_0, \qquad \|u\|_{m,\Omega} = \left(\sum_{|\alpha|\leq m} \|D^\alpha u\|^2_{0,\Omega}\right)^{\frac{1}{2}} \tag{28}$$

Herein (27) and (28), the multi-index notation is used: $\alpha=(\alpha_1, \alpha_2)$, α_i are non-negative integers, and

$$D^\alpha = \frac{\partial^{\alpha_1}}{\partial x_1}\frac{\partial^{\alpha_2}}{\partial x_2}, \qquad |\alpha| = \alpha_1 + \alpha_2 \tag{29}$$

Further, let $L_p(\Omega)$ denote the space of all pth integrable functions defined on Ω, with norm,

$$\|u\|_{L_p(\Omega)} = \left(\int_\Omega [u(x)]^p\,\mathrm{d}x\right)^{1/p} < \infty \tag{30}$$

and let

$$H_0^m(\Omega) = \{u: u \in H^m(\Omega),\ D^\alpha u\,|_{\partial\Omega} = 0,\ |\alpha| < m\} \tag{31}$$

It is well known that for every $u \in H_0^1(\Omega)$, there exists a constant $c_1 = c_1(\Omega)$ such that[31]

$$\|u\|_{0,\Omega} \leq c_1\,|u|_{1,\Omega}, \tag{32}$$

which is known as the *Poincaré–Friedrich inequality*. Hence, the seminorm is a norm over the space $H_0^1(\Omega)$ for all functions defined over the bounded set, Ω. Also, by Sobolev imbedding theorems, it follows that the imbedding $H^1(\Omega) \to L_4(\Omega)$ is compact, and

$$\|u\|_{L_4(\Omega)} \leq c_2\,\|u\|_{1,\Omega} \tag{33}$$

Penalty function formulation of the Navier–Stokes equations

Consider the stationary (steady) Navier–Stokes equations governing an incompressible viscous fluid confined in Ω: Find the velocity field, $u=(u_1, u_2)$ and the pressure, P defined over Ω such that,

$$\begin{aligned} -\nu\nabla^2 u + \sum_{i=1}^{2} u_i \frac{\partial u}{\partial x_i} + \operatorname{grad} P = f \text{ in } \Omega \\ \operatorname{div} u = 0 \text{ in } \Omega, \qquad u = 0 \text{ on } \partial\Omega \end{aligned} \tag{34}$$

where $f=(f_1, f_2)$ is the body force vector, ν is the kinematic viscosity of the fluid, and P is the pressure divided by the density. The weak variational formulation of the Navier–Stokes

equations (34) involves seeking a pair (u, P) $[H_0^1(\Omega)]^2 \times \tilde{H}^0(\Omega)$ such that

$$\int_\Omega \left\{ \nu \frac{\partial u_i}{\partial x_j} \frac{\partial v_i}{\partial x_j} + u_j \frac{\partial u_i}{\partial x_j} v_i - f_i v_i \right\} dx = 0 \tag{35}$$

for all $v \in [H_0^1(\Omega)]^2$, subject to the (constraint) condition,

$$\operatorname{div} u = 0 \tag{36}$$

Here we used the notation,

$$[H_0^1(\Omega)]^2 \equiv H_0^1(\Omega) \times H_0^1(\Omega), \quad \tilde{H}_0(\Omega) = \left\{ P: P \in L_2(\Omega), \int_\Omega P \, dx = 0 \right\} \tag{37}$$

That is, the solution (u, P) of (34) belongs to the space

$$X = \{(u, P): u \in [H_0^1(\Omega)]^2, P \in \tilde{H}^0(\Omega), \operatorname{div} u = 0 \text{ in } \Omega\} \tag{38}$$

Next we define the following bilinear and trilinear forms,

$$\begin{aligned} G(u, v) &= \int_\Omega \operatorname{div} u \operatorname{div} v \, dx \\ B(u, v) &= \int_\Omega \nabla u \cdot \nabla v \, dx \\ N(u, v, w) &= \int_\Omega u_i \frac{\partial v_j}{\partial x_i} w_j \, dx \end{aligned} \tag{41}$$

The penalty formulation of equations (34) involves seeking $u = u_\varepsilon$ such that

$$\text{(a)} \quad u_\varepsilon \in [H_\varepsilon^1(\Omega)]^2 = \{u: u \in [H_0^1(\Omega)]^2, \|u\|_{1,\varepsilon,\Omega} = \|u\|_{1,\Omega} + \varepsilon^{-1} G(u, u)\}, \tag{39}$$

$$\text{(b)} \quad \nu B(u, v) + N(u, u, v) + \varepsilon^{-1} G(u, v) = (f, v), \; v \in [H_\varepsilon^1(\Omega)]^2 \tag{40}$$

where ε is the penalty parameter (inverse of α_n in (23)).

The Lagrange multiplier, $P_\varepsilon \in \tilde{H}^0(\Omega)$, can be computed from,

$$P_\varepsilon = -\varepsilon^{-1} \operatorname{div} u_\varepsilon \tag{42}$$

Existence and Uniqueness

Theorem 2. There exists a solution to the penalty-variational problem (40) for any ε, $0 < \varepsilon < 1$. The solution is unique for sufficiently large ν.

Proof. First it is shown that the auxiliary problem,

$$A_u(w, v) + \varepsilon^{-1} G(w, v) = (f, v), \; A_u(\cdot, \cdot) = \nu B(\cdot, \cdot) + N(u, \cdot, \cdot) \tag{43}$$

has a unique solution $w \in [H_\varepsilon^1(\Omega)]^2$ for fixed $u \in [H_\varepsilon^1(\Omega)]^2$. Then the mapping $Tu = w$ is shown to have a fixed point, thereby the proof of existence is completed. The part of the proof which shows that the mapping T has a fixed point is too long and will not be repeated here. For details see References 32 and 36.

Existence of a unique solution to (43) is guaranteed by the Lax–Milgram theorem. That is, if $A_u(\cdot, \cdot)$ is $[H_0^1(\Omega)]^2$-elliptic (continuity is obvious) it follows that equation (43) has a

solution w in $[H_\varepsilon^1(\Omega)]^2$. To this end note that, for $u, v, w \in [H_0^1(\Omega)]^2$,

$$-[N(u, v, w)+N(u, w, v)]=M(u, v\cdot w),\ M(u, Q)=\int_\Omega (\operatorname{div} u)Q\,dx \tag{44}$$

and the Brezzi condition,[33]

$$\begin{aligned}\sup_{v\in[H_0^1(\Omega)]^2}\frac{-M(v, Q)}{\|v\|_{1,\Omega}} &= \sup_{v\in[H_0^1(\Omega)]^2}\frac{-\int_\Omega (\operatorname{div} v)Q\,dx}{\|v\|_{1,\Omega}}\\ &= \sup_{v\in[H_0^1(\Omega)]^2}\frac{\int_\Omega v\ \operatorname{grad} Q\,dx}{\|v\|_{1,\Omega}}\equiv\|\operatorname{grad} Q\|_{[H^{-1}(\Omega)]^2}\\ &= c_3\|Q\|_{\tilde{H}^0(\Omega)}\end{aligned} \tag{45}$$

where $H^{-1}(\Omega)$ is the dual (space) of $H_0^1(\Omega)$, $H^{-1}(\Omega)\equiv(H_0^1(\Omega))'$. Therefore,

$$|M(u, v\cdot v)|=\left|\int_\Omega |v|^2 \operatorname{div} u\,dx\right|\le\|v\|_{L_4(\Omega)}\|\operatorname{div} u\|_0$$

$$\le c_1^2c_2^2\,|v|_{1,\Omega}^2\,\|\operatorname{div} u\|_0$$

or

$$M(u, v\cdot v)\ge -c_1^2c_2^2\,|v|_{1,\Omega}^2\,\|\operatorname{div} u\|_0$$

Then

$$A_u(v, v)=\nu\,|v|_{1,\Omega}^2-\tfrac{1}{2}M(u, v\cdot v)\ge\left(\nu-\frac{c_1^2c_2^2}{2}\|\operatorname{div} u\|_0\right)|v|_{1,\Omega}^2$$

Choosing α such that, $0<\alpha<\nu$,

$$\nu-\frac{c_1^2c_2^2}{2}\|\operatorname{div} u\|_0\ge\alpha,\ \text{or}\ \|\operatorname{div} u\|_0\le\frac{2(\nu-\alpha)}{c_1^2c_2^2}\equiv l \tag{46}$$

one has

$$A_u(v, v)\ge\alpha\,|v|_{1,\Omega}^2 \tag{47}$$

for $u\in C\equiv\{u: u\in[H_0^1(\Omega)]^2, \|\operatorname{div} u\|_0\le l\}$. That is, $A_u(\cdot,\cdot)$ is $[H_0^1(\Omega)]^2$-elliptic for $u\in C$. Hence there exists a unique solution $w\in[H_\varepsilon^1(\Omega)]^2$ for every $u\in C$. The mapping $T: C\to C$ defined by

$$Tu=w \tag{48}$$

has a fixed point in C. This completes the proof of the theorem.

Next we estimate the error between the solution (u_ε) to the penalty problem and the solution (u) to the original problem. The alternative penalty formulation involves seeking $(u_\varepsilon, P_\varepsilon)\in[H_\varepsilon^1(\Omega)]^2\times\tilde{H}^0(\Omega)\equiv\tilde{X}_\varepsilon$ such that

$$\begin{aligned}\nu B(u_\varepsilon, v)+N(u_\varepsilon, u_\varepsilon, v)-M(v, P_\varepsilon)&=(f, v)_0\\ -M(u_\varepsilon, Q)&=\varepsilon(P_\varepsilon, Q)_0\end{aligned} \tag{49}$$

for every $(v, Q) \in \mathring{X}_\varepsilon$ and $0 < \varepsilon < 1$. The mixed formulation of the Navier–Stokes equations involves determining $(u, P) \in [\dot{H}_0^1(\Omega)]^2 \times \tilde{H}^0(\Omega) \equiv \mathring{X}$ such that

$$\begin{aligned} \nu B(u, v) + N(u, u, v) - M(v, P) &= (f, v)_0 \\ M(u, Q) &= 0 \end{aligned} \tag{50}$$

for every $(v, Q) \in \mathring{X}$. Existence of solution $(u, P) \in [H_0^1(\Omega)]^2 \times \tilde{H}^0(\Omega)$ to (50) can be proved using a generalized Lax–Milgram theorem due to Brezzi.[33] Existence of solution $(u_\varepsilon, P_\varepsilon) \in \mathring{X}_\varepsilon$ to the penalty formulation (49) follows along the same lines. Alternatively, since the solution $u_\varepsilon \in [H_\varepsilon^1(\Omega)]^2$ to (40) exists and is unique for $u_\varepsilon \in C$, it follows from the second equation in (49) that P_ε is unique.

Subtracting equation (49) from equation (50), we get

$$\begin{aligned} \nu B(u - u_\varepsilon, v) + N(u, u, v) - N(u_\varepsilon, u_\varepsilon, v) - M(v, P - P_\varepsilon) &= 0 \\ -M(u - u_\varepsilon, Q) - \varepsilon(P - P_\varepsilon, Q)_0 &= -\varepsilon(P, Q)_0 \end{aligned} \tag{51}$$

for every $(v, Q) \in \mathring{X}$. It can be shown that

$$|u|_{1,\Omega} \le \frac{c_1}{\alpha} \|f\|_0, \qquad |u_\varepsilon|_{1,\Omega} \le \frac{c_1}{\alpha} \|f\|_0 \tag{52}$$

Further,

$$\begin{aligned} \nu \, |u_\varepsilon - u|_{1,\Omega}^2 &= \nu B(u - u_\varepsilon, u - u_\varepsilon) = N(u, u, u_\varepsilon - u) - N(u_\varepsilon, u_\varepsilon, u_\varepsilon - u) - M(u_\varepsilon - u, P - P_\varepsilon), \\ &= [N(u, u, u_\varepsilon - u) - N(u_\varepsilon, u_\varepsilon, u_\varepsilon - u)] - \varepsilon \, \|P - P_\varepsilon\|_0^2 + \varepsilon(P, P - P_\varepsilon)_0 \end{aligned} \tag{53}$$

Now consider the term in the square brackets. Assuming that the trilinear form $N(\cdot, \cdot, \cdot)$ satisfies the condition,

$$\beta = \sup_{u,v,w \in [H_\varepsilon^1(\Omega)]^2} \frac{N(u, v, w)}{|u|_{1,\Omega} \, |v|_{1,\Omega} \, |w|_{1,\Omega}} \tag{54}$$

we write

$$\begin{aligned} |[\cdot - \cdot]| &= |N(u_\varepsilon, u - u_\varepsilon, u_\varepsilon - u) + N(u - u_\varepsilon, u, u_\varepsilon - u)| \\ &\le \tfrac{1}{2} c_1^2 c_2^2 \, l \, |u - u_\varepsilon|_{1,\Omega}^2 + \beta \, |u - u_\varepsilon|_{1,\Omega}^2 \, |u|_{1,\Omega} \end{aligned} \tag{55}$$

Using the elementary inequality $ab \le \dfrac{1}{4\theta} a^2 + \theta b^2$, $\theta > 0$, equation (53) can be written as

$$\left(\nu - \tfrac{1}{2} c_1^2 c_2^2 l - \frac{\beta c_1}{\alpha} \|f\|_0\right) |u - u_\varepsilon|_{1,\Omega}^2 \le \frac{\varepsilon}{4\theta} \|P\|_0^2 + \varepsilon\theta \, \|P - P\|_0^2 - \varepsilon \, \|P - P_\varepsilon\|_0^2$$

Choosing $\theta = 1$ in the above equation and letting $\eta = \alpha - \dfrac{\beta c_1}{\alpha} \|f\|_0$, we obtain

$$|u - u_\varepsilon|_{1,\Omega} \le c\sqrt{\varepsilon} \, \|P_0\|_0 \tag{56}$$

Thus, the solution (u_ε) to the penalty problem in (49) converges to the solution (u) of the mixed problem in (50) as ε goes to zero. The rate of convergence is only 1/2 as opposed to 1 in the Stokes flow.[28,34]

Existence of solutions to the discrete problem (e.g., finite-element approximation) associated with equation (40) can be established only if the Brezzi condition holds for the finite-dimensional spaces. This is an area where much research needs to be done; in this

connection the works of Oden and his colleagues[20,27,35] should provide a starting point. It should be pointed out that the parameter β in equation (54) should be a constant independent of the mesh parameter,[33] or a constant dependent on a positive power of the mesh parameter, $\beta = h^{\sigma}$, $\sigma > 0$ (see Reference 35).

EQUIVALENCE OF THE FINITE ELEMENT MODELS OF FLUID FLOW

In this section we establish the equivalence of the numerical models based on the functionals in equations (23) and (25). Although the present discussion is focused on the Stokes problem, the discussion given below is equally valid, as will be seen, for Navier–Stokes equations.

Let $\mathbb{U}_n \subset [H_{\varepsilon}^1(\Omega)]^2$, $U_h \subset H_{\varepsilon}^1(\Omega)$, and $V_h \subset \tilde{H}^0(\Omega)$ be the finite-dimensional subspaces, and define

$$\begin{aligned} B(u, v) &= \nu \int_{\Omega} \operatorname{grad} u \cdot \operatorname{grad} v \, dx \\ G(u, v) &= \int_{\Omega} \operatorname{div} u \cdot \operatorname{div} v \, dx \\ M(u, Q) &= \int_{\Omega} (\operatorname{div} u) Q \, dx \end{aligned} \tag{57}$$

Now we give the discrete analogues of the two penalty models.

Penalty Model I (conventional)

The approximate problem associated with functional in (23) involves seeking $u_h^{\varepsilon} \in \mathbb{U}_h$ such that

$$B(u_h^{\varepsilon}, v_h) + \alpha_n G(u_h^{\varepsilon}, v_h) = (f, v_h)_0 \tag{58}$$

for all $v_h \in \mathbb{U}_h$.

Penalty Model II (mixed)

The discrete analogue of the variational problem associated with the functional in (25) involves finding the pair $(\bar{u}_h^{\varepsilon}, P_h^{\varepsilon}) \in \mathbb{U}_h \times V_h$ such that

$$\begin{aligned} B(\bar{u}_h^{\varepsilon}, v_h) - M(v_h, P_h^{\varepsilon}) &= (f, v_h)_0 \\ -M(\bar{u}_h^{\varepsilon}, Q_h) &= \alpha_n^{-1} (P_h^{\varepsilon}, Q_h)_0 \end{aligned} \tag{59}$$

for every $v_h \in \mathbb{U}_h$ and $Q_h \in V_h$. The equations (59) have a solution provided $M(\cdot, \cdot)$ satisfies the approximate Brezzi condition:

$$\sup_{v_h \in \mathbb{U}_h} \frac{-M(v_h, Q_h)}{|v_h|_{1,\Omega}} \geq \delta \, \|Q_h\|_0, \qquad Q_h \in V_h \tag{60}$$

It should be pointed out that the continuous Brezzi condition (45) does not imply, in general, the approximate Brezzi condition (60). A sufficient condition for (60) to hold was given by Mercier.[36] If the approximate Brezzi condition is satisfied, then the problem (59) has a unique solution (the proof is similar to that given in Theorem 2).

Equivalence of the models

The equivalence of the two penalty models described above can be established for the following case:

$$\text{div}\,(\mathbb{U}_h) \subset V_h, \mathbb{U}_h = U_h \times U_h. \tag{61}$$

Theorem 3. The conventional and penalty models (Models I and II) are equivalent if the condition in (61) holds.

Proof. We prove the equivalence using ideas similar to those employed by Oden and Reddy.[37] Subtracting the first of equation (59) from (58), we get

$$B(u_h^\varepsilon - \bar{u}_h^\varepsilon, v_h) + \alpha_n G(u_h^\varepsilon, v_h) + M(v_h, P_h^\varepsilon) = 0 \tag{62}$$

Since $\text{div}\,(\mathbb{U}_h) \subset V_h$, every element v_h in $\mathbb{U}_h$ is of the form, $Q_h = \text{div}\, v_h$, $Q_h \in V_h$. Hence, from the second equation in (59), we have

$$\begin{aligned} -M(\bar{u}_h^\varepsilon, \text{div}\, v_h) &= \alpha_n^{-1}(P_h^\varepsilon, \text{div}\, v_h)_0 \\ -\alpha_n G(\bar{u}_h^\varepsilon, v_h) &= M(v_h, P_h^\varepsilon) \end{aligned} \tag{63}$$

Using (63) in (62), we obtain

$$\hat{B}(u_h^\varepsilon - \bar{u}_h^\varepsilon, v_h) \equiv B(u_h^\varepsilon - \bar{u}_h^\varepsilon, v_h) + \alpha_n G(u_h^\varepsilon - \bar{u}_h^\varepsilon, v_h) = 0, \quad \text{for all } v_h \in \mathbb{U}_h \tag{64}$$

Since $\hat{B}(\cdot, \cdot)$ is coercive on $[H_0^1(\Omega)]^2$, it follows that $u_h^\varepsilon = \bar{u}_h^\varepsilon$.

In an independent study Malkus and Hughes[38] (see also Reference 39) discussed the equivalence of certain mixed finite element methods with displacement methods which employ reduced and selective integration techniques. These are similar to the model discussed here. It should be cautioned that the Lagrange multiplier model is *not* equivalent to Model II.

A NUMERICAL EXAMPLE

In this section we illustrate via an example problem some of the ideas discussed in the previous sections. Specifically, we discuss the equivalence in the light of exact and reduced integrations, investigate the influence of the penalty parameter α_n on the accuracy, and verify the error estimates for Stokes's problem.

Finite-element models

Let u, v and P be interpolated, in a typical element Ω^e, by

$$u = \sum_i^r u_i N_i, \qquad v = \sum_i^r v_i N_i, \qquad P = \sum_i^s P_i N_i^\phi \tag{65}$$

where N_i and N_i^ϕ are the element interpolation functions ($r \geq s$).

Substituting (65) for u and v into the first variation of the functional in (23), we obtain

$$\begin{bmatrix} [K^{11}]+\alpha_n[S^{11}] & [K^{12}]+\alpha_n[S^{12}] \\ \text{symm.} & [K^{22}]+\alpha_n[S^{22}] \end{bmatrix} \begin{Bmatrix} \{u\} \\ \{v\} \end{Bmatrix} = \begin{Bmatrix} \{F^1\} \\ \{F^2\} \end{Bmatrix} \tag{66}$$

where

$$[K^{11}]=\mu(2[S^{11}]+[S^{22}]),\ [K^{12}]=\mu[S^{12}]^T$$
$$[K^{22}]=\mu([S^{11}]+2[S^{22}]),\ F_1^\alpha=\oint_{\partial\Omega^e} t_\alpha N_i\,ds+\int_{\Omega^e} f_\alpha N_i\,dx\,dy \tag{67}$$
$$S_{ij}^{\alpha\beta}=\int_{\Omega^e} N_{i,\alpha}N_{j,\beta}\,dx\,dy,\qquad N_{i,\alpha}=\partial N_i/\partial x_\alpha,\quad (\alpha,\beta=0,1,2),$$

$x_1=x$, $x_2=y$, and $t_\alpha(\alpha=1,2)$ denote the surface tractions on the boundary $\partial\Omega^e$ of the element, and f_α denote the body forces.

Substituting (65) into the first variation of the functional in (25), we obtain,

$$\begin{bmatrix} [K^{11}] & [K^{12}] & -[K^{13}] \\ [K^{12}]^T & [K^{22}] & -[K^{23}] \\ -[K^{13}]^T & -[K^{23}]^T & -\dfrac{1}{\alpha_n}[K^{33}] \end{bmatrix} \begin{Bmatrix} \{u\} \\ \{v\} \\ \{P\} \end{Bmatrix} = \begin{Bmatrix} \{F^1\} \\ \{F^2\} \\ \{0\} \end{Bmatrix} \tag{68}$$

where

$$K_{ij}^{13}=\int_{\Omega^e} N_{i,x}N_j^\phi\,dx\,dy,\qquad K_{ij}^{23}=\int_{\Omega^e} N_{i,y}N_j^\phi\,dx\,dy$$
$$K_{ij}^{33}=\int_{\Omega^e} N_1^\phi N_j^\phi\,dx\,dy \tag{69}$$

Note that the element equation (68) is of the same form as that associated with the Lagrange multiplier functional (called mixed model) in (21), except for $[K^{33}]$ which is zero in the mixed model.

Equivalence of models I and II

To establish the equivalence between the two models, we eliminate $\{P\}$ from Model II by solving the third equation in (68) for $\{P\}$:

$$\{P\}=-\alpha_n[K^{33}]^{-1}([K^{13}]^T\{u\}+[K^{23}]^T\{v\}) \tag{70}$$

Substituting equation (70) into the first two equations in (68), we obtain

$$\begin{bmatrix} [K^{11}]+\alpha_n[\hat{K}^{11}] & [K^{12}]+\alpha_n[\hat{K}^{12}] \\ \text{symm.} & [K^{22}]+\alpha_n[\hat{K}^{22}] \end{bmatrix} \begin{Bmatrix} \{u\} \\ \{v\} \end{Bmatrix} = \begin{Bmatrix} \{F^1\} \\ \{F^2\} \end{Bmatrix} \tag{71}$$

where

$$[\hat{K}^{11}]=[K^{13}][K^{33}]^{-1}[K^{13}]^T,\qquad [\hat{K}^{12}]=[K^{13}][K^{33}]^{-1}[K^{23}]^T$$
$$[\hat{K}^{22}]=[K^{23}][K^{33}]^{-1}[K^{23}]^T \tag{72}$$

Now comparing equation (66) with equation (71), we must have, in order Models I and II to be the same,

$$[S^{11}]=[\hat{K}^{11}],\qquad [S^{12}]=[\hat{K}^{12}],\qquad [S^{22}]=[\hat{K}^{22}] \tag{73}$$

Now we wish to numerically verify, for a specific element, that the equivalence holds. We shall consider the following two approximations:

Case 1: Bilinear interpolation of the velocity field and constant (discontinuous) approximation of the pressure.

Case 2: Bilinear interpolation of the velocity field and continuous (bilinear) approximation of the pressure.

In each case the effect (and type) of reduced integration on the equivalence will be discussed.

For simplicity the rectangular element with sides a and b is chosen for the study. Table I shows various matrices computed using exact integration and by two different reduced numerical integrations. Clearly, reduced integration using one method does not yield the same result as other methods (since the quadrature errors are different for different methods).

Case 1 (*FEM*-1). Matrices $[K^{13}]$, $[K^{23}]$, and $[K^{33}]$ are given by (reduced as well as exact integrations give the same result),

$$\underset{4\times 1}{\{K^{13}\}} = \frac{-b}{2}\{-1, 1, 1, -1\}^T, \underset{4\times 1}{\{K^{23}\}} = \frac{-a}{2}\{-1, -1, 1, 1\}^T, \underset{1\times 1}{K^{33}} = ab \tag{74}$$

We compute $[\hat{K}^{11}]$, $[\hat{K}^{12}]$, and $[\hat{K}^{22}]$ and find them to be

$$[\hat{K}^{\alpha\beta}] = [S^{\alpha\beta}], \ (\alpha, \beta = 1, 2). \tag{75}$$

That is, Models I and II are equivalent, for Case 1, if the penalty terms are evaluated numerically using one-point Gauss quadrature. However, the equivalence does not hold if the penalty terms are evaluated using the trapezoidal rule.

Case 2. In this case the matrices $[K^{13}]$, $[K^{23}]$, and $[K^{33}]$ can be identified with

$$[K^{13}] = [S^{01}]^T, \qquad [K^{23}] = [S^{02}]^T, \qquad [K^{33}] = [S^{00}] \tag{76}$$

We have the following five alternatives in this case:

(i) If all of the matrix coefficients are evaluated exactly, the resulting matrices satisfy the identities in equation (73); hence, equivalence of Models I and II holds.

(ii) If the one-point Gauss rule is used to evaluate the penalty terms in both models, we note from Table I that $[K^{33}] = [S^{00}]$ is a singular matrix, therefore, Model II cannot be identified with Model I. In other words, pressure cannot be calculated from (70).

(iii) If the trapezoidal rule is used to evaluate the penalty terms in both models, equivalence of Models I and II holds; however, the matrices are different from those obtained in Case 1, as can be seen from Table I.

(iv) If the trapezoidal rule is used to evaluate the Gram matrix, $[S^{00}]$, and the one-point Gauss rule is used to evaluate the penalty terms in both models, equivalence of Models I and II holds. In this case, the matrices are identical to those obtained in Case 1, but differ from those obtained in (i) and (iii). This alternative is equivalent to (ii), except that the pressure is calculated using $[S^{00}]$ which is computed by the trapezoidal rule.

(v) If the Gram matrix is evaluated exactly, and the one-point Gauss rule is used to evaluate the penalty terms in both models, equivalence of Models I and II holds.

Table I. Exact and reduced numerical integration of the matrices in equations (67)

Matrix	Exact integration*	Reduced integration: Trapezoidal rule	Reduced integration: One-point Gauss rule
$[S^{00}]$	$\frac{ab}{36}\begin{bmatrix}4 & 2 & 1 & 2\\ & 4 & 2 & 1\\ \text{symm.} & & 4 & 2\\ & & & 4\end{bmatrix}$	$\frac{ab}{4}\begin{bmatrix}1 & 0 & 0 & 0\\ & 1 & 0 & 0\\ \text{symm.} & & 1 & 0\\ & & & 1\end{bmatrix}$	$\frac{ab}{16}\begin{bmatrix}1 & 1 & 1 & 1\\ & 1 & 1 & 1\\ \text{symm.} & & 1 & 1\\ & & & 1\end{bmatrix}$
$[S^{01}]$	$\frac{b}{12}\begin{bmatrix}-2 & 2 & 1 & -1\\ -2 & 2 & 1 & -1\\ -1 & 1 & 2 & -2\\ -1 & 1 & 2 & -2\end{bmatrix}$	$\frac{b}{4}\begin{bmatrix}-1 & 1 & 0 & 0\\ -1 & 1 & 0 & 0\\ 0 & 0 & 1 & -1\\ 0 & 0 & 1 & -1\end{bmatrix}$	$\frac{b}{8}\begin{bmatrix}-1 & 1 & 1 & -1\\ -1 & 1 & 1 & -1\\ -1 & 1 & 1 & -1\\ -1 & 1 & 1 & -1\end{bmatrix}$
$[S^{02}]$	$\frac{a}{12}\begin{bmatrix}-2 & -1 & 1 & 2\\ -1 & -2 & 2 & 1\\ -1 & -2 & 2 & 1\\ -2 & -1 & 1 & 2\end{bmatrix}$	$\frac{a}{4}\begin{bmatrix}-1 & 0 & 0 & 1\\ 0 & -1 & 1 & 0\\ 0 & -1 & 1 & 0\\ -1 & 0 & 0 & 1\end{bmatrix}$	$\frac{a}{8}\begin{bmatrix}-1 & -1 & 1 & 1\\ -1 & -1 & 1 & 1\\ -1 & -1 & 1 & 1\\ -1 & -1 & 1 & 1\end{bmatrix}$
$[S^{11}]$	$\frac{b}{6a}\begin{bmatrix}2 & -2 & -1 & 1\\ & 2 & 1 & -1\\ \text{symm.} & & 2 & -2\\ & & & 2\end{bmatrix}$	$\frac{b}{2a}\begin{bmatrix}1 & -1 & 0 & 0\\ & 1 & 0 & 0\\ \text{symm.} & & 1 & -1\\ & & & 1\end{bmatrix}$	$\frac{b}{4a}\begin{bmatrix}1 & -1 & -1 & 1\\ & 1 & 1 & -1\\ \text{symm.} & & 1 & -1\\ & & & 1\end{bmatrix}$
$[S^{12}]$	$\frac{1}{4}\begin{bmatrix}1 & 1 & -1 & -1\\ -1 & -1 & 1 & 1\\ -1 & -1 & 1 & 1\\ 1 & 1 & -1 & -1\end{bmatrix}$	$\frac{1}{4}\begin{bmatrix}1 & 1 & -1 & -1\\ -1 & -1 & 1 & 1\\ -1 & -1 & 1 & 1\\ 1 & 1 & -1 & -1\end{bmatrix}$	$\frac{1}{4}\begin{bmatrix}1 & 1 & -1 & -1\\ -1 & -1 & 1 & 1\\ -1 & -1 & 1 & 1\\ 1 & 1 & -1 & -1\end{bmatrix}$
$[S^{22}]$	$\frac{a}{6b}\begin{bmatrix}2 & 1 & -1 & -2\\ & 2 & -2 & -1\\ \text{symm.} & & 2 & 1\\ & & & 2\end{bmatrix}$	$\frac{a}{2b}\begin{bmatrix}1 & 0 & 0 & -1\\ & 1 & -1 & 0\\ \text{symm.} & & 1 & 0\\ & & & 1\end{bmatrix}$	$\frac{a}{46}\begin{bmatrix}1 & 1 & -1 & -1\\ & 1 & -1 & -1\\ \text{symm.} & & 1 & 1\\ & & & 1\end{bmatrix}$

* 2-point Gauss rule and one-third Simpson's rule also give the same result.

Apparently, this case appears to be equivalent to the piecewise constant pressure model of Case 1.

We shall denote the finite element models discussed in (i)–(v), respectively, by FEM-2, FEM-3, . . . , FEM-6.

Pressure calculation

The pressure in a typical element can be calculated in each of the above approximations using equation (70). For Case 1 (model FEM-1), we have,

$$P_e = -\frac{\alpha_n}{2}\left(\frac{1}{a}\{-1, 1, 1, -1\}\begin{Bmatrix}u_1\\ u_2\\ u_3\\ u_4\end{Bmatrix}_e + \frac{1}{b}\{-1, -1, 1, 1\}\begin{Bmatrix}v_1\\ v_2\\ v_3\\ v_4\end{Bmatrix}\right) \tag{77}$$

For (iv) and (v) of Case 2 (models FEM-5 and FEM-6), we have from (70),

$$\{P\}_e = -\frac{\alpha_n}{2}\left(\frac{1}{a}\begin{bmatrix}-1 & 1 & 1 & -1\\ -1 & 1 & 1 & -1\\ -1 & 1 & 1 & -1\\ -1 & 1 & 1 & -1\end{bmatrix}\begin{Bmatrix}u_1\\u_2\\u_3\\u_4\end{Bmatrix}_e + \frac{1}{b}\begin{bmatrix}-1 & -1 & 1 & 1\\ -1 & -1 & 1 & 1\\ -1 & -1 & 1 & 1\\ -1 & -1 & 1 & 1\end{bmatrix}\begin{Bmatrix}v_1\\v_2\\v_3\\v_4\end{Bmatrix}_e\right) \tag{78}$$

Clearly, equations (77) and (78) are equivalent. Further, they yield the same values since the velocities obtained in both of the models are the same.

For (i) and (iii) of Case 2 (models FEM-2 and FEM-4), we have from (70),

$$\{P\}_e = -\alpha_n\left(\frac{1}{a}\begin{bmatrix}-1 & 1 & 0 & 0\\ -1 & 1 & 0 & 0\\ 0 & 0 & 1 & -1\\ 0 & 0 & 1 & -1\end{bmatrix}\begin{Bmatrix}u_1\\u_2\\u_3\\u_4\end{Bmatrix}_e + \frac{1}{b}\begin{bmatrix}-1 & 0 & 0 & 1\\ 0 & -1 & 1 & 0\\ 0 & -1 & 1 & 0\\ -1 & 0 & 0 & 1\end{bmatrix}\begin{Bmatrix}v_1\\v_2\\v_3\\v_4\end{Bmatrix}_e\right) \tag{79}$$

Although the same equation is valid for both of the models, they do not yield the same pressures since the velocities obtained by models FEM-2 and FEM-4 are different, as pointed out in alternative (iii).

Influence of the penalty parameter on the accuracy

To investigate the influence of the penalty parameter α_n on the solution (u_n, v_n, P_n), the problem of natural convection in a square enclosure (in the presence of a constant thermal gradient between the two vertical walls, while the two horizontal walls are insulated; see Figure 1) is solved for $Ra = Pr = 1$ (see Reference 21 for additional information).

Table II shows the velocities and stream function values for $\alpha_n = 10^3$, 10^8, for models FEM-1, FEM-2 and FEM-4. The stream function is computed from the velocity field by

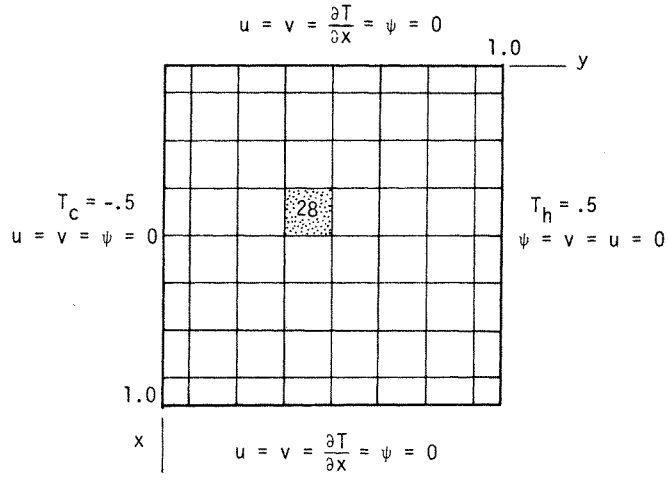

Figure 1. Finite element mesh and boundary conditions

Table II. Comparison of the velocities and stream function as computed by various models

α_n	Variable	x or y	Model FEM-1	Model FEM-2	Model FEM-4
	$u(0{\cdot}5, y)$ $\times 10^2$	0·08	0·20035	$0{\cdot}468\times10^{-6}$	-0.982×10^{-2}
		0·22	0·40263	$0{\cdot}450\times10^{-6}$	$-0{\cdot}319\times10^{-7}$
		0·36	0·24402	$0{\cdot}259\times10^{-6}$	$0{\cdot}982\times10^{-2}$
10^8	$v(x, 0{\cdot}5)$ $\times 10^2$	0·08	−0·29645	$-0{\cdot}440\times10^{-6}$	$0{\cdot}982\times10^{-2}$
		0·22	−0·40379	$-0{\cdot}429\times10^{-6}$	$0{\cdot}791\times10^{-7}$
		0·36	−0·24593	$-0{\cdot}247\times10^{-6}$	$-0{\cdot}982\times10^{-2}$
	$\psi(x, 0{\cdot}5)$ $\times 10^3$	0·08	0·11858	$0{\cdot}188\times10^{-6}$	$-0{\cdot}393\times10^{-2}$
		0·22	0·60875	$0{\cdot}813\times10^{-6}$	$-0{\cdot}108\times10^{-1}$
		0·36	1·0635	$0{\cdot}130\times10^{-6}$	$-0{\cdot}393\times10^{-2}$
		0·50	1·2357	$0{\cdot}147\times10^{-5}$	$0{\cdot}294\times10^{-2}$
	$u(0{\cdot}5, y)$ $\times 10^2$	0·08	0·30105	0·04081	−0·01287
		0·22	0·40387	0·04010	−0·00321
		0·36	0·24490	0·02315	0·01239
10^3	$v(x, 0{\cdot}5)$ $\times 10^2$	0·08	−0·29556	−0·03809	0·01768
		0·22	−0·40243	−0·03797	0·00788
		0·36	−0·24506	−0·02198	−0·01285
	$\psi(x, 0{\cdot}5)$ $\times 10^3$	0·08	0·11937	0·01639	−0·00576
		0·22	0·60965	0·07134	−0·02196
		0·36	1·0640	0·11438	−0·01735
		0·50	1·2359	0·13012	−0·00787

solving the Poisson equation

$$-\nabla^2\psi = \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} \text{ in } \Omega, \qquad \psi = 0 \text{ on } \partial\Omega. \tag{80}$$

Clearly, the solutions obtained by models FEM-2 and FEM-4 are meaningless. In other words, equal interpolation both without reduced integration and with reduced integration by the trapezoidal rule for the penalty terms result in wrong solutions.

Table III shows the influence of the penalty parameter on the velocities, pressure, and stream function as computed by model FEM-1. As can be seen, the 'accuracy' increases with increasing α_n. Also, note that the stream function is relatively less sensitive to the penalty parameter. The solution is unchanged for $\alpha_n = 10^6$, 10^8, 10^{10}. For $\alpha_n > 10^{10}$, the roundoff errors in the computer gradually increased until $\alpha_n = 5\times10^{13}$, for which the coefficient matrix became singular. Figure 2 shows plots of the log of error, $\|\phi_8 - \phi_\alpha\|$ in a typical variable ϕ versus log $\alpha_n = n$. All of the slopes were measured to be unity, thus verifying the theoretical error estimate for the Stokes problem. However, the numerical convergence for $Ra = 10^4$ shows that the theoretical estimate is a conservative one for the Navier–Stokes equations.

Table III. Influence of the Penalty Parameter on the Solution of the Natural Convection Problem

α_n	$u(0{\cdot}5, 0{\cdot}22)\times 10^3$	$-v(0{\cdot}22, 0{\cdot}5)\times 10^3$	$\psi(0{\cdot}5, 0{\cdot}5)\times 10^3$	P (element 28)
1	6·1957	1·5565	1·2824	0·8781
10	4·8646	3·1060	1·2492	0·1552
10^2	4·1449	3·9074	1·2374	0·1233
10^3	4·0387	4·0243	1·2359	0·1070
10^4	4·0276	4·0365	1·2357	0·1050
10^6	4·0264	4·0379	1·2357	0·1048
10^8	4·0263	4·0379	1·2357	0·1048
10^{10}	4·0264	4·0379	1·2357	0·1048
10^{12}	4·0269	4·0387	1·2359	0·1068
10^{13}	3·8141	3·8200	1·1745	0·1523
$0{\cdot}5\times 10^{13}$	Zero appeared on the diagonal			

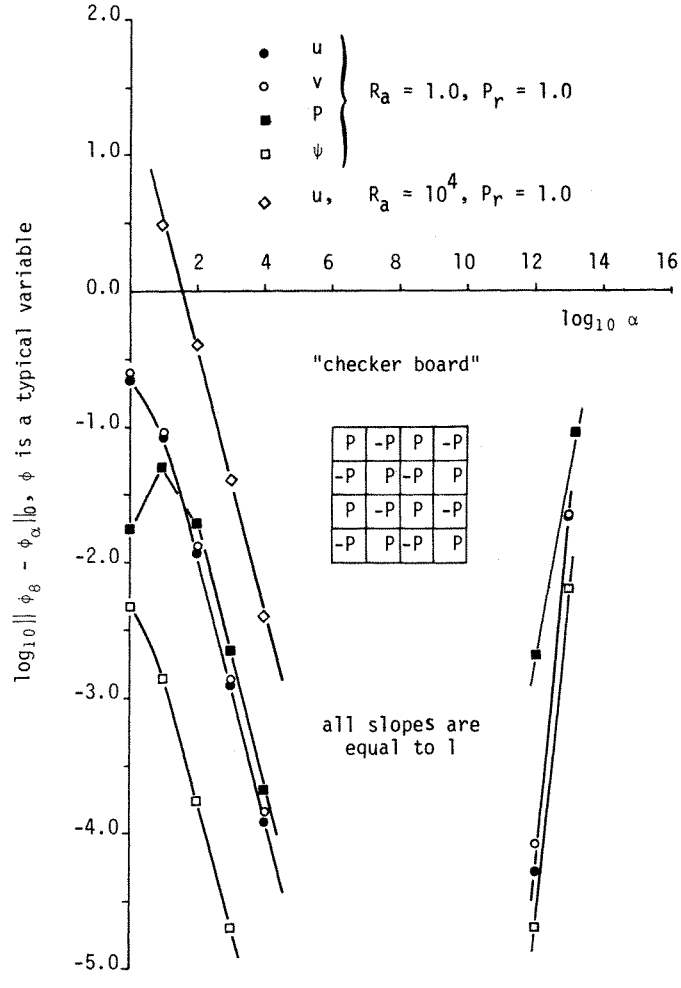

Figure 2. Accuracy (i.e. error estimates) of the penalty-finite element solutions of natural convection in a square cavity. ϕ_n is the solution corresponding to $\alpha = 10^n$

SUMMARY AND CONCLUSIONS

The idea of the penalty function method is reviewed as a variational technique that transforms a constrained minimization problem into a problem of unconstrained minimization(s). Existence and convergence of the solution to the penalty problem are also reviewed. Application of the penalty function method to Stokes problem with the divergence free condition as a constraint is discussed, and existence and uniqueness of solutions to the penalty function formulation of the Navier–Stokes equations are proved. Equivalence of conventional and mixed penalty-finite element models of fluid flow is established for a particular element, and theoretical error estimates for the Stokes flow are verified numerically. It is found from the numerical studies that the error estimate for Navier–Stokes equations is the same order as that in the Stokes problem. Thus, the theoretical error estimate in (56) is not optimal.

The finite element model that employs bilinear interpolation for the velocities and constant interpolation for the pressure is equivalent to the model that employs bilinear interpolation for all of the variables but uses the reduced Gauss rule for the numerical integration of the penalty terms. Further, use of the trapezoidal rule for numerical evaluation of the penalty terms is found to result in erroneous results. Thus, the type of numerical quadrature is crucial (for the penalty terms) for the success of the penalty method for incompressible fluid flow.

In all of the models discussed herein, it was discovered that a 'chequer board' pattern of the pressure was present. Model FEM-1 (equivalent to models FEM-3, FEM-5, and FEM-6, except for the pressure calculation) is found to give most reliable results for the velocities and pressure. It is computationally economical to use Model I over Model II, and then use a desirable approximation scheme[24] to determine the pressure. Further study must be carried out to determine the effect of the type of reduced integration (e.g., the three-point Gauss rule[35]) on the solution. For a penalty-finite element analysis of three-dimensional flows, see a recent study by the author.[40]

ACKNOWLEDGMENTS

Support of this work by the Global Atmospheric Research Program of the U.S. National Science Foundation through Grant ATM77-23111 is gratefully acknowledged. Thanks are also due to Dr. Philip Gresho of Lawrence Livermore National Laboratories for his constructive comments on the paper. The author is grateful to his teacher, Professor John Tinsley Oden, to whom this paper is dedicated.

REFERENCES

1. R. Courant, 'Variational methods for the solution of problems of equilibrium and vibrations', *Bulletin of the American Mathematical Society*, **49,** 1–23 (1943).
2. R. Courant, *Calculus of Variations and Supplementary Notes and Exercises* (Mimeographed Notes), Supplementary Notes by M. Kruskal and H. Rubin, revised and amended by J. Moser, New York University, 1956–57.
3. G. Leitmann, *Optimization Techniques: With Applications to Aerospace Systems*, Academic Press, New York, 1962.
4. K. R. Frisch, 'Principles of linear programming—with particular reference to the double gradient form of the logarithmic potential method', *Memorandum of October* 18, 1954, University Institute of Economics, Oslo.
5. C. W. Carroll, 'The created response surface technique for optimizing nonlinear restrained systems', *Operations Research*, **9,** (2), 169–184 (1961).
6. H. Rubin and P. Unger, 'Motion under a strong constraining force', *Commun. Pure and Applied Mathematics*, **10,** 65–87 (1957).
7. A. V. Fiacco and G. P. McCormick, *Nonlinear Programming: Sequential Unconstrained Methods for Solving Constrained Minimization Techniques*, Wiley, New York, 1968.

8. M. R. Hestenes, *Optimization Theory: The Finite Dimensional Case*, Wiley-Interscience, New York, 1975.
9. Y. K. Sasaki, 'Variational design of finite–difference schemes for initial-value problems with an integral invariant', *Journal of Computational Physics*, **21,** (3), 270–278 (1976).
10. I. Babuska, 'The finite element method with penalty', *Tech. Note BN-710*, The Institute for Fluid Dynamics and Applied Mathematics, University of Maryland, August 1971.
11. O. C. Zienkiewicz, 'Constrained variational principles and penalty function methods in finite element analysis,' in *Lecture Notes in Mathematics: Conference on the Numerical Solution of Differential Equations*, edited by G. A. Watson, Springer-Verlag, Berlin, 207–214, 1974.
12. O. C. Zienkiewicz, R. L. Taylor and J. M. Too, 'Reduced integration technique in general analysis of plates and shells', *International Journal of Numerical Methods in Engineering*, **3,** 575–586, 1971.
13. O. C. Zienkiewicz and P. N. Godbole, 'Viscous, incompressible flow with special reference to non-Newtonian (plastic) fluids' in *Finite elements in Fluids, Vol. 1*, R. H. Gallagher, J. T. Oden, C. Taylor and O. C. Zienkiewicz (eds.) Wiley-Interscience, London, pp. 25–55, 1975.
14. O. C. Zienkiewicz and E. Hinton, 'Reduced integration, function smoothing and non-conformity in finite element analysis', *Journal of the Franklin Institute*, **302,** 443–461 (1976).
15. T. J. R. Hughes, R. L. Taylor and J. F. Levy, 'High Reynolds number steady, incompressible flow by a finite element method' in *Finite Elements in Fluids, Vol. 3*, R. H. Gallagher, J. T. Oden, C. Taylor and O. C. Zienkiewicz (eds.) Wiley-Interscience, London, pp. 55–72, 1979.
16. J. N. Reddy and K. H. Patil, 'Alternate finite element formulations of incompressible fluid flow with applications to geological folding' in *Applications of Computer Methods in Engineering, Vol. 1*, L. C. Wellford, Jr. (ed.), University of Southern California, Los Angeles, pp. 179–190, 1977.
17. R. S. Marshall, J. C. Heinrich and O. C. Zienkiewicz, 'Natural convection in a square enclosure by a finite-element penalty function method using primitive fluid variables', *Numerical Heat Transfer*, **1,** 315–330 (1978).
18. J. N. Reddy, 'Penalty finite element methods for the solution of advection and free convection flows' in *Finite Element Methods in Engineering (Proc. Third Int. Conf. in Australia on Finite Element Methods)*, A. P. Kabaila and V. A. Pulmano (eds.), the University of New South Wales, Sydney, 583–598 (1979).
19. T. J. R. Hughes, W. K. Liu and A. Brooks, 'Finite element analysis of incompressible viscous flows by the penalty function formulation', *Journal of Computational Physics*, **30,** 1–60 (1979).
20. J. T. Oden, N. Kikuchi and Y. J. Song, 'An analysis of exterior penalty methods and reduced integration for finite element approximations of contact problems in incompressible elasticity', *TICOM Report 79–10*, Texas Institute for Computational Mechanics, The University of Texas, Austin, 1979.
21. J. N. Reddy and A. Satake, 'A comparison of various finite-element models of natural convection in enclosures', *Journal of Heat Transfer*, **102,** 659–666 (1980).
22. T. J. Chung and G. R. Karr, 'Significance of convective Term in transport equations', *Finite Element Methods for Convection Dominated Flows*, AMD-Vol. 34, The American Society of Mechanical Engineers, New York, 137–147 (1979).
23. D. S. Malkus, 'Finite elements with penalties in nonlinear elasticity', *International Journal for Numerical Methods in Engineering*, **16,** 121–136 (1980).
24. R. L. Sani, P. M. Gresho, R. L. Lee, D. F. Griffiths and M. Engelman, 'The cause and cure (?) of the spurious pressures generated by certain FEM solutions of the incompressible Navier–Stokes equations: Parts 1 and 2', *International Journal for Numerical Methods in Fluids*, **1,** 17–43 and 171–204 (1981).
25. B. T. Polya, 'The convergence rate of the penalty function method', *Zh. vychisl. Mat. mat. fiz.*, **11,** 3–11 (1971); English translation: *U.S.S.R. Computational Mathematics and Mathematical Physics*, **11,** 1–12 (1971).
26. J. T. Oden, 'A theory of penalty methods for finite element approximations of Highly Nonlinear Problems in Continuum Mechanics', *Computers and Structures*, **8,** 445–449 (1978).
27. N. Kikuchi, 'Convergence of a penalty method for variational inequalities', *TICOM Report 79–16*, October 1979, The Texas Institute for Computational Mechanics, The University of Texas at Austin, Austin Texas.
28. J. N. Reddy, 'On the accuracy and existence of solutions to primitive variable models of viscous incompressible fluids', *International Journal of Engineering Science*, **16,** 921–929 (1978).
29. J. N. Reddy, 'On the finite element method with penalty for incompressible fluid flow problems', *The Mathematics of Finite Elements and Applications III*, J. R. Whiteman (ed.) Academic Press, London, pp. 227–235, 1979.
30. V. Girault and P. A. Raviart, 'An analysis of a mixed finite element method for the Navier–Stokes equations', *Analysis Numerique*, Université Pierre et Marie Curie, Paris, 1978.
31. J. T. Oden and J. N. Reddy, *An Introduction to the Mathematical Theory of Finite Elements*, Wiley-Interscience, New York, 1976.
32. J. N. Reddy, 'On the mathematical theory of penalty-finite elements for Navier–Stokes equations', *Third International Conference on Finite Elements in Flow Problems*, Banff Center, Banff, Alberta, June 10–13, 1980.
33. F. Brezzi, 'On the existence, uniqueness, and approximation of saddle-point problems arising from Lagrange multipliers', *R.A.I.R.O.*, **8,** (R-2) 129–151 (1974).
34. M. Bercovier, 'Perturbations of mixed variational problems, applications to mixed finite element methods', *R.A.I.R.O., Analyze Numerique/Numerical Analysis*, **12,** 211–236 (1978).

35. J. T. Oden, 'Penalty methods and selective reduced integration for Stokesian flows', *Third International Conference on Finite Elements in Flow Problems*, Banff Center, Banff, Alberta, June 10–13, 1980.
36. B. Mercier, 'Topics in finite element solution of elliptic problems', *TIFR Lecture Notes*, Tata Institute of Fundamental Research, Bombay, India, 1979.
37. J. T. Oden and J. N. Reddy, 'Some observations on properties of certain mixed finite element approximations', *International Journal for Numerical Methods in Engineering*, **9**, 933–938 (1975).
38. D. S. Malkus, and T. J. R. Hughes, 'Mixed finite element methods—reduced and selective integration techniques: a unification of concepts', *Computer Methods in Applied Mechanics and Engineering*, **15**, (1), 63–81 (1978).
39. N. Kikuchi and Y. J. Song, 'Remarks on relations between penalty and mixed finite element methods for a class of variational inequalities', *International Journal for Numerical Methods in Engineering*, **15**, 1557–1561 (1980).
40. J. N. Reddy, 'A finite-element analysis of steady incompressible flows in three dimensions by a penalty function formulation', *Research Report No.* VPI-E-81-14, Department of Engineering Science and Mechanics, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061.