

## CONTINUATION-CONJUGATE GRADIENT METHODS FOR THE LEAST SQUARES SOLUTION OF NONLINEAR BOUNDARY VALUE PROBLEMS\*

R. GLOWINSKI†, H. B. KELLER‡ AND L. REINHART§

**Abstract.** We discuss in this paper a new combination of methods for solving nonlinear boundary value problems containing a parameter. Methods of the continuation type are combined with least squares formulations, preconditioned conjugate gradient algorithms and finite element approximations.

We can compute branches of solutions with limit points, bifurcation points, etc.

Several numerical tests illustrate the possibilities of the methods discussed in the present paper; these include the Bratu problem in one and two dimensions, one-dimensional bifurcation and perturbed bifurcation problems, the driven cavity problem for the Navier–Stokes equations.

**Key words.** nonlinear boundary value problems, bifurcation, continuation methods, nonlinear least squares, conjugate gradient, finite elements, Navier–Stokes equations, biharmonic solvers

**1. Introduction.** We present in this paper a powerful combination of techniques that are used to solve a variety of nonlinear boundary value problems containing a parameter. Indeed the resulting method can be employed to study a large class of nonlinear eigenvalue problems. The individual techniques include: arclength or pseudo-arclength continuation, least squares formulation in an appropriate Hilbert space setting, a conjugate gradient iterative method for solving the least squares problem and finite element approximations to yield a finite dimensional problem for computation.

In § 2 the solution techniques are described in some detail. Specifically in § 2.1 the last squares formulation of a broad class of nonlinear problems, say in the form

$$(1.1) \quad AU = T(U),$$

are formulated in an appropriate Hilbert space setting. Then in § 2.2 a conjugate gradient iterative solution technique for solving such least squares problems is presented. In § 2.3 a pseudo-arc length continuation method for nonlinear eigenvalue problems in the form

$$(1.2) \quad Lu = G(u, \lambda)$$

is discussed. This involves adjoining a scalar linear constraint, say

$$(1.3) \quad l(u, \lambda, s) = 0,$$

and with  $U \equiv \{u, \lambda\}$  the previous least squares and conjugate gradient techniques can be applied to the system (1.2), (1.3). One big advantage of our specific continuation method is that simple limit or fold points of the original problem (1.2) are just regular points for our reformulation in the form (1.1). The entire procedure thus enables us to determine large arcs of branches of solutions of (1.2) with no special precautions or change of methods near limit points.

---

\* Received by the editors November 22, 1982, and in revised form June 12, 1984.

† Paris VI University, LA 189, Tour 55.65, 75230 Paris Cedex 05 France, and INRIA, Domaine de Voluceau, Rocquencourt, 78153 Le Chesnay Cedex, France.

‡ Applied Mathematics, California Institute of Technology, Pasadena, California 91125. The research of this author was supported in part by the U.S. Department of Energy under contract EY-76-S-03-0767, Project Agreement 12, and by the Army Research Office under contract DAAG 29-78-C-0011.

§ INRIA, Domaine de Voluceau, Rocquencourt, 78153 Le Chesnay Cedex, France.

These techniques, as described in § 2, apply to the analytical problem. However they go over extremely well when various discrete approximations are applied to yield computational methods of great power and practicality. We illustrate this by considering several nonlinear boundary value problems of some difficulty and current interest. In each of these problems the discretization is obtained by some finite element formulation. The well-known Bratu problem on a square domain is treated in § 3. Several ordinary differential equation examples displaying bifurcation and the effects of perturbed bifurcation are treated in § 4. We show how to use perturbed bifurcation and continuation to obtain the bifurcating solutions. Finally in § 5 the Navier–Stokes equations are solved for the driven cavity problem.

Actually the techniques described in this paper have also been applied to the solution of nonlinear boundary value problems, more complicated than those considered in the following sections. Among these problems, we shall mention the *Von Karman equations for nonlinear plates* and the computation of the multiple solutions of the *full potential equation* modelling *transonic flows for compressible inviscid fluids*.

**2. Solution techniques.** We introduce in this section the methods we shall apply in §§ 3, 4, 5, to the solution of quite general nonlinear boundary value problems. They include *least squares, conjugate gradient and arc length continuation* methods.

Let  $V$  be a *Hilbert space* (real for simplicity) equipped with the scalar product  $(\cdot, \cdot)$  and the corresponding norm  $\|\cdot\|$ . We denote by  $V'$  the *dual space* of  $V$ , by  $\langle \cdot, \cdot \rangle$  the *duality pairing* between  $V$  and  $V'$ , and by  $\|\cdot\|_*$  the corresponding *dual norm*, i.e.

$$(2.1) \quad \|f\|_* = \sup_{v \in V - \{0\}} \frac{|\langle f, v \rangle|}{\|v\|} \quad \forall f \in V'.$$

The problem that we consider is to find  $u \in V$  such that

$$(2.2) \quad S(u) = 0,$$

where  $S$  is a nonlinear operator from  $V$  to  $V'$ .

**2.1. Least squares formulation of problem (2.2).** A least squares formulation of (2.2) is obtained by observing that any solution of (2.2) is also a *global minimizer* over  $V$  of the functional  $J: V \rightarrow \mathbb{R}$  defined by

$$(2.3) \quad J(v) = \frac{1}{2} \|S(v)\|_*^2.$$

Hence a least squares formulation of (2.2) is:

*Find  $u \in V$  such that*

$$(2.4) \quad J(u) \leq J(v) \quad \forall v \in V.$$

In practice we proceed as follows. Let  $A$  be the *duality isomorphism* corresponding to  $(\cdot, \cdot)$  and  $\langle \cdot, \cdot \rangle$ . That is  $\forall v \in V, Av \in V'$  satisfies

$$(2.5) \quad \langle Av, w \rangle = (v, w) \quad \forall w \in V,$$

$$(2.6) \quad \|v\| = \|Av\|_* \quad (\text{or equivalently } \|f\|_* = \|A^{-1}f\|, \forall f \in V').$$

It follows that

$$(2.7) \quad J(v) = \frac{1}{2} \langle A\xi, \xi \rangle \quad (= \frac{1}{2} \|\xi\|^2),$$

where  $\xi$  is a (nonlinear) function of  $v$  obtained via the solution of the *well-posed linear problem*

$$(2.8) \quad A\xi = S(v).$$

We observe that (2.4) has the structure of an *optimal control* problem, where

- (i)  $v$  is the *control vector*,
- (ii)  $\xi$  is the *state vector*,
- (iii) (2.8) is the *state equation*,
- (iv)  $J$  is the *cost function*.

As a final remark we observe that any solution of the minimization problem (2.4) for which  $J$  vanishes is also a solution of the original problem (2.2).

**2.2. Solution by a conjugate gradient algorithm.** We suppose from now on that  $S$  is differentiable implying in turn the differentiability of  $J$  over  $V$ . We denote by  $S'$  and  $J'$  the Fréchet derivatives of  $S$  and  $J$  respectively.

From the differentiability of  $J$  it is quite natural to solve the minimization problem (2.4) by a conjugate gradient algorithm; among the possible conjugate gradient algorithms we have selected the Polak–Ribière variant (cf. Polak [1]) whose very good performance has been discussed by Powell [2] (see also Shanno [28]). The Polak–Ribière method applied to the solution of (2.4) provides the following algorithm.

*Step 0: Initialization.* For some given

$$(2.9) \quad u^0 \in V,$$

compute  $g^0 \in V$  as the solution of

$$(2.10) \quad Ag^0 = J'(u^0),$$

and set

$$(2.11) \quad z^0 = g^0.$$

Then, for  $n \geq 0$ , with  $u^n, g^n, z^n$  known, compute  $u^{n+1}, g^{n+1}, z^{n+1}$  by:

*Step 1: Descent.* Compute:

$$(2.12) \quad \rho_n = \arg \min_{\rho \in \mathbb{R}} J(u^n - \rho z^n),$$

and then set

$$(2.13) \quad u^{n+1} = u^n - \rho_n z^n.$$

*Step 2: New descent direction.* Define  $g^{n+1} \in V$  as the solution of

$$(2.14) \quad Ag^{n+1} = J'(u^{n+1});$$

then compute

$$(2.15) \quad \gamma_n = \frac{\langle A(g^{n+1} - g^n), g^{n+1} \rangle}{\langle Ag^n, g^n \rangle} \left( = \frac{\langle g^{n+1} - g^n, g^{n+1} \rangle}{\langle g^n, g^n \rangle} \right)$$

and set

$$(2.16) \quad z^{n+1} = g^{n+1} + \gamma_n z^n.$$

Set  $n = n + 1$  and return to Step 1.

The two costly steps (because they need some auxiliary computations) of algorithm (2.9)–(2.16) are:

- (i) The solution of the one-dimensional minimization problem (2.12) to obtain  $\rho_n$ . We have done the corresponding line search by *dichotomy* and *parabolic interpolation*, using  $\rho_{n-1}$  as starting value<sup>1</sup> (see [3] for more details). We recall that each evaluation

<sup>1</sup> If the nonlinearity is *polynomial* we can use faster methods.

of  $J(v)$ , for a given argument  $v$ , requires the solution of the linear problem (2.8) to obtain the corresponding  $\xi$ .

(ii) The calculation of  $g^{n+1}$  from  $u^{n+1}$  which requires the solution of two linear problems associated with  $A$  (namely (2.8) with  $v = u^{n+1}$  and (2.14)).

*Calculation of  $J'(u^n)$  and  $g^n$ :* Owing to the importance of Step 2, let us detail the calculation of  $J'(u^n)$  and  $g^n$ .

Let  $v \in V$ ; then  $J'(v)$  may be defined by

$$(2.17) \quad \langle J'(v), w \rangle = \lim_{\substack{t \rightarrow 0 \\ t \neq 0}} \frac{J(v + tw) - J(v)}{t} \quad \forall w \in V.$$

We obtain from (2.7), (2.8), (2.17) that

$$(2.18) \quad \langle J'(v), w \rangle = \langle A\xi, \eta \rangle$$

where  $\xi$  and  $\eta$  are the solutions of (2.8) and

$$(2.19) \quad A\eta = S'(v) \cdot w,$$

respectively. Since  $A$  is *self-adjoint* (from (2.5)) we also have from (2.18), (2.19) that

$$(2.20) \quad \langle J'(v), w \rangle = \langle A\xi, \eta \rangle = \langle A\eta, \xi \rangle = \langle S'(v) \cdot w, \xi \rangle.$$

Therefore  $J'(v) \in V'$  may be identified with the *linear functional*

$$(2.21) \quad w \rightarrow \langle S'(v) \cdot w, \xi \rangle.$$

It follows then from (2.14), (2.20), (2.21) that  $g^n$  is the solution of the following *linear variational problem*:

$$(2.22) \quad \text{Find } g^n \in V \text{ such that} \\ \langle Ag^n, w \rangle = \langle S'(u^n) \cdot w, \xi^n \rangle \quad \forall w \in V,$$

where  $\xi^n$  is the solution of (2.8) corresponding to  $v = u^n$ .

*Remark 2.1.* It is clear from the above observations that an efficient solver for linear problems related to operator  $A$  (in fact to a *finite-dimensional* approximation of  $A$ ) will be a fundamental tool for the solution of problem (2.2) by the conjugate gradient algorithm (2.9)–(2.16).

*Remark 2.2.* The fact that  $J'(v)$  is known through (2.20) is not at all a drawback if a *Galerkin* or a *finite element* method is used to approximate (2.2). Indeed we only need to know the value of  $\langle J'(v), w \rangle$  for  $w$  belonging to a *basis* of the *finite-dimensional subspace* of  $V$  corresponding to the Galerkin or finite element approximation under consideration.

*Convergence of algorithm (2.9)–(2.16):* We introduce the concept of *regular solution* of problem (2.2) by the following definition.

DEFINITION 2.1. A solution  $u$  of (2.2) is said to be regular if the operator  $S'(u)$  ( $\in \mathcal{L}(V, V')$ ) is an isomorphism from  $V$  onto  $V'$ .

Using a modification of the finite-dimensional techniques of Polak [1], it has been proved in Reinhart [3] that if problem (2.2) has a *finite number* of solutions and if these solutions are *regular*, then the conjugate gradient algorithm (2.9)–(2.16) converges to a solution of (2.2), depending upon the initial iterate  $u^0$  in (2.9). This convergence result requires that  $u^0$  is well chosen, as in Newton's method. Hence the role that continuation methods may play is apparent.

**2.3. Arc length continuation methods.** Consider now the solution of nonlinear problems depending upon a real parameter  $\lambda$ ; we would like to follow in the space  $V \times \mathbb{R}$  branches of solutions  $\{u(\lambda), \lambda\}$  when  $\lambda$  belongs to a compact interval of  $\mathbb{R}$ .

These nonlinear eigenvalue problems can be written as follows

$$(2.23) \quad S(u, \lambda) = 0, \quad \lambda \in \mathbb{R}, \quad u \in V.$$

Equation (2.23) reduces quite often to

$$(2.24) \quad Lu = G(u, \lambda), \quad \lambda \in \mathbb{R}, \quad u \in V,$$

where  $L: V \rightarrow V'$  is a linear elliptic operator, and where  $G$  is a *nonlinear Fredholm operator* (see e.g. Berger [4] for the definition of Fredholm operators).

A classical approach is to use  $\lambda$  as the parameter defining arcs of solutions. If for  $\lambda = \lambda_0$  problem (2.23) has a unique solution  $u = u_0$  and if that solution is *isolated*, that is

$$(2.25) \quad S_u^0 = \frac{\partial S}{\partial u}(u_0, \lambda_0) \text{ is an isomorphism from } V \text{ onto } V',$$

and if  $\{u, \lambda\} \rightarrow S(u, \lambda)$  is  $C^1$  in some ball around  $\{u_0, \lambda_0\}$ , then the *implicit function theorem* implies the existence of a smooth arc of regular solutions  $u = u(\lambda)$  for  $|\lambda - \lambda_0| < \rho$ . Therefore, for  $\lambda$  given *sufficiently close to*  $\lambda_0$  we may solve problem (2.23) just as problem (2.2). These procedures, however, may fail or encounter difficulties (slow convergence for example) close to a *nonisolated* solution.

To overcome these difficulties we replace problem (2.23) by the following system

$$(2.26) \quad S(u, \lambda) = 0,$$

$$(2.27) \quad I(u, \lambda, s) = 0,$$

where  $I: V \times \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$  is chosen such that  $s$  is some *arc length* (or a convenient *approximation* to it) on the solution branch. We look then for a solution  $\{u(s), \lambda(s)\}$ ,  $s$  being given (but not  $\lambda$ ). If in addition to  $\{u_0, \lambda_0\}$  we know a tangent vector to the path  $\{\dot{u}(s_0), \dot{\lambda}(s_0)\}$  (where  $\dot{v}$  denotes the derivative of  $v$  with respect to  $s$ ) satisfying:

$$(2.28a) \quad S_u(u_0, \lambda_0)\dot{u}(s_0) + S_\lambda(u_0, \lambda_0)\dot{\lambda}(s_0) = 0,$$

$$(2.28b) \quad \|\dot{u}(s_0)\|^2 + |\dot{\lambda}(s_0)|^2 = 1,$$

then we can use

$$(2.29) \quad I(u, \lambda, s) \equiv (\dot{u}(s_0), u(s) - u(s_0)) + \dot{\lambda}(s_0)(\lambda(s) - \lambda(s_0)) - (s - s_0) = 0,$$

for  $|s - s_0|$  sufficiently small.

Let us define  $U \in V \times \mathbb{R}$  by  $U = \{u, \lambda\}$ ; then problem (2.26), (2.27) can be written as

$$(2.30) \quad T_s(U) = 0,$$

where

$$(2.31) \quad T_s(U) = \begin{pmatrix} T_{1s}(U) \\ T_{2s}(U) \end{pmatrix}$$

with

$$(2.32) \quad T_{1s}(U) = S(u, \lambda), \quad T_{2s}(U) = I(u, \lambda, s).$$

The main interest of this new formulation is that the ordinary *limit points* of (2.26) become *regular* solutions of (2.30) (see H. B. Keller [5], [6] for more details).

Using the notation of § 2.1 a least squares formulation of (2.30) (generalizing (2.4)) is given by:

$$(2.33) \quad \text{Find } U(s) = \{u(s), \lambda(s)\} \text{ such that} \\ J_s(U(s)) \leq J_s(W) \quad \forall W = \{w, \mu\} \in V \times \mathbb{R},$$

with

$$(2.34) \quad J_s(W) = \frac{1}{2} \langle A\tilde{w}, \tilde{w} \rangle + \frac{1}{2} |\tilde{\mu}|^2$$

where, in (2.34),  $\tilde{w}$  and  $\tilde{\mu}$  are (nonlinear) functions of  $\{w, \mu\}$  via the solution of the linear problems

$$(2.35) \quad A\tilde{w} = T_{1s}(w, \mu),$$

$$(2.36) \quad \tilde{\mu} = T_{2s}(w, \mu),$$

respectively.

We consider now a conjugate gradient algorithm (in fact, a variation of algorithm (2.9)–(2.16)) to solve the least squares problem (2.33); this algorithm is defined as follows.

*Step 0: Initialization.* For some given

$$(2.37) \quad U^0 = \{u^0, \lambda^0\}$$

compute  $G^0 = \{g_u^0, g_\lambda^0\} \in V \times \mathbb{R}$  as the solution of

$$(2.38) \quad Ag_u^0 = \frac{\partial J_s}{\partial u}(U^0),$$

$$(2.39) \quad g_\lambda^0 = \frac{\partial J_s}{\partial \lambda}(U^0),$$

and set

$$(2.40) \quad Z^0 = G^0.$$

Then for  $n \geq 0$ , with  $U^n, G^n, Z^n$  known, compute  $U^{n+1}, G^{n+1}, Z^{n+1}$  as follows.

*Step 1: Descent.* Compute

$$(2.41) \quad \rho_n = \text{Arg Min}_{\rho \in \mathbb{R}} J_s(U^n - \rho Z^n),$$

$$(2.42) \quad U^{n+1} = U^n - \rho_n Z^n$$

(i.e.  $u^{n+1} = u^n - \rho_n z_u^n, \lambda^{n+1} = \lambda^n - \rho_n z_\lambda^n$ ).

*Step 2: Calculation of the new descent direction.* Define  $G^{n+1} = \{g_u^{n+1}, g_\lambda^{n+1}\} \in V \times \mathbb{R}$  as the solution of

$$(2.43) \quad Ag_u^{n+1} = \frac{\partial J_s}{\partial u}(U^{n+1}),$$

$$(2.44) \quad g_\lambda^{n+1} = \frac{\partial J_s}{\partial \lambda}(U^{n+1}),$$

then compute

$$\begin{aligned}
 \gamma_n &= \frac{\langle A(g_u^{n+1} - g_u^n), g_u^{n+1} \rangle + (g_\lambda^{n+1} - g_\lambda^n)g_\lambda^{n+1}}{\langle Ag_u^n, g_u^n \rangle + |g_\lambda^n|^2} \\
 (2.45) \qquad &= \frac{(g_u^{n+1} - g_u^n, g_u^{n+1}) + (g_\lambda^{n+1} - g_\lambda^n)g_\lambda^{n+1}}{\|g_u^n\|^2 + |g_\lambda^n|^2},
 \end{aligned}$$

and set

$$(2.46) \qquad Z^{n+1} = G^{n+1} + \gamma_n Z^n.$$

Set  $n \leftarrow n + 1$  and return to Step 1.

The various comments given in § 2.2, concerning algorithm (2.9)–(2.16), still hold for algorithm (2.37)–(2.46). In particular as we pointed out in Remark 2.1 is the importance of efficient methods for solving linear problems related to operator  $A$ . This remark still holds, indeed, since in the context of § 2.3,  $A$  is replaced by the *block-diagonal* isomorphism  $\mathcal{A}: V \times \mathbb{R} \rightarrow V' \times \mathbb{R}$  defined by

$$\mathcal{A} = \begin{pmatrix} A & 0 \\ 0 & 1 \end{pmatrix}.$$

We do not go into the details of the calculation of  $\partial J_s / \partial u$ ,  $\partial J_s / \partial \lambda$  since it is just a trivial modification of the calculation done in § 2.2 to obtain  $J'$ .

A crucial step in the continuation method is a “good” *initialization* choice in (2.37). The obvious choice

$$(2.47) \qquad \{u^0, \lambda^0\} = \{u(s_0), \lambda(s_0)\}$$

is naive and a better choice is provided by using the tangent of (2.28) in the *extrapolation*:

$$\begin{aligned}
 (2.48) \qquad u^0 &= u(s_0) + (s - s_0)\dot{u}(s_0), \\
 \lambda^0 &= \lambda(s_0) + (s - s_0)\dot{\lambda}(s_0).
 \end{aligned}$$

This results in much faster convergence, especially close to the limit points. This initialization technique leads to the so-called *continuation method with incremental load* and is of order 2 (see Deuffhard [38]). We shall return to this initialization problem in § 2.4.

*Convergence of algorithm (2.37)–(2.46).* The fundamental advantage of the arc length continuation approach is that it provides an efficient solution method in the neighborhood of the so-called *limit* (or *fold*) *points* of problem (2.23). A precise formulation of the concept of simple limit points is given by the following definition.

DEFINITION 2.2. Let  $\{u_0, \lambda_0\} \in V \times \mathbb{R}$  be the solution of problem (2.23). We say that  $\{u_0, \lambda_0\}$  is a *simple limit point* if:

$$(2.49) \qquad \dim \text{Ker} \left( \frac{\partial S}{\partial u}(u_0, \lambda_0) \right) = \text{codim Range} \left( \frac{\partial S}{\partial u}(u_0, \lambda_0) \right) = 1,$$

$$(2.50) \qquad \frac{\partial S}{\partial \lambda}(u_0, \lambda_0) \notin \text{Range} \left( \frac{\partial S}{\partial u}(u_0, \lambda_0) \right).$$

We show in Fig. 3.2 an example of such a limit point (located on a solution curve of  $S(u, \lambda) = 0$  where  $\lambda_0 = \lambda_{CR}$ ).

The main justification of arc length continuation methods follows from the next proposition.

PROPOSITION 2.1. *Any simple limit point solution of problem (2.23) is a regular solution of problem (2.30).*

For a proof, see Keller [39], Decker-Keller [40].

From a practical point of view, Proposition 2.1 is of fundamental importance for the following reasons.

(i) Since simple limit points for problem (2.23) are regular points for problem (2.30), the conjugate gradient algorithm can be used to compute these limit points via the least squares formulation (2.33). This property is a direct consequence of the convergence properties of the conjugate gradient algorithm mentioned in § 2.2 and discussed in details in [3].

(ii) Using a perturbation technique, bifurcation points can be approximated by simple limit points. Then the solution methods described in the present § 2 can be applied; several examples of such situations will be discussed in § 4.

**2.4. Implementation of the arc length continuation method.** To help potential users of continuation methods discussed in the previous sections, we summarize here the essentials of these methods. We solve the nonlinear problem (2.23) via the solution of a family (parametrized by  $s$ ) of nonlinear system (2.26), (2.27). In practice we approximate (2.26), (2.27) by the discrete family of nonlinear systems described below, where  $\Delta s$  is an *arc length step, positive or negative* (possibly varying with  $n$ ) and where  $u^n \approx u(n \Delta s)$ ,  $\lambda^n \approx \lambda(n \Delta s)$ :

*Initialization.* We suppose that we know a solution  $\{u^0, \lambda^0\}$  of (2.23); we take it as origin of the arc of solutions, i.e.  $u^0 = u(0)$ ,  $\lambda^0 = \lambda(0)$ . We suppose also that we know the tangent  $(\dot{u}(0), \dot{\lambda}(0))$  satisfying (2.28a, b) (or at least an approximation to it; see Remark 2.3).

*Continuation.* Then for  $n \geq 0$ , with  $u^n, \lambda^n$  known and also  $u^{n-1}, \lambda^{n-1}$  (resp.  $u(0), \lambda(0)$  if  $n = 0$ ), we obtain  $\{u^{n+1}, \lambda^{n+1}\} \in V \times \mathbb{R}$  as the solution of:

$$(2.51) \quad S(u^{n+1}, \lambda^{n+1}) = 0,$$

and

$$(2.52) \quad (u^1 - u^0, \dot{u}(0)) + (\lambda^1 - \lambda^0)\dot{\lambda}(0) = \Delta s \quad \text{if } n = 0,$$

$$(2.53) \quad \left( u^{n+1} - u^n, \frac{u^n - u^{n-1}}{\Delta s} \right) + (\lambda^{n+1} - \lambda^n) \left( \frac{\lambda^n - \lambda^{n-1}}{\Delta s} \right) = \Delta s \quad \text{if } n \geq 1.$$

*Remark 2.3.* It may occur that obtaining  $\{\dot{u}(0), \dot{\lambda}(0)\}$  is by itself a complicated problem; however obtaining a second solution of (2.23), close to  $\{u^0, \lambda^0\}$ , may be easy (using the nonlinear least squares-conjugate gradient methods of §§ 2.1, 2.2, for example). This supposes that we are sufficiently far from a singular point. Let us denote this second solution by  $\{u^{-1}, \lambda^{-1}\}$ ; to approximate  $\{\dot{u}(0), \dot{\lambda}(0)\}$  we compute first  $\Delta s^0$  by

$$(2.54) \quad (\Delta s^0)^2 = \|u^0 - u^{-1}\|^2 + |\lambda^0 - \lambda^{-1}|^2,$$

and approximate  $\dot{u}(0), \dot{\lambda}(0)$  by

$$(2.55) \quad \frac{u^0 - u^{-1}}{\Delta s^0}, \quad \frac{\lambda^0 - \lambda^{-1}}{\Delta s^0},$$

respectively. The sign of  $\Delta s^0$  depends upon the orientation chosen for the arc of solutions and of the relative positions of  $\{u^0, \lambda^0\}$  and  $\{u^{-1}, \lambda^{-1}\}$  on it.

*Remark 2.4.* Relations (2.52), (2.53) employ difference quotients to replace or approximate the tangents used in (2.28b) and (2.29). From this idea we can derive

many other schemes for the approximation of (2.26), (2.27). Methods for the automatic adjustments of  $\Delta s$  are important but we do not discuss them here; see Rheinboldt [41]. A least squares conjugate gradient method for solving in  $V \times \mathbb{R}$  systems very close to (2.51)–(2.53) has been discussed in § 2.3. To start this algorithm we have used  $\{2u^n - u^{n-1}, 2\lambda^h - \lambda^{n-1}\}$  as an initial guess to compute  $\{u^{n+1}, \lambda^{n+1}\}$ . This is just another use of the difference approximation to the tangent, but now in (2.48).

**3. Application to the solution of the Bratu problem.**

**3.1. Formulation of the problem, properties of its solutions.** As a first test problem for the solution techniques discussed in § 2 we consider the numerical solution of a modified *Bratu problem*, i.e. find a solution  $u$  of the nonlinear boundary value problem:

$$(3.1) \quad \begin{aligned} -\Delta u &= \lambda e^u + f \quad \text{in } \Omega, \\ u &= 0 \quad \text{on } \partial\Omega. \end{aligned}$$

Here  $\Omega$  is a *bounded* domain of  $\mathbb{R}^N$  and  $\partial\Omega$  its boundary. We denote by  $x = \{x_i\}_{i=1}^N$  the generic point of  $\mathbb{R}^N$  and define  $dx$  by  $dx = dx_1 \cdots dx_N$ . The (quite classical) Sobolev–Hilbert space

$$(3.2) \quad H_0^1(\Omega) = \left\{ v \mid v \in L^2(\Omega), \frac{\partial v}{\partial x_i} \in L^2(\Omega), \forall i = 1, \dots, N, v = 0 \text{ on } \partial\Omega \right\},$$

equipped with the scalar product

$$(3.3) \quad (u, v)_{H_0^1(\Omega)} = \int_{\Omega} \nabla u \cdot \nabla v \, dx$$

and the corresponding form

$$(3.4) \quad \|v\|_{H_0^1(\Omega)} = \left( \int_{\Omega} |\nabla v|^2 \, dx \right)^{1/2},$$

provides a functional framework well suited to the solution of (3.1) by variational methods, and most particularly by those discussed in § 2.

For simplicity we consider only situations for which  $f$  is a *nonnegative constant* ( $f \equiv 0$  in the Bratu case). We suppose also that  $\lambda \geq 0$ , since problem (3.1) has a unique solution in  $H_0^1(\Omega)$  if  $\lambda \leq 0$ ; such a result can be proved using *monotonicity* methods, like those discussed in e.g. Lions [7], and based on the fact that the operator

$$v \rightarrow -\Delta v - \lambda e^v - f$$

is *monotone* over  $H_0^1(\Omega)$  if  $\lambda < 0$ . If  $\lambda > 0$ , problem (3.1) and closely related nonlinear problems, have been considered by many authors; with regard to recent publications let us mention among others Crandall–Rabinowitz [8], [9], Amann [10], Mignot–Puel [11], Mignot–Murat–Puel [12], Keller–Cohen [30], Keener–Keller [31]; in particular we find in [12] an interesting discussion showing relationships between (3.1) and *combustion phenomena* and in [30] a relationship to joule-heating in conductors. The following has been proved for  $\lambda > 0$ :

*There exists a critical value of  $\lambda$ , say  $\lambda^* > 0$ , such that:*

- (i) *If  $\lambda > \lambda^*$ , then problem (3.1) has no solution.*
- (ii) *If  $\lambda \in ]0, \lambda^*]$  (resp.  $\lambda \in ]0, \lambda^*[$ ), then problem (3.1) has at least one solution (resp. two solutions) belonging to  $H_0^1(\Omega) \cap W^{2,p}(\Omega) \forall p \geq 1$ , where*

$$W^{2,p}(\Omega) = \left\{ v \mid v, \frac{\partial v}{\partial x_i}, \frac{\partial^2 v}{\partial x_i \partial x_j} \in L^p(\Omega), \forall 1 \leq i, j \leq N \right\}.$$

(iii) If  $\lambda = \lambda^*$  there exists a unique  $u^* \in H_0^1(\Omega) \cap W^{2,p}(\Omega)$ ,  $\forall p \geq 1$ ; moreover  $\{u^*, \lambda^*\}$  is a simple limit point for the equation

$$S(u, \lambda) = 0,$$

where the operator  $S$  is defined over  $H_0^1(\Omega) \times \mathbb{R}$  by

$$S(v, \mu) = -\Delta v - \mu e^v - f.$$

In the above theoretical references it is also proved that these solutions which are not limit points are regular solutions. It follows from all these properties that the solution techniques discussed in § 2 can be applied to the solution of (3.1) if  $\lambda > 0$ . Their application to the computational solution of (3.1) requires, however, a *finite dimensional approximation* of this last problem; such an approximation—by *finite element methods*—is considered in the following sections. Problem (3.1) has been investigated numerically by, among others, Kikuchi [13], Simpson [14], Moore–Spence [15], Chan–Keller [16] (by arc length continuation and multigrid finite difference methods), Reinhart [3], to which we refer for more details and further references.

**3.2. Finite element approximation of the Bratu problem.**

**3.2.1. Variational formulation of the Bratu problem. Triangulation of  $\Omega$ . Fundamental discrete spaces.** A variational formulation of the Bratu problem (3.1), well suited to finite element approximations and to the solution techniques of § 2, is given by

(3.5) Find  $\{u, \lambda\} \in H_0^1(\Omega) \times \mathbb{R}$  such that

$$\int_{\Omega} \nabla u \cdot \nabla v \, dx = \int_{\Omega} (\lambda e^u + f)v \, dx \quad \forall v \in H_0^1(\Omega).$$

We describe only the approximation of problem (3.1) for  $N = 2$  (the one-dimensional case,  $N = 1$ , is much simpler); we suppose also for simplicity that  $\Omega$  is a polygonal domain of  $\mathbb{R}^2$ . We consider now a standard family of finite element triangulations  $\{\mathcal{T}_h\}_h$  of  $\Omega$ , i.e. for a given  $h$ ,  $\mathcal{T}_h$  is a finite collection of (*closed*) subtriangles,  $T$ , of  $\Omega$ , such that

- (i)  $\cup_{T \in \mathcal{T}_h} T = \bar{\Omega}$ ,
- (ii)  $\forall T, T' \in \mathcal{T}_h, T \neq T'$ , we have either
  - (\*)  $T \cap T' = \emptyset$ ,
  - (\*\*) or  $T, T'$  have only one vertex in common,
  - (\*\*\*) or  $T, T'$  have only a whole edge in common,
- (iii)  $h$  is the maximal length of the edges of the  $T \in \mathcal{T}_h$ .

An example of such a triangulation is shown in Fig. 3.1, for  $\Omega = ]0, 1[ \times ]0, 1[$ .

We approximate  $H_0^1(\Omega)$  by the finite-dimensional space:

(3.6)  $V_{0h} = \{v_h \mid v_h \in C^0(\bar{\Omega}), v_h|_T \in P_1 \quad \forall T \in \mathcal{T}_h, v_h = 0 \text{ on } \partial\Omega\}$

where  $P_1$  is the space of polynomials in  $x_1, x_2$  of degree  $\leq 1$ . It follows that  $\dim V_{0h} = N_{0h}$  where  $N_{0h}$  is the number of vertices of  $\mathcal{T}_h$  interior to  $\Omega$  and  $V_{0h} \subset H_0^1(\Omega)$ .

**3.2.2. Formulation of the approximate problems.** As an approximate problem it is quite natural to take:

(3.7) Find  $\{u_h, \lambda\} \in V_{0h} \times \mathbb{R}$  such that

$$\int_{\Omega} \nabla u_h \cdot \nabla v_h \, dx = \int_{\Omega} (\lambda e^{u_h} + f)v_h \, dx \quad \forall v_h \in V_{0h}.$$

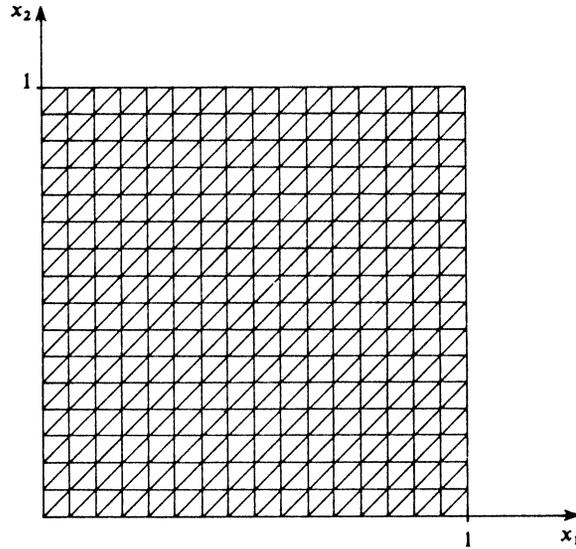


FIG. 3.1. Finite element triangulation for the Bratu problem.

Problem (3.7) is equivalent to a system of nonlinear equations in  $\mathbb{R}^{N_{0h}+1}$ . To obtain this system we suppose that the set  $\Sigma_{0h}$  of the vertices of  $\mathcal{T}_h$  has been ordered so that

$$(3.8) \quad \Sigma_{0h} = \{P_i\}_{i=1}^{N_{0h}},$$

and that to each  $P_i$  of  $\Sigma_{0h}$  we have associated the function  $w_i$  satisfying

$$(3.9) \quad w_i \in V_{0h}, w_i(P_j) = \delta_{ij} \quad \forall 1 \leq i, j \leq N_{0h}.$$

The set  $B_{0h} = \{w_i\}_{i=1}^{N_{0h}}$  is a basis of the vector space  $V_{0h}$  and we clearly have the important relation,  $\forall v_h \in V_{0h}$ :

$$(3.10) \quad v_h = \sum_{i=1}^{N_{0h}} v_h(P_i) w_i.$$

Using (3.10) in (3.7), we get the nonlinear system:

$$(3.11) \quad \sum_{j=1}^{N_{0h}} \left( \int_{\Omega} \nabla w_i \cdot \nabla w_j dx \right) u_h(P_j) = \int_{\Omega} \left( \lambda \exp \left( \sum_{j=1}^{N_{0h}} u_h(P_j) w_j \right) + f \right) w_i dx, \quad 1 \leq i \leq N_{0h}.$$

Here the unknown vector is  $\{ \{u_h(P_i)\}_{i=1}^{N_{0h}}, \lambda \} \in \mathbb{R}^{N_{0h} \times \mathbb{R}}$ .

Since  $\nabla w_i, \nabla w_j$  are piecewise constant functions, the calculation of the left-hand side is an easy task. The integrals occurring on the right-hand side of (3.11) can be calculated exactly. However in order to reduce the computational work, we evaluate these integrals approximately. Two possibilities are as follows:

(i) Calculate  $\int_{\Omega} e^{u_h} w_i dx$  using the *two-dimensional Simpson rule* on each triangle  $T \in \mathcal{T}_h$ , i.e.

$$(3.12) \quad \int_T \phi(x) dx \approx \frac{1}{3} \text{measure}(T) \sum_{j=1}^3 \phi(m_{jT}),$$

where  $m_{1T}, m_{2T}, m_{3T}$  are the midpoints of the three edges of  $T$ . Formula (3.12) is exact

if  $\phi \in P_2$  ( $P_2 =$  space of polynomials of degree  $\leq 2$ ). We need to apply Simpson's rule only on those triangles of  $\mathcal{T}_h$  with  $P_i$  as a common vertex.

(ii) Apply to  $\int_{\Omega} e^{u_h} w_i dx$  the *two-dimensional trapezoidal rule*, i.e.

$$(3.13) \quad \int_T \phi(x) dx \approx \frac{1}{3} \text{measure}(T) \sum_{j=1}^3 \phi(P_{jT}),$$

where  $P_{jT}$ ,  $j = 1, 2, 3$  are the vertices of  $T$ ; formula (3.13) is exact if  $\phi \in P_1$ . If  $\mathcal{T}_h$  is a regular triangulation, like the one of Fig. 3.1, using (3.13) to calculate the right-hand sides of (3.11) we recover classical finite difference schemes for the discretization of (3.1).

**3.3. Numerical solution of the discrete Bratu problem by arc length continuation methods.** We now apply the continuation methods of §§ 2.3, 2.4 to solve the discrete Bratu problem (3.7). This leads to the following algorithm:

(a) *Initialization.* Set

$$(3.14) \quad \lambda^0 = 0.$$

The corresponding  $u_h^0$  is the unique solution of the following discrete linear Dirichlet problem (given in variational form).

Find  $u_h^0 \in V_{0h}$  such that

$$(3.15) \quad \int_{\Omega} \nabla u_h^0 \cdot \nabla v_h dx = \int_{\Omega} f v_h dx \quad \forall v_h \in V_{0h}.$$

This is equivalent to a linear system (obtained by setting  $\lambda = 0$  in (3.11)) whose matrix is *symmetric* and *positive definite*. We take  $\{u_h^0, 0\}$  as the point on the arc of solutions  $\{u_h(s), \lambda(s)\}$  for which  $s = 0$ . Denote  $dX(s)/ds$  by  $\dot{X}(s)$  for  $X = u_h$  or  $X = \lambda$ . Then by differentiation of (3.7), with respect to  $s$ , we obtain at  $s = 0$ :

$$(3.16) \quad \int_{\Omega} \nabla \dot{u}_h(0) \cdot \nabla v_h dx = \dot{\lambda}(0) \int_{\Omega} e^{u_h^0} v_h dx \quad \forall v_h \in V_{0h},$$

$$\dot{u}_h(0) \in V_{0h}.$$

We also require as a definition of  $s$ :

$$(3.17) \quad \int_{\Omega} |\nabla \dot{u}_h(0)|^2 dx + \dot{\lambda}^2(0) = 1.$$

Define  $\hat{u}_h$  as the solution of

$$(3.18) \quad \hat{u}_h \in V_{0h},$$

$$\int_{\Omega} \nabla \hat{u}_h \cdot \nabla v_h dx = \int_{\Omega} e^{u_h^0} v_h dx \quad \forall v_h \in V_{0h}.$$

Then from (3.16)-(3.18) we have

$$(3.19) \quad \dot{u}_h(0) = \dot{\lambda}(0) \hat{u}_h,$$

$$(3.20) \quad \dot{\lambda}^2(0) = \left( 1 + \int_{\Omega} |\nabla \hat{u}_h|^2 dx \right)^{-1}.$$

Since we are interested in solving the Bratu problem for  $\lambda > 0$  and we have set  $\lambda(0) = 0$

we must orient the solution arc in such a way that  $d\lambda/ds (= \dot{\lambda}) \geq 0$ . Thus (3.20) yields:

$$(3.21) \quad \dot{\lambda}(0) = \left( 1 + \int_{\Omega} |\nabla \hat{u}_h|^2 dx \right)^{-1/2}.$$

(b) *Continuation.* With  $\Delta s (> 0)$  as a fixed step in arc length, we define for  $n \geq 0$  an approximation  $\{u_h^{n+1}, \lambda^{n+1}\} (\in V_{0h} \times \mathbb{R})$  of  $\{u_h((n+1)\Delta s), \lambda((n+1)\Delta s)\}$  as the solution of the following nonlinear variational system:

Find  $\{u_h^{n+1}, \lambda^{n+1}\} \in V_{0h} \times \mathbb{R}$  such that

$$(3.22a) \quad \int_{\Omega} \nabla u_h^{n+1} \cdot \nabla v_h dx = \int_{\Omega} (\lambda^{n+1} e^{u_h^{n+1}} + f) v_h dx \quad \forall v_h \in V_{0h},$$

$$(3.22b) \quad \int_{\Omega} \nabla (u_h^1 - u_h^0) \cdot \nabla \dot{u}_h(0) dx + (\lambda^1 - \lambda^0) \dot{\lambda}(0) = \Delta s \quad \text{if } n = 0,$$

$$(3.22c) \quad \int_{\Omega} \nabla (u_h^{n+1} - u_h^n) \cdot \nabla \left( \frac{u_h^n - u_h^{n-1}}{\Delta s} \right) dx + (\lambda^{n+1} - \lambda^n) \left( \frac{\lambda^n - \lambda^{n-1}}{\Delta s} \right) = \Delta s \quad \text{if } n \geq 1.$$

To solve the nonlinear system (3.22), we use the *nonlinear least squares conjugate gradient* techniques of § 2.3. We give a detailed description of the operations involved in the solution process for this first application.

A convenient nonlinear least squares formulation of (3.22) is:

$$(3.23) \quad \text{Find } \{u_h^{n+1}, \lambda^{n+1}\} \in V_{0h} \times \mathbb{R} \text{ such that}$$

$$J_{n+1}(u_h^{n+1}, \lambda^{n+1}) \leq J_{n+1}(w_h, \mu) \quad \forall \{w_h, \mu\} \in V_{0h} \times \mathbb{R}.$$

Here the functional  $J_{n+1}(\cdot, \cdot)$  is defined by

$$J_{n+1}(w_h, \mu) = \frac{1}{2} \int_{\Omega} |\nabla \tilde{w}_h|^2 dx + \frac{1}{2} |\tilde{\mu}|^2$$

where  $\tilde{w}_h$  and  $\tilde{\mu}$  are nonlinear functions of  $\{w_h, \mu\}$  obtained as the solutions of the *linear* problems:

$$(3.24) \quad \tilde{w}_h \in V_{0h},$$

$$\int_{\Omega} \nabla \tilde{w}_h \cdot \nabla v_h dx = \int_{\Omega} \nabla w_h \cdot \nabla v_h dx - \int_{\Omega} (\mu e^{w_h} + f) v_h dx \quad \forall v_h \in V_{0h},$$

$$(3.25) \quad \tilde{\mu} = \int_{\Omega} \nabla (w_h - u_h^n) \cdot \nabla \left( \frac{u_h^n - u_h^{n-1}}{\Delta s} \right) dx + (\mu - \lambda^n) \left( \frac{\lambda^n - \lambda^{n-1}}{\Delta s} \right) + \Delta s.$$

In this particular case, the conjugate gradient algorithm (2.37)-(2.46) reduces to:

*Step 0: Initialization.* For a given

$$(3.26) \quad \{u_h^0, \lambda^0\} \in V_{0h} \times \mathbb{R},$$

compute  $\{g_u^0, g_\lambda^0\} \in V_{0h} \times \mathbb{R}$  as the solution of:

$$(3.27) \quad \int_{\Omega} \nabla g_u^0 \cdot \nabla v_h dx = \left\langle \frac{\partial J_{n+1}}{\partial u_h}(u_h^0, \lambda^0), v_h \right\rangle \quad \forall v_h \in V_{0h},$$

$$(3.28) \quad g_\lambda^0 = \frac{\partial J_{n+1}}{\partial \lambda}(u_h^0, \lambda^0).$$

Then set

$$(3.29) \quad \{z_u^0, z_\lambda^0\} = \{g_u^0, g_\lambda^0\}$$

and for  $m \geq 0$ , assuming that  $\{u_h^m, \lambda^m\}$ ,  $\{g_u^m, g_\lambda^m\}$ ,  $\{z_u^m, z_\lambda^m\}$  are known, compute  $\{u_h^{m+1}, \lambda^{m+1}\}$ ,  $\{g_u^{m+1}, g_\lambda^{m+1}\}$ ,  $\{z_u^{m+1}, z_\lambda^{m+1}\}$  by:

Step 1: *Descent.* Find  $\rho_m \in \mathbb{R}$  such that,  $\forall \rho \in \mathbb{R}$ :

$$(3.30) \quad J_{n+1}(u_h^m - \rho_m z_u^m, \lambda^m - \rho_m z_\lambda^m) \leq J_{n+1}(u_h^m - \rho z_u^m, \lambda^m - \rho z_\lambda^m).$$

Then set

$$(3.31) \quad u_h^{m+1} = u_h^m - \rho_m z_u^m, \lambda^{m+1} = \lambda^m - \rho_m z_\lambda^m.$$

Step 2: *Calculate new descent direction.* Compute  $\{g_u^{m+1}, g_\lambda^{m+1}\} \in V_{0h} \times \mathbb{R}$  as the solution of:

$$(3.32) \quad \int_\Omega \nabla g_u^{m+1} \cdot \nabla v_h \, dx = \left\langle \frac{\partial J_{n+1}}{\partial u_h}(u_h^{m+1}, \lambda^{m+1}), v_h \right\rangle \quad \forall v_h \in V_{0h},$$

$$(3.33) \quad g_\lambda^{m+1} = \frac{\partial J_{n+1}}{\partial \lambda}(u_h^{m+1}, \lambda^{m+1}).$$

Evaluate

$$(3.34) \quad \gamma_m = \frac{\int_\Omega \nabla(g_u^{m+1} - g_u^m) \cdot \nabla g_u^{m+1} \, dx + (g_\lambda^{m+1} - g_\lambda^m) g_\lambda^{m+1}}{\int_\Omega |\nabla g_u^{m+1}|^2 \, dx + |g_\lambda^{m+1}|^2}$$

and

$$(3.35) \quad z_u^{m+1} = g_u^{m+1} + \gamma_m z_u^m, z_\lambda^{m+1} = g_\lambda^{m+1} + \gamma_m z_\lambda^m.$$

Then set  $m = m + 1$  and return to step 1.

As in § 2.4 we can use  $\{2u_h^n - u_h^{n-1}, 2\lambda^n - \lambda^{n-1}\}$  in (3.26). The partial derivatives  $\partial J_{n+1}/\partial u_h, \partial J_{n+1}/\partial \lambda$  (occurring in (3.32), (3.33)) can be evaluated using the derivative calculation technique of § 2.2. At  $\{w_h, \mu\}$  this gives:

$$(3.36) \quad \left\langle \frac{\partial J_{n+1}}{\partial u_h}(w_h, \mu), v_h \right\rangle = \int_\Omega \nabla \tilde{w}_h \cdot \nabla v_h \, dx - \mu \int_\Omega e^{w_h} \tilde{w}_h v_h \, dx + \tilde{\mu} \int_\Omega \nabla \left( \frac{u_h^n - u_h^{n-1}}{\Delta s} \right) \cdot \nabla v_h \, dx \quad \forall v_h \in V_{0h},$$

$$(3.37) \quad \frac{\partial J_{n+1}}{\partial \lambda}(w_h, \mu) = \tilde{\mu} \left( \frac{\lambda^n - \lambda^{n-1}}{\Delta s} \right) - \int_\Omega e^{w_h} \tilde{w}_h \, dx.$$

Here of course  $\{\tilde{w}_h, \tilde{\mu}\}$  are obtained from  $\{w_h, \mu\}$ , through the solution of (3.24), (3.25).

As convergence test we took

$$(3.38) \quad J_{n+1}(u_h^{m+1}, \lambda^{m+1}) \leq \varepsilon.$$

For the examples we used  $\varepsilon = 10^{-6}$ .

**3.4. Numerical examples.** We have employed the above indicated procedures to solve problem (3.1) in three specific cases:

- A.  $f \equiv 0, \quad \Omega \equiv ]0, 1[;$
- B.  $f \equiv 1, \quad \Omega \equiv ]0, 1[;$
- C.  $f \equiv 0, \quad \Omega \equiv ]0, 1[ \times ]0, 1[.$

Cases A and B have been discretized by *one-dimensional finite elements*, using a space discretization step  $h = 0.1$ . Case C has been approximated using a triangulation  $\mathcal{T}_h$  as shown in Fig. 3.1, consisting of 512 triangles. The unknowns are the values taken by the approximate solution  $u_h$  at the interior nodes of  $\mathcal{T}_h$ ; we have 225 such nodes. The continuation algorithm described in § 3.3 has been applied with fixed  $\Delta s = 0.1$ .

We show in Fig. 3.2 the variation of  $u_h(0.5, 0.5)$  (maximal value of  $u_h$ ) as a function of  $\lambda$  for case C. The numerical results agree very well with those of Kikuchi [13], obtained by quite different methods.

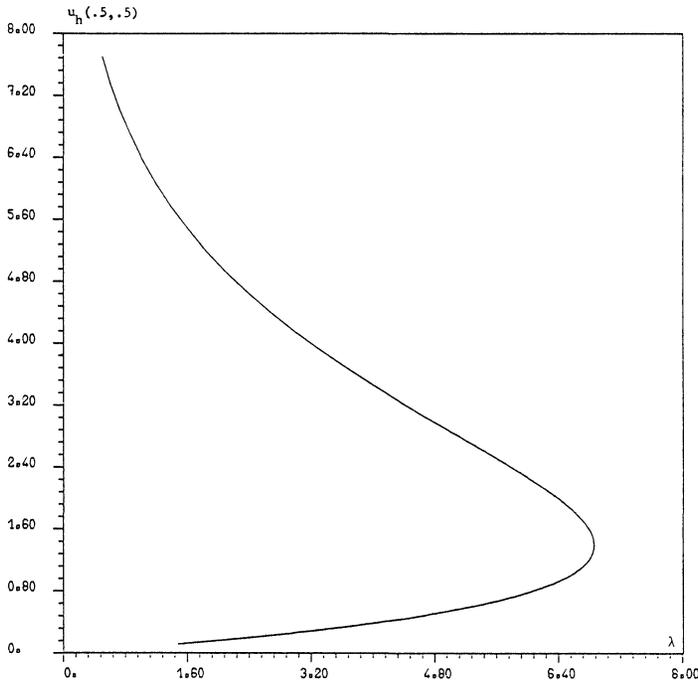


FIG. 3.2. Solution  $u_h$  at  $\{x, y\} = \{0.5, 0.5\}$  of the Bratu problem on the unit square (case C).

With  $\Delta s = 0.1$ , the solution of the above three test problems never required more than 3 to 4 iterations of the conjugate gradient algorithm (3.26), (3.35) to obtain  $\{u_h^{n+1}, \lambda^{n+1}\}$  from  $\{u_h^n, \lambda^n\}$  and  $\{u_h^{n-1}, \lambda^{n-1}\}$  via the solution of the least squares problem (3.23). This efficiency is partly due to the good initialization of algorithm (3.26)–(3.35) provided by the initial guess  $\{2u_h^n - u_h^{n-1}, 2\lambda^n - \lambda^{n-1}\}$  and partly due to the small step size of  $\Delta s$ . Using the above methods there were no difficulties close to and at the limit point.

We point out that each iteration of the conjugate gradient algorithm (3.26)–(3.35), requires the solution of several discrete linear systems with a *fixed* coefficient matrix independent of  $n$  and  $m$ ; since this matrix is symmetric and positive definite we use *only one* Cholesky factorization, taking into account the sparsity of the matrix. The solution procedure is thus quite efficient.

#### 4. Applications to the solution of bifurcation problems via perturbed bifurcations.

**4.1. Synopsis. Generalities.** In this section we discuss the numerical treatment of *nonlinear second order boundary value problems* whose branches of solutions exhibit *bifurcation*. To do this we perturb the original problem into a new one whose branches

of solutions do not bifurcate but instead have *limit or fold points*. The solution branches can then be computed by the continuation methods of § 2.3. The process is based on the use of a simple perturbation method related to the concept of *perturbed bifurcation*, see Keener-Keller [32], Matkowsky-Reiss [33]. By continuing from the perturbed to the unperturbed problem we can recover the bifurcating solution branches.

**4.2. First example: A nonlinear Dirichlet problem.** With  $\Omega$  a bounded domain of  $\mathbb{R}^N$  ( $N \geq 1$ ), we consider the solution in  $V = H_0^1(\Omega)$ , of the nonlinear Dirichlet problem:

$$(4.1) \quad \begin{aligned} -\Delta u &= \lambda u^2 + \delta && \text{in } \Omega, \\ u &= 0 && \text{on } \partial\Omega. \end{aligned}$$

Here  $\delta \in \mathbb{R}$  is the perturbing parameter. The nonperturbed problem ( $\delta = 0$ ) has two solution branches for  $\lambda \geq 0$ :

(i) the trivial branch  $\{u, \lambda\} = \{0, \lambda\}, \forall \lambda \in \mathbb{R}$ .

(ii) a nontrivial branch which never crosses the trivial one (see Fig. 4.1). By symmetry with respect to  $\{u, \lambda\} \rightarrow \{-u, -\lambda\}$  about  $\{0, 0\}$  we easily obtain the unperturbed solutions of (4.1) corresponding to  $\lambda \leq 0$ . Thus in general the unperturbed problem with  $\delta = 0$ , has three disjoint solution branches:  $u \equiv 0, \lambda \in \mathbb{R}$  and two ‘‘hyperbolic’’ branches, one for  $\lambda > 0$ , and one for  $\lambda < 0$ . For the perturbed problem,  $\delta \neq 0$ , it follows from Mignot-Puel [11] that only two distinct branches exist. One of these is a perturbation of one of the hyperbolic branches and it contains only regular solutions. The other is formed from the perturbed trivial branch joining the other hyperbolic branch and it contains one simple fold point. Figure 4.1 shows some of these solution branches for  $\delta \geq 0$ .

This problem with  $\delta = 0$  is not, technically, a bifurcation problem. Rather we may say that it exhibits bifurcation at  $\lambda = +\infty$ . However it furnishes a clear example of our perturbation techniques. We now describe our procedure for computing the nontrivial branch of solutions of the unperturbed problem (4.1) with  $\lambda \geq 0$ . With fixed  $\delta > 0$ , ‘‘sufficiently small’’, we solve (4.1) by a continuation method, as described in § 2.3, using  $\{u, \lambda\} = \{u_\delta^0, 0\}$  as starting point. Here  $u_\delta^0 \in H_0^1(\Omega)$  is the solution of

$$(4.2) \quad \begin{aligned} -\Delta u_\delta^0 &= \delta && \text{in } \Omega, \\ u_\delta^0 &= 0 && \text{on } \partial\Omega. \end{aligned}$$

For  $\delta > 0$  and sufficiently small, the upper part,  $C_\delta^+$  of the branch of solutions of the perturbed problem, away from the limit point  $\{u_{\delta_c}, \lambda_{\delta_c}\}$  is a good approximation of the nontrivial branch of solutions of the nonperturbed problem. We take two distinct points on  $C_\delta^+$ , say  $\{u_{\delta_1}, \lambda_{\delta_1}\}$  and  $\{u_{\delta_2}, \lambda_{\delta_2}\}$  and compute the nontrivial solutions of the unperturbed problem corresponding to values  $\lambda = \lambda_{\delta_1}$  and  $\lambda = \lambda_{\delta_2}$ . These solutions can be obtained using simply the least squares conjugate gradient method of §§ 2.1, 2.2 (i.e. without continuation), taking  $u_{\delta_1}$  and  $u_{\delta_2}$  as starting points. If necessary, however, continuation with respect to  $\delta$  can be used to reach the value  $\delta = 0$ .

Once two distinct solutions (sufficiently close to each other) on the nontrivial branch of solutions of the unperturbed problem have been obtained, we can use continuation, again, to compute the whole unperturbed branch. Fig. 4.1 illustrates the indicated process. Indeed the curves in this figure are the results of computation using  $\Omega \equiv ]0, 1[$  in (4.1). We discuss these calculations below where also the influence of step size  $\Delta s$  and other factors are considered.

The above technique has been applied to compute the nontrivial solutions of more complicated nonlinear boundary value problems. We discuss some such examples in

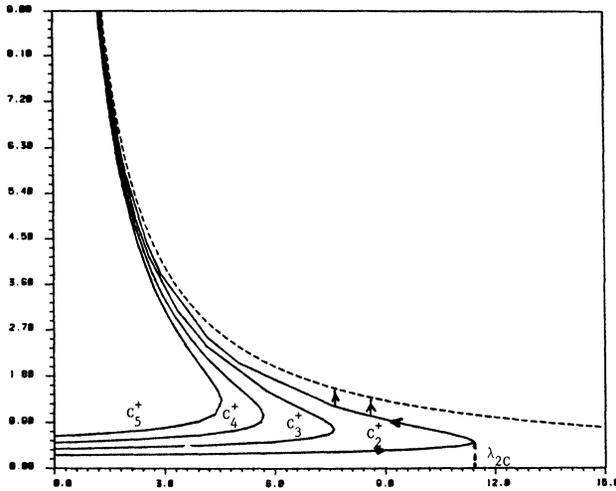


FIG. 4.1

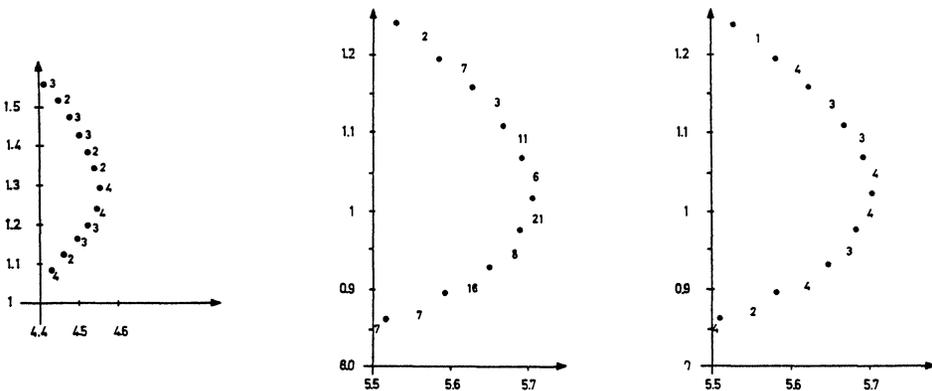
§§ 4.3 and 4.4. The *Von Karman equations* for nonlinear plates are treated in Reinhart [3], [17].

*Computational Results and Tests.* The methods of § 2 were applied to (4.1) with  $\Omega \equiv ]0, 1[$  for  $\delta = 5, 4, 3, 2, 0$ . The nontrivial branch of solutions corresponding to  $\delta = 0$  is obtained by the method indicated on Fig. 4.1. We have used  $h = 0.1$  and  $\Delta s = 0.1$  for the space discretization and the continuation algorithm, respectively. The numerical results are shown on Fig. 4.1, where we have plotted  $\max_{x \in [0,1]} u_h(x) = u_h(.5)$  versus  $\lambda$ .

The computed results agree with those obtained elsewhere by other methods.

Using (4.1) with  $\delta = 5$  as a test problem we show in Fig. 4.2(a) the number of conjugate gradient iterations necessary to solve the least squares problem encountered at each step of the continuation process. The convergence test is as indicated in (3.38), with  $\epsilon = 10^{-6}$ .

If one takes  $\gamma_n = 0$  in algorithm (2.37)-(2.46) (instead of  $\gamma_n$  given by (2.45)) we recover a *steepest descent* algorithm for solving the least squares problem (2.33). In the particular case of (4.1) with  $\delta = 4$  we have done a comparison between the performances of the steepest descent and conjugate gradient algorithms when applied to the continuation solution. The computed results are summarized on Fig. 4.2(b)



(a) Conjugate gradient ( $\delta = 5$ ).

(b) Steepest descent ( $\delta = 4$ ).

(c) Conjugate gradient ( $\delta = 4$ ).

FIG. 4.2

(steepest descent) and 4.2(c) (conjugate gradient). They show clearly the superiority of the conjugate variant in the neighborhood of the limit point. Note that the steepest descent case “oscillates”, each step taking more or less iterates depending upon the previous rate of convergence. This effect is less apparent in the conjugate gradient case. Finally we note some effects of the size of  $\Delta s$  upon the convergence of our continuation method, *particularly in the neighborhood of the limit point.*

(a) If  $\Delta s$  is too large, the algorithm does not converge close to the limit point. We can explain this behavior by the fact that the initial guess at the solution, provided by  $\{2u_h^n - u_h^{n-1}, 2\lambda^n - \lambda^{n-1}\}$  or  $\{u_h^n, \lambda^n\}$  is too far from the branch of solutions.

(b) The smaller the  $\Delta s$ , the smaller is the number of iterations close to the limit points. However if we are sufficiently far from the limit point the number of iterations is quite small and essentially independent of  $\Delta s$ .

(c) The smaller the  $\Delta s$ , the better is the approximation to the location of the limit point.

In conclusion, we should use large  $\Delta s$  if we are sufficiently far from the limit point, and decrease  $\Delta s$  if we are close to the limit point (further details concerning the choice of  $\Delta s$  may be found in [3], [5], in Rheinboldt [41] and in Perozzi [34]).

### 4.3. Bifurcation from a trivial branch.

**4.3.1. Synopsis. Generalities.** In this section we study simple nonlinear eigenvalue problems with bifurcations from the trivial branch. In the perturbed form, these problems are:

*Find  $u \in H_0^1(\Omega)$  such that*

$$(4.3) \quad \begin{aligned} -\Delta u &= \lambda u + f(u, \lambda) + \delta \quad \text{in } \Omega, \\ u &= 0 \quad \text{on } \partial\Omega. \end{aligned}$$

Here  $f$  satisfies:

$$(4.4) \quad f(0, \lambda) = 0 \quad \forall \lambda \in \mathbb{R}$$

and

$$(4.5) \quad f'_u(0, \lambda) = 0 \quad \forall \lambda \in \mathbb{R}.$$

For  $\delta = 0$  we note that  $\{u, \lambda\} \equiv \{0, \lambda\}$  is a solution of (4.3) for all  $\lambda \in \mathbb{R}$ . This is the trivial branch and we seek nontrivial branches bifurcating from it. The linearized problem about  $u = 0$ , reduces to:

$$(4.6) \quad \begin{aligned} -\Delta w &= \lambda w \quad \text{in } \Omega, \\ w &= 0 \quad \text{on } \partial\Omega. \end{aligned}$$

It is well known (i.e. Crandall-Rabinowitz [18] or Keller-Langford [36]) that if  $\lambda_i$  is an eigenvalue of *multiplicity one* of (4.6), then the pair  $\{0, \lambda_i\}$  is a simple *bifurcation point* for solutions of the unperturbed problem:  $\delta \equiv 0$  in (4.3). If  $w_i$  is a corresponding eigenfunction then it is also well known (i.e. Keller-Langford [36] or Brezzi-Rappaz-Raviart [19]) that the bifurcation is symmetric for  $\alpha_i = 0$  and asymmetric or transverse for  $\alpha_i \neq 0$  where:

$$(4.7) \quad \alpha_i \equiv \int_{\Omega} w_i^3(x) \int_{\Omega} G(x, \xi) f''_{uu}(0, \lambda_i) d\xi dx.$$

Here  $G(x, \xi)$  is the Green's function for  $(-\Delta)$  on  $\Omega$ . These cases are illustrated by the curves for  $\delta = 0$  in Figs. 4.3 and 4.4, respectively. If  $\delta \neq 0$  we have the local behavior

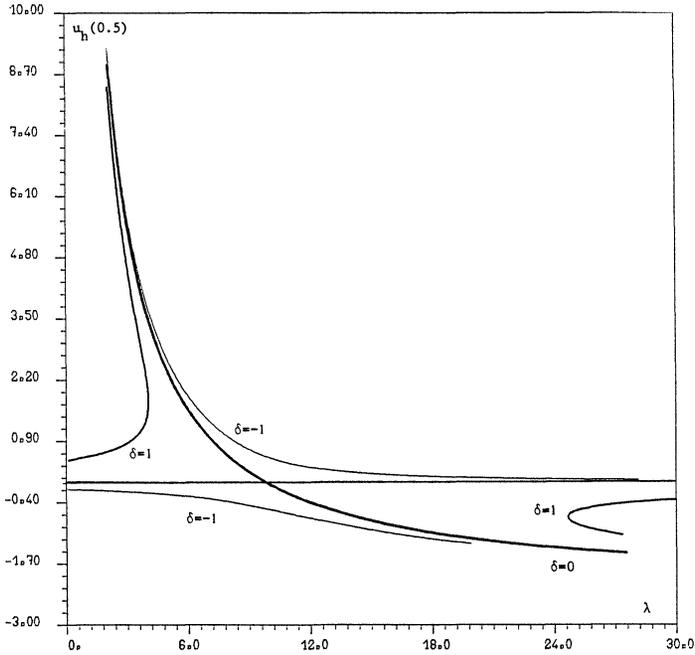


FIG. 4.3.  $-u'' = \lambda u + \lambda u^2/2 + \delta$  in  $]0, 1[$ ,  $u(0) = u(1) = 0$ .

indicated on Figs. 4.3 and 4.4, for the solutions of the perturbed problem (4.3). These configurations are called perturbed bifurcation in Keener-Keller [32] or imperfect bifurcations in Matkowsky-Reiss [33]. We shall study in particular the cases of (4.3) in which

(4.8) 
$$f(u, \lambda) \equiv \frac{\lambda}{2} u^2,$$

and

(4.9) 
$$f(u, \lambda) \equiv -u^3.$$

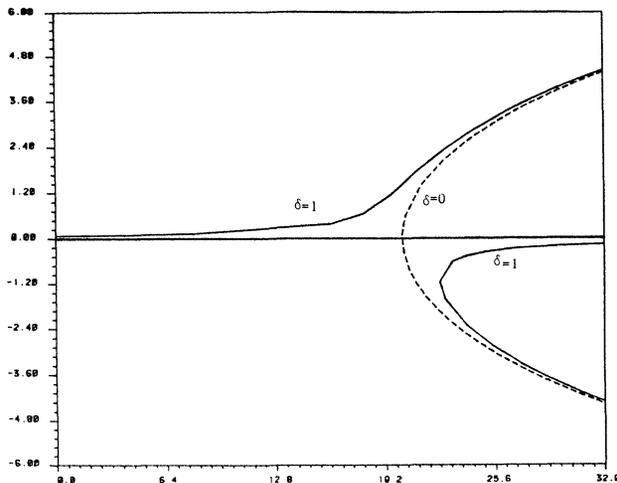


FIG. 4.4.  $-\Delta u = \lambda u - u^3 + \delta$  in  $]0, 1[ \times ]0, 1[$  ( $= \Omega$ ),  $u = 0$  on  $\partial\Omega$ .

It easily follows from (4.6) and (4.7) that for  $\Omega \equiv ]0, 1[$ :

$$(4.10a) \quad \alpha_i = 0, \forall i = 1, 2, \dots \quad \text{when } f \text{ is as in (4.9),}$$

$$(4.10b) \quad \alpha_i = 0 \text{ for } i = 2, 4, 6, \dots \quad \text{when } f \text{ is as in (4.8),}$$

$$\alpha_i \neq 0 \text{ for } i = 1, 3, 5, \dots$$

Similarly for  $\Omega \equiv ]0, 1[ \times ]0, 1[$  it follows that for the smallest eigenvalue  $\lambda_{11}$ :

$$(4.11) \quad \alpha_{11} \begin{cases} = 0 & \text{for (4.9),} \\ \neq 0 & \text{for (4.8).} \end{cases}$$

Using *finite element approximations* and the *continuation methods* discussed in § 4.2, we have computed approximate solutions of the perturbed and unperturbed problems near the *first eigenvalue* of the linearized problems. In Fig. 4.3 we show results for  $\Omega = ]0, 1[$  and  $f(u, \lambda) = \lambda u^2/2$ .

Both perturbed and unperturbed asymmetric bifurcation phenomena at the first eigenvalue of the linearized approximate problem are illustrated. For  $f(u, \lambda) = -u^3$  and  $\Omega \equiv ]0, 1[ \times ]0, 1[$  we have symmetric bifurcation phenomenon ( $\alpha_{11} = 0$ ) at the first eigenvalue of the linearized problem. The results are shown on Fig. 4.4 for both perturbed and unperturbed cases. For more details about the numerical procedure we refer to [3]. We refer also to [17] where it is shown (theoretically and computationally) that the solutions of Von Karman equations for nonlinear plates have the same qualitative behavior as observed here for  $f(u) = -u^3$  (for the first eigenvalue of the linearized problem).

**4.4. Bifurcation from a nontrivial branch.**

**4.4.1. Formulation and properties of the solutions.** We discuss in this section the solution of the nonlinear boundary value problem of *Neumann type*:

*Find  $\{u, \lambda\} \in H^1(0, 1) \times \mathbb{R}$  such that:*

$$(4.12) \quad -\frac{u''}{\pi^2} + u = \lambda e^u \text{ on } ]0, 1[,$$

$$u'(0) = u'(1) = 0.$$

Problem (4.12) has a branch of solutions  $\{u, \lambda\}$  with  $u = \text{const.}$  on  $]0, 1[$ . This constant is any root of

$$(4.13a) \quad u = \lambda e^u.$$

Alternatively each  $u \in \mathbb{R}$  is a root of (4.13a) for the value

$$(4.13b) \quad \lambda = u e^{-u}.$$

The “almost trivial” solution branch of (4.12) given by (4.13) is shown in Fig. 4.5a.

To find solutions bifurcating from the nontrivial branch, we note that the linearized form of (4.12) about the nontrivial branch is

$$(4.14) \quad -\frac{w''}{\pi^2} + w = \lambda e^u w \text{ in } ]0, 1[,$$

$$w'(0) = w'(1) = 0.$$

It easily follows that  $\lambda e^u$  must have one of the values

$$(4.15) \quad \lambda e^u = 1 + k^2, \quad k = 0, 1, \dots$$

Since the relation  $u = \lambda e^u$  holds, we find that the set of "singular" points  $\{u, \lambda\} = \{u_k, \lambda_k\}$  is a *discrete* set defined for  $k = 0, 1, \dots$ , by:

$$(4.16) \quad u_k = 1 + k^2, \quad \lambda_k = (1 + k^2) e^{-(1+k^2)}.$$

We can take as eigenfunctions  $w_k$  in (4.14):

$$(4.17) \quad w_k(x) = \cos k\pi x, \quad k = 0, 1, \dots.$$

The first "singular" point (obtained by taking  $k = 0$  in (4.16)) is  $\{1, e^{-1}\}$ .

We can show that it is a *simple limit point* for problem (4.12). This reduces to showing that  $\pi^2 e$  is not in the range of:  $-d^2/dx^2 \phi$  subject to  $\phi'(0) = \phi'(1) = 0$ . But this follows since  $\pi^2 e \neq 0$ .

On the other hand  $\{u_1, \lambda_1\} = \{2, 2e^{-2}\}$  is a *simple bifurcation* point and it can be proved (using e.g., [36] or [19]) that the bifurcation at  $\{u_1, \lambda_1\}$  is a *symmetric* one. All points  $\{u_k, \lambda_k\}$  for  $k > 1$  are also bifurcation points.

**4.4.2. Numerical results.** To compute the nonconstant solutions of (4.12) we use that combination of *finite element approximation* and *continuation techniques* already used in the previous sections. To avoid difficulties close to the bifurcation points during the continuation process we introduce a *perturbation* of the problem in the *boundary* conditions to get

$$(4.18) \quad \begin{aligned} -\frac{u''}{\pi^2} + u &= \lambda e^u \quad \text{on } ]0, 1[, \\ -u'(0) &= u'(1) = \delta. \end{aligned}$$

We observe that if  $\{u, \lambda\}$  is a solution of (4.18) and if  $u^*$  is defined by  $u^*(x) \equiv u(1-x)$ , then  $\{u^*, \lambda\}$  is a solution with  $\delta$  replaced by  $-\delta$ . This property of course holds if  $\delta = 0$ . It holds also for the approximate problems.

With  $N$  a positive integer and  $h = 1/N$ , we define  $x_i$  by

$$x_i = ih, \quad i = 0, \dots, N,$$

and we use piecewise linear elements and the trapezoidal rule in our variational formulation. The resulting system of  $N+2$  nonlinear equations is identical to the standard finite difference formulation of the Neumann problem (4.18). Our solution algorithm is but a trivial modification of that described in § 3.

Using the continuation strategy summarized in Fig. 4.1 we have computed branches of solutions of the perturbed problem (4.18) and also the nonconstant

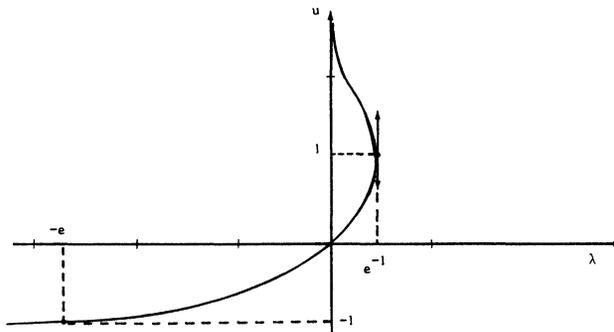


FIG. 4.5a. Constant solutions,  $u = \lambda e^u$ , of (4.12).

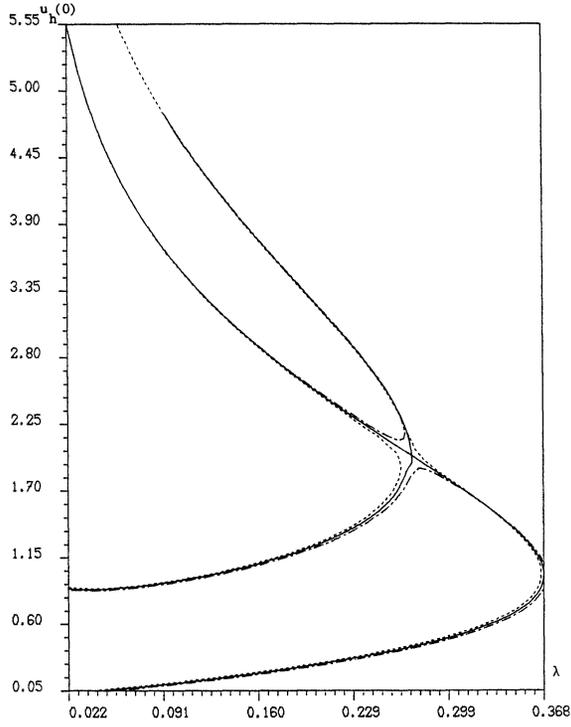


FIG. 4.5b.  $-u''/\pi^2 + u = \lambda e^u$ .  $-u'(0) = u'(1) = \delta$ . ----  $\delta = 0.01$ , —  $\delta = 0$ , -·-  $\delta = -0.01$ .

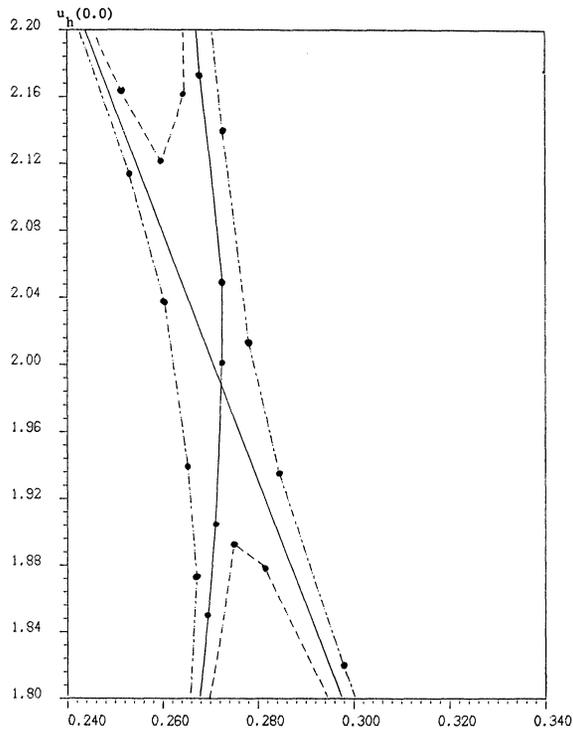


FIG. 4.6.  $-u''/\pi^2 + u = \lambda e^u$ ,  $u'(0) = -u'(1) = \delta$ . ---  $\delta = 0.01$ , —  $\delta = 0$ , -·-  $\delta = -0.01$ .

solutions of the unperturbed problem (4.12). The following results have been obtained using  $N = 20$  (i.e.  $h = 0.05$ ) for the approximate problems, and  $\delta = 0.01$  as perturbation parameter.

The variation of  $u_h(0)$  with  $\lambda$  is shown in Fig. 4.5b. (Clearly the behavior of  $u_h(1)$  is also described by this figure). Since the first bifurcation is *symmetric*, the tangent to the branch of nonconstant solutions of this bifurcation point has to be vertical; it is so with good precision. Actually, using smaller  $\Delta s$  and amplifying the vertical variations, we have shown in Fig. 4.6 the variations of  $u_h(0)$  and the above property of vertical tangent appears even more clearly.

## 5. Application to the Navier–Stokes equations for incompressible viscous fluids.

**5.1. Formulation of the Navier–Stokes equations.** Let  $\Omega$  be a domain of  $\mathbb{R}^N$  ( $N = 2, 3$  in practice) and  $\Gamma$  be its boundary. The *steady* flows of an incompressible and viscous Newtonian fluid, in  $\Omega$ , are modelled by the *Navier–Stokes equations*:

$$(5.1) \quad -\nu \Delta \mathbf{u} + (\mathbf{u} \cdot \nabla) \mathbf{u} + \nabla p = \mathbf{f} \quad \text{in } \Omega,$$

$$(5.2) \quad \nabla \cdot \mathbf{u} = 0 \quad \text{in } \Omega \quad (\text{incompressibility condition}).$$

In (5.1), (5.2):

$\mathbf{u} = \{u_i\}_{i=1}^N$  is the *flow velocity*,

$p$  is the *pressure*,

$\nu$  is a *viscosity* parameter,

$\mathbf{f}$  is a *density of external forces*,

$(\mathbf{u} \cdot \nabla) \mathbf{u}$  is a symbolic notation for the vector-function  $\{u_i \partial u_i / \partial x_j\}_{i=1}^N$ .

Typical *boundary conditions* associated with (5.1), (5.2) are

$$(5.3) \quad \mathbf{u} = \mathbf{u}_\beta \quad \text{on } \Gamma,$$

where  $\mathbf{u}_\beta$  is a given function defined over  $\Gamma$  and satisfying (from the incompressibility condition (5.2))

$$(5.4) \quad \int_{\Gamma} \mathbf{u}_\beta \cdot \mathbf{n} \, d\Gamma = 0$$

where  $\mathbf{n}$  is the outward normal unit vector on  $\Gamma$ .

The Navier–Stokes equations for incompressible viscous fluids have motivated a countless number of papers, reports, books, conferences, workshops, from both the theoretical and numerical points of view. Concentrating on books only, we mention, among others: Lions [7], Ladyzhenskaya [20], Temam [21], Girault–Raviart [22], Rautmann [23], Thomasset [24], Glowinski [42, Chapt. 7]; we refer also to the numerous references contained in these books.

It follows in particular from [7], [20], [21] that if  $\mathbf{f}$  and  $\mathbf{u}_\beta$  are sufficiently smooth, then problem (5.1), (5.2), (5.3) has a solution  $\{\mathbf{u}, p\}$  belonging to  $(H^1(\Omega))^N \times (L^2(\Omega)/\mathbb{R})$  (the pressure  $p$  is clearly determined only to within an arbitrary constant). If we suppose in addition that  $\nu$  is sufficiently large (or equivalently—if  $\nu$  is given—that  $\mathbf{f}$  and  $\mathbf{u}_\beta$  are sufficiently small), then problem (5.1)–(5.3) has unique solution in  $(H^1(\Omega))^N \times (L^2(\Omega)/\mathbb{R})$ .

**5.2. Stream function–vorticity formulation of the Navier–Stokes equations.** We suppose from now on that  $\Omega$  is a bounded domain of  $\mathbb{R}^2$ . We also assume for simplicity that  $\Omega$  is *simply connected* (see e.g. Glowinski–Pironneau [25] for the case where  $\Omega$  is multiply connected). With  $\Gamma$  the boundary of  $\Omega$ , let  $\mathbf{n}, \mathbf{s}$  be respectively the unit vector

of the outward normal at  $\Omega$  on  $\Gamma$  and the unit vector of the corresponding *oriented tangent*.

There exists from (5.2) a stream function  $\psi$  (determined only to within an arbitrary constant) such that

$$(5.5) \quad u_1 = \frac{\partial \psi}{\partial x_2}, \quad u_2 = -\frac{\partial \psi}{\partial x_1},$$

and it follows from (5.1), (5.5) that  $\psi$  satisfies the following (well-known) *nonlinear biharmonic equation*

$$(5.6) \quad \nu \Delta^2 \psi + \frac{\partial \psi}{\partial x_1} \frac{\partial}{\partial x_2} \Delta \psi - \frac{\partial \psi}{\partial x_2} \frac{\partial}{\partial x_1} \Delta \psi = \frac{\partial f_2}{\partial x_1} - \frac{\partial f_1}{\partial x_2} \quad \text{in } \Omega.$$

Concerning the *boundary conditions* we have

$$(5.7) \quad \frac{\partial \psi}{\partial s} = \mathbf{u}_\beta \cdot \mathbf{n} \quad \text{on } \Gamma.$$

Since  $\int_\Gamma \mathbf{u}_\beta \cdot \mathbf{n} \, d\Gamma = 0$ , (5.7) implies that

$$(5.8) \quad \psi(M) = \int_{M_0 M} \mathbf{u}_\beta \cdot \mathbf{n} \, d\Gamma \quad \forall M \in \Gamma,$$

where  $M_0 \in \Gamma$  ( $M_0$  can be arbitrarily chosen and we have prescribed  $\psi(M_0) = 0$ ). We also have

$$(5.9) \quad \frac{\partial \psi}{\partial n} = -\mathbf{s} \cdot \mathbf{u}_\beta \quad \text{on } \Gamma.$$

Actually (5.6), (5.8), (5.9) is a particular case of the more general family of nonlinear biharmonic problems

$$(5.10) \quad \nu \Delta^2 \psi + \frac{\partial \psi}{\partial x_1} \frac{\partial}{\partial x_2} \Delta \psi - \frac{\partial \psi}{\partial x_2} \frac{\partial}{\partial x_1} \Delta \psi = f \quad \text{in } \Omega,$$

$$(5.11) \quad \psi = g_1 \quad \text{on } \Gamma,$$

$$(5.12) \quad \frac{\partial \psi}{\partial n} = g_2 \quad \text{on } \Gamma.$$

An equivalent formulation of (5.10)–(5.12) as a nonlinear system of coupled *second order* elliptic equations is

$$(5.13) \quad -\nu \Delta \omega + \frac{\partial \omega}{\partial x_1} \frac{\partial \psi}{\partial x_2} - \frac{\partial \omega}{\partial x_2} \frac{\partial \psi}{\partial x_1} = f \quad \text{in } \Omega,$$

$$(5.14) \quad -\Delta \psi = \omega \quad \text{in } \Omega,$$

with the boundary conditions (5.11), (5.12). In (5.13), (5.14),  $\omega$  is the *voriticity* of the flow.

**5.3. Variational formulations.** We suppose that  $g = \{g_1, g_2\}$  is sufficiently smooth (see [25] for the precise requirement), so that there exists  $\psi_0$  such that,  $\psi_0|_\Gamma = g_1$ ,  $(\partial \psi_0 / \partial n)|_\Gamma = g_2$ . Let us define  $V_g$  by

$$(5.15) \quad V_g = \left\{ \phi \mid \phi \in H^2(\Omega), \phi|_\Gamma = g_1, \frac{\partial \psi}{\partial n} \Big|_\Gamma = g_2 \right\};$$

then  $V_g$  is a *nonempty, closed, affine* subspace of  $H^2(\Omega)$ , where

$$H^2(\Omega) = \left\{ \phi \mid \phi, \frac{\partial \phi}{\partial x_i}, \frac{\partial^2 \phi}{\partial x_i \partial x_j} \in L^2(\Omega), \forall i, j \right\}.$$

In particular  $V_0 = \{ \phi \mid \phi \in H^2(\Omega), \phi|_\Gamma = (\partial \phi / \partial n)|_\Gamma = 0 \} \equiv H_0^2(\Omega)$  is a *closed* subspace of  $H^2(\Omega)$ . We recall— $\Omega$  being bounded—that  $\phi \rightarrow (\int_\Omega |\Delta \phi|^2 dx)^{1/2}$  is a norm on  $V_0$  equivalent to the  $H^2$ -norm.

A *variational formulation* of (5.10)–(5.12) is then:

Find  $\psi \in V_g$  such that  $\forall \phi \in V_0$

$$(5.16) \quad \nu \int_\Omega \Delta \psi \Delta \phi dx + \int_\Omega \Delta \psi \left( \frac{\partial \psi}{\partial x_2} \frac{\partial \phi}{\partial x_1} - \frac{\partial \psi}{\partial x_1} \frac{\partial \phi}{\partial x_2} \right) dx = \int_\Omega f \phi dx.$$

To obtain a variational formulation of (5.11)–(5.14) seems to be more complicated; in fact, introducing  $\theta \in L^2(\Omega)$  such that  $\theta = -\Delta \phi$  and using (5.14), it follows from (5.16) that the pair  $\{ \omega, \psi \}$  satisfies

$$(5.17) \quad \begin{aligned} &\{ \omega, \psi \} \in W_g, \\ &\text{and } \forall \{ \theta, \phi \} \in W_0 \text{ we have} \\ &\nu \int_\Omega \omega \theta dx + \int_\Omega \omega \left( \frac{\partial \psi}{\partial x_1} \frac{\partial \phi}{\partial x_2} - \frac{\partial \psi}{\partial x_2} \frac{\partial \phi}{\partial x_1} \right) dx = \int_\Omega f \phi dx, \end{aligned}$$

where

$$(5.18) \quad W_0 = \{ \{ \theta, \phi \} \mid \theta \in L^2(\Omega), \phi \in V_0, -\Delta \phi = \theta \text{ in } \Omega \},$$

$$(5.19) \quad W_g = \{ \{ \theta, \phi \} \mid \theta \in L^2(\Omega), \phi \in V_g, -\Delta \phi = \theta \text{ in } \Omega \}.$$

Conversely if a pair  $\{ \omega, \psi \}$  satisfies (5.17), then  $\{ \omega, \psi \}$  is also a solution of the nonlinear boundary value problem (5.11)–(5.14) (and  $\psi$  a solution of (5.10)–(5.12)).

The variational formulation (5.17) of problem (5.11)–(5.14) contains *second order* derivatives in the definition of  $W_0$  and  $W_g$ ; having in view the approximation of (5.11)–(5.14) by simple finite element methods, it is of great interest to have a variational formulation of (5.11)–(5.14) containing first order derivatives, only. Such a goal is easily achieved since  $W_0$  and  $W_g$  have the alternative definitions

$$(5.20) \quad W_0 = \left\{ \{ \theta, \phi \} \in L^2(\Omega) \times H_0^1(\Omega), \int_\Omega \nabla \phi \cdot \nabla q dx = \int_\Omega \theta q dx, \forall q \in H^1(\Omega) \right\},$$

$$(5.21) \quad \begin{aligned} &W_g = \left\{ \{ \theta, \phi \} \in L^2(\Omega) \times H^1(\Omega), \phi = g_1 \text{ on } \Gamma, \right. \\ &\left. \int_\Omega \nabla \phi \cdot \nabla q dx = \int_\Omega \theta q dx + \int_\Gamma g_2 q d\Gamma, \forall q \in H^1(\Omega) \right\}, \end{aligned}$$

respectively. The equivalence between (5.18), (5.19) and (5.20), (5.21) follows easily from the *Green's formula*

$$\int_\Gamma \frac{\partial \phi}{\partial n} q d\Gamma = \int_\Omega \Delta \phi q dx + \int_\Omega \nabla \phi \cdot \nabla q dx \quad \forall q \in H^1(\Omega), \quad \forall \phi \in H^2(\Omega),$$

and the assumption that  $\Gamma (= \partial \Omega)$  is *sufficiently smooth* (or  $\Omega$  *convex*).

A variational formulation such as (5.17), (5.20), (5.21) is usually known as a *mixed variational formulation*.

**5.4. Continuation solution of problem (5.10)–(5.12).**

**5.4.1. Synopsis.** We apply now the solution methods of § 2 to the nonlinear boundary value problem (5.10)–(5.12). As parameter  $\lambda$  we choose  $\lambda = 1/\nu$ ;  $\lambda$  is directly proportional to the *Reynolds number* if we fix the boundary conditions as  $\lambda$  varies.

We consider first (in § 5.4.2) the solution of (5.10)–(5.12) via the variational formulation (5.16); the solution of (5.10)–(5.12) via (5.17) will be discussed in § 5.4.3. A mixed finite element implementation will be discussed in § 5.5.

**5.4.2. Solution of (5.10)–(5.12) via the variational formulation (5.16).** The space  $V_0 (= H_0^2(\Omega))$  which plays a fundamental role in the sequel is equipped with the inner product

$$\{v, w\} \rightarrow \int_{\Omega} \Delta v \Delta w \, dx$$

and the corresponding norm  $v \rightarrow (\int_{\Omega} |\Delta v|^2 \, dx)^{1/2}$ .

Taking  $\lambda$  as parameter the problem to be solved is

$$(5.22) \quad \Delta^2 \psi = \lambda \left( \frac{\partial \psi}{\partial x_2} \frac{\partial}{\partial x_1} \Delta \psi - \frac{\partial \psi}{\partial x_1} \frac{\partial}{\partial x_2} \Delta \psi \right) + \lambda f \quad \text{in } \Omega,$$

$$(5.23) \quad \psi = g_1 \quad \text{on } \Gamma,$$

$$(5.24) \quad \frac{\partial \psi}{\partial n} = g_2 \quad \text{on } \Gamma.$$

A variational formulation of (5.22)–(5.24) is given by:

$$(5.25) \quad \text{Find } \psi \in V_g \text{ such that } \forall \phi \in V_0$$

$$\int_{\Omega} \Delta \psi \Delta \phi \, dx = \lambda \int_{\Omega} \Delta \psi \left( \frac{\partial \psi}{\partial x_1} \frac{\partial \phi}{\partial x_2} - \frac{\partial \psi}{\partial x_2} \frac{\partial \phi}{\partial x_1} \right) dx + \lambda \int_{\Omega} f \phi \, dx.$$

*Description of the continuation procedure.* In the particular case of problem (5.22)–(5.24) the continuation techniques of §§ 2.3, 2.4 lead to the following algorithm:

(a) *Initialization.*

$$(5.26) \quad \text{Take } \lambda^0 = 0;$$

the corresponding  $\psi^0$  is the *unique* solution of the following *linear* variational problem:

$$(5.27) \quad \text{Find } \psi^0 \in V_g \text{ such that}$$

$$\int_{\Omega} \Delta \psi^0 \Delta \phi \, dx = 0 \quad \forall \phi \in V_0.$$

Problem (5.27) is in fact equivalent to the linear biharmonic problem

$$(5.28) \quad \Delta^2 \psi^0 = 0 \quad \text{in } \Omega,$$

$$\psi^0 = g_1 \quad \text{on } \Gamma, \quad \frac{\partial \psi^0}{\partial n} = g_2 \quad \text{on } \Gamma.$$

We take  $\{\psi^0, 0\}$  as the origin of the arc of solutions passing through it, and define the arc length  $s$  by

$$(5.29) \quad (\delta s)^2 = \int_{\Omega} |\Delta \delta \psi|^2 \, dx + (\delta \lambda)^2.$$

Denote  $dX/ds$  by  $\dot{X}$ ; by differentiation of (5.25) with respect to  $s$ , we obtain at  $s=0$

$$(5.30) \quad \int_{\Omega} \Delta \dot{\psi}(0) \Delta \phi \, dx = \dot{\lambda}(0) \int_{\Omega} \Delta \psi^0 \left( \frac{\partial \psi^0}{\partial x_1} \frac{\partial \phi}{\partial x_2} - \frac{\partial \psi^0}{\partial x_2} \frac{\partial \phi}{\partial x_1} \right) dx \\ + \dot{\lambda}(0) \int_{\Omega} f \phi \, dx \quad \forall \phi \in V_0, \quad \dot{\psi}(0) \in V_0.$$

We have also by definition of  $s$

$$(5.31) \quad \int_{\Omega} |\Delta \dot{\psi}(0)|^2 \, dx + \dot{\lambda}^2(0) = 1.$$

Define  $\hat{\psi}$  as the solution of the following problem

$$(5.32) \quad \hat{\psi} \in V_0, \\ \int_{\Omega} \Delta \hat{\psi} \Delta \phi \, dx = \int_{\Omega} \Delta \psi^0 \left( \frac{\partial \psi^0}{\partial x_1} \frac{\partial \phi}{\partial x_2} - \frac{\partial \psi^0}{\partial x_2} \frac{\partial \phi}{\partial x_1} \right) dx + \int_{\Omega} f \phi \, dx \quad \forall \phi \in V_0.$$

We clearly have from (5.30)–(5.32), that

$$(5.33) \quad \dot{\psi}(0) = \dot{\lambda}(0) \hat{\psi},$$

$$(5.34) \quad \dot{\lambda}(0) = \left( 1 + \int_{\Omega} |\Delta \hat{\psi}|^2 \, dx \right)^{-1/2}.$$

b) *Continuation.* With  $\Delta s (>0)$  an increment of arc length, we define for  $n \geq 0$  an approximation  $\{\psi^{n+1}, \lambda^{n+1}\} (\in V_g \times \mathbb{R})$  of  $\{\psi((n+1)\Delta s), \lambda((n+1)\Delta s)\}$  as the solution of the following nonlinear variational system:

Find  $\{\psi^{n+1}, \lambda^{n+1}\} \in V_g \times \mathbb{R}$  such that

$$(5.35a) \quad \int_{\Omega} \Delta \psi^{n+1} \Delta \phi \, dx = \lambda^{n+1} \int_{\Omega} \Delta \psi^{n+1} \left( \frac{\partial \psi^{n+1}}{\partial x_1} \frac{\partial \phi}{\partial x_2} - \frac{\partial \psi^{n+1}}{\partial x_2} \frac{\partial \phi}{\partial x_1} \right) dx \\ + \lambda^{n+1} \int_{\Omega} f \phi \, dx \quad \forall \phi \in V_0,$$

$$(5.35b) \quad \int_{\Omega} \Delta(\psi^1 - \psi^0) \Delta \dot{\psi}(0) \, dx + (\lambda^1 - \lambda^0) \dot{\lambda}(0) = \Delta s \quad \text{if } n=0,$$

$$(5.35c) \quad \int_{\Omega} \Delta(\psi^{n+1} - \psi^n) \Delta \left( \frac{\psi^n - \psi^{n-1}}{\Delta s} \right) dx + (\lambda^{n+1} - \lambda^n) \left( \frac{\lambda^n - \lambda^{n-1}}{\Delta s} \right) = \Delta s \quad \text{if } n \geq 1.$$

With  $V_0$  equipped with the inner product

$$\{v, w\} \rightarrow \int_{\Omega} \Delta v \Delta w \, dx$$

a convenient nonlinear least squares formulation of (5.35) is then:

$$(5.36) \quad \text{Find } \{\psi^{n+1}, \lambda^{n+1}\} \in V_g \times \mathbb{R} \text{ such that} \\ J_{n+1}(\psi^{n+1}, \lambda^{n+1}) \leq J_{n+1}(\chi, \mu) \quad \forall \{\chi, \mu\} \in V_g \times \mathbb{R}.$$

Here the functional  $J_{n+1}(\cdot, \cdot)$  is defined by

$$(5.37) \quad J_{n+1}(\chi, \mu) = \frac{1}{2} \int_{\Omega} |\Delta \tilde{\chi}|^2 \, dx + \frac{1}{2} |\tilde{\mu}|^2;$$

where  $\tilde{\chi}$  and  $\tilde{\mu}$  are nonlinear functions of  $\{\chi, \mu\}$ , obtained as the solutions of the *linear* problems, respectively:

$$(5.38) \quad \tilde{\chi} \in V_0, \quad \forall \phi \in V_0 \text{ we have}$$

$$\int_{\Omega} \Delta \tilde{\chi} \Delta \phi \, dx = \int_{\Omega} \Delta \chi \Delta \phi \, dx - \mu \int_{\Omega} \Delta \chi \left( \frac{\partial \chi}{\partial x_1} \frac{\partial \phi}{\partial x_2} - \frac{\partial \chi}{\partial x_2} \frac{\partial \phi}{\partial x_1} \right) dx - \mu \int_{\Omega} f \phi \, dx,$$

$$(5.39) \quad \tilde{\mu} = \int_{\Omega} \Delta(\chi - \psi^n) \Delta \left( \frac{\psi^n - \psi^{n-1}}{\Delta s} \right) dx + (\mu - \lambda^n) \left( \frac{\lambda^n - \lambda^{n-1}}{\Delta s} \right) - \Delta s.$$

Problem (5.38) is a biharmonic problem.

The least squares problem (5.36) can be solved by the conjugate gradient algorithm described in § 2.3; since the nonlinearity in (5.35a) is *quadratic* one should verify that each iteration requires the solution of “only” 3 *linear biharmonic problems* of the following type:

$$(5.40) \quad \Delta^2 w = f \quad \text{in } \Omega$$

$$w = g_1 \quad \text{on } \Gamma, \quad \frac{\partial w}{\partial n} = g_2 \quad \text{on } \Gamma.$$

*Finite element* solvers for (5.40) will be discussed in § 5.5; they are founded on the mixed variational formulation (5.17).

More details about the conjugate gradient solution of (5.36) are given in [43].

**5.4.3. Solution of (5.10)–(5.12) via the mixed variational formulation (5.17).** Using  $\lambda = 1/\nu$  as parameter, the nonlinear mixed variational problem (5.17) becomes

$$(5.41) \quad \text{Find } \{\omega, \psi\} \in W_g \text{ such that } \forall \{\theta, \phi\} \in W_0 \text{ we have}$$

$$\int_{\Omega} \omega \theta \, dx + \lambda \int_{\Omega} \omega \left( \frac{\partial \psi}{\partial x_1} \frac{\partial \phi}{\partial x_2} - \frac{\partial \psi}{\partial x_2} \frac{\partial \phi}{\partial x_1} \right) dx = \lambda \int_{\Omega} f \phi \, dx,$$

with  $W_0$  and  $W_g$  still defined by (5.20), (5.21), respectively.

*Description of the continuation procedure.* We clearly have from § 5.4.2.

(a) *Initialization.*

$$(5.42) \quad \text{Take } \lambda^0 = 0;$$

then  $\{\omega^0, \psi^0\}$  is the unique solution of the following linear mixed variational problem (equivalent to (5.27)):

$$(5.43) \quad \text{Find } \{\omega^0, \psi^0\} \in W_g \text{ such that}$$

$$\int_{\Omega} \omega^0 \theta \, dx = 0 \quad \forall \{\theta, \phi\} \in W_0.$$

We take then  $\{\{\omega^0, \psi^0\}, 0\}$  as the origin of the arc of solutions passing through it and define the arc length  $s$  by

$$(5.44) \quad (\delta s)^2 = \int_{\Omega} (\delta \omega)^2 \, dx + (\delta \lambda)^2.$$

By differentiation of (5.41) with respect to  $s$ , we obtain at  $s=0$

$$(5.45) \quad \int_{\Omega} \dot{\omega}(0)\theta \, dx = \dot{\lambda}(0) \int_{\Omega} \omega^0 \left( \frac{\partial \psi^0}{\partial x_2} \frac{\partial \phi}{\partial x_1} - \frac{\partial \psi^0}{\partial x_1} \frac{\partial \phi}{\partial x_2} \right) dx \\ + \dot{\lambda}(0) \int_{\Omega} f\phi \, dx \quad \forall \{\theta, \phi\} \in W_0, \quad \{\dot{\omega}(0), \dot{\psi}(0)\} \in W_0.$$

Since we have

$$\int_{\Omega} |\dot{\omega}(0)|^2 \, dx + \dot{\lambda}^2(0) = 1,$$

we obtain from (5.45) that

$$\dot{\lambda}(0) = \left( 1 + \int_{\Omega} |\hat{\omega}|^2 \, dx \right)^{-1/2}, \quad \{\dot{\omega}(0), \dot{\psi}(0)\} = \dot{\lambda}(0) \{\hat{\omega}, \hat{\psi}\},$$

where  $\{\hat{\omega}, \hat{\psi}\}$  is the solution of the linear mixed variational problem

$$(5.46) \quad \{\hat{\omega}, \hat{\psi}\} \in W_0, \\ \int_{\Omega} \hat{\omega}\theta \, dx = \int_{\Omega} \omega^0 \left( \frac{\partial \psi^0}{\partial x_2} \frac{\partial \phi}{\partial x_1} - \frac{\partial \psi^0}{\partial x_1} \frac{\partial \phi}{\partial x_2} \right) dx + \int_{\Omega} f\phi \, dx \quad \forall \{\theta, \phi\} \in W_0.$$

(b) *Continuation.* With  $\Delta s (>0)$  an elementary arc length we define for  $n \geq 0$  an approximation  $\{\{\omega^{n+1}, \psi^{n+1}\}, \lambda^{n+1}\} \in W_g \times \mathbb{R}$  of  $\{\{\omega((n+1)\Delta s), \psi((n+1)\Delta s)\}, \lambda((n+1)\Delta s)\}$  as the solution of the following nonlinear mixed variational system:

Find  $\{\{\omega^{n+1}, \psi^{n+1}\}, \lambda^{n+1}\} \in W_g \times \mathbb{R}$  such that

$$(5.47a) \quad \int_{\Omega} \omega^{n+1}\theta \, dx = \lambda^{n+1} \int_{\Omega} \omega^{n+1} \left( \frac{\partial \psi^{n+1}}{\partial x_2} \frac{\partial \phi}{\partial x_1} - \frac{\partial \psi^{n+1}}{\partial x_1} \frac{\partial \phi}{\partial x_2} \right) dx \\ + \lambda^{n+1} \int_{\Omega} f\phi \, dx \quad \forall \{\theta, \phi\} \in W_0,$$

$$(5.47b) \quad \int_{\Omega} (\omega^1 - \omega^0)\dot{\omega}(0) \, dx + (\lambda^1 - \lambda^0)\dot{\lambda}(0) = \Delta s \quad \text{if } n=0,$$

$$(5.47c) \quad \int_{\Omega} (\omega^{n+1} - \omega^n) \left( \frac{\omega^n - \omega^{n-1}}{\Delta s} \right) dx + (\lambda^{n+1} - \lambda^n) \left( \frac{\lambda^n - \lambda^{n-1}}{\Delta s} \right) = \Delta s \quad \text{if } n \geq 1.$$

The space  $W_0$  can be equipped with the inner product

$$\{ \{\theta_1, \phi_1\}, \{\theta_2, \phi_2\} \} \rightarrow \int_{\Omega} \theta_1 \theta_2 \, dx.$$

A convenient least squares formulation of (5.47) is then

$$(5.48) \quad \text{Find } \{\{\omega^{n+1}, \psi^{n+1}\}, \lambda^{n+1}\} \in W_g \times \mathbb{R} \text{ such that} \\ j_{n+1}(\{\omega^{n+1}, \psi^{n+1}\}, \lambda^{n+1}) \leq j_{n+1}(\{\eta, \chi\}, \mu) \quad \forall \{\{\eta, \chi\}, \mu\} \in W_g \times \mathbb{R},$$

where in (5.48) the functional  $j_{n+1}(\cdot, \cdot)$  is defined by

$$(5.49) \quad j_{n+1}(\{\eta, \chi\}, \mu) = \frac{1}{2} \int_{\Omega} |\tilde{\eta}|^2 \, dx + \frac{1}{2} |\tilde{\mu}|^2.$$

In (5.49),  $\{\tilde{\eta}, \tilde{\chi}\}$  and  $\tilde{\mu}$  are nonlinear functions of  $\eta, \chi, \mu$  obtained as the solutions of the linear problems

$$\begin{aligned} & \{\tilde{\eta}, \tilde{\chi}\} \in W_0, \\ (5.50) \quad & \forall \{\theta, \phi\} \in W_0 \text{ we have} \\ & \int_{\Omega} \tilde{\eta} \theta \, dx = \int_{\Omega} \eta \theta \, dx - \mu \int_{\Omega} \eta \left( \frac{\partial \chi}{\partial x_2} \frac{\partial \phi}{\partial x_1} - \frac{\partial \chi}{\partial x_1} \frac{\partial \phi}{\partial x_2} \right) \, dx - \mu \int_{\Omega} f \phi \, dx, \\ (5.51) \quad & \tilde{\mu} = \int_{\Omega} (\eta - \omega^n) \left( \frac{\omega^n - \omega^{n-1}}{\Delta s} \right) \, dx + (\mu - \lambda^n) \left( \frac{\lambda^n - \lambda^{n-1}}{\Delta s} \right) - \Delta s, \end{aligned}$$

respectively; problem (5.50) is equivalent to the biharmonic problem (5.38).

The conjugate gradient algorithm (2.37)–(2.46) can be applied, again, in this context, each iteration requiring, as in § 5.4.2, the solution of 3 linear biharmonic problems (see [43] for more details).

*Remark 5.1.* The main motivation of the mixed variational formulation discussed in §§ 5.3 and 5.4.3 is that it provides a convenient framework for the approximation of linear and nonlinear biharmonic problems, by very simple finite element methods like those discussed in the following section. Another application is discussed in [17]; it concerns the Von Karman equations for nonlinear plates.

**5.5. Finite element approximation.**

**5.5.1. Triangulation of  $\Omega$ . Fundamental discrete spaces.** We suppose for simplicity that  $\Omega$  is a polygonal domain of  $\mathbb{R}^2$ . With  $\mathcal{T}_h$  a triangulation of  $\Omega$  obeying the conditions given in § 3.2.1 we define the following finite-dimensional functional spaces:

$$\begin{aligned} (5.52) \quad & H_h^1 = \{v_h \mid v_h \in C^0(\bar{\Omega}), v_h|_T \in P_k, \forall T \in \mathcal{T}_h\}, \\ (5.53) \quad & H_{0h}^1 = H_h^1 \cap H_0^1(\Omega) \quad (= \{v_h \mid v_h \in H_h^1, v_h|_{\Gamma} = 0\}) \end{aligned}$$

with  $P_k \equiv$  space of polynomials in  $x_1, x_2$  of degree  $\leq k$ ;  $H_h^1$  and  $H_{0h}^1$  approximate  $H^1(\Omega)$  and  $H_0^1(\Omega)$ , respectively.

We approximate then the spaces  $W_0$  and  $W_g$  (defined by (5.20) and (5.21), respectively) by

$$\begin{aligned} (5.54) \quad & W_{0h} = \left\{ \{\theta_h, \phi_h\} \in H_h^1 \times H_{0h}^1, \int_{\Omega} \nabla \phi_h \cdot \nabla q_h \, dx = \int_{\Omega} \theta_h q_h \, dx, \forall q_h \in H_h^1 \right\}, \\ (5.55) \quad & W_{gh} = \left\{ \{\theta_h, \phi_h\} \in H_h^1 \times H_n^1, \phi_h = g_{1h} \text{ on } \Gamma, \right. \\ & \left. \int_{\Omega} \nabla \phi_h \cdot \nabla q_h \, dx = \int_{\Omega} \theta_h q_h \, dx + \int_{\Gamma} g_{2h} q_h \, d\Gamma, \forall q_h \in H_h^1 \right\}. \end{aligned}$$

Here  $g_{1h}$  and  $g_{2h}$  are convenient approximations to  $g_1$  and  $g_2$ , respectively. We observe that  $W_{0h} \not\subset W_0$ ; similarly  $W_{gh} \not\subset W_g$ , even in the simple case where  $g_{1h} = g_1, g_{2h} = g_2$ .

**5.5.2. Approximation of the Navier–Stokes equations via the  $\{\omega, \psi\}$  formulation.** Using  $\lambda = 1/\nu$  as parameter, a mixed variational formulation of the Navier–Stokes equations was given in § 5.4.3 by (5.41). We approximate (5.41) by:

$$(5.56) \quad \text{Find } \{\omega_h, \psi_h\} \in W_{gh} \text{ such that } \forall \{\theta_h, \phi_h\} \in W_{0h} \text{ we have}$$

$$\int_{\Omega} \omega_h \theta_h \, dx + \lambda \int_{\Omega} \omega_h \left( \frac{\partial \psi_h}{\partial x_1} \frac{\partial \phi_h}{\partial x_2} - \frac{\partial \psi_h}{\partial x_2} \frac{\partial \phi_h}{\partial x_1} \right) \, dx = \lambda \int_{\Omega} f_h \phi_h \, dx,$$

with  $f_h$  a convenient approximation to  $f$ . We refer to Girault–Raviart [22] for the convergence properties of  $\{\omega_h, \psi_h\}$  as  $h \rightarrow 0$ .

Concentrating on the numerical solution of problem (5.56) by *continuation least squares methods* we easily adapt the algorithms of §§ 2.3 and 5.4.3 to the solution of the approximate problem (5.56) (see Reinhart [3] for more details on the solution of (5.56) by the methods of the present paper).

In fact applying the discrete analogues of the methods described in § 5.4.3 to the solution of (5.56) requires an efficient solver for the various *discrete linear biharmonic problems* coming from the mixed finite element approximation. Such a solver is particularly required by the conjugate gradient algorithm in solving the least squares problem encountered at each step of the continuation process (we have to solve 3 linear biharmonic problems at each iteration).

### 5.5.3. On the solution of the discrete linear biharmonic problems.

**5.5.3.1. Generalities. Synopsis.** A careful examination of the algorithms discussed in § 5.4.3 shows that the *discrete linear biharmonic problems* to be solved are in fact mixed finite element approximations of biharmonic problems of the following class:

$$(5.57) \quad \begin{aligned} \Delta^2 \psi &= f_0 - \frac{\partial f_1}{\partial x_1} - \frac{\partial f_2}{\partial x_2} - \Delta f_3 \quad \text{in } \Omega, \\ \psi|_{\Gamma} &= g_1, \quad \frac{\partial \psi}{\partial n} \Big|_{\Gamma} = g_2. \end{aligned}$$

Here  $f_i \in L^2(\Omega)$ ,  $\forall i = 0, 1, 2, 3$ , and the derivatives occurring in (5.57) have to be understood in the sense of distributions. Assuming that  $g_1, g_2$  are sufficiently smooth, problem (5.57) has a unique solution in  $V_g$  (see § 5.3 for the definition of  $V_g$  and  $V_0$ ); this solution  $\psi$  is also the unique solution of the following variational problem:

$$(5.58) \quad \begin{aligned} &\text{Find } \psi \in V_g \text{ such that } \forall \phi \in V_0 \\ &\int_{\Omega} \Delta \psi \Delta \phi \, dx = \int_{\Omega} \left( f_0 \phi + f_1 \frac{\partial \phi}{\partial x_1} + f_2 \frac{\partial \phi}{\partial x_2} - f_3 \Delta \phi \right) dx. \end{aligned}$$

An equivalent *mixed variational formulation* of (5.58) is given by:

$$(5.59) \quad \begin{aligned} &\text{Find } \{\omega, \psi\} \in W_g \text{ such that } \forall \{\theta, \phi\} \in W_0 \\ &\int_{\Omega} \omega \theta \, dx = \int_{\Omega} \left( f_0 \phi + f_1 \frac{\partial \phi}{\partial x_1} + f_2 \frac{\partial \phi}{\partial x_2} + f_3 \theta \right) dx, \end{aligned}$$

where  $W_0$  and  $W_g$  are defined by (5.20) and (5.21), respectively.

Starting from the mixed formulation (5.59) we shall discuss in the following sections the finite element approximation of (5.59) and solution methods for the approximate problems.

**5.5.3.2. Finite element approximation of (5.59).** Following Ciarlet–Raviart [26] and Glowinski–Pironneau [25] we approximate (5.59) by

$$(5.60) \quad \begin{aligned} &\text{Find } \{\omega_h, \psi_h\} \in W_{gh} \text{ such that } \forall \{\theta_h, \phi_h\} \in W_{0h} \\ &\int_{\Omega} \omega_h \theta_h \, dx = \int_{\Omega} \left( f_{0h} \phi_h + f_{1h} \frac{\partial \phi_h}{\partial x_1} + f_{2h} \frac{\partial \phi_h}{\partial x_2} + f_{3h} \theta_h \right) dx, \end{aligned}$$

where  $W_{0h}$  and  $W_{gh}$  are still defined by (5.54) and (5.55), respectively, and where  $f_{ih}$  is, for  $i = 0, 1, 2, 3$ , a convenient approximation to  $f_i$ .

It is quite easy to prove that (5.60) has a unique solution; concerning the convergence of  $\{\omega_h, \psi_h\}$  to  $\{-\Delta\psi, \psi\}$  as  $h \rightarrow 0$ , it follows from Ciarlet-Raviart [26], Scholz [27] that

$$(5.61) \quad \lim_{h \rightarrow 0} \|-\Delta\psi - \omega_h\|_{L^2(\Omega)} = 0, \quad \lim_{h \rightarrow 0} \|\psi - \psi_h\|_{H^1(\Omega)} = 0$$

for all  $k \geq 1$  (in the definition of  $H_h^1$ ; cf. (5.52)). Actually the convergence result (5.61) supposes that some mild assumptions on the angles are satisfied as  $h \rightarrow 0$  (see the two above references for more details).

**5.5.3.3. Decomposition properties of the approximate problem (5.60).** We here follow and extend on some points in Glowinski-Pironneau [25].

A direct solution of (5.60) is a nontrivial task; however taking into account the very special structure of (5.60) we shall be able, *via a decomposition principle*, to reduce its solution to the solution of a family of *discrete Poisson problems which are much easier to solve*.

The starting point of our discussion is the fact that the pair  $\{\omega_h, \psi_h\}$ , solving (5.60), is characterized by the existence of  $p_h$  such that

$$(5.62a) \quad p_h \in H_h^1, \quad \int_{\Omega} \nabla p_h \cdot \nabla \phi_h \, dx = \int_{\Omega} \left( f_{0h} \phi_h + f_{1h} \frac{\partial \phi_h}{\partial x_1} + f_{2h} \frac{\partial \phi_h}{\partial x_2} \right) dx \quad \forall \phi_h \in H_{0h}^1,$$

$$(5.62b) \quad \omega_h \in H_h^1, \quad \int_{\Omega} \omega_h \theta_h \, dx = \int_{\Omega} (f_{3h} + p_h) \theta_h \, dx \quad \forall \theta_h \in H_h^1,$$

$$(5.62c) \quad \psi_h \in H_h^1, \psi_h = g_{1h} \quad \text{on } \Gamma, \quad \int_{\Omega} \nabla \psi_h \cdot \nabla q_h \, dx = \int_{\Omega} \omega_h q_h \, dx + \int_{\Gamma} g_{2h} q_h \, d\Gamma \quad \forall q_h \in H_h^1.$$

To prove the characterization (5.62) we observe that (5.60) is equivalent to the minimization problem:

$$(5.63) \quad \text{Find } \{\omega_h, \psi_h\} \in W_{gh} \text{ such that} \\ j_h(\omega_h, \psi_h) \leq j_h(\theta_h, \phi_h) \quad \forall \{\theta_h, \phi_h\} \in W_{gh},$$

where

$$(5.64) \quad j_h(\theta_h, \phi_h) = \frac{1}{2} \int_{\Omega} \theta_h^2 \, dx - \int_{\Omega} \left( f_{0h} \phi_h + f_{1h} \frac{\partial \phi_h}{\partial x_1} + f_{2h} \frac{\partial \phi_h}{\partial x_2} + f_{3h} \theta_h \right) dx.$$

Hence  $p_h$  appears as a *Lagrange multiplier* for the *linear equality constraints* satisfied by  $\{\omega_h, \psi_h\}$  in (5.62c) (and in the definition of  $W_{gh}$ ; see (5.55)).

To go further into the decomposition properties we introduce a space  $\mathcal{M}_h$  with the following properties:

$$(5.65) \quad \mathcal{M}_h \text{ is a complementary space (not precisely defined for the moment) of } H_{0h}^1 \text{ in } H_h^1, \text{ i.e. } \mathcal{M}_h \subset H_h^1 \text{ and } H_{0h}^1 \oplus \mathcal{M}_h = H_h^1.$$

It follows from (5.65) that the bilinear form  $\mathcal{M}_h \times \mathcal{M}_h \rightarrow \mathbb{R}$  defined by

$$\{\lambda_h, \mu_h\} \rightarrow \int_{\Gamma} \lambda_h \mu_h \, d\Gamma$$

is a *scalar product* over  $\mathcal{M}_h$ . The key step is in fact to introduce a bilinear form  $a_h: \mathcal{M}_h \times \mathcal{M}_h \rightarrow \mathbb{R}$ , defined as follows:

Let  $\lambda_h \in \mathcal{M}_h$  and let  $p_h$ , respectively  $\psi_h$ , be the solutions of the following approximate problems:

$$(5.66a) \quad \int_{\Omega} \nabla p_h \cdot \nabla \phi_h \, dx = 0 \quad \forall \phi_h \in H_{0h}^1, \quad p_h \in H_h^1, \quad p_h - \lambda_h \in H_{0h}^1,$$

$$(5.66b) \quad \int_{\Omega} \nabla \psi_h \cdot \nabla \phi_h \, dx = \int_{\Omega} p_h \phi_h \, dx \quad \forall \phi_h \in H_{0h}^1, \quad \psi_h \in H_{0h}^1.$$

Then we define the bilinear form  $a_h(\cdot, \cdot)$  by

$$(5.66c) \quad a_h(\lambda_h, \mu_h) = \int_{\Omega} p_h \mu_h \, dx - \int_{\Omega} \nabla \psi_h \cdot \nabla \mu_h \, dx \quad \forall \mu_h \in \mathcal{M}_h.$$

It then follows from [25, § 3.5] that  $a_h(\cdot, \cdot)$  is *symmetric* and *positive definite*.

*Application to the decomposition of the approximate problem* (5.60). Let  $\{\omega_h, \psi_h\}$  be the solution of (5.60) and let  $\lambda_h$  be the component in  $\mathcal{M}_h$  of the function  $p_h$  occurring in the characterization (5.62). Let  $\bar{p}_h, \bar{\psi}_h$  be the solutions of

$$(5.67) \quad \int_{\Omega} \nabla \bar{p}_h \cdot \nabla \phi_h \, dx = 0 \quad \forall \phi_h \in H_{0h}^1, \quad \bar{p}_h - \lambda_h \in H_{0h}^1,$$

$$(5.68) \quad \int_{\Omega} \nabla \bar{\psi}_h \cdot \nabla \phi_h \, dx = \int_{\Omega} \bar{p}_h \phi_h \, dx \quad \forall \phi_h \in H_{0h}^1, \quad \bar{\psi}_h \in H_{0h}^1.$$

Let  $p_{0h}$  and  $\psi_{0h}$  be the solutions of

$$(5.69) \quad \int_{\Omega} \nabla p_{0h} \cdot \nabla \phi_h \, dx = \int_{\Omega} \left( f_{0h} \phi_h + f_{1h} \frac{\partial \phi_h}{\partial x_1} + f_{2h} \frac{\partial \phi_h}{\partial x_2} \right) dx \quad \forall \phi_h \in H_{0h}^1, \quad p_{0h} \in H_{0h}^1,$$

$$(5.70) \quad \int_{\Omega} \nabla \psi_{0h} \cdot \nabla \phi_h \, dx = \int_{\Omega} (p_{0h} + f_{3h}) \phi_h \, dx \quad \forall \phi_h \in H_{0h}^1, \quad \psi_h \in H_h^1, \quad \psi_{0h} = g_{1h} \quad \text{on } \Gamma.$$

We clearly have  $p_h = \bar{p}_h + p_{0h}$  and  $\psi_h = \bar{\psi}_h + \psi_{0h}$ .

We shall now show that  $\lambda_h$  is the solution of a variational problem in  $\mathcal{M}_h$ .

**THEOREM 5.1.** *Let  $\{\omega_h, \psi_h\}$  be the solution of (5.60) and let  $\lambda_h$  be the component in  $\mathcal{M}_h$  of  $p_h$  defined from  $\{\omega_h, \psi_h\}$  by (5.62). Then  $\lambda_h$  is the unique solution of the linear variational problem*

$$(5.71) \quad a_h(\lambda_h, \mu_h) = \int_{\Omega} \nabla \psi_{0h} \cdot \nabla \mu_h \, dx - \int_{\Omega} (p_{0h} - f_{3h}) \mu_h \, dx - \int_{\Gamma} g_{2h} \mu_h \, d\Gamma \quad \forall \mu_h \in \mathcal{M}_h, \quad \lambda_h \in \mathcal{M}_h$$

which is equivalent to a linear system with a positive definite matrix.

*Proof.* We have from (5.66)–(5.68) that

$$\begin{aligned}
 a_h(\lambda_h, \mu_h) &= \int_{\Omega} \bar{p}_h \mu_h \, dx - \int_{\Omega} \nabla \bar{\psi}_h \cdot \nabla \mu_h \, dx \\
 &= \int_{\Omega} (p_h - p_{0h}) \mu_h \, dx - \int_{\Omega} \nabla (\psi_h - \psi_{0h}) \cdot \nabla \mu_h \, dx \\
 (5.72) \quad &= \int_{\Omega} \nabla \psi_{0h} \cdot \nabla \mu_h \, dx - \int_{\Omega} (p_{0h} + f_{3h}) \mu_h \, dx \\
 &\quad - \left( \int_{\Omega} \nabla \psi_h \cdot \nabla \mu_h \, dx - \int_{\Omega} (p_h + f_{3h}) \mu_h \, dx \right) \quad \forall \mu_h \in \mathcal{M}_h.
 \end{aligned}$$

But from (5.62b, c) we have

$$\begin{aligned}
 \int_{\Omega} \nabla \psi_h \cdot \nabla \mu_h \, dx - \int_{\Omega} (p_h + f_{3h}) \mu_h \, dx &= \int_{\Omega} \nabla \psi_h \cdot \nabla \mu_h \, dx - \int_{\Omega} \omega_h \mu_h \, dx \\
 &= \int_{\Omega} g_{2h} \mu_h \, d\Gamma \quad \forall \mu_h \in \mathcal{M}_h,
 \end{aligned}$$

which, together with (5.72), proves (5.71). The uniqueness is obvious since  $a_h(\cdot, \cdot)$  is positive definite. The equivalence with a positive definite linear system is a classical result in the approximation of linear variational problems.  $\square$

*Remark 5.2.* To compute the right-hand side of (5.71) it is necessary to solve the two approximate Dirichlet problems (5.69) and (5.70). Similarly if  $\lambda_h$  is known, to compute  $p_h$ ,  $\omega_h$  and  $\psi_h$  it is necessary to solve

$$\begin{aligned}
 (5.73) \quad &p_h \in H_h^1, p_h - \lambda_h \in H_{0h}^1, \\
 &\int_{\Omega} \nabla p_h \cdot \nabla \phi_h \, dx = \int_{\Omega} \left( f_{0h} \phi_h + f_{1h} \frac{\partial \phi_h}{\partial x_1} + f_{2h} \frac{\partial \phi_h}{\partial x_2} \right) dx \quad \forall \phi_h \in H_{0h}^1,
 \end{aligned}$$

$$\begin{aligned}
 (5.74) \quad &\omega_h \in H_h^1, \\
 &\int_{\Omega} \omega_h \theta_h \, dx = \int_{\Omega} (f_{3h} + p_h) \theta_h \, dx \quad \forall \theta_h \in H_h^1,
 \end{aligned}$$

$$\begin{aligned}
 (5.75) \quad &\psi_h \in H_h^1, \psi_h = g_{1h} \text{ on } \Gamma, \\
 &\int_{\Omega} \nabla \psi_h \cdot \nabla \phi_h \, dx = \int_{\Omega} (p_h + f_{3h}) \phi_h \, dx \quad \forall \phi_h \in H_{0h}^1,
 \end{aligned}$$

i.e. two discrete Dirichlet problems, (5.73) and (5.75), and (5.74) which is a much simpler linear problem ( $\omega_h$  is in fact the  $L^2$ -projection on  $H_h^1$  of the function  $p_h + f_{3h}$ ).

*Recapitulation.* It has been shown that solving the discrete biharmonic problem (5.60) is equivalent to solving (5.69), (5.70), (5.71), (5.73), (5.74), (5.75) *sequentially*. Problems (5.69), (5.70), (5.73), (5.75) are discrete Dirichlet problems, for the operator  $-\Delta$ , for which very efficient direct or iterative solvers exist. The variational problem (5.74) is even simpler to solve, since the matrix of the equivalent linear system is very sparse, has a condition number  $O(1)$  and is in fact an approximation to the identity operator. Finally the only nonstandard step is the solution of the variational problem (5.71) which is discussed in the following § 5.5.3.4.

**5.5.3.4. Solution of problem (5.71).** Several methods for the solution of (5.71) have been discussed in [25, §§ 4 and 5]. Let us mention among them a *conjugate gradient method* which yields a solution algorithm for the discrete biharmonic problem (5.60);

the cost per iteration is essentially the solution of *two* discrete Dirichlet problems for the operator  $-\Delta$ ; numerical experiments show a convergence in  $O(N_h^{1/2})$  iterations, where  $N_h = \dim \mathcal{M}_h$ . We find also in [25, § 4] a detailed analysis of a direct method for solving (5.71) requiring the construction of the symmetric, positive definite (and full) matrix  $A_h$  of the linear system equivalent to (5.71). In fact one does not construct  $A_h$ , but (using the *Cholesky factorization* method) a lower triangular-regular matrix  $L_h$  such that  $A_h = L_h L_h^T$ ; since the construction of  $L_h$  requires (cf. [25, § 4]) the solution of  $2N_h$  discrete Dirichlet problems it seems preferable to use the conjugate gradient algorithm. However in practice we prefer direct solvers for the following reasons:

(i) Since the  $2N_h$  discrete Dirichlet problems mentioned above have all the same matrix which is symmetric and positive definite, a Cholesky factorization done once and for all will result in an important saving of computational time.

(ii) If a large number of linear discrete biharmonic problems have to be solved—as in time dependent problems or during an iterative process like those discussed in this paper—the solution method of (5.60), founded on the construction of  $L_h$  offers (from our numerical experiments) a more economical strategy than the conjugate gradient algorithms discussed in [25, § 5].

The above comments justify the choice of the direct solution of (5.71) for the numerical experiments described in § 5.6.

We have given in [43] the description of a *new conjugate gradient algorithm with scaling* (i.e. preconditioning). If the speed of convergence is measured in number of iterations, the new algorithm is faster than those discussed in [25, § 5]. However the new algorithm requires the solution of *three* discrete Dirichlet problems instead of *two*, for each iteration, as in the algorithm (5.26)–(5.33) [25, pp. 197–198].

*Remark 5.3 (On the choice of  $\mathcal{M}_h$ ).* Suppose that  $H_h^1$  is composed of ordinary Lagrangian finite elements of order  $k$  ( $k = 1, 2$  in most applications). It follows then from [25] (for which we refer for more details) that the best choice for  $\mathcal{M}_h$  is given by

$$(5.76) \quad \mathcal{M}_h = \{\mu_h \mid \mu_h \in H_h^1, \mu_h|_T = 0 \quad \forall T \in \mathcal{T}_h \text{ such that } \partial T \cap \Gamma = \emptyset\}.$$

With such a choice the elements of  $\mathcal{M}_h$  are completely determined by the values attained at those nodes of  $\mathcal{T}_h$  belonging to  $\Gamma$ . Thus we should take as basis functions for  $\mathcal{M}_h$  those basis functions of  $H_h^1$  associated with the boundary nodes (again, see [25] for more details).

**5.6. Numerical experiments.**

**5.6.1. Formulation of the test problem.** With  $\Omega = ]0, 1[ \times ]0, 1[$  we consider the following Navier–Stokes test problem:

$$(5.77) \quad \begin{aligned} -\nu \Delta \mathbf{u} + (\mathbf{u} \cdot \nabla) \mathbf{u} + \nabla p &= 0 \quad \text{in } \Omega, \\ \nabla \cdot \mathbf{u} &= 0 \quad \text{in } \Omega, \\ \mathbf{u}(x_1, x_2)|_\Gamma &= \begin{cases} \{1, 0\} & \text{if } x_2 = 1, \\ \{0, 0\} & \text{if } 0 \leq x_2 < 1. \end{cases} \end{aligned}$$

Hence problem (5.77) is the classical *driven cavity* problem. The corresponding  $\{\omega, \psi\}$  formulation is

$$(5.78) \quad \begin{aligned} \nu \Delta \omega + \left( \frac{\partial \psi}{\partial x_2} \frac{\partial \omega}{\partial x_1} - \frac{\partial \psi}{\partial x_1} \frac{\partial \omega}{\partial x_2} \right) &= 0 \quad \text{in } \Omega, \\ -\Delta \psi &= \omega \quad \text{in } \Omega, \\ \psi &= 0 \text{ on } \Gamma; \frac{\partial \psi}{\partial n}(x_1, x_2)|_\Gamma = \begin{cases} 1 & \text{if } x_2 = 1, \\ 0 & \text{if } 0 \leq x_2 < 1. \end{cases} \end{aligned}$$

**5.6.2. Triangulation of  $\Omega$ .** The triangulation  $\mathcal{T}_h$  used to approximate (5.77), (5.78) by the methods of § 5.5, is shown on Fig. 5.1. It contains 800 triangles and since *piecewise quadratic* elements are used (i.e.  $k=2$  in (5.52)), it corresponds to 160 boundary nodes and 1581 interior nodes (vertices and midpoints); we have therefore a nonlinear system of about 3300 unknowns to solve after discretization.

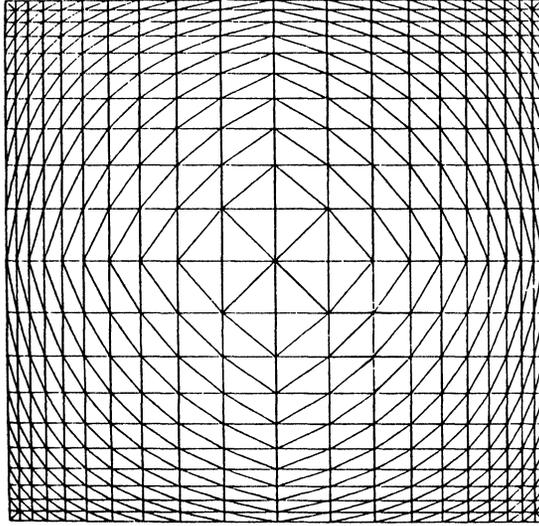


FIG. 5.1

**5.6.3. Numerical results—further comments.** The numerical procedure described in § 5.4.3 has been applied to the solution of the approximate problem (5.56) associated with (5.78) (using  $\lambda = 1/\nu = \text{Re}$ ). The computations have been done on a CRAY-1 computer, with special vectorized subroutines (in particular every subroutine concerning profile matrices (product, Cholesky factorization, resolution of triangular linear systems) has been vectorized).

We have used  $\Delta s = 100$  for  $0 \leq \lambda \leq 1400$ ,  $\Delta s = 200$  for  $1400 \leq \lambda \leq 2600$ ,  $\Delta s = 400$  for  $\lambda = 3000$ .

The conjugate gradient iterations were stopped as soon as the least squares cost functional was less than  $10^{-5}$ . The computations have been done with double precision variables.

Figures 5.2, 5.3, 5.4, 5.5, 5.6 show the variations of the least squares cost functional as a function of the number of conjugate gradient iterations, for  $\lambda = 100, 400, 1600, 2000, 3000$ , respectively; as expected the number of iterations necessary to obtain the convergence is an increasing function of  $\lambda$  ( $= \text{Re}$ ).

For  $\text{Re} = 3000$ , the average CPU for one iteration of conjugate gradient is about 0.9 second on the CRAY-1 computer.

The stream lines for  $\text{Re} = 100, 400, 1600, 2000, 3000$  are shown in Figs. 5.7–5.11 respectively. The values of the stream function along the lines are:

$$\begin{aligned} \psi &= -0.12, -0.1, -0.08, -0.06, -0.04, -0.02, 0.0 \\ &= 0.0025, 0.001, 0.0005, 0.0001, 0.00005. \end{aligned}$$

Even for small values of the Reynolds number, there appear two secondary vortices in the lower upstream and downstream corners. These vortices grow larger as the Reynolds number increases. For values of  $\text{Re}$  beyond 1500, a third secondary vortex

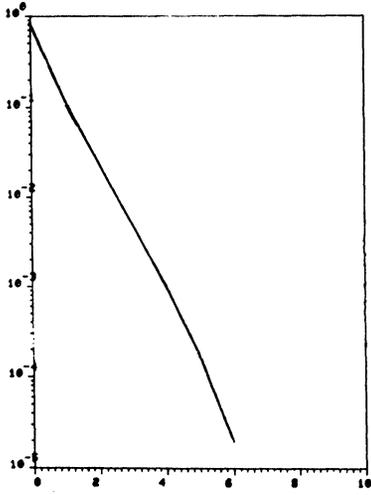


FIG. 5.2.  $Re = 100$ .

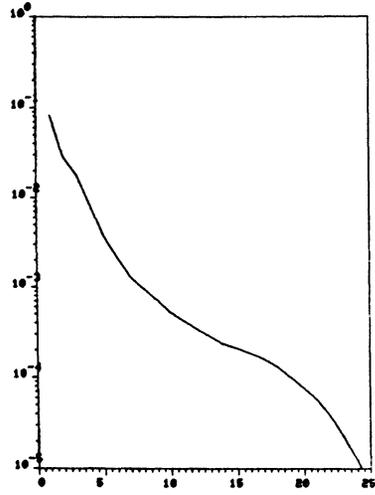


FIG. 5.3.  $Re = 400$ .

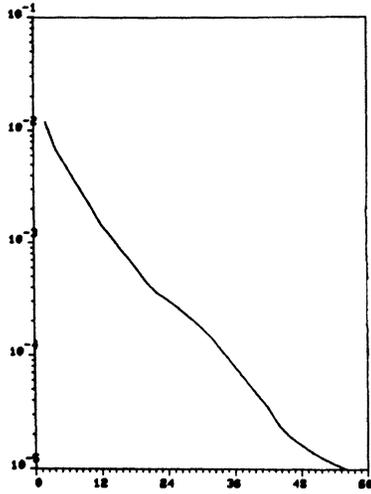


FIG. 5.4.  $Re = 1600$ .

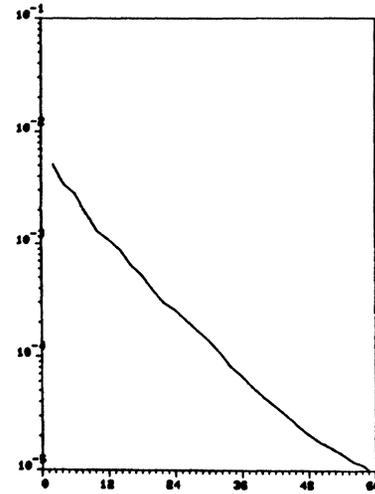


FIG. 5.5.  $Re = 2000$ .

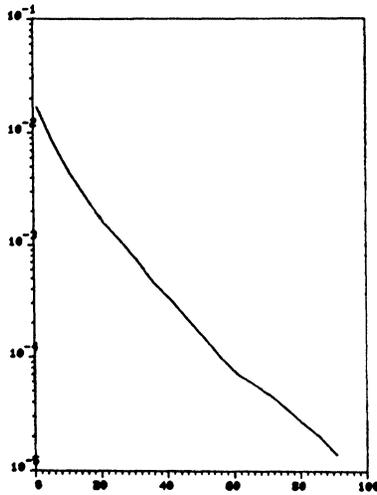
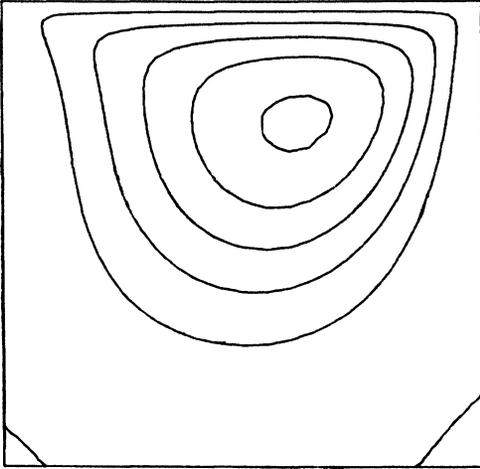
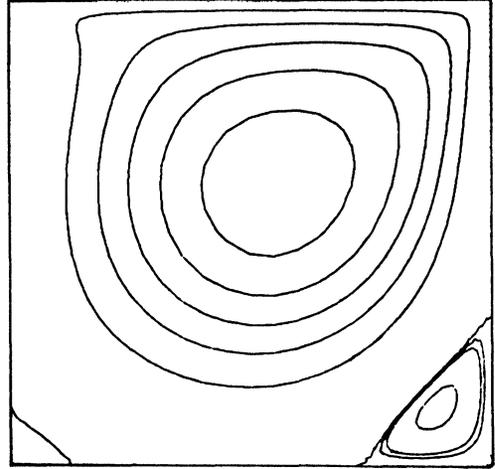
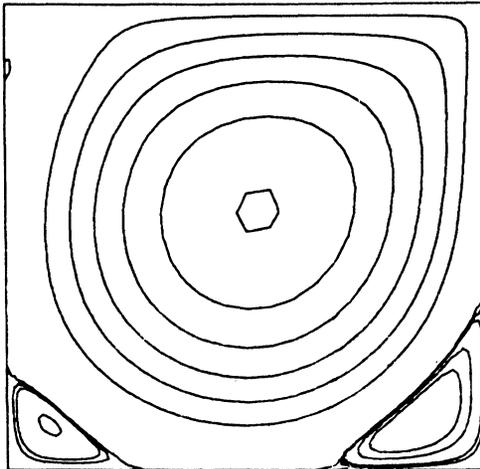
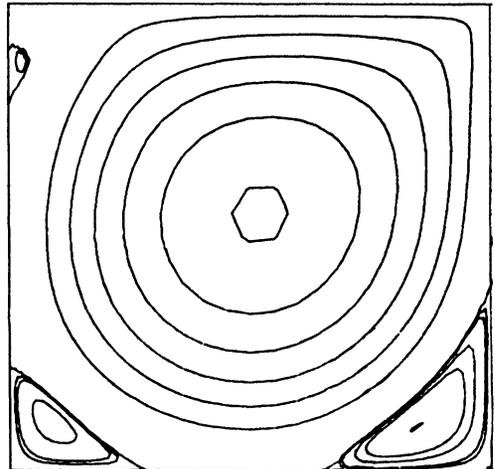
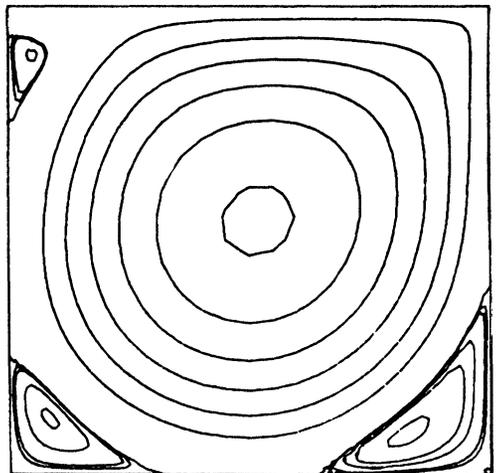


FIG. 5.6.  $Re = 3000$ .

FIG. 5.7.  $Re=100$ .FIG. 5.8.  $Re=400$ .FIG. 5.9.  $Re=1600$ .FIG. 5.10.  $Re=2000$ .FIG. 5.11.  $Re=3000$ .

appeared in the upper, upstream corner. These qualitative results agree with the numerical tests done by Olson–Tuann [29] using other finite element methods, by Schreiber–Keller [35] using continuation and finite difference methods and by Winters–Cliff [44] using finite elements and refinements in the corners.

A most interesting question is the possible occurrence of multiple solutions as the Reynolds number increases beyond some critical value. So far, we did not observe such behavior in the range of  $Re$  that we considered in our computations, i.e.  $0 \leq Re \leq 3000$ . Actually and to our knowledge the computed solutions obtained in the range  $0 \leq Re \leq 5000$  by various authors using different methods agree quite well; this observation suggests that multiple solutions can only appear for greater values of  $Re$ . Nevertheless it would be interesting to refine the numerical techniques in order to detect such a behavior.

**6. Conclusion.** We have discussed in this paper the solution of nonlinear boundary value problems containing a parameter by a combination of arc length continuation methods, least squares—conjugate gradient algorithms and finite element approximations. The resulting methodology is quite general and has been applied to the solution of second order and fourth order nonlinear boundary value problems whose branches of solutions may exhibit limit points and bifurcation.

**Acknowledgments.** The authors would like to thank Professor R. B. Simpson and the referees for most helpful comments and suggestions.

#### REFERENCES

- [1] E. POLAK, *Computational Methods in Optimization*, Academic Press, New York, 1971.
- [2] M. J. D. POWELL, *Restart procedure for the conjugate gradient method*, Math. Programming, 12 (1977), pp. 148–162.
- [3] L. REINHART, *Sur la résolution numérique de problèmes aux limites non linéaires par des méthodes de continuation*, Thèse de 3ème Cycle, Université Pierre et Marie Curie, Paris VI, June 1980.
- [4] M. S. BERGER, *Nonlinearity and Functional Analysis*, Academic Press, New York, 1977.
- [5] H. B. KELLER, *Numerical solution of bifurcation and nonlinear eigenvalue problems*, in Applications of Bifurcation Theory, P. Rabinowitz, ed., Academic Press, New York, 1977, pp. 359–384.
- [6] ———, *Global homotopies and Newton methods*, in Recent Advances in Numerical Analysis, C. de Boor and G. H. Golub, eds., Academic Press, New York, 1978, pp. 73–94.
- [7] J. L. LIONS, *Quelques méthodes de résolution des problèmes aux limites non linéaires*, Dunod, Gauthier-Villars, Paris, 1969.
- [8] M. G. CRANDALL AND P. H. RABINOWITZ, *Bifurcation, perturbation of single eigenvalues and linearized stability*, Arch. Rat. Mech. Anal., 52 (1973), pp. 161–180.
- [9] ———, *Some continuation and variational methods for positive solutions of nonlinear elliptic eigenvalue problems*, Arch. Rat. Mech. Anal., 58 (1975), pp. 207–218.
- [10] H. AMANN, *Fixed point equations and nonlinear eigenvalue problems in ordered Banach spaces*, SIAM Rev., 18 (1976), pp. 620–709.
- [11] F. MIGNOT AND J. P. PUEL, *Sur une classe de problèmes non linéaires avec non linéarité positive, croissante, convexe*, Comptes Rendus du Congrès d'Analyse Non Linéaire, Rome, May 1978, Pitagora Editrice, Bologna, 1979, pp. 45–72.
- [12] F. MIGNOT, F. MURAT AND J. P. PUEL, *Variation d'un point de retournement en fonction du domaine*, Comm. Partial Differential Equations, 4 (1979), pp. 1263–1297.
- [13] F. KIKUCHI, *Finite element approximations to bifurcation problems of turning point type*, in Computing Methods in Applied Sciences and Engineering, 1977, Part 1, R. Glowinski and J. L. Lions, eds., Lecture Notes in Mathematics 704, Springer-Verlag, Berlin, 1979, pp. 252–266.
- [14] R. B. SIMPSON, *A method for the numerical determination of bifurcation states of nonlinear systems of equations*, SIAM J. Numer. Anal., 12 (1975), pp. 439–451.
- [15] G. MOORE AND A. SPENCE, *The calculation of turning points of nonlinear equations*, SIAM J. Numer. Anal., 17 (1980), pp. 567–576.

- [16] T. F. CHAN AND H. B. KELLER, *Arclength continuation and multigrid techniques for nonlinear eigenvalue problems*, this Journal, 3 (1982), pp. 173-194.
- [17] L. REINHART, *On the numerical analysis of the Von Karman equations: mixed finite element approximation and continuation techniques*, Numer. Math., 39 (1982), pp. 371-404.
- [18] M. G. CRANDALL AND P. H. RABINOWITZ, *Bifurcation from simple eigenvalues*, J. Funct. Anal., 8 (1971), pp. 321-340.
- [19] F. BREZZI, J. RAPPAZ AND P. A. RAVIART, *Finite dimensional approximation of nonlinear problems. Part III: Simple bifurcation points*, Numer. Math., 38 (1981), pp. 1-30.
- [20] O. A. LADYZHENSKAYA, *The Mathematical Theory of Viscous Incompressible Flows*, Gordon and Breach, New York, 1969.
- [21] R. TEMAM, *Navier-Stokes Equations*, North-Holland, Amsterdam, 1977.
- [22] V. GIRAULT AND P. A. RAVIART, *Finite element approximation of the Navier-Stokes equations*, Lecture Notes in Mathematics 749, Springer-Verlag, Berlin, 1979.
- [23] R. RAUTMANN, ed., *Approximation Methods for Navier-Stokes Equations*, Lecture Notes in Mathematics 771, Springer-Verlag, Berlin, 1980.
- [24] F. THOMASSET, *Implementation of Finite Element Methods for Navier-Stokes Equations*, Springer-Verlag, New York, 1981.
- [25] R. GLOWINSKI AND O. PIRONNEAU, *Numerical methods for the first biharmonic equation and for the two-dimensional Stokes problem*, SIAM Rev., 17 (1979), pp. 167-212.
- [26] P. G. CIARLET AND P. A. RAVIART, *A mixed finite element method for the biharmonic equation*, in Mathematical Aspects of Finite Element Methods in Partial Differential Equations, C. de Boor, ed., Academic Press, New York, 1974, pp. 125-145.
- [27] R. SCHOLZ, *A mixed method for 4th order problems using linear finite elements*, Rev. Française Autom. Inf. Rech. Op., Anal. Num., 11 (1977), pp. 197-208.
- [28] D. F. SHANNO, *Conjugate gradient methods with inexact searches*, Math. Oper. Res., 13 (1978), pp. 244-255.
- [29] M. D. OLSON AND S. Y. TUANN, *Further finite element results for the square cavity*, Proc. Third International Conference on Finite Elements in Flow Problems, Banff, Alberta, Canada, 10-13 June, 1980, Volume 1, D. H. Norrie, ed., University of Calgary, pp. 143-152.
- [30] H. B. KELLER AND D. S. COHEN, *Some positive problems suggested by nonlinear heat generation*, J. Math. Mech., 16 (1967), pp. 1361-1376.
- [31] J. P. KEENER AND H. B. KELLER, *Positive solutions of convex nonlinear eigenvalue problems*, J. Diff. Eqs., 16 (1974), pp. 103-125.
- [32] ———, *Perturbed bifurcation theory*, Arch. Rat. Mech. Anal., 50 (1973), pp. 159-175.
- [33] B. MATKOWSKY AND E. L. REISS, *Singular perturbations of bifurcations*, SIAM J. Appl. Math., 33 (1977), pp. 230-255.
- [34] D. PEROZZI, *Thesis, Part II: Analysis of optimal step size selection in homotopy and continuation methods*, California Institute of Technology, Applied Math., 1980, pp. 82-156.
- [35] R. SCHREIBER AND H. B. KELLER, *Driven cavity flows by efficient numerical techniques*, J. Comp. Phys., 40 (1983), pp. 310-333.
- [36] H. B. KELLER AND W. F. LANGFORD, *Iterations, perturbations and multiplicities for nonlinear bifurcation problems*, Arch. Rat. Mech. Anal., 48 (1972), pp. 83-108.
- [37] W. RHEINBOLDT, *Solution field of nonlinear equations and continuation methods*, SIAM J. Numer. Anal., 17 (1980), pp. 221-237.
- [38] P. DEUFLHARD, *A stepsize control for continuation methods and its special application to multiple shooting techniques*, Numer. Math., 33 (1979), pp. 115-146.
- [39] H. B. KELLER, *Practical procedures in path following near limit points*, in Computing Methods in Applied Science and Engineering V, R. Glowinski and J. L. Lions, eds., North-Holland, Amsterdam, 1982, pp. 177-183.
- [40] D. W. DECKER AND H. B. KELLER, *Multiple limit point bifurcation*, J. Math. Anal. Appl., 75 (1980), pp. 417-430.
- [41] W. C. RHEINBOLDT, *Solution fields of nonlinear equations and continuation methods*, SIAM J. Numer. Anal., 17 (1980), pp. 221-237.
- [42] R. GLOWINSKI, *Numerical Methods for Nonlinear Variational Problems*, Springer, New York, 1984.
- [43] R. GLOWINSKI, H. B. KELLER AND L. REINHART, *Continuation-conjugate gradient methods for the least square solution of nonlinear boundary value problems*, Rapport de Recherche INRIA No. 141, June 1982.
- [44] K. H. WINTERS AND K. A. CLIFFE, *A finite element study of driven laminar flow in a square cavity*, AERE Report R. 9444, 1979.