DEPARTMENT OF MATHEMATICS

ON THE CONVERGENCE OF THE RITZ-GALERKIN TECHNIQUE FOR SOLVING TWO-DIMENSIONAL OPTIMAL SHAPE PROBLEMS

S. LYLE and N.K. NICHOLS

Numerical Analysis Report no. 15/89

UNIVERSITY OF READING

ON THE CONVERGENCE OF THE RITZ-GALERKIN TECHNIQUE FOR SOLVING TWO-DIMENSIONAL OPTIMAL SHAPE PROBLEMS

Numerical Analysis Report no. 15/89

S. Lyle

School of Mathematics

Kingston Polytechnic

Penrhyn Road

Kingston upon Thames

KT1 2EE

N.K. Nichols

Department of Mathematics

University of Reading

P.O. Box 220

Whiteknights

Reading RG6 2AX

1. Introduction

The last two decades have seen a growing interest in numerical solution of optimal control problems. These are essentially constrained optimization problems where a local minimum of a given cost functional is sought subject to a set of differential equations and possibly some further integral and inequality constraints. While the early efforts have been centered around the solution of linear problems with quadratic cost functional by a variety of numerical techniques, more recently the work has concentrated on non-linear problems [1], [2], [3], [5], [9], [13].

The variational nature of optimal control problems make them particularly suitable for the application of finite element methods. Bosarge et al [3] consider the use of the Ritz-Galerkin method in one-dimensional problems with fixed final time.

In this paper we extend the technique described in [3] to the two-dimensional case. We treat the problem as a standard Calculus of Variations problem and reduce it to unconstrained form using the method of Lagrange. The Ritz-Galerkin finite dimensional approximation is then derived and the local convergence of the approximation is proved under certain reasonable smoothness assumptions. We then consider the application of the technique to optimal design problems. These are distributed parameter problems in which the shape of the boundary is not known and has to be determined as part of the problem. We introduce a transformation which maps the unknown region onto a unit square and then apply the Ritz-Galerkin approximation. The method is finally demonstrated on a simple test example.

2. Statement of the Problem

We consider the following fixed boundary problem:

Find the minimum of the functional:

$$I(\underline{u}) = \iint_R g(x,y,\underline{u})dxdy , \tag{2.1}$$

subject to a set of first order, non-linear, partial differential equations:

$$\underline{f}(x,y,\underline{u},\underline{u}_x,\underline{u}_y) = \underline{0} , \tag{2.2}$$

and the boundary conditions:

$$\underline{u} = \underline{0} \quad \text{on} \quad \Gamma_1 , \tag{2.3}$$

where $\underline{u}(x,y) \in U$ is an n-dimensional vector, $\underline{f}(x,y,\underline{u},\underline{u}_x,\underline{u}_y)$ is an r-dimensional vector and $g(x,y,\underline{u})$ is a scalar-valued function. We assume that the domain R is bounded by the piecewise smooth, closed curve $\Gamma = \Gamma_1 \cup \Gamma_2$, assumed known. The space U is chosen to be the Sobolev space consisting of all n-vector valued functions $\underline{z}$ defined on R, such that $\underline{z}$ and its partial derivatives of order $\alpha + 1$, $(\alpha \geq 1)$, with respect to x and y are square integrable, that is

$$U = \{\underline{z} | \{\underline{z}, \frac{\partial^{i+j}\underline{z}}{\partial x^i \partial y^j} \ (i+j \leq \alpha+1)\} \subset \{L_2[R], E^n\}\} .$$

In subsequent analysis we shall use the following norms:

$$\|\underline{z}\|_2^2 \equiv \|\underline{z}\|_{2,0}^2 = \sum_{i=1}^{n} \iint_R z_i^2 \, dxdy , \tag{2.4}$$

and

$$\|\underline{z}\|^2_{2,1} = \sum_{i=1}^{n} \iint_R [z_i^2 + z_{ix}^2 + z_{iy}^2]dxdy . \quad (2.5)$$

The problem (2.1)-(2.3) can be treated as a standard Calculus of Variation problem. We introduce vector-valued Lagrange multipliers $\underline{\lambda}(x,y)$ and define the Lagrangian $L[\underline{u},\underline{\lambda}]$ by:

$$L[\underline{u},\underline{\lambda}] = \iint_R \{g(x,y,\underline{u}) + \underline{f}^T(x,y,\underline{u},\underline{u}_x,\underline{u}_y)\underline{\lambda}(x,y)\}dxdy \quad (2.6)$$

where

$$\underline{\lambda} \in \Lambda = \{\underline{z}|\underline{z}, \frac{\partial^{i+j}\underline{z}}{\partial x^i \partial y^j} (i+j \leq \alpha+1)\} \subset \{L_2[R], E^r\}\} .$$

The problem can now be reformulated in the following way:

Given the functional (2.1) and the system (2.2) find the optimal quantities $\underline{u}^*$ and $\underline{\lambda}^*$ such that:

$$L[\underline{u}^*,\underline{\lambda}^*] = \sup_{\underline{\lambda}\in\Lambda} \inf_{\underline{u}\in U} L[\underline{u},\underline{\lambda}] . \quad (2.7)$$

The problem is also equivalent to the following one:

Find the optimal quantities $\underline{u}^*$ and $\underline{\lambda}^*$ such that partial Fréchet derivatives:

$$\frac{\partial L}{\partial \underline{u}} [\underline{u}^*,\underline{\lambda}^*] = \underline{0}$$

$$\frac{\partial L}{\partial \underline{\lambda}} [\underline{u}^*,\underline{\lambda}^*] = \underline{0} , \quad (2.8)$$

simultaneously.

The existence and uniqueness of a solution to the original problem is assumed and the equivalence of the three formulations can be proved using standard variational arguments.

Let

$$\underline{\rho}(\underline{u},\underline{\lambda}) \;=\; \frac{\partial g}{\partial \underline{u}} + \frac{\partial \underline{f}}{\partial \underline{u}}\,\underline{\lambda} - \frac{\partial}{\partial x}\left[\frac{\partial \underline{f}}{\partial \underline{u}_x}\,\underline{\lambda}\right] - \frac{\partial}{\partial y}\left[\frac{\partial \underline{f}}{\partial \underline{u}_y}\,\underline{\lambda}\right] , \tag{2.9}$$

and let

$$\begin{aligned} \underline{f}(x,y,\underline{u}^*,\underline{u}_x^*,\underline{u}_y^*) &\equiv \underline{f}^* \\ g(x,y,\underline{u}^*) &\equiv g^* . \end{aligned} \tag{2.10}$$

All three formulations of the problem give the familiar set of necessary conditions:

$$\underline{\rho}^* \equiv \underline{\rho}(\underline{u}^*,\underline{\lambda}^*) \;=\; \underline{0} \tag{2.11}$$

$$\underline{f}^* \;=\; \underline{0} \tag{2.12}$$

$$\underline{u}^* \;=\; \underline{0} \quad \text{on} \quad \Gamma_1 \tag{2.13}$$

$$\left[\frac{\partial \underline{f}^*}{\partial \underline{u}_x}\,dy - \frac{\partial \underline{f}^*}{\partial \underline{u}_y}\,dx\right]\underline{\lambda}^* \;=\; \underline{0} \quad \text{on} \quad \Gamma_2 . \tag{2.14}$$

The following assumptions are made about the problem:

(1) Functions $\underline{f}(x,y,\underline{u},\underline{u}_x,\underline{u}_y)$ and $g(x,y,\underline{u})$ are twice continuously differentiable with respect to $\underline{u}$, $\underline{u}_x$ and $\underline{u}_y$ and continuously differentiable with respect to x and y.

(2) The operators $\frac{\partial^{i+j+k}\underline{f}}{\partial\underline{u}^i\partial\underline{u}_x^j\partial\underline{u}_y^k}$ and $\frac{\partial^i g}{\partial\underline{u}^i}$, with $i+j+k \leq 2$, map bounded neighbourhoods of U into bounded neighbourhoods of $\{L_2[R],E^n\}$ and $\{L_2[R],E^1\}$ respectively.

(3) The Lagrangian is extremised at the pair $(\underline{u}^*,\underline{\lambda}^*)$, which satisfies (2.11)-(2.14).

(4) The second variation of L with respect to $\underline{u}$ is strongly positive in a bounded convex neighbourhood of $(\underline{u}^*,\underline{\lambda}^*)$.

Let $F = g + \underline{f}^T\underline{\lambda}$. The condition (4) can be expressed in the following way:

$$\iint_R [\underline{h}^T,\underline{h}_x^T,\underline{h}_y^T]H(\tilde{\underline{u}},\tilde{\underline{\lambda}})[\underline{h}^T,\underline{h}_x^T,\underline{h}_y^T]^T \, dxdy \geq \sigma\|\underline{h}\|_{2,1}^2 \, , \tag{2.15}$$

where σ = const. > 0, $\underline{h} = \underline{u}-\tilde{\underline{u}}$, $u \in U$, $(\tilde{\underline{u}},\tilde{\underline{\lambda}}) \in N(\underline{u}^*) \times N(\underline{\lambda}^*)$, the neighbourhoods of $\underline{u}^*$ and $\underline{\lambda}^*$ and:

$$H(\underline{u},\underline{\lambda}) = \begin{bmatrix} F_{\underline{u}\underline{u}} & F_{\underline{u}\underline{u}_x} & F_{\underline{u}\underline{u}_y} \\ F_{\underline{u}_x\underline{u}} & F_{\underline{u}_x\underline{u}_x} & F_{\underline{u}_x\underline{u}_y} \\ F_{\underline{u}_y\underline{u}} & F_{\underline{u}_y\underline{u}_x} & F_{\underline{u}_y\underline{u}_y} \end{bmatrix} . \tag{2.16}$$

The following lemma describes the behaviour of the Lagrangian at the extremal and gives a result which will be needed for the proof of subsequent theorems:

Lemma 1. Let $\underline{u} \in N(\underline{u}^*)$ and $\underline{\lambda} \in N(\underline{\lambda}^*)$, then L has a degenerate saddle point at $(\underline{u}^*, \underline{\lambda}^*)$, that is, the following condition is satisfied:

$$L[\underline{u}^*, \underline{\lambda}] = L[\underline{u}^*, \underline{\lambda}^*] \le L[\underline{u}, \underline{\lambda}^*] , \qquad \forall \ \underline{\lambda} \in \Lambda . \tag{2.17}$$

Proof. The left-hand equality follows trivially by noting that $\underline{f}^* = 0$ and thus the condition is satisfied for any $\underline{\lambda}$. To obtain the right-hand inequality we expand $L[\underline{u}, \underline{\lambda}^*]$ into a Taylor series:

$$L[\underline{u}, \underline{\lambda}^*] = L[\underline{u}^*, \underline{\lambda}^*] + \iint_R \{\underline{\rho}^T(\underline{u}^*, \underline{\lambda}^*)\Delta\underline{u}\}dxdy$$

$$+ \oint_\Gamma \left\{ \left[\frac{\partial \underline{f}^*}{\partial \underline{u}_x} \underline{\lambda}^*\right]^T \Delta\underline{u} \; dy - \left[\frac{\partial \underline{f}^*}{\partial \underline{u}_y} \underline{\lambda}^*\right]^T \Delta\underline{u} \; dx \right\}$$

$$+ \int_0^1 (1-\tau) \iint_R [\Delta\underline{u}^T, \Delta\underline{u}_x^T, \Delta\underline{u}_y^T] H(\tilde{\underline{u}}, \underline{\lambda}^*) [\Delta\underline{u}^T, \Delta\underline{u}_x^T, \Delta\underline{u}_y^T]^T dxdy \; d\tau , \tag{2.18}$$

where $\tilde{\underline{u}} = \tau\underline{u} + (1-\tau)\underline{u}^*$ and $\Delta\underline{u} = \underline{u} - \underline{u}^*$. The first and second integrand are zero since (2.11), (2.12) and (2.13) are satisfied, and using the strong positivity assumption (2.15) we obtain:

$$L[\underline{u}, \underline{\lambda}^*] \ge L[\underline{u}^*, \underline{\lambda}^*] + \sigma\|\Delta\underline{u}\|_{2,1}^2 \ge L[\underline{u}^*, \underline{\lambda}^*] , \tag{2.19}$$

which proves the lemma. □

We note that the assumption (4) is a local sufficiency condition.

3. Finite dimensional approximation of the problem

The solution of the original problem can be found by solving the necessary conditions (2.11)-(2.14). The exact solution of these, usually non-linear, partial differential equations is often impossible, and some approximation is needed. We look for the solutions $\underline{u}$ and $\underline{\lambda}$, belonging to finite dimensional subspaces generated by basis functions $w_{i,m}$, $i = 1,2,\ldots m$, which can be assumed orthogonal without loss of generality. Thus $\underline{u} \in C_m \subset U$ and $\underline{\lambda} \in M_m \subset \Lambda$, where C_m and M_m are finite dimensional spaces defined by:

$$C_m = \{\underline{u} | \underline{u} \in C_m(w) \cap \bar{C}\} \tag{3.1}$$

$$M_m = \{\underline{\lambda} | \underline{\lambda} \in M_m(w) \cap \bar{M}\} . \tag{3.2}$$

The spaces $C_m(w)$ and $M_m(w)$ are given by:

$$C_m(w) = \{\underline{u} | \underline{u}(x,y) = \sum_{i=1}^{m} \underline{\gamma}_i w_{i,m}(x,y), (x,y) \in R; \quad \underline{u} = \underline{0} \quad \text{on} \quad \Gamma_1\} \tag{3.3}$$

$$M_m(w) = \{\underline{\lambda} | \underline{\lambda}(x,y) = \sum_{i=1}^{m} \underline{\beta}_i w_{i,m}(x,y), (x,y) \in R\} \tag{3.4}$$

and $\bar{C}$ and $\bar{M}$ denote the closures of:

$$C = \{\underline{u} | \underline{u} \in U \cap N(\underline{u}^*) \cap D_u\} \tag{3.5}$$

$$M = \{\underline{\lambda} | \underline{\lambda} \in \Lambda \cap N(\underline{\lambda}^*) \cap D_\lambda\} , \tag{3.6}$$

where D_u and D_λ are the sets of n-vector functions and r-vector functions which are continuous, piecewise differentiable functions on R. Thus we look for admissible, continuous and piecewise differentiable finite dimensional approximations which are in the neighbourhood of the optimum. The spaces C_m and M_m are closed and convex, and so they possess unique best approximations to $\underline{u}^*$ and $\underline{\lambda}^*$, which we shall denote by $\hat{\underline{u}}$ and $\hat{\underline{\lambda}}$, and which are defined in the following way:

$$\|\hat{\underline{u}}-\underline{u}^*\|_{2,1} = \inf_{\underline{u}\in C_m} \|\underline{u}-\underline{u}^*\|_{2,1} \tag{3.7}$$

$$\|\hat{\underline{\lambda}}-\underline{\lambda}^*\|_{2,1} = \inf_{\underline{\lambda}\in M_m} \|\underline{\lambda}-\underline{\lambda}^*\|_{2,1} \, . \tag{3.8}$$

Let $\epsilon(\hat{\underline{u}}) = \hat{\underline{u}}-\underline{u}^*$ and $\epsilon(\hat{\underline{\lambda}}) = \hat{\underline{\lambda}}-\underline{\lambda}^*$. We assume that C_m and M_m are good approximating spaces in the sense that $\|\epsilon(\hat{\underline{u}})\|_{2,1} \to 0$ and $\|\epsilon(\hat{\underline{\lambda}})\|_{2,1} \to 0$ as $m \to \infty$.

We now state the finite-dimensional analogues for the three equivalent formulations of the problem (2.1)-(2.3):

<u>Ritz Method</u> Find $\bar{\underline{u}} \in C_m$ such that

$$I(\bar{\underline{u}}) = \inf_{\underline{u}\in C_m} I(\underline{u}) \, , \tag{3.9}$$

subject to

$$\iint_R w_{j,m}(x,y)\underline{f}(x,y,\bar{\underline{u}},\bar{\underline{u}}_x,\bar{\underline{u}}_y)\, dxdy = \underline{0} \qquad j = 1,\ldots,m \tag{3.10}$$

$$\bar{\underline{u}}(x,y) = 0 \text{ on } \Gamma_1 \, . \tag{3.11}$$

Alternatively, the problem can be formulated in the following way.

Find the pair $(\bar{\underline{u}},\bar{\underline{\lambda}}) \in C_m \times M_m$ such that

$$L[\bar{\underline{u}},\bar{\underline{\lambda}}] \;=\; \sup_{\underline{\lambda}\in M_m}\; \inf_{\underline{u}\in C_m} L[\underline{u},\underline{\lambda}]\ . \tag{3.12}$$

The third formulation of the problem has the following finite dimensional equivalent:

<u>Galerkin Method</u> Find the pair $[\bar{\underline{u}},\bar{\underline{\lambda}}] \in C_m \times M_m$ such that

$$\frac{\partial L}{\partial \underline{u}}[\bar{\underline{u}},\bar{\underline{\lambda}}] \;=\; \underline{0} \tag{3.13}$$

$$\frac{\partial L}{\partial \underline{\lambda}}[\bar{\underline{u}},\bar{\underline{\lambda}}] \;=\; \underline{0}\ . \tag{3.14}$$

Because of the equivalence of the formulations both Ritz and Galerkin method yield the same set of equations:

$$\iint_R w_{j,m}\, \underline{\rho}(\bar{\underline{u}},\bar{\underline{\lambda}})\, dxdy \;=\; \underline{0}\ . \qquad j = 1,\ldots,m \tag{3.15}$$

$$\iint_R w_{j,m}\, \underline{f}(x,y,\bar{\underline{u}},\bar{\underline{u}}_x,\bar{\underline{u}}_y)\, dxdy \;=\; \underline{0}\ , \qquad j = 1,\ldots,m \tag{3.16}$$

$$\left[\frac{\partial \bar{\underline{f}}}{\partial \underline{u}_x}\, dy \;-\; \frac{\partial \bar{\underline{f}}}{\partial \underline{u}_y}\, dx\right] \bar{\underline{\lambda}} \;=\; \underline{0} \quad \text{on } \Gamma_2\ , \tag{3.17}$$

where $\bar{\underline{g}} = \underline{g}(x,y,\bar{\underline{u}})$ and $\bar{\underline{f}} = \underline{f}(x,y,\bar{\underline{u}},\bar{\underline{u}}_x,\bar{\underline{u}}_y)$. The second formulation is required for the convergence proofs which also need the result established by the following lemma:

<u>Lemma 2</u>. Let $(\bar{\underline{u}},\bar{\underline{\lambda}})$ be the solution of the Ritz-Galerkin equations (3.15)-(3.17), and let $(\underline{u},\underline{\lambda})$ denote an arbitrary pair in $C_m \times M_m$. Then L has a degenerate saddle point at $(\bar{\underline{u}},\bar{\underline{\lambda}})$ such that:

$$L[\bar{\underline{u}},\underline{\lambda}] \;=\; L[\bar{\underline{u}},\bar{\underline{\lambda}}] \le L[\underline{u},\bar{\underline{\lambda}}] \; . \qquad (3.18)$$

<u>Proof</u>. The proof is similar to the proof of Lemma 1. □

Let

$$I^* \;=\; L[\underline{u}^*,\underline{\lambda}^*] \quad (= I(\underline{u}^*)) \qquad (3.19)$$

$$I_m \;=\; L[\bar{\underline{u}},\bar{\underline{\lambda}}] \quad (= I(\bar{\underline{u}})), \qquad (3.20)$$

where $(\bar{\underline{u}},\bar{\underline{\lambda}})$ is the solution of the Ritz-Galerkin equations.

We aim to prove that the method is convergent, that is $\|\underline{u}^*-\bar{\underline{u}}\|_{2,1} \to 0$, $|I_m - I^*| \to 0$ as $m \to \infty$. We now state the theorem which gives the bounds on $\|\underline{u}^*-\bar{\underline{u}}\|_{2,1}$.

<u>Theorem 1</u>. Let $(\bar{\underline{u}},\bar{\underline{\lambda}}) \in C_m \times M_m$ be the solution of the Ritz-Galerkin equations (3.15)-(3.17). Then there exists a non-negative constant c, independent of m, such that

$$\|\underline{u}^*-\bar{\underline{u}}\|_{2,1} \le \eta_m \qquad (3.21)$$

where

$$\eta_m \;=\; c[\|\epsilon(\hat{\underline{\lambda}})\|_2 + \|\epsilon(\hat{\underline{u}})\|_{2,1} + \|\epsilon(\hat{\underline{u}})\|_{2,1}^{1/2}] \; . \qquad (3.22)$$

<u>Proof</u>. From Lemma 2 we have that:

$$L[\bar{\underline{u}},\underline{\lambda}] \;=\; L[\bar{\underline{u}},\bar{\underline{\lambda}}] \le L[\underline{u},\bar{\underline{\lambda}}] \; ,$$

for any $(\underline{u},\underline{\lambda}) \in C_m \times M_m$, and in particular

$$L[\bar{\underline{u}},\hat{\underline{\lambda}}] \;=\; I_m \le L[\hat{\underline{u}},\bar{\underline{\lambda}}] \; . \qquad (3.23)$$

Expanding the right-hand side of (3.23) into a Taylor series about $(\underline{u}^*,\bar{\underline{\lambda}})$ we obtain:

$$I_m \le L[\underline{u}^*,\bar{\underline{\lambda}}] + \int_R\int\Bigg[\left[\frac{\partial g^*}{\partial \underline{u}} + \frac{\partial \underline{f}^*}{\partial \underline{u}}\bar{\underline{\lambda}}\right]^T(\hat{\underline{u}}-\underline{u}^*) + \left[\frac{\partial \underline{f}}{\partial \underline{u}_x}\bar{\underline{\lambda}}\right]^T(\hat{\underline{u}}_x-\underline{u}_x^*)$$

$$+ \left[\frac{\partial \underline{f}^*}{\partial \underline{u}_y}\bar{\underline{\lambda}}\right]^T(\hat{\underline{u}}_y-\underline{u}_y^*)\Bigg]dxdy + \int_0^1(1-\tau)\int_R\int\left[(\hat{\underline{u}}-\underline{u}^*)^T,\ (\hat{\underline{u}}_x-\underline{u}_x^*)^T,\right.$$

$$\left.(\hat{\underline{u}}_y-\underline{u}_y^*)^T\right]H(\tilde{\underline{u}},\bar{\underline{\lambda}})\left[(\hat{\underline{u}}-\underline{u}^*)^T,\ (\hat{\underline{u}}_x-\underline{u}_x^*)^T,\ (\hat{\underline{u}}_y-\underline{u}_y^*)^T\right]^T dxdyd\tau \qquad (3.24)$$

where $\tilde{\underline{u}} = \tau\hat{\underline{u}} + (1-\tau)\underline{u}^*$. After the use of the Schwarz inequality and Lemma 1, (3.24) becomes:

$$I_m \le I^* + \left\|\frac{\partial g^*}{\partial \underline{u}}\right\|_2\|\hat{\underline{u}}-\underline{u}^*\|_2 + \|\bar{\underline{\lambda}}\|_2\left\|\left[\frac{\partial \underline{f}^{*T}}{\partial \underline{u}},\ \frac{\partial \underline{f}^{*T}}{\partial \underline{u}_x},\ \frac{\partial \underline{f}^{*T}}{\partial \underline{u}_y}\right]\left[(\hat{\underline{u}}-\underline{u}^*)^T,(\hat{\underline{u}}_x-\underline{u}_x^*)^T,\right.\right.$$

$$\left.\left.(\hat{\underline{u}}_y-\underline{u}_y^*)^T\right]^T\right\|_2$$

$$+ \tfrac{1}{2}\|\hat{\underline{u}}-\underline{u}^*\|_{2,1}\|H[(\hat{\underline{u}}-\underline{u}^*)^T,(\hat{\underline{u}}_x-\underline{u}_x^*)^T,(\hat{\underline{u}}_y-\underline{u}_y^*)^T]^T\|_2 \ . \qquad (3.25)$$

Using the assumption (2) and the fact that $\|\bar{\underline{\lambda}}\|_2 \le c_\lambda$, (since $\bar{\underline{\lambda}} \in \bar{M}$), we finally obtain:

$$I_m \le I^* + c_1\|\epsilon(\hat{\underline{u}})\|_{2,1} + c_2\|\epsilon(\hat{\underline{u}})\|_{2,1}^2 \ , \qquad (3.26)$$

where c_1 and c_2 are positive constants, independent of m.

A lower bound on I_m is obtained by expanding the left-hand side of (3.23) into a Taylor series about $(\underline{u}^*,\hat{\underline{\lambda}})$:

$$I_m = L[\underline{u}^*, \bar{\underline{\lambda}}] + \iint_R \Big[\Big[\frac{\partial g^*}{\partial \underline{u}} + \frac{\partial \underline{f}^*}{\partial \underline{u}}\hat{\underline{\lambda}}\Big]^T (\bar{\underline{u}}-\underline{u}^*) + \Big[\frac{\partial \underline{f}^*}{\partial \underline{u}_x}\hat{\underline{\lambda}}\Big]^T (\bar{\underline{u}}_x-\underline{u}_x^*)$$

$$+ \Big[\frac{\partial \underline{f}^*}{\partial \underline{u}_y}\hat{\underline{\lambda}}\Big](\bar{\underline{u}}_y-\underline{u}_y^*)\Big]dxdy + \int_0^1 (1-\tau)\iint_R \Big[(\bar{\underline{u}}-\underline{u}^*)^T,\ (\bar{\underline{u}}_x-\underline{u}_x^*)^T,$$

$$(\bar{\underline{u}}_y-\underline{u}_y^*)^T\Big] H(\tilde{\underline{u}},\hat{\underline{\lambda}})\Big[(\bar{\underline{u}}-\underline{u}^*)^T,\ (\bar{\underline{u}}_x-\underline{u}_x^*)^T,\ (\bar{\underline{u}}_y-\underline{u}_y^*)^T\Big]^T dxdyd\tau \qquad (3.27)$$

where $\tilde{\underline{u}} = \tau\bar{\underline{u}} + (1-\tau)\underline{u}^*$. We note that at the optimum the first variation of the Lagrangian vanishes, and thus:

$$\iint_R \Big[\Big[\frac{\partial g^*}{\partial \underline{u}} + \frac{\partial \underline{f}^*}{\partial \underline{u}}\underline{\lambda}^*\Big]^T (\bar{\underline{u}}-\underline{u}^*) + \Big[\frac{\partial \underline{f}^*}{\partial \underline{u}_x}\underline{\lambda}^*\Big]^T (\bar{\underline{u}}_x-\underline{u}_x^*) + \Big[\frac{\partial \underline{f}^*}{\partial \underline{u}_y}\underline{\lambda}^*\Big](\bar{\underline{u}}_y-\underline{u}_y^*)\Big]dxdy = 0.$$

(3.28)

Subtracting (3.28) from (3.27) and using the Schwarz inequality and the strong positivity assumption, we get the following expression:

$$I_m \geq I^* - \|\hat{\underline{\lambda}}-\underline{\lambda}^*\|_2 \Big\| \Big[\frac{\partial \underline{f}^{*T}}{\partial \underline{u}},\ \frac{\partial \underline{f}^{*T}}{\partial \underline{u}_x},\ \frac{\partial \underline{f}^{*T}}{\partial \underline{u}_y}\Big]\Big[(\bar{\underline{u}}-\underline{u}^*)^T,\ (\bar{\underline{u}}_x-\underline{u}_x^*)^T,\ (\bar{\underline{u}}_y-\underline{u}_y^*)^T\Big]^T \Big\|_2$$

$$+ \sigma\|\bar{\underline{u}}-\underline{u}^*\|_{2,1}^2$$

which together with the assumption (2) gives:

$$I_m \geq I^* - c_3\|\epsilon(\hat{\underline{\lambda}})\|_2 \|\bar{\underline{u}}-\underline{u}^*\|_{2,1} + \sigma\|\bar{\underline{u}}-\underline{u}^*\|_{2,1}^2 , \qquad (3.29)$$

where c_3 is a positive constant, independent of m. Combining

(3.26) with (3.29) we obtain:

$$I^* - c_3\|\epsilon(\hat{\underline{\lambda}})\|_2\|\bar{\underline{u}}-\underline{u}^*\|_{2,1} + \sigma\|\bar{\underline{u}}-\underline{u}^*\|_{2,1}^2 \le I^* + c_1\|\epsilon(\hat{\underline{u}})\|_{2,1} + c_2\|\epsilon(\hat{\underline{u}})\|_{2,1}^2 , \tag{3.30}$$

which is equivalent to:

$$\left[\|\bar{\underline{u}}-\underline{u}^*\|_{2,1} - \frac{c_3}{2\sigma}\|\epsilon(\hat{\underline{\lambda}})\|_2\right]^2 \le \frac{c_3^2}{4\sigma^2}\|\epsilon(\hat{\underline{\lambda}})\|_2^2 + \frac{c_1}{\sigma}\|\epsilon(\hat{\underline{u}})\|_{2,1} + \frac{c_2}{\sigma}\|\epsilon(\hat{\underline{u}})\|_{2,1}^2 . \tag{3.31}$$

Thus:

$$\|\bar{\underline{u}}-\underline{u}^*\|_{2,1} \le \frac{c_3}{2\sigma}\|\epsilon(\hat{\underline{\lambda}})\|_2 + \left[\frac{c_3^2}{4\sigma^2}\|\epsilon(\hat{\underline{\lambda}})\|_2^2 + \frac{c_1}{\sigma}\|\epsilon(\hat{\underline{u}})\|_{2,1} + \frac{c_2}{\sigma}\|\epsilon(\hat{\underline{u}})\|_{2,1}^2\right]^{1/2} , \tag{3.32}$$

which immediately gives:

$$\|\bar{\underline{u}}-\underline{u}^*\|_{2,1} \le c\left[\|\epsilon(\hat{\underline{\lambda}})\|_2 + \|\epsilon(\hat{\underline{u}})\|_{2,1} + \|\epsilon(\hat{\underline{u}})\|_{2,1}^{2}\right]^{1/2} , \tag{3.33}$$

the desired bound on $\|\bar{\underline{u}}-\underline{u}^*\|_{2,1}$. □

We next consider the bounds on the cost functional I_m.

<u>Theorem 2</u>. Let $(\bar{\underline{u}},\bar{\underline{\lambda}}) \in C_m \times M_m$ be the solution of the Ritz-Galerkin equations (3.15)-(3.17). Then there exist positive constants K_1 and K_2, independent of m, such that:

$$I^* - K_1\|\epsilon(\hat{\underline{\lambda}})\|_2^2 \le I_m \le I^* + K_2\zeta_m , \tag{3.34}$$

where

$$\zeta_m = \|\epsilon(\hat{\underline{u}})\|_{2,1} + \|\epsilon(\hat{\underline{u}})\|_{2,1}^2 . \tag{3.35}$$

Proof. The upper bound is obtained immediately from (3.26). To obtain the lower bound we introduce the intermediate approximation:

$$J_m = \sup_{\underline{\lambda}\in M_m} \inf_{\underline{u}\in N(\underline{u}^*)} L[\underline{u},\underline{\lambda}] . \tag{3.36}$$

We observe that $I_m \geq J_m$, since $C_m \subset N(\underline{u}^*)$, and that

$$J_m \geq \inf_{\underline{u}\in N(\underline{u}^*)} L[\underline{u},\hat{\underline{\lambda}}] . \tag{3.37}$$

Expanding the right-hand side of (3.37) into a Taylor series about $(\underline{u}^*,\hat{\underline{\lambda}})$ we obtain:

$$L[\underline{u},\hat{\underline{\lambda}}] = I^* + \int_R\int\left[\left[\frac{\partial g^*}{\partial \underline{u}} + \frac{\partial \underline{f}^*}{\partial \underline{u}}\hat{\underline{\lambda}}\right]^T(\underline{u}-\underline{u}^*) + \left[\frac{\partial \underline{f}^*}{\partial \underline{u}_x}\hat{\underline{\lambda}}\right]^T(\underline{u}_x-\underline{u}_x^*) + \left[\frac{\partial \underline{f}^*}{\partial \underline{u}_y}\hat{\underline{\lambda}}\right]^T(\underline{u}_y-\underline{u}_y^*)\right]dxdy$$

$$+ \int_0^1(1-t)\int_R\int[(\underline{u}-\underline{u}^*)^T,(\underline{u}_x-\underline{u}_x^*)^T,(\underline{u}_y-\underline{u}_y^*)^T]H(\tilde{\underline{u}},\hat{\underline{\lambda}})[(\underline{u}-\underline{u}^*)^T,(\underline{u}_x-\underline{u}_x^*)^T,$$

$$(\underline{u}_y-\underline{u}_y^*)^T]^T dxdyd\tau , \tag{3.38}$$

where $\tilde{\underline{u}} = \tau\underline{u} + (1-\tau)\underline{u}^*$. We use the Schwarz inequality, an expression similar to (3.28) and the strong positivity assumption to obtain:

$$L[\underline{u},\hat{\underline{\lambda}}] \geq I^* - \|\hat{\underline{\lambda}}-\underline{\lambda}^*\|_2 \left\| \left[\frac{\partial \underline{f}^{*T}}{\partial \underline{u}}, \frac{\partial \underline{f}^{*T}}{\partial \underline{u}_x}, \frac{\partial \underline{f}^{*T}}{\partial \underline{u}_y}\right][(\underline{u}-\underline{u}^*)^T,(\underline{u}_x-\underline{u}_x^*)^T,(\underline{u}_y-\underline{u}_y^*)^T]^T \right\|_2$$

$$+ \sigma\|\underline{u}-\underline{u}^*\|_{2,1}^2 . \tag{3.39}$$

This, together with the assumption (2), gives

$$L[\underline{u},\hat{\underline{\lambda}}] \geq I^* - K\|\epsilon(\hat{\underline{\lambda}})\|_2\|\Delta\underline{u}\|_{2,1} + \sigma\|\Delta\underline{u}\|_{2,1}^2 \;, \tag{3.40}$$

where $\Delta\underline{u} = \underline{u}-\underline{u}^*$, and K is a positive constant. Thus

$$\begin{aligned} L[\underline{u},\hat{\underline{\lambda}}] &\geq I^* + \sigma(\|\Delta\underline{u}\|_{2,1} - \frac{K}{2\sigma}\|\epsilon(\hat{\underline{\lambda}})\|_2)^2 - \frac{K^2}{4\sigma^2}\|\epsilon(\hat{\underline{\lambda}})\|_2^2 \\ &\geq I^* - K_1\|\epsilon(\hat{\underline{\lambda}})\|_2^2 \;. \end{aligned} \tag{3.41}$$

From (3.37) and (3.41) we arrive at the following result:

$$J_m \geq \inf_{\underline{u}\in N(\underline{u}^*)} \{I^* - K_1\|\epsilon(\hat{\underline{\lambda}})\|_2^2\} = I^* - K_1\|\epsilon(\hat{\underline{\lambda}})\|_2^2 \;. \tag{3.42}$$

The lower bound on I_m is immediately obtained by observing that $I_m \geq J_m$. □

We have assumed that C_m and M_m are good approximating spaces and thus $\|\epsilon(\hat{\underline{\lambda}})\|_2$, $\|\epsilon(\hat{\underline{u}})\|_2$ and $\|\epsilon(\hat{\underline{u}})\|_{2,1}$ each tend to zero as $m \to \infty$. It follows, therefore, that η_m and ζ_m also tend to zero as m tends to infinity. This proves the convergence of $\bar{\underline{u}}$ and I_m.

We note that if $\underline{f}(x,y,\underline{u},\underline{u}_x,\underline{u}_y) = \underline{0}$ is a quasilinear equation of the form

$$a(x,y,\underline{u})\underline{u}_x + b(x,y,\underline{u})\underline{u}_y + e(x,y,\underline{u}) = 0 \;, \tag{3.43}$$

then the conditions on the approximating spaces C_m and M_m may be slightly relaxed. In that case, the results of the theorems hold provided $\underline{u}$ is square integrable with square integrable first partial derivatives.

4. Application to Optimal Design Problems

We now consider the application of the above theory to optimal design problems which involve the computation of an unknown boundary curve. These problems cover a wide area of application in science and engineering. The examples include optimal design of an electromagnet [11], a nozzle [11], the minimum drag problem [11], optimization of a wing [11], an elastic body subject to torsion [8] and many others. All these problems require minimisation of a functional of the type (2.1) subject to constraints (2.2) and some boundary conditions, over a region R bounded by a curve Γ which is either wholly or partly unknown. In particular, we consider the region as described in Figure 1

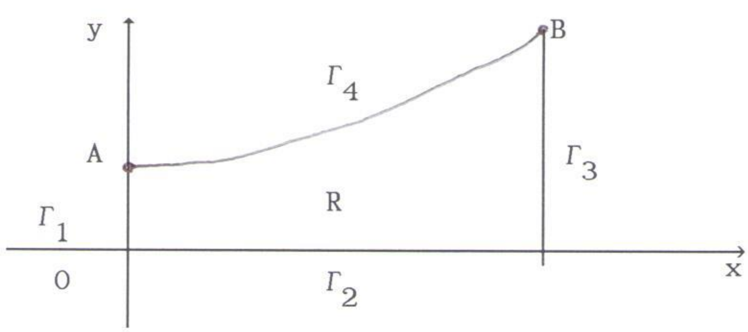

Figure 1

where the shape of Γ_4 is not known and the positions of points A and B (and hence the lengths of Γ_1 and Γ_3) may or may not be specified. It is assumed that the shape of the unknown boundary may be expressed in the form of an equation:

$$y = \ell(x)$$

where the function $\ell(x)$ and its first and second derivatives are square integrable and $\ell(x) > 0, \ \forall x \in [0,1]$.

The necessary conditions for an optimum are given by (2.11), (2.12) and

$$\int_{\Gamma=\Gamma_1 \cup \Gamma_2 \cup \Gamma_3 \cup \Gamma_4} \delta \underline{u}^T \frac{\partial \underline{f}}{\partial \underline{u}_x} \underline{\lambda} \, dy - \delta \underline{u}^T \frac{\partial \underline{f}}{\partial \underline{u}_y} \underline{\lambda} \, dx = 0 \tag{4.1}$$

where $\delta \underline{u}$ denotes the variation in $\underline{u}$. If the positions of A and B are not given we also need to satisfy:

$$\int_{\Gamma_1 \cup \Gamma_3 \cup \Gamma_4} \delta y \left[- \underline{u}_y^T \frac{\partial \underline{f}}{\partial \underline{u}_x} \underline{\lambda} \, dy - \left(- \underline{u}_y^T \frac{\partial \underline{f}}{\partial \underline{u}_y} \underline{\lambda} + g + \underline{f}^T \lambda \right) dx \right] = 0 \tag{4.2}$$

where δy represents the variation in y.

The optimal shape problem is tackled by transforming R to a known region. Thus we introduce new coordinates X and Y such that $X = x$ and $Y = \frac{y}{\ell(x)}$ which has the effect of mapping R onto a unit square. The transformed functions $\underline{u}$ and $\underline{\lambda}$ are denoted by:

$$\begin{aligned} \underline{u}(x,y) &= \underline{u}(X, Y\ell) = \underline{U}(X,Y) \\ \underline{\lambda}(x,y) &= \underline{\lambda}(X, Y\ell) = \Lambda(X,Y) \end{aligned} \tag{4.3}$$

and the derivatives $\underline{u}_x$ and $\underline{u}_y$ become:

$$\begin{aligned} \underline{u}_x &= \underline{U}_X - Y \frac{\ell'}{\ell} \underline{U}_Y \\ \underline{u}_y &= \frac{1}{\ell} \underline{U}_Y \quad . \end{aligned} \tag{4.4}$$

The problem is now to minimise

$$\int_0^1\int_0^1 G(X,Y,\ell,\underline{U})dXdY \ . \tag{4.5}$$

Subject to

$$\underline{F}(X,Y,\ell,\ell',\underline{U},\underline{U}_X,\underline{U}_Y) = \underline{0} \tag{4.6}$$

and appropriate boundary conditions, where:

$$G(X,Y,\ell,\underline{U}) = \ell g(x,y,\underline{u}) \tag{4.7}$$

and

$$F(X,Y,\ell,\ell',\underline{U},\underline{U}_X,\underline{U}_Y) = \ell\underline{f}(x,y,\underline{u},\underline{u}_x,\underline{u}_y) \ . \tag{4.8}$$

It can be shown that the transformed problem is equivalent to the original problem in the sense that they both yield the same set of necessary conditions [7].

The Ritz-Galerkin method, described in section 3, can be applied to the transformed problem. Let $\underline{V} = [\underline{U}^T,\ell]^T$. The proof of convergence, given in the previous section, is valid provided the space $C_m(w)$ is constructed in the following way:

$$C_m(w) = \left\{\left[\begin{array}{c}\underline{U}\\ \hline \ell\end{array}\right] \middle| \underline{U} = \sum_{i=1}^{m} \underline{\gamma}_i w_{i,m}(X,Y), \ (X,Y) \in R; \right.$$

$$\left. \ell = \sum_{i=1}^{n(m)} \ell_i \psi_{i,m}(X) \ , \qquad X \in [0,1]\right\} \tag{4.9}$$

where $w_{i,m}$ and $\psi_{i,m}$ are appropriate basis functions. In this case the Ritz-Galerkin equations (3.15)-(3.17) for the optimal solution

$(\overline{\overline{\underline{U}}},\overline{\overline{\ell}},\overline{\overline{\Lambda}}) \in C_m \times M_m$ take the form

$$\iint_{R'} w_{j,m}\left[\frac{\partial\overline{\overline{G}}}{\partial\underline{U}} + \frac{\partial\overline{\overline{\underline{F}}}}{\partial\underline{U}}\,\overline{\overline{\underline{\Lambda}}} - \frac{\partial}{\partial X}\left[\frac{\partial\overline{\overline{\underline{F}}}}{\partial\underline{U}_X}\,\overline{\overline{\underline{\Lambda}}}\right] - \frac{\partial}{\partial Y}\left[\frac{\partial\overline{\overline{\underline{F}}}}{\partial\underline{U}_Y}\,\overline{\overline{\underline{\Lambda}}}\right]\right] dXdY = \underline{0}$$

$$\iint_{R'} w_{j,m}\overline{\overline{\underline{F}}}(X,Y,\ell,\ell',\underline{U},\underline{U}_X,\underline{U}_Y)\, dXdY = \underline{0} \tag{4.10}$$

$$\iint_{R'} \left\{\psi_{j,n}\left[\frac{\partial\overline{\overline{G}}}{\partial\ell} + \overline{\overline{\underline{\Lambda}}}^T\frac{\partial\overline{\overline{\underline{F}}}}{\partial\ell}\right] + \psi'_{j,n}\,\overline{\overline{\underline{\Lambda}}}^T\frac{\partial\overline{\overline{\underline{F}}}}{\partial\ell'}\right\} dXdY = 0$$

where $\overline{\overline{G}} \equiv G(X,Y,\overline{\overline{\ell}},\overline{\overline{\ell}}',\overline{\overline{\underline{U}}})$ and $\overline{\overline{\underline{F}}}(X,Y,\overline{\overline{\ell}},\overline{\overline{\ell}}',\overline{\overline{\underline{U}}},\overline{\overline{\underline{U}}}_X,\overline{\overline{\underline{U}}}_Y)$. The boundary conditions and the appropriate transversality conditions need also to be satisfied.

5. Example

In this section we present numerical results obtained for a very simple test example. More complex problems have also been solved by the same method [7],[8]. The test example is a degenerate case of the problem described in section 4 where the region of integration (Figure 2) is a rectangle of an unspecified length and unit width.

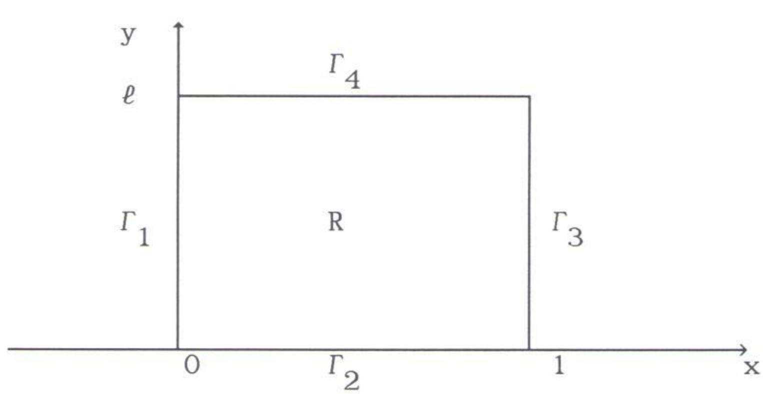

Figure 2

The equation of Γ_4 is therefore

$$y = \ell = \text{const.} \tag{5.1}$$

The aim is to determine ℓ and $u(x,y)$ which minimise the cost functional

$$\iint_R (u^2 - 1)\, dxdy \tag{5.2}$$

subject to

$$\nabla^2 u = 0 \tag{5.3}$$

and

$$u(x,\ell) = 0, \quad u_x(0,y) = u_x(1,y) = 0, \quad u_y(x,0) = 1. \tag{5.4}$$

In order to reduce the problem to the standard form, Laplace's equation (5.3) is treated as a system of first order partial differential equations:

$$\begin{aligned} u_x &= \phi \\ u_y &= \psi \\ \phi_x + \psi_y &= 0 . \end{aligned} \tag{5.5}$$

The problem is now in the form in which the transformation technique and the Ritz-Galerkin method may be applied. We adopt a finite element approach where the region is divided into elements and basis functions are chosen so that they are non-zero over a small patch of elements. Because of the shape of R it is natural to perform the subdivision into a finite number of rectangles. We choose a mesh of $N \times M$ equal rectangles, such that the lengths of the sides of each element are

$h_X = 1/N$ and $h_Y = 1/M$. The simplest basis functions which can be used with such a mesh are the bilinear functions. Let i denote the node with the coordinates (jh_X, kh_Y), $j = 0,1,\dots,N$, $k = 0,1,\dots,M$; then $i = kN + j$. The bilinear basis function $W_{i,m}(X,Y)$ is non-zero only for $(X,Y) \in [(j-1)h_X, (j+1)h_X] \times [(k-1)h_Y, (k+1)h_Y]$. In this region it is of the form:

$$W_{i,m}(X,Y) = a_1 + a_2X + a_3Y + a_4XY \tag{5.6}$$

where $W_{i,m} = 1$ at the ith node and $W_{i,m} = 0$ at neighbouring nodes, and the coefficients a_i, $i = 1,\dots,4$ are determined from the values of $W_{i,m}$ at these nodal points. The approximating space $C_m(w)$ is now in the form

$$C_m(w) = \left\{ \left[\frac{\underline{Z}}{\ell} \right] \middle| \underline{Z} = \sum_{i=1}^{(N+1)(M+1)} \underline{\gamma}_i \, W_{i,m}(X,Y), \ (X,Y) \in [0,1] \times [0,1], \ \ell = \text{const} \right\} \tag{5.7}$$

where $\underline{Z} = (\underline{U},\underline{\Phi},\underline{\Psi})^T$. With this choice of $C_m(w)$ the Ritz-Galerkin equations (4.10) are assembled and solved using the program of J.K. Reid [14] based on the Marquardt algorithm.

The numerical solution has been calculated for various numbers of elements and compared with the analytical solution, which in this case can be easily calculated as

$$u = y - 1, \qquad \ell = 1 . \tag{5.8}$$

The results for 2×2, 2×4 and 4×4 elements are given in tables 2, 3 and 4 while the analytical solution, evaluated at the mesh points, is given in Table 1. The errors in the discrete ℓ_2-norm are given at the end of each table. We observe that even for small meshes the method produces very accurate solutions.

6. Conclusions

In this paper we have devised a strategy for solving two-dimensional distributed parameter control problems and demonstrated convergence for a wide class of non-linear problems. The technique has also been applied to optimal design problems where a transformation technique is used to map the unknown region onto a known one. The results for a simple test example indicate that the technique works well in practice. More complex problems have been also successfully solved in the same way [7], [8]. The transformation introduces the advantage that the domain of integration becomes fixed and thus needs to be discretized only once. The resulting equations are, however, always non-linear, but in most cases this is not an additional disadvantage because of the non-linearity of the original problem. The technique requires all the variables to be approximated simultaneously, which may result in a large system of non-linear equations. The equations are fortunately sparse and the application of sparse solvers, such as one described in [14] overcomes possible storage difficulties.

Table 1: Example 1 - Analytical Solution

$\ell = 1.0000$

u	0.0000	0.2500	0.5000	0.7500	1.0000 x
0.0000	-1.0000	-1.0000	-1.0000	-1.0000	-1.0000
0.2500	-0.7500	-0.7500	-0.7500	-0.7500	-0.7500
0.5000	-0.5000	-0.5000	-0.5000	-0.5000	-0.5000
0.7500	-0.2500	-0.2500	-0.2500	-0.2500	-0.2500
1.0000	0.0000	0.0000	0.0000	0.0000	0.0000
y					

Table 2: Example 1 - Numerical Solution - 2 × 2 Elements

$\bar{\ell} = 1.00000013$

$\bar{u}$	0.0000	0.2500	0.5000	0.7500	1.0000 x
0.0000	-1.00000012		-1.00000014		-1.00000012
0.2500					
0.5000	-0.50000055		-0.500000075		-0.500000056
0.7500					
1.0000	0.00000000		0.00000000		0.00000000
y					

Errors: $\|u-\bar{u}\|_2 = 2.453 \times 10^{-7}$ $\|\ell-\bar{\ell}\|_2 = 1.3 \times 10^{-7}$

Table 3: Example 1 - Numerical Solution - 2 × 4 Elements

$\bar{\ell} = 0.999999828$

$\bar{u}$	0.0000	0.2500	0.5000	0.7500	1.0000 (x)
0.0000	-0.999999820	-0.999999838	-0.999999820	-0.999999837	-0.999999819
0.2500					
0.50000	-0.499999910	-0.499999922	-0.499999910	-0.499999923	-0.499999911
0.7500					
1.0000	0.000000000	0.000000000	0.000000000	0.000000000	0.000000000
y					

Errors: $\|u-\bar{u}\|_2 = 4.319 \times 10^{-7}$ $\|\ell-\bar{\ell}\|_2 = 1.72 \times 10^{-7}$

Table 4: Example 1 - Numerical Solution - 4 × 4 Elements

$\bar{\ell} = 0.999999997$

$\bar{u}$	0.0000	0.2500	0.5000	0.7500	1.0000 (x)
0.0000	-0.999999996	-0.999999999	-0.999999996	-0.999999999	-0.999999996
0.2500	-0.749999999	-0.749999997	-0.749999999	-0.749999997	-0.749999998
0.50000	-0.499999998	-0.499999999	-0.499999998	-0.500000000	-0.499999998
0.7500	-0.250000000	-0.249999999	-0.250000000	-0.249999999	-0.250000000
1.0000	0.000000000	0.000000000	0.000000000	0.000000000	0.000000000
y					

Errors: $\|u-\bar{u}\|_2 = 0.2195 \times 10^{-9}$ $\|\ell-\bar{\ell}\|_2 = 3.0 \times 10^{-9}$

References

1. BOSARGE, W.E. and JOHNSON, O.G.: Error Bounds of High-Order Accuracy for the State Regulator Problem via Piecewise Polynomial Approximation, SIAM J. on Control 9, (1971), pp 15-28.

2. BOSARGE, W.E., JOHNSON, O.G. and SMITH, C.L.: A Direct Method Approximation to the Linear Parabolic Regulator Problem over Multivariate Spline Basis, SIAM J. Numer. Anal. 10, (1973), pp 35-49.

3. BOSARGE, W.E., JOHNSON, O.G., McKNIGHT, R.S. and TIMLAKE, W.P.: The Ritz-Galerkin Procedure for Non-Linear Control Problems, SIAM J. Numer. Anal. 10, (1973), pp 94-111.

4. CHENEY, E.W.: Introduction to Approximation Theory, McGraw-Hill (1966).

5. CORNICK, D.E. and MICHEL, A.N.: Numerical Optimization of Linear Distributed Parameter Systems, JOTA 14, (1974), pp 73-98.

6. GELFAND, I.M. and FOMIN, S.V.: Calculus of Variations, Prentice-Hall, (1964).

7. LYLE, S.: Numerical Methods for Optimal Control Problems, Ph.D. Thesis, Dept. of Mathematics, University of Reading, (1981).

8. LYLE, S. and NICHOLS, N.K.: On the Finite Element Solution of Optimal Design Problems Using Transformation Techniques, The Mathematics of Finite Elements and Applications, IV, Academic Press, (1982), pp. 183-192.

9. McKNIGHT, R.S. and BOSARGE, W.E.: The Ritz-Galerkin Procedure for Parabolic Control Problems, SIAM J. Control 11, (1973), pp 510-523.

10. PIRONNEAU, O.: Variational Methods for the Numerical Solution of Free Boundary Problems and Optimum Design Problems, IRIA Report (1977).

11. PIRONNEAU, O.: Optimal Shape Design for Elliptic Systems, Springer-Verlag, (1984).

12. PONTRYAGIN, C.S., BOLTYANSKII, V.G., GAMKRELIDZE, R.V. and MISHCHENKO, E.F.: The Mathematical Theory of Optimal Processes, John Wiley and Sons, (1962).

13. De WITT, C.: A Survey of Methods of Numerical Solution of Optimal Control Problems, Mathematics and Mathematical Engineering 3, (1978), pp 139-171.

14. REID, J.K.: FORTRAN Subroutines for the Solution of Sparse Systems of Non-Linear Equations, AERE, Harwell Rpt. R7293, (1972).