On Numerical Approximation of an Optimal Control Problem in Linear Elasticity

Sobre la Aproximación Numérica de un Problema de Control Óptimo en Elasticidad Lineal

A. Rincon, (`rincon@dcc.ufrj.br`)
I-Shih Liu, (`liu@dmm.im.ufrj.br`)
Instituto de Matemática,
Universidade Federal do Rio de Janeiro
Caixa Postal 68530,
Rio de Janeiro 21945-970, Brazil

Abstract

We apply the optimal control theory to a linear elasticity problem. It can be formulated as a minimization problem of a cost functional. Numerical approximations are based on the optimality system of the corresponding minimization problem. An iterative method is considered, for which convergence of the approximate solutions is proved provided that a penalization parameter in the cost functional is not too small. Numerical solutions are presented to emphasize the role of this parameter. It is shown that the iterative method suffers a drawback that the approximate solutions are not good enough because the penalization parameter can not be taken sufficiently small. Such a limitation on the parameter is not inherited from the use of a penalization parameter in the cost functional. Indeed, some methods may be free of such a limitation. We have shown by a simple spectral analysis that the solution exists as the parameter tends to zero by the use of eigenfunction representations.

Key words and phrases: Optimal control, Linear elasticity, finite element method, Iterative method, Eigenfunction expansion.

Resumen

Aplicamos la teoría del control óptimo a un problema de elasticidad lineal que puede ser formulado como un problema de minimización de una funcional de costo. Las aproximaciones numéricas se basan en el sistema de optimalidad del correspondiente problema de minimización. Se considera un método iterativo para el cual se prueba la convergencia de las soluciones aproximadas suponiendo que un parámetro de penalización en la funcional de costo no sea muy pequeño. Se presentan soluciones numéricas para enfatizar el rol de este parámetro. Se muestra que el método iterativo tiene el inconveniente de que las soluciones aproximadas no son suficientemente buenas porque el parámetro de penalización no puede ser tomado suficientemente pequeño. Tal limitación sobre el parámetro no es heredada del uso de un parámetro de penalización en la funcional de costo. En verdad, algunos métodos pueden estar libres de esa limitación. Hemos mostrado mediante un simple análisis espectral que la solución existe cuando el parámetro tiende a cero mediante el uso de representaciones con funciones propias.

Palabras y frases clave: Control óptimo, elasticidad lineal, método de los elementos finitos, método iterativo, desarrollo en funciones propias.

1   Introduction

The theory of optimal control of systems governed by partial differential equations was essentially developed by Lions [7]. Many different type of control problems have been considered and solutions by numerical methods have been widely studied in the literature (see for examples, [3, 5, 7, 8]).

Theory of optimal control governed by a scalar equation of elliptic type, such as those involving heat conduction, are well-known in the literature. The problem is usually formulated as a minimization problem of a cost functional, involving a positive parameter for technical reasons. For the practical objective of the optimal control solution this penalization parameter should be taken as small as possible.

In this paper we shall apply the theory of optimal control to equilibrium problems of linear elasticity, governed by an elliptic system of partial differential equations. An iterative method for the optimality system equivalent to the minimization problem is considered and it is shown that the convergence of the iterative solutions is guaranteed if the penalization parameter is not too small. The role of this parameter will be examined in the numerical solutions with finite element method in the iterative procedure. It is found that such solutions are not good enough even with the smallest allowable parameter and

hence the iterative method which does not allow the parameter to be arbitrary small may hardly be able to deliver a satisfactory optimal solution. However, this limitation seems to have been overlooked in many similar problems in the literature [5, 8, 10], where the convergence of the numerical solution is ensured by simply taking this parameter as 1 or a fixed convenient positive number using gradient type methods. Therefore apparently those results may suffer the same drawback as ours in this respect.

Such a limitation is not inherited from the use of penalization parameter in the cost functional. Indeed, there are methods free of such limitation when the parameter tends to zero. For this reason, we shall consider a spectral analysis in which an explicit solution of the optimality system in Fourier series expansion can be obtained in the limit as the parameter tends to zero.

A bold-faced letter stands for a vector quantity and its components are represented by the corresponding normal letter with subindices ranging from 1 to n, the dimension of the physical space. The usual summation convention will be used to the component indices, *i.e.*, the repeated component indices indicate a sum over its range from 1 to n.

2   An Optimal Control Problem in Elasticity

We consider the following boundary value problem of Dirichlet type in linear elasticity:

$$\begin{cases} -\dfrac{\partial}{\partial x_j}\Big(C_{ijkl}\dfrac{\partial u_k}{\partial x_l}\Big) = f_i, & \text{in } \Omega \\[2mm] \qquad\quad u_i = 0, & \text{on } \partial\Omega \end{cases} \tag{1}$$

where the elasticity tensor C_{ijkl} and the external force f_i are given functions of $\boldsymbol{x} = (x_i)$ in a smooth region $\Omega \subset R^n$ with smooth boundary $\partial\Omega$. The problem is to determine the displacement field $\boldsymbol{u} = (u_i)$ satisfying the system (1). It is usually assumed that ([2])
a) The elasticity tensor C_{ijkl} satisfy the condition:

$$C_{ijkl} = C_{jikl} = C_{ijlk} = C_{klij}. \tag{2}$$

b) There are positive constants C and $\bar{C}$, such that for any symmetric matrices S_{ij},

$$CS_{ij}S_{ij} \le C_{ijkl}S_{ij}S_{kl} \le \bar{C}S_{ij}S_{ij}. \tag{3}$$

The assumption (a) follows from the existence of stored-energy function and the symmetry of stress and strain tensors, while the assumption (b) states that the elasticity tensor is bounded and strictly positive definite.

By virtue of these assumptions, the bilinear function σ defined by

$$\sigma(\boldsymbol{u}, \boldsymbol{v}) = \int_{\Omega} C_{ijkl} \frac{\partial u_i}{\partial x_j} \frac{\partial v_k}{\partial x_l} \, dx, \qquad \forall \, \boldsymbol{u}, \boldsymbol{v} \in (H_0^1(\Omega))^n \qquad (4)$$

is an inner product in $(H_0^1(\Omega))^n$, called the energy inner product. The energy norm $\|\boldsymbol{u}\|_{\sigma} = \sigma(\boldsymbol{u}, \boldsymbol{u})^{1/2}$ is equivalent to the standard norm $\|\boldsymbol{u}\|_1$ of the space $(H^1(\Omega))^n$ for $\boldsymbol{u} \in (H_0^1(\Omega))^n$. Indeed, from (3) we have

$$C\|\boldsymbol{u}\|_1 \leq \|\boldsymbol{u}\|_{\sigma} \leq \bar{C}\|\boldsymbol{u}\|_1. \qquad (5)$$

We denote the inner product in $(L^2(\Omega))^n$ by

$$(\boldsymbol{u}, \boldsymbol{v})_0 = \int_{\Omega} \boldsymbol{u} \cdot \boldsymbol{v} \, dv, \qquad \forall \, \boldsymbol{u}, \boldsymbol{v} \in (L^2(\Omega))^n,$$

where $\boldsymbol{u} \cdot \boldsymbol{v} = u_i v_i$ is the usual inner product in R^n. The usual norm in $(L^2(\Omega))^n$ will be denoted by $\|\boldsymbol{u}\|_0 = (\boldsymbol{u}, \boldsymbol{u})_0^{1/2}$.

The weak form of the problem (1) can now be stated as follows: For a given function $\boldsymbol{f} \in (L^2(\Omega))^n$, find the solution $\boldsymbol{u} \in (H_0^1(\Omega))^n$ such that

$$\sigma(\boldsymbol{u}, \boldsymbol{v}) = (\boldsymbol{f}, \boldsymbol{v})_0, \qquad \forall \, \boldsymbol{v} \in (H_0^1(\Omega))^n. \qquad (6)$$

We can verify easily that the bilinear function $\sigma(\cdot)$ is continuous and coercive in $(H_0^1(\Omega))^n$. Therefore, according to Lax-Milgram Theorem ([2, 9]) and by the use of elliptical regularity, for any $\boldsymbol{f} \in (L^2(\Omega))^n$, there is a unique solution $\boldsymbol{u} \in (H_0^1(\Omega))^n \cap (H^2(\Omega))^n$. The unique solution for the given $\boldsymbol{f}$ will be denoted by $\boldsymbol{u}(\boldsymbol{f})$.

Now let us turn to the formulation of an optimal control problem to obtain a prescribed displacement by means of externally applied forces on the body. Suppose that the equilibrium position of the body has some prescribed form given by a function $\boldsymbol{z}^o(\boldsymbol{x}) \in (L^2(\Omega))^n$, called an *objective function*. We can ask the question of how to determine the function $\boldsymbol{f}$ such that $\boldsymbol{u}(\boldsymbol{f})$ is as close to the objective function $\boldsymbol{z}^o$ as possible in $(L^2(\Omega))^n$.

By introducing the *cost functional* defined by

$$J(\boldsymbol{f}) = \|\boldsymbol{u}(\boldsymbol{f}) - \boldsymbol{z}^o\|_0^2 + N\|\boldsymbol{f}\|_0^2, \qquad (7)$$

the optimal control problem can be formulated as a minimization problem of the cost functional:

Determine a function $\boldsymbol{g} \in U_{ad}$ such that

$$J(\boldsymbol{g}) = \inf \big\{ J(\boldsymbol{f}); \ \boldsymbol{f} \in U_{ad} \big\}, \tag{8}$$

where U_{ad} is a convex and closed set of $(L^2(\Omega))^n$.

The real goal of the problem is to minimize the first term of the cost functional. The second term is added to limit the cost of control. The non-negative penalty parameter N is used to change the relative importance of the two terms in the cost functional. Therefore, for real purpose in obtaining the optimal result, the constant N should be taken as small as possible.

3   Optimality System

For numerical calculation of the optimal control we shall establish a more convenient formulation in terms of a system of differential equations. By the use of the Gateaux-differential of the cost functional J, the problem (8) can be characterized by

$$\int_\Omega (\boldsymbol{u}(\boldsymbol{g}) - \boldsymbol{z}^o) \cdot (\boldsymbol{u}(\boldsymbol{v}) - \boldsymbol{u}(\boldsymbol{g})) \, dx + N \int_\Omega \boldsymbol{g} \cdot (\boldsymbol{v} - \boldsymbol{g}) \, dx \geq 0, \qquad \forall \, \boldsymbol{v} \in U_{ad}. \tag{9}$$

In order to obtain a more convenient form for the numerical calculation of the function $\boldsymbol{g}$, we shall reformulate the problem through the adjoint system. Let the operator defined on the left hand side of the system (1) be denoted by $A : (H_0^1(\Omega))^n \to (L^2(\Omega))^n$,

$$A u_i = -\frac{\partial}{\partial x_j} \Big(C_{ijkl} \frac{\partial u_k}{\partial x_l} \Big). \tag{10}$$

Since by assumption, the elasticity C_{ijkl} is symmetric, the operator A is self-adjoint, *i.e.*,

$$(A\boldsymbol{u}, \boldsymbol{v})_0 = (\boldsymbol{u}, A\boldsymbol{v})_0, \qquad \forall \, \boldsymbol{u}, \boldsymbol{v} \in (H_0^1(\Omega))^n.$$

If we define the adjoint system by

$$\begin{cases} A\boldsymbol{p} = \boldsymbol{u}(\boldsymbol{g}) - \boldsymbol{z}^o & \text{in } \ \Omega, \\ \quad \boldsymbol{p} = 0 & \text{on } \ \partial\Omega, \end{cases} \tag{11}$$

then by substituting $(\boldsymbol{u}(\boldsymbol{g}) - \boldsymbol{z}^o)$ from (11) into (9), we have

$$\int_\Omega A\boldsymbol{p} \cdot (\boldsymbol{u}(\boldsymbol{v}) - \boldsymbol{u}(\boldsymbol{g})) + N \int_\Omega \boldsymbol{g} \cdot (\boldsymbol{v} - \boldsymbol{g}) \, dx \geq 0.$$

Since A is self-adjoint and $A\boldsymbol{u}(\boldsymbol{f}) = \boldsymbol{f}$ from the system (1), we have

$$\int_\Omega (\boldsymbol{p} + N\boldsymbol{g}) \cdot (\boldsymbol{v} - \boldsymbol{g})\, dx \geq 0, \qquad \forall\, \boldsymbol{v} \in U_{ad}. \tag{12}$$

We consider the case $U_{ad} = (L^2(\Omega))^n$, for which it follows that

$$\boldsymbol{g} = -\frac{1}{N}\boldsymbol{p}.$$

Therefore, we can define the *optimality system* by

$$\begin{cases} A\boldsymbol{u} = -\frac{1}{N}\,\boldsymbol{p} & \text{in } \Omega \\[2mm] A\boldsymbol{p} = \boldsymbol{u} - \boldsymbol{z}^o & \text{in } \Omega \\[2mm] \boldsymbol{u} = \boldsymbol{p} = 0 & \text{on } \partial\Omega. \end{cases} \tag{13}$$

This is a coupled system, besides some other solving methods, such as conjugate gradient method, least-squared method [5, 10, 4], we shall consider an iterative method which uncouples the system.

4   Iterative Method

In order to get a numerical solution for the problem, the optimality system is uncoupled in the following way: Given $\boldsymbol{p}^0 = 0$ then the values of $\boldsymbol{u}^m$ and $\boldsymbol{p}^m$ are iteratively calculated from the following algorithm:

$$\begin{cases} A\boldsymbol{u}^m = -\frac{1}{N}\,\boldsymbol{p}^{m-1} & \text{in } \Omega, \\[2mm] A\boldsymbol{p}^m = \boldsymbol{u}^m - \boldsymbol{z}^o & \text{in } \Omega, \\[2mm] \boldsymbol{u}^m = \boldsymbol{p}^m = 0 & \text{on } \partial\Omega, \end{cases} \tag{14}$$

We shall prove in the following theorem that the algorithm above is convergent if N is not too small.

Theorem 4.1. *There exists a positive constant δ, such that for $N > \delta$,*

$$\{\boldsymbol{p}^m, \boldsymbol{u}^m\} \to \{\boldsymbol{p}, \boldsymbol{u}\} \text{ strongly in } (H_0^1(\Omega))^n \cap (H^2(\Omega))^n,$$

where $\boldsymbol{p}$ and $\boldsymbol{u}$ are solutions of the optimality system (13).

Proof. With the notation,

$$\boldsymbol{P}^m = \boldsymbol{p}^m - \boldsymbol{p} \qquad \text{and} \qquad \boldsymbol{U}^m = \boldsymbol{u}^m - \boldsymbol{u},$$

from (13) and (14) we have

$$\begin{cases} A\boldsymbol{U}^m = -\frac{1}{N}\boldsymbol{P}^{m-1} & \text{in } \Omega, \\ A\boldsymbol{P}^m = \boldsymbol{U}^m & \text{in } \Omega, \\ \boldsymbol{U}^m = \boldsymbol{P}^m = 0 & \text{on } \partial\Omega. \end{cases} \tag{15}$$

Taking inner product of the first equation with $\boldsymbol{U}^m$ and the second with $\boldsymbol{P}^m$ and integrating over Ω, we obtain by the use of (4) after integration by parts,

$$\sigma(\boldsymbol{U}^m, \boldsymbol{U}^m) = -\frac{1}{N}\int_\Omega \boldsymbol{P}^{m-1}\cdot\boldsymbol{U}^m\,dx, \tag{16}$$

$$\sigma(\boldsymbol{P}^m, \boldsymbol{P}^m) = \int_\Omega \boldsymbol{U}^m\cdot\boldsymbol{P}^m\,dx. \tag{17}$$

Taking the energy norm of (16) and (17), we get

$$\|\boldsymbol{U}^m\|_\sigma^2 \le \frac{1}{2N}\left\{\|\boldsymbol{P}^{m-1}\|_0^2 + \|\boldsymbol{U}^m\|_0^2\right\},$$

$$\|\boldsymbol{P}^m\|_\sigma^2 \le \frac{1}{2}\left\{\alpha\,\|\boldsymbol{U}^m\|_0^2 + \frac{1}{\alpha}\,\|\boldsymbol{P}^m\|_0^2\right\}.$$

In these relations we have used the elementary inequality, $ab \le \frac{1}{2}(\alpha\,a^2+b^2/\alpha)$, where α is real positive number. Since the norms $\|\cdot\|_\sigma$ and $\|\cdot\|_1$ are equivalent in $(H_0^1(\Omega))^n$, by (5) we have

$$C^2\|\boldsymbol{U}^m\|_1^2 \le \frac{1}{2N}\left\{\|\boldsymbol{P}^{m-1}\|_0^2 + \|\boldsymbol{U}^m\|_0^2\right\} \le \frac{K}{2N}\left\{\|\boldsymbol{P}^{m-1}\|_1^2 + \|\boldsymbol{U}^m\|_1^2\right\},$$

$$C^2\|\boldsymbol{P}^m\|_1^2 \le \frac{1}{2}\left\{\alpha\,\|\boldsymbol{U}^m\|_0^2 + \frac{1}{\alpha}\,\|\boldsymbol{P}^m\|_0^2\right\} \le \frac{K}{2}\left\{\alpha\,\|\boldsymbol{U}^m\|_1^2 + \frac{1}{\alpha}\,\|\boldsymbol{P}^m\|_1^2\right\}.$$

In the second part of the above relations, we have used the Poincaré inequality, where the constant K depends on Ω only. Hence we have

$$(2NC^2 - K)\|\boldsymbol{U}^m\|_1^2 \le K\|\boldsymbol{P}^{m-1}\|_1^2,$$

$$(2\alpha\,C^2 - K)\|\boldsymbol{P}^m\|_1^2 \le K\alpha^2\|\boldsymbol{U}^m\|_1^2.$$

Suppose that $\{N,\alpha\} \ge r/2$, where $r = K/C^2$, then by combining the above two inequalities we obtain

$$\|\boldsymbol{P}^m\|_1^2 \le \gamma\|\boldsymbol{P}^{m-1}\|_1^2, \tag{18}$$

where

$$\gamma = \frac{\alpha^2 r^2}{(2\alpha - r)(2N - r)}.$$

The inequality (18) is valid for $m = 1, 2, \ldots$, so we have

$$\|\boldsymbol{P}^m\|_1^2 \le \gamma^m \|\boldsymbol{P}^0\|_1^2. \tag{19}$$

Since $\boldsymbol{P}^0 \in (H_0^1(\Omega))^n$, if $\gamma < 1$ we conclude that as $m \to \infty$, the sequence $\{\boldsymbol{P}^m\}$ converges strongly to zero in $(H_0^1(\Omega))^n \cap (H^2(\Omega))^n$, by the use of elliptic regularity. The convergence of the sequence $\{\boldsymbol{U}^m\}$ also follows in exactly the same argument.

To conclude the proof of the theorem, we need to determine the constants α and N so that the condition $\gamma < 1$ is satisfied. Moreover, the value of α will be chosen in such a way that the lower bound for N is as small as possible. Since $\alpha > r/2$, let $\alpha = r/2 + \varepsilon$ for $\varepsilon > 0$, then the condition $\gamma < 1$ is equivalent to

$$N > \frac{r}{2}\Big(1 + \frac{r}{2\varepsilon}\big(\frac{r}{2} + \varepsilon\big)^2\Big).$$

Let the right hand side be denoted by δ, which takes its minimal value at $\varepsilon = r/2$. With this choice, we have $\alpha = r$ and

$$N > \delta = \frac{r}{2}\big(1 + r^2\big), \qquad r = \frac{K}{C^2}, \tag{20}$$

which ensures the condition $\gamma < 1$ and the theorem is proved. $\square$

Remark. *On the value of δ*:

By the assumptions (2) and (3) on the elasticity tensor C_{ijkl}, the elliptic operator A defined in (10) is self-adjoint and positive definite. Hence by the spectral theorem, the eigenvectors of A form an orthonormal basis of the Hilbert space $(H_0^1(\Omega))^n \cap (H^2(\Omega))^n$, and the eigenvalues $\{\lambda_m\}$ form an increasing sequence of positive real numbers, *i.e.*, $0 < \lambda_m \le \lambda_{m+1}$ for $m = 1, 2, \cdots$. Moreover, one can show that

$$\lambda_1 = \inf_{\boldsymbol{u} \ne 0} \frac{\|\boldsymbol{u}\|_\sigma^2}{\|\boldsymbol{u}\|_0^2}.$$

Hence, it follows from (5) that

$$\|\boldsymbol{u}\|_0^2 \le \frac{1}{\lambda_1}\|\boldsymbol{u}\|_\sigma^2 \le \frac{\bar{C}}{\lambda_1}\|\boldsymbol{u}\|_1^2. \tag{21}$$

In order words, the constant K in the Poincaré inequality can be taken as $\bar{C}/\lambda_1$, and we arrive at an estimated value of δ,

$$N > \delta = \frac{\bar{C}}{2C^2\lambda_1}\left(1 + \frac{\bar{C}}{C^4\lambda_1^2}\right).$$

In particular, for $A = -\Delta$, the Laplace operator, and $\Omega = (0,1) \times (0,1)$, we have $C = \bar{C} = 1$ and $\lambda_1 = 2\pi^2$, hence the constant N can be chosen as small as 0.0254 and since the above estimate is not optimal the lower limit of δ can be even smaller than this value as we shall see in the numerical example later.

4.1 Finite Element Approximation

Let V^h be the finite dimensional subspace of $(H_0^1(\Omega))^n$ in the finite element approximation with maximum mesh size h. The Galerkin formulation of the iterative problem (14) is given as follows:

Given $\boldsymbol{z}^0 \in (L^2(\Omega))^n$ and $\boldsymbol{p}_h^0 = 0$, find $\boldsymbol{u}_h^m \in V^h$ and $\boldsymbol{p}_h^m \in V^h$ such that for all $\boldsymbol{v}_h \in V^h$,

$$\begin{cases} \sigma(\boldsymbol{u}_h^m, \boldsymbol{v}_h) = (F(\boldsymbol{p}_h^{m-1}), \boldsymbol{v}_h)_0, \\ \sigma(\boldsymbol{p}_h^m, \boldsymbol{v}_h) = (\boldsymbol{u}_h^m - \boldsymbol{z}^o, \boldsymbol{v}_h)_0. \end{cases} \quad (22)$$

By employing finite element basis functions in V^h, (22) are systems of linear algebraic equations.

In Theorem 1, we have proved for $N > \delta$ that the numerical solution $\{\boldsymbol{u}^m, \boldsymbol{p}^m\}$ for uncoupled system (14) converges to the solution $\{\boldsymbol{u}, \boldsymbol{p}\}$ of the optimality system. In the next theorem, we claim that the numerical solution of system (22), $\{\boldsymbol{u}_h^m, \boldsymbol{p}_h^m\}$, obtained by the finite element method also converges to $\{\boldsymbol{u}, \boldsymbol{p}\}$.

Theorem 4.2. *Let $N > \delta$, then for $m \to \infty$ and $h \to 0$,*

$$\{\boldsymbol{u}_h^m, \boldsymbol{p}_h^m\} \longrightarrow \{\boldsymbol{u}, \boldsymbol{p}\} \quad in \ (H_0^1(\Omega))^n \cap (H^2(\Omega))^n,$$

where δ is the constant defined in Theorem 4.1.

Since $\boldsymbol{u}^m$ and $\boldsymbol{p}^m$ are solutions in $(H_0^1(\Omega))^n \cap (H^2(\Omega))^n$ of the optimality system, the proof follows from the standard error estimate [1, 9, 11] and the triangular inequality.

5   Eigenfunction Expansion Method

In the previous section, we have proposed an iterative method for which the approximate solution $\{\boldsymbol{u}_h^m, \boldsymbol{p}_h^m\}$ converges strongly to the solution $\{\boldsymbol{u}, \boldsymbol{p}\}$ in $(H_0^1(\Omega))^n \cap (H^2(\Omega))^n$, provided that the constant $N > \delta$, meaning, N is not allowed to be arbitrary small. In our numerical calculations, we have found that such a restriction on N could be quite unsatisfactory in practical solutions. In order to find the solution of the optimality system free from such a restriction, in the following, we shall analyze the problem via the eigenfunction expansion method.

Since the operator A is positive-definite and self-adjoint defined in the Hilbert space $H = (H_0^1(\Omega))^n \cap (H^2(\Omega))^n$. By the spectral theorem, H admits a complete orthonormal basis of eigenfunctions $\{\boldsymbol{\varphi}_m\}$ of A and the corresponding eigenvalues λ_m can be arranged in ascending order,

$$0 < \lambda_m \leq \lambda_{m+1}, \quad m = 1, 2, 3 \ldots.$$

Consequently, $\{\boldsymbol{u}, \boldsymbol{p}\} \in H$ can be expressed in an eigenfunction expansion of the form:

$$\boldsymbol{u}(x) = \sum_{m=1}^{\infty} u_m \boldsymbol{\varphi}_m(x), \qquad \boldsymbol{p}(x) = \sum_{m=1}^{\infty} p_m \boldsymbol{\varphi}_m(x). \tag{23}$$

Therefore, we have

$$A\boldsymbol{u}(x) = \sum_{m=1}^{\infty} u_m \lambda_m \boldsymbol{\varphi}_m(x), \qquad A\boldsymbol{p}(x) = \sum_{m=1}^{\infty} p_m \lambda_m \boldsymbol{\varphi}_m(x). \tag{24}$$

Substituting into the equation $(13)_1$, we obtain

$$\sum_{m=1}^{\infty} u_m \lambda_m \boldsymbol{\varphi}_m(x) = -\frac{1}{N} \sum_{m=1}^{\infty} p_m \boldsymbol{\varphi}_m(x),$$

or

$$\sum_{m=1}^{\infty} \left(u_m \lambda_m + \frac{p_m}{N} \right) \boldsymbol{\varphi}_m(x) = 0,$$

which implies that

$$p_m = -u_m \lambda_m N. \tag{25}$$

Similarly, from the equation $(13)_2$, we obtain

$$\sum_{m=1}^{\infty} p_m \lambda_m \boldsymbol{\varphi}_m(x) = \sum_{m=1}^{\infty} (u_m - z_m) \boldsymbol{\varphi}_m(x),$$

or

$$\sum_{m=1}^{\infty} \big(p_m \lambda_m - (u_m - z_m) \big) \boldsymbol{\varphi}_m(x) = 0,$$

where z_m are the Fourier coefficients in the eigenfunction expansion of the given objective function $\boldsymbol{z}^0 \in (L^2(\Omega))^n$:

$$\boldsymbol{z}^0(x) = \sum_{m=1}^{\infty} z_m \boldsymbol{\varphi}_m(x). \tag{26}$$

Hence, we have

$$u_m = z_m + p_m \lambda_m. \tag{27}$$

From (25) and (27), we obtain the Fourier coefficients u_m and p_m,

$$u_m = \frac{z_m}{N \lambda_m^2 + 1}, \qquad p_m = -\frac{N \lambda_m z_m}{N \lambda_m^2 + 1},$$

and the solutions $\boldsymbol{u}(x)$ and $\boldsymbol{p}(x)$ of the optimality system (13) are given explicitly by

$$\boldsymbol{u}(x) = \sum_{m=1}^{\infty} \frac{z_m}{N \lambda_m^2 + 1} \boldsymbol{\varphi}_m(x), \tag{28}$$

$$\boldsymbol{p}(x) = -\sum_{m=1}^{\infty} \frac{N \lambda_m z_m}{N \lambda_m^2 + 1} \boldsymbol{\varphi}_m(x). \tag{29}$$

From these solutions we can easily see that when $N \to 0$ the function $\boldsymbol{p}(x)$ tends to zero, while the solution $\boldsymbol{u}(x)$ converges to the objective function $\boldsymbol{z}^0(x)$. Moreover, by (28) it follows that $\|\boldsymbol{u}\|_0 < \|\boldsymbol{z}^0\|_0$ for any $N > 0$.

On the other hand, the optimal control $\boldsymbol{g} \in (L^2(\Omega))^n$ admits a representation in Fourier series, and since $\boldsymbol{g} = -\boldsymbol{p}/N$ we obtain

$$\boldsymbol{g}(x) = \sum_{m=1}^{\infty} \frac{\lambda_m z_m}{N \lambda_m^2 + 1} \boldsymbol{\varphi}_m(x), \tag{30}$$

which tends to the expected solution given by the eigenfunction expansion of $A\boldsymbol{z}^0(x)$ in the limit as $N \to 0$.

6 Numerical Results

For convenience, we shall consider the Laplace operator, $A = -\Delta$, in the unit square $\Omega = (0,1) \times (0,1)$ as an example. Mathematically, it is a very special case of the linear elasticity operator, in which the components of the equilibrium equation become independent of each other and the problem (1) can be regarded as two independent problems of scalar equations for the individual components. But for the purpose of examining the qualitative behavior of the optimal control solutions, our numerical calculation, both in finite element iterative approximation and in Fourier series expansion, will be illustrated for the case of Laplace operator. For more general elliptic operators or more general domains, eigenfunctions for spectral representations can be constructed numerically.

Since $u = p = 0$ on $\partial\Omega$ the eigenvalues and the eigenfunctions of the Laplace operator are well-known and they are given by

$$\lambda_{mn} = (n^2 + m^2)\pi^2, \qquad \varphi_{mn} = \sin m\pi x \, \sin n\pi y.$$

The objective function can then be represented by

$$z^0(x,y) = \sum_{m,n=1}^{\infty} z_{mn} \sin m\pi x \, \sin n\pi y.$$

Substituting into (28) and (30), we have the approximation of the objective function,

$$u(x,y) = \sum_{m,n=1}^{\infty} \frac{z_{mn}}{N(m^2+n^2)^2\pi^4 + 1} \sin m\pi x \, \sin n\pi y,$$

and the function of optimal control,

$$g(x,y) = \sum_{m,n=1}^{\infty} \frac{(m^2+n^2)\pi^2 z_{mn}}{N(m^2+n^2)^2\pi^4 + 1} \sin m\pi x \, \sin n\pi y.$$

The exact solution of optimal control is then given by the function in the limit when $N \to 0$,

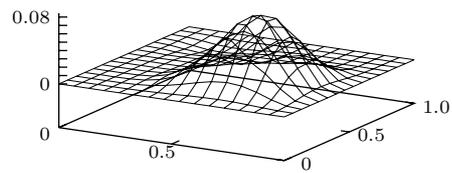$$g(x,y) = \sum_{m,n=1}^{\infty} (m^2+n^2)\pi^2 z_{mn} \sin m\pi x \, \sin n\pi y.$$
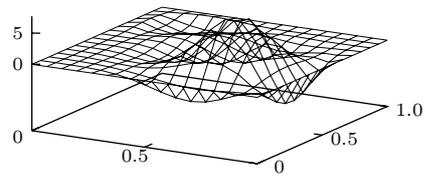
Fig. 1. Objective function $z^0(x, y)$
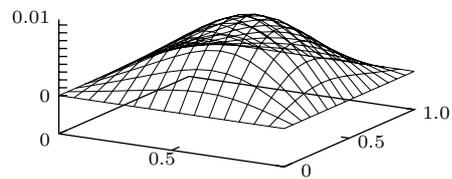

Fig. 2. Exact optimal control $g(x, y)$
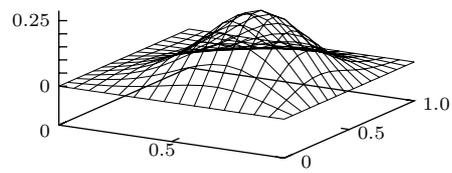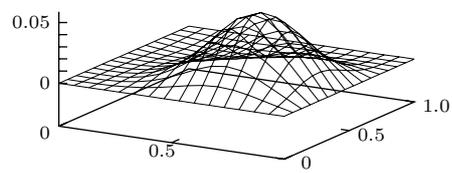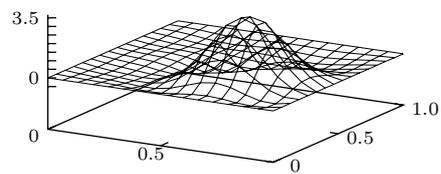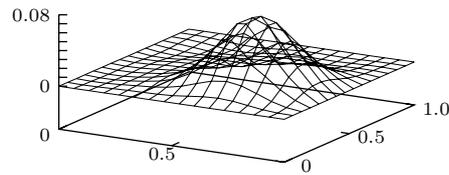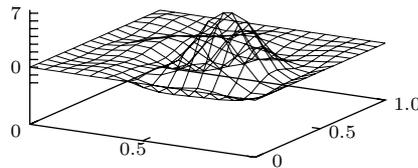

Fig. 3. $u(x, y)$ for $N = 0.005$

Fig. 4. $g(x, y)$ for $N = 0.005$

Fig. 5. $u(x, y)$ for $N = 0.00005$

Fig. 6. $g(x, y)$ for $N = 0.00005$

Fig. 7. $u(x, y)$ for $N = 0$

Fig. 8. $g(x, y)$ for $N = 0$

For numerical calculations, we consider a prescribed objective function given by

$$z^0(x, y) = \left(16xy(1 - x)(1 - y)(x - y^2)\right)^3.$$

In Fig. 1 and Fig. 2, the objective function $z^0(x, y)$ and the exact optimal control $g(x, y)$ given above are shown. In finite element approximation, we have taken a mesh of 15×15 square elements. Remember that there is a lower limit of the parameter N. In the present case we have found that this limit is approximately equal to 5×10^{-3} (Note that it is much smaller than the estimated value given before) for which convergence is ensured in 20 iterations. Convergence is much faster for greater values of N. However, from Fig. 3 and Fig. 4, we can see that with this smallest allowable value of N, the numerical results are still quite far from a good approximation to the exact solutions by comparing the scales shown in the graphics. It is obvious from the graphics

that $\|u\|_0$ is much too small compared to the expected value $\|z^0\|_0$.

On the other hand, there is no restriction on the value of N for the method of Fourier series expansion. Approximation by Fourier series is calculated by a sum of 22 terms in eigenfunctions φ_{mn} for $m^2 + n^2 < 6^2$. Numerical solutions are obtained for values of $N = 5 \times 10^{-3}$, 5×10^{-5} and also for $N = 0$. which represents the exact solution to the problem. The figures for $N = 5 \times 10^{-3}$ are not shown here because they are practically identical to Figures 3 and 4 for the results of iterative approximation. Fig. 5, Fig. 6 and Fig. 7, Fig. 8 show the graphics for $N = 5 \times 10^{-5}$ and $N = 0$ respectively. We can see the gradual improvement of the approximation for decreasing values of N and an excellent agreement with the exact solutions shown in Fig. 1 and Fig. 2 for the case of $N = 0$, for which the series are calculated with a finite sum of the first 22 terms only.

References

[1] Ciarlet, P. G.: The Finite Element Method for Elliptic Problems, North Holand, Amsterdam (1987).

[2] Fichera, G.: Existence Theorems in Elasticity, in Handbuch der Physik, Band VIa/2, Edited by C. Truesdell, Springer-Verlag (1972).

[3] Gill, S. I., Murray, W., Wright, M. H.: Practical Optimization, Academic Press (1988).

[4] Gunzburger, M. D.; Lee, H-C.: Analysis and approximation of optimal control problems for first-order elliptic systems in three dimensions. Appl. Math. Comput., 100, 49-70 (1999).

[5] Haslinger, J.; Neittaanmäki, P.: Finite Element Approximation for Optimal Shape Design: Theory and Applications, John Wiley & sons (1988).

[6] Kelley, C. T., Sachs, E. W.: Multilevel algorithms for constrained compact fixed point problems, SIAM J. Sci. Comput. 15, 645-667 (1994).

[7] Lions, J. L.: Optimal control of systems governed by partial differential equations. Springer-Verlag, New York (1972).

[8] Neittaanmäki, P., Tiba, D.: Optimal Control of Nonlinear Parabolic Systems, Marcel Dekker, Inc. New York (1994).

[9] Oden, J. T., Reddy, J. N.: Variational Methods in Theoretical Mechanics, Spring-Verlag (1976).

[10] Pironneau, O.: Optimal Shape Design for Elliptic Systems, Springer-Verlag, Berlin (1984).

[11] Strang, G., Fix, G. J.: An Analysis of the Finite Element Method, Prentice-Hall (1973).