Optimal control of radiating panels via sequential quadratic programming

Igor Patlashenko
Performance Group, EMC Corporation, Hopkinton, Massachusetts, USA
Dan Givoli
Department of Aerospace Engineering and Asher Space Institute, Technion—Israel Institute of Technology, Haifa, Israel

Received June 2001
Revised October 2001
Accepted October 2001

Keywords *Optimal control, Heat transfer, Finite elements*

Abstract *The optimal control of the steady-state temperature distribution in radiating panels using control heat sources is considered. The problem has important applications in the thermal control of space structures. A mathematical model leads to an elliptic nonlinear optimal control problem. A numerical optimal control method, based on finite element (FE) discretization and sequential quadratic programming (SQP), is employed. Results are presented for some specific examples.*

1 Introduction

The thermal control of space structures is a very important ingredient in space structure design and analysis, from two main reasons. First, space structures such as satellites sometimes carry instrumentation which is effective only in a certain range of temperatures. For example, some batteries cannot function when their temperature is below zero Celsius (Guelman *et al.*, 2000). Second, variations in the temperature field generate an elastic deformation in the structure which must satisfy the allowed working conditions of various mounted orientation-sensitive instruments and antennas. For these reasons, space structures sometimes require an especially detailed and accurate thermal control. Recent work on various aspects of optimal thermal control of space structures includes Lu and Nan (1993), Zhu *et al.* (1995) and Mattei (1998). The heat sources in this case are solar radiation, radiation from other planets, and heat generated inside the structure. Space structures emit thermal radiation to the environment, which is a significant nonlinear mechanism of heat transfer in space.

Various finite element (FE) schemes for the solution of thermal control problems have been used. Work that focuses directly on space structures includes Warren and Arelt (1991) and Chin *et al.* (1992). Related FE schemes for other applications have been proposed, e.g. in Gunzberger *et al.* (1993), Kim *et al.*

This work was partly supported by the Fund for the Promotion of Research at the Technion.

International Journal of Numerical Methods for Heat & Fluid Flow, Vol. 12 No. 1, 2002, pp. 47–64.

DOI 10.1108/09615530210413163

(1996), Ravindran (1997), Suzuki *et al.* (1996) and Manservisi (2000). The main body of work in this area is concerned with the control of time-dependent variables, and assumes linear governing equation and boundary conditions. On the other hand, in the present paper we consider the steady-state (elliptic) nonlinear optimal thermal control of a structure which radiates heat to the environment, such as a panel in a space structure (Givoli and Rand, 1995).

Most of the FE optimal control formulations that have been proposed involve the "adjoint state", which appears when the Pontryagin maximum principle is employed (Knowles, 1981). The problem's variables include, in addition to the primary state variables, the adjoint variables. This leads to a mixed FE formulation involving a linear algebraic system of dimension $2N^y$, where N^y is the number of primary degrees of freedom (see e.g. Gunzberger *et al.*, 1991; Grandhi *et al.*, 1993; Hou and Turner, 1995; Stavroulakis, 1995). Recently, a new general framework has been developed for the FE solution of optimal control problems governed by nonlinear elliptic partial differential equations (Givoli, 1999; Givoli and Patlashenko, 2000). In contrast to the FE schemes mentioned above, the approach in Givoli (1999) is a direct one, which does not involve adjoint variables. Computationally, this has the effect of leading to a simpler formulation and reducing the number of variables by a factor of two. The solution of the final discrete minimization problem is performed via sequential quadratic programming (SQP). This formulation does not employ the standard tools of classical control theory, but fits naturally into the framework of computational continuum mechanics.

In Givoli (1999) the general numerical procedure has been developed. In addition, the method has been applied to a simple scalar one-dimensional model problem. In Givoli and Patlashenko (2000) the method was applied to problems in nonlinear elasticity. In this paper we adapt and apply the FE–SQP optimal control methodology to the problem of optimal thermal control of a two-dimensional radiating panel. Temperature distribution in the panel is controlled via concentrated or distributed control heat sources which must satisfy given constraints.

Following is the outline of this paper. In Section 2 we give the detailed statement of the optimal control problem on the continuous level. In Section 3 we then introduce the FE–SQP scheme for the approximate solution of this problem. We discuss the computational aspects of the method, as well as the simplifications associated with the special unconstrained case. We present numerical results for specific examples in Section 4, and conclude the paper with remarks in Section 5.

2 Statement of the optimal-control problem

Let Ω be a finite two-dimensional spatial domain representing a thin panel, and let Γ be its closed one-dimensional boundary (see Figure 1). Let $y(\boldsymbol{x})$ be the unknown temperature field in the plate. The panel's surface is exposed to

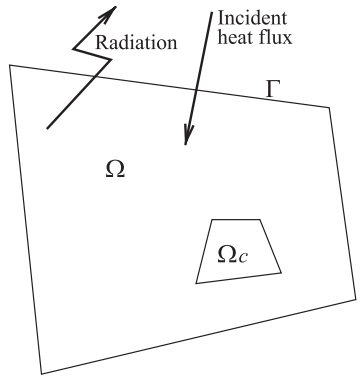

Figure 1.
Setup for the optimal control problem of a radiating panel

incoming heat flux. The panel conducts heat, exchanges heat with other parts of the structure through Γ, and radiates heat out through the surface. The material the panel is made of is assumed to be linear (i.e. its thermal conductivity and radiation coefficient do not depend on temperature) but may be thermally anisotropic. We ignore transient response and consider the steady-state temperature distribution in the panel.

The statement of the problem consists of three ingredients:

(1) governing equation and boundary conditions,

(2) objective functional, and

(3) constraints on the control.

2.1 Governing equation and boundary conditions

The governing equation and boundary conditions are

$$-\nabla \cdot (\boldsymbol{\kappa} \nabla y) + C_R y^4 = f(\boldsymbol{x}) + \beta(\boldsymbol{x}) u(\boldsymbol{x}) \quad \text{in} \quad \Omega, \tag{1}$$

$$y = g \quad \text{on} \quad \Gamma_g, \tag{2}$$

$$\boldsymbol{\kappa} \nabla y \cdot \boldsymbol{n} = h \quad \text{on} \quad \Gamma_h. \tag{3}$$

Here $\boldsymbol{\kappa} = [\kappa_{ij}]$ is the given thermal conductivity tensor for the panel's material, C_R is the given radiation coefficient, f is the given incident flux on the panel's surface, Γ_g and Γ_h are, respectively, the parts of Γ on which temperature and normal heat flux are prescribed, $\boldsymbol{n}$ is the normal unit vector pointing out of Γ_h, and g and h are given functions. The term $\beta(\boldsymbol{x})u(\boldsymbol{x})$ in the right side of (1) is the control term; $u(\boldsymbol{x})$ is an unknown control flux, and $\beta(\boldsymbol{x})$ is a given characteristic function which assumes the values 0 and 1 only. The support of β, namely the region where $\beta = 1$, is denoted Ω_c and is called the control-flux region (see Figure 1). Typically Ω_c is much smaller than the entire domain Ω. The two

primal unknowns of the problem are the temperature $y(\boldsymbol{x})$ for $\boldsymbol{x} \in \Omega$ and the control flux $u(\boldsymbol{x})$ for $\boldsymbol{x} \in \Omega_c$.

2.2 Objective functional

We consider the quadratic objective functional

$$C[y] = \int_{\Omega} \alpha(\boldsymbol{x})(y(\boldsymbol{x}) - y_0(\boldsymbol{x}))^2 d\boldsymbol{x}, \tag{4}$$

where $\alpha(\boldsymbol{x}) \geq 0$ and $y_0(\boldsymbol{x})$ are given functions. The interpretation of (4) is that the temperature distribution $y(\boldsymbol{x})$ is to be as close as possible to the given reference temperature function $y_0(\boldsymbol{x})$, where the relative importance of this requirement in various locations in Ω is determined by the given weighting $\alpha(\boldsymbol{x})$. The closeness of y to y_0 is enforced in (4) in the least-square sense. The simplest case is of course $\alpha(\boldsymbol{x}) \equiv 1$ and $y_0(\boldsymbol{x}) = \text{const.}$, i.e. the temperature in the whole panel is to be as close as possible to the constant temperature y_0.

2.3 Constraints on the control

We defined bounds which limit the size of the control flux u, i.e.

$$u \in \delta_u = u | u_{\min}(\boldsymbol{x}) \leq u(\boldsymbol{x}) \leq u_{\max}(\boldsymbol{x}), \text{for} \boldsymbol{x} \in \Omega_c\}, \tag{5}$$

where $u_{\min}(\boldsymbol{x})$ and $u_{\max}(\boldsymbol{x})$ are given functions. (Either one may attain an infinite value if one of the constraints is to be removed.) One may be interested also in the *unconstrained* case, which simplifies the formulation. We shall relate to this case later when considering the numerical optimal control scheme.

2.4 Statement of the problem

The optimal control problem to be solved is: *find $y(\boldsymbol{x})$, $\boldsymbol{x} \in \Omega$ and $u(\boldsymbol{x})$, $\boldsymbol{x} \in \Omega_c$, which satisfy the nonlinear equation (1), the boundary conditions (2) and (3), and the constraint (5), such that C[y] given by (4) is minimized.*

2.5 Design and closed-loop control

The global design problem involves two aspects that have not been dealt with in the statement of the problem above: (a) determining the location and shape of the control region Ω_c (or equivalently the characteristic function $\beta(\boldsymbol{x})$ in (1)); and (b) closing the control loop. We shall not deal with these two important issues in the present paper. However, the optimal-control scheme proposed here can serve as the basis for a control-location optimization scheme on one hand, and for closed-loop control design on the other hand. The optimal location of the controllers can also be performed manually, using the proposed optimal-control scheme in a repetitive manner. In other words, the analyzer may try a few reasonable control configurations (i.e. a few choices for Ω_c), apply the scheme developed here for each of them separately to obtain the control

functions, temperature distribution and cost $C[y]$ in each case, and finally choose the one configuration which yields the lowest value of the cost.

3 Computational scheme

3.1 Finite element discretization

The Galerkin finite element (FE) method is applied to the problem under consideration. Both y and u are approximated via FE shape function expansions, i.e.

$$y(\boldsymbol{x}) \simeq y^h(\boldsymbol{x}) = \sum_{I \in E_y} d_I \psi_I(\boldsymbol{x}), \tag{6}$$

$$u(\boldsymbol{x}) \simeq u^h(\boldsymbol{x}) = \sum_{A \in E_u} U_A \phi_A(\boldsymbol{x}). \tag{7}$$

Here E_y and E_u are the sets of temperature nodes and control nodes, ψ_I and ϕ_A are the temperature and control shape functions, and d_I and U_A are the temperature and control nodal values, respectively. The global vectors whose entries are all the nodal temperatures and all control fluxes are denoted $\boldsymbol{d}$ and $\boldsymbol{U}$, respectively. We denote the total number of temperature degrees of freedom by N^y, and the total number of control degrees of freedom by N^u.

In order to obtain a finite element discretization for (1)–(3) we first write these equations in a weak form:

$$a(w,y) - (w,f) + (w,h)_\Gamma + b(w,u). \tag{8}$$

Here,

$$a(w,y) \int_\Omega [\boldsymbol{\nabla} w \cdot \boldsymbol{\kappa} \boldsymbol{\nabla} y + w C_R y^4] d\boldsymbol{\Omega}, \tag{9}$$

$$(w,f) = \int_\Omega w f d\boldsymbol{\Omega}, \tag{10}$$

$$(w,h)_\Gamma = \int_{\Gamma_h} w h d\boldsymbol{\Gamma}, \tag{11}$$

$$b(w,u) = \int_\Omega w \beta u d\boldsymbol{\Omega} \equiv \int_{\Omega_c} w u d\boldsymbol{\Omega}. \tag{12}$$

We seek $y \in \mathcal{S}$, and we require that (8) hold for all weighting functions $w \in \mathcal{S}_0$, where $\mathcal{S}$ is the appropriate trial space and $\mathcal{S}_0$ is its homogeneous counterpart (see Hughes, 1987).

Applying the approximations in (6) and (7) to the weak form of the problem results in a system of nonlinear algebraic equations, of the form

$$\boldsymbol{G}_0(\boldsymbol{d}) = \boldsymbol{F} + \boldsymbol{Q}\boldsymbol{U}. \tag{13}$$

Here $\boldsymbol{G}_0$ is the vector of "internal fluxes" (which is a *nonlinear* function of the temperature vector $\boldsymbol{d}$), and $\boldsymbol{F}$ is the background thermal load vector, both standard in nonlinear FE analysis. Their expressions are given by

$$\boldsymbol{G}_0 = (G_0)_I\} \quad (N^y \times 1) \quad (G_0)_I = a(\psi_I, y^h) \equiv a(\psi_I, \sum_{J \in E_y} d_J \psi_J), \tag{14}$$

$$\boldsymbol{F} = F_I\} \quad (N^y \times 1) \quad F_I = (\psi_I, f) + (\psi_I, h)_\Gamma. \tag{15}$$

Of course, in practice these vectors (and other global arrays appearing below) are calculated by applying the assembly operation to the analogous element-level arrays. The term $\boldsymbol{Q}\boldsymbol{U}$ in the right side of (13) is the control flux contribution to the thermal equilibrium equations. Note that both $\boldsymbol{d}$ and $\boldsymbol{U}$ are unknown. The matrix $\boldsymbol{Q}$ has the form

$$\boldsymbol{Q} = [Q_{IB}] \quad (N^y \times N^u) \quad Q_{IB} = \int_\Omega \psi_I \beta \phi_B d\Omega \equiv \int_{\Omega_c} \psi_I \phi_B d\Omega. \tag{16}$$

We substitute (6) into (4), and after some algebra obtain the discrete objective function,

$$C^h[\boldsymbol{d}, \boldsymbol{U}] = \boldsymbol{d}^T \boldsymbol{M} \boldsymbol{d} - 2\boldsymbol{L}^T \boldsymbol{d} + c_0, \tag{17}$$

where

$$\boldsymbol{M} = [M_{IJ}] \quad (N^y \times N^y) \quad M_{IJ} = \int_\Omega \psi_I \alpha \psi_J d\boldsymbol{x}, \tag{18}$$

$$\boldsymbol{L} = L_I\} \quad (N^y \times 1) \quad L_I = \int_\Omega \psi_I \alpha y_0 d\boldsymbol{x}. \tag{19}$$

In (17), the superscript T denotes transposition, and c_0 is a constant scalar which does not affect the minimization of C^h and therefore can be ignored.

The discrete counterpart of the constraint (5) is

$$\boldsymbol{T}\boldsymbol{U} \leq \boldsymbol{Z}, \tag{20}$$

where $\boldsymbol{T}$ is a given constant transformation matrix, $\boldsymbol{Z}$ is a given constant bound vector, and the vector inequality is to be interpreted in the component-wise sense, i.e. $\boldsymbol{a} \leq \boldsymbol{b}$ means $a_i \leq b_i$ for all i.

Note that the discrete optimal control problem can be posed as follows: *find* $\boldsymbol{d}$ *and* $\boldsymbol{U}$ *which satisfy the nonlinear system (13) and the constraint (20), such that* C^h *given by (17) is minimized.*

3.2 Sequential quadratic programming

The Newton iteration procedure is now applied to the nonlinear system (13). We denote the vector $\boldsymbol{d}$ at iteration i by $\boldsymbol{d}^{(i)}$. At iteration i+1 the solution vector is updated via

$$\boldsymbol{d}^{(i+1)} = \boldsymbol{d}^{(i)} + \Delta\boldsymbol{d}^{(i)} \tag{21}$$

The increment $\boldsymbol{\Delta d}^{(i)}$ is found by solving the linear system of equations,

$$\boldsymbol{K}^{(i)}\boldsymbol{\Delta d}^{(i)} = \boldsymbol{R}^{(i)} \tag{22}$$

Here $\boldsymbol{K}^{(i)}$ is the tangent stiffness (or thermal conductivity) matrix,

$$\boldsymbol{K}^{(i)} \equiv \left.\frac{\partial \boldsymbol{G}_0(\boldsymbol{d})}{\partial \boldsymbol{d}}\right|_{\boldsymbol{d}=\boldsymbol{d}^{(i)}} \tag{23}$$

and $\boldsymbol{R}^{(i)}$ is the residual vector obtained from (13),

$$\boldsymbol{R}^{(i)} = \boldsymbol{F} - \boldsymbol{G}_0(\boldsymbol{d}^{(i)}) + \boldsymbol{QU}. \tag{24}$$

Equations (21), (22) and (24) can be written as

$$\boldsymbol{d}^{(i+1)}(\boldsymbol{U}) = \hat{d}^{(i)} + (\boldsymbol{K}^{(i)})^{-1}\boldsymbol{QU} \tag{25}$$

where

$$\hat{d}^{(i)} = \boldsymbol{d}^{(i)} + (\boldsymbol{K}^{(i)})^{-1}(\boldsymbol{F} - \boldsymbol{G}_0(\boldsymbol{d}^{(i)})). \tag{26}$$

The vector $\hat{d}^{(i)}$ as defined by (26) is the current solution *with no control.*

Now we substitute (25) into (17), and after some algebra obtain,

$$\bar{C}^h[\boldsymbol{U}] = \boldsymbol{U}^T\boldsymbol{P}^{(i)}\boldsymbol{U} - 2\boldsymbol{U}^T\boldsymbol{B}^{(i)} + \text{const.} \tag{27}$$

where

$$\boldsymbol{P}^{(i)} = \boldsymbol{A}^{(i)T}\boldsymbol{M}\boldsymbol{A}^{(i)} \quad (N^u \times N^u) \tag{28}$$

$$\boldsymbol{A}^{(i)} = (\boldsymbol{K}^{(i)})^{-1}\boldsymbol{Q} \quad (N^y \times N^u) \tag{29}$$

$$\boldsymbol{B}^{(i)} = \boldsymbol{A}^{(i)T}(\boldsymbol{L} - \boldsymbol{M}\hat{d}^{(i)}) \quad (N^u \times 1) \tag{30}$$

From (27) and (20) we then obtain the quadratic programming (QP) problem:

$$\text{Given} \quad \boldsymbol{d}^{(i)}, \quad \text{find} \quad \min_{\boldsymbol{TU} \le \boldsymbol{Z}} [\boldsymbol{U}^T\boldsymbol{P}^{(i)}\boldsymbol{U} - 2\boldsymbol{U}^T\boldsymbol{B}^{(i)}]. \tag{31}$$

The QP problem (31) can be solved using a standard QP algorithm (Gill *et al.*, 1981; Luenberger, 1984). In the numerical examples of the next section, we shall use the Goldfarb–Idnani QP algorithm (Goldfarb and Idnani, 1983) for this purpose.

To summarize, the proposed method reduces the original optimal control problem into a sequence of QP problems, one in each Newton iteration. These problems are in turn solved by applying a standard QP algorithm. The whole solution process is described in "Solution procedure for the optimal control problem using the FE-SQP approach" below.

3.3 Computational aspects

We now make a few remarks regarding the computational aspects of this formulation.

Remark 1. Equations (26) and (29) involve the inverse of the tangent stiffness matrix $\boldsymbol{K}^{(i)}$. In practice, the inverse is never actually computed, but $\boldsymbol{K}^{(i)}$ is factorised, and then back substitution is performed to obtain $\boldsymbol{K}^{-1}(\boldsymbol{F}-\boldsymbol{G}_0)$ in (26) and $\boldsymbol{K}^{-1}\boldsymbol{Q}$ in (29). The latter involves back substitution for each column of the "right-hand-side vector" $\boldsymbol{Q}$.

Solution procedure for the optimal control problem using the FE–SQP approach:

(1) Calculate the constant matrices $\boldsymbol{Q}$, $\boldsymbol{M}$ and $\boldsymbol{L}$.

(2) Set $i = 0$. Start with an initial guess $\boldsymbol{d}^{(0)}$.

(3) Calculate $\boldsymbol{G}_0^{(i)}$, $\boldsymbol{F}$ and $\boldsymbol{K}^{(i)}$ (standard in nonlinear finite element analysis).

(4) Calculate $\hat{d}^{(i)}$, $\boldsymbol{A}^{(i)}$, $\boldsymbol{P}^{(i)}$ and $\boldsymbol{B}^{(i)}$.

(5) Solve the QP problem (31) to find the current control $\boldsymbol{U}$.

(6) Calculate the current solution $\boldsymbol{d}^{(i+1)} = \hat{d}^{(i)} + \boldsymbol{A}^{(i)}\boldsymbol{U}$.

(7) Check convergence of Newton iteration ($\|\boldsymbol{R}^{(i)}\| < \epsilon$. If converged, stop.

(8) $i \leftarrow i+1$. Return to step 3.

Remark 2. The matrix $\boldsymbol{P}^{(i)}$ appearing in the quadratic form in (31) is *symmetric and positive semidefinite*. Symmetry follows from (28) and from the symmetry of $\boldsymbol{M}$ defined in (18). Positivity is obtained from the simple calculation

$$\boldsymbol{v}^T\boldsymbol{P}\boldsymbol{v} = \boldsymbol{v}^T\boldsymbol{A}^T\boldsymbol{M}\boldsymbol{A}\boldsymbol{v} = (\boldsymbol{A}\boldsymbol{v})^T\boldsymbol{M}(\boldsymbol{A}\boldsymbol{v}) \geq 0. \tag{32}$$

The first inequality in (32) follows from (28), and the last inequality follows from the positivity of $\boldsymbol{M}$, which can easily be shown. Strict positive-definiteness of $\boldsymbol{P}^{(i)}$ is not obtained in general. QP algorithms for the problem (31) with a symmetric positive semidefinite matrix $\boldsymbol{P}$ are widely known (Gill *et al.*, 1981; Luenberger, 1984).

Remark 3. The matrix $\boldsymbol{P}$ and the operations in (28)–(30) are global in nature. This may have an undesirable effect on the computational effort needed in forming $\boldsymbol{P}$ and in the actual solution of the problem (31). However, this becomes a difficulty only when N^u (i.e. the dimension of the discrete control space) is large. Typically, N^u is much smaller than N^y. In other words, the total number of nodal control variables is much smaller than the total number of temperature degrees of freedom. Thus, the computational effort associated with the matrix $\boldsymbol{P}$ in (28) is not necessarily large, even when the discrete problem at hand is large.

Remark 4. It is important to note that operations with the N^y − dimensional arrays are local in nature, and can be performed on the element level. The matrices and vectors, $\boldsymbol{Q}$, $\boldsymbol{M}$, $\boldsymbol{G}_0^{(i)}$, $\boldsymbol{F}$ and $\boldsymbol{K}^{(i)}$, calculated in the proposed scheme, are formed in practice by the assembly of analogous element-level matrices and vectors, as usual in finite element analysis. The calculation of $\boldsymbol{M}\hat{\boldsymbol{d}}$ in (30) is also performed on the element level.

Remark 5. The QP problem (31) is solved *in each* Newton iteration by using a QP algorithm. All QP algorithms are iterative, and include some stopping criteria (Gill *et al.*, 1981; Luenberger, 1984). Since only the QP step in the *last* Newton iteration yields the final optimal control, it is reasonable to modify the QP stopping criterion tolerance during the Newton process, so that it becomes tighter towards the end of this process. This would guarantee that the computational effort associated with the QP step is not too large when the solution $\boldsymbol{d}^{(i)}$ is not sufficiently close to the converged solution.

Remark 6. The shape functions ψ_I are standard C^0 finite element functions, e.g. linear on triangular elements or bilinear on quadrilateral elements. On the other hand, the control shape functions ϕ_A need not be so regular. In the formulation above they appear only in the definition of the matrix $\boldsymbol{Q}$ (see (16)), and thus they are allowed to be *piecewise-continuous*. For example, one may use piecewise-constant ϕ_As where A indicates the midpoint of element A. This enables one to represent, for example, a spatial "bang–bang control" as the approximate solution. In fact, the ϕ_A may even be Dirac delta functions, since the integral in (16) exists in this case. Then $u^h = \Sigma_A U_A \phi_A$ represents "concentrated fluxes" with intensities U_A applied at the control nodal points.

Remark 7. There are four main error sources involved in the FE–SQP algorithm described above:

(a) the finite element discretization error associated with the variable y;

(b) the finite element discretization error associated with the control u;

(c) the error associated with the Newton iteration process;

(d) the error associated with the QP algorithm.

The first two errors are associated with the approximation of the continuous problem by a discrete one, whereas the last two errors are defined purely on the discrete (finite dimensional) level. Each of these errors is discussed separately

in Givoli (1999), which also provides basic convergence testing of the scheme for a one-dimensional model problem. A particularly interesting question is that of the interaction between the Newton iteration process and the QP minimization process; it is not obvious that this integration necessarily leads to convergence. However, in Givoli and Patlashenko (2000) we show, albeit in a simplified situation, that the Newton scheme combined with the QP solver retains its quadratic convergence property. For more details, see Givoli (1999) and Givoli and Patlashenko (2000).

3.4 The unconstrained case

Now we consider the case where there are *no constraints* except those related to the required regularity of the control functions. (Regarding the latter requirement, see Givoli, 1999.) In this case, the controls may be "constrained" through penalty terms in the objective functional. Thus, we replace the objective functional $C[y]$ in (4) by

$$C[y] = \int_\Omega \alpha(\boldsymbol{x})(y(\boldsymbol{x}) - y_0(\boldsymbol{x}))^2 d\boldsymbol{x} + \int_\Omega W(\boldsymbol{x}) u^2(\boldsymbol{x}) d\boldsymbol{x}, \tag{33}$$

where $\alpha(\boldsymbol{x}) \geq 0, y_0(\boldsymbol{x})$ and $W(\boldsymbol{x}) \geq 0$ are given functions. It is assumed that the penalty weight $W(\boldsymbol{x})$ does not vanish identically.

The continuous-level optimal control problem is now: *find* $y(\boldsymbol{x})$, $\boldsymbol{x} \in \Omega$ *and* $u(\boldsymbol{x})$, $\boldsymbol{x} \in \Gamma_c$, *which satisfy the nonlinear equation (1), and the boundary conditions (2) and (3), such that C[y] given by (33) is minimized.*

We introduce the FE approximations (6) and (7). After some algebra, we arrive at the following expressions for the discrete objective function, replacing (17) and (27):

$$C^h[\boldsymbol{d}, \boldsymbol{U}] = \boldsymbol{d}^T \boldsymbol{M} \boldsymbol{d} - 2\boldsymbol{L}^T \boldsymbol{d} + \boldsymbol{U}^T \boldsymbol{N} \boldsymbol{U}, \tag{34}$$

$$\bar{C}^h[\boldsymbol{U}] = \boldsymbol{U}^T \boldsymbol{P}^{(i)} \boldsymbol{U} - 2\boldsymbol{U}^T \boldsymbol{B}^{(i)} + \boldsymbol{U}^T \boldsymbol{N} \boldsymbol{U}. \tag{35}$$

Here

$$\boldsymbol{N} = [N_{AB}] \quad (N^u \times N^u) \quad N_{AB} = \int_\Omega \phi_A W \phi_B d\boldsymbol{x}, \tag{36}$$

and all the other arrays are defined as before. Since no constraints are imposed, the minimization problem (31) becomes:

$$\text{Given} \quad \boldsymbol{d}^{(i)}, \quad \text{find} \quad min[\boldsymbol{U}^T \boldsymbol{P}^{(i)} \boldsymbol{U} - 2\boldsymbol{U}^T \boldsymbol{B}^{(i)} + \boldsymbol{U}^T \boldsymbol{N} \boldsymbol{U}]. \tag{37}$$

A necessary condition for a minimum is $\partial \bar{C}^h / \partial U_A = 0$, for $A = 1, \ldots, N^u$. Hence (37) yields the $N^u -$ dimensional linear system of equations,

$$(\boldsymbol{P}^{(i)} + \boldsymbol{N}) \boldsymbol{U} = \boldsymbol{B}^{(i)}. \tag{38}$$

This linear system has to be solved anew in each Newton iteration.

We now make a few remarks.

Remark 1. The matrix $\boldsymbol{P}^{(i)}+\boldsymbol{N}$ appearing in the linear system (38) is *symmetric and positive definite*. Symmetry follows from the symmetry of $\boldsymbol{P}^{(i)}$ (see Remark 2 in Section 3.3) and from the symmetry of $\boldsymbol{N}$ defined in (36). Positive definiteness follows from the fact that $\boldsymbol{P}^{(i)}$ is positive semi-definite (see Remark 2 in Section 3.3), and from the positive-definiteness of $\boldsymbol{N}$, which can easily be shown. This guarantees that (38) can be solved uniquely.

Remark 2. The solution process is again described by the algorithm above, with two differences. First, in step 1, the constant matrix $\boldsymbol{N}$ (cf. (36)) is calculated too. This matrix is sparse and is formed in practice by the assembly of analogous element-level matrices. Second, in step 5, the linear system (38) is solved rather than a QP problem.

Remark 3. As in the constrained case (see Remark 6 of Section 3.3), the control shape functions ϕ_A may be *piecewise-continuous*, e.g. piecewise constant. However, as opposed to the constrained case, they cannot be as singular as Dirac delta functions, since $\boldsymbol{N}$ defined by (36) precludes this possibility. This does not mean that concentrated control fluxes cannot be applied, but that the objective functional must be modified in that case.

Suppose, for example, that $E_y = E_u$, namely the $y-$ nodes and the $u-$ nodes coincide (and hence $N^y = N^u$), and that the control is represented by nodal concentrated forces. Then (1) is replaced by the equation

$$-\nabla\cdot(\kappa\nabla y)+C_R y^4 = f(\boldsymbol{x})+\sum_{I\in E_y}\beta_I U_I\delta(\boldsymbol{x}-\boldsymbol{x}_I)\quad\text{in}\quad\Omega. \tag{39}$$

Here the U_I are unknown constants, representing the magnitude of the concentrated control fluxes, $\boldsymbol{x}_I$ is the location of node I, $\delta(\boldsymbol{x}-\boldsymbol{x}_I)$ is the Dirac delta with singularity at $\boldsymbol{x}_I$, and β_I are given binary constants defined by

$$\beta_I = \begin{cases} 0, & \text{No control flux is applied at node } I \\ 1, & \text{Control flux is applied at node } I \end{cases} \tag{40}$$

In addition, (33) is replaced by the objective functional

$$C[y] = \int_\Omega \alpha(\boldsymbol{x})(y(\boldsymbol{x})-y_0(\boldsymbol{x}))^2 d\boldsymbol{x} + \sum_{I\in E_u} W_I U_I^2. \tag{41}$$

Here the W_I are given non-negative constants (weights). Repeating the analysis with (39) and (41), it is easy to verify that (38) still holds, with all the matrices and vectors as defined previously, except that $\boldsymbol{Q}$ (cf. (16)) and $\boldsymbol{N}$ (cf. (36)) now become the diagonal matrices $\boldsymbol{Q}=\text{diag}\{\beta_I\}$ and $\boldsymbol{N}=\text{diag}\{W_I\}$. In deriving $\boldsymbol{Q}$, we made use of the "interpolation property" of finite element shape functions, namely $\psi_I(\boldsymbol{x}_J)=\delta_{IJ}$, where δ_{IJ} is the Kronecker delta.

We remark that in practice the use of concentrated controllers may be inefficient, especially in large structures where "precise control" is desired,

since it may require a large number of control points. However, computationally, the resulting scheme is especially simple and may serve for quick checking of candidate control configurations, which may lead to a prudent choice of the control region Ω_c.

4 Numerical examples

We apply the FE–SQP method to the problem of the optimal control of a square radiating panel of dimensions 20×20 made of an isotropic material. A Cartesian coordinate system (x_1,x_2) is used, whose origin is in the lower left corner of the panel. The nonlinear equation (1) holds in the square domain Ω, with $\boldsymbol{\kappa} = \kappa \boldsymbol{I}$, $\boldsymbol{I}$ being the identity matrix, and a constant incident flux f. On the boundary Γ the Dirichlet condition $y = g$ (see (2)) is imposed. The objective functional (4) is to be minimized, with $\alpha \equiv 1$ and y_0 a constant. Thus, the goal of the control is to make the temperature in the panel as close as possible, in the L_2 sense, to the constant temperature y_0. In all the numerical experiments that follow we take $C_R = 5.67 \cdot 10^{-8}$, $g = 300°$ and $y_0 = 300°$.

A finite element mesh with $20 \times 20 = 400$ bilinear square elements is used to discretize the panel. The control function u is assumed to be constant in each element. For the QP solution we use the Goldfarb–Idnani QP algorithm (Goldfarb and Idnani, 1983). Initially we apply no constraints on the control, but later we shall add the control constraint $|u| \leq u_{\max}$ (see (5)).

First we try the trivial case where $\Omega_c = \Omega$ (or $\beta(\boldsymbol{x}) \equiv 1$), namely all the elements are control elements. The result that we get is that the temperature everywhere is uniform and equal to $y_0 = 300°$, that the discrete objective function C^h is zero, and that the control value is uniform in all the elements and is equal to $C_R y_0^4 - f$. This result is indeed expected and coincides with the exact solution of the problem. This is evident from (4), where taking $y \equiv y_0$ makes the objective function vanish, and from (1), where, with $y \equiv y_0$ and $\beta \equiv 1$ we get $C_R y_0^4 = f + u$. Of course, once we reduce the size of the control region Ω_c, the resulting temperature distribution becomes non-uniform and the cost C^h cannot be made to vanish completely.

Now we consider the control configuration described in Figure 2. The control flux region, characterized by the function $\beta(\boldsymbol{x})$ in equation (1), consists of a rectangular "frame" of 32 elements as shown in the figure. Of course, in general the control-region location and shape should be the product of optimization; however, here we fix this region and solve the optimal-control problem as defined previously (see discussion on this issue in Section 2.5). We set $\boldsymbol{\kappa} = 10$ and $f = 140$, and apply no constraints on the control. Figure 3 shows the controlled and uncontrolled temperature variations along the line $x_2 = 10$ which is the central horizontal line in the panel.

As seen in the figure, while the uncontrolled temperature is everywhere lower than 300° and reaches a temperature close to 230° in a large portion of the panel, the temperature in the controlled panel is oscillatory around $y_0 = 300°$.

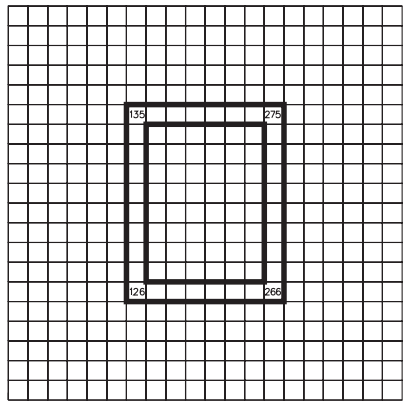

Figure 2.
Basic control flux region. The control function u is assumed to be constant in each of the 32 elements included in this region

The optimal control is effective in that the value of the discrete objective function C^h (see (17)) is reduced from about C^h = 1,302,000 without control to C^h = 437,000 with control, namely reduced by a factor of 3.

One may note the large temperature gradients obtained in the controlled case. This is an "overshoot" effect; the control is quite local, and in bringing the temperature in the panel to a higher level it has the side effect of increasing the temperature in the immediate control region well beyond the target temperature of 300°. This effect is even more pronounced in Figure 4, where the controlled and uncontrolled temperature variations are shown along the line x_2 = 5. The large temperature gradients may be undesired in practice, but this

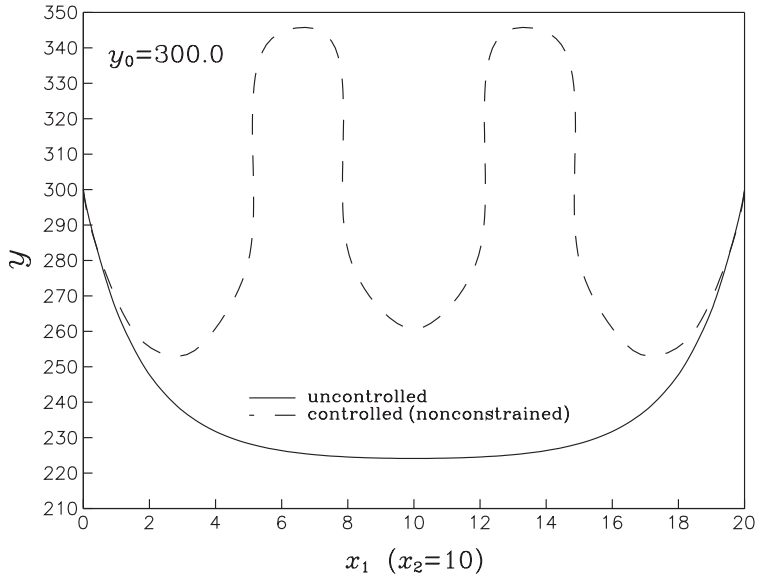

Figure 3.
Controlled and uncontrolled temperature variation along the central horizontal line in the panel, for κ = 10, f = 140, control region as in Figure 2 and without constraints

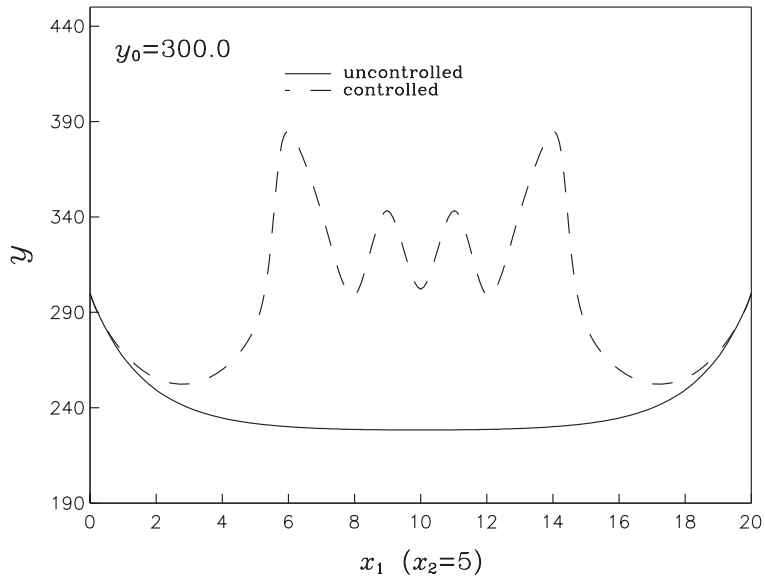

Figure 4.
Controlled and uncontrolled temperature variation along the line $x_2 = 5$, for $\kappa = 10$, $f = 140$, control region as in Figure 2 and without constraints

is not a consideration that has been defined as part of the optimization problem. One may prevent the appearance of large temperature gradients in the optimal solution by modifying the objective functional (4) to include an appropriate gradient term.

Next we set $\kappa = 100$ and $f = 700$, and again use the control configuration of Figure 2. With these parameters the uncontrolled temperature in the panel becomes dominated by the incoming flux rather than by heat radiation. Figures 5 and 6 show the controlled and uncontrolled temperature variations along the lines $x_2 = 10$ and $x_2 = 5$, respectively. As seen in the figures, the optimal control is indeed very effective. The value of the discrete objective function C^h is reduced from $C^h = 114{,}802$ without control to $C^h = 9615$ with control, a reduction by a factor of almost 12. Also, in this case local temperature changes are not as sharp as in the previous case.

Now we keep the parameter values as in Figures 5 and 6, still with no control constraints, and examine four different cases of the flux control region:

(a) control region is the entire rectangular "frame" shown in Figure 2;

(b) control region consists only of the two horizontal strips of the "frame" shown in Figure 2 (between elements 126 and 266, and between elements 135 and 275);

(c) control region consists only of the four corner elements of this "frame" (126, 266, 135 and 275);

(d) no control.

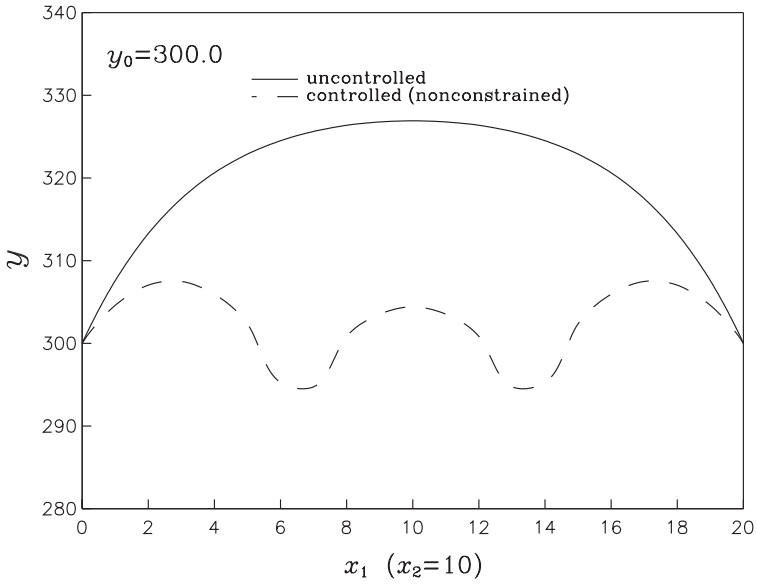

Figure 5.
Controlled and uncontrolled temperature variation along the line $x_2 = 10$, for $\kappa = 100$, $f = 700$, control region as in Figure 2 and without constraints

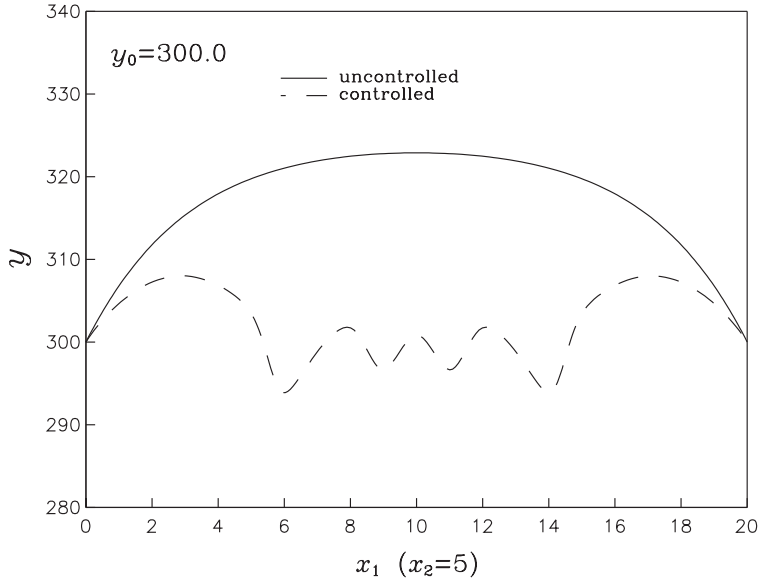

Figure 6.
Controlled and uncontrolled temperature variation along the line $x_2 = 5$, for $\kappa = 10$, $f = 140$, control region as in Figure 2 and without constraints

Table I compares the results obtained for these cases. For each case it gives the amount of control area used (in percentage of the total area of the panel), the value of the discrete objective function C^h, and the temperature interval (minimum and maximum temperature values). It is apparent that C^h is monotonely decreasing with increase of the control area, as expected. The effect on the temperature interval is more subtle. For example, the size of this interval is the same in cases C and D (27°), although C^h is much lower in case C; thus the effect the control has in case C is in "shifting" the average temperatures to a lower level.

Finally, we consider the effect of adding the control constraint $|u| \leq u_{\max}$. We fix the control region to be that of case A (the entire rectangular "frame"). Again the parameter values are as in Figures 5 and 6. Table II summarizes the results obtained for different values of $u_{\max}$. The cases $u_{\max} \rightarrow \infty$ and $u_{\max} \rightarrow 0$ are the previously considered unconstrained and uncontrolled cases, respectively.

It is clear that as the constraint becomes tighter the value of the discrete objective function C^h increases. This is the price that has to be paid for lowering the control energy. The table also shows the number of active constraints, N_{ac}. Altogether there are 32 constraints, one for each element in the control region (see Figure 2). However, only a partial set of the constraints may

Table I.
Comparison of results obtained for the four control-region cases. C^h is the value of the discrete objective function

Case	Control area (%)	C^h	Temperature interval
A	8	9615	[295, 308]
B	4	15,640	[293, 310]
C	1	181,64	[284, 311]
D	0	114,802	[300, 327]

Table II.
Comparison of results obtained for different control constraints

$u_{\max}$	C^h	N_{ac}
$\rightarrow \infty$	9615	0
40	10,351	4
30	10,574	12
15	11,297	24
10	24,837	32
5	85,750	32
$\rightarrow 0$	114,802	32

Note: The case $u_{\max} \rightarrow \infty$ is the unconstrained case, while the case $u_{\max} \rightarrow 0$ is the uncontrolled case. C^h is the value of the discrete objective function, and N_{ac} is the number of active constraints

be active; an active constraint is one for which $|u| \leq u_{\max}$ in the corresponding element. As seen in Table II, the number of active constraints increases gradually when the constraint becomes tighter, till finally all 32 constraints become active. Once this happens, any further decrease of $u_{\max}$ causes a dramatic increase in the objective function value (see last three lines of Table II).

5 Conclusion

Motivated by applications in the thermal control of space structures, we have presented a numerical scheme for the optimal control of the steady-state temperature distribution in radiating panels using control heat sources. It has been demonstrated that the Finite Element Sequential Quadratic Programming (FE–SQP) approach is effective for this type of problem.

The proposed method has the advantage that it does not make use of adjoint variables and thus leads to a simpler and more direct formulation with a reduced number of unknown variables. The underlying formulation does not employ the standard tools of classical control theory, but fits naturally into the framework of computational continuum mechanics.

Future research may include the adaptation and application of the scheme to other nonlinear optimal control problems in heat and fluid flow, in both two and three dimensions. A particularly interesting and important area where the scheme can be useful is that of the optimal control of processes involved in the growth of crystals from the melt (see e.g. Givoli *et al.*, 1996).

References

Chin, J.H., Panczak, T.D. and Fried, L. (1992), "Spacecraft thermal modelling", *Int. J. Numer. Meth. Engng.*, Vol. 35, pp. 641–53.

Gill, P.E., Murray, W. and Wright, M.H. (1981), *Practical Optimization*, Academic Press, New York, NY.

Givoli, D. (1999), "A direct approach to the finite element solution of elliptic optimal control problems", *Numer. Meth. Partial Differ. Equations*, Vol. 15, pp. 371–88.

Givoli, D. and Patlashenko, I. (2000), "Solution of static optimal control problems in nonlinear elasticity via quadratic programming", *Commun. Numer. Meth. Engng.*, Vol. 16, pp. 877–90.

Givoli, D. and Rand, O. (1995), "Minimization of the thermoelastic deformation in space structures undergoing periodic motion", *AIAA J. of Spacecraft and Rockets*, Vol. 32, pp. 662–9.

Givoli, D., Flaherty, J.E. and Shephard, M.S. (1996), "Simulation of Czochralski melt flows using parallel adaptive finite element procedures", *Modelling and Simulation in Material Science and Engineering*, Vol. 4, pp. 623–39.

Goldfarb, D. and Idnani, A. (1983), "A numerically dual method for solving strictly convex quadratic programs", *Mathematical Programming*, Vol. 27, pp. 1–33.

Grandhi, R.V., Kumar, A. and Chaudhary, A. (1993), "State-space representation and optimal control of non-linear material deformation using the finite element method", *Int. J. Num. Meth. Engng.*, Vol. 36, pp. 1967–86.

Guelman, M., Flohr, I., Ortenberg, F., Shachar, M., Shiryaev, A., Volfovsky, A. and Waler, R. (2000), "The Israeli Microsatellite Techsat for Scientific Research; development and in-orbit testing", *Acta Astronautica*, Vol. 46 No. 2–6, pp. 397–404.

Gunzberger, M.D., Hou, L. and Svobodny, T.P. (1991), "Finite element approximations of an optimal control problem associated with scalar Ginzburg–Landau equations", *Comput. & Math. with Applic.*, Vol. 21, pp. 123 ff.

Gunzburger, M.D., Steven, H.L. and Svobodny, T.P. (1993), "Approximation of boundary control problems for fluid flows with an application to control by heating and cooling", *Computers & Fluids*, Vol. 22, pp. 239–51.

Hou, L.S. and Turner, J.C. (1995), "Finite element approximation of optimal control problems for the Von Karman equations", *Numer. Methods Partial Differ. Equations*, Vol. 11, pp. 111–25.

Hughes, T.J.R. (1987), *The Finite Element Method*, Prentice-Hall, Englewood Cliffs, NJ.

Kim, T., Steadman, D., Hanagud, S.V. and Atluri, S.N. (1996), "On the feasibility of using thermal gradients for active control of interlaminar stresses in laminated composites", *Finite Elements Anal. & Design*, Vol. 23, pp. 349–63.

Knowles, G. (1981), *An Introduction to Applied Optimal Control*, Academic Press, New York, NY.

Lu, X. and Nan, Y. (1993), "Optimal return trajectory design of space vehicle", *Adv. Astro. Sci.*, Vol. 84, pp. 1379–86.

Luenberger, D.G. (1984), *Linear and Nonlinear Programming*, Addison-Wesley, Reading, MA.

Manservisi, S. (2000), "Optimal control approach to an inverse nonlinear elastic shell problem applied to car windscreen design", *Comput. Meth. Appl. Mech. Engng.*, Vol. 189, pp. 463–80.

Mattei, M. (1998), "Robust temperature trajectory following control system for space vehicle testing", *Proc. IEEE Conf. Decision & Control, IEEE*, Vol. 4, pp. 4214–19.

Ravindran, S.S. (1997), "Numerical solutions of optimal control for thermally convective flows", *Int. J. Numer. Meth. Fluids*, Vol. 25, pp. 205–23.

Stavroulakis, G.E. (1995), "Optimal prestress of cracked unilateral structures: finite element analysis of an optimal control problem for variational inequalities", *Compt. Meth. Applied Mech. Engng.*, Vol. 123, pp. 231–46.

Suzuki, S., Anju-Akira, A. and Kawahara, M. (1996), "Management of ground temperatures by bang–bang control based on finite element application", *Int. J. Numer. Meth. Engng.*, Vol. 39, pp. 885–901.

Warren, A.H. and Arelt, J.E. (1991), "Space station freedom module to truss system thermal/structural analysis", in, *Space Station ECLSS and Thermal Control*, SAE, Warrendale, PA No. 875, pp. 11–20.

Zhu, G., Grigoriadis, K.M. and Skelton, R.E. (1995), "Covariance control design for Hubble Space Telescope", *J. Guidance, Control & Dynamics*, Vol. 18, pp. 230–6.