Pergamon

Int. J. Heat Mass Transfer. Vol. 40, No. 17, pp. 4115–4127, 1997

Printed in Great Britain
0017–9310/97 $17.00+0.00

PII: S0017-9310(97)00044-6

Transient inverse heat conduction problem solutions via Newton's method

GREGORY A. DORAI and DANIEL A. TORTORELLI†

Department of Mechanical and Industrial Engineering, University of Illinois, Urbana-Champaign, IL-61801, U.S.A.

(*Received* 21 *June* 1996 *and in final form* 24 *January* 1997)

Abstract—Analytical first and second-order sensitivities are derived for a general, transient nonlinear problem and are then used to solve an inverse heat conduction problem (IHCP). The inverse analyses use Newton's method to minimize an error function which quantifies the discrepancy between the experimental and predicted responses. These Newton results are compared to results obtained from the first-order variable metric Broyton–Fletcher–Goldfarb–Shanno (BFGS) method. Inverse analyses are performed for both linear and nonlinear thermal systems. For linear systems, Newton's method converges in one iteration. For nonlinear systems, Newton's method sometimes diverges apparently due to a small radius of convergence. In these cases a combined BFGS–Newton's method is used to solve the IHCP. The unknown data fields are parameterized via the eigen basis of the Hessian to illustrate the need for regularization. Regularization is then incorporated and the IHCP is solved with Newton's method. All heat transfer analyses and sensitivity analyses are performed via the finite element method. © 1997 Elsevier Science Ltd.

1. INTRODUCTION

Inverse analyses have been applied in the fields of solid and structural mechanics [1, 2], heat transfer [3] and nuclear engineering [4], for both analysis and design optimization applications. Here we perform inverse analyses of heat conduction systems to predict immeasurable surface temperature/surface flux boundary conditions. These problems are referred to as inverse heat conduction problems (IHCP).

The existence and uniqueness of a bounded solution to the one-dimensional IHCP has been shown by Cannon [5]. However, it can be shown that this solution is often unstable [3], i.e. large solution fluctuations result from slight variations in the experimental data. Stolz [6] has been among the first to formulate and solve the inverse problem; he used the Duhamel integral [7]. Other methods based on the Duhamel integral have been presented in [8, 9]. These techniques utilize superposition and, thus, they are not valid for nonlinear systems. These earlier methods match temperature data exactly and, consequently, are very sensitive to small experimental data changes. Frank [10] solves the IHCP by parameterizing the unknown data and determining the unknown function coefficients by solving a least-squares problem. This method is further developed in [11], where simultaneous estimation of the parameters occurs and in [12], where sequential estimation of the parameters occurs. The least-squares method, commonly referred to as the function specification method, yields more accurate results as the number of unknown model parameters increase; however, this increase also causes instabilities. In the sequential method Beck *et al.* [12] avoid the instabilities by subdividing the time domain and then solving the least-squares problem over each of the subdomains. Their method makes assumptions about the behavior of the experimental data at future time steps. The stability of the sequential method is rigorously analyzed in [13] and it is extended to two spatial dimensions in [14].

The presence of instabilities necessitates the use of regularization in inverse analyses. The regularization function, which penalizes large variations in the unknown fields, is added to the least-squares error function. Applications of regularization to the IHCP have been discussed in [15–18]. The amount of regularization that is required has recently been studied in [19]. Other methods which are used to solve the IHCP include dynamic programming [20], the mollification method [21, 22] and solutions based on series expansions [23]. A comparison of these methods is summarized in [24].

Iterative techniques are required to solve the least squares problems encountered in the IHCP. The conjugate gradient method is employed to solve the IHCP in [16, 25]. In refs. [26, 27] the conjugate gradient method is coupled with the adjoint sensitivity analysis method to more efficiently solve the IHCP over multidimensional spatial domains. Newton's method is used in [28] to solve a steady-state IHCP and is shown to have superior convergence and stability characteristics than the Broyton–Fletcher–Goldfarb–

† Department of Theoretical and Applied Mechanics. Author to whom correspondence should be addressed.

NOMENCLATURE

A	area	q^n	heat flux
$\mathbf{B}$	matrix of polygon vertices	$\bar{q}$	scaling parameter
E	error	$\mathbf{R}$	residual vector
e	internal energy	t	time step
F	response functional (objective function)	$\mathbf{T}$	discretized temperature field
		T^*	weighting field
$\hat{F}$	augmented response functional	T	temperature
f	heat source	$\mathbf{u}$	discretized response vector
G	response functional	$\mathbf{v}$	eigenvector
$\bar{h}$	prescribed heat flux	v	parametric variable
h	plate height	w	parametric variable
k	thermal conductivity	$\mathbf{x}$	spatial position vector.
L	order of the B-spline curve interpolating time		
M	total number of time steps	Greek symbols	
$\bar{M}$	B-spline weighting function	β	scaling parameter
m	number of polygon vertices comprising the time curve	Γ	boundary of spatial region, Ω
		ν	Lagrange multiplier vector
N	total number of model parameters	Δ	difference operator
$\bar{N}$	B-spline weighting function	δ	Kronecker delta function
$\mathbf{n}$	normal direction	$\varepsilon_{\text{pert}}$	contant
n	number of polygon vertices comprising the position curve	ε_{reg}	constant
		ε_{tol}	constant
P	order of the B-spline curve interpolating position	λ	eigenvalue
		$\boldsymbol{\phi}$	model parameter vector
$\mathbf{q}$	heat flux vector	Ω	spatial region.

Shanno (BFGS) first-order variable metric method [29]. All of the above iterative methods require that the gradients of the error function be evaluated; see [30, 31] for an overview of the available gradient evaluation techniques.

The stability and confidence intervals of the IHCP solution is addressed in [32], where an eigenvalue analysis is performed on an approximation to the error function's Hessian to estimate the stability region. A somewhat similar analysis is presented in [28] where the eigenvalue analysis is performed on the actual Hessian.

In the following, first- and second-order design sensitivities are derived for a general nonlinear transient system and then specialized to a nonlinear heat conduction problem. An IHCP is solved via first- and second-order sensitivities in conjunction with Newton's method. The unknown data fields are then expressed via the error function Hessian's eigen basis. The need for regularization to stabilize the solution is shown via this eigenvector parameterization. Newton's method with regularization is used to solve the IHCP and the results are compared to results obtained from the BFGS method. Both linear and nonlinear conduction systems are studied. All heat transfer analyses and sensitivity analyses are performed via the finite element method. Although the simultaneous estimation procedure is used here, the methodology can also be used in the sequential estimation procedure.

2. TRANSIENT NONLINEAR SYSTEMS

The governing equations for transient nonlinear problems are discretized in time and space to form the residual vector $\mathbf{R}$ at each time step ${}^{n}t$ [31], i.e. the residual ${}^{n}\mathbf{R}$ at time ${}^{n}t$, which is written equivalently as

$$ {}^{n}\mathbf{R}({}^{n}\mathbf{u}, {}^{n-1}\mathbf{u}) = \mathbf{R}(\mathbf{u}({}^{n}t), \quad \mathbf{u}({}^{n-1}t), {}^{n}t) = 0 \qquad (1) $$

where ${}^{n}\mathbf{u}$ is the discretized response vector at time ${}^{n}t$ and ${}^{n-1}\mathbf{u}$ at the discretized response vector at time ${}^{n-1}t$. Here ${}^{n-1}\mathbf{u}$ is known from the prior analysis at time ${}^{n-1}t$, and ${}^{n}\mathbf{u}$ is evaluated from this analysis at time ${}^{n}t$. The Newton–Raphson method is used to perform the primal analysis, i.e. to solve equation (1) iteratively. The solution iterate ${}^{n}\mathbf{u}^{I}$ at iteration I is updated to ${}^{n}\mathbf{u}^{I+1} = {}^{n}\mathbf{u}^{I} + \delta\mathbf{u}$ where $\delta\mathbf{u}$ is determined by solving the linear problem

$$ \frac{\partial^{n}\mathbf{R}}{\partial^{n}\mathbf{u}}({}^{n}\mathbf{u}^{I}, {}^{n-1}\mathbf{u})\delta\mathbf{u} = -{}^{n}\mathbf{R}({}^{n}\mathbf{u}^{I}, {}^{n-1}\mathbf{u}). \qquad (2) $$

In the above $\partial^{n}\mathbf{R}/\partial^{n}\mathbf{u}$ is the tangent stiffness matrix. This process is repeated until convergence is achieved which is asymptotically quadratic. After ${}^{n}\mathbf{u}$ is deter-

mined the time is incremented and the process is repeated to evaluate ${}^{n+1}\mathbf{u}$.

2.1. *First-order sensitivity analysis*

The residual ${}^{n}\mathbf{R}$ and, hence, the system response ${}^{n}\mathbf{u}$, are dependent on the N model parameters which are the components of the vector $\boldsymbol{\phi}$, and so we rewrite equation (2) as

$$\mathbf{R}(\mathbf{u}(\boldsymbol{\phi}, {}^{n}t),\ \mathbf{u}(\boldsymbol{\phi}, {}^{n-1}t),\ \boldsymbol{\phi},\ {}^{n}t) = 0. \tag{3}$$

The parameters in $\boldsymbol{\phi}$ may be used to describe material properties, loads, boundary and initial conditions, and shape. Upon completion of the analysis at time ${}^{M}t$, a general response functional

$$F(\boldsymbol{\phi}) = G(\mathbf{u}(\boldsymbol{\phi}, {}^{M}t), \mathbf{u}(\boldsymbol{\phi}, {}^{M-1}t), \ldots, \mathbf{u}(\boldsymbol{\phi}, {}^{1}t), \boldsymbol{\phi}) \tag{4}$$

is evaluated to, e.g. monitor the cost or constraint functions in an optimization problem or the error function in an inverse analysis.

The objective of the sensitivity analysis is to determine the derivatives of F with respect to the model parameters † i.e. to evaluate

$$\frac{DF}{D\phi_i} = \sum_{n=1}^{M} \left\{ \frac{\partial G}{\partial {}^{n}\mathbf{u}} \frac{D {}^{n}\mathbf{u}}{D\phi_i} \right\} + \frac{\partial G}{\partial \phi_i} \tag{5}$$

where $\partial G/\partial {}^{n}\mathbf{u}$ and $\partial G/\partial \phi_i$ are explicit quantities, whereas the response derivative $D {}^{n}\mathbf{u}/D\phi_i$ is implicitly defined through the residual time ${}^{n}t$ (cf. equation (3)). To resolve this implicit quantity, either the direct differentiation method or the adjoint method may be pursued.

2.1.1. *Direct differentiation method.* In the direct differentiation method, the implicit response derivative $D {}^{n}\mathbf{u}/D\phi_i$ is evaluated by differentiating the residual of equation (3) [31], after which some manipulation yields

$$\frac{\partial {}^{n}\mathbf{R}}{\partial {}^{n}\mathbf{u}} \frac{D {}^{n}\mathbf{u}}{D\phi_i} = -\left(\frac{\partial {}^{n}\mathbf{R}}{\partial {}^{n-1}\mathbf{u}} \frac{D {}^{n-1}\mathbf{u}}{D\phi_i} + \frac{\partial {}^{n}\mathbf{R}}{\partial \phi_i} \right). \tag{6}$$

Equation (6) forms a linear pseudo-problem, that may be solved to evaluate the response derivative $D {}^{n}\mathbf{u}/D\phi_i$, if the derivative $D {}^{n-1}\mathbf{u}/D\phi_i$ is known. This derivative $D {}^{n-1}\mathbf{u}/D\phi_i$ is known if we evaluate equation (6) at the first time step $n = 1$, since the initial condition sensitivity $D {}^{0}\mathbf{u}/D\phi_i$ is known (and assumed zero for our analyses), and march forward in time. At each time step ${}^{n}t$, the Newton–Raphson method is used to evaluate ${}^{n}\mathbf{u}$, then the pseudo-loads are assembled and the response derivative $D {}^{n}\mathbf{u}/D\phi_i$, are calculated with respect to the $i = 1, N$ model parameters. The time is then advanced and this process is repeated. Finally, the first-order sensitivity is calculated from equation (5). Note that the operator in the pseudo-problem is the tangent operator which is used in the Newton–Raphson iteration for the primal analysis at time ${}^{n}t$ and, hence, the response derivatives are efficiently evaluated simultaneously with the primal analysis.

The direct differential method is viable provided the number of model parameters, N is not large compared to the number of response functionals [33]. If N is large, then the N pseudo-load assemblies and back substitutions, which are performed at each time step, may require significant computational effort. In these cases the adjoint method is preferred.

2.1.2. *Adjoint method.* In the adjoint method, the first-order sensitivities are obtained via the Lagrange multiplier method, where the implicit response derivative $D {}^{n}\mathbf{u}/D\phi_i$ is eliminated from equation (5) [31]. Equations (1) and (4) are combined to form the augmented functional

$$\hat{F}(\phi) = G(\mathbf{u}(\boldsymbol{\phi}, {}^{M}t), \mathbf{u}(\boldsymbol{\phi}, {}^{M-1}t), \ldots, \mathbf{u}(\boldsymbol{\phi}, {}^{1}t), \boldsymbol{\phi}) - \sum_{n=1}^{M} {}^{n}\boldsymbol{\nu}^{\mathrm{T}} \mathbf{R}(\mathbf{u}(\boldsymbol{\phi}, {}^{n}t), \mathbf{u}(\boldsymbol{\phi}, {}^{n-1}t), \boldsymbol{\phi}, {}^{n}t) \tag{7}$$

where ${}^{1}\boldsymbol{\nu}, {}^{2}\boldsymbol{\nu}, \ldots, {}^{M}\boldsymbol{\nu}$ are the Lagrange multipliers which equal the discretized solution of an adjoint terminal value problem at times ${}^{n}t$, $n = 1, \ldots, M$. Further, $\hat{F} = F$, since the residuals equal zero, i.e. ${}^{n}\mathbf{R} = 0$ (cf equation (3)) and $D\hat{F}/D\phi_i = DF/D\phi_i$, since $D {}^{n}\mathbf{R}/D\phi_i = 0$ (cf. equation (6)). Differentiation of equation (7) yields

$$\frac{D\hat{F}}{D\phi_i} = \left[\frac{\partial G}{\partial \phi_i} - \sum_{n=1}^{M} {}^{n}\boldsymbol{\nu}^{\mathrm{T}} \frac{\partial {}^{n}\mathbf{R}}{\partial \phi_i} \right] \tag{8}$$

where ${}^{n}\boldsymbol{\nu}$ is defined through the following terminal adjoint problem to annihilate the implicit response derivatives $D {}^{n}\mathbf{u}/D\phi_i$, that would otherwise be present in equation (8)

$$\left(\frac{\partial {}^{M}\mathbf{R}}{\partial {}^{M}\mathbf{u}} \right)^{\mathrm{T}} {}^{M}\boldsymbol{\nu} = \frac{\partial G}{\partial {}^{M}\mathbf{u}}$$

$$\left(\frac{\partial {}^{n}\mathbf{R}}{\partial {}^{n}\mathbf{u}} \right)^{\mathrm{T}} {}^{n}\boldsymbol{\nu} = \frac{\partial G}{\partial {}^{n}\mathbf{u}} - \left(\frac{\partial {}^{n+1}\mathbf{R}}{\partial {}^{n}\mathbf{u}} \right)^{\mathrm{T}} {}^{n-1}\boldsymbol{\nu},$$

$$n = M-1, \ldots, 1. \tag{9}$$

Upon solving equation (9) for the adjoint response ${}^{M}\boldsymbol{\nu}$, ${}^{M-1}\boldsymbol{\nu}, \ldots, {}^{1}\boldsymbol{\nu}$, the sensitivity is evaluated from equation (8). The adjoint analysis requires either the storage or recomputation of the decomposed tangent operators for every time step because the transient response $\mathbf{u}$ must be known for all times before the adjoint analysis begins. Note that one adjoint analysis is required for each response functional.

2.2. *Second-order design sensitivity analysis*

An application of the chain rule to equation (5) gives the symmetric second-order design sensitivity expression [28]

† To simplify this analysis we do not consider F to be a function of the initial conditions ${}^{0}\mathbf{u}$ and we consider the initial conditions ${}^{0}\mathbf{u}$, to be constant, i.e. $(D {}^{0}\mathbf{u})/(D\phi_i) = 0$.

$$\frac{D^2F}{D\phi_i D\phi_j} = \sum_{n=1}^{n=M}\left\{\frac{D({}^n\mathbf{u})}{D\phi_j}\frac{\partial^2 G}{\partial({}^n\mathbf{u})^2}\frac{D({}^n\mathbf{u})}{D\phi_i} + \frac{\partial^2 G}{\partial({}^n\mathbf{u})\partial\phi_j}\frac{D({}^n\mathbf{u})}{D\phi_i} + \frac{\partial G}{\partial({}^n\mathbf{u})}\frac{D^2({}^n\mathbf{u})}{D\phi_i D\phi_j} + \frac{\partial^2 G}{\partial\phi_i\partial({}^n\mathbf{u})}\frac{D({}^n\mathbf{u})}{D\phi_j}\right\} + \frac{\partial^2 G}{\partial\phi_i\partial\phi_j} \quad (10)$$

where the arguments in the above are suppressed for conciseness. The difficulty in evaluating equation (10) once again arises from the presence of the response derivatives, $D({}^n\mathbf{u})/D\phi_i$, $D({}^n\mathbf{u})/D\phi_j$ and $D^2({}^n\mathbf{u})/D\phi_i D\phi_j$.

The hybrid method which combines the previously discussed direct differentiation and adjoint methods is used to compute the second-order design sensitivities [28, 34]. Following the first-order adjoint sensitivity analysis, equation (5) is augmented with the derivative of the residual (cf. equation (6)) to give

$$\frac{D\hat{F}}{D\phi_i} = \sum_{n=1}^{M}\left\{\frac{\partial G}{\partial({}^n\mathbf{u})}\frac{D({}^n\mathbf{u})}{D\phi_i}\right\} + \frac{\partial G}{\partial\phi_i} - \sum_{n=1}^{M} {}^n\boldsymbol{\nu}^{\mathrm{T}}\left\{\frac{\partial({}^n\mathbf{R})}{\partial({}^n\mathbf{u})}\frac{D({}^n\mathbf{u})}{D\phi_i} + \frac{\partial({}^n\mathbf{R})}{\partial({}^{n-1}\mathbf{u})}\frac{\partial({}^{n-1}\mathbf{u})}{D\phi_i} + \frac{\partial({}^n\mathbf{R})}{\partial^n\phi_i}\right\}. \quad (11)$$

Recall that $D\hat{F}/D\phi_i = DF/D\phi_i$, and further that $D^2\hat{F}/D\phi_i D\phi_j = D^2F/D\phi_i D\phi_j$, since $D/D\phi_j\{D^n\mathbf{R}/D\phi_j\} = \mathbf{0}$. Differentiating equation (11) with respect to ϕ_j, gives after some rearranging

$$\begin{aligned}\frac{D^2\hat{F}}{D\phi_i D\phi_j} = \Bigg[&\frac{\partial^2 G}{\partial\phi_i\partial\phi_j} + \sum_{n=1}^{M}\Bigg(-{}^n\boldsymbol{\nu}^{\mathrm{T}}\frac{\partial^2({}^n\mathbf{R})}{\partial\phi_i\partial\phi_j} - {}^n\boldsymbol{\nu}^{\mathrm{T}}\frac{\partial^2({}^n\mathbf{R})}{\partial\phi_j\partial({}^{(n-1)}\mathbf{u})}\frac{D({}^{n-1}\mathbf{u})}{D\phi_i} \\ &- {}^n\boldsymbol{\nu}^{\mathrm{T}}\frac{\partial^2({}^n\mathbf{R})}{\partial\phi_i\partial({}^{n-1}\mathbf{u})}\frac{D({}^{n-1}\mathbf{u})}{D\phi_j} + \frac{\partial^2 G}{\partial({}^n\mathbf{u})\partial\phi_j}\times\frac{D({}^n\mathbf{u})}{D\phi_i} - {}^n\boldsymbol{\nu}^{\mathrm{T}}\frac{\partial^2({}^n\mathbf{R})}{\partial\phi_i\partial({}^n\mathbf{u})}\frac{D({}^n\mathbf{u})}{D\phi_j} \\ &+ \frac{\partial^2 G}{\partial\phi_i\partial({}^n\mathbf{u})}\frac{D({}^n\mathbf{u})}{D\phi_j} - {}^n\boldsymbol{\nu}^{\mathrm{T}}\frac{\partial^2({}^n\mathbf{R})}{\partial\phi_i\partial({}^n\mathbf{u})}\frac{D({}^n\mathbf{u})}{D\phi_j} - {}^n\boldsymbol{\nu}^{\mathrm{T}}\frac{\partial^2({}^n\mathbf{R})}{\partial({}^{n-1}\mathbf{u})^2}\frac{D({}^{n-1}\mathbf{u})}{D\phi_i}\frac{D({}^{n-1}\mathbf{u})}{D\phi_j} \\ &+ {}^n\boldsymbol{\nu}^{\mathrm{T}}\frac{\partial^2({}^n\mathbf{R})}{\partial({}^n\mathbf{u})\partial({}^{n-1}\mathbf{u})}\frac{D({}^n\mathbf{u})}{D\phi_i}\frac{D({}^{n-1}\mathbf{u})}{D\phi_j} - {}^n\boldsymbol{\nu}^{\mathrm{T}}\frac{\partial^2({}^n\mathbf{R})}{\partial({}^{n-1}\mathbf{u})\partial({}^n\mathbf{u})}\frac{D({}^{n-1}\mathbf{u})}{D\phi_j}\frac{D({}^n\mathbf{u})}{D\phi_j} \\ &- {}^n\boldsymbol{\nu}^{\mathrm{T}}\frac{\partial^2({}^n\mathbf{R})}{\partial({}^n\mathbf{u})^2}\frac{D({}^n\mathbf{u})}{D\phi_i}\frac{D({}^n\mathbf{u})}{D\phi_j}\Bigg)\Bigg]. \end{aligned} \quad (12)$$

To evaluate equation (12) we use the direct differentiation method outlined in Section 2.1.1 to compute $D^n\mathbf{u}/D\phi_i$, and $D^n\mathbf{u}/D\phi_j$ and solve the adjoint problem previously discussed (cf. equation (9)) for ${}^n\boldsymbol{\nu}$ to eliminate the products of $D^2({}^n\mathbf{u})/D\phi_i D\phi_j$ which would otherwise be present in equation (12).

3. TRANSIENT NONLINEAR HEAT CONDUCTION PROBLEMS

We now specialize these equations for a transient heat conduction problem. For a more in-depth discussion regarding the derivation of the governing equations and the standard finite element formulations see [35, 36].

3.1. *The initial boundary value problem*

Consider a spatial region Ω bounded by the surface Γ with outward unit normal $\mathbf{n}$. Γ is comprised of two complementary sub-surfaces Γ_1 and Γ_2, and the time domain is given by $I =]0, {}^M t[$. The initial boundary-value problem (IBVP) is expressed in terms of the local energy balance equation, constitutive relation and initial and boundary conditions as

$$\begin{aligned} \dot{e}(T) + \nabla\cdot\mathbf{q} &= f \quad \text{in } \Omega\times I \\ \mathbf{q} &= -k(T)\nabla T \quad \text{in } \Omega\times I \\ -\mathbf{q}\cdot\mathbf{n} &= q^{\mathrm{n}} \quad \text{on } \Gamma\times I \\ T|_{\Omega\times 0} &= T_0 \quad \text{in } \Omega \\ T &= T_{\mathrm{p}} \quad \text{on } \Gamma_1\times I \\ q^{\mathrm{n}} &= \bar{h} \quad \text{on } \Gamma_2\times I \end{aligned} \quad (13)$$

where T is the temperature, $\mathbf{q}$ is the heat flux vector, e is the temperature dependent internal energy, f is the heat source, k is the temperature dependent thermal conductivity for an isotropic material, T_0 is the initial temperature, T_{p} is the prescribed temperature over Γ_1 and $\bar{h}$ is the prescribed surface flux over Γ_2.

3.2. *Primal analysis*

We compute the solution to equation (13) by expressing the IBVP in a weak formulation and discretizing by the finite element method. The time derivative is approximated via

$${}^n\dot{e} = \frac{\partial^n e}{\partial T}\frac{{}^nT - {}^{n-1}T}{\Delta t} \quad (14)$$

where $\Delta t = {}^nt - {}^{n-1}t$. The resulting time discretized weak problem statement corresponding to equation (13) is: find ${}^nT$ such that equations (13.2)–(13.5) hold and

$${}^nR({}^nT, {}^{n-1}T) = 0 = \int_\Omega T^*\frac{\partial^n e}{\partial T}\frac{{}^nT - {}^{n-1}T}{\Delta t}\mathrm{d}\Omega - \int_\Omega T^* f\,\mathrm{d}\Omega + \int_{\Gamma_2} T^{*n}\bar{h}\,\mathrm{d}\Gamma + \int_\Omega (\nabla T^*)^{\mathrm{T}n}k\nabla^n T\,\mathrm{d}\Omega \quad (15)$$

for all sufficiently smooth T^* such that $T^* = 0$ on Γ_1.

In the Galerkin finite element method, the temperature field T is interpolated within each element via the element shape functions and the nodal temperatures† $\mathbf{T}$ to form the discretized residual $\mathbf{R}$ of equation (1). The Newton–Raphson method discussed in Section 2 is then used to solve equation (15) where the discretized tangent operator is obtained from

$$\frac{\partial^n R}{\partial^n T}\delta T = \int_\Omega T^* \frac{\partial^2({}^n e)}{\partial T^2}\delta T \frac{{}^n T - {}^{n-1}T}{\Delta t}\,d\Omega + \int_\Omega T^* \frac{\partial^n e}{\partial T}\frac{\delta T}{\Delta t}\,d\Omega + \int_\Omega (\nabla T^*)^T \times \frac{\partial^n k}{\partial T}\delta T \nabla {}^n T\,d\Omega + \int_\Omega (\nabla T^*)^{Tn} k \nabla(\delta T)\,d\Omega. \quad (16)$$

This formulation is now applied to solve a two-dimensional transient heat conduction problem.

Example 1. We consider a transient version of the steady-state problem solved in [28] and as in [28] this direct problem is used to define the ensuing inverse problem. The two-dimensional rectangular plate, shown in Fig. 1(a) is 1.0×0.5 m, is isotropic and homogeneous, and is subjected to a prescribed temperature on each edge, $T((0, x_2), t) = 0°C$, $T((x_1, 0), t) = 0°C$, $T((L, x_2), t) = 0°C$ and $T((x_1, h), t) = T_p(x_1, t)$, where the components of $\mathbf{x}$ are with respect to the orthonormal basis $\{\hat{\mathbf{e}}_1, \hat{\mathbf{e}}_2\}$ shown. Here $T_p = \bar{T}_p$ is the prescribed temperature profile shown in Fig. 1(b) which is generated via B-spline basis functions. (cf. Section 4.1). Note that $\bar{T}_p$ is not symmetric. A uniform initial temperature field of $T_0 = 1°C$ is prescribed and the material response is given by $k(T) = 10\ \mathrm{Wm^{-1}\,°C^{-1}}$ and $e(T) = 0.5\ T\ \mathrm{Jm^{-3}}$.

The spatial domain is discretized with 800 quadrilateral elements (40×20); the time increment is given by $\Delta t = 0.005$ s; and the analysis is performed for $M = 10$ time steps (${}^M t = 0.05$ s). The isotherm distribution for the last time step, i.e. $t = 0.05$ s, is shown in Fig. 1(d). The transient resultant surface flux, q^n at the bottom edge of the plate, i.e. over the $x_2 = 0$ edge, is calculated using equation (13.3) and is illustrated in Fig. 1(c).

4. INVERSE PROBLEM

In the ill-posed inverse problem, the boundary condition on the top edge, i.e. the $x_2 = h$ edge, is unknown. Rather, both the surface flux distribution q_{data} and the temperature on the $x_2 = 0$ edge are known. To solve this inverse problem we apply a prescribed temperature on the entire surface and solve the well-posed direct problem and then compute the surface flux, q^n on the $x_2 = 0$ edge. An error function F, is introduced to quantify how closely the computed surface flux q^n, resulting from our choice for T_p over the $x_2 = h$ edge, agrees with the over prescribed surface flux q_{data}

$$F(\boldsymbol{\phi}) = \sum_{n=1}^{M} \int_{A_{\text{data}}} \frac{1}{2\bar{q}^2}(q^n(\mathbf{x}, \boldsymbol{\phi}, {}^n t) - q_{\text{data}}(\mathbf{x}, {}^n t))^2\,dA \quad (17)$$

where A_{data} is the area associated with the $x_2 = 0$ edge and $\bar{q} = 1\ \mathrm{Wm^{-2}}$ is a scaling constant. In the above, the temperature profile T_p on the $x_2 = h$ edge has been parameterized by N model parameters which are the components of the vector $\boldsymbol{\phi}$ for our examples, i.e. $T_p = T_p((x_1, h), \boldsymbol{\phi}, {}^n t)$. The objective of the inverse problem is to find the $\boldsymbol{\phi}$ value such that $F(\boldsymbol{\phi}) = 0$ so that the over-prescribed and computed surface fluxes agree, i.e. so that $q_{\text{data}}((x_1, 0), {}^n t) = q^n((x_1, 0), \boldsymbol{\phi}, {}^n t))$. We use numerical optimization to systematically minimize this error function and, hence, to solve the inverse problem.

4.1. *Temperature profile parameterization*

A B-spline surface patch is used to parameterize T_p over the $x_2 = h$ edge and time domain. This surface is the Cartesian product of two B-spline curves [37]; a time curve which interpolates T_p in time and a space curve which interpolates T_p in space (i.e. along the $x_2 = h$ edge). The B-spline basis is chosen because the basis is non-global and it can be of any order. The two B-spline curves are functions of the parametric variables w and v, respectively, which correspond to the time t and position x_1 along the $x_2 = h$ edge, respectively. We relate w to t as

$$w = \bar{w}(t) = \frac{t}{t_{\max}}(n - L + 1) \quad (18)$$

where $t_{\max}$ is the maximum time considered and L and n are the order of the curve and the number of polygon vertices which compromise the time curve, respectively. In a similar manner, v is related to x_1 as

$$v = \bar{v}(x_1) = \frac{x_1}{L_{\max}}(m - P + 1) \quad (19)$$

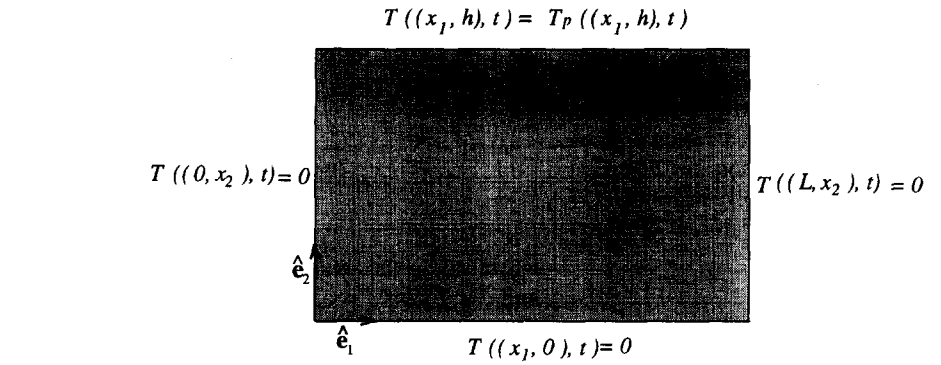
where $L_{\max}$ is the maximum value of x_1 considered and P and m are the order of the curve and the number of vertices which comprise the position curve, respectively.

The temperature T_p at any $(w, v) = (\bar{w}(t), \bar{v}(x_1))$ pair is obtained from

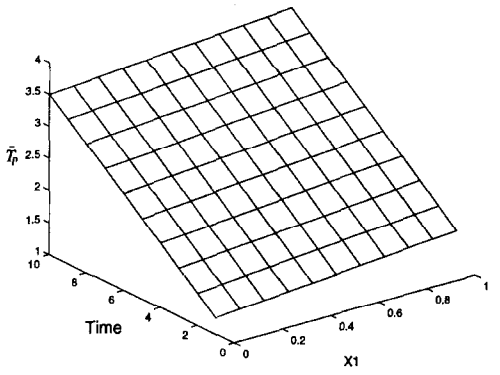
$$T_p((x_1, h), \boldsymbol{\phi}, t) = \sum_{i=0}^{n}\sum_{j=0}^{m} B_{i,j}(\boldsymbol{\phi})\bar{N}_{i,P}(\bar{w}(t))\bar{M}_{j,L}(\bar{v}(x_1)) \quad (20)$$

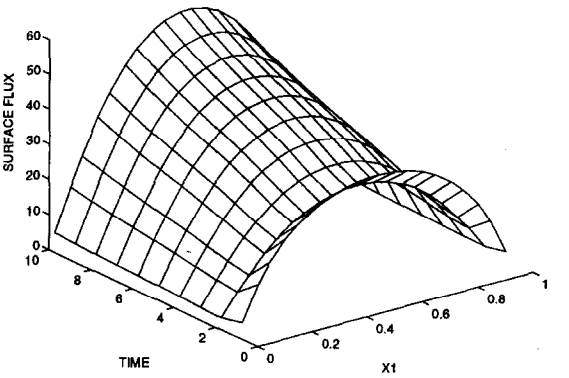
where $\bar{N}$ and $\bar{M}$ are the weighting functions defined below and the matrix $\mathbf{B}$ contains the $n \times m$ defining polygon vertices, $B_{i,j}$ which are the model parameters in $\boldsymbol{\phi}$. The weighting function $\bar{N}$ is defined as

† Here $\mathbf{u}$ in Section 2 is replaced by $\mathbf{T}$ for obvious reasons.

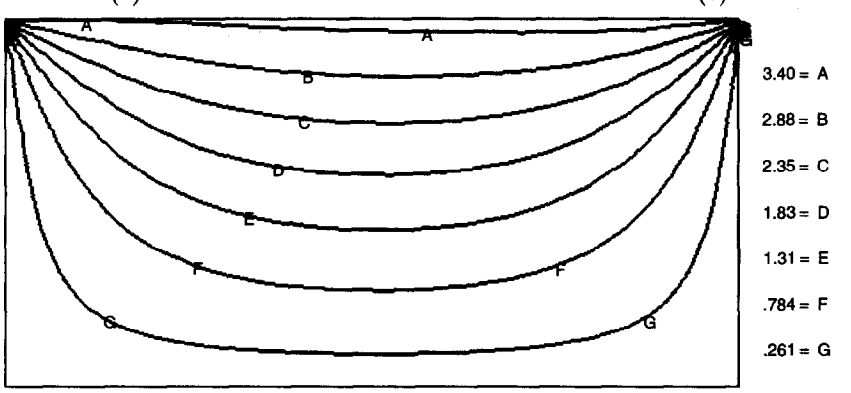


(a)





(b) (c)



(d)

Fig. 1. (a) Plate with applied boundary conditions; (b) prescribed temperature profile $\bar{T}_p$; (c) surface flux history along the $x_2 = 0$ edge; and (d) isotherm distribution for time $t = 0.05$ s.

$$\bar{N}_{i,1}(w) = \begin{cases} 1 & \text{if } kx_i \leqslant w < kx_{i+1} \\ 0 & \text{otherwise} \end{cases} \tag{21}$$

$$\bar{N}_{i,P}(w) = \frac{(w-kx_i)\bar{N}_{i,P-1}(w)}{kx_{i+P-1}-kx_i} + \frac{(kx_{i+P}-w)\bar{N}_{i+1,P-1}(w)}{kx_{i+P}-kx_{i+1}} \tag{22}$$

where kx_i are the knot vector elements for the time curve of order P. The weighting function $\bar{M}$ is similarly defined for the space curve of order L. The temperature profile $\bar{T}_p$ over the $x_2 = h$ edge, shown in Fig. 1(b) is a B-spline surface generated with two fourth-order curves that each contain nine vertices, i.e. $L = P = 4$, and $n = m = 9$. This prescribed temperature distribution is used to generate the

desired surface flux q_{data} for the ensuing inverse problem.

The $(n \times m)$ polygon vertices in $\mathbf{B}$ serve as the N model parameters in $\boldsymbol{\phi}$. The sensitivity of T_p with respect to these parameters is obtained by differentiating equation (20) with respect to $B_{p,q}(=\phi_r)$

$$\frac{\partial T_p}{\partial \phi_r}((x_1,h),\boldsymbol{\phi},t) = \sum_{i=0}^{n}\sum_{j=0}^{m} \delta_{ip}\delta_{jq}\bar{N}_{i,P}(\bar{w}(t))\bar{M}_{j,L}(\bar{v}(x_1)) \tag{23}$$

where δ_{ip} is the Kronecker delta function.

The spatial and temporal derivatives of T_p are required when regularization is used in the inverse analysis. Differentiation of equation (20) with respect to time gives

$$\frac{\partial T_p}{\partial t}((x_1,h),\boldsymbol{\phi},t) = \sum_{i=0}^{n}\sum_{j=0}^{m} B_{i,j}(\boldsymbol{\phi}) \times \frac{\partial \bar{N}_{i,P}}{\partial w}(\bar{w}(t))\bar{M}_{j,L}(\bar{v}(x_1))\frac{\partial w}{\partial t}(t) \tag{24}$$

where

$$\frac{\partial \bar{N}_{i,1}}{\partial w} = 0 \tag{25}$$

$$\frac{\partial \bar{N}_{i,P}}{\partial w}(w) = \frac{(w-kx_i)\dfrac{\partial \bar{N}_{i,P-1}}{\partial w}(w)}{kx_{i+P-1}-kx_i} + \frac{\bar{N}_{i,P-1}(w)}{kx_{i+P-1}-kx_i} + \frac{(kx_{i+p}-w)\dfrac{\partial \bar{N}_{i+1,P-1}}{\partial w}(w)}{kx_{i+P}-kx_{i+1}} + \frac{-\bar{N}_{i+1,P-1}(w)}{kx_{i+P}-kx_{i+1}}. \tag{26}$$

The derivative $\partial T_p/\partial x_i$ is similarly obtained. The sensitivity of $\partial T_p/\partial t$ with respect to $B_{p,q}(=\phi_r)$ is then obtained from

$$\frac{\partial}{\partial \phi_r}\left(\frac{\partial T_p}{\partial t}\right)((x_1,h),\boldsymbol{\phi},t) = \sum_{i=0}^{n}\sum_{j=0}^{m} \delta_{ip}\delta_{jq}\frac{\partial \bar{N}_{i,P}}{\partial w}(\bar{w}(t))\bar{M}_{j,L}\frac{\partial w}{\partial t}(t). \tag{27}$$

Again the derivative $\partial/\partial\phi_r$ $(\partial T_p/\partial x_1)$ is similarly obtained. In comparison to the cubic spline sensitivities used in [28], these derivative expressions for the B-spline surface are considerably simplified.

4.2. *Design sensitivity analysis*

First- and second-order sensitivities are now computed for the previous example using the methods discussed in Sections 2.1 and 2.2. The prescribed surface flux and heat source terms are not included in the following analyses as they do not appear in the example problem.

4.2.1. *First-order design sensitivity analysis.* The first-order design sensitivities of equation (17) are computed by using the direct differentiation and adjoint methods. Differentiation of equation (17) with respect to ϕ_i gives us $\partial G/\partial^n\mathbf{T}\ D^n\mathbf{T}/D\phi_i$ (cf. equation (35) in [28]) and $\partial G/\partial\phi_i = 0$ since G is not an explicit function of $\boldsymbol{\phi}$. In the direct differentiation method, the implicitly defined response sensitivity $D^n\mathbf{T}/D\phi_i$ is evaluated via equation (6). The pseudo-problem corresponding to equation (6) is obtained by differentiation of equation (15), i.e. to find $D^nT/D\phi_i$ where $D^nT/D\phi_i = D^nT_p/D\phi_i$ on Γ_1 and

$$\begin{aligned}\frac{D^nR}{D\phi_i} = 0 = &\int_\Omega T^*\frac{\partial^2({}^ne)}{\partial(T)^2}\frac{D^nT}{D\phi_i} \times \frac{{}^nT - {}^{n-1}T}{\Delta t}\mathrm{d}\Omega + \int_\Omega T^*\frac{\partial^n e}{\partial T}\frac{D^nT}{D\phi_i} \times \frac{1}{\Delta t}\mathrm{d}\Omega + \int_\Omega (\nabla T^*)^T\frac{\partial^n k}{\partial T}\frac{D^nT}{D\phi_i}\nabla^n T\,\mathrm{d}\Omega \\ &+ \int_\Omega (\nabla T^*)^T k\nabla\left(\frac{D^nT}{D\phi_i}\right)\mathrm{d}\Omega - \int_\Omega T^*\frac{\partial^n e}{\partial T}\frac{D^{n-1}T}{D\phi_i}\frac{1}{\Delta t}\mathrm{d}\Omega \\ = &\frac{\partial^n R}{\partial T}\frac{\partial^n T}{\partial \phi_i} - \int_\Omega T^*\frac{\partial^n e}{\partial T}\frac{D^{n-1}T}{D\phi_i}\frac{1}{\Delta t}\mathrm{d}\Omega\end{aligned} \tag{28}$$

for all suitable T^*. In the above, $D^nT_p/D\phi_i$ on the $x_2 = h$ edge is obtained via the B-spline sensitivities discussed in Section 4.1 and $D^nT_p/D\phi_i = 0$ on the remaining edges. In our problem the residual is not an explicit function of $\boldsymbol{\phi}$, i.e. $\partial^n\mathbf{R}/\partial\phi_i = \mathbf{0}$ and, hence, the only contributions to the pseudo-load are obtained by discretizing

$$\int_\Omega T^*\frac{\partial^n e}{\partial T}\frac{D^{n-1}T}{D\phi_i}\frac{1}{\Delta t}\mathrm{d}\Omega$$

and accommodating the non-homogeneous essential boundary conditions on the $x_2 = h$ edge due to $D^nT_p/D\phi_i$. The adjoint method (cf. Section 2.1.2) requires the solution to the adjoint problem of equation (9) in which the derivative $\partial^n G/\partial^n T$ is obtained from equation (35) in [28]. After the adjoint problem is solved, the sensitivity is evaluated from equation (8).

4.2.2. *Second-order sensitivity analysis.* The hybrid method (cf. Section 2.2) is used to compute the second-order sensitivities of the error function. Recall that neither the error function nor the residual are explicit functions of $\boldsymbol{\phi}$, thus all partial derivatives of G and $\mathbf{R}$ with respect to $\boldsymbol{\phi}$ are zero. The terms in $D^2\hat{F}/D\phi_iD\phi_j$ can be found by an application of the chain rule to the first-order sensitivity expressions. The product $D^n\mathbf{T}/D\phi_j\partial^2G/\partial({}^n\mathbf{T})^2D^n\mathbf{T}/D\phi_i$ is evaluated from equations (44)–(48) in [28]. The remaining derivatives are found by discretizing the right-hand sides of

$$\frac{\partial^2({}^n\mathbf{R})}{\partial {}^n\mathbf{T}\partial {}^{n-1}\mathbf{T}}\frac{D^{n-1}\mathbf{T}}{D\phi_i}\frac{D^n\mathbf{T}}{D\phi_j} = -\int_\Omega T^*\frac{\partial^2({}^nc)}{\partial T^2}\frac{D^{n-1}T}{D\phi_i}\frac{D^nT}{D\phi_j}\frac{1}{\Delta t}\mathrm{d}\Omega \quad (29)$$

and

$$\begin{aligned}\frac{\partial^2({}^n\mathbf{R})}{\partial({}^n\mathbf{T})^2}\frac{D^n\mathbf{T}}{D\phi_i}\frac{D^n\mathbf{T}}{D\phi_j} &= \int_\Omega(\nabla T^*)^{\mathrm{T}}\frac{\partial^n k}{\partial T}\frac{D^nT}{D\phi_i}\nabla\frac{D^nT}{D\phi_j}\mathrm{d}\Omega \\ &+\int_\Omega(\nabla T^*)^{\mathrm{T}}\frac{\partial^2({}^nk)}{\partial T^2} \\ &\times\frac{D^nT}{D\phi_i}\frac{D^nT}{D\phi_j}\nabla{}^nT\mathrm{d}\Omega \\ &+\int_\Omega(\nabla T^*)^{\mathrm{T}}\frac{\partial^n k}{\partial T}\frac{D^nT}{D\phi_j}\nabla\frac{D^nT}{D\phi_i}\mathrm{d}\Omega+\frac{2}{\Delta t}\int_\Omega T^* \\ &\times\frac{\partial^2({}^nc)}{\partial T^2}\frac{D^nT}{D\phi_i}\frac{D^nT}{D\phi_j}\mathrm{d}\Omega+ \\ &\times\frac{1}{\Delta t}\int_\Omega T^*\frac{\partial^3({}^nc)}{\partial T^3}\frac{D^nT}{D\phi_i}\frac{D^nT}{D\phi_j}{}^nT\mathrm{d}\Omega \\ &-\frac{1}{\Delta t}\int_\Omega T^*\frac{\partial^3({}^nc)}{\partial T^3}\frac{D^nT}{D\phi_i}\frac{D^nT}{D\phi_j}{}^{n-1}T\mathrm{d}\Omega. \end{aligned} \quad (30)$$

The other derivatives required to evaluate $D^2\hat{F}/D\phi_iD\phi_j$ are either obtained from symmetry of the derivative or from similar computations performed at the $n-1$ time step.

4.3. *Solution strategy*

To solve the previously discussed inverse problem, the prescribed temperature T_p, is parameterized over the $x_2 = h$ edge, via a B-spline patch (cf. Section 4.1) and an optimization algorithm is used to determine the model parameters $\phi_r = B_{p,q}$, i.e. the prescribed temperature T_p, that minimize the error function F of equation (17). The process begins by supplying an initial guess for $\boldsymbol{\phi}$. The B-spline interpolation is used to generate T_0, the resulting well-posed problem is solved, the surface flux $q^n((x_1, 0), \boldsymbol{\phi}, t)$ is evaluated and this flux, ${}^nq$, is compared to the surface flux boundary condition of the original ill-posed inverse problem q_{data}, via the error function F (cf. equation (17)). If the error function equals zero, i.e. $F(\boldsymbol{\phi}) = 0$, then we have found a solution to the inverse problem. Otherwise, the model parameters are updated and the process is repeated.

Two types of unconstrained optimization programs are implemented to solve the inverse problem, a Newton's method which utilizes second-order design sensitivities of the error function F and the BFGS variable metric method which uses first-order sensitivities [29, 38].

4.3.1. *Linear systems. Case number* 1. For this linear thermal system $k = 10.0\ \mathrm{Wm^{-1}{}^\circ C^{-1}}$, and $e = 0.5\ T\ \mathrm{Jm^{-3}}$. The prescribed temperature distribution $\bar{T}_p$ shown in Fig. 1(b) is used to compute q_{data} on the $x_2 = 0$ edge which is used as the over-prescribed surface flux for the inverse problem. The goal is to replicate this $\bar{T}_p$ temperature distribution. Note that since T is linear in $\bar{T}_p$, i.e. $\boldsymbol{\phi}$, F is quadratic in $\boldsymbol{\phi}$ so that the Hessian of F is constant [39].

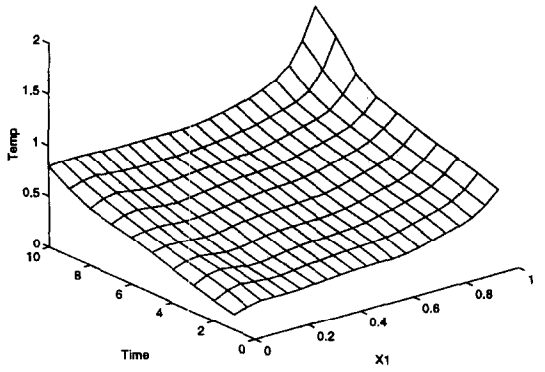
For the inverse analysis T_p is defined as a B-spline surface with a 9×9 array of vertices, i.e. there are $N = 81$ model parameters in $\boldsymbol{\phi}$. The initial T_p distribution for our analysis is illustrated in Fig. 2(a) and clearly does not agree with that of $\bar{T}_p$ illustrated in Fig. 1(b). A preliminary analysis with this T_p yields $F(\boldsymbol{\phi}) = 103.698$.

The BFGS method with polynomial interpolation for the one-dimensional search is first employed to solve the inverse problem. The computed profile after 42 iterations is shown in Fig. 2(b). In theory the BFGS method should converge in 81 iterations assuming that the line search is exact. Each iteration in the BFGS algorithm of [40] requires one gradient and three function evaluations, thus making this method computationally prohibitive. It is noted that the one-dimensional search uses a quadratic polynomial to approximate the error function and since F is quadratic in $\boldsymbol{\phi}$ the polynomial approximation is exact. The results in Fig. 2(b) show that the T_p distribution after 42 iterations does not resemble the over-prescribed $\bar{T}_p$. However, $F(\boldsymbol{\phi}) = 1.051\times10^{-5}$ and the surface flux produced from this T_p is indistinguishable from the q_{data} shown in Fig. 1(c).

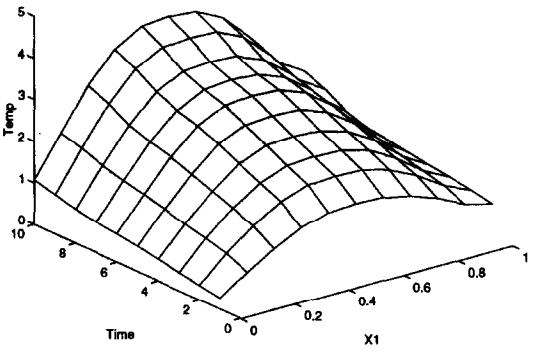
Newton's method is also used to solve the inverse problem. Convergence is obtained in one iteration because the error function is quadratic in $\boldsymbol{\phi}$. Again the computed T_p distribution, shown in Fig. 2(c), does not agree with the $\bar{T}_p$ distribution. However, as with the BFGS method, the surface flux distribution on the $x_2 = 0$ edge agrees with q_{data} and $F(\boldsymbol{\phi}) = 1.452\times10^{-6}$. The ill-posed nature of the IHCP is seen via the disparity between T_p and $\bar{T}_p$ which is attributed to the Hessian's poor condition number, that equals 1×10^{18}. A singularity of the Hessian at convergence indicates that the minimum is not unique, and/or that the model basis is linearly dependent [41].

Case number 2. As described in [28] if the condition number of the Hessian degrades, then the changes in $F(\boldsymbol{\phi})$ due to $\boldsymbol{\phi}$ perturbations in the directions of the Hessian's eigenvectors associated with the smaller eigenvalues are negligible. To see this, the Hessian's† N eigenvalues λ_i and their corresponding normalized eigenvectors $\mathbf{v}_i$, are computed and arranged in ascending order according to the value of λ_i. The eigenvectors are used to define a basis for the model space, i.e. $\boldsymbol{\phi} = \mathbf{S}\boldsymbol{\alpha}$, where $\mathbf{S}$ is the matrix whose columns are the

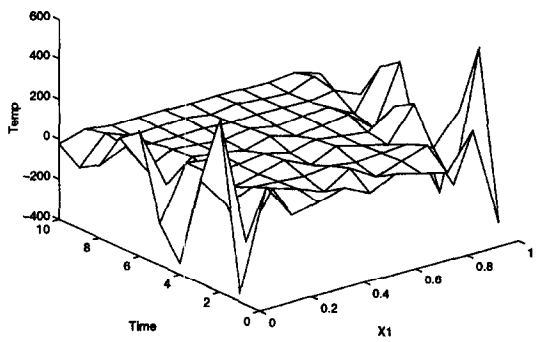
† Recall that the Hessian is independent of $\boldsymbol{\phi}$ for this case.



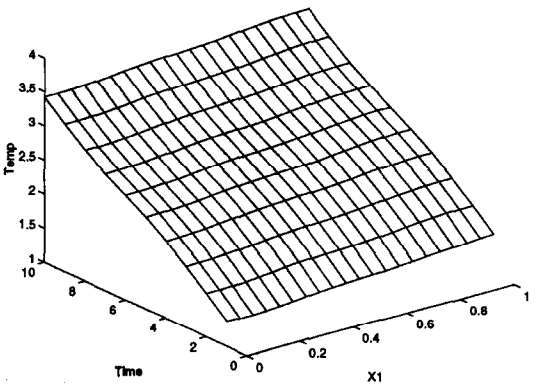
(*a*) Initial Temperature Profile T_p



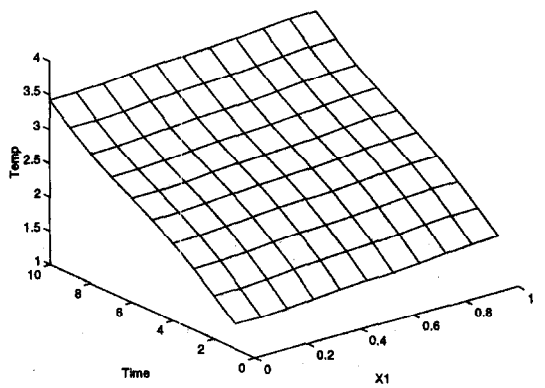
(*b*) BFGS: 42 iterations



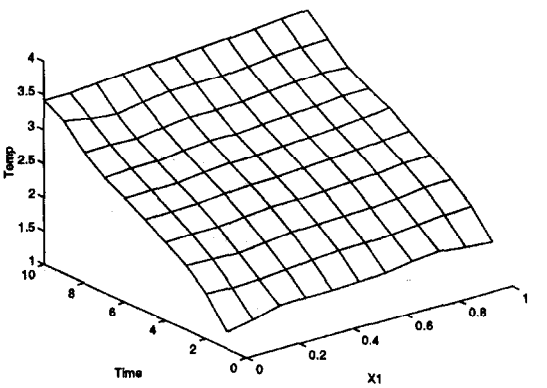
(*c*) Newton's: 1 iteration



(*d*) Smoothing along smaller eigenvalues



(*e*) Newton's with regularization



(*f*) BFGS with regularization

Fig. 2. Initial temperature profile and converged T_p for the different methods.

$i = 1, N$ eigenvectors $\mathbf{v}_i$, and α contains the components of ϕ with respect to the eigen basis. Now we modify $\bar{\phi}$, i.e. the model parameters used to generate $\bar{T}_p$, to $\bar{\phi}+\varepsilon_{\text{pert}}\mathbf{v}_1$, where $\varepsilon_{\text{pert}}$ is the perturbation constant. A modification of $\bar{\phi}$ to $\bar{\phi}+\varepsilon_{\text{pert}}\mathbf{v}_N$, is also performed. Figure 3(a,b) shows the modified $\bar{T}_p$ profiles for these two perturbations which correspond to $\varepsilon_{\text{pert}} = 100$. It is noted that both the profiles differ greatly from the $\bar{T}_p$ profile of Fig. 1(b). Figure 3(c) shows the effects of these modifications on the error function for various values of $\varepsilon_{\text{pert}}$. It is evident that perturbations along $\mathbf{v}_1$ produce negligible changes in the error function, whereas perturbations along $\mathbf{v}_N$

† In [32] the true Hessian is not used, instead a normalized sensitivity matrix is analyzed.

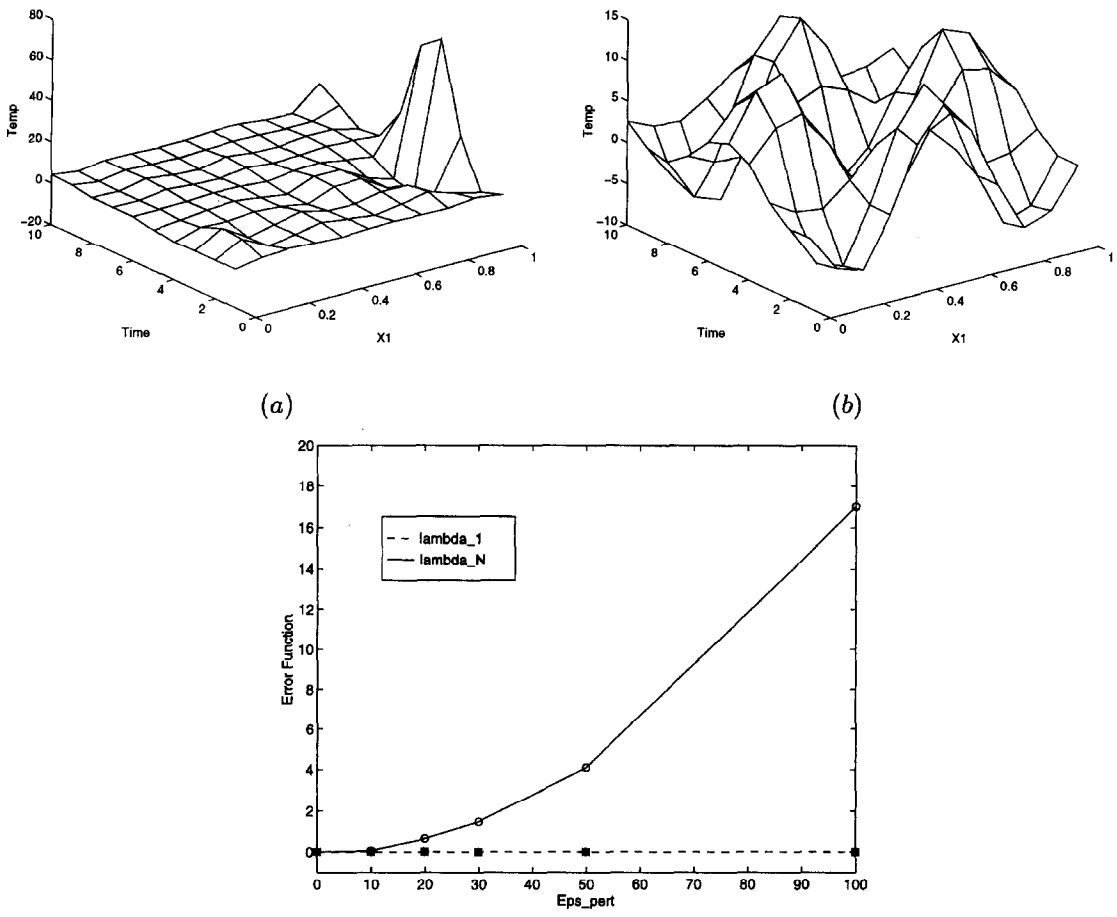


Fig. 3. (a) $\bar{T}_p$ for perturbations along $\mathbf{v}_1$; (b) $\bar{T}_p$ for perturbations along $\mathbf{v}_N$; (c) variations in F for perturbations along $\mathbf{v}_1$ and $\mathbf{v}_2$.

greatly affect the error function value. This information is used in [32] to perform the stability analysis† and to estimate the confidence interval associated with each mode. Note also that modifications along $\mathbf{v}_1$ (and the other eigenvectors associated with near zero eigenvalues) produce localized variations in T_p, whereas more global changes result from variations along $\mathbf{v}_N$ (and the other eigenvectors associated with non-zero eigenvalues); and that the discrepancies between Figs. 1(b) and 2(c) result from localized variations. This result indicates that the discrepancies, between the prescribed temperature profile $\bar{T}_p$ and the computed temperature profiles of Fig. 2(b, c) are due to the arbitrariness of the components of $\boldsymbol{\phi}$ with respect to the eigenvectors associated with the smallest eigenvalues. To illustrate this effect, we minimize the discrepancy between T_p (the distribution obtained from Newton's method) and $\bar{T}_p$ of Fig. 2(b) on the reduced basis corresponding to the 36 eigenvectors associated with the 36 smallest eigenvalues whose magnitudes are less than 10^{-7}. The result $\boldsymbol{\phi}^*$ yields the temperature distribution in Fig. 2(d) and an error value $E(\boldsymbol{\phi}^*) = 1.023 \times 10^{-5}$.

Case number 3. As alluded to above, regularization is necessary to stabilize the inverse problem solution. Effectively, regularization adds a predominantly diagonal matrix to the Hessian, consequently it improves the Hessian's condition number. The regularization function f_{reg} which is added to the error function F, is defined to penalize abrupt variations in the temperature profile, T_p, and as seen through Fig. 3(a), the abrupt variation in T_p are associated with the eigenvectors corresponding to the near zero eigenvalues. The IHCP is now solved by minimizing the modified error function $F_r(\boldsymbol{\phi}) = F(\boldsymbol{\phi}) + f_{reg}((x_1, h), \boldsymbol{\phi}, t)$ where

$$f_{reg}((x_1,h),\boldsymbol{\phi},t) = \varepsilon_{reg}\int_0^{M_t}\int_{A_{T_p}}\left[\left(\beta 1\frac{\partial T_p}{\partial x_1}((x_1,h),\boldsymbol{\phi},t)\right)^2 + \left(\beta_2\frac{\partial T_p}{\partial t}((x_1,h),\boldsymbol{\phi},t)\right)^2\right]dA\,dt \quad (31)$$

where A_{T_p} is the area defined by the $x_2 = h$ edge and the scaling constants β_1 and β_2 are 1 m°C^{-1} and 1 s°C^{-1}, respectively, for all our examples. The amount

of regularization is controlled through the regularization constant ε_{reg} and the derivative $\partial T_p/\partial t$ is obtained from equation (24). ($\partial T^p/\partial x_1$ is similarly obtained.)

To incorporate the regularization into the first- and second-order sensitivity expressions, we differentiate f_{reg} with respect to the model parameters. Thus, the first-order sensitivity is augmented with

$$\frac{\partial f_{reg}}{\partial \phi_i} = \varepsilon_{reg}\int_0^{M_t}\int_{A_{T_p}}\left[2\beta_1\frac{\partial T_p}{\partial x_1}\frac{\partial}{\partial \phi_i}\left(\frac{\partial T_p}{\partial x_1}\right) + 2\beta_2\frac{\partial T_p}{\partial t}\frac{\partial}{\partial \phi_i}\left(\frac{\partial T_p}{\partial t}\right)\right]dA\,dt \quad (32)$$

where $\partial/\partial\phi_i\partial(\partial T_p/\partial t)/\partial\phi_i$ and $\partial/\partial\phi_i\partial(\partial T_p/\partial x_1)/\partial\phi_i$ are obtained via equation (27). Likewise, the second-order sensitivity expression is augmented with

$$\frac{\partial^2 f_{reg}}{\partial\phi_i\partial\phi_j} = \varepsilon_{reg}\int_0^{M_t}\int_{A_{T_p}}\left[2\beta_1\left\{\frac{\partial}{\partial\phi_i}\left(\frac{\partial T_p}{\partial x_1}\right)\times\frac{\partial}{\partial\phi_j}\left(\frac{\partial T_p}{\partial x_1}\right)+\frac{\partial T_p}{\partial x_1}\frac{\partial^2}{\partial\phi_i\partial\phi_j}\left(\frac{\partial T_p}{\partial x_1}\right)\right\} + 2\beta_2\left\{\frac{\partial}{\partial\phi_i}\left(\frac{\partial T_p}{\partial t}\right)\frac{\partial}{\partial\phi_j}\left(\frac{\partial T_p}{\partial t}\right) + \frac{\partial T_p}{\partial t}\frac{\partial^2}{\partial\phi_i\partial\phi_j}\left(\frac{\partial T_p}{\partial t}\right)\right\}\right]dA\,dt \quad (33)$$

where, for all B-spline interpolations, $\partial^2(\partial T_p/\partial T)/\partial\phi_i\partial\phi_j = \partial^2(\partial T_p/\partial x_1)/\partial\phi_i\partial\gamma_j = 0$.

Regularization is now incorporated in the case number 1 IHCP. Figure 2(e) shows the computed T_p profile obtained via Newton's method and $\varepsilon_{reg} = 10^{-6}$. Except for subtle variations this T_p profile agrees with the $\bar{T}_p$ profile. Convergence is, once again, achieved in one iteration after which $F_r(\phi) = 3.36\times10^{-5}$ and $F(\phi) = 1.03\times10^{-5}$. The near zero value of $F(\phi)$ indicates that the regularization did not adversely affect the solution to the IHCP.

The BFGS method is also used to solve the case number 1 IHCP. A larger regularization constant, $\varepsilon_{reg} = 10^{-3}$ is used since the BFGS method fails to converge for $\varepsilon_{reg} = 10^{-6}$. The computed T_p profile after 40 iterations is shown in Fig. 2(f) and yields $F_r(\phi) = 8.87\times10^{-4}$ and $F(\phi) = 1.03\times10^{-5}$. It is observed that this T_p profile once again agrees with $\bar{T}_p$ profile and that $F(\phi)$ is again close to zero. However, the number of iterations, combined with the relatively larger value of ε_{reg} makes this method unattractive when compared to Newton's method.

4.3.2. *Nonlinear systems. Case number* 4. We now proceed to more computationally demanding examples in which the conduction problem is nonlinear because $k(T) = 10+0.25\ T^2$ Wm^{-1} $°C^{-1}$, and $e(T) = 0.5\ T^2$ Jm^{-3}. The transient analysis is performed for $M = 10$ time steps, with an increment of $\Delta t = 0.005$ s. The computational time for the primal analysis is significantly increased because of the nonlinear nature of the system.

We also use a different, more irregular parameterization of the B-spline surface to define the prescribed temperature profile $\bar{T}_p$ which is used to obtain the over-prescribed flux, q_{data} on the $x_2 = 0$ edge. this $\bar{T}_p$ surface uses a 4×4 array of vertices (i.e. $N = 16$) and is illustrated in Fig. 4(a). The initial T_p profile for the inverse analyses, which also uses a 4×4 array of vertices, is illustrated in Fig. 4(b) and clearly does not agree with $\bar{T}_p$. An analysis with this initial T_p yields $F(\phi) = 189.5$. The effectiveness of Newton's method and its considerable computational advantage over the BFGS method has been demonstrated. Therefore, we only use Newton's method to solve this IHCP.

Newton's method, with regularization did not converge after 10 iterations. Moreover, F_r did not even show signs of monotonic convergence ($\varepsilon_{reg} = 10^{-6}$). This behavior is possibly attributed to a small radius of convergence of the IHCP. To rectify this problem the BFGS method is used for the first few iterations, until the error falls to the order of 10^{-1}, i.e. $F_r(\phi) < 10^{-1}$. For all subsequent iterations we use Newton's method as presumably ϕ now lies within the radius of convergence.

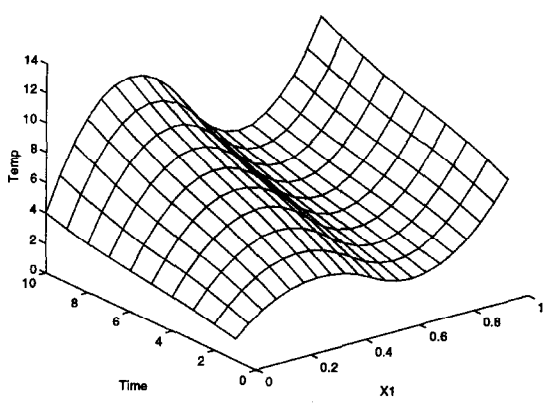
The combined BFGS–Newton's method is used to solve the IHCP with regularization of $\varepsilon_{reg} = 10^{-6}$. Convergence is achieved in 11 iterations (seven BFGS and four Newton iterations) after which $F_r(\phi) = 7.29\times10^{-4}$ and $F(\phi) = 1.07\times10^{-5}$. The combined BFGS–Newton's method failed to converge when the same analysis is repeated for 20 time steps for $\varepsilon_{reg} = 10^{-6}$. This is attributed to the degradation of the Hessian. For such instances the sequential estimation procedure should be implemented to solve the IHCP. In essence, the sequential method reduces the minimization problem of F on $\mathscr{R}^n$ to a series of minimization problems of F_i on $\mathscr{R}^m$ where $m < n$, as only those ϕ_j to which F_i is most sensitive are included in the minimization of each F_i. Thus, the Hessian remains well conditions for each F_i minimization problem and, hence, we can apply our methods to minimize each F_i.

Case number 5. Until now, the over prescribed surface flux distribution, q_{data}, is assumed to be smooth and contain no measurement errors. However, in practice q_{data} is obtained through experimental measurement, and therefore noise and errors are present. To study the effects of noise and errors in the surface flux, q_{data} is replaced with $\hat{q}_{data}$, where

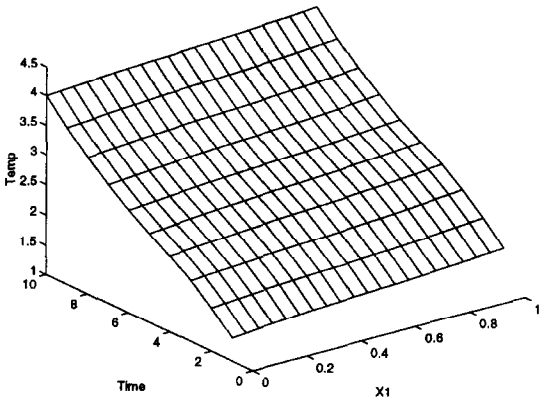
$$\hat{q}_{data} = q_{data} + \varepsilon_{tol}\,\mathrm{RAND}(x) \quad (34)$$

is the 'measured' surface flux, RAND(x) is a random number between -1.0 and 1.0, and ε_{tol} is the error constant. An additive error is assumed for simplicity, since these errors, in contrast to multiplicative errors, do not vary greatly with the independent variables, e.g. position [42].

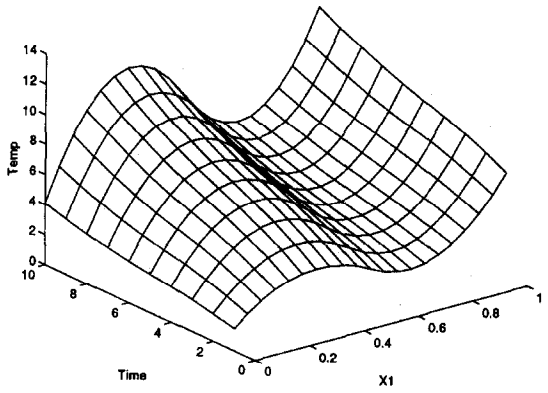
Case number 4 is now recomputed with $\varepsilon_{tol} = 1.0$



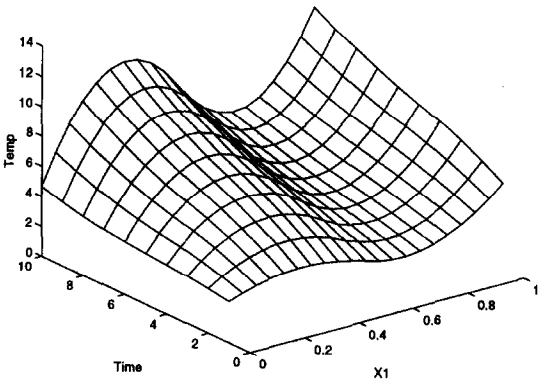
(a): Over prescribed profile



(b): Initial temperature profile



(c): $\varepsilon_{reg} = 10^{-6}$



(d)

Fig. 4. Case number 4: (a) prescribed; (b) initial, converged temperature profiles with (c) $\varepsilon_{reg} = 10^{-6}$ and (d) Case number 5: computed T_p profile.

and $\varepsilon_{reg} = 10^{-6}$. The combined BFGS–Newton's method is used to solve the IHCP and convergence is achieved in 12 iterations (seven BFGS and five Newton iterations) after which $F_r(\phi) = 8.59 \times 10^{-3}$ and $F(\phi) = 1.56 \times 10^{-4}$. The computed profile for T_p shown in Fig. 4(d) agrees with the prescribed $\bar{T}_p$ profile.

5. CONCLUSIONS

Analytical first- and second-order sensitivities are derived for nonlinear, transient systems using the direct/adjoint and hybrid methods, respectively. The sensitivity expressions are implemented in a numerical minimization algorithm based on Newton's method to solve an IHCP. The need for regularization to stabilize the solution is shown via a parameterization of the model parameters in the error function's Hessian eigen basis. Since the Hessian is available, it is noted that stability analysis can easily be performed for these solutions and although the simultaneous estimation procedure is used here, the methodology can also be used with the sequential estimation procedure. Finally, while the IHCP in this work are compared to known numerical solutions, the suggested methods can easily be adapted to solve IHCP based on available experimental data.

The Newton's method results for several examples are compared to those obtained with the first-order variable metric BFGS method. Newton's method is shown to be superior to the BFGS method because of its limited dependence on regularization and faster convergence. For nonlinear thermal systems, Newton's method fails and so a combined BFGS-Newton's method is used to solve the IHCP.

Acknowledgements—The authors would like to acknowledge the financial support provided by the National Science Foundation (grant NYI-DMI 9358132).

REFERENCES

1. Schnur, D. S. and Zabaras, N., Finite element solution of two-dimensional inverse elasticity problems using spatial smoothing. *International Journal for Numerical Methods in Engineering*, 1990, **30**, 57–75.
2. Nishmoura, D. and Kobayashi, S., A boundary integral equation method for an inverse problem related to crack

detection. *International Journal for Numerical Methods in Engineering*, 1991, **32**, 1371–1387.
3. Beck, J. V., Blackwell, B. and St. Clair, C. R. Jr., *Inverse Heat Conduction—Ill Posed Problems*. Wiley, New York, 1985.
4. Reinhardt, H. J. and Valencia, L., The numerical solution of inverse heat conduction problems with applications to reactor technology. In *SMIRT Conference*, 1987, pp. B25–30.
5. Cannon, J. R., *Numerical Solutions of Nonlinear Differential Equations*. Wiley, New York, 1966.
6. Stolz, G. Jr., Numerical solutions to an inverse problem of heat conduction for simple shapes. *Transactions of the ASME, Journal of Heat Transfer*, 1960, **82c**, 20–26.
7. Özişik, M. N., *Heat Conduction*. Wiley, New York, 1980.
8. Beck, J. V., Surface heat flux determination using an integral method. *Nuclear Engineering Design*, 1968, **7**, 170–178.
9. Sparrow, E. M., Haji-Sheihk and Lundgren, T. S., The inverse problem in transient heat conduction. *Transactions of the ASME, Journal of Heat Transfer*, 1964, **86E**, 369–375.
10. Frank, I., An application of least square method to the solution of the inverse problems of heat conduction. *Journal of Heat Transfer, Transactions of the ASME*, 1963, **83**(4), 378–379.
11. Beck, J. V., Nonlinear estimation applied to the nonlinear heat conduction problem. *International Journal of Heat and Mass Transfer*, 1970, **13**, 703–716.
12. Beck, J. V., Litkouhi, B. and St. Clair, C. R., Efficient sequential solution of the nonlinear inverse heat conduction problem. *Numerical Heat Transfer*, 1982, **5**, 275–286.
13. Liu, J., A stability analysis on beck's procedure for inverse heat conduction problems. *Journal of Computational Physics*, 1995, **123**, 65–73.
14. Reinhardt, H. J., A numerical method for the solution of two-dimensional inverse heat conduction problem. *International Journal for Numerical Methods in Engineering*, 1991, **32**, 363–383.
15. Tikhnov, A. N. and Arsenini, V. Y., *Solution of Ill-Posed Problems*. V. H. Winston and Sons, Washington, 1977.
16. Alifanov, O. M. and Artyukhin, E. A., Regularized numerical solution of nonlinear inverse heat conduction problem. *Journal of Engineering Physics*, 1975, **29**(1), 934–938.
17. Alifanov, O. M. and Mikhailov, V. V., Solution of the nonlinear inverse thermal conductivity problem by the iteration method. *Journal of Engineering Physics*, 1978, **35**(6), 1501–1506.
18. Beck, J. V. and Murio, D. A., Combined function specification-regularization procedure for solution of inverse heat conduction problem. *AIAA Journal*, 1986, **24**, 180–185.
19. Maniatty, A. M. and Zabaras, N. J., Investigation of regularization parameters and error estimating in inverse elasticity problems. *International Journal for Numerical Methods in Engineering*, 1994, **37**, 1039–1052.
20. Busby, H. R. and Trujillo, D. M., Numerical solution to a two-dimensional inverse heat conduction problem. *International Journal for Numerical Methods in Engineering*, 1985, **21**, 349–359.
21. Murio, D. A., The mollification method and the numerical solution of an inverse heat conduction problem. *SIAM Journal of Science Statistics in Computing*, 1981, **2**, 17–34.
22. Murio, D. A., The mollification method and the numerical solution of the inverse heat conduction problem by finite differences. *Computing Mathematics Applications*, 1989, **17**, 1385–1396.
23. Burggraf, O. R., An exact solution of the inverse problem in heat conduction theory and application. *Transactions of ASME, Journal of Heat Transfer*, 1964, **86C**, 373–382.
24. Beck, J. V., Criteria for comparison of methods of solution of the inverse heat conduction problem. *Nuclear Engineering and Design*, 1977, **53**, 11–22.
25. Alifanov, O. M., Solution of an inverse problem of heat conduction by iteration methods. *Journal of Engineering Physics*, 1974, **26**(4), 471–476.
26. Jarny, Y., Özişik, M. N. and Bardon, J. P., A general optimization method using adjoint equation for solving multidimensional inverse heat conduction. *International Journal of Heat and Mass Transfer*, 1991, **34**(11), 2911–2919.
27. Huang, C. H. and Özişik, M. N., Inverse problem of determining unknown wall heat flux in laminar flow through a parallel plate duct. *Numerical Heat Transfer—Part A*, 1992, **21**, 55–70.
28. Scheuing, J. E. and Tortorelli, D. A., Inverse heat conduction problem solutions via second-order design sensitivities and Newton's method. *Inverse Problems in Engineering*, 1996, **2**, 227–262.
29. Vanderplaats, G. N., *Numerical Optimization Techniques for Engineering Design: With Applications*. McGraw-Hill, New York, 1984.
30. Tortorelli, D. A. and Michaleris, P., Design sensitivity analysis: overview and review. *Inverse Problems in Engineering*, 1994, **1**, 71–105.
31. Michaleris, P., Tortorelli, D. A. and Vidal, C. A., Tangent operators and design sensitivity formulations for transient nonlinear coupled problems with applications to elasto-plasticity. *International Journal for Numerical Methods in Engineering*, 1995, **37**(2), 71–105.
32. Truffart, B., Jarny, Y. and Bourouga, B., A general approach to solve 2-d inverse heat conduction problems. In *Proceedings of the Seventh Inverse Problems in Engineering*, Ohio, 1995.
33. Belegundu, A. D., The direct and adjoint methods of design sensitivity analysis. *ASME Paper*, 1984, **84**-DET, 211.
34. Haftka, R. T. and Mróz, Z., First- and second-order sensitivity analysis of linear and nonlinear structures. *AIAA Journal*, 1986, **24**(7), 1187–1192.
35. Incropera, F. D. and DeWitt, D. P., *Fundamentals of Heat and Mass Transfer*, 3rd edn. Wiley, New York, 1990.
36. Bathe, K. J., *Finite Element Procedures in Engineering Analysis*. Prentice-Hall, Englewood Cliffs, New Jersey, 1982.
37. Rogers, David F. and Adams, J. Alan., *Mathematical Elements for Computer Graphics*. McGraw-Hill Company, New York, 1937.
38. Gill, P. E., Murray, W. and Wright, M. H., *Practical Optimization*, 9th edn. Academic Press, New York, 1992.
39. Papalambros, P. Y. and Wilde, D. J., *Principles of Optimal Design*. Cambridge University Press, New York, 1988.
40. VMA Engineering. *DOT User's Manual*, Version 3.00. Vanderplaats, Miura and Associates, Goleta, CA, 1992.
41. Strang, G., *Linear Algebra and Its Applications*, 3rd edn. Harcourt Brace Jovanovich College Publishers, New York, 1988.
42. Beck, J. V. and Arnold, K. J., *Parameter Estimation in Engineering and Science*. Wiley, New York, 1977.