Computer Assisted Mechanics and Engineering Sciences, 2: 253–270, 1995.

Parameter sensitivity in nonlinear transient thermal problems

Michał Kleiber
Institute of Fundamental Technological Research, Polish Academy of Sciences
ul. Świętokrzyska 21, 00-049 Warsaw, Poland

Andrzej Służalec
Technical University of Częstochowa
ul. Armii Krajowej 21, 42-200 Częstochowa, Poland

(Received March 30, 1995)

A general problem of parameter sensitivity of non-linear transient thermal systems is considered. The non-linear sensitivity path is followed by a weighted residual method employing the continuum description. The resulting finite element equations are derived. Both the direct differentiation and adjoint system methods are employed to evaluate sensitivity functional increments during the integration time step. Numerical results illustrate the method proposed.

1. Introduction

Having efficient computational means for realistic assessment of the nonlinear response of bodies subject to thermal loadings is crucial for solving very many engineering problems. Sophisticated FEM-based algorithms have therefore been developed to this purpose, see [2, 9] for instance. A natural extension of the analysis capabilities has been the developments in the parameter (or design) sensitivity area for the thermal problems. The so-called design sensitivity analysis (DSA) consists in computing variations in response quantities with respect to parameters (or design variables) entering the theory. The sensitivity information so obtained may be used to assess the effect of uncertainties in the mathematical model, to predict the response changes due to a change in the parameters, and to optimize the system by using appropriate optimization techniques.

Even if in the last decades DSA has attracted considerable attention in the field of solid and structural mechanics, relatively little work has been published on it in the thermal problems literature.

The subject of the design sensitivity analysis in the specific context of thermal systems was apparently first undertaken in [7]. The steady-state and transient linear and nonlinear problems were there discussed in the discrete FEM formalism, and a 1D illustration was provided. The two basic methodologies for solving sensitivity problems, later called the direct differentiation method (DDM) and the adjoint system method (ASM), were identified. This enlightening contribution was followed by the articles [3, 4, 11, 12]. In the two-part paper [3, 4] the author put forward the basic continuum formulation for DSA of linear isotropic thermal systems. Almost every aspect relevant to DSA for such class of problems was included in the formulation: non-shape and shape sensitivity techniques, DDM and ASM, steady-state and transient problems, first-and second-order sensitivities, etc. The shape sensitivity approach was based on the so-called material derivative concept while the ASM took advantage of the Lagrange multiplier technique resulting in a terminal-value problem for adjoint temperature field. A time mapping was introduced to transform the terminal-value problem into an initial-value problem. In papers [11, 12] a formalism was developed for solving thermal sensitivity problems by the boundary element method. In [12] nonlinear steady-state problem was considered and equations describing shape design sensitivity and shape optimization were given using the material derivative concept within the adjoint system methodology.

First-order design sensitivity for transient conduction problems was analyzed in [14] by the adjoint method. The approach was based on the mutual energy principles with a terminal-value problem resulting for the adjoint sensitivity equation. The so-called domain parametrization method was used for the analysis of shape sensitivities. Both continuum and discretized formalism were presented.

Shape and non-shape sensitivity formulations for transient nonlinear thermal problems were subject to scrutinity in [15]. ASM involving the Lagrange multiplier technique was used with convolution replacing multiplication as the operator between the multipliers and the governing equations. This resulted in an initial-value adjoint problem with no need for additional time mappings. Both continuum and FEM formulations were employed. In [13] the authors proposed to use the so-called Kirchhoff transformation for shape sensitivity analysis in nonlinear heat conduction problems.

In some more recent literature [6, 16] the thermal sensitivity analysis is considered as a part of more general thermo-elastic formulations; the review of the work is beyond the scope of this article, though.

It appears on the basis of the above discussion that a uniform treatment of the continuum and discretized, DDM-and ASM-based sensitivity formulations valid for fully nonlinear steady-state and transient heat conduction problems with anisotropic material properties and various boundary conditions has never been explicitly given. Such a review task, along with the discussion of some computational aspects, is undertaken in this article.

2. Continuous formulation

Design of thermal processing systems requires selecting of some design parameters $h_d \in R^D$ which describe the material properties and boundary conditions. The parameters can also represent the geometry (i.e. shape) of the domain Ω in which the problem is defined; only non-shape design problems are considered in this paper, though. Since in the sensitivity analysis the nature of the design dependence of particular functions is of crucial significance, the parameters h_d, $d = 1, 2, \ldots, D$ will be explicitly indicated below when formulating the heat transfer initial-boundary value problem.

The governing equations of heat transfer in a thermally anisotropic 3D region Ω can be written in a differential form as follows $(i, j = 1, 2, 3)$

$$(k_{ij}T_{,j})_{,i} + Q = \tilde{c}\frac{\partial T}{\partial \tau}\,, \qquad (x, \tau, h) \in \Omega \times \mathcal{T} \times R^D\,, \tag{1}$$

with the boundary conditions imposed on the boundary surface temperature,

$$T = \hat{T}\,, \qquad (x, \tau, h) \in \partial\Omega_T \times \mathcal{T} \times R^D\,, \tag{2}$$

and the boundary surface heat flux,

$$-k_{ij}n_jT_{,i} = \hat{q}\,, \qquad (x, \tau, h) \in \partial\Omega_q \times \mathcal{T} \times R^D\,, \tag{3}$$

and the initial condition imposed on the initial temperature distribution

$$T = \hat{T}_0\,, \qquad (x, \tau, h) \in \Omega \times \{t_0\} \times R^D\,, \tag{4}$$

where

$T = T(x, \tau, h)$ is the temperature,
$k_{ij} = k_{ij}(T(x, \tau, h), x, h)$ is the thermal conductivity tensor,
$Q = Q(T(x, \tau, h), T_{,i}(x, \tau, h), x, \tau, h)$ is the rate of heat generated per unit volume,

$\tilde{c} = \tilde{c}(T(x,\tau,h),x,h) = \rho c$	is the material heat capacity,
$\rho = \rho(T(x,\tau,h),x,h)$	is the density of material,
$c = c(T(x,\tau,h),x,h)$	is the specific heat of material,
x	is the position vector which identifies materials particles in the domain Ω,
τ	denotes time in the domain $\mathcal{T} = [t_0, t_f]$, t_0 and t_f being the initial and terminal time instances, respectively,
$\hat{T} = \hat{T}(x,\tau,h)$	is the temperature acting on the boundary surface $\partial\Omega_T$,
$\hat{q} = \hat{q}(T(x,\tau,h),x,\tau,h)$	is the heat flux on the complementary boundary surface $\partial\Omega_q$,
$n_i = n_i(x)$	is the unit outward-drawn vector normal to $\partial\Omega$,
$\hat{T}_0 = \hat{T}_0(x,h)$	is the initial temperature

and for any function g, the notation $g_{,i}$ stands for the partial differentiation of g with respect to the spatial coordinate x_i.

The Fourier's constitutive relation reads

$$q_i = -k_{ij}T_{,j}\,, \tag{5}$$

q_i being the heat flux vector.

Equation (3) may be specified to include convection boundary conditions on a part $\partial\Omega_q^{(1)}$ of $\partial\Omega_q$,

$$-k_{ij}n_jT_{,i} = \xi_{(c)}(T - T_\infty)\,, \qquad (x,\tau,h) \in \partial\Omega_q^{(1)} \times \mathcal{T} \times R^D\,, \tag{6}$$

where $\xi_{(c)}$ is the (possibly temperature dependent) convection coefficient, and radiation boundary conditions on a part $\partial\Omega_q^{(2)}$ of $\partial\Omega_q$,

$$-k_{ij}n_jT_{,i} = \xi_{(r)}\left(T^4 - T_{(r)}^4\right) = \chi\left(T - T_{(r)}\right) \tag{7}$$

in which the coefficient $\xi_{(r)}$ is computed as

$$\xi_{(r)} = \sigma V\left(\frac{1}{\varepsilon} + \frac{1}{\varepsilon_{(r)}} - 1\right)^{-1} \tag{8}$$

and

$$\chi = \xi_{(r)}\left(T^2 + T_{(r)}^2\right)\left(T + T_{(r)}\right) \tag{9}$$

where σ is the Stefan-Boltzmann constant, $T_{(r)}$ is the temperature of known external radiation source, V is the radiation view factor, ε is the surface emissivity and $\varepsilon_{(r)}$ the emissivity of the radiation source.

Looking for an approximate temperature solution to the above initial-boundary value problem we usually form the residuals (generally non-zero if T is only approximate)

$$r_1 = -(k_{ij}T_{,j})_{,i} - Q + \tilde{c}\frac{\partial T}{\partial \tau} \tag{10}$$

$$r_2 = \hat{q} + k_{ij}n_jT_{,i} \tag{11}$$

and then solve the problem (1)–(4) by determining the square integrable temperature field T satisfying the temperature boundary condition and zeroeing the following weighted residual

$$R = \int_\Omega r_1\phi\,\mathrm{d}\Omega + \int_{\partial\Omega_q} r_2\phi\,\mathrm{d}(\partial\Omega) = 0 \tag{12}$$

for all square integrable weighting functions $\phi(x)$ that vanish on $\partial\Omega_T$. The residual (12) can be transformed as follows

$$\begin{aligned} R &= \int_\Omega \left[-(k_{ij}T_{,j})_{,i}\,\phi + \left(\tilde{c}\frac{\partial T}{\partial \tau} - Q\right)\phi \right] \mathrm{d}\Omega + \int_{\partial\Omega_q} (\hat{q} + k_{ij}n_jT_{,i})\,\phi\,\mathrm{d}(\partial\Omega) \\ &= -\int_{\partial\Omega_q} k_{ij}n_jT_{,i}\phi\,\mathrm{d}(\partial\Omega) + \int_\Omega \left[k_{ij}T_{,j}\phi_{,i} + \left(\tilde{c}\frac{\partial T}{\partial \tau} - Q\right)\phi \right] \mathrm{d}\Omega + \int_{\partial\Omega_q} (\hat{q} + k_{ij}n_jT_{,i})\,\phi\,\mathrm{d}(\partial\Omega) \\ &= \int_\Omega \left[k_{ij}T_{,j}\phi_{,i} + \left(\tilde{c}\frac{\partial T}{\partial \tau} - Q\right)\phi \right] \mathrm{d}\Omega + \int_{\partial\Omega_q} \hat{q}\phi\,\mathrm{d}(\partial\Omega) = 0\,. \end{aligned} \tag{13}$$

At any given time instant τ, Eq. (13) is clearly nonlinear in T. To solve it for T we may use the iterative technique of Newton-Raphson which is based on zeroeing the 'next' (k)-th residual written as

$$R^{(k)} \cong R^{(k-1)} + R_T^{(k-1)} * \delta T^{(k)} = 0 \tag{14}$$

in which $R^{(k-1)} = R^{(k-1)}(T(x,\tau,h),x,\tau,h)$ corresponds to the 'last' $(k-1)$-th approximation to the temperature field $T = T^{(k-1)}$ assumed known, $\delta T^{(k)}$ is the iterative correction to be determined from Eq. (14) such that

$$T^{(k)} = T^{(k-1)} + \delta T^{(k)} \tag{15}$$

and $R_T^{(k-1)} = \mathrm{d}R^{(k-1)}/\mathrm{d}T$ is the $(k-1)$-th tangent operator defined by

$$\begin{aligned} R_T = &\int_\Omega \left[\frac{\partial k_{ij}}{\partial T}T_{,j}\phi_{,i} + k_{ij}\phi_{,i}\frac{\partial}{\partial x_j} + \phi\left(\tilde{c}\frac{\partial}{\partial\tau} + \frac{\partial\tilde{c}}{\partial T}\frac{\partial T}{\partial \tau} - \frac{\partial Q}{\partial T} - \frac{\partial Q}{\partial T_{,i}}\frac{\partial}{\partial x_i}\right)\right] \mathrm{d}\Omega \\ &+ \int_{\partial\Omega_q} \phi\frac{\partial\hat{q}}{\partial T}\,\mathrm{d}(\partial\Omega)\,. \end{aligned} \tag{16}$$

The notation $R_T * \delta T$ should be clear from the context with $*$ indicating that δT multiplies the appropriate integrand rather than the whole integrand expression that defines R_T. The operator R_T depends nonlinearly on T, linearly on ϕ and acts linearly on δT; we shall use the notation

$$\begin{aligned} R_T[T;\phi] * \delta T &= \hat{R}_T[T;\phi;\delta T] \\ &= \int_\Omega \left[\frac{\partial k_{ij}}{\partial T}T_{,j}\phi_{,i}\delta T + k_{ij}\phi_{,i}\delta T_{,j} \right. \\ &\qquad \left. +\phi\left(\tilde{c}\frac{\partial\delta T}{\partial\tau} + \frac{\partial\tilde{c}}{\partial T}\frac{\partial T}{\partial\tau}\delta T - \frac{\partial Q}{\partial T}\delta T - \frac{\partial Q}{\partial T_{,i}}\delta T_{,i}\right)\right] \mathrm{d}\Omega \\ &\quad + \int_{\partial\Omega_q} \phi\frac{\partial\hat{q}}{\partial T}\delta T\,\mathrm{d}(\partial\Omega) \end{aligned} \tag{17}$$

or, shorter,

$$R_T[T;\phi] * \delta T = \left(K_T^{(1)}[T;\phi] + K_T^{(2)}[T;\phi] + C[T;\phi]\frac{\partial}{\partial\tau}\right) * \delta T = \left(K_T[T;\phi] + C[T;\phi]\frac{\partial}{\partial\tau}\right) * \delta T \tag{18}$$

where

$$K_T^{(1)} = \int_\Omega \left(k_{ij}\phi_{,i}\frac{\partial}{\partial x_j} + \frac{\partial k_{ij}}{\partial T}T_{,j}\phi_{,i} - \phi\frac{\partial Q}{\partial T} - \phi\frac{\partial Q}{\partial T_{,i}}\frac{\partial}{\partial x_i}\right) \mathrm{d}\Omega + \int_{\partial\Omega_q}\phi\frac{\partial\hat{q}}{\partial T}\,\mathrm{d}(\partial\Omega)\,, \tag{19}$$

$$K_T^{(2)} = \int_\Omega \phi\frac{\partial\tilde{c}}{\partial T}\frac{\partial T}{\partial\tau}\,\mathrm{d}\Omega\,, \tag{20}$$

$$K_T = K_T^{(1)} + K_T^{(2)}\,, \tag{21}$$

$$C = \int_\Omega \phi\tilde{c}\,\mathrm{d}\Omega\,. \tag{22}$$

The iterative procedure can be seen more clearly for a time-discretized formulation in which we assume that:

(a) solution up to a typical time instant t has been obtained,

(b) solution $T_{t+\Delta t}$ at time $t+\Delta t$ is looked for,

(c) a finite difference scheme in time such as the one-step backward Euler scheme is employed so that

$$\dot{T}_{t+\Delta t} = \frac{1}{\Delta t}(T_{t+\Delta t} - T_t) \,. \tag{23}$$

and consequently

$$\dot{T}^{(k)}_{t+\Delta t} = \frac{1}{\Delta t}\left(T^{(k)}_{t+\Delta t} - T_t\right),$$

$$\delta T^{(k)} = \delta T^{(k)}_{t+\Delta t} = \Delta t\, \delta \dot{T}^{(k)}_{t+\Delta t},$$

$$\frac{\partial}{\partial \tau}\delta T^{(k)} = \frac{\partial}{\partial \tau}\delta T^{(k)}_{t+\Delta t} = \frac{1}{\Delta t}\delta T^{(k)}_{t+\Delta t} = \frac{1}{\Delta t}\delta T^{(k)} \,.$$

Equation (17) written at $\tau = t+\Delta t$ then becomes

$$\begin{aligned} R_T^{(k-1)}[T;\phi] * \delta T^{(k)} = \int_\Omega \Bigg[& \frac{\partial k_{ij}^{(k-1)}}{\partial T} T_{,j}^{(k-1)} \phi_{,i} + k_{ij}^{(k-1)} \phi_{,i} \frac{\partial}{\partial x_j} \\ & + \phi\left(\tilde{c}^{(k-1)} \frac{1}{\Delta t} + \frac{\partial \tilde{c}^{(k-1)}}{\partial T} \frac{1}{\Delta t}\left(T^{(k-1)} - T_t\right) \right. \\ & \left. - \frac{\partial Q}{\partial T} - \frac{\partial Q}{\partial T_{,i}} \frac{\partial}{\partial x_i}\right)\Bigg] \delta T^{(k)} \mathrm{d}\Omega + \int_{\partial\Omega_q} \phi \frac{\partial \hat{q}}{\partial T} \delta T^{(k)} \,\mathrm{d}(\partial\Omega) \end{aligned} \tag{24}$$

where all the functions (except for T_t!) are understood to be computed at time $t+\Delta t$ and the temperature value $T^{(k-1)}_{t+\Delta t}$. The operator equation (14) with the term $R_T^{(k-1)}\delta T^{(k)}$ given as Eq. (24) can be solved for $\delta T^{(k)}$ by any of the known techniques in use for solving PDE's with respect to space variables. It should be noted in this context that even though the operator R in Eq. (13) generates the symmetric finite element 'secant' stiffness matrix (dependent on T), the spatial discretization applied to Eq. (14) with (24) results in a non-symmetric tangent stiffness matrix which may unfavorably influence the efficiency of the solution procedure typically based on symmetric linear equations solvers. Therefore, different symmetric approximations to the non-symmetric tangent stiffness are used in practice with the non-symmetry effects accounted for in an iterative fashion; we shall further comment on this aspect in Section 3.

Let us now consider a functional

$$\begin{aligned} \mathcal{G}(h) = \int_{t_0}^{t_f} G \,\mathrm{d}\tau = \int_{t_0}^{t_f} \Bigg[& \int_\Omega G_1(T(x,\tau,h), T_{,i}(x,\tau,h), x, \tau, h)\,\mathrm{d}\Omega \\ & + \int_{\partial\Omega_q} G_2(T(x,\tau,h), q(x,\tau,h), x, \tau, h)\,\mathrm{d}(\partial\Omega)\Bigg] \mathrm{d}\tau \end{aligned} \tag{25}$$

whose design gradient $\mathrm{d}\mathcal{G}/\mathrm{d}h_d$ is of our concern in sensitivity analysis, G_1 and G_2 are assumed to be known functions of their indicated arguments and

$$q = q_i n_i \tag{26}$$

is the surface heat flux. The functional (25) may be a cost or constraint functional typical of optimization methodologies, for instance. We note that even though it is expressed in integral form, time and or space localized quantities may also be represented by proper choice of weighting functions (Dirac delta distributions, in particular). By differentiating $\mathcal{G}$ with respect to the design variable h_d we obtain

$$\frac{\mathrm{d}\mathcal{G}}{\mathrm{d}h_d} = \int_{t_0}^{t_f} \left[\int_\Omega \left(\frac{\partial G_1}{\partial T} T_d + \frac{\partial G_1}{\partial T_{,i}} \frac{\partial T_d}{\partial x_i} + G_{1d} \right) \mathrm{d}\Omega + \int_{\partial\Omega_q} \left(\frac{\partial G_2}{\partial T} T_d + \frac{\partial G_2}{\partial q} q_d + G_{2d} \right) \mathrm{d}(\partial\Omega) \right] \mathrm{d}\tau \tag{27}$$

where for any function g the notation g_d stands for the partial derivative of g with respect to the design variable h_d. By noting that the unit normal n_i is independent of h_d, the design derivative of the surface heat flux becomes

$$q_d = \left(\frac{\partial q_i}{\partial T_{,j}} \frac{\partial T_d}{\partial x_j} + \frac{\partial q_i}{\partial T} T_d + \frac{\partial q_i}{\partial h_d} \right) n_i \,. \tag{28}$$

We may now observe that knowing the time trajectories of T at any point x computed for the nominal value of the design vector $\mathbf{h}$ (i.e. having solved the heat transfer primary problem) the only quantity needed to effectively evaluate the gradient (27) is the gradient of temperature T with respect to the design variable h_d. Therefore we shall concentrate now on ways to compute T_d.

Let us start by differentiating Eq. (13) with respect to h_d. We obtain

$$\frac{\mathrm{d}R}{\mathrm{d}h_d} = \int_\Omega \left[\frac{\mathrm{d}k_{ij}}{\mathrm{d}h_d} T_{,j}\phi_{,i} + k_{ij}\phi_{,i}T_{d,j} + \phi \left(\tilde{c} \frac{\partial T_d}{\partial \tau} + \frac{\mathrm{d}\tilde{c}}{\mathrm{d}h_d} \frac{\partial T}{\partial \tau} - \frac{\mathrm{d}Q}{\mathrm{d}h_d} \right) \right] \mathrm{d}\Omega + \int_{\partial\Omega_q} \phi \frac{\mathrm{d}\hat{q}}{\mathrm{d}h_d} \mathrm{d}(\partial\Omega) \tag{29}$$

which by noting that the residual R should vanish at both the nominal h_d and perturbed $h_d + \delta h_d$ values of the design, and that

$$\frac{\mathrm{d}k_{ij}}{\mathrm{d}h_d} = \frac{\partial k_{ij}}{\partial T} T_d + \frac{\partial k_{ij}}{\partial h_d} \,, \tag{30}$$

$$\frac{\mathrm{d}\tilde{c}}{\mathrm{d}h_d} = \frac{\partial \tilde{c}}{\partial T} T_d + \frac{\partial \tilde{c}}{\partial h_d} \,, \tag{31}$$

$$\frac{\mathrm{d}Q}{\mathrm{d}h_d} = \frac{\partial Q}{\partial T} T_d + \frac{\partial Q}{\partial T_{,i}} T_{d,i} + \frac{\partial Q}{\partial h_d} \,, \tag{32}$$

$$\frac{\mathrm{d}\hat{q}}{\mathrm{d}h_d} = \frac{\partial \hat{q}}{\partial T} T_d + \frac{\partial \hat{q}}{\partial h_d} \,, \tag{33}$$

becomes

$$\begin{aligned} \frac{\mathrm{d}R}{\mathrm{d}h_d} = & \int_\Omega \left[\frac{\partial k_{ij}}{\partial h_d} T_{,j}\phi_{,i}T_d + k_{ij}\phi_{,i}T_{d,j} + \phi \left(\tilde{c} \frac{\partial T_d}{\partial \tau} + \frac{\partial \tilde{c}}{\partial T} \frac{\partial T}{\partial \tau} T_d - \frac{\partial Q}{\partial T} T_d - \frac{\partial Q}{\partial T_{,i}} T_{d,i} \right) \right] \mathrm{d}\Omega \\ & + \int_{\partial\Omega_q} \phi \frac{\partial \hat{q}}{\partial T} T_d \, \mathrm{d}(\partial\Omega) + \underline{\int_\Omega \left[\frac{\partial k_{ij}}{\partial h_d} T_{,j}\phi_{,i} + \phi \left(\frac{\partial \tilde{c}}{\partial h_d} \frac{\partial T}{\partial \tau} - \frac{\partial Q}{\partial h_d} \right) \right] \mathrm{d}\Omega + \int_{\partial\Omega_q} \phi \frac{\partial \hat{q}}{\partial h_d} \mathrm{d}(\partial\Omega)} . \end{aligned} \tag{34}$$

If we now observe that the underlined term in Eq. (34) can easily be computed (at the given time step, after the solution to the primary problem has been obtained) and compare Eq. (34) with Eq. (17), we conclude that Eq. (34) can be compactly rewritten as a linear in T_d equation of the form

$$\begin{aligned} R_T * T_d &= \left(K_T[T;\phi] + C[T;\phi] \frac{\partial}{\partial \tau} \right) * T_d \\ &= - \int_\Omega \left[\frac{\partial k_{ij}}{\partial h_d} T_{,j}\phi_{,i} + \phi \left(\frac{\partial \tilde{c}}{\partial h_d} \frac{\partial T}{\partial \tau} - \frac{\partial Q}{\partial h_d} \right) \right] \mathrm{d}\Omega - \int_{\partial\Omega_q} \phi \frac{\partial \hat{q}}{\partial h_d} \mathrm{d}(\partial\Omega) \end{aligned} \tag{35}$$

or

$$R_T * T_d = -\frac{\partial R}{\partial h_d}, \tag{36}$$

since the right-hand side of Eq. (35) is just the partial (explicit) derivative of the residual (13).

Equation (35) is the result of fundamental significance. It essentially says that having at a given time instant the temperature distribution T and the tangent operator R_T we may solve Eq. (35) for T_d at that instant which is the only unknown needed to compute the sensitivity functional (27). Moreover, while the primary problem requires iterations according to Eq. (14), the sensitivity gradient T_d can be found directly (i.e. with no iteration) from Eq. (35) in which we just need to use the last available value of the tangent operator R_T. The above characteristics of the sensitivity problem has a crucial significance for the efficiency of the appropriate computational process.

The method just discussed is called the direct differentiation method (DDM). An alternative approach known as the adjoint system method (ASM) attempts to compute directly (i.e. not by the way of the temperature sensitivities T_d) the functional sensitivity $\mathrm{d}\mathcal{G}/\mathrm{d}h_d$ by eliminating T_d through the use of Lagrange multipliers. To do so we consider an augmented functional

$$\hat{\mathcal{G}} = \mathcal{G} + R^{(a)} \tag{37}$$

in which $\mathcal{G}$ is given in Eq. (25) while $R^{(a)}$ is defined as

$$R^{(a)} = \int_{t_0}^{t_f} \left\{ \int_\Omega T^{(a)} \left[-(k_{ij}T_{,j})_{,i} + \left(\tilde{c}\frac{\partial T}{\partial \tau} - Q \right) \right] \mathrm{d}\Omega + \int_{\partial\Omega_q} T^{(a)} \left(\hat{q} + k_{ij}n_j T_{,i} \right) \mathrm{d}(\partial\Omega) \right\} \mathrm{d}\tau . \tag{38}$$

We observe that the integral of the time integration is just the residual (12) with the weighting function ϕ replaced by the adjoint temperature $T^{(a)}$; we require that the latter vanishes on $\partial\Omega_T$ but it is arbitrary otherwise. Equation (38) becomes

$$R^{(a)} = \int_{t_0}^{t_f} \left\{ \int_\Omega T^{(a)} \left[k_{ij}T_{,j}T^{(a)}_{,i} + \left(\tilde{c}\frac{\partial T}{\partial \tau} - Q \right) T^{(a)} \right] \mathrm{d}\Omega + \int_{\partial\Omega_q} \hat{q} T^{(a)} \,\mathrm{d}(\partial\Omega) \right\} \mathrm{d}\tau . \tag{39}$$

The variational formalism for the adjoint system method follows here the methodology suggested in [1]. In it, the adjoint variable $T^{(a)}$ is determined from the requirement that the augmented functional (37) be stationary with respect to the primary variable T, i.e.

$$\delta_T \hat{\mathcal{G}} = \frac{\partial \hat{\mathcal{G}}}{\partial T} * \delta T = 0 \tag{40}$$

which, by considering T_d as a possible specific case of δT, becomes

$$\frac{\partial \hat{\mathcal{G}}}{\partial T} * T_d = 0 . \tag{41}$$

According to the ASM variational methodology, the substitution into Eq. (37) of the adjoint variable $T^{(a)}$ computed from Eq. (41) implies that the desired response functional sensitivity $\mathrm{d}\mathcal{G}/\mathrm{d}h_d$ can be computed as

$$\frac{\mathrm{d}\mathcal{G}}{\mathrm{d}h_d} = \frac{\partial \hat{\mathcal{G}}}{\partial h_d} \tag{42}$$

in which the right-hand side involves only the explicit design differentiation of $\hat{\mathcal{G}}$ which is straightforward to compute as

$$\frac{\partial \hat{\mathcal{G}}}{\partial h_d} = \int_{t_0}^{t_f} \left\{ \int_\Omega \frac{\partial G_1}{\partial h_d} \mathrm{d}\Omega + \int_{\partial\Omega_q} \frac{\partial G_2}{\partial h_d} \mathrm{d}(\partial\Omega) \right.$$
$$\left. + \int_\Omega \left[\frac{\partial k_{ij}}{\partial h_d} T_{,j} T^{(a)}_{,i} + \left(\frac{\partial \tilde{c}}{\partial h_d}\frac{\partial T}{\partial \tau} - \frac{\partial Q}{\partial h_d} \right) T^{(a)} \right] \mathrm{d}\Omega + \int_{\partial\Omega_q} \frac{\partial \hat{q}}{\partial h_d} T^{(a)} \,\mathrm{d}(\partial\Omega) \right\} \mathrm{d}\tau . \tag{43}$$

The above result has been obtained by using the combination of Eqs. (25), (37) and (39) which yields

$$\hat{\mathcal{G}} = \int_{t_0}^{t_f} \left\{ \int_\Omega \left[G_1 + k_{ij} T_{,j} T_{,i}^{(a)} + \left(\tilde{c} \frac{\partial T}{\partial \tau} - Q \right) T^{(a)} \right] \mathrm{d}\Omega + \int_{\partial\Omega_q} \left(G_2 + \hat{q} T^{(a)} \right) \mathrm{d}(\partial\Omega) \right\} \mathrm{d}\tau \,. \tag{44}$$

The condition (41) is rewritten as

$$\begin{aligned} \frac{\partial \hat{\mathcal{G}}}{\partial T} * T_d = \int_{t_0}^{t_f} \Bigg\{ \int_\Omega \Bigg[& \frac{\partial G_1}{\partial T} T_d + \frac{\partial G_1}{\partial T_{,i}} \frac{\partial T_d}{\partial x_i} + \frac{\partial k_{ij}}{\partial T} T_d T_{,j} T_{,i}^{(a)} + k_{ij} \frac{\partial T_d}{\partial x_j} T_{,i}^{(a)} \\ & + \left(\frac{\partial \tilde{c}}{\partial T} T_d \frac{\partial T}{\partial \tau} + \tilde{c} \frac{\partial T_d}{\partial \tau} - \frac{\partial Q}{\partial T} T_d - \frac{\partial Q}{\partial T_{,i}} \frac{\partial T_d}{\partial x_i} \right) T^{(a)} \Bigg] \mathrm{d}\Omega \\ & + \int_{\partial\Omega_q} \left[\frac{\partial G_2}{\partial T} T_d + \frac{\partial G_2}{\partial q} n_i \left(\frac{\partial q_i}{\partial T} T_d + \frac{\partial q_i}{\partial T_{,j}} \frac{\partial T_d}{\partial x_j} \right) + \frac{\partial \hat{q}}{\partial T} T_d T^{(a)} \right] \mathrm{d}(\partial\Omega) \Bigg\} \mathrm{d}\tau \end{aligned} \tag{45}$$

which by integrating by parts in time the term $\tilde{c} \dot{T}_d T^{(a)}$ and re-ordering the terms becomes

$$\begin{aligned} \frac{\partial \hat{\mathcal{G}}}{\partial T} * T_d = & \int_{t_0}^{t_f} \Bigg\{ \int_\Omega \left[\frac{\partial G_1}{\partial T} T_d + \frac{\partial G_1}{\partial T_{,i}} \frac{\partial T_d}{\partial x_i} \right] \mathrm{d}\Omega \\ & + \int_{\partial\Omega_q} \left[\frac{\partial G_2}{\partial T} T_d + \frac{\partial G_2}{\partial q} n_i \left(\frac{\partial q_i}{\partial T} T_d + \frac{\partial q_i}{\partial T_{,j}} \frac{\partial T_d}{\partial x_j} \right) \right] \mathrm{d}(\partial\Omega) \\ & + \int_\Omega \Bigg[k_{ij} \frac{\partial T_d}{\partial x_j} \frac{\partial T^{(a)}}{\partial x_i} + \frac{\partial k_{ij}}{\partial T} T_d \frac{\partial T}{\partial x_j} \frac{\partial T^{(a)}}{\partial x_i} + \frac{\partial \tilde{c}}{\partial T} T_d \frac{\partial T}{\partial \tau} - \tilde{c} \frac{\partial T^{(a)}}{\partial \tau} T_d - \dot{\tilde{c}} T^{(a)} T_d \\ & - \left(\frac{\partial Q}{\partial T} T_d + \frac{\partial Q}{\partial T_{,i}} \frac{\partial T_d}{\partial x_i} \right) T^{(a)} \Bigg] \mathrm{d}\Omega \\ & + \int_{\partial\Omega_q} \frac{\partial \hat{q}}{\partial T} T_d T^{(a)} \, \mathrm{d}(\partial\Omega) \Bigg\} \mathrm{d}\tau + \int_\Omega T^{(a)} \tilde{c} T_d \, \mathrm{d}\Omega \Big|_{t_0}^{t_f} . \end{aligned} \tag{46}$$

Consequently, by noting additionally the notation introduced in Eq. (17) and the condition

$$T^{(a)}(t_0) = 0 \,, \tag{47}$$

we require that

$$\left(K_T[T; T^{(a)}] - \dot{C}[T; T^{(a)}] - C\left[T; \frac{\partial}{\partial \tau} T^{(a)}\right] \right) * T_d = - \frac{\partial G}{\partial T} * T_d \,, \tag{48}$$

$$T^{(a)} \Big|_{t_f} = 0 \,, \tag{49}$$

in which we used the following property of the operator C,

$$C[T; f] * \frac{\partial g}{\partial \tau} = C\left[T; \frac{\partial g}{\partial \tau}\right] * f \,, \qquad \text{for any } f,\ g. \tag{50}$$

Equation (48) is the fundamental equation for solving nonlinear, non-stationary thermal sensitivity problem by the ASM. It is a variational equation for the adjoint temperature field $T^{(a)}(x_i)$, which has to be satisfied for any T_d. The partial differential equation generated by Eq. (48) has to be solved backward in time; the time transformation

$$\tau' = t_f - \tau \tag{51}$$

allows to replace it by the initial-boundary value problem of the form

$$\left(K_T^{(1)}[T;T^{(a)}] - K_T^{(2)}[T;T^{(a)}] + \dot{C}[T;T^{(a)}] + C\left[T;\frac{\partial}{\partial\tau}T^{(a)}\right]\right) * T_d = -\frac{\partial G}{\partial T} * T_d\,, \tag{52}$$

$$T^{(a)}\Big|_{\tau'=0} = 0\,, \tag{53}$$

in which all the time derivatives are taken with respect to time τ'.

We may observe that for $\tilde{c} = \text{const}$ which implies $\dot{C} = 0$ and $K_T^{(2)} = 0$ we have

$$\left(K_T^{(1)}[T;T^{(a)}] + C\left[T;\frac{\partial}{\partial\tau}T^{(a)}\right]\right) * T_d = -\frac{\partial G}{\partial T}T_d\,. \tag{54}$$

Identifying the field T_d with the weighting function ϕ of Eq. (18), we obtain

$$\left(K_T^{(1)T}[T;\phi] + C[T;\phi]\frac{\partial}{\partial\tau}\right) * T^{(a)} = -\frac{\partial G}{\partial T} * \phi \tag{55}$$

where the following notation is introduced:

$$K_T^{(1)T}[T;\phi] * T^{(a)} = K_T^{(1)}[T;T^{(a)}] * \phi\,. \tag{56}$$

A remarkable formal similarity of the left-hand sides of Eq. (55) and DDM equation (36) is emphasized.

3. Semi-discretized formulation

By replacing in Eq. (13) the weighting function ϕ by ϕ_α , $\alpha = 1, 2, \ldots, N$, considering for compactness just one design parameter h and using the typical finite element expansion for temperature, which on the system level reads

$$T(x,\tau,h) = \phi_\alpha(x)\vartheta_\alpha(\tau,h)\,, \qquad \alpha = 1, 2, \ldots, N\,, \tag{57}$$

where ϕ_α are the shape functions, ϑ_α are the nodal temperatures and N is the total number of the degrees of freedom in the discretized system, we obtain from Eq. (12)

$$\int_\Omega \left(k_{ij}\phi_{\alpha,i}\phi_{\beta,j}\vartheta_\beta + \tilde{c}\phi_\alpha\phi_\beta\dot{\vartheta}_\beta - Q\phi_\alpha\right) \mathrm{d}\Omega + \int_{\partial\hat{\Omega}_q} \hat{q}\phi_\alpha \,\mathrm{d}(\partial\Omega)$$
$$+ \int_{\partial\Omega_q^{(1)}} \xi_{(c)}\left(\vartheta_\beta\phi_\beta - T_\infty\right)\phi_\alpha \,\mathrm{d}(\partial\Omega) + \int_{\partial\Omega_q^{(2)}} \chi\left(\vartheta_\beta\phi_\beta - T_{(r)}\right)\phi_\alpha \,\mathrm{d}(\partial\Omega) = 0 \tag{58}$$

in which

$$\partial\hat{\Omega}_q = \partial\Omega_q \setminus \left(\partial\Omega_q^{(1)} \cup \partial\Omega_q^{(2)}\right). \tag{59}$$

Denoting

$$K_{\alpha\beta}^{(k)} = \int_\Omega k_{ij}\phi_{\alpha,i}\phi_{\beta,j}\,\mathrm{d}\Omega\,, \tag{60}$$

$$K_{\alpha\beta}^{(c)} = \int_{\partial\Omega_q^{(1)}} \xi_{(c)}\phi_\alpha\phi_\beta \,\mathrm{d}(\partial\Omega)\,, \tag{61}$$

$$K_{\alpha\beta}^{(r)} = \int_{\partial\Omega_q^{(2)}} \chi\phi_\alpha\phi_\beta \,\mathrm{d}(\partial\Omega)\,, \tag{62}$$

$$C_{\alpha\beta} = \int_\Omega \tilde{c}\phi_\alpha\phi_\beta\,\mathrm{d}\Omega\,, \tag{63}$$

$$\hat{Q}_\alpha = \int_\Omega Q\phi_\alpha\,\mathrm{d}\Omega - \int_{\partial\hat{\Omega}_q} \hat{q}\phi_\alpha \,\mathrm{d}(\partial\Omega)\,, \tag{64}$$

$$\tilde{Q}_\alpha = \int_{\partial\Omega_q^{(1)}} \xi_{(c)}T_\infty\phi_\alpha \,\mathrm{d}(\partial\Omega) + \int_{\partial\Omega_q^{(2)}} \chi T_{(r)}\phi_\alpha \,\mathrm{d}(\partial\Omega)\,, \tag{65}$$

we obtain

$$C_{\alpha\beta}\dot{\vartheta}_\beta + \underbrace{\left(K^{(k)}_{\alpha\beta} + K^{(c)}_{\alpha\beta} + K^{(r)}_{\alpha\beta}\right)}_{K_{\alpha\beta}}\vartheta_\beta = \underbrace{\hat{Q}_\alpha + \tilde{Q}_\alpha}_{Q_\alpha} \tag{66}$$

or, shortly,

$$C_{\alpha\beta}\dot{\vartheta}_\beta + K_{\alpha\beta}\vartheta_\beta = Q_\alpha\,, \tag{67}$$

which is the fundamental equation describing nonlinear transient heat transfer in the discretized system. The matrices $K^{(k)}_{\alpha\beta}$, $K^{(c)}_{\alpha\beta}$ and $K^{(r)}_{\alpha\beta}$ may be briefly called the conductivity, convection and radiation stiffness matrices, respectively. It should be noted, however, that only in the linear analysis these names (in fact, only the first two since $K^{(r)}_{\alpha\beta} = 0$ then) appear fully legitimate.

The apparently elegant form of ordinary differential equation (67) is somewhat misleading since the matrices $C_{\alpha\beta}$ and $K_{\alpha\beta}$ (in fact, all the three terms in the latter) and vector Q_α may generally depend on the temperature ϑ_α. Therefore, Eq. (67) has to be first linearized with respect to ϑ_α and only then solved by using an iteration scheme. We also note that the matrices $C_{\alpha\beta}$ and $K_{\alpha\beta}$ are symmetric which is essential for efficiency of computational procedures.

Guided by the linear heat transfer equations we adopt the following linearization assumptions:

$$^{t+\Delta t}\chi\left(^{t+\Delta t}T - {}^{t+\Delta t}T_{(r)}\right) \cong {}^{t}\chi\left(^{t+\Delta t}T - {}^{t+\Delta t}T_{(r)}\right), \tag{68}$$

$$^{t+\Delta t}\xi_{(c)}\left(^{t+\Delta t}T - {}^{t+\Delta t}T_{\infty}\right) \cong {}^{t}\xi_{(c)}\left(^{t+\Delta t}T - {}^{t+\Delta t}T_{\infty}\right), \tag{69}$$

$$^{t+\Delta t}k_{ij}\,{}^{t+\Delta t}T_{,j} \cong {}^{t}k_{ij}\,{}^{t+\Delta t}T_{,j}\,, \tag{70}$$

$$^{t+\Delta t}\tilde{c}\,{}^{t+\Delta t}\dot{T} \cong {}^{t}\tilde{c}\,{}^{t+\Delta t}\dot{T}\,. \tag{71}$$

We note that (i) the above assumptions are automatically satisfied in the linear analysis in which $\xi_{(c)} = \text{const}$, $k_{ij} = \text{const}$ and $\tilde{c} = \text{const}$, (ii) iterative solution schemes may fully restore neglected terms in the nonlinear analysis.

By writing out Eq. (66) for the time instant $t + \Delta t$ and using the linearization (68)–(71) we arrive at

$$^{t}C_{\alpha\beta}\,{}^{t+\Delta t}\dot{\vartheta}_\beta + {}^{t}K_{\alpha\beta}\,{}^{t+\Delta t}\vartheta_\beta = {}^{t+\Delta t}Q_\alpha\,. \tag{72}$$

The right-hand side vector is assumed here as given at $t + \Delta t$ — if either Q or $\hat{q}$ depend in a given way on the temperature ϑ_α, some resulting contribution will appear on the left-hand side of Eq. (72).

Using a time integration algorithm, say the one-step backward Euler scheme of the form

$$^{t+\Delta t}\dot{\vartheta}_\alpha = \frac{^{t+\Delta t}\vartheta_\alpha - {}^{t}\vartheta_\alpha}{\Delta t} \tag{73}$$

we may solve Eq. (72) for $^{t+\Delta t}\vartheta$ in terms of all the quantities defined at time t and thus assumed known using the equation

$$\left(\frac{1}{\Delta t}\,{}^{t}C_{\alpha\beta} + {}^{t}K_{\alpha\beta}\right){}^{t+\Delta t}\vartheta_\beta = {}^{t+\Delta t}Q_\alpha + \frac{1}{\Delta t}\,{}^{t}C_{\alpha\beta}\,{}^{t}\vartheta_\beta \tag{74}$$

or, briefly,

$$^{t}K^{(eff)}_{\alpha\beta}\,{}^{t+\Delta t}\vartheta_\beta = {}^{t+\Delta t}Q_\alpha + \frac{1}{\Delta t}\,{}^{t}C_{\alpha\beta}\,{}^{t}\vartheta_\beta\,. \tag{75}$$

The effective 'stiffness' matrix $^{t}K^{(eff)}_{\alpha\beta}$ is symmetric. The linearization errors can be removed by using an iteration algorithm restoring the nonlinear heat flow equilibrium. A commonly employed iteration scheme is based on using the constant matrix ${}^{t}K^{(eff)}_{\alpha\beta}$ to solve the following equation for the k-th correction to the temperature at $t+\Delta t$

$$ {}^{t}K^{(eff)}_{\alpha\beta}\,\delta\vartheta^{(k)}_{\beta} = {}^{t+\Delta t}R^{(k-1)}_{\alpha}\,, \qquad k = 2,3,\ldots \tag{76} $$

where the $(k-1)$-th residual reads

$$ {}^{t+\Delta t}R^{(k-1)}_{\alpha} = {}^{t+\Delta t}Q_{\alpha} - \left[{}^{t+\Delta t}C^{(k-1)}_{\alpha\beta}\,{}^{t+\Delta t}\dot{\vartheta}^{(k-1)}_{\beta} + {}^{t+\Delta t}K^{(k-1)}_{\alpha\beta}\,{}^{t+\Delta t}\vartheta^{(k-1)}_{\beta}\right] \tag{77} $$

while

$$ {}^{t+\Delta t}\dot{\vartheta}^{(k-1)}_{\alpha} = \frac{{}^{t+\Delta t}\vartheta^{(k-1)}_{\alpha} - {}^{t}\vartheta_{\alpha}}{\Delta t}\,, \tag{78} $$

$$ {}^{t+\Delta t}C^{(k-1)}_{\alpha\beta} = C_{\alpha\beta}\left({}^{t+\Delta t}\vartheta^{(k-1)}_{\gamma}\right), \quad \text{etc.,} \tag{79} $$

and

$$ {}^{t+\Delta t}\vartheta^{(k)}_{\alpha} = {}^{t+\Delta t}\vartheta^{(k-1)}_{\alpha} + \delta\vartheta^{(k)}_{\alpha}\,. \tag{80} $$

The value of ${}^{t+\Delta t}\vartheta^{(1)}_{\alpha}$ is the one obtained from Eq. (75). The operator form of Eq. (76) reads

$$ \left({}^{t}C_{\alpha\beta}\frac{\partial}{\partial\tau} + {}^{t}K_{\alpha\beta}\right)\delta\vartheta^{(k)}_{\beta} = {}^{t+\Delta t}R^{(k-1)}_{\alpha}\,, \qquad k = 2,3,\ldots \tag{81} $$

because

$$ \frac{\partial}{\partial\tau}\delta\vartheta^{(k)}_{\beta} = \delta\dot{\vartheta}^{(k)}_{\beta} = \delta\,\frac{{}^{t+\Delta t}\vartheta^{(k)}_{\beta} - {}^{t}\vartheta_{\beta}}{\Delta t} = \frac{1}{\Delta t}\delta\,{}^{t+\Delta t}\vartheta^{(k)}_{\beta} \tag{82} $$

so that

$$ {}^{t}C_{\alpha\beta}\frac{\partial}{\partial\tau}\delta\vartheta^{(k)}_{\beta} = \frac{1}{\Delta t}\,{}^{t}C_{\alpha\beta}\delta\vartheta^{(k)}_{\beta}\,. \tag{83} $$

A faster convergence can theoretically be achieved by the full Newton-Raphson algorithm, cf. Eq. (14). Let as rewrite Eq. (67) in the residual form as, cf. Eq. (77),

$$ R_{\alpha} = Q_{\alpha} - \left(C_{\alpha\beta}\frac{\partial}{\partial\tau} + K_{\alpha\beta}\right)\vartheta_{\beta} \tag{84} $$

and use it to compute the tangent operator R_T

$$ R_{T\alpha\beta}\delta\vartheta_{\beta} = \frac{\partial R_{\alpha}}{\partial\vartheta_{\beta}}\delta\vartheta_{\beta}\,, \tag{85} $$

as

$$ R_{T\alpha\beta}\delta\vartheta_{\beta} = -\left(C_{\alpha\beta}\frac{\partial}{\partial\tau} + K_{T\alpha\beta}\right)\delta\vartheta_{\beta} \tag{86} $$

in which

$$ K_{T\alpha\beta} = \overbrace{C_{\alpha\gamma,\beta}\dot{\vartheta}_{\gamma}}^{K_T^{(2)}} + \overbrace{K_{\alpha\beta} + K_{\alpha\gamma,\beta}\vartheta_{\gamma} - Q_{\alpha,\beta}}^{K_T^{(1)}} \tag{87} $$

(if Q_{α} does not depend on ϑ_{α}, as assumed previously, then clearly $Q_{\alpha,\beta} = 0$). We note that for some minor technical reasons the discretized and continuous residuals in our formulation differ in sign. Consequently, the adjoint variables $T^{(a)}$ and $\vartheta^{(a)}$ in both the formulations will have opposite signs as well. We also note that even though the matrices $C_{\alpha\beta}$ and $K_{\alpha\beta}$ are symmetric, the matrix

$K_{T\alpha\beta}$ is not. This is a serious drawback when it comes to iterative corrections of the solution at a given time step using the full Newton-Raphson scheme. We have in such a case

$$ {}^{t+\Delta t}R_{\alpha}^{(k)} = {}^{t+\Delta t}R_{\alpha}^{(k-1)} + {}^{t+\Delta t}R_{T_{\alpha\beta}}^{(k-1)} \delta\vartheta_{\beta}^{(k)} = 0 \,, \tag{88} $$

i.e.

$$ \underbrace{\left({}^{t+\Delta t}C_{\alpha\beta}^{(k-1)} \frac{\partial}{\partial\tau} + {}^{t+\Delta t}K_{T_{\alpha\beta}}^{(k-1)} \right)}_{-{}^{t+\Delta t}R_{T_{\alpha\beta}}^{(k-1)}} \delta\vartheta_{\beta}^{(k)} = {}^{t+\Delta t}R_{\alpha}^{(k-1)} \,, \tag{89} $$

where

$$ {}^{t+\Delta t}R_{\alpha}^{(k-1)} = {}^{t+\Delta t}Q_{\alpha} - \left({}^{t+\Delta t}C_{\alpha\beta}^{(k-1)} \frac{\partial}{\partial\tau} + {}^{t+\Delta t}K_{\alpha\beta}^{(k-1)} \right) {}^{t+\Delta t}\vartheta_{\beta}^{(k-1)} \,. \tag{90} $$

Equation (89) should be compared against Eq. (81). The operator $R_{T_{\alpha\beta}}$ in Eq. (89) guarantees the faster (quadratic) convergence of the iteration than the operator in eq. (81). However, the latter is kept constant for every iteration and it generates the symmetric iteration matrix — the factors which may compensate for a slower convergence rate.

Furthermore, in view of our fundamental observation made with respect to the DDM sensitivity equation (35), in which precisely the same tangent operator has to be used (in a non-iterative way, though), the non-symmetry of the tangent matrix presents a similar computational disadvantage. By differentiating Eq. (72) with respect to the design parameter h we obtain

$$ \left({}^{t+\Delta t}C_{\alpha\beta} \frac{\partial}{\partial\tau} + {}^{t+\Delta t}K_{T_{\alpha\beta}} \right) \frac{\mathrm{d}\, {}^{t+\Delta t}\vartheta_{\beta}}{\mathrm{d}h} = \frac{\partial\, {}^{t+\Delta t}Q_{\alpha}}{\partial h} - \frac{\partial\, {}^{t+\Delta t}C_{\alpha\beta}}{\partial h} {}^{t+\Delta t}\dot{\vartheta}_{\beta} + \frac{\partial\, {}^{t+\Delta t}K_{\alpha\beta}}{\partial h} {}^{t+\Delta t}\vartheta_{\beta} \tag{91} $$

which may be presented as, cf. Eqs. (84), (86),

$$ {}^{t+\Delta t}R_{T_{\alpha\beta}} \frac{\mathrm{d}\, {}^{t+\Delta t}\vartheta_{\beta}}{\mathrm{d}h} = \frac{\partial\, {}^{t+\Delta t}R_{\alpha}}{\partial h} \,. \tag{92} $$

Equation (92) is linear in $\mathrm{d}\, {}^{t+\Delta t}\vartheta_{\beta}/\mathrm{d}h$ with its right-hand side known provided the primary problem has been solved in the time interval considered.

Equation (92) can be solved either directly (i.e. using the non-symmetric equation solver), in which case no iterations are required, or it can be transformed to the form

$$ \left({}^{t+\Delta t}C_{\alpha\beta} \frac{\partial}{\partial\tau} + {}^{t+\Delta t}K_{\alpha\beta} \right) \frac{\mathrm{d}\, {}^{t+\Delta t}\vartheta_{\beta}}{\mathrm{d}h} = \frac{\partial\, {}^{t+\Delta t}R_{\alpha}}{\partial h} - {}^{t+\Delta t}K_{\alpha\beta}^{*} \frac{\mathrm{d}\, {}^{t+\Delta t}\vartheta_{\beta}}{\mathrm{d}h} \tag{93} $$

where

$$ {}^{t+\Delta t}K_{\alpha\beta}^{*} = {}^{t+\Delta t}K_{T_{\alpha\beta}} - {}^{t+\Delta t}K_{\alpha\beta} = {}^{t+\Delta t}C_{\alpha\gamma,\beta}\dot{\vartheta}_{\gamma} + {}^{t+\Delta t}K_{\alpha\gamma,\beta}\vartheta_{\gamma} - {}^{t+\Delta t}Q_{\alpha,\beta} \,. \tag{94} $$

Equation (93) can be solved for the nodal temperature sensitivity vector $\mathrm{d}\, {}^{t+\Delta t}\vartheta_{\beta}/\mathrm{d}h$ at the desired time instant by using the direct iteration according to

$$ \left({}^{t+\Delta t}C_{\alpha\beta} \frac{\partial}{\partial\tau} + {}^{t+\Delta t}K_{\alpha\beta} \right) \frac{\mathrm{d}\, {}^{t+\Delta t}\vartheta_{\beta}^{(k)}}{\mathrm{d}h} = \frac{\partial\, {}^{t+\Delta t}R_{\alpha}}{\partial h} - {}^{t+\Delta t}K_{\alpha\beta}^{*} \frac{\mathrm{d}\, {}^{t+\Delta t}\vartheta_{\beta}^{(k-1)}}{\mathrm{d}h} \,, \quad k = 1, 2, \ldots, \tag{95} $$

with $\mathrm{d}\, {}^{t+\Delta t}\vartheta_{\beta}^{(0)}/\mathrm{d}h$ assumed zero, or the last value available (i.e. $\mathrm{d}\, {}^{t}\vartheta_{\alpha}/\mathrm{d}h$). Which of the above approaches should be used for a specific problem depends to a large extent on the solution algorithm employed for solving the primary problem.

Using the backward Euler time integration scheme for the sensitivity equation (95),

$$\frac{\mathrm{d}\,{}^{t+\Delta t}\dot{\vartheta}_\alpha}{\mathrm{d}h} = \frac{\frac{\mathrm{d}\,{}^{t+\Delta t}\vartheta_\alpha}{\mathrm{d}h} - \frac{\mathrm{d}\,{}^{t}\vartheta_\alpha}{\mathrm{d}h}}{\Delta t}, \tag{96}$$

Eq. (95) becomes

$${}^{t+\Delta t}K^{(eff)}_{\alpha\beta}\frac{\mathrm{d}\,{}^{t+\Delta t}\vartheta^{(k)}_\beta}{\mathrm{d}h} = \frac{\partial\,{}^{t+\Delta t}R_\alpha}{\partial h} - {}^{t+\Delta t}K_{\alpha\beta} * \frac{\mathrm{d}\,{}^{t+\Delta t}\vartheta^{(k-1)}_\beta}{\mathrm{d}h} + \frac{1}{\Delta t}\,{}^{t+\Delta t}C_{\alpha\beta}\frac{\mathrm{d}\,{}^{t}\vartheta_\beta}{\mathrm{d}h}\,. \tag{97}$$

Equation (97) can be iteratively solved for $\mathrm{d}\,{}^{t+\Delta t}\vartheta_\alpha/\mathrm{d}h$. It requires the storage of the previous (computed at $\tau = t$) sensitivity vector $\mathrm{d}\,{}^{t}\vartheta_\alpha/\mathrm{d}h$ but no such vectors at any earlier time instants.

For problems with temperature sensitivity slowly changing in time the following non-iterative approximation to Eq. (97) may turn out reasonable

$${}^{t+\Delta t}K^{(eff)}_{\alpha\beta}\frac{\mathrm{d}\,{}^{t+\Delta t}\vartheta^{(k)}_\beta}{\mathrm{d}h} = \frac{\partial\,{}^{t+\Delta t}R_\alpha}{\partial h} - \left({}^{t+\Delta t}K^*_{\alpha\beta} - \frac{1}{\Delta t}\,{}^{t+\Delta t}C_{\alpha\beta}\right)\frac{\mathrm{d}\,{}^{t}\vartheta_\beta}{\mathrm{d}h} \tag{98}$$

which combines the advantages of the symmetric governing matrix and lack of iteration at the cost of a decreased accuracy.

Having an effective technique to compute the temperature sensitivities at every time instant makes it possible to determine sensitivity of any functional such as, quite generally,

$$\mathcal{G}(h) = \int_{t_0}^{t_f} G(\vartheta_\alpha(\tau,h),h)\,\mathrm{d}\tau + g(\vartheta(t_f,h),h)\,. \tag{99}$$

The functional can be specified to describe the so-called *critical time sensitivity* (in which the problem is to find the design gradient of a functional at a pre-set time instant) or the so-called time interval sensitivity (in which the sensitivity gradient is obtained as the time integral over a selected time interval). In fact, both these situations could be described just by the first term in Eq. (99); to be more explicit we have directly included there the term g which is local in time.

Differentiation of Eq. (99) with respect to the parameter h yields

$$\frac{\mathrm{d}\mathcal{G}}{\mathrm{d}h} = \int_{t_0}^{t_f}\left(\frac{\partial G}{\partial\vartheta_\alpha}\frac{\mathrm{d}\vartheta_\alpha}{\mathrm{d}h} + \frac{\partial G}{\partial h}\right)\mathrm{d}\tau + \left(\frac{\partial g}{\partial\vartheta_\alpha}\frac{\mathrm{d}\vartheta_\alpha}{\mathrm{d}h}\right)\bigg|_{t_f} + \frac{\partial g}{\partial h}\bigg|_{t_f} \tag{100}$$

which clearly indicates the algorithm for computing $\mathrm{d}\mathcal{G}/\mathrm{d}h$ in terms of $\mathrm{d}\vartheta_\alpha/\mathrm{d}h$ determined at all the discrete time instants $\tau \in [t_0, t_f]$.

Before moving on to discuss the adjoint system method we shall make another general comment. In discretized nonlinear heat transfer problems the 'secant' formulation is given in terms of the symmetric matrices $C_{\alpha\beta}$ and $K_{\alpha\beta}$, of Eq. (67); the tangent formulation involves $K_{T_{\alpha\beta}}$ which is not symmetric, cf. Eq. (87). In typical nonlinear solid mechanics problems the secant formulation is governed by some non-symmetric matrices while the tangent formulation involves symmetric matrices (problems with configuration dependent loadings, boundary friction effects or non-associative plastic flow rules are notable exceptions in this regard). Since the DDM approach is inherently based on the tangent matrices, its use in nonlinear solid mechanics appears more straightforward than in heat transfer problems.

Let us now consider the alternative way of finding $\mathrm{d}\mathcal{G}/\mathrm{d}h$ by using the adjoint system method (ASM). To this aim we employ again the variational technique of [1] which states that

$$\frac{\mathrm{d}\mathcal{G}}{\mathrm{d}h} = \frac{\partial\hat{\mathcal{G}}}{\partial h} \tag{101}$$

provided

$$\hat{\mathcal{G}} = \mathcal{G} + \lambda_\alpha R_\alpha \tag{102}$$

and λ_α, $\alpha = 1,2,\ldots,N$ are such as to make $\hat{\mathcal{G}}$ stationary with respect to ϑ_α, i.e.

$$\frac{\partial\hat{\mathcal{G}}}{\partial\vartheta_\alpha}\delta\vartheta_\alpha = 0\,. \tag{103}$$

Let us first form the extended functional $\hat{\mathcal{G}}$ as

$$\hat{\mathcal{G}}(h) = \mathcal{G}(h) + \int_{t_0}^{t_f} \lambda_\alpha(\tau)R_\alpha(\tau)\,\mathrm{d}\tau\,, \qquad \alpha = 1,2,\ldots,N\,, \tag{104}$$

in which the residuals R_α are given by Eq. (84). Considering in Eq. (103) only the design variations

$$\bar{\delta\vartheta}_\alpha = \frac{\mathrm{d}\vartheta_\alpha}{\mathrm{d}h}\delta h \tag{105}$$

we compute

$$\begin{aligned}\frac{\partial\hat{\mathcal{G}}}{\partial\vartheta_\gamma}\bar{\delta\vartheta}_\gamma &= \left.\left(\frac{\partial g}{\partial\vartheta_\gamma}\bar{\delta\vartheta}_\gamma\right)\right|_{t_f} + \int_{t_0}^{t_f}\left[\frac{\partial G}{\partial\vartheta_\gamma} + \lambda_\alpha\frac{\partial}{\partial\vartheta_\gamma}\left(Q_\alpha - C_{\alpha\beta}\dot{\vartheta}_\beta - K_{\alpha\beta}\vartheta_\beta\right)\right]\bar{\delta\vartheta}_\gamma\,\mathrm{d}\tau \\ &= \left.\left(\frac{\partial g}{\partial\vartheta_\gamma}\bar{\delta\vartheta}_\gamma\right)\right|_{t_f} + \int_{t_0}^{t_f}\left[\frac{\partial G}{\partial\vartheta_\gamma} + \lambda_\alpha\left(Q_{\alpha,\gamma} - C_{\alpha\beta,\gamma}\dot{\vartheta}_\beta - C_{\alpha\beta}\dot{\vartheta}_{\beta,\gamma}\right.\right. \\ &\qquad \left.\left. - K_{\alpha\beta,\gamma}\vartheta_\beta - K_{\alpha\beta}\delta_{\beta\gamma}\right)\right]\bar{\delta\vartheta}_\gamma\,\mathrm{d}\tau\end{aligned} \tag{106}$$

which by integrating by parts in time the third term in parentheses and observing Eq. (87) becomes

$$\frac{\partial\hat{\mathcal{G}}}{\partial\vartheta_\gamma}\bar{\delta\vartheta}_\gamma = \left.\left(\frac{\partial g}{\partial\vartheta_\gamma}\bar{\delta\vartheta}_\gamma\right)\right|_{t_f} - \left.\left(\lambda_\alpha C_{\alpha_\beta}\bar{\delta\vartheta}_\beta\right)\right|_{t_0}^{t_f} + \int_{t_0}^{t_f}\left[\frac{\partial G}{\partial\vartheta_\gamma} + \lambda_\alpha\left(-K_{T\alpha\gamma} + \dot{C}_{\alpha\gamma}\right) + \dot{\lambda}_\alpha C_{\alpha\gamma}\right]\bar{\delta\vartheta}_\gamma\,\mathrm{d}\tau \tag{107}$$

or (note that $\bar{\delta\vartheta}_\alpha(t_0) = 0$)

$$\frac{\partial\hat{\mathcal{G}}}{\partial\vartheta_\gamma}\bar{\delta\vartheta}_\gamma = \left.\left(\frac{\partial g}{\partial\vartheta_\gamma} - \lambda_\alpha C_{\alpha\gamma}\right)\right|_{t_f}\bar{\delta\vartheta}_\gamma(t_f) + \int_{t_0}^{t_f}\left[\frac{\partial G}{\partial\vartheta_\gamma} - \lambda_\alpha\left(K_{T\alpha\gamma} - \dot{C}_{\alpha\gamma}\right) + \dot{\lambda}_\alpha C_{\alpha\gamma}\right]\bar{\delta\vartheta}_\gamma\,\mathrm{d}\tau = 0\,. \tag{108}$$

The adjoint equation is thus implied to be

$$C_{\alpha\beta}\dot{\lambda}_\beta - \left(K_{T_{\beta\alpha}} + \dot{C}_{\alpha\beta}\right)\lambda_\beta = -\frac{\partial G}{\partial\vartheta_\alpha}\,, \qquad \tau\in[t_0,t_f]\,, \tag{109}$$

with the terminal condition at $\tau = t_f$ resulting as the natural boundary condition in the form

$$C_{\alpha\beta}\lambda_\beta = \frac{\partial g}{\partial\vartheta_\alpha}\,, \qquad \tau = t_f\,. \tag{110}$$

It is emphasized that Eqs. (109), (110) form a terminal-value problem which is to be integrated backward in time once the primary problem has been solved forward in time. The adjoint problem can be formally replaced by an initial-value problem if only we employ the time transformation

$$\tau' = t_f - \tau \tag{111}$$

in which case Eqs. (109) and (110) become

$$C_{\alpha\beta}\dot{\lambda}_\beta + \left(K^{(1)}_{T_{\beta\alpha}} - K^{(2)}_{T_{\beta\alpha}} + \dot{C}_{\alpha\beta}\right)\lambda_\beta = \frac{\partial G}{\partial\vartheta_\alpha}\,, \qquad \tau'\in[0,t_f - t_0]\,, \tag{112}$$

$$C_{\alpha\beta}\lambda_\beta = \frac{\partial g}{\partial\vartheta_\alpha}\,, \qquad \tau' = 0\,, \tag{113}$$

where the dot stands now for differentiation with respect to the variable τ'.

The basic ASM sensitivity relationship, Eq. (100), can be rewritten as

$$\frac{\mathrm{d}\mathcal{G}}{\mathrm{d}h} = \frac{\partial g}{\partial h} + \int_{t_0}^{t_f} \left[\frac{\partial G}{\partial h} + \lambda_\alpha \left(\frac{\partial Q_\alpha}{\partial h} - \frac{\partial K_{\alpha\beta}}{\partial h}\vartheta_\beta - \frac{\partial C_{\alpha\beta}}{\partial h}\dot{\vartheta}_\beta \right) \right] \mathrm{d}\tau \tag{114}$$

which defines the effective way to determine $\mathrm{d}\mathcal{G}/\mathrm{d}h$ once λ_α is known at every discrete $\tau \in [t_0, t_f]$.

Furthermore, we observe that for $\tilde{c} = \text{const}$ which implies $\dot{C}_{\alpha\beta} = 0$ and $K^{(2)}_{T_{\alpha\beta}} = 0$ we obtain from Eq. (112)

$$C_{\alpha\beta}\dot{\lambda}_\beta + K^{(1)}_{T_{\beta\alpha}}\lambda_\beta = \frac{\partial G}{\partial \vartheta_\alpha} \tag{115}$$

which clearly is the discretized counterpart of the Eq. (55) derived in Section 2. This fact has tremendous computational significance.

Writing Eq. (112) for the time $\tau = t + \Delta t$ we arrive at

$${}^{t+\Delta t}C_{\alpha\beta}\,{}^{t+\Delta t}\dot{\lambda}_\beta + \left({}^{t+\Delta t}K^{(1)}_{T_{\beta\alpha}} - {}^{t+\Delta t}K^{(2)}_{T_{\beta\alpha}} + {}^{t+\Delta t}\dot{C}_{\alpha\beta} \right) {}^{t+\Delta t}\lambda_\beta = -\frac{\partial\, {}^{t+\Delta t}G}{\partial \vartheta_\alpha}\,. \tag{116}$$

The right-hand side vector is assumed given at time $t + \Delta t$. Using the time integration scheme in the form of the backward Euler algorithm

$${}^{t+\Delta t}\dot{\lambda}_\beta = \frac{{}^{t+\Delta t}\lambda_\beta - {}^{t}\lambda_\beta}{\Delta t} \tag{117}$$

we can solve Eq. (116) for ${}^{t+\Delta t}\lambda$ in terms of all the quantifies defined at time t

$$\left(\frac{1}{\Delta t}{}^{t+\Delta t}C_{\alpha\beta} + {}^{t+\Delta t}K^{(1)}_{T_{\beta\alpha}} - {}^{t+\Delta t}K^{(2)}_{T_{\beta\alpha}} + {}^{t+\Delta t}\dot{C}_{\alpha\beta} \right) {}^{t+\Delta t}\lambda_\beta = \frac{\partial\, {}^{t+\Delta t}G}{\partial \vartheta_\alpha} + \frac{1}{\Delta t}{}^{t+\Delta t}C_{\alpha\beta}\,{}^{t}\lambda_\beta\,. \tag{118}$$

It requires storage of the previous adjoint temperature vector ${}^{t}\lambda_\beta$.

Equation (118) can be written in a short form as

$${}^{t+\Delta t}K^{(a)(\mathit{eff})}_{\beta\alpha}\,{}^{t+\Delta t}\lambda_\beta = \frac{\partial\, {}^{t+\Delta t}G}{\partial \vartheta_\alpha} + \frac{1}{\Delta t}{}^{t+\Delta t}C_{\alpha\beta}\,{}^{t}\lambda_\beta \tag{119}$$

where ${}^{t+\Delta t}K^{(a)(\mathit{eff})}_{\alpha\beta}$ is the effective stiffness matrix for the adjoint problem. The matrix ${}^{t+\Delta t}K^{(a)(\mathit{eff})}_{\alpha\beta}$ is not symmetric which implies computational disadvantages.

Equation (119) can be solved either directly by using the non-symmetric equation solver or it can be transformed to the form (note the symmetry of $K_{\alpha\beta}$)

$$\left[\frac{1}{\Delta t}{}^{t+\Delta t}C_{\alpha\beta} + \left({}^{t+\Delta t}K_{\alpha\beta} + {}^{t+\Delta t}\dot{C}_{\alpha\beta} \right) \right] {}^{t+\Delta t}\lambda_\beta$$
$$= \frac{\partial\, {}^{t+\Delta t}G}{\partial \vartheta_\alpha} + \frac{1}{\Delta t}{}^{t+\Delta t}C_{\alpha\beta}\,{}^{t}\lambda_\beta + \left({}^{t+\Delta t}K^{(2)}_{T_{\beta\alpha}} - {}^{t+\Delta t}K^{(1)*}_{T_{\beta\alpha}} \right) {}^{t+\Delta t}\lambda_\beta \tag{120}$$

where

$$K^{(1)*}_{T_{\beta\alpha}} = K^{(1)}_{T_{\beta\alpha}} - K_{\alpha\beta}\,. \tag{121}$$

Equation (120) can be solved for the nodal adjoint temperature vector ${}^{t+\Delta t}\lambda$ at the desired time instant by using the direct iteration according to

$$\left[\frac{1}{\Delta t}{}^{t+\Delta t}C_{\alpha\beta} + \left({}^{t+\Delta t}K_{\alpha\beta} + {}^{t+\Delta t}\dot{C}_{\alpha\beta} \right) \right] {}^{t+\Delta t}\lambda^{(k)}_\beta$$
$$= \frac{\partial\, {}^{t+\Delta t}G}{\partial \vartheta_\alpha} + \frac{1}{\Delta t}{}^{t+\Delta t}C_{\alpha\beta}\,{}^{t}\lambda_\beta + \left({}^{t+\Delta t}K^{(2)}_{T_{\beta\alpha}} - {}^{t+\Delta t}K^{(1)*}_{T_{\beta\alpha}} \right) {}^{t+\Delta t}\lambda^{(k-1)}_\beta \tag{122}$$

with $\lambda_\beta^{(0)}$ assumed zero, or the last value available (i.e. ${}^t\lambda_\beta$). Denoting

$$ {}^{t+\Delta t}K_{\alpha\beta}^{(s)(eff)} = \frac{1}{\Delta t}\,{}^{t+\Delta t}C_{\alpha\beta} + \left({}^{t+\Delta t}K_{\alpha\beta} + {}^{t+\Delta t}\dot{C}_{\alpha\beta}\right) \tag{123} $$

and

$$ {}^{t+\Delta t}K_{T\alpha\beta}^{(3)} = {}^{t+\Delta t}K_{T\alpha\beta}^{(2)} - {}^{t+\Delta t}K_{T\alpha\beta}^{(1)*} \tag{124} $$

Eq. (122) becomes

$$ {}^{t+\Delta t}K_{\alpha\beta}^{(s)(eff)}\,{}^{t+\Delta t}\lambda_\beta^{(k)} = \frac{\partial\,{}^{t+\Delta t}G}{\partial\vartheta_\alpha} + \frac{1}{\Delta t}\,{}^{t+\Delta t}C_{\alpha\beta}\,{}^t\lambda_\beta + {}^{t+\Delta t}K_{T\alpha\beta}^{(3)}\,{}^{t+\Delta t}\lambda_\beta^{(k-1)}\,. \tag{125} $$

Equation (125) can be iteratively solved for ${}^{t+\Delta t}\lambda_\beta^{(k)}$. It requires storage of the previous adjoint temperature vector ${}^t\lambda_\beta$.

For problems with temperature sensitivity changing slowly in time the following non-iterative approximation to Eq. (125) may sometimes be used

$$ {}^{t+\Delta t}K_{\alpha\beta}^{(s)(eff)}\,{}^{t+\Delta t}\lambda_\beta = \frac{\partial\,{}^{t+\Delta t}G}{\partial\vartheta_\alpha} + \left(\frac{1}{\Delta t}\,{}^{t+\Delta t}C_{\alpha\beta} + {}^{t+\Delta t}K_{T\alpha\beta}^{(3)}\right){}^t\lambda_\beta\,. \tag{126} $$

We note in closing that contrary to typical nonlinear solid mechanics problems the tangent matrix $\mathbf{K}_T$ in nonlinear heat transfer problem is non-symmetric, and it is $\mathbf{K}_T$ transposed which appears in the adjoint formulation. Nevertheless, once the primary response is known, the adjoint response can be obtained numerically in an efficient manner. This is particularly so when sufficient storage is available and the decomposed stiffness matrix from the primary forward-in-time analysis at each discrete time instant τ can be stored and later utilized to determine the adjoint response at time $t_f - \tau$ during the backward-in-time non-iterative analysis. Such an algorithm requires neither additional stiffness matrix assemblies nor decompositions for the adjoint problem; only right-hand side assemblies and back substitutions are needed to obtain the adjoint solution, [16]. Consequently, the computational cost for the adjoint analysis is much smaller than that for the primary analysis which may generally require several iterations to converge at each step.

4. Example — angular speed as the design parameter in friction welding

The direct differentiation method has been applied to the sensitivity analysis in which the angular speed is considered as the design parameter in the problem of friction welding. The finite element modelling of the friction welding process was undertaken by the authors in [10]. Friction welding is a process in which the heat for welding is produced by direct conversion of mechanical energy to thermal energy at the interface of the workpieces without the application of electrical energy, or heat from other sources, to the workpieces. Friction welds are made by holding a non-rotating workpiece in contact with a rotating workpiece under constant or gradually increasing pressure until the interface reaches welding temperature, and then stopping rotation to complete the weld. The frictional heat developed at the interface rapidly raises the temperature of the workpieces over a very short axial distance to a value approaching from below the melting point and welding occurs under the influence of a pressure that is applied while the heated zone is in the plastic range. Friction welding is classified as a solid-state welding process in which joining occurs at a temperature below the melting point of the work metal. In the example we will analyse the temperature sensitivity to the variation of the angular speed of the workpieces. This problem has a clear practical significance in the welding optimization. Two rods shown schematically in Fig. 1 illustrate the workpieces in friction welding. The problem can be considered as axisymmetric. The simple triangular finite

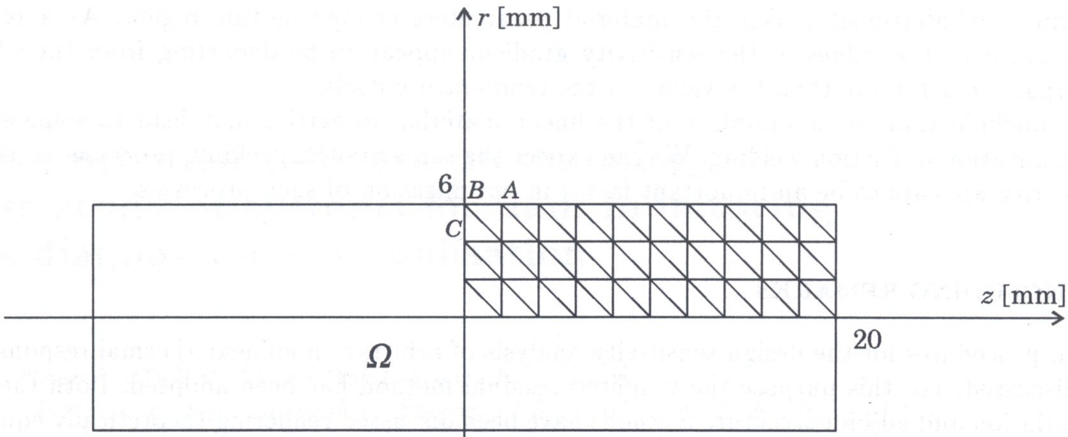

Fig. 1. Geometry and finite element mesh of the workpieces

Table 1. Temperature sensitivity to variation of the angular speed in friction welding

Time	Point	Temperature T		Sensitivity $\mathrm{d}T/\mathrm{d}\omega$	
		linear analysis	nonlinear analysis	linear analysis	nonlinear analysis
1	A	256	254	2.328	2.291
	B	534	531	5.742	5.352
	C	371	369	4.217	3.927
2	A	424	421	3.259	3.127
	B	815	811	8.038	7.814
	C	658	654	5.903	5.643
3	A	582	578	3.957	3.804
	B	1034	1029	9.763	8.901
	C	893	889	7.169	6.825

elements with linear shape functions in the radial cross-section are used. At the place of abutment the heat source is given by the following equation

$$Q = \int_{\Omega} \sigma \mu \omega r \, \mathrm{d}\Omega \tag{127}$$

where σ is the normal stress at the place of contact, μ is the coefficient of friction, ω is the angular speed, r is the radius and Ω is the surface area upon which the heat acts. Steel rods of the diameter $\phi = 12\,\mathrm{mm}$ are considered. In our analysis $\sigma\mu$ is assumed to be equal to $3 \times 10^7\,\mathrm{J/m^3}$ and ω to $100\,\mathrm{sec}^{-1}$. The material properties are assumed as follows: thermal conductivity $k = 50(1{+}0.001T)\,\mathrm{W/mK}$, specific heat $c = 510(1{+}0.02T)\,\mathrm{Jkg/K}$, thermal convection coefficient $0.25\,\mathrm{W/m^2}$, initial temperature $0°C$, surrounding temperature $0°C$. The derivatives $\frac{\partial k}{\partial T}$, $\frac{\partial c}{\partial T}$, $\frac{\partial Q}{\partial \omega}$ are constant throughout the whole history of the process. The results of the sensitivity analysis $\frac{\mathrm{d}T}{\mathrm{d}\omega}$ at the chosen points A, B, C at subsequent time instants $t_1 = 1\,\mathrm{sec}$, $t_2 = 2\,\mathrm{sec}$, $t_3 = 3\,\mathrm{sec}$ are presented in Table 1. In order to compare the results with the sensitivity analysis for the corresponding linear problem the latter results are also displayed in Table 1 to illustrate the influence of nonlinearity on the sensitivity process.

As it is seen, in the modelling of friction welding the assumption of nonlinearity (i.e. temperature dependence of the material parameters) may play some role. The observed temperature rise in the vicinity of abutment makes the material parameters change in this region. As a result of the nonlinearity the values of the sensitivity gradient appear to be departing from their 'linear' counterparts much more than the values of the temperature itself.

We conclude that the assumption of the linear material properties may lead to some error in the optimization of friction welding. We can expect the same in other welding processes so that the nonlinearity appears to be an important factor in optimization of such processes.

5. Concluding remarks

General procedures for the design sensitivity analysis of arbitrary nonlinear thermal response have been discussed. For this purpose the weighted residual method has been adopted. Both the direct differentiation and adjoint structure methods have been discussed rendering theoretically equivalent formulations. Any useful sensitivity analysis method must be effective for large-scale problems — the finite element approach has therefore been employed next. The resulting finite element equations for the sensitivity analysis of nonlinear thermal transient systems have been derived. In finite element formulation partial derivatives of the conductivity and heat capacity matrices as well as of the thermal load vector with respect to the temperature and design variables appear have to be computed. As it is typical of any incremental strategy, these values have to be updated throughout the response time interval.

The finite element formulation lends itself to a straightforward computer implementation.

References

[1] J.S. Arora, J.B. Cardoso. Variational principle for shape design sensitivity analysis. *AIAA J.*, **30**: 538–547, 1992.

[2] K.-J. Bathe. *Finite Element Procedures in Engineering Analysis.* Prentice-Hall, 1982.

[3] K. Dems. Sensitivity analysis in thermal problems — I: Variation of material parameter within a fixed domain. *J. of Thermal Stresses*, **9**: 303–324, 1986.

[4] K. Dems. Sensitivity analysis in thermal problems — II: Structural shape variation. *J. of Thermal Stresses*, **10**: 1–16, 1987.

[5] K. Dems, Z. Mróz. Variational approach to first and second-order sensitivity analysis. *Int. J. Num. Meth. Engng.*, **21**: 637–646, 1985.

[6] K. Dems, Z. Mróz. Variational approach to sensitivity analysis in thermo-elasticity. *J. of Thermal Stresses*, **10**: 283–306, 1987.

[7] R.T. Haftka. Techniques for thermal sensitivity analysis. *Int. J. Num. Meths. Engng.*, **17**: 71–80, 1981.

[8] E.J. Hang, K.K. Choi, V. Komkov. *Design Sensitivity Analysis of Structural Systems.* Series in Math. Sci. Engng., Academic Press, 1986.

[9] M. Kleiber, A. Służalec. Numerical analysis of heat flow in flash welding. *Arch. Mech.*, **35**: 687–699, 1983.

[10] M. Kleiber, A. Służalec. Finite element analysis of heat flow in friction welding. *Engng. Trans.*, **32**: 107–113, 1984.

[11] R.A. Meric. Shape design sensitivity analysis for nonlinear anisotropic heat conducting solids and shape optimization by the BEM. *Int. J. Num. Meths. Engng.*, **26**: 109–120, 1988.

[12] C.W. Park, Y.M. Yoo. Shape design sensitivity analysis of a two-dimensional heat transfer system using the boundary element method. *Comp. Struct.*, **28**: 543–550, 1988.

[13] A. Służalec, M. Kleiber. Shape sensitivity analysis for nonlinear steady-state heat conduction problems. *Int. J. Heat and Mass Transfer* (in print).

[14] D.A. Tortorelli, R.B. Haber. First-order design sensitivity for transient conduction problems by an adjoint method. *Int. J. Num. Meths. Engng.*, **28**: 733–752, 1989.

[15] D.A. Tortorelli, R.B. Haber, S.C.-Y. Lu. Design sensitivity analysis for nonlinear thermal systems. *Comp. Meths. Appl. Mech. Engng.*, **77**: 61–77, 1989.

[16] D.A. Tortorelli, R.B. Haber, S.C.-Y. Lu. Adjoint sensitivity analysis for nonlinear dynamic thermoelastic systems. *AIAA J.*, **29**: 253–263, 1991.