COBEM 99
XV Congresso Brasileiro de Engenharia Mecânica
15th Brazilian Congress of Mechanical Engineering
22 - 26 de Novembro de 1999 / November 22 - 26, 1999 Águas de Lindóia, São Paulo,

FUNCTION ESTIMATION WITH THE CONJUGATE GRADIENT METHOD IN LINEAR AND NONLINEAR HEAT CONDUCTION PROBLEMS

Antônio J. Silva Neto[1]
Francisco D. Moura Neto
Instituto Politécnico (IPRJ), Universidade do Estado do Rio de Janeiro
CP 97282, 28601-970, Nova Friburgo, RJ, Brazil, ajsneto@iprj.uerj.br, fmoura@iprj.uerj.br
Jian Su
Nuclear Engineering Program, COPPE/UFRJ
CP 68509, 21945-970, Rio de Janeiro, RJ, Brazil, sujian@lmn.con.ufrj.br

Abstract: *Various aspects of a methodology to determine unknown functions in inverse heat conduction problems are presented. The methodology is centered on gradient based iterative procedures for optimization problems and we consider specifically Landweber Iteration and Conjugate Gradient Method.*

Key words: *Function estimation, Landweber iteration, Inverse heat conduction problems, Conjugate gradient method.*

1 INTRODUCTION

Inverse heat conduction problems have numerous important applications in various branches of engineering and science, including among others, estimation of unknown boundary heat fluxes (Blanc et al., 1998), thermophysical properties of materials (Artyukhin et al., 1993), and timewise variation of the strength of energy sources located inside a medium (Silva Neto and Özisik, 1994).

These problems are known to be ill-posed, in contrast to the direct heat conduction problems, which are well-posed (that is: the solution exists, it is unique and stable to small changes in the input data.) A variety of numerical and analytical techniques has been proposed for the solution of these inverse heat conduction problems, (Beck et al., 1985, Alifanov et al., 1995).

The inverse problems may be considered in two groups: function estimation and parameter estimation. Function estimation involves an infinite dimensional optimization problem in which we search for the solution in a space of functions with no prior information on the functional form of the quantity to be estimated. Parameter estimation consists of a finite dimensional optimization problem in which a finite number of parameters is estimated.

[1] Also, Nuclear Engineering Program, COPPE/UFRJ, CP 68509, 21945-970, Rio de Janeiro, RJ, Brazil.

The Conjugate Gradient Method is a powerful approach for the solution of function estimation inverse problems because regularization is implicitly built into the algorithm. In the application of this method several steps have to be considered: the direct problem, the sensitivity problem, the adjoint problem, the gradient equation determination, the conjugate gradient method of minimization and the stopping criterion. In this work the formulation of the several steps involved on the application of this procedure for function estimation in linear and nonlinear diffusive processes is presented.

In Section 2 we present a class of direct heat conduction problems (DHCP), and make some general remarks. Inverse heat conduction problems (IHCP) are considered in Section 3 where we establish in which optimal sense one wants to solve it. Section 4 discusses some gradient based iterative algorithms for optimization problems, including Landweber and Conjugate Gradient methods. These methods require the computation of the gradient of energy functionals of a certain type defined in infinite dimensional spaces, and we show how this is done in Section 5. It follows a discussion of so-called sensitivity and adjoint problems. Finally, in Section 6, we present some conclusions and perspectives for future work.

2 LINEAR AND NON-LINEAR HEAT CONDUCTION PROBLEMS

Heat conduction problems study the evolution of the temperature on a material body. Here we will assume an undeformable material body represented by $\Omega \subset I\!R^3$. Let $T = T(\boldsymbol{x}, t)$, $\boldsymbol{x} \in \Omega$, $t \in [0, \infty)$ be the temperature distribution in Ω along the time. Then, T satisfies an initial-boundary value problem for a partial differential equation:

$$\textbf{Evolution equation: } C(T)\frac{\partial T}{\partial t} = \mathsf{div}\,(k(T)\nabla T) + g\ ,\ \boldsymbol{x} \in \Omega,\ t > 0 \quad (1)$$

$$\textbf{Initial temperature distribution: } T(\boldsymbol{x}, 0) = T^*(\boldsymbol{x})\ ,\ \boldsymbol{x} \in \Omega \quad (2)$$

Here $C(T) = \rho c_p(T)$, where ρ is the material density and $c_p(T)$ is the specific heat which may depend on the temperature, $k(T)$ is the heat diffusion coefficient, g is due to a heat source accounting for a rate of heat per unit volume and $T^*(\boldsymbol{x})$ is the initial temperature distribution.

As for the boundary one usually has one of the following types of conditions:

$$T(\boldsymbol{x}, t) = T_e(\boldsymbol{x}, t),\ \ \boldsymbol{x} \in \partial\Omega,\ t > 0 \quad (3)$$

$$-k(T)\boldsymbol{n}(\boldsymbol{x}) \cdot \nabla T(\boldsymbol{x}, t) = f(\boldsymbol{x}, t),\ \ \boldsymbol{x} \in \partial\Omega,\ t > 0 \quad (4)$$

$$-k(T)\boldsymbol{n}(\boldsymbol{x}) \cdot \nabla T(\boldsymbol{x}, t) = h\,(T(\boldsymbol{x}, t) - T_e(\boldsymbol{x}, t))\ \ \boldsymbol{x} \in \partial\Omega,\ t > 0 \quad (5)$$

The first, Eq. (3), is Dirichlet's condition where the temperature at the boundary is prescribed, the second, Eq. (4), is Neumann's condition where the heat flux is prescribed, and the third, Eq.(5), is Robin's condition where a relation between temperature and heat flux is prescribed. Here $\boldsymbol{n}(\boldsymbol{x})$ denotes the exterior unit vector to the boundary of Ω at $\boldsymbol{x} \in \partial\Omega$, T_e is the external temperature, f is the heat flux through the boundary and h is the heat transfer coefficient.

One may, depending on the application, have a mixed type boundary condition. The boundary $\partial\Omega$ can be decomposed in two or more disjoint pieces, say, Γ_1, Γ_2 and Γ_3, and have Dirichlet boundary condition on Γ_1, Neumann on Γ_2 and Robin on Γ_3.

For direct heat conduction problems one assumes that $C = C(T)$, $k = k(T)$, g, the geometry (Ω), and initial condition (T^*) to be known. Depending on the specific type of boundary condition, we also assume that: (a) T_e is known; (b) f is known or (c) h and T_e are known.

Nonlinearity can arise at the level of the equation or the boundary conditions. For one to have a linear problem, $C(T)$ and $k(T)$ should be constant. However, even though we have not allowed that in Eq. (1), for the general case of bodies exhibiting spatial or temporal material variations one may also allow C and k to vary with $\boldsymbol{x}$ and t, *i.e.* $C = C(T, \boldsymbol{x}, t)$, and $k = k(T, \boldsymbol{x}, t)$. In this case, to have a linear problem one should have C and k depending at most on $\boldsymbol{x}$ and t. For a linear problem, non homogeneity of the problem can be given by either g, f, T_e or T^*.

One may exhibit explicitly dependence of the field temperature on the data of the problem. For example if one has Robin boundary condition, we write $T(\boldsymbol{x}, t) = T\left[T^*, g, T_e\right](\boldsymbol{x}, t)$ and say that $T = T\left[T^*, g, T_e\right]$ represents the solution operator.

In linear problems homogeneity becomes a key concept in that the solution operator becomes linear with respect to data. For example, the solution operator of Eq. (1,2,5) is linear if:

- $g = 0$ and $T_e = 0$ with the initial condition (T^*) as the data:

$$T\left[\alpha T^{*,1} + \beta T^{*,2}, 0, 0\right] \;=\; \alpha T\left[T^{*,1}, 0, 0\right] + \beta T\left[T^{*,2}, 0, 0\right], \quad \alpha, \beta \in I\!R$$

- $g = 0$ and $T^* = 0$ with the external temperature (T_e) as the data:

$$T\left[0, 0, \alpha T_e^1 + \beta T_e^2\right] \;=\; \alpha T\left[0, 0, T_e^1\right] + \beta T\left[0, 0, T_e^2\right], \quad \alpha, \beta \in I\!R$$

- $T_e = 0$ and $T^* = 0$ with the source term (g) as the data:

$$T\left[0, \alpha g^1 + \beta g^2, 0\right] \;=\; \alpha T\left[0, g^1, 0\right] + \beta T\left[0, g^2, 0\right], \quad \alpha, \beta \in I\!R$$

For certain linear problems and specific geometries one can write explicit formulas for the solution operator exhibiting the dependence of the field temperature on the data. As one illustrative example, let $\Omega = I\!R^3$, $k = C = 1$, $g = 0$ and zero Dirichlet boundary condition at infinity, then one has:

$$T(\boldsymbol{x}, t) \;=\; \frac{1}{(4\pi t)^{\frac{3}{2}}} \int_{I\!R^3} e^{-\frac{||x-y||^2}{4t}} T^*(\boldsymbol{y})\, d\boldsymbol{y}$$

However, in general, one can only get either qualitative (mathematical) assertions on the solution operator and on the solution, (Friedman, 1964, Folland, 1976), or approximations obtained by some numerical scheme, (Richtmyer and Morton, 1967, Sod, 1985).

3 INVERSE PROBLEMS

To solve for the field temperature T, one has to know some constants or functions, depending on the type of boundary conditions appropriate to the specific problem at hand. An inverse heat conduction problem (IHCP) is characterized by having one or more of these constants/functions unknown. For example, in problem (1,2,5), one may know k,

C, T^* and T_e while the heat source g is unknown. Or else, we know k, C, T_e and g but we do not know the initial temperature distribution, T^*. Other possibilities can also happen.

The objective, then, in an IHCP is to estimate (approximate) unknown constants or functions. In any case, one needs additional information, typically experimental measurements of the temperature field. This can be done in different ways but, for simplicity, we assume here that one has temperature measurements at positions $\boldsymbol{x}_i$, $i = 1, \ldots n$, along the time, and which are represented by $Z_i(t)$.

Say we consider the problem of determining the source g in Eq. (1,2,5), when T^*, and T_e are known. Due to unavoidable experimental errors or slight differences between the mathematical model and physical reality, or also, as is frequently the case, only partial information (incomplete data) is available, one might not have any possible choice of g such that the solution $T = T[g]$ would fit exactly (interpolate) the data,[2] *i.e.*:

$$T[g](\boldsymbol{x}_i, t) \quad = \quad Z_i(t), \quad t > 0,\ i = 1, \ldots n \tag{6}$$

In this case we start by defining a squared residue error function,

$$\begin{aligned} g \mapsto E[g] &= \frac{1}{2}\sum_{i=1}^{n}\int_0^{t_f} \{T[g](\boldsymbol{x}_i, t) - Z_i(t)\}^2 \, dt \\ &= \frac{1}{2}\int_0^{t_f}\int_\Omega \sum_{i=1}^{n} \{T[g](\boldsymbol{x}, t) - Z_i(t)\}^2 \, \delta_{x_i}(\boldsymbol{x}) \, d\boldsymbol{x} \, dt \end{aligned}$$

where t_f is the final time of measurements. We then choose the source term $\hat{g}$ that minimizes E among all possible sources g in an adequate function space, here denoted by V. That is,

$$E[\hat{g}] \quad = \quad \min_{g \in V} E[g] \tag{7}$$

This is an infinite dimensional optimization problem.

4 GRADIENT BASED ITERATIVE PROCEDURES FOR THE MINIMIZATION OF AN ERROR FUNCTION

We shall outline some methods to minimize the error function E based on its gradient. We start by considering the steepest descent method and Landweber iteration, which is a simplication of the former. and then we present the conjugate gradient method.

4.1 Steepest descent and Landweber iteration

Let ϵ be a given tolerance. The steepest descent method is given by the following algorithm.

---

[2]One might want to discard the model if Eq. (6) is not satisfied for any g. This is similar to trying to fit a line $y = ax + b$ to a collection of points $(\boldsymbol{x}_i, \boldsymbol{y}_i)$, $i = 1, \ldots n$ in the plane, which is in general impossible if $n > 2$. However, as one knows, for many practical purposes, it is enough to have a best residue squared approximation. Also, sometimes it is better to have that type of line approximation than to have a higher order interpolating polynomial. A higher order polynomial would increase the mathematical complexity of the model without significantly increasing the quality of it and most probably the interpolating model would be a worse model (wild behaviour) *off* the data.

1^{st} **step:** (Initialization of the method) Choose an initial function[3] $g^o(\boldsymbol{x},t) = g_*(\boldsymbol{x},t) \in L^2(\Omega \times [0,t_f])$. Let $k=0$.

2^{nd} **step:** (Determination of the steepest descent direction) Compute the gradient of E at g^k, a function of $\boldsymbol{x}$ and t, $E'_{g^k} = E'_{g^k}(\boldsymbol{x},t)$, and define the descent direction p^k by

$$p^k \quad = \quad E'_{g^k} \tag{8}$$

3^{rd} **step:** (Determination of the step size in the descent direction) Compute α^k satisfying the critical point equation of the one variable real valued function $I\!\!R \ni \alpha \mapsto E\left[g^k + \alpha p^k\right] \in I\!\!R$, that is, α^k is the solution of

$$\frac{d}{d\alpha} E\left[g^k + \alpha p^k\right] \quad = \quad 0 \tag{9}$$

4^{th} **step:** (Determination of a new approximation) Define

$$g^{k+1} \quad = \quad g^k + \alpha^k p^k \tag{10}$$

5^{th} **step** (Stopping criterion) If

$$\|g^{k+1} - g^k\|_2^2 \;=\; \|\alpha^k p^k\|^2 \quad = \quad \int_0^{t_f}\int_\Omega \left|\alpha^k p^k(\boldsymbol{x},t)\right|^2 \, d\boldsymbol{x}\, dt < \epsilon^2$$

stop. Else, do $k = k+1$ and return to step 2.

We note that, even though, in this general setting, the steepest descent method looks simple, the application of it to IHCP is not. For example:

(i) the determination of E'_{g^k} in Eq. (8) demands the solution of two sets of partial differential equation problems as is shown in Section 5.4.

(ii) finding the solution of the 1-D equation given by Eq. (9) can be time consuming and analytic and numerically complex even considering a Newton type method.

To illustrate some of the difficulties involved in solving Eq. (9) we consider the special case when the problem is linear. In this case,

$$\alpha^k \quad = \quad -\frac{\int_0^{t_f}\int_\Omega \sum_{i=1}^n \left(T[g^k](\boldsymbol{x},t) - Z_i(t)\right)\tilde{T}(\boldsymbol{x},t)\delta_{x_i}(\boldsymbol{x})\, d\boldsymbol{x}\, dt}{\int_0^{t_f}\int_\Omega \sum_{i=1}^n \left(\tilde{T}(\boldsymbol{x},t)\right)^2 \delta_{x_i}(\boldsymbol{x})\, d\boldsymbol{x}\, dt} \tag{11}$$

where $\tilde{T}$ is the solution of the sensitivity equation, a linear partial differential equation which is presented is Section 5.2. Therefore, the determination of α^k, in the linear case, involves solving two sets of partial differential equations to get $T[g^k]$ and $\tilde{T}$ and computing the integrals in Eq. (11). For the nonlinear case we could still use α^k in Eq. (11) as a rough approximation to the solution of Eq. (9). However, the determination of $T[g^k]$ would involve solving a nonlinear partial differential equation. We present the derivation of the step size, Eq (11), in Section 5.5.

Since in fact the solution of Eq. (9) is not what is at stake here, one wants to solve Eq. (7), we can consider a variant of the steepest descent method, known as Landweber iteration, which is obtained by a further simplification and is defined by changing Eq. (10) into $g^{k+1} = g^k + \beta p^k$, where $\beta < 0$ is a constant chosen arbitrarily. By doing that one avoids altogether having to compute α^k in Eq. (11) and even of solving Eq. (9).

[3] As an example, one may set $g^o(\boldsymbol{x},t) = 0$.

4.2 CG: conjugate gradient method

In this method the 2^{nd} step of the algorithm is altered. Equation (8) for the descent direction is modified and the direction at iteration k is constructed as a combination of the gradient E'_{g^k} with the descent direction of the previous iteration $k-1$,

$$p^k \;=\; E'_{g^k} + \gamma^k p^{k-1}, \text{ with } \gamma^0 = 0, \text{ for } k \geq 0 \tag{12}$$

where the conjugate coefficient γ^k may be calculated in several ways being one of the most frequently used,

$$\gamma^k \;=\; \frac{\int_0^{t_f}\int_\Omega \left(E'_{g^k}\right)^2 \, d\boldsymbol{x}\, dt}{\int_0^{t_f}\int_\Omega \left(E'_{g^{k-1}}\right)^2 \, d\boldsymbol{x}\, dt}$$

One then proceeds with the remaining steps as above to determine the sucessive approximations.

5 THE GRADIENT, THE SENSITIVITY AND THE ADJOINT PROBLEMS

5.1 Directional derivatives and the definition of the gradient

We consider now the question of determining the gradient of E. This is of surmount importance in these problems since it is used in defining the search direction as can be seen in Eq. (8) and Eq. (12) for the steepest descent, Landweber or conjugate gradient methods. We tackle this by considering a directional derivative (Gateâux derivative) of the functional E. First, let $g_\epsilon = g_\epsilon(\boldsymbol{x}, t)$ denote a one parameter family of functions such that:

$$g_\epsilon\big|_{\epsilon=0} = g_0(\boldsymbol{x},t) \;=\; g(\boldsymbol{x},t) \quad \text{and} \quad \left.\frac{dg}{d\epsilon}\right|_{\epsilon=0} \;=\; \tilde{g}(\boldsymbol{x},t)$$

We may think of g_ϵ as a small perturbation[4] of g; in fact, $g_\epsilon \sim g + \epsilon\tilde{g}$, as $\epsilon \to 0$. If we denote by $dE_g[\tilde{g}]$ the directional derivative of E at the point g in the direction $\tilde{g}$, then,

$$dE_g[\tilde{g}] = \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} E\left[g_\epsilon\right] = \int_0^{t_f}\int_\Omega \left(\sum_{i=1}^{n} \left\{T[g](\boldsymbol{x},t) - Z_i(t)\right\} \delta_{x_i}(\boldsymbol{x})\right) \tilde{T}(\boldsymbol{x},t)\, d\boldsymbol{x}\, dt \tag{13}$$

Here, $\tilde{T} = \tilde{T}(\boldsymbol{x},t)$ stands for the directional derivative of T at g in the direction $\tilde{g}$,

$$\tilde{T}(\boldsymbol{x},t) \;=\; \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} T[g_\epsilon](\boldsymbol{x},t) \tag{14}$$

and it is the solution of the so-called *sensitivity problem* which we discuss further below in Section 5.2. We observe that the notation used in Eq. (14) does not refer explicitly to the dependence of $\tilde{T}$ on $\tilde{g}$ nor on g. For linear problems, in fact, $\tilde{T}$ does not depend on g,

[4] This setup is most convenient when considering nonlinear problems. For linear problems, one usually writes equations for $T + \Delta T$ due to a source $g + \Delta g$ and gets information for the first order terms in ΔT and Δg, neglecting higher order terms in ΔT and Δg, (Silva Neto and Özisik, 1994).

only on $\tilde{g}$. Nonetheless, $\tilde{T}$ depends linearly on $\tilde{g}$ (either for linear or nonlinear problems) and we write that dependence by means of a linear operator P,

$$\tilde{T} \;=\; P\tilde{g} \tag{15}$$

Then, from Eq. (13),

$$dE_g[\tilde{g}] \;=\; \int_0^{t_f}\int_\Omega \left(\sum_{i=1}^{n} \{T[g](\boldsymbol{x},t) - Z_i(t)\}\,\delta_{x_i}(\boldsymbol{x})\right)(P\tilde{g})(\boldsymbol{x},t)\,d\boldsymbol{x}\,dt \tag{16}$$

We recall that by the definition of the gradient of a functional E at a point g, denoted here by E'_g, the gradient is just the function that represents the (Fréchet) derivative, dE_g, with respect to the inner product, that is, such that:

$$dE_g[\tilde{g}] \;=\; \langle E'_g, \tilde{g}\rangle \;=\; \int_0^{t_f}\int_\Omega E'_g(\boldsymbol{x},t)\tilde{g}(\boldsymbol{x},t)\,d\boldsymbol{x}\,dt \text{ for all } \tilde{g} \tag{17}$$

In order to get E'_g one needs to change,in Eq. (16), the action of the linear operator P from $\tilde{g}$ over to $\sum_{i=1}^{n}\{T[g](\boldsymbol{x},t) - Z_i(t)\}\,\delta_{x_i}(\boldsymbol{x})$ as can be seen by comparing Eq. (16) with Eq. (17). This is done by means of an adjoint operator.

Before we consider further the determination of the gradient, we look at the sensitivity problem.

5.2 Sensitivity problem

By definition of $\tilde{T}$ (given by Eq. (14)), we see that in the problems here considered, it is the solution of the linearization of problem (1,2,5). In fact, writing $T_\epsilon(\boldsymbol{x},t) = T[g_\epsilon](\boldsymbol{x},t)$, we have that:

$$\begin{aligned} C(T_\epsilon)\frac{\partial T_\epsilon}{\partial t} &= \mathsf{div}\,(k(T_\epsilon)\nabla T_\epsilon) + g_\epsilon, \quad \boldsymbol{x}\in\Omega,\ t>0 && (18)\\ T_\epsilon(\boldsymbol{x},0) &= T^*(\boldsymbol{x}), \quad \boldsymbol{x}\in\Omega && (19)\\ -k(T_\epsilon)\boldsymbol{n}\cdot\nabla T_\epsilon(\boldsymbol{x},t) &= h\left(T_\epsilon(\boldsymbol{x},t) - T_e(\boldsymbol{x},t)\right), \quad \boldsymbol{x}\in\partial\Omega,\ t>0 && (20) \end{aligned}$$

One looks for the equation that $\tilde{T} = \frac{d}{d\epsilon}\Big|_{\epsilon=0} T_\epsilon$ has to satisfy. Differentiating Eq. (18-20) with respect to ϵ and letting $\epsilon = 0$, we have that $\tilde{T}$ satisfies[5]:

$$\begin{aligned} \frac{\partial}{\partial t}\left(C(T_0)\tilde{T}\right) &= \triangle\left(k(T_0)\tilde{T}\right) + \tilde{g},\ \boldsymbol{x}\in\Omega, t>0 && (21)\\ \tilde{T}(\boldsymbol{x},0) &= 0,\ \boldsymbol{x}\in\Omega && (22)\\ -\boldsymbol{n}(\boldsymbol{x})\cdot\nabla\left(k(T_0(\boldsymbol{x},t))\tilde{T}(\boldsymbol{x},t)\right) &= h\tilde{T}(\boldsymbol{x},t)\ \boldsymbol{x}\in\partial\Omega,\ t>0 && (23) \end{aligned}$$

Here we recall that $T_0 = T[g_0] = T[g]$. We note that (21)-(23) is a linear problem for $\tilde{T}$ with respect to $\tilde{g}$, independently of whether the original problem is linear or nonlinear. In particular the boundary and the initial conditions are homogeneous in $\tilde{T}$. We are thus justified to write Eq. (15).

---

[5]We have used that

$$C(T_0)\frac{\partial\tilde{T}}{\partial t} + C'(T_0)\tilde{T}\frac{\partial T_0}{\partial t} = \frac{\partial}{\partial t}\left(C(T_0)\tilde{T}\right) \quad\text{and that}\quad k(T_0)\nabla\tilde{T} + k'(T_0)\tilde{T}\nabla T_0 = \nabla\left(k(T_0)\tilde{T}\right)$$

where C', k' are the derivatives of C and k with respect to T.

5.3 Adjoint operator and adjoint problem

Consider the linear differential operator

$$L\phi \;=\; \frac{\partial}{\partial t}\left(C(T_0)\phi\right) - \triangle(k(T_0)\phi) \tag{24}$$

which we have gotten from considering the homogeneous terms in $\tilde{T}$ on Eq. (21). Let it be defined on the set of functions

$$\mathcal{U} = \{\phi \,|\, \phi(\boldsymbol{x}, 0) = 0 \,\text{for}\, \boldsymbol{x} \in \Omega, \text{ and } -\boldsymbol{n}\cdot\nabla(k(T_0)\phi) = h\phi,\; \boldsymbol{x} \in \partial\Omega, 0 \le t \le t_f\} \tag{25}$$

which is suggested by Eq. (22) and Eq. (23) (just replacing $\tilde{T}$ by ϕ).

The formal adjoint operator $\mathcal{L}^*$ will be defined in an appropriate function space, $\mathcal{V}$, in such a way that

$$\langle \psi, \mathcal{L}\phi\rangle \;=\; \langle \mathcal{L}^*\psi, \phi\rangle, \quad \text{for all } \phi \in \mathcal{U}, \psi \in \mathcal{V} \tag{26}$$

By definition of $\mathcal{L}$ and integration by parts we have[6]

$$\begin{aligned}
\langle \psi, \mathcal{L}\phi\rangle &= \int_0^{t_f}\int_\Omega \psi\left\{\frac{\partial}{\partial t}(C(T_0)\phi) - \mathsf{div}\,\left(\nabla(k(T_0)\phi)\right)\right\} d\boldsymbol{x}\,dt \\
&= \int_0^{t_f}\int_\Omega \left[-C(T_0)\frac{\partial\psi}{\partial t} - k(T_0)\triangle\psi\right]\phi\, d\boldsymbol{x}\,dt \qquad (27)\\
&+ \underbrace{\int_\Omega \psi C(T_0)\phi|_0^{t_f}\; d\boldsymbol{x} - \int_0^{t_f}\int_{\partial\Omega}\left[\psi\boldsymbol{n}\cdot\nabla(k(T_0)\phi) + k(T_0)\boldsymbol{n}\cdot(\nabla\psi)\,\phi\right] dS_x\,dt}_{\text{boundary terms}}
\end{aligned}$$

Using the fact that $\phi \in \mathcal{U}$, (see Eq. (25)), we have that:

$$\begin{aligned}
\text{boundary terms} &= \int_\Omega \psi(\boldsymbol{x}, t_f)C(T_0(\boldsymbol{x}, t_f))\phi(\boldsymbol{x}, t_f)\,d\boldsymbol{x} \\
&- \int_0^{t_f}\int_{\partial\Omega}\left[-h\psi(\boldsymbol{x}, t) + k(T_0)\boldsymbol{n}\cdot\nabla\psi\right]\phi(\boldsymbol{x}, t)\,dS_x\,dt
\end{aligned} \tag{28}$$

If we now take the differential operator and function space as

$$\mathcal{L}^* \;=\; -C(T_0)\frac{\partial}{\partial t} - k(T_0)\triangle \tag{29}$$

$$\mathcal{V} \;=\; \{\psi \,|\, \psi(\boldsymbol{x}, t_f) = 0, \text{ for } \boldsymbol{x} \in \Omega, \text{and } k(T_0)\boldsymbol{n}\cdot\nabla\psi = h\psi, \boldsymbol{x} \in \partial\Omega,\; 0 \le t \le t_f\}$$

we have from the definition of $\mathcal{V}$ that the boundary terms in Eq. (28) are null and from Eq. (27) and the definition of $\mathcal{L}^*$ in Eq. (29) that Eq. (26) is satisfied. We then say that $\mathcal{L}^*$ on $\mathcal{V}$ is the *formal adjoint operator* of $\mathcal{L}$ on $\mathcal{U}$.

We are now in a position to define the adjoint problem. This is done by looking at the differential operator $\mathcal{L}^*$ and the function space $\mathcal{V}$, which suggests the equation and

---

[6]Here we have used that:

$$\begin{aligned}
\psi\frac{\partial}{\partial t}\left(C(T_0)\phi\right) &= -\frac{\partial\psi}{\partial t}C(T_0)\phi + \frac{\partial}{\partial t}\left(\psi C(T_0)\phi\right) \text{ and} \\
-\psi\triangle\left(K(T_0)\phi\right) &= -K(T_0)(\triangle\psi)\phi - \mathsf{div}\,\left(\psi\nabla(K(T_0)\phi)\right) - \mathsf{div}\left(K(T_0)(\nabla\psi)\phi\right)
\end{aligned}$$

adequate initial and boundary conditions. We say that θ satisfies the *adjoint problem* (to the sensitivity problem (21-23)) if:

$$-C(T_0)\frac{\partial}{\partial t}\theta = k(T_0) \triangle \theta + l(\boldsymbol{x},t) \ \boldsymbol{x} \in \Omega,\ 0 < t < t_f \tag{30}$$
$$\theta(\boldsymbol{x},t_f) = 0,\ \boldsymbol{x} \in \Omega \tag{31}$$
$$k(T_0)\boldsymbol{n}\cdot\nabla\psi = h\psi,\ \boldsymbol{x} \in \partial\Omega,\ 0 < t < t_f \tag{32}$$

This is a *backward heat transfer problem* (the final 'temperature' distribution is given instead of the initial 'temperature'), see Eq. (31). Here l is a source term.

We denote by Q the solution operator of problem defined by Eq. (30-32), that is,

$$\theta(\boldsymbol{x},t) = (Ql)(\boldsymbol{x},t) \tag{33}$$

and note that it is a linear operator in l.

5.4 A differential equation to determine the gradient

It is not a very trivial fact[7] but, in some cases, the adjoint of the solution operator of the sensitivity problem (P^*, with P defined in Eq. (15)) is the solution operator of the adjoint problem (Q defined in Eq. (33)), that is $P^* = Q$.

From Eq. (16) we have that:

$$dE_g[\tilde{g}] = \int_0^{t_f}\int_\Omega \left(P^*\left[\sum_{i=1}^n (T[g]-Z_i)\,\delta_{x_i}\right]\right)(\boldsymbol{x},t)\tilde{g}(\boldsymbol{x},t)\,d\boldsymbol{x}\,dt \tag{34}$$

Comparing Eq. (34) with Eq. (17), and using that $P^* = Q$, we conclude that:

$$E'_g = P^*\left(\sum_{i=1}^n (T[g]-Z_i)\,\delta_{x_i}\right) \tag{35}$$

Some remarks about Eq. (35) are in order. Say one wants to compute the gradient of E at g, that is, E'_g. Then, from Eq. (35), we see that one has to solve the adjoint problem, Eq. (30-32) with the source term given by

$$l = \sum_{i=1}^n (T[g]-Z_i)\,\delta_{x_i} \tag{36}$$

However, the adjoint problem depends on the knowledge of T_0 (see Eq. (30-32). Also the source term, Eq. (36), depends on it since $T[g] = T_0$. The function T_0 is the solution of the (possibly nonlinear) DHCP defined by Eq. (1,2,5). Summarizing, to determine E'_g one has first to solve a DHCP to get $T[g]$ and secondly an adjoint problem with source term given by Eq. (36). This comprises two sets of partial differential equations.

5.5 Step size

In the linear case, $T[g^k + \alpha p^k] = T[g^k] + \alpha\tilde{T}$ with $\tilde{T}$ denoting the solution of the sensitivity problem with source term given by p^k. Therefore, $E[g^k + \alpha p^k]$ is a quadratic polynomial in α and has a minimum value exactly at α^k given by Eq. (11).

[7] If $\mathcal{L}$ denotes the linear operator defining the linear problem, $\mathcal{L}^{-1}$ will be the solution operator. The assertion that *the adjoint of the solution operator* is *the solution of the adjoint problem* can be written as: $(\mathcal{L}^{-1})^* = (\mathcal{L}^*)^{-1}$. Formally, this is shown as: $\langle u,(\mathcal{L}^{-1})^*v\rangle = \langle\mathcal{L}^{-1}u,\mathcal{L}^*(\mathcal{L}^*)^{-1}v\rangle = \langle\mathcal{L}\mathcal{L}^{-1}u,(\mathcal{L}^*)^{-1}v\rangle = \langle u,(\mathcal{L}^*)^{-1}v\rangle$, $\forall u,v$. The derivation presented is alright for bounded operators in Hilbert spaces. It lacks rigour for typical IHCP, where unbounded operators are involved and the mathematics is much more involved, however it is not our aim here to consider that.

6 CONCLUSIONS AND PERSPECTIVES FOR FUTURE WORK

We presented a general framework do deal with IHCP which applies both to linear and nonlinear problems. We emphasized key concepts such as linearization, adjoints, Fréchet and Gateâux derivatives and the gradient. The determination of the gradient of the error function is central in a broad range of minimization methods needed to solve IHCP. Nonlinear problems are more complex due to the need of dealing with more abstract mathematical concepts. However, we will present elsewhere an operational procedure to get the gradient without explicitely considering the full abstract setup, and we will show that the procedure is mathematicaly sound. The framework discussed is needed to extend to nonlinear problems the investigation initiated in Silva Neto and Özisik, 1994. We intend to extend GMRES methodology, introduced by Saad and Schultz, 1986, for finite dimensional problems, to deal with IHCP. We will investigate and compare the performance of the methods based on the conjugate gradient and GMRES on typical IHCP.

References

Alifanov, O. M., Artyukhin, E. A. and Rumiantsev, S. V., 1995, Extreme Methods for Solving Ill-Posed Problems with Applications to Inverse Heat Transfer Problems, Springer Verlag.

Artyukhin, E. A., Ivanov, G. A. and Nenarokomov, A. V., 1993, Determining the set of thermophysical properties of materials from unsteady-state temperature measurements, High Temperature, Vol. 32 (2), pp. 199-202.

Beck, J. V., Blackwell, B. and St. Clair Jr., C. R., 1985, Inverse Heat Conduction, Wiley, New York.

Blanc, G., Raynaud, M. and Chau, T. H., 1998, A guide for the use of the function specification method for 2D inverse heat conduction problems, Rev. Gén. Therm., Vol. 37, pp. 17-30.

Folland, G. B., 1976, Introduction to Partial Differential Equations, Princeton University Press, New Jersey.

Friedman, A., 1964, Partial Differential Equations of Parabolic Type, Prentice Hall, New York.

Richtmyer, R. D. and Morton, K. W., 1967, Difference Methods for Initial-Value Problems, John Wiley (Interscience), New York.

Saad, Y. and Schultz, M. H., 1986, GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems, SIAM J. Sci. Stat. Comput., Vol. 7(3), pp. 856-869.

Silva Neto, A. J. and Özisik, M. N., 1994, The estimation of space and time dependent strength of a volumetric heat source in a one-dimensional plate, Int. J. Heat and Mass Transfer, Vol. 37(6), pp. 909-915.

Sod, G. A., 1985, Numerical Methods in Fluid Mechanics, Prentice Hall, New York.