J. Austral. Math. Soc. Ser. B 40(1998), 266–287

COMPUTING OPTIMAL CONTROL WITH A HYPERBOLIC PARTIAL DIFFERENTIAL EQUATION

V. REHBOCK[1], S. WANG[1] and K. L. TEO[1]

(Received 19 July 1996; revised 8 November 1996)

Abstract

We present a method for solving a class of optimal control problems involving hyperbolic partial differential equations. A numerical integration method for the solution of a general linear second-order hyperbolic partial differential equation representing the type of dynamics under consideration is given. The method, based on the piecewise bilinear finite element approximation on a rectangular mesh, is explicit. The optimal control problem is thus discretized and reduced to an ordinary optimization problem. Fast automatic differentiation is applied to calculate the exact gradient of the discretized problem so that existing optimization algorithms may be applied. Various types of constraints may be imposed on the problem. A practical application arising from the process of gas absorption is solved using the proposed method.

1. Introduction

Finding optimal controls for problems involving partial differential equations (PDEs) has been the subject of a good proportion of the control literature for many years (see [1]–[7], [9], [11], [12], [16], [17], [20], [22], [23], [25] and the relevant references cited therein). However, unlike the case of optimal control problems involving ordinary differential equations (see [10] and [19]), the majority of this work concentrates on the theoretical aspects of the problem. Comparatively little work has been done to find computational methods for solving such problems effectively (see [4], [5], [20], [26] and the relevant references cited therein). This is mainly due to the computational complexity of the problems.

The purpose of this paper is to present a computational method for solving a general class of optimal control problems subject to a second order linear hyperbolic PDE. Various types of constraints may also be imposed. We present the different aspects of

[1]School of Mathematics and Statistics, Curtin University of Technology, GPO Box U1987, Perth, 6845, Australia

© Australian Mathematical Society, 1998, Serial-fee code 0334-2700/98

https://doi.org/10.1017/S0334270000012510 Published online by Cambridge University Press

the method as follows.

The dynamics of the problem under consideration are governed by a general second order hyperbolic partial differential equation. A discretization method based on the integration of the equation over a rectangular element and on the bilinear approximation of the control on that element is proposed. This yields a second order 2-point implicit scheme (similar to the box or Crank-Nicholson scheme; cf., for example, [24, Chapters 2 and 4]) along both the x and t directions. It turns out that the resulting scheme is explicit and we show that the method is unconditionally stable, provided the coefficients in the equation are positive. The discretization error is shown to be of second order in both the independent variables.

The discretization turns the problem into a standard mathematical programming problem, although the objective and constraint gradients have to be calculated in a roundabout manner. This is usually done by first calculating the gradients of the original (nondiscretized) problem, in the process yielding a continuous PDE for the costates of the problem (which also arises in the application of a maximum principle). For a numerical method, this costate equation is then discretized. Unless care is taken to correctly match the costate discretization with the discretization of the state equations, a small amount of error is introduced into the gradient values. This error usually has a detrimental effect on the optimization process and results in poor convergence.

To overcome the possibility of these numerical difficulties, we bypass the gradient calculations of the original nondiscretized problem and calculate the gradients of the discretized problem directly by fast automatic differentiation (FAD) (see [14] and the relevant references cited therein). FAD has been applied successfully in a variety of problems, including optimal control problems (see [13]). In the optimal control applications, it has been shown to yield the discretized costate equation of a problem directly without first deriving the continuous costate equation [13]. These are the exact equations for the discretized problem (unlike those which can be obtained by possibly carelessly discretizing the continuous version of the costate equation). They yield exact gradient values and hence the convergence of the optimization process is not impeded.

As an application example, we consider the problem of optimizing a gas absorption process. Numerical results are presented.

2. Discretization of the system dynamics

Consider a linear hyperbolic equation of the form

$$Lz := \frac{\partial^2 z}{\partial x \partial t} + a(x,t)\frac{\partial z}{\partial x} + b(x,t)\frac{\partial z}{\partial t} = f(x,t) \quad \text{in } \Omega = (0,1)^2 \tag{2.1}$$

https://doi.org/10.1017/S0334270000012510 Published online by Cambridge University Press

with Darboux boundary conditions

$$z(x,0) = p(x), \quad z(0,t) = q(t), \tag{2.2}$$

where $a(x,t)$, $b(x,t)$, $f(x,t)$, $p(x)$ and $q(t)$ are given functions. The unknown z and all the known functions in the above are either scalar or vector-valued. This problem has been studied extensively in the context of optimal control (cf., for example, [15], [26] and the references cited therein). Despite its importance in chemical processes and optimal control, the numerical solution of this problem, to the best of knowledge, has not been investigated thoroughly.

We now discuss the discretization of (2.1) with the boundary conditions (2.2). Let Ω be partitioned into rectangular elements with mesh node distributions $\{x_i\}_0^I$ and $\{t_j\}_0^J$ respectively on the x-axis and the t-axis. We let $\Delta x_i = x_{i+1} - x_i$ and $\Delta t_j = t_{j+1} - t_j$ for $i = 0, 1, \ldots, I-1$ and $j = 0, 1, \ldots, J-1$. Integrating (2.1) over $\Omega_{i,j} = (x_i, x_{i+1}) \times (t_j, t_{j+1})$, we have

$$\int_{\Omega_{i,j}} Lz(x,t)\, d\Omega = \int_{\Omega_{i,j}} f(x,t)\, d\Omega. \tag{2.3}$$

For any $0 \le i \le I$ and $0 \le j \le J$, let $\phi_{i,j}(x,t)$ be the conventional piecewise bilinear basis function associated with the node (x_i, t_j), and let

$$z_h(x,t) = \sum_{0 \le i \le I, 0 \le j \le J} z_i^j \phi_{i,j}(x,t),$$

where $\{z_i^j\}$ are approximations to the nodal values of z. We also let $z_i(t) = z_h(x_i, t)$ and $z^j(x) = z_h(x, t_j)$. Obviously both $z_i(t)$ and $z^j(x)$ are piecewise linear on [0, 1].

Now, replacing z in (2.3) by z_h we obtain the equation

$$\int_{\Omega_{i,j}} Lz_h\, d\Omega = \int_{\Omega_{i,j}} \left(\frac{\partial^2 z_h}{\partial x \partial t} + a(x,t)\frac{\partial z_h}{\partial x} + b(x,t)\frac{\partial z_h}{\partial t} \right) d\Omega = \int_{\Omega_{i,j}} f(x,t)\, d\Omega \tag{2.4}$$

for all $i = 0, 1, \ldots, I-1$ and $j = 0, 1, \ldots, J-1$. By direct integration on $\Omega_{i,j}$ we obtain

$$\int_{t_j}^{t_{j+1}} \int_{x_i}^{x_{i+1}} \frac{\partial^2 z_h}{\partial x \partial t}\, dx dt = \int_{t_j}^{t_{j+1}} \left(\frac{dz_{i+1}}{dt} - \frac{dz_i}{dt} \right) dt = (z_{i+1}^{j+1} - z_{i+1}^j) - (z_i^{j+1} - z_i^j). \tag{2.5}$$

Because z_h is piecewise bilinear for any fixed x or fixed t, both $\frac{\partial z_h}{\partial x}$ and $\frac{\partial z_h}{\partial t}$ are constant on any line segment within an element parallel respectively to the x-axis and the t-axis. Integrating the second term in (2.4) by parts, we have

$$\int_{t_j}^{t_{j+1}} \int_{x_i}^{x_{i+1}} a \frac{\partial z_h}{\partial x}\, dx dt = \int_{t_j}^{t_{j+1}} \left(\frac{z_{i+1} - z_i}{\Delta x_i} \int_{x_i}^{x_{i+1}} a(x,t)\, dx \right) dt$$

https://doi.org/10.1017/S0334270000012510 Published online by Cambridge University Press

$$
\begin{aligned}
&= \int_{t_j}^{t_{j+1}} \overline{a}_i(t)\,(z_{i+1}(t) - z_i(t))\,dt \\
&= A_i(t)(z_{i+1}(t) - z_i(t))\Big|_{t_j}^{t_{j+1}} - \int_{t_j}^{t_{j+1}} A_i \left(\frac{dz_{i+1}}{dt} - \frac{dz_i}{dt}\right) dt \\
&= A_i^{j+1}\left(z_{i+1}^{j+1} - z_i^{j+1}\right) - A_i^j\left(z_{i+1}^j - z_i^j\right) \\
&\quad - \overline{A}_i^j\left[\left(z_{i+1}^{j+1} - z_{i+1}^j\right) - \left(z_i^{j+1} - z_i^j\right)\right] \\
&= \left(A_i^{j+1} - \overline{A}_i^j\right)\left(z_{i+1}^{j+1} - z_i^{j+1}\right) + \left(\overline{A}_i^j - A_i^j\right)\left(z_{i+1}^j - z_i^j\right),
\end{aligned}
\tag{2.6}
$$

where

$$
\overline{a}_i(t) = \frac{1}{\Delta x_i}\int_{x_i}^{x_{i+1}} a(x,t)\,dx, \quad A_i(t) = \int \overline{a}_i(t)\,dt, \quad \overline{A}_i^j = \frac{1}{\Delta t_j}\int_{t_j}^{t_{j+1}} A_i(t)\,dt
\tag{2.7}
$$

and $A_i^k = A_i(t_k)$, $k = j, j+1$. Similarly we have

$$
\begin{aligned}
\int_{t_j}^{t_{j+1}}\int_{x_i}^{x_{i+1}} b\frac{\partial z_h}{\partial t}\,dx\,dt &= \int_{x_i}^{x_{i+1}} \overline{b}^j(x)(z^{j+1}(x) - z^j(x))\,dx \\
&= B^j(x)\left(z^{j+1}(x) - z^j(x)\right)\Big|_{x_i}^{x_{i+1}} \\
&\quad - \int_{x_i}^{x_{i+1}} B^j\left(\frac{dz^{j+1}}{dx} - \frac{dz^j}{dx}\right) dx \\
&= \left(B_{i+1}^j - \overline{B}_i^j\right)\left(z_{i+1}^{j+1} - z_{i+1}^j\right) + \left(\overline{B}_i^j - B_i^j\right)(z_i^{j+1} - z_i^j),
\end{aligned}
\tag{2.8}
$$

where

$$
\begin{aligned}
\overline{b}^j(x) = \frac{1}{\Delta t_j}\int_{t_j}^{t_{j+1}} b(x,t)\,dt, \qquad B^j(x) = \int \overline{b}^j(x)\,dx, \\
\overline{B}_i^j = \frac{1}{\Delta x_i}\int_{x_i}^{x_{i+1}} B^j(x)\,dx
\end{aligned}
\tag{2.9}
$$

and $B_k^j = B^j(x_k)$, $k = i, i+1$. Substituting (2.5), (2.6) and (2.8) into (2.4) we obtain

$$
\begin{aligned}
&\left(1 + C_i^j + E_i^j\right) z_{i+1}^{j+1} - \left(1 + C_i^j - F_i^j\right) z_i^{j+1} - \left(1 - D_i^j + E_i^j\right) z_{i+1}^j \\
&\quad + \left(1 - D_i^j - F_i^j\right) z_i^j = G_{i,j},
\end{aligned}
\tag{2.10}
$$

where

$$
C_i^j = A_i^{j+1} - \overline{A}_i^j, \quad D_i^j = \overline{A}_i^j - A_i^j, \quad E_i^j = B_{i+1}^j - \overline{B}_i^j, \quad F_i^j = \overline{B}_i^j - B_i^j
\tag{2.11}
$$

https://doi.org/10.1017/S0334270000012510 Published online by Cambridge University Press

and $G_{i,j}$ denotes the integral on the right side of (2.4). From (2.10) we obtain

$$z_{i+1}^{j+1} = \left(1 + C_i^j + E_i^j\right)^{-1} \Big[\left(1 + C_i^j - F_i^j\right) z_i^{j+1} + \left(1 - D_i^j + E_i^j\right) z_{i+1}^j - \left(1 - D_i^j - F_i^j\right) u_i^j + G_{i,j}\Big] \tag{2.12}$$

for all $i = 0, 1, \ldots, I-1$ and $j = 0, 1, \ldots, J-1$. This, together with the boundary conditions in (2.2), yields an explicit scheme for the solution of (2.1). In the case that a and b are constant on Ω, the coefficients in (2.11) reduce to

$$C_i^j = \frac{\Delta t_j}{2} a = D_i^j, \quad E_i^j = \frac{\Delta x_i}{2} b = F_i^j.$$

We now discuss the stability of the method. As mentioned before, the method is equivalent to applying a 2-point implicit scheme to (2.1) along both the x and t directions (see pages 29 and 59 of [24]). We show only that the method is unconditionally stable along the t direction. The stability along x direction then follows because of the symmetry of the problem and the method.

For any $j = 0, 1, \ldots, J-1$, (2.11) can be rewritten as

$$P^j z^{j+1} = R^j z^j + g^j, \tag{2.13}$$

where $z^j = (z_1^j, z_2^j, \ldots, z_I^j)^T$, $g^j = (g_1^j, g_2^j, \ldots, g_I^j)^T$ is a known vector consisting of the terms $G_{i,j}$ in (2.11) and some boundary values, and P^j and R^j are two $I \times I$ lower triangular matrices in the forms

$$P^j = \begin{pmatrix} p_{11} & 0 & 0 & \ldots & 0 & 0 & 0 \\ p_{21} & p_{22} & 0 & \ldots & 0 & 0 & 0 \\ 0 & p_{32} & p_{33} & \ldots & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & p_{I-1I-2} & p_{I-1I-1} & 0 \\ 0 & 0 & 0 & \ldots & 0 & p_{II-1} & p_{II} \end{pmatrix}$$

and

$$R^j = \begin{pmatrix} r_{11} & 0 & 0 & \ldots & 0 & 0 & 0 \\ r_{21} & r_{22} & 0 & \ldots & 0 & 0 & 0 \\ 0 & r_{32} & r_{33} & \ldots & 0 & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ldots & r_{I-1I-2} & r_{I-1I-1} & 0 \\ 0 & 0 & 0 & \ldots & 0 & r_{II-1} & r_{II} \end{pmatrix}.$$

https://doi.org/10.1017/S0334270000012510 Published online by Cambridge University Press

The nonzero entries in P^j and R^j are

$$p_{ii} = 1 + C^j_{i-1} + E^j_{i-1}, \quad p_{ii-1} = -(1 + C^j_{i-1} - F^j_{i-1}), \tag{2.14}$$

$$r_{ii} = 1 - D^j_{i-1} + E^j_{i-1}, \quad r_{ii-1} = -(1 - D^j_{i-1} - F^j_{i-1}). \tag{2.15}$$

From (2.13) we obtain

$$z^{j+1} = K^j z^j + \left(P^j\right)^{-1} g^j \tag{2.16}$$

with $K^j = \left(P^j\right)^{-1} R^j$. Now the stability of (2.16) (Von Neumann condition) requires that all the eigenvalues of K^j and $\left(P^j\right)^{-1}$ lie on the interval $[-1, 1]$. We first show the former. Because both P^j and R^j are lower triangular matrices, K^j is also a lower triangular matrix, and so we need only to show that $|k_{ii}| \leq 1$ for all $i = 1, 2, \dots, I$, where k_{ii} denotes the diagonal element of K^j in row i. From the definitions of P^j and R^j and the relationships (2.14) and (2.15) we see that

$$k_{ii} = \frac{r_{ii}}{p_{ii}} = \frac{1 - D^j_{i-1} + E^j_{i-1}}{1 + C^j_{i-1} + E^j_{i-1}}.$$

So, if we can show that both D^j_{i-1} and C^j_{i-1} are nonnegative, then $|k_{ii}| \leq 1$. In fact, since $a(x, t)$ is positive, from (2.7) we know that $A_i(t)$ is strictly increasing. Thus, from (2.11) we have that C^{j+1}_{i-1} and D^j_{i-1} are positive, and so $|k_{ii}| < 1$. Similarly, it is easy to show that all the eigenvalues of $\left(P^j\right)^{-1}$ lie in $(-1, 1)$ because $p_{ii} > 1$ for all $i = 1, 2, \dots, I$. Therefore we have shown that the scheme is unconditionally stable along the t direction. The stability along the x direction follows similarly because both the problem and the method are symmetric with respect to t and x. We comment that even if D^j_{i-1} and C^j_{i-1} are negative, from the above equality we see that $k_{ii} \to 1$ as $\Delta t_j \to 0$, because both D^j_{i-1} and C^j_{i-1} are of order Δt_j. So, in this case, the method may still be stable when Δx_i and Δt_j are sufficiently small. We now discuss the accuracy of the method. Recall that $z = z(x, t)$ denotes the solution of the undiscretized problem defined by (2.1) and (2.2). Let

$$Z = \sum_{0 \leq i \leq I, 0 \leq j \leq J} z(x_i, t_j) \phi_{i,j}(x, t),$$

that is, Z is the interpolant of z in the space spanned by the piecewise bilinear basis functions $\{\phi_{i,j}\}$. Because Z is bilinear on each $\Omega_{i,j}$, from the deduction of (2.10) we see that

$$\int_{\Omega_{i,j}} LZ(x, t)\, d\Omega = \left(1 + C^j_i + E^j_i\right) z(x_{i+1}, t_{j+1}) - \left(1 + C^j_i - F^j_i\right) z(x_i, t_{j+1})$$

https://doi.org/10.1017/S0334270000012510 Published online by Cambridge University Press

$$- \left(1 - D_i^j + E_i^j\right) z(x_{i+1}, t_j) + \left(1 - D_i^j - F_i^j\right) z(x_i, t_j).$$

Subtracting this equation from (2.10) we obtain

$$\begin{aligned} &\left(1 + C_i^j + E_i^j\right) e_{i+1}^{j+1} - \left(1 + C_i^j - F_i^j\right) e_i^{j+1} - \left(1 - D_i^j + E_i^j\right) e_{i+1}^j \\ &\qquad + \left(1 - D_i^j - F_i^j\right) e_i^j = G_{i,j} - \int_{\Omega_{i,j}} LZ(x,t)\, d\Omega, \end{aligned} \tag{2.17}$$

where $e_i^j = z_i^j - z(x_i, t_j)$. Notice that

$$\int_{t_j}^{t_{j+1}} \int_{x_i}^{x_{i+1}} \frac{\partial^2 (z - Z)}{\partial x \partial t}\, dx dt = 0.$$

The last term in (2.17) becomes

$$\begin{aligned} \left| G_{i,j} - \int_{\Omega_{i,j}} LZ(x,t)\, d\Omega \right| &= \left| \int_{\Omega_{i,j}} L(z - Z(x,t))\, d\Omega \right| \\ &= \left| \int_{\Omega_{i,j}} \left(a(x,t) \frac{\partial (z - Z)}{\partial x} + b(x,t) \frac{\partial (z - Z)}{\partial t} \right) d\Omega \right| \\ &\le \int_{\Omega_{i,j}} (a + b)\, d\Omega \sup_{(x,t) \in \Omega_{i,j}} \left(\left| \frac{\partial (z - Z)}{\partial x} \right| + \left| \frac{\partial (z - Z)}{\partial x} \right| \right). \end{aligned}$$

Let h be the mesh parameter such that $\Delta x_i, \Delta t_j \le h$ for all i and j. Because Z is the bilinear interpolant of z, using the standard argument of interpolation in Sobolev spaces (cf., for example, [8], §4.4) we have

$$\sup_{(x,t) \in \Omega_{i,j}} \left(\left| \frac{\partial (z - Z)}{\partial x} \right| + \left| \frac{\partial (z - Z)}{\partial x} \right| \right) \le Mh,$$

where M is a generic positive constant depending on the second order seminorm of z (which is bounded on $\overline{\Omega}$ because we have assumed that the second order partial derivatives of z are continuous on $\overline{\Omega}$). Combining the last two inequalities and noting that a and b are continuous on $\overline{\Omega}$ we obtain

$$\left| G_{i,j} - \int_{\Omega_{i,j}} LZ(x,t)\, d\Omega \right| \le Mh \Delta x_i \Delta t_j. \tag{2.18}$$

For any positive integer m, we let $\|\cdot\|_\infty$ denote the sup norm on $\mathbb{R}^m$ defined by

$$\|v\|_\infty = \sup_{1 \le j \le m} |v_j| \qquad \forall v = (v_1, v_2, \ldots, v_m) \in \mathbb{R}^m. \tag{2.19}$$

https://doi.org/10.1017/S0334270000012510 Published online by Cambridge University Press

Also, for any $m \times m$ matrix K, we let

$$\|K\| = \sup_{v \in \mathbb{R}^m} \frac{\|Kv\|_\infty}{\|v\|_\infty}.$$

Obviously $\|\cdot\|$ is a norm on the space of all linear transformations from $\mathbb{R}^m$ to $\mathbb{R}^m$. For any j, since all the eigenvalues of K^j and $(P^j)^{-1}$ are in $(-1, 1)$, we have, for all $j = 0, 1, \dots, J-1$,

$$\|K^j\| < 1, \quad \|(P^j)^{-1}\| < 1. \tag{2.20}$$

Rewrite (2.17) in a matrix form as (2.16). Taking the sup norm on both sides and using (2.18) we have

$$\|e^{j+1}\|_\infty \le \|K^j\| \|e^j\|_\infty + Mh\|(P^j)^{-1}\| \|g^j\|_\infty,$$

where K^j and P^j are the same as those in (2.16) and

$$e^j = (e_1^j, e_2^j, \dots, e_l^j), \qquad g^j = (\Delta x_1, \Delta x_2, \dots, \Delta x_l)\Delta t_j.$$

Repeatedly using the above estimate and (2.20) we obtain

$$\begin{aligned}
\|e^{j+1}\|_\infty &\le \|K^j\| \left(\|K^{j-1}\| \|e^{j-1}\|_\infty + Mh\|(P^{j-1})^{-1}\| \|g^{j-1}\|_\infty\right) \\
&\quad + Mh\|(P^j)^{-1}\| \|g^j\|_\infty \\
&= \|K^j\| \|K^{j-1}\| \|e^{j-1}\|_\infty \\
&\quad + Mh\left(\|K^j\| \|(P^{j-1})^{-1}\| \|g^{j-1}\|_\infty + \|(P^j)^{-1}\| \|g^j\|_\infty\right) \\
&\le \|e^0\|_\infty \prod_{l=0}^{j} \|K^l\| + Mh \sum_{l=0}^{j} \left(\prod_{m=l+1}^{j} \|K^l\|\right) \|(P^l)^{-1}\| \|g^l\|_\infty \\
&\le Mh \sum_{l=1}^{j} \|g^l\|_\infty \\
&\le Mhj \sup_{1 \le l \le j} \|g^l\|_\infty \\
&\le Mh^2 j \sup_{1 \le l \le j} \Delta t_l \\
&\le Mh^2
\end{aligned}$$

because $\|e^0\|_\infty = 0$. Thus, from this it follows that

$$|e_i^j| \le \|e^j\|_\infty \le Mh^2 \tag{2.21}$$

for any i and j. Therefore, the method is of second order accuracy.

https://doi.org/10.1017/S0334270000012510 Published online by Cambridge University Press

3. Gradient calculations

Consider the following optimal control problem:
Minimize

$$M(u) = \int_{\Omega} \mathscr{L}(x, t, z, u)\, d\Omega \tag{3.1}$$

subject to

$$\frac{\partial^2 z}{\partial x \partial t} + a(x, t, u(x, t))\frac{\partial z}{\partial x} + b(x, t, u(x, t))\frac{\partial z}{\partial t} = f(x, t, u(x, t)), \tag{3.2}$$

for all $(x, t) \in \operatorname{int}(\Omega)$, where $\Omega = [0, 1] \times [0, 1]$, $\mathscr{L} : \Omega \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is given and $u : \Omega \to \mathbb{R}$ is the control function to be determined. We assume that the control is bounded, that is,

$$u_{\min} \leq u(x, t) \leq u_{\max}, \text{ for all } (x, t) \in \Omega. \tag{3.3}$$

Let $\mathscr{U}$ denote the set of all control functions satisfying (3.3). Note that the system dynamics (3.2) are a slight generalization of those considered in the previous section.

Boundary conditions of the form (2.2) may be imposed and it is possible for these to contain another control function and/or parameters which can be used in the optimization. We do not specify these here for the sake of brevity. Furthermore, we could impose certain types of canonical constraints on the problem at this stage, but we choose not to do so. Both of these complications are dealt with to some extend in the application example.

The task at this stage is to derive the gradient formula of the objective functional (3.1) with respect to the control u. Usually the gradients of the original (nondiscretized) problem are derived first, leading to a PDE in the costate of the problem which also arises when a maximum principle is applied to the problem. It is then possible to discretize both (3.2) and the costate equation, solve them and calculate the required gradient of the discretized problem.

We proceed differently here, firstly discretizing (3.2) and then using the concept of fast automatic differentiation (FAD) to derive the costate equation in an already discretized form. The advantages of this approach are twofold. We avoid the need to calculate the gradient of the original (nondiscretized) problem, the details of which may get quite complicated for a given problem, and the resulting gradients are guaranteed to be exact.

The idea of automatic differentiation and its applications to optimal control problems in particular has been described in [13]. Of the so-called forward and reverse forms of the method, we choose the latter because it is computationally far more

https://doi.org/10.1017/S0334270000012510 Published online by Cambridge University Press

efficient for the type of problem under consideration. The basis of the method can be described as follows. Consider the problem of finding the gradient of a scalar function $\Omega(\boldsymbol{u})$ with respect to $\boldsymbol{u} \in \mathbb{R}^r$. Assume that

$$\Omega(\boldsymbol{u}) = W(z, \boldsymbol{u}), \tag{3.4}$$

where $z \in \mathbb{R}^n$ and $\boldsymbol{u}$ are related by the following set of equations

$$\Phi(z, \boldsymbol{u}) = \boldsymbol{0}, \tag{3.5}$$

where $\Phi : \mathbb{R}^n \times \mathbb{R}^r \mapsto \mathbb{R}^n$. If we assume that the matrix $\Phi_z(z, \boldsymbol{u})$ is nonsingular, then the implicit function theorem guarantees that (3.5) defines z as a differentiable function of $\boldsymbol{u}$ and the required gradient exists. It can be calculated by first forming the Lagrangian

$$L(z, \boldsymbol{u}) = W(z, \boldsymbol{u}) + \boldsymbol{p}^{\top}\Phi(z, \boldsymbol{u}). \tag{3.6}$$

The vector $\boldsymbol{p}$ can then be determined from

$$L_z(z, \boldsymbol{u}) = W_z(z, \boldsymbol{u}) + \boldsymbol{p}^{\top}\Phi_z(z, \boldsymbol{u}) = \boldsymbol{0}^{\top}. \tag{3.7}$$

Finally, the required gradient is given by

$$\frac{d\Omega}{d\boldsymbol{u}} = L_u(z, \boldsymbol{u}) = W_u(z, \boldsymbol{u}) + \boldsymbol{p}^{\top}\Phi_u(z, \boldsymbol{u}). \tag{3.8}$$

Now returning to the original task, we follow exactly the same discretization procedure as in the previous section, except that we also define u_i^j to be the nodal values of the control function. The choice and exact structure of the control functions is arbitrary at this stage.

The discretized problem then ultimately only depends on the set of nodal control values $\{u_i^j\}$ and the cost reduces to a function of these only:

$$\bar{M}(\boldsymbol{u}) = W(z(\boldsymbol{u}), \boldsymbol{u}), \tag{3.9}$$

where z denotes the collection of $\{z_i^j\}$ and $\boldsymbol{u}$ represents the collection of $\{u_i^j\}$. (We represent the right hand side of (3.9) in this way, since it will, in general, also contain the dependent state variable.)

Regardless of what numerical integration scheme we now choose to employ, we can describe it by a general set of equations as follows:

$$z_i^j = F(i, j, Z_i^j, U_i^j), \quad i = 0, \dots, I, ; \; j = 0, \dots, J, \tag{3.10}$$

where the exact form of F depends on the particular integration scheme used and on the prescribed boundary conditions of the problem. Here, Z_i^j denotes the set of all z_l^m

https://doi.org/10.1017/S0334270000012510 Published online by Cambridge University Press

that appear in the equation for z_i^j. Similarly, U_i^j denotes the set of all u_l^m that appear in the equation for z_i^j.

Furthermore, let $D = \{(i, j) : 0 \le i \le I, 0 \le j \le J\}$ and define the sets

$$Q_{i,j} = \{(l, m) \in D : z_i^j \in Z_l^m\}, \tag{3.11}$$

$$K_{i,j} = \{(l, m) \in D : u_i^j \in U_l^m\}. \tag{3.12}$$

Then, applying (3.8), the required gradients are given by

$$\frac{\partial \bar{M}}{\partial u_i^j} = W_{u_i^j}(z, u) + \sum_{(l,m)\in K_{i,j}} p_l^m F_{u_i^j}(l, m, Z_l^m, U_l^m), \tag{3.13}$$

$0 \le i \le I, 0 \le j \le J$, where, according to (3.7), the nodal values of the costate p_i^j satisfy the relations

$$p_i^j = W_{z_i^j}(z, u) + \sum_{(l,m)\in Q_{i,j}} p_l^m F_{z_i^j}(l, m, Z_l^m, U_l^m), \tag{3.14}$$

for $0 \le i \le I, 0 \le j \le J$.

The same technique of automatic differentiation can be applied to find gradients in virtually any type of integration scheme and, in fact, in any type of multistep method.

4. Application example

Consider the problem of purifying a mixture of air and poison gas. The mixture is passed through a tubular filter containing an appropriate absorbent. Let $v(t)$ be the time dependent velocity of the mixture passing through the filter. When $v(t)$ is sufficiently large, the diffusion process is negligible and the quantity of the gas $Q(x, t)$ absorbed per unit volume of the absorbent as well as the concentration $C(x, t)$ of the gas in the pores of the absorbent satisfy the following equations (see, for example, [15] and [21, p. 175]).

$$-v(t)\frac{\partial C}{\partial x} = \frac{\partial Q}{\partial t}, \tag{4.1}$$

$$\frac{\partial Q}{\partial t} = \beta(C - \gamma Q), \tag{4.2}$$

with the boundary and initial conditions

$$C(0, t) = C_0, \tag{4.3}$$

$$Q(x, 0) = 0, \tag{4.4}$$

https://doi.org/10.1017/S0334270000012510 Published online by Cambridge University Press

where x denotes the longitudinal axis of the tube, C_0 is a constant, and β and γ are, respectively, the kinetic and Henry's coefficients which are positive constants. Assume here that x and t are scaled quantities so that the region of interest is $\Omega = [0, 1] \times [0, 1]$.

To eliminate the unknown function $Q(x, t)$, we differentiate (4.1) with respect to t

$$-v\frac{\partial^2 C}{\partial x \partial t} - \frac{\partial v}{\partial t}\frac{\partial C}{\partial x} = \frac{\partial^2 Q}{\partial t^2}.$$

Differentiating (4.2) with respect to t, substituting the result into the above equation and using (4.1) we obtain

$$-v\frac{\partial^2 C}{\partial x \partial t} - \frac{dv}{dt}\frac{\partial C}{\partial x} = \beta\frac{\partial C}{\partial t} - \beta\gamma\frac{\partial Q}{\partial t} = \beta\frac{\partial C}{\partial t} + \beta\gamma v\frac{\partial C}{\partial x}$$

or

$$\frac{\partial^2 C}{\partial x \partial t} + \left(\beta\gamma + \frac{1}{v}\frac{dv}{dt}\right)\frac{\partial C}{\partial x} + \frac{\beta}{v}\frac{\partial C}{\partial t} = 0. \tag{4.5}$$

For simplicity we assume that $\beta\gamma + \dot{v}/v > 0$ for all $t > 0$. This includes the cases of v being nondecreasing and of v decreasing at a moderate rate compared to its magnitude. The boundary condition for C is given in (4.3). Substituting (4.2) into (4.1), putting $t = 0$ and using (4.4) we have

$$-v_0 C_x(x, 0) = \beta C(x, 0),$$

where $v_0 = v(0)$. From this and (4.3) we obtain the initial condition for C

$$C(x, 0) = C_0 \exp\left(-\frac{\beta x}{v_0}\right). \tag{4.6}$$

Equation (4.5) and the boundary and initial conditions (4.3) and (4.5) form a linear second order hyperbolic problem describing the process of gas absorption. As mentioned in [21], similar equations arise in other problems such as the processes of drying with a current of air and heating a tube with a current of water.

Suppose that the nature of the tubular filter is such that it ceases to be effective once the concentration of the gas in the pores of the absorbent at the end of the tube is, say, 5% of C_0. Hence we impose the constraint

$$C(1, t) \le 0.05 C_0, \quad \text{for all } t \in [0, 1]. \tag{4.7}$$

Furthermore, we require the velocity of the mixture to always exceed a given minimum value. Hence we require that

$$v(t) \ge v_{\min}, \quad \text{for all } t \in [0, 1]. \tag{4.8}$$

https://doi.org/10.1017/S0334270000012510 Published online by Cambridge University Press

The optimal control problem then is to determine the optimum velocity profile v such that the volume of gas/air mix passing through the filter is maximised while the constraints (4.7) and (4.8) are satisfied. In other words, the problem in standard form is to minimize

$$M(v) = -\int_0^1 v(t)\,dt \tag{4.9}$$

subject to the system dynamics (4.5) with (4.3) and (4.6) and while not violating (4.7) and (4.8).

The control function in this problem is clearly $v(t)$, but the term $\dot{v}/v$ in (4.5) causes difficulties. A way around this problem is to introduce an additional state variable $y = \ln v$. Then $\dot{y} = \dot{v}/v$, $v = e^y$ and the system dynamics now consist of two equations

$$C_{xt} + (\beta\gamma + u)C_x + \beta e^{-y}C_t = 0, \tag{4.10}$$

$$\dot{y} = u, \tag{4.11}$$

where the function $u(t)$ is the new control function for the problem. We also need to specify an initial condition for (4.11), so put

$$y(0) = \xi, \tag{4.12}$$

where ξ is now also a variable in the optimization process.

The problem now is to minimize

$$\mathscr{M}(u, \xi) = -\int_0^1 e^{y(t)}\,dt \tag{4.13}$$

subject to the system dynamics (4.10) with (4.3) and (4.6), (4.11) with (4.12) while satisfying the constraints (4.7) and (4.8).

5. Numerical solution

In this section we discuss the numerical solution of the application example. We show how the problem is discretized, how we deal with the constraints and how the FAD approach discussed previously applies to this particular problem. Finally we present the numerical results obtained.

5.1. Discretization and Constraint Transcription. If we let z_i^j represent the value of $C(x, t)$ at the node (x_i, t_j), use the integration scheme (2.12) for (4.10) and a simple

https://doi.org/10.1017/S0334270000012510 Published online by Cambridge University Press

Euler step for (4.11), we obtain the following:

$$z_{i+1}^{j+1} = \frac{\left[\left(1 + C_i^j - E_i^j\right) z_i^{j+1} + \left(1 - C_i^j + E_i^j\right) z_{i+1}^j - \left(1 - C_i^j - E_i^j\right) z_i^j\right]}{1 + C_i^j + E_i^j}, \tag{5.1}$$

$$y_{j+1} = y_j + u_j \Delta t_j \tag{5.2}$$

for all $i = 0, 1, \ldots, I-1$ and $j = 0, 1, \ldots, J-1$, where

$$C_i^j = \frac{(\beta\gamma + u_j)\Delta t_j}{2}, \qquad E_i^j = \frac{\beta(e^{-y_j} - e^{-y_{j+1}})}{2u_j \Delta t_j}\Delta x_i.$$

We assume that the control is a piecewise constant function, so (5.2) actually gives the exact knot values of y.

The boundary and initial conditions become

$$z_0^j = C_0, \qquad 0 \le j \le J \tag{5.3}$$

$$z_i^0 = C_0 e^{-\beta x_i e^{-\xi}}, \quad 0 \le i \le I \tag{5.4}$$

$$y_0 = \xi. \tag{5.5}$$

Again, since we are assuming that u is piecewise constant, the objective function (4.9) can be shown to be

$$W_0(\boldsymbol{y}, \boldsymbol{u}) = \sum_{j=0}^{J-1} \frac{e^{y_j} - e^{y_{j+1}}}{u_j}, \tag{5.6}$$

where $\boldsymbol{y} = [y_0, \ldots, y_J]$ and $\boldsymbol{u} = [u_0, \ldots, u_{J-1}]$.

In discretized form, the constraints (4.7) and (4.8) are

$$z_I^j \le 0.05 C_0, \quad 0 \le j \le J, \tag{5.7}$$

$$e^{y_j} \ge v_{\min}, \quad 0 \le j \le J. \tag{5.8}$$

Refer to the discretized problem of minimizing (5.6) subject to (5.1)-(5.5), (5.7) and (5.8) as Problem (P).

We treat both (5.7) and (5.8) using a constraint transcription and penalty function approach developed in [18] and applied successfully to many similarly constrained problems. Consider first (5.7). For $j = 0$, we have $z_I^0 \le 0.05C_0$. By (5.4), this is equivalent to $e^{-\beta e^{-\xi}} \le 0.05 C_0$. Rearranging, we get an upper bound on the parameter ξ:

$$\xi \le -\ln\left(-\frac{\ln(0.05 C_0)}{\beta}\right). \tag{5.9}$$

https://doi.org/10.1017/S0334270000012510 Published online by Cambridge University Press

For the remaining j, we can write (5.4) as

$$0.05C_0 - z_l^j \geq 0, \qquad 1 \leq j \leq J$$

which is equivalent to

$$\sum_{j=1}^{J} \min\left\{0.05C_0 - z_l^j, 0\right\} = 0$$

or

$$\sum_{j=1}^{J} \min\left\{h_1(z_l^j), 0\right\} = 0,$$

where $h_1(z) = 0.05C_0 - z$. Note that $\min\{h, 0\}$ is a nonsmooth function. We approximate it by the smooth function

$$\mathscr{L}_\epsilon\{h\} = \begin{cases} h, & \text{if } h \leq -\epsilon; \\ -\dfrac{(h-\epsilon)^2}{4\epsilon}, & \text{if } -\epsilon < h < \epsilon; \\ 0, & \text{if } h \geq \epsilon; \end{cases}$$

and then treat the smoothed constraint as a penalty function by appending the term

$$W_1(z) = -\gamma_1 \sum_{j=1}^{J} \mathscr{L}_{\epsilon_1}\left\{h_1(z_l^j)\right\}$$

to the objective functional, where γ_1 is an appropriate weighting factor. Treating the constraint in this manner means that we need to construct an iterative algorithm where the value of ϵ_1 is successively decreased and the value of γ_1 is adjusted at each stage to insure constraint satisfaction. The details of the method and a convergence analysis are given in [18].

Now consider the velocity constraint (5.8). For $j = 0$, we have $e^{y_0} \geq v_{\min}$, that is, $e^{\xi} \geq v_{\min}$ or

$$\xi \geq \ln(v_{\min}). \tag{5.10}$$

For the remaining j we write (5.8) as

$$v_j - v_{\min} \geq 0, \qquad 1 \leq j \leq J$$

or

$$e^{y_j} - v_{\min} \geq 0, \qquad 1 \leq j \leq J.$$

Proceeding in exactly the same manner as above, we transcribe the constraint into the form

$$\sum_{j=1}^{J} \min\{e^{y_j} - v_{\min}, 0\},$$

https://doi.org/10.1017/S0334270000012510 Published online by Cambridge University Press

smooth it, and then append the penalty term

$$W_2(y) = -\gamma_2 \sum_{j=1}^{J} \mathscr{L}_{\epsilon_2} \left\{h_2(y_j)\right\}$$

to the objective functional, where ϵ_2 is the smoothing parameter, γ_2 is a penalty weighting factor and $h_2(y) = e^y - v_{\min}$.

By treating the constraints in this manner, we obtain the following approximate problem, referred to as Problem ($P_{\epsilon,\gamma}$):

Minimize the objective functional

$$\bar{M}(\boldsymbol{u}, \xi) = W(z(\boldsymbol{u}, \xi), \boldsymbol{y}(\boldsymbol{u}, \xi), \boldsymbol{u}, \xi) = W_0(\boldsymbol{y}, \boldsymbol{u}) + W_1(z) + W_2(\boldsymbol{y})$$

subject to the discretized dynamics (5.1) and (5.2) with the boundary and initial conditions (5.3), (5.4) and (5.5). We also have bounds on the control and on ξ:

$$-\beta\gamma \le u_{\min} \le u_j \le u_{\max}, \quad 0 \le j \le J-1 \tag{5.11}$$

$$\ln(v_{\min}) \le \xi \le -\ln\left(-\frac{\ln(0.05C_0)}{\beta}\right). \tag{5.12}$$

Recall that we assume $\beta\gamma + \dot{v}/v > 0$ in (4.5). Hence the reason for $-\beta\gamma \le u_{\min}$ in (5.11).

The smoothing approximation in the constraint transcription means that instead of the original Problem (P), we are now looking at solving a sequence of Problems ($P_{\epsilon,\gamma}$) where the ϵ and γ parameters are updated to insure a good approximation of the original problem and to achieve constraint satisfaction.

5.2. Applying Automatic Differentiation. Recall that we introduced the additional state y in Section 4. Consequently, we need to modify the application of automatic differentiation to the problem as it is discussed in Section 3. The interdependence between z, $\boldsymbol{y}$, $\boldsymbol{u}$ and ξ can be described with a number of sets like those defined by (3.11) and (3.12). For the sake of brevity, we only record the resulting costate equations and gradient formulae here.

Firstly, we require the partial derivatives of $W(z, \boldsymbol{y}, \boldsymbol{u}, \xi)$.

$$\begin{aligned} W_{z_i^j}(z, \boldsymbol{y}, \boldsymbol{u}, \xi) &= \underbrace{W_{0\,z_i^j}(\boldsymbol{y}, \boldsymbol{u})}_{=0} + W_{1\,z_i^j}(z) + \underbrace{W_{2\,z_i^j}(\boldsymbol{y})}_{=0} \\ &= \begin{cases} 0, & 0 \le i \le I-1, 0 \le j \le J; \\ 0, & i = I, 1 \le j \le J; \\ \gamma_1 \dfrac{\partial \mathscr{L}_{\epsilon_1}\{h_1(z_i^j)\}}{\partial h_1}, & i = I, 1 \le j \le J; \end{cases} \end{aligned} \tag{5.13}$$

https://doi.org/10.1017/S0334270000012510 Published online by Cambridge University Press

$$W_{y_j}(z, \boldsymbol{y}, \boldsymbol{u}, \xi) = W_{0\, y_j}(\boldsymbol{y}, \boldsymbol{u}) + \underbrace{W_{1\, y_j}(z)}_{=0} + W_{2\, y_j}(\boldsymbol{y})$$

$$= \begin{cases} \dfrac{e^{y_j}}{u_j}, & j = 0; \\ \dfrac{e^{y_j}}{u_j} - \dfrac{e^{y_j}}{u_{j-1}} - \gamma_2 e^{y_j} \dfrac{\partial \mathscr{L}_{\epsilon_2}\{h_2(y_j)\}}{\partial h_2}, & 1 \le j \le J-1; \\ -\dfrac{e^{y_j}}{u_{j-1}} - \gamma_2 e^{y_j} \dfrac{\partial \mathscr{L}_{\epsilon_2}\{h_2(y_j)\}}{\partial h_2}, & j = J; \end{cases} \tag{5.14}$$

$$W_{u_j}(z, \boldsymbol{y}, \boldsymbol{u}, \xi) = W_{0\, u_j}(\boldsymbol{y}, \boldsymbol{u}) + \underbrace{W_{1\, u_j}(z)}_{=0} + \underbrace{W_{2\, u_j}(\boldsymbol{y})}_{=0}$$

$$= \frac{e^{y_{j+1}} - e^{y_j}}{u_j^{\,2}}, \quad 0 \le j \le J-1; \tag{5.15}$$

$$W_{\xi}(z, \boldsymbol{y}, \boldsymbol{u}, \xi) = \underbrace{W_{0\, \xi}(\boldsymbol{y}, \boldsymbol{u})}_{=0} + \underbrace{W_{1\, \xi}(z)}_{=0} + \underbrace{W_{2\, \xi}(\boldsymbol{y})}_{=0} = 0, \tag{5.16}$$

where

$$\frac{\partial \mathscr{L}_{\epsilon}\{h\}}{\partial h} = \begin{cases} 1, & \text{if } h \le -\epsilon; \\ -\dfrac{(h-\epsilon)}{2\epsilon}, & \text{if } -\epsilon < h < \epsilon; \\ 0, & \text{if } h \ge \epsilon. \end{cases}$$

To simplify the notation, note that from (5.1), we may write

$$\begin{aligned} z_i^j &= \frac{\left(1 + C_{i-1}^{j-1} - E_{i-1}^{j-1}\right) z_{i-1}^j + \left(1 - C_{i-1}^{j-1} + E_{i-1}^{j-1}\right) z_i^{j-1}}{1 + C_{i-1}^{j-1} + E_{i-1}^{j-1}} \\ &\quad - \frac{\left(1 - C_{i-1}^{j-1} - E_{i-1}^{j-1}\right) z_{i-1}^{j-1}}{1 + C_{i-1}^{j-1} + E_{i-1}^{j-1}} \\ &= F(i, j), \quad \text{for } 1 \le i \le I,\ 1 \le j \le J. \end{aligned} \tag{5.17}$$

Furthermore,

$$\frac{\partial E_{i-1}^{j-1}}{\partial y_j} = \frac{\beta e^{-y_j}}{2u_{j-1}\Delta t_{j-1}} \Delta x_{i-1}, \text{ and } \frac{\partial E_{i-1}^{j}}{\partial y_j} = -\frac{\beta e^{-y_j}}{2u_j \Delta t_j} \Delta x_{i-1}. \tag{5.18}$$

As well as the costate $\boldsymbol{p}$ mentioned in Section 3, the occurrence of ys in this application leads to an additional costate q. The equations defining both of these are as follows:

$$p_i^j = W_{z_i^j}(z, \boldsymbol{y}, \boldsymbol{u}, \xi)$$

https://doi.org/10.1017/S0334270000012510 Published online by Cambridge University Press

$$
+\begin{cases}
-p_{i+1}^{j+1}\dfrac{\left(1-C_i^j-E_i^j\right)}{\left(1+C_i^j+E_i^j\right)}, & i=0,\ j=0;\\[2ex]
p_{i+1}^{j}\dfrac{\left(1+C_i^{j-1}-E_i^{j-1}\right)}{\left(1+C_i^{j-1}+E_i^{j-1}\right)}-p_{i+1}^{j+1}\dfrac{\left(1-C_i^j-E_i^j\right)}{\left(1+C_i^j+E_i^j\right)}, & i=0,\ 1\le j\le J-1;\\[2ex]
p_{i}^{j+1}\dfrac{\left(1-C_{i-1}^j+E_{i-1}^j\right)}{\left(1+C_{i-1}^j+E_{i-1}^j\right)}-p_{i+1}^{j+1}\dfrac{\left(1-C_i^j-E_i^j\right)}{\left(1+C_i^j+E_i^j\right)}, & 1\le i\le I-1,\ j=0;\\[2ex]
p_{i+1}^{j}\dfrac{\left(1+C_i^{j-1}-E_i^{j-1}\right)}{\left(1+C_i^{j-1}+E_i^{j-1}\right)}+p_{i}^{j+1}\dfrac{\left(1-C_{i-1}^j+E_{i-1}^j\right)}{\left(1+C_{i-1}^j+E_{i-1}^j\right)} & \\
\quad -p_{i+1}^{j+1}\dfrac{\left(1-C_i^j-E_i^j\right)}{\left(1+C_i^j+E_i^j\right)}, & 1\le i\le I-1,\ 1\le j\le J-1;\\[2ex]
p_{i+1}^{j}\dfrac{\left(1+C_i^{j-1}-E_i^{j-1}\right)}{\left(1+C_i^{j-1}+E_i^{j-1}\right)}, & 0\le i\le I-1,\ j=J;\\[2ex]
p_{i}^{j+1}\dfrac{\left(1-C_{i-1}^j+E_{i-1}^j\right)}{\left(1+C_{i-1}^j+E_{i-1}^j\right)}, & i=I,\ 0\le j\le J-1;\\[2ex]
0, & i=I,\ j=J;
\end{cases}
\tag{5.19}
$$

where $W_{z_i^j}(z, y, u, \xi)$ is given by (5.13).

$$
q_j = W_{y_j}(z, y, u, \xi) \tag{5.20}
$$

$$
+\begin{cases}
\displaystyle\sum_{i=1}^{I}\left[p_i^j\frac{\partial F(i,j)}{\partial E_{i-1}^{j-1}}\frac{\partial E_{i-1}^{j-1}}{\partial y_j}+p_i^{j+1}\frac{\partial F(i,j+1)}{\partial E_{i-1}^{j}}\frac{\partial E_{i-1}^{j}}{\partial y_j}\right]+q_{j+1}, & \\
\qquad 0\le j\le J-1; & \\[2ex]
\sum_{i=1}^{I} p_i^j\dfrac{\partial F(i,j)}{\partial E_{i-1}^{j-1}}\dfrac{\partial E_{i-1}^{j-1}}{\partial y_j}, & \\
\qquad j=J; &
\end{cases}
$$

where $W_{y_j}(z, y, u, \xi)$ is given by (5.14), $\dfrac{\partial F(i,j)}{\partial E_{i-1}^{j-1}}$ and $\dfrac{\partial F(i,j+1)}{\partial E_{i-1}^{j}}$ can be easily calculated from (5.17) and $\dfrac{\partial E_{i-1}^{j-1}}{\partial y_j}$ and $\dfrac{\partial E_{i-1}^{j}}{\partial y_j}$ are given by (5.18).

https://doi.org/10.1017/S0334270000012510 Published online by Cambridge University Press

Having calculated the costates, we can finally calculate the required gradients:

$$\frac{\partial \bar{M}(\boldsymbol{u}, \xi)}{\partial u_j} = W_{u_j}(z, \boldsymbol{y}, \boldsymbol{u}, \xi) + \sum_{i=1}^{I} p_i^{j+1} \left[\frac{\partial F(i, j+1)}{\partial C_{i-1}^j} \frac{\partial C_{i-1}^j}{\partial u_j} + \frac{\partial F(i, j+1)}{\partial E_{i-1}^j} \frac{\partial E_{i-1}^j}{\partial u_j} \right], \quad 0 \le j \le J-1; \tag{5.21}$$

$$\frac{\partial \bar{M}(\boldsymbol{u}, \xi)}{\partial \xi} = W_\xi(z, \boldsymbol{y}, \boldsymbol{u}, \xi) + \sum_{i=1}^{I} p_i^0 \beta x_i e^{-\xi - \beta x_i e^{-\xi}} + q_0, \tag{5.22}$$

where $\dfrac{\partial F(i, j+1)}{\partial C_{i-1}^j}$ and $\dfrac{\partial F(i, j+1)}{\partial E_{i-1}^j}$ can be calculated from (5.17), $W_{u_j}(z, \boldsymbol{y}, \boldsymbol{u}, \xi)$ is given by (5.15), $W_\xi(z, \boldsymbol{y}, \boldsymbol{u}, \xi)$ is given by (5.16), and

$$\frac{\partial C_{i-1}^j}{\partial u_j} = \frac{\Delta t_j}{2}, \text{ and } \frac{\partial E_{i-1}^j}{\partial u_j} = -\frac{\beta \left(e^{-y_j} - e^{-y_{j+1}}\right)}{2u_j^2 \Delta t_j} \Delta x_{i-1}$$

Using these gradients, the control problem can now be solved as a standard mathematical programming problem.

5.3. Numerical Results. We consider a problem which is characterized by the following data:

$$\beta = 50, \quad \gamma = 1, \quad C_0 = 1, \quad v_{\min} = 0.5.$$

We choose a uniform mesh node distribution with $I = 100$ and $J = 100$. The control u is assumed to be piecewise constant on the interval $[0, 1]$. Hence, in the discretized problem, we lump together each successive set of five u_js into one parameter, giving a total of 20 control parameters. This helps to reduce the size of the optimization problem. Constant bounds are placed on the control:

$$-5.0 \le u \le 5.0$$

and the bounds for ξ are determined by (5.12).

The parameters ϵ_1, ϵ_2, γ_1 and γ_2 are initialized to 10^{-2}, 10^{-2}, 1 and 1, respectively.

The results obtained are shown in Figures 1 and 2. Figure 1 shows the optimal velocity of the gas/air mixture and Figure 2 plots the corresponding pore concentration $C(x, t)$. Note how the velocity constraint and the constraint on $C(1, t)$ are both satisfied.

The objective function value is -0.734477, that is, a volume of 0.734477 units^3 of the mixture has passed through the filter.

https://doi.org/10.1017/S0334270000012510 Published online by Cambridge University Press

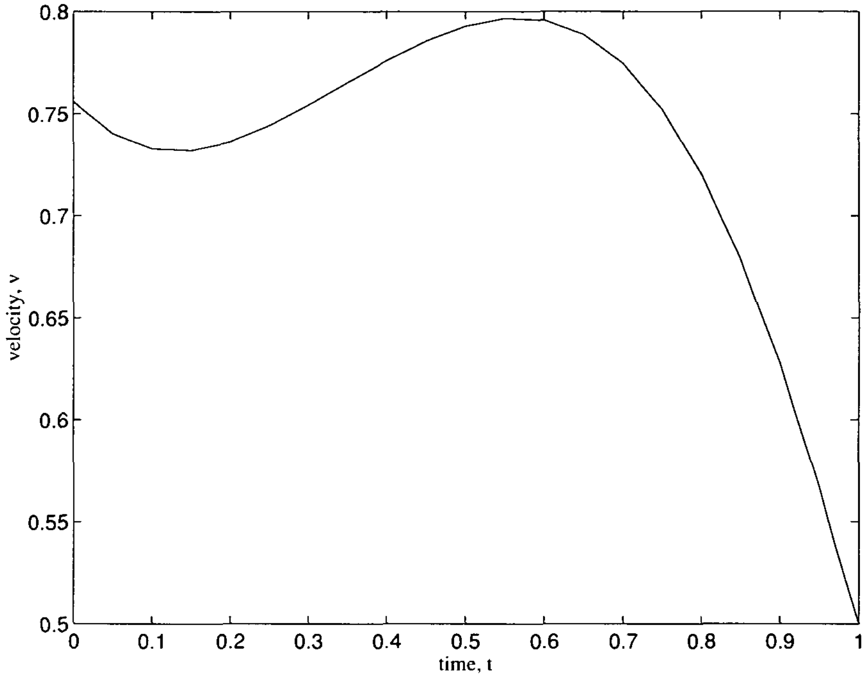

FIGURE 1. Optimal velocity of air/gas mixture.

6. Conclusion

We have presented a numerical solution method to solve a class of optimal control problems involving hyperbolic partial differential equations. The solution of the PDE is obtained using an explicit integration scheme which we showed to be unconditionally stable and of second order accuracy.

The gradient of the objective functional is then obtained by applying the technique of automatic differentiation. As demonstrated, this technique yields the exact discretized costate equations without needing to first calculate the costate equations of the nondiscretized problem.

We considered a practical problem as an application example and demonstrated how the method can be easily expanded to deal with some generalizations of the basic problem, in particular the addition of constraints.

Numerical results were presented to verify our theoretical results and to demonstrate the usefulness of the method.

https://doi.org/10.1017/S0334270000012510 Published online by Cambridge University Press

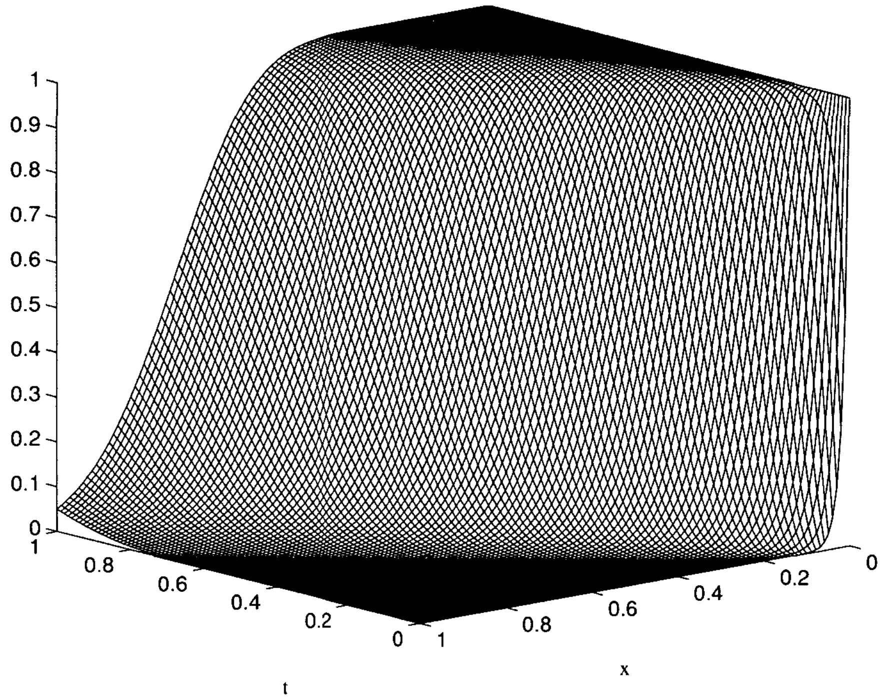

FIGURE 2. Pore Concentration corresponding to optimal velocity.

References

[1] N. U. Ahmed, *Optimization and identification of systems governed by evolution equations on Banach space* (Longman Scientific and Technical, Essex, England, 1988).

[2] N. U. Ahmed, *Semigroup theory with applications to systems and control* (Longman Scientific and Technical, Essex, England, 1991).

[3] N. U. Ahmed and K. L. Teo, *Optimal control of distributed parameter systems* (Elsevier North Holland, New York, 1981).

[4] H. T. Banks (ed.), *Control and estimation in distributed parameter systems*, Frontiers in Applied Mathematics (SIAM, Philadelphia, 1992).

[5] H. T. Banks and K. Kunisch, *Estimation techniques for distributed parameter systems* (Birkhäuser, Boston, 1989).

[6] A. Bensoussan, G. da Prato, M. C. Delfour and S. K. Mitter, *Representation and control of infinite dimensional systems*, Volume 1 (Birkhäuser, Boston, 1992).

[7] A. Bensoussan, G. da Prato, M. C. Delfour and S. K. Mitter, *Representation and control of infinite dimensional systems*, Volume 2 (Birkhäuser, Boston, 1992).

[8] C. Brenner and L. R. Scott, *The mathematical theory of finite element methods* (Springer-Verlag, New York, 1994).

[9] A. G. Butkovsky, *Distributed control systems* (American Elsevier, New York, 1969).

https://doi.org/10.1017/S0334270000012510 Published online by Cambridge University Press

[10] B. D. Craven, *Control and Optimization* (Chapman & Hall, London, 1995).

[11] R. F. Curtain and A. J. Pritchard, *Infinite dimensional linear systems theory*, Lecture Notes in Control and Information Sciences 8 (Springer-Verlag, Berlin, 1978).

[12] R. F. Curtain and H. J. Zwart, *An introduction to infinite dimensional linear systems theory* (Springer-Verlag, New York, 1995).

[13] Y. G. Evtushenko, "Fast automatic differentiation and related topics", submitted.

[14] A. Griewank and G. F. Corliss, *Automatic differentiation of algorithms: theory, implementation and applications* (SIAM, Philadelphia, 1991).

[15] D. Idczak, K. Kibalczyc and S. Walczak, "On an oprimization problem with cost of rapid variation of control", *J. Austral. Math. Soc. Ser. B* **36** (1994) 117–131.

[16] X. J. Li and J. M. Yong, *Optimal control theory for infinite dimensional systems* (Birkhäuser, Boston, 1995).

[17] J. L. Lions, *Optimal control of systems described by partial differential equations* (Springer-Verlag, Berlin, 1971).

[18] V. Rehbock, K. L. Teo and L. S. Jennings, "An exact penalty function approach to all-time-step constrained discrete-time optimal control problems", *Applied Mathematics and Computation* **49** (1992) 215–230.

[19] K. L. Teo, C. J. Goh and K. H. Wong, *A unified computational approach to optimal control problems* (Longman Scientific and Technical, Essex, England, 1991).

[20] K. L. Teo and Z. S. Wu, *Computational methods for optimizing distributed systems* (Academic Press, Orlando, 1984).

[21] A. N. Tikhonov and A. A. Samarski, *Equations of mathematical physics* (Moscow, 1978).

[22] F. Tröltzsch, *Optimality conditions for parabolic control problems and applications* (Teubner-Texte zur Mathematik, Band 62, Leipzig, 1982).

[23] S. G. Tzafestas (ed.), *Distributed parameter systems: theory and applications* (Pergamon Press, New York, 1982).

[24] R. Vichnevetsky and J. B. Bowles, *Fourier analysis of numerical approximations of hyperbolic equations* (SIAM, Philadelphia, 1982).

[25] P. K. C. Wang, "Control of distributed parameter systems", in *Advances in control systems, theory and applications* (ed. C. T. Leondes), Volume 1, (Academic Press, New York, 1964) 75–172.

[26] Z. S. Wu, "On the optimal control of a class of hyperbolic systems"Phd, University of New South Wales, 1981.

https://doi.org/10.1017/S0334270000012510 Published online by Cambridge University Press