APPLICATIONES MATHEMATICAE
27,2 (2000), pp. 239–250

S. H. FARAG (Minia)
M. H. FARAG (Ibri)

ON AN OPTIMAL CONTROL PROBLEM FOR A QUASILINEAR PARABOLIC EQUATION

Abstract. An optimal control problem governed by a quasilinear parabolic equation with additional constraints is investigated. The optimal control problem is converted to an optimization problem which is solved using a penalty function technique. The existence and uniqueness theorems are investigated. The derivation of formulae for the gradient of the modified function is explained by solving the adjoint problem.

1. Introduction. Optimal control problems for partial differential equations are currently of much interest. An extensive literature in this area is devoted to parabolic equations [1, 11, 12, 14, 15]. These problems describe the processes of hydro- and gasdynamics, heat physics, filtration, plasma physics and others [8, 9].

This paper presents an optimal control problem governed by a quasilinear parabolic equation with additional constraints. The optimal control problem is converted to an optimization problem which is solved using a penalty function technique. The existence and uniqueness theorems are investigated. The derivation of formulae for the gradient of the modified function is explained by solving the adjoint problem.

2. The optimal control problem. Let D be a bounded domain of the N-dimensional Euclidean space E_N, let l, T be given positive numbers, and let $\Omega = \{(x,t) : x \in D,\ t \in (0,T)\}$. Let $V = \{v : v = (v_1, \ldots, v_N) \in E_N,\ \|v\|_{E_N} \le R\}$, where $R > 0$ is a given number. We consider the heat

2000 *Mathematics Subject Classification*: 49J20, 49K20, 49M29, 49M30.
Key words and phrases: optimal control, parabolic equations, penalty function methods, existence theory.

exchange process described by the equation

$$(1)\qquad \frac{\partial u}{\partial t}-\frac{\partial}{\partial x}\left(\lambda(u,v)\frac{\partial u}{\partial x}\right)+B(u,v)\frac{\partial u}{\partial x}=f(x,t,u,v),\qquad (x,t)\in\Omega,$$

with initial and boundary conditions

$$(2)\qquad u(x,0)=\phi(x),\qquad x\in D,$$

$$(3)\qquad \lambda(u,v)\frac{\partial u}{\partial x}\bigg|_{x=0}=g_0(t),\qquad \lambda(u,v)\frac{\partial u}{\partial x}\bigg|_{x=l}=g_1(t),\qquad 0\le t\le T,$$

where $\phi(x)\in L_2(D)$, $g_0(t),g_1(t)\in L_2(0,T)$.

The function $f(x,t,u,v)\in L_2(\Omega)$ for every $(u,v)\in[r_1,r_2]\times E_N$ is measurable in $(x,t)\in\Omega$ and for all $(x,t)\in\Omega$ it is continuous in $(u,v)\in[r_1,r_2]\times E_N$. Furthermore, this function has a continuous derivative in u for each $(x,t)\in\Omega$, and for $(u,v)\in[r_1,r_2]\times E_N$, the derivative $\partial f(x,t,u,v)/\partial u$ is bounded. Moreover, the functions $\lambda(u,v),B(u,v)$ are continuous on $[r_1,r_2]\times E_N$, have continuous derivatives in u and for all $(u,v)\in[r_1,r_2]\times E_N$, the derivatives $\partial\lambda(u,v)/\partial u$, $\partial B(u,v)/\partial u$ are bounded, where r_1,r_2 are given numbers.

On the set V, under the conditions (1)–(3) and the additional restrictions

$$(4)\qquad \nu_0\le\lambda(u,v)\le\mu_0,\qquad \nu_0\le B(u,v)\le\mu_0,\qquad r_1\le u(x,t)\le r_2$$

it is required to minimize the function [14]

$$(5)\quad f_\alpha(u,v)=\int_0^T\{\beta_0[u(0,t)-f_0(t)]^2+\beta_1[u(l,t)-f_1(t)]^2\}\,dt+\alpha\|v-\omega\|^2_{E_N}$$

where $f_0(t),f_1(t)\in L_2(0,T)$ are given functions, $\alpha\ge0$, $\nu_0,\mu_0>0$, $\beta_0\ge0$, $\beta_1\ge0$, $\beta_0+\beta_1\ne0$, are given numbers, and $\omega=(\omega_1,\dots,\omega_N)\in E_N$ is a given vector.

Definition 1. The problem of finding a function $u=u(x,t)\in V_2^{1,0}(\Omega)$ from conditions (1)–(4) for a given $v\in V$ is called the *reduced problem.*

Definition 2. A *solution* of the reduced problem (1)–(4) corresponding to $v\in V$ is a function $u(x,t)\in V_2^{1,0}(\Omega)$ that satisfies the integral identity

$$(6)\qquad \int_0^l\int_0^T\left[u\frac{\partial\eta}{\partial t}-\lambda(u,v)\frac{\partial u}{\partial x}\frac{\partial\eta}{\partial x}-B(u,v)\frac{\partial u}{\partial x}\eta+\eta f(x,t,u,v)\right]dx\,dt$$
$$=-\int_0^l\phi(x)\eta(x,0)\,dx-\int_0^T\eta(0,t)g_0(t)\,dt+\int_0^T\eta(l,t)g_1(t)\,dt$$

for all $\eta=\eta(x,t)\in W_2^{1,1}(\Omega)$ with $\eta(x,T)=0$.

A solution of the reduced problem (1)–(4) explicitly depends on the control v, therefore we shall also use the notation $u=u(x,t;v)$.

From the assumptions and the results of [6] it follows that for every $v \in V$ a solution of the problem (1)–(4) exists, it is unique and $|u_x| \le C_0$ for all $(x,t) \in \Omega$ and $v \in V$, where C_0 is a certain constant.

The inequality constrained problem (1) through (5) is converted to a problem without inequality constraints by adding a penalty function [3, 16] to the objective (5) $\{OCP\}$, yielding the following function $\Phi(v) = \Phi_{\alpha,k}(v, A_k)$:

$$\Phi(v) = f_\alpha(u(v), v) + P_k(u(v), v) \tag{7}$$

where

$$Z(u,v) = [\max\{\nu_0 - \lambda(u,v); 0\}]^2 + [\max\{\lambda(u,v) - \mu_0; 0\}]^2,$$
$$Y(u,v) = [\max\{\nu_0 - B(u,v); 0\}]^2 + [\max\{B(u,v) - \mu_0; 0\}]^2,$$
$$Q^1(u) = [\max\{r_1 - u(x,t;v); 0\}]^2, \qquad Q^2(u) = [\max\{u(x,t;v) - r_2; 0\}]^2,$$
$$P_k(v) = A_k \int_0^l \int_0^T [Z(u,v) + Y(u,v) + Q^1(u) + Q^2(u)]\, dx\, dt$$

and A_k, $k = 1, 2, \ldots$, are positive numbers with $\lim_{k\to\infty} A_k = \infty$.

3. Well-posedness of the problem. Optimal control problems for solutions of differential equations do not always have a solution [13]. In this section, we will prove the existence and uniqueness of solution of problem (1)–(5).

LEMMA 3.1. *Under the above assumptions for every solution of the reduced problem* (1)–(5) *the following estimate is valid*:

$$\|\delta u\|_{V_2^{1,0}(\Omega)} \le C\left[\left\|\delta\lambda \frac{\partial u}{\partial x}\right\|^2_{L_2(\Omega)} + \left\|\delta B \frac{\partial u}{\partial x}\right\|^2_{L_2(\Omega)} + \|\delta f\|^2_{L_2(\Omega)}\right]^{1/2} \tag{8}$$

where $C \ge 0$ is a constant not depending on δv.

Proof. Set $\delta u(x,t) = u(x,t;v+\delta v) - u(x,t;v)$, $u = u(x,t;v)$, $u' = u(x,t;v+\delta v)$. From (6) it follows that

$$\begin{aligned}
&\int_0^l \int_0^T \left[-\delta u \frac{\partial \eta}{\partial t} + \lambda' \frac{\partial \delta u}{\partial x}\frac{\partial \eta}{\partial x} + \frac{\partial \lambda(u+\theta_1\delta u, v+\delta v)}{\partial u}\frac{\partial u}{\partial x}\frac{\partial \eta}{\partial x}\delta u + \delta\lambda \frac{\partial u}{\partial x}\frac{\partial \eta}{\partial x}\right] dx\, dt \\
&\quad + \int_0^l \int_0^T \left[B' \frac{\partial \delta u}{\partial x}\eta + \frac{\partial B(u+\theta_2\delta u, v+\delta v)}{\partial u}\frac{\partial u}{\partial x}\eta\delta u + \delta B \frac{\partial u}{\partial x}\eta\right] dx\, dt \\
&\quad - \int_0^l \int_0^T \left[\frac{\partial f(x,t,u+\theta_3\delta u, v+\delta v)}{\partial u}\delta u\eta + \delta f\eta\right] dx\, dt = 0
\end{aligned} \tag{9}$$

for all $\eta = \eta(x,t) \in W_2^{1,1}(\Omega)$ with $\eta(x,T) = 0$. Here $\theta_1, \theta_2, \theta_3 \in (0,1)$ are some numbers and

$$\delta f = f(x,t,u,v+\delta v) - f(x,t,u,v),$$
$$\lambda' = \lambda(u+\delta u, v+\delta v), \qquad \delta\lambda = \lambda(u,v+\delta v) - \lambda(u,v)$$
$$B' = B(u+\delta u, v+\delta v), \qquad \delta B = B(u,v+\delta v) - B(u,v).$$

Let $\eta_h(x,t) = h^{-1}\int_{t-h}^{t} \overline{\eta}(x,\tau)\,d\tau$, $0 < h < \tau$, where $\overline{\eta}(x,t) = \delta u(x,t)$ for $(x,t) \in \Omega_{t_1}$, zero for $t > t_1$ $(t_1 \le T-h)$, and $\Omega_{t_1} = D \times (0,t_1]$. In identity (9) put $\eta(x,t)$ instead of $\eta_h(x,t)$. Following the method of [7, pp. 166–168] we obtain

$$(10)\qquad \frac{1}{2}\int_D \delta u^2(x,t_1)\,dx$$
$$+ \int_{\Omega_{t_1}} \left[\lambda'\left(\frac{\partial \delta u}{\partial x}\right)^2 + \frac{\partial \lambda(u+\theta_1\delta u, v+\delta v)}{\partial u}\frac{\partial u}{\partial x}\frac{\partial \delta u}{\partial x}\delta u + \delta\lambda \frac{\partial u}{\partial x}\frac{\partial \delta u}{\partial x}\right] dx\,dt$$
$$+ \int_{\Omega_{t_1}} \left[B'\frac{\partial u}{\partial x}\delta u + \frac{\partial B(u+\theta_2\delta u, v+\delta v)}{\partial u}\frac{\partial u}{\partial x}(\delta u)^2 + \delta B\frac{\partial u}{\partial x}\delta u\right] dx\,dt$$
$$- \int_{\Omega_{t_1}} \left[\frac{\partial f(x,t,u+\theta_3\delta u, v+\delta v)}{\partial u}(\delta u)^2 + \delta f\,\delta u\right] dx\,dt = 0.$$

Hence, from the above assumptions and applying the Cauchy–Bunyakovskiĭ inequality, we have

$$(11)\qquad \frac{1}{2}\int_D \delta u^2(x,t_1)\,dx + \nu_0 \int_{\Omega_{t_1}} \left(\frac{\partial \delta u}{\partial x}\right)^2 dx\,dt$$
$$\le (C_3 + C_4)\int_{\Omega_{t_1}} \delta u^2\,dx\,dt$$
$$+ (C_1 + C_2)\Big(\int_{\Omega_{t_1}} \delta u^2\,dx\,dt\Big)^{1/2}\Big(\int_{\Omega_{t_1}} \left(\frac{\partial \delta u}{\partial x}\right)^2 dx\,dt\Big)^{1/2}$$
$$+ \Big(\int_{\Omega_{t_1}} \left(\delta B\frac{\partial u}{\partial x}\right)^2 dx\,dt\Big)^{1/2}\Big(\int_{\Omega_{t_1}} \delta u^2\,dx\,dt\Big)^{1/2}$$
$$+ \Big(\int_{\Omega_{t_1}} (\delta f)^2\,dx\,dt\Big)^{1/2}\Big(\int_{\Omega_{t_1}} \delta u^2\,dx\,dt\Big)^{1/2}$$
$$+ \Big(\int_{\Omega_{t_1}} \left(\delta\lambda\frac{\partial u}{\partial x}\right)^2 dx\,dt\Big)^{1/2}\Big(\int_{\Omega_{t_1}} \left(\frac{\partial \delta u}{\partial x}\right)^2 dx\,dt\Big)^{1/2}$$

where C_1, C_2, C_3 and C_4 are positive constants not depending on δv.

Take $\varepsilon_1 = 2C_1/\nu_0$, $\varepsilon_2 = 2C_2/\nu_0$ and apply the Cauchy inequality with ε ($|ab| \le \frac{\varepsilon}{2}|a|^2 + \frac{1}{2\varepsilon}|b|^2$) to the second and third summands on the right hand side of (11); multiplying both sides by two we obtain

$$\begin{aligned}
(12)\qquad & \|\delta u(x,t_1)\|^2_{L_2(D)} + \nu_0 \left\|\frac{\partial \delta u}{\partial x}\right\|^2_{L_2(\Omega_{t_1})} \\
&\le 2\biggl(\frac{C_2^2}{\nu_0} + C_3 + C_4 + \frac{C_1^2}{\nu_0}\biggr)\|\delta u\|^2_{L_2(\Omega_{t_1})} \\
&\quad + 2\biggl(\int_{\Omega_{t_1}} \biggl(\delta B\frac{\partial u}{\partial x}\biggr)^2 dx\,dt\biggr)^{1/2}\biggl(\int_{\Omega_{t_1}} \delta u^2\,dx\,dt\biggr)^{1/2} \\
&\quad + 2\biggl(\int_{\Omega_{t_1}} \delta f^2\,dx\,dt\biggr)^{1/2}\biggl(\int_{\Omega_{t_1}} \delta u^2\,dx\,dt\biggr)^{1/2} \\
&\quad + 2\biggl(\int_{\Omega_{t_1}} \biggl(\delta\lambda\frac{\partial u}{\partial x}\biggr)^2 dx\,dt\biggr)^{1/2}\biggl(\int_{\Omega_{t_1}} \biggl(\frac{\partial \delta u}{\partial x}\biggr)^2 dx\,dt\biggr)^{1/2}
\end{aligned}$$

Applying Cauchy's inequality with ε to the last three summands on the right side of (12) and taking $\varepsilon = \nu_0/2$ we obtain

$$\begin{aligned}
(13)\qquad & \|\delta u(x,t_1)\|^2_{L_2(D)} + \frac{\nu_0}{2}\left\|\frac{\partial \delta u}{\partial x}\right\|^2_{L_2(\Omega_{t_1})} \\
&\le 2\biggl(\frac{C_1^2 + C_2 + \nu_0^2}{\nu_0} + C_3 + C_4\biggr)\|\delta u\|^2_{L_2(\Omega_{t_1})} \\
&\quad + \frac{2}{\nu_0}\left\|\delta\lambda\frac{\partial u}{\partial x}\right\|^2_{L_2(\Omega_{t_1})} + \frac{2}{\nu_0}\left\|\delta B\frac{\partial u}{\partial x}\right\|^2_{L_2(\Omega_{t_1})} + \frac{2}{\nu_0}\|\delta f\|^2_{L_2(\Omega_{t_1})}.
\end{aligned}$$

Now we set

$$y(t_1) = \|\delta u(x,t_1)\|^2_{L_2(\Omega)},$$

$$M = \left\|\delta\lambda\frac{\partial u}{\partial x}\right\|^2_{L_2(\Omega_{t_1})} + \left\|\delta B\frac{\partial u}{\partial x}\right\|^2_{L_2(\Omega_{t_1})} + \|\delta f\|^2_{L_2(\Omega_{t_1})}.$$

Then inequality (13) yields the two inequalities

$$(14)\qquad y(t_1) \le C_5\int_0^{t_1} y(t)\,dt + \frac{2M}{\nu_0},$$

$$(15)\qquad \left\|\frac{\partial \delta u}{\partial x}\right\|^2_{L_2(\Omega_{t_1})} \le \frac{2C_5}{\nu_0}\|\delta u\|^2_{L_2(\Omega_{t_1})} + \frac{4M}{\nu_0^2},$$

where $C_5 = (2C_2^2 + 2C_1^2)/\nu_0 + 2C_3 + 2C_4 + 2\nu_0$ is a positive constant not depending on δv.

From the known estimate [6, pp. 166–167] it follows that

$$y(t_1) \le C_6 M \tag{16}$$

where C_6 is a positive constant not depending on δv. Consequently,

$$\max_{0\le t\le t_1} \|\delta u(x,t)\|_{L_2(D)} \le C_6 M^{1/2}. \tag{17}$$

Similarly we obtain

$$\left\| \frac{\partial \delta u}{\partial x} \right\|_{L_2(\Omega_{t_1})} \le C_7 M^{1/2} \tag{18}$$

where C_7 is a positive constant not depending on δv.

If we combine the estimates for δu and $\partial\delta u/\partial x$, then we obtain

$$\begin{aligned} \|\delta u\|_{V_2^{1,0}(\Omega_{t_1})} &= \max_{0\le t\le t_1} \|\delta u(x,t)\|_{L_2(D)} + \left\| \frac{\partial \delta u}{\partial x} \right\|_{L_2(\Omega_{t_1})} \\ &\le C_8 M^{1/2} \end{aligned} \tag{19}$$

where C_8 is a positive costant not depending on δv. Lemma 3.1 is proved.

Corollary 3.1. *Under the above assumptions the right side of estimate* (8) *converges to zero as* $\|\delta v\|_{E_N} \to 0$, *therefore* $\|\delta u\|_{V_2^{1,0}(\Omega)} \to 0$ *as* $\|\delta v\|_{E_N} \to 0$.

Hence from the trace theorem [10] we get

$$\|\delta u(0,t)\|_{L_2(0,T)} \to 0, \quad \|\delta u(l,t)\|_{L_2(0,T)} \to 0 \quad \text{as } \|\delta v\|_{E_N} \to 0. \tag{20}$$

Now we consider the function $J_0(u,v)$ of the form

$$J_0(u,v) = \beta_0 \int_0^T [u(0,t) - f_0(t)]^2 dt + \beta_1 \int_0^T [u(l,t) - f_1(t)]^2\, dt.$$

Lemma 3.2. *The function $J_0(u,v)$ is continuous on V.*

Proof. Let $\delta v = (\delta v_1, \ldots, \delta v_N)$ be an increment of control on an element $v \in V$ such that $v + \delta v \in V$. For the increment of $J_0(u,v)$ we have

$$\begin{aligned} \delta J_0(u,v) = 2\beta_0 &\int_0^T [u(0,t) - f_0(t)]\delta u(0,t)\, dt \\ &+ 2\beta_1 \int_0^T [u(l,t) - f_1(t)]\delta u(l,t)\, dt \\ &+ \beta_0 \int_0^T [\delta u(0,t)]^2\, dt + \beta_1 \int_0^T [\delta u(l,t)]^2\, dt. \end{aligned} \tag{21}$$

Applying the Cauchy–Bunyakovskiĭ inequality, we obtain

$$
\begin{aligned}
(22)\qquad |\delta J_0(u,v)| \le{}& 2\beta_0\|u(0,t)-f_0(t)\|_{L_2(0,T)}\|\delta u(0,t)\|_{L_2(0,T)}\\
&+2\beta_1\|u(l,t)-f_1(t)\|_{L_2(0,T)}\|\delta u(l,t)\|_{L_2(0,T)}\\
&+\beta_0\|\delta u(0,t)\|^2_{L_2(0,T)}+\beta_1\|\delta u(l,t)\|^2_{L_2(0,T)}.
\end{aligned}
$$

An application of Corollary 3.1 completes the proof.

THEOREM 3.1. *For any $\alpha \ge 0$ problem* (1)–(5) *has at least one solution.*

Proof. The set V is closed and bounded in E_N. Since $J_0(u,v)$ is continuous on V by Lemma 3.2, so is

$$J_\alpha(u,v)=J_0(u,v)+\alpha\|v-\omega\|^2_{E_N}.$$

Then from the Weierstrass theorem [5] it follows that problem (1)–(5) has at least one solution.

THEOREM 3.2. *For $\alpha>0$ and almost all $\omega\in E_N$ problem* (1)–(5) *has a unique solution.*

Proof. The functions $J_0(u,v)$ and $J_\alpha(u,v)$, $\alpha>0$, are continuous on V. Moreover, since E_N is a uniformly convex space, a theorem of [4] yields the existence of a dense subset K of E_N such that for any $\omega\in K$ and $\alpha>0$ problem (1)–(5) has a unique solution. Consequently, for almost all $\omega\in E_N$ and $\omega>0$ problem (1)–(5) has a unique solution.

4. Adjoint problem and gradient formulae

4.1. *The adjoint problem.* We illustrate the adjoint problem for the system (1)–(3). The Lagrangian function $L(x,t,u,v,\Theta)$ for the optimal control problem is defined as

$$
\begin{aligned}
(23)\qquad & L(x,t,u,v,\Theta)\\
&=\beta_0\int_0^T[u(0,t)-f_0(t)]^2\,dt+\beta_1\int_0^T[u(l,t)-f_1(t)]^2\,dt\\
&\quad+\alpha\|v-\omega\|^2_{E_N}+A_k\int_0^l\int_0^T[Z(u,v)+Y(u,v)+Q^1(u)+Q^2(u)]\,dx\,dt\\
&\quad+\int_0^l\int_0^T\Theta\left[\frac{\partial u}{\partial t}-\frac{\partial}{\partial x}\left(\lambda(u,v)\frac{\partial u}{\partial x}\right)+B(u,v)\frac{\partial u}{\partial x}-f(x,t,u,v)\right]dx\,dt.
\end{aligned}
$$

The first variation of the Lagrangian is

$$
\begin{aligned}
(24)\qquad & \delta L(x,t,u,v,\Theta) \\
&= 2\beta_0 \int_0^T [u(0,t)-f_0(t)]\delta u(0,t)\,dt + 2\beta_1 \int_0^T [u(l,t)-f_1(t)]\delta u(l,t)\,dt \\
&\quad + \beta_0 \int_0^T [\delta u(0,t)]^2\,dt + \beta_1 \int_0^T [\delta u(l,t)]^2\,dt + 2\alpha\langle v-\omega, \delta v\rangle_{E_N} + \alpha\|\delta v\|^2_{E_N} \\
&\quad + A_k \int_0^l\int_0^T \left[\frac{\partial Z(u,v)}{\partial u} + \frac{\partial Y(u,v)}{\partial v} + \frac{\partial Q^1(u)}{\partial u} + \frac{\partial Q^2(u)}{\partial u}\right]\delta u(x,t)\,dx\,dt \\
&\quad + \int_0^l\int_0^T \Theta\biggl[\frac{\partial \delta u}{\partial t} - \frac{\partial}{\partial x}\left(\lambda'\frac{\partial \delta u}{\partial x}\right) - \frac{\partial}{\partial x}\left(\frac{\partial \lambda}{\partial u}\frac{\partial u}{\partial x}\delta u\right) - \frac{\partial}{\partial x}\left(\lambda''\frac{\partial u}{\partial x}\right) \\
&\quad + B(u,v)\frac{\partial \delta u}{\partial x} + \frac{\partial B}{\partial u}\frac{\partial u}{\partial x}\delta u + \{f(x,t,u+\delta u, v+\delta v) - f(x,t,u,v)\}\biggr]\,dx\,dt
\end{aligned}
$$

where $\lambda' = \lambda(u+\delta u, v+\delta v)$, $\lambda'' = \lambda(u+\delta u, v)$.

Integrating (24) by parts we obtain

$$
\begin{aligned}
(25)\qquad & \delta L(x,t,u,v,\Theta) \\
&= 2\beta_0 \int_0^T [u(0,t)-f_0(t)]\delta u(0,t)\,dt + 2\beta_1 \int_0^T [u(l,t)-f_1(t)]\delta u(l,t)\,dt \\
&\quad + \beta_0 \int_0^T [\delta u(0,t)]^2\,dt + \beta_1 \int_0^T [\delta u(l,t)]^2\,dt + 2\alpha\langle v-\omega, \delta v\rangle_{E_N} + \alpha\|\delta v\|^2_{E_N} \\
&\quad + A_k \int_0^l\int_0^T \left[\frac{\partial Z(u,v)}{\partial u} + \frac{\partial Y(u,v)}{\partial v} + \frac{\partial Q^1(u)}{\partial u} + \frac{\partial Q^2(u)}{\partial u}\right]\delta u(x,t)\,dx\,dt \\
&\quad + \int_0^l\int_0^T \biggl[-\frac{\partial \Theta}{\partial t} - \frac{\partial}{\partial x}\left(\lambda'\frac{\partial \Theta}{\partial x}\right) + \frac{\partial \lambda}{\partial u}\frac{\partial u}{\partial x}\frac{\partial \Theta}{\partial x} \\
&\quad + \left(\frac{\partial B}{\partial u}\frac{\partial u}{\partial x}\Theta + \frac{\partial (B\Theta)}{\partial x}\right)\biggr]\delta u(x,t)\,dx\,dt \\
&\quad + \int_0^l\int_0^T \frac{\partial f}{\partial u}\Theta\delta u(x,t)\,dx\,dt + \int_0^l (\Theta\delta u)|_{t=T}\,dx + \int_0^T \left(\lambda'\frac{\partial \Theta}{\partial x}\delta u\right)\bigg|_{x=l}\,dt \\
&\quad + \int_0^T \left(\lambda'\frac{\partial \Theta}{\partial x}\delta u\right)\bigg|_{x=0}\,dt + \int_0^T (B\Theta\delta u)|_{x=l}\,dt + \int_0^T (B\Theta\delta u)|_{x=0}\,dt.
\end{aligned}
$$

Setting the variation in the Lagrangian equal to zero (the first order necessary condition for minimizing $L(x,t,u,v,\Theta)$) implies, since (25) must hold for any $\delta u(x,t)$ [11], that we obtain the adjoint problem:

$$\Theta_t + (\lambda(u,v)\Theta_x)_x - \lambda_u(u,v)\Theta_x u_x - [B_u u_x \Theta + (B\Theta)_x] - f_u\Theta = A_k[Z_u(u,v) + Y_u(u,v) + Q_u^2 + Q_u^1], \quad (x,t) \in \Omega, \tag{26}$$

$$\Theta(x,T) = 0, \quad x \in D, \tag{27}$$

$$(\lambda\Theta_x + B\Theta)|_{x=0} = 2\beta_0[u(0,t) - f_0(t)],$$

$$(\lambda\Theta_x + B\Theta)|_{x=l} = -2\beta_1[u(l,t) - f_1(t)], \quad t \in [0,T], \tag{28}$$

where $u = u(x,t)$ is the solution of problem (1)–(3) corresponding to $v \in V$.

Definition 3. A *solution* of the adjoint problem (26)–(28) corresponding to $v \in V$ is a function $\Theta(x,t) \in V_2^{1,0}(\Omega)$ such that the following integral identity is satisfied:

$$\begin{aligned}\int_0^l\int_0^T [\Theta\gamma_t + \lambda(u,v)\Theta_x\gamma_x + \lambda_u(u,v)\Theta_x u_x\gamma]\,dx\,dt \\ + \int_0^l\int_0^T [B_u u_x\Theta + (B\Theta)_x + f_u(x,t,u,v)\Theta]\gamma(x,t)\,dx\,dt \\ = -A_k\int_0^l\int_0^T [Z_u(u,v) + Y_u(u,v) + Q_u^2 + Q_u^1]\gamma(x,t)\,dx\,dt\end{aligned} \tag{29}$$

for all $\gamma = \gamma(x,t) \in W_2^{1,1}(\Omega)$ with $\gamma(x,0) = 0$.

From the above assumptions and the results of [7] it follows that for every $v \in V$ a solution of the adjoint problem (26)–(28) exists, it is unique and $|\Theta_x| \le C_9$ for almost all $(x,t) \in \Omega$ and all $v \in V$, where C_9 is a certain constant.

4.2. *Gradient formulae for $\Phi(v)$.* Sufficient differentiability conditions for $\Phi(v)$ and its gradient formulae will be obtained by defining the Hamiltonian function [2] $H(u,\Theta,v)$ as

$$\begin{aligned}H(u,\Theta,v) \equiv -\int_0^l\int_0^T [\lambda(u,v)\Theta_x u_x + B(u,v)u_x\Theta - f(x,t,u,v)\Theta \\ + A_k\{Z(u,v) + Y(u,v)\}]\,dx\,dt - \alpha\|v-\omega\|_{E_N}^2.\end{aligned} \tag{30}$$

Theorem 4.1. *Assume that*:

(i) *The functions $\lambda(u,v), B(u,v), f(x,t,u,v)$ satisfy the Lipschitz condition for v.*

(ii) *The first derivatives of* $\lambda(u,v), B(u,v), f(x,t,u,v)$ *with respect to* v *are continuous functions and for any* $v \in V$ *such that* $\|v\|_{E_N} \le R$, *the functions* $\lambda_v(u,v), B_v(u,v), f_v(x,t,u,v)$ *belong to* $L_\infty(\Omega)$.

(iii) *The operators*

$$\int_0^l\int_0^T \lambda_v(u,v)\,dx\,dt, \quad \int_0^l\int_0^T B_v(u,v)\,dx\,dt \quad \text{and} \quad \int_0^l\int_0^T f_v(x,t,u,v)\,dx\,dt$$

are bounded in E_N.

Then the function $\Phi(v)$ *is differentiable and its gradient is*

$$\frac{\partial\Phi(v)}{\partial v} = -\frac{\partial H}{\partial v} \equiv \Big(-\frac{\partial H}{\partial v_1},\ldots,-\frac{\partial H}{\partial v_N}\Big). \tag{31}$$

Proof. Suppose that $v \equiv (v_1,\ldots,v_N)$, $\delta v \equiv (\delta v_1,\ldots,\delta v_N)$, $\delta v \in E_N$, $v+\delta v \in V$ and set $\delta u \equiv u(x,t;v+\delta v) - u(x,t;v)$. The increment of $\Phi(v)$ can be expressed as

$$\begin{aligned} \delta\Phi(v) &= \Phi(v+\delta v) - \Phi(v) \\ &= 2\beta_0 \int_0^T [u(0,t) - f_0(t)]\delta u(0,t)\,dt + 2\beta_1 \int_0^T [u(l,t) - f_1(t)]\delta u(l,t)\,dt \\ &\quad + A_k \int_0^l\int_0^T [Z_u(u,v) + Y_u(u,v) + Q_u^1(u) + Q_u^2(u)]\delta u(x,t)\,dx\,dt \\ &\quad + A_k \int_0^l\int_0^T [Z(u,v+\delta v) - Z(u,v) + Y(u,v+\delta v) - Y(u,v)]\,dx\,dt \\ &\quad + 2\alpha\langle v-\omega,\delta v\rangle_{E_N} + R_1(\delta v) \end{aligned} \tag{32}$$

where

$$R_1(\delta v) = \beta_0 \int_0^T [\delta u(0,t)]^2\,dt + \beta_1 \int_0^T [\delta u(l,t)]^2\,dt + \alpha\|\delta v\|_{E_N}^2. \tag{33}$$

Using the estimate (8), we get the inequality $|R_1(\delta v)| \le C_{10}\|\delta v\|_{E_N}$ where C_{10} is a constant not depending on δv.

If we put $\gamma = \delta u(x,t)$ in (29) and $\eta = \Theta(x,t)$ in (9) and subtract the resulting relations, we obtain

$$\begin{aligned} & 2\beta_0 \int_0^T [u(0,t) - f_0(t)]\delta u(0,t)\,dt + 2\beta_1 \int_0^T [u(l,t) - f_1(t)]\delta u(l,t)\,dt \\ &\quad + A_k \int_0^l\int_0^T [Z_u(u,v) + Y_u(u,v) + Q_u^1(u) + Q_u^2(u)]\delta u(x,t)\,dx\,dt \end{aligned} \tag{34}$$

$$= \int_0^l \int_0^T [\delta\lambda u_x \Theta_x + \delta B u_x \Theta - \delta f \Theta]\, dx\, dt + R_2(\delta v)$$

where

(35) $R_2(\delta v)$

$$= \int_0^l \int_0^T \left\{ \lambda' \frac{\partial \delta u}{\partial x}\frac{\partial \Theta}{\partial x} + \left[\frac{\partial \lambda(u+\theta_1 \delta u, v+\delta v)}{\partial u} - \frac{\partial \lambda(u,v)}{\partial u} \right] \frac{\partial u}{\partial x}\frac{\partial \Theta}{\partial x}\delta u \right\} dx\, dt$$
$$+ \int_0^l \int_0^T \left\{ B' \Theta \frac{\partial \delta u}{\partial x} + \left[\frac{\partial B(u+\theta_2 \delta u, v+\delta v)}{\partial u} - \frac{\partial B(u,v)}{\partial u} \right] \Theta \frac{\partial u}{\partial x}\delta u \right\} dx\, dt$$
$$+ \int_0^l \int_0^T \left[\frac{\partial f(x,t,u+\theta_3 \delta u, v+\delta v)}{\partial u} - \frac{\partial f(x,t,u,v)}{\partial u} \right] \delta u(x,t)\Theta(x,t)\, dx\, dt$$

and $\theta_i \in (0,1)$, $i = 1,2,3$.

By assumption (i), $R_2(\delta v)$ is estimated as $|R_2(\delta v)| \le C_{11}\|\delta v\|_{E_N}$, where C_{11} is a constant independent of δv. Using the above assumptions, we can estimate

$$\begin{aligned}
Z(u, v+\delta v) - Z(u,v) &= \langle Z_v(u,v), \delta v\rangle_{E_N} + O(\|\delta v\|_{E_N}),\\
Y(u, v+\delta v) - Y(u,v) &= \langle Y_v(u,v), \delta v\rangle_{E_N} + O(\|\delta v\|_{E_N}),\\
\lambda(u, v+\delta v) - \lambda(u,v) &= \langle \lambda_v(u,v), \delta v\rangle_{E_N} + O(\|\delta v\|_{E_N}),\\
B(u, v+\delta v) - B(u,v) &= \langle B_v(u,v), \delta v\rangle_{E_N} + O(\|\delta v\|_{E_N}),\\
f(x,t,u, v+\delta v) - f(x,t,u,v) &= \langle f_v(x,t,u,v), \delta v\rangle_{E_N} + O(\|\delta v\|_{E_N}).
\end{aligned}$$

By substituting the last five expansions in (32) and (34), we obtain

$$\begin{aligned}
(36)\qquad \delta\Phi(v) = \int_0^l \int_0^T & \langle \lambda_v(u,v) u_x \Theta_x - \{B_v(u,v)u_x - f_v(x,t,u,v)\}\Theta\\
&+ A_k\{Z_v(u,v) + Y_v(u,v)\}, \delta v\rangle_{E_N}\, dx\, dt\\
&+ 2\alpha\langle v - \omega, \delta v\rangle_{E_N} + R_3(\delta v)
\end{aligned}$$

where $R_3(\delta v) = R_1(\delta v) + R_2(\delta v) + O(\|\delta v\|_{E_N})$.

From the formula for $R_3(\delta v)$, we have

$$(37)\qquad |R_3(\delta v)| \le C_{12}\|\delta v\|_{E_N}$$

where C_{12} is a constant independent of δv.

From (36), (37), using the function $H(u,\Theta,v)$ we have

$$(38)\qquad \delta\Phi(v) = \left\langle -\frac{\partial H(u,\Theta,v)}{\partial v}, \delta v \right\rangle_{E_N} + O(\|\delta v\|_{E_N}),$$

which shows the differentiability of $\Phi(v)$ and also gives the gradient formulae for $\Phi(v)$. Theorem 4.1 is proved.

References

[1] M. Bergounioux and F. Tröltzsch, *Optimality conditions and generalized bang-bang principle for a state constrained semilinear parabolic problem*, Numer. Funct. Anal. Optim. 17 (1996), 517–536.

[2] P. Enidr, *Optimal Control and Calculus of Variations*, Oxford Sci. Publ., London, 1993.

[3] M. H. Farag, *Application of the exterior penalty method for solving constrained optimal control problem*, Math. Phys. Soc. Egypt, 1995.

[4] M. Goebel, *On existence of optimal control*, Math. Nachr. 93 (1979), 67–73.

[5] W. Krabs, *Optimization and Approximation*, Wiley, New York, 1979.

[6] O. A. Ladyzhenskaya, V. A. Solonnikov and N. N. Ural'tseva, *Linear and Quasilinear Parabolic Equations*, Nauka, Moscow, 1976 (in Russian).

[7] O. A. Ladyzhenskaya, *Boundary Value Problems of Mathematical Physics*, Nauka, Moscow, 1973 (in Russian).

[8] J.-L. Lions, *Optimal Control by Systems Described by Partial Differential Equations*, Mir, Moscow, 1972 (in Russian).

[9] K. A. Lourie, *Optimal Control in Problems of Mathematical Physics*, Nauka, Moscow, 1975 (in Russian).

[10] V. P. Mikhailov, *Partial Differential Equations*, Nauka, Moscow, 1983 (in Russian).

[11] J. P. Raymond, *Nonlinear boundary control semilinear parabolic equations with pointwise state constraints*, Discrete Contin. Dynam. Systems 3 (1997), 341–370.

[12] J. P. Raymond and F. Tröltzsch, *Second order sufficient optimality conditions for nonlinear parabolic control problems with constraints*, preprint, Fak. Math., Tech. Univ. Chemnitz, 1998.

[13] A. N. Tikhonov and N. Ya. Arsenin, *Methods for the Solution of Ill-Posed Problems*, Nauka, Moscow, 1974 (in Russian).

[14] F. Tröltzsch, *On the Lagrange–Newton–SQP method for the optimal control for semilinear parabolic equations*, preprint, Fak. Math., Tech. Univ. Chemnitz, 1998.

[15] T. Tsachev, *Optimal control of linear parabolic equation: The constrained right-hand side as control function*, Numer. Funct. Anal. Optim. 13 (1992), 369–380.

[16] A.-Q. Xing, *The exact penalty function method in constrained optimal control problems*, J. Math. Anal. Appl. 186 (1994), 514–522.

S. H. Farag
Mathematics Department
Faculty of Science,
Minia University
Minia, Egypt

M. H. Farag
Mathematics Department
Faculty of Education
P.O. Box 14, Ibri 516, Sultanate of Oman
E-mail: farag5358@yahoo.com

Received on 29.4.1999;
revised version on 11.10.1999