Journal of Convex Analysis
Volume 9 (2002), No. 2, 543–561

Sensitivity Analaysis for Parametric Optimal Control of Semilinear Parabolic Equations

Kazimierz Malanowski
Systems Research Institute, Polish Academy of Sciences,
ul. Newelska 6, 01-447 Warszawa, Poland
kmalan@ibspan.waw.pl

Received November 8, 2000
Revised manuscript received April 4, 2001

Parametric optimal control problems for semilinear parabolic equations are considered. Using recent Lipschitz stability results for solutions of such problems, it is shown that, under standard coercivity conditions, the solutions are Bouligand differentiable (in L^p, $p < \infty$) functions of the parameter. The differentials are characterized as the solutions of accessory linear-quadratic problems. A uniform second order expansion of the optimal value function is obtained, as a corollary.

Keywords: Parametric optimal control, semilinear parabolic equations, control constraints, Bouligand differentiability of the solutions

1991 Mathematics Subject Classification: 49K40, 49K20, 49K30

1. Introduction

Stability and sensitivity analysis for parametric mathematical programs in finite dimension has been well developed and the results are fairly complete (see e.g., [3]). Such an analysis for parametric optimal control problems is less advanced. Here the difficulties are connected with infinite dimensionality of the problems. In particular, the so called *two-norm discrepancy* is typical for nonlinear control problems. Namely, as a rule, the Lagrangians of such problems are twice differentiable in a *stronger* norm (of L^∞ type), whereas coercivity conditions are satisfied in a weaker norm (of L^2 type), in which the Lagrangian is not twice differentiable. This phenomenon creates difficulties in stability analysis. However, in case of control constrained problems, these difficulties can be overcome using the structure of optimality conditions. For these problems, Lipschitz stability in stronger norm of L^∞ type has been established both for ODEs (see e.g., [7]) and PDEs (see [11, 12]). These stability results will be the starting point for our sensitivity analysis. It is crucial that the stability holds in L^∞, since it allows to consider general nonlinear control problems, subject to a broad class of perturbations. If only L^2 stability is used, one has to impose some restrictions on the problem (linear-quadratic with respect to control) and on the class of perturbations (see [2, 4]).

In sensitivity analysis of optimal control problems mostly the concept of *directional* differentiability of the solutions has been exploited. This refers in particular to PDEs (see e.g., [3, 20]). In parametric mathematical programs a stronger concept of differentiability, the so called *Bouligand* or B-differentiability has been used (see [6, 16, 17, 18, 19]). Let us recall the definition of Bouligand differentiability (see [6, 16, 18]).

ISSN 0944-6532 / $ 2.50

Definition 1.1. A function ϕ, from an open set G of a normed linear space H into another normed linear space X, is called B-differentiable (or Bouligand differentiable) at a point $h_0 \in G$ if there exists a positively homogeneous map $D_h\phi(h_0) : G \to X$, called B-derivative, such that

$$\phi(h_0 + \Delta h) = \phi(h_0) + D_h\phi(h_0)\Delta h + o(\|\Delta h\|_H). \tag{1.1}$$

$\Diamond$

Clearly, if $D_h\phi(h_0)$ is linear, it becomes the Fréchet derivative. Accordingly, the Bouligand derivative is sometimes called the *directional Fréchet* derivative [3]. Note that B-differentiability is a *stronger* property than conical differentiability, exploited e.g., in [13], where the remainder term in (1.1) is uniform on *compact* subsets of increments Δh (see Remark (2), p.142 in [13]).

To the knowledge of the author, the only B-differentiability result for solutions to parametric optimal control problems was obtained in [10] for nonlinear ODEs. The purpose of this paper is to extend the B-sensitivity result of [10] to systems described by PDEs. We will study an optimal control problem for a semilinear parabolic equation. Stability results for this problem were derived in [11]. As in [10], we use abstract results of Dontchev [6], which allow to deduce differentiability properties of the solutions to general nonlinear parametric optimization problems from the same properties of the solutions to linear-quadratic accessory problems (see Theorem 3.2). Sensitivity analysis for the accessory problem is performed in two steps. First, we prove directional differentiability of the solutions and characterize the differentials. Then, using this characterization, we obtain estimates, which show that the differentials are actually Bouligand.

In the principal Theorem 5.1 we show B-differentiability in L^p, $p < \infty$, of the solutions to our initial nonlinear problem and we characterize the B-differential as the solution to an auxiliary linear-quadratic optimal control problem. As a corollary, we obtain a *uniform* second order expansion of the optimal value function. Throughout the paper, we often refer to [11] and adapt the notation used therein.

Note that the methodology developed in [10] and in this paper can be used in sensitivity analysis for a broad class of control constrained optimal control problems with different dynamics.

2. Preliminaries

We will recall the parametric optimal control problem for semilinear parabolic equations, which was considered in [11]. Let $\Omega \subset I\!R^N$ $(N \geq 2)$ be a bounded domain with boundary $\partial\Omega = \Gamma$. For a fixed $T > 0$, we put $Q = \Omega \times (0, T)$ and $\Sigma = \Gamma \times (0, T)$. By A we denote an elliptic differential operator

$$A\, y := -\sum_{i,j=1}^{N} D_j\,(a_{ij} D_i\, y)$$

with sufficiently smooth coefficients $a_{ij} = a_{ij}(x)$ satisfying the condition of symmetry $a_{ij} = a_{ji}$. By $\partial_\nu y$ the co-normal derivative of y at Γ is denoted, where ν is the outward normal to Γ. Thus we have

$$\partial_\nu y := \sum_{i,j=1} a_{ij} \nu_i \, D_j y.$$

Let H be a Banach space of parameters and $G \subset H$ a bounded open set of feasible parameters. For any feasible h consider the following optimal control problem depending on h:

(P_h) Find $(y,u) \in C(\overline{Q}) \times L^\infty(Q)$ such that

$$\mathcal{J}_h(y_h,u_h) = \min\{\mathcal{J}_h(y,u) := \int_Q \psi(x,t,y,u,h)dxdt\} \tag{2.1}$$

subject to

$$\begin{aligned} y_t(x,t) + A\,y(x,t) + a(x,t,y(x,t),u(x,t),h) &= 0 \text{ in } Q, \\ \partial_\nu y(x,t) + b(x,t,y(x,t),h) &= 0 \text{ in } \Sigma, \\ y(x,0) &= 0 \text{ in } \Omega, \end{aligned} \tag{2.2}$$

$$u \in \mathcal{U} := \{u \in L^\infty(Q) \mid m_1(x,t) \le u(x,t) \le m_2(x,t)\}, \tag{2.3}$$

where ψ, a, b, m_1, m_2 are given functions. Note that we assume the homogeneous initial condition of the state equation (2.2) just to simplify some further evaluations. We could have assumed as well $y(0) = \chi \in C(\overline{\Omega})$.

We assume that at a reference value $h_0 \in G$ of the parameter there exists a solution $(y_0,u_0) := (y_{h_0},u_{h_0})$ of (P_{h_0}), and we are interested in the following problem:

Find conditions under which there are neighborhoods $G_0 \subset G$ and $\mathcal{Y} \subset C(\overline{Q}) \times L^\infty$ of h_0 and (y_0,u_0), respectively, such that for each $h \in G_0$ there exists a unique solution (y_h,u_h) in $\mathcal{Y}$ of (P_h), and the map $h \mapsto (y_h,u_h) \cap \mathcal{Y}$ is B-differentiable.

Our starting point will be the result of [11], where conditions are derived, under which the solutions to (P_h) exist are locally unique and Lipschitz continuous with respect to h. We will show that, virtually the same conditions ensure also B-differentiability of the solutions. As in [11] we assume:

(A1) Γ is of class $C^{2,\alpha}$ for some $\alpha \in (0,1]$. A is uniformly elliptic (see e.g., the definition given in [5]). Its coefficients a_{ij} belong to $C^{1,\alpha}(\bar{\Omega})$.

(A2) The nonlinear real-valued function $a = a(x,t,y,u,h)$, defined on $\overline{Q} \times I\!R^2 \times G$, satisfies the following Carathéodory type condition:

- **(i)** For all $(y,u,h) \in I\!R^2 \times G$, $a(\cdot,\cdot,y,u,h)$ and its first- and second order derivatives $D_y a, D_u a, D^2_{yy} a, D^2_{yu} a, D^2_{uu} a$ (all depending on $(\cdot,\cdot,y,u,h)$) are Lebesgue measurable on $\overline{Q}$.
- **(ii)** For all $h \in G$ and almost all $(x,t) \in Q$, $a(x,t,\cdot,\cdot,h)$ is twice continuously differentiable with respect to $(y,u) \in I\!R^2$ on $I\!R^2$.

(A3) For any fixed $K > 0$, the function a fulfils the following conditions:

- **(i)** *Boundedness*

$$|a(x,t,0,u,h)| \le a_K(x,t) \quad \forall (x,t) \in Q, \ |u| \le K, \ h \in G, \tag{2.4}$$

where $a_K \in L^q(Q)$ and $q > \frac{N}{2} + 1$. There are constants $c^0_K \ge 0$ and c^1_K such that

$$c^0_K \le a_y(x,t,y,u,h) \le c^1_K \tag{2.5}$$

for a.e. $(x,t) \in Q$, all $y \in I\!R$, all $|u| \leq K$ and $h \in G$.

(ii) *Differentiability with respect to the parameter*
For almost all $(x,t) \in Q$, all $|y| \leq K$ and $|u| \leq K$, the functions $a(x,t,y,u,\cdot)$ as well as $D_x a(x,t,y,u,\cdot)$ and $D_u a(x,t,y,u,\cdot)$ are Fréchet differentiable on G.

(A4) The nonlinear real-valued function $b = b(x,t,y,h)$, defined on $\Sigma \times I\!R \times G$, satisfies conditions analogous to **(A2)**, **(A3)**. These conditions are obtained by substituting Q by Σ and deleting u in **(A2)**, **(A3)**.

(A5) The real-valued function ψ satisfies the assumptions **(A2)**, **(A3)** imposed on a, except the growth condition (2.5).

(A6) The functions m_1 and m_2 are of class $L^\infty(Q)$ and $m_1(x,t) < m_2(x,t)$ on Q.

By a *weak solution* of (2.2) we understand a function $y \in L^2(0,T;H^1(\Omega)) \cap C(\overline{Q})$ such that

$$\begin{aligned} &\int_Q (-y \cdot p_t + \langle \nabla_x y, \nabla_x p \rangle) dxdt + \int_Q a(x,t,y,u,h)\; p\; dxdt \\ &\qquad + \int_\Sigma b(x,t,y,h)\; p\; dS_x dt = 0 \end{aligned} \tag{2.6}$$

for all $p \in W_2^{1,1}(Q)$ satisfying $p(x,T) = 0$. Here dS_x denotes the surface measure induced in Γ. The following theorem is a special case of a more general result proved in [5] or [14].

Theorem 2.1. *Suppose that* **(A1)**-**(A4)** *are satisfied and* $u \in L^\infty(Q)$. *Then problem (2.2) has a unique weak solution*

$$y \in L^2(0,T;H^1(\Omega)) \cap C(\overline{Q}).$$

$\Diamond$

Let us introduce the following spaces

$$\begin{aligned} &W(0,T) = \{y \in L^2(0,T;H^1(\Omega)) \mid y_t \in L^2(0,T;(H^1(\Omega))')\}, \\ &\text{where prime denotes the dual space,} \\ &W^s := \{y \in W(0,T) \mid y_t + A\,y \in L^s(Q),\, \partial_\nu y \in L^s(\Sigma),\, y(0) \in C(\overline{\Omega})\} \\ &Z^s = W^s \times L^s(Q), \qquad \text{where } s \in [2,\infty]. \end{aligned} \tag{2.7}$$

In W^s, we shall use the norm

$$\|y\|_{W^s} = \|y_t + A\,y\|_{L^s(Q)} + \|\partial_\nu y\|_{L^s(\Sigma)} + \|y(0)\|_{C(\overline{\Omega})}.$$

For $s > \max\{N/2+1, N+1\}$, this space is continuously embedded in $C(\overline{Q})$. This follows from the results of [5] and [14]. By the definition of the norm in W^s, the operators $y_t + A\,y$ and $\partial_\nu y$ are continuous from W^s to $L^s(Q)$ and $L^s(\Sigma)$, respectively.

Define the following Hamiltonian $\mathcal{H} = \mathcal{H}(x,t,y,u,p,h) : I\!R^{N+4} \times G \to I\!R$,

$$\mathcal{H} = \psi(x,t,y,u,h) - p \cdot a(x,t,y,u,h) \tag{2.8}$$

and the Lagrangian $\mathcal{L} : W^\infty \times L^\infty(Q) \times W(0,T) \times G \to I\!R$,

$$\begin{aligned} \mathcal{L}(y,u,p,h) \;\; &:= \int_Q \mathcal{H}(y,u,p,h)\, dxdt - \int_\Sigma p \cdot b(y,h)\, dS_x dt \\ &\quad - \int_\Omega p(0)y(0) dx - \int_Q (y_t + A\,y)\, p\, dxdt. \end{aligned} \tag{2.9}$$

The stationarity conditions of the Lagrangian have the form:

$$D_y\mathcal{L}(y,u,p,h)\, z = 0 \qquad \text{for all } z \in W^\infty, \tag{2.10}$$

$$D_u\mathcal{L}(y,u,p,h)(v-u) = \int_Q D_u\mathcal{H}(y,u,p,h)(v-u)\,dxdt \geq 0 \quad \text{for all } v \in \mathcal{U}. \tag{2.11}$$

Condition (2.10) yields the adjoint equation

$$\begin{aligned} -p_t(x,t) + A\,p(x,t) &= D_y\mathcal{H}(x,t,y,u,p,h) && \text{in } Q,\\ \partial_\nu p(x,t) + D_y b(x,t,y,h)\,p(x,t) &= 0 && \text{in } \Sigma,\\ p(x,T) &= 0 && \text{in } \Omega. \end{aligned} \tag{2.12}$$

Define the spaces:

$$\begin{aligned} W_0^s &:= \{y \in W^s \mid y(0) = 0\}, \quad W_T^s := \{p \in W^s \mid p(T) = 0\},\\ X^s &:= W_0^s \times L^s(Q) \times W_T^s,\\ \Delta^s &:= L^s(Q) \times L^s(\Sigma) \times L^s(Q) \times L^s(Q) \times L^s(\Sigma). \end{aligned} \tag{2.13}$$

Introduce the following set-valued map with the closed graph

$$\mathcal{N}(u) = \begin{cases} \lambda \in \{L^\infty(Q) \mid \int_Q \lambda(v-u)\,dxdt \leq 0 \quad \forall v \in \mathcal{U}\} & \text{if } u \in \mathcal{U},\\ \emptyset & \text{if } u \notin \mathcal{U}. \end{cases} \tag{2.14}$$

Using (2.14), the optimality system consisting of (2.12) and (2.11) as well as of (2.2) and (2.3) can be expressed in the form of the following *generalized equation*

$$0 \in \mathcal{F}(\xi,h) + \mathcal{T}(\xi),$$

where $\xi = (y,u,p)$, while $\mathcal{F} : X^\infty \times G \to \Delta^\infty$ and $\mathcal{T} : X^\infty \to 2^{\Delta^\infty}$, are, respectively, a function and a set valued map with closed graph, given by

$$\mathcal{F}(\xi,h) = \begin{bmatrix} -p_t + A\,p - D_y\mathcal{H}(y,u,p,h) & \text{in } Q\\ \partial_\nu p + D_y b(y,h)\,p & \text{in } \Sigma\\ D_u\mathcal{H}(y,u,p,h) & \text{in } Q\\ y_t + A\,y + a(y,u,h) & \text{in } Q\\ \partial_\nu y + b(y,h) & \text{in } \Sigma \end{bmatrix}, \tag{2.15}$$

$$\mathcal{T} = [\{0\},\{0\},\mathcal{N}(u),\{0\},\{0\}]^T. \tag{2.16}$$

To simplify notation, the subscript 0 will be used to denote that a given function is evaluated at the reference solution, e.g., $\mathcal{H}_0(x,t) := \mathcal{H}(x,t,y_0,u_0,\, p_0,h_0)$.

We assume:

(A7) For a fixed reference value $h_0 \in G$ of the parameter there exists a solution $(y_0,u_0) := (y_{h_0},u_{h_0}) \in Z^\infty$ of (P_{h_0}) and an associated adjoint state $p_0 := p_{h_0} \in W_T^\infty$. The element $\xi_0 := (y_0,u_0,p_0)$ satisfies the generalized equation

$$0 \in \mathcal{F}(\xi_0,h_0) + \mathcal{T}(\xi_0). \tag{2.17}$$

3. Application of abstract theorems for generalized equations

We are going to investigate conditions of existence, local uniqueness, Lipschitz continuity and differentiability of solutions $\xi_h = (y_h, u_h, p_h)$ to the generalized equation

$$0 \in \mathcal{F}(\xi, h) + \mathcal{T}(\xi). \tag{3.1}$$

We will use certain theorems for abstract generalized equations [15, 6]. Note that, by our assumptions, $\mathcal{F}$ is Fréchet differentiable in ξ for $h \in G$ and in h for $\xi \in X^\infty$. Along with (3.1), let us introduce the following generalized equation, obtained from (3.1) by linearization of $\mathcal{F}$ and by perturbation:

$$\delta \in \mathcal{F}(\xi_0, h_0) + D_\xi \mathcal{F}(\xi_0, h_0)(\zeta - \xi_0) + \mathcal{T}(\zeta), \tag{3.2}$$

where $\delta \in \Delta^\infty$ is the perturbation. Clearly, for $\delta = 0$, ξ_0 is a solution to (3.2).

We will denote by

$$\mathcal{B}^X_\rho(x_0) := \{x \in X \mid \|x - x_0\|_X \leq \rho\}$$

the closed ball of radius ρ centered at x_0 in a Banach space X.

To investigate Lipschitz continuity properties of the solutions to (3.1), the following Robinson's abstract implicit function theorem is used (see Theorem 2.1 and Corollary 2.2 in [15]).

Theorem 3.1. *If there exist $\rho_1 > 0$ and $\rho_2 > 0$ such that, for each $\delta \in \mathcal{B}^{\Delta^\infty}_{\rho_1}(0)$ there is a unique solution ζ_δ in $\mathcal{B}^X_{\rho_2}(\xi_0)$ of (3.2), which is Lipschitz continuous in δ, then there exist $\sigma_1 > 0$ and $\sigma_2 > 0$ such that, for each $h \in \mathcal{B}^H_{\sigma_1}(h_0)$ there is a unique solution ξ_h in $\mathcal{B}^X_{\sigma_2}(\xi_0)$ of (3.1), which is Lipschitz continuous in h.* $\Diamond$

Similarly, to investigate differentiability of the solutions to (3.1) we use the following theorem of Dontchev (see Theorem 2.4 and Remark 2.6 in [6]).

Theorem 3.2. *If the assumptions of Theorem 3.1 are satisfied and, in addition, the solutions ζ_δ of (3.2) are directionally (respectively, Gateaux, Bouligand, Fréchet) differentiable functions of δ in a neighborhood of the origin, with the differential $(D_\delta \zeta_0; \eta)$, then the solutions ξ_h of (3.1) are directionally (respectively, Gateaux, Bouligand, Fréchet) differentiable in a neighborhood of h_0. For a direction $g \in H$, the differential at h_0 is given by*

$$(D_h \xi_0; g) = (D_\delta \zeta_0; -D_h \mathcal{F}(\xi_0, h_0) g). \tag{3.3}$$

$\Diamond$

Note that, the result of Theorem 3.2 is actually contained in Theorem 2.3 in [15].

Remark 3.3. In Theorem 3.1, Lipschitz continuity of ζ and ξ is understood in the sense of this norm in the space X, in which $\mathcal{F}(\cdot, h)$ is differentiable. On the other hand, Theorem 3.2 remains true, if the differentiability is satisfied in a norm in the image space X *weaker* than that in which Lipschitz continuity in Theorem 3.1 holds (see Remark 2.11 in [6]); e.g., in L^p, $(p < \infty)$, rather than in L^∞. This property will be used in Section 4. $\Diamond$

Theorems 3.1 and 3.2 allow us to deduce stability and sensitivity properties of solutions to nonlinear generalized equations (3.1) from the same properties of solutions to linear equation (3.2). Usually, checking these last properties is much easier than the original ones.

We will proceed in the following steps:

1) For $\mathcal{F}$ and $\mathcal{T}$ given in (2.15) and (2.16), we find the generalized equation (LO$_\delta$) being the linearization (3.2) of (3.1).
2) We notice that (LO$_\delta$) constitutes an optimality system for a linear-quadratic accessory problem (QP$_\delta$).
3) We impose coercivity condition **(AC)**. By the results of [11], we find that, for δ sufficiently small, (QP$_\delta$) has a locally unique stationary point, which is a Lipschitz continuous function of δ. Thus, (LO$_\delta$) has a locally unique solution, which is Lipschitz in δ.
4) We show that, under **(AC)**, the solutions to (LO$_\delta$) are B-differentiable functions of δ.
5) Using the abstract theorms we find that the solutions to the nonlinear generalized equation (3.1) are B-differentiable functions of h.
6) By the results of [11], condition **(AC)** imply that for h in a neighborhhod of h_0, solutions to (3.1) correspond to the solutions of (P$_h$) and to the associated adjoint states. Thus, we arrive at our principal differentability result.

We start with point 1). Let $\delta = (\delta^1, \delta^2, \delta^3, \delta^4, \delta^5) \in \Delta^\infty$ be the vector of perturbations. Recall that the subscript 0 is used to denote that a given function is evaluated at the reference solution. In view of (2.15) and (2.16), the generalized equation (3.2) takes the form

$$(\mathrm{LO}_\delta) \qquad \left.\begin{aligned} -q_t + A\,q + D_y a_0\, q &= g_Q^0 + \delta^1 + D_{yy}^2\mathcal{H}_0\, z + D_{yu}^2\mathcal{H}_0 v, \\ \partial_\nu q + D_y b_0\, q &= g_\Sigma^0 + \delta^2 - p_0 \cdot D_{yy}^2 b_0 z, \end{aligned}\right\} \tag{3.4}$$

$$D_{uy}^2\mathcal{H}_0\, z + D_{uu}^2\mathcal{H}_0 v \,-\, D_u a_0\, q - g_u^0 - \delta^3 \in -\mathcal{N}(v), \tag{3.5}$$

$$\left.\begin{aligned} z_t + A\,z + D_y a_0\, z &= d_Q^0 + \delta^4 - D_u a_0\, v, \\ \partial_\nu z + D_y b_0\, z &= d_\Sigma^0 + \delta^5, \end{aligned}\right\} \tag{3.6}$$

where

$$\left.\begin{aligned} g_Q^0 &= D_y\psi_0 - D_{yy}^2\mathcal{H}_0\, y_0 - D_{yu}^2\mathcal{H}_0\, u_0, \\ g_\Sigma^0 &= p_0 \cdot D_{yy}^2 b_0\, y_0, \\ g_u^0 &= -D_u\psi_0 + D_{uy}^2\mathcal{H}_0\, y_0 + D_{uu}^2\mathcal{H}_0\, u_0, \\ d_Q^0 &= -a_0 + D_y a_0\, y_0 + D_u a_0\, u_0, \\ d_\Sigma^0 &= -b_0 + D_y b_0\, y_0. \end{aligned}\right\} \tag{3.7}$$

Note that

$$(z_0, v_0, q_0) = (y_0, u_0, p_0) \tag{3.8}$$

is a solution to (LO)$_\delta$ for $\delta = 0$. An inspection shows that (LO$_\delta$) constitutes an optimality system for the following linear-quadratic *accessory problem*:

$$
(\mathrm{QP}_\delta) \qquad \begin{aligned}
&\text{Find } (z_\delta, v_\delta) \in W(0,T) \times L^2(Q) \text{ that minimizes} \\
&\mathcal{I}_\delta(\zeta) = \tfrac{1}{2}((z,v), D^2_{\zeta\zeta}\mathcal{L}_0(z,v)) + \int_Q (g_Q^0 + \delta^1) z \, dxdt \\
&\qquad + \int_Q (g_u^0 + \delta^3) v \, dxdt + \int_\Sigma (g_\Sigma^0 + \delta^2)\, z \, dS_x dt
\end{aligned} \tag{3.9}
$$

subject to

$$
\begin{aligned}
z_t + A\,z + D_y a_0\, z &= d_Q^0 + \delta^4 - D_u a_0\, v && \text{in } Q \\
\partial_\nu z + D_y b_0\, z &= d_\Sigma^0 + \delta^5 && \text{in } \Sigma \\
z(0) &= 0 && \text{in } \Omega,
\end{aligned} \tag{3.10}
$$

and

$$
v \in \mathcal{U}, \tag{3.11}
$$

where the quadratic form in the cost functional $\mathcal{I}_\delta(\zeta)$ is given by

$$
\begin{aligned}
((z_1,v_1), D^2_{\zeta\zeta}\mathcal{L}_0(z_2,v_2)) &= \int_Q [z_1, v_1] \begin{bmatrix} D^2_{yy}\mathcal{H}_0 & D^2_{yu}\mathcal{H}_0 \\ D^2_{uy}\mathcal{H}_0 & D^2_{uu}\mathcal{H}_0 \end{bmatrix} \begin{bmatrix} z_2 \\ v_2 \end{bmatrix} dxdt \\
&\quad + \int_\Sigma z_1 p_0 \cdot D^2_{yy} b_0 z_2 \; dS_x dt.
\end{aligned} \tag{3.12}
$$

Remark 3.4. Problem (QP_δ) is defined in the Hilbert space $W(0,T) \times L^2(Q)$, rather than in Z^∞. It follows from classical results for parabolic equations (see e.g., Theorem 5.1, Chpt III in [9]) that for $v \in L^2(Q)$ there exists a unique weak solution $z \in W(0,T)$ of (3.10). So, (QP_δ) is well-defined. It can be shown (see e.g., [21]) that, under the coercivity assumption stated below in Section 4, for δ sufficiently small, solutions to (QP_δ) exist are locally unique and belong to Z^∞. $\Diamond$

In order to apply Theorems 3.1 and 3.2, we have to show that the stationary points of (QP_δ) are Lipschitz continuous and differentiable functions of δ.

4. Differentiability of the solutions to accessory problems

In [11] a coercivity condition was introduced, under which, for δ sufficiently small, stationary points of (QP_δ) are Lipschitz continuous. We are going to show that, under the same conditions, the stationary points are also B-differentiable.

To introduce this coercivity condition, define the sets:

$$
I = \{(x,t) \in Q \mid u_0(x,t) = m_1(x,t)\}, \quad J = \{(x,t) \in Q \mid u_0(x,t) = m_2(x,t)\}. \tag{4.1}
$$

Moreover, for $\alpha \geq 0$ define

$$
I^\alpha = \{(x,t) \in I \mid D_u\mathcal{H}_0(x,t) > \alpha\}, \quad J^\alpha = \{(x,t) \in J \mid -D_u\mathcal{H}_0(x,t) > \alpha\}. \tag{4.2}
$$

Assume the following:

(AC) (*Coercivity*) There exist $\alpha > 0$ and $\gamma > 0$ such that

$$\begin{array}{ll} (\zeta, D^2_{\zeta\zeta}\mathcal{L}_0\zeta) \geq \gamma \, \|v\|^2_{L^2(Q)} & \text{for all } \zeta := (z,v) \text{ such that} \\ z_t + A\,z + D_y a_0\, z + D_u a_0\, v = 0 & \text{in } Q \\ \partial_\nu z + D_y b_0\, z = 0 & \text{in } \Sigma \\ v = 0 & \text{in } I^\alpha \cup J^\alpha. \end{array} \tag{4.3}$$

By Lemma 5.1 in [21], **(AC)** implies

$$D^2_{uu}\mathcal{H}^0(x,t) \geq \gamma \quad \text{for a.a. } (x,t) \in Q \setminus (I^\alpha \cup J^\alpha). \tag{4.4}$$

The following result is a slight modification of Theorem 5.1 in [21] (see also Theorem 4.3 in [11]).

Proposition 4.1. *If conditions* **(A1)-(A7)** *and* **(AC)** *hold, then there exist constants $\rho > 0$ and $\varsigma > 0$ such that for each $\delta \in \mathcal{B}^{\Delta^\infty}_\rho(0)$ there is a unique stationary point*

$$(z_\delta, v_\delta, q_\delta) \in Z^\infty \times Y^\infty$$

in $\mathcal{B}^{X^\infty}_\varsigma(\xi_0)$ of (QP_δ). *Moreover, there exists a constant $\ell > 0$ such that*

$$\|z_{\delta'} - z_{\delta''}\|_{W^s}, \|v_{\delta'} - v_{\delta''}\|_{L^s(Q)}, \|q_{\delta'} - q_{\delta''}\|_{W^s} \leq \ell \, \|\delta' - \delta''\|_{\Delta^s}, \tag{4.5}$$

for all $\delta', \delta'' \in \mathcal{B}^{\Delta^\infty}_\rho(0)$ and all $s \in [2, \infty]$. ◊

Remark 4.2. Coercivity condition (4.3) is not satisfied in L^∞ in which problem (P_h) is well posed and differentiable, but in the weaker L^2-norm. This phenomenon is called *two-norm discrepancy* and it is typical for nonlinear optimal control problems. Two-norm discrepancy complicates the stability analysis, since under **(AC)**, the natural norm in which we can get Lipschitz continuity is L^2, and it is too weak to apply Theorem 3.1. In the proof of Theorem 5.1 in [21], the smoothing property of the state equation is exploited to pass from Lipschitz continuity in L^2 to such continuity in L^∞. Below, in the proof of B-differentiability, we will use (4.5) for $s < \infty$. ◊

The proof of B-differentiability of the stationary points of (QP_δ) is performed in two steps. In the first step, directional differentiability is proved and the directional differential is characterized. This characterization is used in the second step to show that the differential is actually Bouligand.

Let us start with the directional differentiability.

Proposition 4.3. *Let* **(A1)-(A7)** *as well as* **(AC)** *be satisfied and let $\rho > 0$, $\varsigma > 0$ be as in Proposition 4.1. Let $\zeta_\delta := (z_\delta, v_\delta, q_\delta) \in \mathcal{B}^{X^\infty}_\varsigma(\xi_0)$ denote the unique stationary point in $\mathcal{B}^{X^\infty}_\varsigma(\xi_0)$ of* (QP_δ). *Then the map*

$$\zeta_\delta := (z_\delta, v_\delta, q_\delta) : \mathcal{B}^{\Delta^\infty}_\rho(0) \to X^2,$$

is directionally differentiable. The directional differential at $\delta = 0$ in a direction $\eta \in \Delta^\infty$ is given by $(\varpi_\eta, w_\eta, r_\eta)$, where (ϖ_η, w_η) is the solution and r_η the associated adjoint state *of the following linear-quadratic optimal control problem:*

(LQ$_\eta$) *Find* $(\varpi_\eta, w_\eta) \in W^2 \times L^2(Q)$ *that minimizes*

$$\begin{aligned} \mathcal{J}_\eta(\varpi, w) &= \tfrac{1}{2}((\varpi, w), D^2_{\zeta\zeta}\mathcal{L}_0(\varpi, w)) + \int_Q \eta^1 \varpi \, dxdt \\ &+ \int_Q \eta^3 w \, dxdt + \int_\Sigma \eta^2 \, \varpi \, dS_x dt \end{aligned} \tag{4.6}$$

subject to

$$\begin{aligned} \varpi_t + A\varpi + D_y a_0 \varpi &= -D_u a_0 \, w + \eta^4 && \text{in } Q, \\ \partial_\nu \varpi + D_y b_0 \varpi &= \eta^5 && \text{in } \Sigma, \\ \varpi(0) &= 0 && \text{in } \Omega, \end{aligned} \tag{4.7}$$

and

$$w(x,t) \begin{cases} = 0 & \text{for } (x,t) \in (I^0 \cup J^0), \\ \geq 0 & \text{for } (x,t) \in (I \setminus I^0), \\ \leq 0 & \text{for } (x,t) \in (J \setminus J^0), \\ \text{free} & \text{for } (x,t) \in Q \setminus (I \cup J). \end{cases} \tag{4.8}$$

$\Diamond$

Proof. Let us choose $\eta \in \Delta^\infty$ and let $\{\tau_k\} \downarrow 0$ be an arbitrary sequence of positive numbers convergent to zero. Denote $\delta_k = \tau_k \eta$. Let (z_k, v_k) and q_k be the solution to (QP$_{\delta_k}$) and the associated adjoint state, respectively. By Proposition 4.1 we have

$$\|z_k - z_0\|_{W^\infty}, \ \|v_k - v_0\|_{L^\infty(Q)}, \ \|q_k - q_0\|_{W^\infty} \leq \ell \, \|\tau_k \eta\|_{\Delta^\infty}, \tag{4.9}$$

which implies that there exists a constant $l > 0$ such that

$$\left\| \frac{z_k - z_0}{\tau_k} \right\|_{W(0,T)}, \ \left\| \frac{v_k - v_0}{\tau_k} \right\|_{L^2(Q)}, \ \left\| \frac{q_k - q_0}{\tau_k} \right\|_{W(0,T)} \leq l. \tag{4.10}$$

Hence there exist a subsequence, still denoted $\{\tau_k\}$, and elements ϖ, $r \in W(0,T)$ such that

$$\begin{aligned} \frac{z_k - z_0}{\tau_k} &\rightharpoonup \varpi && \text{weakly in } W(0,T), \\ \frac{q_k - q_0}{\tau_k} &\rightharpoonup r && \text{weakly in } W(0,T). \end{aligned} \tag{4.11}$$

It is well known (see [1]) that the embedding $W(0,T) \subset L^2(Q)$ is compact. So (4.11) implies

$$\begin{aligned} \frac{z_k - z_0}{\tau_k} &\to \varpi && \text{strongly in } L^2(Q), \\ \frac{q_k - q_0}{\tau_k} &\to r && \text{strongly in } L^2(Q). \end{aligned} \tag{4.12}$$

Hence, in particular

$$\begin{aligned} \frac{z_k(x,t) - z_0(x,t)}{\tau_k} &\to \varpi(x,t) && \text{a.e. in } Q, \\ \frac{q_k(x,t) - q_0(x,t)}{\tau_k} &\to r(x,t) && \text{a.e. in } Q. \end{aligned} \tag{4.13}$$

Let us rewrite (3.5) in the equivalent form of the following variational inequality:

$$(D^2_{uy}\mathcal{H}_0\, z + D^2_{uu}\mathcal{H}_0 v\ - D_u a_0\, q - g^0_u - \delta^3, u - v) \geq 0 \quad \forall u \in \mathcal{U}.$$

In view of the structure of the set $\mathcal{U}$, this inequality implies

$$\begin{aligned}&[D^2_{uy}\mathcal{H}_0(x,t)\, z(x,t) + D^2_{uu}\mathcal{H}_0(x,t)v(x,t)\ - D_u a_0(x,t)\, q(x,t) - g^0_u(x,t)\\ &\quad -\delta^3(x,t)] \times [u - v(x,t)] \geq 0 \qquad (4.14)\\ &\text{for all } u \in [m_1(x,t), m_2(x,t)] \text{ and a.a. } (x,t) \in Q.\end{aligned}$$

The linear variational inequality (4.14) depends on the vector

$$(z(x,t), q(x,t), \delta^3(x,t)) \in I\!R^3,$$

which can be treated as a parameter. Let $v_k(x,t)$ be a solution to (4.14) corresponding to $z(x,t) = z_k(x,t),\ q(x,t) = q_k(x,t),\ \delta^3(x,t) = \alpha_k \eta^3(x,t)$. In view of (4.4) and (4.13), well known sensitivity result for finite dimensional mathematical programs (see e.g., [8]) implies that

$$\frac{v_k(x,t) - v_0(x,t)}{\tau_k} \to w(x,t), \qquad (4.15)$$

where $w(x,t)$ is the solution of the following variational inequality:

$$\begin{aligned}&[D^2_{uy}\mathcal{H}_0(x,t)\,\varpi(x,t) + D^2_{uu}\mathcal{H}_0(x,t)w(x,t)\ - D_u a_0(x,t)\, r(x,t) - \eta^3(x,t)]\times\\ &\quad \times[v - w(x,t)] \geq 0 \quad \text{for all } v \in I\!R \text{ satisfying (4.8).}\end{aligned} \qquad (4.16)$$

By the Lebesgue dominated convergence theorem, the pointwise convergence (4.15) together with the bound (4.10) implies

$$\frac{v_k - v_0}{\tau_k} \to w \qquad \text{strongly in } L^2(Q), \qquad (4.17)$$

where w is the solution of the variational inequality

$$\begin{aligned}&(D^2_{uy}\mathcal{H}_0\,\varpi + D^2_{uu}\mathcal{H}_0 w\ - D_u a_0\, r - \eta^3, v - w) \geq 0\\ &\text{for all } v \in L^2(Q) \text{ satisfying (4.8).}\end{aligned} \qquad (4.18)$$

In view of the definition (2.6) of the weak solutions to parabolic boundary value problems, equations (3.6) and (3.4) together with (4.11) and (4.17) imply that ϖ and r are the solutions of the following equations:

$$\begin{aligned}\varpi_t + A\varpi + D_y a_0 \varpi &= -D_u a_0 w + \eta^4,\\ \partial_\nu \varpi + D_y b_0 \varpi &= \eta^5,\\ \varpi(0) &= 0,\end{aligned} \qquad (4.19)$$

$$\begin{aligned}-r_t + Ar + D_y a_0 r &= D^2_{yy}\mathcal{H}_0 \varpi + D^2_{yu}\mathcal{H}_0 w + \eta^1,\\ \partial_\nu r + D_y b_0 r &= -p_0 D^2_{yy} b_0 \varpi + \eta^2,\\ r(T) &= 0.\end{aligned} \qquad (4.20)$$

The state and adjoint equations (4.19) and (4.20), together with the variational inequality (4.18), constitute the optimality system for problem (LQ$_\eta$). In view of **(AC)**, this problem has a unique solution (ϖ_η, w_η) and a unique adjoint state r_η. This shows that the convergence in (4.11) and (4.17) holds for the whole sequence $\{\tau_k\}$ and completes the proof of the proposition. □

Note that, by the same argument as in Proposition 4.1, we find that the stationary points of (LQ$_\eta$) are Lipschitz continuous functions of η. Since $(\varpi_0, w_0, r_0) = (0, 0, 0)$, we have

$$\|\varpi_\eta\|_{W^s},\ \|w_\eta\|_{L^s(Q)},\ \|r_\eta\|_{W^s} \le \ell\|\eta\|_{\Delta^s}, \quad s \in [2, \infty]. \tag{4.21}$$

We are now going to show that (ϖ_η, w_η) and r_η are actually B-differentials at $\delta = 0$ of (z_δ, v_δ) and q_δ, respectively.

Theorem 4.4. *Let* **(A1)-(A7)** *as well as* **(AC)** *be satisfied and let $\rho > 0$, $\varsigma > 0$ be as in Proposition 4.1. Then the map*

$$\zeta_\delta := (z_\delta, v_\delta, q_\delta) : \mathcal{B}_\rho^{\Delta^\infty}(0) \to X^s, \tag{4.22}$$

where $\zeta_\delta := (z_\delta, v_\delta, q_\delta) \in \mathcal{B}_\varsigma^{X^\infty}(\xi_0)$ denotes the unique stationary point in $\mathcal{B}_\varsigma^{X^\infty}(\xi_0)$ of (QP$_\delta$)*, is B-differentiable for any $s \in [2, \infty)$. The B-differential at $\delta = 0$ in a direction $\eta \in \Delta^\infty$ is given by $\vartheta_\eta := (\varpi_\eta, w_\eta, r_\eta)$, where (ϖ_η, w_η) is the solution and r_η the associated adjoint state of problem* (LQ$_\eta$). ♢

Proof. We have to show that the solution $(\varpi_\eta, w_\eta, r_\eta)$ of (4.18)-(4.20) are B-differentials of the solution to (LO$_\delta$). Clearly, $(\varpi_\eta, w_\eta, r_\eta)$ is a positively homogeneous function of η, so it is enough to show that

$$\begin{aligned} &z_\eta = z_0 + \varpi_\eta + \sigma_1(\eta), \quad v_\eta = v_0 + w_\eta + \sigma_2(\eta), \quad q_\eta = q_0 + r_\eta + \sigma_1(\eta), \\ &\text{where } \frac{\|\sigma_1(\eta)\|_{W^s}}{\|\eta\|_{\Delta^\infty}} \to 0, \quad \frac{\|\sigma_2(\eta)\|_{L^s(Q)}}{\|\eta\|_{\Delta^\infty}} \to 0, \quad \text{as } \|\eta\|_{\Delta^\infty} \to 0, \\ &\text{for any } s \in [2, \infty). \end{aligned} \tag{4.23}$$

Denote

$$(z_\eta - z_0) = \widetilde{\varpi}_\eta, \quad (v_\eta - v_0) = \widetilde{w}_\eta, \quad (q_\eta - q_0) = \widetilde{r}_\eta. \tag{4.24}$$

It follows from (3.6) and (3.4) that $(\widetilde{\varpi}_\eta, \widetilde{w}_\eta, \widetilde{r}_\eta)$ satisfy equations identical with (4.19) and (4.20):

$$\begin{aligned} (\widetilde{\varpi}_\eta)_t + A\widetilde{\varpi}_\eta + D_y a_0 \widetilde{\varpi}_\eta &= -D_u a_0 \widetilde{w}_\eta + \eta^4, \\ \partial_\nu \widetilde{\varpi}_\eta + D_y b_0 \widetilde{\varpi}_\eta &= \eta^5, \\ \widetilde{\varpi}_\eta(0) &= 0, \end{aligned} \tag{4.25}$$

$$\begin{aligned} -(\widetilde{r}_\eta)_t + A\widetilde{r}_\eta + D_y a_0 \widetilde{r}_\eta &= D_{yy}^2 \mathcal{H}_0 \widetilde{\varpi}_\eta + D_{yu}^2 \mathcal{H}_0 \widetilde{w}_\eta + \eta^1, \\ \partial_\nu \widetilde{r}_\eta + D_y b_0 \widetilde{r}_\eta &= -p_0 D_{yy}^2 b_0 \widetilde{\varpi}_\eta + \eta^2, \\ \widetilde{r}_\eta(T) &= 0. \end{aligned} \tag{4.26}$$

Let us choose $\beta \in (0, \alpha)$, where α is given in **(AC)**. Define the sets

$$\begin{aligned} K_1^\beta &= \{(x,t) \in I^0 \mid D^2_{uu}\mathcal{H}_0(x,t) \in (0,\beta)\}, \\ K_2^\beta &= \{(x,t) \in J^0 \mid -D^2_{uu}\mathcal{H}_0(x,t) \in (0,\beta)\}, \\ L^\beta &= \{(x,t) \in Q \\ &\qquad \mid u_0(x,t) \in (m_1(x,t), m_1(x,t)+\beta) \cup (m_2(x,t)-\beta, m_2(x,t))\}. \end{aligned} \tag{4.27}$$

Note that

$$\text{meas } (K_1^\beta \cup K_2^\beta \cup L^\beta) \to 0 \quad \text{as } \beta \to 0. \tag{4.28}$$

Let us split up the set Q into the following subsets

$$\begin{aligned} \mathcal{A} &= Q \setminus (I \cup J \cup L^\beta), & \mathcal{B} &= (I^0 \setminus K_1^\beta) \cup (J^0 \setminus K_2^\beta), \\ \mathcal{C} &= (I \setminus I^0) \cup (J \setminus J^0), & \mathcal{D} &= K_1^\beta \cup K_2^\beta \cup L^\beta, \end{aligned}$$

where I, J, I^0 and J^0 are given in (4.1) and (4.2). We will analyze conditions analogous to (4.14) on each of these subsets successively.

Subset $\mathcal{A}$

Choose $\varrho(\beta) = \ell^{-1}\beta$. Then by (3.8) and (4.1), as well as by Proposition 4.1, for all $\eta \in \mathcal{B}^{\Delta\infty}_{\varrho(\beta)}(0)$ we get

$$v_\eta(x,t) \in (m_1(x,t), m_2(x,t)) \qquad \text{for a.a. } (x,t) \in \mathcal{A}, \tag{4.29}$$

i.e., by (4.14)

$$\begin{aligned} &D^2_{uy}\mathcal{H}_0(x,t)\, z_\eta(x,t) + D^2_{uu}\mathcal{H}_0(x,t) v_\eta(x,t) - D_u a_0(x,t)\, q_\eta(x,t) \\ &\quad -g^0_u(x,t) - \eta^3(x,t) = 0 \quad \text{ for a.a. } (x,t) \in \mathcal{A}. \end{aligned} \tag{4.30}$$

Subtracting from (4.30) the analogous equation for (z_0, v_0, q_0) and using notation (4.24), we obtain

$$\begin{aligned} &D^2_{uy}\mathcal{H}_0(x,t)\, \widetilde{\varpi}_\eta(x,t) + D^2_{uu}\mathcal{H}_0(x,t) \widetilde{w}_\eta(x,t) - D_u a_0(x,t)\, \widetilde{r}_\eta(x,t) \\ &\quad -\eta^3(x,t) = 0 \quad \text{ for a.a. } (x,t) \in \mathcal{A}. \end{aligned} \tag{4.31}$$

Subset $\mathcal{B}$

It follows from Proposition 4.1 that, shrinking $\varrho(\beta) > 0$ if necessary, for all $\eta \in \mathcal{B}^{\Delta\infty}_{\varrho(\beta)}(0)$ we obtain

$$\begin{aligned} &D^2_{uy}\mathcal{H}_0(x,t)\, z_\eta(x,t) + D^2_{uu}\mathcal{H}_0(x,t) v_\eta(x,t) - D_u a_0(x,t)\, q_\eta(x,t) \\ &\quad -g^0_u(x,t) - \eta^3(x,t) \begin{cases} > 0 & \text{for a.a. } (x,t) \in I^0 \setminus K_1^\beta, \\ < 0 & \text{for a.a. } (x,t) \in J^0 \setminus K_2^\beta, \end{cases} \end{aligned} \tag{4.32}$$

which, by (4.14) implies

$$v_\eta(x,t) = \begin{cases} m_1(x,t) & \text{for a.a. } (x,t) \in I^0 \setminus K_1^\beta, \\ m_2(x,t) & \text{for a.a. } (x,t) \in J^0 \setminus K_2^\beta, \end{cases}$$

i.e.,

$$\widetilde{w}_\eta(x,t) = 0 \qquad \text{for a.a. } (x,t) \in \mathcal{B}. \tag{4.33}$$

<u>Subset $\mathcal{C}$</u>
By (3.8), (4.1) and (4.2) we have

$$v_0(x,t) = u_0(x,t) = \begin{cases} m_1(x,t) & \text{for a.a. } (x,t) \in I \setminus I^0, \\ m_2(x,t) & \text{for a.a. } (x,t) \in J \setminus J^0 \end{cases} \tag{4.34}$$

and

$$\begin{aligned} &D^2_{uy}\mathcal{H}_0(x,t)\, z_0(x,t) + D^2_{uu}\mathcal{H}_0(x,t) v_0(x,t) \;-\; D_u a_0(x,t)\, q_0(x,t) = 0 \\ &\text{for a.a. } (x,t) \in (I \setminus I^0) \cup (J \setminus J^0). \end{aligned} \tag{4.35}$$

Proposition 4.1, together with (4.34) implies that, shrinking $\varrho(\beta)$ if necessary, for any $\eta \in \mathcal{B}^{\Delta^\infty}_{\varrho(\beta)}(0)$ we get

$$v_\eta(x,t) \in \begin{cases} [m_1(x,t), m_2(x,t)) & \text{for a.a. } t \in I \setminus I^0, \\ (m_1(x,t), m_2(x,t)] & \text{for a.a. } t \in J \setminus J^0. \end{cases} \tag{4.36}$$

Hence, in view of (4.14) we have

$$\begin{aligned} &D^2_{uy}\mathcal{H}_0(x,t)\, z_\eta(x,t) + D^2_{uu}\mathcal{H}_0(x,t) v_\eta(x,t) \;-\; D_u a_0(x,t)\, q_\eta(x,t) \\ &\qquad -\eta^3(x,t) \begin{cases} \geq 0 & \text{for a.a. } (x,t) \in I \setminus I^0, \\ \leq 0 & \text{for a.a. } (x,t) \in J \setminus J^0. \end{cases} \end{aligned} \tag{4.37}$$

Conditions (4.34)–(4.37) imply:

$$\widetilde{w}_\eta(x,t) \begin{cases} \geq 0 & \text{for a.a. } (x,t) \in I \setminus I^0, \\ \leq 0 & \text{for a.a. } (x,t) \in J \setminus J^0 \end{cases}, \tag{4.38}$$

$$\begin{aligned} &D^2_{uy}\mathcal{H}_0(x,t)\, \widetilde{\varpi}_\eta(x,t) + D^2_{uu}\mathcal{H}_0(x,t)\widetilde{w}_\eta(x,t) \;-\; D_u a_0(x,t)\, \widetilde{r}_\eta(x,t) \\ &\qquad -\eta^3(x,t) \begin{cases} \geq 0 & \text{for a.a. } (x,t) \in I \setminus I^0, \\ \leq 0 & \text{for a.a. } (x,t) \in J \setminus J^0, \end{cases} \end{aligned} \tag{4.39}$$

and

$$\begin{aligned} &(D^2_{uy}\mathcal{H}_0(x,t)\, \widetilde{\varpi}_\eta(x,t) + D^2_{uu}\mathcal{H}_0(x,t)\widetilde{w}_\eta(x,t) \;-\; D_u a_0(x,t)\, \widetilde{r}_\eta(x,t) \\ &\qquad -\eta^3(x,t))(w - \widetilde{w}_\eta(x,t)) \geq 0 \; \begin{cases} \text{for all } w \geq 0 & \text{on } I \setminus I^0, \\ \text{for all } w \leq 0 & \text{on } J \setminus J^0. \end{cases} \end{aligned} \tag{4.40}$$

<u>Subset $\mathcal{D}$</u>
The analysis of subset $\mathcal{D}$ is the most difficult, because we do not know a priori if for $(x,t) \in \mathcal{D}$ the constraints are active or not at v_η, no matter how small η is chosen. Without this information, we can say very little about $\widetilde{w}_\eta(x,t) = v_\eta(x,t) - v_0(x,t)$. Let us denote

$$\begin{aligned} (\widetilde{\eta}^3)'(x,t) &= D^2_{uy}\mathcal{H}_0(x,t)\,(z_\eta(x,t) - z_0(x,t)) + D^2_{uu}\mathcal{H}_0(x,t)(v_\eta(x,t) - v_0(x,t)) \\ &\quad -D_u a_0(x,t)\,(q_\eta(x,t) - q_0(x,t)) \text{ for a.a. } (x,t) \in \mathcal{D}. \end{aligned} \tag{4.41}$$

By definition (4.24) we have

$$
\begin{aligned}
&D^2_{uy}\mathcal{H}_0(x,t)\,\widetilde{\varpi}_\eta(x,t) + D^2_{uu}\mathcal{H}_0(x,t)\widetilde{v}_\eta(x,t) \;-D_u a_0(x,t)\,\widetilde{r}_\eta(x,t)\\
&-(\widetilde{\eta}^3)'(x,t) = 0 \quad \text{for a.a. } (x,t)\in\mathcal{D}.
\end{aligned}
\tag{4.42}
$$

Denote $\eta' = (\eta^1,\eta^2,(\eta^3)',\eta^4,\eta^5)$, where

$$
(\eta^3)'(x,t) = \begin{cases} (\widetilde{\eta}^3)'(x,t) & \text{for } (x,t)\in\mathcal{D},\\ \eta^3(x,t) & \text{otherwise.}\end{cases}
\tag{4.43}
$$

It is easy to see that (4.25) and (4.26) together with (4.31), (4.33), (4.38)–(4.40) and (4.42) can be interpreted as an optimality system for the optimal control problem $(\widetilde{\mathrm{LQ}}_{\eta'})$, where $(\widetilde{\mathrm{LQ}}_\eta)$ is the following slight modification of (LQ_η):

$(\widetilde{\mathrm{LQ}}_\eta)$ Find $(\widetilde{\varpi}_\eta,\widetilde{w}_\eta)\in W^2\times L^2(Q)$ that minimizes $\mathcal{J}_\eta(\varpi,w)$ subject to

$$
\begin{aligned}
\varpi_t + A\,\varpi + D_y a_0\,\varpi &= -D_u a_0\,w + \eta^4 && \text{in } Q,\\
\partial_\nu\varpi + D_y b_0\,\varpi &= \eta^5 && \text{in } \Sigma,\\
\varpi(0) &= 0 && \text{in } \Omega,
\end{aligned}
$$

and

$$
w(x,t)\begin{cases} =0 & \text{for } (x,t)\in(I^0\setminus K_1^\beta)\cup(J^0\setminus K_2^\beta),\\ \ge 0 & \text{for } (x,t)\in(I\setminus I^0),\\ \le 0 & \text{for } (x,t)\in(J\setminus J^0),\\ \text{free} & \text{for } (x,t)\in(Q\setminus(I\cup J))\cup(K_1^\beta\cup K_2^\beta).\end{cases}
$$

Similarly $(\varpi_\eta,w_\eta,r_\eta)$ can be interpreted as a stationary point of $(\widetilde{\mathrm{LQ}}_{\eta''})$, where $\eta'' = (\eta^1,\eta^2,(\eta^3)'',\eta^4,\eta^5)$, with

$$
\begin{aligned}
&(\eta^3)''(x,t) = \begin{cases} (\widetilde{\eta}^3)''(x,t) & \text{for } (x,t)\in\mathcal{D},\\ \eta^3(x,t) & \text{otherwise,}\end{cases}\\
&(\widetilde{\eta}^3)''(x,t) = D^2_{uy}\mathcal{H}_0(x,t)\,\varpi_\eta(x,t) + D^2_{uu}\mathcal{H}_0(x,t)w_\eta(x,t) \;-D_u a_0(x,t)\,r_\eta(x,t).
\end{aligned}
\tag{4.44}
$$

It can be easily checked that, as in the case of (LQ_η), the stationary points of $(\widetilde{\mathrm{LQ}}_\eta)$ are Lipschitz continuous functions of η. Hence, in view of (4.43) and (4.44), we have

$$
\begin{aligned}
&\|\widetilde{\varpi}_\eta-\varpi_\eta\|_{W^s},\|\widetilde{w}_\eta-w_\eta\|_{L^s(Q)},\|\widetilde{r}_\eta-r_\eta\|_{W^s}\\
&\quad\le \ell\,\|\eta'-\eta''\|_{\Delta^s} = \ell\left\{\int\limits_{K_1^\beta\cup K_2^\beta\cup L^\beta} |(\widetilde{\eta}^3)'(x,t)-(\widetilde{\eta}^3)''(x,t)|^s dxdt\right\}^{\frac{1}{s}}
\end{aligned}
\tag{4.45}
$$

Using the definitions (4.41), (4.44) and taking advantage of (4.5) and of (4.21) we get

$$
\begin{aligned}
&|(\widetilde{\eta}^3)'(x,t) - (\widetilde{\eta}^3)''(x,t)| \le |(\widetilde{\eta}^3)'(x,t)| + |(\widetilde{\eta}^3)''(x,t)| \\
&= |D^2_{uy}\mathcal{H}_0(x,t)\,(z_\eta(x,t) - z_0(x,t)) + D^2_{uu}\mathcal{H}_0(x,t)(v_\eta(x,t) - v_0(x,t)) \\
&\quad -D_u a_0(x,t)\,(q_\eta(x,t) - q_0(x,t))| \\
&\quad +|D^2_{uy}\mathcal{H}_0(x,t)\,\varpi_\eta(x,t) + D^2_{uu}\mathcal{H}_0(x,t)v_\eta(x,t) \; - D_u a_0(x,t)\,r_\eta(x,t)| \\
&\quad \le c\;\|\eta\|_{\Delta^\infty} \quad \text{for a.a. } (x,t) \in K_1^\beta \cup K_2^\beta \cup L^\beta.
\end{aligned}
\tag{4.46}
$$

Substituting (4.46) to (4.45) we obtain

$$
\begin{aligned}
&\|\widetilde{\varpi}_\eta - \varpi_\eta\|_{W^s}, \|\widetilde{w}_\eta - w_\eta\|_{L^s(Q)}, \|\widetilde{r}_\eta - r_\eta\|_{W^s} \\
&\qquad \le c\|\eta\|_{\Delta^\infty} \left\{ \text{meas}\ (K_1^\beta \cup K_2^\beta \cup L^\beta) \right\}^{\frac{1}{s}}.
\end{aligned}
\tag{4.47}
$$

In view of (4.24) and(4.28), we find that for any $\epsilon > 0$ and any $s \in [2, \infty)$ we can choose $\beta(\epsilon, s) > 0$ and the corresponding $\varrho(\beta(\epsilon, s))$, so small that

$$
\begin{aligned}
&\|z_\eta - z_0 - \varpi_\eta\|_{W^s}, \|v_\eta - v_0 - w_\eta\|_{L^s(Q)}, \|q_\eta - q_0 - r_\eta\|_{W^s} \\
&\qquad \le \epsilon\;\|\eta\|_{\Delta^\infty} \quad \text{for all } \eta \in \mathcal{B}^{\Delta^\infty}_{\varrho(\beta(\epsilon,s))}(0).
\end{aligned}
$$

This shows that (4.23) holds and completes the proof of the theorem. □

Remark 4.5. The proof of Theorem 4.4 cannot be repeated for $s = \infty$ and the counterexample in [10] shows that B-differentiability of (4.22) cannot be expected for $s = \infty$. ♢

5. Differentiability of the solutions to nonlinear problems

Theorems 3.2 and 4.4 imply that, for h in a neighborhood of h_0, (P_h) has a locally unique stationary point (y_h, u_h, p_h), which is a B-differentiable function of h. On the other hand, by Lammas 5.1 and 5.2 in [11], condition **(AC)** implies that (y_h, u_h) and p_h are a solution and the associated adjoint state of (P_h), respectively. Thus, we arrive at the following principal result of this paper:

Theorem 5.1. *If* **(A1)**–**(A7)** *and* **(AC)** *hold, then there exist constants $\sigma_1 > 0$ and $\sigma_2 > 0$ such that for all $h \in \mathcal{B}^H_{\sigma_1}(h_0)$ there is a unique stationary point (y_h, u_h, p_h) in $\mathcal{B}^{X^\infty}_{\sigma_2}((y_0, u_0, p_0))$ of (P_h), where (y_h, u_h) is a solution to (P_h). The map*

$$
(y_h, u_h, p_h) : \; \mathcal{B}^H_\sigma(h_0) \to X^s, \quad s \in [2, \infty) \tag{5.1}
$$

is B-differentiable, and the B-differential evaluated at h_0 in a direction $g \in H$ is given by the solution and adjoint state of the following linear-quadratic optimal control problem

(L_g) *Find* $(z_g, v_g) \in W^2 \times L^2(Q)$ *that minimizes*

$$\mathcal{K}_g(z,v) = \frac{1}{2}((z,v), D^2_{\zeta\zeta}\mathcal{L}_0(z,v)) + \int_Q D^2_{yh}\mathcal{H}_0 g z \, dxdt + \int_Q D^2_{uh}\mathcal{H}_0 g v \, dxdt + \int_\Sigma p_0 D^2_{yh} b_0 g \, z \, dS_x dt$$

subject to

$$\begin{aligned} z_t + A\,z + D_y a_0\, z &= -D_u a_0\, v - D_h a_0\, g && \text{in } Q, \\ \partial_\nu z + D_y b_0\, z &= -D_h b_0\, g && \text{in } \Sigma, \\ z(0) &= 0 && \text{in } \Omega, \end{aligned}$$

and

$$v(x,t) \begin{cases} = 0 & \text{for } (x,t) \in (I^0 \cup J^0), \\ \geq 0 & \text{for } (x,t) \in (I \setminus I^0), \\ \leq 0 & \text{for } (x,t) \in (J \setminus J^0), \\ \text{free} & \text{for } (x,t) \in Q \setminus (I \cup J)). \end{cases}$$

$\diamondsuit$

As it was noticed in Introduction, Bouligand differential becomes Fréchet if it is linear. Hence from the form of (L_g), we obtain immediately:

Corollary 5.2. *If assumptions of Theorem 5.1 hold and* $\operatorname{meas}(I \setminus I^0) = \operatorname{meas}(J \setminus J^0) = 0$, *then the map (5.1) is Fréchet differentiable.* $\diamondsuit$

In sensitivity analysis of optimization problems an important role is played by the so-called optimal value function, which on $\mathcal{B}^H_\sigma(h_0)$ is defined by:

$$\mathcal{J}^0(h) := \mathcal{J}_h(y_h, u_h),$$

i.e., to each $h \in \mathcal{B}^H_\sigma(h_0)$, $\mathcal{J}^0$ assigns the (local) optimal value of the cost functional. The second order *directional* expansion of the optimal value function has been known from the literature (see e.g., Theorem 3.1 in [2]). The following corollary of Theorem 5.1 shows that Bouligand differentiability of the solutions implies the second order expansion of $\mathcal{J}_0$, *uniform* in a neighborhood of h_0.

Corollary 5.3. *If assumptions of Theorem 5.1 hold, then for each* $h = h_0 + g \in \mathcal{B}^H_\sigma(h_0)$

$$\begin{aligned} \mathcal{J}^0(h) &= \mathcal{J}^0(h_0) + (D_h\mathcal{L}_0, g) \\ &+\tfrac{1}{2}\left((z_g, v_g, g), \begin{pmatrix} D^2_{yy}\mathcal{L}_0 & D^2_{yu}\mathcal{L}_0 & D^2_{yh}\mathcal{L}_0 \\ D^2_{uy}\mathcal{L}_0 & D^2_{uu}\mathcal{L}_0 & D^2_{uh}\mathcal{L}_0 \\ D^2_{hy}\mathcal{L}_0 & D^2_{hu}\mathcal{L}_0 & D^2_{hh}\mathcal{L}_0 \end{pmatrix} (z_g, v_g, g)\right) \\ &+o(\|g\|^2_H), \end{aligned} \tag{5.2}$$

where (z_g, v_g) *is the B-differential of* (y_h, u_h) *at* h_0 *in the direction* g, *i.e., it is given by the solution to* (L_g). $\diamondsuit$

Proof Denote $\mathcal{L}_h := \mathcal{L}(y_h, u_h, p_h, h)$. It follows from the definition (2.9) that

$$\mathcal{J}^0(h) = \mathcal{L}_h. \tag{5.3}$$

From (5.3) and Theorem 4.1 we find the following form of the Bouligand differential of the value function

$$D_h\mathcal{J}^0(h)g = D_y\mathcal{L}_h z_g + D_u\mathcal{L}_h v_g + D_p\mathcal{L}_h q_g + D_h\mathcal{L}_h g, \tag{5.4}$$

where (z_g, v_g, q_g) is the B-differential of (y_h, u_h, p_h) in the direction g. By optimality condition for (P_h)

$$D_y\mathcal{L}_h = 0, \qquad D_p\mathcal{L}_h = 0. \tag{5.5}$$

Moreover, since (z_g, v_g, q_g) is given by the stationary point of the linear–quadratic problem analogous to (L_g), but evaluated at h rather than at h_0, we find that

$$D_u\mathcal{L}_h v_g = 0. \tag{5.6}$$

Equation (5.4) together with (5.5) and (5.6) yields

$$D_h\mathcal{J}^0(h)g = D_h\mathcal{L}_h g. \tag{5.7}$$

By (5.3) and (5.7) we have

$$\mathcal{J}^0(h) = \mathcal{J}^0(h_0) + \int_0^1 D_h\mathcal{L}_{h_\alpha} g\, d\alpha, \tag{5.8}$$

where $h_\alpha = h_0 + \alpha g$. Using Theorem 5.1 we obtain

$$\begin{aligned} D_h\mathcal{L}_{h_\alpha} &= D_h\mathcal{L}_0 + \alpha(D^2_{hy}\mathcal{L}_0 z_g + D^2_{hu}\mathcal{L}_0 v_g + D^2_{hp}\mathcal{L}_0 q_g + D^2_{hh}\mathcal{L}_0 g) \\ &\quad + o(\alpha\|g\|_H). \end{aligned} \tag{5.9}$$

Substituting (5.9) to (5.8) and integrating we get

$$\begin{aligned} &\mathcal{J}^0(h) = \mathcal{J}^0(h_0) + D_h\mathcal{L}_0 g \\ &\quad + \tfrac{1}{2}(D^2_{hy}\mathcal{L}_0 z_g + D^2_{hu}\mathcal{L}_0 v_g + D^2_{hp}\mathcal{L}_0 q_g + D^2_{hh}\mathcal{L}_0 g, g) + o(\|g\|^2_H). \end{aligned} \tag{5.10}$$

Differentiating (5.5) with respect to h, we obtain

$$D^2_{yy}\mathcal{L}_0 z_g + D^2_{yu}\mathcal{L}_0 v_g + D^2_{yp}\mathcal{L}_0 q_g + D^2_{yh}\mathcal{L}_0 g = 0, \tag{5.11}$$

$$D^2_{py}\mathcal{L}_0 z_g + D^2_{pu}\mathcal{L}_0 v_g + D^2_{ph}\mathcal{L}_0 g = 0. \tag{5.12}$$

On the other hand, by the optimality condition for (L_g)

$$(D^2_{uy}\mathcal{L}_0 z_g + D^2_{uu}\mathcal{L}_0 v_g + D^2_{up}\mathcal{L}_0 q_g + D^2_{uh}\mathcal{L}_0 g, v_g) = 0. \tag{5.13}$$

Multiplying (5.11) and (5.12) by z_g and q_g, respectively, combining with (5.13) and substituting to (5.10) we obtain (5.2). □

Acknowledgements. The author would like to express his deep gratitude to one of the anonymous referees for very careful reading of the manuscript and important suggestions that allowed to correct some errors and improve the presentation.

References

[1] J. P. Aubin: Un théorème de compacité, C. R. Acad. Sci. Paris 256 (1963) 5012–5044.

[2] J. F. Bonnans: Second-order analysis for control constrained optimal control problems of semilinear elliptic systems, Appl. Math. Optim. 38 (1998) 303–325.

[3] J. F. Bonnans, A. Shapiro: Perturbation Analysis of Optimization Problems, Springer, New York (2000).

[4] M. Bergounioux, N. Mérabet: Sensitivity analysis and optimal control of problems governed by semilinear parabolic equations, Control. Cybern. 29 (2000) 861–886.

[5] E. Casas: Pontryagin's principle for state-constrained boundary control problems of semilinear parabolic equations, SIAM J. Control Optim. 35 (1997) 1297–1327.

[6] A. L. Dontchev: Implicit function theorems for generalized equations, Math. Program. 70 (1995) 91–106.

[7] A. L. Dontchev, W. W. Hager, A. B. Poore, B. Yang: Optimality, stability and convergence in nonlinear control, Appl. Math. Optim. 31 (1995) 297–326.

[8] K. Jittorntrum: Solution point differentiability without strict complementarity in nonlinear programming, Math. Program. Study 21 (1984) 127–138.

[9] O. A. Ladyženskaya, V. A. Solonnikov, N. N. Ural'ceva: Linear and quasilinear equations of parabolic type, Transl. of Math. Monographs, Vol. 23, Amer. Math. Soc., Providence, R.I. (1968).

[10] K. Malanowski: Bouligand differentiability of solutions to parametric optimal control problems, Working Paper WP-1-2000, System Research Institute, Polish Academy of Sciences, Warsaw (2000).

[11] K. Malanowski, F. Tröltzsch: Lipschitz stability of solutions to parameric optimal control problems for parabolic equations, Zeitschrift Anal. Anwend. 18 (1999) 469–489.

[12] K. Malanowski, F. Tröltzsch: Lipschitz stability of solutions to parameric optimal control problems for elliptic equations, Control Cybern. 29 (2000) 237–256.

[13] F. Mignot: Contrôle dans les inéquations variationnelles, J. Func. Anal. 22 (1976) 130–185.

[14] J. P. Raymond, H. Zidani: Hamiltonian Pontryagin's principles for control problems governed by semilinear parabolic equations, Appl. Math. Optim. 39 (1999) 143–177.

[15] S. M. Robinson: Strongly regular generalized equations, Math. Oper. Res. 5 (1980) 43–62.

[16] S. M. Robinson: Local structure of feasible sets in nonlinear programming, Part III: Stability and sensitivity, Math. Program. Study 30 (1987) 97–116.

[17] S. M. Robinson: An implicit-function theorem for a class of nonsmooth functions, Math. Oper. Res. 16 (1991) 292–309.

[18] A. Shapiro: On concepts of directional differentiability, J. Math. Anal. Appl. 66 (1990) 477–487.

[19] A. Shapiro: Perturbation analysis of optimization problems in Banach spaces, Numer. Funct. Anal. Optimiz. 13 (1992) 97–116.

[20] J. Sokołowski: Sensitivity analysis of control constrained optimal control problems for distributed parameter systems, SIAM J. Control Optimiz. 26 (1987) 1542–1556.

[21] F. Tröltzsch: Lipschitz stability of solutions to linear-quadratic parabolic control problems with respect to perturbations, Discr. Cont. Dynam. Systems 6 (2000) 289–306.