Analysis of the Lagrange–SQP–Newton Method
for the Control of a Phase Field Equation

Matthias Heinkenschloss
Fredi Tröltzsch

November 1998

TR98-23
Analysis of the Lagrange–SQP–Newton Method for the Control of a Phase Field Equation

Matthias Heinkenschloss ∗ Fredi Tröltzsch†

November 1998

Abstract

This paper investigates the local convergence of the Lagrange–SQP–Newton method applied to an optimal control problem governed by a phase field equation with distributed control. The phase field equation is a system of two semilinear parabolic differential equations. Stability analysis of optimization problems and regularity results for parabolic differential equations are used to proof convergence of the controls with respect to the $L^2(Q)$ norm and with respect to the $L^\infty(Q)$ norm.

Key words Sequential quadratic programming method, Lagrange–SQP–Newton method, optimal control, phase field equation, control constraints.

AMS subject classifications 49M37, 49K20

1 Introduction

In this paper we investigate the local convergence of the Lagrange–Sequential–Quadratic–Programming– (SQP)–Newton method for the solution of an optimal control problem governed by a phase field equation. Phase field equations are used to model solidification. They are systems of partial differential equations (PDEs). The unknowns in the system of PDEs are the order parameter $\varphi$ (also called phase function) and the temperature $u$. Unlike in the classical Stefan problem which models a sharp solid-liquid interface, phase field models allow for a mushy region. The phases are identified using the order parameter $\varphi$. Assuming a suitable normalization, $\{x \in \Omega | \varphi(x) = 1\}$ is the liquid region and $\{x \in \Omega | \varphi(x) = -1\}$ is the solid region. The interface is described by points $x \in \Omega$ for which the order parameter takes values in $(-1, 1)$. We consider the phase field model introduced in [7], [13], [18] that consists of the two differential equations

$$\frac{\partial}{\partial t} u + \frac{\ell}{2} \frac{\partial}{\partial t} \varphi = \kappa \Delta u + f$$  \hspace{1cm} (1.1)

in $\Omega \times (0,T]$, $\frac{\tau}{\partial^2} \varphi = \xi^2 \Delta \varphi + g(\varphi) + 2u$  \hspace{1cm} (1.2)

with boundary conditions

$$\frac{\partial}{\partial n} u = 0, \quad \frac{\partial}{\partial n} \varphi = 0, \quad \text{on } \partial \Omega \times (0,T).$$  \hspace{1cm} (1.3)

∗Department of Computational and Applied Mathematics, Rice University, Houston, Texas 77005–1892, E-Mail: heinken@rice.edu. This work was supported in part by the NSF under Grant DMS–9403699 and by the AFOSR under Grant F49620–93–1–0280.

†Technische Universität Chemnitz-Zwickau, Fakultät für Mathematik, D–09114 Chemnitz, Germany, E-Mail: troeltzsch@mathematik.tu-chemnitz.de. This work was supported in part by AFOSR under Grant F49620–93–1–0280 while the author was a visiting scientist at the Interdisciplinary Center for Applied Mathematics, Virginia Polytechnic Institute and State University.
and initial conditions

\[ u = u_0, \quad \varphi = \varphi_0 \quad \text{in } \Omega. \tag{1.4} \]

The function \(-g(z)\) is the derivative of a so-called double well potential \(G(z)\). Often \(G(z) = \frac{1}{4}(z^2 - 1)^2\). We assume that \(g(z) = az + bz^2 - cz^3\) with bounded coefficient functions \(a, b, c\) and strictly positive \(c\). In the model \(\kappa\) denotes the heat conductivity, \(\ell\) the latent heat, \(\tau\) is the relaxation time, and \(\xi\) is the length scale of the interface. Other, more complicated phase field equations have been derived. See, e.g., [28, 29, 38, 39, 41].

In this paper we study a constrained distributed control problem in which the state equation is given by the phase field model. The objective is to find a heat input \(f\) such that the resulting temperature \(u\) and phase \(\varphi\) match desired temperature and phase profiles \(u_d\) and \(\varphi_d\), respectively. Mathematically, the problem is formulated as

\[
\min_{f \in F_{ad}} J(f),
\tag{1.5}
\]

where the objective function is given by

\[
J(u, \varphi, f) = \int_0^T \int_\Omega \frac{\alpha}{2} (u(x,t) - u_d(x,t))^2 + \frac{\beta}{2} (\varphi(x,t) - \varphi_d(x,t))^2 + \frac{\gamma}{2} (f(x,t))^2 \, dx \, dt.
\tag{1.6}
\]

The set of admissible controls has the form

\[
F_{ad} = \{ f \in L^2(\Omega \times (0,T)) \mid f(x,t) \in F \; \text{a.e.} \},
\tag{1.7}
\]

where \(F \subset \mathbb{R}\) is a closed interval, e.g., \(F = [-1,1]\). We do not require \(F\) to be bounded. In the objective function \(u\) and \(\varphi\) are the solutions of (1.1) to (1.4) corresponding to the control \(f\). Thus, the right hand side \(f\) is the control and the pair \((u, \varphi)\) is the state. In the objective function \(\alpha\) and \(\beta\) are weighting parameters, and \(\frac{\gamma}{2} \int_0^T \|f(t)\|^2 \, dt\) is a regularization term. We assume that

\[
u_d, \varphi_d \in L^2(\Omega \times (0,T))
\]

are given functions and that \(\alpha, \beta, \text{ and } \gamma\) are positive constants. Later we will require that \(u_d, \varphi_d \in L^q(\Omega \times (0,T))\) with \(q > 2\) if \(n = 2\) and \(q > 5/2\) if \(n = 3\) to derive regularity estimates for the adjoint variables and our strongest convergence estimates for the optimization algorithm. The requirement \(\gamma > 0\) is important for our second order sufficient optimality condition. It can only be expected to hold if \(\gamma > 0\).

The infinite dimensional phase field model (1.1) to (1.4) has been analyzed in [7, 17, 24, 10]. In [7, 17] the case \(f \equiv 0\) is considered. Numerical investigations of the phase field model (1.1) to (1.4) can be found e.g., in [18, 34, 19, 8, 12, 10]. Numerical simulations using other phase field equations have been performed, e.g., in [9, 25, 29].

The control problem stated above has been analyzed in [10, 11, 24]. In [10, 24] the infinite dimensional control problem is considered. Existence and uniqueness results for the optimal control are derived and the differentiability of the state with respect to the control is analyzed. In [10, 11] a discretization of the control problem is introduced and some of its approximation properties are analyzed. The gradient method for the numerical solution of the control problem (1.5), (1.6), (1.1), (1.2), (1.3), (1.4) is studied in [11]. For optimal control problems governed by the Penrose–Fife phase field model [38] existence of solutions and their characterization is studied in [25, 42, 43].

The purpose of this paper is the analysis of the Lagrange SQP–Newton method for the solution of the above mentioned control problem. SQP methods are used to solve nonlinear constrained optimization problems. Their success for finite dimensional problems has sparked the research on their application to optimal control and other infinite dimensional problems. SQP methods treat states and controls as independent variables. The nonlinear problem is solved using a sequence of linear quadratic problems. In the context of control problems with linear control constraints the constraints of the quadratic program are given by the linearized state equation and the linear control constraints. In our analysis we use exact second order derivative information and, since we are interested in the local convergence analysis, we use the quadratic model of the Lagrangian about the current iterate as the objective function in the quadratic programming subproblems. This method is called Lagrange–SQP–Newton method. If no control constraints are given, then near to a local minimum point that satisfies the second order sufficient optimality conditions the Lagrange–SQP–Newton method is equivalent to the Newton method applied to the necessary optimality conditions. In the presence of control constraints it is equivalent to the generalized Newton method applied to a set of generalized equations [1, 40].
Since states and controls are treated as independent variables the nonlinear state equation does not have to be solved in every iteration, but is part of the constraints and is satisfied in the limit. Another attractive feature is the fast local convergence speed of SQP methods. If exact second derivative information is used these methods show a local q–quadratic convergence behavior. If quasi–Newton approximations for the second derivatives are used, then they show some kind of q–superlinear convergence. SQP methods for finite dimensional nonlinear programming problems are discussed, e.g., in the overview article [5]. Theoretical and numerical studies of SQP methods applied to optimal control problems in an infinite dimensional framework can found, e.g., in [1, 2, 3, 20, 23, 27, 30, 32, 44, 45]. A local convergence analysis for reduced SQP methods in Hilbert spaces using quasi–Newton updates is given in [31]. Studies of the local convergence behavior of the Lagrange–SQP–Newton method for several classes of optimal control problems can be found, e.g., in [1, 2, 4, 44].

The general outline of our convergence proof for the Lagrange–SQP–Newton method for (1.5), (1.6), (1.1), (1.2), (1.3), (1.4) is identical to the ones in [1, 44]. The details of the convergence proof, however, are very different from those in [1, 44]. These differences are due to differences in the governing equations. In [1] the governing equations are ordinary differential equations and in [44] the governing equation is the linear heat equation with a nonlinear boundary condition. Here we carefully use the structure of the phase field equations (1.1), (1.2), (1.3), (1.4) to overcome a two–norm discrepancy and to show convergence of the controls with respect to the $L^2(Q)$ norm and with respect to the $L^\infty(Q)$ norm. In particular, we will show that

$$
\|u_+ - u_*\|_{W^{2,1}_{q}} + \|\varphi_+ - \varphi_*\|_{W^{2,1}_{q}} + \|f_+ - f_*\|_{L^\infty} + \|\lambda_+ - \lambda_*\|\Lambda_q 
\leq C (\|u_+ - u_*\|_{W^{2,1}_{q}} + \|\varphi_+ - \varphi_*\|_{W^{2,1}_{q}} + \|f_+ - f_*\|_{L^\infty} + \|\lambda_+ - \lambda_*\|\Lambda_q)^2
$$

where $C$ is some positive constant and $q > 2$ if $n = 2$ or $q > 5/2$ if $n = 3$. Here the subscripts $*, +, c$ denote optimal solution, new iterate, and current iterate, respectively, and $\Lambda = (p, \psi)$ are the Lagrange multipliers in the dual space $\Lambda_q = W^{2,1}_q(Q) \times W^{2,1}_q(Q)$. A complete review of the notation applied in this paper is given at the end of this section. The surprising feature of the estimate (1.8) is that the $L^\infty$ norm of the error in the new control $f_+ - f_*$ can be estimated using the much weaker $L^p$ norm of the error in the current control $f_c - f_*$. It is necessary to discuss what we mean by a two–norm discrepancy. Often, differentiation of the objective and constraints is only possible with respect to a rather strong norm in the control space, say the $L^\infty$–norm, whereas the second order sufficient optimality conditions hold only with respect to a weaker norm, say the $L^2$–norm. This two–norm discrepancy principle plays an important role in the analysis of nonlinear control problems and we refer to [2, 15, 16, 36] as a selection of references in which various aspects of this and related issues are investigated. More references can be found in those papers. In our case the situation is slightly different in that the nonlinear term $f$ in the state equation is a polynomial of degree three. Using Hölder’s inequality it can be seen that $\varphi \to g(\varphi)$ is infinitely often differentiable as an operator from $L^q(Q)$ to $L^2(Q)$. Using the smoothness of solutions of parabolic equations this enables us to prove convergence of the controls in $L^2(Q)$. This seems to be the natural space if the set of controls $F_{ad}$ is bounded, then the controls are in $L^\infty(Q)$ and one wants to establish convergence with respect to this stronger norm. It is in this case that the two–norm discrepancy issue arises. The difficulties that have to be overcome are the same. However, the reason for the two–norm discrepancy is different. Differentiability can be shown if $\varphi \to g(\varphi)$ is viewed as an operator from $L^q(Q)$ to $L^2(Q)$. For the problem under consideration these norms are "compatible". The desire to have convergence of the controls with respect to the $L^\infty(Q)$–norm, which is important from a numerical point of view, e.g., for the identification of active indices, causes the incompatibility.

In [22] a multilevel Newton method is applied to solve the unconstrained control problem (1.5), (1.6), (1.1), (1.2), (1.3), (1.4) is solved and in [23] the constrained control problem studied in this paper is treated numerically. The multilevel Newton method in [22] is an extension of the SQP method in that it incorporates an efficient solution method for the computation of the steps. Its convergence, however, requires controls in $L^q(Q)$ with $q > 4$. The numerical method in [23] is a combination of the multilevel Newton method in [22] and the projected Newton method in [26]. While no convergence analysis for the algorithm in [23] exists, the strong convergence results for the Lagrange–SQP–Newton method proven in this paper might serve as an indication why the algorithm [23] performed well numerically.

As in [44], our second order sufficient conditions requires the positive definiteness of the Hessian of the Lagrangian on the space of functions satisfying the homogeneous linearized state equations. In the presence of bound constraints on the controls this seems to be too strong, because the active bound constraints limit the space of functions on which...
the Hessian of the Lagrangian has to be positive definite further. See, e.g., [37]. In the papers [20, 45] convergence of the Lagrange–SQP–Newton method for some semilinear parabolic control problems is proven under second order sufficient optimality conditions weaker than the ones we use. Active bound constraints are incorporated into the positive definiteness conditions. The mathematical tool in [20, 45] is the convergence analysis of Newton’s method for generalized nonlinear equations [1, 40]. The price for the relaxation of the second order sufficient optimality condition, however, is that a constraint involving the optimal control has to be introduced into the quadratic–programming (QP) which generates the new iterate. Since the purpose of the Lagrange–SQP–Newton method is the determination of the optimal control such a constraint is not practical. Fortunately, its inclusion does not seem to be necessary in practice [20, 21]. This shows, however, that there still is a gap between practical, efficient algorithms for semilinear control problems and their theoretical justification. This paper is meant to narrow this gap.

The outline of the paper is as follows: In Section 2 we review some results on the existence and uniqueness of solutions of the state equation and of the optimal control problem. Necessary and sufficient optimality conditions are discussed in Section 3. The SQP method and basic properties of the iterates will be discussed in Section 4. The solutions of the state equation and of the optimal control problem. Necessary and sufficient optimality conditions weaker than the ones we use. Active bound constraints are incorporated into the generalized nonlinear equations [1, 40]. The price for the relaxation of the second order sufficient optimality condition, however, is that a constraint involving the optimal control has to be introduced into the quadratic–programming (QP) which generates the new iterate. Since the purpose of the Lagrange–SQP–Newton method is the determination of the optimal control such a constraint is not practical. Fortunately, its inclusion does not seem to be necessary in practice [20, 21]. This shows, however, that there still is a gap between practical, efficient algorithms for semilinear control problems and their theoretical justification. This paper is meant to narrow this gap.

Before we start with the discussion of the control problem, we state the assumptions which are assumed to hold throughout this paper:

(A1) The domain \( \Omega \subset \mathbb{R}^n, n = 2, 3 \), is a bounded \( C^2 \)–domain.

(A2) The coefficients in (1.1), (1.2) satisfy

\[ \kappa, \xi, \ell, \tau > 0 \]

and the function \( g \) is given by \( g(z) = a z + b z^2 - c z^3 \) with

\[ a(x, t) \leq \alpha, \quad b(x, t) \leq \beta, \quad 0 < \varepsilon \leq c(x, t) \leq c. \]

(The results can be generalized to the case where \( \kappa, \xi, \ell, \) and \( \tau \) are strictly positive, sufficiently smooth functions.)

(A3) The initial conditions satisfy

\[ u_0, \varphi_0 \in W^{2}_{\infty}(\Omega) \]

and the compatibility conditions

\[ \frac{\partial}{\partial n} u_0 = \frac{\partial}{\partial n} \varphi_0 = 0. \]

We use the following notation: The space–time–domain is denoted by \( Q = \Omega \times (0, T) \). For \( p \in [1, \infty) \) we define

\[ W^{2,1}_p(Q) = \left\{ u \mid u, \frac{\partial u}{\partial x_i}, \frac{\partial^2 u}{\partial x_i \partial x_j}, \frac{\partial u}{\partial t} \in L^p(Q) \right\}. \]

The space \( W^{2,1}_p(Q) \) equipped with the norm

\[ \| u \|_{W^{2,1}_p(Q)} = \left( \int_Q |u|^p + \sum_{i=1}^n \left| \frac{\partial u}{\partial x_i} \right|^p + \sum_{i,j=1}^n \left| \frac{\partial^2 u}{\partial x_i \partial x_j} \right|^p + \left| \frac{\partial u}{\partial t} \right|^p \ dx \right)^{1/p} \]

is a Banach space. We often omit the space \( Q \) and use \( W^{2,1}_p, L^p \) instead of \( W^{2,1}_p(Q), L^p(Q) \), respectively. By \( \langle \cdot, \cdot \rangle \) and \( \| \cdot \| \) we denote the scalar product and norm in \( L^2(\Omega) \).

In the SQP method we use the following notations: The current iterates are denoted by a subscript \( c \), the new iterates are indicated by a subscript \( + \), and the optimal values have a subscript \( * \). Thus, \( (u_c, \varphi_c, f_c), (u_+, \varphi_+, f_+), (u_*, \varphi_*, f_*) \) denote the current iterate, new iterate, and minimum point, respectively. Moreover, we use the notations

\[ v = (u, \varphi, f), \quad \lambda = (p, \psi), \quad (1.9) \]
for the triple of states and control and the pair of co–states, respectively. Similar notations are used for \(v_*, \hat{v}, \) etc.

Finally, we introduce the product spaces

\[ V_q = W_q^{2,1}(Q) \times W_q^{2,1}(Q) \times L^q(Q), \quad \Lambda_q = W_q^{2,1}(Q) \times W_q^{2,1}(Q). \]  

(1.10)

For \( q = 2 \) we simply write \( V = V_2 \) and \( \Lambda = \Lambda_2. \)

In all our proofs, \( C \) will be a generic positive constant.

\section{Well–Posedness of the State Equation and Existence of Optimal Controls}

Existence and uniqueness of the solution of the state equation (1.1), (1.2), (1.3), (1.4) are proven in [10] and in [24]. Other proofs for the case \( f = 0 \) can be found in [7] and [17]. The following result is taken from [24].

\textbf{Theorem 2.1} If the assumptions (A1)–(A3) are satisfied, then for each \( f \in L^q(Q), q \geq 2, \) there exists a unique solution \((u, \varphi) \in W_q^{2,1}(Q) \times W_p^{2,1}(Q)\) of the state equation (1.1) to (1.4). Moreover, the solution obeys

\[ \|u\|_{W_q^{2,1}} + \|\varphi\|_{W_p^{2,1}} \leq C(\|u_0\|_{W_q^\infty(Q)} + \|\varphi_0\|_{W_q^\infty(Q)} + \|f\|_{L^q}). \]

Here the parameter \( p \) is given by

\[ p = \begin{cases} \frac{5}{2} & \text{if } q \in [2, 5/2) \text{ and } n = 3, \\ \text{any positive number} & \text{if } q > 5/2 \text{ and } n = 3, \text{ or if } q \geq 2 \text{ and } n = 2. \end{cases} \]

Note that since \( 5q/(5 - 2q) > q \) for \( q \in [2, 5/2) \) we may set \( p = q \) in Theorem 2.1. In addition to \( p = q \) we will frequently use Theorem 2.1 with \( q = 2, p = 10. \)

The proof of Theorem 2.1 uses the Leray–Schauder fixed point theorem and the following imbedding results due to Lions and Peetre:

\textbf{Theorem 2.2} If \( \Omega \subset \mathbb{R}^2 \) is a bounded domain having the cone property, then the imbeddings

\[ W_q^{2,1}(Q) \subset L^\infty(Q), \quad q > 2, \]

and

\[ W_q^{2,1}(Q) \subset L^\infty(0,T;L^p(\Omega)) \cap L^q(0,T;L^\infty(\Omega)), \quad p,q \in [1, \infty) \]

are continuous; the imbeddings \( W_q^{2,1}(Q) \subset L^p(Q), p \in [1, \infty), q \in [2, \infty) \) are compact. If \( \Omega \subset \mathbb{R}^2 \) is a bounded domain having the cone property, then the imbeddings

\[ W_q^{2,1}(Q) \subset L^p(Q), \]

where

\[ p = \begin{cases} \infty & \text{if } q > 5/2, \\ \text{any positive number} & \text{if } q = 5/2, \\ \frac{5}{2} & \text{if } q < 5/2, \end{cases} \]

are continuous; the imbeddings \( W_q^{2,1}(Q) \subset L^{p-\epsilon}(Q), \) where \( p \) is given as above and \( \epsilon > 0, \) are compact.

\textbf{Proof:} The assertions are proven in [35, pp. 14,15,24,25]. \( \square \)

We will frequently use this theorem with \( p = q \) or \( q = 2, p = 10 \) or \( q = 10, p = \infty. \)

Theorem 2.1 allows us to define the solution operator

\[ S : L^q(Q) \ni f \mapsto W_q^{2,1}(Q) \times W_p^{2,1}(Q), \quad \varphi, \]

(2.1)
mapping the right-hand side into the solution of the state equations (1.1)–(1.4). It is shown in [24] that the solution operator is continuous and Fréchet differentiable. In particular, it holds that
\[ ||u||_{W^{2,1}_2} + ||\varphi||_{W^{2,1}_2} \leq c \left( ||u_0||_{W^{2,1}_0} + ||\varphi_0||_{W^{2,1}_0} + ||f||_{L^2} \right), \tag{2.2} \]
where \( c = c(\ell, \xi, T), \) and
\[ ||u_1 - u_2||_{W^{2,1}_2} + ||\varphi_1 - \varphi_2||_{W^{2,1}_2} \leq C||f_1 - f_2||_{L^2} \tag{2.3} \]
for \((u_i, \varphi_i) = S(f_i), i = 1, 2,\) where \( C \) depends on \( ||u_i||_{W^{2,1}_2}, ||\varphi_i||_{W^{2,1}_2}, i = 1, 2.\) If \( u_i, \varphi_i, i = 1, 2,\) are contained in bounded subsets \( B \subset (W^{2,1}_2)^2,\) which will be the case in our analysis, then \( C \) can be assumed to depend only on \( B,\) but not to be independent of individual \( u_i, \varphi_i \in B, i = 1, 2.\) In particular, if \( f_1, f_2 \) are in a bounded set \( \mathcal{F}_b \subset \mathcal{F}_{ad},\) then (2.2) implies that (2.3) holds with \( C \) depending only on \( \mathcal{F}_b.\) If \( \mathcal{F}_{ad} \) is already bounded, then there is a \( C \) such that (2.3) is valid for all \( f_1, f_2 \in \mathcal{F}_{ad}.\)

Since the objective function (1.6) includes a regularization term \( \frac{1}{2} \int_0^T ||f(t)||^2 \, dt, \gamma > 0,\) one can show the existence of a minimizing sequence \( \{f_n\}, \) which is bounded in \( L^2(Q).\)

Using the existence of weakly converging subsequences and the weak lower semi-continuity of the objective function one can prove the existence of an optimal control. For details we refer to [10] and [24]. If \( \gamma = 0 \) and if \( F \) is bounded, then one can extract a subsequence out of the minimizing sequence that is weak* convergent in \( L^\infty(Q).\) Using arguments as before, one can establish the existence of an optimal control. In [24] it is also shown that the optimal control is unique if the final time \( T \) is sufficiently small. The existence and uniqueness results for optimal controls are summarized in the following theorem taken from [24]:

**Theorem 2.3** Let the assumptions of Theorem 2.1 be valid. If \( \gamma > 0 \) or if \( F \) is bounded, then there exists an optimal control \( f_* \in L^2(Q).\) The optimal control satisfies \( f_* \in L^\infty(Q) \) if \( F \) is bounded. Moreover, if \( \gamma > 0 \) and if the final time \( T \) is sufficiently small, then the optimal control \( f_* \) is unique.

For our analysis we do not need this strong result. All statements in the following sections hold for local solutions \((u_+, \varphi_+)\) that satisfy the second order sufficient conditions.

Characterizations of the optimal controls are given in the following section.

We introduce the auxiliary system
\[ \begin{align*}
\frac{\partial u}{\partial t} + \ell \frac{\partial \varphi}{\partial x} &= \kappa \Delta u + \beta_1 u + f_1 \\
\frac{\partial \varphi}{\partial t} &= \xi^2 \Delta \varphi + \beta_2 \varphi + 2u + f_2
\end{align*} \]
in \( \Omega \times (0, T], \tag{2.4} \]
with boundary conditions (1.3) and initial conditions (1.4).

This linear system will play an important role. In fact, if \( f_1 = h, f_2 = 0, u_0 = 0, \varphi_0 = 0, \) and \( \beta_1 = 0, \beta_2 = g^*(\varphi(x,t))\), then the solution \((u, \varphi)\) of (2.4), (1.3), (1.4) is the first Fréchet derivative \( (u, \varphi) = (u(f_2), h, \varphi_1(f_2), h) \) of \((u(f), \varphi(f))\) at the control \( f_2.\) Moreover, with \( f_1 = f_1, f_2 = g(\varphi), \beta_1 = 0, \) and \( \beta_2 = g(\varphi) \) the system (2.4), (1.3), (1.4) essentially defines the pair \((u_+, \varphi_+)\) of the new iterate in the SQP method. See Section 4.

**Theorem 2.4** Let the assumptions (A1)–(A3) be valid and suppose that \( f_1, f_2 \in L^q(Q), q \geq 2.\) If \( \beta_1, \beta_2 \in L^3(Q),\) then there exists a unique solution \((u, \varphi) \in W^{2,1}_q(Q) \times W^{2,1}_q(Q)\) of the system (2.4), (1.3), (1.4). The solution obeys
\[ ||u||_{W^{2,1}_q} + ||\varphi||_{W^{2,1}_q} \leq C( ||u_0||_{W^{2,1}_0} + ||\varphi_0||_{W^{2,1}_0} + ||f_1||_{L^q} + ||f_2||_{L^q} ) \]
with constant \( C = C(\beta_1, \beta_2) \) depending only on \( ||\beta_1||_{L^3}, ||\beta_2||_{L^3}.\) Moreover, the function \((\beta_1, \beta_2) \mapsto C(\beta_1, \beta_2)\) maps bounded sets into bounded sets.

**Proof:** The proof is given in the appendix. \( \square \)

The conditions on \( \beta_i, i = 1, 2,\) are clearly satisfied if \( \beta_i \in L^\infty.\) However, they are also satisfied if \( \beta_i \) is of the form
\[ \beta_i(x, t) = a(x, t) + b(x, t)\tilde{\varphi}(x, t) - c(x, t)\tilde{\varphi}(x, t)^2 \]
or, more generally, of the form
\[ \beta_i(x,t) = \sum_{j=1}^{m} c_j(x,t) \hat{\varphi}(x,t)^j, \]
with \( c_j \in L^\infty(Q) \) and \( \hat{\varphi} \in L^3 m \).

3 First and Second Order Optimality Conditions

The first order necessary optimality conditions for the optimal control problem under investigation are established in [10] and [24]. We state the first order necessary conditions in the form needed for our purposes and derive second order sufficient conditions.

The Lagrange function for the optimal control problem is given by
\[
L(u, \varphi, f, p, \psi) = J(u, \varphi, f) - \int_0^T \int_\Omega p [u_t + \ell \varphi_t - \kappa \Delta u - f] \, dx \, dt \\
- \int_0^T \int_\Omega \psi \left[ \tau \varphi_t - \xi^2 \Delta \varphi - g(\varphi) - 2u \right] \, dx \, dt
\]
which, using integration by parts, can be written as
\[
L(u, \varphi, f, p, \psi) = J(u, \varphi, f) - \int_0^T \left( u_t + \frac{\ell}{2} \varphi_t + \kappa (\nabla u \nabla p) - (f, p) \right) \, dt \\
- \int_0^T \tau (\varphi_t, \psi) + \xi^2 (\nabla \varphi, \nabla \psi) - (g(\varphi) + 2u, \psi) \, dt.
\]
(3.1)

For fixed \( p \) and \( \psi \) the Lagrangian splits into a linear part \( L_1 \) and a nonlinear part \( L_2 \), \( L = L_1 + L_2 \), where
\[
L_2(u, \varphi, f, p, \psi) = J(u, \varphi, f) + \int_0^T (g(\varphi), \psi) \, dt.
\]
(3.2)

The first order necessary optimality conditions for the optimal control problem are:
\[
L_{(u, \varphi)}(u_*, \varphi_*, f_*, p_*, \psi_*) = 0, \\
L_{(p, \psi)}(u_*, \varphi_*, f_*, p_*, \psi_*) = 0, \\
L_f(u_*, \varphi_*, f_*, p_*, \psi_*)(f - f_*) \geq 0 \quad \forall f \in \mathcal{F}_{ad},
\]
(3.3) \hspace{1cm} (3.4) \hspace{1cm} (3.5)

The equation (3.4) means that the state equation has to be satisfied. An evaluation of (3.3) shows that the pair \((p_*, \psi_*)\) has to satisfy the adjoint system
\[
-\frac{\partial}{\partial t} p = \kappa \Delta p + 2 \varphi + \alpha(u_* - u_d), \\
-\tau \frac{\partial}{\partial t} \psi - \frac{\ell}{2} \frac{\partial}{\partial t} \psi = \xi^2 \Delta \varphi + g'(\varphi_*) \psi + \beta(\varphi_* - \varphi_d), \\
\text{in } \Omega \times (0, T],
\]
(3.7)

with boundary conditions
\[
\frac{\partial}{\partial n} p = 0, \quad \frac{\partial}{\partial n} \psi = 0, \quad \text{on } \partial \Omega \times (0, T),
\]
(3.8)

and final conditions
\[
p(x, T) = 0, \quad \psi(x, T) = 0 \quad \text{in } \Omega.
\]
(3.9)

The inequality (3.5) is equivalent to the variational inequality
\[
\int_0^T \int_\Omega (p_* + \gamma f_*)(f - f_*) \, dx \, dt \geq 0 \quad \forall f \in \mathcal{F}_{ad}.
\]
Since $\gamma > 0$, a standard discussion gives
\[ f_*(x, t) = P_F(-\gamma^{-1} p_*(x, t)) \quad \text{a.e. on } Q, \]
where $P_F: \mathbb{R} \rightarrow F$ denotes the projection onto the closed set $F$.

**Theorem 3.1** Let the assumptions (A1)–(A2) be valid and suppose that the right hand side function $f$ in (1.1) obeys $f \in L^q(Q)$. If $u_d, \varphi_d \in L^p(Q), \mu \in [2, \infty)$, then there exists a unique solution $(p, \psi) \in W^{2,1}_v(Q) \times W^{2,1}_v(Q)$ of the adjoint system (3.7) to (3.9). Moreover, the solution obeys
\[ \|p\|_{W^{2,1}_v} + \|\psi\|_{W^{2,1}_v} \leq C(\|u_0 - u_d\|_{L^p} + \|\varphi_0 - \varphi_d\|_{L^p}), \]
where $\nu = \min\{\mu, \frac{5q}{q-2}\}$ if $2 \leq q < 5/2$ and $n = 3$, and $\nu = \mu$ if $q \geq 5/2$ or $n = 2$. The constant $C$ depends only on $\varphi_0$.

**Proof:** The assumptions (A1)–(A2) and $f \in L^q$ guarantee that the solution of the state equation obeys $u_*, \varphi_* \in W^{2,1}_v$. See Theorem 2.1. Thus, by Theorem 2.2, $u_0 - u_d, \varphi_0 - \varphi_d \in L^p$, where $\nu$ is defined above.

If one introduces the transformation $t \rightarrow T - t$, then (3.7) to (3.9) is equal to (2.4) with homogeneous initial and boundary conditions and with $f_1 = \alpha(u_0 - u_d), f_2 = \beta(\varphi_0 - \varphi_d)$, and $\beta_1 = g'(\varphi_0), \beta_2 = 0$. Thus the assertion follows from Theorem 2.4. \(\square\)

In the following we use the notations (1.9) and (1.10).

We will show that $v_*$ is a locally optimal point if it satisfies the first order optimality condition and
\[ \mathcal{L}_{vv}(v_*, \lambda_*)[v, v] \geq \sigma_*(\|u\|^2_{L^2} + \|\varphi\|^2_{L^2} + \|f\|^2_{L^2}) \quad (3.11) \]
for all $v = (u, \varphi, f) \in V$ satisfying the linearized state equation
\[
\begin{align*}
\frac{\partial}{\partial t} u + \frac{\partial}{\partial x} \varphi &= \kappa \Delta u + f, \\
\tau \frac{\partial}{\partial t} \varphi &= \xi^2 \Delta \varphi + g'(\varphi_*) \varphi + 2u,
\end{align*}
\]
with homogeneous initial and boundary conditions. Note that Theorem 2.4 shows that the solution of (3.12) obeys
\[ \|f\|_{L^2} \geq C(\|u\|^2_{W^{2,1}_v} + \|\varphi\|^2_{W^{2,1}_v}). \]
Hence,
\[ \mathcal{L}_{vv}(v_*, \lambda_*)[v, v] \geq \tilde{\sigma}_*\|f\|^2_{L^2} \quad (3.13) \]
for all $v = (u, \varphi, f)$ satisfying the linearized state equation (3.12) is necessary and sufficient for (3.11).

**Theorem 3.2** Suppose that the necessary optimality conditions (3.3)–(3.6) are satisfied at $(u_*, \varphi_*, f_*)$ and that (3.11) holds. Then there exist $\epsilon > 0$ and $c > 0$ such that
\[ J(u, \varphi, f) \geq J(u_*, \varphi_*, f_*) + c(\|u - u_*\|^2_{W^{2,1}_v} + \|\varphi - \varphi_*\|^2_{W^{2,1}_v} + \|f - f_*\|^2_{L^2}) \]
for all feasible $(u, \varphi, f)$ with $\|f - f_*\|_{L^2} \leq \epsilon$.

**Proof:** Let $f \in I_{ad}$ be given and let $(u, \varphi)$ be the corresponding solution of the state equation.

First we note that (2.3) implies the estimate
\[ \|u - u_*\|_{W^{2,1}_v} + \|\varphi - \varphi_*\|_{W^{2,1}_v} \leq C\|f - f_*\|_{L^2} \leq C\epsilon. \quad (3.14) \]
The first order necessary conditions (3.3)–(3.5) imply
\[ J(v) = L(v, \lambda) \]
\[ \geq J(v_*) + \frac{1}{2} \mathcal{L}_{vv}(v_*, \lambda_*)[v-v_*, v-v_*] + r(v_*, \lambda_*)[v-v_*, v-v_*], \]
where
\[ \mathcal{L}_{vv}(v_*, \lambda_*)[v-v_*, v-v_*] = \int_0^T (g''(\varphi_*)(\varphi - \varphi_*')^2, \psi_*) \, dt + \mathcal{J}_{vv}(v_*)[v-v_*, v-v_*] \]
\[ = \int_0^T ((2b - 6c\varphi_*)(\varphi - \varphi_*')^2, \psi_*) \, dt \]
\[ + \int_0^T \alpha \|u - u_*\|^2 + \beta \|\varphi - \varphi_*\|^2 + \gamma \|f - f_*\|^2 \, dt \]
(3.15)
and the remainder term is given by
\[ r(v_*, \lambda_*)[v-v_*, v-v_*] = \int_0^T \int_0^T \int_0^T \langle g''(\varphi_* + \eta(\varphi - \varphi_*)) - g''(\varphi_*)(\varphi - \varphi_*')^2, \psi_* \rangle \, d\eta \, d\tau \]
\[ = \int_0^T \int_0^T \int_0^T \langle -6c\eta(\varphi - \varphi_*'), \psi_* \rangle \, d\eta \, d\tau \]
(3.16)
cf. (3.2). Thus, using (2.3) and Theorem 2.2,
\[ |r(v_*, \lambda_*)[v-v_*, v-v_*]| \leq C\|\psi_*\|_{L^2} \|\varphi - \varphi_*\|_{L^2}^2 \leq C\|f - f_*\|_{L^2}^2. \]
(3.17)

Let \((\hat{u}, \hat{\varphi})\) be the solution of (2.4) with \(f_1 = \hat{f} = f - f_*, f_2 = 0, \beta_1 = 0, \beta_2 = g'(\varphi_*)\), and homogeneous initial and boundary conditions, i.e. let \((\hat{u}, \hat{\varphi}, \hat{f})\) solve the linearized state equation. Then Theorem 2.4 and (2.3) give the estimate
\[ \|\hat{u}\|_{L^\infty} + \|\hat{\varphi}\|_{L^\infty} \leq C(\|\hat{u}\|_{W^{2,1}} + \|\hat{\varphi}\|_{W^{2,1}}) \leq C\|f - f_*\|_{L^2}. \]
(3.18)
The pair
\[ (\delta u, \delta \varphi) = (u, \varphi) - (u_*, \varphi_*) - (\hat{u}, \hat{\varphi}) \]
is the solution of (2.4) with \(f_1 = 0, f_2 = g(\varphi) - g(\varphi_*) - g'(\varphi_*)(\varphi - \varphi_*), \beta_1 = 0, \beta_2 = g'(\varphi_*)\) and homogeneous initial and boundary conditions. From Theorem 2.4 and the imbedding Theorem 2.2 we find that
\[ \|\delta u\|_{L^\infty} + \|\delta \varphi\|_{L^\infty} \leq C\|g(\varphi) - g(\varphi_*) - g'(\varphi_*)(\varphi - \varphi_*)\|_{L^2}. \]
(3.19)
The definition of \(g\) yields
\[ g(\varphi) - g(\varphi_*) - g'(\varphi_*)(\varphi - \varphi_*) = b(\varphi - \varphi_*')^2 - c(\varphi + 2\varphi_*)(\varphi - \varphi_*')^2. \]
This equality, (2.3), (3.14), and Theorem 2.1 imply that
\[ \|g(\varphi) - g(\varphi_*) - g'(\varphi_*)(\varphi - \varphi_*)\|_{L^2} \]
\[ = \|b(\varphi - \varphi_*')^2 - c(\varphi + 2\varphi_*)(\varphi - \varphi_*')^2\|_{L^2} \]
\[ \leq C\|\varphi - \varphi_*\|_{L^2}^2 + C\|\varphi + 2\varphi_*\|_{L^2}\|\varphi - \varphi_*\|_{L^2}^2 \]
\[ \leq C\|\varphi - \varphi_*\|_{L^2}^2 \leq C\|f - f_*\|_{L^2}^2, \]
(3.20)
where \(C\) depends only on \(\|f\|_{L^2}, \|f_*\|_{L^2}\) and therefore only on \(\epsilon\). Inserting (3.20) into (3.19) and using (2.3) yields
\[ \|\delta u\|_{L^\infty} + \|\delta \varphi\|_{L^\infty} \leq C\|f - f_*\|_{L^2}^2. \]
(3.21)
We set \( \delta f = 0 \) and \( \delta v = (\delta u, \delta \varphi, \delta f) \).

From (3.18) and (3.21) we obtain the inequality
\[
|L_{ev}(v_*, \lambda_*)[\delta v, \delta \nu]| = \left| \int_0^T \langle (2b - 6c\varphi_*)\delta \varphi, \psi_* \rangle \, dt \right| + \int_0^T \alpha \langle \delta u, \delta u \rangle + \beta \langle \delta \varphi, \delta \varphi \rangle \, dt \\
\leq \|2b - 6c\varphi_*\|_{L^1} \|\delta \varphi\|_{L^1} \|\psi_*\|_{L^1} + \alpha \|\delta \nu\|^2 + \beta \|\delta \varphi\|^2 \\
\leq C\|f - f_*\|_{L^2}^2.
\]

(3.22)

Using similar estimates we find that
\[
|L_{ev}(v_*, \lambda_*)[\delta v, \delta \nu]| \leq \left| \int_0^T \langle (2b - 6c\varphi_*)(\delta \varphi)^2, \psi_* \rangle \, dt \right| \\
+ \int_0^T \alpha \|\delta u\|^2 + \beta \|\delta \varphi\|^2 + \gamma \|\delta f\|^2 \, dt \\
\leq C\|f - f_*\|_{L^2}^2.
\]

(3.23)

Combining the convexity assumption (3.11) and the estimates (3.17), (3.22), (3.23) we can deduce that
\[
J(v) \geq J(v_*) + \frac{1}{2} L_{ev}(v_*, \lambda_*)[v - v_*, v - v_*] + r(v_*, \lambda_*)[v - v_*, v - v_*] \\
= J(v_*) + \frac{1}{2} L_{ev}(v_*, \lambda_*)[\delta \nu, \delta \nu] + L_{ev}(v_*, \lambda_*)[\delta v, \delta \nu] \\
+ \frac{1}{2} L_{ev}(v_*, \lambda_*)[\delta v, \delta \nu] + r(v_*, \lambda_*)[v - v_*, v - v_*] \\
\geq J(v_*) + \sigma_* \|f - f_*\|_{L^2}^2 - C\|f - f_*\|_{L^2}^2.
\]

(3.24)

With (2.3) this gives the desired result.

\[\square\]

We conclude this section with a brief discussion of the first and second order remainder terms arising in the Taylor expansion of the Nemytskiĭ–operator "\( g \)" defined by \( \varphi \rightarrow g(\varphi) \).

Since \( g \) is a polynomial of degree three we can view "\( g \)" as an operator from \( L^6(Q) \) into \( L^3(Q) \). In this setting "\( g \)" is infinitely often differentiable. Note that we apply "\( g \)" to solutions \( \varphi \) of parabolic equations. The regularity of solutions and the imbedding theorem guarantee that \( \varphi \in W^{2,1}_2(Q) \subset L^6(Q) \). We make use of this particular form and this will enable us to derive convergence theorem for \((u, \varphi, f)\) in the space \( W^{2,1}_2(Q) \times W^{2,1}_2(Q) \times L^2(Q) \). This is a reasonable choice if the admissible set \( \mathcal{F}_{ad} \) is unbounded.

However, we also may view the Nemytskiĭ–operator "\( g \)" as an operator in \( L^\infty(Q) \), which is the appropriate setting if \( \mathcal{F}_{ad} \) is bounded. This setting usually also has to be chosen if \( g \) is not a polynomial. In this case one has to overcome the so–called two–norm discrepancy: The second order sufficiency condition (3.11) can only be expected to hold in the \( L^2(Q) \)–norm, whereas differentiability is given only with respect to the \( L^\infty(Q) \)–norm.

In our situation the reason for having to cope with difficulties usually arising in the two–norm discrepancy is not the lack of differentiability in a space with weaker topology than \( L^\infty(Q) \), but the desire to obtain the convergence of controls in \( L^\infty(Q) \).

The first order remainder term \( r^{(1)}_g \) of the Nemytskiĭ–operator "\( g \)" is obtained from
\[
g(\varphi + h) - g(\varphi) = g'(\varphi)h + \int_0^1 (g'(\varphi + \tau h) - g'(\varphi))h \, d\tau
\]
and is given by
\[
r^{(1)}_g(\varphi, h) = \int_0^1 (g'(\varphi + \tau h) - g'(\varphi))h \, d\tau.
\]
From the differentiability of \( g \) in \( L^\infty(Q) \) it follows that \( \|r_g^{(1)}(\varphi, h)\|_{L^\infty}/\|h\|_{L^\infty} \to 0 \) as \( \|h\|_{L^\infty} \to 0 \). However, the weaker condition
\[
\frac{\|r_g^{(1)}(\varphi, h)\|_{L^2}}{\|h\|_{L^2}} \to 0 \quad \text{as} \quad \|h\|_{L^\infty} \to 0
\]
(3.24)
can also easily be shown. Similarly, one can show that the second order remainder term \( r_g^{(2)} \) given by
\[
r_g^{(2)}(\varphi, h) = \frac{1}{2} \int_0^1 (g''(\varphi + \tau h) - g''(\varphi)) h^2 \, d\tau
\]
obeys
\[
\frac{\|r_g^{(2)}(\varphi, h)\|_{L^2}}{\|h\|_{L^2}} \to 0 \quad \text{as} \quad \|h\|_{L^\infty} \to 0 .
\]
(3.25)
The conditions (3.24), (3.25) are the key observations needed to overcome the above mentioned two–norm discrepancy. Similar relations are also to be used if \( g \) were not a polynomial.

4 The Lagrange–SQP–Newton Method

The Lagrange–SQP–Newton method solves the nonlinear, non–convex optimal control problem (1.5), (1.6), (1.1), (1.2), (1.3), (1.4) through a sequence of linear–quadratic control problems.

We continue to use the notations (1.9) and (1.10). Moreover, we denote the current iterate by \((v_c, \lambda_c) = (u_c, \varphi_c, f_c, p_c, \psi_c)\) and the new iterate by \((v_+, \lambda_+) = (u_+, \varphi_+, f_+, p_+, \psi_+)\).

In the Lagrange–SQP–Newton method the new iterate is computed as the solution of the following minimization problem:

\[
\text{Minimize } q_c(v) = J_c(v_c)(v - v_c) + \frac{1}{2} \mathcal{L}_{wv}(v_c, \lambda_c) [v - v_c, v - v_c] \quad \quad (4.1)
\]
subject to the linearized state equation
\[
\frac{\partial}{\partial t} u + \frac{\ell}{2} \frac{\partial}{\partial t} \varphi = \kappa \Delta u + f \quad \quad \text{in } \Omega \times (0,T],
\]
(4.2)
\[
\tau \frac{\partial}{\partial t} \varphi = \xi^2 \Delta \varphi + g(\varphi_c) + \ell g'(\varphi_c) (\varphi - \varphi_c) + 2u
\]
(4.3)
with boundary conditions
\[
\frac{\partial}{\partial n} u = 0, \quad \frac{\partial}{\partial n} \varphi = 0, \quad \text{on } \partial \Omega \times (0,T),
\]
(4.4)
and initial conditions
\[
u = u_0, \quad \varphi = \varphi_0 \quad \text{in } \Omega,
\]
(4.5)
and subject to the control constraints
\[
f \in \mathcal{F}_{ad}.
\]
(4.6)

Lemma 4.1 Let the assumptions (A1)–(A3) be valid. If \((u_c, \varphi_c) \in W^{2,1}_q(Q) \) and \( f_+ \in L^q(Q) \), then there exists a unique solution \( v_+ = (u_+, \varphi_+, f_+) \) of (4.2)–(4.5). The solution obeys
\[
\|u_+\|_{W^{2,1}_q} + \|\varphi_+\|_{W^{2,1}_q} \leq C(\|u_0\|_{W^{2,1}_q(\Omega)} + \|\varphi_0\|_{W^{2,1}_q(\Omega)} + \|f_+\|_{L^q} + \|\varphi_c\|_{W^{2,1}_q}).
\]
(4.7)
The constant \( C \) depends only on \( \varphi_c \) and can be chosen independent of \( \varphi_c \) if \( \varphi_c \) is contained in a \( W^{2,1}_2 \)–bounded set.
Proof: The pair \((u_+, \varphi_+)\) satisfies the system (2.4) with \(f_1 = f_+, f_2 = g(\varphi_+ - g'(\varphi_+))\varphi_+\), \(\beta_1 = 0\), and \(\beta_2 = g'(\varphi_+)\). Since \(q \in [2, 5/2)\) implies \(5q/(5 - 2q) \geq 3q\) the imbedding Theorem 2.2 implies that \((u_+, \varphi_+) \in L^{3q}(Q)\) if \(q \in [2, 5/2)\), that \((u_+, \varphi_+) \in L^r(Q)\) for all \(\nu \in [2, \infty)\) if \(q = 5/2\), and that \((u_+, \varphi_+) \in L^\infty(Q)\) if \(q > 5/2\). Moreover, since \(g\) is a polynomial of degree three with coefficients in \(L^\infty(Q)\) it holds that

\[
\|f_2\|_{L^r} = \|g(\varphi_+ - g'(\varphi_+))\varphi_+\|_{L^r} \leq C\|\varphi_+\|_{L^{3q}} \leq C\|\varphi_+\|_{W^{2,1}}.
\]

Inserting this estimate into Theorem 2.4 we obtain the desired result. \(\square\)

The objective of the quadratic subproblem is given by

\[
q_c(u, \varphi, f) = \int_0^T \int_\Omega \alpha(u - u_d)(u - u_c) + \beta(\varphi - \varphi_d)(\varphi - \varphi_c) + \gamma f_1(f - f_c) \, dxdt + \frac{1}{2} \int_0^T \int_\Omega \alpha(u - u_c)^2 + \beta(\varphi - \varphi_c)^2 + \gamma(f - f_c)^2 \, dxdt + \frac{1}{2} \int_0^T \int_\Omega \psi_c g''(\varphi - \varphi_c)^2 \, dxdt.
\]

Using standard techniques one can show that a solution \((u_+, \varphi_+, f_+)\) of the quadratic subproblem (4.1)–(4.6), if it exists, satisfies (4.2)–(4.6) and

\[
\int_0^T \int_\Omega (p_+(x, t) + \gamma f_+(x, t))(f(x, t) - f_+(x, t)) \, dxdt \geq 0 \quad \forall f \in \mathcal{F}_{ad},
\]

where \((p_+, \psi_+)\) is the solution of the adjoint equation for the linearized problem which is given by

\[
\begin{align*}
-\frac{\partial}{\partial t} p &= \kappa \Delta p + 2\psi + \alpha(u_+ - u_d) \quad \text{in } \Omega \times (0, T], \\
-\tau \frac{\partial}{\partial t} \psi - \frac{\partial}{\partial t} \frac{\partial}{\partial t} p &= \xi^2 \Delta \psi + g'(\varphi_+)\psi + \psi_c g'(\varphi_+ - \varphi_c) + \beta(\varphi_+ - \varphi_d),
\end{align*}
\]

with boundary conditions

\[
\frac{\partial}{\partial n} p = 0, \quad \frac{\partial}{\partial n} \psi = 0, \quad \text{on } \partial\Omega \times (0, T),
\]

and final conditions

\[
p(x, T) = 0, \quad \psi(x, T) = 0 \quad \text{in } \Omega.
\]

Analogously to Theorem 3.1 one can show that if \(u_d, \varphi_d \in L^\mu(Q), \mu \in [2, \infty)\), the adjoint system (4.10) has a unique solution \((p_+, \psi_+) \in W^{2,1}_\mu(Q) \times W^{2,1}_\mu(Q)\) which obeys

\[
\|p_+\|_{W^{2,1}_\mu} + \|\psi_+\|_{W^{2,1}_\mu} \leq C(\|u_+ - u_d\|_{L^\mu} + \|\varphi_+ - \varphi_d\|_{L^\mu} + \|\varphi_+ - \varphi_c\|_{L^\mu}).
\]

The constant \(C\) depends only on \(\varphi_c\) and can be chosen independent of \(\varphi_c\) if \(\varphi_c\) is contained in a \(W^{2,1}_2\)-bounded set. The parameter \(\mu\) is given by \(\mu = \min\{\mu, \frac{5}{2q}\}\) if \(q \in [2, 5/2)\) and \(n = 3\), and \(\mu = \mu\) if \(q > 5/2\) or \(n = 2\). Here \(q\) is determined by the control \(f_+ \in L^\mu(Q)\). Note that \(f_+ \in L^\infty(Q)\) if \(\mathcal{F}_{ad}\) is bounded and \(f_+ \in L^2(Q)\) otherwise.

Since \(\gamma > 0\), then (4.9) is equivalent to

\[
f_+(x, t) = P_T(-\gamma^{-1} p_+(x, t)).
\]

If \(q_c\) is convex, then the first order conditions are not only necessary but also sufficient. In the following we will establish the existence of a unique solution of the quadratic subproblem. This will be done by showing that the objective function \(q_c\) is strictly convex on the null space of the linearized constraints if \(v_t\) is close to a point \(v_*\) at which the second order sufficient optimality conditions are satisfied. The following lemma shows that the positivity of the Hessian of the Lagrangian (3.11) is preserved if \((v_*, \lambda_*)\) is replaced by a sufficiently close point \((v_c, \lambda_c)\).
Lemma 4.2 Let the assumptions (A1)–(A3) be valid and let \( v_* = (u_*, \varphi_*, f_*) \in V \) satisfy the second order sufficient optimality conditions. Then there exist \( \epsilon > 0 \) and \( \sigma > 0 \) such that
\[
\mathcal{L}_{uv}(v_*, \lambda_c)[v, v] \geq \sigma (||u||_{L^2}^2 + ||\varphi||_{L^2}^2 + ||f||_{L^2}^2)
\]  
(4.13)
for all \( v = (u, \varphi, f) \in V \) satisfying
\[
\begin{align*}
\frac{\partial}{\partial t} u + \frac{\partial}{\partial x} \varphi &= \kappa \Delta u + f, \\
\tau \frac{\partial}{\partial x} \varphi &= \xi^2 \Delta \varphi + g'(\varphi_c)\varphi + 2u,
\end{align*}
\]  
in \( \Omega \times (0, T) \),
(4.14)
with homogeneous initial and boundary conditions and for all \( v_* \), \( \lambda_c \) with \( ||v_* - v||_V + ||\lambda_c - \lambda_*||_A \leq \epsilon \).

Proof: Let \( \hat{v} = (\hat{u}, \hat{\varphi}, \hat{f}) \) satisfy the linearized state equation (3.12). Then, by Theorem 2.4,
\[
||\hat{u}||_{W^{1, 2}} + ||\hat{\varphi}||_{W^{1, 2}} \leq C||\hat{f}||_{L^2}.
\]  
(4.15)
The definitions of the Lagrangian and \( g \) give
\[
\mathcal{L}_{uv}(v_*, \lambda_c)[v, v] = \int_0^T \alpha ||u||^2 + \beta ||\varphi||^2 + \gamma ||f||^2 + \langle (2b - 6c\varphi_c)\varphi_*, \psi_\epsilon \rangle \, dt.
\]
cf. (3.1), (3.2). Thus,
\[
\begin{align*}
|\mathcal{L}_{uv}(v_*, \lambda_*)[\hat{v}, \hat{v}] - \mathcal{L}_{uv}(v_*, \lambda_*)[\hat{v}, \hat{v}]| \\
&\leq C\left(||\varphi_* - \varphi_*||_{L^2}||\hat{\varphi}||_{L^2} + ||\psi_* - \hat{\psi}||_{L^2}\right).
\end{align*}
\]
If \( \varphi_* \) is in an neighborhood of \( \varphi_* \), then \( ||2b - 6c\varphi_*||_{L^2} \leq C \). From this bound and (4.15) we find that
\[
|\mathcal{L}_{uv}(v_*, \lambda_*)[\hat{v}, \hat{v}] - \mathcal{L}_{uv}(v_*, \lambda_*)[\hat{v}, \hat{v}]| \\
\leq C\left(||\varphi_* - \varphi_*||_{L^2}||\psi_*||_{L^2} + ||\psi_* - \hat{\psi}||_{L^2}\right)||\hat{f}||_{L^2}.
\]
Thus, there exists \( \epsilon > 0 \) such that
\[
\mathcal{L}_{uv}(v_*, \lambda_*)[\hat{v}, \hat{v}] \geq \frac{3}{4} \sigma \left(||\hat{u}||_{L^2}^2 + ||\hat{\varphi}||_{L^2}^2 + ||\hat{f}||_{L^2}^2\right)
\]  
(4.16)
for all \( \hat{v} = (\hat{u}, \hat{\varphi}, \hat{f}) \) satisfying the linearized state equation (3.12) and for all \( v, \lambda \) with \( ||v_* - v||_V + ||\lambda_* - \lambda_*||_A \leq \epsilon \).

Let \( v = (u, \varphi, f) \) satisfy (4.14). Then \( \hat{v} - v \) satisfies (2.4) with \( f_1 = 0, f_2 = (g'(\varphi_c) - g'(\varphi_*))\varphi, \beta_1 = 0, \beta_2 = g'(\varphi_*), \) and homogeneous initial and boundary conditions. Hence, Theorem 2.4 yields
\[
\begin{align*}
||u - \hat{u}||_{W^{1, 2}} + ||\varphi - \hat{\varphi}||_{W^{1, 2}} &\leq C||g'(\varphi_* - g'(\varphi_*))\varphi||_{L^2} \\
&\leq C||g'(\varphi_* - g'(\varphi_*))||_{L^1}||\varphi||_{L^1} \\
&\leq C||2b(\varphi_* - \varphi_*) + 3c(\varphi_* + \varphi_*)(\varphi_* - \varphi_*)||_{L^1}||\hat{f}||_{L^2} \\
&\leq C\left(||\varphi_* - \varphi_*||_{L^1} + ||\varphi_* + \varphi_*||_{L^1}||\varphi_* - \varphi_*||_{L^1}\right)||\hat{f}||_{L^2} \\
&\leq C||v_* - v||_V ||\hat{f}||_{L^2}.
\end{align*}
\]  
(4.17)
Suppose that \( v_\epsilon, \lambda \) obey \( \| v_\epsilon - v_* \|_V + \| \lambda - \lambda_* \|_\Lambda \leq \tilde{\epsilon} \). Using (4.15), (4.16), (4.17), the imbedding Theorem 2.2, and the definition of \( \mathcal{L}_{ev} \), we find that

\[
\mathcal{L}_{ev}(v_\epsilon, \lambda_\epsilon)[v, v] \\
= \mathcal{L}_{ev}(v_\epsilon, \lambda_\epsilon)[\hat{v} + (v - \hat{v}), \hat{v} + (v - \hat{v})] \\
= \mathcal{L}_{ev}(v_\epsilon, \lambda_\epsilon)[\hat{v}, \hat{v}] + 2\mathcal{L}_{ev}(v_\epsilon, \lambda_\epsilon)[\hat{v}, (v - \hat{v})] + \mathcal{L}_{ev}(v_\epsilon, \lambda_\epsilon)[v - \hat{v}, v - \hat{v}] \\
\geq \frac{3}{4} \sigma_2 (\| \hat{v} \|_L^2 + \| \hat{\varphi} \|_L^2 + \| \hat{f} \|_L^2) \\
- C \int_0^T 2\left| \left( (2b - 6c\varphi_\epsilon)\varphi - \hat{\varphi}, \psi_\epsilon \right) \right| + \left| \left( (2b - 6c\varphi_\epsilon)(\varphi - \hat{\varphi})^2, \psi_\epsilon \right) \right| dt \\
\geq \frac{3}{4} \sigma_2 (\| \hat{v} \|_L^2 + \| \hat{\varphi} \|_L^2 + \| \hat{f} \|_L^2) \\
- C\| \hat{f} \|_L^2 \| v - v_* \|_V \| \hat{f} \|_L^2 - C\| v - v_* \|_V \| \hat{f} \|_L^2 . \tag{4.18}
\]

With the equality \( \hat{v} = v + (\hat{v} - v) \) and the estimate (4.17) the assertion follows from (4.18). \( \square \)

The definition (4.1) of \( q_\epsilon \) and Lemma 4.2 imply the following result:

**Corollary 4.3** Let the assumptions (A1)–(A3) be valid and let \( v_* = (u_*, \varphi_*, f_*) \in V \) satisfy the second order sufficient optimality conditions. Then there exists \( \epsilon > 0 \) such that \( q_\epsilon \) is strictly convex on the null space of the linearized constraints for all \( v_\epsilon, \lambda_\epsilon \) with \( \| v_\epsilon - v_* \|_V + \| \lambda_\epsilon - \lambda_* \|_\Lambda \leq \epsilon \).

The strict convexity of the objective function on the null space of the linearized constraints implies the existence of a unique solution of the linear quadratic control problem. Moreover, one can derive an estimate for the difference between new iterate and optimal point:

**Lemma 4.4** Let the assumptions (A1)–(A3) be valid. If \( v_* = (u_*, \varphi_*, f_*) \in V \) satisfies the second order sufficient optimality conditions, then there exists \( \epsilon > 0 \) such that if \( \| v_\epsilon - v_* \|_V + \| \lambda_\epsilon - \lambda_* \|_\Lambda \leq \epsilon \) the linear quadratic optimal control problem (4.1)–(4.6) has a unique solution \( v_+ = (u_+, \varphi_+, f_+) \) obeying

\[
\| v_+ - v_* \|_V \leq C \left( \| v_\epsilon - v_* \|_V + \| \lambda_\epsilon - \lambda_* \|_\Lambda \right)^{1/2} \tag{4.19}
\]

and

\[
\| \lambda_+ \|_\Lambda \leq C .
\]

**Proof:** (i) The existence and uniqueness of \( v_+ \) follows from standard arguments using the convexity of the problem established in Corollary 4.3.

In the following we consider points \( v_\epsilon, \lambda_\epsilon \) with

\[
\| v_\epsilon - v_* \|_V + \| \lambda_\epsilon - \lambda_* \|_\Lambda \leq \min \{ 1, \tilde{\epsilon} \} , \tag{4.20}
\]

where \( \tilde{\epsilon} \) is determined by Corollary 4.3.

(ii) First, we derive a bound for the value of the objective of the subproblem at the new iterate.

Let \((u, \varphi)\) be the solution of (4.2)–(4.5) with \( f = f_* \). We set \( v = (u, \varphi, f_\epsilon) \). Since \( v \) is a feasible point for (4.1)–(4.6), we find that \( q_\epsilon(v_+) \leq q_\epsilon(v) \). Using the definition (4.8) of \( q_\epsilon \), the imbedding \( W^{2,1}_2(Q) \subset L^{10}(Q) \), and H"older’s inequality we can conclude that

\[
q_\epsilon(v_+) \leq q_\epsilon(v) \leq C (\| v - v_* \|_V + \| v - v_* \|_V^2 ) . \tag{4.21}
\]
Next we estimate \( v - v_* \). The pair \( (u - u_*, \varphi - \varphi_*) \) satisfies (2.4) with \( \beta_1 = 0, \beta_2 = g'(\varphi_*) \), \( f_1 = f_c - f_*, \)
\( f_2 = g(\varphi_*) - g(\varphi_*) + g'(\varphi_*)(\varphi - \varphi_*) \), and homogeneous initial and boundary conditions. Using the definition of \( g \),
Theorem 2.4, and the imbedding Theorem 2.2, we can show that
\[
\|u - u_*\|_{W^{2,1}_2} + \|\varphi - \varphi_*\|_{W^{2,1}_2} \leq C(\|f_c - f_*\|_{L^2} + \|\varphi_c - \varphi_*\|_{W^{2,1}_2}) \leq C\|v - v_*\|_V. \tag{4.22}
\]
Cf. (3.20). Thus,
\[
\|v - v_*\|_V \leq \|v - v_*\|_V + \|v_* - v_c\|_V \leq C\|v_* - v_*\|_V.
\]
Inserting this inequality into (4.21) and using \( \|v_* - v_*\|_V < 1 \) yields
\[
q_c(v_+) \leq C(\|v_* - v_*\|_V + \|v_* - v_*\|_V^2) \leq C\|v_* - v_*\|_V. \tag{4.23}
\]

(iii) Next we prove the uniform boundedness of the new iterates.
If the set of admissible controls \( \mathcal{F}^{ad} \) is bounded, then Lemma 4.1 and (4.20) imply the boundedness of the new iterate.
If the admissible set \( \mathcal{F}^{ad} \) is unbounded, we can use the convexity of \( q_c \) on the null–space of the linearized constraints to establish the boundedness of the iterate.
Let \( (u_+^0, \varphi_+^0) \) be the solution of (4.14) with \( f = f_+ \) and homogeneous initial and boundary conditions. Moreover, let \( (u_+^1, \varphi_+^1) \) solve
\[
\begin{align*}
\frac{\partial u}{\partial t} + \frac{\partial}{\partial t} \varphi &= \kappa \Delta u \\
\tau \frac{\partial \varphi}{\partial t} &= \xi_2 \Delta \varphi + g(\varphi_*) - g(\varphi_* + 2u), \quad \text{in} \ \Omega \times (0, T],
\end{align*}
\]
with initial conditions (4.5) and boundary conditions (4.4). Then
\[
(u_+, \varphi_+, f_+) = (u_+^0, \varphi_+^0, f_+) + (u_+^1, \varphi_+^1, 0).
\]
Using Theorem 2.4 we find that
\[
\|u_+^1\|_{W^{2,1}_2} + \|\varphi_+^1\|_{W^{2,1}_2} \leq C(\|u_0\|_{W^{2,1}_2} + \|\varphi_0\|_{W^{2,1}_2} + \|\varphi_*\|_{W^{2,1}_2}). \tag{4.24}
\]
Using the definition (4.1) of \( q_c \), the convexity of the Lagrangian (4.13), and the bounds (4.23), (4.24) we find that with some \( C > 0 \) depending on \( \xi \), but independent of \( v_* \) the inequalities
\[
\begin{align*}
&\sigma(\|u_0\|^2_{L^2} + \|\varphi_0^2\|_{L^2} + \|f_1\|^2_{L^2}) - C(1 + \|u_+^1\|^2_{L^2} + \|\varphi_+^1\|^2_{L^2} + \|u_*\|^2_{L^2} + \|\varphi_*\|^2_{L^2} + \|f_*\|^2_{L^2}) \\
&\leq q_c(v_+) \\
&\leq C\|v_* - v_*\|_V.
\end{align*}
\]
are valid. This gives the boundedness of \( \|u_+^0\|^2_{L^2}, \|\varphi_+^0\|^2_{L^2}, \) and \( \|f_+\|^2_{L^2} \). With Lemma 4.1 we can deduce the bound
\[
\|v_*\|_V \leq C(1 + \|v_*\|_{V^2}). \tag{4.25}
\]
This shows that the new iterates are bounded, if (4.20) holds true.

(iv) The definition of \( q_c \) yields
\[
q_c(v_+) = J_0(v_+ - v_*) + \frac{1}{2} L_{vv}(v_*, \lambda_*)[v_+ - v_c, v_+ - v_c] = J_0(v_+ - v_*) + J_0(v_+ - v_*) + \frac{1}{2} L_{vv}(v_*, \lambda_*)[v_+ - v_*, v_+ - v_*] \\
+ L_{vv}(v_*, \lambda_*)[v_+ - v_*, v_+ - v_*] + \frac{1}{2} L_{vv}(v_*, \lambda_*)[v_+ - v_*, v_+ - v_*] \\
\geq J_0(v_+ - v_*) + \frac{1}{2} L_{vv}(v_*, \lambda_*)[v_+ - v_*, v_+ - v_*] - C\|v_c - v_*\|_V. \tag{4.26}
\]
In the last estimate of (4.26) we used the boundedness of the new iterate, Hölder’s inequality and the imbedding
Theorem 2.2. These estimates are analogous to (4.18).

We will show in (v) that the estimates

\[ J_e(v_\ast)(v_\ast - v_*) \geq -C\|v_\ast - v_*\|_V, \quad L_{\nu}(v_\ast, \lambda)[v_\ast - v_\ast, v_\ast - v_*] \geq \frac{\sigma_s}{2}\|v_\ast - v_\ast\|^2_V - C(\|v_\ast - v_*\|_V + \|\lambda_e - \lambda_*\|_\Lambda), \]

hold. Using the estimates (4.27), (4.28) in (4.26) yields

\[ q_e(v_\ast) \geq \frac{\sigma_s}{2}\|v_\ast - v_\ast\|^2_V - C\|v_\ast - v_*\|_V - C\|\lambda_e - \lambda_*\|_\Lambda. \]

If we combine (4.23) and (4.29) we obtain the desired Hölder estimate (4.19).

(v) For the proof of (4.27) we proceed as follows:

From the necessary optimality conditions for the nonlinear optimal control problem we find that

\[ J_e(v_\ast)(v - v_\ast) \geq 0 \]

for all \( v = (u, \varphi, f) \) with \( f \in F_{ad} \) and

\[ \frac{\partial}{\partial t}(u - u_\ast) + \frac{\partial}{\partial t}(\varphi - \varphi_\ast) = \kappa \Delta (u - u_\ast) + (f - f_\ast), \]

\[ \tau \frac{\partial}{\partial t}(\varphi - \varphi_\ast) = \xi^2 \Delta (\varphi - \varphi_\ast) + g'(\varphi_\ast)(\varphi - \varphi_\ast) + 2(u - u_\ast), \quad \text{in } \Omega \times (0, T], \]

with homogeneous initial and boundary conditions.

Let \( \hat{v} = (\hat{u}, \hat{\varphi}, \hat{f}_\ast) \) be defined by (4.31) with \( f = f_\ast \).

The triple \( v_\ast - v_\ast = (u_\ast - u_\ast, \varphi_\ast - \varphi_\ast, f_\ast - f_\ast) \) satisfies (2.4) with \( \beta_1 = 0, \beta_2 = g'(\varphi_\ast), f_1 = f_\ast - f_\ast, f_2 = g(\varphi_\ast) - g(\varphi_\ast) - g'(\varphi_\ast)(\varphi_\ast - \varphi_\ast) + g'(\varphi_\ast)(\varphi_\ast - \varphi_\ast), \)

and homogeneous initial and boundary conditions.

The triple \( v_* - \hat{v} = (v_* - \hat{v}) = (\hat{v} - v_*), \hat{f}_\ast \) satisfies (2.4) with \( \beta_1 = 0, \beta_2 = g'(\varphi_\ast), f_1 = 0, f_2 \) given by (4.32), and homogeneous initial and boundary conditions. Using

\[ f_2 = g(\varphi_\ast) - g(\varphi_\ast) - g'(\varphi_\ast)(\varphi_\ast - \varphi_\ast) - g'(\varphi_\ast)(\varphi_\ast - \varphi_\ast) + g'(\varphi_\ast)(\varphi_\ast - \varphi_\ast), \]

the boundedness of \( v_\ast, v_\ast \), the definition of \( g \) and Hölder inequality, we can show that

\[ \|f_2\|_{L^2} \leq C\|\varphi_\ast - \varphi_*\|_{W^{2,1}} \leq C\|v_\ast - v_*\|_V. \]

Thus, Theorem 2.4 yields

\[ \|(v_\ast - v_\ast) - (\hat{u} - v_*))\| = \|v_\ast - \hat{u}\|_V \leq C\|v_\ast - v_*\|_V. \]

Combining (4.30), (4.33) and using (4.25) yields

\[ J_e(v_\ast)(v_\ast - v_\ast) = J_e(v_\ast)(v_\ast - v_\ast) + (J_e(v_\ast) - J_e(v_\ast))(v_\ast - v_\ast), \]

\[ \geq J_e(v_\ast)(v_\ast - v_\ast) - C\|v_\ast - v_*\|_V, \]

\[ = J_e(v_\ast)(v_\ast - v_\ast) + J_e(v_\ast)(v_\ast - v_\ast) - (\hat{v} - v_\ast) - C\|v_\ast - v_*\|_V, \]

\[ \geq J_e(v_\ast)(\hat{v} - v_\ast) - C\|v_\ast - v_*\|_V. \]

This proves (4.27).
The estimate (4.28) can be derived in a similar way. The boundedness of $\|v_+ - v_*\|_V$ yields
\[
L_{vv}(v_*, \lambda_*)[v_+ - v_*] \geq L_{vv}(v_*, \lambda_*)[v_+ - v_*] - C(\|v_+ - v_*\|_V + \|\lambda_+ - \lambda_*\|_\lambda),
\]
(4.34) cf. (4.18). As before, let $\hat{v} = (\hat{u}, \hat{\phi}, \hat{f}_+)$ be defined by (4.31) with $f = f_+$, i.e. $\hat{v} - v_*$ satisfies the linearized state equation. Using (3.11) we can show that
\[
L_{vv}(v_*, \lambda_*)[v_+ - v_*],
\]
\[
\geq L_{vv}(v_*, \lambda_*)[\hat{v} - v_*] - C(\|\hat{v} - v_*\|_V),
\]
(4.35) In the last two estimates of (4.35) we have used the boundedness of $v_+$ and the inequality (4.33). If we use $\hat{v} = v_+ + (\hat{v} - v_+)$, (4.25), and (4.33), the inequality (4.35) yields
\[
L_{vv}(v_*, \lambda_*)[v_+ - v_*] \geq \sigma_+ \|\hat{v} - v_*\|_V - C(\|v_+ - v_*\|_V),
\]
(4.36) This lower bound, together with (4.34) yields the desired estimate (4.28).

(vi) It remains to prove the estimate for the Lagrange multiplier $\lambda_+$. This estimate follows easily from (4.11), (4.20), and (4.25). In fact, we find that
\[
\|p_+\|_{W^{2,1}_x} + \|\psi'_+\|_{W^{2,1}_x} \leq C(\|u_+ - u_d\|_{L^2} + \|\varphi_+ - \varphi_d\|_{L^2} + \|\varphi_+ - \varphi_d\|_{L^2}) \leq C.
\]
This concludes the proof. \hfill \Box

In the following we improve the estimate (4.19). We will show that the error in the new iterate can be bounded even in a stronger norm. The proof is based on the regularity estimates for the system (2.4). We have to require that the iterates, the optimal point and the Lagrange multipliers satisfy $v_+, v_+, v_* \in V_\gamma$ and $\lambda_+, \lambda_+, \lambda_* \in L_\gamma$, where $V_\gamma$ and $L_\gamma$ are defined in (1.10). The parameter $q > 5/2$ if $n = 3$ and $q > 2$ if $n = 2$. The regularity of states and adjoints can be guaranteed if the initial iterate is sufficiently smooth, if the desired temperature and phase function obey $u_d, \varphi_d \in L^6(Q)$, and if $f_+, f_+, f_* \in L^6(Q)$, cf. Theorems 2.1, 3.1 and equations (4.7), (4.11). Since $\gamma > 0$, the conditions $f_+, f_+, f_* \in L^6(Q)$ are implied by the regularity $\lambda_+, \lambda_* \in L_\gamma$ of the adjoint variables, cf. (4.9), (4.12).

**Lemma 4.5** Let the assumptions (A1)–(A3) be valid, let $F_{ad}$ be bounded, $\gamma > 0$, and suppose that the second order sufficient optimality conditions are satisfied at $v_* = (u_*, \varphi_*, f_*) \in V$.

If $q \in [2, \infty)$ the iterates and the optimal point satisfy $v_+, v_+, v_* \in V_\gamma$, then there exists $\epsilon > 0$ such that $\|v_+ - v_*\|_V + \|\lambda_+ - \lambda_*\|_\lambda \leq \epsilon$ implies the estimates
\[
\|v_+ - v_*\|_V \leq C(\|v_+ - v_*\|_V + \|\lambda_+ - \lambda_*\|_\lambda)^{1/q},
\]
(4.37) and
\[
\|\lambda_+ - \lambda_*\|_\lambda \leq C(\|v_+ - v_*\|_V + \|\lambda_+ - \lambda_*\|_\lambda)^{1/q}.
\]
(4.38) Moreover, if $n = 3$ and $q > 5/2$ or if $n = 2$ and $q > 2$, then (4.37) can be replaced by
\[
\|u_+ - u_*\|_{W^{2,1}_x} + \|\varphi_+ - \varphi_*\|_{W^{2,1}_x} + \|f_+ - f_*\|_{L^\infty} \leq C(\|v_+ - v_*\|_V + \|\lambda_+ - \lambda_*\|_\lambda)^{1/q}.
\]
(4.39)
Proof: Using (4.19) and the boundedness of $F_{ad}$ we can conclude that
\[
\|f_+ - f_*\|_{L^2}^2 = \int_0^T \int_{\Omega} |f_+ - f_*|^2 |f_+ - f_*|^{q-2} \, dx \, dt \leq C \|f_+ - f_*\|_{L^2}^2 \leq C \left( \|v_* - v_*\|_V + \|\lambda_* - \lambda_*\|_{\Lambda} \right).
\]
(4.40)
The difference $v_+ - v_*$ satisfies (2.4) with $f_1 = f_+ - f_*$, $f_2 = g(\varphi_*) - g(\varphi_*) + g'(\varphi_*)(\varphi_* - \varphi_*)$, $\beta_1 = 0$, $\beta_2 = \beta$, and homogeneous initial and boundary conditions. Using the definition of $g$, the right hand side $f_2$ can be estimated by
\[
\|f_2\|_{L^q} \leq C \|\varphi_* - \varphi_*\|_{L^q}.
\]
Thus, from Theorem 2.4 and (4.40) we obtain
\[
\|u_+ - u_*\|_{W^{2,1}} + \|\varphi_+ - \varphi_*\|_{W^{2,1}} \leq C(\|u_+\|_{L^q} + \|f_2\|_{L^q})
\]
\[
\leq C(\|f_+ - f_*\|_{L^q} + \|\varphi_* - \varphi_*\|_{L^q})
\]
\[
\leq C \left( \|v_* - v_*\|_V + \|\lambda_* - \lambda_*\|_{\Lambda} \right)^{1/q} + C \|\varphi_* - \varphi_*\|_{L^q}.
\]
(4.41)
Note that $W^{2,1}_q(Q) \subset L^{3q}(Q)$ for $q \in [2, 5/2)$ and $W^{2,1}_q(Q) \subset L^r(Q)$ for arbitrary $\nu \in [2, \infty)$ otherwise, cf. Theorem 2.2. If $\epsilon < 1$, then (4.41) yields
\[
\|u_+ - u_*\|_{W^{2,1}} + \|\varphi_+ - \varphi_*\|_{W^{2,1}} \leq C(\|v_* - v_*\|_V + \|\lambda_* - \lambda_*\|_{\Lambda})^{1/q}
\]
(4.42)
for all $v_*, \lambda_*$ with $\|v_* - v_*\|_V + \|\lambda_* - \lambda_*\|_{\Lambda} < \epsilon$.

The adjoints can be estimated in a similar way. In fact, the difference $\lambda_+ - \lambda_*$ satisfies
\[
-\frac{\partial}{\partial t} p = \kappa \Delta p + 2\psi + \alpha(u_+ - u_*),
\]
\[
-\tau \frac{\partial}{\partial t} \psi + \frac{\partial}{\partial x} p = \xi^2 \Delta \psi + g'(\varphi_*)\psi + (g'(\varphi_*) - g'(\varphi_*))\psi_+ + g''(\varphi_*)(\varphi_+ - \varphi_*) + \beta(\varphi_+ - \varphi_*),
\]
(4.43)
in $\Omega \times (0,T]$ with homogeneous initial and boundary conditions. Using the boundedness of the iterates and of the Lagrange multipliers in $V$ and $\Lambda$, respectively, and Theorem 2.4 this yields
\[
\|p_+ - p_*\|_{W^{2,1}} + \|\psi_+ - \psi_*\|_{W^{2,1}} \leq C(\|f_+ - f_*\|_{L^q} + \|\varphi_* - \varphi_*\|_{L^q})
\]
\[
\leq C \left( \|v_* - v_*\|_V + \|\lambda_* - \lambda_*\|_{\Lambda} \right)^{1/q} + C \|\varphi_* - \varphi_*\|_{L^q}
\]
\[
\leq C \left( \|v_* - v_*\|_V + \|\lambda_* - \lambda_*\|_{\Lambda} \right)^{1/q},
\]
(4.44)
provided $\|v_* - v_*\|_V + \|\lambda_* - \lambda_*\|_{\Lambda} \leq 1$.

Suppose that $n = 3$, $q > 5/2$ or $n = 2, q > 2$. For the estimate of $\|f_+ - f_*\|_{L^q}$ we use the first order optimality conditions. From (3.10), (4.12), (4.42), and the imbedding $W^{2,1}_q(Q) \subset L^\infty(Q)$ we can deduce that
\[
\left| f_+(x,t) - f_+(x,t) \right| = \left| P_F(-\gamma^{-1}p_+(x,t)) - P_F(-\gamma^{-1}p_+(x,t)) \right|
\]
\[
\leq \gamma^{-1} \left| p_+(x,t) - p_+(x,t) \right|
\]
\[
\leq C \left( \|v_* - v_*\|_V + \|\lambda_* - \lambda_*\|_{\Lambda} \right)^{1/q}.
\]
(4.45)
5 A Perturbed Quadratic Problem

In the convergence proof of the Lagrange–SQP–Newton method we will view the SQP subproblem (4.1) to (4.6) as a perturbation of the SQP subproblem at the strict local minimizer \((u_*, \varphi_*, f_*)\). To obtain convergence estimates we have to study the dependence of the solution upon the perturbation \(\pi = (\pi_s, \pi_a) \in (L^2(\Omega))^2\). Here \(\pi_s\) will be a perturbation on the right hand side of the linearized state equation and \(\pi_a\) will be a perturbation in the objective function, which relates to a perturbation in the adjoint equation.

The perturbed quadratic subproblem is given by:

\[
\min \; q_s(v) + \langle \pi_a, \varphi \rangle_{L^2} = J_v(v_*)(v - v_*) + \frac{1}{2} \mathcal{L}_{v,v}(v_*, \lambda_*)[v - v_*] + \langle \pi_a, \varphi \rangle_{L^2}
\]  

subject to the linearized state equation

\[
\frac{\partial}{\partial t} u + \frac{\ell}{2} \frac{\partial}{\partial t} \varphi = \kappa \Delta u + f
\]  
in \(\Omega \times (0,T]\),

\[
\tau \frac{\partial}{\partial t} \varphi = \xi^2 \Delta \varphi + g(\varphi_*) + g'(\varphi_*)(\varphi - \varphi_*) + 2u + \pi_s
\]  

with boundary conditions

\[
\frac{\partial}{\partial n} u = 0, \quad \frac{\partial}{\partial n} \varphi = 0, \quad \text{on} \; \partial \Omega \times (0,T),
\]

and initial conditions

\[
u = u_0, \quad \varphi = \varphi_0 \quad \text{in} \; \Omega,
\]

and subject to the control constraints

\[f \in \mathcal{F}_{ad}.
\]  

We recall, cf. (4.8), that the objective of the quadratic subproblem is given by

\[
q_s(u, \varphi, f) + \langle \pi_a, \varphi \rangle_{L^2} = \int_0^T \int_\Omega \alpha(u_* - u_d)(u - u_*) + \beta(\varphi_* - \varphi_d)(\varphi - \varphi_*) + \gamma f_*(f - f_*) \, dxdt + \frac{1}{2} \int_0^T \int_\Omega (u_*)^2 + \beta(\varphi_*^2) + \gamma(f - f_*)^2 \, dxdt + \frac{1}{2} \int_0^T \int_\Omega \psi_t g''(\varphi_*)(\varphi - \varphi_*)^2 \, dxdt + \int_0^T \int_\Omega \pi_a \varphi \, dxdt.
\]

Since the second order sufficient optimality conditions are satisfied at \((u_*, \varphi_*, f_*)\) the objective function is strictly convex and one can use standard techniques to show the existence of a unique solution \((u_{\pi}, \varphi_{\pi}, f_{\pi})\). Of course, if \(\pi_s = \pi_a = 0\), then \((u_{\pi}, \varphi_{\pi}, f_{\pi}) = (u_*, \varphi_*, f_*)\). Moreover, one can show that the solution \((u_{\pi}, \varphi_{\pi}, f_{\pi})\) of the quadratic subproblem (5.1)–(5.5) satisfies (5.2)–(5.5) and

\[
\int_0^T \int_\Omega \left( p_{\pi}(x,t) + \gamma f_*(x,t) \right) \left( f(x,t) - f_{\pi}(x,t) \right) \, dxdt \geq 0 \quad \forall f \in \mathcal{F}_{ad},
\]

where \(\lambda_{\pi} = (p_{\pi}, \psi_{\pi})\) is the solution of the adjoint equation for the linearized problem which is given by

\[
-\frac{\partial}{\partial t} p = \kappa \Delta p + 2\psi + \alpha(u_* - u_d),
\]  
in \(\Omega \times (0,T]\),

\[
-\tau \frac{\partial}{\partial t} \psi + \frac{\ell}{2} \frac{\partial}{\partial t} p = \xi^2 \Delta \psi + g'(\varphi_*) \psi + \psi_t g''(\varphi_*)(\varphi - \varphi_*) + \beta(\varphi_* - \varphi_d) + \pi_a,
\]

with boundary conditions

\[
\frac{\partial}{\partial n} p = 0, \quad \frac{\partial}{\partial n} \psi = 0, \quad \text{on} \; \partial \Omega \times (0,T),
\]
and final conditions
\[ p(x, T) = 0, \quad \psi(x, T) = 0 \quad \text{in } \Omega. \]

If \( \gamma > 0 \), then (5.7) is equivalent to
\[ f_\pi(x, t) = P_\pi(-\gamma^{-1} p_\pi(x, t)). \]  
(5.9)

Using Theorem 2.4 one can see that the solution \((u_\pi, \varphi_\pi)\) of (5.2) satisfies
\[
\|u_\pi\|_{W^{2,1}_0} + \|\varphi_\pi\|_{W^{2,1}_0} \\
\leq C\left(\|u_0\|_{W^{2,1}_0(\Omega)} + \|\varphi_0\|_{W^{2,1}_0(\Omega)} + \|f_\pi\|_{L^q} + \|\pi_s\|_{L^s}\right),
\]  
(5.10)

provided \(f_\pi, \pi_s \in L^q(Q)\). Moreover, if \(\pi_0 \in L^q(Q)\), then the solution of the adjoint equation (5.8) satisfies
\[
\|p_\pi\|_{W^{2,1}_0} + \|\psi_\pi\|_{W^{2,1}_0} \\
\leq C\left(\|u_\pi - u_*\|_{L^q} + \|\varphi_\pi - \varphi_*\|_{W^{2,1}_0} + \|\varphi_\pi - \varphi_*\|_{L^q} + \|\pi_0\|_{L^s}\right),
\]  
(5.11)

where the parameter \( \nu \) is given as in (4.11).

**Lemma 5.1** Let the assumptions (A1)–(A3) be valid. If \(u_\pi, u_* , \varphi_\pi, \varphi_* \in W^{2,1}_0(Q)\), then there exists a constant \(C > 0\) such that
\[
\|\lambda_\pi - \lambda_*\|_{\Lambda_q} \leq C\left(\|u_\pi - u_*\|_{W^{2,1}_0} + \|\varphi_\pi - \varphi_*\|_{W^{2,1}_0} + \|\pi_0\|_{L^s}\right).
\]

**Proof:** The difference \((p, \psi) = \lambda = (\lambda_\pi - \lambda_*)\) satisfies the system
\[
-\frac{\partial}{\partial t} p + \kappa \Delta p + 2\psi + \alpha(u_\pi - u_*), \\
-\tau \frac{\partial}{\partial t} \psi + \frac{\ell}{2} \frac{\partial}{\partial t} p = \xi^2 \Delta \psi + g'(\varphi_\pi)\psi + \beta(\varphi_\pi - \varphi_*) + \psi_\pi g''(\varphi_\pi - \varphi_*) + \pi_s, 
\]
in \(\Omega \times (0, T]\),

with homogeneous initial and boundary conditions. The assumptions on \(u_\pi, u_* , \varphi_\pi, \varphi_*\) imply \(u_\pi, u_* \in L^q(Q)\) and, since \(5q/(5 - 2q) > 3q\) for \(q \in [2, 5/2)\), \(\|\psi_\pi, g''(\varphi_\pi)\|_{L^q} \leq \|\psi_\pi\|_{L^q} + \|g''(\varphi_\pi)\|_{L^{3q}} \|\varphi_\pi - \varphi_*\|_{L^q}\). Thus, the assertion follows from Theorems 2.2, 2.4.

The next statement is rather standard for linear-quadratic control problems. We present the proof for convenience.

**Lemma 5.2** Let the assumptions (A1)–(A3) be valid. There exists a constant \(C > 0\) such that
\[
\|u_\pi - u_*\|_{V} \leq C\|\pi\|_{(L^2)^2}
\]
for all \(\pi \in (L^2(Q))^2\).

**Proof:** The Lagrange function for the perturbed problem (5.1) to (5.5) is given by
\[
\bar{L}(v, \lambda_\pi) = I_v(v_\pi)(v - v_\pi) + \frac{1}{2} L_{wv}(v_\pi, \lambda_\pi)[v - v_\pi, v - v_\pi] + \langle \pi_s, \varphi \rangle_{L^2} \\
- \int_0^T \langle u_\pi + \frac{\ell}{2} \varphi_\pi, \psi_\pi \rangle + \kappa \langle \nabla u, \nabla \psi_\pi \rangle - \langle f, \psi_\pi \rangle \, dt \\
- \int_0^T \tau \langle \varphi_\pi, \psi_\pi \rangle + \xi^2 \langle \nabla \varphi_\pi, \nabla \psi_\pi \rangle - \langle g(\varphi_\pi), g'(\varphi_\pi)(\varphi - \varphi_*) + 2u + \pi_s, \psi_\pi \rangle \, dt.
\]  
(5.12)

The first order necessary conditions for (5.1) to (5.5) are given by
\[
\bar{L}_v(v_\pi, \lambda_\pi)(v - v_\pi) \geq 0
\]
for all \( v = (u, \varphi, f) \) with \( f \in \mathcal{F}_{ad} \). Using the definition (5.12) and inserting \( v = \upsilon \), this leads to

\[
0 \leq J_v(\upsilon_v)(\upsilon_v - \upsilon) + \mathcal{L}_{uv}(\upsilon_v, \lambda_v)[\upsilon_v - \upsilon_v, \upsilon_v - \upsilon_v] + \langle \pi_v, \varphi_v - \varphi_v \rangle_{L^2} \\
- \int_0^T \left< (u_\pi - u), \frac{\ell}{2} (\varphi - \varphi_v), \frac{\ell}{2} (\varphi - \varphi_v), p \right> + \kappa \left< \nabla (u_\pi - u), \nabla p \right> - \left< (f - f_\pi), p \right> dt \\
- \int_0^T \tau \left< (\varphi_v - \varphi_v), \psi_v \right> + \xi^2 \left< \nabla (\varphi_v - \varphi_v), \nabla \psi_v \right> - \left< g' \left( \varphi_v - \varphi_v \right), 2(\upsilon_v - \upsilon_v), \psi_v \right> dt.
\]

(5.13)

The difference \( \upsilon = \upsilon_v - \upsilon = (u_\pi - u, \varphi_v - \varphi_v, f - f_\pi) \) satisfies

\[
\frac{\partial}{\partial t} \upsilon + \frac{\ell}{2} \frac{\partial}{\partial \varphi} \varphi = \kappa \Delta \varphi + f, \\
\frac{\partial}{\partial t} \varphi = \xi^2 \Delta \varphi + g' \left( \varphi \right) \varphi + 2\upsilon - \pi_v, \quad \text{in } \Omega \times (0, T],
\]

(5.14)

with homogeneous initial and boundary conditions. Hence, (5.13) is equivalent to

\[
0 \leq J_v(\upsilon_v)(\upsilon_v - \upsilon) + \mathcal{L}_{uv}(\upsilon_v, \lambda_v)[\upsilon_v - \upsilon_v, \upsilon_v - \upsilon_v] + \langle \pi_v, \varphi_v - \varphi_v \rangle_{L^2} + \langle \pi_v, \psi_v \rangle_{L^2}.
\]

(5.15)

On the other hand, the first order necessary conditions (3.3)–(3.3) for the nonlinear control problem yield that \( \mathcal{L}_v(\upsilon_v, \lambda_v)[\upsilon_v - \upsilon_v] \geq 0 \), or, using the definition (3.1) of \( \mathcal{L} \),

\[
0 \leq J_v(\upsilon_v)(\upsilon_v - \upsilon_v) \\
- \int_0^T \left< (u_\pi - u), \frac{\ell}{2} (\varphi_v - \varphi_v), \frac{\ell}{2} (\varphi_v - \varphi_v), p \right> + \kappa \left< \nabla (u_\pi - u), \nabla p \right> - \left< (f - f_\pi), p \right> dt \\
- \int_0^T \tau \left< (\varphi_v - \varphi_v), \psi_v \right> + \xi^2 \left< \nabla (\varphi_v - \varphi_v), \nabla \psi_v \right> - \left< g' \left( \varphi_v - \varphi_v \right), 2(\upsilon_v - \upsilon_v), \psi_v \right> dt \\
= J_v(\upsilon_v)(\upsilon_v - \upsilon_v) - \langle \pi_v, \psi_v \rangle_{L^2}.
\]

Thus, \( J_v(\upsilon_v)(\upsilon_v - \upsilon_v) \leq -\langle \pi_v, \psi_v \rangle_{L^2} \). Inserting this into (5.15) and using Lemma 5.1 we find that

\[
\mathcal{L}_{uv}(\upsilon_v, \lambda_v)[\upsilon_v - \upsilon_v, \upsilon_v - \upsilon_v] \\
\leq \langle \pi_v, \varphi_v - \varphi_v \rangle_{L^2} - \langle \pi_v, \psi_v \rangle_{L^2} \\
\leq ||\pi_v||_{L^2} ||\varphi_v - \varphi_v||_{L^2} + ||\pi_v||_{L^2} ||\psi_v - \psi_v||_{L^2} \\
\leq C(||\pi_v||_{L^2} + ||\pi_v||_{L^2}) \left( ||u_\pi - u\|_{W^{2,1}_2} + ||\varphi_v - \varphi_v||_{W^{2,1}_2} \right) \\
+ C(||\pi_v||_{L^2} + ||\pi_v||_{L^2})^2.
\]

(5.16)

Next we derive a lower bound for \( \mathcal{L}_{uv}(\upsilon_v, \lambda_v)[\upsilon_v - \upsilon_v, \upsilon_v - \upsilon_v] \). Let \( (u, \varphi) \) satisfy the linearized state equation (3.12) with \( f = f_\pi - f_v \). The second order sufficient condition (3.11) implies

\[
\mathcal{L}_{uv}(\upsilon_v, \lambda_v)[v, v] \geq \sigma_v (||u||_{L^2}^2 + ||\varphi||_{L^2}^2 + ||f - f_v||_{L^2}^2),
\]

(5.17)

where \( v = (u, \varphi, f_\pi - f_v) \). Moreover, the triple \( \hat{\upsilon} = v - (\upsilon_v - \upsilon_v) \) satisfies

\[
\frac{\partial}{\partial t} \hat{\upsilon} + \frac{\ell}{2} \frac{\partial}{\partial \varphi} \hat{\upsilon} = \kappa \Delta \hat{\upsilon} \\
\frac{\partial}{\partial \varphi} \hat{\upsilon} = \xi^2 \Delta \hat{\upsilon} + g'(\varphi) \hat{\upsilon} + 2\hat{\upsilon} - \pi_v, \quad \text{in } \Omega \times (0, T]
\]

with homogeneous initial and boundary conditions. Theorem 2.4 implies that

\[
||\hat{\upsilon}||_{W^{2,1}_2} \leq C ||\pi_v||_{L^2} \leq C ||\pi_v||_{L^2}^2.
\]
Using this estimate, the representation \( v = \hat{v} + (v_\pi - v_*) \), and (5.17) we find that
\[
\mathcal{L}_{vv}(v_* , \lambda_*) [v_\pi - v_* , v_\pi - v_*] \\
\geq \sigma_* (|\|\theta\|_{L^2}^2 + |\|\varphi\|_{L^2}^2 + |\|f_\pi - f_*\|_{L^2}^2) + 2\mathcal{L}_{vv}(v_* , \lambda_*) [\hat{v}, v] + \mathcal{L}_{vv}(v_* , \lambda_*) [\hat{v}, \hat{v}] \\
\geq \sigma_* (|\|f_\pi - f_*\|_{L^2}^2 - C|\|\hat{v}\|_{V^2} - C|\|\theta\|_{V}^2) \\
\geq \sigma_* (\|f_\pi - f_*\|_{L^2}^2 - C(\|\pi\|_{L^2}^2 \|v_\pi - v_*\|_{V} + \|\pi\|_{(L^2)^2}^2)).
\]  
(5.18)

Combining (5.16) and (5.18) yields
\[
\|f_\pi - f_*\|_{L^2}^2 \leq C(\|\pi\|_{(L^2)^2} \|v_\pi - v_*\|_{V} + \|\pi\|_{(L^2)^2}^2).
\]  
(5.19)

Since \( v = v_* - v_\pi \) satisfies (5.14), Theorem 2.4 implies the estimate
\[
\|u_\pi - u_*\|_{W^{2,1}} + \|\varphi_\pi - \varphi_*\|_{W^{2,1}} \leq C\|f_\pi - f_*\|_{L^2} + \|\pi\|_{(L^2)^2}.
\]  
(5.20)

Inserting this into (5.19) gives
\[
\|f_\pi - f_*\|_{L^2}^2 \leq C\|\pi\|_{(L^2)^2} \|f_\pi - f_*\|_{V} + \|\pi\|_{(L^2)^2}^2).
\]  
(5.21)

If \( \|f_\pi - f_*\|_{L^2} \leq 2C\|\pi\|_{(L^2)^2} \), then the assertion follows instantly from (5.20). If \( \|f_\pi - f_*\|_{L^2}^2 > 2C\|\pi\|_{(L^2)^2} \), then (5.21) yields
\[
\|f_\pi - f_*\|_{L^2} \leq C\|\pi\|_{(L^2)^2}.
\]  
(5.22)

The assertion now follows with (5.20).

**Lemma 5.3** Let the assumptions (A1)–(A3) be valid, let \( F_{ad} \) be bounded, and suppose that the second order sufficient optimality conditions are satisfied at \( v_* = (u_*, \varphi_*, f_*) \) ∈ \( V \).

If for \( q \in [2, \infty) \) the solution of the perturbed problem and the optimal solution satisfy \( v_\pi, v_* \in V_q \), then
\[
\|v_\pi - v_*\|_{V_q} \leq C\|\pi\|_{(L^2)^2}
\]  
for all \( \pi \in (L^q(Q))^2 \).

Moreover, if \( \gamma > 0 \) and if \( n = 3 \) and \( q > 5/2 \) or if \( n = 2 \) and \( q > 2 \), the previous estimate can be replaced by
\[
\|u_\pi - u_*\|_{W^{2,1}} + \|\varphi_\pi - \varphi_*\|_{W^{2,1}} + \|f_\pi - f_*\|_{L^\infty} \leq C\|\pi\|_{(L^2)^2}.
\]

**Proof:** As in the proof of Lemma 4.5 we can conclude that
\[
\|f_\pi - f_*\|_{L^\infty} \leq C\|\pi\|_{(L^2)^2}.
\]

Since \( v = v_* - v_\pi \) satisfies (5.14), we can appeal to Theorem 2.4 to show that
\[
\|u_\pi - u_*\|_{W^{2,1}} + \|\varphi_\pi - \varphi_*\|_{W^{2,1}} \leq C(\|f_\pi - f_*\|_{L^2} + \|\pi\|_{(L^2)^2}) \leq C\|\pi\|_{(L^2)^2}.
\]

This gives the first estimate. We can use this inequality and Lemma 5.1 to conclude that
\[
\|\lambda_\pi - \lambda_*\|_{L_q} \leq C\|\pi\|_{(L^2)^2}.
\]

The desired estimate for \( \|f_\pi - f_*\|_{L^\infty} \) can now be proven analogously to Lemma 4.5.

As we have mentioned at the beginning of this section we interpret the SQP subproblem (4.1) to (4.6) as a perturbed quadratic problem about the solution \( v_* \). The relation between the SQP subproblem (4.1) to (4.6) and the perturbed subproblem (5.1) to (5.5) is exploited in the following lemma:
Lemma 5.4 Let the assumptions (A1)–(A3) be valid. If \( v_* = (u_*, \varphi_*, f_*) \in V \) satisfies the second order sufficient optimality conditions, then there exists \( \epsilon > 0 \) such that if \( \| v_* - v_* \|_V + \| \lambda_* - \lambda_* \|_\Lambda \leq \epsilon \) the solutions of the SQP subproblem (4.1) to (4.6) are solutions of the perturbed subproblem (5.1) to (5.5) with
\[
\pi_s = g(\varphi_e) - g(\varphi_*) + g'(\varphi_*)(\varphi_+ - \varphi_e) - g'(\varphi_*)(\varphi_+ - \varphi_*) ,
\]
\[
\pi_a = g'(\varphi_e)\psi_+ + \psi_* g''(\varphi_e)(\varphi_+ - \varphi_e) - g'(\varphi_*)\psi_+ - \psi_* g''(\varphi_*)(\varphi_+ - \varphi_*) .
\]

Proof: From Lemma 4.4 we know that there exists \( \epsilon > 0 \) such that if \( \| v_* - v_* \|_V + \| \lambda_* - \lambda_* \|_\Lambda \leq \epsilon \) there exists a unique solution \((u_+, \varphi_+, f_+)\) of the SQP subproblem (4.1) to (4.6) which can be characterized by the linearized state equations (4.2)–(4.5), the adjoint equations (4.10), and the conditions (4.9), or (4.12). From a comparison of (4.2)–(4.5) and (5.2)–(5.4) one can see that \((u_*, \varphi_*, f_*)\) satisfies the perturbed system (5.2)–(5.4) with \( \pi_s \) given in (5.23). Moreover, one can see that \((u_+, \varphi_+, p_+, \psi_+ \) satisfies the perturbed adjoint system (5.8) with \( \pi_a \) given in (5.24). This implies the assertion.

Lemma 5.5 If the assumptions of Lemma 5.4 are valid, and if \( \varphi_*, \varphi_+, \psi_*, \psi_+ \in W^{2,1}_0(Q) \), then
\[
\| \pi_s \|_{L^0} \leq C (\| \varphi_* - \varphi_+ \|_{W^{2,1}_0}^2 + \| \varphi_* - \varphi_+ \|_{W^{2,1}_0} \| \varphi_+ - \varphi_* \|_{W^{2,1}_0} ) ,
\]
\[
\| \pi_a \|_{L^0} \leq C (\| \varphi_* - \varphi_+ \|_{W^{2,1}_0}^2 + \| \psi_* \|_{W^{2,1}_0} + \| \varphi_* - \varphi_+ \|_{W^{2,1}_0}) (\| \varphi_+ - \varphi_* \|_{W^{2,1}_0} + \| \varphi_* - \varphi_+ \|_{W^{2,1}_0} ) .
\]
The constant \( C \) depends on \( \varphi_*, \varphi_+, \psi_*, \psi_+ \), but is uniformly bounded if \( \varphi_*, \varphi_+, \psi_*, \psi_+ \in W^{2,1}_0(Q) \) are contained in a bounded set in \( W^{2,1}_0(Q) \).

Proof: From (5.23) we find that
\[
\pi_s = g(\varphi_e) - g(\varphi_*) + g'(\varphi_*)(\varphi_+ - \varphi_e) - g'(\varphi_*)(\varphi_+ - \varphi_*) ,
\]
\[
\pi_a = g'(\varphi_e)\psi_+ + \psi_* g''(\varphi_e)(\varphi_+ - \varphi_e) - g'(\varphi_*)\psi_+ - \psi_* g''(\varphi_*)(\varphi_+ - \varphi_*) .
\]
Using the definition of \( g \), Hölder’s inequality, and the imbedding \( W^{2,1}_0(Q) \subset L^{3q}(Q) \) which holds true since \( 5q/(5-2q) \geq 3q \) for \( q \in [2,5/2) \) or for \( q \geq 5/2 \), this gives the first estimate.

Similarly, (5.24) can be written as
\[
\pi_a = g'(\varphi_e)\psi_+ + \psi_* g''(\varphi_*)(\varphi_+ - \varphi_e) - g'(\varphi_*)\psi_+ + \psi_* g''(\varphi_*)(\varphi_+ - \varphi_*)
\]
\[
= [g'(\varphi_e) - g'(\varphi_*)](\varphi_+ - \varphi_e)]\psi_+ + g''(\varphi_*)(\psi_+ - \psi_*)(\varphi_+ - \varphi_*)
\]
\[
+ [\psi_* g''(\varphi_e) - \psi_* g''(\varphi_*)](\varphi_+ - \varphi_*) .
\]
Applying estimates analogous to the ones above gives the assertion.

6 Local Convergence of the SQP Method

As we have described in Section 4, the Lagrange–SQP–Newton method solves the nonlinear, non–convex optimal control problem (1.5), (1.6), (1.1), (1.2), (1.3), (1.4) through a sequence of linear–quadratic control problems. Given current approximations for control, states and Lagrange multipliers \( (v_*, \lambda_*) = (u_*, \varphi_*, f_*, p_*, \psi_*) \) the new approximations for control, states and Lagrange multipliers \( (v_*; \lambda_*) = (u_+, \varphi_+, f_+, p_+, \psi_+) \) are computed as the solution of (4.1)–(4.6). In the previous section we have shown that this subproblem can be viewed as a perturbation of the linear quadratic optimal control problem (4.1)–(4.6) at \( v_* = v_* \). This observation and the Lipschitz continuous dependence of the solution of the perturbed problem upon the perturbation can be used to establish the quadratic convergence of the Lagrange–SQP–Newton method.
Theorem 6.1 Suppose that the assumptions (A1)–(A3) are satisfied, $\gamma > 0$, and that for $q \in [2, \infty)$ the desired temperature and phase profiles satisfy $u_d, \varphi_d \in L^q(Q)$. Moreover let the current iterate satisfy $(v, \lambda) \in V_q \times \Lambda_q$.

(i) If $q = 2$ and $\gamma > 0$, then there exists $\epsilon > 0$ such that $\|v - v_*\|_V + \|\lambda - \lambda_*\|_A \leq \epsilon$ implies

$$\|v - v_*\|_V + \|\lambda - \lambda_*\|_A \leq C \left( \|v - v_*\|_V + \|\lambda - \lambda_*\|_A \right)^2.$$ 

(ii) If $q > 2$ and $F_d$ is bounded, then there exists $\epsilon > 0$ such that $\|v - v_*\|_V + \|\lambda - \lambda_*\|_A \leq \epsilon$ implies

$$\|v - v_*\|_V + \|\lambda - \lambda_*\|_A \leq C \left( \|v - v_*\|_V + \|\lambda - \lambda_*\|_A \right)^2.$$ 

(iii) If $n = 3$ and $q > 5/2$ or if $n = 2$ and $q > 2$ and if $F_d$ is bounded, then there exists $\epsilon > 0$ such that $\|v - v_*\|_V + \|\lambda - \lambda_*\|_A \leq \epsilon$ implies

$$\|v - v_*\|_V + \|\lambda - \lambda_*\|_A \leq C \left( \|v - v_*\|_V + \|\lambda - \lambda_*\|_A \right)^2.$$ 

Proof: By Lemma 4.4 there exists $\epsilon_1 > 0$ such that for all $v \in V, \lambda \in \Lambda$ with $\|v - v_*\|_V + \|\lambda - \lambda_*\|_A \leq \epsilon_1$ there exists a unique solution $v_* \in V$ of (4.1)–(4.6) with corresponding Lagrange multiplier $\lambda_* \in \Lambda$. Moreover, the assumptions $u_d, \varphi_d \in L^q(Q), (v, \lambda) \in V_q \times \Lambda_q$ imply that $(v_* + \lambda, (v, \lambda) \in V_q \times \Lambda_q$. See Theorems 2.1, 3.1, Lemma 4.7, and equation (4.11).

(i) Lemma 4.4 guarantees the existence of $\epsilon_2 \in (0, \epsilon_1)$ such that $\|v - v_*\|_V + \|\lambda - \lambda_*\|_A \leq \epsilon_2$. The constant $C$ depends only on $\epsilon_2$.

If we define the perturbation $\pi = (\pi_s, \pi_q)$ as in (5.23), (5.24), then Lemma 5.5 and the boundedness of the iterates and Lagrange multipliers imply that

$$\|\pi_s\|_V + \|\pi_q\|_V \leq C \left( \|v - v_*\|_V + \|\lambda - \lambda_*\|_A \right)^2 + \|v - v_*\|_V + \|\lambda - \lambda_*\|_A.$$ 

Since $\|v - v_*\|_V + \|\lambda - \lambda_*\|_A \leq 2 \epsilon_2$ this gives the estimate

$$\|\pi_s\|_V + \|\pi_q\|_V \leq C \left( \|v - v_*\|_V + \|\lambda - \lambda_*\|_A \right). \quad (6.1)$$

Lemma 5.4 shows that $v_*$ is the solution of the perturbed problem with perturbation $\pi = (\pi_s, \pi_q)$ given by (5.23), (5.24) and corresponding Lagrange multiplier $\lambda_*$. Hence, the estimates in Lemmas 5.1, 5.2 and (6.1) show that

$$\|v - v_*\|_V \leq C \left( \|v - v_*\|_V + \|\lambda - \lambda_*\|_A \right)$$

and

$$\|\lambda - \lambda_*\|_A \leq C \left( \|v - v_*\|_V + \|\lambda - \lambda_*\|_A \right).$$

Inserting these estimates into (5.25), (5.26) gives

$$\|\pi_s\|_V + \|\pi_q\|_V \leq C \left( \|v - v_*\|_V + \|\lambda - \lambda_*\|_A \right)^2. \quad (6.2)$$

Using the estimates in Lemmas 5.1, 5.2 and (6.1) again we derive at the desired inequalities

$$\|v - v_*\|_V \leq C \left( \|v - v_*\|_V + \|\lambda - \lambda_*\|_A \right)^2. \quad (6.3)$$

and

$$\|\lambda - \lambda_*\|_A \leq C \left( \|v - v_*\|_V + \|\lambda - \lambda_*\|_A \right)^2. \quad (6.4)$$

(ii), (iii) These assertions can be proven analogously. We have to replace Lemma 4.4 by Lemma 4.5 and Lemma 5.2 by Lemma 5.3. \qed
7 Appendix

For convenience we restate the auxiliary system and the Theorem 2.4. The system is given by

\[
\begin{align*}
\frac{\partial u}{\partial t} + \frac{\ell}{2} \varphi &= \kappa \Delta u + \beta_1 u + f_1, \\
\tau \frac{\partial \varphi}{\partial t} &= \xi^2 \Delta \varphi + \beta_2 \varphi + 2u + f_2
\end{align*}
\]

in \(\Omega \times (0, T]\), \hspace{1cm} (7.1)

with boundary conditions

\[
\begin{align*}
\frac{\partial u}{\partial n} &= 0, \hspace{1cm} \frac{\partial \varphi}{\partial n} = 0, \hspace{1cm} \text{on} \partial \Omega \times (0, T),
\end{align*}
\]

and initial conditions

\[
u = u_0, \hspace{1cm} \varphi = \varphi_0 \hspace{1cm} \text{in} \Omega.\]

(7.2)

\textbf{Theorem 7.1} Let the assumptions (A1)–(A3) be valid and suppose that \(f_1, f_2 \in L^q(\Omega), q \geq 2\). If \(\beta_1, \beta_2 \in L^3(\Omega),\) then there exists a unique solution \((u, \varphi) \in W^{2,1}_0(\Omega) \times W^{2,1}_0(\Omega)\) of the system (7.1), (7.2), (7.3). The solution obeys

\[
\|(u)\|_{W^{2,1}_0(\Omega)} + \|(\varphi)\|_{W^{2,1}_0(\Omega)} \leq C \left(\|u_0\|_{L^\infty(\Omega)} + \|\varphi_0\|_{w^2_\infty(\Omega)} + \|f_1\|_{L^q(\Omega)} + \|f_2\|_{L^q(\Omega)}\right),
\]

with a constant \(C = C(\beta_1, \beta_2)\) depending only on \(\beta_1\|_{L^3(\Omega)}, \|\beta_2\|_{L^3(\Omega)}\). Moreover, the function \((\beta_1, \beta_2) \mapsto C(\beta_1, \beta_2)\) maps bounded sets into bounded sets.

\textbf{Proof:} Again, \(C\) denotes a generic constant independent of \(u\) and \(\varphi\).

(i) An a–priori estimate:

Suppose that \(u, \varphi \in \{v \mid v \in L^2(0, T; H^1(\Omega)), \frac{\partial v}{\partial t} \in L^2(0, T; (H^1(\Omega)^\prime))\}\) solve (7.1), (7.2), (7.3). Multiplying the first equation in (7.1) by \(u(t) + \frac{\ell}{2} \varphi(t)\) and the second one by \(\varphi(t)\), using integration by parts, and applying the Sobolev imbedding \(H^1(\Omega) \subset L^6(\Omega)\) yields

\[
\begin{align*}
\frac{1}{2} \frac{\partial}{\partial t} \int_\Omega (u(t) + \frac{\ell}{2} \varphi(t))^2 \, dx + \int_\Omega \kappa |\nabla u(t)|^2 \, dx + \frac{\ell \kappa}{2} \int_\Omega \nabla u(t) \nabla \varphi(t) \, dx \\
&= \int_\Omega \beta_1(t) u(t)(u(t) + \frac{\ell}{2} \varphi(t)) \, dx + \int_\Omega (u(t) + \frac{\ell}{2} \varphi(t)) f_1(t) \, dx \\
&\leq \|\beta_1(t)\|_{L^3(\Omega)} \|u(t)\|_{L^\infty(\Omega)} \|u(t) + \frac{\ell}{2} \varphi(t)\|_{L^\infty(\Omega)} + \frac{1}{2} \|u(t)\|^2_{L^\infty(\Omega)} + \frac{1}{2} \|f_1(t)\|^2_{L^2(\Omega)} \\
&\leq C \|\beta_1(t)\|_{L^3(\Omega)} \|u(t)\|_{L^\infty(\Omega)} + \frac{1}{2} \|u(t)\|^2_{L^\infty(\Omega)} + \frac{1}{2} \|f_1(t)\|^2_{L^2(\Omega)} \\
&\leq C \|\beta_1(t)\|^2_{L^3(\Omega)} \|u(t)\|^2_{L^\infty(\Omega)} + \frac{\kappa}{2} \|\nabla (u(t) + \frac{\ell}{2} \varphi(t))^2\|_{L^2(\Omega)} \|u(t)\|_{L^\infty(\Omega)} + \frac{1}{2} \|u(t)\|^2_{L^\infty(\Omega)} + \frac{1}{2} \|f_1(t)\|^2_{L^2(\Omega)} \\
&\leq \frac{\kappa}{2} \|\nabla (u(t) + \frac{\ell}{2} \varphi(t))^2\|_{L^2(\Omega)} + C \|\beta_1(t)\|^2_{L^3(\Omega)} \|u(t)\|^2_{L^\infty(\Omega)} + \|u(t)\|^2_{L^\infty(\Omega)} + \|f_1(t)\|^2_{L^2(\Omega)} \hspace{1cm} (7.4)
\end{align*}
\]

and

\[
\begin{align*}
\frac{\tau}{2} \frac{\partial}{\partial t} \int_\Omega \varphi(t)^2 \, dx + \int_\Omega \xi^2 |\nabla \varphi(t)|^2 \, dx \\
&= \int_\Omega \beta_2(t) \varphi(t)^2 \, dx + 2 \int_\Omega u(t) \varphi(t) \, dx + \int_\Omega f_2(t) \varphi(t) \, dx \\
&\leq \|\beta_2(t)\|_{L^3(\Omega)} \|\varphi(t)\|_{L^\infty(\Omega)} \|\varphi(t)\|_{L^\infty(\Omega)} + 2 \int_\Omega (u(t) + \frac{\ell}{2} \varphi(t))^2 \, dx + C \|\varphi(t)\|^2_{L^2(\Omega)} + C \|f_2(t)\|^2_{L^2(\Omega)} \\
&\leq \frac{\xi^2}{2} \|\varphi(t)\|^2_{L^2(\Omega)} + C \|\beta_2(t)\|^2_{L^3(\Omega)} \|u(t) + \frac{\ell}{2} \varphi(t)\|^2_{L^2(\Omega)} + \|\varphi(t)\|^2_{L^2(\Omega)} + C \|f_2(t)\|^2_{L^2(\Omega)}. \hspace{1cm} (7.5)
\end{align*}
\]
Multiplying (7.5) by $A$ and adding the result to (7.4) yields
\[
\frac{1}{2} \frac{\partial}{\partial t} \int_{\Omega} (u(t) + \frac{\ell}{2} \varphi(t))^2 + \tau A \varphi(t)^2 \, dx + \frac{1}{2} \int_{\Omega} k |\nabla u(t)|^2 + (A \frac{\ell^2}{2} - \frac{\kappa \ell^2}{4}) |\nabla \varphi(t)|^2 \, dx
\leq C \left( 1 + \| \beta_1(t) \|_{L^2(\Omega)}^2 + \| \beta_2(t) \|_{L^2(\Omega)}^2 \right) \left( \| u(t) + \frac{\ell}{2} \varphi(t) \|_{L^2(\Omega)}^2 + \| \varphi(t) \|_{L^2(\Omega)}^2 \right) \\
+ C \| f_1(t) \|_{L^2(\Omega)}^2 + C \| f_2(t) \|_{L^2(\Omega)}^2.
\] (7.6)

If we choose $A \geq 2(\kappa \ell^2/4 + 1)/\xi^2$, then (7.6) implies
\[
\frac{1}{2} \frac{\partial}{\partial t} \int_{\Omega} (u(t) + \frac{\ell}{2} \varphi(t))^2 + \tau A \varphi(t)^2 \, dx + \frac{1}{2} \int_{\Omega} k |\nabla u(t)|^2 + |\nabla \varphi(t)|^2 \, dx
\leq C \left( 1 + \| \beta_1(t) \|_{L^2(\Omega)}^2 + \| \beta_2(t) \|_{L^2(\Omega)}^2 \right) \left( \| u(t) + \frac{\ell}{2} \varphi(t) \|_{L^2(\Omega)}^2 + \| \varphi(t) \|_{L^2(\Omega)}^2 \right) \\
+ C \| f_1(t) \|_{L^2(\Omega)}^2 + C \| f_2(t) \|_{L^2(\Omega)}^2.
\] (7.7)

Integration over $t$ and by parts gives
\[
\frac{1}{2} \| u(t) + \frac{\ell}{2} \varphi(t) \|_{L^2(\Omega)}^2 + \frac{\tau A}{2} \| \varphi(t) \|_{L^2(\Omega)}^2 + \frac{1}{2} \int_0^t \kappa |\nabla u(s)|_{L^2(\Omega)}^2 + |\nabla \varphi(s)|_{L^2(\Omega)}^2 \, ds
\leq \frac{1}{2} \left( \| u(0) + \frac{\ell}{2} \varphi(0) \|_{L^2(\Omega)}^2 + \tau A \| \varphi(0) \|_{L^2(\Omega)}^2 \right) + C \| f_1 \|_{L^2(\Omega)}^2 + C \| f_2 \|_{L^2(\Omega)}^2 \\
+ C \int_0^t \left( 1 + \| \beta_1(s) \|_{L^2(\Omega)}^2 + \| \beta_2(s) \|_{L^2(\Omega)}^2 \right) \left( \| u(s) \|_{L^2(\Omega)}^2 + \| \varphi(s) \|_{L^2(\Omega)}^2 \right) \, ds.
\] (7.8)

If $A \geq (1/2 + \ell^2/4)/\tau$, then
\[(a + \frac{\ell}{2} b)^2 + \tau A b^2 = a^2 + \ell ab + \frac{\ell^2}{4} b^2 + \tau A b^2 \geq a^2 - \frac{1}{2} a^2 - \frac{\ell^2}{4} b^2 + \frac{\ell^2}{4} b^2 + \tau A b^2 \geq \frac{1}{2} a^2 + \frac{1}{2} b^2\]

for all $a, b$. Thus, for $A \geq (1/2 + \ell^2/4)/\tau$ equation (7.8) gives the estimate
\[
\| u(t) \|_{L^2(\Omega)}^2 + \| \varphi(t) \|_{L^2(\Omega)}^2 \leq C \left( \| u(0) \|_{L^2(\Omega)}^2 + \| \varphi(0) \|_{L^2(\Omega)}^2 + \| f_1 \|_{L^2(\Omega)}^2 + \| f_2 \|_{L^2(\Omega)}^2 \right) \\
+ C \int_0^t \left( 1 + \| \beta_1(s) \|_{L^2(\Omega)}^2 + \| \beta_2(s) \|_{L^2(\Omega)}^2 \right) \left( \| u(s) \|_{L^2(\Omega)}^2 + \| \varphi(s) \|_{L^2(\Omega)}^2 \right) \, ds.
\] (7.9)

Now we can use the Gronwall–Bellman inequality\(^1\) to derive the inequality
\[
\| u(t) \|_{L^2(\Omega)}^2 + \| \varphi(t) \|_{L^2(\Omega)}^2 \leq C \left( \| u(0) \|_{L^2(\Omega)}^2 + \| \varphi(0) \|_{L^2(\Omega)}^2 + \| f_1 \|_{L^2(\Omega)}^2 + \| f_2 \|_{L^2(\Omega)}^2 \right).
\]

---
\(^1\)We use the version of the Gronwall–Bellman inequality given in [6, L. A.4]: Let $m \in L^1(0,T)$, $m \geq 0$, a.e. $[0,T]$, and $a$ be a positive constant. If $\psi \in C[0,T]$ satisfies the inequality
\[
\psi(t) \leq a + \int_0^t m(s) \psi(s) \, ds, \quad t \in [0,T],
\]
then the following estimate is true:
\[
\psi(t) \leq a \exp \left( \int_0^t m(s) \, ds \right), \quad t \in [0,T].
\]
Inserting this equation into (7.9) yields the a–priori estimate
\[
\|u(t)\|^2_{L^2(\Omega)} + \|\varphi(t)\|^2_{L^2(\Omega)} + \int_0^t \kappa \|\nabla u(s)\|^2_{L^2(\Omega)} + \|\nabla \varphi(s)\|^2_{L^2(\Omega)} \, ds \\
\leq C\left(\|u(0)\|^2_{L^2(\Omega)} + \|\varphi(0)\|^2_{L^2(\Omega)} + \|f_1\|^2_{L^2(\Omega)} + \|f_2\|^2_{L^2(\Omega)}\right).
\]
(7.10)

(ii) Uniqueness of the solution: If \((u_i, \varphi_i), i = 1, 2\) are solutions of (7.1), then \((u, \varphi) = (u_1, \varphi_1) - (u_2, \varphi_2)\) solves (7.1) with \(f_1 = f_2 = 0\) and homogeneous initial and boundary conditions. The a–priori estimate (7.10) implies that \(u = \varphi = 0\).

(iii) Existence of the solution: From the a–priori estimate (7.10) we can deduce the existence of a solution
\[
(u, \varphi) \in \left\{ v \mid v \in L^2(0, T; H^1(\Omega)), \frac{\partial}{\partial t} v \in L^2(0, T; (H^1(\Omega))^*) \right\}^2 \subset \left( C(0, T; L^2(\Omega)) \right)^2
\]
using the Galerkin method. This proof uses standard techniques, see e.g., [14, pp. 509ff], [46, § 23.9]. We omit the details.

(iv) Regularity of the solution: To establish the regularity result, we first consider the second equation in system (7.1) with given \(u \in L^2(0, T; H^1(\Omega))\) on the right hand side. From the \(L^p\) theory of linear parabolic equations in [33, p. 341] we find that the solution \(\varphi\) is in \(W^{2,1}_2(Q)\). Using this regularity estimate, i.e. that \(\frac{\partial}{\partial t} \varphi \in L^2(Q)\), in the right hand side of the first equation in (7.1), we can show by the same arguments that \(u \in W^{2,1}_2(Q)\).

Now we can apply the same arguments to show the desired regularity result. In the second equation in system (7.1) we view \(\beta_2 \varphi + 2u + f_2\) as given. Since \(\varphi \in L^{10}(Q)\), cf. Theorem 2.2, it holds that
\[
\|\beta_2 \varphi\|_{L^{10/3}(Q)} \leq \|\beta_2\|_{L^{3}(Q)} \|\varphi\|_{L^{10}(Q)}.
\]
From this we deduce that \(\varphi \in W^{2,1}_{30/13}(Q) \subset L^{30}(Q)\). Inserting this into the first equation gives \(u \in W^{2,1}_{30/13}(Q) \subset L^{30}(Q)\). Hence,
\[
\|\beta_2 \varphi\|_{L^{2/1}(Q)} \leq \|\beta_2\|_{L^{3}(Q)} \|\varphi\|_{L^{10}(Q)}
\]
and \(u, \varphi \in W^{2,1}_p(Q)\) with \(p = \min(q, 5/2)\). If \(5/2 < q\) we can use the imbedding \(W^{2,1}_{5/2} \subset L^{p}(Q)\) for all \(\mu \in [1, \infty)\) and repeat the previous steps to show that \(u, \varphi \in W^{2,1}_q(Q)\).

Acknowledgements

This work was initiated while both authors were at the Interdisciplinary Center for Applied Mathematics, the first author as a faculty member, the second author as a visiting scientist. This paper is based on an unpublished technical report (Interdisciplinary Center for Applied Mathematics (ICAM) REPORT 95–03–01, Virginia Polytechnic Institute and State University, Blacksburg, VA, 1995). The authors would like to thank their colleagues at ICAM for their generous support and hospitality.

References


