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

Matthias Heinkenschloss *　　**Fredi Tröltzsch**†

October 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 fi eld 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 fi eld equation. Phase fi eld equations are used to model solidifi cation. They are systems of partial differential equations (PDEs). The unknowns in the system of PDEs are the order parameter φ (also called phase function) and the temperature u. Unlike in the classical Stefan problem which models a sharp solid-liquid interface, phase fi eld models allow for a mushy region. The phases are identifi ed using the order parameter φ. 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 fi eld model introduced in [7], [13], [18] that consists of the two differential equations

$$\frac{\partial}{\partial t}u + \frac{\ell}{2}\frac{\partial}{\partial t}\varphi = \kappa \Delta u + f \tag{1.1}$$

$$\text{in } \Omega \times (0, T],$$

$$\tau \frac{\partial}{\partial t}\varphi = \xi^2 \Delta \varphi + g(\varphi) + 2u \tag{1.2}$$

with boundary conditions

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

and initial conditions

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

*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, Facultät für Mathematik, D–09009 Chemnitz, PSF 964, 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.

The function $-g(z)$ is the derivative of a so–called double well potential $G(z)$. Often $G(z) = \frac{1}{8}(z^2-1)^2$. We assume that $g(z) = az + bz^2 - cz^3$ with bounded coefficient functions a, b, c and strictly positive c. In the model κ denotes the heat conductivity, ℓ the latent heat, τ is the relaxation time, and ξ 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 φ match desired temperature and phase profiles u_d and φ_d, respectively. Mathematically, the problem is formulated as

$$\min_{f \in \mathcal{F}_{ad}} J(f), \tag{1.5}$$

where the objective function is given by

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

The set of admissible controls has the form

$$\mathcal{F}_{ad} = \{ f \in L^2(\Omega \times (0,T)) \mid f(x,t) \in F \text{ a.e. } \}, \tag{1.7}$$

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 φ 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,φ) is the state. In the objective function α and β 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 α, β, and γ 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 be 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

$$\begin{aligned} &\|u_+ - u_*\|_{W_q^{2,1}} + \|\varphi_+ - \varphi_*\|_{W_q^{2,1}} + \|f_+ - f_*\|_{L^\infty} + \|\lambda_+ - \lambda_*\|_{\Lambda_q} \\ \leq\ & C\left(\|u_c - u_*\|_{W_q^{2,1}} + \|\varphi_c - \varphi_*\|_{W_q^{2,1}} + \|f_c - f_*\|_{L^\infty} + \|\lambda_c - \lambda_*\|_{\Lambda_q}\right)^2 \end{aligned} \tag{1.8}$$

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 $\Lambda_q = W_q^{2,1}(Q) \times W_q^{2,1}(Q)$. 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^∞ 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^∞–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 g 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 $L^6(Q)$ 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 $\mathcal{F}_{ad}$ is unbounded. However, if $\mathcal{F}_{ad}$ 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 $L^6(Q)$ 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] perfomed 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 definitness 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 convergence proof for the SQP method is based on the observation that the iterates can be interpreted as solutions of a perturbed quadratic problem obtained from a linearization of the original problem around the solution. This relation and some important estimates for the solutions of the perturbed problem are discussed in Section 5. The results of Sections 4 and 5 are used in Section 6 to derive the desired convergence estimates.

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 $\Omega \subset I\!R^n$, $n = 2, 3$, 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 $g(z) = az + bz^2 - cz^3$ with

$$a(x,t) \le \overline{a}, \quad b(x,t) \le \overline{b}, \quad 0 < \underline{c} \le c(x,t) \le \overline{c}.$$

(The results can be generalized to the case where κ, ξ, ℓ, and τ 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

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

The space $W_p^{2,1}(Q)$ equipped with the norm

$$||u||_{W_p^{2,1}(Q)} = \left(\int_Q |u|^p + \sum_{i=1}^n |\frac{\partial u}{\partial x_i}|^p + \sum_{i,j=1}^n |\frac{\partial^2 u}{\partial x_i \partial x_j}|^p + |\frac{\partial u}{\partial t}|^p \, dxdt \right)^{1/p}$$

is a Banach space. We often omit the space Q and use $W_p^{2,1}, L^p$ instead of $W_p^{2,1}(Q), L^p(Q)$, 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, φ_c, f_c), (u_+, φ_+, f_+), (u_*, φ_*, f_*) denote the current iterate, new iterate, and minimum point, respectively. Moreover, we use the notations

$$v = (u, \varphi, f), \quad \lambda = (p, \psi), \tag{1.9}$$

for the triple of states and control and the pair of co–states, respectively. Similar notations are used for $v_*, \widehat{v}$, etc. Finally, we introduce the product spaces

$$V_q = W_q^{2,1}(Q) \times W_q^{2,1}(Q) \times L^q(Q), \quad \Lambda_q = W_q^{2,1}(Q) \times W_q^{2,1}(Q). \tag{1.10}$$

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

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

2 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].

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*

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

Here the parameter p *is given by*

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

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:

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

$$W_2^{2,1}(Q) \subset L^\infty(0, T; L^p(\Omega)) \cap L^q(0, T; L^\infty(\Omega)), \quad p, q \in [1, \infty)$$

are continuous; the imbeddings $W_q^{2,1}(Q) \subset L^p(Q)$, $p \in [1, \infty)$, $q \in [2, \infty)$ *are compact.*

If $\Omega \subset \mathbb{R}^3$ *is a bounded domain having the cone property, then the imbeddings*

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

where

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

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.*

Proof: The assertions are proven in [35, pp. 14,15,24,25]. □

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

$$\begin{array}{llll} S: & L^q(Q) & \longrightarrow & W_q^{2,1}(Q) \times W_p^{2,1}(Q), \\ & f & \mapsto & (u, \varphi), \end{array} \tag{2.1}$$

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

$$||u||_{W_2^{2,1}} + ||\varphi||_{W_2^{2,1}} \leq c\left(||u_0||_{W_\infty^2} + ||\varphi_0||_{W_\infty^2} + ||f||_{L^2}\right), \tag{2.2}$$

where $c = c(\ell, \xi, T)$, and

$$||u_1 - u_2||_{W_2^{2,1}} + ||\varphi_1 - \varphi_2||_{W_2^{2,1}} \leq C||f_1 - f_2||_{L^2} \tag{2.3}$$

for $(u_i, \varphi_i) = S(f_i)$, $i = 1, 2$, where C depends on $||u_i||_{W_2^{2,1}}, ||\varphi_i||_{W_2^{2,1}}$, $i = 1, 2$. If u_i, φ_i, $i = 1, 2$, are contained in bounded subsets $\mathcal{B} \subset (W_2^{2,1})^2$, which will be the case in our analysis, then C can be assumed to depend only on $\mathcal{B}$, but to be independent of individual $u_i, \varphi_i \in \mathcal{B}$, $i = 1, 2$. 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 $\frac{\gamma}{2}\int_0^T \|f(t)\|^2\,dt$, $\gamma > 0$, 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_*, φ_*) that satisfy the second order sufficient conditions.

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

We introduce the auxiliary system

$$\begin{aligned} \frac{\partial}{\partial t}u + \frac{\ell}{2}\frac{\partial}{\partial t}\varphi &= \kappa\Delta u + \beta_1 u + f_1 \\ \tau\frac{\partial}{\partial t}\varphi &= \xi^2\Delta\varphi + \beta_2\varphi + 2u + f_2 \end{aligned} \qquad \text{in } \Omega\times(0,T], \tag{2.4}$$

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 $\beta_1 = 0$, $\beta_2(x,t) = g'(\varphi_c(x,t))$, then the solution (u,φ) of (2.4), (1.3), (1.4) is the first Fréchet derivative $(u,\varphi) = (u_f(f_c)h, \varphi_f(f_c)h)$ of $(u(f),\varphi(f))$ at the control f_c. Moreover, with $f_1 = f, f_2 = g(\varphi_c)$, $\beta_1 = 0$, and $\beta_2 = g'(\varphi_c)$ the system (2.4), (1.3), (1.4) essentially defines the pair (u_+,φ_+) 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 $(u,\varphi) \in W_q^{2,1}(Q)\times W_q^{2,1}(Q)$ of the system* (2.4), (1.3), (1.4). *The solution obeys*

$$\|u\|_{W_q^{2,1}} + \|\varphi\|_{W_q^{2,1}} \leq C(\|u_0\|_{W^2_\infty(\Omega)} + \|\varphi_0\|_{W^2_\infty(\Omega)} + \|f_1\|_{L^q} + \|f_2\|_{L^q})$$

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. □

The conditions on β_i, $i = 1,2$, are clearly satisfied if $\beta_i \in L^\infty$. However, they are also satisfied if β_i is of the form

$$\beta_i(x,t) = a(x,t) + b(x,t)\widehat{\varphi}(x,t) - c(x,t)\widehat{\varphi}(x,t)^2$$

or, more generally, of the form

$$\beta_i(x,t) = \sum_{j=1}^m c_j(x,t)\widehat{\varphi}(x,t)^j,$$

with $c_j \in L^\infty(Q)$ and $\widehat{\varphi} \in L^{3m}$.

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

$$\begin{aligned}\mathcal{L}(u,\varphi,f,p,\psi) &= J(u,\varphi,f) - \int_0^T\int_\Omega p\,[u_t + \frac{\ell}{2}\varphi_t - \kappa\Delta u - f]\,dxdt \\ &\quad - \int_0^T\int_\Omega \psi\,[\tau\varphi_t - \xi^2\Delta\varphi - g(\varphi) - 2u]\,dxdt\end{aligned}$$

which, using integration by parts, can be written as

$$\begin{aligned}\mathcal{L}(u,\varphi,f,p,\psi) &= J(u,\varphi,f) - \int_0^T \langle u_t + \frac{\ell}{2}\varphi_t, p\rangle + \kappa\langle\nabla u,\nabla p\rangle - \langle f,p\rangle\,dt \\ &\quad - \int_0^T \tau\langle\varphi_t,\psi\rangle + \xi^2\langle\nabla\varphi,\nabla\psi\rangle - \langle g(\varphi)+2u,\psi\rangle\,dt\,. \qquad (3.1)\end{aligned}$$

For fixed p and ψ the Lagrangian splits into a linear part $\mathcal{L}_1$ and a nonlinear part $\mathcal{L}_2$, $\mathcal{L} = \mathcal{L}_1 + \mathcal{L}_2$, where

$$\mathcal{L}_2(u,\varphi,f,p,\psi) = J(u,\varphi,f) + \int_0^T \langle g(\varphi),\psi\rangle\,dt\,. \qquad (3.2)$$

The first order necessary optimality conditions for the optimal control problem are:

$$\begin{aligned}\mathcal{L}_{(u,\varphi)}(u_*,\varphi_*,f_*,p_*,\psi_*) &= 0, && (3.3)\\ \mathcal{L}_{(p,\psi)}(u_*,\varphi_*,f_*,p_*,\psi_*) &= 0, && (3.4)\\ \mathcal{L}_f(u_*,\varphi_*,f_*,p_*,\psi_*)(f-f_*) &\geq 0 \qquad \forall\, f\in\mathcal{F}_{ad}, && (3.5)\\ f_* &\in \mathcal{F}_{ad}. && (3.6)\end{aligned}$$

The equation (3.4) means that the state equation has to be satisfied. An evaluation of (3.3) shows that the pair (p_*,ψ_*) has to satisfy the adjoint system

$$\begin{aligned}-\tfrac{\partial}{\partial t}p &= \kappa\Delta p + 2\psi + \alpha(u_* - u_d), \\ -\tau\tfrac{\partial}{\partial t}\psi - \tfrac{\ell}{2}\tfrac{\partial}{\partial t}p &= \xi^2\Delta\psi + g'(\varphi_*)\psi + \beta(\varphi_* - \varphi_d),\end{aligned} \qquad \text{in } \Omega\times(0,T], \qquad (3.7)$$

with boundary conditions

$$\frac{\partial}{\partial n}p = 0, \quad \frac{\partial}{\partial n}\psi = 0, \qquad \text{on } \partial\Omega\times(0,T), \qquad (3.8)$$

and final conditions

$$p(x,T) = 0, \quad \psi(x,T) = 0 \qquad \text{in } \Omega. \qquad (3.9)$$

The inequality (3.5) is equivalent to the variational inequality

$$\int_0^T\int_\Omega (p_* + \gamma f_*)(f - f_*)\,dxdt \geq 0 \qquad \forall\, f\in\mathcal{F}_{ad}.$$

Since $\gamma > 0$, a standard discussion gives

$$f_*(x,t) = P_F(-\gamma^{-1}p_*(x,t)) \quad \text{a.e. on } Q, \qquad (3.10)$$

where $P_F:\ \mathbb{R}\to F$ 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* $u_d,\varphi_d\in L^\mu(Q)$, $\mu\in[2,\infty)$, *then there exists a unique solution* $(p,\psi)\in W_\nu^{2,1}(Q)\times W_\nu^{2,1}(Q)$ *of the adjoint system (3.7) to (3.9). Moreover, the solution obeys*

$$||p||_{W_\nu^{2,1}} + ||\psi||_{W_\nu^{2,1}} \leq C(||u_* - u_d||_{L^\nu} + ||\varphi_* - \varphi_d||_{L^\nu})\,,$$

where $\nu = \min\{\mu, \frac{5q}{5-2q}\}$ *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* φ_*.

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

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

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

$$\mathcal{L}_{vv}(v_*, \lambda_*)[v, v] \geq \sigma_* \big(\|u\|_{L^2}^2 + \|\varphi\|_{L^2}^2 + \|f\|_{L^2}^2 \big) \tag{3.11}$$

for all $v = (u, \varphi, f) \in V$ satisfying the linearized state equation

$$\begin{aligned} \frac{\partial}{\partial t} u + \frac{\ell}{2} \frac{\partial}{\partial t} \varphi &= \kappa \Delta u + f, \\ \tau \frac{\partial}{\partial t} \varphi &= \xi^2 \Delta \varphi + g'(\varphi_*) \varphi + 2u, \end{aligned} \qquad \text{in } \Omega \times (0, T], \tag{3.12}$$

with *homogeneous* initial and boundary conditions. Note that Theorem 2.4 shows that the solution of (3.12) obeys

$$\|f\|_{L^2} \geq C \big(\|u\|_{W_2^{2,1}} + \|\varphi\|_{W_2^{2,1}} \big) .$$

Hence,

$$\mathcal{L}_{vv}(v_*, \lambda_*)[v, v] \geq \tilde{\sigma}_* \|f\|_{L^2}^2 \tag{3.13}$$

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_*, φ_*, f_*) *and that* (3.11) *holds. Then there exist* $\epsilon > 0$ *and* $c > 0$ *such that*

$$J(u, \varphi, f) \geq J(u_*, \varphi_*, f_*) + c \big(\|u - u_*\|_{W_2^{2,1}}^2 + \|\varphi - \varphi_*\|_{W_2^{2,1}}^2 + \|f - f_*\|_{L^2}^2 \big)$$

for all feasible (u, φ, f) *with* $\|f - f_*\|_{L^2} \leq \epsilon$.

Proof: Let $f \in \mathcal{F}_{ad}$ be given and let (u, φ) be the corresponding solution of the state equation.

First we note that (2.3) implies the estimate

$$\|u - u_*\|_{W_2^{2,1}} + \|\varphi - \varphi_*\|_{W_2^{2,1}} \leq C \|f - f_*\|_{L^2} \leq C \epsilon. \tag{3.14}$$

The first order necessary conditions (3.3)–(3.5) imply

$$\begin{aligned} J(v) &= \mathcal{L}(v, \lambda) \\ &\geq J(v_*) + \frac{1}{2} \mathcal{L}_{vv}(v_*, \lambda_*)[v - v_*, v - v_*] + r(v_*, \lambda_*)[v - v_*, v - v_*] , \end{aligned}$$

where

$$\begin{aligned} \mathcal{L}_{vv}(v_*, \lambda_*)[v - v_*, v - v_*] &= \int_0^T \langle g''(\varphi_*)(\varphi - \varphi_*)^2, \psi_* \rangle \, dt + J_{vv}(v_*)[v - v_*, v - v_*] \\ &= \int_0^T \langle (2b - 6c\varphi_*)(\varphi - \varphi_*)^2, \psi_* \rangle \, dt \\ &\quad + \int_0^T \alpha \|u - u_*\|^2 + \beta \|\varphi - \varphi_*\|^2 + \gamma \|f - f_*\|^2 \, dt \end{aligned} \tag{3.15}$$

and the remainder term is given by

$$\begin{aligned} r(v_*, \lambda_*)[v - v_*, v - v_*] &= \int_0^1 \int_0^\tau \int_0^T \langle [g''(\varphi_* + \eta(\varphi - \varphi_*)) - g''(\varphi_*)](\varphi - \varphi_*)^2, \psi_* \rangle \, dt \, d\eta \, d\tau \\ &= \int_0^1 \int_0^\tau \int_0^T \langle -6c\eta(\varphi - \varphi_*)^3, \psi_* \rangle \, dt \, d\eta \, d\tau \end{aligned} \tag{3.16}$$

cf. (3.2). Thus, using (2.3) and Theorem 2.2,

$$\left| r(v_*,\lambda_*)[v-v_*,v-v_*] \right| \le C||\psi_*||_{L^2}||\varphi-\varphi_*||^3_{L^6} \le C||f-f_*||^3_{L^2}\,. \tag{3.17}$$

Let $(\widehat{u},\widehat{\varphi})$ be the solution of (2.4) with $f_1 = \widehat{f} = f - f_*$, $f_2 = 0$, $\beta_1 = 0$, $\beta_2 = g'(\varphi_*)$, and homogeneous initial and boundary conditions, i.e. let $\widehat{v} = (\widehat{u},\widehat{\varphi},\widehat{f})$ solve the linearized state equation. Then Theorem 2.4 and (2.3) give the estimate

$$||\widehat{u}||_{L^{10}} + ||\widehat{\varphi}||_{L^{10}} \le C(||\widehat{u}||_{W_2^{2,1}} + ||\widehat{\varphi}||_{W_2^{2,1}}) \le C||f-f_*||_{L^2}\,. \tag{3.18}$$

The pair

$$(\delta u,\delta\varphi) = (u,\varphi) - (u_*,\varphi_*) - (\widehat{u},\widehat{\varphi})$$

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

$$||\delta u||_{L^{10}} + ||\delta\varphi||_{L^{10}} \le C||g(\varphi) - g(\varphi_*) - g'(\varphi_*)(\varphi-\varphi_*)||_{L^2} \tag{3.19}$$

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

$$\begin{aligned} &||g(\varphi) - g(\varphi_*) - g'(\varphi_*)(\varphi-\varphi_*)||_{L^2} \\ &= ||b(\varphi-\varphi_*)^2 - c(\varphi+2\varphi_*)(\varphi-\varphi_*)^2||_{L^2} \\ &\le C||\varphi-\varphi_*||^2_{L^4} + C||(\varphi+2\varphi_*)||_{L^4}||\varphi-\varphi_*||^2_{L^8} \\ &\le C||\varphi-\varphi_*||^2_{L^8} \le C||f-f_*||^2_{L^2}, \end{aligned} \tag{3.20}$$

where C depends only on $||f||_{L^2}$, $||f_*||_{L^2}$ and therefore only on ϵ. Inserting (3.20) into (3.19) and using (2.3) yields

$$||\delta u||_{L^{10}} + ||\delta\varphi||_{L^{10}} \le C||f-f_*||^2_{L^2}\,. \tag{3.21}$$

We set $\delta f = 0$ and $\delta v = (\delta u,\delta\varphi,\delta f)$.

From (3.18) and (3.21) we obtain the inequality

$$\begin{aligned} |\mathcal{L}_{vv}(v_*,\lambda_*)[\widehat{v},\delta v]\,| &= \Big|\int_0^T \langle(2b-6c\varphi_*)\widehat{\varphi}\,\delta\varphi,\psi_*\rangle\,dt\,\Big| + \int_0^T \alpha\langle\widehat{u},\delta u\rangle + \beta\langle\widehat{\varphi},\delta\varphi\rangle\,dt \\ &\le ||2b-6c\varphi_*||_{L^2}||\widehat{\varphi}||_{L^8}\,||\delta\varphi||_{L^8}||\psi_*||_{L^4} + \alpha||\widehat{u}||\,||\delta u|| + \beta||\widehat{\varphi}||\,||\delta\varphi|| \\ &\le C||f-f_*||^3_{L^2}\,. \end{aligned} \tag{3.22}$$

Using similar estimates we find that

$$\begin{aligned} |\mathcal{L}_{vv}(v_*,\lambda_*)[\delta v,\delta v]\,| &\le \Big|\int_0^T \langle(2b-6c\varphi_*)(\delta\varphi)^2,\psi_*\rangle\,dt\,\Big| \\ &\quad + \int_0^T \alpha||\delta u||^2 + \beta||\delta\varphi||^2 + \gamma||\delta f||^2\,dt \\ &\le C||f-f_*||^4_{L^2}\,. \end{aligned} \tag{3.23}$$

Combining the convexity assumption (3.11) and the estimates (3.17), (3.22), (3.23) we can deduce that

$$\begin{aligned} J(v) &\ge J(v_*) + \frac{1}{2}\mathcal{L}_{vv}(v_*,\lambda_*)[v-v_*,v-v_*] + r(v_*,\lambda_*)[v-v_*,v-v_*] \\ &= J(v_*) + \frac{1}{2}\mathcal{L}_{vv}(v_*,\lambda_*)[\widehat{v},\widehat{v}] + \mathcal{L}_{vv}(v_*,\lambda_*)[\widehat{v},\delta v] \\ &\quad + \frac{1}{2}\mathcal{L}_{vv}(v_*,\lambda_*)[\delta v,\delta v] + r(v_*,\lambda_*)[v-v_*,v-v_*] \\ &\ge J(v_*) + \sigma_*||f-f_*||^2_{L^2} - C||f-f_*||^3_{L^2}\,. \end{aligned}$$

With (2.3) this gives the desired result. □

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

Since g is a polynomial of degree three we can view "g" as an operator from $L^6(Q)$ into $L^2(Q)$. In this setting "g" is infinitely often differentiable. Note that we apply "g" to solutions φ of parabolic equations. The regularity of solutions and the imbedding theorem guarantee that $\varphi \in W_2^{2,1}(Q) \subset L^6(Q)$. We make use of this particular form and this will enable us to derive convergence estimates for (u, φ, f) in the space $W_2^{2,1}(Q) \times W_2^{2,1}(Q) \times L^2(Q)$. This is a reasonable choice if the admissible set $\mathcal{F}_{ad}$ is unbounded.

However, we also may view the Nemytski–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 $r_g^{(1)}$ of the Nemytskií–operator "g" is obtained from

$$g(\varphi + h) - g(\varphi) = g'(\varphi)h + \int_0^1 \big(g'(\varphi + \tau h) - g'(\varphi)\big) h \, d\tau$$

and is given by

$$r_g^{(1)}(\varphi, h) = \int_0^1 \big(g'(\varphi + \tau h) - g'(\varphi)\big) h \, d\tau \, .$$

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

$$\frac{||r_g^{(1)}(\varphi, h)||_{L^2}}{||h||_{L^2}} \to 0 \quad \text{as} \quad ||h||_{L^\infty} \to 0 \tag{3.24}$$

can also easily be shown. Similarly, one can show that the second order remainder term $r_g^{(2)}$ given by

$$r_g^{(2)}(\varphi, h) = \frac{1}{2} \int_0^1 \big(g''(\varphi + \tau h) - g''(\varphi)\big) h^2 \, d\tau$$

obeys

$$\frac{||r_g^{(2)}(\varphi, h)||_{L^2}}{||h||_{L^2}^2} \to 0 \quad \text{as} \quad ||h||_{L^\infty} \to 0 \, . \tag{3.25}$$

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:

$$\text{Minimize} \;\; q_c(v) = J_v(v_c)(v - v_c) + \frac{1}{2}\mathcal{L}_{vv}(v_c, \lambda_c)[v - v_c, v - v_c] \tag{4.1}$$

subject to the linearized state equation

$$\frac{\partial}{\partial t}u + \frac{\ell}{2}\frac{\partial}{\partial t}\varphi = \kappa\Delta u + f \tag{4.2}$$

in $\Omega \times (0,T]$,

$$\tau\frac{\partial}{\partial t}\varphi = \xi^2\Delta\varphi + g(\varphi_c) + g'(\varphi_c)(\varphi - \varphi_c) + 2u \tag{4.3}$$

with boundary conditions

$$\frac{\partial}{\partial n}u = 0, \quad \frac{\partial}{\partial n}\varphi = 0, \quad \text{on } \partial\Omega \times (0,T), \tag{4.4}$$

and initial conditions

$$u = u_0, \quad \varphi = \varphi_0 \qquad \text{in } \Omega, \tag{4.5}$$

and subject to the control constraints

$$f \in \mathcal{F}_{ad}\,. \tag{4.6}$$

Lemma 4.1 *Let the assumptions* (A1)–(A3) *be valid. If* $(u_c, \varphi_c) \in W_q^{2,1}(Q)$ *and* $f_+ \in L^q(Q)$*, then there exists a unique solution* $v_+ = (u_+, \varphi_+, f_+)$ *of* (4.2)–(4.5)*. The solution obeys*

$$\|u_+\|_{W_q^{2,1}} + \|\varphi_+\|_{W_q^{2,1}} \le C(\|u_0\|_{W^2_\infty(\Omega)} + \|\varphi_0\|_{W^2_\infty(\Omega)} + \|f_+\|_{L^q} + \|\varphi_c\|_{W_q^{2,1}}). \tag{4.7}$$

The constant C *depends only on* φ_c *and can be chosen independent of* φ_c *if* φ_c *is contained in a* $W_2^{2,1}$*–bounded set.*

Proof: The pair (u_+, φ_+) satisfies the system (2.4) with $f_1 = f_+$, $f_2 = g(\varphi_c) - g'(\varphi_c)\varphi_c$, $\beta_1 = 0$, and $\beta_2 = g'(\varphi_c)$. Since $q \in [2, 5/2)$ implies $5q/(5-2q) \ge 3q$ the imbedding Theorem 2.2 implies that $(u_c, \varphi_c) \in L^{3q}(Q)$ if $q \in [2, 5/2)$, that $(u_c, \varphi_c) \in L^\nu(Q)$ for all $\nu \in [2, \infty)$ if $q = 5/2$, and that $(u_c, \varphi_c) \in L^\infty(Q)$ if $q > 5/2$. Moreover, since g is a polynomial of degree three with coefficients in $L^\infty(Q)$ it holds that

$$\|f_2\|_{L^q} = \|g(\varphi_c) - g'(\varphi_c)\varphi_c\|_{L^q} \le C\|\varphi_c\|_{L^{3q}} \le C\|\varphi_c\|_{W_q^{2,1}}\,.$$

Inserting this estimate into Theorem 2.4 we obtain the desired result. □

The objective of the quadratic subproblem is given by

$$\begin{aligned} q_c(u,\varphi,f) = & \int_0^T\!\!\int_\Omega \alpha(u_c - u_d)(u - u_c) + \beta(\varphi_c - \varphi_d)(\varphi - \varphi_c) + \gamma f_c(f - f_c)\,dxdt \\ & + \frac{1}{2}\int_0^T\!\!\int_\Omega \alpha(u - u_c)^2 + \beta(\varphi - \varphi_c)^2 + \gamma(f - f_c)^2\,dxdt \\ & + \frac{1}{2}\int_0^T\!\!\int_\Omega \psi_c\, g''(\varphi_c)(\varphi - \varphi_c)^2\,dxdt\,. \end{aligned} \tag{4.8}$$

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

$$\int_0^T\!\!\int_\Omega \big(p_+(x,t) + \gamma f_+(x,t)\big)\big(f(x,t) - f_+(x,t)\big)\,dxdt \ge 0 \qquad \forall\, f \in \mathcal{F}_{ad}\,, \tag{4.9}$$

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

$$\begin{aligned} -\tfrac{\partial}{\partial t}p &= \kappa\Delta p + 2\psi + \alpha(u_+ - u_d) \\ -\tau\tfrac{\partial}{\partial t}\psi - \tfrac{\ell}{2}\tfrac{\partial}{\partial t}p &= \xi^2\Delta\psi + g'(\varphi_c)\psi + \psi_c g''(\varphi_c)(\varphi_+ - \varphi_c) + \beta(\varphi_+ - \varphi_d), \end{aligned} \quad \text{in } \Omega \times (0,T], \tag{4.10}$$

with boundary conditions

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

and final conditions

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

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 $(p_+, \psi_+) \in W_\nu^{2,1}(Q) \times W_\nu^{2,1}(Q)$ which obeys

$$\begin{aligned} &||p_+||_{W_\nu^{2,1}} + ||\psi_+||_{W_\nu^{2,1}} \\ &\le C(||u_+ - u_d||_{L^\mu} + ||\varphi_+ - \varphi_d||_{L^\mu} + ||\varphi_+ - \varphi_c||_{L^\mu}) \,. \end{aligned} \tag{4.11}$$

The constant C depends only on φ_c and can be chosen independent of φ_c if φ_c is contained in a $W_2^{2,1}$–bounded set. The parameter ν is given by $\nu = \min\{\mu, \frac{5q}{5-2q}\}$ if $q \in [2, 5/2)$ and $n = 3$, and $\nu = \mu$ if $q \ge 5/2$ or $n = 2$. Here q is determined by the control $f_+ \in L^q(Q)$. 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

$$f_+(x,t) = P_F(-\gamma^{-1} p_+(x,t)) \,. \tag{4.12}$$

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_c 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_*, λ_*) is replaced by a sufficiently close point (v_c, λ_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*

$$\mathcal{L}_{vv}(v_c, \lambda_c)[v, v] \ge \sigma\big(||u||_{L^2}^2 + ||\varphi||_{L^2}^2 + ||f||_{L^2}^2\big) \tag{4.13}$$

for all $v = (u, \varphi, f) \in V$ satisfying

$$\begin{aligned} \tfrac{\partial}{\partial t} u + \tfrac{\ell}{2} \tfrac{\partial}{\partial t} \varphi &= \kappa \Delta u + f \\ \tau \tfrac{\partial}{\partial t} \varphi &= \xi^2 \Delta \varphi + g'(\varphi_c)\varphi + 2u, \end{aligned} \qquad \textit{in } \Omega \times (0,T], \tag{4.14}$$

with homogeneous initial and boundary conditions and for all v_c, λ_c with $||v_c - v_*||_V + ||\lambda_c - \lambda_*||_\Lambda \le \epsilon$.

Proof: Let $\hat{v} = (\hat{u}, \hat{\varphi}, \hat{f})$ satisfy the linearized state equation (3.12). Then, by Theorem 2.4,

$$||\hat{u}||_{W_2^{2,1}} + ||\hat{\varphi}||_{W_2^{2,1}} \le C||\hat{f}||_{L^2} \,. \tag{4.15}$$

The definitions of the Lagrangian and g give

$$\mathcal{L}_{vv}(v_c, \lambda_c)[v, v] \;=\; \int_0^T \alpha||u||^2 + \beta||\varphi||^2 + \gamma||f||^2 + \langle (2b - 6c\varphi_c)\varphi^2, \psi_c\rangle \, dt,$$

cf. (3.1), (3.2). Thus,

$$\begin{aligned} &\big|\mathcal{L}_{vv}(v_*, \lambda_*)[\hat{v}, \hat{v}] - \mathcal{L}_{vv}(v_c, \lambda_c)[\hat{v}, \hat{v}]\big| \\ &= \Big| \int_0^T \langle 6c(\varphi_c - \varphi_*)\hat{\varphi}^2, \psi_*\rangle + \langle (2b - 6c\varphi_c)\hat{\varphi}^2, (\psi_* - \psi_c)\rangle \, dt \Big| \\ &\le C\Big(||\varphi_c - \varphi_*||_{L^2} ||\hat{\varphi}||_{L^6}^2 ||\psi_*||_{L^6} + ||2b - 6c\varphi_c||_{L^2} ||\hat{\varphi}||_{L^6}^2 ||\psi_* - \psi_c||_{L^6} \Big) \,. \end{aligned}$$

If φ_c is in an neighborhood of φ_*, then $||2b - 6c\varphi_c||_{L^2} \le C$. From this bound and (4.15) we find that

$$\begin{aligned} &\big|\mathcal{L}_{vv}(v_*, \lambda_*)[\hat{v}, \hat{v}] - \mathcal{L}_{vv}(v_c, \lambda_c)[\hat{v}, \hat{v}]\big| \\ &\le C\Big(||\varphi_c - \varphi_*||_{L^2} ||\psi_*||_{L^6} + ||\psi_c - \psi_*||_{L^6} \Big) ||\hat{f}||_{L^2}^2 \,. \end{aligned}$$

Thus, there exists $\hat{\epsilon} > 0$ such that

$$\mathcal{L}_{vv}(v_c, \lambda_c)[\hat{v}, \hat{v}] \geq \frac{3}{4}\sigma_*\big(\|\hat{u}\|_{L^2}^2 + \|\hat{\varphi}\|_{L^2}^2 + \|\hat{f}\|_{L^2}^2\big) \tag{4.16}$$

for all $\hat{v} = (\hat{u}, \hat{\varphi}, \hat{f})$ satisfying the linearized state equation (3.12) and for all v_c, λ_c with $\|v_c - v_*\|_V + \|\lambda - \lambda_c\|_\Lambda \leq \hat{\epsilon}$.

Let $v = (u, \varphi, \hat{f})$ satisfy (4.14). Then $\hat{v} - v$ satisfies (2.4) with $f_1 = 0$, $f_2 = \big(g'(\varphi_*) - g'(\varphi_c)\big)\varphi$, $\beta_1 = 0$, $\beta_2 = g'(\varphi_*)$, and homogeneous initial and boundary conditions. Hence, Theorem 2.4 yields

$$\begin{aligned}
&\|u - \hat{u}\|_{W_2^{2,1}} + \|\varphi - \hat{\varphi}\|_{W_2^{2,1}} \\
&\leq C\|\big(g'(\varphi_c) - g'(\varphi_*)\big)\varphi\|_{L^2} \\
&\leq C\|g'(\varphi_c) - g'(\varphi_*)\|_{L^4}\|\varphi\|_{L^4} \\
&\leq C\|2b\big(\varphi_c - \varphi_*\big) + 3c\big(\varphi_c + \varphi_*\big)\big(\varphi_c - \varphi_*\big)\|_{L^4}\|\hat{f}\|_{L^2} \\
&\leq C\Big(\|\varphi_c - \varphi_*\|_{L^4} + \|\varphi_c + \varphi_*\|_{L^8}\|\varphi_c - \varphi_*\|_{L^8}\Big)\|\hat{f}\|_{L^2} \\
&\leq C\|v_c - v_*\|_V\|\hat{f}\|_{L^2}\,.
\end{aligned} \tag{4.17}$$

Suppose that v_c, λ_c obey $\|v_c - v_*\|_V + \|\lambda - \lambda_c\|_\Lambda \leq \hat{\epsilon}$. Using (4.15), (4.16), (4.17), the imbedding Theorem 2.2, and the definition of $\mathcal{L}_{vv}$ we find that

$$\begin{aligned}
&\mathcal{L}_{vv}(v_c, \lambda_c)[v, v] \\
&= \mathcal{L}_{vv}(v_c, \lambda_c)[\hat{v} + (v - \hat{v}), \hat{v} + (v - \hat{v})] \\
&= \mathcal{L}_{vv}(v_c, \lambda_c)[\hat{v}, \hat{v}] + 2\mathcal{L}_{vv}(v_c, \lambda_c)[\hat{v}, (v - \hat{v})] + \mathcal{L}_{vv}(v_c, \lambda_c)[v - \hat{v}, v - \hat{v}] \\
&\geq \frac{3}{4}\sigma_*\big(\|\hat{u}\|_{L^2} + \|\hat{\varphi}\|_{L^2} + \|\hat{f}\|_{L^2}\big) \\
&\quad -C\int_0^T 2\,|\langle(2b - 6c\varphi_c)\hat{\varphi}(\varphi - \hat{\varphi}), \psi_c\rangle| + |\langle(2b - 6c\varphi_c)(\varphi - \hat{\varphi})^2, \psi_c\rangle|\,dt \\
&\geq \frac{3}{4}\sigma_*\big(\|\hat{u}\|_{L^2}^2 + \|\hat{\varphi}\|_{L^2}^2 + \|\hat{f}\|_{L^2}^2\big) \\
&\quad -C\|2b - 6c\varphi_c\|_{L^2}\|\psi_c\|_{L^6}\big(\|\hat{\varphi}\|_{L^6}\|\varphi - \hat{\varphi}\|_{L^6} + \|\varphi - \hat{\varphi}\|_{L^6}^2\big) \\
&\geq \frac{3}{4}\sigma_*\big(\|\hat{u}\|_{L^2}^2 + \|\hat{\varphi}\|_{L^2}^2 + \|\hat{f}\|_{L^2}^2\big) \\
&\quad -C\|\hat{f}\|_{L^2}\|v_c - v_*\|_V\|\hat{f}\|_{L^2} - C\|v_c - v_*\|_V^2\|\hat{f}\|_{L^2}^2\,.
\end{aligned} \tag{4.18}$$

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

The definition (4.1) of q_c 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_c is strictly convex on the null space of the linearized constraints for all v_c, λ_c with $\|v_c - v_*\|_V + \|\lambda_c - \lambda_*\|_\Lambda \leq \epsilon$.*

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 $\|v_c - v_*\|_V + \|\lambda_c - \lambda_*\|_\Lambda \leq \epsilon$ the linear quadratic optimal control problem (4.1)–(4.6) has a unique solution $v_+ = (u_+, \varphi_+, f_+)$ obeying*

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

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_c, λ_c with

$$||v_c - v_*||_V + ||\lambda_c - \lambda_*||_\Lambda \le \min\{1, \tilde{\epsilon}\}, \tag{4.20}$$

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, φ) be the solution of (4.2)–(4.5) with $f = f_c$. We set $v = (u, \varphi, f_c)$. Since v is a feasible point for (4.1)–(4.6), we find that $q_c(v_+) \le q_c(v)$. Using the definition (4.8) of q_c, the imbedding $W_2^{2,1}(Q) \subset L^{10}(Q)$, and Hölder's inequality we can conclude that

$$q_c(v_+) \le q_c(v) \le C(||v - v_c||_V + ||v - v_c||_V^2)\,. \tag{4.21}$$

Next we estimate $v - v_*$. The pair $(u - u_*, \varphi - \varphi_*)$ satisfies (2.4) with $\beta_1 = 0$, $\beta_2 = g'(\varphi_c)$, $f_1 = f_c - f_*$, $f_2 = g(\varphi_c) - g(\varphi_*) + g'(\varphi_c)(\varphi_* - \varphi_c)$, and homogeneous initial and boundary conditions. Using the definition of g, Theorem 2.4, and the imbedding Theorem 2.2, we can show that

$$||u - u_*||_{W_2^{2,1}} + ||\varphi - \varphi_*||_{W_2^{2,1}} \le C(||f_c - f_*||_{L^2} + ||\varphi_c - \varphi_*||_{W_2^{2,1}}) \le C||v_c - v_*||_V\,, \tag{4.22}$$

cf. (3.20). Thus,

$$||v - v_c||_V \le ||v - v_*||_V + ||v_* - v_c||_V \le C||v_c - v_*||_V\,.$$

Inserting this inequality into (4.21) and using $||v_c - v_*||_V < 1$ yields

$$q_c(v_+) \le C(||v_c - v_*||_V + ||v_c - v_*||_V^2) \le C||v_c - v_*||_V\,. \tag{4.23}$$

(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, φ_+^0) be the solution of (4.14) with $f = f_+$ and homogeneous initial and boundary conditions. Moreover, let (u_+^1, φ_+^1) solve

$$\begin{aligned} \frac{\partial}{\partial t}u + \frac{\ell}{2}\frac{\partial}{\partial t}\varphi &= \kappa\Delta u \\ \tau\frac{\partial}{\partial t}\varphi &= \xi^2\Delta\varphi + g(\varphi_c) - g'(\varphi_c)\varphi_c + 2u, \end{aligned} \qquad \text{in } \Omega \times (0, T],$$

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

$$||u_+^1||_{W_2^{2,1}} + ||\varphi_+^1||_{W_2^{2,1}} \le C(||u_0||_{W_\infty^2} + ||\varphi_0||_{W_\infty^2} + ||\varphi_c||_{W_2^{2,1}}). \tag{4.24}$$

Using the definition (4.1) of q, the convexity of the Lagrangian (4.13), and the bounds (4.23), (4.24) we find that with some $C > 0$ depending on $\tilde{\epsilon}$, but independent of v_+ the inequalities

$$\begin{aligned} &\sigma\big(||u_+^0||_{L^2}^2 + ||\varphi_+^0||_{L^2}^2 + ||f_+||_{L^2}^2\big) - C\big(1 + ||u_+^1||_{L^2}^2 + ||\varphi_+^1||_{L^2}^2 + ||u_+||_{L^2} + ||\varphi_+||_{L^2} + ||f_+||_{L^2}\big) \\ \le\;& q_c(v_+) \\ \le\;& C||v_c - v_*||_V\,. \end{aligned}$$

are valid. This gives the boundedness of $||u_+^0||_{L^2}^2$, $||\varphi_+^0||_{L^2}^2$, and $||f_+||_{L^2}^2$. With Lemma 4.1 we can deduce the bound

$$||v_+||_V \le C(1 + ||v_c||_V^{1/2}). \tag{4.25}$$

This shows that the new iterates are bounded, if (4.20) holds true.

(iv) The definition of q_c yields

$$\begin{aligned} q_c(v_+) &= J_v(v_c)(v_+ - v_c) + \frac{1}{2}\mathcal{L}_{vv}(v_c,\lambda_c)[v_+ - v_c, v_+ - v_c] \\ &= J_v(v_c)(v_+ - v_*) + J_v(v_c)(v_* - v_c) + \frac{1}{2}\mathcal{L}_{vv}(v_c,\lambda_c)[v_+ - v_*, v_+ - v_*] \\ &\quad +\mathcal{L}_{vv}(v_c,\lambda_c)[v_+ - v_*, v_* - v_c] + \frac{1}{2}\mathcal{L}_{vv}(v_c,\lambda_c)[v_* - v_c, v_* - v_c] \\ &\geq J_v(v_c)(v_+ - v_*) + \frac{1}{2}\mathcal{L}_{vv}(v_c,\lambda_c)[v_+ - v_*, v_+ - v_*] - C\|v_c - v_*\|_V \,. \end{aligned} \tag{4.26}$$

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

$$J_v(v_c)(v_+ - v_*) \geq -C\,\|v_c - v_*\|_V , \tag{4.27}$$

$$\mathcal{L}_{vv}(v_c,\lambda_c)[v_+ - v_*, v_+ - v_*] \geq \frac{\sigma_*}{2}\|v_+ - v_*\|_V^2 - C\left(\|v_c - v_*\|_V + \|\lambda_c - \lambda_*\|_\Lambda\right), \tag{4.28}$$

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

$$q_c(v_+) \geq \frac{\sigma_*}{2}\|v_+ - v_*\|_V^2 - C\|v_c - v_*\|_V - C\|\lambda_c - \lambda_*\|_\Lambda \,. \tag{4.29}$$

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

$$J_v(v_*)(v - v_*) \geq 0 \tag{4.30}$$

for all $v = (u, \varphi, f)$ with $f \in \mathcal{F}_{ad}$ and

$$\left.\begin{aligned} \tfrac{\partial}{\partial t}(u - u_*) + \tfrac{\ell}{2}\tfrac{\partial}{\partial t}(\varphi - \varphi_*) &= \kappa\Delta(u - u_*) + (f - f_*), \\ \tau\tfrac{\partial}{\partial t}(\varphi - \varphi_*) &= \xi^2\Delta(\varphi - \varphi_*) + g'(\varphi_*)(\varphi - \varphi_*) + 2(u - u_*), \end{aligned}\right\} \text{ in } \Omega \times (0,T], \tag{4.31}$$

with homogeneous initial and boundary conditions.

Let $\widehat{v} = (\widehat{u}, \widehat{\varphi}, f_+)$ be defined by (4.31) with $f = f_+$.

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

$$f_2 = g(\varphi_c) - g(\varphi_*) - g'(\varphi_*)(\varphi_+ - \varphi_*) + g'(\varphi_c)(\varphi_+ - \varphi_c), \tag{4.32}$$

and homogeneous initial and boundary conditions.

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

$$\begin{aligned} f_2 &= g(\varphi_c) - g(\varphi_*) - g'(\varphi_*)(\varphi_+ - \varphi_*) + g'(\varphi_c)(\varphi_+ - \varphi_c) \\ &= g(\varphi_c) - g(\varphi_*) - g'(\varphi_*)(\varphi_c - \varphi_*) - g'(\varphi_*)(\varphi_+ - \varphi_c) + g'(\varphi_c)(\varphi_+ - \varphi_c), \end{aligned}$$

the boundedness of v_+, v_c, the definition of g and Hölder inequality, we can show that

$$\|f_2\|_{L^2} \leq C\|\varphi_c - \varphi_*\|_{W_2^{2,1}} \leq C\|v_c - v_*\|_V \,.$$

Thus, Theorem 2.4 yields

$$\|(v_+ - v_*) - (\widehat{v} - v_*)\| = \|v_+ - \widehat{v}\|_V \leq C\|v_c - v_*\|_V \,. \tag{4.33}$$

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

$$\begin{aligned} J_v(v_c)(v_+ - v_*) &= J_v(v_*)(v_+ - v_*) + \big(J_v(v_c) - J_v(v_*)\big)(v_+ - v_*), \\ &\geq J_v(v_*)(v_+ - v_*) - C\,\|v_c - v_*\|_V, \\ &= J_v(v_*)(\widehat{v} - v_*) + J_v(v_*)(v_+ - v_* - (\widehat{v} - v_*)) - C\,\|v_c - v_*\|_V, \\ &\geq J_v(v_*)(\widehat{v} - v_*) - C\,\|v_c - v_*\|_V\,. \end{aligned}$$

This proves (4.27).

The estimate (4.28) can be derived in a similar way. The boundedness of $\|v_+ - v_*\|_V$ yields

$$\mathcal{L}_{vv}(v_c,\lambda_c)[v_+ - v_*, v_+ - v_*] \geq \mathcal{L}_{vv}(v_*,\lambda_*)[v_+ - v_*, v_+ - v_*] - C\big(\|v_c - v_*\|_V + \|\lambda_c - \lambda_*\|_\Lambda\big), \tag{4.34}$$

cf. (4.18). As before, let $\widehat{v} = (\widehat{u}, \widehat{\varphi}, f_+)$ be defined by (4.31) with $f = f_+$, i.e. $\widehat{v} - v_*$ satisfies the linearized state equation. Using (3.11) we can show that

$$\begin{aligned} &\mathcal{L}_{vv}(v_*,\lambda_*)[v_+ - v_*, v_+ - v_*] \\ &= \mathcal{L}_{vv}(v_*,\lambda_*)[\widehat{v} - v_*, \widehat{v} - v_*] + 2\mathcal{L}_{vv}(v_*,\lambda_*)[\widehat{v} - v_*, v_+ - \widehat{v}] + \mathcal{L}_{vv}(v_*,\lambda_*)[v_+ - \widehat{v}, v_+ - \widehat{v}] \\ &\geq \mathcal{L}_{vv}(v_*,\lambda_*)[\widehat{v} - v_*, \widehat{v} - v_*] - C\big(\|\widehat{v} - v_*\|_V\|\widehat{v} - v_+\|_V + \|v_+ - \widehat{v}\|_V^2\big) \\ &\geq \sigma_*\|\widehat{v} - v_*\|_V^2 - C\big(\|\widehat{v} - v_*\|_V\|\widehat{v} - v_+\|_V + \|v_+ - \widehat{v}\|_V^2\big) \\ &\geq \sigma_*\|\widehat{v} - v_*\|_V^2 - C\|\widehat{v} - v_+\|_V \\ &\geq \sigma_*\|\widehat{v} - v_*\|_V^2 - C\|v_c - v_*\|_V\,. \end{aligned} \tag{4.35}$$

In the last two estimates of (4.35) we have used the boundedness of v_+ and the inequality (4.33). If we use $\widehat{v} = v_+ + (\widehat{v} - v_+)$, (4.25), and (4.33), the inequality (4.35) yields

$$\begin{aligned} \mathcal{L}_{vv}(v_*,\lambda_*)[v_+ - v_*, v_+ - v_*] &\geq \sigma_*\|\widehat{v} - v_*\|_V^2 - C\|v_c - v_*\|_V \\ &\geq \sigma_*\|v_+ - v_*\|_V^2 - C\|v_c - v_*\|_V\,. \end{aligned} \tag{4.36}$$

This lower bound, together with (4.34) yields the desired estimate (4.28).

(vi) It remains to prove the estimate for the Lagrange multiplier λ_+. This estimate follows easily from (4.11), (4.20), and (4.25). In fact, we find that

$$\|p_+\|_{W_2^{2,1}} + \|\psi_+\|_{W_2^{2,1}} \leq C(\|u_+ - u_d\|_{L^2} + \|\varphi_+ - \varphi_d\|_{L^2} + \|\varphi_+ - \varphi_c\|_{L^2}) \leq C\,.$$

This concludes the proof. □

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 $v_c, v_+, v_* \in V_q$ and $\lambda_c, \lambda_+, \lambda_* \in \Lambda_q$, where V_q and Λ_q 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 $u_d, \varphi_d \in L^q(Q)$, and if $f_c, f_+, f_* \in L^q(Q)$, cf. Theorems 2.1, 3.1 and equations (4.7), (4.11). Since $\gamma > 0$, the conditions $f_c, f_+, f_* \in L^q(Q)$ are implied by the regularity $\lambda_+, \lambda_* \in \Lambda_q$ of the adjoint variables, cf. (4.9), (4.12).

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

If for $q \in [2,\infty)$ the iterates and the optimal point satisfy $v_c, v_+, v_* \in V_q$, then there exists $\epsilon > 0$ such that $\|v_c - v_*\|_V + \|\lambda_c - \lambda_*\|_\Lambda \leq \epsilon$ implies the estimates

$$\|v_+ - v_*\|_{V_q} \leq C\left(\|v_c - v_*\|_V + \|\lambda_c - \lambda_*\|_\Lambda\right)^{1/q} \tag{4.37}$$

and

$$\|\lambda_+ - \lambda_*\|_{\Lambda_q} \leq C\big(\|v_c - v_*\|_V + \|\lambda_c - \lambda_*\|_\Lambda\big)^{1/q}\,. \tag{4.38}$$

Moreover, if $n = 3$ and $q > 5/2$ or if $n = 2$ and $q > 2$, then (4.37) *can be replaced by*

$$\|u_+ - u_*\|_{W_q^{2,1}} + \|\varphi_+ - \varphi_*\|_{W_q^{2,1}} + \|f_+ - f_*\|_{L^\infty} \leq C\left(\|v_c - v_*\|_V + \|\lambda_c - \lambda_*\|_\Lambda\right)^{1/q}. \tag{4.39}$$

Proof: Using (4.19) and the boundedness of $\mathcal{F}_{ad}$ we can conclude that

$$\begin{aligned} \|f_+ - f_*\|_{L^q}^q &= \int_0^T\int_\Omega |f_+ - f_*|^2\, |f_+ - f_*|^{q-2}\, dxdt \;\le\; C\|f_+ - f_*\|_{L^2}^2 \\ &\le C\left(\|v_c - v_*\|_V + \|\lambda_c - \lambda_*\|_\Lambda\right). \end{aligned} \tag{4.40}$$

The difference $v_+ - v_*$ satisfies (2.4) with $f_1 = f_+ - f_*$, $f_2 = g(\varphi_c) - g(\varphi_*) + g'(\varphi_c)(\varphi_* - \varphi_c)$, $\beta_1 = 0$, $\beta_2 = g'(\varphi_c)$, and homogeneous initial and boundary conditions. Using the definition of g, the right hand side f_2 can be estimated by

$$\|f_2\|_{L^q} \le C\|\varphi_c - \varphi_*\|_{L^{3q}}.$$

Thus, from Theorem 2.4 and (4.40) we obtain

$$\begin{aligned} &\|u_+ - u_*\|_{W_q^{2,1}} + \|\varphi_+ - \varphi_*\|_{W_q^{2,1}} \;\le\; C(\|f_1\|_{L^q} + \|f_2\|_{L^q}) \\ &\quad\le C(\|f_+ - f_*\|_{L^q} + \|\varphi_c - \varphi_*\|_{L^{3q}}) \\ &\quad\le C\left(\|v_c - v_*\|_V + \|\lambda_c - \lambda_*\|_\Lambda\right)^{1/q} + C\|\varphi_c - \varphi_*\|_{L^{3q}}. \end{aligned} \tag{4.41}$$

Note that $W_q^{2,1}(Q) \subset L^{3q}(Q)$ for $q \in [2, 5/2)$ and $W_q^{2,1}(Q) \subset L^\nu(Q)$ for arbitrary $\nu \in [2,\infty)$ otherwise, cf. Theorem 2.2. If $\epsilon < 1$, then (4.41) yields

$$\|u_+ - u_*\|_{W_q^{2,1}} + \|\varphi_+ - \varphi_*\|_{W_q^{2,1}} \le C\left(\|v_c - v_*\|_V + \|\lambda_c - \lambda_*\|_\Lambda\right)^{1/q} \tag{4.42}$$

for all v_c, λ_c with $\|v_c - v_*\|_V + \|\lambda_c - \lambda_*\|_\Lambda < \epsilon$.

The adjoints can be estimated in a similar way. In fact, the difference $\lambda_+ - \lambda_*$ satisfies

$$\begin{aligned} -\tfrac{\partial}{\partial t}p &= \kappa\Delta p + 2\psi + \alpha(u_+ - u_*), \\ -\tau\tfrac{\partial}{\partial t}\psi - \tfrac{\ell}{2}\tfrac{\partial}{\partial t}p &= \xi^2\Delta\psi + g'(\varphi_*)\psi + (g'(\varphi_c) - g'(\varphi_*))\psi_+ + g''(\varphi_c)(\varphi_+ - \varphi_c) + \beta(\varphi_+ - \varphi_*), \end{aligned} \tag{4.43}$$

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 Λ, respectively, and Theorem 2.4 this yields

$$\begin{aligned} &\|p_+ - p_*\|_{W_q^{2,1}} + \|\psi_+ - \psi_*\|_{W_q^{2,1}} \\ &\quad\le C(\|f_+ - f_*\|_{L^q} + \|\varphi_c - \varphi_*\|_{L^{3q}}) \\ &\quad\le C\left(\|v_c - v_*\|_V + \|\lambda_c - \lambda_*\|_\Lambda\right)^{1/q} + C\|\varphi_c - \varphi_*\|_{L^{3q}} \\ &\quad\le C\left(\|v_c - v_*\|_V + \|\lambda_c - \lambda_*\|_\Lambda\right)^{1/q}, \end{aligned} \tag{4.44}$$

provided $\|v_c - v_*\|_V + \|\lambda_c - \lambda_*\|_\Lambda \le 1$.

Suppose that $n = 3$, $q > 5/2$ or $n = 2$, $q > 2$. For the estimate of $\|f_+ - f_*\|_{L^\infty}$ we use the first order optimality conditions. From (3.10), (4.12), (4.42), and the imbedding $W_q^{2,1}(Q) \subset L^\infty(Q)$ we can deduce that

$$\begin{aligned} |f_+(x,t) - f_*(x,t)| &= \left|P_F(-\gamma^{-1}p_+(x,t)) - P_F(-\gamma^{-1}p_*(x,t))\right| \\ &\le \gamma^{-1}\left|p_+(x,t) - p_*(x,t)\right| \\ &\le C\left(\|v_c - v_*\|_V + \|\lambda_c - \lambda_*\|_\Lambda\right)^{1/q}. \end{aligned} \tag{4.45}$$

□

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_*, φ_*, f_*). To obtain convergence estimates we have to study the dependence of the solution upon the perturbation $\pi = (\pi_s, \pi_a) \in \big(L^q(Q)\big)^2$. Here π_s will be a perturbation on the right hand side of the linearized state equation and π_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:

$$\text{Minimize } \; q_*(v) + \langle \pi_a, \varphi \rangle_{L^2} = J_v(v_*)(v - v_*) + \frac{1}{2}\mathcal{L}_{vv}(v_*, \lambda_*)[v - v_*, v - v_*] + \langle \pi_a, \varphi \rangle_{L^2} \tag{5.1}$$

subject to the linearized state equation

$$\begin{aligned} \frac{\partial}{\partial t}u + \frac{\ell}{2}\frac{\partial}{\partial t}\varphi &= \kappa \Delta u + f \\ \tau \frac{\partial}{\partial t}\varphi &= \xi^2 \Delta\varphi + g(\varphi_*) + g'(\varphi_*)(\varphi - \varphi_*) + 2u + \pi_s \end{aligned} \qquad \text{in } \Omega \times (0, T], \tag{5.2}$$

with boundary conditions

$$\frac{\partial}{\partial n}u = 0, \quad \frac{\partial}{\partial n}\varphi = 0, \quad \text{ on } \partial\Omega \times (0, T), \tag{5.3}$$

and initial conditions

$$u = u_0, \quad \varphi = \varphi_0 \qquad \text{in } \Omega, \tag{5.4}$$

and subject to the control constraints

$$f \in \mathcal{F}_{ad}\,. \tag{5.5}$$

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

$$\begin{aligned} q_*(u, \varphi, f) + \langle \pi_a, \varphi \rangle_{L^2} \;=\; & \int_0^T\!\!\int_\Omega \alpha(u_* - u_d)(u - u_*) + \beta(\varphi_* - \varphi_d)(\varphi - \varphi_*) + \gamma f_*(f - f_*)\, dxdt \\ & + \frac{1}{2}\int_0^T\!\!\int_\Omega \alpha(u - u_*)^2 + \beta(\varphi - \varphi_*)^2 + \gamma(f - f_*)^2\, dxdt \\ & + \frac{1}{2}\int_0^T\!\!\int_\Omega \psi_*\, g''(\varphi_*)(\varphi - \varphi_*)^2\; dxdt + \int_0^T\!\!\int_\Omega \pi_a \varphi\, dxdt\,. \end{aligned} \tag{5.6}$$

Since the second order sufficient optimality conditions are satisfied at (u_*, φ_*, 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

$$\int_0^T\!\!\int_\Omega \big(p_\pi(x,t) + \gamma f_\pi(x,t)\big)\big(f(x,t) - f_\pi(x,t)\big)\; dxdt \ge 0 \qquad \forall\, f \in \mathcal{F}_{ad}\,, \tag{5.7}$$

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

$$\begin{aligned} -\tfrac{\partial}{\partial t}p &= \kappa\Delta p + 2\psi + \alpha(u_\pi - u_d), \\ -\tau\tfrac{\partial}{\partial t}\psi - \tfrac{\ell}{2}\tfrac{\partial}{\partial t}p &= \xi^2\Delta\psi + g'(\varphi_*)\psi + \psi_* g''(\varphi_*)(\varphi_\pi - \varphi_*) + \beta(\varphi_\pi - \varphi_d) + \pi_a, \end{aligned} \qquad \text{in } \Omega \times (0, T], \tag{5.8}$$

with boundary conditions

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

and final conditions

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

If $\gamma > 0$, then (5.7) is equivalent to

$$f_\pi(x,t) = P_F(-\gamma^{-1}p_\pi(x,t))\,. \tag{5.9}$$

Using Theorem 2.4 one can see that the solution (u_π, φ_π) of (5.2) satisfies

$$\begin{aligned}&\|u_\pi\|_{W_q^{2,1}} + \|\varphi_\pi\|_{W_q^{2,1}} \\ &\le C(\|u_0\|_{W^2_\infty(\Omega)} + \|\varphi_0\|_{W^2_\infty(\Omega)} + \|f_\pi\|_{L^q} + \|\pi_s\|_{L^q}) ,\end{aligned} \tag{5.10}$$

provided $f_\pi, \pi_s \in L^q(Q)$. Moreover, if $\pi_a \in L^q(Q)$, then the solution of the adjoint equation (5.8) satisfies

$$\begin{aligned}&\|p_\pi\|_{W_\nu^{2,1}} + \|\psi_\pi\|_{W_\nu^{2,1}} \\ &\le C\Big(\|u_\pi - u_d\|_{L^\mu} + \|\varphi_\pi - \varphi_*\|_{W_q^{2,1}} + \|\varphi_\pi - \varphi_d\|_{L^\mu} + \|\pi_a\|_{L^q}\Big) .\end{aligned} \tag{5.11}$$

where the parameter ν is given as in (4.11).

Lemma 5.1 *Let the assumptions* (A1)–(A3) *be valid. If* $u_\pi, u_*, \varphi_\pi, \varphi_* \in W_q^{2,1}(Q)$*, then there exists a constant* $C > 0$ *such that*

$$\|\lambda_\pi - \lambda_*\|_{\Lambda_q} \le C\big(\|u_\pi - u_*\|_{W_q^{2,1}} + \|\varphi_\pi - \varphi_*\|_{W_q^{2,1}} + \|\pi_a\|_{L^q}\big) .$$

Proof: The difference $(p, \psi) = \lambda = (\lambda_\pi - \lambda_*)$ satisfies the system

$$\begin{aligned} -\tfrac{\partial}{\partial t}p &= \kappa\Delta p + 2\psi + \alpha(u_\pi - u_*), \\ -\tau\tfrac{\partial}{\partial t}\psi - \tfrac{\ell}{2}\tfrac{\partial}{\partial t}p &= \xi^2\Delta\psi + g'(\varphi_*)\psi + \beta(\varphi_\pi - \varphi_*) + \psi_* g''(\varphi_*)(\varphi_\pi - \varphi_*) + \pi_a, \end{aligned} \quad \text{in } \Omega\times(0,T],$$

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 $q \in [2, 5/2)$, $\|\psi_* g''(\varphi_*)(\varphi_\pi - \varphi_*)\|_{L^q} \le \|\psi_*\|_{L^{3q}}\|g''(\varphi_*)\|_{L^{3q}}\|\varphi_\pi - \varphi_*\|_{L^{3q}}$. 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*

$$\|v_\pi - v_*\|_V \le C\|\pi\|_{(L^2)^2}$$

for all $\pi \in (L^2(Q))^2$.

Proof: The Lagrange function for the perturbed problem (5.1) to (5.5) is given by

$$\begin{aligned}\widetilde{\mathcal{L}}(v, \lambda_\pi) &= J_v(v_*)(v - v_*) + \frac{1}{2}\mathcal{L}_{vv}(v_*, \lambda_*)[v - v_*, v - v_*] + \langle \pi_a, \varphi\rangle_{L^2} \\ &\quad - \int_0^T \langle u_t + \frac{\ell}{2}\varphi_t, p_\pi\rangle + \kappa\langle\nabla u, \nabla p_\pi\rangle - \langle f, p_\pi\rangle\, dt \\ &\quad - \int_0^T \tau\langle\varphi_t, \psi_\pi\rangle + \xi^2\langle\nabla\varphi, \nabla\psi_\pi\rangle - \langle g(\varphi_*) + g'(\varphi_*)(\varphi - \varphi_*) + 2u + \pi_s, \psi_\pi\rangle\, dt .\end{aligned} \tag{5.12}$$

The first oder necessary conditions for (5.1) to (5.5) are given by

$$\widetilde{\mathcal{L}}_v(v_\pi, \lambda_\pi)(v - v_\pi) \ge 0$$

for all $v = (u, \varphi, f)$ with $f \in \mathcal{F}_{ad}$. Using the definition (5.12) and inserting $v = v_*$ this leads to

$$\begin{aligned} 0 \le\ & J_v(v_*)(v_* - v_\pi) + \mathcal{L}_{vv}(v_*, \lambda_*)[v_\pi - v_*, v_* - v_\pi] + \langle\pi_a, \varphi_* - \varphi_\pi\rangle_{L^2} \\ & - \int_0^T \langle (u_* - u_\pi)_t + \frac{\ell}{2}(\varphi_* - \varphi_\pi)_t, p_\pi\rangle + \kappa\langle\nabla(u_* - u_\pi), \nabla p_\pi\rangle - \langle (f_* - f_\pi), p_\pi\rangle\, dt \\ & - \int_0^T \tau\langle(\varphi_* - \varphi_\pi)_t, \psi_\pi\rangle + \xi^2\langle\nabla(\varphi_* - \varphi_\pi), \nabla\psi_\pi\rangle - \langle g'(\varphi_*)(\varphi_* - \varphi_\pi) + 2(u_* - u_\pi), \psi_\pi\rangle\, dt .\end{aligned} \tag{5.13}$$

The difference $v = v_* - v_\pi = (u_* - u_\pi, \varphi_* - \varphi_\pi, f_* - f_\pi)$ satisfies

$$\begin{array}{rcl} \frac{\partial}{\partial t} u + \frac{\ell}{2}\frac{\partial}{\partial t}\varphi &=& \kappa\Delta u + f, \\ \tau\frac{\partial}{\partial t}\varphi &=& \xi^2\Delta\varphi + g'(\varphi_*)\varphi + 2u - \pi_s, \end{array} \quad \text{in } \Omega \times (0,T], \tag{5.14}$$

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

$$0 \leq J_v(v_*)(v_* - v_\pi) + \mathcal{L}_{vv}(v_*, \lambda_*)[v_\pi - v_*, v_* - v_\pi] + \langle \pi_a, \varphi_* - \varphi_\pi \rangle_{L^2} + \langle \pi_s, \psi_\pi \rangle_{L^2} . \tag{5.15}$$

On the other hand the first order necessary conditions (3.3)–(3.3) for the nonlinear control problem yield that $\mathcal{L}_v(v_*, \lambda_*)(v_\pi - v_*) \geq 0$, or, using the definition (3.1) of $\mathcal{L}$,

$$\begin{array}{rcl} 0 &\leq& J_v(v_*)(v_\pi - v_*) \\ && - \displaystyle\int_0^T \langle (u_\pi - u_*)_t + \frac{\ell}{2}(\varphi_\pi - \varphi_*)_t, p_* \rangle + \kappa\langle \nabla(u_\pi - u_*), \nabla p_* \rangle - \langle (f_\pi - f_*), p_* \rangle \, dt \\ && - \displaystyle\int_0^T \tau\langle (\varphi_\pi - \varphi_*)_t, \psi_* \rangle + \xi^2\langle \nabla(\varphi_\pi - \varphi_*), \nabla\psi_* \rangle - \langle g'(\varphi_*)(\varphi_\pi - \varphi_*) + 2(u_\pi - u_*), \psi_* \rangle \, dt \\ &=& J_v(v_*)(v_\pi - v_*) - \langle \pi_s, \psi_* \rangle_{L^2} . \end{array}$$

Thus, $J_v(v_*)(v_* - v_\pi) \leq -\langle \pi_s, \psi_* \rangle_{L^2}$. Inserting this into (5.15) and using Lemma 5.1 we find that

$$\begin{array}{rl} & \mathcal{L}_{vv}(v_*, \lambda_*)[v_\pi - v_*, v_\pi - v_*] \\ \leq & \langle \pi_a, \varphi_* - \varphi_\pi \rangle_{L^2} - \langle \pi_s, (\psi_* - \psi_\pi) \rangle_{L^2} \\ \leq & \|\pi_a\|_{L^2} \|\varphi_* - \varphi_\pi\|_{L^2} + \|\pi_s\|_{L^2} \|\psi_* - \psi_\pi\|_{L^2} \\ \leq & C\big(\|\pi_a\|_{L^2} + \|\pi_s\|_{L^2}\big)\big(\|u_* - u_\pi\|_{W_2^{2,1}} + \|\varphi_* - \varphi_\pi\|_{W_2^{2,1}}\big) \\ & + C\big(\|\pi_a\|_{L^2} + \|\pi_s\|_{L^2}\big)^2 . \end{array} \tag{5.16}$$

Next we derive a lower bound for $\mathcal{L}_{vv}(v_*, \lambda_*)[v_\pi - v_*, v_\pi - v_*]$. Let (u, φ) satisfy the linearized state equation (3.12) with $f = f_\pi - f_*$. The second order sufficient condition (3.11) implies

$$\mathcal{L}_{vv}(v_*, \lambda_*)[v, v] \geq \sigma_*\big(\|u\|_{L^2}^2 + \|\varphi\|_{L^2}^2 + \|f_\pi - f_*\|_{L^2}^2\big), \tag{5.17}$$

where $v = (u, \varphi, f_\pi - f_*)$. Moreover, the triple $\widehat{v} = v - (v_\pi - v_*)$ satisfies

$$\begin{array}{rcl} \frac{\partial}{\partial t}\widehat{u} + \frac{\ell}{2}\frac{\partial}{\partial t}\widehat{\varphi} &=& \kappa\Delta\widehat{u} \\ \tau\frac{\partial}{\partial t}\widehat{\varphi} &=& \xi^2\Delta\widehat{\varphi} + g'(\varphi_*)\widehat{\varphi} + 2\widehat{u} - \pi_s \end{array} \quad \text{in } \Omega \times (0,T]$$

with homogeneous initial and boundary conditions. Theorem 2.4 implies that

$$\|\widehat{v}\|_{W_2^{2,1}} \leq C\|\pi_s\|_{L^2} \leq C\|\pi\|_{(L^2)^2} .$$

Using this estimate, the representation $v = \widehat{v} + (v_\pi - v_*)$, and (5.17) we find that

$$\begin{array}{rl} & \mathcal{L}_{vv}(v_*, \lambda_*)[v_\pi - v_*, v_\pi - v_*] \\ \geq & \sigma_*\big(\|u\|_{L^2}^2 + \|\varphi\|_{L^2}^2 + \|f_\pi - f_*\|_{L^2}^2\big) + 2\mathcal{L}_{vv}(v_*, \lambda_*)[\widehat{v}, v] + \mathcal{L}_{vv}(v_*, \lambda_*)[\widehat{v}, \widehat{v}] \\ \geq & \sigma_*\|f_\pi - f_*\|_{L^2}^2 - C\|\widehat{v}\|_V \|v\| - C\|\widehat{v}\|_V^2 \\ \geq & \sigma_*\|f_\pi - f_*\|_{L^2}^2 - C\big(\|\pi\|_{(L^2)^2}\|v_\pi - v_*\|_V + \|\pi\|_{(L^2)^2}^2\big). \end{array} \tag{5.18}$$

Combining (5.16) and (5.18) yields

$$\|f_\pi - f_*\|_{L^2}^2 \leq C\big(\|\pi\|_{(L^2)^2}\|v_\pi - v_*\|_V + \|\pi\|_{(L^2)^2}^2\big). \tag{5.19}$$

Since $v = v_* - v_\pi$ satisfies (5.14), Theorem 2.4 implies the estimate

$$||u_\pi - u_*||_{W_2^{2,1}} + ||\varphi_\pi - \varphi_*||_{W_2^{2,1}} \le C\big(||f_\pi - f_*||_{L^2} + ||\pi||_{(L^2)^2}\big)\,. \tag{5.20}$$

Inserting this into (5.19) gives

$$||f_\pi - f_*||_{L^2}^2 \le C\big(||\pi||_{(L^2)^2}||f_\pi - f_*||_V + ||\pi||_{(L^2)^2}^2\big). \tag{5.21}$$

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

$$||f_\pi - f_*||_{L^2} \le c\,||\pi||_{(L^2)^2}\,. \tag{5.22}$$

The assertion now follows with (5.20). □

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

If for $q \in [2, \infty)$ the solution of the perturbed problem and the optimal solution satisfy $v_\pi, v_* \in V_q$, then

$$||v_\pi - v_*||_{V_q} \le C||\pi||_{(L^q)^2}$$

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

$$||u_\pi - u_*||_{W_q^{2,1}} + ||\varphi_\pi - \varphi_*||_{W_q^{2,1}} + ||f_\pi - f_*||_{L^\infty} \le C\,||\pi||_{(L^q)^2}\,.$$

Proof: As in the proof of Lemma 4.5 we can conclude that

$$||f_\pi - f_*||_{L^q} \le C\,||\pi||_{(L^2)^2}\,.$$

Since $v = v_* - v_\pi$ satisfies (5.14), we can appeal to Theorem 2.4 to show that

$$||u_\pi - u_*||_{W_q^{2,1}} + ||\varphi_\pi - \varphi_*||_{W_q^{2,1}} \le C\big(||f_\pi - f_*||_{L^q} + ||\pi||_{(L^q)^2}\big) \le C||\pi||_{(L^q)^2}\,.$$

This gives the first estimate. We can use this inequality and Lemma 5.1 to conclude that

$$||\lambda_\pi - \lambda_*||_{\Lambda_q} \le C||\pi||_{(L^q)^2}\,.$$

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 $||v_c - v_*||_V + ||\lambda_c - \lambda_*||_\Lambda \le \epsilon$ the solutions of the SQP subproblem (4.1) to (4.6) are solutions of the perturbed subproblem (5.1) to (5.5) with*

$$\pi_s = g(\varphi_c) - g(\varphi_*) + g'(\varphi_c)(\varphi_+ - \varphi_c) - g'(\varphi_*)(\varphi_+ - \varphi_*)\,, \tag{5.23}$$

$$\pi_a = g'(\varphi_c)\psi_+ + \psi_c g''(\varphi_c)(\varphi_+ - \varphi_c) - g'(\varphi_*)\psi_+ - \psi_* g''(\varphi_*)(\varphi_+ - \varphi_*)\,. \tag{5.24}$$

Proof: From Lemma 4.4 we know that there exists $\epsilon > 0$ such that if $||v_c - v_*||_V + ||\lambda_c - \lambda_*||_\Lambda \le \epsilon$ there exists a unique solution (u_+, φ_+, 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_+, φ_+, f_+) satisfies the perturbed system (5.2)–(5.4) with π_s given in (5.23). Moreover, one can see that $(u_+, \varphi_+, p_+, \psi_+)$ satisfies the perturbed adjoint system (5.8) with π_a given in (5.24). This implies the assertion. □

Lemma 5.5 *If the assumptions of Lemma 5.4 are valid, and if $\varphi_c, \varphi_+, \varphi_*, \psi_c, \psi_+, \psi_* \in W_q^{2,1}(Q)$, then*

$$\|\pi_s\|_{L^q} \leq C\big(\|\varphi_c - \varphi_*\|^2_{W_q^{2,1}} + \|\varphi_c - \varphi_*\|_{W_q^{2,1}} \|\varphi_+ - \varphi_*\|_{W_q^{2,1}}\big), \tag{5.25}$$

$$\begin{aligned} \|\pi_a\|_{L^q} \leq\ & C\big[\|\varphi_c - \varphi_*\|^2_{W_q^{2,1}} \|\psi_+\|_{W_q^{2,1}} + \|\psi_+ - \psi_*\|_{W_q^{2,1}} \|\varphi_c - \varphi_*\|_{W_q^{2,1}} \\ & + \big(\|\psi_c - \psi_*\|_{W_q^{2,1}} + \|\varphi_c - \varphi_*\|_{W_q^{2,1}}\big)\big(\|\varphi_+ - \varphi_*\|_{W_q^{2,1}} + \|\varphi_c - \varphi_*\|_{W_q^{2,1}}\big)\big]. \end{aligned} \tag{5.26}$$

The constant C depends on $\varphi_c, \varphi_+, \varphi_*, \psi_c, \psi_+, \psi_*$, but is uniformly bounded if $\varphi_c, \varphi_+, \varphi_*, \psi_c, \psi_+, \psi_*$ are contained in a bounded set in $W_q^{2,1}(Q)$.

Proof: From (5.23) we find that

$$\begin{aligned} \pi_s &= g(\varphi_c) - g(\varphi_*) + g'(\varphi_c)(\varphi_+ - \varphi_c) - g'(\varphi_*)(\varphi_+ - \varphi_*), \\ &= g(\varphi_c) - g(\varphi_*) - g'(\varphi_*)(\varphi_c - \varphi_*) + \big(g'(\varphi_c) - g'(\varphi_*)\big)(\varphi_+ - \varphi_c). \end{aligned}$$

Using the definition of g, Hölder's inequality, and the imbedding $W_q^{2,1}(Q) \subset L^{3q}(Q)$ 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

$$\begin{aligned} \pi_a &= g'(\varphi_c)\psi_+ + \psi_c g''(\varphi_c)(\varphi_+ - \varphi_c) - g'(\varphi_*)\psi_+ - \psi_* g''(\varphi_*)(\varphi_+ - \varphi_*) \\ &= \big[g'(\varphi_c) - g'(\varphi_*) - g''(\varphi_*)(\varphi_c - \varphi_*)\big]\psi_+ + g''(\varphi_*)(\psi_+ - \psi_*)(\varphi_c - \varphi_*) \\ &\quad + \big[\psi_c g''(\varphi_c) - \psi_* g''(\varphi_*)\big](\varphi_+ - \varphi_c). \end{aligned}$$

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 $(v_c, \lambda_c) = (u_c, \varphi_c, f_c, p_c, \psi_c)$ 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 $v_c = v_*$. 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 $(v_c, \lambda_c) \in V_q \times \Lambda_q$.*

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

$$\|v_+ - v_*\|_V + \|\lambda_+ - \lambda_*\|_\Lambda \leq C\left(\|v_c - v_*\|_V + \|\lambda_c - \lambda_*\|_\Lambda\right)^2.$$

(ii) *If $q > 2$ and if $\mathcal{F}_{ad}$ is bounded, then there exists $\epsilon > 0$ such that $\|v_c - v_*\|_{V_q} + \|\lambda_c - \lambda_*\|_{\Lambda_q} \leq \epsilon$ implies*

$$\|v_+ - v_*\|_{V_q} + \|\lambda_+ - \lambda_*\|_{\Lambda_q} \leq C\left(\|v_c - v_*\|_{V_q} + \|\lambda_c - \lambda_*\|_{\Lambda_q}\right)^2.$$

(iii) *If $n = 3$ and $q > 5/2$ or if $n = 2$ and $q > 2$ and if $\mathcal{F}_{ad}$ is bounded, then there exists $\epsilon > 0$ such that $\|v_c - v_*\|_{V_q} + \|\lambda_c - \lambda_*\|_{\Lambda_q} \leq \epsilon$ implies*

$$\|u_+ - u_*\|_{W_q^{2,1}} + \|\varphi_+ - \varphi_*\|_{W_q^{2,1}} + \|f_+ - f_*\|_{L^\infty} + \|\lambda_+ - \lambda_*\|_{\Lambda_q} \leq C\left(\|v_c - v_*\|_{V_q} + \|\lambda_c - \lambda_*\|_{\Lambda_q}\right)^2.$$

Proof: By Lemma 4.4 there exists $\epsilon_1 > 0$ such that for all $v_c \in V, \lambda_c \in \Lambda$ with $||v_c - v_*||_V + ||\lambda_c - \lambda_*||_\Lambda \leq \epsilon_1$ there exists a unique solution $v_+ \in V$ of (4.1)–(4.6) with corresponding Lagrange multiplier $\lambda_+ \in \Lambda$. Moreover, the assumptions $u_d, \varphi_d \in L^q(Q)$, $(v_c, \lambda_c) \in V_q \times \Lambda_q$ 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 $||v_+||_V, ||\lambda_+||_\Lambda \leq C$ for all $v_c \in V, \lambda_c \in \Lambda$ with $||v_c - v_*||_V + ||\lambda_c - \lambda_*||_\Lambda \leq \epsilon_2$. The constant C depends only on ϵ_2.

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

$$||\pi_s||_{L^2} + ||\pi_a||_{L^2} \leq C\big[\big(||v_c - v_*||_V + ||\lambda_c - \lambda_*||_\Lambda^2\big)^2 + ||v_c - v_*||_V + ||\lambda_c - \lambda_*||_\Lambda\big] .$$

Since $||v_c - v_*||_V + ||\lambda_c - \lambda_*||_\Lambda \leq \epsilon_2$ this gives the estimate

$$||\pi_s||_{L^2} + ||\pi_a||_{L^2} \leq C\big(||v_c - v_*||_V + ||\lambda_c - \lambda_*||_\Lambda\big) . \tag{6.1}$$

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

$$||v_+ - v_*||_V \leq C\big(||v_c - v_*||_V + ||\lambda_c - \lambda_*||_\Lambda\big)$$

and

$$||\lambda_+ - \lambda_*||_\Lambda \leq C\big(||v_c - v_*||_V + ||\lambda_c - \lambda_*||_\Lambda\big) .$$

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

$$||\pi_s||_{L^2} + ||\pi_a||_{L^2} \leq C\big(||v_c - v_*||_V + ||\lambda_c - \lambda_*||_\Lambda\big)^2 . \tag{6.2}$$

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

$$||v_+ - v_*||_V \leq C\big(||v_c - v_*||_V + ||\lambda_c - \lambda_*||_\Lambda\big)^2 \tag{6.3}$$

and

$$||\lambda_+ - \lambda_*||_\Lambda \leq C\big(||v_c - v_*||_V + ||\lambda_c - \lambda_*||_\Lambda\big)^2 . \tag{6.4}$$

(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. □

The previous theorem shows a quadratic reduction of the error for a single iteration. Standard induction arguments can now be used to show that all iterates (v_k, λ_k) satisfy $||v_k - v_*||_{V_q} + ||\lambda_k - \lambda_*||_{\Lambda_q} \leq \epsilon$ which implies the local quadratic convergence of the SQP method.

7 Appendix

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

$$\begin{aligned} \frac{\partial}{\partial t}u + \frac{\ell}{2}\frac{\partial}{\partial t}\varphi &= \kappa\Delta u + \beta_1 u + f_1 \\ \tau\frac{\partial}{\partial t}\varphi &= \xi^2\Delta\varphi + \beta_2\varphi + 2u + f_2 \end{aligned} \qquad \text{in } \Omega \times (0, T], \tag{7.1}$$

with boundary conditions

$$\frac{\partial}{\partial n}u = 0, \quad \frac{\partial}{\partial n}\varphi = 0, \quad \text{on } \partial\Omega \times (0, T), \tag{7.2}$$

and initial conditions

$$u = u_0, \quad \varphi = \varphi_0 \qquad \text{in } \Omega. \tag{7.3}$$

Theorem 7.1 *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* $(u, \varphi) \in W_q^{2,1}(Q) \times W_q^{2,1}(Q)$ *of the system* (7.1), (7.2), (7.3). *The solution obeys*

$$\|u\|_{W_q^{2,1}(Q)} + \|\varphi\|_{W_q^{2,1}(Q)} \leq C(\|u_0\|_{W^2_\infty(\Omega)} + \|\varphi_0\|_{W^2_\infty(\Omega)} + \|f_1\|_{L^q(Q)} + \|f_2\|_{L^q(Q)}),$$

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

Proof: Again, C denotes a generic constant independent of u and φ.

(i) An a–priori estimate:

Suppose that $u, \varphi \in \{v \mid v \in L^2(0,T; H^1(\Omega)), \frac{\partial}{\partial t} v \in L^2(0,T;(H^1(\Omega))')\}$ 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

$$\begin{aligned}
&\frac{1}{2}\frac{\partial}{\partial t}\int_\Omega (u(t) + \frac{\ell}{2}\varphi(t))^2\,dx + \int_\Omega \kappa|\nabla u(t)|^2\,dx + \frac{\ell\kappa}{2}\int_\Omega \nabla u(t)\nabla\varphi(t)\,dx\\
=\;& \int_\Omega \beta_1(t)u(t)\big(u(t) + \frac{\ell}{2}\varphi(t)\big)\,dx + \int_\Omega (u(t) + \frac{\ell}{2}\varphi(t)) f_1(t)\,dx\\
\leq\;& \|\beta_1(t)\|_{L^3(\Omega)}\|u(t)\|_{L^2(\Omega)}\|u(t) + \frac{\ell}{2}\varphi(t)\|_{L^6(\Omega)} + \frac{1}{2}\|u(t) + \frac{\ell}{2}\varphi(t)\|^2_{L^2(\Omega)} + \frac{1}{2}\|f_1(t)\|^2_{L^2(\Omega)}\\
\leq\;& C\|\beta_1(t)\|_{L^3(\Omega)}\|u(t)\|_{L^2(\Omega)}\|u(t) + \frac{\ell}{2}\varphi(t)\|_{H^1(\Omega)} + \frac{1}{2}\|u(t) + \frac{\ell}{2}\varphi(t)\|^2_{L^2(\Omega)} + \frac{1}{2}\|f_1(t)\|^2_{L^2(\Omega)}\\
\leq\;& \frac{C^2}{2\kappa}\|\beta_1(t)\|^2_{L^3(\Omega)}\|u(t)\|^2_{L^2(\Omega)} + \frac{\kappa}{2}\|\nabla(u(t) + \frac{\ell}{2}\varphi(t))\|^2_{L^2(\Omega)} + \frac{1+\kappa}{2}\|u(t) + \frac{\ell}{2}\varphi(t)\|^2_{L^2(\Omega)} + \frac{1}{2}\|f_1(t)\|^2_{L^2(\Omega)}\\
\leq\;& \frac{\kappa}{2}\|\nabla(u(t) + \frac{\ell}{2}\varphi(t))\|^2_{L^2(\Omega)}\\
&+C\big(1 + \|\beta_1(t)\|^2_{L^3(\Omega)}\big)\big(\|u(t) + \frac{\ell}{2}\varphi(t)\|^2_{L^2(\Omega)} + \|u(t)\|^2_{L^2(\Omega)}\big) + C\|f_1(t)\|^2_{L^2(\Omega)} \qquad (7.4)
\end{aligned}$$

and

$$\begin{aligned}
&\frac{\tau}{2}\frac{\partial}{\partial t}\int_\Omega \varphi(t)^2\,dx + \int_\Omega \xi^2|\nabla\varphi(t)|^2\,dx\\
=\;& \int_\Omega \beta_2(t)\varphi(t)^2\,dx + 2\int_\Omega u(t)\varphi(t)\,dx + \int_\Omega f_2(t)\varphi(t)\,dx\\
\leq\;& \|\beta_2(t)\|_{L^3(\Omega)}\|\varphi(t)\|_{L^2(\Omega)}\|\varphi(t)\|_{L^6(\Omega)} + 2\int_\Omega \big(u(t) + \frac{\ell}{2}\varphi(t)\big)\varphi(t)\,dx + C\|\varphi(t)\|^2_{L^2(\Omega)} + C\|f_2(t)\|^2_{L^2(\Omega)}\\
\leq\;& \frac{\xi^2}{2}\|\nabla\varphi(t)\|^2_{L^2(\Omega)} + C\big(1 + \|\beta_2(t)\|^2_{L^3(\Omega)}\big)\big(\|u(t) + \frac{\ell}{2}\varphi(t)\|^2_{L^2(\Omega)} + \|\varphi(t)\|^2_{L^2(\Omega)}\big) + C\|f_2(t)\|^2_{L^2(\Omega)}. \qquad (7.5)
\end{aligned}$$

Multiplying (7.5) by A and adding the result to (7.4) yields

$$\begin{aligned}
&\frac{1}{2}\frac{\partial}{\partial t}\int_\Omega (u(t) + \frac{\ell}{2}\varphi(t))^2 + \tau A\varphi(t)^2\,dx + \frac{1}{2}\int_\Omega \kappa|\nabla u(t)|^2 + (A\frac{\xi^2}{2} - \frac{\kappa\ell^2}{4})|\nabla\varphi(t)|^2\,dx\\
\leq\;& C\big(1 + \|\beta_1(t)\|^2_{L^3(\Omega)} + \|\beta_2(t)\|^2_{L^3(\Omega)}\big)\big(\|u(t) + \frac{\ell}{2}\varphi(t)\|^2_{L^2(\Omega)} + \|u(t)\|^2_{L^2(\Omega)} + \|\varphi(t)\|^2_{L^2(\Omega)}\big)\\
&+C\|f_1(t)\|^2_{L^2(\Omega)} + C\|f_2(t)\|^2_{L^2(\Omega)}. \qquad (7.6)
\end{aligned}$$

If we choose $A \geq 2(\kappa\ell^2/4 + 1)/\xi^2$, then (7.6) implies

$$\begin{aligned}
&\frac{1}{2}\frac{\partial}{\partial t}\int_\Omega (u(t) + \frac{\ell}{2}\varphi(t))^2 + \tau A\varphi(t)^2\,dx + \frac{1}{2}\int_\Omega \kappa|\nabla u(t)|^2 + |\nabla\varphi(t)|^2\,dx\\
\leq\;& C\big(1 + \|\beta_1(t)\|^2_{L^3(\Omega)} + \|\beta_2(t)\|^2_{L^3(\Omega)}\big)\big(\|u(t) + \frac{\ell}{2}\varphi(t)\|^2_{L^2(\Omega)} + \|u(t)\|^2_{L^2(\Omega)} + \|\varphi(t)\|^2_{L^2(\Omega)}\big)\\
&+C\|f_1(t)\|^2_{L^2(\Omega)} + C\|f_2(t)\|^2_{L^2(\Omega)}. \qquad (7.7)
\end{aligned}$$

Integration over t and by parts gives

$$\begin{aligned} &\frac{1}{2}\|u(t)+\frac{\ell}{2}\varphi(t)\|^2_{L^2(\Omega)}+\frac{\tau A}{2}\|\varphi(t)\|^2_{L^2(\Omega)}+\frac{1}{2}\int_0^t \kappa\|\nabla u(s)\|^2_{L^2(\Omega)}+\|\nabla\varphi(s)\|^2_{L^2(\Omega)}\,ds \\ \leq\;& \frac{1}{2}\Big(\|u(0)+\frac{\ell}{2}\varphi(0)\|^2_{L^2(\Omega)}+\tau A\|\varphi(0)^2\|^2_{L^2(\Omega)}\Big)+C\|f_1\|^2_{L^2(Q)}+C\|f_2\|^2_{L^2(Q)} \\ &+C\int_0^t \big(1+\|\beta_1(s)\|^2_{L^3(\Omega)}+\|\beta_2(s)\|^2_{L^3(\Omega)}\big)\big(\|u(s)\|^2_{L^2(\Omega)}+\|\varphi(s)\|^2_{L^2(\Omega)}\big)\,ds. \end{aligned} \tag{7.8}$$

If $A \geq \big(1/2+\ell^2/4\big)/\tau$, then

$$(a+\frac{\ell}{2}b)^2+\tau Ab^2 = a^2+\ell ab+\frac{\ell^2}{4}b^2+\tau Ab^2 \geq a^2-\frac{1}{2}a^2-\frac{\ell^2}{2}b^2+\frac{\ell^2}{4}b^2+\tau Ab^2 \geq \frac{1}{2}a^2+\frac{1}{2}b^2$$

for all a, b. Thus, for $A \geq (1/2+\ell^2/4)/\tau$ equation (7.8) gives the estimate

$$\begin{aligned} &\|u(t)\|^2_{L^2(\Omega)}+\|\varphi(t)\|^2_{L^2(\Omega)}+\int_0^t \kappa\|\nabla u(s)\|^2_{L^2(\Omega)}+\|\nabla\varphi(s)\|^2_{L^2(\Omega)}\,ds \\ \leq\;& C\Big(\|u(0)\|^2_{L^2(\Omega)}+\|\varphi(0)\|^2_{L^2(\Omega)}+\|f_1\|^2_{L^2(Q)}+\|f_2\|^2_{L^2(Q)}\Big) \\ &+C\int_0^t \big(1+\|\beta_1(s)\|^2_{L^3(\Omega)}+\|\beta_2(s)\|^2_{L^3(\Omega)}\big)\big(\|u(s)\|^2_{L^2(\Omega)}+\|\varphi(s)\|^2_{L^2(\Omega)}\big)\,ds. \end{aligned} \tag{7.9}$$

Now we can use the Gronwall–Bellman inequality[1] to derive the inequality

$$\|u(t)\|^2_{L^2(\Omega)}+\|\varphi(t)\|^2_{L^2(\Omega)} \leq C\Big(\|u(0)\|^2_{L^2(\Omega)}+\|\varphi(0)\|^2_{L^2(\Omega)}+\|f_1\|^2_{L^2(Q)}+\|f_2\|^2_{L^2(Q)}\Big).$$

Inserting this equation into (7.9) yields the a–priori estimate

$$\begin{aligned} &\|u(t)\|^2_{L^2(\Omega)}+\|\varphi(t)\|^2_{L^2(\Omega)}+\int_0^t \kappa\|\nabla u(s)\|^2_{L^2(\Omega)}+\|\nabla\varphi(s)\|^2_{L^2(\Omega)}\,ds \\ \leq\;& C\Big(\|u(0)\|^2_{L^2(\Omega)}+\|\varphi(0)\|^2_{L^2(\Omega)}+\|f_1\|^2_{L^2(Q)}+\|f_2\|^2_{L^2(Q)}\Big). \end{aligned} \tag{7.10}$$

(ii) Uniqueness of the solution: If (u_i, φ_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

$$(u,\varphi) \in \Big\{v \mid v \in L^2(0,T;H^1(\Omega)),\ \frac{\partial}{\partial t}v \in L^2(0,T;(H^1(\Omega))^*)\Big\}^2 \subset \Big(C(0,T;L^2(\Omega))\Big)^2$$

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,

[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:

$$\psi(t) \leq a\ \exp\Big(\int_0^t m(s)ds\Big), \quad t \in [0,T].$$

p. 341] we find that the solution φ is in $W_2^{2,1}(Q)$. 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 $u \in W_2^{2,1}(Q)$.

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

$$||\beta_2\varphi||_{L^{30/13}(Q)} \le ||\beta_2||_{L^3(Q)}||\varphi||_{L^{10}(Q)} .$$

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,

$$||\beta_2\varphi||_{L^{5/2}(Q)} \le ||\beta_2||_{L^3(Q)}||\varphi||_{L^{15}(Q)}$$

and $u, \varphi \in W_p^{2,1}(Q)$ with $p = \min(q, 5/2)$. If $5/2 < q$ we can use the imbedding $W^{2,1}_{5/2} \subset L^\mu(Q)$ for all $\mu \in [1, \infty)$ and repeat the previous steps to show that $u, \varphi \in W_q^{2,1}(Q)$. □

References

[1] W. ALT, *The Lagrange–Newton method for infinite dimensional optimization problems*, Numer. Funct. Anal. Optim., 11 (1990), pp. 201–224.

[2] W. ALT AND K. MALANOWSKI, *The Lagrange–Newton method for nonlinear optimal control problems*, Computational Optimization and Applications, 2 (1993), pp. 77–100.

[3] ——, *The Lagrange–Newton method for state-constrained optimal control problems*, Computational Optimization and Applications, 4 (1995), pp. 217–239.

[4] W. ALT, R. SONTAG, AND F. TRÖLTZSCH, *The SQP-method and optimal control of weakly singular Hammerstein integral equations*, Appl. Math. Optimization, 33 (1994), pp. 227–252.

[5] P. T. BOGGS, *Sequential quadratic programming*, in Acta Numerica 1995, A. Iserles, ed., Cambridge University Press, Cambridge, London, New York, 1995, pp. 1–51.

[6] H. BRÉZIS, *Operateurs Maximaux Monotones et Semi–Groups de Contractions Dans les Espaces de Hilbert*, Notas de Matemática, Vol 50, North Holland, Amsterdam, London, New-York, 1973.

[7] G. CAGINALP, *An analysis of a phase field model of a free boundary*, Archive for Rational Mechanics and Analysis, 92 (1986), pp. 205–245.

[8] G. CAGINALP AND J. T. LIN, *A numerical analysis of an anisotropic phase field model*, IMA J. Appl. Math., 39 (1987), pp. 51–66.

[9] G. CAGINALP AND E. SOCOLOVSKI, *Phase field computations of single–needle crystals, crystal growth, and motion by mean curvature*, SIAM J. Sci. Comput., 15 (1994), pp. 106–126.

[10] Z. CHEN, *Das Phasenfeldmodell für das Problem der Phasenübergänge: Analysis, Numerik und Steuerung*, PhD thesis, Institut für Mathematik, Universität Augsburg, Germany, 1991. (in german).

[11] Z. CHEN AND K.-H. HOFFMANN, *Numerical solutions of the optimal control problem governed by a phase field model*, in Optimal Control of Partial Differential Equations, W. Desch, F. Kappel, and K. Kunisch, eds., Basel, Boston, Berlin, 1991, Birkhäuser Verlag.

[12] ——, *Error estimate of a finite element scheme for a phase field model*, IMA Journal on Numerical Analysis, (1993).

[13] J. B. COLLINS AND H. LEVINE, *Diffusive interface model of diffusion limited crystal growth*, Physica Review B, 31 (1985), pp. 6119–6122.

[14] R. DAUTRAY AND J. L. LIONS, *Mathematical Analysis and Numerical Methods for Science and Technology, Volume 5: Evolution Problems I*, Springer Verlag, Berlin, Heidelberg, New York, 1992.

[15] A. L. DONTCHEV, W. W. HAGER, A. B. POORE, AND B. YANG, *Optimality, stability and convergence in nonlinear control*, Applied Mathematics & Optimization, 31 (1995), pp. 297–326.

[16] J. C. DUNN AND T. TIAN, *Variants of the Kuhn–Tucker sufficient conditions in cones of nonnegaive functions*, SIAM J. Control Optim., 30 (1992), pp. 1361–1384.

[17] C. M. ELLIOTT AND Z. SONGMU, *Global existence and stability of solutions to the phase field equations*, in Free Boundary Problems, K.-H. Hoffmann and J. Sprekels, eds., Birkh¨auser Verlag, Basel, 1990, pp. 46–58.

[18] G. FIX, *Numerical simulation of free boundary problems using phase field models*, in The Mathematics of Finite Elements and Applications IV, MAFELAP 1981, J. R. Whiteman, ed., Academic Press, London, New York, ..., 1982, pp. 265–279.

[19] G. J. FIX AND J. T. LIN, *Numerical simulations of a nonlinear phase transition I: The isotrophic case*, Nonlinear Analysis. Theory, Methods & Applications, 12 (1988), pp. 811–823.

[20] H. GOLDBERG AND F. TRÖLTZSCH, *On a Lagrange–Newton method for a nonlinear parabolic boundary control problem*, Optimization Methods and Software, (1998).

[21] ———, *On a SQP-multigrid technique for nonlinear parabolic boundary control problems*, in Optimal Control: Theory, Algorithms, and Applications, W. W. Hager and P. M. Pardalos, eds., Kluwer Academic Publishers B.V., 1998, pp. 154–177.

[22] M. HEINKENSCHLOSS, *The numerical solution of a control problem governed by a phase field model*, Optimization Methods and Software, 7 (1997), pp. 211–263.

[23] M. HEINKENSCHLOSS AND E. W. SACHS, *Numerical solution of a constrained control problem for a phase field model*, in Control and Estimation of Distributed Parameter Systems: Nonlinear Phenomena, W. Desch, F. Kappel, and K. Kunisch, eds., Int. Series of Numer. Math. Vol. 118, Basel, Boston, Berlin, 1994, Birkh¨auser Verlag, pp. 171–187.

[24] K.-H. HOFFMANN AND L. JIANG, *Optimal control problem of a phase field model for solidification*, Numer. Funct. Anal., 13 (1992), pp. 11–27.

[25] W. HORN, *Mathematical aspects of the Penrose–Fife phase-field model*, Control and Cybernetics, 23 (1994), pp. 677–689.

[26] C. T. KELLEY AND E. W. SACHS, *Solution of optimal control problems by a pointwise projected Newton method*, SIAM J. on Control and Optimization, 33 (1995), pp. 1731–1757.

[27] C. T. KELLEY AND S. J. WRIGHT, *Sequential quadratic programming for certain parameter identification problems*, Math. Programming, 51 (1991), pp. 281–305.

[28] N. KENMOCHI, *Systems of nonlinear PDEs arising from dynamical phase transitions*, in Phase Transitions and Hysterisis, Montecatini Terme, 1993. Lecture Notes in Mathematics, Vol.1584, M. Brokate, N. Kenmochi, I. M¨uller, J. F. Rodriguez, and C. Verdi, eds., Berlin, Heidelberg, New-York, 1994, Springer Verlag, pp. 39–86.

[29] R. KOBAYASHI, *Modelling and numerical simulations of dentric crystal growth*, Physica D, 63 (1993), pp. 410–423.

[30] K. KUNISCH AND E. SACHS, *Reduced SQP methods for parameter identification problems*, SIAM J. Numer. Anal., 29 (1992), pp. 1793–1820.

[31] F.-S. KUPFER, *An infinite dimensional convergence theory for reduced SQP methods in Hilbert space*, SIAM J. Optimization, 6 (1996), pp. 126–163.

[32] F.-S. KUPFER AND E. W. SACHS, *Numerical solution of a nonlinear parabolic control problem by a reduced SQP method*, Computational Optimization and Applications, 1 (1992), pp. 113–135.

[33] O. A. LADYZEHNSKAYA, V. A. SOLONNIKOV, AND N. N. URAL'TSEVA, *Linear and Quasilinear Equations of Parabolic Type*, American Mathematical Society, Rhode Island, 1968.

[34] J. T. LIN, *The numerical analysis of a phase field model in moving boundary problems*, SIAM J. Numer. Anal., 25 (1988), pp. 1015–1031.

[35] J. L. LIONS, *Control of Distributed Singular Systems*, Dunod, Gauthier–Villars, Paris, 1985.

[36] H. MAURER, *First and second order sufficient optimality conditions in mathematical programming and optimal control*, Mathematical Programming Study, 14 (1981), pp. 163–177.

[37] H. MAURER AND J. ZOWE, *First and second–order necessary and sufficient optimality conditions for infinite–dimensional programming problems*, Math. Programming, 16 (1979), pp. 98–110.

[38] O. PENROSE AND P. C. FIFE, *Thermodynamically consistent models of phase–field type for kinetics of phase transitions*, Physica D, 43 (1990), pp. 44–62.

[39] ——, *On the relation between the standard phase–field model and a "thermodynamically consistent" phase–field model*, Physica D, 69 (1993), pp. 107–113.

[40] S. M. ROBINSON, *Strongly regular generalized equations*, Math. of Operations Research, 5 (1980), pp. 43–62.

[41] J. SPREKELS AND M. BROKATE, *Hysteresis and Phase Transitions*, Applied Mathematical Sciences, Vol 121, Springer Verlag, Heidelberg, New York, Berlin, 1996.

[42] J. SPREKELS AND S. ZHENG, *Optimal control problems for a thermodynamically consistent model of phase–field type for phase transitions*, Advances in Mathematical Sciences and Applications, 1 (1992), pp. 113–125.

[43] ——, *Global smooth solutions to a thermodynamically consistent model of phase–field type in higher space dimensions*, J. Math. Anal. Appl., (to appear).

[44] F. TRÖLTZSCH, *An SQP method for optimal control of a nonlinear heat equation*, Control and Cybernetics, 23 (1994), pp. 268–288.

[45] ——, *On the Lagrange–Newton–SQP method for the optimal control of semilinear parabolic equations*, Tech. Rep., Fakultät für Mathematik, 1998. `http://www.tu-chemnitz.de/~troeltz/index.html`.

[46] E. ZEIDLER, *Nonlinear Functional Analysis and its Applications II/A: Linear Monotone Operators*, Springer Verlag, Berlin, Heidelberg, New-York, 1990.