J. Nonsmooth Anal. Optim. 4 (2023), 10834, doi: 10.46298/jnsao-2023-10834
Submitted: 2023-01-19, accepted: 2023-07-26

OPTIMAL CONTROL OF A VISCOUS TWO-FIELD DAMAGE MODEL WITH FATIGUE

Livia Betz*

Abstract Motivated by fatigue damage models, this paper addresses optimal control problems governed by a non-smooth system featuring two non-differentiable mappings. This consists of a coupling between a doubly non-smooth history-dependent evolution and an elliptic PDE. After proving the directional differentiability of the associated solution mapping, an optimality system which is stronger than the one obtained by classical smoothening procedures is derived. If one of the non-differentiable mappings becomes smooth, the optimality conditions are of strong stationary type, i.e., equivalent to the primal necessary optimality condition.

Keywords: Damage models with fatigue, non-smooth optimization, evolutionary VIs, optimal control of PDEs, history-dependence, strong stationarity.

MSC (2020): 34G25, 34K35, 49J20, 49J27, 74R99.

1 INTRODUCTION

Fatigue is considered to be the main cause of mechanical failure [28, 34]. It describes the weakening of a material due to repeated applied loads (fluctuating stresses, strains, forces, environmental factors, temperature, etc.), which individually would be too small to cause its malfunction [1, 34]. Whether in association with environmental damage (corrosion fatigue) or elevated temperatures (creep fatigue), fatigue failure is often an unexpected phenomenon. Unfortunately, in real situations, it is very difficult to identify the fatigue degradation state of a material, which sometimes might result in devastating events. Therefore, it is extremely important to find methods which allow us to describe and control the behaviour of materials exposed to fatigue. While there are very few papers [1] (damage in elastic materials) and [11] (cohesive fracture), concerned with a rigorous mathematical examination of models describing fatigue damage, the literature regarding the optimal control of fatigue models is practically nonexistent. All the existing results which include the terminology "optimal control" in the context of fatigue damage do not address theoretical aspects nor involve mathematical tools such as optimal control theory in Banach spaces as in the present work, but focus on design of controllers and simulations instead, see e.g. [18, 27] and the references therein.

In this paper we investigate the optimal control of the following viscous two-field gradient damage problem with fatigue:

$$\left.\begin{aligned} \varphi(t) &\in \operatorname*{arg\,min}_{\varphi \in H^1(\Omega)} \mathcal{E}(t, \varphi, q(t)), \\ -\partial_q \mathcal{E}(t, \varphi(t), q(t)) &\in \partial_{\dot{q}} \mathcal{R}_\epsilon(\mathcal{H}(q)(t), \dot{q}(t)) \text{ in } L^2(\Omega), \quad q(0) = 0, \end{aligned}\right\} \tag{1.1}$$

This work was supported by the DFG grant BE 7178/3-1 for the project "Optimal Control of Viscous Fatigue Damage Models for Brittle Materials: Optimality Systems".

*University of Würzburg, Faculty of Mathematics, livia.betz@uni-wuerzburg.de

a.e. in $(0, T)$. To be more precise, we prove an optimality system that is far stronger than the one obtained by classical smoothening techniques.

The main novelty concerning (1.1) arises from the *highly non-smooth structure*, which is due to the non-differentiability of the dissipation $\mathcal{R}_\epsilon$ in the evolution inclusion, in combination with *an additional non-smooth* fatigue degradation mapping which shall be introduced below. This excludes the application of standard adjoint techniques for the derivation of first-order necessary conditions in form of optimality systems. Not only does the evolution in (1.1) have a highly non-smooth character, but, as we will next see, it is also *history-dependent*. The fact that the differential inclusion is coupled with a minimization problem (which can be reduced to an elliptic PDE) gives rise to additional challenges [5].

The problem describes the evolution of damage under the influence of a time-dependent load $\ell : [0, T] \to H^1(\Omega)^*$ (control) acting on a body occupying the bounded Lipschitz domain $\Omega \subset \mathbb{R}^N$, $N \in \{2, 3\}$. The induced 'local' and 'nonlocal' damage are expressed in terms of the functions $q : [0, T] \to L^2(\Omega)$ and $\varphi : [0, T] \to H^1(\Omega)$, respectively (states).

In (1.1), the stored energy $\mathcal{E} : [0, T] \times H^1(\Omega) \times L^2(\Omega) \to \mathbb{R}$ is given by

$$\mathcal{E}(t, \varphi, q) := \frac{\alpha}{2}\|\nabla\varphi\|^2_{L^2(\Omega)} + \frac{\beta}{2}\|\varphi - q\|^2_{L^2(\Omega)} - \langle \ell(t), \varphi\rangle_{H^1(\Omega)}, \tag{1.2}$$

where $\alpha > 0$ is the gradient regularization and $\beta > 0$ denotes the *penalization parameter*. Thus, the two damage variables are connected through the penalty term β in the stored energy, so that our model becomes a penalized version of the viscous fatigue damage model addressed in [1] (two-dimensional case); note that, for simplicity reasons, we do not take a displacement variable into account. The type of penalization used in (1.2) has already been proven to be successful in the context of classical damage models (without fatigue). Firstly, it approximates the classical single-field damage model, in the sense that, when $\beta \to \infty$, the penalized damage model coincides with the model addressed in [16, 21], cf. [24]. Secondly, the penalization we use is frequently employed in computational mechanics due to the numerical benefits offered by the additional damage variable (see e.g. [13] and the references therein). For more details, we also refer to [23, Sec. 2.1-2.2].

The differential inclusion appearing in (1.1) describes the evolution of the damage variable q under *fatigue* effects. Therein, $\mathcal{H}$ is a so-called *history operator* that models how the damage experienced by the material affects its fatigue level. Thus, as opposed to other well-known damage models, cf. e.g. [15, 16, 21], the dissipation $\mathcal{R}_\epsilon$ in (1.1) is affected by the *history of the evolution*, $\mathcal{H}(q)$. The parameter $\epsilon > 0$ stands for the viscosity parameter, while the symbol $\partial_{\dot q}$ denotes the convex subdifferential of the functional $\mathcal{R}_\epsilon$ in its second argument. Thus, the *non-smooth differential inclusion* is to be understood as follows:

$$(-\partial_q\mathcal{E}(t, \varphi(t), q(t)), \eta - \dot q(t))_{L^2(\Omega)} \le \mathcal{R}_\epsilon(\mathcal{H}(q)(t), \eta) - \mathcal{R}_\epsilon(\mathcal{H}(q)(t), \dot q(t)) \quad \forall\, \eta \in L^2(\Omega).$$

The viscous dissipation $\mathcal{R}_\epsilon : L^2(\Omega) \times L^2(\Omega) \to (-\infty, \infty]$ is defined as

$$\mathcal{R}_\epsilon(\omega, \eta) := \begin{cases} \displaystyle\int_\Omega f(\omega)\,\eta\,dx + \frac{\epsilon}{2}\|\eta\|^2_{L^2(\Omega)}, & \text{if } \eta \ge 0 \text{ a.e. in } \Omega, \\ \infty & \text{otherwise,} \end{cases} \tag{1.3}$$

and features *a second non-smooth component*, namely the fatigue degradation mapping f. This describes in which measure the fatigue affects the fracture toughness of the material. This mapping is non-increasing in applications, since the higher the cumulated damage $\mathcal{H}(q)$, the lower the fracture toughness $f(\mathcal{H}(q))$. Whereas usually the toughness of the material is described by a fixed (nonnegative) constant [15, 16], in the present model it changes at each point in time and space, depending on $\mathcal{H}(q)$. To be more precise, the value of the fracture toughness of the body at (t, x) is given by $f(\mathcal{H}(q))(t, x)$, cf. (1.3). Hence, the model (1.1) takes into account the following crucial aspect: the occurrence of damage is favoured in regions where fatigue accumulates.

We underline that the dissipation $\mathcal{R}_\epsilon$ accounts for the non-smooth nature of the evolution in the first place: even if f is replaced by a (nonnegative) constant, the evolution in (1.1) still describes a non-smooth process. The optimal control thereof is far away from being standard and has been recently addressed in [5, Sec. 4], where strong stationarity for the damage model (1.1) without fatigue is proven. By contrast, in applications which take fatigue into consideration, $f : \mathbb{R}^+ \to \mathbb{R}^+$ is constant until its kink point is achieved, after which it monotonically decreases [2, Sec. 2.6.2]. Thus, it is the fatigue degradation mapping f which accounts for the *highly non-smooth character* of our problem.

Deriving necessary optimality conditions is a challenging issue even in finite dimensions, where a special attention is given to MPCCs (mathematical programs with complementarity constraints). In [30] a detailed overview of various optimality conditions of different strength was introduced, see also [20] for the infinite-dimensional case. The most rigorous stationarity concept is strong stationarity. Roughly speaking, the strong stationarity conditions involve an optimality system, which is equivalent to the purely primal conditions saying that the directional derivative of the reduced objective in feasible directions is nonnegative (which is referred to as B stationarity).

While there are plenty of contributions in the field of optimal control of smooth problems, see e.g. [37] and the references therein, fewer papers are dealing with non-smooth problems. Most of these papers resort to regularization or relaxation techniques to smoothen the problem, see e.g. [3, 17, 19] and the references therein. The optimality systems derived in this way are of intermediate strength and are not expected to be of strong stationary type, since one always loses information when passing to the limit in the regularization scheme. Thus, proving strong stationarity for optimal control of non-smooth problems requires direct approaches, which employ the limited differentiability properties of the control-to-state map. In this context, there are even less contributions. We refer to the pioneering work [25] (strong stationarity for optimal control of elliptic VIs of obstacle type), which was followed by other papers addressing strong stationarity of various types of VIs [7, 8, 12, 26, 38, 39]. Regarding strong stationarity for optimal control of non-smooth PDEs, the literature is rather scarce and the only papers known to the author addressing this issue so far are [5, 6, 9, 10, 22].

Let us point out the main contributions of the present work. This paper aims at deriving optimality conditions which - regarding their strength - lie between the conditions derived by classical regularization techniques and the strong stationary ones. Starting from an optimality system obtained via smoothening, we resort to direct methods from previous works [5, 22], in order to improve our initial optimality conditions as far as we can. This approach has been employed in [5, 22] to arrive at a strong stationary optimality system. However, in the literature concerned with the derivation of optimality conditions which are weaker than strong stationarity, the analysis ends with the passage to the limit in the regularized system. We underline that in the present work we exploit the B-stationarity condition to improve the limit optimality system, though, in the end, optimality conditions of strong stationary type are not established. Indeed, in contrast to [5, 22], our state system features *two* non-differentiable mappings instead of one, so that the methods from the aforementioned works are of limited applicability: Optimality conditions equivalent to the B-stationary ones are not expected in our complex doubly non-smooth setting. If the fatigue degradation mapping is smooth, strong stationarity conditions are available. While control problems featuring non-smooth terms both in the objective and the state equation have been addressed in [4, 6, 36], we point out that, to the best of our knowledge, optimization problems featuring two non-differentiable functions in the state equation have not been tackled so far, not even in the context of classical smoothening methods.

The paper is structured as follows. After an introduction of the notation, section 2 focuses on the analysis of our fatigue damage model (1.1). Here we address the existence and uniqueness of solutions, by proving that (1.1) is in fact equivalent to a PDE system. This consists of an elliptic PDE and a highly non-smooth differential ODE. The latter one is of particular interest. It features two non-differentiable functions, namely max and the fatigue degradation function f; the latter appears in the argument of the initial non-smoothness, cf. (2.2a). The properties of the control-to-state operator associated to (1.1) are investigated. In particular, we are concerned with the *directional differentiability* of the solution mapping of the non-smooth state system.

In section 3 we present the optimal control problem and investigate the existence of optimal minimizers. Then, in subsection 3.1 we derive our first optimality conditions, by resorting to a classical smoothening method. These conditions are of intermediate strength. If the non-smoothness is inactive, they coincide with the classical KKT system. However, our first optimality system does not contain any information in those points (t, x) where the non-differentiable mappings max and f attain their kink points. This is namely the focus of section 3.2, where the main result is proven in Theorem 3.15. Here, the initial optimality system is improved by employing the "surjectivity" trick from [5, 22]. The new and final optimality conditions (3.17) are comparatively strong (but not strong stationary). They contain information in terms of sign conditions on sets where the non-smoothness is active; these are not expected to be obtained if one just smoothens the problem, cf. e.g. [6, Remark 3.9]. Moreover, if the fatigue degradation function f is smooth, then (3.17) is of strong stationary type (Corollary 3.16). For completeness, the expected (not proven) strong stationarity system associated to the doubly non-smooth state system is presented in Section 3.3. Here we include a thorough explanation as to why the methods from [5, 22] fail (Remark 3.22). Finally, we include in Appendix A the proof of Lemma 3.7, for convenience of the reader.

NOTATION

Throughout the paper, $T > 0$ is a fixed final time. If X and Y are linear normed spaces, then the space of linear and bounded operators from X to Y is denoted by $\mathcal{L}(X, Y)$, and $X \overset{d}{\hookrightarrow} Y$ means that X is densely embedded in Y. The dual space of X will be denoted by X^*. For the dual pairing between X and X^* we write $\langle ., . \rangle_X$. The closed ball in X around $x \in X$ with radius $\alpha > 0$ is denoted by $B_X(x, \alpha)$. If X is a Hilbert space, we write $(\cdot, \cdot)_X$ for the associated scalar product. The following abbreviations will be used throughout the paper:

$$H_0^1(0, T; X) := \{z \in H^1(0, T; X) : z(0) = 0\},$$
$$H_T^1(0, T; X) := \{z \in H^1(0, T; X) : z(T) = 0\},$$

where X is a Banach space. The adjoint operator of a linear and continuous mapping A is denoted by $A^\star$. By χ_M we denote the characteristic function associated to the set M. Derivatives w.r.t. time (weak derivatives of vector-valued functions) are frequently denoted by a dot. The symbol ∂ stands for the convex subdifferential, see e.g. [29]. With a little abuse of notation, the Nemytskii-operators associated with the mappings considered in this paper will be denoted by the same symbol, even when considered with different domains and ranges. The mapping $\max\{\cdot, 0\}$ is abbreviated by $\max(\cdot)$. With a little abuse of notation, we use in the following the Laplace symbol for the operator $\Delta : H^1(\Omega) \to H^1(\Omega)^*$ defined by

$$\langle \Delta\eta, \psi \rangle_{H^1(\Omega)} := -\int_\Omega \nabla\eta \nabla\psi \, dx \quad \forall\, \psi \in H^1(\Omega).$$

2 PROPERTIES OF THE CONTROL-TO-STATE MAP

This section is concerned with the investigation of the solvability and differentiability properties of the state system (1.1).

Assumption 2.1. For the mappings associated with fatigue in (1.1) we require the following:

1. The *history operator* $\mathcal{H} : L^2(0, T; L^2(\Omega)) \to L^2(0, T; L^2(\Omega))$ satisfies

$$\|\mathcal{H}(q_1)(t) - \mathcal{H}(q_2)(t)\|_{L^2(\Omega)} \le L_{\mathcal{H}} \int_0^t \|q_1(s) - q_2(s)\|_{L^2(\Omega)} \, ds \quad \text{a.e. in } (0, T),$$

for all $q_1, q_2 \in L^2(0,T;L^2(\Omega))$, where $L_{\mathcal{H}} > 0$ is a positive constant.
Moreover, $\mathcal{H} : L^2(0,T;L^2(\Omega)) \to L^2(0,T;L^2(\Omega))$ is supposed to be Gâteaux-differentiable with continuous derivative on $H^1(0,T;L^2(\Omega))$.

2. The non-linear function $f : \mathbb{R} \to \mathbb{R}$ is assumed to be Lipschitz-continuous with Lipschitz-constant $L_f > 0$ and directionally differentiable.

Remark 2.2. Note that Assumption 2.1.1 is satisfied by the Volterra operator $\mathcal{H} : L^2(0,T;L^2(\Omega)) \to L^2(0,T;L^2(\Omega))$, defined as

$$[0,T] \ni t \mapsto \mathcal{H}(q)(t) := \int_0^t A(t-s)q(s)\,ds + q_0 \in L^2(\Omega),$$

where $A \in C([0,T];\mathcal{L}(L^2(\Omega),L^2(\Omega)))$ and $q_0 \in L^2(\Omega)$. This type of operator is often employed in the study of history-dependent evolutionary variational inequalities, see e.g. [33, Ch. 4.4].

Concerning Assumption 2.1.2, we remark that non-differentiable fatigue degradation functions are very common in applications, since such mappings often display at least one kink point, see [2, Sec. 2.6.2]. This basically means that once the cumulated fatigue $\mathcal{H}(q)$ achieves a certain value, say n_f, the body suddenly starts to become weaker in terms of its fracture toughness (so that n_f is a kink point of f). This abrupt weakening of the material is described by the monotonically decreasing mapping f on the interval $[n_f,\infty)$, see [2, Sec. 2.6.2].

Assumption 2.1 is supposed to hold throughout the paper, without mentioning it every time.

It is not difficult to check that the Nemytskii operator $f : L^2(\Omega) \to L^2(\Omega)$ is Lipschitz continuous with constant L_f. In view of Assumption 2.1.1, we thus have

$$\|(f\circ\mathcal{H})(q_1)(t) - (f\circ\mathcal{H})(q_2)(t)\|_{L^2(\Omega)} \le L_f\, L_{\mathcal{H}} \int_0^t \|q_1(s)-q_2(s)\|_{L^2(\Omega)}\,ds \tag{2.1}$$

a.e. in $(0,T)$, for all $q_1,q_2 \in L^2(0,T;L^2(\Omega))$.

Proposition 2.3 (Control-to-state map). *For every $\ell \in L^2(0,T;H^1(\Omega)^*)$, the fatigue damage problem (1.1) admits a unique solution $(q,\varphi) \in H^1_0(0,T;L^2(\Omega)) \times L^2(0,T;H^1(\Omega))$, which is characterized by the following PDE system*

$$\dot q(t) = \frac{1}{\epsilon}\max\big(-\beta(q(t)-\varphi(t)) - (f\circ\mathcal{H})(q)(t)\big) \ \text{ in } L^2(\Omega), \quad q(0)=0, \tag{2.2a}$$

$$-\alpha\Delta\varphi(t) + \beta\,\varphi(t) = \beta q(t) + \ell(t) \quad \text{in } H^1(\Omega)^* \tag{2.2b}$$

a.e. in $(0,T)$.

Proof. Let $t \in [0,T]$ and $\hat q : [0,T] \to L^2(\Omega)$ be arbitrary, but fixed. Since $\mathcal{E}(t,\cdot,\hat q(t))$ is strictly convex, continuous and radially unbounded (see (1.2)), the minimization problem $\min_{\varphi\in H^1(\Omega)} \mathcal{E}(t,\cdot,\hat q(t))$ admits a unique solution $\hat\varphi(t)$ characterized by $\partial_\varphi\mathcal{E}(t,\hat\varphi(t),\hat q(t)) = 0$ in $H^1(\Omega)^*$. In view of (1.2), this means that

$$\hat\varphi(t) \in \operatorname*{arg\,min}_{\varphi\in H^1(\Omega)} \mathcal{E}(t,\varphi,\hat q(t)) \iff \hat\varphi(t) = \phi(\hat q(t),\ell(t)), \tag{2.3}$$

where $\phi : L^2(\Omega)\times H^1(\Omega)^* \ni (\widetilde q,\widetilde\ell) \mapsto \widetilde\varphi \in H^1(\Omega)$ is the solution operator of

$$-\alpha\Delta\widetilde\varphi + \beta\,\widetilde\varphi = \beta\widetilde q + \widetilde\ell \quad \text{in } H^1(\Omega)^*. \tag{2.4}$$

With the map ϕ at hand, the evolution in (1.1) reads

$$-\partial_q\mathcal{E}(t,\phi(q(t),\ell(t)),q(t)) \in \partial_{\dot q}\mathcal{R}_\epsilon(\mathcal{H}(q)(t),\dot q(t)) \ \text{ a.e. in } (0,T). \tag{2.5}$$

In the light of (1.2), (1.3), and sum rule for convex subdifferentials, (2.5) is equivalent to

$$\begin{aligned}\mathcal{R}(\mathcal{H}(q)(t), v) - \mathcal{R}(\mathcal{H}(q)(t), \dot{q}(t)) + \epsilon\,(\dot{q}(t), v - \dot{q}(t))_{L^2(\Omega)} \\ \geq \beta\big(\phi(q(t), \ell(t)) - q(t), v - \dot{q}(t)\big)_{L^2(\Omega)} \quad \forall\, v \in L^2(\Omega),\ \text{a.e. in } (0,T),\end{aligned} \tag{2.6}$$

where

$$\mathcal{R} : L^2(\Omega) \times L^2(\Omega) \to \mathbb{R}, \quad \mathcal{R}(\omega, \eta) := \begin{cases} \int_\Omega f(\omega)\eta\, dx, & \text{if } \eta \geq 0 \text{ a.e. in } \Omega, \\ \infty & \text{otherwise.} \end{cases} \tag{2.7}$$

Now we use the result in [5, Lemma 3.3] for each time point t and we see that (2.6) is in fact equivalent with

$$\dot{q}(t) = \frac{1}{\epsilon}(\mathbb{I} - P_{\partial_{\dot q}\mathcal{R}(\mathcal{H}(q)(t),0)})\big(g(q(t), \ell(t))\big) \ \text{ a.e. in } (0,T), \tag{2.8}$$

where we abbreviate for convenience

$$g(q(t), \ell(t)) := \beta(\phi(q(t), \ell(t)) - q(t)). \tag{2.9}$$

In (2.8), $P_{\partial_{\dot q}\mathcal{R}(\mathcal{H}(q)(t),0)} : L^2(\Omega) \to L^2(\Omega)$ stands for the (metric) projection onto the set $\partial_{\dot q}\mathcal{R}(\mathcal{H}(q)(t), 0)$, i.e., $P_{\partial_{\dot q}\mathcal{R}(\mathcal{H}(q)(t),0)}\eta$ is the unique solution of

$$\min_{\mu \in \partial_{\dot q}\mathcal{R}(\mathcal{H}(q)(t),0)} \|\eta - \mu\|^2_{L^2(\Omega)}$$

for any $\eta \in L^2(\Omega)$. In order to compute $\partial_{\dot q}\mathcal{R}(\mathcal{H}(q)(t), 0)$, we use the definition of the convex subdifferential and the fact that $\mathcal{R}(\mathcal{H}(q)(t), 0) = 0$, from which we deduce

$$\partial_{\dot q}\mathcal{R}(\mathcal{H}(q)(t), 0) = \{\mu \in L^2(\Omega) |\, (\mu, v)_{L^2(\Omega)} \leq \mathcal{R}(\mathcal{H}(q)(t), v) \quad \forall\, v \in L^2(\Omega)\}.$$

Now, in view of (2.7) combined with the fundamental lemma of the calculus of variations we have

$$\partial_{\dot q}\mathcal{R}(\mathcal{H}(q)(t), 0) = \{\mu \in L^2(\Omega) |\, \mu \leq f(\mathcal{H}(q)(t)) \text{ a.e. in } \Omega\}.$$

This means that $P_{\partial_{\dot q}\mathcal{R}(\mathcal{H}(q)(t),0)}(\eta) = \min\{\eta, f(\mathcal{H}(q)(t))\}$ and since

$$\eta - \min\{\eta, f(\mathcal{H}(q)(t))\} = \max\{\eta - f(\mathcal{H}(q)(t)), 0\},$$

we can finally write (2.8) as

$$\dot{q}(t) = \frac{1}{\epsilon}\max\{g(q(t), \ell(t)) - f(\mathcal{H}(q)(t)), 0\} \quad \text{a.e. in } (0,T). \tag{2.10}$$

To summarize, we have shown that the evolution in (2.5) is equivalent to (2.10).

To solve (2.10), we apply a fixed-point argument. For this, we take a look at the mapping $L^2(0,t;L^2(\Omega)) \ni \eta \mapsto \mathcal{G}(\eta) \in H^1(0,t;L^2(\Omega))$, given by

$$\mathcal{G}(\eta)(\tau) := \int_0^\tau \max(g(\eta(s), \ell(s)) - (f \circ \mathcal{H})(\eta)(s))\, ds \quad \forall\, \tau \in [0,t],$$

where $t \in (0,T]$ is to be determined so that $\mathcal{G} : L^2(0,t;L^2(\Omega)) \to L^2(0,t;L^2(\Omega))$ is a contraction. For all $q_1, q_2 \in L^2(0,t;L^2(\Omega))$ the following estimate is true

$$\begin{aligned}\|\mathcal{G}(q_1)(\tau) - \mathcal{G}(q_2)(\tau)\|_{L^2(\Omega)} &\leq \int_0^\tau \|g(q_1(s), \ell(s)) - g(q_2(s), \ell(s))\|_{L^2(\Omega)}\, ds \\ &\quad + \int_0^\tau \|(f \circ \mathcal{H})(q_1)(s) - (f \circ \mathcal{H})(q_2)(s)\|_{L^2(\Omega)}\, ds \\ &\leq c \int_0^\tau \|q_1(s) - q_2(s)\|_{L^2(\Omega)}\, ds + L_f\, L_{\mathcal{H}} \int_0^\tau \int_0^s \|q_1(\zeta) - q_2(\zeta)\|_{L^2(\Omega)}\, d\zeta\, ds \\ &\leq c\, t^{1/2}\|q_1 - q_2\|_{L^2(0,t;L^2(\Omega))} + tL_f\, L_{\mathcal{H}}\|q_1 - q_2\|_{L^1(0,t;L^2(\Omega))} \\ &\leq (c\, t^{1/2} + L_f\, L_{\mathcal{H}}\, t^{3/2})\|q_1 - q_2\|_{L^2(0,t;L^2(\Omega))} \quad \text{for all } \tau \in [0,t],\end{aligned} \tag{2.11}$$

where $c > 0$ is a positive constant. Here we used the fact that $\max : L^2(\Omega) \to L^2(\Omega)$ is Lipschitzian with constant 1, the definition of g (see (2.9)) combined with the boundedness of ϕ, and the estimate (2.1). From (2.11) we deduce

$$\|\mathcal{G}(q_1) - \mathcal{G}(q_2)\|_{L^2(0,t;L^2(\Omega))} \le (c\,t + L_f\,L_{\mathcal{H}}\,t^2)\|q_1 - q_2\|_{L^2(0,t;L^2(\Omega))}, \tag{2.12}$$

which allows us to conclude that $\frac{1}{\epsilon}\,\mathcal{G}$ is a contraction for a small enough t. Thus, the PDE (2.10) restricted on $(0,t)$ admits a unique solution in $H_0^1(0,t;L^2(\Omega))$(see e.g. [14, Thm. 7.2.3]). Now, the unique solvability of (2.10) on the whole interval $(0,T)$ and the desired regularity of q follow by a concatenation argument.

Finally, we recall that $\varphi(\cdot) = \phi(q(\cdot), \ell(\cdot))$, cf. (2.3) and we deduce from (2.4) that $\varphi \in L^2(0,T;H^1(\Omega))$. To summarize, we obtained that (1.1) admits a unique solution $(q,\varphi) \in H_0^1(0,T;L^2(\Omega)) \times L^2(0,T;H^1(\Omega))$, which, owing to (2.3) and (2.10), is characterized by (2.2). □

Lemma 2.4. *The solution map associated to* (1.1)

$$S : L^2(0,T;H^1(\Omega)^*) \ni \ell \mapsto (q,\varphi) \in H_0^1(0,T;L^2(\Omega)) \times L^2(0,T;H^1(\Omega))$$

is Lipschitz continuous.

Proof. Let $\ell_1, \ell_2 \in L^2(0,T;H^1(\Omega)^*)$ be arbitrary, but fixed. In the following, we abbreviate $(q_i,\varphi_i) := S(\ell_i)$ and $g(q_i(\cdot),\ell_i(\cdot)) := \beta(\phi(q_i(\cdot),\ell_i(\cdot)) - q_i(\cdot))$, $i = 1,2$, where ϕ is the solution operator of (2.4) . In view of Proposition 2.3 combined with (2.1), we obtain

$$\begin{aligned}
\|(q_1 - q_2)(t)\|_{L^2(\Omega)} &\le \frac{1}{\epsilon}\int_0^t \|g(q_1(s),\ell_1(s)) - g(q_2(s),\ell_2(s))\|_{L^2(\Omega)}\,ds \\
&\quad + \frac{1}{\epsilon}\int_0^t \|(f\circ\mathcal{H})(q_1)(s) - (f\circ\mathcal{H})(q_2)(s)\|_{L^2(\Omega)}\,ds \\
&\le c\int_0^t \|q_1(s) - q_2(s)\|_{L^2(\Omega)} + \|\ell_1(s) - \ell_2(s)\|_{H^1(\Omega)^*}\,ds \\
&\quad + \frac{1}{\epsilon}L_f\,L_{\mathcal{H}}\int_0^t\int_0^s \|q_1(\zeta) - q_2(\zeta)\|_{L^2(\Omega)}\,d\zeta\,ds \quad \forall\, t \in [0,T],
\end{aligned}$$

where $c > 0$ is a constant dependent only on the given data. Then, applying Gronwall's inequality leads to

$$\|(q_1 - q_2)(t)\|_{L^2(\Omega)} \le \hat{c}\int_0^t \|\ell_1(s) - \ell_2(s)\|_{H^1(\Omega)^*}\,ds \quad \forall\, t \in [0,T],$$

where $\hat{c} > 0$ is a constant dependent only on the given data. By employing again (2.2a) and by estimating as above without integrating over time, we obtain

$$\|q_1 - q_2\|_{H^1(0,T;L^2(\Omega))} \le \widetilde{c}\,\|\ell_1 - \ell_2\|_{L^2(0,T;H^1(\Omega)^*)}, \tag{2.13}$$

where $\widetilde{c} > 0$ is another constant dependent only on the given data. Now, the desired result follows from $\varphi_i = \phi(q_i,\ell_i)$, $i = 1,2$, $\phi \in \mathcal{L}(L^2(\Omega) \times H^1(\Omega)^*, H^1(\Omega))$ and (2.13). □

Lemma 2.5. *The mapping* $(f\circ\mathcal{H}) : L^2(0,T;L^2(\Omega)) \to L^2(0,T;L^2(\Omega))$ *is Hadamard directionally differentiable with*

$$(f\circ\mathcal{H})'(\eta;\delta\eta) = f'(\mathcal{H}(\eta);\mathcal{H}'(\eta)(\delta\eta)) \quad \forall\, \eta,\delta\eta \in L^2(0,T;L^2(\Omega)). \tag{2.14}$$

Moreover, for all $\eta,\delta\eta_1,\delta\eta_2 \in L^2(0,T;L^2(\Omega))$, *it holds*

$$\|(f\circ\mathcal{H})'(\eta;\delta\eta_1)(t) - (f\circ\mathcal{H})'(\eta;\delta\eta_2)(t)\|_{L^2(\Omega)} \le L_f\,L_{\mathcal{H}}\int_0^t \|\delta\eta_1(s) - \delta\eta_2(s)\|_{L^2(\Omega)}\,ds \tag{2.15}$$

a.e. in $(0,T)$.

Proof. In view of the differentiability properties of $\mathcal{H}$ and f, the mapping $(f\circ\mathcal{H}) : L^2(0,T;L^2(\Omega)) \to L^2(0,T;L^2(\Omega))$ is Hadamard directionally differentiable [31, Def. 3.1.1, Lem. 3.1.2(b)]. To see this, we first note that $f : L^2(0,T;L^2(\Omega)) \to L^2(0,T;L^2(\Omega))$ is Hadamard directionally differentiable, since it is directionally differentiable (by Assumption 2.1.2 and Lebesgue's dominated convergence theorem, see e.g. [35, Lemma A.1]) and Lipschitz-continuous. In view of Assumption 2.1.1, chain rule [32, Prop. 3.6(i)] implies that $(f\circ\mathcal{H})$ is Hadamard directionally differentiable as well, with directional derivative given by (2.14). To prove (2.15), we observe that, as a consequence of (2.1), we have

$$\frac{1}{\tau}\|(f\circ\mathcal{H})(\eta+\tau\delta\eta_1)(t)-(f\circ\mathcal{H})(\eta+\tau\delta\eta_2)(t)\|_{L^2(\Omega)} \le L_f\,L_{\mathcal{H}}\int_0^t \|\delta\eta_1(s)-\delta\eta_2(s)\|_{L^2(\Omega)}\,ds$$

a.e. in $(0,T)$, for all $\eta,\delta\eta_1,\delta\eta_2 \in L^2(0,T;L^2(\Omega))$ and all $\tau>0$. Passing to the limit $\tau\searrow 0$, where one uses the directional differentiability of $f\circ\mathcal{H}$ and the fact that convergence in $L^2(0,T;L^2(\Omega))$ implies a.e. convergence in $L^2(\Omega)$ for a subsequence, then yields the desired estimate. □

Proposition 2.6 (Directional differentiability). *The operator $S : L^2(0,T;H^1(\Omega)^*) \to H^1_0(0,T;L^2(\Omega)) \times L^2(0,T;H^1(\Omega))$ is directionally differentiable. Its directional derivative $(\delta q,\delta\varphi) := S'(\ell;\delta\ell)$ at point $\ell \in L^2(0,T;H^1(\Omega)^*)$ in direction $\delta\ell \in L^2(0,T;H^1(\Omega)^*)$ is the unique solution of*

$$\dot{\delta q}(t) = \frac{1}{\epsilon}\max{}'\big(z(t);-\beta(\delta q(t)-\delta\varphi(t)) - f'(\mathcal{H}(q);\mathcal{H}'(q)(\delta q))(t)\big) \text{ in } L^2(\Omega),\ \ \delta q(0)=0, \tag{2.16a}$$

$$-\alpha\Delta\delta\varphi(t) + \beta\,\delta\varphi(t) = \beta\delta q(t) + \delta\ell(t) \quad \text{in } H^1(\Omega)^* \tag{2.16b}$$

a.e. in $(0,T)$, where we abbreviate $z(t) := -\beta(q(t)-\varphi(t)) - (f\circ\mathcal{H})(q)(t)$.

Proof. We start by examining the solvability of (2.16). To this end, we just check that the mapping $L^2(0,t;L^2(\Omega)) \ni \eta \mapsto \widehat{\mathcal{G}}(\eta) \in H^1(0,t;L^2(\Omega))$, given by

$$\widehat{\mathcal{G}}(\eta)(\tau) := \int_0^\tau \max{}'\big(z(s);-\beta(\eta(s)-\phi(\eta(s),\delta\ell(s)) - f'\big(\mathcal{H}(q);\mathcal{H}'(q)(\eta)\big)(s)\big)\,ds$$

for all $\tau\in[0,t]$, is Lipschitzian from $L^2(0,t;L^2(\Omega))$ to $L^2(0,t;L^2(\Omega))$ with constant smaller than ϵ, for $t\in(0,T]$ small enough. Then, by using the arguments employed at the end of the proof of Proposition 2.3, we can deduce that, for any $\delta\ell \in L^2(0,T;H^1(\Omega)^*)$, (2.16) admits a unique solution $(\delta q,\delta\varphi) \in H^1_0(0,T;L^2(\Omega))\times L^2(0,T;H^1(\Omega))$. For all $\eta_1,\eta_2 \in L^2(0,t;L^2(\Omega))$ the following estimate is true

$$\begin{aligned}
\|\widehat{\mathcal{G}}(\eta_1)(\tau)-\widehat{\mathcal{G}}(\eta_2)(\tau)\|_{L^2(\Omega)} &\le \int_0^\tau \|g(\eta_1(s),\delta\ell(s)) - g(\eta_2(s),\delta\ell(s))\|_{L^2(\Omega)}\,ds \\
&\quad + \int_0^\tau \|f'(\mathcal{H}(q);\mathcal{H}'(q;\eta_1))(s) - f'(\mathcal{H}(q);\mathcal{H}'(q;\eta_2))(s)\|_{L^2(\Omega)}\,ds \\
&\le c\int_0^\tau \|\eta_1(s)-\eta_2(s)\|_{L^2(\Omega)}\,ds + L_f\,L_{\mathcal{H}}\int_0^\tau\int_0^s \|\eta_1(\zeta)-\eta_2(\zeta)\|_{L^2(\Omega)}\,d\zeta\,ds \\
&\le c\,t^{1/2}\|\eta_1-\eta_2\|_{L^2(0,t;L^2(\Omega))} + tL_f\,L_{\mathcal{H}}\|\eta_1-\eta_2\|_{L^1(0,t;L^2(\Omega))} \\
&\le (c\,t^{1/2} + L_f\,L_{\mathcal{H}}\,t^{3/2})\|\eta_1-\eta_2\|_{L^2(0,t;L^2(\Omega))} \qquad \text{for all } \tau\in[0,t],
\end{aligned}$$

where $c>0$ is a positive constant; note that here we abbreviated again $g(\eta_i(\cdot),\delta\ell(\cdot)) := \beta(\phi(\eta_i(\cdot),\delta\ell(\cdot)) - \eta_i(\cdot))$, $i=1,2$. Here we used the fact that $\max'(z(s),\cdot) : L^2(\Omega)\to L^2(\Omega)$ is Lipschitzian with constant 1, the boundedness of ϕ (see (2.4)), and (2.15) in combination with (2.14). Then, we obtain an estimate similar to (2.12) which allows us to conclude the fact that $\frac{1}{\epsilon}\widehat{\mathcal{G}}$ is a contraction.

Next we focus on the convergence of the difference quotients associated with the mapping S. We begin by observing that the operator $\max : L^2(0,T;L^2(\Omega)) \to L^2(0,T;L^2(\Omega))$ is Hadamard directionally differentiable [31, Def. 3.1.1, Lem. 3.1.2(b)], since it is directionally differentiable (by Lebesgue's dominated convergence theorem, see e.g. [35, Lem. A.1]) and Lipschitz-continuous. Moreover,

$$G : (\eta, \psi) \mapsto -\beta(\eta - \phi(\eta, \psi)) - (f \circ \mathcal{H})(\eta)$$

is directionally differentiable from $L^2(0, T; L^2(\Omega)) \times L^2(0, T; H^1(\Omega)^*)$ to $L^2(0, T; L^2(\Omega))$, since ϕ is linear and bounded between these spaces (cf. (2.4)) and as a result of Lemma 2.5. Now chain rule [32, Prop. 3.6(i)] implies that

$$\mathcal{F} := \max \circ G$$

is (Hadamard) directionally differentiable from $L^2(0, T; L^2(\Omega)) \times L^2(0, T; H^1(\Omega)^*)$ to $L^2(0, T; L^2(\Omega))$ with

$$\mathcal{F}'((q, \ell); (\delta q, \delta\ell)) = \max{}'\big(G(q, \ell); G'((q, \ell); (\delta q, \delta\ell))\big)$$

for all $(q, \ell), (\delta q, \delta\ell) \in L^2(0, T; L^2(\Omega)) \times L^2(0, T; H^1(\Omega)^*)$. For simplicity, in the following we abbreviate $q^\tau := S_1(\ell + \tau\,\delta\ell)$, where $\tau > 0$ is arbitrary, but fixed. S_1 denotes the first component of the map S, i.e., $S_1 : L^2(0, T; H^1(\Omega)^*) \to H^1_0(0, T; L^2(\Omega))$ is the solution map associated with (2.10). By combining the equations for q^τ, q and (2.16), we obtain

$$\begin{aligned}
&\frac{d}{dt}\Big(\frac{q^\tau - q}{\tau} - \delta q\Big) = \frac{\mathcal{F}(q^\tau, \ell + \tau\,\delta\ell) - \mathcal{F}(q, \ell)}{\tau} - \mathcal{F}'\big((q, \ell); (\delta q, \delta\ell)\big) \quad \text{a.e. in } (0, T), \\
&\Big(\frac{q^\tau - q}{\tau} - \delta q\Big)(0) = 0.
\end{aligned} \tag{2.17}$$

This implies

$$\begin{aligned}
&\Big\|\Big(\frac{q^\tau - q}{\tau} - \delta q\Big)(t)\Big\|_{L^2(\Omega)} \\
&\leq \int_0^t \Big\|\frac{\mathcal{F}\big(q^\tau, \ell + \tau\,\delta\ell\big)(s) - \mathcal{F}\big((q, \ell) + \tau(\delta q, \delta\ell)\big)(s)}{\tau}\Big\|_{L^2(\Omega)} \\
&\qquad + \underbrace{\Big\|\frac{\mathcal{F}\big((q, \ell) + \tau(\delta q, \delta\ell)\big)(s) - \mathcal{F}(q, \ell)(s)}{\tau} - \mathcal{F}'\big((q, \ell); (\delta q, \delta\ell)\big)(s)\Big\|_{L^2(\Omega)}}_{=:A_\tau(s)} ds \\
&\leq \int_0^t \Big\|\frac{G(q^\tau, \ell + \tau\delta\ell)(s) - G((q, \ell) + \tau(\delta q, \delta\ell))(s)}{\tau}\Big\|_{L^2(\Omega)} ds + \|A_\tau\|_{L^1(0,t;L^2(\Omega))} \\
&\leq c \int_0^t \Big\|\Big(\frac{q^\tau - q}{\tau} - \delta q\Big)(s)\Big\|_{L^2(\Omega)} ds + L_f\, L_{\mathcal{H}} \int_0^t \int_0^s \Big\|\Big(\frac{q^\tau - q}{\tau} - \delta q\Big)(\zeta)\Big\|_{L^2(\Omega)} d\zeta\, ds \\
&\qquad + \|A_\tau\|_{L^1(0,T;L^2(\Omega))} \qquad \forall\, t \in [0, T],
\end{aligned} \tag{2.18}$$

where $c > 0$ is the positive constant appearing in (2.11). In (2.18) we used again the Lipschitz continuity of $\max : L^2(\Omega) \to L^2(\Omega)$, the boundedness of ϕ (cf. (2.3) and (2.4)), and the estimate (2.1). Applying Gronwall's inequality in (2.18) yields

$$\Big\|\Big(\frac{q^\tau - q}{\tau} - \delta q\Big)(t)\Big\|_{L^2(\Omega)} \leq C\, \|A_\tau\|_{L^1(0,T;L^2(\Omega))} \qquad \forall\, t \in [0, T], \tag{2.19}$$

where $C > 0$ is a constant dependent only on the given data. Now, (2.17) and estimating as in (2.18), in combination with (2.19), leads to

$$\Big\|\frac{q^\tau - q}{\tau} - \delta q\Big\|_{H^1(0,T;L^2(\Omega))} \leq \widehat{C}\, \|A_\tau\|_{L^2(0,T;L^2(\Omega))} \quad \forall\, \tau > 0, \tag{2.20}$$

where $\hat{C} > 0$ is a constant dependent only on the given data. On the other hand, we recall the definition of A_τ in (2.18) and the fact that $\mathcal{F}$ is directionally differentiable from $L^2(0,T;L^2(\Omega)) \times L^2(0,T;H^1(\Omega)^*)$ to $L^2(0,T;L^2(\Omega))$, which implies

$$\|A_\tau\|_{L^2(0,T;L^2(\Omega))} \to 0 \quad \text{as } \tau \searrow 0.$$

In view of (2.20), we have shown that $S_1 : L^2(0,T;H^1(\Omega)^*) \to H^1(0,T;L^2(\Omega)$ is directionally differentiable with $S_1'(\ell;\delta\ell) = \delta q$. Further, from (2.3) we have $S_2(\ell) = \phi(S_1(\ell),\ell)$ for all $\ell \in L^2(0,T;H^1(\Omega)^*)$, where S_2 is the second component of the operator S, i.e., $S_2 : L^2(0,T;H^1(\Omega)^*) \ni \ell \mapsto \varphi \in L^2(0,T;H^1(\Omega))$. Thus, S_2 is directionally differentiable as well, since $\phi \in \mathcal{L}(L^2(\Omega) \times H^1(\Omega)^*; H^1(\Omega))$ and S_1 is directionally differentiable. Its directional derivative $S_2'(\ell;\delta\ell)$ is given by $\phi(S_1'(\ell;\delta\ell),\delta\ell)$, i.e., $S_2'(\ell;\delta\ell) = \delta\varphi$, see (2.16). The proof is now complete. □

3 THE OPTIMAL CONTROL PROBLEM

Now, we turn our attention to the optimal control of the fatigue damage model (1.1). In the remainder of the paper, we are concerned with the examination of the following optimal control problem

$$\begin{aligned} \min_{\ell \in H^1(0,T;L^2(\Omega))} \quad & J(q,\varphi,\ell) \\ \text{s.t.} \quad & (q,\varphi) \text{ solves (1.1) with r.h.s. } \ell. \end{aligned}$$

In view of Proposition 2.3, this can also be formulated as

$$\text{(P)} \qquad \left.\begin{aligned} \min_{\ell \in H^1(0,T;L^2(\Omega))} \quad & J(q,\varphi,\ell) \\ \text{s.t.} \quad & (q,\varphi) \text{ solves (2.2) with r.h.s. } \ell. \end{aligned}\right\}$$

Assumption 3.1. The functional J satisfies

$$J(q,\varphi,\ell) = j(q,\varphi) + \frac{1}{2}\|\ell\|^2_{H^1(0,T;L^2(\Omega))},$$

where $j : L^2(0,T;L^2(\Omega)) \times L^2(0,T;H^1(\Omega)) \to \mathbb{R}$ is continuously Fréchet-differentiable.

Note that Assumption 3.1 is satisfied by classical objectives of tracking type such as

$$J_{ex}(q,\varphi,\ell) := \frac{1}{2}\,\|q - q_d\|^2_{L^2(0,T;L^2(\Omega))} + \frac{\kappa}{2}\|\varphi\|^2_{L^2(0,T;H^1(\Omega))} + \frac{1}{2}\|\ell\|^2_{H^1(0,T;L^2(\Omega))},$$

where $q_d \in L^2(0,T;L^2(\Omega))$ and $\kappa \geq 0$.

Proposition 3.2 (Existence of optimal solutions for (P)). *The optimal control problem* (P) *admits at least one solution in* $H^1(0,T;L^2(\Omega))$.

Proof. The assertion follows by standard arguments which rely on the direct method of the calculus of variations combined with the radial unboundedness of the reduced objective

$$H^1(0,T;L^2(\Omega)) \ni \ell \mapsto J(S(\ell),\ell) \in \mathbb{R},$$

the Lipschitz continuity of S on $L^2(0,T;H^1(\Omega)^*)$ (Lemma 2.4), the compact embedding

$$H^1(0,T;L^2(\Omega)) \hookrightarrow\hookrightarrow L^2(0,T;H^1(\Omega)^*)$$

and the continuity of j from Assumption 3.1. □

3.1 Regularization and passage to the limit

In this section, we are concerned with the derivation of a first optimality system for local optima of (P). Based thereon, we shall improve our optimality conditions in the next section.

To obtain a first strong optimality system, see (3.5) below, we need the following rather non-restrictive assumption:

Assumption 3.3. In addition to Assumption 2.1, we require that the mappings associated with fatigue in (1.1) satisfy:

1. The *history operator* $\mathcal{H} : L^2(0,T;L^\infty(\Omega)) \to L^2(0,T;L^\infty(\Omega))$ fulfills

$$\|\mathcal{H}(q_1)(t) - \mathcal{H}(q_2)(t)\|_{L^\infty(\Omega)} \leq \widehat{L}_{\mathcal{H}} \int_0^t \|q_1(s) - q_2(s)\|_{L^\infty(\Omega)}\, ds \quad \text{a.e. in } (0,T),$$

   for all $q_1, q_2 \in L^2(0,T;L^\infty(\Omega))$, where $\widehat{L}_{\mathcal{H}} > 0$ is a positive constant.

2. The non-differentiable function $f : \mathbb{R} \to \mathbb{R}$ is assumed to have one non-smooth point n_f. Moreover, f is assumed to be continuously differentiable on $\mathbb{R} \setminus \{n_f\}$.

Remark 3.4. Similarly to Remark 2.2, we observe that Assumption 3.3.1 is satisfied by classical Volterra operators which are employed in the study of history-dependent evolutionary variational inequalities, i.e., $\mathcal{H} : L^2(0,T;L^\infty(\Omega)) \to L^2(0,T;L^\infty(\Omega))$

$$[0,T] \ni t \mapsto \mathcal{H}(q)(t) := \int_0^t A(t-s)q(s)\, ds + q_0 \in L^\infty(\Omega),$$

where $A \in C([0,T];\mathcal{L}(L^\infty(\Omega), L^\infty(\Omega)))$ and $q_0 \in L^\infty(\Omega)$.

We underline that Assumption 3.3.2 is very reasonable from the point of view of applications, since fatigue degradation functions have at most two kink points in practice [2, Sec. 2.6.2]; moreover, such mappings are always piecewise continuously differentiable. Our mathematical analysis can be carried on in an analogous way if f has a countable number of non-smooth points; since this is rather uncommon in applications and for the sake of a better overview, we stick to the setting where f has a single non-differentiable point.

In the rest of the paper, we will tacitly assume that, in addition to Assumptions 2.1 and 3.1, Assumption 3.3 is always fulfilled, without mentioning it every time.

Definition 3.5 (Regularization of f). For every $\varepsilon > 0$, the differentiable function $f_\varepsilon : \mathbb{R} \to \mathbb{R}$ is defined as

$$f_\varepsilon(v) := \int_{-\infty}^{\infty} f(v - \varepsilon s)\psi(s)\, ds,$$

where $\psi \in C_c^\infty(\mathbb{R})$, $\psi \geq 0$, $\operatorname{supp}\psi \subset [-1,1]$ and $\int_{-\infty}^{\infty} \psi(s)\, ds = 1$.

Lemma 3.6 (Properties of f_ε). *The following assertions are true:*

1. *There exists a constant $C > 0$, independent of ε, such that*

$$|f_\varepsilon(v) - f(v)| \leq C\varepsilon \quad \forall\, v \in \mathbb{R},\ \forall\, \varepsilon > 0.$$

2. *For every $\varepsilon > 0$, f_ε is Lipschitz continuous. Its Lipschitz constant is given by $L_f > 0$, and it is thus independent of ε.*

3. *For every $\delta > 0$ and every $K \geq |n_f| + \delta$, the sequence $\{f_\varepsilon'\}$ converges uniformly towards f' on $[-K, n_f - \delta] \cup [n_f + \delta, K]$ as $\varepsilon \searrow 0$.*

Proof. The first two statements are an immediate consequence of the Lipschitz continuity of f. Note that $f_\varepsilon = \psi_\varepsilon \star f$, where $\psi_\varepsilon(\cdot) = \psi(\cdot/\varepsilon)/\varepsilon$. Then, the third statement is a consequence of the fact that f' is continuous on $[-K, n_f - \delta] \cup [n_f + \delta, K]$ in combination with the fact that $f'_\varepsilon = \psi_\varepsilon \star f'$. □

For an arbitrary local minimizer $\bar\ell$ of (P), consider the following regularization, also known as "adapted penalization", see e.g. [4]:

$$
(\mathrm{P}_\varepsilon)\qquad \left.\begin{aligned}
\min_{\ell \in H^1(0,T;L^2(\Omega))} \quad & J(q,\varphi,\ell) + \frac{1}{2}\|\ell - \bar\ell\|^2_{H^1(0,T;L^2(\Omega))} \\
\text{s.t.} \quad & \dot q(t) = \frac{1}{\epsilon} \max{}_\varepsilon(-\beta(q(t) - \varphi(t)) - (f_\varepsilon \circ \mathcal{H})(q)(t)) \ \text{ in } L^2(\Omega), \\
& q(0) = 0, \\
& -\alpha\Delta\varphi(t) + \beta\,\varphi(t) = \beta q(t) + \ell(t) \quad \text{in } H^1(\Omega)^*, \quad \text{a.e. in } (0,T),
\end{aligned}\right\}
$$

where

$$
\max{}_\varepsilon : \mathbb{R} \to \mathbb{R}, \quad \max{}_\varepsilon(x) := \begin{cases} 0, & x \le 0, \\ \frac{1}{2\varepsilon}x^2, & x \in (0,\varepsilon), \\ x - \frac{\varepsilon}{2}, & x \ge \varepsilon. \end{cases}
$$

Lemma 3.7. *For each local optimum $\bar\ell$ of (P) there exists a sequence of local minimizers $\{\ell_\varepsilon\}$ of (P_ε) such that*

$$
\ell_\varepsilon \to \bar\ell \quad \text{in } H^1(0,T;L^2(\Omega)) \quad \text{as } \varepsilon \searrow 0. \tag{3.1}
$$

Moreover,

$$
S_\varepsilon(\ell_\varepsilon) \to S(\bar\ell) \quad \text{in } H^1_0(0,T;L^2(\Omega)) \times L^2(0,T;H^1(\Omega)) \quad \text{as } \varepsilon \searrow 0, \tag{3.2}
$$

where $S_\varepsilon : L^2(0,T;H^1(\Omega)^*) \ni \ell \mapsto (q_\varepsilon, \varphi_\varepsilon) \in H^1_0(0,T;L^2(\Omega)) \times L^2(0,T;H^1(\Omega))$ is the control-to-state map associated to the state equation in (P_ε).

Proof. see Appendix A. □

The next result is essential for the solvability of the first adjoint equation in (3.5).

Lemma 3.8. *For all $\eta, \delta\eta \in L^2(0,T;L^2(\Omega))$ it holds*

$$
\|[(f_\varepsilon \circ \mathcal{H})'(\eta)]^\star(\delta\eta)(t)\|_{L^2(\Omega)} \le \hat L_f\, L_{\mathcal{H}} \int_t^T \|\delta\eta(s)\|_{L^2(\Omega)}\, ds \quad \text{a.e. in } (0,T), \tag{3.3}
$$

where $[(f_\varepsilon \circ \mathcal{H})'(\eta)]^\star : L^2(0,T;L^2(\Omega)) \to L^2(0,T;L^2(\Omega))$ stands for the adjoint operator of $(f_\varepsilon \circ \mathcal{H})'(\eta)$.

Proof. Let $\psi \in L^2(0,T;L^2(\Omega))$ be arbitrary, but fixed. By virtue of (2.15) (applied for f_ε instead of f), we have

$$
\begin{aligned}
([(f_\varepsilon \circ \mathcal{H})'(\eta)]^\star(\delta\eta), \psi)_{L^2(0,T;L^2(\Omega))} &= ((f_\varepsilon \circ \mathcal{H})'(\eta)(\psi), \delta\eta)_{L^2(0,T;L^2(\Omega))} \\
&\le \int_0^T \hat L_f\, L_{\mathcal{H}} \int_0^T \chi_{[0,t]}(s)\|\psi(s)\|_{L^2(\Omega)}\, ds\, \|\delta\eta(t)\|_{L^2(\Omega)}\, dt \\
&= \hat L_f\, L_{\mathcal{H}} \int_0^T \int_0^T \chi_{[0,s]}(t)\|\delta\eta(s)\|_{L^2(\Omega)}\, ds\, \|\psi(t)\|_{L^2(\Omega)}\, dt \\
&= \hat L_f\, L_{\mathcal{H}} \int_0^T \int_t^T \|\delta\eta(s)\|_{L^2(\Omega)}\, ds\, \|\psi(t)\|_{L^2(\Omega)}\, dt.
\end{aligned}
$$

Note that in the first identity we made use of Fubini's theorem. Now, testing with $\psi := v\rho$, where $v \in L^2(\Omega)$ and $\rho \in L^2(0,T)$, $\rho \geq 0$, are arbitrary, but fixed yields

$$\int_0^T ([(f_\varepsilon \circ \mathcal{H})'(\eta)]^\star(\delta\eta)(t), v)_{L^2(\Omega)}\rho(t)\,dt \\ \leq \hat{L}_f\, L_{\mathcal{H}} \int_0^T \int_t^T \|\delta\eta(s)\|_{L^2(\Omega)}\,ds\,\|v\|_{L^2(\Omega)}\rho(t)\,dt.$$

Applying the fundamental lemma of the calculus of variations then gives in turn

$$([(f_\varepsilon \circ \mathcal{H})'(\eta)]^\star(\delta\eta)(t), v)_{L^2(\Omega)} \leq \hat{L}_f\, L_{\mathcal{H}} \int_t^T \|\delta\eta(s)\|_{L^2(\Omega)}\,ds\,\|v\|_{L^2(\Omega)}$$

a.e. in $(0,T)$. Since $v \in L^2(\Omega)$ was arbitrary, the proof is now complete. □

To show that the relations in (3.12) below are valid, we need to prove that the convergence in (3.2) is true in $L^\infty(0,T;L^\infty(\Omega))$ as well. This is confirmed by the following

Lemma 3.9. *Let $\{\ell_\varepsilon\}$ be the sequence of local minimizers from Lemma 3.7 associated to a local optimum $\bar{\ell}$ of* (P). *Then,*

$$S_\varepsilon(\ell_\varepsilon) \to S(\bar{\ell}) \quad \text{in } L^\infty(0,T;L^\infty(\Omega)) \times L^\infty(0,T;L^\infty(\Omega)) \quad \text{as } \varepsilon \searrow 0. \tag{3.4}$$

Proof. Let us first show that $(\bar{q},\bar{\varphi})$ belongs to $L^\infty(0,T;L^\infty(\Omega)) \times L^\infty(0,T;L^\infty(\Omega))$. The assertion for $(q_\varepsilon,\varphi_\varepsilon)$ follows in a completely analogous way. By taking a look at (2.2), we see that, since $\bar{\ell} \in L^\infty(0,T;L^2(\Omega))$, the mapping $\bar{\varphi}$ belongs to the space $L^\infty(0,T;L^\infty(\Omega))$; this follows by the so-called Stampacchia method, cf. e.g. [37, Chp. 7.2.2]. Then, by arguing as in the proof of Proposition 2.3, where one employs Assumption 3.3.1, one obtains that $\bar{q} \in H^1(0,T;L^\infty(\Omega)) \subset L^\infty(0,T;L^\infty(\Omega))$. Now, to show the convergence (3.4), we subtract the equation associated to $\bar{q}$ (see (2.2a)) from the one associated to q_ε (see (A.1a)). By using the fact that $|\max_\varepsilon(x) - \max(x)| \leq \varepsilon\ \forall\, x \in \mathbb{R}$, and by relying on the Lipschitz continuity of max and f, as well as Lemma 3.6.1, we arrive at

$$\begin{aligned}
\|(q_\varepsilon - \bar{q})(t)\|_{L^\infty(\Omega)} &\leq \frac{2\varepsilon t}{\epsilon} + \frac{c}{\epsilon}\int_0^t \|q_\varepsilon(s) - \bar{q}(s)\|_{L^\infty(\Omega)} + \|\varphi_\varepsilon(s) - \bar{\varphi}(s)\|_{L^\infty(\Omega)}\,ds \\
&\quad + \frac{L_f}{\epsilon}\int_0^t \|\mathcal{H}(q_\varepsilon)(s) - \mathcal{H}(\bar{q})(s)\|_{L^\infty(\Omega)}\,ds \\
&\leq \frac{2\varepsilon t}{\epsilon} + \frac{c}{\epsilon}\int_0^t \|q_\varepsilon(s) - \bar{q}(s)\|_{L^\infty(\Omega)} + \|\ell_\varepsilon(s) - \bar{\ell}(s)\|_{L^2(\Omega)}\,ds \\
&\quad + \frac{L_f\,\widehat{L}_{\mathcal{H}}}{\epsilon}\int_0^t\int_0^s \|q_\varepsilon(\zeta) - \bar{q}(\zeta)\|_{L^\infty(\Omega)}\,d\zeta\,ds \quad \forall\, t \in [0,T],
\end{aligned}$$

where $c > 0$ is a constant dependent only on the given data; note that in the last inequality we used Assumption 3.3.1. Then, applying Gronwall's inequality leads to

$$\|(q_\varepsilon - \bar{q})(t)\|_{L^\infty(\Omega)} \leq \hat{c}\left(\frac{2\varepsilon t}{\epsilon} + \frac{c}{\epsilon}\int_0^t \|\ell_\varepsilon(s) - \bar{\ell}(s)\|_{L^2(\Omega)}\,ds\right) \quad \forall\, t \in [0,T],$$

where $\hat{c} > 0$ is a constant dependent only on the given data. By employing (3.1), we can finally deduce that $q_\varepsilon \to \bar{q}$ in $L^\infty(0,T;L^\infty(\Omega))$. In view of (2.2b), the proof is now complete. □

We are now in the position to state the main result of this subsection.

Proposition 3.10. *Suppose that Assumptions 3.1 and 3.3 are fulfilled. Let $\bar\ell$ be a local optimum of* (P) *with associated state $(\bar q, \bar\varphi) \in H_0^1(0,T;L^2(\Omega)) \times L^2(0,T;H^1(\Omega))$. Then there exist adjoint states*

$$\xi \in H_T^1(0,T;L^2(\Omega)) \text{ and } w \in L^2(0,T;H^1(\Omega))$$

and multipliers $\lambda \in L^\infty(0,T;L^2(\Omega))$ and $\mu \in L^\infty(0,T;L^2(\Omega))$ such that the following optimality system is satisfied

$$-\dot\xi - \beta\big(w-\lambda\big) + \mathcal{H}'(\bar q)^\star(\mu) = \partial_q j(\bar q,\bar\varphi) \ \text{ in } L^2(0,T;L^2(\Omega)), \quad \xi(T)=0, \tag{3.5a}$$

$$-\alpha\Delta w + \beta\big(w-\lambda\big) = \partial_\varphi j(\bar q,\bar\varphi) \ \text{ in } L^2(0,T;H^1(\Omega)^*), \tag{3.5b}$$

$$\lambda(t,x) = \frac{1}{\epsilon}\chi_{\{\bar z>0\}}(t,x)\xi(t,x) \quad \text{a.e. where } \bar z(t,x)\neq 0, \tag{3.5c}$$

$$\mu(t,x) = f'(\mathcal{H}(\bar q)(t,x))\lambda(t,x) \quad \text{a.e. where } \mathcal{H}(\bar q)(t,x) \neq n_f, \tag{3.5d}$$

$$(w,\delta\ell)_{L^2(0,T;L^2(\Omega))} + (\bar\ell,\delta\ell)_{H^1(0,T;L^2(\Omega))} = 0 \quad \forall \delta\ell \in H^1(0,T;L^2(\Omega)), \tag{3.5e}$$

where we abbreviate $\bar z := -\beta(\bar q - \bar\varphi) - (f\circ\mathcal{H})(\bar q)$.

Proof. Let $\{\ell_\varepsilon\}$ be the sequence of local minimizers from Lemma 3.7. Since ℓ_ε is locally optimal for (P_ε) and on account of the differentiability properties of S_ε, cf. Appendix A, and J, see Assumption 3.1, we can write down the necessary optimality condition

$$j'(S_\varepsilon(\ell_\varepsilon))(S_\varepsilon'(\ell_\varepsilon)(\delta\ell)) + (\ell_\varepsilon,\delta\ell)_{H^1(0,T;L^2(\Omega))} + (\ell_\varepsilon - \bar\ell,\delta\ell)_{H^1(0,T;L^2(\Omega))} = 0 \tag{3.6}$$

for all $\delta\ell \in H^1(0,T;L^2(\Omega))$. Now, let us consider the system

$$\begin{aligned} -\dot\xi_\varepsilon(t) - \beta\big(w_\varepsilon(t) - \tfrac{1}{\epsilon}\max{}_\varepsilon{}'(z_\varepsilon(t))\xi_\varepsilon(t)\big) & \\ + \mathcal{H}'(q_\varepsilon)^\star\Big(f_\varepsilon'(\mathcal{H}(q_\varepsilon))\big(\tfrac{1}{\epsilon}\max{}_\varepsilon{}'(z_\varepsilon)\xi_\varepsilon\big)\Big)(t) &= \partial_q j(S_\varepsilon(\ell_\varepsilon))(t), \quad \xi_\varepsilon(T)=0, \end{aligned} \tag{3.7a}$$

$$-\alpha\Delta w_\varepsilon(t) + \beta\big(w_\varepsilon(t) - \frac{1}{\epsilon}\max{}_\varepsilon{}'(z_\varepsilon(t))\xi_\varepsilon(t)\big) = \partial_\varphi j(S_\varepsilon(\ell_\varepsilon))(t) \tag{3.7b}$$

a.e. in $(0,T)$, where we abbreviate $z_\varepsilon := -\beta(q_\varepsilon - \varphi_\varepsilon) - (f_\varepsilon\circ\mathcal{H})(q_\varepsilon)$ and $(q_\varepsilon,\varphi_\varepsilon) := S_\varepsilon(\ell_\varepsilon)$. In (3.7a), $\mathcal{H}'(q_\varepsilon)^\star : L^2(0,T;L^2(\Omega)) \to L^2(0,T;L^2(\Omega))$ stands for the adjoint operator of $\mathcal{H}'(q_\varepsilon)$.

By arguments inspired e.g. from the proof of [35, Lem. 5.7] in combination with the estimate (3.3), one obtains that (3.7) admits a unique solution $(\xi_\varepsilon, w_\varepsilon) \in H_T^1(0,T;L^2(\Omega)) \times L^2(0,T;H^1(\Omega))$. Let us go a little more into detail concerning the solvability of (3.7a). In this context one checks if the mapping $L^2(0,t;L^2(\Omega)) \ni \eta \mapsto \mathcal{G}(\eta) \in H^1(0,t;L^2(\Omega))$, given by

$$\begin{aligned} \mathcal{G}(\eta)(\tau) := \int_0^\tau & \beta\big(w_\varepsilon(T-s,\eta(s)) - \frac{1}{\epsilon}\max{}_\varepsilon{}'(z_\varepsilon(T-s))\eta(s)\big) \\ & - \underbrace{\mathcal{H}'(q_\varepsilon)^\star\Big(f_\varepsilon'(\mathcal{H}(q_\varepsilon))\big(\frac{1}{\epsilon}\max{}_\varepsilon{}'(z_\varepsilon)\eta(T-\cdot)\big)\Big)}_{=[(f_\varepsilon\circ\mathcal{H})'(q_\varepsilon)]^\star\left(\frac{1}{\epsilon}\max{}_\varepsilon{}'(z_\varepsilon)(\eta(T-\cdot))\right)}(T-s) + \partial_q j(S_\varepsilon(\ell_\varepsilon))(T-s)\, ds \end{aligned}$$

for all $\tau \in [0,t]$, is Lipschitzian from $L^2(0,t;L^2(\Omega))$ to $L^2(0,t;L^2(\Omega))$ with constant smaller than 1, for $t \in (0,T]$ small enough; here $w_\varepsilon(t,v)$ denotes the solution of

$$-\alpha\Delta w_\varepsilon(t,v) + \beta\big(w_\varepsilon(t,v) - \frac{1}{\epsilon}\max{}_\varepsilon{}'(z_\varepsilon(t))v\big) = \partial_\varphi j(S_\varepsilon(\ell_\varepsilon))(t)$$

for $t \in [0,T]$ and $v \in L^2(\Omega)$. We observe that, for all $\eta_1, \eta_2 \in L^2(0,t;L^2(\Omega))$, the following estimate is true

$$\begin{aligned}\|\mathcal{G}(\eta_1)(\tau) - \mathcal{G}(\eta_2)(\tau)\|_{L^2(\Omega)} &\leq c \int_0^\tau \|\eta_1(s) - \eta_2(s)\|_{L^2(\Omega)}\,ds \\ &\quad + \int_0^\tau \hat{L}_f\, L_{\mathcal{H}} \int_{T-s}^T \|(\eta_1 - \eta_2)(T-\zeta)\|_{L^2(\Omega)}\,d\zeta\ ds \\ &\leq c\, t^{1/2} \|\eta_1 - \eta_2\|_{L^2(0,t;L^2(\Omega))} + \hat{L}_f\, L_{\mathcal{H}} \int_0^\tau \int_0^s \|(\eta_1 - \eta_2)(\zeta)\|_{L^2(\Omega)}\,d\zeta\ ds \\ &\leq (c\, t^{1/2} + \hat{L}_f\, L_{\mathcal{H}}\, t^{3/2}) \|\eta_1 - \eta_2\|_{L^2(0,t;L^2(\Omega))} \quad \text{for all } \tau \in [0,t],\end{aligned}$$

where in the first inequality we used the global Lipschitz-continuity of $\max_\varepsilon$ with constant 1 and (3.3); now the reader is referred to the first part of the proof of Proposition 2.6 where the exact type of estimate was established in order to obtain that $\eta = \mathcal{G}(\eta)$ admits a solution in $H_0^1(0,T;L^2(\Omega))$; finally, a transformation of the variables yields that $(\xi_\varepsilon, w_\varepsilon) := (\eta(T-\cdot), w_\varepsilon(t, \eta(T-\cdot))$ is the solution of the adjoint system (3.7).

Testing (3.7) with $S_\varepsilon'(\ell_\varepsilon)(\delta\ell)$ and (A.2) with $(\xi_\varepsilon, w_\varepsilon)$ yields

$$(w_\varepsilon, \delta\ell)_{L^2(0,T;L^2(\Omega))} = j'(S_\varepsilon(\ell_\varepsilon))(S_\varepsilon'(\ell_\varepsilon)(\delta\ell)),$$

which inserted in (3.6) gives

$$(w_\varepsilon, \delta\ell)_{L^2(0,T;L^2(\Omega))} + (\ell_\varepsilon, \delta\ell)_{H^1(0,T;L^2(\Omega))} + (\ell_\varepsilon - \bar{\ell}, \delta\ell)_{H^1(0,T;L^2(\Omega))} = 0 \tag{3.8}$$

for all $\delta\ell \in H^1(0,T;L^2(\Omega))$. Further, we observe that

$$\partial_q j(S_\varepsilon(\ell_\varepsilon)) \to \partial_q j(S(\bar{\ell})) \quad \text{in } L^2(0,T;L^2(\Omega)), \tag{3.9a}$$

$$\partial_\varphi j(S_\varepsilon(\ell_\varepsilon)) \to \partial_\varphi j(S(\bar{\ell})) \quad \text{in } L^2(0,T;H^1(\Omega)^*), \tag{3.9b}$$

in the light of (3.2) combined with the continuous Fréchet-differentiability of J (Assumption 3.1). Next we focus on proving uniform bounds for the regularized adjoint states. By employing again a transformation of the variables where this time we abbreviate $\hat{\xi}_\varepsilon := \xi_\varepsilon(T-\cdot)$ and by relying again on the global Lipschitz-continuity of $\max_\varepsilon$ and (3.3), we obtain from (3.7a)

$$\begin{aligned}\|\hat{\xi}_\varepsilon(t)\|_{L^2(\Omega)} &\leq \int_0^t \|\beta\big(w_\varepsilon(T-s, \hat{\xi}_\varepsilon(s)) - \frac{1}{\epsilon} \max{}_\varepsilon{}'(z_\varepsilon(T-s))\hat{\xi}_\varepsilon(s)\big)\|_{L^2(\Omega)}\,ds \\ &\quad + \int_0^t \|[(f_\varepsilon \circ \mathcal{H})'(q_\varepsilon)]^\star \big(\frac{1}{\epsilon} \max{}_\varepsilon{}'(z_\varepsilon)(\hat{\xi}_\varepsilon(T-\cdot))\big)(T-s)\|_{L^2(\Omega)}\,ds \\ &\quad + \int_0^t \|\partial_q j(S_\varepsilon(\ell_\varepsilon))(T-s)\|_{L^2(\Omega)}\,ds \\ &\leq \int_0^t c\,(\|\hat{\xi}_\varepsilon(s)\|_{L^2(\Omega)} + \|\partial_\varphi j(S_\varepsilon(\ell_\varepsilon))(T-s)\|_{H^1(\Omega)^*})\,ds \\ &\quad + \int_0^t \hat{L}_f L_{\mathcal{H}} \underbrace{\int_{T-s}^T \|\hat{\xi}_\varepsilon(T-\zeta)\|_{L^2(\Omega)}\,d\zeta}_{=\int_0^s \|\hat{\xi}_\varepsilon(\zeta)\|_{L^2(\Omega)}\,d\zeta}\ ds \\ &\quad + \int_0^t \|\partial_q j(S_\varepsilon(\ell_\varepsilon))(T-s)\|_{L^2(\Omega)}\,ds \quad \forall\, t \in [0,T].\end{aligned}$$

Now, Gronwall's inequality gives in turn

$$\|\xi_\varepsilon(t)\|_{L^2(\Omega)} \leq \widetilde{c} \int_0^{T-t} \|\partial_\varphi j(S_\varepsilon(\ell_\varepsilon))(T-s)\|_{H^1(\Omega)^*} + \|\partial_q j(S_\varepsilon(\ell_\varepsilon))(T-s)\|_{L^2(\Omega)}\,ds$$

for all $t \in [0,T]$. Thus, by relying on (3.9a)-(3.9b) and by estimating again as above in (3.7a), this time without integrating, one obtains that there exists a constant, independent of ε, such that

$$\|\xi_\varepsilon\|_{H^1(0,T;L^2(\Omega))} \le c.$$

As a consequence,

$$\lambda_\varepsilon := \frac{1}{\epsilon} \max{}_\varepsilon{}'(z_\varepsilon)\xi_\varepsilon$$

and

$$\mu_\varepsilon := f_\varepsilon'(\mathcal{H}(q_\varepsilon))\lambda_\varepsilon$$

are uniformly bounded in $L^\infty(0,T;L^2(\Omega))$ (recall that $\max_\varepsilon$ and f_ε are globally Lipschitz continuous with constants independent of ε). From (3.7b) we can further deduce that there exists a constant $c > 0$, independent of ε, such that $\|w_\varepsilon\|_{L^2(0,T;H^1(\Omega))} \le c$, where we use again (3.9b). Therefore, we can extract weakly convergent subsequences (denoted by the same symbol) so that

$$\begin{aligned} w_\varepsilon \rightharpoonup w \quad &\text{in } L^2(0,T;H^1(\Omega)), \quad \xi_\varepsilon \rightharpoonup \xi \quad \text{in } H^1(0,T;L^2(\Omega)), \\ \lambda_\varepsilon \rightharpoonup^* \lambda, \quad \mu_\varepsilon \rightharpoonup^* \mu \quad &\text{in } L^\infty(0,T;L^2(\Omega)) \quad \text{as } \varepsilon \to 0. \end{aligned} \tag{3.10}$$

Owing to (3.10), (3.9a), (3.9b) and (3.1), we can pass to the limit in (3.7)-(3.8). This results in

$$-\dot{\xi} - \beta\big(w - \lambda\big) + \mathcal{H}'(\bar{q})^\star\mu = \partial_q j(\bar{q},\bar{\varphi}) \ \text{ in } L^2(0,T;L^2(\Omega)), \quad \xi(T) = 0, \tag{3.11a}$$

$$-\alpha\Delta w + \beta\big(w - \lambda\big) = \partial_\varphi j(\bar{q},\bar{\varphi}) \ \text{ in } L^2(0,T;H^1(\Omega)^*), \tag{3.11b}$$

$$(w,\delta\ell)_{L^2(0,T;L^2(\Omega))} + (\bar{\ell},\delta\ell)_{H^1(0,T;L^2(\Omega))} = 0 \quad \forall \delta\ell \in H^1(0,T;L^2(\Omega)), \tag{3.11c}$$

where for the passage to the limit in (3.7a) we also relied on the continuity of the derivative of $\mathcal{H}$ (see Assumption 2.1.1) combined with (3.2).

Now, it remains to prove that (3.5c)-(3.5d) is true. To this end, we show that, for each $\delta > 0$, we have

$$\lambda = \frac{1}{\epsilon} \max{}'(\bar{z})\xi \quad \text{a.e. in } M_\delta, \tag{3.12a}$$

$$\mu = f'(\mathcal{H}(\bar{q}))\lambda \quad \text{a.e. in } \widehat{M}_\delta, \tag{3.12b}$$

where we abbreviate $M_\delta := \{(t,x) : |\bar{z}(t,x)| \ge \delta\}$, $\bar{z} := -\beta(\bar{q} - \bar{\varphi}) - (f \circ \mathcal{H})(\bar{q})$, and $\widehat{M}_\delta := \{(t,x) : |\mathcal{H}(\bar{q})(t,x) - n_f| \ge \delta\}$.

We begin by observing that

$$\|\mathcal{H}(q_\varepsilon)(t) - \mathcal{H}(\bar{q})(t)\|_{L^\infty(\Omega)} \le \widehat{L}_{\mathcal{H}}\|q_\varepsilon - \bar{q}\|_{L^1(0,T;L^\infty(\Omega))} \quad \text{a.e. in } (0,T),$$

in light of Assumption 3.3.1. Thus, as a consequence of (3.4), we have

$$\mathcal{H}(q_\varepsilon) \to \mathcal{H}(\bar{q}) \quad \text{in } L^\infty(0,T;L^\infty(\Omega)), \tag{3.13}$$

which then implies

$$z_\varepsilon \to \bar{z} \quad \text{in } L^\infty((0,T)\times\Omega),$$

by the Lipschitz continuity of f and Lemma 3.6.1. This means that $|z_\varepsilon(t,x)| \ge \delta/2$ f.a.a. $(t,x) \in M_\delta$ for ε small enough, independent of (t,x). In view of the definition of $\max_\varepsilon$ we have

$$\max{}_\varepsilon{}'(z_\varepsilon(\cdot)) = \max{}'(\bar{z}(\cdot)) \quad \text{a.e. in } M_\delta$$

for $\varepsilon \le \delta/2$. The definition of λ_ε and (3.10) now yield (3.12a). To show (3.12b), we proceed in a similar way. Thanks to (3.13), there exists an ε small enough, independent of (t,x), so that $|\mathcal{H}(q_\varepsilon)(t,x) - n_f| \ge \delta/2$ f.a.a. $(t,x) \in \widehat{M}_\delta$. Lemma 3.6.3 applied for $\delta/2$ and

$$K := \max\{\|\mathcal{H}(\bar{q})\|_{L^\infty((0,T)\times\Omega)}, |n_f|\} + \delta$$

then gives in turn the convergence

$$f_\varepsilon'(\mathcal{H}(q_\varepsilon)) - f'(\mathcal{H}(q_\varepsilon)) \to 0 \quad \text{in } L^\infty(\widehat{M}_\delta).$$

Note that, in light of (3.13), it holds $\|\mathcal{H}(q_\varepsilon)\|_{L^\infty((0,T)\times\Omega)} \le \|\mathcal{H}(\bar{q})\|_{L^\infty((0,T)\times\Omega)} + \delta \le K$, for $\varepsilon > 0$ small enough. Since f' is continuous on $[-K, n_f - \delta/2] \cup [n_f + \delta/2, K]$, by Assumption 3.3.2, we deduce in view of (3.13), Assumption 2.1.2 and Lebesgue dominated convergence that

$$f'(\mathcal{H}(q_\varepsilon)) - f'(\mathcal{H}(\bar{q})) \to 0 \quad \text{in } L^2(\widehat{M}_\delta).$$

Finally, the convergence of $\{\lambda_\varepsilon\}$ from (3.10) along with the definition of μ_ε yield that

$$\mu_\varepsilon \rightharpoonup \mu = f'(\mathcal{H}(\bar{q}))\lambda \quad \text{in } L^1(\widehat{M}_\delta),$$

i.e., (3.12b). Since $\delta > 0$ was arbitrary and since $\bigcup_{\delta>0} M_\delta = \{(t,x) : \bar{z}(t,x) \neq 0\}$ and $\bigcup_{\delta>0} \widehat{M}_\delta = \{(t,x) : \mathcal{H}(\bar{q})(t,x) \neq n_f\}$ (up to a set of measure zero), the proof is now complete. □

Remark 3.11. Some remarks concerning (3.5) are in order:

- If $\bar{z}(t,x) \neq 0$ and if $\mathcal{H}(\bar{q})(t,x) \neq n_f$ a.e. in $(0,T)\times\Omega$, then the optimality system in Proposition 3.10 coincides with the very same optimality conditions which one obtains when directly applying the KKT-theory to (P), cf. [37]. Moreover, we observe that (3.5) does not contain any information as to what happens in those (t,x) for which $\bar{z}(t,x)$ and $\mathcal{H}(\bar{q})(t,x)$ are non-smooth points of the mappings max and f, respectively. This is the focus of the next section, where the optimality conditions from Proposition 3.10 shall be improved.
- Indeed, (3.5) is not the best optimality system one could obtain via regularization. Such a system should also contain the relations

  (3.14a) $$\lambda(t,x) \in \frac{1}{\epsilon}\partial \max(0)\xi(t,x) \quad \text{a.e. where } \bar{z}(t,x) = 0,$$

  (3.14b) $$\mu(t,x) \in \partial f(n_f)\lambda(t,x) \quad \text{a.e. where } \mathcal{H}(\bar{q})(t,x) = n_f.$$

  We acknowledge the results [36, Thm. 2.4], [6, Prop. 2.17], [9, Thm. 4.4], where the respective limit optimality systems, though not strong stationary, include such relations between multipliers and adjoint states on the sets where the non-smoothness is active. We cannot expect this to happen in the present paper; by contrast to the aforementioned contributions, our adjoint state $\xi_\varepsilon \in H^1(0,T;L^2(\Omega))$ converges weakly in a space which is not compactly embedded in a Lebesgue space. Although we are able to show

  $$\max{}_\varepsilon{}'(\bar{z}_\varepsilon(\cdot)) \rightharpoonup^* \gamma \in \partial \max(\bar{z}(\cdot)) \text{ in } L^\infty((0,T)\times\Omega),$$

  this does not help us conclude (3.14), in view of the lack of space regularity of the adjoint state.

3.2 Towards strong stationarity

In this section, we aim to derive a stronger optimality system than (3.5). To this end, we will employ arguments from previous works [5, 22], which are entirely based on the limited differentiability properties of the non-smooth mappings involved. We begin by stating the first order necessary optimality conditions in primal form.

Lemma 3.12 (B-stationarity). *If $\bar{\ell} \in H^1(0,T;L^2(\Omega))$ is locally optimal for (P), then there holds*

(3.15) $$j'(S(\bar{\ell}))S'(\bar{\ell};\delta\ell) + (\bar{\ell},\delta\ell)_{H^1(0,T;L^2(\Omega))} \ge 0 \quad \forall\, \delta\ell \in H^1(0,T;L^2(\Omega)).$$

Proof. As a result of Proposition 2.6 and Assumption 3.1 we have that the composite mapping

$$H^1(0,T;L^2(\Omega)) \ni \ell \mapsto J(S(\ell),\ell) \in \mathbb{R}$$

is (Hadamard) directionally differentiable [31, Def. 3.1.1] at $\bar\ell$ in any direction $\delta\ell$ with directional derivative $\partial_{(q,\varphi)}J(S(\bar\ell),\bar\ell)S'(\bar\ell;\delta\ell) + \partial_\ell J(S(\bar\ell),\bar\ell)\delta\ell$; see [31, Lem. 3.1.2(b)] and [32, Prop. 3.6(i)]. The result then follows immediately from the local optimality of $\bar\ell$ and Assumption 3.1. □

In order to improve the optimality conditions from the previous section 3.1, we make use of the following very natural requirement:

Assumption 3.13. The *history operator* $\mathcal{H}$ satisfies the monotonicity condition

$$\mathcal{H}(q_1) \geq \mathcal{H}(q_2) \quad \forall\, q_1, q_2 \in L^2(0,T;L^2(\Omega)) \text{ with } q_1 \geq q_2.$$

Remark 3.14. It is self-evident that the cumulated damage $\mathcal{H}(q)$ (fatigue level of the material) increases as the damage q increases. Hence, the condition in Assumption 3.13 is always satisfied in applications.

As an immediate consequence of Assumption 3.13, we have

$$\mathcal{H}'(q)(\eta) = \lim_{\tau\searrow 0} \frac{\mathcal{H}(q+\tau\eta) - \mathcal{H}(q)}{\tau} \geq 0 \quad \text{a.e. in } (0,T)\times\Omega \tag{3.16}$$

for all $q,\eta \in L^2(0,T;L^2(\Omega))$ with $\eta \geq 0$ a.e. in $(0,T)\times\Omega$.

The main result of this section reads as follows.

Theorem 3.15. *Suppose that Assumptions 3.1, 3.3, and 3.13 are fulfilled. Let $\bar\ell \in H^1(0,T;L^2(\Omega))$ be locally optimal for (P) with associated states*

$$\bar q \in H^1_0(0,T;L^2(\Omega)) \quad \textit{and} \quad \bar\varphi \in L^2(0,T;H^1(\Omega)).$$

Then, there exist adjoint states

$$\xi \in H^1_T(0,T;L^2(\Omega)) \quad \textit{and} \quad w \in L^2(0,T;H^1(\Omega)),$$

and multipliers $\lambda \in L^\infty(0,T;L^2(\Omega))$ and $\mu \in L^\infty(0,T;L^2(\Omega))$ such that the following system is satisfied

$$-\dot\xi - \beta\big(w-\lambda\big) + \mathcal{H}'(\bar q)^\star(\mu) = \partial_q j(\bar q,\bar\varphi) \;\; \textit{in } L^2(0,T;L^2(\Omega)), \quad \xi(T) = 0, \tag{3.17a}$$

$$-\alpha\Delta w + \beta\big(w-\lambda\big) = \partial_\varphi j(\bar q,\bar\varphi) \;\; \textit{in } L^2(0,T;H^1(\Omega)^*), \tag{3.17b}$$

$$\left.\begin{aligned} \lambda(t,x) &= \frac{1}{\epsilon}\chi_{\{\bar z>0\}}(t,x)\xi(t,x) \quad \textit{a.e. where } \bar z(t,x) \neq 0,\\ \mu(t,x) &= f'(\mathcal{H}(\bar q)(t,x))\lambda(t,x) \quad \textit{a.e. where } \mathcal{H}(\bar q)(t,x) \neq n_f, \end{aligned}\right\} \tag{3.17c}$$

$$\left.\begin{aligned} 0 \leq \lambda(t,x) &\leq \frac{1}{\epsilon}(\xi(t,x) + G^+(t,x)) \quad \textit{a.e. where } \bar z(t,x) = 0,\\ G^-(t,x) &\leq 0 \leq G^+(t,x) \quad \textit{a.e. where } \bar z(t,x) > 0, \end{aligned}\right\} \tag{3.17d}$$

$$(w,\delta\ell)_{L^2(0,T;L^2(\Omega))} + (\bar\ell,\delta\ell)_{H^1(0,T;L^2(\Omega))} = 0 \quad \forall \delta\ell \in H^1(0,T;L^2(\Omega)), \tag{3.17e}$$

where we abbreviate $\bar z := -\beta(\bar q - \bar\varphi) - (f\circ\mathcal{H})(\bar q)$. In (3.17d), the mappings $G^+, G^- : [0,T]\times\Omega$ are defined as follows

$$\begin{aligned} G^+(t,x) &:= \int_t^T \mathcal{H}'(\bar q)^\star[\chi_{\{\mathcal{H}(\bar q)=n_f\}}(-\lambda f'_+(n_f) + \mu)](s,x)\, ds,\\ G^-(t,x) &:= \int_t^T \mathcal{H}'(\bar q)^\star[\chi_{\{\mathcal{H}(\bar q)=n_f\}}(-\lambda f'_-(n_f) + \mu)](s,x)\, ds, \end{aligned} \tag{3.18}$$

where, for any $v \in \mathbb{R}$, the right- and left-sided derivative of $f : \mathbb{R} \to \mathbb{R}$ are given by $f'_+(v) := f'(v;1)$ and $f'_-(v) := -f'(v;-1)$, respectively.

Proof. The existence of a tuple $(\xi, w, \lambda, \mu) \in H^1(0,T;L^2(\Omega)) \times L^2(0,T;H^1(\Omega)) \times L^\infty(0,T;L^2(\Omega)) \times L^\infty(0,T;L^2(\Omega))$ satisfying the system (3.17a)-(3.17b)-(3.17c)-(3.17e) is due to Proposition 3.10. Thus, the rest of the proof is focused on showing (3.17d). In this context, we first follow the ideas from [5, Proof of Lem. 2.8] and prove that the set of arguments of $\max{}'(\bar z;\cdot)$ from (2.16a) is dense in $L^2(0,T;L^2(\Omega))$ (step (I)). With this information at hand, we are then able to show the desired result by employing a technique from [5, Proof of Thm. 2.11], see also [22, Proof of Thm. 5.3] (step (II)).

(I) Let $\rho \in L^2(0,T;L^2(\Omega))$ be arbitrary, but fixed. As indicated above, we next show that there exists $\{\delta\ell_n\} \subset H^1(0,T;L^2(\Omega))$ such that

$$\underbrace{-\beta(\delta q_n - \delta\varphi_n) - (f\circ\mathcal{H})'(\bar q;\delta q_n)}_{:=\rho_n} \to \rho \quad \text{in } L^2(0,T;L^2(\Omega)) \quad \text{as } n\to\infty, \tag{3.19}$$

where we abbreviate $(\delta q_n, \delta\varphi_n) := S'(\bar\ell;\delta\ell_n)$ and $\rho_n := -\beta(\delta q_n - \delta\varphi_n) - (f\circ\mathcal{H})'(\bar q)(\delta q_n)$ for all $n \in \mathbb{N}$. To this end, we follow the lines of the proof of [5, Lem. 2.8]. We start by noticing that the mapping

$$[0,T] \ni t \mapsto \hat q(t) \in L^2(\Omega), \quad \hat q(t) := \frac{1}{\epsilon}\int_0^t \max{}'(\bar z(s);\rho(s))\,ds$$

satisfies $\hat q(0) = 0$ and $\hat q \in H^1(0,T;L^2(\Omega))$. Then, we observe that $\hat q$ fulfills

$$\frac{d}{dt}\hat q(t) = \frac{1}{\epsilon}\max{}'\big(\bar z(t); -\beta\hat q(t) - (f\circ\mathcal{H})'(\bar q;\hat q)(t) + \rho(t) + \beta\hat q(t) + (f\circ\mathcal{H})'(\bar q;\hat q)(t)\big) \tag{3.20}$$

a.e. in $(0,T)$. In view of the embedding $H^1(0,T;C_c^\infty(\Omega)) \overset{d}{\hookrightarrow} L^2(0,T;L^2(\Omega))$, there exists a sequence $\{\hat\varphi_n\}_n \subset H^1(0,T;C_c^\infty(\Omega))$ such that

$$\beta\hat\varphi_n \to \rho + \beta\hat q + (f\circ\mathcal{H})'(\bar q;\hat q) \quad \text{in } L^2(0,T;L^2(\Omega)) \ \text{ as } n\to\infty. \tag{3.21}$$

For any $n \in \mathbb{N}$, consider the equation

$$\frac{d}{dt}\hat q_n(t) = \frac{1}{\epsilon}\max{}'\big(\bar z(t); -\beta(\hat q_n - \hat\varphi_n) - (f\circ\mathcal{H})'(\bar q;\hat q_n)\big) \quad \text{a.e. in } (0,T),\ \hat q_n(0) = 0. \tag{3.22}$$

By arguing as in the proof of Lemma 2.6 we see that (3.22) admits a unique solution $\hat q_n \in H_0^1(0,T;L^2(\Omega))$. Now, we define

$$\delta\ell_n := -\alpha\Delta\hat\varphi_n + \beta\big(\hat\varphi_n - \hat q_n\big) \in H^1(0,T;L^2(\Omega)), \tag{3.23}$$

such that $(\hat q_n, \hat\varphi_n)$ solves the system (2.16) associated to $\bar\ell$ with right-hand side $\delta\ell_n \in H^1(0,T;L^2(\Omega))$; note that the regularity of $\delta\ell_n$ in (3.23) is due to the $H^1(0,T;C_c^\infty(\Omega))$-regularity of $\hat\varphi_n$. In view of the unique solvability of (2.16), cf. Proposition 2.6, $(\hat q_n, \hat\varphi_n) = S'(\bar\ell;\delta\ell_n)$. Owing to the Lipschitz-continuity of the directional derivative of max (w.r.t. direction) and (2.15), we further obtain from (3.20) and (3.22)

$$\begin{aligned}
\epsilon\|(\hat q_n - \hat q)(t)\|_{L^2(\Omega)} \le \beta \int_0^t &\|(\hat q - \hat q_n)(s)\|_{L^2(\Omega)}\,ds \\
&+ L_f\,L_{\mathcal{H}}\int_0^t\int_0^s \|(\hat q - \hat q_n)(\zeta)\|_{L^2(\Omega)}\,d\zeta\ ds \\
&+ \int_0^t \|-\beta\hat\varphi_n(s) + \rho(s) + \beta\hat q(s) + (f\circ\mathcal{H})'(\bar q;\hat q)(s)\|_{L^2(\Omega)}\,ds.
\end{aligned}$$

Gronwall's inequality and (3.21) then give in turn

$$\|\hat q_n - \hat q\|_{H^1(0,T;L^2(\Omega))} \le c\,\|-\beta\hat\varphi_n + \rho + \beta\hat q + (f\circ\mathcal{H})'(\bar q;\hat q)\|_{L^2(0,T;L^2(\Omega))} \to 0 \ \text{ as } n\to\infty, \tag{3.24}$$

where $c > 0$ is a constant dependent only on the given data. By relying on the continuity of $(f \circ \mathcal{H})'(\bar{q}; \cdot) : L^2(0,T; L^2(\Omega)) \to L^2(0,T; L^2(\Omega))$, cf. (2.15), we have

$$\beta\hat{q}_n + (f \circ \mathcal{H})'(\bar{q}; \hat{q}_n) \to \beta\hat{q} + (f \circ \mathcal{H})'(\bar{q}; \hat{q}) \quad \text{in } L^2(0,T; L^2(\Omega)) \text{ as } n \to \infty, \tag{3.25}$$

as a result of (3.24). Combining (3.21) and (3.25) finally yields

$$-\beta(\hat{q}_n - \hat{\varphi}_n) - (f \circ \mathcal{H})'(\bar{q}; \hat{q}_n) \to \rho \quad \text{in } L^2(0,T; L^2(\Omega)) \quad \text{as } n \to \infty.$$

Since we established above that $(\hat{q}_n, \hat{\varphi}_n) = S'(\bar{\ell}; \delta\ell_n)$, the proof of this step is now complete.

(II) In the following, $\rho \in L^2(0,T; L^2(\Omega))$ remains arbitrary, but fixed. To prove the desired relations in (3.17d), we first make use of the B-stationarity from (3.15). Here we test with the function $\delta\ell_n \in H^1(0,T; L^2(\Omega))$ which was defined in (3.23).

We test (3.17a), (3.17b), and (3.17e) with $(\delta q_n, \delta\varphi_n) := S'(\bar{\ell}; \delta\ell_n)$ and $\delta\ell_n$, respectively. This leads to

$$\begin{aligned}
0 &\le \partial_q j(\bar{q}, \bar{\varphi})\delta q_n + \partial_\varphi j(\bar{q}, \bar{\varphi})\delta\varphi_n + (\bar{\ell}, \delta\ell_n)_{H^1(0,T;L^2(\Omega))} \\
&= -\int_0^T (\dot{\xi}(t), \delta q_n(t))_{L^2(\Omega)}\, dt - \beta(w - \lambda, \delta q_n)_{L^2(0,T;L^2(\Omega))} \\
&\quad + (\mathcal{H}'(\bar{q})^\star(\mu), \delta q_n)_{L^2(0,T;L^2(\Omega))} + \beta(w - \lambda, \delta\varphi_n)_{L^2(0,T;L^2(\Omega))} \\
&\quad + \alpha(\nabla w, \nabla\delta\varphi_n)_{L^2(0,T;L^2(\Omega))} - (w, \delta\ell_n)_{L^2(0,T;L^2(\Omega))} \\
&= \int_0^T (\xi(t), \dot{\delta q}_n(t))_{L^2(\Omega)}\, dt - (\lambda, -\beta(\delta q_n - \delta\varphi_n))_{L^2(0,T;L^2(\Omega))} \\
&\quad + (\mu, \mathcal{H}'(\bar{q})(\delta q_n))_{L^2(0,T;L^2(\Omega))} - \underbrace{\langle \beta(\delta q_n - \delta\varphi_n) + \alpha\Delta\delta\varphi_n + \delta\ell_n, w\rangle}_{=0, \text{ cf. (2.16b)}}{}_{L^2(0,T;H^1(\Omega))} \\
&\underbrace{=}_{(2.16a)} \int_0^T (\xi(t), \frac{1}{\epsilon}\max{}'(\bar{z}(t); \rho_n(t))_{L^2(\Omega)}\, dt - \int_0^T (\lambda(t), \rho_n(t))_{L^2(\Omega)}\, dt \\
&\quad - \int_0^T (\lambda(t), (f \circ \mathcal{H})'(\bar{q}; \delta q_n)(t))_{L^2(\Omega)}\, dt + \int_0^T (\mu(t), \mathcal{H}'(\bar{q})(\delta q_n)(t))_{L^2(\Omega)}\, dt \quad \forall\, n \in \mathbb{N},
\end{aligned} \tag{3.26}$$

where the second identity follows from integration by parts, $\delta q_n(0) = 0$, and $\xi(T) = 0$; here we also recall the abbreviation $\rho_n := -\beta(\delta q_n - \delta\varphi_n) - (f \circ \mathcal{H})'(\bar{q}; \delta q_n)$, see (3.19). In view of (3.19) and since $\delta q_n(t) = \frac{1}{\epsilon}\int_0^t \max{}'(\bar{z}(s); \rho_n(s))\, ds$, letting $n \to \infty$ in (3.26) leads to

$$\begin{aligned}
0 \le &\int_0^T (\xi(t), \frac{1}{\epsilon}\max{}'(\bar{z}(t); \rho(t))_{L^2(\Omega)}\, dt - \int_0^T (\lambda(t), \rho(t))_{L^2(\Omega)}\, dt \\
&- \int_0^T (\lambda(t), (f \circ \mathcal{H})'(\bar{q}; \widehat{q}_\rho)(t))_{L^2(\Omega)}\, dt + \int_0^T (\mu(t), \mathcal{H}'(\bar{q})(\widehat{q}_\rho)(t))_{L^2(\Omega)}\, dt
\end{aligned} \tag{3.27}$$

for all $\rho \in L^2(0,T; L^2(\Omega))$, where we abbreviate

$$\widehat{q}_\rho(t) := \frac{1}{\epsilon}\int_0^t \max{}'(\bar{z}(s); \rho(s))\, ds \quad \forall\, t \in [0,T]. \tag{3.28}$$

Here we used the fact that $\max{}'(\bar{z}; \cdot) : L^2(0,T; L^2(\Omega)) \to L^2(0,T; L^2(\Omega))$ is continuous, by the Lipschitz-continuity of max, as well as (3.24) in combination with (2.15), and the fact that $\mathcal{H}'(\bar{q})$ belongs to $\mathcal{L}(L^2(0,T; L^2(\Omega)), L^2(0,T; L^2(\Omega)))$.

Next, we take a closer look at the second line in the estimate (3.27). In this context, we first notice that, for all $v, h \in \mathbb{R}$, it holds

$$f'(v; h) = \begin{cases} f'_+(v)h, & \text{if } h \ge 0, \\ f'_-(v)h, & \text{if } h < 0. \end{cases} \tag{3.29}$$

Moreover, we recall that

$$\max{}'(v;h) = \begin{cases} h & \text{if } v > 0, \\ \max\{h,0\} & \text{if } v = 0, \\ 0 & \text{if } v < 0. \end{cases} \tag{3.30}$$

Now, let $\rho \in L^2(0,T;L^2(\Omega))$ with $\rho \geq 0$ a.e. in $(0,T)\times\Omega$ be arbitrary, but fixed. In view of (3.28) and (3.30), we have $\widehat{q}_\rho \geq 0$ a.e. in $(0,T)\times\Omega$ and (3.16) implies

$$\mathcal{H}'(\bar{q})(\widehat{q}_\rho) \geq 0 \quad \text{a.e. in } (0,T)\times\Omega.$$

Then, by recalling (2.14) and by employing Fubini's theorem, we obtain

$$\begin{aligned}
&-\int_0^T (\lambda(t),(f\circ\mathcal{H})'(\bar{q};\widehat{q}_\rho)(t))_{L^2(\Omega)}\,dt + \int_0^T (\mu(t),\mathcal{H}'(\bar{q})(\widehat{q}_\rho)(t))_{L^2(\Omega)}\,dt \\
&\quad= \int_0^T\int_\Omega [-\lambda(t,x)f_+'(\mathcal{H}(\bar{q})(t,x)) + \mu(t,x)]\mathcal{H}'(\bar{q})(\widehat{q}_\rho)(t,x)\,dx\,dt \\
&\quad= \int_0^T\int_\Omega \mathcal{H}'(\bar{q})^\star[-\lambda f_+'(\mathcal{H}(\bar{q})) + \mu](t,x)\widehat{q}_\rho(t,x)\,dx\,dt \\
&\quad\underset{(3.28)}{=} \int_\Omega\int_0^T \mathcal{H}'(\bar{q})^\star[-\lambda f_+'(\mathcal{H}(\bar{q})) + \mu](t,x)\Big(\frac{1}{\epsilon}\int_0^t \max{}'(\bar{z}(s,x);\rho(s,x))\,ds\Big)\,dt\,dx \\
&\quad= \int_\Omega\int_0^T \frac{1}{\epsilon}\max{}'(\bar{z}(t,x);\rho(t,x))\Big(\int_t^T \mathcal{H}'(\bar{q})^\star[-\lambda f_+'(\mathcal{H}(\bar{q})) + \mu](s,x)\,ds\Big)\,dt\,dx \\
&\quad= \int_0^T\int_\Omega \frac{1}{\epsilon}\max{}'(\bar{z}(t,x);\rho(t,x))G^+(t,x)\,dx\,dt \quad \forall\,\rho\in L^2(0,T;L^2(\Omega)),\ \rho\geq 0,
\end{aligned} \tag{3.31}$$

where the last equality is due to the definition of G^+ in (3.18) combined with the second identity in (3.17c). Going back to (3.27), we have

$$\begin{aligned}
0 \leq &\int_0^T\int_\Omega \frac{1}{\epsilon}\max{}'(\bar{z}(t,x);\rho(t,x))\xi(t,x) - \lambda(t,x)\rho(t,x)\,dx\,dt \\
&+ \int_0^T\int_\Omega \frac{1}{\epsilon}\max{}'(\bar{z}(t,x);\rho(t,x))G^+(t,x)\,dx\,dt \quad \forall\,\rho\in L^2(0,T;L^2(\Omega)),\ \rho\geq 0.
\end{aligned} \tag{3.32}$$

By means of the fundamental lemma of calculus of variations in combination with the positive homogeneity of the directional derivative w.r.t. direction, we deduce from (3.32) the inequality

$$\frac{1}{\epsilon}\max{}'(\bar{z}(t,x);1)\xi(t,x) - \lambda(t,x) + \frac{1}{\epsilon}\max{}'(\bar{z}(t,x);1)G^+(t,x) \geq 0 \quad \text{a.e. in } (0,T)\times\Omega. \tag{3.33}$$

By arguing exactly in the same way as above, where one takes into account the fact that $\mathcal{H}'(\bar{q})(\widehat{q}_\rho) \leq 0$ a.e. in $(0,T)\times\Omega$, for $\rho \leq 0$ a.e. in $(0,T)\times\Omega$, we show

$$\begin{aligned}
&-\int_0^T (\lambda(t),(f\circ\mathcal{H})'(\bar{q};\widehat{q}_\rho)(t))_{L^2(\Omega)}\,dt + \int_0^T (\mu(t),\mathcal{H}'(\bar{q})(\widehat{q}_\rho)(t))_{L^2(\Omega)}\,dt \\
&\quad= \int_0^T\int_\Omega \frac{1}{\epsilon}\max{}'(\bar{z}(t,x);\rho(t,x))\underbrace{\Big(\int_t^T \mathcal{H}'(\bar{q})^\star[-\lambda f_-'(\mathcal{H}(\bar{q})) + \mu](s,x)\,ds\Big)}_{=G^-(t,x)}\,dx\,dt \\
&\qquad\qquad\qquad\qquad \forall\,\rho\in L^2(0,T;L^2(\Omega)),\ \rho\leq 0.
\end{aligned} \tag{3.34}$$

This gives in turn

$$\frac{1}{\epsilon}\max{}'(\bar z(t,x);-1)\xi(t,x)+\lambda(t,x)+\frac{1}{\epsilon}\max{}'(\bar z(t,x);-1)G^-(t,x)\geq 0 \tag{3.35}$$

a.e. in $(0,T)\times\Omega$, where we relied again on the fundamental lemma of calculus of variations and the positive homogeneity of the directional derivative w.r.t. direction. From (3.33)- (3.35) and the fact that $\max{}'(0;\cdot)=\max\{\cdot,0\}$ (see (3.30)) we can now conclude the first relation in (3.17d). Finally, the second relation in (3.17d) is a consequence of (3.17c), (3.33)- (3.35) and (3.30). This completes the proof. □

Corollary 3.16 (Strong stationarity in the case that f is smooth). *Suppose that Assumption 3.1 is fulfilled. Let $\bar\ell\in H^1(0,T;L^2(\Omega))$ be locally optimal for (P) with associated states*

$$\bar q\in H^1_0(0,T;L^2(\Omega))\quad\text{and}\quad \bar\varphi\in L^2(0,T;H^1(\Omega)).$$

If the mapping f is differentiable, then there exist unique adjoint states

$$\xi\in H^1_T(0,T;L^2(\Omega))\quad\text{and}\quad w\in L^2(0,T;H^1(\Omega)),$$

and a unique multiplier $\lambda\in L^\infty(0,T;L^2(\Omega))$ such that the following system is satisfied

$$-\dot\xi-\beta\big(w-\lambda\big)+[(f\circ\mathcal{H})'(\bar q)]^\star(\lambda)=\partial_q j(\bar q,\bar\varphi)\ \text{ in } L^2(0,T;L^2(\Omega)),\quad \xi(T)=0, \tag{3.36a}$$

$$-\alpha\Delta w+\beta\big(w-\lambda\big)=\partial_\varphi j(\bar q,\bar\varphi)\quad\text{in } L^2(0,T;H^1(\Omega)^*), \tag{3.36b}$$

$$\left.\begin{aligned}&\lambda(t,x)=\frac{1}{\epsilon}\chi_{\{\bar z>0\}}(t,x)\xi(t,x)\quad\text{a.e. where } \bar z(t,x)\neq 0,\\ &0\leq\lambda(t,x)\leq\frac{1}{\epsilon}\xi(t,x)\quad\text{a.e. where } \bar z(t,x)=0,\end{aligned}\right\} \tag{3.36c}$$

$$(w,\delta\ell)_{L^2(0,T;L^2(\Omega))}+(\bar\ell,\delta\ell)_{H^1(0,T;L^2(\Omega))}=0\quad\forall\delta\ell\in H^1(0,T;L^2(\Omega)), \tag{3.36d}$$

where we abbreviate $\bar z:=-\beta(\bar q-\bar\varphi)-(f\circ\mathcal{H})(\bar q)$. Moreover, (3.36) is of strong stationary type, i.e., if $\bar\ell\in H^1(0,T;L^2(\Omega))$ together with its states $(\bar q,\bar\varphi)\in H^1_0(0,T;L^2(\Omega))\times L^2(0,T;H^1(\Omega))$, some adjoint states $(\xi,w)\in H^1_T(0,T;L^2(\Omega))\times L^2(0,T;H^1(\Omega))$, and a multiplier $\lambda\in L^\infty(0,T;L^2(\Omega))$ satisfy the optimality system (3.36a)–(3.36d), then it also satisfies the variational inequality (3.15).

Proof. The first statement is a consequence of Theorem 3.15. Note that Assumption 3.3 is not required here, as this was employed only in the context of convergence of terms featuring the smoothened function f. As this mapping is already smooth, by assumption, such convergences are no longer needed. Assumption 3.13 is also not necessary here; this was used in the proof of Theorem 3.15 to show (3.31) and (3.34). Since $\{(t,x)\in(0,T)\times\Omega:\mathcal{H}(\bar q)(t,x)=n_f\}$ has measure zero, (3.31) and (3.34) follow immediately from the second relation in (3.17c).

To prove the second assertion, we let $\rho\in L^2(0,T;L^2(\Omega))$ be arbitrary, but fixed and abbreviate $\rho^+:=\max\{\rho,0\}$ and $\rho^-:=\min\{\rho,0\}$. By distinguishing between the sets $\{(t,x)\in(0,T)\times\Omega:\bar z(t,x)>0\}$, $\{(t,x)\in(0,T)\times\Omega:\bar z(t,x)=0\}$ and $\{(t,x)\in(0,T)\times\Omega:\bar z(t,x)<0\}$, we obtain from (3.36c) and (3.30)

$$\begin{aligned}0&\leq\int_0^T\int_\Omega\frac{1}{\epsilon}[\max{}'(\bar z(t,x);\rho^+(t,x))+\max{}'(\bar z(t,x);\rho^-(t,x))]\xi(t,x)\,dx\,dt\\&\qquad-\int_0^T\int_\Omega\lambda(t,x)[\rho^+(t,x)+\rho^-(t,x)]\,dx\,dt\\&=\int_0^T\int_\Omega\frac{1}{\epsilon}\max{}'(\bar z(t,x);\rho(t,x))\xi(t,x)\,dx\,dt-\int_0^T\int_\Omega\lambda(t,x)\rho(t,x)\,dx\,dt\end{aligned} \tag{3.37}$$

for all $\rho \in L^2(0,T;L^2(\Omega))$. Now, let $\delta\ell \in H^1(0,T;L^2(\Omega))$ be arbitrary but fixed and test (3.36a), (3.36b), and (3.36d) with $(\delta q, \delta\varphi) := S'(\bar\ell;\delta\ell)$ and $\delta\ell$, respectively. This leads to

$$\begin{aligned}
&\partial_q j(\bar q,\bar\varphi)\delta q + \partial_\varphi j(\bar q,\bar\varphi)\delta\varphi + (\bar\ell,\delta\ell)_{H^1(0,T;L^2(\Omega))} \\
&= -\int_0^T (\dot\xi(t), \delta q(t))_{L^2(\Omega)}\,dt - \beta(w-\lambda,\delta q)_{L^2(0,T;L^2(\Omega))} \\
&\quad + ([(f\circ\mathcal{H})'(\bar q)]^\star(\lambda), \delta q)_{L^2(0,T;L^2(\Omega))} + \beta(w-\lambda,\delta\varphi)_{L^2(0,T;L^2(\Omega))} \\
&\quad + \alpha(\nabla w, \nabla\delta\varphi)_{L^2(0,T;L^2(\Omega))} - (w,\delta\ell)_{L^2(0,T;L^2(\Omega))} \\
&= \int_0^T (\xi(t), \dot{\delta q}(t))_{L^2(\Omega)}\,dt - (\lambda, -\beta(\delta q-\delta\varphi))_{L^2(0,T;L^2(\Omega))} \\
&\quad + (\lambda, (f\circ\mathcal{H})'(\bar q)(\delta q))_{L^2(0,T;L^2(\Omega))} - \underbrace{\langle \beta(\delta q-\delta\varphi) + \alpha\Delta\delta\varphi + \delta\ell, w\rangle_{L^2(0,T;H^1(\Omega))}}_{=0,\text{ cf. (2.16b)}} \\
&\underbrace{=}_{(2.16a)} \int_0^T \Big(\xi(t), \frac{1}{\epsilon}\max{}'(\bar z(t); (-\beta(\delta q-\delta\varphi) - (f\circ\mathcal{H})'(\bar q;\delta q))(t)\Big)_{L^2(\Omega)}\,dt \\
&\qquad\qquad - \int_0^T (\lambda(t), (-\beta(\delta q-\delta\varphi) - (f\circ\mathcal{H})'(\bar q;\delta q))(t))_{L^2(\Omega)}\,dt \\
&\underset{(3.37)}{\geq} 0,
\end{aligned}$$

where the second identity follows from integration by parts, $\delta q(0) = 0$, and $\xi(T) = 0$. Since $\delta\ell \in H^1(0,T;L^2(\Omega))$ was arbitrary, the proof is now complete. □

Remark 3.17. We remark that if fatigue is not taken into consideration, i.e., if f is replaced by a nonnegative constant, then (3.36) reduces to the strong stationary optimality conditions obtained in [5, Thm. 4.5]; note that therein the control space is $L^2(0,T;L^2(\Omega))$ instead of $H^1(0,T;L^2(\Omega))$.

Remark 3.18. As opposed to (3.36), the optimality system in Theorem 3.15 is not strong stationary, as we will see in the next section. However, we emphasize that (3.17) is a comparatively strong optimality system. While countless non-smooth problems have been addressed by resorting to a smoothening procedure as the one in the proof of Proposition 3.10 (see e.g. [3, 17, 19] and the references therein), we went a step further and improved the optimality conditions from Proposition 3.10 by proving the additional information contained in (3.17d). Let us point out that sign conditions on the sets where the non-smoothness is active, in our case

$$0 \leq \lambda(t,x) \quad \text{a.e. where } \bar z(t,x) = 0$$

are not expected to be obtained by classical regularization techniques, see e.g. [6, Remark 3.9].

3.3 Discussion of the optimality system (3.17). Comparison to strong stationarity

We begin this section by writing down how the strong stationary optimality conditions for the control of (P) should look like.

Proposition 3.19 (An optimality system that implies B-stationarity). *Suppose that Assumption 3.1 is fulfilled. Assume that $\bar\ell \in H^1(0,T;L^2(\Omega))$ together with its states $(\bar q,\bar\varphi) \in H^1_0(0,T;L^2(\Omega)) \times L^2(0,T;H^1(\Omega))$, some adjoint states $(\xi, w) \in H^1_T(0,T;L^2(\Omega)) \times L^2(0,T;H^1(\Omega))$, and some multipliers $\lambda, \mu \in L^\infty(0,T;L^2(\Omega))$*

satisfy the optimality system

$$
\begin{aligned}
&\text{(3.38a)} && -\dot{\xi} - \beta\big(w - \lambda\big) + \mathcal{H}'(\bar{q})^\star(\mu) = \partial_q j(\bar{q}, \bar{\varphi}) \text{ in } L^2(0,T;L^2(\Omega)), \quad \xi(T) = 0,\\
&\text{(3.38b)} && -\alpha\Delta w + \beta\big(w - \lambda\big) = \partial_\varphi j(\bar{q}, \bar{\varphi}) \text{ in } L^2(0,T;H^1(\Omega)^*),\\
&\text{(3.38c)} && \left.\begin{aligned} \lambda(t,x) &= \frac{1}{\epsilon}\chi_{\{\bar{z}>0\}}(t,x)\xi(t,x) \quad \text{a.e. where } \bar{z}(t,x) \neq 0,\\ \mu(t,x) &= f'(\mathcal{H}(\bar{q})(t,x))\lambda(t,x) \quad \text{a.e. where } \mathcal{H}(\bar{q})(t,x) \neq n_f, \end{aligned}\right\}\\
&\text{(3.38d)} && \left.\begin{aligned} 0 \leq \lambda(t,x) &\leq \frac{1}{\epsilon}\xi(t,x) \quad \text{a.e. where } \bar{z}(t,x) = 0,\\ f'_+(n_f)\lambda(t,x) \leq \mu(t,x) &\leq f'_-(n_f)\lambda(t,x) \quad \text{a.e. where } \mathcal{H}(\bar{q})(t,x) = n_f, \end{aligned}\right\}\\
&\text{(3.38e)} && (w, \delta\ell)_{L^2(0,T;L^2(\Omega))} + (\bar{\ell}, \delta\ell)_{H^1(0,T;L^2(\Omega))} = 0 \quad \forall \delta\ell \in H^1(0,T;L^2(\Omega)),
\end{aligned}
$$

where we abbreviate $\bar{z} := -\beta(\bar{q} - \bar{\varphi}) - (f \circ \mathcal{H})(\bar{q})$ *and where, for any* $v \in \mathbb{R}$, *the right- and left-sided derivative of* $f : \mathbb{R} \to \mathbb{R}$ *are given by* $f'_+(v) := f'(v;1)$ *and* $f'_-(v) := -f'(v;-1)$, *respectively. Then,* $\bar{\ell}$ *also satisfies the variational inequality* (3.15).

Proof. Let $\rho \in L^2(0,T;L^2(\Omega))$ be arbitrary, but fixed. In the proof of Corollary 3.16 we saw that the first identity in (3.38c) and the first relation in (3.38d) combined with (3.30) imply

$$
\text{(3.39)} \qquad 0 \leq \int_0^T \int_\Omega \frac{1}{\epsilon} \max{}'(\bar{z}(t,x); \rho(t,x))\xi(t,x)\, dx\, dt - \int_0^T \int_\Omega \lambda(t,x)\rho(t,x)\, dx\, dt
$$

for all $\rho \in L^2(0,T;L^2(\Omega))$. Next we abbreviate $\mathcal{H}'(\bar{q})(\widehat{q}_\rho)^- := \min\{\mathcal{H}'(\bar{q})(\widehat{q}_\rho), 0\}$ and $\mathcal{H}'(\bar{q})(\widehat{q}_\rho)^+ := \max\{\mathcal{H}'(\bar{q})(\widehat{q}_\rho), 0\}$, where

$$
\widehat{q}_\rho(t) := \frac{1}{\epsilon}\int_0^t \max{}'(\bar{z}(s); \rho(s))\, ds \quad \forall\, t \in [0,T].
$$

From the second identity in (3.38c) and the second relation in (3.38d) we deduce that

$$
\text{(3.40)} \qquad
\begin{aligned}
0 &\leq \int_0^T \int_\Omega \underbrace{[-\lambda(t,x)f'_+(\mathcal{H}(\bar{q})(t,x)) + \mu(t,x)]}_{\geq 0} \mathcal{H}'(\bar{q})(\widehat{q}_\rho)^+(t,x)\, dx\, dt\\
&\quad + \int_0^T \int_\Omega \underbrace{[-\lambda(t,x)f'_-(\mathcal{H}(\bar{q})(t,x)) + \mu(t,x)]}_{\leq 0} \mathcal{H}'(\bar{q})(\widehat{q}_\rho)^-(t,x)\, dx\, dt\\
&= \int_0^T \int_\Omega -\lambda(t,x)f'(\mathcal{H}(\bar{q})(t,x); \mathcal{H}'(\bar{q})(\widehat{q}_\rho)^+(t,x)) + \mu(t,x)\mathcal{H}'(\bar{q})(\widehat{q}_\rho)^+(t,x)\, dx\, dt\\
&\quad + \int_0^T \int_\Omega -\lambda(t,x)f'(\mathcal{H}(\bar{q})(t,x); \mathcal{H}'(\bar{q})(\widehat{q}_\rho)^-(t,x)) + \mu(t,x)\mathcal{H}'(\bar{q})(\widehat{q}_\rho)^-(t,x)\, dx\, dt\\
&= -\int_0^T (\lambda(t), (f \circ \mathcal{H})'(\bar{q}; \widehat{q}_\rho)(t))_{L^2(\Omega)}\, dt + \int_0^T (\mu(t), \mathcal{H}'(\bar{q})(\widehat{q}_\rho)(t))_{L^2(\Omega)}\, dt,
\end{aligned}
$$

where in the second identity we relied on (3.29). Adding (3.39) and (3.40) yields (3.27). Now, let $\delta\ell \in H^1(0,T;L^2(\Omega))$ be arbitrary but fixed and abbreviate $(\delta q, \delta\varphi) := S'(\bar{\ell}; \delta\ell)$. By testing (3.27) with $-\beta(\delta q - \delta\varphi) - (f \circ \mathcal{H})'(\bar{q}; \delta q)$ and by arguing step by step backwards as in the proof of (3.26), we finally arrive at the desired result. ☐

Remark 3.20. Some words concerning Proposition 3.19 are in order:

- The optimality system (3.38) differs from (3.17) only regarding the relations in (3.38d) and (3.17d). As expected, the optimality conditions in (3.38d) contain more information than (3.17d). This is also confirmed by Proposition 3.21 below.

- We point out that (3.38) is not of strong stationary type, as we were not able to show (3.15) $\Rightarrow$ (3.38); the optimality conditions in (3.38) just point out the information that is missing in (3.17), namely

$$\begin{aligned} \lambda(t,x) &\leq \frac{1}{\epsilon}\xi(t,x) \quad \text{a.e. where } \bar{z}(t,x)=0, \\ f_+'(n_f)\lambda(t,x) \leq \mu(t,x) &\leq f_-'(n_f)\lambda(t,x) \quad \text{a.e. where } \mathcal{H}(\bar{q})(t,x)=n_f. \end{aligned} \tag{3.41}$$

  Note that the sign condition

$$0 \leq \lambda(t,x) \quad \text{a.e. where } \bar{z}(t,x)=0$$

  is already contained in (3.17d). The proof of Proposition 3.19 shows that (3.41) is indeed needed for the implication (3.17) $\Rightarrow$ (3.15).

- In order to prove that a certain optimality system implies B-stationarity, it is essential that it includes sign conditions for the involved multipliers and/or adjoint states on the sets where the non-smoothness is active. This fact has been observed in many contributions dealing with strong stationarity [22, Rem. 6.9], [6, Rem. 3.9], [5, Rem. 4.8], [9, Rem. 4.15]. In our case, see (3.17d), the information on $\{\bar{z}=0\}$ is incomplete, while the sign conditions on the set $\{\mathcal{H}(\bar{q})=n_f\}$ are non-existent and seem to be hidden in the integral formulations (3.18).

Proposition 3.21 (The optimality system (3.38) is stronger than (3.17)). *Suppose that all the hypotheses in Proposition 3.19 are fulfilled. If, in addition, Assumption 3.13 holds true, then (3.17) is satisfied.*

Proof. We only need to show that (3.38d) implies (3.17d). To this end, we first prove that

$$\mathcal{H}'(\bar{q})^\star(\eta_1) \geq \mathcal{H}'(\bar{q})^\star(\eta_2) \quad \forall\, \eta_1,\eta_2 \in L^2(0,T;L^2(\Omega)) \text{ with } \eta_1 \geq \eta_2. \tag{3.42}$$

We recall that, as a consequence of Assumption 3.13, $\mathcal{H}'(\bar{q})(\rho) \geq 0$ for all $\rho \in L^2(0,T;L^2(\Omega))$, $\rho \geq 0$, cf. (3.16). This leads to

$$\begin{aligned} (\mathcal{H}'(\bar{q})^\star(\eta_1),\rho)_{L^2(0,T;L^2(\Omega))} &= (\eta_1,\mathcal{H}'(\bar{q})(\rho))_{L^2(0,T;L^2(\Omega))} \\ &\geq (\eta_2,\mathcal{H}'(\bar{q})(\rho))_{L^2(0,T;L^2(\Omega))} = (\mathcal{H}'(\bar{q})^\star(\eta_2),\rho)_{L^2(0,T;L^2(\Omega))}, \end{aligned}$$

from which (3.42) follows. Now, the second relation in (3.38d) and the definitions of G^+ and G^- in (3.18) give in turn

$$G^+ \geq 0 \text{ and } G^- \leq 0 \quad \text{a.e. in } (0,T)\times\Omega.$$

Thus, (3.38d) implies (3.17d) and the proof is complete. □

Remark 3.22. The gap between (3.17) and the strong stationary optimality conditions (3.38) is due to the additional non-smooth mapping f appearing in the argument of the initial non-smoothness max, cf. (2.2a). To see this, let us take a closer look at the proof of Theorem 3.15. Therein, (3.17d) is proven by relying on direct methods from previous works [5, 22] which deal with strong stationarity in the context of one non-differentiable map. In these findings it has been observed that the set of directions into which the non-smoothness is differentiated - in the "linearized" state equation - must be dense in a suitable (Bochner) space [5, Remark 2.12], [22, Lem. 5.2]. The density of the set of directions into which max is differentiated, see (2.16a), is indeed available, as the first step of the proof of Theorem

3.15 shows. This allowed us to improve the optimality system (3.5) from the previous section. However, the non-differentiable function f requires a similar density property too, which reads as follows

$$\{\mathcal{H}'(\bar{q}; S_1'(\bar{\ell}; \delta\ell)) : \delta\ell \in H^1(0,T;L^2(\Omega))\} \overset{d}{\hookrightarrow} L^2(0,T;L^2(\Omega)), \tag{3.43}$$

where S_1 denotes the first component of the control-to-state map $S : L^2(0,T;H^1(\Omega)^*) \ni \ell \mapsto (q,\varphi) \in H_0^1(0,T;L^2(\Omega)) \times L^2(0,T;H^1(\Omega))$. By taking a look at the "linearized" state equation (2.16a), we see that (3.43) is not to be expected, due to the lack of surjectivity of the mapping $\max'(\bar{z};\cdot)$. Thus, the methods from [5, 22] are restricted to one non-smoothness and permit us to improve the limit optimality system (3.5) only up to a certain point. Thus, the strong stationarity for the control of (P) remains an open question.

APPENDIX A

Proof of Lemma 3.7.

The arguments are well-known [4] and can be found in [5, App. B] for the case that $(f \circ \mathcal{H})(q)$ is constant and the control space is $L^2(0,T;L^2(\Omega))$ instead of $H^1(0,T;L^2(\Omega))$.

(I) Let $\varepsilon > 0$ be arbitrary, but fixed. We begin by recalling the smooth state equation appearing in (P_ε):

$$\dot{q}(t) = \frac{1}{\epsilon} \max{}_\varepsilon(-\beta(q(t) - \varphi(t)) - (f_\varepsilon \circ \mathcal{H})(q)(t)) \ \text{ in } L^2(\Omega), \quad q(0) = 0, \tag{A.1a}$$

$$-\alpha\Delta\varphi(t) + \beta\,\varphi(t) = \beta q(t) + \ell(t) \quad \text{in } H^1(\Omega)^*, \quad \text{a.e. in } (0,T). \tag{A.1b}$$

By employing the exact same arguments as in the proof of Proposition 2.3, one infers that (A.1) admits a unique solution $(q_\varepsilon, \varphi_\varepsilon) \in H_0^1(0,T;L^2(\Omega)) \times L^2(0,T;H^1(\Omega))$ for every $\ell \in L^2(0,T;H^1(\Omega)^*)$, which allows us to define the regularized solution mapping

$$S_\varepsilon : L^2(0,T;H^1(\Omega)^*) \ni \ell \mapsto (q_\varepsilon, \varphi_\varepsilon) \in H_0^1(0,T;L^2(\Omega)) \times L^2(0,T;H^1(\Omega)).$$

The operator S_ε is Gâteaux-differentiable and its derivative at $\ell \in L^2(0,T;H^1(\Omega)^*)$ in direction $\delta\ell \in L^2(0,T;H^1(\Omega)^*)$, i.e., $(\delta q, \delta\varphi) := S_\varepsilon'(\ell)(\delta\ell)$, is the unique solution of

$$\begin{aligned} \dot{\delta q}(t) &= \frac{1}{\epsilon} \max{}_\varepsilon'(z_\varepsilon(t))\big(-\beta(\delta q(t) - \delta\varphi(t)) - (f_\varepsilon \circ \mathcal{H})'(q_\varepsilon)(\delta q)(t)\big) \ \text{ in } L^2(\Omega), \quad \delta q(0) = 0, \\ -\alpha\Delta\delta\varphi(t) &+ \beta\,\delta\varphi(t) = \beta\delta q(t) + \delta\ell(t) \quad \text{in } H^1(\Omega)^*, \quad \text{a.e. in } (0,T), \end{aligned} \tag{A.2}$$

where we abbreviate $z_\varepsilon := -\beta(q_\varepsilon - \varphi_\varepsilon) - (f_\varepsilon \circ \mathcal{H})(q_\varepsilon)$. By arguing as in the proof of Lemma 2.4 we deduce that $S_\varepsilon : L^2(0,T;H^1(\Omega)^*) \to H_0^1(0,T;L^2(\Omega)) \times L^2(0,T;H^1(\Omega))$ is Lipschitz continuous (with constant independent of ε). Moreover, we have the convergence

$$S_\varepsilon(\ell_\varepsilon) \to S(\ell) \quad \text{in } H_0^1(0,T;L^2(\Omega)) \times L^2(0,T;H^1(\Omega)), \tag{A.3}$$

for $\ell_\varepsilon \to \ell$ in $L^2(0,T;H^1(\Omega)^*)$. To see this, one first shows that $S_\varepsilon(\ell) \to S(\ell)$, which follows by estimating as in the proof of Lemma 2.4 and by using (2.1) applied for f_ε along with Lemma 3.6.1. Then, (A.3) is a consequence of the Lipschitz continuity of S_ε (with constant independent of ε).

(II) Next, we focus on proving that $\bar{\ell}$ can be approximated via local minimizers of optimal control problems governed by (A.1). To this end, let $B_{H^1(0,T;L^2(\Omega))}(\bar{\ell}, \rho)$ be the ball of local optimality of $\bar{\ell}$ and consider the smooth (reduced) optimal control problem

$$\left.\begin{aligned} \min_{\ell \in H^1(0,T;L^2(\Omega))} \quad & J(S_\varepsilon(\ell), \ell) + \frac{1}{2}\|\ell - \bar{\ell}\|^2_{H^1(0,T;L^2(\Omega))} \\ \text{s.t.} \quad & \ell \in B_{H^1(0,T;L^2(\Omega))}(\bar{\ell}, \rho). \end{aligned}\right\} \tag{P_ε^ρ}$$

By arguing as in the proof of Proposition 3.2, we see that (P_ε^ρ) admits a global solution $\ell_\varepsilon \in H^1(0,T;L^2(\Omega))$. Since $\ell_\varepsilon \in B_{H^1(0,T;L^2(\Omega))}(\bar{\ell},\rho)$, we can select a subsequence with

$$\ell_\varepsilon \rightharpoonup \widetilde{\ell} \quad \text{in } H^1(0,T;L^2(\Omega)), \tag{A.4}$$

where $\widetilde{\ell} \in B_{H^1(0,T;L^2(\Omega))}(\bar{\ell},\rho)$. For simplicity, we abbreviate in the following

$$\mathcal{J}(\ell) := J(S(\ell),\ell), \tag{A.5a}$$

$$\mathcal{J}_\varepsilon(\ell) := J(S_\varepsilon(\ell),\ell) + \frac{1}{2}\|\ell - \bar{\ell}\|^2_{H^1(0,T;L^2(\Omega))} \tag{A.5b}$$

for all $\ell \in H^1(0,T;L^2(\Omega))$. Due to (A.3) and Assumption 3.1, it holds

$$\mathcal{J}(\bar{\ell}) \overset{\text{(A.5a)}}{=} J(S(\bar{\ell}),\bar{\ell}) = \lim_{\varepsilon\to 0} J(S_\varepsilon(\bar{\ell}),\bar{\ell}) \overset{\text{(A.5b)}}{=} \lim_{\varepsilon\to 0}\mathcal{J}_\varepsilon(\bar{\ell}) \ge \limsup_{\varepsilon\to 0}\mathcal{J}_\varepsilon(\ell_\varepsilon), \tag{A.6}$$

where for the last inequality we relied on the fact that ℓ_ε is a global minimizer of (P_ε^ρ) and that $\bar{\ell}$ is admissible for (P_ε^ρ). In view of (A.5b), (A.6) can be continued as

$$\begin{aligned}\mathcal{J}(\bar{\ell}) &\ge \limsup_{\varepsilon\to 0} J(S_\varepsilon(\ell_\varepsilon),\ell_\varepsilon) + \frac{1}{2}\|\ell_\varepsilon - \bar{\ell}\|^2_{H^1(0,T;L^2(\Omega))}\\ &\ge \liminf_{\varepsilon\to 0} J(S_\varepsilon(\ell_\varepsilon),\ell_\varepsilon) + \frac{1}{2}\|\ell_\varepsilon - \bar{\ell}\|^2_{H^1(0,T;L^2(\Omega))}\\ &\ge J(S(\widetilde{\ell}),\widetilde{\ell}) + \frac{1}{2}\|\widetilde{\ell} - \bar{\ell}\|^2_{H^1(0,T;L^2(\Omega))} \ge \mathcal{J}(\bar{\ell}),\end{aligned} \tag{A.7}$$

where we used again (A.3), as well as the compact embedding $H^1(0,T;L^2(\Omega)) \hookrightarrow\hookrightarrow L^2(0,T;H^1(\Omega)^*)$, and the continuity of j, see Assumption 3.1; note that for the last inequality in (A.7) we employed the fact that $\widetilde{\ell} \in B_{H^1(0,T;L^2(\Omega))}(\bar{\ell},\rho)$. From (A.7) we obtain that $\widetilde{\ell} = \bar{\ell}$ and

$$\mathcal{J}(\bar{\ell}) = \lim_{\varepsilon\to 0} J(S_\varepsilon(\ell_\varepsilon),\ell_\varepsilon) + \frac{1}{2}\|\ell_\varepsilon - \bar{\ell}\|^2_{H^1(0,T;L^2(\Omega))} = J(S(\widetilde{\ell}),\widetilde{\ell}) + \frac{1}{2}\|\widetilde{\ell} - \bar{\ell}\|^2_{H^1(0,T;L^2(\Omega))}.$$

Since $J(S_\varepsilon(\ell_\varepsilon),\ell_\varepsilon) \to J(S(\widetilde{\ell}),\widetilde{\ell})$, one has the convergence

$$\ell_\varepsilon \to \bar{\ell} \quad \text{in } H^1(0,T;L^2(\Omega)), \tag{A.8}$$

where we also relied on (A.4). As a consequence, (A.3) yields

$$S_\varepsilon(\ell_\varepsilon) \to S(\bar{\ell}) \quad \text{in } H^1_0(0,T;L^2(\Omega)) \times L^2(0,T;H^1(\Omega)). \tag{A.9}$$

A classical argument finally shows that ℓ_ε is a local minimizer of $\min_{\ell\in H^1(0,T;L^2(\Omega))}\mathcal{J}_\varepsilon(\ell)$ for $\varepsilon > 0$ sufficiently small.

REFERENCES

[1] R. Alessi, V. Crismale, and G. Orlando, Fatigue effects in elastic materials with variational damage models: A vanishing viscosity approach, *Journal of Nonlinear Science* 29 (2019), 1041–1094.

[2] R. Alessi, S. Vidoli, and L. De Lorenzis, Variational approach to fatigue phenomena with a phase-field model: the one-dimensional case, *Engineering Fracture Mechanics* 190 (2018), 53–73.

[3] V. Barbu, Necessary conditions for distributed control problems governed by parabolic variational inequalities, *SIAM Journal on Control and Optimization* 19 (1981), 64–86.

[4] V. Barbu, *Optimal Control of Variational Inequalities*, Research notes in mathematics 100, Pitman, Boston-London-Melbourne, 1984.

[5] L. Betz, Strong stationarity for optimal control of a non-smooth coupled system: Application to a viscous evolutionary VI coupled with an elliptic PDE, *SIAM J. on Optimization* 29 (2019), 3069–3099.

[6] L. Betz, Strong stationarity for a highly nonsmooth optimization problem with control constraints, *Mathematical Control and Related Fields* 13 (2023), 1500–1528.

[7] C. Christof, Sensitivity analysis and optimal control of obstacle-type evolution variational inequalities, *SIAM J. Control Optim.* 57 (2019), 192–218.

[8] C. Christof and M. Brokate, Strong stationarity conditions for optimal control problems governed by a rate-independent evolution variational inequality, *SIAM Journal on Control and Optimization* 61 (2023), 2222–2250.

[9] C. Christof, C. Clason, C. Meyer, and S. Walther, Optimal control of a non-smooth, semilinear elliptic equation, *Mathematical Control and Related Fields* 8 (2018), 247–276.

[10] C. Clason, V. Nhu, and A. Rösch, Optimal control of a non-smooth quasilinear elliptic equation, *Mathematical Control and Related Fields* 11 (2021), 521–554.

[11] V. Crismale, G. Lazzaroni, and G. Orlando, Cohesive fracture with irreversibility: quasistatic evolution for a model subject to fatigue, *Math. Models Methods Appl. Sci.* 28 (2018), 1371–1412.

[12] J. C. De los Reyes and C. Meyer, Strong stationarity conditions for a class of optimization problems governed by variational inequalities of the second kind, *Journal of Optimization Theory and Applications* 168 (2016), 375–409.

[13] B. Dimitrijevic and K. Hackl, A method for gradient enhancement of continuum damage models, *Technische Mechanik* 28 (2008), 43–52.

[14] E. Emmrich, *Gewöhnliche und Operator Differentialgleichungen*, Vieweg, Wiesbaden, 2004.

[15] M. Frémond and N. Kenmochi, Damage problems for viscous locking materials, *Adv. Math. Sci. Appl.* 16 (2006), 697–716.

[16] M. Frémond and B. Nedjar, Damage, gradient of damage and principle of virtual power, *Int. J. Solids Struct.* 33 (1996), 1083–1103.

[17] A. Friedman, Optimal control for parabolic variational inequalities, *SIAM Journal on Control and Optimization* 25 (1987), 482–497.

[18] K. Hammerum, P. Brath, and N. K. Poulsen, A fatigue approach to wind turbine control, *J. of Physics: Conf. Series* 75 (2007), 012–081.

[19] Z. He, State constrained control problems governed by variational inequalities, *SIAM Journal on Control and Optimization* 25 (1987), 1119–1144.

[20] M. Hintermüller and I. Kopacka, Mathematical programs with complementarity constraints in function space: C- and strong stationarity and a path-following algorithm, *SIAM Journal on Optimization* 20 (2009), 868–902.

[21] D. Knees, R. Rossi, and C. Zanini, A vanishing viscosity approach to a rate-independent damage model, *Mathematical Models and Methods in Applied Sciences* 23 (2013), 565–616.

[22] C. Meyer and L. Susu, Optimal control of nonsmooth, semilinear parabolic equations, *SIAM Journal on Control and Optimization* 55 (2017), 2206–2234.

[23] C. Meyer and L. Susu, Analysis of a viscous two-field gradient damage model. Part I: Existence and uniqueness, *Z. Anal. Anwend.* 38 (2019), 249–286.

[24] C. Meyer and L. Susu, Analysis of a viscous two-field gradient damage model. Part II: Penalization limit, *Z. Anal. Anwend.* 38 (2019), 439–474.

[25] F. Mignot, Contrôle dans les inéquations variationelles elliptiques, *Journal of Functional Analysis* 22 (1976), 130–185.

[26] F. Mignot and J. P. Puel, Optimal control in some variational inequalities, *SIAM Journal on Control and Optimization* 22 (1984), 466–476.

[27] I. Munteanu, A. I. Bratcu, N. A. Cutululis, and E. Ceanga, *Optimal Control of Wind Energy Systems: Towards a Global Approach*, Springer Science & Business Media, 2008.

[28] R. Ritchie and M. Launey, Crack growth in brittle and ductile solids, in *Wang Q.J., Chung YW. (eds) Encyclopedia of Tribology*, Springer, Boston, 2013.

[29] R. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, 1970.

[30] H. Scheel and S. Scholtes, Mathematical programs with complementarity constraints: stationarity, optimality, and sensitivity, *Mathematics of Operations Research* 25 (2000), 1–22.

[31] W. Schirotzek, *Nonsmooth Analysis*, Springer, Berlin, 2007.

[32] A. Shapiro, On concepts of directional differentiability, *Journal of Optimization Theory and Applications* 66 (1990), 477–487.

[33] M. Sofonea and A. Matei, *Variational Inequalities with Applications*, Springer, New York, 2009.

[34] R. Stephens, A. Fatemi, R. Stephens, and H. Fuchs, *Metal Fatigue in Engineering*, A Wiley-Interscience publication, Wiley, New York, 2000.

[35] L. Susu, Optimal control of a viscous two-field gradient damage model, *GAMM-Mitt.* 40 (2018), 287 – 311.

[36] D. Tiba, *Optimal Control of Nonsmooth Distributed Parameter Systems*, Springer, 1990.

[37] F. Tröltzsch, *Optimal Control of Partial Differential Equations*, volume 112 of Graduate studies in mathematics, American Mathematical Society, Providence, 2010. Theory, methods and applications, Translated from the 2005 German original by Jürgen Sprekels.

[38] G. Wachsmuth, Strong stationarity for optimal control of the obstacle problem with control constraints, *SIAM Journal on Optimization* 24 (2014), 1914–1932.

[39] G. Wachsmuth, Elliptic quasi-variational inequalities under a smallness assumption: uniqueness, differential stability and optimal control, *Calc. Var.* 82 (2020).