SIAM J. CONTROL OPTIM.
Vol. 50, No. 1, pp. 388–418

DISTRIBUTED OPTIMAL CONTROL OF THE CAHN–HILLIARD SYSTEM INCLUDING THE CASE OF A DOUBLE-OBSTACLE HOMOGENEOUS FREE ENERGY DENSITY*

M. HINTERMÜLLER† AND D. WEGNER‡

Abstract. In this paper we study the distributed optimal control for the Cahn–Hilliard system. A general class of free energy potentials is allowed which, in particular, includes the double-obstacle potential. The latter potential yields an optimal control problem of a parabolic variational inequality which is of fourth order in space. We show the existence of optimal controls to approximating problems where the potential is replaced by a mollified version of its Moreau–Yosida approximation. Corresponding first-order optimality conditions for the mollified problems are given. For this purpose a new result on the continuous Fréchet differentiability of superposition operators with values in Sobolev spaces is established. Besides the convergence of optimal controls of the mollified problems to an optimal control of the original problem, we also derive first-order optimality conditions for the original problem by a limit process. The newly derived stationarity system corresponds to a function space version of C-stationarity.

Key words. Cahn–Hilliard system, double-obstacle potential, mathematical programming with equilibrium constraints, distributed optimal control, Yosida regularization, C-stationarity

AMS subject classifications. 35K86, 49J20, 34G25

DOI. 10.1137/110824152

1. Introduction.

In the study of interface dynamics phase field models received a considerable amount of attention in the recent past; see, e.g., [58] for a review on phase field models in material science. Such models have also been applied successfully in fluid dynamics [7], image processing [15, 22], and cancer growth modeling [28].

In [21] Cahn and Hilliard introduced a continuous model for phase transitions in systems of nonuniform compositions capturing spinodal decomposition. Based on the minimization of a Ginzburg–Landau type free energy, the resulting system is of parabolic type with a fourth-order (partial) differential operator in space which describes the evolution of a local phase variable y. Within a spatial domain Ω, the latter is required to take values in $[-1, 1]$ where, for $x \in \Omega$, $y(x) = 1$ represents one of two phases and $y(x) = -1$ the other, respectively. For $-1 < y(x) < 1$ the composition is in a mixed state at $x \in \Omega$. Utilizing a chemical potential w and assuming constant mobility (normalized to 1), the associated mathematical model reads

$$(1.1) \qquad \begin{aligned} &y_t - \Delta w = 0, \quad && w + \gamma\Delta y \in \partial\Psi(y) \quad && \text{in } \Omega, \\ &\nabla w \cdot \vec{n} = 0, && \nabla y \cdot \vec{n} = 0 && \text{on } \partial\Omega \end{aligned}$$

*Received by the editors February 10, 2011; accepted for publication (in revised form) November 10, 2011; published electronically February 2, 2012. The authors gratefully acknowledge support by DFG Research Center MATHEON under project C28. M.H. further acknowledges support through FWF under START-program Y305-N18 "Interfaces and Free Boundaries," SFB F3204-N18 "Mathematical Optimization and Its Applications in Biomedical Sciences," and the DFG SPP 1253 "Optimization with Partial Differential Equations."

http://www.siam.org/journals/sicon/50-1/82415.html

†Department of Mathematics, Humboldt-Universität zu Berlin, D-10099 Berlin, Germany and Department of Mathematics and Scientific Computing, Karl-Franzens-University of Graz, Graz, Austria (hint@math.hu-berlin.de).

‡Department of Mathematics, Humboldt-Universität zu Berlin, D-10099 Berlin, Germany (dwegner@math.hu-berlin.de).

with appropriate initial conditions. Here, $\gamma > 0$ denotes a given parameter, Ψ represents the homogeneous free energy density contained in the Ginzburg–Landau model, and $\partial\Psi$ stands for the generalized derivative from nonsmooth analysis [23]. Note that the latter is single-valued whenever Ψ is differentiable at y. In this case $\partial\Psi(y) = \{\Psi'(y)\}$ with $\Psi'(y)$ the derivative of Ψ at y.

Depending on the application context, different choices of Ψ have been investigated in the literature. Typically, the various versions of Ψ aim at confining the values of y to $[-1, 1]$ or $(-1, 1)$. In this context, a widely studied choice is the double-well potential [25, 27, 54]. Another choice is of logarithmic form and goes back to Cahn and Hilliard's original work [21]; see also [2]. Logarithmic forms of the free energy density are also important in the Flory–Huggins solution theory of the thermodynamics of polymer solutions. While the double-well type free energy allows violations of $y(x) \in [-1, 1]$, the logarithmic potential does not. Both choices, however, share certain differentiability properties such that $\partial\Psi$ becomes single-valued and the second equation in (1.1) becomes an equality with the derivative Ψ' on the right-hand side. On the other hand, in [55] Oono and Puri found that in the case of deep quenches of, e.g., binary alloys, the so-called double-obstacle potential

$$\Psi(y)(x) = \begin{cases} \frac{1}{2}(1 - y(x))^2 & \text{if } |y(x)| \le 1, \\ +\infty & \text{if } |y(x)| > 1, \end{cases} \tag{1.2}$$

$x \in \Omega$, is better suited than the other free energy models mentioned above. A similar observation appears to be true in the case of polymeric membrane formation under rapid wall hardening. For this choice of the free energy, due to the nondifferentiability of the associated function Ψ the second relation in the system (1.1) is indeed a variational inclusion or, equivalently, a variational inequality. For the resulting Cahn–Hilliard system, a comprehensive mathematical analysis can be found in [17, 18]. Concerning numerical solvers we refer the reader to [8, 9, 16, 33, 34, 40] and the references therein.

In many applications, it might be interesting to influence the phase transition in such a way that a prespecified control goal is achieved. In this direction, feedback stabilization, as well as optimal control for the Cahn–Hilliard equation with a double-well type homogeneous free energy density, is studied theoretically in [62]. For a polynomial-type free energy density, in [60] a first-order optimality system is derived for minimizing a tracking-type objective subject to the associated Cahn–Hilliard equation. With the goal of preventing spinodal decomposition in Fe-Al alloys, in [32] the Cahn–Hilliard system with double-well type free energy is controlled near a steady-state of the system. In some applications one might be interested in governing the Cahn–Hilliard system from an initial phase distribution (often a homogeneous mixture) y_0 to some desired phase pattern y_T at a given (final) time T. For the Cahn–Hilliard system with a double-obstacle type homogeneous free energy density, such a problem formulation was mentioned by Garcke in [31]. For instance, in the context of polymeric membrane formation, y_T may describe a desired porosity patter which implies filtration or other membrane qualities.

In this paper we pick up the latter perspective and study the minimization of an objective of the type

$$J(y, u) = \frac{\mu_1}{2}\int_0^T\!\!\int_\Omega |y(x,t) - y_\Omega|^2 dx\, dt + \frac{\mu_2}{2}\int_\Omega |y(x,T) - y_T|^2 dx + \frac{1}{2}\int_0^T\!\!\int_\Omega |u(x,t)|^2 dx\, dt \tag{1.3}$$

subject to (1.1) with the double-obstacle homogeneous free energy density (1.2). Here,

$\mu_1, \mu_2 \geq 0$ are fixed and y_Ω and y_T are given targets, respectively, and u denotes the control variable. In this paper we consider control actions which enter through the right-hand side of the transport equation; i.e., we have

$$y_t - \Delta w = u, \quad w + \gamma \Delta y \in \partial \Psi(y). \tag{1.4}$$

As noted above, the Cahn–Hilliard system with a double-obstacle type homogeneous free energy density admits an equivalent reformulation as a variational inequality such that the resulting minimization problem amounts to an optimal control problem for a parabolic variational inequality which is of fourth order in space.

Recently, in [41, 42] optimal control problems for variational inequalities were linked to so-called mathematical programs with equilibrium constraints (MPECs). The latter problem class is well studied in finite dimensions; see, e.g., the monographs [47, 50, 51, 56] and the many references therein. It is well-known that MPECs are problematic from an optimization-theoretic point of view due to a generic lack of constraint qualification. This fact prevents the application of Karush–Kuhn–Tucker (KKT) type stationarity concepts for mathematical programs in Banach space [49,63]. In function space, the literature on MPECs is much scarcer and the work on the relation between finite and infinite dimensional versions of stationarity notions such as C-stationarity or strong (S) stationarity is only at its beginning. Most problems in function space are formulated in terms of elliptic variational inequalities (see [11,13,14, 38,39,42,44,52] and the references therein, for instance). An account of parabolic-type variational inequalities can be found in the literature on mathematical finance [3,4,43] and also in connection with the Stefan problem [11], as well as in the monograph [53] on the control of parabolic systems. Typically, the differential operator in these applications is of second order in space only, and to the best of our knowledge systems like the Cahn–Hilliard with double-obstacle potential are untreated in the literature.

In this paper our goal is to derive C-stationarity conditions for the minimization of (1.3) subject to (1.4) with the same boundary conditions as in (1.1) and with homogeneous free energy densities which cover the case of the double-obstacle free energy in (1.2). This will be achieved by a regularization process, which allows the application of classical KKT theory in Banach space [63] and the subsequent passage to the limit with respect to the regularization parameter. Concerning the control action, we study the case of distributed control as specified in (1.4) above. A further example for distributed control can be found, e.g., in [37], where a control problem for the Cahn–Hilliard equation with the double-well potential is studied.

Clearly, many applications require the coupling of the Cahn–Hilliard system with another system of partial differential equations describing the underlying physics. This could be the Navier–Stokes system [1] or linear or nonlinear elasticity [19,30,61], to mention only two. Coverage of such a fully coupled system, however, goes beyond the scope of the present work and rather justifies an independent study.

The remainder of the paper is organized as follows. In section 2 we introduce notation and provide existence and regularity results for solutions of parabolic differential inclusions. The optimal control problem is formulated in section 3 together with necessary assumptions on the involved operators. Then we study the existence of a solution of the original and the regularized MPEC as well as the behavior of solutions under a vanishing regularization parameter. Concerning the homogeneous free energy density we particularly highlight the distributed optimal control of the Cahn–Hilliard system with double-obstacle potential in section 4. For this particular case, for vanishing regularization parameter we obtain a function space version of C-stationarity conditions.

2. Preliminaries. In this section we introduce the notation used throughout this work. We also provide an abstract existence result for evolution inclusions and results concerning the regularity of solutions.

2.1. Notation and conventions. For any real Banach space X let us denote by $\|\,.\,\|_X$ its norm, with $\langle\,.\,,\,.\,\rangle_X : X^* \times X \to \mathbb{R}$ its dual pairing and the canonical injection of X into its bidual space by $i_X : X \to X^{**}$, $\langle i_X x, x^* \rangle_{X^*} := x^*(x)$. For $M \subset X$ we define a set $M^+ := \{x^* \in X^* \ : \ \langle x^*, m\rangle \geq 0 \quad \text{for all } m \in M\} \subset X^*$. If X is a Hilbert space, then $(\,.\,|\,.\,)_X : X \times X \to \mathbb{R}$ represents its inner product, and $J_X : X \to X^*$, $\langle J_X x, y\rangle_X := (x|y)_X$ denote the corresponding dual mapping. For two sets $M_1 \subset M_2$ we denote by $I_{M_1\to M_2} : M_1 \to M_2$, $I_{M_1\to M_2}\, x := x$ the identity regarded as a mapping from M_1 into M_2. Finally, the convex indicator function $\imath_{M_1} : M_2 \to \overline{\mathbb{R}}$ of M_1 is defined by

$$\imath_{M_1}(x) := \begin{cases} 0 & \text{if } x \in M_1, \\ \infty & \text{otherwise.} \end{cases}$$

Let Ω be an open, bounded, and connected subset of $\mathbb{R}^N$ with smooth boundary, let $\vec{n}$ be its unit outer normal vector field, and let $\mathcal{T} =]0, T[$ be a bounded interval. For $n \in \{1, \ldots, 4\}$ we define the spaces H, V_n by the following (for the definition of real-valued L^p-spaces, see, e.g., [5]; for vector-valued spaces, we refer the reader to [29]):

$$H := \left\{ u \in L^2(\Omega) \ : \ \int_\Omega u = 0 \right\},$$

$$V_n := \{u \in H^n(\Omega) : \Delta^k u \in H,\ \nabla\Delta^l u \cdot \vec{n}|_{\partial\Omega} = 0 \ \forall\, k, l \in \mathbb{N}_0 \text{ with } 2k \leq n,\ 2l + 2 \leq n\}.$$

In the above definition, Δ^0 equals the identity and there is no condition on $\nabla\Delta^l u$ for $n = 1$. The operators $L_1 : V_1 \to V_1^*$, $L_n : V_n \to V_{n-2}$ are given by

$$\langle L_1 u, v\rangle_{V_1} := \int_\Omega \nabla u \cdot \nabla v, \qquad L_2 u := -\Delta u, \qquad L_n := L_2|_{V_n}.$$

We equip H with the L^2–inner product and $V_1, \ldots, V_4$ with the inner products

$$(u|v)_{V_1} := \langle L_1 u, v\rangle_{V_1}, \qquad (u|v)_{V_n} := (L_n u | L_n v)_{V_{n-2}}.$$

For $n \in \{1, \ldots, 4\}$ the corresponding time-dependent spaces are

$$\mathcal{H} := L^2(\mathcal{T}; H), \quad \mathcal{V}_n := L^2(\mathcal{T}; V_n).$$

The spaces H and H^* will be identified via J_H. Furthermore, for $n \in \{3, 4\}$ we use the abbreviations $I := I_{V_1\to V_1}$, $I_1 := I_{V_1\to V_1^*}$, $I_2 := I_{V_2\to H}$, $I_n := I_{V_n\to V_{n-2}}$, and $I_{V_1} := I_{V_1\to H}$.

For ease of notation we identify a multivalued operator $A : X \rightrightarrows Y$ with its graph $\mathrm{gph}(A) \subset X \times Y$, define its domain $D(A)$ as the set $\{x \in X : Ax \neq \emptyset\}$, and denote the inverse and composition for $A \subset X \times Y$, $B \subset Y \times Z$ by

$$A^{-1} := \{(y, x) \ : \ (x, y) \in A\}, \quad BA := \{(x, z) \in X \times Z \ : \ Ax \cap B^{-1}z \neq \emptyset\}.$$

For multivalued operators from a Banach space X into a Banach space Y we denote by calligraphic letters the corresponding superposition operators mapping from $L^2(\mathcal{T}; X)$ into $L^2(\mathcal{T}; Y)$. For instance, for $A \subset X \times Y$ the operator $\mathcal{A}$ reads as

$$\mathcal{A} := \{(f, g) \in L^2(\mathcal{T}; X) \times L^2(\mathcal{T}; Y) \ : \ g(t) \in Af(t) \text{ for a.e. } t \in \mathcal{T}\}.$$

Moreover, for a functional on X denoted by a lowercase Greek letter, we use an uppercase Greek letter for its L^2-realization with respect to $\mathcal{T}$. If, for example, $\varphi : X \to \overline{\mathbb{R}}$ is given, then Φ is defined as

$$\Phi : L^2(\mathcal{T}; X) \to \overline{\mathbb{R}}, \;\; \Phi(u) := \begin{cases} \int_\Omega \varphi \circ u & \text{if } \varphi \circ u \in L^1(\mathcal{T}), \\ \infty & \text{otherwise.} \end{cases}$$

For two Banach spaces X and Y, we denote by $\mathcal{L}(X;Y)$ the space of continuous linear operators from X to Y. In the case $X = Y$, we also write $\mathcal{L}(X)$ instead of $\mathcal{L}(X;X)$. Finally, assume that $E \in \mathcal{L}(X;Y)$ is injective. Then the extension of φ to Y with respect to E is given by

$$\mathrm{Ext}(\varphi, X, Y, E) : Y \to \overline{\mathbb{R}}, \quad \mathrm{Ext}(\varphi, X, Y, E)(y) := \begin{cases} \varphi(E^{-1}y) & \text{if } y \in E(X), \\ \infty & \text{otherwise.} \end{cases}$$

Remark 2.1. 1. Since all the spaces $H, V_1, \dots, V_4$ are reflexive, they possess the Radon–Nikodým property (cf. [26]). Hence, e.g., $\mathcal{V}_1^*$ can be identified with the space $L^2(\mathcal{T}; V_1^*)$, and we shall do so without explicit use of the identification mapping. An analogous identification will be used for the other spaces.

2. The positive-definiteness of the inner products of $V_1, \dots, V_4$ is a consequence of the Poincaré inequality and the fact that the only solution to the homogeneous Laplace equation with zero mean value and Neumann boundary conditions is the 0-function. Moreover, it is not difficult to prove that these spaces are in fact Hilbert spaces.

3. Furthermore, the operator L_1 corresponds to the negative Laplacian with Neumann boundary conditions in its weak form as an operator from V_1 into V_1^*. By definition of the inner product of V_1, L_1 also coincides with the dual mapping of V_1, and all mappings $L_1, \dots, L_4$ are unitary operators.

4. Since we identify the Hilbert space H with its dual H^* through J_H, and because $I_{V_1} : V_1 \to H$ is linear, continuous, and has a dense range, the space $H \cong H^*$ becomes a subspace of V_1^* (under the identification $I_{V_1}^*$). Additionally, Rellich's lemma implies that I_{V_1} is a linear, compact mapping, and hence these properties transfer to $I_{V_1}^*$.

5. If $A : X \to Y$ is single-valued and defined on all of X, and if it is continuous and satisfies the growth condition

$$||Ax||_Y \leq C(1 + ||x||_X)$$

for a constant $C > 0$, then $\mathcal{A}$ is single-valued and defined on $L^2(\mathcal{T}; X)$.

Definition 2.2. *We define the following spaces that allow a weak (time-)derivative*

$$\mathcal{W}_1 := \{u \in \mathcal{V}_1 \; : \; u \in H^1(\mathcal{T}; V_1^*)\}, \quad \mathcal{W}_3 := \{u \in \mathcal{V}_3 \; : \; u \in H^1(\mathcal{T}; V_1^*)\},$$
$$\mathcal{W}_4 := \{u \in \mathcal{V}_4 \; : \; u \in H^1(\mathcal{T}; H)\}.$$

We fix the following norms:

$$||u||_{\mathcal{W}_1} := (||u||_{\mathcal{V}_1}^2 + ||u'||_{\mathcal{V}_1^*}^2)^{1/2}, \quad ||u||_{\mathcal{W}_3} := (||u||_{\mathcal{V}_3}^2 + ||u'||_{\mathcal{V}_1^*}^2)^{1/2},$$
$$||u||_{\mathcal{W}_4} := (||u||_{\mathcal{V}_4}^2 + ||u'||_{\mathcal{H}}^2)^{1/2}.$$

Most of the operators considered below are subdifferentials of convex functionals on various spaces.

DEFINITION 2.3. *For every real Banach space U and every functional $\varphi : U \to \overline{\mathbb{R}}$ we define a multivalued operator $\partial\varphi \subset U \times U^*$ by*

$$\partial\varphi := \{(u, u^*) \in U \times U^* \;:\; u \in \operatorname{dom}\varphi \;\wedge\; \varphi(v) - \varphi(u) \geq \langle u^*, v - u\rangle_U \;\; \forall\, v \in U\}.$$

If U is a Hilbert space, we further introduce $\partial^*\varphi \subset U \times U$:

$$\partial^*\varphi := J_U^{-1}\,\partial\varphi.$$

2.2. Existence results. Now we establish two properties of the extension of functionals that will be used below. The proofs of this lemma and of other basic results of this section are provided in the appendix.

LEMMA 2.4. *Let X and Y be two Banach spaces, let $E \in \mathcal{L}(X;Y)$ be injective, and let $\varphi : X \to \overline{\mathbb{R}}$ a functional. Then we have*

$$\partial\operatorname{Ext}(\varphi, X, Y, E) = (E^*)^{-1}\partial\varphi E^{-1}.$$

If, furthermore, X is reflexive, φ is proper, convex, and lower-semicontinuous and has bounded lower-level sets (i.e., $\varphi^{-1}(]-\infty, a])$ is bounded in X for every $a \in \mathbb{R}$), then $\operatorname{Ext}(\varphi, X, Y, E) : Y \to \overline{\mathbb{R}}$ is proper, convex, and lower-semicontinuous as well.

The following theorem by Abels/Wilke [2, Theorem 3.1] is our central key to proving existence of solutions to evolution inclusions (cf. also [46]).

THEOREM 2.5. *Assume that W_1 and W_0 are real, separable Hilbert spaces and that $E \in \mathcal{L}(W_1;W_0)$ is injective with $E(W_1)$ dense in W_0. Let $\alpha > 0$, let $\psi_0 := \operatorname{Ext}(\alpha||\,.\,||^2_{W_1}, W_1, W_0, E)$, and let $\psi : W_0 \to \overline{\mathbb{R}}$ be a nonnegative, proper, convex, and lower-semicontinuous functional such that $\operatorname{dom}\psi \cap \operatorname{dom}\psi_0 \neq \emptyset$. We set $\psi_1 := \psi_0 + \psi$ and $R := \partial^*\psi_1$. Furthermore, suppose that $S : W_1 \to W_0$ is Lipschitz continuous. Then, for every $y_0 \in D(R)$ and $f \in L^2(\mathcal{T};W_0)$ there exists a unique $y \in L^\infty(\mathcal{T};W_1)$ with $\mathcal{E}y \in H^1(\mathcal{T};W_0)$ such that*

$$(\mathcal{E}y)' + \mathcal{R}\mathcal{E}y + \mathcal{S}y \ni f,$$
$$(\mathcal{E}y)(0) = y_0.$$

Moreover, it holds that $\Psi_1(\mathcal{E}y) \in L^\infty(\mathcal{T})$.

Remark 2.6. The nonnegativity assumption on ψ is not essential. Indeed, if $\psi : W_0 \to \overline{\mathbb{R}}$ is proper, convex, and lower-semicontinuous, then it is bounded from below by an affine functional, i.e., $\psi(w) \geq (w_0\,|\,w)_{W_0} + r$ for some $w_0 \in W_0$, $r \in \mathbb{R}$. Therefore, we can apply Theorem 2.5 to $\widetilde{\psi}(w) := \psi(w) - (w_0\,|\,w)_{W_0} - r$ and $\widetilde{S}w := Sw + w_0$.

Next, we provide a reformulation of this theorem applied to our situation.

THEOREM 2.7. *Suppose we are given a proper, convex, and lower-semicontinuous functional $\varphi : V_1 \to \overline{\mathbb{R}}$, a Lipschitz mapping $B : V_1 \to V_1$, $y_0 \in D(\partial\operatorname{Ext}(\varphi, V_1, V_1^*, I_1))$, and $f \in \mathcal{V}_1^*$. Then there exists a unique $y \in \mathcal{W}_1 \cap L^\infty(\mathcal{T};V_1)$ with*

$$(\mathcal{I}_1 y)' + \mathcal{L}_1 w = f, \tag{2.1}$$

$$\mathcal{L}_1 y + (\partial\Phi)(y) + \mathcal{I}_1\mathcal{B}y \ni w, \tag{2.2}$$

$$(\mathcal{I}_1 y)(0) = y_0. \tag{2.3}$$

Additionally, y satisfies $\varphi(y) + \frac{1}{2}||y||^2_{V_1} \in L^\infty(\mathcal{T})$.

Proof. We define $\varphi_0, \varphi_1 : V_1 \to \overline{\mathbb{R}}$ by $\varphi_0 := \frac{1}{2}||\,.\,||^2_{V_1}$, $\varphi_1 := \varphi_0 + \varphi$, and $A := \partial\varphi$ and apply Theorem 2.5 to the following setting:

$$W_1 := V_1, \quad W_0 := V_1^*, \quad E := I_1, \quad \alpha := \frac{1}{2}, \quad R := \partial^*\Psi_1, \quad S := L_1 B,$$

where ψ, ψ_0, and ψ_1 denote the extensions of φ, φ_0, and φ_1, to V_1^* with respect to I_1. From Lemma 2.4 it follows that ψ is proper, convex, and lower-semicontinuous. Hence, there exists a unique $y \in L^\infty(\mathcal{T}; V_1)$ with $\mathcal{I}_1 y \in H^1(\mathcal{T}; V_1^*), \Psi_1(\mathcal{I}_1 y) \in L^\infty(\mathcal{T})$ and

$$(\mathcal{I}_1 y)' + \mathcal{R}\mathcal{I}_1 y + \mathcal{S}y \ni f, \quad (\mathcal{I}_1 y)(0) = y_0. \tag{2.4}$$

In order to calculate $\mathcal{R}$ we use two basic facts from functional analysis:

(1) For every Hilbert space W it holds that $i_W = J_{W^*} J_W$, $J_{W^*}^{-1} = J_W^*$.
(2) For Banach spaces X, Y and $F \in \mathcal{L}(X; Y)$ we have $F^{**} i_X = i_Y F$.

Thus, $J_H^* i_H = J_{H^*}^{-1} J_{H^*} J_H = J_H$, and from $I_1 = I_{V_1}^* J_H I_{V_1}$ it follows that

$$I_1^* = I_{V_1}^* J_H^* I_{V_1}^{**} = I_{V_1}^* J_H^* i_H I_{V_1} i_{V_1}^{-1} = I_1 i_{V_1}^{-1} = I_1 J_{V_1}^{-1} J_{V_1^*}^{-1}.$$

Hence, Lemma 2.4 yields

$$R = J_{V_1^*}^{-1} \partial \operatorname{Ext}(\varphi_1, V_1, V_1^*, I_1) = J_{V_1^*}^{-1} (I_1^*)^{-1} \partial \varphi_1 I_1^{-1} = J_{V_1} I_1^{-1} \partial \varphi_1 I_1^{-1} = L_1 I_1^{-1} \partial \varphi_1 I_1^{-1},$$

and since $\varphi_1 = \varphi_0 + \varphi$ with φ_0 being continuous on V_1 and therefore $\partial\varphi_1 = \partial\varphi_0 + \partial\varphi = L_1 + A$, it follows that

$$R I_1 = L_1 I_1^{-1} \partial \varphi_1 = L_1 I_1^{-1}(L_1 + A).$$

Consequently, (2.4) is equivalent to $(\mathcal{I}_1 y)' + \mathcal{L}_1 \mathcal{I}_1^{-1}(\mathcal{L}_1 + \mathcal{A} + \mathcal{I}_1 \mathcal{B}) y \ni f$. Defining $w := \mathcal{L}_1^{-1}(f - (\mathcal{I}_1 y)')$, this can be rewritten as

$$(\mathcal{I}_1 y)' + \mathcal{L}_1 w = f, \quad (\mathcal{L}_1 + \mathcal{A} + \mathcal{I}_1 \mathcal{B}) y \ni \mathcal{I}_1 w,$$

which completes the proof. ∎

2.3. Regularity results. The following proposition provides some basic properties of the interplay between I_1 and L_1.

PROPOSITION 2.8. *The following commutation rules hold true:*

$$I_1 L_3 = L_1 I_3 \quad \text{and} \quad I_3^* L_1 = L_3^* I_1.$$

Moreover, $L_3 : V_3 \to V_1$, $L_1 : V_1 \to V_1^*$, and $L_3^* : V_1^* \to V_3^*$ are unitary operators, and the mappings $L_1^{-1} I_1 \in \mathcal{L}(V_1)$ and $(L_3^*)^{-1} I_3^*$ are positive, symmetric, injective, and compact. If $U_1 \subset V_1$ and $U_1^* \subset V_1^*$ are closed subspaces such that $L_1^{-1} I_1(U_1) = U_1$ and $(L_3^*)^{-1} I_3^*(U_1^*) = U_1^*$ are satisfied, then it holds that

$$I_1 P_{V_1 \to U_1} = P_{V_1^* \to I_1(U_1)} I_1, \; I_3^* P_{V_1^* \to U_1^*} = P_{V_3^* \to I_3^*(U_1^*)} I_3^*,$$
$$L_1 P_{V_1 \to U_1} = P_{V_1^* \to I_1(U_1)} L_1, \; L_3^* P_{V_1^* \to U_1^*} = P_{V_3^* \to I_3^*(U_1^*)} L_3^*$$

with $P_{X \to Y}$ denoting the othogonal projection of X onto the closed subspace $Y \subset X$.

Furthermore, the following generalization of the integration-by-parts formula to functions admitting time-derivatives in different spaces is needed.

PROPOSITION 2.9. *Assume that $y \in \mathcal{V}_3$, $v \in \mathcal{V}_1$ with $y \in H^1(\mathcal{T}; V_1^*)$ and $v \in H^1(\mathcal{T}; V_3^*)$. Then y and v possess representatives in $C(\mathcal{T}; V_1)$ and $C(\mathcal{T}; V_1^*)$, respectively, which can be continuously extended to $\overline{\mathcal{T}}$. Moreover, in this sense it holds that*

$$\langle y', v \rangle_{\mathcal{V}_1} + \langle v', y \rangle_{\mathcal{V}_3} = \langle v(T), y(T) \rangle_{V_1} - \langle v(0), y(0) \rangle_{V_1}.$$

Next, we state a regularity result for the time-dependent bi-Laplace equation ensuring higher space regularity.

THEOREM 2.10. *Suppose that* $y \in \mathcal{V}_1$ *and* $f \in \mathcal{V}_1^*$ *satisfy* $y \in H^1(\mathcal{T}; V_3^*)$, $y(0) \in V_1$ *together with* $y' + \mathcal{L}_3^* \mathcal{L}_1 y = f$ *in* $\mathcal{V}_3^*$. *Then* $y \in \mathcal{V}_3$, $y \in H^1(\mathcal{T}; V_1^*)$ *and it holds that*

$$y' + \mathcal{L}_1 \mathcal{L}_3 y = f \ \text{ in } \mathcal{V}_1^*.$$

Remark 2.11. If we assume that $f \in \mathcal{H}$ and $y(0) \in V_2$ are satisfied, then $y \in \mathcal{V}_4$, $y \in H^1(\mathcal{T}; H)$, and $y' + \mathcal{L}_2 \mathcal{L}_4 y = f \in \mathcal{H}$.

In order to derive an optimality system later on, the continuous Fréchet differentiability of some superposition operators acting on integrable functions with values in Sobolev spaces will be needed.

PROPOSITION 2.12. *Let* $\Omega \subset \mathbb{R}^N$ *be an open, bounded set with smooth boundary, and let* $\gamma : \mathbb{R} \to \mathbb{R}$ *be twice continuously differentiable with bounded first and second derivatives. Let* $\mathcal{A}$ *denote the superposition operator*

$$\mathcal{A} : L^{s_1}(\mathcal{T}; W^{1,s_2}(\Omega)) \to L^2(\mathcal{T}; H^1(\Omega)), \quad (\mathcal{A}y)(t,x) := \gamma(y(t,x)).$$

If $s_1 \geq 4$ *and* $s_2 > 2$ *with* $s_2 \geq \frac{4N}{N+2}$, *then the operator* $\mathcal{A}$ *is continuously Fréchet differentiable, and its derivative* $\mathcal{A}' : L^{s_1}(\mathcal{T}; W^{1,s_2}(\Omega)) \to \mathcal{L}(L^{s_1}(\mathcal{T}; W^{1,s_2}(\Omega)); L^2(\mathcal{T}; H^1(\Omega)))$ *is given by*

$$(\mathcal{A}'(y;r))(t,x) = \gamma'(y(t,x))r(t,x).$$

3. Optimal control for the Cahn–Hilliard system. Now, we turn our attention to a problem of optimal control of the Cahn–Hilliard system. For theoretical as well as numerical purposes, the original problem is approximated by a suitable sequence of associated auxiliary problems which are easier to handle from an optimization-theoretic point of view. We then derive an optimality system for the auxiliary problems which permits us to pass to the limit. This results in first-order optimality conditions of C-stationarity type for the original problem.

In addition to our earlier notation and assumptions we invoke the following assumption.

ASSUMPTION 3.1. *For* $\alpha_0 > 0$ *and* $\Lambda :=]0, \alpha_0]$, $\overline{\Lambda} := [0, \alpha_0]$ *let* $(\varphi_\alpha)_{\alpha \in \overline{\Lambda}}$ *be proper, convex, and lower-semicontinuous functionals on* V_1 *with values in* $\overline{\mathbb{R}}$, *and let* $\mathcal{C} \subset \mathcal{H}$ *be closed, convex, and nonempty. We fix* $\eta \leq 0$ *and define* $\psi := -\frac{\eta}{2} || I_{V_1}(.) ||_H^2$, $\varphi := \varphi_0$, $\theta_\alpha := \frac{1}{2} || \, . \, ||_{V_1}^2 + \varphi_\alpha - \psi : V_1 \to \overline{\mathbb{R}}$ *and the operators* $A_\alpha := \partial \varphi_\alpha$, $A := \partial \varphi$ *for* $\alpha \in \overline{\Lambda}$. *We suppose the following:*

- (i) $(A_\alpha)_{\alpha \in \Lambda}$ *are single-valued operators from* V_1 *into* V_1^*.
- (ii) *There exists some constant* $c_\theta > 0$ *such that the functionals* $(\theta_\alpha - c_\theta || \, . \, ||_{V_1}^2)_{\alpha \in \overline{\Lambda}}$ *are bounded from below by some common constant.*
- (iii) *The restrictions of the superposition operators* $(\mathcal{A}_\alpha)_{\alpha \in \Lambda}$ *of* $(A_\alpha)_{\alpha \in \Lambda}$ *to the subspace* $\mathcal{W}_1 \subset \mathcal{V}_1$ *attain only values in* $\mathcal{V}_1 \subset \mathcal{V}_1^*$. *We define for* $\alpha \in \Lambda$

$$\widehat{\mathcal{A}}_\alpha : \mathcal{W}_1 \to \mathcal{V}_1, \quad \widehat{\mathcal{A}}_\alpha := \mathcal{A}_\alpha.$$

- (iv) *For every* $\alpha \in \Lambda$, *the operator* $(\widehat{\mathcal{A}}_\alpha)_{\alpha \in \Lambda}$ *is continuously Fréchet differentiable from* $\mathcal{W}_3$ *into* $\mathcal{V}_1$. *In addition, for every* $y \in \mathcal{W}_4$ *its derivative* $D\widehat{\mathcal{A}}_\alpha(y)$ *can be continuously extended to a linear, bounded operator* $\overline{D\mathcal{A}_\alpha}(y)$ *from* $\mathcal{V}_1$ *into itself.*
- (v) *For all* $\alpha \in \Lambda$ *and* $y \in \mathcal{W}_3$ *it holds that* $\langle \mathcal{L}_1 y, \widehat{\mathcal{A}}_\alpha y \rangle_{\mathcal{V}_1} \geq 0$.

(vi) *Moreover, if* (α_n) *and* (y_n) *are sequences in* Λ *and* $\mathcal{W}_1$ *such that* $\alpha_n \to 0$, $y_n \rightharpoonup y$ *in* $\mathcal{W}_1$, $y_n \to y$ *in* $\mathcal{H}$, *and* $\alpha \in \Lambda$, *then*

(1) $\mathcal{A}_\alpha y_n \rightharpoonup \mathcal{A}_\alpha y$ *in* $\mathcal{V}_1^*$,

(2) *if* $\widehat{\mathcal{A}}_{\alpha_n} y_n \rightharpoonup h$ *in* $\mathcal{H}$, *then* $(y, h) \in \mathcal{A}$.

We fix $y_\Omega \in \mathcal{H}$, $y_T \in V_1$, $\mu_1, \mu_2 \geq 0$ *and define the functional* $J : \mathcal{W}_1 \times \mathcal{H} \to \mathbb{R}$ *by*

$$J(y,u) := \frac{1}{2}\Big(\mu_1 \|y - y_\Omega\|_{\mathcal{H}}^2 + \mu_2 \|y(T) - y_T\|_H^2 + \|u\|_{\mathcal{H}}^2\Big).$$

Finally, let $y_0 \in V_1$ *be given such that* $y_0 \in D(\partial\,\mathrm{Ext}(\varphi_\alpha, V_1, V_1^*, I_1))$ *for* $\alpha \in \overline{\Lambda}$ *and* $\sup\{\theta_\alpha(y_0) : \alpha \in \overline{\Lambda}\} =: C_\theta < \infty$.

Throughout the rest of the paper, we refer to y as the state and u as the control variable, respectively.

For given $u \in \mathcal{H}$ and $\alpha \in \overline{\Lambda}$ consider the problems of finding $y \in \mathcal{W}_1$ such that there is a $w \in \mathcal{V}_1$ with

$$(\mathrm{Q}_\alpha) \qquad y' + \mathcal{L}_1 w = u, \quad w \in (\mathcal{L}_1 + \mathcal{A}_\alpha + \eta \mathcal{I}_1) y, \quad y(0) = y_0,$$

and $(\mathrm{Q}) := (\mathrm{Q}_0)$. The following theorem shows that the problem (Q_α) admits a unique solution.

Remark 3.2. In section 4 we focus on the special case where φ is the convex indicator function of the set $\{v \in V_1 : v(x) \in [a,b] \text{ a.e. on } \Omega\}$ for $a, b \in \mathbb{R}$, $a < b$, and φ_α for $\alpha > 0$ is some suitable mollified version of its Moreau–Yosida approximation. Note that under these conditions and for $\alpha > 0$, (Q_α) represents a regularized version of the Cahn–Hilliard system with double-obstacle homogeneous free energy density. The regularization acts on the indicator function involved in the potential. For $\alpha = 0$ we arrive at the original Cahn–Hilliard system with double-obstacle potential.

THEOREM 3.3. For every $\alpha \in \overline{\Lambda}$ and every right-hand side $u \in \mathcal{H}$ the problem (Q_α) has a unique solution $y \in \mathcal{W}_1$. If $\alpha \in \Lambda$, the solution satisfies $y \in \mathcal{W}_3$, and with $w := \mathcal{L}_1^{-1}(u - y')$, we have

$$y' + \mathcal{L}_1 w = u, \quad w = (\mathcal{L}_3 + \widehat{\mathcal{A}}_\alpha + \eta \mathcal{I}) y, \quad y(0) = y_0,$$

Proof. The existence and uniqueness of a solution are direct consequences of Theorem 2.7. For given $\alpha \in \Lambda$, let $y \in \mathcal{W}_1$ be the solution to (Q_α) and $w := \mathcal{L}_1^{-1}(u - y')$. Then

$$\mathcal{I}_3^*(y' + \mathcal{L}_1 w) = y' + \mathcal{L}_3^* w = y' + \mathcal{L}_3^*(\mathcal{L}_1 y + \mathcal{A}_\alpha y + \eta \mathcal{I}_1 y),$$

and since $\mathcal{A}_\alpha y, \eta \mathcal{I}_1 y \in \mathcal{V}_1$, it follows that $\mathcal{L}_3^*(\mathcal{A}_\alpha y + \eta \mathcal{I}_1 y) = \mathcal{L}_3^* \mathcal{I}_1 \mathcal{I}_1^{-1}(\mathcal{A}_\alpha y + \eta \mathcal{I}_1 y) = \mathcal{L}_1 \mathcal{I}_1^{-1}(\mathcal{A}_\alpha y + \eta \mathcal{I}_1 y) \in \mathcal{V}_1^*$. Consequently, Theorem 2.10 implies the assertion. ∎

Remark 3.4. Theorem 3.3 shows that the original problem (Q) admits a solution $y \in \mathcal{W}_1$ for every $u \in \mathcal{H}$. Furthermore, the solutions y of the regularized problems (Q_α) belong to $\mathcal{W}_3$ for all $\alpha \in \Lambda$. This holds true even for all $u \in \mathcal{V}_1^*$ and will be used to prove the existence of Lagrange multipliers for the optimal control problem (P_α) below. For $u \in \mathcal{H}$ we can use Remark 2.11 and obtain $y \in \mathcal{W}_4$ for $\alpha > 0$. This higher regularity will be needed only to show that in the linearization of (Q_α) the derivative $D\widehat{\mathcal{A}}_\alpha(y)$ for $y \in \mathcal{W}_4$ extends continuously to a operator in $\mathcal{L}(\mathcal{V}_1; \mathcal{V}_1)$ (which, in general, does not hold for $y \in \mathcal{W}_3$).

DEFINITION 3.5. *For given* $u \in \mathcal{H}$, *let* $S_\alpha u \in \mathcal{W}_1$ *denote the solution of* (Q_α) *given by Theorem* 3.3.

3.1. Regularized optimal control problems. In order to derive an optimality system for problem (P) below, the nonsmooth potential φ is replaced by φ_α leading to the family of smooth optimization problems (P_α), which we study next.

For $\alpha \in \overline{\Lambda}$ and with $(\mathrm{P}) := (\mathrm{P}_0)$ we consider

$$(\mathrm{P}_\alpha) \qquad \inf\{J(y,u) \,:\, (y,u) \in \mathcal{W}_1 \times \mathcal{C},\ y = S_\alpha u\}.$$

For studying $\alpha \to 0$ in (P_α) we need the energy estimate for a solution to problem (Q_α) as given in the next lemma.

LEMMA 3.6. *Let $u \in \mathcal{H}$ and $\alpha \in \overline{\Lambda}$ be given. For the corresponding solution $y := S_\alpha u \in \mathcal{W}_1$ to problem (Q_α) and $w := \mathcal{L}_1^{-1}(u - y')$ it holds that $\theta_\alpha \circ y$ admits an absolutely continuous representative on $\overline{\mathcal{T}}$, and in this sense for $t \in \overline{\mathcal{T}}$ and $\mathcal{T}' := [0,t]$ we have*

$$(\theta_\alpha \circ y)(t) + ||\chi_{\mathcal{T}'} w||_{\mathcal{V}_1}^2 \;=\; (\theta_\alpha \circ y)(0) + \langle u\,, \chi_{\mathcal{T}'} w\rangle_{\mathcal{V}_1}.$$

In particular, it holds that $||\theta_\alpha \circ y||_{L^\infty(\mathcal{T})} + ||w||_{\mathcal{V}_1}^2 \le (\theta_\alpha \circ y)(0) + ||u||_{\mathcal{V}_1^*}||w||_{\mathcal{V}_1}$.

Proof. The chain rule (cf. Proposition 4.2 in [24]) implies for the convex functionals $\sigma = \theta_\alpha + \psi$ and $\sigma = \psi$ that $\sigma \circ y$ is absolutely continuous on $\overline{\mathcal{T}}$ and for every $v \in \mathcal{V}_1$ such that $v(t) \in \partial\sigma(y(t))$ for almost all $t \in \overline{\mathcal{T}}$, and with $\mathcal{T}' := [0,t]$ it holds that

$$(\sigma \circ y)(t) = \int_{\mathcal{T}'} \frac{d}{dt}(\sigma \circ y)\, dt + (\sigma \circ y)(0) \;=\; \left\langle y'\,, \chi_{\mathcal{T}'} v\right\rangle_{\mathcal{V}_1} + (\sigma \circ y)(0).$$

Using $w \in (\mathcal{L}_1 + \mathcal{A}_\alpha + \eta\mathcal{I}_1)y = (\partial(\Theta_\alpha + \Psi) - \partial\Psi)y$, we obtain

$$\begin{aligned}(\theta_\alpha \circ y)(t) &= \left\langle y'\,, \chi_{\mathcal{T}'} w\right\rangle_{\mathcal{V}_1} + (\theta_\alpha \circ y)(0)\\ &= \langle u - \mathcal{L}_1 w\,, \chi_{\mathcal{T}'} w\rangle_{\mathcal{V}_1} + (\theta_\alpha \circ y)(0)\\ &= -||\chi_{\mathcal{T}'} w||_{\mathcal{V}_1}^2 + \langle u, \chi_{\mathcal{T}'} w\rangle_{\mathcal{V}_1} + (\theta_\alpha \circ y)(0),\end{aligned}$$

which yields the desired assertions. □

COROLLARY 3.7. *Let (α_n) and (u_n) be sequences in $\overline{\Lambda}$ and in $\mathcal{H}$, respectively, such that (u_n) is bounded in $\mathcal{V}_1^*$. By (y_n) we denote the sequence of corresponding solution $S_{\alpha_n} u_n$ to the problem (Q_{α_n}) and $w_n := \mathcal{L}_1^{-1}(u_n - y_n')$. Then (y_n) is bounded in $\mathcal{W}_1$ and $L^\infty(\mathcal{T}; V_1)$, and (w_n) is bounded in $\mathcal{V}_1$. If $\alpha_n \in \Lambda$ for all n, then $\widehat{\mathcal{A}}_{\alpha_n} y_n$ is bounded in $\mathcal{H}$.*

Proof. From the energy estimate of Lemma 3.6 it follows that (w_n) is bounded in $\mathcal{V}_1$. Hence, $y_n' = u_n - \mathcal{L}_1 w_n$ is bounded in $\mathcal{V}_1^*$. Moreover, the energy estimate implies that (y_n) is bounded in $L^\infty(\mathcal{T}; V_1)$ since $\theta_\alpha \ge c_\theta||\,.\,||_{V_1}^2 - c$ for all $\alpha \in \overline{\Lambda}$ and some $c \in \mathbb{R}$. Hence, (y_n) is also bounded in $\mathcal{W}_1$. If $\alpha_n \in \Lambda$ for all n, then we obtain from $w_n = (\mathcal{L}_1 + \widehat{\mathcal{A}}_{\alpha_n} + \eta\mathcal{I}_1)y_n$, and with the help of Assumption 3.1(iv)–(v), that

$$\begin{aligned}||\widehat{\mathcal{A}}_{\alpha_n} y_n||_{\mathcal{H}}^2 = \langle \widehat{\mathcal{A}}_{\alpha_n} y_n\,, \widehat{\mathcal{A}}_{\alpha_n} y_n\rangle_{\mathcal{V}_1} \;&=\; \left\langle w_n - (\mathcal{L}_1 + \eta\mathcal{I}_1)y_n\,, \widehat{\mathcal{A}}_{\alpha_n} y_n\right\rangle_{\mathcal{V}_1}\\ &\le (w_n - \eta\mathcal{I}_1 y_n\,|\,\widehat{\mathcal{A}}_{\alpha_n} y_n)_{\mathcal{H}} \;\le\; C\,||\widehat{\mathcal{A}}_{\alpha_n} y_n||_{\mathcal{H}}.\end{aligned}$$

This finishes the proof. □

With these a priori estimates we are able to prove convergence (consistence) results and the existence of minimizers for (P_α).

PROPOSITION 3.8. *Let* (α_n) *and* (u_n) *be sequences in* $\overline{\Lambda}$*, respectively,* $\mathcal{H}$*, with* $\alpha_n \to 0$ *and* $u_n \rightharpoonup u$ *in* $\mathcal{V}_1^*$ *for some* $u \in \mathcal{H}$*. Then there exist subsequences (denoted by the index* m*) such that*

$$S_{\alpha_m} u_m \rightharpoonup Su \quad in\ \mathcal{W}_1, S_{\alpha_m} u_m \to Su \quad in\ \mathcal{V}_1.$$

Proof. From Corollary 3.7 we already know that $y_n := S_{\alpha_n} u_n$ is bounded in $\mathcal{W}_1$ and $w_n := \mathcal{L}_1^{-1}(u_n - y_n')$ in $\mathcal{V}_1$. Due to the compactness of $\mathcal{I}_1 : \mathcal{W}_1 \to \mathcal{H}$ and with $a_n := w_n - (\mathcal{L}_1 + \eta\mathcal{I}_1)y_n \in \mathcal{A}_{\alpha_n} y_n$ we have that

$$\begin{aligned} (y_m, w_m, a_m) &\rightharpoonup (y, w, a) && \text{in } \mathcal{W}_1 \times \mathcal{V}_1 \times \mathcal{V}_1^*, \\ y_m &\to y && \text{in } \mathcal{H} \end{aligned}$$

for some $y \in \mathcal{W}_1$, $w \in \mathcal{V}_1$, and $a \in \mathcal{V}_1^*$ along a subsequence α_m of α_n. Moreover, since $w_n = \mathcal{L}_1 y_n + a_n + \eta y_n$, it follows that $w = \mathcal{L}_1 y + a + \eta y$ as well as

$$\begin{aligned} ||y_m - y||_{\mathcal{V}_1}^2 &= \left\langle \mathcal{L}_1(y_m - y), y_m - y \right\rangle_{\mathcal{V}_1} \\ &= (w_m - w \,|\, y_m - y)_{\mathcal{H}} - \langle a_m - a, y_m - y \rangle_{\mathcal{V}_1} - \langle \eta y_m - \eta y, y_m - y \rangle_{\mathcal{V}_1} \end{aligned}$$

for $m \in \mathbb{N}$. The strong convergence of y_m in $\mathcal{H}$, respectively, $\mathcal{V}_1^*$, shows that the first and the last term on the right-hand side tend to zero. In the case that $\alpha_m = 0$ for infinitely many m, for these indices it holds that $\langle a_m - a, y_m - y \rangle_{\mathcal{V}_1} \geq 0$ by the monotonicity of $\mathcal{A}$, and hence $y_m \to y$ in $\mathcal{V}_1$. Thus, from Proposition 2.5 of [20] we obtain that $(y, a) \in \mathcal{A}$.

Otherwise, we may assume the $\alpha_m > 0$ for all m and that $\widehat{\mathcal{A}}_{\alpha_m} y_m$ converges weakly in $\mathcal{H}$ to an $a_1 \in \mathcal{H}$ by Corollary 3.7. This implies $\langle a_m - a, y_m - y \rangle_{\mathcal{V}_1} = (a_m - a \,|\, y_m - y)_{\mathcal{H}} \to 0$. Consequently, it holds that $y_m \to y$ in $\mathcal{V}_1$ and that $a = a_1$ and $(y, a_1) \in \mathcal{A}$ by Assumption 3.1(vi). □

PROPOSITION 3.9. *For every* $\alpha \in \overline{\Lambda}$ *the problem* (P_α) *admits a minimizer* $(y, u) \in \mathcal{W}_1 \times \mathcal{C}$*.*

Proof. Although the proof technique is standard, we provide the proof for the sake of keeping the paper self-contained. Let $\alpha \in \overline{\Lambda}$ be given, and let $(S_\alpha u_n, u_n)$ be an infimizing sequence for problem (P_α). We set $y_n := S_\alpha u_n \in \mathcal{W}_1$ and $w_n := \mathcal{L}_1^{-1}(u_n - y_n')$. The coercivity of J yields that (u_n) is bounded in $\mathcal{H}$. With the help of Corollary 3.7 we may pass to subsequences (denoted by the index m) such that

$$\begin{aligned} (y_m, u_m) &\rightharpoonup (y, u) && \text{in } \mathcal{W}_1 \times \mathcal{H}, \\ y_m &\to y && \text{in } \mathcal{H}. \end{aligned}$$

The continuity properties of $\mathcal{A}_\alpha$ given in Assumption 3.1(vi) imply that $y = S_\alpha u$ if $\alpha > 0$. In the case of $\alpha = 0$ this is obtained by Proposition 3.8. Moreover, $u \in \mathcal{C}$ since $\mathcal{C}$ is weakly closed. Finally, the weakly lower semicontinuity of $J : \mathcal{W}_1 \times \mathcal{H} \to \mathbb{R}$ implies that (y, u) is in fact a minimizer of (P_α). □

Next, it will be shown that a sequence (y_n, u_n) of minimizers to problem (P_{α_n}) for $\alpha_n \to 0$ admits a cluster point in a suitable topology which is a minimizer of (P). For this purpose we have to pass to the limit in $J(y_n, u_n)$ which, in particular, requires strong convergence of $y(T)$ in H. This is proved next.

LEMMA 3.10. *Let* (α_n) *be a sequence in* Λ *with* $\alpha_n \to 0$*, and let* $(y_n, u_n) \subset \mathcal{W}_1 \times \mathcal{H}$ *be a sequence of solutions to problem* (Q_{α_n}) *such that* (u_n) *is bounded in* $\mathcal{H}$*. Then there exist* $a \in \mathcal{H}$*,* $y \in \mathcal{W}_1$*, and a subsequence (denoted by the index* m*) such that*

$$\begin{aligned} (y_m, \widehat{\mathcal{A}}_{\alpha_m} y_m, u_m) &\rightharpoonup (y, \xi, u) && in\ \mathcal{W}_1 \times \mathcal{H} \times \mathcal{H}, \\ (y_m, y_m(T)) &\to (y, y(T)) && in\ \mathcal{V}_1 \times H. \end{aligned}$$

Proof. Since (u_n) is bounded in $\mathcal{H}$ and due to Corollary 3.7, we can pass to a subsequence such that $u_m \rightharpoonup u$ in $\mathcal{H}$ and $\widehat{\mathcal{A}}_{\alpha_n} y_m \rightharpoonup \xi$ in $\mathcal{H}$ for some $u, \xi \in \mathcal{H}$. Then, Proposition 3.8 shows that $y_m \to y := Su$ weakly in $\mathcal{W}_1$ and strongly in $\mathcal{V}_1$. It remains to show that $y_m(T)$ converges strongly to $y(T)$ in H. For this purpose, notice that $\mathcal{W}_1$ embeds continuously into $C(\overline{\mathcal{T}}; H)$, and hence $y_m(T) \rightharpoonup y(T)$ in H. In order to prove the strong convergence, we show that (y_n) is bounded in $\mathcal{V}_2$, and we apply a compactness argument.

As in the proof of Theorem 2.10, which is given in section A, let (e_n) be a complete orthonormal system of eigenvectors for the eigenvalues $(\lambda_n)_{n\in\mathbb{N}}$ in V_1 of the operator $L_1^{-1} I_1$ and the span of $V_{1,n} := \operatorname{span}\{e_1, \dots, e_n\}$. Defining $\widetilde{y}_{n,m} := P_{V_{1,m}} y_n$ and $\widetilde{u}_{n,m} := P_{V_{1,m}} u_n$, where $P_{V_{1,n}}$ denotes the orthogonal projection of V_1 onto $V_{1,n}$, it follows that $\widetilde{y}_{n,m} \to y_n$ in $\mathcal{V}_1$ as $m \to \infty$. Since (y_n) solves (Q_{α_n}) for the right-hand side $u_n \in \mathcal{H}$ with $\widetilde{w}_{n,m} := \mathcal{L}_1^{-1}(\widetilde{u}_{n,m} - \widetilde{y}'_{n,m})$, we obtain that

$$
\begin{aligned}
&\big(\widetilde{u}_{n,m} \,\big|\, \widetilde{y}_{n,m}\big)_{\mathcal{H}} - \frac{1}{2}(\|\widetilde{y}_{n,m}(T)\|_H^2 - \|\widetilde{y}_{n,m}(0)\|_H^2) \\
&\quad = -\big(\widetilde{y}'_{n,m} - \widetilde{u}_{n,m} \,\big|\, \widetilde{y}_{n,m}\big)_{\mathcal{H}} = \big\langle \mathcal{L}_1 \widetilde{w}_{n,m}, \widetilde{y}_{n,m} \big\rangle_{\mathcal{V}_1} = \big\langle \widetilde{w}_{n,m}, \mathcal{L}_3 \widetilde{y}_{n,m} \big\rangle_{\mathcal{V}_1} \\
&\quad = \|\mathcal{L}_3 \widetilde{y}_{n,m}\|_{\mathcal{H}}^2 + \big\langle \mathcal{A}_{\alpha_n} \widetilde{y}_{n,m}, \mathcal{L}_3 \widetilde{y}_{n,m} \big\rangle_{\mathcal{V}_1} + \big\langle \eta \widetilde{y}_{n,m}, \mathcal{L}_3 \widetilde{y}_{n,m} \big\rangle_{\mathcal{V}_1}.
\end{aligned}
$$

Because $(\eta \widetilde{y}_{n,m})$ is bounded in $\mathcal{H}$ and $\big\langle \mathcal{A}_{\alpha_n} \widetilde{y}_{n,m}, \mathcal{L}_3 \widetilde{y}_{n,m} \big\rangle_{\mathcal{V}_1} = \big\langle \mathcal{L}_1 \widetilde{y}_{n,m}, \widehat{\mathcal{A}}_{\alpha_n} \widetilde{y}_{n,m} \big\rangle_{\mathcal{V}_1} \geq 0$, it follows that

$$
-C\,\|\mathcal{L}_3 \widetilde{y}_{n,m}\|_{\mathcal{H}} + \|\mathcal{L}_3 \widetilde{y}_{n,m}\|_{\mathcal{H}}^2 + \frac{1}{2}\|\widetilde{y}_{n,m}(T)\|_H^2 \leq \|\widetilde{u}_{n,m}\|_{\mathcal{H}} \|\widetilde{y}_{n,m}\|_{\mathcal{H}} + \frac{1}{2}\|\widetilde{y}_{n,m}(0)\|_{\mathcal{H}}^2.
$$

Since the right-hand side is bounded, we conclude the boundedness of $(\mathcal{L}_3 \widetilde{y}_{n,m})$ in $\mathcal{H}$ and therefore of $(\widetilde{y}_{n,m})$ in $\mathcal{V}_2$. Thus, also (y_n) remains bounded in $\mathcal{V}_2$. Applying interpolation arguments, it can be shown that $L^2(\mathcal{T}; V_2) \cap H^1(\mathcal{T}; V_1^*)$ continuously embeds into $C(\overline{\mathcal{T}}; V_{1/2})$ for a Hilbert space $V_{1/2}$ that is compactly embedded in H (cf., e.g., [6]). This means that $y \mapsto y(T)$ is a compact mapping from $L^2(\mathcal{T}; V_2) \cap H^1(\mathcal{T}; V_1^*)$ into H. Consequently, after passing to a subsequence (y_l) of (y_m) we have that $y_l(T)$ converges strongly in H to $y(T)$. This completes the proof. ∎

PROPOSITION 3.11. *Let (α_n) be a sequence in Λ with $\alpha_n \to 0$, and let $(y_n, u_n) \subset \mathcal{W}_1 \times \mathcal{C}$ be a sequence of minimizers to* (P_{α_n}). *Then there exist subsequences (denoted by the index m) and $(y, u) \in \mathcal{W}_1 \times \mathcal{C}$, which is a minimizer of* (P)*, such that*

$$
\begin{aligned}
(y_m, \widehat{\mathcal{A}}_{\alpha_m} y_m) &\rightharpoonup (y, \xi) && \text{in } \mathcal{W}_1 \times \mathcal{H}, \\
(y_m, y_m(T), u_m) &\to (y, y(T), u) && \text{in } \mathcal{V}_1 \times H \times \mathcal{H}.
\end{aligned}
$$

Proof. For fixed $v \in \mathcal{H}$ the sequence $S_{\alpha_n} v$ converges weakly in $\mathcal{W}_1$ to Sv by Proposition 3.8. In particular, $(S_{\alpha_n} v)$ is bounded in $\mathcal{W}_1$, and hence it holds that

$$
J(y_n, u_n) \leq J(S_{\alpha_n} v, v) \leq C
$$

for some constant C, because (y_n, u_n) is a minimizer of (P_{α_n}). The coercivity of J implies that (u_n) is bounded in $\mathcal{H}$. Applying Lemma 3.10 guarantees the existence of sequences with index m such that $(y_m, \widehat{\mathcal{A}}_{\alpha_m} y_m, u_m) \rightharpoonup (y, \xi, u)$ in $\mathcal{W}_1 \times \mathcal{H} \times \mathcal{H}$ and $(y_m, y_m(T)) \to (y, y(T))$ in $\mathcal{V}_1 \times H$. Note that (y, u) satisfies $u \in \mathcal{C}$ and $Su = y$ by Proposition 3.8.

Now, let $(y^*, u^*) \in \mathcal{W}_1 \times \mathcal{C}$ be a minimizer of (P) (which exists by virtue of Proposition 3.9). From Proposition 3.8 and Theorem 3.3 we conclude that $(S_{\alpha_m} u^*, u^*) \rightharpoonup$

(y^*, u^*) in $\mathcal{W}_1 \times \mathcal{H}$ and $S_{\alpha_m} u^* \to y^*$ in $\mathcal{V}_1$ for a suitable subsequence. By Lemma 3.10 and after passing to another subsequence, which we still denote by the index m, we have that $S_{\alpha_m} u^*(T) \to y^*(T)$ in H. The weak lower semicontinuity of J and the convergence properties of $(S_{\alpha_m} u^*, u^*)$ imply that

$$J(y^*, u^*) \le J(y, u) \le \varliminf_{m\to\infty} J(y_m, u_m) \le \varlimsup_{m\to\infty} J(y_m, u_m) \le \varlimsup_{m\to\infty} J(S_{\alpha_m} u^*, u^*) \\ = J(y^*, u^*).$$

Consequently, (y, u) is a minimizer of (P). Moreover, we have that $J(y_m, u_m) \to J(y, u)$ which implies $||u_m||_{\mathcal{H}} \to ||u||_{\mathcal{H}}$. Since we already know that $u_m \rightharpoonup u$ in $\mathcal{H}$, it follows that $u_m \to u$ in $\mathcal{H}$. ∎

3.2. Stationarity system for $(\mathbf{P}_\alpha)$.

THEOREM 3.12. *For every $\alpha \in \Lambda$ and for every minimizer $(y_\alpha, u_\alpha) \in \mathcal{W}_1 \times \mathcal{H}$ of problem (P_α), there exists a $p_\alpha \in \mathcal{W}_3$ with*

$$-p_\alpha' + \mathcal{L}_1\mathcal{L}_3 p_\alpha + \mathcal{R}_\alpha^* \mathcal{L}_3 p_\alpha = \mu_1(y_\alpha - y_\Omega),$$
$$p_\alpha(T) = \mu_2(y_\alpha(T) - y_T),$$

where $\mathcal{R}_\alpha := \overline{D\mathcal{A}_\alpha}(y_\alpha) + \eta\mathcal{I} \in \mathcal{L}(\mathcal{V}_1; \mathcal{V}_1)$. Moreover, for $\mathcal{C}_\alpha := \mathbb{R}^+(\mathcal{C} - u_\alpha)$ we have

$$p_\alpha + u_\alpha \in \mathcal{C}_\alpha^+.$$

Proof. First, recall that by Remark 3.4 a minimizer $(y_\alpha, u_\alpha) \in \mathcal{W}_1 \times \mathcal{H}$ satisfies $y_\alpha \in \mathcal{W}_4$, and therefore $\mathcal{R}_\alpha$ is indeed a linear, bounded operator on $\mathcal{V}_1$ by Assumption 3.1(iv).

In order to prove the assertion, we first apply a theorem of Zowe and Kurcyusz [63] which guarantees the existence of a Lagrange multiplier p_α satisfying a particular partial differential equation. In the second step we show the relation between this multiplier and the control u_α, and in the last step we prove that p_α indeed is a solution to the evolution equation above.

1. For $v_0 \in V_1$ we define $\mathcal{W}_3(v_0) := \{v \in \mathcal{W}_3 : v(0) = v_0\}$ and consider the following setting in order to apply Theorem 3.1 of [63]:

$$X := \mathcal{W}_3 \times \mathcal{H}, \qquad C_X := \mathcal{W}_3(y_0) \times \mathcal{C},$$
$$Y := \mathcal{V}_1^*, \qquad K := \{0\} \subset Y,$$
$$g : X \to Y, \qquad g(y, u) := y' + \mathcal{L}_1(\mathcal{L}_3 + \widehat{\mathcal{A}}_\alpha + \eta\mathcal{I})y - u,$$
$$f : X \to \mathbb{R}, \qquad f(y, u) := J(y, u).$$

From Assumption 3.1 it follows that f and g are continuously Fréchet differentiable with

$$Dg(y, u; \delta_y, \delta_u) = \delta_y{}' + \mathcal{L}_1(\mathcal{L}_3 + D\widehat{\mathcal{A}}_\alpha(y) + \eta\mathcal{I})\delta_y - \delta_u,$$
$$Df(y, u; \delta_y, \delta_u) = \mu_1(y - y_\Omega \,|\, \delta_y)_{\mathcal{H}} + \mu_2\big(y(T) - y_T \,\big|\, \delta_y(T)\big)_H + (u \,|\, \delta_u)_{\mathcal{H}}$$

for $(\delta_y, \delta_u) \in \mathcal{W}_3 \times \mathcal{H}$. For given $v^* \in Y$ we choose $u := 0 \in \mathcal{H}$ and $y \in \mathcal{W}_1$ to be the solution of

$$y' + \mathcal{L}_1 w = v^*, \quad w = (\mathcal{L}_1 + \mathcal{R}_\alpha)y, \quad y(0) = 0.$$

Due to Theorem 2.7 such a solution exists since $\mathcal{R}_\alpha : \mathcal{V}_1 \to \mathcal{V}_1^*$ is Lipschitz continuous. Moreover, Theorem 2.10 shows that $y \in \mathcal{W}_3$ as well as

$$y' + \mathcal{L}_1(\mathcal{L}_3 + \mathcal{R}_\alpha)y - u = v^*.$$

This demonstrates that $g'(\widetilde{x}) \in \mathcal{L}(\mathcal{W}_3(0) \times \{0\}; Y)$ is surjective for $\widetilde{x} := (y_\alpha, u_\alpha)$. Now, using Theorem 3.1 of Zowe and Kurcyusz [63], there is a $\widehat{p_\alpha} \in Y^*$ such that

$$\left\langle \widehat{p_\alpha}, g(\widetilde{x}) \right\rangle_Y = 0, \quad Df(\widetilde{x}; \delta_y, \delta_u) - \left\langle \widehat{p_\alpha}, Dg(\widetilde{x}; \delta_y, \delta_u) \right\rangle_Y = 0$$

for all $(\delta_y, \delta_u) \in X_0 := \mathcal{W}_3(0) \times \mathcal{C}_\alpha$. Since $Y^* \cong \mathcal{V}_1$, the multiplier $\widehat{p_\alpha} \in Y^*$ can be identified with a $p_\alpha \in \mathcal{V}_1$ satisfying

$$\left\langle\, y_\alpha' + \mathcal{L}_1(\mathcal{L}_3 + \widehat{\mathcal{A}}_\alpha + \eta\mathcal{I})y_\alpha - u_\alpha \,,\, p_\alpha \,\right\rangle_{\mathcal{V}_1} \;=\; 0, \tag{3.1}$$

$$DJ(\widetilde{x}; \delta_y, \delta_u) - \left\langle\, {\delta_y}' + \mathcal{L}_1(\mathcal{L}_3 + \mathcal{R}_\alpha)\delta_y - \delta_u \,,\, p_\alpha \,\right\rangle_{\mathcal{V}_1} = 0 \tag{3.2}$$

for all $(\delta_y, \delta_u) \in X_0$.

2. By choosing $\delta_y = 0$ in (3.2) and since $DJ(\widetilde{x}; 0, \delta_u) = (u_\alpha \,|\, \delta_u)_{\mathcal{H}}$ we find

$$DJ(\widetilde{x}; 0, \delta_u) + \langle \delta_u \,, p_\alpha \rangle_{\mathcal{V}_1} \;=\; (u_\alpha + p_\alpha \,|\, \delta_u)_{\mathcal{H}} = 0$$

for all $\delta_u \in \mathcal{C}_\alpha$. This yields $p_\alpha + u_\alpha \in \mathcal{C}_\alpha^+ \subset \mathcal{H}$.

3. Now we show that (3.1) and (3.2) imply the assertion on p_α. For this purpose, first recall that a function $z \in L^2(\mathcal{T}; Z)$ has a weak derivative $v \in L^2(\mathcal{T}; Z)$ for some Banach space Z if and only if

$$\int_{\mathcal{T}} \eta v \;=\; -\int_{\mathcal{T}} \eta' z \quad \forall\, \eta \in C_c^\infty(\mathcal{T}).$$

From (3.2) and the symmetry of $\mathcal{L}_1$ it follows that for all $(\delta_y, \delta_u) \in X_0$ it holds that

$$\begin{aligned}\left\langle {\delta_y}', p_\alpha \right\rangle_{\mathcal{V}_1} &= DJ(\widetilde{x}; \delta_y, \delta_u) - \left\langle \mathcal{L}_1 p_\alpha \,, (\mathcal{L}_3 + \mathcal{R}_\alpha)\delta_y \right\rangle_{\mathcal{V}_1} + \langle \delta_u \,, p_\alpha \rangle_{\mathcal{V}_1} \\ &= DJ(\widetilde{x}; \delta_y, \delta_u) - \left\langle (\mathcal{L}_3 + \mathcal{R}_\alpha)^* \mathcal{L}_1 p_\alpha \,, \delta_y \right\rangle_{\mathcal{V}_3} + \langle \delta_u \,, p_\alpha \rangle_{\mathcal{V}_1}.\end{aligned}$$

Now, for arbitrary $v_0 \in V_3$ and $\eta \in C_c^\infty(\mathcal{T})$, by choosing $\delta_y(t) := \eta(t)v_0$, $\delta_u := 0$ and defining $q := \mu_1(y_\alpha - y_\Omega) - (\mathcal{L}_3 + \mathcal{R}_\alpha)^* \mathcal{L}_1 p_\alpha \in \mathcal{V}_3^*$ we obtain

$$\left\langle \eta' v_0 \,, p_\alpha \right\rangle_{\mathcal{V}_1} = \langle q \,, \eta v_0 \rangle_{\mathcal{V}_3},$$

and therefore

$$\left\langle \int_{\mathcal{T}} \eta' p_\alpha - \int_{\mathcal{T}} \eta q, v_0 \right\rangle_{V_3} \;=\; \left\langle \eta' v_0 \,, p_\alpha \right\rangle_{\mathcal{V}_1} - \langle q \,, \eta v_0 \rangle_{\mathcal{V}_3} = 0.$$

Since $v_0 \in V_3$ was arbitrary, $\int_{\mathcal{T}} \eta' p_\alpha = \int_{\mathcal{T}} \eta q$ holds for every $\eta \in C_c^\infty(\mathcal{T})$. This implies that $p_\alpha \in H^1(\mathcal{T}; V_3^*)$ and

$$p_\alpha' \;=\; (\mathcal{L}_3 + \mathcal{R}_\alpha)^* \mathcal{L}_1 p_\alpha - \mu_1(y_\alpha - y_\Omega).$$

Therefore, in the case $\delta_u = 0$, (3.2) reduces to

$$\left\langle {\delta_y}', p_\alpha \right\rangle_{\mathcal{V}_1} + \left\langle p_\alpha', \delta_y \right\rangle_{\mathcal{V}_3} - \mu_2 \big(y_\alpha(T) - y_T \,\big|\, \delta_y(T) \big)_H = 0.$$

With the help of Proposition 2.9 we infer that

$$\big\langle p_\alpha(T), \delta_y(T) \big\rangle_{V_1} - \big\langle p_\alpha(0), \delta_y(0) \big\rangle_{V_1} - \mu_2 \big\langle y_\alpha(T) - y_T, \delta_y(T) \big\rangle_{V_1} = 0$$

for all $\delta_y \in \mathcal{W}_3(0)$. Since it is possible to find a sequence $\delta_{y_n} \in \mathcal{W}_3(0)$ such that $\delta_{y_n}(T) \to v_1$ in V_1 for arbitrarily given $v_1 \in V_1$, we conclude that

$$p_\alpha(T) \;=\; \mu_2(y_\alpha(T) - y_T).$$

Finally, applying Theorem 2.10 again finishes the proof. ▯

LEMMA 3.13. *Let $y \in \mathcal{W}_3$, $\alpha \in \Lambda$. Then $\mathcal{I}_1 \overline{D\mathcal{A}_\alpha}(y) : \mathcal{V}_1 \to \mathcal{V}_1^*$ is monotone.*

Proof. Let $v \in \mathcal{V}_1$ be given. We have to show that $\big\langle \overline{D\mathcal{A}_\alpha}(y)v, v \big\rangle_{\mathcal{V}_1} \geq 0$. Since the image of $I_{\mathcal{W}_3 \to \mathcal{V}_1}$ is dense in $\mathcal{V}_1$, we may find a sequence (y_n) in $\mathcal{W}_3$ with $y_n \to v$ in $\mathcal{V}_1$. By the continuity of $\overline{D\mathcal{A}_\alpha}(y)$ and the Fréchet differentiability of $\widehat{\mathcal{A}}_\alpha : \mathcal{W}_3 \to \mathcal{V}_1$ in y, it holds that

$$\begin{aligned}
\big\langle \overline{D\mathcal{A}_\alpha}(y)v, v \big\rangle_{\mathcal{V}_1} &= \lim_{n\to\infty} \big\langle D\widehat{\mathcal{A}}_\alpha(y) y_n, y_n \big\rangle_{\mathcal{V}_1} \\
&= \lim_{n\to\infty} \lim_{t\to 0} \frac{1}{t} \big\langle \widehat{\mathcal{A}}_\alpha(y + t y_n) - \widehat{\mathcal{A}}_\alpha y, y_n \big\rangle_{\mathcal{V}_1} \\
&= \lim_{n\to\infty} \lim_{t\to 0} \frac{1}{t^2} \big\langle \mathcal{A}_\alpha(y + t y_n) - \mathcal{A}_\alpha y, (y + t y_n) - y \big\rangle_{\mathcal{V}_1} \\
&\geq 0
\end{aligned}$$

since $A_\alpha = \partial \varphi_\alpha$, which is monotone (cf. [10, Proposition 2.1]). ▯

3.3. The limit problem. We conclude this section by establishing an optimality system for the limit problem. It involves the notion of derivatives of vector-valued distributions which we recall first (cf. [29] and [12] for further details).

DEFINITION 3.14. *Let $\mathcal{D}(\mathcal{T}) := C_c^\infty(\mathcal{T})$ denote the space of test functions with the usual locally convex topology. For a Banach space X, the space of X-valued distributions on $\mathcal{T}$ is the space of all linear, continuous mappings $\mathcal{D}^*(\mathcal{T}; X) := \mathcal{L}(\mathcal{D}(\mathcal{T}); X_w)$ from $\mathcal{D}(\mathcal{T})$ into X equipped with its weak topology X_w. The topology of $\mathcal{D}^*(\mathcal{T}; X)$ is chosen to be the topology induced by the system $\{p_{w,x^*} \mid w \in \mathcal{D}(\mathcal{T}), x^* \in X^*\}$ of seminorms given for $v \in \mathcal{D}^*(\mathcal{T}; X)$ by*

$$p_{w,x^*}(v) \;:=\; |\big\langle x^*, v(w) \big\rangle_X|.$$

Moreover, the derivative $' : \mathcal{D}^*(\mathcal{T}; X) \to \mathcal{D}^*(\mathcal{T}; X)$ of X is defined for $w \in \mathcal{D}(\mathcal{T})$ by

$$v'(w) \;:=\; -v(w').$$

Remark 3.15. As usual, a locally integrable function $f \in L^1_{\mathrm{loc}}(\mathcal{T}; X)$ can be identified with the distribution $F_f \in \mathcal{D}^*(\mathcal{T}; X)$ given by

$$F_f(w) \;:=\; \int_\Omega w f \qquad \text{for } w \in \mathcal{D}(\mathcal{T}).$$

THEOREM 3.16. *Let (α_n) be a sequence in Λ with $\alpha_n \to 0$. We assume that $(y_n, u_n) \in \mathcal{W}_3 \times \mathcal{C}$ is a sequence of minimizers to problem (P_{α_n}), that $p_n \in \mathcal{W}_3$ and $\mathcal{C}_\alpha \subset \mathcal{H}$ are given as in Theorem 3.12, $w_n := \mathcal{L}_1^{-1}(u_n - y_n')$, $\xi_n := \widehat{\mathcal{A}}_{\alpha_n} y_n \in \mathcal{V}_1$, $\lambda_n :=$*

$\overline{D\mathcal{A}_{\alpha_n}}(y_n)^*\mathcal{L}_3 p_n \in \mathcal{V}_1^*$, *and* $\kappa_n := \eta\mathcal{I}_1\mathcal{L}_3 p_n \in \mathcal{V}_1^*$. *Then there exist subsequences (denoted by the index* m*) and a minimizer* $(y,u) \in \mathcal{W}_1 \times \mathcal{C}$ *of* (P) *such that*

$$\begin{aligned}
(y_m, w_m, p_m, p_m', p_m(0)) &\rightharpoonup (y, w, p, p', p_0) && \text{in } \mathcal{W}_1 \times \mathcal{V}_1 \times \mathcal{V}_3 \times \mathcal{W}_1^* \times V_1,\\
(\lambda_m, \xi_m, \kappa_m) &\rightharpoonup (\lambda, \xi, \kappa) && \text{in } \mathcal{W}_1^* \times \mathcal{H} \times \mathcal{V}_1^*,\\
(y_m, u_m) &\to (y, u) && \text{in } \mathcal{V}_1 \times \mathcal{H},\\
p_m &\overset{*}{\rightharpoonup} p && \text{in } L^\infty(\mathcal{T}; V_1),
\end{aligned}$$

where p' *denotes the distributional derivative of* $p \in \mathcal{D}^*(\mathcal{T}; V_1)$*. It holds that* $y_n(0) = y_0$, $p_n(T) = \mu_2(y_n(T) - y_T)$, $p_n + u_n \in \mathcal{C}_\alpha^+$,

$$y_n' + \mathcal{L}_1 w_n = u_n, \qquad w_n = \mathcal{L}_3 y_n + \xi_n + \eta\mathcal{I} y_n, \tag{3.3}$$

$$-p_n' + \mathcal{L}_1\mathcal{L}_3 p_n + \lambda_n + \kappa_n = \mu_1(y_n - y_\Omega), \tag{3.4}$$

as well as $y(0) = y_0$, $p(T) = \mu_2(y(T) - y_T)$, $p + u \in \mathcal{C}_0^+$ *for* $\mathcal{C}_0 := \mathbb{R}^+(\mathcal{C} - u)$*, and*

$$y' + \mathcal{L}_1 w = u, \qquad w = \mathcal{L}_1 y + \xi + \eta\mathcal{I}_1 y, \tag{3.5}$$

$$-p' + \mathcal{L}_1\mathcal{L}_3 p + \lambda + \kappa = \mu_1(y - y_\Omega). \tag{3.6}$$

Proof. Since (y_n) and y are solutions to (P_{α_n}) and (P), respectively, (3.3) and (3.5) are satisfied together with the corresponding initial conditions. Theorem 3.12 guarantees (3.4) and $p_n(T) = \mu_2(y_n(T) - y_T)$. The boundedness and convergence properties of (y_m), (w_m), (ξ_m), and (u_m) are obtained from Corollary 3.7 and Proposition 3.11. Let $\eta_- := \max(1, -\eta)$. Then

$$\langle \eta\mathcal{I}_1 v, v\rangle_{\mathcal{V}_1} \;\geq\; -\eta_- \|v\|_{\mathcal{H}}^2$$

holds for all $v \in \mathcal{V}_1$. For the time being, let us fix $t_0 \in \overline{\mathcal{T}}$ and define $\varphi : \overline{\mathcal{T}} \to \mathbb{R}$, $\varphi(t) := \exp(\eta_-^2 t)$ together with $\widetilde{p_n} := \varphi p_n \in \mathcal{W}_3$, $\widetilde{q_n} := \chi_{[t_0,T]}\widetilde{p_n} \in \mathcal{V}_3$, and $f_n := \varphi(y_n - y_\Omega) \in \mathcal{H}$. Using (3.4) we obtain that

$$\begin{aligned}
\widetilde{p_n}' &= \varphi\Big(\eta_-^2 p_n + p_n'\Big)\\
&= \varphi\Big(\eta_-^2 p_n + \Big[\mathcal{L}_1 + \overline{D\mathcal{A}_{\alpha_n}}(y_n)^* + \eta\mathcal{I}_1\Big]\mathcal{L}_3 p_n\Big) - f_n\\
&= \eta_-^2\widetilde{p_n} + \Big[\mathcal{L}_1 + \overline{D\mathcal{A}_{\alpha_n}}(y_n)^* + \eta\mathcal{I}_1\Big]\mathcal{L}_3\widetilde{p_n} - f_n.
\end{aligned}$$

The integration-by-parts formula yields

$$\Big\langle \widetilde{p_n}', \mathcal{L}_3\widetilde{q_n}\Big\rangle_{\mathcal{V}_1} \;=\; \big\langle (\mathcal{I}_3^*\mathcal{L}_1\mathcal{I}_3\widetilde{p_n})', \widetilde{q_n}\big\rangle_{\mathcal{V}_3} \;=\; \frac{1}{2}\Big(\|\widetilde{p_n}(T)\|_{V_1}^2 - \|\widetilde{p_n}(t_0)\|_{V_1}^2\Big),$$

and therefore, after testing $\widetilde{p_n}'$ with $\mathcal{L}_3\widetilde{q_n}$, we obtain with the help of Lemma 3.13 that

$$\begin{aligned}
&\frac{1}{2}\Big(\|\widetilde{p_n}(T)\|_{V_1}^2 - \|\widetilde{p_n}(t_0)\|_{V_1}^2\Big)\\
&\quad = \Big\langle \widetilde{p_n}', \mathcal{L}_3\widetilde{q_n}\Big\rangle_{\mathcal{V}_1}\\
&\quad \geq \eta_-^2\big\langle \widetilde{p_n}, \mathcal{L}_3\widetilde{q_n}\big\rangle_{\mathcal{V}_1} + \big\langle \mathcal{L}_1\mathcal{L}_3\widetilde{p_n}, \mathcal{L}_3\widetilde{q_n}\big\rangle_{\mathcal{V}_1} - \eta_-\|\mathcal{L}_3\widetilde{q_n}\|_{\mathcal{H}}^2 - \|f_n\|_{\mathcal{V}_1^*}\|\mathcal{L}_3\widetilde{q_n}\|_{\mathcal{V}_1}\\
&\quad = \eta_-^2\|\widetilde{q_n}\|_{\mathcal{V}_1}^2 + \|\widetilde{q_n}\|_{\mathcal{V}_3}^2 - \eta_-\|\mathcal{L}_3\widetilde{q_n}\|_{\mathcal{H}}^2 - \|f_n\|_{\mathcal{V}_1^*}\|\widetilde{q_n}\|_{\mathcal{V}_3}.
\end{aligned}$$

With the estimation

$$
\begin{aligned}
||\mathcal{L}_3\widetilde{q_n}||_{\mathcal{H}}^2 &= \big\langle \mathcal{I}_1\mathcal{L}_3\widetilde{q_n}, \mathcal{L}_3\widetilde{q_n}\big\rangle_{\mathcal{V}_1} = \big\langle \mathcal{L}_3^*\mathcal{I}_1\mathcal{L}_3\widetilde{q_n}, \widetilde{q_n}\big\rangle_{\mathcal{V}_1} = \big\langle \mathcal{L}_1\mathcal{L}_3\widetilde{q_n}, \mathcal{I}_3\widetilde{q_n}\big\rangle_{\mathcal{V}_1} \\
&\le ||\mathcal{L}_1\mathcal{L}_3\widetilde{q_n}||_{\mathcal{V}_1^*}||\widetilde{q_n}||_{\mathcal{V}_1} \le \frac{1}{2\eta_-}||\widetilde{q_n}||_{\mathcal{V}_3}^2 + \frac{\eta_-}{2}||\widetilde{q_n}||_{\mathcal{V}_1}^2
\end{aligned}
$$

we finally arrive at

$$
||\widetilde{p_n}(T)||_{V_1}^2 \ge ||\widetilde{p_n}(t_0)||_{V_1}^2 + \eta_-^2||\widetilde{q_n}||_{\mathcal{V}_1}^2 + ||\widetilde{q_n}||_{\mathcal{V}_3}^2 - 2||f_n||_{\mathcal{V}_1^*}||\widetilde{q_n}||_{\mathcal{V}_3}.
$$

By Corollary 3.7, y_n is bounded in $L^\infty(\mathcal{T}; V_1)$. Hence $p_n(T) = \mu_2(y_n(T) - y_T)$ in V_1. Since (f_n) is bounded in $\mathcal{V}_1^*$, it follows that $(\widetilde{p_n})$ and thus also (p_n) are bounded in $\mathcal{V}_3$ (by choosing $t_0 = 0$) as well as $(p_n(t_0))$ in V_1 . For $v \in \mathcal{W}_1$ we have that

$$
\begin{aligned}
\big\langle p_n', v\big\rangle_{\mathcal{W}_1} = \big\langle p_n', v\big\rangle_{\mathcal{V}_1} &= -\big\langle v', p_n\big\rangle_{\mathcal{V}_1} + \big(p_n(T)\big|v(T)\big)_H - \big(p_n(0)\big|v(0)\big)_H \\
&\le C\Big(||p_n||_{\mathcal{V}_1} + ||p_n(T)||_H + ||p_n(0)||_H\Big)||v||_{\mathcal{W}_1},
\end{aligned}
$$

which shows that p_n' is bounded in $\mathcal{W}_1^*$. The boundedness of (p_n) in $\mathcal{V}_3$, together with Assumption 3.1, implies that (κ_n) is bounded in $\mathcal{V}_1^*$. Moreover, the embedding of $\mathcal{V}_1^*$ into $\mathcal{W}_1^*$ is continuous, and hence it follows from (3.4) that also (λ_n) is bounded in $\mathcal{W}_1^*$.

Hence, we can pass to a (weakly) convergent subsequence (denoted by the index m) such that all the convergences given in the assertion are fulfilled. Passing to the limit as $m \to \infty$ in (3.4) yields (3.6). To show that p' is indeed the distributional derivative of p, we observe that

$$
\begin{aligned}
\big\langle p'(\varphi), v\big\rangle_{V_1} = -\big\langle p(\varphi'), v\big\rangle_{V_1} &= -\Big\langle \int_{\mathcal{T}} \varphi' p, v\Big\rangle_{V_1} = -\big\langle p, \varphi' v\big\rangle_{\mathcal{V}_1} \\
&= -\lim \big\langle p_m, \varphi' v\big\rangle_{\mathcal{V}_1} = \lim \big\langle p_m', \varphi v\big\rangle_{\mathcal{V}_1} = \lim \big\langle p_m', \varphi v\big\rangle_{\mathcal{W}_1} \\
&= \big\langle \lim p_m', \varphi v\big\rangle_{\mathcal{W}_1}
\end{aligned}
$$

for all $\varphi \in \mathcal{D}(\mathcal{T})$ and $v \in V_1$. This completes the proof. ∎

4. The Cahn–Hilliard system with double-obstacle homogeneous free energy density. In this section we highlight the special case where φ is given as the indicator function of some convex subset of V_1. This corresponds to the Cahn–Hilliard system with double-obstacle potential. Moreover, the φ_α are defined as mollified versions of the Moreau–Yosida approximations of φ. In this setting, a function space version of C-stationarity is obtained.

Example 4.1 (double-obstacle potential). Suppose that $N \le 3$, $M = [a, b] \subset \mathbb{R}$, $a < 0 < b$, is a bounded interval and

$$
\begin{aligned}
&K_H := \{v \in H : v(x) \in M \text{ for a.e. } x \in \Omega\}, \quad K_V := I_{V_1}^{-1} K_H, \\
&\mathcal{C} \subset \mathcal{H} \quad \text{closed, convex, and nonempty}, \\
&\tilde{\gamma} := \imath_M : \mathbb{R} \to \overline{\mathbb{R}}, \qquad \tilde{\beta} := \partial\tilde{\gamma}, \\
&\varphi := \imath_{K_V} : V_1 \to \overline{\mathbb{R}}, \qquad \eta := -1, \quad \psi = \frac{1}{2}||\ .\ ||_H^2 : V_1 \to \overline{\mathbb{R}}.
\end{aligned}
$$

Let $\rho \in C^2(\mathbb{R})$ be a fixed mollifier with $\operatorname{supp}\rho \subset [-1, 1]$, $\int_{\mathbb{R}} \rho = 1$, $\overline{\rho} := \int_{\mathbb{R}} r\rho(r)\,dr$, $0 \le \rho(r) \le 1$ for all $r \in \mathbb{R}$. Moreover, $\varepsilon : \mathbb{R}^+ \to \mathbb{R}^+$ is a function with $\varepsilon(\alpha) > 0$, $\varepsilon(\alpha) \to 0$, and $\frac{\varepsilon(\alpha)}{\alpha} \to 0$ as $\alpha \to 0$. By $\tilde{\gamma}_\alpha$ and $\tilde{\varphi_\alpha}$ we denote the Moreau–Yosida approximations of $\tilde{\gamma}$ and φ with the parameter $\alpha > 0$, respectively, and the Yosida approximation of the operator $\tilde{\beta}$ by $\tilde{\beta}_\alpha$. For the general definitions of the Moreau–Yosida and the Yosida approximation we refer the reader to [10]. We set

$$\rho_\varepsilon(r) := \frac{1}{\varepsilon}\rho\Big(\frac{r}{\varepsilon}\Big), \qquad \beta_\alpha := \tilde{\beta}_\alpha * \rho_{\varepsilon(\alpha)}, \qquad \gamma_\alpha(r) := \int_0^r \beta_\alpha, \qquad \varphi_\alpha(u) := \int_\Omega \gamma_\alpha \circ u,$$

where $*$ denotes the convolution operator and $f \circ g$ denotes the superposition of f and g. Furthermore, $q : \mathbb{R} \to \mathbb{R}$ is given by $q := \tilde{\beta}_1 = \alpha\tilde{\beta}_\alpha$, $Q : H \to H$ denotes its superposition operator with respect to Ω, and $\mathcal{Q} : \mathcal{H} \to \mathcal{H}$ denotes the superposition operator of Q with respect to $\mathcal{T}$. Finally, let $y_0 \in V_1 \cap L^\infty(\Omega)$ be such that $a < \operatorname{ess\,inf}_\Omega y_0 \le \operatorname{ess\,sup}_\Omega y_0 < b$.

In order to prove that Example 4.1 falls into the framework of Assumption 3.1, we first collect some basic properties.

Lemma 4.2. *Given the setting of Example* 4.1 *we have*

$$\tilde{\beta}_\alpha(r) = \begin{cases} \frac{1}{\alpha}(r-a) & \text{if } r < a, \\ 0 & \text{if } a \le r \le b, \\ \frac{1}{\alpha}(r-b) & \text{if } b < r, \end{cases} \qquad \tilde{\beta}'_\alpha(r) = \begin{cases} \frac{1}{\alpha} & \text{if } r < a, \\ 0 & \text{if } a < r < b, \\ \frac{1}{\alpha} & \text{if } b < r, \end{cases}$$

$$\tilde{\gamma}_\alpha(r) = \begin{cases} \frac{1}{2\alpha}(r-a)^2 & \text{if } r < a, \\ 0 & \text{if } a \le r \le b, \\ \frac{1}{2\alpha}(r-b)^2 & \text{if } b < r. \end{cases}$$

Moreover, $\tilde{\beta}_\alpha$ and β_α are Lipschitz continuous with constant $\frac{1}{\alpha}$ and monotone, and $\beta_\alpha(0) = 0$ if $\varepsilon(\alpha) \le \min(-a, b)$. Furthermore, $\beta_\alpha(r) = \tilde{\beta}_\alpha(r) - \frac{\varepsilon(\alpha)}{\alpha}\overline{\rho}$ for $r \le a - \varepsilon(\alpha)$ and $r \ge b + \varepsilon(\alpha)$, and

$$(4.1)\quad |\beta'_\alpha(r)| \le \frac{1}{\alpha}, \quad |\beta_\alpha(r) - \tilde{\beta}_\alpha(r)| \le \frac{\varepsilon(\alpha)}{\alpha}(2 + |\overline{\rho}|), \quad |\beta_\alpha(r) - \beta'_\alpha(r)q(r)| \le C\,\frac{\varepsilon(\alpha)}{\alpha}$$

for all $r \in \mathbb{R}$ and some constant C, which does not depend on α. Finally, if $a_\alpha : L^2(\Omega) \to L^2(\Omega)$, $a_\alpha(u)(x) := \beta_\alpha(u(x))$ denotes the superposition operator of β_α, and $P_H : L^2(\Omega) \to H$ denotes the orthogonal projection of $L^2(\Omega)$ onto H, then it holds that $A_\alpha = P_H a_\alpha$.

Proof. 1. The definition of the Yosida approximation directly provides the formula for $\tilde{\beta}_\alpha$ and its derivative. Hence, by integration we obtain the relation for $\tilde{\gamma}_\alpha$.

2. The Lipschitz continuity and monotonicity of $\tilde{\beta}_\alpha$ and β_α, respectively, are direct consequences of the properties of the Yosida approximation and the convolution with $\rho \ge 0$.

3. The inequality $|\beta'_\alpha(r)| \le \frac{1}{\alpha}$ is a consequence of the definition of β_α and the Lipschitz continuity of $\tilde{\beta}_\alpha$ with constant $\frac{1}{\alpha}$. In order to prove the second estimate of (4.1), we first notice that for an affine function $g(r) := c_1 r + c_2$ it holds that $(g * \rho_{\varepsilon(\alpha)})(r) = g(r) - c_1\varepsilon(\alpha)\overline{\rho}$, as readily seen by a simple calculation. This implies the assertion $\beta_\alpha(r) = \tilde{\beta}_\alpha(r) - \frac{\varepsilon(\alpha)}{\alpha}\overline{\rho}$ for $r \le a - \varepsilon(\alpha)$ and $r \ge b + \varepsilon(\alpha)$ (note that outside the interval $[a - \varepsilon(\alpha), \beta + \varepsilon(\alpha)]$ the convolution $\tilde{\beta}_\alpha * \rho_{\varepsilon(\alpha)}$ only touches affine parts of $\tilde{\beta}_\alpha$). Since $\tilde{\beta}_\alpha$ and β_α are monotone, we conclude for $r \in [a - \varepsilon(\alpha), b + \varepsilon(\alpha)]$

that

$$|\beta_\alpha(r) - \tilde{\beta}_\alpha(r)| \le \max\Big(|\tilde{\beta}_\alpha(b+\varepsilon(\alpha)) - \beta_\alpha(a-\varepsilon(\alpha))|, |\beta_\alpha(b+\varepsilon(\alpha)) - \tilde{\beta}_\alpha(a-\varepsilon(\alpha))|\Big)$$
$$\le \frac{\varepsilon(\alpha)}{\alpha}(2+|\overline{\rho}|).$$

This also holds true for r outside the interval. Now let us show the third inequality $|\beta_\alpha(r) - \beta'_\alpha(r)q(r)| \le C\,\frac{\varepsilon(\alpha)}{\alpha}$. Similar to q we define

$$q_\alpha(r) := \begin{cases} r-(a+2\varepsilon(\alpha)) & \text{if } r \le a+2\varepsilon(\alpha),\\ 0 & \text{if } a+2\varepsilon(\alpha) < r < b-2\varepsilon(\alpha),\\ r-(b-2\varepsilon(\alpha)) & \text{if } r \ge b-2\varepsilon(\alpha) \end{cases}$$

and assume that α is sufficiently small (such that $4\varepsilon(\alpha) < b-a$). We have that $|(q-q_\alpha)(r)| \le 2\varepsilon(\alpha)$, and with $\sigma_\varepsilon(r) := r\rho_\varepsilon(r)$ we find

$$|(\tilde{\beta}'_\alpha * \sigma_{\varepsilon(\alpha)})(r)| \le \frac{1}{\alpha}\int_{\mathbb{R}} |\sigma_{\varepsilon(\alpha)}(r)|\,dr \;=\; \frac{1}{\alpha}\int_{\mathbb{R}} \left|\frac{r}{\varepsilon(\alpha)}\,\rho\!\left(\frac{r}{\varepsilon(\alpha)}\right)\right| dr$$
$$= \frac{\varepsilon(\alpha)}{\alpha}\int_{\mathbb{R}} |s\rho(s)|\,ds \;=\; C\,\frac{\varepsilon(\alpha)}{\alpha}.$$

Furthermore, for an affine function $g(r) := c_1 r + c_2$ and any integrable function f, it holds that $[(f*\rho_\varepsilon)g](r) = [(gf)*\rho_\varepsilon](r) - (f*\sigma_\varepsilon)(r)$. Hence, we obtain

$$\begin{aligned}
|\beta_\alpha(r) - \beta'_\alpha(r)q(r)| &\le |\beta_\alpha(r) - \beta'_\alpha(r)q_\alpha(r)| + 2\frac{\varepsilon(\alpha)}{\alpha}\\
&= \Big|\big[\tilde{\beta}_\alpha * \rho_{\varepsilon(\alpha)} - (q_\alpha\tilde{\beta}'_\alpha) * \rho_{\varepsilon(\alpha)}\big](r) + (\tilde{\beta}'_\alpha * \sigma_{\varepsilon(\alpha)})(r)\Big| + 2\frac{\varepsilon(\alpha)}{\alpha}\\
&\le \Big|\big[(\tilde{\beta}_\alpha - q\tilde{\beta}'_\alpha) * \rho_{\varepsilon(\alpha)}\big](r)\Big| + 2\frac{\varepsilon(\alpha)}{\alpha} + C\,\frac{\varepsilon(\alpha)}{\alpha} + 2\frac{\varepsilon(\alpha)}{\alpha}\\
&= (C+4)\frac{\varepsilon(\alpha)}{\alpha}
\end{aligned}$$

due to $\tilde{\beta}_\alpha = q\tilde{\beta}'_\alpha$ and the fact that the convolution $\beta'_\alpha q_\alpha = (\tilde{\beta}'_\alpha * \rho_{\varepsilon(\alpha)})q_\alpha$ involves only affine parts of q_α.

4. Finally, it is easy to show that for $\alpha > 0$ we have

$$\langle A_\alpha u, v\rangle_{V_1} = \int_\Omega v\beta_\alpha(u)$$

for $u, v \in V_1$. Hence, A_α may be written as $P_H a_\alpha$. ∎

PROPOSITION 4.3. *Assume the setting of Example* 4.1 *is fulfilled. Then there exists a constant* $\alpha_0 > 0$ *such that the functionals* $(\varphi_\alpha)_{\alpha\in\overline{\Lambda}}$*,* $\varphi_0 := \varphi$*, and* ψ *satisfy Assumption* 3.1*.*

Proof. We establish the statements (i)–(vi) of Assumption 3.1 step by step.

(i) The fact that A_α is single-valued is a consequences of Lemma 4.2.

(ii) From the properties of β_α given in Lemma 4.2 it follows that there are $0 < \alpha_0 < 1$ and $r_0 > 0$ such that $\gamma_\alpha(r) \ge \frac{1}{2}r^2$ and $r\beta_\alpha(r) \ge \frac{1}{2}r^2$ for all $0 < \alpha \le \alpha_0$ and

$|r| \geq r_0$. Consequently, it holds that

$$\begin{aligned}\theta_\alpha(v) - \frac{1}{2}||\,.\,||^2_{V_1} &\geq \varphi_\alpha(v) - \psi(v)\\ &\geq \int_{\{|v|\leq r_0\}} \gamma_\alpha(v) - \frac{1}{2}|v|^2\,dx + \int_{\{|v|>r_0\}} \gamma_\alpha(v) - \frac{1}{2}|v|^2\,dx\\ &\geq -\frac{r_0^2}{2}|\Omega|.\end{aligned}$$

This proves the desired boundedness from below.

(iii) Since $\beta_\alpha \in W^{1,\infty}(\Omega)$ it follows that $\beta_\alpha \circ u \in H^1(\Omega)$ for $u \in V_1$. Since $A_\alpha = P_H a_\alpha$, A_α can also be written as $A_\alpha u = P_V(\beta'_\alpha(u))$, where $P_V : H^1(\Omega) \to V_1$ denotes the orthogonal projection of $H^1(\Omega)$ onto V_1 with respect to the H^1-inner product. Consequently, we obtain $A_\alpha(\mathcal{V}_1) \subset \mathcal{V}_1$.

(iv) Next we show that $\widehat{\mathcal{A}}_\alpha : \mathcal{W}_3 \to \mathcal{V}_1$ is continuously Fréchet differentiable. For this purpose, first notice that β_α is twice continuously differentiable with bounded first and second derivatives. Furthermore, by the Sobolev embedding theorem (cf. [5]) the space V_3 is continuously embedded into $W^{1,q_1}(\Omega)$ for arbitrary $q_1 < \infty$ if $N \leq 4$. This implies that $\mathcal{W}_3$ can be continuously embedded into $L^{p_1}(\mathcal{T}; W^{1,q_1}(\Omega))$ for $p_1 = 2$ as well as into $C(\overline{\mathcal{T}}; V_1)$ by Proposition 2.9, and hence in particular into $L^{p_2}(\mathcal{T}; W^{1,q_2}(\Omega))$ for $q_2 = 2$ and arbitrary $p_2 < \infty$. Applying standard interpolation arguments (cf. Triebel [59]), the space $L^{p_1}(\mathcal{T}; W^{1,q_1}(\Omega)) \cap L^{p_2}(\mathcal{T}; W^{1,q_2}(\Omega))$ can be embedded into $L^p(\mathcal{T}; W^{1,q}(\Omega))$ for all $1 \leq p, q < \infty$ with

$$\frac{1}{p} = \frac{\theta}{p_1} + \frac{1-\theta}{p_2}, \qquad \frac{1}{q} = \frac{\theta}{q_1} + \frac{1-\theta}{q_2}$$

for $\theta \in]0,1[$. Let $\beta > 1$ be given and set $\theta := \frac{1}{1+\beta}$, $p_2 := \frac{4\theta\beta}{1-2\theta}$, and $q_1 = \frac{p_2}{\beta^2}$. Then we obtain $1-\theta = \theta\beta$, and hence

$$\frac{1}{p} = \frac{\theta}{2} + \frac{\theta\beta(1-2\theta)}{4\theta\beta} = \frac{1}{4}, \qquad \frac{1}{q} = \frac{\theta\beta^2}{p_2} + \frac{\theta\beta}{2} = \beta\left(\frac{\theta}{2} + \frac{1-\theta}{p_2}\right) = \frac{\beta}{p}.$$

Thus, $p = 4$ and $q = \frac{p}{\beta}$. Since $\beta > 1$ was arbitrary, Proposition 2.12 implies the continuous Fréchet differentiability of $\widehat{\mathcal{A}}_\alpha$.

Next, we prove that for each $y \in \mathcal{W}_4$ there exists an operator $\overline{D\mathcal{A}_\alpha}(y) \in \mathcal{L}(\mathcal{V}_1; \mathcal{V}_1)$ with $\overline{D\mathcal{A}_\alpha}(y)v = D\widehat{\mathcal{A}}_\alpha(y)v$ for all $v \in \mathcal{W}_3$. This assertion is equivalent to saying that $D\widehat{\mathcal{A}}_\alpha(y) \in \mathcal{L}(\mathcal{W}_3; \mathcal{V}_1)$ can be continuously extended to $\mathcal{V}_1$, or equivalently, that there exists a constant $C \geq 0$ with

$$||D\widehat{\mathcal{A}}_\alpha(y)v||_{\mathcal{V}_1} \leq C||v||_{\mathcal{V}_1}$$

for all $v_1 \in \mathcal{W}_3$. Let us assume that $y \in \mathcal{W}_4$ and $v \in \mathcal{W}_3$. We have

$$||D\widehat{\mathcal{A}}_\alpha(y)v||^2_{\mathcal{V}_1} \leq ||\nabla(\beta'_\alpha(y)v)||^2_{\mathcal{H}} \leq 2\Big(||\beta''_\alpha(y)v\nabla y||^2_{\mathcal{H}} + ||\beta'_\alpha(y)\nabla v||^2_{\mathcal{H}}\Big)$$

and want to show that both terms on the right-hand side are bounded from above by $C\,||v||_{\mathcal{V}_1}$. Since β'_α and β''_α have bounded range, it remains to be checked that $||v\nabla y||_{\mathcal{H}} \leq C\,||v||_{\mathcal{V}_1}$. For this purpose, notice that $\mathcal{W}_4$ can be continuously embedded into $C(\overline{\mathcal{T}}; V_2)$; in particular $\nabla y \in L^\infty(\mathcal{T}; H^1(\Omega))$. By the Sobolev embedding theorem and since $N \le 3$, $H^1(\Omega)$ embeds continuously into $L^4(\Omega)$. Therefore, it follows that

$$\begin{aligned} ||v\nabla y||^2_{\mathcal{H}} &= \int_{\mathcal{T}} ||v\nabla y||^2_H \;\le\; \int_{\mathcal{T}} ||v||^2_{L^4(\Omega)} ||\nabla y||^2_{L^4(\Omega)} \\ &\le C\, ||\nabla y||^2_{L^\infty(\mathcal{T};L^4(\Omega))}\, ||v||^2_{L^2(\mathcal{T};L^4(\Omega))} \\ &\le C\, ||v||^2_{\mathcal{V}_1}. \end{aligned}$$

(v) Let $\alpha \in \Lambda$ be given. Since $\mathcal{A}_\alpha u = P_V(\beta_\alpha(u))$ and $0 \le \beta'_\alpha(r)$ for all $r \in \mathbb{R}$, we have that

$$\langle \mathcal{L}_1 y, \widehat{\mathcal{A}}_\alpha y\rangle_{\mathcal{V}_1} \;=\; \int_\Omega \nabla y \cdot \nabla(\beta_\alpha(y)) \;=\; \int_\Omega \beta'_\alpha(y)|\nabla y|^2 \;\ge\; 0.$$

(vi) Assuming $y_n \to y$ in $\mathcal{H}$, it follows from the Lipschitz continuity of β_α that $\mathcal{A}_\alpha y_n$ converges to $\mathcal{A}_\alpha y$ in $\mathcal{V}_1^*$. This proves (1).

In order to show (2), from $|\beta_\alpha(r) - \tilde{\beta}_\alpha(r)| \le \frac{\varepsilon(\alpha)}{\alpha}(2+\overline{\rho})$ it follows that $||\beta_{\alpha_n} \circ y_n - \tilde{\beta}_{\alpha_n} \circ y_n||_{\mathcal{H}} \le C\frac{\varepsilon(\alpha)}{\alpha} \to 0$ as $\alpha \to 0$. Let $\widetilde{A}_{\alpha_n} : H \to H$, $\widetilde{A}_{\alpha_n} v := \beta_{\alpha_n}(v)$ denote the Yosida approximation of $\partial\tilde{\gamma}$ with parameter α_n. Consequently, it holds that $\widetilde{\mathcal{A}}_{\alpha_n} y_n \rightharpoonup h$ in $\mathcal{H}$. Since $\partial\tilde{\gamma}$ is maximal-monotone, we invoke Proposition 2.5 of [20] in order to conclude $(y,h) \in \partial\tilde{\gamma}$ and hence $(y,h) \in \mathcal{A}$.

Let us choose $\alpha_0 > 0$ so small that $\varepsilon(\alpha) < \min\{\inf_\Omega y_0 - a,\, b - \sup_\Omega y_0\}$ for all $\alpha \in [0,\alpha_0]$. Then $\beta_\alpha(y_0(x)) = 0$, and y_0 satisfies the conditions given in Assumption 3.1. This completes the proof. ◻

Remark 4.4. For α sufficiently small, β_α vanishes identically in a neighborhood of 0. We could choose different mollifiers ρ^1 and ρ^2 instead of ρ in the definition of $\beta_\alpha(r)$ for either positive r or negative r, respectively. Thus, the assertion of Proposition 4.3 remains true also in this case.

Convention 4.5. *In the last theorem below we will use projection and superposition operators which do not preserve the mean value. In order to simplify the notation, we extend the inner product of H to a semi-inner product on $L^2(\Omega)$ by*

$$(u|v)_H := (Qu|Qv)_H$$

for $u, v \in L^2(\Omega)$, where $Q : L^2(\Omega) \to H$ denotes the orthogonal projection of $L^2(Q)$ onto H. Likewise, we define

$$\langle u, v\rangle_{V_1} := \langle u|_{V_1}, Qv\rangle_{V_1}$$

for $u \in (H^1(\Omega))^*$, $v \in H^1(\Omega)$. This is well defined since Q maps $H^1(\Omega)$ onto V_1. Accordingly, the inner product of $\mathcal{H}$ and the dual pairing between $\mathcal{V}_1$ and its dual are extended to $L^2(\mathcal{T}; L^1(\Omega))$, respectively, $L^2(\mathcal{T}; H^1(\Omega))$ and its dual.

Finally, for the double-obstacle potential according to Example 4.1 we study further properties of various dual quantities involved in the system established in Theorem 3.16. This corresponds to a function space version of C-stationarity; cf. [41, 57].

Theorem 4.6. *Let the setting of Example* 4.1 *and the assumptions of Theorem* 3.16 *be satisfied. Using the notation $\xi_n := \widehat{\mathcal{A}}_{\alpha_n} y_n$ and $\lambda_n := \overline{D\mathcal{A}_{\alpha_n}}(y_n)^* \mathcal{L}_3 p_n$ and the results of Theorem* 3.16*, it follows that*

$$y' + \mathcal{L}_1 w = u, \qquad w = \mathcal{L}_1 y + \xi + \eta \mathcal{I}_1 y,$$

$$-\,p' + \mathcal{L}_1\mathcal{L}_3 p + \lambda + \eta\mathcal{I}_1\mathcal{L}_3 p = \mu_1(y - y_\Omega),$$

$$p + u \in \left(\mathbb{R}^+(\mathcal{C} - u)\right)^+.$$

If furthermore (λ_m) *is bounded in* $\mathcal{V}_1^*$ *and* $h : \mathbb{R} \to \mathbb{R}$ *is a Lipschitz function which satisfies* $h(a) = h(b) = 0$*, then the following relations hold true:*

$$\begin{aligned}(\xi\,|\,y-a)_{\mathcal{H}} = 0,\ (\xi\,|\,y-b)_{\mathcal{H}} = 0,\\ \lim\big(\lambda_m\,\big|\,h(y_m)\big)_{\mathcal{H}} = 0,\ \lim\,(\xi_m\,|\,\mathcal{L}_3 p_m)_{\mathcal{H}} = 0,\\ \underline{\lim}\,(\lambda_m\,|\,\mathcal{L}_3 p_m)_{\mathcal{H}} \geq 0,\end{aligned}$$

$$\lambda_k \to 0 \quad \textit{almost everywhere on } \{z \in \mathcal{T}\times\Omega \,:\, a < y(z) < b\}$$

for a subsequence (λ_k) *of* (λ_m).

Proof. 1. We start by showing the complementarity conditions $(\xi\,|\,y-a)_{\mathcal{H}} = 0,\ (\xi\,|\,y-b)_{\mathcal{H}} = 0$. For this purpose, recall that Assumption 3.1 ensures that $(y,\xi) \in \mathcal{A}$ and that $\mathcal{A}$ is the superposition operator of the subdifferential of $\varphi = \imath_{K_V} : V_1 \to \overline{\mathbb{R}}$. Since $\xi \in \mathcal{H}$, we therefore may conclude that $(y(z), \xi(z)) \in \partial\imath_M$ for almost all $z \in \overline{\mathcal{T}}\times\Omega$, which implies $\xi(z)(y(z)-a) = 0$ and $\xi(z)(y(z)-b) = 0$. Integration yields the desired complementarity conditions.

2. Next, we prove $\lim\big(\lambda_m\,\big|\,h(y_m)\big)_{\mathcal{H}} = 0$. Denoting the metric projection of $\mathbb{R}$ onto $M = [a,b]$ by p_M, the metric projection of H onto K_H by P (which is the superposition operator of p_M), and the superposition operator of P with respect to $\mathcal{T}$ by $\mathcal{P}$, and taking advantage of the continuity of the superposition operator of h on V_1 (cf. [48]), it follows that $\mathcal{P}(\mathcal{V}_1) \subset \mathcal{V}_1$ and $\lim \mathcal{P}y_m = \mathcal{P}y = y$, $\lim h(\mathcal{P}y_m) = h(\mathcal{P}y) = h(y) = \lim h(y_m)$ in $L^2(\mathcal{T}; H^1(\Omega))$.We know that $|\beta'_\alpha(r)| \leq \frac{1}{\alpha}$ for all r and $\beta'_\alpha(r) = 0$ for $a + \varepsilon(\alpha) \leq r \leq b - \varepsilon(\alpha)$. If L_h is the Lipschitz constant of h, then $|h(r)| \leq L_h \min(|r-a|, |r-b|)$ for $r \in \mathbb{R}$. Consequently, it follows that

$$\begin{aligned}|\big(\lambda_m\,\big|\,h(\mathcal{P}y_m)\big)_{\mathcal{H}}|^2 &= |\big(\mathcal{L}_3 p_m\,\big|\,\overline{D\mathcal{A}_{\alpha_m}}(y_m) h(\mathcal{P}y_m)\big)_{\mathcal{H}}|^2\\ &\leq ||\mathcal{L}_3 p_m||^2_{\mathcal{H}} \int_{\mathcal{T}\times\Omega} |\beta'_{\alpha_m}(y_m) h(\mathcal{P}y_m)|^2\\ &\leq \left(|\mathcal{T}\times\Omega|\,||\mathcal{L}_3 p_m||_{\mathcal{H}}\, L_h \frac{\varepsilon(\alpha_m)}{\alpha_m}\right)^2 \ \to\ 0\end{aligned}$$

as $m \to \infty$. Moreover, since (λ_m) is bounded in $\mathcal{V}_1^*$ we have that

$$\lim\big(\lambda_m\,\big|\,h(y_m)\big)_{\mathcal{H}} \;=\; \lim\big(\lambda_m\,\big|\,h(\mathcal{P}y_m)\big)_{\mathcal{H}} + \lim\big\langle \lambda_m\,, h(y_m) - h(\mathcal{P}y_m)\big\rangle_{\mathcal{V}_1} \;=\; 0.$$

3. We set $g_m(r) := \beta_{\alpha_m}(r) - \beta'_{\alpha_m}(r) q(r)$. With $r - p_M(r) = q(r)$ we find

$$\begin{aligned}(\xi_m\,|\,\mathcal{L}_3 p_m)_{\mathcal{H}} &= \big(\mathcal{L}_3 p_m\,\big|\,\beta_{\alpha_m}(y_m)\big)_{\mathcal{H}}\\ &= \big(\mathcal{L}_3 p_m\,\big|\,g_m(y_m)\big)_{\mathcal{H}} + (\lambda_m\,|\,y_m - \mathcal{P}y_m)_{\mathcal{H}}.\end{aligned}$$

By Lemma 4.2, for m sufficiently large it holds that $|g_m(r)| = |\beta_{\alpha_m}(r) - \beta'_{\alpha_m}(r) q(r)| \leq C\,\frac{\varepsilon(\alpha_m)}{\alpha_m}$. Hence, the first term on the right-hand side converges to 0. So does the second, since (λ_m) is bounded in $\mathcal{V}_1^*$ and since (y_m) and $(\mathcal{P}y_m)$ both converge to y in $\mathcal{V}_1$.

4. The fact that $\underline{\lim}\,\langle \lambda_m\,, \mathcal{L}_3 p_m\rangle_{\mathcal{V}_1} \geq 0$ is an obvious consequence of Lemma 3.13.

5. Let us fix representatives of the equivalence classes $y, (y_m)$ and denote Z the set $\{z \in \mathcal{T}\times\Omega : a < y(z) < b\}$. Since y_m converges to y in $\mathcal{V}_1$, a subsequence (y_k) of (y_m) converges almost everywhere on $\mathcal{T}\times\Omega$ to y. Moreover, we know that $\varepsilon(\alpha_m) \to 0$.

Hence, for almost all $z \in Z$ there exists $k_0(z)$ such that $a + \varepsilon(\alpha_k) < y_k(z) < b - \varepsilon(\alpha_k)$ for all $k \geq k_0(z)$.

From the properties of β_α it therefore follows that $\lambda_k(z) = 0$ for almost all $z \in Z$ and $k \geq k_0(z)$. Consequently, λ_k converges to 0 almost everywhere on Z. ∎

Combining the results of Theorems 3.16 and 4.6 and considering the sign condition satisfied by λ_m and p_m in the limit, we find that our stationarity system corresponds to a function space version of C-stationarity for MPECs; see [41, 43, 57].

We end this section by briefly discussing the relevance of C-stationarity in theory and numerical practice. Compared to weaker forms of stationarity, for instance, those contained in [11], C-stationarity represents a sharper stationarity notion avoiding spurious stationarity points. On the numerical level, it appears possible to extend the algorithms in [43] to the Cahn–Hilliard setting, which would then yield C-stationary points of the associated discrete problems.

Appendix A. Proofs for section 2. In this appendix we provide proofs of results given in section 2. We start with the basic Lemma 2.4.

Proof of Lemma 2.4. Let $(y, y^*) \in Y \times Y^*$ be given and define $\psi := \operatorname{Ext}(\varphi, X, Y, E)$. Then we have that

$$\begin{aligned}
&(y, y^*) \in \partial\psi \\
&\quad\iff y \in \operatorname{dom}\psi,\ \ \langle y^*, v - y\rangle_Y \leq \psi(v) - \psi(y) \quad \forall\, v \in Y \\
&\quad\iff y \in \operatorname{dom}\psi,\ \ u := E^{-1}y,\ \ \langle E^* y^*, x - u\rangle_X \leq \varphi(x) - \varphi(u) \quad \forall\, x \in X \\
&\quad\iff (E^{-1}y, E^* y^*) \in \partial\varphi.
\end{aligned}$$

This implies $\partial\psi = (E^*)^{-1}\partial\varphi E^{-1}$. Now assume that φ is proper, convex, and lower-semicontinuous and has bounded lower-level sets. It is clear that ψ is proper and convex. Let (y_n) be a sequence in Y that converges strongly to $y \in Y$. We have to show that $\psi(y) \leq \underline{\lim}\, \psi(y_n)$.

If $\underline{\lim}\, \psi(y_n) = \infty$, then the assertion is trivial. So let us assume that $\underline{\lim}\, \psi(y_n) < \infty$. Hence, it is possible to extract a subsequence (y_m) of (y_n) such that $y_m \in \operatorname{dom}\psi$ and $\underline{\lim}\, \psi(y_n) = \lim \psi(y_m)$. We set $x_m := E^{-1}y_m$, and since $(\varphi(x_m))$ is bounded in $\mathbb{R}$, (x_m) has to be bounded in X. Consequently, there exists a subsequence (x_k) of (x_m) that converges weakly in X. The continuity and injectivity of E therefore imply that this weak limit is $E^{-1}y$. Since φ is convex and lower-semicontinuous, it is even weakly lower-semicontinuous. This implies

$$\psi(y) = \varphi(E^{-1}y) \leq \underline{\lim}_{k\to\infty}\, \varphi(x_k) = \underline{\lim}_{n\to\infty}\, \psi(y_n)$$

and therefore finishes the proof. ∎

Next, the commutation rules and the regularity result for the bi-Laplace equations are established.

Proof of Proposition 2.8. Since $\partial\Omega$ is supposed to be sufficiently smooth, standard regularity results imply that $L_3 : V_3 \to V_1$ is an isomorphism and that $I_1 L_3 = L_1 I_3$ (cf. [35]). Furthermore, the unitarity of L_1 and L_3 is easily verified. Now let $v_0 \in V_3$ and $v_1 \in V_1$ be given. From the symmetry of I_1 and L_1 it follows that

$$\begin{aligned}
\langle I_3^* L_1 v_1, v_0\rangle_{V_3} &= \langle L_1 v_1, I_3 v_0\rangle_{V_1} = \langle L_1 I_3 v_0, v_1\rangle_{V_1} = \langle I_1 L_3 v_0, v_1\rangle_{V_1} = \langle I_1 v_1, L_3 v_0\rangle_{V_1} \\
&= \langle L_3^* I_1 v_1, v_0\rangle_{V_3}.
\end{aligned}$$

Thus, we have that $I_3^* L_1 = L_3^* I_1$. Consider $y, v \in V_1$ and $y^*, v^* \in V_1^*$. Since

$$\left(L_1^{-1} I_1 y \,\middle|\, v\right)_{V_1} = \langle I_1 y, v\rangle_{V_1},$$

$$\left((L_3^*)^{-1} I_3^* y^* \,\middle|\, v^*\right)_{V_1^*} = \left(I_1 L_1^{-1} y^* \,\middle|\, L_1 L_1^{-1} v^*\right)_{V_1^*} = \left\langle I_1 L_1^{-1} y^*, L_1^{-1} v^*\right\rangle_{V_1},$$

the symmetry, positivity, and injectivity of $L_1^{-1} I_1$ and $(L_3^*)^{-1} I_3^*$ are obvious. By Rellich's lemma, I_{V_1} is compact and therefore also are I_1 and $L_1^{-1} I_1$. The same applies to I_3 and $(L_3^*)^{-1} I_3^*$ by the Sobolev embedding theorem.

Assume that U_1 is a closed subspace of V_1 satisfying $L_1^{-1} I_1(U_1) = U_1$. The continuity of L_1^{-1} implies that $I_1(U_1)$ is closed in V_1^* as well. Since U_1 is a Hilbert space with the induced inner product of V_1, we conclude from the spectral theory of self-adjoint, compact operators the existence of a finite or countable sequence of orthonormal eigenvectors $(v_n)_{n\in N}$ with $N \subset \mathbb{N}$ in U_1 of $L_1^{-1} I_1$ for the eigenvalues $(\lambda_n)_{n\in N}$ such that $U_1 = \bigoplus_{n\in N} \mathbb{R} v_n$. Moreover, it holds that $\lambda_n > 0$ for all $n \in N$ and $\lambda_n \to 0$ for $n \to \infty$, $n \in N$. Since $L_1 : V_1 \to V_1^*$ is a unitary operator and $I_1(U_1) = L_1(U_1)$, the orthogonal projections $P_{V_1 \to U_1}$ and $P_{V_1^* \to I_1(U_1)}$ from V_1 onto U_1 and from V_1^* onto $I_1(U_1)$, respectively, are given by

$$P_{V_1\to U_1} v = \sum_{n\in N} (v|v_n)_{V_1} v_n \quad \text{and} \quad P_{V_1^*\to I_1(U_1)} v^* = \sum_{n\in N} (v^*|L_1 v_n)_{V_1^*} L_1 v_n.$$

We have $L_1^{-1} I_1 v_n = \lambda_n v_n$, and hence $\lambda_n L_1 v_n = I_1 v_n$. Furthermore, it holds that

$$(I_1 v|L_1 v_n)_{V_1^*} = \langle I_1 v, v_n\rangle_{V_1} = \langle I_1 v_n, v\rangle_{V_1} = \lambda_n \langle L_1 v_n, v\rangle_{V_1} = \lambda_n (v_n|v)_{V_1}.$$

Altogether, this yields

$$\begin{aligned}
I_1 P_{V_1\to U_1} v &= I_1\left(\sum_{n\in N} (v|v_n)_{V_1} v_n\right) = \sum_{n\in N} (v|v_n)_{V_1} I_1 v_n,\\
P_{V_1^*\to I_1(U_1)} I_1 v &= \sum_{n\in N} (I_1 v|L_1 v_n)_{V_1^*} L_1 v_n = \sum_{n\in N} \lambda_n (v|v_n)_{V_1} \frac{1}{\lambda_n} I_1 v_n,\\
L_1 P_{V_1\to U_1} v &= L_1\left(\sum_{n\in N} (v|v_n)_{V_1} v_n\right) = \sum_{n\in N} (v|v_n)_{V_1} L_1 v_n,\\
P_{V_1^*\to I_1(U_1)} L_1 v &= \sum_{n\in N} (L_1 v|L_1 v_n)_{V_1^*} L_1 v_n = \sum_{n\in N} (v|v_n)_{V_1} L_1 v_n,
\end{aligned}$$

and therefore $I_1 P_{V_1\to U_1} = P_{V_1^*\to I_1(U_1)} I_1$ and $L_1 P_{V_1\to U_1} = P_{V_1^*\to I_1(U_1)} L_1$. Similar arguments can be used to prove $I_3^* P_{V_1^*\to U_1^*} = P_{V_3^*\to I_3^*(U_1^*)} I_3^*$ and $L_3^* P_{V_1^*\to U_1^*} = P_{V_3^*\to I_3^*(U_1^*)} L_3^*$. ∎

The following general integration by parts formula is a result of Gröger [36]; cf. [61].

Proposition A.1. *Let V be a reflexive Banach space, let H be an arbitrary Hilbert space, and let $K \in \mathcal{L}(V; H)$ with dense range. We define $E := K^* J_H K \in \mathcal{L}(V; V^*)$ and $\mathcal{H} := L^2(\mathcal{T}; H)$, $\mathcal{V} := L^2(\mathcal{T}; V)$, $\mathcal{W} := \{y \in \mathcal{V} : (\mathcal{E} y)' \in \mathcal{V}^*\}$ with the standard norms and $\|y\|_{\mathcal{W}} := (\|y\|_{\mathcal{V}}^2 + \|(\mathcal{E} y)'\|_{\mathcal{V}^*}^2)^{1/2}$, respectively. Then the operator $\mathcal{K}$ maps $\mathcal{W}$ continuously into the space $C(\overline{\mathcal{T}}; H)$, meaning that every class of equivalent* *functions in* $\mathcal{K}(\mathcal{W}) \subset L^2(\mathcal{T};H)$ *possesses a representative that is continuous from* $\mathcal{T}$ *into* H *with continuous extension onto* $\overline{\mathcal{T}}$*. Furthermore, in this sense the formulas*

$$\begin{aligned}&\big((\mathcal{K}y)(t_2)\,\big|\,(\mathcal{K}v)(t_2)\big)_H \;-\; \big((\mathcal{K}y)(t_1)\,\big|\,(\mathcal{K}v)(t_1)\big)_H \\ &\qquad = \int_{t_1}^{t_2} [\big\langle (\mathcal{E}y)'(t), v(t)\big\rangle_V \;+\; \big\langle (\mathcal{E}v)'(t), y(t)\big\rangle_V]\, dt, \end{aligned} \tag{A.1}$$

and in particular

$$||(\mathcal{K}y)(t_2)||_H^2 \;-\; ||(\mathcal{K}y)(t_1)||_H^2 \;=\; 2\int_{t_1}^{t_2} \big\langle (\mathcal{E}y)'(t), y(t)\big\rangle_V\, dt, \tag{A.2}$$

hold for all $y, v \in \mathcal{W}$ *and* $t_1, t_2 \in \overline{\mathcal{T}}$*.*

Proof. 1. First, assume that $y, v \in C_c^1(\mathbb{R};V)$. Let us define the functional $\varphi(t) := \big((\mathcal{K}y)(t)\,\big|\,(\mathcal{K}v)(t)\big)_H = \big(Ky(t)\,\big|\,Kv(t)\big)_H$. It is easy to see that $\varphi \in C^1(\mathbb{R})$ and

$$\varphi'(t) = \big(Ky'(t)\,\big|\,Kv(t)\big)_H + \big(Kv'(t)\,\big|\,Ky(t)\big)_H = \big\langle (\mathcal{E}y)'(t), v(t)\big\rangle_V + \big\langle (\mathcal{E}v)'(t), y(t)\big\rangle_V.$$

Hence, (A.1) and (A.2) follow from the fundamental theorem of calculus.

2. Again, let us suppose that $y \in C_c^1(\mathbb{R};V)$ and $t, s \in \mathbb{R}$. By virtue of (A.2) and Hölder's inequality, we conclude that

$$||(\mathcal{K}y)(t)||_H^2 \;\le\; ||(\mathcal{K}y)(s)||_H^2 + 2||(\mathcal{E}y)'||_{\mathcal{V}^*}||y||_{\mathcal{V}} \;\le\; ||(\mathcal{K}y)(s)||_H^2 + 2||y||_{\mathcal{W}}^2.$$

After applying the square root on both sides and integrating over $s \in \operatorname{supp} y$, we can estimate

$$|\operatorname{supp} y|\; ||(\mathcal{K}y)(t)||_H \;\le\; ||\mathcal{K}y||_{L^1(\mathbb{R};H)} + \sqrt{2}\,|\operatorname{supp} y|\; ||y||_{\mathcal{W}}.$$

Assume that $|\operatorname{supp} y| > 0$. Since $\mathcal{K} \in L(L^2(\mathbb{R};V);L^2(\mathbb{R};V))$, we obtain

$$\sup_{t\in\mathbb{R}} ||(\mathcal{K}y)(t)||_H \;\le\; C(||y||_{\mathcal{V}} + ||y||_{\mathcal{W}}) \;\le\; C||y||_{\mathcal{W}} \tag{A.3}$$

for constant C depending only on $|\operatorname{supp} y|$. If $|\operatorname{supp} y| = 0$, this inequality obviously holds as well.

3. Let $y \in \mathcal{W}$ be arbitrary. Since the restriction to $\mathcal{T}$ of the set of continuously differentiable functions on $\mathbb{R}$ with values in V is dense in $\mathcal{W}$, there exists a sequence (y_n) in $C_c^1(\mathbb{R};V)$ such that $y_n|_{\mathcal{T}} \to y$ in $\mathcal{W}$. From (A.3) it follows that $\mathcal{K}y_n$ is a Cauchy sequence in $C(\overline{\mathcal{T}};H)$ and therefore converges to some $h \in C(\overline{\mathcal{T}};H)$. It is not difficult to see that $\mathcal{K}y = h$ almost everywhere on $\mathcal{T}$. Hence, h is the continuous representative of $\mathcal{K}y \in L^\infty(\overline{\mathcal{T}};H)$. Moreover, in the estimation in steps 1 and 2 we can pass to the limit in order to obtain the assertion for arbitrary $y, v \in \mathcal{W}$. This completes the proof. ◻

Proof of Proposition 2.9. From Proposition A.1 and with $L_1 = J_V$ we obtain that for functions $f, g \in \mathcal{V}_3$ with $\mathcal{E}f, \mathcal{E}g \in H^1(\mathcal{T};V_3^*)$ for $E = I_3^* L_1 I_3$ it holds that f, g admit continuous representatives defined on $\overline{\mathcal{T}}$ with values in V_1 and

$$\big\langle (\mathcal{E}f)', g\big\rangle_{\mathcal{V}_3} + \big\langle (\mathcal{E}g)', f\big\rangle_{\mathcal{V}_3} \;=\; \big(f(T)\,\big|\,g(T)\big)_{V_1} - \big(f(0)\,\big|\,g(0)\big)_{V_1}.$$

Proposition 2.8 implies

$$E \;=\; I_3^* L_1 I_3 \;=\; L_3^* I_1 I_3 \;=\; I_3^* I_1 L_3.$$

For y, v as in the assertion, let us choose $f := y \in \mathcal{V}_3$ and $g := L_3^{-1}v \in \mathcal{V}_3$. Thus, we have $H^1(\mathcal{T}; V_1^*) \ni y = (\mathcal{L}_3^*)^{-1}\mathcal{E}f$ and $H^1(\mathcal{T}; V_3^*) \ni v = \mathcal{E}g$ and therefore $\mathcal{E}f, \mathcal{E}g \in H^1(\mathcal{T}; V_3^*)$. This implies that $f = y$ and $g = \mathcal{L}_1^{-1}v$ can be regarded as continuous functions on $\overline{\mathcal{T}}$ with values in V_1 (and hence also v with values in V_1^*). Moreover, it follows that

$$
\begin{aligned}
&\langle y', v\rangle_{\mathcal{V}_1} + \langle v', y\rangle_{\mathcal{V}_3} \\
&\quad = \langle (\mathcal{L}_3^*)^{-1}(\mathcal{E}f)', \mathcal{L}_3 g\rangle_{\mathcal{V}_1} + \langle (\mathcal{E}g)', f\rangle_{\mathcal{V}_3} \;=\; \langle (\mathcal{E}f)', g\rangle_{\mathcal{V}_3} + \langle (\mathcal{E}g)', f\rangle_{\mathcal{V}_3} \\
&\quad = \big(f(T)\,\big|\,g(T)\big)_{V_1} - \big(f(0)\,\big|\,g(0)\big)_{V_1} \\
&\quad = \langle (\mathcal{L}_1 g)(T), f(T)\rangle_{V_1} - \langle (\mathcal{L}_1 g)(0), f(0)\rangle_{V_1} \\
&\quad = \langle v(T), y(T)\rangle_{V_1} - \langle v(T), y(0)\rangle_{V_1}.
\end{aligned}
$$

This completes the proof. ▯

Proof of Theorem 2.10. Choose a complete orthonormal system of eigenvectors $(e_n)_{n\in\mathbb{N}}$ for the eigenvalues $(\lambda_n)_{n\in\mathbb{N}}$ in V_1 of the operator $L_1^{-1}I_1$. We denote the span of $e_1, \ldots, e_n$ by $V_{1,n}$ and $V_{1,n}^* := I_1(V_{1,n}) \subset V_1^*$, $V_{3,n}^* := I_3^*(V_{1,n}^*) \subset V_3^*$. The orthogonal projection of V_1, V_1^*, and V_3^* onto the subspaces $V_{1,n}, V_{1,n}^*$, and $V_{3,n}^*$ are denoted by $P_{V_{1,n}}, P_{V_{1,n}^*}$, and $P_{V_{3,n}^*}$, respectively. Then it is not hard to show that $P_{V_{1,n}}y =: y_n \to y$ in $\mathcal{V}_1$, $y_n \in H^1(\mathcal{T}; V_{1,n})$, and

$$\mathcal{P}_{V_{3,n}^*}y' \;=\; y_n' \to y' \quad \text{in } \mathcal{V}_3^*.$$

Moreover, the continuity of L_3 and L_1 implies that $\mathcal{L}_3^*\mathcal{L}_1 y_n \to \mathcal{L}_3^*\mathcal{L}_1 y$ in $\mathcal{V}_3^*$. With $\lambda_n e_n = L_1^{-1}I_1 e_n = I_3 L_3^{-1} e_n$, it holds that $e_n \in V_3$. Hence $y_n \in L^2(\mathcal{T}; V_3)$.

In order to prove the assertion, it suffices to show that (y_n) and (y_n') remain bounded in $\mathcal{V}_3$ and $\mathcal{V}_1^*$, respectively. From $y' + \mathcal{L}_3^*\mathcal{L}_1 y = f$ and Proposition 2.8 we obtain

$$y_n' + \mathcal{L}_1\mathcal{L}_3 y_n \;=\; \mathcal{P}_{V_{3,n}^*}[y' + \mathcal{L}_3^*\mathcal{L}_1 y] \;=\; \mathcal{P}_{V_{1,n}^*}f,$$

and therefore using $\|f\|_{\mathcal{V}_1^*} \geq \|\mathcal{P}_{V_{1,n}^*}f\|_{\mathcal{V}_1^*}$ also, we obtain

$$\|f\|_{\mathcal{V}_1^*}^2 \;\geq\; \|y_n' + \mathcal{L}_1\mathcal{L}_3 y_n\|_{\mathcal{V}_1^*}^2 \;=\; \|y_n'\|_{\mathcal{V}_1^*}^2 + 2\big(y_n'\,\big|\,\mathcal{L}_1\mathcal{L}_3 y_n\big)_{\mathcal{V}_1^*} + \|\mathcal{L}_1\mathcal{L}_3 y_n\|_{\mathcal{V}_1^*}^2.$$

The term on the right-hand side can be transformed into

$$
\begin{aligned}
\big(y_n'\,\big|\,\mathcal{L}_1\mathcal{L}_3 y_n\big)_{\mathcal{V}_1^*} = \langle y_n', \mathcal{L}_3 y_n\rangle_{\mathcal{V}_1} \;&=\; \langle \mathcal{L}_1 y_n, y_n'\rangle_{\mathcal{V}_1} \;=\; \big(y_n\,\big|\,y_n'\big)_{\mathcal{V}_1} \\
&= \frac{1}{2}\Big(\|y_n(T)\|_{V_1}^2 - \|y_n(0)\|_{V_1}^2\Big).
\end{aligned}
$$

The sequence $\|y_n(0)\|_{V_1}$ is bounded by $\|y(0)\|_{V_1}$ since by Proposition 2.8 it holds that $y_n(0) = (\mathcal{P}_{V_{1,n}}y)(0)$. Because of $\|\mathcal{L}_1\mathcal{L}_3 y_n\|_{\mathcal{V}_1^*} = \|\mathcal{L}_3 y_n\|_{\mathcal{V}_1} = \|y_n\|_{\mathcal{V}_3}$, the proof is finished. ▯

Proof of Remark 2.11. First, notice that the eigenvectors (e_n) of $L_1^{-1}I_1$ are elements of V_4. Then proceeding as above we obtain

$$\|f\|_{\mathcal{H}}^2 \;\geq\; \|y_n' + \mathcal{L}_2\mathcal{L}_4 y_n\|_{\mathcal{H}}^2 \;=\; \|y_n'\|_{\mathcal{H}}^2 + 2\big(y_n'\,\big|\,\mathcal{L}_2\mathcal{L}_4 y_n\big)_{\mathcal{H}} + \|\mathcal{L}_2\mathcal{L}_4 y_n\|_{\mathcal{H}}^2.$$

Again, using analogous commutation rules as given in Proposition 2.8, we conclude that

$$\big(y_n' \,\big|\, \mathcal{L}_2\mathcal{L}_4 y_n\big)_{\mathcal{H}} = \big(\mathcal{L}_2 y_n' \,\big|\, \mathcal{L}_2 y_n\big)_{\mathcal{H}} = \frac{1}{2}\Big(||y_n(T)||_{V_2}^2 - ||y_n(0)||_{V_2}^2\Big).$$

Since $||y_n(0)||_{V_2}$ is bounded by $||y(0)||_{V_2}$, the assertion follows. □

In order to prove Proposition 2.12 we make use of a general result by Kampowsky [45] on the continuous Fréchet differentiability in Lebesgue spaces and reduce the problem to this case.

THEOREM A.2. *Let $(U_i)_{1\le i\le n}$, V be normed vector spaces, let $\Omega \subset \mathbb{R}^N$ be a measurable subset, and let $q, (p_i)_{1\le i\le n}$ be numbers with $1 \le q < p_i \le \infty$ for all $1 \le i \le n$. We define $r_i := (\frac{1}{q} - \frac{1}{p_i})^{-1}$ (with $\frac{1}{\infty} := 0$) and the spaces $U := U_1 \times \cdots \times U_n$, $\mathcal{U}_i := L^{p_i}(\Omega; U_i)$, $\mathcal{U} := \mathcal{U}_1 \times \cdots \times \mathcal{U}_n$, $\mathcal{V} := L^q(\Omega; V)$, and $\mathcal{W}_i := L^{r_i}(\Omega; \mathcal{L}(U_i; V))$ for all $1 \le i \le n$. Assume that the mapping $\gamma : \Omega \times U \to V$ possesses all partial derivatives $\partial_{u_i}\gamma =: \gamma_i : \Omega \times U \to \mathcal{L}(U_i; V)$ in the U_i-variables as Fréchet derivatives, that these fulfill the Carathéodory condition, and that their superposition operators*

$$(A_i y)(x) := \gamma_i(x, y(x))$$

map $\mathcal{U}$ continuously into $\mathcal{W}_i$. Then the superposition operator $A : \mathcal{U} \to \mathcal{V}$ of a is continuously Fréchet differentiable, and its derivative $A' : \mathcal{U} \times \mathcal{L}(\mathcal{U}; \mathcal{V})$ is given by

$$(A'(y; v))(x) \;=\; \sum_{i=1}^{n} A_i'(y)(x)[v(x)]$$

for all $y, v \in \mathcal{U}$ and almost all $x \in \Omega$.

Proof of Proposition 2.12. 1. In the first step we show that the superposition operator A of γ

$$(Ay)(x) := \gamma(x, y(x))$$

mapping $W^{1,s_2}(\Omega)$ into $H^1(\Omega)$ is continuously Fréchet differentiable, and its derivative $A' : W^{1,s_2}(\Omega) \to \mathcal{L}(W^{1,s_2}(\Omega); H^1(\Omega))$ is given by $A'(y; r)(x) = \gamma'(y(x))r(x)$.

For this purpose, choose $2 < p_1 < \infty$ such that $\frac{1}{2} - \frac{1}{s_2} \ge \frac{1}{p_1} \ge \frac{1}{s_2} - \frac{1}{N}$. The assumption guarantees that this is possible. Then, $W^{1,s_2}(\Omega)$ embeds continuously into $L^{p_1}(\Omega)$. Moreover,

$$G : H^1(\Omega) \to L^2(\Omega) \times L^2(\Omega; \mathbb{R}^N), \quad r \mapsto Gr := (y, \nabla y)$$

is an isometric isomorphism from $H^1(\Omega)$ onto a closed subspace H_1 of the space $L^2(\Omega) \times L^2(\Omega; \mathbb{R}^N)$, and the gradient $\nabla(\gamma \circ y)$ is given by $\gamma'(y)\nabla y$. Instead of analyzing A directly, we define the operator

$$\widetilde{A} : L^{p_1}(\Omega) \times L^{p_2}(\Omega; \mathbb{R}^N) \to L^2(\Omega) \times L^2(\Omega; \mathbb{R}^N), \quad (y, v) \mapsto (\gamma(y), \gamma'(y)v)$$

with $p_2 := s_2$ and use the fact that A can be decomposed into

$$\begin{array}{ccccccc} W^{1,s_2}(\Omega) & \to & L^{p_1}(\Omega) \times L^{p_2}(\Omega; \mathbb{R}^N) & \to & H_1 & \to & H^1(\Omega), \\ y & \mapsto & (y, \nabla y) & \mapsto & \widetilde{A}(y, \nabla y) & \mapsto & Ay = G^{-1}\widetilde{A}(y, \nabla y). \end{array}$$

Since $W^{1,s_2}(\Omega) \ni y \mapsto (y, \nabla y) \in L^{p_1}(\Omega) \times L^{p_2}(\Omega;\mathbb{R}^N)$ and $G^{-1} : H_1 \to H^1(\Omega)$ are linear, continuous operators, it suffices to prove that $\widetilde{A}$ is continuously Fréchet differentiable, and $\widetilde{A}' : L^{p_1}(\Omega) \times L^{p_2}(\Omega;\mathbb{R}^N) \to \mathcal{L}(L^{p_1}(\Omega) \times L^{p_2}(\Omega;\mathbb{R}^N); L^2(\Omega) \times L^2(\Omega;\mathbb{R}^N))$ is given by

$$(\widetilde{A}'(y,v)[r,s])(x) \;=\; (\gamma'(y)r,\, \gamma''(y)rv + \gamma'(y)s),$$

because then we have

$$\widetilde{A}'(y,\nabla y)[r,\nabla r] \;=\; (\gamma'(y)r,\, \nabla(\gamma'(y)r)) \;=\; G(\gamma'(y)r).$$

From the fact that γ is differentiable and Lipschitz continuous, it follows that $y \mapsto a(y)$ is continuously Fréchet differentiable from $L^{p_1}(\Omega)$ into $L^2(\Omega)$ by Theorem A.2 with derivative $\gamma'(y)$. Moreover, the mapping $b : (f,h) \mapsto \gamma'(f)h$ from $\mathbb{R} \times \mathbb{R}^N$ into $\mathbb{R}^N$ possesses the total differential $(Db)(e,g) \in \mathcal{L}(\mathbb{R} \times \mathbb{R}^N; \mathbb{R}^N)$, $(Db)(e,g)[f,h] = \gamma''(e)fg + \gamma'(e)h$. We know that $\gamma''(\mathbb{R})$ and $\gamma'(\mathbb{R})$ are bounded subsets of $\mathbb{R}$, that Db is continuous and

$$\begin{aligned} ||(\partial_u b)(e,g)||_{\mathcal{L}(\mathbb{R};\mathbb{R}^N)} &= |\gamma''(e)g| \;\le\; C|g|, \\ ||(\partial_v b)(e,g)||_{\mathcal{L}(\mathbb{R}^N;\mathbb{R}^N)} &= |\gamma'(e)| \;\le\; C. \end{aligned}$$

Consequently, the corresponding superposition operators B_u, B_v satisfy

$$\begin{aligned} B_u &: L^{p_1}(\Omega) \times L^{p_2}(\Omega;\mathbb{R}^N) \to L^{p_2}(\Omega;\mathcal{L}(\mathbb{R};\mathbb{R}^N)) \;\hookrightarrow\; L^{r_1}(\Omega;\mathcal{L}(\mathbb{R};\mathbb{R}^N)), \\ B_v &: L^{p_1}(\Omega) \times L^{p_2}(\Omega;\mathbb{R}^N) \to L^{\infty}(\Omega;\mathcal{L}(\mathbb{R};\mathbb{R}^N)) \;\hookrightarrow\; L^{r_2}(\Omega;\mathcal{L}(\mathbb{R}^N;\mathbb{R}^N)) \end{aligned}$$

and are continuous with $p_2 = s_2 \ge r_1 := (\frac{1}{2} - \frac{1}{p_1})^{-1}$ and $\infty \ge r_2 := (\frac{1}{2} - \frac{1}{p_2})^{-1}$. Theorem A.2 now implies the continuous Fréchet differentiability of A from $W^{1,s_2}(\Omega)$ into $H^1(\Omega)$ with the derivative as desired.

2. Now let us consider the superposition operator $\mathcal{A}$ of A itself and show its continuous Fréchet differentiability. The superposition operator of A' will be denoted by B, i.e., $(Bu)(t) = A'(u(t))$. If we are able to show that B maps $L^{s_1}(\mathcal{T}; W^{1,s_2}(\Omega))$ continuously into $L^r(\mathcal{T}; \mathcal{L}(W^{1,s_2}(\Omega); H^1(\Omega)))$ with $\frac{1}{s_1} + \frac{1}{r} = \frac{1}{2}$, then Theorem A.2 applies and we are done.

Consider $M := \{y \in W^{1,s_2}(\Omega) : ||y||_{W^{1,s_2}(\Omega)} \le 1\}$, and for $y, r \in W^{1,s_2}(\Omega)$

$$\begin{aligned} ||A'(y;r)||^2_{H^1(\Omega)} &= ||\gamma'(y)r||^2_{H^1(\Omega)} \\ &= ||\gamma'(y)r||^2_{L^2(\Omega)} + ||\gamma''(y)r\nabla y||^2_{L^2(\Omega)} + ||\gamma'(y)\nabla r||^2_{L^2(\Omega)} \\ &\le C\Big(||r||^2_{L^2(\Omega)} + ||r\nabla y||^2_{L^2(\Omega)} + ||\nabla r||^2_{L^2(\Omega)}\Big). \end{aligned}$$

The second summand can be estimated with the help of the generalized Hölder's inequality ($||fg||_{L^p} \le ||f||_{L^q}||g||_{L^r}$ for $\frac{1}{p} = \frac{1}{q} + \frac{1}{r}$) and the Sobolev embedding $W^{1,s_2}(\Omega) \hookrightarrow L^{p_1}(\Omega)$ by

$$||r\nabla y||_{L^2(\Omega)} \;\le\; ||r||_{L^{p_1}(\Omega)}||\nabla y||_{L^{q_1}(\Omega)} \;\le\; C||r||_{W^{1,s_2}(\Omega)}||y||_{W^{1,s_2}(\Omega)}$$

with $q_1 := (\frac{1}{2} - \frac{1}{p_1})^{-1} \le (\frac{1}{2} - (\frac{1}{2} - \frac{1}{s_2}))^{-1} = s_2$ by the choice of p_1. Consequently, it holds that

$$||A'(y)||_{\mathcal{L}(W^{1,s_2}(\Omega);H^1(\Omega))} \;=\; \sup\Big\{||A'(y;r)||_{H^1(\Omega)} :\; r \in M\Big\} \;\le\; C(1 + ||y||_{W^{1,s_2}(\Omega)}).$$

Finally, we have $r \leq s_1$ by assumption, which implies that the operator

$$B : L^{s_1}(\mathcal{T}; W^{1,s_2}(\Omega)) \to L^r(\mathcal{T}; \mathcal{L}(W^{1,s_2}(\Omega); H^1(\Omega)))$$

is continuous. This finishes the proof. □

REFERENCES

[1] H. Abels, *Existence of weak solutions for a diffuse interface model for viscous, incompressible fluids with general densities*, Comm. Math. Phys., 289 (2009), pp. 45–73.

[2] H. Abels and M. Wilke, *Convergence to equilibrium for the Cahn-Hilliard equation with a logarithmic free energy,* Nonlinear Anal. Theory Methods Appl., 67 (2007), pp. 3176–3193.

[3] Y. Achdou, *An inverse problem for a parabolic variational inequality arising in volatility calibration with American options*, SIAM J. Control Optim., 43 (2005), pp. 1583–1615.

[4] Y. Achdou and O. Pironneau, *Computational Methods for Option Pricing*, Frontiers Appl. Math. 30, SIAM, Philadelphia, 2005.

[5] R.A. Adams and J.J.F. Fournier, *Sobolev Spaces*, 2nd ed., Pure Appl. Math. 140, Academic Press, New York, 2003.

[6] H. Amann, *Compact embeddings of vector-valued Sobolev and Besov spaces*, Glas. Mat. Ser. III, 35 (2000), pp. 161–177.

[7] D.M. Anderson, G.B. McFadden, and A.A. Wheeler, *Diffuse-interface methods in fluid mechanics*, An. Rev. Fluid Mech., 30 (1998), pp. 139–165.

[8] L. Bañas and R. Nürnberg, *Adaptive finite element methods for Cahn-Hilliard equations*, J. Comput. Appl. Math., 218 (2008), pp. 2–11.

[9] L. Bañas and R. Nürnberg, *A multigrid method for the Cahn-Hilliard equation with obstacle potential,* Appl. Math. Comput., 213 (2009), pp. 290–303.

[10] V. Barbu, *Nonlinear Semigroups and Differential Equations in Banach Spaces*, Editura Academiei Republicii Socialiste România, Bucharest, 1976.

[11] V. Barbu, *Optimal Control of Variational Inequalities*, Res. Notes Math. 100, Pitman (Advanced Publishing Program), Boston, 1984.

[12] V. Barbu and Th. Precupanu, *Convexity and Optimization in Banach Spaces*, 2nd rev. and extended ed., D. Reidel Publishing Company, Dordrecht/Boston/Lancaster, 1986.

[13] M. Bergounioux, *Optimal control of problems governed by abstract elliptic variational inequalities with state constraints*, SIAM J. Control Optim., 36 (1998), pp. 273–289.

[14] M. Bergounioux, *Optimal control of semilinear elliptic obstacle problems*, J. Nonlinear Convex Anal., 3 (2002), pp. 25–39.

[15] A. Bertozzi, S. Esedoḡlu, and A. Gillette, *Analysis of a two-scale Cahn–Hilliard model for binary image inpainting*, Multiscale Model. Simul., 6 (2007), pp. 913–936.

[16] L. Blank, M. Butz, and H. Garcke, *Solving the Cahn-Hilliard Variational Inequality with a Semi-Smooth Newton Method*, ESAIM Control Optim. Calc. Var., 17 (2011), pp. 931–954.

[17] J.F. Blowey and C.M. Elliott, *The Cahn-Hilliard gradient theory for phase separation with non-smooth free energy.* I: *Mathematical analysis*, Eur. J. Appl. Math., 2 (1991), pp. 233–280.

[18] J.F. Blowey and C.M. Elliott, *The Cahn-Hilliard gradient theory for phase separation with non-smooth free energy.* II: *Numerical analysis*, Eur. J. Appl. Math., 3 (1992), pp. 147–179.

[19] E. Bonetti, P. Colli, W. Dreyer, G. Gilardi, G. Schimperna, and J. Sprekels, *On a model for phase separation in binary alloys driven by mechanical effects*, Phys. D, 165 (2002), pp. 48–65.

[20] H. Brézis, *Opérateurs Maximaux Monotones et Semi-groupes de Contractions dans les Espaces de Hilbert*, North-Holland Mathematics Studies 5, North-Holland, Amsterdam, 1973.

[21] J.W. Cahn and J.E. Hilliard, *Free energy of a nonuniform system.* I. *Interfacial free energy*, J Chem. Phys., 28 (1958), pp. 258–266.

[22] I. Capuzzo Dolcetta, S. Finzi Vita, and R. March, *Area preserving curve-shortening flows: From phase separation to image processing*, Interfaces and Free Bound., 4 (2002), pp. 325–343.

[23] F.H. Clarke, *Optimization and Nonsmooth Analysis*, 2nd ed., Classics Appl. Math. 5, SIAM, Philadelphia, 1990.

[24] P. Colli, P. Krejčí, E. Rocca, and J. Sprekels, *Nonlinear evolution inclusions arising from phase change models*, Czech. Math. J., 57 (2007), pp. 1067–1098.

[25] M. Copetti and C.M. Elliott, *Kinetics of phase decomposition processes: Numerical solutions to the Cahn-Hilliard equation*, Materials Science and Technology, 6 (1990), pp. 273–283.

[26] J. Diestel and J.J. jun. Uhl, *Vector Measures*, Math. Surveys 15, AMS, Providence, RI, 1977.

[27] C.M. Elliott and S. Zheng, *On the Cahn-Hilliard equation*, Arch. Ration. Mech. Anal., 96 (1986), pp. 339–357.

[28] H.B. Frieboes et al., *Towards multiscale modeling of nanovectored delivery of therapeutics to cancerous lesions*, Biomed. Dev., to appear.

[29] H. Gajewski, K. Gröger, and K. Zacharias, *Nichtlineare Operatorgleichungen und Operatordifferentialgleichungen*, Mathematische Lehrbücher und Monographien 38, Akademie-Verlag, Berlin, 1974.

[30] H. Garcke, *On Cahn-Hilliard systems with elasticity*, Proc. R. Soc. Edinburgh Sect. A Math., 133 (2003), pp. 307–331.

[31] H. Garcke, *Optimization Problems and Cahn-Hilliard Systems*, Technical report, Mathematisches Institut Oberwolfach, Oberwolfach-Walke, Germany, 2007.

[32] R. Glowinski and A.M. Ramos, *A numerical approach to the Neumann control of the Cahn–Hilliard equation*, in Computational Methods for Control and Applications, Gakuto International Series: Mathematical Sciences and Applications 16, Gakkotosho Co., Tokyo, 2002, pp. 111–155.

[33] C. Gräser and R. Kornhuber, *On preconditioned Uzawa-type iterations for a saddle point problem with inequality constraints*, in Domain Decomposition Methods in Science and Engineering XVI, Lect. Notes Comput. Sci. Eng. 55, Springer, Berlin, 2007, pp. 91–102.

[34] C. Gräser and R. Kornhuber, *Multigrid methods for obstacle problems*, J. Comput. Math., 27 (2009), pp. 1–44.

[35] P. Grisvard, *Elliptic Problems in Nonsmooth Domains,* Monogr. Stud. Math. 24, Pitman (Advanced Publishing Program), Boston, 1985.

[36] K. Gröger, *Lecture on 'Evolutionsgleichungen*,' Humboldt-Universität Berlin, Berlin, 2001.

[37] F. Hausser, S. Rasche, and A. Voigt, *The influence of electric fields on nanostructures—simulation and control*, Math. Comput. Simulation, 80 (2010), pp. 1449–1457.

[38] M. Hintermüller, *Inverse coefficient problems for variational inequalities: Optimality conditions and numerical realization*, M2AN Math. Model. Numer. Anal., 35 (2001), pp. 129–152.

[39] M. Hintermüller, *An active-set equality constrained Newton solver with feasibility restoration for inverse coefficient problems in elliptic variational inequalities*, Inverse Problems, 24 (2008), 034017.

[40] M. Hintermüller, M. Hinze, and M.H. Tber, *An adaptive finite element Moreau-Yosida-based solver for a non-smooth Cahn-Hilliard problem*, Optim. Methods Softw., 26 (2011), pp. 777–811.

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

[42] M. Hintermüller and I. Kopacka, *A smooth penalty approach and a nonlinear multigrid algorithm for elliptic MPECs*, Comput. Optim. Appl., 50 (2011), pp. 111–145.

[43] M. Hintermüller and M.H. Tber, *An inverse problem in American options as a mathematical program with equilibrium constraints: C-stationarity and an active-set-Newton solver*, SIAM J. Control Optim., 48 (2010), pp. 4419–4452.

[44] K. Itŏ and K. Kunisch, *Optimal control of obstacle problems by H^1-obstacles*, Appl. Math. Optim., 56 (2007), pp. 1–17.

[45] W. Kampowsky, *Notwendige Optimalitätsbedingungen für Prozesse in Evolutionsgleichungen,* Inaugural-Dissertation, Universität Greifswald, Greifswald, Germany, 1979.

[46] N. Kenmochi, M. Niezgódka, and I. Pawłow, *Subdifferential operator approach to the Cahn–Hilliard equation with constraint*, J. Differential Equations, 117 (1995), pp. 320–356.

[47] Z.-Q. Luo, J.-S. Pang, and D. Ralph, *Mathematical Programs with Equilibrium Constraints*, Cambridge University Press, Cambridge, UK, 1997.

[48] M. Marcus and V.J. Mizel, *Every superposition operator mapping one Sobolev space into another is continuous*, J. Funct. Anal., 33 (1979), pp. 217–229.

[49] H. Maurer and J. Zowe, *First and second order necessary and sufficient optimality conditions for infinite-dimensional programming problems*, Math. Programming, 16 (1979), pp. 98–110.

[50] B.S. Mordukhovich, *Variational Analysis and Generalized Differentiation,* I *Basic Theory*, Grundlehren Math. Wiss. 330, Springer-Verlag, Berlin, 2006.

[51] B.S. Mordukhovich, *Variational Analysis and Generalized Differentiation.* II *Applications*, Grundlehren Math. Wiss. 331, Springer-Verlag, Berlin, 2006.

[52] P. Neittaanmaki, J. Sprekels, and D. Tiba, *Optimization of Elliptic Systems. Theory and Applications*, Springer Monogr. Math., Springer, New York, 2006.
[53] P. Neittaanmäki and D. Tiba, *Optimal Control of Nonlinear Parabolic Systems: Theory, Algorithms, and Applications*, Pure Appl. Math., Marcel Dekker, New York, 1994.
[54] A. Novick-Cohen and L.A. Segel, *A gradually slowing travelling band of chemotactic bacteria*, J. Math. Biol., 19 (1984), pp. 125–132.
[55] Y. Oono and S. Puri, *Study of phase-separation dynamics by use of cell dynamical systems.* I. *Modeling*, Phys. Rev. A, 38 (1988), pp. 434–453.
[56] J. Outrata, M. Kočvara, and J. Zowe, *Nonsmooth Approach to Optimization Problems with Equilibrium Constraints. Theory, Applications and Numerical Results*, Nonconvex Optim. Appl. 28, Kluwer Academic Publishers, Dordrecht, 1998.
[57] H. Scheel and S. Scholtes, *Mathematical programs with complementarity constraints: Stationarity, optimality, and sensitivity*, Math. Oper. Res., 25 (2000), pp. 1–22.
[58] I. Steinbach, *Phase-field models in materials science*, Modelling and Simulation in Materials Science and Engineering, 17 (2009), 073001.
[59] H. Triebel, *Interpolation Theory, Function Spaces, Differential Operators*, 2nd revised and enlarged ed., Johann Ambrosius Barth, Heidelberg, 1995.
[60] Q.-F. Wang and S. Nakagiri, *Weak solutions of Cahn-Hilliard equations having forcing terms and optimal control problems*, in Mathematical Models in Functional Equations (Kyoto, 1999), Research Institute for Mathematical Sciences 1128, Kyoto University, Kyoto, 2000, pp. 172–180 (in Japanese).
[61] D. Wegner, *Solvability of Pseudomonotone-Strongly Monotone Systems and Application to a Model of Phase Separation*, Ph.D. thesis, Humboldt Universität zu Berlin, Berlin, 2010.
[62] J. Yong and S. Zheng, *Feedback stabilization and optimal control for the Cahn-Hilliard equation*, Nonlinear Anal., 17 (1991), pp. 431–444.
[63] J. Zowe and S. Kurcyusz, *Regularity and stability for the mathematical programming problem in Banach spaces*, Appl. Math. Optim., 5 (1979), pp. 49–62.