Universität Konstanz
Fakultät für Mathematik und Statistik

WS 2024/25

Master Thesis in Numerical Optimization

A Class of Optimal Control Problems Governed by Caginalp Phase Field Models

Author: Florian Schilberth
florian.schilberth@uni-konstanz.de

Supervisor and 1st Reviewer: Jun.-Prof. Dr. Behzad Azmi
behzad.azmi@uni-konstanz.de

2nd Reviewer: Prof. Dr. Stefan Volkwein

submitted at the

Universität Konstanz

September 14, 2025

Contents

Scope of the Thesis

The goal of this thesis is to set the theoretical framework for a numerical realization of an optimal control problem governed by Caginalp's phase field model. More precisely, let $\Omega \in \mathbb{R}^n$, $n = 2, 3$, be a bounded domain with sufficiently smooth boundary and $(0, T) \subset \mathbb{R}$, $T < 0$, be a finite time interval. Setting

$$\begin{aligned} Q_T &:= (0,T) \times \Omega, \\ \Sigma_T &:= (0,T) \times \partial\Omega, \end{aligned}$$

we consider the following problem for $\beta > 0$, the control $u = (g, h)$ and state $y = (\theta, \phi)$.

Optimal Control Problem (OP)

$$\min_{(u,y)\in(L^2(Q_T)^2)^2} J(u,y) := \frac{1}{2}\|y - y_d\|^2_{L^2(Q_T)^2} + \frac{\beta}{2}\|u\|^2_{L^2(Q_T)^2},$$

subject to Caginalp's phase field model

$$\begin{aligned} \partial_t(\theta + \phi) - \Delta\theta &= g \qquad \text{in } Q_T, \\ \partial_t\phi - \Delta\mu &= h \qquad \text{in } Q_T, \\ \mu = -\Delta\phi + F'(\phi) - \theta & \qquad \text{in } Q_T, \end{aligned}$$

with Dirichlet boundary conditions

$$\theta = \phi = \Delta\phi = 0 \qquad \text{on } \Sigma_T,$$

and initial data

$$\theta(0) = \theta_0, \quad \phi(0) = \phi_0 \qquad \text{in } \Omega.$$

For sure, what is meant by setting the theoretical framework for OP leaves room for interpretation. In this thesis, we dedicate one chapter to each of the following three aspects:

- ▷ Deriving Caginalp's model from phase separation phenomena using basic thermodynamics (Chapter 1).
- ▷ Showing well-posedness of Caginalp's model (Chapter 2).
- ▷ Proving the existence of an optimal control and deriving a first order necessary optimality condition (Chapter 3).

Caginalp's phase field model was introduced to describe phase transition processes such as melting and solidification [16], but can also be more generally viewed as a model for non-isothermal phase separation processes [5]. Core of the model is the Cahn-Hilliard equation, which was proposed to describe the so-called spinodal decomposition of binary alloys in a rapid cooling process, but can be utilized for phase-separating mixtures in a much broader range [22]. Chapter 1 starts with the investigation of a binary mixture (Section 1.1) that undergoes the isothermal phase separation process called spinodal decomposition (Section 1.2). With the gathered results, we can quite easily derive the Cahn-Hilliard equation by using basic laws of thermodynamics (Section 1.3). Coupling the Cahn-Hilliard equation with a heat balance equation, we arrive at Caginalp's model (Section 1.4).

In Chapter 2, we dive into pure mathematics and show that for arbitrary initial condition and right hand side, Caginalp's model admits a unique solution. From a mathematical point of view, a first glance at Caginalp's model already reveals two major obstacles for the proof. Firstly, we see that there is a nonlinear term $F'(\phi)$ and secondly, we spot the Bilaplacian Δ^2, both together making the equation nonlinear and of fourth order. Hence, after stating helpful results from functional analysis (Section 2.1), we tackle these problems in two steps. We start by omitting the nonlinearity $F'(\phi)$, resulting in a linear model of fourth order (Section 2.2), and try to find appropriate function spaces, or regularities, for which we can show well-posedness (Section 2.2.1). Then, we try to adapt these function spaces such that they allow to make suitable estimates for our nonlinearity $F'(\phi)$ (Section 2.2.2). With the right spaces and estimates at hand, we add the nonlinearity and show well-posedness of Caginalp's model (Section 2.3).

Turning to Chapter 3 and considering OP, it is not enough to know that for each control we have a unique state, but we moreover need to find a control-state pair minimizing the cost functional J. Hence, to solve OP, we prove the existence of an optimal control (Section 3.1). A numerical algorithm that intends to find this solution depends on conditions distinguishing the optimal control from any control. The well-posedness of the linearized Caginalp model is of crucial importance to state such a condition (Section 3.2). We conclude the thesis by deriving a first order necessary optimality condition for OP (Section 3.3).

1

Spinodal Decomposition, the Cahn-Hilliard Equation and Caginalp's Model

In this chapter, we want to present the Cahn-Hilliard equation as a continuous description of the so-called spinodal decomposition of a binary mixture in a rapid cooling process. To do so, we start by modeling a binary mixture and derive a measure of stability for it. After that, we establish the notion of spinodal decomposition as an isothermal phase separation process for binary mixtures and derive the Cahn Hilliard equation by using basic thermodynics [22]. In the final section, we present Caginalp's phase field model as a generalization of the Cahn-Hilliard equation to non-isothermal phase separation processes [5].

1.1 Binary Mixtures

The spinodal decomposition is a phenomenon of phase separation, describing for example when two liquids are miscible at high temperature, but separate into two distinct phases when the temperature is lowered [15]. To explore this phenomenon in more detail, we start by modeling the mixture and deriving a measure for its stability. This section is mainly based on [11].

The following model is oversimplified, but captures most of the important physics while remaining mathematically easy to handle [15]. We consider a binary mixture of two liquids A and B as a lattice of N_0 sites of equal volume V_0, where N_A and N_B particles of component A and B occupy the positions (see Figure 1.1). We assume that type A particles have the same volume as type B particles and that the total volume of the mixture is constant, independent of the composition. Moreover, neighboring sites are independent of each other, meaning that the type of adjacent particles does not affect the probability whether a site is occupied by a type A or a type B particle. The concentration of particles of type A is given by $\phi_A = \frac{N_A}{N_0}$, the one for type B analogously, so by assumption we have

$$\phi_A + \phi_B = 1.$$

Helmholtz free energy

The reason why a phase separation in our system occurs is because the initial state of the mixture is unstable relative to the final state [18]. In thermodynamics, the relative stability of a system at constant volume and pressure can be determined by its *Helmholtz free energy* (F)

$$F = U - TS,$$

where U is the internal energy, T temperature and S entropy. A system strives for the state where the Helmholtz free energy is lowest [15]. The internal energy arises from the total kinetic and potential energies of the particles, the entropy can be understood as a measure of the randomness of the system [18].

Entropy

The entropy S of our system can be expressed with the total number of configurations

$$W = \frac{N_0!}{N_A!(N_0 - N_A)!} = \frac{N_0!}{N_A!N_B!},$$

and is given by the *Boltzmann* formula

$$S = k_B \ln W,$$

where k_B is the Boltzmann constant. Applying the approximation of Stirling $\ln(N!) \approx N\ln(N) - N$ yields after some basic calculation

$$S = S_0 - k_B[N_B \ln \phi_B + N_A \ln \phi_A],$$

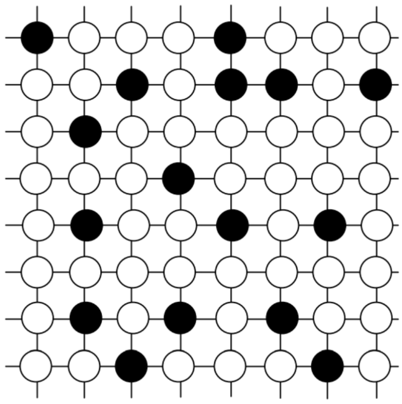

Figure 1.1
Binary mixture of A (white) and B (black) species. The particles are considered to have the same volume. Each particle can occupy only one of the 64 lattice sites and each lattice site is occupied by only one particle [11].

and we note that the entropy can be calculated by a logarithmic structure.

Internal energy

To approximate the internal energy U, we consider the interaction between particles at adjacent positions. Denoting the interaction energy of neighbors as ϵ_{AA}, ϵ_{AB} and ϵ_{BB}, depending on the type of particles interacting, the system's internal energy is given by

$$U = \epsilon_{AA}N_{AA} + \epsilon_{AB}N_{AB} + \epsilon_{BB}N_{BB},$$

where N_{AA} is the number of AA-bonds, N_{BB} and N_{AB} are defined analogously. We denote the coordination number of our lattice by C, so each site is directly bonded to C adjacent sites. Then, disregarding variations at the boundary, the total number of bonds in the lattice is given by $C \cdot N_0/2$, since for each pair of neighbors their bond is counted twice. Similarly, since an AA-bond is counted twice in $C \cdot N_A$, we can state the relation

$$CN_A = 2N_{AA} + N_{AB},$$

and analogously

$$CN_B = 2N_{BB} + N_{AB}.$$

With these relations, we can eliminate N_{AA} and N_{BB} in our expression for U getting

$$U = U_0 + \Big(\epsilon_{AB} - \frac{1}{2}\epsilon_{AA} - \frac{1}{2}\epsilon_{BB}\Big)N_{AB},$$

where $U_0 = C(\epsilon_{AA}N_A/2 + \epsilon_{BB}N_B/2)$. Next, we make the mean field assumption that a given site has $C\phi_A$ type A neighbors and $C\phi_B$ type B neighbors, no matter whether the site is itself occupied by an A or B particle. Then, again by avoiding counting bounds twice, we have

$$N_{AB} = \frac{N_A C\phi_B + N_B C\phi_A}{2} = N_0 C\phi_A\phi_B,$$

which allows us to eliminate N_{AB} [15].

Setting $\phi := \phi_A$, i.e. $\phi_B = 1-\phi$, and plugging U and S into the Helmholtz free energy equation $F = U - TS$, we arrive at

$$F = F_0 + N_0 C\Big(\epsilon_{AB} - \frac{1}{2}\epsilon_{AA} - \frac{1}{2}\epsilon_{BB}\Big)\phi(1-\phi) + k_B T N_0[(1-\phi)\ln(1-\phi) + \phi\ln\phi],$$

where $F_0 = \phi F_A + (1-\phi)F_B$ represents the free energy of a system with $N_A = \phi N_0$ sites filled by A and $N_B = (1-\phi)N_0$ sites filled by B particles, neglecting their interaction. The additional terms describe the change due to mixing.

Flory-Huggins parameter

To shorten the expression, we use the *Flory-Huggins parameter*

$$\chi = \frac{1}{k_B T}C\Big(\epsilon_{AB} - \frac{\epsilon_{AA} + \epsilon_{BB}}{2}\Big),$$

which can be viewed as an interaction parameter: $\chi > 0$ corresponds to $\epsilon_{AB} > \frac{\epsilon_{AA}+\epsilon_{BB}}{2}$, which promotes a tendency towards phase separation to reduce the internal energy of the system. On the other hand, $\chi < 0$ supports a single-phase mixture for all compositions.

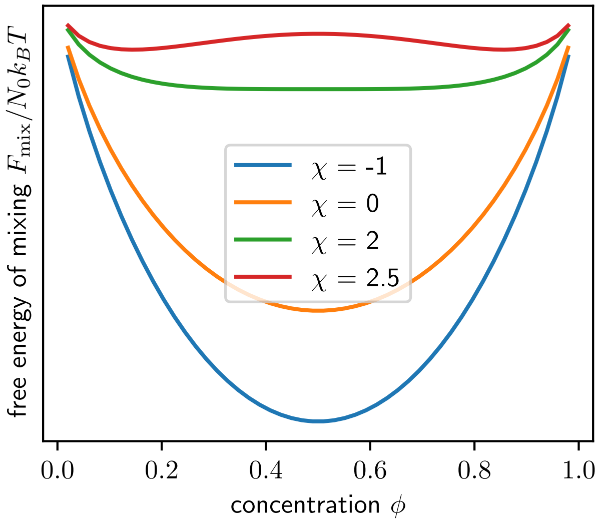

Figure 1.2
Free energy of mixing against composition for different values of χ. For negative χ, we find a single minimum, whereas for $\chi > 2$, there are two minima and a maximum.

Free energy of mixing

Finally, we omit the translation by F_0, which leads to the free energy of mixing

$$\frac{F_{\mathrm{mix}}}{N_0 k_B T} = \chi\phi(1-\phi) + (1-\phi)\ln(1-\phi) + \phi\ln\phi,$$

serving as our measure for stability of the mixture [15].

1.2 Spinodal Decomposition

Now that we have a model for a binary mixture, given by our lattice in Figure 1.1, and know how to quantify the stability of the mixture by computing the free energy of mixing, we can investigate the phenomenon of spinodal decomposition, following [15].

We start by plotting the curves of free energy F_{mix} against composition ϕ for $\chi = -1, 0, 2, 2.5$. The result is shown in Figure 1.2. For negative χ, we find a single minimum, whereas for $\chi > 2$, there are two minima and a maximum. This difference is of crucial physical significance.

Suppose that we have a homogeneous mixture of N_0 particles, i.e. the concentration of type A particles is ϕ_0 everywhere. If the mixture separates into a volume V_1 of N_1 particles, where the concentration of A particles is ϕ_1, and a volume V_2 of N_2 particles, where it is ϕ_2, then to conserve the total amount of type A particles, we must have $N_0\phi_0 = N_1\phi_1 + N_2\phi_2$. Dividing by N_0, we can write

$$\phi_0 = \alpha_1\phi_1 + \alpha_2\phi_2, \qquad \alpha_1 + \alpha_2 = 1. \tag{1.1}$$

The total free energy of the phase separated system can be approximated by $F_{\mathrm{sep}}(\phi_0) = \alpha_1 F_{\mathrm{mix}}(\phi_1) + \alpha_2 F_{\mathrm{mix}}(\phi_2)$, which we can rewrite using (1.1) as

$$F_{\mathrm{sep}} = \frac{\phi_0 - \phi_2}{\phi_1 - \phi_2} F_{\mathrm{mix}}(\phi_1) + \frac{\phi_1 - \phi_0}{\phi_1 - \phi_2} F_{\mathrm{mix}}(\phi_2).$$

As a function of ϕ_0, this expression describes the straight line joining $F_{\mathrm{mix}}(\phi_1)$ and $F_{\mathrm{mix}}(\phi_2)$, hence we can read off the free energy of the separated system $F_{\mathrm{sep}}(\phi_0)$ on that line at ϕ_0. An exemplary situation is shown in Figure 1.3. The initial composition $\phi_0 = 0.4$ separates into phases of concentration $\phi_1 = 0.2$ and $\phi_2 = 0.6$, such that the total concentration of A particles in the system remains ϕ_0. For $\chi = -1$, the free energy of the separated system $F_{\mathrm{sep}}(\phi_0)$, read off on the orange line, is higher than the free energy of the homogeneous mixture $F_{\mathrm{mix}}(\phi_0)$. Thus, phase separation does not lower the free energy so the mixture is stable against fluctuations in composition.

On the other hand, if $\chi = 2.5$, we have $F_{\mathrm{sep}}(\phi_0) < F_{\mathrm{mix}}(\phi_0)$. Here, any small fluctuation in composition results in an immediate phase separation, lowering the free energy of the system. The initial mixture is unstable.

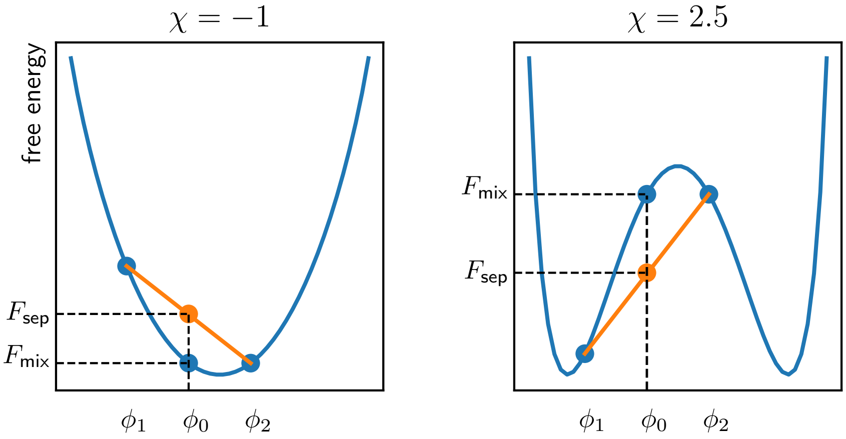

Figure 1.3
Free energy of the separated system against composition for different values of χ. We can read off $F_{\mathrm{sep}}(\phi_0)$ on the orange line joining $F_{\mathrm{mix}}(\phi_1)$ and $F_{\mathrm{mix}}(\phi_2)$ at ϕ_0. The value $F_{\mathrm{mix}}(\phi_0)$ is the free energy of the homogeneous mixture of concentration ϕ_0 without phase separation.

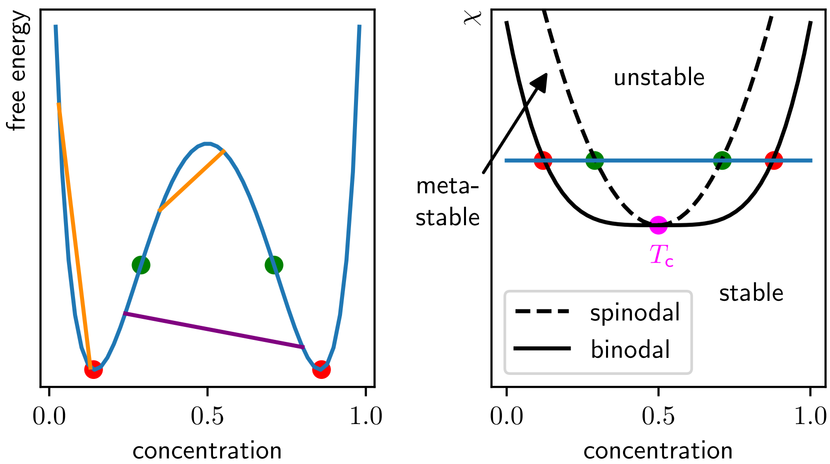

Figure 1.4
Stability analysis of the mixture. For compositions ϕ_0 outside the minimizers (red dots), any fluctuation in concentration increases the free energy (left orange line). For ϕ_0 between the locus of the inflection points (green dots) any small fluctuation decreases the free energy (right orange line). In between a minimizer and an inflection point small fluctuations increase, but bigger fluctuations can decrease the free energy (purple line). Plotting the limiting points as χ changes, we get stable, unstable and metastable regions, bounded by the binodal and spinodal. The critical temperature T_c marks the point (magenta dot), where phase separation starts to occur.

Binodal and spinodal

From a mathematical point of view, illustrated in Figure 1.4, the region where phase separation can lower the free energy is limited by the two minimizers, located at $F'_{\mathrm{mix}}(\phi) = 0$ (red dots). The curve of these minimizers as the temperature, and thus χ, changes is known as the *binodal*. In the region where $F_{\mathrm{mix}}(\phi)$ is concave, limited by the locus of the two inflection points at $F''_{\mathrm{mix}}(\phi) = 0$ (green dots), any small fluctuation in composition leads to phase separation. The locus of the inflection points as temperature changes is called *spinodal*. Outside the binodal, the initial mixture is stable and inside the spinodal, it is unstable. Between a minimizer and the locus of an inflection point, $F_{\mathrm{mix}}(\phi)$ is convex, so small fluctuations in composition increase the free energy. On the other hand, bigger fluctuations can lower the free energy (purple line), so in this region the mixture is called *metastable*. The point where binodal and spinodal meet marks a critical χ, respectively a critical temperature T_c, where phase separation starts to be energetically favorable for some compositions.

In the right diagram in Figure 1.4, we can see that if the interaction parameter χ is less than two, the mixture is completely miscible in all proportions, not showing a tendency towards phase separation. If χ is greater than two, there are initial proportions where the mixture is unstable or metastable. We mentioned that considering the internal energy, $\chi < 0$ means that mixing is energetically favourable, adding up to the increase of entropy. For $0 < \chi < 2$, mixing is energetically unfavourable, but the gain in entropy is large enough to offset this.

Interfacial tension

In the simplest interpretation, we have χ varies as $1/T$, so above the critical temperature T_c, the mixture will form a single phase for all ϕ_0. Below T_c, we find that a mixture will phase-separate for some ϕ_0 as soon as fluctuations appear. Between separated regions, there are interfaces possessing a certain property of tension, which contributes to the amount of free energy of the system. Hence, the equation

$$F_{\mathrm{sep}}(\phi_0) = \alpha_1 F_{\mathrm{mix}}(\phi_1) + \alpha_2 F_{\mathrm{mix}}(\phi_2)$$

can only be viewed as an approximation of the free energy of the separated system, since it omits the interfacial energy. We will improve our expression for the free energy accounting for the interfacial tension after summarizing the previous observations. At this point, note that when we approach T_{c}, we expect the distinction between two phases to vanish and thus their interface to disappear. Hence, as we approach T_{c}, the interfacial tension must approach zero.

Spinodal decomposition

To summarize our preparatory work, we consider a homogeneous binary mixture of concentration ϕ_0. We cool the system rapidly slightly below the critical temperature T_c, such that ϕ_0 lies within the spinodal, i.e. the system is unstable. Then, a small fluctuation in composition is immediately amplified driven by the reduction of the free energy. This process is called *spinodal decomposition*. Important observations and remarks are:

- Spinodal decomposition is a process of *conserved dynamics*. No material can leave or enter the system.
- Spinodal decomposition is a continuous process. In contrast, when we start with a concentration ϕ_0 within the metastable region, a large fluctuation must take place to start a phase separation, corresponding to high energy costs. This nucleus can then grow in size, which is known as *homogeneous nucleation*. In a real system, it is usually found that there are some impurity particles on which the new phase may be nucleated with lower activation energy. This is known as *heterogeneous nucleation*. In both cases, it is not possible for the mixture to phase-separate by a process in which the composition in a region changes continuously.
- In spinodal decomposition, material flows from regions of low concentration to regions of high concentration, which is a reversal of the usual situation, where we expect material to diffuse from regions of high concentration to regions of low concentration.
- During the process of spinodal decomposition, we assume the temperature of the system to be constant in time and space [5] and the latent heat to be zero. So limited growth due to latent heat has no chance to take effect. We will see later how we can generalize these assumptions [7].
- **Order parameter (phase field)** In general, the derivation of a phenomenological theory for a particular phase separation in a thermodynamical system requires two steps. Firstly, we have to identify a physical quantity that characterizes the difference between the two phases, which we call *order parameter* ϕ (also *phase field*). In our previous case, the order parameter was the concentration $\phi(t, x)$. Secondly, the equations of state are determined by constructing the Helmholtz free energy F as a function of ϕ and the absolute temperature T [5].

1.3 The Cahn-Hilliard Equation

Ginzburg-Landau free energy

We want to derive an equation describing the phenomenon of spinodal decomposition, referring to [22]. From now on, we denote the temperature by Θ instead of T, to preserve T as the upper bound for a time interval. So far, we have not accounted for the energy associated with the interfaces between the growing domains [15]. The theory of *Ginzburg* suggests that we should not consider sharp interfaces with vanishing interfacial energy, but instead small transition layers of finite but positive thickness carrying an amount of interfacial energy [5]. This amount depends on the gradient of composition $\nabla\phi$ [15], i.e. we add the gradient as a penalty term for the phase field ϕ, penalizing sudden changes with respect to the spatial variable x. So we start the derivation with an isotropic binary mixture of nonuniform composition at a fixed temperature Θ and consider the following Ginzburg-Landau type free energy,

$$E(\phi) = \int_\Omega \frac{\nu}{2}|\nabla\phi|^2 + F(\phi)\,\mathrm{d}x,$$

where $\Omega \subset \mathbb{R}^n$, $n = 2, 3$, is a bounded domain with sufficiently smooth boundary and F is the Helmholtz free energy. The constant $\nu > 0$ represents the thickness of the transition layers. We consider the local concentration order parameter $\phi(t,x) = \phi_A - \phi_B \subset [-1,1]$ over time $t \in (0,T)$, $T < \infty$, which is equivalent to $\phi = \phi_A$. The Helmholtz free energy F is then given by

$$F(s) = \frac{\Theta}{2}\big[(1+s)\ln(1+s) + (1-s)\ln(1-s)\big] - \frac{\Theta_c}{2}s^2,$$

where $\Theta_c > 0$ is the critical temperature of the system. We remember that if $0 < \Theta < \Theta_c$, F has a double-well structure with two minimizers in $(-1,1)$. When Θ is close to Θ_c, which we call *shallow quenching*, then using a Taylor expansion, F is often approximated by

$$F(s) \approx \Big(\frac{\Theta}{2} - \frac{\Theta_c}{2}\Big)s^2 + \frac{\Theta}{12}s^4.$$

The special case $\Theta = 3$ and $\Theta_c = 4$ and adding $1/4$ yields the mathematically convenient form

$$F(s) \approx \frac{1}{4}(1-s^2)^2.$$

Note that in this case, the two minimizers are shifted to ± 1.

General conserved dynamics

To state a first equation, we mention again that spinodal decomposition is a process of conserved dynamics, mathematically expressed by

$$\partial_t \int_\Omega \phi(t,x)\,\mathrm{d}x = 0 \qquad \text{for all } t \in (0,T). \tag{1.2}$$

In general, conserved dynamics is described by the continuity equation [5]

$$\partial_t\phi + \nabla\cdot J = 0 \qquad \text{in } (0,T)\times\Omega, \tag{1.3}$$
$$J\cdot n = 0 \qquad \text{on } (0,T)\times\partial\Omega, \tag{1.4}$$
$$\partial_n\phi = 0 \qquad \text{on } (0,T)\times\partial\Omega, \tag{1.5}$$

where J denotes the mass flux. Equation (1.3) can be viewed as a differential, i.e. local, form for the law of mass balance. In fact, integrating (1.3) over space and applying the divergence theorem yields

$$\partial_t \int_\Omega \phi\,\mathrm{d}x = -\int_\Omega \nabla\cdot J\,\mathrm{d}x = -\int_{\partial\Omega} J\cdot n\,\mathrm{d}S,$$

so the total change of concentration over time is given by the net amount of material passing through $\partial\Omega$. Due to boundary condition (1.4), which is usually called *no-flux condition*, we get the mass conservation (1.2). The physical interpretation of the homogeneous Neumann condition (1.5) is that the interface between two components intersects $\partial\Omega$ at a perfect angle of $\pi/2$.

Chemical potential

The driving force for the dynamics of a phase separation process is the chemical potential μ. Material will diffuse down the gradient of chemical potential, from regions of high chemical potential to regions of low chemical potential [15]. Hence, we may postulate the constitutive equation

$$J = -M(\phi)\nabla\mu, \tag{1.6}$$

where $M(\phi) > 0$ represents the diffusion mobility. We can find the chemical potential as the variational derivative of the free energy with respect to ϕ, and after a simple calculation using Green's first identity and (1.5), we get

$$\mu = \frac{\delta E(\phi)}{\delta\phi} = -\nu\Delta\phi + F'(\phi). \tag{1.7}$$

Cahn-Hilliard equation

Combining equations (1.3), (1.6) and (1.7), we arrive at the *Cahn-Hilliard equation*

$$\partial_t\phi = \nabla\cdot\big[M(\phi)\nabla\mu\big] \qquad \text{in } (0,T)\times\Omega,$$
$$\mu = -\nu\Delta\phi + F'(\phi) \qquad \text{in } (0,T)\times\Omega,$$

with the initial and boundary conditions

$$\phi(0,x) = \phi_0(x) \qquad \text{in } \Omega,$$
$$M(\phi)\nabla\mu\cdot n = 0 \qquad \text{on } (0,T)\times\partial\Omega,$$
$$\partial_n\phi = 0 \qquad \text{on } (0,T)\times\partial\Omega.$$

1.4 Caginalp's Phase-Field Model

As already mentioned, spinodal decomposition is an isothermal phase separation process, i.e. we assume the temperature of the system during the process to be constant in Ω and that there are no effects due to latent heat. So far, we started with a binary mixture at temperature Θ and cooled it down slightly below a critical temperature Θ_c such that the mixture is unstable and phase-separates. The separated phases were volumes V_1 of high concentration of type A and V_2 of high concentration of type B. But what if instead we started with a pure type A liquid at a temperature Θ and cooled it down slightly below its freezing point Θ_c? We could consider separated phases as volumes V_1 of solid and V_2 of liquid type A, with the phase field values $\phi = \pm 1$ indicating pure solid or liquid. Then, it clearly would not seem reasonable to assume that latent heat does not influence the growth of phases. We need to find a non-isothermal generalization of the Cahn-Hilliad equation, where we consider $\Theta = \Theta(t,x)$ as a function in time and space [5].

Caginalp proposes a phase field model that starts with the Ginzburg-Landau type free energy

$$E_\theta(\phi) = \int_\Omega \frac{\nu}{2}|\nabla\phi|^2 + F(\phi) - \gamma_0\theta\phi \,\mathrm{d}x,$$

where $\theta(t,x) = \Theta(t,x) - \Theta_c$ is the time and space dependent reduced temperature [6] and γ_0 is a positive constant. We choose again the prototype double-well potential $F = \frac{1}{4}(1-\phi^2)^2$. The last term introduces the coupling between temperature and phase field and may be understood best contributing to the TS-part of the free energy [6]. The variational derivative of E_θ is given by

$$\mu = \frac{\delta E_\theta(\phi)}{\delta\phi} = -\nu\Delta\phi + F'(\phi) - \gamma_0\theta,$$

which yields the evolution equation [7]

$$\partial_t\phi = \nabla\cdot\big[M(\phi)\nabla\mu\big] \qquad \text{in } (0,T)\times\Omega, \tag{1.8}$$

$$\mu = -\nu\Delta\phi + F'(\phi) - \gamma_0\theta \qquad \text{in } (0,T)\times\Omega. \tag{1.9}$$

The heat balance for the system is given by

$$\partial_t(\theta + l_0\phi) = \Delta\theta \qquad \text{in } (0,T)\times\Omega, \tag{1.10}$$

where we add the latent heat of fusion $l_0 > 0$ multiplied with $\partial_t\phi$ to the heat equation [6]. Hence, changes in temperature over time are balanced with energy provided by or used for phase transitions.

Equations (1.8), (1.9) and (1.10) form *Caginalp's phase field model* (see [6, 5, 7] for the derivation, we state the version that can be found in [3]). We consider the special case $M(\phi) = 1$.

Caginalp's phase field model

Caginalp's phase field model

$$\partial_t(\theta + l_0\phi) - \Delta\theta = 0 \qquad \text{in } (0,T)\times\Omega,$$
$$\partial_t\phi - \Delta\mu = 0 \qquad \text{in } (0,T)\times\Omega,$$
$$\mu = -\nu\Delta\phi + F'(\phi) - \gamma_0\theta \qquad \text{in } (0,T)\times\Omega,$$

with Neumann-boundary conditions

$$\partial_n\theta = \partial_n\phi = \partial_n\mu = 0 \qquad \text{on } (0,T)\times\partial\Omega,$$

and initial data

$$\theta(0) = \theta_0, \quad \phi(0) = \phi_0 \qquad \text{in } \Omega.$$

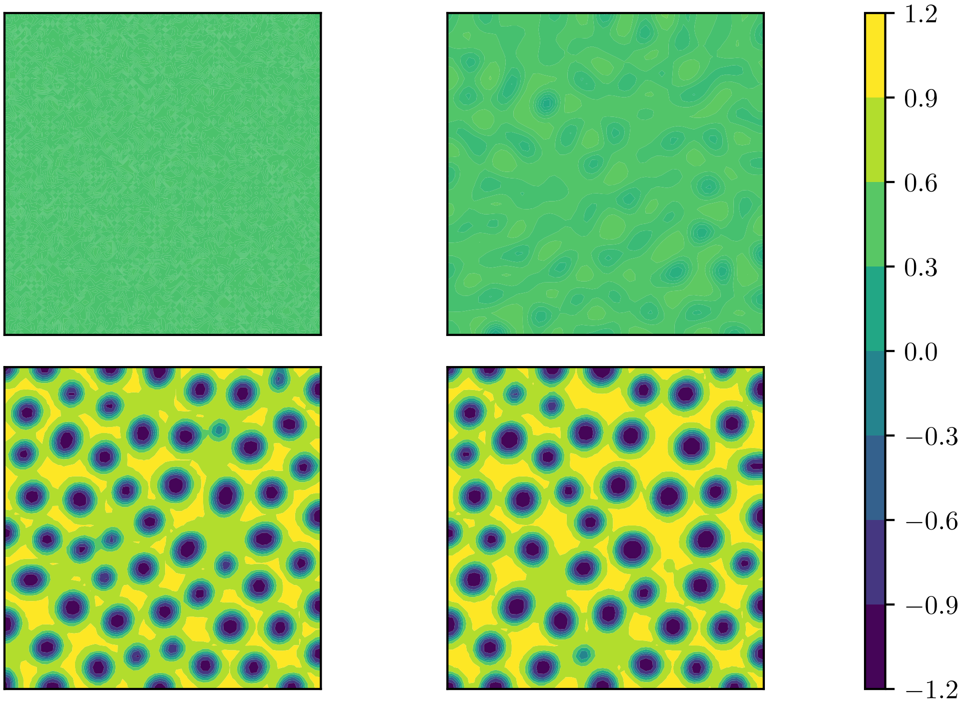

Figure 1.5
Evolution of phase field $\phi(t)$ described by Caginalp's model. At the beginning in the upper left plot, the phase field is almost constant (green) with only small fluctuations (yellow and blue). After some time (upper right and lower left plots), we can see that the fluctuations are expanding over time, generating regions of blue and yellow. In the lower right picture, the green transition areas from yellow to blue diminish.

Similarly to the derivation of conservation of mass, integrating the heat balance over space yields

$$\partial_t \int_\Omega (\theta + l_0\phi)\,\mathrm{d}x = \int_\Omega \Delta\theta\,\mathrm{d}x = \int_{\partial\Omega} \nabla\theta\cdot n\,\mathrm{d}S = \int_{\partial\Omega} \partial_n\theta\,\mathrm{d}S = 0.$$

Hence, due to (1.2), we get conservation of temperature [16].

With guidance of [19], we want to close this chapter by modeling a phase separation process with Caginalp's model. To do so, we consider a homogeneous binary mixture of $\hat{\phi}_0(x) = 0.4$ for every $x \in \Omega$. Small fluctuations are expressed by adding a random number $r(x) \in [-0.01, 0.01]$, so

$$\phi_0(x) = \hat{\phi}_0(x) + r(x).$$

In the same way, we set an initial temperature of

$$\theta_0(x) = 0.6 + r(x).$$

Moreover, as suggested before, we use the approximation $F(\phi) \approx \frac{1}{4}(1-\phi^2)^2$ with minimizers at ± 1. The resulting dynamics is shown in Figure 1.5. We see that the fluctuations at the beginning are amplified continuously, forming regions of yellow and blue. The concentrations within these regions are ± 1, corresponding to the minimizers of F. The green transition areas between the regions diminish, reflecting the presence of the penalty term $\frac{\nu}{2}|\nabla\phi|^2$ in the Ginzburg-Landau free energy.

2

Well-Posedness of Caginalp's Model

In this chapter, we proof existence and uniqueness of a weak solution to Caginalp's phase field model. To do so, we start by stating some fundamental results from functional analysis that we will need throughout the chapter. Then, we approach Caginalp's model in two steps: Firstly, we omit the nonlinearity and choose reasonable regularities for the inital data and right hand side to show existence and uniqueness for the resulting fourth order linear system. Secondly, we change the regularities in the linear model to the ones we need for the nonlinear model and again show existence and uniqueness. Afterwards, we are all set to prove existence and uniqueness for Caginalp's model. In all cases, the proofs consist of finding a suitable energy estimate and perform a Galerkin-approximation. We mainly refer to [17] and for general theory to [12] and [10]. Some inspiration is from [16].

Recall that we are always considering a bounded domain $\Omega \subset \mathbb{R}^n$, $n = 2, 3$, with sufficiently smooth boundary $\partial\Omega$, and a finite time interval $(0, T)$, $0 < T < \infty$. In the whole chapter, C is a positive constant, independent of (t, x), that may change from line to line, or even in the same line. For a Hilbert space X with dual space X', we denote inner product, norm and dual pairing as

$$\begin{aligned} &\langle \cdot\,, \cdot \rangle_X, \\ &\| \cdot \|_X, \\ &\langle \cdot\,, \cdot \rangle_{X',X}. \end{aligned}$$

Moreover, we write $\subset$ for subsets and embeddings and emphasize with $\hookrightarrow$ a continuous and with $\Subset$ a compact embedding. This does not mean that $\subset$ can't be continuous or compact. Furthermore, we denote the space of X-valued distributions on the set M by $\mathcal{D}'(M; X)$.

To simplify the mathematical analysis, we consider the Caginalp model with Dirichlet boundary conditions, although they do not provide the conservation of mass. Yet, one can obtain the same results for Neumann boundary conditions [16]. Moreover, we add a right hand side (g, h), which will serve as our control later. A suitable function space is to be determined. Lastly, we set all parameters equal to one. That being said, we proceed by dealing with the following model.

Caginalp's model with Dirichlet boundary conditions (CM)

$$\partial_t(\theta + \phi) - \Delta\theta = g \qquad \text{in } (0, T) \times \Omega, \tag{2.1}$$

$$\partial_t\phi - \Delta\mu = h \qquad \text{in } (0, T) \times \Omega, \tag{2.2}$$

$$\mu = -\Delta\phi + F'(\phi) - \theta \qquad \text{in } (0, T) \times \Omega, \tag{2.3}$$

with Dirichlet boundary conditions

$$\theta = \phi = \mu = 0 \qquad \text{on } (0, T) \times \partial\Omega, \tag{2.4}$$

and initial data

$$\theta(0) = \theta_0, \quad \phi(0) = \phi_0 \qquad \text{in } \Omega. \tag{2.5}$$

2.1 Preliminaries

The operator A If not mentioned otherwise, the following theory can be found in [17]. Let $V := H_0^1(\Omega)$, $H := L^2(\Omega)$ and define the linear operator

$$A : V \longrightarrow V', \qquad u \longmapsto \langle Au, \cdot \rangle_{V' \times V} = \langle \nabla u, \nabla \cdot \rangle_H = \langle u, \cdot \rangle_V.$$

Then, A is an isomorphism from V to V' and we can view A as the distributional version of $-\Delta$. We set

$$D(A) := A^{-1}(H) = \{u \in V \mid Au \in H\} = H^2(\Omega) \cap V,$$

where the last equality is non-trivial, and call it the domain of A. Well-known results from functional analysis equip A with the following spectral properties:

- The H-realization $A : H \supset D(A) \to H$ is self-adjoint, since $\langle \cdot, \cdot \rangle_V$ is symmetric, and positive, i.e. $\langle Au, u \rangle_H \geq 0$ for all $u \in D(A)$.
- Due to the compact embedding $D(A) \Subset H$, the inverse operator $A^{-1}\big|_H : H \to H$ is compact, furthermore self-adjoint and positive.
- The spectrum of A consists exclusively of eigenvalues $(\lambda_j)_{j \in \mathbb{N}}$ [10], which satisfy

  $$0 < \lambda_1 \;\leq\; \lambda_2 \;\leq\; \ldots \to \infty.$$

  There is an orthonormal basis $(w_j)_{j \in \mathbb{N}} \subset D(A)$ of $(H, \langle \cdot, \cdot \rangle_H)$ of corresponding eigenvectors, which are also orthogonal with respect to $\langle \cdot, \cdot \rangle_V$. Hence, for all $u \in H$, we get a Fourier series

  $$u = \sum_{j=1}^{\infty} u_j w_j, \;\; u_j \in \mathbb{R},$$

  and for all $u \in D(A)$, we have the representation

  $$Au = \sum_{j=1}^{\infty} \lambda_j u_j w_j.$$

- Considering the regularity of the eigenvectors, it holds that $w_k \in C^\infty(\Omega)$, $k \in \mathbb{N}$, and if $\partial\Omega$ is smooth, we even have $w_k \in C^\infty(\overline{\Omega})$ [12].
- For $u = \sum_{j=1}^{\infty} u_j w_j$, it holds that
  - if $u \in H$, then $\|u\|_H^2 = \sum_{j=1}^{\infty} |u_j|^2$.
  - if $u \in V$, then $\|u\|_V^2 = \sum_{j=1}^{\infty} \lambda_j |u_j|^2$.
  - if $u \in D(A)$, then for the graph norm, we have $\|u\|_{D(A)}^2 = \|Au\|_H^2 = \sum_{j=1}^{\infty} \lambda_j^2 |u_j|^2$.

The operators A^α In CM, we need to deal with the Bilaplacian, so in addition to define a distributional version of $-\Delta$, we also require a notion for Δ^2. Motivated by the previous results, we define operators A^α on domains $D(A^\alpha)$,

$$A^\alpha : H \subset D(A^\alpha) \longrightarrow H,$$

where we first consider $\alpha \geq 0$. For $u = \sum_{j=1}^{\infty} u_j w_j$, $v = \sum_{j=1}^{\infty} v_j w_j$, we set $A^0 = I$ on $D(A^0) = H$ and for $\alpha > 0$,

$$A^\alpha u := \sum_{j=1}^{\infty} \lambda_j^\alpha u_j w_j \qquad \text{on} \qquad D(A^\alpha) := \Big\{u \in H \;\Big|\; \sum_{j=1}^{\infty} \lambda_j^{2\alpha} |u_j|^2 < +\infty\Big\}.$$

We endow $D(A^\alpha)$ with the graph norm and scalar product

$$\|u\|_{D(A^\alpha)} := \|A^\alpha \cdot\|_H = \Big(\sum_{j=1}^{\infty} \lambda_j^{2\alpha} |u_j|^2\Big)^{\frac{1}{2}} \qquad \text{and} \qquad \langle u, v \rangle_{D(A^\alpha)} := \sum_{j=1}^{\infty} \lambda_j^{2\alpha} u_j v_j,$$

which makes $D(A^\alpha)$ a Hilbert space. With this definition, one can show that

$$D(A^{\frac{1}{2}}) = V,$$
$$D(A^2) = \{u \in D(A) \mid Au \in D(A)\} = \{u \in H^4(\Omega) \cap V \mid Au = 0 \text{ on } \partial\Omega\}.$$

Moreover, on $D(A^{\frac{k}{2}})$, $k = 1, 2, 4$, we have

$$\|\cdot\|_{D(A^{\frac{k}{2}})} \equiv \|\cdot\|_{H^k(\Omega)},$$

in the sense of equivalence of norms.

For $\alpha < 0$, we define $D(A^\alpha)$ as the topological dual of $D(A^{-\alpha})$, i.e.

$$D(A^\alpha) := D(A^{-\alpha})',$$

and A^α as the inverse of the adjoint operator of $A^{-\alpha}$, i.e. $A^\alpha := (A^{-\alpha})^*)^{-1}$. For example, we get the inverse mappings

$$A^{-1} : D(A^{-1}) \longrightarrow H,$$
$$A^{-1} : V' \longrightarrow V,$$
$$A^{-\frac{1}{2}} : V' \longrightarrow H.$$

Remark If $\alpha > \alpha' \in \mathbb{R}$, then $D(A^\alpha) \Subset D(A^{\alpha'})$ with dense and compact embedding. In particular, we have

$$D(A) \Subset V \Subset H \Subset V' \Subset D(A^{-1})$$

To finish the preliminares, we want to state useful results that we will use intensively in the following. Throughout the thesis, the statement of the first bullet will be called *Hölder's inequality*, and the one of the second bullet *Young's inequality.*

- [12] Let $1 \leq p_1, ..., p_m \leq \infty$, with $\frac{1}{p_1} + \cdots + \frac{1}{p_m} = \frac{1}{r}$, and assume $u_k \in L^{p_k}(M)$ for $k = 1, ..., m$. Then

$$\|u_1 \cdots u_m\|_{L^r(M)} \leq \prod_{k=1}^m \|u_k\|_{L^{p_k}(M)}.$$

- [12] Let $a, b > 0$. Then, for arbitrary $\epsilon > 0$, we have

$$ab \leq \epsilon a^2 + \frac{b^2}{4\epsilon}.$$

This version of the Hölder inequality is a slight generalization of the one in [12], but can be shown straightforward.

Theorem 2.1 Integration by parts [13] Let $\mathcal{V} \subset \mathcal{H} \subset \mathcal{V}'$ be a Gelfand triple of Hilbert spaces. Then, we have the continuous embedding

$$L^2(0, T; \mathcal{V}) \cap H^1(0, T; \mathcal{V}') \hookrightarrow C([0, T]; \mathcal{H}).$$

Moreover, for all $u, v \in L^2(0, T; \mathcal{V}) \cap H^1(0, T; \mathcal{V}')$, the integration by parts formula

$$\langle u(t), v(t)\rangle_{\mathcal{H}} - \langle u(s), v(s)\rangle_{\mathcal{H}} = \int_s^t \langle \partial_t u(\tau), v(\tau)\rangle_{\mathcal{V}' \times \mathcal{V}} + \langle \partial_t v(\tau), u(\tau)\rangle_{\mathcal{V}' \times \mathcal{V}} \, \mathrm{d}\tau$$

holds for all $t, s \in (0, T)$.

Immediate consequences of Theorem 2.1 for $u \in L^2(0, T; \mathcal{V}) \cap H^1(0, T; \mathcal{V}')$ are

- $\frac{1}{2}\partial_t \|u(t)\|_{\mathcal{H}}^2 = \langle \partial_t u(t), u(t)\rangle_{\mathcal{V}' \times \mathcal{V}}$ and
- $\partial_t \langle u(t), v\rangle_{\mathcal{H}} = \langle \partial_t u(t), v\rangle_{\mathcal{V}' \times \mathcal{V}}$ for all $v \in \mathcal{V}$.

Lemma 2.2 Agmon inequality [17] For $n = 2, 3$, we have for all $u \in D(A)$

$$\|u\|_{L^\infty(\Omega)} \leq C\|u\|_V^{\frac{1}{2}}\|u\|_{D(A)}^{\frac{1}{2}}.$$

Lastly, we state a version of Gronwall's lemma that can be found in [17], and always refer to this version within the thesis.

Lemma 2.3 Gronwall lemma [17] Let g, h and y be locally integrable functions such that y' is locally integrable and, for $t \geq t_0$, $t_0 \in \mathbb{R}$,

$$y' \leq gy + h.$$

Then, the following holds for $t \geq t_0$:

$$y(t) \leq e^{\int_{t_0}^t g(s)\,\mathrm{d}s} y(t_0) + e^{\int_{t_0}^t g(s)\,\mathrm{d}s} \int_{t_0}^t e^{-\int_{t_0}^s g(\xi)\,\mathrm{d}\xi} h(s)\,\mathrm{d}s.$$

2.2 The Linear Model

There are at least two difficulties to prove existence and uniqueness of solutions to CM. Firstly, we need to find suitable regularities for the right hand side (g, h) and initial data (θ_0, ϕ_0), insuring the existence of a solution to the fourth order equation governed by the Bilaplacian A^2. Secondly, we need to find properties and estimates handling the nonlinearity $F' = f$.

Transformed linear Caginalp model

To tackle these problems, we start by investigating the linear model associated with CM, i.e. we omit the nonlinear term in (2.3). Introducing the change of variable $z := \theta + \phi$ instead of θ, we transform (2.1)-(2.3) such that there remains only one time derivative in the heat balance equation (2.1). Regarding CM, ϕ has to be more regular than z, which means that we can retransform without loosing regularity. Keeping in mind the definition of the operator A of the preliminaries, we state the following model.

Transformed linear Caginalp model (TLCM)

$$\partial_t z + Az - A\phi = g \qquad \text{in } (0,T) \times \Omega, \tag{2.6}$$
$$\partial_t \phi + A^2\phi - Az + A\phi = h \qquad \text{in } (0,T) \times \Omega, \tag{2.7}$$

with Dirichlet boundary conditions

$$z = \phi = A\phi = 0 \qquad \text{on } (0,T) \times \partial\Omega, \tag{2.8}$$

and initial data

$$z(0) = \theta_0 + \phi_0 =: z_0, \qquad \phi(0) = \phi_0 \qquad \text{in } \Omega. \tag{2.9}$$

2.2.1 A First Version of Regularities

Suitable Hilbert spaces - version 1

We choose a first version of regularities such that we can use properties of Gelfand triples like Theorem 2.1. Hence, we consider the Gelfand triple of Hilbert spaces $W \Subset Z \Subset W'$, where

$$W := V \times D(A),$$
$$Z := H \times H.$$

We seek a solution (z, ϕ) of (2.6)-(2.9) in the space

$$L^2(0,T;W) \cap H^1(0,T;W'),$$

which means, according to Theorem 2.1, that $(z,\phi) \in C([0,T],Z)$. In regard of these regularities, by testing (2.6) with $v \in V$ and (2.7) with $w \in D(A)$, equations (2.6)-(2.7) are associated with the weak formulation

Weak formulation of TLCM - version 1

$$\langle \partial_t z, v\rangle_{V'\times V} + \langle A^{\frac12} z, A^{\frac12} v\rangle_H - \langle A\phi, v\rangle_H = \langle g, v\rangle_{V'\times V} \tag{2.10}$$

$$\langle \partial_t \phi, w\rangle_{D(A^{-1})\times D(A)} + \langle A\phi, Aw\rangle_H - \langle A^{\frac12} z, A^{\frac12} w\rangle_H + \langle A\phi, w\rangle_H = \langle h, w\rangle_{D(A^{-1})\times D(A)}, \tag{2.11}$$

as equations in $\mathcal{D}'(0,T;\mathbb{R})$.

Remark We could also consider (2.6) and (2.7) as a single equation

$$\partial_t \begin{pmatrix} z \\ \phi \end{pmatrix} + \begin{pmatrix} A & -A \\ -A & A^2 + A \end{pmatrix} \begin{pmatrix} z \\ \phi \end{pmatrix} = \begin{pmatrix} g \\ h \end{pmatrix}. \tag{2.12}$$

The weak formulation is attained by testing (2.12) with $(v,w) \in W$, and is given by adding (2.10) and (2.11), as an equation in $D'(0,T;\mathbb{R})$. Consequently, we have two points of view:

- TLCM consists of the two equations (2.6) in $D'(0,T;V')$ (or in $L^2(0,T;V')$) and (2.7) in $D'(0,T;D(A^{-1}))$ (or in $L^2(0,T;D(A^{-1}))$).
- TLCM consists of the equation (2.12) in $D'(0,T;W')$ (or in $L^2(0,T;W')$).

The points of view are equivalent and we switch our view frequently in the following.

To prove the existence of a solution, we rely on the Galerkin method. Thus, we start by showing the following energy estimate.

Energy estimate - version 1

Theorem 2.4 Let $(z_0,\phi_0) \in Z$ and $(g,h) \in L^2(0,T;W')$. Assume there exists a solution

$$(z,\phi) \in L^\infty(0,T;Z) \cap L^2(0,T;W) \cap H^1(0,T;W')$$

of (2.6)-(2.9). Then, it holds that

$$\begin{aligned}\|(\partial_t z, \partial_t \phi)\|^2_{L^2(0,T;W')} + \|(z,\phi)\|^2_{L^\infty(0,T;Z)} + \|(z,\phi)\|^2_{L^2(0,T;W)} \;&\le \\ \le\; C\big(\|(z_0,\phi_0)\|^2_Z + \|(g,h)\|^2_{L^2(0,T;W')}\big).&\end{aligned}$$

Proof. We split the proof in two steps.

Step 1: Under the assumptions of the theorem, it holds that

Estimate for (z,ϕ) - version 1

$$\|(z,\phi)\|^2_{L^\infty(0,T;Z)} + \|(z,\phi)\|^2_{L^2(0,T;W)} \;\le\; C\big(\|(z_0,\phi_0)\|^2_Z + \|(g,h)\|^2_{L^2(0,T;W')}\big).$$

To show this estimate, we test (2.10) with z and (2.11) with ϕ and use Theorem 2.1 to get

$$\frac12 \partial_t \|z\|^2_H + \|A^{\frac12} z\|^2_H = \langle A\phi, z\rangle_H + \langle g, z\rangle_{V'\times V},$$

$$\frac12 \partial_t \|\phi\|^2_H + \|A\phi\|^2_H + \|A^{\frac12}\phi\|^2_H = \langle A^{\frac12} z, A^{\frac12}\phi\rangle_H + \langle h, \phi\rangle_{D(A^{-1})\times D(A)}.$$

To the terms on the right hand side of the equalities, we apply Cauchy-Schwarz and

Young, which yields

$$\begin{aligned}
\langle A\phi, z\rangle_H &\le \frac{1}{2}\|A\phi\|_H^2 + \frac{1}{2}\|z\|_H^2,\\
\langle A^{\frac{1}{2}}z, A^{\frac{1}{2}}\phi\rangle_H &\le \frac{1}{2}\|A^{\frac{1}{2}}z\|_H^2 + \frac{1}{2}\|A^{\frac{1}{2}}\phi\|_H^2,\\
\langle g, z\rangle_{V'\times V} &\le C\|g\|_{V'}^2 + \frac{1}{4}\|A^{\frac{1}{2}}z\|_H^2,\\
\langle h, \phi\rangle_{D(A^{-1})\times D(A)} &\le C\|h\|_{D(A^{-1})}^2 + \frac{1}{4}\|A\phi\|_H^2.
\end{aligned}$$

Hence, adding both equalities, collecting terms and omitting the $A^{\frac{1}{2}}\phi$-term leads to

$$\frac{1}{2}\partial_t\|(z,\phi)\|_Z^2 + \frac{1}{4}\|A^{\frac{1}{2}}z\|_H^2 + \frac{1}{4}\|A\phi\|_H^2 \le \frac{1}{2}\|z\|_H^2 + C\|(g,h)\|_{W'}^2.$$

Adapting constants and adding $C\|\phi\|_H^2$ to the right hand side, we get

$$\partial_t\|(z,\phi)\|_Z^2 + \|(z,\phi)\|_W^2 \le C\big(\|(z,\phi)\|_Z^2 + \|(g,h)\|_{W'}^2\big). \tag{2.13}$$

This means in particular that

$$\partial_t\|(z,\phi)\|_Z^2 \le C\big(\|(z,\phi)\|_Z^2 + \|(g,h)\|_{W'}^2\big).$$

According to the version of the Gronwall lemma stated in Lemma 2.3, we can estimate

$$\|(z(t),\phi(t))\|_Z^2 \le e^{Ct}\|(z_0,\phi_0)\|_Z^2 + e^{Ct}\int_0^t e^{-Cs}C\|(g(s),h(s))\|_{W'}^2\,\mathrm{d}s.$$

Since $T<\infty$, it follows that

$$\|(z(t),\phi(t))\|_Z^2 \le C\big(\|(z_0,\phi_0)\|_Z^2 + \|(g,h)\|_{L^2(0,T;W')}^2\big). \tag{2.14}$$

Returning to (2.13), integration over time yields

$$\begin{aligned}
&\|(z(t),\phi(t))\|_Z^2 + \|(z,\phi)\|_{L^2(0,t;W)}^2 \le\\
&\quad\le C\Big(\int_0^t \|(z(s),\phi(s))\|_Z^2\,\mathrm{d}s + \|(g,h)\|_{L^2(0,t;W')}^2\Big) + \|(z_0,\phi_0)\|_Z^2.
\end{aligned}$$

Using (2.14) on the right hand side, we get for almost every $t\in(0,T)$ the estimate

$$\|(z(t),\phi(t))\|_Z^2 + \|(z,\phi)\|_{L^2(0,t;W)}^2 \le C\big(\|(z_0,\phi_0)\|_Z^2 + \|(g,h)\|_{L^2(0,T;W')}^2\big), \tag{2.15}$$

where we note that the right hand side is independent of t. Thus, we have by omitting $\|(z,\phi)\|_{L^2(0,t;W)}^2$ in (2.15),

$$\|(z,\phi)\|_{L^\infty(0,T;Z)}^2 \le C\big(\|(z_0,\phi_0)\|_Z^2 + \|(g,h)\|_{L^2(0,T;W')}^2\big),$$

and moreover, by omitting $\|(z(t),\phi(t))\|_Z^2$ in (2.15),

$$\|(z,\phi)\|_{L^2(0,t;W)}^2 \le C\big(\|(z_0,\phi_0)\|_Z^2 + \|(g,h)\|_{L^2(0,T;W')}^2\big).$$

Hence, for almost every $t\in(0,T)$, it holds that

$$\|(z,\phi)\|_{L^\infty(0,T;Z)}^2 + \|(z,\phi)\|_{L^2(0,t;W)}^2 \le 2C\big(\|(z_0,\phi_0)\|_Z^2 + \|(g,h)\|_{L^2(0,T;W')}^2\big).$$

Step 1 follows by passing $t\to T$.

Estimate for $(\partial_t z, \partial_t\phi)$ - version 1

Step 2: It holds that

$$\|(\partial_t z, \partial_t\phi)\|_{L^2(0,T;W')}^2 \le C\big(\|(z_0,\phi_0)\|_Z^2 + \|(g,h)\|_{L^2(0,T;W')}^2\big).$$

To show this, we solve (2.6) for $\partial_t z$ and (2.7) for $\partial_t\phi$ and apply the V'- and $D(A^{-1})$-norm, which leads to

$$\begin{aligned}
\|\partial_t z\|_{V'}^2 &\le C\big(\|Az\|_{V'}^2 + \|A\phi\|_{V'}^2 + \|g\|_{V'}^2\big),\\
\|\partial_t\phi\|_{D(A^{-1})}^2 &\le C\big(\|A^2\phi\|_{D(A^{-1})}^2 + \|Az\|_{D(A^{-1})}^2 + \|A\phi\|_{D(A^{-1})}^2 + \|h\|_{D(A^{-1})}^2\big).
\end{aligned}$$

Adding both inequalities, using the properties of the operators A^α and the continuous embedding $D(A) \hookrightarrow V$, we get

$$\|(\partial_t z, \partial_t \phi)\|_{W'}^2 \;\leq\; C\big(\|(z,\phi)\|_W^2 + \|(z,\phi)\|_Z^2 + \|(g,h)\|_{W'}^2\big). \tag{2.16}$$

Noting that

$$\int_0^T \|(z(s),\phi(s))\|_Z^2 \,\mathrm{d}s \;\leq\; \|(z,\phi)\|_{L^\infty(0,T;Z)}^2 T,$$

we carry out time integration on (2.16), which yields

$$\|(\partial_t z, \partial_t \phi)\|_{L^2(0,T;W')}^2 \;\leq\; C\big(\|(z,\phi)\|_{L^2(0,T;W)}^2 + \|(z,\phi)\|_{L^\infty(0,T;Z)}^2 + \|(g,h)\|_{L^2(0,T;W')}^2\big).$$

Then Step 2 immediately follows from Step 1.

Now, according to Step 1 and Step 2, we have

$$\begin{aligned}\|(\partial_t z, \partial_t \phi)\|_{L^2(0,T;W')}^2 + \|(z,\phi)\|_{L^\infty(0,T,Z)}^2 + \|(z,\phi)\|_{L^2(0,T;W)}^2 \;\leq\\ \leq\; C\big(\|(z_0,\phi_0)\|_Z^2 + \|(g,h)\|_{L^2(0,T;W')}^2\big),\end{aligned}$$

which completes the proof of Theorem 2.4. □

The idea of the Galerkin approximation is to find solutions of (2.6)-(2.9) on finite dimensional subspaces of Z exploiting Z. For each of these solutions, the above energy estimate holds, so we get a uniformly bounded sequence of solutions, which admits a weakly convergent subsequence. Its limit solves the problem on the whole space.

Theorem 2.5 Let $(z_0,\phi_0) \in Z$ and $(g,h) \in L^2(0,T;W')$. Then, (2.6)-(2.9) possesses a unique solution (z,ϕ) such that

$$(z,\phi) \in L^\infty(0,T;Z) \cap L^2(0,T;W) \cap H^1(0,T;W').$$

Proof. We split the proof in four steps.

Step A: Construction of solutions (z_m,ϕ_m) on finite-dimensional subspaces W_m.

Approximate solutions - version 1

Let $m \in \mathbb{N}$ and $(w_k)_{k\in\mathbb{N}} \subset D(A)$ be the set of eigenfunctions of A described in the preliminaries. We consider the following approximated problem: Find $(z_m,\phi_m) \in AC([0,T],W)$ of the form

$$z_m(t) = \sum_{k=1}^m c_k^m(t) w_k \qquad \text{and} \qquad \phi_m(t) = \sum_{k=1}^m d_k^m(t) w_k,$$

where $c_k^m, d_k^m : [0,T] \to \mathbb{R}$, such that

$$\partial_t \langle z_m, v\rangle_H + \langle A^{\frac12} z_m, A^{\frac12} v\rangle_H - \langle A\phi_m, v\rangle_H = \langle g, v\rangle_{V'\times V}, \tag{2.17}$$

$$\partial_t \langle \phi_m, w\rangle_H + \langle A\phi_m, Aw\rangle_H - \langle A^{\frac12} z_m, A^{\frac12} w\rangle_H + \langle A\phi_m, w\rangle_H = \langle h, w\rangle_{D(A^{-1})\times D(A)}, \tag{2.18}$$

for all $(v,w) \in \mathrm{Span}(w_1,...,w_m)^2 =: W_m$ as equations in $\mathcal{D}'(0,T;\mathbb{R})$. To get a well-posed problem, we have to adapt the initial condition. We project (z_0,ϕ_0) orthogonally in H onto W_m, i.e.

$$z_{0,m} := \sum_{k=1}^m \langle z_0, w_k\rangle_H w_k \qquad \text{and} \qquad \phi_{0,m} := \sum_{k=1}^m \langle \phi_0, w_k\rangle_H w_k,$$

and set the initial condition as

$$(z_m(0),\phi_m(0)) = (z_{0,m},\phi_{0,m}). \tag{2.19}$$

Note that a basis of W_m is given by the $2m$ vectors

$$\mathcal{B}_m := \{b_1, ..., b_{2m}\} = \left\{ \begin{pmatrix} w_1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ w_1 \end{pmatrix}, \begin{pmatrix} w_2 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ w_2 \end{pmatrix}, ..., \begin{pmatrix} w_m \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ w_m \end{pmatrix} \right\}.$$

Testing (2.17) and (2.18) with each vector of $\mathcal{B}_m$, we get for each $k = 1, ..., m$ the system of ODEs

$$\begin{aligned} \partial_t c_k^m(t) + \lambda_k c_k^m(t) - \lambda_k d_k^m(t) &= g_k(t), \\ \partial_t d_k^m(t) + \lambda_k^2 d_k^m(t) - \lambda_k c_k^m(t) + \lambda_k d_k^m(t) &= h_k(t), \\ (c_k^m(0), d_k^m(0)) &= (\langle z_0, w_k\rangle_H, \langle \phi_0, w_k\rangle_H), \end{aligned}$$

where $g_k(t) = \langle g, w_k\rangle_{V'\times V}$ and $h_k(t) = \langle h, w_k\rangle_{D(A^{-1})\times D(A)}$. Since $g_k(t) \in L^2(0,T;\mathbb{R}) \subset L^1(0,T;\mathbb{R})$, we can apply the theorem of Picard-Lindelöf in the global version for the homogeneous system and the variation of parameters for the inhomogeneous system. Hence, for each k we get a global solution $(c_k^m(t), d_k^m(t)) \in AC([0,T];\mathbb{R}^2) \cap H^1(0,T;\mathbb{R}^2)$ (see [14], Theorem 4 on page 16, in combination with [21], Lemma 1.1 on page 250). Collecting all these solutions, we find

$$\begin{aligned} \big(c^m(t), d^m(t)\big) := \big(c_1^m(t), ..., c_m^m(t), d_1^m(t), ..., d_m^m(t)\big) &\in AC([0,T], \mathbb{R}^{2m}), \\ \big(\partial_t c^m(t), \partial_t d^m(t)\big) &\in L^2(0,T;\mathbb{R}^{2m}), \end{aligned}$$

and then $(z_m, \phi_m) \in AC([0,T], W_m)$, $(\partial_t z_m, \partial_t \phi_m) \in L^2(0,T;W_m)$ defined as above solves (2.17)-(2.19) because of a simple linearity argument.

Convergence of approximate solutions - version 1

Step B: Convergence of the sequence of solutions $((z_m, \phi_m))_{m\in\mathbb{N}}$.

We can easily adapt the proof of Theorem 2.4 such that we can use the same energy estimate for (z_m, ϕ_m). Thus, the sequence $((z_m, \phi_m))_{m\in\mathbb{N}}$ is uniformly bounded with

$$\begin{aligned} \|(\partial_t z_m, \partial_t \phi_m)\|^2_{L^2(0,T;W')} + \|(z_m, \phi_m)\|^2_{L^\infty(0,T;Z)} + \|(z_m, \phi_m)\|^2_{L^2(0,T;W)} &\le \\ \le C\big(\|(z_0, \phi_0)\|_Z^2 + \|(g,h)\|^2_{L^2(0,T;W')}\big). \end{aligned}$$

Consequently, there exists a subsequence $((z_{m_l}, \phi_{m_l}))_{l\in\mathbb{N}}$, which we rename $((z_m, \phi_m))_{m\in\mathbb{N}}$, and functions $(z, \phi) \in L^2(0,T;W)$ and $(a,b) \in L^2(0,T,W')$ such that

$$(z_m, \phi_m) \rightharpoonup (z, \phi) \quad \text{weakly in } L^2(0,T;W), \tag{2.20}$$

$$(\partial_t z_m, \partial_t \phi_m) \rightharpoonup (a, b) \quad \text{weakly in } L^2(0,T;W'). \tag{2.21}$$

We need to show that $(a,b) = (\partial_t z, \partial_t \phi)$. Let $\varphi \in C_c^\infty(0,T;\mathbb{R})$ and $b_n \in \mathcal{B}_N$, $N \in \mathbb{N}$. Then, because of (2.21) and since W is reflexive, we have

$$\begin{aligned} &\langle (\partial_t z_m, \partial_t \phi_m), \varphi(t) b_n\rangle_{L^2(0,T;W')\times L^2(0,T;W)} \overset{m\to\infty}{\longrightarrow} \\ &\overset{m\to\infty}{\longrightarrow} \langle (a,b), \varphi(t) b_n\rangle_{L^2(0,T;W')\times L^2(0,T;W)}. \end{aligned}$$

On the other hand, noting that $\langle \cdot, u\rangle_{L^2(0,T;H)}$, $u \in L^2(0,T;H)$, defines an element of $L^2(0,T;W')$, the convergence (2.20) implies

$$\begin{aligned} &\langle (\partial_t z_m, \partial_t \phi_m), \varphi(t) b_n\rangle_{L^2(0,T;W')\times L^2(0,T;W)} = -\langle (z_m, \phi_m), \varphi'(t) b_n\rangle_{L^2(0,T;H)} \overset{m\to\infty}{\longrightarrow} \\ &\overset{m\to\infty}{\longrightarrow} -\langle (z,\phi), \varphi'(t) b_n\rangle_{L^2(0,T;H)} = \langle (\partial_t z, \partial_t \phi), \varphi(t) b_n\rangle_{L^2(0,T;W')\times L^2(0,T;W)}. \end{aligned}$$

Hence, using the linearity of the limit, it holds that $(a,b) = (\partial_t z, \partial_t \phi)$ tested with every function of the form

$$\sum_{n=1}^{2N} \varphi_n(t) b_n, \; \varphi_n \in C_c^\infty(0,T;\mathbb{R}), \; b_n \in \mathcal{B}_N.$$

Since N was arbitrary, these functions are dense in $L^2(0,T;W)$ [20], so we have $(a,b) = (\partial_t z, \partial_t \phi)$ in $L^2(0,T;W')$.

Existence of a weak solution - version 1

Step C: Existence of a weak solution (z,ϕ).

We show that the weak limit (z,ϕ) solves the original equation. Let $N\in\mathbb{N}$ and $\varphi_k^1,\varphi_k^2\in C_c^\infty(0,T;\mathbb{R})$, $k=1,...,N$. For $m\geq N$, we know that (z_m,ϕ_m) is a solution to (2.17)-(2.19) tested with every $b_n\in\mathcal{B}_N$. Hence, for all $k=1,...,N$, it holds for almost all $t\in(0,T)$ that

$$\begin{aligned}
&\varphi_k^1(t)\langle\partial_t z_m,w_k\rangle_{V'\times V}+\varphi_k^1(t)\langle A^{\frac12}z_m,A^{\frac12}w_k\rangle_H-\varphi_k^1(t)\langle A\phi_m,w_k\rangle_H=\\
&\qquad=\varphi_k^1(t)\langle g,w_k\rangle_{V'\times V},\\
&\varphi_k^2(t)\langle\partial_t\phi_m,w_k\rangle_{D(A^{-1})\times D(A)}+\varphi_k^2(t)\langle A\phi_m,Aw_k\rangle_H-\varphi_k^2(t)\langle A^{\frac12}z_m,A^{\frac12}w_k\rangle_H+\\
&\qquad\qquad+\varphi_k^2(t)\langle A\phi_m,w_k\rangle_H=\varphi_k^2(t)\langle h,w_k\rangle_{D(A^{-1})\times D(A)}.
\end{aligned}$$

Now, integrating over time yields

$$\begin{aligned}
&\int_0^T\varphi_k^1(t)\langle\partial_t z_m,w_k\rangle_{V'\times V}\,\mathrm{d}t+\int_0^T\varphi_k^1(t)\langle A^{\frac12}z_m,A^{\frac12}w_k\rangle_H\,\mathrm{d}t-\int_0^T\varphi_k^1(t)\langle A\phi_m,w_k\rangle_H\,\mathrm{d}t=\\
&\qquad=\int_0^T\varphi_k^1(t)\langle g,w_k\rangle_{V'\times V}\,\mathrm{d}t, &(2.22)\\
&\int_0^T\varphi_k^2(t)\langle\partial_t\phi_m,w_k\rangle_{D(A^{-1})\times D(A)}\,\mathrm{d}t+\int_0^T\varphi_k^2(t)\langle A\phi_m,Aw_k\rangle_H\,\mathrm{d}t-\\
&\qquad-\int_0^T\varphi_k^2(t)\langle A^{\frac12}z_m,A^{\frac12}w_k\rangle_H\,\mathrm{d}t\int_0^T\varphi_k^2(t)\langle A\phi_m,w_k\rangle_H\,\mathrm{d}t=\\
&\qquad=\int_0^T\varphi_k^2(t)\langle h,w_k\rangle_{D(A^{-1})\times D(A)}\,\mathrm{d}t. &(2.23)
\end{aligned}$$

Using that continuous operators are weakly continuous, we can pass to the limits (2.20) and (2.21). Again by a linearity argument, we see that both (2.22) and (2.23) hold for functions of the form

$$\sum_{k=1}^N\psi_k(t)w_k,\ \ \psi_k\in C_c^\infty(0,T;\mathbb{R}).$$

Since N was arbitrary, these functions are dense in $L^2(0,T;V)$ respectively $L^2(0,T;D(A))$, so (z,ϕ) solves (2.6) in $L^2(0,T;V')$ and (2.7) in $L^2(0,T;D(A^{-1}))$. By definition,

$$(z,\phi)\in L^2(0,T;W)\cap H^1(0,T;W'),$$

and the intersection is continuously embedded into $C([0,T];Z)$. As a consequence, we also have $(z,\phi)\in L^\infty(0,T;Z)$.

To complete step C, we need to show that (z,ϕ) satisfies the initial condition, i.e. $(z(0),\phi(0))=(z_0,\phi_0)$. Let $(v_N,y_N)\in W_N$ and $\varphi_1,\varphi_2\in C^1([0,T];\mathbb{R})$ with $\varphi_1(T)=\varphi_2(T)=0$ and $\varphi_1(0),\varphi_2(0)\neq0$. Then, we have for $m\geq N$ due to integration by parts

$$\begin{aligned}
&-\int_0^T\varphi_1'(t)\langle z_m,v_N\rangle_H\,\mathrm{d}t-\varphi_1(0)\langle z_m(0),v_N\rangle_H+\int_0^T\varphi_1(t)\langle A^{\frac12}z_m,A^{\frac12}v_N\rangle_H\,\mathrm{d}t-\\
&\qquad-\int_0^T\varphi_1(t)\langle A\phi_m,v_N\rangle_H\,\mathrm{d}t=\int_0^T\varphi_1(t)\langle g,v_N\rangle_{V'\times V}\,\mathrm{d}t, &(2.24)\\
&-\int_0^T\varphi_2'(t)\langle\phi_m,y_N\rangle_H\,\mathrm{d}t-\varphi_2(0)\langle\phi_m(0),y_N\rangle_H+\int_0^T\varphi_2(t)\langle A\phi_m,Ay_N\rangle_H\,\mathrm{d}t-\\
&\qquad-\int_0^T\varphi_2(t)\langle A^{\frac12}z_m,A^{\frac12}y_N\rangle_H\,\mathrm{d}t+\int_0^T\varphi_2(t)\langle A\phi_m,y_N\rangle_H\,\mathrm{d}t=\\
&\qquad=\int_0^T\varphi_2(t)\langle h,y_N\rangle_{D(A^{-1})\times D(A)}\,\mathrm{d}t. &(2.25)
\end{aligned}$$

Since orthogonal projections are self-adjoint, we have

$$\langle z_m(0),v_N\rangle_H=\langle z_{0,m},v_N\rangle_H=\langle z_0,v_{N,m})\rangle_H=\langle z_0,v_N\rangle_H,$$

and similarly

$$\langle \phi_m(0), y_N\rangle_H = \langle \phi_0, y_N\rangle_H.$$

On the other hand, we know that (z,ϕ) is a solution of (2.6)-(2.7) in $L^2(0,T;W')$, so testing (2.6) with $\varphi_1(t)v_N$ and (2.7) with $\varphi_2(t)y_N$ and integration by parts yields

$$\begin{aligned}
&-\int_0^T \varphi_1'(t)\langle z, v_N\rangle_H \,\mathrm{d}t - \varphi_1(0)\langle z(0), v_N\rangle_H + \int_0^T \varphi_1(t)\langle A^{\frac{1}{2}}z, A^{\frac{1}{2}}v_N\rangle_H \,\mathrm{d}t- \\
&\qquad - \int_0^T \varphi_1(t)\langle A\phi, v_N\rangle_H \,\mathrm{d}t = \int_0^T \varphi_1\langle g, v_N\rangle_{V'\times V}\,\mathrm{d}t, \\
&-\int_0^T \varphi_2'(t)\langle \phi, y_N\rangle_H \,\mathrm{d}t - \varphi_2(0)\langle \phi(0), y_N\rangle_H + \int_0^T \varphi_2(t)\langle A\phi, Ay_N\rangle_H \,\mathrm{d}t- \\
&\qquad - \int_0^T \varphi_2(t)\langle A^{\frac{1}{2}}z, A^{\frac{1}{2}}y_N\rangle_H \,\mathrm{d}t + \int_0^T \varphi_2(t)\langle A\phi, y_N\rangle_H \,\mathrm{d}t = \\
&\quad = \int_0^T \varphi_2(t)\langle h, y_N\rangle_{D(A^{-1})\times D(A)}\,\mathrm{d}t.
\end{aligned}$$

Passing $m\to\infty$ in (2.24) and (2.25) and comparing, we get

$$\langle z_0, v_N\rangle_H = \langle z(0), v_N\rangle_H \qquad \text{and} \qquad \langle \phi_0, y_N\rangle_H = \langle \phi(0), y_N\rangle_H,$$

for every $(v_N, y_N)\in W_N$. Since $\bigcup_{N=1}^\infty W_N$ is dense in W and hence in Z, we have $(z(0),\phi(0)) = (z_0,\phi_0)$ in Z.

Uniqueness of the weak solution - version 1

Step D: Uniqueness of the weak solution (z,ϕ).

Let $(z_1,\phi_1),(z_2,\phi_2)\in L^\infty(0,T;Z)\cap L^2(0,T;W)\cap H^1(0,T;W')$ be two solutions of (2.6)-(2.9). Then $(z,\phi) := (z_1 - z_2, \phi_1-\phi_2)$ is a solution with $(g,h) = (0,0)$ and $(z_0,\phi_0) = (0,0)$. According to the estimate of Theorem 2.4, it follows that $(z,\phi) = (0,0)$ and hence uniqueness of the solution. □

> **Remark** Note that we could omit the weak convergence (2.21) for the time derivatives in the proof. This is done by using integration by parts in time in Step C, to shove the time derivatives on the test functions.

To recap, Theorem 2.5 states that the action of the Bilaplacan A^2 shifts the regularity in space for the right hand side $D(A^{-1})$ to $D(A)$ and we are granted with a unique solution.

2.2.2 A Second Version of Regularities

It turns out that the regularities of version one are suitable for the linear case, but are not appropriate to deal with the nonlinearity. To illustrate this, we recall that the approximate nonlinear term is given by $F'(\phi) = \phi^3 - \phi$, which is a polynomial of third degree. Now, assuming our solution ϕ is in $L^\infty(0,T;H)$, the H-norm of the highest order term satisfies

$$\|\phi^3\|_H^2 = \|\phi\|_{L^6(\Omega)}^6,$$

hence we loose the L^∞-integrability in time.

Suitable Hilbert spaces - version 2

Apart from this example, we will see many other reasons why it is necessary to consider other Hilbert spaces. Proceeding in the same manner as for the first version, we set

$$\begin{aligned}
X_0 &:= V\times D(A^{\frac{3}{2}}), \\
Y &:= H\times V, \\
X_1 &:= V'\times V'.
\end{aligned}$$

Then, we seek a solution (z,ϕ) of (2.6)-(2.9) in the space

$$L^2(0,T;X_0)\cap H^1(0,T;X_1).$$

Since $X_0 \Subset Y \Subset X_1$ is not a Gelfand triple, we need a generalization of the first part of Theorem 2.1 to guarantee the continuity of the solution in time.

Theorem 2.6 [17] Let $\alpha \geq \alpha' \in \mathbb{R}$. Then, we have the continuous embedding

$$L^2(0,T;D(A^\alpha) \cap H^1(0,T;D(A^{\alpha'})) \hookrightarrow C([0,T];D(A^{\frac{\alpha+\alpha'}{2}})).$$

In particular, it follows that

$$L^2(0,T;X_0) \cap H^1(0,T;X_1) \hookrightarrow C([0,T];Y).$$

Weak formulation of TLCM - version 2

By testing (2.6) with $v \in V$ and (2.7) with $w \in V$, equations (2.6)-(2.7) are associated with the weak formulation

$$\langle \partial_t z, v\rangle_{V'\times V} + \langle A^{\frac{1}{2}} z, A^{\frac{1}{2}} v\rangle_H - \langle A\phi, v\rangle_H = \langle g, v\rangle_{V'\times V}, \tag{2.26}$$

$$\langle \partial_t \phi, w\rangle_{V'\times V} + \langle A^{\frac{3}{2}} \phi, A^{\frac{1}{2}} w\rangle_H - \langle A^{\frac{1}{2}} z, A^{\frac{1}{2}} w\rangle_H + \langle A\phi, w\rangle_H = \langle h, w\rangle_{V'\times V}, \tag{2.27}$$

as equations in $\mathcal{D}'(0,T;\mathbb{R})$.

Energy estimate - version 2

Now we are ready to execute the Galerkin method and start again by showing the following energy estimate.

Theorem 2.7 Let $(z_0,\phi_0) \in Y$ and $(g,h) \in L^2(0,T;X_1)$. Assume there exists a solution

$$(z,\phi) \in L^\infty(0,T;Y) \cap L^2(0,T;X_0) \cap H^1(0,T;X_1)$$

of (2.6)-(2.9). Then, it holds that

$$\|(\partial_t z, \partial_t \phi)\|^2_{L^2(0,T;X_1)} + \|(z,\phi)\|^2_{L^\infty(0,T;Y)} + \|(z,\phi)\|^2_{L^2(0,T;X_0)} \leq$$
$$\leq C\big(\|(z_0,\phi_0)\|^2_Y + \|(g,h)\|^2_{L^2(0,T;X_1)}\big)$$

Proof. We use the proof of Theorem 2.4 as guidance and adapt the new regularities. Thus, we split the proof in the same two steps.

Estimate for (z,ϕ) - version 2

Step 1: Under the assumptions of the theorem, it holds that

$$\|(z,\phi)\|^2_{L^\infty(0,T;Y)} + \|(z,\phi)\|^2_{L^2(0,T;X_0)} \leq C\big(\|(z_0,\phi_0)\|^2_Y + \|(g,h)\|^2_{L^2(0,T;X_1)}\big).$$

To show this estimate, we test (2.26) with z and (2.27) with $A\phi$ and use Theorem 2.1 to get

$$\frac{1}{2}\partial_t \|z\|^2_H + \|A^{\frac{1}{2}} z\|^2_H = \langle A\phi, z\rangle_H + \langle g, z\rangle_{V'\times V},$$

$$\frac{1}{2}\partial_t \|A^{\frac{1}{2}}\phi\|^2_H + \|A^{\frac{3}{2}}\phi\|^2_H + \|A\phi\|^2_H = \langle A^{\frac{1}{2}} z, A^{\frac{3}{2}}\phi\rangle_H + \langle h, A\phi\rangle_{V'\times V}.$$

We handle the terms on the right hand side of the equalities by applying Cauchy-Schwarz and Young, yielding

$$\begin{aligned}
\langle A\phi, z\rangle_H &\leq \frac{1}{2}\|A\phi\|^2_H + \frac{1}{2}\|z\|^2_H,\\
\langle A^{\frac{1}{2}} z, A^{\frac{3}{2}}\phi\rangle_H &\leq \frac{1}{2}\|A^{\frac{1}{2}} z\|^2_H + \frac{1}{2}\|A^{\frac{3}{2}}\phi\|^2_H,\\
\langle g, z\rangle_{V'\times V} &\leq C\|g\|^2_{V'} + \frac{1}{4}\|A^{\frac{1}{2}} z\|^2_H,\\
\langle h, A\phi\rangle_{V'\times V} &\leq C\|h\|^2_{V'} + \frac{1}{4}\|A^{\frac{3}{2}}\phi\|^2_H + C\|A^{\frac{1}{2}}\phi\|^2_H.
\end{aligned}$$

Hence, adding both equalities, collecting terms and omitting the $A\phi$-term leads to

$$\frac{1}{2}\partial_t \|(z,\phi)\|^2_Y + \frac{1}{4}\|A^{\frac{1}{2}} z\|^2_H + \frac{1}{4}\|A^{\frac{3}{2}}\phi\|^2_H \leq C\big(\|(z,\phi)\|^2_Y + \|(g,h)\|^2_{X_1}\big).$$

Adapting constants, we get

$$\partial_t \|(z,\phi)\|_Y^2 + \|(z,\phi)\|_{X_0}^2 \;\le\; C\big(\|(z,\phi)\|_Y^2 + \|(g,h)\|_{X_1}^2\big). \tag{2.28}$$

As before, applying the Gronwall Lemma yields

$$\|(z(t),\phi(t))\|_Y^2 \;\le\; C\big(\|(z_0,\phi_0)\|_Y^2 + \|(g,h)\|_{L^2(0,T;X_1)}^2\big).$$

Hence, after integrating (2.28) over time, we get for almost every $t \in (0,T)$

$$\begin{aligned}
&\|(z(t),\phi(t))\|_Y^2 + \|(z,\phi)\|_{L^2(0,t;X_0)}^2 \;\le \\
&\quad\le\; C\Big(\int_0^t \|(z(s),\phi(s))\|_Y^2 \,\mathrm{d}s + \|(g,h)\|_{L^2(0,t;X_1)}^2\Big) + \|(z_0,\phi_0)\|_Y^2 \;\le \\
&\quad\le\; C\big(\|(z_0,\phi_0)\|_Y^2 + \|(g,h)\|_{L^2(0,T;X_1)}^2\big).
\end{aligned}$$

The right hand side is independent of t, so we have

$$\|(z,\phi)\|_{L^\infty(0,T;Y)}^2 + \|(z,\phi)\|_{L^2(0,t;X_0)}^2 \;\le\; 2C\big(\|(z_0,\phi_0)\|_Y^2 + \|(g,h)\|_{L^2(0,T;X_1)}^2\big).$$

Step 1 follows by passing $t \to T$.

Estimate for $(\partial_t z, \partial_t \phi)$ - version 2

Step 2: It holds that

$$\|(\partial_t z, \partial_t \phi)\|_{L^2(0,T;X_1)}^2 \;\le\; C\big(\|(z_0,\phi_0)\|_Y^2 + \|(g,h)\|_{L^2(0,T;X_1)}^2\big).$$

As in the proof of Theorem 2.4, we solve (2.6) for $\partial_t z$ and (2.7) for $\partial_t \phi$, but this time apply the V'-norm to both equations, which leads to

$$\begin{aligned}
\|\partial_t z\|_{V'}^2 \;&\le\; C\big(\|Az\|_{V'}^2 + \|A\phi\|_{V'}^2 + \|g\|_{V'}^2\big), \\
\|\partial_t \phi\|_{V'}^2 \;&\le\; C\big(\|A^2\phi\|_{V'}^2 + \|Az\|_{V'}^2 + \|A\phi\|_{V'}^2 + \|h\|_{V'}^2\big).
\end{aligned}$$

Then, we can finish the proof with exactly the same reasoning as before. □

Having shown the energy estimate (2.7), we are set to execute the Galerkin approximation, resulting in the following analogon to Theorem 2.5.

Theorem 2.8 Let $(z_0,\phi_0) \in Y$ and $(g,h) \in L^2(0,T;X_1)$. Then, (2.6)-(2.9) possesses a unique solution (z,ϕ) such that

$$(z,\phi) \in L^\infty(0,T;Y) \cap L^2(0,T;X_0) \cap H^1(0,T;X_1).$$

Proof. The proof is done by going through the proof of Theorem 2.5 line by line and implementing the new regularities. So we consider the same four steps and proceed by making comments on each step.

Approximate solutions - version 2

Step A: Construction of solutions (z_m,ϕ_m) on finite-dimensional subspaces W_m.

As $(w_k)_{k\in\mathbb{N}} \subset D(A^{\frac{3}{2}})$, we can use the same basis funtions as before. The resulting systems of ODEs are the same, hence we find a solution $(z_m,\phi_m) \in C([0,T],W_m)$, $(\partial_t z_m, \partial_t \phi_m) \in L^2(0,T;W_m)$ of the approximated problem.

Convergence of approximate solutions - version 2

Step B: Convergence of the sequence of solutions $((z_m,\phi_m))_{m\in\mathbb{N}}$.

Due to Theorem 2.7, there exists a subsequence $((z_m,\phi_m))_{m\in\mathbb{N}}$ and a function $(z,\phi) \in L^2(0,T;X_0) \cap H^1(0,T;X_1)$ such that

$$(z_m,\phi_m) \rightharpoonup (z,\phi) \qquad \text{weakly in } L^2(0,T;X_0), \tag{2.29}$$

$$(\partial_t z_m, \partial_t \phi_m) \rightharpoonup (\partial_t z, \partial_t \phi) \qquad \text{weakly in } L^2(0,T;X_1). \tag{2.30}$$

Step C: Existence of a weak solution (z, ϕ).

Existence of a weak solution - version 2

This step is almost a copy and paste. Apart from changing the regularities, note the slight difference in the weak formulations (2.11) and (2.27), where we have

$$\langle A\phi, Aw\rangle_H \qquad \text{instead of} \qquad \langle A^{\frac{3}{2}}\phi, A^{\frac{1}{2}}w\rangle_H.$$

Step D: Uniqueness of the weak solution (z, ϕ).

Uniqueness of the weak solution - version 2

As before, we conclude uniqueness from the energy estimate of Theorem (2.7), using that for linear systems, the difference of two solutions solves TLCM with right hand side zero and initial data zero. □

As it will turn out, the second version of Hilbert spaces and regularities for the right hand side (g, h) and initial data (z_0, ϕ_0) is a suitable choice to deal with the whole nonlinear model CM. Hence, our next step is to include the nonlinearity in our investigations.

2.3 The Nonlinear Model

In Chapter 1, we discussed the role of the Helmholtz free energy F for the phase separation process. A system tends to phase-separate, if F has a double-well structure with two minima and the initial concentration ϕ_0 lies within the two minimizers. We can model F by the polynomial

$$F(s) = \frac{1}{4}(1-s^2)^2, \qquad \text{i.e.} \qquad f(s) := F'(s) = s^3 - s,$$

and immediately observe the following properties,

- $f'(s) = 3s^2 - 1 \geq -1$,
- $f(s)s = s^4 - s^2 \geq \frac{1}{2}(s^4 - 1)$.

Adding the nonlinearity f to TLCM, we state the following model, which is equivalent to CM, and for which we want to prove existence and uniqueness of solutions.

Transformed Caginalp model

Transformed Caginalp Model (TCM)

$$\partial_t z + Az - A\phi = g \qquad \text{in } (0,T)\times\Omega, \tag{2.31}$$
$$\partial_t \phi + A^2\phi - Az + A\phi + Af(\phi) = h \qquad \text{in } (0,T)\times\Omega, \tag{2.32}$$

with Dirichlet boundary conditions

$$z = \phi = A\phi = 0 \qquad \text{on } (0,T)\times\partial\Omega, \tag{2.33}$$

and initial data

$$z(0) = \theta_0 + \phi_0, \qquad \phi(0) = \phi_0 \qquad \text{in } \Omega. \tag{2.34}$$

We choose the same spaces and regularities as for the second version of the linear case, which means we set

Suitable Hilbert spaces - TCM

$$\begin{aligned} X_0 &:= V \times D(A^{\frac{3}{2}}), \\ Y &:= H \times V, \\ X_1 &:= V' \times V' \end{aligned}$$

and consider $(z_0, \phi_0) \in Y$ and $(g, h) \in L^2(0,T;X_1)$. Then, we seek a solution (z, ϕ) of (2.31)-(2.34) in the space

$$L^2(0,T;X_0) \cap H^1(0,T;X_1) \hookrightarrow C([0,T];Y).$$

In order to state well-defined expressions in the weak formulation and in the Galerkin method, we need to understand how the action of f affects the regularity of ϕ. Let $\phi \in L^2(0,T;D(A^{\frac{3}{2}})) \cap H^1(0,T;V')$, then it holds that

▷ $f(\phi) \in L^2(0,T;V)$: On the one hand, due to the continuous embeddings $H_0^1(\Omega) \hookrightarrow H^1(\Omega) \hookrightarrow L^6(\Omega)$ (see [4], Corollary 9.14 on page 284), we have

$$\|f(\phi)\|_H^2 \;\le\; \|\phi^3\|_H^2 + \|\phi\|_H^2 = \|\phi\|_{L^6(\Omega)}^6 + \|\phi\|_H^2 \;\le\; C(\|\phi\|_V^6 + \|\phi\|_V^2) \;\le\; C,$$

where C is independent of t. On the other hand, owing to the Hölder and Young, it holds that

$$\|f(\phi)\|_V^2 = \|\nabla f(\phi)\|_H^2 = \|(3\phi^2-1)\nabla\phi\|_H^2 \;\le\; C\big(\|\phi\|_{L^6(\Omega)}^4\|\nabla\phi\|_{L^6(\Omega)}^2 + \|\nabla\phi\|_H^2\big) \;\le\; \\ \le\; C\big(\|\phi\|_V^4\|A\phi\|_H^2 + \|\phi\|_V^2\big). \tag{2.35}$$

Integrating both estimates over time yields $f(\phi) \in L^2(0,T;V)$. Note that $f'(\phi) = 3\phi^2 - 1$ as a F-derivative is yet to justify. This will be done in Section 3.2.

▷ $f'(\phi) \in L^2(0,T;L^\infty(\Omega))$: Due to the Agmon inequality, we get

$$\|f'(\phi)\|_{L^\infty(\Omega)}^2 = \|3\phi^2-1\|_{L^\infty(\Omega)}^2 \;\le\; C\big(\|\phi\|_V^2\|\phi\|_{D(A)}^2 + \mathrm{Vol}(\Omega)\big). \tag{2.36}$$

Integrating over time yields $f'(\phi) \in L^2(0,T;L^\infty(\Omega))$.

Weak formulation - TCM Now, by testing (2.31) with $v \in V$ and (2.32) with $w \in V$, equations (2.31)-(2.32) are associated with the weak formulation

$$\langle \partial_t z, v\rangle_{V'\times V} + \langle A^{\frac12} z, A^{\frac12} v\rangle_H - \langle A\phi, v\rangle_H = \langle g, v\rangle_{V'\times V} \tag{2.37}$$

$$\langle \partial_t \phi, w\rangle_{V'\times V} + \langle A^{\frac32}\phi, A^{\frac12} w\rangle_H - \langle A^{\frac12} z, A^{\frac12} w\rangle_H + \langle A\phi, w\rangle_H + \langle A^{\frac12} f(\phi), A^{\frac12} w\rangle_H = \\ = \langle h, w\rangle_{V'\times V} \tag{2.38}$$

as equations in $\mathcal{D}'(0,T;\mathbb{R})$. Furthermore, we multiply (2.31) and (2.32) with A^{-1} to get

$$A^{-1}\partial_t z + z - \phi = A^{-1}g, \tag{2.39}$$

$$A^{-1}\partial_t\phi + A\phi - z + \phi + f(\phi) = A^{-1}h, \tag{2.40}$$

in $\mathcal{D}'(0,T;V\times V)$.

Before we start with the Galerkin method, we show an estimate that will help us significantly passing limits for sequences of approximate solutions.

Lemma 2.9 Let $\psi, \phi \in D(A)$. Then it holds that

$$\|f(\psi)-f(\phi)\|_V^2 \;\le\; C\big(1+\|(\psi,\phi)\|_{V\times V}^2\big)\big(1+\|(\psi,\phi)\|_{D(A)\times D(A)}^2\big)\|\psi-\phi\|_V^2,$$

for almost every $t \in (0,T)$.

Proof. For given $\psi, \phi \in D(A)$, we start by calculating

$$\|\nabla f(\psi) - \nabla f(\phi)\|_H^2 = \|f'(\psi)\nabla\psi - f'(\phi)\nabla\phi\|_H^2 = \\ = \|(f'(\psi)+f'(\phi))(\nabla\psi-\nabla\phi) - f'(\phi)\nabla\psi + f'(\psi)\nabla\phi\|_H^2. \tag{2.41}$$

To deal with the last two terms, we recall that $f'(\phi) = 3\phi^2 - 1$ and compute

$$-f'(\phi)\nabla\psi = -\big[(3\phi^2-1)\nabla\psi\big] = -\big[3\phi^2(\nabla\psi-\nabla\phi) + 3\phi^2\nabla\phi - \nabla\psi\big],$$

$$f'(\psi)\nabla\phi = (3\psi^2-1)\nabla\phi = 3(\psi^2-\phi^2)\nabla\phi + 3\phi^2\nabla\phi - \nabla\phi.$$

Plugging into (2.41) and adapting constants yields

$$\|\nabla f(\psi)-\nabla f(\phi)\|_H^2 \;\le\; C\big(\|(f'(\psi)+f'(\phi))(\nabla\psi-\nabla\phi)\|_H^2 + \|\phi^2(\nabla\psi-\nabla\phi)\|_H^2 + \\ + \|(\psi^2-\phi^2)\nabla\phi\|_H^2 + \|\psi-\phi\|_V^2\big). \tag{2.42}$$

We try to find estimates for each of the first three terms separately.

▷ Using (2.36) on the first term, we have

$$\|(f'(\psi)+f'(\phi))(\nabla\psi-\nabla\phi)\|_H^2 \;\le\; \big(\|f'(\psi)\|_{L^\infty(\Omega)}^2+\|f'(\phi)\|_{L^\infty(\Omega)}^2\big)\|\psi-\phi\|_V^2 \;\le\; C\big(\|\psi\|_V^2\|\psi\|_{D(A)}^2+\|\phi\|_V^2\|\phi\|_{D(A)}^2+1\big)\|\psi-\phi\|_V^2.$$

▷ Due to the Agmon inequality, the second term can be estimated by

$$\|\phi^2(\nabla\psi-\nabla\phi)\|_H^2 \;\le\; C\|\phi\|_V^2\|\phi\|_{D(A)}^2\|\psi-\phi\|_V^2.$$

▷ For the third term, we use Hölder to get

$$\|(\psi^2-\phi^2)\nabla\phi\|_H^2 \;\le\; \|\psi^2-\phi^2\|_{L^3(\Omega)}^2\|\nabla\phi\|_{L^6(\Omega)}^2 = \|(\psi-\phi)(\psi+\phi)\|_{L^3(\Omega)}^2\|\nabla\phi\|_{L^6(\Omega)}^2.$$

Again according to Hölder and the continuous embedding $V \subset L^6(\Omega)$, we can proceed by

$$\|(\psi-\phi)(\psi+\phi)\|_{L^3(\Omega)}^2\|\nabla\phi\|_{L^6(\Omega)}^2 \;\le\; \|\psi-\phi\|_{L^6(\Omega)}^2\|\psi+\phi\|_{L^6(\Omega)}^2\|\nabla\phi\|_{L^6(\Omega)}^2 \;\le\; \le\; C\|\psi-\phi\|_V^2\|\psi+\phi\|_V^2\|\phi\|_{D(A)}^2 \;\le\; C\big(\|\psi\|_V^2+\|\phi\|_V^2\big)\|\phi\|_{D(A)}^2\|\psi-\phi\|_V^2$$

Hence, turning back to (2.42), we can continue with

$$\begin{aligned}\|(f'(\psi)+f'(\phi))(\nabla\psi-\nabla\phi)\|_H^2+\|\phi^2(\nabla\psi-\nabla\phi)\|_H^2+\|(\psi^2-\phi^2)\nabla\phi\|_H^2+\|\psi-\phi\|_V^2 \;\le\;&\\ \le\; C\Big(\|\psi\|_V^2\|\psi\|_{D(A)}^2+\|\phi\|_V^2\|\phi\|_{D(A)}^2+1+\big(\|\psi\|_V^2+\|\phi\|_V^2\big)\|\phi\|_{D(A)}^2\Big)\|\psi-\phi\|_V^2 \;\le&\\ \le\; C\big(1+\|(\psi,\phi)\|_{V\times V}^2\big)\big(1+\|(\psi,\phi)\|_{D(A)\times D(A)}^2\big)\|\psi-\phi\|_V^2,&\end{aligned}$$

which completes the proof of Lemma 2.9. □

Now we have collected enough properties of the nonlinearity f to find solutions of the transformed Caginalp model. This time, our energy estimate reads as follows.

Theorem 2.10 Let $(z_0,\phi_0)\in Y$ and $(g,h)\in L^2(0,T;X_1)$. Assume there exists a solution

$$(z,\phi)\in L^\infty(0,T;Y)\cap L^2(0,T;X_0)\cap H^1(0,T;X_1)$$

of (2.31)-(2.34). Then, it holds that

$$\|(\partial_t z,\partial_t\phi)\|_{L^2(0,T;X_1)}^2+\|(z,\phi)\|_{L^\infty(0,T;Y)}^2+\|(z,\phi)\|_{L^2(0,T;X_0)}^2 \;\le\; C,$$

where C is independent of the solution and stays bounded with respect to bounded sequences of initial data $((z_0^m,\phi_0^m))_{m\in\mathbb{N}}\subset Y$ and controls $((g_m,h_m))_{m\in\mathbb{N}}\subset L^2(0,T;X_1)$.

Proof. We split the proof in three steps.

Auxiliary estimate for (z,ϕ) - TCM

Step 1: It holds that

$$\|(z,\phi)\|_{L^\infty(0,T;Z)}^2+\|(z,\phi)\|_{L^2(0,T;W)}^2 \;\le\; C.$$

We test (2.37) with z and (2.38) with ϕ, which yields

$$\frac{1}{2}\partial_t\|z\|_H^2+\|z\|_V^2 = \langle(A^{\frac12}\phi,A^{\frac12}z)\rangle_H+\langle g,z\rangle_{V'\times V},$$
$$\frac{1}{2}\partial_t\|\phi\|_H^2+\|A\phi\|_H^2+\|\phi\|_V^2+\langle A^{\frac12}f(\phi),A^{\frac12}\phi\rangle_H = \langle(A^{\frac12}z,A^{\frac12}\phi)\rangle_H+\langle h,\phi\rangle_{V'\times V}.$$

Since $f' \geq -1$ and applying Cauchy-Schwarz and Young to the terms on the right hand side of the equalities, we can estimate

$$\begin{aligned}
\langle A^{\frac{1}{2}} f(\phi), A^{\frac{1}{2}}\phi\rangle_H = \langle f'(\phi)A^{\frac{1}{2}}\phi, A^{\frac{1}{2}}\phi\rangle_H &\geq -\|\phi\|_V^2,\\
\langle (A^{\frac{1}{2}}\phi, A^{\frac{1}{2}}z)\rangle_H = \langle (A^{\frac{1}{2}}z, A^{\frac{1}{2}}\phi)\rangle_H &\leq \frac{1}{6}\|z\|_V^2 + C\|\phi\|_V^2,\\
\langle g, z\rangle_{V'\times V} &\leq C\|g\|_{V'}^2 + \frac{1}{6}\|z\|_V^2,\\
\langle h, \phi\rangle_{V'\times V} &\leq \frac{1}{2}\|h\|_{V'}^2 + \frac{1}{2}\|\phi\|_V^2.
\end{aligned}$$

Hence, adding both equalitites and collecting terms leads to

$$\partial_t\|(z,\phi)\|_Z^2 + \|z\|_V^2 + \|A\phi\|_H^2 \leq C\big(\|(g,h)\|_{X_1}^2 + \|\phi\|_V^2\big). \tag{2.43}$$

Noting that $\|\phi\|_V^2 \leq \frac{1}{2}\|Au\|_H^2 + \frac{1}{2}\|\phi\|_H^2$, adding $\|z\|_H^2$ on the right hand side and using the Gronwall lemma, we have

$$\|(z,\phi)\|_Z^2 \leq C\big(\|(z_0,\phi_0)\|_Z^2 + \|(g,h)\|_{L^2(0,T;X_1)}^2\big).$$

Thus, integrating (2.43) over time, we get

$$\|(z(t),\phi(t))\|_Z^2 + \|(z,\phi)\|_{L^2(0,T;W)}^2 \leq C\big(\|(z_0,\phi_0)\|_Z^2 + \|(g,h)\|_{L^2(0,T;X_1)}^2\big)$$

for almost every $t \in (0,T)$, showing Step 1.

Estimate for $(\partial_t z, \partial_t\phi)$ - TCM

Step 2: It holds that

$$\|(\partial_t z, \partial_t\phi)\|_{L^2(0,T;X_1)}^2 + \|\phi\|_{L^\infty(0,T;V)}^2 \leq C.$$

We test (2.39) with $\partial_t z$ and (2.40) with $\partial_t\phi$, which yields

$$\begin{aligned}
\|\partial_t z\|_{V'}^2 + \frac{1}{2}\partial_t\|z\|_H^2 &= \langle \partial_t z, \phi\rangle_{V'\times V} + \langle \partial_t z, A^{-1}g\rangle_{V'\times V},\\
\|\partial_t\phi\|_{V'}^2 + \frac{1}{2}\partial_t\|\phi\|_V^2 + \frac{1}{2}\partial_t\|\phi\|_H^2 + \langle \partial_t\phi, f(\phi)\rangle_{V'\times V} &= \langle \partial_t\phi, z\rangle_{V'\times V} + \langle \partial_t\phi, A^{-1}h\rangle_{V'\times V}.
\end{aligned}$$

Applying Cauchy-Schwarz and Young to the terms on the right hand side of the equalities, we get

$$\begin{aligned}
\langle \partial_t z, \phi\rangle_{V'\times V} &\leq \frac{1}{4}\|\partial_t z\|_{V'}^2 + C\|\phi\|_V^2,\\
\langle \partial_t\phi, z\rangle_{V'\times V} &\leq \frac{1}{4}\|\partial_t\phi\|_{V'}^2 + C\|z\|_V^2,\\
\langle \partial_t z, A^{-1}g\rangle_{V'\times V} &\leq \frac{1}{4}\|\partial_t z\|_{V'}^2 + C\|g\|_{V'}^2,\\
\langle \partial_t\phi, A^{-1}h\rangle_{V'\times V} &\leq \frac{1}{4}\|\partial_t\phi\|_{V'}^2 + C\|h\|_{V'}^2.
\end{aligned}$$

Hence, adding both equalities, collecting terms and using $\|\phi\|_V^2 \leq \frac{1}{2}\|A\phi\|_H^2 + \frac{1}{2}\|\phi\|_H^2$ leads to

$$\begin{aligned}
\|(\partial_t z, \partial_t\phi)\|_{X_1}^2 + \partial_t\|(z,\phi)\|_Z^2 + \partial_t\|\phi\|_V^2 + 2\langle \partial_t\phi, f(\phi)\rangle_{V'\times V} &\leq\\
\leq C\big(\|(g,h)\|_{X_1}^2 + \|(z,\phi)\|_W^2 + \|\phi\|_H^2\big).
\end{aligned} \tag{2.44}$$

Now, we set $\hat{F}(s) := \int_0^s f(\xi)\,\mathrm{d}\xi$. Then, it holds that

$$\langle \partial_t\phi, f(\phi)\rangle_{V'\times V} = \partial_t \int_\Omega \hat{F}(\phi)\,\mathrm{d}x.$$

Moreover, we know that $\hat{F}(s) = \frac{1}{4}s^4 - \frac{1}{2}s^2 \geq -\frac{1}{4}$ is bounded from below, and since Ω is bounded, it follows that $\int_\Omega \hat{F}(\phi(t))\,\mathrm{d}x \geq K$, where K is a lower bound independent of t.

Integrating (2.44) over time yields

$$\begin{aligned}\|(\partial_t z, \partial_t \phi)\|^2_{L^2(0,T;X_1)} + \|(z(t),\phi(t))\|^2_Z + \|\phi(t)\|^2_V + 2\int_\Omega \hat{F}(\phi(t))\,\mathrm{d}x \ \le \\ \le\ C\big(\|(z_0,\phi_0)\|^2_Z + \|\phi_0\|^2_V + 2\int_\Omega \hat{F}(\phi_0)\,\mathrm{d}x + \|(g,h)\|^2_{L^2(0,T;X_1)} + \\ + \|(z,\phi)\|^2_{L^2(0,T;W)} + \|(z,\phi)\|^2_{L^\infty(0,T;Z)}\big).\end{aligned}$$

for almost every $t \in (0,T)$. Furthermore, according to the theorem of Rellich-Kondrachov (see [4], Theorem 9.16 on page 285), we know that the embedding $V \Subset L^4(\Omega)$ is compact, so we have

$$\int_\Omega \hat{F}(\phi_0)\,\mathrm{d}x = \frac{1}{4}\|\phi_0^2\|^2_H - \frac{1}{2}\|\phi_0\|^2_H \ \le\ C(\|\phi_0\|^4_{L^4(\Omega)} + \|\phi_0\|^2_H)\ \le\ C(\|\phi_0\|^4_V + \|\phi_0\|^2_V).$$

Then, after canceling $2\int_\Omega \hat{F}(\phi(t))\,\mathrm{d}x$ by adding $2|K|$ on the right, Step 2 is a direct consequence of Step 1.

Step 3: It holds that **Estimate for (z,ϕ) - TCM**

$$\|(z,\phi)\|^2_{L^\infty(0,T;Y)} + \|(z,\phi)\|^2_{L^2(0,T;X_0)}\ \le\ C.$$

We test (2.37) with z and (2.38) with Au, which yields

$$\frac{1}{2}\partial_t\|z\|^2_H + \|z\|^2_V = \langle A\phi, z\rangle_H + \langle g, z\rangle_{V'\times V},$$

$$\frac{1}{2}\partial_t\|\phi\|^2_V + \|A^{\frac{3}{2}}\phi\|^2_H + \|A\phi\|^2_H + \langle A^{\frac{1}{2}} f(\phi), A^{\frac{3}{2}}\phi\rangle_H = \langle A^{\frac{1}{2}} z, A^{\frac{3}{2}}\phi\rangle_H + \langle h, A\phi\rangle_{V'\times V}.$$

Applying Cauchy-Schwarz and Young to the right hand sides of the equalities, we have

$$\begin{aligned}\langle A\phi, z\rangle_H &\le \frac{1}{2}\|A\phi\|^2_H + \frac{1}{2}\|z\|^2_H,\\ \langle A^{\frac{1}{2}} z, A^{\frac{3}{2}}\phi\rangle_H &\le \frac{1}{2}\|z\|^2_V + \frac{1}{2}\|A^{\frac{3}{2}}\phi\|^2_H,\\ \langle g, z\rangle_{V'\times V} &\le C\|g\|^2_{V'} + \frac{1}{4}\|z\|^2_V,\\ \langle h, A\phi\rangle_{V'\times V} &\le C\|h\|^2_{V'} + \frac{1}{8}\|A^{\frac{3}{2}}\phi\|^2_H.\end{aligned}$$

Moreover, we estimate the nonlinear term by

$$\langle A^{\frac{1}{2}} f(\phi), A^{\frac{3}{2}}\phi\rangle_H\ \le\ C\|f(\phi)\|^2_V + \frac{1}{8}\|A^{\frac{3}{2}}\phi\|^2_H.$$

Hence, adding both equalities and collecting terms yields

$$\partial_t\|(z,\phi)\|^2_Y + \|(z,\phi)\|^2_{X_0} + \|A\phi\|^2_H\ \le\ C\big(\|(g,h)\|^2_{X_1} + \|z\|^2_H + \|f(\phi)\|^2_V\big).$$

Using the estimate (2.35), we get

$$\partial_t\|(z,\phi)\|^2_Y + \|(z,\phi)\|^2_{X_0} + \|A\phi\|^2_H\ \le\ C\big(\|(g,h)\|^2_{X_1} + \|(z,\phi)\|^2_Y + \|\phi\|^4_V\|A\phi\|^2_H\big). \tag{2.45}$$

In Step 2, we have shown that $\|\phi\|_{L^\infty(0,T;V)}\ \le\ C$, and due to Step 1, we know that $\|\phi\|^2_{L^2(0,T;D(A))}\ \le\ C$, which means that we can apply the Gronwall lemma, yielding

$$\|(z(t),\phi(t))\|^2_Y\ \le\ C\big(\|(z_0,\phi_0)\|^2_Y + \|(g,h)\|^2_{L^2(0,T;X_1)} + \|\phi\|^4_{L^\infty(0,T;V)}\|\phi\|^2_{L^2(0,T;D(A))}\big).$$

Thus, integrating (2.45) over time, we have

$$\begin{aligned}\|(z(t),\phi(t))\|^2_Y + \|(z,\phi)\|^2_{L^2(0,T;X_0)} + \|A\phi\|^2_{L^2(0,T;D(A))}\ \le \\ \le\ C\big(\|(z_0,\phi_0)\|^2_Y + \|(g,h)\|^2_{L^2(0,T;X_1)} + \|\phi\|^4_{L^\infty(0,T;V)}\|\phi\|^2_{L^2(0,T;D(A))}\big),\end{aligned}$$

for almost every $t \in (0,T)$. This shows Step 3, hence the proof of Theorem 2.10 is complete, where the boundedness property of C can be seen keeping track of the estimates within the steps. □

We note that the energy estimate of Theorem 2.10 is different to the ones for the linear cases regarding the upper bound. For the linear cases, the upper bound can be expressed in terms of the right hand side of the equation (g,h) and the initial data (z_0,ϕ_0). This means that our solution is continuously dependent on the choices of (g,h) and (z_0,ϕ_0) and that the proof of uniqueness is immediate.

Theorem 2.11 Let $(z_0,\phi_0)\in Y$ and $(g,h)\in L^2(0,T;X_1)$. Then, (2.31)-(2.34) possesses a unique solution (z,ϕ) such that

$$(z,\phi)\in L^\infty(0,T;Y)\cap L^2(0,T;X_0)\cap H^1(0,T;X_1).$$

Proof. As already mentioned, we can not prove uniqueness of the solution in the same way as for the linear cases and there are other obstacles making the proof considerably more complex. Nevertheless, we use the proofs of the Theorems 2.5 and 2.8 as blueprints and split the proof in the same four steps.

Approximate solutions - TCM

Step A: Construction of solutions (z_m,ϕ_m) on finite-dimensional subspaces W_m.

Let $m\in\mathbb{N}$ and $(w_j)_{j\in\mathbb{N}}\subset D(A^{\frac{3}{2}})$ be the set of eigenfunctions of A described in the preliminaries. We consider the following approximated problem: Find $(z_m,\phi_m)\in C^1([0,T];W)$ of the form

$$z_m=\sum_{k=1}^m c_k^m(t)w_k \qquad\text{and}\qquad \phi_m=\sum_{k=1}^m d_k^m(t)w_k,$$

where $c_k^m, d_k^m:[0,T]\to\mathbb{R}$, such that

$$\partial_t\langle z_m,v\rangle_H+\langle A^{\frac{1}{2}}z_m,A^{\frac{1}{2}}v\rangle_H-\langle A\phi_m,v\rangle_H=\langle g,v\rangle_{V'\times V}, \tag{2.46}$$

$$\partial_t\langle \phi_m,w\rangle_H+\langle A^{\frac{3}{2}}\phi_m,A^{\frac{1}{2}}w\rangle_H-\langle A^{\frac{1}{2}}z_m,A^{\frac{1}{2}}w\rangle_H+\langle A\phi_m,w\rangle_H+\langle A^{\frac{1}{2}}f(\phi_m),A^{\frac{1}{2}}w\rangle_H= \\ =\langle h,w\rangle_{V'\times V} \tag{2.47}$$

for all $(v,w)\in\operatorname{Span}(w_1,...,w_m)^2=:W_m$ as equations in $D'(0,T;\mathbb{R})$. To get a well-posed problem, we have to adapt the initial condition. We project (z_0,ϕ_0) orthogonally in H onto W_m, i.e.

$$z_{0,m}:=\sum_{k=1}^m\langle z_0,w_k\rangle_H w_k \qquad\text{and}\qquad \phi_{0,m}:=\sum_{k=1}^m\langle \phi_0,w_k\rangle_H w_k,$$

and set the initial condition as

$$(z_m(0),\phi_m(0))=(z_{0,m},\phi_{0,m}). \tag{2.48}$$

Testing (2.46) and (2.47) with w_k, $k=1,...,m$, we get

$$\partial_t c_k^m(t)+\lambda_k c_k^m(t)-\lambda_k d_k^m(t)=g_k(t),$$
$$\partial_t d_k^m(t)+\lambda_k^2 d_k^m(t)-\lambda_k c_k^m(t)+\lambda_k d_k^m(t)+\langle f'(\phi_m)A^{\frac{1}{2}}\phi_m,A^{\frac{1}{2}}w_k\rangle_H=h_k(t),$$

where $g_k(t)=\langle g,w_k\rangle_{V'\times V}$ and $h_k(t)=\langle h,w_k\rangle_{V'\times V}$, together with the initial condition

$$(c_k^m(0),d_k^m(0))=(\langle z_0,w_k\rangle_H,\langle\phi_0,w_k\rangle_H).$$

Considering this system of ODEs, we note certain differences to the linear cases.

- ▷ The ODE is nonlinear, so we can not use the variation of parameters to find a global solution in AC.
- ▷ We can not use the theorem of Picard-Lindelöf, since the right hand sides are not continuous.

▷ In the linear cases, we had a collection of decoupled systems, meaning that we could solve a system of two ODEs for each k separately. Now, each term $\langle f'(\phi_m)A^{\frac{1}{2}}\phi_m, A^{\frac{1}{2}}w_k\rangle_H$ depends on every d_k^m, $k = 1, ..., m$.

Since $f'(\phi) = 3\phi^2 - 1$, we can write

$$\langle f'(\phi_m)A^{\frac{1}{2}}\phi_m, A^{\frac{1}{2}}w_k\rangle_H = 3\langle \phi_m^2 A^{\frac{1}{2}}\phi_m, A^{\frac{1}{2}}w_k\rangle_H - \lambda_k d_k^m(t).$$

Due to the continuous embedding $V \subset L^6(\Omega)$, we know that $w_k \in L^6(\Omega)$, hence $w_i w_j w_k \in H$, $i, j, k = 1, ..., m$. Thus, the term

$$P_k^m(d_1^m(t), ..., d_m^m(t)) := \langle \phi_m^2 A^{\frac{1}{2}}\phi_m, A^{\frac{1}{2}}w_k\rangle_H$$

is a polynomial of degree 3 in $d_k^m(t)$, $k = 1, ..., m$, which means locally Lipschitz continuous. Let $(g^m)_{m\in\mathbb{N}}, (h^m)_{m\in\mathbb{N}} \subset C([0,T]; V')$ such that

$$g^m \overset{m\to\infty}{\longrightarrow} g \qquad \text{and} \qquad h^m \overset{m\to\infty}{\longrightarrow} h \qquad \text{in } L^2(0,T;V').$$

Then, we consider the system of ODEs for $k = 1, ..., m$,

$$\begin{aligned} \partial_t c_k^m(t) + \lambda_k c_k^m(t) - \lambda_k d_k^m(t) &= g_k^m(t), \\ \partial_t d_k^m(t) + \lambda_k^2 d_k^m(t) - \lambda_k c_k^m(t) + P_k^m(d_1^m(t), ..., d_m^m(t)) &= h_k^m(t), \end{aligned}$$

where $g_k^m(t) = \langle g^m, w_k\rangle_{V'\times V}$ and $h_k^m(t) = \langle h^m, w_k\rangle_{V'\times V}$, together with the initial condition

$$(c_k^m(0), d_k^m(0)) = (\langle z_0, w_k\rangle_H, \langle \phi_0, w_k\rangle_H).$$

After this modification, the right hand sides are contionuous in time, so we can apply the theorem of Picard-Lindelöf in the global version to get a solution

$$\big(c^m(t), d^m(t)\big) = \big(c_1^m(t), ..., c_m^m(t), d_1^m(t), ..., d_m^m(t)\big) \in C^1([0,T_m), \mathbb{R}^{2m}).$$

Thus, $(z_m, \phi_m) \in C^1([0,T_m), W_m)$ defined as above solves (2.46)-(2.48) in $D'(0,T_m;\mathbb{R}^2)$, because of a simple linearity argument.

To show that $T_m = T$, we note that we can easily adapt the proof of Theorem 2.10 such that we can use the same energy estimate for (z_m, ϕ_m), with T_m instead of T. Here, we used in particular that $((g^m, h^m))_{m\in\mathbb{N}}$ is bounded in $L^2(0,T;V')$. Now, since all inner products on $\mathbb{R}^m$ are equivalent, we have

$$\begin{aligned} |c^m(t)|^2 &\le C\|z_m(t)\|_H^2, \\ |d^m(t)|^2 &\le C\|\phi_m(t)\|_V^2, \end{aligned}$$

which means that $|(c^m(t), d^m(t))|$ is uniformly bounded with respect to $t \in (0, T_m)$. Assuming that $T_m < T$, then by standard ODE theory we must have $|(c^m(t), d^m(t))| \to \infty$ as $t \to T_m$. This is a contradiction, hence $T_m > T$ and $(z_m, \phi_m) \in C^1([0,T], W_m)$.

Convergence of approximate solutions - TCM

Step B: Convergence of the sequence of solutions $((z_m, \phi_m))_{m\in\mathbb{N}}$.

The sequence $((z_m, \phi_m))_{m\in\mathbb{N}}$ satisfies the energy estimate of Theorem 2.10 and hence is uniformly bounded with

$$\|(\partial_t z_m, \partial_t \phi_m)\|_{L^2(0,T;X_1)}^2 + \|(z_m, \phi_m)\|_{L^\infty(0,T;Y)}^2 + \|(z_m, \phi_m)\|_{L^2(0,T;X_0)}^2 \le C.$$

Consequently, there exists a subsequence $((z_{m_l}, \phi_{m_l}))_{l\in\mathbb{N}}$, which we rename $((z_m, \phi_m))_{m\in\mathbb{N}}$, and functions $(z, \phi) \in L^\infty(0,T;Y) \cap L^2(0,T;X_0) \cap H^1(0,T;X_1)$ such that

$$(z_m, \phi_m) \rightharpoonup (z, \phi) \qquad \text{weakly in } L^2(0,T;X_0), \tag{2.49}$$

$$(\partial_t z_m, \partial_t \phi_m) \rightharpoonup (\partial_t z, \partial_t \phi) \qquad \text{weakly in } L^2(0,T;X_1), \tag{2.50}$$

$$(z_m, \phi_m) \overset{*}{\rightharpoonup} (z, \phi) \qquad \text{weak star in } L^\infty(0,T;Y). \tag{2.51}$$

In the linear cases, we have seen that the convergences (2.49) and (2.50) are enough to show that the limit (z, ϕ) is a solution to the equation. Here, the nonlinearity generates a term that is not a functional in $L^2(0,T;X_0')$, so for this term we need to find other ways to pass the limit. As it will turn out, the following convergences are helpful,

$$\phi_m \to \phi \qquad \text{strongly in } L^4(0,T;L^4(\Omega)), \tag{2.52}$$
$$\phi_m \to \phi \qquad \text{almost everywhere in } (0,T)\times\Omega, \tag{2.53}$$
$$\phi_m \to \phi \qquad \text{strongly in } L^2(0,T;V). \tag{2.54}$$

To prove them, we state the following lemma.

Lemma 2.12 AUBIN-LIONS COMPACTNESS RESULTS [17] Let $\mathcal{X}_0, \mathcal{X}$ and $\mathcal{X}_1$ be three Banach spaces such that $\mathcal{X}_0 \hookrightarrow \mathcal{X} \hookrightarrow \mathcal{X}_1$ with dense and continuous embeddings, where $\mathcal{X}_0 \Subset \mathcal{X}$ is compact. Let $1 \leq p < \infty$ and $1 \leq q \leq \infty$, and set

$$\mathcal{W} := \{u \in L^p(0,T;\mathcal{X}_0) \mid \partial_t u \in L^q(0,T;\mathcal{X}_1)\}.$$

Then, the embedding $\mathcal{W} \Subset L^p(0,T;\mathcal{X})$ is compact.

Since the embedding $V \Subset L^4(\Omega)$ is compact and $L^4(\Omega) \hookrightarrow L^2(\Omega) \hookrightarrow V'$ continuously, using Lemma 2.12, we have the compact embedding

$$L^4(0,T;V) \cap H^1(0,T,V') \Subset L^4(0,T;L^4(\Omega)). \tag{2.55}$$

We already know that the embedding

$$L^2(0,T;D(A^{\frac{3}{2}})) \cap H^1(0,T;V') \hookrightarrow C([0,T],V)$$

is continuous, so (2.49) and (2.50) yields

$$\phi_m \rightharpoonup \phi \qquad \text{weakly in } C([0,T],V).$$

Since $C([0,T],V) \hookrightarrow L^4(0,T;V)$ continuously, we get by (2.50)

$$\phi_m \rightharpoonup \phi \qquad \text{weakly in } L^4([0,T],V) \cap H^1(0,T;V').$$

Thus, due to (2.55) and properties of compact operators (see [9], Satz 5.4 on page 45), it holds that

$$\phi_m \to \phi \qquad \text{in } L^4(0,T;L^4(\Omega)).$$

Then, (2.53) is a direct consequence of (2.52) (see e.g. [8], Korollar 6.9 on page 43).

To prove (2.54), we note that since $D(A^{\frac{3}{2}}) \Subset V$ compactly and $V \hookrightarrow V'$ continuously, the embedding

$$L^2(0,T;D(A^{\frac{3}{2}})) \cap H^1(0,T;V') \Subset L^2(0,T;V)$$

is compact due to Lemma 2.12. Thus, by (2.49) and (2.50) we have the strong convergence

$$\phi_m \to \phi \qquad \text{in } L^2(0,T;V).$$

Existence of a weak solution - TCM

Step C: Existence of a weak solution (z,ϕ).

Proceeding in the same way as for the linear case, one term appears for which it is problematic to pass the limit,

$$\int_0^T \varphi(t)\langle A^{\frac{1}{2}} f(\phi_m), A^{\frac{1}{2}} w_k\rangle_H \,\mathrm{d}t,$$

where $\varphi(t) \in C_c^\infty(0,T;\mathbb{R})$ and w_k as before. We want to present two approaches to solve this problem.

▷ Theorem of Lebesgue: Recall that $f(\phi) = \phi^3 - \phi \in L^2(0,T;H)$, so we can write

$$\int_0^T \varphi(t)\langle A^{\frac{1}{2}} f(\phi_m), A^{\frac{1}{2}} w_k\rangle_H \,\mathrm{d}t = \int_0^T \varphi(t)\langle \phi_m^3, A w_k\rangle_H \,\mathrm{d}t - \int_0^T \varphi(t)\langle A^{\frac{1}{2}}\phi_m, A^{\frac{1}{2}} w_k\rangle_H \,\mathrm{d}t.$$

The second term on the right hand side defines a functional, hence we only need to show that

$$\int_0^T \varphi(t)\langle \phi_m^3, Aw_k\rangle_H \,\mathrm{d}t \overset{m\to\infty}{\longrightarrow} \int_0^T \varphi(t)\langle \phi^3, Aw_k\rangle_H \,\mathrm{d}t.$$

Applying the theorem of Hölder twice, we get

$$\begin{aligned}\left|\int_0^T \varphi(t)\langle \phi_m^3-\phi^3, Aw_k\rangle_H \,\mathrm{d}t\right| &\le \int_0^T \|\phi_m^3-\phi^3\|_{L^{\frac{4}{3}}(\Omega)}\|\varphi(t)Aw_k\|_{L^4(\Omega)}\,\mathrm{d}t\\ = \le\ &\|\phi_m^3-\phi^3\|_{L^{\frac{4}{3}}(0,T;L^{\frac{4}{3}}(\Omega))}\|\varphi(t)Aw_k\|_{L^4(0,T;L^4(\Omega))}.\end{aligned}$$

Thus, it suffices to show that ϕ_m^3 converges to ϕ^3 in the $L^{\frac{4}{3}}(0,T;L^{\frac{4}{3}}(\Omega))$-norm, i.e.

$$\Big(\int_0^T\int_\Omega |\phi_m^3-\phi^3|^{\frac{4}{3}}\,\mathrm{d}x\,\mathrm{d}t\Big)^{\frac{3}{4}} \overset{m\to\infty}{\longrightarrow} 0.$$

To do so, we want to use the theorem of Lebesgue (see [10], Satz B.6 on page 106). According to (2.53), we have $\phi_m^3\to\phi^3$ almost everywhere in $(0,T)\times\Omega$. Moreover, due to (2.52) and an inverse version of the theorem of Lebesgue (see [2], Theorem 1.2.7 on page 10), up to a subsequence we name $(\phi_m)_{m\in\mathbb{N}}$, we find a function $g\in L^4(0,T;L^4(\Omega))$ such that

$$|\phi_m| \le g \qquad \text{almost everywhere in } (0,T)\times\Omega.$$

Hence, for $m:=g^3\in L^{\frac{4}{3}}(0,T;L^{\frac{4}{3}}(\Omega))$, it holds that

$$|\phi_m^3| \le m \qquad \text{almost everywhere in } (0,T)\times\Omega.$$

Thus, the prerequisites of the theorem of Lebesgue are satisfied, so it holds that

$$\phi_m^3\to\phi^3 \qquad \text{in } L^{\frac{4}{3}}(0,T;L^{\frac{4}{3}}(\Omega)).$$

▷ Lemma 2.9: We have to show that

$$\int_0^T \varphi(t)\langle A^{\frac{1}{2}}f(\phi_m), A^{\frac{1}{2}}w_k\rangle_H\,\mathrm{d}t \overset{n\to\infty}{\longrightarrow} \int_0^T \varphi(t)\langle A^{\frac{1}{2}}f(\phi), A^{\frac{1}{2}}w_k\rangle_H\,\mathrm{d}t. \tag{2.56}$$

Using Cauchy-Schwarz, we get the estimate

$$\begin{aligned}&\left|\int_0^T \varphi(t)\langle A^{\frac{1}{2}}f(\phi_m)-A^{\frac{1}{2}}f(\phi), A^{\frac{1}{2}}w_k\rangle_H\,\mathrm{d}t\right|^2 \le\\ &\le \Big(\int_0^T \|A^{\frac{1}{2}}f(\phi_m)-A^{\frac{1}{2}}f(\phi)\|_H^2\,\mathrm{d}t\Big)\Big(\int_0^T \|\varphi(t)A^{\frac{1}{2}}w_k\|_H^2\,\mathrm{d}t\Big) \le\\ &\le C\int_0^T \|A^{\frac{1}{2}}f(\phi_m)-A^{\frac{1}{2}}f(\phi)\|_H^2\,\mathrm{d}t.\end{aligned} \tag{2.57}$$

Applying Lemma 2.9 yields

$$\|f(\phi_m)-f(\phi)\|_V^2 \le C\big(1+\|(\phi_m,\phi)\|_{V\times V}^2\big)\big(1+\|(\phi_m,\phi)\|_{D(A)\times D(A)}^2\big)\|\phi_m-\phi\|_V^2.$$

Due to the weak-*-convergence (2.51), we know that (ϕ_m) is bounded in $L^\infty(0,T;V)$ (see [9], Satz 2.7 on page 8), hence we can estimate (2.57) by

$$\begin{aligned}\int_0^T \|A^{\frac{1}{2}}f(\phi_m)-A^{\frac{1}{2}}f(\phi)\|_H^2\,\mathrm{d}t &\le C\int_0^T \big(1+\|(\phi_m,\phi)\|_{D(A)\times D(A)}^2\big)\|\phi_m-\phi\|_V^2\,\mathrm{d}t \le\\ &\le C\Big(\|\phi_m-\phi\|_{L^2(0,T;V)}^2+\int_0^T \|(\phi_m,\phi)\|_{D(A)\times D(A)}^2\|\phi_m-\phi\|_V^2\,\mathrm{d}t\Big) \le\\ &\le C\Big(\|\phi_m-\phi\|_{L^2(0,T;V)}^2+\\ &\quad+\|\phi_m-\phi\|_{L^\infty(0,T;V)}\int_0^T \|(\phi_m,\phi)\|_{D(A)\times D(A)}^2\|\phi_m-\phi\|_V\,\mathrm{d}t\Big).\end{aligned} \tag{2.58}$$

Now, using Cauchy-Schwarz, we have

$\|(\phi_m,\phi)\|^2_{D(A)\times D(A)} = \|\phi_m\|^2_{D(A)} + \|\phi\|^2_{D(A)} \;\le\; \|\phi_m\|_V\|\phi_m\|_{D(A^{\frac{3}{2}})} + \|\phi\|_V\|\phi\|_{D(A^{\frac{3}{2}})}$.

Thus, we get according to Cauchy-Schwarz

$$\begin{aligned}\int_0^T \|(\phi_m,\phi)\|^2_{D(A)\times D(A)}\|\phi_m-\phi\|_V\,\mathrm{d}t \;&\le\\ \le\; \|\phi_m\|_{L^\infty(0,T;V)}&\|\phi_m\|_{L^2(0,T;D(A^{\frac{3}{2}}))}\|\phi_m-\phi\|_{L^2(0,T;V)}+\\ &+\|\phi\|_{L^\infty(0,T;V)}\|\phi\|_{L^2(0,T;D(A^{\frac{3}{2}}))}\|\phi_m-\phi\|_{L^2(0,T;V)}.\end{aligned}$$

The weak convergence (2.49) implies that (ϕ_m) is bounded in $L^2(0,T;D(A^{\frac{3}{2}}))$ (see [9], Satz 2.7 on page 8), so continuing with (2.58), we arrive at

$$\int_0^T \|A^{\frac{1}{2}}f(\phi_m) - A^{\frac{1}{2}}f(\phi)\|^2_H\,\mathrm{d}t \;\le\; C\big(\|\phi_m-\phi\|^2_{L^2(0,T;V)} + \|\phi_m-\phi\|_{L^2(0,T;V)}\big).$$

Due to the strong convergence (2.54), this means that

$$\int_0^T \|A^{\frac{1}{2}}f(\phi_m) - A^{\frac{1}{2}}f(\phi)\|^2_H\,\mathrm{d}t \overset{m\to\infty}{\longrightarrow} 0,$$

which proves (2.56).

Choosing one of these approaches, we can finish Step C as demonstrated in the proof of the linear case in Theorem 2.5.

Uniqueness of the weak solution - TCM

Step D: Uniqueness of the weak solution (z,ϕ).

As already mentioned, this time we can not use the energy estimate of Theorem 2.10 to get uniqueness immediatley, so we have to find another way.

Let $(z_1,\phi_1),(z_2,\phi_2) \in L^\infty(0,T;Y)\cap L^2(0,T;X_0)\cap H^1(0,T;X_1)$ be two weak solutions of (2.31)-(2.34), as a system in $L^2(0,T;X_1)$. Then $(z,\phi) := (z_1-z_2,\phi_1-\phi_2)$ is a weak solution with $(g,h)=(0,0)$, $(z_0,\phi_0)=(0,0)$ and nonlinearity $f(\phi_1)-f(\phi_2)$. We consider again the reduced formulation (2.39) and (2.40) and test with (z,ϕ), which yields

$$\begin{aligned}\frac{1}{2}\partial_t\|z\|^2_{V'} + \|z\|^2_H \;&\le\; \langle\phi,z\rangle_H,\\ \frac{1}{2}\partial_t\|\phi\|^2_{V'} + \|\phi\|^2_V + \|\phi\|^2_H + \langle f(\phi_1)-f(\phi_2),\phi\rangle_H \;&\le\; \langle z,\phi\rangle_H.\end{aligned}$$

Applying Cauchy-Schwarz and Young, we have

$$\langle z,\phi\rangle_H = \langle\phi,z\rangle_H \;\le\; \frac{1}{2}\|z\|^2_H + \frac{1}{2}\|\phi\|^2_H,$$

so adding the equalities and collecting terms yields

$$\partial_t\|(z,\phi)\|^2_{X_1} + \|\phi\|^2_V + \langle f(\phi_1)-f(\phi_2),\phi\rangle_H \;\le\; 0. \tag{2.59}$$

Now, using the mean value theorem and $f'\ge -1$, we get

$$\langle f(\phi_1)-f(\phi_2),\phi\rangle_H = \langle f'(\xi_{t,x})\phi,\phi)\rangle_H \ge -\|\phi\|^2_H.$$

Furthermore, we have

$$\|\phi\|^2_H = \langle A^{-\frac{1}{2}}\phi, A^{\frac{1}{2}}\phi\rangle_H \;\le\; \frac{1}{2}\|\phi\|^2_{V'} + \frac{1}{2}\|\phi\|^2_V \;\le\; \frac{1}{2}\|(z,\phi)\|^2_{X_1} + \frac{1}{2}\|\phi\|^2_V.$$

Thus, plugging into (2.59) and omitting the $\|\phi\|^2_V$-term leads to

$$\partial_t\|(z,\phi)\|^2_{X_1} \;\le\; C\|(z,\phi)\|^2_{X_1}.$$

Applying the Gronwall lemma yields

$$\|(z,\phi)\|^2_{L^\infty(0,T;X_1)} \;\le\; C\|(z_0,\phi_0)\|^2_{X_1}.$$

Hence, it holds that $(z,\phi)=(0,0)$, which means uniqueness of the solution. □

Theorem 2.11 shows well-posedness of Caginalp's model, but we want to emphasize again that this does not include the continuous dependence on the right hand side and initial data. Nevertheless, with this knowledge at hand we can move forward to the optimal control theory part of this thesis.

3

Existence of an Optimal Control and Optimality Condition

In this chapter, we want to provide the theoretical background for a numerical realization of an optimal control problem governed by Caginalp's phase field model. We consider a quadratic cost functional, which means that we have properties like convexity and differentiability, and Caginalp's model serves as our only constraint. Then, in a first step, we prove well-posedness of the optimal control problem, or in other words, we prove the existence of an optimal control. To state a numerical algorithm that intends to find this optimal control, it is crucial to know necessary or sufficient optimality conditions. Hence, in a second step, we show well-posedness of the linearized Caginalp model. This model turns out to be a significant part in the derivation of a necessary first order optimality condition, which is done in a third step.

As before, let $\Omega \subset \mathbb{R}^n$, $n = 2, 3$ be a bounded domain with sufficiently smooth boundary $\partial\Omega$ and $0 < T < \infty$. We introduce the following notations for the space-time cylinder,

$$\begin{aligned} Q_T &:= (0,T) \times \Omega, \\ \Sigma_T &:= (0,T) \times \partial\Omega, \end{aligned}$$

Optimal control problem

Then, for $\beta > 0$, a control $u := (g,h) \in L^2(Q_T)^2$ and a state $y := (\theta,\phi) \in L^2(Q_T)^2$ we state the following optimal control problem.

Optimal Control Problem (OP)

$$\min_{(u,y)\in(L^2(Q_T)^2)^2} J(u,y) := \frac{1}{2}\|y - y_d\|^2_{L^2(Q_T)^2} + \frac{\beta}{2}\|u\|^2_{L^2(Q_T)^2}, \tag{3.1}$$

subject to Caginalp's model

$$\begin{aligned} \partial_t(\theta + \phi) - \Delta\theta &= g && \text{in } Q_T, \\ \partial_t\phi - \Delta\mu &= h && \text{in } Q_T, \\ \mu = -\Delta\phi + F'(\phi) - \theta & && \text{in } Q_T, \end{aligned}$$

with Dirichlet boundary conditions

$$\theta = \phi = \Delta\phi = 0 \qquad \text{on } \Sigma_T,$$

and initial data

$$\theta(0) = \theta_0, \quad \phi(0) = \phi_0 \qquad \text{in } \Omega.$$

To simplify our notations in the following, recall that $H = L^2(\Omega)$, $V = H^1_0(\Omega)$ and

$$\begin{aligned} X_0 &= V \times D(A^{\frac{3}{2}}), \\ Y &= H \times V, \\ X_1 &= V' \times V'. \end{aligned}$$

Moreover, we will denote the solution space as

$$\mathcal{W} := L^\infty(0,T;Y) \cap L^2(0,T;X_0) \cap H^1(0,T;X_1),$$

where $\mathcal{W}$ consists of the two components

$$\mathcal{W}_1 := L^\infty(0,T;H) \cap L^2(0,T;V) \cap H^1(0,T;V'),$$
$$\mathcal{W}_2 := L^\infty(0,T;V) \cap L^2(0,T;D(A^{\frac{3}{2}})) \cap H^1(0,T;V').$$

3.1 Existence of an Optimal Control

We want to start by showing well-posedness of OP and the ideas are adapted from [1]. Now, our work of Chapter 2 pays off for the first time, since according to Theorem 2.11, for fixed $(\theta_0, \phi_0) \in Y$, we can define an injective solution operator

$$S(y_0) : L^2(Q_T)^2 \longrightarrow \mathcal{W},$$
$$u \longmapsto y(u).$$

Furthermore, due to Theorem 2.10, we know that $y(u)$ satisfies the energy estimate

$$\|\partial_t y(u)\|^2_{L^2(0,T;X_1)} + \|y(u)\|^2_{L^\infty(0,T;Y)} + \|y(u)\|^2_{L^2(0,T;X_0)} \ \le\ C, \tag{3.2}$$

with a uniform C for bounded sequences of controls $(u_n)_{n\in\mathbb{N}}$. With these two results, we can prove the following theorem.

Theorem 3.1 For every finite horizon $T > 0$ and $(\theta_0, \phi_0) \in Y$, the optimal control problem OP admits a solution $(u, y) \in (L^2(Q_T)^2)^2$.

Proof. Theorem 2.11 guarantees that the admissible set of control state pairs (u, y) is nonempty and that is sufficient to consider the restriction

$$\min_{(u,y(u))\in L^2(Q_T)^2\times\mathcal{W}} J(u, y(u)).$$

Setting $y := y(u)$, since $J \ge 0$ is bounded from below, there exists an infimum

$$\overline{J} := \inf_{(u,y)\in L^2(Q_T)^2\times\mathcal{W}} J(u, y).$$

Hence, we can find a minimizing sequence $((u_n, y_n))_{n\in\mathbb{N}} \subset L^2(Q_T)^2 \times \mathcal{W}$ such that $J(u_n, y_n) \to \overline{J}$ as $n \to \infty$. In particular, due to convergence, the sequence $J(u_n, y_n)$ is bounded from above, i.e.

$$J(u_n, y_n) \ \le\ C.$$

Considering the structure of J, we can see that

$$J(u_n, y_n) \ge \frac{\beta}{2}\|u_n\|^2_{L^2(Q_T)^2}.$$

Moreover, since $(\|u_n\|_{L^2(Q_T)^2})_{n\in\mathbb{N}}$ is a bounded sequence of controls, the energy estimate (3.2) yields

$$\|y_n\|^2_{\mathcal{W}} \ \le\ C.$$

Combining the three inequalities, we have

$$\|(u_n, y_n)\|^2_{L^2(Q_T)^2\times\mathcal{W}} = \|u_n\|^2_{L^2(Q_T)^2} + \|y_n\|^2_{\mathcal{W}} \ \le\ C,$$

which means that $((u_n, y_n))_{n\in\mathbb{N}}$ is bounded in $L^2(Q_T)^2 \times \mathcal{W}$. Thus, there exists a subsequence renamed $((u_n, y_n))_{n\in\mathbb{N}} \subset L^2(Q_T)^2 \times \mathcal{W}$ and $(u^*, y^*) \in L^2(Q_T)^2 \times \mathcal{W}$ such that

$$u_n \rightharpoonup u^* \qquad \text{weakly in } L^2(Q_T)^2, \tag{3.3}$$
$$y_n \rightharpoonup y^* \qquad \text{weakly in } L^2(0,T;X_0), \tag{3.4}$$
$$\partial_t y_n \rightharpoonup \partial_t y^* \qquad \text{weakly in } L^2(0,T;X_1), \tag{3.5}$$
$$y_n \overset{*}{\rightharpoonup} y^* \qquad \text{weak star in } L^\infty(0,T;Y). \tag{3.6}$$

In Step B of the proof of Theorem 2.11, we have already shown that this implies

$$\phi_n \to \phi^* \qquad \text{strongly in } L^4(0,T;L^4(\Omega)), \tag{3.7}$$
$$\phi_n \to \phi^* \qquad \text{almost everywhere in } (0,T)\times\Omega, \tag{3.8}$$
$$\phi_n \to \phi^* \qquad \text{strongly in } L^2(0,T;V). \tag{3.9}$$

Now, there remain two things to show. Firstly, we need to prove that y^* is the solution of Caginalp's model corresponding to the control u^*, i.e. $y^* = y(u^*)$, and secondly, that (u^*, y^*) in fact solves OP. Starting with the first, as we have used in all of the existence proofs in Chapter 2, it is sufficient to show that y^* solves Caginalp's model with control u^* for all functions of the form

$$\{(\varphi_1(t)v, \varphi_2(t)w) \mid (v,w) \in V \times V,\ \varphi_1, \varphi_2 \in C_c^\infty(0,T;\mathbb{R})\},$$

which is a dense subset of $L^2(0,T;V\times V)$. By defintion, we know that each (u_n, y_n), $n \in \mathbb{N}$, solves

$$\begin{aligned}
&\langle \partial_t\theta_n, \varphi_1(t)v\rangle_{L^2(0,T,V')\times L^2(0,T;V)} + \langle \partial_t\phi_n, \varphi_1(t)v\rangle_{L^2(0,T,V')\times L^2(0,T;V)} + \\
&\quad + \langle \nabla\theta_n, \varphi_1(t)\nabla v\rangle_{L^2(0,T;H)} = \langle g_n, \varphi_1(t)v\rangle_{L^2(0,T;H)}, \\
&\langle \partial_t\phi_n, \varphi_2(t)w\rangle_{L^2(0,T,V')\times L^2(0,T;V)} - \langle (-\Delta)^{\frac{3}{2}}\phi_n, \varphi_2(t)\nabla w\rangle_{L^2(0,T;H)} + \\
&\quad + \langle \nabla f(\phi_n), \varphi_2(t)\nabla w\rangle_{L^2(0,T;H)} - \langle \nabla\theta_n, \nabla w\rangle_{L^2(0,T;H)} = \\
&\quad = \langle h_n, \varphi_2(t)v\rangle_{L^2(0,T;H)},
\end{aligned}$$

and we need to prove that we can pass the limit $n \to \infty$. Since continuous operators are weakly continuous, we can pass the limit in all the linear terms according to (3.3), (3.4) and (3.5). What is left is the already well-known problematic term

$$\langle \nabla f(\phi_n), \varphi_2(t)\nabla w\rangle_{L^2(0,T;H)} = \int_0^T \varphi_2(t)\langle f(\phi_n), w\rangle_V \,\mathrm{d}t,$$

for which we have to show that

$$\int_0^T \varphi_2(t)\langle f(\phi_n), w\rangle_V \,\mathrm{d}t \overset{n\to\infty}{\longrightarrow} \int_0^T \varphi_2(t)\langle f(\phi^*), w\rangle_V \,\mathrm{d}t. \tag{3.10}$$

In Step C of the proof of Theorem 2.11, we presented two ways to pass this limit. Choosing one, we have shown that $y^* = y(u^*)$. Note that for the first way, we need to consider more regular functions instead of (v, w), for example (w_j, w_k).

To show that (u^*, y^*) solves OP, we note that due to $V \Subset H \hookrightarrow V'$, Lemma 2.12 provides us with the compact embedding

$$L^2(0,T;V) \cap H^1(0,T;V') \Subset L^2(0,T;H).$$

Together with (3.9), this means that $y_n \to y^*$ strongly in $L^2(0,T;Y) \subset L^2(Q_T)^2$ and by (3.3), we have $u_n \rightharpoonup u^*$ weakly in $L^2(Q_T)^2$. Hence, we get the estimate (see [9], Satz 2.7 on page 8)

$$J(u^*, y^*) \;\le\; \lim_{n\to\infty} \frac{1}{2}\|y_n - y_d\|^2_{L^2(Q_T)^2} + \liminf_{n\to\infty} \frac{\beta}{2}\|u_n\|^2_{L^2(Q_T)^2} = \liminf_{n\to\infty} J(u_n, y_n) = \overline{J},$$

which means that (u^*, y^*) is optimal. □

3.2 The Linearized Model

Now that we have shown the existence of an optimal control for OP, we want to derive an optimality condition. To do so, we start by proving well-posedness of the linearized Caginalp model. Why we do that is not obvious at this point, roughly speaking, this model serves as a derivative of our constraint. The details will be presented later, for now, we can view this section as a lemma.

To state the linearized model, we have to compute the Fréchet derivative of the nonlinearity

$$f : \mathcal{W}_2 \longrightarrow L^2(Q_T), \qquad \phi \longmapsto \phi^3 - \phi,$$

and we show that the derivative is continuous. Let $h \in \mathcal{W}_2$. Using Hölder and the continuous embedding $V \hookrightarrow L^6(\Omega)$, we compute

$$\begin{aligned} \|(\phi+h)^3 - \phi^3 - 3\phi^2 h\|^2_{L^2(Q_T)} &= \int_0^T \|3\phi h^2 + h^3\|^2_H \,\mathrm{d}t \ \leq \\ &\leq\ 3\int_0^T \|\phi\|^2_{L^6(\Omega)}\|h\|^4_{L^6(\Omega)}\,\mathrm{d}t + \int_0^T \|h\|^6_{L^6(\Omega)}\,\mathrm{d}t \ \leq \\ &\leq\ C\big(\|\phi\|^2_{L^\infty(0,T;V)} + \|h\|^2_{L^\infty(0,T;V)}\big)\|h\|^4_{L^\infty(0,T;V)}. \end{aligned}$$

Hence, we have

$$\|(\phi+h)^3 - \phi^3 - 3\phi^2 h\|_{L^2(Q_T)} = o(\|h\|_{\mathcal{W}_2}).$$

Due to the sum rule, this means that f is F-differentiable on $\mathcal{W}$ with derivative

$$f' : \mathcal{W}_2 \longrightarrow L(\mathcal{W}_2, L^2(Q_T)), \qquad f'(\phi)v = (3\phi^2 - 1)v. \tag{3.11}$$

To prove that f' is sequentially continuous, let $\phi, \epsilon, h \in \mathcal{W}_2$. Using similar arguments, we compute

$$\begin{aligned} \|(f'(\phi+\epsilon) - f'(\phi))h\|^2_{L^2(Q_T)} &= 3\int_0^T \|(2\phi+\epsilon)\epsilon h\|^2_H \,\mathrm{d}t \ \leq \\ &\leq\ \|2\phi+\epsilon\|^2_{L^\infty(0,T;V)}\|\epsilon\|^2_{L^\infty(0,T;V)}\|h\|^2_{L^\infty(0,T;V)}. \end{aligned}$$

This means that for $\|\epsilon\|_{\mathcal{W}_2} \to 0$, we have $f'(\phi+\epsilon) \to f'(\phi)$ in the operator norm on $L(\mathcal{W}_2, L^2(Q_T))$, so f' is continuous.

Linearized transformed Caginalp model

The computation of the F-derivatives of linear operators is straightforward. As before, introducing $z := \theta + \phi$ instead of θ, we transform the system (2.1)-(2.3) such that we only have one time derivative in the heat balance equation. Hence, after passing to the F-derivatives, we state the linearized transformed Caginalp model at the point $(z, \phi) \in \mathcal{W}$ for (η, ξ) as follows.

Linearized Transformed Caginalp Model

$$\partial_t \eta + A\eta - A\xi = \tilde{g} \qquad \text{in } (0,T)\times\Omega, \tag{3.12}$$

$$\partial_t \xi + A^2\xi - A\eta + A\xi + A(f'(\phi)\xi) = \tilde{h} \qquad \text{in } (0,T)\times\Omega, \tag{3.13}$$

with Dirichlet boundary conditions

$$\eta = \xi = A\xi = 0 \qquad \text{on } (0,T)\times\partial\Omega, \tag{3.14}$$

and initial data

$$\eta(0) = \eta_0, \qquad \xi(0) = \xi_0 \qquad \text{in } \Omega. \tag{3.15}$$

As for the nonlinear model TCM in Section 2.3, we want to see how the action of $f'(\phi)$ affects the regularity of ξ. We prove the following lemma.

Lemma 3.2 Let $\phi, \xi \in \mathcal{W}_2$. Then, it holds that $f'(\phi)\xi \in L^2(0,T;V)$. In particular, we make use of the estimate

$$\|f'(\phi)\xi\|^2_V \ \leq\ C\left(\|\phi\|^2_{D(A)} + 1\right)\|\xi\|^2_V. \tag{3.16}$$

Proof. Let $\phi, \xi \in \mathcal{W}_2$. Then, it holds that

- ▷ $f'(\phi)\xi \in L^2(0,T;H)$: Hölder and the continuous embedding $V \hookrightarrow L^6(\Omega)$ yield

$$\begin{aligned}\|f'(\phi)\xi\|^2_{L^2(0,T;H)} &\leq 3\int_0^T \|\phi^2\xi\|_H^2 \,\mathrm{d}t + \int_0^T \|\xi\|_H^2 \,\mathrm{d}t \leq \\ &\leq C\int_0^T \|\phi\|_V^4\|\xi\|_V^2 \,\mathrm{d}t + \int_0^T \|\xi\|_H^2 \,\mathrm{d}t \leq C,\end{aligned}$$

  where we have used that $\phi, \xi \in L^\infty(0,T;V)$.

- ▷ $\nabla(f'(\phi)\xi) \in L^2(0,T;H)$: Using product and chain rule yields

$$\begin{aligned}\|\nabla(f'(\phi)\xi)\|_H^2 &= \|f''(\phi)\nabla\phi\xi + f'(\phi)\nabla\xi\|_H^2 \leq \\ &\leq C\left(\|\phi\nabla\phi\xi\|_H^2 + \|\phi^2\nabla\xi\|_H^2 + \|\xi\|_V^2\right). \end{aligned} \tag{3.17}$$

  Here, we used that $f''(\phi) = 6\phi$, which is shown in the same way as before. Estimating the first two terms of the last expression separately, we get due to Hölder and $V \hookrightarrow L^6(\Omega)$ that

$$\|\phi\nabla\phi\xi\|_H^2 \leq C\|\phi\|_V^2\|\phi\|_{D(A)}^2\|\xi\|_V^2 \leq C\|\phi\|^2_{L^\infty(0,T;V)}\|\phi\|^2_{D(A)}\|\xi\|_V^2.$$

  Moreover, it holds due to the Agmon inequality that

$$\begin{aligned}\|\phi^2\nabla\xi\|_H^2 &\leq C\|\phi\|^4_{L^\infty(\Omega)}\|\xi\|_V^2 \leq C\|\phi\|_V^2\|\phi\|^2_{D(A)}\|\xi\|_V^2 \leq \\ &\leq C\|\phi\|^2_{L^\infty(0,T;V)}\|\phi\|^2_{D(A)}\|\xi\|_V^2.\end{aligned}$$

  Thus, returning to (3.17), we have shown that

$$\|f'(\phi)\xi\|_V^2 \leq C\left(\|\phi\|^2_{D(A)} + 1\right)\|\xi\|_V^2,$$

  which is the equation (3.16) of the second part of the lemma. Now, integrating (3.16) over time yields

$$\|f'(\phi)\xi\|^2_{L^2(0,T;V)} \leq C\|\xi\|^2_{L^\infty(0,T;V)}(\|\phi\|^2_{\mathcal{W}_2} + 1) \leq C.$$

□

Weak formulation of the linearized model

By testing (3.12) with $v \in V$ and (3.13) with $w \in V$, equations (3.12)-(3.13) are associated with the weak formulation

$$\langle\partial_t\eta, v\rangle_{V'\times V} + \langle A^{\frac12}\eta, A^{\frac12}v\rangle_H - \langle A\xi, v\rangle_H = \langle\tilde{g}, v\rangle_{V'\times V}, \tag{3.18}$$

$$\begin{aligned}&\langle\partial_t\xi, w\rangle_{V'\times V} + \langle A^{\frac32}\xi, A^{\frac12}w\rangle_H - \langle A^{\frac12}\eta, A^{\frac12}w\rangle_H + \langle A\xi, w\rangle_H + \langle A^{\frac12}(f'(\phi)\xi), A^{\frac12}w\rangle_H = \\ &\quad = \langle\tilde{h}, w\rangle_{V'\times V},\end{aligned} \tag{3.19}$$

as equations in $\mathcal{D}'(0,T;\mathbb{R})$.

Now we are ready to prove the following energy estimate. Note that we regain the continuous dependence of the solution on the control and the initial data.

Theorem 3.3 Let $(\eta_0, \xi_0) \in Y$ and $(\tilde{g}, \tilde{h}) \in L^2(0,T;X_1)$. Assume there exists a solution

$$(\eta, \xi) \in L^\infty(0,T;Y) \cap L^2(0,T;X_0) \cap H^1(0,T;X_1)$$

of (3.12)-(3.15). Then, it holds that

$$\begin{aligned}&\|(\partial_t\eta, \partial_t\xi)\|^2_{L^2(0,T;X_1)} + \|(\eta,\xi)\|^2_{L^\infty(0,T;Y)} + \|(\eta,\xi)\|^2_{L^2(0,T;X_0)} \leq \\ &\quad \leq C\big(\|(\eta_0,\xi_0)\|_Y^2 + \|(\tilde{g},\tilde{h})\|^2_{L^2(0,T;X_1)}\big).\end{aligned}$$

Proof. If it ain't broken, don't fix it, so we proceed with the same steps as in the proofs of Theorems 2.4 and 2.7 and use analog arguments.

Estimate for (η, ξ) - linearized model

Step 1: Under the assumptions of the theorem, it holds that

$$\|(\eta,\xi)\|^2_{L^\infty(0,T;Y)} + \|(\eta,\xi)\|^2_{L^2(0,T;X_0)} \;\le\; C\big(\|(\eta_0,\xi_0)\|^2_Y + \|(\tilde g,\tilde h)\|^2_{L^2(0,T;X_1)}\big).$$

To show this estimate, we test (3.18) with η and (3.19) with $A\xi$ and use Theorem 2.1 to get

$$\frac12\partial_t\|\eta\|^2_H + \|A^{\frac12}\eta\|^2_H = \langle A\xi,\eta\rangle_H + \langle\tilde g,\eta\rangle_{V'\times V},$$
$$\frac12\partial_t\|A^{\frac12}\xi\|^2_H + \|A^{\frac32}\xi\|^2_H + \|A\xi\|^2_H + \langle A^{\frac12}(f'(\phi)\xi), A^{\frac32}\xi\rangle_H = \langle A^{\frac12}\eta, A^{\frac32}\xi\rangle_H + \langle\tilde h, A\xi\rangle_{V'\times V}.$$

Applying Cauchy-Schwarz and Young to the inner product terms of the equalities and using Lemma 3.2, we have

$$\begin{aligned}
\langle A\xi,\eta\rangle_H &\le \frac12\|A\xi\|^2_H + \frac12\|\eta\|^2_H,\\
\langle A^{\frac12}\eta, A^{\frac32}\xi\rangle_H &\le \frac12\|A^{\frac12}\eta\|^2_H + \frac12\|A^{\frac32}\xi\|^2_H,\\
\langle\tilde g,\eta\rangle_{V'\times V} &\le C\|\tilde g\|^2_{V'} + \frac14\|A^{\frac12}\eta\|^2_H,\\
\langle\tilde h, A\xi\rangle_{V'\times V} &\le C\|\tilde h\|^2_{V'} + \frac18\|A^{\frac32}\xi\|^2_H + C\|A^{\frac12}\xi\|^2_H,\\
\langle A^{\frac12}(f'(\phi)\xi), A^{\frac32}\xi\rangle_H &\le C(\|A^{\frac32}\phi\|^2_H+1)\|A^{\frac12}\xi\|^2_H + \frac18\|A^{\frac32}\xi\|^2_H.
\end{aligned}$$

Hence, adding both equalities, collecting terms, omitting the $A\xi$-term and adapting constants yields

$$\partial_t\|(\eta,\xi)\|^2_Y + \|(\eta,\xi)\|^2_{X_0} \;\le\; C\big(\|\eta\|^2_H + (\|A^{\frac32}\phi\|^2_H+1)\|\xi\|^2_V + \|(\tilde g,\tilde h)\|^2_{X_1}\big). \tag{3.20}$$

Next, since $\phi\in\mathcal W_2$, we define the $L^2(0,T;\mathbb R)$-integrable function

$$m(t) := \max(1\,,\ \|A^{\frac32}\phi\|^2_H + 1),$$

which turns (3.20) into

$$\partial_t\|(\eta,\xi)\|^2_Y + \|(\eta,\xi)\|^2_{X_0} \;\le\; C\big(m(t)\|(\eta,\xi)\|^2_Y + \|(\tilde g,\tilde h)\|^2_{X_1}\big). \tag{3.21}$$

Since m is a positive function, applying the Gronwall Lemma leads to the estimate

$$\|(\eta(t),\xi(t))\|^2_Y \;\le\; C\big(\|(\eta_0,\xi_0)\|^2_Y + \|(\tilde g,\tilde h)\|^2_{L^2(0,T;X_1)}\big).$$

Hence, integrating (3.21) over time, we get for almost every $t\in(0,T)$

$$\begin{aligned}
&\|(\eta(t),\xi(t))\|^2_Y + \|(\eta,\xi)\|^2_{L^2(0,t;X_0)} \;\le\\
&\quad\le\; C\Big(\int_0^t m(s)\|(\eta(s),\xi(s))\|^2_Y\,\mathrm ds + \|(\tilde g,\tilde h)\|^2_{L^2(0,t;X_1)}\Big) + \|(\eta_0,\xi_0)\|^2_Y\\
&\quad\le\; C\big(\|(\eta_0,\xi_0)\|^2_Y + \|(\tilde g,\tilde h)\|^2_{L^2(0,T;X_1)}\big).
\end{aligned}$$

The right hand side is independent of t, so we have

$$\|(\eta,\xi)\|^2_{L^\infty(0,T;Y)} + \|(\eta,\xi)\|^2_{L^2(0,t;X_0)} \;\le\; 2C\big(\|(\eta_0,\xi_0)\|^2_Y + \|(\tilde g,\tilde h)\|^2_{L^2(0,T;X_1)}\big).$$

Step 1 follows by passing $t\to T$.

Estimate for $(\partial_t\eta, \partial_t\xi)$ - linearized model

Step 2: It holds that

$$\|(\partial_t\eta,\partial_t\xi)\|^2_{L^2(0,T;X_1)} \;\le\; C\big(\|(\eta_0,\xi_0)\|^2_Y + \|(\tilde g,\tilde h)\|^2_{L^2(0,T;X_1)}\big).$$

We solve (3.12) for $\partial_t\eta$ and (3.13) for $\partial_t\xi$ and apply the V'-norm to both equations, which leads to

$$\begin{aligned}\|\partial_t\eta\|_{V'}^2 &\le C\left(\|A\eta\|_{V'}^2 + \|A\xi\|_{V'}^2 + \|g\|_{V'}^2\right),\\ \|\partial_t\xi\|_{V'}^2 &\le C\left(\|A^2\xi\|_{V'}^2 + \|A\eta\|_{V'}^2 + \|A\xi\|_{V'}^2 + \|A(f'(\phi)\xi)\|_{V'}^2 + \|h\|_{V'}^2\right).\end{aligned}$$

The difference to Step 2 of the proof of Theorem 2.7 is the $\|A(f'(\phi)\xi)\|_{V'}^2$-term. We can easily deal with this term applying Lemma 3.2, yielding

$$\int_0^T \|A(f'(\phi)\xi)\|_{V'}^2\,\mathrm{d}t \le C\int_0^T \|f'(\phi)\xi\|_V^2\,\mathrm{d}t \le C\|\xi\|_{L^\infty(0,T;V)}^2(\|\phi\|_{\mathcal{W}_2}^2+1) \le$$
$$\le C\|\xi\|_{L^\infty(0,T;V)}^2.$$

Then, we can finish with exactly the same reasoning as before. which completes the proof of Theorem 3.3. □

Theorem 3.3 provides us with an energy estimate that we use for the Galerkin approximation to prove well-posedness of the linearized model.

Theorem 3.4 Let $(\eta_0,\xi_0) \in Y$ and $(\tilde{g},\tilde{h}) \in L^2(0,T;X_1)$. Then, (3.12)-(3.15) possess a unique solution (η,ξ) such that

$$(\eta,\xi) \in L^\infty(0,T;Y) \cap L^2(0,T;X_0) \cap H^1(0,T;X_1).$$

Proof. The proof is very similar to the linear case of Theorem 2.5, so we proceed in the same steps but online describe which changes are to be made in each step.

Approximate solutions - linearized model

Step A: Construction of solutions (η_m,ξ_m) on finite-dimensional subspaces W_m.

We start with the ansatz

$$\eta_m = \sum_{k=1}^m c_k^m(t)w_k \qquad \text{and} \qquad \xi_m = \sum_{k=1}^m d_k^m(t)w_k,$$

and consider the approximated problem

$$\partial_t\langle\eta_m,v\rangle_H + \langle A^{\frac12}\eta_m, A^{\frac12}v\rangle_H - \langle A\phi_m,v\rangle_H = \langle\tilde{g},v\rangle_{V'\times V}, \tag{3.22}$$

$$\begin{aligned}\partial_t\langle\xi_m,w\rangle_H + \langle A^{\frac32}\xi_m, A^{\frac12}w\rangle_H - \langle A^{\frac12}\eta_m, A^{\frac12}w\rangle_H+&\\ +\langle A\xi_m,w\rangle_H + \langle A^{\frac12}(f'(\phi)\xi_m), A^{\frac12}w\rangle_H = \langle\tilde{h},w\rangle_{V'\times V},&\end{aligned} \tag{3.23}$$

for all $(v,w) \in \mathrm{Span}(w_1,...,w_m)^2 =: W_m$ as equations in $\mathcal{D}'(0,T;\mathbb{R})$ and with projected initial conditions. Testing both equations of the approximated problem with w_j, $j = 1,...,m$, the only problematic terms in the resulting system of ODEs are given by

$$\langle A^{\frac12}(f'(\phi)\xi_m), A^{\frac12}w_j\rangle_H = \langle f'(\phi)\xi_m, Aw_j\rangle_H = \sum_{k=1}^m \lambda_j d_k^m(t)\langle f'(\phi)w_k,w_j\rangle_H. \tag{3.24}$$

We know that $\phi \in C([0,T];V)$ and we claim that

$$f'(\phi) = 3\phi^2 - 1 \in C([0,T];H). \tag{3.25}$$

To show this, let $t,s \in [0,T]$ such that $t+s \in [0,T]$. Then, due to Hölder and the continuous embeddings $V \hookrightarrow L^6(\Omega) \hookrightarrow L^3(\Omega)$, we have

$$\|\phi^2(t+s)-\phi^2(t)\|_H \le \|\phi(t+s)-\phi(t)\|_{L^3(\Omega)}\|\phi(t+s)+\phi(t)\|_{L^6(\Omega)} \le$$
$$\le \|\phi(t+s)-\phi(t)\|_V\|\phi(t+s)+\phi(t)\|_V$$

As $s \to 0$, the first term of the last expression converges to zero, the second to $\|2\phi(t)\|_V$, which shows sequential continuity of $\phi(t)^2$ on $[0,T]$. Since $[0,T]$ is compact, we have $\phi(t)^2 \in C([0,T];H)$, hence $f'(\phi) \in C([0,T];H)$.

Regarding (3.24) in the presence of (3.25), we see immediately due to $w_k \in L^\infty(\Omega)$ that

$$\langle f'(\phi)w_k, w_j\rangle_H \in C([0,T];\mathbb{R}).$$

Thus, we can apply the theorem of Picard-Lindelöf to the homogeneous system of ODEs, which gives us a unique solution globally on $(0,T)$ because of the energy estimate of Theorem 3.3. Then, the variation of parameters serves to get a unique solution $(c^m(t), d^m(t)) \in AC([0,T];\mathbb{R}^{2m}) \cap H^1(0,T;\mathbb{R}^{2m})$ of the inhomogeneous system of ODEs and hence to get a unique solution $(\eta_m, \xi_m) \in AC([0,T], W_m)$, $(\partial_t\eta_m, \xi_m) \in L^2(0,T;W_m)$ of the approximated problem.

Convergence of approximate solutions - linearized model

Step B: Convergence of the sequence of solutions $((\eta_m, \xi_m))_{m\in\mathbb{N}}$.

Following the same reasoning as in Step B of Theorem 2.5, we get

$$\begin{aligned}
(\eta_m, \xi_m) &\rightharpoonup (\eta, \xi) && \text{weakly in } L^2(0,T;X_0)\\
(\partial_t\eta_m, \partial_t\xi_m) &\rightharpoonup (\partial_t\eta, \partial_t\xi) && \text{weakly in } L^2(0,T;X_1),\\
(\eta_m, \xi_m) &\overset{*}{\rightharpoonup} (\eta, \xi) && \text{weak star in } L^\infty(0,T;Y), && (3.26)\\
\xi_m &\to \xi && \text{strongly in } L^2(0,T;V), && (3.27)
\end{aligned}$$

where we have used the results of Step B of Theorem 2.11 for (3.27). We have to show that (η, ξ) solves (3.12)-(3.15).

Existence of a weak solution - linearized model

Step C: Existence of a weak solution (η, ξ).

Each of the (η_m, ξ_m) solves the approximated problem (3.22)-(3.23) with projected initial conditions and we have to show that we can pass the limit $m \to \infty$. This works exactly as in Theorem 2.5 except for the term

$$\int_0^T \varphi(t)\langle A^{\frac{1}{2}}(f'(\phi)\xi_m), A^{\frac{1}{2}}w_k\rangle_H \,\mathrm{d}t,$$

for which we have to show that

$$\int_0^T \varphi_k(t)\langle f'(\phi)\xi_m, w_k\rangle_V \,\mathrm{d}t \overset{m\to\infty}{\longrightarrow} \int_0^T \varphi(t)\langle f'(\phi)\xi, w_k\rangle_V \,\mathrm{d}t. \qquad (3.28)$$

Using Cauchy-Schwarz, we can estimate

$$\begin{aligned}
\left|\int_0^T \varphi(t)\langle f'(\phi)(\xi_m - \xi), w_k\rangle_V \,\mathrm{d}t\right|^2 &\le C\left|\int_0^T \varphi(t)\|f'(\phi)(\xi_m - \xi)\|_V\|w_k\|_V \,\mathrm{d}t\right|^2 \le\\
&\le C\|\varphi(t)w_k\|^2_{L^2(0,T;V)}\int_0^T \|f'(\phi)(\xi_m - \xi)\|_V^2 \,\mathrm{d}t = C\int_0^T \|f'(\phi)(\xi_m - \xi)\|_V^2 \,\mathrm{d}t.
\end{aligned}$$

According to Lemma 3.2, we get

$$\int_0^T \|f'(\phi)(\xi_m - \xi)\|_V^2 \,\mathrm{d}t \le C\int_0^T \|\phi\|^2_{D(A)}\|\xi_m - \xi\|_V^2 \,\mathrm{d}t + \|\xi_m - \xi\|^2_{L^2(0,T;V)}. \qquad (3.29)$$

The first term on the right hand side can be processed by

$$\begin{aligned}
&\int_0^T \|\phi\|^2_{D(A)}\|\xi_m - \xi\|_V^2 \,\mathrm{d}t \le\\
&\quad\le C\|\phi\|_{L^\infty(0,T;V)}\|\xi_m - \xi\|_{L^\infty(0,T;V)}\int_0^T \|\phi\|_{D(A^{\frac{3}{2}})}\|\xi_m - \xi\|_V \,\mathrm{d}t \le\\
&\quad\le C\|\phi\|_{L^2(0,T;D(A^{\frac{3}{2}}))}\|\xi_m - \xi\|_{L^2(0,T;V)},
\end{aligned}$$

where we have used the boundedness of (ξ_m) in $L^\infty(0,T;V)$ due to (3.26). Returning to (3.29), we have shown that

$$\int_0^T \|f'(\phi)(\xi_m - \xi)\|_V^2 \,\mathrm{d}t \;\le\; C(\|\xi_m - \xi\|_{L^2(0,T;V)} + \|\xi_m - \xi\|^2_{L^2(0,T;V)}).$$

Then, (3.28) follows from the strong convergence (3.27).

Step D: Uniqueness of the weak solution (η, ξ).

Uniqueness of the weak solution - linearized model

This is exactly the same argument as in Theorem 2.5, using the fact that the right hand side of the energy estimate of Theorem 3.3 only depends on the initial data and the control. □

3.3 First Order Necessary Optimality Condition

Based on the ideas of [13], in this section we want to derive a first order optimality condition for the problem OP. Knowing that it is sufficient to search for states y in $\mathcal{W}$, we recall the slightly changed OP as the following minimization problem.

Optimal Control Problem (OP)

$$\min_{(u,y)\in L^2(Q_T)^2\times\mathcal{W}} J(u,y) := \frac{1}{2}\|y - y_d\|^2_{L^2(Q_T)^2} + \frac{\beta}{2}\|u\|^2_{L^2(Q_T)^2}, \tag{3.30}$$

subject to Caginalp's model

$$\begin{aligned} \partial_t(\theta+\phi) - \Delta\theta &= g \qquad \text{in } Q_T,\\ \partial_t\phi - \Delta\mu &= h \qquad \text{in } Q_T,\\ \mu = -\Delta\phi + F'(\phi) - \theta & \qquad\quad \text{in } Q_T, \end{aligned}$$

with Dirichlet boundary conditions

$$\theta = \phi = \Delta\phi = 0 \qquad \text{on } \Sigma_T,$$

and initial data

$$\theta(0) = \theta_0, \quad \phi(0) = \phi_0 \qquad \text{in } \Omega.$$

We consider again the control $u = (g,h) \in L^2(Q_T)^2$, state $y = (\theta,\phi) \in \mathcal{W}$ and $y_0 = (\theta_0, \phi_0)$ and view Caginalp's model as an equality constraint,

Equality constraint $e(u,y)$

$$e(u,y) : L^2(Q_T)^2 \times \mathcal{W} \longrightarrow L^2(0,T;X_1) \times Y$$
$$(u,y) \longmapsto \begin{pmatrix} \begin{pmatrix} \partial_t - \Delta & \partial_t \\ \Delta & \partial_t + \Delta^2 \end{pmatrix} y - \begin{pmatrix} 0 \\ \Delta f(\phi) \end{pmatrix} - u \\ y(0) - y_0 \end{pmatrix} =: \begin{pmatrix} e_1(u,y) \\ e_2(u,y) \end{pmatrix}.$$

Note that we can evaluate e at general $y \in \mathcal{W}$ and that we have $e(u,y) = 0$ if and only if $y = y(u)$. Furthermore, using the solution operator $y(u)$, we can define the reduced cost functional

$$\hat{J}(u) := J(u, y(u)).$$

The first order optimality condition will be stated for the derivative of the reduced cost functional, so we need to find a representation for $\hat{J}'(u) \in (L^2(Q_T)^2)' = L^2(Q_T)^2$. Let $v \in L^2(Q_T)^2$, then we formally differentiate

$$\begin{aligned} \langle \hat{J}'(u), v\rangle_{L^2(Q_T)^2} &= \langle \partial_u J(u,y(u))\,,\, v\rangle_{L^2(Q_T)^2} + \langle \partial_y J(u,y(u))\,,\, y'(u)v\rangle_{\mathcal{W}'\times\mathcal{W}} \\ &= \langle \partial_u J(u,y(u))\,,\, v\rangle_{L^2(Q_T)^2} + \langle y'(u)^*\partial_y J(u,y(u))\,,\, v\rangle_{L^2(Q_T)^2}, \end{aligned}$$

where * denotes the adjoint operator. Consequently, it holds that

$$\hat{J}'(u) = y'(u)^* \partial_y J(u, y(u)) + \partial_u J(u, y(u)). \tag{3.31}$$

Since J is continuously F-differentiable, to prove that this expression is well-defined we have to show that the solution operator $S(y_0)$ is continuously F-differentiable.

Differentiability of the solution operator

We prove that $y'(u)$ as a continuous F-derivative exists for each $u \in L^2(Q_T)^2$. To do so, we use the following Banach space version of the implicit function theorem.

> **Lemma 3.5** Implicit function theorem [13] Let $\mathcal{X}, \mathcal{Y}$ and $\mathcal{Z}$ be Banach spaces and let $F : G \to Z$ be a continuously F-differentiable map from an open set $G \subset X \times Y$ to Z. Let $(\bar{x}, \bar{y}) \in G$ such that $F(\bar{x}, \bar{y}) = 0$ and that $\partial_y F(\bar{x}, \bar{y}) \in L(Y, Z)$ has a bounded inverse.
>
> Then, there exists an open neighborhood $U_X(\bar{x}) \times U_Y(\bar{y}) \subset G$ of $(\bar{x}, \bar{y})$ and a unique continuous function $w : U_X(\bar{x}) \to Y$ such that
>
> - $w(\bar{x}) = \bar{y}$,
> - for all $x \in U_X(\bar{x})$ there exists exactly one $y \in U_Y(\bar{y})$ with $F(x, y) = 0$, namely $y = w(x)$.
>
> Moreover, the mapping $w : U_X(\bar{x}) \to Y$ is continuously F-differentiable with derivative
>
> $$w'(x) = -\partial_y F(x, w(x))^{-1} \partial_x F(x, w(x)).$$

Let $\bar{u} \in L^2(Q_T)^2$. Adapting the notations of Lemma 3.5 to our situation, we have

$$\begin{aligned} \mathcal{X} &= L^2(Q_T)^2, \\ \mathcal{Y} &= \mathcal{W}, \\ \mathcal{Z} &= L^2(0, T; X_1) \times Y, \\ G &= L^2(Q_T)^2 \times \mathcal{W}, \\ F &= e, \\ (\bar{x}, \bar{y}) &= (\bar{u}, y(\bar{u})), \\ w(\bar{u}) &= y(\bar{u}). \end{aligned}$$

Thus, we see that to show the F-differentiability of the solution operator at $\bar{u}$, it remains to show that

- e is continuously F-differentiable,
- $\partial_y e(\bar{u}, y(\bar{u})) \in L(\mathcal{W}, L^2(0, T; X_1) \times Y)$ has a bounded inverse.

Then, Lemma 3.5 also provides us with a representation of $y'(\bar{u})$ given by

$$y'(\bar{u}) = -\partial_y e(\bar{u}, y(\bar{u}))^{-1} \partial_u e(\bar{u}, y(\bar{u})). \tag{3.32}$$

To prove the two remainders, we compute the partial derivatives $\partial_y e$ and $\partial_u e$ first. Let $h = (\eta, \xi) \in \mathcal{W}$, then we have

$$e(u, y+h) - e(u, y) = \begin{pmatrix} \begin{pmatrix} \partial_t - \Delta & \partial_t \\ \Delta & \partial_t + \Delta^2 \end{pmatrix} h - \begin{pmatrix} 0 \\ \Delta(f(\phi + \xi) - f(\phi)) \end{pmatrix} \\ h(0) \end{pmatrix}.$$

So we set

$$\partial_y e(u, y) h := \begin{pmatrix} \begin{pmatrix} \partial_t - \Delta & \partial_t \\ \Delta & \partial_t + \Delta^2 - \Delta f'(\phi) \end{pmatrix} h \\ h(0) \end{pmatrix}$$

and we see immediately due to (3.11) that this is the F-derivative of e at (u, y) with respect to y. For the concrete choice $(\bar{u}, y(\bar{u}))$, we observe that $\partial_y e(\bar{u}, y(\bar{u}))h$ represents exactly the linearized model of Section 3.2. Thus, $\partial_y e(\bar{u}, y(\bar{u}))$ has a bounded inverse if and only if the linearized Caginalp model admits a unique solution that is continuously dependent on the control and the initial data. This was shown in Theorems 3.3 and 3.4, so the second remainder holds.

Next, we show the sequential continuity of

$$\begin{aligned} \partial_y e : L^2(Q_T)^2 \times \mathcal{W} &\longrightarrow L(\mathcal{W}, L^2(0,T;X_1) \times Y), \\ (u,y) &\longmapsto \partial_y e(u,y). \end{aligned}$$

Let $\delta \in L^2(Q_T)^2$, $\epsilon = (\epsilon_1, \epsilon_2) \in \mathcal{W}$ and $h = (\eta, \xi) \in \mathcal{W}$. Then

$$\partial_y e(u+\delta, y+\epsilon)h = \begin{pmatrix} \begin{pmatrix} \partial_t - \Delta & \partial_t \\ \Delta & \partial_t + \Delta^2 - \Delta f'(\phi + \epsilon_2) \end{pmatrix} h \\ h(0) \end{pmatrix} \in L^2(0,T;X_1) \times Y,$$

and we have to prove that

$$\sup_{\|h\|_{\mathcal{W}}=1} \|(\partial_y e(u+\delta, y+\epsilon) - \partial_y e(u,y))h\|_{L^2(0,T;X_1)\times Y} \to 0,$$

as $\|(\delta,\epsilon)\|_{L^2(Q_T)^2\times\mathcal{W}} \to 0$. Since all but one terms are independent of (δ, ϵ), it suffices to show that

$$\sup_{\|\xi\|_{\mathcal{W}_2}=1} \|\Delta(f'(\phi+\epsilon_2) - f'(\phi))\xi\|_{L^2(0,T;V')} \to 0, \tag{3.33}$$

as $\|\epsilon_2\|_{\mathcal{W}_2} \to 0$. Due to the equivalence of norms $\|\Delta\,\cdot\,\|_{V'} \equiv \|\cdot\|_V$, we get the estimate

$$\begin{aligned} \|\Delta\,\cdot\,\|^2_{L^2(0,T;V')} &= \int_0^T \|\Delta\,\cdot\,\|^2_{V'}\,\mathrm{d}t \;\le\; C\int_0^T \|\cdot\|^2_V\,\mathrm{d}t = C\int_0^T \|\nabla\,\cdot\,\|^2_H\,\mathrm{d}t = \\ &= C\|\nabla\,\cdot\,\|^2_{L^2(Q_T)}. \end{aligned}$$

Thus, instead of (3.33), we can also show that

$$\sup_{\|\xi\|_{\mathcal{W}_2}=1} \|\nabla(f'(\phi+\epsilon_2) - f'(\phi))\xi\|_{L^2(Q_T)} \to 0, \tag{3.34}$$

as $\|\epsilon_2\|_{\mathcal{W}_2} \to 0$. Using product and chain rule and noting that $f'(\phi) = 3\phi^2 - 1$ and $f''(\phi) = 6\phi$, we compute

$$\begin{aligned} &\|\nabla(f'(\phi+\epsilon_2) - f'(\phi))\xi\|^2_{L^2(Q_T)} = \\ &\quad = \int_0^T \|\big(f''(\phi+\epsilon_2)(\nabla\phi + \nabla\epsilon_2) - f''(\phi)\nabla\phi\big)\xi + \big(f'(\phi+\epsilon_2) - f'(\phi)\big)\nabla\xi\|^2_H\,\mathrm{d}t = \\ &\quad = \int_0^T \|6\big(\phi\nabla\epsilon_2 + \epsilon_2\nabla\phi + \epsilon_2\nabla\epsilon_2\big)\xi + 3\big(2\phi\epsilon + \epsilon^2\big)\nabla\xi\|_H\,\mathrm{d}t. \end{aligned}$$

Applying the triangle inequality, we can deal with five terms separately. Due to Hölder and the continuous embedding $V \hookrightarrow L^6(\Omega)$, for the first term we have

$$\begin{aligned} \int_0^T \|\phi\nabla\epsilon_2\xi\|^2_H\,\mathrm{d}t \;&\le\; C\int_0^T \|\phi\|^2_V\|\epsilon_2\|^2_{D(A)}\|\xi\|^2_V\,\mathrm{d}t \;\le \\ &\le\; C\|\phi\|^2_{L^\infty(0,T;V)}\|\epsilon_2\|^2_{L^2(0,T;D(A^{\frac{3}{2}}))}\|\xi\|^2_{L^\infty(0,T;V)} \;\le\; C\|\phi\|^2_{\mathcal{W}_2}\|\epsilon_2\|^2_{\mathcal{W}_2}\|\xi\|^2_{\mathcal{W}_2}. \end{aligned}$$

For the other four terms, we get analog estimates in the same way. Thus, we have shown that (3.34) holds, so $\partial_y e$ is sequentially continuous and hence continuous.

The computation of $\partial_u e$ is straightforward. At each point $(u,y) \in L^2(Q_T)^2 \times \mathcal{W}$, we have

$$\partial_u e(u,y) \in L(L^2(Q_T))^2, L^2(0,T;X_1) \times Y), \qquad \partial_u e(u,y)(\tilde{g}, \tilde{h}) = ((-\tilde{g}, -\tilde{h}), 0), \tag{3.35}$$

and the mapping

$$\begin{aligned} \partial_u e : L^2(Q_T)^2 \times \mathcal{W} &\longrightarrow L(L^2(Q_T)^2, L^2(0,T;X_1) \times Y), \\ (u,y) &\longmapsto \partial_u e(u,y), \end{aligned}$$

is constant, so obviously continuous.

Hence, we know that e is continuously partially F-differentiable and thus continuously F-differentiable, so the first remainder holds.

The adjoint equation Returning to (3.31), we know that we have found a well-defined expression for $\hat{J}'(u)$, which we recall as

$$\hat{J}'(u) = y'(u)^* \partial_y J(u, y(u)) + \partial_u J(u, y(u)).$$

We want to improve this expression with the result (3.32) provided by the implicit function theorem, given by

$$y'(\bar{u}) = -\partial_y e(\bar{u}, y(\bar{u}))^{-1} \partial_u e(\bar{u}, y(\bar{u})).$$

Obviously, we can replace the first summand,

$$y'(u)^* \partial_y J(u, y(u)) = -\partial_u e(u, y(u))^* \partial_y e(u, y(u))^{-*} \partial_y J(u, y(u)).$$

The right hand side includes an inverse, which we handle by introducing a new variable (p, q), as depicted in the following theorem.

Theorem 3.6 About the adjoint equation Let $u \in L^2(Q_T)^2$ be a control and $y(u) \in \mathcal{W}$ the correspoding state. Then, with J and e defined as above, we have

$$y'(u)^* \partial_y J(u, y(u)) = \partial_u e(u, y(u))^* (p, q)(u) \qquad \text{in } \left(L^2(Q_T)^2\right)' = L^2(Q_T)^2,$$

where the *adjoint state* $(p, q) = (p, q)(u) \in L^2(0, T; V \times V) \times Y'$ solves the *adjoint equation*

$$\partial_y e(u, y(u))^* (p, q) = -\partial_y J(u, y(u)) \qquad \text{in } \mathcal{W}'.$$

Using Theorem 3.6 to reforumulate (3.31), the derivative $\hat{J}'(u)$ is given by

$$\hat{J}'(u) = \partial_u e(u, y(u))^* (p, q) + \partial_u J(u, y(u)). \tag{3.36}$$

Hence, we shifted the problem of finding a suitable expression for $\hat{J}'(u)$ to finding a concrete expression for the adjoint eqaution. Let $(u, y = y(u)) \in L^2(Q_T)^2 \times \mathcal{W}$ be a control and corresponding solution for fixed y_0. Investigating the adjoint equation of Theorem 3.6, we first note that since $\partial_y e(u, y)$ has a bounded inverse, there exists a unique adjoint state

$$(p, q) = (p, q)(u) = ((p_1, p_2), (q_1, q_2)) \in L^2(0, T; V \times V) \times Y'.$$

Let $h = (\eta, \xi) \in \mathcal{W}$ and we assume at this point that $p \in \mathcal{W}$, which will be justified later. Then, by definition of the adjoint operator, the adjoint equation satisfies

$$\begin{aligned} 0 &= \langle \partial_y e(u, y)^* (p, q) \,,\, h \rangle_{\mathcal{W}' \times \mathcal{W}} + \langle \partial_y J(u, y) \,,\, h \rangle_{\mathcal{W}' \times \mathcal{W}} = \\ &= \langle \partial_y e(u, y) h \,,\, (p, q) \rangle_{(L^2(0,T;X_1) \times Y) \times (L^2(0,T;V \times V) \times Y')} + \langle \partial_y J(u, y) \,,\, h \rangle_{\mathcal{W}' \times \mathcal{W}}. \end{aligned} \tag{3.37}$$

The first term of the last expression in (3.37) is given by

$$\left\langle \left(\begin{pmatrix} \partial_t - \Delta & \partial_t \\ \Delta & \partial_t + \Delta^2 - \Delta f'(\phi) \end{pmatrix} h \atop h(0) \right) , \begin{pmatrix} p \\ q \end{pmatrix} \right\rangle_{(L^2(0,T;X_1) \times Y) \times (L^2(0,T;V \times V) \times Y')} =$$

$$\begin{aligned} &= \int_0^T \langle \partial_t \eta, p_1 \rangle_{V' \times V} + \langle \nabla \eta, \nabla p_1 \rangle_H + \langle \partial_t \xi, p_1 \rangle_{V' \times V} - \langle \nabla \eta, \nabla p_2 \rangle_H + \langle \partial_t \xi, p_2 \rangle_{V' \times V} + \\ &\quad + \langle \Delta \eta, \Delta p_2 \rangle_H + \langle \xi, f'(\phi) \Delta p_2 \rangle_H \,\mathrm{d}t + \langle h(0), q \rangle_{Y \times Y'}. \end{aligned}$$

Integrating the ∂_t-terms by parts yields

$$\left\langle \left(\begin{pmatrix} \partial_t - \Delta & \partial_t \\ \Delta & \partial_t + \Delta^2 - \Delta f'(\phi) \end{pmatrix} h \\ h(0) \right), \begin{pmatrix} p \\ q \end{pmatrix} \right\rangle_{(L^2(0,T;X_1)\times Y)\times(L^2(0,T;V\times V)\times Y')} =$$

$$= \left\langle \begin{pmatrix} -\partial_t - \Delta & \Delta \\ -\partial_t & -\partial_t + \Delta^2 - f'(\phi)\Delta \end{pmatrix} p \, , \, h \right\rangle_{L^2(0,T;X_1)\times L^2(0,T;V\times V)} +$$
$$+ \left\langle \begin{pmatrix} p_1(T) \\ p_2(T) \end{pmatrix}, \begin{pmatrix} \eta(T) + \xi(T) \\ \xi(T) \end{pmatrix} \right\rangle_{H\times H} - \left\langle \begin{pmatrix} p_1(0) \\ p_2(0) \end{pmatrix}, \begin{pmatrix} \eta(0) + \xi(0) \\ \xi(0) \end{pmatrix} \right\rangle_{H\times H} +$$
$$+ \langle h(0), q\rangle_{Y\times Y'}. \tag{3.38}$$

The second term of (3.37) is

$$\langle \partial_y J(u,y) \, , \, h\rangle_{\mathcal{W}'\times\mathcal{W}} = \langle y - y_d, h\rangle_{L^2(Q_T)^2}, \tag{3.39}$$

where we see that $\partial_y J(u,y) \in L^2(0,T;X_1)$. Combining (3.38) and (3.39) and since $\mathcal{W}$ is dense in $L^2(0,T;V\times V)$, we conclude on the one hand that q satisfies

$$\langle h(0), q\rangle_{Y\times Y'} = \left\langle \begin{pmatrix} p_1(0) \\ p_2(0) \end{pmatrix}, \begin{pmatrix} \eta(0) + \xi(0) \\ \xi(0) \end{pmatrix} \right\rangle_{H\times H},$$

and on the other that p is the solution of the following PDE.

Adjoint Equation (AE)

$$\begin{aligned} -\partial_t p_1 - \Delta p_1 + \Delta p_2 &= -(\theta - \theta_d) && \text{in } L^2(0,T;V'), \\ -\partial_t(p_1 + p_2) + \Delta^2 p_2 - f'(\phi)\Delta p_2 &= -(\phi - \phi_d) && \text{in } L^2(0,T;V'), \end{aligned}$$

with Dirichlet boundary conditions

$$p_1 = p_2 = \Delta p_2 = 0 \qquad \text{on } \Sigma_T,$$

and initial data

$$p_1(T) = 0, \quad p_2(T) = 0 \qquad \text{in } \Omega.$$

Now, we can also see that the assumption $p \in \mathcal{W}$ was reasonable, since we can prove analogously as various times before that solutions of AE are found in $\mathcal{W}$.

Necessary first order optimality condition

In the adjoint equation AE, we have found a representation of $\hat{J}'(u)$ with which we are now able to state a necessary first order optimality condition concerning the optimal control problem OP.

Theorem 3.7 Let $(\bar{u}, y(\bar{u})) \in L^2(Q_T)^2 \times \mathcal{W}$ be an optimal solution of OP. Then, it holds that

$$\hat{J}'(\bar{u}) = 0 \in L^2(Q_T)^2,$$

where we can compute $\hat{J}'(\bar{u})$ as follows.

- ▷ Compute the first component of the adjoint state $\bar{p}$ by solving the adjoint equation AE.
- ▷ Compute $\hat{J}'(\bar{u})$ via

$$\hat{J}'(\bar{u}) = -\bar{p} + \beta\bar{u}.$$

Proof. Since $\bar{u}$ is an optimal solution of the reduced cost functional $\hat{J}$, it holds for the directonal derivative in direction $0 \neq h \in L^2(Q_T)^2$ that

$$0 \leq \lim_{t \to 0^+} \frac{\hat{J}(\bar{u} + th) - \hat{J}(\bar{u})}{t} = \hat{J}'(\bar{u})h.$$

The same holds for $-h$ and we conclude that $\hat{J}(\bar{u}) = 0$. Recalling (3.35), the second part of the theorem follows by plugging in

$$\begin{aligned} \partial_u e(\bar{u}, y(\bar{u}))^*(p, q) &= -\bar{p}, \\ \partial_u J(\bar{u}, y(\bar{u})) &= \beta \bar{u}, \end{aligned}$$

into (3.36). □

Conclusions

Recapitulating what was already mentioned in the introduction, from a mathematical point of view, we were faced with two main difficulties in the analysis of Caginalp's phase field model. Firstly, the Bilaplacian Δ^2 sets a system of fourth order and secondly, we have a nonlinear term $F'(\phi) = \phi^3 - \phi$.

It turned out that finding suitable function spaces for the Bilaplacian can be adapted from well-known rules of thumb for second order equations. Considering the generic problem

$$-\Delta\phi = g \text{ in } \Omega, \qquad \partial_n \phi = 0 \text{ on } \partial\Omega,$$

we can say

$$\text{if } g \in H^m(\Omega),\ m \in \mathbb{Z}, \text{ then } \phi \in H^{m+2}(\Omega).$$

Adapting to our situation, the generic problem reads

$$(-\Delta)^2\phi = g \text{ in } \Omega, \qquad \partial_n\phi = \partial_n\Delta\phi = 0 \text{ on } \partial\Omega,$$

and we found out that we can simply generalize to

$$\text{if } g \in H^m(\Omega),\ m \in \mathbb{Z}, \text{ then } \phi \in H^{m+4}(\Omega).$$

This is reflected in the fact that within our investigation of the linear model in Section 2.2, we could show well-posedness for right hand sides

$$\begin{aligned} g \in D(A^{-1}), &\qquad \text{and then} \qquad \phi \in D(A), \\ g \in V', &\qquad \text{and then} \qquad \phi \in D(A^{\frac{3}{2}}). \end{aligned}$$

Results from functional analysis provided us with spaces in between, H respectively V, onto which we had continuity in time. This allowed us to consider initial conditions $\phi_0 \in H$ respectively $\phi_0 \in V$, hence suitable function spaces to deal with the Bilaplacian were found in the triplets

$$\begin{aligned} D(A) \subset H \subset D(A^{-1}), \\ D(A^{\frac{3}{2}}) \subset V \subset V'. \end{aligned}$$

We had to discard the first triplet, since it indeed serves for the Bilaplacian, but does not handle the nonlinearity. Amongst many other reasons, the fact that $V \hookrightarrow L^6(\Omega)$ continuously is very illustrative to point out the necessity of the second triplet. Regarding the structure of F', we see a third order polynomial, and for the highest order term, it holds that

$$\|\phi^3\|_H^2 = \|\phi\|_{L^6(\Omega)}^6 \ \leq\ C\|\phi\|_V^6.$$

Hence, we can treat the H-norm of $F'(\phi)$ with the V-norm of ϕ, but not with the H-norm, motivating the middle space of the second triplet. Apart from that, we noticed that adding the nonlinearity really made life harder compared to the linear models. Considering the energy estimate for the nonlinear model of Theorem 2.10, not only estimating the nonlinear term in the proof was much more elaborate, but we also lost the continuous dependence on the data. As a consequence, the uniqueness of the solution was not straightforward. To show existence in the Galerkin approximation for the linear cases, we relied on the fact that weak convergence is exactly what we need to pass the limit from the approximated problem to the full problem. Let $\varphi \in C_c^\infty(0, T; \mathbb{R})$, $w \in H$, then terms of the form

$$\int_0^T \varphi(t)\langle\,\cdot\,, w\rangle_H \,\mathrm{d}t$$

define functionals in $L^2(0,T;H) \subset L^2(0,T;V')$. Hence, if $(z_m)_{m\in\mathbb{N}}$ converges weakly to z in $L^2(0,T;V)$, we know that

$$\int_0^T \varphi(t)\langle z_m, w\rangle_H \,\mathrm{d}t \overset{m\to\infty}{\longrightarrow} \int_0^T \varphi(t)\langle z, w\rangle_H \,\mathrm{d}t.$$

Compared to that, the nonlinearity generates terms of the form

$$\int_0^T \varphi(t)\langle F'(\,\cdot\,), w\rangle_H \,\mathrm{d}t,$$

which are not functionals and forced us to find other arguments to pass the limit.

After finishing the analysis of Caginalp's model, we had at hand a well-defined solution operator $u \to y(u)$. Showing its differentiability was the most difficult part in the remainder. We want to emphasize the role of the implicit function theorem again, additionally providing us with a representation for $y'(u)$, out of which we could derive the adjoint equation.

References

[1] Behzad Azmi, Anne-Céline Boulanger, and Karl Kunisch. “On the semi-global stabilizability of the Korteweg-de Vries Equation via model predictive control”. In: *ESAIM: COCV* 24 (1 2018), pp. 237–263.

[2] Marino Badiale and Enrico Serra. *Semilinear Elliptic Equations for Beginners: Existence Results via the Variational Approach*. Universitext. London: Springer, 2011.

[3] Viorel Barbu et al. “Feedback stabilization of the Cahn–Hilliard type system for phase separation”. In: *Journal of Differential Equations* 262 (3 2017), pp. 2286–2334.

[4] Haim Brezis. *Functional Analysis, Sobolev Spaces and Partial Differential Equations*. Universitext. New York: Springer, 2011.

[5] Martin Brokate and Jürgen Sprekels. *Hysteresis and Phase Transitions*. Applied Mathematical Sciences. New York: Springer, 1996.

[6] Gunduz Caginalp. “An Analysis of a Phase Field Model of a Free Boundary”. In: *Archive for Rational Mechanics and Analysis* 92 (3 1986), pp. 205–245.

[7] Gunduz Caginalp. “The Dynamics of a Conserved Phase Field System: Stefan-like, Hele-Shaw, and Cahn-Hilliard Models as Asymptotic Limits”. In: *IMA Journal of Applied Mathematics* 44 (1 1990), pp. 77–94.

[8] Robert Denk. *Skript zur Vorlesung Analysis III - Maßtheorie*. Lecture notes. University of Konstanz, 2021.

[9] Robert Denk. *Skript zur Vorlesung Funktionalanalysis 2 (Operatortheorie)*. Lecture notes. University of Konstanz, 2023.

[10] Robert Denk. *Skript zur Vorlesung Theorie partieller Differentialgleichungen 2*. Lecture notes. University of Konstanz, 2023.

[11] Kay-Robert Dormann et al. *The Cahn-Hilliard Equation: Computer Simulations and Analytical Investigations*. Advanced theoretical lab work. Technische Universität Darmstadt, 2024.

[12] Lawrence C. Evans. *Partial Differential Equations*. 2nd ed. Vol. 19. Graduate Studies in Mathematics. Providence: American Mathematical Society, 2010.

[13] Michael Hinze et al. *Optimization with PDE Constraints*. Mathematical Modelling: Theory and Applications. Dordrecht: Springer, 2009.

[14] Simcha Horowitz. *Fundamental Theorem of Calculus for the Lebesgue Integral*. Lecture notes. Bar-Ilan University, 2019.

[15] Richard Jones. *Soft Condensed Matter*. Reprinted with corr. Vol. 6. Oxford Master Series in Condensed Matter Physics. Oxford: Oxford University Press, 2002.

[16] Alain Miranville. “On the conserved phase-field model”. In: *Journal of Mathematical Analysis and Applications* 400 (1 2013), pp. 143–152.

[17] Alain Miranville. *The Cahn-Hilliard Equation: Recent Advances and Applications*. CBMS-NSF Regional Conference Series in Applied Mathematics. Philadelphia: Society for Industrial and Applied Mathematics, 2019.

[18] David A. Porter and Kenneth E. Easterling. *Phase Transformations in Metals and Alloys*. Wokingham: Van Nostrand Reinhold, 1981.

[19] FEniCS Project. *Cahn-Hilliard equation*. 2017. URL: https://olddocs.fenicsproject.org/dolfin/latest/python/demos/cahn-hilliard/demo_cahn-hilliard.py.html.

[20] Michael Renardy and Robert C. Rodgers. *An Introduction to Partial Differential Equations*. 2nd ed. Texts in Applied Mathematics. New York: Springer, 1993.

[21] Roger Temam. *Navier-Stokes Equations: Theory and Numerical Analysis*. Vol. 2. Studies in Mathematics and its Applications. Amsterdam etc.: North-Holland Publishing Company, 1977.

[22] Hao Wu. "A review on the Cahn–Hilliard equation: classical results and recent advances in dynamic boundary conditions". In: *Electronic Research Archive* 30 (8 2022), pp. 2788–2832.