DISCRETE AND CONTINUOUS
DYNAMICAL SYSTEMS
Volume **15**, Number **2**, June **2006**

Website: http://AIMsciences.org

pp. **505–528**

A DESCENT METHOD FOR THE FREE ENERGY OF MULTICOMPONENT SYSTEMS

Herbert Gajewski and Jens A. Griepentrog

Weierstrass Institute for Applied Analysis and Stochastics
Mohrenstrasse 39, D-10117 Berlin, Germany

Abstract. Equilibrium distributions of multicomponent systems minimize the free energy functional under the constraint of mass conservation of the components. However, since the free energy is not convex in general, usually one tries to characterize and to construct equilibrium distributions as steady states of an adequate evolution equation, for example, the nonlocal Cahn–Hilliard equation for binary alloys. In this work a direct descent method for nonconvex functionals is established and applied to phase separation problems in multicomponent systems and image segmentation.

1. **Introduction.** To describe the phase separation model underlying this work we consider a closed multicomponent system with interacting particles of type $i \in \{0, 1, \dots, m\}$ occupying a spatial domain $\Omega \subset \mathbb{R}^n$. We assume that the particles jump around on a given microscopically scaled lattice following a stochastic exchange process (see [8]). On each lattice site sits exactly one particle (exclusion principle). Two particles of type i and ℓ change their sites x and y with a certain probability $p_{i\ell}(x, y)$ due to diffusion and interaction. The hydrodynamical limit leads to a system of conservation laws for $i \in \{0, 1, \dots, m\}$,

$$u_i' + \nabla \cdot j_i = 0 \quad \text{in } \mathbb{R}_+ \times \Omega, \quad \nu \cdot j_i = 0 \quad \text{on } \mathbb{R}_+ \times \partial\Omega, \quad u_i(0) = u_{0i} \quad \text{in } \Omega, \quad (1)$$

for (scaled) mass densities $u_0, u_1, \dots, u_m$, their initial values $u_{00}, u_{01}, \dots, u_{0m}$, and current densities $j_0, j_1, \dots, j_m$. In general we can assume $\sum_{i=0}^m u_i = 1$ due to the exclusion principle, that means, only m of the $m + 1$ equations in (1) are independent of each other. Hence, we can drop out one equation, say that one for the zero component, and describe the state of the system by m-component vectors $u = (u_1, \dots, u_m)$ and $u_0 = 1 - \sum_{i=1}^m u_i$.

Equilibrium distributions $u^* = (u_1^*, \dots, u_m^*) : \Omega \longrightarrow \mathbb{R}^m$ of the multicomponent system and, more generally, steady states of (1) can be supposed to be (local) minimizers of the free energy functional F under the constraint of mass conservation:

$$F(u^*) = \min \left\{ F(u) : \textstyle\int_\Omega (u_i - u_{0i})\, dx = 0 \text{ for all } i \in \{1, \dots, m\} \right\},$$

or solutions (u^*, μ^*) of the corresponding Euler–Lagrange equations including Lagrange multipliers $\mu^* \in \mathbb{R}^m$:

$$\sum_{i=1}^m \mu_i^* g_i = DF(u^*), \quad \langle g_i, u \rangle = \int_\Omega u_i \, dx, \quad \langle g_i, u^* - u_0 \rangle = 0, \quad i \in \{1, \dots, m\}. \quad (2)$$

2000 *Mathematics Subject Classification.* 90C26, 82B26, 94A08.

Key words and phrases. Nonconvex functionals, Lyapunov function, asymptotic behaviour, Cahn–Hilliard equation, phase separation, image segmentation.

Research partially supported by BMBF grant 03GANGB5.

In many applications one is originally interested in u^*. However, F is in general not convex, so it seems to be difficult to solve (2) directly. By this reason one tries to construct u^* as steady state of the evolution equation (1). That approach rests on the following consideration: Having in mind that the Lagrange multipliers μ_i^* should be constant, one assumes their antigradients to be driving forces towards equilibrium. This leads to the evolution system (1) with current densities $j_i = -\sum_{\ell=1}^{m} a_{i\ell}(u)\nabla\mu_\ell$ and positively semidefinite mobility matrix $(a_{i\ell})$ (see [9, 10, 13]). Evidently, F is a Lyapunov function of (1). It can be expected and is proved in some cases (see [9]) that solutions of (1) satisfy

$$\lim_{t\to\infty} F(u(t)) = F(u^*), \quad \lim_{t\to\infty} u(t) = u^*,$$

where u^* is a solution to the Euler–Lagrange equations (2). However, from the practical point of view that approach becomes questionable if meta-stable states occur.

In this work we establish a direct method to solve (2). For a relevant class of nonconvex free energies F we define iteration sequences (u_k, μ_k) as solutions of auxiliary Euler–Lagrange equations such that $(F(u_k))$ decreases and μ_{ki} are constants. Moreover, we prove convergence results

$$\lim_{k\to\infty} F(u_k) = F(u^*), \quad \lim_{k\to\infty} u_k = u^*, \quad \lim_{k\to\infty} \mu_k = \mu^*,$$

where the strong limit (u^*, μ^*) of the sequence (u_k, μ_k) satisfies (2).

In Section 2 we formulate assumptions and the constrained minimum problem in a more general functional analytic setting. The assumptions will be verified in the sections concerned with applications. In Section 3 we establish the direct method. Local phase separation problems in binary alloys are considered in Section 4. Section 5 is devoted to nonlocal phase separation problems in multicomponent systems. In Section 6 we describe an image segmentation algorithm. Finally, in Section 7 we conclude with simulation results for ternary systems.

2. **The Constrained Minimum Problem.**

Let $(H, \|\ \|_H)$ be a separable Hilbert space, $(H^*, \|\ \|_{H^*})$ its dual, and $\langle\ ,\ \rangle$ the dual pairing between H and H^*. In addition to that, we denote by $J \in \mathcal{L}(H; H^*)$ the duality map between H and H^* and by $R \in \mathcal{L}(H^*; H)$ its inverse. We consider functionals $\Phi : H \longrightarrow \mathbb{R} \cup \{+\infty\}$ and $\Psi : H \longrightarrow \mathbb{R}$ satisfying

Assumption 1. Let $\Phi : H \longrightarrow \mathbb{R} \cup \{+\infty\}$ be a proper, lower semicontinuous, and strongly convex functional with closed effective domain $\mathrm{dom}(\Phi) \subset H$. That means, there exists some $\alpha > 0$ such that for all u, $v \in \mathrm{dom}(\Phi)$ and $\tau \in [0, 1]$ we have

$$\tau\Phi(u) + (1-\tau)\Phi(v) \geq \Phi(\tau u + (1-\tau)v) + \tfrac{\alpha}{2}\tau(1-\tau)\|u - v\|_H^2. \tag{3}$$

Let $\Psi : H \longrightarrow \mathbb{R}$ be bounded from below on $\mathrm{dom}(\Phi) \subset H$ and Fréchet differentiable on H with Lipschitz continuous and compact Fréchet derivative $D\Psi : H \longrightarrow H^*$, that means, there exists some constant $\beta > 0$ such that for all u, $v \in H$,

$$\|D\Psi(u) - D\Psi(v)\|_{H^*} \leq \beta\, \|u - v\|_H. \tag{4}$$

Remark 1. As a consequence, the subdifferential $\partial\Phi \subset H \times H^*$ is both strongly monotone and maximal monotone. Furthermore, under the above general assumptions the sum $F = \Phi + \Psi : H \longrightarrow \mathbb{R} \cup \{+\infty\}$ is a well-defined functional with nonempty, closed and convex effective domain $\mathrm{dom}(F) = \mathrm{dom}(\Phi)$.

As a proper, lower semicontinuous, and convex functional $\Phi : H \longrightarrow \mathbb{R} \cup \{+\infty\}$ is weakly lower semicontinuous. Moreover, the complete continuity of the potential operator $D\Psi : H \longrightarrow H^*$ implies the strong continuity of its potential $\Psi : H \longrightarrow \mathbb{R}$. Hence, the sum $F = \Phi + \Psi$ is weakly lower semicontinuous, too.

We are interested in (local) minimizers $u^* \in K$ of $F : H \longrightarrow \mathbb{R} \cup \{+\infty\}$, where $K \subset \mathrm{dom}(F)$ represents a nonempty, closed, and convex set of given constraints.

Lemma 1 (Existence of minimizers)**.** *Assumption 1 implies that the functional $F : H \longrightarrow \mathbb{R} \cup \{+\infty\}$ is bounded from below. There exists a solution $u^* \in K$ of the constrained minimum problem*

$$F(u^*) = \min\{F(u) : u \in K\}. \tag{5}$$

Proof. Because of the boundedness of $\Psi : H \longrightarrow \mathbb{R}$ from below on $\mathrm{dom}(F)$ we can find a constant $c \in \mathbb{R}$ such that $\Psi(u) \geq c$ for all $u \in \mathrm{dom}(F)$. We fix $v \in K$ and $d \in \mathbb{R}$ with $F(v) < d + c$. The strong convexity of $\Phi : H \longrightarrow \mathbb{R} \cup \{+\infty\}$ implies the existence of some $r > 0$ such that $\Phi(u) \geq d$ for all $u \in H$, $\|u\|_H \geq r$. This yields

$$F(v) < d + c \leq \Phi(u) + \Psi(u) = F(u) \quad \text{for all } u \in \mathrm{dom}(F),\ \|u\|_H \geq r.$$

Hence, we have found $v \in K$ and $r > 0$ such that $F(v) < F(u)$ for all $u \in H$, $\|u\|_H \geq r$. That means, it suffices to look for a minimum of F on the nonempty, bounded, closed, and convex subset $K \cap \{u \in H : \|u\|_H \leq r\}$. Using the weak lower semicontinuity of F (see Remark 1) the generalized Weierstrass theorem yields both the existence of a solution $u^* \in K$ to the minimum problem (5) and the boundedness of F from below. □

3. **The Descent Method.** Knowing about the solvability of the constrained minimum problem (5) we want to establish a direct and constructive solution algorithm to find (local) minimizers of F. Our plan is to approximate (local) minimizers of the original problem (5) by a sequence of solutions of constrained minimum problems (7) for partially linearized functionals $F_u : H \longrightarrow \mathbb{R} \cup \{+\infty\}$ defined as

$$F_u(v) = \Phi(v) + \langle D\Psi(u), v\rangle, \quad u, v \in H. \tag{6}$$

Lemma 2. *Assumption 1 implies that $F_u : H \longrightarrow \mathbb{R} \cup \{+\infty\}$ is bounded from below for every $u \in H$. There exists a unique solution $v^* \in K$ of the constrained minimum problem*

$$F_u(v^*) = \min\{F_u(v) : v \in K\}. \tag{7}$$

Proof. For all $u \in H$ the functional F_u is proper, lower semicontinuous, and strongly convex. Hence, it is both weakly lower semicontinuous and weakly coercive. The desired result is a consequence of the generalized Weierstrass theorem. □

Lemma 3 (Descent property)**.** *Let Assumption 1 be satisfied and $v^* \in K$ be the solution of problem* (7) *for fixed $u \in K$. Then for all parameters $\tau \in (0, 1]$ and $u_\tau = \tau v^* + (1 - \tau)u \in K$ we have*

$$F(u) - F(u_\tau) \geq \left(\tfrac{\alpha}{2\tau} - \tfrac{\beta}{2}\right)\|u - u_\tau\|_H^2.$$

Proof. 1. For all u, $v \in H$ we can use Lagrange's formula and (4) to get

$$\begin{aligned}\Psi(v) - \Psi(u) - \langle D\Psi(u), v - u\rangle &= \int_0^1 \langle D\Psi(u + s(v - u)) - D\Psi(u), v - u\rangle\, ds \\ &\leq \int_0^1 \beta s \|v - u\|_H^2\, ds = \tfrac{\beta}{2}\|v - u\|_H^2.\end{aligned}$$

2. Let $u \in K$ be fixed and $v^* \in K$ the solution of problem (7). Then for $\tau \in (0,1]$ and $u_\tau = \tau v^* + (1-\tau)u \in K$ the estimate

$$\Phi(v^*) + \langle D\Psi(u), v^* \rangle \le \Phi(u_\tau) + \langle D\Psi(u), u_\tau \rangle$$

holds true. Together with the strong convexity of Φ (see (3)) this yields

$$\begin{aligned}(1-\tau)\big(\Phi(u) - \Phi(u_\tau)\big) &\ge \tau\big(\Phi(u_\tau) - \Phi(v^*)\big) + \tfrac{\alpha}{2}\tau(1-\tau)\|v^* - u\|_H^2 \\ &\ge \tau\langle D\Psi(u), v^* - u_\tau\rangle + \tfrac{\alpha}{2}\tau(1-\tau)\|v^* - u\|_H^2 \\ &= (1-\tau)\langle D\Psi(u), u_\tau - u\rangle + \tfrac{\alpha}{2\tau}(1-\tau)\|u_\tau - u\|_H^2.\end{aligned}$$

Let $\tau \in (0,1)$. Dividing both sides by $1-\tau$ and adding $\Psi(u) - \Psi(u_\tau)$ this implies

$$\begin{aligned}F(u) - F(u_\tau) &\ge \Psi(u) - \Psi(u_\tau) + \langle D\Psi(u), u_\tau - u\rangle + \tfrac{\alpha}{2\tau}\|u_\tau - u\|_H^2 \\ &\ge \big(\tfrac{\alpha}{2\tau} - \tfrac{\beta}{2}\big)\|u_\tau - u\|_H^2,\end{aligned}$$

where we have used the estimate presented in Step 1 of the proof. The desired estimate remains true for $\tau = 1$ since the lower semicontinuity of the functional $F : H \longrightarrow \mathbb{R} \cup \{+\infty\}$ allows us to take the limit $\tau \uparrow 1$. □

Lemma 4 (Descent method)**.** *Let Assumption 1 and $\alpha > \beta\tau$ be satisfied for the parameter $\tau \in (0,1]$. Let $u_0 \in K$ and define the sequences $(v_k), (u_k) \subset K$ by*

$$u_{k+1} = \tau v_k + (1-\tau)u_k, \quad F_{u_k}(v_k) = \min\big\{F_{u_k}(v) : v \in K\big\} \tag{8}$$

Then $(F(u_k))$ is decreasing and convergent. In fact, we have the estimate

$$\|u_k - u_{k+1}\|_H^2 \le \tfrac{2\tau}{\alpha - \beta\tau}\big(F(u_k) - F(u_{k+1})\big) \quad \textit{for all } k \in \mathbb{N}. \tag{9}$$

Proof. In view of Lemma 2 the sequences $(v_k), (u_k) \subset K$ are correctly defined. Using Lemma 3 with $u_\tau = u_{k+1}$, $u = u_k$, $v^* = v_k$ we see that $(F(u_k))$ is decreasing and

$$F(u_k) - F(u_{k+1}) \ge \big(\tfrac{\alpha}{2\tau} - \tfrac{\beta}{2}\big)\|u_k - u_{k+1}\|_H^2 \quad \text{for all } k \in \mathbb{N}.$$

If $F(u_{k+1}) = F(u_k)$ for some $k \in \mathbb{N}$ then the estimate yields $u_{k+1} = u_k$, and the sequence arrives at a stationary point. Otherwise we have $F(u_k) > F(u_{k+1})$. By Lemma 1 the sequence $(F(u_k))$ is bounded from below which implies its convergence. □

Assumption 2. Let $(V, \|\ \|_V)$ and $(W, \|\ \|_W)$ be Banach spaces densely and continuously embedded into the Hilbert space $(H, \|\ \|)$ and its dual $(H^*, \|\ \|_*)$, respectively. We assume that the restriction $J|V$ of the duality map $J \in \mathcal{L}(H; H^*)$ to V is an isomorphism from V onto $W = J[V]$. Moreover, let $H = H_0 + H_1$ be a Hilbert sum representation of H where $H_1 \subset V$ is a finite dimensional subspace and $H_0 \subset H$ is its orthogonal complement in H. Let $P_1 \in \mathcal{L}(H; H_1)$ be the orthogonal projector onto H_1 and consider the annihilator of H_0:

$$H_0^0 = \{f \in H^* : \langle f, v\rangle = 0 \text{ for all } v \in H_0\} = J[H_1].$$

Assumption 3. Here, we specify the set K of constraints under consideration: Let $K \subset \operatorname{dom}(F)$ be a nonempty, closed, and convex set in H such that $u, v \in K$ implies $u - v \in H_0$. Moreover, we impose the following condition: For all $u \in K$ the Euler–Lagrange equation

$$f \in \partial\Phi(v^*) + D\Psi(u), \tag{10}$$

corresponding to (7), has a solution $(v^*, f) \in C \times M$ where $C \subset K \cap \operatorname{dom}(\partial\Phi)$ and $M \subset H_0^0$ are some bounded, closed, and convex sets in H and H^*, respectively.

Remark 2. Let Assumptions 1, 2, 3 be satisfied and $(v^*, f) \in \big(K \cap \mathrm{dom}(\partial\Phi)\big) \times H_0^0$ be a solution of (10). By definition of $\partial\Phi(v^*) \subset H^*$ for all $v \in K$ we have

$$\Phi(v) - \Phi(v^*) \geq \langle f, v - v^* \rangle - \langle D\Psi(u), v - v^* \rangle = \langle D\Psi(u), v^* - v \rangle.$$

That means, $v^* \in K$ is the solution of the constrained minimum problem (7) which is unique by Lemma 2. Hence, we can reformulate our descent method as follows:

Definition 1 (Descent method)**.** Let the Assumptions 1, 2, 3 and $\alpha > \beta\tau$ be satisfied for some $\tau \in (0, 1]$ and $u_0 \in K$ be some given start element. Then we define the sequences $(u_k) \subset K$ and $(v_k, f_k) \subset C \times M$ by

$$u_{k+1} = \tau v_k + (1 - \tau) u_k, \quad f_k \in \partial\Phi(v_k) + D\Psi(u_k). \tag{11}$$

Theorem 5 (Convergence of a subsequence)**.** *Under the Assumptions 1, 2, and 3 the sequence $(u_k, f_k) \subset K \times M$ constructed in Definition 1 contains a subsequence (u_{k_ℓ}, f_{k_ℓ}) which converges to some solution $(u^*, f^*) \in C \times M$ of the Euler–Lagrange equation*

$$f^* \in \partial\Phi(u^*) + D\Psi(u^*), \tag{12}$$

in the sense of

$$\lim_{k\to\infty} F(u_k) = F(u^*), \quad \lim_{\ell\to\infty} \|u_{k_\ell} - u^*\|_H = 0, \quad \lim_{\ell\to\infty} \|f_{k_\ell} - f^*\|_{H^*} = 0. \tag{13}$$

Proof. 1. Because of Remark 2 the solutions of the Euler–Lagrange equation (11) are solutions of the constrained minimum problem (8). Thus, Lemma 4 yields that $(F(u_k))$ is a decreasing and convergent sequence and $\lim_{k\to\infty} \|u_{k+1} - u_k\|_H = 0$.

Since $u_{k+1} - u_k = \tau(v_k - u_k)$ for all $k \in \mathbb{N}$ both sequences $(v_k) \subset C$ and $(u_k) \subset K$ are bounded in H. Together with the boundedness of $(f_k) \subset M$ in the finite dimensional subspace $H_0^0 = J[H_1]$ this implies the precompactness of the sequences $(D\Psi(u_k))$ and (f_k) in H^*. Hence, there exist a subsequence $(u_{k_\ell}, f_{k_\ell}) \subset (u_k, f_k)$ and accumulation points $u^* \in C$, $f^* \in M$, and $h^* \in H^*$ such that both $(v_{k_\ell}) \subset C$ and $(u_{k_\ell}) \subset K$ converge weakly to u^* in H, and $(f_{k_\ell}) \subset M$ and $(D\Psi(u_{k_\ell})) \subset H^*$ converge strongly to f^* and h^* in H^*, respectively. In view of the Euler–Lagrange equations

$$f_{k_\ell} - D\Psi(u_{k_\ell}) \in \partial\Phi(v_{k_\ell}) \quad \text{for all } \ell \in \mathbb{N}, \tag{14}$$

the maximal monotonicity of $\partial\Phi \subset H \times H^*$ allows us to take the limit $\ell \to \infty$ to get $f^* - h^* \in \partial\Phi(u^*)$.

2. The Euler–Lagrange equations (14) and the definition of $\partial\Phi(u_{k_\ell}) \in H^*$ yield

$$\Phi(v_{k_\ell}) - \Phi(u^*) \leq \big\langle f_{k_\ell}, v_{k_\ell} - u^* \big\rangle - \big\langle D\Psi(u_{k_\ell}), v_{k_\ell} - u^* \big\rangle = \big\langle D\Psi(u_{k_\ell}), u^* - v_{k_\ell} \big\rangle.$$

In the limit process $\ell \to \infty$ we can use the lower semicontinuity of Φ and the convergence results of Step 1 to get $\lim_{\ell\to\infty} \Phi(v_{k_\ell}) = \Phi(u^*)$ because of

$$0 \leq \liminf_{\ell\to\infty} \Phi(v_{k_\ell}) - \Phi(u^*) \leq \limsup_{\ell\to\infty} \Phi(v_{k_\ell}) - \Phi(u^*) \leq \lim_{\ell\to\infty} \big\langle D\Psi(u_{k_\ell}), u^* - v_{k_\ell} \big\rangle = 0.$$

On the other hand, the convexity of Φ and the identity $u_{k_\ell+1} = \tau v_{k_\ell} + (1 - \tau) u_{k_\ell}$ imply

$$\Phi(u_{k_\ell+1}) \leq \tau\, \Phi(v_{k_\ell}) + (1 - \tau)\, \Phi(u_{k_\ell}) \quad \text{for all } \ell \in \mathbb{N}.$$

Because of Step 1 both sequences (u_{k_ℓ}) and $(u_{k_\ell+1})$ converge weakly to u^* in H. Consequently, using the complete continuity of $D\Psi : H \longrightarrow H^*$ both sequences $(\Psi(u_{k_\ell}))$ and $(\Psi(u_{k_\ell+1}))$ tend to $\Psi(u^*)$. Due to the convergence of $(F(u_k))$ the sequences $(\Phi(u_{k_\ell}))$ and $(\Phi(u_{k_\ell+1}))$ converge to the same limit. In view of $\lim_{\ell\to\infty} \Phi(v_{k_\ell}) = \Phi(u^*)$ the limit process $\ell \to \infty$ in the last estimate yields

$$\lim_{\ell\to\infty} \Phi(u_{k_\ell}) = \lim_{\ell\to\infty} \Phi(u_{k_\ell+1}) \leq \tau\, \Phi(u^*) + (1 - \tau) \lim_{\ell\to\infty} \Phi(u_{k_\ell}),$$

that means, $\lim_{\ell\to\infty}\Phi(u_{k_\ell})\le\Phi(u^*)$. In fact, this implies $\lim_{\ell\to\infty}\Phi(u_{k_\ell})=\Phi(u^*)$ because of the lower semicontinuity of Φ. Together with the convergence of $(\Psi(u_{k_\ell}))$ to $\Psi(u^*)$ we get $\lim_{k\to\infty}F(u_k)=\lim_{\ell\to\infty}F(u_{k_\ell})=F(u^*)$.

3. Let $\ell\in\mathbb{N}$ be fixed and $u_s=su^*+(1-s)u_{k_\ell}\in K$ for $s\in(0,1)$. Due to the definition of $\partial\Phi(u^*)\in H^*$ and the results of Step 1 we have

$$\Phi(u_s)-\Phi(u^*)\ge\langle f^*,u_s-u^*\rangle-\langle h^*,u_s-u^*\rangle=\langle h^*,u^*-u_s\rangle.$$

In view of the strong convexity of Φ (see (3)) we get

$$\begin{aligned}(1-s)\big(\Phi(u_{k_\ell})-\Phi(u_s)\big)&\ge s\big(\Phi(u_s)-\Phi(u^*)\big)+\tfrac{\alpha}{2}s(1-s)\|u^*-u_{k_\ell}\|_H^2\\&\ge s\langle h^*,u^*-u_s\rangle+\tfrac{\alpha}{2}s(1-s)\|u^*-u_{k_\ell}\|_H^2\\&=(1-s)\langle h^*,u_s-u_{k_\ell}\rangle+\tfrac{\alpha}{2s}(1-s)\|u_s-u_{k_\ell}\|_H^2.\end{aligned}$$

Dividing both sides by $1-s$ this implies

$$\tfrac{\alpha}{2s}\|u_{k_\ell}-u_s\|_H^2\le\Phi(u_{k_\ell})-\Phi(u_s)+\langle h^*,u_{k_\ell}-u_s\rangle.$$

Using the lower semicontinuity of Φ the limit process $s\uparrow 1$ yields

$$\tfrac{\alpha}{2}\|u_{k_\ell}-u^*\|_H^2\le\big(\Phi(u_{k_\ell})-\Phi(u^*)\big)+\langle h^*,u_{k_\ell}-u^*\rangle\quad\text{for all }\ell\in\mathbb{N}.$$

In the limit process $\ell\to\infty$ both terms of the right hand side tend to zero: The first term due to the results of Step 2 and the last term because of the weak convergence of (u_{k_ℓ}) to u^* in H. Hence, we have shown $\lim_{\ell\to\infty}\|u_{k_\ell}-u^*\|_H=0$. Finally, the continuity of $D\Psi:H\longrightarrow H^*$ implies $h^*=D\Psi(u^*)$ and the desired Euler–Lagrange equation (12). □

In the case of strong convexity of the functional F the whole sequence converges to the uniquely determined limit point $(u^*,f^*)\in C\times M$. However, in general F is not convex, and we cannot apply this standard argument. Instead of this we follow the ideas of [6, 7, 15] using an appropriate Łojasiewicz–Simon type inequality. To ensure the validity of such an inequality for $F=\Phi+\Psi$ we impose sufficient conditions on the functionals Φ and Ψ suitable for our applications:

Assumption 4. Let $T\in\mathcal{L}(H;H^*)$ be a self-adjoint and completely continuous operator such that its restriction $T|V$ to V is a completely continuous operator in $\mathcal{L}(V;W)$. For fixed $l\in W$ and $d\in\mathbb{R}$ we consider the quadratic functional $\Psi:H\longrightarrow\mathbb{R}$ given by

$$\Psi(u)=\tfrac{1}{2}\langle Tu,u\rangle+\langle l,u\rangle+d,\quad u\in H.\tag{15}$$

Assumption 5. Let U be an open subset in V and $\Phi:U\longrightarrow\mathbb{R}$ be a Fréchet differentiable functional. Additionally, we assume that the Fréchet derivative $D\Phi:U\longrightarrow W$ is a real analytic operator (see [17] and Remark 4) which satisfies

$$\langle D\Phi(u)-D\Phi(v),u-v\rangle\ge\alpha\,\|u-v\|_H^2,\quad\|D\Phi(u)-D\Phi(v)\|_{H^*}\le\gamma\,\|u-v\|_H,\tag{16}$$

for all $u,v\in U$ and some constants α, $\gamma>0$. Moreover, the second Fréchet derivative $D^2\Phi(u)\in\mathcal{L}(V;W)$ is assumed to be an isomorphism for all $u\in U$.

Remark 3. If Assumptions 2 and 5 are satisfied then $D\Phi:U\longrightarrow W$ is injective. Therefore, the inverse mapping theorem for real analytic operators (see [17]) implies that for every $u\in U$ we can find an open neighbourhood $U_0\subset U$ of u in V such that the inverse $D\Phi^{-1}:D\Phi[U]\longrightarrow U$ is real analytic in the open neighbourhood $D\Phi[U_0]$ of $D\Phi(u)$ in W and, hence, in $D\Phi[U]$.

Remark 4. Let Assumptions 2 and 5 be satisfied and $u \in U$ be fixed such that $u+v \in U$ for all $v \in V$, $\|v\|_V < 2\delta$ and some $\delta > 0$. Because of the real analyticity we can use the Taylor expansion of the operator $D\Phi : U \longrightarrow W$ near $u \in U$. That means, there exist symmetric bounded k-linear forms $B_k(u) \in \mathcal{L}^k(V;W)$ such that both the power series

$$D\Phi(u+v) - D\Phi(u) = \sum_{k=1}^{\infty} \frac{1}{k!} B_k(u)[v,\dots,v], \quad \sum_{k=1}^{\infty} \frac{1}{k!} \|B_k(u)\|_{\mathcal{L}^k(V;W)} \|v\|_V^k \quad (17)$$

converge uniformly for $v \in V$, $\|v\|_V \le \delta$. For all $k \in \mathbb{N}$ we can define symmetric bounded $(k+1)$-linear forms $A_{k+1}(u) \in \mathcal{L}^{k+1}(V;\mathbb{R})$ by

$$A_{k+1}(u)[v_1,\dots,v_k,v] = \big\langle B_k(u)[v_1,\dots,v_k], v\big\rangle, \quad v_1,\dots,v_k,v \in V,$$

because the continuous embeddings of V in H and W in H^* imply an estimate

$$\|A_{k+1}(u)\|_{\mathcal{L}^{k+1}(V;\mathbb{R})} \le c\,\|B_k(u)\|_{\mathcal{L}^k(V;W)} \quad \text{for all } k \in \mathbb{N},$$

and some constant $c > 0$. Together with (17) this yields that for all $v \in V$, $\|v\|_V < \delta$ the Taylor expansion of $\Phi : U \longrightarrow \mathbb{R}$ near $u \in U$ has the form

$$\begin{aligned}\Phi(u+v) - \Phi(u) - \langle D\Phi(u), v\rangle &= \int_0^1 \langle D\Phi(u+sv) - D\Phi(u), v\rangle\, ds \\ &= \int_0^1 \sum_{k=1}^{\infty} \frac{1}{k!} \big\langle B_k(u)[v,\dots,v], v\big\rangle s^k\, ds \\ &= \sum_{k=1}^{\infty} \frac{1}{(k+1)!} A_{k+1}(u)[v,\dots,v],\end{aligned}$$

where we have used the uniform convergence of the power series

$$\sum_{k=1}^{\infty} \frac{1}{(k+1)!} \|A_{k+1}(u)\|_{\mathcal{L}^{k+1}(V;\mathbb{R})} \|v\|_V^{k+1} \le \sum_{k=1}^{\infty} \frac{c}{(k+1)!} \|B_k(u)\|_{\mathcal{L}^k(V;W)} \|v\|_V^{k+1}$$

for $v \in V$, $\|v\|_V \le \delta$. Hence, we have shown that $\Phi : U \longrightarrow \mathbb{R}$ is real analytic, too.

Theorem 6 (Łojasiewicz–Simon inequality)**.** *Let Assumptions 2, 4, and 5 be satisfied and $(u^*, f^*) \in U \times H_0^0$ a solution of the Euler–Lagrange equation $DF(u^*) = f^*$. Then we can find constants δ, $\lambda > 0$, and $\theta \in \big(0, \frac{1}{2}\big]$ such that for all $u \in U$ which satisfy $u - u^* \in H_0$ and $\|u - u^*\|_H \le \delta$ we have the following inequality:*

$$|F(u) - F(u^*)|^{1-\theta} \le \lambda \inf\big\{\|DF(u) - f\|_{H^*} : f \in H_0^0\big\}. \quad (18)$$

Proof. 1. Our proof closely follows the ideas of [7], but for our purpose we need a slightly more general Łojasiewicz–Simon inequality (18) suitable for the case when affine constraints have to be taken into account.

We introduce the spaces $\mathcal{H} = H \times H_0^0$, $\mathcal{K} = H^* \times H_1$, $\mathcal{V} = V \times H_0^0$, and $\mathcal{W} = W \times H_1$ equipped with the Euclidean norms of the corresponding product spaces. By virtue of Assumption 2 the spaces $\mathcal{V}$ and $\mathcal{W}$ are densely and continuously embedded in $\mathcal{H}$ and $\mathcal{K}$, respectively. Moreover, we set $\mathcal{U} = U \times H_0^0$ and define an augmented functional $\Lambda : \mathcal{U} \longrightarrow \mathbb{R}$ by

$$\Lambda(u,f) = F(u) - \langle f, u - u^*\rangle, \quad (u,f) \in \mathcal{U}.$$

Because of Assumptions 2, 4, 5, and Remark 4 the functional Λ is real analytic in $\mathcal{U}$. Its Fréchet derivative $A = D\Lambda : \mathcal{U} \longrightarrow \mathcal{W}$, given by the formula

$$\langle A(u,f), (v,g)\rangle = \langle DF(u) - f, v\rangle - \langle g, u - u^*\rangle, \quad (u,f) \in \mathcal{U},\ (v,g) \in \mathcal{V},$$

is a real analytic operator, and there exists a constant $c_1 > 0$ such that

$$\|A(u,f) - A(v,g)\|_{\mathcal{K}} \le c_1 \|(u,f) - (v,g)\|_{\mathcal{H}} \quad \text{for all } (u,f),\, (v,g) \in \mathcal{U}. \tag{19}$$

Obviously, $(u^*, f^*) \in \mathcal{U}$ is a critical point of Λ, that means, we have $A(u^*, f^*) = 0$.

2. Let $E \subset \mathbb{R}$ be the set of eigenvalues of the symmetric and completely continuous operator $RT \in \mathcal{L}(H;H)$. By virtue of the Riesz spectral theory there cannot exist nonzero accumulation points of the at most countable set $E \subset \mathbb{R}$. Hence, if we consider the decomposition of E into the subsets

$$E_1 = \bigl\{\omega \in E : \omega \ge -\tfrac{\alpha}{2}\bigr\}, \quad E_2 = \bigl\{\omega \in E : \omega < -\tfrac{\alpha}{2}\bigr\},$$

then E_2 is a finite subset of E. Consequently, the Hilbert sum $H_2 \subset H$ of orthogonal eigenspaces to the eigenvalues $\omega \in E_2$ of RT is a finite dimensional subspace of H. Let $P_2 \in \mathcal{L}(H;H_2)$ be the orthogonal projector onto H_2. Then we get a splitting of T into a sum $T = T_1 + T_2$ of the finite rank operator $T_2 = TP_2 \in \mathcal{L}(H;J[H_2])$ and the completely continuous operator $T_1 = T - T_2 \in \mathcal{L}(H;H^*)$.

Following Assumption 4 the restriction $T|V$ of T to V is a completely continuous operator in $\mathcal{L}(V;W)$. Together with the dense and continuous embedding of V in H the Riesz spectral theory yields that both operators $RT \in \mathcal{L}(H;H)$ and $RT|V \in \mathcal{L}(V;V)$ have the same nonzero eigenvalues and corresponding eigenspaces (see [5]). That means, we have $H_2 \subset V$ and $J[H_2] \subset W$.

3. In view of Step 1 and 2 it turns out to be convenient to write A as a difference $A = A_1 - A_2$ of the real analytic operator $A_1 : \mathcal{U} \longrightarrow \mathcal{W}$ given by

$$\langle A_1(u,f),(v,g)\rangle = \langle D\Phi(u) + T_1 u + l, v\rangle + \langle g, Rf - Rf^*\rangle, \quad (u,f) \in \mathcal{U},\ (v,g) \in \mathcal{V},$$

and the linear finite rank operator $A_2 \in \mathcal{L}(\mathcal{V};\mathcal{W})$ defined as

$$\langle A_2(u,f),(v,g)\rangle = \langle f - T_2 u, v\rangle + \langle g, P_1 u - P_1 u^* + Rf - Rf^*\rangle, \quad (u,f),(v,g) \in \mathcal{V}.$$

By virtue of Assumption 4 and 5 and the construction of T_1 we observe that $A_1 : \mathcal{U} \longrightarrow \mathcal{W}$ is injective because it satisfies

$$\begin{aligned}\min\bigl\{1, \tfrac{\alpha}{2}\bigr\}\|(u,f) - (v,g)\|^2_{\mathcal{H}} &\le \langle A_1(u,f) - A_1(v,g), (u,f) - (v,g)\rangle \\ &\le \|A_1(u,f) - A_1(v,g)\|_{\mathcal{K}}\, \|(u,f) - (v,g)\|_{\mathcal{H}}\end{aligned}$$

for all $(u,f),(v,g) \in \mathcal{U}$. Hence, the inverse operator $A_1^{-1} : A_1[\mathcal{U}] \longrightarrow \mathcal{U}$ exists, and we can find a constant $c_2 > 0$ such that for all (f,u), $(g,v) \in A_1[\mathcal{U}]$ we have

$$\|A_1^{-1}(f,u) - A_1^{-1}(g,v)\|_{\mathcal{H}} \le c_2 \|(f,u) - (g,v)\|_{\mathcal{K}}. \tag{20}$$

The Fréchet derivative $DA_1(u,f) \in \mathcal{L}(\mathcal{V};\mathcal{W})$ is symmetric and has the form

$$\langle DA_1(u,f)(w,h),(v,g)\rangle = \bigl(\langle D^2\Phi(u)w, v\rangle + \langle g, Rh\rangle\bigr) + \langle T_1 w, v\rangle,$$

for all $(u,f) \in \mathcal{U}$, (w,h), $(v,g) \in \mathcal{V}$. It can be interpreted as a sum of an isomorphism and a completely continuous operator. Furthermore, $DA_1(u,f)$ is injective because Assumptions 4, 5, and Step 2 imply

$$\langle DA_1(u,f)(v,g),(v,g)\rangle \ge \min\bigl\{1, \tfrac{\alpha}{2}\bigr\}\|(v,g)\|^2_{\mathcal{H}} \quad \text{for all } (u,f) \in \mathcal{U},\ (v,g) \in \mathcal{V}.$$

Hence, $DA_1(u,f) \in \mathcal{L}(\mathcal{V};\mathcal{W})$ itself is an isomorphism. The inverse mapping theorem for real analytic operators (see [17]) yields that for every $(u,f) \in \mathcal{U}$ there exists an open neighbourhood $\mathcal{U}_0 \subset \mathcal{U}$ of (u,f) in $\mathcal{V}$ such that the inverse $A_1^{-1} : A_1[\mathcal{U}] \longrightarrow \mathcal{U}$ is real analytic in the open neighbourhood $A_1[\mathcal{U}_0]$ of $A_1(u,f)$ in $\mathcal{W}$ and, consequently, in $A_1[\mathcal{U}]$.

4. Next, we define the real analytic functional $G : A_1[\mathcal{U}] \cap A_2[\mathcal{V}] \longrightarrow \mathbb{R}$ by

$$G(g,v) = \Lambda\bigl(A_1^{-1}(g,v)\bigr), \quad (g,v) \in A_1[\mathcal{U}] \cap A_2[\mathcal{V}].$$

Because of $DA_1^{-1}(g,v) \in \mathcal{L}(\mathcal{W};\mathcal{V})$ the chain rule yields $DG(g,v) \in \mathcal{V}$ and

$$\big\langle (f,u), DG(g,v)\big\rangle = \big\langle AA_1^{-1}(g,v), DA_1^{-1}(g,v)(f,u)\big\rangle \tag{21}$$

for all $(g,v) \in A_1[\mathcal{U}] \cap A_2[\mathcal{V}]$, $(f,u) \in \mathcal{W}$. Hence, $A_1(u^*,f^*) = A_2(u^*,f^*) = (f^* - T_2u^*, 0)$ is a critical point of G. Since $A_2[\mathcal{V}]$ is a finite dimensional subspace of $\mathcal{W}$ there exist constants $\lambda_1 > 0$, $\theta \in \big(0, \frac{1}{2}\big]$ and some open neighbourhood $\mathcal{V}_0 \subset \mathcal{U}$ of (u^*,f^*) in $\mathcal{V}$ such that G satisfies the classical Łojasiewicz inequality (see [4, 14]):

$$|G(g,v) - G(A_2(u^*,f^*))|^{1-\theta} \le \lambda_1 \|DG(g,v)\|_{\mathcal{H}} \quad \text{for all } (g,v) \in A_1[\mathcal{V}_0] \cap A_2[\mathcal{V}].$$

In view of $T_2 \in \mathcal{L}(H; J[H_2])$ and $P_1 \in \mathcal{L}(H; H_1)$ we can find some constants δ, $\delta_* > 0$ such that the image $A_2[\mathcal{U}(\delta,\delta_*)]$ of

$$\mathcal{U}(\delta,\delta_*) = \big\{(u,f) \in \mathcal{U} : \|u - u^*\|_H < \delta,\ \|f - f^*\|_{H^*} < \delta_*\big\}$$

is contained in the open neighbourhood $A_1[\mathcal{V}_0]$ of $A_2(u^*,f^*)$ in $\mathcal{W}$. Hence, for all $(u,f) \in \mathcal{U}(\delta,\delta_*)$ we arrive at

$$|G(A_2(u,f)) - G(A_2(u^*,f^*))|^{1-\theta} \le \lambda_1 \|DG(A_2(u,f))\|_{\mathcal{H}}. \tag{22}$$

5. To estimate the right hand side of the last inequality let $(u,f) \in \mathcal{U}(\delta,\delta_*)$. The symmetry of $DA_1^{-1}(A_2(u,f)) \in \mathcal{L}(\mathcal{W};\mathcal{V})$ and the dense and continuous embeddings of $\mathcal{V}$ and $\mathcal{W}$ in $\mathcal{H}$ and $\mathcal{K}$, respectively, yields that the norm of the extension of $DA_1^{-1}(A_2(u,f))$ in $\mathcal{L}(\mathcal{K};\mathcal{H})$ is not greater than the norm of $DA_1^{-1}(A_2(u,f))$ in $\mathcal{L}(\mathcal{W};\mathcal{V})$ (see [5]). Together with (21) this implies the following estimate:

$$\|DG(A_2(u,f))\|_{\mathcal{H}} \le \|DA_1^{-1}(A_2(u,f))\|_{\mathcal{L}(\mathcal{W};\mathcal{V})}\, \|AA_1^{-1}A_2(u,f)\|_{\mathcal{K}}.$$

By virtue of the real analyticity of $A_1^{-1} : A_1[\mathcal{U}] \longrightarrow \mathcal{U}$ the first factor is uniformly bounded on $\mathcal{U}(\delta,\delta_*)$ by a constant $c_3 > 0$, that means, we have

$$\|DG(A_2(u,f))\|_{\mathcal{H}} \le c_3 \|AA_1^{-1}A_2(u,f)\|_{\mathcal{K}} \quad \text{for all } (u,f) \in \mathcal{U}(\delta,\delta_*). \tag{23}$$

Applying the estimates (20) and (19) we get

$$\|A_1^{-1}A_2(u,f) - (u,f)\|_{\mathcal{H}} = \|A_1^{-1}A_2(u,f) - A_1^{-1}A_1(u,f)\|_{\mathcal{H}} \le c_2 \|A(u,f)\|_{\mathcal{K}}, \tag{24}$$

and, consequently,

$$\|AA_1^{-1}A_2(u,f) - A(u,f)\|_{\mathcal{K}} \le c_1 \|A_1^{-1}A_2(u,f) - A_1^{-1}A_1(u,f)\|_{\mathcal{H}} \le c_1c_2 \|A(u,f)\|_{\mathcal{K}},$$

which implies for all $(u,f) \in \mathcal{U}(\delta,\delta_*)$ the relation

$$\begin{aligned} \|AA_1^{-1}A_2(u,f)\|_{\mathcal{K}} &\le \|A(u,f)\|_{\mathcal{K}} + \|AA_1^{-1}A_2(u,f) - A(u,f)\|_{\mathcal{K}} \\ &\le (1 + c_1c_2)\|A(u,f)\|_{\mathcal{K}}. \end{aligned}$$

In view of (22), (23) this yields the existence of some constant $c_4 > 0$ such that

$$|\Lambda(A_1^{-1}A_2(u,f)) - \Lambda(u^*,f^*)|^{1-\theta} \le c_4^{1-\theta} \|A(u,f)\|_{\mathcal{K}} \quad \text{for all } (u,f) \in \mathcal{U}(\delta,\delta_*). \tag{25}$$

6. Using (19) and Lagrange's formula we get

$$\begin{aligned} |\Lambda(v,g) - \Lambda(u,f)| &= \left| \int_0^1 \langle A(s(v,g) + (1-s)(u,f)), (v,g) - (u,f)\rangle\, ds \right| \\ &\le c_1 \|(v,g) - (u,f)\|_{\mathcal{H}}^2 + \|A(u,f)\|_{\mathcal{K}}\, \|(v,g) - (u,f)\|_{\mathcal{H}} \end{aligned}$$

for all (u,f), $(v,g) \in \mathcal{U}$ and, hence, by virtue of (24)

$$|\Lambda(A_1^{-1}A_2(u,f)) - \Lambda(u,f)| \le c_1c_2^2 \|A(u,f)\|_{\mathcal{K}}^2 + c_2 \|A(u,f)\|_{\mathcal{K}}^2$$

for all $(u,f) \in \mathcal{U}(\delta,\delta_*)$. Together with (25) this yields

$$|\Lambda(u,f) - \Lambda(u^*,f^*)| \le |\Lambda(A_1^{-1}A_2(u,f)) - \Lambda(u,f)| + |\Lambda(A_1^{-1}A_2(u,f)) - \Lambda(u^*,f^*)| \le (c_1c_2^2 + c_2)\|A(u,f)\|_{\mathcal{K}}^2 + c_4\|A(u,f)\|_{\mathcal{K}}^{1/(1-\theta)}.$$

Due to (19) we can choose δ, $\delta_* > 0$ small enough such that for all $(u,f) \in \mathcal{U}(\delta,\delta_*)$ we have $\|A(u,f)\|_{\mathcal{K}} \le c_1\|(u,f) - (u^*,f^*)\|_{\mathcal{H}} \le 1$ which implies

$$|\Lambda(u,f) - \Lambda(u^*,f^*)|^{1-\theta} \le (c_1c_2^2 + c_2 + c_4)^{1-\theta}\|A(u,f)\|_{\mathcal{K}}.$$

In view of the estimate

$$\|A(u,f)\|_{\mathcal{K}} \le \|DF(u) - f\|_{H^*} + \sup\big\{|\langle g, u-u^*\rangle| : g \in H_0^0, \|g\|_{H^*} \le 1\big\},$$

(see Step 1) and the identity

$$\sup\big\{|\langle g, w\rangle| : g \in H_0^0, \|g\|_{H^*} \le 1\big\} = \inf\big\{\|w - v\|_H : v \in H_0\big\} \quad \text{for all } w \in H,$$

we can find some constant $\lambda > 0$ such that for all $u \in U$, $\|u-u^*\|_H \le \delta$, and $f \in H_0^0$, $\|f - f^*\|_{H^*} \le \delta_*$, $v \in H_0$ we have

$$|F(u) - F(u^*) - \langle f, u-u^*\rangle|^{1-\theta} \le \lambda\big(\|DF(u) - f\|_{H^*} + \|u - u^* - v\|_H\big).$$

For all $u \in U$, $u - u^* \in H_0$, $\|u-u^*\|_H \le \delta$, and $f \in H_0^0$, $\|f-f^*\|_{H^*} \le \delta_*$ we get

$$|F(u) - F(u^*)|^{1-\theta} \le \lambda\,\|DF(u) - f\|_{H^*}. \tag{26}$$

We can choose $\beta \ge \|T\|_{\mathcal{L}(H;H^*)}$ and $\delta > 0$ small enough such that $(\gamma+\beta)\delta \le \delta_*$. Let $u \in U$ be such that $u - u^* \in H_0$ and $\|u-u^*\|_H \le \delta$, and let $f \in H_0^0$ satisfy

$$\|DF(u) - f\|_{H^*} = \inf\big\{\|DF(u) - g\|_{H^*} : g \in H_0^0\big\}.$$

Then the Lipschitz continuity of $D\Psi$ and $D\Phi$ on U (see (15) and (16)) yields

$$\|f-f^*\|_{H^*}^2 + \|DF(u) - f\|_{H^*}^2 = \|DF(u) - DF(u^*)\|_{H^*}^2 \le (\gamma+\beta)^2\|u-u^*\|_H^2 \le \delta_*^2.$$

Having in mind (26) this implies the desired result. □

Theorem 7 (Convergence of the whole sequence)**.** *Let Assumptions 1, 2, 3, 4, 5 be satisfied. If we assume that U is bounded in V, $K \subset V$, and*

$$\tau C + (1-\tau)K \subset U \subset \operatorname{dom}(\partial\Phi),$$

then the sequence $(u_k, f_k) \subset K \times M$ constructed in Definition 1 converges to some solution $(u^*, f^*) \in C \times M$ of the Euler–Lagrange equation

$$DF(u^*) = f^*, \tag{27}$$

in the sense of

$$\lim_{k\to\infty} F(u_k) = F(u^*), \quad \lim_{k\to\infty}\|u_k - u^*\|_V = 0, \quad \lim_{k\to\infty}\|f_k - f^*\|_W = 0. \tag{28}$$

Proof. 1. Theorem 5 ensures the convergence of a subsequence $(u_{k_\ell}, f_{k_\ell}) \subset (u_k, f_k)$ to some solution $(u^*, f^*) \in C \times M$ of problem (12) in the sense of (13). By virtue of Theorem 6 we can choose constants δ, $\lambda > 0$, and $\theta \in \big(0, \frac{1}{2}\big]$ such that for all $u \in K \cap U$, $\|u-u^*\|_H \le \delta$, the Łojasiewicz–Simon inequality (18) holds true. Following the ideas of [15], for every $\varepsilon \in \big(0, \frac{\delta}{2}\big)$ we define numbers $k(\varepsilon)$, $m(\varepsilon) \in \mathbb{N}$, and $n(\varepsilon) \in \mathbb{N} \cup \{+\infty\}$ by

$$k(\varepsilon) = \min\big\{k \in \mathbb{N} : \|u_\ell - u_{\ell+1}\|_H \le \tfrac{\varepsilon}{2} \text{ for all } \ell \ge k\big\}, \tag{29}$$

$$m(\varepsilon) = \min\big\{m \ge k(\varepsilon) : \|u_m - u^*\|_H \le \tfrac{\varepsilon}{2},\ [F(u_m) - F(u^*)]^\theta \le \tfrac{(\alpha-\beta\tau)\theta\varepsilon}{4\lambda\gamma}\big\}, \tag{30}$$

$$n(\varepsilon) = \sup\big\{m \ge m(\varepsilon) : \|u_\ell - u^*\|_H \le \delta \text{ for all } m(\varepsilon) \le \ell \le m\big\}, \tag{31}$$

and, furthermore, the subset $N(\varepsilon) = \{\ell \in \mathbb{N} : m(\varepsilon) \le \ell \le n(\varepsilon)\}$ of $\mathbb{N}$.

If we have $F(u_{\ell+1}) = F(u_\ell)$ for some $\ell \in \mathbb{N}$, then the descent property (9) yields $u_{\ell+1} = u_\ell$, and the sequence (u_ℓ, f_ℓ) arrives at a stationary point. Hence, it is sufficient to consider the case where

$$F(u_\ell) > F(u_{\ell+1}) > F(u^*), \quad \|u_\ell - u_{\ell+1}\|_H > 0 \quad \text{for all } \ell \in \mathbb{N}.$$

Due to the above construction it is easy to see that each number defined in (29), (30), and (31) goes to infinity if $\varepsilon \in \left(0, \frac{\delta}{2}\right)$ tends to zero, and that $m(\varepsilon) + 1 \le n(\varepsilon)$ holds true.

2. In view of Step 1 of the proof the Łojasiewicz–Simon inequality (18) yields

$$[F(u_\ell) - F(u^*)]^{1-\theta} \le \lambda \inf\left\{\|DF(u_\ell) - f\|_{H^*} : f \in H_0^0\right\} \quad \text{for all } \ell \in N(\varepsilon).$$

To estimate the right hand side we make use of the identity

$$\inf\left\{\|g - f\|_{H^*} : f \in H_0^0\right\} = \sup\left\{|\langle g, v\rangle| : v \in H_0,\ \|v\|_H \le 1\right\} \quad \text{for all } g \in H^*.$$

Because of $D\Phi(v_\ell) + D\Psi(u_\ell) = f_\ell \in H_0^0$ for all $\ell \in \mathbb{N}$ and $v \in H_0$ we have

$$\begin{aligned}\langle DF(u_\ell), v\rangle &= \langle D\Phi(u_\ell) - D\Phi(v_\ell), v\rangle + \langle D\Phi(v_\ell) + D\Psi(u_\ell), v\rangle \\ &= \langle D\Phi(u_\ell) - D\Phi(v_\ell), v\rangle.\end{aligned}$$

The Lipschitz continuity of $D\Phi$ on U (see (16)) and $u_{\ell+1} - u_\ell = \tau(v_\ell - u_\ell)$ imply

$$[F(u_\ell) - F(u^*)]^{1-\theta} \le \lambda\gamma \|u_\ell - v_\ell\|_H = \tfrac{\lambda\gamma}{\tau} \|u_{\ell+1} - u_\ell\|_H \quad \text{for all } \ell \in N(\varepsilon).$$

Hence, using the descent property (9) for all $\ell \in N(\varepsilon)$ we get

$$\begin{aligned}\|u_{\ell+1} - u_\ell\|_H &\le \tfrac{\lambda\gamma}{\tau} [F(u_\ell) - F(u^*)]^{\theta-1} \|u_{\ell+1} - u_\ell\|_H^2, \\ &\le \tfrac{2\lambda\gamma}{\alpha-\beta\tau} [F(u_\ell) - F(u^*)]^{\theta-1} \left([F(u_\ell) - F(u^*)] - [F(u_{\ell+1}) - F(u^*)]\right).\end{aligned}$$

Applying the elementary inequality $\theta a^{\theta-1}(a-b) \le a^\theta - b^\theta$ for $0 < b \le a$ it follows

$$\|u_{\ell+1} - u_\ell\|_H \le \tfrac{2\lambda\gamma}{(\alpha-\beta\tau)\theta} \left([F(u_\ell) - F(u^*)]^\theta - [F(u_{\ell+1}) - F(u^*)]^\theta\right) \quad \text{for all } \ell \in N(\varepsilon).$$

Summing up and using the definition of $m(\varepsilon)$ (see (30)) for all $k \in N(\varepsilon)$ we obtain

$$\|u_k - u_{m(\varepsilon)}\|_H \le \tfrac{2\lambda\gamma}{(\alpha-\beta\tau)\theta} \left([F(u_{m(\varepsilon)}) - F(u^*)]^\theta - [F(u_k) - F(u^*)]^\theta\right) \le \tfrac{\varepsilon}{2},$$

and, hence,

$$\|u_k - u^*\|_H \le \|u_k - u_{m(\varepsilon)}\|_H + \|u_{m(\varepsilon)} - u^*\|_H \le \tfrac{\varepsilon}{2} + \tfrac{\varepsilon}{2} \le \varepsilon \quad \text{for all } k \in N(\varepsilon). \tag{32}$$

As a consequence, there cannot exist sequences $(\varepsilon_\ell) \subset \left(0, \frac{\delta}{2}\right)$, $(k_\ell) \subset \mathbb{N}$ such that

$$\lim_{\ell\to\infty} \varepsilon_\ell = 0, \quad k_\ell \in N(\varepsilon_\ell), \quad \|u_{k_\ell} - u^*\|_H > \tfrac{\delta}{2} \quad \text{for all } \ell \in \mathbb{N}.$$

That means, we can find some $\varepsilon^* \in \left(0, \frac{\delta}{2}\right)$ such that for every $\varepsilon \in (0, \varepsilon^*)$ we have

$$\|u_k - u^*\|_H \le \tfrac{\delta}{2} \quad \text{for all } k \in N(\varepsilon).$$

In addition to that, due to (29) we also get

$$\|u_{k+1} - u^*\|_H \le \|u_{k+1} - u_k\|_H + \|u_k - u^*\|_H \le \tfrac{\varepsilon}{2} + \tfrac{\delta}{2} \le \delta \quad \text{for all } k \in N(\varepsilon).$$

Because of (31) this implies $n(\varepsilon) = +\infty$. In view of (32) for all $\varepsilon \in (0, \varepsilon^*)$, $k \ge m(\varepsilon)$ the estimate $\|u_k - u^*\|_H \le \varepsilon$ holds true, which means $\lim_{k\to\infty} \|u_k - u^*\|_H = 0$ and $\lim_{k\to\infty} \|v_k - u^*\|_H = 0$. The Lipschitz continuity of $D\Psi$ and $D\Phi$ on U yields

$$\begin{aligned}\|f_k - f^*\|_{H^*} &\le \|D\Phi(v_k) - D\Phi(u^*)\|_{H^*} + \|D\Psi(u_k) - D\Psi(u^*)\|_{H^*} \\ &\le \gamma \|v_k - u^*\|_H + \beta \|u_k - u^*\|_H,\end{aligned}$$

and, therefore, $\lim_{k\to\infty} \|f_k - f^*\|_{H^*} = 0$.

3. According to Assumption 4 the restriction $T|V$ is a completely continuous operator in $\mathcal{L}(V;W)$. The boundedness of $(u_k) \subset \tau C + (1-\tau)K \subset U$ in V yields that $(D\Psi(u_k))$ is precompact in W. In addition to that, the sequence $(f_k) \subset H_0^0$ converges to $f^* \in H_0^0$ in H^* (see Step 2) and, hence, in W because H_0^0 is a finite dimensional subspace of W. Therefore, $(f_k - D\Psi(u_k))$ is precompact in W, too.

Due to Assumption 5 and Remark 3 the inverse $(D\Phi)^{-1} : D\Phi[U] \longrightarrow U$ is real analytic in $D\Phi[U]$. By virtue of $(v_k) \subset C \subset U$ we have $f_k - D\Psi(u_k) = D\Phi(v_k) \in D\Phi[U]$ for all $k \in \mathbb{N}$. Therefore, the image of $(f_k - D\Psi(u_k))$ under $(D\Phi)^{-1} : D\Phi[U] \longrightarrow U$ is precompact in V. Hence, $(v_k) \subset C$ converges to $u^* \in C$ not only in H (see Step 2) but also in V.

4. It remains to show that $(u_k) \subset K$ converges to $u^* \in C$ in V. In view of the definition $u_{k+1} = \tau v_k + (1-\tau)u_k$ and the elementary identity

$$\sum_{\ell=1}^{k} \tau(1-\tau)^{k-\ell} = 1 - (1-\tau)^k$$

we get the representation

$$u_k - u^* = \sum_{\ell=1}^{k} \tau(1-\tau)^{k-\ell}(v_\ell - u^*) + (1-\tau)^k(u_0 - u^*) \quad \text{for all } k \in \mathbb{N}.$$

Let $c > 0$ be some constant such that $\|u_0\|_V \le c$, $\|u^*\|_V \le c$, and $\|v_\ell\|_V \le c$ for all $\ell \in \mathbb{N}$, and let $\varepsilon > 0$ be fixed. Due to Step 3 we can find some $k_0 \in \mathbb{N}$ such that $2c(1-\tau)^{k_0} \le \varepsilon$ and $\|v_\ell - u^*\|_V \le \varepsilon$ for all $\ell \in \mathbb{N}$, $\ell \ge k_0$. Hence, for all $k \in \mathbb{N}$, $k \ge 2k_0$ we arrive at

$$\begin{aligned}\|u_k - u^*\|_V &\le \sum_{\ell=1}^{k} \tau(1-\tau)^{k-\ell}\|v_\ell - u^*\|_V + (1-\tau)^k\|u_0 - u^*\|_V \\ &\le \sum_{\ell=1}^{k_0} 2c\tau(1-\tau)^{k-\ell} + \sum_{\ell=k_0+1}^{k} \varepsilon\tau(1-\tau)^{k-\ell} + 2c(1-\tau)^k,\end{aligned}$$

that means, $\|u_k - u^*\|_V \le 2c(1-\tau)^{k-k_0} + \varepsilon + 2c(1-\tau)^k \le 3\varepsilon$ which yields the result. □

4. **Phase Separation in Binary Alloys.** We consider a closed binary system of particles interacting in a bounded domain $\Omega \subset \mathbb{R}^n$ with Lipschitz boundary (see [11, 12]). We describe the state of the system by the density $u : \Omega \longrightarrow [0,1]$ of one component. Naturally, $1-u$ is the density of the other component.

Our plan is to apply the descent method to the free energy functional of the classical Cahn–Hilliard phase field theory (see [2]). Usually, it is defined as a sum of a double-well potential and an interface energy term. In contrast to that we split $F : H \longrightarrow \mathbb{R} \cup \{+\infty\}$ into a sum of a convex functional $\Phi : H \longrightarrow \mathbb{R} \cup \{+\infty\}$ and a concave functional $\Psi : H \longrightarrow \mathbb{R}$.

Assumption 6. To satisfy Assumption 2 we introduce the spaces $H = H^1(\Omega)$ and $V = H^1(\Omega) \cap L^\infty(\Omega)$, $W = J[V]$ equipped with the norms

$$\|u\|_H^2 = \int_\Omega \left(|\nabla u|^2 + |u|^2\right) dx, \quad u \in H,$$

$$\|u\|_V^2 = \int_\Omega \left(|\nabla u|^2 + |u|^2\right) dx + \operatorname*{ess\,sup}_{x\in\Omega} |u(x)|^2, \quad u \in V, \qquad \|f\|_W = \|Rf\|_V, \quad f \in W.$$

We consider the Hilbert sum decomposition $H = H_0 + H_1$ into the closed subspace

$$H_0 = \{u \in H : \int_\Omega u\, dx = 0\},$$

and the one-dimensional subspace $H_1 \subset V$ of constant functions. Then the annihilator $H_0^0 = J[H_1]$ is the one-dimensional space $\{\mu g \in W : \mu \in \mathbb{R}\}$ where $g \in W$ is given by $\langle g, u\rangle = \int_\Omega u\, dx$, $u \in H$.

Assumption 7. Let $\kappa > 0$ be a constant. We consider the proper, lower semicontinuous, and strongly convex functional $\varphi : \mathbb{R} \longrightarrow \mathbb{R} \cup \{+\infty\}$ defined by

$$\varphi(s) = \begin{cases} s\log(s) + (1-s)\log(1-s) & \text{if } s \in [0,1], \\ +\infty & \text{otherwise.} \end{cases}$$

The lower semicontinuous and strongly convex functional $\Phi : H \longrightarrow \mathbb{R} \cup \{+\infty\}$ given by

$$\Phi(u) = \begin{cases} \int_\Omega \left(\frac{\kappa}{2}|\nabla u|^2 + \varphi(u)\right) dx & \text{if } u \in H,\ 0 \le u \le 1, \\ +\infty & \text{otherwise,} \end{cases} \tag{33}$$

has the closed effective domain $\operatorname{dom}(\Phi) = \{u \in H : 0 \le u \le 1\}$. Finally, for some constant $\varkappa > 0$ we define the concave functional $\Psi : H \longrightarrow \mathbb{R}$ by setting

$$\Psi(u) = \int_\Omega \varkappa u(1-u)\, dx, \quad u \in H. \tag{34}$$

Remark 5. Note, that $\varphi'(s) = \log(s) - \log(1-s)$ and $\varphi''(s) = 1/s(1-s) \ge 4$ for all $s \in (0,1)$. Hence, $\Phi : H \longrightarrow \mathbb{R} \cup \{+\infty\}$ satisfies (3) with the constant $\alpha = \min\{\kappa, 4\}$. In addition to that, $\Psi : H \longrightarrow \mathbb{R}$ is bounded on $\operatorname{dom}(\Phi) \subset H$, and its Fréchet derivative $D\Psi : H \longrightarrow H^*$ satisfies the Lipschitz condition (4) for $\beta = 2\varkappa$. According to Assumption 4 we set $d = 0$ and define $T \in \mathcal{L}(H; H^*)$ and $l \in W$ by

$$\langle Tu, v\rangle = -\int_\Omega 2\varkappa uv\, dx, \quad \langle l, v\rangle = \int_\Omega \varkappa v\, dx, \quad u,\ v \in H,$$

and the operator $S \in \mathcal{L}(L^2(\Omega); H^*)$ by $\langle Sw, v\rangle = \int_\Omega wv\, dx$, $v \in H$. Then both $S|H \in \mathcal{L}(H; H^*)$ and $T \in \mathcal{L}(H; H^*)$ are completely continuous operators because of the compact embedding of H in $L^2(\Omega)$. Due to results of elliptic regularity theory (see [12]) there exists a Hölder exponent $\nu \in (0,1)$ such that $RS|L^\infty(\Omega) \in \mathcal{L}(L^\infty(\Omega); H^1(\Omega) \cap C^\nu(\overline{\Omega}))$. Using the compact embedding of $C^\nu(\overline{\Omega})$ in $L^\infty(\Omega)$, the restrictions $S|V$ and $T|V$ are completely continuous operators in $\mathcal{L}(V; W)$. Hence, Assumptions 1 and 4 are satisfied, and the sum $F = \Phi + \Psi : H \longrightarrow \mathbb{R} \cup \{+\infty\}$ is a well-defined functional with nonempty, closed, and convex effective domain $\operatorname{dom}(F) = \operatorname{dom}(\Phi) \subset H$.

Lemma 8 (Uniform boundedness)**.** *Assumptions 6 and 7 imply the following:*

(i) *For all $\bar{u} \in (0,1)$ and $w \in L^\infty(\Omega)$ there exists a uniquely determined solution $(u, \mu) \in \operatorname{dom}(\partial\Phi) \times \mathbb{R}$ of the constrained problem*

$$\langle \partial\Phi(u), v\rangle = \int_\Omega (\mu - w)v\, dx \quad \text{for all } v \in H, \qquad \int_\Omega (u - \bar{u})\, dx = 0. \tag{35}$$

(ii) *Let $\check{w} \le w \le \hat{w}$ for some $\check{w}$, $\hat{w} \in \mathbb{R}$. There exist constants $\check{u}$, $\hat{u} \in (0,1)$, $\check{\mu}$, $\hat{\mu} \in \mathbb{R}$, $c > 0$ depending only on $\check{w}$, $\hat{w}$, $\bar{u}$ such that the solution (u, μ) satisfies*

$$\check{u} \le u \le \hat{u}, \quad \|u - \check{u}\|_H \le c, \quad \check{\mu} \le \mu \le \hat{\mu}.$$

Proof. 1. Let $\bar{u} \in (0,1)$ and $w \in L^\infty(\Omega)$ be given such that $\check{w} \le w \le \hat{w}$ for some bounds $\check{w}$, $\hat{w} \in \mathbb{R}$. By virtue of the strong monotonicity and surjectivity of $\varphi' : (0,1) \longrightarrow \mathbb{R}$ we can find numbers $\check{u} \in (0, \bar{u}]$ and $\hat{u} \in [\bar{u}, 1)$ such that

$$\varphi'(\check{u}) = \check{w} - \hat{w} + \varphi'(\bar{u}), \quad \varphi'(\hat{u}) = \hat{w} - \check{w} + \varphi'(\bar{u}).$$

Now, we take a regularization $\phi : \mathbb{R} \longrightarrow \mathbb{R}$ of φ such that $\phi' : \mathbb{R} \longrightarrow \mathbb{R}$ is Lipschitz continuous and strongly monotone, and ϕ' coincides with φ' on the interval $[\check{u}, \hat{u}]$. Hence, the regularization $A : H \longrightarrow H^*$ of $\partial\Phi \subset H \times H^*$ defined by

$$\langle Au, \psi \rangle = \int_\Omega \big(\kappa \nabla u \cdot \nabla \psi + \phi'(u)\psi\big)\, dx, \quad u, \psi \in H,$$

is a Lipschitz continuous and strongly monotone operator, too.

2. We continue the proof with a comparison principle: If u, $v \in H$ satisfy

$$\langle Au - Av, \psi \rangle \le 0 \quad \text{for all } \psi \in H,\ \psi \ge 0,$$

then $u \le v$ holds true. Indeed, taking the test function $\psi = (u - v)^\oplus \in H$ we get $u \le v$ due to the estimate

$$\begin{aligned} 0 &\ge \int_\Omega \big(\kappa \nabla(u - v) \cdot \nabla(u - v)^\oplus + \big(\phi'(u) - \phi'(v)\big)(u - v)^\oplus\big)\, dx \\ &= \kappa \int_\Omega |\nabla(u - v)^\oplus|^2\, dx + \int_{\{x \in \Omega:\, u(x) \ge v(x)\}} \big(\phi'(u) - \phi'(v)\big)(u - v)\, dx \\ &\ge \kappa \int_\Omega |\nabla(u - v)^\oplus|^2\, dx + \tilde{\alpha} \int_\Omega |(u - v)^\oplus|^2\, dx, \end{aligned}$$

where $\tilde{\alpha} > 0$ is a monotonicity constant of $\phi' : \mathbb{R} \longrightarrow \mathbb{R}$.

3. Because of the Lipschitz continuity and the strong monotonicity of the operator $A : H \longrightarrow H^*$ the inverse $A^{-1} : H^* \longrightarrow H$ has the same properties, too. To find a solution $(u, \mu) \in H \times \mathbb{R}$ of the constrained problem

$$Au = \mu g - h, \quad \langle g, u \rangle = r, \tag{36}$$

for given $h \in H^*$ and $r \in \mathbb{R}$, we use the properties of A^{-1} to define the Lipschitz continuous and strongly monotone function $a : \mathbb{R} \longrightarrow \mathbb{R}$ by

$$a(\mu) = \big\langle g, A^{-1}(\mu g - h)\big\rangle, \quad \mu \in \mathbb{R}.$$

Hence, the equation $a(\mu) = r$ has a solution $\mu \in \mathbb{R}$. Setting $u = A^{-1}(\mu g - h) \in H$ we have found a solution $(u, \mu) \in H \times \mathbb{R}$ of problem (36) which is in fact unique because of the strong monotonicity of A.

4. Specifying the data of problem (36), we see that there exists a uniquely determined solution $(u, \mu) \in H \times \mathbb{R}$ of the constrained problem

$$\langle Au, \psi \rangle = \int_\Omega (\mu - w)\psi\, dx \quad \text{for all } \psi \in H, \quad \int_\Omega (u - \bar{u})\, dx = 0. \tag{37}$$

To prove estimates for the solution $(u, \mu) \in H \times \mathbb{R}$ of (37) assume that $\mu - \check{w} < \phi'(\bar{u})$. Then there exists an $\varepsilon > 0$ such that $v = \bar{u} - \varepsilon$ still satisfies $\mu - \check{w} < \phi'(v)$ which yields

$$\langle Au, \psi \rangle = \int_\Omega (\mu - w)\psi\, dx \le \int_\Omega \phi'(v)\psi\, dx = \langle Av, \psi \rangle \quad \text{for all } \psi \in H,\ \psi \ge 0.$$

Now, the comparison principle (see Step 2 of the proof) implies $u \le v = \bar{u} - \varepsilon$ which contradicts to the fact $\int_\Omega (u - \bar{u})\, dx = 0$. Hence, we have shown $\mu - \check{w} \ge \phi'(\bar{u})$, and

$\mu - \hat{w} \le \phi'(\bar{u})$ follows by an analogous argument. Using the fact, that $\bar{u} \in (0,1)$ belongs to the interval $[\check{u}, \hat{u}]$ of coincidence between ϕ' and φ', we set

$$\check{\mu} = \check{w} + \varphi'(\bar{u}), \quad \hat{\mu} = \hat{w} + \varphi'(\bar{u}),$$

to get $\check{\mu} \le \mu \le \hat{\mu}$ and $\check{w} - \hat{w} + \varphi'(\bar{u}) \le \mu - w \le \hat{w} - \check{w} + \varphi'(\bar{u})$. Having in mind the definition of $\check{u} \in (0, \bar{u}]$ and $\hat{u} \in [\bar{u}, 1)$ in Step 1 for all $\psi \in H$, $\psi \ge 0$ we arrive at

$$\langle A\check{u}, \psi\rangle = \int_\Omega \varphi'(\check{u})\psi\, dx \le \int_\Omega (\mu - w)\psi\, dx = \langle Au, \psi\rangle,$$

$$\langle A\hat{u}, \psi\rangle = \int_\Omega \varphi'(\hat{u})\psi\, dx \ge \int_\Omega (\mu - w)\psi\, dx = \langle Au, \psi\rangle.$$

The comparison principle implies $\check{u} \le u \le \hat{u}$. Hence, $(u, \mu) \in \mathrm{dom}(\partial\Phi) \times \mathbb{R}$ is not only a solution of the regularized problem (37) but also of the original problem (35) which is uniquely solvable because of the strong monotonicity of $\partial\Phi \subset H \times H^*$. Finally,

$$\tilde{\alpha}\, \|u - \check{u}\|_H^2 \le \langle Au - A\check{u}, u - \check{u}\rangle \le \int_\Omega \big(\varphi'(\hat{u}) - \varphi'(\check{u})\big)(u - \check{u})\, dx \le 2\, \|\hat{w} - \check{w}\|_H \|u - \check{u}\|_H,$$

that means, we get an estimate of the form $\|u - \check{u}\|_H \le \frac{2}{\tilde{\alpha}} \|\hat{w} - \check{w}\|_H$. ☐

Lemma 9 (Analyticity)**.** *Let Assumptions 6, 7 and $r \in \left(0, \frac{1}{2}\right)$ be satisfied. Then Φ is real analytic in every subset U which is open in V and contained in*

$$U(r) = \{u \in V : r \le u \le 1 - r\}.$$

The Fréchet derivative $D\Phi : U \longrightarrow W$ is a real analytic operator, $D^2\Phi(u) \in \mathcal{L}(V; W)$ is an isomorphism for all $u \in U$, and there exists a constant $\gamma > 0$ depending on r such that (16) *holds true.*

Proof. 1. Let $r \in \left(0, \frac{1}{2}\right)$ and $\phi : (0,1) \longrightarrow \mathbb{R}$ be a real analytic function. Because $[r, 1-r]$ is a compact subset of $(0,1)$ this implies Cauchy's inequalities (see [5]):

$$\left|\phi^{(k)}(s)\right| \le c_1 k!\, \delta^{-k} \quad \text{for all } k \in \mathbb{N},\ s \in [r, 1-r],$$

for some constants $c_1 > 0$, $\delta \in \left(0, \frac{r}{2}\right)$ depending on r.

2. Let $U \subset U(r)$ be open in V. If we define $\phi : [0,1] \longrightarrow \mathbb{R}$ by $\phi(s) = \varphi(s) - \frac{\kappa}{2}|s|^2$, $s \in (0,1)$, then we can rewrite Φ as follows:

$$\Phi(u) = \tfrac{\kappa}{2}\langle Ju, u\rangle + \Lambda(u), \quad \Lambda(u) = \textstyle\int_\Omega \phi(u)\, dx \quad \text{for all } u \in \mathrm{dom}(\Phi). \tag{38}$$

Obviously, the functional defined by $u \longmapsto \frac{\kappa}{2}\langle Ju, u\rangle$ and its Fréchet derivative $\kappa J \in \mathcal{L}(V; W)$ are real analytic in V. Remark 5 and Step 1 of the proof yield that the second summand Λ is also Fréchet differentiable on U, and that the derivatives $D\Phi(u) \in W$ and $D^2\Phi(u) \in \mathcal{L}(V; W)$ have the form

$$\big\langle D\Phi(u), v_1\big\rangle = \int_\Omega \big(\kappa \nabla u \cdot \nabla v_1 + \varphi'(u) v_1\big)\, dx,$$

$$\big\langle D^2\Phi(u) v_1, v_2\big\rangle = \int_\Omega \big(\kappa \nabla v_1 \cdot \nabla v_2 + \varphi''(u) v_1 v_2\big)\, dx,$$

for all $u \in U$ and v_1, $v_2 \in V$. In view of Remark 5 this implies that $D^2\Phi(u) \in \mathcal{L}(V; W)$ is an isomorphism for all $u \in U$, and that there exists a constant $\gamma > 0$ depending on r such that (16) is satisfied.

3. It remains to show that $D\Lambda : U \longrightarrow W$ is real analytic. For all $u \in U$ and $k \in \mathbb{N}$ we can define symmetric bounded k-linear forms $B_k(u) \in \mathcal{L}^k(V;W)$ by

$$\big\langle B_k(u)[v_1,\ldots,v_k], v\big\rangle = \int_\Omega \phi^{(k+1)}(u) v_1 \cdots v_k v\, dx, \quad v_1,\ldots,v_k, v \in V,$$

because Remark 5 and Step 1 yield the existence of a constant $c_2 > 0$ depending on r such that $\|B_k(u)\|_{\mathcal{L}^k(V;W)} \le c_2 (k+1)!\, \delta^{-k-1}$ for all $u \in U$, $k \in \mathbb{N}$. Consequently, for all $u \in U$ and $\varrho \in (0,\delta)$ both the power series

$$\sum_{k=1}^\infty \frac{1}{k!} \|B_k(u)\|_{\mathcal{L}^k(V;W)} \|v\|_V^k, \quad D\Lambda(u+v) - D\Lambda(u) = \sum_{k=1}^\infty \frac{1}{k!} B_k(u)[v,\ldots,v],$$

converge uniformly for $v \in V$, $\|v\|_V \le \varrho$. Hence, $D\Lambda : U \longrightarrow W$ is a real analytic operator, which implies the real analyticity of Λ on U (see Remark 4). □

Theorem 10 (Convergence)**.** *Let Assumptions 6, 7, and $\alpha > \beta\tau$ be satisfied for some $\tau \in (0,1]$. If $\bar u \in (0,1)$ is given and if we set*

$$K = \big\{u \in \mathrm{dom}(\Phi) : \textstyle\int_\Omega (u - \bar u)\, dx = 0\big\},$$

then there exist constants $\check u$, $\hat u \in (0,1)$, $\check\mu$, $\hat\mu \in \mathbb{R}$, and $c > 0$ depending only on $\bar u$ and the data of the problem such that for all initial values $u_0 \in K$ the sequence $(u_k, f_k) \subset K \times M$ defined by (11) *converges to a solution $(u^*, \mu^*) \in C \times M$ of the Euler–Lagrange equation* (27) *in the sense of* (28)*, where*

$$C = \big\{u \in K : \check u \le u \le \hat u,\ \|u - \check u\|_H \le c\big\}, \quad M = \big\{\mu g \in H_0^0 : \check\mu \le \mu \le \hat\mu\big\}.$$

Proof. We have shown in Remark 5 that Assumptions 1 and 4 are satisfied. In view of the definition of $\Psi : H \longrightarrow \mathbb{R}$ in Assumption 7 its Fréchet derivative $D\Psi : H \longrightarrow H^*$ has the form

$$\langle D\Psi(v), \psi\rangle = \int_\Omega \varkappa(1-2v)\psi\, dx \quad \text{for all } v, \psi \in H.$$

Clearly, we have $-\varkappa \le \varkappa(1-2v) \le \varkappa$ for all $v \in \mathrm{dom}(\Phi)$. By Assumption 6, Lemma 8, and its proof for all $v \in K$ the solution $(u, \mu g)$ of the Euler–Lagrange equation $\mu g - D\Psi(v) \in \partial\Phi(u)$ belongs to $C \times M$, if we set

$$\varphi'(\check u) = \varphi'(\bar u) - 2\varkappa,\ \varphi'(\hat u) = \varphi'(\bar u) + 2\varkappa,\ \check\mu = \varphi'(\bar u) - \varkappa,\ \hat\mu = \varphi'(\bar u) + \varkappa,\ c = \tfrac{4}{\alpha}\|\varkappa\|_H,$$

which gives Assumption 3. Following Lemma 9 we can choose $r \in \big(0, \frac12\big)$ and a subset $U \subset U(r)$ depending on $\check u$, $\hat u \in (0,1)$ such that $\tau C + (1-\tau)K \subset U$ and Assumption 5 is satisfied. Now, the application of Theorem 7 yields the desired convergence result. □

5. **Phase Separation in Multicomponent Systems.** As mentioned in the introduction, we consider a closed multicomponent system with interacting particles of type $i \in \{0,1,\ldots,m\}$ occupying a bounded domain $\Omega \subset \mathbb{R}^n$ with Lipschitz boundary (see [11, 12]). Due to the exclusion principle we assume $\sum_{i=0}^m u_i = 1$ for the densities $u_0, u_1, \ldots, u_m : \Omega \longrightarrow [0,1]$. Hence, in the following we describe the states of the system by m-component vectors $u = (u_1,\ldots,u_m)$ and $u_0 = 1 - \sum_{i=0}^m u_i$.

In contrast to the classical Cahn–Hilliard theory (see [2]) we consider diffuse interface models and free energy functionals with nonlocal expressions (see [1, 3]). As a straight-forward generalization of the nonlocal phase separation model for binary systems (see [9]) we split the free energy functional $F : H \longrightarrow \mathbb{R} \cup \{+\infty\}$ into the sum of an entropy part $\Phi : H \longrightarrow \mathbb{R} \cup \{+\infty\}$ and a nonlocal interaction part $\Psi : H \longrightarrow \mathbb{R}$ (see [13]).

Assumption 8. To satisfy Assumption 2 we set $H = L^2(\Omega;\mathbb{R}^m)$, $V = L^\infty(\Omega;\mathbb{R}^m)$ and $W = J[V]$; their norms are defined as usual by

$$\|u\|_H^2 = \int_\Omega |u|^2\,dx, \quad u \in H,$$

$$\|u\|_V = \operatorname*{ess\,sup}_{x\in\Omega} |u(x)|, \quad u \in V, \qquad \|f\|_W = \|Rf\|_V, \quad f \in W.$$

We consider the Hilbert sum decomposition $H = H_0 + H_1$ into the closed subspace

$$H_0 = \big\{u \in H : \textstyle\int_\Omega u\,dx = 0\big\},$$

and the m-dimensional subspace $H_1 \subset V$ of constant functions. Then the annihilator $H_0^0 = J[H_1] \subset W$ is the m-dimensional subspace of elements $f = \sum_{i=1}^m \mu_i g_i$ where $\mu \in \mathbb{R}^m$ and the functionals $g_1, \dots, g_m \in W$ are given by $\langle g_i, u\rangle = \int_\Omega u_i\,dx$, $u \in H$, $i \in \{1,\dots,m\}$.

Assumption 9. We consider the simplex $\Sigma = \{z \in \mathbb{R}^m : 0 \le z_0, z_1, \dots, z_m \le 1\}$ and, furthermore, a proper, lower semicontinuous, and strongly convex functional $\varphi : \mathbb{R}^m \longrightarrow \mathbb{R} \cup \{+\infty\}$ with effective domain $\operatorname{dom}(\varphi) = \Sigma$, that means, for all y, $z \in \Sigma$, $t \in [0,1]$ and some $\alpha > 0$ we have

$$t\varphi(y) + (1-t)\varphi(z) \ge \varphi(ty + (1-t)z) + \tfrac{\alpha}{2} t(1-t)|y-z|^2.$$

Moreover, we assume that φ is real analytic in $\operatorname{int}\Sigma$ and $\operatorname{dom}(\partial\varphi) = \operatorname{int}\Sigma$ holds true for its subdifferential. Now, we introduce a lower semicontinuous and strongly convex functional $\Phi : H \longrightarrow \mathbb{R} \cup \{+\infty\}$ by

$$\Phi(u) = \begin{cases} \int_\Omega \varphi(u)\,dx & \text{if } u \in H,\ 0 \le u_0, u_1, \dots, u_m \le 1, \\ +\infty & \text{otherwise,} \end{cases} \tag{39}$$

with closed effective domain $\operatorname{dom}(\Phi) = \{u \in H : 0 \le u_0, u_1, \dots, u_m \le 1\}$.

Remark 6. As a consequence, the convex conjugate $\varphi^* : \mathbb{R}^m \longrightarrow \mathbb{R}$ of φ is Fréchet differentiable with the derivative $D\varphi^* : \mathbb{R}^m \longrightarrow \operatorname{int}\Sigma$. Moreover, the Young–Fenchel inequality yields $\varphi^*(\xi) + \varphi(z) \ge \xi \cdot z$ for all $\xi \in \mathbb{R}^m$, $z \in \Sigma$. Hence, considering the extremal points z of the simplex Σ this implies

$$\varphi^*(\xi) + \sup_{z\in\Sigma} \varphi(z) \ge \sup_{1\le i\le m} \xi_i^{\oplus} \quad \text{for all } \xi \in \mathbb{R}^m. \tag{40}$$

The conjugate functional $\Phi^* : H^* \longrightarrow \mathbb{R}$ has the form $\Phi^*(h) = \int_\Omega \varphi^*(Rh)\,dx$ for all $h \in H^*$. The Fenchel–Moreau theorem implies that $(u,h) \in \partial\Phi$ if and only if $(h,u) \in \partial\Phi^*$.

Lemma 11 (Uniform boundedness)**.** *Assumptions 8 and 9 imply the following:*

(i) *For all $\bar u \in \operatorname{int}\Sigma$ and $w \in V$ there exists a uniquely determined solution $(u,\mu) \in \operatorname{dom}(\partial\Phi) \times \mathbb{R}^m$ of the constrained problem*

$$\langle \partial\Phi(u), v\rangle = \int_\Omega (\mu - w)\cdot v\,dx \quad \textit{for all } v \in H, \qquad \int_\Omega (u - \bar u)\,dx = 0. \tag{41}$$

(ii) *Let $\check w_i \le w_i \le \hat w_i$ for all $i \in \{1,\dots,m\}$ and some $\check w$, $\hat w \in \mathbb{R}^m$. Then there exist $\check u$, $\hat u \in \operatorname{int}\Sigma$ and $\check\mu$, $\hat\mu \in \mathbb{R}^m$ depending only on $\check w$, $\hat w$, $\bar u$ such that the solution (u,μ) satisfies*

$$\check u_i \le u_i \le \hat u_i \quad \textit{for all } i \in \{0,1,\dots,m\}, \qquad \check\mu_i \le \mu_i \le \hat\mu_i \quad \textit{for all } i \in \{1,\dots,m\}.$$

Proof. 1. Set $c = \sup_{z\in\Sigma}\varphi(z) \in \mathbb{R}$ and let $\check{w}_i \le w_i \le \hat{w}_i$ hold true for all $i \in \{1,\dots,m\}$ and some $\check{w}$, $\hat{w} \in \mathbb{R}^m$. Following Remark 6 for fixed $\bar{u} \in \operatorname{int}\Sigma$ and $w \in V$ we can define a Fréchet differentiable convex function $\Lambda : \mathbb{R}^m \longrightarrow \mathbb{R}$ by

$$\Lambda(\lambda) = \int_\Omega \big(\varphi^*(\lambda - w) + c - \bar{u}\cdot(\lambda - w)\big)\,dx, \quad \lambda \in \mathbb{R}^m. \tag{42}$$

In view of (40) for all $\lambda \in \mathbb{R}^m$ we can find the following estimate:

$$\begin{aligned}\Lambda(\lambda) &= \int_\Omega \big(\varphi^*(\lambda - w) + c - \bar{u}\cdot(\lambda - w)\big)\,dx \\ &\ge \int_\Omega \bigg(\sup_{1\le i\le m}(\lambda_i - w_i)^\oplus - \sum_{i=1}^m \bar{u}_i(\lambda_i - w_i)^\oplus\bigg)\,dx + \int_\Omega \sum_{i=1}^m \bar{u}_i(\lambda_i - w_i)^\ominus\,dx \\ &\ge \int_\Omega \bigg(1 - \sum_{i=1}^m \bar{u}_i\bigg)\sup_{1\le i\le m}(\lambda_i - \hat{w}_i)^\oplus\,dx + \int_\Omega \sum_{i=1}^m \bar{u}_i(\lambda_i - \check{w}_i)^\ominus\,dx.\end{aligned}$$

Because of $\bar{u} \in \operatorname{int}\Sigma$ this yields $\lim_{|\lambda|\to+\infty}\Lambda(\lambda) = +\infty$, that means, $\Lambda : \mathbb{R}^m \longrightarrow \mathbb{R}$ is weakly coercive. Hence, Λ attains its minimum which implies the existence of some $\mu \in \mathbb{R}^m$ such that $D\Lambda(\mu) = 0$. Therefore, using (42) and setting

$$f = \sum_{i=1}^m \mu_i g_i \in H_0^0, \quad u = \partial\Phi^*(f - Jw) \in \operatorname{dom}(\partial\Phi),$$

we get $\int_\Omega (u - \bar{u})\,dx = 0$. Therefore, $(u,\mu) \in \operatorname{dom}(\partial\Phi)\times\mathbb{R}^m$ is a solution of (41) which is uniquely determined because of the strong monotonicity of $\partial\Phi \subset H \times H^*$.

2. Due to the definition of $(u,\mu) \in \operatorname{dom}(\partial\Phi)\times\mathbb{R}^m$ in Step 1 of the proof we can apply the Fenchel–Moreau theorem to get the relations

$$\int_\Omega \big(\varphi^*(\mu - w) + \varphi(u)\big)\,dx = \int_\Omega u\cdot(\mu - w)\,dx,$$
$$\int_\Omega \big(\varphi^*(-w) + \varphi(u)\big)\,dx \ge -\int_\Omega u\cdot w\,dx.$$

Together with (41) and (42) this yields

$$\Lambda(\mu) = \int_\Omega \big(c + (u - \bar{u})\cdot(\mu - w) - \varphi(u)\big)\,dx \le \int_\Omega \big(c + \varphi^*(-w) + \bar{u}\cdot w\big)\,dx.$$

The weak coercivity estimate for Λ (see Step 1) implies the existence of $\check{\mu}$, $\hat{\mu} \in \mathbb{R}^m$ depending only on $\check{w}$, $\hat{w}$, $\bar{u}$ such that $\check{\mu}_i \le \mu_i \le \hat{\mu}_i$ for all $i \in \{1,\dots,m\}$.

Furthermore, the image of $\big\{\lambda \in \mathbb{R}^m : \check{\mu}_i - \hat{w}_i \le \lambda_i \le \hat{\mu}_i - \check{w}_i,\ i \in \{1,\dots,m\}\big\}$ under $D\varphi^* : \mathbb{R}^m \longrightarrow \operatorname{int}\Sigma$ is a compact subset of $\operatorname{int}\Sigma$. Hence, we can find some $\check{u}$, $\hat{u} \in \operatorname{int}\Sigma$ depending on $\check{w}$, $\hat{w}$, $\bar{u}$ such that $\check{u}_i \le u_i \le \hat{u}_i$ for all $i \in \{0,1,\dots,m\}$. □

Lemma 12 (Analyticity). *If Assumptions 8 and 9 are satisfied, then Φ is real analytic in every subset U which is open in V and contained in*

$$U(r) = \{u \in V : r \le u_0, u_1, \dots, u_m \le 1 - r\}$$

for some $r \in \big(0, \frac{1}{m}\big)$. Moreover, the Fréchet derivative $D\Phi : U \longrightarrow W$ is a real analytic operator, the second derivative $D^2\Phi(u) \in \mathcal{L}(V;W)$ is an isomorphism for all $u \in U$, and (16) holds true for some constant $\gamma > 0$ depending on r.

Proof. 1. Let $r \in \big(0, \frac{1}{m}\big)$ be arbitrarily fixed and consider the compact subset

$$\Sigma(r) = \{z \in \mathbb{R}^m : r \le z_0, z_1, \dots, z_m \le 1 - r\}$$

of $\operatorname{int}\Sigma$. By the real analyticity of φ in $\operatorname{int}\Sigma$ we can find constants $c_1 > 0$, $\delta \in \left(0, \frac{r}{2}\right)$ depending on r such that for all $k \in \mathbb{N}$, $z \in \Sigma(r)$, $\zeta_1, \dots, \zeta_k \in \mathbb{R}^m$, Cauchy's inequalities (see [5]) hold true:

$$\left|D^k\varphi(z)\zeta_1 \cdots \zeta_k\right| \le c_1 k!\, \delta^{-k}\, |\zeta_1| \cdots |\zeta_k|.$$

2. Let $U \subset U(r)$ be open in V. Due to Step 1 of the proof Φ is Fréchet differentiable on U; the derivatives $D\Phi(u) \in W$, $D^2\Phi(u) \in \mathcal{L}(V;W)$ have the form

$$\langle D\Phi(u), v_1\rangle = \int_\Omega D\varphi(u) \cdot v_1\, dx,$$

$$\langle D^2\Phi(u)v_1, v_2\rangle = \int_\Omega D^2\varphi(u)v_1 \cdot v_2\, dx,$$

for all $u \in U$ and v_1, $v_2 \in V$. Together with Assumption 9 this yields that $D^2\Phi(u) \in \mathcal{L}(V;W)$ is an isomorphism for all $u \in U$, and that there exists a constant $\gamma > 0$ depending on r such that (16) is satisfied.

3. Moreover, for all $u \in U$ and $k \in \mathbb{N}$ we can define symmetric bounded k-linear forms $B_k(u) \in \mathcal{L}^k(V;W)$ by

$$\left\langle B_k(u)[v_1, \dots, v_k], v\right\rangle = \int_\Omega D^{k+1}\varphi(u)v_1 \cdots v_k \cdot v\, dx, \quad v_1, \dots, v_k, v \in V,$$

because Step 1 yields the estimate $\|B_k(u)\|_{\mathcal{L}^k(V;W)} \le c_2(k+1)!\,\delta^{-k-1}$ for all $u \in U$, $k \in \mathbb{N}$ and some $c_2 > 0$. Hence, for all $u \in U$ and $\varrho \in (0, \delta)$ both the power series

$$\sum_{k=1}^{\infty} \frac{1}{k!}\|B_k(u)\|_{\mathcal{L}^k(V;W)}\|v\|_V^k, \quad D\Phi(u+v) - D\Phi(u) = \sum_{k=1}^{\infty}\frac{1}{k!}B_k(u)[v, \dots, v],$$

converge uniformly for $v \in V$, $\|v\|_V \le \varrho$. Consequently, $D\Phi : U \longrightarrow W$ is a real analytic operator which implies the real analyticity of Φ on U (see Remark 4). ☐

Theorem 13 (Convergence)**.** *Let Assumptions 4, 8, 9, and $\alpha > \beta\tau$ be satisfied for some $\tau \in (0,1]$. If we take $\bar{u} \in \operatorname{int}\Sigma$ and*

$$K = \left\{u \in \operatorname{dom}(\Phi) : \textstyle\int_\Omega (u - \bar{u})\, dx = 0\right\},$$

then there exist constants $\check{u}$, $\hat{u} \in \operatorname{int}\Sigma$, $\check{\mu}$, $\hat{\mu} \in \mathbb{R}^m$ depending only on $\bar{u}$ and the data of the problem such that for all initial values $u_0 \in K$ the sequence $(u_k, f_k) \subset K \times M$ defined by (11) *converges to a solution $(u^*, f^*) \in C \times M$ of the Euler–Lagrange equation* (27) *in the sense of* (28)*, where*

$$C = \left\{u \in K : \check{u}_i \le u_i \le \hat{u}_i,\ i \in \{0, 1, \dots, m\}\right\},$$
$$M = \left\{\textstyle\sum_{i=1}^m \mu_i g_i \in H_0^0 : \check{\mu}_i \le \mu_i \le \hat{\mu}_i,\ i \in \{1, \dots, m\}\right\}.$$

Proof. 1. In view of Assumptions 4, 8, and 9 also Assumption 1 is satisfied and $F = \Phi + \Psi : H \longrightarrow \mathbb{R} \cup \{+\infty\}$ is a well-defined functional with nonempty, closed, and convex effective domain $\operatorname{dom}(F) = \operatorname{dom}(\Phi) \subset V$.

2. By virtue of Assumption 4 we have $D\Psi(v) = Tv + l$ for all $v \in H$, and we can find constants $\check{w}$, $\hat{w} \in \mathbb{R}^m$ such that

$$\check{w}_i \le (RTv)_i + (Rl)_i \le \hat{w}_i \quad \text{for all } i \in \{1, \dots, m\},\ v \in \operatorname{dom}(\Phi).$$

Now, Assumption 8 and Lemma 11 yield that the solution $\left(u, \sum_{i=1}^m \mu_i g_i\right)$ of the Euler–Lagrange equation $\sum_{i=1}^m \mu_i g_i - D\Psi(v) \in \partial\Phi(u)$ belongs to $C \times M$ for all $v \in K$, if we take the constants $\check{u}$, $\hat{u} \in \operatorname{int}\Sigma$, $\check{\mu}$, $\hat{\mu} \in \mathbb{R}^m$ as in the proof of Lemma 11. This is the contents of Assumption 3. Because of Lemma 12 we can choose $r \in \left(0, \frac{1}{m}\right)$ and a subset $U \subset U(r)$ depending on $\check{u}$, $\hat{u} \in \operatorname{int}\Sigma$ such that

$\tau C + (1-\tau)K \subset U$ and Assumption 5 is satisfied. Summing up we can apply Theorem 7 to get the desired convergence result. □

6. **The Image Segmentation Algorithm.**

Various approaches to local image segmentation have been introduced in the literature (see [16]). In contrast to these methods we want to establish a nonlocal image segmentation algorithm based on the descent method. To do so, let all the assumptions of the previous section be satisfied. We consider functions $c \in L^\infty(\Omega)$, $0 \le c \le 1$ representing (normalized) gray scaled images. To segment c with respect to given gray levels

$$a_0, a_1, \ldots, a_m \in [0,1], \quad 0 = a_0 < a_1 < \cdots < a_m = 1,$$

we introduce the following algorithm:

Step 1 (Decomposition into phases)**.** We transform c into an m-component distribution $u_0 = (u_{01}, \ldots, u_{0m}) \in K$ such that the i-th component corresponds to the level $a_i \in [0,1]$:

$$0 \le u_{00}, u_{01}, \ldots, u_{0m} \le 1, \quad u_{00} = 1 - \sum_{i=1}^m u_{0i}.$$

To that end, we consider a continuous partition of unity $(\eta_0, \ldots, \eta_m) \subset C([0,1])$ with weights b_0, $b_1, \ldots, b_m \in \mathbb{R}$ such that

$$0 \le \eta_i \le 1, \quad a_i \in \operatorname{supp}(\eta_i), \quad \sum_{i=0}^m \eta_i = 1, \quad i \in \{0,1,\ldots,m\}, \tag{43}$$

$$0 < b_i < 1, \quad \sum_{i=0}^m b_i = 1, \quad \int_0^1 \eta_i(s)\,ds = b_i, \quad i \in \{0,1,\ldots,m\}. \tag{44}$$

Now, we are ready to define the transformation

$$c \longmapsto u_0 = (u_{01}, \ldots, u_{0m}) = (\eta_1(c), \ldots, \eta_m(c)) \in K.$$

Step 2 (Nonlocal phase separation)**.** Given $u_0 \in K$, we solve the nonlocal phase separation problem (27) for the m-component system to find the corresponding critical point $u^* \in C$ of the energy functional $F : H \longrightarrow \mathbb{R} \cup \{+\infty\}$.

Step 3 (Composition of segmented phases)**.** Finally, we calculate the segmented version $c^* \in L^\infty(\Omega)$, $0 \le c^* \le 1$ of $c \in L^\infty(\Omega)$ as a convex combination of the levels $a_0, a_1, \ldots, a_m \in [0,1]$ with respect to the weight functions u_0^*, $u_1^*, \ldots, u_m^*$, that means,

$$u^* \longmapsto c^* = \sum_{i=0}^m u_i^* a_i.$$

Before we present our simulation results we choose a partition of unity and a special class of segmentation entropy and nonlocal interaction energy functionals.

Example 1 (Partition of unity)**.** To construct a partition of unity we choose numbers b_0, $b_1, \ldots, b_m \in (0,1)$ such that

$$\sum_{j=0}^m b_j = 1, \quad b_i^* = \sum_{j=0}^{i-1} b_j \in (a_{i-1}, a_i), \quad i \in \{1, \ldots, m\}.$$

For $i \in \{1, \ldots, m\}$ we define exponents $\omega_i > 0$ and functions $h_i \in C([a_{i-1}, a_i])$ by

$$\omega_i = \frac{a_i - b_i^*}{b_i^* - a_{i-1}}, \quad h_i(s) = \left(\frac{a_i - s}{a_i - a_{i-1}}\right)^{\omega_i}, \quad s \in [a_{i-1}, a_i].$$

Now, we get a continuous partition of unity $(\eta_0, \dots, \eta_m) \subset C([0,1])$ with the properties (43) and (44) by setting

$$\eta_0(s) = \begin{cases} h_1(s) & \text{if } s \in [a_0, a_1], \\ 0 & \text{otherwise}, \end{cases} \qquad \eta_m(s) = \begin{cases} 1 - h_m(s) & \text{if } s \in [a_{m-1}, a_m], \\ 0 & \text{otherwise}, \end{cases}$$

for the boundary cases $i = 0$ and $i = m$ and

$$\eta_i(s) = \begin{cases} 1 - h_i(s) & \text{if } s \in [a_{i-1}, a_i], \\ h_{i+1}(s) & \text{if } s \in [a_i, a_{i+1}], \\ 0 & \text{otherwise}, \end{cases}$$

for $i \in \{1, \dots, m-1\}$, respectively.

Example 2 (Segmentation entropy)**.** According to (39) we specify the lower continuous and strongly convex function $\varphi : \mathbb{R}^m \longrightarrow \mathbb{R} \cup \{+\infty\}$ as logarithmic potential

$$\varphi(z) = \begin{cases} \sum_{i=0}^m z_i \log(z_i) & \text{if } z \in \Sigma, \\ +\infty & \text{otherwise}. \end{cases}$$

Obviously, $\operatorname{dom}(\varphi) = \Sigma$ and φ is real analytic in $\operatorname{dom}(\partial\varphi) = \operatorname{int}\Sigma$ with partial derivatives $D_i\varphi(z) = \log(z_i) - \log(z_0)$ for all $z \in \operatorname{int}\Sigma$ and $i \in \{1, \dots, m\}$. We can interpret the value $\Phi(u) = \int_\Omega \varphi(u)\,dx$ as the segmentation entropy of the state $u \in \operatorname{dom}(\Phi)$.

Example 3 (Nonlocal interaction energy)**.** In the following we describe the nonlocal interaction by means of inverse operators corresponding to second order elliptic operators with appropriate regularity properties. To do so, for $r > 0$ we consider the family of elliptic operators $E_r \in \mathcal{L}\big(H^1(\Omega); H^1(\Omega)^*\big)$ (including Neumann boundary conditions) given by

$$\langle E_r v, h\rangle = \int_\Omega \big(r^2 \nabla v \cdot \nabla h + vh\big)\,dx, \quad v, h \in H^1(\Omega).$$

We want to emphasize that the inverse operators $E_r^{-1} \in \mathcal{L}\big(H^1(\Omega)^*; H^1(\Omega)\big)$ are completely continuous from $L^2(\Omega)$ into $L^2(\Omega)$ as well as from $L^\infty(\Omega)$ into $L^\infty(\Omega)$ (see Remark 5).

To control the qualitative behaviour of nonlocal interaction we prescribe effective ranges ϱ, $r > 0$ and intensities $\sigma_{i\ell}$, $s_{i\ell} \in \mathbb{R}$ of interaction forces between particles of type i and $\ell \in \{0, 1, \dots, m\}$, respectively. Clearly, both matrices are assumed to be symmetric. The cases $\sigma_{i\ell} > 0$ and $\sigma_{i\ell} < 0$ represent repulsive and attractive interaction, respectively. According to Assumption 4 we define the quadratic functional $\Psi : H \longrightarrow \mathbb{R}$ for $u \in H$ by

$$\Psi(u) = \frac{1}{2}\sum_{i=0}^m \sum_{\ell=0}^m \int_\Omega u_i \sigma_{i\ell} E_\varrho^{-1} u_\ell\,dx + \frac{1}{2}\sum_{i=0}^m \sum_{\ell=0}^m \int_\Omega (u_i - \tilde{u}_i) s_{i\ell} E_r^{-1}(u_\ell - \tilde{u}_\ell)\,dx. \tag{45}$$

Note, that by choosing the matrix $(s_{i\ell})$ appropriately, it is possible to get final states $u^* \in C$ close to some prescribed state $\tilde{u} \in K$.

7. **Simulation Results for Ternary Systems.**

We apply our image segmentation algorithm to different situations in image processing. For simplicity, in all of

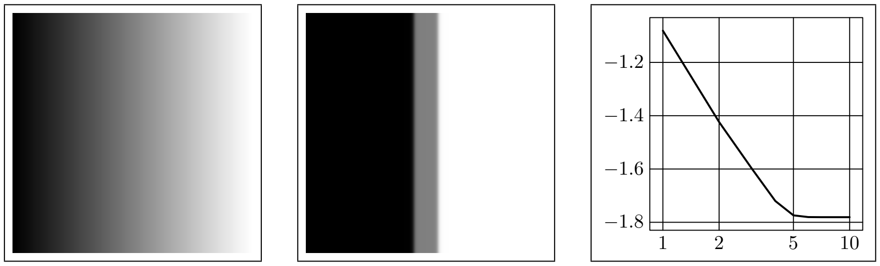

FIGURE 1. Phase separation: (a) initial value (constant in vertical direction); (b) final state (stripe pattern); (c) decay to the corresponding local minimum of the free energy functional F under the constraint of mass conservation after 10 iteration steps.

our following examples we consider ternary systems of three colored components $(m = 2)$ and (3×3)-matrices of the form

$$(\sigma_{i\ell}) = \begin{pmatrix} -\sigma & +\sigma & +\sigma \\ +\sigma & -\sigma & +\sigma \\ +\sigma & +\sigma & -\sigma \end{pmatrix}, \quad (s_{i\ell}) = \begin{pmatrix} +s & -s & -s \\ -s & +s & -s \\ -s & -s & +s \end{pmatrix}. \tag{46}$$

From the structure of $(\sigma_{i\ell})$ it follows, that particles of the same type attract and particles of different type repel each other with the same range $\varrho > 0$ and intensity $\sigma > 0$ of interaction.

In a first example we consider the case of pure phase separation $(s = 0)$ without stabilization of the initial value. The other examples deal with the segmentation of a perfect image and the reconstruction of a noisy image, respectively. Here, the nontrivial choice of $(s_{i\ell})$, $s > 0$ and $\tilde{u} = u_0$ enables us to get final states close to the corresponding initial values $u_0 \in K$ (see (45)).

Remark 7. Naturally, planar images are represented by bounded rectangular domains $\Omega \subset \mathbb{R}^2$. The ranges of interaction are given in the natural length unit of the problem, that means, the edge length of one (square) pixel. Of course, our method can be applied also to voxel images defined in a domain $\Omega \subset \mathbb{R}^n$ of arbitrary space dimension $n \in \mathbb{N}$.

Example 4 (Phase separation)**.** We separate two gray scaled images with respect to three equally weighted gray levels,

$$a_0 = 0,\ a_1 = \frac{1}{2},\ a_2 = 1, \quad b_0 = b_1 = b_2 = \frac{1}{3},$$

and interaction parameters (according to (46))

$$\varrho = 5,\ \sigma = 4, \quad r = 5,\ s = 0.$$

Figures 1 and 2 show simulation results for two very similar 256 by 256 pixel images. Both initial configurations contain equal numbers of black, white, and medium gray particles, respectively. Obviously, the final states do not depend only on these integral quantities.

Example 5 (Image segmentation)**.** We segment the well-known *Lena image* with respect to three equally weighted gray levels,

$$a_0 = 0,\ a_1 = \frac{1}{2},\ a_2 = 1, \quad b_0 = b_1 = b_2 = \frac{1}{3},$$

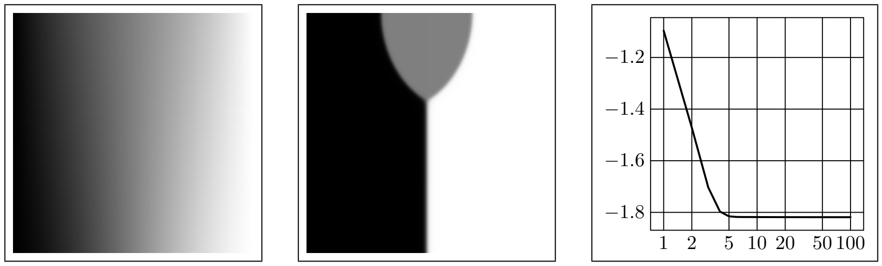

FIGURE 2. Phase separation: (a) mirror-symmetric initial value; (b) final state (phases separated by two arcs and a straight line); (c) decay to the global minimum of the free energy functional F under the constraint of mass conservation after 100 iteration steps.

FIGURE 3. Image segmentation applied to the *Lena image*: (a) initial value (original image); (b) final state of the two-component black and white segmentation; (c) final state of the three-component segmentation.

and interaction parameters (see (46))

$$\varrho = 1,\, \sigma = 4, \quad r = 1,\, s = \frac{24}{5}.$$

In Figure 3 we present simulation results for the 256 by 256 pixel *Lena image*. Here, we compare the above mentioned three-component case with a two-component black and white segmentation (with similar parameters).

Example 6 (Image reconstruction)**.** Finally, we reconstruct a noisy image with respect to three weighted gray levels

$$a_0 = 0,\, a_1 = \frac{49}{100},\, a_2 = 1, \quad b_0 = \frac{39}{100},\, b_1 = \frac{22}{100},\, b_2 = \frac{39}{100},$$

and interaction parameters (according to (46))

$$\varrho = 2,\, \sigma = 10, \quad r = 2,\, s = 12.$$

Figure 4 shows numerical results for a noisy 200 by 200 pixel image. The advantage of the three-component case compared with the two-component black and white reconstruction (with similar parameters) is obvious.

REFERENCES

[1] Bates, P.W., Chmaj, A.: An integrodifferential model for phase transitions: Stationary solutions in higher space dimensions. *J. Statist. Phys. 95*, 1119–1139 (1999).

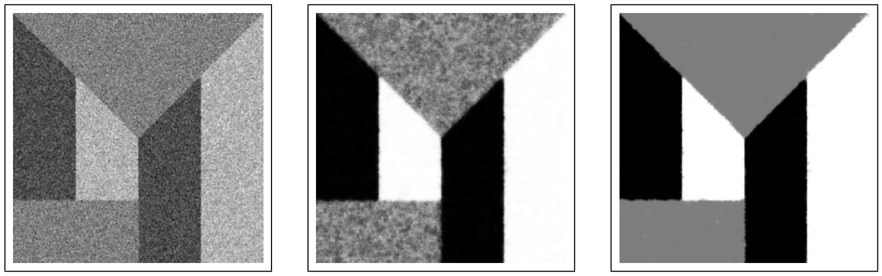

FIGURE 4. Image reconstruction: (a) initial value (noisy image); (b) final state of the two-component black and white reconstruction (gray region still noisy); (c) final state of the three-component reconstruction.

[2] Cahn, J.W., Hilliard, J.E.: Free energy of a nonuniform system. I. Interfacial free energy. *J. Chem. Phys. 28*, 258–267 (1958).

[3] Chen, C.-K., Fife, P.C.: Nonlocal models of phase transitions in solids. *Adv. Math. Sci. Appl. 10*, 821–849 (2000).

[4] Chill, R.: On the Łojasiewicz–Simon gradient inequality. *J. Funct. Anal. 201*, 572–601 (2003).

[5] Dieudonné, J.: *Grundzüge der modernen Analysis, Band 1.* Dritte Auflage. Braunschweig, Wiesbaden: Vieweg & Sohn, 1985.

[6] Feireisl, E., Issard-Roch, F., Petzeltova, H.: Long-time behaviour and convergence towards equilibria for a conserved phase field model. *Discrete Contin. Dynam. Systems 10*, 239–252 (2004).

[7] Feireisl, E., Issard-Roch, F., Petzeltova, H.: A non-smooth version of the Łojasiewicz–Simon theorem with applications to non-local phase-field systems. *J. Differential Equations 199*, 1–21 (2004).

[8] Giacomin, G., Lebowitz, J.L., Marra, R.: Macroscopic evolution of particle systems with short- and long-range interactions. *Nonlinearity 13*, 2143–2162 (2000).

[9] Gajewski, H., Zacharias, K.: On a nonlocal phase separation model. *J. Math. Anal. Appl. 286*, 11–31 (2003).

[10] Gajewski, H., Gärtner, K.: On a nonlocal model of image segmentation. *Z. Angew. Math. Phys. 56*, 572–591 (2005).

[11] Giusti, E.: *Metodi diretti nel calcolo delle variazioni.* Bologna: Unione Matematica Italiana, 1994.

[12] Griepentrog, J.A.: Linear elliptic boundary value problems with non-smooth data: Campanato spaces of functionals. *Math. Nachr. 243*, 19–42 (2002).

[13] Griepentrog, J.A.: On the unique solvability of a nonlocal phase separation problem for multicomponent systems. *Banach Center Publ. 66*, 153–164 (2004).

[14] Łojasiewicz, S.: Une propriété topologique des sous-ensembles analytiques réels. *Colloques internationaux du C.N.R.S.: Les équations aux dérivées partielles 117.* Paris: Editions du C.N.R.S., 87–89 (1963).

[15] Miranville, A., Rougirel, A.: Local and asymptotic analysis of the flow generated by the Cahn–Hilliard–Gurtin equation. *Preprint*, Université de Poitiers, 1–24 (2004).

[16] Morel, J.-M., Solimini, S.: *Variational Methods in Image Segmentation.* Boston, Basel, Berlin: Birkhäuser, 1994.

[17] Wainberg, M.M., Trenogin, W.A.: *Theorie der Lösungsverzweigung bei nichtlinearen Gleichungen.* Berlin: Akademie-Verlag, 1973.

Received March 2005; 1st revision June 2005; 2nd revision December 2005.

E-mail address: gajewski@wias-berlin.de
E-mail address: griepent@wias-berlin.de