technische universität
dortmund

On gradient flows of nonconvex functionals in Hilbert spaces with Riemannian metric and application to Cahn-Hilliard equations

Martin Heida

Preprint 2012-19 November 2012

Fakultät für Mathematik
Technische Universität Dortmund
Vogelpothsweg 87
44227 Dortmund

tu-dortmund.de/MathPreprints

ON GRADIENT FLOWS OF NONCONVEX FUNCTIONALS IN HILBERT SPACES WITH RIEMANNIAN METRIC AND APPLICATION TO CAHN-HILLIARD EQUATIONS

MARTIN HEIDA

Abstract. In Hilbert spaces with a densely defined Riemannian metric, we study gradient flows (curves of maximal slope) of the form

$$\partial_t u + \nabla_l \mathcal{S}(u) \ni f$$

where $\mathcal{S}$ is a nonconvex functional, $\nabla_l \mathcal{S}(u)$ is the strong-weak closure of the subgradient of $\mathcal{S}$ and f is a time dependent right hand side. The article generalizes the results by Rossi and Savaré to this setting and provides some examples from multiphase systems. In particular, we treat Allen-Cahn- and Cahn-Hilliard equations with mobility depending nonlinear on the concentration and its gradient. We also study systems of multiple phases derived by Heida, Málek and Rajagopal [20, 19] in a simplified form. In particular, we will show that a certain class of reaction-diffusion equations coming from a modeling approach by Rajagopal and Srinivasa [27] are automatically subject to the theory of curves of maximal slope.

1. Introduction

In this work, we treat gradient flow equations of the form

$$\partial_t u \in -\nabla_{l,u} \mathcal{S}(u) + f(t) \tag{1.1}$$

with $\mathcal{S}$ being a (possible nonconvex) lower semicontinuous entropy functional on a Hilbert space $\mathcal{H}$, $\nabla_{l,u}\mathcal{S}$ being the *limiting subgradient* with respect to a *densely defined metric structure* $g_\bullet$ and $f \in L^2(0,T;\mathcal{H})$. To this aim, we will work in between the approaches by Rossi and Savaré [30] (in Hilbert spaces) and Ambrosio, Gigli, Savaré [2] (complete metric spaces), where we will stay more closely related to the Hilbert structure of [30]. The main results generalize three existence results in [30] to a further degree of nonlinearity. We will state the main results below as theorems 1.7-1.9. But first we need to introduce the major concepts and notations. After formulating the main results we will shortly discuss the necessity of these generalizations and compare to literature.

Notations and Concepts. Consider Hilbert spaces $\mathcal{H}_0 \hookrightarrow \tilde{\mathcal{H}} \hookrightarrow \mathcal{H}$ with the set $B(\mathcal{H})$ of positive definite continuous bilinear forms. We then use the following terms and notations:

Definition 1.1. We call any tuple $(\mathcal{H}_0, \tilde{\mathcal{H}}, \mathcal{H}, g)$ of Hilbert spaces $\mathcal{H}_0$, $\tilde{\mathcal{H}}$, $\mathcal{H}$ and a mapping $g_\bullet : \tilde{\mathcal{H}} \to B(\mathcal{H})$ satisfying 1 and 2 an entropy space:

(1) $\mathcal{H}_0 \hookrightarrow \tilde{\mathcal{H}} \hookrightarrow \mathcal{H}$, where the imbeddings are dense, and the imbedding $\mathcal{H}_0 \hookrightarrow \tilde{\mathcal{H}}$ is compact. We denote $\|\cdot\|_{\mathcal{H}}$, $\|\cdot\|_{\tilde{\mathcal{H}}}$, $\|\cdot\|_{\mathcal{H}_0}$ the respective norms and by $\langle\cdot,\cdot\rangle_{\mathcal{H}}$ the scalar product on $\mathcal{H}$.
(2) g is a densely defined metric in the following sense: There are positive constants $1 \le G^* < +\infty$ such that

$$\sqrt{G^*}^{-1} \left|\langle x,y\rangle_{\mathcal{H}}\right| \le |g_u(x,y)| \le \sqrt{G^*} \left|\langle x,y\rangle_{\mathcal{H}}\right| \quad \forall u \in \tilde{\mathcal{H}}, \quad \forall x,y \in \mathcal{H}\,, \tag{1.2}$$

and $g_\bullet$ is strong-weak-continuous in the following sense: if $u_n \to u$ strongly in $\tilde{\mathcal{H}}$ and $\varphi_n \rightharpoonup \varphi$ weakly in $\mathcal{H}$ as $n \to \infty$, then

$$g_{u_n}(\varphi_n, \psi) \to g_u(\varphi, \psi) \quad \text{as } n \to \infty \qquad \forall \psi \in \mathcal{H}\,. \tag{1.3}$$

Date: 29.7.2012.
2000 *Mathematics Subject Classification.* 35D30, 35K57,47J35, 80A220.

Formally, we write $g_u(\phi,\psi)=+\infty$ for all $\phi,\psi\in\mathcal{H}$ whenever $u\notin\tilde{\mathcal{H}}$. This means that to every point $u\in\tilde{\mathcal{H}}$ we associate a local scalar product and local norm

$$\langle x,y\rangle_{g(u)}:=g_u(x,y)\,,\qquad \|x\|_{g(u)}:=\sqrt{g_u(x,x)}\qquad\forall x,y\in\mathcal{H}\,.$$

We denote by $\tilde{g}_u$ the unique automorphism on $\mathcal{H}$ such that

$$g_u(v,\varphi)=\langle\tilde{g}_u(v),\varphi\rangle_{\mathcal{H}}\qquad\forall v,\varphi\in\mathcal{H}\,. \tag{1.4}$$

Furthermore, we will assume that $\mathcal{S}:\,\mathcal{H}\to(-\infty,+\infty]$ is a proper functional. Then, we define the set valued subdifferential $d\mathcal{S}(u)$ at $u\in D(\mathcal{S})$ through

$$\delta\in d\mathcal{S}(u)\quad\Leftrightarrow\quad\langle\delta,v\rangle_{\mathcal{H}}\leq\liminf_{h\searrow0}\frac{\mathcal{S}(u+hv)-\mathcal{S}(u)}{h}\quad\forall v\in\mathcal{H} \tag{1.5}$$

and the subgradient $\nabla_u\mathcal{S}(u)$ of $\mathcal{S}$ in $u\in\mathcal{H}$ through

$$\delta\in\nabla_u\mathcal{S}(u)\quad\Leftrightarrow\quad\exists\tilde{\delta}\in d\mathcal{S}(u)\,:\,g_u\left(\delta,v\right):=\left\langle\tilde{\delta},v\right\rangle_{\mathcal{H}}\qquad\forall v\in\mathcal{H}\,, \tag{1.6}$$

where the index u refers to the local metric. To make this notation more clear, note that more generally, equivalent definitions of $\nabla_v\mathcal{S}(u)$ are given through

$$\tilde{g}_v\left(\nabla_v\mathcal{S}(u)\right)=d\mathcal{S}(u)\,,\quad v\in\tilde{\mathcal{H}},\,u\in D(d\mathcal{S}) \tag{1.7}$$

or the condition that for all curves $\gamma:[-1,1]\to\mathcal{H}$ with $\gamma(0)=u$ holds

$$\delta\in\nabla_v\mathcal{S}(u)\quad\Leftrightarrow\quad\langle\delta,\partial_t\gamma(0)\rangle_{g(v)}\leq\left.\frac{d}{dt}\mathcal{S}(\gamma(t))\right|_{t=0}\,. \tag{1.8}$$

If no confusion occurs, we write $\nabla\mathcal{S}(u)=\nabla_u\mathcal{S}(u)$. In what follows, we denote the *local slope* by

$$\left|\partial\mathcal{S}\right|(u):=\limsup_{w\to u,\,w\in D(\mathcal{S})}\frac{|\mathcal{S}(u)-\mathcal{S}(w)|}{\|u-w\|_{g(u)}}\,, \tag{1.9}$$

which coincides with the definition for Riemannian manifolds, as we will show in section 7.2. It is natural to assume

$$\sup_{\delta\in\nabla_u\mathcal{S}(u)}\|\delta\|_{g(u)}\leq\left|\partial\mathcal{S}\right|(u),\qquad\forall u\in D(d\mathcal{S})\,, \tag{1.10}$$

but in some cases, like in example 3.1, it may not be clear whether $\mathcal{S}(u+hd\mathcal{S}(u))<+\infty$ for some $h>0$ and in these cases, (1.10) might fail.

Finally, for every subset $A\subset\mathcal{H}$ we define the affine hull aff A and its minimal section A° through

$$\operatorname{aff}A:=\left\{\sum_i t_ia_i\,:\,a_i\in A,\,t_i\in\mathbb{R},\,\sum_i t_i=1\right\}\,,$$
$$|A^\circ|:=\inf_{\xi\in A}\|\xi\|_{\mathcal{H}}\,,\qquad A^\circ:=\{\xi\in A\,:\,\|\xi\|_{\mathcal{H}}=|A^\circ|\}\,.$$

Definition 1.2. We say that for any $u\in\mathcal{H}$, $\xi\in\mathcal{H}$ is an element of the limiting subdifferential $d_l\mathcal{S}(u)$ of $\mathcal{S}$ in u if there are $u_n\in\mathcal{H}$ with $u_n\to u$ strongly and $\xi_n\in d\mathcal{S}(u_n)$ such that $\xi_n\rightharpoonup\xi$ weakly in $\mathcal{H}$. The limiting subgradient and the weakly lower semicontinuous envelope of $|\partial\mathcal{S}|$ are defined through

$$\nabla_{l,u}\mathcal{S}(u)=\tilde{g}_u^{-1}(d_l\mathcal{S}(u))\,,$$
$$|\nabla_l\mathcal{S}(u)^\circ|:=\inf_{\xi\in\nabla_l\mathcal{S}(u)}\|\xi\|_{g(u)}\qquad\nabla_l\mathcal{S}(u)^\circ:=\left\{\xi\in\nabla_l\mathcal{S}(u)\,:\,\|\xi\|_{g(u)}=|\nabla_l\mathcal{S}(u)^\circ|\right\}\,.$$

Thus, equation (1.1) has to be understood in the sense of

$$g_u(\partial_tu,\varphi)\in\langle d_lS(u),\varphi\rangle_{\mathcal{H}}+g_u(f,\varphi)\qquad\forall\varphi\in L^2(0,T;\mathcal{H}) \tag{1.11}$$

For the rest of the paper, we assume that $\mathcal{S}$ is an entropy functional in the following sense:

Definition 1.3. Let $(\mathcal{H}_0,\tilde{\mathcal{H}},\mathcal{H},g)$ be an entropy space with $G^*>1$. We say that $\mathcal{S}:\,\mathcal{H}\to(-\infty,+\infty]$ is an entropy functional on $(\mathcal{H}_0,\tilde{\mathcal{H}},\mathcal{H},g)$ if it satisfies :

(1) $D(\mathcal{S})\subset\tilde{\mathcal{H}}$ and $\mathcal{S}:\,\mathcal{H}\to\mathbb{R}$ being proper, lower semicontinuous, i.e. the domain $D(\mathcal{S})$ of $\mathcal{S}$ is non-empty.

(2) $\mathcal{S}+\|\cdot\|_{\mathcal{H}}$ has compact sublevels, i.e. there exists $\tau_*>0$ such that sets

$$\left\{v\in\mathcal{H}\,:\,\mathcal{S}(v)+\frac{1}{2\tau}\sqrt{G^*}^{-1}\|v\|^2_{\mathcal{H}}<C\right\}$$

are compact for any $\tau<\tau_*$ and any $C>0$ and there is a constant $S_0>0$ such that

$$\mathcal{S}(v)+\frac{1}{2\tau_*}\sqrt{G^*}^{-1}\|v\|^2_{\mathcal{H}}\geq -S_0 \tag{1.12}$$

(3) $\mathcal{S}$ satisfies the estimate

$$\|u\|_{\mathcal{H}_0}\leq C\left(\mathcal{S}(u)+|\partial S|^2(u)+\|d\mathcal{S}(u)\|^2_{\mathcal{H}}+1\right)$$

Remark 1.4. Due to the structure of (1.11), it is now evident that $\tilde{\mathcal{H}}$ is the right support for $g_\bullet$: $\mathcal{H}_0$ might be to small, while we know for sure that $D(d_l\mathcal{S})\subset\tilde{\mathcal{H}}$. Also, the compactness of the embedding $\mathcal{H}_0\hookrightarrow\tilde{\mathcal{H}}$ will provide compactness of the approximating sequences in the time discretization sceme. On the other hand, in most cases of interest (i.e. $\tilde{\mathcal{H}}\neq\mathcal{H}$), $\mathcal{H}$ would be far to big.

Following Rossi and Savaré [30], we chose the following approximation scheme: Let $\left(\mathcal{H}_0,\tilde{\mathcal{H}},\mathcal{H},g\right)$ be an entropy space and $\mathcal{S}$ a corresponding entropy functional defined by 1.3. We introduce a modified Moreau-Yosida approximation defining

$$J_\sigma(v,f):=\mathrm{argmin}_{w\in\mathcal{H}}\left(\frac{\|v-w\|^2_{g(v)}}{2\sigma}+\mathcal{S}(w)-\langle f,w\rangle_{g(v)}\right),\qquad v\in\tilde{\mathcal{H}}\,,$$

where the infimum is attained due to 1.3-(2). For fixed $0<T<\infty$ and time step $0<\tau<\tau_*$, there corresponds a partition of $(0,T)$ as

$$t_0:=0<t_1<\cdots<t_j<\cdots<\cdots<t_{N-1}<T\leq t_N,\qquad t_j:=j\tau,\quad N\in\mathbb{N}\,.$$

We set

$$\overline{F}_\tau(t):=F^j_\tau=\frac{1}{\tau}\int_{t_{j-1}}^{t_j}f(s)ds\quad\text{for }t\in(t_{j-i},t_j],\quad j=1,\dots,N$$

and $U^0_\tau:=u_0$ for all τ and for any $j=1,\dots,N$

$$U^j_\tau\in J_\tau(U^{j-1}_\tau,F^j_\tau)\,,\qquad \overline{U}_\tau(t):=U^j_\tau,\quad U_\tau(t):=\frac{t_j-t}{\tau}U^{j-1}_\tau+\frac{t-t_{j-t}}{\tau}U^j_\tau,\qquad t\in(t_{j-1},t_j]\,,$$

We will repeat this construction more detailed in section 5.

Definition 1.5. $u\in H^1(0,T;\mathcal{H})$ is a generalized minimizing movement, provided there is a sequence $\tau_k\to 0$ such that $\bar{U}_{\tau_k}(t)\to u(t)$ in $\mathcal{H}$ for almost all $t\in(0,T)$. We denote the set of all generalized minimizing movements by $GMM(\mathcal{S},u_0,f)$.

Abstract results. The most important Lemma, which will be proved below is the following:

Lemma 1.6. *Let $\mathcal{H}_0$, $\tilde{\mathcal{H}}$, $\mathcal{H}$, g and $\mathcal{S}$ satisfy definitions 1.1 and 1.3. Then, $GMM(\mathcal{S},u_0,f)\neq\emptyset$ and for all $u\in GMM(\mathcal{S},u_0,f)$ holds $u\in H^1(0,T;\mathcal{H})$.*

This lemma will be proved in section 5. One fundamental assumption which is used in the statements of all three main theorems is the continuity assumption:

$$v_n\to v,\ \sup_n\left(|\partial\mathcal{S}(v_n)|\,,\mathcal{S}(v_n)\right)<+\infty\ \Rightarrow\ \mathcal{S}(v_n)\to\mathcal{S}(v)\quad\text{as }n\nearrow\infty \tag{1.13}$$

We are thus ready to state the three main theorems of this paper:

Theorem 1.7. *Let $\mathcal{H}_0$, $\tilde{\mathcal{H}}$, $\mathcal{H}$, g and $\mathcal{S}$ satisfy definitions 1.1 and 1.3 with $d_l\mathcal{S}(u)$ being convex and closed for all $u\in\mathcal{H}$.*

$$\mathcal{S}(u)=\mathcal{S}_{\mathcal{H}}(u)+\mathcal{S}_{\tilde{\mathcal{H}}}(u) \tag{1.14}$$

with functionals $\mathcal{S}_{\mathcal{H}} : \mathcal{H} \to \mathbb{R}$ being proper, lower semicontinuous, and $\mathcal{S}_{\tilde{\mathcal{H}}} : D(\mathcal{S}) \subset \tilde{\mathcal{H}} \to \mathbb{R}$ being continuous w.r.t. $\tilde{\mathcal{H}}$. Furthermore, let $f \in L^2(0,T;\mathcal{H})$. Then, for each $u_0 \in \mathcal{H}_0$ and every $0 < T \in \mathbb{R}$, $GMM(\mathcal{S}, u_0, f) \neq \emptyset$ and each $u \in GMM(\mathcal{S}, u_0, f)$ is a solution to (1.11)*, satisfying the Lyapunov inequality*

$$\frac{1}{2}\int_0^t \|u'\|^2_{g(u)} + \frac{1}{2}\int_0^t |(f - \nabla_l \mathcal{S}(u))^\circ|^2 + \mathcal{S}(u(t)) \leq \mathcal{S}(u(0)) \qquad \text{for a.e. } t \in (0,T)\,. \tag{1.15}$$

If $\mathcal{S}$ additionally fulfills the continuity assumption (1.13) *then, there is a negligible set $\mathcal{N} \subset (0,T)$ such that*

$$\frac{1}{2}\int_s^t |u'|^2 + \frac{1}{2}\int_s^t |(f - \nabla_l \mathcal{S}(u))^\circ|^2 + \mathcal{S}(u(t)) \leq \mathcal{S}(u(s)) \qquad \forall t \in (s,T)\backslash\mathcal{N},\ \forall s \in (0,T)\backslash\mathcal{N}\,.$$

For the next theorem, we assume the following chain rule (see also [30]): If $v \in H^1(0,T;\mathcal{H})$, $\xi \in L^2(0,T;\mathcal{H})$ with $\xi(t) \in d_l\mathcal{S}(v(t))$ for a.e. $t \in (0,T)$, and $\mathcal{S} \circ v$ is a.e. equal to a function s of bounded variation, then

$$\frac{d}{dt}s(t) = \langle \xi, v'(t)\rangle_{\mathcal{H}}\,. \tag{1.16}$$

Theorem 1.8. *Let $\mathcal{H}_0$, $\tilde{\mathcal{H}}$, $\mathcal{H}$, g and $\mathcal{S}$ satisfy definitions 1.1 and 1.3 and assume* (1.13) *and* (1.16) *hold. Furthermore, let $f \in L^2(0,T;\mathcal{H})$. Then, for each $u_0 \in \mathcal{H}_0$ and every $0 < T \in \mathbb{R}$, $GMM(\mathcal{S}, u_0, f) \neq \emptyset$ and each $u \in GMM(\mathcal{S}, u_0, f)$ is a solution to* (1.11) *with $u(0) = u_0$. Furthermore, for almost all $t \in (0,T)$,*

$$\left\{\begin{array}{l} u' \text{ is the projection of the origin on the affine hull } \mathit{aff}(f - \nabla_l \mathcal{S}) \text{ and} \\ \text{fulfills the minimal section principle } u'(t) = (f - \nabla_l \mathcal{S}(u(t)))^\circ \text{ and} \end{array}\right\} \tag{1.17}$$

$$u'(t) \text{ belongs to the strong closure of } f - \nabla\mathcal{S}(u(t))\,. \tag{1.18}$$

Finally, the energy inequality

$$\int_s^t \|u'(\sigma)\|^2_{g(u(\sigma))}\, d\sigma + \mathcal{S}(u(t)) \leq \mathcal{S}(u(s)) \tag{1.19}$$

holds for all $t \in (0,T)$, $s \in (0,t)\backslash\mathcal{N}$, $\mathcal{N}$ being a set of measure 0 in $(0,T)$, and $\mathcal{S} \circ u$ coincides a.e. in $(0,T)$ with a function $s \geq \mathcal{S} \circ u$ of bounded variation satisfying

$$\frac{d}{dt}s(t) = -\|u'(t)\|^2_{g(u(t))} \quad \text{a.e. in } (0,T)\,. \tag{1.20}$$

As already stated by Rossi and Savaré in [30], the Lyapunov inequalities (1.15) and (1.19) hold almost everywhere but $\frac{d}{dt}s(t)$ might not be absolutely continuous with respect to the Lebesgue measure, and $\mathcal{S} \circ u$ might have essential jumps along u. Like in [30], this phenonema can be circumvented using slightly stronger assumptions on the chain rule condition, in particular:

$$\begin{gathered} \text{If } v \in H^1(0,T;\mathcal{H}), \xi \in L^2(0,T;\mathcal{H}) \text{ with } \xi(t) \in \nabla_l\mathcal{S}(v(t)) \text{ for a.e. } t \in (0,T)\,, \text{ and} \\ \mathcal{S} \circ v \text{ is bounded,} \quad \text{then } \mathcal{S} \circ v \in AC(0,T) \text{ and} \frac{d}{dt}\mathcal{S}(v(t)) = \langle \xi, v'(t)\rangle_{g(v(t))}\,. \end{gathered} \tag{1.21}$$

Using this improved chainrule condition, we generalize [30, theorem 3]:

Theorem 1.9. *Let $\mathcal{H}_0$, $\tilde{\mathcal{H}}$, $\mathcal{H}$, g and $\mathcal{S}$ satisfy definitions 1.1 and 1.3 and assume* (1.13) *and* (1.21) *hold. Furthermore, let $f \in L^2(0,T;\mathcal{H})$. Then, for each $u_0 \in \mathcal{H}_0$ and every $0 < T \in \mathbb{R}$, $GMM(\mathcal{S}, u_0, f) \neq \emptyset$ and each $u \in GMM(\mathcal{S}, u_0, f)$ is a solution to* (1.11) *with $u(0) = u_0$. Furthermore, for almost all $t \in (0,T)$,* (1.17)-(1.18) *hold. Finally, we get the energy equality*

$$\int_s^t \|u'(\sigma)\|^2_{g(u(\sigma))}\, d\sigma + \mathcal{S}(u(t)) = \mathcal{S}(u(s)) \qquad \forall s,t: \quad 0 \leq s \leq t \leq T\,. \tag{1.22}$$

Comparison with former results. Up to [30] one of the major ingredients to the existence theory of gradient flow was the assumption that the graph of $(\mathcal{S}, d\mathcal{S})$ was strongly-weakly closed in $\mathcal{H} \times \mathcal{H} \times \mathbb{R}$, i.e.

$$\left.\begin{array}{l} \xi_n \in d\mathcal{S}(v_n), \quad r_n = \mathcal{S}(v_n) \\ v_n \to v, \quad \xi_n \rightharpoonup \xi, \quad r_n \to r \end{array}\right\} \quad \Rightarrow \quad \xi \in d\mathcal{S}(v), \quad r = \mathcal{S}(v), \tag{1.23}$$

yielding $d_l\mathcal{S} = d\mathcal{S}$. As explained by Rossi and Savaré this condition yields closedness and convexity of $d\mathcal{S}$, chainrule (1.16) and the continuity condition (1.13). Note that in case $\mathcal{S}$ is convex, condition (1.23) is fulfilled. For a review on results earlier than [2, 30] the reader is referred to these sources.

However, the results of [30] generalized former existence results to the case that only one of the assumptions (1.16) or closedness and convexity of $d\mathcal{S}$ holds. As in the present paper, the continuity assumption (1.13) is needed in both cases. A further important contribution by Rossi and Savaré is a result on the class of functionals satisfying the continuity conditions (1.13) or (1.21), compare e.g. for [30] theorem 4. This result will be cited as theorem 2.2 below.

Another recent result due to Rossi, Mielke and Savaré [29], generalizes classical results as in [2] to time dependent functionals and to quasi-metrics based on asymmetric distances. They also construct solutions for distances defined through Finstlerian structures. As one may be tempted in interpreting g as a classical Finstlerian structure, note that these structures need to be defined in a unique way such that in definition 1.1 $\tilde{\mathcal{H}} = \mathcal{H}$. In general, note that a topological approach using metric spaces, as in [2] or [29] would not lead to above results: As stated clearly in [30, 29], results in metric spaces only apply to functionals satisfying (1.23). Another recent and conceptionally related result is due to Mielke, Rossi and Savaré [23]: They considered a Finstler structure generalizing 1.1-(2) but claiming convergence in (1.3) only for $u_n \to u$ in $\mathcal{H}$.

The approximation scheme, introduced above, provides the advantage that the theory of Young measures, one of the major ingredients to [30], is still applicable. This makes the proofs similar to [30] but with some non trivial generalizations. Note that particularly, properties 2 and 3 of g and $\mathcal{S}$ are crucial, but as will be shown in sections 3 and 8 they are fulfilled by a large class of metrics and functionals.

Finally, compared to [30], the major non-trivial steps are the following: [30] benefits a lot from the strong-weak closure of $d_l\mathcal{S}$ by linear convergence arguments. However, proving convergence of the approximation scheme, we will have to work with $(\nabla_{v_n}\mathcal{S})(u_n) = \tilde{g}_{v_n}^{-1}(d\mathcal{S}(u_n))$, thus with a partially nonlinear dependence of $\tilde{g}$ on v_n. Furthermore, we have to make sure the chain rule condition (1.16) implies $s' = \langle \nabla_{l,v}\mathcal{S}(v), v' \rangle_{g(v)}$ at least for generalized minimizing movements. Having shown these two important properties, the remaining steps are similar to standard methods in gradient flows.

Physical Insight. It is commonly known that many partial differential equations, in particular describing phase transitions, can be described as gradient flows with respect to a functional $\mathcal{S}$. We refer to [6, 29, 30] for further references. From the applied point of view, this is the major reason for the huge interest in such systems. Starting from an initial article by Otto [26], the theory seems to have developed a new dynamics over the last couple of years and the partially formal calculations below are surely inspired by what Villani calls in his book "Otto calculus" [32], though we will not work in Wasserstein spaces.

Most authors assume that $\mathcal{S}$ is the total energy or the free energy of the systems. However, the author of the present article believes that $\mathcal{S}$ should reflect the entropy of the system. This is for several reasons: First, entropy is *the key* to thermodynamics, the fundamental variable which distinguishes thermodynamics from any other physical discipline. Second, a recently developed method by Rajagopal and Srinivasa [27], based on a maximization of the rate of entropy production, is able to provide models to various problems in fluid dynamics and other fields. Third, the author formally showed in [16, 17] that these assumptions can also be formulated in an integral setting. It will be main part of section 7 to introduce the ideas behind this method and to demonstrate at least for a special class of models, that the resulting equations will always lead to a gradient flow.

When choosing this approach, one has to be careful with the choice of variables:

(1) Entropy does not depend on temperature ϑ, but on the internal energy e. Furthermore, if $\mathcal{S} = \int \eta(e)$, we find $\vartheta^{-1} = \frac{\partial \eta}{\partial e}$. Thus, the relevant equation would be formulated in terms of e in form of

$$\partial_t e - \operatorname{div}\left(\kappa \nabla \frac{1}{\vartheta(e)}\right) = 0\,.$$

(2) Entropy is concave in e (see e.g. the most interesting work by Lieb and Yngvason [21]) but also in all other variables. For the study in terms of gradient flows, it is thus more comfortable to study the evolution of $\tilde{\mathcal{S}} = -\mathcal{S}$. The (physically unavoidable) increase of $\mathcal{S}$ with time is then reflected in a decrease of $\tilde{\mathcal{S}}$ with time. Note that concavity in combination with $\vartheta^{-1} = \frac{\partial \eta}{\partial e}$ in particular implies the monotonicity of $\vartheta(e)$.

We will come back to this point later in sections 7-8.

Structure of the article. In section 2 we will recall some fundamental Banach spaces and notations for the main proofs but also the fundamental Sobolev spaces that will be used in the examples in sections 3 and 8. Section 3 will deal with some examples, among which can be found the Allen-Cahn and the Cahn-Hilliard equations with additional nonlinearities but also generalizations of Examples given in [30]. Furthermore, we will shortly highlight, why the metric approach may fail even in case (1.23) is given. In section 4, some results from theory of Young measures will be recalled and a new convergence result will be proved. In section 5 we introduce the approximation scheme with more detail and provide uniform a priori estimates, which will be used in section 6 to proof the main existence results 1.7-1.9. Finally, in sections 7 and 8 the assumption of maximum rate of entropy production will be introduced for reaction diffusion equations and it will be shown that the method is implicitly equivalent to the theory of curves of maximal slope, while finally we prove general existence results for the obtained equations.

2. Notations and Preliminaries

For any Hilbert space $\mathcal{H}$, we denote $L^p(0,T;\mathcal{H})$ the Bochner space of L^p-functions over $(0,T]$ having values in $\mathcal{H}$ and by $H^1(0,T;\mathcal{H})$ the space of functions $u \in L^2(0,T;\mathcal{H})$ having $u' \in L^2(0,T;\mathcal{H})$. Furthermore, by $C([0,T],\mathcal{H})$ we denote the continuous functions from $[0,T]$ to $\mathcal{H}$, by $C^k([0,T],\mathcal{H})$ the k-times continuously differentiable functions and by $AC([0,T];\mathcal{H})$ the set of absolutely continuous functions over $[0,T]$.

Theorem 2.1 (Egorov's theorem for $L^2(0,T;\mathcal{H})$)**.** *Let $\mathcal{H}$ be a Hilbert space and $(v_n)_{n\in\mathbb{N}} \subset L^2(0,T;\mathcal{H})$ be a sequence such that $v_n \to v \in L^2(0,T;\mathcal{H})$ strongly and pointwise for a.e. $t \in (0,T)$. Then, for any $\varepsilon > 0$ there is $K_\varepsilon \subset (0,T)$ compact with $\mathcal{L}((0,T)\backslash K_\varepsilon) < \varepsilon$ such that $v_n \to v$ uniformly on K_ε.*

Proof. For any $\varepsilon > 0$, Egorov's theorem yields existence of measurable $\tilde{K}_\varepsilon$ with $\mathcal{L}((0,T)\backslash\tilde{K}_\varepsilon) < \frac{\varepsilon}{2}$ such that $v_n \to v$ uniformly on $\tilde{K}_\varepsilon$. Inner regularity of the Lebesgue measure yields existence of compact $K_\varepsilon \subset \tilde{K}_\varepsilon$ and $\mathcal{L}(\tilde{K}_\varepsilon\backslash K_\varepsilon) < \varepsilon/2$. □

Theorem 2.2. [30] *Let $\mathcal{S} : \mathcal{H} \to (-\infty,+\infty]$ be a functional satisfying the compactness property* (1.13) *and admitting the following decomposition:*

$$\mathcal{S} = \psi_1 - \psi_2 \quad \text{in} \quad D(\mathcal{S}), \quad \text{with}$$
$$\psi_1 \,:\, D(\mathcal{S}) \to \mathbb{R} \quad \text{l.s.c and satisfying (1.16) or (1.21)},$$
$$\psi_2 \,:\, co(D(\mathcal{S})) \to \mathbb{R} \quad \text{convex and l.s.c. in } D(\mathcal{S}), \quad D(d_l\psi_1) \subset D(d\psi_2)\,.$$

If

$$\forall M \geq 0 \quad \exists \rho < 1,\, \gamma \geq 0 \quad s.t. \quad \sup_{\xi_2 \in d\psi_2(u)} \|\xi_2\|_{\mathcal{H}} \leq \rho\,\|d_l\psi_1(u)^\circ\|_{\mathcal{H}} + \gamma$$
$$\text{for every } u \in D(d_l\psi_1) \text{ with } \max(\mathcal{S}(u), \|u\|_{\mathcal{H}}) \leq M$$

then $\mathcal{S}$ satisfies the corresponding chain rule property (1.16) *resp.* (1.21).

Frequently used Hilbert spaces. In order to study the examples below, we will frequently make use of the following Hilbert spaces: We consider an open, bounded domain $\Omega \subset \mathbb{R}^n$ with smooth boundary $\Gamma = \partial\Omega$ and outer normal vector $\boldsymbol{n}_\Gamma$. $H^m(\Omega)$ denotes the usual Sobolev space of order m and $H^m_0(\Omega)$ denotes the space of $H^m(\Omega)$-functions having zero boundary values. $H^{-1}(\Omega)$ is the dual of $H^1_0(\Omega)$. Furthermore, we introduce

$$H^1_{(0)}(\Omega) := \left\{\phi \in H^1(\Omega) \,:\, \int_\Omega \phi = 0\right\}$$

with the scalar product

$$\langle\phi,\psi\rangle_{H^1_{(0)}} := \int_\Omega \nabla\phi\cdot\nabla\psi \qquad \forall\phi,\psi \in H^1_{(0)}(\Omega)$$

and its dual space $H^{-1}_{(0)}(\Omega)$ with scalar product

$$\langle\phi,\psi\rangle_{H^{-1}_{(0)}} := \langle\nabla\Delta_N^{-1}\phi, \nabla\Delta_N^{-1}\psi\rangle_{L^2} \qquad \forall\phi,\psi \in H^{-1}_{(0)}(\Omega)\,,$$

where Δ_N is the Laplace operator with Neumann boundary conditions. More generally, define

$$L^2_{(m)}(\Omega) := \left\{f \in L^2(\Omega) \,:\, \int_\Omega f = m\right\}\,,$$

and

$$P_0 :\; L^2(\Omega) \to L^2_{(0)}(\Omega), \qquad f \mapsto f - \int_\Omega f$$

the orthogonal projection on $L^2_{(0)}(\Omega)$. For simplicity, we may sometimes omit the (Ω) if the context is clear. Then, $-\Delta_N : H^1_{(0)}(\Omega) \to H^{-1}_{(0)}(\Omega)$ is the Riesz isomorphism. We then get the dual product as the formal product

$$\langle \phi, \psi \rangle_{H^1_{(0)}, H^{-1}_{(0)}} = \int_\Omega \phi\psi \qquad \forall \phi \in H^1_{(0)}(\Omega)\,,\ \psi \in H^{-1}_{(0)}(\Omega)\,.$$

If $H^*(\Omega) := H^1(\Omega)^{-1}$ is the dual space of $H^1(\Omega)$, the Riesz isomorphism $\mathcal{R} : H^1(\Omega) \to H^*(\Omega)$ is given by $\mathcal{R} := -\Delta_N + 1$ in a sense that $(-\Delta_N + 1)^{-1}\phi$ is the unique solution $p \in H^1(\Omega)$ satisfying

$$-\Delta_N p + p = \phi\,.$$

The scalar product on $H^*(\Omega)$ is then given through

$$\langle \phi, \psi \rangle_{H^*} := \left\langle \nabla\left(-\Delta_N + 1\right)^{-1}\phi, \nabla\left(-\Delta_N + 1\right)^{-1}\psi \right\rangle_{L^2} + \left\langle \left(-\Delta_N + 1\right)^{-1}\phi, \left(-\Delta_N + 1\right)^{-1}\psi \right\rangle_{L^2} \qquad \forall \phi, \psi \in H^*(\Omega)\,,$$

with the dual pairing

$$\langle \phi, \psi \rangle_{H^1, H^*} = \int_\Omega \phi\psi \qquad \forall \phi \in H^1(\Omega)\,,\ \psi \in H^*(\Omega)\,.$$

Lemma 2.3. *Let $A \in L^\infty(\Omega; \mathbb{R}^{n\times n})$, $B \in L^\infty(\Omega)$ having the property that there is $0 < C \leq 1$ such that $C\,|\xi|^2 \leq \xi A(x)\xi \leq C^{-1}\,|\xi|^2$, $C < B(x) < C^{-1}$ for a.e. $x \in \Omega$ and for all $\xi \in \mathbb{R}^n$. For $\phi \in H^{-1}_{(0),n}(\Omega)$, $\psi \in H^*(\Omega)$ let $p_\phi \in H^1_{(0)}(\Omega)$, $p_\psi \in H^1(\Omega)$ solve*

$$\begin{aligned} -\mathit{div}\,(A\nabla p_\phi) &= \phi \ \text{ on } \Omega, \qquad (A\nabla p_\phi)\cdot \boldsymbol{n}_\Gamma = 0 \ \text{ on } \Gamma\,,\\ -\mathit{div}\,(A\nabla p_\psi) + B\,p_\psi &= \psi \ \text{ on } \Omega, \qquad (A\nabla p_\phi)\cdot \boldsymbol{n}_\Gamma = 0 \ \text{ on } \Gamma\,. \end{aligned}$$

Then, there is $0 < G \leq 1$ only depending on A, B, C and Ω such that for all $\phi \in H^{-1}_{(0)}(\Omega)$, $\psi \in H^*(\Omega)$ holds

$$\begin{aligned} G\,\|\phi\|^2_{H^{-1}_{(0)}} &\leq \int_\Omega \nabla p_\phi \cdot (A\nabla p_\phi) \leq G^{-1}\,\|\phi\|^2_{H^{-1}_{(0)}}\,,\\ G\,\|\psi\|^2_{H^*} &\leq \int_\Omega \nabla p_\psi \cdot (A\nabla p_\psi) + B\,p_\psi^2 \leq G^{-1}\,\|\psi\|^2_{H^*}\,. \end{aligned}$$

Proof. Let $p^*_\phi \in H^1_{(0)}(\Omega)$ solve

$$-\mathrm{div}\,(\frac{C}{2}\nabla p^*_\phi) = \phi \text{ on } \Omega\,, \qquad \nabla p^*_\phi \cdot \boldsymbol{n}_\Gamma = 0 \text{ on } \Gamma$$

and define $\tilde{p}_\phi := p_\phi - p^*_\phi$. Through

$$\int_\Omega \nabla\tilde{p}_\phi A \nabla\varphi + \int_\Omega \nabla p^*_\phi \left(A - \frac{C}{2}\right)\nabla\varphi = 0$$

we can estimate $\tilde{p}_\phi$ by p^*_ϕ by testing with $\tilde{p}_\phi$. By the same argument, the right inequality is trivial. A similar argument applies to the second chain of inequalities. □

Definition 2.4. Let $\mathcal{S}$ be proper functional $\mathcal{S} : \mathcal{H} \to (-\infty, +\infty]$ for $\mathcal{H} = H^{-1}_{(0)}(\Omega)$ or $\mathcal{H} = H^*(\Omega)$. Then, we consider the restriction of $\tilde{\mathcal{S}} := \mathcal{S}\big|_{L^2}$ of $\mathcal{S}$ to $L^2(\Omega)$ and define the set valued L^2-subdifferentials $\frac{\delta\mathcal{S}}{\delta u}(u) \subset L^2(\Omega)$ and $\frac{\delta^0\mathcal{S}}{\delta u}(u) \subset L^2_{(0)}(\Omega)$ at $u \in D(\tilde{\mathcal{S}})$ through:

$$\begin{aligned} u \in D(\tilde{\mathcal{S}}) : \quad & \delta \in \frac{\delta\mathcal{S}}{\delta u}(u) \quad \Leftrightarrow \quad \langle \delta, v \rangle_{L^2} \leq \lim_{h\searrow 0} \frac{\tilde{\mathcal{S}}(u + hv) - \tilde{\mathcal{S}}(u)}{h} \quad \forall v \in L^2(\Omega)\\ u \in D(\tilde{\mathcal{S}}) \cap L^2_{(0)}(\Omega) : \quad & \delta \in \frac{\delta^0\mathcal{S}}{\delta u}(u) \quad \Leftrightarrow \quad \langle \delta, v \rangle_{L^2} \leq \lim_{h\searrow 0} \frac{\tilde{\mathcal{S}}(u + hv) - \tilde{\mathcal{S}}(u)}{h} \quad \forall v \in L^2_{(0)}(\Omega) \end{aligned}$$

3. Examples

3.1. **Cahn-Hilliard equation with diffusion coefficient depending on ∇u.** On an open and bounded set $\Omega \subset \mathbb{R}^n$ with smooth boundary Γ and outer normal vector $\boldsymbol{n}_\Gamma$, consider the Hilbert space $\mathcal{H} := H^{-1}_{(0)}(\Omega)$. On $L^2_{(0)}(\Omega)$, respectively $\mathcal{H}$, consider the functional

$$\mathcal{S}(u) = \begin{cases} \int_\Omega \left(s_0(u) + \frac{1}{2} |\nabla u|^2 \right) & \text{for } u \in H^1(\Omega) \\ +\infty & \text{otherwise} \end{cases}, \tag{3.1}$$

where $s_0 : [a, b] \to \mathbb{R}$ is continuous and convex and $s_0(x) = +\infty$ for $x \notin [a, b]$.

In this section, we proof existence of solutions to the following generalized Cahn-Hilliard equation:

$$\partial_t u + \operatorname{div} \left(A(u, \nabla u) \nabla \left(\Delta u - s_0'(u) \right) \right) \ni 0 \qquad \text{on } (0, T] \times \Omega\,, \tag{3.2}$$

$$\left(A(u, \nabla u) \nabla \left(\Delta u - s_0'(u) \right) \right) \cdot \boldsymbol{n}_\Gamma = \nabla u \cdot \boldsymbol{n}_\Gamma = 0 \qquad \text{on } (0, T] \times \Omega\,, \tag{3.3}$$

$$u(0) = u_0 \qquad \text{for } t = 0\,. \tag{3.4}$$

Historically, the first existence proof for the Cahn-Hilliard equation with smooth function s_0 and constant A was given in [14]. Though there is a huge literature on Cahn-Hilliard equation (refer to [1, 6] and references therein), there seems to be only few results on concentration dependent mobility, among the most cited being Cahn et. al. [8]. Other works are by Liu [11], the one dimensional treatments by Dal Passo et. al.[12] and Liu [22] and the work by Novick-Cohen [24, 25] which both treat very special cases, but which are both not covered by our approach. Rossi [28] and Grasselli et. al. [15] deal with a Cahn-Hilliard equation of the form

$$\partial_t u - \Delta \alpha(w) = 0\,, \qquad w = s_0'(u) - \Delta u\,,$$

where w represents the analogue of curvature of the interface in the phase field setting. A dependence of the mobility on w is subject to future investigation. Finally, the author is not aware of any work on the mobility depending on ∇u.

Lemma 3.1. [1, Lemma 4.1, Corollary 4.4] *Let $s_0 : [a, b] \to \mathbb{R}$ be a continuous and convex function. Then, $\mathcal{S} : L^2_{(m)}(\Omega) \to \mathbb{R}$ and $\mathcal{S} : \mathcal{H} \to \mathbb{R}$ are proper, lower semicontinuous and convex.*

Additionally, Abels and Wilke [1] identified the L^2-subgradient of $\mathcal{S}$:

Lemma 3.2. [1] *Let $s_0 : [a, b] \to \mathbb{R}$ be a continuous and convex function that is twice continuously differentiable in (a, b) and satisfies $\lim_{x\to a} s_0'(x) = -\infty$, $\lim_{x\to b} s_0'(x) = +\infty$. Moreover, we set $s_0' = +\infty$ for $x \notin (a, b)$ and let $\mathcal{S}$ be defined as in (3.1). Then, for the L^2-subdifferential holds*

$$D\left(\frac{\delta^0 \mathcal{S}}{\delta u}\right) = \left\{ c \in H^2(\Omega) \cap L^2_{(0)}(\Omega) \,:\, s_0'(c) \in L^2(\Omega),\, s_0''(c) |\nabla c|^2 \in L^1(\Omega),\, \partial_n c\Big|_{\partial\Omega} = 0 \right\} \tag{3.5}$$

and

$$\frac{\delta^0 \mathcal{S}}{\delta u}(\tilde{u}) = -\Delta \tilde{u} + P_0 s_0'(\tilde{u})\,. \tag{3.6}$$

Moreover,

$$\|\tilde{u}\|^2_{H^2(\Omega)} + \|s_0'(\tilde{u})\|^2_{L^2(\Omega)} + \int_\Omega s_0''(\tilde{u}) |\nabla \tilde{u}|^2 \le C \left(\left\| \frac{\delta^0 \mathcal{S}}{\delta u}(\tilde{u}) \right\|^2_{L^2(\Omega)} + \|\tilde{u}\|^2_{L^2(\Omega)} \right) \tag{3.7}$$

for some constant C independent of $\tilde{u}$.

For the $\mathcal{H}$-Subdifferential holds

$$D(d\mathcal{S}) = \left\{ c \in D\left(\frac{\delta^0 \mathcal{S}}{\delta u}\right) \,:\, \frac{\delta^0 \mathcal{S}}{\delta u}(c) \in H^1_{(0)}(\Omega) \right\} \tag{3.8}$$

$$d\mathcal{S}(\tilde{u}) = \Delta \left(-\Delta \tilde{u} + P_0 s_0'(\tilde{u}) \right)\,, \tag{3.9}$$

and in particular, $d\mathcal{S}$ is single valued and

$$\|\tilde{u}\|^2_{H^2(\Omega)} \le C \, \|d\mathcal{S}(\tilde{u})\|_{\mathcal{H}}\,. \tag{3.10}$$

Remark 3.3. Note that the proof in [1] yields for any $\tilde u \in D(\frac{\delta\mathcal{S}}{\delta u})$:

$$\frac{\delta\mathcal{S}}{\delta u}(\tilde u) = -\Delta\tilde u + s_0'(\tilde u)\,,$$

where estimate (3.7) reads

$$\|\tilde u\|^2_{H^2(\Omega)} + \|s_0'(\tilde u)\|^2_{L^2(\Omega)} + \int_\Omega s_0''(\tilde u)\,|\nabla\tilde u|^2 \le C\left(\left\|\frac{\delta\mathcal{S}}{\delta u}(\tilde u)\right\|^2_{L^2(\Omega)} + \|\tilde u\|^2_{L^2(\Omega)}\right).$$

By estimate (3.7), the compact embedding $H^2_{(0)} \hookrightarrow H^1_{(0)}$, convexity and lower semicontinuity of $\mathcal{S}$, we have strong-weak closedness of the graph of $(d\mathcal{S},\mathcal{S})$ in the sense of (1.23) (see [4]). In particular, this implies the chain-rule condition (1.16).

Abels and Wilke used these results to study the quasilinear Cahn-Hilliard equation

$$\partial_t u + \Delta\left(\Delta u - s_0'(u)\right) = 0\,.$$

Of course, it is not natural to claim convexity of $\mathcal{S}$ in u. Thus, we assume

$$\mathcal{S}(u) = \begin{cases} \int_\Omega \left(s_0(u) + s_1(u) + \frac{1}{2}|\nabla u|^2\right) & \text{for } u \in L^2_{(0)}(\Omega)\cap H^1(\Omega) \\ +\infty & \text{otherwise}\end{cases},$$

where s_0 has the properties claimed in lemmata 3.1, 3.2 and $s_1 : \mathbb{R} \to \mathbb{R}$ is concave and continuously differentiable, i.e. it has bounded derivatives on $[a,b]$. Then, (3.5)-(3.9) still hold with modified constants for $s_0 \rightsquigarrow s_0 + s_1$. It is easy to check that $d\mathcal{S}$ is still strong-weak closed, in particular, that chain-rule (1.16) is fulfilled due to Theorem 2.2. Note that even for arbitrary $s_1 \in C^1(\mathbb{R}^n)$, we would preserve convexity of $d\mathcal{S}$ and the continuity condition. Thus, theorem 1.7 would remain applicable.

As we will see in section 7, it is physically justifiable to generalize the above Cahn-Hilliard equation to (3.2)-(3.4) and the question arises how to get existence for this equation. First, note that with

$$\tilde{\mathcal{H}} := H^1_{(0)}(\Omega) \quad \mathcal{H}_0 := H^2(\Omega)$$

we find $\mathcal{H}_0 \hookrightarrow \tilde{\mathcal{H}} \hookrightarrow L^2(\Omega) \hookrightarrow \mathcal{H}$ with all embeddings being dens and compact.

For $u \in \tilde{\mathcal{H}}$, we define for $s_1, s_2 \in \mathcal{H}$:

$$g_u(s_1,s_2) = \int_\Omega \nabla p_1^u A(u,\nabla u)\nabla p_2^u = \langle p_1, s_2\rangle_{H^1_{(0)},H^{-1}_{(0)}} = \int_\Omega s_1 p_2^u = \int_\Omega s_2 p_1^u\,, \tag{3.11}$$

where $A : \mathbb{R}\times\mathbb{R}^n \to \mathbb{R}^{3\times 3}$ is bounded and Lipschitz continuous with $\xi\cdot A(\cdot)\xi \ge a_0|\xi|^2$ for all parameters and $\xi \in \mathbb{R}^3$ and p_i^u solves

$$\left.\begin{aligned} -\text{div}\,\left(A(u,\nabla u)\nabla p_i^u\right) &= s_i \quad \text{on } \Omega \\ \left(A(u,\nabla u)\nabla p_i^u\right)\cdot \boldsymbol{n}_\Gamma &= 0 \quad \text{on } \Gamma \end{aligned}\right\} \quad \text{for } i = 1,2\,. \tag{3.12}$$

It is immediate to check that g is a densely defined metric in the sense of definition 1.1 once we realize that for sequences $(s_{1,n}, u_n)_{n\in\mathbb{N}} \subset \mathcal{H}\times\tilde{\mathcal{H}}$ with $s_{1,n} \rightharpoonup s_1$ weakly in $\mathcal{H}$ and $u_n \to u$ strongly in $\tilde{\mathcal{H}}$ as $n\to\infty$ holds $p_1^{u_n} \rightharpoonup p_1^u$ weakly in $H^1_{(0)}(\Omega)$ and thus $g_{u_n}(s_{1,n}, s_2) \to g_u(s_1,s_2)$ for all $s_2 \in \mathcal{H}$.

Then, above considerations together with (1.10) yield that $\mathcal{S}$ fulfills all requirements of definition 1.3. As a consequence of theorem 1.8 we get existence of a solution $u \in H^1(0,T;\mathcal{H}) \cap L^2(0,T;H^2(\Omega))$ to (1.11) and it remains to reconstruct an expression of the form (1.1):

For any $u \in D(\mathcal{S})$, $\delta \in \nabla_u\mathcal{S}(u)$, $s \in H^{-1}_{(0)}(\Omega)$ with p_δ^u and p_s^u solutions of (3.12) for δ and s respectively, we formally write

$$\liminf_{t\to 0}\frac{\mathcal{S}(u+ts)-\mathcal{S}(u)}{t} \ge g_u(\delta,s) = \int_\Omega p_\delta^u s = \int -\text{div}\,\left(A(u,\nabla u)\nabla p_\delta^u\right)p_s^u = \int_\Omega \delta p_s^u \tag{3.13}$$

as the first equality holds for all $s \in L^2_{(0)}(\Omega)$, we find $p_\delta^u = \frac{\delta^0\mathcal{S}}{\delta u}(u)$. Thus, using lemma 3.2 the gradient flow

$$g_u(\partial_t u,\varphi) \in -\langle d_l\mathcal{S}(u),\varphi\rangle_{\mathcal{H}} \qquad \forall\varphi \in L^2(0,T;\mathcal{H})\,, \tag{3.14}$$

is the weak formulation of (3.2)-(3.4) and we have shown the following result:

Theorem 3.4. *Let $\mathcal{S}$ and g be as above. Then, for any $u_0 \in H^1(\Omega)$ there exists $u \in H^1(0,T;\mathcal{H}) \cap L^2(0,T;H^2(\Omega))$ satisfying* (3.14) *and $u(0) = u_0$.*

Remark 3.5.

3.2. **A note on the applicability of the metric approach.** Example 3.1 nicely demonstrates, why it is hard to demonstrate applicability of the metric approach even in case (1.23) is satisfied by the functional $\mathcal{S}$: The metric g_u is only well defined for $u \in H^1(\Omega)$. Defining the length of any curve $\gamma \in AC(0,T;\mathcal{H})$ with $\gamma(t) \in H^1(\Omega)$ for a.e. $t \in (0,T)$ as

$$L(\gamma) := \int_0^T \sqrt{g_{\gamma(t)}(\gamma'(t),\gamma'(t))}dt\,,$$

it is the essence of [7] that this Length structure

(1) is extendable to arbitrary $\gamma \in AC(0,T;\mathcal{H})$ such that
(2) L complies with the length structure induced by the intrinsic metric

if and only if $\gamma \to L(\gamma)$ is lower semicontinuous with respect to pointwise convergence. This means for γ_i such that $\gamma_i(t) \to \gamma(t)$ for all t, $L(\gamma_i) \to L(\gamma)$. However, with regard to the last example, for $\gamma_i(t) \to \gamma(t)$ in $\mathcal{H}$, the best convergence in $H^1_{(0)}(\Omega)$ (or even in L^2) we can expect is weak convergence. On the other hand, due to the above two conditions, for $\gamma_i \in C^1([0,T];\mathcal{H})$ with $\gamma_i \to \gamma \in C^1([0,T];\mathcal{H})$ such that $\gamma_i(t), \gamma(t) \in H^1(\Omega)$ for all t we would need $A(\gamma_i(t), \nabla\gamma_i(t)) \to A(\gamma(t), \nabla\gamma(t))$ for a.e. $t \in (0,T)$. Even if L would be lower semicontinuous, it is not clear whether the resulting gradient flow would be a weak solution to the classical formulation.

3.3. **Allen-Cahn equation with nonlinear diffusion coefficient.** As the Cahn-Hilliard-equation, the Allen-Cahn equation has been subject to various studies, refer to [6]. We consider the functional

$$\mathcal{S}_b(u) := \begin{cases} \int_\Omega \left(s_0(u) + \frac{1}{2}\left|\nabla u\right|^2 \right) & \text{for } u \in H^1(\Omega) \\ +\infty & \text{otherwise} \end{cases} \tag{3.15}$$

on the Hilbert space $L^2(\Omega)$, and we want to find $d\mathcal{S}_b$, being now the subdifferential with respect to $L^2(\Omega)$. From Lemma 3.2, resp. remark 3.3, we know that for any $\gamma \in C^1(-T,T;C^\infty(\Omega))$ holds

$$\frac{d}{dt}\left(\mathcal{S}_b \circ \gamma(t)\right) = \int_\Omega \left(-\Delta\gamma(t) + s_0'(\gamma(t))\right)\gamma'(t)\,,$$

or, differently:

$$\langle d\mathcal{S}_b(u), \varphi\rangle_{L^2(\Omega)} = \int_\Omega \left(-\Delta u + s_0'(u)\right)\varphi \qquad \forall \varphi \in C^\infty(\Omega)\,.$$

Thus, if we introduce for $A \in C^1(\mathbb{R}\times\mathbb{R}^n)$ with $0 < a_0 \le A(\cdot,\cdot) \le a_0^{-1} < \infty$ the metric

$$\begin{aligned} g_u : H^1(\Omega) &\to B(L^2(\Omega)) \\ u &\mapsto \left((\varphi,\psi) \mapsto \int_\Omega A(u,\nabla u)^{-1}\varphi\psi \qquad \forall \varphi,\psi \in L^2(\Omega) \right), \end{aligned}$$

we find that $\mathcal{H}_0 = H^2(\Omega)$, $\tilde{\mathcal{H}} = H^1(\Omega)$ and $\mathcal{H} = L^2(\Omega)$ such that we find the following result:

Proposition 3.6. *For any $u_0 \in D(d\mathcal{S}_b)$ There exists $u \in H^1(0,T;\mathcal{H}) \cap L^2(0,T;\mathcal{H}_0)$ such that u solves*

$$\begin{aligned} \partial_t u + A(u,\nabla u)\left(-\Delta u + s_0'(u)\right) &= 0 && \text{on } (0,T)\times\Omega \\ \boldsymbol{n}_\Gamma \cdot \nabla u &= 0 && \text{on } (0,T)\times\partial\Omega \\ u(0) - u_0 &= 0 && \text{on } \Omega\,. \end{aligned}$$

3.4. **A finite dimensional example.** Let $F, G \in C^1(\mathbb{R}^n)$ be two given functions satisfying

$$\liminf_{|u|\to+\infty} \frac{F(u)}{|u|^2} > -\infty\,, \qquad \liminf_{|\chi|\to+\infty} \frac{G(\chi)}{|\chi|} = +\infty\,.$$

Rossi and Savaré [30] considered the system

$$\begin{aligned} u'(t) + \nabla F(u(t)) &= \chi(t) + f(t)\,, \\ \nabla G(\chi(t)) &= u(t)\,, \end{aligned}$$

and showed that this system can be reinterpreted as gradient flow with respect to the functional

$$\phi(u) := \min_{\sigma\in\mathbb{R}^n} \left(F(u) + G(\sigma) - \langle u, \sigma\rangle\right).$$

Due to superlinear growth of G, the set

$$M(u) := \operatorname{argmin}\left(G(\sigma) - \langle u, \sigma\rangle\right)$$

is not empty and we find $d\phi(u) = \nabla F(u) - M(u)$, which transforms above system into the shape

$$u'(t) \in -d\phi(u) + f(t)\,.$$

Eventually using theorem 1.8, we may now also find solutions to the problem

$$\begin{aligned} u'(t) &= A(u)\left(\chi(t) - \nabla F(u(t))\right) + f(t)\,, \\ \nabla G(\chi(t)) &= u(t)\,, \end{aligned}$$

where $A \in C^{0,1}(\mathbb{R}^n)^{n\times n}$ is bounded, Lipschitz continuous, uniformly strictly positive definite and uniformly elliptic.

Note that Rossi and Savaré also give an explicit example for G such that $M(u)$ is not convex.

3.5. **The Stefan problem.** A similar problem to 3.3 that was provided by Rossi and Savaré is the Stefan problem

$$\begin{aligned} \partial_t u - \Delta\beta(u) &= f && \text{on } (0,T)\times\Omega \\ \beta(u) &= 0 && \text{on } (0,T)\times\partial\Omega \\ u(0,x) &= u_0(x) && \text{on } \Omega \end{aligned}$$

where $f \in L^2(0,T;H^{-1}(\Omega))$ and $\beta(u) = (u-1)^+ - (u-1)^-$ for $u \in \mathbb{R}$. In [30], it is shown that the above system is a gradient flow in $H^{-1}(\Omega)$ with respect to the functional

$$\phi(u) := \min_{\sigma\in L^2(\Omega)} \int_\Omega \left(\frac{1}{2}|u-\sigma|^2 + I_{[-1,1]}(\sigma)\right),$$

where

$$I_{[-1,1]}(\sigma) = \begin{cases} 0 & \text{if } -1 \le x \le 1 \\ +\infty & \text{else} \end{cases}.$$

However, in view of the above considerations, the generalization to

$$\begin{aligned} \partial_t u - \operatorname{div}\left(A(u)\nabla\beta(u)\right) &= f && \text{on } (0,T)\times\Omega \\ \beta(u) &= 0 && \text{on } (0,T)\times\partial\Omega \\ u(0,x) &= u_0(x) && \text{on } \Omega \end{aligned}$$

is now obvious.

4. Young measures

For a separable metric space E, we denote by $\mathcal{B}(E)$ the Borel-σ-algebra, where $\mathcal{L}(0,T)$ is the Lebesgue-σ-algebra on $(0,T)$ and $\mathcal{L}(0,T)\otimes\mathcal{B}(E)$ is the product σ-algebra. $\mathcal{M}(0,T;E)$ denotes the set of measurable functions over $(0,T)$ with values in E. A $\mathcal{L}(0,T)\otimes\mathcal{B}(E)$-measurable function $h:(0,T)\times E\to(-\infty,+\infty]$ is a normal integrand if $v\mapsto h(t,v)$ is lower semicontinuous for all $t\in(0,T)$. We denote the set of measurable normal integrands by $\mathcal{G}(0,T;E)$ and the set of positive normal integrands by $\mathcal{G}^+(0,T;E)$. The subset $\mathcal{K}(0,T;E)$ of $\mathcal{G}^+(0,T;E)$ is the set of all $h\in\mathcal{G}^+(0,T;E)$ such that for all $t\in(0,T)$, $0\le\gamma<+\infty$

$$\{e\in E\,:\,h(t,e)\le\gamma\}\ \text{ is compact}\,.$$

Definition 4.1. A sequence $(\boldsymbol{v_k})_{k\in\mathbb{N}}\subset\mathcal{M}(0,T;E)$ is said to be *tight* if there exists $h\in\mathcal{K}(0,T;E)$ such that

$$\sup_k\int_0^T h(t,\boldsymbol{v_k})\,dt<+\infty\,.$$

For a Hilbert space $\mathcal{H}$, let $\mathcal{B}(\mathcal{H})$ denote the Borel-sigma-algebra with respect to $\|\cdot\|_{\mathcal{H}}$. We say that a $\mathcal{L}\otimes\mathcal{B}(\mathcal{H})$-measurable functional $h:\,(0,T)\times\mathcal{H}\to(-\infty,+\infty]$ is a weakly normal integrand if

$$v\mapsto h_t(v):=h(t,v)\quad\text{is sequentially weakly l.s.c. for a.e. } t\in(0,T)\,.$$

Definition 4.2. (Time dependent parameterized measures) A parameterized measure in E is a family $\boldsymbol{\nu}:=\{\nu_t\}_{t\in(0,T)}$ of Borel probability measures on $\mathcal{H}$ such that

$$t\in(0,T)\mapsto\nu_t(B)\quad\text{is }\mathcal{L}-\text{measurable for all } B\in\mathcal{B}(E)\,.$$

We denote by $\mathcal{Y}(0,T;E)$ the set of all parameterized measures.

For computations below, the most important result on parameterized measures is a generalization of Fubini's theorem [13]: For every parameterized measure $\boldsymbol{\nu}=\{\nu_t\}_{t\in(0,T)}$, there exists a unique measure ν on $\mathcal{L}(0,T)\otimes\mathcal{B}(E)$ defined by

$$\nu(I\times A)=\int_I\nu_t(A)dt\qquad\forall I\in\mathcal{L}(0,T),\ A\in\mathcal{B}(E)\,.$$

Moreover, for every $\mathcal{L}(0,T)\otimes\mathcal{B}(E)$-measurable function $h:(0,T)\times E\to[0,+\infty]$, the function

$$t\mapsto\int_E h(t,\xi)d\nu_t(\xi)$$

is $\mathcal{L}(0,T)$-measurable and the Fubini integral representation holds:

$$\int_{(0,T)\times E}h(t,\xi)d\nu(t,\xi)=\int_0^T\left(\int_E h(t,\xi)d\nu_t(\xi)\right)dt\,. \tag{4.1}$$

If ν is concentrated on the graph of a measurable function $u:(0,T)\to E$, then $\nu_t=\delta_{u(t)}$ for a.e. $t\in(0,T)$, where $\delta_{u(t)}$ denotes the Dirac's measure carried by $\{u(t)\}$. In this case, by (4.1):

$$\int_{(0,T)\times E}h(t,\xi)d\nu(t,\xi)=\int_0^T h(t,u(t))dt\,.$$

The following theorem is due to Balder and adapted to the specific case of an interval $(0,T)$:

Theorem 4.3. [3, Theorem 1]*Suppose $(\boldsymbol{v_k})_{k\in\mathbb{N}}\subset\mathcal{M}(0,T;E)$ is tight. Then, there exists a subsequence still denoted k and a parameterized measure $\boldsymbol{\nu}=\{\nu_t\}_{t\in(0,T)}\in\mathcal{Y}(0,T;E)$ such that for all $h\in\mathcal{G}(0,T;E)$*

$$\lim_k\int_0^T h(t,\boldsymbol{v_k}(t))\,dt\ge\int_0^T\int_E h(t,\xi)\,d\nu_t(\xi)\,dt$$

provided that

$$\left\{t\mapsto h^-(t,\boldsymbol{v_k}(t))\right\}\quad\textit{is uniformly integrable.}$$

Moreover, for a.e. $t\in(0,T)$, the measure ν_t is carried by the set

$$\bigcap_{p=1}^{\infty}\overline{\{\boldsymbol{v_k}(t)\,:\,k\ge p\}}^E$$

of all limit points of $\boldsymbol{v_k}(t)$.

This theorem is the key ingredient in the following fundamental result by Rossi and Savaré, which will be generalized in theorem 4.7 below:

Theorem 4.4 (The fundamental theorem for weak topologies)**.** [30] *Let* $\{v_n\}_{n\in\mathbb{N}}$ *be a bounded sequence in* $L^p(0,T;\mathcal{H})$*, for some* $+\infty > p > 1$*. Then there exists a subsequence* $k \mapsto v_{n_k}$ *and a parameterized measure* $\boldsymbol{\nu} = \{\nu_t\}_{t\in(0,T)} \in \mathcal{Y}(0,T;\mathcal{H})$ *such that for a.e.* $t \in (0,T)$

$$\limsup_{k\to\infty} \|v_{n_k}(t)\|_{\mathcal{H}} < +\infty, \quad \nu_t \text{ is concentrated on } L(t) := \bigcap_{q=1}^{\infty} \overline{\{v_{n_k}(t)\,:\,k \geq q\}}^{w}$$

of weak limit points of $\{v_n\}_{n\in\mathbb{N}}$ *, and*

$$\liminf_{k\to\infty} \int_0^T h(t, v_{n_k}(t))dt \geq \int_0^T \left(\int_{\mathcal{H}} h(t,\xi) d\nu_t(\xi) \right) dt$$

for every weakly normal integrand h *such that* $h^-(\cdot, v_{n_k}(\cdot))$ *is uniformly integrable. In particular,*

$$\int_0^T \left(\int_{\mathcal{H}} \|\xi\|_{\mathcal{H}}^p \, d\nu_t(\xi) \right) \leq \liminf_{k\to\infty} \int_0^T \|v_{n_k}\|_{\mathcal{H}}^p \, dt \,,$$

and, setting

$$v(t) := \int_{\mathcal{H}} \xi d\nu_t(\xi), \quad \text{we have } v_{n_k} \rightharpoonup v \text{ in } L^p(0,T;\mathcal{H})\,.$$

Finally, if $\nu_t = \delta_{v(t)}$ *for a.e.* $t \in (0,T)$*, then*

$$\langle v_{n_k}, w\rangle_{\mathcal{H}} \to \langle v, w\rangle_{\mathcal{H}} \quad \text{in } L^1(0,T) \quad \forall w \in L^q(0,T;\mathcal{H}), \quad \frac{1}{p} + \frac{1}{q} = 1\,,$$

and up to extraction of a further subsequence independent of t *(still denoted by* v_{n_k}*)*

$$v_{n_k}(t) \rightharpoonup v(t) \quad \text{for a.e. } t \in (0,T)\,.$$

For the proofs of theorems 1.15-1.9 below, we need the additional limit behavior

$$\liminf_{k\to\infty} \int_0^T g_{u_{n_k}(t)}(v_{n_k}(t), v_{n_k}(t))dt \geq \int_0^T \left(\int_{\mathcal{H}} g_{u(t)}(\xi,\xi) d\nu_t(\xi) \right) dt\,. \tag{4.2}$$

as $k \to \infty$, where $u_{n_k}(t) \to u(t)$ for a.e. $t \in (0,T)$. Theorem 4.4 gives us no information whether this holds, but the original proof gives us strong hints. Thus, in the proof of the following theorem we will shortly address the main steps of the original proof of 4.4 in [30] as far as needed, while we skip the parts that are proved identically to 4.4.

Corollary 4.5. *As a consequence of definition 1.1-*(2)*, we find for* $u_n \to u$ *strongly in* $\tilde{\mathcal{H}}$ *and* $\varphi_n \rightharpoonup \varphi$ *weakly in* $\mathcal{H}$*:*

$$g_u(\varphi,\varphi) \leq \liminf_{n\to\infty} g_{u_n}(\varphi_n,\varphi_n)\,.$$

Proof.

$$\|\varphi\|_{g(u)}^2 := g_u(\varphi,\varphi) \overset{1.1-(2)}{=} \liminf_{n\to\infty} g_{u_n}(\varphi_n,\varphi) \leq \liminf_{n\to\infty} \|\varphi_n\|_{g(u_n)} \|\varphi\|_{g(u_n)}$$

$$\leq \liminf_{n\to\infty} \|\varphi_n\|_{g(u_n)} \limsup_{n\to\infty} \sqrt{g_{u_n}(\varphi,\varphi)} \overset{1.1-(2)}{=} \liminf_{n\to\infty} \|\varphi_n\|_{g(u_n)} \|\varphi\|_{g(u)} \,.$$

□

Corollary 4.6. *For a bounded sequence* $\varphi_n \in \mathcal{H}$ *and* $u_n \to u$ *strongly in* $\tilde{\mathcal{H}}$*, we find* $\varphi_n \rightharpoonup \varphi$ *weakly in* $\mathcal{H}$ *iff* $\tilde{g}_{u_n}(\varphi_n) \rightharpoonup \tilde{g}_u(\varphi)$ *weakly in* $\mathcal{H}$*, where* $\tilde{g}_u$ *is defined through* (1.4)*.*

Proof. Let $\tilde{g}_{u_n}(\varphi_n) \rightharpoonup a$ weakly in $\mathcal{H}$ and note that bijectivity of $\tilde{g}_u$ yields $\varphi \in \mathcal{H}$ such that $a = \tilde{g}_u(\varphi)$, i.e. $g_{u_n}(\varphi_n,\psi) \to g_u(\varphi,\psi)$ for all $\psi \in \mathcal{H}$. As φ_n is bounded, there is a subsequence such that $\varphi_n \rightharpoonup \tilde{\varphi}$ weakly in $\mathcal{H}$. The convergence properties of $g_\bullet$ yield $\lim_{n\to\infty} g_{u_n}(\varphi_n,\psi) = g_u(\varphi,\psi) = g_u(\tilde{\varphi},\psi)$ for all $\psi \in \mathcal{H}$. □

Theorem 4.7. *Let $(\mathcal{H}_0, \tilde{\mathcal{H}}, \mathcal{H}, g)$ be an entropy space in sense of definition 1.1, $\{v_n\}_{n\in\mathbb{N}}$ be a bounded sequence in $L^p(0,T;\mathcal{H})$, for some $p>1$, and let $\{u_n\}_{n\in\mathbb{N}}$ be a sequence in $L^p(0,T;\tilde{\mathcal{H}})$, $u \in L^p(0,T;\tilde{\mathcal{H}})$ with $u_n(t) \to u(t)$ for a.e. $t \in (0,T)$. Then there exists a subsequence $k \mapsto v_{n_k}$ and a parameterized measure $\boldsymbol{\nu} = \{\nu_t\}_{t\in(0,T)} \in \mathcal{Y}(0,T;\mathcal{H})$ such that for a.e. $t \in (0,T)$*

$$\limsup_{k\to\infty} \|v_{n_k}(t)\|_{\mathcal{H}} < +\infty, \quad \nu_t \text{ is concentrated on } L(t) := \bigcap_{q=1}^{\infty} \overline{\{v_{n_k}(t)\,:\, k \geq q\}}^{w}$$

of weak limit points of $\{v_n\}_{n\in\mathbb{N}}$. The sequence v_{n_k} and Young measure $\boldsymbol{\nu}$ having the properties from theorem 4.4 and the limit behavior (4.2) *holds.*

Remark 4.8. It is immediate to see that for $\phi_n \in L^2(0,T;\mathcal{H})$ with $\phi_n \rightharpoonup \phi$ weakly in $L^2(0,T;\mathcal{H})$ and u_n as in theorem 4.7, that for $v_n(t) := \tilde{g}^{-1}_{u_n(t)}(\phi_n(t))$ holds

$$L(t) = \bigcap_{q=1}^{\infty} \overline{\left\{\tilde{g}^{-1}_{u_n(t)}(\phi_n(t))\,:\, k \geq q\right\}}^{w} = \tilde{g}^{-1}_{u_n(t)}\left(\bigcap_{q=1}^{\infty} \overline{\{\phi_n(t)\,:\, k \geq q\}}^{w}\right)$$

for a.e. $t \in (0,T)$. This is, as for any sequence $\phi_{n_k(t)} \rightharpoonup \phi(t)$ for $t \in (0,T)$ s.t. $u_{n_k}(t) \to u(t)$ there holds

$$\tilde{g}^{-1}_{u_{n_k}(t)}(\phi_{n_k}(t)) \rightharpoonup \tilde{g}^{-1}_{u(t)}(\phi(t))$$

due to corollary 4.6.

Proof. In order to proof theorem 4.4, Rossi and Savaré [30] used a trick to circumvent the non-metrizability of $\mathcal{H}$ endowed with the weak topology. The construction can also be found in the book by Brézis [5]. We will make use of this trick defining the metric space

$$\tilde{E} := \left\{\boldsymbol{v} = (u,v,w) \in \tilde{\mathcal{H}} \times \mathcal{H} \times \mathbb{R}\,:\, \|v\|_{\mathcal{H}} \leq w\right\}$$

with the distance defined through

$$\tilde{d}(\boldsymbol{v_1}, \boldsymbol{v_2}) := \|u_1 - u_2\|_{\tilde{\mathcal{H}}} + |||v_1 - v_2||| + |w_1 - w_2|,$$

where $|||\cdot|||$ is defined with help of a fixed orthonormal basis $\{e_m\}_{m\in\mathbb{N}}$ of $\mathcal{H}$ and

$$|||v|||^2 := \sum_{m=1}^{\infty} 2^{-m} \left|\langle e_m, v\rangle_{\mathcal{H}}\right|^2 \quad \forall v \in \mathcal{H}\,.$$

$\tilde{E}$ then is a separable complete metric space. We show that any intersection of closed balls in $\tilde{\mathcal{H}} \times \mathcal{H} \times \mathbb{R}$ with $\tilde{E}$ are Borel-subsets of $\tilde{E}$, implying

$$B \in \mathcal{B}(\tilde{\mathcal{H}} \times \mathcal{H} \times \mathbb{R}) \quad \Rightarrow \quad B \cap \tilde{E} \in \mathcal{B}(\tilde{E})\,.$$

Thus, any Borel measure on E can be trivially extended to a Borel measure on $\tilde{\mathcal{H}} \times \mathcal{H} \times \mathbb{R}$.

To this aim, let $B \subset \tilde{\mathcal{H}} \times \mathcal{H} \times \mathbb{R}$ be a closed ball and consider any sequence $(\boldsymbol{v}_k)_{k\in\mathbb{N}} \subset B \cap \tilde{E}$ such that $\tilde{d}(\boldsymbol{v}_k, \boldsymbol{v}) \to 0$. If $\boldsymbol{v}_k = (a_k, b_k, c_k)$, $\boldsymbol{v} = (a,b,c)$, we immediately get $a_k \to a$ in $\tilde{\mathcal{H}}$, $b_k \rightharpoonup b$ weakly in $\mathcal{H}$ and $c_k \to c$. As B is convex, this implies $(a,b,c) \in B \cap \tilde{E}$.

We will show that the sequence $\boldsymbol{v}_n := (u_n, v_n, \|v_n\|_{\mathcal{H}})$ is tight. To this aim, let $\mathcal{N}$ be a set of Lebesgue-measure zero, such that $u_n(t) \to u(t)$ for all $t \notin \mathcal{N}$ and consider the function

$$h(t,(u,v,w)) = \begin{cases} \|v(t)\|^2_{\mathcal{H}} + |w(t)|^2 & \text{if } t \notin \mathcal{N},\ u \in \{u_n(t)\}_{n\in\mathbb{N}} \cup \{u(t)\} \\ \|v(t)\|^2_{\mathcal{H}} + |w(t)|^2 & \text{if } t \in \mathcal{N},\ u \in \{u_1(t)\}_{n\in\mathbb{N}} \\ +\infty & \text{else} \end{cases}.$$

Then, $h \in \mathcal{K}(0,T;\tilde{E})$, as is easy to verify, and

$$\sup_n \int_0^T h(t,(u_n, v_n, \|v_n\|_{\mathcal{H}}) = \sup_n \int_0^T 2\,\|v_n\|^2_{\mathcal{H}}\ ,$$

in particular, $\boldsymbol{v}_n = (u_n, v_n, \|v_n\|_{\mathcal{H}})$ is tight. Applying theorem 4.3, we find a subsequence $k \mapsto \boldsymbol{v}_{n_k} = \big(u_{n_k}, v_{n_k}, \|v_{n_k}\|_{\mathcal{H}}\big)$ and a parameterized measure $\tilde{\boldsymbol{\mu}} = \{\tilde{\mu}_t\}_{t\in(0,T)} \in \mathcal{Y}(0,T;\tilde{E})$ such that for a.e. $t \in (0,T)$ $\tilde{\mu}_t$ is concentrated on

$$\tilde{\boldsymbol{L}}(t) = \bigcap_{q=1}^{\infty} \overline{\{\boldsymbol{v}_{n_k}(t) \,:\, k \geq q\}}^{\tilde{E}} .$$

Defining the complete separable metric space

$$E := \{\boldsymbol{v} = (v,w) \in \mathcal{H}\times\mathbb{R} \,:\, |v| \leq w\} \quad \text{with distance}$$
$$d(\boldsymbol{v_1},\boldsymbol{v_2}) := |||v_1 - v_2||| + |w_1 - w_2|,$$

from the proof of theorem 3.2 in [30], we find for any measurable set $A \in \mathcal{B}(E)$ that $\tilde{\mathcal{H}} \times A \in \mathcal{B}(\tilde{E})$ and the measure $\boldsymbol{\mu} = \{\mu_t\}_{t\in(0,T)} \in \mathcal{Y}(0,T;E)$ with

$$\mu_t(A) := \tilde{\mu}_t(\tilde{\mathcal{H}} \times A)$$

is well defined and concentrated on

$$\boldsymbol{L}(t) = \bigcap_{q=1}^{\infty} \overline{\{\big(v_{n_k}(t), |v_{n_k}(t)|\big) \,:\, k \geq q\}}^{E} .$$

As $\tilde{\boldsymbol{L}}(t) \subset \{u(t)\} \times \boldsymbol{L}(t)$ almost surely, we find $\tilde{\mu}(t) = \delta_{u(t)} \times \mu_t$ a.s. and thus for any $h \in \mathcal{G}(0,T;E)$

$$\lim_k \int_0^T h(t,\boldsymbol{v_k}(t))\,dt \geq \int_0^T \int_E h(t,\xi)\,d\nu_t(\xi)\,dt$$

provided

$$\big\{t \mapsto h^-(t,\boldsymbol{v_k}(t))\big\} \quad \text{is uniformly integrable.}$$

Defining $\nu_t(A) := \mu_t(A \times [0,+\infty))$ for $A \subset \mathcal{H}$, we are back in the setting of theorem 4.4 and all other statements of theorem 4.4 except for (4.2) can be proved like in [30]. Finally, corollary 4.5 yields that $(a,b,c) \mapsto g_a(b,b)$ is a normal integrand on $\tilde{E}$ and theorem 4.3 together with the relation $\nu_t(A) = \mu_t(A \times [0,\infty))$ yields

$$\begin{aligned}\liminf_{k\to\infty} \int_0^T g_{u_{n_k}}(v_{n_k}(t), v_{n_k}(t))dt &\geq \int_0^T \left(\int_{\tilde{E}} g_\zeta(\xi,\xi)d\tilde{\mu}_t(\zeta,\xi,\eta)\right)dt \\ &\geq \int_0^T \left(\int_E g_u(\xi,\xi)d\mu_t(\xi,\eta)\right)dt \geq \int_0^T \left(\int_{\mathcal{H}} g_u(\xi,\xi)d\nu_t(\xi)\right)dt\,.\end{aligned}$$

□

Finally, we will need the following theorem by Rossi and Savaré:

Theorem 4.9. [30] *Let us suppose $\mathcal{S}$ satisfies the chain rule condition* (1.16)*, let $v \in H^1(0,T;\mathcal{H})$ be such that $\mathcal{S}\circ v$ is a.e. equal to a function s of bounded variation and $v(t) \in D(d_l\mathcal{S})$ for a.e. $t \in (0,T)$.*

(1) *If*

$$\int_0^T |d_l^\circ \mathcal{S}(v(t)|^2\,dt < +\infty\,,$$

then

$$s'(t) = \langle \xi, v'(t)\rangle_{\mathcal{H}} \quad \forall \xi \in \overline{aff}(d_l^\circ \mathcal{S}(v(t)))\,. \tag{4.3}$$

(2) *If $\boldsymbol{\mu} = \{\mu_t\}_{t\in(0,T)}$ is a Young measure in $\mathcal{H}$ satisfying*

$$\int_0^T \int_{\mathcal{H}} \|\xi\|^2_{\mathcal{H}}\,d\mu_t(\xi)dt < +\infty, \quad \mu_t(\mathcal{H}\backslash d_l\mathcal{S}(v(t)) = 0 \quad \textit{for a.e. } t \in (0,T)\,,$$

then

$$s'(t) = \int_{\mathcal{H}} \langle \xi, v'(t)\rangle_{\mathcal{H}}\,d\mu_t(\xi) \quad \textit{for a.e.} \quad t \in (0,T)\,.$$

5. Approximation scheme and Apriori Estimates

We want to construct solutions to (1.11) using the approximation scheme going back to De Giorgi, though being slightly modified in order to be placed in between the general frameworks [2] and [30].

Thus, let $\left(\mathcal{H}_0, \tilde{\mathcal{H}}, \mathcal{H}, g\right)$ be an entropy space and $\mathcal{S}$ a corresponding entropy functional according to 1.3. We introduce a modified Moreau-Yosida approximation defining

$$\begin{cases} \mathcal{S}^*(\sigma, v; w, \mathfrak{f}) := \dfrac{\|v-w\|^2_{g(v)}}{2\sigma} + \mathcal{S}(w) - \langle \mathfrak{f}, w\rangle_{g(v)} \,, \\ \mathcal{S}_\sigma(v;\mathfrak{f}) := \inf\limits_{w\in\mathcal{H}} \mathcal{S}^*(\sigma, v; w) \qquad\qquad v\in\tilde{\mathcal{H}}, \quad \sigma>0 \end{cases}$$

and note that the infimum is attained due to 1.3-(2), i.e. we can define

$$J_\sigma(v;\mathfrak{f}) := \operatorname{argmin}_{w\in\mathcal{H}} \mathcal{S}^*(\sigma, v; w, \mathfrak{f}), \qquad v\in\tilde{\mathcal{H}}\,. \tag{5.1}$$

Like in [30], we denote v_σ the generic element of $J_\sigma(v)$ and define

$$\mathrm{d}^+_\sigma(v;\mathfrak{f}) := \sup_{v_\sigma\in J_\sigma(v;\mathfrak{f})} \|v_\sigma - v\|_{g(v)}\,, \qquad \mathrm{d}^-_\sigma(v;\mathfrak{f}) := \inf_{v_\sigma\in J_\sigma(v;\mathfrak{f})} \|v_\sigma - v\|_{g(v)}\,.$$

Then, we can get similar estimates on $\mathcal{S}$ as in [2, 30], namely

Lemma 5.1. *For every $v\in\mathcal{H}$ and every $0<\sigma_1<\sigma_2$ there holds*

$$\mathcal{S}_{\sigma_2}(v;\mathfrak{f}) \le \mathcal{S}_{\sigma_1}(v;\mathfrak{f}) \le \mathcal{S}(v) - \langle\mathfrak{f}, v\rangle_{g(v)}\,, \qquad d(v_{\sigma_1}, v) \le d(v_{\sigma_2}, v)$$
$$d^+_{\sigma_1}(v;\mathfrak{f}) \le d^-_{\sigma_2}(v;\mathfrak{f}) \le d^+_{\sigma_2}(v;\mathfrak{f})$$

In particular, for every $v\in\mathcal{H}$ there exists an (at most) countable set $\mathcal{N}_v\subset(0,\tau_*)$ such that

$$d^+_\sigma(v;\mathfrak{f}) = d^-_\sigma(v;\mathfrak{f}) \qquad \forall\sigma\in(0,\tau_*)\backslash\mathcal{N}_v\,.$$

Finally, for every $v\in D(\mathcal{S})$ we have

$$\lim_{\sigma\searrow 0} d^+_\sigma(v;\mathfrak{f}) = 0, \qquad \lim_{\sigma\searrow 0}\mathcal{S}_\sigma(v;\mathfrak{f}) = \lim_{\sigma\searrow 0}\inf_{v_\sigma\in J_\sigma(v;\mathfrak{f})} \mathcal{S}(v_\sigma) = \mathcal{S}(v) - \langle\mathfrak{f}, v\rangle_{g(v)}$$

$$g\in\nabla_v\mathcal{S}(v) \quad\Rightarrow\quad \lim_{\sigma\to 0}\frac{d^+_\sigma(v;\mathfrak{f})}{\sigma} = 0\,.$$

The proof is the same as for Lemma 4.1 in [30] replacing $|a-b|$ with $\|a-b\|_{g(v)}$, $\langle a,b\rangle$ with $\langle a,b\rangle_{g(v)}$ and $\partial\mathcal{S}$ with $\nabla\mathcal{S}$. For similar reasons it is immediate consequence of the proof of Lemma 4.2 in [30], that also the following Lemma holds:

Lemma 5.2. *Under the present assumptions, we have that for every $v\in D(\mathcal{S})$, the map $\sigma\mapsto\mathcal{S}_\sigma(v)$ is locally Lipschitz on $(0,\tau_*)$ and for $\mathcal{N}_v$ as in Lemma 5.1*

$$\frac{d}{d\sigma}\mathcal{S}_\sigma(v;\mathfrak{f}) = -\frac{d^+_\sigma(v;\mathfrak{f})^2}{2\sigma^2} = -\frac{d^-_\sigma(v;\mathfrak{f})^2}{2\sigma^2} \qquad \forall\sigma\in(0,\tau_*)\backslash\mathcal{N}_v\,.$$

In particular, we have

$$\frac{\|v_{\sigma_0} - v\|^2_{g(v)}}{2\sigma_0} + \frac{1}{2}\int_0^{\sigma_0}\frac{d^+_\sigma(v;\mathfrak{f})^2}{\sigma^2}d\sigma = \mathcal{S}(v) - \mathcal{S}(v_{\sigma_0}) \tag{5.2}$$

for all $\sigma_0\in(0,\tau_*)$ and $v_{\sigma_0}\in J_{\sigma_0}(v;\mathfrak{f})$.

5.1. **The approximation scheme.** For fixed $0<T<\infty$ and time step $0<\tau<\tau_*$, there corresponds a partition of $(0,T)$ as

$$t_0 := 0 < t_1 < \cdots < t_j < \cdots < \cdots < t_{N-1} < T \le t_N, \qquad t_j := j\tau, \quad N\in\mathbb{N}\,.$$

We set

$$\overline{F}_\tau(t) := F^n_\tau = \frac{1}{\tau}\int_{t_{j-1}}^{t_j} f(s)ds \quad \text{for } t\in(t_{j-i}, t_j], \quad j = 1,\dots,N$$

and note that

$$\tau\left|\overline{F}_\tau(t)\right|^2 \le \|f\|^2_{L^2(t_{j-1},t_j;\mathcal{H})} \qquad \forall t\in(t_{j-1},t_j)$$
$$\left\|\overline{F}_\tau\right\|^2_{L^2(t_m,t_n;\mathcal{H})} \le \|f\|^2_{L^2(t_m,t_n;\mathcal{H})} \qquad \forall 1\le m<n\le N$$
$$\overline{F}_\tau \to f \quad \text{as } \tau\to 0 \text{ strongly in } L^2(0,T;\mathcal{H})\,.$$

As for the approximation of u, consider $U_\tau^0 := u_0$ for all τ and for any $j=1,\dots,N$

$$U_\tau^j \in J_\tau(U_\tau^{j-1};F_\tau^j)\,. \tag{5.3}$$

We introduce the piecewise constant interpolant $\overline{U}_\tau$ and the linear interpolant U_τ through

$$\overline{U}_\tau(t) := U_\tau^j,\quad U_\tau(t) := \frac{t_j-t}{\tau}U_\tau^{j-1} + \frac{t-t_{j-t}}{\tau}U_\tau^j, \qquad t\in(t_{j-1},t_j]\,, \tag{5.4}$$

as well as the so called De Giorgi variational interpolant $\tilde{U}_\tau$ through $\tilde{U}_\tau(0)=u_0$ and

$$\tilde{U}_\tau(t)\in J_\sigma(U_\tau^{j-1},F_\tau^j) \qquad \text{for } t=t_{j-1}+\sigma\in(t_{j-1},t_j] \tag{5.5}$$

in a way that $\tilde{U}_\tau$ is Lebesgue-measurable. The existence of such measurable selection is guarantied by [10] Corollary III.3, Theorem III.6, as well as by upper continuity of $\sigma\mapsto J_\sigma(U_\tau^{j-1},F_\tau^j)$ and compacticity of its values.

Note that this minimality implies

$$\frac{1}{\tau}\left\langle U_\tau^j - U_\tau^{j-1},\varphi\right\rangle_{g(U_\tau^{j-1})} + \langle d\mathcal{S},\varphi\rangle_{\mathcal{H}} - \left\langle F_\tau^j,\varphi\right\rangle_{g(U_\tau^{j-1})} \ni 0 \qquad \forall\varphi\in\mathcal{H}$$

or, differently:

$$U_\tau'(t) + \tilde{g}^{-1}_{\bar{U}_\tau(t_{j-1})}\left(d\mathcal{S}(\bar{U}(t))\right) - \overline{F}_\tau(t) \ni 0 \qquad \forall t\in(t_{j-1},t_j],\ j=1,\dots N-1 \tag{5.6}$$

For convenience of notation, we write

$$|U_\tau'|(t) := \frac{1}{\tau}\left\|U_\tau^j - U_\tau^{j-1}\right\|_{g(U_\tau^{j-1})} \qquad \text{for } t\in(t_{j-1},t_j]$$

We finally introduce

$$\Theta_\tau(t) := \frac{\tilde{U}_\tau(t)-U_\tau^{j-1}}{t-t_{j-1}} \qquad \text{for } t\in(t_{j-1},t_j] \tag{5.7}$$

satisfying

$$g_{U_\tau^{j-1}}(\Theta_\tau(t),\varphi) + g_{U_\tau^{j-1}}(\nabla_{U_\tau^{j-1}}\mathcal{S}(\tilde{U}_\tau(t)),\varphi) - g_{U_\tau^{j-1}}(F_\tau^j,\varphi) \ni 0 \qquad \text{for } t\in(t_{j-1},t_j],\quad \forall\varphi\in\mathcal{H} \tag{5.8}$$

and

$$G_\tau(t) := \frac{\mathrm{d}^+_\sigma(U_\tau^{j-1},F_\tau^j)}{\sigma} \ge \frac{\left\|\tilde{U}_\tau(t)-U_\tau^{j-1}\right\|_{g(U_\tau^{j-1})}}{t-t_{j-1}} = \|\Theta_\tau(t)\|_{g(U_\tau^{j-1})} \qquad \text{for } t=t_{j-1}+\sigma\in(t_{j-1}t_j]\,. \tag{5.9}$$

5.2. **A priori estimates.** Recalling definition 1.5, it is the aim of this section to show that $GMM(\mathcal{S},u_0)$ is not empty, provided $\mathcal{S}$ is "good enough". This will be topic of Lemma 5.5, which is a generalization of lemma 1.6. First, we find a priori estimates for the discrete solutions according to Lemma 5.4 which will be proved using the following estimate

Lemma 5.3. [30, 4.5] *Let B, b and κ be positive constants fulfilling $1-b\ge\frac{1}{\kappa}>0$ and let $\{a_n\}\subset[0,\infty)$ be a sequence satisfying*

$$a_n \le B + b\sum_{k=1}^n a_k \qquad \forall n\in\mathbb{N}\,.$$

Then, $\{a_n\}$ can be bounded by $a_n\le\kappa B\exp(\kappa bn)$.

Lemma 5.4. *Let $\tau \in (0, \tau_*/10)$, let $\left\{U_\tau^j\right\}_{j\in\mathbb{N}}$, U_τ and G_τ be defined through (5.3), (5.4) and (5.9). Then, for each couple of integers $1 \le i \le j$ we have*

$$\frac{1}{2}\int_{t_i}^{t_j} \left|U_\tau'\right|^2 + \frac{1}{2}\int_{t_i}^{t_j} G_\tau^2 + \mathcal{S}(U_\tau^j) = \mathcal{S}(U_\tau^i) + \int_{t_i}^{t_j} \left\langle \overline{F}_\tau, \varphi\right\rangle_{g(\overline{U}_\tau)} . \tag{5.10}$$

Moreover, there is a constant C only depending on u_0, τ_* and S_0 such that

$$\left\|\bar{U}_\tau\right\|_{L^\infty(0,T;\mathcal{H})} + \left\|U_\tau\right\|_{L^\infty(0,T;\mathcal{H})} + \left\|\tilde{U}_\tau\right\|_{L^\infty(0,T;\mathcal{H})} \le C\,, \tag{5.11}$$

$$\sup_{[0,T]} \left(\mathcal{S}(\bar{U}_\tau), \mathcal{S}(\tilde{U}_\tau)\right) \le C\,, \tag{5.12}$$

$$\left\|U_\tau'\right\|_{L^2(0,T;\mathcal{H})} + \left\|\tilde{\Theta}_\tau\right\|_{L^2(0,T;\mathcal{H})} \le C\,, \tag{5.13}$$

$$\left\|\bar{U}_\tau - U_\tau\right\|_{L^\infty(0,T;\mathcal{H})} = C\tau^{1/2}\,, \quad \left\|\tilde{U}_\tau - U_\tau\right\|_{L^\infty(0,T;\mathcal{H})} = C\tau^{1/2}\,. \tag{5.14}$$

$$\left\|U_\tau\right\|_{L^2(0,T;\mathcal{H}_0)} \le C \tag{5.15}$$

Proof. Equation (5.2) with definition of G_τ yields by choosing $t \in (t_{j-1}, t_j]$, $\sigma_0 = \tau$, $v = U_\tau^{j-1}$, $v_\sigma = \tilde{U}(t_{j-1} + \sigma)$, $0 < \sigma < \tau$ and $u_{\sigma_0} = U_\tau^j$:

$$\frac{\left\|U_\tau^j - U_\tau^{j-1}\right\|_{g(U_\tau^{j-1})}^2}{2\tau} + \frac{1}{2}\int_{t_{j-1}}^{t_j} G_\tau^2(t)\,dt = \mathcal{S}(U_\tau^{j-1}) - \mathcal{S}(U_\tau^j) + \int_{t_{j-1}}^{t_j} \left\langle \overline{F}_\tau, \varphi\right\rangle_{g(U_\tau^{j-1})} . \tag{5.16}$$

In particular, the last equation yields (5.10) by adding up over the particular subintervals of the partition.

Note that (5.16) also implies

$$\frac{\left\|U_\tau^j - U_\tau^{j-1}\right\|_{\mathcal{H}}^2}{4\tau} \le \sqrt{G^*}\left(\mathcal{S}(U_\tau^{j-1}) - \mathcal{S}(U_\tau^j) + \tau\left\|F_\tau^j\right\|_{\mathcal{H}}^2\right)$$

and similar to [30], for every $0 < \tau < \tau_*$ and every $n = 1, \dots, N$ holds

$$\begin{aligned}
\frac{1}{2}\left\|U_\tau^n\right\|_{\mathcal{H}}^2 - \frac{1}{2}\left\|u_0\right\|_{\mathcal{H}}^2 &= \sum_{k=1}^n \left(\frac{1}{2}\left\|U_\tau^k\right\|_{\mathcal{H}}^2 - \frac{1}{2}\left\|U_\tau^{k-1}\right\|_{\mathcal{H}}^2\right) \le \sum_{k=1}^n \left(\left\|U_\tau^k\right\|_{\mathcal{H}}^2 - \left\|U_\tau^k\right\|_{\mathcal{H}}\left\|U_\tau^{k-1}\right\|_{\mathcal{H}}\right)\\
&\le \sum_{k=1}^n \left\|U_\tau^k\right\|_{\mathcal{H}}\left\|U_\tau^k - U_\tau^{k-1}\right\|_{\mathcal{H}} \le \frac{\eta}{2\tau}\sum_{k=1}^n \left\|U_\tau^k - U_\tau^{k-1}\right\|_{\mathcal{H}}^2 + \frac{\tau}{2\eta}\sum_{k=1}^n \left\|U_\tau^k\right\|_{\mathcal{H}}^2\\
&\le 2\eta\sqrt{G^*}\left(\mathcal{S}(u_0) + S + \frac{1}{2\tau_*}\left\|U_\tau^n\right\|_{\mathcal{H}}^2 + \left\|f\right\|_{L^2(0,T;\mathcal{H})}^2\right) + \frac{\tau}{2\eta}\sum_{k=1}^n \left\|U_\tau^k\right\|_{\mathcal{H}}^2
\end{aligned}$$

where we used Young's inequality for some $\eta > 0$ and (1.12). Choosing $\eta = \tau_*/4$ in the last inequality yields

$$\left\|U_\tau^n\right\|_{\mathcal{H}}^2 \le 2\left\|u_0\right\|_{\mathcal{H}}^2 + 2\tau_*\sqrt{G^*}\left(\mathcal{S}(u_0) + S + \left\|f\right\|_{L^2(0,T;\mathcal{H})}^2\right) + \frac{8\tau}{\tau_*}\sum_{k=1}^n \left\|U_\tau^k\right\|_{\mathcal{H}}^2 .$$

Applying Lemma 5.3 with $\kappa = 5$ (i.e. $\tau < \tau_*/10$) yields the L^∞-bounds on $\bar{U}_\tau$ and U_τ.

It is straight forward to conclude (5.13) and estimate (5.12) on $\mathcal{S}(\bar{U}_\tau)$ from (5.10), (1.12), (5.11) and (5.9).

Then, (5.9) and (5.13) yield for all $t \in (t_{j-1}, t_j]$

$$\begin{aligned}
&\left\|U_\tau^{j-1} - U_\tau(t)\right\|_{g(U_\tau^{j-1})}^2 \le \tau\int_{t_{j-1}}^{t_j} \left|U_\tau'(s)\right|^2 ds \le C\tau\\
&\left\|U_\tau^{j-1} - \tilde{U}_\tau(t)\right\|_{g(U_\tau^{j-1})}^2 \le C\tau\,.
\end{aligned}$$

This yields (5.16), the L^∞ bound on $\tilde{U}$ as well as (5.12) for $\mathcal{S}(\tilde{U})$.

Finally, note that we obtain by means of (5.6), (5.13), lemma 5.6 and definitions 1.1-(2), 1.3-(3):

$$2\left(G^*\right)^2\left(\left\|\overline{F}_\tau^j\right\|_{g(U_\tau^{j-1})}^2 + \frac{\left\|U_\tau^j - U_\tau^{j-1}\right\|_{g(U_\tau^{j-1})}^2}{\tau^2}\right) + \left(\mathcal{S}(U_\tau^j) + 1\right) \gtrsim \left(\left\|d\mathcal{S}(U_\tau^j)\right\|_{\mathcal{H}}^2 + \left|\partial\mathcal{S}\right|^2(U_\tau^j) + \mathcal{S}(U_\tau^j) + 1\right) \gtrsim \left\|U_\tau^j\right\|_{\mathcal{H}_0}^2 ,$$

where we write $\lesssim, \gtrsim$ instead of $\leq, \geq$ up to a constant independent on τ and U_0. Eventually using (5.6) and the regularity of $g_\bullet$, this implies for the linear interpolation U_τ and $j > 1$:

$$
\begin{aligned}
\int_{t_{j-1}}^{t_j} \|U_\tau\|_{\mathcal{H}_0}^2 &\lesssim 2\tau \left(\left\|U_\tau^j\right\|_{\mathcal{H}_0}^2 + \left\|U_\tau^{j-1}\right\|_{\mathcal{H}_0}^2 \right) \\
&\lesssim 2\tau \left(\frac{\left\|U_\tau^j - U_\tau^{j-1}\right\|_{g(U_\tau^{j-1})}^2}{\tau^2} + \left\|\overline{F}_\tau^j\right\|_{g(U_\tau^{j-1})}^2 + \left(\left\|d\mathcal{S}(U_\tau^j)\right\|_{\mathcal{H}}^2 + \mathcal{S}(U_0) + 1 \right) \right) \\
&\quad + 2\tau \left(\frac{\left\|U_\tau^{j-1} - U_\tau^{j-2}\right\|_{g(U_\tau^{j-2})}^2}{\tau^2} + \left\|\overline{F}_\tau^{j-1}\right\|_{g(U_\tau^{j-2})}^2 + \left(\left\|d\mathcal{S}(U_\tau^{j-1})\right\|_{\mathcal{H}}^2 + \mathcal{S}(U_0) + 1 \right) \right) \\
&\lesssim \tau \left(\frac{1}{\tau} \int_{t_{j-1}}^{t_j} |U_\tau'|^2 + \frac{1}{\tau} \int_{t_{j-2}}^{t_{j-1}} |U_\tau'|^2 + \left\|\overline{F}_\tau^{j-1}\right\|_{\mathcal{H}}^2 + \left\|\overline{F}_\tau^{j-2}\right\|_{\mathcal{H}} + (\mathcal{S}(U_0) + 1) \right)
\end{aligned}
$$

and therefore:

$$
\int_0^T \|U_\tau\|_{\mathcal{H}_0}^2 \leq C \left(\int_0^T |U_\tau'|^2 + \tau \left\|U_\tau^0\right\|_{\mathcal{H}_0}^2 + (T + \tau)\left(S(U_0) + 1\right) + \|f\|_{L^2(0,T;\mathcal{H})} \right),
$$

with C not depending on τ or U_0, yielding (5.15) through (5.10). □

Lemma 1.6 is a consequence of the following more detailed result:

Lemma 5.5. *Let $\left(\mathcal{H}_0, \tilde{\mathcal{H}}, \mathcal{H}, g\right)$ be an entropy space and $\mathcal{S}$ a corresponding entropy functional and let the data satisfy $u_0 \in \mathcal{H}_0$. Then, $GMM(\mathcal{S}, u_0)$ is not empty. In particular, we find a sequence $k \mapsto \tau_k \searrow 0$ of timesteps, $u \in H^1(0,T;\mathcal{H})$, a non-increasing function $\varphi : [0,T] \to \mathbb{R}$ and two parameterized Young-measures $\boldsymbol{\mu} = \{\mu_t\}_{t\in(0,T)}$, $\boldsymbol{\nu} = \{\nu_t\}_{t\in(0,T)}$ on $\mathcal{H}$, such that as $k \to \infty$:*

$$\bar{U}_{\tau_k}, U_{\tau_k}, \tilde{U}_{\tau_k} \to u \quad \text{strongly in } L^\infty(0,T;\mathcal{H}), \tag{5.17}$$

$$U_{\tau_k}' \rightharpoonup u' \quad \text{weakly in } L^2(0,T;\mathcal{H}), \tag{5.18}$$

$$U_{\tau_k} \rightharpoonup u \quad \text{weakly in } L^2(0,T;\mathcal{H}_0), \quad U_{\tau_k} \to u \quad \text{strongly in } L^2(0,T;\tilde{\mathcal{H}}) \tag{5.19}$$

$$\bar{U}_{\tau_k}(t) \to u(t) \quad \text{in } \tilde{H} \text{ for a.e.} t \in (0,T). \tag{5.20}$$

$$\begin{gathered} \mathcal{S}(\bar{U}_{\tau_k}(t)) \to \varphi(t) \geq \mathcal{S}(u(t)) \quad \text{for a.e. } t \in [0,T] \\ \varphi(0) = \mathcal{S}(u_0) \end{gathered} \tag{5.21}$$

$$\begin{gathered} \boldsymbol{\mu} \text{ resp. } \boldsymbol{\nu} \text{ are the limit Young measures associated with } U_{t_k}', \text{ resp. } \tilde{\Theta}_{\tau_k}, \\ \text{and } \mu_t,\, \nu_t \text{ are concentrated on } f - \tilde{g}_u^{-1}(d_l\mathcal{S}(u(t))) \text{ for a.e. } t \in (0,T), \end{gathered} \tag{5.22}$$

(5.23)
$$\frac{1}{2}\int_s^t \left(\int_{\mathcal{H}} g_u(\xi,\xi)\, d\mu_r(\xi) + \int_{\mathcal{H}} g_u(\xi,\xi)\, d\nu_r(\xi) \right) dr + \varphi(t) \leq \varphi(s) + \int_s^t \langle f, u' \rangle_{g(u)} \qquad \text{for a.e. } 0 \leq s < t \leq T$$
(5.24)
$$\frac{1}{2}\int_s^t \left(\int_{\mathcal{H}} \liminf_{k\to\infty} \left|U_{\tau_k}'(r)\right|^2 dr + \int_{\mathcal{H}} g_u(\xi,\xi)\, d\nu_r(\xi) \right) dr + \varphi(t) \leq \varphi(s) + \int_s^t \langle f, u' \rangle_{g(u)} \qquad \text{for a.e. } 0 \leq s < t \leq T$$

Proof. Equation (5.13) together with

$$\|U_\tau(t) - U_\tau(s)\|_{\mathcal{H}} \leq |t-s|^{1/2} \|U_\tau'\|_{L^2(0,T;\mathcal{H})}$$

yields equicontinuity of $\{U_\tau\}$ for $0 < \tau \leq \tau_*/10$. As $u \mapsto \mathcal{S}(u) + \frac{1}{2\tau}\|u\|_{\mathcal{H}}^2$ has compact sublevels, estimates (5.11), (5.12) yield relative compactness of $\{U_\tau(t)\}_\tau$ for all $t \in (0,T]$. The Arzela-Ascoli theorem then yields relative compactness of $\{U_\tau\}_\tau$ in $C([0,T];\mathcal{H})$.

Estimates (5.11) and (5.14) then clearly yield convergence of (a subsequence of) U_{τ_k}, $\bar{U}_{\tau_k}$ and $\tilde{U}_{\tau_k}$ in $L^2(0,T;\mathcal{H})$ to a function u with $u \in H^1(0,T;\mathcal{H})$ where $U_{\tau_k}' \rightharpoonup u'$ weakly in $L^2(0,T;\mathcal{H})$. Also, due to (5.15), we get for a further subsequence $U_{\tau_k} \rightharpoonup u$ weakly in $L^2(0,T;\mathcal{H}_0)$ and by the Aubin-Lions theorem $U_{\tau_k} \to u$ strongly in $L^2(0,T;\tilde{\mathcal{H}})$. Passing to a further subsequence, we find $U_{\tau_k}(t) \to u(t)$ pointwise in $\tilde{\mathcal{H}}$.

Due to the pointwise convergence in $\tilde{\mathcal{H}}$ almost everywhere of the subsequence U_{τ_k}, Egorov's theorem 2.1 yields for all $\varepsilon > 0$ a compact set $K_\varepsilon \subset (0,T)$ with $\mathcal{L}((0,T)\backslash K_\varepsilon) < \varepsilon$ such that $\|U_{\tau_k} - u\|_{\tilde{\mathcal{H}}} \to 0$ uniformly on K_ε. Note that by construction, $U_{\tau_k} \in C([0,T];\tilde{\mathcal{H}})$ and thus $U_{\tau_k} \to u$ in $C(K_\varepsilon;\tilde{\mathcal{H}})$.

Using the Arzela-Ascoli theorem, we find for each $\tilde{\varepsilon} > 0$ and each $t \in K_\varepsilon$ a constant $\delta(t) > 0$ such that $\|U_{\tau_k}(t_1) - U_{\tau_k}(t_2)\|_{\tilde{\mathcal{H}}} < \frac{\tilde{\varepsilon}}{2}$ for $t_1, t_2 \in B_{\delta(t)}(t) \cap K_\varepsilon$ and all $k \in \mathbb{N}$, where $B_{\delta(t)}(t) = (t - \delta(t), t + \delta(t))$. We find $\tilde{t}_1, \dots, \tilde{t}_m$, $m < \infty$, such that $K_\varepsilon \subset \bigcup_{i=1}^m B_{\delta(\tilde{t}_i)/2}(\tilde{t}_i)$ and note that for $\tau_k < \min_{i=1,\dots,m} \delta(\tilde{t}_i)/2$ $\left\|\bar{U}_{\tau_k}(t) - U_{\tau_k}(t)\right\|_{\tilde{\mathcal{H}}} < \tilde{\varepsilon}$ for all $t \in K_\varepsilon$. Thus, $\bar{U}_{\tau_k}(t) \to u(t)$ in $\tilde{\mathcal{H}}$ for all $t \in K_\varepsilon$ and $\bar{U}_{\tau_k} \to u$ in $L^2(K_\varepsilon;\tilde{\mathcal{H}})$. As $\varepsilon \to 0$, we therefore find $\bar{U}_{\tau_k}(t) \to u(t)$ in $\tilde{\mathcal{H}}$ for a.e. $t \in (0,T)$ as $k \to \infty$.

For later purpose, we note that

$$\lim_{k\to\infty} \int_0^T g_{\bar{U}_{\tau_k}(t)}(U'_{\tau_k}(t), \psi(t))\, dt = \int_0^T g_u(u'(t), \psi(t))\, dt \quad \forall \psi \in L^2(0,T;\mathcal{H})\,, \tag{5.25}$$

which holds as $\bigcup_{\varepsilon>0} \left\{\chi_{K_\varepsilon}\psi \,:\, \psi \in L^2(0,T;\mathcal{H})\right\}$ is dense in $L^2(0,T;\mathcal{H})$ and the functionals $\int_0^T g_{\bar{U}_{\tau_k}(t)}(U'_{\tau_k}(t), \psi(t))\, dt$ are uniformly bounded.

Like in the proof of proposition 4.7 in [30] Helly's theorem (compare e.g. for [2] Lemma 3.3.3) and the lower semicontinuity of $\mathcal{S}$ yield the existence of φ satisfying (5.21).

As for the definitions (5.4) and (5.7), we find with the above estimates and (5.9) as well as theorem 4.7 two Young measures $\boldsymbol{\mu}, \boldsymbol{\nu} \in \mathcal{Y}(0,T;\mathcal{H})$ associated with U'_{τ_k} and Θ_{τ_k} such that $U'_{\tau_k} \rightharpoonup \int_{\mathcal{H}} \xi d\mu_t(\xi)$ and $\Theta_{\tau_k} \rightharpoonup \int_{\mathcal{H}} \xi d\nu_t(\xi)$ weakly in $L^2(0,T;\mathcal{H})$. Our final aim is now to identify the sets of concentration of $\boldsymbol{\mu}, \boldsymbol{\nu}$ and to proof estimates (5.23)-(5.24):

We find with help of theorem 4.7 and corollary 4.5 that

$$\liminf_{k\to\infty} \int_0^T g_{\bar{U}_{\tau_k}}(U'_{\tau_k}, U'_{\tau_k}) \geq \int_0^T \int_{\mathcal{H}} g_u(\xi,\xi)\, d\mu_t(\xi) \quad \text{and} \quad \liminf_{k\to\infty} \int_0^T G^2_{\tau_k} \geq \int_0^T \int_{\mathcal{H}} g_u(\xi,\xi)\, d\nu_t(\xi)\,.$$

Also, with help of (5.6) and (5.8) as well as corollary 4.6 and remark 4.8, we find that μ_t, ν_t are concentrated on $f(t) - \tilde{g}^{-1}_{u(t)}(d_l\mathcal{S}(u(t)))$ for $t \in K_\varepsilon$ for all $\varepsilon > 0$. As ε was arbitrary, we find that (5.23)-(5.24) hold. □

Finally, we state and prove the technical lemma:

Lemma 5.6. *If $u_\sigma \in J_\sigma(v;\mathfrak{f})$, then $u_\sigma \in D\left(|\partial\mathcal{S}|\right)$,*

$$|\partial\mathcal{S}|^2(u_\sigma) \leq 2\,(G^*)^2 \left(\frac{\|u_\sigma - v\|^2_{g(v)}}{\sigma^2} + \|\mathfrak{f}\|^2_{\mathcal{H}}\right).$$

Proof. Starting from (5.1), we check that

$$\begin{aligned}\mathcal{S}(u_\sigma) - \mathcal{S}(w) &\leq \frac{1}{2\sigma}\left(\|v - w\|^2_{g(v)} - \|u_\sigma - v\|^2_{g(v)} + 2\sigma\,\langle\mathfrak{f}, u_\sigma - w\rangle_{g(v)}\right)\\ &\leq \frac{1}{2\sigma}\,\|u_\sigma - w\|_{g(v)}\left(\|w - v\|_{g(v)} + \|u_\sigma - v\|_{g(v)} + 2\sigma\,\|\mathfrak{f}\|\right)\end{aligned}$$

for all $w \in D(\mathcal{S})$. Dividing this equation by $\|u_\sigma - w\|_{g(v)}$ yields

$$\begin{aligned}|\partial\mathcal{S}|(u_\sigma) &= \limsup_{w\to u_\sigma} \frac{|\mathcal{S}(u_\sigma) - \mathcal{S}(w)|}{\|u_\sigma - w\|_{g(u_\sigma)}} = \limsup_{w\to u_\sigma} \frac{|\mathcal{S}(u_\sigma) - \mathcal{S}(w)|}{\|u_\sigma - w\|_{g(v)}} \limsup_{w\to u_\sigma} \frac{\|u_\sigma - w\|_{g(v)}}{\|u_\sigma - w\|_{g(u_\sigma)}}\\ &\leq \left(\frac{\|u_\sigma - v\|_{g(v)}}{\sigma} + \|\mathfrak{f}\|_{\mathcal{H}}\right) \limsup_{w\to u_\sigma} \frac{\|u_\sigma - w\|_{g(v)}}{\|u_\sigma - w\|_{g(u_\sigma)}} \leq \left(\frac{\|u_\sigma - v\|_{g(v)}}{\sigma} + \|\mathfrak{f}\|_{\mathcal{H}}\right)(G^*)\,.\end{aligned}$$

□

6. Proof of Existence Theorems

Proof of theorem 1.7. Suppose $\mathcal{S}_{\tilde{\mathcal{H}}} \equiv 0$. Take an arbitrary limit function u of lemma 5.5 and let U_{τ_k}, $\bar{U}_{\tau_k}$ be the corresponding sequence of discrete solutions. Due to theorem 4.4, the Young measure $\boldsymbol{\mu}$ satisfies

$$U'_\tau \rightharpoonup \int_{\mathcal{H}} \xi d\mu_t(\xi) \quad \text{weakly in } L^2(0,T;\mathcal{H})\,.$$

As $U'_\tau \rightharpoonup u'$ weakly in $L^2(0,T;\mathcal{H})$ and $\boldsymbol{\mu}$ is concentrated on $f - \tilde{g}_u^{-1}(d_l\mathcal{S})$, we find $u' \in f - \tilde{g}_u^{-1}(d_l\mathcal{S})$, as this is a convex set. As $g_u(\xi,\xi)$ is convex in ξ for any $u \in \tilde{\mathcal{H}}$, we furthermore have by Jensens' inequality

$$g_u(u',u') \leq \int_{\mathcal{H}} g_u(\xi,\xi) d\mu_t(\xi)\,.$$

Also, since $\boldsymbol{\nu}$ is concentrated on $f - \tilde{g}_u^{-1}(d_l\mathcal{S})$, we find

$$|(f-\nabla_l\mathcal{S}(u))^\circ|^2 \leq \inf_{\varphi\in f-\nabla_l\mathcal{S}(u)} g_u(\varphi,\varphi) \leq g_u(\int_{\mathcal{H}} \xi d\nu_t(\xi), \int_{\mathcal{H}} \xi d\nu_t(\xi)) \leq \int_{\mathcal{H}} g_u(\xi,\xi) d\nu_t(\xi)\,.$$

Finally, if continuity condition 1.13 holds, we find $\varphi(t) = \mathcal{S}(u(t))$ for a.e. $t \in (0,T)$. This concludes the proof in the simple case.

On the other hand, suppose now, $\mathcal{S} \not\equiv 0$. It only remains to note that by (5.20)

$$\lim_{\tau\to 0} \mathcal{S}_{\tilde{\mathcal{H}}}(\overline{U}_\tau(t)) = \mathcal{S}_{\tilde{\mathcal{H}}}(u(t)) \quad \text{for a.e. } t \in (0,T)\,.$$

□

Proof of theorem 1.8. Like in the proof of lemma 5.5, using the notations

$$U^g_{\tau_k} := \tilde{g}_{\bar{U}_\tau(t_{j-1})}(U'_{\tau_k}), \qquad \Theta^g_{\tau_k} := \tilde{g}_{\bar{U}_\tau(t_{j-1})}(\Theta_{\tau_k}), \quad \forall t \in (t_{j-1}, t_j],\, j = 1,\dots N-1$$

inclusions (5.6) and (5.8) can equally be interpreted as

$$U^g_{\tau_k} \in \tilde{g}_{\overline{U}_\tau}(\overline{F}_\tau) - d\mathcal{S}(U_{\tau_k}(t_j)), \qquad \Theta^g_{\tau_k} \in \tilde{g}_{\overline{U}_\tau}(\overline{F}_\tau) - d\mathcal{S}(\tilde{U}_{\tau_k}(t_j))\,.$$

Due to theorems 4.4 and 4.9, corollary 4.6 and lemma 5.5, we find weak limit functions

$$U^g(t) = \int_{\mathcal{H}} \xi d\tilde{\mu}_t(\xi)\,, \qquad \Theta^g(t) = \int_{\mathcal{H}} \xi d\tilde{\nu}_t(\xi)$$

such that $U^g_{\tau_k} \rightharpoonup U^g$ and $\Theta^g_{\tau_k} \rightharpoonup \Theta^g$ weakly in $L^2(0,T;\mathcal{H})$ with Young measures $\tilde{\boldsymbol{\mu}} = \{\tilde{\mu}_t\}$ and $\tilde{\boldsymbol{\nu}} = \{\tilde{\nu}_t\}$ being concentrated on $\tilde{g}_{u(t)}(f(t)) - d_l\mathcal{S}(u(t))$. Furthermore, (5.25) yields $U^g(t) = \tilde{g}_{u(t)}u'(t)$ and similar for all weak limit points. Thus, we evidently find

$$\int_{\mathcal{H}} \tilde{g}_{u(t)}(\xi) d\mu_t(\xi) = \int_{\mathcal{H}} \xi d\tilde{\mu}_t(\xi)\,, \quad \int_{\mathcal{H}} \tilde{g}_{u(t)}(\xi) d\nu_t(\xi) = \int_{\mathcal{H}} \xi d\tilde{\nu}_t(\xi) \tag{6.1}$$

We use inequality (5.23) together with (1.2) and the Lebesgue differentiation theorem to get for a.e. $s \in (0,T)$

$$\frac{1}{2}\sqrt{G^*}^{-1}\left(\int_{\mathcal{H}} \|\xi\|^2_{\mathcal{H}}\, d\tilde{\mu}_t(\xi) + \int_{\mathcal{H}} \|\xi\|^2_{\mathcal{H}}\, d\tilde{\nu}_t(\xi)\right) \leq -\varphi'(t)\,.$$

Since (1.13) holds, we know $\varphi = \mathcal{S}\circ u$ a.e. in $(0,T)$ and (1.16) yields through theorem 4.9 and (5.22)

$$\begin{aligned} \varphi'(t) = \int_{\mathcal{H}} \left\langle \tilde{g}_{u(t)}(f) - \xi, u'(t)\right\rangle_{\mathcal{H}} d\tilde{\nu}_t(\xi) &= \int_{\mathcal{H}} \left\langle \tilde{g}_{u(t)}(f-\xi), u'(t)\right\rangle_{\mathcal{H}} d\nu_t(\xi) \\ &= \int_{\mathcal{H}} \left\langle f-\xi, u'(t)\right\rangle_{g(u(t))} d\nu_t(\xi) = \left\langle f-\theta(t), u'(t)\right\rangle_{g(u(t))} \end{aligned} \tag{6.2}$$

with $\theta(t) = \int_{\mathcal{H}} \xi d\nu_t(\xi)$. Combining (6.2) with the localized version of (5.23):

$$\frac{1}{2}\left(\int_{\mathcal{H}} g_u(\xi,\xi)\, d\mu_t(\xi) + \int_{\mathcal{H}} g_u(\xi,\xi)\, d\nu_t(\xi)\right) + \varphi'(t) - f \leq 0$$

yields

$$\frac{1}{2}\left(\int_{\mathcal{H}} g_u(\xi,\xi)\, d\mu_t(\xi) + \int_{\mathcal{H}} g_u(\xi,\xi)\, d\nu_t(\xi)\right) \leq \left\langle \theta(s), u'(s)\right\rangle_{g(u(s))} \quad \text{for a.e. } t \in (0,T)\,.$$

We furthermore have $u'(t) = \int_{\mathcal{H}} \eta d\mu_t(\eta)$ and

$$\int_{\mathcal{H}} \|\xi - \theta\|_{g(u)}^2 \, d\nu_t(\xi) = \int_{\mathcal{H}} \|\xi\|_{g(u)}^2 \, d\nu_t(\xi) - \|\theta\|_{g(u)}^2$$
$$\int_{\mathcal{H}} \|\eta - u'\|_{g(u)}^2 \, d\mu_t(\eta) = \int_{\mathcal{H}} \|\eta\|_{g(u)}^2 \, d\mu_t(\eta) - \|u'\|_{g(u)}^2$$

yielding finally

$$\int_{\mathcal{H}} \|\xi - \theta\|_{g(u)}^2 \, d\nu_t(\xi) + \int_{\mathcal{H}} \|\eta - u'\|_{g(u)}^2 \, d\mu_t(\eta) + \|u' - \theta\|_{g(u)}^2 \leq 0\,.$$

In particular,

$$u'(t) = \theta(t) \in f - \nabla_l \mathcal{S}(u(t))\,, \quad \mu_t = \nu_t = \delta_{u'(t)} \quad \text{for a.e. } t \in (0,T)$$

since μ_t, ν_t are concentrated on $f - \nabla_l \mathcal{S}(u(t))$ for a.e. $t \in (0,T)$ and u is a strong solution to (1.11) and satisfies (1.19).

Now, (4.3) and (6.1) yield

$$\varphi'(t) = \langle f(t) - u'(t), u'(t)\rangle_{g(u(t))} = \langle \xi, u'(t)\rangle_{g(u(t))} \quad \forall \xi \in f(t) - \tilde{g}_u^{-1}\left(\overline{\mathrm{aff}}(d_l \mathcal{S}(u(t))\right) \quad \text{for a.e. } t \in (0,T)\,.$$

Thus, denoting $K_t := f - \tilde{g}_u^{-1}\left(\overline{\mathrm{aff}}(d_l \mathcal{S}(u(t))\right)$, $u'(t)$ satisfies

$$u'(t) \in K_t, \qquad \langle u'(t), u'(t) - \eta\rangle_{g(u(t))} = 0 \quad \forall \eta \in K_t,$$

which yields (1.17). The remaining statements are proved the same way as Theorem 2 in [30] replacing $|\cdot|$ by $\|\cdot\|_{g(u)}$. □

Proof of theorem 1.9. We make use of theorem 1.8 to get that u is a solution to (1.11) with $u(0) = u_0$ and for almost all $t \in (0,T)$, (1.17)-(1.18). Finally, (1.22) follows from (1.20) similar to the proof of theorem 3 in [30] with $\|\cdot\|_{\mathcal{H}}$ replaced by $\|\cdot\|_{g(u(t))}$ and $\langle\cdot,\cdot\rangle_{\mathcal{H}}$ replaced by $\langle\cdot,\cdot\rangle_{g(u(t))}$. □

7. The Assumption of Maximum Rate of Entropy Production for Reaction Diffusion Systems

We will now summarize the ideas by Rajagopal and Srinivasa [27] of the maximization of the rate of entropy production and follow the outline of [17] by applying the method directly to reaction diffusion systems. Note that this has been carried out in much greater detail in Heida, Málek and Rajagopal [19, 20] as well as in Heida [16, 17].

Then, we will provide a theoretical link between these formal calculations and the theory of gradient flows and show that such systems automatically fall into the pattern of curves of maximal slope. To this aim, we will, however, stay in a more simple setting than in [16, 19, 20, 17], in particular, the energy flux will be of much simpler structure.

7.1. **Formal calculations.** Thus, following Callen [9], we assume the existence of a specific entropy η as a differentiable function of the internal energy e and the state variables $\boldsymbol{C} = (c_i)_{i=1,\dots,J}$. In particular, we assume $\eta = \tilde{\eta}(e, \boldsymbol{C})$ with $\tilde{\eta}$ being increasing with respect to e. The temperature is given through $\vartheta := \left(\frac{\partial \tilde{\eta}}{\partial e}\right)^{-1}$ and the time derivative of η is

$$\partial_t \eta = \vartheta^{-1} \partial_t e + \sum_{i=0}^{J} \frac{\partial \tilde{\eta}}{\partial c_i} \partial_t c_i \tag{7.1}$$

Inserting the energy and mass balance equations of the form

$$\partial_t c_i + \operatorname{div} \boldsymbol{j}_i = \overset{+}{c}_i, \qquad \partial_t e - \operatorname{div} \boldsymbol{h} = 0$$

into (7.1) we obtain an equation describing the evolution of the entropy with time and which is of the form

$$\dot{\eta} - \operatorname{div} \boldsymbol{q} = \xi\,, \tag{7.2}$$

where

$$\xi = -\sum_i \boldsymbol{j}_i \cdot \nabla\mu_i - \sum_i \overset{+}{c}_i\ \mu_i + \boldsymbol{h} \cdot \frac{\nabla\vartheta}{\vartheta^2}\,, \tag{7.3}$$

$$\boldsymbol{q} = \boldsymbol{h} + \sum_i \mu_i\, \boldsymbol{j}_i\,. \tag{7.4}$$

In (7.2), $\boldsymbol{q}$ is an entropy flux and ξ is the rate of entropy production. Although this is physically slightly improper, we also denote ξ as the rate of entropy production. The second law of thermodynamics then in particular implies

$$\xi \geq 0$$

for all times.

Brought into the language of the present article, Rajagopal and Srinivasa [27] assumed that the rate of entropy production can be prescribed by a nonnegative function $\tilde{\xi}$ such that the second law of thermodynamics is automatically fulfilled:

$$\xi = \tilde{\xi}(\mathcal{J},\mathcal{C}) \geq 0, \qquad \mathcal{J} = (\boldsymbol{j}_1,\dots \boldsymbol{j}_J,\boldsymbol{h}), \qquad \mathcal{C} = (\overset{+}{c}_1,\dots \overset{+}{c}_J)\,. \tag{7.5}$$

The choice of such a constitutive relation is not an easy task and requires apriori knowledge (or at least well justified hypotheses) on the dissipative processes in the particular material under consideration. Of course, with respect to (7.3), $\tilde{\xi}$ also has to fulfill the equality

$$\tilde{\xi}(\mathcal{J},\mathcal{C}) = -\sum_i \boldsymbol{j}_i \cdot \nabla\mu_i - \sum_i \overset{+}{c}_i\ \mu_i + \boldsymbol{h} \cdot \frac{\nabla\vartheta}{\vartheta^2}\,, \tag{7.6}$$

Rajagopal and Srinivasa claimed that constitutive equations for $\mathcal{J}$ can be obtained maximizing with respect to the thermodynamical fluxes, the maximization problem reads

$$\max_{\mathcal{J},\mathcal{C}} \tilde{\xi}\,(\mathcal{J},\mathcal{C}) \quad \text{provided (7.6) holds.}$$

We will not dig into the details of a rather trivial calculation to note that for a simple choice

$$\tilde{\xi}(\mathcal{J},\mathcal{C}) = \sum_i \alpha_i^{-1}\,|\boldsymbol{j}_i|^2 + \sum_i \beta_i^{-1}\left(\overset{+}{c}_i\right)^2 + \kappa^{-1}\,|\boldsymbol{h}|^2 \tag{7.7}$$

would lead to the set of constitutive equations

$$\boldsymbol{j}_i = -\alpha_i\nabla\mu_i, \qquad \overset{+}{c}_i = -\beta_i\mu_i, \qquad \boldsymbol{h} = \frac{\kappa}{\vartheta}\nabla\vartheta \tag{7.8}$$

Now, for two sets of variables (c_i^1, e^1) and (c_i^2, e^2) satisfying equations of the form

$$\partial_t c_i^j + \operatorname{div}\boldsymbol{j}_i^j = \overset{+j}{c}_i, \qquad \partial_t e^j - \operatorname{div}\boldsymbol{h}^j = 0 \qquad \text{for } j = 1, 2$$

with boundary conditions

$$\boldsymbol{q}\cdot\boldsymbol{n}_\Gamma = 0\,, \quad \boldsymbol{j}_i\cdot\boldsymbol{n}_\Gamma = 0\,, \quad \boldsymbol{h}\cdot\boldsymbol{n}_\Gamma = 0 \tag{7.9}$$

we define the (not necessarilly uniquely defined) bilinear form

$$g\left((\partial_t c_i^1, \partial_t e^1), (\partial_t c_i^2, \partial_t e^2)\right) := \int_\Omega \sum_i \alpha_i^{-1}\,|\boldsymbol{j}_i|^2 + \sum_i \beta_i^{-1}\left(\overset{+}{c}_i\right)^2 + \kappa^{-1}\,|\boldsymbol{h}|^2\ .$$

It is interesting to note that with

$$\mathcal{S} := \int_\Omega \eta\,,$$

integrating (7.2) over Ω with boundary conditions $\boldsymbol{q}\cdot\boldsymbol{n}_\Gamma = 0$, $\boldsymbol{j}_i\cdot\boldsymbol{n}_\Gamma = 0$, $\boldsymbol{h}\cdot\boldsymbol{n}_\Gamma = 0$ we obtain the evolution equation

$$\frac{d}{dt}\mathcal{S} = -g\left((\partial_t c_i, \partial_t e), (\partial_t c_i, \partial_t e)\right)\ .$$

It is important to note that in the above quadratic setting, the choice of $\tilde{\xi}(\dots)$ is not limited to (7.7). Indeed, let $\tilde{\xi}(\mathcal{J},\mathcal{C}) = \tilde{\xi}_1(\mathcal{J}) + \tilde{\xi}_2(\mathcal{C})$ any positive definite matrices $\mathbb{A}_i = \mathbb{Q}_i^T \mathbb{D}_i \mathbb{Q}_i$ (of correct dimension) with orthogonal matrices $\mathbb{Q}_i$ and a diagonal matrices $\mathbb{D}_i$ for $i = 1, 2$ would define strictly positive $\tilde{\xi}_i(\cdot)$ via

$$\tilde{\xi}_1(\mathcal{J}) = \frac{1}{2}\mathcal{J}^T \mathbb{A}_1 \mathcal{J} = \frac{1}{2}\mathcal{J}_{\mathbb{Q}}^T \mathbb{D}_1 \mathcal{J}_{\mathbb{Q}}\,, \qquad \tilde{\xi}_2(\mathcal{C}) = \frac{1}{2}\mathcal{C}_{\mathbb{Q}}^T \mathbb{D}_2 \mathcal{C}_{\mathbb{Q}}$$

where $\mathcal{J}_{\mathbb{Q}} = \mathbb{Q}_1 \mathcal{J}$, $\mathcal{C}_{\mathbb{Q}} = \mathbb{Q}_2 \mathcal{C}$. We may then proceed following the above ansatz.

Example 7.1. (See [17, Appendix A]) We consider a single diffusion system without reaction terms, i.e. $\overset{+}{c}_i = 0$. In what follows, we write $\mathbb{I}_3$ for the identity in $\mathbb{R}^3$. Let $\boldsymbol{a}_1, \boldsymbol{a}_2 \in \mathbb{R}^3$ be two vectors in the three dimensional space. Then, by $\boldsymbol{a} := \begin{pmatrix} \boldsymbol{a}_1 \\ \boldsymbol{a}_2 \end{pmatrix}$ we denote the element $\boldsymbol{a} \in \mathbb{R}^6$ such that the first three coordinates are given by $\boldsymbol{a}_1$ and the last three coordinates are given by $\boldsymbol{a}_2$. Similar, for $\mathbb{A}_i \in \mathbb{R}^{3\times 3}$, $i \in \{1, \dots, 4\}$, we identify

$$\begin{pmatrix} \mathbb{A}_1 & \mathbb{A}_2 \\ \mathbb{A}_3 & \mathbb{A}_4 \end{pmatrix} \in \mathbb{R}^{6\times 6}$$

in an obvious way. Now, we assume that the system is given by two incompressible fluids of equal density, i.e. we restrict to one single component in the equations: $c := c_1$. Then, we are interested in a system

$$\partial_t c + \operatorname{div} \boldsymbol{j} = 0, \qquad \partial_t E - \operatorname{div} \boldsymbol{h} = 0$$

where we assume for E a constitutive equation of the form $E = \tilde{\vartheta}(\eta, c)$. Thus, with $\mu := \frac{\partial E}{\partial c}$, the rate entropy production takes the form

$$\xi = \boldsymbol{h} \cdot \nabla\vartheta - \boldsymbol{j} \cdot \nabla\mu, \qquad \boldsymbol{q} = \boldsymbol{h} - \mu \boldsymbol{j}\,.$$

Now, we assume that

$$\tilde{\xi}(\boldsymbol{j}, \boldsymbol{q}) = \begin{pmatrix} \boldsymbol{j} \\ \boldsymbol{q} \end{pmatrix}^T \mathbb{Q}^T \begin{pmatrix} \mathbb{I} & 0 \\ 0 & 2\mathbb{I} \end{pmatrix} \mathbb{Q} \begin{pmatrix} \boldsymbol{j} \\ \boldsymbol{q} \end{pmatrix}, \qquad \mathbb{Q} = \begin{pmatrix} \sin\varphi\mathbb{I} & \cos\varphi\mathbb{I} \\ -\cos\varphi\mathbb{I} & \sin\varphi\mathbb{I} \end{pmatrix},$$

such that $\mathbb{Q}$ is an orthogonal matrix. Thus, with

$$\boldsymbol{J} = \begin{pmatrix} \boldsymbol{J}_1 \\ \boldsymbol{J}_2 \end{pmatrix} := \begin{pmatrix} \sin\varphi \boldsymbol{j} + \cos\varphi \boldsymbol{q} \\ -\cos\varphi \boldsymbol{j} + \sin\varphi \boldsymbol{q} \end{pmatrix}, \quad \boldsymbol{A} = \begin{pmatrix} \boldsymbol{A}_1 \\ \boldsymbol{A}_2 \end{pmatrix} := \begin{pmatrix} -\sin\varphi\nabla\mu + \cos\varphi\nabla\vartheta \\ +\cos\varphi\nabla\mu + \sin\varphi\nabla\vartheta \end{pmatrix}$$

we find

$$\xi = \boldsymbol{A} \cdot \boldsymbol{J} \qquad \text{and} \qquad \tilde{\xi}(\boldsymbol{J}) = \boldsymbol{J}^T \begin{pmatrix} \mathbb{I} & 0 \\ 0 & 2\mathbb{I} \end{pmatrix} \boldsymbol{J}\,.$$

Finally, we find

$$\boldsymbol{J} = \begin{pmatrix} \mathbb{I} & 0 \\ 0 & \frac{1}{2}\mathbb{I} \end{pmatrix} \boldsymbol{A}\,, \quad \text{and} \quad \begin{pmatrix} \boldsymbol{j} \\ \boldsymbol{q} \end{pmatrix} = \mathbb{Q}^T \boldsymbol{J} = \begin{pmatrix} -\nabla\mu + \frac{1}{2}\cos\varphi\sin\varphi\nabla\vartheta \\ \nabla\vartheta - \frac{1}{2}\cos\varphi\sin\varphi\nabla\mu \end{pmatrix}. \tag{7.10}$$

These relations show the structure of Fick's original equations (see also Truesdell [31] Appendix 5B, equation (5B.4.1)).

Remark 7.2. Note that a shift $\mathcal{S} \rightsquigarrow \mathcal{S} + \int_\Omega Ce$ with $C \in \mathbb{R}$ constant does not affect the constitutive equations in (7.8) or (7.10) as (7.1) would be replaced by

$$\partial_t \eta = \left(\vartheta^{-1} + C\right) \partial_t e + \sum_{i=0}^{J} \frac{\partial\tilde{\eta}}{\partial c_i} \partial_t c_i\,.$$

7.2. **The Link to Gradient Flows.** The following theory will provide the theoretical link between gradient flows and the above maximization of the rate of entropy production. The link may not be understandable that easily, so we will carry it out in full detail in section 8.1 using a concrete example. We consider a Hilbert space $\mathcal{H}$ with scalar product $\langle\cdot,\cdot\rangle_{\mathcal{H}}$ and norm $\|\cdot\|_{\mathcal{H}}$ and assume the existence of a lower semicontinuous mapping

$$g_\bullet : \quad \mathcal{H} \to B(\mathcal{H}), \qquad u \mapsto g_u(\cdot,\cdot)$$

having the property (1.2) with G^* being independent on u.

We define for any $a < b \in \mathbb{R}$ and any function $\gamma \in C^1([a,b];\mathcal{H})$ the length $L(\gamma)$ through

$$L(\gamma) := \int_a^b \sqrt{g_{\gamma(t)}(\partial_t\gamma(t), \partial_t\gamma(t))}\, dt$$

and find that for any $a' < b' \in \mathbb{R}$ and any strictly monotone continuously differentiable mapping $s : [a', b'] \to [a, b]$ with $\gamma^* := \gamma \circ s$ there holds $L(\gamma^*) = L(\gamma)$.

This makes it possible to define a corresponding metric

$$d(x, y) := \inf_{\substack{\gamma \in C^1([0,1];\mathcal{H}) \\ \gamma(0)=x, \gamma(1)=y}} L(\gamma) \,.$$

Note that the definition of $L(\cdot)$ and $d(\cdot, \cdot)$ allow for $\gamma \in AC(0, 1; \mathcal{H})$. In particular, the path of minimal length connecting two elements of $\mathcal{H}$ is in $AC(0, 1; \mathcal{H})$.

As a consequence of the lower semicontinuity of $g_\bullet$, lower semicontinuity of $g_u(\cdot, \cdot)$ for all $u \in \mathcal{H}$, as well as Theorems 2.4.3 and 2.7.6 of [7], the space $\mathcal{H}$ together with $d(\cdot, \cdot)$ forms a length space in the sense of the following definition:

Definition 7.3. A length space is a metric space with a length function $L(\cdot)$ and a distance function $d(\cdot, \cdot)$ having the following properties

$$L : \qquad AC([0, 1], \mathcal{H}) \to \mathbb{R}$$

$$\gamma \mapsto \int_0^1 |\gamma'(t)| \; dt, \qquad |\gamma'(t)| := \limsup_{\varepsilon \to 0} \frac{d(\gamma(t), \gamma(t + \varepsilon))}{|\varepsilon|}$$

$$d(x, y) = \inf \{L(\gamma) \, : \, \gamma(0) = x, \, \gamma(1) = y\} \,.$$

The most important formal consequence of this abstract result is the equivalence

$$|\gamma'(t)| = \sqrt{g_{\gamma(t)}(\partial_t \gamma(t), \partial_t \gamma(t))} = \|\partial_t \gamma(t)\|_{g(\gamma(t))} \,.$$

Remark 7.4. As stated in section 3.2, this does not necessarily apply to $g_\bullet$ from the examples in section 3 as those may not be lower semicontinuous with respect to $\|\cdot\|_{\mathcal{H}}$.

Corollary 7.5.

$$\lim_{w \to u} \frac{\|u - w\|_{g(u)}}{d(u, w)} = 1$$

Proof. This is a consequence of the Lipschitz continuity of $g_\bullet$, definition of $L(\cdot)$ and

$$\lim_{w \to u} \frac{\|u - w\|_{g(u)}}{d(u, w)} = \lim_{w \to u} \frac{L(\gamma_w)}{d(u, w)} \cdot \lim_{w \to u} \frac{\|u - w\|_{g(u)}}{L(\gamma_w)} = 1 \cdot 1 \,,$$

where $\gamma_w(t) = u + (w - u)t$ and the first limit on the right hand side is due to [7], Corollary 2.7.5. □

Let $\mathcal{S} : \mathcal{H} \to (-\infty, +\infty]$ be a proper lower semicontinuous functional. The last corollary shows that the following definition of the local slope $|\partial \mathcal{S}|$ of $\mathcal{S}$ at $u \in D(\mathcal{S})$ coincides with (1.9)

$$|\partial \mathcal{S}| \, (u) := \limsup_{w \to u} \frac{(\mathcal{S}(u) - \mathcal{S}(w))^+}{d(u, w)} \,, \tag{7.11}$$

As we can easily calculate [2] that for any curve $\gamma \in AC(-1, 1; \mathcal{H})$

$$|(\mathcal{S} \circ \gamma)'(t)| = \left| \lim_{h \searrow 0} \frac{\mathcal{S}(\gamma(t + h)) - \mathcal{S}(\gamma(t))}{h} \right| \le |\partial \mathcal{S}| \, (\gamma(t)) \, |\gamma'(t)| \,, \tag{7.12}$$

we find

$$-\,|(\mathcal{S} \circ \gamma)'(t)| \ge -\frac{1}{2} \, |\partial \mathcal{S}|^2 - \frac{1}{2} \, |\gamma'(t)|^2 \,. \tag{7.13}$$

Definition 7.6. We say that a locally absolutely continuous curve $\gamma : (a, b) \to \mathcal{H}$ is curve of maximal slope for the functional $\mathcal{S}$ with respect to $|\partial \mathcal{S}|$ if $\mathcal{S} \circ \gamma$ is a.e. equal to a non-increasing map s and

$$s'(t) \le -\frac{1}{2} \, |\partial \mathcal{S}|^2 - \frac{1}{2} \, |\gamma'(t)|^2 \tag{7.14}$$

As a consequence of (7.13) and (7.14), any curve of maximal slope satisfies

$$-s'(t) = |\partial\mathcal{S}|^2 = |\gamma'(t)|^2 \ . \tag{7.15}$$

For the mathematical theory of curves of maximal slope, refer to [2].

Furthermore, under the additional assumption that

$$d\mathcal{S}(u) \quad \text{is single valued for all } u \in \mathcal{H} \tag{7.16}$$

$|\gamma'(t)|^2$ is maximal in the sense that if there is another absolutely continuous curve $\tilde{\gamma} : (a,b) \to \mathcal{H}$ with $\gamma(t_0) = \tilde{\gamma}(t_0)$ such that the second equality in (7.15) still holds for $|\tilde{\gamma}'(t)|$ (i.e. $-|(\mathcal{S}\circ\tilde{\gamma})'(t)| = -|\tilde{\gamma}'(t)|^2$), but with $|\tilde{\gamma}'(t)|^2 > |\gamma'(t)|^2$, we naturally recover (7.13), i.e.

$$-|(\mathcal{S}\circ\tilde{\gamma})'(t)| \geq -\frac{1}{2}|\partial\mathcal{S}|^2 - \frac{1}{2}|\tilde{\gamma}'(t)|^2 \tag{7.17}$$

but on the other hand

$$-|(\mathcal{S}\circ\tilde{\gamma})'(t)| = -\frac{1}{2}|\tilde{\gamma}'(t)|^2 - \frac{1}{2}|\tilde{\gamma}'(t)|^2 < -\frac{1}{2}|\tilde{\gamma}'(t)|^2 - \frac{1}{2}|\gamma'(t)|^2 = -\frac{1}{2}|\tilde{\gamma}'(t)|^2 - \frac{1}{2}|\partial\mathcal{S}|^2 \ ,$$

contradicting (7.17). Thus, γ satisfies

$$\partial_t\gamma(t) = \operatorname{argmax}_{s\in\mathcal{H}}\left\{g_{\gamma(t)}(s,s) \,:\, g_{\gamma(t)}(s,s) = -\langle d\mathcal{S}(\gamma(t)), s\rangle_{\mathcal{H}}\right\} \qquad \forall t \in (-1,1)\,. \tag{7.18}$$

Now, assume that $\gamma \in AC(-1,1;\mathcal{H})$ satisfies (7.18): We clearly find for $\lambda = 2$ that

$$g_{\gamma(t)}(s,s) - \lambda\left(g_{\gamma(t)}(s,s) + \langle d\mathcal{S}(\gamma(t)), s\rangle_{\mathcal{H}}\right)$$

has a local extremum in $\partial_t\gamma(t)$. Thus, $\partial_t\gamma(t)$ is the global minimizer of

$$\frac{1}{2}g_{\gamma(t)}(s,s) + \langle d\mathcal{S}(\gamma(t)), s\rangle_{\mathcal{H}}$$

meaning

$$g_{\gamma(t)}(\partial_t\gamma(t), \phi) = -\langle d\mathcal{S}(\gamma(t)), \phi\rangle_{\mathcal{H}}$$

which by (7.12) finally yields (7.15) and (7.14). Note that definition (1.6) enables us to write the last equality as

$$\partial_t\gamma(t) = -\nabla\mathcal{S}(\gamma(t)) \tag{7.19}$$

Thus, we have shown the following Lemma:

Lemma 7.7. *Assuming* (7.16)*, a curve* $\gamma \in AC(-1,1;\mathcal{H})$ *is a solution of the gradient flow equation*

$$\partial_t\gamma = -\nabla\mathcal{S}(\gamma)$$

(or equivalently a curve of maximal slope) if and only if its derivative satisfies (7.18)*.*

The last lemma is important in order to understand the relation between the formal calculations from section 7 and the rigorous calculations from section 8 below.

Note that in case $\nabla\mathcal{S}$ is set-valued, (7.19) is reformulated into

$$\partial_t\gamma(t) \in -\nabla\mathcal{S}(\gamma(t)) \tag{7.20}$$

8. Multiphase Allen-Cahn and Cahn-Hilliard Equations

We will now provide a functional analytic approach to the method sketched in section 7. The approach is closely related to recent works by the author [17] and Heida, Málek and Rajagopal [20, 19].

The necessity to study combinations of Allen-Cahn and Cahn-Hilliard equations for multiphase flow stems from physics: The Allen-Cahn equation describes phase transitions, e.g. water-ice, whereas the Cahn-Hilliard equation describes phase separation, e.g. air-water. Coupled to transport and Fouriers law, these equations enable us to model complex interactions like water-ice-air-vapor, even with dynamic boundary conditions and on multiple scales (see Heida [16, 18]).

We will start with some preliminary calculations in order to keep the proofs of the theorems below as readable as possible. In particular, we will prove a generalized version of 3.2 on the subgradients of $\mathcal{S}$ to the n components. However, for simplicity, we have to restrict to some very special classes of functionals. As mentioned in the introduction, physical entropy is concave, which is why we deal with the negative of the entropy. We then face the problem that negative entropy $\mathcal{S}$ of the form

$$\mathcal{S} = \int_\Omega \eta(e, \dots)$$

shows the sublinear behavior $\eta(e, \cdots) \to -\infty$, $\frac{\partial\eta}{\partial e}(e, \dots) \to 0$, thus not necessarily showing compactness of sublevels claimed in 1.3. As announced in remark 7.2, in a L^1-setting, we eventually could circumvent this problem by shifting $\mathcal{S} \rightsquigarrow \mathcal{S} + \int_\Omega Ce$ with $C \in \mathbb{R}$ being constant.

However, for L^2-theory, this is not enough, as the L^2-norm of $d\mathcal{S}$ could not provide sufficient regularity for e to determine the set $d_l\mathcal{S}$ or an estimate on $\|e\|_{L^2(0,T;H^1_{(0)}(\Omega))}$. Therfore, we have to make the unphysical assumption that $\lim_{e\to+\infty} e^{-2}\frac{\partial\eta}{\partial e} > 0$. Finally, note that all calculations in 8.1-8.2 will be rigorous, though they might look formal.

8.1. **A simple case.** In order to keep calculations as readable as possible, we first restrict to a dependence of $\mathcal{S}$ on energy e and mass concentration c of one of two chemical components. We assume that chemical reactions take place, i.e. it makes sense to consider $H^1(\Omega)$ instead of $H^1_{(0)}(\Omega)$. Thus, following section 3, we place ourselves in $\mathcal{H} := H^*(\Omega) \times H^{-1}_{(0)}$ with $D(\mathcal{S}) = H^1(\Omega) \times L^2(\Omega)$. Below, we will see that the best choice for $\tilde{\mathcal{H}}$ and $\mathcal{H}_0$ are

$$\tilde{\mathcal{H}} = H^1(\Omega) \times L^2_{(0)}(\Omega)\,, \qquad \mathcal{H}_0 = H^2(\Omega) \times H^1_{(0)}(\Omega)\,.$$

In particular, we think of

$$\mathcal{S} : \mathcal{H} \to \mathbb{R}, \quad (c, e) \mapsto \int_\Omega \eta(c, \nabla c, e), \qquad \eta(c, \nabla c, e) = s_0(c) + s_1(c, e) + e + \frac{\sigma}{2}|\nabla c|^2\,,$$

s_0 having the properties like in definition (3.1) and $s_1 \in C^1(\mathbb{R} \times \mathbb{R}_+; \mathbb{R}_+)$ with

$$\text{(8.1)} \qquad \begin{gathered} \liminf_{y\to+\infty} \frac{s_1(x,y)}{y^2} > 0\,, \quad \text{and} \quad \lim_{y\searrow 0} \partial_y s_1(x,y) = -\infty \qquad \forall x \in \mathbb{R} \\ |\partial_1 s_1(x,y)| + \left|\partial_1^2 s_1(x,y)\right| < C, \quad \partial_1^2 s_1(x,y) \le 0, \quad \partial_2^2 s_1(x,y) \ge 0 \quad \forall (x,y) \in \mathbb{R}^2 \end{gathered}$$

for all $(x, y) \in \mathbb{R} \times \mathbb{R}_+$.

Furthermore, note that due to the regularity of s_1 and the considerations in section 3.1, we find that $\mathcal{S}$ is strongly-weakly closed in the sense of (1.23) and therefore, theorem 1.8 is applicable for any suitable metric g.

Existence of compact sublevels in the sense of 1.3-(2) is easy to check from the definition of $\mathcal{S}$.

Definition 8.1. Let $\mathbb{L}^2 := L^2(\Omega) \times L^2_{(0)}(\Omega)$. We denote $D_{L^2}\mathcal{S}(u)$ the set of all $\delta_u \in \mathbb{L}^2$ such that for all $\phi \in \mathbb{L}^2$ and all $\gamma \in H^1(-1, 1; \mathbb{L}^2)$ with $\gamma(0) = u$ and $\gamma'(0) = \phi$ holds

$$\text{(8.2)} \qquad \langle \delta_u, \phi \rangle_{L^2(\Omega)} \le \lim_{t\to 0} \frac{1}{t}\left(\mathcal{S}(\gamma(t)) - \mathcal{S}(u)\right)\,.$$

Moreover, we denote by $\mathbb{P}_e$ and $\mathbb{P}_c$ the projections on the e and c coordinates of $L^2(\Omega)^2$ and say $\delta_c \in \frac{\delta S}{\delta c}(u)$ iff (8.2) holds for all $\phi \in \mathbb{P}_c(\mathbb{L}^2)$ and likewise for $\delta_e \in \frac{\delta S}{\delta e}$. Clearly, $\frac{\delta S}{\delta c} = \mathbb{P}_c(D_{L^2}\mathcal{S})$ and $\frac{\delta S}{\delta e} = \mathbb{P}_e(D_{L^2}\mathcal{S})$.

In terms of section 7, we search for solutions

$$\partial_t c + \operatorname{div} \boldsymbol{j} = \overset{+}{c}\,, \qquad \partial_t e - \operatorname{div} \boldsymbol{h} = 0\,,$$

where

$$\boldsymbol{j} = -A(\dots)\nabla\frac{\delta\mathcal{S}}{\delta c}, \quad \overset{+}{c} = -B(\dots)\frac{\delta\mathcal{S}}{\delta c}, \quad \boldsymbol{h} = \kappa(\dots)\nabla\frac{\delta\mathcal{S}}{\delta e}\,.$$

Clearly, $\frac{\delta S}{\delta c} = \mathbb{P}_c(D_{L^2}\mathcal{S})$ and $\frac{\delta S}{\delta e} = \mathbb{P}_e(D_{L^2}\mathcal{S})$. Note that

$$\frac{\delta\mathcal{S}}{\delta c} = \frac{\partial s_0(c)}{\partial c} + \frac{\partial s_1(c,e)}{\partial c} - \sigma\Delta c, \qquad \frac{\delta\mathcal{S}}{\delta e} = \frac{\partial\eta_0(c,e)}{\partial e}\,.$$

Note that lemma 3.2 as well as the properties of s_0 and s_1 yield $\frac{\delta S}{\delta c} \in L^2(\Omega)$. Now, let $A, \kappa \in C^{0,1}(\mathbb{R} \times \mathbb{R}^3 \times \mathbb{R}; \mathbb{R}^{3\times 3})$ be Lipschitz-continuous with constants $0 < C_- \le C_+$ such that

$$C_-|\xi|^2 \le \xi \cdot (A(\cdot)\xi) \le C_+|\xi|^2\,, \qquad C_-|\xi|^2 \le \xi \cdot (\kappa(\cdot)\xi) \le C_+|\xi|^2\,, \quad \text{for all parameters}\,.$$

Furthermore, let $B \in C^{0,1}(\mathbb{R}\times\mathbb{R}^3\times\mathbb{R};\mathbb{R})$ be Lipschitz with $0 < C_- \le B \le C_+$. Then, we define A, κ, B as functionals

$$\begin{aligned} A_\bullet, \kappa_\bullet : \quad & \tilde{\mathcal{H}} \to L^\infty(\Omega)^{3\times 3} \\ & (c,e) \mapsto A_{(c,e)}(\cdot) := A(c(\cdot), \nabla c(\cdot), e(\cdot)) \\ & \qquad \mapsto \kappa_{(c,e)}(\cdot) := \kappa(c(\cdot), \nabla c(\cdot), e(\cdot)) , \\ B_\bullet : \quad & \tilde{\mathcal{H}} \to L^\infty(\Omega) \\ & (c,e) \mapsto B_{(c,e)}(\cdot) := B(c(\cdot), \nabla c(\cdot), e(\cdot)) . \end{aligned}$$

For any $\tilde{s} := (s_c, s_e) \in \mathcal{H}$, we then solve the equations

$$\begin{aligned} -s_c &= \operatorname{div}\left(A_{(c,e)}\nabla p^c_{\tilde{s}}\right) - B_{(c,e)} p^c_{\tilde{s}} \quad \text{on } \Omega , \qquad \left(A_{(c,e)}\nabla p^c_{\tilde{s}}\right)\cdot \boldsymbol{n}_\Gamma = 0 \quad \text{on } \Gamma , \\ -s_e &= \operatorname{div}\,\left(\kappa_{(c,e)}\nabla p^e_{\tilde{s}}\right) \qquad\qquad\quad \text{on } \Omega , \qquad \left(\kappa_{(c,e)}\nabla p^e_{\tilde{s}}\right)\cdot \boldsymbol{n}_\Gamma = 0 \quad \text{on } \Gamma , \end{aligned} \tag{8.3}$$

with $p^c_{\tilde{s}} \in H^1(\Omega)$, $p^e_{\tilde{s}} \in H^1_{(0)}(\Omega)$.

Corollary 8.2. *For any sequences $(c_n, e_n) \to (c,e)$ strongly in $\tilde{\mathcal{H}}$ and $\tilde{s}_n \rightharpoonup \tilde{s}$ weakly in $\mathcal{H}$, the corresponding sequence $\left(p^c_{\tilde{s}_n}, p^e_{\tilde{s}_n}\right) \in H^1(\Omega)\times H^1_{(0)}(\Omega)$ is bounded by $C\left\|\tilde{s}_n\right\|_{\mathcal{H}}$ with some constant C, and weakly converges towards $(p^c_{\tilde{s}}, p^e_{\tilde{s}})$.*

As $p^c_{\tilde{s}}$ and $p^e_{\tilde{s}}$ are uniquely determined, so is the expression

$$g_{(c,e)}\left(\tilde{s}_1, \tilde{s}_2\right) := \int_\Omega \left(\nabla p^c_{\tilde{s}_1}\cdot\left(A_{(c,e)}\nabla p^c_{\tilde{s}_2}\right) + B_{(c,e)} p^c_{\tilde{s}_1} p^c_{\tilde{s}_2} + \nabla p^e_{\tilde{s}_1}\cdot\left(\kappa_{(c,e)}\nabla p^e_{\tilde{s}_2}\right)\right) .$$

As trivial consequence of lemma 2.3 we get:

Corollary 8.3. *There are constants G_- and G_+ such that $G_-\left\|\tilde{s}\right\|_{\mathcal{H}} \le g_{(c,e)}\left(\tilde{s},\tilde{s}\right) \le G_+\left\|\tilde{s}\right\|_{\mathcal{H}}$ independent on $(c,e)\in\tilde{\mathcal{H}}$.*

Next, note that

$$\begin{aligned} g_{(c,e)}\left(\tilde{s}_1, \tilde{s}_2\right) &= \int_\Omega \left(-\nabla\cdot\left(A_{(c,e)}\nabla p^c_{\tilde{s}_1}\right) p^c_{\tilde{s}_2} + B_{(c,e)} p^c_{\tilde{s}_1} p^c_{\tilde{s}_2} + \nabla\cdot\left(\kappa_{(c,e)}\nabla p^e_{\tilde{s}_1}\right) p^e_{\tilde{s}_2}\right) \\ &= \int_\Omega \left(s_{1,c} p^c_{\tilde{s}_2} + s_{1,e} p^e_{\tilde{s}_2}\right) \end{aligned} \tag{8.4}$$

In order to verify 1.1-(2) note that for $(c_n, e_n) \to (c,e)$ strongly in $\tilde{\mathcal{H}}$ and $\tilde{s}_n \rightharpoonup \tilde{s}$ weakly in $\mathcal{H}$, we get for any fixed $\tilde{s}_* \in \mathcal{H}$ due to corollary 8.2:

$$\lim_{n\to\infty} g_{(c_n,e_n)}\left(\tilde{s}_*, \tilde{s}_n\right) = \lim_{n\to\infty}\int_\Omega \left(s_{*,c} p^c_{\tilde{s}_n} + s_{*,e} p^e_{\tilde{s}_n}\right) = \int_\Omega \left(s_{*,c} p^c_{\tilde{s}} + s_{*,e} p^e_{\tilde{s}}\right) = g_{(c,e)}\left(\tilde{s}_*, \tilde{s}\right) .$$

We will now recover the gradient flow of $\mathcal{S}$ in $\left(\mathcal{H}_0, \tilde{\mathcal{H}}, \mathcal{H}, g\right)$: First, in order to recover $\nabla\mathcal{S}$, we follow the calculations and argumentation of section 3 to find with help of lemma 3.2 :

$$\left\langle \nabla S(\gamma(\tau)), \tilde{s}\right\rangle_{g(\gamma(\tau))} = \int_\Omega \left(p^c_{\nabla\mathcal{S}} s_c + p^e_{\nabla\mathcal{S}} s_e\right) \tag{8.5}$$

with $(p^c_{\nabla\mathcal{S}}, p^e_{\nabla\mathcal{S}}) \in D_{L^2}\mathcal{S}(\gamma(\tau))$, meaning $p^c_{\nabla\mathcal{S}}$ has the form $p^c_{\nabla\mathcal{S}} := \left(-\operatorname{div}\,(\sigma\nabla c) + \frac{\partial\eta_0}{\partial c}\right)$ and $p^e_{\nabla\mathcal{S}} = P_0\frac{\partial\eta_0}{\partial e}$. We may use these insights to reformulate the equation

$$\partial_t u(t) = \nabla\mathcal{S}(u) \qquad \Leftrightarrow \qquad g_u\left(\partial_t u, \phi\right) = g_u(\nabla\mathcal{S}(u), \phi) \qquad \forall\phi\in L^2(0,T;\mathcal{H})$$

into

$$\begin{aligned} \partial_t c &\in \operatorname{div}\,\left(A_{(c,e)}\nabla\left(-\operatorname{div}\,(\sigma\nabla c) + \frac{\partial\eta_0}{\partial c}\right)\right) - B_{(c,e)}\left(-\operatorname{div}\,(\sigma\nabla c) + \frac{\partial\eta_0}{\partial c}\right) \\ \partial_t e &\in \nabla\cdot\left(\kappa_{(c,e)}\nabla\frac{\partial s_1}{\partial e}\right) . \end{aligned}$$

Note that defining

$$\boldsymbol{j}_{\tilde{s}} := -A_{(c,e)}\nabla p^c_{\tilde{s}}, \quad \overset{+}{c}_{\tilde{s}} := -B_{(c,e)} p^c_{\tilde{s}}, \quad \boldsymbol{h} := \kappa_{(c,e)}\nabla p^e_{\tilde{s}}$$

the definition for g can be rewritten as

$$g\left(\tilde{s}_1,\tilde{s}_2\right):=\int_\Omega\left(\boldsymbol{j}_{\tilde{s}_1}\cdot\left(A^{-1}_{(c,e)}\boldsymbol{j}_{\tilde{s}_2}\right)+B^{-1}_{(c,e)}\overset{+}{c}_{\tilde{s}_1}\overset{+}{c}_{\tilde{s}_2}+\boldsymbol{h}_{\tilde{s}_1}\cdot\left(\kappa^{-1}_{(c,e)}\boldsymbol{h}_{\tilde{s}_2}\right)\right).$$

Due to the linear dependence of $\boldsymbol{j}_{\tilde{s}}$, $\overset{+}{c}_{\tilde{s}}$ and $\boldsymbol{h}_{\tilde{s}}$ on $\tilde{s}$, with regard to lemma 7.7, it is natural that the formal calculations in 7 and the last calculations yield the same model equations. In particular, the linear dependence of $\boldsymbol{j}$,$\boldsymbol{h}$ and $\overset{+}{c}$ on $\partial_t c$ and $\partial_t e$ yields equivalence of the maximization principle from section 7 and the maximum characterization of gradient flows from lemma 7.7.

8.2. **The general case.** In a more general setting with $J=m+k+l$ constituents, assume $\mathcal{H}:=L^2(\Omega)^m\times H^*(\Omega)^k\times H^{-1}_{(0)}(\Omega)^l\times H^{-1}_{(0)}(\Omega)$ with $\tilde{\mathcal{H}}=L^2(\Omega)^m\times H^1(\Omega)^k\times H^1(\Omega)^l\times L^2(\Omega)$. We indicate elements of $\mathcal{H}$ by (c,e) with $c=(c_1,\dots,c_m,c_{m+1},\dots c_{m+k},c_{m+k+1},\dots,c_{m+k+l})\in H^1(\Omega)^m\times H^1(\Omega)^k\times H^1_{(0)}(\Omega)^l$ and $e\in L^2(\Omega)$.

Physically, the first $m+k$ constituents are those that might take part in chemical reactions, while the l next constituents are only subject to diffusive transport.

We then consider a functional

$$\mathcal{S}:\quad \mathcal{H}\to\mathbb{R},\qquad (c,e)\mapsto\int_\Omega s(c,e)+\nabla c\cdot(\sigma\nabla c)\,, \tag{8.6}$$

where σ is strictly positive uniformly elliptic and we write for simplicity $\sigma=(\sigma_{ij})_{i,j\in\{1,\dots,m+k+l\}}$ with $\sigma_{ij}\in L^\infty(\Omega)^{3\times3}$ for all $i,j\in\{1,\dots,m+k+l\}$. Concerning $s(\cdot,\cdot)$ we will be more precise below. For the moment, let us state that s is chosen such that one of theorems 1.7-1.9 is applicable. In particular, we find $\mathcal{S}(c,e)<+\infty$ implies $\|c\|_{H^1(\Omega)^{m+k+l}}<+\infty$.

Definition 8.4. Let $\mathbb{L}^2:=L^2(\Omega)^{m+k}\times L^2_{(0)}(\Omega)^{l+1}$. We denote $D_{L^2}\mathcal{S}(u)$ the set of all $\delta_u\in\mathbb{L}^2$ such that for all $\phi\in\mathbb{L}^2$ and all $\gamma\in H^1(-1,1;\mathbb{L}^2)$ with $\gamma(0)=u$ and $\gamma'(0)=\phi$ holds

$$\langle\delta_u,\phi\rangle_{\mathbb{L}^2}\le\lim_{t\to0}\frac{1}{t}\left(\mathcal{S}(\gamma(t))-\mathcal{S}(u)\right)\,. \tag{8.7}$$

Moreover, we denote by $\mathbb{P}_i$ resp. $\mathbb{P}_e$ the projections on the i-th resp. the e- coordinates of $\mathbb{L}^2$ and say $\delta_i\in\frac{\delta S}{\delta c_i}(c,e)$ resp. $e\in\frac{\delta S}{\delta e}(c,e)$ iff (8.7) holds for all $\phi\in\mathbb{P}_i(\mathbb{L}^2)$ resp. $\phi\in\mathbb{P}_e(\mathbb{L}^2)$.

We set $n=m+k+l+1$ and consider Lipschitz continuous functions

$$A:\mathbb{R}^n\times\mathbb{R}^{3(n-1)}\to\left(\mathbb{R}^{3\times3}\right)^{(n-m)\times(n-m)},\qquad B:\mathbb{R}^n\times\mathbb{R}^{3(n-1)}\to\mathbb{R}^{(m+k)\times(m+k)}$$

such that A and B are symmetric, strictly positive definite and uniform elliptic. For simplicity, we write $A_{(c,e)}=\left(A_{(c,e),ij}\right)_{i,j=m+1,\dots n}:=\left(A_{ij}(c,e,\nabla c)\in\mathbb{R}^{3\times3}\right)_{i,j=m+1,\dots n}$ and $B_{(c,e)}=\left(B_{(c,e),ij}\right)_{i,j=1,\dots m+k}:=\left(B_{ij}(c,e,\nabla c)\right)_{i,j=1,\dots m+k}$ and for any $\tilde{s}:=\left((s_{c,i})_{i=1,\dots,n-1},s_e\right)\in\mathcal{H}$, we solve the equations

$$\begin{aligned}
s_{c,i}&=\sum_{j=1\dots m+k}B_{(c,e)i,j}p^j_{\tilde{s}}\quad i=1\dots m\\
s_{c,i}&=-\sum_{j=m+1\dots n}\operatorname{div}\left(A_{(c,e),i,j}\nabla p^j_{\tilde{s}}\right)+\sum_{j=1\dots m+k}B_{(c,e),i,j}p^j_{\tilde{s}}\quad i=m+1\dots m+k\\
s_{c,i}&=-\sum_{j=m+1\dots n}\operatorname{div}\left(A_{(c,e),i,j}\nabla p^j_{\tilde{s}}\right)\quad i=m+k+1\dots m+k+l\\
s_e&=-\sum_{j=m+1\dots n}\operatorname{div}\left(A_{(c,e),n,j}\nabla p^j_{\tilde{s}}\right)
\end{aligned}\tag{8.8}$$

with $p^j_{\tilde{s}}\in H^1_n(\Omega)$ for all $j=1,\dots m$ and $p^e_{\tilde{s}}\in H^1_{(0),n}(\Omega)$. We define $\tilde{\nabla}p:=(\left(\nabla p^j_{\tilde{s}}\right)_{i=m+1\dots n})$ and $\tilde{p}:=\left(p^j_{\tilde{s}}\right)_{i=1\dots m+k}$, to define g through

$$g\left(\tilde{s}_1,\tilde{s}_2\right):=\int_\Omega\left(\tilde{\nabla}p_1\left(A\tilde{\nabla}p_2\right)+\tilde{p}_1B\tilde{p}_2\right).$$

Defining

$$\boldsymbol{j}_i := \begin{cases} \sum_{j=m+1\dots n} A_{(c,e),i,j} \nabla p_{\tilde{s}}^j & \text{for } i = m+1, \dots n-1 \\ \boldsymbol{0} & \text{else} \end{cases}$$

$$\overset{+}{c}_i := \begin{cases} \sum_{j=1\dots m+k} B_{(c,e),i,j} p_{\tilde{s}}^j & \text{for } i = 1, \dots m+k \\ 0 & \text{else} \end{cases}$$

$$\boldsymbol{h} := -\sum_{j=1\dots m} A_{(c,e),n,j} \nabla p_{\tilde{s}}^j$$

and

$$\mathcal{J}_a := \left(\boldsymbol{j}_{a,m+1}, \dots, \boldsymbol{j}_{a,n-1}, \boldsymbol{h}_a\right), \qquad \mathcal{C}_a = \left(\overset{+}{c}_{a,1}, \dots \overset{+}{c}_{a,m+k}\right), \qquad a \in \{1,2\}$$

the last definition for g can be rewritten as

$$g\left(\tilde{s}_1, \tilde{s}_2\right) := \int_\Omega \left(\mathcal{J}_1 \left(A_{(c,e)}^{-1} \mathcal{J}_2\right) + \mathcal{C}_1 B_{(c,e)}^{-1} \mathcal{C}_2\right) .$$

The calculation of the effective equations is analogous with the simple case and leads to

$$\partial_t c_i \in \overset{+}{c}_i - \operatorname{div} \boldsymbol{j}_i$$
$$\partial_t e + \operatorname{div} \boldsymbol{h} \ni 0 .$$

With regard to the previous paragraph or with regard to example 3.1, it is easy to check that $p_{\tilde{s}}^j \in \frac{\delta \mathcal{S}}{\delta c_j}$ for $j \in \{1, \dots, n-1\}$ and $p_{\tilde{s}}^j \in \frac{\delta \mathcal{S}}{\delta e}$ for $j = n$.

It finally remains to identify a broader class of physically relevant functionals $\mathcal{S}$, resp. functions $s(\cdot,\cdot)$, for which one of theorems 1.7-1.9 can be applied and for which we can verify relation

$$\frac{\delta \mathcal{S}}{\delta u_i} = \partial_{u_i} \phi(u) - \operatorname{div} \left(\sum_j \sigma_{ij} \nabla u_j + \sigma_{ii} \nabla u_i \right) . \tag{8.9}$$

8.3. **Choice of the functional $\mathcal{S}$.** The rest of this section is devoted to the identification of a class of functionals $\mathcal{S}$ that satisfies the claims of either theorem 1.7 or 1.8.

We consider functionals

$$\tilde{\mathcal{S}}(u) := \int_\Omega \phi(u) + \int_\Omega \nabla u \cdot (\sigma \nabla u)$$

with σ uniformly elliptic and $\sigma = (\sigma_{ij})_{i,j \in \{1,\dots,m+k\}}$ with $\sigma_{ij} \in L^\infty(\Omega)^{3\times 3}$ for all $i,j \in \{1,\dots,m+k\}$ and $\phi : \mathbb{R}^{m+k} \to \mathbb{R}$ to be specified below. Note that in contrary with definition (8.4) we restrict to the first $m+k$ coordinates and $D_{L^2}(\tilde{\mathcal{S}})$ has to be understood in the sense of pointwise $L^2(\Omega)$-derivatives.

We start by recalling some basic facts on convex sets and convex functions and by introducing our notations. For simplicity, we write

$$\phi' := \nabla \phi \qquad \text{and} \qquad \phi'' := \nabla \left(\nabla \phi\right) .$$

Definition 8.5. For any $n \in \mathbb{N}$ let $K^\circ \subset \mathbb{R}^n$ be open and convex with $K := K^\circ \cup \partial K$. We say that a convex function $\phi \in C^2(K; [0, +\infty))$ is *admissible* if there is a sequence of convex functions $\phi_\varepsilon \in H^2(K; [0, +\infty)) \cap C^2(\overline{K})$ with $\phi_\varepsilon(x) \le \phi(x)$, $|\phi_\varepsilon'(x)| \le C_\varepsilon$ for some $C_\varepsilon > 0$ and $\phi_\varepsilon'(x) \cdot \phi_\varepsilon'(x) \le \phi'(x) \cdot \phi_\varepsilon'(x) \le \phi'(x) \cdot \phi'(x)$ for all $x \in K$, and $\phi_\varepsilon(x) \to \phi(x)$, $\phi_\varepsilon'(x) \to \phi'(x)$ for all $x \in K$. Furthermore, for any $x \in K$, $y \notin K$ and $s \in [0,1]$ s.t. $x + s(y-x) \in \partial K$

$$\phi_\varepsilon'(x + s(y-x)) \cdot (y-x) \ge 0 \tag{8.10}$$

and for any ε there is $t_* > 0$ such that for any $t < t_*$, $u \in K$ holds

$$\lim_{|v| \to +\infty} \frac{1}{2t} |u - v| + \phi_\varepsilon(v) \to \infty .$$

Corollary 8.6. *Let ϕ be an admissible convex function on the closed convex set K with $K^\circ := K \backslash \partial K$ and assume*

$$(x_j)_{j \in \mathbb{N}} \in K^\circ,\ |x_j| < \mathit{const}\ \forall j\,, x_j \to x \in \partial K \text{ as } j \to \infty \quad \Rightarrow \quad |\nabla \phi(x_j)| \to +\infty \text{ as } j \to \infty . \tag{8.11}$$

Then, for any $\varepsilon > 0$ and any $x \notin K$ and any $t > 0$ holds $x + t \nabla \phi_\varepsilon(x) \notin K$.

Proof. We find for any α such that $y := x + \alpha\nabla\phi_\varepsilon(x)$ satisfies $y \in K$ that $\psi(t) := \phi_\varepsilon(y + t(x-y))$ is convex and $\frac{d}{dt}\psi(t) = \phi_\varepsilon'(y + t(x-y)) \cdot (x-y)$. In $t = 1$ we thus find $\frac{d}{dt}\psi(t) = -\alpha\left|\nabla\phi_\varepsilon(x)\right|^2$. For t such that $y + t(x-y) \in \partial K$ we already find $\frac{d}{dt}\psi(t) \geq 0$ and thus $\alpha < 0$. □

Lemma 8.7. *Let $K \subset \mathbb{R}^{m+k}$ be a convex and closed subset with $K^\circ := K\backslash\partial K$ and $\phi : K \to \mathbb{R}$ be an admissible convex function in K° or a sum of admissible convex functions. Furthermore, suppose* (8.11) *holds and set $|\nabla\phi(x)| = +\infty$ for $x \notin K^\circ$. Then,*

$$D(D_{L^2}(\tilde{\mathcal{S}})) \subset L^2_{(K)}(\Omega) := \left\{ fu \in L^2(\Omega) \,:\, \int_\Omega u \in K^\circ\,,\ \left(\sum_j \sigma_{ij}\nabla u_j\right)\cdot \boldsymbol{n}_\Gamma = 0\ \forall i \leq m+k \right\}$$

and for $w \in D_{L^2}(\tilde{\mathcal{S}})$ holds

$$\int_\Omega \left|\phi'(u)\right|^2 + \int_\Omega \nabla u(\sigma(\phi''(u)\nabla u)) \leq C\left\|w\right\|_{L^2(\Omega)}\,.$$

Furthermore, $D_{L^2}(\tilde{\mathcal{S}})$ is single valued and

$$D_{L^2}(\tilde{\mathcal{S}}) = \phi' - div(\sigma\nabla u)\,.$$

Remark 8.8. We have $\nabla u(\sigma(\phi''(u)\nabla u)) > 0$ and thus $\phi'(u) \in L^2(\Omega)$: Due to properties of ϕ, ϕ'' is diagonizable with positive eigenvalues, thus let λ be an eigenvalue of ϕ'' and e the corresponding eigenvector and $E := (e,0,0) \in \mathbb{R}^{n\times 3}$, then $\phi''(u)E = \lambda E$ and $E(\sigma(\phi''(u)E)) = \lambda E\sigma E > 0$.

Proof. We consider the case ϕ is an admissible function. For $\varepsilon > 0$ let ϕ_ε be the approximation from definition 8.5. Now, let $u \in L^2_{(K)}(\Omega)$ and solve for $u_t = u - t\phi_\varepsilon'(u_t)$ which is the unique minimizer of

$$\Phi_\varepsilon(v) := \frac{1}{2t}\left|u - v\right|^2 + \phi_\varepsilon(v)\,,$$

as the Hessian matrix of Φ_ε is strictly positive. Corollary 8.6 yields $u_t \in K$ for all $u \in K$ and due to the inverse function theorem, we can write $u_t = F_t^\varepsilon(u)$ where $F_t^\varepsilon : K \to K$ is continuously differentiable with bounded derivative (note that the Hessian of ϕ_ε is bounded). Thus, as for $\mathcal{F}_t : v \mapsto v + t\phi_\varepsilon'(v)$ with $\mathcal{F}_t'(v) = Id + t\phi_\varepsilon''(v)$ we find $\mathcal{F}_t \to Id$ and $\mathcal{F}_t' \to Id$ uniformly on K, $u_t, \phi_\varepsilon'(u_t) \in H^1(\Omega)$ and

$$u_t \to u, \quad \phi_\varepsilon'(u_t) \to \phi_\varepsilon'(u) \quad \text{strongly in } H^1(\Omega) \quad \text{as } t \to 0\,.$$

Thus, setting $v = -\phi_\varepsilon'(u_t)$, we find for each $w \in D_{L^2}(\tilde{\mathcal{S}}(u))$

$$\begin{aligned}\langle w, v\rangle &\geq \lim_{t\to 0}\frac{1}{t}\left(\tilde{\mathcal{S}}(u) - \tilde{\mathcal{S}}(u+tv)\right)\\ &= \int_\Omega \phi_\varepsilon'(u_t)\cdot\phi'(u) + \int_\Omega \nabla u\sigma\nabla(\phi_\varepsilon'(u_t))\,,\end{aligned}$$

which yields for $t \to 0$:

$$\int_\Omega \left|\phi_\varepsilon'(u)\right|^2 + \int_\Omega \nabla u(\sigma(\phi_\varepsilon''(u)\nabla u)) \leq C\left\|w\right\|_{L^2(\Omega)}\,.$$

As this holds for all $\varepsilon > 0$ and $\phi_\varepsilon' \to \phi'$, $\phi_\varepsilon'' \to \phi''$ pointwise as $\varepsilon \to 0$, we find

$$\int_\Omega \left|\phi'(u)\right|^2 + \int_\Omega \nabla u(\sigma(\phi''(u)\nabla u)) \leq C\left\|w\right\|_{L^2(\Omega)}\,,$$

and by the same time, as $|\phi'(u)| = \infty$ for $u \in \partial K$, we find $u(x) \in K^\circ$ for a.e. $x \in \Omega$ and $u \in L^2_{(K)}(\Omega)$.

In order to calculate the i-th partial derivative $\frac{\delta\tilde{\mathcal{S}}}{\delta u_i}(u)$ of $\mathcal{S}$, we define

$$K_i^\circ(u) := \{y \in \mathbb{R} \,:\, (u_1, \dots, u_{i-1}, y, u_{i+1}, \dots u_n) \in K^\circ\}$$

and note that $u \in L^2_{(K)}(\Omega)$ implies $u(x) \in K_i^\circ(u(x))$ for a.e. $x \in \Omega$. However, in order to derive $\frac{\delta S}{\delta u_i}(u)$ we need to make sure for the test functions ψ holds $u(x) + t\psi(x) \in K$ for a.e. $x \in \Omega$. Thus, for $M > 0$ consider the sets $K_M := \{u \in K \,:\, |\phi'(u)| < M+1\}$. Then, K_M is boundedly away from ∂K and choosing functions $\chi_M : \mathbb{R}^{(m+k)} \to [0,1]$ with $\chi_M(x) = 1$ for $|x| < M$, $\chi_M(x) = 0$ for $|x| > M+1$ and $\chi_M' \leq 2$ we get for any $\psi \in C_0^\infty(\Omega)$ and for t small enough:

$$\psi_M := \chi_M\left(\phi'(u)\right)\psi \in H^1(\Omega), \qquad u(x) + t\psi_M(x) \in K$$

for a.e. $x \in \Omega$. In particular, for any $\psi \in C_0^\infty$ the function $u_t := u - t\psi_M \boldsymbol{e}_i$ satisfies $\tilde{\mathcal{S}}(u_t) < +\infty$. Now, we find for $w \in \frac{\delta \tilde{\mathcal{S}}}{\delta u_i}(u)$

$$\langle w, \psi_M \rangle \geq \lim_{t\to 0} \frac{1}{t}\left(\tilde{\mathcal{S}}(u) - \tilde{\mathcal{S}}(u_t)\right) = \int_\Omega \phi'(u) \cdot \boldsymbol{e}_i\, \psi_M - (\sum_j \sigma_{ij} \nabla u_j) \cdot \nabla \left(\boldsymbol{e}_i \psi_M\right) \tag{8.12}$$

and replacing ψ with $-\psi$, we obtain equality in the last relation. Note that $\chi_M\left(\phi'(u)\right) \to 1$ pointwise as $M \to +\infty$. Thus, it remains to search for the limit of $\nabla\psi_M$.

We find

$$\begin{aligned}
&\int_\Omega (\sum_j \sigma_{ij} \nabla u_j) \cdot (\boldsymbol{e}_i \otimes \nabla \psi_M) \\
&\qquad = \int_\Omega (\sum_j \sigma_{ij} \nabla u_j) \cdot \chi_M\left(\phi'(u)\right) (\boldsymbol{e}_i \otimes \nabla\psi) + \int_\Omega \psi (\sum_j \sigma_{ij} \nabla u_j) \cdot \chi_M'\left(\phi'(u)\right) \phi''(u) \left(\boldsymbol{e}_i \otimes \nabla u\right)
\end{aligned}$$

As $\nabla u(\sigma(\phi_\varepsilon''(u)\nabla u)) \in L^1(\Omega)$ and $\chi_M'\left(\phi'(u)\right) \to 0$ pointwise in Ω, the second term on the left hand side vanishes as $M \to \infty$. Thus, the limit reads

$$\lim_{M\to\infty} \int_\Omega (\sum_j \sigma_{ij} \nabla u_j) \cdot (\boldsymbol{e}_i \otimes \nabla \psi_M) = \int_\Omega (\sum_j \sigma_{ij} \nabla u_j) \cdot (\boldsymbol{e}_i \otimes \nabla \psi)\ .$$

We use this result and conclude in (8.12) that

$$\begin{aligned}
\langle w, \psi \rangle &= \int_\Omega \phi'(u) \cdot \boldsymbol{e}_i\, \psi - +(\sum_j \sigma_{ij} \nabla u_j) \cdot \nabla (\boldsymbol{e}_i \psi) \\
&= \int_\Omega \left(\phi'(u) - \operatorname{div} (\sum_j \sigma_{ij} \nabla u_j) \right) \cdot \boldsymbol{e}_i\, \psi\,.
\end{aligned}$$

□

Remark 8.9. Note that above calculations do not hold for subdifferentials in $L^2_{(0)}(\Omega)$. This is for the simple reason that the derivation of the L^2-bound on $\phi'(u)$ is not valid in that case.

Example 8.10. We will now give some examples of admissible convex functions:

(1) Let $K_1 \subset \mathbb{R}$ be convex and $f : K_1 \times \mathbb{R}^{n-1} \to \mathbb{R}$, $x = (x_1, \dots, x_n) \mapsto f(x)$ be constant in $x_2, \dots, x_n$ and convex and twice continuously differentiable in x_1 and fulfill (8.10). Then, f is admissible.
(2) For $i = 1, \dots m$ let $K_i \subset \mathbb{R}^n$ be convex sets and $f_i : K_i \to \mathbb{R}$ be admissible functions. Then, $f := \sum_i f_i$ is an admissible function on $K := \bigcap_i K_i$.
(3) Convex functions with bounded derivatives are admissible.
(4) Let $K \subset \mathbb{R}^n$ be convex with C^∞ -boundary. Then, the distance function $d(x) := \operatorname{dist}(x, \partial K)$ is C^∞. Now, for $f \in C^2((0, +\infty)$ convex with $f'(t) \to +\infty$ as $t \to 0$ and $f'(t) \to 0$ as $t \to +\infty$, the function $\phi(x) := f(d(x))$ is admissible: It is easy to see that $d(x)$ in concave and f is monotonically decreasing. Thus, a simple calculation leads to

$$f(d(\lambda x + (1-\lambda)y)) \leq f(\lambda d(x) + (1-\lambda) d(y)) \leq \lambda f(d(x)) + (1-\lambda) f(d(y))\,.$$

As f is admissible on $(0, +\infty)$, so is $f \circ d$.

Using this toolbox, some more complicated examples may be constructed by summing up. However, there remains the question for what functions ϕ we can show (8.9) for $i > m+k$. The results from [1] do not apply as the proof strongly relies on the one dimensional structure.

However, we give one particular example where (8.9) holds: Let $\tilde{K}_i \subset \mathbb{R}$ be convex sets with $\tilde{f}_i : \tilde{K}_i \to \mathbb{R}$ admissible functions. Then, defining $K_i := \mathbb{R}^{i-1} \times \tilde{K}_i \times \mathbb{R}^{n-i-1}$, $K := \bigcap_i K_i \subset \mathbb{R}^{n-1}$, the proof of Abels and Wilke [1] of above Lemma 3.2 yields validity of (8.9) for

$$\phi(u) := \begin{cases} \sum_i \tilde{f}_i(u_i) & \text{for } u \in K \\ +\infty & \text{else} \end{cases}\ .$$

Of course, we can use any invertible matrix $Q : \mathbb{R}^{n-1} \to \mathbb{R}^{n-1}$ and (8.9) would still hold for $\phi \circ Q$ with K replaced by $Q^{-1}K$. Also, we can add any term $\phi_1(u)$ with $\phi_1(u) \in C^1(\overline{K}) \cap C^2(K^\circ)$ being a convex function with bounded derivatives.

For multiphase systems in terms of concentrations, a system of $J + 1 = m + k + l + 1$ constituents is represented by J mass concentrations $c = (c_1, \dots, c_J)$ with the properties

$$c \in K := \left\{ a \in \mathbb{R}^J \, : \, a_i \geq 0 \quad \forall i, \qquad \sum_i a_i \leq 1 \right\}$$

reflecting the idea, that the mass concentration is always nonnegative and that the sum of all $J + 1$-mass concentrations add up to 1.

Using above considerations, we may quickly construct admissible functions $\tilde{f}_i : [0,1] \to \mathbb{R}$, $i = 1, \dots, J$, $\tilde{f}_{J+1} : [0,1] \to \mathbb{R}$ and an orthogonal matrix $Q \in \mathbb{R}^{J \times J}$ such that for the standard basis $\boldsymbol{e}_i$ of $\mathbb{R}^J$ holds $\boldsymbol{e}_1 Q \sum_i c_i \boldsymbol{e}_i = \sum_i c_i$. We then define

$$s_0 : K \to \mathbb{R}, \quad c = (c_1, \dots, c_J) \mapsto \sum_{i=1}^{J} \tilde{f}_i(c_i) + \tilde{f}_{J+1}(\boldsymbol{e}_1 Q c)\,, \tag{8.13}$$

which is an admissible function on K satisfying (8.9).

Furthermore, we use any admissible function s_1 on $\mathbb{R}^{m+k}$, any admissible function s_2 on $\mathbb{R}_{\geq 0}$ satisfying (8.1) and any bounded, Lipschitz continuous function s_3 on $\mathbb{R}^n$ to get that

$$s(c,e) := s_0(c) + s_1(c_1, \dots, c_{m+k}) + s_2(e) + e + s_3(c,e)$$

is such that $\mathcal{S}$ fulfills the requirements of theorem 1.7.

References

[1] Helmut Abels and Mathias Wilke. Convergence to equilibrium for the cahn-hilliard equation with a logarithmic free energy. *Nonlinear Analysis: Theory, Methods & Applications*, 67(11):3176–3193, 2007.

[2] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient Flows.* Birkhäuser, 2005.

[3] E.J. Balder. A general approach to lower semicontinuity and lower closure in optimal control theory. *SIAM Journal on Control and Optimization*, 22:570–598, 1984.

[4] H. Brézis. *Opérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert.* North-Holland Publishing Co., Amsterdam, 1973.

[5] H. Brézis. *Analyse functionelle - Théorie et applications.* Masson, Paris, 1983.

[6] M. Brokate, N. Kenmochi, I. Müller, J.F. Rodriguez, and C. Verdi. *Phase transitions and hysteresis. Lectures given at the 3rd session of the Centro Internazionale Matematico Estivo (C.I.M.E.) held in Montecatini Terme, Italy, July 13-21, 1993.* Lecture Notes in Mathematics. 1584. Berlin: Springer-Verlag. vii, 1994.

[7] D. Burago, Y. Burago, and S. Ivanov. *A course in metric geometry.* American Mathematical Society, 2001.

[8] JW Cahn, CM Elliott, and A. Novick-Cohen. The cahn-hilliard equation with a concentration dependent mobility: motion by minus the laplacian of the mean curvature. *European Journal of Applied Mathematics*, 7(3):287–302, 1996.

[9] H. Callen. *Thermodynamics and an introduction to thermostatics.* Wiley, London, 1985.

[10] C. Castaing and M. Valadier. *Convex Analysis and measurable multifunctions.* Springer, 1977.

[11] L. Changchun. Cahn-hilliard equation with terms of lower order and non-constant mobility. *Electronic Journal of Qualitative Theory of Differential Equations*, 10, 2003.

[12] R. Dal Passo, L. Giacomelli, and A. Novick-Cohen. *Existence for an Allen-Cahn/Cahn-Hilliard system with degenerate mobility.* Carnegie Mellon University, Department of Mathematical Sciences [Center for Nonlinear Analysis, 1999.

[13] C. Dellacherie and P.A. Meyer. *Probabilities and potential.* North-Holland Publishing Co., Amsterdam, 1978.

[14] C.M. Elliott and S. Zheng. On the Cahn-Hilliard equation. *Archive for Rational Mechanics and Analysis*, 96(4):339–357, 1986.

[15] M. Grasselli, A. Miranville, R. Rossi, and G. Schimperna. Analysis of the Cahn–Hilliard equation with a chemical potential dependent mobility. *Communications in Partial Differential Equations*, 36(7):1193–1238, 2011.

[16] M. Heida. *Modeling multiphase flow in porous media with an application to permafrost soil.* PhD thesis, University of Heidelberg, 2011.

[17] M. Heida. On the derivation of thermodynamically consistent boundary conditions for the Cahn-Hilliard-Navier-Stokes system. *submitted to IJES*, 2011.

[18] M. Heida. A two-scale model of two-phase flow in porous media ranging from porespace to the macro scale. *submitted to SIAM MMS*, 2012.

[19] M. Heida, J. Màlek, and K.R. Rajagopal. On the development and generalizations of Allen-Cahn and Stefan equations within a thermodynmic framework. *to be submitted to Zeitschrift für Angewandte Mathematik und Physik (ZAMP)*, 2011.

[20] M. Heida, J. Màlek, and K.R. Rajagopal. On the development and generalizations of Cahn-Hilliard equations within rational thermodynamic framework. *accepted by Zeitschrift für Angewandte Mathematik und Physik (ZAMP)*, 2011.

[21] E. H. Lieb and J. Yngvason. The physics and mathematics of the second law of thermodynamics. *Physics Reports*, 310:1, 1999.
[22] C. Liu. On the convective cahn–hilliard equation with degenerate mobility. *Journal of Mathematical Analysis and Applications*, 344(1):124–144, 2008.
[23] A. Mielke, R. Rossi, and G. Savaré. Nonsmooth analysis of doubly nonlinear evolution equations. *Calculus of Variations and Partial Differential Equations*, pages 1–58, 2011.
[24] A. Novick-Cohen. The cahn–hilliard equation. *Handbook of differential equations: evolutionary equations*, 4:201–228, 2008.
[25] A. Novick-Cohen. *The Cahn-Hilliard Equation: From Backwards Diffusion to Surface Diffusion.* Cambridge University Press, to appear.
[26] F. Otto. The geometry of dissipative evolution equations: the porous medium equation. *Communications in Partial Differential Equations*, 26:101–174, 2001.
[27] K. R. Rajagopal and A. R. Srinivasa. On thermomechanical restrictions of continua. *Procedings of the Royal Society in London, Series A, Mathematical and Physical and Engineering Sciences*, 460:631–651, 2004.
[28] R. Rossi. On two classes of generalized viscous Cahn-Hilliard equations. *Commun. Pure Appl. Anal*, 4(2):405–430, 2005.
[29] R. Rossi, A. Mielke, and G. Savaré. A metric approach to a class of doubly nonlinear evolution equations and applications. *Annali della Scuola Normale Superiore di Pisa-Classe di Scienze-Serie IV*, 7(1):97, 2008.
[30] R. Rossi and G. Savaré. Gradient flows of non convex functionals in Hilbert spaces and applications. *ESAIM: Control, Optimisation and Calculus of Variations*, 12(03):564–614, 2006.
[31] C. Truesdell. *Rational thermodynamics*, volume 53. Springer, Berlin, 1985.
[32] C. Villani. *Optimal transport: old and new*, volume 338. Springer Verlag, 2009.

University of Dortmund, Faculty of Mathematics, Vogelpothsweg 87, 44227 Dortmund
E-mail address: martin.heida@tu-dortmund.de

Preprints ab 2010/16

2012-19 **Martin Heida**
On gradient flows of nonconvex functionals in Hilbert spaces with
Riemannian metric and application to Cahn-Hilliard equations

2012-18 **R.V. Kohn, J. Lu, B. Schweizer, and M.I. Weinstein**
A variational perspective on cloaking by anomalous localized resonance

2012-17 **Margit Rösler and Michael Voit**
Olshanski spherical functions for infinite dimensional motion groups of fixed rank

2012-16 **Selim Esedoḡlu, Andreas Rätz, Matthias Röger**
Colliding Interfaces in Old and New Diffuse-interface Approximations of Willmore-flow

2012-15 **Patrick Henning, Mario Ohlberger and Ben Schweizer**
An adaptive multiscale finite element method

2012-14 **Andreas Knauf, Frank Schulz, Karl Friedrich Siburg**
Positive topological entropy for multi-bump magnetic fields

2012-13 **Margit Rösler, Tom Koornwinder, and Michael Voit**
Limit transition between hypergeometric functions of type BC and Type A

2012-12 **Alexander Schnurr**
Generalization of the Blumenthal-Getoor Index to the Class of Homogeneous
Diffusions with Jumps and some Applications

2012-11 **Wilfried Hazod**
Remarks on pseudo stable laws on contractible groups

2012-10 **Waldemar Grundmann**
Limit theorems for radial random walks on Euclidean spaces of high dimensions

2012-09 **Martin Heida**
A two-scale model of two-phase flow in porous media ranging from porespace
to the macro scale

2012-08 **Martin Heida**
On the derivation of thermodynamically consistent boundary
conditions for the Cahn-Hilliard-Navier-Stokes system

2012-07 **Michael Voit**
Uniform oscillatory behavior of spherical functions of GL_n/U_n at the identity
and a central limit theorem

2012-06 **Agnes Lamacz and Ben Schweizer**
Effective Maxwell equations in a geometry with flat rings of arbitrary shape

2012-05 **Frank Klinker and Günter Skoruppa**
Ein optimiertes Glättungsverfahren motiviert durch eine technische Fragestellung

2012-04 **Patrick Henning, Mario Ohlberger, and Ben Schweizer**
Homogenization of the degenerate two-phase flow equations

2012-03 **Andreas Rätz**
A new diffuse-interface model for step flow in epitaxial growth

2012-02 **Andreas Rätz and Ben Schweizer**
Hysteresis models and gravity fingering in porous media

2012-01 **Wilfried Hazod**
Intrinsic topologies on H-contraction groups with applications to semistability

2011-14 **Guy Bouchitté and Ben Schweizer**
Plasmonic waves allow perfect transmission through sub-wavelength metallic gratings

2011-13 **Waldemar Grundmann**
Moment functions and Central Limit Theorem for Jacobi hypergroups on $[0, \infty[$

2011-12 **J. Koch, A. Rätz, and B. Schweizer**
Two-phase flow equations with a dynamic capillary pressure

2011-11 **Michael Voit**
Central limit theorems for hyperbolic spaces and Jacobi processes on $[0, \infty[$

2011-10 **Ben Schweizer**
The Richards equation with hysteresis and degenerate capillary pressure

2011-09 **Andreas Rätz and Matthias Röger**
Turing instabilities in a mathematical model for signaling networks

2011-08 **Matthias Röger and Reiner Schätzle**
Control of the isoperimetric deficit by the Willmore deficit

2011-07 **Frank Klinker**
Generalized duality for k-forms

2011-06 **Sebastian Aland, Andreas Rätz, Matthias Röger, and Axel Voigt**
Buckling instability of viral capsides - a continuum approach

2011-05 **Wilfried Hazod**
The concentration function problem for locally compact groups revisited: Non-dissipating space-time random walks, τ-decomposable laws and their continuous time analogues

2011-04 **Wilfried Hazod, Katrin Kosfeld**
Multiple decomposability of probabilities on contractible locally compact groups

2011-03 **Alexandra Monzner* and Frol Zapolsky†**
A comparison of symplectic homogenization and Calabi quasi-states

2011-02 **Stefan Jäschke, Karl Friedrich Siburg and Pavel A. Stoimenov**
Modelling dependence of extreme events in energy markets using tail copulas

2011-01 **Ben Schweizer and Marco Veneroni**
The needle problem approach to non-periodic homogenization

2010-16 **Sebastian Engelke and Jeannette H.C. Woerner**
A unifying approach to fractional Lévy processes